De Gruyter. All rights reserved. May not be reproduced in any form without permission from the publisher, permitted under U.S. or applicable copyright law.

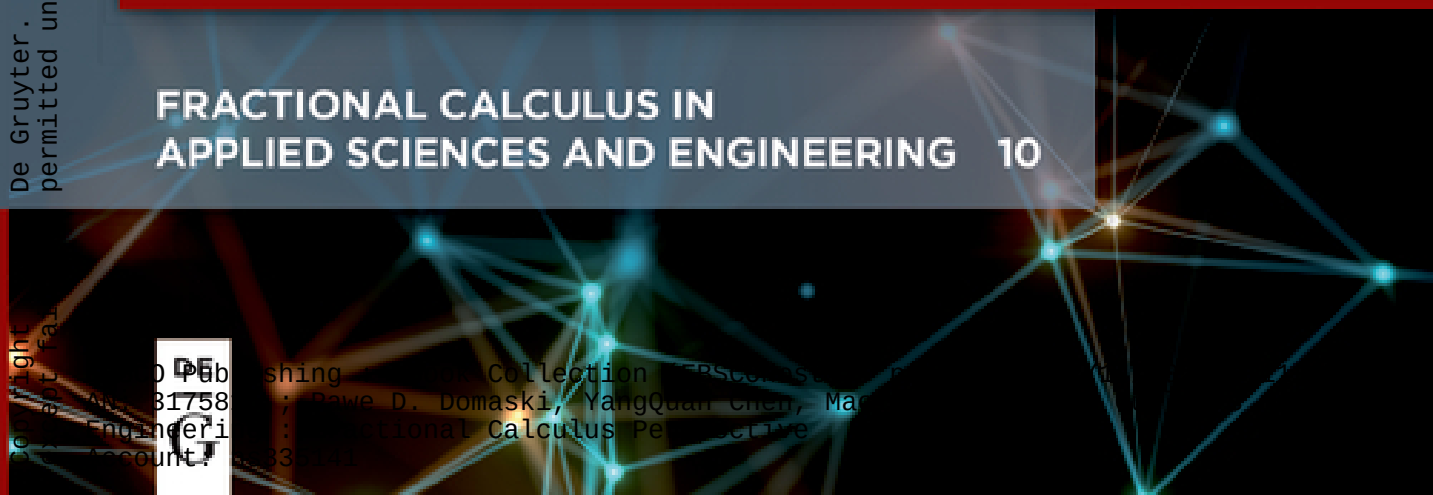
DE GRUYTER

OUTLIERS IN CONTROL ENGINEERING

FRACTIONAL CALCULUS PERSPECTIVE

Edited by Paweł D. Domański, YangQuan Chen, Maciej Ławryńczuk

FRACTIONAL CALCULUS IN APPLIED SCIENCES AND ENGINEERING 10



D Publishing Collection ERSc
31758 ; Pawe D. Domaski, YangQuan Chen, Ma
Engineeri : Fractional Calculus Pe ective
Account: 33

Paweł D. Domański, YangQuan Chen, and Maciej Ławryńczuk (Eds.)
Outliers in Control Engineering

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Fractional Calculus in Applied Sciences and Engineering

Editor-in Chief
Changpin Li

Editorial Board
Virginia Kiryakova
Francesco Mainardi
Dragan Spasic
Bruce Ian Henry
YangQuan Chen

Volume 10

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Outliers in Control Engineering

Fractional Calculus Perspective

Edited by
Paweł D. Domański, YangQuan Chen, and
Maciej Ławryńczuk

DE GRUYTER

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mathematics Subject Classification 2020
26A33, 93C83, 60Gxx, 62G32, 28A80

Editors
Prof. Paweł D. Domański
Warsaw University of Technology
Institute of Control and Computation Engineering
ul. Nowowiejska 15/19
00-665 Warsaw
Poland
pawel.domanski@pw.edu.pl

Prof. Maciej Ławryńczuk
Warsaw University of Technology
Institute of Control and Computation Engineering
ul. Nowowiejska 15/19
00-665 Warsaw
Poland
maciej.lawrynczuk@pw.edu.pl

Prof. YangQuan Chen
University of California
Dept. of Mechanical Engineering Electrical Eng and Computer Sciences
5200 North Lake Rd
Merced 95343
USA
ychen53@ucmerced.edu

ISBN 978-3-11-072907-8
e-ISBN (PDF) 978-3-11-072912-2
e-ISBN (EPUB) 978-3-11-072913-9
ISSN 2509-7210

Library of Congress Control Number: 2021950261

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

© 2022 Walter de Gruyter GmbH, Berlin/Boston
Cover image: naddi/iStock/thinkstock
Typesetting: VTeX UAB, Lithuania
Printing and binding: CPI books GmbH, Leck

www.degruyter.com

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

to my beloved Princess, unceasingly
Paweł D. Domański

mom, family, mentors
YangQuan Chen

to my children, young scientists
Maciej Ławryńczuk

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Preface

The origin of this edited book can be traced back to the IFAC 2017 World Congress, Toulouse, France. The 20th World Congress of the International Federation of Automatic Control happened from 9–14 July, 2017. Y.C. was visiting the poster session and was attracted to a poster entitled "Non-Gaussian Assessment of the Benefits from Improved Control"[1] presented by P.D. Y.C. immediately recognized this work is valuable and instantaneously related this paper to fractional calculus thus he started conversation with P.D. and tried to inform him this connection. This paper abstract says that "Control quality significantly contributes to the process technical and financial performance. Productivity, environmental issues and energy management push systems towards their technological constraints calling for better regulation closer to process limitations. Any control improvement initiative should be predated with the estimation of the potential benefits associated with control rehabilitation project. The assessment is based on the appropriate measures. Classical methods are based on Gaussian approach. However, investigation of industrial data frequently is not compliant with normal assumption on control signals. This paper proposes approach using non-Gaussian probabilistic distributions like Cauchy, Laplace and Lévy. The methodology is illustrated on the exemplary industrial data." with keywords "monitoring; performance assessment; statistical data analysis; non-Gaussian distributions; same limit method; control benefits."

Control engineering perspective of this work is enhanced by M.Ł., whose expertise in advanced control together with the participation in several industrial projects have increased the domain of our research. Since then, with extensive exchange of information and several joint works, we have gained deeper understanding on the role of outliers in control engineering as the outliers are connected to fractional calculus (FC) via heavytailedness (FT).

We wish to thank all contributors of this book who are directly or indirectly dealing with outliers in control engineering and also control theory. Furthermore, special thanks go to Professor Changpin Li of Shanghai University, China, the editor-in-chief of the De Gruyter book series on "Fractional Calculus in Applied Sciences and Engineering"[2] for encouraging us to move forward with a book proposal. We are grateful to Leonardo Milla, Editor Engineering/Computer Science, De Gruyter for peer review of our book proposal and the final offer of book project contract during the pandemic!

1 See full paper at https://doi.org/10.1016/j.ifacol.2017.08.753

2 https://www.degruyter.com/serial/FCASE-B/html

https://doi.org/10.1515/9783110729122-201

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Last but not least, we are thankful to our family members for their patience and support in this trying time.

January 2022

Paweł D. Domański, Warsaw
YangQuan Chen, Merced
Maciej Ławryńczuk, Warsaw

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Contents

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Part I: **Foundations**

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Paweł D. Domański, YangQuan Chen, and Maciej Ławryńczuk

1 Outliers in control engineering—they exist, like it or not

Abstract: Major achievements in control engineering, like the least squares estimation, minimum variance, LQG (Linear Quadratic Gaussian), MPC (Model Predictive Control) or adaptive control strategies rely on the assumption that process disturbances exhibit Gaussian properties. Though this applies to the original algorithms' formulations and modifications relaxing this assumptions exist, these conditions are rarely verified. On the other hand, anyone who has worked in industry is very often confronted with signals and processes that do not meet these conditions. Analysis of real industrial signals very often discloses their non-Gaussian properties, which manifests itself in more or less frequent anomalies. The easiest method to identify them is through statistical analysis as they are responsible for the tails in distributions. Therefore, statistical analysis should be considered as the appropriate analytical tool. Fat-tail properties are often closely related to the fractional properties of the process behind this data and fractional calculus complements statistical methods, giving further insight into the generating mechanism. Outliers do not have to be artificial exceptions. Their observation and analysis may enrich our knowledge and give new, and open up further, perspectives for our research curiosity.

Keywords: outliers, control engineering, fat tails, tail index, fractional calculus

> "Corollary to Moore's Law: every ten years, collective wisdom degrades by half."
>
> Nassim N. Taleb, *The Bed of Procrustes*

1.1 Introduction

Outliers exist, not only in the social or natural sciences. We also witness them in the technical sciences, and control engineering is not an exception. In the beginning they were considered as strange, occasional events. The fact that they are rare might cause them to go unnoticed. Unfortunately, there is another feature associated with the outlier. The relevance of such an event might be abnormally large, carrying a dispropor-

Paweł D. Domański, Maciej Ławryńczuk, Warsaw University of Technology, Institute of Control and Computation Engineering, Warsaw, Poland, e-mails: P.Domanski@ia.pw.edu.pl, M.Lawrynczuk@ia.pw.edu.pl
YangQuan Chen, Department of Mechanical Engineering, University of California, Merced, USA, e-mail: ychen53@ucmerced.edu

https://doi.org/10.1515/9783110729122-001

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

tionately high cost. Sporadic high-magnitude earthquakes and tsunamis, devastating tornadoes, damaging earthquakes or volcanic eruptions, economic crises or pandemic outbreaks are extremely rare events from the historical perspective, but they are well remembered by mankind for their destructive effects. An outlier is often called by other names, such as: anomaly, deviant, oddity, contamination, aberration, an exception or fringelier reflecting "*unusual events which occur more often than seldom*" [91]. Whatever name is used, they are rare and of unknown origin.

An outlier in control engineering is not so spectacular as the eruption of Vesuvius, however its impact still might be significant. Erroneous measurement observation in a control system may cause system malfunctioning, a plant shutdown or device destruction. It's only a matter of proportionality.

Human perception is constructed in such a way that we tend towards simplifications, closing our eyes to rare and strange phenomena. We reject extremes and prefer more comfortable surroundings. Our attention is raised only after the fact, and usually by their consequences. We prefer simple linear time-invariant transfer functions over nonlinear partial differential equations, despite the fact that the process is highly nonlinear and dispersed. Similarly, we use the Gaussian normality assumption, though we certainly know that the empirical dataset distribution is not even close to the bell-shape. Yet, we do that. Why? There are many reasons, among which may be that it is a quick theory ready at hand, because everyone does the same, due to education, best practices and procedures, *the power of tradition*, etc.

Actually, academia rarely prepares us for their existence. Once we put our hands on real projects, we start to notice that something does not fit, that theory does not want to suit the practice. As our experience accumulates, we notice that theory aligns with practice only in theory. Closing our eyes is no longer a good option. The ad hockery is too close to ignorance as we start to dig into the subject. What happens? What's the reason? What are the artifacts doing here? It seems that we have to leave our comfort zone.

We see strange observations in data, and we begin to wonder about their cause and source. We start thinking about what to do with them and, most importantly, we try to understand whether they are just random errors or whether they carry important information. The research begins. This study aims at giving just a general picture of the subject. It tries to identify the most significant aspects that require deeper attention, showing the beauty of research challenges. The story begins in Section 1.2 with the very fact of observing an outlying observation. Once we have seen it, we may try to identify its sources and origins (Section 1.3) and further to find a way how to detect it (Section 1.4). Once the outlier is known, i. e., identified and labeled, we may propose an action appropriate to the situation, like the ones proposed in Section 1.5. The study concludes with Section 1.6 containing practical comments identifying open research areas for further research.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

1.2 Observation: to remove or not to remove

Outliers in data cause various reactions among observers. Some people pretend that everything happens according to the initial *normal and safe* assumptions, i. e., as if nothing had happened. We notice an unknown observation, some unknown unknowns, and we decide to remove it from our comfort zone because it may disturb our approach by challenging the results obtained so far. Actually, this is not a solution. It sets us back rather than pushing the research forward. Should we always try to falsify existing hypotheses rather than to confirm them by force [71] We always have to be aware that things exist that have never been seen before, and if we did not see them previously, it does not mean that they do not exist. Once aware of them, we mustn't deny their existence and remove them from the analysis.

Once we see and accept outliers, we may take them into account in two ways. They can be treated as erroneous observations resulting from system or human errors, in which case they shall be found and removed. This happens in the case of bad measurement, human errors, system breaks, etc. In such a situation outliers' detection serves to label and remove them [29, 73, 94]. Actually, this approach is the most popular and considered a common reaction. However, one important should be raised here—the removal of outliers impacts further analysis and obtained results [70].

In the latter case they are considered informative as a potential focus for further analysis. Data contamination is considered as a source of important information, and one should focus on their analysis. Such a situation happens in the case of fraud or leakage detection, medical diagnosis, cybersecurity, etc. [58, 64, 72].

As was briefly just mentioned, we may treat outliers in various different ways: we may remove them or not; we may ignore them or investigate. There are dozens of policies for each decision. There are essential arguments for their rejection. On the contrary, if we suspect that they carry on important information or that they present relevant part of the population, they mustn't be removed. Nonetheless, they should be isolated for further investigation. As one can see, the analysis of outliers includes challenging activities: detection, labelling, interpretation and treatment.

The story of their investigation begins with the works of Bernoulli (1777) who pointed out the practice of deleting them. Probably, this tradition is still alive as it is customary to remove them. The first statistical approach to treat outliers appeared in 1850 [7], but two opposing practices have existed from the beginning. In the 19th century, Boscovitch suggested to delete them through the so-called *ad hoc adjustments* similar to Pierce (1852), Chauvenet (1863) or Wright (1884). The removal practice exists today and is still used and proposed [12].

An opposing group argues that these anomalies provide useful information and should be kept. Legendre (1805) recommended not to erase the extreme observations "*adjusted too large to be admissible*". Bessel and Baeuer (1838) claimed that deleting leaves a gap in the data that is artificial. As one can see, the dilemma to remove or not to remove outliers is still an issue today, as it was 200 years ago.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

The first formal outlier definitions were proposed by Dixon [14], who called them "*dubious in the eyes of the researcher*',' or by Wainer [91], who called them *contaminants*. The most popular definition was formulated by Hawking [33], who called an outlier "*an observation which deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism*". Johnson [45] defined it "*as an observation in a data set which appears to be inconsistent with the remainder of that set of data*", while Barnett [5] noted that "*an outlying observation, or outlier, is one that appears to deviate markedly from other members of the sample in which it occurs*".

Although some of the above comments seem to be obvious, especially for practitioners, it is worth to highlight these issues. Once we agree, we should first, before any further data analysis, check whether our simplifying assumptions hold and whether the data exhibit outliers. The next paragraph brings us closer to the sources and causes of outliers showing up and how they manifest themselves in the data, while opening gates to the methods for their detection and labelling.

1.3 Sources and origins

An outlier is a strange being and as the name suggests it comes from outside the system. There are as many meanings and interpretations of outliers as there are perspectives for their observation and analysis. Outlying data might have a significant impact on data analysis, whatever it may be. Outliers increase the data variance and reduce the power of statistical tests [63]. They deteriorate data statistical properties, like Gaussianity, and introduce tails into the distribution [82] and bias regression analysis [77].

Once we agree that outliers exist and we notice them, then we can start to investigate their origins [8, 63]. Generally, we have two options. The first supposition assumes that they originate from erroneous observations. Such data artifacts do not come from the process itself, rather their generating mechanism is unknown and artificial to the data source. However, it is still an open question whether they carry information or not. Aberrant observations come out of human errors, intentional or motivated activities, wrong computer systems (data measurement, control, sampling or collection) operation or from standardization failures. Identification and treatment of such incorrect observations is not simple, but procedures, double checking, modelling, validation or recalculation may help.

The second option is that they belong inherently to some complex, nonlinear and not fully known process. We have to remember that the characteristics of the process is such that outliers can be originated by the process itself. Process complexity, nonlinearity and their not fully understood nature can additionally cause incorrect interpretations and assumptions about the data properties that possibly cause outliers [43]. Such systems can cause multimodal, skewed, asymmetric, heavy-tailed, flat or very unconventional distributions, depending not only on the generating mechanism, but

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

also on data sampling. Complex processes may generate data with properties other than originally assumed and characterized by tails [6]. Furthermore, there might be more than one underlying generating mechanism. Data can be infected by nonstationarities and trends, external cross-correlations with variable delays, and fractal or multifractal properties [56]. In fact, observed outliers might be just rare, but natural, consequence of the process. All these reasons cause problems and mislead control engineers who are accustomed to linear, quadratic and Gaussian simplifications.

It is worth to investigate how these general reflections manifest themselves in control engineering. First of all, we have to remember that control cannot exist without the process and that each process property is reflected. Moreover, physical control system is just an IT or embedded system being subject to anything that might happen to it.

Erroneous unintentional outliers that are subject to the forthcoming removal can be caused by the control system (e. g., measurement unit breaks, data communication jams, calculation failures, numerical errors, actuator equipment breakdowns) or by human impacts (misspelled data, wrong interpretation of displayed data or misuse of the equipment). Apart from random incidental aberrations of data originating from unknown reasons, intentional anomalies caused by cyberattacks might occur. They are intended to be very dangerous, and they differ from previously mentioned causes. Unlike unintentional erroneous data, which can be just deleted and do not require in-depth attention, contamination intentionally caused by cyberattacks must be identified, labeled, isolated, traced and immediately counter-reacted.

The second group of outliers, i. e., those that are generated by internal process mechanisms, are not so straightforward to interpret. It might be that the process itself exhibits non-Gaussian properties that frequently generate data lying far from the bulk of the data. Another reason may be that two different mechanisms are interacting, even if each of them witnesses only normal Gaussian properties. Cross-correlations with varying delays may also cause fat and heavy tails containing outlying observations. One has to be aware that negative feedback, which is fundamental to each control system, also contributes to the tails [15, 16, 22].

1.4 Detection and labeling

The simplest method of outlier detection is just visual inspection. Process time trends can be plotted and visually inspected. Any outlier can be identified using human expert knowledge and delayed. The review in the time domain can be supported by the investigation of the statistical data histograms [21]. Such manual methods can be improved by automated analytical methods. Initially, statistical approaches have been investigated [31, 44, 65, 68]. The evaluated methods have applied Gaussian normality assumptions exploiting different properties of the normal probabilistic distribution function (PDF) [33, 34, 74, 77]. Formally, a statistical approach follows three steps [43]:

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

labeling (flagging for further investigation), accommodation with robust statistics and finally identification.

There are numerous statistical algorithms, and the decision whether the observation can be flagged as an outlier depends on the underlying data distribution. The majority of research concentrates on univariate Gaussian analysis. Control engineering is not an exception. Robust regression research [40, 77] shows that classical least squares regression is sensitive to even a single outlier. This observation impacts further research using robust regression estimators, such as Z-scores with median and MADAM (median of the absolute deviations about the median) [43], Hampel filter or power-law tail index estimator [81]. Finally, the assumption that an α-stable distribution is responsible for the underlying generation mechanism opens new perspectives [19] because it may be considered as a signal-generation mechanism in control systems [15]. It is characterized not only by scale and location, but also by stability and skewness coefficients.

Recently, research has started to exploit artificial intelligence (AI). In this regard, there are three different types of algorithms:

- supervised methods apply in training historical normal and abnormal patterns,
- semi-supervised methods use in learning only normal or abnormal patterns,
- while training examples are not used by the unsupervised methods.

Other classification takes into account the resolution of the reference, meaning the difference between local versus global ranges. However, many methods use hybrid approaches and cannot be clearly classified. Next, classification takes into consideration algorithm output. Labeling exhibits binary output, naming the objects normal or abnormal. Scoring approaches give continuous output, as for instance the outlying probability.

The issue of statistical detection of outliers traces back to the 19th century and works by Peirce (1852) [68]. The research has continued, and the literature is rich. However, currently this approach is not so popular. Nowadays, AI-based algorithms are most frequently used [32, 96]. However, it must be remembered that statistics shares formal simplicity and a rigor that is very attractive. Especially interesting are recent findings in the area of non-Gaussian and robust methods. [37, 38, 76, 84] gave a new impulse by improving robustness and reliability. Reviews of statistical outlier analysis have been assessed by many authors [4, 32, 51, 87].

Statistical approach depends on assumed model of the underlying probabilistic process. Because there are so many methods, one has to select those that are the most appropriate in situation under consideration [20, 21]. Selected methods are listed next, including well-known algorithms and recent modifications:

(1) Z-Scores methods [43], also called MDist [51], considered the earliest and the most common, may use different scoring approaches as, for instance: 3σ, modified $0.6745 \cdot 3\sigma$ or $0.6745 \cdot 3.5\sigma$, application of robust scale estimators, like MAD, MADAM, robust M-estimators [40, 88], L-moments [78].

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

(2) higher-order moments [3] or joint statistical moments [1],
(3) InterQuartile Range (IQR) method [93] using standard Q1 and Q3 quantiles or modified ones like 5 % [13] or 0.5 % [2].
(4) Minimum covariance determinant (MCD) analysis [41],
(5) Extreme Studentized deviate (ESD) test [75],
(6) Grubbs' [31] and Tietjen–Moore test [86],
(7) Hampel filter [66],
(8) Thompson Tau test [85],
(9) Cupola-based method [54],
(10) analytical tail modelling method using Gaussian or power-law tail starting point estimators [81] or α-stable distribution coefficients [21],
(11) numerical tail index estimators [9, 79].

This list includes methods used for symmetric or not-too-much skewed two-sided distributions. Such assumptions could have been made because the majority of control engineering applications and data subject to the analysis exhibit such properties. In general this is the unwritten supposition. Nonetheless, we have to remember that there might exist engineering tasks where this assumptions does not hold. Such data exhibit a single-sided distribution model, like for instance Gamma, lognormal, exponential, Weibull, Generalized Extreme Value, generalized Pareto or the four-parameter Kappa probabilistic density function. In such cases outliers' investigation, often referred to as extreme value analysis, may use other methods and aim at other goals [11, 47], as for instance extreme values frequency and repeatability, peaks-over-threshold, flow duration, return period analysis, etc.

The conclusion is that outliers deteriorate Gaussian assumptions. They do so in two ways: They change properties of the stochastic process and simultaneously they bias classical statistical moments estimators. This directs the research towards fat- or heavy-tailed approaches. This observation naturally opens new research opportunities. One should take into account observations and results from other research areas and contexts, as for instance persistence and fractal analysis, non-Gaussian statistics, robust estimators, various approaches to the tail analysis [28], and fractional calculus, among others.

1.5 Action

As one can see, the subject is vast, and what is the most challenging, the mathematical definition of the outlier per se seems to be not formulated yet [49]. Due to the natural text volume limitation and authors personal preferences and dislikes, further considerations focus on two issues: fat- and heavy-tails (FT and HT), and fractional-order calculus (FC).

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

1.5.1 Fat- and heavy-tails

The distribution's tails are built out of very small values on the left side and very large ones to the right, finally defining the distribution shape. According to the traditional definition, these are PDFs with fatter tails than the normal Gaussian, which means that there more realizations within one σ and with kurtosis higher than 3. Once we take into account the fact that classical moments are biased by extreme values (outliers) and many heavy-tailed functions do not have finite moments, we need robust formulation. L-moments give us such an opportunity [37]. They are formulated as linear combinations of order statistics. In contrast to product moments, they produce almost unbiased statistics, even for very small samples simultaneously less sensitive to outliers [67]. L-kurtosis $\tau_4 \in (-1/4, 1)$ might be used as an alternative. The value for L-kurtosis for the Gaussian distribution equals $\tau_4 = 0.1226$. Therefore, we can assume that distributions with $\tau_4 > 0.1226$ are leptokurtic and can be called fat-tailed. Many distribution classes belong to this group, like the power law, slowly varying and the Pareto tail class.

Heavy-tailed distributions are more heavy than fat-tailed ones Al their moments are infinite above some order. A random variable with some PDF can be called heavy-tailed if its tail Pareto-like decay is above some threshold [10]. Gaussian distribution is considered not to be fat-tailed. Fatter tails appear with the subexponential class, which has all moments (like for instance lognormal distribution). For the supercubic class (Lévy-stable distribution), there is no variance nor higher moments. We call these functions power or Pareto laws. In more extreme cases, even a mean does not exist [82].

An interesting notion of ostensible heavy tails was introduced in [50]. Klebanov and Volchenkova use this name for a distribution that exhibits "*high pike near mean value, thin or truncated "far" tail and The part of distributions body outside of the pike changes not too essentially with the remotion from the pike on a distance, but turns into thin tail after that*". Sample distributions exhibiting various tail fatness are depicted in Fig. 1.1. The figure shows only the right side for clearness.

It is interesting to notice that observation of the control-engineering data reveals the fact that Gaussian cases are in the minority. Reviews shows that the share of variables' normal properties is significantly low (below 10 %) [15]. Therefore the analysis of tails is justified, and HT distribution may be considered as natural models for control-engineering problems. The tails and observations responsible for their appearance might have a negative nature, like erroneous observations or intentionally injected contamination. On the contrary, other outliers might just reflect some specific properties of the stochastic data generating process, and they should not be considered negatively. They are be just a natural consequence of the process behind the data.

Another issue raised in [81] draws attention to the implications of the central limit theorem and limit distributions. Unknowingly, we tend to assume that the empirical data distribution obtained from available observations is the right one, i. e., equal to

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

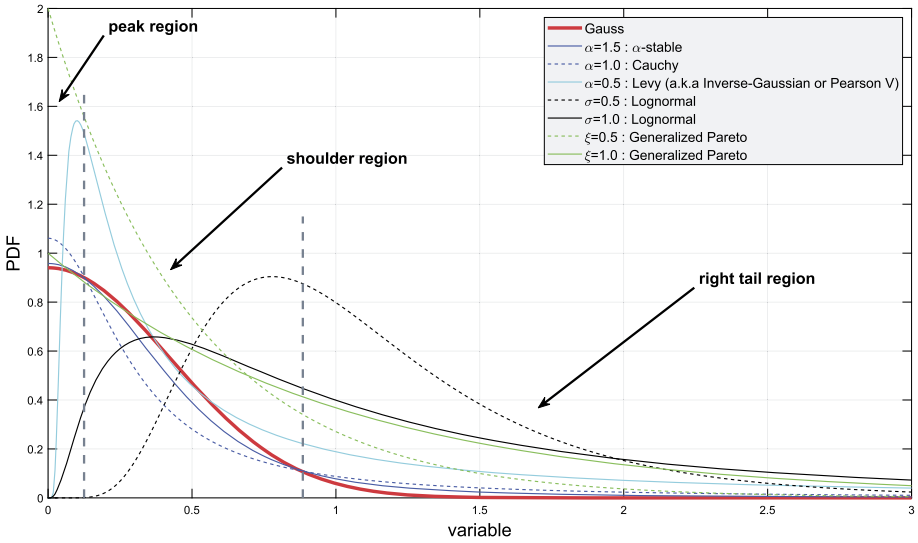


Figure 1.1: The tails heaviness versus reference no-tail Gaussian distribution for sample distributions. The regions are named according to [82].

the underlying theoretical limiting distribution. However, this does not have to be the true. We never know whether the available number of data points is high enough from the perspective of the limit theorem. This is especially due to the fact that the rate of convergence for fat-tailed distributions is not fast and slower than for the normal one. Asymmetries even make the problem harder. Therefore, the scientist or data analyst has to be very cautious because she/he has to make several proper decisions that may significantly affect the result.

The literature shows many different ways to analyze the tails. The most common statistical approach using the kurtosis can be applied in a narrow class of problems, i. e., when all moments are finite. It enables simple comparison and difference measures from Gaussian. For fat-tailed cases, one may use tail index or tail exponent. Another approach is to apply extreme value theory and use peaks-over-threshold or exceedance probability methods. Finally, one can try to measure the distance from the limiting distribution, assumed as the α-stable one [82].

The tail index, denoted α, is the shape parameter determining how fat the tail is [13]. Its higher values correspond to thinner tails. The use of the tail index gives us a lot of freedom. The Hill quasi-maximum likelihood estimator [35] is the most often used, especially due to its simplicity. As we never know how close we are to the limiting values, it is unclear what should be the number of observations because it may bias the estimation of α. The wrong choice may cause a trade-off between its variance and the bias of the estimator.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1.2 shows a sample Hill plot (of the discharge level of a river in Poland), i. e., dependence of the tail index on the number of highest observations. It is clearly visible that the estimate varies with the data number quite substantially. Estimator volatility decreases with number of observations.The literature shows many different methods for selecting the proper number of observation, thereby determining the tail index α [13]. Apart from the Hill estimator, there are many others, like for instance minimal distance, bias-reduced, generalized class, stable distribution tail index, small sample and robust, peaks-over-random threshold, parametric methods or estimators allowing negative values. Comprehensive description and comparison of these methods can be found in [28].

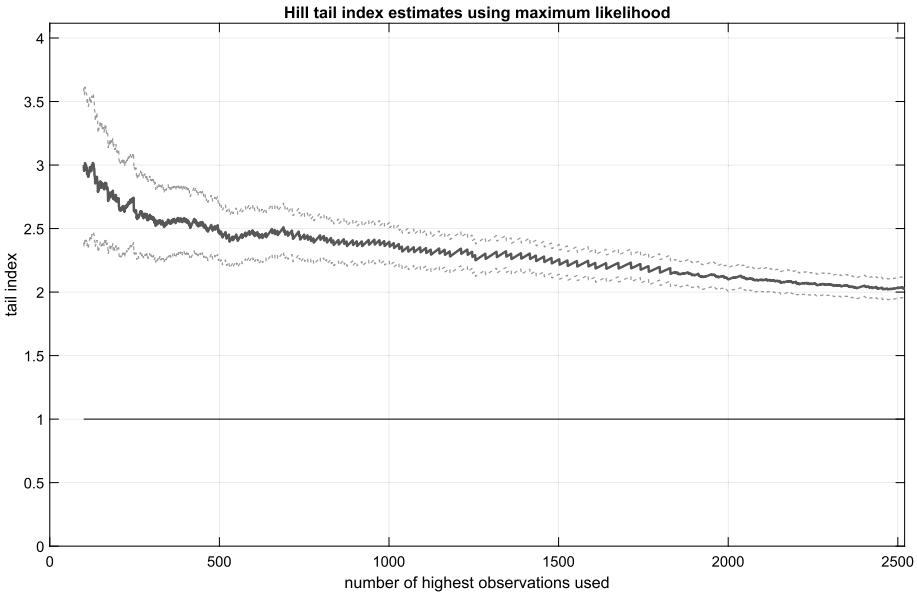


Figure 1.2: Sample Hill plot showing tail index α estimates for various numbers of the highest observations.

Although the literature on the tail index theory and applications is vast, there are only a few works on its application to the task of outlier detection. Therefore, it seems that this direction might happen to be a promising opportunity for practical and challenging research.

1.5.2 Fractional calculus

Fractional-order calculus (FC) is about differentiation and integration of non-integer orders. Using "fractional" is actually a misnomer, but for historical reasons we con-

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

tinue to use FC, meaning that the orders can be any non-integers that can be real or complex numbers. Denying fractional calculus is like denying there are non-integers in between integers. This type of "in-between thinking" is called "Fractional-Order Thinking" when we study complex dynamic systems. Fractional calculus-based fractional-order thinking (FOT) has been shown to help us to better understand complex systems, better process complex signals, better control complex systems, better perform optimizations and even better enable creativity [92].

Operators definitions, which are recalled in the following [48, 69]:

– Grunwald–Letnikov's fractional-order derivative/integral definition:

$$ {}_a^G D_t^\alpha f(t) := \lim_{h\to 0} \frac{1}{h^\alpha} \sum_{j=0}^{[(t-a)/h]} (-1)^j \binom{\alpha}{j} f(t-jh), \quad (\alpha \in \mathrm{R}) $$

– Riemann–Liouville's fractional-order integral definition:

$$ {}_a^R D_t^{-\alpha} f(t) := \frac{1}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} f(\tau) d\tau, \quad (\alpha > 0) $$

– Riemann–Liouville's fractional-order derivative definition:

$$ {}_a^R D_t^\alpha f(t) := \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \left[\int_a^t (t-\tau)^{n-\alpha-1} f(\tau) d\tau \right], \quad (n-1 < \alpha < n) $$

– Caputo's fractional-order derivative definition:

$$ {}_a^C D_t^\alpha f(t) := \frac{1}{\Gamma(n-\alpha)} \left[\int_a^t (t-\tau)^{n-\alpha-1} f^{(n)}(\tau) d\tau \right], \quad (n-1 < \alpha < n) $$

– The Riesz fractional derivative is expressed as

$$ \frac{\partial^\alpha}{\partial |x|^\alpha} f(x) = -\frac{1}{2\cos(\pi\alpha/2)} [I_+^{-\alpha} f(x) + I_-^{-\alpha} f(x)], \quad 0 < \alpha \leqslant 2, $$

where

$$ \begin{cases} I_+^\alpha f(x) = \dfrac{1}{\Gamma(\alpha)} \displaystyle\int_{-\infty}^{x} (x-\xi)^{\alpha-1} f(\xi) d\xi, \\ I_-^\alpha f(x) = \dfrac{1}{\Gamma(\alpha)} \displaystyle\int_{x}^{+\infty} (\xi-x)^{\alpha-1} f(\xi) d\xi. \end{cases} $$

Based on these definitions, fractional-calculus equations that can more accurately describe behaviors of real physical phenomenon and systems have become a hot topic

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

in recent decades [80]. A fractional derivative provides a perfect tool when it is used to describe the memory and hereditary properties of various materials and processes. This is the main reason that fractional differential equations are being used in modeling mechanical and electrical properties of real materials, rheological properties of rocks and many other fields. As an important application field of fractional calculus, fractional-order control and systems has become the work of many researchers. A traditional fractional-order differential equation that can describe the fractional-order system's dynamical properties has the following form:

$$F(x, {}_0D_t^{\alpha_1}y, {}_0D_t^{\alpha_2}y, \ldots, {}_0D_t^{\alpha_n}y) = 0,$$

where ${}_0D_t^{\alpha_i}$, $(i = 1, \ldots, n)$ can adopt Riemann-Liouville's or Caputo's definition. For fractional order systems, control and signal processing, we suggest the books by Monje et al. (2010) and Sheng et al. (2012), respectively, and more fractional calculus books can be found in the link.[1]

The deterministic world has almost everywhere the smell of fractional calculus when the world is complex. But, this is only one side of the coin. The other side of the coin is the stochastic world. In the previous subsection, we encountered a specific class of fractional randomness in which we can observe the inverse power law in the sense of PDF (probabilistic density function), ACF (autocorrelation function) or PSD (power spectral density). In deterministic models, it is not hard to understand the connection of the inverse power law (IPL) and fractional calculus due to the asymptotic behavior of Mittag–Leffler functions. It should be pointed out that integer-order models correspond to the "exponential law." However, it is not quite straightforward to understand the connection between heavy-tailedness and fractional calculus.

Fortunately, Gorenflo and Mainardi's excellent tutorial paper [30] explains in detail the connection between fractional calculus and stable probability distributions, including most of the cases in the previous subsection. Even more interestingly, the reader is suggested to consult the monumental review for more insights into the connection between fractional calculus and fractional (or complex) randomness [59].

It is quite sure that fractional calculus can help us to better characterize outliers in general.

1.5.3 Case study: the MPC case

Analysis of industrial data indicates that the majority of controlled variables exhibit non-Gaussian behavior, and significant outliers are very frequent. To solve these problems, it is recommended to carefully choose and tune the control algorithm and appropriately select the control-performance indicators.

1 https://mechatronics.ucmerced.edu/fcbooks

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

When the classical PI or PID controllers are used, the objective of control performance assessment is to determine whether or not they are tuned properly. Let us stress that, even in the case of multiple-input multiple-output processes, the typical industrial approach is to use a set of single-loop PI or PID controllers. Such an approach greatly simplifies the controllers' performance assessment because each control loop is analyzed separately. For each controller, we have only two or three tuning parameters. Typically, as a result of the analysis, we obtain a set of quite simple to implement conclusions:

- reduce/increase the controller's proportional gain,
- reduce/increase the controller's integration time-constant,
- reduce/increase the controller's derivative time-constant,
- switch on/off the derivation.

One advanced control approach, namely Model Predictive Control (MPC) [55, 83], is becoming more and more popular in practice. In contrast to the classical PI and PID controllers, a dynamical model of the process and an optimization procedure are used online in MPC. The model calculates predictions of controlled variables (and state variables if necessary) over some time horizon, called the prediction horizon. The numerical optimization procedure calculates the best possible control policy that minimizes a predefined performance cost-function. Prediction and optimization are repeated at each sampling instant. Assuming that the number of process manipulated variables (inputs) is denoted by n_{u} and the number of controlled variables (outputs) is denoted by n_{y}, the classical cost-function is

$$J(k) = \sum_{p=1}^{N} \sum_{n=1}^{n_{\mathrm{y}}} \psi_{p,n} \left(y_n^{\mathrm{sp}}(k+p|k) - \hat{y}_n(k+p|k) \right)^2 + \sum_{p=0}^{N_{\mathrm{u}}-1} \sum_{n=1}^{n_{\mathrm{u}}} \lambda_{p,n} \left(\triangle u_n(k+p|k) \right)^2. \tag{1.1}$$

The first part of the cost-function measures the control errors predicted over the prediction horizon N. The set-points for the future sampling instant $k+p$, known at the current sampling instant k, are denoted by $y_n^{\mathrm{sp}}(k+p|k)$, where the index n indicates the output number. Similarly, the predicted values of the process outputs for the future sampling instant $k+p$, determined at the current sampling instant k, are denoted by $\hat{y}_n(k+p|k)$. The role of the second part of the cost-function, i. e., the penalty term, is to minimize unwanted large changes of the manipulated variables over the so-called control horizon N_{u}.

As far as the MPC optimization carried out at each sampling instant k is concerned, the cost-function $J(k)$ is minimized subject to some constraints. Typically, the constraints are imposed on the magnitude and the rate of change of the manipulated variables. Additionally, in some applications, it is beneficial to impose constraints put

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

on predicted values of the output variables. All things considered, the typical MPC optimization problem is

$$\begin{aligned}
&\min_{\triangle \boldsymbol{u}(k)} \{J(k)\} \\
&\text{subject to} \\
&u_n^{\min} \le u_n(k+p|k) \le u_n^{\max}, \quad n = 1, \ldots, n_{\mathrm{u}},\ p = 0, \ldots, N_{\mathrm{u}} - 1 \\
&\triangle u_n^{\min} \le \triangle u_n(k+p|k) \le \triangle u_n^{\max}, \quad n = 1, \ldots, n_{\mathrm{u}},\ p = 0, \ldots, N_{\mathrm{u}} - 1 \\
&y_n^{\min} \le \hat{y}_n(k+p|k) \le y_n^{\max}, \quad n = 1, \ldots, n_{\mathrm{y}},\ p = 1, \ldots, N,
\end{aligned} \tag{1.2}$$

where $u_n^{\min}$ and $u_n^{\max}$ define the constraints imposed on the magnitude of the manipulated variables, respectively; $\triangle u_n^{\min}$ and $\triangle u_n^{\max}$ specify the limitations related to changes of the manipulated variables; $y_n^{\min}$, and $y_n^{\max}$ defines the constraints put on the predicted values of the controlled variables. The decision vector, $\triangle \boldsymbol{u}(k)$, is defined as increments of the manipulated variables over the control horizon, which means that the actual number of computed decision variables is $n_{\mathrm{u}} N_{\mathrm{u}}$. Although in MPC at each sampling instant the whole decision vector $\triangle \boldsymbol{u}(k)$ is found as a result of online optimization, only its first n_{u} elements, i. e., the increments for the current sampling instant k, are actually applied to the process. In the consecutive sampling instants, the entire prediction and optimization procedure is repeated.

Let us briefly mention two essential advantages of MPC. First, due to problem formulation, i. e., the calculation of the control policy from an optimization task, it is possible to control multivariable processes with numerous manipulated and controlled variables. Interestingly, the numbers of process inputs and outputs may be different. Let us mention that such processes are always difficult to handle when a set of single-loop PI or PID controllers are used. Secondly, it is very easy to enforce satisfaction of constraints in MPC since they are simply parts of the optimization task solved at each sampling instant. Although simple limiters may be used in the case of the classical PI and PID structures, satisfaction of some constraints imposed on the predicted variables is not possible. As a result of their advantages, MPC algorithms have been applied to many processes. Typically, MPC methods have been used in industrial control; example applications are chemical reactors [57] and distillation columns [42]. In addition to that, due to the availability of fast hardware platforms, MPC algorithms are more and more popular in embedded control systems; example applications are fuel cells [53], combustion engines [46], robots [39], electromagnetic mills [62], electromechanical systems [95] and servomotors [36].

On the one hand, in general, MPC algorithms have numerous advantages over the classical PI/PID structures and a great potential. On the other hand, it is necessary to emphasize the difficulties of control performance assessment of MPC-based control systems. The reasons for that are the following:

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. In the case of the PI and PID controllers, we have only two or three tuning parameters, respectively. Conversely, in MPC, we have many more parameters that affect the control quality:
   - the prediction horizon N,
   - the control horizon N_u,
   - the weighting coefficients $\psi_{p,n} \geq 0$ that prioritize the predicted control errors of the consecutive controlled variables, $n = 1, \ldots, n_y$, and for the consecutive sampling instant over the prediction horizon, $p = 1, \ldots, N$,
   - the weighting coefficients $\lambda_{p,n} > 0$ that prioritize the future increments of the consecutive manipulated variables, $n = 1, \ldots, n_u$, and for the consecutive sampling instant over the control horizon, $p = 0, \ldots, N_u - 1$.
2. Let us recall that the dynamical model of the controlled process has a fundamental role in MPC since it is used for prediction. MPC algorithms are likely to work successfully, and in many difficult cases they are able to achieve much better control performance than the classical PI/PID-based control structures, but the model utilized for prediction must be precise. Unfortunately, when the model is only a rough approximation of the process, the predictions computed from the model and the real values of the process state and/or output variables differ significantly. As a result, MPC algorithms may produce poor control quality. Of course, all MPC algorithms have a negative feedback mechanism and the integral action that compensate for model inaccuracies and external process disturbances. Typically, when the model–process mismatch is not significant, no steady-state error is obtained. Unfortunately, when the model–process mismatch is huge, the predictions are wrong, and the values of the manipulated variables calculated using such erroneous predictions do not yield good control quality.
3. The process–model mismatch may be observed not only when the model is wrong but also when the nonlinear model is successively linearized online to formulate a simple-to-solve quadratic optimisation task rather than a complex nonlinear one [52].
4. The optimization procedure must compute a new vector of decision variables at each sampling instant online. Therefore, any weakness of the optimization process, e. g., too early termination resulting from a short sampling instant and a slow hardware platform, results in a deterioration of control quality.

All things considered, the control quality possible in MPC-based control system is influenced by many factors: the horizons, the tuning coefficients, the model accuracy and the effectiveness of the optimization procedure. There are many more parameters than in the case of PI/PID-based control systems.

In addition to typically used control-performance indicators, such as the sum of squared errors, overshoot and setting time, for control performance assessment of MPC-based control systems, more sophisticated indices, including fractal and entropy measures [17], are recommended. This is because they can better measure the

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

controller's performance, in particular, when the control errors have a non-Gaussian nature and outliers are present. The effectiveness of fractal and entropy control-performance assessment methods is discussed in [23, 24, 26] (for MPC algorithms based on linear models) and in [25, 27] (for nonlinear MPC algorithms based on nonlinear models). A review of control performance assessment methods for MPC is given in [18]. It must be stressed that analysis of poor control performance is much more difficult than in the case of PI/PID-based control systems. It is also necessary to point out that quite frequently it is impossible to determine exactly the reason for the unsatisfactory control quality because different reasons may give very similar results. Let us give two simple examples. First, when control is too slow, the reasons may be the following:

- too short or too long horizons,
- too low coefficients $\psi_{p,n}$ or too large coefficients $\lambda_{p,n}$,
- model errors, typically its static gain is much greater than the gain of the process or the model delay is too long.

Second, when control is too fast, which means large changes of the manipulated variables, the reasons may be the following:

- too short or too long horizons,
- too low coefficients $\lambda_{p,n}$ in comparison with the coefficients $\psi_{p,n}$,
- model errors, typically its static gain is much lower than the gain of the process or the model delay is too short.

The recommended procedure to properly tune the MPC algorithm is to verify the model first. In particular, since typical processes are in some regions of operation nonlinear, the model must be validated for different operating points and different scenarios of disturbances. Next, the adequate lengths of the prediction and control horizons must be adjusted, taking into account the speed of the process dynamics and the sampling time used. Finally, the tuning coefficients are tuned. Although, for simplicity, in many applications, the same coefficients $\psi_{p,n}$ over the whole prediction horizon are assumed, it is beneficial to consider a more sophisticated approach as discussed in [60]. Although tuning takes some time, it is not recommended to use equal coefficients $\psi_{p,n}$. Similarly, it is discouraged to use the same weights $\lambda_{p,n}$.

MPC, as an algorithm that is more complex than common PID control law, enables achieving very high system performance. On the other hand, we must not forget that exorbitant controller tuning and efficiency is vulnerable to the situations that have not been taken into account during the design and tuning phase. Usually, MPC sacrifices robustness at the expense of high performance. And this is where the door opens for the analysis of outliers, their impact, and of the associated risk. It is not an art to come up with a cutting-edge control algorithm that achieves sensational quality in a given specific situation. The trick is to come up with a control law that improves ef-

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

ficiency regardless of changing or even previously unforeseen (unknown unknowns) operating conditions. Outliers may help in that regard.

1.6 Conclusions, what's next and beyond

The key message from this chapter is clear: outliers exist and matter. In control- engineering practice, we should be aware of, and informed about, outliers. In particular, we have explicitly pointed out the connection between outliers and fractional calculus, thus the title of this edited volume. In the following chapters, this theme will unfold with interesting new results and insights.

Looking towards the future, we believe in the Big Data era. While Big Data generators are usually complex dynamic systems, it is inevitable to characterize the "variabilities" hidden in the big data and thus, frequently, we will meet the heavy-tailedness or inverse power law (IPL) in a certain sense. So, we wish to make the community aware that tail information in the outliers, due to complex stochasticity, may be the key enabling insight towards better data analytics for Big Data. Therefore, fractional calculus should go hand-in-hand with statistical analysis in the analysis of the outliers.

Outliers are neither anything artificial nor alien. They simply surround us. Closing our eyes will not make them disappear. We should not be afraid of outliers, deleting them in panic. They often carry very important information that ought to be used because it informs us about complexity and richness of the surrounding environment.

From a control-systems point of view, however, it remains unclear in control-engineering practice that, when a system is driven by a fractional order noise, what is the best tail-index matched feedback control law to optimally mitigate the fractional noises. Some initial attempt in this regard can be found in [89, 90]. The more complex the control law (APC, MPC, optimization-based, etc.), the greater the importance of outliers. Future opportunities to harness the knowledge of outliers are only limited by our imagination, which could range from human–automation interaction to health-aware smart control systems, from attack awareness and attack resilience to cognitive prognostics, etc. We should point out that outliers could also be beneficially utilized, for example, to enhance system identifiability, more optimal search, more optimal learning, etc. See [61].

Bibliography

[1] K. Aditya, H. Kolla, W. P. Kegelmeyer, T. M. Shead, J. Ling, and W. L. Davis. Anomaly detection in scientific data using joint statistical moments. *Journal of Computational Physics*, 387:522–538, 2019.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[2] A. Alfons, M. Templ, and P. Filzmoser. Robust estimation of economic indicators from survey samples based on Pareto tail modelling. *Journal of the Royal Statistical Society Series C*, 62(2):271–286, 2013.
[3] I. Arismendi, S. L. Johnson, and J. B. Dunham. Higher-order statistical moments and a procedure that detects potentially anomalous years as two alternative methods describing alterations in continuous environmental data. *Hydrology and Earth System Sciences*, 19:1169–1180, 2015.
[4] G. Barbato, E. M. Barini, G. Genta, and R. Levi. Features and performance of some outlier detection methods. *Journal of Applied Statistics*, 38(10):2133–2149, 2011.
[5] V. D. Barnett and T. Lewis. *Outliers in Statistical Data*, 3rd edition. Wiley, Chichester, 1994.
[6] J. Barunik, T. Aste, T. Di Matteo, and R. Liu. Understanding the source of multifractality in financial markets. *Physica A: Statistical Mechanics and Its Applications*, 391(17):4234–4251, 2012.
[7] R. J. Beckman and R. D. Cook. Outlier … … …. s. *Technometrics*, 25(2):119–149, 1983.
[8] I. Ben-Gal. Outlier detection. In O. Maimon and L. Rockach, editors, *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*, pages 131–146. Springer, Boston, MA, 2005.
[9] S. Bhattacharya, M. Kallitsis, and S. Stoev. Data-adaptive trimming of the hill estimator and detection of outliers in the extremes of heavy-tailed data. arXiv:1808.07704v1, Cornell University Library, 2018.
[10] E. Cohen and J. C. Beck. Fat- and heavy-tailed behavior in satisficing planning. In S. A. McIlraith and K. Q. Weinberger, editors, *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2–7, 2018*, pages 6136–6143. AAAI Press, 2018.
[11] S. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics. Springer Verlag London, 2001.
[12] D. Cousineau and S. Chartier. Outliers detection and treatment: a review. *International Journal of Psychological Research*, 3:58–67, Jun 2010.
[13] J. Danielsson, L. M. Ergun, L. de Haan, and C. G. de Vries. Tail index estimation: Quantile-driven threshold selection. Bank of Canada, Staff Working Paper 2019-28, 2019.
[14] W. J. Dixon. Analysis of extreme values. *The Annals of Mathematical Statistics*, 21(4):488–506, 12 1950.
[15] P. D. Domański. Non-Gaussian properties of the real industrial control error in SISO loops. In *Proceedings of the 19th International Conference on System Theory, Control and Computing*, pages 877–882, 2015.
[16] P. D. Domański. Non-Gaussian and persistence measures for control loop quality assessment. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 26(4):043105, 2016.
[17] P. D. Domański. *Control Performance Assessment: Theoretical Analyses and Industrial Practice*, volume 245 of *Studies in Systems, Decision and Control*. Springer, Cham, 2020.
[18] P. D. Domański. Performance assessment of predictive control—A survey. *Algorithms*, 13:97, 2020.
[19] P. D. Domański. Statistical measures. In *Control Performance Assessment: Theoretical Analyses and Industrial Practice*, pages 53–74. Springer International Publishing, Cham, 2020.
[20] P. D. Domański. Statistical outlier labelling—A comparative study. In *Preprints of 7th International Conference on Control, Decision and Information Technologies (CoDIT'20)*, pages 439–444, Prague, Czech Republic, 2020.
[21] P. D. Domański. Study on statistical outlier detection and labelling. *International Journal of Automation and Computing*, 17(6):788–811, 2020.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[22] P. D. Domański, S. Golonka, R. Jankowski, P. Kalbarczyk, and B. Moszowski. Control rehabilitation impact on production efficiency of ammonia synthesis installation. *Industrial & Engineering Chemistry Research*, 55(39):10366–10376, 2016.
[23] P. D. Domański and M. Ławryńczuk. Assessment of predictive control performance using fractal measures. *Nonlinear Dynamics*, 89:773–790, 2017.
[24] P. D. Domański and M. Ławryńczuk. Assessment of the GPC control quality using non-Gaussian statistical measures. *International Journal of Applied Mathematics and Computer Science*, 27:291–307, 2017.
[25] P. D. Domański and M. Ławryńczuk. Control quality assessment for processes with asymmetric properties and its application to pH reactor. *IEEE Access*, 8:94535–94546, 2020.
[26] P. D. Domański and M. Ławryńczuk. Multi-criteria control performance assessment method for a multivariate MPC. In *Proceedings of the American Control Conference (ACC 2020)*, pages 1968–1973, Denver, Colorado, USA, 2020.
[27] P. D. Domański and M. Ławryńczuk. Quality assessment of nonlinear model predictive control using fractal and entropy measures. In W. Lacarbonara, B. Balachandran, J. Ma, J. Tenreiro Machado, and G. Stepan, editors, *Nonlinear Dynamics and Control*, pages 147–156. Springer, Cham, 2020.
[28] I. Fedotenkov. A review of more than one hundred Pareto-tail index estimators. *Statistica*, 80(3):245–299, 2020.
[29] A. Gaddam, T. Wilkin, M. Angelova, and J. Gaddam. Detecting sensor faults, anomalies and outliers in the Internet of things: A survey on the challenges and solutions. *Electronics*, 9(3):511, 2020.
[30] R. Gorenflo and F. Mainardi. Fractional calculus and stable probability distributions. *Archives of Mechanics*, 50(3):377–388, 1998.
[31] F. E. Grubbs. Sample criteria for testing outlying observations. *The Annals of Mathematical Statistics*, 21(1):27–58, 03 1950.
[32] M. Gupta, J. Gao, C. Aggarwal, and J. Han. *Outlier Detection for Temporal Data*. Morgan & Claypool Publishers, 2014.
[33] D. M. Hawkins. *Identification of Outliers*. Chapman and Hall, London; New York, 1980.
[34] N. A. Heckert, J. J. Filliben, C. M. Croarkin, B. Hembree, W. F. Guthrie, P. Tobias, and J. Prinz. NIST/SEMATECH e-handbook of statistical methods. NIST, US Department of Commerce, [08-February-2020], 2012.
[35] B. M. Hill. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174, 1975.
[36] D. Horla. Experimental results on actuator/sensor failures in adaptive GPC position control. *Actuators*, 10:43, 2021.
[37] J. R. M. Hosking. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(1):105–124, 1990.
[38] J. R. M. Hosking. Moments or l moments? An example comparing two measures of distributional shape. *The American Statistician*, 46(3):186–189, 1992.
[39] X. Hou, S. Guo, L. Shi, H. Xing, H. Yin, Z. Li, and M. Zhou. Improved model predictive-based underwater trajectory tracking control for the biomimetic spherical robot under constraints. *Applied Sciences*, 10:8106, 2020.
[40] P. J. Huber and E. M. Ronchetti. *Robust Statistics*, 2nd edition. Wiley, 2009.
[41] M. Hubert and M. Debruyne. Minimum covariance determinant. *WIREs Computational Statistics*, 2(1):36–43, 2010.
[42] B. Huyck, J. De Brabanter, B. De Moor, J. F. Van Impe, and F. Logist. Online model predictive control of industrial processes using low level control hardware: A pilot-scale distillation column case study. *Control Engineering Practice*, 28:34–48, 2014.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[43] B. Iglewicz and D. C. Hoaglin. *How to Detect and Handle Outliers*. ASQC Quality Press, Milwaukee, WI, 1993.
[44] J. O. Irvin. On a criterion for the rejection of outlying observation. *Biometrika*, 17:238–250, 1925.
[45] R. A. Johnson and D. W. Wichern. *Applied Multivariate Statistical Analysis*, 3 edition. Prentice-Hall, Englewood Cliffs, NJ, USA, 1992.
[46] A. Kaleli. Development of the predictive based control of an autonomous engine cooling system for variable engine operating conditions in SI engines: design, modeling and real-time application. *Control Engineering Practice*, 100:104424, 2020.
[47] R. W. Katz, M. B. Parlange, and P. Naveau. Statistics of extremes in hydrology. *Advances in Water Resources*, 25(8):1287–1304, 2002.
[48] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo. *Theory and Applications of Fractional Differential Equations*, volume 204. Elsevier, 2006.
[49] L. B. Klebanov. Big outliers versus heavy tails: what to use? arXiv:1611.05410v1, Department of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic, 2016.
[50] L. B. Klebanov and I. Volchenkova. Outliers and the ostensibly heavy tails. arXiv:1807.08715v1, Department of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic, 2018.
[51] H. P. Kriegel, P. Kroger, and A. Zimek. Outlier detection techniques. Tutorial, 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, 2010.
[52] M. Ławryńczuk. *Computationally Efficient Model Predictive Control Algorithms: A Neural Network Approach*, volume 3 of *Studies in Systems, Decision and Control*. Springer, Cham, 2014.
[53] M. Ławryńczuk. Constrained computationally efficient nonlinear predictive control of Solid Oxide Fuel Cell: Tuning, feasibility and performance. *ISA Transactions*, 99:270–289, 2020.
[54] Z. Li, Y. Zhao, N. Botta, C. Ionescu, and X. Hu: Copod: Copula-based outlier detection. In *2020 IEEE International Conference on Data Mining (ICDM)*, pages 1118–1123, 2020.
[55] J. M. Maciejowski. *Predictive Control with Constraints*. Prentice Hall, Harlow, 2002.
[56] B. B. Mandelbrot and R. L. Hudson. *The Misbehavior of Markets: A Fractal View of Financial Turbulence*. Basic Books, New York, 2004.
[57] P. M. Marusak. A numerically efficient fuzzy mpc algorithm with fast generation of the control signal. *International Journal of Applied Mathematics and Computer Science*, 31:59–71, 2021.
[58] K. G. Mehrotra, C. K. Mohan, and H. M. Huang. *Anomaly Detection Principles and Algorithms*. Terrorism, Security, and Computation. Springer Publishing Company, Inc., 2017.
[59] Ralf Metzler and Joseph Klafter. The random walk's guide to anomalous diffusion: A fractional dynamics approach. *Physics Reports*, 339(1):1–77, 2000.
[60] R. Nebeluk and M. Ławryńczuk. Tuning of multivariable model predictive control for industrial tasks. *Algorithms*, 14:10, 2021.
[61] H. Niu, Y. Q. Chen, and B. J. West. Why do big data and machine learning entail the fractional dynamics? *Entropy*, 23(3):297 https://doi.org/10.3390/e23030297, 2021.
[62] S. Ogonowski, D. Bismor, and Z. Ogonowski. Control of complex dynamic nonlinear loading process for electromagnetic mill. *Archives of Control Sciences*, 30:471–500, 2020.
[63] J. W. Osborne and A. Overbay. The power of outliers (and why researchers should always check for them). *Practical Assessment, Research, and Evaluation*, 9(6):1–8, 03 2004.
[64] T. J. Papadimos. Stoicism, the physician, and care of medical outliers. *BMC Medical Ethics*, 5:8, 2004.
[65] E. S. Pearson and C. Chandra Sekar. The efficiency of statistical tools and a criterion for the rejection of outlying observations. *Biometrika*, 28(3/4):308–320, 12 1936.
[66] R. K. Pearson. *Mining Imperfect Data; Dealing with Contamination and Incomplete Records*. SIAM, Philadelphia, PA, USA, 2005.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[67] M. Peel, Q. Wang, and T. Mcmahon. The utility l-moment ratio diagrams for selecting a regional probability distribution. *Hydrological Sciences Journal*, 46:147–155, 02 2001.
[68] B. Peirce. Criterion for the rejection of doubtful observations. *Astronomical Journal*, 2(45):161–163, 7 1852.
[69] I. Podlubny. *Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications*. Elsevier, 1998.
[70] L. Pollet, T. V. van der Meij. To remove or not to remove: the impact of outlier handling on significance testing in testosterone data. *Adaptive Human Behavior and Physiology*, 3:43–60, 2017.
[71] K. R. Popper. *The Logic of Scientific Discovery*, 2nd edition. Routledge, 2002.
[72] U. Porwal and S. Mukund. Credit card fraud detection in e-commerce: An outlier detection approach. arXiv:1811.02196v2, eBay Inc, San Jose, California, 2019.
[73] M-J. Rakotosaona, V. La Barbera, P. Guerrero, N. J. Mitra, and M. Ovsjanikov. Pointcleannet: Learning to denoise and remove outliers from dense point clouds. arXiv:1901.01060, Cornell University Library, 2020.
[74] F. Rosado. Outliers: The strength of minors. In A. Pacheco, R. Santos, M. R. Oliveira, and C. D. Paulino, editors, *New Advances in Statistical Modeling and Applications*, pages 17–27. Springer International Publishing, Cham, 2014.
[75] B. Rosner. Percentage points for a generalized ESD many-outlier procedure. *Technometrics*, 25(2):165–172, 1983.
[76] P. J. Rousseeuw and M. Hubert. Anomaly detection by robust statistics. *WIREs Data Mining and Knowledge Discovery*, 8(2):e1236, 2018.
[77] P. J. Rousseeuw and A. M. Leroy. *Robust Regression and Outlier Detection*. John Wiley & Sons, Inc., New York, NY, USA, 1987.
[78] A. P. Sarraf. Flood outlier detection using PCA and effect of how to deal with them in regional flood frequency analysis via l-moment method. *Water Resources*, 42:448–459, 2015.
[79] C. Schluter and M. Trede. Identifying multiple outliers in heavy-tailed distributions with an application to market crashes. *Journal of Empirical Finance*, 15(4):700–713, 2008.
[80] H. G. Sun, Y. Zhang, D. Baleanu, W. Chen, and Y. Q. Chen. A new collection of real world applications of fractional calculus in science and engineering. *Communications in Nonlinear Science and Numerical Simulation*, 64:213–231, 2018.
[81] N. N. Taleb. *Real-world Statistical Consequences of Fat Tails: Papers and Commentary*. Technical Incerto Collection. STEM Academic Press, 2018.
[82] N. N. Taleb. Statistical consequences of fat tails: Real world preasymptotics, epistemology, and applications. arXiv:2001.10488, Cornell University Library, 2020.
[83] P. Tatjewski. *Advanced Control of Industrial Processes, Structures and Algorithms*. Springer, London, 2007.
[84] M. Templ, J. Gussenbauer, and P. Filzmoser. Evaluation of robust outlier detection methods for zero-inflated complex data. *Journal of Applied Statistics*, 47(7):1144–1167, 2019.
[85] R. Thompson. A note on restricted maximum likelihood estimation with an alternative outlier model. *Journal of the Royal Statistical Society: Series B (Methodological)*, 47(1):53–55, 1985.
[86] G. L. Tietjen and R. H. Moore. Some Grubbs-type statistics for the detection of several outliers. *Technometrics*, 14(3):583–597, 1972.
[87] M. P. J. van der Loo. Distribution based outlier detection in univariate data. Technical Report Discussion paper 10003, Statistics Netherlands, The Hague/Heerlen, Netherlands, 2010.
[88] S. Verboven and M. Hubert. LIBRA: A Matlab library for robust analysis. *Chemometrics and Intelligent Laboratory Systems*, 75:127–136, 02 2005.
[89] J. Viola, P. Oziablo, and Y. Q. Chen. An experimental networked control system with fractional order delay dynamics. In *2019 7th International Conference on Control, Mechatronics and Automation (ICCMA)*, pages 226–231, 2019.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[90] J. Viola, P. Oziablo, and Y. Q. Chen. A study of the influence of stochastic fractional-order delay dynamics in a networked control system. In *2020 IFAC World Congress. IFAC-PapersOnLine vol. 53 no. 2*, pages 5789–5794, 2020.
[91] H. Wainer. Robust statistics: A survey and some prescriptions. *Journal of Educational Statistics*, 1(4):285–312, 1976.
[92] B. J West. *Fractional Calculus View of Complexity: Tomorrow's Science*. CRC Press, 2016.
[93] D. L. Whaley. *The Interquartile Range: Theory and Estimation*. PhD thesis, Faculty of the Department of Mathematics, East Tennessee State University, 8 2005. Electronic Theses and Dissertations. Paper 1030.
[94] C. S. C. Woolley, I. G. Handel, B. M. Bronsvoort, J. J. Schoenebeck, and D. N. Clements. Is it time to stop sweeping data cleaning under the carpet? A novel algorithm for outlier management in growth data. *PLOS ONE*, 15:1–21, 01 2020.
[95] K. Zarzycki and M. Ławryńczuk. Fast real-time model predictive control for a ball-on-plate process. *Sensors*, 21:3959, 2021.
[96] A. Zimek and P. Filzmoser. There and back again: Outlier detection between statistical reasoning and data mining algorithms. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(6):e1280, 11 2018.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Part II: **Theory**

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Stefan Domek

2 On the possibilities of using fractional-order differential calculus in linear and nonlinear model predictive control

Abstract: Biological, sociocognitive and economic phenomena, as well as transport, information and technological systems, often need to be described by means of nonlinear, non-integer order differential equations. As a consequence, control systems should often also realize nonlinear and non-integer order control algorithms. In this paper a basic methods of differential calculus and methods for its use to model complex processes are reviewed. The chapter presents developing methods of linear Model Predictive Control (MPC) with an internal fractional-order process model and a fractional-order cost function (FO MPC). Further, the concepts of three suboptimal nonlinear MPC (NMPC) methods for nonlinear, non-integer order systems are described. In the first fractional-order NMPC method, the linearized fractional-order model of the process is used. In the second method the fractional-order model is successively linearized online and used only for optimization, but the prediction is made with a nonlinear model. In the third a linear approximation along the future process trajectory is also used. In all cases, as a result of linearization, the future control policy is calculated by means of quadratic optimization. The proposed FO NMPC algorithms offer an additional degree of freedom in tuning a control loop for higher efficiency. In the paper the discussion is illustrated with results from some numerical experiments.

Keywords: fractional-order differential calculus, fractional-order dynamic models, model predictive control, fractional-order nonlinear model predictive control

2.1 Introduction

The idea of MPC, which was put forward several dozen years ago and has been intensely developed since then, is considered to be, after many years of operating experience in industry, one of the most universally used and effective control methods. In MPC the future control actions $u(t+j|t)$ are to be found at each instant $t \in \mathbb{Z}_+$ within the control horizon from $j = 0$ to $j = N_u - 1$ to minimize the differences between the reference values $y^r(t+j|t)$ and the predicted values $y^p(t+j|t)$ within the prediction horizon from $j = N_1$ to $j = N_2$. The optimal control sequence is computed in the open loop mode with the use of the plant model. Only the first value of the computed sequence is

Stefan Domek, West Pomeranian University of Technology in Szczecin, Szczecin, Poland, e-mail: stefan.domek@zut.edu.pl

https://doi.org/10.1515/9783110729122-002

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

fed as the input into the real plant, and the whole procedure is repeated at the following discrete time instants (receding horizon principle) [2, 16, 29]. MPC can handle in a natural way multivariable systems, and, moreover, it can take into account explicitly various signals constraints and various kinds of disturbances. However, especially in the case of processes with particularly complex properties, its effectiveness depends on the accuracy of the process model that is utilized directly to compute the manipulated variable [8, 18, 30, 31]. The use of nonlinear models is an effective and natural modeling method, but unfortunately in MPC this leads to a very complex nonlinear optimization problem that must be solved at each sampling step [13, 22, 29]. One of the methods to mitigate the numerical complexity of a nonlinear MPC algorithm is the use of methods with nonlinear trajectory prediction and cyclic model linearization for the needs of ongoing optimization. There are several ways to accomplish this general principle of simplifying the implementation of nonlinear predictive control, but the most advanced methods seems to be that labeled with the acronyms NPL (Nonlinear Prediction of the free response component with model Linearization for optimization) and NPLPT (Nonlinear Prediction and Linearization along the Predicted Trajectory) [4, 13, 14, 29].

From the opposing perspective, it has been known for several years that the actual properties of many complex phenomena and nonlinear industrial processes can be effectively modeled using fractional-order differential calculus [10, 11, 19, 26]. The theory of fractional-order systems (FOS) has been intensively developing over recent decades [11, 12, 19–21, 28]. In automation and robotics, as in the case of integer-order models, the description by means of fractional-order models can be used indirectly for tuning or directly for synthesis of linear control algorithms [3, 24, 25, 27]. In the second case, fractional-order differential calculus has been applied to control theory, which, in its turn, should contribute to the development of new control algorithms significantly different from the well-known integer-order algorithms, and thus, by implication, provide potentially new opportunities for control performance and robustness [32]. Allowing integration/differentiation of arbitrary orders in classic control algorithms results in increasing the number of degrees of freedom in control-parameter tuning, and thus creates new potentialities for control performance and robustness. Excellent examples here are the CRONE algorithm [20] and fractional-order digital $PI^{\lambda}D^{\mu}$ algorithm [21], already regarded as standard, but also the fractional-order iterative learning control, linear-quadratic control, model reference adaptive control, dead-beat control and sliding mode control, which have been proposed in subsequent years [1, 15, 17]. New possibilities can also naturally create the fractional-order model predictive control (FO MPC) methods [6, 24].

As is well known, to determine the manipulated variable in MPC, three elements need to be defined—the reference trajectory, the plant model and the cost function. In analyzing the possibility and the sense of the use of the fractional-order differential calculus in predictive control, one can readily see that the reference trajectory is

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

not dependent on whether the algorithm of integer or fractional order is considered. However, fractional-order differential calculus may be applied both to plant model selection and to defining the cost function. Thus, taking this circumstance into account at the stage of synthesizing the MPC system may increase naturally the applicability of the controller, which can be regarded as a fractional-order model predictive controller in such a case. A possibility to include the concept of a fractional-order performance index into the FO MPC algorithm was proposed in [23]. The multiple employing of the fractional-order differential calculus in FO MPC, first as a fractional-order performance index and second one as an internal fractional-order model of the controlled process, was proposed in [7]. The proposition provides potentially new opportunities for control performance and robustness in the case of fractional-order linear processes. However, if controlled fractional-order process is nonlinear, a direct use in FO MPC of a nonlinear model leads to a nontrivial, time-consuming, nonlinear optimization problem, which must be solved on-line. Thus, special fractional-order nonlinear model predictive control (FO NMPC) methods are required.

One of the methods to overcome these difficulties consists of replacing a complex nonlinear model by a set of local linear submodels valid for small regions around various operating points. The idea of this approach boils down to switching over active models in time so that the generalized modeling error does not exceed specified bounds (e. g., in terms of a chosen norm) and to obtain a reduction in the computational complexity of the description at the same time. By this means a nonlinear time-invariant process (NLTI) can be treated as a linear time-variant (LTV) one. Instantaneous properties of the modeled nonlinear process are described then by a quasilinear switched model, composed of a set of local linear submodels. Selected discrete-time switched models of non-integer order in the state space are defined, their basic properties are discussed and examples of simulation results are given in [5, 12]. In [4] employment of discrete-time switched models of non-integer order for the synthesis of FO NMPC algorithms has been proposed. The introduced nonlinear fractional-order predictive control methods are based on integer-order MPC NPL and MPC NPLPT and simplifying the implementation of FO NMPC because the future control policy is calculated by means of well-known quadratic optimization [9, 13, 14, 29].

The chapter is structured as follows: in Section 2.2 the basics of fractional-order differential and difference calculus are reviewed; in Section 2.3 linear and nonlinear dynamic state-space models of non-integer order based on them is described; in Section 2.4 a synthesis of linear fractional-order MPC with multiple employment of fractional-order differential calculus is presented. Also, the effects of various orders of performance index and various model–plant mismatches results are recalled from [8]; Finally, in Section 2.5 the concepts of three FO NMPC methods are described; The entire chapter content is summarized in the conclusions section.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.2 Basics of fractional-order differential and difference calculus

Differential calculus of non-integer order is a generalization of the classic differential calculus. For a derivative of non-integer order $\alpha \in \mathbb{R}$ of a real-valued function $f(t),\ t \in \mathbb{R}$ on the interval $[t_0,\ t]$, denoted by the operator ${}_{t_0}D_t^{\alpha} f(t) = \frac{d^{\alpha} f(t)}{dt^{\alpha}}$, there exist many definitions proposed by various researchers, for example, by Riemann and Liouville, Caputo, Weyl, Fourier, Cauchy and Abel. The definitions differ in their properties and/or area of applicability [12, 19, 21]. In practical applications, especially in digital control systems where discrete values of the function $f(t)$ taken with the sampling interval h are used in a natural way for computations, the most commonly encountered is the definition introduced by Grünwald and Letnikov [19]:

A derivative of order $\alpha \in \mathbb{R}$ of the function $f(t)$, $t \in \mathbb{R}$, according to Grünwald and Letnikov, is given by

$$ {}_{t_0}^{\mathrm{GL}}D_t^{\alpha} f(t) = \lim_{h \to 0} h^{-\alpha} \sum_{j=0}^{\lfloor \frac{t-t_0}{h} \rfloor} c_j^{\alpha} f(t - jh), \tag{2.1} $$

where the symbol $\lfloor \cdot \rfloor$ denotes the integer part, $c_j^{\alpha} = (-1)^j \binom{\alpha}{j}$ and the so-called generalized Newton symbol is defined by

$$ \binom{\alpha}{j} = \begin{cases} 1 & \text{for } j = 0, \\ \frac{\alpha(\alpha-1)\cdots(\alpha-j+1)}{j!} & \text{for } j = 1, 2, 3, \ldots. \end{cases} \tag{2.2} $$

Remark. It follows from the definition of the generalized derivative that ${}_{t_0}D_t^0 f(t) = f(t)$ and the fractional order derivative of the function $f(t)$ for $\alpha < 0$ is an integral of the order $-\alpha$ frequently symbolized by the operator ${}_{t_0}I_t^{-\alpha} f(t)$.

For discrete-time functions $f(t)$ the difference calculus of non-integer order is the counterpart of the differential calculus of non-integer order. Taking for simplicity the normalized sampling period $h = 1$, by analogy with (2.1), the following discrete difference of fractional order $\alpha \in \mathbb{R}$ of the function $f(t)$, $t \in \mathbb{Z}$ can be formulated [12]:

$$ {}_{t_0}\Delta_t^{\alpha} f(t) = \sum_{j=0}^{t-t_0} c_j^{\alpha} f(t - j), \tag{2.3} $$

with $t_0 = 0$ being most commonly adopted. If so, (2.3) takes a simpler form

$$ \Delta^{\alpha} f(t) = \sum_{j=0}^{t} c_j^{\alpha} f(t - j). \tag{2.4} $$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.3 Discrete-time dynamic state-space models of non-integer order

A fractional-order generalized model of a discrete-time nonlinear process may be defined in the state space as [9]

$$\Delta^{\Upsilon} x(t+1) = \boldsymbol{f}\big(x(t), u(t)\big) + v(t+1), \tag{2.5}$$

$$y(t) = \boldsymbol{g}\big(x(t)\big) + d(t), \tag{2.6}$$

where

$$\Delta^{\Upsilon} x(t) = \begin{bmatrix} \Delta^{\alpha_1} x_1(t) & \cdots & \Delta^{\alpha_n} x_n(t) \end{bmatrix} = \sum_{i=0}^{t} (-1)^i \Upsilon_i x(t-i), \tag{2.7}$$

$$\Upsilon_i = \operatorname{diag}\begin{bmatrix} \binom{\alpha_1}{i} & \cdots & \binom{\alpha_n}{i} \end{bmatrix} \in \mathbb{R}^{n \times n} \tag{2.8}$$

is the generalized backward difference of the state vector $x(t) \in \mathbb{R}^n$ with different orders for individual state variables $\Upsilon = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, while $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^p$ denote the input and output vectors and $v(t) \in \mathbb{R}^n$, $d(t) \in \mathbb{R}^p$ are the state and output disturbance vectors, respectively.

The model (2.5) can be rewritten in the following form:

$$x(t+1) = \boldsymbol{f}\big(x(t), u(t)\big) - \sum_{i=1}^{t+1} (-1)^i \Upsilon_i x(t+1-i) + v(t+1). \tag{2.9}$$

Note that the unknown vector $v(t)$ may be assessed as the difference between the measured or estimated state at the current and previous sampling instants $0, 1, 2, \ldots, t$ and the state calculated from the state equation (2.9) for the sampling instant t at the sampling instant $t-1$:

$$v(t|t) = x(t) + \sum_{i=1}^{t} (-1)^i \Upsilon_i x(t-i) - \boldsymbol{f}\big(x(t-1), u(t-1)\big). \tag{2.10}$$

The unmeasured output disturbance vector can be calculated just as the difference between the measured output at the current sampling instants and the state calculated from the output equation (2.6) for the sampling instant t:

$$d(t|t) = y(t) - \boldsymbol{g}\big(\boldsymbol{f}\big(x(t-1), u(t-1)\big) + v(t)\big). \tag{2.11}$$

Assuming that the functions $\boldsymbol{f}$ and $\boldsymbol{g}$ in (2.5) and (2.6) are differentiable, the non-integer order linearized model at the current operating point of a nonlinear process

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

may be given in the state space using the Taylor series expansion method as:

$$x(t+1) = \mathbf{A}x(t) + \mathbf{B}u(t) - \sum_{i=1}^{t+1}(-1)^i \Upsilon_i x(t+1-i) + v(t+1), \tag{2.12}$$

$$y(t) = \mathbf{C}x(t) + d(t), \tag{2.13}$$

where the complemented state matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, input matrix $\mathbf{B} \in \mathbb{R}^{n\times m}$ and output matrix $\mathbf{C} \in \mathbb{R}^{p\times n}$ are the corresponding derivatives of functions $\boldsymbol{f}(x,u)$ and $\boldsymbol{g}(x)$ at the current operating point of a nonlinear process. For such a linearized model, the consecutive predictions of the state vectors can be calculated from the following formula [4, 12]:

$$x^p(t+j|t) = \Phi^\Upsilon(j)x(t) + \sum_{i=0}^{j-1}\Phi^\Upsilon(j-i-1)\mathbf{B}u(t+i) + v(t+j|t), \tag{2.14}$$

where the matrix $\Phi^\Upsilon(j)$ is determined by the recurrence relation

$$\Phi^\Upsilon(j+1) = \Phi^\Upsilon(j)(\mathbf{A} + \Upsilon_1) + \sum_{i=2}^{j+1}(-1)^{i+1}\Upsilon_i\Phi^\Upsilon(j-i+1), \quad \Phi^\Upsilon(0) = \mathbf{I}_n. \tag{2.15}$$

For the non-integer order linearized model (2.12), (2.13) the unknown state disturbance vector may be assessed as:

$$v(t|t) = x(t) - \mathbf{A}x(t-1) - \mathbf{B}u(t-1) + \sum_{i=1}^{t}(-1)^i\Upsilon_i x(t-i). \tag{2.16}$$

Similarly, the unmeasured output disturbance vector $d(t)$ can be calculated just as the difference between the measured output at the current sampling instant and the output predicted for the sampling instant t at the sampling instant $t-1$:

$$d(t|t) = y(t) - \mathbf{C}x^p(t|t-1). \tag{2.17}$$

In practical applications of models (2.12), (2.13) it is not possible to take for numerical calculations the state vector $x(t+1-i)$ samples, the number of which grows rapidly with increasing discrete time t. One of the methods to cope with the problem is to adopt a finite-length memory in which the instantaneous values of the state vector are stored. Such an approach, taken from the theory of digital filters with a finite impulse response, leads to the so-called finite fractional differences (FFD) [4, 28]. In view of this, the discrete difference of fractional order $\alpha \in \mathbb{R}$ of the state vector $x(t)$, $t \in Z$ (2.4) takes the form:

$$\Delta^\alpha x(t+1) = \sum_{i=0}^{L} c_i^\alpha x(t+1-i), \tag{2.18}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

where L is the adopted memory length. However, in the process of calculations, it should be taken into account that the upper limit of summation must be reduced to the value of $t+1$, until enough samples are accumulated. A similar shortening applies to the sum in the generalized models (2.5)–(2.8) [6, 12].

2.4 Linear fractional-order MPC with multiple uses of the fractional-order differential calculus

In the predictive control algorithm of integer order, the cost function depends on the sum of the weighted squared prediction errors over the prediction horizon and on the sum of the weighted squared control signal increments to be sought within the control horizon

$$J(t) = \sum_{j=N_1}^{N_2} \mu(j)\left[e^p(t+j|t)\right]^2 + \sum_{j=0}^{N_u-1} \lambda(j)\left[\Delta u(t+j|t)\right]^2, \tag{2.19}$$

with $\Delta u(t+j|t) = 0$ for $j \geq N_u$, and

$$e^p(t+j|t) = y^r(t+j|t) - y^p(t+j|t). \tag{2.20}$$

In [23] it was proposed to introduce a formal generalization of integer-order sums in the cost function (2.19) as those of the fractional order

$$J(t) = {}_{t+N_1}I^{\beta_1}_{t+N_2}\left[e^p(t)\right]^2 + \lambda_t I^{\beta_2}_{t+N_u-1}\left[\Delta u(t)\right]^2, \tag{2.21}$$

where $\beta_1, \beta_2 \in \mathbb{R}_+$ are fractional orders of integration (summation) applied to squared prediction errors and squared control increments, respectively. By this means the principles of a fractional-order discrete predictive algorithm have been defined by adding fractional orders β_1, β_2 as two new tuning parameters, in addition to those typical ones, namely, the ranges of the prediction horizon N_1, N_2 and the length of the control horizon N_u. On the other hand, it was proposed in [4] to consider the fractional nature of the controlled plant at the synthesizing stage. In the case of predictive control, this is particularly justified because the model of the controlled plant is utilized here directly to compute the manipulated variable. Hence, the effectiveness of the predictive control largely depends on to what extent the employed model matches the controlled plant. In the case of difficult plants exhibiting properties that can be modeled well by models of fractional order, taking this circumstance into account at the synthesizing stage increases naturally the applicability of the controller [6, 8].

Consider a more general case of the fractional-order cost function (2.21) with weighting coefficients $\mu(j)$ and $\lambda(j)$ as in (2.19).

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

In such a case the cost function (2.21) may be rewritten in vector-matrix form [7] as

$$\begin{aligned} J(t) = &\left[E^p(t)_{\leftrightarrow}\right]^T \operatorname{diag}[\mathbf{M}_{\leftarrow} \quad \mathbf{M}_{\rightarrow}]\left[E^p(t)_{\leftrightarrow}\right] \\ &+ \left[\Delta U(t)_{\leftrightarrow}\right]^T \operatorname{diag}[\mathbf{\Lambda}_{\leftarrow} \quad \mathbf{\Lambda}_{\rightarrow}]\left[\Delta U(t)_{\leftrightarrow}\right], \end{aligned} \tag{2.22}$$

where

$$E^p(t)_{\leftrightarrow} = \begin{bmatrix} E^p(t)_{\leftarrow} \\ E^p(t)_{\rightarrow} \end{bmatrix} = \begin{bmatrix} \vdots \\ e^p(t-1) \\ e^p(t) \\ e^p(t+1|t) \\ \vdots \\ e^p(t+N_1-1|t) \\ e^p(t+N_1|t) \\ e^p(t+N_1+1|t) \\ \vdots \\ e^p(t+N_2|t) \end{bmatrix} \in \mathbb{R}^{\infty}, \tag{2.23}$$

$$\Delta U(t)_{\leftrightarrow} = \begin{bmatrix} \Delta U(t)_{\leftarrow} \\ \Delta U(t)_{\rightarrow} \end{bmatrix} = \begin{bmatrix} \vdots \\ \Delta u(t-1) \\ \Delta u(t|t) \\ \Delta u(t+1|t) \\ \vdots \\ \Delta u(t+N_u-1|t) \end{bmatrix} \in \mathbb{R}^{\infty}. \tag{2.24}$$

Vectors $\Delta U(t)_{\leftarrow}$ and $\Delta U(t)_{\rightarrow}$ denote the past and the unknown future increments of the manipulated variable, respectively. Similarly, $Y^p(t)_{\leftarrow}$ denotes the past plant output values, and vectors $Y^p(t)_{\rightarrow}$, $Y^r(t)_{\rightarrow}$ stand for future values of the natural response of the model and those of the reference trajectory, respectively.

