Multivariate CARMA processes, continuous-time state space models and complete regularity of the innovations of the sampled processes

ECKHARD SCHLEMM* , ROBERT STELZER

TUM Institute for Advanced Study & Zentrum Mathematik
Technische Universität München
Boltzmannstr. 3
85748 Garching, Germany

E-mail: schlemm@ma.tum.de; *url:* http://www-m4.ma.tum.de

E-mail: rstelzer@ma.tum.de; *url:* http://www-m4.ma.tum.de

The class of multivariate Lévy-driven auto-regressive moving-average (MCARMA) processes, the continuous-time analogues of the classical vector ARMA processes, is shown to be equivalent to the class of continuous-time state space models. The linear innovations of the weak ARMA process arising from sampling an MCARMA process at an equidistant grid are proved to be exponentially completely regular (β-mixing) under a mild continuity assumption on the driving Lévy process. It is verified that this continuity assumption is satisfied in most practically relevant situations including the case when the driving Lévy process has a non-singular Gaussian component, is compound Poisson with an absolutely continuous jump size distribution or has an infinite Lévy measure admitting a density around zero.

AMS 2000 subject classifications: Primary 60G10; secondary 60G51, 94A20.
Keywords: complete regularity, linear innovations, multivariate CARMA process, sampling, state space representation, strong mixing, vector ARMA process.

1. Introduction

CARMA processes are the continuous-time analogue of the widely known discrete-time ARMA processes (see e.g. [8] for a comprehensive introduction); they have first been defined in [12] in the univariate Gaussian setting and have stimulated a considerable amount of research in recent years (see e.g. [7] and references therein). In particular, the restriction of the driving process to Brownian motion was relaxed and [5] allowed for Lévy processes with a finite logarithmic moment. Because of their applicability to irregularly spaced observations and high-frequency data they have turned out to be a versatile and powerful tool in the modeling of phenomena from the natural sciences, engineering and finance. Recently, [18] extended the concept to multivariate CARMA (MCARMA) processes with the intention to be able to model the joint behavior of several dependent time series. MCARMA processes are thus the continuous-time analogue of discrete-time vector ARMA (VARMA) models (see for instance [17]).

The aim of this paper is twofold: first we establish the equivalence between MCARMA and multivariate continuous-time state space models, a correspondence which is well known in the discrete-time setting ([14]). The second topic is the investigation of the probabilistic properties of the discrete-time process obtained by recording the values of an MCARMA process at discrete, equally spaced points in time. A detailed understanding of the innovations of the arising weak VARMA process is a prerequisite for proving asymptotic properties of statistics of a discretely observed MCARMA process. One notion of asymptotic independence which is very useful in this context is complete regularity (see section 4 for a precise definition) and we show that the innovations of a discretized MCARMA process have this desirable property. Our results therefore not only add important insight into the probabilistic structure of CARMA processes bur are also fundamental to the development of an estimation theory for non-Gaussian continuous-time state space models based on equidistant observations.

In this paper we show that a sampled MCARMA process is a discrete-time VARMA process with dependent innovations. While the mixing behaviour of ARMA and more generally linear processes is fairly well understood (see e.g. [1, 19, 20]), the mixing properties of the innovations of a sampled continuous-time process have received very little attention in the past. From [9] it is only known that the innovations of a discretized univariate Lévy-driven CARMA process are weak white noise, which, by itself, is typically of little help in applications. We show that the linear innovations of a sampled MCARMA process satisfy a set of VARMA equations and we conclude that under a mild continuity assumption on the driving Lévy process they are geometrically completely regular and in particular geometrically strongly mixing. This continuity assumption is further shown to be satisfied for most of the practically relevant choices of the driving Lévy process including processes with a non-singular Gaussian component as well as compound Poisson processes with an absolutely continuous jump size distribution and infinite-activity processes whose Lévy measure admits a density in a neighborhood of zero.

This paper is structured as follows: In section 2 we review some well-known properties of Lévy processes, which we use later. The class of multivariate CARMA processes, in a slightly more general form than in the original definition of [18], is described in detail in section 3 and shown to be equivalent to the class of continuous-time state space models. In section 4, the main result about the mixing properties of the sampled processes is stated and demonstrated to be applicable in many practical situations. The proofs of the results are presented in section 5.1.

We use the following notation: The space of $m \times n$-matrices with entries in the ring $\mathbb{K}$ is denoted by $M_{m,n}(\mathbb{K})$ or $M_m(\mathbb{K})$ if $m = n$. A^T is the transpose of the matrix A, the matrices $\mathbb{I}_m$ and 0_m are the identity and the zero element of $M_m(\mathbb{K})$ and $A \otimes B$ stands for the Kronecker product of the matrices A and B. The zero vector in $\mathbb{R}^m$ is denoted by $\mathbf{0}_m$ and $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ represent the Euclidean norm and inner product. Finally, $\mathbb{K}[z]$ ($\mathbb{K}\{z\}$) is the ring of polynomial (rational) expressions in z over $\mathbb{K}$ and $I_B(\cdot)$ is the indicator function of the set B.

2. Multivariate Lévy processes

In this section we review the definition and some elementary facts about multivariate Lévy processes which we will use later. More details and proofs can be found in e.g. [22].

***Definition* 2.1.** *A (one-sided) $\mathbb{R}^m$-valued* Lévy process $(\boldsymbol{L}(t))_{t\geq 0}$ *is a stochastic process with stationary, independent increments, continuous in probability and satisfying $\boldsymbol{L}(0) = \mathbf{0}_m$ almost surely.*

Every $\mathbb{R}^m$-valued Lévy process$(\boldsymbol{L}(t))_{t\geq 0}$ can be assumed to be càdlàg and is completely characterized by its characteristic function in the Lévy-Khintchine form $\mathbb{E}e^{i\langle \boldsymbol{u},\boldsymbol{L}(t)\rangle} = \exp\{t\psi^{\boldsymbol{L}}(\boldsymbol{u})\}$, $\boldsymbol{u} \in \mathbb{R}^m$, $t \geq 0$, where $\psi^{\boldsymbol{L}}$ has the special form

$$\psi^{\boldsymbol{L}}(\boldsymbol{u}) = i\langle \boldsymbol{\gamma}, \boldsymbol{u}\rangle - \frac{1}{2}\langle \boldsymbol{u}, \Sigma^{\mathcal{G}}\boldsymbol{u}\rangle + \int_{\mathbb{R}^m}\left[e^{i\langle \boldsymbol{u},\boldsymbol{x}\rangle} - 1 - i\langle \boldsymbol{u}, \boldsymbol{x}\rangle I_{\{\|x\|\leq 1\}}\right]\nu^{\boldsymbol{L}}(\mathrm{d}\boldsymbol{x}).$$

The vector $\gamma \in \mathbb{R}^m$ is called the drift, the non-negative definite, symmetric $m\times m$-matrix $\Sigma^{\mathcal{G}}$ is the Gaussian covariance matrix and $\nu^{\boldsymbol{L}}$ is a measure on $\mathbb{R}^m$, referred to as the Lévy measure, satisfying

$$\nu^{\boldsymbol{L}}(\{\mathbf{0}_m\}) = 0, \qquad \int_{\mathbb{R}^m}\min(\|\boldsymbol{x}\|^2, 1)\nu^{\boldsymbol{L}}(\mathrm{d}\boldsymbol{x}) < \infty.$$

We will work with two-sided Lévy processes $\boldsymbol{L} = (\boldsymbol{L}(t))_{t\in\mathbb{R}}$. These are obtained from two independent copies $(\boldsymbol{L}_1(t))_{t\geq 0}$, $(\boldsymbol{L}_2(t))_{t\geq 0}$ of a one-sided Lévy process via the construction

$$\boldsymbol{L}(t) = \begin{cases}\boldsymbol{L}_1(t) & t \geq 0\\ -\lim_{s\nearrow -t}\boldsymbol{L}_2(s) & t < 0\end{cases}.$$

Throughout the paper we restrict attention to Lévy processes with mean zero and finite second moments.

***Assumption* L1.** The Lévy process $\boldsymbol{L}$ satisfies $\mathbb{E}\boldsymbol{L}(1) = 0$ and $\mathbb{E}\|\boldsymbol{L}(1)\|^2 < \infty$.

The assumption $\mathbb{E}\boldsymbol{L}(1) = 0$ is made only for notational convenience and is not essential for our results to hold. The premise that $\boldsymbol{L}$ have finite variance, in contrast, is a true restriction, which is very often made in the analysis of (C)ARMA processes. The treatment of the infinite variance case requires different techniques and often does not lead to comparable results. It is well-known that $\boldsymbol{L}$ has finite second moments if and only if $\int_{\|\boldsymbol{x}\|\geq 1}\|\boldsymbol{x}\|^2\nu(\mathrm{d}\boldsymbol{x})$ is finite and that $\Sigma^{\boldsymbol{L}} = \mathbb{E}\boldsymbol{L}(1)\boldsymbol{L}(1)^T$ is then given by $\int_{R^m}\boldsymbol{x}\boldsymbol{x}^T\nu^{\boldsymbol{L}}(\mathrm{d}\boldsymbol{x})$.

