Whittle estimation and the functional integrated periodogram for MCARMA processes

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der KIT-Fakultät für Mathematik des
Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

Celeste Cecile Mayer
aus Karlsruhe

Tag der mündlichen Prüfung: 07.07.2021

Referent: Prof. Dr. Vicky Fasen-Hartmann

Korreferent: Prof. Dr. Claudia Klüppelberg

To my grandmother, who never got to know how much I love her.

Acknowledgments

First of all, I want to express my deepest gratitude towards my advisor Prof. Dr. Vicky Fasen-Hartmann. I will be forever thankful for all the time she took to help me. Without our discussions, her various suggestions and her investment, this thesis would surely not exist. I truly can not imagine a more engaged supervisor.

I am also grateful for Prof. Dr. Claudia Klüppelberg, who I once met in Munich and who agreed to act as a second referee.

Moreover, I want to thank all the people who shared my joy in good times and helped me to overcome the difficulties of the last years. In particular, I wholeheartedly thank my natural family Ingrid, Bernd and Zoe and my chosen family Johannes for all their love and assistance.

I also want to thank my colleagues and my former colleagues at the Karlsruhe Institute of Technology and all my friends for many happy moments and emotional support.

Abstract

In many real-world applications data is generated by continuous-time processes but due to inadequate measurements can only be observed on a discrete-time grid. In this thesis, we consider the class of equidistantly sampled multivariate continuous-time ARMA (MCARMA) processes. To obtain most flexibility, the driving process is assumed to be a general centered Lévy process.

In the first part, we investigate parameter estimation procedures. We start with the asymptotic behavior of the Whittle estimator when the driving process has at least existing second moments. The Whittle estimator is based on a frequency domain approach. Namely, it is the minimizing argument of the Whittle function which measures the distance between the periodogram and its theoretical counterpart, the spectral density, corresponding to a fixed parameter. Under some identifiability conditions, we obtain strong consistency and asymptotic normality. We then introduce an adjusted version of this estimator for a univariate setting. Thereby, we adapt the Whittle function in a way which makes it independent of the variance parameter of the driving process. This step is motivated by the desire to find an estimation procedure for processes with non-existent second moments. Consequently, we investigate the minimizing argument of this adjusted function in two settings. In the first one, light-tailed processes are again in our focus. As before, we obtain strong consistency and asymptotic normality. Subsequently, we also consider the estimator for processes without existing second moments, namely, for the class of symmetric α-stable CARMA processes. Unfortunately, consistency can only be derived when the underlying process is an Ornstein-Uhlenbeck (CAR(1)) process. Actually, we also give processes for which consistency does not hold.

In the second part, we return to the class of light-tailed sampled MCARMA processes and investigate the function-indexed normalized integrated periodogram. Under different sets of conditions concerning the driving process and the set of index functions, we prove a functional central limit theorem. We also consider several applications. In particular, we derive the asymptotic normality of some spectral goodness-of-fit test statistics. Popular examples are the Grenander-Rosenblatt and the Cramér-von Mises statistic. Based on

lemmata of the previous chapters, the asymptotic normality of the Whittle estimator can also be concluded from the functional central limit theorem.

Finally, a simulation study demonstrates the behavior of the estimators as well as the behavior of the goodness-of-fit test statistics for finite sample sizes.

Prior Publications

Parts of all chapters, but mainly of Chapter 3, 4 and 6, are direct quotes from the prior publications:

Fasen-Hartmann and Mayer (2021+)
Fasen-Hartmann and Mayer (2020)

Contents

Chapter 1

Introduction

Autoregressive moving average (ARMA) processes are doubtless the most prominent process class in the field of time series analysis and are well studied, see for example Brockwell and Davis (1991) or Box et al. (2015). Their popularity might be traced back to Wolds Theorem which states that any univariate second-order stationary time series can be represented as the sum of a deterministic process and an MA(∞) process. Nowadays, technical advances enable us to gather and access data more easily and therefore, big data sets and multivariate data come more and more into focus. Consequently, the interaction between components of the data gets significant interest. To model these relations realistically, it is essential to use a framework which includes dependencies between different variables. Generalizing the ARMA class to multivariate processes yields the class of vector ARMA (VARMA) processes. An m-dimensional VARMA(p, q) process $(Z_n)_{n\in\mathbb{N}}$ is the solution of a difference equation of the form

$$P^{(D)}(\mathfrak{B})Z_n = Q^{(D)}(\mathfrak{B})e_n, \quad n \in \mathbb{N}, \tag{1.1}$$

where $\mathfrak{B}$ is the backshift operator $\mathfrak{B}Z_n = Z_{n-1}$, $P^{(D)}$ and $Q^{(D)}$ are $m \times m$-dimensional and $m \times d$-dimensional polynomials of order p and q, respectively, and $(e_n)_{n\in\mathbb{Z}}$ is a d-dimensional white noise, see, e.g., the monographs of Brockwell and Davis (1991) and Lütkepohl (2005) for more details.

However, in various practical applications, the use of a time-continuous framework is favorable. Even though data is often only gathered at discrete-time points, most underlying processes are time-continuous in real-world situations. In the fields of physics and finance, many applications already use continuous-time settings. For example, the Black-Scholes

option-pricing model (Hull and White (1987)) is based on stochastic differential equations and one of the most popular foundations of modern finance. In mechanics, continuous-time processes are in the focus for many years, see Fowler (1929), and Bergstrom (1990) presents numerous econometric applications. The continuous-time analogues to (V)ARMA processes are the (multivariate) continuous-time ARMA ((M)CARMA) processes. They date back to 1944, where Doob firstly mentioned the univariate Gaussian continuous-time ARMA processes, see Doob (1944). In 2001, Brockwell (2001b) extended the definition to Lévy-driven CARMA processes. This generalization allows to model processes with various marginal distributions. For example, assuming the driving process to be α-stable yields the class of α-stable CARMA processes. Stable CARMA processes were already applied for future pricing in electricity markets (Benth et al. (2014)) and signal extraction (McElroy (2013)). We return to this class of processes later. Applications of CARMA processes with existing second moments in finance were discussed in Brockwell (2009) and, in particular, Andresen et al. (2014) investigated the CARMA interest rate model. Later, Marquardt and Stelzer (2007) generalized the class of Lévy-driven CARMA processes to a multivariate setting. Obviously, choosing a process in this broader class might yield more accuracy in many applications. However, they are also used as a starting point for constructing an even richer process class, for example the class of continuous-time threshold ARMA processes (Brockwell (2001a)) or the class of cointegrated MCARMA processes (Fasen-Hartmann and Scholz (2020)). Just recently, the class of MCARMA processes was further generalized. Namely, Brockwell and Matsuda (2017) considered the class of isotropic CARMA random fields whereas Pham (2020) introduced the class of causal Lévy-driven CARMA random fields. Furthermore, (M)CARMA processes were already investigated for high-frequency data, see Brockwell et al. (2013), Ferrazzano and Fuchs (2013) and Fasen and Fuchs (2013b) for light-tailed CARMA processes and Fasen and Fuchs (2013a) for stable CARMA processes, or irregular spaced data (Jones (1981)).
However, in this thesis, we solely consider Lévy-driven MCARMA processes which are sampled equidistantly and at a low frequency. We therefore introduce the two-sided d-dimensional Lévy process $L = (L_t)_{t\in\mathbb{R}}$ which is constructed by a sum of two independent and identically distributed (i.i.d.) one-sided Lévy processes $L^{(1)}$ and $L^{(2)}$ via $L_t = L_t^{(1)}\mathbf{1}_{\{t\geq 0\}} - \lim_{s\uparrow -t} L_s^{(2)}\mathbf{1}_{\{t<0\}}$. A one-sided Lévy process is a stochastic process with stationary, independent increments, continuous in probability with càdlàg sample paths and satisfying $L_0 = 0$ almost surely. Then, by Marquardt and Stelzer (2007), a MCARMA(p, q) process $(Y_t)_{t\in\mathbb{R}}$ is the stationary solution of a continuous-time state space model (A, B, C, L) of the form

$$dX_t = AX_tdt + BdL_t, \quad Y_t = CX_t, \quad t \in \mathbb{R},$$

where

$$A = \begin{pmatrix} 0 & I_m & 0 & \dots & 0 \\ \cdots & & & & I_m \\ -A_p & -A_{p-1} & \dots & \dots & -A_1 \end{pmatrix}, \; B = (\beta_1^\top \dots \beta_p^\top)^\top, \; C = (I_m \; 0 \; \dots \; 0)\,,$$

$$\beta_{p-j} = -\mathbf{1}_{\{0,\dots,q\}}(j) \left[\sum_{i=1}^{p-j-1} A_i \beta_{p-j-i} - B_{q-j} \right], \quad j = 0, \dots, p-1,$$

for some matrices $A_1, \dots, A_p \in \mathbb{R}^{m \times m}, B_0, \dots, B_q \in \mathbb{R}^{m \times d}$. Strictly speaking, Marquardt and Stelzer (2007) only investigated processes where the dimensions m and d are the same. However, their results can easily be carried out to this general setting. Here and in the following, I_m denotes the m-dimensional identity matrix. The stationary solution of this model also has the differential equation representation

$$\begin{aligned} P(D)Y_t &= Q(D)DL_t, \quad D = \frac{d}{dt}, \\ P(z) &:= I_m z^p + A_1 z^{p-1} + \ldots + A_p, \\ Q(z) &:= B_0 z^q + \ldots + B_q. \end{aligned}$$

Note that p and q are the degrees of the so-called autoregressive and moving average polynomial, respectively. However, this representation is purely formal since the paths of a Lévy process are not differentiable in general. Accordingly, it only serves us to depict the similarities between MCARMA processes and their discrete-time counterparts. Sampling an MCARMA process equidistantly with a fixed distance $\Delta > 0$ yields a discrete-time process which has an MA(∞) representation as well as a discrete-time state space representation, see Schlemm and Stelzer (2012a). Since the white noise process in the MA(∞) representation of the sampled process is not independent and identically distributed in general, we only obtain a weak MA(∞) representation. The properties of the sampled process are elaborated in more detail in Chapter 2. For the sake of completeness, note that Fasen-Hartmann and Scholz (2021) also found a weak VARMA($p, p-1$) representation of the sampled MCARMA process, recently.
In the introduced setting, we consider several spectral applications which have diverse purposes in time series analysis. In general, frequency domain approaches are used in the context of model fitting and prediction. Furthermore, in physical applications, they can be interpreted directly and might therefore be a better option than a time domain procedure. In this thesis, the main focus lies on model fitting.
We start with an investigation of the Whittle estimator, which was first presented by Peter Whittle (1951) and later used in various different models to estimate some parameter in a given setting. Therefore, the class of MCARMA processes has to be parameterized. In our context, the parameters p and q are fixed and given. Obviously, this assumption is quite strong since the degrees of the polynomials are not known in general applications. To

determine the parameters p and q beforehand, information criteria could be applied. In the context of sampled MCARMA processes Fasen and Kimmig (2017) already investigated some information criteria. We now consider a parameter set $\Theta \subset \mathbb{R}^r$ and let $Y(\vartheta) = (A(\vartheta), B(\vartheta), C(\vartheta), L(\vartheta))$ be the MCARMA(p, q) process corresponding to the parameter $\vartheta \in \Theta$. This means that the components of the matrices $A_1, \ldots, A_p, B_0, \ldots, B_q$ are potentially unknown as well as the distribution and, in particular, the covariance Σ_L of the driving process. The aim is to estimate the true parameter $\vartheta_0 \in \Theta$ based on a sample $(Y_\Delta(\vartheta_0), \ldots Y_{n\Delta}(\vartheta_0))$ of the sampled process. The Whittle estimator is then defined as a minimizing argument of the so-called Whittle function which exact definition depends on the particular setting. However, in each scenario, the Whittle function at a parameter ϑ is based on some kind of distance between the periodogram of the sampled process and the spectral density corresponding to this ϑ. It therefore measures the distance between the Fourier transform of the sample autocovariance function and the Fourier transform of the theoretical autocovariance function of a process with the parameter ϑ. In our setting, the Whittle function W_n is specifically defined as

$$W_n(\vartheta) = \frac{1}{2n} \sum_{j=-n+1}^{n} \Big[\operatorname{tr}\left(f_Y^{(\Delta)}(\omega_j, \vartheta)^{-1} I_{n,Y}(\omega_j) \right) + \log\left(\det\left(f_Y^{(\Delta)}(\omega_j, \vartheta) \right) \right) \Big], \quad \vartheta \in \Theta, \tag{1.2}$$

where $I_{n,Y}$ is the periodogram of the sample $(Y_\Delta(\vartheta_0), \ldots Y_{n\Delta}(\vartheta_0))$, $f_Y^{(\Delta)}(\cdot, \vartheta)$ is the spectral density of the sampled process corresponding to $\vartheta \in \Theta$ and $\omega_j = \frac{\pi j}{n}$, $j = -n+1, \ldots, n$, are the Fourier frequencies. Note that it is possible to exchange the $\log(\det(f_Y^{(\Delta)}(\cdot, \vartheta)))$ terms for $\log(\det(V^{(\Delta)}(\vartheta)))$ without changing the minimizing argument. Thereby, $V^{(\Delta)}(\vartheta)$ is the covariance matrix of the one-step linear prediction error which generally depends on the parameters of the autoregressive and moving average polynomials. In contrast, in case of Whittle estimation for VARMA processes, Dunsmuir and Hannan (1976) used a Whittle function in which the $\log(\det(f_Y^{(\Delta)}(\cdot, \vartheta)))$ terms were replaced by $\log(\det(\Sigma(\vartheta)))$ where $\Sigma(\vartheta)$ denotes the covariance matrix of the white noise process corresponding to ϑ. Obviously, this term does not depend on the moving average and autoregressive polynomials of the underlying process. Consequently, it differs from our function which explains why their ideas can not be directly carried out to our setting. In some articles, the asymptotic behavior of the Whittle estimator is a direct conclusion of a functional central limit theorem, see for example Dahlhaus (1988) or Bardet et al. (2008) for such a result in the context of weakly dependent time series. Since the sampled MCARMA processes have a weak VARMA representation, it seems fruitful to apply a functional central limit theorem for the integrated periodogram of VARMA processes. For these processes, Bardet et al. (2008), Dahlhaus (1988) and Mikosch and Norvaiša (1997) proved several versions of spectral functional central limit theorems. Unfortunately, their results all assume the white noise process to have quite strong properties which are not given in the sampled MCARMA setting. However, in Chapter 4, we prove a functional central limit theorem for the integrated periodogram in the context of the sampled MCARMA process and will then broadly outline the asymptotic normality of the Whittle estimator as a conclusion of that result. Having said this, the conclusion heavily relies on the properties which we treat beforehand. For example, for most statistical procedures, it is necessary to guarantee identifiability. In our case, this means it has to be theoretically possible to learn the true value after observing an infinite number of data points. In the context of sampled MCARMA processes, identifiability contains different areas. On the one hand, we have to prevent that different parametrizations yield the same process. On the other hand, conditions have to be imposed to eliminate the so-called aliasing effect. Aliasing occurs when two distinguishable processes become indistinguishable when sampled. In Figure 1.1, the effect is illustrated. In this example, observing the process with a smaller sampling distance would prevent the aliasing. In the same way, we bypass the effect in our setting by assuming a small enough sampling distance. More details concerning the aliasing can be found in Hansen and Sargent (1983).

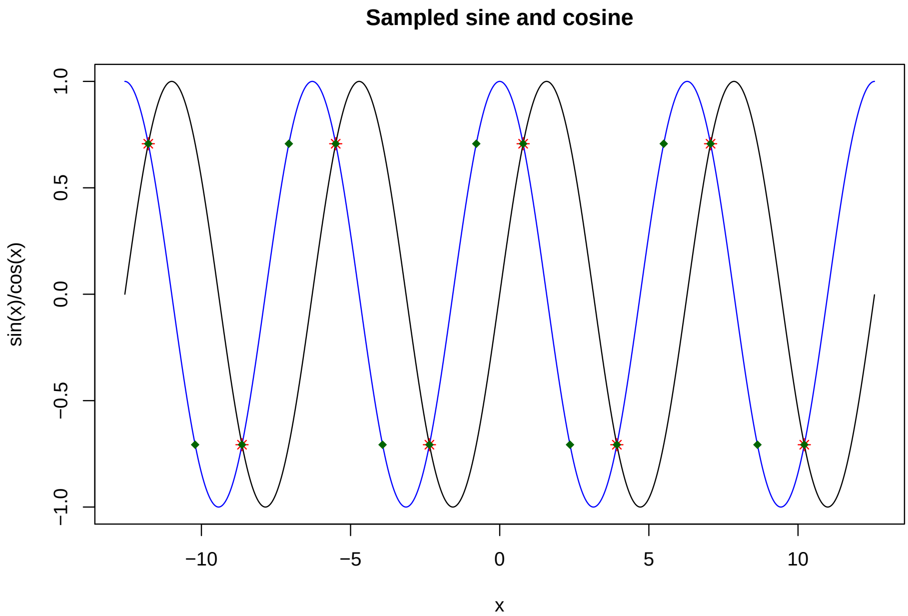

Figure 1.1.: Sine and cosine function. When sampled at $x = k + 1/4\pi$, $k \in \mathbb{Z}$, the functions can not be distinguished (red stars). When sampled at a higher frequency, it is possible to identify the original function (green rectangles).

To tackle the problem that different parametrizations might yield the same continuous-time process, we repeat the findings of Schlemm and Stelzer (2012b) who considered Echelon state space realizations. Schlemm and Stelzer (2012b) dealt with identification of MCARMA processes for the same reasons as we do: they aimed to derive an estimation procedure for the sampled MCARMA processes with existing second moments. In their article, they introduce a quasi maximum likelihood estimator and prove its strong consistency and asymptotic normality. In contrast to us, they have to assume existing $(4+\delta)$th moments of the driving Lévy process for the asymptotic normality to hold. In case of the Whittle estimator, we see in Chapter 3 that it suffices to assume the slightly weaker condition of existing 4th moments. As an alternative for univariate CARMA processes, we introduce the adjusted Whittle estimator in Chapter 4 which is a minimizing argument of the adjusted Whittle function. Essentially, this function results from (1.2) when we omit the $\log(\det(f_Y^{(\Delta)}(\cdot,\vartheta)))$ terms and multiply the remaining function with the covariance matrix $V^{(\Delta)}(\vartheta)$ of the one step linear prediction error. The motivation for doing so is that the resulting function and therefore also the minimizing argument, is independent of the covariance matrix of the driving Lévy process. Our hope was to thereby construct an estimator for symmetric α-stable CARMA processes which do not have existing second moments. However, even though we can show strong consistency and asymptotic normality in a slightly adapted light-tailed setting, in general, the same does not hold for the α-stable setting. We prove that the adjusted Whittle estimator is consistent for the class of univariate Ornstein-Uhlenbeck (CAR(1)) processes, but since the adjusted Whittle function converges to an α-stable random variable which in general does not have a minimum in the true parameter, we can easily construct an example for which consistency does not hold. Therefore, the sampled MCARMA case is fundamentally different to the case of α-stable ARMA processes. For that case, Mikosch et al. (1995) investigated the Whittle estimator and obtained a convergence result. On top, in their central limit theorem the convergence rate is $n^{1/\alpha}$ which is even better than in the light-tailed case. Consequently, the VARMA results can not be carried out to the sampled MCARMA case. We trace this back to the fact that the sampled MCARMA processes have indeed a weak VARMA representation but with a white noise which is not i.i.d. in general. Only when considering the class of CAR(1) processes, the sampled processes have a strong VARMA representation and in this case we also have the desired convergence. Consequently, we indeed constructed a suitable estimator for parameter estimation of Ornstein-Uhlenbeck processes, but we did not find one for the general class of CARMA processes. While the problem of parameter estimation for α-stable Ornstein-Uhlenbeck processes was already considered in some articles, see Fasen and Fuchs (2013b), Hu and Long (2007), Hu and Long (2009), Ljungdahl and Podolskij (2020), Zhang and Zhang (2013), parameter estimation procedures for the general α-stable models are not investigated in detail yet. Only García et al. (2011) introduced an indirect quasi maximum likelihood estimation procedure, but they do not present a mathematical analysis. Even though their simulation study suggests that their method has desirable asymptotic properties, we assume that their estimator is not consistent as well. In a simulation study, we consider the setting of their simulation study as well as an α-stable setting in which the adjusted Whitte estimator does not work. In both situations, the adjusted Whittle estimator and the estimator of García et al. (2011) behave similarly. This means that the adjusted Whittle estimator also seems to converge in the setting introduced in García et al. (2011) and that we can not observe convergence of their estimator in our setting. In addition, we also present a simulation study for the light-tailed setting. We therefore simulate different CARMA processes and study the behavior of the Whittle estimator and the adjusted Whittle estimator. Additionally, we compare their behaviors to those of the quasi maximum likelihood estimator of Schlemm and Stelzer (2012b) which seems to be the only alternative to our procedures. As we will see, their estimator seems to behave similar to the Whittle estimator and much better than the adjusted Whittle estimator. Accordingly, we also do not recommend the adjusted Whittle estimator in the light-tailed setting. However, the quasi-maximum likelihood estimator and the Whittle estimator are quite good. At this point, we want to mention, that even though no estimator should be preferred based on the simulation studies, the Whittle estimator has two advantages. As said, the moment conditions concerning the driving process are slightly weaker. Furthermore, an analytical representation of the covariance matrix of the limit distribution in the asymptotic normality result is only known in case of the Whittle estimator. This covariance matrix can be used to derive confidence bands, for example.

Finally, we also investigate the function-indexed integrated periodogram separately. A functional central limit theorem for the function-indexed periodogram is of interest, since one can easily derive the asymptotic behavior of various statistical applications. As mentioned earlier, the asymptotic normality of the Whittle estimator can be concluded as well as the asymptotic behavior of some spectral goodness-of-fit test statistics. Similar results are already derived for many settings. For example, Dahlhaus (1988) proved a functional central limit theorem for the integrated periodogram for a general class of multivariate time series with appropriately bounded higher order cumulant spectra of the white noise process and a weak entropy condition. Under the same entropy condition, Mikosch and Norvaiša (1997) showed such a theorem for univariate linear processes with finite fourth moments, whereas Can et al. (2010) proved weak convergence under a stronger condition in the context of α-stable linear processes. Under a fundamentally different structured function space condition, Bardet et al. (2008) derived a functional central limit theorem for a broad class of weakly dependent time series. In light of these articles, we investigate the integrated periodogram under similar conditions in the context of sampled MCARMA processes. Consequently, we obtain a central limit theorem under three different sets of conditions. Namely, in a first setting, we assume that the function space is of a given structure similar to that of Can et al. (2010) whereas the driving Lévy process has to have just existing 4th moments. Even though one can not say if the function set condition is stronger than the one of our second setting, the set of indicator functions does not satisfy the condition of the first setting. This set is of special interest since it yields the asymptotic normality of the spectral goodness-of-fit test statistics. In the second setting, we impose a covering number condition similar to that of Bardet et al. (2008) and also assume existing fourth moments of the driving process. The set of indicator functions satisfies this covering number condition. However, the scenario does not enable us to derive the asymptotic normality of the Whittle estimator. Finally, we suppose an entropy condition in the third setting which is strictly weaker than the condition of the second setting but we therefore have to assume that all moments of the driving process exist and are appropriately bounded. Note that this setting is motivated by Dahlhaus (1988) and Dahlhaus and Polonik (2009). It allows us to directly conclude the asymptotic normality of the Whittle estimator as well as the asymptotic normality of the spectral goodness-of-fit test statistics. Therefore, we also give the limit behavior of some popular spectral test statistics. As we will see, the limit statistics are functionals of a Gaussian process. For Brownian motion driven MCARMA processes, they correspond to the limits of the Gaussian ARMA case, see Priestley (1981), but they differ from the limit statistic of the heavy-tailed ARMA process class which was investigated in Klüppelberg and Mikosch (1996). Lastly, we also tackle the question how the test statistics perform in practical applications. Therefore, we explicitly consider the Grenander-Rosenblatt statistic and the Cramér-von Mises statistic for finite sample sizes in a simulation study. Thereby, we first compare the behavior of the quantiles of the empirical processes for different sample sizes with the theoretical quantiles of the limit process. Since the empirical quantiles are close to the theoretical ones even for small and moderate sample sizes, we also perform some testing. When simulating the processes under various misspecifications, the corresponding tests reject quite often. We therefore conclude that they both are suited for practical applications. Furthermore, since the theoretical limit process depends on the distribution of the underlying Lévy process which is not known in general, we also do some Bootstrap testing. In a small study, the results are promising.

Outline of the thesis

We want to give a short overview about what follows. In Chapter 2 we introduce the fundamentals of this thesis. Namely, in Section 2.1, we first recall the definition of general Lévy-driven univariate CARMA processes of Brockwell (2001b) and we then introduce the class of MCARMA processes of Marquardt and Stelzer (2007) in Section 2.2. Since our multivariate extension does not cover α-stable processes, it is necessary to consider the general univariate class separately. In Section 2.3 we state different results concerning the representation of the sampled process. Thereby, a distinction between sampled processes with existing second moments and symmetric α-stable processes is done. In the same way, we treat the following subsection. Obviously, when we talk about second order properties, it is necessary to assume the existence of the second moments. However, in Section 2.4, we first consider MCARMA processes with existing second moments and investigate the behavior of the sample autocovariance function of the sampled processes and then deal with the appropriately normalized sample autocovariance function for univariate α-stable processes. We thereby also introduce the periodogram and its theoretical counterpart, the spectral density which are the base of the later treated estimators and the integrated periodogram.

In Chapter 3, we then consider the Whittle estimation procedure for sampled multivariate Lévy-driven CARMA processes with existing second moments. We start with a brief introduction of the estimator in Section 3.1. To provide the strong consistency and the asymptotic normality of this estimator, we address an identifiability problem. Schlemm and Stelzer (2012b) already tackled this topic so that we just revisit the broad connections in Section 3.2. The consequential setting is then given in Section 3.3 and we directly state the main results concerning the Whittle estimator in Section 3.4. Thereby, a simplification of the covariance matrix of the limit process of the asymptotic normality result is presented for the Brownian motion driven MCARMA process and a comparison to the VARMA setting is made. We conclude this chapter by the proofs of the strong consistency and the asymptotic normality in Section 3.5. Note that a central limit theorem for the integrated periodogram is a fundamental part of the proof of the asymptotic normality.

In Chapter 4, we leave the multivariate setting and solely consider univariate CARMA processes. An adaption of the procedure of Chapter 3, yields an estimator which is independent of the variance parameter of the driving process. We motivate this adjustment in Section 4.1. In Section 4.2, we investigate the resulting estimator in the case that the driving process has an existing positive variance. We thereby introduce an appropriate setting and derive strong consistency and asymptotic normality of the adjusted Whittle estimator. Note that the setting which is considered here is similar to the one of Chapter 3. We then directly see that the variance of the limit process is at least as big as the limit variance of the Whittle estimator for univariate CAR(1) processes. Looking back, this could be a first indication that our adjusted Whittle estimator is not as promising as the original one. However, the main objective of constructing this adjusted version was to obtain an estimator which is suitable for estimation in a light-tailed and a heavy-tailed setting. Therefore, as a second step, we also consider the procedure for symmetric α-stable CARMA processes in Section 4.3. We derive uniform convergence of the adjusted Whittle function to a random variable and conclude that this random variable only has a unique minimum at the true parameter with certainty when the underlying process is an Ornstein-Uhlenbeck (CAR(1)) process. Since it is necessary for the consistency that the limit of the adjusted Whittle function only takes its minimum at the true parameter, we deduce that the adjusted Whittle estimator is not consistent in general. Furthermore, we explicitly give examples for which the estimator is not consistent.

In Chapter 5, we return to multivariate processes with existing second moments and focus on the integrated periodogram. In contrast to Section 3.5 where we index the integrated periodogram with the class of parameterized inverted spectral densities, we now allow more general index functions. In Section 5.1 we state three different sets of conditions and then directly give a functional central limit theorem in the resulting settings. Generally, the limit process is a centered Gaussian process. We determine its parameters and also notice that in the case of a Brownian motion driven MCARMA process, the limit process has a simplified representation. As a conclusion of the functional central limit theorem, we directly obtain a major part of the proof of the asymptotic normality of the Whittle estimator of Chapter 3. Therefore, we quickly sketch this approach in Section 5.2. Additionally, we conclude the asymptotic normality of various spectral goodness-of-fit test statistics. Finally, we prove the functional central limit theorem in Section 5.3. We start to prove an analogue theorem for the white noise of the sampled process. In our corresponding metric spaces, weak convergence can be shown by proving weak convergence of the finite-dimensional distributions and tightness of the empirical process of the white noise. A universal proof for the weak convergence of the finite-dimensional distributions is made, whereas the tightness has to be shown for each setting separately. Finally, we bind the error which results from approximating the original process by the process corresponding to the white noise and conclude the proof of the functional central limit theorem for the general class of MCARMA processes.

The suitability of the theoretical results of Chapter 3-Chapter 5 for practical applications is the topic of Chapter 6. Here, practical means that we investigate how the estimators and the goodness-of-fit test statistics behave for finite sample sizes in different settings. Therefore, the chapter is divided in two parts. In the first part, Section 6.1, we compare the estimation procedures. Since our initial simulation study is concerned with parameter estimation in the multivariate setting, we start with the introduction of the Echelon state space representation. Thereby, a suitable parametrization for avoiding redundancies in the MCARMA model is derived. Subsequently, we consider the quasi maximum likelihood estimator of Schlemm and Stelzer (2012b) which serves as the only alternative procedure for parameter estimation in the general multivariate setting. Finally, we are able to present our simulation results for the multivariate setting, where the performances of the quasi maximum likelihood estimator and the Whittle estimator are compared as well as for the univariate setting. In the light-tailed univariate setting, the adjusted Whittle estimator is also applied. As expected, we see that it often yields the worst results. Concluding this section, a simulation study for the α-stable case is presented as well. As an alternative procedure to the adjusted Whittle estimator, we shortly introduce the estimator of García et al. (2011). Matching the theory, the simulations indicate consistency for the Ornstein-Uhlenbeck process but there are settings in which neither of the estimators seems to converge. For the adjusted Whittle estimator, this property also matches the theoretical results obtained in Section 4.3. Furthermore, it allows us to assume the same for the estimator of García et al. (2011).

In the second part of this chapter, we also investigate the finite sample behavior of the Grenander-Rosenblatt statistic and the Cramér-von Mises test statistic in a correctly specified setting. Additionally, we do some testing under fixed alternatives. Since the limit processes of the test statistics depend on the driving process, see Section 5.2, we also apply a Bootstrap procedure.

Finally, Chapter 7 concludes the main part of the thesis by giving a short summary and an outlook to open problems. Appendix A contains some analytical foundations. Namely, all the basic Fourier analytical results are given as well as two results concerning the approximation of integrals by sums.

CHAPTER 2

Multivariate Lévy-driven CARMA processes

In this chapter we introduce the fundamental process classes of this thesis. In particular, we define Lévy-driven multivariate CARMA processes with existing second moments as well as univariate symmetric Lévy-driven α-stable CARMA processes. We start with a brief introduction of Lévy processes whose increments can be interpreted as the continuous-time analogue of the white noise process in discrete-time settings. For more details concerning Lévy processes, we refer to Applebaum (2004) and Sato (1999).

Definition 2.1 (Lévy process, see Sato (1999))**.**
A one-sided d-dimensional Lévy process $(L_t)_{t\geq 0}$ is a stochastic process with stationary, independent increments, continuous in probability with càdlàg sample paths satisfying $L_0 = 0$ almost surely.

Due to the Lévy-Khinchin formula, the characteristic function of a Lévy process $(L_t)_{t\geq 0}$ has the representation

$$\mathbb{E}\left[e^{ix^\top L_t}\right] = e^{t\Psi(x)}, \quad x \in \mathbb{R}^d,$$

where

$$\Psi(x) = -\frac{1}{2}\langle x, \Sigma x\rangle + i\langle \gamma, x\rangle + \int_{\mathbb{R}^d} \left(e^{i\langle x,t\rangle} - 1 - i\langle x, t\rangle \mathbb{1}_{|t|\leq 1}\right) \nu(dx)$$

with $\Sigma \in \mathbb{R}^{d\times d}$ symmetric and non-negative definite, $\gamma \in \mathbb{R}^d$ and ν a measure on $\mathbb{R}^d$ with

$$\nu(\{0\}) = 0 \text{ and } \int_{\mathbb{R}^d} \min\{\|x\|^2, 1\} d\nu(x) < \infty.$$

Accordingly, we define the characteristic triplet (Σ, ν, γ). The naming is appropriate since it uniquely characterizes a Lévy process and therefore, we identify any Lévy process by its characteristic triplet, see Theorem 8.1 of Sato (1999). Note that, by Sato (1999), Corollary

25.8,

$$\mathbb{E}[\|L_t\|^k] < \infty \ \forall t \geq 0 \iff \int_{\|x\|\geq 1} \|x\|^k \nu(dx) < \infty.$$

Since we aim to define processes indexed by the set of real numbers, we have to construct the driving process on the whole real line as well. A two-sided Lévy process $L := (L_t)_{t\in\mathbb{R}}$ can be built from two independent one-sided Lévy processes $L^{(1)}$ and $L^{(2)}$ through

$$L_t = L_t^{(1)} \mathbf{1}_{\{t\geq 0\}} - \lim_{s\uparrow -t} L_s^{(2)} \mathbf{1}_{\{t<0\}}.$$

In the following, we only deal with two-sided d-dimensional Lévy processes.

2.1. Univariate CARMA processes

In this section, we mainly consider univariate processes, which are driven by a univariate Lévy process, i.e., $d = 1$. Following the structure of ARMA processes, we define for $p > q$ and $a_1, \ldots, a_p, c_0, \ldots, c_q \in \mathbb{R}, a_p \neq 0, c_j \neq 0$ for some $j \in \{0, \ldots, q\}$, the autoregressive polynomial a and the moving average polynomial c as

$$a(z) := z^p + a_1 z^{p-1} + \ldots + a_p \qquad \text{and} \qquad c(z) := c_0 z^q + c_1 z^{q-1} + \ldots + c_q,$$

respectively. The CARMA(p, q) process $(Y_t)_{t\in\mathbb{R}}$ can now be interpreted as the stationary solution of the pth order differential equation

$$a(D)Y_t = c(D)DL_t, \quad t \geq 0, \tag{2.1}$$

where $D = \frac{d}{dt}$ is the differential operator. But since the paths of a Lévy process are not differentiable in general, this definition is purely formal and we need an applicable one. Introductory, we consider the class of general (N-dimensional) linear state space models. Since we return to this definition when we introduce multivariate processes, we now allow $d \neq 1$ and $m \neq 1$.

Definition 2.2 (continuous-time linear state space model, see Schlemm and Stelzer (2012b))**.**
Let L be a d-dimensional Lévy process. For matrices $A \in \mathbb{R}^{N\times N}$, $B \in \mathbb{R}^{N\times d}$ and $C \in \mathbb{R}^{m\times N}$ a continuous-time linear state space model (A, B, C, L) is defined by

$$dX_t = AX_t dt + BdL_t, \quad Y_t = CX_t, \quad t \in \mathbb{R}. \tag{2.2}$$

The processes $(X_t)_{t\in\mathbb{R}}$ and $(Y_t)_{t\in\mathbb{R}}$ are called state- and output process, respectively.

The univariate CARMA(p, q) process corresponding to (2.1) can now be defined by its controller canonical linear state space representation. Therefore, let

$$A := \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 1 \\ -a_p & -a_{p-1} & \dots & \dots & -a_1 \end{pmatrix} \in \mathbb{R}^{p\times p}, \tag{2.3a}$$

$$B := e_p := (0, \dots, 0, 1)^\top \in \mathbb{R}^p$$
$$C := c^\top := (c_q,\, c_{q-1},\, \dots,\, c_0, 0, \dots, 0) \in \mathbb{R}^{1\times p}.$$

Finally, let $(X_t)_{t\in\mathbb{R}}$ be a strictly stationary solution of the first equation in (2.2). Then, the output process $(Y_t)_{t\in\mathbb{R}}$ of Definition Definition 2.2 is the above CARMA(p, q) process, i.e. Y_t satisfies

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + e_p\,\mathrm{d}L_t, \quad t \in \mathbb{R}, \tag{2.3b}$$

$$Y_t := c^\top X_t, \quad t \in \mathbb{R}. \tag{2.3c}$$

In general, this process exists, if we assume that the eigenvalues of A have nonzero real parts and $\mathbb{E}[\log^+|L_1|] < \infty$, see Theorem 3.3 of Brockwell and Lindner (2009). Here, $\log^+|L_1|$ stands for the positive part of the logarithm of $|L_1|$. Under the stricter condition that the eigenvalues of A have strictly negative real parts, the solution can be assumed to be causal, i.e. Y_t is independent of the σ-algebra generated by $(L_s)_{s>t}$. Note that the eigenvalues of A correspond to the zeroes of a. Therefore, we could alternatively impose an equivalent condition on those zeroes.

Additionally, the process can be seen as a special continuous-time moving average (CMA) process, see Remark 3.2 of Marquardt and Stelzer (2007). A CMA process $(Y_t)_{t\in\mathbb{R}}$ is a process which can be represented as

$$Y_t = \int_{-\infty}^{\infty} g(t-s)dL_s, \quad t \in \mathbb{R},$$

for some $g \in L^2(\mathbb{R})$. Thereby, $L^2(\mathbb{R})$ denotes the space of square-integrable functions on $\mathbb{R}$. In the the context of a centered CARMA process, the state process $(X_t)_{t\in\mathbb{R}}$ has the multivariate Ornstein-Uhlenbeck representation

$$X_t = \int_{-\infty}^{t} e^{A(t-u)}e_p dL_u, \quad t \in \mathbb{R},$$

and accordingly,

$$Y_t = \int_{-\infty}^{\infty} g(t-s)\,dL_s, \quad t \in \mathbb{R}, \quad \text{with} \quad g(t) = c^\top \mathrm{e}^{At} e_p 1_{[0,\infty)}(t). \tag{2.4}$$

Note that in contrast to the discrete-time ARMA and MA processes, the class of CMA processes is richer than the class of CARMA processes.

To now model a CARMA process with non-existent second moments, we choose as driving processes the class of α-stable Lévy processes which we tag with the superscript α. Those processes are Lévy processes with the restriction that the increments $(L_t^{(\alpha)} - L_s^{(\alpha)})_{s<t}$ have an α-stable distribution. Therefore, we introduce α-stable random variables.

Definition 2.3.
A random variable Z is called α-stable distributed, $\alpha \in (0, 2]$, if Z has the characteristic function $\mathbb{E}(\mathrm{e}^{izZ}) = \exp(\varphi_Z(z))$, $z \in \mathbb{R}$, with

$$\varphi_Z(z) = \begin{cases} -\sigma^\alpha |z|^\alpha (1 - i\beta(\operatorname{sign}(z)) \tan\left(\pi \frac{\alpha}{2}\right) + i\mu z, & \text{for } \alpha \neq 1, \\ -\sigma|z|(1 + i\beta(\operatorname{sign}(z)) \log|z| \left(\frac{2}{\pi}\right) + i\mu z, & \text{for } \alpha = 1, \end{cases} \qquad z \in \mathbb{R},$$

and $\beta \in [-1, 1], \sigma > 0$ and $\mu \in \mathbb{R}$. The parameters α and σ are the index of stability and scale parameter, respectively. We write $Z \sim S_\alpha(\sigma, \beta, \mu)$.

For $\alpha = 2$, we obtain a normally distributed random variable with mean μ and variance $2\sigma^2$. So even though the scale parameter differs from the standard deviation in the 2-stable case, we denote both parameters as σ. The context prevents confusion. Note that for $\alpha < 2$ not all moments of Z exist. To be more precise

$$\mathbb{E}|Z|^j < \infty \quad \text{for } 0 < j < \alpha \quad \text{and} \quad \mathbb{E}|Z|^j = \infty \quad \text{for } j \geq \alpha. \tag{2.5}$$

Since we do not restrict ourselves to processes with finite first moments, we assume that the increments are symmetrically distributed. Thereby, we obtain a condition which corresponds to a centeredness condition for processes with existing first moments. Note that for an α-stable random variable symmetry is equivalent to $\mu = 0 = \beta$. In this setting,

$$\lim_{n \to \infty} n\mathbb{P}(|Z| > n^{1/\alpha}) = C_\alpha \sigma^\alpha$$

with

$$C_\alpha = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)}, & \text{if } \alpha \neq 1 \\ \frac{2}{\pi}, & \text{if } \alpha = 1 \end{cases}, \tag{2.6}$$

see Property 1.2.15 of Samorodnitsky and Taqqu (1994), where more details on stable distributions can be found as well. Consequently, the symmetric α-stable Lévy process $L^{(\alpha)}$ satisfies

$$L_t^{(\alpha)} - L_s^{(\alpha)} \sim S_\alpha(\sigma(t-s)^{1/\alpha}, 0, 0), \quad s < t,$$

for some scale parameter $\sigma > 0$ and stability parameter $\alpha \in (0, 2)$.
For more details relating to general CARMA processes, we refer to Brockwell and Lindner (2009) and Brockwell (2001b).

2.2. Multivariate CARMA processes with existing second moments

We now extend the previous definitions to define m-dimensional MCARMA(p,q) processes with finite second moments. Therefore, it is necessary to assume that the covariance matrix Σ_L of the driving Lévy process $L = (\Sigma, \nu, \gamma)$ exists. Equivalently, suppose $\int_{\|x\|\geq 1} \|x\|^2 \nu(dx) < \infty$. Then, by Example 25.11 of Sato (1999), the covariance matrix Σ_L is defined as

$$\Sigma_L = \mathbb{E}[L_1 L_1^\top] = \int_{\|x\|\geq 1} x x^\top \nu(dx) + \Sigma.$$

For this section, we assume

Assumption L1.
$L := (L_t)_{t\in\mathbb{R}}$ is a centered Lévy process with positive definite covariance matrix Σ_L.

Following the ideas of the previous section, we define the MCARMA(p,q) process for $p > q$ via a continuous-time linear state space model.

Definition 2.4 (stationary MCARMA(p,q) process)**.**
Let Assumption $L1$ hold. For integers $p > q$ let further be $A_1, \dots, A_p \in \mathbb{R}^{m\times m}$, $B_0, \dots, B_q \in \mathbb{R}^{m\times d}$. Define

$$\begin{aligned}
A &= \begin{pmatrix} 0 & I_m & 0 & \dots & 0 \\ \cdots & & & & I_m \\ -A_p & -A_{p-1} & \dots & \dots & -A_1 \end{pmatrix} \in \mathbb{R}^{mp\times mp}, \\
B &= (\beta_1^\top \dots \beta_p^\top)^\top \in \mathbb{R}^{mp\times d}, \\
\beta_{p-j} &= -\mathbf{1}_{\{0,\dots,q\}}(j) \left[\sum_{i=1}^{p-j-1} A_i \beta_{p-j-i} - B_{q-j} \right], \quad j = 0, \dots, p-1, \\
C &= (I_m\ 0\ \dots\ 0) \in \mathbb{R}^{m\times mp},
\end{aligned} \tag{2.7}$$

and assume that the eigenvalues of A have non-zero reals parts. Then, the stationary solution $(Y_t)_{t\in\mathbb{R}}$ of the continuous-time linear state space model (2.2) with (A, B, C, L) is called multivariate CARMA(p,q) process (MCARMA(p,q) process).

Note that by Proposition 3.12 of Marquardt and Stelzer (2007), the existence of the second moment of the driving process is inherited by the MCARMA process. Strictly speaking, Marquardt and Stelzer (2007) only considered MCARMA processes where $d = m$. However, their definitions and results can be easily extended to processes whose dimension differs from the dimension of the driving process. As in the univariate setting, the assumption that the eigenvalues of A are a subset of $\mathbb{R}\backslash\{0\} + i\mathbb{R}$, yields that the processes given in Definition 2.4 exist and are strictly stationary, see Proposition 3.26 of Marquardt and Stelzer (2007). In contrast to the univariate case, the existence of the second moment of the Lévy process is necessary for the general definition of potentially non-causal processes. Alternatively, proposing that the eigenvalues of A have strictly negative real parts would enable us to define multivariate heavy-tailed CARMA processes which are naturally causal. In this case, it would be possible to loosen the moment condition for the Lévy process and assume the weaker logarithmic condition $\int_{\|x\|\geq 1} \log(\|x\|)\nu(dx) < \infty$, see Definition 3.20 of Marquardt and Stelzer (2007). However, as in the univariate case, the MCARMA(p,q) process can be interpreted as the stationary solution of a pth order differential equation. Given a state space model (A,B,C,L), we define the matrix-valued autoregressive polynomial P and moving average polynomial Q as

$$P(z) := I_m z^p + A_1 z^{p-1} + \ldots + A_p,$$
$$Q(z) := B_0 z^q + \ldots + B_q, \quad z \in \mathbb{R},$$

for the $A_1,\ldots,A_p \in \mathbb{R}^{m\times m}$, $B_0,\ldots,B_q \in \mathbb{R}^{m\times d}$ of (2.7). Then, the multivariate continuous-time ARMA(p,q) can be interpreted as solution of the pth order differential equation

$$P(D)Y_t = Q(D)DL_t. \tag{2.8}$$

Even though Definition 2.4 assumes matrices of a special structure Schlemm and Stelzer (2012a) proved that the classes of causal MCARMA processes and of causal linear state space models coincide under Assumption $L1$. This means, that any process $(Y_t)_{t\in\mathbb{R}}$ which satisfies (2.2) for some A,B,C where A has only eigenvalues with negative real parts, is a MCARMA process as well. Since the class of causal continuous linear state space models is better suited for the purpose of estimation, we mostly use this representation. Conclusively, note that two different decompositions (A_1,B_1,C_1,L) and (A_2,B_2,C_2,L) may lead to the same state space model in the multivariate setting. This problem does not occur in the univariate setting if we assume the controller canonical state space representation (2.3) and fix $c_0 = 1$. Therefore, we have to derive further conditions to avoid undesired redundancy for multivariate processes later.

2.3. Discrete sampling

In our setting, statistical inference is done for some MCARMA(p,q) process which is observed at a discrete-time grid with a fixed distance $\Delta > 0$ between the observations. Therefore, we revisit the results of Schlemm and Stelzer (2012b) who already investigated this setting for processes with existing second moments. Until further mentioning, we assume Assumption $L1$.

Theorem 2.5 (Theorem 3.6 of Schlemm and Stelzer (2012b)).
Assume that $(Y_t)_{t\in\mathbb{R}}$ is the output process of the continuous-time state space model given in (2.2). Then the sampled process $\left(Y_k^{(\Delta)}\right)_{k\in\mathbb{Z}} := (Y_{k\Delta})_{k\in\mathbb{Z}}$ has the state space representation

$$Y_k^{(\Delta)} = CX_k^{(\Delta)}, \quad X_k^{(\Delta)} = e^{A\Delta}X_{k-1}^{(\Delta)} + N_k^{(\Delta)}, \quad k \in \mathbb{Z}, \tag{2.9}$$

where

$$N_k^{(\Delta)} = \int_{(k-1)\Delta}^{k\Delta} e^{A(k\Delta-u)} B dL_u.$$

The sequence $\left(N_k^{(\Delta)}\right)_{k\in\mathbb{Z}}$ *is i.i.d. with mean zero and covariance matrix*

$$\Sigma_N^{(\Delta)} = \int_0^{\Delta} e^{Au} B \Sigma_L B^\top e^{A^\top u} du.$$

Furthermore, under the assumption that the eigenvalues of A *have strictly negative real parts,* $Y_k^{(\Delta)}$ *has the vector MA*(∞) *representation*

$$Y_k^{(\Delta)} = \sum_{j=0}^{\infty} \Phi_j N_{k-j}^{(\Delta)}, \quad k \in \mathbb{Z},$$

where $\Phi_j = Ce^{A\Delta j} \in \mathbb{R}^{m\times N}$.

In the following, $\|M\|$ denotes the Frobenius norm of a matrix M of arbitrary dimensions. However, it can often be replaced by any sub-multiplicative matrix norm.

Remark 2.6.

a) If $\mathbb{E}\|L_1\|^s < \infty$, then the same arguments as in Marquardt and Stelzer (2007), Proposition 3.30, lead to $\mathbb{E}\|N_1^{(\Delta)}\|^s < \infty$. The sub-multiplicity of $\|\cdot\|$ and the monotone convergence theorem imply

$$\mathbb{E}\|Y_k^{(\Delta)}\|^s \leq \sum_{j=0}^{\infty} \|\Phi_j\|^s \mathbb{E}\|N_1\|^s.$$

Therefore, $\mathbb{E}\|Y_k^{(\Delta)}\|^s < \infty$ holds as well.

b) Under the assumption that the eigenvalues of A have strictly negative real parts, the coefficients Φ_j of the polynomial Φ are exponentially decreasing which directly implies

$$\sum_{s=0}^{\infty} s^r \|\Phi_s\| < \infty \quad \text{for any } r \in \mathbb{N}_0. \tag{2.10}$$

Later, we need a special type of state space representation of the sampled process to show the consistency of Whittle‘s estimate and to construct the adjusted Whittle estimator. In view of this model, we have to bring the so called linear innovations up.

Definition 2.7. (linear innovations, see Schlemm and Stelzer Schlemm and Stelzer (2012b)) Let $(Y_k)_{k\in\mathbb{Z}}$ be a (discrete) $\mathbb{R}^m$-valued stochastic process with finite second moments. The linear innovations $(\varepsilon_k)_{k\in\mathbb{Z}}$ of $(Y_k)_{k\in\mathbb{Z}}$ are defined by

$$\varepsilon_k = Y_k - \mathrm{Pr}_{k-1} Y_k,$$

$$\Pr_k = \text{ orthogonal projection onto } \overline{\text{span}}\{Y_\nu : -\infty < \nu \leq k\},$$

where the closure is taken in the Hilbert space of square-integrable random variables with inner product $(X, Y) \to \mathbb{E}X^\top Y$.

Adjusted to our notation, Proposition 2.1 of Schlemm and Stelzer (2012b) gives the following result, which implies that almost all sampled stationary linear state space models have a vector MA(∞) representation in which the white noise are the linear innovations.

Theorem 2.8 (Proposition 2.1 of Schlemm and Stelzer (2012b))**.**
In the situation of Theorem 2.5 assume that the eigenvalues of A have strictly negative real parts and Σ_L is positive definite. Then, the following holds:

(a) The Riccati equation

$$\begin{aligned}\Omega^{(\Delta)} =& e^{A\Delta}\Omega^{(\Delta)}\left(e^{A\Delta}\right)^\top + \Sigma_N^{(\Delta)} \\ &- \left(e^{A\Delta}\Omega^{(\Delta)}C^\top\right)\left(C\Omega^{(\Delta)}C^\top\right)^{-1}\left(e^{A\Delta}\Omega^{(\Delta)}C^\top\right)^\top\end{aligned}$$

has a unique positive semidefinite solution $\Omega^{(\Delta)}$.

(b) Let $\mathfrak{B}$ be the backshift operator, i.e.,

$$\mathfrak{B}Y_k^{(\Delta)} = Y_{k-1}^{(\Delta)},$$

and

$$K^{(\Delta)} = \left(e^{A\Delta}\Omega^{(\Delta)}C^\top\right)\left(C\Omega^{(\Delta)}C^\top\right)^{-1}$$

be the Kalman gain matrix. Furthermore, define the polynomial Π as

$$\Pi(z) := \Pi^{(\Delta)}(z) := \left(I_m - C\left(I_N - (e^{A\Delta} - K^{(\Delta)}C)z\right)^{-1}K^{(\Delta)}z\right).$$

Then, the linear innovations are

$$\begin{aligned}\varepsilon_k^{(\Delta)} &= \left(I_m - C\left(I_N - (e^{A\Delta} - K^{(\Delta)}C)\mathfrak{B}\right)^{-1}K^{(\Delta)}\mathfrak{B}\right)Y_k^{(\Delta)} \\ &= Y_k^{(\Delta)} - \sum_{j=1}^{\infty} C(e^{A\Delta} - K^{(\Delta)}C)^{j-1}K^{(\Delta)}Y_{k-j}^{(\Delta)} \\ &= \Pi^{(\Delta)}(\mathfrak{B})Y_k^{(\Delta)}.\end{aligned}$$

Furthermore, the absolute value of any eigenvalue of $e^{A\Delta} - K^{(\Delta)}C$ is less than one and $Y^{(\Delta)} := (Y_k^{(\Delta)})_{k\in\mathbb{Z}}$ has the moving average representation

$$Y_k^{(\Delta)} = \varepsilon_k^{(\Delta)} + C\sum_{j=1}^{\infty}\left(e^{A\Delta}\right)^{j-1}K^{(\Delta)}\varepsilon_{k-j}^{(\Delta)}, \quad k \in \mathbb{Z}.$$

(c) The covariance matrix $V^{(\Delta)}$ of the linear innovations $(\varepsilon_k^{(\Delta)})_{k\in\mathbb{Z}}$ has the representation

$$V^{(\Delta)} = C\Omega^{(\Delta)}C^\top.$$

If $\Omega^{(\Delta)}$ is positive definite and C has full rank, $V^{(\Delta)}$ is invertible.

Remark 2.9.

(a) In the situation of Theorem 2.8

$$\operatorname{tr}(V^{(\Delta)}) = \min_{X\in\mathcal{M}_{k-1}} \mathbb{E}\left[\left(Y_k^{(\Delta)} - X\right)^\top \left(Y_k^{(\Delta)} - X\right)\right],$$

where $\operatorname{tr}(V^{(\Delta)})$ denotes the trace of $V^{(\Delta)}$ and $\mathcal{M}_{k-1} := \overline{\operatorname{span}}\left\{Y_\nu^{(\Delta)}, \nu \le k-1\right\}$ holds.

(b) If A has strictly negative real parts, then the sampled state space process is stable which provides the representations of Theorem 2.8 (b). Thereby, $\Pi(z)$ is invertible for $z \in \mathbb{C}$ with $|z| = 1$ and

$$\Pi^{-1}(z) := \Pi^{(\Delta)-1}(z) = \left(I_m + C\left(I_N - e^{A\Delta}z\right)^{-1} K^{(\Delta)}z\right). \tag{2.11}$$

In the case of an α-stable CARMA process, sampling properties were discovered by García et al. (2011). For the sake of completeness, we state Proposition 3.1 of García et al. (2011). Therefore, we have to introduce Proposition 2 of Brockwell et al. (2011) which yields another representation for the continuous-time α-stable CARMA(p,q) process.

Proposition 2.10 (Proposition 2 of Brockwell et al. (2011))**.**
Assume that $(Y_t)_{t\in\mathbb{R}}$ is a (causal) symmetric α-stable CARMA(p,q) process with driving process $L^{(\alpha)}$ and assume further that the polynomials a and c have no common factors and all zeros $\lambda_1,\dots,\lambda_p$ of the polynomial a are distinct. Then, $(Y_t)_{t\in\mathbb{R}}$ can be represented as a sum of dependent and possibly complex-valued CAR(1) processes

$$Y_t = \sum_{j=1}^p Y_t^{(j)}, \quad t\in\mathbb{R},$$

where

$$Y_t^{(j)} = \kappa_j \int_{-\infty}^t e^{\lambda_j(t-u)} dL_u^{(\alpha)}, \quad \kappa_j = \frac{b(\lambda_j)}{a'(\lambda_j)}, \quad j = 1,\dots,p.$$

Here a' denotes the first derivative of a.

Note that the following result is adapted for the symmetric α-stable CARMA process. The original result includes α-stable processes with $\beta \neq 0$.

Proposition 2.11 (Proposition 3.1 of García et al. (2011))**.**
Under the assumptions of Proposition 2.10, the following holds:

a) For given $\Delta > 0$, the sampled process $(Y_k^{(\Delta)})_{k\in\mathbb{Z}}$ has the representation

$$Y_k^{(\Delta)} = \sum_{j=1}^{p} Y_k^{(j,\Delta)}, \quad k \in \mathbb{Z},$$

where the discrete-time processes $(Y_k^{(j,\Delta)})_{k\in\mathbb{Z}}$, $j = 1,\ldots,p$, are obtained by sampling the component CAR(1) processes $(Y_t^{(1)})_{t\in\mathbb{R}},\ldots,(Y_t^{(p)})_{t\in\mathbb{R}}$ of Proposition 2.10 at spacing Δ. Since $(Y_t)_{t\in\mathbb{R}}$ is assumed to be strictly stationary by definition, we consider

$$Y_k^{(j,\Delta)} = e^{\lambda_j\Delta} Y_{k-1}^{j,\Delta} + Z_k^{(j,\Delta)}, \quad k \in \mathbb{Z},$$

with white noise

$$Z_k^{(j,\Delta)} = \kappa_j \int_{(k-1)\Delta}^{k\Delta} e^{\lambda_j(k\Delta-u)} dL_u^{(\alpha)}, \quad k \in \mathbb{Z}.$$

The random variables $Z_k^{(j,\Delta)}$ are $S_\alpha(\sigma|\kappa_j|\left[\int_0^\Delta e^{\alpha\lambda_j(\Delta-u)}du\right]^{1/\alpha}, 0, 0)$-distributed. For fixed Δ, j but varying k, they are i.i.d., for fixed Δ, k but varying j they are dependent.

b) For given $\Delta > 0$ the sampled process $(Y_k^{(\Delta)})_{k\in\mathbb{Z}}$ satisfies the equations

$$a_D^{(\Delta)}(\mathfrak{B})Y_k^{(\Delta)} = \prod_{j=1}^{p}(1 - e^{\lambda_j\Delta}\mathfrak{B})Y_k^{(\Delta)} = U_k^{(\Delta)}, \quad k \in \mathbb{Z},$$

where

$$a_D^{(\Delta)}(x) = \prod_{j=1}^{p}(1 - e^{\lambda_j\Delta}x) = 1 - \phi_1 x - \ldots - \phi_p x^p, \quad x \in \mathbb{R}.$$

The process $(U_k^{(\Delta)})_{k\in\mathbb{Z}}$ has the representation

$$U_k^{(\Delta)} = W_{k,0}^{(\Delta)} + W_{k,1}^{(\Delta)} + \ldots + W_{k,n-1}^{(\Delta)}, \quad k \in \mathbb{Z},$$

where

$$W_{k,\ell}^{(\Delta)} = \sum_{j=1}^{p} \eta_\ell^{(j,\Delta)} Z_{k-\ell}^{(j,\Delta)}$$

and $Z_{k-\ell}^{(j,\Delta)}$ are as in a). For all $k \in \mathbb{Z}$ the vector $(W_{k,0}^{(\Delta)},\ldots,W_{k,p-1}^{(\Delta)})$ has independent components, which for $\ell = 0,\ldots,p-1$ are $S_\alpha(\sigma_\ell^{(\Delta)}, 0, 0)$-distributed with

$$\sigma_\ell^{(\Delta)} = \sigma\left(\int_0^\Delta |f_\ell^{(\Delta)}(u)|^\alpha du\right)^{1/\alpha},$$

for $f_\ell^{(\Delta)}(u) = \sum_{j=1}^{p} \eta_\ell^{(j,\Delta)} \kappa_j e^{\lambda_j(\Delta-u)}$ for $0 \le u \le \Delta$. The $\eta_\ell^{(j,\Delta)}$ have the representa-

tion

$$\eta_0^{(j,\Delta)} = 1, \quad \eta_1^{(j,\Delta)} = -\sum_{r\neq j} e^{\lambda_r \Delta}, \quad \eta_2^{(j,\Delta)} = \sum_{r_1,r_2\neq j}{}' e^{(\lambda_{r_1}+\lambda_{r_2})\Delta},$$
$$\eta_3^{(j,\Delta)} = -\sum_{r_1,r_2,r_3\neq j}{}' e^{(\lambda_{r_1}+\lambda_{r_2}+\lambda_{r_3})\Delta}, \ldots, \quad \eta_{p-1}^{(j,\Delta)} = (-1)^{p-1} e^{\sum_{r\neq j}\lambda_r \Delta},$$

where the sum $\sum'$ *is taken over the different indices only.*

c) *The process* $(U_k^{(\Delta)})_{k\in\mathbb{Z}}$ *is* $(p-1)$*-dependent, and for all* $k\in\mathbb{Z}$ *the random variable* $U_k^{(\Delta)}$ *is* $S_\alpha(\sigma_\star^{(\Delta)},0,0)$*-distributed with*

$$\sigma_\star^{(\Delta)} = \left(\sum_{\ell=0}^{p-1}\left(\sigma_\ell^{(\Delta)}\right)^\alpha\right)^{1/\alpha}.$$

2.4. Second order properties

Firstly, we assume Assumption $L1$. Later, we aim to estimate a general linear state space model where the matrices A, B, C and the covariance matrix Σ_L of the driving Lévy process are parameterized. These parameters model the dependency structure of the MCARMA process. Therefore, it naturally comes to mind, to construct an estimator based on statistical figures which depict the dependency structure of the underlying process. From the theoretical perspective, for the multivariate stationary discrete-time process $(Z_k)_{k\in\mathbb{Z}}$ the most popular one is the autocovariance function Γ_Z which is defined by

$$\Gamma_Z(h) = \operatorname{Cov}(Z_h, Z_0),\ h\in\mathbb{N}_0, \quad \Gamma_Z(h) = \Gamma_Z(-h)^\top,\ h<0.$$

In the same way, the definition can be extended to continuous-time processes. Doing a Fourier transformation, let us leave the time domain and enter the frequency domain. As a result, we obtain the spectral density.

