

LIMIT THEORY FOR HIGH FREQUENCY SAMPLED MCARMA MODELS

VICKY FASEN,* *ETH Zürich*

Abstract

We consider a multivariate continuous-time ARMA (MCARMA) process sampled at a high-frequency time grid $\{h_n, 2h_n, \dots, nh_n\}$, where $h_n \downarrow 0$ and $nh_n \rightarrow \infty$ as $n \rightarrow \infty$, or at a constant time grid where $h_n = h$. For this model, we present the asymptotic behavior of the properly normalized partial sum to a multivariate stable or a multivariate normal random vector depending on the domain of attraction of the driving Lévy process. Furthermore, we derive the asymptotic behavior of the sample variance. In the case of finite second moments of the driving Lévy process the sample variance is a consistent estimator. Moreover, we embed the MCARMA process in a cointegrated model. For this model, we propose a parameter estimator and derive its asymptotic behavior. The results are given for more general processes than MCARMA processes and contain some asymptotic properties of stochastic integrals.

Keywords: Central limit theorem; cointegration; domain of attraction; high-frequency data; multivariate CARMA process; regular variation; Ornstein–Uhlenbeck process; sample variance

2010 Mathematics Subject Classification: Primary 62M10; 60F05
Secondary 91B84

1. Introduction

Multivariate continuous-time ARMA (MCARMA) processes $V = (V(t))_{t \geq 0}$ are the continuous-time versions of the well-known multivariate ARMA processes in discrete time having short memory. They are important for stochastic modeling in many areas of applications, such as, e.g. signal processing and control (cf. [17] and [24]), econometrics (cf. [2]), high-frequency financial econometrics (cf. [41]), and financial mathematics (cf. [1]). Starting at least with Doob [10] in 1944, Gaussian CARMA processes, under the name Gaussian processes with rational spectral density, appeared, where the driving force is a Brownian motion. To obtain more flexible marginal distributions and dynamics, Brockwell [4] analyzed Lévy-driven CARMA models, which were extended by Marquardt and Stelzer [25] to the multivariate setting; see [5] for an overview and a comprehensive list of references.

Lévy processes are defined to have independent and stationary increments, and are characterized by their Lévy–Khintchine representation. An \mathbb{R}^m -valued Lévy process $(L(t))_{t \geq 0}$ has the Lévy–Khintchine representation $\mathbb{E}(e^{i\Theta^\top L(t)}) = \exp(-t\Psi(\Theta))$ for $\Theta \in \mathbb{R}^m$, where Θ^\top is the transpose of Θ and

$$\Psi(\Theta) = -i\gamma_L^\top \Theta + \frac{1}{2}\Theta^\top \Sigma_L \Theta + \int_{\mathbb{R}^m} (1 - e^{ix^\top \Theta} + ix^\top \Theta \mathbf{1}_{\{\|x\|^2 \leq 1\}}) \nu_L(dx)$$

Received 31 August 2012; revision received 3 October 2013.

* Postal address: Institute of Stochastics, Karlsruhe Institute of Technology, Kaiserstrasse 89, 76133 Karlsruhe, Germany. Email address: vicky.fasen@kit.edu

with $\gamma_L \in \mathbb{R}^m$, Σ_L a positive semidefinite matrix in $\mathbb{R}^{m \times m}$, and ν_L a measure on $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$, called the *Lévy measure*, which satisfies $\int_{\mathbb{R}^m} \min\{\|x\|^2, 1\} \nu_L(dx) < \infty$ and $\nu_L(\{0_m\}) = 0$. The triplet $(\gamma_L, \Sigma_L, \nu_L)$ is called the *characteristic triplet*, because it characterizes completely the distribution of the Lévy process. A two-sided Lévy process $(L(t))_{t \in \mathbb{R}}$ is then a composition of two independent and identically distributed Lévy processes, $(L^{(1)}(t))_{t \geq 0}$ and $(L^{(2)}(t))_{t \geq 0}$, in $L(t) = L^{(1)}(t)\mathbf{1}_{[0, \infty)}(t) + L^{(2)}(-t)\mathbf{1}_{(-\infty, 0)}(t)$. We refer the reader to the excellent monograph [38] for more details on Lévy processes. In this paper the driving Lévy process is very general. It is allowed to have either a finite variance, $\mathbb{E}\|L(1)\|^2 < \infty$, or an infinite variance, $\mathbb{E}\|L(1)\|^2 = \infty$, which is modeled by a multivariate regularly varying Lévy process. CARMA processes driven by infinite variance Lévy processes are particularly relevant in modeling energy markets; see [16] for instance. We will investigate MCARMA processes (see Definition 2.1) observed not only at a constant frequency h but also, and especially, for high frequencies, as found in finance (cf. [41]) and turbulence (cf. [6]). Then the observation grid is $\{h_n, 2h_n, \dots, nh_n\}$, where $h_n \downarrow 0$ and $\lim_{n \rightarrow \infty} nh_n = \infty$. For the statistical inference of a MCARMA process, e.g. parameter estimation and hypothesis testing, it is crucial to know the asymptotic behavior of the partial sum (cf. [14] and [15]). We will show the convergence of the properly normalized partial sum to an α -stable random vector, $\alpha \in (0, 2]$, where $\alpha = 2$ reflects the multivariate normal distribution. In the high-frequency setting the limit distribution includes a random factor independent of the MCARMA parameters and a deterministic factor, which is determined by the model parameters (the integral over the kernel function). This is the same pattern as for multivariate ARMA models. However, the normalization differs in the continuous-time and discrete-time cases. The grid distance h_n has an influence on the convergence rate and, hence, determines the normalization in the continuous-time model. Furthermore, we study the asymptotic behavior of the sample variance. The results show that in the finite second moment case the sample variance is a consistent estimator for the variance. In the infinite second moment case it converges to an $\alpha/2$ -stable random matrix. Again, the convergence rate depends on the sampling distance h_n .

Another aim of this paper is to provide an estimation of a cointegrated model in continuous time, where the MCARMA process is embedded. Cointegration plays an important role in financial econometrics, see, e.g. [11], and is well understood in discrete time if second moments exist (cf. the monograph [21]). Most of the literature on cointegrated models in continuous time is restricted to Gaussian processes; see, e.g. [7], [23], and [40]. First approaches to drop the Gaussian assumption go back to Phillips [30]; see also [12], [13] and the references therein. Let $L_1 = (L_1(t))_{t \in \mathbb{R}}$ be the \mathbb{R}^m -valued driving Lévy process of the \mathbb{R}^d -valued MCARMA process V , and let $L_2 = (L_2(t))_{t \in \mathbb{R}}$ be an \mathbb{R}^v -valued Lévy process independent of L_1 . Then, for $A \in \mathbb{R}^{d \times v}$, we investigate the *multivariate cointegrated model*

$$\begin{aligned} Y(t) &= AX(t) + V(t), & t \geq 0, \text{ in } \mathbb{R}^d, \\ X(t) &= L_2(t), & t \geq 0, \text{ in } \mathbb{R}^v. \end{aligned} \tag{1.1}$$

The observation scheme is

$$Y_n^\top = (Y(h_n), \dots, Y(nh_n)) \in \mathbb{R}^{d \times n}, \quad X_n^\top = (X(h_n), \dots, X(nh_n)) \in \mathbb{R}^{v \times n}. \tag{1.2}$$

However, in this paper we investigate a more general model. Let $(\xi_{n,k})_{k \in \mathbb{Z}}$ and $(\varepsilon_{n,k})_{k \in \mathbb{Z}}$ be independent sequences of independently and identically distributed (i.i.d.) random vectors in \mathbb{R}^m and \mathbb{R}^v , respectively, for any $n \in \mathbb{N}$, and let $(C_{n,k})_{k \in \mathbb{N}}$ be a sequence of deterministic matrices in $\mathbb{R}^{d \times m}$ satisfying some general constraints. Then we may define, for any $n \in \mathbb{N}$, the

\mathbb{R}^d -valued stationary moving average process

$$\mathbf{Z}_{n,k} = \sum_{j=0}^{\infty} \mathbf{C}_{n,j} \boldsymbol{\xi}_{n,k-j} \quad \text{for } k \in \mathbb{N}_0, \tag{1.3}$$

and the cointegrated model as

$$\begin{aligned} \mathbf{Y}_{n,k} &= \mathbf{A}\mathbf{X}_{n,k} + \mathbf{Z}_{n,k} && \text{for } n, k \in \mathbb{N}, \text{ in } \mathbb{R}^d, \\ \mathbf{X}_{n,k} &= \mathbf{X}_{n,k-1} + \boldsymbol{\varepsilon}_{n,k} && \text{for } n, k \in \mathbb{N}, \text{ in } \mathbb{R}^v. \end{aligned} \tag{1.4}$$

In this case the observation scheme is

$$\mathbb{Y}_n^\top = (\mathbf{Y}_{n,1}, \dots, \mathbf{Y}_{n,n}) \in \mathbb{R}^{d \times n}, \quad \mathbb{X}_n^\top = (\mathbf{X}_{n,1}, \dots, \mathbf{X}_{n,n}) \in \mathbb{R}^{v \times n}. \tag{1.5}$$

Since the high-frequency sampled MCARMA process $(\mathbf{V}(kh_n))_{k \in \mathbb{Z}}$ has a representation as in (1.3) and

$$\mathbf{L}_2(kh_n) = \mathbf{L}_2((k - 1)h_n) + [\mathbf{L}_2(kh_n) - \mathbf{L}_2((k - 1)h_n)],$$

where $(\mathbf{L}_2(kh_n) - \mathbf{L}_2((k - 1)h_n))_{k \in \mathbb{N}}$ is an i.i.d. sequence by the independent and stationary increment property of a Lévy process, (1.2) can be interpreted as a special case of (1.5). As an estimator for \mathbf{A} , we use the least-squares estimator

$$\hat{\mathbf{A}}_n = \mathbb{Y}_n^\top \mathbb{X}_n (\mathbb{X}_n^\top \mathbb{X}_n)^{-1}. \tag{1.6}$$

The paper is organized as follows. In Section 2 we present some preliminaries on MCARMA processes, regular variation, and the model assumptions. The main results of this paper on limit theory for high-frequency sampled MCARMA processes and equidistant sampled MCARMA processes are the topic of Section 3. We show that the properly normalized partial sum $\sum_{k=1}^n \mathbf{V}(kh_n)$ and the sample variance $\sum_{k=1}^n \mathbf{V}(kh_n) \mathbf{V}(kh_n)^\top$ of the MCARMA process with either $h_n \downarrow 0$ and $nh_n \rightarrow \infty$ as $n \rightarrow \infty$ or $h_n = h$ (but with different normalization) converge weakly, and we completely characterize their limit distributions. Moreover, we investigate the cointegrated model (1.1)–(1.2). All results are compared to multivariate ARMA models in discrete time. The proofs of this section are based on some general limit theorems as constituted in Section 4. There we present, under some general assumptions, the asymptotic behavior of $\hat{\mathbf{A}}_n$ for the multivariate cointegrated model (1.4)–(1.5). Finally, in Section 5 we present the proofs of the stated results, and in Appendix A we examine the asymptotic behavior of stochastic integrals where the driving Lévy process has either a finite second moment or is multivariate regularly varying. These results are interesting in their own right, but they also act as preliminaries to the results in this paper.

We use the notation ‘ \Rightarrow ’ for weak convergence, ‘ \xrightarrow{P} ’ for convergence in probability, and ‘ \xrightarrow{v} ’ for vague convergence. Let $\mathbb{R} = \mathbb{R} \cup \{-\infty, \infty\}$ be the compactification of \mathbb{R} , and let $\mathcal{B}(\cdot)$ be the Borel σ -algebra. For two random vectors \mathbf{X} and \mathbf{Y} , the notation $\mathbf{X} \stackrel{D}{=} \mathbf{Y}$ means equality in distribution. We use as norms the Euclidean norm $\|\cdot\|$ in \mathbb{R}^d and the corresponding operator norm $\|\cdot\|$ for matrices, which is submultiplicative. Recall that two norms on a finite-dimensional linear space are always equivalent and, hence, our results remain true if we replace the Euclidean norm by any other norm. For a measurable function $f: (0, \infty) \rightarrow (0, \infty)$ and $\alpha \in \mathbb{R}$, we say that f is regularly varying of index $-\alpha$ if $\lim_{x \rightarrow \infty} f(xu)/f(x) = u^{-\alpha}$ for any $u > 0$, and we write $f \in \mathcal{R}_{-\alpha}$. The set of $d \times m$ matrices over \mathbb{R} is denoted by $M_{d \times m}(\mathbb{R})$. The matrix $\mathbf{0}_{d \times m}$ is the zero matrix in $M_{d \times m}(\mathbb{R})$ and $\mathbf{I}_{d \times d}$ is the identity matrix in $M_{d \times d}(\mathbb{R})$. For a vector $\mathbf{x} \in \mathbb{R}^d$,

we write \mathbf{x}^\top for its transpose, and, for $x \in \mathbb{R}$, we write $\lfloor x \rfloor = \sup\{k \in \mathbb{Z} : k \leq x\}$. The space $(\mathbb{D}[0, 1], \mathbb{R}^d)$ denotes the space of all càdlàg (right continuous with left limits) functions on $[0, 1]$ with values in \mathbb{R}^d equipped with the Skorokhod J_1 topology. Finally, for a semimartingale $\mathbf{W} = (\mathbf{W}_1(t), \dots, \mathbf{W}_d(t))_{t \geq 0}$ in \mathbb{R}^d , we denote by $[\mathbf{W}, \mathbf{W}]_t = ([\mathbf{W}_i, \mathbf{W}_j]_t)_{i,j=1,\dots,d}$ for $t \geq 0$ the quadratic variation process.