The weighting matrices are defined by [4, 23]:

$$\operatorname{diag}\left[\mathbf{M}_{\leftarrow} \quad \mathbf{M}_{\rightarrow}\right] = \operatorname{diag}\left[\cdots \quad w_\mu(-1) \quad w_\mu(0) \quad \cdots \quad w_\mu(N_1) \quad \cdots \quad w_\mu(N_2)\right], \tag{2.25}$$

$$\operatorname{diag}\left[\mathbf{\Lambda}_{\leftarrow} \quad \mathbf{\Lambda}_{\rightarrow}\right] = \operatorname{diag}\left[\cdots \quad w_\lambda(-1) \quad w_\lambda(0) \quad w_\lambda(1) \quad \cdots \quad w_\lambda(N_u-1)\right] \tag{2.26}$$

and

$$w_\mu(j) = \begin{cases} c_{N_2-j}^{-\beta_1} - c_{N_1-j}^{-\beta_1} & \text{for } j < N_1, \\ \mu(j) c_{N_2-j}^{-\beta_1} & \text{for } N_1 \le j \le N_2, \end{cases} \tag{2.27}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$w_\lambda(j) = \begin{cases} c_{N_u-1-j}^{-\beta_2} - c_{-j}^{-\beta_2} & \text{for } j < 0, \\ \lambda(j) c_{N_u-1-j}^{-\beta_2} & \text{for } 0 \le j \le N_u - 1. \end{cases} \tag{2.28}$$

According to (2.20) the future prediction error can be written in the form

$$E^p(t)_{\rightarrow} = Y^r(t)_{\rightarrow} - Y^p(t)_{\rightarrow} = \begin{bmatrix} y^r(t+N_1|t) \\ \vdots \\ y^r(t+N_2|t) \end{bmatrix} - \begin{bmatrix} y^p(t+N_1|t) \\ \vdots \\ y^p(t+N_2|t) \end{bmatrix}, \tag{2.29}$$

assuming that in linear systems the output prediction is composed as a sum of two components—the prediction of the forced response $y^c(t)$, which depends only on the future control moves, and of the natural process response $y^0(t)$ resulting from equation (2.14) that takes into account only the past:

$$Y^p(t)_{\rightarrow} = Y^c(t)_{\rightarrow} + Y^0(t)_{\rightarrow} = \begin{bmatrix} y^c(t+N_1|t) \\ \vdots \\ y^c(t+N_2|t) \end{bmatrix} + \begin{bmatrix} y^0(t+N_1|t) \\ \vdots \\ y^0(t+N_2|t) \end{bmatrix}. \tag{2.30}$$

To predict the natural process response vector in (2.31), the approach known from the DMC algorithm [2, 16] is adopted, i. e., constant disturbances (2.16) and (2.17) in the prediction horizon are assumed:

$$V(t)_{\rightarrow} = \begin{bmatrix} v(t|t) \\ v(t|t) \\ \vdots \\ v(t|t) \end{bmatrix}, \tag{2.31}$$

$$D(t)_{\rightarrow} = \begin{bmatrix} d(t|t) \\ d(t|t) \\ \vdots \\ d(t|t) \end{bmatrix}. \tag{2.32}$$

In view of equations (2.14), (2.15) and (2.30), the output prediction vector assumes the following form:

$$Y^p(t)_{\rightarrow} = \mathbf{E}_Y \Delta U(t)_{\rightarrow} + Y^0(t)_{\rightarrow}, \tag{2.33}$$

with a output dynamic matrix of the model

$$\mathbf{E}_Y = \operatorname{diag}[\mathbf{C}, \ldots, \mathbf{C}] \cdot \mathbf{E}_X \tag{2.34}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

and a state dynamic matrix

$$\mathbf{E}_X = \begin{bmatrix} \sum_{i=0}^{N_1-1} \Phi^{\Upsilon}(i)\mathbf{B} & \cdots & \mathbf{0}_{n\times m} & \cdots & \mathbf{0}_{n\times m} \\ \sum_{i=0}^{N_1} \Phi^{\Upsilon}(i)\mathbf{B} & \cdots & \mathbf{B} & \cdots & \mathbf{0}_{n\times m} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \sum_{i=0}^{N_u-1} \Phi^{\Upsilon}(i)\mathbf{B} & \cdots & \cdots & (\mathbf{A}+\Upsilon_1)\mathbf{B} & \mathbf{B} \\ \vdots & \vdots & \vdots & & \vdots \\ \sum_{i=0}^{N_2-1} \Phi^{\Upsilon}(i)\mathbf{B} & \cdots & \cdots & \cdots & \sum_{i=0}^{N_2-N_u} \Phi^{\Upsilon}(i)\mathbf{B} \end{bmatrix}. \tag{2.35}$$

Thus, the minimum of the cost function (2.22) with respect to the control increment $\Delta U_{\rightarrow}(t)$ in the future (within the control horizon) can be found as [4]:

$$\begin{aligned} \Delta U_{\text{opt}\rightarrow}(t) &= \left(\mathbf{E}_Y{}^T \mathbf{M}_{\rightarrow} \mathbf{E}_Y + \mathbf{\Lambda}_{\rightarrow}\right)^{-1} \mathbf{E}_Y{}^T \mathbf{M}_{\rightarrow} [Y^r(t)_{\rightarrow} - Y^0(t)_{\rightarrow}] \\ &= \mathbf{K}_{\rightarrow} [Y^r(t)_{\rightarrow} - Y^0(t)_{\rightarrow}]. \end{aligned} \tag{2.36}$$

It is pertinent to note that the proposed fractional-order predictive algorithm can be naturally and easily extended to the case with constraints imposed on process variables. In such a case, taking into account (2.22)–(2.32), the general FO MPC optimization problem takes the form [9, 16]:

$$\begin{aligned} \min_{\substack{\Delta U(t)_{\rightarrow} \\ E^{\min}(t)_{\rightarrow} \\ E^{\max}(t)_{\rightarrow}}} J(t) = \{ &\left\| Y^r(t)_{\rightarrow} - Y^p(t)_{\rightarrow} \right\|^2_{\mathbf{M}_{\rightarrow}} + \left\| \Delta U(t)_{\rightarrow} \right\|^2_{\mathbf{\Lambda}_{\rightarrow}} \\ &+ \rho^{\min} \left\| E^{\min}(t)_{\rightarrow} \right\|^2 + \rho^{\max} \left\| E^{\max}(t)_{\rightarrow} \right\|^2 \}, \end{aligned} \tag{2.37}$$

where $\|V\|^2_{\mathbf{W}} = V^T \mathbf{W} V$ denotes the vector V norm with weighting matrix $\mathbf{W} \geq \mathbf{0}$, with rigid constraints imposed on amplitudes of the manipulated variable and its increments

$$u^{\min} \leq u(t+j|t) \leq u^{\max}, \quad j = 0, 1, \ldots, N_u - 1, \tag{2.38}$$

$$\Delta u^{\min} \leq \Delta u(t+j|t) \leq \Delta u^{\max}, \quad j = 0, 1, 2, \ldots, N_u - 1, \tag{2.39}$$

and soft constraints on process outputs, as well as

$$y^{\min} - \varepsilon^{\min}(t+j|t) \leq y^p(t+j|t) \leq y^{\max} + \varepsilon^{\max}(t+j|t), \quad j = N_1, \ldots, N_2. \tag{2.40}$$

The vectors

$$E^{\min}(t)_{\rightarrow} = \begin{bmatrix} \varepsilon^{\min}(t+N_1|t) \\ \vdots \\ \varepsilon^{\min}(t+N_2|t) \end{bmatrix},$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$E^{\max}(t)_{\rightarrow} = \begin{bmatrix} \varepsilon^{\max}(t+N_1|t) \\ \vdots \\ \varepsilon^{\max}(t+N_2|t) \end{bmatrix} \tag{2.41}$$

denote additional decision variables that determine the degree of constraint easing in the case that restrictions (2.38)–(2.40) must be violated to solve the optimization problem. In practical applications of MPC, the matrices and weighting coefficients in the cost function (2.37) usually take the following values: $\mathbf{M}_{\rightarrow} = \mathbf{I}_{(N_2-N_1+1)\cdot p}$ or $\mathbf{M}_{\rightarrow} = \mathbf{0}_{(N_2-N_1+1)\cdot p}$, $\mathbf{\Lambda}_{\rightarrow} = \lambda \cdot \mathbf{I}_{N_u \cdot m}$, $\lambda > 0$, $\rho^{\min}, \rho^{\max} > 0$.

Solution (2.36) resembles the solution for predictive control of integer order in its structure and clearly shows the dual use of fractional differential calculus in FO MPC:

1. In the integer-order predictive controller, the non-negative weighting coefficients $\mu(j)$, $\lambda(j)$ are chosen by the designer. In the fractional-order predictive controller, the values of weighting matrices $\mathbf{M}_{\rightarrow}$, $\mathbf{\Lambda}_{\rightarrow}$ depend only partly on $\mu(j)$, $\lambda(j)$ and more, as can be seen in (2.27), (2.28) on fractional orders β_1, β_2. Thus, the fractional orders of discrete summation β_1, β_2 offer the controller designer new possibilities, but at the same time complicate the tuning of FO MPC;
2. The effectiveness of predictive control largely depends on the extent to which the employed model matches the controlled plant since the model is used directly in (2.33) the computation of the control signal. If the plant is of non-integer order, using an internal fractional-order model with $\alpha_1, \alpha_2, \ldots, \alpha_n$ other than ones, can in a natural way improve the control quality.

Remark. For fractional orders β_1, β_2 and $\alpha_1, \alpha_2, \ldots, \alpha_n$ equal to one, the equation (2.36) is reduced to the formula known for integer-order predictive control.

Example 2.4.1. FO MPC control of the integer-order plant with various fractional orders of summation β_1 and β_2 in the cost function.

Let us consider a non-integer order linear plant (2.12), (2.13) with

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 0.9744 & -1.976 \cdot 10^{-4} & -4.113 \cdot 10^{-7} \\ 0.9872 & 0.9999 & -2.065 \cdot 10^{-7} \\ 0.4957 & 1 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0.9872 \\ 0.4957 \\ 0.1656 \end{bmatrix}, \\ \mathbf{C} &= \begin{bmatrix} 0 & 0 & 4.167 \cdot 10^{-7} \end{bmatrix}, \quad \mathbf{D} = 0, \\ \alpha_1 &= \alpha_2 = \alpha_3 = 0.9. \end{aligned} \tag{2.42}$$

For the set-point tracking, the FO MPC controller (2.36) with the following parameters have been used: $N_1 = 1$, $N_2 = 100$, $N_u = 8$, $\mu_i = 1$ $\lambda_i = 0.1$ $L = 2000$, $\beta_1 \in \{0.6,\ 1.0,\ 1.3\}$, $\beta_2 \in \{0.6,\ 1.0,\ 1.3\}$.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2.1 shows step responses in the control system with a selected plant for all tested values β_1 and β_2.

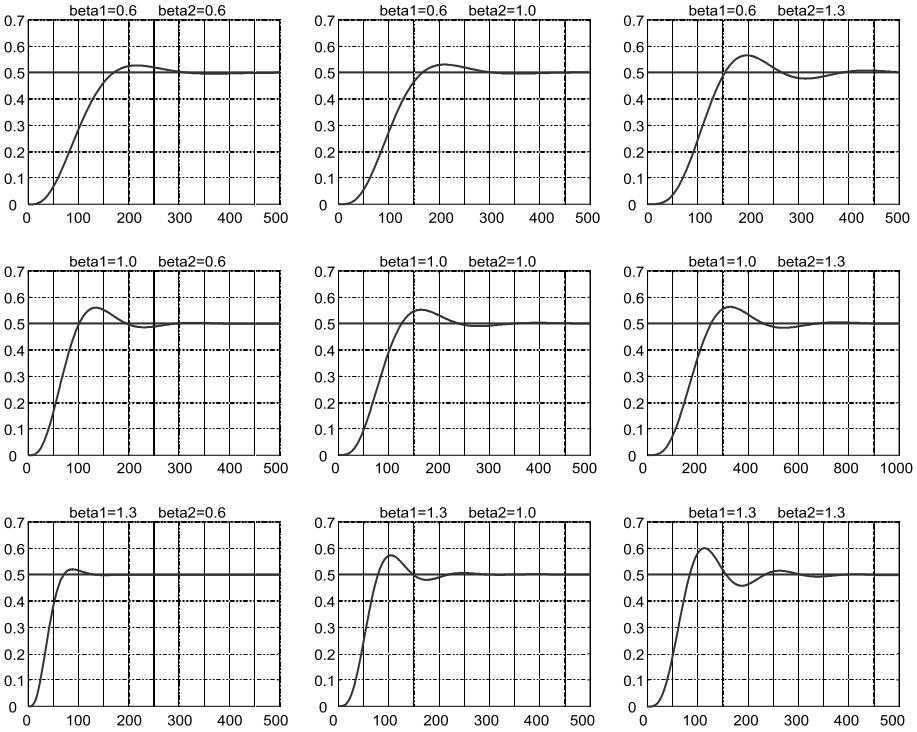


Figure 2.1: Step responses in the control system with a selected plant for several values β_1 and β_2.

Example 2.4.2. The effect of the plant–model mismatch on the fractional orders exhibited by the plant α_i and the model α_{Mi}.

Let us consider a non-integer order linear plant (2.12), (2.13) with

$$\mathbf{A} = \begin{bmatrix} 2.7756 & -1.2876 & 0.7985 \\ 2 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0.0313 \\ 0 \\ 0 \end{bmatrix}, \qquad (2.43)$$
$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{D} = 0,$$

with various combinations of the fractional orders of the plant and the model:

$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha \in \{0.80,\ 0.85,\ 0.90,\ 0.95,\ 1.00,\ 1.05,\ 1.10,\ 1.15,\ 1.20\}. \qquad (2.44)$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

For the set-point tracking the FO MPC controller (2.36) with the following parameters have been used: $N_1 = 1$, $N_2 = 10$, $N_u = 2$, $\mu_i = 0.1$ for $i = 1, 2, \ldots, 10$, $\lambda_i = 1$ for $i = 1, 2$, $L = 2000$, $\beta_1 = \beta_2 = 1$.

Figure 2.2 shows examples of step responses in the control system with a selected plant for several values $\Delta\alpha$, and Fig. 2.3 gives IAE values for the same cases.

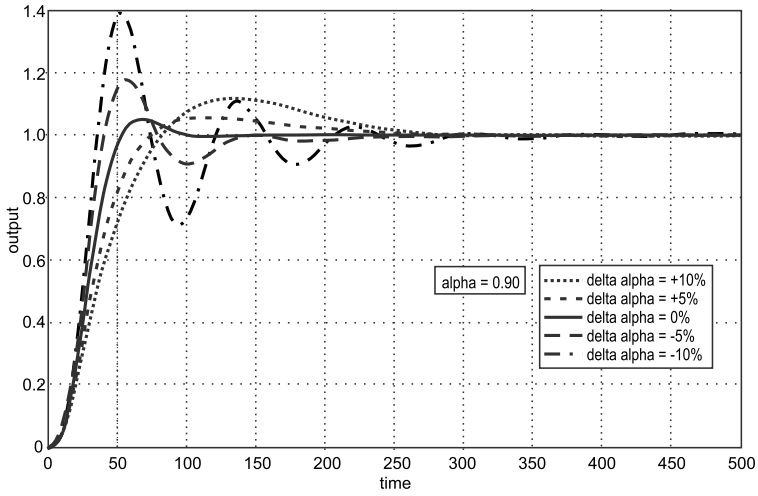


Figure 2.2: Step responses in the control system with a selected plant for several values $\Delta\alpha$.

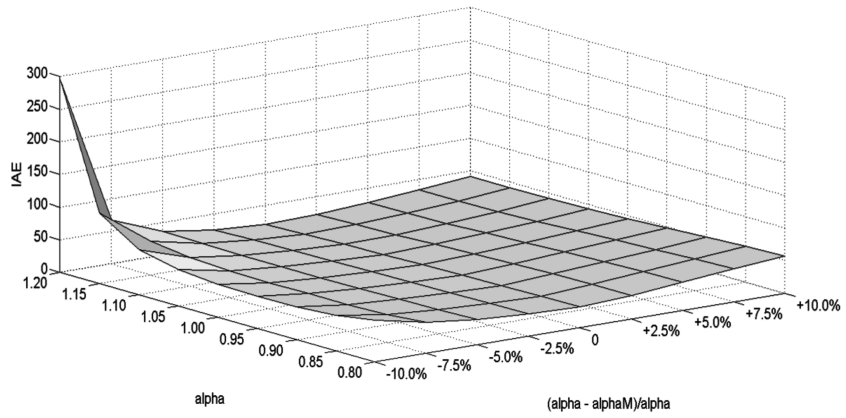


Figure 2.3: IAE indices in the control system with a selected plant for several values $\Delta\alpha$.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.5 Nonlinear suboptimal fractional-order MPC

As mentioned in previous sections, the idea of model predictive controls is general and true also for nonlinear, as well as fractional-order, processes. Unfortunately, the direct use of the nonlinear model (2.5)–(2.8) in a cost function (2.19) in FO NMPC leads to a nontrivial, time-consuming, nonlinear, and also potentially non-convex optimization problem that must be solved successively in real-time. Although, the NMPC method is the most accurate and theoretically the most valid, the computational complexity of nonlinear optimization is usually too high for today's available automation hardware platforms [13]. One of the methods to overcome these difficulties, mentioned earlier, consists of replacing a complex nonlinear fractional-order model (2.5)–(2.8) by a set of S local piecewise linear fractional-order submodels $\mathcal{M}_j$, $j \in \{1, 2, \ldots, S\}$, valid for small regions around various operating points [9, 10]:

$$x(t+1) = \mathbf{A}_j x(t) + \mathbf{B}_j u(t) - \sum_{i=1}^{t+1} (-1)^i \Upsilon_{i,j} x(t+1-i) + v(t+1), \tag{2.45}$$

$$y_j(t) = \mathbf{C}_j x(t) + d(t+1), \quad j \in \{1, 2, \ldots, S\},\ t \in \mathbb{Z}, \tag{2.46}$$

$$\Upsilon_{i,j} = \operatorname{diag}\left[\binom{\alpha_{1,j}}{i} \quad \cdots \quad \binom{\alpha_{n,j}}{i}\right], \tag{2.47}$$

where $\mathbf{A}_j$ is the complemented state matrix of the jth local model, $\Upsilon_{i,j}$ are matrices of its generalized fractional orders and $\mathbf{B}_j$, $\mathbf{C}_j$ are input and output local matrices of the jth submodel. Therefore, according to (2.42), (2.44), instantaneous properties of the modeled nonlinear process are described then by a quasilinear switched model, and hence a suboptimal MPC algorithm with a quadratic optimization problem can be developed [9].

To select piecewise linear submodels $\mathcal{M}_j$, the linear space $P = U \times Y$ made up of the input space $u(t) \in U \subseteq \mathbb{R}^m$ and the output space $y(t) \in Y \subseteq \mathbb{R}^p$ of the nonlinear process are divided into S convex polyhedrons $P_1, P_2, \ldots, P_S$, $P_j = U_j \times Y_j \subset P$, such that

$$P = \bigcup_{j=1}^{S} P_j \quad \text{and} \quad P_j \cap P_{i,i\neq j} = \varnothing \quad \forall i, j \in \{1, 2, \ldots, S\}, \tag{2.48}$$

defined by the matrices $\mathbf{S}_j^y$, $\mathbf{S}_j^u$, $\mathbf{S}_j^0$ and the following linear matrix inequalities:

$$\mathbf{S}_j^y y(t) + \mathbf{S}_j^u u(t) \leq \mathbf{S}_j^0, \quad j = 1, 2, \ldots, S. \tag{2.49}$$

It is assumed that inside the polyhedron P_j the linear submodel $\mathcal{M}_j$ is the best local linearization of the nonlinear process (2.5)–(2.8) and that only one submodel can be active at the same time.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Remark. For FO NMPC with state-space submodels (2.45)–(2.47), an alternative formulation is possible with the linear space $P = U \times X$ made up of the input space $u(t) \in U \subseteq \mathbb{R}^m$ and the state space $x(t) \in X \subseteq \mathbb{R}^n$ of the nonlinear process.

Using the concept of switched models, various strategies of suboptimal non-linear fractional-order model predictive control can be proposed. All of them are based on the fact that according to (2.33)–(2.35), for switched models the output prediction vector (2.30) assumes the following form:

$$Y^p(t)_{\rightarrow} = \mathbf{E}_{Y,j}\Delta U(t)_{\rightarrow} + Y^0(t)_{\rightarrow}, \tag{2.50}$$

with a output dynamic matrix of the model

$$\mathbf{E}_{Y,j} = \begin{bmatrix} \mathbf{C}_j \sum_{i=0}^{N_1-1} \Phi^{Y_j}(i)\mathbf{B}_j & \cdots & \mathbf{0}_{n\times m} \\ \mathbf{C}_j \sum_{i=0}^{N_1} \Phi^{Y_j}(i)\mathbf{B}_j & \cdots & \mathbf{0}_{n\times m} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_j \sum_{i=0}^{N_u-1} \Phi^{Y_j}(i)\mathbf{B}_j & \cdots & \mathbf{C}_j\mathbf{B}_j \\ \vdots & \ddots & \vdots \\ \mathbf{C}_j \sum_{i=0}^{N_2-1} \Phi^{Y_j}(i)\mathbf{B}_j & \cdots & \mathbf{C}_j \sum_{i=0}^{N_2-N_u} \Phi^{Y_j}(i)\mathbf{B}_j \end{bmatrix}. \tag{2.51}$$

Thus, adopting strategies known from classic integer-order suboptimal NMPC algorithms [4, 14, 29], for fractional-order nonlinear model predictive control the following suboptimal algorithms can be proposed:

1. with successive linearization for current operating point—FO MPC SL;
2. with nonlinear prediction and linearization for current operating point—FO MPC NPL;
3. with nonlinear prediction and linearization for current operating point along the predicted trajectory—FO MPC NPLPT.

In the FO MPC SL algorithm, the dynamic matrix (2.51) in (2.50) is successively determined at each sampling instant according to the actual solution of (2.49). The linearized free trajectory in (2.50) is also computed successively, according to (2.13)–(2.17) and (2.31), (2.32) [4].

In te FO MPC NPL algorithm, it is assumed that, as for linear systems, where the superposition principle in (2.33) applies, one can write (2.50), but only the first, forced response component is calculated from the linearized model in the form of the dynamic matrix (2.51) successively determined as in FO MPC SL. The second component, including prediction of the free trajectory depending on the past, as well as estimates of the unmeasured state disturbance and output disturbance vectors, is calculated on the basis of a more accurate nonlinear model (2.5)–(2.11).

In the FO MPC NPLPT algorithm, the linearization is carried out not only for the current operating point, as in the FO MPC NPL, but also along the predicted trajec-

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

tory, i. e., the quadratic optimization and nonlinear prediction of the free trajectory and disturbance vectors are repeated few times at each sampling instant [4]. Once the membership of the input vector $u(t-1)$ and the output vector $y(t)$ in the polyhedron $P_j = U_j \times Y_j$ has been established at each sampling instant t on the basis of the matrix inequality system (2.49), the initial model dynamics matrix $\mathbf{E}^1_{Y,j}$ for the whole prediction horizon and then the initial predicted trajectory $Y^p(t)^1_{\rightarrow}$ and the initial solution $\Delta U(t)^1_{\rightarrow}$ to the quadratic optimization problem (2.37) are determined. Further search for the optimal solution $\Delta U(t)^i_{\rightarrow}$ is done iteratively with the index $i \geq 2$. Based on $Y^p(t)^{i-1}_{\rightarrow}$ and $\Delta U(t)^{i-1}_{\rightarrow}$, switching sequences for local submodels, according to the set of matrix inequalities (2.49), are determined by:

$$\begin{gathered} \mathcal{M}^i_{k,K^i_k|j}, \quad k = 1,2,\ldots,S^i,\ j \in \{1,2,\ldots,S\}, \\ 1 \leq K^i_k \leq N_2 - N_1 + 1, \quad \sum_{k=1}^{S^i} K^i_k = N_2 - N_1 + 1, \end{gathered} \tag{2.52}$$

i. e., the number of models $S^i \leq S$ switched at the ith iteration step, an ordered list of switched models $\mathcal{M}^i_k$ from the set of all submodels $\{\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_S\}$ and the length $\{K^i_1, K^i_2, \ldots, K^i_{S^i}\}$ of the portion of the prediction horizon where these submodels are active. On this basis, according to (2.51), the dynamics matrix is created at the ith iteration step

$$\mathbf{E}^i_Y = \begin{bmatrix} \mathbf{E}^i_{Y,j|1} \in \mathbb{R}^{K^i_1 \cdot n \times N_u \cdot m} \\ \mathbf{E}^i_{Y,j|2} \in \mathbb{R}^{K^i_2 \cdot n \times N_u \cdot m} \\ \vdots \\ \mathbf{E}^i_{Y,j|k} \in \mathbb{R}^{K^i_k \cdot n \times N_u \cdot m} \\ \vdots \\ \mathbf{E}^i_{Y,j|S^i} \in \mathbb{R}^{K^i_{S^i} \cdot n \times N_u \cdot m} \end{bmatrix}, \tag{2.53}$$

and the suboptimal vector of future sequences of increments in the manipulated variable is determined at the current sampling instant $\Delta U(t)^i_{\rightarrow}$. Iterations are carried out until one of the two conditions that interrupt the iterative search for the suboptimal vector of future sequences of increments in the manipulated variable at the current sampling instant t is fulfilled:

$$\left\| \Delta U(t)^i_{\rightarrow} - \Delta U(t)^{i-1}_{\rightarrow} \right\| \leq \delta_u \left\| \Delta U(t)^{i-1}_{\rightarrow} \right\| \quad \text{or} \quad i > \delta_i, \tag{2.54}$$

which means a small improvement in the sought vector of increments or exceeding the time limit allotted for iterative searches. The parameters δ_u, δ_i are to be chosen experimentally.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figures 2.4, 2.5, and 2.6 show the general principle for computing a manipulated variable in the mentioned algorithms.

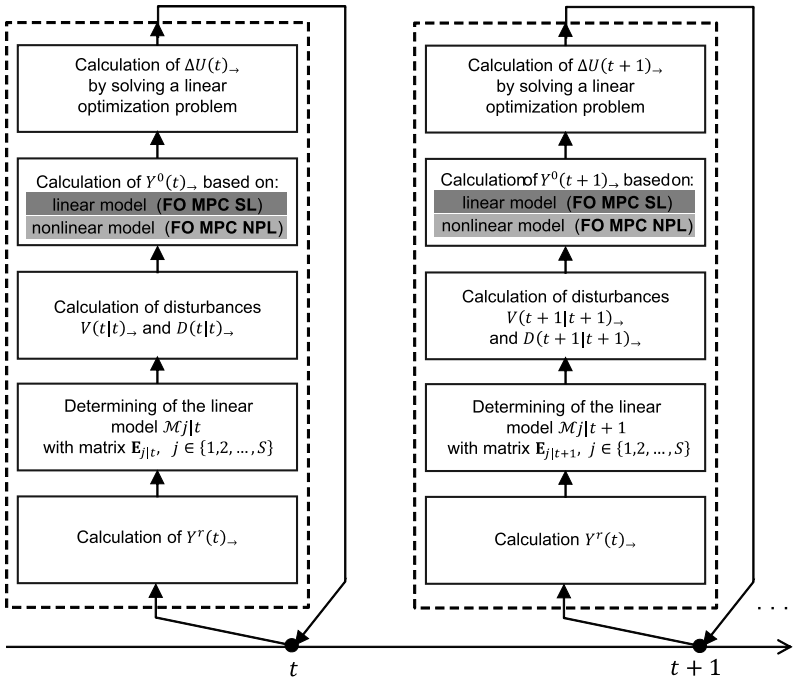


Figure 2.4: Calculation schemes of FO MPC SL and FO MPC NPL algorithms.

2.6 Conclusion

This work discusses the possibilities of using the fractional differential calculus in model predictive control. It has been known for several years that the actual properties of many complex phenomena and nonlinear industrial processes can be effectively modeled using fractional-order differential or difference calculus. This is of particular importance in the case of control methods that make direct use of the process model, as in the model predictive control. Similarly, the fractional orders of discrete summation in the cost function have a pronounced effect on the control signal and offer the controller designer new possibilities.

The article describes the two-fold use of fractional-order differential calculus in linear model predictive control. Next, the fractional-order nonlinear model predictive control algorithms were considered, and three types of suboptimal, computationally

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

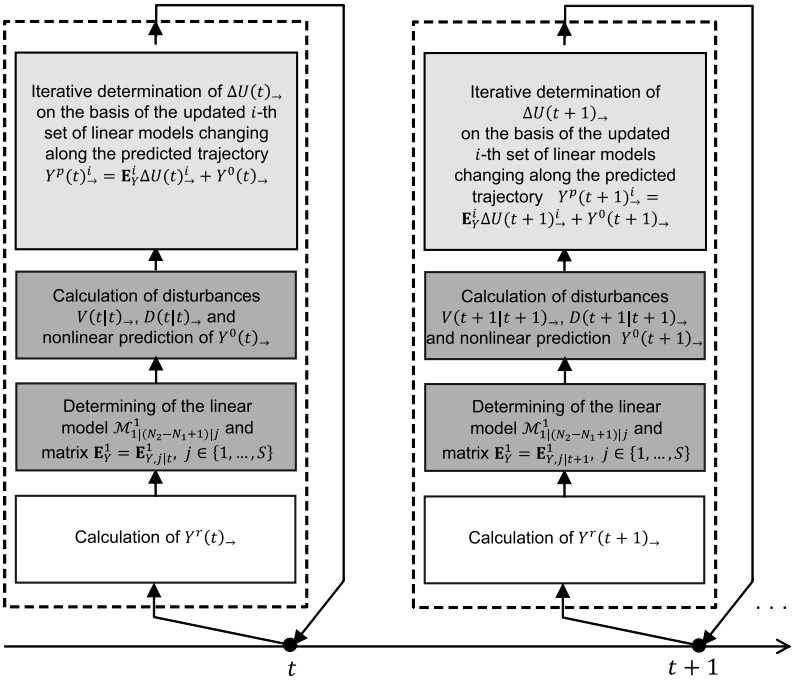


Figure 2.5: Calculation schemes of the switched FO MPC NPLPT algorithm at the sampling instants.

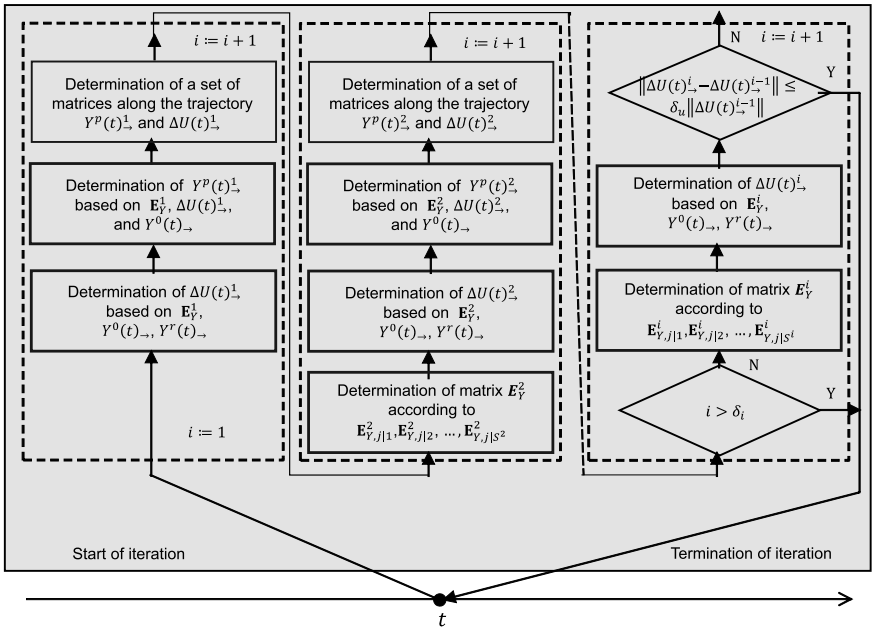


Figure 2.6: Calculation schemes of the switched FO MPC NPLPT algorithm–iterative part.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

efficient linearized FO NMPC algorithms with precise calculation schemes were proposed.

For the synthesis of proposed algorithms, quasi-linear switched models of non-integer order were used. On this basis, it is possible to implement the algorithms in PAC programmable controllers.

Bibliography

[1] O. P. Agrawal and D. Baleanu. A Hamiltonian formulation and a direct numerical scheme for fractional optimal control problems. *Journal of Vibration and Control*, 13:1269–1281, 2007.

[2] E. F. Camacho and C. Bordons. *Model Predictive Control*. Springer, London, 2004.

[3] Y. Q. Chen, I. Petráš, and D. Xue. Fractional order control. pages 1397–1410, St. Louis, 2009. American Control Conference, ACC.

[4] S. Domek. *Fractional-Order Differential Calculus in Model Predictive Control* (in Polish). West Pomeranian University of Technology Academic Press, Szczecin, 2013.

[5] S. Domek. Piecewise affine representation of discrete in time, non-integer order systems. *Lecture Notes in Electrical Engineering*, 257:149–159, 2013.

[6] S. Domek. Switched state model predictive control of fractional-order nonlinear discrete-time systems. *Asian Journal of Control, Special Issue*, 15(3):658–668, 2013.

[7] S. Domek. Multiple use of the fractional-order differential calculus in the model predictive control. pages 359–362, 2014.

[8] S. Domek. Model-plant mismatch in fractional order model predictive control. *Lecture Notes in Electrical Engineering*, 357:281–291, 2016.

[9] S. Domek. Switched fractional state-space predictive control methods for non-linear fractional systems. *Lecture Notes in Electrical Engineering*, 559:113–127, 2020.

[10] S. Domek. Switched models of non-integer order. *Studies in Systems, Decision and Control*, 296:153–185, 2021.

[11] A. Dzieliński, D. Sierociuk, and G. Sarwas. Some applications of fractional order calculus. *Bulletin of the Polish Academy of Sciences, Technical Sciences*, 58(4):583–592, 2010.

[12] T. Kaczorek. *Selected Problems of Fractional Systems Theory*. Springer, Berlin, 2011.

[13] M. Ławryńczuk. *Computationally Efficient Model Predictive Control Algorithms. A Neural Network Approach*, volume 3. Springer, London, 2014.

[14] M. Ławryńczuk. Nonlinear state-space predictive control with on-line linearisation and state estimation. *International Journal of Applied Mathematics and Computer Science*, 25(4):833–847, 2015.

[15] Y. Li, Y. Q. Chen, and H.-S. Ahn. Fractional-order iterative learning control for fractional-order linear systems. *Asian Journal of Control*, 13(1):54–63, 2011.

[16] J. M. Maciejowski. *Predictive Control with Constraints*. Prentice Hall, Englewood Cliffs, 2002.

[17] E. Mehmet. Fractional fuzzy adaptive sliding-mode control of a 2-DOF direct drive robot arm. *IEEE Transactions on Systems, Man and Cybernetics*, 38:1561–1570, 2008.

[18] A. I. Nafsun and N. Yusoff. Effect of model-plant mismatch on MPC controller performance. *Journal of Applied Sciences*, 21(11):3579–3585, 2011.

[19] P. Ostalczyk. The non-integer difference of the discrete-time function and its application to the control system synthesis. *International Journal of Systems Science*, 31(12):1551–1561, 2000.

[20] A. Oustaloup. *La Derivation Non Entiere: Theorie, Synthese et Applications*. Hermes, Paris, 1995.

[21] I. Podlubny. *Fractional Differential Equations*. Academic Press, San Diego, 1999.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[22] J. Rawlings and D. Mayne. *Model Predictive Control: Theory and Design*. Nob Hill Publishing, Madison, 2009.
[23] M. Romero, Á. P. De Madrid, C. Mañoso, and B. M. Vinagre. Fractional-order generalized predictive control: Formulation and some properties. page 1495–1500, Singapore, 2010. 11th Int. Conf. Control, Automation, Robotics and Vision.
[24] M. Romero, Á. P. De Madrid, C. Mañoso, and B. M. Vinagre. A survey of fractional-order generalized predictive control. pages 6867–6872, Maui, 2012. IEEE 51st annual conference on decision and control, CDC.
[25] D. Shantanu. *Functional Fractional Calculus for System Identification and Controls*. Springer Verlag, Berlin, 2008.
[26] C. J. Silva and D. F. M. Torres. Stability of a fractional HIV/AIDS model. *Mathematics and Computers in Simulation*, 164:180–190, 2019.
[27] P. Sopasakis and H. Sarimveis. Stabilising model predictive control for discrete-time fractional-order systems. *Automatica*, 75:24–31, 2017.
[28] R. Stanisławski and K. Latawiec. Normalized finite fractional differences—the computational and accuracy breakthroughs. *International Journal of Applied Mathematics and Computer Science*, 22(4):907–919, 2012.
[29] P. Tatjewski. *Advanced Control of Industrial Processes*. Springer, London, 2007.
[30] P. Tatjewski. Disturbance modeling and state estimation for offset-free predictive control with state-space models. *International Journal of Applied Mathematics and Computer Science*, 24(2):313–323, 2014.
[31] Y. Wang and X. Wang. Performance diagnosis of mpc with model-plant mismatch. pages 78–82, Xuzhou, 2010. Chinese Control and Decision Conf.
[32] D. Xue and Y. Chen. A comparative introduction of four fractional order controllers. pages 3228–3235, Shanghai, 2002. 4th IEEE World Congress on Intelligent Contr. and Autom.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Tyrone E. Duncan and Bozenna Pasik-Duncan

3 Stochastic control systems with long-range dependent noise

Abstract: The importance and analysis of noise models other than Brownian motion for stochastic control systems are described. These processes include fractional Brownian motions, Gauss–Volterra processes and Rosenblatt processes. The first two are Gaussian processes with long-range dependence and the third comprises non-Gaussian processes with long-range dependence. Systems with these noise processes are important for models of many physical systems where Brownian motions have been shown to be inadequate. An ergodic optimal control for a two- dimensional linear stochastic system with a quadratic cost and a Rosenblatt noise process is explicitly solved to demonstrate the solvability of some of these models.

Keywords: Gauss–Volterra processes, fractional Brownian motions, Rosenblatt processes, linear-quadratic control

3.1 Introduction

Historically, Brownian motions have been used for models of noise in physical systems primarily because of their mathematical tractability. The most important mathematical properties of these processes are that Brownian motions are Gaussian processes and martingales. Both of these properties are very useful for the solutions of control problems. However the appropriateness of Brownian motion models for modeling physical control systems is often difficult or impossible to justify from the physical data. Domanski [6] has demonstrated this need for other noise models for many control systems. Furthermore, there are mathematical justifications from the appropriateness of non-Gaussian models from non-Central Limit theorems [5].

In this paper some processes are described that are either Gaussian or non-Gaussian that have a long-range dependence property. They include fractional Brownian motions, Gauss–Volterra processes, and Rosenblatt processes. The Rosenblatt processes are non-Gaussian processes that have a stochastic calculus that allows for the formulation and solution of some stochastic control problems that lead to explicit optimal controls. An ergodic optimal control problem for a two-dimensional linear system and a quadratic cost functional is explicitly solved to demonstrate the feasibility of these models for control problems.

Tyrone E. Duncan, Bozenna Pasik-Duncan, Department of Mathematics, University of Kansas, Lawrence, KS 66045, USA, e-mails: teduncan@ku.edu, bozenna@ku.edu

https://doi.org/10.1515/9783110729122-003

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.2 Some noise processes

Initially, some definitions are given to introduce the various noise processes that are considered to clarify the subsequent discussion and to fix notation. A fractional Brownian motion is considered for the Hurst parameter, $H \in (\frac{1}{2}, 1)$. It is well known that for $H = \frac{1}{2}$ the Gaussian process is Brownian motion and for $H \in (\frac{1}{2}, 1)$ the Gaussian process is a fractional Brownian motion that has a long-range dependence, which is evidenced by the relatively slow decay of the covariance function. For the discussion here of the noise processes, they are subdivided into two families, Gaussian and non-Gaussian. The Gaussian processes are obtained from Brownian motions by single singular integrals, i. e., Wiener integrals having singular integrands or kernels. The general family of these processes are called Gauss–Volterra processes because they are Gaussian processes that are defined via Volterra kernels. They include fractional Brownian motions with the Hurst parameter $H \in (\frac{1}{2}, 1)$ and some other Gaussian processes that have been specifically investigated in the literature. The Gauss–Volterra processescan be viewed as a natural generalization of fractional Brownian motions as they are now described. The scalar process $(b(t), t \geq 0)$ is a Gauss–Volterra process with zero mean that can be described by its covariance function, R, as

$$R(t,s) = \mathbb{E}\big[b(t)b(s)\big] := \int_0^{\min(t,s)} K(t,r)K(s,r)dr, \tag{3.1}$$

where the kernel $K : \mathbb{R}_+^2 \to \mathbb{R}$ satisfies the following four conditions:

- (K1) $K(t,s) = 0$ for $s > t$, $K(0,0) = 0$, and $K(t,\cdot) \in L^2(0,t)$ for each $t \in \mathbb{R}_+$.
- (K2) For each $T > 0$ there are positive constants C, β such that

$$\int_0^T \big(K(t,r) - K(s,r)\big)^2 dr \leq C|t-s|^\beta, \quad t,s \in (0,T]. \tag{3.2}$$

- (K3)
  - (i) $K = K(t,s)$ is differentiable in the first variable in $\{0 < s < t < \infty\}$, both K and $\frac{\partial}{\partial t}K$ are continuous and $K(s+,s) = 0$ for each $s \in [0,\infty)$
  - (ii) $|\frac{\partial K}{\partial t}(t,s)| \leq c_T(t-s)^{\alpha-1}(\frac{t}{s})^\alpha$
  - (iii) $\int_0^t (K(t,u))^2 du \leq c_T(t-s)^{1-2\alpha}$ on the set $\{0 < s < t < T\}$, $T < \infty$, for some constants $c_T > 0$ and $\alpha \in (0, \frac{1}{2})$.
- (K4) Let $\alpha(t) := \frac{\partial}{\partial t}(\int_0^t (\mathcal{K}_t^*\sigma)^2(r)dr)$, and assume that $\alpha \in C(\mathbb{R}_+)$, where

$$(\mathcal{K}_T^*\sigma)(s) := K(s+,s)\sigma(s) + \int_s^T \sigma(r)K(dr,s) \tag{3.3}$$

  and $\mathcal{K}_T^*$ is injective.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

It is assumed that there is a real-valued standard Wiener process $(W(t), t \geq 0)$ such that

$$b(t) = \int_0^t K(t,r)dW(r), \quad t \in \mathbb{R}_+ \tag{3.4}$$

(conditions when a Volterra process admits such a representation (3.4) have been obtained, cf. [13]). From (K2) it easily follows by the Kolmogorov sample-path continuity test that $(b(t), t \geq 0)$ has a continuous modification, which is the version that is chosen for the subsequent discussion here. It is assumed that, for all $s \in [0,T)$, $T > 0$, $K(\cdot, s)$ has bounded variation on the interval (s,T) and

$$\int_0^T |K|^2((s,T],s)ds < \infty, \tag{3.5}$$

where $|K|$ denotes the variation of K. Three examples of Gauss–Volterra processes satisfying the above conditions are:

(i) A fractional Brownian motion (FBM) with the Hurst parameter $H \in (\frac{1}{2}, 1)$. In this case

$$\begin{aligned} K(t,s) &= C_H s^{1/2-H} \int_s^t (u-s)^{H-3/2} u^{H-1/2} du, \quad s < t \\ &= 0, \quad t \leq s. \end{aligned} \tag{3.6}$$

The kernel satisfies conditions (K1)–(K3) with $\alpha = H - \frac{1}{2}$.

(ii) The Liouville fractional Brownian motion (LFBM, cf. [2]) for $H \in (\frac{1}{2}, 1)$, in which case

$$K(t,s) = C_H (t-s)^{H-\frac{1}{2}} 1_{(0,t]}(s), \quad t > s,\ t,s \in \mathbb{R}_+ \tag{3.7}$$

satisfies (K1)–(K3) with $\alpha = H - \frac{1}{2}$.

(iii) The multifractional Brownian motion (MBM). A simplified version analogous to LFBM in Example (ii) is considered. The kernel $K : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is defined as

$$K(t,s) = (t-s)^{H(t)-\frac{1}{2}}\, 1_{(0,t]}(s), \quad t,s \in \mathbb{R}_+,$$

where $H : \mathbb{R}_+ \to [\frac{1}{2}, 1)$ is the "time-dependent Hurst parameter".

The other family of processes considered here are called Rosenblatt processes because M. Rosenblatt [14] was apparently the first one to consider them. Note that they are a natural generalization of fractional Brownian motions for $H \in (\frac{1}{2}, 1)$ though they are

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

non-Gaussian because they are double Wiener integrals. The definition of the Rosenblatt processes indicates how one can define other non-Gaussian processes with a long-range dependence by taking higher-order multiple Wiener integrals with singular kernels. The Rosenblatt processes are defined as follows.

Let $H \in (1/2, 1)$. The *Rosenblatt process* $R_H = (R_H(t), t \geq 0)$ is defined as

$$\begin{aligned} R_H(t) &= C_H^R \int_{\mathbb{R}^2} \left(\int_0^t h_2^H(u, y_1, y_2) du \right) dW(y_1) dW(y_2) \\ &= C_H^R \int_0^s u_{(\frac{H}{2}-\frac{1}{4})}(y_1) I_{s-}^{-(\frac{H}{2}-\frac{1}{4})} I_{t-}^{(\frac{H}{2}-\frac{1}{4})} u_{(\frac{H}{2}-\frac{1}{4})}(y_1) \\ &\quad \times 1_{[s,t)}(y_1) u_{(\frac{H}{2}-\frac{1}{4})}(y_2) I_{s-}^{-(\frac{H}{2}-\frac{1}{4})} I_{t-}^{(\frac{H}{2}-\frac{1}{4})} \\ &\quad \times u_{(\frac{H}{2}-\frac{1}{4})}(y_2) 1_{[s,t)}(y_2) du \, dW(y_1) dW(y_2), \end{aligned}$$

and the fractional Brownian motion $(B_H(t), t \geq 0)$ is defined as

$$B_H(t) = C_H^B \int_{\mathbb{R}} \left(\int_0^t h_1^H(u, y) du \right) dW_y, \tag{3.8}$$

where for $(u)_+ = \max\{u, 0\}$ (the positive part of u) and h_k^H is defined as

$$h_k^H(u, y) = \prod_{j=1}^k (u - y_j)_+^{\frac{H}{k} - (\frac{1}{k} + \frac{1}{2})}. \tag{3.9}$$

The constants C_H^R and C_H^B are chosen so that the second moments of R_H and B_H at $t = 1$ are 1.

The normalizing constants C_H^B and C_H^R in the previous two definitions are given explicitly as

$$C_H^B = \sqrt{\frac{H(2H-1)}{\mathrm{B}(2-2H, H-\frac{1}{2})}}, \quad C_H^R = \frac{\sqrt{2H(2H-1)},}{2\mathrm{B}(1-H, \frac{H}{2})}$$

where B is the Beta function. For the subsequent Itô-type formula (change of variables), it is also convenient to define the following constants

$$c_H^B = C_H^B \, \Gamma\left(H - \frac{1}{2}\right), \quad c_H^R = C_H^R \, \Gamma\left(\frac{H}{2}\right)^2,$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

and

$$c_H^{B,R} = \frac{c_R^H}{c^B_{\frac{H}{2}+\frac{1}{2}}} = \sqrt{\frac{(2H-1)}{(H+1)}\frac{\Gamma(1-\frac{H}{2})\Gamma(\frac{H}{2})}{\Gamma(1-H)}}, \tag{3.10}$$

where Γ is the Gamma function.

Some approaches to a stochastic integration for Rosenblatt processes are given in [1, 4, 15, 16]. The following result is a change of variables (Itô formula) for a Rosenblatt process [4]. Note that in general there is a third derivative term in the result contrary to the case for Brownian motion.

Proposition 3.2.1. *Let $y_0 \in \mathbb{R}$ be a deterministic constant, $\vartheta \in L^1(0,T;\mathbb{D}^{2,2})$, and define the process $(y_t)_{t\geq 0}$ by*

$$y_t \overset{\text{Def.}}{=} y_0 + \int_0^t \vartheta_s ds + R_t^H.$$

Let f be a function in $\mathscr{C}^3(\mathbb{R}_+ \times \mathbb{R})$ such that its second and third partial derivatives in the second variable are locally bounded. Consider the following three conditions for $T > 0$, where $\mathbb{D}^{m,n}$ are suitable Malliavin spaces:

1. *The process $\frac{\partial f}{\partial x}(s, y_s)$ belongs to $L^{\frac{1}{H}}(0,T;\mathbb{D}^{2,2})$.*
2.
   (a) *It holds that*

$$\int_0^T \left\| \left(\nabla^{\frac{H}{2}} \frac{\partial f}{\partial x}(s, y_s) \right)(s) \right\|_{\mathbb{D}^{1,2}}^{\frac{2}{1+H}} ds < \infty.$$

   (b) *There is a non-negative function p_1 that belongs to the space $L^{\frac{2}{1+H}}(0,T)$ such that, for almost every $s \in [0,T]$, the estimate*

$$\left\| \left(\nabla^{\frac{H}{2}} \frac{\partial f}{\partial x}(s, y_s) \right)(u) \right\|_{\mathbb{D}^{1,2}} \leq p_1(s)$$

   holds for almost every $u \in [s,T]$.

   (c) *For almost every $s \in [0,T]$, the map from $\mathbb{R}$ to $\mathbb{D}^{1,2}$ given by*

$$u \mapsto \left(\nabla^{\frac{H}{2}} \frac{\partial f}{\partial x}(s, y_s) \right)(u)$$

   is right continuous at the point s.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.

(a) *It holds that*

$$\int_0^T \left\| \left(\nabla^{\frac{H}{2},\frac{H}{2}} \frac{\partial f}{\partial x}(s, y_s) \right)(s,s) \right\|_{L^2(\Omega)} ds < \infty.$$

(b) *There is a non-negative function p_2 that belongs to the space $L^1(0,T)$ such that, for almost every $s \in [0,T]$, the estimate*

$$\left\| \left(\nabla^{\frac{H}{2},\frac{H}{2}} \frac{\partial f}{\partial x}(s, y_s) \right)(u,u) \right\|_{L^2(\Omega)} \leq p_2(s)$$

holds for almost every $u \in [s,T]$.

(c) *For almost every $s \in [0,T]$, the map from $\mathbb{R}$ to $L^2(\Omega)$ given by*

$$u \mapsto \left(\nabla^{\frac{H}{2},\frac{H}{2}} \frac{\partial f}{\partial x}(s, y_s) \right)(u,u)$$

is right continuous at the point s.

If conditions (1)–(3c) *are satisfied for every $T > 0$, then the process $(Y_t)_{t\geq 0}$ defined by $Y_t = f(t, y_t)$ satisfies the equation*

$$Y_t = Y_0 + \int_0^t \tilde{\vartheta}_s ds + 2c_H^{B,R} \int_0^t \tilde{\varphi}_s \delta B_s^{\frac{H}{2}+\frac{1}{2}} + \int_0^t \tilde{\psi}_s \delta R_s^H \tag{3.11}$$

for every $t \geq 0$, where

$$\begin{aligned}
\tilde{\vartheta}_s &= \frac{\partial f}{\partial s}(s, y_s) + \frac{\partial f}{\partial x}(s, y_s)\vartheta_s \\
&\quad + c_H^R \frac{\partial^2 f}{\partial x^2}(s, y_s)(\nabla^{\frac{H}{2},\frac{H}{2}} y_s)(s,s) \\
&\quad + c_H^R \frac{\partial^3 f}{\partial x^3}(s, y_s)\left[(\nabla^{\frac{H}{2}} y_s)(s)\right]^2, \\
\tilde{\varphi}_s &= \frac{\partial^2 f}{\partial x^2}(s, y_s)(\nabla^{\frac{H}{2}} y_s)(s), \\
\tilde{\psi}_s &= \frac{\partial f}{\partial x}(s, y_s).
\end{aligned}$$

A proof of this result is given in [4]. The stochastic integrals given here are Skorokhod integrals so they have expectation zero.

The change of variables formula contains the following two differential operators,

$$\nabla^{\frac{H}{2}} = I_+^{\frac{H}{2}} D, \tag{3.12}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$\nabla^{\frac{H}{2},\frac{H}{2}} = I_{+,+}^{\frac{H}{2},\frac{H}{2}} D^2, \tag{3.13}$$

where D is the Malliavin derivative and

$$I_{+}^{\alpha}(f(x)) = \int_{-\infty}^{x} f(u,v)(x-u)^{\alpha-1}du, \tag{3.14}$$

$$(I_{+,+}^{\alpha_1,\alpha_2} f)(x_1,x_2) \overset{\text{Def.}}{=} \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f(u,v) \times (x_1-u)^{\alpha_1-1}(x_2-v)^{\alpha_2-1}dudv.$$

These operators reflect the singular integral definition of a Rosenblatt process.

Similar to the case for Brownian motion, the previous change of variables formula is a basic tool for solving problems of control and estimation for models that contain a Rosenblatt noise process. Unfortunately, some important structure is lost when the noise is not a Brownian motion, especially the property of martingales. The controls are assumed to be feedback functions of only the current state. This approach has also been effective for the solution of control problems with a fractional Brownian motion for $H \in (\frac{1}{2},1)$ [8] and for more general Gaussian noise processes [9]. Some results are also available for stochastic differential games with general noise processes [7]. The properties of non-Gaussian and long-range dependence from Rosenblatt processes should make them an attractive alternative to Brownian motion or fractional Brownian motion noise models, especially given the results of Domanski [6].

To indicate the usefulness of the model with Rosenblatt noise for explicit control solutions, a two-dimensional stochastic control system is described with a quadratic cost functional. The controlled stochastic system satisfies the following linear stochastic equation:

$$dX(t) = AX(t)dt + BU(t)dt + dR_H(t), \tag{3.15}$$

$$X(0) = x_0, \tag{3.16}$$

where $X(t) \in \mathbb{R}^2$, $A \in \mathcal{L}(\mathbb{R}^2,\mathbb{R}^2)$, $B \in \mathcal{L}(\mathbb{R}^2,\mathbb{R}^2)$ is $B = I$, $(R_H(t), t \geq 0)$ is a standard two-dimensional Rosenblatt process with parameter $H \in (\frac{1}{2},1)$ for both independent components of the two-dimensional Rosenblatt process. It is noted that more generally the two noise components can be correlated, and the components can have different H values. These extensions are fairly straightforward. All of the random variables are defined on the complete probability space $(\Omega,\mathcal{F},\mathbb{P})$.

The quadratic cost, $J_T(U)$, is

$$J_T(U) = \mathbb{E}\int_0^T (\langle QX(t),X(t)\rangle + \langle RU(t),U(t)\rangle)dt, \tag{3.17}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

where Q and R are symmetric and positive definite linear transformations, and $T > 0$ is initially fixed. The problem that is solved is an ergodic control problem from this finite-time problem, i. e.,

$$J_\infty = \limsup_{T\to\infty} \frac{1}{T} J_T. \tag{3.18}$$

The family of admissible controls, $\mathcal{U}$, is the collection of constant linear feedbacks of the state X, i. e.,

$$\mathcal{U} = \{U(t) = KX(t) \mid K \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)\}. \tag{3.19}$$

This family of feedback controls is quite natural from the result for a Brownian motion noise and to obtain explicit and implementable optimal controls. However, allowing the controls to be adapted to the past of the state process in this case would imply functional optimal controls with functional dependence on the past of the state because the control would be predicting the future of the state process, as is the case with fractional Brownian motions [10]. Such controls are not easily implementable. Furthermore, the approach in this paper has been successful for scalar linear systems driven by a Rosenblatt process [3, 12] and for multidimensional linear and bilinear systems driven by fractional Brownian motions and Gauss–Volterra processes [3, 8, 9, 11, 12].

3.3 Optimal feedback control

The Riccati equation, which is used for some computations here, is the one used for a Brownian motion noise so it is not intrinsic for a Rosenblatt noise but it suffices for some computations. It is the following equation:

$$\frac{dP}{dt} = -A^T P - PA + B^T P R^{-1} B P - Q, \tag{3.20}$$

$$P(T) = 0. \tag{3.21}$$

The dependence of the solution P on the interval $[0, T]$ is suppressed because it is well known that these solutions converge to the solution of the algebraic Riccati equation on the interval $[0, \infty)$. The solution of the optimal feedback control for the finite-time horizon control problem described by (3.15) and (3.17) is given in the following theorem.

Theorem 3.3.1. *The stochastic control problem with the stochastic equation* (3.15) *and the quadratic cost* (3.17) *has an optimal feedback control, K^*, given by the minimum of the following expression, which can be obtained by differentiation. The expression is*

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

strictly convex in K so the optimal K is determined by the unique zero of the derivative of g.

$$g(K) = \lim_{T\to\infty} \frac{1}{T}\left[\int_0^T \left|R^{-\frac{1}{2}}(RKX + B^T PX)\right|^2 dt + \tilde{C}_H \int_0^T \mathrm{e}^{(A+BK+A^T+K^T B^T)r} r^{2H-2} dr\right], \tag{3.22}$$

where P is the unique solution of the Riccati equation (3.20).

Proof. The change of variables formula for a Rosenblatt process is applied to the real-valued process $(\langle P(t)X(t), X(t)\rangle, t \geq 0)$, where P satisfies the aforementioned Riccati equation (3.20). Initially, a change of variables formula for Rosenblatt processes is applied to $(\langle P(t)X(t), X(t)\rangle, t \in [0,T])$, using the result that was just stated. The result applied here is the following:

$$\begin{aligned}
&\langle P(T)X(T), X(T)\rangle - \langle P(0)x_0, x_0\rangle \\
&= \int_0^T [\langle P(A + BK + A^T + K^T B^T)X, X\rangle \\
&\quad + 2c_H \operatorname{tr}(\nabla^{\frac{H}{2},\frac{H}{2}} X_s(s,s))]ds \\
&\quad + 2\int_0^T \langle \nabla^{\frac{H}{2}} X_s(s), dB_H\rangle + 2\int_0^T \langle X, dR_H\rangle \\
&\quad + \int_0^T \left\langle \frac{dP}{dt} X(s), X(s)\right\rangle ds \\
&= \int_0^T [\langle P(A + BK + A^T + K^T B^T)X, X\rangle \\
&\quad + 2c_H \operatorname{tr} \nabla^{\frac{H}{2},\frac{H}{2}} X_s(s,s))]ds \\
&\quad + 2\int \langle \nabla^{\frac{H}{2}} X_s(s), dB_H\rangle + 2\int_0^T \langle X, dR_H\rangle \\
&\quad + \int_0^T \left\langle \frac{dP}{dt} X(s), X(s)\right\rangle ds.
\end{aligned} \tag{3.23}$$

The two stochastic integrals in the equality are Skorokhod integrals so they have expectation zero. It is necessary to compute $\nabla^{\frac{H}{2},\frac{H}{2}} X_t(u,u)$. This term is the analog of the

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

second derivative in the change of variables formula for a Brownian motion noise. Initially, the process X is replaced by the Rosenblatt process R_H to determine $\nabla^{\frac{H}{2},\frac{H}{2}} X_t(u,u)$ because X to a linear transformation of R_H. Recall that the noise process in the stochastic equation (3.15) is R_H, so it is necessary to compute in the change of variables formula, $\nabla^{\frac{H}{2},\frac{H}{2}} R_H$, where R_H is a two-vector of independent real-valued Rosenblatt processes. It follows from computations in [3] that

$$\nabla^{\frac{H}{2},\frac{H}{2}} R_{H,t}(u,u) = \tilde{C}_H \int_0^t |u-r|^{2H-2} dr, \tag{3.24}$$

where the constant, $\tilde{C}_H$, is given by

$$\tilde{C}_H = 2c_H^R \frac{\mathrm{B}(\frac{H}{2}, 1-H)^2}{\Gamma(\frac{H}{2})^2}, \tag{3.25}$$

and B is the Beta function. Note that the integral on the RHS of (3.24) is a two-vector, each element having the same integrand. Let $\Xi_t(u) = \nabla^{\frac{H}{2},\frac{H}{2}} X_t(u,u)$. Then, it follows from the solution of (3.15) by the linearity of the differential operator that

$$\begin{aligned}\Xi_t(u) = &\int_0^t [(A+BK) + (A^T + K^T B^T)]\Xi_s(u) ds \\ &+ \nabla^{\frac{H}{2},\frac{H}{2}} R_{H.t}(u,u)\end{aligned} \tag{3.26}$$

because the operator $\nabla^{\frac{H}{2},\frac{H}{2}} X_t(u,u)$ is symmetric. Solving this integral equation, it follows directly from the linearity of (3.26) using (3.24) that

$$\Xi_t(u) = \tilde{C}_H \int_0^t \mathrm{e}^{(A+BK+A^T+K^T B^T)(t-r)} |u-r|^{2H-2} dr, \tag{3.27}$$

which, by an elementary change of variables, letting $t = u = s$, that

$$\Xi_s(s) = \nabla^{\frac{H}{2},\frac{H}{2}} X_s(s,s) = \tilde{C}_H \int_0^s \mathrm{e}^{(A+BK+A^T+K^T B^T)r} r^{2H-2} dr. \tag{3.28}$$

Note that the term $|u-r|^{2H-2}$ in (3.27) is a two-vector which has this same scalar term in both elements. Since $A + KB + A^T + B^T K^T$ is symmetric, it can be diagonalized. Fix K, and diagonalize the linear operator $A + KB + A^T + K^T B^T$ which will be denoted $\mathrm{diag}(a_1, \ldots, a_n)$.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Substituting the Riccati equation (3.20) in (3.23) and applying expectation, the following equation results:

$$\begin{aligned}
&\mathbb{E}\langle P(T)X(T), X(T)\rangle + \mathbb{E}\int_0^T \langle QX, X\rangle dt \\
&\quad + \mathbb{E}\int_0^T \langle RKX, KX\rangle dt \\
&= \mathbb{E}\langle P(0)x_0, x_0\rangle + \mathbb{E}\int_0^T \Bigg(\langle RKX, KX\rangle dt \\
&\quad + \langle P(BK + K^T B^T)X, X\rangle dt \\
&\quad + \int_0^T \operatorname{tr}\Bigg(\tilde{C}_H \int_0^t \mathrm{e}^{(A+BK+A^T+K^T B^T)r} r^{2H-2} dr \Bigg)\Bigg) dt \\
&= \mathbb{E}\Bigg[\langle P(0)x_0, x_0\rangle + \int_0^T \left| R^{-\frac{1}{2}}(RKX + B^T PX)\right|^2 \Bigg] dt \\
&\quad + \tilde{C}_H \int_0^T \operatorname{tr}\Bigg(\int_0^t \mathrm{e}^{(A+BK+A^T+K^T B^T)r} r^{2H-2} dr \Bigg) dt. \qquad (3.29)
\end{aligned}$$

Recall that $P(T) = 0$ from the Riccati equation (3.20).

Now, consider a limit of the inner integral for the last term on the RHS, i. e.,

$$\begin{aligned}
&\lim_{t\to\infty} \tilde{C}_H \operatorname{tr}\Bigg(\int_0^t \mathrm{e}^{(A+BK+A^T+K^T B^T)r} r^{2H-2} dr \Bigg) \\
&\quad = \sum_{i=1}^{2} \frac{\Gamma(2H-1)}{a_i^{2H-1}}, \qquad (3.30)
\end{aligned}$$

where $(a_i, i = 1, 2)$ are the eigenvalues of the symmetric transformation $(A + BK + A^T + K^T B^T)$. Clearly, averaging this result as $\frac{1}{T}\int_o^T$ converges to the same value. Now, divide the previous equality by T and let $T \to \infty$.

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T}\mathbb{E}J_T(K) \\
&\quad = \lim_{T\to\infty} \frac{1}{T}\Bigg[\mathbb{E}\int_0^T \langle QX, X\rangle dt + \mathbb{E}\int_0^T \langle RKX, KX\rangle dt \Bigg]
\end{aligned}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$= \lim_{T\to\infty}\left[\frac{1}{T}\int_0^T \left|R^{-1}(RKX + B^T PX)\right|^2 dt\right]$$
$$+ \sum_{i=1}^{2} \frac{\Gamma(2H-1)}{a_i^{2H-1}}. \tag{3.31}$$

Clearly, the minimum K can be determined from this last equality.
The same method can be used for higher-dimensional systems. □

3.4 Concluding remarks

The solution of the ergodic control problem (3.18) shows that some non-Gaussian, long-range dependent processes can be used as models for noise in control systems, and optimal controls can be determined. It is important to determine the appropriateness of the Rosenblatt processes for some noise that occurs in physical control systems. Domanski [6] has demonstrated by empirical studies of control systems the need to use non-Gaussian noise models, and Rosenblatt processes can be considered as an important candidate for modeling noise.

Bibliography

[1] B. Arras. From forward integrals to Wick–Itô integrals: The fractional Brownian motion and the Rosenblatt process cases. preprint, 2016.
[2] Z. Brzeźniak, J. van Neerven, and D. Salopek. Stochastic evolution equations driven by Liouville fractional Brownian motion. *Czechoslovak Mathematical Journal*, 62:1–27, 2012.
[3] P. Čoupek, T. E. Duncan, B. Maslowski, and B. Pasik-Duncan. An infinite time horizon linear-quadratic control with a Rosenblatt process. In *Proc. IEEE Conf. on Decision and Control*, pages 99–104, Miami, FL, 2018.
[4] P. Čoupek, T. E. Duncan, and B. Pasik-Duncan. A stochastic calculus for Rosenblatt processes. *Stochastic Processes and Their Applications*, 2012. 10.1016/j.spa.2020.01.004, to appear in journal in 2021.
[5] R. L. Dobrushin and P. Major. Non-central limit theorems for non-linear functional of Gaussian fields. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 50:27–52, 1979.
[6] P. D. Domański. Non-Gaussian properties of the real industrial control error in SISO loops. In *Proceedings of the 19th International Conference on System Theory, Control and Computing (ICSTCC)*, pages 877–882, 2015.
[7] T. E. Duncan. Linear-quadratic stochastic differential games with general noise processes. In F. El Ouardighi and K. Kogan, editors, *Models and Methods in Economics and Management Science: Essays in Honor of Charles S. Tapiero*, volume 198 of *Operations Research and Management Series*, pages 17–26. Springer International Publishing, Switzerland, 2014.
[8] T. E. Duncan, B. Maslowski, and B. Pasik-Duncan. Linear-quadratic control for stochastic equations in a Hilbert space with fractional Brownian motions. *SIAM Journal on Control and Optimization*, 50:207–231, 2012.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[9] T. E. Duncan, B. Maslowski, and B. Pasik-Duncan. Stochastic linear-quadratic control for bilinear evolution operators driven by Gauss–Volterra processes. *Applied Mathematics & Optimization*, 80:369–389, 2019.
[10] T. E. Duncan and B. Pasik-Duncan. Linear-quadratic fractional Gaussian control. *SIAM Journal on Control and Optimization*, 51:4604–4619, 2013.
[11] T. E. Duncan and B. Pasik-Duncan. Some linear-quadratic stochastic differential games driven by state dependent Gauss–Volterra processes. In G. Yin and Q. Zhang, editors, *IMA Stochastic Control, Computational Methods and Applications*, pages 173–180. Springer, 2019.
[12] T. E. Duncan and B. Pasik-Duncan. A scalar linear-quadratic two player stochastic differential game with a Rosenblatt process noise. In *Proc. 59th IEEE Conf. Decision and Control*, 2020.
[13] M. Erraoui and E. H. Essaky. Canonical representation for Gaussian processes. *Séminaire de Probabilités*, XLII:365–381, 2009.
[14] M. Rosenblatt. Independence and dependence. In *Proceedings of the 4th Berkeley Symposium Mathematical Statistics and Probability*, pages 431–443, University of California Press, 1961.
[15] F. Russo and P. Vallois. Forward, backward and symmetric stochastic integration. *Probability Theory and Related Fields*, 97:403–421, 1993.
[16] C. A. Tudor. Analysis of the Rosenblatt process. *ESAIM Probabilités Et Statistique*, 12:230–257, 2008.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Yulong Huang, Fengchi Zhu, Yonggang Zhang, Yuxin Zhao,
Peng Shi, and Jonathon Chambers

4 Outlier-robust Kalman filtering framework based on statistical similarity measure

Abstract: This chapter introduces the statistical similarity measure to quantify the similarity between two random vectors, which is then employed to develop an outlier-robust Kalman filtering framework. The approximation errors and the stability of the filter are analyzed and discussed. To implement the filter, a fixed-point iterative algorithm and a separate iterative algorithm are given, their local convergent conditions are also provided, and comparisons are made. In addition, selection of the similarity function is considered, and four exemplary similarity functions are established. Then the relationships between the filter and other existing outlier-robust Kalman filters are revealed. Simulation examples are used to illustrate the effectiveness and potential of the filtering scheme.

Keywords: Kalman filter, statistical similarity measure, outliers, heavy-tailed noise, separate iterative algorithm

4.1 Introduction

The estimation problem has been one of the most important issues from industrial applications to research areas including signal processing, optimal control, navigation, etc. The actual applications include the parameter estimate [36], system identification [38], target tracking [42], simultaneous localization and mapping [10] and many others [35]. The Kalman filter is best-known as an optimal recursive state estimator in the sense of minimum variance for a linear system with Gaussian noises [7, 25]. To solve nonlinear filtering problems, researchers have made many additions to the Kalman family filters, including extended Kalman filters (EKF) [2], second-order extended Kalman filters (SEKF) [32], unscented Kalman filters (UKF) [24], cubature

Yulong Huang, Fengchi Zhu, Yonggang Zhang, Yuxin Zhao, College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China, e-mails: heuedu@163.com, zfchiggins@163.com, zhangyg@hrbeu.edu.cn, zhaoyuxin@hrbeu.edu.cn
Peng Shi, School of Electrical and Electronic Engineering, The University of Adelaide, Adelaide, SA 5005, Australia; and College of Engineering and Science, Victoria University, Melbourne, VIC 8001, Australia, e-mail: peng.shi@adelaide.edu.au
Jonathon Chambers, College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China; and Department of Engineering, University of Leicester, Leicester LE1 7RH, UK, e-mail: Jonathon.Chambers@le.ac.uk

https://doi.org/10.1515/9783110729122-004

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kalman filters (CKF) [3] and many others. EKF is a basic extension of the KF, which approximates the nonlinear system by its first-order linearization [2]. SEKF maintains the second-order terms of the Taylor expansions of the state and measurement equations [32]. UKF was derived based on the unscented transform for approximation [24], and CKF was developed by using a spherical–radial cubature rule for approximating the Gaussian filter [3]. However, the estimation accuracy of such Kalman filters degrades dramatically when the system is eroded by state and measurement outliers caused by external interference or unreliable sensors [17]. The state and measurement noises are non-Gaussian distributed in such cases [5].

For such non-Gaussian filtering problems, robust filters have been studied. Examples include the fixed-lag doubly robust smoothing (DRS) [12], sum-of-norms regularization [29], non-smooth quadratic support (QS) penalties [4], and so on. Moreover, H_∞ filters are robust filters that can solve efficiently the model uncertainty problems [14, 28, 40]. The finite impulse response (FIR) state estimation are also proposed, which estimates the current state using the measurements over only an interval of N most recent neighboring points called a horizon [43]. The particle filter (PF) can achieve an approximate state estimate by modelling the noises as non-Gaussian heavy-tailed distributed and approximating the posterior probability density function (PDF) as a set of weighted random samples using the sequential Monte Carlo sampling technique [5, 27]. The Gaussian sum filter (GSF) can also be used to address the non-Gaussian filtering problem by running a group of Kalman filters, in which a finite sum of Gaussian distributions are employed to model non-Gaussian noises, and the posterior PDF can be updated as a weighted sum of Gaussian PDFs [1]. Moreover, the interacting multiple model (IMM) filter is a promising approach to address model uncertainty, in which several sub-filters are performed in parallel based on the preselected model set and the corresponding model transition probability matrix, and the sub-filters interact with each other by fusing the state estimates and corresponding estimation error covariance matrices based on the recursive estimates of the mode probabilities [6]. Normally, the performances of the PF, GSF and IMM filters rely heavily on the preselected distributions to model non-Gaussian state and measurement noises. Unfortunately, it is very difficult to select accurate distributions to model unknown and time-varying non-Gaussian noises that are often induced by outliers so that the estimation accuracy of the PF, GSF and IMM filter degrades significantly when inaccurate or even wrong noise distributions are used. The contribution of the work introduced in this chapter is therefore to provide a unified theoretic framework to solve the non-Gaussian filtering problem for a linear state-space model with unknown non-Gaussian heavy-tailed noises.

A large number of outlier-robust Kalman filters have been proposed to achieve a tradeoff between estimation accuracy and computational complexity. As a classic robust regression technique, the M-estimator is robust to measurement outliers and has been successfully extended to the Kalman filter setting [23]. By employing the influ-

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

ence function approach on the prediction error and residual error, many outlier-robust Kalman filters have been proposed based on the M-estimate method [13]. The Huber Kalman filter (HKF) is the most famous extension of the M-estimator to the Kalman filter setting, which utilizes a combined l_1 and l_2 norm as a robust cost function, and a generalized robust maximum likelihood estimate is achieved by minimizing the Huber cost function [26]. The maximum correntropy Kalman filter (MCKF) is an alternative method to handle state and measurement outliers. The sum of the Gaussian kernel functions of the prediction error and residual error are selected as the robust cost function, and many MCKFs have been proposed by maximizing such robust cost functions [8, 11, 39]. Motivated by the fact that the outlier contaminated state and measurement noises often have non-Gaussian heavy-tailed distributions, many outlier-robust filters have been proposed by modelling the state and measurement noises as Student's t distributed [16, 21, 30, 33, 37, 44]. These robust filters can be divided into two categories: Student's t filters and robust Student's t based Kalman filters (RSTKF). In the Student's t filter, the predicted state and observation are jointly modelled as Student's t-distributed, and then the posterior filtering PDF is recursively approximated by a Student's t-PDF with fixed degrees of freedom (dof) parameter based on the Bayesian rule and moment matching method [16, 33]. To address the approximate errors induced by moment matching method in the Student's t filter, a Kullback–Leilber divergence (KLD) minimization based adaptive Student's t filter has been proposed, where the scale matrices and the state are jointly estimated by minimizing the upper bound of the KLD [15]. Nevertheless, in the RSTKF, the one-step prediction PDF and the observation likelihood PDF are modelled by Student's t-PDFs and formulated as Gaussian–Gamma mixture forms, based on which the posterior filtering PDF is approximated by a Gaussian PDF with adaptively selected covariance matrix using the variational Bayesian (VB) approach [18, 21]. For a linear system with moderately heavy-tailed state and measurement noises, the RSTKF can achieve better estimation accuracy than the Student's t filter, but at the cost of higher computational complexity. Also, the adaptive Kalman filter based on the VB approach can to some extent address the heavy-tailed state and measurement noises induced by outliers based on the adaptive modifications of the one-step prediction error covariance matrix and measurement noise covariance matrix [20].

Although the HKF, MCKF and RSTKF can all achieve better estimation accuracy than the standard Kalman filter for a linear system with outlier-corrupted state and measurement noises, their interrelationships have as yet not been revealed. In addition, they still have the following theoretical problems:

– Essentially, the HKF and MCKF are both generalized maximum-likelihood estimators. In the HKF, random state vector is considered determinate, and the optimal point estimate of state vector is obtained by minimizing the Huber's cost function. In MCKF, the correntropy is approximated as a finite sum of Gaussian kernel

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

functions, which leads to the MCKF being a sort of point estimator [15]. It can be observed from these discussions that the uncertainty of the state vector has no effect on the cost function in the HKF and MCKF, which makes the estimation accuracy of HKF and MCKF limited.
– In the RSTKF, the Student's t distribution is used to model the non-Gaussian outlier-corrupted state and measurement noises. The estimation accuracy of the RSTKF depends heavily on the modeling accuracy of non-Gaussian noise by Student's t distribution. If the Student's t distribution is mismatched with the real noise distribution, the estimation accuracy of the RSTKF will degrade greatly.

An advanced outlier-robust Kalman filtering framework based on statistical similarity measure (SSMKF) is therefore proposed in [22] to reveal the interrelationships between the HKF, MCKF and RSTKF and further improve their estimation accuracy. The SSMKF framework can be thought of as improved versions of the HKF and MCKF and a generalized version of RSTKF. Also, the advantages of the framework are two-fold:
– The statistical similarity measure (SSM) is chosen as an objective function, by maximizing which, not only the posterior mean but also covariance matrix of the state vector, can be obtained. Since the randomness of the state vector is considered, the SSMKF achieves better estimation accuracy, as will be shown in Section 4.7.
– When the non-Gaussian noise cannot be well modelled by Student's t-distribution, the SSMKF can obtain better estimation accuracy by selecting the appropriate similarity function.

The SSMKF proposed in [22] is introduced in this chapter. The SSM is firstly presented to quantify the similarity between two random vectors and then used to develop the SSMKF framework, in which the lower bounds of the SSM between the state vector and the predicted state vector and that between the measurement vector and the predicted measurement vector are maximized, and the posterior PDF of the state is approximated as Gaussian. To illustrate the effectiveness of the SSMKF, the effects of the approximation errors on the SSMKF are analyzed in detail, and the numerical and filtering stabilities of the SSMKF are also discussed. The fixed-point iterative algorithm and the separate iterative algorithm are used to implement the SSMKF, and their local convergence conditions are also provided and compared in this chapter. Furthermore, the selections of the similarity functions are presented, and four exemplary similarity functions are provided. Then the relationships between the SSMKF framework and other existing outlier-robust Kalman filters are revealed. Simulation results of a manoeuvring target-tracking example illustrate that, by selecting appropriate similarity functions, the SSMKFs have improved estimation accuracy but higher computational complexities than other existing state-of-the-art filters.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.2 Statistical similarity measure

We introduce the SSM to quantify the similarity between two random vectors. The SSM $s(\mathbf{x}, \mathbf{y})$ for random vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as follows:

$$s(\mathbf{x}, \mathbf{y}) = \mathrm{E}[f(\|\mathbf{x} - \mathbf{y}\|^2)] = \int\int f(\|\mathbf{x} - \mathbf{y}\|^2) p(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y}, \tag{4.1}$$

where $\mathrm{E}[\cdot]$ denotes the expectation operation, and $p(\mathbf{x}, \mathbf{y})$ denotes the joint PDF of random vectors $\mathbf{x}$ and $\mathbf{y}$. The scalar function $f(\cdot)$ is named as the similarity function, and it must satisfy the following three conditions:

- Condition 1: $f(\cdot)$ is a continuous function defined on $[0, +\infty)$;
- Condition 2: $f(\cdot)$ is a strictly monotonically decreasing function: $\dot{f}(t) < 0$ for $t \in [0, +\infty)$;
- Condition 3: The second derivative of $f(\cdot)$ is non-negative: $\ddot{f}(t) \geq 0$ for $t \in [0, +\infty)$.

The monotonically decreasing property of the similarity function $f(\cdot)$ can guarantee that the SSM $s(\mathbf{x}, \mathbf{y})$ is increasing as the distance between $\mathbf{x}$ and $\mathbf{y}$ decreases. As a result, the SSM conforms to the usual definition of a similarity measure that is in some sense the inverse of distance metrics. The greater the similarity between random vectors $\mathbf{x}$ and $\mathbf{y}$, the larger SSM becomes. The SSM has some basic properties as follows:

- Property 1: The SSM is symmetric: $s(\mathbf{x}, \mathbf{y}) = s(\mathbf{y}, \mathbf{x})$;
- Property 2: The SSM achieves the maximum value if and only if the random vectors $\mathbf{x}$ and $\mathbf{y}$ are identical;
- Property 3: The SSM includes all the even-order moments of the random vector $\mathbf{x} - \mathbf{y}$: $s(\mathbf{x}, \mathbf{y}) = \sum_{l=0}^{+\infty} \frac{f^{(l)}(0)}{l!} \mathrm{E}[\|\mathbf{x} - \mathbf{y}\|^{2l}]$ if the Taylor expansion of the similarity function $f(t)$ exists when $t \geq 0$.

Property 1 can be easily verified using the definition of the SSM in (4.1), and Property 3 can be directly derived by exploiting the Taylor series expansion of the similarity function $f(t)$ at $t = 0$. Next, we will prove Property 2 in the following Theorem 4.2.1.

Theorem 4.2.1. *If the similarity function $f(\cdot)$ satisfies Condition 2, then Property 2 holds.*

Proof. Using $\dot{f}(t) < 0$ and $\|\mathbf{x} - \mathbf{y}\|^2 \geq 0$ yields

$$f(\|\mathbf{x} - \mathbf{y}\|^2) \leq f(0). \tag{4.2}$$

Substituting (4.2) in (4.1) gives

$$s(\mathbf{x}, \mathbf{y}) = \mathrm{E}[f(\|\mathbf{x} - \mathbf{y}\|^2)] \leq \mathrm{E}[f(0)] = f(0). \tag{4.3}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Considering that the inequality (4.3) holds for arbitrary random vectors $\mathbf{x}$ and $\mathbf{y}$, we have

$$\max s(\mathbf{x}, \mathbf{y}) = f(0). \tag{4.4}$$

It is evident that the SSM $s(\mathbf{x}, \mathbf{y})$ is identical to $f(0)$ when $\mathbf{x} = \mathbf{y}$, and then $\mathbf{x} = \mathbf{y}$ is a maximum point of the SSM $s(\mathbf{x}, \mathbf{y})$. □

Property 2 guarantees that the SSM has a maximum point, i. e., $\mathbf{x} = \mathbf{y}$. It is noteworthy that the SSM $s(\mathbf{x}, \mathbf{y})$ is an upper bound of $f(\mathrm{E}[\|\mathbf{x}-\mathbf{y}\|^2])$, and they have the same maximum point, i. e., $\mathbf{x} = \mathbf{y}$.

The SSM is a generalized similarity measure between two random vectors and encompasses existing similarity measures. For example, the SSM $s(\mathbf{x}, \mathbf{y})$ is the negative mean squared error between random vectors $\mathbf{x}$ and $\mathbf{y}$ when the similarity function is chosen as $f(t) = -t$. The SSM $s(\mathbf{x}, \mathbf{y})$ is the correntropy between random vectors $\mathbf{x}$ and $\mathbf{y}$ when the similarity function is selected as $f(t) = \exp(-\frac{1}{2\sigma^2}t)$ [11, 39]. More interestingly, the generalized correntropy [9] between random vectors $\mathbf{x}$ and $\mathbf{y}$ is also a special case of the SSM $s(\mathbf{x}, \mathbf{y})$ when the similarity function is chosen as $f(t) = \frac{\alpha}{2\beta\Gamma(1/\alpha)}\exp(-t^{\frac{\alpha}{2}}\beta^{-\alpha})$ and the shape parameter satisfies the constraint $0 < \alpha \leq 2$.

For the SSM, except for the first two conditions, the similarity function $f(\cdot)$ has to satisfy the third condition. The third condition not only can facilitate the design of an outlier-robust Kalman filter but also can guarantee robustness to outliers and local convergence of fixed point iterations, as will be shown in the following sections.

The SSM can be used in Bayesian inference. Different SSMs are achieved when different similarity functions $f(\cdot)$ are selected, based on which different state estimates can be obtained by maximizing the corresponding SSMs. Next, an outlier-robust Kalman filtering framework can be derived based on the SSM.

4.3 Outlier-robust Kalman filtering framework based on SSM

Consider a linear dynamical system described by a linear discrete-time state-space model as follows

$$\begin{cases} \mathbf{x}_k = \mathbf{F}_k\mathbf{x}_{k-1} + \mathbf{w}_k & \text{(state equation)}, \\ \mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k & \text{(measurement equation)}, \end{cases} \tag{4.5}$$

where k is the discrete time index, $\mathbf{x}_k \in \mathbb{R}^n$ is the state vector, $\mathbf{z}_k \in \mathbb{R}^m$ is the measurement vector, $\mathbf{F}_k \in \mathbb{R}^{n\times n}$ and $\mathbf{H}_k \in \mathbb{R}^{m\times n}$ are, respectively, the known state transition matrix and measurement matrix and $\mathbf{w}_k \in \mathbb{R}^n$ and $\mathbf{v}_k \in \mathbb{R}^m$ are, respectively, state and

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

measurement noise vectors. In this chapter, the state and measurement noises are assumed to have non-Gaussian distributions that are, respectively, induced by state and measurement outliers.