Definition 2.12 (spectral density)**.**
For a centered stationary discrete time series $(Z_k)_{k\in\mathbb{Z}}$ with an absolute summable autocovariance function $\Gamma_Z = (\Gamma_Z(h))_{h\in\mathbb{Z}}$ the spectral density f_Z is defined as

$$f_Z(\omega) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} \Gamma_Z(h) e^{-ih\omega}, \quad \omega\in[-\pi,\pi].$$

Vice versa, the autocovariance function can be represented as

$$\Gamma_Z(h) = \int_{-\pi}^{\pi} f_Z(\omega) e^{ih\omega} d\omega, \quad h\in\mathbb{Z}. \tag{2.12}$$

Obviously, the autocovariance function as well as the spectral density function are purely theoretical measures. Therefore, we approximate them by their empirical counterparts, the sample autocovariance function and the periodogram, respectively. Based on a sample of size n the autocovariance function of the centered stationary time series $(Z_k)_{k\in\mathbb{Z}}$ can be estimated evidently by

$$\overline{\Gamma}_{n,Z}(h) := \frac{1}{n}\sum_{k=1}^{n-h} Z_{k+h}Z_k^\top,\ h \geq 0, \quad \overline{\Gamma}_{n,Z}(h) := \overline{\Gamma}_{n,Z}(-h)^\top,\ h < 0.$$

A convenient way to estimate the spectral density is the periodogram.

Definition 2.13. (periodogram)
For a sample $Z_1, \ldots, Z_n$ of n observations of a (real-valued) m-dimensional stationary time series $(Z_k)_{k\in\mathbb{Z}}$ the periodogram $I_{n,Z} : [-\pi, \pi] \to \mathbb{R}^{m\times m}$ is defined as

$$I_{n,Z}(\omega) = \frac{1}{2\pi n}\left(\sum_{j=1}^{n} Z_j e^{-ij\omega}\right)\left(\sum_{k=1}^{n} Z_k e^{ik\omega}\right)^\top, \quad \omega \in [-\pi, \pi]. \tag{2.13}$$

Considering

$$I_{n,Z}(\omega_j) = \frac{1}{2\pi}\sum_{h=-n+1}^{n-1} \overline{\Gamma}_{n,Z}(h)e^{-ih\omega_j}$$

for $\omega_j = \frac{\pi j}{n}$, $j = -n+1, \ldots, n$, we see that the periodogram is just the Fourier transform of the sample autocovariance function. Therefore, we get an obvious connection between the periodogram and the spectral density. It is well known that the periodogram is not a consistent estimator for the spectral density in the equidistantly sampled MCARMA setting, see Theorem 3.1 of Fasen (2013). However, for a causal MCARMA process, the sample autocovariance is strongly consistent. In the following, we denote the sample autocovariance of $(Y_k^{(\Delta)})_{k\in\mathbb{Z}}$ and $(N_k^{(\Delta)})_{k\in\mathbb{Z}}$ of size n as $\overline{\Gamma}_{n,Y}$ and $\overline{\Gamma}_{n,N}$, respectively. In the same way, the corresponding autocovariances, spectral densities and periodograms are named. Note that we often use small letters for the corresponding univariate counterparts.

Lemma 2.14.
Suppose Assumption L1 hold and that the eigenvalues of A have strictly negative real parts. Then,

$$\overline{\Gamma}_{n,Y}(h) \xrightarrow{a.s.} \Gamma_Y(h) \quad \forall h \in \mathbb{Z}$$

and $\sum_{h=-\infty}^{\infty} \|\Gamma_Y(h)\| < \infty$.

Proof. Due to Proposition 3.34 of Marquardt and Stelzer (2007) the process $(Y_t)_{t\in\mathbb{R}}$ is ergodic. Therefore, Theorem 2.5 and Theorem 4.3 of Krengel (2011) imply that the sampled process $\left(Y_k^{(\Delta)}\right)_{k\in\mathbb{Z}}$ is ergodic as well. Moreover, by Proposition 3.13 of Marquardt and Stelzer (2007) the autocovariance function $\Gamma_Y^{(c)}$ of the continuous-time process satisfies

$$\Gamma_Y^{(c)}(h) = Ce^{Ah}\Sigma_N^{(\Delta)}C^\top,$$

with $\Sigma_N^{(\Delta)}$ as defined in Theorem 2.5. Since the eigenvalues of A have strictly negative real parts

$$\sum_{h\in\mathbb{Z}} \|\Gamma_Y(h)\| = \sum_{h\in\mathbb{Z}} \|\Gamma_Y^{(c)}(\Delta h)\| < \infty.$$

Birkhoff's Ergodic Theorem now leads to

$$\overline{\Gamma}_{n,Y}(h) \xrightarrow{a.s.} \mathbb{E}\left[Y_h^{(\Delta)} Y_0^{(\Delta)\top}\right] = \Gamma_Y(h) \quad \forall\, h \in \mathbb{Z}.$$

□

Remark 2.15.
Similarly, one can show that in the situation of Lemma 2.14 the estimated autocovariance of $N^{(\Delta)}$ as introduced in Remark 3.7 behaves in the same way, i.e.,

$$\overline{\Gamma}_{n,N}(h) \xrightarrow{a.s.} \Gamma_N(h), \qquad h \in \mathbb{Z}.$$

Obviously, $\Gamma_N(h) = 0$ for $h \neq 0$ and $\Gamma_N(0) = \Sigma_N^{(\Delta)}$ holds.

Considering an i.i.d. sequence, the sample autocovariance even has an asymptotic normal distribution.

Lemma 2.16.
Let $(Z_k)_{k\in\mathbb{Z}}$ be a N-dimensional i.i.d. sequence with $\mathbb{E}\|Z_1\|^4 < \infty$, and covariance matrix Σ_Z. Define

$$\overline{\Gamma}_{n,Z}(h) = \frac{1}{n}\sum_{j=1}^{n-h} Z_{j+h} Z_j^\top, \qquad n \geq h \geq 0,$$

Then, for fixed $\ell \in \mathbb{N}$,

$$\sqrt{n}\left(\begin{bmatrix} \operatorname{vec}\left(\overline{\Gamma}_{n,Z}(0)\right) \\ \operatorname{vec}\left(\overline{\Gamma}_{n,Z}(1)\right) \\ \vdots \\ \operatorname{vec}\left(\overline{\Gamma}_{n,Z}(\ell)\right) \end{bmatrix} - \begin{bmatrix} \operatorname{vec}(\Sigma_Z) \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma_{\Gamma_Z}(\ell)),$$

where

$$\Sigma_{\Gamma_Z}(\ell) = \left(\begin{array}{c|c} \mathbb{E}[Z_1 Z_1^\top \otimes Z_1 Z_1^\top] - \Sigma_Z \otimes \Sigma_Z & 0_{N^2\times \ell N^2} \\ \hline 0_{\ell N^2 \times N^2} & I_\ell \otimes \Sigma_Z \otimes \Sigma_Z \end{array}\right).$$

Proof. The proof is similar to the proof of Proposition 4.4 in Lütkepohl (2005) and is therefore omitted. □

The assumptions of Lemma 2.16 are obviously satisfied when considering an i.i.d. white noise. In that case, we can also show that the supremum over the expectation of the centered squared norm of the sample autocovariance multiplied by the sample size is bounded. Hereby, as always, $\mathfrak{C} > 0$ denotes a generic constant which may change from line to line.

Lemma 2.17.
Under Assumption L1,

$$\sup_{j\in\mathbb{Z}} n\mathbb{E}\left[\left\|\overline{\Gamma}_{n,N}(j) - \mathbb{E}\left[\overline{\Gamma}_{n,N}(j)\right]\right\|^2\right] \leq \mathfrak{C}$$

holds.

Proof. On the one hand, we have

$$\begin{aligned}
&n\mathbb{E}[\|\overline{\Gamma}_{n,N}(0) - \mathbb{E}[\overline{\Gamma}_{n,N}(0)]\|^2] \\
&= n\mathbb{E}\left[\sum_{s,t=1}^{N}\left|\overline{\Gamma}_{n,N}(0)[s,t] - \mathbb{E}[\overline{\Gamma}_{n,N}(0)[s,t]]\right|^2\right] \\
&= n\sum_{s,t=1}^{N}\mathbb{E}\left[\overline{\Gamma}_{n,N}(0)[s,t]^2\right] - \mathbb{E}\left[\overline{\Gamma}_{n,N}(0)[s,t]\right]^2 \\
&= n\sum_{s,t=1}^{N}\mathbb{E}\left[\frac{1}{n^2}\sum_{j,k=1}^{n}(N_j^{(\Delta)}N_j^{(\Delta)\top})[s,t](N_k^{(\Delta)}N_k^{(\Delta)\top})[s,t]\right] - \mathbb{E}\left[\frac{1}{n}\sum_{j=1}^{n}(N_j^{(\Delta)}N_j^{(\Delta)\top})[s,t]\right]^2 \\
&= \sum_{s,t=1}^{N}\left(\mathbb{E}\left[(N_1^{(\Delta)}N_1^{(\Delta)\top})[s,t]^2\right] - \mathbb{E}\left[(N_1^{(\Delta)}N_1^{(\Delta)\top})[s,t]\right]^2\right) \\
&= \sum_{s,t=1}^{N}\operatorname{Var}\left((N_1^{(\Delta)}N_1^{(\Delta)\top})[s,t]\right).
\end{aligned}$$

Since $\mathbb{E}\left[\overline{\Gamma}_{n,N}(j)\right] = 0$ for $n > j \neq 0$, we obtain for $j > 0$

$$\begin{aligned}
&n\mathbb{E}[\|\overline{\Gamma}_{n,N}(j) - \mathbb{E}[\overline{\Gamma}_{n,N}(j)]\|^2] \\
&= n\sum_{s,t=1}^{N}\mathbb{E}\left[\overline{\Gamma}_{n,N}(j)[s,t]^2\right] \\
&= n\sum_{s,t=1}^{N}\mathbb{E}\left[\frac{1}{n^2}\sum_{k,\ell=1}^{n-j}(N_{k+j}^{(\Delta)}N_k^{(\Delta)\top})[s,t](N_{\ell+j}^{(\Delta)}N_\ell^{(\Delta)\top})[s,t]\right] \\
&= n\sum_{s,t=1}^{N}\mathbb{E}\left[\frac{1}{n^2}\sum_{k=1}^{n-j}(N_{k+j}^{(\Delta)}N_k^{(\Delta)\top})[s,t]^2\right] \\
&= \frac{n-j}{n}\sum_{s,t=1}^{N}\mathbb{E}\left[(N_{1+j}^{(\Delta)}N_1^{(\Delta)\top})[s,t]^2\right] \\
&\leq \sum_{s,t=1}^{N}\mathbb{E}\left[(N_1^{(\Delta)}N_2^{(\Delta)\top})[s,t]^2\right]
\end{aligned}$$

$$= \sum_{s,t=1}^{N} \mathrm{Var}\left((N_1^{(\Delta)} N_2^{(\Delta)\top})[s,t]\right)$$

and with similar calculations we obtain the same bound for $j < 0$. Therefore,

$$\begin{aligned}
&\sup_{j\in\mathbb{Z}} n\mathbb{E}\left[\left\|\overline{\Gamma}_{n,N}(j) - \mathbb{E}\left[\overline{\Gamma}_{n,N}(j)\right]\right\|^2\right] \\
&\leq \max\left\{\sum_{s,t=1}^{N} \mathrm{Var}\left((N_1^{(\Delta)} N_2^{(\Delta)\top})[s,t]\right), \sum_{s,t=1}^{N} \mathrm{Var}\left((N_1^{(\Delta)} N_1^{(\Delta)\top})[s,t]\right)\right\} \leq \mathfrak{C}.
\end{aligned}$$

□

Changing to the frequency domain, an application of Theorem 11.8.3 Brockwell and Davis (1991) and Remark 2.9 already give a direct representation of the spectral density of the sampled process. By means of the definition

$$\Phi(x) := \sum_{j=0}^{\infty} \Phi_j x^j, \quad x \in \mathbb{C} : |x| = 1,$$

the spectral density $f_Y^{(\Delta)}$ of $Y^{(\Delta)}$ has the representation

$$\begin{aligned}
f_Y^{(\Delta)}(\omega) &= \frac{1}{2\pi}\Phi(e^{-i\omega})\Sigma_N^{(\Delta)}\Phi(e^{i\omega})^\top \\
&= \frac{1}{2\pi} C\left(e^{i\omega} I_N - e^{A\Delta}\right)^{-1} \Sigma_N^{(\Delta)} \left(e^{-i\omega} I_N - e^{A^\top \Delta}\right)^{-1} C^\top, \quad \omega \in [-\pi, \pi]. \qquad (2.14)
\end{aligned}$$

Alternatively, the same theorem and Theorem 2.8 (b) yield

$$f_Y^{(\Delta)}(\omega) = \Pi^{-1}(e^{-i\omega}) V^{(\Delta)} \Pi^{-1}(e^{i\omega})^\top, \quad \omega \in [-\pi, \pi]. \qquad (2.15)$$

We furthermore have the representation

$$\begin{aligned}
f_Y^{(\Delta)}(\omega) &= \frac{\Sigma_L}{2\pi}\int_0^\Delta \left|\sum_{j=-\infty}^{\infty} g(u + j\Delta)\mathrm{e}^{i\omega j}\right|^2 du \\
&= \frac{\Sigma_L}{2\pi}\int_0^\Delta \left|c^\top \mathrm{e}^{Au}(I_p - \mathrm{e}^{A\Delta + i\omega I_p})^{-1} e_p\right|^2 du, \quad \omega \in [-\pi, \pi],
\end{aligned}$$

for univariate CARMA processes, see Fasen (2013), Example 2.4.

We now change to a heavy-tailed setting. Naturally, if second moments do not exist, it is not possible to define the autocovariance function and the spectral density, respectively. However, the sample autocovariance function and the periodogram can be defined nonetheless, even though it is questionable how to interpret them. In view of our previous results and what follows, it would be desirable to obtain a similar property as the almost sure convergence of Lemma 2.14 for a symmetric α-stable CARMA process. Fortunately, with an appropriate normalization, an asymptotic result can be proven. However, in contrast to

Lemma 2.14, the limit remains random.

Theorem 2.18.
Let Y be a symmetric α-stable CARMA process with kernel function $g(t) = c^\top e^{At} e_p \mathbb{1}_{[0,\infty)}(t)$ as given in (2.4) and let $\overline{\gamma}_{n,Y}(h)$, $h = -n+1, \ldots, n-1$, be the sample autocovariance function. Then, for fixed $m \in \mathbb{N}$ and as $n \to \infty$,

$$\frac{1}{n^{2/\alpha-1}} \left(\overline{\gamma}_{n,Y}(0), \ldots, \overline{\gamma}_{n,Y}(m)\right)$$
$$\xrightarrow{\mathcal{D}} \left(\int_0^\Delta \sum_{j=-\infty}^{\infty} g(\Delta j - s)^2 dL_s^{(\alpha/2)}, \ldots, \int_0^\Delta \sum_{j=-\infty}^{\infty} g(\Delta j - s) g(\Delta(j+m) - s) dL_s^{(\alpha/2)}\right),$$

where $L^{(\alpha/2)} = (L_t^{(\alpha/2)})_{t\geq 0}$ is an $\alpha/2$-stable Lévy process with

$$L_1^{(\alpha/2)} \sim S_{\alpha/2}\left(\sigma^2 \left(C_\alpha / C_{\alpha/2}\right)^{2/\alpha}, 1, 0\right)$$

and the constants C_α and $C_{\alpha/2}$ are defined as in (2.6).

To prove this result, we need a generalized version of Theorem 3 of Drapatz (2017). Since the proof of the following proposition is mostly the same as the proof Theorem 3 of Drapatz (2017), it is therefore omitted.

Proposition 2.19.
Let $g_1, g_2 : \mathbb{R} \to \mathbb{R}$ be bounded functions with $g_1, g_2 \in L^\delta(\mathbb{R})$ for some $\delta < \min\{\alpha, 1\}$ and $0 < \int_{-\infty}^{\infty} |g_1(s) g_2(s)|\, ds < \infty$. Suppose $L^{(\alpha)}$ is a symmetric α-stable Lévy process with $\alpha \in (0,2)$ and $L_1^{(\alpha)} \sim S_\alpha(\sigma, 0, 0)$. Define the continuous-time MA processes

$$Y_t^{[1]} = \int_{-\infty}^{\infty} g_1(t-s) dL_s^{(\alpha)} \quad \text{and} \quad Y_t^{[2]} = \int_{-\infty}^{\infty} g_2(t-s) dL_s^{(\alpha)}, \quad t \geq 0.$$

Furthermore, $G_{g_1,g_2} : [0,\Delta] \to \mathbb{R}$ is given as $s \to \sum_{j=-\infty}^{\infty} g_1(\Delta j - s) g_2(\Delta j - s)$ and suppose $G_{g_1,g_2} \in \mathrm{L}^{\alpha/2}[0,\Delta]$. Then, as $n \to \infty$,

$$\frac{1}{n^{2/\alpha}} \sum_{k=1}^{n} Y_{k\Delta}^{[1]} Y_{k\Delta}^{[2]} \xrightarrow{\mathcal{D}} \int_0^\Delta G_{g_1,g_2}(s) dL_s^{(\alpha/2)},$$

where $L^{(\alpha/2)}$ is the $\alpha/2$-stable Lévy process of Theorem 2.18.

Remark 2.20.
Note that

$$\int_0^\Delta G_{g_1,g_2}(s) dL_s^{(\alpha/2)} \sim S_{\alpha/2}(\sigma_{g_1,g_2}, \beta_{g_1,g_2}, 0),$$

has an $\alpha/2$-stable distribution with parameters

$$\beta_{g_1,g_2} = \frac{\int_0^\Delta (G_{g_1,g_2}^+(s))^{\alpha/2} - (G_{g_1,g_2}^-(s))^{\alpha/2} ds}{\int_0^\Delta |G_{g_1,g_2}(s)|^{\alpha/2} ds},$$

$$(\sigma_{g_1,g_2})^{\alpha/2} = \frac{\sigma^\alpha C_\alpha}{C_{\alpha/2}} \int_0^\Delta |G_{g_1,g_2}(s)|^{\alpha/2} ds,$$

see Property 1.2.3 and 3.2.2 of Samorodnitsky and Taqqu (1994). Here, $G^+_{g_1,g_2}(s) = \max\{0, G_{g_1,g_2}(s)\}$ and $G^-_{g_1,g_2}(s) = \max\{0, -G_{g_1,g_2}(s)\}$ denote the positive and negative part of $G_{g_1,g_2}(s)$.

Proof of Theorem 2.18.
Let $c_0, \ldots, c_m \in \mathbb{R}$. Then,

$$\begin{aligned}
&\frac{n}{n^{2/\alpha}} \left(c_0 \overline{\gamma}_{n,Y}(0) + \ldots + c_m \overline{\gamma}_{n,Y}(m)\right) \\
&= \frac{1}{n^{2/\alpha}} \left(c_0 \sum_{j=1}^{n} Y_j^{(\Delta)2} + \ldots + c_m \sum_{j=1}^{n-m} Y_j^{(\Delta)} Y_{j+m}^{(\Delta)} \right) \\
&= \frac{1}{n^{2/\alpha}} \left(\sum_{j=1}^{n} Y_j^{(\Delta)} \left(\sum_{k=0}^{m} c_k Y_{k+j}^{(\Delta)} \right) - \sum_{j=n-m+1}^{n} \sum_{k=n-j+1}^{m} c_k Y_j^{(\Delta)} Y_{k+j}^{(\Delta)} \right) \\
&=: J_n^{[1]} + J_n^{[2]}.
\end{aligned} \tag{2.16}$$

We obtain for $\delta < \alpha/2$

$$\begin{aligned}
\mathbb{E} \left| J_n^{[2]} \right|^\delta &= \mathbb{E} \left| n^{-2/\alpha} \sum_{j=n-m+1}^{n} \sum_{k=n-j+1}^{m} c_k Y_j^{(\Delta)} Y_{k+j}^{(\Delta)} \right|^\delta \\
&\leq n^{-2\delta/\alpha} m^\delta \max_{k=0,\ldots,m} |c_k|^\delta \mathbb{E} \left| Y_1^{(\Delta)} Y_{1+k}^{(\Delta)} \right|^\delta \stackrel{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Therefore, the second term in (2.16), $J_n^{[2]}$, is negligible. For the first term in (2.16), $J_n^{[1]}$, we define

$$Y_t^{[1]} := \int_{-\infty}^{\infty} \sum_{k=0}^{m} c_k g(t + k\Delta - s) dL_s^{(\alpha)} \quad \text{and} \quad Y_t^{[2]} := \int_{-\infty}^{\infty} g(t-s) dL_s^{(\alpha)}, \quad t \geq 0.$$

Thereby, we have

$$\begin{aligned}
\sum_{k=0}^{m} c_k Y_{k+j}^{(\Delta)} &= \int_{-\infty}^{\infty} \sum_{k=0}^{m} c_k g((k+j)\Delta - s) dL_s^{(\alpha)} = Y_{j\Delta}^{[1]}, \\
Y_j^{(\Delta)} &= \int_{-\infty}^{\infty} g(j\Delta - s) dL_s^{(\alpha)} = Y_{j\Delta}^{[2]}.
\end{aligned}$$

An application of Proposition 2.19 leads for $n \to \infty$ to

$$\begin{aligned}
\frac{n}{n^{2/\alpha}} \sum_{k=0}^{m} c_k \overline{\gamma}_{n,Y}(k) &= \frac{1}{n^{2/\alpha}} \sum_{k=1}^{n} Y_{k\Delta}^{[1]} Y_{k\Delta}^{(2)} + J_n^{[2]} \\
&\stackrel{\mathcal{D}}{\longrightarrow} \int_0^\Delta \sum_{j=-\infty}^{\infty} \left(\sum_{k=0}^{m} c_k g(\Delta(k+j) - s) g(\Delta j - s) \right) dL_s^{(\alpha/2)}
\end{aligned}$$

$$= \sum_{k=0}^{m} \int_0^{\Delta} \sum_{j=-\infty}^{\infty} c_k g(\Delta(k+j) - s) g(\Delta j - s) dL_s^{(\alpha/2)}.$$

Cramér-Wold completes the proof. □

Chapter 3

Whittle estimation for MCARMA Processes

In this chapter, we investigate the Whittle estimator for an equidistantly sampled MCARMA(p, q) process with existing second moments. We assume that the orders p and q of the autoregressive and moving average polynomial, respectively, are fixed and that the polynomials P and Q and the covariance matrix Σ_L of the driving process have to be estimated. In view of Chapter 2, we alternatively estimate the parameters of the corresponding continuous-time linear state space model. We consider the parameterized linear state space models $(A(\vartheta), B(\vartheta), C(\vartheta), L(\vartheta))_{\vartheta\in\Theta}$ for some parameter space $\Theta \subset \mathbb{R}^r$. In general, we emphasize all dependencies on the parameter $\vartheta \in \Theta$ by using notations such as $f_Y^{(\Delta)}(\cdot, \vartheta), \Phi(\cdot, \vartheta), \Pi(\cdot, \vartheta), \ldots$. In the following, we denote the true parameter as ϑ_0 and, for better readability, we do not carry it in the remaining part of this thesis, i.e. $f_Y^{(\Delta)}(\cdot), \Phi(\cdot), \Pi(\cdot), \ldots$ stand for $f_Y^{(\Delta)}(\cdot, \vartheta_0), \Phi(\cdot, \vartheta_0), \Pi(\cdot, \vartheta_0), \ldots$ in all our parameterized settings.
We start with the introduction of the Whittle estimator. To later derive the strong consistency and the asymptotic normality of the procedure, we then present some results concerning the identifiability of the sampled processes. Applying these findings yields our setting. Finally we state and proof the strong consistency and the asymptotic normality result.

3.1. The Whittle estimator

We turn towards the Whittle estimator. In 1951, this estimation procedure was firstly introduced by Peter Whittle (1951) who searched for a simple approximation for the Gaussian maximum likelihood method. His approach is based on the minimizing argument of the so-called Whittle function which measures the distance between the periodogram and the spectral density. Over the years, the original procedure was adapted for different settings and found various applications. In our setting, we define the Whittle function W_n as

$$W_n(\vartheta) = \frac{1}{2n} \sum_{j=-n+1}^{n} \left[\operatorname{tr}\left(f_Y^{(\Delta)}(\omega_j, \vartheta)^{-1} I_{n,Y}(\omega_j) \right) + \log\left(\det\left(f_Y^{(\Delta)}(\omega_j, \vartheta) \right) \right) \right], \quad \vartheta \in \Theta,$$

with $\omega_j = \frac{\pi j}{n}$ for $j = -n+1, \ldots, n$. Accordingly, the Whittle estimator is

$$\widehat{\vartheta}_n^{(\Delta)} := \arg\min_{\vartheta \in \Theta} W_n(\vartheta).$$

In this definition the terms $\log(\det(f_Y^{(\Delta)}(\omega_j, \vartheta)))$ can be replaced by $\log(\det V^{(\Delta)}(\vartheta))$ where $V^{(\Delta)}(\vartheta)$ is the covariance matrix of the linear innovations of Theorem 2.8 corresponding to the parameter ϑ. Thereby, the minimizing argument does not change, see Theorem 3''' of Chapter 3 of Hannan (2009) for a motivation. Therefore, if the covariance matrix $V^{(\Delta)}(\vartheta)$ of the linear prediction error does not depend on ϑ, we can neglect the penalty term $\log(\det V^{(\Delta)}(\vartheta))$ completely since it then is constant for all ϑ. However, in the case of state space models, $V^{(\Delta)}(\vartheta)$ depends on ϑ and has to be computed additionally (cf. Theorem 2.8). Conversely, for VARMA models, the Whittle function with penalty term $\log(\det \Sigma_e(\vartheta))$ of Dunsmuir and Hannan (1976) differs from our Whittle function. In their case Σ_e is the covariance matrix of the white noise which is obviously independent of the moving average and autoregressive polynomial, respectively.

3.2. Identifiability

As already mentioned, in the general multivariate setting, it is possible to obtain multiple linear state space representations (A, B, C, L) of the same process. Furthermore, the Whittle estimator is based on second order properties. Therefore, it is necessary that the true process can be identified from its second order properties. Additionally, since we only observe the continuous-time processes at a discrete-time grid, the so-called aliasing effect might appear. Loosely speaking, aliasing describes the possibility to observe the same discrete data from different continuous-time processes. Consequently, when aliasing occurs, it is impossible to uniquely identify the generating continuous-time process from just the discrete observations. This is a well-known problem when identifying parameters from discrete-time data, see e.g. Hansen and Sargent (1983). Obviously, the ability to uniquely identify the underlying process from discrete observations is necessary to estimate the true parameter correctly. Consequently, we have to restrict our setting in a way which renders a bijectiveness between the continuous-time state space models and the observable discrete-time processes. Therefore, we revisit the results of Schlemm and Stelzer (2012b) who already found conditions to avoid these problems. There are different approaches to determine the parametrization in a way to forgo multiple linear state space representations of the continuous-time process. Naturally, it is advantageous to choose a setting which is somewhat minimal. In our case, the minimality is referenced to the McMillan degree of the state space model.

Definition 3.1.

a) Let H be an $m \times d$-dimensional rational matrix function, i.e., the components $H[s,t]$ of H are rational functions of some variable $z \in \mathbb{R}$ for all $s \in \{1, \ldots m\}, t \in \{1, \ldots, d\}$. A matrix triplet (A, B, C) is called an algebraic realization of H of dimension N if $H(z) = C(zI_N - A)^{-1}B$ for $z \in \mathbb{R}$ and $A \in \mathbb{R}^{N \times N}, B \in \mathbb{R}^{N \times d}$ and $C \in \mathbb{R}^{m \times N}$.

b) Let H be a $m \times d$-dimensional rational matrix function. A minimal realization of H is an algebraic realization of dimension smaller than or equal to the dimension of every other algebraic realization of H. The dimension of a minimal realization of H is the McMillan degree of H.

Theorem 2.3.4 in Hannan and Deistler (1988) shows that assuming a minimal McMillan degree guarantees the uniqueness of the state space representation up to a change of basis. Hence, a fixed minimal McMillan degree reduces redundancies in the continuous-time model. Additionally, assuming Assumption $L1$ and a minimal McMillan degree of all investigated processes yield the positive definiteness of the covariance matrices $\Sigma_N^{(\Delta)}(\vartheta)$, $\vartheta \in \Theta$, in (2.2).

Lemma 3.2 (Corollary 3.9 of Schlemm and Stelzer (2012b))**.**
If the triplet (A, B, C) is minimal of dimension N and Σ is positive definite, then the $N \times N$ matrix $\not{\Sigma} = \int_0^\Delta e^{Au} B \Sigma B^\top e^{A^\top u} du$ has full rank N.

We now consider a result which serves to eventually bypass the aliasing effect with a practical assumption.

Lemma 3.3.
Assume that the matrices $A_1, A_2 \in \mathbb{R}^{N \times N}$ satisfy $e^{A_1 \Delta} = e^{A_2 \Delta}$ for some $\Delta > 0$. If all $\lambda \in \mathbb{C}$ which are either eigenvalues of A_1 or A_2 satisfy $|\Im(\lambda)| < \pi/\Delta$, then $A_1 = A_2$.

Finally, we do not only want to guarantee that the process can be identified from its second-order properties, but we further want to identify it by the second-order properties of its sampled process.

Definition 3.4.
Two stochastic processes are L^2-observationally equivalent if their spectral densities are the same. A family of continuous-time stochastic processes $(Y(\vartheta))_{\vartheta \in \Theta}$ is identifiable from the spectral density if, for every $\vartheta_1 \neq \vartheta_2$, the two processes $Y(\vartheta_1)$ and $Y(\vartheta_2)$ are not L^2-observationally equivalent. It is Δ-identifiable from the spectral density for some $\Delta > 0$ if for every $\vartheta_1 \neq \vartheta_2$ the two sampled processes $Y(\Delta \cdot, \vartheta_1)$ and $Y(\Delta \cdot, \vartheta_2)$ are not L^2-observationally equivalent.

Imposing conditions adjusted to the previous results, it now suffices to assume Δ-identifiability from the spectral density to forgo all identification difficulties.

Theorem 3.5 (Theorem 3.13 of Schlemm and Stelzer (2012b)).
Let $\Theta \subset \mathbb{R}^r$ be a parameter space and $\vartheta \to (A(\vartheta), B(\vartheta), C(\vartheta), L(\vartheta))$ be a parametrization of the continuous-time state space model (2.2). Assume that for $\vartheta \in \Theta$ the driving process $L(\vartheta)$ is centered and has a positive definite covariance matrix. Assume further that for $\vartheta \in \Theta$ the eigenvalues of the matrix $A(\vartheta)$ have strictly negative real parts and are a subset of $\{z \in \mathbb{C} : -\pi/\Delta < \Im(z) < \pi/\Delta\}$ and that $(A(\vartheta), B(\vartheta), C(\vartheta))$ is minimal with McMillan degree N. Finally assume that the collection of output processes corresponding to $(A(\vartheta), B(\vartheta), C(\vartheta), L(\vartheta))$ is identifiable from the spectral density. Then, it is Δ-identifiable as well.

3.3. Setting

In view of the previous section, we work under the following assumptions.

Assumption A.
For all $\vartheta \in \Theta$ the following holds:

(A1) The parameter space Θ is a compact subset of $\mathbb{R}^r$.

(A2) $L(\vartheta) = (L_t(\vartheta))_{t\in\mathbb{R}}$ is a centered Lévy process with positive definite covariance matrix $\Sigma_L(\vartheta)$.

(A3) The eigenvalues of $A(\vartheta)$ have strictly negative real parts.

(A4) The functions $\vartheta \mapsto \Sigma_L(\vartheta)$, $\vartheta \mapsto A(\vartheta)$, $\vartheta \mapsto B(\vartheta)$ and $\vartheta \mapsto C(\vartheta)$ are continuous. In addition, $C(\vartheta)$ has full rank.

(A5) The linear state space model $(A(\vartheta), B(\vartheta), C(\vartheta), L(\vartheta))$ is minimal with McMillan degree N.

(A6) For any $\vartheta_1, \vartheta_2 \in \Theta$ with $\vartheta_1 \neq \vartheta_2$ there exists an $\omega \in [-\pi, \pi]$ such that $f_Y(\omega, \vartheta_1) \neq f_Y(\omega, \vartheta_2)$, where $f_Y(\omega, \vartheta)$ is the spectral density of $Y(\vartheta)$.

(A7) The spectrum of $A(\vartheta) \in \mathbb{R}^{N\times N}$ is a subset of $\{z \in \mathbb{C} : -\frac{\pi}{\Delta} < \Im(z) < \frac{\pi}{\Delta}\}$.

Remark 3.6.

(a) As mentioned, Lemma 3.3 yields that the covariance matrix $\Sigma_N^{(\Delta)}(\vartheta)$ has full rank under Assumption $(A2)$ and $(A5)$. Furthermore, note that Assumptions $(A2)$ and $(A3)$ allow us to calculate the linear innovations. In this case, the covariance matrix $V^{(\Delta)}(\vartheta)$ of the linear innovations is non-singular (cf. Lemma 3.14 in Schlemm and Stelzer (2012b)) as well.

(b) Under Assumption A and representation (2.15) of the spectral density, the inverse $f_Y^{(\Delta)}(\omega, \vartheta)^{-1}$ of the spectral density exists and the mapping $(\omega, \vartheta) \mapsto f_Y^{(\Delta)}(\omega, \vartheta)^{-1}$ is continuous.

For the asymptotic normality of the Whittle estimator some further assumptions are required. In view of what follows, we introduce an abridging notation of the gradient vector and the Hessian matrix of some function. Namely, for some matrix function $g(\vartheta)$ in $\mathbb{R}^{m\times s}$ with ϑ in $\mathbb{R}^r$, the gradient matrix with respect to the parameter vector ϑ is denoted as $\nabla_\vartheta g(\vartheta) = \frac{\partial \operatorname{vec}(g(\vartheta))}{\partial \vartheta} \in \mathbb{R}^{ms\times r}$ and $\nabla_\vartheta g(\vartheta_0)$ is the shorthand for $\nabla_\vartheta g(\vartheta)|_{\vartheta=\vartheta_0}$. If $g: \mathbb{R}^r \to \mathbb{R}$, then $\nabla^2_\vartheta g(\vartheta) \in \mathbb{R}^{r\times r}$ denotes the Hessian matrix of $g(\vartheta)$.

Assumption B.

(B1) The true parameter value ϑ_0 is in the interior of Θ.

(B2) $\mathbb{E}\|L_1\|^4 < \infty$.

(B3) The functions $\vartheta \mapsto A(\vartheta)$, $\vartheta \mapsto B(\vartheta)$, $\vartheta \mapsto C(\vartheta)$ and $\vartheta \mapsto \Sigma_L(\vartheta)$ are three times continuously differentiable.

(B4) For any $c \in \mathbb{C}^r$, there exists an $\omega^* \in [-\pi, \pi]$ such that $\nabla_\vartheta f_Y^{(\Delta)}(\omega^*, \vartheta_0)c \neq 0_{m^2}$.

Remark 3.7.
Under Assumption A and $(B3)$ the mapping $\vartheta \mapsto f_Y^{(\Delta)}(\omega, \vartheta)$ is three times continuously differentiable due to representation (2.14) of the spectral density.

3.4. Strong consistency and asymptotic normality

Theorem 3.8.
Let Assumption A hold. Then, as $n \to \infty$,

$$\widehat{\vartheta}_n^{(\Delta)} \xrightarrow{a.s.} \vartheta_0.$$

Theorem 3.9.
Let Assumptions A and B hold. Furthermore, let $\Sigma_{\nabla W}$ be defined as

$$\begin{aligned}
\Sigma_{\nabla W} = \frac{1}{\pi} \int_{-\pi}^{\pi} &\nabla_\vartheta f_Y^{(\Delta)}(-\omega, \vartheta_0)^\top \left[f_Y^{(\Delta)}(-\omega)^{-1} \otimes f_Y^{(\Delta)}(\omega)^{-1} \right] \nabla_\vartheta f_Y^{(\Delta)}(\omega, \vartheta_0) d\omega \\
&+ \frac{1}{16\pi^4} \left[\int_{-\pi}^{\pi} \left[\Phi(e^{i\omega})^\top f_Y^{(\Delta)}(\omega)^{-1} \otimes \Phi(e^{-i\omega})^\top f_Y^{(\Delta)}(-\omega)^{-1} \right] \nabla_\vartheta f_Y^{(\Delta)}(-\omega, \vartheta_0) d\omega \right]^\top \\
&\cdot \left[\mathbb{E}\left[N_1^{(\Delta)} N_1^{(\Delta)\top} \otimes N_1^{(\Delta)} N_1^{(\Delta)\top} \right] - 3\Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)} \right] \\
&\cdot \left[\int_{-\pi}^{\pi} \left[\Phi(e^{-i\omega})^\top f_Y^{(\Delta)}(-\omega)^{-1} \otimes \Phi(e^{i\omega})^\top f_Y^{(\Delta)}(\omega)^{-1} \right] \nabla_\vartheta f_Y^{(\Delta)}(\omega, \vartheta_0) d\omega \right].
\end{aligned} \tag{3.1}$$

and

$$\Sigma_{\nabla^2 W} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \nabla_\vartheta f_Y^{(\Delta)}(-\omega, \vartheta_0)^\top \left[f_Y^{(\Delta)}(-\omega)^{-1} \otimes f_Y^{(\Delta)}(\omega)^{-1} \right] \nabla_\vartheta f_Y^{(\Delta)}(\omega, \vartheta_0) d\omega. \tag{3.2}$$

Then, as $n \to \infty$,

$$\sqrt{n}\left(\widehat{\vartheta}_n^{(\Delta)} - \vartheta_0 \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma_W),$$

where Σ_W has the representation $\Sigma_W = [\Sigma_{\nabla^2 W}]^{-1} \Sigma_{\nabla W} [\Sigma_{\nabla^2 W}]^{-1}$.

In contrast to the quasi maximum likelihood estimator of Schlemm and Stelzer (2012b), the limit covariance matrix of the Whittle estimator has an analytic representation. It can be used for the calculation of confidence bands.

Remark 3.10.
We want to compare our outcome with an analogue result for stationary discrete-time VARMA(p,q) processes $(Z_n)_{n\in\mathbb{N}}$ of the form (1.1) with finite fourth moments. In our setting we have the drawback that the autoregressive and the moving average polynomial influence the covariance matrix $\Sigma_N^{(\Delta)}$ of $(N_k^{(\Delta)})_{k\in\mathbb{N}}$. In the setting of stationary VARMA(p,q) processes of Dunsmuir and Hannan (1976) the covariance matrix Σ_e of the white noise $(e_n)_{n\in\mathbb{Z}}$ is not affected by the AR and MA polynomials. It was shown in Dunsmuir and Hannan (1976) that under very general assumptions for $d=m$ the resulting limit covariance matrix of the Whittle estimator for the VARMA parameters has the representation

$$\begin{aligned}\Sigma_W^{\text{VARMA}} &= \left[\frac{1}{4\pi}\int_{-\pi}^{\pi}\nabla_\vartheta f_Z(-\omega,\vartheta_0)^\top\left[f_Z(-\omega)^{-1}\otimes f_Z(\omega)^{-1}\right]\nabla_\vartheta f_Z(\omega,\vartheta_0)d\omega\right]^{-1}\\ &= 2\cdot[\Sigma_{\nabla^2 W}^{\text{VARMA}}]^{-1},\end{aligned}$$

which is simpler than our Σ_W. This can be traced back to $\Sigma_{\nabla W}^{\text{VARMA}} = 2\cdot\Sigma_{\nabla^2 W}^{\text{VARMA}}$, which is motivated on p. 51. In particular, for a Gaussian VARMA model, Σ_W^{VARMA} is the inverse of the Fisher information matrix.

Remark 3.11.

(a) Let the driving Lévy process be a Brownian motion. Then, the matrix $\Sigma_{\nabla W}$ reduces to

$$\begin{aligned}\Sigma_{\nabla W} =& \frac{1}{\pi}\int_{-\pi}^{\pi}\nabla_\vartheta f_Y^{(\Delta)}(-\omega,\vartheta_0)^\top\left[f_Y^{(\Delta)}(-\omega)^{-1}\otimes f_Y^{(\Delta)}(\omega)^{-1}\right]\nabla_\vartheta f_Y^{(\Delta)}(\omega,\vartheta_0)d\omega\\ =&2\cdot[\Sigma_{\nabla^2 W}]^{-1},\end{aligned}$$

see Remark 3.19, and hence, $\Sigma_W = 2\cdot[\Sigma_{\nabla W}]^{-1}$ is the inverse of the Fisher information matrix and corresponds to Σ_W^{VARMA} as in the previously mentioned discrete-time VARMA setting.

(b) Let $d=m=N$ and $C(\vartheta)=I_m$. Then, the state space model is a multivariate Ornstein-Uhlenbeck process (MCAR(1) process). In this example, $\Sigma_{\nabla W} = 2\cdot[\Sigma_{\nabla^2 W}]^{-1}$ holds as well. Because of $\Phi(z,\vartheta)=\sum_{j=0}^\infty e^{A(\vartheta)\Delta j}z^j = (1-e^{A(\vartheta)\Delta}z)^{-1} = \Pi^{-1}(z,\vartheta)$, the arguments are very similar to the arguments for VARMA models in Remark 3.10.

Remark 3.12.
For Gaussian state space processes

$$J = \left[2\mathbb{E}\left[\left(\frac{\partial}{\partial\vartheta_i}\varepsilon_1^{(\Delta)}(\vartheta_0)\right)^\mathsf{T} V^{(\Delta)-1}\left(\frac{\partial}{\partial\vartheta_j}\varepsilon_1^{(\Delta)}(\vartheta_0)\right)\right]\right.$$

$$+\operatorname{tr}\left(\left(\frac{\partial}{\partial\vartheta_i}V^{(\Delta)}(\vartheta_0)\right)V^{(\Delta)-1}\left(\frac{\partial}{\partial\vartheta_j}V^{(\Delta)}(\vartheta_0)\right)V^{(\Delta)-1}\right)\Bigg]_{i,j=1,\ldots,r}$$

is the Fisher information matrix (cf. Schlemm and Stelzer (2012b)). Since $W(\vartheta)=\mathcal{L}(\vartheta)$ due to Lemma 3.15, and $\nabla_\vartheta f_Y^{(\Delta)}(\omega,\vartheta)$ is uniformly bounded by an integrated dominant, we get by some straightforward applications of dominated convergence and some arguments of the proof of Schlemm and Stelzer (2012b), Lemma 2.17, that

$$\begin{aligned} J[i,j] &= \lim_{n\to\infty}\mathbb{E}\left[\tfrac{\partial}{\partial\vartheta_i}\tfrac{\partial}{\partial\vartheta_j}\mathcal{L}_n(\vartheta_0)\right]=\tfrac{\partial}{\partial\vartheta_i}\tfrac{\partial}{\partial\vartheta_j}\lim_{n\to\infty}\mathbb{E}[\mathcal{L}_n(\vartheta_0)]\\ &=\tfrac{\partial}{\partial\vartheta_i}\tfrac{\partial}{\partial\vartheta_j}W(\vartheta_0)=\lim_{n\to\infty}\mathbb{E}\left[\tfrac{\partial}{\partial\vartheta_i}\tfrac{\partial}{\partial\vartheta_j}W_n(\vartheta_0)\right]=\Sigma_{\nabla^2 W}[i,j], \end{aligned}$$

where $\mathcal{L}_n(\vartheta)$ is the quasi-Gaussian likelihood function. Furthermore, Schlemm and Stelzer (2012b), Lemma 2.17, show that if Assumption A holds and if there exists an $j_0\in\mathbb{N}$ such that the $((j_0+2)m^2)\times r$-matrix

$$\nabla\left[\begin{array}{c}\left[I_{j_0+1}\otimes K^{(\Delta)}(\vartheta_0)^\top\otimes C(\vartheta_0)\right]\left[\left(\operatorname{vec}\left(e^{I_N\Delta}\right)\right)^\top\left(\operatorname{vec}\left(e^{A(\vartheta_0)\Delta}\right)\right)^\top\cdots\left(\operatorname{vec}\left(e^{A^j(\vartheta_0)\Delta}\right)\right)^\top\right]^\top\\ \operatorname{vec}\left(V^{(\Delta)}(\vartheta_0)\right)\end{array}\right]$$

has rank r, then the matrix J is positive definite. Thus, our Assumption (B4) can be replaced by this condition.

3.5. Proofs of Theorem 3.8 and Theorem 3.9

3.5.1. Proof of the strong consistency

We start to prove some auxiliary results which we need for the proof of the consistency of Whittle‘s estimator. The following proposition states that the Whittle function W_n converges almost surely uniformly.

Proposition 3.13.
Let Assumptions (A1)–(A4) hold and

$$W(\vartheta):=\frac{1}{2\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(f_Y^{(\Delta)}(\omega,\vartheta)^{-1}f_Y^{(\Delta)}(\omega)\right)+\log\left(\det\left(f_Y^{(\Delta)}(\omega,\vartheta)\right)\right)d\omega,\quad \vartheta\in\Theta.$$

Then,

$$\sup_{\vartheta\in\Theta}|W_n(\vartheta)-W(\vartheta)|\overset{n\to\infty}{\longrightarrow}0\quad \mathbb{P}\text{-}a.s.$$

Proof We divide W_n in two parts and investigate them separately. Therefore, define

$$W_n^{(1)}(\vartheta):=\frac{1}{2n}\sum_{j=-n+1}^{n}\operatorname{tr}\left(f_Y^{(\Delta)}(\omega_j,\vartheta)^{-1}I_{n,Y}(\omega_j)\right)$$

and

$$W_n^{(2)}(\vartheta) = \frac{1}{2n} \sum_{j=-n+1}^{n} \log\left(\det\left(f_Y^{(\Delta)}(\omega_j, \vartheta)\right)\right),$$

such that $W_n(\vartheta) = W_n^{(1)}(\vartheta) + W_n^{(2)}(\vartheta)$. Since $(A1)$ and $(A4)$ are satisfied, we can apply Proposition A.6, which gives the uniform convergence

$$\sup_{\vartheta\in\Theta}\left|W_n^{(2)}(\vartheta) - \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left(\det\left(f_Y^{(\Delta)}(\omega, \vartheta)\right)\right) d\omega\right| \xrightarrow{n\to\infty} 0. \tag{3.3}$$

It remains to prove the appropriate convergence of $W_n^{(1)}$. Therefore, it is sufficient to show that

$$\sup_{\vartheta\in\Theta}\left\|\frac{1}{2n}\sum_{j=-n+1}^{n} f_Y^{(\Delta)}(\omega_j, \vartheta)^{-1} I_{n,Y}(\omega_j) - \frac{1}{2\pi}\int_{-\pi}^{\pi} f_Y^{(\Delta)}(\omega, \vartheta)^{-1} f_Y^{(\Delta)}(\omega) d\omega\right\| \xrightarrow{a.s.} 0 \tag{3.4}$$

holds. We approximate $f_Y^{(\Delta)}(\omega_j, \vartheta)^{-1}$ by the Cesàro sum of its Fourier series of size M for M sufficiently large. For better readability, we denote the kth Fourier coefficient of $f_Y^{(\Delta)}(\cdot, \vartheta)^{-1}$ by $\widehat{(f^{-1}(\vartheta))}_k$, i.e.,

$$\widehat{(f^{-1}(\vartheta))}_k := \frac{1}{2\pi}\int_{-\pi}^{\pi} f_Y^{(\Delta)}(\omega, \vartheta)^{-1} e^{-ik\omega} d\omega, \quad k \in \mathbb{Z},$$

and define

$$q_M(\omega, \vartheta) := \frac{1}{M}\sum_{j=0}^{M-1}\left(\sum_{|k|\le j} \widehat{(f^{-1}(\vartheta))}_k e^{ik\omega}\right) = \sum_{|k|<M}\left(1 - \frac{|k|}{M}\right)\widehat{(f^{-1}(\vartheta))}_k e^{ik\omega}$$

The inverse $f_Y^{(\Delta)}(\omega, \vartheta)^{-1}$ exists, is continuous and 2π-periodic in the first component. Thus, an application of Theorem A.3 gives that for any $\epsilon > 0$ there exists an $M_0(\epsilon) \in \mathbb{N}$ such that for $M \ge M_0(\epsilon)$

$$\sup_{\omega\in[-\pi,\pi]}\sup_{\vartheta\in\Theta}\left\|f_Y^{(\Delta)}(\omega, \vartheta)^{-1} - q_M(\omega, \vartheta)\right\| < \epsilon. \tag{3.5}$$

Let $\epsilon > 0$. In view of (3.5), we get

$$\begin{aligned} &\left\|\frac{1}{2n}\sum_{j=-n+1}^{n} f_Y^{(\Delta)}(\omega_j, \vartheta)^{-1} I_{n,Y}(\omega_j) - \frac{1}{2n}\sum_{j=-n+1}^{n} q_M(\omega_j, \vartheta) I_{n,Y}(\omega_j)\right\| \\ \le\ & \frac{\epsilon}{2n}\sum_{j=-n+1}^{n} \left\|I_{n,Y}(\omega_j)\right\|. \end{aligned} \tag{3.6}$$

Since all matrix norms are equivalent, we replace the Frobenius norm by the 1-norm and obtain

$$\frac{\epsilon}{2n}\sum_{j=-n+1}^{n}\|I_{n,Y}(\omega_j)\| \leq \frac{\epsilon\mathfrak{C}}{2n}\sum_{j=-n+1}^{n}\sum_{k=1}^{m}\sum_{\ell=1}^{m}|I_{n,Y}(\omega_j)[k,\ell]|. \tag{3.7}$$

The representation (2.13) of the periodogram and the non-negativeness of any one dimensional periodogram imply that $a^\top I_{n,Y}(\omega_j)a = I_{n,a^\top Y}(\omega_j) \geq 0$ so that $I_{n,Y}(\omega_j)$ is a positive semi-definite and Hermitian matrix. Therefore, for $k,\ell \in \{1,\ldots,m\},\ j \in \{-n+1,\ldots,n\}$,

$$\det\begin{pmatrix} I_{n,Y}(\omega_j)[k,k] & I_{n,Y}(\omega_j)[k,\ell] \\ I_{n,Y}(\omega_j)[\ell,k] & I_{n,Y}(\omega_j)[\ell,\ell]\end{pmatrix} \geq 0,$$

which implies

$$|I_{n,Y}(\omega_j)[k,\ell]| \leq \sqrt{I_{n,Y}(\omega_j)[k,k]I_{n,Y}(\omega_j)[\ell,\ell]} \leq I_{n,Y}(\omega_j)[k,k] + I_{n,Y}(\omega_j)[\ell,\ell]. \tag{3.8}$$

Combining (3.6), (3.7), (3.8) and Lemma A.1 give for $M \geq M_0(\epsilon)$

$$\begin{aligned}
&\left\|\frac{1}{2n}\sum_{j=-n+1}^{n} f_Y^{(\Delta)}(\omega_j,\vartheta)^{-1}I_{n,Y}(\omega_j) - \frac{1}{2n}\sum_{j=-n+1}^{n} q_M(\omega_j,\vartheta)I_{n,Y}(\omega_j)\right\| \\
&\leq q\frac{\epsilon\mathfrak{C}}{2n}\sum_{j=-n+1}^{n}\sum_{k=1}^{m}\sum_{\ell=1}^{m}[I_{n,Y}(\omega_j)[k,k] + I_{n,Y}(\omega_j)[\ell,\ell]] \\
&\leq \frac{\epsilon\mathfrak{C}m}{n}\sum_{j=-n+1}^{n}\sum_{k=1}^{m} I_{n,Y}(\omega_j)[k,k] \\
&\leq 2\epsilon\mathfrak{C}m\sum_{k=1}^{m}\overline{\Gamma}_{n,Y}(0)[k,k].
\end{aligned}$$

Since $\sum_{k=1}^m \overline{\Gamma}_{n,Y}(0)[k,k] \xrightarrow{a.s.} \sum_{k=1}^m \Gamma_Y(0)[k,k] < \infty$ due to Lemma 2.14, we obtain for $M \geq M_0(\epsilon)$ and n large

$$\sup_{\vartheta\in\Theta}\left\|\frac{1}{2n}\sum_{j=-n+1}^{n}\left(f_Y^{(\Delta)}(\omega_j,\vartheta)^{-1}I_{n,Y}(\omega_j)\right) - \frac{1}{2n}\sum_{j=-n+1}^{n} q_M(\omega_j,\vartheta)I_{n,Y}(\omega_j)\right\| \leq \varepsilon\mathfrak{C}$$

almost surely. Consequently, for the proof of (3.4) it is sufficient to show that

$$\sup_{\vartheta\in\Theta}\left\|\frac{1}{2n}\sum_{j=-n+1}^{n} q_M(\omega_j,\vartheta)I_{n,Y}(\omega_j) - \frac{1}{2\pi}\int_{-\pi}^{\pi} f_Y^{(\Delta)}(\omega,\vartheta)^{-1}f_Y^{(\Delta)}(\omega)d\omega\right\| \xrightarrow{a.s.} 0. \tag{3.9}$$

On the one hand, Lemma A.1 yields

$$\frac{1}{2n}\sum_{j=-n+1}^{n} q_M(\omega_j,\vartheta)I_{n,Y}(\omega_j)$$

$$
\begin{aligned}
&= \frac{1}{2\pi}\sum_{|k|<M}\sum_{|h|<n}\left(\left(1-\frac{|k|}{M}\right)\widehat{(f^{-1}(\vartheta))}_k\overline{\Gamma}_{n,Y}(h)\left(\frac{1}{2n}\sum_{j=-n+1}^{n}e^{-i(h-k)\omega_j}\right)\right)\\
&= \frac{1}{2\pi}\sum_{|k|<M}\left(1-\frac{|k|}{M}\right)\widehat{(f^{-1}(\vartheta))}_k\overline{\Gamma}_{n,Y}(k)\\
&\xrightarrow{a.s.} \frac{1}{2\pi}\sum_{|k|<M}\left(1-\frac{|k|}{M}\right)\widehat{(f^{-1}(\vartheta))}_k\Gamma_Y(k) \qquad (3.10)
\end{aligned}
$$

uniformly in ϑ, since $\widehat{(f^{-1}(\vartheta))}_k$ is uniformly bounded in ϑ for all k. The reason is that $f_Y^{(\Delta)}(\omega,\vartheta)^{-1}$ is continuous on the compact set $[-\pi,\pi]\times\Theta$ and

$$
\sup_{\substack{\vartheta\in\Theta\\ k\in\mathbb{Z}}}\left\|\widehat{(f^{-1}(\vartheta))}_k\right\| = \sup_{\substack{\vartheta\in\Theta\\ k\in\mathbb{Z}}}\left\|\frac{1}{2\pi}\int_{-\pi}^{\pi} f_Y^{(\Delta)}(\omega,\vartheta)^{-1}e^{-ik\omega}d\omega\right\| \leq \max_{\vartheta\in\Theta}\max_{\omega\in[-\pi,\pi]}\left\|f_Y^{(\Delta)}(\omega,\vartheta)^{-1}\right\|.
$$

On the other hand, due to (2.12), we get

$$
\begin{aligned}
&\left\|\frac{1}{2\pi}\sum_{|k|<M}\left(1-\frac{|k|}{M}\right)\widehat{(f^{-1}(\vartheta))}_k\Gamma_Y(k) - \frac{1}{2\pi}\int_{-\pi}^{\pi} f_Y^{(\Delta)}(\omega,\vartheta)^{-1}f_Y^{(\Delta)}(\omega)d\omega\right\|\\
&= \left\|\frac{1}{2\pi}\sum_{|k|<M}\left(1-\frac{|k|}{M}\right)\widehat{(f^{-1}(\vartheta))}_k\int_{-\pi}^{\pi} f_Y^{(\Delta)}(\omega)e^{ik\omega}d\omega - \frac{1}{2\pi}\int_{-\pi}^{\pi} f_Y^{(\Delta)}(\omega,\vartheta)^{-1}f_Y^{(\Delta)}(\omega)d\omega\right\|\\
&= \left\|\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(q_M(\omega,\vartheta)-f_Y^{(\Delta)}(\omega,\vartheta)^{-1}\right)f_Y^{(\Delta)}(\omega)d\omega\right\|\\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}\left\|q_M(\omega,\vartheta)-f_Y^{(\Delta)}(\omega,\vartheta)^{-1}\right\|\left\|f_Y^{(\Delta)}(\omega)\right\|d\omega \leq \epsilon\mathfrak{C}, \qquad (3.11)
\end{aligned}
$$

where we used (3.5) and the continuity of $f_Y^{(\Delta)}(\omega)$ for the last inequality. Combining (3.10) and (3.11) gives (3.9). □

Obviously, it is necessary that ϑ_0 is a global minimum of W to guarantee the consistency of the Whittle estimator.

