Department of Mathematics and Systems Analysis

New approaches for modeling and estimation of discrete and continuous time stationary processes

Marko Voutilainen

Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 209/2020

New approaches for modeling and estimation of discrete and continuous time stationary processes

Marko Voutilainen

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Science, https://aalto.zoom.us/j/61650076687, on 18 December 2020 at 12.

Aalto University
School of Science
Department of Mathematics and Systems Analysis
Stochastics and Statistics

Supervising professor
Assistant Professor Pauliina Ilmonen, Aalto University School of Science, Department of Mathematics and Systems Analysis

Thesis advisor
Lecturer Lauri Viitasaari, Aalto University School of Business, Department of Information and Service Management

Preliminary examiners
Professor Yuliya Mishura, Taras Shevchenko National University of Kyiv, Ukraine
Professor Paavo Salminen, Åbo Akademi, Finland

Opponent
Professor Paavo Salminen, Åbo Akademi, Finland

Aalto University publication series
DOCTORAL DISSERTATIONS 209/2020

ISBN 978-952-64-0182-9 (printed)
ISBN 978-952-64-0183-6 (pdf)
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-64-0183-6

Unigrafia Oy
Helsinki 2020

Finland

Printed matter
4041-0619

Aalto University

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Marko Voutilainen

Name of the doctoral dissertation
New approaches for modeling and estimation of discrete and continuous time stationary processes

Publisher School of Science

Unit Department of Mathematics and Systems Analysis

Series Aalto University publication series DOCTORAL DISSERTATIONS 209/2020

Field of research Mathematics

Manuscript submitted 23 September 2020 **Date of the defence** 18 December 2020

Permission for public defence granted (date) 11 November 2020 **Language** English

☐ **Monograph** ☒ **Article dissertation** ☐ **Essay dissertation**

Abstract

Stationary processes form an important class of stochastic processes that has been extensively studied in the literature, and widely applied in many fields of science. Applications include modeling and forecasting various real-life phenomena such as stock market behavior, sales of a company, natural disasters and velocity of a Brownian particle under the influence of friction, to mention a few.

In this dissertation, we consider new methods for modeling and estimation of discrete and continuous time stationary processes. We characterize discrete and continuous time strictly stationary processes by AR(1) and Langevin equations, respectively. From these characterizations, we derive quadratic (matrix) equations for the corresponding model parameter (matrix) in terms of autocovariance of the stationary process. Based on the equations, we construct an estimator for the model parameter. Furthermore, we show that the estimator inherits consistency and the rate of convergence from the chosen autocovariance estimators. Moreover, its limiting distribution is given by a linear function of the limiting distribution of the autocovariance estimators. In addition, we apply the presented general theory in modeling and estimation of a generalization of the ARCH model with stationary liquidity.

Keywords

ISBN (printed) 978-952-64-0182-9 **ISBN (pdf)** 978-952-64-0183-6

ISSN (printed) 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2020

Pages 180 **urn** http://urn.fi/URN:ISBN:978-952-64-0183-6

Aalto-yliopisto

Aalto-yliopisto, PL 11000, 00076 Aalto www.aalto.fi

Tiivistelmä

Tekijä
Marko Voutilainen

Väitöskirjan nimi
Uusia menetelmiä diskreetti- ja jatkuva-aikaisten stationaaristen prosessien mallintamiseksi ja estimoimiseksi

Julkaisija Perustieteiden korkeakoulu

Yksikkö Matematiikan ja systeemianalyysin laitos

Sarja Aalto University publication series DOCTORAL DISSERTATIONS 209/2020

Tutkimusala Matematiikka

Käsikirjoituksen pvm 23.09.2020 **Väitöspäivä** 18.12.2020

Väittelyluvan myöntämispäivä 11.11.2020 **Kieli** Englanti

☐ **Monografia** ☒ **Artikkeliväitöskirja** ☐ **Esseeväitöskirja**

Tiivistelmä

Stationaariset prosessit muodostavat merkittävän stokastisten prosessien luokan, jota on tutkittu laajalti ja jolle löytyy sovelluksia monilta tieteen eri osa-alueilta. Sovelluskohteita ovat esimerkiksi monien reaalimaailman ilmiöiden mallintaminen ja niiden ennustaminen, kuten pörssikurssit, yrityksen liikevaihto, luonnonkatastrofit ja liikevastuksen vaikutuksen alaisen Brownin hiukkasen nopeus.

Tässä väitöskirjassa esitellään uusia menetelmiä diskreetti- ja jatkuva-aikaisten stationaaristen prosessien mallintamiseksi ja estimoimiseksi. Diskreetti- ja jatkuva-aikaiset vahvasti stationaariset prosessit karakterisoidaan AR(1) ja Langevin yhtälöiden avulla. Kyseisten karakterisaatioiden pohjalta johdetaan stationaarisen prosessin autokovarianssin avulla ilmaistavat toisen asteen (matriisi)yhtälöt mallin (matriisi)parametrille. Perustuen näihin yhtälöihin, mallin parametrille määritellään estimaattori. Estimaattorin tarkentuvuuden ja suppenemisnopeuden osoitetaan seuraavan suoraan valittujen autokovarianssiestimaattoreiden vastaavista ominaisuuksista. Tämän lisäksi estimaattorin rajajakauma voidaan esittää lineaarisen funktion avulla autokovarianssiestimaattoreiden rajajakaumasta. Esitettyä yleistä teoriaa sovelletaan myös ARCH-mallin erään yleistyksen estimoimiseksi.

Avainsanat

ISBN (painettu) 978-952-64-0182-9 **ISBN (pdf)** 978-952-64-0183-6

ISSN (painettu) 1799-4934 **ISSN (pdf)** 1799-4942

Julkaisupaikka Helsinki **Painopaikka** Helsinki **Vuosi** 2020

Sivumäärä 180 **urn** http://urn.fi/URN:ISBN:978-952-64-0183-6

Preface

First of all, I want to express my gratitude to my supervisor Prof. Pauliina Ilmonen and advisor Adj. Prof. Lauri Viitasaari. Their dedication and enthusiasm towards mathematics, and life in general, have been tremendously inspiring. I could not have wished for better mentors for this project.

I would like to thank the preliminary examiners Prof. Paavo Salminen and Prof. Yuliya Mishura for reviewing this thesis and for their valuable feedback. In addition, I wish to thank Prof. Paavo Salminen for agreeing to act as my opponent.

I am grateful that I have had the opportunity to collaborate with such established academics as Prof. Soledad Torres and Prof. Ciprian Tudor during this project. It was a valuable experience to participate in the Conference of the Romanian Society of Probability and Statistics and visit the University of Valparaíso. I acknowledge the financial support from the Magnus Ehrnrooth foundation that enabled the research visit to Chile.

I feel proud that I have been a part of the Department of Mathematics and Systems Analysis, Aalto University School of Science these past years. In particular, I would like to thank the Stochastic and Statistics group for creating a professional but laid-back working atmosphere. My special thanks go to Niko Lietzén for giving me the possibility to observe the defence mechanisms of the human mind when confronting the stress of constant defeat in heated table tennis duels.

Last but not least, I would like to thank my close ones for their unconditional love and support. My friends have been an essential asset whenever I have needed a break from sometimes overly intensive thought processes. Thank you Katja for bringing joy back into my life, and also for assisting me in English related issues when I was polishing this thesis. Above all I want to thank my parents. Without your patience and support I would not be here today.

Espoo, November 20, 2020,

Marko Voutilainen

Contents

List of publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I M. Voutilainen, L. Viitasaari, and P. Ilmonen. On model fitting and estimation of strictly stationary processes. *Modern Stochastics: Theory and Applications*, 4(4), 381–406, December 2017.

II M. Voutilainen, L. Viitasaari, and P. Ilmonen. Note on AR(1)-characterisation of stationary processes and model fitting. *Modern Stochastics: Theory and Applications*, 6(2), 195–207, March 2019.

III M. Voutilainen, P. Ilmonen, S. Torres, C. Tudor, L. Viitasaari. On the ARCH model with stationary liquidity. *Metrika*, DOI: 10.1007/s00184-020-00779-x, June 2020.

IV M. Voutilainen, L. Viitasaari, P. Ilmonen, S. Torres, C. Tudor. Vector-valued generalised Ornstein-Uhlenbeck processes. *Submitted to a journal*, arXiv: 1909.02376v2, November 2020.

V M. Voutilainen. Modeling and estimation of multivariate discrete and continuous time stationary processes. *Frontiers in Applied Mathematics and Statistics*, DOI: 10.3389/fams.2020.00043, September 2020.

Author's contribution

Publication I: "On model fitting and estimation of strictly stationary processes"

The idea of the characterization of stationary processes is adapted by the author from [73] by Adj. Prof. Viitasaari. The consequent estimation method is the author's idea. All the theoretical results are derived by the author with the guidance of Adj. Prof. Viitasaari. The numerical experiments were designed as a joint effort and implemented by the author. The author wrote the first version of the manuscript. All authors contributed to finalizing the results and to writing the final version of the article.

Publication II: "Note on AR(1)-characterisation of stationary processes and model fitting"

The idea for the paper came from the author and Adj. Prof. Viitasaari. All the theoretical results are derived by the author with the guidance of Adj. Prof. Viitasaari. The simulations are designed and implemented by the author. The graphical illustrations are created by the author. The paper is mainly written by the author. All authors contributed to finalizing the results and to writing the final version of the article.

Publication III: "On the ARCH model with stationary liquidity"

The idea for the paper came from Prof. Torres and a few ideas for theoretical results are adapted from a draft of [5] by Bahamonde, Torres and Tudor. All the theoretical results are derived by the author with the guidance of Adj. Prof. Viitasaari. The simulations are designed by the author, Adj. Prof. Viitasaari and Prof. Torres, and they are implemented by the author. The computations and the writing of Subsection 3.3. (Examples) have been carried out by Prof. Tudor. The introduction is largely written by Prof. Torres. Otherwise, the paper is mainly written by the author. All authors contributed to finalizing the results and to writing the final version of the article.

Publication IV: "Vector-valued generalised Ornstein-Uhlenbeck processes"

The idea for the paper came from the author and Adj. Prof. Viitasaari. The theoretical results are derived by the author with some guidance from Adj. Prof. Viitasaari. The proofs related to Subsection 2.2. (Application to Gaussian processes) are the result of a joint effort by the author and Adj. Prof. Viitasaari. The introduction and Subsection 2.2. are written by Adj. Prof. Viitasaari. Otherwise, the paper is mainly written by the author. All authors contributed to finalizing the results and to writing the final version of the article.