3. MCARMA processes and state space models

If $\boldsymbol{L}$ is a two-sided Lévy process with values in $\mathbb{R}^m$ and $p > q$ are positive integers, then the d-dimensional $\boldsymbol{L}$-driven auto-regressive moving-average (*MCARMA*) process with auto-regressive polynomial

$$z \mapsto P(z) := \mathbb{I}_d z^p + A_1 z^{p-1} + \ldots + A_p \in M_d(\mathbb{R}[z]) \tag{3.1a}$$

and moving-average polynomial

$$z \mapsto Q(z) := B_0 z^q + B_1 z^{q-1} + \ldots + B_q \in M_{d,m}(\mathbb{R}[z]) \tag{3.1b}$$

is thought of as the solution to the formal differential equation

$$P(D)\boldsymbol{Y}(t) = Q(D)D\boldsymbol{L}(t), \quad D \equiv \frac{\mathrm{d}}{\mathrm{d}t}, \tag{3.2}$$

which is the continuous-time analogue of the discrete-time ARMA equations. We note that we allow for the driving Lévy process $\boldsymbol{L}$ and the $\boldsymbol{L}$-driven MCARMA process to have different dimensions and thus slightly extend the original defintion of [18]. All the results we need from [18] are easily seen to hold true also in this more general setting. Since, in general, Lévy processes are not differentiable equation (3.2) is purely formal and, as usual, interpreted as being equivalent to the *state space representation*

$$\mathrm{d}\boldsymbol{G}(t) = \mathcal{A}\boldsymbol{G}(t)\mathrm{d}t + \mathcal{B}\mathrm{d}\boldsymbol{L}(t), \quad \boldsymbol{Y}(t) = C\boldsymbol{G}(t), \quad t \in \mathbb{R}, \tag{3.3}$$

where $\mathcal{A}, \mathcal{B}, C$ are given by

$$\mathcal{A} = \begin{pmatrix} 0 & \mathbb{I}_d & 0 & \ldots & 0 \\ 0 & 0 & \mathbb{I}_d & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \mathbb{I}_d \\ -A_p & -A_{p-1} & \ldots & \ldots & -A_1 \end{pmatrix} \in M_{pd}(\mathbb{R}), \tag{3.4a}$$

$$\mathcal{B} = \begin{pmatrix} \beta_1^T & \cdots & \beta_p^T \end{pmatrix}^T \in M_{pd,m}(\mathbb{R}), \quad \beta_{p-j} = -I_{\{0,\ldots,q\}}(j)\left[\sum_{i=1}^{p-j-1} A_i\beta_{p-j-i} + B_{q-j}\right] \quad \text{and} \tag{3.4b}$$

$$C = (\mathbb{I}_d, 0_d, \ldots, 0_d) \in M_{d,pd}(\mathbb{R}). \tag{3.4c}$$

In view of representation (3.3), MCARMA processes are linear continuous-time state space models. We will consider this class of processes and see that it is in fact equivalent to the class of MCARMA models.

***Definition* 3.1.** *An $\mathbb{R}^d$-valued* continuous-time linear state space model $(A, B, C, \boldsymbol{L})$ *of dimension N is characterized by an $\mathbb{R}^m$-valued driving Lévy process* $\mathbf{L}$*, a state transition matrix $A \in M_N(\mathbb{R})$, an input matrix $B \in M_{N,m}(\mathbb{R})$ and an observation matrix $C \in M_{d,N}(\mathbb{R})$. It consists of a state equation of Ornstein-Uhlenbeck type*

$$\mathrm{d}\boldsymbol{X}(t) = A\boldsymbol{X}(t)\mathrm{d}t + B\mathrm{d}\boldsymbol{L}(t) \tag{3.5a}$$

and an observation equation

$$\boldsymbol{Y}(t) = C\boldsymbol{X}(t). \tag{3.5b}$$

The $\mathbb{R}^N$-valued process $\boldsymbol{X} = (\boldsymbol{X}(t))_{t\in\mathbb{R}}$ is the state vector process and $\boldsymbol{Y} = (\boldsymbol{Y}(t))_{t\in\mathbb{R}}$ is the output process.

A solution $\boldsymbol{Y}$ to (3.5) is called *causal* if, for all t, $\boldsymbol{Y}(t)$ is independent of the σ-algebra generated by $\{\boldsymbol{L}(s) : s > t\}$. Every solution to (3.5a) satisfies

$$\boldsymbol{X}(t) = e^{A(t-s)}\boldsymbol{X}(s) + \int_s^t e^{A(t-u)}B\mathrm{d}\boldsymbol{L}(u), \quad s,t \in \mathbb{R}, \quad s < t, \tag{3.6}$$

The independent-increment property of Lévy processes implies that $\boldsymbol{X}$ is a Markov process. We always work under a standard causal stationarity assumption:

***Assumption* E1.** The eigenvalues of A have strictly negative real parts.

The following is well-known ([24]) and recalls conditions for the existence of a stationary causal solution of the state equation (3.5a) for easy reference.

Proposition 3.2. *If* **L1**, **E1** *hold then* (3.5a) *has a unique strictly stationary, causal solution* $\boldsymbol{X}$ *given by*

$$\boldsymbol{X}(t) = \int_{-\infty}^t e^{A(t-u)}B\mathrm{d}\boldsymbol{L}(u), \quad t \in \mathbb{R}, \tag{3.7}$$

which has the same distribution as $\int_0^\infty e^{Au}B\mathrm{d}\boldsymbol{L}(u)$*. Moreover,* $\boldsymbol{X}(t)$ *has mean zero,*

$$\mathrm{Var}(\boldsymbol{X}(t)) = \mathbb{E}\boldsymbol{X}(t)\boldsymbol{X}(t)^T =: \Gamma_0 = \int_0^\infty e^{Au}B\Sigma^{\boldsymbol{L}}B^Te^{A^Tu}\mathrm{d}u, \tag{3.8a}$$

$$\mathrm{Cov}\,(\boldsymbol{X}(t+h), \boldsymbol{X}(t)) = \mathbb{E}\boldsymbol{X}(t+h)\boldsymbol{X}(t)^T = e^{Ah}\Gamma_0, \quad h \geq 0, \tag{3.8b}$$

and Γ_0 *satisfies* $A\Gamma_0 + \Gamma_0A^T = -B\Sigma^{\boldsymbol{L}}B^T$.

It is an immediate consequence that the output process $\boldsymbol{Y}$ has mean zero and autocovariance function $h \mapsto \gamma_{\boldsymbol{Y}}(h) = Ce^{Ah}\Gamma_0C^T$, and that $\boldsymbol{Y}$ can be written as a moving average of the driving Lévy process as

$$\boldsymbol{Y}(t) = \int_{-\infty}^\infty g(t-u)\mathrm{d}\boldsymbol{L}(u), \quad t \in \mathbb{R}; \quad g(t) = Ce^{At}BI_{[0,\infty)}(t). \tag{3.9}$$

This equation serves, in conjunction with (3.4), as the definition of an MCARMA process with auto-regressive and moving-average polynomials given by (3.1). It shows that the behaviour of the process $\boldsymbol{Y}$ depends on the values of the individual matrices A, B, C only through the products $Ce^{At}B$, $t \in \mathbb{R}$. These products in turn are intimately related to the rational matrix function $H : z \mapsto C(z\mathbb{I}_N - A)^{-1}B$, which is called the *transfer function* of the state space model (3.5). A pair (P, Q), $P \in M_d(\mathbb{R}[z])$, $Q \in M_{d,m}(\mathbb{R}[z])$, of rational matrix functions is a *left matrix-fraction description* for the rational matrix function $H \in M_d(\mathbb{R}\{z\})$ if $P(z)^{-1}Q(z) = H(z)$, for all $z \in \mathbb{C}$. The next theorem gives an answer to the question what other state space representations beside (3.3) can be used to define an MCARMA process. The proof is given in section 5.1.

