Continuous-time locally stationary time series models

Annemarie Bitter, Robert Stelzer and Bennet Ströh[*]

Abstract

We adapt the classical definition of locally stationary processes in discrete-time (see e.g. [14]) to the continuous-time setting and obtain equivalent representations in the time and frequency domain. From this, a unique time-varying spectral density is derived using the Wigner-Ville spectrum. As an example, we investigate time-varying Lévy-driven state space processes, including the class of time-varying Lévy-driven CARMA processes. First, the connection between these two classes of processes is examined. Considering a sequence of time-varying Lévy-driven state space processes, we then give sufficient conditions on the coefficient functions that ensure local stationarity with respect to the given definition.

1 Introduction

To model non-stationary time series that behave locally in a stationary manner, Dahlhaus and others developed, starting with the seminal paper [11], a comprehensive theory and powerful estimation procedures, using a parameterized sequence of processes for the definition of local stationarity (see e.g. [12, 13, 15], or [14] for an overview). Noticeable examples include, for instance, ARMA processes with continuous coefficient functions (see [11]). More recently, also non-parametric approaches that allow for linear and non-linear locally stationary models were introduced and investigated in [3, 16, 17, 38, 39].

Despite this success, the above approaches have just been carried out for models defined on $\mathbb{Z}$, i.e. in a discrete-time framework. Surprisingly, there is so far no theory for locally stationary models defined on $\mathbb{R}$, i.e. in a continuous-time framework, available.

In this paper, we tackle this issue and define local stationarity for continuous-time models following the original definition from Dahlhaus [11]. We establish such a definition in the frequency and time domain and, as we consistently use L^2-integration theory (see e.g. [1] for an introduction), we readily obtain that both definitions are equivalent. Based on the definition in the frequency domain, we define a time-varying spectral density and show that it can be uniquely determined by a sequence of locally stationary processes, using the Wigner-Ville spectrum (see also [11]). This uniqueness is a powerful property, as it is known to pave the way for a likelihood approximation (comparable to the Whittle likelihood for stationary processes), leading to powerful estimation methods (see [13]).

---

[*]Ulm University, Institute of Mathematical Finance, Helmholtzstraße 18, 89069 Ulm, Germany. Emails: bitter.annemarie@gmail.com, robert.stelzer@uni-ulm.de, bennet.stroeh@uni-ulm.de.

As an example, we consider time-varying Lévy-driven state space processes, which include the continuous-time analog of time-varying ARMA (time-varying CARMA) processes. Lévy-driven CARMA processes are known to provide a flexible yet analytically tractable class of processes that have been applied to model a variety of phenomena from different areas [4, 23, 25].

In the time-invariant setting, it is known from [35] that the class of CARMA processes is equivalent to the class of Lévy-driven state space processes. While it is easy to see that also every time-varying CARMA process is a time-varying Lévy-driven state space process, we show that the inverse inclusion fails to hold, at least for non-continuous coefficient functions. This motivates to look at the class of time-varying Lévy-driven state space models.

More precisely, the paper is structured as follows. In Section 2, we first review the definition of local stationarity in the discrete-time framework. Then in Section 2.2, we summarize basic facts about Lévy processes and orthogonal random measures, including integration with respect to them.

The novel definition of local stationarity for continuous-time models both in the frequency and time domain is given in Section 3. Moreover, we investigate asymptotic distributional properties of such models and show that the autocovariance function evaluated at distinct points tends to zero.

In Section 4, we investigate time-varying state space processes in the context of local stationarity. We start with a simple example in Section 4.1, where we consider a sequence of time-varying CAR(1) processes and give sufficient conditions on the coefficient function such that the sequence is locally stationary according to the given definition. Section 4.2 and 4.3 are dedicated to general time-varying state space processes. First, in Section 4.2, the connection between the class of time-varying CARMA processes and time-varying state space processes is examined. Then, we give sufficient conditions for a sequence of time-varying state space processes to be locally stationary.

Finally, in Section 5 we investigate the time-varying spectral density and the Wigner-Ville spectrum of locally stationary processes.

2 Preliminaries

Throughout this paper, we denote the set of positive integers by $\mathbb{N}$, non-negative real numbers by $\mathbb{R}_0^+$, the set of $m \times n$ matrices over a ring R by $M_{m \times n}(R)$ and $\mathbf{1}_n$ stands for the $n \times n$ identity matrix. Given a complex number z, we denote the complex conjugate of z by $\bar{z}$. For square matrices $A, B \in M_{n \times n}(R)$, $[A, B] = AB - BA$ denotes the commutator of A and B, $Rank(A)$ the rank of A and $\sigma(B)$ the spectrum of B. We shortly write the transpose of a matrix $A \in M_{m \times n}(\mathbb{R})$ as A' and the adjoint of a matrix $B \in M_{m \times n}(\mathbb{C})$ as B^*. Norms of matrices and vectors are denoted by $\|\cdot\|$. If the norm is not further specified, we take the Euclidean norm or its induced operator norm, respectively. For a complex number $z \in \mathbb{C}$, the real part of z is denoted by $\mathfrak{Re}(z)$. The Borel σ-algebras are denoted by $\mathcal{B}(\cdot)$ and λ stands for the Lebesgue measure, at least in the context of measures. In the following, we will assume all stochastic processes and random variables to be defined on a common complete probability space $(\Omega, \mathcal{F}, P)$ equipped with an appropriate filtration if necessary. We simply write L^p to denote the space $L^p(\Omega, \mathcal{F}, P)$ and $L^p(X)$ to denote the space $L^p(X, \mathcal{B}(X), \lambda)$ for some set $X \subset \mathbb{R}$ with corresponding norms $\|\cdot\|_{L^p}$. The ring of continuous functions in t from $\mathbb{R}$ to $\mathbb{R}$ is denoted by $\mathcal{R}[t]$.

2.1 Locally stationary time series in discrete time

We follow the concept of local stationarity as established in [14] for discrete-time locally stationary time series models. There, the authors considered a parametric representation of a sequence of non-stationary time-varying processes either in the time or frequency domain, which has to satisfy certain regularity conditions.

In the following we briefly review the mathematical details of the aforementioned concepts as well as the most important results. To this end, we define the total variation of a function g on $[0, 1]$, denoted by $V(g)$, as

$$V(g) := \sup \left\{ \sum_{k=1}^{m} |g(x_k) - g(x_{k-1})|, 0 \leq x_0 < \ldots < x_m \leq 1, m \in \mathbb{N} \right\}$$

and for $\kappa > 0$ we define

$$\ell_\kappa(j) := \begin{cases} 1, & |j| \leq 1, \\ |j| \log^{1+\kappa} |j|, & |j| > 1 \end{cases}$$

for all $j \in \mathbb{Z}$. For further details on the following two definitions we refer to [14].

Definition 2.1. *Let $\{X_{t,T}, t = 1, \ldots, T\}_{T \in \mathbb{N}}$ be a sequence of stochastic processes. Then, $X_{t,T}$ is called locally stationary in the time domain, if there exists a representation*

$$X_{t,T} = \sum_{j=-\infty}^{\infty} a_{t,T,j} \varepsilon_{t-j}, \qquad T \in \mathbb{N}, \ t = 1, ..., T$$

where

(a) $\{\varepsilon_t, t \in \mathbb{Z}\}$ is an i.i.d. (independent identically distributed) sequence with $E[\varepsilon_t] = 0$ and $Var(\varepsilon_t) = 1$,

(b) for all $j \in \mathbb{Z}$ it holds

$$\sup_{\substack{t=1,\ldots,T \\ T \in \mathbb{N}}} |a_{t,T,j}| \leq \frac{K}{\ell_\kappa(j)},$$

where $\kappa, K > 0$ are constants and

(c) there exist functions $a_j(\cdot) : (0, 1] \to \mathbb{R}$, $j \in \mathbb{Z}$, satisfying

$$\sup_{u \in (0,1]} |a_j(u)| \leq \frac{K}{\ell_\kappa(j)}, \qquad \sup_{j \in \mathbb{Z}} \sum_{t=1}^{T} |a_{t,T,j} - a_j(\tfrac{t}{T})| \leq K \ and \qquad V(a_j(\cdot)) \leq \frac{K}{\ell_\kappa(j)} \tag{1}$$

for some constant K.

Definition 2.2. *Let $\{X_{t,T} : t = 1, \ldots, T\}_{T \in \mathbb{N}}$ be a sequence of stochastic processes. Then, $X_{t,T}$ is called locally stationary in the frequency domain with transfer functions $A^0_{t,T} : [-\pi, \pi] \to \mathbb{C}$, $T \in \mathbb{N}$, $t = 1, \ldots, T$, if it has the representation*

$$X_{t,T} = \int_{-\pi}^{\pi} e^{i\lambda t} A^0_{t,T}(\lambda) \xi(d\lambda) \qquad for \ all \ T \in \mathbb{N}, \ t = 1, \ldots, T,$$

(with the integrals existing in L^2) where

3

(a) $\xi(\lambda)$ is a stochastic process on $[-\pi, \pi]$ with mean zero and orthogonal increments,

(b) there exists a constant K and a function $A : [0,1] \times [-\pi, \pi] \to \mathbb{C}$, which is continuous in the first component satisfying $\overline{A(u, \lambda)} = A(u, -\lambda)$ and

$$\sup_{\substack{t=1,\ldots,T, \\ \lambda \in [-\pi,\pi]}} |A^0_{t,T}(\lambda) - A(\tfrac{t}{T}, \lambda)| \leq \frac{K}{T}, \qquad T \in \mathbb{N}. \tag{2}$$

Remark 2.3. *(a) Due to the smoothness conditions on the coefficient functions $a_j(u)$ and the transfer function $A(u, \lambda)$, the sequence $X_{t,T}$ shows a locally stationary behavior (see e.g. [12, Definition 2.1]).*

(b) For a comprehensive introduction to orthogonal increment processes, orthogonal random measures and the related L^2-integration theory we refer to [8].

(c) We note that the given definitions of local stationarity in the time and frequency domain are not equivalent.
However, using the spectral representation of the noise $\varepsilon_t = \int_{(-\pi,\pi]} \frac{1}{\sqrt{2\pi}} e^{i\lambda t} \xi(d\lambda)$ (see [8]), the Fourier transform allows for the following connections (see [13, Remark 2.2]) between the two concepts. It holds

$$A^0_{t,T}(\lambda) = \frac{1}{\sqrt{2\pi}} \sum_{j=-\infty}^{\infty} a_{t,T,j} e^{-i\lambda j}, \qquad\qquad A(u, \lambda) = \frac{1}{\sqrt{2\pi}} \sum_{j=-\infty}^{\infty} a_j(u) e^{-i\lambda j},$$

$$a_{t,T,j} = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} A^0_{t,T}(\lambda) e^{i\lambda j} d\lambda \text{ and} \qquad\qquad a_j(u) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} A(u, \lambda) e^{i\lambda j} d\lambda,$$

since $A^0_{t,T}(\cdot) \in L^2([-\pi, \pi], \mathbb{C})$ and $a_{t,T,j} \in \ell^2$.
Necessary conditions for Definition 2.1 and 2.2 to be equivalent can be found in [15, Remark 2.2]. In particular, this includes additional smoothness assumptions on $A(u, \lambda)$ and a stronger version of the second condition in (1).

The following two propositions give further insight into Definition 2.2 and the notion of local stationarity.

Proposition 2.4. *Let $X_{t,T}$ be a locally stationary process in the frequency domain and $\{T_n\}_{n\in\mathbb{N}} \subset \mathbb{N}$ an increasing sequence. If $sT_n \in \{1,\ldots,T_n\}$ for some fixed $s \in [0,1]$ and all $n > n_0$, $n_0 \in \mathbb{N}$, then it holds*

$$A^0_{sT_n, T_n}(\cdot) \xrightarrow[n\to\infty]{L^2} A(s, \cdot).$$

Proof. Follows directly from (2). □

For instance, the choice $T_n = 2^n$ and $s = k/2^{n_0}$ for some $n_0 \in \mathbb{N}$ and $k \in \{1,\ldots,T_{n_0}\}$ suits the conditions of Proposition 2.4.

Proposition 2.5. *Let $X_{t,T}$ be a locally stationary process in the frequency domain with associated orthogonal increment process $\{\xi(\lambda), \lambda \in [-\pi, \pi]\}$ corresponding to an i.i.d. noise (i.e.*

$\varepsilon_t = \int_{(-\pi,\pi]} \frac{1}{\sqrt{2\pi}} e^{i\lambda t} \xi(d\lambda)$ *defines an i.i.d. noise) and* $\{T_n\}_{n\in\mathbb{N}} \subset \mathbb{N}$ *an increasing sequence. If* $sT_n \in \{1, ..., T_n\}$ *for some fixed* $s \in [0, 1]$ *and all* $n > n_0$, $n_0 \in \mathbb{N}$, *then it holds*

$$X_{sT_n,T_n} \xrightarrow[n\to\infty]{d} \int_{-\pi}^{\pi} A(s, \lambda) \xi(d\lambda).$$

Proof. First observe that every time series of the form $\int_{-\pi}^{\pi} e^{i\lambda t} A(\lambda) \xi(d\lambda)$, $t \in \mathbb{Z}$, where ξ is an orthogonal increment process coming from an i.i.d. noise, is strictly stationary. Thus,

$$\int_{-\pi}^{\pi} e^{i\lambda t_1} A^0_{sT_n,T_n}(\lambda) \xi(d\lambda) \overset{d}{=} \int_{-\pi}^{\pi} e^{i\lambda t_0} A^0_{sT_n,T_n}(\lambda) \xi(d\lambda)$$

for all $t_0, t_1 \in \mathbb{Z}$. In particular, for $t_1 = sT_n$, where $s \in [0, 1]$ such that $sT_n \in \{1, \ldots, T\}$, and $t_0 = 0$ we obtain

$$X_{sT_n,T_n} = \int_{-\pi}^{\pi} e^{i\lambda sT_n} A^0_{sT_n,T_n}(\lambda) \xi(d\lambda) \overset{d}{=} \int_{-\pi}^{\pi} A^0_{sT_n,T_n}(\lambda) \xi(d\lambda).$$

The remainder follows from Proposition 2.4 and the continuity of the stochastic integral in mean square and thus in distribution with respect to the integrand. $\square$

Remark 2.6. *A noticeable class of processes that are locally stationary in the time and frequency domain are time-varying AR(p) processes with continuous coefficient functions. For the mathematical details of this result we refer to [11, p. 147].*

Among the variety of different concepts for local stationarity in the literature, we mention the results from [16] and [38]. In [38], the author considers a triangular array $X_{t,T}$, $T \in \mathbb{N}$, $t = 1, \ldots, T$ *to be locally stationary, if for each* $u \in [0, 1]$ *there exists a strictly stationary process* $\{X_t(u), t = 1, \ldots, T\}$ *such that almost surely*

$$|X_{t,T} - X_t(u)| \leq (|\tfrac{t}{T} - u| + \tfrac{1}{T}) U_{t,T}(u),$$

where $U_{t,T}(u)$ *are positive random variables satisfying* $E[(U_{t,T}(u))^\rho] < \infty$ *for some* $\rho > 0$ *uniformly in* u, t *and* T. *Time-varying AR(p) processes with continuous coefficient functions can also be embedded in this framework using similar arguments as in [17].*

More recently, the authors in [16] developed a general theory for locally stationary processes based on stationary approximations. Similarly to [38], it is assumed that there exists a strictly stationary process $\{X_t(u), t = 1, \ldots, T\}$ *such that for some* $q, C > 0$

$$\|X_t(u) - X_t(v)\|_{L^q} \leq C |u - v| \quad \text{and} \quad \|X_{t,T} - X_t(\tfrac{t}{T})\|_{L^q} \leq \tfrac{C}{T}, \tag{3}$$

uniformly in $t = 1, \ldots, T$ *and* $u, v \in [0, 1]$. *Based on these approximations the authors established asymptotic results as a law of large numbers and a central limit theorem, which, in turn, are used to derive asymptotic results for a maximum likelihood estimator (see [16, Section 5]). Again, time-varying AR(p) processes with continuous coefficient functions can be embedded in this framework. Recently, this work has been extended to models with infinite memory in [3]. In view of the statistical results obtained from the approximations (3), a possible characterization of local stationarity in terms of similar approximations for continuous-time models will be the topic of future work.*

2.2 Lévy processes and orthogonal random measures

In this section we lay the foundation for the definition of continuous-time locally stationary processes and briefly review Lévy processes, orthogonal random measures and cover basic results including stochastic integration with respect to Lévy processes and orthogonal random measures. For further insight we refer to [1] and [33].