4.3.1 Design of outlier-robust Kalman filtering framework

Similar to the standard Kalman filter, the SSMKF is also composed of time and measurement updates. In the time update, the one-step predicted state vector $\hat{\mathbf{x}}_{k|k-1}$ and corresponding nominal prediction error covariance matrix $\mathbf{P}_{k|k-1}$ are calculated as follows

$$\begin{cases} \hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1}, \\ \mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^{\mathrm{T}} + \mathbf{Q}_k, \end{cases} \tag{4.6}$$

where $\hat{\mathbf{x}}_{k-1|k-1}$ and $\mathbf{P}_{k-1|k-1}$ are, respectively, the state estimate and corresponding estimation error covariance matrix at time $k-1$, and $\mathbf{Q}_k$ denotes the nominal state noise covariance matrix. Note that $\mathbf{P}_{k|k-1}$ is called the nominal prediction error covariance matrix since the used nominal state noise covariance matrix $\mathbf{Q}_k$ is inaccurate in the presence of state outliers.

The measurement update aims to achieve an approximate posterior PDF $q^*(\mathbf{x}_k) \approx p(\mathbf{x}_k|\mathrm{z}_{1:k})$ through maximizing the sum of the SSM between $\mathbf{S}_{k|k-1}^{-1}\mathbf{x}_k$ and $\mathbf{S}_{k|k-1}^{-1}\hat{\mathbf{x}}_{k|k-1}$ and the SSM between $\mathbf{S}_{\mathbf{R}_k}^{-1}\mathbf{z}_k$ and $\mathbf{S}_{\mathbf{R}_k}^{-1}\mathbf{H}_k\mathbf{x}_k$, i. e.,

$$q^*(\mathbf{x}_k) = \arg\max_{q(\mathbf{x}_k)} \{ s(\mathbf{S}_{k|k-1}^{-1}\mathbf{x}_k, \mathbf{S}_{k|k-1}^{-1}\hat{\mathbf{x}}_{k|k-1}) + s(\mathbf{S}_{\mathbf{R}_k}^{-1}\mathbf{z}_k, \mathbf{S}_{\mathbf{R}_k}^{-1}\mathbf{H}_k\mathbf{x}_k) \}, \tag{4.7}$$

where $\mathbf{S}_{k|k-1}$ and $\mathbf{S}_{\mathbf{R}_k}$ are, respectively, the square roots of the nominal predicted error covariance matrix $\mathbf{P}_{k|k-1}$ and the nominal measurement noise covariance matrix $\mathbf{R}_k$, i. e.,

$$\mathbf{P}_{k|k-1} = \mathbf{S}_{k|k-1}\mathbf{S}_{k|k-1}^{\mathrm{T}}, \quad \mathbf{R}_k = \mathbf{S}_{\mathbf{R}_k}\mathbf{S}_{\mathbf{R}_k}^{\mathrm{T}}, \tag{4.8}$$

where $\mathbf{R}_k$ denotes the nominal measurement noise covariance matrix. Note that the one-step statistical similarity measure based cost function in (4.7) is sufficient for designing an outlier-robust Kalman filtering framework, in which the one-step predicted state vector $\hat{\mathbf{x}}_{k|k-1}$ and measurement vector $\mathbf{z}_k$ at the current time are used to construct the cost function. Such an idea can be extended to derive an outlier-robust Kalman smoothing framework based on a multiple-steps SSM cost function by resorting to some standard techniques for designing a maximum a-posterior estimator in a Bayesian framework [34].

Considering that the one-step predicted state vector $\hat{\mathbf{x}}_{k|k-1}$ and measurement vector $\mathbf{z}_k$ are known and deterministic quantities in the measurement update of the

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kalman filter, the maximization problem in (4.7) can be reformulated as

$$q^*(\mathbf{x}_k) = \arg\max_{q(\mathbf{x}_k)} \Big\{ \int f_x(\|\mathbf{S}_{k|k-1}^{-1}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\|^2) q(\mathbf{x}_k) d\mathbf{x}_k + \int f_z(\|\mathbf{S}_{\mathbf{R}_k}^{-1}(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_k)\|^2) q(\mathbf{x}_k) d\mathbf{x}_k \Big\}, \tag{4.9}$$

where $f_x(\cdot)$ and $f_z(\cdot)$ denote the similarity functions of state and measurement models, respectively.

It is very difficult to achieve an optimal solution for the maximization problem in (4.9) since both the explicit form of posterior PDF $q(\mathbf{x}_k)$ and closed-form solutions for the integrals are unavailable. To solve this problem, the posterior PDF $q(\mathbf{x}_k)$ is approximated as Gaussian, and the lower bound of the cost function is maximized.

Firstly, the posterior PDF $q(\mathbf{x}_k)$ can be approximated as a Gaussian PDF and given by

$$q(\mathbf{x}_k) \approx \mathrm{N}(\mathbf{x}_k; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \tag{4.10}$$

where $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ are, respectively, the mean vector and covariance matrix of the posterior PDF $q(\mathbf{x}_k)$.

Substituting (4.10) into (4.9), the maximization problem with respect to the posterior PDF $q(\mathbf{x}_k)$ in (4.9) is approximately transformed into a maximization problem with respect to the posterior mean vector and covariance matrix, i. e.,

$$\{\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*\} \approx \arg\max_{\{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}} \Big\{ \int f_x(\|\mathbf{S}_{k|k-1}^{-1}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\|^2) \mathrm{N}(\mathbf{x}_k; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) d\mathbf{x}_k + \int f_z(\|\mathbf{S}_{\mathbf{R}_k}^{-1}(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_k)\|^2) \mathrm{N}(\mathbf{x}_k; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) d\mathbf{x}_k \Big\} \quad \text{s.t.} \quad \boldsymbol{\Sigma}_k > \mathbf{0}, \tag{4.11}$$

where $\boldsymbol{\mu}_k^*$ and $\boldsymbol{\Sigma}_k^*$ denote the optimal posterior mean vector and covariance matrix. Next, the cost function in (4.11) will be approximated as its lower bound, and then an approximate solution can be obtained.

Theorem 4.3.1. *If the similarity functions satisfy Condition 3, i. e., $\ddot{f}_x(t) \geq 0$ and $\ddot{f}_z(t) \geq 0$ for $t \in [0, +\infty)$, the maximization problem in* (4.11) *can be transformed as follows*

$$\{\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*\} \approx \arg\max_{\{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}} J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad \text{s.t.} \quad \boldsymbol{\Sigma}_k > \mathbf{0}, \tag{4.12}$$

where $J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is the lower bound of the cost function in (4.11) *and formulated as*

$$J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = f_x(\mathrm{tr}(\mathbf{A}_k\mathbf{P}_{k|k-1}^{-1})) + f_z(\mathrm{tr}(\mathbf{B}_k\mathbf{R}_k^{-1})), \tag{4.13}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

where $\mathrm{tr}(\cdot)$ *denotes the trace operation of a matrix, and* $\mathbf{A}_k$ *and* $\mathbf{B}_k$ *are, respectively given by*

$$\begin{aligned}\mathbf{A}_k &= \int (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^{\mathrm{T}} \mathrm{N}(\mathbf{x}_k; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) d\mathbf{x}_k \\ &= \boldsymbol{\Sigma}_k + (\boldsymbol{\mu}_k - \hat{\mathbf{x}}_{k|k-1})(\boldsymbol{\mu}_k - \hat{\mathbf{x}}_{k|k-1})^{\mathrm{T}}, \end{aligned} \tag{4.14}$$

$$\begin{aligned}\mathbf{B}_k &= \int (\mathbf{z}_k - \mathbf{H}_k \mathbf{x}_k)(\mathbf{z}_k - \mathbf{H}_k \mathbf{x}_k)^{\mathrm{T}} \mathrm{N}(\mathbf{x}_k; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) d\mathbf{x}_k \\ &= (\mathbf{z}_k - \mathbf{H}_k \boldsymbol{\mu}_k)(\mathbf{z}_k - \mathbf{H}_k \boldsymbol{\mu}_k)^{\mathrm{T}} + \mathbf{H}_k \boldsymbol{\Sigma}_k \mathbf{H}_k^{\mathrm{T}}. \end{aligned} \tag{4.15}$$

Proof. Since $\ddot{f}_x(t) \geq 0$ and $\ddot{f}_z(t) \geq 0$ for $t \in [0, +\infty)$, $f_x(\cdot)$ and $f_z(\cdot)$ are convex functions. Using Jensen's inequality, we have

$$\begin{aligned}&\int f_x(\left\|\mathbf{S}_{k|k-1}^{-1}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\right\|^2) \mathrm{N}(\mathbf{x}_k; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) d\mathbf{x}_k \\ &\quad \geq f_x\left(\int \left\|\mathbf{S}_{k|k-1}^{-1}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\right\|^2 \mathrm{N}(\mathbf{x}_k; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) d\mathbf{x}_k \right), \end{aligned} \tag{4.16}$$

$$\begin{aligned}&\int f_z(\left\|\mathbf{S}_{\mathbf{R}_k}^{-1}(\mathbf{z}_k - \mathbf{H}_k \mathbf{x}_k)\right\|^2) \mathrm{N}(\mathbf{x}_k; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) d\mathbf{x}_k \\ &\quad \geq f_z\left(\int \left\|\mathbf{S}_{\mathbf{R}_k}^{-1}(\mathbf{z}_k - \mathbf{H}_k \mathbf{x}_k)\right\|^2 \mathrm{N}(\mathbf{x}_k; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) d\mathbf{x}_k \right), \end{aligned} \tag{4.17}$$

where the equalities hold if and only if the similarity functions $f_x(\cdot)$ and $f_z(\cdot)$ are linear or the covariance matrix of $q(\mathbf{x}_k)$ is zero.

Substituting (4.16)–(4.17) into (4.11), the maximization problem with respect to the posterior mean vector and covariance matrix can be approximated as (4.12)–(4.15). □

Define four auxiliary variables ξ_k, λ_k, $\tilde{\xi}_k$ and $\tilde{\lambda}_k$ as follows

$$\begin{cases} \xi_k \triangleq -2\dot{f}_x(\mathrm{tr}(\mathbf{A}_k \mathbf{P}_{k|k-1}^{-1})), \quad \lambda_k \triangleq -2\dot{f}_z(\mathrm{tr}(\mathbf{B}_k \mathbf{R}_k^{-1})), \\ \tilde{\xi}_k \triangleq 2\ddot{f}_x(\mathrm{tr}(\mathbf{A}_k \mathbf{P}_{k|k-1}^{-1})), \quad \tilde{\lambda}_k \triangleq 2\ddot{f}_z(\mathrm{tr}(\mathbf{B}_k \mathbf{R}_k^{-1})) \end{cases} \tag{4.18}$$

and three auxiliary matrices as follows

$$\begin{cases} \boldsymbol{\Delta}_{\mu_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{\partial J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\partial \boldsymbol{\mu}_k}, \\ \boldsymbol{\Delta}_{\Sigma_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{\partial J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\partial \boldsymbol{\Sigma}_k}, \\ \boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{\partial^2 J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\partial \boldsymbol{\mu}_k \partial \boldsymbol{\mu}_k^{\mathrm{T}}}, \end{cases} \tag{4.19}$$

where $\boldsymbol{\Delta}_{\mu_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ and $\boldsymbol{\Delta}_{\Sigma_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ denote the Jacobian matrices of the approximate cost function $J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ with respect to the posterior mean vector and covariance matrix, respectively, and $\boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ denotes the Hessian matrix of the approximate cost function $J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ with respect to the posterior mean vector.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Theorem 4.3.2. *The optimal solution $\boldsymbol{\mu}_k^*$ of the approximate cost function $J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is formulated as follows:*

$$\boldsymbol{\mu}_k^* = \hat{\mathbf{x}}_{k|k-1} + \tilde{\mathbf{K}}_k^*(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1}), \tag{4.20}$$

$$\tilde{\mathbf{K}}_k^* = \tilde{\mathbf{P}}_{k|k-1}^*\mathbf{H}_k^{\mathrm{T}}(\mathbf{H}_k\tilde{\mathbf{P}}_{k|k-1}^*\mathbf{H}_k^{\mathrm{T}} + \tilde{\mathbf{R}}_k^*)^{-1}, \tag{4.21}$$

where $\tilde{\mathbf{P}}_{k|k-1}^*$ and $\tilde{\mathbf{R}}_k^*$ are, respectively, the modified one-step prediction error covariance matrix and measurement noise covariance matrix given by

$$\tilde{\mathbf{P}}_{k|k-1}^* = \mathbf{P}_{k|k-1}/\xi_k^*, \quad \tilde{\mathbf{R}}_k^* = \mathbf{R}_k/\lambda_k^*, \tag{4.22}$$

and the auxiliary parameters ξ_k^* and λ_k^* are given by

$$\xi_k^* = -2\dot{f}_x(\mathrm{tr}(\mathbf{A}_k^*\mathbf{P}_{k|k-1}^{-1})), \quad \lambda_k^* = -2\dot{f}_z(\mathrm{tr}(\mathbf{B}_k^*\mathbf{R}_k^{-1})), \tag{4.23}$$

and the auxiliary parameters $\mathbf{A}_k^*$ and $\mathbf{B}_k^*$ are given by

$$\mathbf{A}_k^* = \boldsymbol{\Sigma}_k^* + (\boldsymbol{\mu}_k^* - \hat{\mathbf{x}}_{k|k-1})(\boldsymbol{\mu}_k^* - \hat{\mathbf{x}}_{k|k-1})^{\mathrm{T}}, \tag{4.24}$$

$$\mathbf{B}_k^* = (\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k^*)(\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k^*)^{\mathrm{T}} + \mathbf{H}_k\boldsymbol{\Sigma}_k^*\mathbf{H}_k^{\mathrm{T}}. \tag{4.25}$$

Proof. Using (4.13)–(4.15) and (4.18), the Jacobian matrix $\boldsymbol{\Delta}_{\mu_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is calculated as

$$\boldsymbol{\Delta}_{\mu_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = -\xi_k\mathbf{P}_{k|k-1}^{-1}(\boldsymbol{\mu}_k - \hat{\mathbf{x}}_{k|k-1}) + \lambda_k\mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}(\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k). \tag{4.26}$$

According to the maximum criterion, the maximum point $\boldsymbol{\mu}_k^*$ satisfies the following equation

$$\boldsymbol{\Delta}_{\mu_k}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*) = \mathbf{0}. \tag{4.27}$$

Utilizing (4.23)–(4.25) and (4.26)–(4.27) yields

$$-\xi_k^*\mathbf{P}_{k|k-1}^{-1}(\boldsymbol{\mu}_k^* - \hat{\mathbf{x}}_{k|k-1}) + \lambda_k^*\mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}(\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k^*) = \mathbf{0}. \tag{4.28}$$

Solving equation (4.28), we can obtain the maximum point $\boldsymbol{\mu}_k^*$ as follows:

$$\boldsymbol{\mu}_k^* = (\tilde{\mathbf{P}}_{k|k-1}^{*-1} + \mathbf{H}_k^{\mathrm{T}}\tilde{\mathbf{R}}_k^{*-1}\mathbf{H}_k)^{-1}(\tilde{\mathbf{P}}_{k|k-1}^{*-1}\hat{\mathbf{x}}_{k|k-1} + \mathbf{H}_k^{\mathrm{T}}\tilde{\mathbf{R}}_k^{*-1}\mathbf{z}_k). \tag{4.29}$$

Substituting (4.22) into (4.29) and employing the matrix inversion lemma [35, pp. 11–12], we can obtain (4.20)–(4.21). □

Next, we will further confirm the extreme point $\boldsymbol{\mu}_k^*$ in (4.20) is a maximum point or a minimum point, and present the monotonicity of the approximate cost function $J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ with respect to the posterior covariance matrix $\boldsymbol{\Sigma}_k$.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Theorem 4.3.3. *If the similarity functions satisfy Condition 2 and Condition 3 and the following inequalities hold:*

$$-\left(\xi_k^*\right)^2 + 2\tilde{\xi}_k^* < 0, \quad -\left(\lambda_k^*\right)^2 + 2\tilde{\lambda}_k^* < 0, \tag{4.30}$$

then both the Hessian matrix $\boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$ *and the Jacobian matrix* $\boldsymbol{\Delta}_{\Sigma_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ *are negative definite, i. e.,*

$$\boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*) < \mathbf{0}, \quad \boldsymbol{\Delta}_{\Sigma_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) < \mathbf{0}. \tag{4.31}$$

Proof. Using (4.13)–(4.15) and (4.18), the Jacobian matrix $\boldsymbol{\Delta}_{\Sigma_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ and the Hessian matrix $\boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ can be calculated as

$$\boldsymbol{\Delta}_{\Sigma_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = -0.5\xi_k \mathbf{P}_{k|k-1}^{-1} - 0.5\lambda_k \mathbf{H}_k^{\mathrm{T}} \mathbf{R}_k^{-1} \mathbf{H}_k, \tag{4.32}$$

$$\boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = -\xi_k \mathbf{P}_{k|k-1}^{-1} - \lambda_k \mathbf{H}_k^{\mathrm{T}} \mathbf{R}_k^{-1} \mathbf{H}_k + \mathbf{D}_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \tag{4.33}$$

where $\mathbf{D}_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ denotes the second-order term given by

$$\begin{aligned} \mathbf{D}_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = {} & 2\tilde{\xi}_k \mathbf{P}_{k|k-1}^{-1} (\boldsymbol{\mu}_k - \hat{\mathbf{x}}_{k|k-1})(\boldsymbol{\mu}_k - \hat{\mathbf{x}}_{k|k-1})^{\mathrm{T}} \mathbf{P}_{k|k-1}^{-1} \\ & + 2\tilde{\lambda}_k \mathbf{H}_k^{\mathrm{T}} \mathbf{R}_k^{-1} (\mathbf{z}_k - \mathbf{H}_k \boldsymbol{\mu}_k)(\mathbf{z}_k - \mathbf{H}_k \boldsymbol{\mu}_k)^{\mathrm{T}} \mathbf{R}_k^{-1} \mathbf{H}_k. \end{aligned} \tag{4.34}$$

Utilizing $\xi_k > 0$ and $\lambda_k > 0$ in (4.32), we can obtain $\boldsymbol{\Delta}_{\Sigma_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) < \mathbf{0}$, and then the maximum point $\boldsymbol{\Sigma}_k^*$ can be given by (4.44). Substituting (4.22)–(4.25) into (4.33) yields

$$\boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*) = -\xi_k^* \mathbf{P}_{k|k-1}^{-1} - \lambda_k^* \mathbf{H}_k^{\mathrm{T}} \mathbf{R}_k^{-1} \mathbf{H}_k + \mathbf{D}_1(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*). \tag{4.35}$$

Employing (4.20)–(4.21) in (4.34), $\mathbf{D}_1(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$ is calculated as

$$\begin{aligned} \mathbf{D}_1(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*) = {} & 2\tilde{\xi}_k^* \mathbf{P}_{k|k-1}^{-1} \tilde{\mathbf{K}}_k^* \bar{\mathbf{P}}_{k|k-1}^{zz} (\tilde{\mathbf{K}}_k^*)^{\mathrm{T}} \mathbf{P}_{k|k-1}^{-1} \\ & + 2\tilde{\lambda}_k^* \mathbf{H}_k^{\mathrm{T}} \mathbf{R}_k^{-1} (\mathbf{I}_m - \mathbf{H}_k \tilde{\mathbf{K}}_k^*) \bar{\mathbf{P}}_{k|k-1}^{zz} (\mathbf{I}_m - \mathbf{H}_k \tilde{\mathbf{K}}_k^*)^{\mathrm{T}} \mathbf{R}_k^{-1} \mathbf{H}_k, \end{aligned} \tag{4.36}$$

where $\bar{\mathbf{P}}_{k|k-1}^{zz}$ denotes an approximate innovation matrix given by

$$\bar{\mathbf{P}}_{k|k-1}^{zz} = (\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1})(\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1})^{\mathrm{T}}. \tag{4.37}$$

It can be seen from (4.36) that the second-order term $\mathbf{D}_1(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$ depends on the real-time measurement $\mathbf{z}_k$. As a result, it is very difficult to compare the term $-\xi_k^* \mathbf{P}_{k|k-1}^{-1} - \lambda_k^* \mathbf{H}_k^{\mathrm{T}} \mathbf{R}_k^{-1} \mathbf{H}_k$ and the second-order term $\mathbf{D}_1(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$. To solve this problem, a reasonable approximation is introduced as follows:

$$\bar{\mathbf{P}}_{k|k-1}^{zz} \approx \tilde{\mathbf{P}}_{k|k-1}^{zz*} = \mathbf{H}_k \tilde{\mathbf{P}}_{k|k-1}^* \mathbf{H}_k^{\mathrm{T}} + \tilde{\mathbf{R}}_k^*, \tag{4.38}$$

where $\tilde{\mathbf{P}}_{k|k-1}^{zz*}$ denotes the modified innovation matrix.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Exploiting (4.21), (4.44) and (4.38) yields

$$\begin{cases} \tilde{\mathbf{K}}_k^* \tilde{\mathbf{P}}_{k|k-1}^{zz} (\tilde{\mathbf{K}}_k^*)^{\mathrm{T}} \approx \tilde{\mathbf{P}}_{k|k-1}^* - \boldsymbol{\Sigma}_k^* < \tilde{\mathbf{P}}_{k|k-1}^*, \\ (\mathbf{I}_m - \mathbf{H}_k \tilde{\mathbf{K}}_k^*) \tilde{\mathbf{P}}_{k|k-1}^{zz} (\mathbf{I}_m - \mathbf{H}_k \tilde{\mathbf{K}}_k^*)^{\mathrm{T}} \approx \tilde{\mathbf{R}}_k^* (\mathbf{H}_k \tilde{\mathbf{P}}_{k|k-1}^* \mathbf{H}_k^{\mathrm{T}} + \tilde{\mathbf{R}}_k^*)^{-1} \tilde{\mathbf{R}}_k^* < \tilde{\mathbf{R}}_k^*. \end{cases} \tag{4.39}$$

Substituting (4.36) into (4.35) and using (4.22) and (4.39) yields

$$\boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*) = (-\xi_k^* + 2\tilde{\xi}_k^* / \xi_k^*) \mathbf{P}_{k|k-1}^{-1} - (\lambda_k^* - 2\tilde{\lambda}_k^* / \lambda_k^*) \mathbf{H}_k^{\mathrm{T}} \mathbf{R}_k^{-1} \mathbf{H}_k. \tag{4.40}$$

Employing (4.30) in (4.40), we can obtain $\boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*) < \mathbf{0}$. □

Theorem 4.3.3 implies that the extreme point $\boldsymbol{\mu}_k^*$ in (4.20) is a maximum point of the approximate cost function $J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$. It can be also observed from Theorem 4.3.3 that the approximate cost function $J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is monotonically decreasing with respect to the posterior covariance matrix $\boldsymbol{\Sigma}_k$, and then the approximate cost function $J_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ can achieve a unique optimal solution $\boldsymbol{\Sigma}_k^*$ at the lower bound of the posterior covariance matrix. Next, we will determine the optimal posterior covariance matrix $\boldsymbol{\Sigma}_k^*$.

To obtain the maximum point $\boldsymbol{\Sigma}_k^*$, we need to find a reasonable constraint to apply upon $\boldsymbol{\Sigma}_k$. Motivated by the fact that the covariance matrix of the posterior PDF is the negative inverse of the Hessian matrix of the least squares cost function in the traditional maximum a posteriori estimation framework, we introduce a heuristic assumption that $\boldsymbol{\Sigma}_k$ is not less than the negative inverse of the Hessian matrix $\boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$, i. e.,

$$\boldsymbol{\Sigma}_k \geq -\boldsymbol{\Theta}_{\mu_k}^{-1}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*). \tag{4.41}$$

It is worth noting that the well-known M-estimate employs a similar way to deal with the posterior covariance matrix, which is set as the negative inverse of the Hessian matrix of the robust cost function [23, 26]. Since the cost function $J(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is monotonically decreasing with respect to $\boldsymbol{\Sigma}_k$, the optimal covariance matrix should be the negative inverse of the Hessian matrix, i. e., $-\boldsymbol{\Theta}_{\mu_k}^{-1}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$. Unfortunately, the Hessian matrix can easily lose its negative definiteness during the iterative computation, and then the filter was often found to halt its operation due to the numerical problem. To keep the negative definiteness, the positive semi-definite term $\mathbf{D}_1(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$ is mandatorily subtracted from the Hessian matrix $\boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$. That is to say, the lower bound of the posterior covariance matrix is further reduced, i. e.,

$$\boldsymbol{\Sigma}_k \geq -\boldsymbol{\Theta}_{\mu_k}^{-1}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*) \geq -\left[\boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*) - \mathbf{D}_1(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)\right]^{-1}. \tag{4.42}$$

Exploiting (4.42) and (4.33), the maximum point $\boldsymbol{\Sigma}_k^*$ can be formulated as

$$\boldsymbol{\Sigma}_k^* = \left(\xi_k^* \mathbf{P}_{k|k-1}^{-1} + \lambda_k^* \mathbf{H}_k^{\mathrm{T}} \mathbf{R}_k^{-1} \mathbf{H}_k\right)^{-1}. \tag{4.43}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Employing the matrix inversion lemma [35, pp. 11–12] and equations (4.21)–(4.22) and (4.43) gives

$$\boldsymbol{\Sigma}_k^* = \tilde{\mathbf{P}}_{k|k-1}^* - \tilde{\mathbf{K}}_k^* \mathbf{H}_k \tilde{\mathbf{P}}_{k|k-1}^*. \tag{4.44}$$

4.3.2 Error analyses of the SSMKF

In Section 4.3.1, three approximations are employed to derive an analytical solution for the original maximization problem in (4.9), which are listed as follows:

- Approximation 1: The posterior PDF $q(\mathbf{x}_k)$ is approximated as a Gaussian PDF in (4.10);
- Approximation 2: The original cost function is approximated as its lower bound by Theorem 4.3.1;
- Approximation 3: The original Hessian matrix is approximated as $\boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*) - \mathbf{D}_1(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$.

First, we discuss the reasonability of Approximation 1. The outlier-contaminated state and measurement noises often have non-Gaussian distributions, and then the true posterior PDF $p(\mathbf{x}_k|\mathrm{z}_{1:k})$ also has a non-Gaussian distribution [31, 33]. Unfortunately, there is often not a mathematical formulation for a general non-Gaussian distribution. As a result, it is not possible to achieve an optimal solution of the maximization problem in (4.9) for a general non-Gaussian linear system. Motivated by the fact that the Gaussian approximation to the posterior PDF has been widely accepted in designing a cost-effective filter for a linear system with non-Gaussian noises, the posterior PDF is also approximated as a Gaussian PDF in the SSMKF, as shown in (4.10), based on which an approximately analytical solution can be obtained. Although such an approximation may impose an error on the posterior PDF, it often exhibits good estimation accuracy with reasonable computational complexity. Thus, the Gaussian approximation to the posterior PDF can provide a trade-off between estimation accuracy and computational complexity.

Second, we analyze the effects of Approximation 2 on the optimal solution. To this end, the relationship between the maximization problem in (4.11) and the maximization problem in (4.12)–(4.15) will be further revealed. Define four auxiliary variables as follows:

$$\begin{cases} Y_{1k} = \|\mathbf{S}_{k|k-1}^{-1}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\|^2, \quad Y_{1k}^* = \mathrm{tr}(\mathbf{A}_k^* \mathbf{P}_{k|k-1}^{-1}), \\ Y_{2k} = \|\mathbf{S}_{\mathbf{R}_k}^{-1}(\mathbf{z}_k - \mathbf{H}_k \mathbf{x}_k)\|^2, \quad Y_{2k}^* = \mathrm{tr}(\mathbf{B}_k^* \mathbf{R}_k^{-1}), \end{cases} \tag{4.45}$$

where Y_{1k}^* and Y_{2k}^* are, respectively, the expectations of Y_{1k} and Y_{2k} with respect to the approximate posterior PDF $q^*(\mathbf{x}_k) = \mathrm{N}(\mathbf{x}_k; \boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Taking the first-order Taylor series expansions of the similarity functions $f_x(t)$ and $f_z(t)$ at $t = Y_{1k}^*$ and $t = Y_{2k}^*$, respectively, we have

$$\begin{cases} f_x(t) = f_x(Y_{1k}^*) + \dot{f}_x(Y_{1k}^*)(t - Y_{1k}^*) + \mathrm{o}(t - Y_{1k}^*), \\ f_z(t) = f_z(Y_{2k}^*) + \dot{f}_z(Y_{2k}^*)(t - Y_{2k}^*) + \mathrm{o}(t - Y_{2k}^*), \end{cases} \tag{4.46}$$

where $\mathrm{o}(t - Y_{1k}^*)$ and $\mathrm{o}(t - Y_{2k}^*)$ denote the first-order remainder terms of the similarity functions $f_x(t)$ and $f_z(t)$ at $t = Y_{1k}^*$ and $t = Y_{2k}^*$, respectively.

Using (4.46) yields

$$\begin{cases} f_x(Y_{1k}) \approx f_x(Y_{1k}^*) + \dot{f}_x(Y_{1k}^*)(Y_{1k} - Y_{1k}^*), \\ f_z(Y_{2k}) \approx f_z(Y_{2k}^*) + \dot{f}_z(Y_{2k}^*)(Y_{2k} - Y_{2k}^*). \end{cases} \tag{4.47}$$

Proposition 4.3.4. *The maximization problem in Theorem* 4.3.1 *and the maximization problem in* (4.11) *with the first-order Taylor approximations* (4.45) *and* (4.47) *have the same optimal solution* $\{\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*\}$.

Proof. Using (4.47), the cost function in (4.11) can be approximated as

$$\tilde{J}_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \int [\dot{f}_x(Y_{1k}^*)Y_{1k} + \dot{f}_z(Y_{2k}^*)Y_{2k}]\mathrm{N}(\mathbf{x}_k; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)d\mathbf{x}_k + c_{\{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}}. \tag{4.48}$$

Substituting (4.45) into (4.48) and using (4.14)–(4.15), (4.18) and (4.23) yields

$$\tilde{J}_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = -0.5\xi_k^* \operatorname{tr}(\mathbf{A}_k\mathbf{P}_{k|k-1}^{-1}) - 0.5\lambda_k^* \operatorname{tr}(\mathbf{B}_k\mathbf{R}_k^{-1}) + c_{\{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}}. \tag{4.49}$$

Employing (4.49), the Jacobian matrices of $\tilde{J}_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ with respect to $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ and the Hessian matrix of $\tilde{J}_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ with respect to $\boldsymbol{\mu}_k$ are, respectively, formulated as:

$$\begin{cases} \frac{\partial \tilde{J}_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\partial \boldsymbol{\mu}_k} = -\xi_k^* \mathbf{P}_{k|k-1}^{-1}(\boldsymbol{\mu}_k - \hat{\mathbf{x}}_{k|k-1}) + \lambda_k^* \mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}(\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k), \\ \frac{\partial \tilde{J}_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\partial \boldsymbol{\Sigma}_k} = -0.5\xi_k^* \mathbf{P}_{k|k-1}^{-1} - 0.5\lambda_k^* \mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}\mathbf{H}_k, \\ \frac{\partial \tilde{J}_1^2(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\partial \boldsymbol{\mu}_k \partial \boldsymbol{\mu}_k^{\mathrm{T}}} = -\xi_k^* \mathbf{P}_{k|k-1}^{-1} - \lambda_k^* \mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}\mathbf{H}_k. \end{cases} \tag{4.50}$$

According to the maximum criterion and utilizing (4.50), the maximum points $\boldsymbol{\mu}_k^*$ and $\boldsymbol{\Sigma}_k^*$ can be formulated as (4.20) and (4.44). Thus, the maximization problem in Theorem 4.3.1 and the maximization problem in (4.11) with the first-order Taylor approximations (4.45) and (4.47) have the same optimal solution. □

Proposition 4.3.4 means that the lower bound of the original cost function and the first-order approximation of the original cost function have the same maximum solution. Thus, the effects of Approximation 2 on the optimal solution are determined by the higher-order approximation errors $\mathrm{o}(Y_{1k} - Y_{1k}^*)$ and $\mathrm{o}(Y_{2k} - Y_{2k}^*)$. Since Y_{1k}^* and Y_{2k}^* are the mean values of the random variables Y_{1k} and Y_{2k}, the difference values

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

$Y_{1k} - Y_{1k}^*$ and $Y_{2k} - Y_{2k}^*$ depend heavily on the variances of the random variables Y_{1k} and Y_{2k}. That is to say, the higher are the variances of the random variables Y_{1k} and Y_{2k}, the larger are the fluctuation ranges of the difference values $Y_{1k} - Y_{1k}^*$ and $Y_{2k} - Y_{2k}^*$ that will be generated and then result in larger higher-order approximation errors, and vice versa. Thus, the approximation accuracy of Theorem 4.3.1 is mainly dominated by the variances of the random variables Y_{1k} and Y_{2k}. Therefore, next, we will study the variances of the random variables Y_{1k} and Y_{2k}.

Proposition 4.3.5. *The variances of the random variables Y_{1k} and Y_{2k} satisfy the following upper bound constraints*:

$$\begin{cases} \mathrm{Var}[Y_{1k}] \le (n^2+2n-1)[\mathrm{tr}(\boldsymbol{\Sigma}_k\mathbf{P}_{k|k-1}^{-1})]^2 + \|2\boldsymbol{\Sigma}_k^{\frac{\mathrm{T}}{2}}\mathbf{P}_{k|k-1}^{-1}(\boldsymbol{\mu}_k - \hat{\mathbf{x}}_{k|k-1})\|^2, \\ \mathrm{Var}[Y_{2k}] \le (n^2+2n-1)[\mathrm{tr}(\mathbf{H}_k\boldsymbol{\Sigma}_k\mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1})]^2 + \|2\boldsymbol{\Sigma}_k^{\frac{\mathrm{T}}{2}}\mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}(\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k)\|^2, \end{cases} \tag{4.51}$$

where $\mathrm{Var}[\cdot]$ *denotes the variance operation.*

Proof. Using (4.45) yields

$$\begin{cases} Y_{1k} = \|\mathbf{S}_{k|k-1}^{-1}\boldsymbol{\Sigma}_k^{\frac{1}{2}}\boldsymbol{\tau}_x\|^2 + 2\mathbf{a}_x^{\mathrm{T}}\mathbf{S}_{k|k-1}^{-1}\boldsymbol{\Sigma}_k^{\frac{1}{2}}\boldsymbol{\tau}_x + \|\mathbf{a}_x\|^2, \\ \boldsymbol{\tau}_x = \boldsymbol{\Sigma}_k^{-\frac{1}{2}}(\mathbf{x}_k - \boldsymbol{\mu}_k), \quad \mathbf{a}_x = \mathbf{S}_{k|k-1}^{-1}(\boldsymbol{\mu}_k - \hat{\mathbf{x}}_{k|k-1}). \end{cases} \tag{4.52}$$

Since the posterior mean vector and covariance matrix of the Gaussian distributed random vector $\mathbf{x}_k$ are, respectively, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$, the random vector $\boldsymbol{\tau}_x$ has a standard Gaussian distribution, i. e., $\boldsymbol{\tau}_x \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_n)$. Employing (4.52) and $\boldsymbol{\tau}_x \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_n)$, the variance of Y_{1k} is calculated as

$$\mathrm{Var}[Y_{1k}] = \mathrm{E}[\|\mathbf{S}_{k|k-1}^{-1}\boldsymbol{\Sigma}_k^{\frac{1}{2}}\boldsymbol{\tau}_x\|^4] - [\mathrm{tr}(\boldsymbol{\Sigma}_k\mathbf{P}_{k|k-1}^{-1})]^2 + \|2\boldsymbol{\Sigma}_k^{\frac{\mathrm{T}}{2}}\mathbf{P}_{k|k-1}^{-1}(\boldsymbol{\mu}_k - \hat{\mathbf{x}}_{k|k-1})\|^2, \tag{4.53}$$

in which note that the cross variance between $\|\mathbf{S}_{k|k-1}^{-1}\boldsymbol{\Sigma}_k^{\frac{1}{2}}\boldsymbol{\tau}_x\|^2$ and $2\mathbf{a}_x^{\mathrm{T}}\mathbf{S}_{k|k-1}^{-1}\boldsymbol{\Sigma}_k^{\frac{1}{2}}\boldsymbol{\tau}_x$ is zero since the odd origin moments of $\boldsymbol{\tau}_x$ are all zeros. According to the compatibility of matrix and vector norms, we have

$$\mathrm{E}[\|\mathbf{S}_{k|k-1}^{-1}\boldsymbol{\Sigma}_k^{\frac{1}{2}}\boldsymbol{\tau}_x\|^4] \le \mathrm{E}[\|\boldsymbol{\tau}_x\|^4][\mathrm{tr}(\boldsymbol{\Sigma}_k\mathbf{P}_{k|k-1}^{-1})]^2. \tag{4.54}$$

Considering that the random vector $\boldsymbol{\tau}_x$ has a standard Gaussian distribution, then the random variable $\|\boldsymbol{\tau}_x\|^2$ is a sum of the squares of n independent Gaussian random variables and has a chi-square distribution with the dof parameter n, i. e., $\|\boldsymbol{\tau}_x\|^2 \sim \chi^2(n)$. According the property of the chi-square distribution, the second- order origin moment of $\|\boldsymbol{\tau}_x\|^2$ can be calculated as $\mathrm{E}[\|\boldsymbol{\tau}_x\|^4] = n^2 + 2n$. Employing $\mathrm{E}[\|\boldsymbol{\tau}_x\|^4] = n^2 + 2n$ and (4.53)–(4.54), we can obtain the upper-bound constraint of $\mathrm{Var}[Y_{1k}]$ in (4.51). Similarly, we can also derive the upper-bound constraint of $\mathrm{Var}[Y_{2k}]$ in (4.51). □

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

It is observed from Proposition 4.3.5 that both the variances of the random variables Y_{1k} and Y_{2k} depend on the state dimension n and the posterior covariance matrix $\mathbf{\Sigma}_k$. It can be observed from (4.51) that the higher the state dimension, the larger will be the variances of the random variables Y_{1k} and Y_{2k}. Such a result is consistent with the intuition that the higher the state dimension, the larger the errors will be when the original cost function is approximated by its lower bound in Theorem 4.3.1. We can also observe from (4.51) that the larger the posterior covariance matrix $\mathbf{\Sigma}_k$, the larger will be the variances of the random variables Y_{1k} and Y_{2k}. This result is also consistent with the fact that the approximation errors are mainly dominated by the randomness of state vector $\mathbf{x}_k$ (i. e., the covariance matrix) when the original cost function is approximated by its lower bound based on Jensen's inequality. In conclusion, the effects of Approximation 2 on the optimal solution relies mainly on the state dimension and the posterior covariance matrix of the state vector. Fortunately, the posterior uncertainty is gradually reduced as the filter converges, which can to some extent mitigate the effects of Approximation 2 on the optimal solution. More importantly, both the errors $Y_{1k} - Y_{1k}^*$ and $Y_{2k} - Y_{2k}^*$ can be deemed as small terms as the filter converges, and the higher-order derivatives of the similarity functions $f_x(t)$ and $f_z(t)$ are significantly smaller than the first-order derivatives for the exemplary similarity functions in Table 4.1. As a result, the higher-order approximation errors $\mathrm{o}(Y_{1k} - Y_{1k}^*)$ and $\mathrm{o}(Y_{2k} - Y_{2k}^*)$ are significantly smaller than the first-order terms for the exemplary similarity functions, which contributes to the effectiveness and reasonability of Approximation 2.

Finally, we discuss the effects of Approximation 3 on the optimal solution. The original Hessian matrix is forcibly reduced by subtracting a positive semi-definite matrix, and the resultant posterior covariance matrix is less than the optimal value. Although such approximation imposes an error on the posterior covariance matrix, it is beneficial to guarantee the positive definiteness of the posterior covariance matrix and then improve the numerical and filtering stabilities, as will be shown in next section. Moreover, the reduction of the posterior covariance matrix is also beneficial to mitigate the effects of Approximation 2 on the optimal solution.

4.3.3 Stability discussions of the SSMKF

For the application of the SSMKF in practical engineering, it is necessary to guarantee the numerical and filtering stabilities of the SSMKF. These will be discussed next.

It is seen from (4.23) that, if the similarity functions $f_x(\cdot)$ and $f_z(\cdot)$ satisfy Condition 2, i. e., $\dot{f}_x(t) < 0$ and $\dot{f}_z(t) < 0$ for $t \in [0, +\infty)$, then both the auxiliary parameters ξ_k^* and λ_k^* are greater than zero, i. e.,

$$\xi_k^* > 0, \quad \lambda_k^* > 0. \tag{4.55}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Using (4.22), (4.43) and (4.55) yields

$$\tilde{\mathbf{P}}_{k|k-1}^{*} > \mathbf{0}, \quad \tilde{\mathbf{R}}_{k}^{*} > \mathbf{0}, \quad \mathbf{\Sigma}_{k}^{*} > \mathbf{0}. \tag{4.56}$$

We can see from (4.56) that, if $\dot{f}_x(t) < 0$ and $\dot{f}_z(t) < 0$ for $t \in [0, +\infty)$, then the modified one-step prediction error and measurement noise covariance matrices and the posterior covariance matrix are all positive definite. Thus, the SSMKF is numerically stable if the similarity function similarity functions $f_x(\cdot)$ and $f_z(\cdot)$ satisfy Condition 2.

In this chapter, the filtering stability means that the state estimation error $\tilde{\mathbf{x}}_{k|k}$ is bounded in the sense of the mean square, i. e., $\mathrm{E}\{\|\tilde{\mathbf{x}}_{k|k}\|^2\} < +\infty$ [41]. According to the theoretical analysis of the filtering stability in [41], if the modified one-step prediction error and measurement noise covariance matrices and the posterior covariance matrix satisfy the following constraints

$$\begin{cases} \tilde{\mathbf{P}}_{k|k-1}^{*} \le q_{\max}\mathbf{I}_n, \quad \tilde{\mathbf{R}}_{k}^{*} \ge r_{\min}\mathbf{I}_m, \\ p_{\min}\mathbf{I}_n \le \mathbf{\Sigma}_{k}^{*} \le p_{\max}\mathbf{I}_n, \end{cases} \tag{4.57}$$

where $q_{\max}$, $r_{\min}$, $p_{\min}$ and $p_{\max}$ are all positive real numbers, then the state estimation error $\tilde{\mathbf{x}}_{k|k}$ of the SSMKF is bounded in the sense of the mean square, i. e., $\mathrm{E}\{\|\tilde{\mathbf{x}}_{k|k}\|^2\} < +\infty$. Note that the other conditions for guaranteeing the filtering stability in [41] hold naturally for linear systems.

Proposition 4.3.6. *If there exists positive real numbers $\xi_{\min}$, $\xi_{\max}$, $\lambda_{\min}$ and $\lambda_{\max}$ such that the following constraints are fulfilled*

$$\xi_{\min} \le \xi_{k}^{*} \le \xi_{\max}, \quad \lambda_{\min} \le \lambda_{k}^{*} \le \lambda_{\max}, \tag{4.58}$$

then the inequalities in (4.57) *hold.*

Proof. Using (4.58) in (4.22) and (4.43) yields

$$\tilde{\mathbf{P}}_{k|k-1}^{*} \le \mathbf{P}_{k|k-1}/\xi_{\min}, \quad \tilde{\mathbf{R}}_{k}^{*} \ge \mathbf{R}_{k}/\lambda_{\max}, \tag{4.59}$$

$$\begin{cases} \mathbf{\Sigma}_{k}^{*} \ge (\xi_{\max}\mathbf{P}_{k|k-1}^{-1} + \lambda_{\max}\mathbf{H}_{k}^{\mathrm{T}}\mathbf{R}_{k}^{-1}\mathbf{H}_{k})^{-1}, \\ \mathbf{\Sigma}_{k}^{*} \le (\xi_{\min}\mathbf{P}_{k|k-1}^{-1} + \lambda_{\min}\mathbf{H}_{k}^{\mathrm{T}}\mathbf{R}_{k}^{-1}\mathbf{H}_{k})^{-1}. \end{cases} \tag{4.60}$$

Choosing $q_{\max}$ and $p_{\max}$ as the maximum eigenvalues of $\mathbf{P}_{k|k-1}/\xi_{\min}$ and $(\xi_{\min}\mathbf{P}_{k|k-1}^{-1} + \lambda_{\min}\mathbf{H}_{k}^{\mathrm{T}}\mathbf{R}_{k}^{-1}\mathbf{H}_{k})^{-1}$ and selecting $r_{\min}$ and $p_{\min}$ as the minimum eigenvalues of $\mathbf{R}_k/\lambda_{\max}$ and $(\xi_{\max}\mathbf{P}_{k|k-1}^{-1} + \lambda_{\max}\mathbf{H}_{k}^{\mathrm{T}}\mathbf{R}_{k}^{-1}\mathbf{H}_{k})^{-1}$, we can obtain (4.57). □

Proposition 4.3.6 demonstrates that, if both the auxiliary parameters ξ_k^* and λ_k^* have lower and upper bounds, then the SSMKF is always stable, which will impose additional constraints on the similarity functions $f_x(\cdot)$ and $f_z(\cdot)$.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.4 Novel iterative algorithms for the SSMKF

4.4.1 Fixed-point iterative algorithm

In general, it is very difficult to find analytical solutions for $\boldsymbol{\mu}_k^*$ and $\boldsymbol{\Sigma}_k^*$ through solving equations (4.20)–(4.25) and (4.44) directly when the similarity functions $f_x(\cdot)$ and $f_z(\cdot)$ are nonlinear functions. To solve this problem, we employ fixed-point iterations to solve these equations approximately. The detailed implementation of the SSMKF based on fixed-point iterations is listed in Algorithm 4.1, where ϵ denotes the iteration threshold, N_m denotes the maximum number of iterations and i^* denotes the cyclic variable at the end of the loop, and δ denotes the lower bounds of the auxiliary parameters ξ_k and λ_k that is beneficial to guarantee the filtering stability, as will be discussed in Section 4.5.

Algorithm 4.1: One time step of the SSMKF based on fixed-point iterations.

Inputs: $\hat{\mathbf{x}}_{k-1|k-1}$, $\mathbf{P}_{k-1|k-1}$, $\mathbf{F}_k$, $\mathbf{H}_k$, $\mathbf{z}_k$, $\mathbf{Q}_k$, $\mathbf{R}_k$, $f_x(\cdot)$, $f_z(\cdot)$, ϵ, N_m, δ.

Time update:

1. $\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k\hat{\mathbf{x}}_{k-1|k-1}$
2. $\mathbf{P}_{k|k-1} = \mathbf{F}_k\mathbf{P}_{k-1|k-1}\mathbf{F}_k^T + \mathbf{Q}_k$

Iterative measurement update:

3. Initialization: $\xi_k^{(0)} = 1$, $\lambda_k^{(0)} = 1$

for $i = 0 : N_m - 1$

Calculate $\tilde{\mathbf{P}}_{k|k-1}^{(i+1)}$ and $\tilde{\mathbf{R}}_k^{(i+1)}$

4. $\tilde{\mathbf{P}}_{k|k-1}^{(i+1)} = \mathbf{P}_{k|k-1}/\xi_k^{(i)}$, $\tilde{\mathbf{R}}_k^{(i+1)} = \mathbf{R}_k/\lambda_k^{(i)}$

Evaluate $\boldsymbol{\mu}_k^{(i+1)}$ and $\boldsymbol{\Sigma}_k^{(i+1)}$

5. $\tilde{\mathbf{K}}_k^{(i+1)} = \tilde{\mathbf{P}}_{k|k-1}^{(i+1)}\mathbf{H}_k^T(\mathbf{H}_k\tilde{\mathbf{P}}_{k|k-1}^{(i+1)}\mathbf{H}_k^T + \tilde{\mathbf{R}}_k^{(i+1)})^{-1}$
6. $\boldsymbol{\mu}_k^{(i+1)} = \hat{\mathbf{x}}_{k|k-1} + \tilde{\mathbf{K}}_k^{(i+1)}(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1})$
7. $\boldsymbol{\Sigma}_k^{(i+1)} = \tilde{\mathbf{P}}_{k|k-1}^{(i+1)} - \tilde{\mathbf{K}}_k^{(i+1)}\mathbf{H}_k\tilde{\mathbf{P}}_{k|k-1}^{(i+1)}$

Check the convergence of iteration

8. If $\frac{\|\boldsymbol{\mu}_k^{(i+1)} - \boldsymbol{\mu}_k^{(i)}\|}{\|\boldsymbol{\mu}_k^{(i)}\|} \le \epsilon$, stop iteration.

Calculate $\mathbf{A}_k^{(i+1)}$ and $\mathbf{B}_k^{(i+1)}$

9. $\mathbf{A}_k^{(i+1)} = \boldsymbol{\Sigma}_k^{(i+1)} + (\boldsymbol{\mu}_k^{(i+1)} - \hat{\mathbf{x}}_{k|k-1})(\boldsymbol{\mu}_k^{(i+1)} - \hat{\mathbf{x}}_{k|k-1})^{\mathrm{T}}$
10. $\mathbf{B}_k^{(i+1)} = (\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k^{(i+1)})(\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k^{(i+1)})^{\mathrm{T}} + \mathbf{H}_k\boldsymbol{\Sigma}_k^{(i+1)}\mathbf{H}_k^{\mathrm{T}}$

Evaluate $\xi_k^{(i+1)}$ and $\lambda_k^{(i+1)}$

11. $\xi_k^{(i+1)} = -2\dot{f}_x(\mathrm{tr}(\mathbf{A}_k^{(i+1)}\mathbf{P}_{k|k-1}^{-1}))$
12. $\lambda_k^{(i+1)} = -2\dot{f}_z(\mathrm{tr}(\mathbf{B}_k^{(i+1)}\mathbf{R}_k^{-1}))$

Check the constraints in (4.36)

13. If $\xi_k^{(i+1)} < \delta$, then $\xi_k^{(i+1)} = \delta$.
14. If $\lambda_k^{(i+1)} < \delta$, then $\lambda_k^{(i+1)} = \delta$.

end for

15. $\hat{x}_{k|k} = \boldsymbol{\mu}_k^{(i^*)}$, $\mathbf{P}_{k|k} = \boldsymbol{\Sigma}_k^{(i^*)}$

Outputs: $\hat{x}_{k|k}$ and $\mathbf{P}_{k|k}$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

The premise of applying the SSMKF in practical engineering is to guarantee the convergence of the fixed-point iterations. To this end, the relationship between the fixed-point iterative algorithm and the existing nonlinear optimization algorithm is first revealed.

Proposition 4.4.1. *The fixed-point iterative algorithm is identical to the existing Newton's method with the modified Hessian matrix as follows*:

$$\tilde{\boldsymbol{\Theta}}_{\mu_k}(\boldsymbol{\mu}_k^{(i)}, \boldsymbol{\Sigma}_k^{(i)}) = \boldsymbol{\Theta}_{\mu_k}(\boldsymbol{\mu}_k^{(i)}, \boldsymbol{\Sigma}_k^{(i)}) - \mathbf{D}_1(\boldsymbol{\mu}_k^{(i)}, \boldsymbol{\Sigma}_k^{(i)}). \tag{4.61}$$

Proof. Using (4.61) and (4.33), the modified Hessian matrix is calculated as

$$\tilde{\boldsymbol{\Theta}}_{\mu_k}(\boldsymbol{\mu}_k^{(i)}, \boldsymbol{\Sigma}_k^{(i)}) = -\xi_k^{(i)}\mathbf{P}_{k|k-1}^{-1} - \lambda_k^{(i)}\mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}\mathbf{H}_k. \tag{4.62}$$

According to Newton's iterative scheme, we have

$$\begin{cases} \boldsymbol{\mu}_k^{(i+1)} = \boldsymbol{\mu}_k^{(i)} - [\tilde{\boldsymbol{\Theta}}_{\mu_k}(\boldsymbol{\mu}_k^{(i)}, \boldsymbol{\Sigma}_k^{(i)})]^{-1}\boldsymbol{\Delta}_{\mu_k}(\boldsymbol{\mu}_k^{(i)}, \boldsymbol{\Sigma}_k^{(i)}), \\ \boldsymbol{\Sigma}_k^{(i+1)} = -[\tilde{\boldsymbol{\Theta}}_{\mu_k}(\boldsymbol{\mu}_k^{(i)}, \boldsymbol{\Sigma}_k^{(i)})]^{-1}. \end{cases} \tag{4.63}$$

Employing (4.26) and (4.62) in (4.63) results in

$$\begin{cases} \boldsymbol{\mu}_k^{(i+1)} = \boldsymbol{\Sigma}_k^{(i+1)}(\xi_k^{(i)}\mathbf{P}_{k|k-1}^{-1}\hat{\mathbf{x}}_{k|k-1} + \lambda_k^{(i)}\mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}\mathbf{z}_k), \\ \boldsymbol{\Sigma}_k^{(i+1)} = (\xi_k^{(i)}\mathbf{P}_{k|k-1}^{-1} + \lambda_k^{(i)}\mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}\mathbf{H}_k)^{-1}. \end{cases} \tag{4.64}$$

By utilizing the matrix inversion lemma [35, pp. 11–12], $\boldsymbol{\mu}_k^{(i+1)}$ and $\boldsymbol{\Sigma}_k^{(i+1)}$ in (4.64) can be written as the 6th and 7th equations in Algorithm 4.1. Thus, we can obtain Proposition 4.4.1. □

Proposition 4.4.1 means that the fixed-point iterative algorithm is an improved version of the existing Newton's method. The positive semi-definite term $\mathbf{D}_1(\boldsymbol{\mu}_k^{(i)}, \boldsymbol{\Sigma}_k^{(i)})$ is subtracted from the original Hessian matrix so that the negative definiteness can be preserved. As a result, the fixed-point iterative algorithm has better numerical stability than the standard Newton's method. Normally, to guarantee the local convergence of Newton's method, the Hessian matrix needs to satisfy the Lipschitz condition, and the initial value is sufficiently close to the optimal value. Next, we will provide the convergence conditions of the fixed-point iterative algorithm.

Theorem 4.4.2. *If the initial mean vector $\mu_k^{(0)}$ is sufficiently close to the optimal mean vector μ_k^* and there are positive and bounded real numbers α_1 and α_2 making the following inequalities hold*

$$\begin{cases} 0 \le \ddot{f}_x(t^2)t \le \alpha_1, \\ 0 \le \ddot{f}_z(t^2)t \le \alpha_2, \end{cases} \quad \forall t \ge 0, \tag{4.65}$$

then the fixed-point iterative algorithm has local convergence.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Proof. Let $\boldsymbol{\Delta}_{\tilde{\boldsymbol{\Theta}}}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{\partial \tilde{\boldsymbol{\Theta}}_{\mu_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\partial \boldsymbol{\mu}_k}$. Using (4.14), (4.15), (4.18), (4.61) and (4.33) yields

$$\begin{aligned}\boldsymbol{\Delta}_{\tilde{\boldsymbol{\Theta}}}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = {} & 2\tilde{\xi}_k[\mathbf{P}_{k|k-1}^{-1}(\boldsymbol{\mu}_k - \hat{\mathbf{x}}_{k|k-1})] \otimes \mathbf{P}_{k|k-1}^{-1} \\ & - 2\tilde{\lambda}_k[\mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}(\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k)] \otimes [\mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}\mathbf{H}_k],\end{aligned} \tag{4.66}$$

where $\otimes$ denotes the Kronecker product.

Taking the norm on both sides of (4.66) and utilizing the properties of matrix norms results in

$$\left\|\boldsymbol{\Delta}_{\tilde{\boldsymbol{\Theta}}}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)\right\|_F \leq 2\beta_1\tilde{\xi}_k\left\|\mathbf{S}_{k|k-1}^{-1}(\boldsymbol{\mu}_k - \hat{\mathbf{x}}_{k|k-1})\right\|_F + 2\beta_2\tilde{\lambda}_k\left\|\mathbf{S}_{\mathbf{R}_k}^{-1}(\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k)\right\|_F, \tag{4.67}$$

where $\|\cdot\|$ denotes the Frobenius norm and β_1 and β_2 are, respectively, given by

$$\begin{cases}\beta_1 = \|\mathbf{S}_{k|k-1}^{-\mathrm{T}}\|_F\|\mathbf{P}_{k|k-1}^{-1}\|_F, \\ \beta_2 = \|\mathbf{H}_k^{\mathrm{T}}\mathbf{S}_{\mathbf{R}_k}^{-\mathrm{T}}\|_F\|\mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}\mathbf{H}_k\|_F.\end{cases} \tag{4.68}$$

Employing (4.8) and (4.14)–(4.15), we have

$$\begin{cases}\|\mathbf{S}_{k|k-1}^{-1}(\boldsymbol{\mu}_k - \hat{\mathbf{x}}_{k|k-1})\|_F \leq \sqrt{\mathrm{tr}(\mathbf{A}_k\mathbf{P}_{k|k-1}^{-1})}, \\ \|\mathbf{S}_{\mathbf{R}_k}^{-1}(\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k)\|_F \leq \sqrt{\mathrm{tr}(\mathbf{B}_k\mathbf{R}_k^{-1})}.\end{cases} \tag{4.69}$$

Exploiting (4.18), (4.65) and (4.68)–(4.69) in (4.67) gives

$$\left\|\boldsymbol{\Delta}_{\tilde{\boldsymbol{\Theta}}}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)\right\|_F \leq 4\alpha_1\beta_1 + 4\alpha_2\beta_2. \tag{4.70}$$

Define an auxiliary function as follows:

$$\varphi(\tau) = \tilde{\boldsymbol{\Theta}}_{\mu_k}(\boldsymbol{\mu}_k^2 + \tau(\boldsymbol{\mu}_k^1 - \boldsymbol{\mu}_k^2), \boldsymbol{\Sigma}_k) \quad \text{s.t.} \quad \tau \in [0,\ 1], \tag{4.71}$$

where $\boldsymbol{\mu}_k^1$ and $\boldsymbol{\mu}_k^2$ are arbitrary two posterior mean vectors.

Taking the first-order derivative of $\varphi(\tau)$ obtains

$$\dot{\varphi}(\tau) = \boldsymbol{\Delta}_{\tilde{\boldsymbol{\Theta}}}(\boldsymbol{\mu}_k^2 + \tau(\boldsymbol{\mu}_k^1 - \boldsymbol{\mu}_k^2), \boldsymbol{\Sigma}_k)(\boldsymbol{\mu}_k^1 - \boldsymbol{\mu}_k^2). \tag{4.72}$$

According to the Lagrange mean value theorem, there is a variable $\theta \in [0,\ 1]$ such that the following equation is fulfilled:

$$\varphi(1) - \varphi(0) = \dot{\varphi}(\theta)(1 - 0). \tag{4.73}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Substituting (4.71)–(4.72) in (4.73) yields

$$\tilde{\boldsymbol{\Theta}}_{\mu_k}(\boldsymbol{\mu}_k^1, \boldsymbol{\Sigma}_k) - \tilde{\boldsymbol{\Theta}}_{\mu_k}(\boldsymbol{\mu}_k^2, \boldsymbol{\Sigma}_k) = \boldsymbol{\Delta}_{\tilde{\boldsymbol{\Theta}}}(\boldsymbol{\mu}_k^2 + \theta(\boldsymbol{\mu}_k^1 - \boldsymbol{\mu}_k^2), \boldsymbol{\Sigma}_k)(\boldsymbol{\mu}_k^1 - \boldsymbol{\mu}_k^2). \tag{4.74}$$

Taking the norm on both sides of (4.74) and using (4.70), we have

$$\|\tilde{\boldsymbol{\Theta}}_{\mu_k}(\boldsymbol{\mu}_k^1, \boldsymbol{\Sigma}_k) - \tilde{\boldsymbol{\Theta}}_{\mu_k}(\boldsymbol{\mu}_k^2, \boldsymbol{\Sigma}_k)\|_F \leq \beta \|\boldsymbol{\mu}_k^1 - \boldsymbol{\mu}_k^2\|, \tag{4.75}$$

where β is given by

$$\beta = 4\alpha_1\beta_1 + 4\alpha_2\beta_2. \tag{4.76}$$

It can be seen from (4.75)–(4.76) that the modified Hessian matrix $\tilde{\boldsymbol{\Theta}}_{\mu_k}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ satisfies the Lipschitz condition. Thus, if the initial mean vector $\mu_k^{(0)}$ is sufficiently close to the optimal mean vector μ_k^*, then the fixed-point iterative algorithm has local convergence. □

Generally, the initial mean vector $\mu_k^{(0)}$ is selected as the one-step predicted state vector $\hat{\mathbf{x}}_{k|k-1}$, i. e., $\mu_k^{(0)} = \hat{\mathbf{x}}_{k|k-1}$, since $\hat{\mathbf{x}}_{k|k-1}$ is the only available information for the mean vector before the iterative measurement update. It is seen from Theorem 4.4.2 that the initial value $\mu_k^{(0)}$, i. e., $\hat{\mathbf{x}}_{k|k-1}$, needs to be sufficiently close to the optimal value μ_k^* to guarantee the local convergence of the fixed-point iterative algorithm. Unfortunately, the one-step predicted state vector $\hat{\mathbf{x}}_{k|k-1}$ may be far away from the optimal value μ_k^* when the linear system suffers from large process uncertainty or a state outlier. As a result, the fixed-point iterative algorithm may not converge to a local optimum when the linear system suffers from large process uncertainty or a state outlier, which will degrade the filtering accuracy of the SSMKF dramatically.

4.4.2 Separate iterative algorithm

Before introducing the separate iterative algorithm, we first present the problems existing in the fixed-point iterative algorithm. It can be seen from Algorithm 4.1 that the Kalman gain $\tilde{\mathbf{K}}_k^{(i+1)}$ plays an important role in the iterative measurement update because it can adjust the weights of one-step predicted state vector $\hat{\mathbf{x}}_{k|k-1}$ and measurement innovation $\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1}$ adaptively. Next, we will discuss the behavior of the Kalman gain $\tilde{\mathbf{K}}_k^{(i+1)}$ during the iterative measurement update.

Using the 4th and 5th equations in Algorithm 4.1, the Kalman gain $\tilde{\mathbf{K}}_k^{(i+1)}$ can be rewritten as

$$\tilde{\mathbf{K}}_k^{(i+1)} = \mathbf{P}_{k|k-1}\mathbf{H}_k^T\left(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \frac{\xi_k^{(i)}}{\lambda_k^{(i)}}\mathbf{R}_k\right)^{-1}. \tag{4.77}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

It is observed from (4.77) that the behavior of the Kalman gain $\tilde{\mathbf{K}}_k^{(i+1)}$ depends only on the ratio of auxiliary parameters, i. e., $\frac{\xi_k^{(i)}}{\lambda_k^{(i)}}$, during the iterative measurement update. It can be also observed from the 9th–12th equations in Algorithm 4.1 that the ratio of auxiliary parameters $\frac{\xi_k^{(i)}}{\lambda_k^{(i)}}$ relies on not only the similarity functions $f_x(\cdot)$ and $f_z(\cdot)$ but also the parameters $\mathbf{A}_k^{(i)}$ and $\mathbf{B}_k^{(i)}$, and the parameters $\mathbf{A}_k^{(i)}$ and $\mathbf{B}_k^{(i)}$ are adaptively adjusted by the iterative posterior mean vector $\boldsymbol{\mu}_k^{(i)}$ and covariance matrix $\boldsymbol{\Sigma}_k^{(i)}$ at the same time. As a result, the behavior of the Kalman gain $\tilde{\mathbf{K}}_k^{(i+1)}$ is likely to be indefinite and depends heavily on the relative size of state and measurement outliers. The indefinite behavior of the Kalman gain $\tilde{\mathbf{K}}_k^{(i+1)}$ can easily cause the iterative posterior mean vector $\boldsymbol{\mu}_k^{(i)}$ and covariance matrix $\boldsymbol{\Sigma}_k^{(i)}$ to not converge to local optimums.

To solve this problem, [22] propose a heuristic idea that the iterations of auxiliary parameters are separated to guarantee the definite behavior of the Kalman gain. That is to say, the auxiliary parameter $\lambda_k^{(i)(0)}$ is first iterated with fixed auxiliary parameter $\xi_k^{(0)(0)} = 1$ until convergence, i. e., $\lim_{i\to+\infty} \lambda_k^{(i)(0)} = \bar{\lambda}_k$, and then the auxiliary parameter $\xi_k^{(0)(j)}$ is iterated with fixed auxiliary parameter $\bar{\lambda}_k$ until convergence, i. e., $\lim_{j\to+\infty} \xi_k^{(0)(j)} = \bar{\xi}_k$, where $\bar{\xi}_k$ and $\bar{\lambda}_k$ denote the local optimums of the auxiliary parameters. The detailed implementation of the SSMKF based on the separate iterative algorithm is listed in Algorithm 4.2, where ϵ_1 and ϵ_2 denote the iteration thresholds of ξ_k and λ_k, respectively, N_1 and N_2 denote the maximum numbers of iterations of ξ_k and λ_k, respectively, and i_1^* and j_2^* denote the cyclic variables at the end of the first and second loops, respectively.

Theorem 4.4.3. *If the similarity functions $f_z(\cdot)$ and $f_x(\cdot)$ satisfy Condition* 3 *and $\dot{f}_z(0)$ and $\dot{f}_x(0)$ have lower bounds, then the iterative auxiliary parameters $\lambda_k^{(i)(0)}$ and $\xi_k^{(0)(j)}$ will converge to local optimums $\bar{\lambda}_k$ and $\bar{\xi}_k$, respectively, i. e.,*

$$\lim_{i\to+\infty} \lambda_k^{(i)(0)} = \bar{\lambda}_k, \quad \lim_{j\to+\infty} \xi_k^{(0)(j)} = \bar{\xi}_k. \tag{4.78}$$

Proof. First, we prove the convergence of $\lambda_k^{(i)(0)}$. To this end, we consider the two cases: $\lambda_k^{(1)(0)} \geq \lambda_k^{(0)(0)}$ and $\lambda_k^{(1)(0)} \leq \lambda_k^{(0)(0)}$.

Case 1: We first assume $\lambda_k^{(i)(0)} \geq \lambda_k^{(i-1)(0)}$. Then, using the 4th–8th equations in Algorithm 4.2 yields

$$\begin{cases} \mathbf{B}_k^{(i+1)(0)} = (\mathbf{I}_m - \mathbf{H}_k\tilde{\mathbf{K}}_k^{(i+1)(0)})(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1})(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1})^{\mathrm{T}}(\mathbf{I}_m - \mathbf{H}_k\tilde{\mathbf{K}}_k^{(i+1)(0)})^{\mathrm{T}}, \\ \mathbf{I}_m - \mathbf{H}_k\tilde{\mathbf{K}}_k^{(i+1)(0)} = \mathbf{R}_k(\mathbf{H}_k\lambda_k^{(i)(0)}\mathbf{P}_{k|k-1}\mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k)^{-1}. \end{cases} \tag{4.79}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Algorithm 4.2: One time step of the SSMKF based on the separate iterative algorithm.

Inputs: $\hat{\mathbf{x}}_{k-1|k-1}$, $\mathbf{P}_{k-1|k-1}$, $\mathbf{F}_k$, $\mathbf{H}_k$, $\mathbf{z}_k$, $\mathbf{Q}_k$, $\mathbf{R}_k$, $f_x(\cdot)$, $f_z(\cdot)$, ϵ_1, ϵ_2, N_1, N_2, δ.

Time update:

1. $\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k\hat{\mathbf{x}}_{k-1|k-1}$
2. $\mathbf{P}_{k|k-1} = \mathbf{F}_k\mathbf{P}_{k-1|k-1}\mathbf{F}_k^T + \mathbf{Q}_k$

Iterative measurement update:

3. Initialization: $\xi_k^{(0)(0)} = 1$, $\lambda_k^{(0)(0)} = 1$

for $i = 0 : N_1 - 1$

Calculate $\tilde{\mathbf{P}}_{k|k-1}^{(i+1)(0)}$ and $\tilde{\mathbf{R}}_k^{(i+1)(0)}$

4. $\tilde{\mathbf{P}}_{k|k-1}^{(i+1)(0)} = \mathbf{P}_{k|k-1}/\xi_k^{(0)(0)}$, $\tilde{\mathbf{R}}_k^{(i+1)(0)} = \mathbf{R}_k/\lambda_k^{(i)(0)}$

Evaluate $\boldsymbol{\mu}_k^{(i+1)(0)}$ and $\boldsymbol{\Sigma}_k^{(i+1)(0)}$

5. $\tilde{\mathbf{K}}_k^{(i+1)(0)} = \tilde{\mathbf{P}}_{k|k-1}^{(i+1)(0)}\mathbf{H}_k^T(\mathbf{H}_k\tilde{\mathbf{P}}_{k|k-1}^{(i+1)(0)}\mathbf{H}_k^T + \tilde{\mathbf{R}}_k^{(i+1)(0)})^{-1}$
6. $\boldsymbol{\mu}_k^{(i+1)(0)} = \hat{\mathbf{x}}_{k|k-1} + \tilde{\mathbf{K}}_k^{(i+1)(0)}(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1})$
7. $\boldsymbol{\Sigma}_k^{(i+1)(0)} = \tilde{\mathbf{P}}_{k|k-1}^{(i+1)(0)} - \tilde{\mathbf{K}}_k^{(i+1)(0)}\mathbf{H}_k\tilde{\mathbf{P}}_{k|k-1}^{(i+1)(0)}$

Calculate $\mathbf{B}_k^{(i+1)(0)}$

8. $\mathbf{B}_k^{(i+1)(0)} = (\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k^{(i+1)(0)})(\mathbf{z}_k - \mathbf{H}_k\boldsymbol{\mu}_k^{(i+1)(0)})^{\mathrm{T}} + \mathbf{H}_k\boldsymbol{\Sigma}_k^{(i+1)(0)}\mathbf{H}_k^{\mathrm{T}}$

Evaluate $\lambda_k^{(i+1)(0)}$

9. $\lambda_k^{(i+1)(0)} = -2\dot{f}_z(\mathrm{tr}(\mathbf{B}_k^{(i+1)(0)}\mathbf{R}_k^{-1}))$

Check the constraints in (4.36)

10 If $\lambda_k^{(i+1)(0)} < \delta$, then $\lambda_k^{(i+1)(0)} = \delta$.

Check the convergence of iteration

11. If $\frac{|\lambda_k^{(i+1)(0)} - \lambda_k^{(i)(0)}|}{|\lambda_k^{(i)(0)}|} \le \epsilon_1$, stop iteration.

end for

for $j = 0 : N_2 - 1$

Calculate $\tilde{\mathbf{P}}_{k|k-1}^{(i_1^*)(j+1)}$ and $\tilde{\mathbf{R}}_k^{(i_1^*)(j+1)}$

12. $\tilde{\mathbf{P}}_{k|k-1}^{(i_1^*)(j+1)} = \mathbf{P}_{k|k-1}/\xi_k^{(0)(j)}$, $\tilde{\mathbf{R}}_k^{(i_1^*)(j+1)} = \mathbf{R}_k/\lambda_k^{(i_1^*)(0)}$

Evaluate $\boldsymbol{\mu}_k^{(i_1^*)(j+1)}$ and $\boldsymbol{\Sigma}_k^{(i_1^*)(j+1)}$

13. $\tilde{\mathbf{K}}_k^{(i_1^*)(j+1)} = \tilde{\mathbf{P}}_{k|k-1}^{(i_1^*)(j+1)}\mathbf{H}_k^T(\mathbf{H}_k\tilde{\mathbf{P}}_{k|k-1}^{(i_1^*)(j+1)}\mathbf{H}_k^T + \tilde{\mathbf{R}}_k^{(i_1^*)(j+1)})^{-1}$
14. $\boldsymbol{\mu}_k^{(i_1^*)(j+1)} = \hat{\mathbf{x}}_{k|k-1} + \tilde{\mathbf{K}}_k^{(i_1^*)(j+1)}(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1})$
15. $\boldsymbol{\Sigma}_k^{(i_1^*)(j+1)} = \tilde{\mathbf{P}}_{k|k-1}^{(i_1^*)(j+1)} - \tilde{\mathbf{K}}_k^{(i_1^*)(j+1)}\mathbf{H}_k\tilde{\mathbf{P}}_{k|k-1}^{(i_1^*)(j+1)}$

Calculate $\mathbf{A}_k^{(0)(j+1)}$

16. $\mathbf{A}_k^{(0)(j+1)} = (\boldsymbol{\mu}_k^{(i_1^*)(j+1)} - \hat{\mathbf{x}}_{k|k-1})(\boldsymbol{\mu}_k^{(i_1^*)(j+1)} - \hat{\mathbf{x}}_{k|k-1})^{\mathrm{T}} + \boldsymbol{\Sigma}_k^{(i_1^*)(j+1)}$

Evaluate $\xi_k^{(0)(j+1)}$

17. $\xi_k^{(0)(j+1)} = -2\dot{f}_x(\mathrm{tr}(\mathbf{A}_k^{(0)(j+1)}\mathbf{P}_{k|k-1}^{-1}))$

Check the constraints in (4.36)

18. If $\xi_k^{(0)(j+1)} < \delta$, then $\xi_k^{(0)(j+1)} = \delta$.

Check the convergence of iteration

19. If $\frac{|\xi_k^{(0)(j+1)} - \xi_k^{(0)(j+1)}|}{|\xi_k^{(0)(j)}|} \le \epsilon_2$, stop iteration.

end for

20. $\hat{x}_{k|k} = \boldsymbol{\mu}_k^{(i_1^*)(j_2^*)}$, $\mathbf{P}_{k|k} = \boldsymbol{\Sigma}_k^{(i_1^*)(j_2^*)}$

Outputs: $\hat{x}_{k|k}$ and $\mathbf{P}_{k|k}$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Employing (4.79) and $\lambda_k^{(i)(0)} \geq \lambda_k^{(i-1)(0)}$ results in

$$\mathrm{tr}(\mathbf{B}_k^{(i)(0)}\mathbf{R}_k^{-1}) \geq \mathrm{tr}(\mathbf{B}_k^{(i+1)(0)}\mathbf{R}_k^{-1}). \tag{4.80}$$

Since $\ddot{f}_z(t) \geq 0$, $-\dot{f}_z(t)$ is a monotonically decreasing function, utilizing (4.80) and the 9th equation in Algorithm 4.2 yields

$$\lambda_k^{(i+1)(0)} \leq \lambda_k^{(i)(0)}. \tag{4.81}$$

According to the mathematical induction method, we have

$$\lambda_k^{(0)(0)} \leq \lambda_k^{(1)(0)} \leq \cdots \leq \lambda_k^{(i)(0)} \leq \cdots \leq \lambda_k^{(\infty)(0)}. \tag{4.82}$$

Considering that $-\dot{f}_z(t)$ is a monotonically decreasing function and $\mathrm{tr}(\mathbf{B}_k^{(i)(0)}\mathbf{R}_k^{-1}) \geq 0$, we obtain

$$\lambda_k^{(i)(0)} = -2\dot{f}_z(\mathrm{tr}(\mathbf{B}_k^{(i)(0)}\mathbf{R}_k^{-1})) \leq -2\dot{f}_z(0). \tag{4.83}$$

It can be seen from (4.82)–(4.83) that $\{\lambda_k^{(i)(0)}\}$ is a monotonically increasing sequence with an upper bound $-2\dot{f}_z(0)$. Thus, the sequence $\{\lambda_k^{(i)(0)}\}$ converges when $\lambda_k^{(1)(0)} \geq \lambda_k^{(0)(0)}$ and $\dot{f}_z(0)$ has a lower bound.

Case 2: Similar to the Case 1, if $\lambda_k^{(1)(0)} \leq \lambda_k^{(0)(0)}$, we have

$$\lambda_k^{(0)(0)} \geq \lambda_k^{(1)(0)} \geq \cdots \geq \lambda_k^{(i)(0)} \geq \cdots \geq \lambda_k^{(\infty)(0)} > 0. \tag{4.84}$$

It can be seen from (4.84) that $\{\lambda_k^{(i)(0)}\}$ is a monotonically decreasing sequence with a lower bound 0. Thus, the sequence $\{\lambda_k^{(i)(0)}\}$ also converges when $\lambda_k^{(1)(0)} \leq \lambda_k^{(0)(0)}$. Above all, the sequence $\{\lambda_k^{(i)(0)}\}$ converges if the similarity function $f_z(\cdot)$ satisfies Condition 3 and $\dot{f}_z(0)$ has lower bound. Similarly, we can also prove that the sequence $\xi_k^{(0)(j)}$ converges if the similarity function $f_x(\cdot)$ satisfies Condition 3 and $\dot{f}_x(0)$ has lower bound. □

We can see from Theorem 4.4.3 that, if the similarity functions $f_x(\cdot)$ and $f_z(\cdot)$ satisfy Condition 3 and $\dot{f}_z(0)$ and $\dot{f}_x(0)$ have lower bounds, the iterative auxiliary parameters $\lambda_k^{(i)(0)}$ and $\xi_k^{(0)(j)}$ will converge to local optima, and then the corresponding posterior mean vector and covariance matrix also converge to local optima, which guarantees the local convergence of the separate iterative algorithm. As compared with the fixed-point iteration algorithm, the separate iterative algorithm does not require an initial mean vector to be sufficiently close to the optimal mean vector. Thus, the local convergence conditions of the fixed-point iterative algorithm are more harsh than those of the separate iterative algorithm.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.5 Selections of the similarity functions

It is seen from Algorithm 4.1 and Algorithm 4.2 that the similarity functions $f_x(\cdot)$ and $f_z(\cdot)$ are necessary to implement the SSMKF. Next, we will provide the selections of the similarity functions to guarantee that the SSMKF is identical to the standard Kalman filter for the case of Gaussian noises and robust to the state and measurement outliers.

Firstly, the relationships between the approximate one-step prediction error covariance matrix $\mathbf{A}_k^*$ and the true one-step prediction error covariance matrix $\mathbf{P}_{k|k-1}$ and that between the approximate measurement noise covariance matrix $\mathbf{B}_k^*$ and the true measurement noise covariance matrix $\mathbf{R}_k$ are studied for a linear state-space model with Gaussian state and measurement noises.

Proposition 4.5.1. *For a linear state-space model with Gaussian state and measurement noises, $\mathbf{A}_k^*$ and $\mathbf{B}_k^*$ can be, respectively, approximated as the one-step prediction error covariance matrix and measurement noise covariance matrix, i. e.,*

$$\mathbf{A}_k^* \approx \mathbf{P}_{k|k-1}, \quad \mathbf{B}_k^* \approx \mathbf{R}_k. \tag{4.85}$$

Proof. Using the Kalman measurement update equations in (4.24)–(4.25) yields

$$\begin{cases} \mathbf{A}_k^* = \mathbf{P}_{k|k} + \mathbf{K}_k \bar{\mathbf{P}}_{k|k-1}^{zz} \mathbf{K}_k^{\mathrm{T}}, \\ \mathbf{B}_k^* = (\mathbf{I}_m - \mathbf{H}_k \mathbf{K}_k) \bar{\mathbf{P}}_{k|k-1}^{zz} (\mathbf{I}_m - \mathbf{H}_k \mathbf{K}_k)^{\mathrm{T}} + \mathbf{H}_k \mathbf{P}_{k|k} \mathbf{H}_k^{\mathrm{T}}. \end{cases} \tag{4.86}$$

Considering that $\bar{\mathbf{P}}_{k|k-1}^{zz} \approx \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k$ for the case of Gaussian noises and employing (4.86) results in

$$\begin{cases} \mathbf{A}_k^* \approx \mathbf{P}_{k|k} + \mathbf{K}_k (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k) \mathbf{K}_k^{\mathrm{T}}, \\ \mathbf{B}_k^* \approx \mathbf{R}_k - \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^{\mathrm{T}} + \mathbf{H}_k \mathbf{K}_k (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k) \mathbf{K}_k^{\mathrm{T}} \mathbf{H}_k^{\mathrm{T}} + \mathbf{H}_k \mathbf{P}_{k|k} \mathbf{H}_k^{\mathrm{T}}. \end{cases} \tag{4.87}$$

Substituting the Kalman measurement update equation in (4.87), we can obtain (4.85). □

Second, we study the conditions of the similarity functions $f_x(\cdot)$ and $f_z(\cdot)$ to guarantee that the SSMKF is identical to the standard Kalman filter for the case of Gaussian noises. Substituting (4.85) in (4.23), we have

$$\xi_k^* = -2\dot{f}_x(n), \quad \lambda_k^* = -2\dot{f}_z(m) \tag{4.88}$$

where n and m are, respectively, the dimensions of the state vector and the measurement vector.

It is seen from Algorithm 4.1 and Algorithm 4.2 that the SSMKF is identical to the standard Kalman filter when the modified parameters are unity, i. e., $\xi_k^* = \lambda_k^* = 1$. Therefore, to guarantee that the SSMKF is identical to the standard Kalman filter when

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

there are no state and measurement outliers, the similarity functions $f_x(\cdot)$ and $f_z(\cdot)$ need to satisfy the following equations:

$$\dot{f}_x(n) = -0.5, \quad \dot{f}_z(m) = -0.5. \tag{4.89}$$

Finally, we discuss the conditions of the similarity functions $f_x(\cdot)$ and $f_z(\cdot)$ to guarantee that the SSMKF is robust to outliers. Define two auxiliary matrices as follows:

$$\mathbf{\Psi}_{k1} = \mathbf{A}_k^* - \mathbf{P}_{k|k-1}, \quad \mathbf{\Psi}_{k2} = \mathbf{B}_k^* - \mathbf{R}_k. \tag{4.90}$$

In general, if there are state and measurement outliers, the approximate one-step prediction error covariance matrix $\mathbf{A}_k^*$ is not less than the nominal one-step prediction error covariance matrix $\mathbf{P}_{k|k-1}$, and the approximate measurement noise covariance matrix $\mathbf{B}_k^*$ is also not less than the nominal measurement noise covariance matrix $\mathbf{R}_k$, i. e.,

$$\mathbf{\Psi}_{k1} \geq \mathbf{0}, \quad \mathbf{\Psi}_{k2} \geq \mathbf{0}. \tag{4.91}$$

Furthermore, the auxiliary matrices $\mathbf{\Psi}_{k1}$ and $\mathbf{\Psi}_{k2}$ depend on the magnitudes of the state and measurement outliers, respectively, and the larger the state and measurement outliers, the larger the auxiliary matrices $\mathbf{\Psi}_{k1}$ and $\mathbf{\Psi}_{k2}$ that will be generated.

Proposition 4.5.2. *For a linear state-space model with outlier-contaminated state and measurement noises, if the similarity functions $f_x(\cdot)$ and $f_z(\cdot)$ satisfy the following conditions*

$$\begin{cases} \dot{f}_x(t) < 0, \quad \ddot{f}_x(t) \geq 0, \quad \dot{f}_x(n) = -0.5, \quad t \in [0, +\infty), \\ \dot{f}_z(t) < 0, \quad \ddot{f}_z(t) \geq 0, \quad \dot{f}_z(m) = -0.5, \quad t \in [0, +\infty), \end{cases} \tag{4.92}$$

then the modified auxiliary parameters ξ_k^* and λ_k^* satisfy the following equations:

$$0 < \xi_k^* \leq 1, \quad 0 < \lambda_k^* \leq 1, \tag{4.93}$$

and the larger are the state and measurement outliers, the smaller will be the modified auxiliary parameters ξ_k^* and λ_k^* obtained.