2. Preliminaries

2.1. MCARMA process

Let $\mathbf{L}_1 = (\mathbf{L}_1(t))_{t \in \mathbb{R}}$ be a two-sided \mathbb{R}^m -valued Lévy process, and let $p > q$ be positive integers. Then the d -dimensional MCARMA(p, q) model can be interpreted as the solution to the stochastic differential equation

$$\mathbf{P}(D)\mathbf{V}(t) = \mathbf{Q}(D)D\mathbf{L}_1(t) \quad \text{for } t \in \mathbb{R},$$

where D is the differential operator,

$$\mathbf{P}(z) := \mathbf{I}_{d \times d} z^p + \mathbf{P}_1 z^{p-1} + \dots + \mathbf{P}_{p-1} z + \mathbf{P}_p \tag{2.1}$$

with $\mathbf{P}_1, \dots, \mathbf{P}_p \in M_{d \times d}(\mathbb{R})$ the autoregressive polynomial, and

$$\mathbf{Q}(z) := \mathbf{Q}_0 z^q + \mathbf{Q}_1 z^{q-1} + \dots + \mathbf{Q}_{q-1} z + \mathbf{Q}_q \tag{2.2}$$

with $\mathbf{Q}_0, \dots, \mathbf{Q}_q \in M_{d \times m}(\mathbb{R})$ the moving average polynomial. Since a Lévy process is not differentiable, this definition cannot be used; however, it can be interpreted to be equivalent to the following definition.

Definition 2.1. Let $(\mathbf{L}_1(t))_{t \in \mathbb{R}}$ be an \mathbb{R}^m -valued Lévy process, and let the polynomials $\mathbf{P}(z)$ and $\mathbf{Q}(z)$ be defined as in (2.1) and (2.2) with $p, q \in \mathbb{N}_0, q < p$, and $\mathbf{Q}_0 \neq \mathbf{0}_{d \times m}$. Moreover, define

$$\mathbf{\Lambda} = - \begin{pmatrix} \mathbf{0}_{d \times d} & \mathbf{I}_{d \times d} & \mathbf{0}_{d \times d} & \cdots & \mathbf{0}_{d \times d} \\ \mathbf{0}_{d \times d} & \mathbf{0}_{d \times d} & \mathbf{I}_{d \times d} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathbf{0}_{d \times d} \\ \mathbf{0}_{d \times d} & \cdots & \cdots & \mathbf{0}_{d \times d} & \mathbf{I}_{d \times d} \\ -\mathbf{P}_p & -\mathbf{P}_{p-1} & \cdots & \cdots & -\mathbf{P}_1 \end{pmatrix} \in M_{pd \times pd}(\mathbb{R}),$$

$\mathbf{E} = (\mathbf{I}_{d \times d}, \mathbf{0}_{d \times d}, \dots, \mathbf{0}_{d \times d}) \in M_{d \times pd}(\mathbb{R})$, and $\mathbf{B} = (\mathbf{B}_1^\top \dots \mathbf{B}_p^\top)^\top \in M_{pd \times m}(\mathbb{R})$ with

$$\begin{aligned} \mathbf{B}_1 &:= \cdots := \mathbf{B}_{p-q-1} := \mathbf{0}_{d \times m} \\ \text{and } \mathbf{B}_{p-j} &:= - \sum_{i=1}^{p-j-1} \mathbf{P}_i \mathbf{B}_{p-j-i} + \mathbf{Q}_{q-j}, \quad j = 0, \dots, q. \end{aligned}$$

Assume that $\mathcal{N}(\mathbf{P}) = \{z \in \mathbb{C} : \det(\mathbf{P}(z)) = 0\} \subseteq (-\infty, 0) + i\mathbb{R}$. Furthermore, the Lévy measure ν_{L_1} of \mathbf{L}_1 satisfies $\int_{\|\mathbf{x}\| > 1} \log \|\mathbf{x}\| \nu_{L_1}(d\mathbf{x}) < \infty$. Then the \mathbb{R}^d -valued causal MCARMA(p, q) process $(\mathbf{V}(t))_{t \in \mathbb{R}}$ is defined by the state space equation

$$\mathbf{V}(t) = \mathbf{E}\mathbf{Z}(t) \quad \text{for } t \in \mathbb{R}, \tag{2.3}$$

where

$$\mathbf{Z}(t) = \int_{-\infty}^t e^{-\Lambda(t-s)} \mathbf{B} \, d\mathbf{L}_1(s) \quad \text{for } t \in \mathbb{R} \tag{2.4}$$

is the stationary unique solution to the pd -dimensional stochastic differential equation $d\mathbf{Z}(t) = -\Lambda\mathbf{Z}(t) \, dt + \mathbf{B} \, d\mathbf{L}(t)$. The function $\mathbf{f}(t) = \mathbf{E}e^{-\Lambda t} \mathbf{B}\mathbf{1}_{(0,\infty)}(t)$ for $t \in \mathbb{R}$ is called the kernel function.

In particular, the MCARMA(1, 0) process and \mathbf{Z} in (2.4) are multivariate Ornstein–Uhlenbeck processes. To see that the MCARMA(p, q) process is well defined, we refer the reader to [25]. Moreover, [25, Lemma 3.8] says that the set $\mathcal{N}(\mathbf{P})$ is equal to the set of eigenvalues of $-\Lambda$, which means that, for an MCARMA(p, q) process, the eigenvalues of Λ have strictly positive real parts. The class of MCARMA processes is huge. Schlemm and Stelzer [39, Corollary 3.4] showed that the class of state space models of the form

$$d\tilde{\mathbf{Z}}(t) = -\tilde{\Lambda}\tilde{\mathbf{Z}}(t) \, dt + \tilde{\mathbf{B}} \, d\mathbf{L}(t) \quad \text{and} \quad \tilde{\mathbf{V}}(t) = \tilde{\mathbf{C}}\tilde{\mathbf{Z}}(t),$$

where $\tilde{\Lambda} \in \mathbb{R}^{N \times N}$ has only eigenvalues with strictly positive real parts, $\tilde{\mathbf{B}} \in \mathbb{R}^{N \times m}$, and $\tilde{\mathbf{C}} \in \mathbb{R}^{d \times N}$, and the class of causal MCARMA processes are equivalent if $\mathbb{E}\|\mathbf{L}(1)\|^2 < \infty$ and $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_m$.

2.2. Multivariate regular variation and assumptions

Multivariate regular variation plays a basic part in our model assumptions. First, we recall the definition.

Definition 2.2. A random vector $\mathbf{U} \in \mathbb{R}^d$ is multivariate regularly varying with index $-\alpha < 0$ if and only if there exists a nonzero Radon measure μ on $(\overline{\mathbb{R}^d} \setminus \{\mathbf{0}_d\}, \mathcal{B}(\overline{\mathbb{R}^d} \setminus \{\mathbf{0}_d\}))$ with $\mu(\overline{\mathbb{R}^d} \setminus \mathbb{R}^d) = 0$ and a sequence $(a_n)_{n \in \mathbb{N}}$ of positive numbers increasing to ∞ such that

$$n\mathbb{P}(a_n^{-1}\mathbf{U} \in \cdot) \xrightarrow{v} \mu(\cdot) \quad \text{as } n \rightarrow \infty \text{ on } \mathcal{B}(\overline{\mathbb{R}^d} \setminus \{\mathbf{0}_d\}),$$

where the limit measure μ is homogeneous of order $-\alpha$, i.e. $\mu(uB) = u^{-\alpha}\mu(B)$ for $u > 0$ and $B \in \mathcal{B}(\overline{\mathbb{R}^d} \setminus \{\mathbf{0}_d\})$. We write $\mathbf{U} \in \mathcal{R}_{-\alpha}(a_n, \mu)$.

If the representation of the limit measure μ or the norming sequence $(a_n)_{n \in \mathbb{N}}$ does not matter, we also write $\mathcal{R}_{-\alpha}(a_n)$ and $\mathcal{R}_{-\alpha}$, respectively. For further information regarding multivariate regular variation of random vectors, we refer the reader to [36].

Definition 2.3. Let \mathbf{U} be an \mathbb{R}^d -valued random vector, $\alpha \in (0, 2]$, $(a_n)_{n \in \mathbb{N}}$ be an increasing sequence of positive constants tending to ∞ , μ be a Radon measure on $(\overline{\mathbb{R}^d} \setminus \{\mathbf{0}_d\}, \mathcal{B}(\overline{\mathbb{R}^d} \setminus \{\mathbf{0}_d\}))$ with $\mu(\overline{\mathbb{R}^d} \setminus \mathbb{R}^d) = 0$, and let $\Sigma \in M_{d \times d}(\mathbb{R})$ be a positive semidefinite matrix. We write $\mathbf{U} \in \text{DA}(\alpha, a_n, \Sigma, \mu)$ if either

- (a) $\alpha < 2$, $\Sigma = \mathbf{0}_{d \times d}$, μ is nonzero, and $\mathbf{U} \in \mathcal{R}_{-\alpha}(a_n, \mu)$; or
- (b) $\alpha = 2$, $a_n = n^{1/2}$, $\mu = 0$, and $\mathbb{E}\|\mathbf{U}\|^2 < \infty$ with $\mathbb{E}(\mathbf{U}\mathbf{U}^\top) = \Sigma$.

The abbreviation DA stands for *domain of attraction* because of the following argument. Let $(\mathbf{U}_k)_{k \in \mathbb{N}}$ be a sequence of i.i.d. \mathbb{R}^d -valued random vectors with $\mathbf{U}_1 \in \text{DA}(\alpha, a_n, \mu, \Sigma)$, $\alpha \neq 1$, and $\mathbf{S} = (\mathbf{S}(t))_{t \geq 0}$ an \mathbb{R}^d -valued α -stable Lévy process with characteristic triplet

$(\int_{\|x\| \leq 1} x \mu(dx), \Sigma, \mu)$ if $\alpha \in (0, 1)$ and $(-\int_{\|x\| > 1} x \mu(dx), \Sigma, \mu)$ if $\alpha > 1$. Assume that $\mathbb{E}(U_1) = \mathbf{0}_d$ if $\alpha > 1$. Then

$$a_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} U_k \Rightarrow S \quad \text{as } n \rightarrow \infty \text{ in } \mathbb{D}([0, 1], \mathbb{R}^d).$$

This means that the triplet (α, μ, Σ) characterizes completely the limit distribution and $(a_n)_{n \in \mathbb{N}}$ the convergence rate. For $\alpha = 1$, we additionally need a deterministic shift factor to obtain the convergence, which we can neglect if U_1 is symmetric. In general, the only possible limit of a normalized partial sum of i.i.d. random vectors is an α -stable distribution with $\alpha \in (0, 2]$ (cf. [27, Theorem 8.2.1.8] and [37]). The limit distribution is an α -stable random vector with $\alpha < 2$ if and only if U_1 is multivariate regularly varying of index $-\alpha$. Then $\mathbb{E}\|U_1\|^2 = \infty$ also. On the other hand, $\mathbb{E}\|U_1\|^2 < \infty$ is only a sufficient assumption to be in the domain of attraction of a normal distribution.

3. Main results

We start with a central limit theorem for MCARMA processes.

Theorem 3.1. *Let $(V(t))_{t \in \mathbb{R}}$ be an \mathbb{R}^d -valued causal MCARMA(p, q) process as given in Definition 2.1 driven by the \mathbb{R}^m -valued Lévy process $(L_1(t))_{t \in \mathbb{R}}$ with $L_1(1) \in \text{DA}(\alpha, a_n, \mu_1, \Sigma_1)$ and $\mathbb{E}(L_1(1)) = \mathbf{0}_m$ if $\alpha > 1$. Set $a_t := a_{\lfloor t \rfloor}$ for $t \geq 0$. If $\alpha = 1$, we additionally assume that $L_1(1)$ is symmetric.*

- (a) *Let $(S_1(t))_{t \geq 0}$ be an \mathbb{R}^m -valued α -stable Lévy process with characteristic triplet $(\int_{\|x\| \leq 1} x \mu_1(dx), \Sigma_1, \mu_1)$ if $\alpha \in (0, 1]$ and $(-\int_{\|x\| > 1} x \mu_1(dx), \Sigma_1, \mu_1)$ if $\alpha > 1$. Suppose that the sequence of positive constants $(h_n)_{n \in \mathbb{N}}$ satisfies $h_n \downarrow 0$ as $n \rightarrow \infty$ and $\lim_{n \rightarrow \infty} nh_n = \infty$. Then, as $n \rightarrow \infty$,*

$$h_n a_{nh_n}^{-1} \sum_{k=1}^n V(kh_n) \Rightarrow \left(\int_0^\infty f(s) ds \right) S_1(1).$$

- (b) *Let $h > 0$, and let $(S_{f,h}(t))_{t \geq 0}$ be an \mathbb{R}^d -valued α -stable Lévy process with characteristic triplet $(\int_{\|x\| \leq 1} x \mu_{f,h}(dx), \Sigma_{f,h}, \mu_{f,h})$ if $\alpha \in (0, 1]$ and $(-\int_{\|x\| > 1} x \mu_{f,h}(dx), \Sigma_{f,h}, \mu_{f,h})$ if $\alpha > 1$, where*

$$\mu_{f,h}(B) = \int_0^h \int_{\mathbb{R}^m} \mathbf{1}_B \left(\sum_{k=0}^\infty f(kh+s)x \right) \mu_1(dx) ds \quad \text{for } B \in \mathcal{B}(\mathbb{R}^d \setminus \{\mathbf{0}_d\}), \quad (3.1)$$

$$\Sigma_{f,h} = \int_0^h \left(\sum_{k=0}^\infty f(kh+s) \right) \Sigma_1 \left(\sum_{k=0}^\infty f(kh+s) \right)^\top ds. \quad (3.2)$$

Suppose that $\mathbb{E}\|L_1(1)\|^r < \infty$ for some $r > 2$ if $\alpha = 2$. Then, as $n \rightarrow \infty$,

$$a_n^{-1} \sum_{k=1}^n V(kh) \Rightarrow S_{f,h}(1).$$

We will compare this result to the limit results for ARMA models and present a motivation for the normalization.