Definition 2.7. *A real-valued stochastic process $L = \{L(t), t \in \mathbb{R}_0^+\}$ is called Lévy process if*

(a) $L(0) = 0$ *almost surely,*

(b) *for any $n \in \mathbb{N}$ and $t_0 < t_1 < t_2 < \cdots < t_n$, the random variables $(L(t_0), L(t_1) - L(t_0), \ldots, L(t_n) - L(t_{n-1}))$ are independent,*

(c) *for all $s, t \geq 0$, the distribution of $L(s + t) - L(s)$ does not depend on s and*

(d) L *is stochastically continuous.*

Theorem 2.8. *Let $L = \{L(t), t \geq 0\}$ be a real-valued Lévy process. Then, $L(1)$ is an infinitely divisible real-valued random variable with characteristic triplet (γ, Σ, ν), where $\gamma \in \mathbb{R}$, $\Sigma > 0$ and ν is a Lévy measure on $\mathbb{R}$, i.e. $\nu(0) = 0$ and $\int_{\mathbb{R}} (1 \wedge |x|^2) \nu(dx) < \infty$. The characteristic function of $L(t)$ is given by*

$$\varphi_{L(t)}(z) = E[e^{izL(t)}] = \exp(t\Psi_L(z)),$$
$$\Psi_L(z) = \left(i\gamma z - \frac{\Sigma z^2}{2} + \int_{\mathbb{R}} \left(e^{izx} - 1 - izx \mathbb{1}_Z(x) \right) \nu(dx) \right), \tag{4}$$

where $z \in \mathbb{R}$ and $Z = \{x \in \mathbb{R}, |x| \leq 1\}$.

In the remainder we work with two-sided Lévy process, i.e. $L(t) = L_1(t) \mathbb{1}_{\{t \geq 0\}} - L_2(-t) \mathbb{1}_{\{t < 0\}}$, where L_1 and L_2 are independent copies of a one-sided Lévy process. Throughout this paper, it will be assumed that

$$E[L(1)] = \gamma + \int_{|x| > 1} x \nu(dx) = 0 \quad \text{and} \quad E[L(1)^2] = \Sigma + \int_{x \in \mathbb{R}} x^2 \nu(dx) < \infty. \tag{5}$$

Thus, the above assumptions on the Lévy process imply that $\int_{\mathbb{R}} x^2 \nu(dx) < \infty$. Occasionally, we will denote $\Sigma_L := Var(L(1)) = \Sigma + \int_{x \in \mathbb{R}} x^2 \nu(dx)$.

If the Lévy process satisfies (5) and $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a $\mathcal{B}(\mathbb{R} \times \mathbb{R}) - \mathcal{B}(\mathbb{R})$-measurable function satisfying $f(t, \cdot) \in L^2(\mathbb{R})$, then the integral $X(t) = \int_{\mathbb{R}} f(t, s) L(ds)$, $t \in \mathbb{R}$ exists in L^2 (see e.g. [25]).

Definition 2.9 ([21, Definition 2.3.5]). *A family $\{\xi(\Delta)\}_{\Delta \in \mathcal{B}(\mathbb{R})}$ of $\mathbb{C}$-valued random variables is called an orthogonal random measure (ORM) if*

(a) $\xi(\Delta) \in L^2(\mathcal{B}(\mathbb{R}), \mathbb{C})$ *for all bounded $\Delta \in \mathcal{B}(\mathbb{R})$,*

(b) $\xi(\emptyset) = 0$,

(c) $\xi(\Delta_1 \cup \Delta_2) = \xi(\Delta_1) + \xi(\Delta_2)$ *a.s. whenever $\Delta_1 \cap \Delta_2 = \emptyset$ and*

(d) $F : \mathcal{B}(\mathbb{R}) \to \mathbb{C}$ *such that $F(\Delta) = E[\xi(\Delta)\overline{\xi(\Delta)}]$ defines a σ-additive positive definite measure and it holds that $E[\xi(\Delta_1)\overline{\xi(\Delta_2)}] = F(\Delta_1 \cap \Delta_2)$ for all $\Delta_1, \Delta_2 \in \mathcal{B}(\mathbb{R})$.*

F is referred to as the spectral measure of ξ.

Theorem 2.10 ([25, Theorem 3.5]). *Let L be a two-sided Lévy process satisfying (5). Then, there exists an ORM Φ_L with spectral measure F_L, such that*

(a) $E[\Phi_L(\Delta)] = 0$ *for any bounded $\Delta \in \mathcal{B}(\mathbb{R})$,*

(b) $F_L(dt) = \frac{\Sigma_L}{2\pi} dt$ *and*

(c) Φ_L *is uniquely determined by $\Phi_L([a,b)) := \int_{-\infty}^{\infty} \frac{e^{-i\mu a} - e^{-i\mu b}}{2\pi i \mu} L(d\mu)$.*

In the proof of the above theorem the standard theory of Fourier transforms on $L^2(\mathbb{R})$ (see e.g. [10, Chapter 2] for an introduction) is used to show

$$\int_{-\infty}^{\infty} \varphi(\mu) \Phi_L(d\mu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{\varphi}(u) L(du) \tag{6}$$

for all complex functions $\varphi \in L^2(\mathbb{R})$ and their (inverse) Fourier transforms $\widehat{\varphi}$, where

$$\widehat{\varphi}(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\mu u} \varphi(\mu) d\mu \quad \text{and} \quad \varphi(\mu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\mu u} \widehat{\varphi}(u) du.$$

We also recall that for two complex functions $f, g \in L^2(\mathbb{R})$ and their Fourier transforms $\widehat{f}, \widehat{g}$, it follows that $\hat{f}, \hat{g} \in L^2(\mathbb{R})$ and, due to [31, p. 189],

$$\int_{-\infty}^{\infty} f(\mu)\overline{g(\mu)} d\mu = \int_{-\infty}^{\infty} \widehat{f}(u)\overline{\widehat{g}(u)} du.$$

3   Locally stationary processes in continuous-time

Analogously to Section 2.1 one can define a (stationary) stochastic process $\{Y(t)\}_{t \in \mathbb{R}}$ via the representation as a linear process or the spectral representation, i.e.

$$Y(t) = \int_{\mathbb{R}} g(t-u) L(du) \quad \text{or} \quad Y(t) = \int_{\mathbb{R}} e^{i\mu t} A(\mu) \Phi_L(d\mu), \qquad t \in \mathbb{R},$$

where g and A are square integrable functions and L is a two-sided Lévy process with corresponding ORM Φ_L. As we consistently use L^2-integrals to define the process both in the time and the frequency domain and the Fourier transform is an isometry on L^2 the two definitions are equivalent. Hence, from (6) it follows that the transfer function A and the kernel g satisfy

$$g(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\mu u} A(\mu) d\mu \quad \text{and} \quad A(\mu) = \int_{-\infty}^{\infty} e^{-i\mu v} g(v) dv.$$

Now, we allow the kernel function and the transfer function be time dependent, leading to

$$Y(t) = \int_{\mathbb{R}} e^{i\mu t} A(t, \mu) \Phi_L(d\mu) = \int_{\mathbb{R}} g(t, t-u) L(du), \qquad t \in \mathbb{R},$$

where

$$g(t, u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\mu u} A(t, \mu) d\mu \quad \text{and} \quad A(t, \mu) = \int_{-\infty}^{\infty} e^{-i\mu u} g(t, u) du.$$

As we are interested in real-valued processes we demand g to be real-valued or equivalently that $\overline{A(\mu)} = A(-\mu)$ for all $\mu \in \mathbb{R}$.

To be able to define local stationarity analogously to Section 2.1, not only a time varying representation, but also a sequence of stochastic processes is needed. The intuitive idea is to take a limiting kernel g and a sequence of kernels g_N^0 defining the processes in the time domain such that

$$\left\| g_N^0(t, \cdot) - g(\tfrac{t}{N}, \cdot) \right\|_{L^2} \xrightarrow[N \to \infty]{} 0.$$

However, for the limiting (stationary) process we prefer to fix a time $t \in \mathbb{R}$ rather than dealing with fractions $\frac{t}{N}$. By replacing t by Nt this leads to the following definition.

Definition 3.1. *A sequence of stochastic processes $\{Y_N(t), t \in \mathbb{R}\}_{N \in \mathbb{N}}$ is said to be locally stationary in the time domain, if it can be represented as*

$$Y_N(t) = \int_{\mathbb{R}} g_N^0(Nt, Nt - u) L(du), \quad \text{for all } t \in \mathbb{R}, N \in \mathbb{N},$$

where L is a two-sided Lévy process and the kernel functions $g_N^0 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ satisfy

  (a) *$g_N^0(Nt, \cdot) \in L^2(\mathbb{R})$ for all $t \in \mathbb{R}, N \in \mathbb{N}$ and*

  (b) *there exists a (local/limiting kernel) function $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that the mapping $\mathbb{R} \to L^2(\mathbb{R})$, $t \mapsto g(t, \cdot)$ is continuous and*

$$g_N^0(Nt, \cdot) \xrightarrow[N \to \infty]{L^2} g(t, \cdot) \quad \text{for all } t \in \mathbb{R}.$$

Definition 3.2. *A sequence of stochastic processes $\{Y_N(t), t \in \mathbb{R}\}_{N \in \mathbb{N}}$ is said to be locally stationary in the frequency domain, if it can be represented as*

$$Y_N(t) = \int_{\mathbb{R}} e^{i\mu Nt} A_N^0(Nt, \mu) \Phi_L(d\mu), \quad \text{for all } t \in \mathbb{R}, N \in \mathbb{N}, \tag{7}$$

where Φ_L is the ORM of a two-sided Lévy process L and the transfer functions $A_N^0 : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{C}$ satisfy

  (a) *$A_N^0(Nt, \cdot) \in L^2$ for all $t \in \mathbb{R}, N \in \mathbb{N}$,*

  (b) *$\overline{A_N^0(\cdot, \cdot)} = A_N^0(\cdot, -\cdot)$ and*

  (c) *there exists a (local/limiting transfer) function $A : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ with $\overline{A(\cdot, \cdot)} = A(\cdot, -\cdot)$ such that the mapping $\mathbb{R} \mapsto L^2(\mathbb{R})$, $t \mapsto A(t, \cdot)$ is continuous and*

$$A_N^0(Nt, \cdot) \xrightarrow[N \to \infty]{L^2} A(t, \cdot), \quad \text{for all } t \in \mathbb{R}.$$

In contrast to the discrete time case it is now irrelevant whether we use the definition in the time or the frequency domain. Therefore, we will just speak of "locally stationary" in both cases.

Proposition 3.3. *The Definitions 3.1 and 3.2 are equivalent. Moreover, the relationship between the (limiting) transfer function and the (limiting) kernel, using their Fourier transforms, is given by*

$$A_N^0(Nt, \mu) = \int_{-\infty}^{\infty} e^{-i\mu u} g_N^0(Nt, u) du, \qquad\qquad A(t, \mu) = \int_{-\infty}^{\infty} e^{-i\mu u} g(t, u) du,$$

$$g_N^0(Nt, u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\mu u} A_N^0(Nt, \mu) d\mu \text{ and} \qquad\qquad g(t, u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\mu u} A(t, \mu) d\mu.$$

Proof. The result follows immediately from the Definitions 3.1 and 3.2 using Plancherel's theorem. $\qquad\square$

The following lemma provides sufficient conditions for the continuity conditions on the mappings $t \mapsto A(t, \cdot)$ and $t \mapsto g(t, \cdot)$ from Definition 3.1 and 3.2.

Lemma 3.4. *Let $A : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ be a function, which is continuous in the first argument such that for all $t \in \mathbb{R}$ there exists an $\epsilon_t > 0$ and a real function $f_t \in L^2(\mathbb{R})$ such that $|A(s, \cdot)| \leq f_t(\cdot)$ for all $s \in [t - \epsilon_t, t + \epsilon_t]$. Then, the mapping $\mathbb{R} \to L^2(\mathbb{R})$, $t \mapsto A(t, \cdot)$ is continuous.*

Proof. Straightforward application of the dominated convergence theorem. $\qquad\square$

In principle it is possible to replace the Lévy process by a process with weakly stationary uncorrelated increments and the ORM induced by the Lévy process by an arbitrary ORM. The resulting processes would be (locally) weakly stationary. However, to us it seems at the moment not worthwhile to pursue this any further for the following reason.

To derive a continuous-time analogue of Proposition 2.5 the stationary and independent increments of the driving Lévy process L are essential. Therefore, also the orthogonal random measure in Definition 3.2 has to be generated by a stochastic process on $\mathbb{R}$ with independent and stationary increments, i.e by a Lévy process.