Theorem 3.3. *If* (P, Q) *is a left matrix-fraction description for the transfer function* $z \mapsto C(z\mathbb{I}_N - A)^{-1}B$*, then the stationary solution* $\mathbf{Y}$ *of the state space model* $(A, B, C, \boldsymbol{L})$ *defined in* (3.5) *is an* $\boldsymbol{L}$*-driven MCARMA process with auto-regressive polynomial* P *and moving-average polynomial* Q*.*

Corollary 3.4. *The classes of MCARMA and continuous-time state space models are equivalent.*

Proof. By definition, every MCARMA process is the output process of a state space model. Conversely, given any state space model $(A, B, C, \boldsymbol{L})$ with output process $\boldsymbol{Y}$, [10, Appendix 2, Theorem 8] shows that the transfer function $H : z \mapsto C(zI_N - A)^{-1}B$ possesses a left matrix-fraction description $H(z) = P(z)^{-1}Q(z)$. Hence, by theorem 3.3, $\boldsymbol{Y}$ is an MCARMA process. □

4. Complete regularity of the innovations of a sampled MCARMA process

For a continuous-time stochastic process $\boldsymbol{Y} = (\boldsymbol{Y}(t))_{t\in\mathbb{R}}$ and a positive constant h, the corresponding sampled process $\boldsymbol{Y}^{(h)} = (\boldsymbol{Y}^{(h)}_n)_{n\in\mathbb{Z}}$ is defined by $\boldsymbol{Y}^{(h)}_n = \boldsymbol{Y}(nh)$. A common problem in applications is the estimation of a set of model parameters based on observations of the values of a realization of a continuous-time process at equally spaced points in time. In order to make MCARMA processes amenable to parameter inference from equidistantly sampled observations it is important to have a good understanding of the probabilistic properties of $\boldsymbol{Y}^{(h)}$. One such property which has turned out useful for the derivation of asymptotic properties of estimators is *mixing*, for which there are several different notions (see, e.g., [4] for a detailed exposition). Let I denote $\mathbb{Z}$ or $\mathbb{R}$. For a stationary process $\boldsymbol{X} = (X_n)_{n\in I}$ on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$ one writes $\mathscr{F}_n^m = \sigma(X_j : n < j < m)$, $-\infty \leq n < m \leq \infty$. The α-mixing coefficients $(\alpha(m))_{m\in I}$ are then defined by

$$\alpha(m) = \sup_{A\in\mathscr{F}^0_{-\infty}, B\in\mathscr{F}^\infty_m} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)|.$$

If $\lim_{m\to\infty}\alpha(m) = 0$ the process $\boldsymbol{X}$ is called *strongly mixing* and if there exist constants $C > 0$ and $0 < \lambda < 1$ such that $\alpha_m < C\lambda^m$, $m \geq 1$, it is *exponentially strongly mixing*. The β-mixing coefficients $(\beta(m))_{m\in I}$ are similarly defined as

$$\beta(m) = \mathbb{E}\sup_{B\in\mathscr{F}^\infty_m} |\mathbb{P}(B|\mathscr{F}^0_{-\infty}) - \mathbb{P}(B)|.$$

If $\lim_{m\to\infty}\beta(m) = 0$ the process $\boldsymbol{X}$ is called *completely regular* or *β-mixing* and if there exist constants $C > 0$ and $0 < \lambda < 1$ such that $\beta_m < C\lambda^m$, $m \geq 1$, it is *exponentially completely regular*. It is clear from these definitions that $\alpha(m) \leq \beta(m)$ and that (exponential) complete regularity implies (exponential) strong mixing. It has been shown in [18, Proposition 3.34] that every causal MCARMA process $\boldsymbol{Y}$ with a finite κ^{th} moment, $\kappa > 0$, is strongly mixing and this naturally carries over to the sampled process $\boldsymbol{Y}^{(h)}$. In this paper we therefore do not investigate the mixing properties of the process $\boldsymbol{Y}^{(h)}$ itself but rather of its linear innovations.

***Definition* 4.1.** *Let $(\boldsymbol{Y}_n)_{n\in\mathbb{Z}}$ be an $\mathbb{R}^d$-valued stationary stochastic process with finite second moments. Then the linear innovations $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{Z}}$ of $(\boldsymbol{Y}_n)_{n\in\mathbb{Z}}$ are defined by*

$$\boldsymbol{\varepsilon}_n = \boldsymbol{Y}_n - P_{n-1}\boldsymbol{Y}_n, \quad P_n = \text{orthogonal projection onto } \overline{\text{span}}\{\boldsymbol{Y}_\nu : -\infty < \nu \leq n\}, \tag{4.1}$$

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

From now on we work under an additional assumption, which is standard in the univariate case.

***Assumption* E2.** The eigenvalues $\lambda_1, \ldots, \lambda_N$ of the state transition matrix A in equation (3.5) are distinct.

A polynomial $p \in M_d(\mathbb{C}[z])$ is called *monic* if its leading coefficient is equal to $\mathbb{I}_d$ and *Schur-stable* if the zeros of $z \mapsto \det p(z)$ all lie in the complement of the closed unit disk. We first give a semi-explicit construction of a weak VARMA representation of $\boldsymbol{Y}^{(h)}$ with complex-valued coefficient matrices, a generalization of [6, Proposition 3].

Theorem 4.2. *Assume $\boldsymbol{Y}$ is the output process of the state space system (3.5) satisfying* **L1**, **E1**, **E2** *and $\boldsymbol{Y}^{(h)}$ is its sampled version with linear innovations $\boldsymbol{\varepsilon}^{(h)}$; define the Schur-stable polynomial $\varphi \in \mathbb{C}[z]$ by*

$$\varphi(z) = \prod_{\nu=1}^{N}\left(1 - e^{h\lambda_\nu}z\right) =: \left(1 - \varphi_1 z - \ldots - \varphi_N z^N\right). \tag{4.2}$$

Then there exists a monic Schur-stable polynomial $\Theta \in M_d(\mathbb{C}[z])$ of degree at most $N-1$ such that

$$\varphi(B)\boldsymbol{Y}^{(h)}_n = \Theta(B)\boldsymbol{\varepsilon}^{(h)}_n, \quad n \in \mathbb{Z}, \tag{4.3}$$

where B denotes the backshift operator, i.e. $B^j\boldsymbol{Y}^{(h)}_n = \boldsymbol{Y}^{(h)}_{n-j}$ for every non-negative integer j.

This result shows that a good understanding of the mixing properties of $\boldsymbol{\varepsilon}^{(h)}$ is theoretically interesting and practically relevant for the purpose of statistical inference for multivariate CARMA processes. Since the eigenvalues $\lambda_1, \ldots, \lambda_N$ of A are the roots of the characteristic polynomial $z \mapsto \det(z\mathbb{I}_N - A)$ the fundamental theorem of algebra implies that they are either real or come in complex conjugate pairs. We can therefore assume that they are ordered in such a way that for some $r \in \{0, \ldots, N\}$

$$\lambda_\nu \in \mathbb{R}, \quad 1 \leq \nu \leq r, \qquad \lambda_\nu = \overline{\lambda_{\nu+1}} \in \mathbb{C}\backslash\mathbb{R}, \quad \nu = r+1, r+3, \ldots, N-1.$$

By Lebesgue's decomposition theorem [15, Theorem 7.33], every measure μ on $\mathbb{R}^d$ can be uniquely decomposed as $\mu = \mu_c + \mu_s$, where μ_c and μ_s are absolutely continuous and singular, respectively, with respect to d-dimensional Lesbesgue measure. If μ_c is not the zero measure we say that μ has a non-trivial absolutely continuous component.

Theorem 4.3. *Assume $\boldsymbol{Y}$ is the output process of the continuous-time state space model $(A, B, C, \boldsymbol{L})$ satisfying* **L1**, **E1** *and* **E2**. *Denote by $\boldsymbol{\varepsilon}^{(h)}$ the innovations of the sampled process $\boldsymbol{Y}^{(h)}$ and further assume that the law of the $\mathbb{R}^{mN}$-valued random variable*

$$\mathscr{M}^{(h)} = \left[\begin{array}{cccccccc} \boldsymbol{M}_1^{(h)T} & \cdots & \boldsymbol{M}_r^{(h)T} & \underline{\boldsymbol{M}}_{r+1}^{(h)}{}^T & \underline{\boldsymbol{M}}_{r+3}^{(h)}{}^T & \cdots & \underline{\boldsymbol{M}}_{N-1}^{(h)}{}^T \end{array}\right]^T, \tag{4.4}$$

where

$$\underline{\boldsymbol{M}}_\nu^{(h)} = \left[\begin{array}{cc} \operatorname{Re} \boldsymbol{M}_\nu^{(h)T} & \operatorname{Im} \boldsymbol{M}_\nu^{(h)T} \end{array}\right]^T, \quad \boldsymbol{M}_\nu^{(h)} = \int_0^h e^{(h-u)\lambda_\nu} \mathrm{d}\boldsymbol{L}(u), \quad \nu = 1, \ldots, N, \tag{4.5}$$

has a non-trivial absolutely continuous component with respect to the mN-dimensional Lebesgue measure. Then $\boldsymbol{\varepsilon}^{(h)}$ is exponentially completely regular.

The assumption on the distribution of $\boldsymbol{M}^{(h)}$ made in theorem 4.3 is not very restrictive. Its verification is based on the following lemma which allows us to derive sufficient conditions in terms of the Lévy process $\boldsymbol{L}$ which show that it is indeed satisfied in most practical situations.

