Statistica Sinica Preprint No: SS-2015-0314.R3	
Title	Limit behaviour of the truncated pathwise Fourier-transformation of L\'evy-driven CARMA processes for non-equidistant discrete time observations
Manuscript ID	SS-2015-0314.R3
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202015.0314
Complete List of Authors	Robert Stelzer and Zywilla Fechner
Corresponding Author	Robert Stelzer
E-mail	robert.stelzer@uni-ulm.de
Notice: Accepted version subject to English editing.	

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

Limit behaviour of the truncated pathwise Fourier-transformation of Lévy-driven CARMA processes for non-equidistant discrete time observations

Żywilla Fechner and Robert Stelzer

University of Silesia in Katowice and Ulm University

Abstract: This paper considers a continuous time analogue of the classical autoregressive moving average processes, Lévy-driven CARMA processes. First we describe limiting properties of the periodogram by means of the so-called truncated Fourier transform if observations are available continuously. The obtained results are in accordance with their counterparts from the discrete-time case. Then we discuss the numerical approximation of the truncated Fourier transform based on non-equidistant high frequency data. In order to ensure convergence of the numerical approximation to the true value of the truncated Fourier transform a certain control on the maximal distance between observations and the number of observations is needed. We obtain both convergence to the continuous time quantity and asymptotic normality under a high-frequency infinite time horizon limit.

Key words and phrases: CARMA Process, Frequency Domain, High-frequency Data, Lévy Process, Trapezoidal Rule

AMS Subject Classification 2010: Primary 62M10, 62M15, Secondary 60G10, 60G51

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

1 Introduction

The classical autoregressive moving average process ARMA has been broadly discussed in the literature. For a comprehensive discussion see e.g. the monograph by Brockwell and Davis (2006) and references therein. In the discrete time models we restrict ourselves to observations at fixed equidistant points in time. In many cases these observations made at discrete times come from an underlying continuous process, thus the natural question arises: can we model also the time series in continuous time? One of the earliest results dealing with properties of such processes can be found in Doob (1944). Later this problem was discussed in Brockwell (2001b) for continuous time ARMA processes driven by Gaussian noise. The next step was to extend these ideas to the models with noise modelled by jump processes, so-called Lévy-driven CARMA models introduced in Brockwell (2001a). In these papers time series are modelled as continuous time processes with continuous time noises (with or without jumps) and the inference is based mainly on discrete equidistant data. One of the latest results can be found in the paper Brockwell et al. (2011), which consideres QML estimations of the AR and MA parameters based on equidistant observations.

The estimation procedure of Lévy-driven CARMA processes in high-frequency settings has been discussed by Fasen and Fuchs (2013), where the authors deal with the limit behaviour of the periodogram of CARMA processes

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

under equidistant sampling when the sampling interval tends to 0. The results are analogous to ARMA processes: the periodogram for CARMA processes is not a consistent estimator of the spectral density function, but after appropriate smoothing the consistency can be obtained. Some related results were discussed by Fasen and Fuchs, where asymptotic distributions of periodograms of CARMA processes driven by a symmetric α-stable Lévy noise are obtained and where it is shown that the vector composed of periodograms for various frequencies converges in distribution to a function of a multidimensional stable random vector. Likewise, Fasen (2013) considers the behaviour of the periodogram for an equidistantly sampled continuous time moving average process when only the number of observations goes to infinity. Marquardt and Stelzer (2007) or Schlemm and Stelzer (2012) are, for instance, papers considering multivariate CARMA processes.

The problem of statistical analysis of such processes has been studied further for example in the dissertation Gillberg (2006), where different approaches to the estimation of CARMA processes with Gaussian noise are discussed both using equidistant and non-equidistant observations. The author works mainly in the frequency domain. He describes the properties of the truncated Fourier transform of a CARMA process with Gaussian noise on a fixed interval $[0,T]$ based on equidistant frequencies. In the non-equidistant case he has used a method

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

based on splines in order to find an approximation of the spectral density.

Another approach for the estimation of a stationary process $(Y(t))_{t\in\mathbb{R}}$ with mean zero, finite second-order moments and continuous covariance function has been discussed by Lii and Masry (1992, 1994), where they described some properties of a smoothed periodogram. Here observations are assumed to be given on a random grid (τ_k) of an interval $[0,T]$, where τ_k is a stationary point process on the real line which is independent of $(Y(t))_{t\in\mathbb{R}}$.

In the present paper we are going to describe the asymptotic behaviour of the so-called truncated Fourier transform of a CARMA process, which is a building block for an estimation of the spectral density of a CARMA process. We are going to use some of the ideas from Gillberg (2006) to prove results in more general settings.

The paper is structured as follows: first we recall second order Lévy-driven CARMA models and summarize the results needed later in Section 2. Then we define in Section 3 the truncated Fourier transform of a CARMA process and we investigate its asymptotic properties at a fixed frequency: for a non-zero frequency we obtain that the limiting law of the real and imaginary part is the two dimensional normal distribution with mean zero and the covariance matrix depending on the spectral density of the CARMA process. If we consider the truncated Fourier transform at zero, we obtain a one dimensional normal law

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

with mean zero and variance depending only on two parameters of the CARMA process. We show that the limiting law of the joint distribution of the squared modulus of the truncated Fourier transform at different positive frequencies converges to a vector of independent and exponentially distributed random variables with mean depending on the values of the spectral density. All these results can be interpreted as the limiting behaviour of the truncated Fourier transform when the CARMA process is observed continuously. The next step in Section 3.2 is to approximate the truncated Fourier transform when the CARMA process is observed on a non-equidistant deterministic grid. In order to find a numerical approximation value of the truncated Fourier transform we apply the trapezoidal rule. We are interested in the convergence of the truncated Fourier transform when the length of the interval T goes to infinity and the mesh of the grid to zero. Since the interplay of the length of the interval, of the number of elements of the grid and of the maximal distance between the elements of the grid plays a crucial role, in order to ensure the convergence of the approximating sum to the true value of the truncated Fourier transform we have to impose some limiting conditions on these quantities. In the last Section 4 we look at some illustrative simulations of the truncated Fourier transform based on non-equidistant observations. We consider Ornstein-Uhlenbeck type (CAR(1)) and CARMA(2,1) processes driven by a standard Brownian motion, a Variance Gamma process

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

and a "two-sided Poisson process" and we compare our simulations with the theoretical asymptotic distributions described earlier.

Note that all proofs are relegated to the online supplement to the present paper where some additional results on the moments of the truncated Fourier transform can be found as well.

Notation

The symbol $\mathbb{N} := \{1, 2, 3, \dots\}$ denotes the set of positive integers, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, $\mathbb{R}$ is the set of real numbers and $\mathbb{C}$ denotes the set of complex numbers. The symbol $\mathbb{R}^{m\times n}$, resp. $\mathbb{C}^{m\times n}$ denotes the space of real- (resp. complex-) valued matrices with m rows and n columns. For $A \in \mathbb{C}^{m\times n}$ the symbol A^T denotes the transposed of a matrix A. We are working on a given filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ satisfying the usual hypothesis (cf. Protter (2004), Chapter 1). Moreover, by $X \stackrel{d}{=} Y$ we denote that the random variables X and Y are equal in distribution.