We note that this also ensures for all $\tilde{t} \in \mathbb{R}$ that the limiting process $Y_{\tilde{t}}(t) = \int_{\mathbb{R}} g(\tilde{t}, t - u) L(du)$ is strictly stationary. The next proposition provides the aforementioned continuous-time analogue of Proposition 2.5.

Proposition 3.5. *Let $\{Y_N(t), t \in \mathbb{R}\}_{N \in \mathbb{N}}$ be a locally stationary process. Then, for fixed $t \in \mathbb{R}$*

$$Y_N(t) \xrightarrow[n \to \infty]{d} \int_{\mathbb{R}} A(t, \mu) \Phi_L(d\mu) = \int_{\mathbb{R}} g(t, -u) L(du).$$

Proof. For $t \in \mathbb{R}$ we obtain, using a stationarity argument,

$$Y_N(t) = \int_{\mathbb{R}} e^{i\mu Nt} A_N^0(Nt, \mu) \Phi_L(d\mu) \stackrel{d}{=} \int_{\mathbb{R}} A_N^0(Nt, \mu) \Phi_L(d\mu).$$

The remainder follows from the continuity of the stochastic integral in mean square and thus in distribution with respect to the integrand. $\qquad\square$

Proposition 3.6. *Let $\{Y_N(t), t \in \mathbb{R}\}_{N \in \mathbb{N}}$ be a locally stationary sequence and $t_1, t_2 \in \mathbb{R}$ such that $t_1 \neq t_2$. Then, $Y_N(t_1)$ and $Y_N(t_2)$ are asymptotically uncorrelated, i.e. $Cov(Y_N(t_1), Y_N(t_2)) \to 0$ as $N \to \infty$.*

The intuition behind this proposition is that the kernel functions $g_N^0(Nt, Nt - \cdot)$ are square integrable and therefore roughly vanish if the second argument tends to infinity. For $t_1 \neq t_2$ the difference between Nt_1 and Nt_2 increases for $N \to \infty$. Therefore, for large N, the bulks of the kernels for t_1 and t_2 rest on far apart segments of the Lévy process, which has independent increments.

Proof. Let $Y_N(t) = \int_{\mathbb{R}} g_N^0(Nt, Nt - u) L(du)$ be a sequence of locally stationary processes. Without loss of generality we assume that $t_1 > t_2$ and set $h = t_1 - t_2 > 0$. It is sufficient to show that for all $t_1 \in \mathbb{R}$ and $\varepsilon > 0$ there exists an $N_0 \in \mathbb{N}$ such that for all $N > N_0$

$$|Cov(Y_N(t_2), Y_N(t_2 + h))| = \Sigma_L \left| \int_{\mathbb{R}} g_N^0(Nt_2, Nt_2 - u) g_N^0(N(t_2 + h), N(t_2 + h) - u) du \right| < \varepsilon.$$

Let $t \in \mathbb{R}$ and define $\mathcal{E}$ as the set of all elementary real functions in $L^2(\mathbb{R})$, i.e.

$$\mathcal{E} = \left\{ f \in L^2(\mathbb{R}) : f = \sum_{i=1}^n c_i \mathbb{1}_{[a_i, b_i)}, n \in \mathbb{N}, c_i \in \mathbb{R}, -\infty < a_i < b_i < \infty, i = 1, \dots, n \right\}.$$

Then for all $\eta > 0$ there exists $N_1 \in \mathbb{N}$ and elementary functions $\hat{g}(t_2, \cdot), \hat{g}(t_2 + h, \cdot) \in \mathcal{E}$ such that

$$\|g(t_2, \cdot) - \hat{g}(t_2, \cdot)\|_{L^2} < \eta, \qquad\qquad \|g(t_2 + h, \cdot) - \hat{g}(t_2 + h, \cdot)\|_{L^2} < \eta,$$
$$\left\|g_N^0(Nt_2, \cdot) - g(t_2, \cdot)\right\|_{L^2} < \eta \text{ and} \qquad \left\|g_N^0(N(t_2 + h), \cdot) - g((t_2 + h), \cdot)\right\|_{L^2} < \eta,$$

using [30, Prop. 6.8]. For the remainder of the proof, it will be assumed that $N > N_1$. Thus,

$$\left\|g_N^0(N(t_2 + h), \cdot)\right\|_{L^2} \leq \eta + \|g(t_2 + h, \cdot)\|_{L^2} \text{ and} \qquad \left\|g_N^0(N(t_2), \cdot)\right\|_{L^2} \leq \eta + \|g(t_2, \cdot)\|_{L^2}.$$

We define the constant $K = \eta + \max\{\|g(t_2, \cdot)\|_{L^2}, \|g(t_2 + h, \cdot)\|_{L^2}\} < \infty$. Then, using the triangle and Cauchy-Schwartz's inequality shows

$$
\begin{aligned}
|Cov(Y_N(t_2), Y_N(t_2 + h))| &= \Sigma_L \left| \int_{\mathbb{R}} g_N^0(Nt_2, Nt_2 - u) g_N^0(N(t_2 + h), N(t_2 + h) - u) du \right| \\
&\leq \Sigma_L \Big(\|g_N^0(Nt_2, \cdot) - g(t_2, \cdot)\|_{L^2} \|g_N^0(N(t_2 + h), \cdot)\|_{L^2} \\
&\quad + \|g(t_2, \cdot)\|_{L^2} \|g_N^0(N(t_2 + h), \cdot) - g(t_2 + h, \cdot)\|_{L^2} \\
&\quad + \|g(t_2, \cdot) - \hat{g}(t_2, \cdot)\|_{L^2} \|g(t_2 + h, \cdot)\|_{L^2} \\
&\quad + \|\hat{g}(t_2, \cdot)\|_{L^2} \|g(t_2 + h, \cdot) - \hat{g}(t_2 + h, \cdot)\|_{L^2} \\
&\quad + \int_{\mathbb{R}} |\hat{g}(t_2, Nt_2 - u)| \, |\hat{g}(t_2 + h, N(t_2 + h) - u)| du \Big) \\
&\leq \Sigma_L \Big(4\eta K + \int_{\mathbb{R}} |\hat{g}(t_2, Nt_2 - u)| \, |\hat{g}(t_2 + h, N(t_2 + h) - u)| du \Big),
\end{aligned}
$$

where the last integral tends to zero for $N \to \infty$ by using the dominated convergence theorem and noting that the elementary functions $\hat{g}$ have bounded support. $\qquad\square$

4 Classes of locally stationary processes in continuous-time

In this section, we consider sequences of time-varying CARMA processes, for which we derive sufficient conditions for local stationarity.

4.1 Locally stationary CAR(1) processes

The simplest Lévy-driven CARMA process is the Lévy-driven CAR(1) or Ornstein-Uhlenbeck type process.

For a constant coefficient $a > 0$ a CAR(1) process is the stationary solution to the stochastic differential equation $dY(t) = -aY(t)dt + L(dt)$, which can be expressed as

$$Y(t) = \int_{-\infty}^{t} e^{-a(t-u)} L(du).$$

We replace the constant a by a time-varying function $a(t)$ and arrive at a so called time-varying CAR(1) process, which is given by

$$Y(t) = \int_{-\infty}^{t} e^{-\int_{u}^{t} a(s)ds} L(du).$$

Additional rescaling results in a sequence of time-varying CAR(1) processes that could be locally stationary. We consider the sequence of stochastic processes $\{Y_N(t), t \in \mathbb{R}\}_{N \in \mathbb{N}}$ defined by

$$Y_N(t) = \int_{-\infty}^{Nt} e^{-\int_{u}^{Nt} a(\frac{s}{N})ds} L(du), \tag{8}$$

where $a : \mathbb{R} \to \mathbb{R}$ is a continuous coefficient function such that $u \mapsto e^{-\int_{u}^{Nt} a(\frac{s}{N})ds}$ is square integrable for all $t \in \mathbb{R}$, $N \in \mathbb{N}$ and L is a two-sided Lévy process. Recall that the Lévy process satisfies (5). In view of Definition 3.1, we obtain from (8) that

$$
\begin{aligned}
g_N^0(Nt, Nt - u) &= \mathbb{1}_{\{Nt-u \geq 0\}} e^{-\int_{u}^{Nt} a(\frac{s}{N})ds} &&= \mathbb{1}_{\{Nt-u \geq 0\}} e^{-\int_{-(Nt-u)}^{0} a(\frac{s}{N}+t)ds} \quad \text{and} \\
A_N^0(Nt, \mu) &= \int_{\mathbb{R}} e^{-i\mu u} g_N^0(Nt, u) du &&= \int_{\mathbb{R}} e^{-i\mu v} \mathbb{1}_{\{v \geq 0\}} e^{-\int_{-v}^{0} a(\frac{s}{N}+t)ds} dv.
\end{aligned} \tag{9}
$$

Proposition 4.1. *Let $\{Y_N(t), t \in \mathbb{R}\}_{N \in \mathbb{N}}$ be a sequence of time-varying CAR(1) processes as defined in (8). If*

(C1) *$a(\cdot)$ is continuous and*

(C2) *for every $T \in \mathbb{R}^+$ there exists $\varepsilon_T > 0$ such that $a(s) \geq \varepsilon_T$ for all $s \leq T$,*

then $Y_N(t)$ is locally stationary, where the limiting kernel g and limiting transfer function are given by

$$g(t, u) = \mathbb{1}_{\{u \geq 0\}} e^{-a(t)u} \quad \text{and} \quad A(t, \mu) = \int_{\mathbb{R}} e^{-i\mu u} \mathbb{1}_{\{u \geq 0\}} e^{-a(t)u} du.$$

Proof. For all $t \in \mathbb{R}$ it holds

$$
\begin{aligned}
\left\| g_N^0(Nt, \cdot) - g(t, \cdot) \right\|_{L^2}^2 &= \left\| g_N^0(Nt, Nt - \cdot) - g(t, Nt - \cdot) \right\|_{L^2}^2 \\
&= \int_{\mathbb{R}} \left| \mathbb{1}_{\{Nt-u \geq 0\}} e^{-\int_{-(Nt-u)}^{0} a(\frac{s}{N}+t)ds} - \mathbb{1}_{\{Nt-u \geq 0\}} e^{-a(t)(Nt-u)} \right|^2 du \\
&= \int_{\mathbb{R}} \mathbb{1}_{\{u \leq 0\}} \left| e^{-\int_{u}^{0} a(\frac{s}{N}+t)ds} - e^{a(t)u} \right|^2 du \xrightarrow[N \to \infty]{} 0,
\end{aligned}
$$

using the dominated convergence theorem. For the inner integral the continuity of a on a compact set is sufficient for an application of the dominated convergence theorem. As majorant for the outer integral we consider $u \mapsto 4\mathbb{1}_{\{u \leq 0\}} e^{2\varepsilon_t u} \in L^2$. The demanded L^2-continuity of the limiting kernel g can be obtained similarly, using Lemma 3.4. $\qquad \square$

11

Remark 4.2. *Condition (C1) is intrinsically related to the continuity of the limiting kernel demanded in the definition of local stationarity. (C2) is obviously satisfied if $a(\cdot)$ is bounded away from zero. However, as time goes to infinity $a(\cdot)$ may go to zero arbitrarily fast. The latter is clearly connected to the fact that our time-varying CAR processes are causal by definition. It may be possible to weaken (C2) and allow $a(\cdot)$ also to approach 0 as time goes to minus infinity. Then, the convergence to zero must be slow enough for all integrals to exist in L^2. Carrying this out in detail appears rather intricate and not of relevance for the applications of locally stationary CAR(1) processes.*

4.2 Time-varying CARMA(p,q) processes and time-varying state space models

Consider $p, q \in \mathbb{N}$, where $p > q$. The formal differential equation for a time-varying Lévy-driven CARMA(p, q) process is given by

$$p(t, D)Y(t) = q(t, D)DL(t), \text{ i.e.}$$

$$D^p Y(t) + a_1(t)D^{p-1}Y(t) + \ldots + a_p(t)Y(t) = b_0(t)DL(t) + b_1(t)D^2 L(t) + \ldots + b_q(t)D^{q+1}L(t),$$

where D denotes the differential operator with respect to time and $L(t)$ is a two sided Lévy process satisfying (5). For continuous functions $a_i(t), b_i(t)$, $i = 1, \ldots, p$, where $b_i(t) = 0$ for all $i > q$, the polynomials

$$p(t, z) = z^p + a_1(t)z^{p-1} + \ldots + a_{p-1}(t)z + a_p(t) \text{ and}$$
$$q(t, z) = b_0(t) + b_1(t)z + \ldots + b_{q-1}(t)z^{q-1} + b_q(t)z^q \tag{10}$$

are called autoregressive (AR) and moving average (MA) polynomials. For a rigorous definition we interpret the differential equations to be equivalent to the state space representation

$$Y(t) = \mathcal{B}(t)'\mathcal{X}(t), \text{ and}$$
$$d\mathcal{X}(t) = \mathcal{A}(t)\mathcal{X}(t)dt + \mathcal{C}L(dt), \quad t \in \mathbb{R} \tag{11}$$

with

$$\mathcal{A}(t) = \begin{pmatrix} 0 & 1 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & & 1 \\ -a_p(t) & -a_{p-1}(t) & \ldots & -a_1(t) \end{pmatrix} \in M_{p \times p}(\mathcal{R}[t]) \text{ and}$$

$$\mathcal{B}(t) = \begin{pmatrix} b_0(t) \\ b_1(t) \\ \vdots \\ b_{p-1}(t) \end{pmatrix} \in M_{p \times 1}(\mathcal{R}[t]), \qquad \mathcal{C} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in M_{p \times 1}(\mathbb{R}),$$

where $\mathcal{R}[t]$ denotes the ring of continuous functions in t from $\mathbb{R}$ to $\mathbb{R}$.

It is obvious that (11) has a unique solution when one fixes the value $X(t_0)$ at some point $t_0 \in \mathbb{R}$. For a Brownian motion as driving noise such equations were investigated in [37, Section 2.1.1.]. Provided the integrals exist in L^2, it can be shown that a solution is given by

$$\mathcal{X}(t) = \int_{-\infty}^{t} \Psi(t, s)\mathcal{C}L(ds) \quad \text{and} \quad Y(t) = \mathcal{B}(t)' \int_{-\infty}^{t} \Psi(t, s)\mathcal{C}L(ds), \tag{12}$$

12

where $\Psi(t, t_0)$ is the unique matrix solution of the homogeneous initial value problem (IVP) $\frac{d}{dt}\Psi(t, t_0) = \mathcal{A}(t)\Psi(t, t_0)$, where $\Psi(t_0, t_0) = \mathbf{1}_p$ for all $t > t_0$ (see [7, Section 3 and 4]). The transition matrix Ψ satisfies $\Psi(t, t_0) = \Psi(t, u)\Psi(u, t_0)$ for all $t > u > t_0$ (see [7, Section 4, Theorem 2]). In particular, the integrals in (12) are well-defined (see Section 2.2), if there exist $\gamma, \lambda > 0$, such that

$$\|\Psi(t, t_0)\| \le \gamma e^{-\lambda(t-t_0)} \qquad \text{for all } t, t_0 \text{ with } t \ge t_0.$$

This condition corresponds to uniform exponential stability of the state space model in (11) and will be explained more detailed in Section 4.3.