Lemma 4.4. *There exist matrices $G \in M_{mN}(\mathbb{R})$ and $H \in M_{mN,m}(\mathbb{R})$ such that $\mathscr{M}^{(h)} = \mathscr{M}(h)$ where $(\mathscr{M}(t))_{t\geq 0}$ is the unique solution to the stochastic differential equation*

$$\mathrm{d}\mathscr{M}(t) = G\mathscr{M}(t)\mathrm{d}t + H\mathrm{d}\boldsymbol{L}(t), \quad \boldsymbol{M}(0) = \boldsymbol{0}_{mN}. \tag{4.6}$$

Moreover, rank $H = m$ *and the $mN \times mN$-matrix $\left[\begin{array}{cccc} H & GH & \cdots & G^{N-1}H \end{array}\right]$ is non-singular.*

The last part of the statement is referred to as *controllability* of the pair (G, H) and is essential in the proofs of the following explicit sufficient conditions for theorem 4.3 to hold.

Proposition 4.5. *Assume that the Lévy process $\boldsymbol{L}$ has a non-singular Gaussian covariance matrix $\Sigma^{\mathcal{G}}$. Then theorem 4.3 holds.*

Proof. By [23, Corollary 2.19], the law of $\boldsymbol{M}^{(h)}$ is infinitely divisible with Gaussian covariance matrix given by $\int_0^h e^{Gu} H \Sigma^{\mathcal{G}} H^T e^{G^T u} \mathrm{d}u$. By the controllability of (G, H) and [3, Lemma 12.6.2] this matrix is non-singular and [22, E 29.14.] concludes the proof. □

A simple Lévy process of practical importance which does not have a non-singular Gaussian covariance matrix is the *compound Poisson Process* which is defined by $\boldsymbol{L}(t) = \sum_{n=1}^{N(t)} \boldsymbol{J}_n$, where $(N(t))_{t\in\mathbb{R}^+}$ is a Poisson process and $(\boldsymbol{J}_n)_{n\in\mathbb{Z}}$ is an i.i.d sequence independent of $(N(t))_{t\in\mathbb{R}^+}$; the law of $\boldsymbol{J}_n$ is called the jump size distribution. The proof of [21, Theorem 1.1] in conjuncton with Lemma 4.4 implies the following result.

Proposition 4.6. *Assume $\boldsymbol{L}$ is a compound Poisson process with absolutely continuous jump size distribution. Then Theorem 4.3 holds.*

Under a similar smoothness assumption the conclusion of Theorem 4.3 also holds in the case of infinite-activity Lévy processes. The statement follows from applying [21, Theorem 1.1] to (4.6).

Proposition 4.7. *Assume that the Lévy measure $\nu^{\boldsymbol{L}}$ of $\boldsymbol{L}$ satisfies $\nu^{\boldsymbol{L}}(\mathbb{R}^m) = \infty$ and that there exists a positive constant ρ such that $\nu^{\boldsymbol{L}}$ restricted to the ball $\{\boldsymbol{x} \in \mathbb{R}^m : \|\boldsymbol{x}\| \leq \rho\}$ has a density with respect to the m-dimensional Lebesgue measure. Then Theorem 4.3 holds.*

While the preceeding three propositions already cover a wide range of Lévy processes encountered in practice, there are some relevant cases which are not yet taken care of, in particular the construction of the Lévy process as a vector of independent univariate Lévy processes (upcoming corollary 4.11). To also cover this and related choices we employ the polar decomposition for Lévy measures [2, Lemma 2.1.]. By this result, for every Lévy measure $\nu^{\boldsymbol{L}}$ there exists a probability measure α on the $(m-1)$-sphere $S^{m-1} := \{\boldsymbol{x} \in \mathbb{R}^m : \|\boldsymbol{x}\| = 1\}$ and a family $\{\nu_{\boldsymbol{\xi}} : \boldsymbol{\xi} \in S^{m-1}\}$ of measures on $\mathbb{R}^+$ such that for each Borel set $B \in \mathcal{B}(\mathbb{R}^+)$, the function $\boldsymbol{\xi} \mapsto \nu_{\boldsymbol{\xi}}(B)$ is measurable and

$$\nu^{\boldsymbol{L}}(B) = \int_{S^{m-1}} \int_0^\infty I_B(\lambda\boldsymbol{\xi})\nu_{\boldsymbol{\xi}}(\mathrm{d}\lambda)\alpha(\mathrm{d}\boldsymbol{\xi}), \quad B \in \mathcal{B}(\mathbb{R}^m \backslash \{\boldsymbol{0}_m\}). \tag{4.7}$$

A hyperplane in a finite dimensional vector space is a linear subspace of codimension one.

Proposition 4.8. *If the Lévy measure $\nu^{\boldsymbol{L}}$ has a polar decomposition $(\alpha, \nu_{\boldsymbol{\xi}} : \boldsymbol{\xi} \in S^{m-1})$ such that for any hyperplane $\mathcal{H} \subset \mathbb{R}^m$ it holds that $\int_{S^{m-1}} I_{\mathbb{R}^m\backslash\mathcal{H}}(\boldsymbol{\xi}) \int_0^\infty \nu_{\boldsymbol{\xi}}(\mathrm{d}\lambda)\alpha(\mathrm{d}\boldsymbol{\xi}) = \infty$, then Theorem 4.3 holds.*

Proof. The proof rests upon the main Theorem of [25]. We denote by $\operatorname{im} H$ the image of the linear operator associated to the matrix H. Since $\operatorname{rank} H = m$ and the pair (G, H) is controllable we only have to show that $\nu^{\boldsymbol{L}}(\{\boldsymbol{x} \in \mathbb{R}^m : H\boldsymbol{x} \in \operatorname{im} H \backslash \mathcal{H}\}) = \infty$ for all hyperplanes $\mathcal{H} \subset \operatorname{im} H$, and since $\mathbb{R}^m \cong \operatorname{im} H$ the last condition is equivalent to $\nu^{\boldsymbol{L}}(\mathbb{R}^m \backslash \mathcal{H}) = \infty$ for all hyperplanes $\mathcal{H} \subset \mathbb{R}^m$. Using (4.7) and the fact that for every $\boldsymbol{\xi} \in S^{m-1}$ and every $\lambda \in \mathbb{R}^+$ the vector $\lambda\boldsymbol{\xi}$ is in $\mathcal{H}$ if and only if the vector $\boldsymbol{\xi}$ is, this is seen to be equivalent to the assumption of the proposition. □

Corollary 4.9. *If the Lévy measure $\nu^{\boldsymbol{L}}$ has a polar decomposition $(\alpha, \nu_{\boldsymbol{\xi}} : \boldsymbol{\xi} \in S^{m-1})$ such that $\alpha(S^{m-1} \backslash \mathcal{H})$ is positive for all hyperplanes $\mathcal{H} \in \mathbb{R}^m$ and $\nu_{\boldsymbol{\xi}}(\mathbb{R}^+) = \infty$ for α-almost every $\boldsymbol{\xi}$, then Theorem 4.3 holds.*

Corollary 4.10. *If the Lévy measure $\nu^{\boldsymbol{L}}$ has a polar decomposition $(\alpha, \nu_{\boldsymbol{\xi}} : \boldsymbol{\xi} \in S^{m-1})$ such that for some linearly independent vectors $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_m \in S^{m-1}$ it holds that $\alpha(\boldsymbol{\xi}_k) > 0$ and $\nu_{\boldsymbol{\xi}_k}(\mathbb{R}^+) = \infty$ for $k = 1, \ldots, m$, then Theorem 4.3 holds.*

Corollary 4.11. *Assume that $l \geq m$ is an integer and that the matrix $R \in M_{m,l}(\mathbb{R})$ has full rank m. If $\boldsymbol{L} = R\begin{pmatrix} L_1 & \cdots & L_l \end{pmatrix}^T$, where L_k, $k = 1, \ldots, l$, are independent univariate Lévy processes with Lévy measures ν_k^L satisfying $\nu_k^L(\mathbb{R}) = \infty$, then Theorem 4.3 holds.*

5. Proofs

5.1. Proofs for section 3

Proof of Theorem 3.3. The first step in the proof is to show that the pair (P, Q) is a left-matrix fraction description of $C(z\mathbb{I}_{pd} - \mathcal{A})^{-1}\mathcal{B}$, where $\mathcal{A}$, $\mathcal{B}$ and C are defined in (3.4). We first show the relation

$$(z\mathbb{I}_{pd} - \mathcal{A})^{-1}\mathcal{B} = \begin{bmatrix} w_1(z)^T & \cdots & w_p^T(z) \end{bmatrix}^T, \tag{5.1}$$

where $w_j(z) \in M_{d,m}(\mathbb{R}\{z\})$, $j = 1, \ldots p$, are defined by the equations

$$w_j(z) = \frac{1}{z}(w_{j+1}(z) + \beta_j), \quad j = 1, \ldots, p-1, \quad \text{and} \tag{5.2a}$$

$$w_p(z) = \frac{1}{z}\left(-\sum_{k=0}^{p-1} A_{p-k}w_{k+1}(z) + \beta_p\right). \tag{5.2b}$$