Proposition 3.14.
Let Assumptions (A1)–(A4) and (A6) hold. Then, W has a unique global minimum in ϑ_0.

The proof is based on an alternative representation of W. Namely, the function W is exactly the limit function of the quasi maximum likelihood estimator of Schlemm and Stelzer (2012b).

Lemma 3.15.
Let Assumptions (A1)–(A4) hold and let $\xi_k^{(\Delta)}(\vartheta)=\Pi(\mathfrak{B},\vartheta)Y_k^{(\Delta)}$ with $\Pi(z,\vartheta)$ as given in Theorem 2.8. Furthermore, define

$$
\mathcal{L}(\vartheta) := \mathbb{E}\left[\operatorname{tr}\left(\xi_1^{(\Delta)}(\vartheta)^\top V^{(\Delta)}(\vartheta)^{-1}\xi_1^{(\Delta)}(\vartheta)\right)\right] + \log(\det(V^{(\Delta)}(\vartheta))) - m\log(2\pi), \quad \vartheta\in\Theta.
$$

Then, $W(\vartheta) = \mathcal{L}(\vartheta)$ *for* $\vartheta \in \Theta$.

Proof. In view of Theorem 2.8, we express the linear innovations as

$$\varepsilon_k^{(\Delta)}(\vartheta) = \Pi(\mathfrak{B}, \vartheta) Y_k^{(\Delta)}(\vartheta), \quad k \in \mathbb{N},$$

and define the pseudo innovations as

$$\xi_k^{(\Delta)}(\vartheta) := \Pi(\mathfrak{B}, \vartheta) Y_k^{(\Delta)}(\vartheta_0), \quad k \in \mathbb{N}.$$

An application of Theorem 11.8.3 of Brockwell and Davis (1991) leads to the spectral densities of $(\varepsilon_k^{(\Delta)}(\vartheta))_{k\in\mathbb{N}}$ and $(\xi_k^{(\Delta)}(\vartheta))_{k\in\mathbb{N}}$ as

$$\begin{aligned} f_\varepsilon^{(\Delta)}(\omega, \vartheta) &= \Pi(e^{-i\omega}, \vartheta) f_Y^{(\Delta)}(\omega, \vartheta) \Pi(e^{i\omega}, \vartheta)^\top, \quad \omega \in [-\pi, \pi], \\ f_\xi^{(\Delta)}(\omega, \vartheta) &= \Pi(e^{-i\omega}, \vartheta) f_Y^{(\Delta)}(\omega) \Pi(e^{i\omega}, \vartheta)^\top, \quad \omega \in [-\pi, \pi], \end{aligned}$$

respectively. Consequently,

$$\begin{aligned} &\frac{1}{2\pi} \int_{-\pi}^{\pi} \operatorname{tr}\left(f_Y^{(\Delta)}(\omega, \vartheta)^{-1} f_Y^{(\Delta)}(\omega) \right) d\omega \\ &\quad = \frac{1}{2\pi} \operatorname{tr}\left(\int_{-\pi}^{\pi} 2\pi \Pi(e^{i\omega}, \vartheta)^\top V^{(\Delta)}(\vartheta)^{-1} \Pi(e^{-i\omega}, \vartheta) f_Y^{(\Delta)}(\omega) d\omega \right) \\ &\quad = \operatorname{tr}\left(V^{(\Delta)}(\vartheta)^{-1} \int_{-\pi}^{\pi} f_\xi^{(\Delta)}(\omega, \vartheta) d\omega \right) \\ &\quad = \mathbb{E}\left[\operatorname{tr}\left(\xi_1^{(\Delta)}(\vartheta)^\top V^{(\Delta)}(\vartheta)^{-1} \xi_1^{(\Delta)}(\vartheta) \right) \right] \end{aligned}$$

holds. Finally,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left(\det\left(f_Y^{(\Delta)}(\omega, \vartheta) \right) \right) d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left(\det\left(2\pi f_Y^{(\Delta)}(\omega, \vartheta) \right) \right) d\omega - m \log(2\pi),$$

and an application of Theorem 3''' of Chapter 3 of Hannan (2009) results in

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left(\det\left(2\pi f_Y^{(\Delta)}(\omega, \vartheta) \right) \right) d\omega - m \log(2\pi) = \log\left(\det\left(V^{(\Delta)}(\vartheta) \right) \right) - m \log(2\pi),$$

which completes the proof. □

Proof of Proposition 3.14 Considering Lemma 3.15 we get $W(\vartheta) = \mathcal{L}(\vartheta)$. Schlemm and Stelzer (2012b), Lemma 2.10, proved that $\mathcal{L}$ has a unique global minimum in ϑ_0 under conditions which are fulfilled in our setting (see Lemma 2.3 and Lemma 3.14 of Schlemm and Stelzer (2012b)). □

Proof of Theorem 3.8. Due to Proposition 3.13 and Proposition 3.14, we know that the Whittle function W_n converges almost surely uniformly to W and that W has a unique global minimum in ϑ_0. It remains to show that the minimizing arguments of W_n converge almost surely to the minimizer of W. To that effect, we first prove

$$W_n(\widehat{\vartheta}_n^{(\Delta)}) \xrightarrow{a.s.} W(\vartheta_0) \tag{3.12}$$

and deduce that for every neighborhood U of ϑ_0 Whittle‘s estimate $\widehat{\vartheta}_n^{(\Delta)}$ lies in U almost surely for n large enough.

In view of Proposition 3.13, for all $\epsilon > 0$ there exists some $n_0 \in \mathbb{N}$ with

$$\sup_{\vartheta\in\Theta} |W_n(\vartheta) - W(\vartheta)| \leq \epsilon \quad \forall\, n \geq n_0 \quad \mathbb{P}\text{-a.s.} \tag{3.13}$$

Therefore, using the definition of $\widehat{\vartheta}_n^{(\Delta)}$ and Proposition 3.14, we get for $n \geq n_0$

$$\begin{aligned}
&W_n(\widehat{\vartheta}_n^{(\Delta)}) \leq W_n(\vartheta_0) \leq W(\vartheta_0) + \epsilon \quad \mathbb{P}\text{-a.s.} \quad \text{and} \\
&W_n(\widehat{\vartheta}_n^{(\Delta)}) \geq W(\widehat{\vartheta}_n^{(\Delta)}) - \epsilon \geq W(\vartheta_0) - \epsilon \quad \mathbb{P}\text{-a.s.}
\end{aligned}$$

and hence,

$$\sup_{n\geq n_0} |W_n(\widehat{\vartheta}_n^{(\Delta)}) - W(\vartheta_0)| \leq \epsilon \quad \mathbb{P}\text{-a.s.}$$

follows. This gives the desired convergence (3.12). Now, define $\delta(U) := \inf_{\vartheta\in\Theta\setminus U} W(\vartheta) - W(\vartheta_0) > 0$ for any neighborhood U of ϑ_0. The inequalities

$$\begin{aligned}
&\mathbb{P}\left(\lim_{n\to\infty} \widehat{\vartheta}_n^{(\Delta)} = \vartheta_0\right) = \mathbb{P}\left(\forall\, U\ \exists\, n_0(U) \in \mathbb{N}:\ \widehat{\vartheta}_n^{(\Delta)} \in U\ \forall\, n \geq n_0(U)\right) \\
&\geq\ \mathbb{P}\Bigg(\forall\, U\ \exists\, n_0(U) \in \mathbb{N}:\ |W_n(\widehat{\vartheta}_n^{(\Delta)}) - W(\vartheta_0)| < \frac{\delta(U)}{2} \\
&\qquad \text{and } |W_n(\widehat{\vartheta}_n^{(\Delta)}) - W(\widehat{\vartheta}_n^{(\Delta)})| < \frac{\delta(U)}{2}\ \forall\, n \geq n_0(U)\Bigg) = 1,
\end{aligned}$$

where the last equality follows from (3.12) and Proposition 3.13, complete the proof. □

3.5.2. Proof of the asymptotic normality

The proof of the asymptotic normality of the Whittle estimator is based on a Taylor expansion of the gradient function $\nabla_\vartheta W_n$ around $\widehat{\vartheta}_n^{(\Delta)}$ in ϑ_0, i.e.,

$$\sqrt{n}\,[\nabla_\vartheta W_n(\vartheta_0)] = \sqrt{n}\left[\nabla_\vartheta W_n(\widehat{\vartheta}_n^{(\Delta)})\right] - \sqrt{n}(\widehat{\vartheta}_n^{(\Delta)} - \vartheta_0)^\top \left[\nabla_\vartheta^2 W_n(\vartheta_n^*)\right] \tag{3.14}$$

for an appropriate $\vartheta_n^* \in \Theta$ with $\|\vartheta_n^* - \vartheta_0\| \leq \|\widehat{\vartheta}_n^{(\Delta)} - \vartheta_0\|$. Since $\widehat{\vartheta}_n^{(\Delta)}$ minimizes W_n and converges almost surely to ϑ_0, which is in the interior of Θ (Assumption $(B1)$), $\nabla_\vartheta W_n(\widehat{\vartheta}_n^{(\Delta)}) = 0$. Hence, in the case of an invertible matrix $\nabla_\vartheta^2 W_n(\vartheta_n^*)$ we can rewrite (3.14) and obtain

$$\sqrt{n}(\widehat{\vartheta}_n^{(\Delta)} - \vartheta_0)^\top = -\sqrt{n}\,[\nabla_\vartheta W_n(\vartheta_0)]\left[\nabla_\vartheta^2 W_n(\vartheta_n^*)\right]^{-1}. \tag{3.15}$$

Therefore, we receive the asymptotic normality of the Whittle estimator from the asymptotic behavior of the individual components in (3.15). First, we investigate the asymptotic behavior of the Hessian matrix $\nabla^2_\vartheta W_n(\vartheta^*_n)$.

Proposition 3.16.
Let Assumptions (A1)–(A4) and (B3) hold and $\Sigma_{\nabla^2 W}$ be defined as in (3.2). Furthermore, let $(\vartheta^*_n)_{n\in\mathbb{N}}$ be a sequence in Θ with $\vartheta^*_n \xrightarrow{a.s.} \vartheta_0$ as $n\to\infty$. Then, as $n\to\infty$,

$$\nabla^2_\vartheta W_n(\vartheta^*_n) \xrightarrow{a.s.} \Sigma_{\nabla^2 W}.$$

Proof. Under the Assumptions (A1)–(A4) and (B3) the spectral density $f_Y^{(\Delta)}(\omega,\vartheta)$ and its inverse $f_Y^{(\Delta)}(\omega,\vartheta)^{-1}$ are three times continuously differentiable in ϑ (see Remark 3.6 and Remark 3.7). Furthermore,

$$\frac{\partial^2}{\partial\vartheta_k\partial\vartheta_\ell}\operatorname{tr}\left(f_Y^{(\Delta)}(\omega,\vartheta)^{-1}I_n(\omega)\right) = \operatorname{tr}\left(\frac{\partial^2}{\partial\vartheta_k\partial\vartheta_\ell}\left(f_Y^{(\Delta)}(\omega,\vartheta)^{-1}\right)I_n(\omega)\right), \quad k,\ell\in\{1,\ldots,r\}.$$

Therefore, the proof of

$$\sup_{\vartheta\in\Theta}\left\|\nabla^2_\vartheta W_n(\vartheta)-\nabla^2_\vartheta W(\vartheta)\right\| \xrightarrow{a.s.} 0$$

goes in the same way as the proof of Proposition 3.13. It remains to show that $\nabla^2_\vartheta W(\vartheta_0) = \Sigma_{\nabla^2 W}$. First, note that

$$\nabla^2_\vartheta W(\vartheta_0) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\nabla^2_\vartheta \operatorname{tr}(f_Y^{(\Delta)}(\omega,\vartheta_0)^{-1}f_Y^{(\Delta)}(\omega)) + \nabla^2_\vartheta\log\left(\det\left(f_Y^{(\Delta)}(\omega,\vartheta_0)\right)\right)d\omega. \quad (3.16)$$

On the one hand,

$$\begin{aligned}&\frac{1}{2\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(\frac{\partial^2}{\partial\vartheta_k\partial\vartheta_\ell}\left(f_Y^{(\Delta)}(\omega,\vartheta_0)^{-1}\right)f_Y^{(\Delta)}(\omega)\right)d\omega\\ &\quad=\frac{1}{2\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(2f_Y^{(\Delta)}(\omega)^{-1}\left(\frac{\partial}{\partial\vartheta_k}f_Y^{(\Delta)}(\omega,\vartheta_0)\right)f_Y^{(\Delta)}(\omega)^{-1}\left(\frac{\partial}{\partial\vartheta_\ell}f_Y^{(\Delta)}(\omega,\vartheta_0)\right)\right.\\ &\qquad\qquad\left.-f_Y^{(\Delta)}(\omega)^{-1}\left(\frac{\partial^2}{\partial\vartheta_k\partial\vartheta_\ell}f_Y^{(\Delta)}(\omega,\vartheta_0)\right)\right)d\omega \end{aligned} \quad (3.17)$$

holds. On the other hand, Jacobi's formula leads to

$$\begin{aligned}&\frac{\partial^2}{\partial\vartheta_k\partial\vartheta_\ell}\log(\det(f_Y^{(\Delta)}(\omega,\vartheta_0)))\\ &\quad=\operatorname{tr}\left(-f_Y^{(\Delta)}(\omega)^{-1}\left(\frac{\partial}{\partial\vartheta_k}f_Y^{(\Delta)}(\omega,\vartheta_0)\right)f_Y^{(\Delta)}(\omega)^{-1}\left(\frac{\partial}{\partial\vartheta_\ell}f_Y^{(\Delta)}(\omega,\vartheta_0)\right)\right)\\ &\qquad+\operatorname{tr}\left(f_Y^{(\Delta)}(\omega)^{-1}\left(\frac{\partial^2}{\partial\vartheta_k\partial\vartheta_\ell}f_Y^{(\Delta)}(\omega,\vartheta_0)\right)\right).\end{aligned} \quad (3.18)$$

Combining (3.16), (3.17), (3.18) and the property

$$\operatorname{vec}\left(A^H\right)^H\left(B^T\otimes C\right)\operatorname{vec}(D)=\operatorname{tr}\left(BACD\right) \tag{3.19}$$

for appropriate matrices A, B, C, D (see Brewer (1978), properties T2.4, T3.4 and T3.8) gives

$$\begin{aligned}\nabla^2_\vartheta W(\vartheta_0) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\nabla_\vartheta f_Y^{(\Delta)}(-\omega,\vartheta_0)^\top\left[f_Y^{(\Delta)}(-\omega)^{-1}\otimes f_Y^{(\Delta)}(\omega)^{-1}\right]\nabla_\vartheta f_Y^{(\Delta)}(\omega,\vartheta_0)d\omega\\ &=\Sigma_{\nabla^2 W}.\end{aligned}$$

□

Furthermore, we require that for large n the random matrix $\nabla^2_\vartheta W_n(\vartheta_n^*)$ is invertible. Therefore, we show the positive definiteness of the limit matrix $\Sigma_{\nabla^2 W}$.

Lemma 3.17.
Let Assumptions A and (B4) hold. Then, $\Sigma_{\nabla^2 W}$ is positive definite.

Proof. Let $c\in\mathbb{C}^r$ be fixed and ω^* as in $(B4)$. The continuity of $f_Y^{(\Delta)}(\omega)$ and its regularity imply for any ω in a neighborhood of ω^* that

$$\left\|\left(f_Y^{(\Delta)}(-\omega)^{-1/2}\otimes f_Y^{(\Delta)}(\omega)^{-1/2}\right)\nabla_\vartheta f_Y^{(\Delta)}(\omega,\vartheta_0)c\right\|_2>0$$

where $\|\cdot\|_2$ is the Euclidean norm. Consequently,

$$\begin{aligned}c^\top\Sigma_{\nabla^2 W}c &= \frac{1}{2\pi}\int_{-\pi}^{\pi}c^\top\nabla_\vartheta f_Y^{(\Delta)}(\omega,\vartheta_0)^H\left[f_Y^{(\Delta)}(-\omega)^{-1}\otimes f_Y^{(\Delta)}(\omega)^{-1}\right]\nabla_\vartheta f_Y^{(\Delta)}(\omega,\vartheta_0)c\,d\omega\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left\|\left(f_Y^{(\Delta)}(-\omega)^{-1/2}\otimes f_Y^{(\Delta)}(\omega)^{-1/2}\right)\nabla_\vartheta f_Y^{(\Delta)}(\omega,\vartheta_0)c\right\|_2^2 d\omega>0.\end{aligned}$$

Therefore, $\Sigma_{\nabla^2 W}$ is positive definite. □

Next, we investigate the asymptotic behavior of the second term in (3.15). Since the components of the score $\nabla_\vartheta W_n(\vartheta_0)$ can be written as an integrated periodogram, we first derive the asymptotic behavior of the integrated periodogram and state the asymptotic normality afterwards.

Proposition 3.18.
Let Assumptions $(A2)$–$(A4)$ and $(B2)$ hold. Suppose $\eta:[-\pi,\pi]\to\mathbb{C}^{m\times m}$ is a symmetric matrix-valued continuous function, i.e. $\eta(\omega)=\eta(-\omega)^\top$. Assume further that the Fourier coefficients $(\widehat{\eta}_u)_{u\in\mathbb{Z}}$ of η satisfy $\sum_{u=-\infty}^{\infty}\|\widehat{\eta}_u\||u|^{1/2}<\infty$. Then, as $n\to\infty$,

$$\frac{1}{2\sqrt{n}}\sum_{j=-n+1}^{n}\operatorname{tr}\left(\eta(\omega_j)I_{n,Y}(\omega_j)-\eta(\omega_j)f_Y^{(\Delta)}(\omega_j)\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,\Sigma_\eta),$$

where

$$\begin{aligned}\Sigma_\eta &= \frac{1}{\pi}\int_{-\pi}^{\pi} \operatorname{tr}\left(\eta(\omega) f_Y^{(\Delta)}(\omega)\eta(\omega) f_Y^{(\Delta)}(\omega)\right) d\omega \\ &+ \frac{1}{16\pi^4}\int_{-\pi}^{\pi} \operatorname{vec}\left(\Phi(e^{-i\omega})^\top \eta(\omega)^\top \Phi(e^{i\omega})\right)^\top d\omega \left(\mathbb{E}\left[N_1^{(\Delta)} N_1^{(\Delta)\top} \otimes N_1^{(\Delta)} N_1^{(\Delta)\top}\right]\right. \\ &\left. -3\Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)}\right) \int_{-\pi}^{\pi} \operatorname{vec}\left(\Phi(e^{i\omega})^\top \eta(\omega)\Phi(e^{-i\omega})\right) d\omega.\end{aligned}$$

The asymptotic behavior of the integrated periodogram is interesting for its own. In the fifth chapter, we modify it and investigate a generalized version. Thereby, we can not only derive the asymptotic normality of the Whittle estimator, but we also obtain the behavior of some goodness-of-fit tests in the frequency domain.

Remark 3.19.
Let the driving Lévy process be a Brownian motion. Since the fourth moment of a centered normal distribution is equal to three times its second moment and $N_1^{(\Delta)} \sim \mathcal{N}(0, \Sigma_N^{(\Delta)})$, we get $\mathbb{E}[N_1^{(\Delta)} N_1^{(\Delta)\top} \otimes N_1^{(\Delta)} N_1^{(\Delta)\top}] = 3\Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)}$. Therefore, the matrix Σ_η in Proposition 3.18 reduces to

$$\Sigma_\eta = \frac{1}{\pi}\int_{-\pi}^{\pi} \operatorname{tr}\left(\eta(\omega) f_Y^{(\Delta)}(\omega)\eta(\omega) f_Y^{(\Delta)}(\omega)\right) d\omega,$$

which is for $m = 1$ equal to $\Sigma_\eta = \frac{1}{\pi}\int_{-\pi}^{\pi} \eta(\omega)^2 f_Y^{(\Delta)}(\omega)^2 d\omega$.

For the proof of Proposition 3.18 we require some auxiliary result.

Lemma 3.20.
Let Assumptions (A2)–(A4) hold and $\eta : [-\pi, \pi] \to \mathbb{C}^{m\times m}$ be a symmetric matrix-valued continuous function with Fourier coefficients $(\widehat{\eta}_k)_{k\in\mathbb{Z}}$ satisfying $\sum_{k=-\infty}^{\infty} \|\widehat{\eta}_k\| < \infty$. Then,

$$\lim_{n\to\infty} \mathbb{E}\left|\frac{1}{2\sqrt{n}}\sum_{j=-n+1}^{n} \operatorname{tr}\left(\eta(\omega_j) I_{n,Y}(\omega_j) - \eta(\omega_j)\Phi(e^{-i\omega_j}) I_{n,N}(\omega_j)\Phi(e^{i\omega_j})^\top\right)\right| = 0.$$

Proof. Define $R_n(\omega) = I_{n,Y}(\omega) - \Phi(e^{-i\omega}) I_{n,N}(\omega)\Phi(e^{i\omega})^\top$ for $\omega \in [-\pi, \pi]$. We get

$$\begin{aligned}R_n(\omega_j) &= \frac{1}{2\pi n}\left(\sum_{k=1}^{n}\sum_{s=0}^{\infty} \Phi_s N_{k-s}^{(\Delta)}\right)\left(\sum_{\ell=1}^{n}\sum_{t=0}^{\infty} \Phi_t N_{\ell-t}^{(\Delta)}\right)^\top e^{-i(k-\ell)\omega_j} \\ &\quad - \frac{1}{2\pi n}\left(\sum_{k=1}^{n}\sum_{s=0}^{\infty} \Phi_s N_{k}^{(\Delta)}\right)\left(\sum_{\ell=1}^{n}\sum_{t=0}^{\infty} \Phi_t N_{\ell}^{(\Delta)}\right)^\top e^{-i(k+s-\ell-t)\omega_j} \\ &= \frac{1}{2\pi n}\left(\sum_{s=0}^{\infty}\sum_{t=0}^{\infty} \Phi_s \left(\left(\sum_{k=1}^{n}\sum_{\ell=1-t}^{0} - \sum_{k=1}^{n}\sum_{\ell=n-t+1}^{n} + \sum_{k=1-s}^{0}\sum_{\ell=1}^{n} + \sum_{k=1-s}^{0}\sum_{\ell=1-t}^{0}\right.\right.\right. \\ &\quad \left. - \sum_{k=1-s}^{0}\sum_{\ell=n-t+1}^{n} - \sum_{k=n-s+1}^{n}\sum_{\ell=1}^{n} - \sum_{k=n-s+1}^{n}\sum_{\ell=1-t}^{0} + \sum_{k=n-s+1}^{n}\sum_{\ell=n-t+1}^{n}\right) \\ &\quad \left.\left. N_k^{(\Delta)} N_\ell^{(\Delta)\top} e^{-i(k+s-\ell-t)\omega_j}\right) \Phi_t^\top\right)\end{aligned}$$

$$=: \sum_{i=1}^{8} R_n^{(i)}(\omega_j).$$

Thus,

$$\mathbb{E}\left|\frac{1}{2\sqrt{n}}\sum_{j=-n+1}^{n} \operatorname{tr}\left(\eta(\omega_j)R_n(\omega_j)\right)\right| \leq \sum_{i=1}^{8}\mathbb{E}\left|\frac{1}{2\sqrt{n}}\sum_{j=-n+1}^{n} \operatorname{tr}\left(\eta(\omega_j)R_n^{(i)}(\omega_j)\right)\right|.$$

We have to show that these 8 components converge to zero. Since we can treat each component similarly, we only give the detailed proof for the convergence of the first term.

Due to $\operatorname{tr}(A) \leq \|A\|_1$ for all quadratic matrices A, we get an upper bound for the trace of any quadratic matrix. Once again, the equivalence of all matrix norms and $\eta(\omega_j) = \sum_{u=-\infty}^{\infty} \widehat{\eta}_u e^{i\omega_j u}$ yield

$$\begin{aligned}
&\mathbb{E}\left|\frac{1}{2\sqrt{n}}\sum_{j=-n+1}^{n} \operatorname{tr}\left(R_n^{(1)}(\omega_j)\eta(\omega_j)\right)\right| \\
&\quad\leq \mathfrak{C}\mathbb{E}\left\|\sum_{j=-n+1}^{n}\frac{1}{\sqrt{n}}\frac{1}{n}\sum_{s=0}^{\infty}\sum_{t=0}^{\infty}\Phi_s\sum_{k=1}^{n}\sum_{\ell=1-t}^{0} N_k^{(\Delta)}N_\ell^{(\Delta)\top}\Phi_t^\top\sum_{u=-\infty}^{\infty}\widehat{\eta}_u e^{-i(k+s-\ell-t-u)\omega_j}\right\| \\
&\quad\leq \mathfrak{C}\frac{1}{\sqrt{n}}\sum_{s=0}^{\infty}\sum_{t=0}^{\infty}\|\Phi_s\|\sum_{k=1}^{n}\sum_{\ell=1-t}^{0}\mathbb{E}\|N_1^{(\Delta)}\|^2\|\Phi_t\|\sum_{u=-\infty}^{\infty}\|\widehat{\eta}_u\|\frac{1}{n}\left\|\sum_{j=-n+1}^{n} e^{-i(k+s-\ell-t-u)\omega_j}\right\|.
\end{aligned}$$

Due to $(A2)$, $\mathbb{E}\|N_1^{(\Delta)}\|^2 < \infty$. Further, an application of Lemma A.1 gives

$$\mathbb{E}\left|\frac{1}{2\sqrt{n}}\sum_{j=-n+1}^{n} \operatorname{tr}\left(R_n^{(1)}(\omega_j)\eta(\omega_j)\right)\right| \leq \mathfrak{C}\frac{1}{\sqrt{n}}\sum_{s=0}^{\infty}\|\Phi_s\|\sum_{t=0}^{\infty} t\|\Phi_t\|\sum_{u=-\infty}^{\infty}\|\widehat{\eta}_u\| \overset{n\to\infty}{\longrightarrow} 0.$$

□

This lemma helps to deduce Proposition 3.18, which can be seen as the main part of the proof of the asymptotic normality of the Whittle estimator.

Proof of Proposition 3.18. Due to Lemma 3.20, we get

$$\begin{aligned}
&\frac{1}{2\sqrt{n}}\sum_{j=-n+1}^{n} \operatorname{tr}\left(\eta(\omega_j)I_{n,Y}(\omega_j) - \eta(\omega_j)f_Y^{(\Delta)}(\omega_j)\right) \\
&= \frac{1}{2\sqrt{n}}\sum_{j=-n+1}^{n}\left\{\operatorname{tr}\left(I_{n,N}(\omega_j)\Phi(e^{i\omega_j})^\top\eta(\omega_j)\Phi(e^{-i\omega_j})\right) - \operatorname{tr}\left(\eta(\omega_j)f_Y^{(\Delta)}(\omega_j)\right)\right\} + o_{\mathbb{P}}(1).
\end{aligned}$$

We define

$$q(\omega) := \Phi(e^{i\omega})^\top\eta(\omega)\Phi(e^{-i\omega}), \quad \omega \in [-\pi, \pi],$$

and approximate q by its Fourier series of degree M, namely,

$$q_M(\omega) = \sum_{|k|\leq M} \widehat{q}_k e^{ik\omega} \quad \text{where} \quad \widehat{q}_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} q(\omega)e^{-ik\omega}d\omega, \quad k \in \mathbb{Z}. \tag{3.20}$$

The coefficients $\widehat{q}_k$, $k \in \mathbb{Z}$, satisfy

$$\begin{aligned}
&\sum_{k=-\infty}^{\infty} \|\widehat{q}_k\||k|^{1/2} \\
=\ & \sum_{k=-\infty}^{\infty} \left\| \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi(e^{i\omega})^\top \eta(\omega)\Phi(e^{-i\omega})e^{-ik\omega}d\omega \right\| |k|^{1/2} \\
=\ & \sum_{k=-\infty}^{\infty} \left\| \frac{1}{2\pi}\sum_{j=0}^{\infty}\sum_{\ell=0}^{\infty}\sum_{u=-\infty}^{\infty} \Phi_j^\top \widehat{\eta}_u \Phi_\ell \int_{-\pi}^{\pi} e^{-i(k-j+u+\ell)\omega}d\omega \right\| |k|^{1/2} \\
\leq\ & \sum_{j=0}^{\infty}\sum_{\ell=0}^{\infty}\sum_{u=-\infty}^{\infty} \|\Phi_j\|\|\widehat{\eta}_u\|\|\Phi_\ell\||j-u-\ell|^{1/2} \\
\leq\ & \mathfrak{C}\sum_{j=0}^{\infty} \|\Phi_j\| \left(\max\{1,|j|\}\right)^{1/2} \sum_{u=-\infty}^{\infty} \|\widehat{\eta}_u\| \left(\max\{1,|u|\}\right)^{1/2} \sum_{\ell=0}^{\infty} \|\Phi_\ell\| \left(\max\{1,|\ell|\}\right)^{1/2} \\
<\ & \infty,
\end{aligned} \tag{3.21}$$

and therefore $\sum_{k=-\infty}^{\infty} \|\widehat{q}_k\| < \infty$ as well. An application of Theorem A.2 leads to

$$q_M(\omega) \stackrel{M\to\infty}{\longrightarrow} q(\omega) \quad \text{uniformly in } \omega \in [-\pi,\pi].$$

Step 1: We show

$$\lim_{M\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\frac{1}{\sqrt{n}} \left| \sum_{j=-n+1}^{n} \operatorname{tr}(I_{n,N}(\omega_j)(q(\omega_j) - q_M(\omega_j)) \right| > \epsilon \right) = 0 \quad \forall\, \epsilon > 0. \tag{3.22}$$

Consider

$$\begin{aligned}
&\frac{1}{\sqrt{n}} \sum_{j=-n+1}^{n} \operatorname{tr}\left(I_{n,N}(\omega_j)(q(\omega_j) - q_M(\omega_j))\right) \\
&\quad = \frac{\sqrt{n}}{\pi} \sum_{|k|>M} \operatorname{tr}\left(\sum_{h=-n+1}^{n-1} \overline{\Gamma}_{n,N}(h)\widehat{q}_k \left(\frac{1}{2n}\sum_{j=-n+1}^{n} e^{-i(h-k)\omega_j} \right)\right).
\end{aligned} \tag{3.23}$$

We investigate the terms with $h = 0$ and $h \neq 0$ separately. For $h = 0$ and $n > M$ we get

$$\left| \frac{\sqrt{n}}{\pi} \sum_{|k|>M} \operatorname{tr}\left(\overline{\Gamma}_{n,N}(0)\widehat{q}_k \mathbf{1}_{\{\exists z\in\mathbb{Z}\setminus\{0\}\,:\,k=2nz\}}\right)\right| \leq \mathfrak{C}\sqrt{n}\|\overline{\Gamma}_{n,N}(0)\| \sum_{|k|\geq 2n} \|\widehat{q}_k\| \stackrel{n\to\infty}{\longrightarrow} 0 \quad \mathbb{P}\text{-a.s.,} \tag{3.24}$$

since Remark 2.15 and the continuous mapping theorem imply $\|\overline{\Gamma}_{n,N}(0)\| \stackrel{a.s.}{\to} \|\Sigma_N^{(\Delta)}\|$. Now, we investigate the terms with $h \neq 0$. The independence of the sequence $(N_k^{(\Delta)})_{k\in\mathbb{Z}}$ leads to

$$\mathbb{E}\left[\overline{\Gamma}_{n,N}(h)\right] = 0 \quad \text{for } h \neq 0$$

and therefore,

$$\mathbb{E}\left[\sqrt{n}\sum_{|k|>M}\operatorname{tr}\left(\left(\sum_{h=1}^{n-1}\overline{\Gamma}_{n,N}(h) + \sum_{h=-n+1}^{-1}\overline{\Gamma}_{n,N}(h)\right)\widehat{q}_k\mathbf{1}_{\{\exists z\in\mathbb{Z}\,:\,h=k+2nz\}}\right)\right] = 0. \tag{3.25}$$

Due to (3.23)-(3.25) and the Tschebycheff inequality, for the proof of (3.22) it is sufficient to show that

$$\lim_{M\to\infty}\lim_{n\to\infty}\operatorname{Var}\left(\sqrt{n}\sum_{|k|>M}\operatorname{tr}\left(\left(\sum_{h=1}^{n-1}\overline{\Gamma}_{n,N}(h) + \sum_{h=-n+1}^{-1}\overline{\Gamma}_{n,N}(h)\right)\widehat{q}_k\mathbf{1}_{\{\exists z\in\mathbb{Z}\,:\,h=k+2nz\}}\right)\right) = 0. \tag{3.26}$$

First, property (3.19) and $\mathbb{E}\left\|\operatorname{vec}\left(\overline{\Gamma}_{n,N}(h)\right)\operatorname{vec}\left(\overline{\Gamma}_{n,N}(h)\right)^\top\right\| \leq \frac{\mathfrak{C}}{n}$ result in

$$\begin{aligned}
&\operatorname{Var}\left(\sqrt{n}\sum_{|k|>M}\operatorname{tr}\left(\left(\sum_{h=1}^{n-1}\overline{\Gamma}_{n,N}(h) + \sum_{h=-n+1}^{-1}\overline{\Gamma}_{n,N}(h)\right)\widehat{q}_k\mathbf{1}_{\{\exists z\in\mathbb{Z}\,:\,h=k+2nz\}}\right)\right)\\
&= \operatorname{Var}\left(2\sqrt{n}\sum_{|k|>M}\operatorname{tr}\left(\sum_{h=1}^{n-1}\overline{\Gamma}_{n,N}(h)\widehat{q}_k\mathbf{1}_{\{\exists z\in\mathbb{Z}\,:\,h=k+2nz\}}\right)\right)\\
&= \operatorname{Var}\left(2\sqrt{n}\sum_{h=1}^{n-1}\operatorname{vec}\left(\sum_{|k|>M}\widehat{q}_k^\top\mathbf{1}_{\{\exists z\in\mathbb{Z}\,:\,h=k+2nz\}}\right)^\top (I_N\otimes I_N)\operatorname{vec}\left(\overline{\Gamma}_{n,N}(h)\right)\right)\\
&\leq 4n\sum_{h=1}^{n-1}\left\|\operatorname{vec}\left(\sum_{|k|>M}\widehat{q}_k^\top\mathbf{1}_{\{\exists z\in\mathbb{Z}\,:\,h=k+2nz\}}\right)\right\|^2\|I_N\otimes I_N\|^2\\
&\qquad\left\|\mathbb{E}\left[\operatorname{vec}\left(\overline{\Gamma}_{n,N}(h)\right)\operatorname{vec}\left(\overline{\Gamma}_{n,N}(h)\right)^\top\right]\right\|\\
&\leq \mathfrak{C}\sum_{h=1}^{n-1}\left\|\sum_{|k|>M}\widehat{q}_k\mathbf{1}_{\{\exists z\in\mathbb{Z}\,:\,h=k+2nz\}}\right\|^2 \leq \mathfrak{C}\left(\sum_{|k|>M}\|\widehat{q}_k\|\right)^2 \overset{M\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Step 2: We show

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{j=-n+1}^{n}\left(\operatorname{tr}\left(I_{n,N}(\omega_j)q_M(\omega_j)\right) - \operatorname{tr}\left(\eta(\omega_j)f_Y^{(\Delta)}(\omega_j)\right)\right)\\
&\quad= \frac{\sqrt{n}}{\pi}\operatorname{tr}\left(\sum_{h=-M}^{M}\left(\overline{\Gamma}_{n,N}(h) - \Gamma_N(h)\right)\widehat{q}_h\right) + o_{\mathbb{P}}(1).
\end{aligned} \tag{3.27}$$

Let $M > n$. Then, due to Lemma A.7 and Parseval's equality, we obtain

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{j=-n+1}^{n}\left(\operatorname{tr}\left(I_{n,N}(\omega_j)q_M(\omega_j)\right)-\operatorname{tr}\left(\eta(\omega_j)f_Y^{(\Delta)}(\omega_j)\right)\right)\\
&= \frac{\sqrt{n}}{\pi}\operatorname{tr}\left(\sum_{h=-M}^{M}\overline{\Gamma}_{n,N}(h)\widehat{q}_h\right)-\frac{\sqrt{n}}{\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(\eta(\omega)f_Y^{(\Delta)}(\omega)\right)d\omega\\
&\quad+\frac{\sqrt{n}}{\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(\eta(\omega)f_Y^{(\Delta)}(\omega)\right)d\omega-\operatorname{tr}\left(\frac{1}{\sqrt{n}}\sum_{j=-n+1}^{n}\eta(\omega_j)f_Y^{(\Delta)}(\omega_j)\right)\\
&= \frac{\sqrt{n}}{\pi}\operatorname{tr}\left(\sum_{h=-M}^{M}\overline{\Gamma}_{n,N}(h)\widehat{q}_h\right)-\frac{\sqrt{n}}{\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(\eta(\omega)f_Y^{(\Delta)}(\omega)\right)d\omega+o(1). \qquad (3.28)
\end{aligned}$$

Taking $\Gamma_N(h) = 0_{N\times N}$ for $h \neq 0$ into account, we receive

$$\begin{aligned}
&\frac{\sqrt{n}}{\pi}\operatorname{tr}\left(\sum_{h=-M}^{M}\overline{\Gamma}_{n,N}(h)\widehat{q}_h\right)-\frac{\sqrt{n}}{\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(\eta(\omega)f_Y^{(\Delta)}(\omega)\right)d\omega\\
&= \frac{\sqrt{n}}{\pi}\operatorname{tr}\left(\sum_{h=-M}^{M}\left(\overline{\Gamma}_{n,N}(h)-\Gamma_N(h)\right)\widehat{q}_h\right)\\
&\quad+\frac{\sqrt{n}}{\pi}\left(\operatorname{tr}(\Gamma_N(0)\widehat{q}_0)-\int_{-\pi}^{\pi}\operatorname{tr}\left(\eta(\omega)f_Y^{(\Delta)}(\omega)\right)d\omega\right). \qquad (3.29)
\end{aligned}$$

Using the representation $f_Y^{(\Delta)}(\omega) = \frac{1}{2\pi}\Phi(e^{-i\omega})\Sigma_N^{(\Delta)}\Phi(e^{i\omega})^\top$ and $q(\omega) = \Phi(e^{i\omega})^\top\eta(\omega)\Phi(e^{-i\omega})$ for $\omega \in [-\pi, \pi]$, yield

$$\begin{aligned}
&\frac{\sqrt{n}}{\pi}\operatorname{tr}\left(\Gamma_N(0)\widehat{q}_0\right)-\operatorname{tr}\left(\int_{-\pi}^{\pi}\eta(\omega)f_Y^{(\Delta)}(\omega)d\omega\right)\\
&= \frac{\sqrt{n}}{\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(\frac{1}{2\pi}\Sigma_N^{(\Delta)}q(\omega)\right)-\operatorname{tr}\left(\eta(\omega)f_Y^{(\Delta)}(\omega)\right)d\omega\\
&= \frac{\sqrt{n}}{\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(\eta(\omega)\frac{1}{2\pi}\Phi(e^{-i\omega})\Sigma_N^{(\Delta)}\Phi(e^{i\omega})^\top-\eta(\omega)f_Y^{(\Delta)}(\omega)\right)d\omega = 0. \qquad (3.30)
\end{aligned}$$

Then, (3.28)-(3.30) result in (3.27).

Step 3: Next, we prove the asymptotic normality

$$\frac{\sqrt{n}}{2\pi}\operatorname{tr}\left(\sum_{h=-M}^{M}\left(\overline{\Gamma}_{n,N}(h)-\Gamma_N(h)\right)\widehat{q}_h\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,\Sigma_\eta(M)), \qquad (3.31)$$

where $\Sigma_\eta(M)$ is defined as

$$\begin{aligned}
\Sigma_\eta(M) &= \frac{1}{\pi^2}\sum_{h=1}^{M}\operatorname{tr}\left(\widehat{q}_h\Sigma_N^{(\Delta)}\widehat{q}_h^H\Sigma_N^{(\Delta)}\right)\\
&\quad+\frac{1}{4\pi^2}\operatorname{vec}(\widehat{q}_0^\top)^\top\left(\mathbb{E}\left[N_1^{(\Delta)}N_1^{(\Delta)\top}\otimes N_1^{(\Delta)}N_1^{(\Delta)\top}\right]-\Sigma_N^{(\Delta)}\otimes\Sigma_N^{(\Delta)}\right)\operatorname{vec}(\widehat{q}_0^H).
\end{aligned}$$

Therefore, we consider

$$\begin{aligned}&\frac{\sqrt{n}}{2\pi}\operatorname{tr}\left(\sum_{h=-M}^{M}\left(\overline{\Gamma}_{n,N}(h)-\Gamma_N(h)\right)\widehat{q}_h\right)\\&\quad=\frac{1}{\pi}\sum_{h=1}^{M}\sqrt{n}\operatorname{tr}\left(\left(\overline{\Gamma}_{n,N}(h)-\Gamma_N(h)\right)\widehat{q}_h\right)+\frac{\sqrt{n}}{2\pi}\operatorname{tr}\left(\left(\overline{\Gamma}_{n,N}(0)-\Gamma_N(0)\right)\widehat{q}_0\right).\end{aligned}\tag{3.32}$$

Writing

$$\sqrt{n}\operatorname{tr}\left(\left(\overline{\Gamma}_{n,N}(h)-\Gamma_N(h)\right)\widehat{q}_h\right)=\sqrt{n}\operatorname{vec}(\widehat{q}_h^\top)^\top\operatorname{vec}\left(\overline{\Gamma}_{n,N}(h)-\Gamma_N(h)\right),$$

an application of Lemma 2.16 leads to

$$\sqrt{n}\operatorname{tr}\left(\left(\overline{\Gamma}_{n,N}(h)-\Gamma_N(h)\right)\widehat{q}_h\right)\xrightarrow{\mathcal{D}}\mathcal{N}_h,$$

where $(\mathcal{N}_h)_{h\in\mathbb{N}_0}$ is an independent centered normally distributed sequence of random vectors with covariance matrices

$$\Sigma_{\mathcal{N}_h}:=\operatorname{vec}(\widehat{q}_h^\top)^\top\left(\Sigma_N^{(\Delta)}\otimes\Sigma_N^{(\Delta)}\right)\operatorname{vec}(\widehat{q}_h^H)=\operatorname{tr}\left(\widehat{q}_h\Sigma_N^{(\Delta)}\widehat{q}_h^H\Sigma_N^{(\Delta)}\right)\quad\text{for } h\neq 0$$

and

$$\Sigma_{\mathcal{N}_0}:=\operatorname{vec}(\widehat{q}_0^\top)^\top\left(\mathbb{E}\left[N_1^{(\Delta)}N_1^{(\Delta)\top}\otimes N_1^{(\Delta)}N_1^{(\Delta)\top}\right]-\Sigma_N^{(\Delta)}\otimes\Sigma_N^{(\Delta)}\right)\operatorname{vec}(\widehat{q}_0^H).$$

Finally,

$$\begin{aligned}&\frac{\sqrt{n}}{2\pi}\operatorname{tr}\left(\sum_{h=-M}^{M}\left(\overline{\Gamma}_{n,N}(h)-\Gamma_N(h)\right)\widehat{q}_h\right)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\frac{1}{\pi^2}\sum_{h=1}^{M}\operatorname{tr}\left(\widehat{q}_h\Sigma_N^{(\Delta)}\widehat{q}_h^H\Sigma_N^{(\Delta)}\right)\right.\\&\quad\left.+\frac{1}{4\pi^2}\operatorname{vec}(\widehat{q}_0^\top)^\top\left(\mathbb{E}\left[N_1^{(\Delta)}N_1^{(\Delta)\top}\otimes N_1^{(\Delta)}N_1^{(\Delta)\top}\right]-\Sigma_N^{(\Delta)}\otimes\Sigma_N^{(\Delta)}\right)\operatorname{vec}(\widehat{q}_0^H)\right).\end{aligned}$$

Step 4: We show

$$\begin{aligned}&\frac{1}{\pi^2}\sum_{h=1}^{M}\operatorname{tr}\left(\widehat{q}_h\Sigma_N^{(\Delta)}\widehat{q}_h^H\Sigma_N^{(\Delta)}\right)\\&\qquad+\frac{1}{4\pi^2}\operatorname{vec}(\widehat{q}_0^\top)^\top\left(\mathbb{E}\left[N_1^{(\Delta)}N_1^{(\Delta)\top}\otimes N_1^{(\Delta)}N_1^{(\Delta)\top}\right]-\Sigma_N^{(\Delta)}\otimes\Sigma_N^{(\Delta)}\right)\operatorname{vec}(\widehat{q}_0^H)\\&\xrightarrow{M\to\infty}\frac{1}{\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(\eta(\omega)f_Y^{(\Delta)}(\omega)\eta(\omega)f_Y^{(\Delta)}(\omega)\right)d\omega\\&\quad+\frac{1}{16\pi^4}\int_{-\pi}^{\pi}\operatorname{vec}\left(\Phi(e^{-i\omega})^\top\eta(\omega)^\top\Phi(e^{i\omega})\right)^\top d\omega\\&\qquad\qquad\left(\mathbb{E}\left[N_1^{(\Delta)}N_1^{(\Delta)\top}\otimes N_1^{(\Delta)}N_1^{(\Delta)\top}\right]-3\Sigma_N^{(\Delta)}\otimes\Sigma_N^{(\Delta)}\right)\\&\qquad\int_{-\pi}^{\pi}\operatorname{vec}\left(\Phi(e^{i\omega})^\top\eta(\omega)\Phi(e^{-i\omega})\right)d\omega.\end{aligned}\tag{3.33}$$

Therefore, note that

$$\begin{aligned}
&\frac{1}{\pi^2}\sum_{h=1}^{M}\operatorname{tr}\left(\widehat{q}_h\Sigma_N^{(\Delta)}\widehat{q}_h^H\Sigma_N^{(\Delta)}\right)\\
&\qquad+\frac{1}{4\pi^2}\operatorname{vec}(\widehat{q}_0^\top)^\top\left(\mathbb{E}\left[N_1^{(\Delta)}N_1^{(\Delta)\top}\otimes N_1^{(\Delta)}N_1^{(\Delta)\top}\right]-\Sigma_N^{(\Delta)}\otimes\Sigma_N^{(\Delta)}\right)\operatorname{vec}(\widehat{q}_0^H)\\
\overset{M\to\infty}{\longrightarrow}\;&\frac{1}{\pi^2}\sum_{h=1}^{\infty}\operatorname{tr}\left(\widehat{q}_h\Sigma_N^{(\Delta)}\widehat{q}_h^H\Sigma_N^{(\Delta)}\right)\\
&\qquad+\frac{1}{4\pi^2}\operatorname{vec}(\widehat{q}_0^\top)^\top\left(\mathbb{E}\left[N_1^{(\Delta)}N_1^{(\Delta)\top}\otimes N_1^{(\Delta)}N_1^{(\Delta)\top}\right]-\Sigma_N^{(\Delta)}\otimes\Sigma_N^{(\Delta)}\right)\operatorname{vec}(\widehat{q}_0^H).
\end{aligned}$$

But

$$\begin{aligned}
\frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\operatorname{tr}\left(\widehat{q}_h\Sigma_N^{(\Delta)}\widehat{q}_h^H\Sigma_N^{(\Delta)}\right) &= \frac{1}{4\pi^2}\sum_{h=-\infty}^{\infty}\sum_{\ell=-\infty}^{\infty}\operatorname{tr}\left(\widehat{q}_h\Sigma_N^{(\Delta)}\widehat{q}_\ell^H\Sigma_N^{(\Delta)}\right)\int_{-\pi}^{\pi}e^{i(h-\ell)\omega}d\omega\\
&= \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\operatorname{tr}\left(q(\omega)\Sigma_N^{(\Delta)}q(\omega)^H\Sigma_N^{(\Delta)}\right)d\omega\\
&= \int_{-\pi}^{\pi}\operatorname{tr}\left(\eta(\omega)f_Y^{(\Delta)}(\omega)\eta(\omega)^H f_Y^{(\Delta)}(\omega)\right)d\omega,
\end{aligned}$$

where we plugged in the definition of q in the last equality. Eventually, due to the representation of $\widehat{q}_0$, we receive

$$\begin{aligned}
&\frac{1}{\pi^2}\sum_{h=1}^{\infty}\operatorname{tr}\left(\widehat{q}_h\Sigma_N^{(\Delta)}\widehat{q}_h^H\Sigma_N^{(\Delta)}\right)\\
&+\frac{1}{4\pi^2}\operatorname{vec}(\widehat{q}_0^\top)^\top\left(\mathbb{E}\left[N_1^{(\Delta)}N_1^{(\Delta)\top}\otimes N_1^{(\Delta)}N_1^{(\Delta)\top}\right]-\Sigma_N^{(\Delta)}\otimes\Sigma_N^{(\Delta)}\right)\operatorname{vec}(\widehat{q}_0^H)\\
=\;&\frac{1}{\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(\eta(\omega)f_Y^{(\Delta)}(\omega)\eta(\omega)f_Y^{(\Delta)}(\omega)\right)d\omega+\frac{1}{16\pi^4}\int_{-\pi}^{\pi}\operatorname{vec}\left(\Phi(e^{-i\omega})^\top\eta(\omega)^\top\Phi(e^{i\omega})\right)^\top d\omega\\
&\left(\mathbb{E}\left[N_1^{(\Delta)}N_1^{(\Delta)\top}\otimes N_1^{(\Delta)}N_1^{(\Delta)\top}\right]-3\Sigma_N^{(\Delta)}\otimes\Sigma_N^{(\Delta)}\right)\int_{-\pi}^{\pi}\operatorname{vec}\left(\Phi(e^{i\omega})^\top\eta(\omega)\Phi(e^{-i\omega})\right)d\omega.
\end{aligned}$$

Finally, Step 3, Step 4 and a multivariate version of Problem 6.16 of Brockwell and Davis (1991) give

$$\frac{\sqrt{n}}{2\pi}\operatorname{tr}\left(\sum_{h=-M}^{M}\left(\overline{\Gamma}_{n,N}(h)-\Gamma_N(h)\right)\widehat{q}_h\right)\overset{\mathcal{D},n\to\infty}{\longrightarrow}\mathcal{N}(0,\Sigma_\eta(M))\overset{\mathcal{D},M\to\infty}{\longrightarrow}\mathcal{N}(0,\Sigma_\eta).$$

Along with Step 1, Step 2 and Proposition 6.3.9 of Brockwell and Davis (1991), the statement follows. □

Finally, we obtain the asymptotic behavior of the score function.

Proposition 3.21.

Let Assumptions (*A*2)–(*A*4) *and* (*B*2)–(*B*3) *hold. Let* $\Sigma_{\nabla W}$ *be defined as in* (3.1). *Then,* *as* $n \to \infty$,

$$\sqrt{n}\,[\nabla_\vartheta W_n(\vartheta_0)]^\top \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma_{\nabla W}).$$

Proof The proof is based on the Cramér-Wold Theorem and Proposition 3.18. Therefore, let $\lambda = (\lambda_1, \dots, \lambda_r)^\top \in \mathbb{R}^r$. We obtain

$$\begin{aligned}
&\sqrt{n}\,[\nabla_\vartheta W_n(\vartheta_0)]\,\lambda \\
&= \frac{1}{2\sqrt{n}} \sum_{j=-n+1}^{n} \nabla_\vartheta \left[\operatorname{tr}\left(f_Y^{(\Delta)}(\omega_j, \vartheta_0)^{-1} I_{n,Y}(\omega_j)\right) + \log(\det(f_Y^{(\Delta)}(\omega_j, \vartheta_0)))\right] \lambda \\
&= \frac{1}{2\sqrt{n}} \sum_{j=-n+1}^{n} \left[\sum_{t=1}^{r} \operatorname{tr}\left(-\lambda_t f_Y^{(\Delta)}(\omega_j)^{-1} \left(\frac{\partial}{\partial \vartheta_t} f_Y^{(\Delta)}(\omega_j, \vartheta_0)\right) f_Y^{(\Delta)}(\omega_j)^{-1} I_{n,Y}(\omega_j)\right)\right] \\
&\quad + \frac{1}{2\sqrt{n}} \sum_{j=-n+1}^{n} \nabla_\vartheta [\operatorname{tr}(\log(f_Y^{(\Delta)}(\omega_j, \vartheta_0)))]\lambda.
\end{aligned}$$

We define the matrix function $\eta_\lambda : [-\pi, \pi] \to \mathbb{C}^{m\times m}$ as

$$\eta_\lambda(\omega) = -\sum_{t=1}^{r} \lambda_t f_Y^{(\Delta)}(\omega)^{-1} \left(\frac{\partial}{\partial \vartheta_t} f_Y^{(\Delta)}(\omega, \vartheta_0)\right) f_Y^{(\Delta)}(\omega)^{-1}, \quad \omega \in [-\pi, \pi]. \tag{3.34}$$

Furthermore,

$$\operatorname{tr}\left(\frac{\partial}{\partial \vartheta_t} \log\left(f_Y^{(\Delta)}(\omega, \vartheta_0)\right)\right) = \operatorname{tr}\left(f_Y^{(\Delta)}(\omega)^{-1} \left(\frac{\partial}{\partial \vartheta_t} f_Y^{(\Delta)}(\omega, \vartheta_0)\right)\right).$$

Then,

$$\sqrt{n}\,[\nabla_\vartheta W_n(\vartheta_0)]\,\lambda = \frac{1}{2\sqrt{n}} \sum_{j=-n+1}^{n} \operatorname{tr}\left(\eta_\lambda(\omega_j)\left(I_{n,Y}(\omega_j) - f_Y^{(\Delta)}(\omega_j)\right)\right).$$

Apparently, η_λ is two times continuously differentiable by Remark 3.7 and 2π-periodic. Moreover, every component of the Fourier coefficients $(\widehat{(\eta_\lambda)}_u)_{u\in\mathbb{Z}}$ of η_λ satisfies $\sum_{u=-\infty}^{\infty} |\widehat{(\eta_\lambda)}_u[k,\ell]||u|^{1/2} < \infty, k, \ell \in \{1, \dots, m\}$ (see Brockwell and Davis (1991), Exercise 2.22 applied to η_λ and its derivative η'_λ), and therefore, $\sum_{u=-\infty}^{\infty} \|\widehat{(\eta_\lambda)}_u\||u|^{1/2} < \infty$ follows. Then, due to Proposition 3.18, we get as $n \to \infty$,

$$\sqrt{n}\,[\nabla_\vartheta W_n(\vartheta_0)]\,\lambda \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma_{\nabla_\vartheta W \lambda}),$$

where

$$\begin{aligned}
\Sigma_{\nabla_\vartheta W \lambda} &= \frac{1}{\pi} \int_{-\pi}^{\pi} \operatorname{tr}\left(\eta_\lambda(\omega) f_Y^{(\Delta)}(\omega) \eta_\lambda(\omega) f_Y^{(\Delta)}(\omega)\right) d\omega \\
&+ \frac{1}{16\pi^4} \int_{-\pi}^{\pi} \operatorname{vec}\left(\Phi(e^{-i\omega})^\top \eta_\lambda(\omega)^\top \Phi(e^{i\omega})\right)^\top d\omega \\
&\quad \left(\mathbb{E}\left[N_1^{(\Delta)} N_1^{(\Delta)\top} \otimes N_1^{(\Delta)} N_1^{(\Delta)\top}\right] - 3\Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)}\right) \int_{-\pi}^{\pi} \operatorname{vec}\left(\Phi(e^{i\omega})^\top \eta_\lambda(\omega) \Phi(e^{-i\omega})\right) d\omega \\
&=: \Sigma_{\lambda,1} + \Sigma_{\lambda,2} + \Sigma_{\lambda,3}.
\end{aligned}$$

We investigate the three terms separately. With (3.19), the first term fulfills the representation

$$
\begin{aligned}
&\Sigma_{\lambda,1}\\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \operatorname{tr}\left(\eta_\lambda(\omega) f_Y^{(\Delta)}(\omega)\eta_\lambda(\omega) f_Y^{(\Delta)}(\omega)\right) d\omega\\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \operatorname{tr}\left(\left(\sum_{t=1}^{r}\lambda_t \frac{\partial}{\partial\vartheta_t} f_Y^{(\Delta)}(\omega,\vartheta_0)\right) f_Y^{(\Delta)}(\omega)^{-1}\left(\sum_{s=1}^{r}\lambda_s \frac{\partial}{\partial\vartheta_s} f_Y^{(\Delta)}(\omega,\vartheta_0)\right) f_Y^{(\Delta)}(\omega)^{-1}\right) d\omega\\
&= \lambda^\top\left[\frac{1}{\pi}\int_{-\pi}^{\pi}\left(\nabla_\vartheta f_Y^{(\Delta)}(-\omega,\vartheta_0)\right)^\top\left(f_Y^{(\Delta)}(-\omega)^{-1}\otimes f_Y^{(\Delta)}(\omega)^{-1}\right)\nabla_\vartheta f_Y^{(\Delta)}(\omega,\vartheta_0)\right]\lambda.
\end{aligned}
$$

Similarly, we get the representation

$$
\begin{aligned}
&\Sigma_{\lambda,2}\\
&= \frac{\lambda^\top}{16\pi^4}\left[\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\mathbb{E}\left[\nabla_\vartheta f_Y^{(\Delta)}(-\omega,\vartheta_0)^\top\left(f_Y^{(\Delta)}(-\omega)^{-1}\Phi(e^{i\omega})\otimes f_Y^{(\Delta)}(\omega)^{-1}\Phi(e^{-i\omega})\right)\right.\right.\\
&\qquad \left(N_1^{(\Delta)}N_1^{(\Delta)\top}\right)\otimes\left(N_1^{(\Delta)}N_1^{(\Delta)\top}\right)\\
&\qquad \left.\left.\left(\Phi(e^{-i\tau})^\top f_Y^{(\Delta)}(-\tau)^{-1}\otimes\Phi(e^{i\tau})^\top f_Y^{(\Delta)}(\tau)^{-1}\right)\nabla_\vartheta f_Y^{(\Delta)}(\tau,\vartheta_0)\right] d\omega d\tau\right]\lambda
\end{aligned}
$$

for the second term, and analogously

$$
\begin{aligned}
&\Sigma_{\lambda,3}\\
&= -\frac{3\lambda^\top}{16\pi^4}\left[\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\nabla_\vartheta f_Y^{(\Delta)}(-\omega,\vartheta_0)^\top\left(f_Y^{(\Delta)}(-\omega)^{-1}\Phi(e^{i\omega})\Sigma_N^{(\Delta)}\Phi(e^{-i\tau})^\top f_Y^{(\Delta)}(-\tau)^{-1}\right)\right.\\
&\qquad \left.\otimes\left(f_Y^{(\Delta)}(\omega)^{-1}\Phi(e^{-i\omega})\Sigma_N^{(\Delta)}\Phi(e^{i\tau})^\top f_Y^{(\Delta)}(\tau)^{-1}\right)\nabla_\vartheta f_Y^{(\Delta)}(\tau,\vartheta_0) d\omega d\tau\right]\lambda
\end{aligned}
$$

for the third term. □

Proof of Theorem 3.9. Since $\widehat{\vartheta}_n^{(\Delta)} \xrightarrow{a.s.} \vartheta_0$ (see Theorem 3.8) and $\Sigma_{\nabla^2 W}$ is positive definite (see Lemma 3.17) the conclusion follows from (3.15), Proposition 3.16 and Proposition 3.21.