The usual integral representation of stationary causal CARMA processes motivates the following definition.

Definition 4.3. *A solution $\{Y(t)\}_{t \in \mathbb{R}}$ of the observation and state equations (11) in the form (12) is called a time-varying Lévy-driven CARMA(p,q) process (tvCARMA(p,q)).*

For some initial time $t_0 \in \mathbb{R}$ the process satisfies the relation (see [37, Section 2.1.1.])

$$\mathcal{X}(t) = \Psi(t, t_0)\left(\mathcal{X}(t_0) + \int_{t_0}^{t} \Psi(u, t_0)^{-1}\mathcal{C}L(du)\right). \tag{13}$$

From [2, Remark 2] it follows that if for all $t, t_0 \in \mathbb{R}$ and $t > t_0$

$$\mathcal{A}(t)\int_{t_0}^{t} \mathcal{A}(s)ds = \int_{t_0}^{t} \mathcal{A}(s)ds\mathcal{A}(t), \tag{14}$$

then $\Psi(t, t_0) = e^{\int_{t_0}^{t} \mathcal{A}(s)ds}$.

If assumption (14) does not hold, $\Psi(t, t_0)$ can be expressed by the Peano-Baker series (see [2, Section 2])

$$\Psi(t, t_0) = \mathbf{1}_p + \int_{t_0}^{t} \mathcal{A}(\tau_1)d\tau_1 + \int_{t_0}^{t} \mathcal{A}(\tau_1)\int_{t_0}^{\tau_1} \mathcal{A}(\tau_2)d\tau_2 d\tau_1 + \ldots = \sum_{n=0}^{\infty} \mathcal{I}_n(t),$$

where $\mathcal{I}_0(t) = \mathbf{1}_p$ and $\mathcal{I}_n(t) = \int_{t_0}^{t} \mathcal{A}(\tau_1)\int_{t_0}^{\tau_1} \mathcal{A}(\tau_2)\cdots\int_{t_0}^{\tau_{n-1}} \mathcal{A}(\tau_n)d\tau_n \ldots d\tau_2 d\tau_1$.

If the commutativity assumption (14) holds, the equations (12) and (13) simplify to

$$\mathcal{X}(t) = e^{\int_{t_0}^{t} \mathcal{A}(s)ds}\mathcal{X}(t_0) + \int_{t_0}^{t} e^{\int_{u}^{t} \mathcal{A}(s)ds}\mathcal{C}L(du) \qquad\qquad = \int_{-\infty}^{t} e^{\int_{u}^{t} \mathcal{A}(s)ds}\mathcal{C}L(du) \text{ and}$$

$$Y(t) = \mathcal{B}(t)'e^{\int_{t_0}^{t} \mathcal{A}(s)ds}\mathcal{X}(t_0) + \int_{t_0}^{t} \mathcal{B}(t)'e^{\int_{u}^{t} \mathcal{A}(s)ds}\mathcal{C}L(du) \qquad = \int_{-\infty}^{t} \mathcal{B}(t)'e^{\int_{u}^{t} \mathcal{A}(s)ds}\mathcal{C}L(du)$$

for $t, t_0 \in \mathbb{R}$, where $t > t_0$.

Remark 4.4. *If $\mathcal{A}(s)$ and $\mathcal{A}(t)$ commute, i.e. $[\mathcal{A}(s), \mathcal{A}(t)] = 0$ for all $s, t \in \mathbb{R}$, then the commutativity assumption (14) holds. However, the matrices $\mathcal{A}(t)$, $t \in \mathbb{R}$ are in companion form and are not in general commutative (see also Proposition 4.18). For further insight into the commutativity of some matrices $\mathcal{A}(t)$ and $\int_{t_0}^{t} \mathcal{A}(s)ds$ as well as $\mathcal{A}(s)$ and $\mathcal{A}(t)$, we refer to [32, Exercise 4.8] and [41].*

The previous remark shows that, when considering time-varying CARMA(p,q) processes, it is in general not possible to describe the solution of the state space equations explicitly in form of a matrix exponential. Instead one has to use the Peano-Baker series.

In [35, Corollary 3.4] it is proved that, in the time-invariant case, the class of CARMA processes is equivalent to the class of continuous-time state space models. This motivates looking at time-varying state space processes. We consider the observation and state equations

$$Y(t) = B(t)'X(t) \text{ and}$$
$$dX(t) = A(t)X(t)dt + C(t)L(dt), \tag{15}$$

where $t \in \mathbb{R}$, $A(t) \in M_{p \times p}(\mathcal{R}[t])$ and $B(t), C(t) \in M_{p \times 1}(\mathcal{R}[t])$ are arbitrary continuous coefficient functions and L is a two-sided Lévy process satisfying (5).

Now, the representation of a time-varying CARMA processes as given in (12) can be adapted to (general) state space processes. Provided the integrals exist in L^2, it can be shown that a solution of (15) is given by

$$X(t) = \int_{-\infty}^{t} \Psi(t,u)C(u)L(du) \quad \text{and} \quad Y(t) = B(t)\int_{-\infty}^{t} \Psi(t,u)C(u)L(du), \tag{16}$$

where $\Psi(t,t_0)$ is the unique matrix solution of the IVP $\frac{d}{dt}\Psi(t,t_0) = A(t)\Psi(t,t_0)$, $\Psi(t_0,t_0) = \mathbf{1}_p$ for $t > t_0$. In particular, in the representation of (16), the integrals are well-defined, if $C(\cdot) \in L^2(\mathbb{R})$ and there exist $\gamma, \lambda > 0$, such that

$$\|\Psi(t,t_0)\| \le \gamma e^{-\lambda(t-t_0)} \qquad \text{for all } t \ge t_0 \text{ (uniform exponential stability)}.$$

For some initial time $t_0 \in \mathbb{R}$ the process satisfies the relation

$$X(t) = \Psi(t,t_0)\left(X(t_0) + \int_{t_0}^{t} \Psi(u,t_0)^{-1}C(u)L(du)\right). \tag{17}$$

Finally, we define

Definition 4.5. *A solution $\{Y(t)\}_{t \in \mathbb{R}}$ of the observation and state equations (15) in the form (16) is called a time-varying Lévy-driven state space process.*

The natural question arises, whether all time-varying state space processes are tvCARMA processes, as in the time-invariant case. A comprehensive investigation of this question seems beyond the scope of this work. Below we present a result indicating that this is probably not the case in general (definitely not when allowing the coefficient functions to have a discontinuity). Moreover, we give sufficient conditions for a positive answer.

Proposition 4.6. *The class of time-varying Lévy-driven CARMA models (11) and time-varying Lévy-driven state space models (15) with not necessarily continuous coefficient functions do not coincide in general.*

Proof. Consider a two dimensional time-varying state space model as defined in (15) with a structural break at $t = 1$. As coefficient functions we consider the step functions

$$B(t) = \begin{cases} B_1 & \text{if } t \le 1 \\ B_2 & \text{if } t > 1 \end{cases}, \quad A(t) = \begin{cases} A_1 & \text{if } t \le 1 \\ A_2 & \text{if } t > 1 \end{cases} \text{ and } C(t) = \begin{cases} C_1 & \text{if } t \le 1 \\ C_2 & \text{if } t > 1 \end{cases}, \tag{18}$$

14

which satisfy the uniform exponential stability assumption for the solution of (15).

We assume that the system is in the form of a CARMA process for $t \leq 1$ and assume (for contradiction) that there exists an equivalent CARMA model as defined in (11) for all $t \in \mathbb{R}$. Then, the CARMA model shows the same structural resemblance as the corresponding state space model. In the following we denote the coefficients of the CARMA model by $\mathcal{B}(t), \mathcal{A}(t)$ and $\mathcal{C}(t)$. Using the same notation as in (18) we obtain

$$
\begin{aligned}
\mathcal{B}_1 = B_1, && \mathcal{A}_1 = A_1, && \mathcal{C}_1 = C_1, \\
\mathcal{B}_2 = \begin{pmatrix} * \\ * \end{pmatrix}, && \mathcal{A}_2 = \begin{pmatrix} 0 & 1 \\ * & * \end{pmatrix} && \text{and} && \mathcal{C}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.
\end{aligned}
\tag{19}
$$

Since the structural break divides the model in two separate linear models, the CARMA representations $(\mathcal{B}_1, \mathcal{A}_1, \mathcal{C}_1)$ and $(\mathcal{B}_2, \mathcal{A}_2, \mathcal{C}_2)$ are unique. From the proof of [35, Theorem 3.3] we obtain that $\mathcal{B}_2 e^{\mathcal{A}_2(t-u)} \mathcal{C}_2 = B_2 e^{\mathcal{A}_2(t-u)} C_2$ for all $t > 1$ and $u \in (1, t]$.

On the one hand, we have

$$
\begin{aligned}
Y(t) = {}& \mathbb{1}_{\{t \leq 1\}} \left(\int_{-\infty}^t B(t)' \Psi_A(t, u) C(u) L(du) \right) \\
& + \mathbb{1}_{\{t > 1\}} \left(B(t)' \Psi_A(t, 1) X(1) + B(t)' \Psi_A(t, 1) \int_1^t \Psi_A(u, 1)^{-1} C(u) L(du) \right) \\
= {}& \mathbb{1}_{\{t \leq 1\}} \left(\int_{-\infty}^t B_1' e^{A_1(t-u)} C_1 L(du) \right) \\
& + \mathbb{1}_{\{t > 1\}} \left(B_2' e^{A_2(t-1)} X(1) + \int_1^t B_2' e^{A_2(t-u)} C_2 L(du) \right),
\end{aligned}
\tag{20}
$$

where $\Psi_A(s, s_0)$ denotes the solution of the aforementioned IVP with respect to A. On the other hand, $Y(t)$ can be written as

$$
\begin{aligned}
Y(t) = {}& \mathbb{1}_{\{t \leq 1\}} \left(\int_{-\infty}^t \mathcal{B}_1' e^{\mathcal{A}_1(t-u)} \mathcal{C}_1 L(du) \right) \\
& + \mathbb{1}_{\{t > 1\}} \left(\mathcal{B}_2' e^{\mathcal{A}_2(t-1)} \mathcal{X}(1) + \int_1^t \mathcal{B}_2' e^{\mathcal{A}_2(t-u)} \mathcal{C}_2 L(du) \right).
\end{aligned}
\tag{21}
$$

From (19) it follows that

$$
X(1) = \int_{-\infty}^1 e^{A_1(1-u)} C_1 L(du) = \int_{-\infty}^1 e^{\mathcal{A}_1(1-u)} \mathcal{C}_1 L(du) = \mathcal{X}(1).
$$

Thus, combining (20) and (21), using (19) and the independent increments of the Lévy process, the equality $B_2' e^{A_2(t-1)} X(1) = \mathcal{B}_2' e^{\mathcal{A}_2(t-1)} X(1)$ has to hold almost surely for all $t > 1$. Therefore, for almost all x in the support of $X(1)$ we obtain

$$
B_2' e^{A_2(t-1)} x = \mathcal{B}_2' e^{\mathcal{A}_2(t-1)} x.
\tag{22}
$$

In the sequel we give a particular Lévy process and coefficient functions that lead to a contradiction in (22).

Assume that the Lévy process is a Brownian motion. Thus, it has the triplet $(0, \Sigma, 0)$ for some $\Sigma > 0$. From [34] it follows that $X(t)$ is a Lévy process with triplet $(0, \Sigma_X^t, 0)$, where $\Sigma_X^1 = \int_{-\infty}^1 e^{A_1(1-u)} C_1 \Sigma C_1' e^{A_1'(1-u)} du = \Sigma \int_0^\infty e^{A_1 u} C_1 C_1' e^{A_1' u} du \in M_{2\times2}(\mathbb{R})$. The regularity of

Σ_X^1 can be shown by investigating $Im(\Sigma_X^1)$, where $Im(D) = \{Dx : x \in \mathbb{R}^d\}$ denotes the image of a matrix $D \in M_{d \times d}(\mathbb{R})$. Using [6, Lemma 12.6.2] (see also [35, p. 54]) we obtain

$$Im\left(\int_0^\infty e^{A_1 u} C_1 C_1' e^{A_1' u} du\right) = Im\left([C_1 \quad A_1 C_1 \quad \cdots \quad A_1^{p-1} C_1]\right).$$

Therefore, in our setting, it is sufficient to find A_1, C_1 such that $[C_1 \ A_1 C_1]$ is regular, which also implies that Σ_X^1 is positive definite. Then, $X(1)$ has characteristic function $E[e^{i\langle z, X(1)\rangle}] = e^{-\frac{1}{2}\langle z, \Sigma_X^1 z\rangle}$, which corresponds to a two dimensional $N(0, \Sigma_X^1)$ distributed random variable, having positive density for all values $x \in \mathbb{R}^2$. To contradict (22), it is enough to show that for some $t > 1$

$$B_2' e^{A_2(t-1)} x \neq \mathcal{B}_2' e^{\mathcal{A}_2(t-1)} x \qquad \text{for all } x \in I, \text{ where } I \subset \mathbb{R}^2 \text{ with } \lambda(I) > 0.$$

We define

$$B_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathcal{B}_2 = \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -2 & 0 \\ 0 & -3 \end{pmatrix}, \quad \mathcal{A}_2 = \begin{pmatrix} 0 & 1 \\ -6 & -5 \end{pmatrix} \text{ and}$$

$$C_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathcal{C}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

From this we obtain that the CARMA model has the same transfer function as the state space model, since $B_2'(z\mathbf{1}_2 - A_2)^{-1} C_2 = \frac{2z+5}{z^2+5z+6} = \mathcal{B}_2'(z - \mathbf{1}_2 \mathcal{A}_2)^{-1} \mathcal{C}_2$. Moreover, $[C_1 \ A_1 C_1] = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ is regular. Given a vector $x = (x_1, x_2)'$ it is left to investigate $\mathcal{B}_2' e^{\mathcal{A}_2(t-1)} x - B_2' e^{A_2(t-1)} x$. For a matrix $D \in \mathbb{C}^{2 \times 2}$ with eigenvalues $\sigma(A) = \{\mu, \lambda\}$, [6, Proposition 11.3.2] gives that

$$e^D = \begin{cases} e^\lambda((1-\lambda)\mathbf{1}_2 + D) & \text{if } \mu = \lambda \\ \frac{\mu e^\lambda - \lambda e^\mu}{\mu - \lambda} \mathbf{1}_2 + \frac{e^\mu - e^\lambda}{\mu - \lambda} D & \text{if } \mu \neq \lambda. \end{cases}$$

Since $\sigma(\mathcal{A}_2(t-1)) = \{-2(t-1), -3(t-1)\}$, we obtain

$$\mathcal{B}_2' e^{\mathcal{A}_2(t-1)} x - B_2' e^{A_2(t-1)} x = \begin{pmatrix} 5 & 2 \end{pmatrix} \left(\frac{-3(t-1)e^{-2(t-1)} - (-2)(t-1)e^{-3(t-1)}}{-3(t-1) - (-2)(t-1)} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right.$$

$$+ \frac{e^{-3(t-1)} - e^{-2(t-1)}}{-3(t-1) - (-2)(t-1)} \begin{pmatrix} 0 & 1 \\ -6 & -5 \end{pmatrix} (t-1) \left. \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$- \begin{pmatrix} 1 & 1 \end{pmatrix} \left(\begin{pmatrix} e^{-2(t-1)} & 0 \\ 0 & e^{-3(t-1)} \end{pmatrix} \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= 2x_1 \left(e^{-2(t-1)} + e^{-3(t-1)} \right) + x_2 \left(e^{-2(t-1)} \right) > 0$$

for all $x \in I = \{x \in \mathbb{R}^2 : x_1 > 0, x_2 > 0\}$ and $t > 1$. $\qquad \square$

Under more rigorous conditions on the coefficient functions, the concept of controllability from linear system theory allows for a characterization for special *canonical* forms, which occur in the state space representation of CARMA processes ($\mathcal{A}$ is in companion matrix form). The following results summarize the key aspects of this characterization, which is mainly based on [36], but also [5, 29, 28].