Publication V: "Modeling and estimation of multivariate discrete and continuous time stationary processes"

The article is a personal work by the author.

1. Introduction

Stationary processes, a notion defined by Aleksander Khintchine in his paper [42] in 1934, form undoubtedly one of the most important and widely studied classes of stochastic processes. Some notable results from the early stages of development of the field include spectral representation theorems of stationary processes and their autocovariance functions, factorization of spectra, Wold's decomposition, linear least squares forecasting and filtering, and ergodic theorems. Details on these topics together with historical and bibliographical remarks can be found in [28]. For a brief account of scientific history of time series analysis and stationary processes, we refer to [19].

Probably the simplest example of a non-degenerate stationary process is a sequence of independent and identically distributed random variables. These type of time series are encountered e.g. by flipping a coin or tossing a dice consecutively. As a more non-artificial example, one could consider stock market returns to be realizations of stationary processes possessing memory. Although most real-life time series are not stationary, many of them can be stationarized via a suitable transformation of the original data. Due to the capitalist nature of the modern world based on constant growth, various economy related time series exhibit a linear upward trend. In addition, many environmental time series are of seasonal nature. One classic example on stationarizable time series is given by carbon dioxide concentration measurements at Mauna Loa volcano. The data indicates not only a linearly rising trend, but also seasonal behavior. After eliminating these expressions of non-stationarity by differencing the original data, the transformed time series can be adequately modeled by applying stationary processes [33].

The foregoing discussion connects stationary stochastic processes closely to temporal phenomena. It is worth to mention that stochastic processes and stationarity can be easily defined using more general parameter sets without time interpretations.

However, in this dissertation, we concentrate on stationary processes indexed by the two most commonly used parameter sets, namely the integers and the real numbers. Throughout this dissertation, we refer to these cases as ”discrete time” and ”continuous time”, respectively.

Discrete time

In addition to temporal phenomena that are discrete at their core, typically measurements of continuous time variables also result in discrete time series. Conventionally, the statistical modeling of discrete time data has been done by applying models of the vast ARMA (autoregressive moving-average) family constituting the foundation of the modern time series analysis. The origin of the ARMA processes dates back to studies by Udny Yule and Eugene Slutsky from 1920s. Yule introduced pure autoregressive processes in his paper [78] from year 1927, and the ideas were extended a few years later by Gilbert Walker [74]. Moving-average processes were constructed by Slutsky in [66], a paper which was originally written in Russian also in 1927. We also mention Herman Wold, who established ARMA models in his work [77] renowned for the decomposition theorem named after him. As an important consequence of the theorem, any purely non-deterministic weakly stationary process can be approximated by ARMA processes highlighting their generality. A reader interested in the early scientific history of the ARMA processes may turn to [57]. To this date, ARMA processes have given rise to an exhaustive number of models of stationary processes. For example, various ARCH (introduced by Engle [24]) and GARCH (generalized autoregressive conditional heteroskedasticity models introduced by Bollerslev [12]) have been widely applied to describe fluctuations of volatility in financial data. Different members of the ARMA family together with estimation in these models have been considered e.g. in [46], [26], [30], [35], [45] and [6], to mention a few. For a glossary on ARCH related models, see [13].

When applying ARMA models, there are a few issues that have to be overcome by the practitioner. The first question to be addressed is the stationarity of the given data together with the model selection problem. After fixing the used model, the practitioner has to select a method for estimation of the related model parameters. The final step is to apply model validation methods in order to evaluate the quality of the estimated model. The foregoing discussion summarizes the three stages of the Box-Jenkins method introduced in [14] giving explicit instructions for the modeling procedure, and consequently causing a boom of applications of ARMA processes. Next, we give a brief overview on these stages of the Box-Jenkins method.

The approach to model identification taken in [14] relied heavily on the study of the sample autocorrelation and partial autocorrelation functions that can lead to equivocal interpretations especially in the case of more complicated ARMA processes. Thenceforth, an array of more sophisticated methods have been proposed to be based, for example, on canonical correlation analysis (see e.g. [2] and [71]). We also mention different information criteria, such as AIC by Akaike [1] and BIC by Schwartz [65], that are widely applied today in practice. However, there does not exist a single model selection strategy superior to others in every given situation.

The estimation of the parameters of the chosen model is conventionally carried out by applying different maximum likelihood methods such as optimizing exact, conditional or quasi likelihood functions. However, since there exists no closed form exact maximum likelihood estimators even for the simplest ARMA models, the related maximization problem has to be solved by applying some numerical method. Also, the computation of the maximum likelihood function can already be involving for a general ARMA process. In addition, the maximum likelihood method requires knowledge of the exact distribution of the noise process, although Gaussian likelihood approach may yield asymptotically consistent and asymptotically normal estimators also for a non-Gaussian noise. Furthermore, also different least squares and moment based methods have been applied in the estimation of ARMA processes. Contrary to the maximum likelihood method, the ordinary least squares estimators of a pure autoregressive process admit closed form representations. However, the least squares method is restricted to ARMA processes that are causal and invertible, since otherwise the related minimization problem may not have a finite solution. For details on estimation of ARMA processes, and on modern time series analysis in general, we refer to [16] and [29]. We also mention the recognized paper [31] by Hannan on asymptotic properties of estimators of linear processes.

There exists an abundance of diagnostics for assessing the goodness of the estimated model that are typically designed to recognize whether the estimated residuals support the underlying assumptions of the noise process. That is, the residuals should embody a realization of a white noise or an IID process, depending on whether strict or weak stationary is under consideration. Perhaps the most well-known tool for testing zero-autocorrelation of the residual series and adequacy of the model is the Ljung-Box test [47], a modification of the test proposed by Box and Pierce in [15]. If the applied tests suggest that the estimated model is not satisfactory, then the practitioner is prompted to begin a new iteration of the Box-Jenkins method.

Continuous time

Time series of continuous time arise naturally in many physical applications, and even if the observations are made discretely, it is justified to apply continuous time models to the underlying phenomena. In addition, it might also be convenient to apply continuous time models to irregularly spaced time series of discrete processes (see e.g. [39]). The time evolution of a continuous time variable is often modeled by stochastic differential equations (SDEs) that can be regarded as continuous time analogs of stochastic difference equations such as ARMA equations. Moreover, in many cases, introducing a suitable initial condition yields a stationary solution to a SDE. However, the transition from discrete to continuous time is not straightforward. For example, it is not obvious what should constitute the continuous time counterpart of the discrete time white noise process. Attempts to preserve discrete time white noise properties face several difficulties leading to a mathematically improper process with autocovariance function given by the Dirac delta function. Nevertheless, continuous time (Gaussian) white noise may be regarded as the formal derivative of Brownian motion (for details, see e.g. [60] and [50]), which is the famous Gaussian process with continuous paths and independent stationary increments.

The mathematical history of Brownian motion and SDEs date back to studies of Thorvald Thiele [70], Louis Bachelier [4], Marian Smoluchowski [67] and Albert Einstein [22] in the turn of the 19th and 20th centuries. From the aforementioned, the paper by Einstein became the most celebrated and influential, and in many occasions, it is regarded as the initiator of stochastic calculus. In his paper, Einstein derived a Fokker-Planck differential equation for the probability density function of displacement of a Brownian particle immersed in liquid noting that the mean-square of the displacement grows linearly in time. Inspired by the work of Einstein, Paul Langevin took apparently a more simple approach to the same problem based on the classical laws of physics [44]. His model for a Brownian particle was presumably the first differential equation involving a stochastic driving force (equivalent to continuous time white noise process), but a rigorous mathematical treatment of the equation had to wait for two major advancements.

The first of these advancements is largely due to Wiener, who gave a mathematical construction of Brownian motion [75, 76] and hence, honoring his achievements, the mathematical formulation of Brownian motion is known as Wiener process. One of the many key results of Wiener was that the sample paths of Brownian motion are almost surely nowhere differentiable. Consequently, stochastic differential equations involving Brownian motion are usually interpreted as formal representations of integral equations. However, since Brownian motion has almost surely infinite variation on compact sets, the usual integration machinery, such as Riemann-Stieltjes, is not available creating a demand for stochastic integration theory. Although Wiener contributed significantly also to the theory of integration of deterministic functions with respect to Brownian motion, the problem was thoroughly addressed by Kiyoshi Itô by developing the modern stochastic calculus initiated in [36]. Nowadays, stochastic differential equations are typically interpreted as Itô integral equations. However, other approaches, such as Stratonovich and pathwise interpretations, exist as well. For details on the topics discussed in this paragraph, we refer to [51], [37], [55] and [61]. On the scientific history of Brownian motion and stochastic calculus, see e.g. [38], [27] and [41].

Next, we turn to continuous time ARMA (CARMA) processes that can be regarded as limits of discrete time ARMA processes as the time interval between successive observations tends to zero. CARMA processes can be formally expressed as solutions to stochastic differential equations together with a given interpretation. Usually, higher order CARMA differential equations are interpreted via observation and state equations. Furthermore, under some restrictions on the AR polynomial and a suitable (random) initial condition, they yield stationary solutions. For more details on CARMA processes, we refer to [17] and references therein.

The CAR(1) process, better known as the Ornstein-Uhlenbeck process, is in fact given by the Langevin differential equation for a Brownian particle, and it has been extensively covered in the literature. The Langevin equation has been generalized e.g. by replacing the driving Brownian force with other Lévy or stationary increment processes. It is a well-known fact that the Langevin equation driven by a stationary increment process satisfying mild integrability conditions admits a unique stationary solution (see e.g. [9]). Furthermore, it was shown in [73] that a generalization of the Langevin equation characterizes all stationary processes possessing continuous sample paths. Extensions of the Langevin equation and estimation in such models, and related models, have been considered e.g. in [3], [49], [58], [20], [68], [48] and [54], to name a few. However, to the best of our knowledge, most of the literature regarding the estimation related to the Langevin equation is restricted to specific, often Gaussian or Lévy, driving forces. Especially fractional Ornstein-Uhlenbeck processes, which were introduced by [18], recovered by fractional Brownian motion driven Langevin equations have received plenty of attention recently (see e.g. [34], [7], [69], [8] and references therein). Finally, we mention the paper [43] by John Lamperti, where he proved that there exists a one-to-one correspondence between stationary and self-similar processes given by a transformation named after him. In particular, the transformation of Brownian motion as a $\frac{1}{2}$-self-similar process recovers the stationary Ornstein-Uhlenbeck process.