Proof. Using (4.90)–(4.91) yields

$$\begin{cases} \eta_1 = \mathrm{tr}(\mathbf{A}_k^* \mathbf{P}_{k|k-1}^{-1}) = \mathrm{tr}(\mathbf{\Psi}_{k1} \mathbf{P}_{k|k-1}^{-1}) + n \geq n, \\ \eta_2 = \mathrm{tr}(\mathbf{B}_k^* \mathbf{R}_k^{-1}) = \mathrm{tr}(\mathbf{\Psi}_{k2} \mathbf{R}_k^{-1}) + m \geq m. \end{cases} \tag{4.94}$$

Substituting (4.94) into (4.23) results in

$$\xi_k^* = -2\dot{f}_x(\eta_1), \quad \lambda_k^* = -2\dot{f}_z(\eta_2). \tag{4.95}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Employing (4.92) in (4.95), we can obtain (4.93). Moreover, the larger are the state and measurement outliers, the larger are the auxiliary matrices $\mathbf{\Psi}_{k1}$ and $\mathbf{\Psi}_{k2}$ generated. Then, the larger η_1 and η_2 are obtained, based on which the smaller modified auxiliary parameters ξ_k^* and λ_k^* are achieved by using (4.92). □

Employing (4.93) in (4.22) results in

$$\tilde{\mathbf{P}}_{k|k-1}^* - \mathbf{P}_{k|k-1} \geq \mathbf{0}, \quad \tilde{\mathbf{R}}_k^* - \mathbf{R}_k \geq \mathbf{0}. \tag{4.96}$$

It is observed from (4.96) that the modified prediction error covariance matrix and the modified measurement noise covariance matrix are, respectively, not less than the nominal prediction error covariance matrix and the nominal measurement noise covariance matrix when there are, respectively, state and measurement outliers. Moreover, according to Proposition 4.5.2 and (4.22), the larger are the state and measurement outliers, the smaller the modified auxiliary parameters ξ_k^* and λ_k^* will be, and the larger are the modified prediction error covariance matrix and the modified measurement noise covariance matrix will become.

Using (4.21)–(4.22), the Kalman gain $\tilde{\mathbf{K}}_k^*$ can be reformulated as

$$\tilde{\mathbf{K}}_k^* = \mathbf{P}_{k|k-1}\mathbf{H}_k^T\left(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \frac{\xi_k^*}{\lambda_k^*}\mathbf{R}_k\right)^{-1}. \tag{4.97}$$

According to Proposition 4.5.2 and (4.97), the Kalman gain $\tilde{\mathbf{K}}_k^*$ depends heavily on the relative magnitudes of the state and measurement outliers. Specifically, the Kalman gain is increased if the state outlier has a larger magnitude than the measurement outlier, and vice versa. As a result, the negative effects of outliers on the SSMKF can be resisted through adjusting the Kalman gain, adaptively.

As an example, some exemplary similarity functions $f(\cdot)$ are listed in Table 4.1, where p denotes the dimension of the state vector when $f_x(\cdot)$ is selected as $f(\cdot)$ and denotes the dimension of the measurement vector when $f_z(\cdot)$ is selected as $f(\cdot)$, and σ and v are, respectively, named as the kernel size and the dof parameter to be consistent with the existing MCKF [8] and RSTKF [18], and ω is also named as the dof parameter.

It is easy to verify that the exemplary similarity functions listed in Table 4.1 satisfy the conditions of Theorems 4.2.1 and 4.3.1 and Proposition 4.5.2. Theorem 4.3.2 and

Table 4.1: Exemplary similarity functions $f(\cdot)$ and their first and second derivatives.

$f(t)$	$\dot{f}(t)$	$\ddot{f}(t)$
$-0.5t$	-0.5	0
$\sigma^2\exp(\frac{p-t}{2\sigma^2})$	$-0.5\exp(\frac{p-t}{2\sigma^2})$	$\frac{1}{4\sigma^2}\exp(\frac{p-t}{2\sigma^2})$
$-0.5(v+p)\log(1+\frac{t}{v})$	$-0.5\frac{v+p}{v+t}$	$0.5\frac{v+p}{(v+t)^2}$
$-\sqrt{(\omega+p)(\omega+t)}$	$-0.5\sqrt{\frac{\omega+p}{\omega+t}}$	$0.25\frac{\sqrt{\omega+p}}{\sqrt[3]{\omega+t}}$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Propositions 4.3.4, 4.3.5, 4.4.1 and 4.5.1 do not need the exemplary similarity functions to satisfy the additional conditions. Next, we will further confirm whether the exemplary similarity functions satisfy the conditions of Theorems 4.3.3, 4.4.2 and 4.4.3 and Proposition 4.3.6.

Corollary 4.5.3. *If the kernel size σ and the dof parameter ν satisfy the following constraints*:

$$\sigma^2 \exp\left(\frac{n - Y_{1k}^*}{2\sigma^2}\right) > 1, \quad \sigma^2 \exp\left(\frac{m - Y_{2k}^*}{2\sigma^2}\right) > 1, \quad \nu > 2 - p, \tag{4.98}$$

then the exemplary similarity functions in Table 4.1 *satisfy the equation* (4.30) *in Theorem* 4.3.3.

Proof. Using (4.18) and (4.45), (4.30) can be rewritten as

$$\left[\dot{f}_x(Y_{1k}^*)\right]^2 \geq \ddot{f}_x(Y_{1k}^*), \quad \left[\dot{f}_z(Y_{2k}^*)\right]^2 \geq \ddot{f}_z(Y_{2k}^*), \tag{4.99}$$

where Y_{1k}^* and Y_{2k}^* are given in (4.45).

Substituting the exemplary similarity functions in Table 4.1 in (4.99), we can obtain (4.98). □

It is worth noting that there is no constraint on the dof parameter ω to guarantee that Theorem 4.3.3 holds. It is seen from Corollary 4.5.3 the constraint on the kernel size σ depends on the auxiliary parameters Y_{1k}^* and Y_{2k}^*. As a result, the constraint on the kernel size σ may change for various application scenarios. To address this problem, a reasonable scheme is choosing a sufficiently large kernel size σ so that the constraint on the kernel size always holds.

It can be seen from (4.93) that ξ_k^* and λ_k^* have positive upper bounds $\xi_{\max} = \lambda_{\max} = 1$. Since the second derivatives of the similarity functions are nonnegative, the minimum values of the negative derivatives of the similarity functions $-\dot{f}_x(t)$ and $-\dot{f}_z(t)$ are achieved at $t = +\infty$. It can be seen from Table 4.1 that the negative derivatives of the exemplary similarity functions approach 0 as t tends to $+\infty$. As a result, the modified auxiliary parameters ξ_k^* and λ_k^* do not have positive lower bounds, and then Proposition 4.3.6 does not hold, which may lead to filtering instability. To address this problem, we can impose a very small lower bound δ on the modified auxiliary parameters ξ_k^* and λ_k^* to guarantee filtering stability, as shown in the 13th and 14th equations of Algorithm 4.1 and the 10th and 18th equations of Algorithm 4.2.

Corollary 4.5.4. *For the exemplary similarity functions, if the kernel size $\sigma \nrightarrow 0$ and the dof parameters $\nu \nrightarrow 0$ and $\omega \nrightarrow 0$, then there exists positive and bounded real numbers α_1 and α_2, making equation* (4.65) *in Theorem* 4.4.2 *hold.*

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Proof. Let $g(t) = \ddot{f}(t^2)t$. For the case of exponential similarity function, the first-order derivative of $g(t)$ is formulated as follows:

$$\dot{g}(t) = \frac{1}{4\sigma^2}\exp\left(\frac{p-t^2}{2\sigma^2}\right)\left(1-\frac{t^2}{\sigma^2}\right). \tag{4.100}$$

It can be seen from (4.100) that $\dot{g}(t) \geq 0$ when $t \in [0,\sigma]$ and $\dot{g}(t) \leq 0$ when $t \in [\sigma,+\infty]$. Then, $g(t)$ achieves the maximum value at $t = \sigma$, and its maximum value is $\frac{1}{4\sigma}\exp(0.5p/\sigma^2 - 0.5)$. Thus, if the kernel size $\sigma \nrightarrow 0$, then there is a positive and bounded $\alpha = \frac{1}{4\sigma}\exp(0.5p/\sigma^2 - 0.5)$, making (4.65) hold.

Similarly, (4.65) also holds for logarithmic and square-root similarity functions if the dof parameters satisfy the constraints $\nu \nrightarrow 0$ and $\omega \nrightarrow 0$. □

Finally, we discuss the conditions of Theorem 4.4.3. It is easy to verify that, if the kernel size $\sigma \nrightarrow 0$ and the dof parameters $\nu \nrightarrow 0$ and $\omega \nrightarrow 0$, then $\dot{f}(0)$ has a lower bound, which guarantees the local convergence of the separate iterative algorithm. Also, according to Theorem 4.4.2 and Corollary 4.5.4, if the kernel size $\sigma \nrightarrow 0$ and the dof parameters $\nu \nrightarrow 0$ and $\omega \nrightarrow 0$, then the fixed-point iterative algorithm has local convergence when the initial mean vector $\mu_k^{(0)}$ is sufficiently close to the optimal mean vector μ_k^*. Thus, for the exemplary similarity functions, the convergence conditions of the separate iterative algorithm is easier to be satisfied as compared with that of the fixed-point iterative algorithm.

4.6 Relationships to other outlier-robust Kalman filters

In this section, we investigate the relationships between SSMKF and M-estimate based Kalman filters, RSTKFs and standard Kalman filters. In the M-estimate based Kalman filter, the state estimate is obtained by solving the following minimization problem [23]

$$\hat{\mathbf{x}}_{k|k} = \arg\min_{\mathbf{x}_k} J(\mathbf{x}_k), \tag{4.101}$$

where $J(\mathbf{x}_k)$ is formulated as

$$J(\mathbf{x}_k) = \sum_{i=1}^{n}\rho\big(\mathbf{N}_{ki}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\big) + \sum_{j=1}^{m}\rho\big(\mathbf{M}_{kj}(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_k)\big). \tag{4.102}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4.2: Exemplary similarity functions $f(\cdot)$ in SSMKF and their corresponding robust cost functions $\rho(\cdot)$ in M-estimation.

Function type	$f(t)$	$\rho(t)$
Exponential function	$\sigma^2 \exp(\frac{p-t}{2\sigma^2})$	$-\sigma^2 \exp(-\frac{t^2}{2\sigma^2})$
Logarithmic function	$-0.5(v+p)\log(1+\frac{t}{v})$	$\frac{v^2}{2}\log(1+\frac{t^2}{v})$
Square root function	$-\sqrt{(\omega+p)(\omega+t)}$	$\sqrt{\omega(\omega+t^2)}$

where $\rho(\cdot)$ is the robust cost function, and the vectors $\mathbf{N}_{ki}$ and $\mathbf{M}_{kj}$ are, respectively, defined as follows:

$$\begin{cases} \mathbf{S}_{k|k-1}^{-1} = [\mathbf{N}_{k1}^{\mathrm{T}}, \ldots, \mathbf{N}_{ki}^{\mathrm{T}}, \ldots, \mathbf{N}_{kn}^{\mathrm{T}}]^{\mathrm{T}}, \\ \mathbf{S}_{\mathbf{R}_k}^{-1} = [\mathbf{M}_{k1}^{\mathrm{T}}, \ldots, \mathbf{M}_{kj}^{\mathrm{T}}, \ldots, \mathbf{M}_{km}^{\mathrm{T}}]^{\mathrm{T}}, \\ \mathbf{P}_{k|k-1} = \mathbf{S}_{k|k-1}\mathbf{S}_{k|k-1}^{\mathrm{T}}, \quad \mathbf{R}_k = \mathbf{S}_{\mathbf{R}_k}\mathbf{S}_{\mathbf{R}_k}^{\mathrm{T}}. \end{cases} \tag{4.103}$$

Interestingly, when the robust cost function in (4.102) is chosen as $\rho(t) \propto -f(t^2)$, as shown in Table 4.2, we can see the connection and difference between the SSMKF framework and the M-estimator. When the robust cost function is set as $\rho(t) = -\sigma^2 \exp(-\frac{t^2}{2\sigma^2})$, the M-estimator turn out to be the MCKF. And the M-estimator with cost function $\rho(t) = \sqrt{\omega(\omega+t^2)}$ is similar to the Huber-based Kalman filter. In addition, We can obtain the Cauchy-based M-estimator by choosing the cost function as $\rho(t) = \frac{v^2}{2}\log(1+\frac{t^2}{v})$. Nonetheless, there are obvious difference between the two theoretical frameworks. It can be spotted from (4.101) and (4.102) that the posterior covariance matrix is independent of the minimization of the robust cost function in the M-estimator, and only a point estimate of the state vector can be obtained. Different from the M-estimate based Kalman filter, in the SSMKF, the posterior PDF of the state can be calculated by maximizing the lower-bound of the SSM. Thanks to the expectation operations in (4.16) and (4.17), the SSMKF takes advantage of the randomness inherent in the state vector, which is not utilized in the M-estimator. Therefore, the SSMKF framework is an improved version of the M-estimate based Kalman filter.

The RSTKF [18] and standard Kalman filter are both special cases of the SSMKF framework. It can be found from Algorithm 4.1 that the RSTKF is equivalent to the SSMKF when the similarity functions are, respectively, selected as $f_x(t) = -0.5(v+n)\log(1+\frac{t}{v})$ and $f_z(t) = -0.5(v+m)\log(1+\frac{t}{v})$. Hence, the SSMKF can be looked on as a generalized version of the RSTKF. The SSMKF becomes the standard Kalman filter when the similarity functions are, respectively, chosen as $f_x(t) = -0.5t$ and $f_z(t) = -0.5t$. Besides, it is observed from Table 4.1 that the derivatives of the exemplary similarity functions are all -0.5 when the kernel size σ and the dof parameters v and ω tend to infinity, i. e., $\sigma \to +\infty$, $v \to +\infty$ and $\omega \to +\infty$. As a result, the SSMKF based

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

on the exemplary similarity functions all reduce to the standard Kalman filter when the parameters σ, ν and ω tend to infinity.

4.7 Simulation study

4.7.1 Simulation setup and description

We consider a problem of tracking an agile target whose positions are measured in clutter, and the horizontal positions and corresponding velocities are chosen as elements of the state vector. The state transition matrix and measurement matrix are, respectively, $\mathbf{F}_k = \begin{bmatrix} \mathbf{I}_2 & T\mathbf{I}_2 \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix}$ and $\mathbf{H}_k = [\mathbf{I}_2 \ \ \mathbf{0}]$, where $T = 1\,\mathrm{s}$ and $\mathbf{I}_2$ denote the sampling interval and 2-D identity matrix, respectively. The outlier contaminated state and measurement noises are generated according to [19]

$$\begin{cases} \mathbf{w}_k \sim \begin{cases} N(\mathbf{0}, \mathbf{Q}) & \text{w.p.} \quad 0.95, \\ N(\mathbf{0}, 100\mathbf{Q}) & \text{w.p.} \quad 0.05, \end{cases} \\ \mathbf{v}_k \sim \begin{cases} N(\mathbf{0}, \mathbf{R}) & \text{w.p.} \quad 0.95, \\ N(\mathbf{0}, 500\mathbf{R}) & \text{w.p.} \quad 0.05, \end{cases} \end{cases} \tag{4.104}$$

where the nominal state and measurement noise covariance matrices are, respectively, selected as $\mathbf{Q} = \begin{bmatrix} \frac{T^3}{3}\mathbf{I}_2 & \frac{T^2}{2}\mathbf{I}_2 \\ \frac{T^2}{2}\mathbf{I}_2 & T\mathbf{I}_2 \end{bmatrix}$ and $\mathbf{R} = 100\mathbf{I}_2$. The true initial state vector is chosen as $\mathbf{x}_0 = [0, \ 0, \ 10, \ 10]^{\mathrm{T}}$, the initial estimation error variance is set as $\mathbf{P}_0 = \mathrm{diag}([10000, \ 10000, \ 100, \ 100])$ and the initial state estimate is randomly selected from a Gaussian distribution, i. e., $\hat{\mathbf{x}}_{0|0} \sim \mathrm{N}(\mathbf{x}_0, \mathbf{P}_0)$.

As an example, the similarity functions $f_x(\cdot)$ and $f_z(\cdot)$ are, respectively, selected as exponential, logarithmic and square-root functions as in Table 4.1, and the separate iterative algorithm is employed to implement the SSMKF. Then, three outlier-robust Kalman filters can be obtained including SSMKF-exp-S, SSMKF-log-S, and SSMKF-sqrt-S, where SSMKF-exp-S denotes the exponential similarity function and the separate iterative algorithm-based Kalman filter, and the explanations of the other two abbreviations are similar to the SSMKF-exp-S.

The SSMKFs are compared with the standard Kalman filter with true noise covariance matrices (KFTNCM), the HKF [26], the MCKF [8], the RSTKF [18], the IMM filter [6], and the PF [5], where the true noise covariance matrices are used to obtain filtering estimates in the KFTNCM. The tuning parameter of the existing HKF is set as a common value of $\gamma = 1.345$ [26], and the kernel size of the SSMKF-exp-S and the existing MCKF is selected as $\sigma = 10$ to achieve a trade-off between estimation accuracy and stability [8], and the dof parameter of the SSMKF-log-S and the existing RSTKF is

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

set as $\nu = 10$, and the dof parameter of the SSMKF-sqrt-S is set as $\omega = 5$. To guarantee the convergence of the iterations, the iteration threshold and the maximum number of iterations are, respectively, set as $\epsilon = 10^{-16}$ and $N_m = 50$ in all outlier-robust Kalman filters. To guarantee the filtering stability of the SSMKFs, the lower bounds of the auxiliary parameters are set as $\delta = 10^{-8}$. To better show the advantages of the SSMKFs, two IMM filters and three particle filters are performed. In the first IMM filter (IMM-1), the true instantaneous values of state and measurement noise covariance matrices are used, and the four corresponding noise models are, respectively, $\{\mathbf{Q}, \mathbf{R}\}$, $\{\mathbf{Q}, 500\mathbf{R}\}$, $\{100\mathbf{Q}, \mathbf{R}\}$, and $\{100\mathbf{Q}, 500\mathbf{R}\}$, and the model transition probability matrix of the first IMM filter is set as $\mathbf{\Pi}_1$, where $\mathbf{\Pi}_1(i,i) = 0.85$ and $\mathbf{\Pi}_1(i,j) = 0.05(i \neq j)$. In the second IMM filter (IMM-2), the inaccurate instantaneous state and measurement noise covariance matrices are employed, and the four corresponding noise models are, respectively, $\{\mathbf{Q}, \mathbf{R}\}$, $\{\mathbf{Q}, 100\mathbf{R}\}$, $\{1000\mathbf{Q}, \mathbf{R}\}$, and $\{1000\mathbf{Q}, 100\mathbf{R}\}$, and the model transition probability matrix of the first IMM filter is selected as $\mathbf{\Pi}_2$, where all elements of $\mathbf{\Pi}_2$ are 0.25. The initial model probability vectors of the IMM-1 and IMM-2 are both chosen as $[0.25,\ 0.25,\ 0.25,\ 0.25]$. In the first PF (PF-1) and the third PF (PF-3), the true Gaussian mixture PDFs of state and measurement noises given in (4.104) are used, and the particle numbers are, respectively, selected as 1000 and 500 in the PF-1 and PF-3. In the second PF (PF-2), the inaccurate Gaussian mixture PDFs of state and measurement noises are employed, where the used state and measurement noise PDFs are, respectively, $p(\mathbf{w}_k) = 0.98N(\mathbf{w}_k; \mathbf{0}, \mathbf{Q}) + 0.02N(\mathbf{w}_k; \mathbf{0}, 1000\mathbf{Q})$ and $p(\mathbf{v}_k) = 0.98N(\mathbf{v}_k; \mathbf{0}, \mathbf{R}) + 0.02N(\mathbf{v}_k; \mathbf{0}, 100\mathbf{R})$. Note that the IMM-1 and PF-1 are only used as filtering references since the true instantaneous values of state and measurement noise covariance matrices, and the true state and measurement noise PDFs are all unavailable in the presence of random and unknown state and measurement outliers. All filtering algorithms are coded with MATLAB and are executed on a computer with Intel Core i7-6900K CPU @ 3.20 GHz. The MATLAB codes of this chapter will be open access and can be freely downloaded from the link https://www.researchgate.net/profile/Yulong_Huang3.

In this simulation, the simulation time is set as 200 s, and the total number of Monte Carlo runs is selected as 1,000. The root mean square errors (RMSEs) and averaged RMSEs (ARMSEs) of position and velocity are chosen as performance metrics to compare the estimation accuracy, whose definitions are given in the literature [18]. To better exhibit the RMSEs of position and velocity of all filters in Figs. 4.1–4.4, the RMSEs are smoothed using a moving average method with span of 10 s.

4.7.2 Simulation results and comparisons

The RMSEs and ARMSEs (40–200 s) of position and velocity and single-step run time from the SSMKF-exp-S, SSMKF-log-S and SSMKF-sqrt-S and the existing outlier-robust Kalman filters are, respectively, illustrated in Figs. 4.1–4.2 and Table 4.3. It can be seen from Figs. 4.1–4.2 and Table 4.3 that the RMSEs and ARMSEs of position and velocity

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

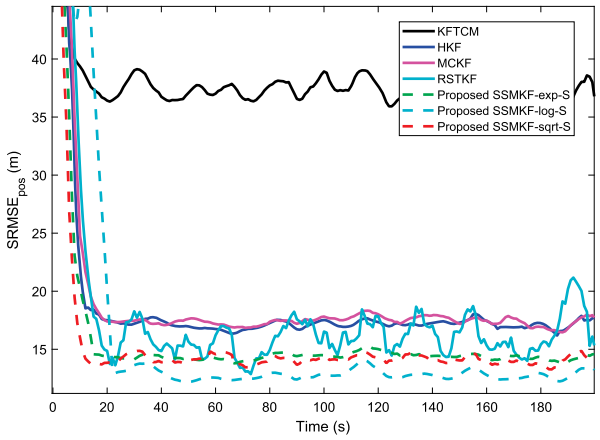


Figure 4.1: RMSEs of positions of the SSMKF-exp-S, SSMKF-log-S and SSMKF-sqrt-S and other outlier-robust Kalman filters.

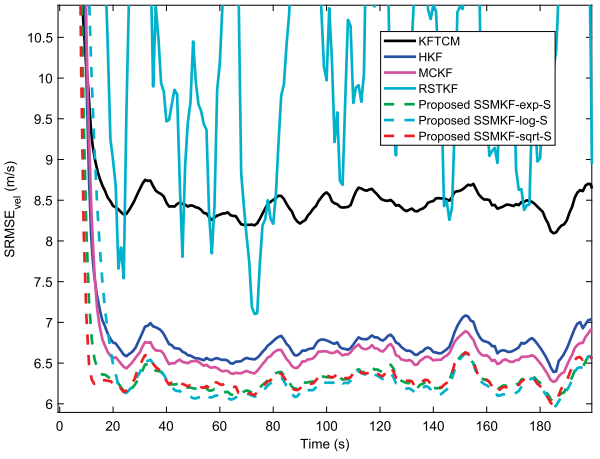


Figure 4.2: RMSEs of velocities of the SSMKF-exp-S, SSMKF-log-S and SSMKF-sqrt-S and other outlier-robust Kalman filters.

from the SSMKF-exp-S, SSMKF-log-S and SSMKF-sqrt-S are all smaller than those from the existing KFTNCM, HKF, MCKF and RSTKF. We can also see from Table 4.3 that the SSMKFs all require more run time than the existing outlier-robust Kalman filters. As compared with the best ARMSE$_{pos}$ from the existing RSTKF and the best ARMSE$_{vel}$ from the existing MCKF, the ARMSEs of position and velocity from the SSMKF-log-S improve 20.91 % and 5.18 %, respectively. Thus, the SSMKF-exp-S, SSMKF-log-S and SSMKF-sqrt-S all have better estimation accuracy but higher computational complexities than the existing KFTNCM, HKF, MCKF and RSTKF.

Figures 4.3–4.4 and Table 4.3, respectively, show the RMSEs and ARMSEs (40 s–200 s) of position and velocity and single-step run time from the SSMKF-exp-S, SSMKF-log-S and SSMKF-sqrt-S and the existing IMM filters and PFs. It is observed from Figs. 4.3–4.4 and Table 4.3 that the SSMKF-exp-S, SSMKF-log-S and SSMKF-sqrt-S all

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4.3: Single-step run time and ARMSEs over 40–200 s.

Filters	$ARMSE_{pos}$ (m)	$ARMSE_{vel}$ (m/s)	Time (ms)
KFTNCM	37.17	8.44	0.019
HKF	17.11	6.71	0.445
MCKF	17.43	6.56	0.465
RSTKF	16.12	10.48	0.505
IMM-1	10.73	5.91	0.4337
IMM-2	17.04	11.53	0.4337
PF-1	15.00	5.80	80.365
PF-2	16.65	8.95	80.365
PF-3	21.54	6.30	40.183
SSMKF-exp-S	14.38	6.28	0.563
SSMKF-log-S	12.75	6.22	1.359
SSMKF-sqrt-S	14.11	6.29	1.220

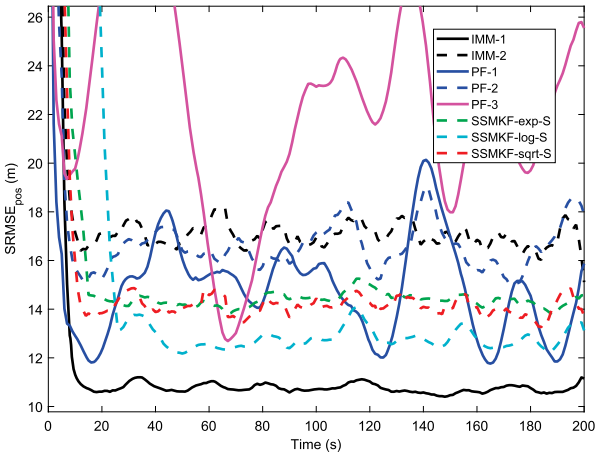


Figure 4.3: RMSEs of positions of the SSMKF-exp-S, SSMKF-log-S and SSMKF-sqrt-S and the IMM filters and PFs.

have smaller RMSEs and ARMSEs of position and velocity than the existing IMM-2 (inaccurate noise models), PF-2 (inaccurate noise PDFs and 1,000 particles), and PF-3 (accurate noise PDFs and 500 particles). As compared with the best $ARMSE_{pos}$ from the PF-2 and the best $ARMSE_{vel}$ from the PF-3, the ARMSEs of position and velocity from the SSMKF-log-S improve 23.42 % and 1.27 %, respectively. It can be also observed from Figs. 4.3–4.4 and Table 4.3 that the RMSEs and ARMSEs of position of the SSMKF-log-S are close to those of the IMM-1 (filtering reference), and the RMSEs and ARMSEs of velocity of the SSMKF-log-S are close to those of the PF-1 (filtering reference), and the filters all have smaller RMSEs and ARMSEs of position than the PF-1. The reason why PF-1 exhibits poor estimation accuracy of position may be that the heavy-tailed features of posterior PDFs are easily lost during the particle-filtering process when a limited number of particles are used. Furthermore, we can observe

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

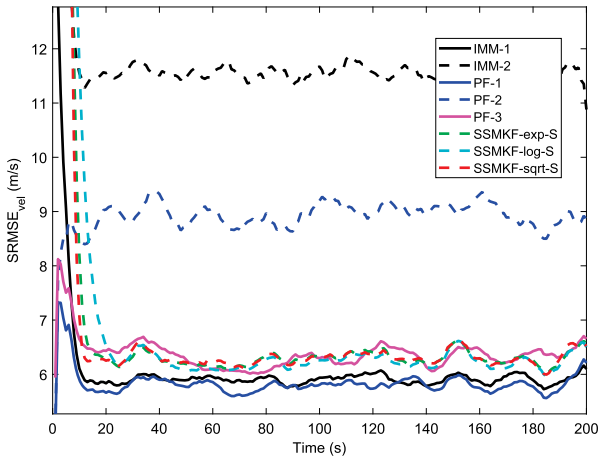


Figure 4.4: RMSEs of velocities of the SSMKF-exp-S, SSMKF-log-S and SSMKF-sqrt-S and the IMM filters and PFs.

from Table 4.3 that the SSMKFs have a slightly longer run time than the IMM filters, but significantly shorter run time than the PFs. Although the computational time of the PF can be significantly reduced if it is implemented in a parallel fashion, it still requires accurate knowledge of the probability distributions of the state and measurement noises. Thus, the SSMKF-exp-S, SSMKF-log-S and SSMKF-sqrt-S all have better estimation accuracy than the IMM-2, PF-2 and PF-3, and slightly higher computational complexities than the IMM filters but significantly lower computational complexities than the standard PFs.

4.8 Conclusions

This chapter introduced the SSM to quantify the similarity between two random vectors. The measure was then employed to develop the SSMKF. Some theoretical analyses and discussions about the approximate errors and the numerical and filtering stabilities were provided to illustrate the effectiveness of the SSMKF. The fixed-point iterative algorithm and the separate iterative algorithm were used to implement the SSMKF, and their local convergence conditions were also provided and compared. In addition, the selections of the similarity functions were presented, and four exemplary similarity functions were provided. Then the relationships between the SSMKF and other existing outlier-robust Kalman filters were revealed. Simulation results illustrated that, by selecting appropriate similarity functions, the SSMKF can achieve improved estimation accuracy but have higher computational complexities than other existing outlier-robust Kalman filters. Also, as compared with the IMM filter and PF, the SSMKF is more suitable for addressing the filtering problem of a linear system with outlier-contaminated state and measurement noises.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bibliography

[1] D. Alspach and H. Sorenson. Nonlinear Bayesian estimation using Gaussian sum approximations. *IEEE Transactions on Automatic Control*, 17(4):439–448, 1972.

[2] B. D. O. Anderson. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, 1979.

[3] I. Arasaratnam and S. Haykin. Cubature Kalman filters. *IEEE Transactions on Automatic Control*, 54(6):1254–1269, 2009.

[4] A. Y. Aravkin, J. V. Burke, and G. Pillonetto. Sparse/robust estimation and Kalman smoothing with nonsmooth log-concave densities: Modeling, computation, and theory. *Journal of Machine Learning Research*, 14:2689–2728, 2013.

[5] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, 2002.

[6] H. A. P. Blom and Y. Bar-Shalom. The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Transactions on Automatic Control*, 33(8):780–783, 1988.

[7] A. E. Bryson and Y. C. Ho. *Applied Optimal Control: Optimization, Estimation, and Control*. CRC Press, 1975.

[8] B. Chen, X. Liu, H. Zhao, and J. C. Principe. Maximum correntropy Kalman filter. *Automatica*, 76(2):70–77, 2017.

[9] B. Chen, L. Xing, H. Zhao, N. Zheng, and J. C. Prıńcipe. Generalized correntropy for robust adaptive filtering. *IEEE Transactions on Signal Processing*, 64(13):3376–3387, 2016.

[10] H. Chou, M. Traonmilin, E. Ollivier, and M. Parent. A simultaneous localization and mapping algorithm based on Kalman filtering. In *IEEE Intelligent Vehicles Symposium, 2004*, volume 1, pages 631–635. IEEE, 2004.

[11] G. T. Cinar and J. C. Príncipe. Hidden state estimation using the correntropy filter with fixed point update and adaptive kernel size. In *The 2012 International Joint Conference on Neural Networks (IJCNN)*, pages 1–6, 2012.

[12] S. Farahmand, G. B. Giannakis, and D. Angelosante. Doubly robust smoothing of dynamical processes via outlier sparsity constraints. *IEEE Transactions on Signal Processing*, 59(10):4529–4543, 2011.

[13] M. A. Gandhi and L. Mili. Robust
Kalman filter based on a generalized maximum-likelihood-type estimator. *IEEE Transactions on Signal Processing*, 58(5):2509–2520, 2010.

[14] H. Gao and X. Li. h_∞ filtering for discrete-time state-delayed systems with finite frequency specifications. *IEEE Transactions on Automatic Control*, 56(12):2935–2941, 2011.

[15] Y. Huang, Y. Zhang, and J. Chambers. A novel Kullback–Leibler divergence minimization-based adaptive student's t-filter. *IEEE Transactions on Signal Processing*, 67(20):5417–5432, 2019.

[16] Y. Huang, Y. Zhang, N. Li, and J. Chambers. Robust student's t based nonlinear filter and smoother. *IEEE Transactions on Aerospace and Electronic Systems*, 52(5):2586–2596, 2016.

[17] Y. Huang, Y. Zhang, N. Li, and J. Chambers. A robust Gaussian approximate fixed-interval smoother for nonlinear systems with heavy-tailed process and measurement noises. *IEEE Signal Processing Letters*, 23:468–472, 2016.

[18] Y. Huang, Y. Zhang, N. Li, Z. Wu, and J. A. Chambers. A novel robust Student's t-based Kalman filter. *IEEE Transactions on Aerospace and Electronic Systems*, 53(3):1545–1554, 2017.

[19] Y. Huang, Y. Zhang, P. Shi, Z. Wu, J. Qian, and J. A. Chambers. Robust Kalman filters based on Gaussian scale mixture distributions with application to target tracking. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(10):2082–2096, 2020.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[20] Y. Huang, Y. Zhang, Z. Wu, N. Li, and J. Chambers. A novel adaptive Kalman filter with inaccurate process and measurement noise covariance matrices. *IEEE Transactions on Automatic Control*, 63(2):594–601, 2018.
[21] Y. Huang, Y. Zhang, B. Xu, Z. Wu, and J. Chambers. A new outlier-robust student's t based Gaussian approximate filter for cooperative localization. *IEEE/ASME Transactions on Mechatronics*, 22(5):2380–2386, 2017.
[22] Y. Huang, Y. Zhang, Y. Zhao, P. Shi, and J. Chambers. A novel outlier-robust Kalman filtering framework based on statistical similarity measure. *IEEE Transactions on Automatic Control*, 66(6):2677–2692, 2021.
[23] P. J. Huber. *Robust Statistics*. Springer, Berlin, Heidelberg, 2011.
[24] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45(3):477–482, 2000.
[25] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82D:35–45, 1960.
[26] C. D. Karlgaard and H. Schaub. Huber-based divided difference filtering. *Journal of Guidance, Control, and Dynamics*, 30(3):885–891, 2007.
[27] S. Li, H. Wang, and T. Chai. A t-distribution based particle filter for target tracking. In *2006 American Control Conference*, pages 2191–2196, 2006.
[28] X. Li and H. Gao. Robust finite frequency h_∞ filtering for uncertain 2-D Roesser systems. *Automatica*, 48(6):1163–1170, 2012.
[29] H. Ohlsson, F. Gustafsson, L. Ljung, and S. Boyd. Smoothed state estimates under abrupt changes using sum-of-norms regularization. *Automatica*, 48(4):595–605, 2012.
[30] R. Piché, S. Särkkä, and J. Hartikainen. Recursive outlier-robust filtering and smoothing for nonlinear systems using the multivariate student-t distribution. In *2012 IEEE International Workshop on Machine Learning for Signal Processing*, pages 1–6, 2012.
[31] M. Roth, T. Ardeshiri, E. Özkan, and F. Gustafsson. Robust Bayesian filtering and smoothing using Student's t distribution. arXiv preprint, arXiv:1703.02428.
[32] M. Roth and F. Gustafsson. An efficient implementation of the second order extended Kalman filter. In *14th International Conference on Information Fusion*, volume 1, pages 1–6. IEEE, 2011.
[33] M. Roth, E. Özkan, and F. Gustafsson. A Student's t filter for heavy tailed process and measurement noise. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 5770–5774, 2013.
[34] D. Shi, T. Chen, and L. Shi. Event-triggered maximum likelihood state estimation. *Automatica*, 50(1):247–254, 2014.
[35] D. Simon. *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. John Wiley & Sons.Inc, 2006.
[36] G.Tambini, G. C. Montanari, and M. Cacciari. The Kalman filter as a way to estimate the life-model parameters of insulating materials and system. In *2002 8th International Conference on Conduction & Breakdown in Solid Dielectrics*, volume 1, pages 523–527. IEEE, 1992.
[37] J. Ting, E. Theodorou, and S. Schaal. Learning an outlier-robust Kalman filter. In *18th European Conference on Machine Learning Warsaw*, 2007.
[38] D. Unsal and M Dogan. Implementation of identification system for imus based on Kalman filtering. In *2002 8th International Conference on Conduction & Breakdown in Solid Dielectrics*, volume 1, pages 236–240. IEEE, 2014.
[39] Y. Wang, W. Zheng, S. Sun, and L. Li. Robust information filter based on maximum correntropy criterion. *Journal of Guidance, Control, and Dynamics*, 39(5):1126–1131, 2016.
[40] L. Xie, Ced Souza, and M. Fu. h_∞ estimation for discrete-time linear uncertain systems. *International Journal of Robust & Nonlinear Control*, 1:111–123, 2010.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[41] K. Xiong, H. Zhang, and C. W. Chan. Performance evaluation of UKF-based nonlinear filtering. *Automatica*, 42(2):261–270, 2006.
[42] A. Yadav, N. Naik, M. R. Ananthasayanam, A. Gaur, and Y. N. Singh. A constant gain Kalman filter approach to target tracking in wireless sensor networks. In *2012 8th Industrial and Information Systems (ICIIS), 2012 7th IEEE International Conference*, volume 1, pages 1–7. IEEE, 2012.
[43] S.Zhao, Y. Ma, and B. Huang. Robust fir state estimation of dynamic processes corrupted by outliers. *IEEE Transactions on Industrial Informatics*, 15(1):139–147, 2019.
[44] H. Zhu, H. Leung, and Z. He. A variational Bayesian approach to robust sensor fusion based on student-t distribution. *Information Science*, 221(2013):201–214, 2012.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Paweł Oświęcimka and Ludovico Minati

5 Multifractal characteristics of singular signals

Abstract: Multifractal formalism is one of the most potent tools for characterizing the singular behavior of signals encountered across various scientific and engineering situations. This conceptually advanced methodology is made accessible to experimentalists via the development of algorithms capable of determining a time-serie's local scaling properties through a set of scaling exponents and an associated singularity spectrum. By determining a signal's multifractal properties, the temporal organization of the underlying various amplitude fluctuations can be quantitatively described within a unique scheme of the correlation structure. In this work, we demonstrate that the degree of complexity of a signal and its hierarchical organization's character are reflected in the shape of the singularity spectrum. A stark example of this is offered by financial time series, wherein well-developed multifractal spectra quantify the hierarchical structure of the data and the heterogeneity of singular behavior. On the other hand, for signals without a cascade-like organization of the singularities, the singularity spectrum's interpretation must be undertaken with extreme care. In such cases, artifactual singularities can be produced by processes that are not firmly interrelated. Thus, the fluctuation structure does not truly reveal the hierarchy of the organization. We show examples of such time series produced by nonlinear dynamical systems, particularly the Saito chaos generator, within which twofold dynamics, one related to periodic component and the other related to hysteresis, are not hierarchically nested. Consequently, a local scaling analysis based on the wavelet transform must be adequately applied to identify the structure of isolated singularities.

Keywords: multifractal analysis, Hölder exponent, singularity, complexity, dynamical system, time-series analysis

5.1 Introduction

Multifractality is a concept that has found its application across many different areas of science [2, 6, 10, 12, 13, 24, 26]. Within the framework of this methodology, the set of

Paweł Oświęcimka, Complex Systems Theory Department, Institute of Nuclear Physics Polish Academy of Sciences, 31-342 Kraków, Poland; and Institute of Theoretical Physics, Jagiellonian University, Kraków, 30-348, Poland, e-mail: pawel.oswiecimka@ifj.edu.pl
Ludovico Minati, Center for Mind/Brain Sciences (CIMeC), University of Trento, 38122 Trento, Italy; and Institute of Innovative Research, Tokyo Institute of Technology, 226-8503 Yokohama, Japan; and Institute of Theoretical Physics, Jagiellonian University, Kraków, 30-348, Poland, e-mail: lminati@ieee.org

https://doi.org/10.1515/9783110729122-005

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

scaling exponents is used to describe the stochastic processes characterized by multiscaling properties. Hence, multifractal formalism is especially useful when singular measures or functions, numerously represented in experimental data, are considered. By decomposing the data's correlation structure, the subsets related to various signal amplitudes are quantitatively described within the uniform methodology. This is particularly important when temporal data organization is dependent on the data amplitude, as in the multifractal case.

The fundamental concepts of multifractal formalism are the singularity exponent and the singularity spectrum. The local regularity of a function can be estimated via the following relationship:

$$|f(x_0 + \delta) - f(x_0)| \sim C\delta^{\alpha(x_0)}, \tag{5.1}$$

where $\alpha(x_0)$, known as the Hölder exponent, determines the singularity strength around x_0; the more singular the function, the lower the Hölder exponent α. The statistical distribution of α refers to the Hausdorff dimension of the data support with a particular α and is defined as follows:

$$f(\alpha) = d_f(\{x_0, \alpha(x_0) = \alpha\}). \tag{5.2}$$

The shape of $f(\alpha)$ resembles an inverted parabola with a maximum corresponding to the most well-represented singularity in the data. To assess the singularity spectrum of a time-series, two common approaches were proposed, i. e., multifractal detrended fluctuation analysis, MFDFA [16], and wavelet transform modulus maxima, WTMM [18]. Within the framework of these algorithms, the q-filtering technique is used to estimate a partition function, and, on this basis, the singularity spectrum. However, when isolated singularities are present, these techniques have to be used carefully. This is due to the limitation of these methods in forming a spectrum consisting of isolated points. Thus, considering singular but not fractal functions through the MFDFA and WTMM algorithms, we obtain broad spectra mistakenly suggesting a complex and hierarchical structure [21].

In this contribution, we demonstrate the results of the analysis of some systems considered as multifractals, and of signals being 'only' singular but not fractally organized. In the former case, we consider synthetic multifractals, i. e., the binomial cascade and time series generated by the Ikeda chaotic system. Then, we present the analysis of representative financial data: logarithmic returns of the S&P 500 index, and time intervals between consecutive transactions for Bayer (BAY) company listed on the German stock market DAX. Finally, we present an analysis of the chaotic signal generated by an electronic realization of the Saito system, an example of a signal with isolated singularities.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

5.2 Methodology

5.2.1 Multifractal detrended fluctuation analysis

The possibly multiscale temporal organization of a time series may be characterized employing multifractal detrended fluctuation analysis (MFDFA) [16]. This method is beneficial for complex, nonstationary time-series when more than a single scaling exponent is needed to characterize the analyzed system sufficiently. The algorithm consists of the following steps. Given a time series x_i of length N, $i = 1, 2, \ldots, N$, the profile $Y(t)$ is calculated according to the formula:

$$Y(j) = \sum_{i=1}^{j} [x_i - \langle x \rangle], \tag{5.3}$$

where $\langle\ \rangle$ denotes the average over the entire time series. Due to the possible fractal nature of the data, the profile has to be analyzed on various timescales s. Thus, $Y(t)$ is divided into N_s non-overlapping segments ν of length s ($N_s = \lfloor \mathrm{int}(N/s) \rfloor$). Moreover, the length of the time series is not necessarily a multiple of the scale s. Thus, to avoid neglecting data from the analysis, the division is performed starting both from the beginning and from the end of the time series. Finally, the $2N_s$ segments are considered in the analysis. The next step is to remove the possible trend from the data. Therefore, for each segment, the assumed trend, usually a polynomial $P_\nu^{(m)}$ of order m, is fitted and then subtracted from the data. The potential of this procedure to eradicate the trend is dependent on the degree of the polynomial m [22]. In many numerical and empirical experiments, the reasonable choice is $m = 2$, which is also used in the study presented in this chapter. In each segment, the detrended variance is calculated:

$$F^2(\nu, s) = \frac{1}{s} \sum_{k=1}^{s} \left(X\big((\nu - 1)s + k\big) - P_\nu^{(m)}(k) \right)^2. \tag{5.4}$$

To assess the scaling properties of the data with respect to the size of their amplitude, the q-order filtering technique is applied. The fluctuations are filtered according to the q, which amplifies segments with large ($q > 0$) or small fluctuations ($q < 0$). On this basis, the q-order fluctuation functions are calculated with this equation:

$$F_q(s) = \left\{ \frac{1}{2N_s} \sum_{\nu=1}^{2N_s} \left[F^2(\nu, s) \right]^{q/2} \right\}^{1/q}, \quad q \in \mathbb{R} \setminus \{0\}. \tag{5.5}$$

If a time series exhibits fractal patterns, the power-law behavior is observed as

$$F_q(s) \sim s^{h(q)}, \tag{5.6}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

where $h(q)$ denotes the generalized Hurst exponents. A time series that exhibits monofractal scaling behavior can be described by a single scaling exponent $h(q) = H$ equal to the Hurst exponent [8, 11]. The value of H indicates the character and strength of the linear correlation in the time series. Here, $0.5 < H \leq 1$ means persistent data that are characterized by long-range positive autocorrelations. Data with $0 \leq H < 0.5$ exhibits antipersistent behavior, i. e., negative autocorrelations. For $H = 0.5$, only short-range temporal dependencies are present. However, in the case of multifractal behavior, the set of scaling exponents is needed to characterize the time series's temporal organization. Thus, $h(q)$ is a concave function of q, and the Hurst exponent is retrieved for $h(q = 2)$. According to the fractal formalism, the generalized Hurst exponents are recalculated on the singularity spectrum using the formula

$$\alpha_G = h(q) + qh'(q), \quad f(\alpha_G) = q[\alpha_G - h(q)] + 1, \tag{5.7}$$

where α_G is a local version of the Hurst exponent called the Hölder exponent, and $f(\alpha_G)$ refers to the fractal dimension of the data support with a particular α_G. Moreover, we use subscript G to emphasize that the presented methodology uses a "global" measure (5.5) to assess the set of Hölder exponents [1]. The complexity of the data can be quantified by the width of the singularity spectrum $\Delta\alpha_G$. Thus, the more complex a time series is (the more developed multifractality), the wider the multifractal spectrum, and vice versa. Information about the hierarchical organization of the data is also provided by the degree of spectrum asymmetry, which can be assessed by the equation [5]

$$A_\alpha = (\Delta\alpha_{GL} - \Delta\alpha_{GR})/(\Delta\alpha_{GL} + \Delta\alpha_{GR}), \tag{5.8}$$

where $\Delta\alpha_{GL}$ and $\Delta\alpha_{GR}$ denote the distance of the spectrum maximum to the smallest and the largest values of α_G, respectively. Thus, positive values of A_α are an attribute of the left-sided asymmetry of the spectrum, reflecting a much more complex organization of the large fluctuations than the smaller ones. On the contrary, in the case of right-sided asymmetry, one has $A_\alpha < 0$, and the support of the multifractality consists mainly of small fluctuations.

5.2.2 Wavelet leader multifractal analysis

One of the most powerful tools for unveiling the data's hierarchical structure and fractal organization is the wavelet transform (WT) [15, 18]. Within this methodology, employing the elementary function $\psi_{a,s}(x) = s^{-1/2}\psi(\frac{x-a}{s})$, the so-called mother wavelet, the analyzed signal is decomposed in a scale-space(time) half-plane. s is the scale parameter (frequency band), and a is the wavelet's space (time) position ($a, s \in \mathbb{R}, s > 0$).

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

The wavelet transform of the function $f(x)$ is defined in the following way [1]:

$$Wf(a,s) = \frac{1}{s^{-1/2}} \int_{-\infty}^{\infty} f(x)\psi\left(\frac{x-a}{s}\right)\mathrm{d}x. \tag{5.9}$$

Singular behavior of the function f can be characterized through the scaling behavior of $Wf(a,s)$. Thus, the strength of singularity (Hölder exponent α) around the point x_0 is estimated by the expression

$$Wf(x_0,s) \sim s^{\alpha_L(x_0)}, \quad s \to 0^+, \tag{5.10}$$

where subscript L denotes that the algorithm uses only local measures to estimate the Hölder exponent. These properties make the wavelet formalism ideal in estimating the singularity spectrum $f(\alpha)$ for fractal functions. To improve the stability of the calculations, which can distort the results of the analysis for the most complex signals, the wavelet leader approach was introduced [14, 15]. In this algorithm, the wavelet coefficients $c_{j,k}$ for a function f are calculated as [28]

$$c_{j,k} = 2^{-j} \int_{\mathbb{R}} f(x)\psi(2^{-j}x - k)\mathrm{d}x, \tag{5.11}$$

where $s_j = 2^{-j}$ and $a_{j,k} = 2^{-jk}$ $(j,k \in Z)$ denote the discrete scale and time parameters, respectively. The wavelet leader $L_j(x_0)$ in x_0 and level j is defined as the largest wavelet coefficient in spatial the neighborhood of x_0 at finer scales

$$L_j(x_0) = \sup_{\lambda' \subset 3\lambda_{j,k}(x_0)} \left|c_{j,k}(\lambda')\right|, \tag{5.12}$$

where only dyadic intervals $\lambda_{j,k} = [2^j k, 2^j(k+1)]$ are considered, and $3\lambda_{j,k}(x_0) = \lambda_{j,k-1} \cup \lambda_{j,k} \cup \lambda_{j,k+1} = [2^j(k-1, 2^j(k+2))]$ contains x_0. Based on this definition, the structure function $S(q,j)$ can be defined as

$$S(q,j) = 2^j \sum_{\lambda \in \Lambda_j} L_j^q, \tag{5.13}$$

where q is a real number and Λ_j is a set of dyadic intervals at scale j. For fractal time-series, power-law behavior of the structure functions is expected:

$$S(q,j) \approx C_q 2^{j\zeta(q)}, \quad (2^j \to 0). \tag{5.14}$$

Hence, the estimated scaling exponents $\zeta(q)$ are directly used to calculate the multifractal spectrum according to the formula

$$f(\alpha_G) = \inf_{q \in R}\left(q\alpha_G - \zeta(q)\right) + 1. \tag{5.15}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

5.3 Multifractal synthetic processes

5.3.1 Binomial cascade

The potential of multifractal methodology can be demonstrated when non-homogeneous measures, i. e., structures whose singularity spectrum $f(\alpha)$ is supported by more than a single point, are analyzed [18]. The instructive example in this respect is a binomial cascade, some variants of which can be regarded as a generalization of the Cantor set construction. The procedure of generation of the cascade is defined in the following way [17, 20]: The initial segment with a homogeneously distributed measure $[0, 1]$ is divided into two parts, and an a and $(1 - a)$ portion of the initial measure is assigned to the left and right subinterval, respectively. In the following steps, this procedure is repeated for each subinterval independently, *ad infinitum*. In practical application, the number of the cascade levels k is finite, and the size of each subinterval j is 2^{-k}. The measure M_j at the final stage is a product of k multipliers a_j^i according to the equation [17]

$$M_j = \prod_{i=1}^{k} a_j^i, \tag{5.16}$$

where i is the particular level of the cascade. In Fig. 5.1a, the generated cascade is depicted. It is clear that the measure is not homogeneously distributed for the presented data, and thus a different degree of singularities α can be identified. The hierarchical structure of the singularities visualized utilizing the wavelet transform is presented in Fig. 5.1b. The maxima of the wavelet coefficients form a tree with branches containing information about the hierarchical construction. For this deterministic structure, the singularity spectrum is derived straightforwardly as [16]:

$$\alpha = -\frac{1}{\log(2)} \frac{a^q \log(a) + (1-a)^q \log(1-a)}{a^q + (1-a)^q}, \tag{5.17}$$

$$f(\alpha) = -\frac{q}{\log(2)} \frac{a^q \log(a) + (1-a)^q \log(1-a)}{a^q + (1-a)^q} - \frac{-\log(a^q + (1-a)^q)}{\log(2)}, \tag{5.18}$$

where $q \in \mathbb{R}$. The estimated singularity spectrum, and also its theoretical counterparts, are depicted in Fig. 5.1c. The width of the spectrum $\Delta\alpha_G \approx 0.8$ unequivocally indicates that the analyzed structure is multifractal. Moreover, the symmetry of the spectrum implies that strong and weak singularities contribute equally to the multifractal organization of the data. This result is confirmed by local analysis of the singularities. Through Eq. 5.10, one can detect singular behaviors and assess their strength. The outcome of this procedure applied to the binomial cascade is presented in Fig. 5.1d. The continuous distribution of the estimated Hölder exponent and its symmetry confirm a well-developed multifractality and undiversified distribution of the correlations over fluctuations having different amplitudes.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

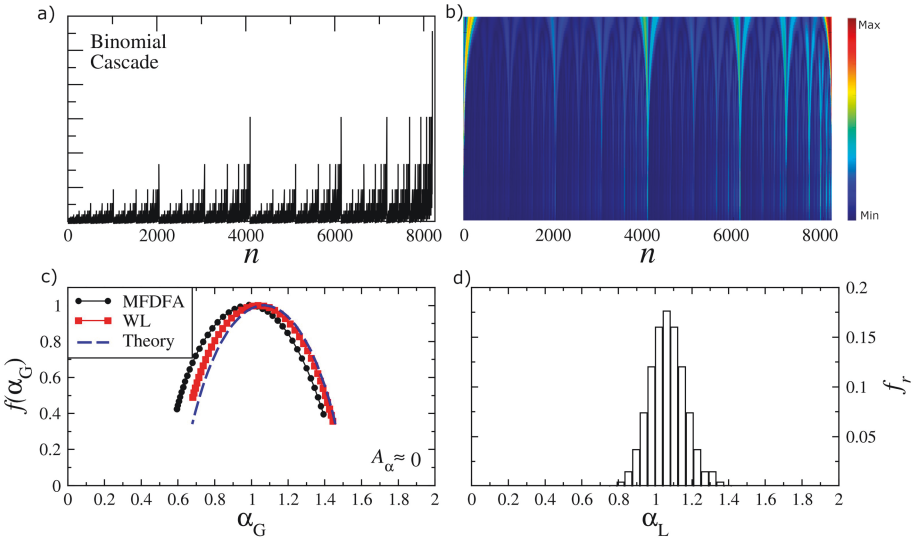


Figure 5.1: Multifractal analysis of the binomial cascade. a) Graph of the binomial cascade generated with $a = 0.65$. b) Wavelet transform of the data. c) Multifractal spectrum of the cascade. d) Relative frequency f_r histograms of the Hölder exponents.

5.3.2 Ikeda chaotic system

Another example of a synthetic system that can reproduce multifractal organization is the chaotic metronome derived from the Ikeda map [27]. Multifractal analysis of this dataset is shown in Fig. 5.2. For the binomial cascade, the estimated singularity spectrum is broad ($\Delta\alpha_G \approx 0.6$), revealing the heterogeneous distribution of the measure and nonlinear dependencies among the data (Fig. 5.2b). However, in contrast to the binomial cascade, the spectrum's shape is markedly asymmetrical, with the left wing more developed than the right one. This implies that the primary support of multifractality is related to large fluctuations, which generate a more complex temporal organization than the small ones. Analysis of the local scaling properties confirms this conclusion (Fig. 5.2c). The histogram of Hölder exponents presents a continuous shape, hallmarking a "rich" multifractality. Moreover, left-sided skewness of the distribution is visible. It is worth noting that distortions in the hierarchical organization of the correlations expressed by the skewness of the multifractal spectrum are often observed in experimental data, and their examples are presented in the next section.

5.4 Multifractality of singular financial data

Economics data are an example of time series with extremely complex temporal organization [7]. On the one hand, the distribution is leptokurtic, with power-law tails (fat

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

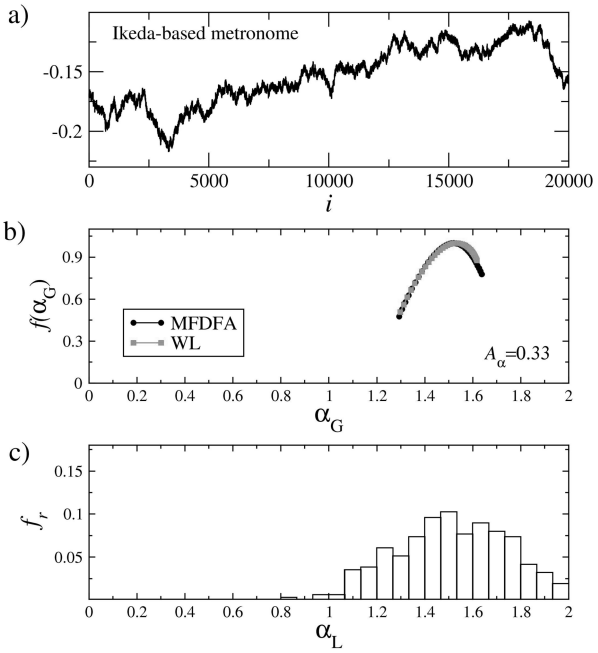


Figure 5.2: Analysis of the Ikeda map. a) Graph of the Ikeda time series. b) Multifractal spectra of the data. c) Relative frequency f_r histograms of the Hölder exponents.

tails) characterized by a scaling exponent ≈ 3 [3, 9]. Hence, large fluctuations are more likely compared to a Gaussian distribution. On the other hand, the data display nonlinear dependencies with a multiscale hierarchy of correlations visible to the multifractal formalism [4, 23]. Thus, a set of various singularities, even those related to sudden trend reversal (crashes), is determined by the unique multifractal theory framework. In Fig. 5.3 we present an example of the multifractal analysis of financial time series. On the left, the analysis of logarithmic returns ($r(t) = \log(P(t)) - \log(P(t - \Delta t))$), where $P(t)$ denotes the index value in time t) of the American index S&P 500 is presented. On the right, we show the study of the time intervals ($\Delta t_i = t_{i+1} - t_i$) between consecutive transactions i of Bayer (BAY) company listed on the German stock market DAX. In both cases, the estimated spectra are broad, which confirms multifractality. However, the asymmetry of the spectrum is dependent on the dataset. For the S&P 500 index, the spectrum demonstrates left-sided asymmetry ($A_\alpha = 0.69$), indicating that the large fluctuations' complex organization is the primary support of the multifractal behavior. In contrast, the small ones reveal a much poorer organization [5]. The opposite conclusions can be drawn when results for waiting times Δt_i of Bayer company are analyzed. The right-sided asymmetry of the spectrum ($A_\alpha = -0.3$) suggests a dominant role of small fluctuations in forming a multifractal data organization. The large ones are arranged more homogeneously, with singularities of a similar degree.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

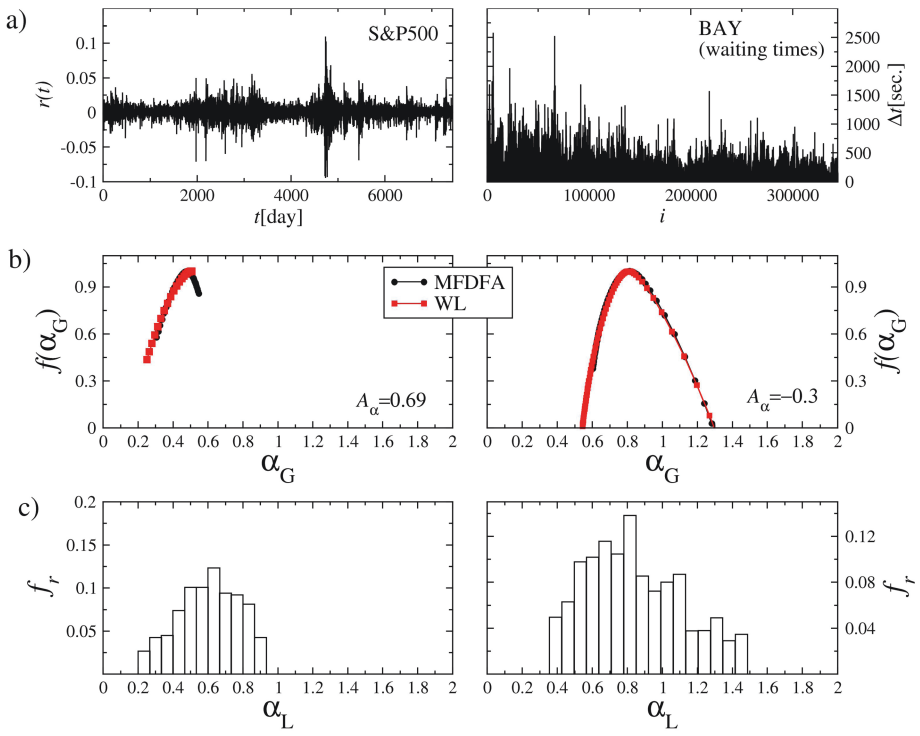


Figure 5.3: Analysis of financial time-series. a) Logarithmic returns of S&P 500 index and time intervals between consecutive transactions for Bayer (BAY) company listed on the German stock market. b) Multifractal spectra of the corresponding data. c) Relative frequency f_r histograms of the Hölder exponents.

These results are confirmed by an analysis of the local scaling properties (Fig. 5.3c). The distribution of the Hölder exponents α_L resembles the broad multifractal spectrum $f(\alpha_G)$ with left- and right-skewness for S&P 500 and BAY data, respectively. This implies that the singularity spectra identified through MFDFA and WL methodologies are a genuine signature of data multifractality. As demonstrated in the next section, distinguishing between true and artifactual multifractality may be problematic and has to be done with sufficient care.

5.5 Apparent multifractality of the Saito system

The Saito chaos generator is a four-dimensional nonlinear oscillator defined by the following state equations [25]:

$$\dot{x} = -z - w, \tag{5.19}$$

$$\dot{y} = y(2\delta y + z), \tag{5.20}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$\dot{z} = \rho(x - y), \tag{5.21}$$

$$\dot{w} = \frac{1}{\epsilon}(x - h(w)), \tag{5.22}$$

wherein

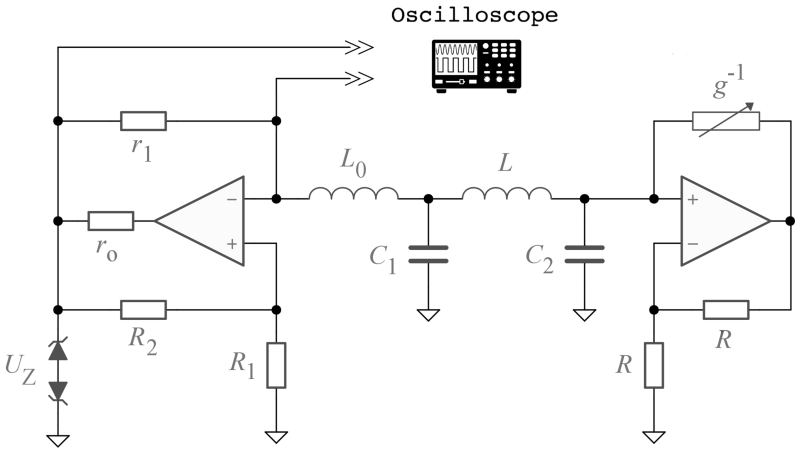
$$h(w) = \begin{cases} w - (1 + \eta) & \text{if } w \geq \eta, \\ -\eta^{-1} w & \text{if } |w| < \eta, \\ w + (1 + \eta) & \text{if } w \leq -\eta. \end{cases} \tag{5.23}$$

The dynamics of the system, spanning periodicity, quasi-periodicity, and several types of chaotic behavior, is determined by the control parameters γ, δ, ϵ, ν and η. From the perspective of singular behavior, the w variable is the most interesting one. The presence of the hysteresis function $h(w)$ (and $\epsilon \to 0$) results in slow and fast motions corresponding to a slow manifold alongside sudden jumps. In comparison with w, the behavior of the x, y and z variables is considerably less volatile and smoother [21].

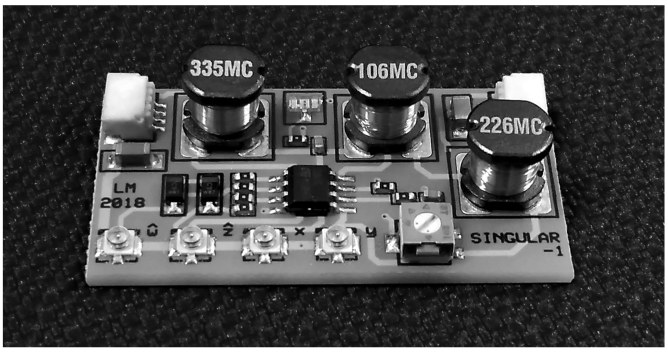
The system can be built experimentally. Results are shown from a circuit including two operational amplifiers (type TL082) and a nonlinearity made with two opposing Zener diodes (type BZT52-C5V1). The corresponding diagram is visible in Fig. 5.4a, where $r_1 = r_2 = R_1 = R_2 = R = 10\,\text{k}\Omega$, $r_o = 820\ \Omega$, $C_1 = C_2 = 3.9\,\text{nF}$, $L_0 = 3.3\,\text{mH}$, $L = 32\,\text{mH}$ (two inductors in series) and $U_Z = 5.1\,\text{V}$. For these component values, one has $\gamma = C_1/C_2 = 1$, $\varepsilon = L_0/(r_1^2 C_1) = 0.0085$, $\eta = r_1/r_2 = 1$ and $\rho = r_1^2 C_1/L = 12.2$. The signal from variable w was recorded from the physical circuit board (Fig. 5.4b) at 1 MSa/s, tuning g^{-1} to obtain data with different δ.

As already stated, the dynamics of the system strongly depend on the control parameters. This chapter presents results from the analysis of time series from quasiperiodic and hyperchaos regimes. The corresponding records for the w variable are depicted in Fig. 5.5a. Even a cursory inspection of the data shows that the data organization is heterogeneous, with sharp upward and downward jumps in both cases. Thus, it is anticipated that the multifractal properties of the time series may be strongly influenced by the presence of strong singularities in the data. In Fig. 5.5b, the corresponding singularity spectra are shown. The spectra are broad ($\Delta\alpha = 2.5$) and asymmetrical ($A_\alpha \approx 0.3$), suggesting a multifractal organization of the data and distortions in a hierarchy of correlations. However, closer inspection of their shape reveals more intriguing details. The distribution of the points along the spectrum concentrates mainly towards its ends, similarly to the characteristics of a bifractal structure [19]. Moreover, the analysis of the local scaling properties discloses more subtleties of the data. The histogram of the local Hölder exponents forms two separable peaks whose locations coincide with the ends of the multifractal spectrum. The position in time of these singularities can be determined through the wavelet transform. In Fig. 5.6, the skeleton

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use



a)



b)

Figure 5.4: Experimental implementation of the Saito chaos generator. a) Circuit diagram, and b) representative example of physical realization. Reproduced from [21].

of the multiscale organization of the data is presented. The maxima of the wavelet transform delineate separable lines related to individual singularities. Such a structure indicates nonfractal data organization, with two kinds of the isolated singularities, in contrast to the true multifractality where a bifurcation-like structure is visible on the time-scale wavelet transform map. The observed broad singularity spectrum obtained using the MFDFA and WL methodology is an effect—in fact, an artifact of the q-filtering method used by these methods to decompose the data concerning the amplitude size. When distinct isolated singularities are presented in the data, the filtering technique produces a smooth singularity spectrum, and the wider is the spectrum, the more difference between the singularity degree. Thus, in this case, the width of the spectrum is not a measure of data complexity but of the diversity of the isolated singularities.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

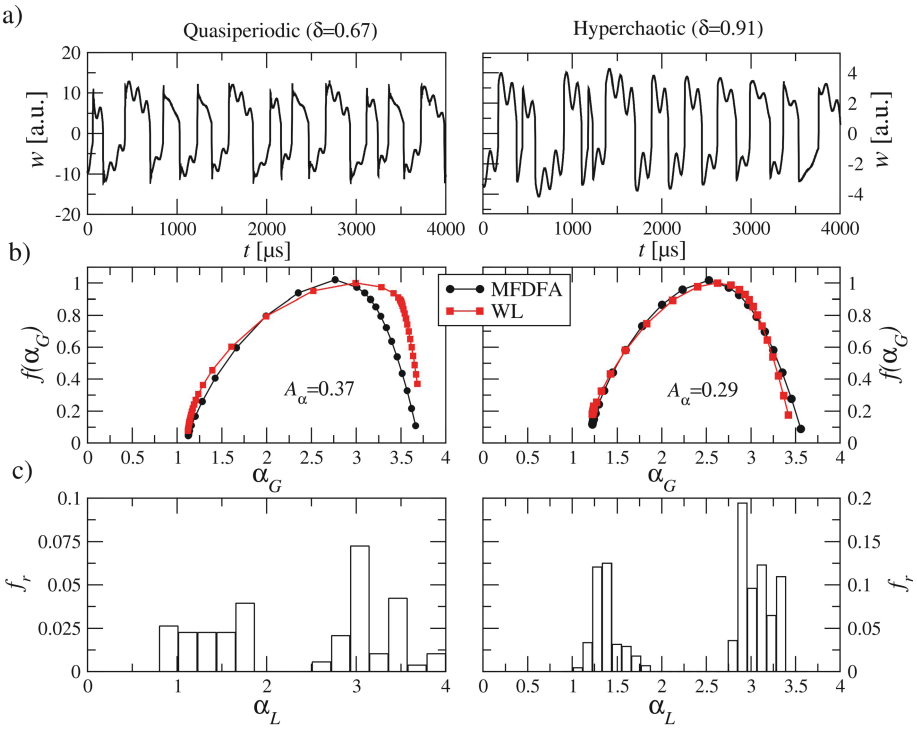


Figure 5.5: a) Experimental time series of variable *w* recorded from the physical Saito chaos generator in the hyperchaotic and quasiperiodic regime. b) Corresponding multifractal spectra and c) relative frequency histograms for the Hölder exponents. Adapted from [21].

5.6 Summary

Development of multifractal methodology creates an opportunity to quantitatively describe the organization of complex structures within a unique and concise theory. Thus, it is not surprising that this interdisciplinary concept has found application across diverse areas of science. It is especially applicable to structures that are characterized by strongly singular behavior. An example of such a process is a mathematically constructed binomial cascade, for which a set of singularities of different degree is determined by a singularity spectrum. Such singular behavior is ubiquitous in experimental signals, and one of the stark examples in this respect is financial data. Their complex and hierarchical organization has been identified by means of multifractal algorithms and constitutes one of the stylized financial facts. However, the results of multifractal analysis have to be interpreted cautiously. This is particularly important when isolated singularities can be present in the data. In this respect, the Saito chaos generator, for which, apart from the complicated data organization and singular signal attributes, the singularities are not hierarchically nested, is a good ex-

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

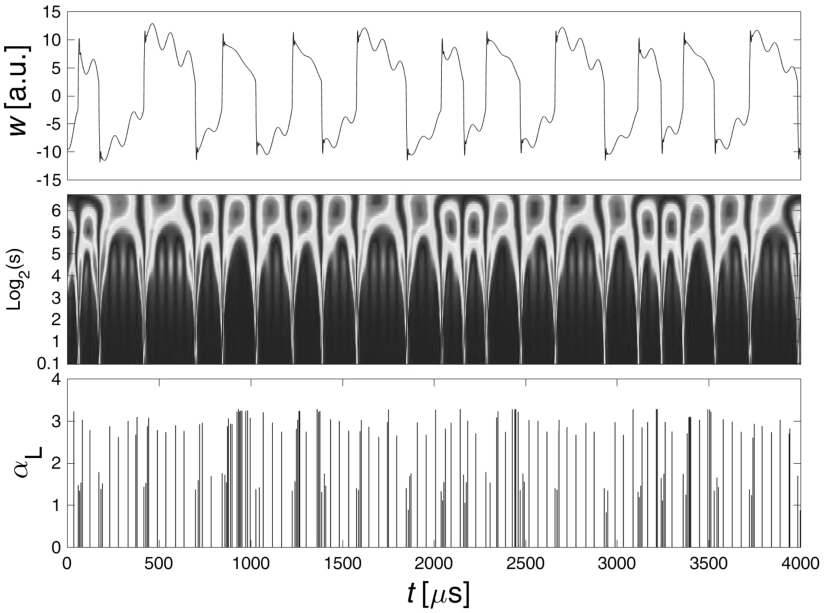


Figure 5.6: Local scaling properties of the *w* time series from the Saito chaos generator, simulated in the hyperchaotic regimes. Top: time courses. Middle: wavelet transform obtained via the fourth derivative of the Gaussian wavelet. Color coding denotes the magnitude of the wavelet coefficients $Wf(x_0, s)$ ranging from dark blue (the smallest $Wf(x_0, s)$) to red (the largest one). Bottom: corresponding time localized Hölder exponents.

ample. While it does not lead to multifractal data organization, the algorithms based on the *q*-filtering technique, MFDFA and WL, produce a broad singularity spectrum. A cursory examination of the data may thus lead to overestimating structural complexity. However, the data's true nature can be uncovered by analyzing the wavelet transform map where the lack of hierarchical organization is clearly detectable.

Bibliography

[1] A. Arneodo, E. Bacry, and J. F Muzy. Oscillating singularities in locally self-similar functions. *Physical Review Letters*, 74:4823–4826, 1995.

[2] M. Ausloos and K. Ivanova. Multifractal nature of stock exchange prices. *Computer Physics Communications*, 147:582–585, 2002.

[3] S. Drożdż, M. Forczek, J. Kwapień, P. Oświęcimka, and R. Rak. Stock market return distributions: from past to present. *Physica A*, 383:59-–64, 2007.

[4] S. Drożdż, R. Kowalski, P. Oświęcimka, R. Rak, and R Gębarowski. Dynamical variety of shapes in financial multifractality. *Complexity*, 2018:7015721, 2018.

[5] S. Drożdż and P. Oświęcimka. Detecting and interpreting distortions in hierarchical organization of complex time series. *Physical Review E*, 91:030902(R), 2015.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[6] S. Drożdż, P. Oświęcimka, A Kulig, J. Kwapień, K. Bazarnik, I. Grabska-Gradzińka, J Rybicki, and M. Stanuszek. Quantifying origin and character of long-range correlations in narrative texts. *Information Sciences*, 331:32–44, 2016.
[7] Z. Eisler and J. Kertész. Size matters: some stylized facts of the stock market revisited. *The European Physical Journal B*, 51:145–-154, 2006.
[8] J. Feder. *Fractals*. Plenum Press, New York, 1998.
[9] P. Gopikrishnan, M. Meyer, L. A. N Amaral, and H. E. Stanley. Inverse cubic law for the distribution of stock price variations. *The European Physical Journal B*, 3:139–-140, 1998.
[10] D. Grech and G. Pamuła. Multifractality of nonlinear transformations with application in finances. *Acta Physica Polonica A*, 123:529–537, 2013.
[11] H. E. Hurst. Long-term storage capacity of reservoirs. *Transactions of American Society of Civil Engineers*, 116:770, 1951.
[12] P. Ch. Ivanov, L. A. N. Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik, and H. E. Stanley. Multifractality in human heartbeat dynamics. *Nature*, 399:461–465, 1999.
[13] G. R. Jafari, P. Pedram, and L. Hedayatifar. Long-range correlation and multifractality in Bach's inventions pitches. *Journal of Statistical Mechanics: Theory and Experiment*, 2007:P04012, 2007.
[14] S. Jaffard. Oscillation spaces: Properties and applications to fractal and multifractal functions. *Journal of Mathematical Physics*, 39:4129–-4141, 1998.
[15] S. Jaffard. Wavelet techniques in multifractal analysis. *Proceedings of Symposia in Pure Mathematics (AMS)*, 72:91–-151, 2004.
[16] J. W. Kantelhardt, S. A. Zschiegner, and H. E. Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A*, 316:87–114, 2002.
[17] T. Lux. The Markov-switching multifractal model of asset returns: Gmm estimation and linear forecasting of volatility. *Journal of Business & Economic Statistics*, 26:194–210, 2008.
[18] J. F. Muzy, E. Bacry, and A. Arneodo. The multifractal formalism revisited with wavelets. *International Journal of Bifurcation and Chaos*, 4:245, 1994.
[19] H Nakao. Multi-scaling properties of truncated Lévy flights. *Physics Letters A*, 266:282-–289, 2000.
[20] P. Oswięcimka. A brief introduction to DFA-based multiscale analysis. In D. Grech and J Miśkiewicz, editors, *Simplicity of Complexity in Economic and Social Systems*, Springer Proceedings in Complexity, pages 89–106. Springer, Cham, 2021.
[21] P. Oświęcimka, S. Drożdż, M. Frasca, R Gębarowski, N. Yoshimura, L. Zunino, and L. Minati. Wavelet-based discrimination of isolated singularities masquerading as multifractals in detrended fluctuation analyses. *Nonlinear Dynamics*, 100:689–-1704, 2020.
[22] P. Oświęcimka, S. Drożdż, J. Kwapień, and A. Z. Górski. Effect of detrending on multifractal characteristics. *Acta Physica Polonica A*, 123:597–603, 2013.
[23] P. Oświęcimka, J. Kwapień, and S. Drożdż. Multifractality in the stock market: price increments versus waiting times. *Physica A*, 347:626–-638, 2005.
[24] A Rosas, E. Nogueira Jr., and J. F. Fontanari. Multifractal analysis of DNA walks and trails. *Physical Review E*, 66:061906, 2002.
[25] T Saito. An approach toward higher dimensional hysteresis chaos generator. *IEEE Transactions on Circuits and Systems*, 37:399–-409, 1990.
[26] H. E. Stanley and P. Meakin. Multifractal phenomena in physics and chemistry. *Nature*, 335:405–409, 1988.
[27] D. G. Stephen and J. A Dixon. Strong anticipation: multifractal cascade dynamics modulate scaling in synchronization behaviors. *Chaos Solitons Fractals*, 44:160–-168, 2011.
[28] H. Wendt, S. Roux, S. Jaffard, and P. Abry. Wavelet leaders and bootstrap for multifractal analysis of images. *Signal Processing*, 89:1100–1114, 2009.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Part III: **Practice**

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Paweł D. Domański and Maciej Ławryńczuk

6 Study on robustness of nonlinear model predictive control performance assessment

Abstract: Model Predictive Control (MPC) has gained the position of a reliable element of industrial control systems. It is considered as a very efficient solution in the case of complex, nonlinear, multivariate or other challenging applications. Sustainability of high operational quality requires close monitoring of MPC performance because industrial applications are often impeded by disturbances with unknown and complex properties. Quality indicators should be robust against such impacts to allow detection of performance degradation. The measures originating from robust statistics and the factors of α-stable distribution are used in this work. It is shown that they are able to detect effectively MPC detuning in the case of complex non-linear control. Neutralization benchmark simulations compare various approaches and prove their reliability. Model Predictive Control (MPC) has gained the position of a reliable element of industrial control systems. It is considered as a very efficient solution in the case of complex, nonlinear, multivariate or other challenging applications. Sustainability of high operational quality requires close monitoring of MPC performance because industrial applications are often impeded by disturbances with unknown and complex properties. Quality indicators should be robust against such impacts to allow detection of performance degradation. The measures originating from robust statistics and the factors of α-stable distribution are used in this work. It is shown that they are able to detect effectively MPC detuning in the case of complex non-linear control. Neutralization benchmark simulations compare various approaches and help to assess their relative reliabilities.

Keywords: control performance assessment, robust regression, stable distribution, MPC, pH neutralization

6.1 Introduction

Model Predictive Control (MPC) is an advanced control technique. It exhibits a specific feature because it utilizes an online mathematical model of the controlled process to predict its future behavior and to determine an optimal control strategy [23]. Optimization is repetitively run for each sampling period. MPC is renowned for its high accuracy and unique ability to take into account process variables' constraints during control

Paweł D. Domański, Maciej Ławryńczuk, Warsaw University of Technology, Institute of Control and Computation Engineering, Warsaw, Poland, e-mails: P.Domanski@ia.pw.edu.pl, M.Lawrynczuk@ia.pw.edu.pl

https://doi.org/10.1515/9783110729122-006

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

rule evaluation. In such a way MPC may deliver tangible benefits. From the start, predictive control was mostly applied in process industry or production systems. Nowadays, due to the availability of fast microcontrollers, MPCs are successfully utilized for fast embedded systems [1, 13, 17, 25, 26].

Precise control of neutralization processes is crucial in chemical engineering, biotechnology and waste-water treatment. As the process is significantly non-linear, it requires an adequate control strategy. Application of the gain scheduling scheme appears to be the most straightforward [7]. When a nonlinear model is applied inside the MPC, we obtain a constrained nonlinear optimization problem, which has to be solved every sampling period. Therefore, MPC with successive model linearization can be applied [11] to reduce computational complexity because it needs quadratic optimization only.

Maintenance and the sustainability of high MPC performance are primary industrial concerns. Experience shows that unsupervised operation can quickly lead to significant performance degradation that contests confidence in the entire implementation [20]. There are several works addressing MPC quality assessment [3]. A knowledge-based system for Dynamic Matrix Control (DMC) [18] was followed by benchmarking approaches [8] and data-driven approaches [28]. DMC has been tested in various application configurations as a single controller or the supervisory level over PID regulatory control [9]. The model-based method [15] has been accompanied by a minimum variance approach [27]. Even simple linear MPC might require alternative solutions, like fractal or persistence factors [2].

Nonlinear industrial control generates challenges for reliable MPC monitoring. It is nonlinear so linear approaches are not suitable. Real-time applications have to cope with frequent disturbances of unknown origin and complex properties. Moreover, the systems are often impacted by unexpected events like failures, maintenance activities, sensor calibration, system errors and personnel interventions. They cause outliers in the data, which may significantly alter a proper analysis. Our presented analysis focuses on outliers that pose significant challenges for control. Commonly used mean square error (MSE) is highly sensitive to any kind of outlier [16] because it is characterized by a 0 % breakdown point. Similar breakdown values characterize the Gaussian standard deviation. An integral of the absolute error (IAE) index is only slightly better and similar to the Laplace distribution scaling coefficient. Other research directions incorporate the concept of general robust statistics [12], proposing robust M-estimators with a 50 % breakdown. Specific non-Gaussian properties of control systems, especially fat tails, are rarely addressed in the research.

An outlier is a strange occurrence in the data [6, 24]. In general, a single outlying observation may come out of an erroneous observation (exogenous outliers), i. e., an unknown contaminating mechanism, or can be an intrinsic symptom of some underlying process-generating mechanism [10] (so-called endogenous outliers). Outliers seriously impede analysis: They enlarge the variance and reduce the power of statistical tests [14], deteriorate Gaussian properties, introduce fat and heavy tails in the

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

histogram [22] and bias regression [16]. An α-stable distribution is a good and robust approach [19] to model such behavior.

This work proposing using the pH neutralization nonlinear MPC control problem as a benchmark. Such a known plant enables observation of the impact of non-Gaussian disturbances on control quality and its assessment. Various scenarios are introduced to measure the sensitivity of control performance assessment (CPA) indicators that are confronted with a wide spectrum of disturbances.

Results show that standard key performance indicators (KPIs) are not effective. The novelty of this work is that the robust M-estimator, with the logistic ψ function and α-stable distribution scale factors, possess high effectiveness and robustness. Model-based and benchmarking indexes are intentionally not included in the analysis because they require *a priori* plant knowledge. Section 6.2 includes the presentation of the MPC problem, while Section 6.3 describes CPA measures. Theoretical introduction is followed by simulations (Section 6.4). Section 6.5 concludes the paper with observations and further research directions.

6.2 Nonlinear MPC

6.2.1 Control formulation

A Manipulated Variable (MV) is denoted by u and a Controlled Variable (CV) by y. The vector of decision variables determined at each sampling ($k = 0, 1, 2, \ldots$) by MPC [23] is denoted as

$$\triangle \boldsymbol{u}(k) = \left[\triangle u(k|k) \quad \triangle u(k+1|k) \quad \cdots \quad \triangle u(k+N_{\mathrm{u}}-1|k)\right]^{\mathrm{T}}, \tag{6.1}$$

where N_{u} is the control horizon, i. e., the number of future control increments $\triangle u(k|k) = u(k|k) - u(k-1)$ and $\triangle u(k+p|k) = u(k+p|k) - u(k+p-1|k)$ for $p = 1, \ldots, N_{\mathrm{u}}-1$. For $p \geq N_{\mathrm{u}}$ it is assumed that the MV is constant. Decision variables (6.1) are calculated from the optimization problem

$$\begin{aligned}
&\min_{\triangle \boldsymbol{u}(k)} \left\{ \sum_{p=1}^{N} \left(y^{\mathrm{sp}}(k+p|k) - \hat{y}(k+p|k)\right)^2 + \lambda \sum_{p=0}^{N_{\mathrm{u}}-1} \left(\triangle u(k+p|k)\right)^2 \right\} \\
&\text{subject to} \\
&u^{\min} \leq u(k+p|k) \leq u^{\max}, \quad p = 0, \ldots, N_{\mathrm{u}}-1 \\
&-\triangle u^{\max} \leq \triangle u(k+p|k) \leq \triangle u^{\max}, \quad p = 0, \ldots, N_{\mathrm{u}}-1 \\
&y^{\min} \leq \hat{y}(k+p|k) \leq y^{\max}, \quad p = 1, \ldots, N.
\end{aligned} \tag{6.2}$$

The first element of the performance index (6.2) minimizes predicted control errors over some prediction horizon N. The setpoint and predicted values of the process out-

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

put for future sampling moment $k + p$ known/calculated for current time k are denoted by $y^{sp}(k + p|k)$ and $\hat{y}(k + p|k)$. The predicted values of the process output are calculated using process mathematical model. The second part of the cost function mitigates excessive MV changes. Constraints may be imposed on future values of the manipulated variable, on the minimal and maximal allowed values u^{min} and u^{max}, on future changes of that variable, on the maximal value is $\triangle u^{max}$, and on predicted values of the controlled variable, the minimal and maximal values of which are y^{min} and y^{max}, respectively. Although the whole sequence of decision variables (6.1) is calculated at each and every sampling, only its first element is applied. During the next sampling, $k + 1$, the CV measurement is updated, and the procedure is repeated.