Definition 4.7 ([32, Chapter 9 and 10]). *Let $Y(t)$ be a state space model as defined in (15), where $A(t)$ is $(p-1)$-times continuously differentiable and $C(t)$ p-times. We define the controllability matrix $W_p(t)$, as*

$$W_p(t) = [K_0(t)\ K_1(t)\ \cdots\ K_{p-1}(t)],\ \text{where}$$
$$K_0(t) = C(t), \quad K_{i+1}(t) = -A(t)K_i(t) + \tfrac{d}{dt}K_i(t), \quad i = 1,\ldots,p-2.$$

Then, the state process $X(t)$ is called

(a) *controllable on $[t_0, t_1]$, $t_0 < t_1$, if there exists $t \in [t_0, t_1]$ with $Rank(W_p(t)) = p$ and*

(b) *instantaneously controllable, if $Rank(W_p(t)) = p$ for all $t \in \mathbb{R}$.*

Proposition 4.8 ([36, Theorem 1]). *Consider a state space process satisfying (15) such that A is $(p-1)$-times continuously differentiable and C p-times. Then, it is equivalent to a CARMA process satisfying (11) if and only if it is instantaneously controllable. Equivalence means that there exists a regular matrix $T(t) \in M_{n \times n}(\mathcal{R}[t])$, which is continuously differentiable and satisfies*

$$\mathcal{X}(t) = T(t)X(t)$$

almost surely. The relationship between both systems is given by $T(t) = \mathcal{W}_p(t)W_p(t)^{-1}$, $\mathcal{A}(t) = \left(T(t)A(t) + \tfrac{d}{dt}T(t)\right)T(t)^{-1}$ and $\mathcal{C} = T(t)C(t)$, where $\mathcal{W}_p(t)$ and $W_p(t)$ are the controllability matrices of the state space model and the CARMA process.

Corollary 4.9. *The class of time-varying Lévy-driven state space models as defined in (15) with $(p-1)$-times continuously differentiable coefficient functions A, p-times continuously differentiable coefficient functions C and controllability matrices $W_p(t)$ that have rank p everywhere, is equivalent to the class of time-varying CARMA(p,q) processes as defined in (11) with $(p-1)$-times continuously differentiable coefficient functions A and controllability matrices $W_p(t)$ that have rank p everywhere.*

Proof. Any time-varying CARMA(p,q) process is obviously also a time-varying state space process. On the contrary, let $Y(t)$ be a time-varying state space process defined by (15), which is instantaneously controllable with controllability matrix $W_p(t)$. Then, due to Proposition 4.8, the state system $dX(t) = A(t)X(t)dt + C(t)L(dt)$ is equivalent to the CARMA system $d\mathcal{X}(t) = \mathcal{A}(t)\mathcal{X}(t)dt + \mathcal{C}L(dt)$ with

$$\mathcal{X}(t) = T(t)X(t), \quad \mathcal{C} = T(t)C(t) \quad \text{and} \quad \mathcal{A}(t) = \left(T(t)A(t) + \tfrac{d}{dt}T(t)\right)T(t)^{-1},$$

where $T(t) = \mathcal{W}_p(t)W_p(t)^{-1}$ is regular. Thus

$$Y(t) = B(t)'X(t) = B(t)'T(t)^{-1}\mathcal{X}(t) = \mathcal{B}(t)'\mathcal{X}(t) \quad \text{and}$$
$$d\mathcal{X}(t) = \mathcal{A}(t)\mathcal{X}(t)dt + \mathcal{C}L(dt),$$

which is a representation for $Y(t)$ as a time-varying CARMA(p,q) process in (11). $\qquad\square$

4.3 Locally stationary linear state space models - Peano-Baker series

We investigate sufficient conditions for sequences of time-varying state space processes, which obviously also includes sequences of time-varying CARMA processes, to be locally stationary. Let $\{Y_N(t), t \in \mathbb{R}\}_{N \in \mathbb{N}}$ be a sequence of time-varying linear state space processes defined by

$$Y_N(t) = B(t)' X_N(Nt) \quad \text{and}$$

$$X_N(Nt) = \Psi_N(Nt, 0) \int_{-\infty}^{Nt} \Psi_N(u, 0)^{-1} C(\tfrac{u}{N}) L(du),$$

where $\Psi_N(s, s_0)$ is the solution of the matrix differential equation

$$\Psi_N(s_0, s_0) = \mathbf{1}_p$$
$$\tfrac{d}{ds} \Psi_N(s, s_0) = A(\tfrac{s}{N}) \Psi_N(s, s_0), \quad \text{for all } s, s_0 \in \mathbb{R} \text{ with } s > s_0,$$

which can be expressed as (see [2] Section 2)

$$\Psi_N(s, s_0) = \mathbf{1}_p + \int_{s_0}^{s} A(\tfrac{\tau_1}{N}) d\tau_1 + \int_{s_0}^{s} A(\tfrac{\tau_1}{N}) \int_{s_0}^{\tau_1} A(\tfrac{\tau_2}{N}) d\tau_2 d\tau_1 + \dots .$$

The substitution $s \mapsto s + Nt$ in (9), is necessary to achieve a dependence of the kernel function $g_N^0(t, \cdot)$ on $Nt - u$. Therefore, we define $\widetilde{\Psi}_{N,t}(0, -(Nt - u))$ for a fixed point $t \in \mathbb{R}$ as the solution of the matrix differential equation

$$\widetilde{\Psi}_{N,t}(s_0, s_0) = \mathbf{1}_p$$
$$\tfrac{d}{ds} \widetilde{\Psi}_{N,t}(s, s_0) = A(\tfrac{s}{N} + t) \widetilde{\Psi}_{N,t}(s, s_0), \quad \text{for all } s, s_0 \in \mathbb{R} \text{ with } s > s_0,$$

which can again be expressed as

$$\widetilde{\Psi}_{N,t}(s, s_0) = \mathbf{1}_p + \int_{s_0}^{s} A(\tfrac{\tau_1}{N} + t) d\tau_1 + \int_{s_0}^{s} A(\tfrac{\tau_1}{N} + t) \int_{s_0}^{\tau_1} A(\tfrac{\tau_2}{N} + t) d\tau_2 d\tau_1 + \dots$$

From [7, Theorem 4.2.] we obtain $\Psi_N(Nt, 0) \Psi_N(u, 0)^{-1} = \Psi_N(Nt, 0) \Psi_N(0, u) = \Psi_N(Nt, u)$. Since,

$$\Psi_N(Nt, u) = \mathbf{1}_p + \int_{u}^{Nt} A(\tfrac{\tau_1}{N}) d\tau_1 + \int_{u}^{Nt} A(\tfrac{\tau_1}{N}) \int_{u}^{\tau_1} A(\tfrac{\tau_2}{N}) d\tau_2 d\tau_1 + \dots$$
$$= \mathbf{1}_p + \int_{u-Nt}^{0} A(\tfrac{\tau_1}{N} + t) d\tau_1 + \int_{u-Nt}^{0} A(\tfrac{\tau_1}{N} + t) \int_{u-Nt}^{\tau_1} A(\tfrac{\tau_2}{N} + t) d\tau_2 d\tau_1 + \dots$$
$$= \widetilde{\Psi}_{N,t}(0, -(Nt - u)),$$

we neglect the superscript tilde and define a sequence of time-varying linear state space processes as follows.

Definition 4.10. *A sequence of time-varying linear state space processes $\{Y_N(t), t \in \mathbb{R}\}_{N \in \mathbb{N}}$ is defined as*

$$Y_N(t) = \int_{\mathbb{R}} \not{1}_{\{Nt - u \geq 0\}} B(t)' \Psi_{N,t}^0(0, -(Nt - u)) C(\tfrac{u}{N}) L(du) \tag{23}$$

18

with (limiting) kernel function (in view of Definition 3.1)

$$g_N^0(Nt, Nt - u) = \mathbb{1}_{\{Nt-u \geq 0\}} B(t)' \Psi_{N,t}^0(0, -(Nt - u)) C\left(\frac{-(Nt-u)}{N} + t\right) \quad and$$
$$g(t, Nt - u) = \mathbb{1}_{\{Nt-u \geq 0\}} B(t)' \Psi_t(0, -(Nt - u)) C(t),$$

where $\Psi_{N,t}^0(0, -(Nt - u))$ and $\Psi_t(0, -(Nt - u))$ are the solutions of the matrix differential equations

$$
\begin{aligned}
\Psi_{N,t}^0(s_0, s_0) &= \mathbf{1}_p, & \frac{d}{ds}\Psi_{N,t}^0(s, s_0) &= A\left(\frac{s}{N} + t\right)\Psi_{N,t}^0(s, s_0), \\
\Psi_t(s_0, s_0) &= \mathbf{1}_p \quad and & \frac{d}{ds}\Psi_t(s, s_0) &= A(t)\Psi_t(s, s_0)
\end{aligned}
\tag{24}
$$

for $s > s_0$.

Using the Peano-Baker series, if necessary, the solutions of the above matrix differential equations are given by $\Psi_t(s, s_0) = e^{A(t)(s-s_0)}$ and

$$\Psi_{N,t}^0(s, s_0) = \mathbf{1}_p + \int_{s_0}^s A\left(\frac{\tau_1}{N} + t\right)d\tau_1 + \int_{s_0}^s A\left(\frac{\tau_1}{N} + t\right)\int_{s_0}^{\tau_1} A\left(\frac{\tau_2}{N} + t\right)d\tau_2 d\tau_1 + \dots.$$

Proposition 4.11. *Let $\{Y_N(t), t \in \mathbb{R}\}_{N \in \mathbb{N}}$ be a sequence of time-varying state space processes as in Definition 4.10. If*

(C1) the coefficient functions $A(\cdot)$, $B(\cdot)$ and $C(\cdot)$ are continuous,

(C2) $\|B(s)\| < \infty$ for all $s \in \mathbb{R}$, $\sup_{s \in \mathbb{R}} \|C(s)\| < \infty$ and

(C3) $\left\|\Psi_{N,t}^0(0, u)\right\| \leq F_t(u)$ for some real function $F_t \in L^2((-\infty, 0])$ for all $N \in \mathbb{N}$ and $t \in \mathbb{R}$,

then $Y_N(t)$ is locally stationary.

Proof. Consider $Y_N(t)$, g_N^0, g, $\Psi_{N,t}^0$ and Ψ_t as defined above. For fixed $u, t \in \mathbb{R}$ it holds

$$
\begin{aligned}
\left| g_N^0(Nt, -u) - g(t, -u) \right| &= \mathbb{1}_{\{u \leq 0\}} \Big| B(t)' \Big(\Psi_{N,t}^0(0, u) - \Psi_t(0, u) \Big) C\left(\frac{u}{N} + t\right) \\
&\qquad + B(t)' \Psi_t(0, u) \Big(C\left(\frac{u}{N} + t\right) - C(t) \Big) \Big| \\
&\leq \mathbb{1}_{\{u \leq 0\}} \Big(\|B(t)\| \left\| \Psi_{N,t}^0(0, u) - \Psi_t(0, u) \right\| \Big(\sup_{s \in \mathbb{R}} \|C(s)\| \Big) \\
&\qquad + \|B(t)\| \|\Psi_t(0, u)\| \left\| C\left(\frac{u}{N} + t\right) - C(t) \right\| \Big) =: P_1 + P_2.
\end{aligned}
$$

Since $C(\cdot)$ is continuous, we immediately obtain $P_2 \to 0$ as $N \to \infty$. In view of P_1 it is sufficient to show that for all $\varepsilon > 0$ and sufficiently large N

$$\left\| \Psi_{N,t}^0(0, u) - \Psi_t(0, u) \right\| \leq \varepsilon. \tag{25}$$

Due to the equivalence of all norms on $M_{p \times p}(\mathbb{R})$, it is sufficient to show (25) for the norm of each column. By $\Psi_{N,t}^{0(j)}(0, u)$ and $\Psi_t^{(j)}(0, u)$ we denote the j-th column, $j = 1, \dots, p$ of $\Psi_{N,t}^0(0, u)$ and $\Psi_t(0, u)$. Then, for functions $f_{N,t}(s, x) = A\left(\frac{s}{N} + t\right)x$ and $\tilde{f}_t(s, x) = A(t)x$ we obtain

$$\Psi_{N,t}^{0(j)}(u, u) = e_j, \qquad\qquad \frac{d}{ds}\Psi_{N,t}^{0(j)}(s, u) = f_{N,t}\left(s, \Psi_{N,t}^{0(j)}(s, u)\right),$$

$$\Psi_t^{(j)}(u, u) = e_j \quad \text{and} \qquad \frac{d}{ds}\Psi_t^{(j)}(s, u) = \tilde{f}_t\left(s, \Psi_t^{(j)}(s, u)\right),$$

where e_j denotes the j-th unit vector. Note that $f_{N,t}$ and $\tilde{f}_t$ are Lipschitz continuous in the second argument with Lipschitz constant $L = \sup_{s \in [u,0]} \|A(\frac{s}{N} + t)\| + A(t) < \infty$. Moreover,

$$\left\| f_t\left(s, \Psi_t^{(j)}(s, u)\right) - f_{N,t}\left(s, \Psi_t^{(j)}(s, u)\right) \right\| \leq \delta \left\| \Psi_t^{(j)}(s, u) \right\| \leq \delta c, \quad s \in [u, 0],$$

since $\|A(\frac{s}{N} + t) - A(t)\| < \delta$ for any $\delta > 0$ for sufficiently large N and $\Psi_t^{(j)}(\cdot, u)$ is continuous and thus bounded on $[u, 0]$. An application of [40, §12.V.] gives (25). Finally, by using the dominated convergence theorem with majorant

$$\left| g_N^0(Nt, -u) \right| \leq \mathcal{W}_{\{u \leq 0\}} \|B(t)\| \left(\sup_{s \in \mathbb{R}} \|C(s)\| \right) \left\| \Psi_{N,t}^0(0, u) \right\|$$

$$\leq \mathcal{W}_{\{u \leq 0\}} c_t F_t(u) \in L^2(\mathbb{R})$$

for some constant $c_t > 0$, where the last inequality follows from $(C1)$ and $(C2)$, we can deduce that $\|g_N^0(Nt, \cdot) - g(t, \cdot)\|_{L^2} \to 0$ as $N \to \infty$. $\qquad \square$

In fact, assumption $(C3)$ in Proposition 4.11 is an immediate consequence if the state space system is uniformly exponentially stable.