On this dissertation

The common theme of this dissertation is to introduce new statistical models for stationary processes providing us with novel estimation methods. In Publication I, Publication II, Publication III and Publication V, we study discrete time stationary processes, whereas Publication IV is devoted to continuous time stationary processes. The rest of this document is organized as follows.

In Chapter 2, we first introduce stationary processes in general, before discussing the discrete and continuous time cases separately in the contexts of ARMA and Ornstein-Uhlenbeck processes. In Chapter 3, we present some basic concepts and results of asymptotic theory that are central for the publications of the dissertation. Chapter 4 provides summaries of the publications.

In Publication I, by applying a discrete time Lamperti theorem, we show that all strictly stationary processes are characterized by the $\mathrm{AR}(1)$ equation when the conventional assumptions related to the noise process are relaxed. Under the assumption of square integrability, the characterization lays the foundation for a novel estimation method based on autocovariance estimators of the modeled stationary process. Furthermore, we show that our closed-form model parameter estimator inherits consistency and asymptotic normality from the autocovariance estimators. It is worth to mention that the approach covers not only all ARMA processes but also essentially any stationary processes.

In Publication I, we discussed how the proposed estimation method can be applied to any square integrable stationary process, excluding a class exhibiting some specific characteristics. In Publication II, we provide a detailed analysis of these special cases. In particular, we show that the class consists of degenerate or approximately degenerate processes and hence, they do not provide very useful models in general.

In Publication III, we study a variant of the ARCH model, whose proper estimation has proven to be a challenging task in the past (see e.g. [5]). Most importantly, by applying the methods introduced in Publication I and Publication II, we define closed-form estimators for the model parameters. Furthermore, we give sufficient conditions under which the estimators are consistent.

In Publication IV, in order to extend the characterization of strictly stationary processes of [73] to cover also multivariate settings, we first state a multidimensional counterpart of the Lamperti theorem. After this, the characterization leads to a novel estimation method of square integrable stationary processes and the unknown parameter matrix of the Langevin equation with a remarkable general driving noise process. The method is based on continuous time algebraic Riccati equations (CAREs) and by applying the related perturbation theory, we show that our estimator inherits consistency from the chosen autocovariance (function) estimator of the stationary process. Furthermore, the asymptotic distribution of the estimator is a linear function of the limiting random process of the autocovariance estimator.

In Publication V, we complete our research of stationary processes by concerning the multivariate discrete time setting. In a similar manner as in Publication IV, we begin by extending the characterization of Publication I to cover multivariate settings as well. Supposing square integrability, this leads to a CARE-based estimation method for the parameter matrix of the characterization. The asymptotic properties of the estimator are inherited as in the continuous time case treated in Publication IV.

Potential continuative topics for future research could be as follows: extensions of the characterization to stationary processes indexed by different parameter sets and the related estimation methods, prediction laws in the models proposed in this dissertation, and assumptions on the noise such that the asymptotic results of the introduced parameter estimators apply.

2. Stationary processes

Stationary processes form a focal class of stochastic processes, and they have been extensively studied in the literature and widely applied in various fields of science. Also, in many cases, a non-stationary time-series can be stationarized via a suitable transformation enabling the utilization of stationary models.

In general terms, a stochastic process is a collection $(X_t)_{t\in T}$ of random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in a common measurable space $(G, \mathcal{G})$. That is, $X_t : \Omega \to G$ is a measurable function for every $t \in T$. However, sometimes it is more convenient to view a stochastic process as a function $X : \Omega \times T \to G$ of two variables. For an introduction to stochastic processes, we refer to [62]. A more in-depth discussion of the topic including details on stationary processes is provided in [28].

In this dissertation, we consider stochastic processes indexed by the real numbers or the integers, and taking values in an n-dimensional real space. Hence, in the sequel, we assume that either $T = \mathbb{R}$ or $T = \mathbb{Z}$. Moreover, in this connection, we discuss the one-dimensional case and hereby $(G, \mathcal{G}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, where $\mathcal{B}(\mathbb{R})$ denotes the Borel sigma-algebra. However, the vast majority of the content of this chapter can be extended to multivariate settings in a straightforward manner. We start by providing the definitions of the two most commonly used types of stationarity.

Definition 2.1. Let $X = (X_t)_{t\in T}$ and $Y = (Y_t)_{t\in T}$ be stochastic processes. If for every $m \in \mathbb{N}$ and $t = [t_1, t_2, \dots, t_m]^\top \in T^m$ the random vectors $[X_{t_1}, X_{t_2}, \dots, X_{t_m}]^\top$ and $[Y_{t_1}, Y_{t_2}, \dots, Y_{t_m}]^\top$ have identical probability distributions, we write

$$X \overset{\text{law}}{=} Y.$$

Definition 2.2. Let $X = (X_t)_{t\in T}$ be a stochastic process. If

$$(X_{t+s})_{t\in T} \overset{\text{law}}{=} (X_t)_{t\in T}$$

for every $s \in T$, then X is strictly stationary.

The above definition states that multidimensional distributions of a strictly stationary process are invariant under uniform shifts in time. In general, stationarity of the one-dimensional distributions (or n-dimensional with a fixed n) is not a sufficient condition for strict stationarity, as the following example illustrates.

Example 2.3. Let $X = (X_t)_{t\in\mathbb{Z}}$ be such that $X_{2k+1} = Y$ for $k \in \mathbb{Z}$, and $(X_{2k})_{k\in\mathbb{Z}}$ is an IID process independent of Y and $X_{2k} \overset{\text{law}}{=} Y$. Then the one-dimensional distributions of X are stationary, but

$$\begin{bmatrix} X_1 \\ X_3 \end{bmatrix} \overset{\text{law}}{=} \begin{bmatrix} Y \\ Y \end{bmatrix} \overset{\text{law}}{\neq} \begin{bmatrix} X_0 \\ X_2 \end{bmatrix}.$$

Definition 2.4. Let $X = (X_t)_{t\in T}$ be a stochastic process. If

(1) $\mathbb{E}(X_t) = \mu, \quad t \in T$

(2) $\operatorname{var}(X_t) = \sigma^2 < \infty, \quad t \in T$

(3) $\operatorname{cov}(X_t, X_{t+s}) = \gamma(s), \quad t, s \in T,$

then X is weakly stationary. In this case, the function γ is called the autocovariance function of X.

Corollary 2.5. *From Definitions 2.2 and 2.4, it follows that every strictly stationary process with a finite second moment is also weakly stationary. In addition, since Gaussian distributions are characterized by the mean vector and the covariance matrix, every Gaussian weakly stationary process is also strictly stationary.*

Next, we turn briefly to autocovariance functions that can be characterized via the property of positive semidefiniteness.

Definition 2.6. Let $f : T \to \mathbb{R}$. In addition, let $m \in \mathbb{N}$, $t \in T^m$ and $z \in \mathbb{R}^m$ be arbitrary. If

$$\sum_{i,j=1}^{m} z_j f(t_i - t_j) z_i \geq 0,$$

then f is a positive semidefinite function.

Theorem 2.7. *Let γ be the autocovariance function of a stationary process. Then*

(1) $\gamma(0) \geq 0$

(2) $|\gamma(t)| \leq \gamma(0), \quad t \in T$

(3) $\gamma(-t) = \gamma(t), \quad t \in T$

(4) γ *is positive semidefinite.*

Conversely, if $f : T \to \mathbb{R}$ *is symmetric and positive semidefinite, then it is the autocovariance function of a (Gaussian) stationary process.*

Proof. The first two properties are given by non-negativeness of variance and the Cauchy-Schwarz inequality. Positive semidefiniteness of γ follows from the fact that $\operatorname{cov}(X)$ is positive semidefinite and setting $X = [X_{t_1}, \ldots, X_{t_m}]^\top$. A proof for the other direction in the case of $T = \mathbb{Z}$ can be found in [16]. However, the same proof is applicable also when $T = \mathbb{R}$. □

Remark 2.8. *In fact, positive semidefiniteness together with symmetricity of* f *gives* (1) *and* (2) *with suitable choices of* t *and* z *in Definition 2.6.*

In Publication II we apply the following result regarding the rank of a covariance matrix as we study a certain class of stationary processes with cyclical type of auto-covariance functions.

Lemma 2.9. *Let* X *be an* n*-dimensional random vector. Then the algebraic multiplicity* $m_a(0)$ *of zero as an eigenvalue of* $\Sigma = \operatorname{cov}(X)$ *equals to the number of linear dependencies within* X*. More precisely,* $m_a(0) = k$ *if and only if all elements of* X *can be expressed linearly from (and no less than)* $n - k$ *fixed elements of* X*.*

Proof. Let us suppose without loss of generality that X is centered. First, assume the redundancy of X. Then it is straightforward to find k linearly independent non-zero vectors such that $a_i^\top X = 0$ almost surely. Furthermore

$$\Sigma a_i = \mathbb{E}(XX^\top)a_i = \mathbb{E}(XX^\top a_i) = 0$$

and thus, a_i is an eigenvector of Σ associated with the eigenvalue zero. Next, assume that a_i is an eigenvector of Σ associated with the eigenvalue zero. Then

$$a_i^\top \Sigma a_i = a_i^\top \mathbb{E}(XX^\top)a_i = \mathbb{E}\left[(a_i^\top X)^2\right] = 0$$

and thus, $a_i^\top X = 0$ almost surely. We conclude the proof by recalling that the eigenvectors of a covariance matrix are orthogonal ensuring that the linear system

$$\begin{bmatrix} a_1^\top \\ \vdots \\ a_k^\top \end{bmatrix} X = 0$$

of equations is of full-rank k. □

2.0.1 Stationary increment processes

Many well-known processes, such as Brownian motion and the Poisson process, exhibit stationarity in their increments. These so-called stationary increment processes are imperative for this dissertation motivating the following short overview of the topic.

Definition 2.10. Let $X = (X_t)_{t\in T}$ be a stochastic process. Let $m \in \mathbb{N}$ and $t, s \in T^m$ be arbitrary. If, for every $h \in T$,

$$\begin{bmatrix} X_{t_1+h} - X_{s_1+h} \\ \vdots \\ X_{t_m+h} - X_{s_m+h} \end{bmatrix} \overset{\text{law}}{=} \begin{bmatrix} X_{t_1} - X_{s_1} \\ \vdots \\ X_{t_m} - X_{s_m} \end{bmatrix}, \tag{2.1}$$

then X is a stationary increment process.

The next lemma shows that the stationary increment property can be compactly written in terms of laws of incremental processes.