Since it has been shown in [18, Theorem 3.12] that $w_1(z) = P(z)^{-1}Q(z)$ this will prove the assertion. Equation (5.1) is clearly equivalent to $\mathcal{B} = (z\mathbb{I}_{pd} - \mathcal{A})\begin{bmatrix} w_1(z)^T & \cdots & w_p^T(z) \end{bmatrix}^T$, which explicitly reads

$$\begin{aligned}\beta_j &= zw_j(z) - w_{j+1}(z), \quad j = 1, \dots p-1,\\ \beta_p &= zw_p(z) + A_p w_1(z) + \dots + A_1 w_p(z)\end{aligned}$$

and is thus equivalent to (5.2). Using the spectral representation [16, Theorem 17.5]

$$e^{At} = \frac{1}{2\pi i}\int_\Gamma e^{zt}(z\mathbb{I}_N - A)^{-1}\mathrm{d}z, \quad t \in \mathbb{R}, \tag{5.3}$$

where Γ is some closed contour in $\mathbb{C}$ winding around each eigenvalue of A exactly once, it follows that

$$\begin{aligned}\boldsymbol{Y}(t) &= \int_{-\infty}^t Ce^{A(t-u)}B\mathrm{d}\boldsymbol{L}(u) = \frac{1}{2\pi i}\int_{-\infty}^t\int_\Gamma e^{z(t-u)}C(z\mathbb{I}_N - A)^{-1}B\mathrm{d}z\mathrm{d}\boldsymbol{L}(u)\\ &= \frac{1}{2\pi i}\int_{-\infty}^t\int_\Gamma e^{z(t-u)}P(z)^{-1}Q(z)\mathrm{d}z\mathrm{d}\boldsymbol{L}(u) = \frac{1}{2\pi i}\int_{-\infty}^t\int_\Gamma e^{z(t-u)}C(z\mathbb{I}_{pd} - \mathcal{A})^{-1}\mathcal{B}\mathrm{d}z\mathrm{d}\boldsymbol{L}(u)\\ &= \int_{-\infty}^t Ce^{\mathcal{A}(t-u)}\mathcal{B}\mathrm{d}\boldsymbol{L}(u),\end{aligned}$$

which is the defintion of an MCARMA process with auto-regressive polynomial P and moving-average polynomial Q. □

5.2. Proofs for section 4

In this section we present the proofs of our main results, Theorem 4.2, Theorem 4.3 and Lemma 4.4, as well as several auxiliary results. The first is a generalization of [6, Proposition 2] expressing MCARMA processes as a sum of multivariate Ornstein-Uhlenbeck processes.

Proposition 5.1. *Let $\boldsymbol{Y}$ be the the output process of the state space system (3.5) and assume that* **E2** *holds. Then, there exist vectors $\boldsymbol{s}_1, \dots, \boldsymbol{s}_N \in \mathbb{C}^m\backslash\{\boldsymbol{0}_m\}$ and $\boldsymbol{b}_1, \dots, \boldsymbol{b}_N \in \mathbb{C}^d\backslash\{\boldsymbol{0}_d\}$ such that $\boldsymbol{Y}$ can be decomposed into a sum of dependent, complex-valued Ornstein-Uhlenbeck processes as $\boldsymbol{Y}(t) = \sum_{\nu=1}^N \boldsymbol{Y}_\nu(t)$, where*

$$\boldsymbol{Y}_\nu(t) = e^{\lambda_\nu(t-s)}\boldsymbol{Y}_\nu(s) + \boldsymbol{b}_\nu\int_s^t e^{\lambda_\nu(t-u)}\mathrm{d}\langle\boldsymbol{s}_\nu, \boldsymbol{L}(u)\rangle, \quad s,t \in \mathbb{R}, \quad s < t. \tag{5.4}$$

Proof. We first choose a left matrix fraction description (P, Q) of the transfer function $z \mapsto C(z\mathbb{I}_N - A)^{-1}B$ such that $z \mapsto \det P(z)$ and $z \mapsto \det Q(z)$ have no common zeros and $z \mapsto \det P(z)$ has no multiple zeros. This is always possible by assumption **E2**. Inserting the spectral representation (5.3) of e^{At} in the kernel $g(t)$ (equation (3.9)) we write $g(t) = \frac{1}{2\pi i}\int_\Gamma e^{zt}C(z\mathbb{I}_N - A)^{-1}B\mathrm{d}zI_{[0,\infty)}(t)$ and, by construction, the integrand equals $e^{zt}P(z)^{-1}Q(z)I_{[0,\infty)}(t)$. After writing $P(z)^{-1} = \frac{1}{\det P(z)}\operatorname{adj} P(z)$, where adj denotes the adjugate of a matrix, an elementwise application of the residue theorem from complex analysis ([11, 9.16.1]) shows that

$$g(t) = \sum_{\nu=1}^N e^{\lambda_\nu t}\frac{1}{(\det P)'(\lambda_\nu)}\operatorname{adj} P(\lambda_\nu)Q(\lambda_\nu)I_{[0,\infty)}(t),$$

where $(\det P)'(\lambda_\nu) := \frac{d}{dz}\det P(z)|_{z=\lambda_\nu}$ is non-zero because $z \mapsto \det P(z)$ has only simple zeros. The same fact, in conjunction with the Smith decomposition of P ([3, Theorem 4.7.5.]), also implies that $\operatorname{rank} P(\lambda_\nu) = d-1$ and thus $\operatorname{rank}\operatorname{adj} P(\lambda_\nu) = 1$ ([3, Fact 2.14.7. *ii)*]). Since $\det P$ and $\det Q$ have no common zeros $[(\det P)'(\lambda_\nu)]^{-1}\operatorname{adj} P(\lambda_\nu)Q(\lambda_\nu)$ has rank one as well and can thus be written as $\boldsymbol{b}_\nu\boldsymbol{s}_\nu^T$ for some non-zero $\boldsymbol{s}_\nu \in \mathbb{C}^m$ and $\boldsymbol{b}_\nu \in \mathbb{C}^d$ ([13, §51, Theorem 1]). □

Lemma 5.2. *Assume $\boldsymbol{Y}$ is the output process of the state space model (3.5). Then the sampled process $\boldsymbol{Y}^{(h)}$ has the state space representation*

$$\boldsymbol{X}_n = e^{Ah}\boldsymbol{X}_{n-1} + \boldsymbol{N}_n, \quad \boldsymbol{N}_n = \int_{(n-1)h}^{nh} e^{A(nh-u)} B \mathrm{d}\boldsymbol{L}_u, \quad \boldsymbol{Y}_n^{(h)} = C\boldsymbol{X}_n^{(h)}. \tag{5.5}$$

The sequence $(\boldsymbol{N}_n)_{n\in\mathbb{Z}}$ is i.i.d. with mean zero and covariance matrix

$$\check{\Sigma} = \mathbb{E}\boldsymbol{N}_n\boldsymbol{N}_n^T = \int_0^h e^{Au} B\Sigma^{\boldsymbol{L}} B^T e^{A^T u} \mathrm{d}u. \tag{5.6}$$

Proof. Equations (5.5) follow from setting $t = nh$, $s = (n-1)h$ in equation (3.6). It is an immediate consequence of the Lévy process $\boldsymbol{L}$ having independent, homegeneous increments that the sequence $(\boldsymbol{Z}_n)_{n\in\mathbb{Z}}$ is i.i.d. and that its covariance matrix $\check{\Sigma}$ is given by equation (5.6). □

From this we can now proceed to prove the weak vector ARMA representation of the process $\boldsymbol{Y}^{(h)}$.