Sketch of the proof of Remark 3.10

Let Φ_Z be the polynomial of the (existing) VAR(∞) of the VARMA(p,q) process. Proposition 3.18 can be formulated for VARMA processes. As in the proof of Theorem 3.9 we have to plug in there for η the function η_λ as given in (3.34). Then, $\widehat{q}_0$ in (3.20) has for the VARMA process $(Z_n)_{n\in\mathbb{N}}$ the form

$$
\begin{aligned}
\widehat{q}_0 &= \int_{-\pi}^{\pi} -2\pi\sum_{t=1}^{r}\lambda_t\Sigma_e^{-1}\Phi_Z(e^{-i\omega})^{-1}\left(\frac{\partial}{\partial\vartheta_t} f_Z(\omega,\vartheta_0)\right)\Phi_Z(e^{i\omega})^{\top-1}\Sigma_e^{-1} d\omega\\
&= -\Sigma_e^{-1}\int_{-\pi}^{\pi}\sum_{t=1}^{r}\lambda_t\frac{\partial}{\partial\vartheta_t}\log\left(\Phi_Z(e^{-i\omega},\vartheta_0)\right) d\omega
\end{aligned}
$$

$$-\int_{-\pi}^{\pi}\left(\sum_{t=1}^{r}\lambda_t\frac{\partial}{\partial\vartheta_t}\log\Big(\Phi_Z(e^{i\omega},\vartheta_0)\Big)\right)^{\top}d\omega\;\Sigma_e^{-1}.$$

If Φ_Z is two times differentiable, the Leibniz rule yields

$$\begin{aligned}\widehat{q}_0 = &-\Sigma_e^{-1}\sum_{t=1}^{r}\lambda_t\frac{\partial}{\partial\vartheta_t}\int_{-\pi}^{\pi}\log\Big(\Phi_Z(e^{-i\omega},\vartheta_0)\Big)\,d\omega\\ &-\left[\sum_{t=1}^{r}\lambda_t\frac{\partial}{\partial\vartheta_t}\int_{-\pi}^{\pi}\log\Big(\Phi_Z(e^{i\omega},\vartheta_0)\Big)\,d\omega\right]^{\top}\Sigma_e^{-1}.\end{aligned}$$

Similarly to the proof of Theorem 5.8.1 of Brockwell and Davis (1991), one can show that the integrals are constant and therefore, that $\widehat{q}_0 = 0$. For a more detailed approach, we refer to Dunsmuir and Hannan (1976). □

CHAPTER 4

The adjusted Whittle estimator

In this chapter, we solely consider state space models where $Y(\vartheta)$ and $L(\vartheta)$ are one-dimensional for every $\vartheta \in \Theta$, i.e., $A(\vartheta) \in \mathbb{R}^{N\times N}$, $B(\vartheta) \in \mathbb{R}^{N\times 1}$ and $C(\vartheta) \in \mathbb{R}^{1\times N}$. This includes, in particular, univariate CARMA processes, see Section 2.1. When considering light-tailed processes, we assume that the variance parameter $\sigma_L^2 := \Sigma_L > 0$ of the driving Lévy process does not depend on ϑ and has not to be estimated. In this context, we consider an adjusted Whittle estimator which takes into account that we do not have to estimate the variance. In some settings, such an adjustment is useful for the estimation of processes with non-existing variance. For example, Mikosch et al. (1995) estimated the parameters of ARMA models in discrete time where the white noise has a symmetric stable distribution and obtained desirable properties. Actually, our hope is that the adaption yields good results for the estimation of α-stable CARMA models. We therefore start with the construction of an estimator which is independent of the variance of the driving process. Subsequently, we derive its asymptotic behavior in two different settings. First, the considered processes are assumed to have finite second moments. We see that the estimator then behaves as desired. Unfortunately, we find out that those properties can not be carried out to a general setting in which the considered processes are symmetric α-stable CARMA processes.

4.1. Motivation

We now adapt the Whittle function in a way which makes it independent of the variance of the driving Lévy process. Therefore, we use the representation of the spectral density in (2.15) which depends on $\Pi^{-1}(\cdot, \vartheta)$ and $V^{(\Delta)}(\vartheta)$. Although the variance σ_L^2 goes linearly in $\Omega^{(\Delta)}(\vartheta)$ and $V^{(\Delta)}(\vartheta)$, both $K^{(\Delta)}(\vartheta)$ and $\Pi(\cdot,\vartheta)$ do not depend on σ_L^2 anymore. The second summand of the Whittle function W_n is removed and the first term is adjusted so that we obtain the *adjusted Whittle function*

$$W_n^{(A)}(\vartheta) = \frac{\pi}{n} \sum_{j=-n+1}^{n} |\Pi(e^{i\omega_j},\vartheta)|^2 I_{n,Y}(\omega_j) = \frac{V^{(\Delta)}(\vartheta)}{2n} \sum_{j=-n+1}^{n} f_Y^{(\Delta)}(\omega_j,\vartheta)^{-1} I_{n,Y}(\omega_j). \tag{4.1}$$

The corresponding minimizer

$$\widehat{\vartheta}_n^{(\Delta,A)} = \arg\min_{\vartheta\in\Theta} W_n^{(A)}(\vartheta)$$

is the *adjusted Whittle estimator*.

4.2. Adjusted Whittle estimation for CARMA processes with finite second moments

4.2.1. Setting

Since the estimation procedure is different to that of the previous chapter, we also have to adapt Assumption A.

Assumption $\widetilde{\mathbf{A}}$.
Let Assumptions $(A1)$–$(A5)$ and $(A7)$ hold. Furthermore, assume

($\widetilde{A}6$) For any $\vartheta_1,\ \vartheta_2 \in \Theta$, $\vartheta_1 \neq \vartheta_2$, there exists some $z \in \mathbb{C}$ with $|z| = 1$ and $\Pi(z,\vartheta_1) \neq \Pi(z,\vartheta_2)$.

It is needless to say that conditions as those for the function $\vartheta \to \sigma_L^2$ are fulfilled naturally. In addition to Remark 3.6, which remains mostly applicable, we stress that, under Assumption $\widetilde{A}$, Π^{-1} as defined in (2.11) exists for all $\vartheta \in \Theta$ and that the mapping $(\omega,\vartheta) \to \Pi^{-1}(e^{i\omega},\vartheta)$ is continuous.

For the asymptotic normality of the adjusted Whittle estimator we have to adjust Assumption B.

Assumption $\widetilde{\mathbf{B}}$.
Let Assumptions (B1)-(B3) hold. Furthermore, assume

($\widetilde{B}4$) For any $c \in \mathbb{C}^r$ there exists an $\omega^* \in [-\pi,\pi]$ such that $\nabla_\vartheta |\Pi(e^{i\omega^*},\vartheta_0)|^{-2} c \neq 0$.

Remark 4.1.
Under Assumption $\widetilde{A}$ and Assumption $\widetilde{B}$ the mapping $\vartheta \to \Pi(e^{i\omega},\vartheta)$ is three times continuously differentiable. Similarly to Lemma 3.17, ($\widetilde{B}4$) guarantees the invertibility of

$$\Sigma_{\nabla^2 W^{(A)}} := \frac{V^{(\Delta)}}{2\pi} \int_{-\pi}^{\pi} \nabla_\vartheta \log\left(|\Pi(e^{i\omega},\vartheta_0)|^{-2}\right)^\top \nabla_\vartheta \log\left(|\Pi(e^{i\omega},\vartheta_0)|^{-2}\right) d\omega. \tag{4.2}$$

4.2.2. Consistency and asymptotic normality

Theorem 4.2.
Let Assumption $\widetilde{A}$ hold. Then, as $n \to \infty$,

$$\widehat{\vartheta}_n^{(\Delta,A)} \xrightarrow{a.s.} \vartheta_0.$$

The proof follows the same steps as the proof of the consistency of the Whittle estimator in Theorem 3.8.

Theorem 4.3.
Let Assumption $\widetilde{A}$ and $\widetilde{B}$ hold. Further, let $\Sigma_{\nabla^2 W^{(A)}}$ be defined as in (4.2) and

$$\begin{aligned}
\Sigma_{\nabla W^{(A)}} &= \frac{V^{(\Delta)2}}{\pi}\int_{-\pi}^{\pi} \nabla_\vartheta \log\left(|\Pi(e^{i\omega},\vartheta_0)|^{-2}\right)^\top \nabla_\vartheta \log\left(|\Pi(e^{i\omega},\vartheta_0)|^{-2}\right) d\omega \\
&\quad + \frac{1}{4\pi^2}\left[\int_{-\pi}^{\pi} \nabla_\vartheta |\Pi(e^{i\omega},\vartheta_0)|^{2\top}\left[\Phi(e^{i\omega})\otimes\Phi(e^{-i\omega})\right] d\omega\right] \\
&\quad \cdot\left[\mathbb{E}\left[N_1^{(\Delta)}N_1^{(\Delta)\top}\otimes N_1^{(\Delta)}N_1^{(\Delta)\top}\right] - 3\Sigma_N^{(\Delta)}\otimes\Sigma_N^{(\Delta)}\right] \\
&\quad \cdot\left[\int_{-\pi}^{\pi} \nabla_\vartheta |\Pi(e^{i\omega},\vartheta_0)|^{2\top}\left[\Phi(e^{-i\omega})\otimes\Phi(e^{i\omega})\right] d\omega\right]^\top .
\end{aligned}$$

Then, as $n \to \infty$,

$$\sqrt{n}\left(\widehat{\vartheta}_n^{(\Delta,A)} - \vartheta_0\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\Sigma_{W^{(A)}}),$$

where $\Sigma_{W^{(A)}}$ has the representation $\Sigma_{W^{(A)}} = [\Sigma_{\nabla^2 W^{(A)}}]^{-1}\Sigma_{\nabla W^{(A)}}[\Sigma_{\nabla^2 W^{(A)}}]^{-1}$.

Remark 4.4.
For the one-dimensional Ornstein-Uhlenbeck process, for which $m = d = N = 1$ and $C(\vartheta) = B(\vartheta) = 1$ holds, the limit covariance matrix $\Sigma_{W^{(A)}}$ of Theorem 4.3 reduces due to Remark 4.8 and Theorem 3''', Chapter 3, of Hannan (2009) to

$$\begin{aligned}
\Sigma_{W^{(A)}} &= 4\pi\left[\int_{-\pi}^{\pi} \nabla_\vartheta \log\left(|\Pi(e^{i\omega},\vartheta_0)|^{-2}\right)^\top \nabla_\vartheta \log\left(|\Pi(e^{i\omega},\vartheta_0)|^{-2}\right) d\omega\right]^{-1} \\
&= 4\pi\left[\int_{-\pi}^{\pi} \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega,\vartheta_0))^\top \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega,\vartheta_0)) d\omega\right. \\
&\quad \left. -\frac{1}{2\pi}\left(\int_{-\pi}^{\pi} \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega,\vartheta_0)) d\omega\right)^\top \left(\int_{-\pi}^{\pi} \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega,\vartheta_0)) d\omega\right)\right]^{-1}.
\end{aligned}$$

Due to Remark 3.11 (b)

$$\Sigma_W = 2\cdot[\Sigma_{\nabla^2 W}]^{-1} = 4\pi\left[\int_{-\pi}^{\pi} \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega,\vartheta_0))^\top \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega,\vartheta_0)) d\omega\right]^{-1}$$

and hence, $\Sigma_{W^{(A)}} \geq \Sigma_W$. Thus, the adjusted Whittle estimator has a higher variance than the Whittle estimator. Let $\vartheta_0 < 0$ be the zero of the AR polynomial in the CAR(1) model, i.e., $A(\vartheta_0) = \vartheta_0$. Simple calculations show that $\Sigma_{W^{(A)}} = e^{-2\vartheta_0} - 1$ which is equal to the asymptotic variance of the maximum likelihood estimator of Brockwell and Lindner

(2019). However, it is not possible to make this conclusion for general CARMA processes. There exist CARMA processes for which the adjusted Whittle estimator has a different asymptotic variance than the maximum likelihood estimator of Brockwell and Lindner (2019).

4.2.3. PROOFS

PROOF OF THEOREM 4.2

The proof of Theorem 4.2 is similar to the proof of Theorem 3.8. Therefore, we simply adapt the parts which are not the same, namely Proposition 3.13 and Proposition 3.14. We start by stating that $W_n^{(A)}$ converges almost surely uniformly to

$$W^{(A)}(\vartheta) := \int_{-\pi}^{\pi} |\Pi(e^{i\omega}, \vartheta)|^2 f_Y^{(\Delta)}(\omega) d\omega$$

which can be shown in the same way as the uniform convergence of $W_n^{(1)}$ in Proposition 3.13.

Proposition 4.5.
Let Assumptions (A1)–(A4) hold. Then, as $n \to \infty$,

$$\sup_{\vartheta \in \Theta} |W_n^{(A)}(\vartheta) - W^{(A)}(\vartheta)| \overset{n\to\infty}{\longrightarrow} 0 \quad \mathbb{P}\text{-}a.s.$$

Proposition 4.6.
Let Assumptions (A1)–(A4) and $(\widetilde{A6})$ hold. Then, $W^{(A)}$ has a unique minimum in ϑ_0.

Proof. Let $\vartheta \neq \vartheta_0$. Due to the definition of the linear innovations and Assumption $(\widetilde{\text{A6}})$, we have

$$\begin{aligned} V^{(\Delta)} &= \mathbb{E}\left[\varepsilon_k^{(\Delta)\,2}\right] = \mathbb{E}\left[\left(\Pi(\mathfrak{B})Y_k^{(\Delta)}\right)^2\right] \\ &< \mathbb{E}\left[\left(\Pi(\mathfrak{B}, \vartheta)Y_k^{(\Delta)}\right)^2\right] = \int_{-\pi}^{\pi} |\Pi(e^{i\omega}, \vartheta)|^2 f_Y^{(\Delta)}(\omega) d\omega = W^{(A)}(\vartheta), \end{aligned}$$

where for the second last equality we used Theorem 11.8.3 of Brockwell and Davis (1991) as well. Furthermore, $V^{(\Delta)} = \mathbb{E}[(\Pi(\mathfrak{B})Y_k^{(\Delta)})^2] = W^{(A)}(\vartheta_0)$ holds. ☐

PROOF OF THEOREM 4.3

The proof of the asymptotic normality of the adjusted Whittle estimator is also similar to the proof of the asymptotic normality of the original Whittle estimator. We start to prove an adapted version of Proposition 3.18.

Proposition 4.7.
Let Assumptions (A2)–(A4) and (B2) hold. Suppose $\eta : [-\pi, \pi] \to \mathbb{C}$ is a symmetric

function with Fourier coefficients $(\widehat{\eta}_u)_{u\in\mathbb{Z}}$ *satisfying* $\sum_{u=-\infty}^{\infty}|\widehat{\eta}_u||u|^{1/2}<\infty$ *and*

$$\int_{-\pi}^{\pi}\left|\Pi^{-1}(e^{i\omega})\right|^2\eta(\omega)d\omega=0.$$

Then, as $n\to\infty$*,*

$$\frac{\pi}{\sqrt{n}}\sum_{j=-n+1}^{n}\eta(\omega_j)I_{n,Y}(\omega_j)\xrightarrow{\mathcal{D}}\mathcal{N}(0,\Sigma_\eta),$$

where

$$\Sigma_\eta=4\pi\int_{-\pi}^{\pi}\eta(\omega)^2f_Y^{(\Delta)}(\omega)^2d\omega+\frac{1}{4\pi^2}\int_{-\pi}^{\pi}\eta(\omega)\operatorname{vec}\left(\Phi(e^{-i\omega})^\top\Phi(e^{i\omega})\right)^\top d\omega$$
$$\cdot\left(\mathbb{E}\left[N_1^{(\Delta)}N_1^{(\Delta)\top}\otimes N_1^{(\Delta)}N_1^{(\Delta)\top}\right]-3\Sigma_N^{(\Delta)}\otimes\Sigma_N^{(\Delta)}\right)\int_{-\pi}^{\pi}\eta(\omega)\operatorname{vec}\left(\Phi(e^{i\omega})^\top\Phi(e^{-i\omega})\right)d\omega.$$

Proof. Note that

$$\sqrt{n}\int_{-\pi}^{\pi}f_Y^{(\Delta)}(\omega)\eta(\omega)d\omega=\frac{\sqrt{n}V^{(\Delta)}}{2\pi}\int_{-\pi}^{\pi}\left|\Pi^{-1}(e^{i\omega})\right|^2\eta(\omega)d\omega=0.$$

Therefore, an application of Lemma A.7 gives

$$\frac{\pi}{\sqrt{n}}\sum_{j=-n+1}^{n}\eta(\omega_j)I_{n,Y}(\omega_j)=\frac{\pi}{\sqrt{n}}\sum_{j=-n+1}^{n}\eta(\omega_j)\left(I_{n,Y}(\omega_j)-f_Y^{(\Delta)}(\omega_j)\right)+o(1)$$

and Proposition 3.18 leads to the statement. □

Remark 4.8.
For an Ornstein-Uhlenbeck process, Σ_η reduces to

$$\Sigma_\eta=4\pi\int_{-\pi}^{\pi}\eta(\omega)^2f_Y^{(\Delta)}(\omega)^2d\omega,$$

since $\Pi^{-1}(e^{i\omega},\vartheta)=\Phi(e^{i\omega},\vartheta)$ for all $(\omega,\vartheta)\in[-\pi,\pi]\times\Theta$ implies

$$\int_{-\pi}^{\pi}\eta(\omega)\operatorname{vec}\left(\Phi(e^{-i\omega})^\top\Phi(e^{i\omega})\right)^\top d\omega=\int_{-\pi}^{\pi}\left|\Pi^{-1}(e^{i\omega})\right|^2\eta(\omega)d\omega=0.$$

Proposition 4.9.
Let Assumptions $(A2)$*–*$(A4)$*,* $(\widetilde{A6})$ *and* $(B2)$*–*$(B3)$ *hold. Then, as* $n\to\infty$*,*

$$\sqrt{n}\left[\nabla_\vartheta W_n^{(A)}(\vartheta_0)\right]^\top\xrightarrow{\mathcal{D}}\mathcal{N}(0,\Sigma_{\nabla W^{(A)}}).$$

Proof. Similar to the proof of Proposition 3.21, we make use of the Cramér-Wold Theorem. For $\lambda=(\lambda_1,\ldots,\lambda_r)^\top\in\mathbb{R}^r$, we get

$$\sqrt{n}\left[\nabla_\vartheta W_n^{(A)}(\vartheta_0)\right]\lambda=\frac{\pi}{\sqrt{n}}\sum_{j=-n+1}^{n}\sum_{t=1}^{r}\lambda_t\frac{\partial}{\partial\vartheta_t}\left|\Pi(e^{i\omega_j},\vartheta_0)\right|^2I_{n,Y}(\omega_j)$$

$$= \frac{\pi}{\sqrt{n}} \sum_{j=-n+1}^{n} \sum_{t=1}^{r} \lambda_t \frac{\partial}{\partial \vartheta_t} \left(f_Y^{(\Delta)}(\omega, \vartheta_0)^{-1} \frac{V^{(\Delta)}(\vartheta_0)}{2\pi} \right) I_{n,Y}(\omega_j).$$

We define η_λ by

$$\eta_\lambda(\omega) = \sum_{t=1}^{r} \lambda_t \frac{\partial}{\partial \vartheta_t} \left(f_Y^{(\Delta)}(\omega, \vartheta_0)^{-1} \frac{V^{(\Delta)}(\vartheta_0)}{2\pi} \right), \quad \omega \in [-\pi, \pi],$$

and obtain

$$\begin{aligned} &\int_{-\pi}^{\pi} \eta_\lambda(\omega) \left| \Pi^{-1}(e^{i\omega}) \right|^2 d\omega \\ = &\int_{-\pi}^{\pi} \sum_{t=1}^{r} \lambda_t \left(\frac{\frac{\partial}{\partial \vartheta_t} V^{(\Delta)}(\vartheta_0)}{2\pi} f_Y^{(\Delta)}(\omega)^{-1} - \frac{\frac{\partial}{\partial \vartheta_t} f_Y^{(\Delta)}(\omega, \vartheta_0)}{f_Y^{(\Delta)}(\omega, \vartheta_0)^2} \frac{V^{(\Delta)}(\vartheta_0)}{2\pi} \right) \left| \Pi^{-1}(e^{i\omega}) \right|^2 d\omega \\ = &\left[2\pi \nabla_\vartheta \log(V^{(\Delta)}(\vartheta_0)) - \int_{-\pi}^{\pi} \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega, \vartheta_0)) d\omega \right] \lambda. \end{aligned}$$

Under Assumption (B3), the Leibniz rule and Theorem 3''', Chapter 3, of Hannan (2009) can be applied, which results in

$$\begin{aligned} &\left[2\pi \nabla_\vartheta \log(V^{(\Delta)}(\vartheta_0)) - \int_{-\pi}^{\pi} \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega, \vartheta_0)) d\omega \right] \lambda \\ = \ &\nabla_\vartheta \left[2\pi \log(V^{(\Delta)}(\vartheta_0)) - 2\pi \log(V^{(\Delta)}(\vartheta_0)) + 2\pi \log(2\pi) \right] \lambda = 0. \end{aligned}$$

As in Proposition 3.21, this transformation leads to the applicability of Proposition 4.7. Therefore, we get

$$\sqrt{n} \left[\nabla_\vartheta W_n^{(A)}(\vartheta_0) \right] \lambda \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma_{\nabla W^{(A)} \lambda})$$

with

$$\begin{aligned} &\Sigma_{\nabla W^{(A)} \lambda} \\ = \ &4\pi \int_{-\pi}^{\pi} \eta_\lambda(\omega)^2 f_Y^{(\Delta)}(\omega)^2 d\omega + \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \eta_\lambda(\omega) \operatorname{vec}\left(\Phi(e^{-i\omega})^\top \Phi(e^{i\omega}) \right)^\top d\omega \\ &\left(\mathbb{E}\left[N_1^{(\Delta)} N_1^{(\Delta)\top} \otimes N_1^{(\Delta)} N_1^{(\Delta)\top} \right] - 3\Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)} \right) \int_{-\pi}^{\pi} \eta_\lambda(\omega) \operatorname{vec}\left(\Phi(e^{i\omega})^\top \Phi(e^{-i\omega}) \right) d\omega. \end{aligned}$$

The representation $\eta_\lambda(\omega) = \left[\nabla_\vartheta |\Pi(e^{i\omega}, \vartheta_0)|^2 \right] \lambda$ completes the proof. □

To prove Theorem 4.3, we need an analogue result to Proposition 3.16. Since the following proposition can be shown completely analogously, the proof will be restricted to the transformation of the limit matrix.

Proposition 4.10.
Let Assumptions (A1)–(A4), $(\widetilde{A6})$ *and (B3) hold. Furthermore, let* $(\vartheta_n^*)_{n\in\mathbb{N}}$ *be a sequence*

in Θ *with* $\vartheta_n^* \xrightarrow{a.s.} \vartheta_0$ *as* $n \to \infty$*. Then, as* $n \to \infty$*,*

$$\nabla_\vartheta^2 W_n^{(A)}(\vartheta_n^*) \xrightarrow{a.s.} \Sigma_{\nabla^2 W^{(A)}}.$$

Proof. Some straightforward calculation yields

$$\Sigma_{\nabla^2 W^{(A)}} = \int_{-\pi}^{\pi} \left[\nabla_\vartheta^2 |\Pi(e^{i\omega}, \vartheta_0)|^2\right] f_Y^{(\Delta)}(\omega) d\omega.$$

Theorem 3''' in Chapter 3 of Hannan (2009) states

$$V^{(\Delta)}(\vartheta) = 2\pi \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log(f_Y^{(\Delta)}(\omega, \vartheta_0)) d\omega\right).$$

Along with (2.15) and the Leibniz rule this give the representation

$$\begin{aligned}
&\Sigma_{\nabla^2 W^{(A)}} \\
&= \int_{-\pi}^{\pi} \nabla_\vartheta^2 \left[\frac{V^{(\Delta)}(\vartheta_0)}{2\pi} f_Y^{(\Delta)}(\omega, \vartheta_0)^{-1}\right] f_Y^{(\Delta)}(\omega) d\omega \\
&= \nabla_\vartheta^2 V^{(\Delta)}(\vartheta_0) - 2\nabla_\vartheta V^{(\Delta)}(\vartheta_0)^\top \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega, \vartheta_0)) d\omega\right) \\
&\quad - \frac{V^{(\Delta)}}{2\pi}\int_{-\pi}^{\pi} \nabla_\vartheta^2 \log(f_Y^{(\Delta)}(\omega, \vartheta_0)) d\omega \\
&\quad + \frac{V^{(\Delta)}}{2\pi}\int_{-\pi}^{\pi} \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega, \vartheta_0))^\top \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega, \vartheta_0)) d\omega \\
&= \frac{V^{(\Delta)}}{2\pi}\int_{-\pi}^{\pi} \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega, \vartheta_0))^\top \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega, \vartheta_0)) d\omega \\
&\quad - V^{(\Delta)} \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega, \vartheta_0)) d\omega\right)^\top \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \nabla_\vartheta \log(f_Y^{(\Delta)}(\omega, \vartheta_0)) d\omega\right) \\
&= \frac{V^{(\Delta)}}{2\pi}\int_{-\pi}^{\pi} \nabla_\vartheta \log(|\Pi(e^{i\omega}, \vartheta_0)|^{-2})^\top \nabla_\vartheta \log(|\Pi(e^{i\omega}, \vartheta_0)|^{-2}) d\omega.
\end{aligned}$$

□

The proof of Theorem 4.3 now matches the proof of Theorem 3.9, where Proposition 3.21 is replaced by Proposition 4.9 and Proposition 3.16 is replaced by Proposition 4.10.

4.3. Adjusted Whittle estimation for α-stable CARMA processes

The topic of this section is the adjusted Whittle estimation of the parameters of symmetric α-stable CARMA processes. There are several arguments for the conjecture that the estimator might converge:

- The adjusted Whittle estimator for light-tailed CARMA processes is strongly consistent and asymptotically normally distributed by Theorem 4.2 and Theorem 4.3.

- In several simulations for symmetric α-stable CARMA processes as, e.g., in the setup of García et al. (2011) (see Section 6.1), it seems that the adjusted Whittle estimator converges to the true parameter.
- In the context of ARMA processes the ideas of Whittle estimation for ARMA processes with finite second moments could be transferred to ARMA processes with infinite second moments (see Mikosch et al. (1995)). Since equidistant sampled CARMA processes with finite second moments have a weak ARMA representation, see Theorem 2.8, it is plausible that similar results hold for the CARMA setting.

However, primary the last argument also yields a reason for the hypothesis that the adjusted Whittle estimator might not be consistent. Namely, by Proposition 2.11, the sampled α-stable CARMA process does not exhibit a weak ARMA representation in general. In fact, we will actually see that the proposed estimation procedure is not suited for parameter estimation of these processes. As an exception, only the Ornstein-Uhlenbeck process can be estimated consistently.
Therefore, we assume:

Assumption L2.
The driving process $L^{(\alpha)}$ of the CARMA process is a symmetric α-stable Lévy process with $L_1^{(\alpha)} \sim S_\alpha(\sigma, 0, 0)$ for some $\sigma > 0,\ \alpha \in (0,2)$,

and that Y is a symmetric α-stable CARMA process with kernel function $g(t) = c^\top \mathrm{e}^{At} e_p \mathbb{1}_{[0,\infty)}(t)$ as given in (2.4). In analogy to (4.1) for light-tailed CARMA processes, for symmetric α-stable CARMA processes the appropriately normalized adjusted Whittle function is

$$W_n^{(\alpha)}(\vartheta) := \frac{\pi}{n^{2/\alpha}} \sum_{j=-n+1}^{n} |\Pi(e^{i\omega_j}, \vartheta)|^2 I_n(\omega_j), \quad \vartheta \in \Theta,$$

and accordingly the adjusted Whittle estimator is

$$\widehat{\vartheta}_n^{(\Delta,\alpha)} := \arg\min_{\vartheta \in \Theta} W_n^{(\alpha)}(\vartheta).$$

Theorem 4.11.
Assume Assumption $\widetilde{A}$ with the second moment condition $(A2)$ replaced by Assumption L2. Suppose further $(B1)$ and $(B3)$. Define

$$W^{(\alpha)}(\vartheta) := \frac{1}{2\pi} \int_0^\Delta \left[\int_{-\pi}^{\pi} |\Pi(e^{i\omega}, \vartheta)|^2 \left| \sum_{j=-\infty}^{\infty} g(\Delta j - s) e^{-ij\omega} \right|^2 d\omega \right] dL_s^{(\alpha/2)},$$

where $L^{(\alpha/2)} = (L_t^{(\alpha/2)})_{t \geq 0}$ is an $\alpha/2$-stable Lévy process with

$$L_1^{(\alpha/2)} \sim S_{\alpha/2}(\sigma^2 \left(C_\alpha / C_{\alpha/2}\right)^{2/\alpha}, 1, 0)$$

and the constants C_α and $C_{\alpha/2}$ are defined as in (2.6)*. Then, as $n \to \infty$,*

$$(W_n^{(\alpha)}(\vartheta))_{\vartheta\in\Theta} \xrightarrow{\mathcal{D}} (W^{(\alpha)}(\vartheta))_{\vartheta\in\Theta} \quad \text{in} \quad (\mathcal{C}(\Theta), \|\cdot\|_\infty),$$

where $\mathcal{C}(\Theta)$ is the space of continuous functions on Θ with the supremum norm $\|\cdot\|_\infty$.

Proof. We approximate $|\Pi(e^{i\omega_j},\vartheta)|^2$ by the Cesàro sum of its Fourier series of size M for M sufficiently large. Define

$$\widehat{(|\Pi(\vartheta)|^2)}_k := \widehat{(|\Pi(e^{i\cdot},\vartheta)|^2)}_k := \frac{1}{2\pi}\int_{-\pi}^{\pi} |\Pi(e^{i\omega},\vartheta)|^2 e^{-ik\omega} d\omega,$$

$$q_M(\omega,\vartheta) := \frac{1}{M}\sum_{j=0}^{M-1}\left(\sum_{|k|\le j}\widehat{(|\Pi(\vartheta)|^2)}_k e^{ik\omega},\right) = \sum_{|k|<M}\left(1-\frac{|k|}{M}\right)\widehat{(|\Pi(\vartheta)|^2)}_k e^{ik\omega}$$

and thereby

$$W_{M,n}^{(\alpha)}(\vartheta) := \frac{\pi}{n^{2/\alpha}}\sum_{j=-n+1}^{n} q_M(\omega_j,\vartheta) I_{n,Y}(\omega_j), \quad \vartheta\in\Theta.$$

Let $\varepsilon_1 > 0$. A conclusion of Proposition 3.21 is that there exists an $M_0(\varepsilon_1)\in\mathbb{N}$ such that for $M \ge M_0(\varepsilon_1)$

$$\sup_{\omega\in[-\pi,\pi]}\sup_{\vartheta\in\Theta}|q_M(\omega,\vartheta) - |\Pi(e^{i\omega},\vartheta)|^2| < \varepsilon_1. \tag{4.3}$$

Similar arguments as in the proof of Proposition 3.13 yield

$$\sup_{\vartheta\in\Theta}\left|W_n^{(\alpha)}(\vartheta) - W_{M,n}^{(\alpha)}(\vartheta)\right| \le \frac{\varepsilon_1}{n^{2/\alpha-1}}\overline{\gamma}_{n,Y}(0) \quad \text{for } M \ge M_0(\epsilon_1).$$

Due to Theorem 2.18

$$\frac{1}{n^{2/\alpha-1}}\overline{\gamma}_{n,Y}(0) \xrightarrow{\mathcal{D}} \int_0^\Delta \sum_{j=-\infty}^{\infty} g(\Delta j - s)^2 dL_s^{(\alpha/2)} \quad \text{as } n\to\infty.$$

Therefore, we have for any $\epsilon_2 > 0$

$$\lim_{M\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(\sup_{\vartheta\in\Theta}\left|W_n^{(\alpha)}(\vartheta) - W_{M,n}^{(\alpha)}(\vartheta)\right| > \varepsilon_2\right) = 0. \tag{4.4}$$

Furthermore, representation (2.13) gives

$$\begin{aligned} &W_{M,n}^{(\alpha)}(\vartheta) \\ &= \sum_{|k|<M}\left(\left(1-\frac{|k|}{M}\right)\widehat{(|\Pi(\vartheta)|^2)}_k\left(n^{-2/\alpha+1}\sum_{|h|<n}\overline{\gamma}_{n,Y}(h)\right)\left(\frac{1}{2n}\sum_{j=-n+1}^{n} e^{i(k-h)\omega_j}\right)\right) \\ &= \sum_{|k|<M}\left(1-\frac{|k|}{M}\right)\widehat{(|\Pi(\vartheta)|^2)}_k n^{-2/\alpha+1}\overline{\gamma}_{n,Y}(k). \end{aligned} \tag{4.5}$$

We define

$$W_M^{(\alpha)}(\vartheta) := \sum_{|k|<M} \left(1 - \frac{|k|}{M}\right) \widehat{(|\Pi(\vartheta)|^2)}_k \int_0^\Delta \sum_{j=-\infty}^{\infty} g(\Delta(j+k) - s) g(\Delta j - s) dL_s^{(\alpha/2)}. \quad (4.6)$$

Due to Assumption $(B3)$ and the definition of Π, there exists a constant $\mathfrak{C} > 0$ such that for any $\delta > 0$

$$\begin{aligned}
&\sup_{\substack{|\vartheta_1-\vartheta_2|<\delta \\ \vartheta_1,\vartheta_2\in\Theta, k\in\mathbb{Z}}} \left| \widehat{(|\Pi(e^{i\cdot}, \vartheta_1)|^2)}_k - \widehat{(|\Pi(e^{i\cdot}, \vartheta_2)|^2)}_k \right| \\
&= \sup_{\substack{|\vartheta_1-\vartheta_2|<\delta \\ \vartheta_1,\vartheta_2\in\Theta, k\in\mathbb{Z}}} \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(|\Pi(e^{i\omega}, \vartheta_1)|^2 - |\Pi(e^{i\omega}, \vartheta_2)|^2 \right) e^{ik\omega} d\omega \right| \\
&\leq \max_{\substack{|\vartheta_1-\vartheta_2|<\delta \\ \vartheta_1,\vartheta_2\in\Theta}} \max_{\omega\in[-\pi,\pi]} ||\Pi(e^{i\omega}, \vartheta_1)|^2 - |\Pi(e^{i\omega}, \vartheta_2)|^2| \leq \mathfrak{C}\delta.
\end{aligned}$$

This means that $(\widehat{(|\Pi(\vartheta)|^2)}_k)_{\vartheta\in\Theta}$ is uniformly continuous. By Theorem 2.18, we have the joint convergence of the random vector $(\overline{\gamma}_{n,Y}(-M+1), \ldots, \overline{\gamma}_{n,Y}(M-1))$ implying with the representations (4.5), (4.6) and the continuous mapping theorem that

$$(W_{M,n}^{(\alpha)}(\vartheta))_{\vartheta\in\Theta} \xrightarrow{\mathcal{D}} (W_M^{(\alpha)}(\vartheta))_{\vartheta\in\Theta} \quad \text{in} \quad (\mathcal{C}(\Theta), \|\cdot\|_\infty). \quad (4.7)$$

Furthermore,

$$\begin{aligned}
&W_M^{(\alpha)}(\vartheta) \\
&= \int_0^\Delta \sum_{|k|<M} \left(1 - \frac{|k|}{M}\right) \widehat{(|\Pi(\vartheta)|^2)}_k \sum_{j=-\infty}^{\infty} g(\Delta(j+k) - s) g(\Delta j - s) dL_s^{(\alpha/2)} \\
&= \int_0^\Delta \sum_{|k|<M} \left(1 - \frac{|k|}{M}\right) \widehat{(|\Pi(\vartheta)|^2)}_k \sum_{j,\ell=-\infty}^{\infty} g(\Delta j - s) g(\Delta \ell - s) \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j+k-\ell)\omega} d\omega \right] dL_s^{(\alpha/2)} \\
&= \frac{1}{2\pi} \int_0^\Delta \left[\int_{-\pi}^{\pi} q_M(\omega, \vartheta) \left| \sum_{j=-\infty}^{\infty} g(\Delta j - s) e^{-ij\omega} \right|^2 d\omega \right] dL_s^{(\alpha/2)}.
\end{aligned}$$

By this,

$$\begin{aligned}
&W_M^{(\alpha)}(\vartheta) - W^{(\alpha)}(\vartheta) \\
&= \frac{1}{2\pi} \int_0^\Delta \left[\int_{-\pi}^{\pi} \left[q_M(\omega, \vartheta) - |\Pi(e^{i\omega}, \vartheta)|^2 \right] \left| \sum_{j=-\infty}^{\infty} g(\Delta j - s) e^{-ij\omega} \right|^2 d\omega \right] dL_s^{(\alpha/2)}
\end{aligned}$$

holds. Since the process $L^{(\alpha/2)}$ is positive and increasing we obtain

$$\sup_{\vartheta\in\Theta} |W_M^{(\alpha)}(\vartheta) - W^{(\alpha)}(\vartheta)|$$

$$
\begin{aligned}
&\leq \frac{1}{2\pi}\int_0^\Delta \left[\int_{-\pi}^{\pi} \sup_{\vartheta\in\Theta}\left|q_M(\omega,\vartheta)-|\Pi(e^{i\omega},\vartheta)|^2\right| \left|\sum_{j=-\infty}^{\infty} g(\Delta j-s)e^{-ij\omega}\right|^2 d\omega\right] dL_s^{(\alpha/2)} \\
&=: \widetilde{W}_M^{(\alpha/2)}.
\end{aligned}
$$

Note that by Property 3.2.2 of Samorodnitsky and Taqqu (1994), $\widetilde{W}_M^{(\alpha/2)} \sim S_{\alpha/2}(\sigma_M,\beta_M,\mu_M)$ where $\beta_M = 1,\ \mu_M = 0$ and

$$
\sigma_M^{\alpha/2} = \frac{\sigma^\alpha C_\alpha}{C_{\alpha/2}}\int_0^\Delta \left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \sup_{\vartheta\in\Theta}\left|q_M(\omega,\vartheta)-|\Pi(e^{i\omega},\vartheta)|^2\right| \left|\sum_{j=-\infty}^{\infty} g(\Delta j-s)e^{-ij\omega}\right|^2 d\omega\right]^{\alpha/2} ds.
$$

Due to Assumption $(A3)$ there exists a constant $\mathfrak{C}>0$ such that $\left(\sum_{j=0}^{\infty}\left\|e^{A(\Delta j-s)}\right\|\right)^2 < \mathfrak{C}$ for any $s\in[0,\Delta]$. Thus,

$$
\begin{aligned}
&\int_0^\Delta \left[\int_{-\pi}^{\pi} \left|\sum_{j=-\infty}^{\infty} g(\Delta j-s)e^{-ij\omega}\right|^2 d\omega\right]^{\alpha/2} ds \\
&\leq \int_0^\Delta \left[\int_{-\pi}^{\pi} \|c\|^2\|e_p\|^2 \left(\sum_{j=0}^{\infty}\left\|e^{A(\Delta j-s)}\right\|\right)^2 d\omega\right]^{\alpha/2} ds \\
&< \infty.
\end{aligned}
$$

A conclusion of this and (4.3) is that $\sigma_M^{\alpha/2} \overset{M\to\infty}{\longrightarrow} 0$ and hence, the characteristic function $\varphi_{\widetilde{W}_M^{\alpha/2}}$ of $\widetilde{W}_M^{(\alpha/2)}$ converges pointwise to $\varphi_{\widetilde{W}^{\alpha/2}} \equiv 1$. An application of Lévys continuity theorem results then in $\widetilde{W}_M^{(\alpha/2)} \overset{\mathbb{P}}{\longrightarrow} 0$ as $M\to\infty$. Finally,

$$
\sup_{\vartheta\in\Theta}|W_M^{(\alpha)}(\vartheta)-W^{(\alpha)}(\vartheta)| \overset{\mathbb{P}}{\longrightarrow} 0 \quad \text{as } M\to\infty \tag{4.8}
$$

as well. In view of (4.4)-(4.8), Theorem 3.2 of Billingsley (1968) completes the proof. □

Corollary 4.12.
Let the assumptions of Theorem 4.11 hold. Define $G_{\vartheta,\vartheta_0}:[0,\Delta]\to\mathbb{R}$ *as*

$$
G_{\vartheta,\vartheta_0}(u) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[|\Pi(e^{i\omega},\vartheta)|^2-|\Pi(e^{i\omega},\vartheta_0)|^2\right]\left|\sum_{j=-\infty}^{\infty} g(\Delta j-u)e^{-ij\omega}\right|^2 d\omega.
$$

Then,

$$
W^{(\alpha)}(\vartheta)-W^{(\alpha)}(\vartheta_0) \sim S_{\alpha/2}(\sigma_{\vartheta,\vartheta_0},\beta_{\vartheta,\vartheta_0},0)
$$

is an $\alpha/2$-stable random variable with parameters

$$\begin{aligned}\beta_{\vartheta,\vartheta_0} &= \frac{\int_0^\Delta (G^+_{\vartheta,\vartheta_0}(s))^{\alpha/2} - (G^-_{\vartheta,\vartheta_0}(s))^{\alpha/2} ds}{\int_0^\Delta |G_{\vartheta,\vartheta_0}(s)|^{\alpha/2} ds},\\ \sigma^{\alpha/2}_{\vartheta,\vartheta_0} &= \frac{\sigma^\alpha C_\alpha}{C_{\alpha/2}} \int_0^\Delta |G_{\vartheta,\vartheta_0}(s)|^{\alpha/2} ds.\end{aligned}$$

4.3.1. Whittle estimation for symmetric α-stable Ornstein-Uhlenbeck processes

An Ornstein-Uhlenbeck process $Y_t(\vartheta) = \int_{-\infty}^t e^{\vartheta(t-s)} dL_s^{(\alpha)}$, $t \geq 0$, sampled equidistantly has the AR(1) representation

$$Y_k^{(\Delta)}(\vartheta) = \mathrm{e}^{\vartheta\Delta} Y_{k-1}^{(\Delta)}(\vartheta) + \xi_k^{(\Delta)}(\vartheta), \qquad k \in \mathbb{Z},$$

where $\xi_k^{(\Delta)}(\vartheta) = \int_{(k-1)\Delta}^{k\Delta} \mathrm{e}^{\vartheta(k\Delta - s)}\, dL_s^{(\alpha)}$, $k \in \mathbb{N}$, is an i.i.d. symmetric α-stable sequence. Since the distribution of the white noise $\xi_k^{(\Delta)}(\vartheta)$ depends on ϑ, the theory of Mikosch et al. (1995) can not be applied directly to estimate ϑ in this setting even though we have an AR(1) representation. Thus, in this subsection we derive the consistency of the Whittle estimator for symmetric α-stable Ornstein-Uhlenbeck processes.

Proposition 4.13.
Let $\Theta \subseteq (-\infty, 0)$ be compact. Consider the family $(Y_t(\vartheta))_{\vartheta \in \Theta}$ with $Y_t(\vartheta) = \int_{-\infty}^t e^{\vartheta(t-s)} dL_s^{(\alpha)}$, $t \geq 0$, of symmetric α-stable Ornstein-Uhlenbeck processes. Then, as $n \to \infty$,

$$W_n^{(\alpha)}(\vartheta) \xrightarrow{\mathcal{D}} W_{OU}(\vartheta) S^*_{\alpha/2} \quad \text{in} \quad (\mathcal{C}(\Theta), \|\cdot\|_\infty),$$

where $S^*_{\alpha/2}$ is a positive $\alpha/2$-stable random variable and

$$W_{OU}(\vartheta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |1 - \mathrm{e}^{\vartheta\Delta + i\omega}|^2 |1 - \mathrm{e}^{\vartheta_0\Delta + i\omega}|^{-2}\, d\omega, \quad \vartheta \in \Theta.$$

Proof. The Ornstein-Uhlenbeck process $Y(\vartheta)$ has the kernel function $g_\vartheta(t) = e^{\vartheta t} \mathbf{1}_{[0,\infty)}(t)$ and the transfer function $\Pi(z, \vartheta) = 1 - e^{\vartheta\Delta} z$. Therefore, an application of Theorem 4.11 yields as $n \to \infty$,

$$\begin{aligned} W_n^{(\alpha)}(\vartheta) &\xrightarrow{\mathcal{D}} \frac{1}{2\pi} \int_0^\Delta \left[\int_{-\pi}^{\pi} |\Pi(e^{i\omega}, \vartheta)|^2 \left| \sum_{j=-\infty}^{\infty} g_{\vartheta_0}(\Delta j - s) e^{-ij\omega} \right|^2 d\omega \right] dL_s^{(\alpha/2)} \\ &= \frac{1}{2\pi} \int_0^\Delta \left[\int_{-\pi}^{\pi} |1 - e^{\vartheta\Delta + i\omega}|^2 \left| \sum_{j=1}^{\infty} \mathrm{e}^{\vartheta_0(\Delta j - s)} e^{-ij\omega} \right|^2 d\omega \right] dL_s^{(\alpha/2)} \\ &= \frac{1}{2\pi} \int_0^\Delta \left[\int_{-\pi}^{\pi} |1 - e^{\vartheta\Delta + i\omega}|^2 \left| e^{-\vartheta_0 s} \left(\sum_{j=0}^{\infty} \mathrm{e}^{\vartheta_0 \Delta j} e^{-ij\omega} - 1 \right) \right|^2 d\omega \right] dL_s^{(\alpha/2)} \end{aligned}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} |1-\mathrm{e}^{\vartheta\Delta+i\omega}|^2|1-\mathrm{e}^{\vartheta_0\Delta+i\omega}|^{-2}\,d\omega\int_0^{\Delta} e^{2\vartheta_0(\Delta-s)}dL_s^{(\alpha/2)}$$

in $(\mathcal{C}(\Theta),\|\cdot\|_\infty)$. Define $S^*_{\alpha/2} := \int_0^\Delta e^{2\vartheta_0(\Delta-s)}dL_s^{(\alpha/2)}$. Due to Property 3.2.2 of Samorodnitsky and Taqqu (1994)

$$S^*_{\alpha/2} \sim S_{\alpha/2}\left(\left(\frac{\sigma^\alpha C_\alpha}{C_{\alpha/2}}\int_0^\Delta e^{\alpha\vartheta_0 s}ds\right)^{2/\alpha},1,0\right)$$

which implies that $S^*_{\alpha/2}$ is positive (see Proposition 1.2.11 of Samorodnitsky and Taqqu (1994)). □

Proposition 4.14.
Let the assumptions of Proposition 4.13 hold. Then, W_{OU} has a unique minimum in ϑ_0.

Proof. Proposition 4.6 implies that under the Assumptions $(A1),(A3),(A4)$ and $(\widetilde{A}6)$

$$W_{OU}(\vartheta_0) = 1 < \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{|\Pi(e^{i\omega},\vartheta)|^2}{|\Pi(e^{i\omega},\vartheta_0)|^2}d\omega = W_{OU}(\vartheta) \quad \text{ for } \vartheta\neq\vartheta_0.$$

Hence, ϑ_0 is indeed the unique minimum. □

Theorem 4.15.
Let the assumptions of Proposition 4.13 hold. Then, as $n\to\infty$,

$$\widehat{\vartheta}_n^{(\Delta,\alpha)} \xrightarrow{\mathbb{P}} \vartheta_0.$$

Proof. Proposition 4.13 and Skorokhods representation theorem give that there exists a probability space with processes $(W_n^*(\vartheta))_{\vartheta\in\Theta}$ and $(W^*(\vartheta))_{\vartheta\in\Theta}$ having the same distributions as $(W_n^{(\alpha)}(\vartheta))_{\vartheta\in\Theta}$ and $(W^{(\alpha)}(\vartheta))_{\vartheta\in\Theta}$, respectively, with

$$\sup_{\vartheta\in\Theta}|W_n^*(\vartheta)-W^*(\vartheta)| \xrightarrow{a.s.} 0, \quad n\to\infty.$$

With the same arguments as in the proof of Theorem 3.8, we can show that the minimizing arguments $\widehat{\vartheta}_n^*$ and $\widehat{\vartheta}_0^*$ of $(W_n^*(\vartheta))_{\vartheta\in\Theta}$ and $(W^*(\vartheta))_{\vartheta\in\Theta}$, respectively, satisfy, as $n\to\infty$,

$$\vartheta_n^* \xrightarrow{a.s.} \vartheta_0,$$

which then implies $\widehat{\vartheta}_n^{(\Delta,\alpha)} \xrightarrow{\mathcal{D}} \vartheta_0$. Since ϑ_0 is a constant, convergence in distribution implies convergence in probability. □

4.3.2. Whittle estimation for general symmetric α-stable CARMA processes

Theorem 4.16.
Consider the setting of Theorem 4.11 for a symmetric α-stable CARMA(p,q) process with

$p \geq 2$. Then, in general, the limit function $W^{(\alpha)}$ of the Whittle function does not have a unique minimum in ϑ_0 and hence, the Whittle estimator is not consistent.

Proof. A necessary condition for the Whittle function $W^{(\alpha)}$ to have a unique minimum in ϑ_0 is that $W^{(\alpha)}(\vartheta) - W^{(\alpha)}(\vartheta_0)$ is a positive random variable for $\vartheta \neq \vartheta_0$ and, hence $\beta_{\vartheta,\vartheta_0}$ as defined in Corollary 4.12 is equal to 1.

However, this is not the case in general as can be seen in Example 4.17 and Example 4.18, which implies that the Whittle estimator is in general not consistent. □

Example 4.17.
We tackle the question, whether it is possible to find a model where $\beta_{\vartheta,\vartheta_0}$ is not equal to 1 for some $\vartheta \neq \vartheta_0$. Therefore, we consider symmetric 1.5-stable CARMA(2,0) processes with autoregressive and moving average polynomial

$$a_\vartheta(z) = z^2 - (\vartheta - 2)z - 2\vartheta \quad \text{and} \quad c_\vartheta(z) = \vartheta - 2,$$

respectively. These CARMA processes have the state space representation

$$\mathrm{d}X_t(\vartheta) = A(\vartheta)X_t(\vartheta)\,\mathrm{d}t + e_2\,\mathrm{d}L_t^{3/2} \quad \text{and} \quad Y_t(\vartheta) := c(\vartheta)^\top X_t(\vartheta), \quad t \geq 0,$$

where

$$A(\vartheta) = \begin{pmatrix} 0 & 1 \\ 2\vartheta & \vartheta - 2 \end{pmatrix} \quad \text{and} \quad c(\vartheta)^\top = (\vartheta - 2, 0).$$

Let the true parameter be $\vartheta_0 = -3$. The behavior of $\beta_{\vartheta,\vartheta_0}$ as defined in Corollary 4.12, the behavior of the non-normalized positive part

$$\beta^+_{\vartheta,\vartheta_0} := \int_0^\Delta (G^+_{\vartheta,\vartheta_0}(s))^{\alpha/2} ds$$

and the negative part

$$\beta^-_{\vartheta,\vartheta_0} := \int_0^\Delta (G^-_{\vartheta,\vartheta_0}(s))^{\alpha/2} ds,$$

respectively, are plotted as functions of ϑ for $\alpha = 1.5$ in Figure 4.1. As one can see, $\beta^-_{\vartheta,\vartheta_0} > 0$ for all $\vartheta \in (-\infty, -3) \cup (-3, -2)$, and hence, $\beta_{\vartheta,\vartheta_0} < 1$. Of course, this holds independent of the choice of α. Thus, $W^{(\alpha)}(\vartheta) - W^{(\alpha)}(\vartheta_0)$ is not a strictly positive random variable for $\vartheta \in (-\infty, -3) \cup (-3, -2)$ and hence, has not almost surely a unique minimum in ϑ_0. Especially, $\beta_{\vartheta,\vartheta_0} \to 0.8$ for $\vartheta \to -\infty$ in the case $\alpha = 1.5$.

Example 4.18.
In view of Example 3.3 of García et al. (2011), we consider CARMA(2,1) processes with the parametrization (2.3) and $a_1(\vartheta) = \vartheta_1$, $a_2(\vartheta) = \vartheta_2$ and $c_0(\vartheta) = \vartheta_3$, $c_1(\vartheta) = 1$, $\vartheta = (\vartheta_1, \vartheta_2, \vartheta_3) \in \Theta$. The kernel function g_ϑ in (2.4) has the representation

$$g_\vartheta(t) = \frac{(\vartheta_3 - \lambda^+(\vartheta))}{\sqrt{\vartheta_1^2 - 4\vartheta_2}} e^{-\lambda^+(\vartheta)t} - \frac{(\vartheta_3 - \lambda^-(\vartheta))}{\sqrt{\vartheta_1^2 - 4\vartheta_2}} e^{-\lambda^-(\vartheta)t}, \quad t \geq 0,$$

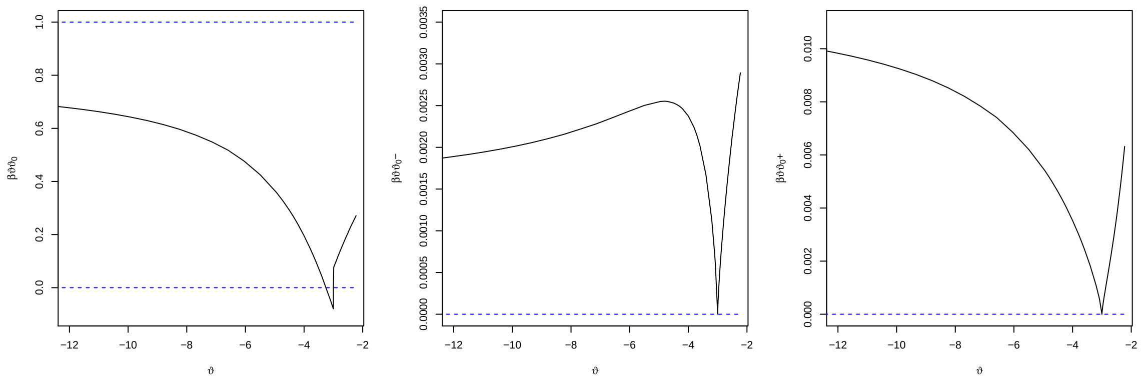

Figure 4.1.: Behavior of $\beta_{\vartheta,\vartheta_0}$, $\beta^-_{\vartheta,\vartheta_0}$ and $\beta^+_{\vartheta,\vartheta_0}$ in the CARMA$(2,0)$ model of Example 4.17. We set $\beta_{\vartheta_0,\vartheta_0} = 0$ to guarantee that $\beta_{\vartheta,\vartheta_0}$ is continuous.

with $\lambda^+(\vartheta) = \frac{\vartheta_1+\sqrt{\vartheta_1^2-4\vartheta_2}}{2}$ and $\lambda^-(\vartheta) = \frac{\vartheta_1-\sqrt{\vartheta_1^2-4\vartheta_2}}{2}$. As in García et al. (2011), the true parameter is chosen as

$$\vartheta_0 = (1.9647, 0.0893, 0.1761).$$

Therefore, the kernel function is

$$g_{\vartheta_0}(t) \approx 0.0692e^{-0.0465t} + 0.9307e^{-1.9181t}, \quad t \geq 0,$$

which is non-negative and we take $\alpha = 1.5$. In this setting, we calculate $\beta_{\vartheta,\vartheta_0}$ as a function of the components ϑ_1, ϑ_2 and ϑ_3, respectively, where we fix the other two variables. Then the functions $\beta_{\vartheta,\vartheta_0}$, $\beta^-_{\vartheta,\vartheta_0}$ and $\beta^+_{\vartheta,\vartheta_0}$ are plotted in Figure 4.2. In all three cases, the plots show that $\beta^-_{\vartheta,\vartheta_0} > 0$ for some $\vartheta \neq \vartheta_0$ implying $\beta_{\vartheta,\vartheta_0} < 1$. Therefore, if we only allow a single parameter to vary, the Whittle estimator converges to a function which has not an unique minimum in the true parameter. Hence, the Whittle estimator is not consistent. Again this statement is independent of the choice of α.