Remarks 3.1. (a) Let $(\xi_k)_{k \in \mathbb{Z}}$ be a sequence of i.i.d. random vectors in \mathbb{R}^m with $\xi_1 \in \mathcal{R}_{-\alpha}(a_n, \mu_1)$ for some $0 < \alpha < 2$. If $\alpha > 1$ then suppose that $\mathbb{E}(\xi_1) = \mathbf{0}_m$, and if $\alpha = 1$ then suppose that ξ_1 is symmetric. Furthermore, let $(C_k)_{k \in \mathbb{N}}$ be a sequence of matrices in $M_{d \times m}(\mathbb{R})$ with $\sum_{k=0}^{\infty} k \|C_k\|^\theta < \infty$ for some $0 < \theta < \alpha, \theta \leq 1$. The \mathbb{R}^d -valued stationary moving average (MA) process $(X_k)_{k \in \mathbb{Z}}$ is defined as $X_k = \sum_{j=0}^{\infty} C_j \xi_{k-j}$ for $k \in \mathbb{Z}$. Then a special case of Theorem 4.1 (from below) is that, as $n \rightarrow \infty$,

$$a_n^{-1} \sum_{k=1}^n X_k \Rightarrow \left(\sum_{k=0}^{\infty} C_k \right) S_1(1).$$

On the one hand, we observe similar structures for the limit distributions $(\int_0^\infty f(s) ds) S_1(1)$ and $(\sum_{k=0}^{\infty} C_k) S_1(1)$ in the continuous-time high-frequency and discrete-time models. On the other hand, the normings are different. To explain the different normings, we consider an α -stable Lévy process $(L_1(t))_{t \geq 0}$ and an α -stable random variable ξ_1 . Then the idea in the continuous-time model is that, as $n \rightarrow \infty$,

$$\begin{aligned} h_n a_{nh_n}^{-1} \sum_{k=1}^n V(kh_n) &= \left(\sum_{j=0}^{\infty} f(jh_n) h_n \right) \left(a_{nh_n}^{-1} \sum_{k=1}^n [L_1(kh_n) - L_1((k-1)h_n)] \right) + o_{\mathbb{P}}(1) \\ &\stackrel{D}{=} \left(\int_0^\infty f(s) ds \right) S_1(1) + o_{\mathbb{P}}(1), \end{aligned} \tag{3.3}$$

and in the discrete-time model that, as $n \rightarrow \infty$,

$$a_n^{-1} \sum_{k=1}^n X_k = \left(\sum_{j=0}^{\infty} C_j \right) \left(a_n^{-1} \sum_{k=1}^n \xi_k \right) + o_{\mathbb{P}}(1) \stackrel{D}{=} \left(\sum_{j=0}^{\infty} C_j \right) \xi_1 + o_{\mathbb{P}}(1). \tag{3.4}$$

In (3.3) and (3.4) we see where the different normings have their origin. In the continuous-time model, the h_n of the norming $h_n a_{nh_n}^{-1}$ goes into the first factor of (3.3), which converges to $(\int_0^\infty f(s) ds)$, and the norming $a_{nh_n}^{-1}$ goes into the second, random factor.

(b) Representation (3.3) also justifies the fact that the classical techniques of Davis and Resnick [8] used to prove the asymptotic behavior of one-dimensional MA processes by using truncated MA processes will not work for the high-frequency case, because $\lim_{n \rightarrow \infty} \sum_{j=0}^M f(jh_n) h_n = \mathbf{0}_{d \times m}$ for $M > 0$.

Remark 3.2. A straightforward extension is the convergence of the finite-dimensional distribution for any $l \in \mathbb{N}$ as $n \rightarrow \infty$:

$$h_n a_{nh_n}^{-1} \left(\sum_{k=1}^n V(kh_n), \dots, \sum_{k=1}^n V((k+l)h_n) \right) \Rightarrow \left(\int_0^\infty f(s) ds \right) (S_1(1), \dots, S_1(1)).$$

Next we investigate the cointegrated model (1.1)–(1.2).

Theorem 3.2. *Let model (1.1)–(1.2) be given where \mathbb{X}_n has full rank, and let the assumptions of Theorem 3.1 hold. Furthermore, let $(L_2(t))_{t \in \mathbb{R}}$ be an \mathbb{R}^v -valued Lévy process independent of $(L_1(t))_{t \in \mathbb{R}}$, where $L_2(1) \in DA(\beta, b_n, \mu_2, \Sigma_2)$ and $\mathbb{E}(L_2(1)) = \mathbf{0}_v$ if $\beta > 1$. If $\beta = 1$, assume additionally that $L_2(1)$ is symmetric. Set $a_t := a_{\lfloor t \rfloor}$ and $b_t = b_{\lfloor t \rfloor}$ for $t \geq 0$. Moreover, let $(S_2(t))_{t \geq 0}$ be an \mathbb{R}^v -valued β -stable Lévy process independent of $(S_1(t))_{t \geq 0}$ with characteristic*

triplet $(\int_{\|x\| \leq 1} x \mu_2(dx), \Sigma_2, \mu_2)$ if $\beta \in (0, 1]$ and $(-\int_{\|x\| > 1} x \mu_2(dx), \Sigma_2, \mu_2)$ if $\beta > 1$, and suppose that

$$\mathbb{P}\left(\det\left(\int_0^1 S_2(s)S_2(s)^\top ds\right) = 0\right) = 0.$$

- (a) Suppose that the sequence of positive constants $(h_n)_{n \in \mathbb{N}}$ satisfies $h_n \downarrow 0$ as $n \rightarrow \infty$ and $\lim_{n \rightarrow \infty} nh_n = \infty$. If $\min(\alpha, \beta) < 2$ and either $\nu_{L_2}(\mathbb{R}^v) = \infty$ or $\Sigma_{L_2} \neq \mathbf{0}_{v \times v}$, we additionally assume that, for some $\epsilon > 0$,

$$\begin{aligned} \lim_{n \rightarrow \infty} n^{1/\min(\alpha, \beta) + \epsilon} h_n^{1/2} a_{nh_n}^{-1} b_{nh_n}^{-1} &= 0 \quad \text{if } \min(\alpha, \beta) \leq 1, \\ \lim_{n \rightarrow \infty} nh_n^{1/2} a_{nh_n}^{-1} b_{nh_n}^{-1} &= 0 \quad \text{if } 1 < \min(\alpha, \beta) < 2. \end{aligned} \tag{3.5}$$

Then \widehat{A}_n as given in (1.6) satisfies, as $n \rightarrow \infty$,

$$\begin{aligned} &nh_n a_{nh_n}^{-1} b_{nh_n} (\widehat{A}_n - A) \\ &\Rightarrow \left(\int_0^\infty f(s) ds\right) \left(S_1(1)S_2(1)^\top - \int_0^1 S_1(s-) dS_2(s)^\top\right) \left(\int_0^1 S_2(s)S_2(s)^\top ds\right)^{-1}. \end{aligned}$$

In particular, $\widehat{A}_n \xrightarrow{p} A$ as $n \rightarrow \infty$ if $\alpha > \beta/(\beta + 1)$, i.e. \widehat{A}_n is a consistent estimator.

- (b) Let $h > 0$ and $h_n = h$ for any $n \in \mathbb{N}$. Suppose that $\mathbb{E}\|L_1(1)\|^r < \infty$ for some $r > 2$ if $\alpha = 2$. Then \widehat{A}_n as given in (1.6) satisfies, as $n \rightarrow \infty$,

$$\begin{aligned} na_n^{-1} b_n (\widehat{A}_n - A) &\Rightarrow \left(S_{f,h}(1)S_2(1)^\top - \int_0^1 S_{f,h}(s-) dS_2(s)^\top\right) \\ &\quad \times \left(\int_0^1 S_2(s)S_2(s)^\top ds\right)^{-1}. \end{aligned}$$

In particular, $\widehat{A}_n \xrightarrow{p} A$ as $n \rightarrow \infty$ if $\alpha > \beta/(\beta + 1)$, i.e. \widehat{A}_n is a consistent estimator.

Remark 3.3. If $\alpha = \beta < 2$, sufficient conditions for (3.5) are that, for some $\epsilon > 0$,

$$\begin{aligned} \lim_{n \rightarrow \infty} nh_n^{2-\alpha/2+\epsilon} &= \infty \quad \text{if } \alpha \leq 1, \\ \lim_{n \rightarrow \infty} nh_n^{1/2+1/(2-\alpha)+\epsilon} &= \infty \quad \text{if } 1 < \alpha < 2 \end{aligned}$$

hold. A conjecture is that assumption (3.5) is not necessary.

Finally, we investigate the asymptotic behavior of the sample variance. Both Theorem 3.1 and Theorem 3.3 are used in [14] and [15] to derive the asymptotic behavior of the normalized, the self-normalized, and the smoothed periodogram as well as for statistical inference of CARMA processes.

Theorem 3.3. Let $(V(t))_{t \geq 0}$ be an \mathbb{R}^d -valued MCARMA(p, q) process as given in Definition 2.1 driven by the \mathbb{R}^m -valued Lévy process $(L_1(t))_{t \in \mathbb{R}}$ with $L_1(1) \in DA(\alpha, a_n, \mu_1, \Sigma_1)$. Set $a_t := a_{\lfloor t \rfloor}$ for $t \geq 0$.

- (a) Let $(S_1(t))_{t \geq 0}$ be an \mathbb{R}^m -valued α -stable Lévy process with characteristic triplet $(\mathbf{0}_m, \Sigma_1, \mu_1)$. Suppose that the sequence of positive constants $(h_n)_{n \in \mathbb{N}}$ satisfies $h_n \downarrow 0$ as $n \rightarrow \infty$

and $\lim_{n \rightarrow \infty} nh_n = \infty$. Then, as $n \rightarrow \infty$,

$$h_n a_{nh_n}^{-2} \sum_{k=1}^n V(kh_n) V(kh_n)^\top \Rightarrow \int_0^\infty f(s) [S_1, S_1]_1 f(s)^\top ds,$$

which is equal to $\mathbb{E}(V(0)V(0)^\top)$ if $\alpha = 2$.

- (b) Let $h > 0$, and let $(S_{f,h}(t))_{t \geq 0}$ be an \mathbb{R}^d -valued α -stable Lévy process with characteristic triplet $(\mathbf{0}_d, \Sigma_{f,h}, \mu_{f,h})$, where $\mu_{f,h}$ and $\Sigma_{f,h}$ are given as in (3.1) and (3.2), respectively. Then, as $n \rightarrow \infty$,

$$a_n^{-2} \sum_{k=1}^n V(kh) V(kh)^\top \Rightarrow [S_{f,h}, S_{f,h}]_1,$$

which is equal to $\Sigma_{f,h}$ if $\alpha = 2$.

Thus, if $\mathbb{E}\|L_1(1)\|^2 < \infty$, the sample variance is a consistent estimator. Furthermore, we want to point out that in contrast to Theorem 3.1, Theorem 3.3 does not require $\mathbb{E}(L_1(1)) = \mathbf{0}_d$ if $1 < \alpha < 2$ and the symmetry of $L_1(1)$ if $\alpha = 1$. Also, the drift term of S_1 can be chosen arbitrarily since it does not have any influence on $[S_1, S_1]_1$. Using the Kronecker product ‘ \otimes ’ and the vec operator ‘vec’, which transforms a matrix into a vector by stacking the columns, the representation $\text{vec}(\int_0^\infty f(s) [S_1, S_1]_1 f(s)^\top ds) = [\int_0^\infty f(s) \otimes f(s) ds] \text{vec}([S_1, S_1]_1)$ holds. This representation shows that the limiting distribution again decomposes into a parametric and a random part.

As in Remark 3.1, we will make a comparison to the discrete-time case.

Remark 3.4. Let a discrete-time MA process as in Remark 3.1 be given. Then, by [9, Theorem 2.1] for the two-dimensional case (see also [26, Equation (4.7)]), as $n \rightarrow \infty$,

$$a_n^{-2} \sum_{k=1}^n X_k X_k^\top \Rightarrow \sum_{k=0}^\infty C_k [S_1, S_1]_1 C_k^\top.$$

Again, we see the similarity between the continuous-time high-frequency and discrete-time models. Considering an α -stable Lévy process $(L_1(t))_{t \geq 0}$ and an α -stable random variable ξ_1 , the normings can be understood in the continuous-time high-frequency model by

$$\begin{aligned} & h_n a_{nh_n}^{-2} \sum_{k=1}^n V(kh_n) V(kh_n)^\top \\ &= \sum_{j=0}^\infty f(jh_n) \left(a_{nh_n}^{-2} \sum_{k=1}^n [L_1(kh_n) - L_1((k-1)h_n)] \right. \\ & \quad \left. \times [L_1(kh_n) - L_1((k-1)h_n)]^\top \right) f(jh_n)^\top h_n + o_{\mathbb{P}}(1) \\ & \stackrel{D}{=} \sum_{j=0}^\infty f(jh_n) [L_1, L_1]_1^\top f(jh_n)^\top h_n + o_{\mathbb{P}}(1) \\ &= \int_0^\infty f(s) [S_1, S_1]_1 f(s)^\top ds + o_{\mathbb{P}}(1). \end{aligned}$$

The first factor h_n of $h_n a_{nh_n}^{-2}$ is required for the convergence of the integral and $a_{nh_n}^{-2}$ for the random part. In the discrete-time model we have

$$a_n^{-2} \sum_{k=1}^n \mathbf{X}_k \mathbf{X}_k^\top = \sum_{j=0}^{\infty} \mathbf{C}_j \left(a_n^{-2} \sum_{k=1}^n \boldsymbol{\xi}_k \boldsymbol{\xi}_k^\top \right) \mathbf{C}_j^\top + o_{\mathbb{P}}(1) \stackrel{D}{=} \sum_{j=0}^{\infty} \mathbf{C}_j [\mathbf{S}_1, \mathbf{S}_1]_1 \mathbf{C}_j^\top + o_{\mathbb{P}}(1).$$

Remark 3.5. An obvious extension is that the finite-dimensional distribution of the sample autocovariance function has, for any $l \in \mathbb{N}$, the asymptotic behavior

$$h_n a_{nh_n}^{-2} \left(\sum_{k=1}^n \mathbf{V}(kh_n) \mathbf{V}(kh_n)^\top, \dots, \sum_{k=1}^n \mathbf{V}(kh_n) \mathbf{V}((k+l)h_n)^\top \right) \Rightarrow \left(\int_0^\infty \mathbf{f}(s) [\mathbf{S}_1, \mathbf{S}_1]_1 \mathbf{f}(s)^\top ds, \dots, \int_0^\infty \mathbf{f}(s) [\mathbf{S}_1, \mathbf{S}_1]_1 \mathbf{f}(s)^\top ds \right) \text{ as } n \rightarrow \infty.$$

4. Multivariate high-frequency model

In this section we derive the properties of the least-squares estimator given in (1.6) for model (1.4)–(1.5). As mentioned in the introduction and used in the proof of Theorem 3.1, the cointegrated MCARMA model can be seen as a special case of this more general model.