Lemma 2.11. *Let $X = (X_t)_{t\in T}$ be a stochastic process. Then X is a stationary increment process if and only if*

$$(X_{t+h} - X_h)_{t\in T} \overset{law}{=} (X_t - X_0)_{t\in T}, \quad h \in T.$$

Proof. Let the above equality of laws hold for every $h \in T$ and choose
$t = [t_1, \dots, t_m, s_1, \dots, s_m]$. Then

$$\begin{bmatrix} X_{t_1+h} - X_h \\ \vdots \\ X_{t_m+h} - X_h \\ X_{s_1+h} - X_h \\ \vdots \\ X_{s_m+h} - X_h \end{bmatrix} \overset{\text{law}}{=} \begin{bmatrix} X_{t_1} - X_0 \\ \vdots \\ X_{t_m} - X_0 \\ X_{s_1} - X_0 \\ \vdots \\ X_{s_m} - X_0 \end{bmatrix}$$

giving (2.1). To prove the other implication, we simply choose $s = 0$ in (2.1). □

In the discrete case $T = \mathbb{Z}$, we obtain even more convenient characterization of stationary increment processes.

Lemma 2.12. *Let $X = (X_t)_{t\in\mathbb{Z}}$ be a stochastic process. Let $\Delta X = (\Delta_t X)_{t\in\mathbb{Z}}$, where $\Delta_t X = X_t - X_{t-1}$. Then X is a stationary increment process if and only if ΔX is strictly stationary.*

Proof. Let ΔX be strictly stationary and without loss of generality assume that $t \geq s$. Then

$$\begin{bmatrix} X_{t_1+h} - X_{s_1+h} \\ \vdots \\ X_{t_m+h} - X_{s_m+h} \end{bmatrix} = \begin{bmatrix} \sum_{k=s_1+1}^{t_1} \Delta_{k+h} X \\ \vdots \\ \sum_{k=s_m+1}^{t_m} \Delta_{k+h} X \end{bmatrix} \overset{\text{law}}{=} \begin{bmatrix} \sum_{k=s_1+1}^{t_1} \Delta_k X \\ \vdots \\ \sum_{k=s_m+1}^{t_m} \Delta_k X \end{bmatrix} = \begin{bmatrix} X_{t_1} - X_{s_1} \\ \vdots \\ X_{t_m} - X_{s_m} \end{bmatrix}.$$

The other direction is straightforward. □

In Publication I and Publication IV, we consider centred stationary increment processes with $X_0 = 0$. In this case, we obtain the following useful representation for the covariance function.

Lemma 2.13. *Let $X = (X_t)_{t\in T}$ be a centred stationary increment process with a finite variance function $v(\cdot)$ and $X_0 = 0$. Then*

$$\operatorname{cov}(X_t, X_s) = \frac{1}{2}(v(t) + v(s) - v(t-s))$$

Proof. Since $X_0 = 0$, and by stationarity of increments

$$\mathbb{E}X_{t-s}^2 = \mathbb{E}(X_t - X_s)(X_t - X_s) = \mathbb{E}X_t^2 + \mathbb{E}X_s^2 - 2\mathbb{E}X_t X_s$$

yielding the result after a rearrangement of the terms. □

2.0.2 Self-similar processes

Self-similar processes are stochastic processes whose distributions are invariant under appropriate scalings of time and space. Moreover and most importantly for us, self-similar processes can be characterized via bijective transformations of strictly stationary processes. For details on self-similar processes, see [23]. We start an introduction to self-similar processes by presenting the original definition concerning the continuous time setting.

Definition 2.14. Let $Y = (Y_t)_{t\geq 0}$ be a stochastic process. If for every $a > 0$ there exists $b > 0$ such that

$$(Y_{at})_{t\geq 0} \overset{\text{law}}{=} (bY_t)_{t\geq 0},$$

then Y is self-similar.

In his seminal paper [43], Lamperti proved that b has a particular functional form under a few assumptions.

Theorem 2.15. *Suppose that $Y = (Y_t)_{t\geq 0}$ is continuous in probability at zero, non-trivial and self-similar. Then there exists $\theta \geq 0$ such that $b = b(a) = a^\theta$.*

Above, by triviality it is meant that Y_t is almost surely a constant for every $t \geq 0$. In the present-day literature, it is customary to define self-similarity according to Theorem 2.15. Moreover, it can be shown that Y is completely degenerate ($Y_t = Y_0$ a.s. for every t) if and only if $\theta = 0$ (see e.g. [23]) and hence, it is common to consider only the case $\theta > 0$. Since in Publication I we introduce discrete time θ-similar processes and the related Lamperti transformation, we use a slightly modified definition that covers both discrete and continuous time cases simultaneously.

Definition 2.16. Let $\theta > 0$ and let $Y = (Y_{e^t})_{t\in T}$ be a stochastic process. If

$$(Y_{e^{t+s}})_{t\in T} \overset{\text{law}}{=} (e^{s\theta} Y_{e^t})_{t\in T}$$

for every $s \in T$, then Y is θ-self-similar.

Remark 2.17. *The only difference between our definition and the standard definition of the continuous time setting (see e.g. [64]) is that we do not consider the process Y at the origin. Otherwise, our definition is obtained through a change of variable from the conventional one. Moreover, from 2.16 it follows that $\lim_{t\to-\infty} Y_{e^t} = 0$ in probability. Furthermore, if Y is stochastically continuous at zero, then $Y_0 = 0$ almost surely coinciding with the standard definition.*

One well-known example of a self-similar process is the fractional Brownian motion (fBm) that is widely applied e.g. in mathematical in order to model long-range dependencies evident in financial time-series.

Example 2.18. The fractional Brownian motion $B^H = (B_t^H)_{t\geq 0}$ with Hurst index $H \in (0,1)$ is the zero mean Gaussian process with the covariance function

$$\operatorname{cov}(B_t^H, B_s^H) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}).$$

Equivalently, it is the unique zero mean Gaussian H-self-similar process with stationary increments. Particularly, with $H = \frac{1}{2}$ we obtain the Brownian motion. For more details on fBm, an interested reader may consider e.g. [59] or [53].

In [43], Lamperti also gave a transformation providing a one-to-one correspondence between θ-similar processes and strictly stationary processes. Again, we state a corresponding result in our terms taking into account of the discrete and continuous cases at the same time. Proofs of multivariate counterparts of the theorem can be found in Publication IV and Publication V.

Definition 2.19. Let $\theta > 0$, and let $X = (X_t)_{t\in T}$ and $Y = (Y_{e^t})_{t\in T}$ be stochastic processes. We define the Lamperti transformation and its inverse by

$$(\mathcal{L}_\theta X)_{e^t} = e^{t\theta} X_t, \quad t \in T$$
$$\left(\mathcal{L}_\theta^{-1} Y\right)_t = e^{-t\theta} Y_{e^t}, \quad t \in T.$$

Theorem 2.20. *The Lamperti transformation $\mathcal{L}_\theta$, together with its inverse $\mathcal{L}_\theta^{-1}$, defines a bijection between the sets of strictly stationary processes and θ-self-similar processes.*

2.1 ARMA processes

In this section, we discuss the discrete case, where $T = \mathbb{Z}$. In discrete time, stationary time-series are often modeled by applying different ARMA (autoregressive moving-average) models, or their extensions. One significant feature of the ARMA processes is the following approximation property. Let X be an arbitrary stationary process with $\gamma_X(t) \to 0$ as $t \to \infty$, and let $k \in \mathbb{N}$. Then there exists an ARMA process with the autocovariance function $\gamma_k(\cdot)$ such that $\gamma_k(t) = \gamma_X(t)$ for every $|t| \leq k$.

Heuristically, an ARMA process is a stochastic process dependent on its own history and on the history of a noise process. In the literature, the noise is conventionally assumed to be white. That is, a sequence of uncorrelated random variables with a common mean and variance. It would also be possible to consider strictly, but not weakly stationary noise processes (see e.g. [52]). However, since the main results of this dissertation are estimation methods based on the second moments, we assume that the IID processes serving as noises are also weakly stationary with variance $\sigma^2 < \infty$.

The principal references for this section are [16], [56] and [29].

Definition 2.21. Let $\epsilon = (\epsilon_t)_{t\in\mathbb{Z}}$ be a white noise or an IID process. If

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}, \quad t \in \mathbb{Z}, \tag{2.2}$$

then $X = (X_t)_{t\in\mathbb{Z}}$ is an ARMA(p, q) process.

The left-hand side and the right-hand side of 2.2 are known as the autoregressive and moving-average parts of the ARMA process, respectively. Moreover, $\{\phi_i\}_{i=1}^p$ and $\{\theta_i\}_{i=1}^q$ are the respective model parameters.

Definition 2.22. Let $X = (X_t)_{t\in\mathbb{Z}}$ be an ARMA(p, q) process. Then X is a causal (with respect to ϵ) if there exists a sequence $(\psi_t)_{t\in\mathbb{N}\cup\{0\}}$ such that

$$\sum_{t=0}^{\infty} |\psi_t| < \infty$$

and

$$X_t = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k}, \quad t \in \mathbb{Z}. \tag{2.3}$$

Remark 2.23. *It is clear from* (2.3) *that if X is causal, then X is also stationary, where the type of stationarity is inherited from the noise process ϵ. However, the converse is not true in general. It is possible to construct stationary ARMA processes for which X_t is correlated with the future values ϵ_s, $s > t$, of the noise. Since these type of solutions of ARMA equations seemingly violate causality, it is customary that only solutions of the form* (2.3) *are considered. For more details on the topic, see e.g. [16] or [56].*

The ARMA equation (2.2) can be compactly written by employing polynomials of the lag operator.

Definition 2.24. Let $X = (X_t)_{t\in\mathbb{Z}}$. The lag operator L is defined by

$$LX_t = X_{t-1}, \quad t \in \mathbb{Z}.$$

Definition 2.25. The lag polynomial representation of an ARMA(p, q) process X is given by

$$\phi(L)X_t = \theta(L)\epsilon_t, \quad t \in \mathbb{Z},$$

where

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p$$
$$\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q$$

are the lag polynomials of the autoregressive and moving-average parts of X, respectively.

The existence of a causal solution to an ARMA equation can be discussed in terms of the lag polynomials. A proof for the following theorem can be found e.g. in [56].