2 Preliminaries

We begin with the model set-up given by Brockwell (2001a,b). A second-order Lévy-driven continuous-time ARMA(p,q) process is defined in terms of a state-

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

space representation of the formal differential equation

$$a(D)Y(t) = b(D)DL(t), \quad t \geq 0. \tag{2.1}$$

Here, D denotes differentiation with respect to t, non-negative integers p, q satisfying $p > q$ and $(L(t))_{t\geq 0}$ is a one dimensional Lévy process (i.e. a continuous time process with stationary and independent increments and $L(0) = 0$ a.s.; see e.g. Applebaum (2009)) with $\mathbb{E}L(1)^2 < \infty$. The polynomials

$$a(z) := z^p + a_1 z^{p-1} + \cdots + a_p, \quad b(z) := b_0 + b_1 z + \cdots + b_{p-1} z^{p-1},$$

are called the *autoregressive*- and *moving average* polynomial, respectively. We assume that $b_q \neq 0$ and $b_j = 0$ for $q < j < p$. The *state-space representation* consists of the *observation* and *state equations*:

$$Y(t) = \mathbf{b}^T \mathbf{X}(t), \tag{2.2}$$

$$d\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t)dt + \mathbf{e}dL(t), \tag{2.3}$$

where

$$\mathbf{A} := \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_p & -a_{p-1} & -a_{p-2} & \dots & -a_1 \end{bmatrix}, \quad \mathbf{X}(t) := \begin{bmatrix} X(t) \\ X^{(1)}(t) \\ \vdots \\ X^{(p-2)}(t) \\ X^{(p-1)}(t) \end{bmatrix},$$

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

$$\mathbf{e} := [0,\dots 0,1]^T, \quad \mathbf{b} := \left[b_0, b_1, \dots b_{p-1}\right]^T,$$

i.e. $\mathbf{A} \in \mathbb{R}^{p\times p}$, $\mathbf{X}(t) \in \mathbb{R}^{p\times 1}$ $\mathbf{e} \in \mathbb{R}^p$, $\mathbf{b} \in \mathbb{R}^p$. If $p = 1$, we set $\mathbf{A} = -a_1$.

Assumption 2.1. $\mathbb{E}L(1) = 0$ *and* $\mathbb{E}|L(1)|^2 = \sigma^2 < \infty$.

Observe that $\mathbb{E}[L(s)L(t)] = \min\{s,t\}\mathbb{E}|L(1)|^2$. It is shown e.g. in Brockwell (2009) that the solution $\mathbf{X}(t)$ of (2.3) satisfies

$$\mathbf{X}(t) = e^{\mathbf{A}t}\mathbf{X}(0) + \int_0^t e^{\mathbf{A}(t-u)}\mathbf{e}dL(u). \tag{2.4}$$

Assumption 2.2. $\mathbf{X}(0)$ *is independent of* $(L(t))_{t\geq 0}$.

From now on let us assume that Assumption 2.2 holds. It is well-known from Proposition 2 of Brockwell (2009) that under Assumptions 2.1 and 2.2 the process $\{\mathbf{X}(t)\}_{t\geq 0}$ is strictly stationary and causal iff $\mathbf{X}(0)$ has the same distribution as $\int_0^\infty e^{\mathbf{A}u}\mathbf{e}dL(u)$ and the p (not necessarily distinct) eigenvalues $\lambda_1,\dots,\lambda_p$ of $\mathbf{A}$ have negative real parts, i.e. $\Re(\lambda_i) < 0, \quad i = 1,\dots,p$. Now we extend the Lévy process $(L(u))_{u\geq 0}$ to the whole line in the usual way: Let $\widetilde{L} = (\widetilde{L}(t))_{t\geq 0}$ be an independent copy of $(L(t))_{t\geq 0}$. For $t \in \mathbb{R}$ we define $L^*(t) := L(t)\mathbb{1}_{[0,\infty)}(t) + \widetilde{L}(-t-)\mathbb{1}_{(-\infty,0]}(t)$. In order to get stationary solutions of (2.3) we need the following assumptions:

Assumption 2.3. *All eigenvalues of* $\mathbf{A}$ *have strictly negative real parts.*

Assumption 2.4. $\mathbf{X}(0) \stackrel{d}{=} \int_{-\infty}^0 e^{-\mathbf{A}u}\mathbf{e}dL^*(u)$

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

In Brockwell (2009) it was shown that if Assumptions 2.3 and 2.4 are satisfied the process $(\mathbf{X}(t))_{t\in\mathbb{R}}$ given by

$$\mathbf{X}(t) = \int_{-\infty}^{t} e^{\mathbf{A}(t-u)}\mathbf{e}dL^*(u) \tag{2.5}$$

is a strictly stationary solution of (2.3) (with L replaced by L^*) for $t \in \mathbb{R}$ with corresponding CARMA process

$$Y(t) = \int_{-\infty}^{t} \mathbf{b}^T e^{\mathbf{A}(t-u)}\mathbf{e}dL^*(u). \tag{2.6}$$

For $t \geq 0$ one can rewrite it in the following form

$$Y(t) = \mathbf{b}^T e^{\mathbf{A}t}\mathbf{X}(0) + \int_{0}^{t} \mathbf{b}^T e^{\mathbf{A}(t-u)}\mathbf{e}dL(u). \tag{2.7}$$

In the present paper the spectral density of a CARMA process plays a crucial role. The spectral density is the Fourier transform of the autocovariance function $\gamma_Y(h) := \mathbb{C}\mathrm{ov}(Y(0), Y(h))$ for $h \in \mathbb{R}$. The spectral density of a CARMA process is

$$f_Y(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \gamma_Y(h)e^{-ih\omega}dh = \frac{\sigma^2}{2\pi}\frac{|b(i\omega)|^2}{|a(i\omega)|^2}, \quad \omega \in \mathbb{R}. \tag{2.8}$$

3 Limit behaviour of the Fourier transform

In this section we are going to deal with the Fourier transform of the CARMA process assuming that the observations are given continuously on the time inter-

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

val $[0,T]$. A similar idea for Gaussian CARMA processes was presented in Gillberg (2006) for equidistant observations. The truncated continuous-time Fourier transform of the process Y at a fixed frequency $\omega \in \mathbb{R}$ is given by

$$\mathscr{F}_T(Y)(\omega) := \frac{1}{\sqrt{T}} \int_0^T Y(t) e^{-i\omega t} dt.$$

Observe that the norming constant $\frac{1}{\sqrt{T}}$ is taken as this ensures convergence in distribution for $T \to \infty$ as will be shown later.