Definition 4.12 ([32, Chapter 6, Definition 6.5 and Theorem 6.7)]). *A sequence of linear state space models as in Definition 4.10 is called uniformly exponentially stable, if there exist $\gamma > 0$ and $\lambda > 0$, such that*

$$\left\| \Psi_{N,t}^0(s, s_0) \right\| \leq \gamma e^{-\lambda(s - s_0)}$$

for all s, s_0, where $s > s_0$, $N \in \mathbb{N}$ and $t \in \mathbb{R}$.

Corollary 4.13. *If a linear state space model as in Definition 4.10 is uniformly exponentially stable, then*

$$\left\| \Psi_{N,t}^0(0, u) \right\| \leq \gamma e^{\lambda u} =: F_t(u) \in L^2((-\infty, 0]), \text{ for all } N, t, u \leq 0,$$

which is $(C3)$ in Proposition 4.11.

Proposition 4.14. *Each of the following two conditions is sufficient for a state space model $\{Y_N(t), t \in \mathbb{R}\}_{N \in \mathbb{N}}$ as in Definition 4.10 to be uniformly exponentially stable.*

(a) *Let $\lambda_{max}(t)$, $t \in \mathbb{R}$ denote the largest eigenvalue of $A(t) + A(t)'$. If there exist positive constants γ and λ, such that*

$$\int_{s_0}^{s} \lambda_{max}(\tfrac{\nu}{N} + t) d\nu \leq -\lambda(s - s_0) + \gamma$$

for all s, s_0, t and N with $s \geq s_0$, then, due to [32, Corollary 8.4], $Y_N(t)$ is uniformly exponentially stable.

(b) *Suppose $A(t)$ is continuously differentiable and there exist positive constants α, μ, and β such that $\|A(t)\| \leq \alpha$, $\left\| \frac{d}{dt} A(t) \right\| \leq \beta$ and the eigenvalues $\lambda_j(t)$ of $A(t)$ for $j = 1, \ldots, p$ satisfy $\mathfrak{Re}(\lambda_j(t)) \leq -\mu$ for all t. Then, due to [32, Theorem 8.7], $Y_N(t)$ is uniformly exponentially stable.*

Remark 4.15. *Part (b) of Proposition 4.14 corresponds to condition (C2) in Proposition 4.1 for sequences of tvCAR(1) processes.*

Remark 4.16. *There is a high structural resemblance of the above results to known results from the theory on locally stationary processes in discrete time. Indeed, conditions for time-varying AR(p) processes to be locally stationary as discussed in [11, Theorem 2.3] and [22] are closely related to the conditions in Proposition 4.14.*

If the commutativity assumption (14) holds, the transition matrix is given by

$$\Psi_{N,t}^0(0,u) = e^{\int_u^0 A(\frac{s}{N}+t)ds}.$$

Then, $Y_N(t)$ simplifies to

$$Y_N(t) = \int_{-\infty}^{Nt} B(t)' e^{\int_{-(Nt-u)}^0 A(\frac{s}{N}+t)ds} C(\frac{u}{N}) L(du).$$

Proposition 4.17. *Let $\{Y_N(t), t \in \mathbb{R}\}_{N\in\mathbb{N}}$ be a sequence of time-varying state space processes as in Definition 4.10. If (C1) and (C2) from Proposition 4.11 hold, $\{A(t)\}_{t\in\mathbb{R}}$ is mutually commutative, the eigenvalues $\lambda_j(t)$ of $A(t)$ for $j = 1, \dots, p$ satisfy $\mathfrak{Re}(\lambda_j(t)) \leq -\mu$ for all $t \in \mathbb{R}$ and some $\mu > 0$ and either*

(D1) $A(t)$ is diagonalizable for all $t \in \mathbb{R}$ or

(D2) there exists $\tau > 0$ such that $\sup_{\tau \leq x < \infty} \left\| \frac{1}{x} \int_\tau^x A\left(\frac{s}{N}+t\right) \right\| < C$ for all N and a constant $C > 0$,

then $Y_N(t)$ is locally stationary.

Proof. It is sufficient to check $(C3)$ from Proposition 4.11. We start by assuming that $(D1)$ holds. Then, due to [19, Theorem 1.3.12], $\{A(t)\}_{t\in\mathbb{R}}$ is simultaneously diagonalizable. Thus, there exists a non-singular matrix S such that $S^{-1}A(\frac{s}{N}+t)S = diag(\lambda_1(\frac{s}{N}+t), \dots, \lambda_p(\frac{s}{N}+t)) =: D(\frac{s}{N}+t)$. Considering the spectral norm, we obtain for all $u \leq 0$

$$\left\| \Psi_{N,t}^0(0,u) \right\| = \left\| e^{\int_u^0 A(\frac{s}{N}+t)ds} \right\| = \left\| e^{\int_u^0 SD(\frac{s}{N}+t)ds S^{-1}} \right\| \leq \|S\| \left\| S^{-1} \right\| \left\| e^{\int_u^0 D(\frac{s}{N}+t)ds} \right\|$$

$$\leq C \max \left\{ \sqrt{\mu} : \mu \in \sigma \left(e^{\int_u^0 D(\frac{s}{N}+t)^*+D(\frac{s}{N}+t)ds} \right) \right\} = C \max_{j=1,\dots,p} \sqrt{e^{2\int_u^0 \mathfrak{Re}(\lambda_j(\frac{s}{N}+t))ds}}$$

$$= C \max_{j=1,\dots,p} e^{\int_u^0 \mathfrak{Re}(\lambda_j(\frac{s}{N}+t))ds} \leq Ce^{\int_u^0 -\mu ds} = Ce^{\mu u}$$

for some constant $C > 0$.
In the case where $(D2)$ holds, we have

$$\left\| \Psi_{N,t}^0(0,u) \right\| = \left\| e^{\int_u^0 A(\frac{s}{N}+t)ds} \right\| = \left\| e^{\int_0^{-u} A(\frac{-s}{N}+t)ds} \right\|$$

$$\leq \mathbb{1}_{\{-u\in[0,\tau]\}} e^{|\tau| \sup_{s\in[0,\tau]} \|A(\frac{-s}{N}+t)\|} + \mathbb{1}_{\{-u>\tau\}} e^{|\tau| \sup_{s\in[0,\tau]} \|A(\frac{-s}{N}+t)\|} \left\| e^{\int_\tau^{-u} A(\frac{-s}{N}+t)ds} \right\|,$$

where we used that the integrals $\int_0^\tau A\left(\frac{-s}{N}+t\right)ds$ and $\int_\tau^{-u} A\left(\frac{-s}{N}+t\right)ds$ commute. Therefore, it is sufficient to bound $\left\| e^{\int_\tau^{-u} A(\frac{-s}{N}+t)ds} \right\|$. In the following we use [24, Theorem 7.7.1]. Since

the family $\{A(t)\}_{t\in\mathbb{R}}$ is mutually commutative, the family can be reduced simultaneously to an upper triangular form by a single unitary transformation, i.e. there exists a unitary matrix $U \in M_{p\times p}(\mathbb{C})$ such that $U^*A(t)U = T(t)$ is an upper triangular matrix for all $t \in \mathbb{R}$ (see [19, Theorem 2.3.3]). For each $x > \tau$, the diagonal entries (and hence also the eigenvalues) of $\frac{1}{x}\int_\tau^x T\left(\frac{s}{N}+t\right)ds$ are $\frac{1}{x}\int_\tau^x \lambda_1\left(\frac{s}{N}+t\right)ds,\ldots,\frac{1}{x}\int_\tau^x \lambda_p\left(\frac{s}{N}+t\right)ds$. These are also the eigenvalues of $\frac{1}{x}\int_\tau^x A\left(\frac{s}{N}+t\right)ds$ since

$$\frac{1}{x}\int_\tau^x T\left(\tfrac{s}{N}+t\right)ds = \tfrac{1}{x}U^*\int_\tau^x A\left(\tfrac{s}{N}+t\right)ds\,U.$$

For the real part of the eigenvalues we obtain

$$\mathfrak{Re}\left(\frac{1}{x}\int_\tau^x \lambda_i\left(\tfrac{s}{N}+t\right)ds\right) = \frac{1}{x}\int_\tau^x \mathfrak{Re}\left(\lambda_i\left(\tfrac{s}{N}+t\right)\right)ds \le -\mu\left(\frac{x-\tau}{x}\right) \le -\mu$$

for all $i = 1,\ldots,p$, $N \in \mathbb{N}$ and $x \in \mathbb{R}$. Hence

$$\overline{\bigcup_{\tau\le t<\infty} \tilde{\sigma}\left(\frac{1}{x}\int_\tau^x A\left(\tfrac{s}{N}+t\right)ds\right)} \subset \overline{\bigcup_{\tau\le t<\infty} \sigma\left(\frac{1}{x}\int_\tau^x A\left(\tfrac{s}{N}+t\right)ds\right)} \subset \{z\in\mathbb{C}: \mathfrak{Re}(z) \le -\mu\}$$

for all $N \in \mathbb{N}$, where $\tilde{\sigma}(B)$ denotes the collection of all distinct eigenvalues of the matrix B and $\sigma(B)$ the spectrum B. Finally, an application of [24, Theorem 7.7.1] gives

$$\left\|e^{\int_\tau^{-u} A\left(\frac{-s}{N}+t\right)ds}\right\| \le Ce^{\mu u}$$

for some constant $C > 0$. $\qquad\square$

Sequences of time-varying CARMA processes, i.e. where the family $\{\mathcal{A}(t)\}_{t\in\mathbb{R}}$ forms a family of companion matrix, cannot be covered by Proposition 4.17, since companion matrices are in general not commutative. The following Proposition brings further insight when a family of companion matrices is mutually commutative.

Proposition 4.18. *Let $\{\mathcal{A}(t)\}_{t\in\mathbb{R}}$ be a family of companion matrices and $\tau \in \mathbb{R}$ fixed. For any $t \in \mathbb{R}$ the matrix $\mathcal{A}(t)$ commutes with $\mathcal{A}(\tau)$ if and only if it is a polynomial of $\mathcal{A}(\tau)$ over $\mathbb{C}$.*

Proof. It is clear that any polynomial of $\mathcal{A}(\tau)$ commutes with $\mathcal{A}(\tau)$. For the other direction we refer to [19, Exercise 3.3P17]. $\qquad\square$

If, for a sequence of time-varying CARMA processes, the family $\{\mathcal{A}(t)\}_{t\in\mathbb{R}}$ is not mutually commutative, Proposition 4.11 provides sufficient conditions for local stationarity, where condition $(C3)$ can be derived from Proposition 4.14 and Corollary 4.13.

5 Time-varying spectrum

For a stationary processes $\{Y(t), t \in \mathbb{R}\}$ the autocovariance function $\gamma_Y(h) := Cov(Y(t+h), Y(t))$ is related to the spectral density $f_Y(\lambda)$ by

$$\gamma_Y(h) = \int_{-\infty}^\infty e^{ih\lambda} f_Y(\lambda)d\lambda \qquad \text{and} \qquad f_Y(\lambda) = \frac{1}{2\pi}\int_{-\infty}^\infty e^{-ih\lambda}\gamma_Y(h)dh.$$

To describe the time-varying spectrum of a discrete-time locally stationary time series, [11] used the Wigner-Ville spectrum (see also [9, 18, 26]). A comparable approach was presented in [27, Section 11.2], where the author used the evolutionary spectrum. However, in contrast to this approach, the Wigner-Ville spectrum has the important consequence of a unique spectral representation as discussed in [9, p. 74] and [11, p. 143]. In view of this property, we follow the approach of [11] and define the Wigner-Ville spectrum and time-varying spectral density for a continuous-time locally stationary process as follows.

Definition 5.1. *Let* $\{Y_N(t), t \in \mathbb{R}\}_{N \in \mathbb{N}}$ *be a sequence of locally stationary processes. For* $N \in \mathbb{N}$ *we define the Wigner-Ville spectrum as*

$$f_N(t, \lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda s} Cov\left(Y_N(t + \tfrac{s}{2N}), Y_N(t - \tfrac{s}{2N})\right) ds,$$

and the (time-varying) spectral density of the process $Y_N(t)$ *as*

$$f(t, \lambda) = \frac{\Sigma_L}{2\pi} |A(t, \lambda)|^2,$$

where $A(t, \lambda)$ *denotes the limiting transfer function from Definition 3.2.*

The following theorem is a continuous-time analogue to [11, Theorem 2.2].