Proof of theorem 4.2. It follows from setting $t = nh$, $s = (n-1)h$ in (5.4) that $\boldsymbol{Y}_n^{(h)}$ can be decomposed as $\boldsymbol{Y}_n^{(h)} = \sum_{\nu=1}^N \boldsymbol{Y}_{\nu,n}^{(h)}$, where $\boldsymbol{Y}_\nu^{(h)}$, satisfying

$$\boldsymbol{Y}_{\nu,n}^{(h)} = e^{\lambda_\nu h}\boldsymbol{Y}_{\nu,n-1}^{(h)} + \boldsymbol{Z}_{\nu,n}^{(h)}, \quad \boldsymbol{Z}_{\nu,n}^{(h)} = \boldsymbol{b}_\nu \int_{(n-1)h}^{nh} e^{\lambda_\nu(nh-u)} \mathrm{d}\langle \boldsymbol{s}_\nu, \boldsymbol{L}(u)\rangle,$$

are the sampled versions of the component MCAR(1) processes from Proposition 5.1. Analogously to [9, Lemma 2.1.] one can show by induction that for each $k \in \mathbb{N}_0$ and every complex $d \times d$ matrices $c_1, \ldots, c_k$ it holds that

$$\boldsymbol{Y}_{\nu,n}^{(h)} = \sum_{r=1}^k c_r \boldsymbol{Y}_{\nu,n-r}^{(h)} + \left[e^{\lambda_\nu hk} - \sum_{r=1}^k c_r e^{\lambda_\nu h(k-r)}\right]\boldsymbol{Y}_{\nu,n-k}^{(h)} + \sum_{r=0}^{k-1}\left[e^{\lambda_\nu hr} - \sum_{j=1}^r c_j e^{\lambda_\nu h(r-j)}\right]\boldsymbol{Z}_{\nu,n-r}^{(h)}. \tag{5.7}$$

If we then use the fact that $e^{-h\lambda_\nu}$ is a root of $z \mapsto \varphi(z)$, which means $e^{Nh\lambda_\nu} - \varphi_1 e^{(N-1)h\lambda_n u} - \ldots - \varphi_N = 0$, and set $k = N$, $c_r = \mathbb{I}_d\varphi_r$, equation (5.7) becomes

$$\varphi(B)\boldsymbol{Y}_{\nu,n}^{(h)} = \sum_{r=0}^{N-1}\left[e^{rh\lambda_\nu} - \sum_{j=1}^r \varphi_j e^{\lambda_\nu h(r-j)}\right]\boldsymbol{Z}_{\nu,n-r}^{(h)}.$$

Summing over ν and rearranging shows that this can be written as

$$\varphi(B)\boldsymbol{Y}_n^{(h)} = \sum_{\nu=1}^N \boldsymbol{V}_{\nu,n-\nu+1}^{(h)}, \tag{5.8}$$

where the i.i.d sequences $\left(\boldsymbol{V}_{\nu,n}^{(h)}\right)_{n\in\mathbb{Z}}$, $\nu \in \{1, \ldots, N\}$, are defined by

$$\boldsymbol{V}_{\nu,n}^{(h)} = \int_{(n-1)h}^{nh} \sum_{\mu=1}^N \boldsymbol{b}_\mu \left[e^{\lambda_\mu h(\nu-1)} - \sum_{\kappa=1}^{\nu-1}\varphi_\kappa e^{\lambda_\mu h(\nu-\kappa-1)}\right] e^{\lambda_\mu(nh-u)}\mathrm{d}\langle \boldsymbol{s}_\mu, \boldsymbol{L}(u)\rangle. \tag{5.9}$$

By a straightforward generalization of [8, proposition 3.2.1] there exists a monic Schur-stable poynomial $\Theta(z) = \mathbb{I}_d + \Theta_1 z + \ldots + \Theta_{N-1}z^{N-1}$ and a white noise sequence $\tilde{\boldsymbol{\varepsilon}}$ such that the $(N-1)$-dependent sequence $\varphi(B)\boldsymbol{Y}^{(h)}$ has the moving average representation $\varphi(B)\boldsymbol{Y}_n^{(h)} = \Theta(B)\tilde{\boldsymbol{\varepsilon}}_n$. Since both φ and Θ are monic and φ is Schur stable (by assumption **E1**) $\tilde{\boldsymbol{\varepsilon}}$ is the innovation process of $\boldsymbol{Y}^{(h)}$ and so it follows that $\tilde{\boldsymbol{\varepsilon}} = \boldsymbol{\varepsilon}^{(h)}$ because the innovations of a stochastic process are uniquely determined. □

As a corollary we obtain that the innovations sequence $\boldsymbol{\varepsilon}^{(h)}$ itself satisfies a set of strong VARMA equations, the attribute *strong* referring to the fact that the noise sequence is i.i.d., not merely white noise.

Corollary 5.3. *Assume $\boldsymbol{Y}$ is the output process of the state space system (3.5) satisfying* **L1**, **E1** *and* **E2***; further that $\boldsymbol{\varepsilon}^{(h)}$ is the innovations sequence of the sampled process $\boldsymbol{Y}^{(h)}$. Then there exists a monic, Schur-stable polynomial $\Theta \in M_d(\mathbb{C}[z])$ of degree at most $N-1$, a polynomial $\theta \in M_{d,dN}(\mathbb{R}[z])$ of degree $N-1$ and a C^{dN}-valued i.i.d sequence $\boldsymbol{W}^{(h)} = (\boldsymbol{W}_n^{(h)})_{n\in\mathbb{Z}}$, such that*

$$\Theta(B)\boldsymbol{\varepsilon}_n^{(h)} = \theta(B)\boldsymbol{W}_n^{(h)}, \quad n \in \mathbb{Z}. \tag{5.10}$$

Proof. Combining equation (4.3) and equation (5.8) gives

$$\boldsymbol{\varepsilon}_n^{(h)} + \Theta_1^{(h)}\boldsymbol{\varepsilon}_{n-1} + \ldots + \Theta_{N-1}^{(h)}\boldsymbol{\varepsilon}_{n-N+1} = \boldsymbol{V}_{1,n}^{(h)} + \boldsymbol{V}_{2,n-1}^{(h)} + \ldots + \boldsymbol{V}_{N,n-N+1}^{(h)}, \quad n \in \mathbb{Z}, \tag{5.11}$$

and with the definitions

$$\boldsymbol{W}_n^{(h)} = \begin{bmatrix} \boldsymbol{V}_{1,n}^{(h)T} & \cdots & \boldsymbol{V}_{N,n}^{(h)\,T} \end{bmatrix}^T \in \mathbb{C}^{dN}, \quad n \in \mathbb{Z}, \tag{5.12a}$$

$$\theta(z) = \sum_{j=1}^{N} \theta_j z^{j-1}, \quad \theta_\nu = [\underbrace{0_d \ \cdots \ 0_d}_{\nu-1 \text{ times}} \ \mathbb{I}_d \ \underbrace{0_d \ \cdots \ 0_d}_{N-\nu \text{ times}}] \in M_{d,dN}(\mathbb{R}), \quad \nu = 1,\ldots,N, \tag{5.12b}$$

equation (5.11) becomes $\Theta(B)\boldsymbol{\varepsilon}_n^{(h)} = \theta(B)\boldsymbol{W}_n^{(h)}$, showing that $\boldsymbol{\varepsilon}^{(h)}$ is indeed a vector ARMA process. □

This corollary is the central step in establishing complete regularity of the innovations process $\boldsymbol{\varepsilon}^{(h)}$.

Proof of Theorem 4.3. We define the $\mathbb{R}^{mN}$-valued random variables

$$\mathscr{M}_n^{(h)} = \begin{bmatrix} \boldsymbol{M}_{n,1}^{(h)T} & \cdots & \boldsymbol{M}_{n,r}^{(h)T} & \underline{\boldsymbol{M}}_{n,r+1}^{(h)\ T} & \underline{\boldsymbol{M}}_{n,r+3}^{(h)\ T} & \cdots & \underline{\boldsymbol{M}}_{n,N-1}^{(h)\ T} \end{bmatrix}^T, \quad n \in \mathbb{Z},$$

where

$$\underline{\boldsymbol{M}}_{n,\nu}^{(h)} = \begin{bmatrix} \operatorname{Re} \boldsymbol{M}_{n,\nu}^{(h)T} & \operatorname{Im} \boldsymbol{M}_{n,\nu}^{(h)T} \end{bmatrix}^T, \quad \boldsymbol{M}_{n,\nu}^{(h)} = \int_{(n-1)h}^{nh} e^{\lambda_\nu(nh-u)} \mathrm{d}\boldsymbol{L}(u), \quad \nu = 1,\ldots,N, \quad n \in \mathbb{Z}.$$