3.1 Properties of the Truncated Fourier Transform of a CARMA Process

First we derive an alternative representation.

Lemma 3.1. *Let* $\mathbf{X}$ *and* Y *be processes given by the state-space representation* (2.2) *and* (2.3)*. Suppose that Assumptions 2.1, 2.2 and 2.3 are satisfied. Then the truncated Fourier transform of the* CARMA *process* Y *at a fixed frequency* $\omega \in \mathbb{R}$ *is of the form*

$$\begin{aligned} \mathscr{F}_T(Y)(\omega) = &\frac{1}{\sqrt{T}} \frac{b(i\omega)}{a(i\omega)} \int_0^T e^{-i\omega t} dL(t) \\ &+ \frac{1}{\sqrt{T}} \mathbf{b}^T (i\omega I - A)^{-1} \left(\mathbf{X}(0) - e^{-i\omega T} \mathbf{X}(T) \right), \end{aligned} \tag{3.9}$$

or equivalently

$$\mathscr{F}_T(Y)(\omega) = \frac{1}{\sqrt{T}} \mathbf{b}^T (i\omega I - A)^{-1} \tag{3.10}$$

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

$$\times \left[\int_0^T \left(e^{-i\omega u} - e^{-i\omega T} e^{\mathbf{A}(T-u)} \right) \mathbf{e} dL(u) + \left(I - e^{(-i\omega I + \mathbf{A})T} \right) \mathbf{X}(0) \right].$$

Now we are going to investigate asymptotic properties of the truncated Fourier transform. First, we will show that the second summand of (3.9) converges in probability to zero.

Lemma 3.2. *Let* $\mathbf{X}$ *and* Y *be processes given by the state-space representation* (2.2) *and* (2.3)*. Suppose that Assumptions 2.1, 2.2 and 2.3 are satisfied. Let*

$$\tilde{Z}(T) := \mathscr{F}_T(Y)(\omega) - \frac{1}{\sqrt{T}} \frac{b(i\omega)}{a(i\omega)} \int_0^T e^{-i\omega t} dL(t).$$

Then $\mathbb{P} - \lim_{T\to\infty} |\tilde{Z}(T)| = 0$.

Now we will show that the first summand of formula (3.9) converges in distribution. Thus, together with Lemma 3.2 we obtain the limit in distribution of the truncated Fourier transform. We have two cases: the first case is if the frequency $\omega = 0$. Then the truncated Fourier transform is a real valued function. In the second case for frequencies $\omega \neq 0$ the truncated Fourier transform is a complex valued function. In both cases we first give the description of the distribution of the truncated Fourier transform and afterwards we describe the distribution of the squared modulus of the truncated Fourier transform.

Theorem 3.3. *Let* $\mathbf{X}$ *and* Y *be processes given by the state-space representation*

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

(2.2) *and* (2.3). *Suppose that Assumptions 2.1 and 2.3 are satisfied. Let*

$$Z(T) := \frac{1}{\sqrt{T}} \frac{b(0)}{a(0)} \int_0^T dL(t).$$

Then

$$d - \lim_{T\to\infty} Z(T) \sim \mathcal{N}\left(0, \left(\frac{b(0)}{a(0)}\right)^2 \sigma^2\right), \quad d - \lim_{T\to\infty} \frac{1}{\sigma^2} \left|\frac{a(0)Z(T)}{b(0)}\right|^2 \sim \chi^2(1)$$

In order to find the asymptotic distribution of the truncated Fourier transform we use the multivariate Central Limit Theorem. Note that we state all results for positive frequencies as the corresponding results for negative can be obtained by taking the complex conjugate.

Theorem 3.4. *Let* **X** *and Y be processes given by the state-space representation* (2.2) *and* (2.3). *Suppose that Assumptions 2.1 and 2.3 are satisfied. Assume that* $\omega > 0$. *Put* $Z(T) := \frac{1}{\sqrt{T}} \frac{b(i\omega)}{a(i\omega)} \int_0^T e^{-i\omega t} dL(t)$ *and* $Z(T) = [\,\Re Z(T), \Im Z(T)]^T$. *Then* $d - \lim_{T\to\infty} Z(T) \sim \mathcal{N}(0, \Sigma)$, *where* $\Sigma = \frac{\sigma^2}{2} \left|\frac{b(i\omega)}{a(i\omega)}\right|^2 I_{2\times 2}$.

Now we apply this theorem to find the asymptotic distribution of the truncated Fourier transform squared.

Theorem 3.5. *Let* **X** *and Y be processes given by the state-space representation* (2.2) *and* (2.3). *Suppose that Assumptions 2.1 and 2.3 are satisfied. Let Z be defined as in Theorem 3.4. Then* $|Z|^2 \sim \text{Exp}\left(\sigma^2 \left|\frac{b(i\omega)}{a(i\omega)}\right|^2\right)$, *where* $\text{Exp}(\lambda)$ *denotes the exponential distribution with mean* λ.

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

Now we are going to give the description of the convergence of the random vector consisting of the truncated Fourier transform at different frequencies.

Theorem 3.6. *Let* $\mathbf{X}$ *and* Y *be processes given by the state-space representation* (2.2) *and* (2.3). *Suppose that Assumptions 2.1, 2.2 and 2.3 are satisfied. Let* $0 < \omega_1 < \cdots < \omega_d$ *be fixed frequencies.*

Then $\left[\Re\left(\mathscr{F}_T(Y)(\omega_j)\right), \Im\left(\mathscr{F}_T(Y)(\omega_j)\right)\right]_{j=1,\ldots,d}^T$ *converges to* $\mathscr{N}\left(0, \frac{\sigma^2}{2}\mathbf{B}\right)$, *with* $\mathbf{B} = \operatorname{diag}\left(\left|\frac{b(i\omega_1)}{a(i\omega_1)}\right|^2, \left|\frac{b(i\omega_1)}{a(i\omega_1)}\right|^2, \ldots, \left|\frac{b(i\omega_d)}{a(i\omega_d)}\right|^2, \left|\frac{b(i\omega_d)}{a(i\omega_d)}\right|^2\right)$. $\left[|\mathscr{F}_T(\omega_j)|^2\right]_{j=1,\ldots,d}^T$ *converges to a random vector of independent* $\operatorname{Exp}\left(\sigma^2\left|\frac{b(i\omega_j)}{a(i\omega_j)}\right|^2\right)$ *distributed random variables for* $j = 1, \ldots, d$.

Note that Theorem 3.4 is basically a special case of Theorem 3.6. However, it seems worthwhile to state this special case and the proof in the online supplement in the case of several frequencies is much more complicated and so it is enlightening to look at the way more elementary reasoning given there for Theorem 3.6.

The limiting result is the analogue of the one for discrete time ARMA models. (See e.g. Brockwell and Davis (2006), Chapter 10.)