Theorem 2.26. *Let $X = (X_t)_{t\in\mathbb{Z}}$ be an ARMA(p,q) process such that the corresponding lag polynomials $\phi(L)$ and $\theta(L)$ do not share roots. Then X is causal (with respect to ϵ) if and only if the roots of $\phi(L)$ lie outside the closed unit disk of the complex plane. Furthermore, the coefficients of the causal representation* (2.3) *are given by*

$$\psi(z) = \sum_{k=0}^{\infty} \psi_k z^k = \frac{\theta(z)}{\phi(z)}, \quad |z| \le 1. \tag{2.4}$$

Remark 2.27. *The structure of the proof of Theorem 2.26 in [56] shows that if the roots of $\phi(L)$ lie outside the closed unit disk, then X is causal with* (2.4) *regardless of whether the two polynomials share roots or not.*

Remark 2.28. *From* (2.4) *it is obvious that X is the unique solution to the ARMA equation* (2.2).

Corollary 2.29. *An AR(1) process*

$$X_t - \phi_1 X_{t-1} = \epsilon_t, \quad t \in \mathbb{Z} \tag{2.5}$$

is causal if and only $|\phi_1| < 1$.

The following theorem encapsulates the existence of a unique stationary solution to a non-redundant ARMA equation.

Theorem 2.30. *Assume that the lag polynomials $\phi(L)$ and $\theta(L)$ do not share roots. Then there exists a unique stationary solution to* (2.2) *if and only if $\phi(z) \neq 0$ for all $|z| = 1$.*

In the case of a pure autoregressive ARMA process, there exists so-called Yule-Walker equations for the parameters in terms of the autocovariance function that can also be applied in estimation. Furthermore, in this dissertation, we define estimators for square integrable stationary processes inspired by the classical result. On how to derive the following classical system of equations, we refer to [29].

Lemma 2.31. *Let $X = (X_t)_{t\in\mathbb{Z}}$ be a causal AR(p) process and let σ^2 be the variance of the noise process ϵ. Set $\Phi = [\phi_1 \dots \phi_p]^\top$, $\gamma_p = [\gamma(1) \dots \gamma(p)]^\top$ and*

$$\Gamma_p = \begin{bmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(p-1) \\ \gamma(1) & \gamma(0) & \dots & \gamma(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(p-1) & \gamma(p-2) & \dots & \gamma(0) \end{bmatrix}.$$

Then the Yule-Walker equations

$$\gamma(0) - \Phi^\top \gamma_p = \sigma^2$$
$$\Gamma_p \Phi = \gamma_p$$

are satisfied.

2.1.1 On ARMA family

There exists an abundance of extensions and variants of ARMA models from which we next present a few that are selected with varying criteria, e.g. relevance to articles of the dissertation or recognition in the literature in general. First, we define the difference operator providing us with a straightforward way to stationarize processes exhibiting e.g. seasonal behavior.

Definition 2.32. Let $X = (X_t)_{t\in\mathbb{Z}}$ and $s \in \mathbb{N}$. The seasonal difference operator D_s is defined by

$$D_s X_t = (1 - L^s)X_t = X_t - X_{t-s}, \quad t \in \mathbb{Z}.$$

For $s = 1$ we use a simpler notation $D \coloneqq D_1$.

Definition 2.33. Let $\epsilon = (\epsilon_t)_{t\in\mathbb{Z}}$ be a white noise or an IID process. In addition, let h and H be non-negative integers. If

$$\phi(L)\Phi(L^s)D_s^H D^h X_t = \theta(L)\Theta(L^s)\epsilon_t, \quad t \in \mathbb{Z}, \tag{2.6}$$

where $\phi(L)$ and $\theta(L)$ are as in Definition 2.25, and

$$\Phi(L^s) = 1 - \Phi_1 L^s - \Phi_2 L^{2s} - \ldots - \Phi_P L^{Ps}$$
$$\Theta(L^s) = 1 + \Theta_1 L^s + \Theta_2 L^{2s} + \ldots + \Theta_Q L^{Qs},$$

then X is a SARIMA$(p, h, q)(P, H, Q)_s$ (seasonal autoregressive integrated moving-average) process with period s.

Remark 2.34. *If the roots of $\phi(L)$ and $\Phi(L^s)$ lie outside the closed unit disk, then by Remark 2.27 the process $Y_t \coloneqq D_s^H D^h X_t$ is causal with respect to ϵ. Furthermore, if there is no redundant factors in* (2.6) *and* $\max(h, H) > 0$*, then by Theorem 2.30 there exists no stationary solution X.*

The next member of the ARMA family has plenty of applications e.g. in mathematical finance, since it enables the modeling of time-varying variance. The study of such dynamic volatility models was initiated by [24] and [12].

Definition 2.35. Let $\alpha_0, \dots, \alpha_{q-1}, \beta_1, \dots \beta_{p-1} \geq 0$ and $\alpha_q, \beta_p > 0$. If

$$X_t = \sigma_t \epsilon_t \qquad \text{and} \qquad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i X_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2,$$

where $\epsilon = (\epsilon_t)_{t\in\mathbb{Z}}$ is an IID sequence with ϵ_t independent of X_{t-i} for $i \geq 1$, then X is a GARCH(p, q) (generalized autoregressive conditional heteroskedasticity) process.

In addition to the models introduced above, we mention fractionally integrated ARMA (FARIMA) models, where the difference operators D and D_s may also admit non-integer exponents h and H. Here $D^h = (1 - L)^h$ can be interpreted through the binomial expansion together with the gamma function. For more details, we refer to [16]. Furthermore, in non-linear ARMA (NARMA) models the dependence of X_t on the history is given by a non-linear function (the interested reader may consult e.g. [25]). Also exogenous variables have been incorporated to ARMA models (ARMAX) and their variants, from which we next give an example that we investigate in Publication III. The model can been seen as a GARCH$(0, 1)$-X model capturing effects of a liquidity process on the time-dependent volatility.

Example 2.36 (GARCH$(0, 1)$-X)**.** Let $\alpha_0 \geq 0$ and $\alpha_1, l_1 > 0$. Then, we set

$$X_t = \sigma_t \epsilon_t \qquad \text{and} \qquad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + l_1 L_{t-1},$$

where $\epsilon = (\epsilon_t)_{t\in\mathbb{Z}}$ is an IID process with $\mathbb{E}(\epsilon_0) = 0$ and $\operatorname{var}(\epsilon_0) = 1$. Furthermore, $L = (L_t)_{t\in\mathbb{Z}}$ is a positive process and independent of ϵ.

2.2 Ornstein-Uhlenbeck processes

In this section, we discuss the continuous case, where $T = \mathbb{R}$. In continuous time, the well-known Ornstein-Uhlenbeck process can be regarded as the analog of the discrete time AR(1) process. Ornstein-Uhlenbeck processes (of the first kind) arise out of the Langevin differential equation (2.7) evidently similar to the AR(1) equation (2.5). For example, in (2.5) the noise is given by the increments of a random walk, whereas in (2.7) the shocks are given by the infinitesimal changes of the Brownian motion, which is the scaling limit of random walks by Donsker's theorem (see e.g. [10]). Moreover, in Publication I, we show that Equation (2.5) characterizes all discrete time strictly stationary processes, when the corresponding noise is allowed to be possibly correlated. Furthermore, in Publication IV we show that (2.7) characterizes all continuous time strictly stationary processes possessing continuous paths, when the concept of the noise is extended in a similar manner. Thereby, we give a short introduction to Ornstein-Uhlenbeck processes and some of their extensions. For more details on the Ornstein-Uhlenbeck process, see e.g. [27].

Definition 2.37. Let $\theta > 0$ and let $W = (W_t)_{t\in\mathbb{R}}$ be a two-sided Brownian motion. If

$$dU_t = -\theta U_t dt + dW_t, \quad t \in \mathbb{R}, \tag{2.7}$$

then $U = (U_t)_{t\in\mathbb{R}}$ is the Ornstein-Uhlenbeck process (of the first kind).

Above, the two-sided Brownian motion is simply the concatenation of two independent copies of Brownian motion indexed by the positive and negative integers, respectively. By the next theorem, the classical Langevin equation admits a (strictly) stationary solution.

Theorem 2.38. *The solution to* (2.7) *with a given random "initial" condition* U_0 *reads*

$$U_t = e^{-\theta t}\left(U_0 + \int_0^t e^{\theta s} dW_s\right).$$

Moreover, the unique stationary solution is given by

$$U_0 = \int_{-\infty}^0 e^{\theta s} dW_s.$$

Remark 2.39. *Technically speaking,* U_0 *is not an initial value in the classical sense as we are considering Equation* (2.7) *in* $\mathbb{R}$. *However, we use this loose notion in the continuation as it does not leave any room for confusion.*

Remark 2.40. *The stochastic integrals in Theorem 2.38 can be interpreted e.g. as path-wise Riemann-Stieltjes integrals, Wiener integrals or Itô integrals. Nevertheless, in our case, these integrals exist and coincide. An overview on Riemann-Stieltjes integrals can be found in [63]. For details on stochastic integration, we refer to [37] and [55].*

Theorem 2.41. *The unique stationary solution to* (2.7) *is the zero mean Gaussian process with the following autocovariance function*

$$\gamma(s) = \frac{e^{-\theta|s|}}{2\theta}.$$

Equivalently, it is the unique stationary Gauss-Markov process with continuous paths.

Equation (2.7) can be generalized by replacing the Brownian motion e.g. with a two-sided fractional Brownian motion recovering the so-called fractional Ornstein-Uhlenbeck process in Theorem 2.38. A two-sided fBm is the zero-mean Gaussian process with the covariance function given in Example 2.18. However, due to the dependency of increments with $H \neq \frac{1}{2}$, a two-sided fBm can not be constructed by concatenating two independent and identical copies of one-sided fractional Brownian motions. Moreover, since fBm is not a semimartingale for $H \neq \frac{1}{2}$, the classical Itô calculus is not available. However, since the paths of fBm are almost surely (α-Hölder for $\alpha < H$) continuous, one can still interpret the stochastic integrals of Theorem 2.38 as path-wise Riemann-Stieltjes integrals. This fact is encapsulated by Theorem 2.42 (see e.g. [63] and [32]) together with Theorem 2.43 [73]. For an exhaustive discussion on integration with respect to fractional Brownian motion, we refer to [53]. For details on fractional Ornstein-Uhlenbeck processes, see e.g. [40].

Theorem 2.42. *Let f be a continuous and g be a monotonically increasing function on the interval $[0,t]$. Then the Riemann-Stieltjes integral*

$$\int_0^t f(s)dg(s)$$

exists. Moreover, the integration by parts formula

$$\int_0^t f(s)dg(s) = f(t)g(t) - f(0)g(0) - \int_0^t g(s)df(s)$$

is valid and the integral on the right-hand side exists as a Riemann-Stieltjes integral.