6.2.2 Computationally efficient nonlinear MPC

Future CV values $\hat{y}(k + p|k)$ are nonlinear functions of calculated decision variables (6.1) when the nonlinear model is used for prediction. As a result, the original optimization problem (6.2) becomes a constrained nonlinear task. To reduce computation effort, an algorithm with a Nonlinear Prediction and Linearization Along the Predicted Trajectory (MPC-NPLPT) can be taken into consideration. Unlike the case in simple MPC formulations with successive model linearization, a linear approximation of the future CV trajectory predicted on the prediction horizon N is determined online every sampling period. Linearization is carried out for a future trajectory of MV increments (6.1). This ends up with a simple quadratic MPC-NPLPT optimization problem [11].

6.3 Control quality assessment

Incessant striving for perfect performance is the ultimate *raison d'être* of any control system. The relationship is simple. The better the control system is, the higher process quality is reached. Although it is obvious, the majority of industrial controllers are not well tuned, nor properly set up [21]. The observation is true for any kind of applied control strategy. Engineers search for tools that enable better assessment of control system efficiency. Furthermore, they require indications of how to improve an imperfect system. The CPA research started in the 1960s and continues, and many approaches have been investigated [2]. Almost each control strategy, from the univariate PID loop to advanced multivariate MPC and adaptive controllers, has been assessed with appropriate methodologies. Control performance assessment was initiated by industry and is continually used and validated in practice.

We may distinguish two types of approaches: data-driven and model-driven. The difference is crucial: Model-driven methods require *a priori* process knowledge (step response, frequency characteristics, model orders, delay, etc). Data-driven approaches use only raw plant-operating data. The methods originate from different

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

domains and research areas. Classical indexes are not fully efficient, and industry still demands robust measures that would work under various conditions. This work compares KPIs utilizing control error ($\epsilon(k) = y^{\mathrm{sp}}(k) - y(k)$) time series [2]. We consider the following indicators:

– the Mean Square Error (MSE)

$$\mathrm{MSE} = \frac{1}{N_{\mathrm{p}}} \sum_{k=1}^{N_{\mathrm{p}}} \epsilon^2(k) = \frac{1}{N_{\mathrm{p}}} \sum_{k=1}^{N_{\mathrm{p}}} \left[y^{\mathrm{sp}}(k) - y(k)\right]^2, \tag{6.3}$$

– the Integral Absolute Error (IAE)

$$\mathrm{IAE} = \frac{1}{N_{\mathrm{p}}} \sum_{k=1}^{N_{\mathrm{p}}} \left|\epsilon(k)\right| = \frac{1}{N_{\mathrm{p}}} \sum_{k=1}^{N_{\mathrm{p}}} \left|y^{\mathrm{sp}}(k) - y(k)\right|, \tag{6.4}$$

– the Least Median Square (LMS)

$$\mathrm{LMS} = \underset{k}{\mathrm{med}}\, \epsilon(k)^2 = \underset{k}{\mathrm{med}}\, \left[y^{\mathrm{sp}}(k) - y(k)\right]^2, \tag{6.5}$$

– the Gaussian standard deviation σ_{G},
– the α-stable distribution scaling factor—γ,
– the robust scale M-estimator with logistic ψ function—σ_{H},
– the differential entropy—H_{diff},

$$H_{\mathrm{diff}} = -\int_{-\infty}^{\infty} \gamma(x) \ln \gamma(x) dx, \tag{6.6}$$

– the rational entropy—H_{rat}

$$H_{\mathrm{rat}} = -\int_{-\infty}^{\infty} \gamma(x) \log\left(\frac{\gamma(x)}{1+\gamma(x)}\right) dx. \tag{6.7}$$

6.4 Results

A pH neutralization reactor [5] is used as the simulated benchmark plant (Figs. 6.1 and 6.2). An unmeasured disturbance is generated as an α-stable process to introduce outliers. Gaussian noise is added to the output as a measurement noise. A base (NaOH) q_1, a buffer ($NaHCO_3$) q_2 and an acid (HNO_3) q_3 stream are mixed in a tank. The process has one MV, which is the base q_1 [ml/s] and one CV—a pH value. The continuous-time model comprises two state equations (6.8)–(6.9) and an algebraic output equa-

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

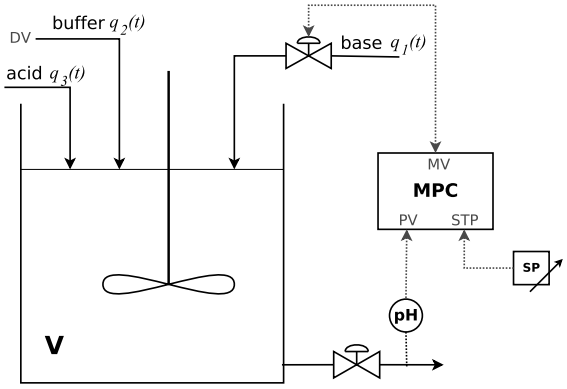


Figure 6.1: The pH-neutralization reactor.

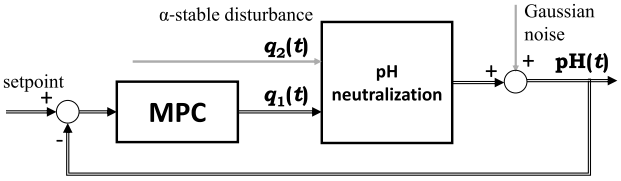


Figure 6.2: The control-loop simulation environment.

tion (6.10), namely

$$\frac{\mathrm{d}W_{\mathrm{a}}(t)}{\mathrm{d}t} = \frac{q_1(t)(W_{\mathrm{a}_1} - W_{\mathrm{a}}(t))}{V} + \frac{q_2(W_{\mathrm{a}_2} - W_{\mathrm{a}}(t))}{V} + \frac{q_3(W_{\mathrm{a}_3} - W_{\mathrm{a}}(t))}{V}, \tag{6.8}$$

$$\frac{\mathrm{d}W_{\mathrm{b}}(t)}{\mathrm{d}t} = \frac{q_1(t)(W_{\mathrm{b}_1} - W_{\mathrm{b}}(t))}{V} + \frac{q_2(W_{\mathrm{b}_2} - W_{\mathrm{b}}(t))}{V} + \frac{q_3(W_{\mathrm{b}_3} - W_{\mathrm{b}}(t))}{V}, \tag{6.9}$$

$$W_{\mathrm{a}}(t) + 10^{\mathrm{pH}(t)-14} - 10^{-\mathrm{pH}(t)} + W_{\mathrm{b}}(t)\frac{1 + 2 \times 10^{\mathrm{pH}(t)-K_2}}{1 + 10^{K_1-\mathrm{pH}(t)} + 10^{\mathrm{pH}(t)-K_2}} = 0. \tag{6.10}$$

Constants W_{a} and W_{b} are reaction invariants. Model parameters are sketched in Table 6.1 and the nominal operating point in Table 6.2. The inflow $q_2(t)$ is treated as the disturbance, while $q_3(t)$ is kept constant.

Table 6.1: Parameters of the pH-neutralization reactor model.

$W_{\mathrm{a}_1} = -3.05 \times 10^{-3}$ mol	$W_{\mathrm{b}_1} = 5 \times 10^{-5}$ mol	$K_1 = 6.35$
$W_{\mathrm{a}_2} = -3 \times 10^{-2}$ mol	$W_{\mathrm{b}_2} = 3 \times 10^{-2}$ mol	$K_2 = 10.25$
$W_{\mathrm{a}_3} = 3 \times 10^{-3}$ mol	$W_{\mathrm{b}_3} = 0$ mol	$V = 2900$ ml

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6.2: Nominal operating point of a pH reactor.

$q_1 = 15.55$ ml/s	$q_3 = 16.60$ ml/s	$W_a = -4.32 \times 10^{-4}$ mol
$q_2 = 0.55$ ml/s	pH = 7	$W_b = 5.28 \times 10^{-4}$ mol

6.4.1 Nonlinear MPC tuning scenarios

This described model is used as the simulated plant, while a Wiener process model is used for prediction. It consists of a linear second-order dynamic part, followed by a nonlinear static element. The dynamic part is characterized by equation

$$v(k) = b_1 u(k-1) + b_2 u(k-2) - a_1 v(k-1) - a_2 v(k-2), \tag{6.11}$$

where v denotes the auxiliary variable, i. e., the output of the first model block. The steady-state part is a continuous function

$$y(k) = g(v(k)). \tag{6.12}$$

A neural network is used as the nonlinear steady-state part of the Wiener model. For modeling purposes. the process variables are scaled: $u = (q_1 - q_{10})/15$, $y = (\mathrm{pH} - \mathrm{pH}_0)/5$, where $q_{10} = 15.55$ ml/s, $\mathrm{pH}_0 = 7$ correspond to the initial operating point. Nominal parameters of the MPC are: $N = 10$, $N_u = 3$, $\lambda = 0.5$; the constraints imposed on the manipulated variable are: $q_1^{\min} = 0$ ml/s, $q_1^{\max} = 30$ ml/s. Apart from a perfect tuning scenario (Sc0), six other MPC settings scenarios are considered to represent poor controller tuning:

Sc0: ideal model and tuning: $N = 10$, $N_u = 3$, $\lambda = 0.5$,
Sc1: the horizons too short: $N = 1$, $N_u = 1$, ($\lambda = 0.5$),
Sc2: horizon too long: $N = 20$ ($N_u = 3$, $\lambda = 0.5$),
Sc3: weighting too small: $\lambda = 0.025$ ($N = 10$, $N_u = 3$),
Sc4: weighting too large: $\lambda = 10.0$ ($N = 10$, $N_u = 3$),
Sc5: model gain 50 % smaller than the nominal one,
Sc6: model gain 50 % bigger than the nominal one.

6.4.2 Disturbance scenarios

Apart from the varying MPC tuning, the analysis takes into account the impac of fatness tails on the mistuning detection. The buffer inflow $q_2(t)$ is considered as the disturbance and is simulated as an α-stable stochastic process. The distribution stability index α of the stable distribution is responsible for the tails. Therefore, testing different α values enables us to investigate the impact of the tails fatness, i. e., of the share of introduced outliers. Eleven values are simulated: $\alpha = 1.0, 1.1, \ldots, 1.9, 2.0$. The smallest one $\alpha = 1.0$ reflects the largest tail fatness (Cauchy function), while the largest one

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

$\alpha = 2.0$ represents the Gaussian normal case. The other distribution coefficients remain unchanged: $\beta = 0.0$, $\delta = 0.55$, $\gamma = 0.01$. Finally, the signal is constrained and truncated into the reactor limitations $q_2(t) \in [0.3, 0.8]$.

Altogether, 77 simulation runs were conducted. The analysis starts from the control error time series review followed by the preparation of respective histograms and their analysis. Exemplary disturbance time trends are shown in Fig. 6.3. The presented time series are limited to the first 50,000 observations, though the analysis uses a full set of 250,000 samples.

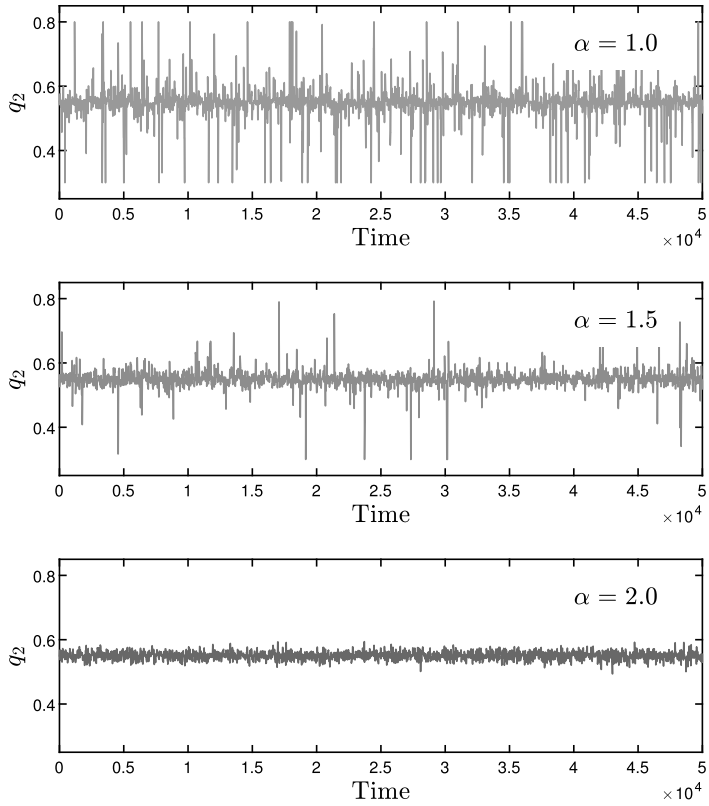


Figure 6.3: Example time series of the $q_2(t)$ disturbance for various values of α.

Statistical properties of the disturbances were verified using data histograms. Figure 6.4 shows a comparison of control-error exemplary data obtained for seven tuning scenarios (**Sc0**, ..., **Sc6**) for a selected single disturbance realization: $\alpha = 1.7$, $\gamma = 0.01$, $\beta = 0.0$, $\delta = 0.0$. Only the first 50, 000 simulation samples are plotted for the sake of the readability. Generally, the higher the fluctuation of the control-error signal is,

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

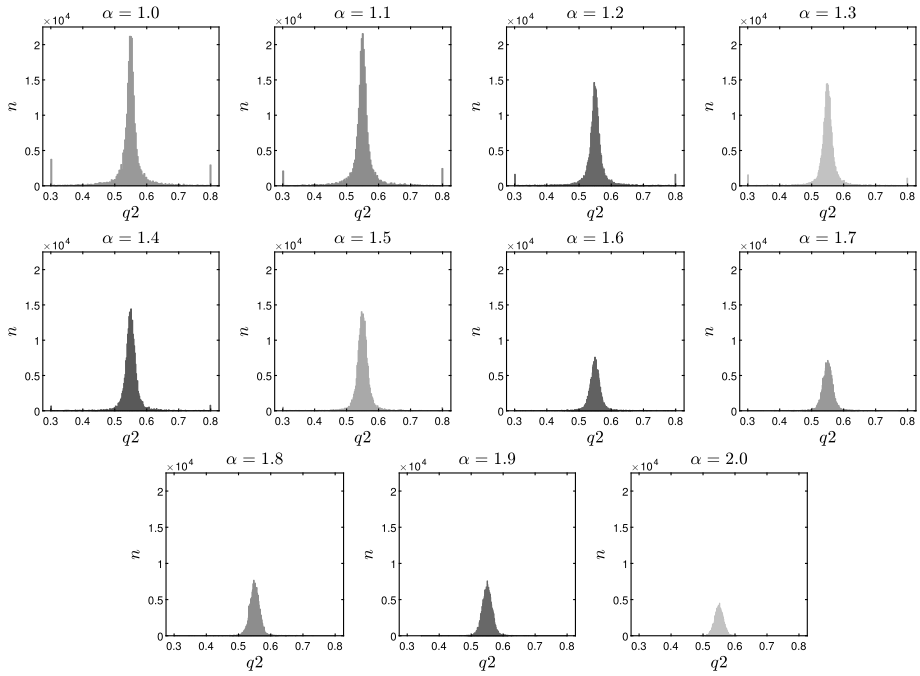


Figure 6.4: Histograms for all q_2 disturbance tail indexes; n is the number of occurrences.

the broader histogram is obtained, and the controller exhibits poorer performance. Scenario **Sc1** shows the most aggressive operation, while the **Sc3** is the most sluggish.

Control-error signal data are obtained after each of the simulations (examples are sketched in Fig. 6.5). Selected KPIs are calculated for each control-error dataset. The histogram for each time series enables us to evaluate its statistical factors. The simulation process produces a table consisting of eight performance indicators for each simulation. The obtained numbers differ by several orders of magnitude, so they have been rescaled.

Robustness review of selected indicators against the tail is measured. Various tail indexes (reflecting the outliers' impact) are simulated. The stability index changes from $\alpha = 1.0$ to $\alpha = 2.0$, incremented by 0.1. The other parameters are constant: $\gamma = 0.01$, $\beta = 0.0$, $\delta = 0.0$. Two main perspectives are investigated:

– **A1**: At first the impact of the tail fatness on the detectability obtained by each performance measure is analyzed. The index values are scaled against the reference, i. e., the realization for the normally distributed disturbance ($\alpha = 2.0$) KPI(2.0, Sc). The ratio (6.13) is calculated for each measure and the selected disturbance. The value varies with each tuning Sc0, . . . , Sc6.

$$\eta_{A1}^{\mathrm{KPI}}(\alpha, \mathrm{Sc}) = \frac{\mathrm{KPI}(\alpha, \mathrm{Sc})}{\mathrm{KPI}(2.0, \mathrm{Sc})}. \tag{6.13}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

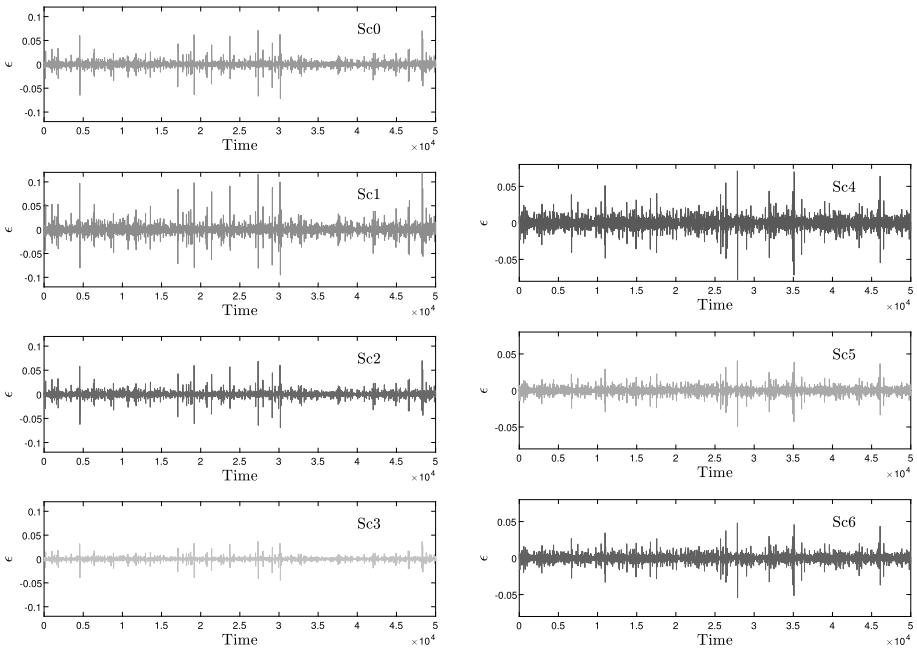


Figure 6.5: Control errors in considered scenarios of MPC tuning (Sc0, ..., Sc6): $\alpha = 1.7$, $\gamma = 0.01$, $\beta = 0.0$, $\delta = 0.0$.

– **A2**: The impact of the tail index on the ability to detect properly poor tuning by each of the indicators is the second analyzed perspective. The index value for each disturbance and for the nominal controller tuning (scenario Sc0) is considered as nominal. The ratio (6.14) between each indicator and its nominal value for the selected disturbance is calculated, showing how the KPI varies against its nominal scenario *Sc*0 value in each considered disturbance:

$$\eta_{A2}^{\mathrm{KPI}}(\alpha, \mathrm{Sc}) = \frac{\mathrm{KPI}(\alpha, \mathrm{Sc})}{\mathrm{KPI}(\alpha, \mathrm{Sc0})}. \tag{6.14}$$

Perspective A1, (the impact of outliers on the poor tuning detectability by each KPI) forms the main consideration, while the A2 is treated as an auxiliary observation.

6.4.2.1 Perspective A1

The analysis consists of two sets of plots. At first, the relationship between the scaled index ratio $\eta_{A1}^{\mathrm{KPI}}(\alpha, \mathrm{Sc})$ and the disturbance tail index is evaluated for each KPI. The resulting diagrams show seven curves presenting the dependence for each simulated MPC tuning. The following behavior is expected. First of all, the ratio should not

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

change with the tail-characteristic exponent. Moreover, the curves for each tuning should be distinguishable. Figure 6.6 depicts the robustness against the tail index for each KPI.

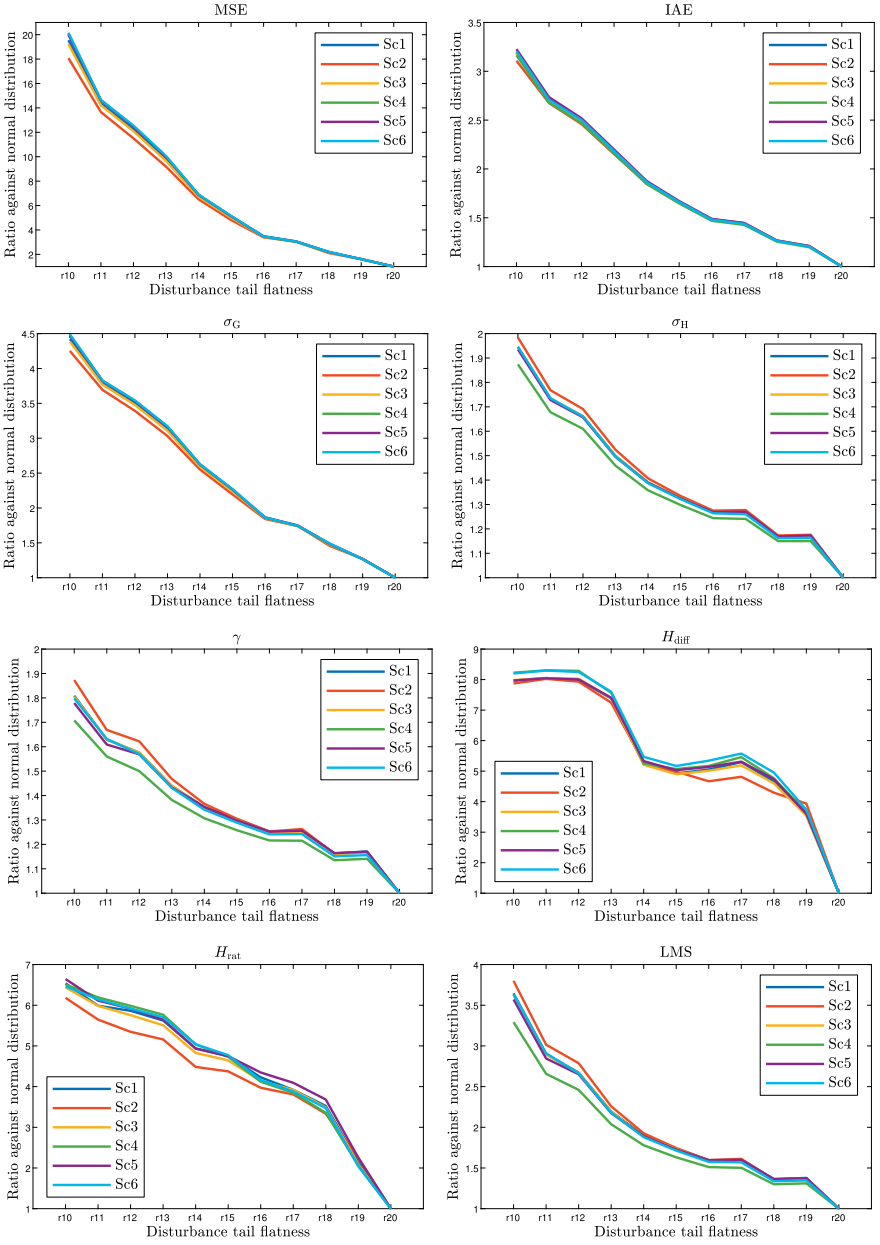


Figure 6.6: Diagrams showing robustness against tail index for each KPI.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Let us discuss the MSE index. Labels r10 to r20 denote the respective values of the disturbance stability factor α. Label r20 means $\alpha = 2.0$, r19 $\alpha = 1.9$ and so on. The index value significantly increases, starting from the normally distributed disturbance (r20) towards fatter tails (more outliers), i. e., as the stability index of the disturbance α decreases. Simultaneously, the curves for various MPC tuning are hardly distinguishable. Actually, only the evidently badly tuned scenario Sc1 (a too short prediction horizon) is easily distinguishable from the others.

Let us discuss the IAE index. It differs from the MSE. The diagrams are characterized by a much smaller difference between a normal Gaussian disturbance (r20) and Cauchy (depicted as r10), which exhibits the fattest tails and the largest number of outliers. It is almost ten times smaller. It shows higher IAE robustness. The observation is in compliance with the theory [16]. Simultaneously, MPC tuning scenarios are distinguishable similarly. The diagram for normal (Gaussian) standard deviation is very similar to that of the IAE, however the variability is smaller, which means larger robustness. The discernibility between the scenarios is on the same level as for the IAE.

Let us discuss the robust standard deviation M-estimator. The properties of the relationship seem to be much better than those of all the previous ones. Robustness is the highest with only about half the variability of the other curves. Moreover, the distinguishability between the curves depicting different tuning scenarios is the highest.

Let us discuss the γ scale factor of the stable distribution. It seems to be even better than for the robust standard deviation. Both parameters, i. e., the robustness and the discernibility, are the highest.

The diagram depicting the differential entropy H_{diff} shows the worst index properties. Apart from the high sensibility against the fat tails, the curves character is not monotonous, which may cause incorrect or ambivalent detection. This property decisively discards the differential entropy. The rational entropy H_{rat} exhibits better properties. Though its robustness is deficient, the distinguishability seems to be good enough.

Let us discuss the LMS index. Its properties are mediocre. Robustness, measured with the curves variability, is of the same magnitude as for normal standard deviation σ_{G}. On the contrary, the MPC tuning distinguishability is higher. It must be noted that the expectations about the LMS index were higher. Statistical analysis shows that the least median square index exhibits a 50 % breakdown, which should make it more robust than, for instance, the normal standard deviation or IAE (having 0 % breakdown point). The simulations do not confirm the expectations.

It can be noticed that CPA points are arranged according to some shape. Second-order polynomial approximations have been fitted, and the respective plots for the MSE and the robust standard deviation σ_{H} are sketched in Fig. 6.7.

Comparison of the variability (robustness against fat tails) of the considered CPA indicators is summarized in Table 6.3. We clearly see that the scale factor γ of stable

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

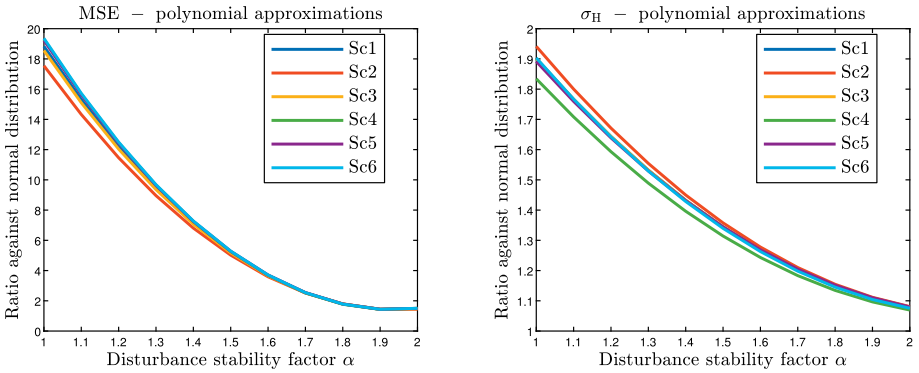


Figure 6.7: Robustness against fat tails for MSE (left) and σ_H (right) approximated by second- order polynomials.

Table 6.3: Indexes variability: green—smallest, red—highest.

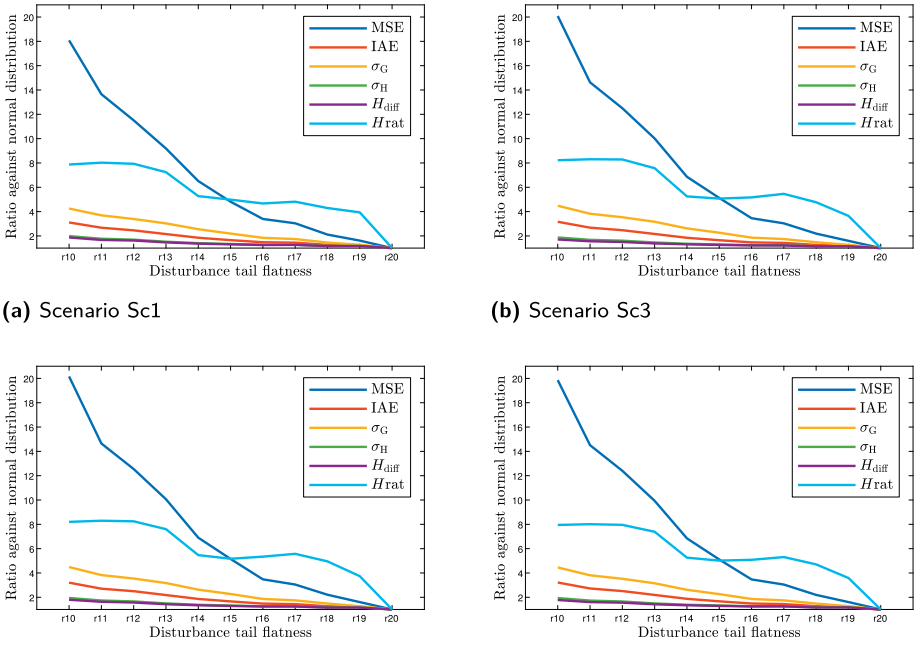
KPI	MSE	IAE	σ_G	σ_H	γ	H_{diff}	H_{rat}	LMS
range	19.16	2.23	3.49	0.98	0.87	7.23	5.71	2.80

distribution holds the highest robustness, with good enough discernibility between different MPC loop-control qualities.

The further diagrams use the same numbers, but show the results in a different way. They confront all the indicators in a single diagram, but they are grouped separately for each tuning scenario. Four scenarios (Sc1, Sc3, Sc5, Sc6), reflecting good curves representatives, are presented for the sake of readability. Robustness exhibited for the scenario Sc1 (the horizons are too short) is presented in Fig. 6.8a. We clearly see that two robust indicators of the scale factor γ for the stable distribution and robust scale M-estimator with the logistics ψ function share very similar, and the highest, robustness. The diagram presented in Fig. 6.8b shows the relationship for scenario Sc3, i. e., with a too small weighting coefficient $\lambda = 0.025$. The curves character and the results are very similar to the previous example. Finally, Figs. 6.8c and 6.8d present the tuning scenarios, for which the embedded model gain is 50 % smaller than the nominal one and 50 % larger, respectively. The curves' shape confirm the previous results and observations about the most robust indexes.

A review of the plots enables us to also determine the least robust indexes. It is evident that the MSE one is the worst selection. This observation might be confusing because the mean square error is the most popular and frequent selection utilized by almost everybody. But, one should take into account a reconsideration of the MSE selection. Two entropies also exhibit very low robustness and should not be considered as reliable CPA selections. Gaussian standard deviation seems to be a better solution

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use



(a) Scenario Sc1

(b) Scenario Sc3

(c) Scenario Sc5

(d) Scenario Sc6

Figure 6.8: KPIs robustness for various tuning scenarios.

than MSE and the entropies, but still it is four times more sensitive than two robust indicators. The LMS one LAO seems to be a disappointment. The remaining IAE measure seems to be relatively good, especially once one considers its much easier calculation complication and a common understanding.

Finally, review of the diagrams brings about one more observation. There is strong equivalence between two indicators. Scale factor γ for stable distribution and the robust scale M-estimator with logistics ψ function behave very similarly and share very close properties for all simulations and perspectives.

6.4.2.2 Perspective A2

The second set of analyzes includes plots of the relationship between the selected MPC tuning scenario (Sc0, . . . , Sc6) and the disturbance tail index (α). It is required that the gap between the improperly tuned MPC algorithms and the nominal one is visible as being disturbance invariant. The analysis shows that the behavior varies for each KPI. A typical relationship for the MSE is shown in Fig. 6.9.

A relative measurement of the relationship-curve variability for each index and scenario has been made in the form of a percentage ratio between the index range

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

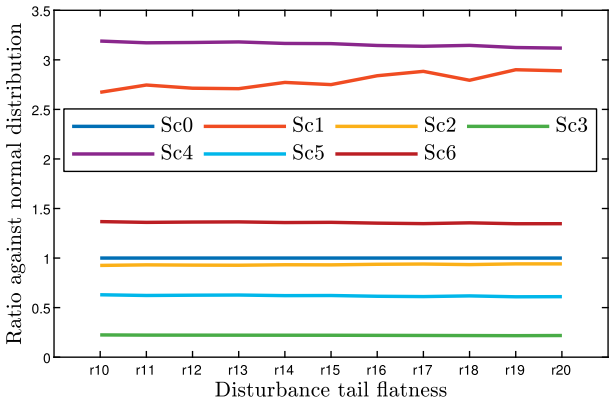


Figure 6.9: MSE ability to detect poor tuning as a function of the impeded disturbance.

Table 6.4: Indexes variability: green—smallest, red—highest.

	MSE	IAE	σ_G	σ_H	γ	H_{diff}	H_{rat}	LMS
Sc1	7.9 %	2.9 %	4.0 %	2.0 %	3.6 %	17.9 %	13.4 %	4.4 %
Sc2	1.7 %	0.4 %	0.8 %	0.1 %	0.5 %	2.6 %	2.6 %	0.9 %
Sc3	3.4 %	1.3 %	1.7 %	3.7 %	5.6 %	3.6 %	8.0 %	9.7 %
Sc4	2.2 %	1.0 %	1.1 %	0.8 %	1.8 %	2.2 %	5.0 %	2.3 %
Sc5	3.2 %	0.6 %	1.6 %	0.7 %	1.5 %	6.0 %	7.2 %	2.5 %
Sc6	1.6 %	0.7 %	0.8 %	0.4 %	1.7 %	2.3 %	5.3 %	3.2 %

and its maximum value. Table 6.4 presents the combined results for all the simulations involved. It is clear that the robust M-estimator and the absolute error IAE are the most robust indices, while both entropy measures H_{diff} and H_{rat} exhibit the worst properties. The MSE index is sensitive and its efficacy is significantly biased.

6.5 Conclusions

This work presents a sensitivity analysis of various CPA indexes against fat tails for nonlinear MPC control. The rationale of this work comes from industry. The analysis covers various sizes of fat tails, reflected by a stable distribution characteristic exponent (called the stability factor or stability index). Furthermore, distribution tails represent outliers in the considered dataset. Various tails are analyzed, starting from a normal distribution ($\alpha = 2.0$) up to the Cauchy function ($\alpha = 1.0$). It is shown that the considered indexes behave in different ways. The scaling factor of α-stable distribution and robust scale M-estimator using logistic ψ-function are the most robust KPIs. They exhibit similar properties, despite the size of the tail. On the contrary, both kinds

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

of entropy H_{diff} and H_{rat} give the worst results. It is also shown that the most popular mean square error is highly affected by the tails, and the assessment with MSE index is biased. The standard deviation, IAE and LMS indexes give mediocre performance.

The results confirm observations that commonly used indexes (MSE, IAE, Gaussian standard deviation) are sensitive to fat tails and outliers [4]. A decision to use the proposed approach should be done after the review of control system properties. When the loop variables are clearly Gaussian and fat-tails are not relevant, there is no need to apply robust indexes. Common MSE or standard deviation are enough then. However, once the tails appear, it is suggested to consider robust indexes (Huber standard deviation or stable distribution dispersion). Robust indicators work not only for fat-tail disturbances, but also for standard Gaussian ones. On the contrary, MSE or normal standard deviation works properly only in Gaussian cases, and fails once fat tails appear. Applicability depends on the control system properties (statistical analysis) and the character of loop variables.

There is one additional observation to be made. Since the considered robust indexes improve control performance assessment of nonlinear MPC control, they can also improve the operation of the controller when applied inside the MPC formulation, replacing the quadratic norm. Although this is a challenging task, nonetheless, it enables having a new degree of freedom in the MPC design.

Extrapolation of these results to the industrial plants requires incorporating visual inspection of the system variables, identification of the possible outliers, verification of the statistical properties of the loop signals and eventual addition of the robust indexes to the assessment procedure. There is no single, universal index. An assessment engineer has to carefully observe variables' time trends and histograms and always use a hybrid approach with some set of indexes because each of the measures addresses a different control feature.

Further research should assess other properties of the industrial disturbances. Signals may exhibit non-stationary, asymmetrical, nonlinear or other distributions. Especially non-stationarity causes unexpected variations to the quality observations. More complex multivariate systems should be also addressed. It is expected that the existing problems might be only magnified in MIMO cases. Multivariability elicits additional effects of the interactions between channels requiring multicriteria assessment and root-cause analysis.

Bibliography

[1] P. Chaber. Fast nonlinear model predictive control algorithm with neural approximation for embedded systems: Preliminary results. In A. Bartoszewicz, J. Kabziński, and J. Kacprzyk, editors, *Advanced, Contemporary Control*, pages 1067–1078, Springer Cham, 2020.

[2] P. D. Domański. *Control Performance Assessment: Theoretical Analyses and Industrial Practice*. Springer International Publishing, Cham, 2020.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[3] P. D. Domański. Performance assessment of predictive control—A survey. *Algorithms*, 13(4):97, 2020.
[4] P. D. Domański and M. Ławryńczuk. Control quality assessment of nonlinear model predictive control using fractal and entropy measures. In W. Lacarbonara, B. Balachandran, J. Ma, J. A. Tenreiro Machado, and G. Stepan, editors, *Nonlinear Dynamics and Control*, pages 147–156. Springer International Publishing, Cham, 2020.
[5] J. C. Gómez, A. Jutan, and E. Baeyens. Wiener model identification and predictive control of a pH neutralisation process. *IEE Proceedings Part D Control Theory and Applications*, 151:329–338, 2004.
[6] D. M. Hawkins. *Identification of Outliers*. Chapman and Hall, London; New York, 1980.
[7] A. W. Hermansson and S. Syafiie. Model predictive control of ph neutralization processes: A review. *Control Engineering Practice*, 45:98–109, 2015.
[8] R. H. Julien, M. W. Foley, and W. R. Cluett. Performance assessment using a model predictive control benchmark. *Journal of Process Control*, 14(4):441–456, 2004.
[9] M. Khan, M. Tahiyat, S. Imtiaz, M. A. A. S. Choudhury, and F. Khan. Experimental evaluation of control performance of MPC as a regulatory controller. *ISA Transactions*, 70:512–520, 2017.
[10] L. B. Klebanov and I. Volchenkova. Outliers and the ostensibly heavy tails. arXiv:1807.08715v1, Department of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic, 2018.
[11] M. Ławryńczuk. *Computationally Efficient Model Predictive Control Algorithms: A Neural Network Approach*. Studies in Systems, Decision and Control, volume 3. Springer, Cham, 2014.
[12] S. Morgenthaler. A survey of robust statistics. *Statistical Methods and Applications*, 15:271–293, 2007.
[13] H.-N. Nguyen, R. Bourdais, and P.-O. Gutman. Fast model predictive control for linear periodic systems with state and control constraints. *International Journal of Robust and Nonlinear Control*, 27(17):3703–3726, 2017.
[14] J. W. Osborne and A. Overbay. The power of outliers (and why researchers should always check for them). *Practical Assessment, Research, and Evaluation*, 9(6):1–8, 03 2004.
[15] G. Pannocchia, A. De Luca, and M. Bottai. Prediction error based performance monitoring, degradation diagnosis and remedies in offset-free MPC: Theory and applications. *Asian Journal of Control*, 16(4):995–1005, 2014.
[16] P. J. Rousseeuw and A. M. Leroy. *Robust Regression and Outlier Detection*. John Wiley & Sons, Inc., New York, NY, USA, 1987.
[17] J. Sawulski and M. Ławryńczuk. Optimization of control strategy for a low fuel consumption vehicle engine. *Information Sciences*, 493:192–216, 2019.
[18] J. Schäfer and A. Cinar. Multivariable MPC system performance assessment, monitoring, and diagnosis. *Journal of Process Control*, 14(2):113–129, 2004.
[19] M. Shao and C. L. Nikias. Signal processing with fractional lower order moments: Stable processes and their applications. *Proceedings of the IEEE*, 81(7):986–1010, 1993.
[20] J. F. Smuts and A. Hussey. Requirements for successfully implementing and sustaining advanced control applications. In *Proceedings of the 54th ISA POWID Symposium*, pages 89–105, 2011.
[21] K. D. Starr, H. Petersen, and M. Bauer. Control loop performance monitoring—ABB's experience over two decades. *IFAC-PapersOnLine*, 49(7):526–532, 2016. 11th IFAC Symposium on Dynamics and Control of Process Systems Including Biosystems DYCOPS-CAB 2016.
[22] N. N. Taleb. Statistical consequences of fat tails: Real world preasymptotics, epistemology, and applications. arXiv:2001.10488, Cornell Univesity Library, 2020.
[23] P. Tatjewski. *Advanced Control of Industrial Processes, Structures and Algorithms*. Springer, London, 2007.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[24] H. Wainer. Robust statistics: A survey and some prescriptions. *Journal of Educational Statistics*, 1(4):285–312, 1976.

[25] P. Wang, X. Feng, W. Li, X. Ping, and W. Yu. Robust rhc for wheeled vehicles with bounded disturbances. *International Journal of Robust and Nonlinear Control*, 29(7):2063–2081, 2019.

[26] A. Wojtulewicz. Implementation of dynamic matrix control algorithm using field programmable gate array: Preliminary results. In W. Mitkowski, J. Kacprzyk, K. Oprzędkiewicz, and P. Skruch, editors, *Trends in Advanced Intelligent Control, Optimization and Automation*, pages 325–334, Springer International Publishing, Cham, 2017.

[27] F. Xu, B. Huang, and E. C. Tamayo. Assessment of economic performance of model predictive control through variance/constraint tuning. *IFAC Proceedings Volumes*, 39(2):899–904, 2006. 6th IFAC Symposium on Advanced Control of Chemical Processes.

[28] Y. Xu, G. Zhang, N. Li, J. Zhang, S. Li, and L. Wang. Data-driven performance monitoring for model predictive control using a mahalanobis distance based overall index. *Asian Journal of Control*, 21(2):891–907, 2019.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Michał Falkowski

7 Causality analysis incorporating outliers information

Abstract: This paper presents the results of deep analysis of causality detection research in multi-control loop system. Considerations focus on the Transfer Entropy (TE) approach application, which has not been widely used during precise construction of information and material-flow pathways. Data knowledge of an industrial nonlinear process is not required. Calculations are performed on simple control loop data, obtained from real data objects, whose structure is *a priori* known. Process data is biased by a factor in the form of outliers. Methods of determining entropy between process variables are defined, basined on different variants for calculating probability distribution.

Keywords: transfer entropy, fault propagation, outliers, causality, large-scale industrial systems, process data decomposition, Darbellay–Vajda algorithm, ranking of process variables

7.1 Introduction

One of the most difficult issues regarding analysis of large-scale industrial processes is to find the root causes of faults. Faults that are commonly caused by inappropriate control loops operation may lead to low productivity of whole system, can increase operational costs and, in the most dangerous cases, to result in an unwanted system shutdown or its destruction. However, to begin this type of analysis, causality between variables of a given process should be determined first. One may use several different approaches, which have been developed in different domains. None of them are applied in industry. They are mainly used in medicine or chemistry, and they are based on a model (e. g., Granger Causality [16]). It has been shown that building a model is a tedious process, and its accuracy depends on many factors [3]. This is a complex and time-consuming issue; thus, the universality of these methods is negligible. Problems also appears because these approaches work well only for linear systems.

A key solution can be found in the group of data-driven model-free methods (e. g., TR entropy and its derivatives). These approaches are based on historical data of given processes. TR entropy enables determination of causality (pathways) between variables with minimal efforts.

Michał Falkowski, Warsaw University of Technology, Institute of Control and Computation Engineering, Warsaw, Poland, e-mail: 01105966@pw.edu.pl

https://doi.org/10.1515/9783110729122-007

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

This paper presents the results of calculating entropy with a classic approach [12] based on the Gaussian distribution, using the Darbellay–Vajda algorithm and where probability density estimation is determined by bin counting with fixed and equally-spaced bins. Analysis is performed on a raw dataset, detrended signals and noise with and without outliers, obtained during decomposition process. The impact of different data types on the effectiveness of the TE approach and pathway determination is shown.

7.2 Data decomposition

Decomposition and further analysis are performed on raw datasets of closed control loops collected from real large-scale industrial process of ammonia distillation. System is strongly nonlinear and a priori known for calculation of TEs between given variables. The outputs of given control loop are shown in form of trends in Fig. 7.1.

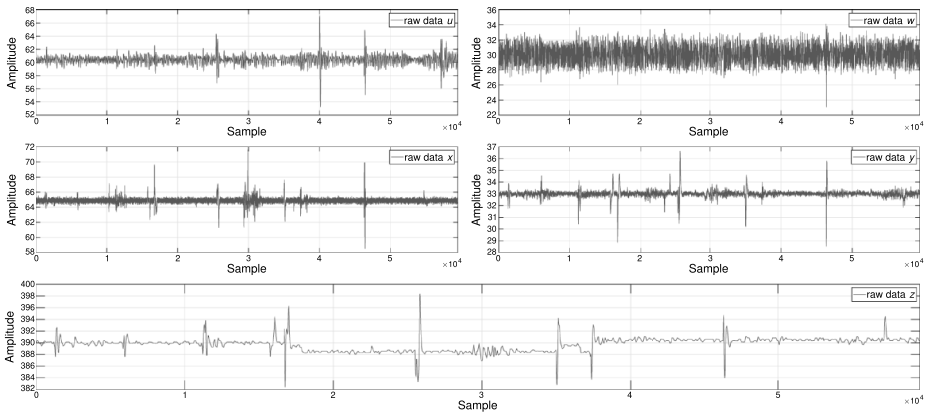


Figure 7.1: Raw dataset of the closed control loop from an ammonia distillation system.

Data is characterized by fairly stable trends for signals u, w, x and y, with occurring oscillations and multiple outliers. Signal z differs from the others by a visible, significant change in values.

7.2.1 Trends removal

Trends identification and removal is accomplished by using the most common polynomial interpolation. Choosing the right ith polynomial order is a contentious issue. Objective evaluation of selection of appropriate polynomial order is carried out by applying the methodology presented in [13]. An increasingly complex trend is set, then,

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

after its removal, the mean absolute deviation (MAD) of a signal is calculated. If the MAD converges to a relatively small and constant value, then the order of the polynomial is determined. Table 7.1 shows MAD calculations results up to the 9th polynomial order. Regardless of the selected order for the u, w, x, y signals, their MADs do not change significantly and only the removal of the constant value from the data is found.

Table 7.1: Mean absolute deviation (MAD) for ith polynomial order.

	0	**1st**	**2nd**	**3rd**	**...**	**6th**	**7th**	**8th**	**9th**
u	**0.5084**	0.5098	0.5099	0.5104	...	0.5105	0.5123	0.5113	0.5110
w	**0.8920**	0.8920	0.8919	0.8918	...	0.8917	0.8917	0.8918	0.8917
x	**0.2082**	0.2082	0.2082	0.2082	...	0.2082	0.2082	0.2082	0.2082
y	**0.1562**	0.1563	0.1563	0.1563	...	0.1561	0.1558	0.1555	0.1554
z	0.8699	0.8297	0.6656	0.6273	...	0.4239	0.4085	**0.4027**	0.4027

According to signal z, the fitted 8th order polynomial is unacceptable. Therefore, only in this case is it decided to use the spline interpolation that is defined piecewise by polynomials. The result of such operation is shown in Fig. 7.2. The trend is clearly better removed from the z signal using spline interpolation for which the MAD converged to the 0.3625 (2nd order polynomial).

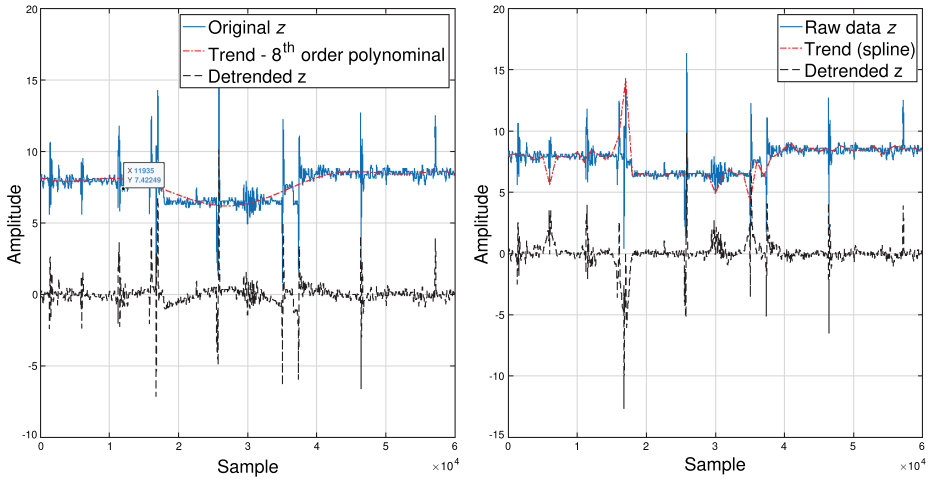


Figure 7.2: Comparison of trend identification and its removal from data z using polynomial and spline interpolation.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

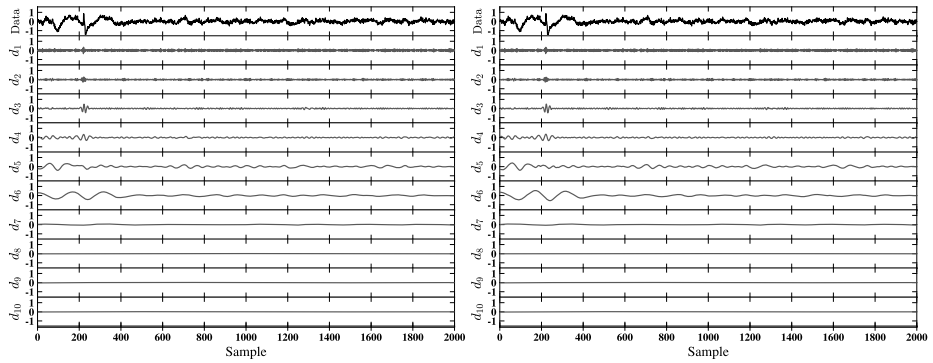
7.2.2 Noise identification

After removing trends from the dataset, the next step is to identify noise. Due to the amount of analyzed data and limitations posed by selected identification methods, this process cannot be performed on the entire dataset. For this purpose, the data is divided into 30 subsets of equal length. Two approaches were used: Ensemble Empirical Mode Decomposition (EEMD) and its extension Median Ensemble Empirical Mode Decomposition (MEEMD). This will make possible allow separation of the noise from the detrended dataset.

7.2.2.1 Ensemble empirical mode decomposition

Since there is usually noise and signal intermittency in real-world data, it causes mode mixing and mode splitting (MS) using the EMD method. To avoid this issue, EEMD [15] is proposed. EEMD belongs to a class of noise-assisted EMD methods which are the most powerful tools for performing time-frequency analysis [8, 11]. They are aimed at alleviating mode mixing caused by noise and signal intermittency and can be successfully used in noise and oscillation identification process.

An example of using the EEMD algorithm on one of the x signal subsets is presented in Fig. 7.3a. Decomposition strongly depends on the data length and fails during operations with large datasets. Figure 7.3a can be divided into the following sections: an analyzed subset (Data) and intrinsic mode function (IMFs), where from d_1 to d_4 there is noise and from d_5 to d_{10} are oscillations. The mail goal is to achieve not only noise but also oscillations for further analysis.



(a) EEMD oscillation identification **(b)** MEEMD oscillation identification

Figure 7.3: Example of signal x subset decomposition via: EEMD (a) and MEEMD (b).

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

7.2.2.2 Median Ensemble Empirical Mode Decomposition

MMEEMD [6] is a variation of the EEMD algorithm, which uses the median operator instead of the mean operator to the ensemble noisy intrinsic mode function (IMF) trials. An example of using the MEEMD algorithm on one of the x signal subsets is presented in Fig. 7.3b.

Despite the different characteristics of both algorithms, results are similar. It is justified to consider decomposed signals via MEEMD for further analysis. The addition of white Gaussian noise in the EEMD method alleviates the mode mixing problem, but it inevitably creates new MS.

Due to multiple number of subsets, it is impossible to show all the calculated noise signals for each variable. Based on the identified noise, TE calculation results are presented in Sec. 7.4.3 and 7.4.4.

7.2.3 Outliers removal

The majority of data is not normal enough to be considered as drawn from a normal distribution. A possible statistic is the IQR approach [14] in such cases. It is evaluated as the difference between the upper 75th (Q_3) and the lower 25th (Q_1) percentile of a dataset. IQR may be used to find outliers. They are considered as observations that fall below $LL = Q_1 - 1.5$ or above $HH = Q_3 + 1.5$ thresholds. The detected outliers are replaced by the threshold level values.

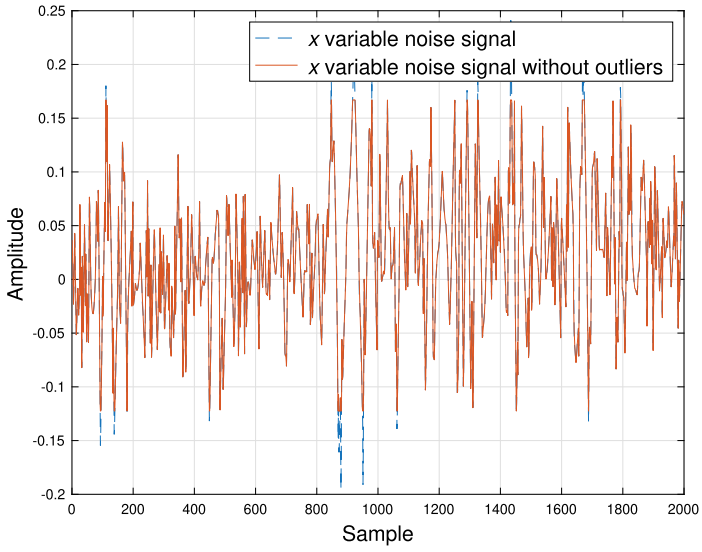


Figure 7.4: Example of outlier removal using the IQR approach on x noise signal.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

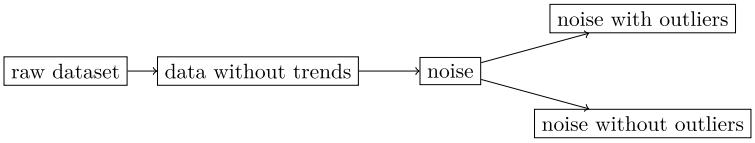


Figure 7.5: Diagram of dataset decomposition process.

MEEMD algorithm allows us to decompose data without the trends and obtain noise signals for individual variables u, w, x, y and z. It was already shown [4] that the IQR approach has the best performance in finding and removing outliers of a given dataset. Figure 7.4 shows the results of such an operation on one of the x variable subset.

The final data decomposition operation is represented by the diagram shown in Fig. 7.5. The following datasets are obtained for analysis using the TE approach: raw datasets, data without trends, noise with outliers and noise without outliers.

7.3 Transfer entropy

TE is an information-theoretic interpretation of Wiener's causality definition. TE is a measure of information transfer from x to y by measuring the reduction of uncertainty while assuming predictability. According to information theory, TE can be described by the equation

$$T_{x\to y} = \sum_{y_{i+h}, Y_i, X_j} p(y_{i+h}, Y_i, X_j) \cdot \log \frac{p(y_{i+h}|Y_i, X_j)}{p(y_{i+h}|Y_i)}, \tag{7.1}$$

where p means the complete or conditional Probability Density Function (PDF), variables Y_i and X_j respectively as $Y_i = [y_i, y_{i-\tau}, \ldots, y_{i-(l-1)\tau}]$, $X_j = [x_j, x_{j-\tau}, \ldots, x_{j-(k-1)\tau}]$, τ is a sampling and h is a prediction horizon. To summarize, it is the difference between information about a future observation of x obtained from the simultaneous observation of past values of both x and y, and the information about the future of x obtained from the past values of x alone. The phenomenon of entropy in both directions is highly probable, which is why a measure described as $T_{x\to y} = T_{x|y} - T_{y|x}$ is decisively due to quantity and direction, which is causality.

Due to the fact that TE is based on probability and the usage of the classical Gaussian distribution is not a necessary condition, the TE approach was implemented using Huber and α-stable distributions [4]. However the applied change does not significantly affect the results of determining the relationships between given variables.

To improve the efficiency of calculating entropy coefficients, another probability density estimation method based on the Darbellay–Vajda (DV) algorithm [2] is presented. This adaptive histogram-generating process is determined by partitioning the

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

observation space into a finite number of non-overlapping rectangular cells that are obtained in the recursive process [9].

Finally, the simplest estimation approach to obtain the PDFs is implemented, namely, the Fixed Bins algorithm. It allocates data points to fixed, equally-spaced bins. To enhance robustness against outliers and sparse regions in the underlying distribution, there is combined fixed binning with ordinal sampling (ranking). In ordinal sampling, the two analyzed time-series values are substituted with their ranks in those time series, similar to most non-parametric statistical tests [5]. The ranks are integers ranging from the smallest to the largest values.

7.4 Calculation results

In the following subsections the pathway between five process variables u, w, x, y and z of the ammonia-distillation closed control loop is determined using the classic TE approach and its derivatives described in Section 7.3. All the calculations are performed on datasets obtained during decomposition, divided into 30 subsets of equal length.

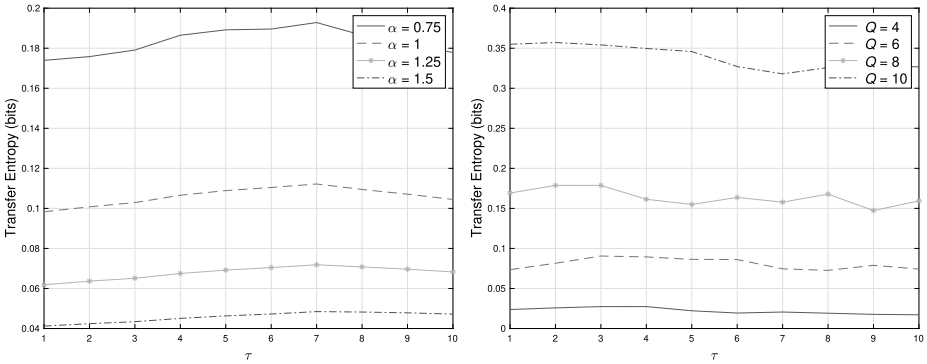
The practical implementation of thee TE approach between pairs of variables according to equation (7.1) requires its simplification to the form presented in equation (7.2):

$$T_{x\to y} = \sum_{y_i, y_{i-t}, x_{i-\tau}} p(y_i, y_{i-t}, x_{i-\tau}) \cdot \log \frac{p(y_i, y_{i-t}, x_{i-\tau}) p(y_{i-t})}{p(y_{i-t}, x_{i-\tau}) p(y_i, y_{i-t})}, \tag{7.2}$$

where p means the conditional (PDF) and τ and t are the time lags in x and y, respectively. If the time-series length is short, the t is set to 1 under the assumption that the maximum auto-transfer of information occurs from the data point immediately before the target value in y. Due to each subset length equalling 2,000 samples, t is set to 5. The same assumptions apply to the Darbellay–Vajda (DV) algorithm. The difficulty arises in selecting α parameter, which is a multiplier for scaling in PDF based on Gaussian Kernel Density Estimation and the optimal number of quantization levels Q for analyzed pair of signals (i. e. x and y) in the Fixed Bins algorithm due to the lack of a priori knowledge regarding coupling time lag [7].

To generalize the assumptions for the given 30 subsets and make the results comparable, respectively, for the raw dataset, data without trends, noise with outliers and noise without outliers, a short analysis of changes in the TE coefficient value as a function of α and Q parameters for KDE and Fixed Bins algorithm is performed, while increasing τ. Figure 7.6 illustrate TE estimation for various values of α and Q for the KDE and fixed-bins methods, respectively. Rejecting the overestimated values of TE for $\alpha = 0.75$ and $Q = 10$, the significant TE can be found at $\tau = 7$. Therefore, arbitrary, but reasonable, choices of $Q = 8$, $\alpha = 1$ and $\tau = 7$ are made.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use



(a) KDE: various bandwidths **(b)** Fixed-bin: various numbers of bins

Figure 7.6: Parameter selection for the KDE (a) and the fixed-bin (b) methods.

7.4.1 Raw dataset

Due to the large amount of data, it is impossible to illustrate the causality for each individual dataset. An exemplary TE coefficients calculation results for a random raw dataset with the KDE approach is presented in Table 7.2.

Table 7.2: Calculated TE coefficients for a random raw dataset with the KDE approach.

$T_{\text{row}\rightarrow\text{column}}$	*u*	*w*	*x*	*y*	*z*
u	*NA*	**0.1003**	0.0790	0.0690	0.0629
w	0.0636	*NA*	**0.0772**	0.0616	0.0436
x	0.0266	0.0463	*NA*	**0.0488**	0.0181
y	0.0634	**0.0900**	0.0754	*NA*	0.0435
z	0.0863	**0.1186**	0.0773	0.0786	*NA*

The fact is that so far no universal threshold method has been developed for the TE approach and an unambiguous definition of the relationship between the given variables. The following assumptions are made for the raw dataset and for each of the subsequent subsets: If the value of TE is determined in the form of *Nan* or *Inf* , it equals 0; the relationship between the variables in a dataset is determined on the basis of the highest value in a given row, as indicated in the example in Table 7.2. After determining the TE coefficients for a raw dataset, calculated with the KDE, DV and fixed-bins algorithms, in Table 7.3, Table 7.4 and Table 7.5, respectively, the results are presented in the form of a sum of cases where the value of the TE coefficient in a row for a given variables pair is the highest for all 30 subsets.

It can be seen that, both for KDE and fixed-bins approaches, the results are almost identical. Despite the different ratios of the number of datasets for which there are the

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 7.3: Sum of cases where the value of the TE coefficient in a row for a given variables pair is the highest—the KDE approach.

$T_{\text{row}\to\text{column}}$	*u*	*w*	*x*	*y*	*z*
u	*NA*	**14**	5	9	2
w	3	*NA*	11	**14**	2
x	3	7	*NA*	**19**	1
y	1	12	**14**	*NA*	3
z	5	**12**	6	7	*NA*

Table 7.4: Sum of cases where the value of the TE coefficient in a row for a given variables pair is the highest—DV approach.

$T_{\text{row}\to\text{column}}$	*u*	*w*	*x*	*y*	*z*
u	*NA*	0	0	0	**30**
w	1	*NA*	0	0	**29**
x	0	0	*NA*	0	**30**
y	2	1	0	*NA*	**27**
z	**18**	12	0	0	*NA*

Table 7.5: Sum of cases where the value of the Transfer Entropy coefficient in a row for a given variables pair is the highest—fixed-bins approach.

$T_{\text{row}\to\text{column}}$	*u*	*w*	*x*	*y*	*z*
u	*NA*	**26**	2	2	0
w	7	*NA*	10	**12**	1
x	1	9	*NA*	**20**	0
y	0	**20**	10	*NA*	0
z	10	**18**	0	2	*NA*

highest values of entropy coefficients in a given row, both methods show causality between the same pairs of variables. The only difference is in the y data pair: for KDE it is $y \to x$ and for fixed-bins it is $y \to w$. However, it can be noted that, in the case of the KDE approach, the number of datasets for which the y dependency is found is similar—12 and 14. Results obtained with the DV algorithm are completely different from the other algorithms. It can be stated with certainty that they do not reflect the actual causality between the control-loop variables. There may be many reasons: wrong parameters for the DV algorithm, data type (raw datasets) or just inappropriate algorithm for such calculations.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

7.4.2 Data without trends

Subsequent calculations are performed on datasets from which the undesirable trend has been removed, according to the methodology presented in Section 7.2.1. The results are shown in Table 7.6, Table 7.7 and Table 7.8 for three approaches introduced in this paper.

Table 7.6: Sum of cases where the value of the TE coefficient in a row for a given variables pair is the highest—the KDE approach.

$T_{\text{row}\to\text{column}}$	u	w	x	y	z
u	*NA*	**15**	5	6	4
w	2	*NA*	12	**15**	1
x	3	7	*NA*	**19**	1
y	1	13	**14**	*NA*	2
z	2	**11**	8	9	*NA*

Table 7.7: Sum of cases where the value of the TE coefficient in a row for a given variables pair is the highest—the DV approach.

$T_{\text{row}\to\text{column}}$	u	w	x	y	z
u	*NA*	0	0	0	**30**
w	0	*NA*	0	0	**30**
x	0	0	*NA*	0	**30**
y	0	0	0	*NA*	**30**
z	**20**	10	0	0	*NA*

Table 7.8: Sum of cases where the value of the TE coefficient in a row for a given variables pair is the highest—fixed-bins approach.

$T_{\text{row}\to\text{column}}$	u	w	x	y	z
u	*NA*	**26**	2	1	1
w	7	*NA*	10	**12**	1
x	1	9	*NA*	**20**	0
y	0	**20**	10	*NA*	0
z	4	**25**	0	1	*NA*

The results for the KDE and fixed-bins approaches are similar with the same differences as for the raw datasets analysis. Causality occurs for the following pairs of control loop signals: $u \to w$, $w \to y$, $x \to y$, $z \to w$ and ambiguously for y data—$y \to x$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

or $y \rightarrow w$, depending on the methods used. This is mainly because of how trends are removed from the data. Due to the characteristics of z data, a spline interpolation is used; for the other signals it is only a shift by the constant value (1st order polynomial). Also in this case, the DV algorithm does not bring the expected effects—the determined causality cannot be accepted, and the results must be discarded.

7.4.3 Noise with outliers

The choice of noise from the signals for analysis is not accidental—the additive noise models [10] are often used in many areas, including automation. As it is shown in Table 7.9, Table 7.10 and Table 7.11 results for the KDE and fixed-bins algorithms have

Table 7.9: Sum of cases where the value of the TE coefficient in a row for a given variables pair is the highest—the KDE approach.

$T_{\text{row}\rightarrow\text{column}}$	u	w	x	y	z
u	*NA*	**22**	1	6	1
w	**16**	*NA*	1	12	1
x	**11**	9	*NA*	8	2
y	10	**15**	2	*NA*	3
z	**19**	6	2	3	*NA*

Table 7.10: Sum of cases where the value of the TE coefficient in a row for a given variables pair is the highest—the DV approach.

$T_{\text{row}\rightarrow\text{column}}$	u	w	x	y	z
u	*NA*	0	0	0	**30**
w	0	*NA*	0	0	**30**
x	0	0	*NA*	0	**30**
y	0	0	0	*NA*	**30**
z	6	**21**	0	3	*NA*

Table 7.11: Sum of cases where the value of the TE coefficient in a row for a given variables pair is the highest—the fixed-bins approach.

$T_{\text{row}\rightarrow\text{column}}$	u	w	x	y	z
u	*NA*	**21**	2	7	0
w	**15**	*NA*	1	13	1
x	8	**13**	*NA*	9	0
y	5	**22**	3	*NA*	0
z	11	**15**	2	2	*NA*

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

negligible repeatability, and there is no reference to calculations performed for the raw datasets and those after trends removal.

The values of the highest entropy coefficient in a given row occur randomly. In case of determining causality diagrams for the obtained results, no logical consensus can be reached in any of them. The DV algorithm along with the noise data is also not working properly and still as in Subsections 7.4.1 and 7.4.2 does not indicate proper cause–effect relationships. This is only confirmed by the fact that the mentioned algorithm is not applicable in this case.

7.4.4 Noise without outliers

As a result of operation of decomposing the raw dataset and obtaining noise from the data, the consequence is to determine the entropy coefficients for noise without outliers. Despite the methodology used as described in the Section 7.2.3 and the satisfactory results of removing outliers from the datasets, it does not have a significant impact on the quality of the end effects. The KDE approach indicates the causality of variables with the data u in a given control loop and the fixed-bins with the data w. The DV algorithm invariably determines the relationship of variables with the given data z. The causality specified for this type of data is completely incorrect and should not be analyzed further which is confirmed by results given in Table 7.12, Table 7.13 and Table 7.14.

Table 7.12: Sum of cases where the value of the TE coefficient in a row for a given variables pair is the highest—the KDE approach.

$T_{\text{row}\to\text{column}}$	u	w	x	y	z
u	*NA*	**17**	2	10	1
w	**16**	*NA*	0	14	0
x	**15**	1	*NA*	14	0
y	**13**	**13**	4	*NA*	0
z	**17**	6	2	5	*NA*

Table 7.13: Sum of cases where the value of the TE coefficient in a row for a given variables pair is the highest—the DV approach.

$T_{\text{row}\to\text{column}}$	u	w	x	y	z
u	*NA*	0	0	0	**30**
w	0	*NA*	0	0	**30**
x	0	0	*NA*	0	**30**
y	0	0	0	*NA*	**30**
z	6	**21**	0	3	*NA*

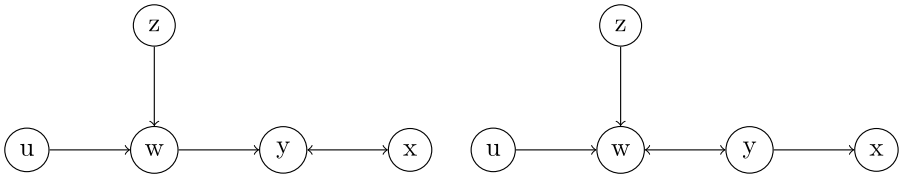
EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 7.14: Sum of cases where the value of the TE coefficient in a row for a given variables pair is the highest—the fixed-bins approach.

$T_{\text{row}\to\text{column}}$	*u*	*w*	*x*	*y*	*z*
u	*NA*	**21**	2	7	0
w	**15**	*NA*	1	13	1
x	8	**13**	*NA*	9	0
y	5	**22**	3	*NA*	0
z	11	**15**	2	2	*NA*

7.5 Conclusion and further research

The TE approach makes possible determination of the causality between variables, without any process knowledge. It is a tool with great potential for cause–effect analysis never used in automation before. In this paper the classic KDE approach with Gaussian distribution and its modifications using the DV and fixed-bins algorithms are introduced in Section 7.3. The calculations of the entropy coefficients are performed on the control-loop dataset obtained from the real object. The raw dataset is divided into 30 subsets of equal length and decomposed using various methodologies presented in Section 7.2.1, Section 7.2.2 and Section 7.2.3—therefore the TE index coefficients are calculated for datasets without trends, noise with outliers and noise without outliers, respectively. Due to the number of subsets for analysis, it is impossible to show the values of the entropy coefficients for each one. This also leads to the fact that it is impossible to adopt one threshold value for all the algorithms used, due to the range of the coefficient values that differ depending on the given approach; the only option is to indicate the pair of variables for which the entropy coefficient value is the highest in a given row, as presented in Table 7.3 to Table 7.14. As a result of the conducted tests and calculations, we can generate the causality diagrams in Fig. 7.7.



(a) KDE: causality diagram **(b)** Fixed-bins: causality diagram

Figure 7.7: Causality diagrams for KDE (a) and fixed-bins (b) methods.

Reasonable results are achieved only for the KDE and fixed-bins approaches. Due to the aforementioned limitations, this is the most general form of a causality diagram. It shows the relationships between the variables of the control-loop outputs. The only

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

difference between KDE and fixed-bins approaches is the relationships of signal y. The others are compatible.

The TE calculation method using the DV algorithm is not suitable for presented data and should be rejected. In this case causality coefficients for each data type and dataset are repeatable but do not coincide with the results for the other methods. This is probably caused by the DV algorithm's properties. It operates and depends only on two variable parameters: the delays τ and t for given pair of signals. A second hypothesis is as follows: If no τ value resulted in a significant information flow for subjectively selected constant delay t, that particular time series should be removed from further analysis. According to this statement, all datasets should be discarded from analysis. The inconsistency of the results with the KDE and fixed-bins methods confirms this hypothesis and rejection of the results obtained for the DV algorithm.

Based on the these considerations, further research should focus on the analysis of oscillation signals extracted from the data using the MEEMD decomposition algorithm. There are highly probable indications that, thanks to the properties of this type of signal, there will be a possibility of a sufficiently unambiguous definition of threshold value during calculation of TE coefficients and determination of causality diagrams. In conjunction with traditional statistical analysis and a fairly new approach based on L-moments analysis [1], subsequent findings will be obtained in the ongoing problem of causality detection and also as root causes of errors in multiple control-loop systems.

Bibliography

[1] W. H. Asquith. Univariate distributional analysis with l-moment statistics using R, 2011.

[2] G. A. Darbellay and I. Vajda. Estimation of the information by an adaptive partitioning of the observation space. *IEEE Transactions on Information Theory*, 45(4):1315–1321, 1999.

[3] M. Falkowski and P. D. Domański. *Nested NARIMA Model of the Atmospheric Distillation Column*. Springer, 2019.

[4] M. Falkowski and P. D. Domański. Impact of outliers on determining relationships between variables in large-scale industrial processes using transfer entropy. In *7th International Conference on Control, Decision and Information Technologies (CoDIT)*, 2020.

[5] A. Kaiser and T. Schreiber. Information transfer in continuous processes. *Physica D*, 166:43–62, 2002.

[6] X. Lang, N. Rehman, Y. Zhang, L. Xie, and H. Su Median ensemble empirical mode decomposition. *Signal Processing*, 176:107686, 2020.

[7] J. Lee, S. Nemati, I. Silva, B. A. Edwards, J. P. Butler, and A. Malhotra. Transfer entropy estimation and directional coupling change detection in biomedical time series. *BioMedical Engineering OnLine*, 11:19, 2012.

[8] D. P. Mandic, N. Rehman, Z. Wu, and N. E. Huang. Empirical mode decomposition-based time-frequency analysis of multivariate signals: the power of adaptive data analysis. *IEEE Signal Process*, 30(6):74–86, 2013.

[9] T. Marek and P. Tichavsky. On the estimation of mutual information. *ROBUST*, 2008.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[10] J. Peters, D. Janzing, and B. Scholkopf. Causal inference on discrete data using additive noise models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2436–2450, 2011.
[11] N. Rehman and D. Mandic. Multivariate empirical mode decomposition. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, The Royal Society*, 466(2117):1291–1302, 2010.
[12] T. Schreiber. Measuring information transfer. *Physical Review Letters*, 85:461–464, 2000.
[13] N. N. Taleb. How much data do you need? A pre-asymptotic metric for fat-tailedness. *International Journal of Forecasting*, 35(2):677–686 2019.
[14] D. L. Whaley. The interquartile range: Theory and estimation. Master's thesis, Faculty of the Department of Mathematics, East Tennessee State University, 2005.
[15] Z. Wu and N. E. Huang. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1(1):1–41, 2009.
[16] T. Yuan and S. J. Qin. Root cause diagnosis of plant-wide oscillations using granger causality. *Journal of Process Control*, 24(2):450–459, 2014.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Furkan Guc and YangQuan Chen

8 Backlash quantification in control systems using noises with outliers: a benchmark study

Abstract: Identification and quantification of mechanical backlash in dynamical systems has been studied for a long time in the literature. One of the key topics in the field is smart random-noise injection used to analyze the behavior and system metrics and correlate them with the mechanical backlash level. In this context, utilization of random noises with outliers offers the fulfillment of the gap in the linear dependence analysis of the backlash level with the corresponding system metrics. In this study, a set of noise characteristics is utilized with both classical definitions like Gaussian and uniform distributions; and random noises with outliers like fractional Gaussian noise distributions with various Hurst exponents and Pareto–Lévy stable distribution. Then, linear dependence of the backlash level and corresponding system metrics are analyzed with respect to various noise characteristics.

Keywords: backlash quantification, random noise injection, fractional Gaussian distribution, Hurst exponent, Pareto–Lévy stable distribution, fractional lower order moments (FLOM)

8.1 Introduction

The ball-and-beam structure is one of the key position reference tracking systems that is studied thoroughly in literature. It has served as a benchmark study for a variety of applications in nonlinear control [4, 16] and stability analysis [1, 17]. Moreover, it is highly utilized in many control laboratories for educational purposes due to the nature of the experimental setup [2, 11, 15]. As a common nonlinear effect in motion control and reference tracking applications, the ball-and-beam structure also suffers from mechanical backlash in various levels. Although the backlash compensation concepts are well-studied over many years of applications [5, 6, 9], most of the methodology rely on the model-based approaches that require backlash quantification at a certain level.

One of the most important points of the concept of backlash quantification is the utilized methodology. Model-based identification techniques include both parameter complexity and model dependency [12]. On the other hand, utilization of the noise

Furkan Guc, YangQuan Chen, Department of Mechanical Engineering, University of California, Merced, USA, e-mails: fguc@ucmerced.edu, ychen53@ucmerced.edu

https://doi.org/10.1515/9783110729122-008

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

techniques has become a standard tool in system identification [10, 13] and machine learning [3, 8]. Thereafter, noise study is extended to the methodology of smart noise injection [18].

In this study, backlash quantification in a control benchmark problem is introduced along with the random noise with outliers and the ball-and-beam structure. The basic Simscape model for the ball-and-beam system from the Control Tutorials for MATLAB and Simulink (CTMS) [7] is used. Then, backlash dynamics is introduced to the model. To quantify the backlash level, a set of random sensor noises is injected for various levels of backlash in the system, respectively. Finally, the linear dependence between the backlash level and corresponding system metric is analyzed under the influence of dedicated noise with the aid of fractional lower-order moments [14].

8.2 System definition

The ball-and-beam system is defined as a ball that has one degree of freedom as it rolls along a beam. To control the position of the ball, one end of the beam is pinned, and a lever arm is placed at the other end. The lever arm is controlled with a servo gear which has a turning angle of θ which causes a change in the angle of the beam α with respect to the reference. The main components of the system can be defined as the ball with a defined mass, inertia and radius; a beam with a defined length; and a lever-arm offset. The corresponding benchmark parameters of the ball-and-beam system are defined in Table 8.1.

Table 8.1: System parameters.

Parameter	Definition	Value	Unit
m	Ball Mass	0.1100	kg
R	Ball Radius	0.0150	m
d	Lever-Arm Offset	0.0300	m
L	Beam Length	1.0000	m
J	Ball Inertia	9.9900×10^{-6}	kgm^2

The ball-and-beam system is visualized in Fig. 8.1, and the Simulink block diagram of the closed-loop structure is given in Fig. 8.2, along with the noise injection subsystem given in Fig. 8.3. Details of the controller and ball-and-beam subsystem are defined in the CTMS [7].

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

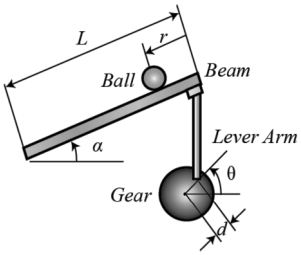


Figure 8.1: The benchmark ball-and-beam system from the CTMS [7].

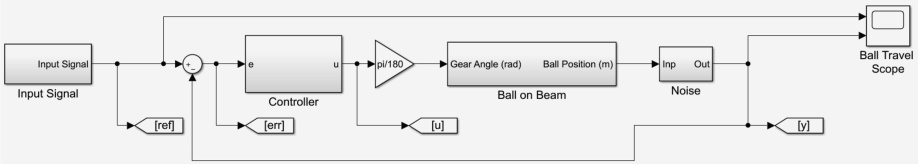


Figure 8.2: Simulink block diagram for the closed-loop structure for the benchmark ball-and-beam system.

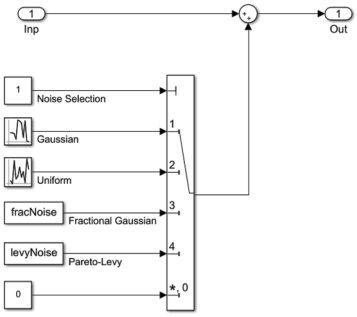


Figure 8.3: Simulink block diagram for the noise subsystem.

8.3 Methodology

To analyze the correlation of the underlying dynamics with various backlash levels, various definitions of noise distributions are utilized with both classical definitions like Gaussian and uniform distributions, while random noises with outliers, like fractional Gaussian noise distributions with various Hurst exponents (0.3, 0.5 and 0.8) and Pareto–Lévy stable distributions. Therefore, the need for a systematic way arises to compare the effect of various injected noise distributions. For each noise scenario, backlash levels of 0 %, 2 %, 4 %, 6 %, 8 % and 10 % are studied with purposely injected noise distributions. Then, the corresponding system metrics are defined as the mean

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

of the ball-position sensor measurements. The final step of the methodology is defined as the investigation of linear dependence between the backlash level and corresponding system metrics under the influence of dedicated noise distribution, along with the fractional lower-order moments. The overall workflow of the presented methodology is summarized in Fig. 8.4.

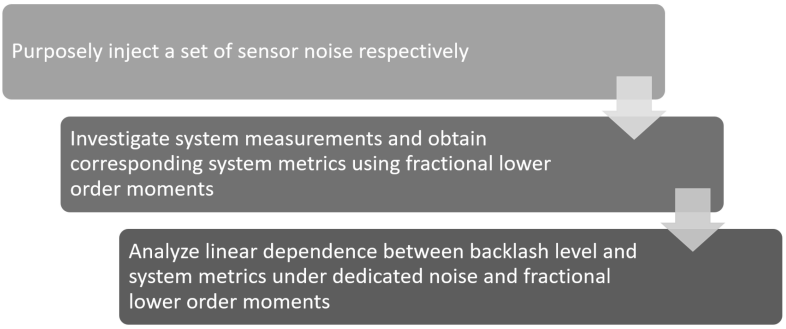


Figure 8.4: Summary of the overall workflow of the presented methodology.

Classical definitions like Gaussian and uniform distributions, along with random noises with outliers like fractional Gaussian noise distributions with various Hurst exponents (0.3, 0.5 and 0.8) and Pareto–Lévy stable distributions, are visualized in Fig. 8.5 and Fig. 8.6.

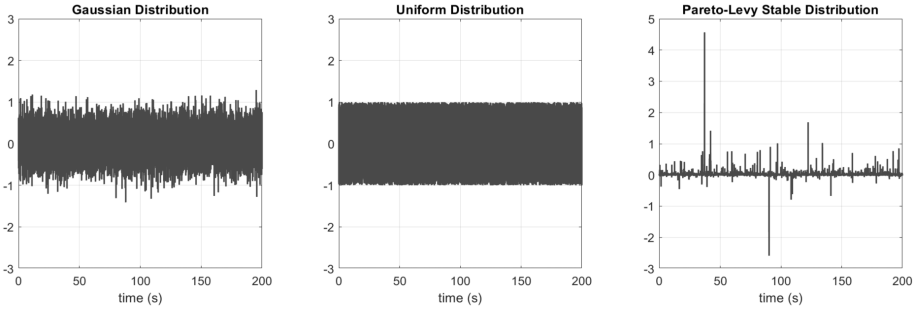


Figure 8.5: Purposely injected noises for Gaussian, uniform and Pareto–Lévy stable distributions.