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

3.2 Numerical Approximation and Limiting Behaviour of the Truncated Pathwise Fourier Transform Based on Non-equidistant Discrete Grids

In this section we deal with the numerical approximation of the integral

$$\mathscr{F}_T(Y)(\omega) := \frac{1}{\sqrt{T}} \int_0^T Y(t) e^{-i\omega t} dt. \tag{3.11}$$

Our aim is to describe conditions under which we are able to calculate numerically the truncated Fourier transform of a CARMA process based on non-equidistant observations.

To this end we consider trapezoidal approximations of integrals over non-equidistant grids and their convergence rate in L^2. Related results on equidistant grids can be found in Brockwell and Schlemm (2013). Our general result is the following:

Theorem 3.7. *Let* $\mathbf{X}$ *and* Y *be processes given by the state-space representation* (2.2) *and* (2.3). *Suppose that Assumptions 2.1, 2.2, 2.3 and 2.4 are satisfied. Assume that* $F\colon \mathbb{R} \to \mathbb{R}^d$ *is a twice continuously differentiable function with* $\|F''\|_\infty < \infty$. *Let* $\left(x_i^{(T)}\right)_{i=0,\dots,N(T)-2}$ *be a partition of the interval* $[a,b]$ *with* $x_0^{(T)} = a$ *and* $x_{N(T)-1}^{(T)} = b$ *and let* $h_{\max}(T) = \max_{j=0,\dots,N(T)-1}\left(x_{j+1}^{(T)} - x_j^{(T)}\right)$. *Put*

$$\alpha_0^{(N(T))} = \frac{x_1^{(T)} - x_0^{(T)}}{2} F\left(x_0^{(T)}\right),$$

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

$$\alpha_{N(T)-1}^{(N(T))} = \frac{x_{N(T)-1}^{(T)} - x_{N(T)-2}^{(T)}}{2} F\left(x_{N(T)-1}^{(T)}\right), \tag{3.12}$$

$$\alpha_j^{(N(T))} = \frac{x_{j+1}^{(T)} - x_{j-1}^{(T)}}{2} F\left(x_j^{(T)}\right), \quad j = 1, \ldots, N(T)-2. \tag{3.13}$$

Then there exist positive constants C_1, C_2 such that

$$\mathbb{E}\left[\left\| \sum_{j=0}^{N(T)-1} \alpha_j^{(N(T))} Y\left(x_j^{(T)}\right) - \int_a^b Y(t)F(t)dt \right\|^2\right] \leq C_1(C_2+T)N(T)^2 h_{\max}^6(T)$$

and thus if $\lim_{T\to\infty} TN(T)^2 h_{\max}^6(T) = 0$*, then*

$$\lim_{T\to\infty} \left\| \sum_{j=0}^{N(T)-1} \alpha_j^{(N(T))} F\left(x_j^{(T)}\right) - \int_a^b Y(t)F(t)dt \right\|_{L^2} = 0.$$

Now we are applying this result to find a numerical approximation of the truncated Fourier transform. Using the notation of Theorem 3.7 we denote the trapezoidal approximation of

$$\mathscr{F}_T(Y)(\omega) = \frac{1}{\sqrt{T}} \int_0^T Y(t) e^{-i\omega t} dt$$

by $\mathscr{T}_T(Y)(\omega)$, i.e.

$$\mathscr{T}_T(Y)(\omega) = \frac{1}{\sqrt{T}} \sum_{j=0}^{N-1} \alpha_j^{(N)} Y\left(x_j^{(N)}\right),$$

where the grid points $\left(x_j^{(N)}\right)_{j=0,\ldots,N(T)-1}$ are given as in Theorem 3.7 and

$$\alpha_0^{(N(T))} = \frac{x_1^{(T)} - x_0^{(T)}}{2} F\left(x_0^{(T)}\right),$$

$$\alpha_{N(T)-1}^{(N(T))} = \frac{x_{N(T)-1}^{(T)} - x_{N(T)-2}^{(T)}}{2} F\left(x_{N(T)-1}^{(T)}\right),$$

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

$$\alpha_j^{(N(T))} = \frac{x_{j+1}^{(T)} - x_{j-1}^{(T)}}{2} F\left(x_j^{(T)}\right), \quad j = 1, \ldots, N(T) - 2,$$

with $F(x) = e^{-i\omega x}$.

Theorem 3.8. *Let* $\mathbf{X}$ *and* Y *be processes given by the state-space representation* (2.2) *and* (2.3). *Suppose that Assumptions 2.1, 2.2, 2.3 and 2.4 are satisfied and that the process* Y *is observed at not necessarily equidistant points* $0 = x_0^{(T)} < x_1^{(T)} < \cdots < x_{N(T)-1}^{(T)} = T$. *Let* $h_{\max}(T) := \max_{j=0,\ldots,N(T)-2}\left(x_{j+1}^{(T)} - x_j^{(T)}\right)$. *If* $\lim_{T\to\infty} N(T) h_{\max}^3(T) = 0$, *then* $\lim_{T\to\infty} \|\mathscr{T}_T(Y)(\omega) - \mathscr{F}_T(Y)(\omega)\|_{L^2} = 0$ *and thus also* $\mathbb{P} - \lim_{T\to\infty} [\mathscr{T}_T(Y)(\omega) - \mathscr{F}_T(Y)(\omega)] = 0$.

Now we can obtain a central limit theorem for the truncated Fourier transform:

Theorem 3.9. *Let* $\mathbf{X}$ *and* Y *be processes given by the state-space representation* (2.2) *and* (2.3). *Suppose that Assumptions 2.1, 2.2, 2.3 and 2.4 are satisfied and the process* Y *is observed at not necessarily equidistant points* $0 = x_0^{(T)} < x_1^{(T)} < \cdots < x_{N(T)-1}^{(T)} = T$. *Let* $h_{\max}(T) := \max_{j=0,\ldots,N(T)-2}(x_{j+1}^{(T)} - x_j^{(T)})$. *Let* $\alpha_j^{(N)}$ *be defined as in Theorem 3.7. Assume that* $\lim_{T\to\infty} N(T) h_{\max}^3(T) = 0$. *Put* $\Sigma = \frac{\sigma^2}{2}\left|\frac{b(i\omega)}{a(i\omega)}\right|^2 I_{2\times 2}$. *If* $\omega \neq 0$, *then*

$$d - \lim_{T\to\infty} \begin{bmatrix} \Re(\mathscr{T}_T Y(\omega)) \\ \Im(\mathscr{T}_T Y(\omega)) \end{bmatrix} = \mathscr{N}(0, \Sigma),$$

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

$$d - \lim_{T\to\infty} \left(\Re(\mathscr{T}_T Y(\omega))^2 + \Im(\mathscr{T}_T Y(\omega))^2\right) = \mathrm{Exp}\left(\sigma^2 \left|\frac{b(i\omega)}{a(i\omega)}\right|^2\right).$$

If $\omega = 0$*, then* $d - \lim_{T\to\infty} \mathscr{T}_T Y(0) = \mathscr{N}\left(0, \left(\frac{b(0)}{a(0)}\right)^2 \sigma^2\right)$ *and*

$$d - \lim_{T\to\infty} \frac{1}{\sigma^2} \left|\frac{a(0)\mathscr{T}_T Y(0)}{b(0)}\right|^2 \sim \chi^2(1).$$

Clearly, an analogous statement using Theorem 3.6 holds for the joint distribution when the truncated Fourier transform is taken at different frequencies.