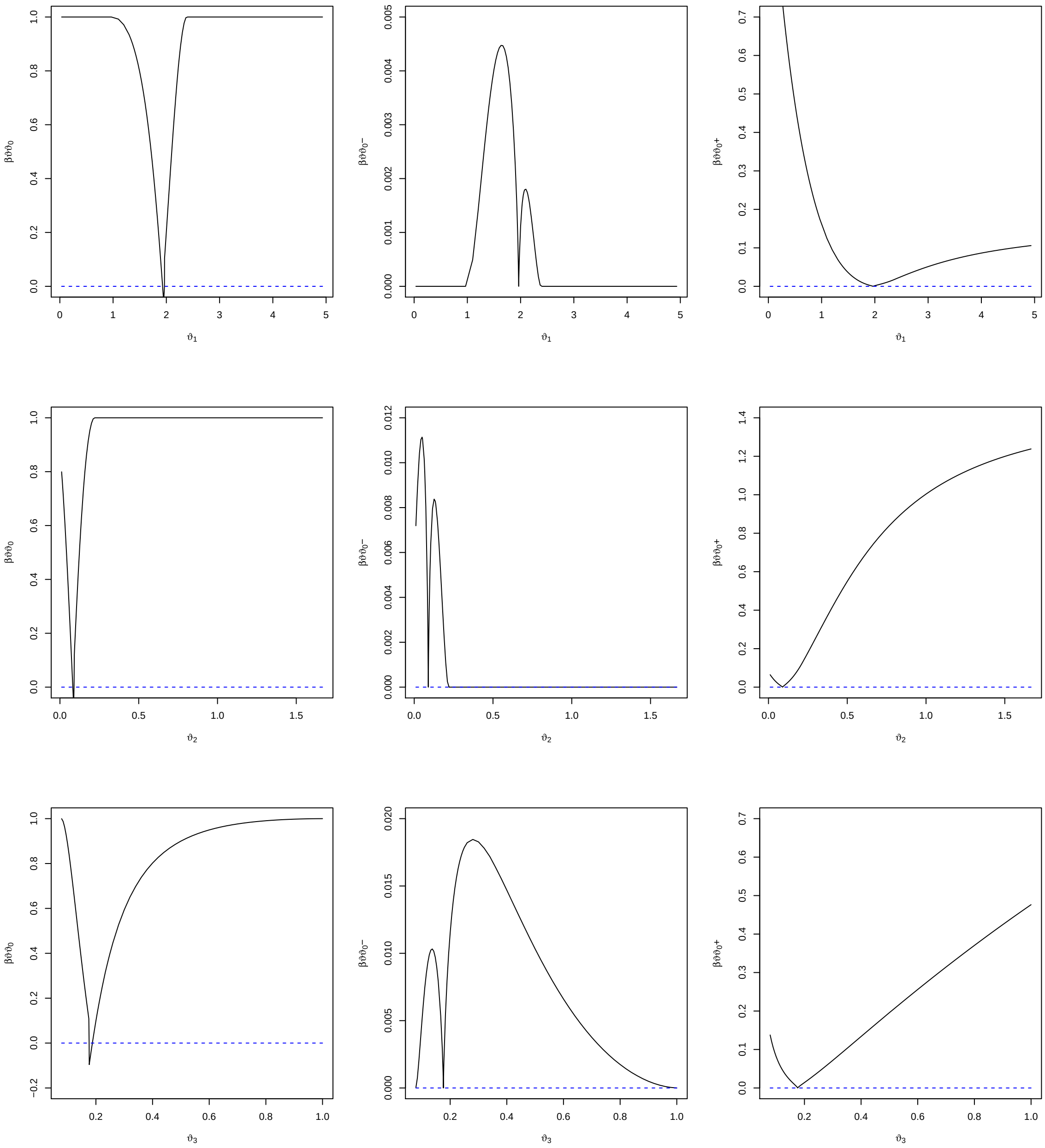

Figure 4.2.: Behavior of $\beta_{\vartheta,\vartheta_0}, \beta^-_{\vartheta,\vartheta_0}$ and $\beta^+_{\vartheta,\vartheta_0}$ in the CARMA(2, 1) model of Example 4.18 where ϑ originates from ϑ_0 when we fix two components and vary the third one. We set $\beta_{\vartheta_0,\vartheta_0} = 0$ to guarantee that $\beta_{\vartheta,\vartheta_0}$ is continuous.

Chapter 5

The integrated periodogram

Looking back to Section 3.5, the asymptotic normality of the integrated periodogram is a main part of the proof of the asymptotic normality of the Whittle estimator. Furthermore, as mentioned in Dahlhaus (1988) and Klüppelberg and Mikosch (1996), many spectral goodness-of-fit test statistics are based on the integrated periodogram as well. Consequently, from practical point of view, even though the asymptotic normality of the Whittle estimator is already derived, a functional central limit theorem for the integrated periodogram is desirable to obtain the limit behavior of the test statistics. Therefore, we define the function-indexed normalized integrated periodogram

$$\begin{aligned} E_n(g) :=& \sqrt{n} \int_{-\pi}^{\pi} g(\omega) \left(I_{n,Y}(\omega) - \frac{1}{2\pi} \Phi(e^{-i\omega}) \Sigma_N^{(\Delta)} \Phi(e^{i\omega})^\top \right) d\omega \\ =& \sqrt{n} \int_{-\pi}^{\pi} g(\omega) \left(I_{n,Y}(\omega) - f_Y^{(\Delta)}(\omega) \right) d\omega. \end{aligned}$$

Assume that a central limit theorem holds for a function class $\mathcal{G}_m$. Then, to conclude the limit behavior of statistical indices of interest such as the spectral goodness-of-fit test statistics, $\mathcal{G}_m$ has to consist of appropriate functions. Therefore, to directly derive asymptotic results for various different applications, including rich enough sets of functions is essential. We start this chapter by introducing three different sets of conditions under which a functional central limit theorem for the trace of the function-indexed normalized integrated periodogram holds. Obviously, it is a limitation to only consider a functional limit theorem for the trace. However, for most practical applications it is sufficient. Our resulting settings are rich enough to directly conclude the asymptotic distribution of some applications. Consequently, we first state the main result of this chapter and then present the limit behavior of the spectral goodness-of-fit tests. Even though we already investigated the Whittle estimator in detail in Chapter 3, we also sketch how to quickly derive its asymptotic normality with focus on the functional central limit theorem. Finally, the proofs are given.

5.1. A functional central limit theorem

In this chapter, we always investigate the trace of the process E_n indexed by subsets of 2π-periodic integrable functions. Therefore, we define

$$\mathcal{H}_m := \{g : [-\pi, \pi] \to \mathbb{C}^{m\times m} \mid g(\pi) = g(-\pi),\ \|g\|_m < \infty\},$$

where the norm $\|\cdot\|_m$ is defined by

$$\|g\|_m^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \|g(x)\|^2 dx.$$

The corresponding metric is then given by $d_m(f, g) = \|f - g\|_m$ for $f, g \in \mathcal{H}_m$. In view of what follows, we also consider the N-dimensional counterparts

$$\mathcal{H}_N := \{g : [-\pi, \pi] \to \mathbb{C}^{N\times N} \ |g(\pi) = g(-\pi),\ \|g\|_N < \infty\}$$

along with the norm and metric defined by

$$\|g\|_N^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \|g(x)\|^2 dx \quad \text{and} \quad d_N(f, g) = \|f - g\|_N,$$

respectively. However, we now mainly turn to the behavior of the trace of E_n in the metric space $(\mathcal{C}(\mathcal{G}_m), \|\cdot\|_\infty)$, where $\mathcal{C}(\mathcal{G}_m)$ is the space of univariate continuous functions on the index function set $\mathcal{G}_m \subset \mathcal{H}_m$ and $\|\cdot\|_\infty$ is the norm defined by

$$\|F\|_\infty = \sup_{g\in\mathcal{G}_m} |F(g)|.$$

Additionally, we state a set of conditions which guarantees the convergence of the trace of E_n in a linear space, namely in $(\mathcal{G}_m^{s\prime}, \|\cdot\|_{\mathcal{G}_m^{s\prime}})$ where

$$\mathcal{G}_m^{s\prime} = \{F : \mathcal{G}_m^s \to \mathbb{C}^{m\times m} \mid F \text{ linear }\} \quad \text{with} \quad \|F\|_{\mathcal{G}_m^{s\prime}} = \sup_{\substack{g\in\mathcal{G}_m^s \\ \|g\|_{\mathcal{G}_m^s}\le 1}} \|F(g)\|.$$

All in all, we start with the specification of various index classes to lay the foundations for the central limit theorem. Thereby, inter alia, we consider totally bounded spaces with different restrictions. Note that a metric space $(\mathcal{G}, d)$ is totally bounded iff its covering numbers

$$N(\varepsilon, \mathcal{G}, d) := \inf\{u |\ \exists g_1, \ldots, g_u \in \mathcal{G} : \inf_{i=1,\ldots,u} d(g - g_i) \le \varepsilon\ \forall g \in \mathcal{G}\}$$

are finite for every $\varepsilon > 0$. Furthermore, it is necessary to assume that the set $\mathcal{G}_m$ is permissible to guarantee that the supremum over the uncountable many measurable functions is measurable, see as well Pollard (1984), Appendix C.

Definition 5.1 (permissible, Definition 1 of Appendix C of Pollard (1984))**.**
Let $(\Omega, \mathcal{A}, \mathcal{P})$ be a probability space and $(S, \mathcal{S})$ be a measure space. Let further $\mathcal{F} = \{f(\cdot, t)|\ t \in T\}$, where T is a separable metric space with Borel σ-field $\mathcal{B}(T)$ and

a) the functions $f(\cdot, \cdot) \in \mathcal{F}$ are $\mathcal{S} \otimes \mathcal{B}(T)$-measurable as functions from $S \otimes T$ into $\mathbb{R}^m$,

b) T is an analytic subset of a compact metric space $\overline{T}$ (from which it inherits its metric and Borel σ-field).

Then, $\mathcal{F}$ is called permissible.

In the following, we strengthen the Lévy process assumption by assuming existing fourth moments. Since all settings assume the MCARMA process to be causal and that C has full rank, we suppose

Assumption L3.
Assumption L1 holds with $E\|L_1\|^4 < \infty$. Furthermore, C has full rank and the eigenvalues of A have strictly negative real parts.

In view of what follows, we define the ℓth Fourier coefficient of some function f as $\widehat{f}_\ell$. We additionally suppose:

Assumption C.
Assume that $\mathcal{G}_m$ is permissible and $\sup_{g \in \mathcal{G}_m} \|g\|_m < \infty$. Let further one of the following sets of conditions hold:

(C1) For some $s > 1/2$ define

$$\mathcal{G}_m := \mathcal{G}_m^s := \{g \in \mathcal{H}_m \mid \|g\|_{\mathcal{G}_m^s} < \infty\}$$

with

$$\|g\|_{\mathcal{G}_m^s}^2 = \sum_{\ell \in \mathbb{Z}} (1 + |\ell|)^{2s} \|\widehat{(g(\cdot)\Phi(e^{-i\cdot}))}_\ell\|^2.$$

(C2) Suppose $(\mathcal{G}_m, \|\cdot\|_m)$ is totally bounded. Define for $f, g \in \mathcal{G}_m$

$$d_\Phi^{(k)}(f, g) = \sup_{2^k \leq j < 2^{k+1}} d_{j,\Phi}(f, g), \quad k \in \mathbb{N}_0, \tag{5.1}$$

where

$$d_{j,\Phi}(f, g) = j \|\widehat{\left(\Phi(e^{i\cdot})^\top (f - g)(\cdot) \Phi(e^{-i\cdot})\right)}_j\|_N \quad \text{for } j \in \mathbb{N}.$$

Assume that, for some $\beta \in (0, 2)$ the covering number $N(\varepsilon, \mathcal{G}_m, d_\Phi^{(k)})$ satisfies

$$N(\varepsilon, \mathcal{G}_m, d_\Phi^{(k)}) \leq \mathfrak{C}(1 + (2^k/\varepsilon)^\beta), \quad \forall \varepsilon \in (0, 1),\ k \in \mathbb{N}_0.$$

(C3) The driving process $(L_t)_{t\in\mathbb{R}}$ satisfies $\mathbb{E}\|L_1\|^j \leq K^j$ for all $j \in \mathbb{N}$ and some $K > 0$ which is independent of j. Furthermore, suppose $(\mathcal{G}_m, \|\cdot\|_m)$ is pointwise bounded and totally bounded with

$$\int_0^1 \log(N(\varepsilon, \mathcal{G}_m, d_m))^2 d\varepsilon < \infty.$$

Finally, assume that there exists a $\widetilde{g} \in \mathcal{G}_m$ with $\|g(x)\| \leq \|\widetilde{g}(x)\|$ for all $g \in \mathcal{G}_m$ and $x \in [-\pi, \pi]$.

Remark 5.2.

a) Under Assumption C, $d_\Phi^{(k)}$ as defined in (5.1) is a metric for $k \geq 0$.

b) The condition concerning the functional space in $(C1)$ is structurally different than the ones concerning the functional spaces in $(C2)$ and $(C3)$ and can not be compared to the second and third one. However, if we consider index functions g where the component functions $g[i,j]$, $i,j \in \{1,\ldots,m\}$, form Vapnik-Chervonenkis classes (VC-classes), the condition of $(C1)$ is not satified in general. But those classes satisfy the conditions of $(C2)$ and $(C3)$, see Example 3.9 of Can et al. (2010) for the univariate case and condition $(C2)$, and Example 3.1 of Dahlhaus (1988) for $(C3)$. Even though Can et al. (2010) considered only univariate index functions, the ideas of their Example 3.9 can be transferred to coherent sets of multivariate index functions. The inclusion of the VC-classes is necessary to guarantee that the goodness-of-fit tests of Section 5.2 are included in the setting. Whereas the functional space condition in $(C3)$ is weaker than the one in $(C2)$, Assumption $(C3)$ requires that all moments of the driving Lévy process exist and are appropriately bounded. Therefore, under $(C3)$, the conditions concerning the driving process are stronger than under $(C2)$.

c) Note that by Theorem A.5, Parsevals equality can be carried out to the multivariate case with the Frobenius norm. Along with $\sup_{\omega\in[-\pi,\pi]} \|\Phi(e^{i\omega})\| \leq \mathfrak{C}$ and $\sup_{g\in\mathcal{G}_m} \|g\|_m < \infty$, this implies

$$\begin{aligned} &\sup_{g\in\mathcal{G}_m} \sum_{h\in\mathbb{Z}} \left\| \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi(e^{i\omega})^\top g(\omega) \Phi(e^{-i\omega}) e^{-ih\omega} d\omega \right\|^2 \\ = &\sup_{g\in\mathcal{G}_m} \sum_{h\in\mathbb{Z}} \left\| \widehat{\left(\Phi(e^{i\cdot})^\top g(\cdot)\Phi(e^{-i\cdot})\right)}_h \right\|^2 \\ = &\sup_{g\in\mathcal{G}_m} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\| \Phi(e^{i\omega})^\top g(\omega) \Phi(e^{-i\omega}) \right\|^2 d\omega \leq \mathfrak{C} \sup_{g\in\mathcal{G}_m} \|g\|_m^2 < \infty. \end{aligned}$$

Therefore, under the Assumptions L3 and C, the limit process of Theorem 5.3 is well-defined. The same arguments yield

$$\sup_{g\in\mathcal{G}_m} \sum_{h\in\mathbb{Z}} \|\widehat{(g(\cdot)\Phi(e^{i\cdot}))}_h\|^2 < \infty. \tag{5.2}$$

Theorem 5.3.
Suppose Assumption L3 and C hold. Then, under (C2) or (C3),

$$\operatorname{tr}(E_n) \xrightarrow{\mathcal{D}} \operatorname{tr}(E) \quad in\ (\mathcal{C}(\mathcal{G}_m), \|\cdot\|_\infty),$$

whereas, under (C1) the convergence holds in $(\mathcal{G}_m^{s\prime}, \|\cdot\|_{\mathcal{G}_m^{s\prime}})$. *The limit process* $\operatorname{tr}(E)$ *is defined by*

$$\begin{aligned}\operatorname{tr}(E(g)) = & \operatorname{tr}\left(\frac{W_0}{2\pi}\int_{-\pi}^{\pi}\Phi(e^{i\omega})^\top g(\omega)\Phi(e^{-i\omega})d\omega\right) \\ & + \operatorname{tr}\left(\sum_{h=1}^{\infty}\frac{W_h}{2\pi}\int_{-\pi}^{\pi}\Phi(e^{i\omega})^\top(g(\omega)+g(-\omega)^\top)\Phi(e^{-i\omega})e^{-ih\omega}d\omega\right)\end{aligned}$$

and W_i, $i \in \mathbb{N}_0$, *are independent Gaussian random matrices with*

$$\operatorname{vec}(W_0) \sim \mathcal{N}(0, \mathbb{E}[N_1^{(\Delta)}N_1^{(\Delta)\top} \otimes N_1^{(\Delta)}N_1^{(\Delta)\top}] - \Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)})$$

and

$$\operatorname{vec}(W_i) \sim \mathcal{N}(0, \Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)}),\ i \in \mathbb{N}.$$

Remark 5.4.

a) In case of a Brownian motion driven MCARMA(p,q) process, the limit process $\operatorname{tr}(E)$ of Theorem 5.3 can be represented by

$$\operatorname{tr}(E(g)) = \operatorname{tr}\left(\sum_{h\in\mathbb{Z}}\frac{\widetilde{W}_h}{4\pi}\int_{-\pi}^{\pi}\Phi(e^{i\omega})^\top(g(\omega)+g(-\omega)^\top)\Phi(e^{-i\omega})e^{-ih\omega}d\omega\right)$$

where $(\widetilde{W}_h)_{h\in\mathbb{Z}}$ is an i.i.d. sequence with $\widetilde{W}_1 \sim W_1$.

b) The limit process $\operatorname{tr}(E)$ is a centered Gaussian process with covariance function defined by

$$\begin{aligned}&\operatorname{Cov}(\operatorname{tr}(E(g_1)), \operatorname{tr}(E(g_2))) \\ &= \frac{1}{4\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(f_Y^{(\Delta)}(\omega)(g_1(\omega)+g_1(-\omega)^\top)^H f_Y^{(\Delta)}(\omega)(g_2(\omega)+g_2(-\omega)^\top)\right)d\omega \\ &\quad + \operatorname{vec}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi(e^{i\omega})^\top g_1(\omega)\Phi(e^{-i\omega})d\omega\right)^H \Big(\mathbb{E}[N_1^{(\Delta)}N_1^{(\Delta)\top} \otimes N_1^{(\Delta)}N_1^{(\Delta)\top}] \\ &\quad -3\Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)}\Big)\operatorname{vec}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi(e^{i\omega})^\top g_2(\omega)\Phi(e^{-i\omega})d\omega\right),\end{aligned}$$

see Section 5.3. Again, in the Brownian motion driven case, this reduces to

$$\begin{aligned}&\operatorname{Cov}(\operatorname{tr}(E(g_1)), \operatorname{tr}(E(g_2))) \\ &= \frac{1}{4\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(f_Y^{(\Delta)}(\omega)(g_1(\omega)^\top+g_1(-\omega))f_Y^{(\Delta)}(\omega)(g_2(\omega)+g_2(-\omega)^\top)\right)d\omega.\end{aligned}$$

5.2. Applications

5.2.1. The Whittle estimator

We revisit the problem of parameter estimation with the Whittle estimator. In this section, we define the Whittle function as

$$W_n^{(\star)}(\vartheta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{tr}\left(f_Y^{(\Delta)}(\omega,\vartheta)^{-1} I_{n,Y}(\omega)\right) + \log\left(\det\left(f_Y^{(\Delta)}(\omega,\vartheta)\right)\right) d\omega.$$

The minimizing argument $\widehat{\vartheta}_n^{(\star)} = \arg\min_{\vartheta\in\Theta} W_n^{(\star)}(\vartheta)$ is the Whittle estimator. Obviously, this estimator is asymptotically equivalent to the one defined in Chapter 3 which justifies that we give it the same name. Slightly stricter assumptions are necessary to apply Theorem 5.3 to obtain the asymptotic normality of the Whittle estimator. Again, we consider an equidistantly sampled parameterized MCARMA(p,q) process with fixed $0 \le q < p$. As before, we let Assumptions A and B hold and we furthermore assume:

(W1) For every $\vartheta \in \Theta$, $\mathbb{E}\|L_1(\vartheta)\|^j \le K(\vartheta)^j$ where $K(\vartheta) > 0$ is a positive constant which depends on ϑ but is independent of j.

(W2) The eigenvalues of $A(\vartheta)$ are distinct for every $\vartheta \in \Theta$.

(W3) $\{\frac{\partial}{\partial\vartheta_i} f_Y^{(\Delta)}(\cdot,\vartheta)^{-1} : \vartheta \in \Theta\}$ and $\{\frac{\partial^2}{\partial\vartheta_i\partial\vartheta_j} f_Y^{(\Delta)}(\cdot,\vartheta)^{-1} : \vartheta \in \Theta\}$ satisfy Assumption $(C1), (C2)$ or $(C3)$ for $i, j \in \{1,\ldots,r\}$.

(W4) For the gradient vector and the Hessian matrix of $W_n^{(\star)}$ the order of integration and differentiation can be interchanged, i.e.,

$$\nabla_\vartheta W_n^{(\star)}(\vartheta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \nabla_\vartheta \left(\operatorname{tr}\left(f_Y^{(\Delta)}(\omega,\vartheta)^{-1} I_{n,Y}(\omega)\right) + \log\left(\det\left(f_Y^{(\Delta)}(\omega,\vartheta)\right)\right)\right) d\omega$$

and

$$\nabla_\vartheta^2 W_n^{(\star)}(\vartheta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \nabla_\vartheta^2 \left(\operatorname{tr}\left(f_Y^{(\Delta)}(\omega,\vartheta)^{-1} I_{n,Y}(\omega)\right) + \log\left(\det\left(f_Y^{(\Delta)}(\omega,\vartheta)\right)\right)\right) d\omega$$

hold.

Now, we can directly conclude the asymptotic normality of the Whittle estimator by the following observations: By Theorem 3.8 of Fasen-Hartmann and Scholz (2021), the sampled processes $Y^{(\Delta)}(\vartheta)$ have weak VARMA representations. Note that $A(\vartheta)$ has to have distinct eigenvalues to apply this theorem. Therefore, the corresponding inverse spectral densities form a finite-dimensional vector space and which gives that the sets $\mathcal{G}_m[i,j] := \{f_Y^{(\Delta)}(\cdot,\vartheta)^{-1}[i,j] | \vartheta \in \Theta\}$ are VC-classes. Furthermore, as before, Assumption $(A4)$ implies that $f_Y^{(\Delta)}(\cdot)^{-1}$ is continuous on the compact set $[-\pi,\pi]\times\Theta$ which yields

$$\max_{\vartheta\in\Theta}\max_{\omega\in[-\pi,\pi]} \|f_Y^{(\Delta)}(\omega,\vartheta)^{-1}\| \le \mathfrak{C}.$$

Consequently, a measurable envelope function exists and we can apply Lemma 2.25 of Pollard (1984). Therefore, Assumption $(C3)$ holds. An application of Theorem 5.3 gives

$$\sup_{\vartheta\in\Theta}\sqrt{n}\left|W_n^{(\star)}(\vartheta)-\frac{1}{2\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left(f_Y^{(\Delta)}(\omega,\vartheta)^{-1}f_Y^{(\Delta)}(\omega)\right)+\log\left(\det\left(f_Y^{(\Delta)}(\omega,\vartheta)\right)\right)d\omega\right|\xrightarrow{\mathcal{D}}0.$$

As in the original proof of the asymptotic normality of the Whittle estimator, a Taylor expansion of order one yields

$$\sqrt{n}(\widehat{\vartheta}_n^{(\star)}-\vartheta_0)^\top=-\sqrt{n}\nabla_\vartheta W_n^{(\star)}(\vartheta_0)\left[\nabla_\vartheta^2 W_n^{(\star)}(\vartheta_n')\right]^{-1},$$

for $\|\vartheta_n'-\vartheta_0\|\leq\|\widehat{\vartheta}_n^{(\star)}-\vartheta_0\|$. Since Assumptions $(W3)$ and $(W4)$ hold, two applications of Theorem 5.3 and $\vartheta_n'\overset{a.s.}{\to}\vartheta_0$ complete the proof. Obviously, these considerations need the consistency of $\widehat{\vartheta}_n^{(\star)}$ which can be proven as in Theorem 3.8 as well as the invertibility of the matrix $\nabla_\vartheta^2 W_n^{(\star)}(\vartheta_n^*)$ which is given by Lemma 3.17.

Remark 5.5.

If we consider univariate CARMA processes or the class of MCAR(1) processes, the assumptions can be weakened. Namely, it is possible to omit the moment condition of the driving process $(W1)$ and replace $(W2)$ by

$(\widetilde{W}2)$ $\{\frac{\partial}{\partial\vartheta_i}f_Y^{(\Delta)}(\cdot,\vartheta)^{-1}:\vartheta\in\Theta\}$ and $\{\frac{\partial^2}{\partial\vartheta_i\partial\vartheta_j}f_Y^{(\Delta)}(\cdot,\vartheta)^{-1}:\vartheta\in\Theta\}$ satisfy Assumption $(C1)$ or $(C2)$ for $i,j\in\{1,\ldots,r\}$.

For the proof, define the set of index functions $\mathcal{G}_m=\{f_Y^{(\Delta)}(\cdot,\vartheta)^{-1}|\vartheta\in\Theta\}$.

Case 1: univariate CARMA(p,q) processes:

The spectral densities $f_Y^{(\Delta)}(\cdot,\vartheta)$, $\vartheta\in\Theta$, are symmetric. Therefore, an application of the Cauchy-Schwarz inequality and (2.10) yield

$$\begin{aligned}
&\sum_{\ell\in\mathbb{Z}}(1+|\ell|)^{2s}\left\|\widehat{\left(f_Y^{(\Delta)}(\cdot,\vartheta)^{-1}\Phi(e^{-i\cdot})\right)_\ell}\right\|^2\\
&=\sum_{\ell\in\mathbb{Z}}(1+|\ell|)^{2s}\left\|\frac{1}{2\pi}\int_{-\pi}^{\pi}f_Y^{(\Delta)}(\omega,\vartheta)^{-1}\sum_{j=1}^{\infty}\Phi_j e^{-i\omega(j+\ell)}d\omega\right\|^2\\
&=\sum_{\ell\in\mathbb{Z}}(1+|\ell|)^{2s}\left\|\sum_{j=0,j=-\ell}^{\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}f_Y^{(\Delta)}(\omega,\vartheta)^{-1}d\omega\Phi_j\right\|^2\\
&\leq\mathfrak{C}\sum_{\ell\in\mathbb{Z}}(1+|\ell|)^{2s}\sum_{j=0,j=-\ell}^{\infty}\frac{1}{(1+|j|)^{4s}}\left\|\frac{1}{2\pi}\int_{-\pi}^{\pi}f_Y^{(\Delta)}(\omega,\vartheta)^{-1}d\omega\right\|^2<\infty.
\end{aligned}$$

Consequently, Assumption $(C1)$ holds.

Case 2: MCAR(1) processes:

Notice that $\Phi^{-1}(\cdot,\vartheta)$ exists and has the representation $\Phi^{-1}(x,\vartheta)=(I_N-e^{A(\vartheta)\Delta}x)$.

Therefore,

$$\sum_{\ell\in\mathbb{Z}}(1+|\ell|)^{2s}\left\|\overline{\left(f_Y^{(\Delta)}(\cdot,\vartheta)^{-1}\Phi(e^{-i\cdot})\right)}_\ell\right\|^2$$

$$= \sum_{\ell\in\mathbb{Z}}(1+|\ell|)^{2s}\left\|\frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{j=0}^{\infty}\left(\Sigma_N^{(\Delta)-1}-e^{A(\vartheta)^\top\Delta}\Sigma_N^{(\Delta)-1}e^{i\omega}\right.\right.$$
$$\left.\left.-\Sigma_N^{(\Delta)-1}e^{A(\vartheta)\Delta}e^{-i\omega}+e^{A(\vartheta)^\top\Delta}\Sigma_N^{(\Delta)-1}e^{A(\vartheta)\Delta}\right)\Phi_j e^{-i\omega(j+\ell)}d\omega\right\|^2.$$

Again, the integral vanishes for $\ell\notin\{-j-1,-j,-j+1\}$. Consequently, the same ideas as in the first case imply that $(C1)$ holds.
The same reasoning as before now yields the asymptotic normality in both cases.

5.2.2. Goodness-of-fit tests

In this section, we investigate the behavior of some empirical goodness-of-fit test statistics which are based on the integrated periodogram. Note that by Example 3.9 of Can et al. (2010), the set of indicator functions satisfies the function class condition of $(C2)$. Since $\sup_{\omega\in[-\pi,\pi]}\|f_Y^{(\Delta)}(\omega)^{-1}\|\le\mathfrak{C}$ holds, minor adaptions yield that $\{\lambda\in[-\pi,\pi]:\mathbb{1}\{\cdot\le\lambda\}f_Y^{(\Delta)}(\cdot)^{-1}\}$ satisfies the conditions for a fixed spectral density $f_Y^{(\Delta)}$ as well. Consequently, the asymptotic behavior of the subsequent test statistics follows from an application of Theorem 5.3 and the continuous mapping theorem.

Corollary 5.6.
Let Assumption C hold and let $W_i,\widetilde{W}_i$, $i\in\mathbb{Z}$, be the Gaussian random matrices of Theorem 5.3 and Remark 5.4 a). Then:

a) The Grenander-Rosenblatt statistic converges in distribution, i.e.,

$$\sqrt{n}\sup_{x\in[-\pi,\pi]}\left|\operatorname{tr}\left(\int_{-\pi}^{x}I_{n,Y}(\omega)-\frac{1}{2\pi}\Phi(e^{-i\omega})\Sigma_N^{(\Delta)}\Phi(e^{i\omega})^\top d\omega\right)\right|$$
$$\xrightarrow{\mathcal{D}}\sup_{x\in[-\pi,\pi]}\left|\operatorname{tr}\left(\frac{W_0}{2\pi}\int_{-\pi}^{x}\Phi(e^{i\omega})^\top\Phi(e^{-i\omega})d\omega\right)\right.$$
$$\left.+\operatorname{tr}\left(\sum_{h=1}^{\infty}\frac{W_h}{2\pi}\left(\int_{-\pi}^{x}\Phi(e^{i\omega})^\top\Phi(e^{-i\omega})e^{-ih\omega}d\omega+\int_{-x}^{\pi}\Phi(e^{i\omega})^\top\Phi(e^{-i\omega})e^{-ih\omega}d\omega\right)\right)\right|.$$

If the driving process is a Brownian motion, the limit process has the representation

$$\sup_{x\in[-\pi,\pi]}\left|\operatorname{tr}\left(\sum_{h=-\infty}^{\infty}\frac{\widetilde{W}_h}{4\pi}\left(\int_{-\pi}^{x}\Phi(e^{i\omega})^\top\Phi(e^{-i\omega})e^{-ih\omega}d\omega\right.\right.\right.$$
$$\left.\left.\left.+\int_{-x}^{\pi}\Phi(e^{i\omega})^\top\Phi(e^{-i\omega})e^{-ih\omega}d\omega\right)\right)\right|.$$

b) The Cramér-von Mises statistic converges in distribution, i.e.,

$$
\begin{aligned}
& n \int_{-\pi}^{\pi} \operatorname{tr}\left(\int_{-\pi}^{x} I_{n,Y}(\omega) - \frac{1}{2\pi}\Phi(e^{-i\omega})\Sigma_N^{(\Delta)}\Phi(e^{i\omega})^\top d\omega\right)^2 dx \\
& \quad \xrightarrow{\mathcal{D}} \int_{-\pi}^{\pi} \operatorname{tr}\left(\frac{W_0}{2\pi}\int_{-\pi}^{x} \Phi(e^{i\omega})^\top \Phi(e^{-i\omega}) d\omega \right.\\
& \qquad \left. + \sum_{h=1}^{\infty} \frac{W_h}{2\pi}\left(\int_{-\pi}^{x} \Phi(e^{i\omega})^\top \Phi(e^{-i\omega}) e^{-ih\omega} d\omega + \int_{-x}^{\pi} \Phi(e^{i\omega})^\top \Phi(e^{-i\omega}) e^{-ih\omega} d\omega\right)\right)^2 dx.
\end{aligned}
$$

If the driving process is a Brownian motion, the limit process has the representation

$$
\begin{aligned}
\int_{-\pi}^{\pi} \operatorname{tr}\left(\sum_{h=-\infty}^{\infty} \frac{\widetilde{W}_h}{4\pi}\left(\int_{-\pi}^{x} \Phi(e^{i\omega})^\top \Phi(e^{-i\omega}) e^{-ih\omega} d\omega \right.\right.\\
\left.\left. + \int_{-x}^{\pi} \Phi(e^{i\omega})^\top \Phi(e^{-i\omega}) e^{-ih\omega} d\omega\right)\right)^2 dx.
\end{aligned}
$$

c) The test statistic of Dahlhaus (1988), Example 3.4, converges in distribution, i.e.,

$$
\begin{aligned}
& \frac{\sqrt{n}}{\sqrt{m}} \sup_{x\in[-\pi,\pi]} \left| \int_{-\pi}^{x} \operatorname{tr}\left(I_{n,Y}(\omega) f_Y^{(\Delta)}(\omega)^{-1}\right) d\omega - (x+\pi)m \right| \\
& \quad \xrightarrow{\mathcal{D}} \sup_{x\in[-\pi,\pi]} \left| \operatorname{tr}\left(\frac{W_0}{2\pi\sqrt{m}} \int_{-\pi}^{x} \Phi(e^{i\omega})^\top f_Y^{(\Delta)}(\omega)^{-1} \Phi(e^{-i\omega}) d\omega\right) \right.\\
& \qquad + \operatorname{tr}\left(\sum_{h=1}^{\infty} \frac{W_h}{2\pi\sqrt{m}} \left(\int_{-\pi}^{x} \Phi(e^{i\omega})^\top f_Y^{(\Delta)}(\omega)^{-1} \Phi(e^{-i\omega}) e^{-ih\omega} d\omega \right.\right.\\
& \qquad \left.\left.\left. + \int_{-x}^{\pi} \Phi(e^{i\omega})^\top f_Y^{(\Delta)}(\omega)^{-1} \Phi(e^{-i\omega}) e^{-ih\omega} d\omega\right)\right) \right|.
\end{aligned}
$$

If the driving process is a Brownian motion, the limit process has the representation

$$
\begin{aligned}
\sup_{x\in[-\pi,\pi]} \left| \operatorname{tr}\left(\sum_{h=-\infty}^{\infty} \frac{\widetilde{W}_h}{4\pi\sqrt{m}} \left(\int_{-\pi}^{x} \Phi(e^{i\omega})^\top f_Y^{(\Delta)}(\omega)^{-1} \Phi(e^{-i\omega}) e^{-ih\omega} d\omega \right.\right.\right.\\
\left.\left.\left. + \int_{-x}^{\pi} \Phi(e^{i\omega})^\top f_Y^{(\Delta)}(\omega)^{-1} \Phi(e^{-i\omega}) e^{-ih\omega} d\omega\right)\right) \right|.
\end{aligned}
$$

d) The Cramér-von Mises statistic with self-normalization converges in distribution, i.e.,

$$
\begin{aligned}
& \frac{n}{m} \int_{-\pi}^{\pi} \left(\int_{-\pi}^{x} \operatorname{tr}\left(I_{n,Y}(\omega) f_Y^{(\Delta)}(\omega)^{-1}\right) d\omega - (x+\pi)m\right)^2 dx \\
& \quad \xrightarrow{\mathcal{D}} \frac{1}{m} \int_{-\pi}^{\pi} \left(\operatorname{tr}\left(\frac{W_0}{2\pi} \int_{-\pi}^{x} \Phi(e^{i\omega})^\top f_Y^{(\Delta)}(\omega)^{-1} \Phi(e^{-i\omega}) d\omega\right) \right.\\
& \qquad + \operatorname{tr}\left(\sum_{h=1}^{\infty} \frac{W_h}{2\pi} \left(\int_{-\pi}^{x} \Phi(e^{i\omega})^\top f_Y^{(\Delta)}(\omega)^{-1} \Phi(e^{-i\omega}) e^{-ih\omega} d\omega \right.\right.\\
& \qquad \left.\left.\left. + \int_{-x}^{\pi} \Phi(e^{i\omega})^\top f_Y^{(\Delta)}(\omega)^{-1} \Phi(e^{-i\omega}) e^{-ih\omega} d\omega\right)\right)\right)^2 dx.
\end{aligned}
$$

If the driving process is a Brownian motion, the limit process has the representation

$$\frac{1}{m}\int_{-\pi}^{\pi}\left(\operatorname{tr}\left(\sum_{h=-\infty}^{\infty}\frac{\widetilde{W}_h}{4\pi}\left(\int_{-\pi}^{x}\Phi(e^{i\omega})^\top f_Y^{(\Delta)}(\omega)^{-1}\Phi(e^{-i\omega})e^{-ih\omega}d\omega\right.\right.\right.$$
$$\left.\left.\left.+\int_{-x}^{\pi}\Phi(e^{i\omega})^\top f_Y^{(\Delta)}(\omega)^{-1}\Phi(e^{-i\omega})e^{-ih\omega}d\omega\right)\right)\right)^2 dx.$$

Remark 5.7.

a) Consider the class of MCAR(1) processes and the test statistics of Corollary 5.6 part c) and d). Note that $m = N$. For W_i $i \in \mathbb{N}_0$, defined as in Theorem 5.3,

$$Z_0 := \operatorname{tr}(W_0\Sigma_N^{(\Delta)-1}) \sim \mathcal{N}\left(0, \mathbb{E}\left[\operatorname{tr}\left(N_1^{(\Delta)}N_1^{(\Delta)\top}\Sigma_N^{(\Delta)-1}N_1^{(\Delta)}N_1^{(\Delta)\top}\Sigma_N^{(\Delta)-1}\right)\right] - m\right),$$
$$\text{and } Z_i := \operatorname{tr}(W_i\Sigma_N^{(\Delta)-1}) \sim \mathcal{N}\left(0, m\right), i \in \mathbb{N},$$

hold by properties T2.4, T3.4 and T3.8 of Brewer (1978). By the Karhunen-Loève expansion,

$$\sqrt{2}\sum_{h=1}^{\infty}\frac{\sin(\pi h x)}{h\pi}N_h, \quad N_h \overset{iid}{\sim} \mathcal{N}(0,1),$$

is a Brownian bridge on $[0,1]$. Therefore, the process $(G(x))_{x\in[-\pi,\pi]}$ with

$$G(x) = \frac{x+\pi}{\sqrt{m}}Z_0 + \sum_{h=1}^{\infty}\frac{2\sin(hx)}{h\sqrt{m}}Z_h, \quad x \in [-\pi,\pi],$$

is a centered Gaussian process with covariance function

$$\Sigma_G(s,t) = \operatorname{Var}(Z_0)\frac{(t+\pi)(s+\pi)}{m} + 2\left(s(\operatorname{sign}(t)\pi - t)\right), \quad 0 \le |s| \le |t|.$$

Accordingly, the limit processes have the representations

$$\sup_{x\in[-\pi,\pi]}|G(x)| \quad \text{and} \quad \int_{-\pi}^{\pi}G(x)^2dx,$$

respectively. In case of the Brownian motion driven MCAR(1) process, $Z_0 \sim \mathcal{N}(0,2m)$, which implies that the limit distributions are independent of parameters of A.

b) Several spectral goodness-of-fit tests were already investigated for ARMA processes. By Section 6.2.6 of Priestley (1981), the limit distribution of the Grenander-Rosenblatt statistic for ARMA processes with normally distributed white noise is similar to ours. In contrast, the convergence rate and limit distributions of the goodness-of-fit statistics of Corollary 5.6 differ from ours when the ARMA process is assumed to be α-stable. It should be noted, that the term corresponding to W_0 vanishes in that case, see Section 4 of Klüppelberg and Mikosch (1996).

5.3. Proof of Theorem 5.3

We now prove Theorem 5.3. As a first step, we prove a central limit theorem like Theorem 5.3 but with $(Y_k^{(\Delta)})_{k\in\mathbb{Z}}$ replaced by the white noise sequence $(N_k^{(\Delta)})_{k\in\mathbb{Z}}$ as introduced in Theorem 2.5 and adapted sets of index functions. Then, we show that the error which occurs by approximating the original process by the white noise process is sufficiently small.

5.3.1. The functional central limit theorem for the white noise process

We introduce the assumptions concerning the function-indexed periodogram of the white noise process which correspond to those of Assumption C.

Assumption N.
Assume that $\mathcal{G}_N$ is permissible. Either of the following conditions hold:

(N1) Define for some $s > 1/2$

$$\mathcal{G}_N := \mathcal{G}_N^s := \{g \in \mathcal{H}_N : \|g\|_{\mathcal{G}_N^s} < \infty\}$$

where

$$\|g\|_{\mathcal{G}_N^s}^2 := \sum_{\ell\in\mathbb{Z}} (1+|\ell|)^{2s} \|\widehat{g}_\ell\|^2.$$

(N2) Assume that the set $\mathcal{G}_N$ satisfies $\sup_{g\in\mathcal{G}_N} \|g\|_N < \infty$. For $f, g \in \mathcal{G}_N$ and $j \in \mathbb{N}$ we define

$$d_j(f,g) = j\|\widehat{f}_j - \widehat{g}_j\|$$

and thereby

$$d^{(k)}(f,g) = \sup_{2^k \le j < 2^{k+1}} d_j(f,g).$$

Then, let for some $\beta \in (0,2)$ the covering number $N(\varepsilon, \mathcal{G}_N, d^{(k)})$ satisfy

$$N(\varepsilon, \mathcal{G}_N, d^{(k)}) \le \mathfrak{C}(1 + (2^k/\varepsilon)^\beta), \quad \forall \varepsilon \in (0,1), k \in \mathbb{N}_0.$$

(N3) Suppose that the driving process $(L_t)_{t\in\mathbb{R}}$ satisfies $\mathbb{E}\|L_1\|^j \le K^j$ for some $K > 0$ and $j \in \mathbb{N}$. Furthermore, let $(\mathcal{G}_N, d_N)$ be pointwise bounded and totally bounded with

$$\int_0^1 \log(N(\varepsilon, \mathcal{G}_N, d_N))^2 d\varepsilon < \infty.$$

Finally, assume that there exists $\widetilde{g} \in \mathcal{G}_N$ with $\|g(x)\| \le \|\widetilde{g}(x)\|$ for all $g \in \mathcal{G}_N$.

For the white noise process, we consider a setting which is more general. Namely, under $(N2)$ and $(N3)$, we investigate a multivariate process in the space $(\mathcal{C}(\mathcal{G}_N), \|\cdot\|_{N,\infty})$ where $\mathcal{C}(\mathcal{G}_N)$ is the space of complex-valued $N \times N$-dimensional continuous functions on $\mathcal{G}_N$ and

$\|\cdot\|_{N,\infty}$ is defined by

$$\|F\|_{N,\infty} = \sup_{g\in\mathcal{G}_N} \|F(g)\|.$$

In the same way, the space $(\mathcal{C}(\mathcal{G}_m), \|\cdot\|_{m,\infty})$ is defined. Under the condition $(N1)$, we again consider weak convergence in a linear space. Therefore, we define

$$\mathcal{G}_N^{s\prime} = \{F : \mathcal{G}_N^s \to \mathbb{C}^{N\times N} \mid F \text{ linear }\} \quad \text{with} \quad \|F\|_{\mathcal{G}_N^{s\prime}} = \sup_{\substack{g\in\mathcal{G}_N^s \\ \|g\|_{\mathcal{G}_N^s}\leq 1}} \|F(g)\|.$$

Theorem 5.8.
Let Assumption L3 and N hold. Define

$$E_{n,N}(g) := \sqrt{n}\int_{-\pi}^{\pi} g(\omega)\left(I_{n,N}(\omega) - \frac{1}{2\pi}\Sigma_N^{(\Delta)}\right) d\omega.$$

Then, if $(N2)$ or $(N3)$ hold,

$$E_{n,N} \xrightarrow{\mathcal{D}} E_N \quad \text{in } (\mathcal{C}(\mathcal{G}_N), \|\cdot\|_{N,\infty}),$$

whereas, under $(N1)$ the process converges in $(\mathcal{G}_N'\|\cdot\|_{\mathcal{G}_N^{s\prime}})$. The process E_N is defined by

$$E_N(g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(\omega)W_0 d\omega + \sum_{h=1}^{\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi} g(\omega)W_h e^{-ih\omega} + g(\omega)W_h^\top e^{ih\omega} d\omega$$

with W_i, $i\in\mathbb{N}_0$, as in Theorem 5.3.

It is well known, that a sequence of probability measures in some Banach space converges weakly if it is tight in the weak topology and if the finite dimensional distributions converge. Therefore, we first prove the convergence in distribution of the finite dimensional distributions of $E_{n,N}$ under the conditions of Theorem 5.8.

Convergence of the finite dimensional distributions

Lemma 5.9.
Let Assumption N hold and $\mathbb{E}\|L_1\|^4 < \infty$. Then, for all $k\in\mathbb{N}$, $g_1,\ldots,g_k\in\mathcal{G}_N$

$$(E_{n,N}(g_1),\ldots,E_{n,N}(g_k)) \xrightarrow{\mathcal{D}} (E_N(g_1),\ldots,E_N(g_k)).$$

Proof. By construction, $\mathcal{G}_N$ is a linear space. Therefore, for fixed $k\in\mathbb{N}$, $g_1,\ldots,g_k\in\mathcal{G}_N$ and $c_1\ldots,c_k$ in $\mathbb{R}$, $c_1g_1+\ldots+c_kg_k\in\mathcal{G}_N$. By Cramér-Wold, it is sufficient to prove

$$E_{n,N}(g) \xrightarrow{\mathcal{D}} E_N(g)$$

for fixed $g \in \mathcal{G}_N$. Note that

$$\begin{aligned}
&\left\| \sqrt{n} \int_{-\pi}^{\pi} g(\omega) \left(I_{n,N}(\omega) - \frac{1}{2\pi} \Sigma_N^{(\Delta)} \right) d\omega \right\| \\
= \quad & \left\| \sqrt{n} \int_{-\pi}^{\pi} \sum_{\ell \in \mathbb{Z}} \widehat{g}_\ell e^{i\omega\ell} \left(\frac{1}{2\pi n} \sum_{j,k=1}^{n} N_j^{(\Delta)} N_k^{(\Delta)\top} e^{-i\omega(j-k)} - \frac{1}{2\pi} \Sigma_N^{(\Delta)} \right) d\omega \right\| \\
= \quad & \left\| \frac{1}{\sqrt{n}} \sum_{j,k=1}^{n} \widehat{g}_{j-k} N_j^{(\Delta)} N_k^{(\Delta)\top} - \sqrt{n} \widehat{g}_0 \Sigma_N^{(\Delta)} \right\| \\
= \quad & \left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \sum_{c=1-k}^{n-k} \widehat{g}_c N_{c+k}^{(\Delta)} N_k^{(\Delta)\top} - \sqrt{n} \widehat{g}_0 \Sigma_N^{(\Delta)} \right\| \\
= \quad & \left\| \frac{1}{\sqrt{n}} \sum_{c=1}^{n-1} \sum_{k=1}^{n-c} \widehat{g}_c N_{c+k}^{(\Delta)} N_k^{(\Delta)\top} + \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \widehat{g}_0 \left(N_k^{(\Delta)} N_k^{(\Delta)\top} - \Sigma_N^{(\Delta)} \right) \right. \\
& \left. + \frac{1}{\sqrt{n}} \sum_{c=1-n}^{-1} \sum_{k=1-c}^{n} \widehat{g}_c N_{c+k}^{(\Delta)} N_k^{(\Delta)\top} \right\| \\
= \quad & \left\| \sqrt{n} \sum_{c=1-n}^{n-1} \widehat{g}_c \left(\overline{\Gamma}_{n,N}(c) - \mathbb{E}[\overline{\Gamma}_{n,N}(c)] \right) \right\|. \qquad (5.3)
\end{aligned}$$

Therefore,

$$E_{n,N}(g) = \sqrt{n} \widehat{g}_0 (\overline{\Gamma}_{n,N}(0) - \Sigma_N^{(\Delta)}) + \sqrt{n} \left(\sum_{h=1}^{n-1} \widehat{g}_h \overline{\Gamma}_{n,N}(h) + \widehat{g}_{-h} \overline{\Gamma}_{n,N}(h)^\top \right)$$

holds. We fix an upper bound for h, say M, to apply Lemma 2.16. Thus, we have

$$\begin{aligned}
\sqrt{n} \widehat{g}_0 (\overline{\Gamma}_{n,N}(0) - \Sigma_N^{(\Delta)}) + \sqrt{n} \sum_{h=1}^{M} \left(\widehat{g}_h \overline{\Gamma}_{n,N}(h) + \widehat{g}_{-h} \overline{\Gamma}_{n,N}(h)^\top \right) \\
\xrightarrow{\mathcal{D}} \widehat{g}_0 W_0 + \sum_{h=1}^{M} \left(\widehat{g}_h W_h + \widehat{g}_{-h} W_h^\top \right),
\end{aligned}$$

where $W_j, j = 0, \ldots, h$, are the Gaussian random matrices as defined in Theorem 5.8. In view of Proposition 6.3.9 of Brockwell and Davis (1991), it remains to prove

$$\lim_{M \to \infty} \limsup_{n \to \infty} \mathbb{P} \left(\left\| \sqrt{n} \sum_{h=M+1}^{n-1} \widehat{g}_h \overline{\Gamma}_{n,N}(h) + \widehat{g}_{-h} \overline{\Gamma}_{n,N}(h)^\top \right\| > \varepsilon \right) = 0. \qquad (5.4)$$

Tschebycheffs inequality leads to

$$\mathbb{P} \left(\left\| \sqrt{n} \sum_{h=M+1}^{n-1} \widehat{g}_h \overline{\Gamma}_{n,N}(h) + \widehat{g}_{-h} \overline{\Gamma}_{n,N}(h)^\top \right\| > \varepsilon \right)$$

$$\leq \frac{n}{\varepsilon^2}\mathbb{E}\left[\left\|\sum_{h=M+1}^{n-1} \widehat{g}_h\overline{\Gamma}_{n,N}(h) + \widehat{g}_{-h}\overline{\Gamma}_{n,N}(h)^\top\right\|^2\right].$$

Since $(N_k^{(\Delta)})_{k\in\mathbb{Z}}$ is i.i.d., $\overline{\Gamma}_{n,N}(h)$ and $\overline{\Gamma}_{n,N}(j)$ are uncorrelated for $h \neq j$. We therefore get

$$\begin{aligned}
&\mathbb{E}\left[\left\|\sum_{h=M+1}^{n-1} \left(\widehat{g}_h\overline{\Gamma}_{n,N}(h) + \widehat{g}_{-h}\right)\overline{\Gamma}_{n,N}(h)^\top\right\|^2\right] \\
&= \sum_{h=M+1}^{n-1} \mathbb{E}\left[\left\|\widehat{g}_h\overline{\Gamma}_{n,N}(h) + \widehat{g}_{-h}\overline{\Gamma}_{n,N}(h)^\top\right\|^2\right] \\
&\leq 2\sum_{h=M+1}^{n-1} \left(\|\widehat{g}_h\|^2 + \|\widehat{g}_{-h}^\top\|^2\right)\mathbb{E}\left[\|\overline{\Gamma}_{n,N}(h)\|^2\right].
\end{aligned}$$

We obtain

$$\mathbb{E}\left[n\|\overline{\Gamma}_{n,N}(h)\|^2\right] = \frac{1}{n}\sum_{s,t=1}^{N}\sum_{k,\ell=1}^{n-h}\mathbb{E}\left[N_{k+h}^{(\Delta)}N_k^{(\Delta)\top}N_{\ell+h}^{(\Delta)}N_\ell^{(\Delta)\top}[s,t]\right] \leq \frac{n-h}{n}K, \tag{5.5}$$

where K is a constant which is independent of h. On the other hand, Parsevals equality yields for $s,t \in \{1,\ldots,N\}$

$$\sum_{h\in\mathbb{Z}} |\widehat{g}_h[s,t]|^2 = \int_{-\pi}^{\pi} |g(x)[s,t]|^2 dx.$$

Since $\int_{-\pi}^{\pi} \|g(x)\|^2 dx < \infty$ due to $g \in \mathcal{H}_N$ the proof is completed. □

TIGHTNESS

To prove Theorem 5.8 it remains to show the tightness of $(E_{n,N})_{n\in\mathbb{N}}$ under the different conditions. We do separate investigations for $(N1), (N2)$ and $(N3)$, respectively. For the set $(N1)$ we make use of Proposition 2.2 of Bharucha-Reid and Römisch (1985) which gives a criterion for proving the tightness. Therefore, we need the definitions of uniform boundedness and flat concentration of a sequence of measures $(P_n)_{n\in\mathbb{N}}$ on a separable Banach space (S,d). By Definition 2.1 of Bharucha-Reid and Römisch (1985) $(P_n)_{n\in\mathbb{N}}$ is called uniformly bounded iff for all $\varepsilon > 0$ there exists a constant $K(\varepsilon)$ with

$$\inf_{n\in\mathbb{N}} P_n(\{x \in S \mid \|x\| \leq K(\varepsilon)\}) \geq 1-\varepsilon.$$

The sequence of measures $(P_n)_{n\in\mathbb{N}}$ is called flatly concentrated iff for all $\varepsilon, \delta > 0$ there exists a finite-dimensional subspace $L \subset S$ with

$$\inf_{n\in\mathbb{N}} P_n(\{x \in S \mid d(x,L) \leq \delta\}) \geq 1-\varepsilon.$$

Proposition 5.10 (Proposition 2.2 of Bharucha-Reid and Römisch (1985))**.**
Let S be a separable Banach space and $\mathcal{P}(S)$ the set of all probability measures on $(S, \mathcal{B}(S))$. A sequence $(P_n)_{n\in\mathbb{N}} \in \mathcal{P}(S)$ is tight if and only if $(P_n)_{n\in\mathbb{N}}$ is uniformly bounded and flatly concentrated.

Following the ideas of Bardet et al. (2008), we now show the tightness in the space $(\mathcal{G}_N^{s\prime}, \|\cdot\|_{\mathcal{G}_N^{s\prime}})$.

Lemma 5.11.
Suppose $(N1)$ holds. Then, the sequence $(E_{n,N})_{n\in\mathbb{N}}$ is tight in $(\mathcal{G}_N'\|\cdot\|_{\mathcal{G}_N^{s\prime}})$.

Proof. In view of Proposition 5.10, we show that $(\mathbb{P}^{E_{n,N}})_{n\in\mathbb{N}}$ is uniformly bounded and flatly concentrated. For the uniform boundedness we have to prove that for every $\varepsilon > 0$ there exists some $K = K(\varepsilon)$ such that

$$\inf_{n\in\mathbb{N}} \mathbb{P}(\|E_{n,N}\|_{\mathcal{G}_N^{s\prime}} \leq K) \geq 1 - \varepsilon. \tag{5.6}$$

Note that by representation (5.3) the equivalences

$$\begin{aligned}
\|E_{n,N}\|_{\mathcal{G}_N^{s\prime}} \leq K \quad &\Longleftrightarrow \quad \sup_{\substack{g\in\mathcal{G}_N^s\\ \|g\|_{\mathcal{G}_N^s}\leq 1}} \|E_{n,N}(g)\| \leq K \\
&\Longleftrightarrow \quad \sup_{\substack{g\in\mathcal{G}_N^s\\ \|g\|_{\mathcal{G}_N^s}\leq 1}} \left\|\sqrt{n}\int_{-\pi}^{\pi} g(\omega)\left(I_{n,N}(\omega) - \frac{1}{2\pi}\Sigma_N^{(\Delta)}\right) d\omega\right\| \leq K \\
&\Longleftrightarrow \quad \sup_{\substack{g\in\mathcal{G}_N^s\\ \|g\|_{\mathcal{G}_N^s}\leq 1}} \left\|\sqrt{n}\sum_{h=1-n}^{n-1} \widehat{g}_h \left(\overline{\Gamma}_{n,N}(h) - \mathbb{E}[\overline{\Gamma}_{n,N}(h)]\right)\right\| \leq K
\end{aligned}$$

hold. An application of the Cauchy-Schwarz inequality yields for g with $\|g\|_{\mathcal{G}_N^s} \leq 1$

$$\begin{aligned}
&\left\|\sqrt{n}\sum_{h=1-n}^{n-1} \widehat{g}_h \left(\overline{\Gamma}_{n,N}(h) - \mathbb{E}[\overline{\Gamma}_{n,N}(h)]\right)\right\| \\
&\leq \sqrt{n}\mathfrak{C} \sum_{h=1-n}^{n-1} \|\widehat{g}_h\| (1+|h|)^s \left\|\overline{\Gamma}_{n,N}(h) - \mathbb{E}[\overline{\Gamma}_{n,N}(h)]\right\| (1+|h|)^{-s} \\
&\leq \mathfrak{C}\sqrt{n\sum_{h=1-n}^{n-1} \left\|\overline{\Gamma}_{n,N}(h) - \mathbb{E}[\overline{\Gamma}_{n,N}(h)]\right\|^2 (1+|h|)^{-2s}},
\end{aligned}$$

so that by Tschebycheffs inequality and $\sum_{h\in\mathbb{Z}}(1+|h|)^{-2s} \leq \mathfrak{C}$,

$$\mathbb{P}\left(\sup_{\substack{g\in\mathcal{G}_N^s\\ \|g\|_{\mathcal{G}_N^s}\leq 1}} \|E_{n,N}(g)\| \leq K\right) \geq 1 - \mathbb{P}\left(\sup_{\substack{g\in\mathcal{G}_N^s\\ \|g\|_{\mathcal{G}_N^s}\leq 1}} \|E_{n,N}(g)\| > K\right)$$

$$\geq 1 - \frac{\mathfrak{C} \sup_{h\in\mathbb{Z}} n\mathbb{E}\left[\left\|\overline{\Gamma}_{n,N}(h) - \mathbb{E}\left[\overline{\Gamma}_{n,N}(h)\right]\right\|^2\right]}{K^2},$$

holds. Since

$$\sup_{h\in\mathbb{Z}} n\mathbb{E}\left[\left\|\overline{\Gamma}_{n,N}(h) - \mathbb{E}\left[\overline{\Gamma}_{n,N}(h)\right]\right\|^2\right] \leq \mathfrak{C}$$

see (2.17), we find a K such that (5.6) is satisfied.
For the flat concentration, we have to show that for every $\varepsilon > 0$ and every $\delta > 0$ there exists a finite-dimensional subspace $L \subset \mathcal{G}_N^{s\prime}$ such that

$$\inf_{n\in\mathbb{N}} \mathbb{P}(d_{\mathcal{G}_N^{s\prime}}(E_{n,N}, L) \leq \delta\}) \geq 1 - \varepsilon.$$

By Markov's inequality

$$\mathbb{P}\left(d_{\mathcal{G}_N^{s\prime}}(E_{n,N}, L) \leq \delta\}\right) \geq 1 - \frac{\mathbb{E}[d_{\mathcal{G}_N^{s\prime}}(E_{n,N}, L)]}{\delta}.$$

Since

$$\mathbb{E}[d_{\mathcal{G}_N^{s\prime}}(E_{n,N}, L)] = \mathbb{E}\left[\inf_{P\in L} \|E_{n,N} - P\|_{\mathcal{G}_N^{s\prime}}\right] = \mathbb{E}\left\|\mathrm{Pr}_{L^\perp}(E_{n,N})\right\|_{\mathcal{G}_N^{S\prime}},$$

where $\mathrm{Pr}_{L^\perp}$ is the orthogonal projection onto $L^\perp$, we prove that $\mathbb{E}\left\|\mathrm{Pr}_{L^\perp}(E_{n,N})\right\|_{\mathcal{G}_N^{S\prime}}$ can be made arbitrary small by choosing L appropriately. Therefore, we choose a sequence $L_k \subset \mathcal{G}_N^{s\prime}$ with

$$\mathbb{E}\left\|\mathrm{Pr}_{L_k^\perp}(E_{n,N})\right\|_{\mathcal{G}_N^{s\prime}} = \mathbb{E}\left[\sup_{\substack{g\in\mathcal{G}_N^s \\ \|g\|_{\mathcal{G}_N^s}\leq 1}} \left\|\mathrm{Pr}_{L_k^\perp}(E_{n,N})(g)\right\|\right] \overset{k\to\infty}{\longrightarrow} 0. \tag{5.7}$$

For $k \in \mathbb{N}$ we define L_k as the linear subspace generated by $(e_\ell)_{|\ell|\leq k}$ where e_ℓ is defined by $e_\ell(g) = \widehat{g}_\ell$. For (5.7), it is sufficient to show that

$$\mathbb{E}\left[\sup_{\substack{g\in\mathcal{G}_N^s,\ \|g\|_{\mathcal{G}_N^s}\leq 1 \\ \widehat{g}_j=0, |j|\leq k}} \|E_{n,N}(g)\|^2\right] \overset{k\to\infty}{\longrightarrow} 0.$$

We use representation (5.3) and the ideas above and obtain

$$\begin{aligned}
\sup_{\substack{g\in\mathcal{G}_N^s,\ \|g\|_{\mathcal{G}_N^s}\leq 1 \\ \widehat{g}_j=0, |j|\leq k}} \|E_{N,n}(g)\|^2 &= \sup_{\substack{g\in\mathcal{G}_N^s,\ \|g\|_{\mathcal{G}_N^s}\leq 1 \\ \widehat{g}_j=0, |j|\leq k}} \left\|\sqrt{n} \sum_{n>|h|>k} \widehat{g}_h \left(\overline{\Gamma}_{n,N}(h) - \mathbb{E}[\overline{\Gamma}_{n,N}(h)]\right)\right\|^2 \\
&\leq n\mathfrak{C} \sum_{n>|h|>k} \left\|\overline{\Gamma}_{n,N}(h) - \mathbb{E}[\overline{\Gamma}_{n,N}(h)]\right\|^2 (1 + |h|)^{-2s}.
\end{aligned}$$

Property (2.17) and $\sum_{|h|>k}(1+|h|)^{-2s} \overset{k\to\infty}{\longrightarrow} 0$ yield the assertion. $\square$

Lemma 5.12.
Let Assumption (N2) hold. Then, $(E_{n,N})_{n\in\mathbb{N}}$ is tight in $(\mathcal{C}(\mathcal{G}_N), \|\cdot\|_{N,\infty})$.