Considering the system measurement X as $x_1, x_2, \ldots, x_N$ after random excitation, standardized moments are defined as

$$\alpha_r = \frac{\mu_r}{\sigma_r}, \tag{8.1}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

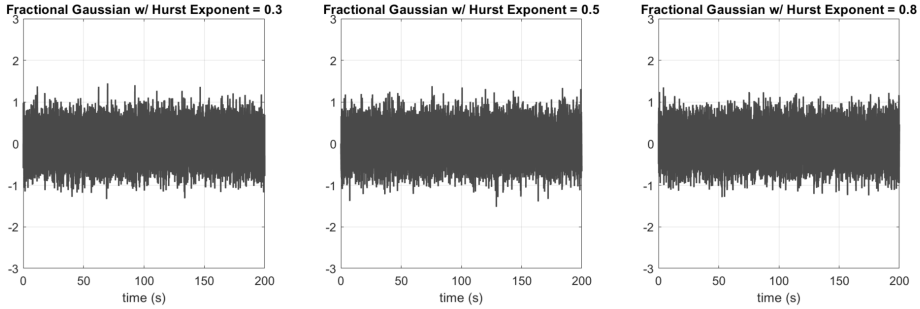


Figure 8.6: Purposely injected noises for fractional Gaussian noise distributions with various Hurst exponents (0.3, 0.5 and 0.8).

where

$$\mu_r = E\{(X-\mu)^r\}, \tag{8.2}$$

$$\sigma_r = \left(E\{(X-\mu)^2\}\right)^{r/2}, \tag{8.3}$$

with μ as the mean of the measurements and r as the order. Then, fractional lower-order moments are defined as

$$\mu_{rF} = E\{(X-\mu)^r\} = \frac{1}{N}\sum_{i=1}^{N}\left((x_i-\mu)^2\right)^{r/2}. \tag{8.4}$$

For each scenario of the noise distribution, fractional lower-order moments from 0.5 to 4 are calculated to find the best combination of the noise injection and fractional-moment order for the linear correlation between the backlash level and output metric. To quantify the linear correlation, the R-squared value of the linear fit for each scenario is calculated.

8.4 Results

Tto analyze and quantify the linear correlation between the backlash level and system metrics, the R-squared statistical metric is employed. In this context, R-squared is a measure of explained variance in the linear regression within the range of $[0, 1]$. An R-squared value of 1 indicates the best possible linear correlation between the backlash level and corresponding metric. To demonstrate the linear correlation and corresponding calculated R-squared values for various noise distributions, the R-squared values are presented in Table 8.2.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 8.2: R-squared values for various noise distributions and mean ball positions.

Noise distribution	R-squared—mean ball position
Gaussian	0.9572
Uniform	0.9991
fGn with H = 0.3	0.9937
fGn with H = 0.5	0.9980
fGn with H = 0.8	0.9726
Pareto–Lévy	0.1344

Moreover, the increasing backlash level of 0 %, 2 %, 4 %, 6 %, 8 % and 10 % vs. mean ball position for various noise distributions are visualized in Fig. 8.7 to show the linear correlation and R-squared values.

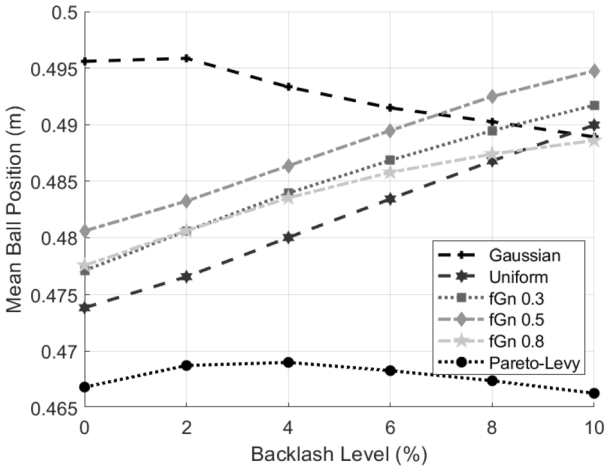


Figure 8.7: Mean ball position with respect to increasing backlash level.

Then, to quantify and find the best combination of the purposely injected noise distribution and fractional moment order, R-squared values for each scenario are calculated and presented in Fig. 8.8.

From the results that are shown with the R-squared values and various noise distributions for increasing fractional moment orders, the fractional Gaussian noise distribution performs better to quantify the linear correlation between the backlash level and the corresponding system metrics, while lower values for the Hurst exponent like 0.3 and 0.5 perform best in the sense of the R-squared metric. Among these fractional Gaussian noise distributions with lower values for the Hurst exponent like 0.3 and 0.5, lower fractional orders performs better. It can be concluded that a trend reversal

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

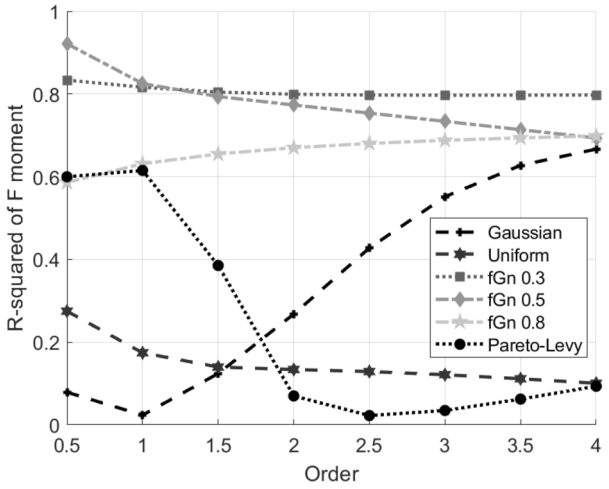


Figure 8.8: Fractional moment order vs. corresponding R-squared values for each noise distribution.

is more likely, and this may play a critical role in understanding the linear correlation between the backlash level and the corresponding system metrics.

8.5 Conclusion

In this chapter, backlash quantification in a ball-and-beam benchmark control study is investigated using both classical definitions like Gaussian and uniform distributions, as well as random noises with outliers like fractional Gaussian noise distributions with various Hurst exponents (0.3, 0.5 and 0.8) and Pareto–Lévy stable distribution. To analyze the linear correlation between the backlash level and respective system metrics of the fractional lower-order moments, a statistical definition of the R-squared metric is used. By using a corresponding cost function that utilizes both system metrics, it is shown that intentional noise injection with fractional Gaussian noise distribution performs better for the backlash quantification when compared to other noise characteristics, along with small fractional moment orders. Results presented in this work are reproduceable since the base Simscape model for the ball-and-beam system is used from the CTMS [7].

Bibliography

[1] C. Barbu, R. Sepulchre, W. Lin, and P. V. Kokotovic. Global asymptotic stabilization of the ball-and-beam system. In *Proceedings of the 36th IEEE Conference on Decision and Control*, volume 3, pages 2351–2355. IEEE, 1997.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[2] L. De La Torre, M. Guinaldo, R. Heradio, and S. Dormido. The ball and beam system: A case study of virtual and remote lab enhancement with moodle. *IEEE Transactions on Industrial Informatics*, 11(4):934–945, 2015.
[3] Y. Grandvalet, S. Canu, and S. Boucheron. Noise injection: Theoretical prospects. *Neural Computation*, 9(5):1093–1108, 1997.
[4] J. Hauser, S. Sastry, and P. Kokotovic. Nonlinear control via approximate input-output linearization: The ball and beam example. *IEEE Transactions on Automatic Control*, 37(3):392–398, 1992.
[5] I. Kolnik and G. Agranovich. Backlash compensation for motion system with elastic transmission. In *2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel*, pages 1–5. IEEE, 2012.
[6] L. Márton and B. Lantos. Control of mechanical systems with stribeck friction and backlash. *Systems & Control Letters*, 58(2):141–147, 2009.
[7] MATLAB. *Control Tutorials for MATLAB and Simulink (CTMS)*. The MathWorks Inc., Natick, MA, 2020.
[8] K. Matsuoka. Noise injection into inputs in back-propagation learning. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(3):436–440, 1992.
[9] M. Nordin and P.-O. Gutman. Controlling mechanical systems with backlash—A survey. *Automatica*, 38(10):1633–1649, 2002.
[10] R. Pintelon. Frequency-domain subspace system identification using non-parametric noise models. *Automatica*, 38(8):1295–1311, 2002.
[11] F. A. Salem. Mechatronics design of ball and beam system: Education and research. *Mechatronics*, 3(4), 2013.
[12] N. Sarkar, R. E. Ellis, and T. N. Moore. Backlash detection in geared mechanisms: Modeling, simulation, and experimentation. *Mechanical Systems and Signal Processing*, 11(3):391–408, 1997.
[13] J. Schoukens, R. Pintelon, G. Vandersteen, and P. Guillaume. Frequency-domain system identification using non-parametric noise models estimated from a small number of data sets. *Automatica*, 33(6):1073–1086, 1997.
[14] H. Sheng, Y. Q. Chen, and T. S. Qiu. *Fractional Processes and Fractional-Order Signal Processing: Techniques and Applications*. Springer Science & Business Media, 2011.
[15] G. Takács, E. Mikuláš, A. Vargová, T. Konkoly, P. Šíma, L. Vadovič, M. Bíro, M. Michal, M. Šimovec, and M. Gulan. Bobshield: An open-source miniature “ball and beam” device for control engineering education. In *2021 IEEE Global Engineering Education Conference (EDUCON)*, pages 1155–1161. IEEE, 2021.
[16] W. Yu. Nonlinear PD regulation for ball and beam system. *International Journal of Electrical Engineering Education*, 46(1):59–73, 2009.
[17] W. Yu and F. Ortiz. Stability analysis of PD regulation for ball and beam system. In *Proceedings of 2005 IEEE Conference on Control Applications, 2005. CCA 2005*, pages 517–522. IEEE, 2005.
[18] C.-B. Yun and E. Y. Bahng. Substructural identification using neural networks. *Computers & Structures*, 77(1):41–52, 2000.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kai Liu, YangQuan Chen, and Paweł D. Domański

9 Control performance assessment of the system with fractional-order dynamics

Abstract: Thousands of control loops comprise a standard industrial operation. The assessment of control loops plays a significant role for product engineers or system engineers. The statistical control quality measures are widely used in current industry to assess control loops, which are based on the classical Gaussian approach. The fractional-order system is the generalization of the conventional integer-order control system. The task becomes more challenging when the system behaves like fractional-order dynamics. In this chapter, we demonstrate that controller tuning has an effect on the Hurst exponents and also as crossover points at various time scales. Error signals from the control system subject to the disturbance with fractional-order dynamics are analyzed by multiple Hurst fitting lines with crossovers. Simulation results and real industry data are given to assess the performance of the control system with proposed the detrended fluctuation analysis (DFA) method with crossover points. The proposed method is a promising quantitative description to characterize control systems for both the short-term and long-term.

Keywords: control performance assessment, fractional-order dynamics, disturbance, Hurst exponent, detrended fluctuation analysis

9.1 Introduction

In current industry, statistical control quality tests based on the classical Gaussian method are commonly used to test control loops. Long-range dependence (LRD), self-similarity, the power law of autocorrelation, infinite variance, and spiky signals are examples of non-Gaussian signals and noises that cause large-amplitude deviations from the average value more frequently than Gaussian ones. In fact, most industrial data contradict the widely held belief that variables are Gaussian. Fractional calculus is the source of these complex phenomena.

Due to the complexity, correlation, time-varying delays and human intervention, the majority of real-life industrial process data has the fat-tailed property, after reviewing such data from industrial processes [1]. Liu et al. proposed a novel control performance assessment (CPA) method with fractional-order signal processing (FOSP) tech-

Kai Liu, YangQuan Chen, Mechatronics, Embedded Systems and Automation (MESA) Lab, University of California, Merced, USA, e-mails: liukai5005@gmail.com, ychen53@ucmerced.edu
Paweł D. Domański, Warsaw University of Technology, Institute of Control and Computation Engineering, Warsaw, Poland, e-mail: p.domanski@ia.pw.edu.pl

https://doi.org/10.1515/9783110729122-009

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

niques in [6]. Hurst exponent fitting with FLOM and multifractal detrended fluctuation analysis (MFDFA) with fractional lower-order moments was proposed in [5]. Strong correlations between process variables at different time instants characterize scenarios of long-range dependence, nonstationary and spiky signals. The non-Gaussian spiky statistical property is conventionally considered as the source of outliers. However, such non-Gaussian behavior can be precisely described and modelled by autoregressive fractionally integrated moving average (ARFIMA) models [7].

9.2 DFA method with crossover points

9.2.1 Hurst exponent

The Hurst exponent H indicates the smoothness of the time series in general: The lower the H, the rougher the time series [8]. The Hurst exponent, which characterizes LRD, indicates that there is a clear coupling effect between values at different time intervals. Process engineers are generally interested in the prospect of using significantly shorter time series for practical purposes. While this asymptotic scaling exponent can be useful for some diagnostic purposes, it has the disadvantage of requiring very large data sets for statistically robust performance. It is probably due to the fact that transient control signals are dominated by the relatively smooth fluctuation, thus resulting in a high local value H. The repeatable signals represent the intrinsic dynamics of a complex system at larger scales. As a result, the Hurst exponent H is a good predictor for describing a system's dynamics over multiple time scales.

9.2.2 Algorithm of the DFA with crossover points

The DFA algorithm is a robust and powerful tool for detecting long-range correlations buried in nonstationary data [8]. For a series x_i with the length N, the DFA process is as follows:

1. Construct a cumulative series $X(i)$

$$\begin{aligned} X(i) &= \sum_{j=1}^{i} (x_j - \bar{x}), \\ \bar{x} &= \frac{1}{N} \sum_{i=1}^{N} x_i. \end{aligned} \tag{9.1}$$

2. Divide the whole series into non-overlapping segments of equal length s. Since the total length N may not be divisible by s, some portion of $X(i)$ will be left unused. To take the advantage of the whole range series, the same operation will be carried out from the end side of the series $X(i)$. Accordingly, $2K$ segments are obtained.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. Use the least squares regression algorithm to fit the local polynomial trend for each of the $2K$ segments and calculate the variance:

$$F^2(v,s) = \frac{1}{s}\sum_{i=1}^{s}\big(X[(v-1)s+i] - x_v(i)\big)^2. \tag{9.2}$$

Thus, for the vth segment, $v = 1, 2, 3, \ldots, K$:

$$F^2(v,s) = \frac{1}{s}\sum_{i=1}^{s}\big(X[N-(v-K)s+i] - x_v(i)\big)^2. \tag{9.3}$$

4. Compute the root-mean-square deviation (RMSD) fluctuation function $F(s)$

$$F(s) = \sqrt{\frac{1}{2K}\sum_{v=1}^{2K} F^2(v,s)}. \tag{9.4}$$

5. Modify the scale s and repeat the above four steps. If x_i is long-range dependence (correlated), the Hurst exponent can be illustrated by the following power law:

$$F(s) \sim s^H. \tag{9.5}$$

6. Use the polynomial curves to fit the two parts of $F(s)$ separated by the moving crossover point c in a log–log plot, respectively. Then, select the crossover point c^* with the minimum error sum of squares E of two polynomial-fitting parts:

$$E_{c^*} = \min\left(\sum_{k=1}^{c} e_k^2 + \sum_{k=c}^{s} e_k^2\right). \tag{9.6}$$

9.2.3 Fractional Gaussian noise

In time-series analysis, the auto-regressive (AR) model constructs a short memory of the previous data, and moving average (MA) models the random noise. The fractional integrated (FI) characterizes the long-memory properties. The ARFIMA (p, d, q) model of a time series X_t is defined as follows [3]:

$$\left(1 - \sum_{i=1}^{p}\phi_i B^i\right)(1-B)^d(X_t - \mu) = \left(1 + \sum_{i=1}^{q}\theta_i B^i\right)\varepsilon_t, \tag{9.7}$$

where $\phi_1, \ldots, \phi_p$ are AR parameters, $\theta_1, \ldots, \theta_q$ are MA parameters, μ is the expectation of X_t, ε_t is the white noise and B denotes the unit delay. In contrast to the ordinary auto-regressive integrated moving average (ARIMA) process, the difference operator

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

d in (p, d, q) is allowed to take non-integer (fractional) values. Hosking defined the fractional difference operator by the binomial series ∇^d [4]:

$$\nabla^d = (1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k = \sum_{k=0}^{\infty} \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d+1-k)} (-B)^k. \tag{9.8}$$

To the best of our knowledge, the control system under the LRD noise has not yet been studied in the CPA. In this chapter, we will construct a control-system simulation with fractional-order Gaussian noise by changing the value of d in the fractional-order difference filter $(1-B)^d$ in Eq. (9.7). In the following sections, we focus on two major questions: How does the LRD noise affect the control system?; and how does one assess the control performances based on various different tunings of the PID controller?

9.3 Simulation analysis

In the simulation, the first-order plus time delay (FOPTD) system $\frac{1}{2s+1}e^{-0.5s}$ is used to simulate the process model in Fig. 9.1. The discrete PID controller parameters are: P = 1.5, I = 1.0, D = 0.1; and the sampling rate is 50 Hz. Control errors $e_t = y^* - y_t$ are used to assess the control performance, where y^* is the setpoint (target) and y_t is the feedback signal. The fractional-order Gaussian noise is modeled and simulated by the ARFIMA$(0, d, 0)$ with LRD disturbances $d \in [-0.5, 0.5]$. $d = 0$ corresponds to the white noise, and the fractional-order integral (derivative) coefficient d changes the slope H in the DFA plot.

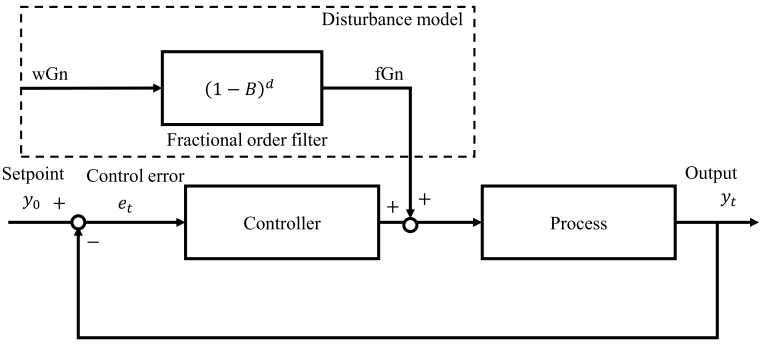


Figure 9.1: The control system under the fractional-order disturbances. The disturbance is modelled by the fractional Gaussian noise (fGn), which is driven by the white Gaussian noise (wGn).

Since the original DFA Hurst fitting method does not consider the LRD (coupling) property of the time series, it should be modified to the multiple-Hurst fitting algorithm

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

with multiple polynomial lines. The fractional-order difference parameter d affects the Hurst exponent H shown in the DFA plot. A larger fractional order d results in a higher H, meaning control becomes more sluggish. For an LRD disturbance with larger d values, CPA should choose the multiple-Hurst fitting results in the various different ranges. The crossover points shift to the smaller range with increasing PID gains in Fig. 9.2. The DFA results in multiple Hurst exponents under different PID tunings, as summarized in Table 9.1. This explains the response of the systems' become

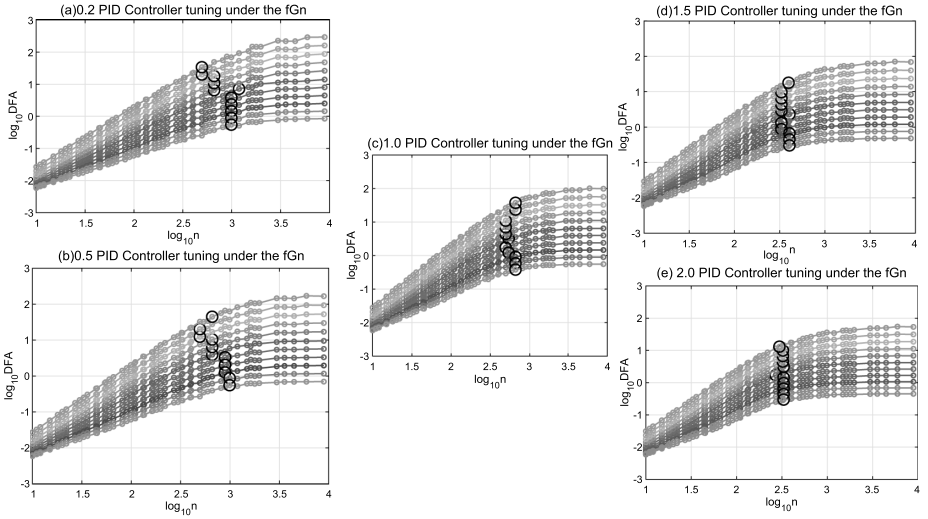


Figure 9.2: DFA plots with crossover points in various PID tunings: From (a) to (e), with the increasing PID gains, the crossover points shift to the small range. In each plot $d \in [-0.5, 0.5]$, the upper fitting line represents the DFA plot with $d = -0.5$, while the lower fitting line is the DFA plot with $d = 0.5$.

Table 9.1: PID tunings results with multiple Hurst exponents.

d	0.2 PID		0.5 PID		1.0 PID		1.5 PID		2.0 PID	
	H_1	H_2	H_1	H_2	H_1	H_2	H_1	H_2	H_1	H_2
−0.5	1.829	0.836	1.791	0.513	1.749	0.360	1.787	0.448	1.817	0.436
−0.4	1.777	0.787	1.777	0.588	1.709	0.326	1.767	0.436	1.775	0.371
−0.3	1.694	0.641	1.721	0.550	1.701	0.398	1.726	0.396	1.738	0.329
−0.2	1.624	0.594	1.627	0.411	1.642	0.368	1.674	0.361	1.693	0.292
−0.1	1.545	0.546	1.553	0.379	1.574	0.340	1.613	0.328	1.631	0.246
0.0	1.421	0.323	1.472	0.332	1.496	0.264	1.525	0.231	1.597	0.257
0.1	1.354	0.346	1.370	0.245	1.417	0.288	1.467	0.271	1.498	0.194
0.2	1.266	0.305	1.283	0.215	1.322	0.237	1.383	0.245	1.420	0.173
0.3	1.177	0.270	1.194	0.188	1.217	0.164	1.280	0.185	1.336	0.154
0.4	1.085	0.241	1.099	0.143	1.129	0.144	1.190	0.165	1.247	0.138
0.5	0.992	0.220	1.008	0.122	1.038	0.125	1.098	0.145	1.154	0.123

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 9.2: Crossover positions at different PID tunings.

d	0.2 PID	0.5 PID	1.0 PID	1.5 PID	2.0 PID
−0.5	2.996, −0.260	2.996, −0.257	2.820, −0.419	2.607, −0.517	2.519, −0.525
−0.4	2.996, −0.048	2.996, −0.058	2.820, −0.231	2.607, −0.347	2.607, −0.347
−0.3	2.996, 0.163	2.950, 0.102	2.820, −0.039	2.607, −0.175	2.607, −0.175
−0.2	2.996, 0.372	2.950, 0.310	2.732, 0.085	2.519, −0.058	2.519, −0.058
−0.1	2.996, 0.580	2.950, 0.520	2.695, 0.237	2.519, 0.117	2.519, 0.117
0.0	3.075, 0.856	2.820, 0.593	2.732, 0.506	2.607, 0.365	2.607, 0.365
0.1	2.820, 0.814	2.820, 0.807	2.695, 0.643	2.519, 0.466	2.519, 0.466
0.2	2.820, 1.032	2.820, 1.018	2.695, 0.844	2.519, 0.640	2.519, 0.640
0.3	2.820, 1.252	2.695, 1.094	2.695, 1.040	2.519, 0.812	2.519, 0.812
0.4	2.695, 1.302	2.695, 1.306	2.820, 1.372	2.519, 0.986	2.519, 0.986
0.5	2.695, 1.529	2.820, 1.651	2.820, 1.573	2.598, 1.254	2.473, 1.116

faster with high PID tunings. From Table 9.2, it clearly shows that the crossover points move towards shorter memories with aggressive tunings. In addition, the crossover points move upwards with the increasing fractional integrated-order filter d, since the Hurst exponents (i. e., the slope of DFA plot) increase accordingly in the small range. The sluggish control relates to the larger crossover, which indicates longer control and longer transient periods.

9.4 Case study: real industry data

9.4.1 Non-stationarity tests

Process industry data show that the majority of the available measurements and process-disturbance time series exhibit nonlinear, non-Gaussian, nonstationary and self-similar properties. The authors have reviewed a great deal of industrial data for the validation of this. Some exemplary time series for real process data are presented in Fig. 9.3.

Visual inspection of the sketched time series shows rapid changes, unexpected significant spikes (outlier values), oscillations and noise. Since first impressions can be deceiving, a statistical study has been conducted. The prepared diagrams show histograms, along with fitted probabilistic density functions (PDF) for three different distributions: classic bell-shaped Gaussian, fat-tailed *alpha*-stable and robust (Huber).

The fractional-dynamics hypothesis may be validated using stationarity tests for one or many samples, estimation of the memory parameter based on sample using wavelets and re-sampling, estimation of the self-similarity index based on sample p-variation or by ARFIMA parameter estimation.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

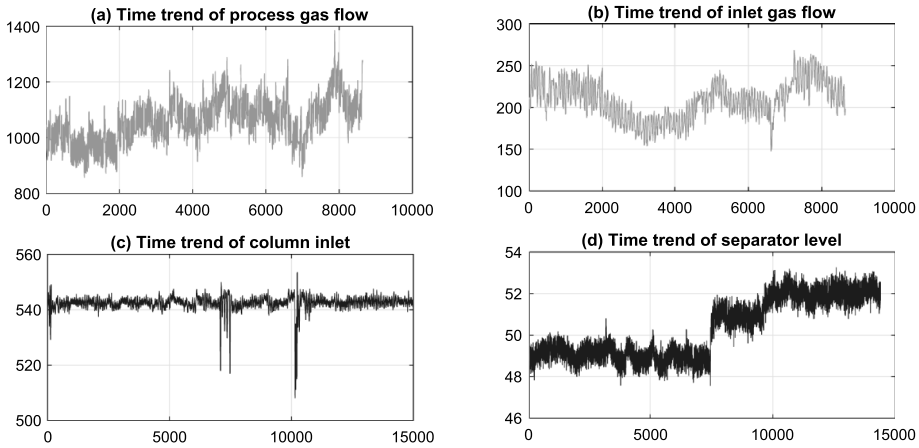


Figure 9.3: Examples of real industrial process variables' time series.

The stationarity tests were run to verify the non-stationarity hypothesis for the considered exemplary time series. Running the augmented Dickey–Fuller test (DF–GLS) and Kwiatkowski–Phillips–Schmidt–Shin test (KPSS) tests over the considered data confirms in all the cases the non-stationarity hypothesis. As these tests are important tools useful in the context of the ARFIMA model, the hypothesis that industrial variables perform similarly to the fractional processes is true. This observation forms a strong rationale for the presented work.

9.4.2 DFA method with multiple Hurst exponents applied to real data

Following the algorithm introduced in Section 9.2, the DFA method with multiple Hurst exponents was applied to the real data in Fig. 9.4. In this regard, it should be noted that apparent crossovers are exhibited for the scaling behavior indicated with black circles.

Figure 9.4(a) shows that the DFA fitting line is straight with no significant crossover point, indicating the consistency of the control performances. From the perspective of control engineers, this control behavior is acceptable. The most interesting result is Fig. 9.4(b). It seems that the DFA fitting plot should be separated into three segments since periodicities or repeated signals may be contained in the original series. This can be observed from the original series plot Fig. 9.3(b).

In Fig. 9.4(c), the overall DFA results are of the convex form, i. e., $H_1 > H_2$, meaning the system controller performance is going stabilize in the long run. In contrast, the badly controlled data set shows a very different crossover pattern. In Fig. 9.4(d), the overall DFA results are of the concave form, i. e., $H_1 < H_2$, meaning the control per-

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

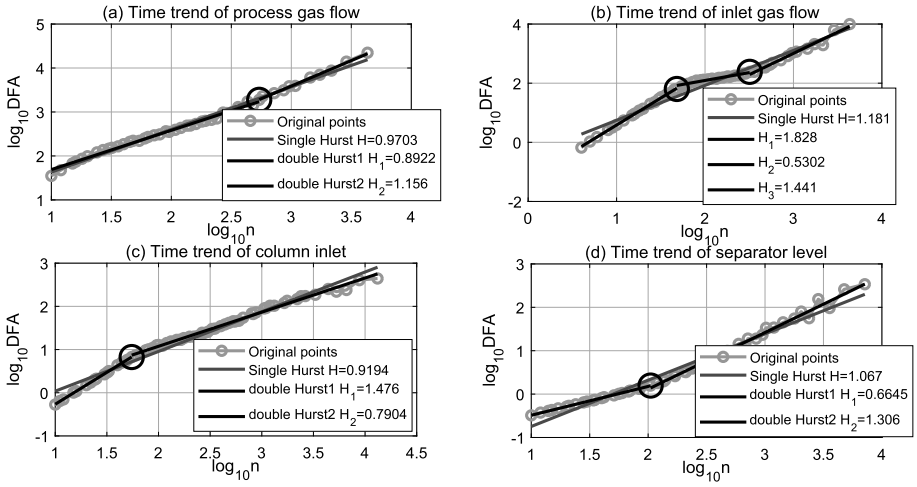


Figure 9.4: DFA method with multiple Hurst exponents applied to real data.

formance is degrading (getting worse). For very short time scales, the fluctuation is quite random since $H \approx 0.5$, while for the larger time scales, the fluctuation becomes smoother and asymptotically approaches Brownian noise, $H \approx 1.5$. These findings are consistent with the previous report in [2]. Thus, the proposed method in this chapter is a good quantitative description that can probably monitor the short-term and long-term of the control system.

9.5 Conclusions

Using FOSP methodology, this chapter analyzes a control system that is subject to a disturbance with fractional dynamics. Multiple Hurst fittings with crossover points are added to the original DFA method. We found that the controller tuning has an impact on the crossover points and Hurst exponents in the simulation performance. The industrial data show that the proposed CPA approach for fractional-order dynamics systems is useful and can be applied in a variety of situations.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bibliography

[1] P. D. Domański. Non-Gaussian properties of the real industrial control error in SISO loops. In *2015 19th International Conference on System Theory, Control and Computing (ICSTCC)*, pages 877–882. IEEE, 2015.

[2] P. D. Domański, Sebastian Golonka, Piotr M Marusak, and Bartosz Moszowski. Robust and asymmetric assessment of the benefits from improved control-industrial validation. *IFAC-PapersOnLine*, 51(18):815–820, 2018.

[3] C. W. J. Granger and R. Joyeux. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–29, 1980.

[4] J. R. M. Hosking. Fractional differencing. *Biometrika*, 68(1):165–176, 1981.

[5] K. Liu, Y. Q. Chen, and P. D. Domański. Control performance assessment of the disturbance with fractional order dynamics. In *Nonlinear Dynamics and Control*, pages 255–264. Springer, 2020.

[6] K. Liu, Y. Q. Chen, P. D. Domański, and X. Zhang. A novel method for control performance assessment with fractional order signal processing and its application to semiconductor manufacturing. *Algorithms*, 11(7):90, 2018.

[7] K. Liu, Y. Q. Chen, and X. Zhang. An evaluation of ARFIMA (autoregressive fractional integral moving average) programs. *Axioms*, 6(2):16, 2017.

[8] C.-K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1):82–87, 1995.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kai Liu, YangQuan Chen, and Paweł D. Domański

10 A novel method for control performance assessment with fractional-order signal processing

Abstract: Control loops are the most significant components in automation systems and subsystems. The essential task for control performance assessment (CPA) is to review and evaluate the performance of the control system. The control system in the semiconductor industry exhibits a complex dynamic behavior, which is hard or even impossible to analyze. One of the challenging problems is that control loops are often multivariable and exhibit nonlinear dynamics, stemming either from the plant, the transducers, the actuators or, even in some cases, the controllers themselves in industrial applications. This chapter investigates the interesting crossover properties of Hurst exponent estimations and proposes a novel method for feature extraction of the nonlinear multi-input multi-output (MIMO) systems. At first, real coupled data from the industry are analyzed by multifractal detrended fluctuation analysis (MFDFA), and the resultant multifractal spectrum is obtained. Second, the crossover points with spline fit in the scale-law curve are located and then employed to segment the entire scale-law curve into several different scaling regions, in which a single Hurst exponent can be estimated accordingly. Third, to further ascertain the origin of the multifractality of control signals, the generalized Hurst exponents of the original series are compared with shuffled data. Finally, non-Gaussian statistical properties, multifractal properties and Hurst exponents of the process control variables are derived and compared with different sets of tuning parameters. The results show that CPA of the MIMO system can be better employed with the help of fractional-order signal processing (FOSP).

Keywords: control performance assessment, fractional-order signal processing, multifractal detrended fluctuation analysis, MIMO, Hurst exponent

10.1 Introduction

In automation systems, control loops are the most important components. Control system efficiency is directly or indirectly related to product quality, operation protection, material and energy use and, thus, financial performance. Control Performance As-

Kai Liu, YangQuan Chen, Mechatronics, Embedded Systems and Automation (MESA) Lab, University of California, Merced, USA, e-mails: liukai5005@gmail.com, ychen53@ucmerced.edu
Paweł D. Domański, Warsaw University of Technology, Institute of Control and Computation Engineering, Warsaw, Poland, e-mail: p.domanski@ia.pw.edu.pl

https://doi.org/10.1515/9783110729122-010

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

sessment (CPA) is a benchmark for identifying and diagnosing the root causes of bad performance, assessing and analyzing current performance and enhancing or preventing performance deterioration [11].

However, the widely held belief that process signals are Gaussian contributes to the foundation of CPA analysis algorithms and methods. On the other hand, non-Gaussian behaviors, such as long-range dependence (LRD), self-similarity, power law autocorrelation, infinite variance, and spiky signals have been observed in a typical industrial process with thousands of control loops.

Non-Gaussian signals and noises tend to produce large-amplitude fluctuations from the average value more frequently than Gaussian ones do [29]. Process control systems are extremely complex, typically spanning several hierarchical levels, making routine maintenance impossible for plant personnel. In industrial applications, for example, control loops are often multivariable and exhibit nonlinear dynamics due to the factory, transducers, actuators or even the controllers themselves.

To address these issues, fractional-order signal processing (FOSP) techniques have been proposed in recent years to better characterize the control process with the notions of fractional calculus [29]. FOSP techniques include fractional-order linear systems, the autoregressive fractional integrated moving average (ARFIMA), Hurst parameter estimation, fractional-order Fourier transformation (FrFT), fractals, multifractal detrended fluctuation analysis (MFDFA), etc. [1]. In addition, fractional-order thinking is indispensable to gain more insights to characterize complex objects [33].

CPA is an important asset-management technology to maintain highly efficient operation performance of automation systems in production plants [10]. There are many classic performance-assessment approaches, such as mean squared error (MSE), integral absolute error (IAE), statistical indexes, fractal indexes, etc. Furthermore, they can vary depending on the system under various circumstances, such as control action restrictions, deterministic disruptions, and setpoint adjustments. Process engineers must determine how to assess the quality of the control system's performance. After some researchers used it with CPA, multifractal analysis has become a hot topic in automation and controls engineering in recent years. The use of non-Gaussian statistical and fractal measures to track on-line control loop output was proposed in [3]. Domański presents results of the research on alternative CPA measures applied to control quality assessment for SISO loop with generalized predictive control (GPC) controller in [2, 4].

However, since the process complexity, cross-dependencies with varying delays, LRD and human factors are not expressed in simulations, the already cited CPA analyses and observations are done with a simple linear SISO scenario, in which the monofractal and multifractal properties are not captured. As a result, various thought-provoking queries, such as multiple-input multiple-output (MIMO) and systems with substantial time variant delays, are left unanswered in the cited papers. Consequently, the roots and explanations of multifractal properties with crossover phenomena need to be investigated further. As a result, more complicated situations,

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

such as nonlinearity, MIMO and processes with large delays, should be investigated to ensure process applicability and effectiveness.

This chapter targets on the new directions of the fractional-order signal processing technique to assess the system performance for hardware engineers such as system engineers, process engineers, reliability engineers and machine-learning engineers, etc. Brief summarized, the purposes of this chapter are:

1. Use MFDFA to analyze semiconductor data, derive the multifractal spectrum and select the characteristic parameters sensitive to changes of the control system;
2. Extract the characteristic parameters from the multifractal spectrum of reference data to form future reference sets;
3. Develop the standard single-Hurst exponent-estimation with the multiple- Hurst exponent-fitting method with crossover points;
4. Select multifractal properties and modified Hurst exponents to distinguish different types of PID tuning performances.

Section 10.1 starts with a brief introduction to the current FOSP techniques in CPA. Section 10.2 provides the definition of the fractional Gaussian noise, Hurst exponent and α-stable distribution for readers with zero knowledge. Section 10.3 proposes the MFDFA algorithm and Hurst spline fit with crossovers. Data analysis with MFDFA for each loop of the real MIMO system is carried out in Section 10.4. The results are analyzed in Section 10.5, followed by the conclusion in Section 10.6.

10.2 Preliminaries

Fractal analysis is especially useful when the data show self-similarity, power-law, scale-invariant and nonlinear properties [5, 27]. From the fractal theory introduced by Mandelbrot in [19], the traditional concept of three-dimensional space can be extended to the fractal (fractional) dimension (FD).

Based on the definitions of fractional-order differential operators, many complex dynamic systems with complex memory behaviors can be more properly described by the fractional calculus. Some researchers have found interesting the analytical results of the linear fractional-order differential equations as represented by the Mittag–Leffler function, which exhibits a power-law asymptotic behavior [13]. Therefore, the fractional calculus is widely used to analyze the random signals with power-law size distributions or a power-law decay of correlations [18, 25].

The fractional-order integral of the function $f(x)$ with $\alpha \in \mathbb{R}+$ is defined as

$$ {}_aD_t^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_a^t \frac{f(\tau)}{(t-\tau)^{1-\alpha}} d\tau, \tag{10.1} $$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

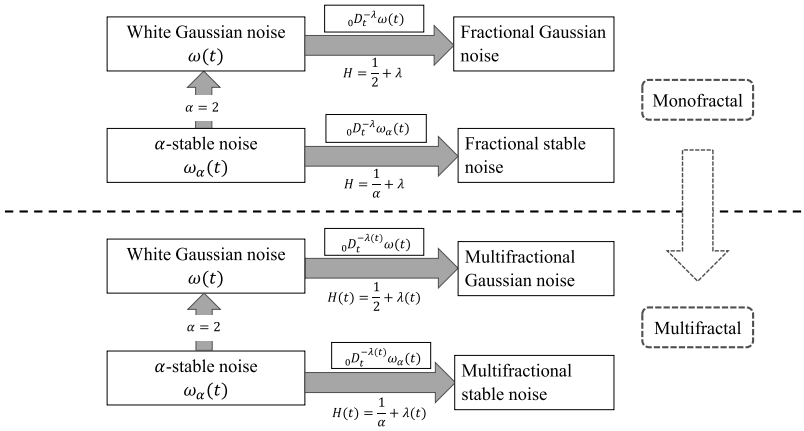


Figure 10.1: Fractional Gaussian noise and fractional Brown noise in monofractal analysis and multifractal analysis.

where $\Gamma(\cdot)$ is the Gamma function and ${}_aD_t^{-\alpha}$ is the fractional integral of order α in $[a,t]$ [25].

Based on the fractional-order integral, the relationship of white Gaussian noise wGn $\omega(t)$, fractional Gaussian noise (fGn), α-stable noise $\omega_\alpha(t)$, fractional stable noise, multifractional Gaussian noise and multifractional stable noise can be described as shown in Fig. 10.1. Besides, the fGn can be expressed as the αth order integration of wGn [20, 29]:

$$X_t = {}_0D_t^{-d}\omega(t). \tag{10.2}$$

Therefore, from the perspective of fractional signals and fractional-order systems, the fGn can be simulated by the dth integrator with wGn as the input in Fig. 10.1.

From that figure, the major difference between monofractal and multifractal properties is the time-dependent variable H. The linear multifractional stable motion is obtained by generalizing the constant Hurst parameter H to a time-dependent local Hölder exponent $H(t)$.

10.2.1 Hurst parameter

LRD is characterized by the Hurst exponent H. This means that there is a strong coupling between values at various different times. The coupling effect characterized by the autocorrelation functions (ACF) and power spectrum density (PSD) obeys the power law, hyperbolically decaying, while the conventional assumptions are based on the exponentially decaying effects. To quantify the level of the coupling, the Hurst

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

exponent and many valuable Hurst exponent estimators have been provided to more accurately characterize the LRD time series.

The rescaled range (R/S) method is one of the time-domain analyses of the Hurst parameter defined as follows [7]:

$$E\left[\frac{R(n)}{S(n)}\right]_{n\to\infty} = Cn^H, \tag{10.3}$$

where $E(\cdot)$ denotes the expected value of the observations, $R(n)$ is the range of the first n values, $S(n)$ is their standard deviation and C is a constant. Alternatively, Whittle's maximum likelihood estimator (MLE) and wavelet analysis use periodogram-based analysis in the frequency domain [26].

LRD processes appear in many contexts, as characterized by the Hurst parameter H ($0 < H < 1$). The phenomenon of LRD can be observed in hydrology, finance, economics, etc. Unlike a stationary process, the autocorrelations between observations of a long memory series slowly decay to zero. However, according to the survey paper [35], Hurst estimators can be significantly affected by trends and seasonality.

LRD or a long-memory process can be defined by the autocovariance $\gamma(k)$ in the time domain or by the power spectrum $P(f)$ in the frequency domain. The power law is observed in the long-memory ARFIMA(p, d, q) process $\{X_t\}$ with $d \in (-1/2, 1/2) \setminus \{0\}$ since the asymptotic behavior of the autocovariance function $\gamma(\cdot)$ is given by [9]

$$\gamma(k) \simeq Ck^{2d-1} \quad (k \to \infty). \tag{10.4}$$

In particular, if $0 < d < 0.5$, then $\{X_t\}$ is a long-memory process, or long- range positive dependence, since $\sum_{k=0}^{\infty} |\gamma(k)| = \infty$ [34]. The process is said to exhibit intermediate memory (anti-persistence) (also known as a mean-reverting process), or long-range negative dependence, for $-0.5 < d < 0$. In addition, the process is nonstationary for $|d| \geq 0.5$ because it possesses infinite variance, and first-order difference is needed to obtain the stationary series. Therefore, the fractional-order difference filter is the first step for processing long-memory ARFIMA models [17].

Besides, the fractional-order difference (integral) coefficient d has a closed relationship with the Hurst parameter and characteristic exponent α:

$$d = H - 1/\alpha, \tag{10.5}$$

where H is the Hurst parameter and α is from the α-stable distribution, which will be introduced in the next subsection. The Hurst parameter of fGn is related to d by Equation (10.5) to identify the long-memory process.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

10.2.2 α-stable Distribution

In real systems, nonstationary, non-Gaussian, spiky signals are usually regarded as outliers and thus discarded by engineers during signal processing. In fact, however, each point of data from the real system could tell a "story',' such as malfunctioning, an inappropriate control philosophy, abnormal human intervention, etc. Non-Gaussian signals and noises tend to produce large-amplitude fluctuations from the average value more frequently than Gaussian ones do. Where did non-Gaussian signals come from? Many researchers have found that they are from fractional calculus or even more specifically, α-stable processes [6].

The α stable distribution, which is the generalization of the Gaussian distribution, is used to model the non-Gaussian distribution. α-stable distribution-based techniques have been applied to describe many natural or man-made phenomena in various fields, such as physics, hydrology, biology, financial and network traffic. Stable distributions provide a useful theoretical tool for this type of signal and noise [23]. The α-stable characteristic function (or distribution) is determined by four parameters: α, β, γ and δ. A univariate distribution function $F(x)$ is stable if and only if its characteristic function has the form

$$\phi(t) = \exp\{j\delta t - \gamma|t|^{\alpha}[1 + j\beta\,\mathrm{sign}(t)\omega(t,\alpha)]\}, \tag{10.6}$$

where

$$\omega(t,\alpha) = \begin{cases} \tan\frac{\alpha\pi}{2}, & \text{if } \alpha \neq 1, \\ \frac{2}{\pi}\log|t|, & \text{if } \alpha = 1, \end{cases} \tag{10.7}$$

$$\mathrm{sign}(t) = \begin{cases} 1, & \text{if } t > 0, \\ 0, & \text{if } t = 0, \\ -1, & \text{if } t < 0, \end{cases} \tag{10.8}$$

and

$$-\infty < \delta < \infty, \quad \gamma > 0, \quad 0 < \alpha \leqslant 2, \quad -1 \leqslant \beta \leqslant 1. \tag{10.9}$$

α is called the characteristic exponent. A small value of α implies a considerable probability mass in the tails of the distribution, that is, the smaller α is, the heavier are the tails. $\alpha = 2$ corresponds to the Gaussian distribution (for any β). γ is a scaling parameter called the dispersion. It is similar to the variance of the Gaussian distribution. β is a symmetry parameter. $\beta = 0$ indicates a distribution symmetric about δ.

There are two algorithms that provide consistent estimators for all four parameters—Koutrouvelis's method, which is based on empirical characteristic function methods of Koutrouvelis [14] and McCulloch's method, which is a simpler and faster,

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

but less accurate, method [21]. In this chapter, we apply the Koutrouvelis method to estimate these parameters in Section 10.4.

10.3 MFDFA algorithm

In early research studies, many researchers were trying to remove periodicity and trends in a time series to determine the true-scale exponents. However, removal of periodicity inevitably leads to unintended modification or a smoothing of the fluctuation. MFDFA is an efficient FOSP technique to detect multifractality in a time series since it does not require any knowledge about the process time delay or other process parameters [22]. It is a well-established method for determining the scaling behavior of noisy data in the presence of trends, without knowing the origin and shape. MFDFA can estimate the multifractal spectrum of the generalized Hurst exponent from a time series and it does not require any knowledge about the process time delay or other process parameters [30]. Currently, MFDFA is being successfully applied to analyze various data, such as hydrographic data [36], wind records [31], financial time series [32], traffic time series [28], mechanical vibration signals [15], etc. It is a powerful tool for uncovering the multifractality of nonstationary time series in the complex systems. In this study, MFDFA is applied to detect the presence of multifractal and monofratal properties in the MIMO systems with external disturbance and noise. Then, the origin of the multifractality of the control signals and the generalized Hurst exponents of the original series are compared with shuffled data.

10.3.1 Basic MFDFA algorithm

Detrended fluctuation analysis (DFA) was first proposed by Peng et al. for DNA analysis in 1995 [24], and then MFDFA was proposed by Kantelhardt et al. in 2002 [12]. For the calculation methodology, Ihlen developed MATLAB code in [8], and Domański modified the algorithm in [2].

The MFDFA method starts with a possibly nonstationary time series $\{e_i\}$ for $i = 1, \ldots, N$, where N indicates its length.

1. Define the "profile" E and transform the original data into mean-reduced cumulative sums,

$$E_j = \sum_{i=1}^{j} (e_i - \bar{e}), \quad j = 1, \ldots, N, \tag{10.10}$$

where $\bar{e}$ is the mean of series, such that the aggregated time series have a zero mean.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. Divide time series E_j into $N_s = \text{int}(N/s)$ non-overlapping segments of equal length s, starting from the beginning. Since the length N of the series is often not a multiple of the considered time scale s, so as not to miss any piece of data, another set of segments starting from the end of data is made from the end progressing to the beginning. As a result, $2N_s$ segments are obtained covering the whole dataset.
3. Calculate the local trend p_k for each of the segments $k = 1, \ldots, 2N_s$ by a least-square fit of the series.
4. Calculate the mean square error $F^2(k,s)$ for the estimate of each segment k of length s.

$$F^2(k,s) = \frac{1}{s}\sum_{i=1}^{s}\left(E\left[(k-1)s+i\right] - p_k[i]\right)^2, \tag{10.11}$$

for each segment $k = 1, \ldots, N_s$ and

$$F^2(k,s) = \frac{1}{s}\sum_{i=1}^{s}\left(E\left[N-(k-N_s)s+i\right] - p_k[i]\right)^2, \tag{10.12}$$

for each segment $k = N_s + 1, \ldots, 2N_s$.
5. Average over all segments to obtain the qth order variance (or fluctuation) function $F_q(s)$ for each size s:

$$F_q(s) = \left(\frac{1}{2N_s}\sum_{k=1}^{2N_s}\left[F^2(k,s)\right]^{q/2}\right)^{1/q}. \tag{10.13}$$

For $q = 0$ use

$$F_0(s) = \exp\left\{\frac{1}{4N_s}\sum_{k=1}^{2N_s}\ln\left(F^2(k,s)\right)\right\}. \tag{10.14}$$

6. Repeat steps (2)–(5) for different s, evaluating new sets of variances $F_q(s)$.
7. Plot $F_q(s)$ for each q in log–log scale and estimate the linear fit with least squares. If slope $h(q)$ varies with q, multifractality is suspected. A single slope indicates monofractal scaling.
8. Calculate multifractal exponent $t(q)$ as

$$t(q) = qh(q) - 1. \tag{10.15}$$

9. Use the Legendre transform to evaluate the Hölder exponent $\alpha(q)$ and multifractal spectrum $f(\alpha)$:

$$\begin{cases}\alpha(q) = H(q) + qH'(q),\\ f(\alpha) = q[\alpha(q) - H(q)] + 1.\end{cases} \tag{10.16}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

The slope H_q of scaling function F_q with various q orders demonstrates the LRD behavior of signals. The multifractal spectrum indicates how dominant are the various fractal exponents present in the series. The width of the singularity spectrum denotes the range of the generalized Hurst (Hölder) exponent, which is defined as

$$\Delta(\alpha) = \max(h_q) - \min(h_q). \tag{10.17}$$

It is often used to quantitatively measure the degree of multifractality of the series. Thus, the wider is the spectrum means, the more multifractality exists inside the series.

10.3.2 Defining the source of the multifractality

The origin of the multifractality of a time series can be distinguished as two different types, i. e., the multifractality due to: (i) the different long-range correlations of the number fluctuations; and due to (ii) the broadness of probability density function (PDF) of the distributions [12].

The easiest way to eliminate the correlations for (i) is shuffling the original series into a random order since the multifractality is due to the probability density, which is not affected by the shuffling. For (ii), the surrogate process of data, defined as replacing the phase of discrete Fourier transform (DFT) coefficients of the original data with a set of pseudo independent and identically distributed quantities in $(-\pi, \pi)$, can change the broad PDF of the original data into the Gaussian distribution, while seldomly destroying the intrinsic long-range correlations of the original data [12].

10.3.3 Plot fitting Hurst exponents with crossovers

Generally, the conventional R/S plot fitting is performed by Equation (10.3). The standard estimation with least squares using QR factorization is used for single Hurst exponent estimation. In the current study, the curve is assumed to be piece-wise linear, which is achieved by solving the first-order spline least squares (LSQ) fitting problem with graph $\log(F_q)$ versus $\log(s)$ identifying crossovers. Each plot includes one, two, and three linear approximations. In the next section, the red line represents a single line and estimation of one memory scale, which is the conventional single-Hurst exponent H_1. The green line represents two memory scales H_1, H_2 with one crossover. The blue line represents three memory scales H_1, H_2, H_3 separated by two crossovers. The short-term is important from the perspective of the controller tuning, while the long-term scale indicates the control system structure. Therefore, the asymptotic and transient behavior of the system can be characterized with multiple Hurst exponents.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

10.4 Case studies

In today's increasingly competitive global marketplace, semiconductor manufacturers are facing a slew of high-level production challenges to remain profitable. Key among these challenges are yield and throughput optimization. Big-data solutions will be critical in the industry as we scale Advanced Process Control (APC) solutions to finer levels of control and diagnostics. The main impact, however, will be to better enable more efficient predictive technologies like Predictive Maintenance (PdM), Virtual Metrology (VM), run-to-tun (R2R) control and yield prediction, which all rely on data from traditional APC capabilities like fault detection [16].

The challenging parts in front of the big data solution are:

1. Semiconductor products are characterized by rapid changes in both the improvement and deterioration of the quality.
2. The process is nonlinear and nonstationary with unpredictable behaviors.
3. The repeatable or R2R property of the high-power plasma etching makes it difficult to analyze the real-time data

This chapter will show how to use Hurst exponents as the temperature control system's main index/matrix. For each recipe, tens of thousands of runs are conducted overnight, and they may be subject to unidentified interventions. The etching process relies on the precise regulation of coolant temperature to ensure wafer uniformity. Data from hundreds of control loops was secretly selected from the phase of dynamic coolant-supply temperature. The industrial part of the research is focused on real data from semiconductor manufacturing etching equipment.

For thousands of data logs from different chambers, how to assess the control performance is a thorny problem for process engineers in the semiconductor industry. In addition, the tuning of the MIMO system is a challenging task because the optimization of PID parameters can be unachievable under coupled loops and nonlinearities. In Fig. 10.2, $r_{1,2,3}$, $y_{1,2,3}$ are set-points and outputs of the control system, respectively. Accordingly, the three PID controllers $C_{1,2,3}$ are designed to control the temperature, flow and coolant level in the tank. Therefore, it is a MIMO system with internal coupling loops and external disturbance and noise with significant delays. In the article,

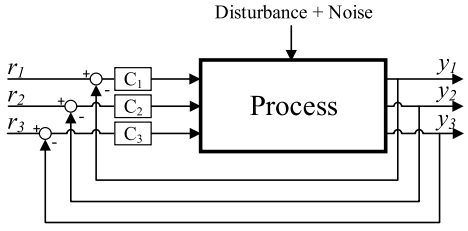


Figure 10.2: Control structure of the MIMO system.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

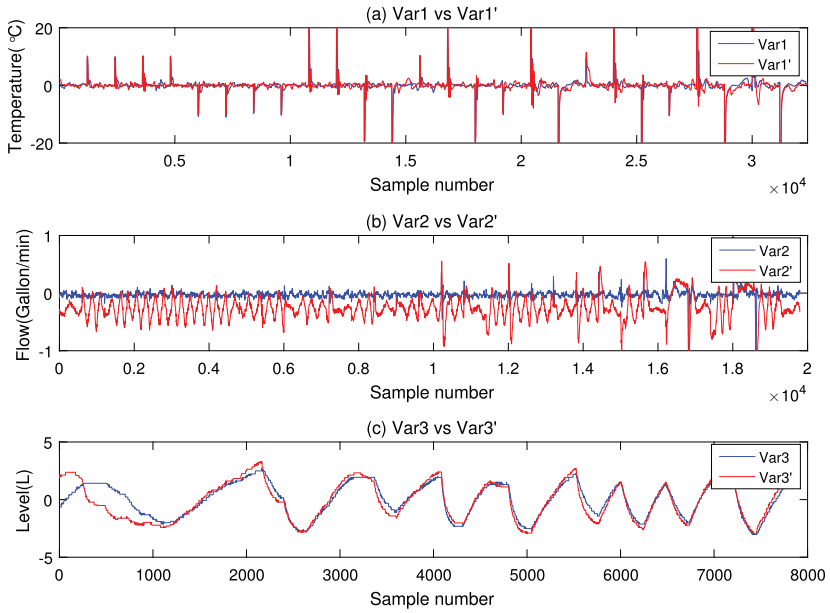


Figure 10.3: Output errors of three loops: `Var1`, `Var2`, `Var3` with first tuning parameters in blue; `Var1'`, `Var2'`, `Var3'` with second tuning parameters in red.

we use `Var1`, `Var2`, `Var3` to represent the error signals ($r_{1,2,3} - y_{1,2,3}$) from each loop as depicted in Fig. 10.3.

10.4.1 Non-Gaussian statistical analysis

First, the histogram plot `Var1`, `Var2`, `Var3` with a probabilistic distribution function (PDF) fitting is frequently carried out to detect if the data are impacted by non-Gaussian noises or subject to human interventions. This is often used to measure the heavy-tailedness underlying process during assessment, and frequently it is a suitable moment to call for further insight. Histogram fitting is significant since some parts of the process: can be more impacted by the disturbances or can be more exposed to the process nonstationarity; can be cross-correlated; or might be subject to human interventions. In real situations it is a challenge to have similar properties for all the loops.

The etching process during semiconductor manufacturing requires uniform temperature behaviors from batch to batch, therefore `Var1` is the major concern in the current MIMO system. As introduced in Section 10.2.2, the α-stable distribution is performed to detect whether the distributions are Gaussian, Lévy or something else. In Fig. 10.4, the time series `Var1` of the MIMO system has significant fat-tail properties. Therefore, it can be inferred that there are non-Gaussian noises entering the system.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

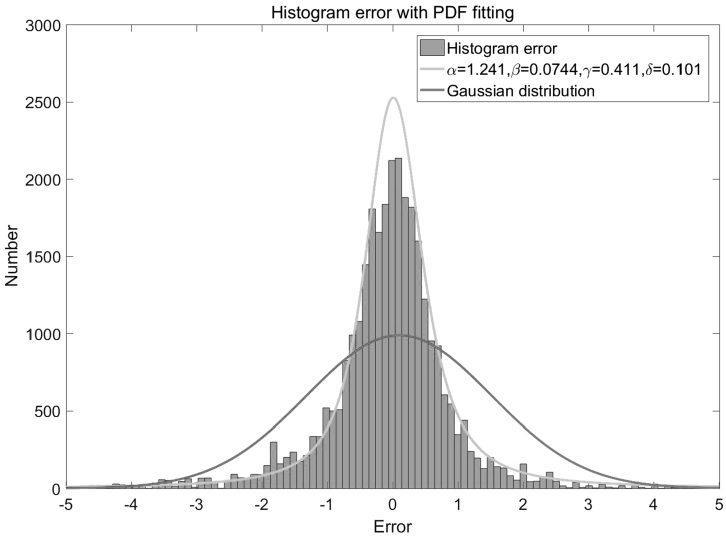


Figure 10.4: Histogram of `Var1` with PDF fitting.

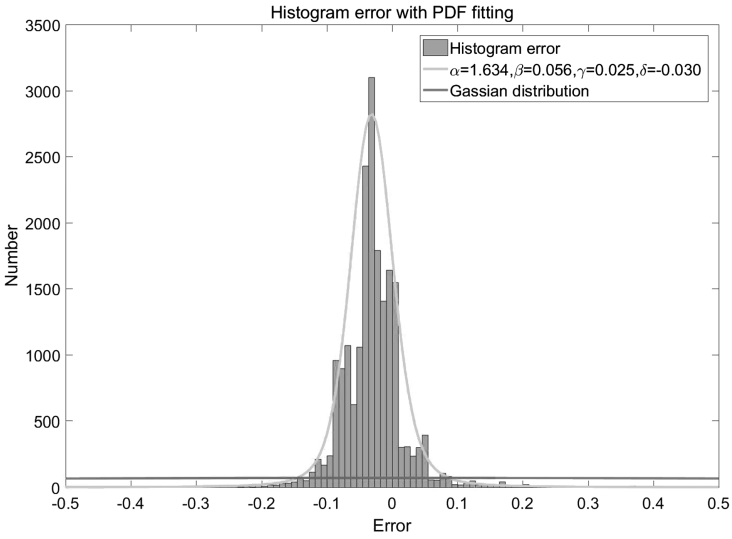


Figure 10.5: Histogram of `Var2` with PDF fitting.

`Var2` is the variable of errors from the flow loop. The higher flow rate can exchange heat efficiently, whereas more turbulence could be introduced at the same time. Therefore, the flow loop is interacted with the temperature loop. From Figs. 10.4 and 10.5, we can see the clear correlations between `Var1` and `Var2`.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

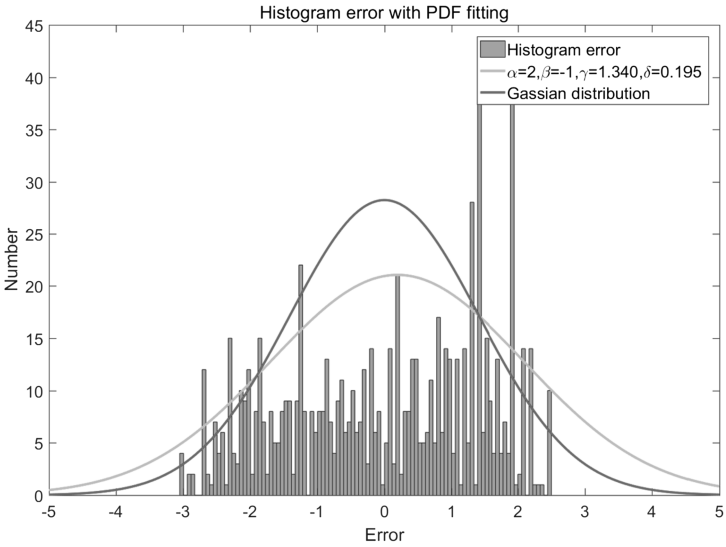


Figure 10.6: Histogram of `Var3` with PDF fitting.

The histogram of the `Var3` (see Fig. 10.6) explicitly shows very poor or even not functioning control. The histogram is highly scattered, almost uniform. The loop is uncontrolled, with possible oscillations, operating on the edge of the stability region or manually operated.

10.4.2 Hurst exponents fitting with crossovers

Second, in Figs. 10.7, 10.8 and Fig. 10.9, the red line represents single memory-scale estimation, i. e., one Hurst exponent fitting without crossover H_1. The green line represents two memory scales $H_{1,2}$ with one crossover. The blue line represents three memory scales $H_{1,2,3}$ with two crossovers. All the crossover points are marked in circles. The phenomenon of the multiple memory scales in `Var1` and `Var2` are observed in Figs. 10.7 and 10.8. Therefore, the variation of the memory scales with crossovers cannot be captured by the conventional Hurst R/S fitting method with the constant value. After processing the data with MFDFA, the multiple Hurst exponents, along with crossover point positions, might carry information about the multifractality of the control performance.

`Var3` is the variable of errors from the coolant level loop, which in reality is designed to keep the coolant away from the alarm level. It should be noted that the control priority is comparatively lower in the design of the MIMO system. Compared with `Var1`, `Var2`, however, the monofractal behavior of `Var3` can be validated through Figs. 10.6, 10.9 and Fig. 10.12.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

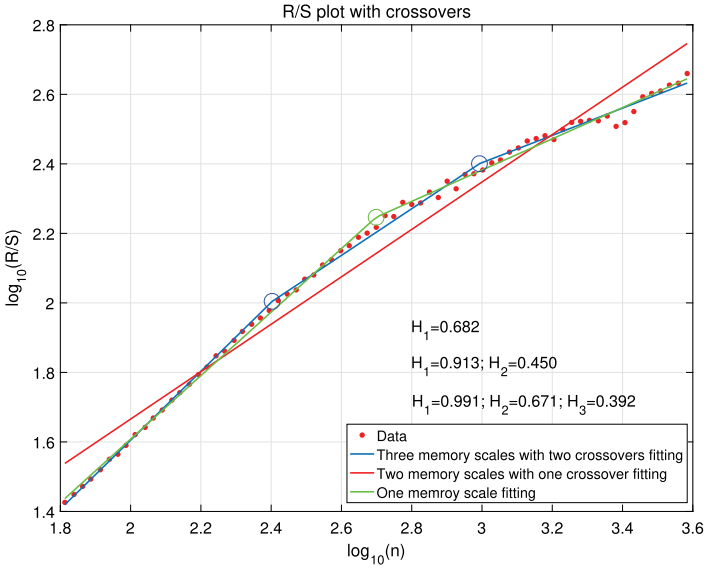


Figure 10.7: R/S plot `Var1` with crossovers.

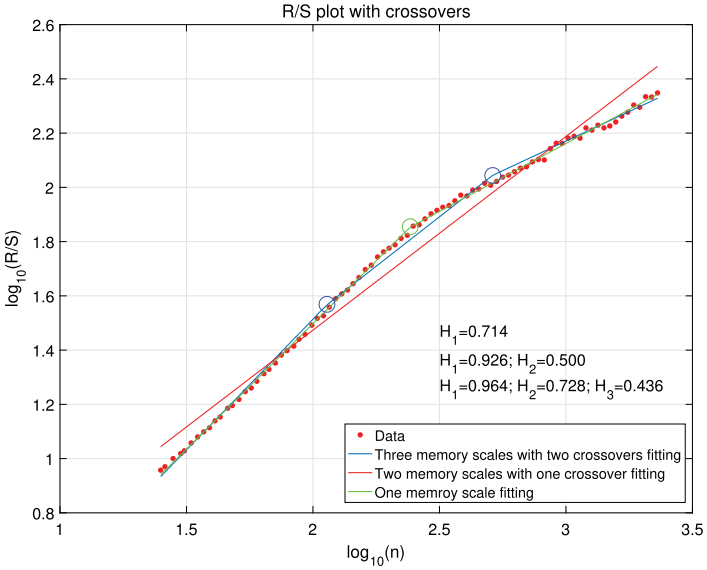


Figure 10.8: R/S plot `Var1` with crossovers.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

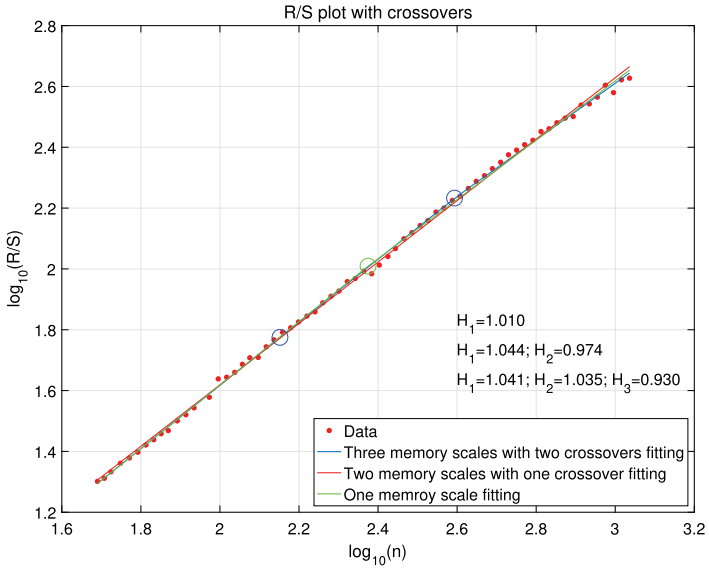


Figure 10.9: R/S plot `Var1` with crossovers.

Some further conclusions about the control quality of the loops could be reached. `Var1` and `Var2` show multi-persistence behavior with different Hurst exponents in different memory scales. On the contrary, `Var3` is clearly mono-persistent. This means that this loop isn't necessarily connected with the others. It has a Hurst exponent of approximately 1. Its tuning is extremely slow, to the point of being nonexistent. It validates the initial observation made with the histogram. The other two loops are the ones that are most likely to be closely coupled. The R/S plots and the shortest Hurst exponent also show dynamically slow power, but the control improves and returns to neutral on the longer memory scale.

10.4.3 Multifractal analysis

To clearly derive the origin of the multifractal property, the shuffled datasets of `Var1`, `Var2`, `Var3` can remove correlations from the data, and any remaining scaling is caused by probability density function broadness. It is shown that the multifractal behaviors can be captured by the multifractal spectrum analysis with the shuffled datasets in Fig. 10.10.

From the multifractal spectrs shown in Figs. 10.10, 10.11 and 10.12, the multifractal spectrum width is significantly different. Notice that the tiny arcs in Fig. 10.12 demonstrate the constant H_q for monofractal property in `Var3`, while the wider arcs of `Var1`, `Var2` shows the multifractality in the MIMO system. Moreover, the multifractality

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

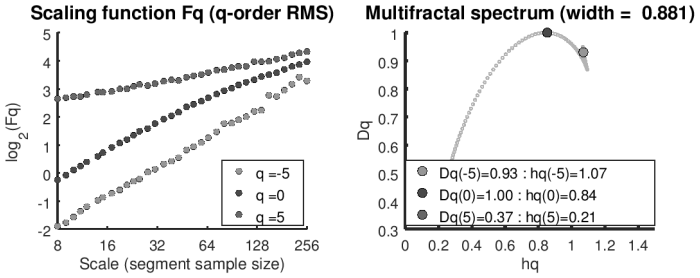


Figure 10.10: Multifractal spectrum for shuffled `Var1`. The nonparallel behavior with different scales (left plot) and the arch with large width (right plot) indicate the multifractal behaviors of `Var1` in the MIMO system.

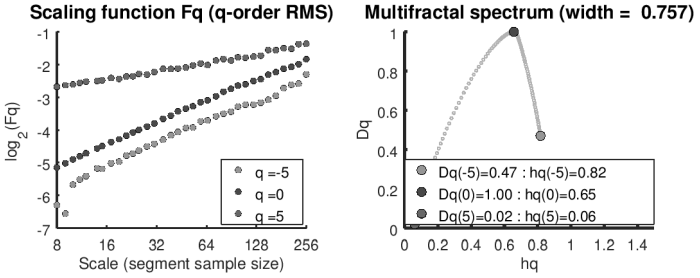


Figure 10.11: Multifractal spectrum for shuffled `Var2`. The nonparallel behavior with different scales (left plot) and the arch with large width (right plot) indicate the multifractal behaviors of `Var2` in the MIMO system.

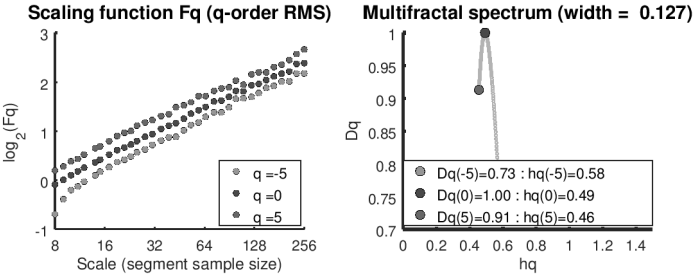


Figure 10.12: Multifractal spectrum for shuffled `Var3`. The nonparallel behavior with different scales (left plot) and the arch with large width (right plot) indicate the multifractal behaviors of `Var3` in the MIMO system.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

of the MIMO process has been detected and analyzed for each loop with shuffled datasets `Var1, Var2, Var3`.

10.5 Results and discussion

This chapter presented a CPA method with non-Gaussian statistical factors. First, we extracted four parameters $\alpha, \beta, \gamma, \delta$ with α stable distribution and calculated the multifractal spectrum width $\Delta(\alpha)$.

It is well known that the PID tuning of the MIMO system is a time consuming and arduous work. Two sets of real data `Var1, Var2, Var3` and `Var1', Var2', Var3'` with different PID parameters are compared at the end of article, of which the initial settings are conservative and latter ones are more aggressive. In Table 10.1, the proposed FOSP method was conducted according to the previous sections. The aggressive tuning with small α depicts the broadness and fat tails of the distribution of the underlying process. Moreover, the absolute values of β, γ, δ indicate the poorer behavior with the latter tuning.

Table 10.1: Indexes for variables with various PID tunings.

Variables	α	β	γ	δ	$h_q(-5)$	$h_q(5)$	$\Delta(\alpha)$
Var1	1.241	0.074	0.411	0.101	1.07	0.21	0.881
Var2	1.634	0.056	0.025	−0.030	0.82	0.06	0.757
Var3	2.000	−1.000	0.456	−0.235	0.58	0.46	0.127
Var1'	1.171	−0.156	0.456	−0.235	1.04	0.25	0.798
Var2'	1.695	0.287	0.123	−0.246	0.71	0.41	0.307
Var3'	2.000	1.000	1.567	−0.208	0.57	0.53	0.069

In addition, the multi-scale Hurst spline fits with crossovers are located and compared with single-Hurst estimation. Using multiple-Hurst exponent fittings at different scales represents the dynamic behavior of the system with noise in Table 10.2.

A review of the persistence properties of the data reveals visible effects of tuning. We observe changes in H_1. For the first loop it moved from a more sluggish value of 0.682 towards neutral tuning, while the second loop `Var2` is slightly more sluggish. Loop `Var3` is unchanged. However, the effect of tuning (whatever the set is) is evident in comparison with the original MIMO system presented in R/S plots in Figs. 10.7, 10.8 and 10.9. `Var1, Var2` require disturbance decoupling compared with aggressive tuning, while `Var3` possibly requires redesign of the control philosophy and fine tuning.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 10.2: Hurst exponents at different memory scales.

Variables	H_1	H_1, H_2	H_1, H_2, H_3
Var1	0.682	0.913, 0.450	0.991, 0.671, 0.392
Var2	0.714	0.926, 0.500	0.964, 0.728, 0.436
Var3	1.010	1.044, 0.974	1.041, 1.035, 0.930
Var1′	0.622	0.827, 0.416	0.905, 0.560, 0.415
Var2′	0.824	1.099, 0.546	1.062, 0.951, 0.350
Var3′	0.968	1.057, 0.878	1.006, 1.066, 0.751

There is no single universal measure that manages and detects all possible features that may be encountered in industrial control, as is well documented. Each approach is well suited to the loop's unique characteristics. Traditional methods, such as integral, mathematical and minimal-variance methods, can be used in a broad range of situations. They can, however, encounter difficulties when dealing with nonlinear and complex processes. The FOSP approach allows for the evaluation of controller output, which is dependent on controller tuning and control theory suitability. It also allows the control error histogram's stable properties to be incorporated into the analysis. Due to the inefficiency of distribution fitting, the multifractal spectrum may reveal additional details about control properties. The overall system's complex cascaded properties, human effects or superposition of different impacts and oscillations (nonconcave spectra) can all be evaluated.

10.6 Conclusions

Fractional-order system dynamics can be seen in the multifractal properties of process control variables. We discussed the origins of multifractal properties and how FOSP techniques contribute to complex phenomena in the control system in this chapter. More importantly, the CPA of the MIMO control system is first proposed using FOSP and then applied to semiconductor manufacturing using data from the industry. Indicators for control improvement are established, and conclusions about the MIMO system's control output are then drawn. The FOSP analysis allows for analysis on various time scales, allowing for controller dynamics (short scale) and device cross-dependencies to be addressed (mostly reflected in longer scales). Similar phenomena have previously been studied in time-delayed systems and networks, allowing for the evaluation of both system parameterization and structure. As a result, it is expected that FOSP will take on a growing number of specific roles that traditional signal processing approaches will be unable to address.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bibliography

[1] Y. Q. Chen, R. Sun, and A. Zhou. An overview of fractional order signal processing (FOSP) techniques. In *Proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pages 1205–1222. American Society of Mechanical Engineers, 2007.

[2] P. D. Domański. Multifractal properties of process control variables. *International Journal of Bifurcation and Chaos*, 27(06):1750094, 2017.

[3] P. D. Domański. On-line control loop assessment with non-Gaussian statistical and fractal measures. In *Proceedings of the American Control Conference (ACC)*, pages 555–560. IEEE, 2017.

[4] P. D. Domański and M. Ławryńczuk. Assessment of predictive control performance using fractal measures. *Nonlinear Dynamics*, 89(2):773–790, 2017.

[5] J. Feder. *Fractals*. Springer Science & Business Media, 2013.

[6] R. Gorenflo and F. Mainardi. Fractional calculus and stable probability distributions. *Archives of Mechanics*, 50(3):377–388, 1998.

[7] H. E. Hurst. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–808, 1951.

[8] E. A. F. Ihlen. Introduction to multifractal detrended fluctuation analysis in Matlab. *Frontiers in Physiology*, 3:141, 2012.

[9] A. Inoue. Asymptotic behavior for partial autocorrelation functions of fractional ARIMA processes. *Annals of Applied Probability*, 12(4):1471–1491, 2002.

[10] M. Jelali. An overview of control performance assessment technology and industrial applications. *Control Engineering Practice*, 14(5):441–466, 2006.

[11] M. Jelali. *Control Performance Management in Industrial Automation: Assessment, Diagnosis and Improvement of Control Loop Performance*. Springer Science & Business Media, 2012.

[12] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, 316(1):87–114, 2002.

[13] A. A. Kilbas, M. Saigo, and R. K. Saxena. Generalized Mittag-Leffler function and generalized fractional calculus operators. *Integral Transforms and Special Functions*, 15(1):31–49, 2004.

[14] I. A. Koutrouvelis. Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association*, 75(372):918–928, 1980.

[15] J. Lin and Q. Chen. Fault diagnosis of rolling bearings based on multifractal detrended fluctuation analysis and Mahalanobis distance criterion. *Mechanical Systems and Signal Processing*, 38(2):515–533, 2013.

[16] K. Liu, Y. Q. Chen, T. Zhang, S. Tian, and X. Zhang. A survey of run-to-run control for batch processes. *ISA Transactions*, 83:107–125, 2018.

[17] K. Liu, Y. Q. Chen, and X. Zhang. An evaluation of ARFIMA (autoregressive fractional integral moving average) programs. *Axioms*, 6(2):16, 2017.

[18] F. Mainardi. *Fractional Calculus and Waves in Linear Viscoelasticity: An Introduction to Mathematical Models*. World Scientific, 2010.

[19] B. B. Mandelbrot. *Fractals*. Wiley Online Library, 1977.

[20] B. B. Mandelbrot and J. W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10(4):422–437, 1968.

[21] J. H. McCulloch. Simple consistent estimators of stable distribution parameters. *Communications in Statistics-Simulation and Computation*, 15(4):1109–1136, 1986.

[22] M. S. Movahed, G. R. Jafari, F. Ghasemi, S. Rahvar, and M. R. R. Tabar. Multifractal detrended fluctuation analysis of sunspot time series. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(02):P02003, 2006.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[23] C. L. Nikias and M. Shao. *Signal Processing with Alpha-Stable Distributions and Applications*. Wiley-Interscience, 1995.
[24] C.-K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1):82–87, 1995.
[25] I. Podlubny. *Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications*. Elsevier, 1998.
[26] G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, volume 1. CRC press, 1994.
[27] D. T. Schmitt and P. Ch. Ivanov. Fractal scale-invariant and nonlinear properties of cardiac dynamics remain stable with advanced age: A new mechanistic picture of cardiac control in healthy elderly. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 293(5):R1923–R1937, 2007.
[28] P. Shang, Y. Lu, and S. Kamae. Detecting long-range correlations of traffic time series with multifractal detrended fluctuation analysis. *Chaos, Solitons & Fractals*, 36(1):82–90, 2008.
[29] H. Sheng, Y. Q. Chen, and T. S. Qiu. *Fractional Processes and Fractional-Order Signal Processing: Techniques and Applications*. Springer Science & Business Media, 2011.
[30] B. Srinivasan, T. Spinner, and R. Rengaswamy. Control loop performance assessment using detrended fluctuation analysis (DFA). *Automatica*, 48(7):1359–1363, 2012.
[31] L. Telesca and M. Lovallo. Analysis of the time dynamics in wind records by means of multifractal detrended fluctuation analysis and the Fisher–Shannon information plane. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(07):P07001, 2011.
[32] Yu. Wang, Y. Wei, and C. Wu. Analysis of the efficiency and multifractality of gold markets based on multifractal detrended fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, 390(5):817–827, 2011.
[33] B. J. West. *Fractional Calculus View of Complexity: Tomorrow's Science*. CRC Press, 2016.
[34] W. A. Woodward, H. L. Gray, and A. C. Elliott. *Applied Time Series Analysis with R*, 2nd edition. CRC Press, 2016.
[35] X. Ye, X. Xia, J. Zhang, and Y. Q. Chen. Effects of trends and seasonalities on robustness of the Hurst parameter estimators. *IET Signal Processing*, 6(9):849–856, 2012.
[36] Q. Zhang, C.-Y. Xu, Y. D. Chen, and Z. Yu. Multifractal detrended fluctuation analysis of streamflow series of the Yangtze River basin, China. *Hydrological Processes*, 22(26):4997–5003, 2008.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Michał W. Okoński and Paweł D. Domański

11 Study on oscillation detection robustness and outlier filtering impact

Abstract: Accurate oscillation analysis is of vital importance to industrial plants' profitability and safety. Most of the oscillation detection and characterization procedures are based on strict assumptions regarding signal and noise properties. Their robustness to outliers is seldom validated and considered. This paper proposes an improved oscillation detection and characterization procedure that is robust to outliers; the robustness to single-point outlying observations (deviants), of endogenous or exogenous origin is considered. The proposed improvements use initial signal preprocessing to identify and replace the outliers. In the first part of the paper, the Hampel filter's ability to detect outliers in oscillating signals is evaluated. The best-performing Hampel filter is selected in a thorough simulation study. Furthermore, the limit to a MA filter's window size, in the context of oscillation characterization, is established. In the second part of the paper, the original oscillation detection procedure's robustness to outliers, with and without initial signal preprocessing, is assessed on simulated examples. A comparison of preprocessing methods reveals that Hampel filter is superior to MA filter. Finally, the proposed improvements are validated on various simulated benchmark signals and real industrial case studies.

Keywords: oscillation detection and characterization, outlier identification, exogenous and endogenous outliers, EMD, data preprocessing

11.1 Introduction

The oscillation detection and characterization procedure is a vital step in control system performance assessment. Its results serve both process tuning and diagnosis and thus have far-reaching influence on a plants' profitability and safety. The unstated assumption behind most modern oscillation detection procedures is that the analyzed datasets exhibit Gaussian properties. The reality however, hardly matches the theory. The majority of industrial data is driven by nonlinear and nonstationary processes exhibiting abnormal properties [4], of which the most troublesome and distinct are the outliers. This paper addresses the issue of outliers in the context of the oscillation detection Hampel filter and its characterization.

Michał W. Okoński, Paweł D. Domański, Warsaw University of Technology, Institute of Control and Computation Engineering, ul. Nowowiejska 15/19, 00-665 Warsaw, Poland, e-mails: michal.wiktor.okonski@gmail.com, pawel.domanski@pw.edu.pl

https://doi.org/10.1515/9783110729122-011

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Over the years, multiple oscillation detection and characterization procedures have been developed. In general, these methods may be classified into two main categories: based on time and frequency domain analysis or based on data mining. The classic time and frequency domain techniques include: detection of spectral peaks in either Fourier spectrum or a periodogram, the integral of absolute error (IAE), decay-ratio approach of autocorrelation function (ACF) [14], the regularity of zero-crossings of ACF [25], wavelets [20] and the discrete cosine transform [13]. More recently, multiple approaches based on data mining were proposed: hidden Markov models [29], artificial neural networks (ANN) [31] or linear predictive coding (LPC) analysis [19].

The aforementioned methods suffer from certain limitations; Their successful operation is either based on strong assumptions, namely, signal linearity and stationarity, or requires large amounts of data. This is why the various time–frequency domain hybrid methods have been gaining researchers' attention. These methods start with an initial step of signal decomposition that splits the signal (information) from noise and nonstationarity (trend). Such signal decomposition methods are: empirical mode decomposition (EMD) [9], local mean decomposition (LMD) [21] and ensemble empirical mode decomposition (EEMD) [27]. The oscillation characterization procedures based on these methods are: EMD with Thornhill's index [22] with numerous further enhancements [15, 24], LMD with Lempel–Ziv complexity [28] and EEMD [26].

The research on oscillation detection and characterization rarely acknowledges the need to account for the distortions in the analyzed data. White-noise robustness is usually the only considered aspect: as in noise-assisted multivariate EMD [1] or the previously mentioned methods based on ACF. Specific non-Gaussian properties of the signal, especially "fat tails", are seldom considered.

An outlier is a strange occurrence in a dataset [7]. In general, two types of outliers can be distinguished [6]: *deviant*, i. e., a single outlying observation that might be frequent, but still appears as an individual observation, and *discord* in the form of an anomalous time-series subsequence substituting for the original process over some given time period.

Outliers may originate from an erroneous observation, an unknown contaminating mechanism or can be an intrinsic symptom of the main underlying stochastic process [11]. Outliers strongly influence data analysis: They enlarge the observed variance and reduce the power of statistical tests [16], destroy Gaussianity, introduce histogram tails [23] and bias regression [18].