4 Illustrative simulations

We now turn to a numerical illustration of the theoretical convergence results given in Section 3.2. We are looking at simulations of CARMA processes and their numerically approximated truncated Fourier transform over different time horizons and maximal grid widths. To illustrate the convergence to the asymptotic normal distribution we shall look at several frequencies and different driving Lévy processes, standard Brownian motion, a Variance Gamma process and a "two sided Poisson process". Of course, the truncated Fourier transform of $(Y(t))_{t\in[0,T]}$ is obtained using the trapezoidal rule based on non-equidistant observations of the CARMA process $(Y(t))_{t\in[0,T]}$ given on the interval $[0,T]$. On an interval $[0,T]$ we generate a non-equidistant grid in the following way: we fix the maximal distance $h_{\max}(T)$ between elements of the grid and from each

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

interval $\left[i\frac{1}{2}h_{\max}(T),(i+1)\frac{1}{2}h_{\max}(T)\right)$ for $i=0,1,\ldots,N-1$ we draw a number according to the uniform distribution. This results in a non-equidistant grid with the number of points being $N(T)=2T/h_{\max}+1$.

For our simulations we used the R Project for Statistical Computing. For the simulation we first generate the non-equidistant grid by the above procedure and then join it with a regular grid of mesh 0.001, which is still on average five times finer than the non-equidistant grid of the largest time horizon considered in order to ensure that the discretisation error for the CARMA SDE is very small in the simulated data. On this joint non-equidistant grid the CARMA process Y is simulated with a standard Euler scheme for the state space representation. Afterwards only the simulated values at the times of the original non-equidistant grid are used to compute the approximation of the truncated Fourier transform with the trapezoidal rule. In all cases we simulate 2000 independent paths of the CARMA process and compute the associated values of the truncated Fourier transform at the following frequencies: $[\omega_1,\omega_2,\omega_3,\omega_4]=[0,\,0.1,\,1,\,10]$. For the non-zero frequencies real and imaginary part have to be considered separately. However, in the following we look only at the real parts as the behaviour of the imaginary parts is most similar. Mainly, the results are presented via QQ-plots where the theoretical values follow the (limiting) law described in Theorem 3.9.

We are going to consider CARMA processes with the following autoregres-

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

sive and moving average orders: $(p,q) = (1,0)$, i.e. an Ornstein-Uhlenbeck type process, and $(p,q) = (2,1)$. For the time horizon T and the maximum distance of the non-equidistant observation times we consider the pairs $(T = 10, h_{\max} = 0.1)$, $(T = 50, h_{\max} = 0.05)$ and $(T = 100, h_{\max} = 0.01)$.

For each case we consider three different driving Lévy noises: standard Brownian Motion, a Variance Gamma process and a "two sided Poisson process". For the definition and properties of the Variance Gamma process we refer to Madan et al. (1998) and references therein. We construct the process in the following way: $V_t = G_t^1 - G_t^2$, where G_t^1 and G_t^2 are independent Gamma processes with shape parameter 1 and scale parameter 4. Likewise the "two sided Poisson process" is the difference of two independent Poisson processes with rate 10, i.e. a compound Poisson process with rate 20 and jumps $+1$ and -1 both with probability $1/2$.

Example 4.1. We consider the CAR(1) model. Then $\mathbf{A} = -a_1$ and $a(z) := z + a_1, \quad b(z) = b_0$. So the spectral density is $f(\omega) = \frac{\sigma^2}{2\pi}\left|\frac{b(i\omega)}{a(i\omega)}\right|^2 = \frac{\sigma^2}{2\pi}\frac{b_0^2}{\omega^2+a_1^2}$. For the simulations we take $[b_0, a_1] = [1,2]$.

QQ-plots of the results for 2000 simulated paths for the four frequencies and three combinations of time horizon and maximum grid width can be found in Figures 1, 2 and 3 for the driving Lévy process being a standard Brownian motion, a Variance Gamma and a two-sided Poisson process, respectively.

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

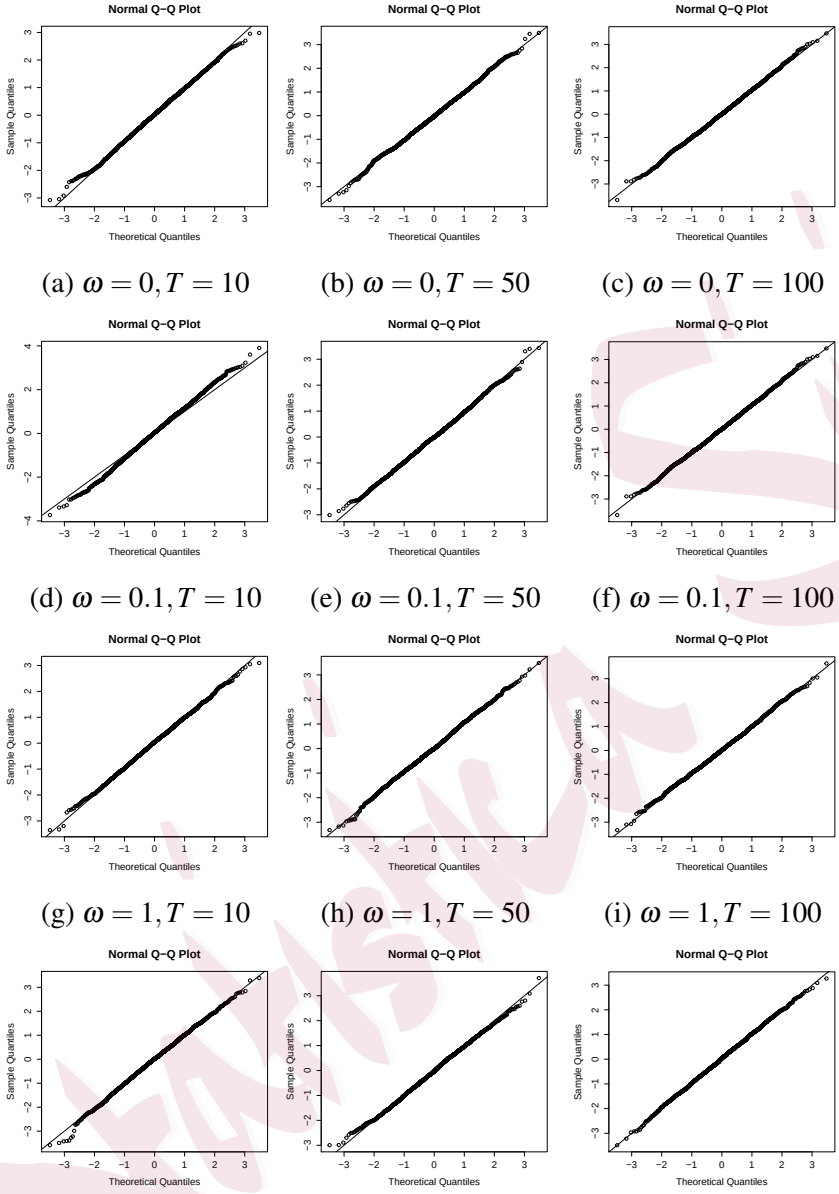
Example 4.2. We consider the CARMA$(2,1)$ model. We have