Theorem 2.43. *Let $X = (X_t)_{t\in\mathbb{R}}$ be a continuous stationary increment process with $X_0 = 0$. Assume that*

$$\sup_{t\in[0,1]} \mathbb{E}\left(\log|X_t|\, \mathbb{1}_{\{|X_t|>1\}}\right)^{2+\delta} < \infty \tag{2.8}$$

for some $\delta > 0$. Then, for every $\theta > 0$,

$$\lim_{u\to\infty} \int_{-u}^0 e^{\theta s} dX_s$$

exists almost surely defining an almost surely finite random variable.

As a consequence, we obtain the next elementary corollary.

Corollary 2.44. *Let $X = (X_t)_{t\in\mathbb{R}}$ be a continuous stationary increment process with $X_0 = 0$. Assume that $\sup_{t\in[0,1]} \mathbb{E}|X_t|^\alpha < \infty$ for some $\alpha > 0$. Then the assertion of Theorem 2.43 is satisfied.*

Proof. Let $\delta > 0$ be fixed. Then there exists $M > 1$ such that $x^{\frac{\alpha}{2+\delta}} > \log(x)$ for all $x \geq M$. Now, we write

$$\mathbb{E}\left(\log|X_t|\, \mathbb{1}_{\{|X_t|>1\}}\right)^{2+\delta} = \mathbb{E}\left(\log|X_t|\, \mathbb{1}_{\{1<|X_t|<M\}}\right)^{2+\delta} + \mathbb{E}\left(\log|X_t|\, \mathbb{1}_{\{M\leq|X_t|\}}\right)^{2+\delta},$$

where the first term is clearly bounded. For the latter, we obtain

$$\mathbb{E}\left(\log|X_t|\,\mathbb{1}_{\{M\leq|X_t|\}}\right)^{2+\delta} \leq \mathbb{E}\left(|X_t|^{\frac{\alpha}{2+\delta}}\,\mathbb{1}_{\{M\leq|X_t|\}}\right)^{2+\delta} \leq \mathbb{E}|X_t|^{\alpha}$$

completing the proof. □

Finally, we present a theorem elaborating the connection between the Langevin equation (2.7) and stationary processes. The related integral is well-defined by Theorems 2.42 and 2.43. Details can be found, for example, in Publication IV.

Theorem 2.45. *Let $\theta > 0$ and let $X = (X_t)_{t\in\mathbb{R}}$ be a continuous stationary increment process with $X_0 = 0$ satisfying* (2.8) *for some $\delta > 0$. Then the generalized Langevin equation*

$$dU_t = -\theta U_t dt + dX_t, \quad t \in \mathbb{R}$$

admits a unique stationary solution given by

$$U_t = e^{-\theta t}\int_{-\infty}^{t} e^{\theta s} dX_s.$$

3. Asymptotic theory

In this chapter, we recall some basic concepts and theorems that are central for the publications of this dissertation. First, we take a relative general approach to weak convergence of random elements enabling the treatment of function spaces considered in Publication IV.

3.1 On weak convergence

We start by giving a definition for weak convergence of random elements taking values in a general metric space (M, d). The metric induces a topology on M with a basis given by the open balls $\{y \in M : d(x, y) < r\}$, where $x \in M$ and $r > 0$. Furthermore, if the metric space is separable, then the Borel sigma-algebra $\mathcal{B}(M)$ equals to the sigma-algebra generated by the open balls [11].

Definition 3.1. Let (M, d) be a metric space and let $(X_T)_{T\geq 1}$, $X_T : \Omega \to M$ be a sequence of random elements. If for every bounded continuous $f : M \to \mathbb{R}$ it holds that $\mathbb{E}f(X_T) \to \mathbb{E}f(X)$, then X_T converges weakly to X and we write $X_T \xrightarrow{\text{law}} X$.

In Publication IV, we consider weak convergence in the space $\mathcal{C} = \mathcal{C}([0, t], \mathbb{R}^m)$ of m-dimensional continuous functions with a fixed $t > 0$. The space $\mathcal{C}$ equipped with the sup-norm $\|f\|_\infty = \sup_{s\in[0,t]} \|f(s)\|$, where $\|\cdot\|$ is the usual l^2 vector norm, is a separable metric space. Furthermore, it can be shown that sufficient conditions for weak convergence are given by tightness of the induced measures $\mu_T(A) = \mathbb{P}(X_T \in A)$ on $(\mathcal{C}, \mathcal{B}(\mathcal{C}))$ and convergence of the finite dimensional distributions (see [10] or [21])

$$\mathbb{P}(X_T(s_1) \in A_1, \dots, X_T(s_n) \in A_n), \quad n \in \mathbb{N}, s \in [0, t]^n \text{ and } A_i \in \mathcal{B}(\mathbb{R}^m). \quad (3.1)$$

Next, we provide the definition of tightness in a general metric space.

Definition 3.2. Let Π be a family of probability measures on a metric space (M, d). If for every $\epsilon > 0$ there exists a compact $K_\epsilon \subset M$ such that $1 - \mu(K_\epsilon) \leq \epsilon$ for all $\mu \in \Pi$, then the family Π is tight.

By applying the Arzela-Ascoli theorem and Kolmogorov's continuity criterion, one can obtain the following result on tightness in the space of continuous functions. For details, see e.g. [10] and [21].

Theorem 3.3. *Let $(X_T)_{T\geq 1}$ be a sequence of random elements $X_T : \Omega \to \mathcal{C}$. Then the corresponding sequence of induced probability measures is tight if*

(1) For every $\epsilon > 0$ there exists N such that

$$\mathbb{P}\left(|X_T(0)| > N\right) \leq \epsilon \quad \textit{for } T \geq 1.$$

(2) For some $\alpha, \beta, C > 0$ it holds that

$$\mathbb{E}|X_T(s_1) - X_T(s_2)|^\beta \leq C|s_1 - s_2|^{1+\alpha} \quad \textit{for } T \geq 1 \textit{ and } s_1, s_2 \in [0, t].$$

Furthermore, for $X_T \xrightarrow{\text{law}} X$ it suffices to show convergence of the finite dimensional distributions (3.1) and item (2) of Theorem 3.3. Indeed, from convergence of the finite dimensional distributions it follows that

$$\mathbb{P}(|X_T(0)| > N) \leq \mathbb{P}(|X(0)| > N) + \frac{\epsilon}{2} \leq \epsilon,$$

when M and T are chosen to be large enough.

3.2 On estimation

The models of stationary processes introduced in this dissertation lead to autocovariance based estimators for the related model parameters in the spirit of the classical Yule-Walker estimators. The Yule-Walker estimators can be defined through Lemma 2.31 by replacing the autocovariances with estimated values providing us with a moment based estimation method of pure autoregressive ARMA models. Conventionally, a general ARMA model is estimated by applying the maximum likelihood method. However, in this case, the practitioner has to be aware of the exact distribution of the noise process. For details on estimation of ARMA models, we refer to [16].

Moreover, in this dissertation, we show that our estimators inherit their consistency from the autocovariance function of the concerned stationary process. Also, we derive the limiting distributions of the estimators in terms of the limiting distribution of the autocovariance function. For this, we utilize the well-known convergence theorems presented in this section. First, we recall the continuous mapping theorem (for a proof, see[21]).

Theorem 3.4. *Assume that $(X_T)_{T\geq 1}$ is a sequence of random elements on (M, d) converging weakly to X. Let $(\tilde{M}, \tilde{d})$ be another metric space and $f : M \to \tilde{M}$ have discontinuity set D_f with $\mathbb{P}(X \in D_f) = 0$. Then*

$$f(X_T) \xrightarrow{\text{law}} f(X).$$

Remark 3.5. *Theorem 3.4 is also valid for convergence in probability and almost sure convergence. However, without any additional assumptions on f, it does not hold true for L^p convergence in general.*

Corollary 3.6. *Assume that $(X_T)_{T\geq 1}$ converges weakly to X and $(Y_T)_{T\geq 1}$ converges in probability to a constant c. Then*

$$\begin{aligned}
&X_T + Y_T \xrightarrow{\text{law}} X + c \\
&X_T Y_T \xrightarrow{\text{law}} cX \\
&\frac{X_T}{Y_T} \xrightarrow{\text{law}} \frac{X}{c}, \quad \text{if } c \text{ is invertible}.
\end{aligned}$$

Moreover, if the convergence of X_T takes place in probability, then the foregoing results are valid in probability.

Proof. The results follow from the fact that $Z_T := [X_T, Y_T]^\top$ converges weakly (or in probability) to $[X, c]^\top$, and applying the continuous mapping theorem to Z_T with functions $f_1(x, y) = x + y$, $f_2(x, y) = xy$ and $f_3(x, y) = \frac{x}{y}$. □

Corollary 3.6 is also known as Slutsky's theorem. Next, we state the delta method. For a proof, we refer to [72] .

Theorem 3.7. *Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be differentiable at $\theta \in \mathbb{R}^m$, and let $(X_T)_{T\geq 1}$ be a sequence of random vectors in $\mathbb{R}^m$. If $l(T)(X_T - \theta) \xrightarrow{\text{law}} X$ for some rate function $l(T) \to \infty$, then*

$$l(T)(f(X_T) - f(\theta)) \xrightarrow{\text{law}} f'_\theta X,$$

where f'_θ is the Jacobian matrix of f at θ.

In Publication I, we make use of the following simple corollary regarding normally distributed limits.

Corollary 3.8. *Let $f : \mathbb{R}^m \to \mathbb{R}$ be differentiable at $\theta \in \mathbb{R}^m$, and let $(X_T)_{T\geq 1}$ be a sequence of random vectors in $\mathbb{R}^m$. Suppose that $l(T)(X_T - \theta) \xrightarrow{law} X$, where $X \sim \mathcal{N}(\mu, \Sigma)$. Then f'_θ is the gradient of f at θ and*

$$l(T)(f(X_T) - f(\theta)) \xrightarrow{law} \mathcal{N}\left(f'_\theta \mu, f'_\theta \Sigma (f'_\theta)^\top\right)$$

The following theorem, also known as the Cramér-Wold device, can be proved by applying Lévy's continuity theorem. For details, see e.g. [72].