Several works propose an initial data preprocessing step that is meant to account for the distorted data, i. e., identification of corrupted data for high-frequency oscillations detection in EEG data [8], a Gaussian M-Dist median filter applied to data preprocessing for detecting low-frequency oscillations in power systems [12] or time-series filtering to reduce false alarms in power-system data [2]. This paper proposes an improved robust oscillation detection procedure initially described in [22]. The rationale behind this work originates from the authors' industrial experience: Outliers in signals are frequent, and process analysis needs to be performed using highly contaminated

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

raw process data. The proposed improvements are based on initial signal preprocessing that is meant to identify and replace the outliers. Although this work focuses on the first type of outliers previously mentioned—individual observations in a time series—it makes a clear distinction between outlier sources and their generation mechanisms. We believe that the generic nature of this study is a vital extension of subject knowledge.

The remainder of this paper is organized as follows. The outlier generation mechanisms and signal preprocessing methods are described in Section 11.2. The proposed improvements and their benchmarking are presented in Section 11.3. Section 11.4 provides a comprehensive simulation study of preprocessing methods and their influence on the performance of the oscillation characterization procedure. Industrial case studies are discussed in Section 11.5, which is being followed by a summary in Section 11.6.

11.2 Preliminaries

This section is structured in a way that resembles the consecutive steps of experiments discussed in Section 11.4. First of all, noise carrying outliers need to be generated. Outlier detection in noise, which will serve during Hampel filter's benchmarking, is followed by the initial signal preprocessing, signal decomposition and the original oscillation characterization procedure.

11.2.1 Outlier generation

As mentioned in Section 11.1, points lying far from the bulk of the data can have different origins. When an underlying generation process has fat-tailed properties, these points are classified as endogenous outliers. If they are a result of erroneous observation, they are referred to as exogenous outliers. The methods used to simulate these two distinct types of outliers are described next.

11.2.1.1 Endogenous outliers

Outliers generated by industrial fat-tailed processes can be simulated using noise generated by α-stable distribution since it exhibits many attractive features [5]. It is expressed by the characteristics equation (11.1)

$$F^{\text{stab}}_{\alpha,\beta,\delta,\gamma}(x) = \exp\{i\delta x - |\gamma x|^{\alpha}(1 - i\beta l(x))\}, \tag{11.1}$$

$$l(x) = \begin{cases} \operatorname{sgn}(x)\tan(\frac{\pi\alpha}{2}) & \text{for } \alpha \neq 1, \\ -\operatorname{sgn}(x)\frac{2}{\pi}\ln|x| & \text{for } \alpha = 1, \end{cases}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

where: $0 < \alpha \leq 2$ is the *stability* index, $|\beta| \leq 1$ is *skewness*, $\delta \in \mathbb{R}$ is *location*, and $\gamma > 0$ is the *scale* parameter. Selected combinations of these parameters result in a closed form of the probabilistic density function (PDF):

- $\alpha = 2$ and $\beta = 0$ denote Gaussian distribution;
- $\alpha = 1$ and $\beta = 0$ denote Cauchy distribution;
- $\alpha = 0.5$ and $\beta = \pm 1$ denote Lévy distribution.

In this paper, symmetric α-stable (SαS) noise, with $\alpha < 2$, $\beta = 0$ and $\delta = 0$, is used to generate endogenous outliers.

11.2.1.2 Exogenous outliers

The second type of outliers is needed to describe signal contamination caused by erroneous observations such as measurement errors, data transfer artifacts or human interventions. These outliers are simulated by white noise (11.2) with randomly injected distant points (11.3).

$$F^{\text{Gauss}}_{x_0,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x_0)^2}{2\sigma^2}}, \tag{11.2}$$

$$F(\pm D) = p, \tag{11.3}$$

where: x_0 is location (mean), σ is scale (standard deviation), p is the contamination level and D is the scale of exogenous outliers.

11.2.2 Outlier detection

The research on outlier detection can be traced back to early 1950s. It started with the use of statistical methods, whereas nowadays data-mining approaches are gaining popularity. Despite the recent advancements, the simplicity, clearness and model-free formulation of statistical techniques make them still attractive.

A classic outlier detection method is based on the Interquartile Range (IQR). The IQR is calculated as the difference between upper 75th (denoted as Q_3) and lower 25th (denoted as Q_1) percentiles of the dataset. Data points are considered to be outliers when they lie below $Q_1 - 1.5\,\text{IQR}$ or above $Q_3 + 1.5\,\text{IQR}$. The method's results are often presented in a box-plot: The highest and lowest occurring values are indicated by whiskers, while the outliers are highlighted by ticks. The breakdown point for IQR is equal to 25 %. The IQR method is especially useful for detecting outliers in bulks of data, such as real or simulated noise realizations.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

11.2.3 Signal preprocessing

11.2.3.1 Applied estimators

The choice of appropriate statistical estimators of location and scale is of importance in the context of outlier identification using nonlinear identifiers. A classic approach assumes Gaussian properties of the noise-generating process. Under such assumptions, the sample mean x_0 estimates location, while the sample standard deviation σ estimates the scale.

Unfortunately, these estimators are biased by outliers and are nonrobust [18]. Their breakdown point, i. e., the maximum allowable share of contaminated data not affecting estimation, equals 0 %. Since industrial datasets are commonly contaminated with outliers, other robust location and scale estimators need to be used. The sample median is a simple robust location estimator; it points to the middle of the dataset of length N. The median absolute deviation (MAD) is a simple robust scale estimator (11.4):

$$\mathrm{MAD} = \underset{k}{\mathrm{median}}\,|\epsilon_k| = \underset{k}{\mathrm{median}}\,|x_k - x_0|. \tag{11.4}$$

To refine MAD, the median can be used as the expected value estimator. Median absolute deviation around median (MADAM) is obtained then (11.5) as

$$\mathrm{MADAM} = \underset{k}{\mathrm{median}}\,|x_k - \overline{x}_{\mathrm{med}}|. \tag{11.5}$$

The research on robust statistics offers other, more elaborate robust estimators, namely, M-estimators with Huber ψ-function or logistic function [3]. M-estimator of location is defined as the solution of the following equation:

$$\sum_{i=1}^{N} \psi\left(\frac{x_i - x_0}{\hat{\sigma_0}}\right) = 0, \tag{11.6}$$

where ψ is any nondecreasing odd function, x_0 is a location estimator and $\hat{\sigma_0}$ is a preliminary scale estimator, i. e., MADAM. M-estimators are affine equivariant, and the equation (11.6) can be solved using the Newton–Raphson algorithm with the sample median as a starting value. Similarly, M-estimator of scale is defined as the solution of the equation

$$\frac{1}{N}\sum_{i=1}^{N} \rho\left(\frac{x_i - \hat{\mu_0}}{\sigma}\right) = \kappa, \quad \text{with } 0 < \kappa < \rho(\infty), \tag{11.7}$$

where ρ is an even, differentiable and nondecreasing (on the positive numbers) loss function, σ is a scale estimator and $\hat{\mu_0}$ is a preliminary location estimator, i. e., the

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

sample median. M-Estimators can use the Huber or logistic ψ function. The Huber ψ_H is defined as

$$\psi_H(x, b) = \max(-b, \min(x, b)), \tag{11.8}$$

where b is a user-defined constant. For $b \to 0$ the location estimator becomes the median, whereas for $b \to \infty$ it becomes the sample mean. For normally distributed data, b equal to 1.345 gives 95 % efficiency. The efficiency of the M-Estimator increases together with b, while the robustness is inversely proportional to b. The logistic smooth $\psi_{\log}$ function is defined as:

$$\psi_{\log}(x) = \frac{e^x - 1}{e^x + 1}, \tag{11.9}$$

which may be rewritten as $\psi_{\log}(x) = 2F(x) - 1$, where $F(x) = 1/(1+e^{-x})$ is the cumulative distribution function of the logistic PDF, also known as the sigmoid function. In this paper, M-Estimators using ψ_H function are called M-Huber estimators, whereas those using $\psi_{\log}$ are referred to as M-Logistic estimators.

11.2.3.2 Outlier detection in signal

A classic method for online outlier identification and replacement is a moving-window Hampel filter [17], whose response is given by

$$y_k = \begin{cases} x_k & |x_k - m_k| \le t \cdot s_k, \\ m_k & |x_k - m_k| > t \cdot s_k, \end{cases} \tag{11.10}$$

where m_k and s_k are the mean and standard deviation estimators from moving data window W. Two tunable parameters of the Hampel filter are w (window half-width) and t (threshold value).

In a typical setup, median is used for m_k and $s_k = 1.4826 \times \text{MADAM}$ as an unbiased estimator of the standard deviation for normally distributed data. However, due to the abnormal, fat-tailed character of outlier generating distributions and a potentially short window W, various other robust estimators presented in Section 11.2.3.1 might be used. The resulting variants of the Hampel identifier $H_1 \dots H_6$ are shown in Table 11.1.

Another possible refinement of the Hampel filter is to designate the task of outlier identification to a different domain. Original signal $x(t)$ can be transformed in a way that emphasizes the influence outliers with respect to the original signal. Two such transformations are: *absolute changes* (11.11) and *log returns* (11.12), the latter originat-

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 11.1: Hampel identifiers for various different estimators.

Hampel identifier	Location estimator m	Scale estimator s
H_1	median	$1.4826 \times$ MADAM
H_2	M-Huber	$1.4826 \times$ MADAM
H_3	M-Logistic	$1.4826 \times$ MADAM
H_4	median	M-Logistic
H_5	M-Huber	M-Logistic
H_6	M-Logistic	M-Logistic

ing from the field of economics.

$$z(k) = x(k+1) - x(k) \quad \forall_{k\in\langle 1,N-1\rangle}, \tag{11.11}$$

$$z_W(k) = \log\big(x(k+1)\big) - \log\big(x(k)\big) \quad \forall_{W\in\langle 1,N-1\rangle, k\in W}. \tag{11.12}$$

The subscript in z_W is meant to emphasize the fact that *log returns* transformation is only valid for a local window W. If, in a given window, any $x(k) < 1$, the following step is needed (11.13):

$$x'(k) = x(k) + 1 - \min_W(x) \quad \forall_{k\in W}, \tag{11.13}$$

where $\min_W(x)$ is the minimum value of x over window W. In this paper the three outlier identification domains are referred to as A (original), B (absolute changes) and C (log returns). For domains B, C, the regular A filter's response (11.10) needs to be modified as shown in (11.14) and (11.15).

$$y'_k = \begin{cases} 0 & |z_k - m_{k,z}| \le t \cdot s_{k,z}, \\ 1 & |z_k - m_{k,z}| > t \cdot s_{k,z}, \end{cases} \tag{11.14}$$

$$y_k = \begin{cases} x_k & y'_{k-1} \wedge y'_k = 0, \\ m_k & y'_{k-1} \wedge y'_k = 1. \end{cases} \tag{11.15}$$

11.2.3.3 Signal denoising

A centered moving average (MA) filter is an optimal solution for the problem of reducing random white noise, while keeping the sharpest step response. It provides the lowest possible noise for a given edge sharpness, which is of importance when analyzing time domain modulated signals. While working with oscillatory signals, it's convenient to relate window size to the oscillating period T and refer to it as relative window size w_r (11.16)

$$w_r = \frac{w}{T}. \tag{11.16}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

11.2.4 Signal decomposition—EMD

Empirical mode decomposition (EMD) adaptively decomposes a signal $x(t)$ into a finite number of oscillatory components called intrinsic mode functions (IMF) [9]. The IMFs satisfy two conditions: (i) The number of extrema and the number of zero crossings must be either equal or differ at most by one; (ii) at any time, the mean value of the envelopes defined by the local extrema equals zero. The IMFs are generated in a sifting process that results in the decomposition (11.17)

$$x(k) = \sum_{i=1}^{M} c_i(k) + r(k), \tag{11.17}$$

where $c_i(k)$ is the ith IMF and $r(k)$ is the residual (nonstationary component). The EMD algorithm can be outlined with the following steps:

Step 1: Find the location of local extrema of the signal $x(k)$;

Step 2: Interpolate all the maxima (minima) to obtain the upper (lower) envelope, $e_{\max}(k)$ ($e_{\min}(k)$);

Step 3: Calculate the local mean, $m(k) = [e_{\min}(k) + e_{\max}(k)]/2$;

Step 4: Subtract the local mean from the signal: $h(k) = x(k) - m(k)$;

Step 5: If $h(k)$ satisfies the stopping criteria, set $c_i(k) = h(k)$, the ith IMF component, and proceed to step 6. Otherwise, set $x(k) = h(k)$ and repeat the procedure starting from step 1;

Step 6: Calculate the residual $r(k) = x(k) - c_i(k)$;

Step 7: If $r(k)$ doesn't satisfy the termination criteria, set $x(k) = r(k)$ and repeat the procedure starting from step 1; otherwise, finish the decomposition.

Sifting in *Step 5* stops when the following limits are reached: the maximum number of sifting iterations and the relative tolerance of consecutive $h(k)$. Possible EMD termination conditions in *Step 7* are evaluated based on the number of IMFs, the number of extrema in residual, the energy ratio of signal and the residual.

11.2.5 Oscillation characterization procedure

The following oscillation detection and characterization procedure is suitable for the detection of multiple, regular oscillations in closed-loop systems:

1. Perform empirical mode decomposition of the process output (PV) signal $x(k)$ to obtain the IMFs—$c_i(k)$;
2. Calculate the normalized correlation coefficient λ_i between the individual IMFs and the original signal (11.18). Check if $\lambda_i > 0.5$. Retain only those IMFs that obey

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

this criterion. In the following sections these IMFs are referred to as *significant*.

$$\lambda_i = \frac{\mu_i}{\max(\mu_i)} \quad \forall i = 1 \dots M, \tag{11.18}$$

$$\mu_i = \mathrm{Corr}(x(t), c_i(t)), \tag{11.19}$$

where $\mathrm{Corr}(a, b)$ is Pearson's correlation coefficient and M is the number of IMFs resulting from EMD;

3. For every significant IMF, compute the amplitude spectrum using a Fourier transform. Calculate the sparseness index Spar_i of the computed amplitude spectra (11.20). Check if $\mathrm{Spar}_i > 0.5$. Retain only those IMFs that obey this criterion.

$$\mathrm{Spar}_i = \frac{\sqrt{N} - (\sum_{j=1}^{N} |f_j|) / \sqrt{\sum_{j=1}^{N} |f_j|^2}}{\sqrt{N} - 1}, \tag{11.20}$$

where f_j is the component of the amplitude spectrum f and N is the length of the vector f;

4. Compute the autocorrelation function (ACF) of the IMFs obtained in step 3. Calculate the 'r' statistic (Thornhill index) for each ACF. Check if $r_i > 1$. The IMFs that obey this criterion are considered to be *oscillatory*.

$$r_i = \frac{1}{3} \frac{\mu_T}{\sigma_T}, \tag{11.21}$$

where μ_T and σ_T are the mean and standard deviation of the time between successive zero crossings of the ACF;

5. Calculate the amplitude of oscillations for oscillatory IMFs.

11.3 Proposed oscillation detection approach

11.3.1 Improved oscillation detection procedure

During our initial studies on the subject of oscillation characterization in signals of industrial origin, we have noticed that results obtained using the oscillation detection algorithm described in Section 11.2.5 were unreliable for data characterized by a low signal-to-noise ratio (SNR). The outcomes are even worse if the signal is contaminated with outliers. Preprocessing of data is the obvious solution. Although it would be tempting to use a linear filter such as the moving average (MA), this approach might be insufficient. A linear filter alters every point in a data set, thus creating the risk of data loss or mode damping. The proposed approach is to filter selectively out of outliers using a nonlinear filter and maintain as much of the original raw data as un-

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

altered as possible. We therefore propose the following 0th step to precede the EMD oscillation detection algorithm:

Step 0: Signal preprocessing—identify and replace the outliers using a nonlinear Hampel filter.

11.3.2 Improvements benchmarking

To compare the performance of different Hampel identifiers on a simulated data set, it is necessary to: (i) A priori define the outliers and (ii) compare the outlier detection algorithm's results with (i). Since a test data set $x(t)$ is composed of signal and noise

$$x(k) = \text{signal}(k) + \text{noise}(k), \tag{11.22}$$

an outlier can be *a priori* defined using the IQR method for the noise realization. (i). The Hampel identifier's performance benchmarks: TPR and FPR, based on the error matrix are used for (ii).

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{TN}}, \quad \text{FPR} = \frac{\text{FP}}{\text{FP} + \text{FN}}, \tag{11.23}$$

where TP (FP) is the number of correctly (incorrectly) detected outliers and TN (FN) is the number of correctly (incorrectly) detected nonoutliers. The effectiveness of improvements proposed in Section 11.3.1 will be assessed by comparing the oscillation characterization procedure's results with and without initial data preprocessing. The following benchmarks, corresponding to 4th and 5th steps of oscillation detection procedure, are compared:

- *Correct detection*—the share of experiments with detected oscillations (4th step in Section 11.2.5) and correctly estimated oscillation period (11.24) is

$$\mu_T - \sigma_T < T < \mu_T + \sigma_T, \tag{11.24}$$

  where T is the actual period, μ_T is the estimated period and σ_T is the standard deviation (error) of the estimated period.
- *Incorrect detection*—the share of experiments with detected oscillations and incorrectly estimated oscillation period.
- *Median* σ_T—median uncertainty of correctly detected period. If a given oscillation period is detected in more than one IMF, minimum σ_T is selected.
- *MAE of amplitude*—median absolute error (11.25) of estimated amplitude, for correctly detected oscillation,

$$\text{MAE}(k) = \underset{k}{\text{median}} |A_k - A|, \tag{11.25}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

where A is the actual amplitude and A_k is the amplitude estimated in the kth experiment.

Correct and incorrect detection are referred to as *detection* benchmarks, while the median σ_T and MAE of amplitude are referred to as *estimation* benchmarks. It's worth noting that, since multiple oscillating components (correct and incorrect) can be detected in a signal containing just a single oscillation, the sum of correct and incorrect detection can exceed the value of 100 %.

11.4 Simulation study

In Sections 11.4.1 and 11.4.3, the aforementioned signal $x(k)$, consisting of an oscillating component and noise (containing outliers), is used for simulation purposes.

$$x(k) = A \sin\left(\frac{2\pi}{T} kh\right) + \text{noise}(k), \tag{11.26}$$

where T and A are the oscillation period and amplitude, k is the sample number and h is the sampling period. Without loss of generality, $h = 1$ and $A = 1$ in all examples.

The level of noise (SNR) of SαS noise (generating endogenous outliers) is calculated with the assumption that its power equals the squared M-Logistic estimate of the standard deviation. The SNR of white noise with injected erroneous observations (generating exogenous outliers) equals the variance of the underlying normal distribution. To account for the slow convergence (experiment's repeatability) of SαS distribution, either the number of samples or the number of experiments needs to be appropriately high.

Experiments in Section 11.4.1 use only SαS noise (stability index $\alpha = 1.7$) because this noise produces outliers in a more general way; the underlying endogenous outlier generation mechanism has a single adjustable parameter: stability α, whereas exogenous outliers have two: intensity p and scale D. To ensure that the experiments are exhaustive, $x(k)$ is studied over a wide range of periods and noise levels. So as not to lose the local character of the signal, the maximum window half-width $w = 10$. Thresholds are evaluated up to $t = 6$. The experiments' setup is listed in Table 11.2. The number of used samples is $n = 1e5$.

Section 11.4.3 uses both SαS and white noise with injected erroneous observations to generate outliers. Without loss of generality, scale parameter D of exogenous outliers is held constant and equals ten times the signal's amplitude. $x(k)$ is studied over a wide range of noise levels, outlier intensities, and for various preprocessing methods. The experiments' setup is listed in Table 11.3. The results of each simulation are averaged over $M = 1e3$ noise realizations. In every experiment, the number of samples equals ten periods of oscillations, i. e., for $T = 100$, the number of samples and $n = 1e3$.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 11.2: Outlier detection experiments' setup.

	T	w	t	Identifier	Domain	SNR
Example 1	100	1:10	0.1:0.1:6	H_1	A, B, C	10
Example 2	100	1:10	0.1:0.1:6	$H_1 \cdots H_6$	A, B, C	10
Example 3	10–1000	1:10	0.1:0.1:6	H_6	A	10
				H_6	B	
				H_4	C	
Example 4	100	1:10	0.1:0.1:6	H_6	A	0:10:20
				H_6	B	
				H_4	C	
Example 5	100	8	3.4	H_6	A	10
		10	2.7	H_6	B	
		10	2.8	H_4	C	

Note: All the experiments are carried out for endogenous outliers—$\alpha = 1.7$.

Table 11.3: Setup of the oscillation characterization experiments.

	T	Benchmark	Preprocessing		
			Without	w_r of MA	TPR of Hampel
Example 1	100	Detection	Yes	30	30
Example 2	100	Detection	No	30, 40, 50	30, 40, 50
Example 3	100	Estimation	No	30	30, 50
Example 4	50 200	Detection	No	30	50

Note: All the experiments are carried out for endogenous outliers—$\alpha = 1.5, 1.7, 1.9$, and exogenous outliers with contamination $p = 0.5, 1, 1.5$ and scale parameter $D = 10$. In every experiment SNR = −5:5:25.

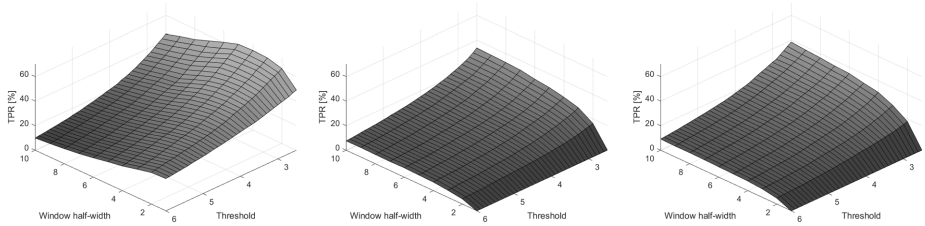
11.4.1 Hampel identifier performance

11.4.1.1 Example 1

For a better initial understanding of the relationship between Hampel identifiers' adjustable parameters, the window half-width and threshold, TPR and FPR surface plots are shown in Figs. 11.1 and 11.2. To focus attention, the plots' threshold axes are limited between 2.5 and 6.

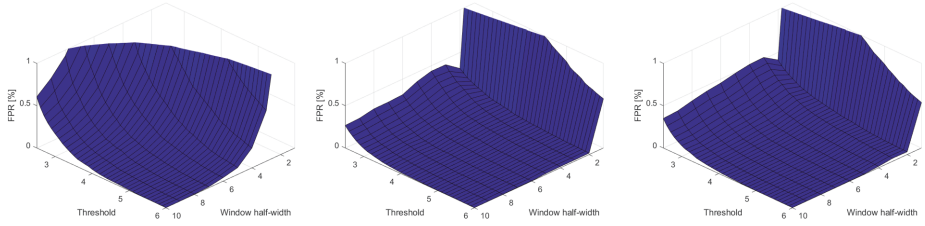
It's clearly visible that in all domains, both TPR and FPR decrease with the increase of the threshold. This observation indicates that there is a trade-off between TPR and FPR, with respect to the threshold value. The window half-width's influence is more subtle and cannot be easily generalized. It's also worth noting that the surface plots in transformed domains B and C are similar, and they differ significantly from

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use



(a) H_1 in A (b) H_1 in B (c) H_1 in C

Figure 11.1: TPR surface plots of H_1 identifier's performance in domains *A*, *B*, *C*. Threshold axes limits are set to 2.5 and 6.



(a) H_1 in A (b) H_1 in B (c) H_1 in C

Figure 11.2: FPR surface plots of H_1 identifier's performance in domains *A*, *B*, *C*. Threshold axes limits are set to 2.5 and 6, while FPR axes limits to 0 and 1. The threshold axis swapped places with the window half-width axis with respect to Fig. 11.1.

the plots in domain A. This observation reveals that the Hampel identifier's optimal parameters are domain-dependent.

11.4.1.2 Example 2

In this example the effect of appropriate estimator selection on the Hampel identifiers' performance is studied. For $H_1 \dots H_6$, FPR is plotted as a function of TPR. Plots comprised of only the optimal (lowest FPR for a given TPR) Hampel identifier's parameters (window half-width and threshold) are shown in Figs. 11.3, 11.4 and 11.5. Since poorly performing identifiers are of no interest, the FPR axes are limited between 0 and 1.

All the identifiers shown in Figs. 11.3a, 11.4a and 11.5a use scaled MADAM as the estimator of standard deviation. The plotted curves overlap one another in every figure, which means that the choice of a right mean estimator has a small influence on outlier detection performance.

Identifiers using the M-Logistic estimator of standard deviation are shown in Figs. 11.3b, 11.4b and 11.5b. A "classic" H_1 identifier is added to the figures as a reference. For every identification domain, identifiers $H_4 \dots H_6$ perform better than

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

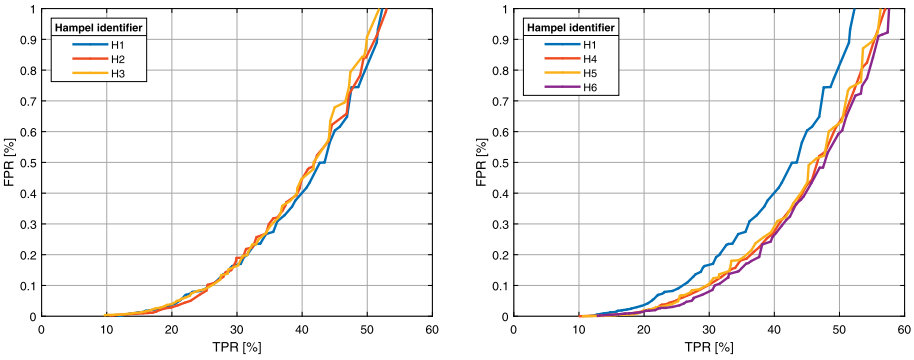


Figure 11.3: FPR vs. TPR plots of $H_1 \ldots H_6$ identifiers' performance in domain *A*. FPR axes limits are set to 0 and 1.

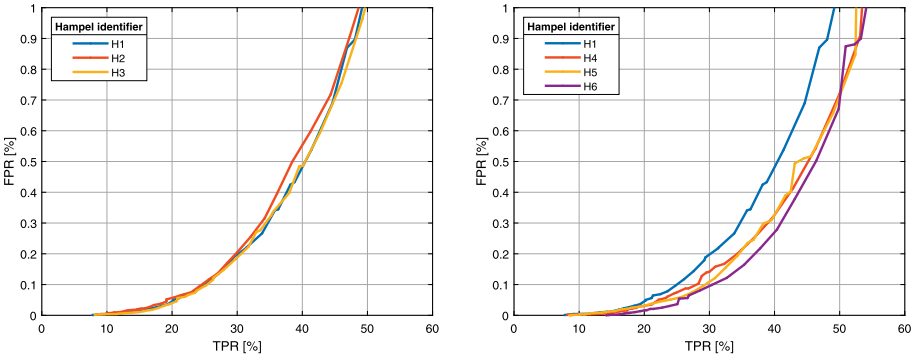


Figure 11.4: FPR vs. TPR plots of $H_1 \ldots H_6$ identifiers' performance in domain *B*. FPR axes limits are set to 0 and 1.

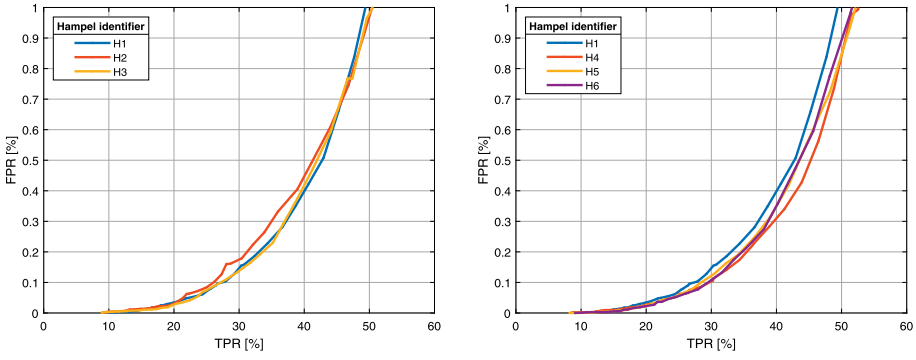


Figure 11.5: FPR vs. TPR plots of $H_1 \ldots H_6$ identifiers' performance in domain *C*. FPR axes limits are set to 0 and 1.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

$H_1 \dots H_3$, meaning that the M-Logistic estimator of standard deviation is superior to scaled MADAM.

As previously, the $H_4 \dots H_6$ curves also mostly overlap, and thus closer study is necessary to select the best pair of estimators. Table 11.4 shows their performance (FPR decrease relative to H_1) for an exemplary TPR = 30. Since this ranking holds for nearly the whole TPR range, the best estimators for each domain are selected. In the following sections H_6 is used in A, H_6 in B and H_4 in C.

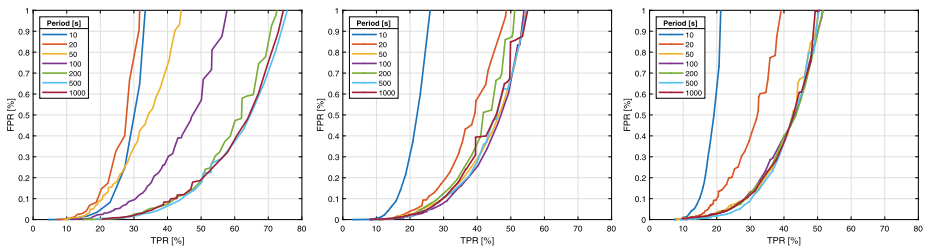
Table 11.4: FPR decrease relative to H_1 for TPR = 30.

Domain	A			B			C		
Identifier	H_4	H_5	H_6	H_4	H_5	H_6	H_4	H_5	H_6
Decrease [%]	39	37	52	28	45	52	30	17	27

Note: For each identification domain, the highest relative decrease is highlighted in gray.

11.4.1.3 Example 3

The effectiveness of outlier detection for various oscillation periods is analyzed in this example. The best FPR vs. TPR plots (as in Section 11.4.1.2) for various oscillation periods are shown in Fig. 11.6. It's worth noting the similarity between outlier detection performance in B and C. In these domains, the effectiveness of outlier detection is mostly independent of the oscillation period. In domain A, however, the identifiers behave differently. In this domain, their effectiveness increases with the increase of the oscillation period. It should also be emphasized that identifiers in domain A are especially suited for slow oscillations, due to their ability to reach TPR > 60 (while maintaining low FPR) for periods greater than 200.



(a) H_6 in A (b) H_6 in B (c) H_4 in C

Figure 11.6: FPR vs. TPR plots of Hampel identifiers' performance for various oscillation periods. FPR axes limits are set to 0 and 1.

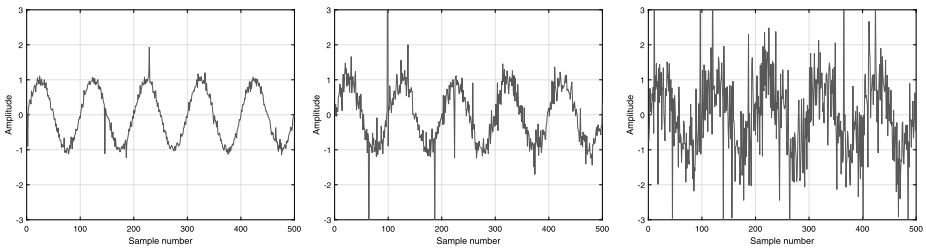
EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

11.4.1.4 Example 4

In this example the effectiveness of outlier detection for various noise levels is studied, and the simulated signals are shown in Fig. 11.7. For all the SNRs, the oscillating character of the signal is not hidden in noise, and the outliers are clearly visible.

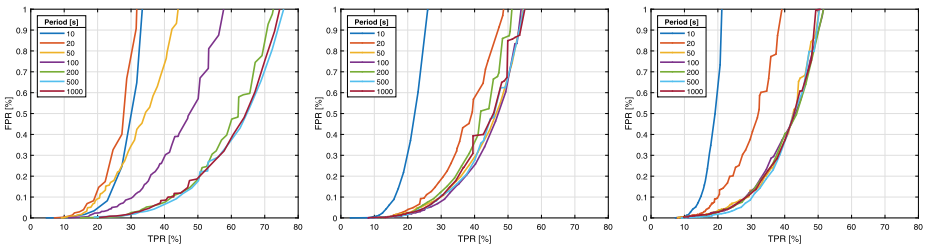
The best FPR vs. TPR curves (as in Section 11.4.1) for various signal-to-noise ratios are plotted in Fig. 11.8. Figure 11.8a shows that, for low noise levels, Hampel identifier computed in A performs worse than in B and C. For an exemplary TPR = 30, the FPR in A is more than five times larger than in the transformed domains. For medium noise levels, outlier identification performance in A is slightly better than in B and C, as shown in Fig. 11.8b. Figure 11.8c shows that, for high noise levels, filtering in A yields much better results than in B or C.

To conclude, outlier detection in domain A fails at low noise levels, while identification in domains B and C proves to be satisfactory over a wide range of SNRs. The performance of outlier detection in transformed domains is much more consistent, in comparison to the original unchanged domain.



(a) SNR = 20 **(b)** SNR = 10 **(c)** SNR = 0

Figure 11.7: Simulated oscillating signals for various levels of noise.



(a) SNR = 20 **(b)** SNR = 10 **(c)** SNR = 0

Figure 11.8: FPR vs. TPR plots of Hampel identifiers' performance in domains A, B, C for various noise levels. FPR axes limits are set to 0 and 1.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

11.4.1.5 Example 5

Although it's convenient to describe outlier detection performance in terms of benchmarks: TPR and FPR, it's also interesting to see which outliers are exactly detected. The studied signal with marked outliers is shown in Fig. 11.9. The Hampel identifier's parameters, window half-width and threshold values, are optimized for TPR = 30 (FPR < 0.2).

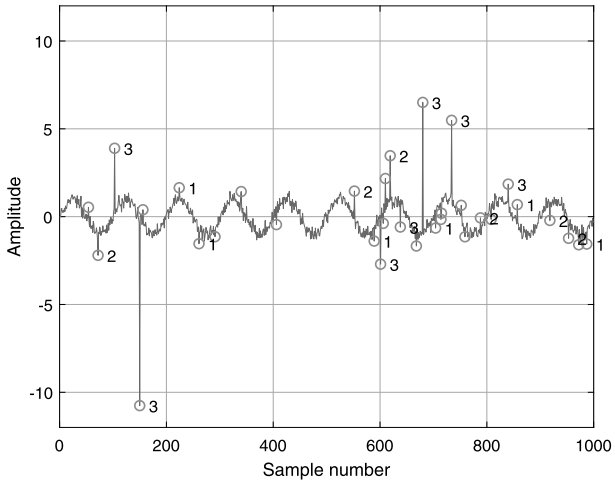


Figure 11.9: Simulated oscillating signal with marked outliers. Red circle–outlier defined using IQR method; 1, 2 and 3—number of domains in which the Hampel filter identified a given point as an outlier.

First of all, it's worth noting that the most extreme outliers are detected in all three domains. It can also be observed that the number of domains in which a particular outlier is detected is proportional to the outlier's magnitude. It was verified that the shares of outliers detected in single, two and three domains are nearly equal.

11.4.1.6 Selecting filtering domain

Sections 11.4.1.4 and 11.4.1.5 showed that transformed domains outperform the original domain in terms of outlier detection performance. More specifically, when compared to *A*, filtering in *B* and *C* is more consistent. It's suitable for a wider range of oscillation parameters. Of the two transformed domains, filtering in *B* is more effective than in *C* due to its ability to reach lower FPR for a given TPR, as shown in Section 11.4.1.5.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the subsequent sections, the best Hampel identifier using M-Logistic estimators of mean and standard deviation (H_6), computed in the incremental domain, (B) is used. Its preprocessing performance is evaluated for TPR = $[10, 30, 50]$.

11.4.2 Moving average filter

11.4.2.1 Selecting window size

Preprocessing a signal using a MA filter affects the oscillation characterization procedure's performance in a twofold way. First of all (i), filtering reduces the noise in the signal, which results in a more accurate EMD decomposition and thus increases the chance of oscillation detection. Second (ii), it distorts the spectrum of the true oscillating component in the signal, thus deteriorating the accuracy of oscillation parameters' estimation. There is a clear trade-off between the requirement of (i), where long window improves noise suppression, and (ii), where short window results in low distortion. Since any non-zero window size improves (i), the limit on window size is set by (ii).

To establish the maximum limit on the relative window size, it's necessary to recall how the amplitude of the oscillating component is calculated—the 5th step of the procedure in Section 11.2.5. To calculate the amplitude, it is necessary to compute the running sum of the amplitude spectrum in the estimated period's region of uncertainty (11.27). An inequality (11.28) stems from the assumption that the signal is oscillating ($r > 1$ in the 4th step of the procedure in Section 11.2.5).

$$\left(\frac{1}{\mu_T + \sigma_T}, \frac{1}{\mu_T - \sigma_T}\right), \tag{11.27}$$

$$\sigma_T < \frac{1}{3}\mu_T. \tag{11.28}$$

Figure 11.10 shows the amplitude spectrum plot of a MA filter, together with the mentioned region of uncertainty. A horizontal line intersecting the 1st peak is added. So that the part of the amplitude spectrum lying in the uncertainty region is above the horizontal line, the intersection of the horizontal line and the amplitude spectrum needs to be on the right side of the region of uncertainty. We therefore propose to select the maximum w_r in such a way that the mentioned intersection lies exactly on the right limit of the region. This "equilibrium" is reached for $w_r \approx 55$. In accordance with this result, in the following sections, MA filters with relative window sizes $w_r = [10, 30, 50]$ are evaluated.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

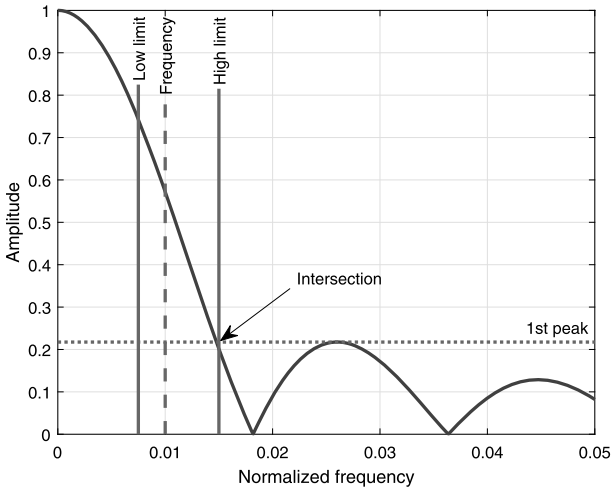
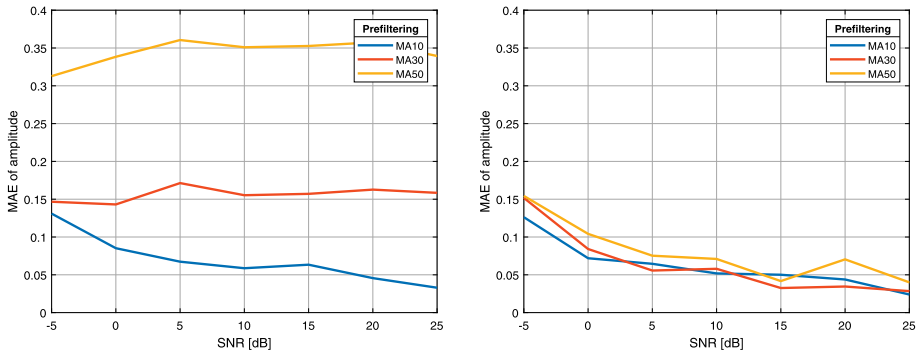


Figure 11.10: MA filter's amplitude spectrum plot for $w_r = 55$. A horizontal line intersects the 1st peak of the amplitude spectrum. Vertical lines represent the oscillation frequency's uncertainty range. Normalized frequency axis limits are set to 0 and 0.05.

11.4.2.2 Amplitude correction factor

As mentioned in Section 11.4.2.1, an MA filter reduces the estimated amplitude of oscillation. Figure 11.11a shows the corresponding MAE for a simulated signal. The error increases significantly with the increase of the relative window size. To counter this negative effect, we propose to apply the inverse of the MA filter in the frequency domain, before estimating the oscillation amplitude. The results of the proposed correction are shown in Fig. 11.11b. The MAE becomes significantly lower and this is nearly irrespective of the relative window size.



(a) Without correction **(b)** With correction

Figure 11.11: MAE of estimated amplitude for various relative window sizes and various noise levels, for endogenous outliers with stability index $\alpha = 1.7$.

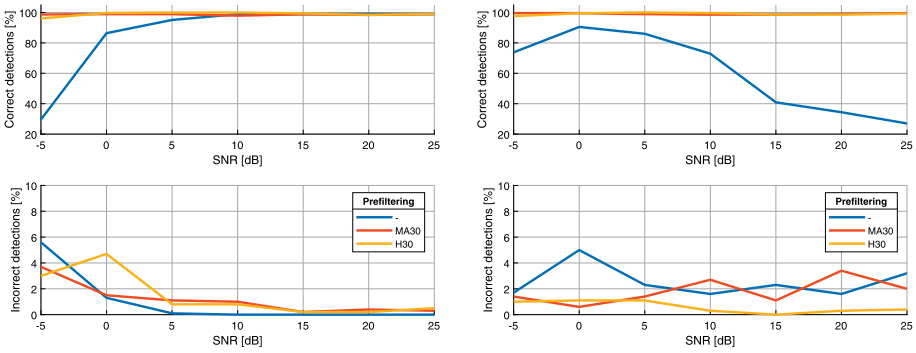
EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

11.4.3 Oscillation characterization performance

11.4.3.1 Importance of filtering

This example studies the influence of initial signal preprocessing on oscillation detection performance. The results for endogenous and exogenous outliers are shown in Figs. 11.12a and 11.12b, respectively.

Signal preprocessing has a significant influence on the share of correct detection for both outlier generation mechanisms. The share of correct detection for unfiltered



(a) Endogenous, stability $\alpha = 1.7$ **(b)** Exogenous, contamination $p = 1\%$

Figure 11.12: Oscillation detection performance with and without preprocessing, for various outlier generation mechanisms. Correct detection axes limits are set to 20 and 105, while incorrect detection axes limits are 0 and 10.

Table 11.5: Correct detection performance with and without preprocessing.

Outlier intensity	Preprocessing	SNR						
		−5	0	5	10	15	20	25
	Hampel	97.9	99.7	99.9	99.6	98.9	99.1	99.3
$\alpha = 1.9$	MA	98.8	98.9	98.8	98.4	99.5	98.6	99.7
	–	81.9	98.3	100.0	99.3	99.3	99.1	99.7
	Hampel	91.4	98.8	99.8	99.7	99.5	98.8	98.9
$\alpha = 1.5$	MA	91.7	98.2	98.0	98.6	98.8	98.8	99.0
	–	3.2	51.1	81.6	95.6	96.1	98.1	99.0
	Hampel	98.8	99.2	99.7	99.7	99.2	98.8	98.9
$p = 0.5$	MA	99.9	99.2	99.1	98.2	98.9	98.5	99.4
	–	90.5	97.4	95.9	91.7	72.1	75.8	70.6
	Hampel	98.1	99.8	99.7	99.6	98.5	98.8	99.4
$p = 1.5$	MA	99.4	99.2	99.0	98.9	98.5	97.9	98.9
	–	54.2	84.2	75.5	59.6	28.8	20.9	20.7

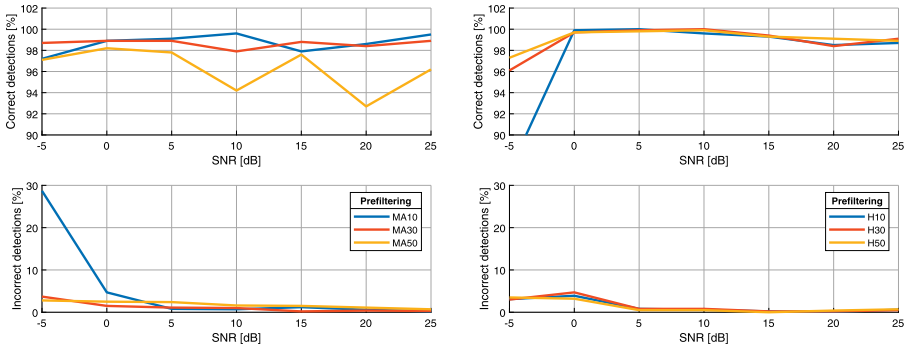
Note: MA filter with $w_r = 30$ and Hampel filter with TPR = 30 are used. For each outlier intensity and SNR, the lowest value (if under 95 %) is highlighted in gray.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

data is remarkably lower than 95 %, for a wide range of noise levels. For the prefiltered data, though, the correct detection curves are nearly flat. Thanks to prefiltering, the correct detection ratio is greater than 95 % in the whole SNR range. Moreover, filtering not only increases the number of correct detection, but it can also reduce the share of incorrect detection. Also verified was the necessity of preprocessing holds for various outlier intensities (various α and p), as shown in Table 11.5.

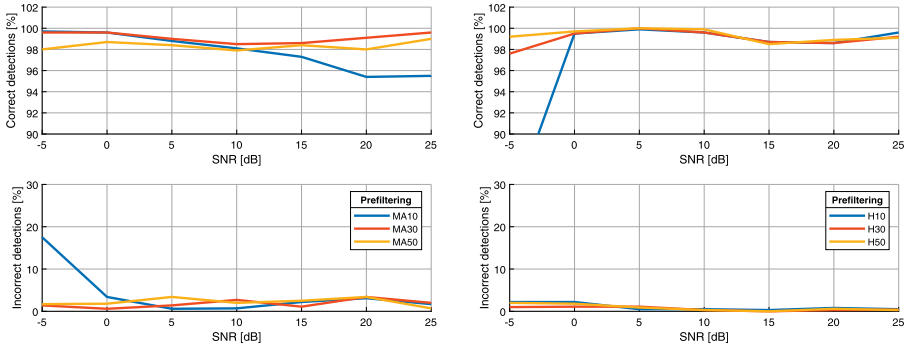
11.4.3.2 Oscillation detection performance

This example studies the influence of preprocessing methods' parameters, i. e., relative window size of MA filter and TPR of Hampel filter, on oscillation detection per-



(a) MA filter (b) Hampel filter

Figure 11.13: Oscillation detection performance for different preprocessing filters' parameters and endogenous outliers with stability index $\alpha = 1.7$. Correct detection axes limits are set to 90 and 102, while incorrect detection axes limits to 0 and 30.



(a) MA filter (b) Hampel filter

Figure 11.14: Oscillation detection performance for various preprocessing methods' parameters and exogenous outliers with contamination level $p = 1\,\%$. Correct detection axes limits are set to 90 and 102, while incorrect detection axes limits are 0 and 30.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

formance. The results for endogenous and exogenous outliers are shown in Figs. 11.13 and 11.14, respectively.

Figure 11.13a shows that short windows result in an increased number of incorrect detection, whereas long windows result in a small decrease in the share of correct ones. Figure 11.13b shows that the correct detection ratio is proportional to the TPR, while the share of incorrect detection is nearly insensitive to this parameter. Similar observations can be made for exogenous outliers, as based in Fig. 11.14.

The observations made in the previous paragraph are also valid for a wider range of outlier intensities, as shown in Table 11.6. It should be highlighted that, for $w_r = 30$, none of the table's cells is colored gray. This observation reveals that $w_r = 30$ is the optimal parameter for a MA filter since it never significantly decreases the correct

Table 11.6: Correct detection performance for various preprocessing methods' parameters.

Outlier intensity	Preprocessing	w_r or TPR	SNR						
			−5	0	5	10	15	20	25
$\alpha = 1.9$	MA	10	99.0	99.5	99.4	98.9	98.1	98.7	98.9
		30	98.8	98.9	98.8	98.4	99.5	98.6	99.7
		50	97.6	98.3	95.3	93.1	98.0	90.8	97.3
	Hampel	10	95.7	99.7	99.7	99.7	99.3	97.9	98.9
		30	97.9	99.7	99.9	99.6	98.9	99.1	99.3
		50	99.1	99.5	99.9	99.6	98.9	98.8	99.0
$\alpha = 1.5$	MA	10	82.6	95.5	96.8	99.1	99.0	98.3	99.3
		30	91.7	98.2	98.0	98.6	98.8	98.8	99.0
		50	92.3	96.7	96.7	95.9	95.7	95.4	92.9
	Hampel	10	57.5	98.6	99.7	99.7	99.1	99.4	99.0
		30	91.4	98.8	99.8	99.7	99.5	98.8	98.9
		50	94.6	99.7	99.9	99.8	99.4	99.4	98.9
$p = 0.5$	MA	10	99.8	99.5	98.7	98.5	97.8	98.6	97.8
		30	99.9	99.2	99.1	98.2	98.9	98.5	99.4
		50	98.2	97.4	96.6	95.9	97.6	96.3	99.3
	Hampel	10	93.3	100.0	99.9	99.6	99.1	98.8	99.1
		30	98.8	99.2	99.7	99.7	99.2	98.8	98.9
		50	98.7	99.7	99.7	99.9	98.7	98.8	99.6
$p = 1.5$	MA	10	99.9	99.7	98.6	97.4	95.5	94.2	90.8
		30	99.4	99.2	99.0	98.9	98.5	97.9	98.9
		50	99.0	98.8	99.4	98.1	98.6	97.2	99.5
	Hampel	10	71.5	99.5	99.9	99.2	98.9	98.8	99.5
		30	98.1	99.8	99.7	99.6	98.5	98.8	99.4
		50	98.1	99.9	99.9	99.4	98.8	99.1	99.7

Note: For each outlier intensity, preprocessing method and SNR, the lowest value (if under 95 %) is highlighted in gray.

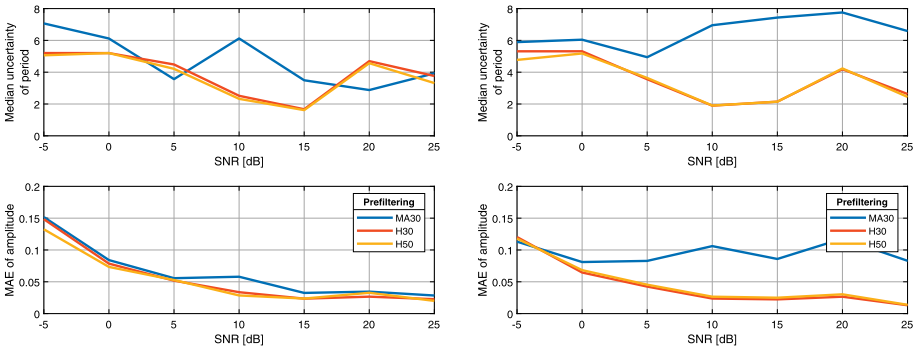
EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

detection performance. For the very same reason, Hampel identifiers with TPR = 30 and TPR = 50 are the best-performing ones. Filters with these parameters are studied if the following section.

11.4.3.3 Oscillation estimation performance

This example studies the performance of oscillation estimation for the best filters' parameters selected in the previous section. The results for endogenous and exogenous outliers are shown in Figs. 11.15a and 11.15b, respectively.

Although for endogenous outliers both types of filters perform similarly, a significant difference in their performance can be observed for exogenous outliers. When a



(a) Endogenous, stability $\alpha = 1.7$ (b) Exogenous, contamination $p = 1\%$

Figure 11.15: Oscillation estimation performance for a range of preprocessing methods' parameters and outlier generation mechanisms.

Table 11.7: MAE of estimated amplitude for various preprocessing methods.

Outlier intensity	Preprocessing	SNR						
		−5	0	5	10	15	20	25
$\alpha = 1.9$	Hampel	0.13	0.06	0.046	0.025	0.024	0.027	0.014
	MA	0.1	0.06	0.045	0.045	0.018	0.039	0.017
$\alpha = 1.5$	Hampel	0.15	0.083	0.053	0.037	0.024	0.031	0.027
	MA	0.23	0.13	0.077	0.072	0.052	0.044	0.045
$p = 0.5$	Hampel	0.11	0.063	0.043	0.025	0.024	0.024	0.012
	MA	0.1	0.073	0.07	0.078	0.064	0.091	0.054
$p = 1.5$	Hampel	0.13	0.067	0.045	0.025	0.029	0.036	0.015
	MA	0.12	0.097	0.096	0.11	0.1	0.12	0.1

Note: MA filter with $w_r = 30$ and Hampel filter with TPR = 50 are used. For each outlier intensity and SNR, the highest MAE (if twice the lowest) is highlighted in gray.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

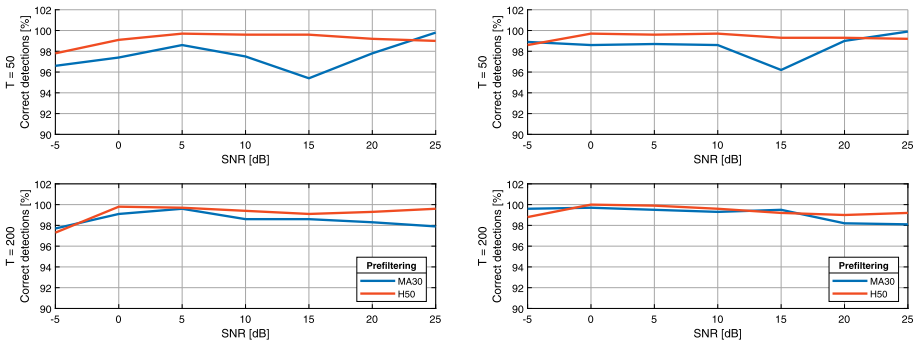
signal is contaminated with this type of outliers, the Hampel filter outperforms the MA filter significantly since, for the latter, both σ_T and MAE of the amplitude are remarkably higher. It's also worth noting that of the two Hampel identifiers, the one with a higher TPR achieves slightly better overall performance.

Table 11.7 shows the MAE of estimated amplitude for various outlier intensities. Only the better Hampel filter (TPR = 50) is included. It turns out that the observations from the previous paragraph hold for the whole studied range of outlier intensities. The cells for which the error of MA filter is more than twice the error of Hampel filter are highlighted in gray.

11.4.3.4 Oscillation detection under uncertainty

In previous sections, the preprocessing filters' parameters were optimized for the known oscillating component of the signal. In this section, however, the filters' robustness to uncertainty in oscillation period is studied. In this example, two signals whose periods' equal 0.5 and 2 of the filters' optimal period are used. The results for endogenous and exogenous outliers are shown in Figs. 11.16a and 11.16b, respectively.

Although, for slow oscillations, both filters perform similarly, differences can be observed for oscillations faster than the filter's optimal. For these frequencies, the Hampel filter outperforms the MA filter since it's correct detection curve lies mostly above the MA filter's curve. These observations are valid for a range of outlier intensities, as shown in Table 11.8.



(a) Endogenous, stability $\alpha = 1.7$ **(b)** Exogenous, contamination $p = 1\%$

Figure 11.16: Oscillation detection performance under uncertainty for various outlier generation mechanisms. Correct detection axes limits are set to 90 and 102.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 11.8: Correct detection performance under uncertainty.

Outlier intensity	Preprocessing	SNR						
		−5	0	5	10	15	20	25
$\alpha = 1.9$	Hampel	99.1	99.5	99.7	99.9	99.5	99.0	99.0
	MA	98.4	97.3	98.2	97.6	93.4	99.6	100.0
$\alpha = 1.5$	Hampel	96.8	99.3	99.6	99.7	99.7	98.7	99.0
	MA	90.9	95.4	96.1	97.2	97.1	96.8	99.8
$p = 0.5$	Hampel	99.2	99.2	99.9	99.5	99.3	98.5	98.8
	MA	99.4	97.4	98.2	98.1	95.6	99.1	100.0
$p = 1.5$	Hampel	99.2	99.5	99.7	99.9	99.4	98.8	98.9
	MA	98.8	98.5	98.4	97.7	96.9	99.0	99.68

Note: MA filter with $w_r = 30$ and Hampel filter with TPR = 50 are used. For each outlier intensity and SNR, the lowest value (if under 95 %) is highlighted in gray.

11.4.3.5 Selecting preprocessing method

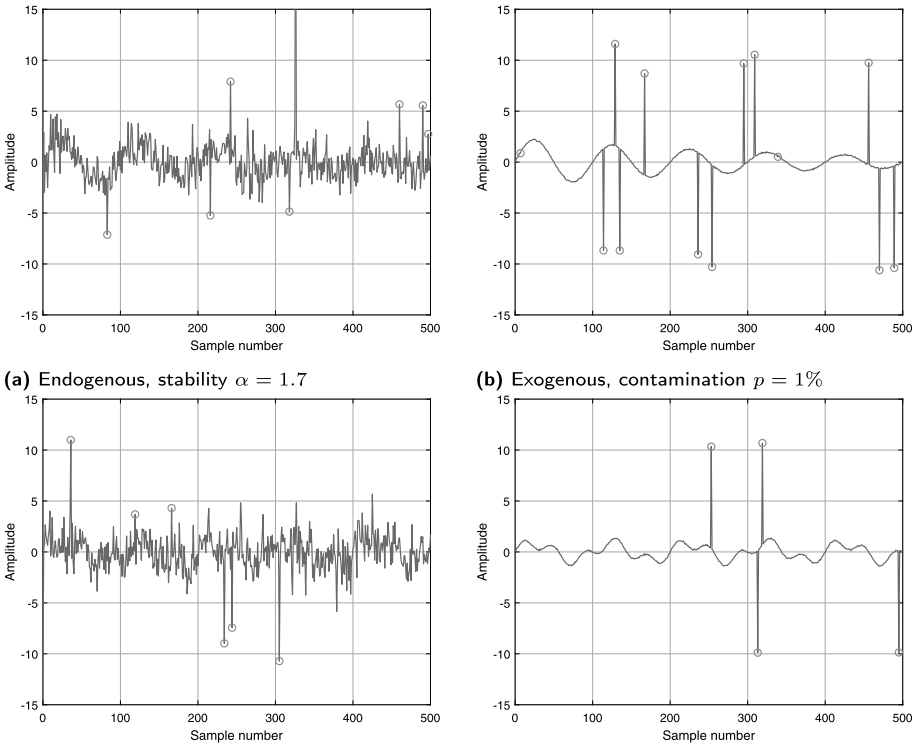
Section 11.4.3.2 showed that both types of filters (if properly tuned) can achieve a similar performance of oscillation detection. The simulation in Section 11.4.3.3 showed that the Hampel filter is better suited for estimating an oscillation's parameters when a signal is contaminated with exogenous outliers. Section 11.4.3.4 showed that the performance of the MA filter's oscillation detection deteriorates when the filter operates outside its optimal range. In accordance with these findings, the Hampel filter is selected as the best preprocessing method and is utilized in subsequent sections.

11.4.4 Literature benchmarks

While Section 11.4.3 was mostly concerned with generic examples, this section studies specific scenarios resembling real-life industrial signals. The need for general examples is obvious: With their help, quantitative conclusions can be made. On the other hand, the specific examples' main goal is to familiarize the reader with typical situations that could benefit from the proposed improvements. Examples in this section are inspired by benchmarks from [10]. Damped oscillation and multiple oscillations signals are studied. In these examples, endogenous outliers with SNR = −5 and exogenous outliers with SNR = 25 are added to the signals. Noise levels are selected based on Fig. 11.12 so that the need for preprocessing is emphasized. Simulated signals are shown in Fig. 11.17, whereas oscillation characterization results in Table 11.9.

For both classes of studied signals, the oscillation characterization procedure accurately detects and estimates oscillation parameters. Without prefiltering, though, either no oscillations are detected (no *significant* IMFs or $r < 1$), the period is estimated

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use



(a) Endogenous, stability $\alpha = 1.7$

(b) Exogenous, contamination $p = 1\%$

(c) Endogenous, stability $\alpha = 1.7$

(d) Exogenous, contamination $p = 1\%$

Figure 11.17: Benchmark signals with outliers generated by various mechanisms. Top: damped oscillations; bottom: multiple oscillations. Detected outliers marked with a red circle.

incorrectly or there is a high amplitude error. It should be stressed that, although the procedure's outcomes are dependent on the noise realization, signal prefiltering improves the oscillation characterization procedure's results in most of the experiments.

11.5 Industrial data validation

In this section, the improved oscillation characterization procedure's superiority over the original version of thereof is validated with real data sets. Analyzed signals are shown in Fig. 11.18. The data sets have different origins: the power industry (power plant megawatts), control laboratory experiments (mechanical ball and plate system) [30] and medicine (muscle electromyography—EMG) and suffer from both types of outliers. The best preprocessing method selected in Sect 11.4.3.5, namely, the Hampel filter with TPR = 50, is used for prefiltering. The oscillation characterization procedure's results, with and without preprocessing, are shown in Table 11.10.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 11.9: Oscillation characterization results for simulated benchmark signals.

Signal	Outlier intensity	Preprocessing	Significant IMF	μ_T	σ_T	'r' statistic	Amplitude	Correct detection	Amplitude error [%]
Damped oscillation	$\alpha = 1.7$	Yes	1	100	6.2	5.4	0.89	Yes	–
		No	1	15	3.6	1.4	3.9	No	–
	$p = 1$	Yes	1	140	65	0.71	–	–	–
			2	100	1.7	20	0.83	Yes	–
		No	–	–	–	–	–	–	–
Multiple oscillations	$\alpha = 1.7$	Yes	1	40	2.4	5.6	0.65	Yes	17
			2	101	8.9	3.8	0.92	Yes	11
		No	1	19	10	0.63	–	–	–
	$p = 1$	Yes	1	40	5.2	2.6	0.62	Yes	12
			2	40	3.1	1.8	0.6	Yes	8
			3	100	3	11	0.82	Yes	1
		No	1	51	34	0.5	–	–	–
			2	106	20	1.8	1.8	Yes	116

Note: Selected cells showing the root cause of oscillation characterization procedure's failure are highlighted in gray.

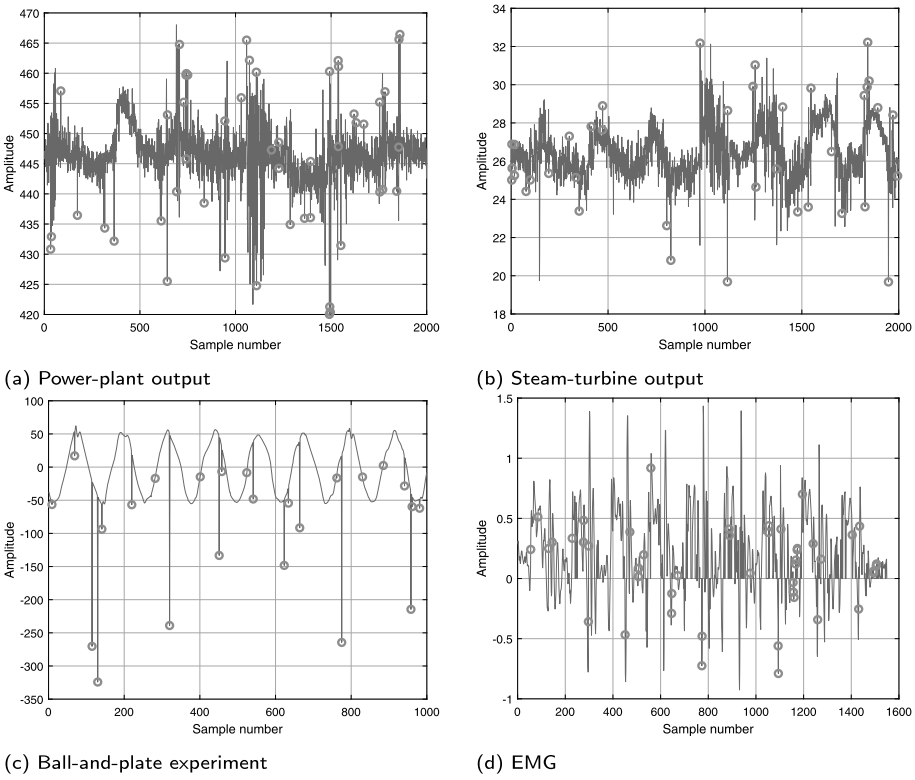


Figure 11.18: Real data procedure's validation with detected outliers marked by a red circle.

Figure 11.18a shows the total power output of a combined-cycle power plant. The data set is contaminated with endogenous outliers, whose parameters are estimated giving a stability index $\alpha = 1.3$ and SNR $= 2$. Without prefiltering, the original procedure fails to detect oscillation since no *significant* IMF is detected. Initial filtering enables a correct detection of period and a coarse (correct order of magnitude) estimation of amplitude.

Figure 11.18b shows the power output of a steam turbine, while Fig. 11.18c shows the data from a ball-and-plate system. The first data set is contaminated with endogenous outliers: $\alpha = 1.3$ and SNR $= 6$, whereas the second one with exogenous outliers: $p = 1.2$ and SNR $= 28$. Again, the original procedure fails without prefiltering. In this case, however, the root cause of failure is a high standard deviation (uncertainty) of the estimated period. Prefiltering significantly lowers the uncertainty and thus enables a correct detection and estimation (correct order of magnitude in the case of power output) of the oscillation.

Figure 11.18d shows the EMG of a left tibialis anterior muscle. In this example, the underlying outlier generation mechanism is unclear. The original procedure detects only a single oscillating component. It skips the faster oscillation, due to a high

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 11.10: Oscillation characterization results for real signals.

Signal	Preprocessing	Significant IMF	μ_T	σ_T	'r' statistic	Amplitude	Correct detection
Power-plant output	Yes	1	224	113	0.66	–	–
		2	270	41	2.2	1.4	Yes
	No	–	–	–	–	–	–
Steam-turbine output	Yes	1	50	30	0.56	–	–
		2	213	102	0.7	–	–
		3	268	66	1.35	1	Yes
	No	1	201	120	0.55	–	–
		2	311	176	0.59	–	–
Ball-and-plate experiment	Yes	1	121	5.9	6.83	57	Yes
	No	1	105	42	0.83	–	–
		2	139	97	0.47	–	–
EMG	Yes	1	18	2.6	2.3	0.5	Yes
		2	41	23	0.59	–	–
		3	64	26	0.82	–	–
		4	164	8.3	6.5	0.2	Yes
	No	1	26	10	0.86	–	–
		2	54	32	0.56	–	–
		3	159	33	1.6	0.23	Yes
		4	167	40	1.39	0.3	Yes

Note: Selected cells showing the root cause of oscillation characterization procedure's failure are highlighted in gray.

uncertainty of the period. Prefiltering significantly lowers the standard deviation of the faster-oscillating component, and thus results in a more detailed analysis of the oscillation's nature.

11.6 Conclusions and further research

This paper presents an improved oscillation characterization procedure. The proposed improvement uses initial signal preprocessing to increase procedure's robustness to outliers.

In the first part of the paper, various versions of Hampel identifier are assessed based on their ability to detect outliers in the presence of oscillations. Based on an exhaustive simulation study, the Hampel filter computed in the incremental domain, using M-Logistic estimators of mean and standard deviation, is selected as the best outlier identifier.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the second part of the paper, the performance of the original oscillation characterization procedure in the presence of outliers is evaluated. Various outlier generation mechanisms are analyzed. It is observed that, for endogenous outliers, simulated using SαS distribution with stability index $\alpha < 2$, the original procedure fails at low SNRs, whereas, for exogenous outliers, simulated using normal distribution contaminated with intensity p, the procedure's performance deteriorates at high SNRs. It is shown that initial signal prefiltering using a MA or Hampel filter significantly increases the procedure's performance. A closer comparison of these two preprocessing methods reveals that the Hampel filter is superior at (i) estimating oscillation amplitude and (ii) performing in nonoptimal conditions. The improved oscillation characterization procedure is validated using both simulated and real data sets.

In our further research, an automatic procedure for Hampel filter parameters' adjustment will be studied. It is also worth analyzing whether the application of a different signal decomposition method (i. e., local mean decomposition (LMD), ensemble empirical mode decomposition (EEMD) or oscillation detection statistic (based on Lempel–Ziv complexity) has any influence on the procedure's robustness to outliers.

Bibliography

[1] M. F. Aftab, M. Hovd, and S. Sivalingam. Improved oscillation detection via noise-assisted data analysis. *Control Engineering Practice*, 81:162–171, 2018.

[2] D. Bian, Z. Yu, D. Shi, R. Diao, and Z. Wang. A robust real-time low-frequency oscillation detection and analysis (lfoda) system with innovative ensemble filtering. *CSEE Journal of Power and Energy Systems*, 6(1):174–183, 2020.

[3] C. Croux and C. Dehon. Robust estimation of location and scale. Wiley StatsRef: Statistics Reference Online, 2014.

[4] P. D. Domański. Statistical measures for proportional–integral–derivative control quality: Simulations and industrial data. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 232(4):428–441, 2018.

[5] P. D. Domański. *Control Performance Assessment: Theoretical Analyses and Industrial Practice*. Springer Nature Switzerland AG, Cham, 2020.

[6] M. Gupta, J. Gao, C. Aggarwal, and J. Han. *Outlier Detection for Temporal Data*. Morgan & Claypool Publishers, 2014.

[7] D. M. Hawkins. *Identification of Outliers*. Chapman and Hall, London; New York, 1980.

[8] L. Huang, X. Ni, W. L. Ditto, M. Spano, P. R. Carney, and Y. C. Lai. Detecting and characterizing high-frequency oscillations in epilepsy: A case study of big data analysis. *Royal Society Open Science*, 4(1):160741, 2017. https://doi.org/10.1098/rsos.160741.

[9] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, 1998.

[10] M. Jelali and B. Huang. *Detection and Diagnosis of Stiction in Control Loops: State of the Art and Advanced Methods*. Springer-Verlag, London, 2010.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[11] L. B. Klebanov and I. Volchenkova. Outliers and the ostensibly heavy tails. arXiv:1807.08715v1, Department of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic, 2018.
[12] P. Y. Kovalenko. WAMS measurements pre-processing for detecting low-frequency oscillations in power systems. *Journal of Physics: Conference Series*, 870:012011, 2017.
[13] X. Li, J. Wang, B. Huang, and S. Lu. The DCT-based oscillation detection method for a single time series. *Journal of Process Control*, 20(5):609–617, 2010.
[14] T. Miao and D. E. Seborg. Automatic detection of excessively oscillatory feedback control loops. In *Proceedings of the 1999 IEEE International Conference on Control Applications*, volume 1, pages 359–364, 1999.
[15] E. Naghoosi and B. Huang. Automatic detection and frequency estimation of oscillatory variables in the presence of multiple oscillations. *Industrial & Engineering Chemistry Research*, 53(22):9427–9438, 2014.
[16] J. W. Osborne and A. Overbay. The power of outliers (and why researchers should always check for them). *Practical Assessment, Research, and Evaluation*, 9(6):1–8, 03 2004.
[17] R. K. Pearson, Y. Neuvo, J. Astola, and M. Gabbouj. Generalized Hampel filters. *EURASIP Journal on Advances in Signal Processing*, 2016:87, 2016.
[18] P. J. Rousseeuw and A. M. Leroy. *Robust Regression and Outlier Detection*. John Wiley & Sons, Inc., New York, NY, 1987.
[19] S. Sharma, V. Kumar, and K. P. S. Rana. Automatic oscillations detection and quantification in process control loops using linear predictive coding. *Engineering Science and Technology, an International Journal*, 23(1):123–143, 2020.
[20] S. Sivalingam and M. Hovd. Use of cross wavelet transform for diagnosis of oscillations due to multiple sources. In *Proceedings of the 18th International Conference on Process Control*, pages 443–451, 2011.
[21] J. S. Smith. The local mean decomposition and its application to EEG perception data. *Journal of the Royal Society Interface*, 2(5):443–454, 2005.
[22] B. Srinivasan and R. Rengaswamy. Automatic oscillation detection and characterization in closed-loop systems. *Control Engineering Practice*, 20(8):733–746, 2012.
[23] N. N. Taleb. Statistical consequences of fat tails: Real world preasymptotics, epistemology, and applications. arXiv:2001.10488, Cornell University Library, 2020.
[24] M. R. Thirumalaisamy and P. J. Ansell. Fast and adaptive empirical mode decomposition for multidimensional, multivariate signals. *IEEE Signal Processing Letters*, (10):1550–1554, 2018.
[25] N. F. Thornhill, B. Huang, and H. Zhang. Detection of multiple oscillations in control loops. *Journal of Process Control*, 13(1):91–100, 2003.
[26] J. Wang, X. Xu, and X. Meng. The modified ensemble empirical mode decomposition method and extraction of oceanic internal wave from synthetic aperture radar image. *Journal of Shanghai Jiaotong University*, 20:240–250, 2015.
[27] Z. Wu and N. E. Huang. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1(01):1–41, 2009.
[28] L. Xie, X. Lang, J. Chen, A. Horch, and H. Su. Time-varying oscillation detector based on improved LMD and robust Lempel–Ziv complexity. *Control Engineering Practice*, 51(C):48–57, 2016.
[29] Z. Yan, J. Chen, and Z. Zhang. Using hidden Markov model to identify oscillation temporal pattern for control loops. *Chemical Engineering Research and Design*, 119:117–129, 2017.
[30] K. Zarzycki. Construction of nonlinear, multidimensional laboratory stand—Ball on plate system. Control algorithms synthesis. Master's thesis, Warsaw University of Technology, Institute of Control and computation Engineering, Warsaw, Poland, 2020.
[31] R. Zuo, J. Wei, X. Li, C. Li, C. Zhao, Z. Ren, Y. Liang, X. Geng, C. Jiang, X. Yang, and X. Zhang. Automated detection of high-frequency oscillations in epilepsy based on a convolutional neural network. *Frontiers in Computational Neuroscience*, 13:6, 2019.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dariusz Rocki

12 3D outliers in BIM supported electrical cable tracing

Abstract: One of the advantages of BIM technology is the ability to transfer construction risks to the design stage. It enables the detection of collisions not only between designed installations, but also between new and existing ones. The most useful tool in this regard is the Point Cloud—the actual 3D scanned image of the existing building. Although it is still an expensive solution, designers and investors more and more often employ this emerging opportunity to mitigate construction costs with 3D mapping. This document compares the traditional cable-trace design and the opportunities offered by BIM technology, in particular, 3D point clouds. The presented solutions demonstrate the possibilities related to the optimization of time and cost. This work discusses two main methods for point cloud generation: the photogrammetric method and laser scanning. It highlights outliers in both methods not only due to user errors, but also caused by hardware and method limitations. The impact of scanning outliers on further cable optimization is identified and assessed.

Keywords: BIM, point cloud, Revit, electrical installations, cable tracing

12.1 Introduction

For several years, there has been a dramatic increase in interest in BIM methodology. At the same time, each participant in the construction process defines this concept in a slightly different way [8]:

- Building Information Model,
- Building Information Modeling,
- Building Information Management.

The common element in these definitions is one, consistent database containing the necessary geometric (dimensions), as well as nongeometric information (data sheets, technical inspection dates, etc.), referred to as the BIM Digital Model [19, 21]. This model is constantly updated throughout the entire building's life cycle, from conceptual design, construction and management phases of the facility to its demolition. Thus, BIM makes possible the ordering and coordination of information exchanged between participants in the construction process [3]. An important advantage of the

Dariusz Rocki, Warsaw University of Technology, Institute of Control and Computation Engineering, Warsaw, Poland, e-mail: dariusz.rocki.dokt@pw.edu.pl

https://doi.org/10.1515/9783110729122-012

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

BIM model is the ability to detect and control risk, in particular, installation and logistics collisions, at an early stage, which significantly reduces construction costs [21].

Currently, the Polish Ministry of Infrastructure is conducting public consultations in connection with plans to implement BIM in public procurement. During the consultations, designers have identified the lack of preparation for working with this technology. Among the issues are the lack of BIM specialists and the costs associated with the purchase of new software [21]. Moreover, an important aspect is the necessity of increasing the amount of work needed to prepare documentation in BIM. In addition to the 2D CAD drawings that prepared earlier, it will be necessary to include additional information that may result in profit reduction for the design firms [21, 26]. The answer to these concerns is undoubtedly the automation of the required work. Available computer programs for creating 3D models are equipped with programming environments thanks to which it is possible to create customized solutions for optimizing the design [21].

In the design of all types of building structures, both infrastructure and cubature, point clouds are increasingly being used. Despite the high costs, it is more and more often obvious that laser scanning of a building at the design stage results in measurable benefits during the construction phase. This allows for the detection of collisions of the designed installations with the existing ones, and thus significantly simplifies the execution and shortens its time.

12.2 Literature review

The existing literature studies, to a large extent, can be divided into two main groups. The first focuses on obtaining a point cloud with a photogrammetric method or laser scanning and the scope of application [9, 10, 17] used to transform the photos or laser scans into the point cloud.

The second group comprises articles on obtaining a digital model based on a point cloud.

12.2.1 Point cloud

Modern construction uses point clouds obtained primarily by two methods:

- Photogrammetry [6, 7, 16]—used mainly in linear infrastructure objects, but also wherever the use of a laser scanner is difficult or impossible (e. g., underwater [18])
- Laser scanning [16]—other buildings.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Both this methods enable generation of a set of points defined by spatial coordinates. [16] Since the number of points depends mainly on the applied resolution, it determines the accuracy of the obtained point cloud and the size of the object.

12.2.1.1 Photogrammetry method

The photogrammetry method consists of taking multiple 2D photos and combining them into the 3D point cloud using computer software [16]. In photogrammetry, digital autographs and digital photogrammetric stations are very often used [10]. Experience in working with point clouds obtained with the photogrammetric method shows that, in most cases, nonmetric cameras also are also able to produce point clouds of sufficient quality [20]. This is also confirmed in the literature—many authors report that taking photos of appropriate quality is possible using most of the available nonmetric cameras, making photogrammetry a relatively low-cost method [2, 15, 16]. Software developers suggest using a high-resolution camera (5 Mpix or more [28]) equipped with 50 mm-focal length lens [1]. According to the results of research carried out at the Academy of Mining and Metallurgy, camera resolution above 3 Mpix does not significantly affect the obtained results [7]. In such a case, camera calibration is required and can be used with free software such as AeroSys [9].

The steps of the photogrammetry method are shown on Fig. 12.1. Virtual points on the facades are usually associated with characteristic places, such as the corners of doors, balconies or other point elements, that can be easily observed on the photograph. On each photograph, there should be at least three virtual points visible, but every additional point is an advantage. All photos are uploaded to the computer software, where the user detects the virtual points and the point cloud is generated [6].

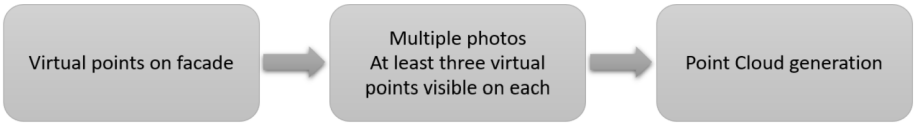


Figure 12.1: Photogrammetry method workflow.

The second advantage of this method is the relatively small demand for hard disk space. The photogrammetry method uses RAW format source files [1], which are much smaller than laser scans. This is particularly important for linear infrastructure facilities such as roads and railroads [17].

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

12.2.1.2 Laser scanning

Laser scanning is the preferred method for cubature construction. It is often used as an inventory aid tool for existing buildings [28] and also during their reconstruction. In such cases, it is particularly difficult to prepare the scanned space so that only the elements that should be in the final point cloud are visible.

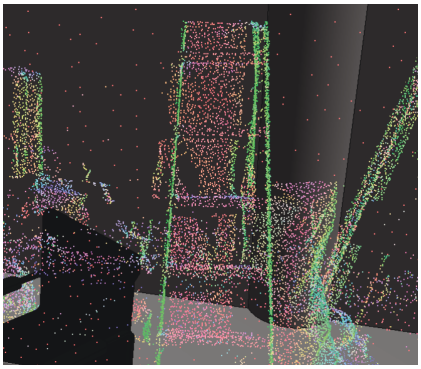
To perform a laser scan, the measuring points where the be set up are first determined. Points are distributed over the entire scanned space so that two other points are visible from each point [14]. Next, calibration points are placed in the form of a checkerboard (Fig. 12.2) or standardized measuring spheres on the scanned space. The next step is to make scans at each measurement point. In the end, all scans are combined into a point cloud using computer software.

Figure 12.2: Calibration points.

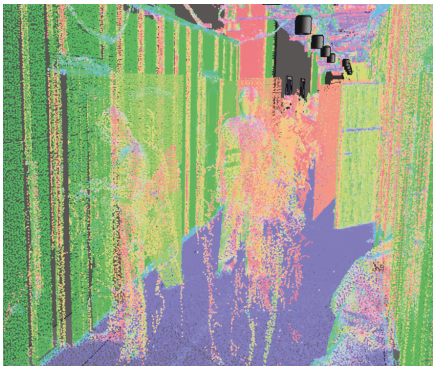
Laser scanning is a method very sensitive to errors, in particular, to errors caused by inadequate preparation of the scanned space and people walking across the scanned area (see Fig. 12.3). Such errors may render the obtained point cloud useless for comparing the model with the actual state.

The existing literature indicates problems in scanning glass surfaces, which partially absorb the scanner rays and result in incorrect information about the object [1]. However, as shown in Fig. 12.4, while there are some disturbances observed in the obtained point cloud, their identification is relatively simple.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use



(a) Ladder and other equipment



(b) Moving people

Figure 12.3: Various unwanted objects in the range of laser scanner.

Figure 12.4: Glass surface scanning.

12.2.2 Obtaining model from a 3D point cloud

Designers usually use a point cloud to visually compare the modeled objects with its real image. This makes it a highly valuable tool, but its use is often difficult due to its large size (see Section 12.4). For this reason, solutions are sought to enable the automatic creation of an object model based on a point cloud.

12.2.2.1 Planes detection

In 2007, the Hough transform and RANSAC algorithm were introduced as tools for automatic detection of roof planes from Lidar Data [23]. Tarsha-Kurdi, Landes and Grussen-

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

meyer compared the Hugh transform and RANSAC algorithm, and, since the processing time differences are negligible, they proposed an extension of the RANSAC algorithm. Their work resulted in an increase in the detection efficiency from 70 % to 80 % [23].

12.2.2.2 Lines detection

Lines detection is possible using the local feature descriptor LSSHOT 3D proposed by Xu, Tuttas and Stilla in 2018 [27]. The accuracy of presented descriptor reaches more than 70 % (tests) and enables tubes, toeboards and deck detection (50–70 % accuracy for the real object) [27].

12.3 Point cloud outliers

In terms of their type, point cloud outliers can be divided into several groups:
- Additional components
- Missing components [4]
- Incorrect geometry [4]
- Incorrect location and orientation [4]

12.3.1 Additional components

Scanning of existing buildings is very often carried out during renovation, reconstruction or during working hours. All this makes it very difficult or even impossible to properly prepare a building for a scan. As a result, the obtained point cloud contains many undesirable elements, such as people, equipment and other elements not strictly related to the scanned building object (see Fig. 12.3). Anil et al. in [4] named this type of outlier.

Figure 12.3b, apart from the moving people, shows also the walls and doors, which will be dismantled during further renovation works. To improve the clash detection, as well as to increase the value of the point cloud for the project team, more and more often the scan of the buildings is performed repeatedly at specific time intervals in such cases.

12.3.2 Missing components

Another example of very common *data collection errors* [4] are missing components, which can be divided into two main groups:

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Inventory errors
  Situation C in Fig. 12.7 shows a very common phenomenon, namely, the omission of an important element during the inventory. In this case, it is the missing structural column.
- Scanner limitations
  One of the basic limitations when creating a point cloud is the range of visibility, both for the laser scanning and the photogrammetric method. Incorrect placement of scanning points and the inability to make the scanned installation visible due to a lack of time or access results in a lack of information in the final point cloud.