Assumption 4.1. Let model (1.4)–(1.5) be given.

- (a) Suppose that there exist sequences of positive constants $\tilde{a}_n, \tilde{b}_n \uparrow \infty$ as $n \rightarrow \infty$ such that, as $n \rightarrow \infty$,

$$\left(\tilde{a}_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{\xi}_{n,k}^\top, \tilde{b}_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{\varepsilon}_{n,k}^\top \right)_{t \geq 0} \Rightarrow (\mathbf{S}_1(t)^\top, \mathbf{S}_2(t)^\top)_{t \geq 0} \text{ in } \mathbb{D}([0, 1], \mathbb{R}^{m+v}),$$

where $\mathbf{S}_1 = (\mathbf{S}_1(t))_{t \geq 0}$ is a càdlàg stochastic process in \mathbb{R}^m and $\mathbf{S}_2 = (\mathbf{S}_2(t))_{t \geq 0}$ is a càdlàg stochastic process in \mathbb{R}^v . Furthermore, suppose that

$$\mathbb{P} \left(\det \left(\int_0^1 \mathbf{S}_2(s) \mathbf{S}_2(s)^\top ds \right) = 0 \right) = 0. \tag{4.1}$$

- (b) Define

$$\tilde{\mathbf{Z}}_{n,k} := \sum_{j=0}^{\infty} \left(\sum_{l=j+1}^{\infty} \mathbf{C}_{n,l} \right) \boldsymbol{\xi}_{n,k-j} \text{ for } k \in \mathbb{N}_0, n \in \mathbb{N}.$$

Suppose that there exist a sequence of positive constants $(h_n)_{n \in \mathbb{N}}$ and a positive bounded decreasing function g with either $g \in \mathcal{R}_{-\alpha}$, $\alpha \in (0, 2)$, or $\int_0^\infty xg(x) dx < \infty$ and $\alpha := 2$, such that

$$\mathbb{P}(h_n \|\tilde{\mathbf{Z}}_{n,0}\| > x) \leq g(x) \text{ for } x \geq 0, n \in \mathbb{N}.$$

- (c) Let $\sum_{k=0}^{\infty} k \|\mathbf{C}_{n,k}\|^\theta < \infty$ for some $0 < \theta < \alpha$ and $\theta \leq 1$. Furthermore, there exists a matrix $\mathbf{C} \in M_{d \times m}(\mathbb{R})$ for $(h_n)_{n \in \mathbb{N}}$ in (b) such that

$$\lim_{n \rightarrow \infty} h_n \sum_{k=0}^{\infty} \mathbf{C}_{n,k} = \mathbf{C}.$$

(d) *There exist constants $K_1, K_2, K_3 < \infty$ and some $0 < \delta < \alpha$ with $\delta \leq 1$ such that the following assertions hold:*

- (i) $n\tilde{b}_n^{-2}\mathbb{E}(\|\boldsymbol{\varepsilon}_{n,1}\|^2\mathbf{1}_{\{\|\boldsymbol{\varepsilon}_{n,1}\|\leq\tilde{b}_n\}}) \leq K_1$ for all $n \in \mathbb{N}$;
- (ii) $n\tilde{b}_n^{-1}\|\mathbb{E}(\boldsymbol{\varepsilon}_{n,1}\mathbf{1}_{\{\|\boldsymbol{\varepsilon}_{n,1}\|\leq\tilde{b}_n\}})\| \leq K_2$ for all $n \in \mathbb{N}$;
- (iii) $n\tilde{b}_n^{-\delta}\mathbb{E}(\|\boldsymbol{\varepsilon}_{n,1}\|^\delta\mathbf{1}_{\{\|\boldsymbol{\varepsilon}_{n,1}\|>\tilde{b}_n\}}) \leq K_3$ for all $n \in \mathbb{N}$.

Furthermore, one of the following conditions is satisfied for g in (b):

- (iv.1) $g \in \mathcal{R}_{-\alpha}$ with $\alpha \in (0, 2)$ and $\lim_{n \rightarrow \infty} n\tilde{a}_n^{-\delta}\tilde{b}_n^{-\delta}\mathbb{E}\|\boldsymbol{\varepsilon}_{n,1}\|^\delta = 0$;
- (iv.2) $\int_0^\infty xg(x) dx < \infty$ and $\lim_{n \rightarrow \infty} n\tilde{a}_n^{-2}\tilde{b}_n^{-2}\mathbb{E}\|\boldsymbol{\varepsilon}_{n,1}\|^2 = 0$.

Note that if g is a positive bounded decreasing function with $g \in \mathcal{R}_{-\alpha}$, $\alpha \in (0, 2)$, then $\int_0^\infty x^{\gamma-1}g(x) dx < \infty$ for any $0 < \gamma < \alpha$ (apply Karamata’s theorem (cf. [36, Theorem 2.1])). Moreover, $\lim_{n \rightarrow \infty} g(\tilde{a}_n) = 0$.

We state the first limit result.

Theorem 4.1. *Let model (1.4)–(1.5) be given where \mathbb{X}_n has full rank, and let Assumption 4.1 hold. Define*

$$\mathbf{S}_{1,n}(t) := h_n\tilde{a}_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \mathbf{Z}_{n,k} \quad \text{and} \quad \mathbf{S}_{2,n}(t) := \tilde{b}_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{\varepsilon}_{n,k} \quad \text{for } t \geq 0, n \in \mathbb{N}.$$

Then, as $n \rightarrow \infty$,

$$\begin{aligned} & \left(\mathbf{S}_{1,n}(1), \mathbf{S}_{2,n}(1), \int_0^1 \mathbf{S}_{2,n}(s)\mathbf{S}_{2,n}(s)^\top ds, \int_0^1 \mathbf{S}_{1,n}(s-) d\mathbf{S}_{2,n}(s)^\top \right) \\ & \Rightarrow \left(\mathbf{C}\mathbf{S}_1(1), \mathbf{S}_2(1), \int_0^1 \mathbf{S}_2(s)\mathbf{S}_2(s)^\top ds, \mathbf{C} \int_0^1 \mathbf{S}_1(s-) d\mathbf{S}_2(s)^\top \right) \end{aligned}$$

in $\mathbb{R}^d \times \mathbb{R}^v \times \mathbb{R}^{v \times v} \times \mathbb{R}^{d \times v}$.

Based on this theorem we are able to derive the asymptotic behavior of the least-squares estimator in the cointegrated model.

Theorem 4.2. *Let model (1.4)–(1.5) be given, and let Assumption 4.1 hold. Then $\widehat{\mathbf{A}}_n$ as given in (1.6) satisfies, as $n \rightarrow \infty$,*

$$nh_n\tilde{a}_n^{-1}\tilde{b}_n(\widehat{\mathbf{A}}_n - \mathbf{A}) \Rightarrow \mathbf{C} \left(\mathbf{S}_1(1)\mathbf{S}_2(1)^\top - \int_0^1 \mathbf{S}_1(s-) d\mathbf{S}_2(s)^\top \right) \left(\int_0^1 \mathbf{S}_2(s)\mathbf{S}_2(s)^\top ds \right)^{-1}.$$

In particular, $\widehat{\mathbf{A}}_n \xrightarrow{P} \mathbf{A}$ as $n \rightarrow \infty$ if $\lim_{n \rightarrow \infty} nh_n\tilde{a}_n^{-1}\tilde{b}_n = \infty$, i.e. $\widehat{\mathbf{A}}_n$ is a consistent estimator.

5. Proofs

5.1. Proofs of Section 4

The proofs of this section are very similar to [13]. However, we mimic them to show where the different assumptions hold. An essential piece of the proof will be that, as $n \rightarrow \infty$,

$$h_n\tilde{a}_n^{-1} \sum_{k=1}^n \mathbf{Z}_{n,k} = \left(h_n \sum_{j=0}^\infty \mathbf{C}_{n,j} \right) \left(\tilde{a}_n^{-1} \sum_{k=1}^n \boldsymbol{\xi}_{n,k} \right) + o_{\mathbb{P}}(1). \tag{5.1}$$

We require the next lemma for the proofs of Theorem 3.3 and Theorem 4.1.

Lemma 5.1. *Let $(\boldsymbol{\varepsilon}_{n,k})_{k \in \mathbb{N}}$ be an i.i.d. sequence of random vectors in \mathbb{R}^v , and let $(\mathbf{W}_{n,k})_{k \in \mathbb{N}}$ be a sequence of random vectors in \mathbb{R}^d for any $n \in \mathbb{N}$, where $(\mathbf{W}_{n,k-j})_{j=1, \dots, k-1}$ is independent of $(\boldsymbol{\varepsilon}_{n,k+j})_{j \in \mathbb{N}}$ for any $n, k \in \mathbb{N}$. Suppose that there exists a positive, bounded, decreasing function g such that*

$$\mathbb{P}(\|\mathbf{W}_{n,k}\| > x) \leq g(x) \quad \text{for any } x \geq 0, n \in \mathbb{N}, k \in \mathbb{N}.$$

Assume that one of the following conditions is satisfied:

- (a) $g \in \mathcal{R}_{-\alpha}$, $0 < \alpha < 2$, and, for some $0 < \delta \leq 1$, $\delta < \alpha$, $\lim_{n \rightarrow \infty} n \tilde{a}_n^{-\delta} \tilde{b}_n^{-\delta} \mathbb{E}\|\boldsymbol{\varepsilon}_{n,1}\|^\delta = 0$ holds;
- (b) $\int_0^\infty xg(x) dx < \infty$, $\mathbb{E}(\boldsymbol{\varepsilon}_{n,1}) = \mathbf{0}_v$, for $n \in \mathbb{N}$, and $\lim_{n \rightarrow \infty} n \tilde{a}_n^{-2} \tilde{b}_n^{-2} \mathbb{E}\|\boldsymbol{\varepsilon}_{n,1}\|^2 = 0$. Then, as $n \rightarrow \infty$,

$$\tilde{a}_n^{-1} \tilde{b}_n^{-1} \sum_{k=1}^n \mathbf{W}_{n,k-1} \boldsymbol{\varepsilon}_{n,k}^\top \xrightarrow{P} \mathbf{0}_{d \times v}.$$

Proof. Case (a). Taking $\delta \leq 1$ into account we have

$$\tilde{a}_n^{-\delta} \tilde{b}_n^{-\delta} \mathbb{E} \left\| \sum_{k=1}^n \mathbf{W}_{n,k-1} \boldsymbol{\varepsilon}_{n,k}^\top \right\|^\delta \leq n \tilde{a}_n^{-\delta} \tilde{b}_n^{-\delta} \left(\delta \int_0^\infty x^{\delta-1} g(x) dx \right) \mathbb{E}\|\boldsymbol{\varepsilon}_{n,1}\|^\delta \rightarrow 0$$

as $n \rightarrow \infty$.

Case (b). We investigate the sequence of random matrices componentwise, and denote by (l, m) the component in the l th row and m th column. Since $((\mathbf{W}_{n,k-1} \boldsymbol{\varepsilon}_{n,k}^\top)_{(l,m)})_{k \in \mathbb{N}}$ is an uncorrelated sequence,

$$\begin{aligned} \tilde{a}_n^{-2} \tilde{b}_n^{-2} \mathbb{E} \left(\left(\sum_{k=1}^n \mathbf{W}_{n,k-1} \boldsymbol{\varepsilon}_{n,k}^\top \right)_{(l,m)}^2 \right) &\leq C_1 \tilde{a}_n^{-2} \tilde{b}_n^{-2} \sum_{k=1}^n \mathbb{E}\|\mathbf{W}_{n,k-1}\|^2 \mathbb{E}\|\boldsymbol{\varepsilon}_{n,k}\|^2 \\ &\leq C_2 n \tilde{a}_n^{-2} \tilde{b}_n^{-2} \mathbb{E}\|\boldsymbol{\varepsilon}_{n,1}\|^2. \end{aligned}$$

The last expression tends to 0 as $n \rightarrow \infty$ by assumption.

We will prove Theorem 4.1 by an application of [20, Theorem VI.6.22]. Therefore, we need the definition of predictably uniformly tightness (P-UT) given in [20, Section VI.6a, p. 377] and the next lemma.