Proof. By representation (5.3), we have to prove that for every $\varepsilon > 0$ there exists some constant $K > 0$ such that

$$\mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sqrt{n}\sum_{h=1-n}^{n-1}\widehat{g}_h\left(\overline{\Gamma}_{n,N}(h) - \mathbb{E}[\overline{\Gamma}_{n,N}(h)]\right)\right\| > K\right) \le \varepsilon \quad \forall n\in\mathbb{N}.$$

We prove some stronger condition. Namely, we show that for every $\varepsilon, \delta > 0$ there exist M, $\varepsilon_1, \varepsilon_2 > 0$ with $\varepsilon_1 + \varepsilon_2 \le \varepsilon$ and $K > 0$ such that

$$\mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sqrt{n}\sum_{h=-M+1}^{M-1}\widehat{g}_h\left(\overline{\Gamma}_{n,N}(h) - \mathbb{E}[\overline{\Gamma}_{n,N}(h)]\right)\right\| > K\right) \le \varepsilon_1$$
$$\text{and} \qquad \mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sqrt{n}\sum_{M\le|h|\le n-1}\widehat{g}_h\left(\overline{\Gamma}_{n,N}(h) - \mathbb{E}[\overline{\Gamma}_{n,N}(h)]\right)\right\| > \delta\right) < \varepsilon_2 \quad \forall n\in\mathbb{N}.$$

Note that for fixed $M\in\mathbb{N}$, we have by the Cauchy-Schwarz inequality

$$\begin{aligned}&\sup_{g\in\mathcal{G}_N}\left\|\sqrt{n}\sum_{h=-M+1}^{M-1}\widehat{g}_h\left(\overline{\Gamma}_{n,N}(h) - \mathbb{E}[\Gamma_{n,N}(h)]\right)\right\|^2\\ \le\;& \sup_{g\in\mathcal{G}_N}\left(\sum_{h=-M+1}^{M-1}\|\widehat{g}_h\|^2\right)\left(\sum_{h=-M+1}^{M-1}\left\|\sqrt{n}\left(\Gamma_{n,N}(h) - \mathbb{E}[\Gamma_{n,N}(h)]\right)\right\|^2\right).\end{aligned}$$

Since Lemma 2.16 implies the tightness of $\left(\sum_{h=-M+1}^{M-1}\left\|\sqrt{n}(\overline{\Gamma}_{n,N}(h) - \mathbb{E}[\Gamma_{n,N}(h)])\right\|^2\right)_{n\in\mathbb{N}}$, we obtain for $\varepsilon > 0$

$$\mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sqrt{n}\sum_{h=-M+1}^{M-1}\widehat{g}_h\left(\overline{\Gamma}_{n,N}(h) - \mathbb{E}[\overline{\Gamma}_{n,N}]\right)\right\| \le K\right) \ge 1-\varepsilon \quad \forall n\in\mathbb{N}$$

for fixed M and some $K > 0$. Therefore, it suffices to show

$$\lim_{M\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sum_{n-1\ge|h|\ge M}\widehat{g}_h\overline{\Gamma}_{n,N}(h)\right\| > \delta\right) = 0 \qquad \forall\delta > 0 \tag{5.8}$$

to prove the tightness of $(E_{n,N})_{n\in\mathbb{N}}$. Since

$$\mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sum_{n-1\ge|h|\ge M}\widehat{g}_h\overline{\Gamma}_{n,N}(h)\right\| > \varepsilon\right) \le \mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sum_{h=M}^{n-1}\widehat{g}_h\overline{\Gamma}_{n,N}(h)\right\| > \frac{\varepsilon}{2}\right)$$

$$+ \mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sum_{h=-n+1}^{-M}\widehat{g}_h\overline{\Gamma}_{n,N}(h)\right\| > \frac{\varepsilon}{2}\right),$$

we prove

$$\lim_{M\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sum_{h=M}^{n-1}\widehat{g}_h\overline{\Gamma}_{n,N}(h)\right\| > \varepsilon\right) = 0 \qquad \forall\varepsilon > 0.$$

The convergence

$$\lim_{M\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sum_{h=-n+1}^{-M}\widehat{g}_h\overline{\Gamma}_{n,N}(h)\right\| > \varepsilon\right) = 0 \qquad \forall\varepsilon > 0$$

can be shown in the same way. As in the proof of Theorem 3.7 of Can et al. (2010), we assume that n and M are of the forms $M = 2^a$, $n = 2^{b+1}$ for some $a < b$ to ease the notation. We have

$$\mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sum_{h=M}^{n-1}\widehat{g}_h\overline{\Gamma}_{n,N}(h)\right\| > \varepsilon\right) \le \sum_{k=a}^{b}\mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sum_{h=2^k}^{2^{k+1}-1}\widehat{g}_h\overline{\Gamma}_{n,N}(h)\right\| > \varepsilon_k\right) \tag{5.9}$$

where $\varepsilon_k = 2^{-\theta k}$ for some $\theta > 0$ with $\varepsilon/(b-a+1) > 2^{-\theta}$. For $k \ge 0$, let $(\varepsilon_{k,\ell})_{k\in\mathbb{N},\ell\ge 0}$ be the array of positive numbers defined by $\varepsilon_{k,\ell} = 2^{-\gamma_1\ell-\gamma_2 k}$ for $\gamma_1 > \frac{2}{2-\beta}$ and $1+\gamma_2 > \frac{1+2\theta}{2-\beta}$. Then, for $k \in \mathbb{N}_0$, let $u := u_{k,0} := N(\varepsilon_{k,0}, \mathcal{G}_N, d_N^{(k)})$ and $g^{(1)}, \ldots, g^{(u)} \in \mathcal{G}_N$ with

$$\sup_{g\in\mathcal{G}_N}\min_{j=1,\ldots,u} d_N^{(k)}(g, g^{(j)}) < \varepsilon_{k,0}.$$

We have

$$\begin{aligned}
&\mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sum_{h=2^k}^{2^{k+1}-1}\widehat{g}_h\overline{\Gamma}_{n,N}(h)\right\| > \varepsilon_k\right)\\
&\le \mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\min_{j=1,\ldots,u}\left\|\sum_{h=2^k}^{2^{k+1}-1}\left(\widehat{g}_h - \widehat{g^{(j)}}_h\right)\overline{\Gamma}_{n,N}(h)\right\| > \frac{\varepsilon_k}{2}\right)\\
&\quad + \mathbb{P}\left(\max_{j=1,\ldots,u}\left\|\sum_{h=2^k}^{2^{k+1}-1}\widehat{g^{(j)}}_h\overline{\Gamma}_{n,N}(h)\right\| > \frac{\varepsilon_k}{2}\right)\\
&\le \mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\min_{j=1,\ldots,u}\left\|\sum_{h=2^k}^{2^{k+1}-1}\left(\widehat{g}_h - \widehat{g^{(j)}}_h\right)\overline{\Gamma}_{n,N}(h)\right\| > \frac{\varepsilon_k}{2}\right)\\
&\quad + \sum_{j=1}^{u}\mathbb{P}\left(\left\|\sum_{h=2^k}^{2^{k+1}-1}\widehat{g^{(j)}}_h\overline{\Gamma}_{n,N}(h)\right\| > \frac{\varepsilon_k}{2}\right)\\
&\le \mathbb{P}\left(\sup_{\substack{g\in\mathcal{G}_N\\ d_N^{(k)}(f,g)\le\varepsilon_{k,0}}}\left\|\sum_{h=2^k}^{2^{k+1}-1}\left(\widehat{g}_h - \widehat{f}_h\right)\overline{\Gamma}_{n,N}(h)\right\| > \frac{\varepsilon_k}{2}\right)
\end{aligned}$$

$$+u \sup_{g\in\mathcal{G}_N} \mathbb{P}\left(\left\|\sum_{h=2^k}^{2^{k+1}-1} \widehat{g}_h \overline{\Gamma}_{n,N}(h)\right\| > \frac{\varepsilon_k}{2}\right).$$

By induction and since for fixed $n \in \mathbb{N}$

$$\begin{aligned}
&\mathbb{P}\left(\sup_{\substack{g\in\mathcal{G}_N\\ d_N^{(k)}(f,g)\le\varepsilon_{k,\ell}}} \left\|\sum_{h=2^k}^{2^{k+1}-1} \left(\widehat{g}_h - \widehat{f}_h\right) \overline{\Gamma}_{n,N}(h)\right\| > \frac{\varepsilon_k}{2^{\ell+1}}\right)\\
&\le \mathbb{P}\left(\sup_{\substack{g\in\mathcal{G}_N\\ d_N^{(k)}(f,g)\le\varepsilon_{k,\ell}}} \sum_{h=2^k}^{2^{k+1}-1} \left\|\overline{\Gamma}_{n,N}(h)\right\| \left\|\widehat{g}_h - \widehat{f}_h\right\| > \frac{\varepsilon_k}{2^{\ell+1}}\right)\\
&\le \mathbb{E}\left[\frac{1}{2^k}\sum_{h=2^k}^{2^{k+1}-1} \left\|\overline{\Gamma}_{n,N}(h)\right\|\right] \frac{\varepsilon_{k,\ell}}{\varepsilon_k} 2^{\ell+1} \overset{\ell\to\infty}{\longrightarrow} 0,
\end{aligned}$$

we obtain

$$\begin{aligned}
&\mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sum_{h=2^k}^{2^{k+1}-1} \widehat{g}_h \overline{\Gamma}_{n,N}(h)\right\| > \varepsilon_k\right)\\
&\le N(\varepsilon_{k,0}, \mathcal{G}_N, d_N^{(k)}) \sup_{g\in\mathcal{G}_N} \mathbb{P}\left(\left\|\sum_{h=2^k}^{2^{k+1}-1} \widehat{g}_h \overline{\Gamma}_{n,N}(h)\right\| > \frac{\varepsilon_k}{2}\right)\\
&+ \sum_{\ell=1}^{\infty} N(\varepsilon_{k,\ell}, \mathcal{G}_N, d_N^{(k)}) \sup_{\substack{f,g\in\mathcal{G}_N\\ d_N^{(k)}(f,g)\le\varepsilon_{k,\ell-1}}} \mathbb{P}\left(\left\|\sum_{h=2^k}^{2^{k+1}-1} \left(\widehat{f}_h - \widehat{g}_h\right) \overline{\Gamma}_{n,N}(h)\right\| > \frac{\varepsilon_k}{2^{\ell+1}}\right) \qquad (5.10)
\end{aligned}$$

The first term on the right hand-side can be treated similar to (5.4). For the sum, note that Tschbeycheffs inequality and the independence of $(N_j^{(\Delta)})_{j\in\mathbb{Z}}$ yield for $k \in \mathbb{N}_0$

$$\begin{aligned}
&\sum_{\ell=1}^{\infty} N(\varepsilon_{k,\ell}, \mathcal{G}_N, d_N^{(k)}) \sup_{\substack{f,g\in\mathcal{G}_N\\ d_N^{(k)}(f,g)\le\varepsilon_{k,\ell-1}}} \mathbb{P}\left(\left\|\sum_{h=2^k}^{2^{k+1}-1} \left(\widehat{f}_h - \widehat{g}_h\right) \overline{\Gamma}_{n,N}(h)\right\| > \frac{\varepsilon_k}{2^{\ell+1}}\right)\\
&\le \sum_{\ell=1}^{\infty} N(\varepsilon_{k,\ell}, \mathcal{G}_N, d_N^{(k)}) \sup_{\substack{f,g\in\mathcal{G}_N\\ d_N^{(k)}(f,g)\le\varepsilon_{k,\ell-1}}} \frac{2^{2\ell+2}}{\varepsilon_k^2} \sum_{h=2^k}^{2^{k+1}-1} \mathbb{E}\left[\left\|\overline{\Gamma}_{n,N}(h)\right\|^2\right] \left\|\widehat{f}_h - \widehat{g}_h\right\|^2\\
&\le \mathfrak{C}\sum_{\ell=1}^{\infty} N(\varepsilon_{k,\ell}, \mathcal{G}_N, d_N^{(k)}) \frac{2^{2\ell+2}}{\varepsilon_k^2} \sum_{h=2^k}^{2^{k+1}-1} \frac{\varepsilon_{k,\ell-1}^2}{h^2}\\
&\le \mathfrak{C}\sum_{\ell=1}^{\infty} N(\varepsilon_{k,\ell}, \mathcal{G}_N, d_N^{(k)}) \frac{2^{2\ell+2}}{\varepsilon_k^2} \frac{1}{2^k} \varepsilon_{k,\ell-1}^2. \qquad (5.11)
\end{aligned}$$

Then, (5.9)-(5.11), Assumption $(N2)$ and $\varepsilon_k = 2^{-\Theta k}$ give

$$
\begin{aligned}
&\mathbb{P}\left(\sup_{g\in\mathcal{G}_N}\left\|\sum_{h=M}^{n-1}\widehat{g}_h\overline{\Gamma}_{n,N}(h)\right\|>\varepsilon\right)\\
\leq\ & \sum_{k=a}^{b}\sum_{\ell=1}^{\infty}N(\varepsilon_{k,\ell},\mathcal{G}_N,d_N^{(k)})\sup_{\substack{f,g\in\mathcal{G}_N\\ d_N^{(k)}(f,g)\leq\varepsilon_{k,\ell-1}}}\mathbb{P}\left(\left\|\sum_{h=2^k}^{2^{k+1}-1}\left(\widehat{f}_h-\widehat{g}_h\right)\overline{\Gamma}_{n,N}(h)\right\|>\frac{\varepsilon_k}{2^{\ell+1}}\right)\\
\leq\ & \mathfrak{C}\sum_{k=a}^{b}\sum_{\ell=1}^{\infty}N(\varepsilon_{k,\ell},\mathcal{G}_N,d_N^{(k)})\frac{2^{2\ell+2}}{\varepsilon_k^2}\frac{1}{2^k}\varepsilon_{k,\ell-1}^2\\
\leq\ & \mathfrak{C}\sum_{k=a}^{b}\sum_{\ell=1}^{\infty}\left(1+\frac{2^{k\beta}}{\varepsilon_{k,\ell}^{\beta}}\right)\frac{2^{2\ell+2}}{\varepsilon_k^2}\frac{1}{2^k}\varepsilon_{k,\ell-1}^2\\
\leq\ & \mathfrak{C}\left(\sum_{k=a}^{b}2^{-k+2\Theta k}\sum_{\ell=1}^{\infty}2^{2\ell+2}\varepsilon_{k,\ell-1}^2+\sum_{k=a}^{b}2^{-k+2\Theta k+k\beta}\sum_{\ell=1}^{\infty}2^{2\ell+2}\frac{\varepsilon_{k,\ell-1}^2}{\varepsilon_{k,\ell}^{\beta}}\right).
\end{aligned}
$$

Taking the limits yields

$$
\begin{aligned}
&\lim_{a\to\infty}\lim_{b\to\infty}\left(\sum_{k=a}^{b}2^{-k+2\Theta k}\sum_{\ell=1}^{\infty}2^{2\ell+2}\varepsilon_{k,\ell-1}^2+\sum_{k=a}^{b}2^{-k+2\Theta k+k\beta}\sum_{\ell=1}^{\infty}2^{2\ell+2}\frac{\varepsilon_{k,\ell-1}^2}{\varepsilon_{k,\ell}^{\beta}}\right)\\
=&\lim_{a\to\infty}\lim_{b\to\infty}\left(\sum_{k=a}^{b}2^{-k+2\Theta k}\sum_{\ell=1}^{\infty}2^{2\ell+2}2^{-2\gamma_1(\ell-1)}2^{-2\gamma_2 k}\right.\\
&\qquad\left.+\sum_{k=a}^{b}2^{-k+2\Theta k+k\beta}\sum_{\ell=1}^{\infty}2^{2\ell+2}\frac{2^{-2\gamma_1(\ell-1)}2^{-2\gamma_2 k}}{2^{-\beta\gamma_1\ell}2^{-\beta\gamma_2 k}}\right)\\
=&\lim_{a\to\infty}\lim_{b\to\infty}\left(\sum_{k=a}^{b}2^{-k+2\Theta k-2\gamma_2 k}\sum_{\ell=1}^{\infty}2^{2\ell+2-2\gamma_1(\ell-1)}\right.\\
&\qquad\left.+\sum_{k=a}^{b}2^{-k+2\Theta k+k\beta-2\gamma_2 k+\beta\gamma_2 k}\sum_{\ell=1}^{\infty}2^{2\ell+2}\frac{2^{-2\gamma_1(\ell-1)}}{2^{-\beta\gamma_1\ell}}\right)\\
=&0,
\end{aligned}
$$

which completes the proof. □

Lemma 5.13.
Let Assumption $(N3)$ hold. Then, $(E_{n,N})_{n\in\mathbb{N}}$ is tight in $(\mathcal{C}(\mathcal{G}_N),\|\cdot\|_{N,\infty})$.

Proof. The assertion follows from Theorem 2.5 of Dahlhaus (1988). Note that by

$$
\begin{aligned}
&\int_0^1\log(N(\varepsilon,\mathcal{G}_m,d_m)^2/\varepsilon)^2d\varepsilon<\infty\\
\iff&\int_0^1 4\log(N(\varepsilon,\mathcal{G}_m,d_m))^2-4\log(\varepsilon)\log(N(\varepsilon,\mathcal{G}_m,d_m))+\log(\varepsilon)^2d\varepsilon<\infty\\
\iff&\int_0^1\log(N(\varepsilon,\mathcal{G}_m,d_m))^2d\varepsilon<\infty
\end{aligned}
$$

our entropy condition is equivalent to the entropy condition therein. Thereby, the second equivalence follows from $\int_0^1 \log(\varepsilon)^2 d\varepsilon = 2$ and the Cauchy-Schwarz inequality. Furthermore, Lemma 2.2 of Dahlhaus (1988) can be proven similar to Lemma 5.7 (ii) of Dahlhaus and Polonik (2009). Therefore, the cumulant spectra assumption of Dahlhaus (1988) can be exchanged by $\mathbb{E}[\|N_1^{(\Delta)}\|^k] < \widetilde{K}^k$ where $\widetilde{K} > 0$. Due to $\mathbb{E}[\|L_1\|^k] \leq K^k$, this follows in the same way as in Lemma 3.15 of Schlemm and Stelzer (2012b). □

Theorem 5.8 now follows from Lemma 5.9, Lemma 5.11, Lemma 5.12 and Lemma 5.13.

5.3.2. Proof of Theorem 5.3

To deduce Theorem 5.3 from Theorem 5.8, we have to check that the error which is made by approximating the periodogram of the sampled process by the periodogram of the sampled white noise is sufficiently small. For $g \in \mathcal{G}_m$, we define

$$E_{n,R}(g) := \sqrt{n} \int_{-\pi}^{\pi} g(\omega) \left(I_{n,Y}(\omega) - \Phi(e^{-i\omega}) I_{n,N}(\omega) \Phi(e^{i\omega})^\top \right) d\omega.$$

In the following, we can interchange summation, integration and taking the expectation since all the necessary integrals and expectations exist. Therefore, we do so without further references.

Lemma 5.14.
Let the assumptions of Theorem 5.3 hold. Then, under $(C1)$, $E_{n,R} \xrightarrow{\mathbb{P}} 0$ in $(\mathcal{G}_m^{s\prime}, \|\cdot\|_{\mathcal{G}_m^{s\prime}})$. Under $(C2)$ or $(C3)$, $E_{n,R} \xrightarrow{\mathbb{P}} 0$ in $(\mathcal{C}(\mathcal{G}_m), \|\cdot\|_{m,\infty})$ holds.

Proof. Define $R_n(\omega) = I_{n,Y}(\omega) - \Phi(e^{-i\omega}) I_{n,N}(\omega) \Phi(e^{i\omega})^\top$ for $\omega \in [-\pi, \pi]$. We get

$$\begin{aligned}
R_n(\omega) &= \frac{1}{2\pi n} \left(\sum_{k=1}^{n} \sum_{r=0}^{\infty} \Phi_r N_{k-r}^{(\Delta)} \right) \left(\sum_{\ell=1}^{n} \sum_{t=0}^{\infty} \Phi_t N_{\ell-t}^{(\Delta)} \right)^\top e^{-i(k-\ell)\omega} \\
&\quad - \frac{1}{2\pi n} \left(\sum_{k=1}^{n} \sum_{r=0}^{\infty} \Phi_r N_{k}^{(\Delta)} \right) \left(\sum_{\ell=1}^{n} \sum_{t=0}^{\infty} \Phi_t N_{\ell}^{(\Delta)} \right)^\top e^{-i(k+r-\ell-t)\omega_j} \\
&= \frac{1}{2\pi n} \left(\sum_{r=0}^{\infty} \sum_{t=0}^{\infty} \Phi_r \left(\sum_{k=1-r}^{0} \sum_{\ell=1-t}^{0} - \sum_{k=1}^{n} \sum_{\ell=n-t+1}^{n} + \sum_{k=1-r}^{0} \sum_{\ell=1}^{n} - \sum_{k=1-r}^{0} \sum_{\ell=n-t+1}^{n} \right.\right. \\
&\quad \left. + \sum_{k=n-r+1}^{n} \sum_{\ell=n-t+1}^{n} - \sum_{k=n-r+1}^{n} \sum_{\ell=1}^{n} + \sum_{k=1}^{n} \sum_{\ell=1-t}^{0} - \sum_{k=n-r+1}^{n} \sum_{\ell=1-t}^{0} \right) \\
&\quad \left. N_k^{(\Delta)} N_\ell^{(\Delta)\top} e^{-i(k+r-\ell-t)\omega} \Phi_t^\top \right) \\
&=: \sum_{i=1}^{8} R_n^{(i)}(\omega).
\end{aligned}$$

Thus, we show that

$$\sup_{g \in \mathcal{G}_m} \left\| \sqrt{n} \int_{-\pi}^{\pi} g(\omega) R_n^{(i)}(\omega) d\omega \right\| \xrightarrow{\mathbb{P}} 0 \quad \text{and}$$

$$\sup_{\substack{g\in\mathcal{G}_m^s\\ \|g\|_{\mathcal{G}_m^s}\le 1}} \left\| \sqrt{n}\int_{-\pi}^{\pi} g(\omega) R_n^{(i)}(\omega)d\omega \right\| \xrightarrow{\mathbb{P}} 0, \quad i=1,\ldots,8,$$

respectively, hold depending on the imposed conditions. By symmetry, the proofs for $i=6,7,8$ are the same as those for $i=2,3,4$ respectively. Note that the proofs for $i=4,5$ are based on the same ideas as the one for $i=1$ and the case $i=3$ goes very similar to the case of $i=2$. Consequently, we only investigate the terms corresponding to $i=1,2$.
As a first step we consider the case $\underline{i=1}$ which can be shown for either set of conditions in the same way:

$$\begin{aligned}
&\sup_{g\in\mathcal{G}_m} \left\| \sqrt{n}\int_{-\pi}^{\pi} g(\omega) R_n^{(1)}(\omega)d\omega \right\| \\
=\ & \sup_{g\in\mathcal{G}_m} \left\| \frac{1}{\sqrt{n}}\int_{-\pi}^{\pi} \frac{1}{2\pi}\sum_{r=0}^{\infty}\sum_{t=0}^{\infty}\sum_{k=1-r}^{0}\sum_{\ell=1-t}^{0}\sum_{u=-\infty}^{\infty} \widehat{g}_u \Phi_r N_k^{(\Delta)} N_\ell^{(\Delta)\top}\Phi_t^\top e^{-i\omega(k+r-\ell-t-u)}d\omega \right\| \\
=\ & \sup_{g\in\mathcal{G}_m} \left\| \frac{1}{\sqrt{n}}\sum_{r=0}^{\infty}\sum_{t=0}^{\infty}\sum_{\ell=1-t}^{0}\sum_{u=-\ell-t+1}^{-\ell-t+r} \widehat{g}_u \Phi_r N_{\ell+t+u-r}^{(\Delta)} N_\ell^{(\Delta)\top}\Phi_t^\top \right\| \\
=\ & \sup_{g\in\mathcal{G}_m} \left\| \frac{1}{\sqrt{n}}\sum_{r=0}^{\infty}\sum_{t=0}^{\infty}\sum_{\ell=1-t}^{0}\sum_{u=1}^{r} \widehat{g}_{u-\ell-t} \Phi_r N_{u-r}^{(\Delta)} N_\ell^{(\Delta)\top}\Phi_t^\top \right\| \\
\le\ & \sup_{g\in\mathcal{G}_m} \frac{1}{\sqrt{n}}\sum_{r=0}^{\infty}\sum_{t=0}^{\infty}\sum_{\ell=1-t}^{0}\sum_{u=1}^{r} \|\Phi_r\|\|N_{u-r}^{(\Delta)}\|\|N_\ell^{(\Delta)}\|\|\Phi_t\|\|\widehat{g}_{u-\ell-t}\| \\
\le\ & \sup_{g\in\mathcal{G}_m} \frac{1}{\sqrt{n}}\sum_{r=0}^{\infty}\sum_{t=0}^{\infty}\sum_{u=1}^{r} \|\Phi_r\|\|\Phi_t\|\|N_{u-r}^{(\Delta)}\| \sqrt{\sum_{\ell_1=1-t}^{0}\|N_{\ell_1}^{(\Delta)}\|^2 \sum_{\ell_2=1-t}^{0}\|\widehat{g}_{u-\ell_2-t}\|^2} \\
\le\ & \frac{\mathfrak{C}}{\sqrt{n}}\sum_{r=0}^{\infty}\|\Phi_r\|(r+1)\sum_{t=0}^{\infty}\sqrt{t+1}\|\Phi_t\| \left(\frac{1}{1+r}\sum_{u=1-r}^{0}\|N_u^{(\Delta)}\|\right)\sqrt{\frac{1}{t+1}\sum_{\ell_1=1-t}^{0}\|N_{\ell_1}^{(\Delta)}\|^2} \\
\xrightarrow{\mathbb{P}}\ & 0, \text{ as } n\to\infty.
\end{aligned}$$

Thereby, we used the assumption $\sup_{g\in\mathcal{G}_m}\|g\|_m^2<\infty$ along with $\sum_{r=0}^{\infty}(1+r)\|\Phi_r\|<\infty$, see (2.10), and the strong law of large numbers, in particular

$$\frac{1}{1+r}\sum_{u=1-r}^{0}\|N_u^{(\Delta)}\| \overset{a.s.}{\to} \mathbb{E}\|N_1^{(\Delta)}\|.$$

Next, we investigate $\underline{i=2}$.
Case 1: Under $(C1)$, we again only consider appropriately normalized functions, i.e., we prove $\|\sqrt{n}\int_{-\pi}^{\pi}\cdot R_n^{(2)}(\omega)d\omega\|_{\mathcal{G}_m^{s\prime}} \xrightarrow{\mathbb{P}} 0$. The Cauchy-Schwarz inequality yields

$$\|\sqrt{n}\int_{-\pi}^{\pi}\cdot R_n^{(2)}(\omega)d\omega\|_{\mathcal{G}_m^{s\prime}}$$

$$
\begin{aligned}
&= \sup_{\substack{g\in\mathcal{G}_m^s\\ \|g\|_{\mathcal{G}_m^s}\le 1}} \left\| \sqrt{n}\int_{-\pi}^{\pi} g(\omega) R_n^{(2)}(\omega) d\omega \right\| \\
&= \sup_{\substack{g\in\mathcal{G}_m^s\\ \|g\|_{\mathcal{G}_m^s}\le 1}} \left\| \frac{1}{\sqrt{n}} \int_{-\pi}^{\pi} \frac{1}{2\pi} \sum_{r=0}^{\infty}\sum_{t=0}^{\infty}\sum_{k=1}^{n}\sum_{\ell=n+1-t}^{n}\sum_{u=-\infty}^{\infty} \widehat{g}_u \Phi_r N_k^{(\Delta)} N_\ell^{(\Delta)\top} \Phi_t^\top e^{-i\omega(k+r-\ell-t-u)} d\omega \right\| \\
&= \sup_{\substack{g\in\mathcal{G}_m^s\\ \|g\|_{\mathcal{G}_m^s}\le 1}} \left\| \frac{1}{\sqrt{n}} \sum_{t=0}^{\infty}\sum_{\ell=n+1-t}^{n}\sum_{r=0}^{\infty}\sum_{u=1+r-t-\ell}^{n+r-t-\ell} \widehat{g}_u \Phi_r N_{\ell+t+u-r}^{(\Delta)} N_\ell^{(\Delta)\top} \Phi_t^\top \right\| \\
&= \sup_{\substack{g\in\mathcal{G}_m^s\\ \|g\|_{\mathcal{G}_m^s}\le 1}} \left\| \frac{1}{\sqrt{n}} \sum_{t=0}^{\infty}\sum_{\ell=n+1-t}^{n}\sum_{u=1-\ell-t}^{n-\ell-t}\sum_{r=0}^{\infty} \widehat{g}_{u+r} \Phi_r N_{\ell+u+t}^{(\Delta)} N_\ell^{(\Delta)\top} \Phi_t^\top \right\| \\
&\le \sup_{\substack{g\in\mathcal{G}_m^s\\ \|g\|_{\mathcal{G}_m^s}\le 1}} \frac{1}{\sqrt{n}} \sum_{t=0}^{\infty}\sum_{\ell=n+1-t}^{n}\sum_{u=1-\ell-t}^{n-\ell-t} \left\| \sum_{r=0}^{\infty} \widehat{g}_{u+r} \Phi_r N_{\ell+u+t}^{(\Delta)} N_\ell^{(\Delta)\top} \Phi_t^\top \right\| \\
&\le \sup_{\substack{g\in\mathcal{G}_m^s\\ \|g\|_{\mathcal{G}_m^s}\le 1}} \sum_{t=0}^{\infty}\sum_{\ell=n+1-t}^{n} \frac{1}{\sqrt{n}} \left(\sum_{u_1=1-\ell-t}^{n-\ell-t} (1+|u_1|)^{2s} \left\| \sum_{r=0}^{\infty} \widehat{g}_{u_1+r}\Phi_r \right\|^2 \right)^{1/2} \\
&\qquad \left(\sum_{u_2=1-\ell-t}^{n-\ell-t} (1+|u_2|)^{-2s} \left\| N_{\ell+u_2+t}^{(\Delta)} N_\ell^{(\Delta)\top} \Phi_t^\top \right\|^2 \right)^{1/2} \\
&\le \sup_{\substack{g\in\mathcal{G}_m^s\\ \|g\|_{\mathcal{G}_m^s}\le 1}} \sum_{t=0}^{\infty}\sum_{\ell=n+1-t}^{n} \frac{1}{\sqrt{n}} \left(\sum_{u_1=-\infty}^{\infty} (1+|u_1|)^{2s} \left\| \sum_{r=0}^{\infty} \widehat{g}_{u_1+r}\Phi_r \right\|^2 \right)^{1/2} \\
&\qquad \left(\sum_{u_2=1-\ell-t}^{n-\ell-t} (1+|u_2|)^{-2s} \left\| N_{\ell+u_2+t}^{(\Delta)} N_\ell^{(\Delta)\top} \Phi_t^\top \right\|^2 \right)^{1/2} \\
&\le \sum_{t=0}^{\infty}\sum_{\ell=n+1-t}^{n} \left(\frac{1}{n} \sum_{u_2=1-\ell-t}^{n-\ell-t} (1+|u_2|)^{-2s} \left\| N_{\ell+u_2+t}^{(\Delta)} N_\ell^{(\Delta)\top} \Phi_t^\top \right\|^2 \right)^{1/2} \\
&\le \frac{1}{\sqrt{n}} \max_{k=1,\dots,n} \|N_k^{(\Delta)}\| \sum_{t=0}^{\infty} \|\Phi_t\| \sum_{\ell=1-t}^{0} \|N_\ell^{(\Delta)}\| \left(\sum_{u_2=1-\ell-t}^{n-\ell-t} (1+|u_2|)^{-2s} \right)^{1/2},
\end{aligned}
$$

where we used that

$$
\|g\|_{\mathcal{G}_m^s}^2 = \sum_{u\in\mathbb{Z}} (1+|u|)^{2s} \left\| \sum_{t=0}^{\infty} \widehat{g}_{u+t}\Phi_t \right\|^2 \le 1.
$$

Therefore, since $\sum_{u_2=-\infty}^{\infty}(1+|u_2|)^{-2s} < \infty$, $\sum_{t=0}^{\infty} \|\Phi_t\||t| < \infty$, $\frac{1}{t}\sum_{\ell=n+1-t}^{n} \|N_\ell^{(\Delta)}\| = O_{\mathbb{P}}(1)$ and

$$
\mathbb{P}\left(\frac{1}{\sqrt{n}} \sup_{k=1,\dots,n} \|N_k^{(\Delta)}\| \ge \varepsilon \right) \le 1 - (1 - \mathbb{P}(\|N_1^{(\Delta)}\| > \varepsilon\sqrt{n})^n
$$

$$\leq 1 - \left(1 - \frac{\mathbb{E}[\|N_1^{(\Delta)}\|^4]}{\varepsilon^4 n^2}\right)^n \overset{n\to\infty}{\longrightarrow} 0,$$

the second term converges as desired.

Case 2: Under $(C2)$ and $(C3)$, the space $(\mathcal{G}_m, \|\cdot\|_m)$ is totally bounded. Therefore, we can approximate the supremum over the potentially uncountable many functions in $\mathcal{G}_m$ by one of finitely many functions. Namely, let $\delta > 0$. Then, there exists $v \in \mathbb{N}$ and $g_1, \ldots, g_v \in \mathcal{G}_m$ such that $\sup_{g\in\mathcal{G}_m} \min_{j=1,\ldots,v} d_m(g, g_j) < \delta$. Therefore, for fixed $\delta > 0$ and appropriately chosen $g_1, \ldots, g_v$, we can approximate the first error term by

$$\begin{aligned}
&\sup_{g\in\mathcal{G}_m} \left\| \sqrt{n} \int_{-\pi}^{\pi} g(\omega) R_n^{(2)}(\omega) d\omega \right\| \\
=\ & \left\| \frac{\sqrt{n}}{2\pi} \int_{-\pi}^{\pi} g(\omega)\Phi(e^{-i\omega}) \sum_{t=0}^{\infty}\sum_{k=1}^{n}\sum_{\ell=n+1-t}^{n} N_k^{(\Delta)} N_\ell^{(\Delta)\top} \Phi_t^\top e^{-i(k-\ell-t)\omega} d\omega \right\| \\
\leq\ & \sup_{g\in\mathcal{G}_m} \min_{j=1,\ldots,v} \left\| \sqrt{n} \int_{-\pi}^{\pi} (g(\omega) - g_j(\omega)) R_n^{(2)}(\omega) d\omega \right\| + \max_{j=1,\ldots,v} \left\| \sqrt{n} \int_{-\pi}^{\pi} g_j(\omega) R_n^{(2)}(\omega) d\omega \right\| \\
\leq\ & \sup_{g\in\mathcal{G}_m} \min_{j=1,\ldots,v} \sqrt{n} \int_{-\pi}^{\pi} \left\| R_n^{(2)}(\omega) \right\| \|g(\omega) - g_j(\omega)\| \, d\omega + \max_{j=1,\ldots,v} \left\| \sqrt{n} \int_{-\pi}^{\pi} g_j(\omega) R_n^{(2)}(\omega) d\omega \right\| \\
\leq\ & \sqrt{n} \left(\int_{-\pi}^{\pi} \left\| R_n^{(2)}(\omega) \right\|^2 d\omega \right)^{1/2} \sup_{g\in\mathcal{G}_m} \min_{j=1,\ldots,v} \left(\int_{-\pi}^{\pi} \|g(\omega) - g_j(\omega)\|^2 \, d\omega \right)^{1/2} \\
&\quad + \max_{j=1,\ldots,v} \left\| \sqrt{n} \int_{-\pi}^{\pi} g_j(\omega) R_n^{(2)}(\omega) d\omega \right\| \\
\leq\ & \sqrt{\delta} \left(n \int_{-\pi}^{\pi} \left\| R_n^{(2)}(\omega) \right\|^2 d\omega \right)^{1/2} + \max_{j=1,\ldots,v} \left\| \sqrt{n} \int_{-\pi}^{\pi} g_j(\omega) R_n^{(2)}(\omega) d\omega \right\|.
\end{aligned}$$

Since δ can be chosen arbitrary small, we have to prove

$$n \int_{-\pi}^{\pi} \|R_n^{(2)}(\omega)\|^2 d\omega = O_{\mathbb{P}}(1), \quad \left\| \sqrt{n} \int_{-\pi}^{\pi} g(\omega) R_n^{(2)}(\omega) d\omega \right\| = o_{\mathbb{P}}(1) \text{ for any } g \in \mathcal{G}_m. \tag{5.12}$$

On the one hand, we have

$$\begin{aligned}
&\mathbb{E}\left[n \int_{-\pi}^{\pi} \|R_n^{(2)}(\omega)\|^2 d\omega \right] \\
=\ & \sum_{S,T=1}^{N} \mathbb{E}\Bigg[n \int_{-\pi}^{\pi} \frac{1}{4\pi^2 n^2} \sum_{r_1,r_2=0}^{\infty} \sum_{t_1,t_2=0}^{\infty} \sum_{k_1,k_2=1}^{n} \sum_{\ell_1=n+1-t_1}^{0} \sum_{\ell_2=n+1-t_2}^{n} \Phi_{r_1} N_{k_1}^{(\Delta)} N_{\ell_1}^{(\Delta)\top} \\
&\qquad \Phi_{t_1}^\top[S,T] \Phi_{r_2} N_{k_2}^{(\Delta)} N_{\ell_2}^{(\Delta)\top} \Phi_{t_2}^\top[S,T] e^{i\omega(k_1+r_1-\ell_1-t_1-k_2-r_2+\ell_2+t_2)} d\omega \Bigg] \\
=\ & \sum_{S,T=1}^{N} \int_{-\pi}^{\pi} \frac{1}{4\pi^2 n} \sum_{r_1,r_2=0}^{\infty} \sum_{t_1,t_2=0}^{\infty} \sum_{k_1,k_2=1}^{n} \sum_{\ell_1=n+1-t_1}^{n} \sum_{\ell_2=n+1-t_2}^{0} \Phi_{r_1} \mathbb{E}\Big[N_{k_1}^{(\Delta)} N_{\ell_1}^{(\Delta)\top} \\
&\qquad \Phi_{t_1}^\top[S,T] \Phi_{r_2} N_{k_2}^{(\Delta)} N_{\ell_2}^{(\Delta)\top} \Big] \Phi_{t_2}^\top[S,T] e^{i\omega(k_1+r_1-\ell_1-t_1-k_2-r_2+\ell_2+t_2)} d\omega
\end{aligned}$$

$$
\begin{aligned}
= & \sum_{S,T=1}^{N} \int_{-\pi}^{\pi} \frac{1}{4\pi^2 n} \sum_{r_1,r_2=0}^{\infty} \sum_{t_1,t_2=0}^{\infty} \sum_{k=1}^{n} \sum_{\ell=n+1-\min\{t_1,t_2\}}^{n} \Phi_{r_1} \mathbb{E}\Big[N_k^{(\Delta)} N_\ell^{(\Delta)\top} \\
& \qquad \Phi_{t_1}^\top [S,T] \Phi_{r_2} N_k^{(\Delta)} N_\ell^{(\Delta)\top}\Big] \Phi_{t_2}^\top [S,T] e^{i\omega(r_1-t_1-r_2+t_2)} d\omega \\
& + \sum_{S,T=1}^{N} \int_{-\pi}^{\pi} \frac{1}{4\pi^2 n} \sum_{r_1,r_2=0}^{\infty} \sum_{t_1,t_2=0}^{\infty} \sum_{k=\max\{1,n+1-t_1\}}^{n} \sum_{\substack{\ell=\max\{1,n+1-t_2\},\\ \ell\neq k}}^{n} \Phi_{r_1} \mathbb{E}\Big[N_1^{(\Delta)} N_1^{(\Delta)\top}\Big] \\
& \qquad \Phi_{t_1}^\top [S,T] \Phi_{r_2} \mathbb{E}\Big[N_1^{(\Delta)} N_1^{(\Delta)\top}\Big] \Phi_{t_2}^\top [S,T] e^{i\omega(r_1-t_1-r_2+t_2)} d\omega \\
& + \sum_{S,T=1}^{N} \int_{-\pi}^{\pi} \frac{1}{4\pi^2 n} \sum_{r_1,r_2=0}^{\infty} \sum_{t_1,t_2=0}^{\infty} \sum_{k=\max\{1,n+1-t_2\}}^{n} \sum_{\substack{\ell=\max\{1,n+1-t_1\},\\ \ell\neq k}}^{n} \Phi_{r_1} \mathbb{E}\Big[N_k^{(\Delta)} N_\ell^{(\Delta)\top} \\
& \qquad \Phi_{t_1}^\top [S,T] \Phi_{r_2} N_\ell^{(\Delta)} N_k^{(\Delta)\top}\Big] \Phi_{t_2}^\top [S,T] e^{i\omega(2k-2\ell+r_1-t_1-r_2+t_2)} d\omega \\
\leq & \sum_{S,T=1}^{N} \frac{1}{2\pi n} \left(\sum_{r=0}^{\infty} \|\Phi_r\|\right)^2 \sum_{t_1,t_2=0}^{\infty} (t_1+1)\|\Phi_{t_1}\|(t_2+1)\|\Phi_{t_2}\| \\
& \frac{1}{(t_1+1)(t_2+1)} \left(\sum_{\ell=n+1-\min\{t_1,t_2\}}^{n} \sum_{k=1}^{n} + 2 \sum_{\substack{k=\max\{1,n+1-t_1\}\\ \ell=\max\{1,n+1-t_2\},\\ \ell\neq k}}^{n} \right) \mathbb{E}\Big[\|N_k^{(\Delta)}\|^2 \|N_\ell^{(\Delta)}\|^2\Big] \\
\leq & \ \mathfrak{C}.
\end{aligned}
$$

On the other hand, we have

$$
\begin{aligned}
& \left\| \sqrt{n} \int_{-\pi}^{\pi} g(\omega) R_n^{(2)}(\omega) d\omega \right\| \\
& \leq \left\| \frac{1}{2\pi\sqrt{n}} \int_{-\pi}^{\pi} g(\omega) \sum_{t=0}^{n-1} \sum_{k=1}^{n-t} \sum_{\ell=1+n-t}^{n} \Phi(e^{-i\omega}) N_k^{(\Delta)} N_\ell^{(\Delta)\top} \Phi_t^\top e^{-i\omega(k-\ell-t)} d\omega \right\| \\
& + \left\| \frac{1}{2\pi\sqrt{n}} \int_{-\pi}^{\pi} g(\omega) \sum_{t=0}^{n} \sum_{k=n+1-t}^{n} \sum_{\ell=1+n-t}^{n} \Phi(e^{-i\omega}) N_k^{(\Delta)} N_\ell^{(\Delta)\top} \Phi_t^\top e^{-i\omega(k-\ell-t)} d\omega \right\| \\
& + \left\| \frac{1}{2\pi\sqrt{n}} \int_{-\pi}^{\pi} g(\omega) \sum_{t=n+1}^{\infty} \sum_{k=1}^{n} \sum_{\ell=1}^{n} \Phi(e^{-i\omega}) N_k^{(\Delta)} N_\ell^{(\Delta)\top} \Phi_t^\top e^{-i\omega(k-\ell-t)} d\omega \right\| \\
& =: R_{n,1}^{(2)} + R_{n,2}^{(2)} + R_{n,3}^{(2)}.
\end{aligned}
$$

We investigate the terms separately. For the first one, the independency of $(N_k^{(\Delta)})_{k\in\mathbb{Z}}$, the Cauchy-Schwarz inequality and (5.2) yield

$$
\begin{aligned}
& \mathbb{E}\left[\left(R_{n,1}^{(2)}\right)^2\right] \\
& = \mathbb{E}\left[\frac{1}{4\pi^2 n} \left\| \int_{-\pi}^{\pi} \sum_{k=1}^{n-1} \sum_{\ell=1+k}^{n} \sum_{t=1+n-\ell}^{n-k} g(\omega) \Phi(e^{-i\omega}) N_k^{(\Delta)} N_\ell^{(\Delta)\top} \Phi_t^\top e^{-i\omega(k-\ell-t)} \right\|^2 \right]
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{4\pi^2 n}\sum_{k=1}^{n-1}\sum_{\ell=1+k}^{n}\mathbb{E}\left[\left\|\int_{-\pi}^{\pi}\sum_{t=1+n-\ell}^{n-k} g(\omega)\Phi(e^{-i\omega})N_k^{(\Delta)}N_\ell^{(\Delta)\top}\Phi_t^\top e^{-i\omega(k-\ell-t)}\right\|^2\right]\\
&\le \frac{\mathfrak{C}}{n}\sum_{k=1}^{n-1}\sum_{\ell=1+k}^{n}\left(\sum_{t_1=1+n-\ell}^{n-k}\|\Phi_{t_1}\|^2 t_1^2\right)\left(\sum_{t_2=1+n-\ell}^{n-k} t_2^{-2}\left\|\int_{-\pi}^{\pi} g(\omega)\Phi(e^{-i\omega})e^{-i\omega(k-\ell-t_2)}d\omega\right\|^2\right)\\
&\le \frac{\mathfrak{C}}{n}\sum_{k=1}^{n-1}\sum_{\ell=1+k}^{n}\left(\sum_{t_1=1+n-\ell}^{n-k}\|\Phi_{t_1}\|^2 t_1^2\right)\left(\sum_{t_2=1+n-\ell}^{n-k} t_2^{-2}\left\|\int_{-\pi}^{\pi} g(\omega)\Phi(e^{-i\omega})e^{-i\omega(k-\ell-t_2)}d\omega\right\|^2\right)\\
&\le \frac{\mathfrak{C}}{n}\sum_{k=1}^{n-1}\sum_{\ell=1+k}^{n}\left(\sum_{t_1=1+n-\ell}^{n-k}\|\Phi_{t_1}\|^2 t_1^2\right)\\
&\qquad\left(\sum_{t_2=1+n}^{n-k+\ell}(t_2-\ell)^{-2}\left\|\int_{-\pi}^{\pi} g(\omega)\Phi(e^{-i\omega})e^{-i\omega(k-t_2)}d\omega\right\|^2\right)\\
&\le \frac{\mathfrak{C}}{n}\sum_{k=1}^{n-1}\left(\sum_{t_1=1}^{n-k}\sum_{\ell=1+n-t_1}^{n}\|\Phi_{t_1}\|^2 t_1^2\right)\left(\sum_{t_2=1}^{\infty} t_2^{-2}\left\|\int_{-\pi}^{\pi} g(\omega)\Phi(e^{-i\omega})e^{-i\omega(k-t_2-n)}d\omega\right\|^2\right)\\
&\le \frac{\mathfrak{C}}{n}\left(\sum_{t=0}^{\infty}\|\Phi_t\|^2 t^3\right)\sum_{t=1}^{\infty}\frac{1}{t^2}\sum_{k=-\infty}^{\infty}\left\|\widehat{\left(g(\cdot)\Phi(e^{-i\cdot})\right)}_k\right\|^2\\
&\le \frac{\mathfrak{C}}{n}.
\end{aligned}
\tag{5.13}
$$

For the second one, consider

$$
\begin{aligned}
&\mathbb{E}\left[\left(R_{n,2}^{(2)}\right)^2\right]\\
&\le \frac{1}{4\pi^2 n}\mathbb{E}\left[\left\|\sum_{k=1}^{n}\sum_{\ell=1}^{n}\sum_{t=1+n-\min\{k,\ell\}}^{n}\int_{-\pi}^{\pi} g(\omega)\Phi(e^{-i\omega})N_k^{(\Delta)}N_\ell^{(\Delta)\top}\Phi_t^\top e^{-i\omega(k-\ell-t)}\right\|^2\right]\\
&\le \frac{1}{4\pi^2 n}\sum_{k=1}^{n}\sum_{\ell=1}^{n}\mathbb{E}\left[\left\|\sum_{t=1+n-\min\{k,\ell\}}^{n}\int_{-\pi}^{\pi} g(\omega)\Phi(e^{-i\omega})N_k^{(\Delta)}N_\ell^{(\Delta)\top}\Phi_t^\top e^{-i\omega(k-\ell-t)}\right\|^2\right]\\
&+ \frac{1}{4\pi^2 n}\left\|\sum_{k=1}^{n}\sum_{t=1+n-k}^{n}\int_{-\pi}^{\pi} g(\omega)\Phi(e^{-i\omega})\Sigma_N^{(\Delta)}\Phi_t^\top e^{i\omega t}\right\|^2\\
&+ \sum_{S,T=1}^{N}\frac{1}{4\pi^2 n}\sum_{k=1}^{n}\sum_{\substack{\ell=1\\ \ell\neq k}}^{n}\sum_{\substack{t_1,t_2=1+n\\ -\min\{k,\ell\}}}^{n}\mathbb{E}\left[\int_{-\pi}^{\pi} g(\omega)\Phi(e^{-i\omega})N_k^{(\Delta)}N_\ell^{(\Delta)\top}\Phi_{t_1}^\top[S,T]e^{-i\omega(k-\ell-t_1)}d\omega\right.\\
&\qquad\left.\int_{-\pi}^{\pi} g(\omega)\Phi(e^{i\omega})N_\ell^{(\Delta)}N_k^{(\Delta)\top}\Phi_{t_2}^\top[S,T]e^{-i\omega(k-\ell+t_2)}d\omega\right]\\
&\le \frac{\mathfrak{C}}{n}\sum_{k=1}^{n}\sum_{\ell=1}^{n}\left(\sum_{\substack{t=1+n\\ -\min\{k,\ell\}}}^{n}\frac{1}{t^2}\left\|\int_{-\pi}^{\pi} g(\omega)\Phi(e^{-i\omega})e^{-i\omega(k-\ell-t)}\right\|^2\right)\left(\sum_{\substack{t=1+n\\ -\min\{k,\ell\}}}^{n}\|\Phi_t\|^2 t^2\right)\\
&+ \frac{\mathfrak{C}}{n}\left(\sum_{t=1}^{n}\sum_{k=1-t+n}^{n}\|\Phi_t\|^2\right)\left(\sum_{t=1}^{n}\left\|\int_{-\pi}^{\pi} g(\omega)\Phi(e^{-i\omega})e^{i\omega t}\right\|^2\right)
\end{aligned}
$$

$$
\begin{aligned}
&+ \frac{\mathfrak{C}}{n}\sum_{k=1}^{n}\sum_{\substack{\ell=1\\ \ell\neq k}}^{n}\sum_{\substack{t_1,t_2=1+n\\-\min\{k,\ell\}}}^{n}\left\|\int_{-\pi}^{\pi} g(\omega)\Phi(e^{-i\omega})e^{-i\omega(k-\ell-t_1)}d\omega\right\|\|\Phi_{t_1}\| \\
&\qquad\qquad\qquad \left\|\int_{-\pi}^{\pi} g(\omega)\Phi(e^{i\omega})e^{-i\omega(k-\ell+t_2)}d\omega\right\|\|\Phi_{t_2}\| \\
&\leq \frac{\mathfrak{C}}{n}\sum_{k=1}^{n}\sum_{\ell=1}^{n}\left(\sum_{t=1+n}^{n+\ell}\frac{1}{(t-\ell)^2}\left\|\widehat{(g(\cdot)\Phi(e^{-i\cdot}))}_{k-t}\right\|^2\right)\left(\sum_{t=1+n-\min\{k,\ell\}}^{n}\|\Phi_t\|^2t^2\right) \\
&+ \frac{\mathfrak{C}}{n}\left(\sum_{t=1}^{n}t\|\Phi_t\|^2\right)\left(\sum_{t=1}^{n}\left\|\widehat{g(\cdot)\Phi(e^{-i\cdot})}_{-t}\right\|^2\right) \\
&+ \frac{\mathfrak{C}}{n}\max_{t\in\mathbb{Z}}\left\|\widehat{(g(\cdot)\Phi(e^{-i\cdot}))}_t\right\| + \max_{t\in\mathbb{Z}}\left\|\widehat{(g(\cdot)\Phi(e^{i\cdot}))}_t\right\|\sum_{t_1,t_2=1}^{n}t_1\|\Phi_{t_1}\|t_2\|\Phi_{t_2}\| \\
&\leq \frac{\mathfrak{C}}{n},
\end{aligned}
$$

where in the last inequality we used similar arguments as in (5.13),

$$\left\|\widehat{(g(\cdot)\Phi(e^{i\cdot}))}_t\right\| \leq \mathfrak{C} \quad \text{and} \quad \left\|\widehat{(g(\cdot)\Phi(e^{-i\cdot}))}_t\right\| \leq \mathfrak{C} \quad \forall t\in\mathbb{Z}.$$

The convergence of $R_{n,3}^{(2)}$ can be proven similarly. □

Proof of Theorem 5.3

Note that the representation

$$\operatorname{tr}(E_n(g)) = \operatorname{tr}\left(E_{n,N}(\Phi(e^{i\cdot})^\top g(\cdot)\Phi(e^{-i\cdot})) + E_{n,R}(g)\right)$$

holds. Due to Lemma 5.14, $E_{n,R}\xrightarrow{\mathbb{P}} 0$ in $(\mathcal{C}(\mathcal{G}_m), \|\cdot\|_m)$ and $(\mathcal{G}_m^{s\prime}, \|\cdot\|_{\mathcal{G}_m^{s\prime}})$, respectively. It suffices to show, that $\operatorname{tr}(E_{n,N}(\Phi(e^{i\cdot})^\top g(\cdot)\Phi(e^{-i\cdot})))$ converges weakly inthe appropriate space under therespective assumptions by an application of the continuous mapping theorem. Therefore, define $\mathcal{G}_N := \{\Phi(e^{i\cdot})^\top g(\cdot)\Phi(e^{-i\cdot}) : g\in\mathcal{G}_m\}\subset\mathcal{H}_N$.
Under $(C1)$

$$
\begin{aligned}
\|\widetilde{g}\|_{\mathcal{G}_N^s} &= \sum_{\ell\in\mathbb{Z}}(1+|\ell|)^{2s}\|\widehat{\left(\Phi(e^{i\cdot})^\top g(\cdot)\Phi(e^{-i\cdot})\right)}_\ell\|^2 \\
&\leq \mathfrak{C}\sum_{k\in\mathbb{Z}}\frac{1}{|k|+1}\left(\sum_{\ell\in\mathbb{Z}}(1+|\ell|)^{2s}\|\widehat{\left(\Phi(e^{i\cdot})^\top g(\cdot)\right)}_{\ell+k}\|^2\right) < \infty
\end{aligned}
$$

holds.
Under $(C2)$, the covering numbers $N(\varepsilon, d_N^{(k)}, \mathcal{G}_N)$ satisfy $N(\varepsilon, d_N^{(k)}, \mathcal{G}_N) = N(\varepsilon, d_{m,\Phi}^{(k)}, \mathcal{G}_m)$ for all $k\in\mathbb{Z}$.
Under $(C3)$ the boundedness $\|\Phi(e^{i\cdot})\|\leq\mathfrak{C}$ implies for $\widetilde{g}\in\mathcal{G}_N$ and for some $g\in\mathcal{G}_m$ $\|g\|_m = \mathfrak{C}\|\widetilde{g}\|_N$ and therefore, $N(\varepsilon,\mathcal{G}_N,d_N)\leq N(\mathfrak{C}\varepsilon,\mathcal{G}_m,d_m)$ for fixed $\varepsilon>0$.
Consequently, an application of Theorem 5.8 yields the assertion under each set of conditions. □

Proof of Remark 5.4 a)

Note that $\operatorname{tr}(E)$ is centered. Consider the Fourier coefficients $\left(\widehat{(\mathfrak{g}_j)}_h\right)_{h\in\mathbb{Z}}$ of $\mathfrak{g}_j(\cdot) = \Phi(e^{i\cdot})^\top \left(g_j(\cdot) + g_j(-\cdot)^\top\right) \Phi(e^{-i\cdot})$ for $j = 1, 2$. Then, since $(W_h)_{h\in\mathbb{N}}$ is i.i.d. and the sum is well-defined, we obtain with (3.19)

$$\begin{aligned}
&\operatorname{Cov}\left(\operatorname{tr}\left(\sum_{h=1}^{\infty} W_h \widehat{(\mathfrak{g}_1)}_h\right), \operatorname{tr}\left(\sum_{h=1}^{\infty} W_h \widehat{(\mathfrak{g}_2)}_h\right)\right)\\
&= \operatorname{Cov}\left(\sum_{h=1}^{\infty} \operatorname{vec}\left(\widehat{(\mathfrak{g}_1)}_h^T\right)^T \operatorname{vec}(W_h), \sum_{h=1}^{\infty} \operatorname{vec}\left(\widehat{(\mathfrak{g}_2)}_h^\top\right)^\top \operatorname{vec}(W_h)\right)\\
&= \sum_{h=1}^{\infty}\left(\operatorname{vec}\left(\widehat{(\mathfrak{g}_1)}_h^T\right)\right)^T \left(\Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)}\right) \operatorname{vec}\left(\widehat{(\mathfrak{g}_2)^H}_h\right)\\
&= \sum_{h=1}^{\infty} \operatorname{tr}\left(\Sigma_N^{(\Delta)} \widehat{(\mathfrak{g}_1)}_h^H \Sigma_N^{(\Delta)} \widehat{(\mathfrak{g}_2)}_h\right)\\
&= \frac{1}{2}\sum_{h\in\mathbb{Z}} \operatorname{tr}\left(\Sigma_N^{(\Delta)} \widehat{(\mathfrak{g}_1)}_h^H \Sigma_N^{(\Delta)} \widehat{(\mathfrak{g}_2)}_h\right) - \frac{1}{2}\operatorname{tr}\left(\Sigma_N^{(\Delta)} \widehat{(\mathfrak{g}_1)}_0^H \Sigma_N^{(\Delta)} \widehat{(\mathfrak{g}_2)}_0\right)\\
&= \frac{1}{4\pi}\sum_{h\in\mathbb{Z}}\sum_{\ell\in\mathbb{Z}} \operatorname{tr}\left(\Sigma_N^{(\Delta)} \widehat{(\mathfrak{g}_1)}_\ell^H \Sigma_N^{(\Delta)} \widehat{(\mathfrak{g}_2)}_h\right) \int_{-\pi}^{\pi} e^{i(h-\ell)\omega} d\omega - \frac{1}{2}\operatorname{tr}\left(\Sigma_N^{(\Delta)} \widehat{(\mathfrak{g}_1)}_0^H \Sigma_N^{(\Delta)} \widehat{(\mathfrak{g}_2)}_0\right)\\
&= \frac{1}{4\pi}\int_{-\pi}^{\pi} \operatorname{tr}\left(\frac{\Sigma_N^{(\Delta)}}{2\pi}\Phi(e^{i\omega})^\top (g_1(\omega) + g_1(-\omega)^\top)^H \Phi(e^{-i\omega})\right.\\
&\qquad \left.\frac{\Sigma_N^{(\Delta)}}{2\pi}\Phi(e^{i\omega})^\top (g_2(\omega) + g_2(-\omega)^\top) \Phi(e^{-i\omega})\right) d\omega - \frac{1}{2}\operatorname{tr}\left(\Sigma_N^{(\Delta)} \widehat{(\mathfrak{g}_1)}_0^H \Sigma_N^{(\Delta)} \widehat{(\mathfrak{g}_2)}_0\right)\\
&= \frac{1}{4\pi}\int_{-\pi}^{\pi} \operatorname{tr}\left(f_Y^{(\Delta)}(\omega)(g_1(\omega) + g_1(-\omega)^\top)^H f_Y^{(\Delta)}(\omega)(g_2(\omega) + g_2(-\omega)^\top)\right) d\omega\\
&- \operatorname{vec}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi(e^{i\omega})^\top g_1(\omega)\Phi(e^{-i\omega}) d\omega\right)^H \left(\Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)}\right)\\
&\quad \operatorname{vec}\left(\frac{1}{\pi}\int_{-\pi}^{\pi} \Phi(e^{i\omega})^\top g_2(\omega)\Phi(e^{-i\omega}) d\omega\right).
\end{aligned}$$

Thereby, and since W_0 is independent from $(W_h)_{h\in\mathbb{N}}$,

$$\begin{aligned}
&\operatorname{Cov}(\operatorname{tr}(E(g_1), \operatorname{tr}(E(g_2)))\\
&= \frac{1}{4\pi}\int_{-\pi}^{\pi} \operatorname{tr}\left(f_Y^{(\Delta)}(\omega)(g_1(\omega) + g_1(-\omega)^\top)^H f_Y^{(\Delta)}(\omega)(g_2(\omega) + g_2(-\omega)^\top)\right) d\omega\\
&- 2\operatorname{vec}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi(e^{i\omega})^\top g_1(\omega)\Phi(e^{-i\omega}) d\omega\right)^H \left(\Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)}\right)\\
&\quad \operatorname{vec}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi(e^{i\omega})^\top g_2(\omega)\Phi(e^{-i\omega}) d\omega\right) + \operatorname{vec}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi(e^{i\omega})^\top g_1(\omega)\Phi(e^{-i\omega}) d\omega\right)^H\\
&\quad \left(\mathbb{E}[N_1^{(\Delta)} N_1^{(\Delta)\top} \otimes N_1^{(\Delta)} N_1^{(\Delta)\top}] - \Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)}\right) \operatorname{vec}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi(e^{i\omega})^\top g_2(\omega)\Phi(e^{-i\omega}) d\omega\right)\\
&= \frac{1}{4\pi}\int_{-\pi}^{\pi} \operatorname{tr}\left(f_Y^{(\Delta)}(\omega)(g_1(\omega) + g_1(-\omega)^\top)^H f_Y^{(\Delta)}(\omega)(g_2(\omega) + g_2(-\omega)^\top)\right) d\omega\\
&+ \operatorname{vec}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi(e^{i\omega})^\top g_1(\omega)\Phi(e^{-i\omega}) d\omega\right)^H
\end{aligned}$$

$$\left(\mathbb{E}[N_1^{(\Delta)} N_1^{(\Delta)\top} \otimes N_1^{(\Delta)} N_1^{(\Delta)\top}] - 3\Sigma_N^{(\Delta)} \otimes \Sigma_N^{(\Delta)}\right) \operatorname{vec}\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi(e^{i\omega})^\top g_2(\omega) \Phi(e^{-i\omega}) d\omega\right)$$

holds. The assertion follows. Note that a Gaussian process is uniquely defined by its expectation function and covariance function.