$$\mathbf{A} := \begin{bmatrix} 0 & 1 \\ -a_2 & -a_1 \end{bmatrix}, \quad \mathbf{b} := \begin{bmatrix} b_0 \\ 1 \end{bmatrix}, \quad \mathbf{e} := \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{X}(t) := \begin{bmatrix} X(t) \\ X^{(1)}(t) \end{bmatrix}$$

The autoregressive and moving-average polynomials are of the form $a(z) = z^2 + a_1 z + a_2, \quad b(z) = z + b_0$. We have $\frac{b(i\omega)}{a(i\omega)} = \frac{i\omega + b_0}{(i\omega)^2 + (i\omega)a_1 + a_2}$, $f(\omega) = \frac{\sigma^2}{2\pi}\left|\frac{b(i\omega)}{a(i\omega)}\right|^2 = \frac{\sigma^2}{2\pi}\frac{b_0^2+\omega^2}{\omega^4+(a_1^2-2a_2)\omega^2+a_2^2}$. For the simulation procedure we take $[b_0, b_1, a_1, a_2] = [1, 1, 1, 2]$. The plots for this simulation study can be found in the online appendix.

In Figure 1 we notice at first a pretty good fit of the empirical quantiles from the simulations with the theoretical ones of the asymptotic distribution across all time horizons and frequencies. Looking more carefully, the fit in the tails clearly improves when the time horizon/fineness of the grid increases, but it is never bad. For the longest time horizon and finest grid the fit is clearly very good. Of course, it should not be forgotten that in this case the distribution of the (trapezoidal approximation of the) truncated Fourier transform is always exactly Gaussian and not only asymptotically. When looking across the non-zero frequencies one notes that for the shortest time horizon the quantiles for the smallest frequency 0.1 appear to lie on a line which is somewhat different from the line of the theoretical quantiles. This indicates that the quantiles of the simulated paths come from a normal distribution, but one with a different variance then the asymptotic one. It is no surprise that this occurs for the lowest

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)



(a) $\omega = 0, T = 10$ (b) $\omega = 0, T = 50$ (c) $\omega = 0, T = 100$

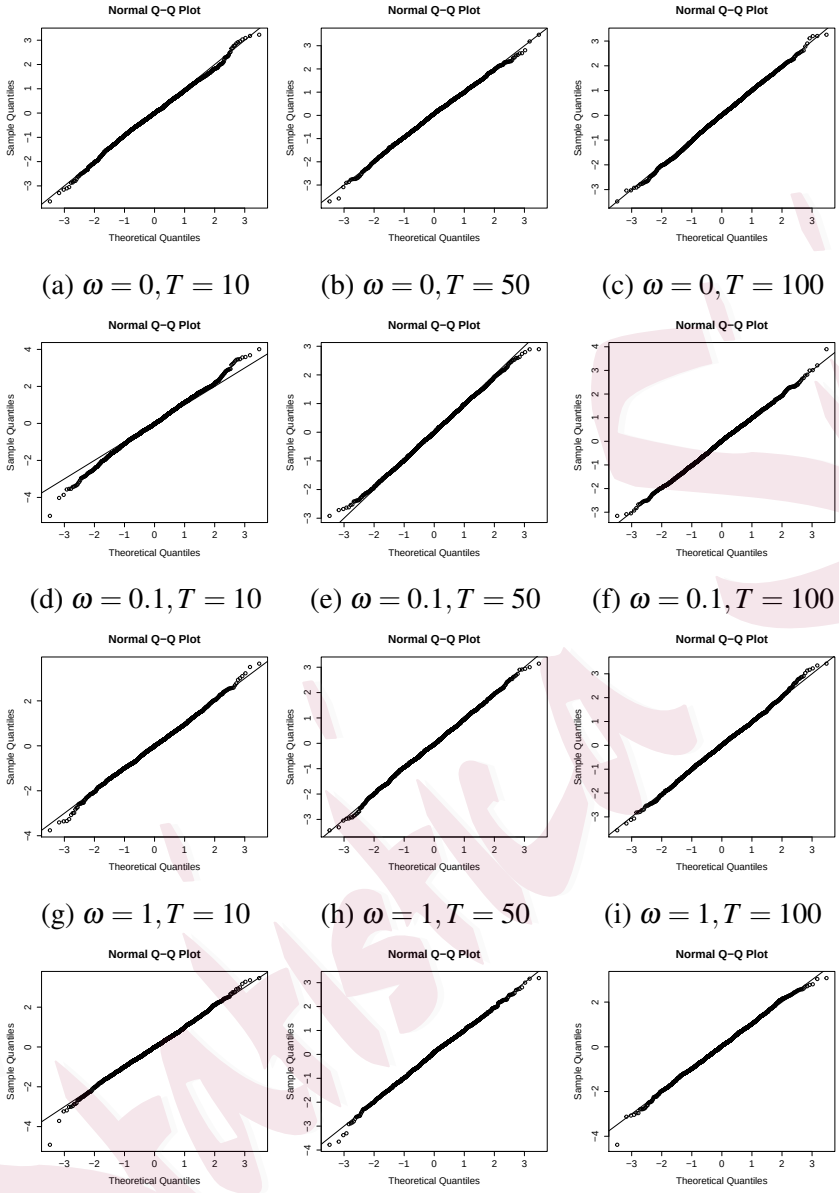
(d) $\omega = 0.1, T = 10$ (e) $\omega = 0.1, T = 50$ (f) $\omega = 0.1, T = 100$

(g) $\omega = 1, T = 10$ (h) $\omega = 1, T = 50$ (i) $\omega = 1, T = 100$

(j) $\omega = 10, T = 10$ (k) $\omega = 10, T = 50$ (l) $\omega = 10, T = 100$

Figure 1: Normal QQ plots for the real part of the truncated Fourier transform of the Ornstein-Uhlenbeck type process driven by standard Brownian Motion for the frequencies $0, 0.1, 1, 10$ (rows) and time horizons/maximum non-equidistant grid sizes $10/0.1, 50/0.05, 100/0.01$ (columns). The theoretical quantiles are coming from the (limiting) law described in Theorem 3.9.