Theorem 3.9. *Let $(X_T)_{T\geq 1}$ be a sequence of random vectors in $\mathbb{R}^m$. Then X_T converges weakly to X if and only if*

$$a^\top X_T \xrightarrow{law} a^\top X \quad \text{for all } a \in \mathbb{R}^m.$$

Finally, in Publication III, we utilize the next non-conventional variant of the law of large numbers. For the reader's convenience, we give also the proof, which was omitted from the original publication.

Theorem 3.10. *Let $(X_T)_{T\in\mathbb{N}}$ be a sequence of random variables with a mutual expectation. In addition, assume that $\operatorname{var}(X_j) \leq C$ and $|\operatorname{cov}(X_j, X_k)| \leq g(|k-j|)$, where $g(i) \to 0$ as $i \to \infty$. Then*

$$\frac{1}{n}\sum_{k=1}^{n} X_k \to \mathbb{E}(X_1)$$

in $L^2(\Omega)$ (and hence, also in probability).

Proof. We have that

$$\mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} X_k - \mathbb{E}(X_1)\right)^2 = \frac{1}{n^2}\operatorname{var}\left(\sum_{k=1}^{n} X_k\right),$$

where

$$\begin{aligned}
\operatorname{var}\left(\sum_{k=1}^{n} X_k\right) &= \sum_{k,j=1}^{n} \operatorname{cov}(X_k, X_j) \\
&= \sum_{k=1}^{n} \operatorname{var}(X_k) + 2\sum_{k=1}^{n}\sum_{j=1}^{k-1} \operatorname{cov}(X_k, X_j) \\
&\leq nC + 2\sum_{k=1}^{n}\sum_{j=1}^{k-1} |\operatorname{cov}(X_k, X_j)|.
\end{aligned}$$

Fix $\delta > 0$. Then, there exists $N_\delta \in \mathbb{N}$ such that $g(|k-j|) < \delta$ whenever $|k-j| \geq N_\delta$. Note also that by Cauchy-Schwarz $|\mathrm{cov}(X_k, X_j)| \leq C$. Assume that $n > N_\delta$. Now

$$\begin{aligned}\sum_{k=1}^{n}\sum_{j=1}^{k-1}|\mathrm{cov}(X_k, X_j)| &\leq \sum_{k=1}^{n}\sum_{j=1}^{k-N_\delta} g(|k-j|) + \sum_{k=1}^{n}\sum_{j=k-N_\delta+1}^{k-1} C \\ &\leq n^2\delta + nN_\delta C.\end{aligned}$$

Hence

$$\mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} X_k - \mathbb{E}(X_1)\right)^2 \leq \frac{nC + 2n^2\delta + 2nN_\delta C}{n^2} = 2\delta + \mathcal{O}\left(\frac{1}{n}\right)$$

concluding the proof, since δ was arbitrary small. □

4. Summaries of the articles

Publication I It was shown in [73] that a generalization of the Langevin equation (2.7) characterizes continuous time strictly stationary processes having continuous sample paths. Motivated by this result, we show that in discrete time the corresponding characterization is given by the AR(1) equation

$$X_t = \phi X_{t-1} + Z_t, \quad t \in \mathbb{Z}, \tag{4.1}$$

where now $0 < \phi < 1$ and the noise Z belonging to a certain class of processes is not necessarily white. Based on Equation (4.1) and under the assumption of finite second moments, we show that ϕ satisfies

$$\phi^2\gamma(n) - \phi(\gamma(n+1) + \gamma(n-1)) + \gamma(n) = r(n), \quad n \in \mathbb{Z}, \tag{4.2}$$

where γ and r are the autocovariance functions of X and Z, respectively. We discuss how to find the correct solution to (4.2) and show that for this, it suffices to know two values of r such that

$$\frac{r(n)}{\gamma(n)} \neq \frac{r(m)}{\gamma(m)}.$$

Consequently, Equations (4.2) provide us with a set of natural estimators for the model parameter ϕ of Yule-Walker type. We prove that consistency and asymptotic normality of these estimators are inherited from the chosen autocovariance estimators of the observed stationary process X. Moreover, since the delta method is employed in derivation of the asymptotic distribution of the estimators, it is obvious that the same technique is applicable also when the limiting distribution of the autocovariances differs from normal.

Finally, we provide a simulation study illustrating convergence of the estimators in the case of AR(1) and ARMA$(1, 2)$ processes.

Publication II In Publication I, we showed that the system (4.2) of equations admits a unique solution and consequently the proposed estimation method can be applied, except within some special class of functions γ. In Publication II, we provide a comprehensive analysis of these special cases. We prove that the functions belonging to the class are indeed autocovariance functions of some (strictly) stationary processes. In addition, we show that such function γ is either dense in $[-\gamma(0), \gamma(0)]$ or periodic. The latter means that there exists N_γ such that $\gamma(m+N_\gamma) = \gamma(m)$ for every $m \in \mathbb{Z}$. Moreover, in the periodic case, the corresponding stationary process is driven linearly by only two random variables. That is, X_t is a linear combination of two fixed random variables with possibly time-dependent coefficients. Furthermore, if γ is dense, then it can be approximated with autocovariance functions of such stationary processes. As an important consequence, we obtain that if the autocovariance function r of the noise Z satisfies the natural assumption $r(n) \to 0$ as $n \to \infty$, then the corresponding stationary process in (4.1) can not be of the special type. Thus, the estimation method of Publication I is applicable.

Publication III We study a variant of the ARCH model in the context of estimation of the three model parameters by applying the method introduced in Publication I and Publication II.

The model is defined by

$$X_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + l_1 L_{t-1}, \quad t \in \mathbb{Z},$$

where α_0 is a non-negative, and α_1 and l_1 are positive parameters. Moreover, $(\epsilon_t)_{t\in\mathbb{Z}}$ is a centred IID noise process with unit variance and $(L_t)_{t\in\mathbb{Z}}$ is a positive process independent of the noise. The model is widely applied in mathematical finance and hence, σ and L are often referred as volatility and liquidity, respectively.

We start by analyzing the existence and uniqueness of a solution with both, non-stationary and stationary liquidities. If the liquidity process is (strictly) stationary, then there exists a unique (strictly) stationary solution X^2 and in this case, we give sufficient and necessary conditions for existence of the autocovariance function. We write the process X^2 in the form (4.1) that under the assumption of finite autocovariance leads to a set of quadratic equations of type (4.2). As in Publication I, these equations can be used to define estimators for the three model parameters α_0, α_1 and l_1.

We also give sufficient conditions for consistency of the typical autocovariance estimators

$$\hat{\gamma}_N(n) = \frac{1}{N} \sum_{t=1}^{N-n} (X_t^2 - \mu)(X_{t+n}^2 - \mu),$$

where μ is the sample mean of the observations $(X_1^2, X_2^2, \ldots, X_N^2)$, expressed in terms of the AR(1) parameter α_1 and the liquidity process L. Consequently, consistency of the parameter estimators follows. We provide several interesting examples of the liquidity process that satisfy the abovementioned requirements. Finally, we illustrate our results by using squared increments of a fractional Brownian motion and a compensated Poisson process as liquidities in a simulation study.

Publication IV Inspired by the results of previous publications and the demand for multivariate models arising out of real-life phenomena, we investigate modeling of multivariate continuous time strictly stationary processes together with estimation of the parameter matrix of the multidimensional Langevin equation.

In order to extend the characterization of [73] to multiple dimensions, we define Θ-self-similar processes for a positive definite matrix Θ. Furthermore, we state the Lamperti theorem in our setting giving one-to-one correspondence between multivariate strictly stationary and Θ-self-similar processes. Consequently, we show that the Langevin equation

$$dU_t = -\Theta U_t + dG_t, \quad t \in \mathbb{R},$$

where G is a stationary increment process of a certain class, characterizes multivariate strictly stationary processes possessing continuous paths. As in Publication I (cf. Equation (4.2)), the characterization provides us with quadratic equations

$$B_t^\top \Theta + \Theta B_t - \Theta C_t \Theta + D_t = 0, \quad t \geq 0 \tag{4.3}$$

for the parameter matrix Θ, where the coefficients are expressed in terms of the autocovariance $\gamma(s) = \mathbb{E}(U_s U_0^\top)$ of the stationary solution and $\operatorname{cov}(G_t)$ for a fixed t. These type of matrix equations are known as continuous time algebraic Riccati equations (CAREs) that are extensively studied in the literature. Especially the existence and uniqueness of a solution is a well-studied topic when (4.3) takes a symmetric form with positive semidefinite C_t and D_t. We construct an estimator based on CAREs, and show that the estimator is consistent under consistency of the chosen autocovariance (function) estimator of the stationary process. Moreover, we derive the asymptotic distribution of the estimator as a linear function of the asymptotic distribution of the autocovariance estimator. Furthermore, we provide expressions for the autocovariance γ in terms of the variance function of the noise and exemplify how these expressions can be utilized in order to verify that the forementioned asymptotic results are valid in a Gaussian case.

Publication V We complete our research of stationary processes by treating the multivariate discrete time case. We start by proving that an $\mathrm{AR}(1)$ type equation characterizes all multidimensional strictly stationary processes by providing a discrete time multivariate version of the Lamperti theorem. Consequently, we derive quadratic equations for the model parameter matrix Φ that are similar to (4.2). However, the obtained equations are not of the symmetric form and hence, the uniqueness of a solution is ambiguous. On the other hand, by applying the approach adapted from Publication IV, we derive symmetric CAREs for $\Theta = I - \Phi$ of the form (4.3). Moreover, by using similar techniques as in Publication IV, we show that the corresponding estimator $\hat{\Theta}_T$ inherits consistency and the rate of convergence from the chosen autocovariance estimators of the stationary process. Furthermore, its limiting distribution is given by a linear function of the limiting distribution of the autocovariance estimators.

In addition, in order to highlight the analogy between the discrete and continuous time cases, we derive equations of the type (4.2) in continuous time, and present the main results of Publication IV paralleling them to the obtained discrete time results.

References

[1] Hirotugu Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

[2] Hirotugu Akaike. Canonical correlation analysis of time series and the use of an information criterion. In *Mathematics in Science and Engineering*, volume 126, pages 27–96. Elsevier, 1976.

[3] David Applebaum. Infinite dimensional Ornstein-Uhlenbeck processes driven by Lévy processes. *Probability Surveys*, 12:33–54, 2015.

[4] Louis Bachelier. Théorie de la spéculation. *Annales scientifiques de l'École normale supérieure*, 17:21–86, 1900.