CHAPTER 6

Simulation study

In this chapter, we investigate the applicability of the theoretical results of Chapter 3-Chapter 5 for finite sample sizes via simulations. First, we start with a comparison of different estimation procedures in various settings. We distinguish between multivariate CARMA processes with finite second moments, univariate CARMA processes with finite second moments and univariate α-stable CARMA processes. In the case of estimating the parameters of multivariate CARMA processes, we introduce a parametrization under which Assumption A holds. Since most of those conditions were assumed to guarantee identifiability of the underlying processes, they are a necessary prerequisite for other estimation procedures in a sampled MCARMA setting as well. We then compare the performances of the Whittle estimator and the quasi maximum likelihood estimator of Schlemm and Stelzer (2012b) in a bivariate setting. For this, we first introduce the procedure and revisit some basic properties. In the univariate setting, we compare the Whittle estimator, the adjusted Whittle estimator and the quasi maximum likelihood estimator under the assumption of existing second moments. In view of Section 4.3, the adjusted Whittle estimator might not be suitable for estimation when we suppose that the second moments of the driving process do not exist and the underlying process is not of order $(p, q) = (1, 0)$. However, we also consider the procedure in this setting and additionally investigate the behavior of the estimator of García et al. (2011). To the best of our knowledge, this estimator is the only one which was proposed for parameter estimation in a general α-stable setting.
Finally, we also examine the Grenander-Rosenblatt test statistic and the Cramér-von Mises test statistic. In particular, we compare the empirical quantiles in different settings to the corresponding theoretical one. Thereby, we determine the quantiles by Monte-Carlo simulations. The quantiles of the limit processes of the statistics are then used to do some

testing under the hypothesis of a correctly specified process and under various alternatives. For the alternatives, we simulate data which is generated by distributions corresponding to spectral densities different from the one is plugged in in the test statistics. We conclude this chapter with a bootstrap variant of the tests for a univariate CARMA(2,1) setting. The background for doing so is that the limit processes in Corollary 5.6 generally depend on the 4th moments of the driving process. In applications, the distribution of the Lévy process is not known. Obviously, we then also have no knowledge about its 4th moments. This problem can be bypassed by solely approximating it by the given data. In particular, based on the observed process, we use the Whittle procedure to estimate the parameters of the generating process and then simulate a fixed number of samples to approximate the desired moment by the empirical counterpart. As we will see, this approach yields a testing procedure which depicts nice properties in a small simulation study.

6.1. Parameter estimation

6.1.1. Parametrization in the multivariate setting

As already mentioned in Section 2.2, it is generally possible to obtain the same process by two different causal stationary linear state space models (A_1, B_1, C_1, L) and (A_2, B_2, C_2, L). We aim to just consider L^2-observationally different processes with distinguishable sampled processes. Therefore, we already introduced Assumption $(A4) - (A7)$ which, along with the causality Assumption $(A3)$, guarantee the desired identifiability. To implement these assumptions, we now present a parametrization for which the conditions are satisfied. Therefore, we introduce the transfer function of a minimal system (A, B, C, L).

Definition 6.1.
Let (A, B, C, L) be a linear state space model. The rational matrix function $H : \mathbb{R} \to \mathbb{R}^{m\times d}$ with $H(z) = C(zI_N - A)^{-1}B$ is called transfer function.

By Schlemm and Stelzer (2012a), Section 3.2, two minimal systems (A_1, B_1, C_1, L) and (A_2, B_2, C_2, L) are observationally equivalent, if they create the same transfer function. Consequently, we can identify linear state space models by their transfer functions if we assume their McMillan degree to be minimal. The following results are from Section 4.1 of Schlemm and Stelzer (2012b). They investigated the quasi maximum likelihood estimator in a similar setting and were therefore already confronted with our problem. We explicitly restate them since they are essential to derive the exact parametrization for the subsequent simulations.

Theorem 6.2 (Bernstein (2005), Theorem 4.7.5)**.**
Let H be a $m \times d$-dimensional rational matrix function of rank ℓ. Then, there exist matrix functions S_1 and S_2 with polynomial entries and constant determinant of dimensions $m\times m$ and $d \times d$, respectively, with $H = S_1 M S_2$. Denoting the ℓ-dimensional diagonal matrix with

diagonal elements $(a_i)_{i=1,\ldots,\ell}$ *as diag*$(a_i)_{i=1,\ldots,\ell}$,*the function* M *is an* $m \times d$*-dimensional function with*

$$M = \begin{pmatrix} \operatorname{diag}(\varepsilon_i/\psi_i)_{i=1,\ldots,\ell} & 0_{\ell\times d-\ell} \\ 0_{m-\ell\times\ell} & 0_{m-\ell\times d-\ell} \end{pmatrix},$$

with monic polynomials $\varepsilon_1, \ldots, \varepsilon_\ell, \psi_1, \ldots, \psi_\ell$*. These polynomials are uniquely determined by* H *and satisfy:*

(i) the polynomials ε_i *and* ψ_i *have no common roots for each* $i \in \{1, \ldots, \ell\}$,

(ii) the polynomial $\varepsilon_i \circ \psi_{i+1}$ *divides the polynomial* $\varepsilon_{i+1} \circ \psi$.

(S_1, M, S_2) *is called the Smith-McMillan decomposition of* H*.*

We now define the Kronecker indices of the polynomial H as $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_m) \in \mathbb{N}^m$. Thereby, $\eta_1, \ldots, \eta_\ell$ denote the degrees of the polynomials ψ in the Smith-McMillan decomposition of H and $\eta_i = 0$ for $\ell < i \leq m$. It should be mentioned that the McMillan degree of some rational matrix function equals the sum over its Kronecker indices. We furthermore define $\eta_{i,j} = \min\{\eta_i + \mathbb{1}_{\{i>j\}}, \eta_j\}$ and are now able to approach the desired parametrization.

Theorem 6.3 (Echelon state space representation, Guidorzi (1975), Section 3)**.** *Let* H *be an* $m \times d$*-dimensional rational matrix function with Kronecker indices* $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_m)$*. Then, a unique realization* (A, B, C) *of* H *of dimension* $N = \sum_{i=1}^{\ell} \eta_i$ *is given by the following structure:*

(i) The matrix $A = (A_{ij})_{i,j=1,\ldots m} \in \mathbb{R}^{N\times N}$ *is a block matrix with blocks* $A_{ij} \in \mathbb{R}^{\eta_i,\eta_j}$ *given by*

$$A_{ij} = \begin{pmatrix} 0 & \ldots & \ldots & \ldots & \ldots & 0 \\ \vdots & & & & & \vdots \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 \\ \alpha_{ij,1} & \ldots & \alpha_{ij,\eta_{i,j}} & 0 & \ldots & 0 \end{pmatrix} + \delta_{i,j} \begin{pmatrix} 0 & \\ \vdots & I_{\eta_i-1} \\ 0 & \\ 0 & \ldots \end{pmatrix},$$

(ii) B *is an unrestricted* $N \times d$ *dimensional real matrix,*

(iii) if $\eta_i > 0$*, for* $i = 1, \ldots m$*, then*

$$C = \begin{pmatrix} 1 & 0 & \ldots & 0 & \vdots & 0 & 0 & \ldots & 0 & \vdots & \vdots & & & \\ & 0_{(m-1)\times\eta_1} & & & \vdots & 1 & 0 & \ldots & 0 & \vdots & \vdots & & 0_{(m-1)\times\eta_m} & \\ & & & & \vdots & & & 0_{(m-2)\times\eta_2} & & \vdots & \vdots & 1 & 0 & \ldots & 0 \end{pmatrix},$$

where $0_{i\times j}$ *denotes the* $i \times j$*-dimensional* $=$*-matrix.*

In general, components of η could be 0. Then, the corresponding row of C would also freely vary. However, this case is not relevant for our chosen settings. In view of the formal differential equation representation (2.8) of the process, we obtain the polynomials P and Q and therefore an equivalent MCARMA representation by the next result.

Theorem 6.4 (Echelon MCARMA realization, Guidorzi (1975), Section 3)**.**
Let H be an $m \times d$-dimensional rational matrix function with Kronecker indices $\boldsymbol{\eta}$. Assume that (A, B, C) is a realization of H which is parameterized as in Theorem 6.3. Then, a unique left matrix fraction description $P^{-1}Q$ of H is given by $P(z) = (P(z)[i,j])_{i,j=1,\ldots,m}, Q(z) = (Q(z)[i,j])_{i=1,\ldots,m,j=1,\ldots,d}$, where

$$P(z)[i,j] = \delta_{i,j} z^{\eta_i} - \sum_{k=1}^{\eta_{i,j}} \alpha_{ij,k} z^{k-1}, \ Q(z)[i,j] = \sum_{k=1}^{\eta_i} \kappa[\eta_1 + \ldots + \eta_{i-1} + k, j] z^{k-1}$$

and the coefficient $\kappa[i,j]$ is the $[i,j]$th entry of the matrix $K = TB$. The matrix $T = (T_{ij})_{i,j=1,\ldots,m} \in \mathbb{R}^{N \times N}$ is a block matrix with blocks $T_{ij} \in R^{\eta_i \times \eta_j}$ given by

$$\begin{pmatrix} -\alpha_{ij,2} & \ldots & -\alpha_{ij,\eta_{i,j}} & 0 & \ldots & 0 \\ \vdots & \iddots & & & & \vdots \\ -\alpha_{ij,\eta_{i,j}} & & & & & \vdots \\ 0 & & & & & \vdots \\ \vdots & & & & & \vdots \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 \end{pmatrix} + \delta_{i,j} \begin{pmatrix} 0 & 0 & \ldots & \ldots & 0 & 1 \\ 0 & 0 & \ldots & \ldots & 1 & 0 \\ \vdots & \vdots & & \iddots & \vdots & \vdots \\ \vdots & \vdots & \iddots & & \vdots & \vdots \\ 0 & 1 & \ldots & \ldots & 0 & 0 \\ 1 & 0 & \ldots & \ldots & 0 & 0 \end{pmatrix}.$$

Note that the orders of the polynomials P and Q satisfy $p = \max\{\eta_1, \ldots, \eta_d\}$ and $0 \leq q \leq p - 1$. As it is, this parametrization does not guarantee identifiability since there is an orthogonal invariance in the spectral factorization of the rational matrix functions, see Theorem 3.5 of Schlemm and Stelzer (2012b). Therefore, we have to assume a normalization condition as $H(0) = H_0$ where H_0 is an $m \times d$-dimensional matrix. There are many possibilities for choosing H_0. When $d = m$ it is an often used condition to set $H_0 = -I_d$. Independent of the choice of H_0, we then have to determine which parameters are free and which are functionally dependent on those. Following Schlemm and Stelzer (2012a), we keep the $\alpha_{ij,k}$ as free parameters and fix some of the entries of B. More precisely, replacing the $[\eta_1 + \ldots \eta_{i-1} + 1, t]$th entry of K by the $[i,j]$th entry of the matrix $-(\alpha_{st,1})_{s,t} H_0$ makes some entries of B functionally dependent, since $B = T^{-1}K$ holds. The resulting parametrization with $H_0 = -I_2$ is the one which we choose in all simulations of 2-dimensional CARMA processes. By Schlemm and Stelzer (2012b), the canonical state space representations and canonical MCARMA realizations are as in Table 6.1 and Table 6.2.

6.1.2. Estimation of MCARMA processes with existing second moments

In our multivariate simulations, we compare the behavior of the Whittle estimator to the performance of the quasi maximum likelihood estimator (QMLE) of Schlemm and Stelzer (2012b). For univariate processes with existing second moments, we also investigate the

$\boldsymbol{\eta}$	A	B	C
(1,1)	$\begin{pmatrix} \vartheta_1 & \vartheta_2 \\ \vartheta_3 & \vartheta_4 \end{pmatrix}$	$\begin{pmatrix} \vartheta_1 & \vartheta_2 \\ \vartheta_3 & \vartheta_4 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
(1,2)	$\begin{pmatrix} \vartheta_1 & \vartheta_2 & 0 \\ 0 & 0 & 1 \\ \vartheta_3 & \vartheta_4 & \vartheta_5 \end{pmatrix}$	$\begin{pmatrix} \vartheta_1 & \vartheta_2 \\ \vartheta_6 & \vartheta_7 \\ \vartheta_3 + \vartheta_5\vartheta_6 & \vartheta_4 + \vartheta_5\vartheta_7 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$
(2,1)	$\begin{pmatrix} 0 & 1 & 0 \\ \vartheta_1 & \vartheta_2 & \vartheta_3 \\ \vartheta_4 & \vartheta_5 & \vartheta_6 \end{pmatrix}$	$\begin{pmatrix} \vartheta_7 & \vartheta_8 \\ \vartheta_1 + \vartheta_2\vartheta_7 & \vartheta_3 + \vartheta_2\vartheta_8 \\ \vartheta_4 + \vartheta_5\vartheta_7 & \vartheta_6 + \vartheta_5\vartheta_8 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
(2,2)	$\begin{pmatrix} 0 & 1 & 0 & 0 \\ \vartheta_1 & \vartheta_2 & \vartheta_3 & \vartheta_4 \\ 0 & 0 & 0 & 1 \\ \vartheta_5 & \vartheta_6 & \vartheta_7 & \vartheta_8 \end{pmatrix}$	$\begin{pmatrix} \vartheta_9 & \vartheta_{10} \\ \vartheta_1 + \vartheta_4\vartheta_{11} + \vartheta_2\vartheta_9 & \vartheta_3 + \vartheta_2\vartheta_{10} + \vartheta_4\vartheta_{12} \\ \vartheta_{11} & \vartheta_{12} \\ \vartheta_5 + \vartheta_8\vartheta_{11} + \vartheta_6\vartheta_9 & \vartheta_7 + \vartheta_6\vartheta_{10} + \vartheta_8\vartheta_{12} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

Table 6.1.: Canonical state space realizations (A, B, C) of rational transfer functions with normalization $H_0 = -I_2$ and different Kronecker indices $\boldsymbol{\eta}$.

$\boldsymbol{\eta}$	$P(z)$	$Q(z)$	(p, q)
(1,1)	$\begin{pmatrix} z - \vartheta_1 & -\vartheta_2 \\ -\vartheta_3 & z - \vartheta_4 \end{pmatrix}$	$\begin{pmatrix} \vartheta_1 & \vartheta_2 \\ \vartheta_3 & \vartheta_4 \end{pmatrix}$	(1,0)
(1,2)	$\begin{pmatrix} z - \vartheta_1 & -\vartheta_2 \\ -\vartheta_3 & z^2 - \vartheta_4 z - \vartheta_5 \end{pmatrix}$	$\begin{pmatrix} \vartheta_1 & \vartheta_2 \\ \vartheta_6 z + \vartheta_3 & \vartheta_7 z + \vartheta_5 \end{pmatrix}$	(2,1)
(2,1)	$\begin{pmatrix} z^2 - \vartheta_1 z - \vartheta_2 & -\vartheta_3 \\ -\vartheta_4 z - \vartheta_5 & z - \vartheta_6 \end{pmatrix}$	$\begin{pmatrix} \vartheta_7 z + \vartheta_2 & \vartheta_8 z + \vartheta_3 \\ \vartheta_5 & \vartheta_6 \end{pmatrix}$	(2,1)
(2,2)	$\begin{pmatrix} z^2 - \vartheta_1 z - \vartheta_2 & -\vartheta_3 z - \vartheta_4 \\ -\vartheta_5 z - \vartheta_6 & z^2 - \vartheta_7 z - \vartheta_8 \end{pmatrix}$	$\begin{pmatrix} \vartheta_9 z + \vartheta_2 & \vartheta_{10} z + \vartheta_4 \\ \vartheta_{11} z + \vartheta_6 & \vartheta_{12} z + \vartheta_8 \end{pmatrix}$	(2,1)

Table 6.2.: Canonical MCARMA realizations (P, Q) of rational transfer functions with normalization $H_0 = -I_2$ and different Kronecker indices $\boldsymbol{\eta}$. The order of the polynomials is denoted as (p, q).

behavior of the adjusted Whittle estimator. Before presenting our simulation studies, the QMLE has to be introduced.

The quasi maximum likelihood estimator

The quasi maximum likelihood procedure for equidistantly sampled stationary continuous-time linear state space models is based on similar ideas as in the discrete-time setting. In a VARMA setting, Kalman filtering yields the linear innovations and their covariance matrices. Those are then taken to construct the log-likelihood function, see (11.5.4) of Brockwell and Davis (1991). In the same way, we approach our MCARMA setting. First, by Theorem 2.5, the sampled processes have an MA(∞) representation where the white noise are the linear innovations. More precisely, considering the parameterized setting with appropriate assumptions, we have for every $\vartheta \in \Theta$ by Theorem 2.8

$$\varepsilon_k^{(\Delta)}(\vartheta) = \Pi(\mathfrak{B}, \vartheta) Y_k^{(\Delta)}(\vartheta).$$

However, when estimating the true parameter, we only observe the sampled process corresponding to ϑ_0. Therefore, plugging in some different parameter yields the sequence of the pseudo-innovations, which we already considered in the proof of Proposition 3.14:

$$\xi_k^{(\Delta)}(\vartheta) = \Pi(\mathfrak{B}, \vartheta) Y_k^{(\Delta)}(\vartheta_0), \quad k \in \mathbb{Z}.$$

Obviously, this sequence coincides with the linear innovations for $\vartheta = \vartheta_0$ and generally differs otherwise. However, since we just observe a sample $Y_1^{(\Delta)}, \ldots, Y_n^{(\Delta)}$ of size n and therefore do not have the full history of $Y^{(\Delta)}$, the pseudo-innovations can not be calculated. Consequently, we have to approximate them. In view of the state space representation

$$\begin{aligned} \widehat{X}_k^{(\Delta)}(\vartheta) &= \left(e^{A(\vartheta)\Delta} - K^{(\Delta)}(\vartheta) C(\vartheta)\right) \widehat{X}_{k-1}^{(\Delta)}(\vartheta) + K^{(\Delta)}(\vartheta) Y_{k-1}^{(\Delta)}, \\ \xi_k^{(\Delta)}(\vartheta) &= Y_k^{(\Delta)}(\vartheta) - C(\vartheta) \widehat{X}_k^{(\Delta)}(\vartheta), \quad k \in \mathbb{Z}, \end{aligned}$$

we adapt the pseudo-innovations by fixing an initial value $\widehat{X}_1^{(\Delta)}(\vartheta) = \widehat{X}_{\text{init}}^{(\Delta)}(\vartheta)$. In our simulations, we choose $\widehat{X}_{\text{init}}^{(\Delta)} \equiv 0$ but we could alternatively use a sample from the stationary distribution of the state process $X(\vartheta)$ or take any other deterministic value. Finally, we define the quasi likelihood function

$$\mathcal{L}_n(\vartheta) = \frac{1}{n} \sum_{k=1}^{n} d \log(2\pi) + \log(\det(V^{(\Delta)}(\vartheta))) + \xi_k^{(\Delta)}(\vartheta)^\top V^{(\Delta)}(\vartheta)^{-1} \xi_k^{(\Delta)}(\vartheta).$$

Hence, the quasi maximum likelihood estimator ϑ_n^* is obtained by minimizing $\mathcal{L}_n$, i.e.,

$$\vartheta_n^* = \arg\min_{\vartheta \in \Theta} \mathcal{L}_n(\vartheta).$$

Note that $\mathcal{L}_n(\vartheta)$ originates from the Gaussian likelihood of ϑ by exchanging the pseudo-innovations by the approximate pseudo-innovations and taking $-2/n$ times the logarithm. This is where the name of the quasi maximum likelihood estimator originates.

Simulations

We simulate continuous-time state space models with an Euler-Maruyama scheme for differential equations with initial value $X(0) = Y(0) = 0$ and step size 0.01. Using $\Delta = 1$ and the interval $[0, 500]$, we therefore get $n_1 = 500$ discrete observations. Furthermore, we investigate how the results change qualitatively when we consider the intervals $[0, 2000]$ and $[0, 5000]$, which imply $n_2 = 2000$ and $n_3 = 5000$ observations, respectively. In each sample, we use 500 replicates. We investigate the estimation procedure based on two different driving Lévy processes. Since the Brownian motion is the most common Lévy process, we examine Whittle‘s estimator based on a Brownian motion. As a second case, we analyze its performance based on a bivariate normal-inverse Gaussian (NIG) Lévy process, which is often used in modeling stochastic volatility or stock returns, see Barndorff-Nielsen (1997). The resulting increments of this process are characterized by the density

$$f(x,\mu,\alpha,\beta,\delta_{NIG},\Delta_{NIG}) = \frac{\delta_{NIG}}{2\pi}\frac{(1+\alpha g(x))}{g(x)^3}\exp(\delta_{NIG}\kappa+\beta^\top x-\alpha g(x)), \quad x\in\mathbb{R}^2,$$

with

$$g(x)=\sqrt{\delta_{NIG}^2+\langle x-\mu,\Delta_{NIG}(x-\mu)\rangle},\ \kappa^2=\alpha^2-\langle\beta,\Delta_{NIG}\beta\rangle>0.$$

Thereby, $\beta\in\mathbb{R}^2$ is a symmetry parameter, $\delta_{NIG}\geq 0$ is a scale parameter and the positive definite matrix Δ_{NIG} models the dependency between the two components of the bivariate Lévy process $(L_t)_{t\in\mathbb{R}}$. We set $\mu=-(\delta_{NIG}\Delta_{NIG}\beta)/\kappa$ to guarantee that the resulting Lévy process is centered, see, e.g., Øigård et al. (2005) or Barndorff-Nielsen (1997) for more details. For better comparability of the Brownian motion driven case and the NIG Lévy-driven case, we choose the parameters of the NIG Lévy process in a way that the resulting covariance matrices of both the Lévy processes are the same.
In the multivariate setting, we consider bivariate MCARMA(2,1) processes of the form

$$dX_t(\vartheta)=A(\vartheta)X_t(\vartheta)dt+B(\vartheta)dL_t(\vartheta) \quad \text{and} \quad Y_t(\vartheta)=C(\vartheta)X_t(\vartheta), \quad t\geq 0,$$

with

$$A(\vartheta)=\begin{pmatrix}\vartheta_1 & \vartheta_2 & 0\\ 0 & 0 & 1\\ \vartheta_3 & \vartheta_4 & \vartheta_5\end{pmatrix}, \qquad B(\vartheta)=\begin{pmatrix}\vartheta_1 & \vartheta_2\\ \vartheta_6 & \vartheta_7\\ \vartheta_3+\vartheta_5\vartheta_6 & \vartheta_6+\vartheta_5\vartheta_7\end{pmatrix},$$
$$C(\vartheta)=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}, \qquad \Sigma_L(\vartheta)=\begin{pmatrix}\vartheta_8 & \vartheta_9\\ \vartheta_9 & \vartheta_{10}\end{pmatrix}.$$

This is the parametrization which is given in Table 6.1 and the representations of the corresponding AR polynomial P and MA polynomial Q are by Table 6.2

$$P(z)=\begin{pmatrix}z-\vartheta_1 & -\vartheta_2\\ -\vartheta_3 & z^2-\vartheta_4 z-\vartheta_5\end{pmatrix}, \quad Q(z)=\begin{pmatrix}\vartheta_1 & \vartheta_2\\ \vartheta_6 z+\vartheta_3 & \vartheta_7 z+\vartheta_5\end{pmatrix}.$$

Furthermore, we get the order $(2,1)$ of the MCARMA process from there as well. In our example, the true parameter value is

$$\vartheta_0^{(1)}=(-1,-2,1,-2,-3,1,2,0.4751,-0.1622,0.3708).$$

Note that the last three parameters are the variance parameters of the driving process and are chosen this way to obtain nice parameters for the NIG Lévy process parametrization.

In particular, we rely on the parameters

$$\delta_{NIG}^{(1)} = 1, \quad \alpha^{(1)} = 3, \quad \beta^{(1)} = (1,1)^T, \quad \Delta_{NIG}^{(1)} = \begin{pmatrix} 5/4 & -1/2 \\ -1/2 & 1 \end{pmatrix}$$

to generate this process. The estimation results are summarized in Table 6.3 and Table 6.4 for the Brownian motion driven model and the NIG driven model, respectively. The consistency can be observed in all simulations, namely the bias and the standard deviations are decreasing for increasing sample size for the Whittle estimator and the quasi maximum likelihood estimator. The performance of the estimators is very similar.

In addition, we investigate bivariate MCAR(1) processes for both the Brownian motion and the NIG driven setting. The parametrization of the MCAR(1) model is again given in Table 6.1 and it is

$$A(\vartheta) = \begin{pmatrix} \vartheta_1 & \vartheta_2 \\ \vartheta_3 & \vartheta_4 \end{pmatrix} = B(\vartheta), \quad C(\vartheta) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Sigma_L(\vartheta) = \begin{pmatrix} \vartheta_5 & \vartheta_6 \\ \vartheta_6 & \vartheta_7 \end{pmatrix}.$$

We choose the parameter

$$\vartheta_0^{(2)} = (1, -2, 3, -4, 0.7513, -0.3536, 0.3536).$$

The results of this simulation study are summarized in Table 6.5 and Table 6.6, respectively. Likewise, as for the MCARMA(2,1) model in Table 6.3 and Table 6.4, the Whittle estimator and the QMLE converge very fast. To see whether the estimation procedures work for small sample sizes, we also investigate the Whittle estimator and the quasi maximum likelihood estimator for $n_4 = 50$ in the MCARMA(2, 1) and the MCAR(1) settings. The results are in Table 6.7. As before, the procedures behave similarly. However, they work well and are therefore also suited for small sample sizes.

Additionally, in the MCAR(1) setting, we use the parameter

$$\vartheta_0^{(3)} = (-0.01, 0, 7, -1, 0.7513, -0.3536, 0.3536).$$

The background is that then one eigenvalue of $A(\vartheta_0^{(3)})$ is close to zero. An eigenvalue equal to zero results in a non-stationary MCARMA process. Table 6.8 shows the results for this setting for $n_2 = 2000$, and both the Brownian motion and the NIG driven model. The Whittle estimator and the QMLE estimate the parameters very well. But it is striking that the bias of several parameters of the QMLE even vanishes.

Since we introduced an alternative estimator for the univariate setting, we perform an additional simulation study concerning one dimensional CARMA processes. In accordance to Assumption $\widetilde{A}$, the variance parameter σ_L^2 of the Lévy process is fixed in this study and

$n_1 = 500$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
-1	-0.9969	0.0031	0.0325	-1.0012	0.0012	0.0572
-2	-2.0218	0.0218	0.0582	-2.0128	0.0128	0.0689
1	0.9980	0.0020	0.0520	1.0075	0.0075	0.0722
-2	-2.0498	0.0498	0.1060	-1.9797	0.0203	0.0758
-3	-2.9840	0.0160	0.0498	-2.9913	0.0087	0.0907
1	1.0062	0.0062	0.1309	0.8034	0.1966	0.3896
2	1.9983	0.0017	0.0532	2.0036	0.0036	0.0768
0.4751	0.4746	0.0005	0.0407	0.4693	0.0048	0.0691
-0.1622	-0.1629	0.0007	0.0134	-0.1624	0.0002	0.0405
0.3708	0.3706	0.0002	0.0064	0.3712	0.0004	0.0328
$n_2 = 2000$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
-1	-0.9970	0.0030	0.0155	-0.9957	0.0043	0.0260
-2	-2.0062	0.0062	0.0252	-2.0047	0.0047	0.0350
1	0.9909	0.0091	0.0266	1.0038	0.0038	0.0399
-2	-2.0394	0.0394	0.0501	-2.0122	0.0122	0.0481
-3	-2.9857	0.0143	0.0371	-3.0350	0.0350	0.0583
1	1.0775	0.0775	0.1030	0.9572	0.0428	0.2583
2	2.0033	0.0033	0.0205	2.0452	0.0452	0.0463
0.4751	0.4731	0.0020	0.0092	0.4719	0.0032	0.0321
-0.1622	-0.1620	0.0002	0.0059	-0.1632	0.0010	0.0197
0.3708	0.3708	0	0.0037	0.3731	0.0023	0.0167
$n_3 = 5000$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
-1	-1.0028	0.0028	0.0172	-0.9960	0.0040	0.0174
-2	-1.9954	0.0146	0.0041	-2.0059	0.0059	0.0196
1	0.9972	0.0028	0.0133	1.0052	0.0052	0.0268
-2	-2.0202	0.0202	0.0210	-2.0043	0.0043	0.0284
-3	-3.0091	0.0091	0.0441	-3.0013	0.0013	0.0261
1	1.0585	0.0585	0.0409	1.0253	0.0253	0.1249
2	2.0109	0.0109	0.0318	2.0479	0.0479	0.0346
0.4751	0.4759	0.0008	0.0100	0.4735	0.0016	0.0200
-0.1622	-0.1652	0.0030	0.0088	-0.1634	0.0012	0.0135
0.3708	0.3904	0.0196	0.0079	0.3727	0.0019	0.0109

Table 6.3.: Estimation results for a Brownian motion driven bivariate MCARMA(2,1) process with parameter $\vartheta_0^{(1)}$.

has not to be estimated. We consider a CARMA(2,1) model where

$$A(\vartheta) = \begin{pmatrix} 0 & 1 \\ \vartheta_1 & \vartheta_2 \end{pmatrix}, \qquad B(\vartheta) = \begin{pmatrix} \vartheta_3 \\ \vartheta_1 + \vartheta_2\vartheta_3 \end{pmatrix} \quad \text{and} \quad C(\vartheta) = (1\ 0).$$

$n_1 = 500$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
-1	-0.9555	0.0445	0.1559	-0.9651	0.0349	0.1854
-2	-1.8822	0.1178	0.2653	-1.6978	0.3022	0.3452
1	0.8746	0.1254	0.1888	1.1479	0.1479	0.2526
-2	-2.0981	0.0981	0.2273	-2.0066	0.0066	0.2962
-3	-3.1833	0.1833	0.2517	-3.0578	0.0578	0.4076
1	1.0533	0.0533	0.3614	1.0272	0.0272	1.2301
2	2.0461	0.0461	0.5710	2.0490	0.0490	1.6673
0.4751	0.4992	0.0241	0.1061	0.4645	0.0106	0.8220
-0.1622	-0.1520	0.0102	0.1130	-0.1669	0.0047	0.3317
0.3708	0.4100	0.0392	0.1081	0.3748	0.0040	0.6100
$n_2 = 2000$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
-1	-1.0351	0.0351	0.1224	-0.9673	0.0327	0.0243
-2	-1.8779	0.1221	0.1894	-1.0564	0.0426	0.0713
1	0.9457	0.0543	0.2620	1.1331	0.1331	0.1214
-2	-1.9586	0.0414	0.2573	-1.9494	0.0506	0.0827
-3	-3.1682	0.1682	0.2238	-3.1990	0.1990	0.4911
1	1.1234	0.1234	0.3120	1.1720	0.1720	0.5933
2	2.0842	0.0842	0.4842	2.0432	0.0432	0.1817
0.4751	0.5010	0.0259	0.1000	0.5237	0.0486	0.2726
-0.1622	-0.1740	0.0118	0.0992	-0.0856	0.0766	0.1413
0.3708	0.3908	0.0200	0.0758	0.3220	0.0488	0.0049
$n_3 = 5000$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
-1	-1.0238	0.0238	0.1182	-0.9844	0.0156	0.0194
-2	-1.9954	0.0046	0.2048	-2.0139	0.0139	0.0246
1	0.9942	0.0058	0.1517	1.0102	0.0102	0.0299
-2	-2.2202	0.2202	0.2210	-2.0043	0.0043	0.0284
-3	-3.0104	0.0104	0.2463	-3.0015	0.0015	0.2291
1	1.0585	0.0585	0.2409	1.0655	0.0655	0.1347
2	2.1169	0.1169	0.0866	2.0400	0.0400	0.0355
0.4751	0.4855	0.0104	0.1180	0.4737	0.0018	0.0206
-0.1622	-0.1682	0.0060	0.0408	-0.1634	0.0012	0.0145
0.3708	0.3908	0.0200	0.0842	0.3730	0.0022	0.0139

Table 6.4.: Estimation results for a NIG driven bivariate MCARMA(2,1) process with parameter $\vartheta_0^{(1)}$.

Note that this parametrization differs slightly from the one introduced in Equation (2.3). However, since the output process $Y(\vartheta)$ of this minimal state space model is of dimension

$n_1 = 500$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
1	1.0018	0.0018	0.0301	1.0045	0.0045	0.0362
-2	-2.0063	0.0063	0.0321	-2.0068	0.0068	0.0357
3	2.9966	0.0034	0.0399	3.0055	0.0055	0.0604
-4	-3.9980	0.0020	0.0399	-4.0019	0.0019	0.0565
0.7513	0.7543	0.0030	0.0516	0.7522	0.0009	0.0923
-0.3536	-0.3573	0.0037	0.0463	-0.3531	0.0005	0.0674
0.3536	0.3685	0.0149	0.0510	0.3704	0.0168	0.0714
$n_2 = 2000$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
1	1.0035	0.0035	0.0150	1.0039	0.0039	0.0181
-2	-2.0067	0.0067	0.0165	-2.0066	0.0066	0.0192
3	2.9991	0.0009	0.0192	3.0021	0.0021	0.0286
-4	-3.9987	0.0013	0.0223	-4.0003	0.0003	0.0302
0.7513	0.7532	0.0019	0.0257	0.7514	0.0001	0.0401
-0.3536	-0.3603	0.0067	0.0248	-0.3574	0.0038	0.0352
0.3536	0.3675	0.0139	0.0280	0.3706	0.0170	0.0376
$n_3 = 5000$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
1	1.0042	0.0042	0.0101	1.0050	0.0050	0.0117
-2	-2.0062	0.0062	0.0106	-2.0074	0.0074	0.0111
3	-2.9996	0.0004	0.0114	3.0021	0.0021	0.0169
-4	-3.9965	0.0035	0.0158	-4.0013	0.0013	0.0196
0.7513	0.7537	0.0024	0.0173	0.7549	0.0036	0.0258
-0.3536	-0.3596	0.0060	0.0166	-0.3559	0.0023	0.0201
0.3536	0.3663	0.0027	0.0169	0.3693	0.0157	0.0200

Table 6.5.: Estimation results for a Brownian motion driven bivariate MCAR(1) process with parameter $\vartheta_0^{(2)}$.

one, the order of the AR polynomial p is equal to $N = 2$ and the order of the MA polynomial is $q = p - 1 = 1$. This confirms that we really have a CARMA$(2, 1)$ process. In our simulation study the true parameter is

$$\vartheta_0^{(4)} = (-2, -2, -1).$$

The simulation results for the Brownian motion driven and the NIG driven CARMA(2,1) process are given in Table 6.9 and Table 6.10, respectively. For all sample sizes, the Whittle estimator and the QMLE behave very similar and give excellent estimation results. Whereas for small sample sizes the adjusted Whittle estimator is remarkably worse, for increasing sample sizes it performs much better and seems to converge. This behavior

$n_1 = 500$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
1	0.9905	0.0095	0.0407	0.9806	0.0194	0.0460
-2	-1.9871	0.0129	0.0531	-2.0038	0.0038	0.0579
3	2.9920	0.0080	0.0579	2.9240	0.0760	0.0842
-4	-3.9409	0.0591	0.1027	-3.9918	0.0082	0.0894
0.7513	0.7281	0.0232	0.1869	0.7125	0.0388	0.0568
-0.3536	-0.3366	0.0170	0.0302	-0.3251	0.0285	0.0497
0.3536	0.3381	0.0155	0.0335	0.3182	0.0354	0.0486
$n_2 = 2000$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
1	0.9916	0.0084	0.0261	0.9839	0.0161	0.0316
-2	-1.9892	0.0110	0.0321	-2.0072	0.0072	0.0320
3	2.9797	0.0203	0.0416	2.9377	0.0623	0.0576
-4	-3.9700	0.0300	0.0767	-4.0051	0.0051	0.0561
0.7513	0.7489	0.0024	0.1392	0.7210	0.0303	0.0351
-0.3536	-0.3603	0.0067	0.0241	-0.3224	0.0312	0.0312
0.3536	0.3417	0.0119	0.0224	0.3352	0.0184	0.0300
$n_3 = 5000$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
1	0.9952	0.0048	0.0186	0.9810	0.0190	0.0240
-2	-1.9890	0.0110	0.0253	-2.0086	0.0086	0.0289
3	2.9789	0.0211	0.0365	2.9341	0.0659	0.0478
-4	-3.9849	0.0151	0.0611	-4.0064	0.0064	0.0516
0.7513	0.7500	0.0013	0.0749	0.6912	0.0601	0.0428
-0.3536	-0.3600	0.0064	0.0148	-0.3412	0.0124	0.0237
0.3536	0.3499	0.0037	0.0201	0.3208	0.0328	0.0238

Table 6.6.: Estimation results for a NIG driven bivariate MCAR(1) process with parameter $\vartheta_0^{(2)}$.

corresponds to the theoretical results of Section 4.2.

We also considered a CAR(3) process in the univariate setting.

For the univariate CAR(3) processes with parametrization

$$A(\vartheta) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \vartheta_1 & \vartheta_2 & \vartheta_3 \end{pmatrix}, \qquad B(\vartheta) = \begin{pmatrix} 0 \\ 0 \\ \vartheta_1 \end{pmatrix}, \quad C(\vartheta) = (1\ 0\ 0).$$

and

$$\vartheta_0^{(5)} = (-6, -11, -6),$$

we once again choose the Brownian motion and the NIG Lévy process as driving processes.

The results are documented in Table 6.11 and Table 6.12. They correspond to the results of Table 6.9 and Table 6.10, respectively for CARMA(2,1) processes.

6.1.3. Estimation of α-stable CARMA processes

We now leave the light-tailed setting and investigate the performance of the adjusted Whittle estimator for finite samples in the α-stable setting. Since we compare it with the behavior of the estimator introduced in García et al. (2011), we quickly revisit the procedure which was presented therein.

The estimator of García et al. (2011)

The estimator of García et al. (2011) is based on an indirect approach. Denoting the zeros of the AR(p) polynomial a as $\lambda_1, \ldots, \lambda_p$ which are assumed to be distinct and defining $a_D^{(\Delta)}(z) = \prod_{j=1}^{p}(1 - e^{\lambda_j \Delta} z)$, we have by Proposition 2.11 that the sampled process $Y^{(\Delta)}$ satisfies the equation

$$a_D^{(\Delta)}(\mathfrak{B}) Y_k^{(\Delta)} = U_k^{(\Delta)}, \quad k \in \mathbb{N}, \tag{6.1}$$

where $(U_k^{(\Delta)})_{k\in\mathbb{N}}$ is a $(p-1)$-dependent sequence. For CARMA processes with finite second moments, $(U_k^{(\Delta)})_{k\in\mathbb{N}}$ is a MA$(p-1)$ process such that $Y^{(\Delta)}$ is an ARMA$(p, p-1)$ process with an uncorrelated but not independent white noise, see Proposition 3 of Brockwell et al. (2011). García et al. (2011) proposed to fit an ARMA$(p, p-1)$ model to the observations $Y_1^{(\Delta)}, \ldots, Y_n^{(\Delta)}$ by standard maximum likelihood estimation for Gaussian ARMA models. The estimated autoregressive part of that ARMA model in discrete time is denoted by $\widehat{a}_D^{(\Delta)}$ and the estimated moving average part is $\widehat{c}_D^{(\Delta)}$. The logarithmic zeros of $\widehat{a}_D^{(\Delta)}$ divided by $-\Delta$ are then estimators $\widehat{\lambda}_1, \ldots, \widehat{\lambda}_p$ for the zeros $\lambda_1, \ldots, \lambda_p$ of a. Hence, we obtain an estimator $\widehat{a}$ for the autoregressive polynomial a. In a final step, the MA polynomial c of the CARMA process is determined. Therefore the parameter $\vartheta = (\vartheta_1, \vartheta_2)$ is divided in two parts where ϑ_1 models the AR coefficients and ϑ_2 the MA coefficients of the CARMA process. Now the autocorrelation function $\rho^{(\mathrm{MA})}_{\widehat{a},\vartheta_2}$ of $\widehat{a}_D^{(\Delta)}(\mathfrak{B}) Y^{(\Delta)}(\widehat{\vartheta}_1, \vartheta_2)$ and the autocorrelation function $\rho^{(\mathrm{MA})}_{\widehat{c}_D^{(\Delta)}}$ of a discrete-time moving average process with moving average polynomial $\widehat{c}_D^{(\Delta)}$ is calculated and $\widehat{\vartheta}_2$ is derived numerically as solution of $\rho^{(\mathrm{MA})}_{\widehat{a}_D^{(\Delta)},\widehat{\vartheta}_2}(k) = \rho^{(\mathrm{MA})}_{\widehat{c}_D^{(\Delta)}}(k)$ for $k = 1, \ldots, q$.

Simulations

To simulate α-stable CARMA processes, we also use an Euler-Maruyama scheme for differential equations with initial value $Y_0 = 0$ and step size 0.01. As before, we set $\Delta = 1$ as the distance between the discrete observations and $\alpha = 1.5$ for the stable index of the driving symmetric α-stable Lévy process. Again, we investigate the behavior of the

adjusted Whittle estimator and the estimator of García et al. (2011) for $n = 500, 2000, 5000$ based on 500 replications.
As a first example, we simulate an Ornstein-Uhlenbeck process with $\vartheta_0 = -1$. The resulting sample mean, bias and sample standard deviation are given in Table 6.13. It seems that both the adjusted Whittle estimator and the estimator of García et al. (2011) converge to the true value. For the adjusted Whittle estimator this is consistent with Theorem 4.15. To compare the behavior in the heavy-tailed setting with the behavior in the light-tailed setting, we present the simulation results of a study where we use for the driving Lévy process of the Ornstein-Uhlenbeck model a Brownian motion. The results are given in Table 6.14. As we can see, the behavior of the Whittle estimator and the behavior of the estimator of García et al. (2011) are similar for the light-tailed and for the heavy-tailed Ornstein-Uhlenbeck process.

Next, we simulate the CARMA(2,0) process of Example 4.17. Accordingly, the true value is $\vartheta_0 = -3$. The results are given in Table 6.15. As already argued in Example 4.17 the adjusted Whittle estimator is not a consistent estimator in this situation. This is confirmed by the simulation study. For $n = 5000$ the bias and standard deviation are even higher than for $n = 2000$. The estimator of García et al. (2011) behaves even worse. On the one hand, the bias and standard deviation of García et al. (2011) are quite high and not decreasing with increasing sample size. On the other hand, the estimation procedure of García et al. (2011) stops for every sample size for more than 1/5th of the replications. This can be traced back to an inadequate estimate of the zero of the AR polynomial, namely the real part of the estimated zero of the AR polynomial is less than 0 which means that the logarithm of this zero is not defined.

Finally, we investigate the CARMA(2,1) process of Example 4.18, see Table 6.16. Our simulation results show the same findings as García et al. (2011); both estimators perform very well in this parameter setting. However, most of the time there is one parameter which has a slightly higher bias or standard deviation such that it is not apparent if the estimator is converging. Indeed, for the adjusted Whittle estimator we already showed in Example 4.18 that this is not the case and we guess that the same holds true for the estimator of García et al. (2011), although at the first view this seems to contradict the simulation study. But from the behavior of $\beta_{\vartheta,\vartheta_0}$ in Figure 4.2 we know that only in a small neighborhood of ϑ_0, the random variables $W^{(\alpha)}(\vartheta) - W^{(\alpha)}(\vartheta_0)$ are not positive and outside this neighborhood they are positive with probability one because $\beta_{\vartheta,\vartheta_0} = 1$. Although $W^{(\alpha)}(\vartheta)$ has not a unique minimum in ϑ_0, ϑ_0 is close to the minimum of $W^{(\alpha)}(\vartheta)$. Thus, the Whittle estimator is close to the true value ϑ_0 as well.

MCARMA(2,1)						
normal	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
-1	-0.9592	0.0408	0.0978	-1.0034	0.0034	0.1900
-2	-1.9436	0.0564	0.1411	-2.0007	0.0007	0.2164
1	1.0649	0.0649	0.2217	1.0370	0.0370	0.2014
-2	-1.8407	0.1593	0.2411	-1.9200	0.0800	0.2568
-3	-2.9698	0.0302	0.1601	-3.0540	0.0540	0.3301
1	1.1332	0.1332	0.2357	0.6574	0.3426	0.4765
2	2.0029	0.0029	0.1098	1.9746	0.0254	0.2461
0.4751	0.4826	0.0075	0.0889	0.4521	0.0230	0.2248
-0.1622	-0.1619	0.0003	0.0565	-0.1623	0.0001	0.1416
0.3708	0.3583	0.0125	0.0467	0.3606	0.0102	0.0963
NIG	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
-1	-0.9729	0.0271	0.1909	-0.9684	0.0316	0.2236
-2	-1.7647	0.2353	0.2479	-1.6184	0.3816	0.3417
1	0.9311	0.0689	0.1804	1.0670	0.0670	0.4061
-2	-1.7890	0.2110	0.2684	-1.8875	0.1125	0.3455
-3	-3.0537	0.0537	0.1427	-3.1586	0.1586	0.4754
1	0.9512	0.0488	0.1641	0.6770	0.3230	0.8610
2	1.8434	0.1566	0.1475	1.5463	0.4537	0.8284
0.4751	0.4389	0.0362	0.1496	0.6328	0.1577	0.2349
-0.1622	-0.1220	0.0402	0.0504	0.0055	0.1677	0.1296
0.3708	0.2970	0.0738	0.0724	0.3061	0.0647	0.1644
MCAR(1)						
normal	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
1	0.9880	0.0120	0.1237	0.9939	0.0061	0.1162
-2	-2.0270	0.0270	0.1208	-2.0064	0.0064	0.1155
3	3.0031	0.0031	0.1394	2.9957	0.0043	0.1384
-4	-4.0631	0.0631	0.1872	-4.0058	0.0058	0.1887
0.7513	0.7261	0.0252	0.1664	0.7329	0.0184	0.2106
-0.3536	-0.3179	0.0357	0.1425	-0.3175	0.0361	0.1522
0.3536	0.3987	0.0451	0.1426	0.3883	0.0347	0.1654
NIG	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
1	0.9916	0.0084	0.1099	0.9782	0.0218	0.1301
-2	-2.0266	0.0266	0.1010	-1.9847	0.0153	0.1424
3	-2.9731	0.0269	0.1089	2.9690	0.0310	0.1797
-4	-3.9993	0.0007	0.2148	-3.9971	0.0029	0.2368
0.7513	0.4721	0.2792	0.2222	0.1926	0.5587	0.1597
-0.3536	-0.1976	0.1560	0.1349	0.0443	0.3979	0.1424
0.3536	0.4588	0.1052	0.1431	0.2195	0.1341	0.1383

Table 6.7.: Estimation results for bivariate MCARMA(2,1) and MCAR(1) processes with parameter $\vartheta_0^{(1)}$ and $\vartheta_0^{(2)}$, respectively. The sample size is $n_4 = 50$ in all simulations.

Brownian motion driven, $n_2 = 2000$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
-0.01	-0.0099	0.0001	0.0005	-0.0103	0.0003	0
0	0	0	0	0	0	0.1891
7	6.9245	0.0755	0.0853	7	0	0.0012
-1	-1.0442	0.0442	0.1915	-1	0	0.0019
0.7513	0.8574	0.1061	0.2193	0.7513	0	0.0031
-0.3536	-0.3492	0.0044	0.0587	-0.3535	0.0001	0.0013
0.3536	0.7958	0.4422	0.4160	0.3536	0	0.0005
NIG driven, $n_2 = 2000$						
	Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.
-0.01	-0.0125	0.0025	0.0534	-0.099	0.0001	0.0001
0	-0.0084	0.0084	0.0507	0	0	0.1805
7	7.0137	0.0137	0.1081	7	0	0.0180
-1	-0.8731	0.1269	0.1354	-1	0	0.0049
0.7513	1.4557	0.7045	0.0959	0.7513	0	0.0027
-0.3536	0.1189	0.4724	0.1675	-0.3536	0	0.0017
0.3536	0.7397	0.3862	0.0524	0.3535	0.0001	0.0008

Table 6.8.: Estimation results for a bivariate MCAR(1) process with parameter $\vartheta_0^{(3)}$ close to the non-stationary case.

$n_1 = 500$									
	Whittle			adjusted Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-2	-2.0951	0.0951	0.7766	3.1063	1.1063	3.4195	-2.0880	0.0880	0.7628
-2	-2.0482	0.0482	0.6500	-2.9233	0.9233	2.9957	-2.0449	0.0449	0.5889
-1	-0.9731	0.0269	0.1186	-0.9028	0.0972	0.3710	-0.9729	0.0271	0.1779
$n_2 = 2000$									
	Whittle			adjusted Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-2	-2.0204	0.0204	0.0755	-2.0816	0.0816	1.0399	-2.0015	0.0015	0.1926
-2	-1.9975	0.0025	0.0637	-2.0732	0.0732	0.9199	-1.9948	0.0052	0.1466
-1	-0.9933	0.0067	0.0547	-0.9965	0.0035	0.1267	-0.9993	0.0007	0.0674
$n_3 = 5000$									
	Whittle			adjusted Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-2	-2.0046	0.0046	0.0117	-1.9854	0.0146	0.0860	-2.0068	0.0068	0.0997
-2	-1.9914	0.0086	0.0149	-1.9840	0.0160	0.0821	-1.9942	0.0058	0.0772
-1	-1.0004	0.0004	0.0153	-1.0070	0.0070	0.0488	-1.0009	0.0009	0.0408

Table 6.9.: Estimation results for a Brownian motion driven CARMA(2,1) process with parameter $\vartheta_0^{(4)}$.

$n_1 = 500$									
	Whittle			adjusted Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-2	-2.3278	0.3278	1.7598	-3.0174	1.0174	3.2090	-2.3175	0.3175	1.0862
-2	-2.2612	0.2612	1.4892	-2.8550	0.8550	2.8684	-2.2047	0.2047	0.8023
-1	-0.9855	0.0145	0.1652	-0.9445	0.0555	0.3376	-0.9243	0.0757	0.2938
$n_2 = 2000$									
	Whittle			adjusted Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-2	-2.0261	0.0261	0.1038	-1.9996	0.0004	0.5351	-2.0122	0.0122	0.2526
-2	-1.9977	0.0023	0.0784	-1.9988	0.0012	0.4552	-2.0034	0.0034	0.1845
-1	-0.9968	0.0032	0.0607	-1.0153	0.0153	0.0961	-1.0037	0.0037	0.0848
$n_3 = 5000$									
	Whittle			adjusted Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-2	-2.0138	0.0138	0.0575	-1.9842	0.0158	0.0902	-1.9938	0.0062	0.1093
-2	-1.9948	0.0052	0.0466	-1.9866	0.0134	0.0825	-1.9917	0.0083	0.0906
-1	-0.9991	0.0009	0.0339	-1.0097	0.0097	0.0508	-1.0059	0.0059	0.0415

Table 6.10.: Estimation results for a NIG driven CARMA(2,1) process with parameter $\vartheta_0^{(4)}$.

$n_1 = 500$									
	Whittle			adjusted Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-6	-5.9230	0.0770	0.2074	-6.2266	0.2266	0.6347	-6.4357	0.4357	1.3266
-11	-10.839	0.1610	0.4119	-11.276	0.2759	0.9351	-11.607	0.6067	1.6706
-6	-5.8267	0.1733	0.3585	-6.0575	0.0575	0.4800	-6.3039	0.3039	1.2821
$n_2 = 2000$									
	Whittle			adjusted Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-6	-5.9886	0.0114	0.1117	-6.0410	0.0410	0.2391	-6.0549	0.0549	0.4510
-11	-10.934	0.0664	0.2372	-11.068	0.0680	0.4126	-11.042	0.0422	0.6005
-6	-5.8855	0.1145	0.1755	-5.9460	0.0540	0.1924	-5.9542	0.0458	0.4464
$n_3 = 5000$									
	Whittle			adjusted Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-6	-5.9856	0.0144	0.0884	-6.0455	0.0455	0.1444	-5.9861	0.0139	0.1120
-11	-10.934	0.0665	0.1471	-11.0349	0.0349	0.1298	-10.926	0.0741	0.1877
-6	-5.9123	0.0877	0.1262	-5.9303	0.0697	0.1104	-5.8937	0.1063	0.1406

Table 6.11.: Estimation results for a Brownian motion driven CAR(3) process with $\vartheta_0^{(5)}$.