Lemma 5.2. *Let Assumptions 4.1(d) hold. Then the sequence of stochastic processes $(\mathbf{S}_{2,n})_{n \in \mathbb{N}}$ as given in Theorem 4.1 is P-UT on $(\Omega, \mathcal{F}, ((\mathcal{F}_t^n)_{t \geq 0})_{n \in \mathbb{N}}, \mathbb{P})$ with $\mathcal{F}_t^n = \sigma(\boldsymbol{\varepsilon}_{n,k} : k \leq \lfloor nt \rfloor)$, $t \geq 0, n \in \mathbb{N}$.*

Proof. We define, for $t \geq 0$ and $n \in \mathbb{N}$,

$$\begin{aligned} \mathbf{M}_n(t) &:= \tilde{b}_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} (\boldsymbol{\varepsilon}_{n,k} \mathbf{1}_{\{\|\boldsymbol{\varepsilon}_{n,k}\| \leq \tilde{b}_n\}} - \mathbb{E}(\boldsymbol{\varepsilon}_{n,1} \mathbf{1}_{\{\|\boldsymbol{\varepsilon}_{n,1}\| \leq \tilde{b}_n\}})), \\ \mathbf{D}_n^{(1)}(t) &:= \lfloor nt \rfloor \tilde{b}_n^{-1} \mathbb{E}(\boldsymbol{\varepsilon}_{n,1} \mathbf{1}_{\{\|\boldsymbol{\varepsilon}_{n,1}\| \leq \tilde{b}_n\}}), \quad \mathbf{D}_n^{(2)}(t) := \tilde{b}_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{\varepsilon}_{n,k} \mathbf{1}_{\{\|\boldsymbol{\varepsilon}_{n,k}\| > \tilde{b}_n\}}. \end{aligned}$$

It is obvious that $(M_n(t))_{t \geq 0}$ is a martingale with respect to $(\mathcal{F}_t^n)_{t \geq 0}$ and, in particular, a local martingale. All three processes are adapted with respect to $(\mathcal{F}_t^n)_{t \geq 0}$ and we have the semimartingale decomposition

$$S_{2,n}(t) = M_n(t) + D_n^{(1)}(t) + D_n^{(2)}(t).$$

If $(M_n)_{n \in \mathbb{N}}$, $(D_n^{(1)})_{n \in \mathbb{N}}$, and $(D_n^{(2)})_{n \in \mathbb{N}}$ are P-UT, then it follows from [20, Property VI.6.4] that $(S_{2,n})_{n \in \mathbb{N}}$ is P-UT as well.

Let $VT_s(W) = \sup_{i=1, \dots, v} VT_s(W_i)$ for $s \geq 0$ denote the variation process of the càdlàg stochastic process $(W(s))_{s \geq 0} = (W_1(s), \dots, W_v(s))_{s \geq 0}$. To prove the ‘P-UTness’ of $(D_n^{(1)})_{n \in \mathbb{N}}$ and $(D_n^{(2)})_{n \in \mathbb{N}}$, it is sufficient to show that $(VT_t(D_n^{(1)}))_{n \in \mathbb{N}}$ and $(VT_t(D_n^{(2)}))_{n \in \mathbb{N}}$ are tight for any $t \geq 0$; see [20, Property VI.6.6]. Let $t \geq 0$ be fixed. We start with the verification of the tightness of $(VT_t(D_n^{(1)}))_{n \in \mathbb{N}}$ by showing that it is uniformly bounded. Assumption 4.1(d)(ii) gives the uniform bound

$$\sup_{n \in \mathbb{N}} VT_t(D_n^{(1)}) \leq C_1 \sup_{n \in \mathbb{N}} nt \tilde{b}_n^{-1} \|\mathbb{E}(\epsilon_{n,1} \mathbf{1}_{\{\|\epsilon_{n,1}\| \leq \tilde{b}_n\}})\| \leq C_2 t,$$

which results in the tightness of $(VT_t(D_n^{(1)}))_{n \in \mathbb{N}}$.

For the proof of the tightness of $(VT_t(D_n^{(2)}))_{n \in \mathbb{N}}$, we use the fact that $(VT_t(D_n^{(2)}))^\delta \leq C_3 \tilde{b}_n^{-\delta} \sum_{k=1}^{\lfloor nt \rfloor} \|\epsilon_{n,k}\|^\delta \mathbf{1}_{\{\|\epsilon_{n,k}\| > \tilde{b}_n\}}$ for $\delta \leq 1$. Then, using Assumption 4.1(d)(iii) and Markov’s inequality, we conclude that

$$\sup_{n \in \mathbb{N}} \mathbb{P}(VT_t(D_n^{(2)}) > \eta) \leq C_4 \eta^{-\delta} \sup_{n \in \mathbb{N}} \tilde{b}_n^{-\delta} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E}(\|\epsilon_{n,k}\|^\delta \mathbf{1}_{\{\|\epsilon_{n,k}\| > \tilde{b}_n\}}) \leq C_5 \eta^{-\delta} t,$$

which tends to 0 as $\eta \rightarrow \infty$. Hence, $(VT_t(D_n^{(2)}))_{n \in \mathbb{N}}$ is also tight. If we show that $([M_n, M_n]_t)_{n \in \mathbb{N}}$ is tight for any $t \geq 0$, then the ‘P-UTness’ of $(M_n)_{n \in \mathbb{N}}$ follows by [20, Proposition VI.6.13]. Here, we use Assumption 4.1(d)(i) to conclude that

$$\sup_{n \in \mathbb{N}} \mathbb{P}(\|[M_n, M_n]_t\| > \eta) \leq \eta^{-1} \sup_{n \in \mathbb{N}} n \tilde{b}_n^{-2} \mathbb{E}(\|\epsilon_{n,1}\|^2 \mathbf{1}_{\{\|\epsilon_{n,1}\| \leq \tilde{b}_n\}}) \leq C_6 \eta^{-1} \rightarrow 0$$

as $\eta \rightarrow \infty$. Finally, $([M_n, M_n]_t)_{n \in \mathbb{N}}$ is tight as well.

Proof of Theorem 4.1. The Beveridge–Nelson decomposition (cf. [3]) has the representation

$$Z_{n,k} = \left(\sum_{j=0}^{\infty} C_{n,j} \right) \xi_{n,k} + \tilde{Z}_{n,k-1} - \tilde{Z}_{n,k} \quad \text{for } k, n \in \mathbb{N}.$$

Thus,

$$S_{1,n}(t) = h_n \tilde{a}_n^{-1} \left(\sum_{j=0}^{\infty} C_{n,j} \right) \sum_{k=1}^{\lfloor nt \rfloor} \xi_{n,k} + h_n \tilde{a}_n^{-1} [\tilde{Z}_{n,0} - \tilde{Z}_{n,\lfloor nt \rfloor}] \quad \text{for } t \geq 0. \tag{5.2}$$

Therefore, we define

$$\tilde{S}_{1,n}(t) := \left(h_n \sum_{j=0}^{\infty} C_{n,j} \right) \tilde{a}_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \xi_{n,k} \quad \text{for } t \geq 0.$$

By Assumption 4.1(a) and (c) we have, as $n \rightarrow \infty$,

$$(\tilde{\mathbf{S}}_{1,n}(t)^\top, \mathbf{S}_{2,n}(t)^\top)_{t \geq 0}^\top \Rightarrow ((\mathbf{C}\mathbf{S}_1(t))^\top, \mathbf{S}_2(t)^\top)_{t \geq 0}^\top \quad \text{in } \mathbb{D}([0, 1], \mathbb{R}^{d+v}).$$

A straightforward conclusion of the continuous mapping theorem is then, as $n \rightarrow \infty$,

$$\begin{aligned} & \left(\tilde{\mathbf{S}}_{1,n}(1), \mathbf{S}_{2,n}(1), \int_0^1 \mathbf{S}_{2,n}(s)\mathbf{S}_{2,n}(s)^\top ds, (\tilde{\mathbf{S}}_{1,n}(t)^\top, \mathbf{S}_{2,n}(t)^\top)_{t \geq 0}^\top \right) \\ & \Rightarrow \left(\mathbf{C}\mathbf{S}_1(1), \mathbf{S}_2(1), \int_0^1 \mathbf{S}_2(s)\mathbf{S}_2(s)^\top ds, ((\mathbf{C}\mathbf{S}_1(t))^\top, \mathbf{S}_2(t)^\top)_{t \geq 0}^\top \right) \end{aligned}$$

in $\mathbb{R}^d \times \mathbb{R}^v \times \mathbb{R}^{v \times v} \times (\mathbb{D}[0, 1], \mathbb{R}^{d+v})$. Since $(\mathbf{S}_{2,n})_{n \in \mathbb{N}}$ is P-UT by Lemma 5.2, a result of [20, Theorem VI.6.22] is that, as $n \rightarrow \infty$,

$$\begin{aligned} & \left(\tilde{\mathbf{S}}_{1,n}(1), \mathbf{S}_{2,n}(1), \int_0^1 \mathbf{S}_{2,n}(s)\mathbf{S}_{2,n}(s)^\top ds, \int_0^1 \tilde{\mathbf{S}}_{1,n}(s-) d\mathbf{S}_{2,n}(s)^\top \right) \\ & \Rightarrow \left(\mathbf{C}\mathbf{S}_1(1), \mathbf{S}_2(1), \int_0^1 \mathbf{S}_2(s)\mathbf{S}_2(s)^\top ds, \mathbf{C} \int_0^1 \mathbf{S}_1(s-) d\mathbf{S}_2(s)^\top \right) \end{aligned} \tag{5.3}$$

in $\mathbb{R}^d \times \mathbb{R}^v \times \mathbb{R}^{v \times v} \times \mathbb{R}^{d \times v}$. On the one hand, by (5.2) we have

$$\begin{aligned} & \int_0^1 \mathbf{S}_{1,n}(s-) d\mathbf{S}_{2,n}(s)^\top \\ & = \int_0^1 \tilde{\mathbf{S}}_{1,n}(s-) d\mathbf{S}_{2,n}(s)^\top + \left[h_n \tilde{a}_n^{-1} \tilde{\mathbf{Z}}_{n,0} \mathbf{S}_{2,n}(1) - h_n \tilde{a}_n^{-1} \tilde{b}_n^{-1} \sum_{k=1}^n \tilde{\mathbf{Z}}_{n,k-1} \mathbf{e}_{n,k}^\top \right]. \end{aligned} \tag{5.4}$$

Applying Lemma 5.1, $h_n \tilde{a}_n^{-1} \tilde{\mathbf{Z}}_{n,0} \xrightarrow{P} \mathbf{0}_d$ as $n \rightarrow \infty$ (by Assumption 4.1(b)), and $\mathbf{S}_{2,n}(1) \Rightarrow \mathbf{S}_2(1)$ as $n \rightarrow \infty$ gives, on the other hand,

$$h_n \tilde{a}_n^{-1} \tilde{b}_n^{-1} \tilde{\mathbf{Z}}_{n,0} \mathbf{S}_{2,n}(1) - h_n \tilde{a}_n^{-1} \tilde{b}_n^{-1} \sum_{k=1}^n \tilde{\mathbf{Z}}_{n,k-1} \mathbf{e}_{n,k}^\top \xrightarrow{P} \mathbf{0}_{d \times v} \quad \text{as } n \rightarrow \infty. \tag{5.5}$$

Finally, the statement follows from (5.3)–(5.5).

Proof of Theorem 4.2. (a) Since $\mathbb{Y}_n^\top = \mathbf{A}\mathbb{X}_n^\top + \mathbb{Z}_n^\top$ with \mathbb{X}_n and \mathbb{Y}_n as given in (1.5), and $\mathbb{Z}_n^\top = (\mathbf{Z}_{n,1}, \dots, \mathbf{Z}_{n,n})$, we have

$$nh_n \tilde{a}_n^{-1} \tilde{b}_n (\hat{\mathbf{A}}_n - \mathbf{A}) = (h_n \tilde{a}_n^{-1} \mathbb{Z}_n^\top \mathbb{X}_n \tilde{b}_n^{-1}) (n^{-1} \tilde{b}_n^{-1} \mathbb{X}_n^\top \mathbb{X}_n \tilde{b}_n^{-1})^{-1}. \tag{5.6}$$

Now

$$\begin{aligned} & (h_n \tilde{a}_n^{-1} \tilde{b}_n^{-1} \mathbb{Z}_n^\top \mathbb{X}_n, n^{-1} \tilde{b}_n^{-2} \mathbb{X}_n^\top \mathbb{X}_n) \\ & = \left(\mathbf{S}_{1,n}(1)\mathbf{S}_{2,n}(1)^\top - \int_0^1 \mathbf{S}_{1,n}(s-) d\mathbf{S}_{2,n}(s)^\top, \int_0^1 \mathbf{S}_{2,n}(s)\mathbf{S}_{2,n}(s)^\top ds \right) \\ & \Rightarrow \left(\mathbf{C}\mathbf{S}_1(1)\mathbf{S}_2(1)^\top - \mathbf{C} \int_0^1 \mathbf{S}_1(s-) d\mathbf{S}_2(s)^\top, \int_0^1 \mathbf{S}_2(s)\mathbf{S}_2(s)^\top ds \right) \end{aligned} \tag{5.7}$$

in $\mathbb{R}^{d \times v} \times \mathbb{R}^{v \times v}$ as $n \rightarrow \infty$ by Theorem 4.1. Finally, (5.6)–(5.7) yield the claim by the continuous mapping theorem, since (4.1) holds.