Clearly, the sequence $(\mathscr{M}_n^{(h)})_{n\in\mathbb{Z}}$ is i.i.d. and $\mathscr{M}^{(h)}$ is equal to $\mathscr{M}_1^{(h)}$. We now argue that the vector $\boldsymbol{W}_n^{(h)}$, as defined in equation (5.12a), is equal to a linear transformation of $\mathscr{M}_n^{(h)}$. By equation (5.9), $\boldsymbol{W}_n^{(h)} = [\Gamma^T \otimes \mathbb{I}_d] \begin{bmatrix} (\boldsymbol{b}_1\boldsymbol{s}_1^T\boldsymbol{M}_{n,1}^{(h)})^T & \cdots & (\boldsymbol{b}_N\boldsymbol{s}_N^T\boldsymbol{M}_{n,N}^{(h)})^T \end{bmatrix}^T$ where $\Gamma = (\gamma_{\mu,\nu}) \in M_N(\mathbb{C})$ is given by $\gamma_{\mu,\nu} = e^{\lambda_\mu h(\nu-1)} + \sum_{\kappa=1}^{\nu-1}\varphi_\kappa e^{\lambda_\mu h(\nu-\kappa-1)}$. With the notation

$$B = \begin{pmatrix} \boldsymbol{b}_1 & \boldsymbol{0}_d & \ldots & \boldsymbol{0}_d \\ \boldsymbol{0}_d & \boldsymbol{b}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0}_d \\ \boldsymbol{0}_d & \ldots & \boldsymbol{0}_d & \boldsymbol{b}_N \end{pmatrix} \in M_{dN,N}(\mathbb{C}), \quad S = \begin{pmatrix} \boldsymbol{s}_1^T & \boldsymbol{0}_d^T & \ldots & \boldsymbol{0}_d^T \\ \boldsymbol{0}_d^T & \boldsymbol{s}_2^T & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0}_d^T \\ \boldsymbol{0}_d^T & \ldots & \boldsymbol{0}_d^T & \boldsymbol{s}_N^T \end{pmatrix} \in M_{N,mN}(\mathbb{C})$$

we get $\begin{bmatrix} \left(\boldsymbol{b}_1\boldsymbol{s}_1^T\boldsymbol{M}_{n,1}^{(h)}\right)^T & \cdots & \left(\boldsymbol{b}_N\boldsymbol{s}_N^T\boldsymbol{M}_{n,N}^{(h)}\right)^T \end{bmatrix}^T = BS\begin{bmatrix} \boldsymbol{M}_{n,1}^{(h)T} & \cdots & \boldsymbol{M}_{n,N}^{(h)\,T} \end{bmatrix}^T$. Keeping only those $\boldsymbol{M}_{n,\nu}^{(h)}$ that correspond to the real and the first of each pair of complex conjugate eigenvalues of A and separating the latter into real and imaginary parts we obtain $\begin{bmatrix} \boldsymbol{M}_{n,1}^{(h)T} & \cdots & \boldsymbol{M}_{n,N}^{(h)\,T} \end{bmatrix}^T = [K \otimes \mathbb{I}_m]\mathscr{M}_n^{(h)}$ where

$$K = \begin{pmatrix} \mathbb{I}_r & & & \\ & J & & \\ & & \ddots & \\ & & & J \end{pmatrix} \in M_N(\mathbb{C}), \quad J = \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix},$$

so that in total $\boldsymbol{W}_n^{(h)} = F\mathscr{M}_n^{(h)}$ with $F = [\Gamma^T \otimes \mathbb{I}_d]BS[K \otimes \mathbb{I}_m] \in M_{dN,mN}(\mathbb{C})$. It follows that the VARMA equation (5.10) for $\boldsymbol{\varepsilon}^{(h)}$ becomes $\Theta(B)\boldsymbol{\varepsilon}_n^{(h)} = \tilde{\theta}(B)\mathscr{M}_n^{(h)}$, where $\tilde{\theta}(z) = \theta(z)F$. By the invertibility of Θ, the transfer function $k : z \mapsto \Theta(z)^{-1}\tilde{\theta}(z)$ is analytic in a disk containing the unit disk and permits a power series expansion $k(z) = \sum_{j=0}^{\infty} \Psi_j z^j$. We next argue that the impulse responses Ψ_j are necessarily *real* $(d \times mN)$ matrices. Since both $\boldsymbol{\varepsilon}_n^{(h)}$ and $\mathscr{M}_n^{(h)}$ are real-valued it follows from taking the imaginary part of the equation $\boldsymbol{\varepsilon}_n^{(h)} = k(B)\mathscr{M}_n^{(h)}$ that $\mathbf{0}_d = \sum_{j=0}^{\infty} \operatorname{Im}\Psi_j \mathscr{M}_{n-j}^{(h)}$. Consequently, $0 = \operatorname{Cov}(\mathbf{0}_d) = \sum_{j=0}^{\infty} \operatorname{Im}\Psi_j \operatorname{Cov}(\mathscr{M}_{n-j}^{(h)}) \operatorname{Im}\Psi_j^T$ and since each term in the sum is a positive semi-definite matrix it follows that $\operatorname{Im}\Psi_j \operatorname{Cov}(\mathscr{M}_{n-j}^{(h)}) \operatorname{Im}\Psi_j^T = 0$ for every j. The existence of an absolutely continuous component of the law of $\mathscr{M}_{n-j}^{(h)}$ with respect to the mN-dimensional Lebesgue measure implies that $\operatorname{Cov}(\mathscr{M}_{n-j}^{(h)})$ is non-singular and it thus follows that $\operatorname{Im}\Psi_j = 0$ for every j. Hence $k(z) \in M_{d,mN}(\mathbb{R})$ for all real z and consequently $k \in M_{d,mN}(\mathbb{R}\{z\})$. [14, Theorem 1.2.1, *(iii)*] then implies that there exists a stable $(\mathscr{M}_n^{(h)})_{n\in\mathbb{N}}$-driven VARMA model for $\boldsymbol{\varepsilon}^{(h)}$ with real-valued coefficient matrices. It has been shown in [19, Theorem 1] that a stable vector ARMA process is geometrically completely regular provided that the driving noise sequence is i.i.d. and absolutely continuous with respect to the Lebesgue measure. A careful analysis of the proof of this result shows that the existence of an absolutely continuous component of the law of the driving noise is already sufficient for the conclusion to hold. We briefly comment on the necessary modifications in the argument. One first notes that under these weaker assumptions the proof of [19, Lemma 3] implies that the n-step transition probabilities $P^n(\boldsymbol{x}, \cdot)$ of the Markov chain X associated to a vector ARMA model via its state space representation have an absolutely continuous component for all n greater than or equal to some n_0. This immediately implies aperiodicity and ϕ-irreducibility of X, where ϕ can be taken as the Lebesgue measure restricted to the support of the continuous component of $P^{n_0}(\boldsymbol{x}, \cdot)$. The rest of the proof, in particular the verification of the Foster-Lyapunov drift condition for complete regularity, is unaltered. This shows that $\boldsymbol{\varepsilon}^{(h)}$ is geometrically completely regular and in particular strongly mixing with exponentially decyaing mixing coefficients. □

Proof of Lemma 4.4. By definition, $\boldsymbol{M}_\nu^{(h)} = \boldsymbol{M}_\nu(h)$, where $(\boldsymbol{M}_\nu(t))_{t\geq 0}$ is the solution to

$$\mathrm{d}\boldsymbol{M}_\nu(t) = \lambda_\nu \boldsymbol{M}_\nu(t)\mathrm{d}t + \mathrm{d}\boldsymbol{L}(t), \quad \boldsymbol{M}_\nu(0) = \mathbf{0}_m.$$

Taking the real and imaginary part of this equation gives

$$\begin{aligned}
\mathrm{d}\operatorname{Re}\boldsymbol{M}_\nu(t) &= \operatorname{Re}\lambda_\nu \boldsymbol{M}_\nu(t)\mathrm{d}t + \mathrm{d}\boldsymbol{L}(t) = [\operatorname{Re}\lambda_\nu \operatorname{Re}\boldsymbol{M}_\nu(t) - \operatorname{Im}\lambda_\nu \operatorname{Im}\boldsymbol{M}_\nu(t)]\,\mathrm{d}t + \mathrm{d}\boldsymbol{L}(t),\\
\mathrm{d}\operatorname{Im}\boldsymbol{M}_\nu(t) &= \operatorname{Im}\lambda_\nu \boldsymbol{M}_\nu(t)\mathrm{d}t \qquad\qquad\;\; = [\operatorname{Re}\lambda_\nu \operatorname{Im}\boldsymbol{M}_\nu(t) + \operatorname{Im}\lambda_\nu \operatorname{Re}\boldsymbol{M}_\nu(t)]\,\mathrm{d}t
\end{aligned}$$

and consequently

$$\mathrm{d}\begin{pmatrix}\operatorname{Re}\boldsymbol{M}_\nu(t)\\ \operatorname{Im}\boldsymbol{M}_\nu(t)\end{pmatrix} = [\Lambda_\nu \otimes \mathbb{I}_m]\begin{pmatrix}\operatorname{Re}\boldsymbol{M}_\nu(t)\\ \operatorname{Im}\boldsymbol{M}_\nu(t)\end{pmatrix}\mathrm{d}t + \begin{pmatrix}\mathbb{I}_m\\ 0_m\end{pmatrix}\mathrm{d}\boldsymbol{L}(t), \quad \Lambda_\nu = \begin{pmatrix}\operatorname{Re}\lambda_\nu & -\operatorname{Im}\lambda_\nu\\ \operatorname{Im}\lambda_\nu & \operatorname{Re}\lambda_\nu\end{pmatrix}.$$