Theorem 5.2. *Let* $\{Y_N(t), t \in \mathbb{R}\}_{N \in \mathbb{N}}$ *be a sequence of locally stationary processes in the form (7). If*

(a) $\left\|A_N^0(N(t \pm \tfrac{s}{2N}), \cdot) - A(t, \cdot)\right\|_{L^2} \xrightarrow[N \to \infty]{} 0$ *for all* $s, t \in \mathbb{R}$,

(b) A_N^0 *as well as* A *are uniformly bounded in* L^2, *i.e* $\left\|A_N^0(Nt, \cdot)\right\|_{L^2} \leq K$, $\left\|A(t, \cdot)\right\|_{L^2} \leq K$ *for all* $t \in \mathbb{R}$, $N \in \mathbb{N}$ *and a constant* $K > 0$ *and*

(c) $A_N^0(Nt, \cdot)$ *as well as* $A(t, \cdot)$ *are differentiable for all* $t \in \mathbb{R}$, $N \in \mathbb{N}$ *and the derivatives* $\frac{d}{d\mu} A_N^0(Nt, \mu)$, $\frac{d}{d\mu} A(t, \mu)$ *are uniformly bounded in* L^2, *i.e.* $\left\|\frac{d}{d\mu} A_N^0(Nt, \cdot)\right\|_{L^2} \leq K$ *and* $\left\|\frac{d}{d\mu} A(t, \cdot)\right\|_{L^2} \leq K$, *for all* t, N *and a constant* $K > 0$,

then the Wigner-Ville spectrum tends pointwise for each $t \in \mathbb{R}$ *in mean square to the time-varying spectral density, i.e.*

$$\int_{\mathbb{R}} |f_N(t, \lambda) - f(t, \lambda)|^2 \, d\lambda \xrightarrow{N \to \infty} 0.$$

Remark 5.3. *Since* A_N^0 *and* A *are defined as Fourier transforms of* g_N^0 *and* g *in* L^2, *they exist as elements in* L^2, *i.e. as representatives of equivalence classes. As usual, this does not allow for taking derivatives in the usual sense, but would lead to the concept of weak derivatives.*
However, for a function $f \in L^1$ *such that* $uf(u) \in L^1$, *the derivative of the Fourier transform* $\widehat{f}(\mu)$ *can be expressed as* $\frac{d}{d\mu} \widehat{f}(\mu) = (-\widehat{iuf(u)})(\mu)$, *due to [20, Theorem 1.6, Chapter VI]. An application of this theorem to the Fourier transform pairs* A_N^0 *and* g_N^0 *as well as* A *and* g, *ensures the existence of the pointwise derivatives in (c). The conditions on the the kernel functions can be readily obtained for instance if the considered sequence of locally stationary state space models is uniformly exponentially stable, since then the kernel functions are of exponential decay (see Corollary 5.5).*

Lemma 5.4. *A continuous function $h : \mathbb{R} \to \mathbb{R}$ in $L^2(\mathbb{R})$ lies also in $L^p(\mathbb{R})$ for all $p > 2$.*

Proof. Define the set $J = \{x, |f(x)| \geq 1\}$. Since $\lim_{x \to \pm\infty} f(x) = 0$, J is a compact set. Thus, there exists a constant C such that $|f(x)| \leq C$ for all $x \in J$. Finally,

$$\int_{\mathbb{R}} |f(x)|^p \, dx \leq \int_J C^p dx + \int_{\mathbb{R} \setminus J} |f(x)|^2 \, dx < \infty.$$

$\square$

Proof of Theorem 5.2. In the following, K denotes the constant from the conditions (b) and (c). First, we note that the covariance of $Y_N(t)$ is given by

$$Cov\left(Y_N(t_1), Y_N(t_2)\right) = \frac{\Sigma_L}{2\pi} \int_{\mathbb{R}} e^{i\mu(t_1 - t_2)N} A_N^0(Nt_1, \mu) \overline{A_N^0(Nt_2, \mu)} d\mu.$$

Then, we obtain for the Wigner-Ville spectrum and the (time-varying) spectral density

$$f_N(t, \lambda) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\lambda s} \frac{\Sigma_L}{2\pi} \int_{\mathbb{R}} e^{i\mu\left((t + \frac{s}{2N}) - (t - \frac{s}{2N})\right)N} A_N^0(N(t + \tfrac{s}{2N}), \mu) \overline{A_N^0(N(t - \tfrac{s}{2N}), \mu)} d\mu \, ds$$

$$= \frac{\Sigma_L}{(2\pi)^2} \int_{\mathbb{R}} e^{-i\lambda s} \int_{\mathbb{R}} e^{i\mu s} A_N^0(N(t + \tfrac{s}{2N}), \mu) \overline{A_N^0(N(t - \tfrac{s}{2N}), \mu)} d\mu \, ds \text{ and}$$

$$f(t, \lambda) = \frac{\Sigma_L}{2\pi} |A(t, \lambda)|^2 = \frac{\Sigma_L}{(2\pi)^2} \int_{\mathbb{R}} e^{-i\lambda s} \int_{\mathbb{R}} e^{i\mu s} |A(t, \lambda)|^2 d\mu \, ds.$$

We note that, due to the differentiability condition (c), the function $A(t, \cdot)$ is continuous and in L^2. Then, an application of Lemma 5.4 gives $A(t, \cdot) \in L^4$, which implies $f(t, \cdot) \in L^2$. Moreover, for all $t \in \mathbb{R}$, $N \in \mathbb{N}$ we obtain from Plancherel's theorem, the Cauchy–Schwarz inequality and the integration by parts formula for some $C > 0$

$$\int_{\mathbb{R}} |f_N(t, \lambda)|^2 \, d\lambda = \frac{\Sigma_L}{4\pi^2} \int_{\mathbb{R}} \left| \int_{\mathbb{R}} e^{-i\lambda s} \int_{\mathbb{R}} e^{i\mu s} A_N^0(N(t + \tfrac{s}{2N}), \mu) \overline{A_N^0(N(t - \tfrac{s}{2N}), \mu)} d\mu \, ds \right|^2 d\lambda$$

$$= C \int_{\mathbb{R}} \left| \int_{\mathbb{R}} e^{i\mu s} A_N^0(N(t + \tfrac{s}{2N}), \mu) \overline{A_N^0(N(t - \tfrac{s}{2N}), \mu)} d\mu \right|^2 ds$$

$$\leq C \int_{|s| < 1} \|A_N^0(N(t + \tfrac{s}{2N}), \cdot)\|_{L^2}^2 \|A_N^0(N(t - \tfrac{s}{2N}), \cdot)\|_{L^2}^2 ds$$

$$+ C \int_{|s| \geq 1} \left| \int_{\mathbb{R}} e^{i\mu s} A_N^0(N(t + \tfrac{s}{2N}), \mu) \overline{A_N^0(N(t - \tfrac{s}{2N}), \mu)} d\mu \right|^2 ds$$

$$\leq 2CK^4 + C \int_{|s| < 1} \left| 0 - \int_{\mathbb{R}} \frac{e^{i\mu s}}{is} \left(A_N^0(N(t + \tfrac{s}{2N}), \mu) \frac{d}{d\mu} \overline{A_N^0(N(t + \tfrac{s}{2N}), \mu)} \right. \right.$$

$$\left. \left. + \frac{d}{d\mu} A_N^0(N(t + \tfrac{s}{2N}), \mu) \overline{A_N^0(N(t + \tfrac{s}{2N}), \mu)} \right) d\mu \right|^2 ds$$

$$\leq 2CK^4 + 4CK^4 \int_{|s| \geq 1} \frac{1}{s^2} ds < \infty,$$

where $\lim_{\mu \to \pm\infty} A_N^0(t, \mu) = 0$, since $A_N^0(t, \cdot)$ is continuous and in $L^2(\mathbb{R})$ for all $t \in \mathbb{R}$. Finally, the integral $\int_{\mathbb{R}} |f_N(t, \lambda) - f(t, \lambda)|^2 d\lambda$ is well defined, since $f_N(t, \cdot)$ and $f(t, \cdot)$ are both in $L^2(\mathbb{R})$ for all $t \in \mathbb{R}$. From Plancherel's theorem we obtain

$$(2\pi)^2 \int_{\mathbb{R}} |f_N(t, \lambda) - f(t, \lambda)|^2 \, d\lambda$$

$$= \frac{\Sigma_L^2}{2\pi} \int_{\mathbb{R}} \left| \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\lambda s} \left(\int_{\mathbb{R}} e^{i\mu s} \left(A_N^0(N(t+\tfrac{s}{2N}),\mu) \overline{A_N^0(N(t-\tfrac{s}{2N}),\mu)} - A(t,\mu)\overline{A(t,\mu)} \right) d\mu \right) ds \right|^2 d\lambda$$

$$= \frac{\Sigma_L^2}{2\pi} \int_{\mathbb{R}} \left| \int_{\mathbb{R}} e^{i\mu s} \left(A_N^0(N(t+\tfrac{s}{2N}),\mu) \overline{A_N^0(N(t-\tfrac{s}{2N}),\mu)} - A(t,\mu)\overline{A(t,\mu)} \right) d\mu \right|^2 ds$$

$$=: \frac{\Sigma_L^2}{2\pi} \int_{\mathbb{R}} |\widehat{a}_N(\tfrac{s}{2N})|^2 \, ds,$$

where $\widehat{a}_N(\tfrac{s}{2N}) = \int_{\mathbb{R}} e^{i\mu s} a_N(\tfrac{s}{2N},\mu) d\mu$. It is left to show

$$\int_{\mathbb{R}} |\widehat{a}_N(\tfrac{s}{2N})|^2 \, ds \xrightarrow[N\to\infty]{} 0. \tag{26}$$

The proof of (26) consists of several steps. We start by showing $\widehat{a}_N(\tfrac{s}{2N}) \xrightarrow[N\to\infty]{} 0$. Indeed, for fixed $s,t \in \mathbb{R}$ we obtain

$$|\widehat{a}_N(\tfrac{s}{2N})| \leq \int_{\mathbb{R}} \left| A_N^0(N(t+\tfrac{s}{2N}),\mu) \overline{A_N^0(N(t-\tfrac{s}{2N}),\mu)} - A(t,\mu)\overline{A(t,\mu)} \right| d\mu$$

$$\leq \int_{\mathbb{R}} \left| A_N^0(N(t+\tfrac{s}{2N}),\mu) - A(t,\mu) \right| \left| \overline{A_N^0(N(t-\tfrac{s}{2N}),\mu)} \right|$$

$$+ |A(t,\mu)| \left| \overline{A_N^0(N(t-\tfrac{s}{2N}),\mu)} - \overline{A(t,\mu)} \right| d\mu$$

$$\leq \left\| A_N^0(N(t+\tfrac{s}{2N}),\cdot) - A(t,\cdot) \right\|_{L^2} \left\| A_N^0(N(t-\tfrac{s}{2N}),\cdot) \right\|_{L^2}$$

$$+ \|A(t,\cdot)\|_{L^2} \left\| A_N^0(N(t-\tfrac{s}{2N}),\cdot) - A(t,\cdot) \right\|_{L^2}.$$

Now, due to condition (a), it holds $\left\| A_N^0(N(t-\tfrac{s}{2N}),\cdot) - A(t,\cdot) \right\|_{L^2} \to 0$ as $N \to \infty$. Moreover, using the conditions (b) and (c), there exists a constant D, which may depend on s and t, such that

$$\|A(t,\cdot)\|_{L^2} \leq D \text{ and}$$

$$\left\| A_N^0(N(t\pm\tfrac{s}{2N}),\cdot) \right\|_{L^2} \leq \left\| A_N^0(N(t\pm\tfrac{s}{2N}),\cdot) - A(t,\cdot) \right\|_{L^2} + \|A(t,\cdot)\|_{L^2} \leq D$$

for sufficiently large N. Thus

$$|\widehat{a}_N(\tfrac{s}{2N})| \leq \left\| A_N^0(N(t+\tfrac{s}{2N}),\cdot) - A(t,\cdot) \right\|_{L^2}^2 \left\| A_N^0(N(t-\tfrac{s}{2N}),\cdot) \right\|_{L^2}^2$$

$$+ \|A(t,\cdot)\|_{L^2}^2 \left\| A_N^0(N(t-\tfrac{s}{2N}),\cdot) - A(t,\cdot) \right\|_{L^2}^2 \to 0, \tag{27}$$

as $N \to \infty$.

Next, we show $|\widehat{a}_N(\tfrac{s}{2N})| \leq \frac{E}{|s|}$, for all $s \in \mathbb{R}$, sufficiently large $N \in \mathbb{N}$ and some constant $E > 0$, which may depend on t. On the one hand, we have

$$\int_{\mathbb{R}} \left| \tfrac{d}{d\mu} a_N(\tfrac{s}{2N},\mu) \right| d\mu = \int_{\mathbb{R}} \left| \left(\tfrac{d}{d\mu} A_N^0(N(t+\tfrac{s}{2N}),\mu) \right) \overline{A_N^0(N(t-\tfrac{s}{2N}),\mu)} \right.$$

$$+ A_N^0(N(t+\tfrac{s}{2N}),\mu) \left(\tfrac{d}{d\mu} \overline{A_N^0(N(t-\tfrac{s}{2N}),\mu)} \right)$$

$$\left. - \left(\tfrac{d}{d\mu} A(t,\mu) \right) \overline{A(t,\mu)} - A(t,\mu) \left(\tfrac{d}{d\mu} \overline{A(t,\mu)} \right) \right| d\mu \tag{28}$$

$$\leq \left\| \tfrac{d}{d\mu} A_N^0(N(t+\tfrac{s}{2N}),\cdot) \right\|_{L^2} \left\| A_N^0(N(t-\tfrac{s}{2N}),\cdot) \right\|_{L^2}$$

$$+ \left\| A_N^0(N(t+\tfrac{s}{2N}),\cdot) \right\|_{L^2} \left\| \tfrac{d}{d\mu} A_N^0(N(t-\tfrac{s}{2N}),\cdot) \right\|_{L^2}$$

$$+ 2 \|A(t,\cdot)\|_{L^2} \left\| \tfrac{d}{d\mu} A(t,\cdot) \right\|_{L^2} \leq E,$$

25

where the last inequality follows from (b) and (c). On the other hand, $\frac{d}{d\mu}a_N(\frac{s}{2N},\mu) \in L^1(\mathbb{R})$, since $A_N^0(t,\cdot), A(t,\cdot), \frac{d}{d\mu}A_N^0(t,\cdot), \frac{d}{d\mu}A(t,\cdot) \in L^2(\mathbb{R})$, such that

$$
\int_{\mathbb{R}} e^{i\mu s}\left(\frac{d}{d\mu}a_N(\tfrac{s}{2N},\mu)\right) d\mu = \left[e^{i\mu s}a_N(\tfrac{s}{2N},\mu)\right]\Big|_{\mu=-\infty}^{\infty} - \int_{\mathbb{R}}(is)e^{i\mu s}a_N(\tfrac{s}{2N},\mu)d\mu
$$
$$
= (-is)\widehat{a}_N(\tfrac{s}{2N}), \tag{29}
$$

where the limit in the first term of the partial integration is zero, because $a_N(\frac{s}{2N},\mu)$ is continuous and in $L^1(\mathbb{R})$. Combining (28) and (29) we obtain

$$
|\widehat{a}_N(\tfrac{s}{2N})| \leq \frac{1}{|s|}\int_{\mathbb{R}}\left|\frac{d}{d\mu}a_N(\tfrac{s}{2N},\mu)\right|d\mu \leq \frac{E}{|s|}.
$$