5.2. Proof of Theorem 3.1

It is well known that the stationary Ornstein–Uhlenbeck process \mathbf{Z} given in (2.4) observed at the time grid $h_n\mathbb{Z}$ has the MA process representation

$$\mathbf{Z}(kh_n) = \sum_{j=0}^{\infty} e^{-\Lambda h_n j} \boldsymbol{\xi}_{n,k-j} \quad \text{for } k \in \mathbb{Z},$$

where

$$\boldsymbol{\xi}_{n,k} = \int_{(k-1)h_n}^{kh_n} e^{-\Lambda(kh_n-s)} \mathbf{B} \, d\mathbf{L}_1(s) \quad \text{for } k \in \mathbb{Z}, n \in \mathbb{N}. \tag{5.8}$$

As already suggested in (5.1), as $n \rightarrow \infty$,

$$h_n a_{nh_n}^{-1} \sum_{k=1}^n \mathbf{V}(kh_n) = \left(h_n \sum_{j=0}^{\infty} \mathbf{E} e^{-\Lambda h_n j} \right) \left(a_{nh_n}^{-1} \sum_{k=1}^n \boldsymbol{\xi}_{n,k} \right) + o_{\mathbb{P}}(1).$$

The convergence of $a_{nh_n}^{-1} \sum_{k=1}^n \boldsymbol{\xi}_{n,k}$ is based on central limit results for arrays and the properties of the sequence of i.i.d. random vectors $(\boldsymbol{\xi}_{n,k})_{k \in \mathbb{Z}}$ as presented in Appendix A.

Before we state the proof of Theorem 3.1, we present the analogous result for the state process \mathbf{Z} which is essential for the proof of Theorem 3.1.

Lemma 5.3. *Let the assumptions of Theorem 3.1 hold. Then $h_n a_{nh_n}^{-1} \sum_{k=1}^n \mathbf{Z}(kh_n) \Rightarrow \boldsymbol{\Lambda}^{-1} \mathbf{B} \mathbf{S}_1(1)$ as $n \rightarrow \infty$.*

Proof. First, we define $\tilde{a}_n := a_{nh_n}$, $\mathbf{C}_{n,k} := e^{-\Lambda h_n k}$, and $(\boldsymbol{\xi}_{n,k})$ as given in (5.8). Then

$$\mathbf{Z}_{n,k} := \mathbf{Z}(kh_n) = \sum_{j=0}^{\infty} \mathbf{C}_{n,j} \boldsymbol{\xi}_{n,k-j} \quad \text{for } k \in \mathbb{Z}, n \in \mathbb{N}.$$

We will show that Assumption 4.1(a)–(d) with $\boldsymbol{\varepsilon}_{n,k} := 0$ are satisfied because then the result follows from Theorem 4.1 (it does not matter that (4.1) is not satisfied for $\boldsymbol{\varepsilon}_{n,k} = 0$).

(a) Consider the $0 < \alpha < 2$ case. By Proposition A.2(a), (c), and (d), $\mathbb{E}(\boldsymbol{\xi}_{n,0}) = \mathbf{0}_{pd}$ if $\alpha > 1$, $\boldsymbol{\xi}_{n,0}$ symmetric for $\alpha = 1$, and [36, Theorem 7.1], we have

$$\left(\tilde{a}_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{\xi}_{n,k} \right)_{t \geq 0} \Rightarrow (\mathbf{B} \mathbf{S}_1(t))_{t \geq 0} \quad \text{as } n \rightarrow \infty \text{ in } \mathbb{D}([0, 1], \mathbb{R}^{pd}). \tag{5.9}$$

Consider $\alpha = 2$. Then Proposition A.1(c)–(g) and [22, Corollary 15.16] give (5.9).

(b) Since

$$\tilde{\mathbf{Z}}_{n,k} = \sum_{j=0}^{\infty} \left(\sum_{l=j+1}^{\infty} e^{-\Lambda h_n l} \right) \boldsymbol{\xi}_{n,k-j} = (\mathbf{I}_{d \times d} - e^{-\Lambda h_n})^{-1} e^{-\Lambda h_n} \mathbf{Z}(kh_n),$$

the inequality

$$\mathbb{P}(h_n \|\tilde{\mathbf{Z}}_{n,0}\| > x) \leq \mathbb{P}(2\|\boldsymbol{\Lambda}^{-1}\| \|\mathbf{Z}(0)\| > x) =: g(x) \quad \text{for } x \geq 0$$

holds, where, for $\alpha < 2$, the function $g \in \mathcal{R}_{-\alpha}$ due to [28, Theorem 3.2], such that, by Karamata’s theorem, $\int_0^{\infty} x^{\gamma-1} g(x) \, dx < \infty$ for any $0 < \gamma < \alpha$, and, for $\alpha = 2$, we have $2 \int_0^{\infty} x g(x) \, dx = 8\|\boldsymbol{\Lambda}^{-1}\|^2 \mathbb{E}\|\mathbf{Z}(0)\|^2 < \infty$.

(c) We have $\sum_{k=0}^{\infty} k \|e^{-\Lambda h_n k}\|^\theta < \infty$ for any $\theta > 0$, $n \in \mathbb{N}$, and $\lim_{n \rightarrow \infty} h_n \sum_{k=0}^{\infty} \mathbf{C}_{n,k} = \Lambda^{-1}$.

Part (d) is obviously satisfied since $\boldsymbol{\varepsilon}_{n,k} = 0$.

Proof of Theorem 3.1. (a) Owing to Lemma 5.3, $h_n a_{nh_n}^{-1} \sum_{k=1}^n \mathbf{Z}(kh_n) \Rightarrow \Lambda^{-1} \mathbf{B} \mathbf{S}_1(1)$ as $n \rightarrow \infty$. Hence, by (2.3) as $n \rightarrow \infty$,

$$h_n a_{nh_n}^{-1} \sum_{k=1}^n \mathbf{V}(kh_n) = h_n a_{nh_n}^{-1} \sum_{k=1}^n \mathbf{E} \mathbf{Z}(kh_n) \Rightarrow \mathbf{E} \Lambda^{-1} \mathbf{B} \mathbf{S}_1(1) = \left(\int_0^\infty \mathbf{f}(s) \, ds \right) \mathbf{S}_1(1).$$

(b) Define $\mathbf{g}(s) := e^{-\Lambda s} \mathbf{B} \mathbf{1}_{(0,\infty)}(s)$. It follows from [13, Proposition 2.1] that

$$a_n^{-1} \sum_{k=1}^n \mathbf{Z}(kh) \Rightarrow \mathbf{S}_{\mathbf{g},h}(1) \quad \text{as } n \rightarrow \infty.$$

Thus, $a_n^{-1} \sum_{k=1}^n \mathbf{V}(kh) \Rightarrow \mathbf{E} \mathbf{S}_{\mathbf{g},h} \stackrel{\text{D}}{=} \mathbf{S}_{\mathbf{f},h}(1)$ as $n \rightarrow \infty$.

5.3. Proof of Theorem 3.2

Again, for the proof of Theorem 3.2, we use the following similar result for the state process \mathbf{Z} .

Lemma 5.4. *Let model (1.1)–(1.2) be given with $\mathbf{V} = \mathbf{Z}$ and $\mathbf{A} \in \mathbb{R}^{pd \times v}$, and let the assumptions of Theorem 3.2 hold. Then $\widehat{\mathbf{A}}_n$ as given in (1.6) satisfies, as $n \rightarrow \infty$,*

$$\begin{aligned} & n h_n a_{nh_n}^{-1} b_{nh_n} (\widehat{\mathbf{A}}_n - \mathbf{A}) \\ & \Rightarrow \Lambda^{-1} \mathbf{B} \left(\mathbf{S}_1(1) \mathbf{S}_2(1)^\top - \int_0^1 \mathbf{S}_1(s) \, d\mathbf{S}_2(s)^\top \right) \left(\int_0^1 \mathbf{S}_2(s) \mathbf{S}_2(s)^\top \, ds \right)^{-1}. \end{aligned}$$

In particular, $\widehat{\mathbf{A}}_n \xrightarrow{\text{P}} \mathbf{A}$ as $n \rightarrow \infty$ if $\alpha > \beta/(\beta + 1)$, i.e. $\widehat{\mathbf{A}}_n$ is a consistent estimator.

Proof. We use the same notation as in the proof of Lemma 5.3, only we define $\widetilde{b}_n := b_{nh_n}$ and $\boldsymbol{\varepsilon}_{n,k} := \mathbf{L}_2(kh_n) - \mathbf{L}_2((k - 1)h_n)$. Again, we will show that Assumption 4.1(a)–(d) are satisfied because then the result follows from Theorem 4.2.

(a) If $\alpha < 2$ due to the independence of $(\boldsymbol{\xi}_{n,k})$ and $(\boldsymbol{\varepsilon}_{n,k})$, Proposition A.2 and [36, Theorem 7.1], the limit result

$$\begin{aligned} & \left(\widetilde{a}_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{\xi}_{n,k}^\top, \widetilde{b}_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{\varepsilon}_{n,k}^\top \right)_{t \geq 0} \\ & \Rightarrow (\mathbf{S}_1(t)^\top, \mathbf{S}_2(t)^\top)_{t \geq 0} \quad \text{as } n \rightarrow \infty \text{ in } \mathbb{D}([0, 1], \mathbb{R}^{pd+v}) \end{aligned} \tag{5.10}$$

holds; see also [29]. If $\alpha = 2$, (5.10) follows from Proposition A.1 and [22, Corollary 15.15].

Parts (b) and (c) are satisfied by the proof of Lemma 5.3.

(d) Part (i) follows from Proposition A.2(c) and Proposition A.1(e). Part (ii) follows from Proposition A.2(e) and Proposition A.1(f). Only for $\alpha = 1$ does it follow by symmetry. We obtain (iii) by Proposition A.2(d) and Proposition A.1(d).

Let $\min(\alpha, \beta) < 2$. Then using $\mathbb{E} \|\mathbf{L}_2(h_n)\|^\delta \leq C_1 h_n^{\delta/2}$ and (3.5) gives (iv.1). In the case of a compound Poisson process, Lemma A.2 says that $\mathbb{E} \|\mathbf{L}_2(h_n)\|^\delta \leq C_2 h_n$, such that no additional assumption is necessary.

If $\alpha = \beta = 2$ then $\lim_{n \rightarrow \infty} n(nh_n)^{-2} \mathbb{E} \|\mathbf{L}_2(h_n)\|^2 = \lim_{n \rightarrow \infty} n(nh_n)^{-2} h_n \mathbb{E} \|\mathbf{L}_2(1)\|^2 = 0$; therefore, (iv.2) holds.

Proof of Theorem 3.2. The proof follows the same lines as that of Theorem 3.1 using only Lemma 5.4 and [13, Theorem 3.4].

5.4. Proof of Theorem 3.3

The main idea of the proof is to show that, as $n \rightarrow \infty$,

$$h_n a_{nh_n}^{-2} \sum_{k=1}^n V(kh_n) V(kh_n)^\top = E \sum_{j=0}^\infty e^{-\Lambda h_n j} \left(a_{nh_n}^{-2} \sum_{k=1}^n \xi_{n,k} \xi_{n,k}^\top \right) e^{-\Lambda^\top h_n j} E^\top h_n + o_{\mathbb{P}}(1).$$

The convergence of $a_{nh_n}^{-2} \sum_{k=1}^n \xi_{n,k} \xi_{n,k}^\top$ follows by the limit results of [36, Theorem 7.1] for the heavy-tailed case, and by the law of large numbers for arrays of independent random vectors and the properties of $(\xi_{n,k})_{k \in \mathbb{Z}}$ as given in Appendix A in the light-tailed case. In the same spirit as before we start with the result for Z .

Lemma 5.5. *Let the assumptions of Theorem 3.3 hold. Then, as $n \rightarrow \infty$,*

$$h_n a_{nh_n}^{-2} \sum_{k=1}^n Z(kh_n) Z(kh_n)^\top \Rightarrow \int_0^\infty e^{-\Lambda s} [B S_1, B S_1] e^{-\Lambda^\top s} ds.$$

Proof. A multivariate version of the second-order Beveridge–Nelson decomposition given in [31, Equation (28)] gives the representation

$$\begin{aligned} Z(kh_n) Z(kh_n)^\top &= \sum_{j=0}^\infty e^{-\Lambda h_n j} \xi_{n,k} \xi_{n,k}^\top e^{-\Lambda^\top h_n j} + (F_{n,k-1}^{(1)} - F_{n,k}^{(1)}) + \sum_{r=1}^\infty (F_{n,k,r}^{(2)} + F_{n,k,-r}^{(2)}) \\ &\quad + \sum_{r=1}^\infty (F_{n,k-1,r}^{(3)} + F_{n,k-1,-r}^{(3)} - F_{n,k,r}^{(3)} - F_{n,k,-r}^{(3)}), \end{aligned}$$

where

$$\begin{aligned} F_{n,k}^{(1)} &= \sum_{j=0}^\infty \sum_{s=j+1}^\infty e^{-\Lambda h_n s} \xi_{n,k-j} \xi_{n,k-j}^\top e^{-\Lambda^\top h_n s}, \\ F_{n,k,r}^{(2)} &= \sum_{j=\max(0,-r)}^\infty e^{-\Lambda h_n j} \xi_{n,k} \xi_{n,k-r}^\top e^{-\Lambda^\top h_n (j+r)}, \\ F_{n,k,r}^{(3)} &= \sum_{j=0}^\infty \sum_{s=\max(j+1,-r)}^\infty e^{-\Lambda h_n s} \xi_{n,k-j} \xi_{n,k-j-r}^\top e^{-\Lambda^\top h_n (s+r)}. \end{aligned}$$

Then

$$\begin{aligned} \sum_{k=1}^n Z(kh_n) Z(kh_n)^\top &= \sum_{j=0}^\infty e^{-\Lambda h_n j} \left(\sum_{k=1}^n \xi_{n,k} \xi_{n,k}^\top \right) e^{-\Lambda^\top h_n j} + (F_{n,0}^{(1)} - F_{n,n}^{(1)}) \\ &\quad + \sum_{k=1}^n \sum_{r=1}^\infty (F_{n,k,r}^{(2)} + F_{n,k,-r}^{(2)}) \\ &\quad + \sum_{r=1}^\infty (F_{n,0,r}^{(3)} + F_{n,0,-r}^{(3)} - F_{n,n,r}^{(3)} - F_{n,n,-r}^{(3)}) \\ &=: J_{n,1} + J_{n,2} + J_{n,3} + J_{n,4}. \end{aligned} \tag{5.11}$$