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)



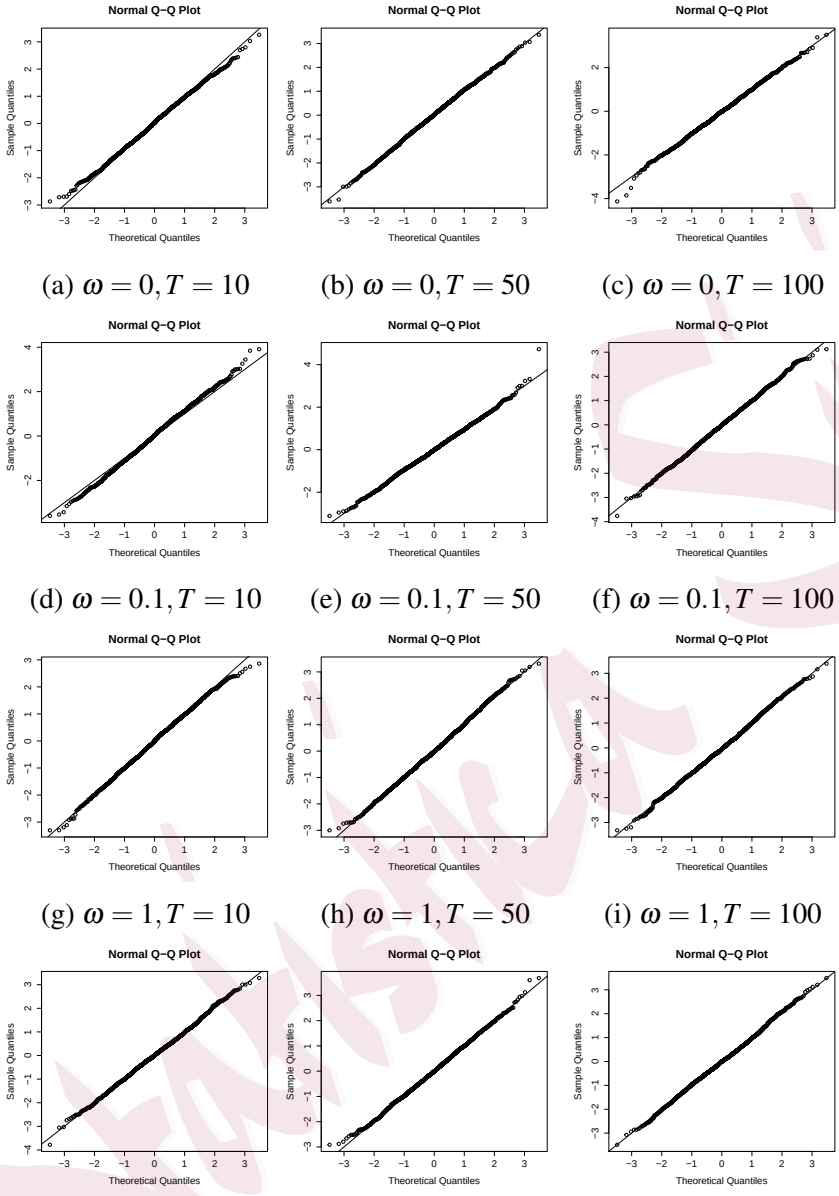
(a) $\omega = 0, T = 10$ (b) $\omega = 0, T = 50$ (c) $\omega = 0, T = 100$

(d) $\omega = 0.1, T = 10$ (e) $\omega = 0.1, T = 50$ (f) $\omega = 0.1, T = 100$

(g) $\omega = 1, T = 10$ (h) $\omega = 1, T = 50$ (i) $\omega = 1, T = 100$

(j) $\omega = 10, T = 10$ (k) $\omega = 10, T = 50$ (l) $\omega = 10, T = 100$

Figure 2: Normal QQ plots for the real part of the truncated Fourier transform of the Ornstein-Uhlenbeck type process driven by a Variance Gamma process for the frequencies $0, 0.1, 1, 10$ (rows) and time horizons/maximum non-equidistant grid sizes $10/0.1, 50/0.05, 100/0.01$ (columns). The theoretical quantiles are coming from the (limiting) law described in Theorem 3.9.



(a) $\omega = 0, T = 10$ (b) $\omega = 0, T = 50$ (c) $\omega = 0, T = 100$

(d) $\omega = 0.1, T = 10$ (e) $\omega = 0.1, T = 50$ (f) $\omega = 0.1, T = 100$

(g) $\omega = 1, T = 10$ (h) $\omega = 1, T = 50$ (i) $\omega = 1, T = 100$

(j) $\omega = 10, T = 10$ (k) $\omega = 10, T = 50$ (l) $\omega = 10, T = 100$

Figure 3: Normal QQ plots for the real part of the truncated Fourier transform of the Ornstein-Uhlenbeck type process driven by a two-sided Poisson process for the frequencies $0, 0.1, 1, 10$ (rows) and time horizons/maximum non-equidistant grid sizes $10/0.1, 50/0.05, 100/0.01$ (columns). The theoretical quantiles are coming from the (limiting) law described in Theorem 3.9.

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

frequency and the smallest time interval, as for low frequencies one observes – regardless of the fineness of the sampling – the fewest full cycles over a time interval of fixed length. For this combination of time horizon and frequency we see only one full cycle.

Turning to Figure 2, we first notice that the fit in the tails improves again clearly with increasing $(T, 1/h_{\max})$. Especially, for the highest $(T, 1/h_{\max})$ one sees that the fit in the tails is a bit worse now for a driving Variance Gamma process compared with the driving Brownian motion in Figure 1. Of course, now the simulated values are indeed only asymptotically following a Gaussian distribution. Looking at the different non-zero frequencies one again sees that the fit improves for the higher frequencies. Most notably for the lowest frequency one sees for the smallest $T = 10$ again that the points do seem to lie on a straight line in the normal QQ-plot, but one with a different slope than for the theoretical quantiles. Hence, the variance is clearly different from the asmpytotic one. Obviously, this effect is now more pronounced than in the case of the driving Brownian motion.

Moving on to the case of the driving process being a two-sided Poisson process in Figure 3 we first of all note that again the fit in particular in the tails clearly improves with increasing $(T, 1/h_{\max})$. For the highest $(T, 1/h_{\max})$ the simulated and theoretical asymptotic quantiles agree again extremely well.

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

Again it is certainly a bit worse than in the case of a driving Brownian motion, but it seems to be very similar to the Variance Gamma case, although maybe for frequency 0 the agreement of the quantiles is slightly worse. Turning to the behaviour across non-zero frequencies, we see again that the quantiles are closer for higher frequencies and that for the smallest non-zero frequency and time horizon the empirical quantiles seem to be in line with a normal distribution with a somewhat different variance compared to the asymptotic one. The size of this effect seems to be rather similar to the Variance Gamma case. Interestingly, also at frequency 0 the QQ-plot seems to indicate for $T = 10$ that the empirical quantiles are close to the ones of a normal distribution with a slightly different variance than the asymptotic one.