Finally, for $s^* \in \mathbb{R}$

$$
(2\pi)^2\int_{\mathbb{R}}|f_N(t,\lambda)-f(t,\lambda)|^2\,d\lambda = \frac{\Sigma_L^2}{2\pi}\int_{s\in\mathbb{R}}\left|\widehat{a}_N(\tfrac{s}{2N})\right|^2 ds
$$
$$
= \frac{\Sigma_L^2}{2\pi}\int_{|s|\geq s^*}\left|\widehat{a}_N(\tfrac{s}{2N})\right|^2 ds + \frac{\Sigma_L^2}{2\pi}\int_{|s|<s^*}\left|\widehat{a}_N(\tfrac{s}{2N})\right|^2 ds
$$
$$
\leq \frac{\Sigma_L^2 E^2}{\pi s^*} + \frac{\Sigma_L^2}{2\pi}\int_{|s|<s^*}\left|\widehat{a}_N(\tfrac{s}{2N})\right|^2 ds.
$$

The second term converges to zero by the dominated convergence theorem, where pointwise convergence follows from (27) and a convergent majorant can be obtained from the boundedness conditions in (b) and the Cauchy-Schwarz inequality, noting that the support of the integral is compact. Therefore, for all $\varepsilon > 0$ and sufficiently large s^* and N it holds

$$
\int_{\mathbb{R}}|f_N(t,\lambda)-f(t,\lambda)|^2\,d\lambda \leq \frac{\Sigma_L^2 E^2}{\pi s^*} + \frac{\Sigma_L^2}{2\pi}\int_{|s|<s^*}\left|\widehat{a}_N(\tfrac{s}{2N})\right|^2 ds \leq \varepsilon,
$$

which concludes the proof. $\qquad\square$

Corollary 5.5. *Let $Y_N(t)$ be a sequence of time-varying linear state space processes as in Definition 4.10, such that both IVPs in (24) are uniformly exponentially stable, the conditions (C1)-(C3) from Proposition 4.11 hold and $\sup_{t\in\mathbb{R}}\|B(t)\| < \infty$. Then, the sequence of Wigner-Ville spectra tends in mean square to the time-varying spectral density.*

Proof. It is sufficient to check the conditions (a), (b) and (c) from Theorem 5.2.

(a) For $s,t \in \mathbb{R}$ we obtain from Plancherel's theorem

$$
\left\|A_N^0(N(t\pm\tfrac{s}{2N}),\cdot) - A(t,\cdot)\right\|_{L^2}^2 = 4\pi^2\left\|g_N^0(N(t\pm\tfrac{s}{2N}),\cdot) - g(t,\cdot)\right\|_{L^2}^2
$$
$$
= 4\pi^2\int_{\mathbb{R}}\mathbb{K}_{\{u\leq0\}}\left|B(t\pm\tfrac{s}{2N})'\Psi_{N,t\pm\frac{s}{2N}}^0(0,u)C(\tfrac{u}{N}+t\pm\tfrac{s}{2N}) - B(t)'\Psi_t(0,u)C(t)\right|^2 du,
$$

which tends to zero as $N \to \infty$ by the dominated convergence theorem. Pointwise convergence is secured by the continuity of A, B and C in (C1) and the continuity of the solution of an IVP on the input (see the proof of Proposition 4.11). Since the sequence $Y_N(t)$ is uniformly exponentially stable, we have $\left\|\Psi_{N,t}^0(s_1,s_0)\right\| \leq \gamma e^{-\lambda(s_1-s_0)}$ for some $\gamma,\lambda > 0$ and all $s_1 > s_0$. Therefore, a convergent majorant can be obtained by noting that

$$
\mathbb{K}_{\{u\leq0\}}\left|B(t\pm\tfrac{s}{2N})'\Psi_{t\pm\frac{s}{2N}}(0,u)C(\tfrac{u}{N}+t\pm\tfrac{s}{2N})\right| \leq \mathbb{K}_{\{u\leq0\}}\left(\sup_{t\in\mathbb{R}}\|B(t)\|\right)\gamma e^{\lambda u}\left(\sup_{t\in\mathbb{R}}\|C(t)\|\right).
$$

(b) For $t \in \mathbb{R}$ and $N \in \mathbb{N}$ it holds $\left\| \Psi_{N,t}^0(s_1, s_0) \right\| \le \gamma e^{-\lambda(s_1 - s_0)}$ and $\left\| \Psi_t(s_1, s_0) \right\| \le \gamma e^{-\lambda(s_1 - s_0)}$ for some $\gamma, \lambda > 0$ and all $s_1 > s_0$. Thus

$$\left\| A_N^0(Nt, \cdot) \right\|_{L^2}^2 = 4\pi^2 \left\| g_N^0(Nt, \cdot) \right\|_{L^2}^2 = 4\pi^2 \int_{\mathbb{R}} \mathbb{K}_{\{u \le 0\}} \left| B(t)' \Psi_{N,t}^0(0, u) C(\tfrac{u}{N} + t) \right|^2 du$$

$$\le \frac{2\pi^2 \gamma^2}{\lambda} \left(\sup_{t \in \mathbb{R}} \|B(t)\| \right)^2 \left(\sup_{t \in \mathbb{R}} \|C(t)\| \right)^2 < \infty \text{ and}$$

$$\|A(t, \cdot)\|_{L^2}^2 = 4\pi^2 \|g(t, \cdot)\|_{L^2}^2 = 4\pi^2 \int_{\mathbb{R}} \mathbb{K}_{\{u \le 0\}} \left\| B(t)' \Psi_t(0, u) C(u) \right\|^2 du$$

$$\le \frac{2\pi^2 \gamma^2}{\lambda} \left(\sup_{t \in \mathbb{R}} \|B(t)\| \right)^2 \left(\sup_{t \in \mathbb{R}} \|C(t)\| \right)^2 < \infty.$$

(c) Since $Y_N(t)$ is uniformly exponentially stable, [20, Theorem 1.6] implies that

$$\tfrac{d}{d\mu} A_N^0(Nt, \mu) = \int_{\mathbb{R}} e^{-i\mu u}(-iu) g_N^0(t, u) du \text{ and}$$

$$\tfrac{d}{d\mu} A(t, \mu) = \int_{\mathbb{R}} e^{-i\mu u}(-iu) g(t, u) du,$$

which are again in $L^2(\mathbb{R})$, since

$$\left\| \tfrac{d}{d\mu} A_N^0(Nt, \cdot) \right\|_{L^2}^2 = 4\pi^2 \left\| (-i\cdot) g_N^0(Nt, \cdot) \right\|_{L^2}^2$$

$$= 4\pi^2 \int_{\mathbb{R}} \mathbb{K}_{\{u \le 0\}} \left| (-iu) B(t)' \Psi_{N,t}^0(0, u) C(\tfrac{u}{N} + t) \right|^2 du$$

$$\le 4\pi^2 \left(\sup_{t \in \mathbb{R}} \|B(t)\| \right)^2 \left(\gamma^2 \int_{-\infty}^{0} u^2 e^{2\lambda u} du \right) \left(\sup_{t \in \mathbb{R}} \|C(t)\| \right)^2$$

$$= \frac{\gamma^2 \pi^2}{\lambda^3} \left(\sup_{t \in \mathbb{R}} \|B(t)\| \right)^2 \left(\sup_{t \in \mathbb{R}} \|C(t)\| \right)^2 < \infty \text{ and analogously}$$

$$\left\| \tfrac{d}{d\mu} A(t, \cdot) \right\|_{L^2}^2 \le \frac{\gamma^2 \pi^2}{\lambda^3} \left(\sup_{t \in \mathbb{R}} \|B(t)\| \right)^2 \left(\sup_{t \in \mathbb{R}} \|C(t)\| \right)^2 < \infty.$$

$\square$

Acknowledgements

References

[1] APPLEBAUM, D (2009). *Lévy processes and stochastic calculus, 2nd ed.* Cambridge University Press, Cambridge.

[2] BAAKE, M. AND SCHLÄGEL, U. (2011). The Peano-Baker series. *Proc. Steklov Inst. Math.* **275** 155–159.

[3] BARDET, J. M., DOUKHAN, P. AND WINTENBERGER, O. (2020). Contrast estimation of general locally stationary processes using coupling. *arXiv:2005.07397.*

[4] BENTH, F. E., KLÜPPELBERG, C., MÜLLER, G. AND VOS, L. (2014). Futures pricing in electricity markets based on stable CARMA spot models. *Energy Econ.* **44** 392–406.

[5] BENMAHAMMED, K. (1987). Model Reduction of Uniformily Controllable Continuous Time Varying Linear Systems. *American Control Conference* 1500–1503.

[6] BERNSTEIN, D. S. (2009). *Matrix mathematics: Theory, facts, and formulas, 2nd ed.* Princeton University Press.

[7] BROCKETT, R. W. (1970). *Finite dimensional linear systems.* John Wiley & Sons Inc.

[8] BROCKWELL, P. J. AND DAVIS, R. A. (1996). *Time series: theory and methods.* Springer, New York.

[9] BRUSCATO, A. AND TOLOI, C. M. C. (2004). Spectral analysis of non-stationary processes using the Fourier transform. *Braz. J. Probab.* **18** 69–102.

[10] CHANDRASEKHARAN, K. (1989). *Classical Fourier Transforms.* Springer, Berlin.

[11] DAHLHAUS, R. (1996). On the Kullback-Leibler information divergence of locally stationary processes. *Stoch. Proc. Appl.* **62** 139–168.

[12] DAHLHAUS, R. (1997). Fitting time series models to nonstationary processes. *Ann. Stat.* **25** 1–37.

[13] DAHLHAUS, R. (2000). A likelihood approximation for locally stationary processes. *Ann. Stat.* **28** 1762–1794.

[14] DAHLHAUS, R. (2012). Locally stationary processes. In *Handbook of statistics: Time series analysis: Methods and applications* **30** 351–413.

[15] DAHLHAUS, R. AND POLONIK, W. (2009). Empirical spectral processes for locally stationary time series. *Bernoulli* **15** 1–39.

[16] DAHLHAUS, R., RICHTER, S. AND WU, W. B. (2019). Towards a general theory for nonlinear locally stationary processes. *Bernoulli* **25** 1013–1044.

[17] DAHLHAUS, R. AND SUBBA RAO, T. (2006). Statistical inference for time-varying ARCH processes. *Ann. Stat.* **34** 1075–1114.

[18] FLANDRIN, P. AND MARTIN, W. (1984). A general class of estimators for the Wigner-Ville spectrum of non-stationary processes. In *Analysis and Optimization of Systems. Lecture Notes in Control and Information Sciences.* 15–23. Springer, Berlin.

[19] HORN, R. A. AND JOHNSON C. R. (1990). *Matrix Analysis.* Cambridge University Press, Cambridge.

[20] KATZNELSON, Y. (2004). *An introduction to harmonic analysis, 3rd ed.* Cambridge Mathematical Library, Cambridge.

[21] KRYLOV, N. V. (2002). *Introduction to the Theory of Random Processes.* American Mathematical Soc., Providence.

[22] KÜNSCH, H. R. (1995). A note on causal solutions for locally stationary AR-processes. Preprint, ftp://ess.r-project.org/users/hkuensch/localstat-ar.pdf.

[23] LARSSON, E. K. AND MOSSBERG, M. (2004). Fast and approximative estimation of continuous-time stochastic signals from discrete-time data. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, Proceedings, ICASSP 04, 2* 529–532.

[24] LUKES, D. N. (1982). *Differential Equations: Classical to Controlled.* Academic Press, London.

[25] MARQUARDT, T. AND STELZER, R. (2007). Multivariate CARMA processes. *Stoch. Proc. Appl.* **117** 96–120.

[26] MARTIN, W. AND FLANDRIN, P. (1985). Wigner-Ville spectral analysis of nonstationary processes. *IEEE Trans. Acoust. Speech Signal Process.* **33** 1461–1470.

[27] PRIESTLEY, M. B. (1994). *Spectral Analysis and Time Series. Volumes I and II in 1 book.* Academic Press, London.

[28] RAMAR, K. AND RAMASWAMI, B. (1971). Transformation of time-variable multi-input systems to a canonical form. *IIEEE Trans. Autom. Control* **16** 371–374.

[29] RAMASWAMI, B. AND RAMAR, K. (1969). On the transformation of time-variable systems to the phase-variable canonical form. *IIEEE Trans. Autom. Control* **14** 417–419.

[30] ROYDEN, H. L. (1988). *Real Analysis, 3rd ed.* Macmillan, New York.

[31] RUDIN, W. (1991). *Functional Analysis, 2nd ed.* McGraw-Hill Science/Engineering/Math, New York.

[32] RUGH, W. J. (1996). *Linear System Theory, 2nd ed.* Prentice-Hall, New Jersey.

[33] SATO, K. I. (2013). *Lévy processes and infinitely divisible distributions, Cambridge studies in advanced mathematics 68.* Cambridge University Press, Cambridge.

[34] SATO, K.-I. (2014). Stochastic integrals with respect to Lévy processes and infinitely divisible distributions. *Sugaku Expositions* **27** 19–42.

[35] SCHLEMM, E. AND STELZER, R. (2012). Multivariate CARMA processes, continuous-time state space models and complete regularity of the innovations of the sampled processes. *Bernoulli* **18** 46–63.

[36] SILVERMAN, L. (1966). Transformation of time-variable systems to canonical (phase-variable) form. *IEEE Trans. Autom. Control* **11** 300–303.

[37] SURULESCU, N. M. (2010). *On some classes of continuous-time series models and their use in financial economics.* PhD thesis, Ruprecht-Karls-Universität Heidelberg.

[38] VOGT, M. (2012). Nonparametric regression for locally stationary time series. *Ann. Stat.* **40** 2601–2633.

[39] VOGT, M. AND DETTE, H. (2015). Detecting gradual changes in locally stationary processes. *Ann. Stat.* **43** 713–740.

[40] WALTER, W. (2000). *Gewöhnliche Differentialgleichungen, 7th ed.* Springer, Berlin.

[41] WU, M.-A. AND SHERIF, A. (1976). On the commutative class of linear time-varying systems. *Int. J. Control.* **23** 433–444.