Step 1. Let $\alpha \in (0, 2)$, and assume that L_1 is a compound Poisson process as given in (A.5) with characteristic triplet $(\mathbf{0}_m, \mathbf{0}_{m \times m}, \nu_{L_1})$. On the one hand, by Lemma 5.6 below we have, for $i = 2, 3, 4$,

$$h_n a_{nh_n}^{-2} J_{n,i} \xrightarrow{P} \mathbf{0}_{pd \times pd} \quad \text{as } n \rightarrow \infty. \tag{5.12}$$

On the other hand, by Proposition A.2(a), (c) and [36, Theorem 7.1], we have

$$S_n := a_{nh_n}^{-2} \sum_{k=1}^n \xi_{n,k} \xi_{n,k}^\top \Rightarrow [BS_1, BS_1]_1 \quad \text{as } n \rightarrow \infty.$$

We denote by g_n and g maps from $M_{pd \times pd}(\mathbb{R}) \rightarrow M_{pd \times pd}(\mathbb{R})$ with

$$g_n(C) = h_n \sum_{j=0}^\infty e^{-\Lambda h_n j} C e^{-\Lambda^\top h_n j} \quad \text{and} \quad g(C) = \int_0^\infty e^{-\Lambda s} C e^{-\Lambda^\top s} ds. \tag{5.13}$$

Since g_n and g are continuous with $\lim_{n \rightarrow \infty} g_n(C_n) = g(C)$ for any sequence $C_n, C \in M_{pd \times pd}(\mathbb{R})$ with $\lim_{n \rightarrow \infty} \|C_n - C\| = 0$, we can apply a generalized version of the continuous mapping theorem (cf. [36, Theorem 3.1]) to obtain $g_n(S_n) \Rightarrow g([BS_1, BS_1]_1)$ as $n \rightarrow \infty$, which means that, as $n \rightarrow \infty$,

$$h_n a_{nh_n}^{-2} J_{n,1} = g_n(S_n) \Rightarrow g([BS_1, BS_1]_1) = \int_0^\infty e^{-\Lambda s} [BS_1, BS_1]_1 e^{-\Lambda^\top s} ds. \tag{5.14}$$

Then the result follows by (5.11)–(5.14).

Step 2. Let $\alpha \in (0, 2)$, and let L_1 be some Lévy process. We use the decompositions $L_1 = L_1^{(1)} + L_1^{(2)}$ and $\xi_{n,k} = \xi_{n,k}^{(1)} + \xi_{n,k}^{(2)}$ as given in (A.3) and (A.4) such that

$$\begin{aligned} Z(t) &= \int_{-\infty}^t e^{-\Lambda(t-s)} B dL^{(1)}(s) + \int_{-\infty}^t e^{-\Lambda(t-s)} B dL^{(2)}(s) \\ &=: Z_1(t) + Z_2(t) \quad \text{for } t \geq 0 \end{aligned}$$

and

$$\begin{aligned} \sum_{k=1}^n Z(kh_n) Z(kh_n)^\top &= \sum_{k=1}^n Z_1(kh_n) Z_1(kh_n)^\top + \sum_{k=1}^n Z_1(kh_n) Z_2(kh_n)^\top \\ &\quad + \sum_{k=1}^n Z_2(kh_n) Z_1(kh_n)^\top + \sum_{k=1}^n Z_2(kh_n) Z_2(kh_n)^\top \\ &=: I_{n,1} + I_{n,2} + I_{n,3} + I_{n,4}. \end{aligned} \tag{5.15}$$

Applying step 1 we obtain, as $n \rightarrow \infty$,

$$h_n a_{nh_n}^{-2} \sum_{k=1}^n Z_1(kh_n) Z_1(kh_n)^\top \Rightarrow \int_0^\infty e^{-\Lambda s} [BS_1, BS_1]_1 e^{-\Lambda^\top s} ds. \tag{5.16}$$

Furthermore, Hölder’s inequality results in the decomposition

$$\begin{aligned} h_n a_{nh_n}^{-2} \max(\|I_{n,2}\|, \|I_{n,3}\|) &\leq \left(h_n a_{nh_n}^{-2} \sum_{k=1}^n \|Z_1(kh_n)\|^2 \right)^{1/2} \left(h_n a_{nh_n}^{-2} \sum_{k=1}^n \|Z_2(kh_n)\|^2 \right)^{1/2} \end{aligned} \tag{5.17}$$

of independent factors. Now we use the fact that $L_1^{(1)}$ has representation (A.5) and we define

$$L^*(t) := \|B\| \sum_{k=1}^{N(t)} \|J_k\|, \quad \xi_{n,k}^* := \int_{(k-1)h_n}^{kh_n} e^{-\lambda(kh_n-s)} dL^*(s), \tag{5.18}$$

$$Z^*(t) := \int_{-\infty}^t e^{-\lambda(t-s)} dL^*(s).$$

Hence, $\|BL_1^{(1)}(t)\| \leq L^*(t)$, $\|\xi_{n,k}^{(1)}\| \leq \xi_{n,k}^*$, and $\|Z_1(t)\| \leq Z^*(t)$. Then a consequence of step 1 is

$$h_n a_{nh_n}^{-2} \sum_{k=1}^n \|Z_1(kh_n)\|^2 \leq h_n a_{nh_n}^{-2} \sum_{k=1}^n Z^*(kh_n)^2 \Rightarrow \frac{1}{2\lambda} [S, S]_1 \quad \text{as } n \rightarrow \infty, \tag{5.19}$$

where S is an α -stable Lévy process. Since $\lim_{n \rightarrow \infty} h_n a_{nh_n}^{-2} \sum_{k=1}^n \mathbb{E} \|Z_2(kh_n)\|^2 = 0$, we obtain

$$\left(h_n a_{nh_n}^{-2} \sum_{k=1}^n \|Z_2(kh_n)\|^2 \right)^{1/2} \xrightarrow{P} 0 \quad \text{as } n \rightarrow \infty. \tag{5.20}$$

Hence, (5.17)–(5.20) give $h_n a_{nh_n}^{-2} \|I_{n,2}\| \xrightarrow{P}$ and $h_n a_{nh_n}^{-2} \|I_{n,3}\| \xrightarrow{P} 0$ as $n \rightarrow \infty$. It follows from (5.20) that $h_n a_{nh_n}^{-2} \|I_{n,4}\| \xrightarrow{P} 0$ as $n \rightarrow \infty$ as well. Finally, the result follows by (5.15) and (5.16).

Step 3. Let $\alpha = 2$. On the one hand, by Lemma 5.7 below we have, for $i = 2, 3, 4$, as $n \rightarrow \infty$,

$$h_n a_{nh_n}^{-2} J_{n,i} \xrightarrow{P} \mathbf{0}_{pd \times pd}.$$

On the other hand, by Proposition A.1(g), as $n \rightarrow \infty$,

$$S_n := a_{nh_n}^{-2} \sum_{k=1}^n \xi_{n,k} \xi_{n,k}^\top \xrightarrow{P} B \Sigma_1 B^\top = [BS_1, BS_1]_1.$$

The same arguments as in step 1 complete the proof.

Lemma 5.6. *Let the assumptions of Lemma 5.5 hold with $\alpha \in (0, 2)$, and suppose that L_1 is a compound Poisson process as given in (A.5) with characteristic triplet $(\mathbf{0}_m, \mathbf{0}_{m \times m}, \nu_{L_1})$.*

- (a) $F_{n,0}^{(1)} \stackrel{D}{=} F_{n,n}^{(1)}$ and $h_n a_{nh_n}^{-2} F_{n,0}^{(1)} \xrightarrow{P} \mathbf{0}_{pd \times pd}$ as $n \rightarrow \infty$.
- (b) $h_n a_{nh_n}^{-2} \sum_{k=1}^n \sum_{r=1}^\infty (F_{n,k,r}^{(2)} + F_{n,k,-r}^{(2)}) \xrightarrow{P} \mathbf{0}_{pd \times pd}$ as $n \rightarrow \infty$.
- (c) $h_n a_{nh_n}^{-2} \sum_{r=1}^\infty (F_{n,0,r}^{(3)} + F_{n,0,-r}^{(3)} - F_{n,n,r}^{(3)} - F_{n,n,-r}^{(3)}) \xrightarrow{P} \mathbf{0}_{pd \times pd}$ as $n \rightarrow \infty$.

Proof. (a) We use the notation given in (5.18). Then

$$\|F_{n,0}^{(1)}\| \leq (1 - e^{-2\lambda h_n})^{-1} \sum_{j=0}^\infty e^{-2\lambda h_n j} \xi_{n,-j}^{*2} \leq (1 - e^{-2\lambda h_n})^{-1} Z^*(0)^2.$$

Hence, $\mathbb{P}(\|F_{n,0}^{(1)}\| > a_{nh_n}^2 h_n^{-1}) \leq \mathbb{P}(Z^*(0)^2 > C_1 a_{nh_n}^2) \rightarrow 0$ as $n \rightarrow \infty$.

(b) The upper bound

$$\left\| \sum_{k=1}^n \sum_{r=1}^{\infty} \mathbf{F}_{n,k,r}^{(2)} \right\| \leq \sum_{k=1}^n \sum_{j=0}^{\infty} e^{-\lambda h_n(2j+1)} \xi_{n,k}^* \left(\sum_{r=0}^{\infty} \xi_{n,k-1-r}^* e^{-\lambda h_n r} \right) \tag{5.21}$$

holds. Applying Lemma 5.1 (here we require that, for a compound Poisson process, $\mathbb{E}\|\xi_{n,0}^*\|^\delta \leq C_2 h_n$ by Lemma A.2, which is used to show Lemma 5.1(a) for some $0 < \delta < 1$, $\delta < \alpha$, and $2\delta > \alpha$) gives

$$\begin{aligned} \left\| \sum_{k=1}^n \sum_{r=1}^{\infty} \mathbf{F}_{n,k,r}^{(2)} \right\| &\leq h_n (1 - e^{-2\lambda h_n})^{-1} a_{nh_n}^{-2} \sum_{k=1}^n \xi_{n,k}^* Z^*((k-1)h_n) \\ &\xrightarrow{p} 0 \quad \text{as } n \rightarrow \infty. \end{aligned} \tag{5.22}$$

On the other hand, if we define $W^*(kh_n) := \sum_{r=0}^{\infty} e^{-\lambda h_n r} \xi_{n,k+r}^*$ then $W^*(kh_n) \stackrel{D}{=} Z^*(0)$ and

$$\left\| \sum_{k=1}^n \sum_{r=1}^{\infty} \mathbf{F}_{n,k,-r}^{(2)} \right\| \leq (1 - e^{-2\lambda h_n})^{-1} \sum_{k=1}^n \xi_{n,k}^* W^*((k+1)h_n). \tag{5.23}$$

Again, using Lemma 5.1 yields

$$h_n (1 - e^{-2\lambda h_n})^{-1} a_{nh_n}^{-2} \sum_{k=1}^n \xi_{n,k}^* W^*((k+1)h_n) \xrightarrow{p} 0 \quad \text{as } n \rightarrow \infty. \tag{5.24}$$

Hence, (5.21)–(5.24) give the statement.

(c) We will show that, as $n \rightarrow \infty$,

$$h_n a_{nh_n}^{-2} \sum_{r=1}^{\infty} \mathbf{F}_{n,0,r}^{(3)} \xrightarrow{p} \mathbf{0}_{pd \times pd} \quad \text{and} \quad h_n a_{nh_n}^{-2} \sum_{r=1}^{\infty} \mathbf{F}_{n,0,-r}^{(3)} \xrightarrow{p} \mathbf{0}_{pd \times pd}.$$

Since $\sum_{r=1}^{\infty} \mathbf{F}_{n,0,r}^{(3)} \stackrel{D}{=} \sum_{r=1}^{\infty} \mathbf{F}_{n,n,r}^{(3)}$ and $\sum_{r=1}^{\infty} \mathbf{F}_{n,0,-r}^{(3)} \stackrel{D}{=} \sum_{r=1}^{\infty} \mathbf{F}_{n,n,-r}^{(3)}$, the proof will then be finished. Again, we use the notation given in (5.18). For the first term, we derive the upper bound

$$\left\| \sum_{r=1}^{\infty} \mathbf{F}_{n,0,r}^{(3)} \right\| \leq (1 - e^{-2\lambda h_n})^{-1} \sum_{j=0}^{\infty} e^{-2\lambda h_n j} \xi_{n,-j}^* Z^*((-j-1)h_n).$$

Applying, for $0 < \delta < \alpha$, $\delta \leq 1$,

$$\mathbb{E} \left(\left(\sum_{j=0}^{\infty} e^{-2\lambda h_n j} \xi_{n,-j}^* Z^*((-j-1)h_n) \right)^\delta \right) \leq \sum_{j=0}^{\infty} e^{-2\delta \lambda h_n j} \mathbb{E}(\xi_{n,0}^* \delta) \mathbb{E}(Z^*(0)^\delta),$$

where we have used the independence of $\xi_{n,-j}^*$ and $Z^*((-j-1)h_n)$, and Lemma A.2 results in

$$h_n^\delta a_{nh_n}^{-2\delta} \mathbb{E} \left\| \sum_{r=1}^{\infty} \mathbf{F}_{n,0,r}^{(3)} \right\|^\delta \leq C_3 a_{nh_n}^{-2\delta} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

For the second term, we have the upper bound

$$\begin{aligned} \left\| \sum_{r=1}^{\infty} \mathbf{F}_{n,0,-r}^{(3)} \right\| &\leq (1 - e^{-2\lambda h_n})^{-1} \sum_{j=0}^{\infty} \sum_{r=1}^j e^{-2\lambda h_n(j+1)} \xi_{n,-j}^* \xi_{n,-j+r}^* e^{\lambda h_n r} \\ &\quad + (1 - e^{-2\lambda h_n})^{-1} \sum_{j=0}^{\infty} \sum_{r=j+1}^{\infty} e^{-2\lambda h_n r} \xi_{n,-j}^* \xi_{n,-j+r}^* e^{\lambda h_n r} \\ &=: I_{n,1} + I_{n,2}. \end{aligned} \tag{5.25}$$

Moreover,

$$I_{n,1} \leq (1 - e^{-2\lambda h_n})^{-1} \mathbf{Z}^*(0)^2 \tag{5.26}$$

and

$$I_{n,2} \stackrel{D}{=} (1 - e^{-2\lambda h_n})^{-1} e^{-\lambda h_n} \mathbf{Z}^*(0) \tilde{\mathbf{Z}}(0), \tag{5.27}$$

where $\tilde{\mathbf{Z}}(0)$ is an independent copy of $\mathbf{Z}^*(0)$. We conclude from (5.25)–(5.27) that, for any $\epsilon > 0$,

$$\mathbb{P} \left(h_n a_{nh_n}^{-2} \left\| \sum_{r=1}^{\infty} \mathbf{F}_{n,0,-r}^{(3)} \right\| > \epsilon \right) \leq \mathbb{P}(\mathbf{Z}^*(0)^2 + \mathbf{Z}^*(0) \tilde{\mathbf{Z}}(0) > C_4 a_{nh_n}^2) \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

as required.