[5] Natalia Bahamonde, Soledad Torres, and Ciprian A Tudor. ARCH model and fractional Brownian motion. *Statistics & Probability Letters*, 134:70–78, 2018.

[6] Richard T Baillie, Tim Bollerslev, and Hans O Mikkelsen. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30, 1996.

[7] Salwa Bajja, Khalifa Es-Sebaiy, and Lauri Viitasaari. Least squares estimator of fractional Ornstein-Uhlenbeck processes with periodic mean. *Journal of the Korean Statistical Society*, 46(4):608–622, 2017.

[8] Maoudo F Balde, Khalifa Es-Sebaiy, and Ciprian A Tudor. Ergodicity and drift parameter estimation for infinite-dimensional fractional Ornstein-Uhlenbeck process of the second kind. *Applied Mathematics & Optimization*, 81:785–814, 2020.

[9] Ole E Barndorff-Nielsen and Andreas Basse-O'Connor. Quasi Ornstein-Uhlenbeck processes. *Bernoulli*, 17(3):916–941, 2011.

[10] Patrick Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, 1999.

[11] Vladimir I Bogachev. *Weak convergence of measures*. American Mathematical Society, 2018.

[12] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

[13] Tim Bollerslev. Glossary to ARCH (GARCH). *CREATES Research paper*, 49, 2008.

[14] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, 2015.

[15] George EP Box and David A Pierce. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332):1509–1526, 1970.

[16] Peter J Brockwell and Richard A Davis. *Time Series: Theory and Methods*. Springer-Verlag New York, 1991.

[17] Peter J Brockwell and Richard A Davis. *Introduction to Time Series and Forecasting*. Springer, 2016.

[18] Patrick Cheridito, Hideyuki Kawaguchi, and Makoto Maejima. Fractional Ornstein-Uhlenbeck processes. *Electronic Journal of Probability*, 8, 2003.

[19] Manfred Deistler. Stationary processes and linear systems. In *Harmonic Analysis and Rational Approximation*, pages 159–179. Springer-Verlag Berlin Heidelberg, 2006.

[20] Soukaina Douissi, Khalifa Es-Sebaiy, and Ciprian A Tudor. Hermite Ornstein-Uhlenbeck processes mixed with a Gamma distribution. *Publicationes Mathematicae Debrecen*, 96:1–22, 2020.

[21] Richard Durrett. *Stochastic Calculus: A Practical Introduction*. CRC Press, 1996.

[22] Albert Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 322(8):549–560, 1905.

[23] Paul Embrechts and Makoto Maejima. *Selfsimilar Processes*. Princeton University Press, 2002.

[24] Robert F Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

[25] Jianqing Fan and Qiwei Yao. *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer-Verlag New York, 2003.

[26] Christian Francq and Jean-Michel Zakoïan. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10(4):605–637, 2004.

[27] Crispin W Gardiner. *Handbook of Stochastic Methods*. Springer-Verlag Berlin-Heidelberg, 1985.

[28] Iosif I Gikhman and Anatoli V Skorokhod. *The Theory of Stochastic Processes I*. Springer-Verlag Berlin Heidelberg, 2004.

[29] James D Hamilton. *Time Series Analysis*, volume 2. Princeton University Press, 1994.

[30] Heejoon Han and Dennis Kristensen. Asymptotic theory for the QMLE in GARCH-X models with stationary and nonstationary covariates. *Journal of Business & Economic Statistics*, 32(3):416–429, 2014.

[31] Edward J Hannan. The asymptotic theory of linear time-series models. *Journal of Applied Probability*, 10(1):130–145, 1973.

[32] Einar Hille and Ralph S Phillips. *Functional Analysis and Semi-Groups*, volume 31. American Mathematical Society, 1996.

[33] Keith W Hipel and A Ian McLeod. *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, 1994.

[34] Yaozhong Hu, David Nualart, and Hongjuan Zhou. Parameter estimation for fractional Ornstein-Uhlenbeck processes of general Hurst parameter. *Statistical Inference for Stochastic Processes*, 22(1):111–142, 2019.

[35] Marie Hušková, Natalie Neumeyer, Tobias Niebuhr, and Leonie Selk. Specification testing in nonparametric AR-ARCH models. *Scandinavian Journal of Statistics*, 46(1):26–58, 2019.

[36] Kiyosi Itô. Stochastic integral. *Proceedings of the Imperial Academy*, 20(8):519–524, 1944.

[37] Svante Janson. *Gaussian Hilbert Spaces*. Cambridge University Press, 1997.

[38] Robert Jarrow and Philip Protter. A short history of stochastic integration and mathematical finance: The early years, 1880–1970. In *A Festschrift for Herman Rubin*, pages 75–91. Institute of Mathematical Statistics, 2004.

[39] Richard H Jones. Maximum likelihood fitting of ARMA models to time series with missing observations. *Technometrics*, 22(3):389–395, 1980.

[40] Terhi Kaarakka and Paavo Salminen. On fractional Ornstein-Uhlenbeck processes. *Communications on Stochastic Analysis*, 5(1):121–133, 2011.

[41] Mark Kac. Wiener and integration in function spaces. *Bulletin of the American Mathematical Society*, 72(1):52–68, 1966.

[42] Alexander Khintchine. Korrelationstheorie der Stationären Stochastischen Prozesse. *Mathematische Annalen*, 109(1):604–615, 1934.

[43] John Lamperti. Semi-stable stochastic processes. *Transactions of the American Mathematical Society*, 104(1):62–78, 1962.

[44] Paul Langevin. Sur la théorie du mouvement brownien. *Comptes rendus de l'Académie des Sciences*, 146:530–533, 1908.

[45] Shiqing Ling and Wai K Li. On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity. *Journal of the American Statistical Association*, 92(439):1184–1194, 1997.

[46] Shiqing Ling and Michael McAleer. Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory*, 19(2):280–310, 2003.

[47] Greta M Ljung and George EP Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, 1978.

[48] Stanislav Lohvinenko and Kostiantyn Ralchenko. Asymptotic distribution of the maximum likelihood estimator in the fractional Vašíček model. *Theory of Probability and Mathematical Statistics*, 99:149–168, 2019.

[49] Marcin Magdziarz. Fractional Ornstein–Uhlenbeck processes. Joseph effect in models with infinite variance. *Physica A: Statistical Mechanics and its Applications*, 387(1):123–133, 2008.

[50] Vasilis Z Marmarelis. *Nonlinear Dynamic Modeling of Physiological Systems*, volume 10. John Wiley & Sons, 2004.

[51] Henry P McKean. *Stochastic Integrals*. Academic Press, 1969.

[52] Thomas Mikosch, Tamar Gadrich, Claudia Kluppelberg, and Robert J Adler. Parameter estimation for ARMA models with infinite variance innovations. *The Annals of Statistics*, 23(1):305–326, 1995.

[53] Yuliya Mishura. *Stochastic Calculus for Fractional Brownian Motion and Related Processes*. Springer-Verlag Berlin-Heidelberg, 2008.

[54] Yuliya Mishura, Kostiantyn Ralchenko, and Sergiy Shklyar. Parameter estimation for Gaussian processes with application to the model with two independent fractional Brownian motions. In *Stochastic Processes and Applications*, pages 123–146. Springer, Cham, 2018.

[55] Yuliya Mishura and Georgiy Shevchenko. *Theory and Statistical Applications of Stochastic Processes*. ISTE and John Wiley & Sons, 2017.

[56] Klaus Neusser. *Time Series Econometrics*. Springer International Publishing, 2016.

[57] Shu-yuan Nie and Xin-qian Wu. A historical study about the developing process of the classical linear time series models. In *Proceedings of The Eighth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA), 2013*, pages 425–433. Springer, 2013.

[58] Ivan Nourdin and TT Diu Tran. Statistical inference for Vasicek-type model driven by Hermite processes. *Stochastic Processes and their Applications*, 129(10):3774–3791, 2019.

[59] David Nualart. *The Malliavin Calculus and Related Topics*. Springer-Verlag Berlin-Heidelberg, 2006.

[60] Maurice B Priestley. *Spectral Analysis and Time Series: Probability and Mathematical Statistics*. Academic Press, 1981.

[61] Philip Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag Berlin Heidelberg, 2005.

[62] Sheldon M Ross. *Stochastic Processes*, volume 2. John Wiley & Sons, 1996.

[63] Walter Rudin. *Principles of Mathematical Analysis*, volume 3. McGraw-Hill, 1964.

[64] Gennady Samorodnitsky and Murad S Taqqu. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman & Hall, 1994.

[65] Gideon Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978.

[66] Eugen Slutzky. The summation of random causes as the source of cyclic processes. *Econometrica*, 5(2):105–146, 1937.

[67] Marian Smoluchowski. Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Annalen der Physik*, 326(14):756–780, 1906.

[68] Tommi Sottinen and Lauri Viitasaari. Parameter estimation for the Langevin equation with stationary-increment Gaussian noise. *Statistical Inference for Stochastic Processes*, 21:569–601, 2018.

[69] Katsuto Tanaka. Maximum likelihood estimation for the non-ergodic fractional Ornstein-Uhlenbeck process. *Statistical Inference for Stochastic Processes*, 18(3):315–332, 2015.

[70] Thorvald N Thiele. *Sur la compensation de quelques erreurs quasi-systématiques par la méthode des moindres carrés*. CA Reitzel, 1880.

[71] Ruey S Tsay and George C Tiao. Use of canonical analysis in time series model identification. *Biometrika*, 72(2):299–315, 1985.

[72] Aad W Van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998.

[73] Lauri Viitasaari. Representation of stationary and stationary increment processes via Langevin equation and self-similar processes. *Statistics & Probability Letters*, 115:45–53, 2016.

[74] Gilbert T Walker. On periodicity in series of related terms. *Proceedings of the Royal Society of London. Series A*, 131(818):518–532, 1931.

[75] Norbert Wiener. Differential-space. *Journal of Mathematics and Physics*, 2(1-4):131–174, 1923.

[76] Norbert Wiener. Generalized harmonic analysis. *Acta Mathematica*, 55:117–258, 1930.

[77] Herman Wold. *A study in the analysis of stationary time series*. PhD thesis, Almqvist & Wiksell, 1938.

[78] G Udny Yule. On a method of investigating periodicities disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A*, 226(636-646):267–298, 1927.

Aalto-DD 209/2020

ISBN 978-952-64-0182-9 (printed)
ISBN 978-952-64-0183-6 (pdf)
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Aalto University
School of Science
Department of Mathematics and Systems Analysis
www.aalto.fi

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

DOCTORAL
DISSERTATIONS