$n_1 = 500$									
	Whittle			adjusted Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-6	-5.9449	0.0551	0.4322	-5.9238	0.0762	0.4799	-6.8247	0.8247	1.9413
-11	-10.922	0.0778	0.5765	-10.905	0.0951	0.6813	-12.186	1.1860	2.3377
-6	-5.8492	0.1508	0.3455	-5.8000	0.2000	0.4239	-6.6137	0.6137	1.6559
$n_2 = 2000$									
	Whittle			adjusted Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-6	-5.9611	0.0389	0.1287	-6.0737	0.0737	0.3438	-6.01035	0.1035	0.6401
-11	-10.901	0.0989	0.2590	-11.050	0.0504	0.4832	-11.105	0.1053	0.8271
-6	-5.8879	0.1121	0.1988	-5.9692	0.0308	0.2175	-6.0036	0.0036	0.5522
$n_3 = 5000$									
	Whittle			adjusted Whittle			QMLE		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-6	-6.0313	0.0313	0.0825	-6.0622	0.0622	0.1883	-6.0087	0.0087	0.2748
-11	-10.888	0.1118	0.1274	-11.035	0.0345	0.1490	-10.954	0.0459	0.3830
-6	-5.9110	0.0190	0.0885	-5.8438	0.1562	0.2144	-5.9164	0.0836	0.2513

Table 6.12.: Estimation results for a NIG driven CAR(3) process with parameter $\vartheta_0^{(5)}$.

adjusted Whittle, $\alpha = 1.5$									
	$n = 500$			$n = 2000$			$n = 5000$		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-1	-1.0132	0.0132	0.1118	-1.0082	0.0082	0.0528	-1.0071	0.0071	0.0367
Estimator of García et al., $\alpha = 1.5$									
	$n = 500$			$n = 2000$			$n = 5000$		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-1	-1.0162	0.0162	0.1018	-0.9948	0.0052	0.0522	-0.9942	0.0058	0.0333

Table 6.13.: Estimation results for a symmetric 1.5-stable Ornstein-Uhlenbeck process with parameter $\vartheta_0 = -1$.

adjusted Whittle, $\alpha = 2$									
	$n = 500$			$n = 2000$			$n = 5000$		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-1	-1.0143	0.0143	0.1183	-1.0082	0.0082	0.0528	-1.0002	0.0002	0.0349
Estimator of García et al., $\alpha = 2$									
	$n = 500$			$n = 2000$			$n = 5000$		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-1	-1.0007	0.0007	0.1133	-1.0011	0.0011	0.0568	-1.0012	0.0012	0.0351

Table 6.14.: Estimation results for a Brownian motion driven Ornstein-Uhlenbeck process with parameter $\vartheta_0 = -1$.

adjusted Whittle, $\alpha = 1.5$									
	$n = 500$			$n = 2000$			$n = 5000$		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-3	-3.4762	0.4762	1.2741	-3.2902	0.2902	0.9367	-3.3002	0.3002	0.9568
Estimator of García et al., $\alpha = 1.5$									
	$n = 500$			$n = 2000$			$n = 5000$		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
-3	-3.2473	0.2473	1.2220	-3.8184	0.8164	1.1089	-4.0770	1.0770	0.9238

Table 6.15.: Estimation results for the symmetric 1.5-stable CARMA(2,0) process of Example 4.17.

adjusted Whittle, $\alpha = 1.5$									
	$n = 500$			$n = 2000$			$n = 5000$		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
1.9647	1.9520	0.0127	0.0516	1.9592	0.0055	0.0321	2.0069	0.0422	1.1890
0.0893	0.1031	0.0138	0.0377	0.0940	0.0047	0.0224	0.0987	0.0094	0.0288
0.1761	-0.0144	0.1905	0.1836	-0.0389	0.215	0.1681	0.1735	0.0026	0.0224
Estimator of García et al., $\alpha = 1.5$									
	$n = 500$			$n = 2000$			$n = 5000$		
ϑ_0	mean	bias	std.	mean	bias	std.	mean	bias	std.
1.9647	2.0947	0.1300	0.4480	2.0138	0.0491	0.2405	2.0036	0.0389	0.1543
0.0893	0.1462	0.0569	0.2160	0.0939	0.0046	0.0323	0.0930	0.0037	0.0300
0.1761	0.2196	0.0435	0.1333	0.1877	0.0116	0.0487	0.1920	0.0159	0.0484

Table 6.16.: Estimation results for the symmetric 1.5-stable CARMA(2,1) process of Example 4.18.

6.2. Goodness-of-fit tests

Our final simulation study has two major purposes. As before, we want to find out if the theoretical results can be observed for finite sample sizes. In view of general testing procedures, it would be desirable if the quantiles of the test statistics are similar to the quantiles of the limit processes for small or moderate sample sizes. Therefore, we first investigate how the empirical and limit quantiles of the spectral goodness-of-fit test statistics behave. Subsequently, we use the quantiles of the limit process to construct some tests. These will then be applied in different scenarios. In the following, we focus on the Grenander-Rosenblatt and the Cramér-von Mises statistic. As before, we start with investigating their behavior in the case of a univariate CARMA(2,1) process defined by

$$dX_t = AX_t dt + BdL_t \quad \text{and} \quad Y_t = CX_t, \quad t \geq 0,$$

with

$$A = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 \end{pmatrix}, \qquad \Sigma_L = 1.$$

and a bivariate Ornstein-Uhlenbeck process with

$$A = \begin{pmatrix} -1 & -0.5 \\ 1 & -1 \end{pmatrix} = B, \qquad C = I_2 = \Sigma_L.$$

As in all the preceding simulations, the processes are simulated with an Euler-Maruyama scheme with initial values $Y(0) = X(0) = 0$, step size 0.01 and observation distance $\Delta = 1$. Again, we take the Brownian motion and the normal-inverse Gaussian process as driving processes. The estimation results for finite sample size are based on 5000 replicates each, whereas the estimation results corresponding to the limit process are based on 10.000 replicates. For each setting, we compute the empirical quantiles to $\alpha = 0.9, 0.95, 0.975$ and $\alpha = 0.99$. The results of the CARMA(2,1) settings can be found in Table 6.17, those of the MCAR(1) settings in Table 6.18. Furthermore, for the Brownian motion driven CARMA(2,1) process, four sample paths of the Grenander-Rosenblatt and Cramér-von Mises statistic with added 95% quantile are made, see Figure 6.1 and Figure 6.2.

As we can see, the quantiles of the test statistics are really similar to those of the limit process even for small sample sizes in all investigated settings. We now consider the hypothesis that the data originates from a process with the spectral density which is plugged in in the test statistic. Under the hypothesis, it is to expect that the corresponding level α-test behaves as desired. Therefore, we do some testing to the 5% level. For both underlying processes and both test statistics, we investigate the behavior under the hypothesis. In the CARMA(2,1) setting, we also consider the test statistics when the data is generated by various CARMA(2,1) processes but the spectral density remains the same.

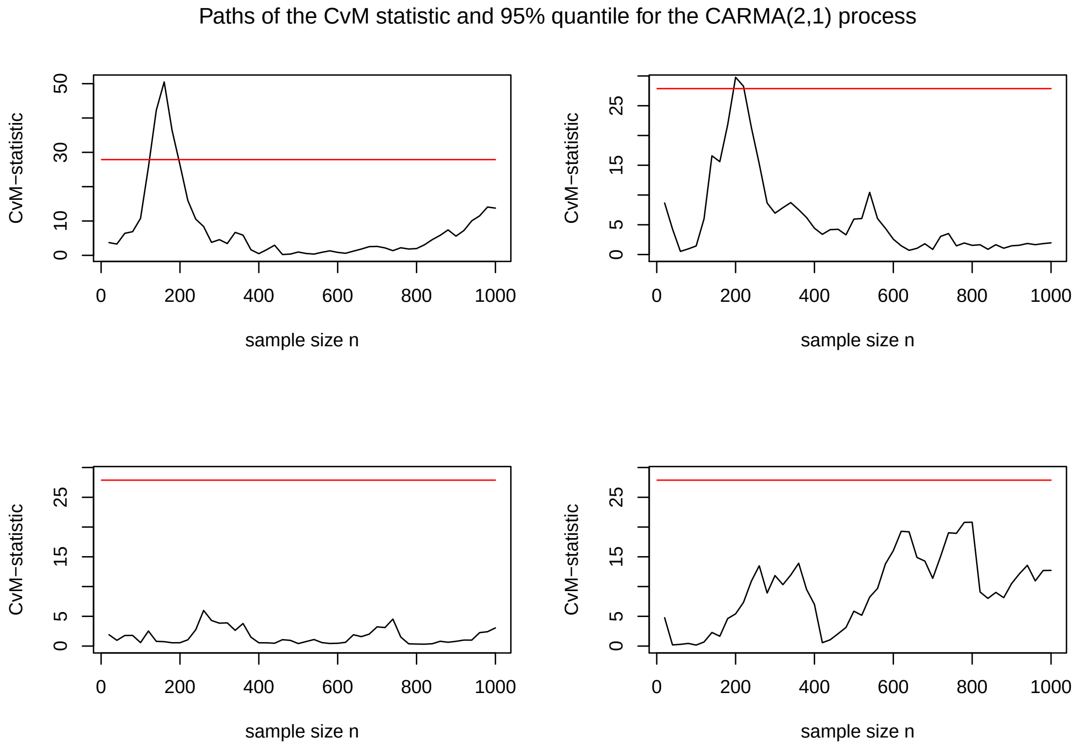

Figure 6.1.: Paths of the Cramér-von Mises statistic with underlying CARMA(2,1) process

Namely, we consider the parametrization of before, i.e.

$$A = \begin{pmatrix} 0 & 1 \\ \vartheta_1 & \vartheta_2 \end{pmatrix}, \qquad B = \begin{pmatrix} \vartheta_3 \\ \vartheta_1 + \vartheta_2 \vartheta_3 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \Sigma_L = 1$$

and choose as alternative generating processes the processes defined by the parameters

$$(C1)\ \vartheta = (-1, -2, 1), \qquad (C2)\ \vartheta = (-2, -3, 5), \qquad (C3)\ \vartheta = (-1, -2, -3),$$
$$(C4)\ \vartheta = (-2, -1, 2), \qquad (C5)\ \vartheta = (-0.5, -0.5, 1), \qquad (C6)\ \vartheta = (-0.5, -1, -1).$$

In the same way, we consider the parametrization

$$A = \begin{pmatrix} \vartheta_1 & \vartheta_2 \\ \vartheta_3 & \vartheta_4 \end{pmatrix} = B, \qquad C = I_2 = \Sigma_L$$

in the MCAR(1) setting and take the parameters

$$(M1)\ \vartheta = (-1, -0.5, 0.5, -1), \qquad (M2)\ \vartheta = (-1, 0, 0, -1),$$
$$(M3)\ \vartheta = (-1, -2, 1, -1), \qquad (M4)\ \vartheta = (-1, -0.5, 1, -2)$$

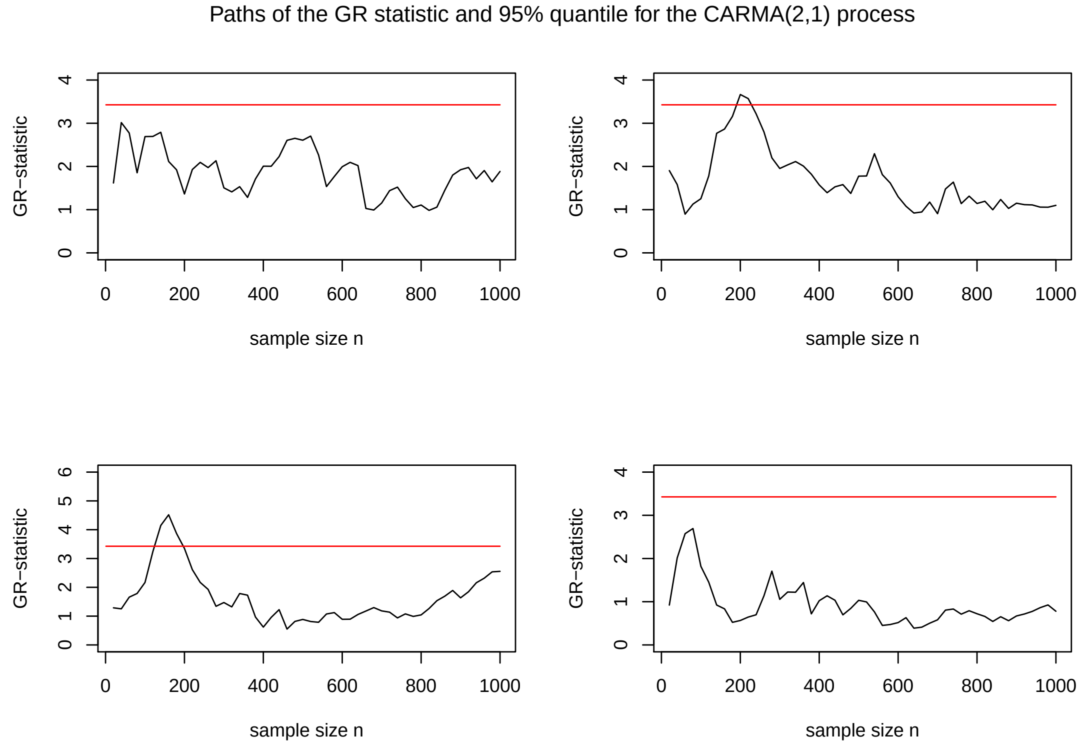

Figure 6.2.: Paths of the Grenander-Rosenblatt statistic with underlying CARMA(2,1) process

to generate data under different alternatives. The results are in Table 6.19 and Table 6.20. As suspected, in the correct specified setting, the statistics hold the given level for most sample sizes. Under the alternatives, the statistics reject quite often for moderate sample sizes and detect every alternative with certainty for $n = 1000$ and higher. Based on the simulations, it is not possible to say if the Cramér-von Mises or the Grenander-Rosenblatt statistic is favorable since their performances are qualitatively similar. Finally, since the limit statistics depend on the true parameter, we also do a bootstrap test in the CARMA(2,1) settings. More precisely, given a fixed sample, we estimate the parameters A, B and C of the state space model with the Whittle estimation procedure of Chapter 3. The estimates $\widehat{A}, \widehat{B}, \widehat{C}$ are then used to generate a sample $\widehat{N}_1^{(\Delta)}, \ldots, \widehat{N}_{500}^{(\Delta)}$ of the estimated white noise process. Eventually, based on 1000 samples we estimate the quantiles of the limit process where W_i, $i \in \mathbb{N}_0$ replaced by $\widehat{W}_i$, $i \in \mathbb{N}_0$ with

$$\operatorname{vec}\left(\widehat{W}_k\right) \sim \mathcal{N}\left(0, \frac{1}{n}\sum_{j=1}^{n} \widehat{N}_j^{(\Delta)} \widehat{N}_j^{(\Delta)\top} \otimes \frac{1}{n}\sum_{j=1}^{n} \widehat{N}_j^{(\Delta)} \widehat{N}_j^{(\Delta)\top}\right), \quad k \in \mathbb{N},$$

CARMA(2,1) process								
Grenander-Rosenblatt statistic								
	Normal distribution				NIG distribution			
n	90%	95%	97.5%	99%	90%	95%	97.5%	99%
50	2.8654	3.3713	3.9420	5.1020	2.7852	3.2047	3.5637	4.3470
100	2.9009	3.3557	3.9185	4.5609	2.8851	3.4118	3.9302	4.6944
200	2.9459	3.4778	3.8956	4.5802	3.0007	3.5462	4.0324	4.6067
500	2.9825	3.5225	3.9959	4.5224	3.0478	3.5056	3.9494	4.4670
1000	2.9753	3.4740	3.9439	4.6561	3.0240	3.5188	3.9614	4.4840
2500	2.9986	3.5068	4.0009	4.5487	3.0253	3.6052	4.0868	4.6110
limit	2.9040	3.4263	3.9305	4.5133	2.9455	3.4609	3.9637	4.5547
Cramér-von Mises statistic								
	Normal distribution				NIG distribution			
n	90%	95%	97.5%	99%	90%	95%	97.5%	99%
50	20.5368	27.9626	38.5060	60.4273	18.7985	24.8263	31.4444	45.8557
100	19.4151	26.8832	37.4596	49.8628	19.4187	27.5602	37.9928	54.0262
200	19.9334	28.5887	36.2754	49.8312	20.7667	29.3701	39.2179	52.0016
500	20.2071	29.3122	38.6322	49.8635	21.1896	29.0540	38.1886	49.7864
1000	20.0263	28.3322	37.3100	51.0001	20.8816	28.6483	37.5922	52.7372
2500	20.3190	27.8820	38.3735	51.2512	21.3228	30.5439	40.3677	52.6218
limit	19.7667	27.8781	37.7293	48.4942	20.6293	28.9460	37.9811	51.1914

Table 6.17.: Empirical quantiles of the Grenander-Rosenblatt and the Cramér-von Mises statistics for the CARMA$(2,1)$ process. The estimation results to the limit process are denoted as “limit”.

and

$$\operatorname{vec}\left(\widehat{W}_0\right) \sim \mathcal{N}\left(0, \frac{1}{n}\sum_{j=1}^{n}\left(\widehat{N}_j^{(\Delta)}\widehat{N}_j^{(\Delta)\top} \otimes \widehat{N}_j^{(\Delta)}\widehat{N}_j^{(\Delta)\top}\right) - \frac{1}{n^2}\sum_{\substack{j=1\\k=1}}^{n}\left(\widehat{N}_j^{(\Delta)}\widehat{N}_j^{(\Delta)\top}\right)\otimes\left(\widehat{N}_k^{(\Delta)}\widehat{N}_k^{(\Delta)\top}\right)\right).$$

Since the procedure is computationally expensive, we only use the sample sizes $n = 50, 100, 200$ and do 500 replicates in each setting. The findings are in Table 6.21. It is remarkable, that under the hypothesis, the test hardly rejects. Furthermore, under the alternatives the bootstrap procedure behaves even better than the original test. This is an unexpected but pleasant finding.

MCAR(1) process								
Grenander-Rosenblatt statistic								
	Normal distribution				NIG distribution			
n	90%	95%	97.5%	99%	90%	95%	97.5%	99%
50	3.0535	3.5157	3.8073	4.5385	3.0735	3.6421	4.3578	5.18897
100	3.0929	3.6552	4.1876	4.6781	3.1855	3.7608	4.2405	4.8650
200	3.1211	3.6824	4.2016	4.9505	3.2628	3.8365	4.3261	5.1105
500	3.1439	3.7339	4.1935	4.7652	3.1961	3.7491	4.2584	4.9095
1000	3.1946	3.7512	4.2244	4.8877	3.3017	3.8345	4.4087	5.1175
2500	3.1256	3.7995	4.2733	4.8794	3.3735	3.9468	4.4847	5.0303
limit	3.1593	3.7113	4.2966	4.8712	3.2938	3.9441	4.4129	5.0938
Cramér-von Mises statistic								
	Normal distribution				NIG distribution			
n	90%	95%	97.5%	99%	90%	95%	97.5%	99%
50	20.2115	27.0517	31.3571	41.4468	21.4855	30.8772	45.0900	64.1383
100	21.6175	30.8035	40.8099	51.0412	22.7657	31.8846	42.3185	55.4154
200	21.5474	30.4375	41.0050	57.3692	23.6207	33.0851	44.0739	60.8951
500	21.8753	30.9503	39.9908	52.7037	22.3696	31.9319	41.7223	56.2380
1000	22.3494	31.5963	41.2221	56.5119	23.7143	33.2557	44.7966	61.1908
2500	22.8001	32.1742	41.7642	55.0025	24.1312	33.9940	44.2733	60.1516
limit	22.6324	32.1829	42.8200	56.2604	23.4122	33.6122	43.9967	59.7828

Table 6.18.: Empirical quantiles of the Grenander-Rosenblatt and the Cramér-von Mises statistics for the MCAR(1) process. The estimation results to the limit process are denoted as “limit”.

CARMA(2,1) process							
Grenander-Rosenblatt statistic							
Normal distribution							
n	T	(C1)	(C2)	(C3)	(C4)	(C5)	(C6)
50	3.34	50.40	100	100	100	40.98	5.50
100	3.16	94.4	100	100	100	66.16	20.02
200	3.32	100	100	100	100	90.16	48.12
500	3.82	100	100	100	100	99.90	92.36
1000	3.50	100	100	100	100	100	100
2500	5.50	100	100	100	100	100	100
NIG distribution							
50	2.88	39.36	100	100	100	37.86	3.22
100	2.48	91.26	100	100	100	63.86	14.14
200	2.92	100	100	100	100	89.38	41.20
500	2.54	100	100	100	100	99.92	88.96
1000	2.54	100	100	100	100	100	100
2500	5.60	100	100	100	100	100	100
Cramér-von Mises statistic							
Normal distribution							
n	T	(C1)	(C2)	(C3)	(C4)	(C5)	(C6)
50	4.78	65.12	100	100	100	43.90	11.02
100	4.98	97.02	100	100	100	68.38	27.30
200	5.20	100	100	100	100	90.62	55.62
500	5.60	100	100	100	100	100	94.52
1000	5.30	100	100	100	100	100	100
2500	5.68	100	100	100	100	100	100
NIG distribution							
50	3.78	63.70	100	100	100	41.88	9.64
100	4.20	96.46	100	100	100	67.62	25.08
200	5.30	100	100	100	100	90.68	54.22
500	5.04	100	100	100	100	99.90	93.80
1000	4.90	100	100	100	100	100	100
2500	5.50	100	100	100	100	100	100

Table 6.19.: Percentages of rejection for the test statistics in the CARMA(2,1) settings based on 5000 replications and the significance level $\alpha = 0.05$. Thereby, “T” stands for the correct specified model whereas “(C1)-(C6)” denote the misspecifications in the CARMA(2,1) setting.

MCAR(1) process										
Grenander-Rosenblatt statistic										
	Normal distribution					NIG distribution				
n	T	(M1)	(M2)	(M3)	(M4)	T	(M1)	(M2)	(M3)	(M4)
50	6.02	23.76	73.94	100	24.78	2.92	13.31	49.68	100	14.44
100	6.86	45.84	97.30	100	42.42	3.92	31.64	88.90	100	29.88
200	6.80	76.40	100	100	70.46	4.30	64.64	100	100	60.78
500	7.94	99.10	100	100	97.56	3.78	97.94	100	100	98.60
1000	7.72	100	100	100	100	4.46	100	100	100	100
2500	5.65	100	100	100	100	5.06	100	100	100	100
Cramér-von Mises statistic										
	Normal distribution					NIG distribution				
n	T	(M1)	(M2)	(M3)	(M4)	T	(M1)	(M2)	(M3)	(M4)
50	3.64	17.98	67.68	100	21.34	4.68	14.74	55.60	100	16.26
100	4.20	39.60	96.08	100	38.12	4.90	32.74	80.98	100	31.38
200	4.50	71.40	100	100	66.34	5.22	63.88	100	100	55.52
500	4.56	98.66	100	100	96.76	4.94	97.40	100	100	92.66
1000	4.88	100	100	100	100	5.40	100	100	100	100
2500	5.00	100	100	100	100	5.40	100	100	100	100

Table 6.20.: Percentages of rejection for the test statistics in the MCAR(1) settings based on 5000 replications and the significance level $\alpha = 0.05$. Thereby, “T” stands for the correct specified model whereas “(M1)-(M4)” denote the misspecifications in the CARMA(2,1) setting.

CARMA(2,1) process						
Grenander-Rosenblatt						
	Normal distribution			NIG distribution		
n	T	(C1)	(C2)	T	(C1)	(C2)
50	0	81.60	90.20	0.20	99.60	56.60
100	2.40	100	99.80	0	100	100
200	0	100	100	0	100	100
Cramér-von Mises						
	Normal distribution			NIG distribution		
n	T	(C1)	(C2)	T	(C1)	(C2)
50	0.80	77.60	92.8	0	88.40	40.40
100	2.0	100	100	0.40	100	96.8
200	2.0	100	100	1.20	100	100

Table 6.21.: Percentages of rejection for the Bootstrap version of the test statistics based on 500 replications and the significance level $\alpha = 0.05$. Thereby, “T” stands for the correct specified model whereas “(C1)-(C2)” denote the misspecifications in the CARMA(2,1) setting.

CHAPTER 7

Conclusion and outlook

The main focus of this thesis was to investigate estimation procedures for discretely observed Lévy-driven causal MCARMA processes. We achieved to prove that the Whittle estimator has desirable asymptotic properties under weak identifiability conditions and under the assumption that the driving process has existing fourth moments. In contrast, the quasi maximum likelihood estimator which might be seen as the only yet investigated alternative, needs a slightly stronger moment condition to work. A further advantage of the Whittle estimator is that the covariance matrix of the limit distribution has a known analytical representation and can be computed in theory. This gives the opportunity to build confidence bands which might be used for different reasons. For example, testing could be done. Not only are the performances of the Whittle estimator and the quasi maximum likelihood estimator really similar for finite sample sizes, they are also quite good for small and moderate sample sizes. Consequently, another useful estimation procedure for equidistantly observed light-tailed Lévy-driven MCARMA processes was found.
For univariate processes, we then adapted the estimator to obtain a procedure which is independent of the variance of the driving process. Motivated by Mikosch (1991) who proved that a Whittle estimator depicts desirable asymptotic properties in a heavy-tailed ARMA setting, we had the hope to thereby construct an estimator which is suited for parameter estimation in a symmetric α-stable CARMA setting. Unfortunately, this did not turn out to be true. Even though we were able to show that the adjusted Whittle estimator is also strongly consistent and asymptotically normally distributed in a setting in which the fourth moment of the driving process exists, the results can not be carried out to a general stable setting. Just for the class of symmetric α-stable Ornstein-Uhlenbeck processes, the estimator is consistent. However, for this setting, there are already some

suitable estimation procedures like the ones presented in Fasen and Fuchs (2013b), Hu and Long (2007), Hu and Long (2009), Ljungdahl and Podolskij (2020), Zhang and Zhang (2013). Accordingly, our findings for heavy-tailed parameter estimation are disenchanting. Additionally, considerations in the light-tailed Ornstein-Uhlenbeck setting and the simulation studies suggest that albeit the estimator is asymptotically normally distributed for light-tailed processes, it is inferior to the quasi maximum likelihood estimator and the original Whittle estimator. Consequently, we would not recommend using this estimator at all since the alternatives probably perform better.

Returning to the problem of parameter estimation for α-stable CARMA processes, we conclude that a suitable estimation procedure has to be found in future. Therefore, this problem remains open for further investigations. There are also many more directions for potential research relating to parameter estimation. Considering classes which are related to the class of MCARMA processes, one could investigate the Whittle estimator for sampled cointegrated MCARMA processes and for sampled (causal) CARMA random fields. For both classes, there is only little research done in the context of parameter estimation based on the sampled processes. Similar as in the MCARMA case, solely the quasi maximum likelihood estimator has already been investigated for the sampled cointegrated MCARMA process, see Fasen-Hartmann and Scholz (2019). For the sampled CARMA random fields, a least-squares approach is the only yet investigated estimation procedure, see Klüppelberg and Pham (2019). Furthermore, the Whittle estimator could also be considered in the context of MCARMA processes with light tails which are sampled with high-frequency or when the MCARMA process is sampled irregularly.

In the second part of the thesis, we investigated the normalized function-indexed periodogram for the sampled light-tailed MCARMA process in different settings. Mainly, we derived a functional central limit theorem for a broad class of index functions. A direct application of this result enabled us to obtain the limit behavior of various spectral goodness-of-fit test statistics. In case of a heavy-tailed ARMA process, Klüppelberg and Mikosch (1996) already proved a central limit theorem for the integrated periodogram indexed by a class of indicator functions and concluded the asymptotic behavior of the spectral goodness-of-fit test statistics in a similar manner. Their limit distributions differ from the ones which are obtained in the MCARMA setting. However, for the Grenander-Rosenblatt statistic, our limit distribution corresponds to the one which is obtained when considering Gaussian ARMA processes, see Section 6.2.6 of Priestley (1981). Although we just contemplated the statistics under the assumption of a correct specified setting, we also simulated misspecified processes and did some testing. The corresponding tests are good even for small sample sizes. Since the limit processes depend on the fourth moment of the driving Lévy process, we also implemented a Bootstrap procedure to bypass the necessity of knowing the distribution of this process. In our small simulation study, this procedure depicted a nice performance. However, we did not investigate it mathematically. This might also be a topic for future research. Additionally, it would be of interest to see how an appropriately normalized function-indexed periodogram behaves in a heavy-tailed

(M)CARMA setting, in a cointegrated (M)CARMA setting or in a CARMA random field setting. In particular, it is an open question if our results can be carried out to a heavy-tailed setting at all. Obviously, since the dependency structure in the sampled heavy-tailed setting leads to problems in the case of the Whittle estimator, it is questionable if there is any normalization under which desirable results can be derived in a heavy-tailed setting. Furthermore, our central limit theorem is based on spectral analysis. Therefore, nice results for spectral goodness-of-fit test statistics were obtained, but there are many open topics in the field of testing for MCARMA processes like goodness-of-fit tests in the time domain. The obvious connection between MCARMA processes and the class of (V)ARMA processes makes these problems not only interesting for the sake of theoretical knowledge. We also expect there to be a high interest for future investigations from diverse scientific applications.

APPENDIX A

Analytical foundations

A.1. Fourier analysis

The frequency domain plays a fundamental part in this thesis: the estimators of Chapter 3 and Chapter 4 and the normalized integrated periodogram of Chapter 5 are all based on the distance between the spectral density of the sampled process and its empirical counterpart, the periodogram. Since the spectral density is just the Fourier transform of the autocovariance function, it is obviously helpful to get familiar with the basics of Fourier analysis. Therefore, we now state the Fourier analytical fundamentals which are needed in this thesis. The first property that we prove, can be seen as a foundation of most of our proofs of Chapter 3 and Chapter 4. It also clarifies why the (adjusted) Whittle estimator is based on the frequencies $\{-\frac{\pi(n-1)}{n}, \dots, \pi\}$.

Lemma A.1.
Let $h \in \mathbb{Z}$. Then

$$\frac{1}{2n} \sum_{j=-n+1}^{n} e^{-ih\omega_j} = \mathbb{1}_{\{\exists z \in \mathbb{Z}:\ h=2zn\}}.$$

Proof. If h is an even multiple of n, say $h = 2zn$ for some $z \in \mathbb{Z}$, $e^{-i2\pi z} = 1$ implies

$$\frac{1}{2n} \sum_{j=-n+1}^{n} e^{-ih\omega_j} = \frac{1}{2n} \sum_{j=-n+1}^{n} e^{-i2\pi z} = 1.$$

We prove that the sum vanishes otherwise. Therefore,

$$e^{-\frac{ih\pi j}{n}} = e^{-\frac{ih\pi j}{n} - 2\pi hi} = e^{-\frac{ih\pi(j+2n)}{n}}$$

implies

$$\sum_{j=-n+1}^{n} e^{-\frac{ih\pi j}{n}} = \sum_{j=-n+1}^{0} e^{-\frac{ih\pi j}{n}} + \sum_{j=1}^{n} e^{-\frac{ih\pi j}{n}} = \sum_{j=-n+1}^{0} e^{-\frac{ih\pi(j+2n)}{n}} + \sum_{j=1}^{n} e^{-\frac{ih\pi j}{n}}$$
$$= \sum_{j=1}^{2n} e^{-\frac{ih\pi j}{n}} = \sum_{j=0}^{2n-1} e^{-\frac{ih\pi j}{n}} = \frac{1-\left(e^{-\frac{ih\pi j}{n}}\right)^{2n}}{1-e^{-\frac{ih\pi j}{n}}} = 0.$$

□

We now introduce results which show that an appropriate approximation of the Fourier series exhibit useful convergence properties.

Theorem A.2.
Let $g : [-\pi, \pi] \to \mathbb{C}$ be continuous. Define

$$\widehat{g}_k := \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\omega) e^{-ik\omega} d\omega \quad \text{and} \quad q_M(\omega) = \sum_{|k|\le M} \widehat{g}_k e^{ik\omega}.$$

Suppose that $\sum_{|k|\le n} |\widehat{g}_k|$ converges. Then

$$\sup_{\omega\in[-\pi,\pi]} |q_M(\omega) - g(\omega)| \overset{M\to\infty}{\longrightarrow} 0.$$

Proof. Körner (1989), Theorem 3.1. □

The assumptions of the previous result are quite strong. If we replace the truncated Fourier series by its Cesàro sum, we receive an approximation which exhibits uniform convergence without assuming that the Fourier coefficients are absolute summable. This result is known as Fejérs Theorem. Since we want to approximate a parameterized function in Chapter 3 and Chapter 4, we have to adjust Fejérs Theorem to a setting which allows a dependency on a second parameter.

Theorem A.3 (Fejérs Theorem, adjusted to a parametrized setting)**.**
Let Θ be a compact parameter space and g be a continuous real-valued function on $[-\pi, \pi]\times\Theta$. Then, the Fourier series of g is Cesàro summable at every point of $[-\pi, \pi]$ for any $\vartheta \in \Theta$. Further, define the Fourier coefficients $\widehat{g}_k(\vartheta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(\omega, \vartheta)e^{-ik\omega}d\omega$ and

$$q_M(\omega, \vartheta) = \frac{1}{M}\sum_{j=0}^{M-1}\left(\sum_{|k|\le j} \widehat{g}_k(\vartheta)e^{ik\omega}\right) = \sum_{|k|<M}\left(1-\frac{|k|}{M}\right)\widehat{g}_k(\vartheta)e^{ik\omega}.$$

Then

$$\lim_{M\to\infty} \sup_{\omega\in[-\pi,\pi]} \sup_{\vartheta\in\Theta} |q_M(\omega, \vartheta) - g(\omega, \vartheta)| = 0.$$

Proof. The proof is similar to the proof of Theorem 2.11.1 of Brockwell and Davis (1991).

We first define the Fejér Kernel K_M by

$$K_M(y) = \frac{1}{2\pi M} \sum_{j=0}^{M-1} \sum_{|k| \leq j} e^{-iky}.$$

The kernel has the properties

a) $K_M(y) \geq 0$,

b) $K_M(\cdot)$ has period 2π,

c) $K_M(\cdot)$ is even,

d) $\int_{-\pi}^{\pi} K_M(y)dy = 1$,

e) for each $\delta > 0, \int_{-\delta}^{\delta} K_M(y)dy \to 1$, as $M \to \infty$,

see, e.g., Brockwell and Davis (1991), p. 71. By defining $g(x, \vartheta) = g(x + 2\pi, \vartheta),\ x \in \mathbb{R},\ \vartheta \in \Theta$, we can represent q_M as

$$\begin{aligned} q_M(\omega, \vartheta) &= \frac{1}{M} \sum_{j=0}^{M-1} \left(\sum_{|k| \leq j} \frac{1}{2\pi} \int_{-\pi}^{\pi} g(x, \vartheta) e^{-ikx} dx e^{ik\omega} \right) \\ &= \int_{-\pi}^{\pi} g(x, \vartheta) \left(\frac{1}{2\pi M} \sum_{j=0}^{M-1} \sum_{|k| \leq j} e^{-ik(x-\omega)} \right) dx \\ &= \int_{-\pi}^{\pi} g(\omega - y, \vartheta) K_M(y) dy. \end{aligned}$$

We denote the distance between the Cesàro sum of degree M and g as $\Delta_M(\cdot, \cdot)$. In view of property d), we obtain

$$\begin{aligned} \Delta_M(\omega, \vartheta) &= \left| \int_{-\pi}^{\pi} g(\omega - y, \vartheta) K_M(y) dy - g(\omega, \vartheta) \right| \\ &= \left| \int_{-\pi}^{\pi} \left(g(\omega - y, \vartheta) - g(\omega, \vartheta) \right) K_M(y) dy \right|. \end{aligned}$$

Let $\varepsilon > 0$. Since $[-\pi, \pi] \times \Theta$ is compact and since g is continuous, g is uniformly continuous on $[-\pi, \pi] \times \Theta$. By definition, $g(\omega, \cdot)$ is 2π-periodic for any $\vartheta \in \Theta$. Therefore, g is uniformly continuous on $\mathbb{R} \times \Theta$. Hence, we can find a δ such that

$$\sup_{(\omega, \vartheta) \in \mathbb{R} \times \Theta} |g(\omega - y, \vartheta) - g(\omega, \vartheta)| < \varepsilon, \quad \text{if } |y| < \delta.$$

A decomposition of Δ_M and property d) yield

$$\begin{aligned} \Delta_M(\omega, \vartheta) \leq & \left| \int_{[-\pi, \pi] \setminus [-\delta, \delta]} \left(g(\omega - y, \vartheta) - g(\omega, \vartheta) \right) K_M(y) dy \right| \\ & + \left| \int_{-\delta}^{\delta} \left(g(\omega - y, \vartheta) - g(\omega, \vartheta) \right) K_M(y) dy \right| \end{aligned}$$

$$\leq \sup_{(\omega,\vartheta)\in\mathbb{R}\times\Theta} |g(\omega-y,\vartheta)-g(\omega,\vartheta)| \left(1-\int_{-\delta}^{\delta} K_M(y)dy\right)+\varepsilon$$
$$\overset{M\to\infty}{\longrightarrow} \varepsilon.$$

□

Remark A.4.
If we investigate the Cesàro sum of a Fourier series of a matrix-valued continuous function $g : [-\pi,\pi]\times\Theta \to \mathbb{R}^{N\times N}$ defined by

$$q_M(\omega,\vartheta) := \frac{1}{M}\sum_{j=0}^{M-1}\left(\sum_{|k|\leq j}\widehat{g(\vartheta)}_k e^{ik\omega}\right), \text{ where } \widehat{g(\vartheta)}_k := \frac{1}{2\pi}\int_{-\pi}^{\pi} g(\omega,\vartheta)e^{-ik\omega}d\omega,$$

Fejérs Theorem gives the uniform convergence of each component of q_M to g on $[-\pi,\pi]\times\Theta$. Since g consists of finite components, q_M also converges to g uniformly. Obviously, the same holds true for any matrix-valued continuous 2π periodic function $g : \mathbb{R}\times\Theta \to \mathbb{R}^{N\times N}$ Similarly, we can transfer Theorem A.2 to matrix-valued functions.

Finally, we often use Parsevals equality in various settings. We introduce a multivariate version which obviously includes the original univariate variant. Note that it is essential to use the Frobenius norm and that the result does not hold when we choose an arbitrary norm.

Theorem A.5 (Parsevals equlality for multivariate functions)**.**
Let $g : [-\pi,\pi] \to \mathbb{R}^{j\times k}$ *satisfy* $\int_{-\pi}^{\pi}\|g(\omega)\|^2 d\omega < \infty$. *Define the Fourier coefficients* $\widehat{g}_\ell = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(\omega)e^{-i\ell\omega}d\omega$. *Then,*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\|g(\omega)\|^2 d\omega = \sum_{\ell\in\mathbb{Z}}\|\widehat{g}_\ell\|^2.$$

Proof. The assertion directly follows from the representations

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\|g(\omega)\|^2 d\omega = \sum_{s=1}^{j}\sum_{t=1}^{k}\frac{1}{2\pi}\int_{-\pi}^{\pi}|g(\omega)[s,t]|^2 d\omega,$$
$$\sum_{\ell\in\mathbb{Z}}\|\widehat{g}_\ell\|^2 = \sum_{s=1}^{j}\sum_{t=1}^{k}\sum_{\ell\in\mathbb{Z}}|\widehat{g[s,t]}_\ell|^2,$$

and Parsevals equality for univariate functions, see for example Theorem 3.4.1 of Simon (2015). □

A.2. Rate of convergence of the integral approximation

To prove the uniform convergence of the Whittle function, it is necessary to guarantee that the deterministic part of the Whittle function converges uniformly.

Proposition A.6.
Let Θ be a compact parameter space and let $g : [-\pi, \pi] \times \Theta \to \mathbb{C}$ be differentiable in $\omega \in [-\pi, \pi]$ for all $\vartheta \in \Theta$. Assume further that $\frac{\partial}{\partial \omega} g$ is continuous in $(\omega, \vartheta) \in [-\pi, \pi] \times \Theta$. Then,

$$\sup_{\vartheta \in \Theta} \left| \frac{1}{2n} \sum_{j=-n+1}^{n} g(\omega_j, \vartheta) - \int_{-\pi}^{\pi} g(\omega, \vartheta) d\omega \right| \overset{n\to\infty}{\longrightarrow} 0.$$

Proof. Since the assertion is equivalent to

$$\sup_{\vartheta \in \Theta} \left| \sum_{j=-n+1}^{n} \left(g(\omega_j, \vartheta)(\omega_j - \omega_{j-1}) - \int_{\omega_{j-1}}^{\omega_j} g(\omega, \vartheta) d\omega \right) \right| \overset{n\to\infty}{\longrightarrow} 0,$$

we show

$$\lim_{n\to\infty} \sum_{j=-n+1}^{n} \sup_{\vartheta \in \Theta} \left| g(\omega_j, \vartheta)(\omega_j - \omega_{j-1}) - \int_{\omega_{j-1}}^{\omega_j} g(\omega, \vartheta) d\omega \right| = 0$$

which therefore implies the statement. With an application of the mean value theorem, we obtain for some $\xi_j(\vartheta) \in [\omega_{j-1}, \omega_j]$

$$\begin{aligned} &\sum_{j=-n+1}^{n} \sup_{\vartheta \in \Theta} \left| g(\omega_j, \vartheta)(\omega_j - \omega_{j-1}) - \int_{\omega_{j-1}}^{\omega_j} g(\omega, \vartheta) d\omega \right| \\ &= \sum_{j=-n+1}^{n} \sup_{\vartheta \in \Theta} \left| \frac{\pi}{n} \left(g(\omega_j, \vartheta) - g(\xi_j(\vartheta), \vartheta) \right) \right| \\ &\leq \sup_{j=-n+1,\ldots,n} \sup_{\vartheta \in \Theta} \pi \left| \left(g(\omega_j, \vartheta) - g(\xi_j(\vartheta), \vartheta) \right) \right|. \end{aligned}$$

Again, the mean value theorem yields for some $\widetilde{\xi}_j(\vartheta)$ with $|\widetilde{\xi}_j(\vartheta) - \omega_j| \leq |\xi_j(\vartheta) - \omega_j|$

$$\sup_{\vartheta \in \Theta} \sup_{j=-n+1,\ldots,n} \pi \left| \left(g(\omega_j, \vartheta) - g(\xi_j(\vartheta), \vartheta) \right) \right| \leq \sup_{\vartheta \in \Theta} \sup_{j=-n+1,\ldots,n} \frac{\pi^2}{n} \left| \frac{\partial}{\partial \omega} g(\widetilde{\xi}_j(\vartheta), \vartheta) \right|.$$

The compactness of $[-\pi, \pi] \times \Theta$ and the continuity of $\frac{\partial}{\partial \omega} g$ complete the proof. □

Since the main results of Chapter 3 and Chapter 4 make use of the normalizing factor $\sqrt{n}$, it is necessary to prove that any investigated sum which converges to an appropriate integral also converges when multiplied with $\sqrt{n}$.

Lemma A.7.
Let $g : [-\pi, \pi] \to \mathbb{C}$ be continuously differentiable. Then,

$$\frac{1}{\sqrt{n}} \sum_{j=-n+1}^{n} g(\omega_j) - \frac{\sqrt{n}}{\pi} \int_{-\pi}^{\pi} g(\omega) d\omega \overset{n\to\infty}{\longrightarrow} 0$$

holds.

Proof. In view of the equivalences

$$\begin{aligned}
&\sum_{j=-n+1}^{n} g(\omega_j)(\omega_j - \omega_{j-1}) \overset{n\to\infty}{\longrightarrow} \int_{-\pi}^{\pi} g(\omega)d\omega \\
\iff &\sum_{j=-n+1}^{n} g(\omega_j)(\omega_j - \omega_{j-1}) - \sum_{j=-n+1}^{n} \int_{\omega_{j-1}}^{\omega_j} g(\omega)d\omega \overset{n\to\infty}{\longrightarrow} 0 \\
\iff &\sum_{j=-n+1}^{n} (g(\omega_j) - g(\xi_j)) \frac{\pi}{n} \overset{n\to\infty}{\longrightarrow} 0,
\end{aligned}$$

for some $\xi_j \in [\omega_{j-1}, \omega_j]$, $j = -n+1, \ldots, n$, it suffices to show

$$\sum_{j=-n+1}^{n} (g(\omega_j) - g(\xi_j)) \frac{\pi}{\sqrt{n}} \overset{n\to\infty}{\longrightarrow} 0$$

for any $\xi_j \in [\omega_{j-1}, \omega_j]$, $j = -n+1, \ldots, n$. Thus, since g is continuously differentiable

$$\left| \sum_{j=-n+1}^{n} (g(\omega_j) - g(\xi_j)) \frac{\pi}{\sqrt{n}} \right| \leq \sum_{j=-n+1}^{n} \mathfrak{C}|\omega_j - \xi_j| \frac{\pi}{\sqrt{n}} \leq \mathfrak{C}\pi^2 \sum_{j=-n+1}^{n} \frac{1}{n^{3/2}} \overset{n\to\infty}{\longrightarrow} 0$$

completes the proof. □

Notation

Symbols

$\mathbb{1}_{\{A\}}$	indicator function of the set A
$\mathbb{C}$	set of complex numbers
Γ	Gamma function
$\mathcal{B}(A)$	Borel σ-algebra of a set A
$\mathfrak{B}$, D	backshift operator, differential operator
$\mathfrak{C}$	positive constant with context-depending value
$\mathbb{N}, \mathbb{N}_0, \mathbb{Z}, \mathbb{R}, \mathbb{R}^+$	set of natural numbers, set of natural numbers and 0, set of whole numbers, set of real numbers, set of positive real numbers
$\nabla_\vartheta g(\vartheta)$, $\nabla_\vartheta g(\vartheta_0)$	$\nabla_\vartheta g(\vartheta) = \frac{\partial \operatorname{vec}(g(\vartheta))}{\partial \vartheta}$, $\nabla_\vartheta g(\vartheta_0) = \nabla_\vartheta g(\vartheta)_{\mid \vartheta = \vartheta_0}$
Pr_L	orthogonal projection onto L
$\xrightarrow{\mathcal{D}}$	convergence in distribution
$\xrightarrow{\mathcal{P}}$	convergence in probability
$\xrightarrow{a.s.}$	almost sure convergence
$\mathbb{P}, \mathbb{E}, \mathrm{Var}, \mathrm{Cov}$	probability, expectation, variance and covariance
$\Re(z), \Im(z)$	real and imaginary part of a complex valued z
σ	standard deviation parameter if $\mathbb{E}X^2 < \infty$ and scale parameter for an α-stable distribution with $\alpha < 2$, repectively
Σ_X	covariance matrix of a random vector X

$\mathrm{sign}(z)$	$\mathrm{sign}(z) = \mathbb{1}_{\{z>0\}} - \mathbb{1}_{\{z<0\}}$
$\widehat{f}_k$	kth Fourier coefficient of the function f
$A[s,t]$	(s,t)th-component of the matrix A
$A \otimes B$	Kronecker product of A and B
$A^\top, A^H$	transpose of A, conjugate transpose of A
e^A, $\log(A)$	matrix exponential of A, matrix logarithm of A
f^+, f^-	positive and negative part of f
g', $\nabla^2_\vartheta g(\vartheta)$	first derivative and Hessian matrix of a univariate function g
I_N	N-dimensional identity matrix
$o_{\mathbb{P}}(1), O_{\mathbb{P}}(1)$	a term which converges to 0 in probability, a term which is tight
$S_\alpha(\sigma,\beta,\mu)$	α-stable random variable with stability index α, scale parameter σ skewness parameter β and location parameter μ
$\det(A)$, $\mathrm{tr}(A)$	determinant of A, trace of a quadratic matrix A
$\mathrm{diag}(a_i)_k$	diagonal matrix of dimension k with diagonal entries $a_1,\ldots,a_k$

Specific variables

$(\varepsilon_k^{(\Delta)})_{k\in\mathbb{Z}}$	linear innovations of the sampled process
(A,B,C,L)	state space representation of a process
$(N_k^{(\Delta)})_{k\in\mathbb{Z}}$	white noise in the CMA(∞) representation of the sampled process
Δ	distance between observations
$f_Y^{(\Delta)}$	spectral density of the sampled process
$\mathcal{G}_m$, $\mathcal{G}_N$	index function sets of Chapter 5
$\overline{\Gamma}_{n,Y}, \overline{\gamma}_{n,Y}, \overline{\Gamma}_{n,N}$	sample autocovariance of $Y_1^{(\Delta)},\ldots Y_n^{(\Delta)}$ in the multivariate and univariate settings, sample autocovariance of $N_1^{(\Delta)},\ldots,N_n^{(\Delta)}$
ϑ, Θ	parameter, parameter space
$\widehat{\vartheta}_n^{(\Delta)}, \widehat{\vartheta}_n^{(\Delta,star)}, \widehat{\vartheta}_n^{(\Delta,A)}, \vartheta_n^*$	Whittle estimator of Chapter 3 and Section 5.2, adjusted Whittle estimator, quasi maximum likelihood estimator
$I_{n,Y}, I_{n,N}$	periodogram of $Y_1^{(\Delta)},\ldots Y_n^{(\Delta)}$, periodogram of $N_1^{(\Delta)},\ldots,N_n^{(\Delta)}$
L	Lévy process

m, d, N, r	dimensions of $Y, L, N^{(\Delta)}$ and Θ
p, q	orders of the autoregressive and moving average polynomial of the MCARMA/VARMA representation
P, Q, a, c	autoregressive and moving average polynomials of the multivariate and univariate CARMA representation
$V^{(\Delta)}$	covariance matrix of the linear innovations of the sampled process
$Y^{(\Delta)}$	sampled process

List of abbreviations

(M)CARMA	(multivariate) continuous-time autoregressive moving average process
(V)ARMA	(vector) autoregressive moving average process
i.i.d.	independent and identically distributed
NIG	normal inverse Gaussian
QMLE	quasi maximum likelihood estimator
VC	Vapnik–Chervonenkis

Bibliography

Andresen, A., Benth, F. E., Koekebakker, S. and Zakamulin, V. (2014). The CARMA interest rate model, *Int. J. Theor. Appl. Finance* **17**(02): 1–27.

Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus*, Cambridge University Press, Cambridge.

Bardet, J.-M., Doukhan, P. and León, J. R. (2008). Uniform limit theorems for the integrated periodogram of weakly dependent time series and their applications to Whittle's estimate, *J. Time Ser. Anal.* **29**(5): 906–945.

Barndorff-Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling, *Scand. J. Statist.* **24**(1): 1–13.

Benth, F. E., Klüppelberg, C., Müller, G. and Vos, L. (2014). Futures pricing in electricity markets based on stable CARMA spot models, *Energy Economics* **44**: 392–406.

Bergstrom, A. R. (1990). *Continuous time econometric modelling*, Oxford University Press.

Bernstein, D. S. (2005). *Matrix mathematics: theory, facts, and formulas with applications to linear systems*, 1 edn, Princeton University Press, Princeton.

Bharucha-Reid, A. and Römisch, W. (1985). Projective Schemes for Random Operator Equations. I. Weak Compactness of Approximate Solution Measures, *J. Integral Equations Appl.* **8**: 95–111.

Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley Series in Probability and Mathematical Statistics, Wiley.

Box, G. E., Jenkins, G. M., Reinsel, G. C. and Ljung, G. M. (2015). *Time series analysis: forecasting and control*, 5 edn, John Wiley & Sons.

Brewer, J. (1978). Kronecker products and matrix calculus in system theory, *IEEE Trans. Circuits Syst. I. Regul. Pap.* **25**(9): 772–781.

Brockwell, P., Davis, R. and Yang, Y. (2011). Estimation for non-negative Lévy-driven CARMA processes, *J. Bus. Econom. Statist.* **29**(2): 250–259.

Brockwell, P. J. (2001a). Continuous-time ARMA processes, *Handbook of Statist.* **19**: 249–276.

Brockwell, P. J. (2001b). Lévy-driven CARMA processes, *Ann. Inst. Statist. Math.* **53**(1): 113–124.

Brockwell, P. J. (2009). Lévy–driven continuous–time ARMA processes, *Handbook of financial time series*, Springer, pp. 457–480.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*, Springer Series in Statistics, New York.

Brockwell, P. J., Ferrazzano, V. and Klüppelberg, C. (2013). High-frequency sampling and kernel estimation for continuous-time moving average processes, *J. Time Ser. Anal.* **34**(3): 385–404.

Brockwell, P. J. and Lindner, A. (2009). Existence and uniqueness of stationary Lévy-driven CARMA processes, *Stochastic Process. Appl.* **119**(8): 2660–2681.

Brockwell, P. J. and Lindner, A. (2019). Sampling, Embedding and Inference for CARMA Processes, *J. Time Ser. Anal.* **40**(2): 163–181.

Brockwell, P. J. and Matsuda, Y. (2017). Continuous auto-regressive moving average random fields on $\mathbb{R}^n$, *J. R. Stat. Soc. Ser. B Stat. Methodol.* **79**(3): 833–857.

Can, S. U., Mikosch, T. and Samorodnitsky, G. (2010). Weak convergence of the function-indexed integrated periodogram for infinite variance processes, *Bernoulli* **16**(4): 995–1015.

Dahlhaus, R. (1988). Empirical spectral processes and their applications to time series analysis, *Stochastic Process. Appl.* **30**: 69–83.

Dahlhaus, R. and Polonik, W. (2009). Empirical spectral processes for locally stationary time series, *Bernoulli* **15**(1): 1–39.

Doob, J. L. (1944). The elementary Gaussian processes, *Ann. Math. Stat.* **15**(3): 229–282.

Drapatz, M. (2017). Limit theorems for the sample mean and sample autocovariances of continuous time moving averages driven by heavy-tailed Lévy noise, *ALEA Lat. Am. J. Probab. Math. Stat.* **14**(1): 403–426.

Dunsmuir, W. and Hannan, E. J. (1976). Vector linear time series models, *Adv. in Appl. Probab.* **8**(2): 339–364.

Fasen-Hartmann, V. and Mayer, C. (2020). A note on estimation of α-stable CARMA processes sampled at low frequencies. arXiv:2009.04823.

Fasen-Hartmann, V. and Mayer, C. (2021+). Whittle estimation of state space models with finite second moments, *Ann. Inst. Statist. Math.* . To appear.

Fasen-Hartmann, V. and Scholz, M. (2019). Quasi-maximum likelihood estimation for cointegrated solutions of continuous-time state space models observed at discrete time points, *Electron. J. Stat.* **13**(2): 5151–5212.

Fasen-Hartmann, V. and Scholz, M. (2020). Cointegrated continuous-time linear state-space and MCARMA models, *Stochastics* **92**(7): 1064–1099.

Fasen-Hartmann, V. and Scholz, M. (2021). Factorization and discrete-time representation of multivariate CARMA processes.

Fasen, V. (2013). Statistical inference of spectral estimation for continuous-time MA processes with finite second moments, *Math. Methods Statist.* **22**(4): 283–309.

Fasen, V. and Fuchs, F. (2013a). On the limit behavior of the periodogram of high-frequency sampled stable CARMA processes, *Stochastic Process. Appl.* **123**(1): 229–273.

Fasen, V. and Fuchs, F. (2013b). Spectral estimates for high-frequency sampled continuous-time autoregressive moving average processes, *J. Time Ser. Anal.* **34**(5): 532–551.

Fasen, V. and Kimmig, S. (2017). Information criteria for multivariate CARMA processes, *Bernoulli* **23**(4A): 2860–2886.

Ferrazzano, V. and Fuchs, F. (2013). Noise recovery for Lévy-driven CARMA processes and high-frequency behaviour of approximating Riemann sums, *Electron. J. Stat.* **7**: 533–561.

Fowler, R. H. (1929). *Statistical mechanics, the theory of the properties of matter in equilibrium*, Cambridge University Press.

García, I., Klüppelberg, C. and Müller, G. (2011). Estimation of stable CARMA models with an application to electricity spot prices, *Stat. Model.* **11**(5): 447–470.

Guidorzi, R. (1975). Canonical structures in the identification of multivariable systems, *Automatica J.* **11**(4): 361–374.

Hannan, E. J. (2009). *Multiple Time Series*, Wiley Ser. Probab. Stat. , New York.

Hannan, E. J. and Deistler, M. (1988). *The statistical theory of linear systems*, Wiley Ser. Probab. Stat., New York.

Hansen, L. P. and Sargent, T. J. (1983). The dimensionality of the aliasing problem in models with rational spectral densities, *Econometrica* pp. 377–387.

Hu, Y. and Long, H. (2007). Parameter estimation for Ornstein-Uhlenbeck processes driven by α-stable Lévy motions, *Commun. Stoch. Anal.* **1**(2): 1.

Hu, Y. and Long, H. (2009). Least squares estimator for Ornstein–Uhlenbeck processes driven by α-stable motions, *Stochastic Process. Appl.* **119**(8): 2465–2480.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities, *The journal of finance* **42**(2): 281–300.

Jones, R. H. (1981). Fitting a continuous time autoregression to discrete data, *Applied time series analysis II*, Elsevier, pp. 651–682.

Klüppelberg, C. and Mikosch, T. (1996). The integrated periodogram for stable processes, *Ann. Statist.* **24**(5): 1855 – 1879.

Klüppelberg, C. and Pham, V. S. (2019). Estimation of causal CARMA random fields, *arXiv:1902.04962* .

Körner, T. W. (1989). *Fourier analysis*, Cambridge University Press, Cambridge.

Krengel, U. (2011). *Ergodic theorems*, De Gruyter Studies in Mathematics, Berlin.

Ljungdahl, M. M. and Podolskij, M. (2020). Multi-dimensional parameter estimation of heavy-tailed moving averages, *arXiv:2007.15301* .

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer, Berlin.

Marquardt, T. and Stelzer, R. (2007). Multivariate CARMA processes, *Stochastic Process. Appl.* **117**(1): 96–120.

McElroy, T. S. (2013). Forecasting continuous-time processes with applications to signal extraction, *Ann. Inst. Statist. Math.* **65**(3): 439–456.

Mikosch, T. (1991). Functional limit theorems for random quadratic forms, *Stochastic Process. Appl.* **37**: 81–98.

Mikosch, T., Gadrich, T., Klüppelberg, C. and Adler, R. J. (1995). Parameter estimation for ARMA models with infinite variance innovations, *Ann. Statist.* **23**(1): 305–326.

Mikosch, T. and Norvaiša, R. (1997). Uniform convergence of the empirical spectral distribution function, *Stochastic Process. Appl.* **70**(1): 85–114.

Øigård, T. A., Hanssen, A., Hansen, R. E. and Godtliebsen, F. (2005). EM-estimation and modeling of heavy-tailed processes with the multivariate normal inverse Gaussian distribution, *Signal Process J.* **85**(8): 1655–1673.

Pham, V. S. (2020). Lévy-driven causal CARMA random fields, *Stochastic Process. Appl.* **130**(12): 7547–7574.

Pollard, D. (1984). *Convergence of Stochastic Processes*, Springer Series in Statistics, Springer, New York.

Priestley, M. B. (1981). *Spectral analysis and time series*, Vol. 1 and 2, Academic Press, Inc., London-New York.

Samorodnitsky, G. and Taqqu, M. S. (1994). *Stable non-Gaussian random processes*, Stochastic Modeling, Chapman & Hall, New York.

Sato, K.-I. (1999). *Lévy processes and infinitely divisible distributions*, Cambridge University Press, Cambridge.

Schlemm, E. and Stelzer, R. (2012a). Multivariate CARMA processes, continuous-time state space models and complete regularity of the innovations of the sampled processes, *Bernoulli* **18**(1): 46–63.

Schlemm, E. and Stelzer, R. (2012b). Quasi maximum likelihood estimation for strongly mixing state space models and multivariate Lévy-driven CARMA processes, *Electron. J. Stat.* **6**: 2185–2234.

Simon, B. (2015). *Real Analysis*, A comprehensive course in analysis, American Mathematical Society.

Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*, PhD thesis, Uppsala University.

Zhang, S. and Zhang, X. (2013). A least squares estimator for discretely observed Ornstein–Uhlenbeck processes driven by symmetric α-stable motions, *Ann. Inst. Statist. Math.* **65**(1): 89–103.