Using that $\lambda_\nu \in \mathbb{R}$ for $\nu = 1, \ldots, r$ and $\lambda_\nu = \overline{\lambda_{\nu+1}} \in \mathbb{C}\backslash\mathbb{R}$ for $\nu = r+1, r+3, \ldots, N-1$ it follows that $\mathscr{M}^{(h)} = \mathscr{M}(h)$ where $(\mathscr{M}(t))_{t\geq 0}$ satisfies $\mathrm{d}\mathscr{M}(t) = G\mathscr{M}(t)\mathrm{d}t + H\mathrm{d}\boldsymbol{L}(t)$ and $G = \tilde{G} \otimes \mathbb{I}_m \in M_{mN}(\mathbb{R})$ and $H = \tilde{H} \otimes \mathbb{I}_m \in M_{mN,m}$ are given by

$$\begin{aligned}
\tilde{G} &= \operatorname{diag}(\lambda_1, \ldots, \lambda_r, \Lambda_{r+1}, \Lambda_{r+3}, \ldots, \Lambda_{N-1}),\\
\tilde{H} &= (\underbrace{1 \;\; \cdots \;\; 1}_{r \text{ times}} \;\; 1 \;\; 0 \;\; 1 \;\; 0 \;\; \cdots \;\; 1 \;\; 0\,)^T.
\end{aligned}$$

Since $\operatorname{rank} H = m$ the first claim of the lemma is proved. Next we show that the controllability matrix $\mathscr{C} := \begin{bmatrix} H & GH & \cdots & G^{N-1}H \end{bmatrix} \in M_{mN}(\mathbb{R})$ is non-singular. With $\widetilde{\mathscr{C}} := \begin{bmatrix} \tilde{H} & \tilde{G}\tilde{H} & \cdots & \tilde{G}^{N-1}\tilde{H} \end{bmatrix}$ and by the properties of the Kronecker product it follows that $\mathscr{C} = \widetilde{\mathscr{C}} \otimes \mathbb{I}_m$ and thus $\det\mathscr{C} = [\det\widetilde{\mathscr{C}}]^m$. The matrix $\widetilde{\mathscr{C}}$ is given explicitly by

$$\widetilde{\mathscr{C}} = \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{N-1} \\ \vdots & & & & \vdots \\ 1 & \lambda_r & \lambda_r^2 & \cdots & \lambda_r^{N-1} \\ 1 & \operatorname{Re}\lambda_{r+1} & \operatorname{Re}\lambda_{r+1}^2 & \cdots & \operatorname{Re}\lambda_{r+1}^{N-1} \\ 0 & \operatorname{Im}\lambda_{r+1} & \operatorname{Im}\lambda_{r+1}^2 & \cdots & \operatorname{Im}\lambda_{r+1}^{N-1} \\ \vdots & & & & \vdots \\ 1 & \operatorname{Re}\lambda_{N-1} & \operatorname{Re}\lambda_{N-1}^2 & \cdots & \operatorname{Re}\lambda_{N-1}^{N-1} \\ 0 & \operatorname{Im}\lambda_{N-1} & \operatorname{Im}\lambda_{N-1}^2 & \cdots & \operatorname{Im}\lambda_{N-1}^{N-1} \end{pmatrix} = T \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{N-1} \\ \vdots & & & & \vdots \\ 1 & \lambda_r & \lambda_r^2 & \cdots & \lambda_r^{N-1} \\ 1 & \lambda_{r+1} & \lambda_{r+1}^2 & \cdots & \lambda_{r+1}^{N-1} \\ i & i\overline{\lambda_{r+1}} & i\overline{\lambda_{r+1}^2} & \cdots & i\overline{\lambda_{r+1}^{N-1}} \\ \vdots & & & & \vdots \\ 1 & \lambda_{N-1} & \lambda_{N-1}^2 & \cdots & \lambda_{N-1}^{N-1} \\ i & i\overline{\lambda_{N-1}} & i\overline{\lambda_{N-1}^2} & \cdots & i\overline{\lambda_{N-1}^{N-1}} \end{pmatrix}$$

with $T \in M_N(\mathbb{R})$ given by $T = \operatorname{diag}(1, \dots, 1, R, \dots, R)$, $R = \frac{1}{2}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}$. Hence the formula for the determinant of a Vandermonde matrix ([3, Fact 5.13.3.]) implies

$$\det \mathscr{C} = \left[(-1)^{\frac{N-r}{2}} \prod_{1\leq\mu<\nu\leq r} (\lambda_\mu - \lambda_\nu) \prod_{\substack{\mu,\nu\in I_{r,N} \\ \mu<\nu}} \operatorname{Im}\lambda_\mu |\lambda_\mu - \lambda_\nu|^2 |\overline{\lambda_\mu} - \lambda_\nu|^2 \prod_{\substack{1\leq\mu\leq r \\ \nu\in I_{r,N}}} |\lambda_\mu - \lambda_\nu|^2 \right]^m,$$

where $I_{r,N} = \{r+1, r+3, \dots, N-1\}$. Hence, $\det \mathscr{C}$ is not zero by assumption **E2** and the proof is complete. □

Acknowledgements

Financial support from the International Graduate School of Science and Engineering of the Technische Universität München and the TUM Institute for Advanced Study funded by the German Excellence Initiative is appreciated.

References

[1] Athreya, K. B. and Pantula, S. G. (1986). Mixing properties of Harris chains and autoregressive processes. *J. Appl. Probab.* **23** 880–892. MR867185

[2] Barndorff-Nielsen, O. E., Maejima, M. and Sato, K.-I. (2006). Some classes of multivariate infinitely divisible distributions admitting stochastic integral representations. *Bernoulli* **12** 1–33. MR2202318

[3] Bernstein, D. S. (2005). *Matrix mathematics*, First ed. Princeton University Press, Princeton, NJ. MR2123424

[4] Bradley, R. C. (2007). *Introduction to strong mixing conditions. Vol. 1*. Kendrick Press, Heber City, UT. MR2325294

[5] Brockwell, P.J., (2001). Lévy-Driven CARMA Processes. *Ann. Inst. Stat. Math.* **53** 113–124.

[6] Brockwell, P.J. and Davis, R.A. and Yang, Y., (2010). Estimation for nonnegative Levy-driven CARMA processes. To appear in *J. Bus. Econ. Stat.* .

[7] Brockwell, P. J. (2001). Continuous-time ARMA processes. In *Stochastic processes: theory and methods. Handbook of Statist.* **19** 249–276. North-Holland, Amsterdam. MR1861726

[8] Brockwell, P. J. and Davis, R. A. (1991). *Time series: theory and methods*, Second ed. *Springer Series in Statistics*. Springer-Verlag, New York. MR1093459

[9] Brockwell, P. J. and Lindner, A. (2009). Existence and uniqueness of stationary Lévy-driven CARMA processes. *Stochastic Process. Appl.* **119** 2660–2681. MR2532218

[10] Caines, P. E. (1988). *Linear stochastic systems*. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. John Wiley & Sons Inc., New York. MR944080

[11] Dieudonné, J. (1968). *Éléments d'analyse. Tome I: Fondements de l'analyse moderne*. Gauthier-Villars, Éditeur, Paris. MR0235945

[12] Doob, J. L. (1944). The elementary Gaussian processes. *Ann. Math. Statistics* **15** 229–282. MR0010931

[13] Halmos, P. R. (1974). *Finite-dimensional vector spaces*, second ed. *Undergraduate Texts in Mathematics*. Springer-Verlag, New York. MR0409503

[14] Hannan, E. J. and Deistler, M. (1988). *The statistical theory of linear systems*. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. John Wiley & Sons Inc., New York. MR940698

[15] Klenke, A. (2008). *Probability theory*. *Universitext*. Springer-Verlag London Ltd., London. MR2372119 (2009f:60001)

[16] Lax, P. D. (2002). *Functional analysis*. *Pure and Applied Mathematics (New York)*. Wiley-Interscience [John Wiley & Sons], New York. MR1892228

[17] Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer-Verlag, Berlin. MR2172368

[18] Marquardt, T. and Stelzer, R. (2007). Multivariate CARMA processes. *Stochastic Process. Appl.* **117** 96–120. MR2287105

[19] Mokkadem, A. (1988). Mixing properties of ARMA processes. *Stochastic Process. Appl.* **29** 309–315. MR958507

[20] Pham, T. D. and Tran, L. T. (1985). Some mixing properties of time series models. *Stochastic Process. Appl.* **19** 297–303. MR787587

[21] Priola, E. and Zabczyk, J. (2009). Densities for Ornstein-Uhlenbeck processes with jumps. *Bull. Lond. Math. Soc.* **41** 41–50. MR2481987

[22] Sato, K.-i. (1999). *Lévy processes and infinitely divisible distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge University Press, Cambridge. MR1739520

[23] Sato, K.-I. (2006). Additive processes and stochastic integrals. *Illinois J. Math.* **50** 825–851 (electronic). MR2247848

[24] Sato, K.-i. and Yamazato, M. (1984). Operator-self-decomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stochastic Process. Appl.* **17** 73–100. MR738769

[25] Simon, T., (2010). On the absolute continuity of multidimensional Ornstein-Uhlenbeck processes. To appear in *Probab. Theory Relat. Field* .