To summarize the simulation study in the CAR(1)/OU-type case we can clearly conclude that the asymptotic distribution result approximates the finite-sample distribution of the trapezoidal approximation of the truncated Fourier transform in our simulations very well and that the convergence to the asymptotic distribution is fast. For small frequencies, especially when one has about one full cycle or less over the time horizon considered, one has to be careful, as then the distribution tends to be somewhat different from the asymptotic one which is – as discussed – not surprising. This effect seems to be more pronounced when one considers a Lévy process with jumps compared to a Brownian

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

motion. In general the quality of the approximation of the simulated quantiles by the asymptotic ones is somewhat better in the case of a Brownian motion than for a pure jump process. Comparing the driving jump processes, the finite activity rather discrete two-sided Poisson process with the infinite activity Variance Gamma process, we do not see any significant differences. It should be noted that both jump processes are, however, light-tailed in the sense that they have exponential moments. It would not be surprising if this picture changes when considering a really heavily tailed driving Lévy process. Note that our theoretical results are valid also in rather heavily-tailed cases. For the asymptotic normality of the (trapezoidal approximation of the) truncated Fourier transform we only needed finite second moments.

Turning to the simulations of CARMA(2,1) processes, most of the findings of the CAR(1)/OU case remain valid, so we only point out the differences. In the case of a driving Brownian motion, depicted in Figure S.1, the only difference seems to be that for $T = 10$ and $\omega = 0.1$ the empirical quantiles are now appearing to lie on a line farther away from the theoretical quantiles which implies that in the CARMA(2,1) case the variance in the simulations is clearly farther away from the asymptotic one than in the OU case. The same applies for the Variance Gamma case of Figure S.3 and the two-sided Poisson case of Figure S.5.

Hence, from our simulations of CARMA(2,1) processes we can conclude

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

that the orders of the CARMA processes and the particular autoregressive and moving average parameters appear not to really matter for the (qualitative) behaviour of the (trapezoidal approximation of the) truncated Fourier transform.

5 Conclusion and Outlook

We have obtained an asymptotic normality result for the (trapezoidal approximation of the) truncated Fourier transform under essentially minimal assumptions (i.e. second moments) and seen via a simulation study that this result approximates the finite sample behaviour very well, unless the frequency is too low compared to the length of the considered time interval. This suggests clearly that it should be very promising to develop statistical inference techniques for non-equidistantly sampled CARMA processes by considering continuous observation techniques and using numerical approximation schemes to compute the quantities of interest based on the observed non-equidistant data. The appropriate set-up to get convergence and asymptotic distribution results is to send the time horizon to infinity and to send at the same time the maximum distance of observation time points to zero.

Based on our results in this paper it seems natural to locally smooth the trapezoidal approximation of the truncated Fourier transform to get consistent

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

estimators of the spectral density and to use it in a Whittle type estimator for the AR and MA parameters. Our results suggest that under appropriate conditions one should be able to approximate a Whittle likelihood (and the resulting estimator) based on continuos observations by observations on a non-equidistant discrete grid. Considering this is beyond the scope of the present paper.

Supplementary Materials

As refered to in the main text, an online supplement is available containing most of the proofs and auxiliary results, as well as further details of the simulations studies.

Acknowledgements

The authors are very grateful to the Associate Editor and anonymous reviewers for their helpful comments, which considerably improved the paper. Furthermore, they would like to cordially thank Włodzimierz Fechner for careful reading of the manuscript and his helpful remarks.

The authors gratefully acknowledge the support of Deutsche Forschungsgemeinschaft (DFG) by research grant STE 2005/1-2.

References

D. Applebaum. *Lévy processes and stochastic calculus*. Cambridge University Press, Cambridge, 2nd edition, 2009.

REFERENCES

P. J. Brockwell. Lévy-driven CARMA processes. *Ann. Inst. Statist. Math.*, 53(96-120):113–124, 2001a.

P. J. Brockwell. *Continuous-time ARMA processes*, volume 19 of *in: Handbook of Statistics, Eds.: D.N. Shanbhag, C.R. Rao*. North-Holland, Amsterdam, 2001b.

P. J. Brockwell. *Lévy driven continuous-time ARMA processes*. in: Handbook of financial time series, Eds.: T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch. Heidelberg: Springer, 2009.

P. J. Brockwell and R. A. Davis. *Time series: theory and methods*. Springer, New York, 2006. Reprint of the 2nd (1991) edition.

P. J. Brockwell and E. Schlemm. Parametric estimation of the driving Lévy process of multivariate CARMA processes from discrete observations. *J. Multivariate Anal.*, 115:217–251, 2013.

P. J. Brockwell, R. A. Davis, and Y. Yang. Estimation for non-negative Lévy-driven CARMA processes. *J. Bus. Econom. Statist.*, 29(2):250–259, 2011.

J. L. Doob. The elementary Gaussian processes. *Ann. Math. Statistics*, 15:229–282, 1944.

V. Fasen. Statistical inference of spectral estimation for continuous-time MA processes with finite second moments. *Math. Methods Statist.*, 22(4):283–309, 2013.

V. Fasen and F. Fuchs. On the limit behavior of the periodogram of high-frequency sampled stable CARMA processes. *Stochastic Process. Appl.*, 123(1):229–273.

V. Fasen and F. Fuchs. Spectral estimates for high-frequency sampled continuous-time autoregressive moving average processes. *J. Time Series Anal.*, 34(5):532–551, 2013.

J. Gillberg. Frequency domain identification of continuous-time systems: Reconstruction and robustness.

Linköping Studies in Science and Technology. Dissertations, (1031), 2006.

K.-S. Lii and E. Masry. Model fitting for continuous-time stationary processes from discrete-time data. *J. Multivariate Anal.*, 41(1):56–79, 1992.

K.-S. Lii and E. Masry. Spectral estimation of continuous-time stationary processes from random sampling. *Stochastic Process. Appl.*, 52(1):39–64, 1994.

D. B. Madan, P. P. Carr, and E. C. Chang. The variance gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.

T. Marquardt and R. Stelzer. Multivariate CARMA processes. *Stochastic Process. Appl.*, 117(1):96–120, 2007.

P. E. Protter. *Stochastic integration and differential equations*. Springer-Verlag, Berlin, 2nd edition, 2004.

E. Schlemm and R. Stelzer. Multivariate CARMA processes, continuous-time state space models and complete regularity of the innovations of the sampled processes. *Bernoulli*, 18(1):46–63, 2012.

Institute of Mathematics, University of Silesia in Katowice, Bankowa 14, 40-007 Katowice, Poland and

Institute of Mathematical Finance, Ulm University, Helmholtzstraße 18, 89069 Ulm, Germany

E-mail: zfechner@gmail.com

Institute of Mathematical Finance, Ulm University, Helmholtzstraße 18, 89069 Ulm, Germany

E-mail: robert.stelzer@uni-ulm.de