12.3.3 Incorrect geometry, location and orientation

Anil et al. in [4] indicate incorrect geometry, location and orientation as three separate complicating issues, however all of these effects have similar causes:

- Limitations related to the device itself
  Scanner resolution or the camera lens may cause a shift of the point by a few millimeters, even as much as several centimeters.
- User errors
  Incorrect placement of scan points may cause disturbances in the geometry of the scanned space and missing elements described in the point cloud (see Section 12.3.2).
  Incompetent use of the scanner, or errors in measuring the coordinates of characteristic points, may result in significant deviations of the coordinates of the scanned points.
- Material characteristics of the elements in the scanned space
  It is very difficult to scan glass and mirror surfaces since the reflections can cause significant disturbances in the measurement [1] (see Fig. 12.4).

12.4 Theoretical office space arrangement

Based on the as-built documentation for the shell and core phase, as well as a site visit and inventory of one sample Warsaw building and a theoretical arrangement of the office space is created. Cost analyzes of electric cables are carried out for three different locations of the electrical distribution board supplying the arranged space. In each case the same number of electric circuits and the same distribution board equipment are maintained. The dimensions of the designed distribution board are 120 cm × 20.5 cm × 183 cm (length × width × height). Designed distribution board elevation is shown in Fig. 12.5.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

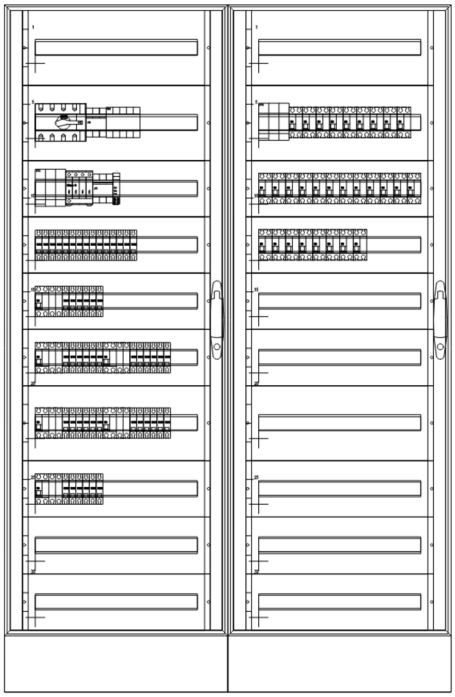


Figure 12.5: Designed distribution board.

The point cloud that was used during the described research concerned a fragment of one floor of an office building (approx. 2,000 m^2) and with the number of points exceeding 300 million, and its size exceeded 1 GB. The point cloud was provided by ARUP Poland.

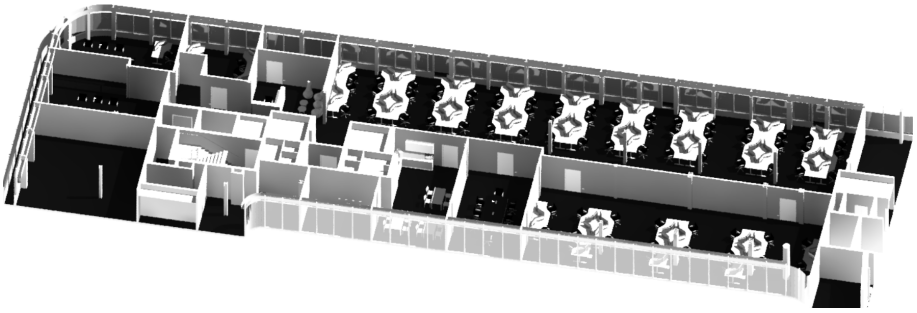
In the next step, a point cloud obtained with a laser scanner is superimposed on the created model of the building. Then the feasibility analysis is performed, and the results obtained in the first step were corrected.

12.4.1 Distribution board location—case study

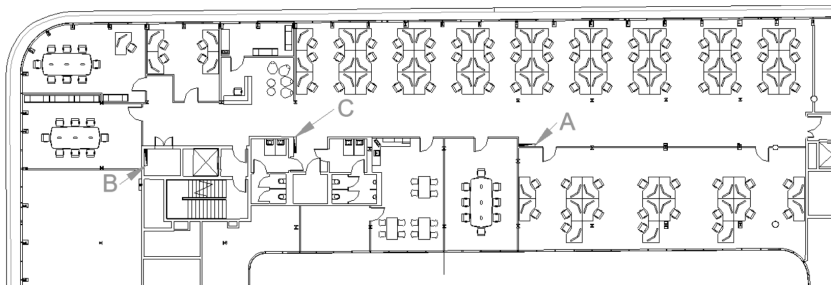
The designed office space (Fig. 12.6) includes the director's room, reception area, two open-space areas, kitchen facilities and two conference rooms. The existing electrical shaft is located in the western part of the arranged space in the immediate vicinity of the staircase. The electrical shaft is equipped with an existing busbar from which the tenant's distribution board will be supplied. This is also the default location for the tenant's distribution board, however, as will be shown subsequently, it is not an optimal location.

The designed distribution board is divided into two main sections. The first section supplies lighting installation, common sockets, hand dryers, kitchen equipment, etc.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use



(a) 3D view



(b) Two-dimensional projection

Figure 12.6: Possible location of a distribution board during the design phase: A—distribution board located in the corridor between two open space areas, next to the conference room, B—default distribution board location (electrical shaft), C—distribution board located next to restroom area.

The second section is designed to supply computer loads and rack cabinets in the IT room. All electrical connections have been modeled using Autodesk Revit software.

The locations of the tenant's distribution board shown in Fig. 12.6. Average cable prices were estimated on the basis of data available on the manufacturers' websites (see Table 12.1). It should be noted that actual prices may vary due to discounts agreed individually with each electrical contractor. The given prices refer to a running meter of the cable.

Table 12.1: Average cable cost.

Cable type	Bitner [5]	Technokabel [24]	Lapp Poland [12]	Average
N2XH 3×1.5	5.547	9.570	11.15	8.75
N2XH 3×2.5	13.327	12.440	16.07	13.95
N2XH 5×35	73.125	196.750	271.84	180.57

Table 12.2 shows that the most advantageous place for the electrical distribution board is at point A or C.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 12.2: Cable cost summary.

	A		B		C	
	Length	**Cost**	**Length**	**Cost**	**Length**	**Cost**
N2XH 3 × 1.5	645 m	5,650 PLN	890 m	7,796 PLN	716 m	6,272 PLN
N2XH 3 × 2.5	1845 m	25,738 PLN	2714 m	37,860 PLN	2024 m	28,234 PLN
N2XH 5 × 35	37 m	6,681 PLN	9 m	1,625 PLN	23 m	4,153 PLN
Total cost		*38,069 PLN*		*47,282 PLN*		*38,660 PLN*

12.4.2 Distribution board location—point cloud confrontation

According to the previousanalysis, points A and C should be considered as the location for the distribution board (see Table 12.2).

Figure 12.7 shows the electrical distribution boards placed in points A, B and C and the point cloud. In point A there is collision between the designed distribution board and the visible existing partition wall. However, the designed office-space arrangement requires the demolition of this wall, and therefore the visible collision is apparent.

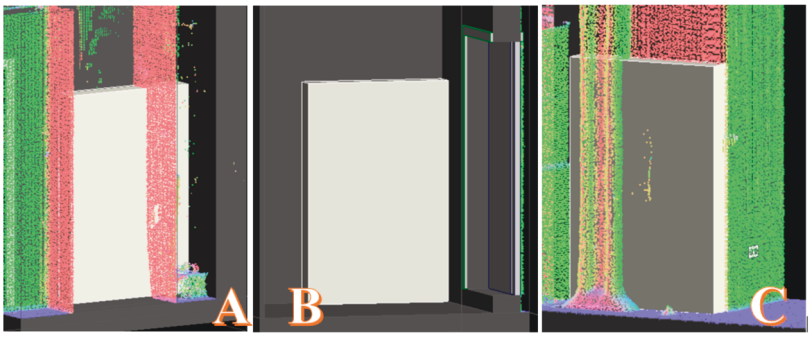


Figure 12.7: Point cloud and distribution board locations A, B and C.

No collision at point B results from the omission of the electric shaft at the stage of laser scanning of the building. However, an on-site inspection confirmed no collision at this point.

Figure 12.7c shows the collision between the designed distribution board and the existing structural column, which has not been included in the as-built documentation of the facility nor during the inventory, which resulted in its absence in the architectural and structural models. This collision indicates that placing the designed distribution board at this point is impossible. Unfortunately, this is a common mistake in design practice that could result in serious and costly problems on the construction site. In such a situation, point C is rejected as a possible distribution board location.

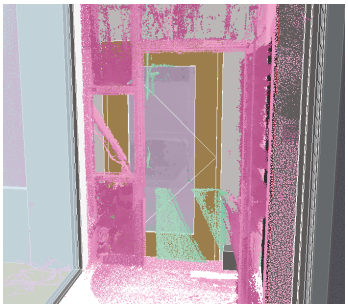
EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

The obtained architectural model also did not take into account the steel truss under the slab. The observed structural elements will affect the locations of the final cable traces. It will also affect the costs of installation and construction.

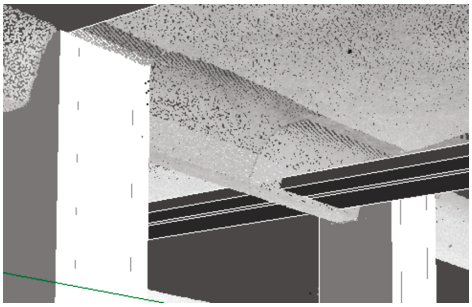
12.4.3 Outliers samples

The as-built documentation, prepared in a hurry, very often has significant shortcomings. Especially in older buildings, where renovation is relatively frequent, the documentation that does not include new or rebuilt elements. In such cases, an on-site inspection, combined with an inventory is carried out, which, in turn, for large areas is often imprecise.

A point cloud is a relatively easy way to identify discrepancies between a digital model and the actual state. Figure 12.8 shows examples taken from authentic models made in Autodesk Revit, with the point clouds overlaid.



(a)



(b)

Figure 12.8: Incorrect location of the door in the architectural model in relation to the actual state (a). The structural beam omitted during the inventory. Collisions of cable trays with the existing beam visible after laser scanning (b).

12.5 Point cloud study

Sample wall point clouds (see Fig. 12.9) shows some of a point cloud's limitations. Laser scanner detects just a surface, not a wall itself. And what is more important, the final points locations depends on the reflection intensity factor, which can be described as the possibility that the scanned point was detected in the right position. Therefore the point cloud surfaces are not flat.

The theoretical space arrangement described in Section 12.4 was prepared on the basis of a real architectural model and a real point cloud, the size of which is over 1 GB and contains over 300 millions of points, while the scanned area is only half of one

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

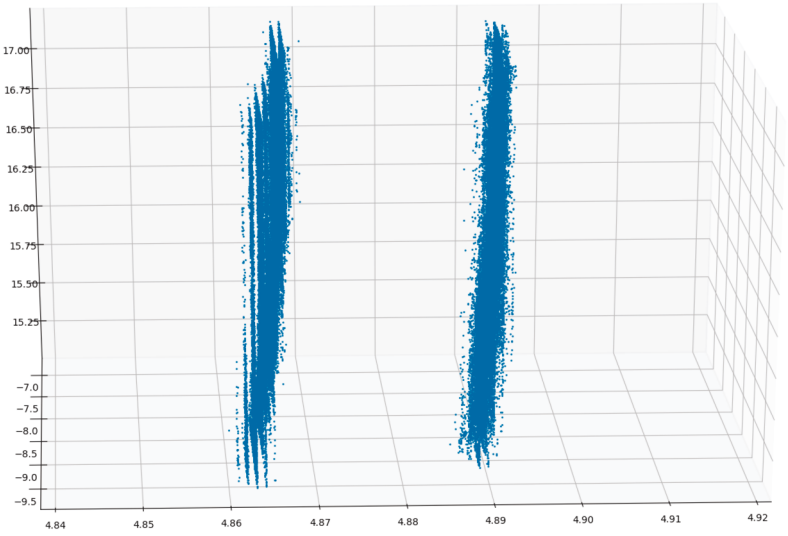


Figure 12.9: Sample wall point cloud—*57,975* points.

floor in a five-story office building. Therefore, it is easy to imagine the size of point clouds used in the building digital model. So, one of the basic assumptions of working with a point cloud should be to limit the number of points and use only points characteristic for a given surface in further work.

The laser scanners only detect obstacles in the path of the rays, therefore element categories (walls, lighting luminaries or furniture) detection is impossible. For example, scanned wall appear as two surfaces—see Fig. 12.9. In addition, the distances between points on the same surface can reach up to 3 cm, therefore the obtained surface is not flat.

BIM uses point clouds prepared as *.rcp files contain the seven-dimensional point description (compare with [13]):

$$[x, y, z, i, r, g, b], \tag{12.1}$$

where x, y, z indicate the 3D location, i is the reflection intensity factor and r, g, b represent the color in RGB.

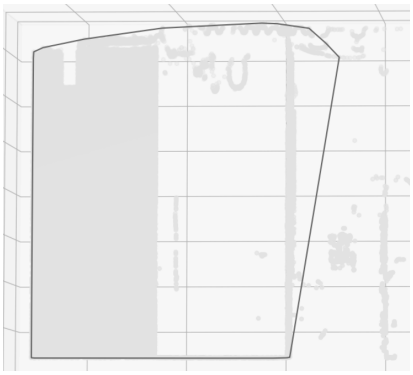
12.5.1 Surface detection, clustering and weighted median filter

Surface detection is performed in three planes: x, y, z. Each point was scanned with a certain reflection intensity factor i, which can be used as a weight in weighted median filter. RGB parameters are important only for the graphical representation, so they are omitted in further considerations.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

The first stage of surface detection was cutting the point cloud successively along the OX, OY and OZ axes into 10 cm-wide fragments. Experimentally, the minimum number of points in each fragment is set at 30,000 points and each fragment containing a lower number of points is omitted in further studies, as in Fig. 12.10.

(a) *6,845* points



(b) *45,197* points

Figure 12.10: Points clouds with different numbers of filtered points.

For each fragment, using the *sklearn* python library, clustering was carried out, allowing to define groups of points located on the same surfaces. Each cluster was saved in a separate *.txt file, the name of which contained information about the cutting axis, fragment number and cluster number.

12.5.1.1 Weighted median filter

A point cloud can be thought of as a 3D set of pixels analogous to graphic images. In this approach, noise reduction is possible through the use of a *median filter*—one of the nonlinear filters [22]. However, in the case of a point cloud, in addition to the x, y, z coordinates, we also obtain information about the intensity of the reflection i, which can be interpreted as the certainty with which the laser scanner indicates the correct position of a given point. Therefore, this additional parameter should be used as a weight in a *weighted median filter*.

The reflection intensity factor ranges from 0 to 255, therefore, each weight should be divided by 255.

$$\widetilde{i} = \frac{i}{255}. \tag{12.2}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

To eliminate the shift towards the center of the coordinate system, three possibilities are considered for each point:

$$\forall k \in (1,n) \begin{cases} a_k * \widetilde{i}_k, \\ a_k, \\ a_k * (1 + (1 - \widetilde{i}_k)), \end{cases} \tag{12.3}$$

n—number of points; a—x, y or z—coordinates depending on the cutting axes.

Unlike the graphic images, where the median filter is applied sequentially to small parts of the whole, the point cloud uses a median filter for the entire detected cluster. The weighted median was calculated for the coordinate consistent with the cutting axis, and then the coordinates along that axis for all points were aligned with the calculated median value. This makes it possible to obtain surfaces that most accurately reflect the location of real planes—walls, ceilings or floors. But, the summary files weight is still not acceptable from a modeling point of view. Therefore, the surfaces have to be limited to the edges that define them.

Initially, the *ConvexHull* [11] function from the *scipy.spatial* [25] library was used to determine the surface edges. Its unquestionable advantage is quite fast operating time. It turned out, however, that this algorithm did not cope well with the detection of edge points in concave polygons. Figure 12.10 (right) shows the yellow points and a line around the cluster. The figure shows a wall in the form of a yellow patch, and a fragment of a corridor bounded by another wall perpendicular to the cutting axis. Meanwhile, the *ConvexHull* algorithm framed also around the corridor. This makes it necessary to use a different algorithm.

12.5.2 Normalized point cloud

The total number of points shown in Fig. 12.10 right is 45,197, while there are 41,935 points in the cluster. The other 3,262 points are the outliers, which should be omitted in further consideration.

Based on the extreme values of the points coordinates, a normalized mesh is created, which enables generation of a normalized point cloud. As a result of this operation, some repeated values are omitted, and therefore the number of considered points is limited to 36,426. A fragment of a point mesh and a normalized point cloud are shown on Fig. 12.11.

Each point from the normalized point cloud is compared to the adjacent mesh points, and, in the case where an empty spot was found in its vicinity, it is treated as a boundary point. The last step is to find the corner points in the obtained edges and arrange them in order—shown as yellow points on Fig. 12.11.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

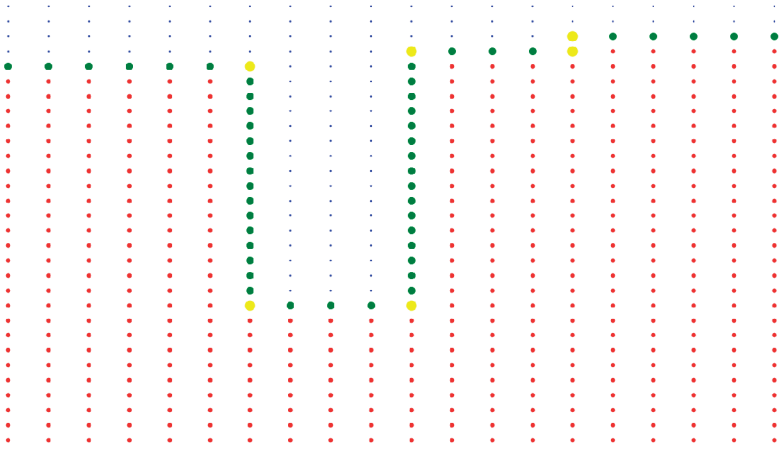


Figure 12.11: Fragment of a point mesh (blue), normalized point cloud (red), boundary points (green) and corner points (yellow).

12.5.3 Revit implementation

12.5.3.1 Clash detection

The corner points obtained as a result of the actions described in Section 12.5 are imported into the Dynamo development environment cooperating with Autodesk Revit. On their basis, surfaces are created using the default Dynamo nodes *Point.ByCoordinates* and *Surface.ByPerimeterPoints*, on which the elements of electrical installations exported from Revit, such as cables and cable trays, are then applied. The collision between Revit elements and surfaces are detected by use of the default Dynamo node *Geometry.Intersect*.

12.5.3.2 Cable traces changes

Th detected collisions are removed by adjusting the location of cable traces to the surfaces obtained from the point cloud. The task is done manually with constant monitoring of changes in Dynamo. After the collision has been removed, another cable length analysis is carried out.

The change of routes resulting from the necessity to eliminate collisions of cables and cable trays with the surfaces of the point cloud, slightly affected the cost:

- Location A: increase from 38,069 PLN to 40,265 PLN,
- Location B: increase from 47,282 PLN to 49,582 PLN.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

The main reason for the cost increase is the need to bypass the collision areas, and in one case a major change of route was necessary. However, the comparison of the values listed in Tables 12.2 and 12.3 confirms that the best location for the distribution board is location A (Fig. 12.6).

Table 12.3: Final cable-cost summary.

	A		B	
	Length	**Cost**	**Length**	**Cost**
N2XH 3 × 1.5	674 m	5,897 PLN	942 m	8,242 PLN
N2XH 3 × 2.5	1920 m	26,784 PLN	2821 m	39,353 PLN
N2XH 5 × 35	42 m	7,584 PLN	11 m	1,986 PLN
Total cost		*40,265 PLN*		*49,582 PLN*

12.6 Summary

This study aimed to combine the classic design of electrical installations with an approach using new technologies, in this case, point clouds. Developing an exact architectural model based solely on measurements turns out to be an extremely difficult task. In the presented example, the difference in costs between the installation designed on the basis of the digital model and the installation compared to the real image of the building in the form of a point cloud is 5 % of the installation cost. However, it should be noted that the presented considerations relate to a relatively small office space.

The method of eliminating outliers seems to be sufficient for the needs of electrical installations, in particular, the main cable traces. The created computer program, which divided the point clouds into the fragments, and then detected clusters of points and made their selection based on the number of points in the cluster, fulfilled its task. The resulting output data made it possible to work with Dynamo software, which is very sensitive to large amounts of data. However, it is worth noting that the running time of the program was unsatisfactory because the point-cloud analysis took seven hours, and the manual collision removal took additional time. Therefore, the possibility of optimizing the code and the algorithms used should be considered.

Bibliography

[1] Agisoft Photoscan User Manual: Professional edition, version 1.3. Agisoft LLC, 2017.
[2] E. Alby, E. Smigiel, P. Assali, P. Grussenmeyer, and I. Kauffmann-Smigiel. Low cost solutions for dense point clouds of small objects: Photomodeler scanner vs. David laserscanner. In *22nd CIPA Symposium, October 11–15*. Kyoto, 2009.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[3] E. Andrejczyk. Technologia BIM—Antidotum na chaos? *Materialy budowlane*, 2019.
[4] E Anil, P. Tang, B. Akinci, and D. Huber. Deviation analysis method for the assessment of the quality of the as-is building information models generated from point cloud data. *Automation in Construction*, 35:507–516, 2013.
[5] Bitner. Bitner cables—The price-list valid from 15.12.2020. 2020.
[6] A. Braun, S. Tuttas, A. Borrmann, and U. Stilla. A concept for automated construction progress monitoring using BIM-based geometric contracints and photogrammetric point clouds. *Journal of Information Technology in Construction*, 20:68–79, 2015.
[7] E. Glowienka, B. Jankowicz, B. Kwoczynska, P. Kuras, K. Michalowska, S. Mikrut, A. Moskal, I. Piech, M. Strach, and J. Sroka. Modelowanie 3D z kamer niemetrycznych w analizach przestrzennych. In *Fotogrametria i skaning laserowy w modelowaniu 3D*, pages 115–135. Wyzsza Szkola Inzynieryjno-Ekonomiczna z siedziba w Rzeszowie, 2015.
[8] D. Kasznia, J. Magiera, and P. Wierzchowski. *BIM w praktyce. Standardy, wdrozenie, case study*. Wydawnictwo Naukowe PWN SA, Warsaw, 2017.
[9] B. Kwoczynska. Elaboration of architectural objects using methods of digital photogrammetry. *Infrastructure and ecology of rural areas*, 3/2010 Polska Akademia Nauk:65–74, 2010.
[10] B. Kwoczynska. Surveying and visualization of architectural objects prepared on the basis of metric and non-metric photos. *Infrastructure and ecology of rural areas*, 1/II/2012 Polska Akademia Nauk:53–66, 2012.
[11] H. Langer, S. Falsaperla, and C. Hammer. *Advantages and Pitfalls of Pattern Recognition, Selected Cases in Geophysics*. Computational Geophysics, volume 3. Elsevier, 2020.
[12] Lapp Poland. Lapp Poland—Price-list valid from 10.05.2020. 2020.
[13] M. Levoy and T. Whitted. The use of points as a display primitive. Technical Report Technical Report 85-022, Computer Science Department, University of North Carolina, Chapel Hill, NC, 1985.
[14] H. Macher, T. Landes, and P. Grussenmeyer. From point clouds to building information models: 3d semi-automatic reconstruction of indoors of existing buildings. *Applied sciences*, 7:1030, 2017.
[15] R. Mlambo, I. Woodhouse, F. Gerard, and K. Anderson. Structure from motion (SfM) photogrammetry with drone data: A low cost method for monitoring greenhouse gas emissions from forests in developing countries. *Forests*, 8:68, 2017.
[16] D. Moon, S. Chung, S. Kwon, J. Seo, and J. Shin. Comparison and utilization of point cloud generated from photogrammetry and laser scanning: 3d world model for smart heavy equipment planning. *Automation in Construction*, 98:322–331, 2019.
[17] A. Paixao, R. Resende, and E. Fortunato. Photogrammetry for digital reconstruction of railway ballast particles—A cost-efficient method. *Construction in Building Materials*, 19:963–976, 2018.
[18] C. Perry. A quick, easy and non-invasive method to quantify coral growth rates using photogrammetry and 3d model comparisons. *Methods in Ecology and Evolution*, 11:714–726, 2020.
[19] M Rahmani, A. Stoupine, S. Zarrinmehr, and W. Yan. Optimo: A BIM-based multi-objective optimization tool utilizing visual programming for high performance building design. In *Education and Research in Computer Aided Architectural Design in Europe*, 2015.
[20] I. Reljic, I. Dunder, and S. Selan. Photogrammetric 3D scanning of physical objects: Tools and workflow. *TEM Journal*, 8:383–388, 2019.
[21] D Rocki. Automatyzacja prac projektowych przy użyciu programu panel connector. *Pomiary Automatyka Robotyka*, 3/2020 DOI: 10.14313/PAR 237/43 Siec Badawcza Lukasiewicz—Przemyslowy Instytut Automatyki i Pomiarow PIAP:43–49, 2020.
[22] L. Tan and J. Jiang. Chapter 14.2 Image Filtering Enhancement. Median Filtering. In *Digital Signal Processing. Fundamentals and Applications*, 3rd edition. Academic Press, An imprint of Elsevier, 2019.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[23] F. Tarsha-Kurdi, T. Landes, and P. Grussenmeyer. Hough-transform and extended RANSAC algorithms for automatic detection of 3d building roof planes from Lidar data. In *ISPRS Workshop on Laser Scanning 2007 and SilviLaser2007, September 2007, Espoo, Finland*, pages 407–412, 2007.
[24] Technokabel. Technokabel—Price-list valid from 11.03.2020. 2020.
[25] The SciPy community. SciPy documentation, SciPy User Guide—Spatial data structures and algorithms (scipy.spatial). 2020.
[26] A. Tomana. Projektowanie BIM—Tansze czy drozsze? *Materialy budowlane*, 2019.
[27] Y. Xu, S. Tuttas, L. Hoegner, and U. Stilla. Reconstruction of scaffolds from a photogrammetric point cloud of construction sites using a novel 3D local feature descriptor. *Automation in Construction*, 85:76–95, 2018.
[28] N. Yastikli. Documentation of cultural heritage using digital photogrammetry and laser scanning. *Journal of Cultural Heritage*, 8:423–427, 2007.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jairo Viola, Carlos Rodriguez, Derek Hollenbeck, and YangQuan Chen

13 A radio frequency impedance matching control benchmark and optimal fractional-order stochastic extremum seeking method

Abstract: An impedance-matching network is mandatory between the electric source and the load to obtain the maximum power transfer in high power applications like plasma etching technology used on semiconductor wafer processing. This chapter presents the problem definition and guidelines for a radio frequency (RF) impedance matching benchmark and its control using an optimal fractional-order stochastic extremum seeking (FO-SESC) method. The benchmark system (files available online) allows the quantitative study and evaluation of multiple RF impedance matching algorithms. The goal of the benchmark is matching the source and load impedances real and complex conjugate terms in a minimum time. The reflection coefficient is used as an evaluation metric to measure the reflected power, which has to be zero to guarantee the maximum power transfer. Some requirements for closed-loop controllers proposed in the literature are tested as sample applications on this benchmark to check and evaluate their stability performance against various initial conditions and load impedance values. A novel Feedforward Control (FF) profile generation method, with a primitive sample feedback control strategy (proportional-integral (PI) controller), is used as a baseline for the benchmark.

Keywords: impedance matching benchmark, feedforward control, fractional order extremum seeking, dithering noise

13.1 Introduction

Impedance matching is the practice of designing load and source impedances to maximize the power transfer minimizing the reflected energy from the load. There are two main impedance matching techniques employed in semiconductor industry. One is called an automatic matching network, where the value of electrical components like

Jairo Viola, Derek Hollenbeck, YangQuan Chen, Department of Mechanical Engineering, University of California, Merced, USA, e-mails: jviola@ucmerced.edu, dhollenbeck@ucmerced.edu, ychen53@ucmerced.edu
Carlos Rodriguez, Ensenada Center for Scientific Research and Higher Education (CICESE), Ensenada, Mexico, e-mail: crodriguez@cicese.edu.mx

https://doi.org/10.1515/9783110729122-013

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

"loss-less" passive electrical elements (such as capacitors and inductors) are adjusted automatically using a control algorithm. The second technique is the so-called generator frequency tuning, where the RF frequency at the generator is modified to reduce the reflected power back to the generator. There are cases when both techniques are used together to obtain a larger window of tune space [23].

The matching network can be designed using various configurations. The most used architectures are the L [19], Π, and T networks. In the case of the L-Type network shown in Fig. 13.1, it can provide a unique solution using only two variables, a series capacitor C_m and a shunt capacitor C_t. Thus, the network can be tuned easily for any feasible load impedance [18].

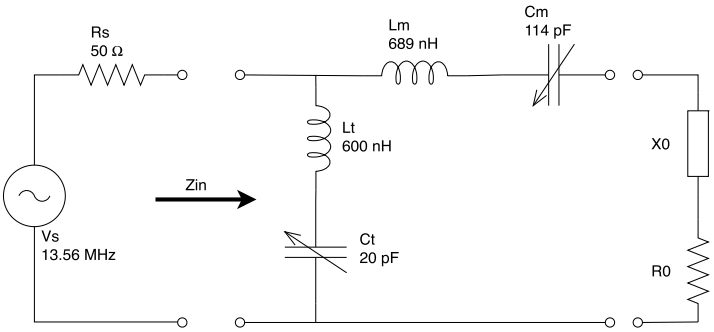


Figure 13.1: L-Type network configuration.

The load impedance of the network is $R_0 + jX_0$. Thus, by changing the capacitance of the variable capacitors C_t and C_m, the input impedance Z_{in} can be controlled. Usually, the impedance measurement through the voltage and current on the load side employs an expensive sensor due to the high voltage and current values. For this reason, the best point to measure the impedance of the network is on the source side. It means that the impedance value Z_{in} is equivalent to the right side of the network as shown in Fig. 13.1. The matching is achieved usually when $Z_{\text{in}} = R_s = 50\,\Omega$. For a general case, $Z_{\text{in}} = Z_s^*$ where Z_s is the source impedance. In [19] is presented a generic detector that provides complete information of complex parameters, impedance (Z), admittance (Y) and reflection coefficient (Γ). However, for some cases, partial information on those parameters can be sufficient for proper control, and simplified versions of the detector can be applied.

Each time the plasma parameters or set-point change (i. e., RF power, pressure, gas mixture), the plasma impedance also changes [1]. Consequently, the reflected power P_r increases, and the controller has to tune the matching box to revert this effect at the minimum time. The reflection coefficient Γ can measure this effect by (13.1),

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

where P_{in} is the input power.

$$|\Gamma| = \sqrt{\frac{P_r}{P_{\text{in}}}} = \left|\frac{Z_{\text{in}} - 50\Omega}{Z_{\text{in}} + 50\Omega}\right|. \tag{13.1}$$

Beginning in 1993, the research literature has been reported work on various control strategies intended to solve the requirements of a full-time dedication problem using neural networks, genetic algorithms, deterministic tuning with look-up tables and adaptive systems and nonlinear control systems [1]. In all these cases, the load impedance is not affected by the matching conditions. In [1], a hierarchical structure controller was proposed that is composed of a higher level coarse controller that drives the system close to the matching point and a lower level feedback controller for fine tuning. Also, Firrao [7] used two steps: First the imaginary part of the load impedance is tuned to (almost) zero using a series (or shunt) reactance, and, second, the resulting real part is transformed to the target real value with a tunable transformer. Hirose in [9] used a Seek + Follow control for a robust behavior using the phase and amplitude of the impedance as reference variables. An adaptive requirement for full-time dedication was proposed by [19] and presented an overview of adaptive controlled matching networks. Ishida [11] presented the same idea as Hirose, but with different conditions, and assumed knowledge of the exact value of the network components. The results showed a controllable region but did not analyze the impedance region corresponding to the series and shunt capacitors values. A tracking controller is shown in [12] with initial conditions close to the desired value.

The extremum control also is applied to solve this problem in [23]. Besides, a centralized controller that uses feedback compensation to regulate the power, Feedforward correction of impairments to the delivery of RF power is presented in [4, 5]. Finally, binary search can be applied with a great improvement in the convergence time, but it does not guarantee a monotonic decrease of the reflected power over time [20, 21].

In addition, Extremum Seeking Control (ESC) is an adaptive model-free technique that allows optimizing smooth convex function problems, in which the extremum or optimal point may be unknown and variable [2, 14, 23]. For the requirements of a full-time dedication problem, [23] showed its application in an L-Type matching network, using sliding mode and perturb and observe ESC controllers. This approach is extended into fractional-order ESC by [12, 22], where the fractional-order sliding mode extremum seeking control was proposed for a class of nonlinear systems, including the requirements for full-time dedication for single-capacitor L-Type networks. Likewise, a fractional-order version of the stochastic perturb and observe ESC controller introduced by Liu in [13] was presented by [10]. In that case, the perturbation signal is replaced by a bounded Fractional-order dithering Random Noise (Gaussian (fGn) or Symmetric Alpha-Stable (SαS)) with Long-Range Dependence (LRD) determined by the noise Hurst exponent.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

This chapter presents an RF impedance matching benchmark that can be employed to evaluate the performance of different control strategies. The system employs an L-Type network as matching circuit. The model considers not only the electrodynamics of the network electrical components but also the mechanical behavior and dynamics of the motor actuators that adjust the variable capacitors C_m and C_t. It include the effects like the rate limit and the normalized motor initial conditions (0% to 100%) for a closer approach to an industrial environment. A Feedforward + PI Controller is introduced to analyze the robustness and convergence time of the matching network, which acts as a baseline for the benchmark. Likewise, the control strategies presented above are tested using this benchmark. The controllers are assessed using as performance criteria the convergence time, the integral of the reflection coefficient, and reflected power mean and peak power.

On the other hand, a Stochastic Perturb and Observe FO-ESC controller (P&O FO-SESC) is employed for the impedance matching problem using the reflected power as an optimization variable and dithering noise (fractional Gaussian random noise) as perturbation signal. The P&O FO-SESC controller performance is compared with the Perturb and Observe sinusoidal ESC controller (P&O ESC) for different plasma loads and capacitor initial conditions (IC).

13.2 L-type network matching description and modeling

The RF power generator presented in Fig. 13.1 can deliver a maximum of 15 kW at a frequency of 13.56 MHz with a 50 – Ω generator output impedance. Two stepping motors control the series (C_m) and shunt (C_t) capacitors, where the speed of the motor limits the change rate. In this case, the rate limit for the motors is 100 %/s with a second order dynamic and a time constant of 20 ms. The current profile of the motor uses the maximum rate limit that does not guarantee a global monotonic decrease of the reflected power. The details of the matching box shown in Table 13.1, and the range of

Table 13.1: Parameters of the matching box.

	Minimum	Maximum
C_m	114 pF	445 pF
C_t	20 pF	205 pF
L_m	689 nH	
L_t	600 nH	

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

load impedance in the real and imaginary axes for the capacitor values are shown in Fig. 13.2.

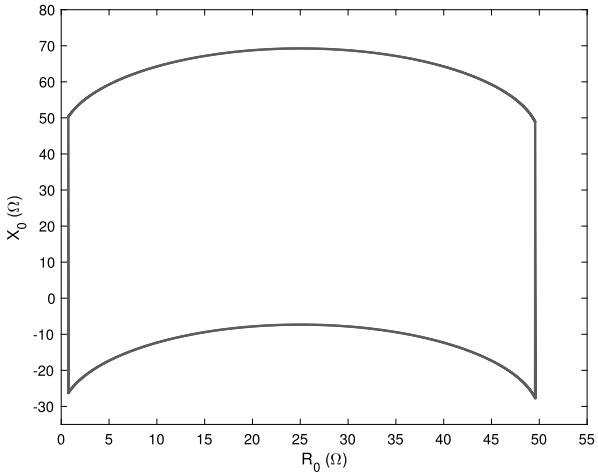


Figure 13.2: Range of the load impedance (real and imaginary parts) for the L-type network.

As mentioned before, there are different configurations a of matching box to match the impedance between the source and the load. The reference signal to control the matching box with proper behavior of the system depends on the network configuration. According to [19, 21] for T and L down-converting networks, the correct variable is the impedance. Nonetheless, for the Π and L up-converting, it is the admittance. The model of the matching network, including the load impedance, is given by (13.6). Initially, the impedance in the section $Z_1 = f(L_m, C_m, X_0, R_0)$ is

$$Z_1 = R_0 + j\left(\omega L_m - \frac{1}{\omega C_m} + X_0\right), \tag{13.2}$$

$$Y_1 = \frac{1}{Z_1} = \frac{R_0 - j(\omega L_m - \frac{1}{\omega C_m} + X_0)}{R_0^2 + (\omega L_m - \frac{1}{\omega C_m} + X_0)^2}, \tag{13.3}$$

$$Y_1 = G_1 - jB_1. \tag{13.4}$$

Thus, the input admittance can be obtained as:

$$Y_{\text{in}} = Y_1 - j\left(\frac{\omega C_t}{\omega^2 C_t L_t - 1}\right), \tag{13.5}$$

$$Y_{\text{in}} = G_1 - j\left(B_1 + \frac{\omega C_t}{\omega^2 C_t L_t - 1}\right) \tag{13.6}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

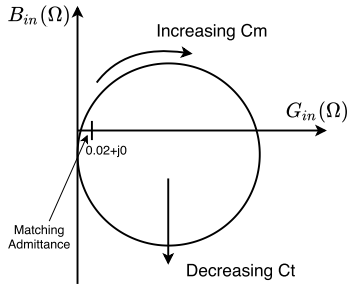


Figure 13.3: Plot of phase plane with possible trajectories of Y_{in} as C_m and C_t are changed.

with

$$G_1 = \frac{R_0}{R_0^2 + (\omega L_m - \frac{1}{\omega C_m} + X_0)^2}, \tag{13.7}$$

$$B_1 = \frac{\omega L_m - \frac{1}{\omega C_m} + X_0}{R_0^2 + (\omega L_m - \frac{1}{\omega C_m} + X_0)^2}. \tag{13.8}$$

Taking the equations (13.7), (13.8) and applying Linear Fractional Transformation (LFT), the phase plane of the system is shown in Fig. 13.3. The global behavior of the matching network can be understood by plotting the trajectories of Y_{in} in phase space because C_m and C_t vary. Thus, the manipulation leads to

$$\left(G_{\text{in}} - \frac{1}{2R_L}\right)^2 + \left(B_{\text{in}} + \frac{\omega C_t}{\omega^2 C_t L_t - 1}\right)^2 = \left(\frac{1}{2R_L}\right)^2, \tag{13.9}$$

indicating that the possible values of G_{in} and B_{in} lie on the circle of radius $1/2R_L$, centered on the point $(1/2R_L, -\omega C_t/(\omega^2 C_t L_t - 1))$. The radius is determined by the load impedance. Analyzing the geometric of the phase plane, if the pair $(G_{\text{in}}, B_{\text{in}})$ is in the upper arc of the circumference $(B_{\text{in}} > -\omega C_t/(\omega^2 C_t L_t - 1))$, the capacitor C_m keeps increasing because the match point can only be achieved with the intersection of the lower arc and the x-axis. In Section 13.4, it is explained how the Feedforward Control helps the Feedback Control keeping the correct path.

Figure 13.4 is a Smith Chart illustrating the value of the admittance Y_{in} for all IC of the matching box. Therefore, the controller has to be able to drive the admittance value to the center of the Smith Chart that represents $Y = 0.02\,\text{S}$ or $Z = 50\,\Omega$. The red circle and the x-axis represent the unitary circle of admittance ($G_{\text{in}} = 0.02\,\text{S}$) and the imaginary reference ($B_{\text{in}} = 0\,\text{S}$). The white region represents the unfeasibility region according to the value of the capacitors of the network.

The benchmark block diagram is shown in Fig. 13.5. The blocks *Rate Limit*, *Motor Dynamic*, *Network Model*, *Sensor Distortion* and *Amp & Phase Noise* remain the same for all the strategies, and, by replacing the *Controller* and *Gamma* blocks, it is possible to evaluate multiple strategies. The MATLAB code of the benchmark system is avail-

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

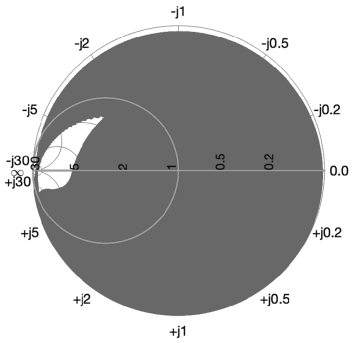


Figure 13.4: Admittance Smith Chart for the L-type network.

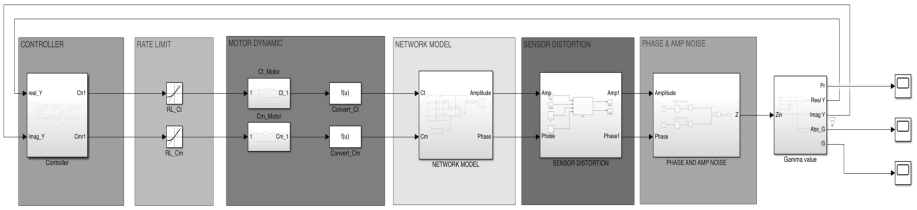


Figure 13.5: Block diagram for the RF impedance matching with L-Type Network.

able at [16]. The second-order dynamic of the motor with a time constant of 20 ms appears in the block *Motor Dynamic*. After that, the *Network Model* block takes as input the capacitors values to calculate the current impedance of the network described by (13.6). Then, the impedance amplitude and phase values can be disturbed on the *Network Distortion* and *Phase & Amp Noise* blocks by introducing distortion and random noise. Finally, the *Gamma* block computes the value of Γ using the impedance value of the network.

13.3 Feedforward control

Initially, a feedforward control is proposed as the baseline for the RF impedance benchmark. It uses the high rate limit of the motors, called (MAX-MAX). This profile can achieve a low convergence time with zero reflected power but does not guarantee a monotonic decreasing behavior for different IC. Therefore, a new profile design is proposed using three or four points as shown in Fig. 13.6. The knobs (r_i and r_j), respectively, with $i = 1, 2, 3$ and $1, \ldots, 5$, were used to build the 3-point and 4-point profiles.

To find the optimal values, we have the cost function

$$\min_{r_{1\ldots n}} J = w_1 T_{\text{tune}} + w_2 \max\left\{\frac{d}{dt}|\Gamma|\right\}, \tag{13.10}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

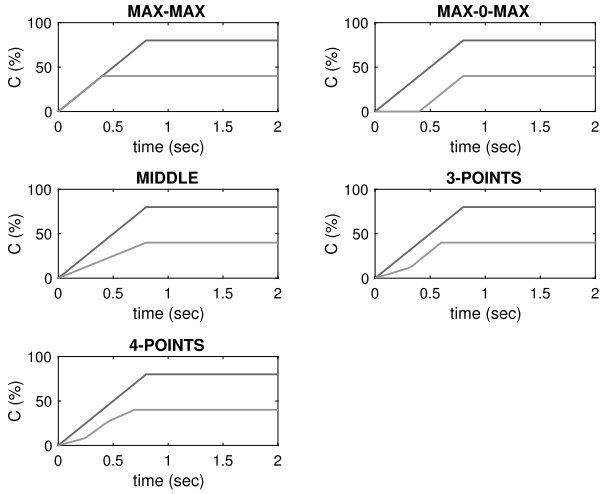


Figure 13.6: Comparison of the profiles used in the Feedforward.

where T_{tune} is the convergence time to 5 % of the reflection coefficient and the derivative of $|\Gamma|$ guarantee a monotonic decreasing behavior of the reflected power in the network. The best performance was obtained for $w_1 = 0.7$ and $w_2 = 0.3$. For each IC and impedance load value, the optimal values of the knobs were obtained, but storage of all this data requires a great amount of space and commutations, so ensemble values were found for the profiles given by (13.11) and (13.12).

$$r_1 = 0.6425, \quad r_2 = 0.3865, \quad r_3 = 0.7556, \tag{13.11}$$

$$r_1 = 0.6425, \quad r_2 = 0.3865, \quad r_3 = 0.7556, \quad r_4 = 0.6448, \quad r_5 = 0.3668. \tag{13.12}$$

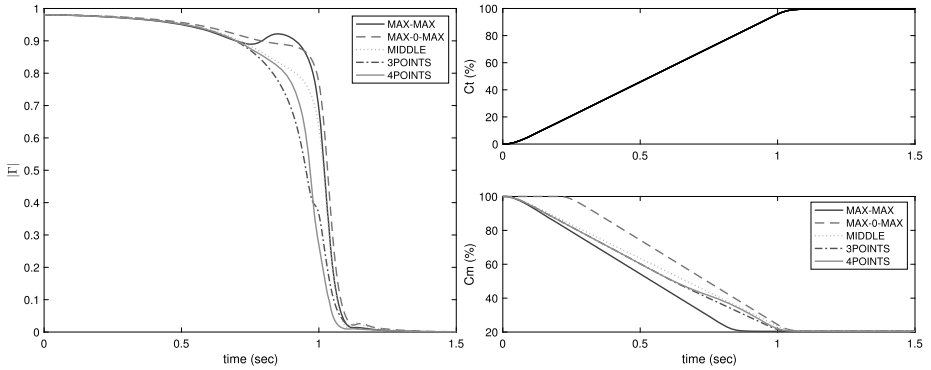
Figure 13.7 shows how the 3-points and 4-points profiles guarantee a monotonically decreasing behavior in the reflection coefficient. in comparison with the (MAX-MAX) profile.

In Table 13.2, the feedforward controller performance is compared using the profiles (MAX-MAX, MAX-0-MAX, MIDDLE, 3 POINTS, and 4-POINTS). The 3-points profile presents the lowest global index; therefore, the application of this profile over the (MAX-MAX) avoids power peaks and reduces energy consumption. The performance of the feedforward controller is evaluated using a set of performance criteria defined by (13.13)–(13.16). These indices are the convergence time of the controller, the integral of the reflection coefficient magnitude and the maximum and mean reflected power.

$$T_{\text{tune}} = \text{time to fall within 5 \% of } |\Gamma|_{\text{final}}, \tag{13.13}$$

$$G_{\text{int}} = \int_0^t |\Gamma(t)| dt, \tag{13.14}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use



(a) Performance of $|\Gamma|$ (b) Profile of the C_t and C_m capacitor

Figure 13.7: Performance of |Γ| using different profiles.

Table 13.2: Performance of the system for different profiles in the Feedforward Control.

PROFILES		T_{tune}		G_{int}		mean(P_r)		max(P_r)
		mean	max	mean	max	mean	max	mean
MAX-MAX		0.613	1.132	178.962	536	5.468	14.203	8.783
MAX-0-MAX		0.627	1.16	190.205	537.476	5.973	14.040	8.773
MIDDLE		0.618	1.158	182.079	537.177	5.535	13.753	8.754
3-P	Optimal	0.612	1.128	180.982	536.256	5.555	14.118	8.753
	Ensemble	0.610	1.14	179.875	536.359	5.527	14.118	8.759
4-P	Optimal	0.611	1.128	181.357	536.362	5.582	14.118	8.764
	Ensemble	0.611	1.148	180.229	536.350	5.533	14.118	8.757

$$\max(P_r) = \max_{i \in n} P_r(i), \tag{13.15}$$

$$\text{mean}(P_r) = \text{mean}(P_r). \tag{13.16}$$

13.4 Impedance matching feedback control

13.4.1 Proposed controllers in the literature

The RF impedance-matching benchmark compares the controllers proposed by Bacelli [1], Hirose [9], Ishida [11], Bezooijen [19], Cottee [3] and our proposed PI + FF. Bacelli used the reflection coefficient as a controllable variable with the control rule with τ as a certain time delay. In his work, one IC of the system is analyzed and show its robustness to variations (steps in the RF power, the gate valve position of O_2 and

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

A_r flow).

$$\begin{aligned} u_{C_t} &= \mathrm{Im}(\Gamma), \\ u_{C_m} &= -\left(1 - |\Gamma|\right)^2 \mathrm{Re}(\Gamma) + 0.1|\Gamma|\,\mathrm{sgn}\left(\Gamma(t) - \Gamma(t-\tau)\right) \end{aligned} \tag{13.17}$$

Hirose and Ishida [9] used a Seek+ Follow strategy with different switching conditions. Hirose proposes a switch between the controllers under the conditions

$$\begin{aligned} \text{Follow} &= \left\{35 \le |Z_{\text{in}}| \le 65\,,\ -30 \le \phi \le 30\right\}, \\ \text{Seek} &= \left\{|Z_{\text{in}}| < 35 \text{ or } 65 < |Z_{\text{in}}|\,,\ \phi < -30 \text{ or } 30 < \phi\right\}, \end{aligned} \tag{13.18}$$

while Ishida the conditions

$$\begin{aligned} \text{Follow} &= \left\{40 \le |Z_{\text{in}}| \le 60\,,\ -18 \le \phi \le 18\right\}, \\ \text{Seek} &= \left\{|Z_{\text{in}}| < 40 \text{ or } 60 < |Z_{\text{in}}|\,,\ \phi < -18 \text{ or } 18 < \phi\right\}. \end{aligned} \tag{13.19}$$

The Follow controller consists of

$$\begin{aligned} \phi > 0, &\quad \text{decrease } C_m, \\ \phi < 0, &\quad \text{increase } C_m, \\ |Z| > 50, &\quad \text{decrease } C_t, \\ |Z| < 50, &\quad \text{increase } C_t, \end{aligned} \tag{13.20}$$

and the Seek controller uses the inverse model of the system to compute the desirable value of the capacitors to match the impedance value.

On the other hand, Bezooijen [19] analyzed the behavior of the real and imaginary parts of the impedance with the displacement of the capacitors C_t and C_m. Then, they proposed a control rule for each impedance matching network type, resulting for the L-network in the one given by (13.21),

$$\begin{aligned} u_{C_t} &= K_t(B_{\text{ref}} - B_{\text{in}}), \\ u_{C_m} &= -K_m(G_{\text{ref}}\,\mathrm{sign}(-B_{\text{in}}) - G_{\text{in}}). \end{aligned} \tag{13.21}$$

The problem with the previous strategies is that it only analyzed the behavior of the network under a closed region of the matching value and not globally. Cottee [3] did a global analysis of the network and proved the local and global stability of the network under the geometric analysis of the phase plane (see Fig. 13.3). Following his methodology for the L-Type network, the control action is defined by (13.22)

$$\begin{aligned} u_{C_t} &= -A\,\mathrm{sign}(s_1), \\ u_{C_m} &= -B\,\mathrm{sign}(s_2), \end{aligned} \tag{13.22}$$

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

with

$$s_1 = B_{\text{in}}, \tag{13.23}$$

$$s_2 = \begin{cases} -B_{\text{in}} - \frac{1}{\beta_{\min}+\omega L_t}, & -B_{\text{in}} - \frac{1}{\beta_{\min}+\omega L_t} \leq 0, \\ 0.02 - G_{\text{in}}, & -B_{\text{in}} - \frac{1}{\beta_{\min}+\omega L_t} > 0; \end{cases} \tag{13.24}$$

and $\beta_{\min} = -\frac{1}{\omega C_{\min}}$.

13.4.2 Feedforward plus PI control

In this section, a new variation of (13.18) is presented. Based on the information of the reflection coefficient amplitude ($|\Gamma| \leq$ threshold), the system switch between the Feedforward and Feedback Control, and, using different values (threshold $=$ $0.1, 0.3, 0.5, 0.7, 0.9, 1$), the performance of the system is studied. When the threshold is equal to one, the system is under Feedback Control only. All the strategies are available at [16] by *Bacelli.xls, Hirose.slx, Ishida.slx, Bezoo.xls, Cottee.xls* and *PI.xls*. The proposed control action is defined as

$$\begin{aligned} \text{Follow} &= \{|\Gamma| \leq \text{threshold}\}, \\ \text{Seek} &= \{|\Gamma| > \text{threshold}\}, \end{aligned} \tag{13.25}$$

and the Follow Controller is

$$\begin{aligned} u_{C_t} &= Kp_{C_t} e_I + Ki_{C_t} \int_0^t e_I \, dt, \\ u_{C_m} &= -Kp_{C_m} e_R + Ki_{C_m} \int_0^t e_R \, dt, \end{aligned} \tag{13.26}$$

with $Kp_x > 0$, $Ki_x > 0$, $e_I = -B_{\text{in}}$, and $e_R = 0.02 - G_{\text{in}}$.

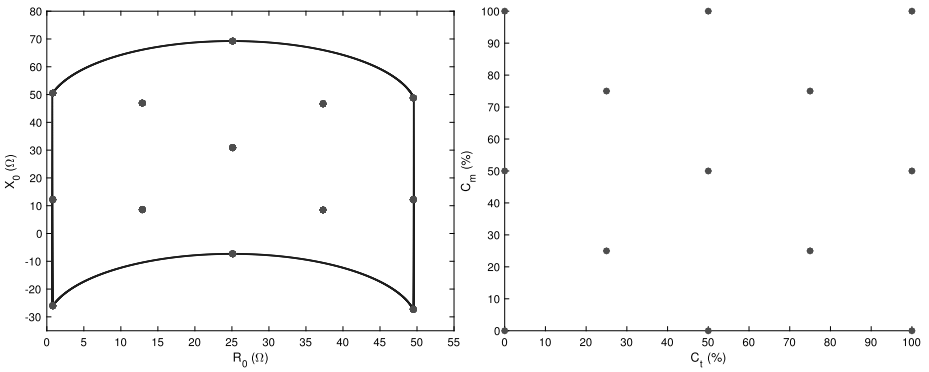
Finally, Table 13.3 shows the performance indices of the controllers evaluated for all the initial conditions described in Fig. 13.8. For the 13 extremum values of capacitance and load impedance, a total of 169 cases were obtained. For each case, the optimal values of the controllers were obtained using the cost function (13.10).

As can be seen, for the proposed controller PI + FF with threshold = 0.1, a global convergence (100 %) of precision is reached but decreases as the threshold increases. This is because the Feedforward Control does not reach the lower section of the phase plane—see Fig. 13.3—before switching to the Feedback Controller, therefore the controller decreases the value of the capacitor C_m driving the pair $(G_{\text{in}}, B_{\text{in}})$ in the wrong direction. Nevertheless, the PI + FF presents a lower convergence time thanks to the Feedforward Controller and more convergence cases than the other strategies.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 13.3: Performance of the controllers with optimal parameters.

CONTROLLERS		T_{tune}		G_{int}		mean(P_r)		max(P_r)	% Conv.
		mean	max	mean	max	mean	max	mean	
Bacelli [1]		2.89	19.99	736.73	7.38	3.01	12.75	5.63	61.54
Hirose [9]		7.51	14.99	1063	4660	3.01	14.45	8.91	71.01
Ishida [11]		7.11	14.99	1107	7175	3.16	14.10	8.78	85.21
Bezooijen [19]		1.9	4.99	507.32	2407	4.72	14.30	8.82	84.62
Cottee [3]		5.31	47.37	2,188.24	22,290.33	5.28	13.87	8.75	100
PI + FF	0.1	0.64	1.38	180.82	536.41	5.46	14.09	8.75	100
	0.3	0.70	1.92	185.69	537.81	5.23	13.81	8.80	99.41
	0.5	0.72	3.14	191.09	541.20	5.28	14.03	8.95	96.45
	0.7	0.75	3.70	196.61	657.09	5.27	14.211	9.01	94.08
	0.9	0.72	2.13	197.08	535.27	5.32	14.102	8.92	91.12
	1.0	0.73	1.98	198.53	936.62	4.77	13.634	8.29	82.24



(a) Load impedance values (b) Capacitor values

Figure 13.8: Extremum values of load impedance and capacitor.

13.5 Extremum seeking control with dithering noise

Extremum Seeking Control (ESC) is an adaptive model-free technique that allows optimizing smooth convex function problems, which extremum or optimal point may be unknown and variable on time [2, 14, 23]. There are several formulations for controllers, such as perturb and observe [2], sliding mode [6], and relay based [8]. In this chapter, the sinusoidal P&O ESC and P&O FO-SESC controllers are employed for the impedance matching problem.

13.5.1 Sinusoidal perturb and observe ESC controller

The structure of the sinusoidal perturb and observe controller is shown in Fig. 13.9 [2]. Initially, a sinusoidal perturbation signal with amplitude a and frequency ω is added

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

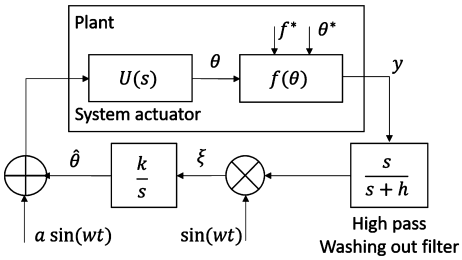


Figure 13.9: Sinusoidal P&O ESC controller.

to the estimated optimal parameter $\hat{\theta}$. Then, it is applied to the plant by the actuator $U(\theta)$, producing a new value of the cost function $f(\theta)$ (13.27), where f^* is the optimal value of f produced by the optimal parameter θ^*. Thus, the produced system output y (13.28) is passed through a high-pass filter with pole h, acting like a function gradient multiplied by the perturbation signal producing $\eta = \dot{\theta}$ that is then integrated, producing a new $\hat{\theta}$ with error dynamics $\dot{\tilde{\theta}}$ stable for $k, a, f'' > 0$ as demonstrated by [2], where k is the integrator gain and a the disturbance scaling factor. As a condition for the controller stability, f'' must exist, and f should be a smooth function.

$$f(\theta) = f^* + \frac{f''}{2}(\theta^* - \theta)^2, \tag{13.27}$$

$$y = f^* + \frac{a^2 f''}{4} + \frac{f''}{2}\tilde{\theta}^2 - af''\tilde{\theta}\sin(wt) + \frac{a^2 f''}{4}\cos(2wt), \tag{13.28}$$

$$\dot{\tilde{\theta}} = -\frac{ka^2 f''}{4}\tilde{\theta}. \tag{13.29}$$

13.5.2 Stochastic fractional-order perturb and observe controller

The P&O FO-SESC controller proposed by [10] is shown in Fig. 13.10. Its structure is similar to the P&O IO-ESC controller. The main difference is that the perturbation signal v corresponds to a bounded stochastic dithering noise that can be Integer-Order Gaussian (Brownian motion), Symmetric Alpha Stable SαS, or Fractional-order Gaussian (fGN). The disturbance signal is passed through a low-pass filter to cut off the high-frequency components of the random noise to perform better optimum seeking.

Fractional Gaussian noise can be represented as a change in Brownian motion steps, which is defined using the Riemann–Liouville fractional integral, as shown by (13.30), where $dB(s)$ is the general definition of white noise, $\Gamma(\cdot)$ is the gamma function and H is the Hurst exponent, which indicates the LRD property of the random disturbance signal [15, 17]. According to the value of H, the dithering-noise time series can represent a Brownian motion if $H = 0.5$, positively correlated if $0.5 < H < 1$ and negatively correlated if $0 < H < 0.5$. In this case, the dithering noise is considered

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

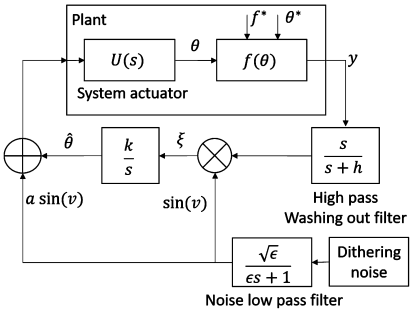


Figure 13.10: Fractional-Order P&O SESC controller.

a fractional-order Brownian motion if $H \neq 0.5$. The exponential stability of the P&O FO-SESC controller is demonstrated in [10] for $f''(\hat{\theta}) > 0$ and $ka > 0$.

$$B_H(t) = \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-0.5} dB(s). \tag{13.30}$$

For the analysis, an L-type network (see Fig. 13.1) is employed with the parameters $425\,\text{pF} \leq C_t \leq 2240\,\text{pF}$, $114\,\text{pF} \leq C_m \leq 445\,\text{pF}$, $L_t = 600\,\text{nH}$, $L_m = 689\,\text{nH}$, and an RF source frequency of 13.56 MHz. The matching space for the L-type network is shown in Fig. 13.11. As can be observed, under the current configuration, the matching network match for the load $Z_{\text{load}} = R_0 + jX_0\,\Omega$ with real and complex parts between $9 < R_0 < 22.8\,\Omega$ and $-57.2 < X_0 < 24\,\Omega$, respectively.

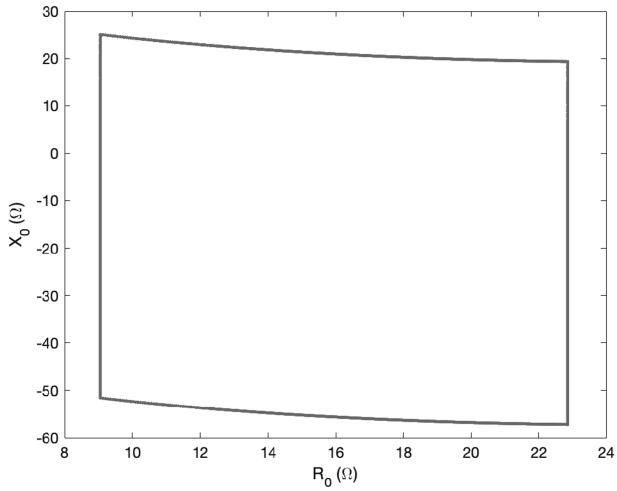


Figure 13.11: Feasible matching region for L-type network.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

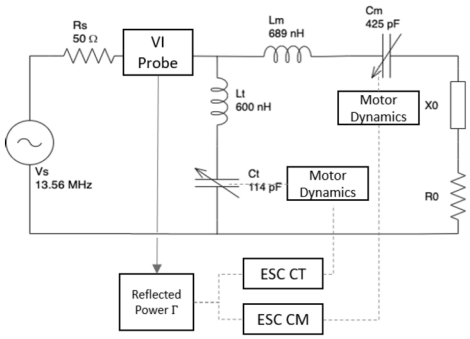


Figure 13.12: MIMO ESC Controller with L-type matching network.

The ESC control architecture is shown in Fig. 13.12. It employs two independent ESC controllers, one for the shunt and series capacitors C_t and C_m, respectively. The motor response should be included in the system model, considering that the electrical dynamic response of the system is much faster compared with the motor dynamics of the variable capacitors.

The cost function employed for the P&O ESC and P&O FO-SESC controllers is the reflected power of the impedance matching network defined by (13.1). In a real L-type network, the input impedance Z_{in} is measured by a Voltage-Current (VI) probe. Here, Z_{in} is calculated using (13.6) according to the values of C_t and C_m. The tuning parameters employed for the P&O ESC and P&O FO-SESC controllers are shown in Table 13.4. In the case of P&O ESC controller, the parameters are selected to ensure the correct extremum tracking. In addition, the P&O FO-SESC controller parameters include a zero mean and a small standard deviation dithering noise for C_t and C_m ES controllers. Larger correction gains and perturbation signal amplitudes are selected for C_m due to its capacitor range. The dithering noise low-pass filter noise is set to 1, and the Hurst exponent value range is 0.1 to 0.9.

Table 13.4: P&O ESC and P&O FO-SESC controller parameters.

Parameter	P&O ESC Controller		FO-SESC Controller	
	C_m	C_t	C_m	C_t
High-pass filter pole (h)	2π	2π	2π	2π
Integrator gain (k)	100	3,000	300	100
Amplitude correction (a)	1.5	2	0.1	0.1
Sinusoidal disturbance frequency (rad/s)	10	7	–	–
Noise mean (μ)	–	–	0	0
Noise Std (σ)	–	–	0.2	0.3
Noise-signal amplitude	1	1	150	70
Low-pass noise-filter pole (*epsilon*)	–	–	1	1
Hurst exponent (H)	0.1:0.1:0.9			

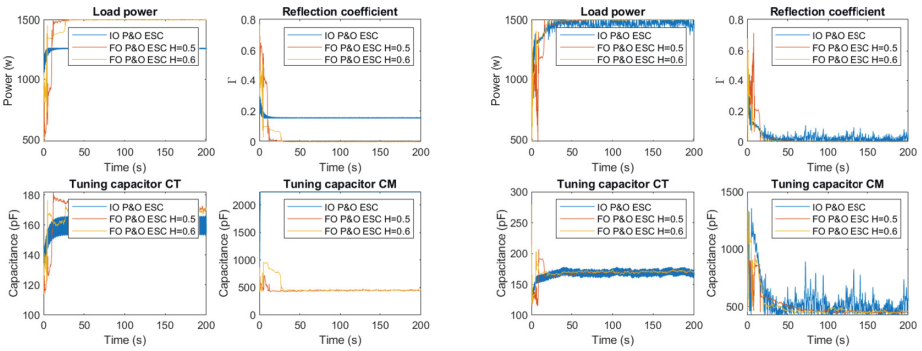
EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

13.6 Simulation study: L-type network

Three tests are performed to evaluate the P&O-ESC and P&O FO-SESC controllers. In the first test, both controllers perform the impedance matching for a power set-point of 1,500 W with a fixed load of $10-j10\,\Omega$ under multiple initial conditions of C_t and C_m and Hurst exponents in the range $H = 0.1{:}0.1{:}0.9$ to analyze the influence of anti-persistent and LRD behaviors on the SESC searching. The second test evaluates the P&O FO-SESC performance for the total load matching space given by Fig. 13.11 with $H = 0.1{:}0.1{:}0.9$ under a different capacitor IC. Finally, the third test evaluates the robustness of the P&O FO-SESC and P&O ESC controllers, analyzing its convergence for all capacitor initial conditions with $Z_{\text{load}} = 7.95 + j5.2\,\Omega$ and the loads in the matching space given by Fig. 13.11.

13.6.1 Test 1: P&O ESC and FO-SESC controllers time responses for a single given load impedance

The time responses of the P&O ESC and the P&O FO-SESC controllers are shown in Fig. 13.13a for the load $10 - j10\,\Omega$ with capacitor initial conditions at IC $= [0, 0]$. It can be observed that the P&O FO-SESC controller reaches the steady-state response, with a smooth control effort given by C_t and C_m for either $H = 0.5$ and $H = 0.6$. On the other hand, the sinusoidal P&O ESC controller is not able to converge, reaching a minimum reflected power coefficient of 0.2. Considering the influence of the capacitor's initial conditions on the search, the test is repeated now for IC $= [50, 50]$, and its results are shown in Fig. 13.13b. It can be appreciated that, under the new initial conditions, the sinusoidal P&O ESC controller now is able to converge, as well as the P&O FO-SESC controller. The control action for the P&O ESC controller exhibits a more oscillatory

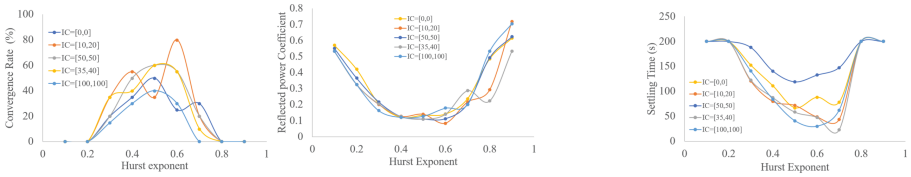


(a) IC=(0,0) (b) IC=(50,50)

Figure 13.13: P&O ESC and P&O FO-SESC controller with load $10 - 10j\,\Omega$.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

behavior compared with the P&O FO-SESC controller. Figure 13.14 shows the evolution of the average Gamma RMS, convergence events, and the settling time for the FO-SESC controller. As can be observed in Fig. 13.14b, for $H = 0.6$, the reflected power RMS is lower for most of the ICs. Likewise, Figs. 13.14a and 13.14c show that the convergence rate and settling time can reach reasonable values when $0.5 < H < 0.6$. Thus, we can say that, using dithering noises between $0.5 \leq H \leq 0.6$, the convergence rate of the ESC controller can be improved.



(a) Convergence rate (b) Reflected power coefficient (c) Settling time

Figure 13.14: Test 1: P&O FO-SESC controller average vs the Hurst exponent.

13.6.2 Test 2: P&O ESC and FO-SESC controllers global convergence analysis

This test analyzes the convergence rate, time and reflected power RMS among the matching region of the L-type network presented in Fig. 13.11. For the ESC controllers' assessment, an extended set of capacitor initial conditions are defined, which corresponds to IC $= [0, 0], [10, 20], [35, 40], [50, 50], [100, 100]$, and each load is tested for various Hurst exponents $H = 0.1{:}0.1{:}0.9$. As an example, Fig. 13.15 shows the comparison of the P&O ESC controller matching zone, regarding the evolution of the P&O FO-SESC controller under the initial condition IC $= [0, 0]$ for $H = 0.1{:}0.1{:}0.9$. As can be observed, the P&O ESC controller convergence region contains 151 positive events. Moreover, the P&O FO-SESC controller exhibits a variation of its convergence region according to the dithering-noise Hurst exponent. If $H < 0.3$ or $H > 0.8$, the convergence region is significantly smaller compared with the P&O ESC controller. However, for $0.5 \leq H \leq 0.7$, the convergence region shows an increasing trend, reaching its maximum at $H = 0.7$ with 171 positive convergence events.

Test 2 is repeated ten times to analyze the average behavior of the P&O FO-SESC controller. The average convergence rate, settling time and reflection power coefficient for the ten tests against its Hurst exponent is shown in Fig. 13.16. As can be observed, the convergence rate in Fig. 13.16a of the P&O FO-SESC controller increases to above 60 % when $0.5 <= H < 0.6$, with a minimum convergence rate at $H < 0.1$. On the other hand, the reflected power RMS and settling time in Fig. 13.16b and 13.16c shows a decreasing behavior as the Hurst exponent increases; however, the convergence rate

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

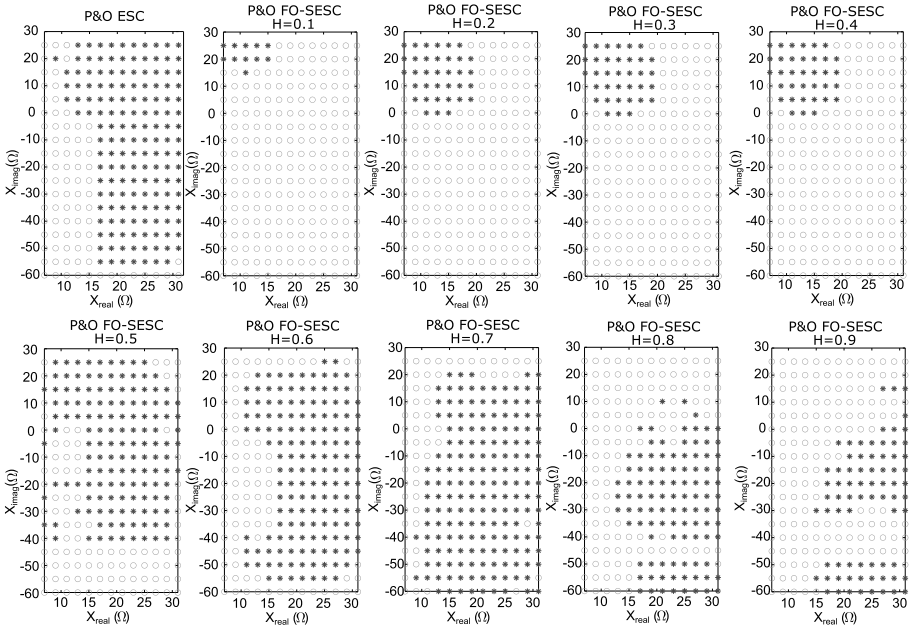
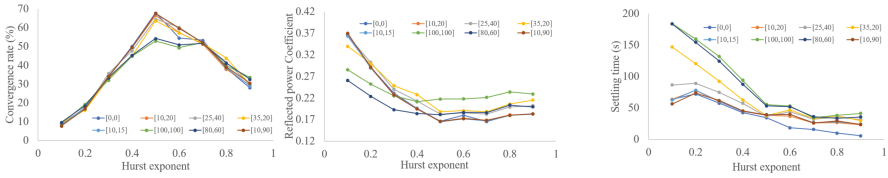


Figure 13.15: Test 2: P&O FO-SESC controller convergence region evolution against the Hurst exponent (blue is success; red is unsuccessful).



(a) Average convergence rate (b) Reflected power RMS (c) Settling time

Figure 13.16: Indexes vs Hurst exponent on Test 2 for the overall matching region.

for $H > 0.6$ suffers a significant decrease, introducing a trade-off between the convergence rate, settling time and reflected power. Thus, we can say that the best operation range for the P&O FO-SESC controller can be reached for $0.5 \leq H \leq 0.6$ in the presence of various loads.

13.6.3 Test 3: ESC controllers capacitor initial condition robustness analysis

In this test, the ESC controller convergence is evaluated in terms of the capacitor's initial condition space. Thus, a set of 25 loads inside the matching space are evaluated

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

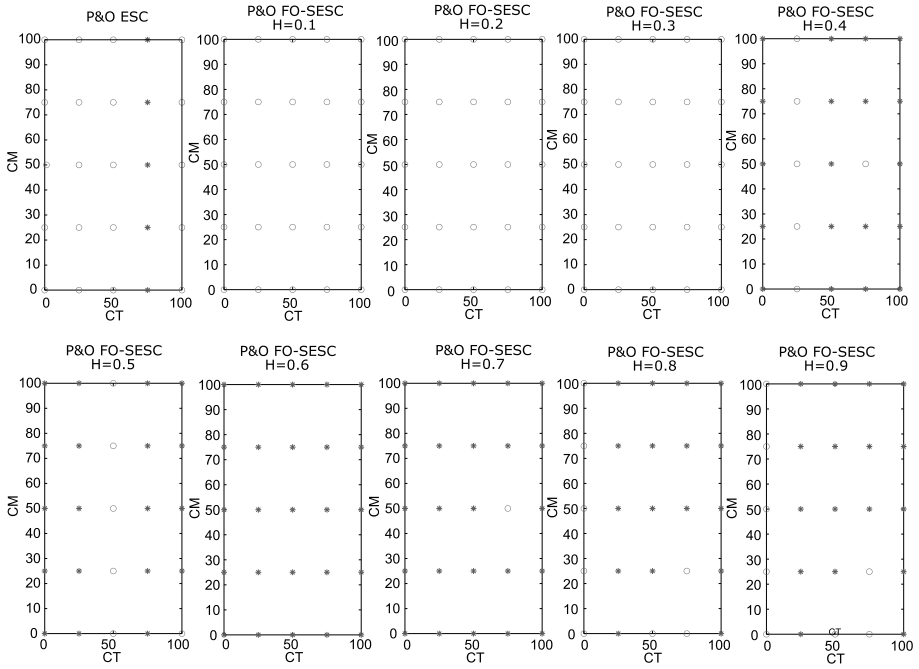


Figure 13.17: Test 3: P&O FO-SESC controller capacitor initial conditions space vs the Hurst exponent for $Z_l = 15 - j25\,\Omega$ (blue is success; red is unsuccessful).

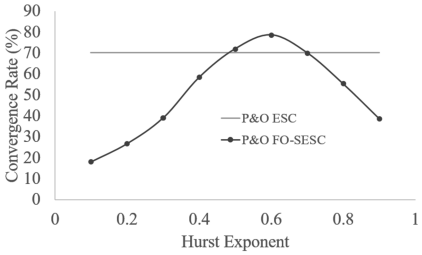


Figure 13.18: Test 3: P&O ESC and P&O FO-SESC controller average settling time vs the Hurst exponent.

with 25 ICs in the range C_t = 0:25:100 and C_m = 0:25:100. Figure 13.17 shows the initial condition capacitor space evolution using the P&O ESC and the P&O FO-SESC controllers for the load $Z_l = 15 - j25\Omega$. As can be seen, increasing the Hurst exponent can improve the convergence rate of the impedance matching. After repeating this test ten times, the global average convergence rate of the P&O FO-SESC controller is shown in Fig. 13.18. As can be observed, the convergence rate can be increased up to 80 % with $0.5 \leq H \leq 0.6$. Likewise, this behavior is consistent with the obtained results from Tests 1 and 2. Therefore, we can say that, using fractional-order dithering noise, it is possible to improve the global performance of plasma impedance matching without

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

feedforward or additional compensation mechanisms in the presence of a variable load inside the matching and capacitor spaces.

13.7 Conclusions and future works

This chapter presented a benchmark system for RF impedance matching controller design and evaluation. A comparison between different impedance matching controller architectures is performed. In the case of the novel Feedforward controller, the selected profiles improves the convergence of the reflection coefficient with a monotonic decreasing behavior. Likewise, the performance of the PI + FF controller was contrasted with the other strategies. Using a switching $|\Gamma| \leq 0.1$ the controller shows a better response; however, it does not achieve global convergence as the Cottee proposal without the Feedforward Control. The advantage of the Feedforward is to speed the convergence of the system, but it is sensitive to distortion in the sensor.

Likewise, a Perturb and Observe Fractional-Order Stochastic Extremum Seeking Control (P&O FO-SESC) with dithering noise was designed for plasma impedance matching. Using the benchmark, the P&O FO-SESC controller was tested for single and multiple loads under various IC and various levels of LRD with the dithering noise given by the Hurst exponent. Obtained results show that using a dithering noise with the P&O FO-SESC controller in the range $0.5 \leq H \leq 0.6$ can improve the system convergence rate, settling time, and minimize the total reflected power during the impedance matching process. The P&O FO-SESC controller is robust under various load and capacitor initial conditions, indicating that the P&O FO-SESC controller is suitable for plasma impedance matching, where the plasma load impedance is time variant and unknown.

Thus, this benchmark marks a starting point for the analysis of L-Type RF impedance matching networks. As future work, more control strategies such as Model Predictive Control (MPC), Model-Free Adaptive Control (MFAC), P&O FO-SESC, etc. can be tested to enhance the system performance and evaluate its performance in the presence of external disturbances and uncertainties, using the admittance model of the matching network. Likewise, introducing feedforward mechanisms based on magnitude and phase to improve ESC control performance can be considered to accelerate the impedance matching speed.

Bibliography

[1] G. Bacelli, J. V. Ringwood, and P. Iordanov. Impedance controller for an inductively coupled plasma chamber—L-type matching network automatic controller. In *Proceedings of the Fourth International Conference on Informatics in Control, Automation and Robotics*, volume SPSMC, pages 202–207. SciTePress—Science and Technology Publications, 2007.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

[2] K. Bariyur and M. Krstic. *Real-Time Optimization by Extremum-Seeking Control*. Wiley-Interscience, 2003.
[3] C. J. Cottee and S. R. Duncan. Design of matching circuit controllers for radio-frequency heating. *IEEE Transactions on Control Systems Technology*, 11(1):91–100, 2003.
[4] D. J. Coumou. Feedforward power distortion correction in RF power delivery systems for plasma processing systems. In *2012 18th IEEE-NPSS Real Time Conference*, pages 1–5. IEEE, Jun 2012.
[5] D. J. Coumou. Coherent feedforward impedance correction and feedback power regulation in a plasma processing RF power delivery system. In *2013 IEEE 14th Workshop on Control and Modeling for Power Electronics (COMPEL)*, pages 1–4. IEEE, Jun 2013.
[6] P. Dix, B. Ashrafi, S. Drakunov, and U. Özguner. ABS control using optimum search via sliding modes. *IEEE Transactions on Control Systems Technology*, 3(1):79–85, 1995.
[7] E. L. Firrao, A.-J. Annema, and B. Nauta. An automatic antenna tuning system using only RF signal amplitudes. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 55(9):833–837, Sep 2008.
[8] G. A. Frantsuzova. Use of a relay controller for automatic extremum seeking in nonlinear systems. *Avtometriya*, 47(3):274–280, 2011.
[9] Y. Hirose, A. Kawamura, A. Takayanagi, and H. Takada. Analysis of impedance matching control. In *2009 IEEE 6th International Power Electronics and Motion Control Conference*, volume 3, pages 1188–1191. IEEE, May 2009.
[10] D. Hollenbeck and Y. Chen. A more optimal stochastic extremum seeking control using fractional dithering for a class of smooth convex functions. In *2020 IFAC World Congress*, pages 1–6, 2020.
[11] S. Ishida, A. Kawamura, A. Takayanagi, and H. Takada. Robust convergence of the impedance matching process based on seek control. In *IECON 2011—37th Annual Conference of the IEEE Industrial Electronics Society*, pages 1554–1559. IEEE, Nov 2011.
[12] Y. F. Li. Analysis and design of tracking control for the RF matching box of plasma system. In *ACM International Conference Proceeding Series*, pages 18–23, 2019.
[13] S.-J. Liu and M. Krstic. Stochastic averaging in continuous time and its applications to extremum seeking. *IEEE Transactions on Automatic Control*, 55(10):2235–2250, 2010.
[14] S.-J. Liu and M. Krstic. *Stochastic Averaging and Stochastic Extremum Seeking*, volume 53. Springer, 2013.
[15] R. L. Magin. Fractional calculus in bioengineering. *Critical Reviews in Biomedical Engineering*, 32(1):1–104, 2004.
[16] C. Rodriguez. RF impedance matching control benchmark. URL: https://is.gd/5aB0kz, 2021.
[17] H. Sheng, Y. Q. Chen, and T. S. Qiu. *Fractional Processes and Fractional-Order Signal Processing Techniques and Applications*. Springer, 2012.
[18] M. Thompson and J. K. Fidler. Determination of the impedance matching domain of impedance matching networks. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 51(10):2098–2106, Oct 2004.
[19] A. van Bezooijen, M. A. de Jongh, F. van Straten, R. Mahmoudi, and A. van Roermund. Adaptive impedance-matching techniques for controlling L networks. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 57(2):495–505, Feb 2010.
[20] B. Xiong and K. Hofmann. Binary search algorithm for adaptive impedance matching network. *Electronics Letters*, 52(9):714–716, 2016.
[21] B. Xiong, L. Yang, and T. Cao. A novel tuning method for impedance matching network based on linear fractional transformation. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 67(6):1039–1043, Jun 2020.
[22] C. Yin, Y. Chen, and S. M. Zhong. Fractional-order sliding mode based extremum seeking control of a class of nonlinear systems. *Automatica*, 50(12):3173–3181, 2014.
[23] C. Zhang and R. Ordóñez. *Extremum-Seeking Control and Applications*, 1st edition. Springer-Verlag, London, 2012.

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Index

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

Fractional Calculus in Applied Sciences and Engineering

Volume 9
Ming Li
Theory of Fractional Engineering Vibrations, 2021
ISBN 978-3-11-072652-7, e-ISBN (PDF) 978-3-11-072615-2,
e-ISBN (EPUB) 978-3-11-072616-9

Volume 8
Vasily E. Tarasov, Valentina V. Tarasov
Economic Dynamics with Memory, 2021
ISBN 978-3-11-062460-1, e-ISBN (PDF) 978-3-11-062460-1,
e-ISBN (EPUB) 978-3-11-062481-6

Volume 7
Manuel Duarte Ortigueira, Duarte Valério
Fractional Signals and Systems, 2020
ISBN 978-3-11-062129-7, e-ISBN (PDF) 978-3-11-062458-8,
e-ISBN (EPUB) 978-3-11-062132-7

Volume 6
Yingjie Liang, Wen Chen, Wei Cai
Hausdorff Calculus. Applications to Fractal Systems, 2018
ISBN 978-3-11-060692-8, e-ISBN (PDF) 978-3-11-060852-6,
e-ISBN (EPUB) 978-3-11-060705-5

Volume 5
JinRong Wang, Michal Fečkan
Fractional Hermite-Hadamard Inequalities, 2018
ISBN 978-3-11-052220-4, e-ISBN (PDF) 978-3-11-052362-1,
e-ISBN (EPUB) 978-3-11-052244-0

Volume 4
Kecai Cao, YangQuan Chen
Fractional Order Crowd Dynamics. Cyber-Human System Modeling and Control, 2018
ISBN 978-3-11-047281-3, e-ISBN (PDF) 978-3-11-047398-8,
e-ISBN (EPUB) 978-3-11-047283-7

www.degruyter.com

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/14/2023 6:11 AM via . All use subject to https://www.ebsco.com/terms-of-use