Lemma 5.7. *Suppose that the assumptions of Lemma 5.5 hold with $\alpha = 2$.*

- (a) $\mathbf{F}_{n,0}^{(1)} \stackrel{D}{=} \mathbf{F}_{n,n}^{(1)}$ and $h_n a_{nh_n}^{-2} \mathbf{F}_{n,0}^{(1)} \xrightarrow{P} \mathbf{0}_{pd \times pd}$ as $n \rightarrow \infty$.
- (b) $h_n a_{nh_n}^{-2} \sum_{k=1}^n \sum_{r=1}^{\infty} (\mathbf{F}_{n,k,r}^{(2)} + \mathbf{F}_{n,k,-r}^{(2)}) \xrightarrow{P} \mathbf{0}_{pd \times pd}$ as $n \rightarrow \infty$.
- (c) $h_n a_{nh_n}^{-2} \sum_{r=1}^{\infty} (\mathbf{F}_{n,0,r}^{(3)} + \mathbf{F}_{n,0,-r}^{(3)} - \mathbf{F}_{n,n,r}^{(3)} - \mathbf{F}_{n,n,-r}^{(3)}) \xrightarrow{P} \mathbf{0}_{pd \times pd}$ as $n \rightarrow \infty$.

Proof. (a) Rewrite

$$\mathbf{F}_{n,0}^{(1)} = \sum_{s=0}^{\infty} e^{-\Lambda h_n(s+1)} \left(\sum_{j=0}^{\infty} e^{-\Lambda h_n j} \xi_{n,-j} \xi_{n,-j}^{\top} e^{-\Lambda^{\top} h_n j} \right) e^{-\Lambda^{\top} h_n(s+1)}.$$

With \mathbf{g}_n and \mathbf{g} as defined in (5.13) and $\mathbf{S}_n := a_{nh_n}^{-2} \sum_{j=0}^{\infty} e^{-\Lambda h_n j} \xi_{n,-j} \xi_{n,-j}^{\top} e^{-\Lambda^{\top} h_n j}$, the equality $h_n a_{nh_n}^{-2} \mathbf{F}_{n,0}^{(1)} = e^{-\Lambda h_n} \mathbf{g}_n(\mathbf{S}_n) e^{-\Lambda^{\top} h_n}$ holds. If we are able to prove that $\mathbf{S}_n \xrightarrow{P} \mathbf{0}_{pd \times pd}$ as $n \rightarrow \infty$ then, using a generalized continuous mapping theorem (the same arguments as in the proof of Lemma 5.5), we can conclude that $h_n a_{nh_n}^{-2} \mathbf{F}_{n,0}^{(1)} \xrightarrow{P} \mathbf{0}_{pd \times pd}$ as $n \rightarrow \infty$. Finally, due to Proposition A.1(a),

$$\mathbb{E} \|\mathbf{S}_n\| \leq a_{nh_n}^{-2} \sum_{j=0}^{\infty} e^{-2\lambda h_n j} \mathbb{E} \|\xi_{n,0}\|^2 \leq C_1 a_{nh_n}^{-2} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

and, hence, $\mathbf{S}_n \xrightarrow{P} \mathbf{0}_{pd \times pd}$ as $n \rightarrow \infty$.

(b) The representation

$$h_n a_{nh_n}^{-2} \sum_{k=1}^n \sum_{r=1}^{\infty} F_{n,k,r}^{(2)} = \sum_{j=0}^{\infty} e^{-\Lambda h_n j} \left(a_{nh_n}^{-2} \sum_{k=1}^n \xi_{n,k} \mathbf{Z}((k-1)h_n)^\top \right) e^{-\Lambda^\top h_n(j+1)h_n}$$

holds. Using the same arguments as in (a), it is sufficient to prove that, as $n \rightarrow \infty$,

$$a_{nh_n}^{-2} \sum_{k=1}^n \xi_{n,k} \mathbf{Z}((k-1)h_n)^\top \xrightarrow{P} \mathbf{0}_{pd \times pd}.$$

However, this follows from Proposition A.1 and Lemma 5.1. Similarly,

$$h_n a_{nh_n}^{-2} \sum_{k=1}^n \sum_{r=1}^{\infty} F_{n,k,-r}^{(2)} = \sum_{j=0}^{\infty} e^{-\Lambda^\top h_n j} \left(a_{nh_n}^{-2} \sum_{k=1}^n \sum_{r=1}^{\infty} e^{-\Lambda h_n r} \xi_{n,k} \xi_{n,k+r}^\top \right) e^{-\Lambda^\top h_n j} h_n$$

and, hence, it is sufficient to show that

$$a_{nh_n}^{-2} \sum_{k=1}^n \sum_{r=1}^{\infty} e^{-\Lambda h_n r} \xi_{n,k} \xi_{n,k+r}^\top \xrightarrow{P} \mathbf{0}_{pd \times pd}. \tag{5.28}$$

We prove the convergence of (5.28) componentwise. The sequence of (l, m) -components $((e^{-\Lambda h_n r} \xi_{n,k} \xi_{n,k+r}^\top)_{(l,m)})_{k,r \in \mathbb{N}}$ is uncorrelated, giving

$$\mathbb{E} \left(\left(\sum_{k=1}^n \sum_{r=1}^{\infty} e^{-\Lambda h_n r} \xi_{n,k} \xi_{n,k+r}^\top \right)_{(l,m)} \right)^2 \leq C_2 \sum_{k=1}^n \sum_{r=1}^{\infty} e^{-2\lambda h_n r} (\mathbb{E} \|\xi_{n,0}\|^2)^2 \leq C_3 n h_n.$$

Thus, (5.28) holds.

(c) Let us start with

$$\begin{aligned} & h_n a_{nh_n}^{-2} \sum_{r=1}^{\infty} F_{n,0,r}^{(3)} \\ &= \sum_{s=0}^{\infty} e^{-\Lambda h_n(s+1)} \left(a_{nh_n}^{-2} \sum_{j=0}^{\infty} e^{-\Lambda h_n j} \xi_{n,-j} \mathbf{Z}((-j-1)h_n)^\top e^{-\Lambda^\top h_n j} \right) e^{-\Lambda^\top h_n(s+2)h_n}. \end{aligned}$$

As before, it is sufficient to show that

$$a_{nh_n}^{-2} \sum_{j=0}^{\infty} e^{-\Lambda h_n j} \xi_{n,-j} \mathbf{Z}((-j-1)h_n)^\top e^{-\Lambda^\top h_n j} \xrightarrow{P} \mathbf{0}_{pd \times pd} \quad \text{as } n \rightarrow \infty,$$

which we prove componentwise using the uncorrelation of the sequence of the (l, m) -components $((\xi_{n,-j} \mathbf{Z}((-j-1)h_n)^\top)_{(l,m)})_{j \in \mathbb{N}}$. For the (l, m) -component, we have

$$\begin{aligned} & a_{nh_n}^{-4} \mathbb{E} \left(\left(\sum_{j=0}^{\infty} e^{-\Lambda h_n j} \xi_{n,-j} \mathbf{Z}((-j-1)h_n)^\top e^{-\Lambda^\top h_n j} \right)_{(l,m)} \right)^2 \\ & \leq a_{nh_n}^{-4} C_4 \sum_{j=0}^{\infty} e^{-4\lambda h_n j} \mathbb{E} \|\xi_{n,0}\|^2 \mathbb{E} \|\mathbf{Z}(0)\|^2 \\ & \leq C_5 a_{nh_n}^{-4} \rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned}$$

We now investigate

$$\begin{aligned} \sum_{r=1}^{\infty} F_{n,0,-r}^{(3)} &= \sum_{j=0}^{\infty} \sum_{r=1}^j \sum_{s=j+1}^{\infty} e^{-\Lambda h_n s} \xi_{n,-j} \xi_{n,-j+r}^{\top} e^{-\Lambda^{\top} h_n (s-r)} \\ &\quad + \sum_{j=0}^{\infty} \sum_{r=j+1}^{\infty} \sum_{s=r}^{\infty} e^{-\Lambda h_n s} \xi_{n,-j} \xi_{n,-j+r}^{\top} e^{-\Lambda^{\top} h_n (s-r)} \\ &=: I_{n,1} + I_{n,2}. \end{aligned}$$

Then

$$I_{n,1} = \sum_{s=0}^{\infty} e^{-\Lambda h_n (s+1)} \left(\sum_{j=0}^{\infty} e^{-\Lambda h_n j} \xi_{n,-j} \left(\sum_{r=1}^j \xi_{n,-j+r}^{\top} e^{-\Lambda^{\top} h_n (j-r)} \right) \right) e^{-\Lambda^{\top} h_n (s+1)}.$$

For the convergence $h_n a_{nh_n}^{-2} I_{n,1} \xrightarrow{p} \mathbf{0}_{pd \times pd}$ as $n \rightarrow \infty$, it is again sufficient to show that

$$a_{nh_n}^{-2} \sum_{j=0}^{\infty} e^{-\Lambda h_n j} \xi_{n,-j} \left(\sum_{u=0}^{j-1} \xi_{n,-u}^{\top} e^{-\Lambda^{\top} h_n u} \right) \xrightarrow{p} \mathbf{0}_{pd \times pd} \quad \text{as } n \rightarrow \infty, \tag{5.29}$$

which we will prove componentwise. By uncorrelation we obtain

$$\begin{aligned} &\mathbb{E} \left(\sum_{j=0}^{\infty} e^{-\Lambda h_n j} \xi_{n,-j} \left(\sum_{u=0}^{j-1} \xi_{n,-u}^{\top} e^{-\Lambda^{\top} h_n u} \right)_{(l,m)} \right)^2 \\ &= \sum_{j=0}^{\infty} \mathbb{E} \left(e^{-\Lambda h_n j} \xi_{n,-j} \left(\sum_{u=0}^{j-1} \xi_{n,-u}^{\top} e^{-\Lambda^{\top} h_n u} \right)_{(l,m)} \right)^2. \end{aligned} \tag{5.30}$$

Furthermore, by Proposition A.1(a) we obtain

$$\mathbb{E} \| e^{-\Lambda h_n j} \xi_{n,-j} \|^2 \leq C_6 e^{-2\lambda h_n j} \mathbb{E} \| \xi_{n,0} \|^2 \leq C_7 h_n e^{-2\lambda h_n j} \tag{5.31}$$

and

$$\mathbb{E} \left\| \sum_{u=0}^{j-1} \xi_{n,-u}^{\top} e^{-\Lambda^{\top} h_n u} \right\|^2 \leq C_8 \sum_{u=0}^{j-1} e^{-2\lambda h_n u} \mathbb{E} \| \xi_{n,0} \|^2 \leq C_9. \tag{5.32}$$

Hence, (5.30)–(5.32) and the independence of $e^{-\Lambda h_n j} \xi_{n,-j}$ and $\sum_{u=0}^{j-1} \xi_{n,-u}^{\top} e^{-\Lambda^{\top} h_n u}$ give

$$\mathbb{E} \left(\sum_{j=0}^{\infty} e^{-\Lambda h_n j} \xi_{n,-j} \left(\sum_{u=0}^{j-1} \xi_{n,-u}^{\top} e^{-\Lambda^{\top} h_n u} \right)_{(l,m)} \right)^2 \leq C_{10} h_n \sum_{j=0}^{\infty} e^{-2\lambda h_n j} \leq C_{11},$$

which results in (5.29).

Next we have to show that $h_n a_{nh_n}^{-2} I_{n,2} \xrightarrow{p} \mathbf{0}_{pd \times pd}$ as $n \rightarrow \infty$. Therefore, we use the representation

$$I_{n,2} = \sum_{s=0}^{\infty} e^{-\Lambda h_n s} \left(\sum_{j=0}^{\infty} \sum_{r=j+1}^{\infty} e^{-\Lambda h_n r} \xi_{n,-j} \xi_{n,-j+r}^{\top} \right) e^{-\Lambda^{\top} h_n s}$$

and prove that, as $n \rightarrow \infty$,

$$a_n^{-2} \sum_{j=0}^{\infty} \sum_{r=j+1}^{\infty} e^{-\Lambda h_n r} \xi_{n,-j} \xi_{n,-j+r}^\top = a_n^{-2} \sum_{j=0}^{\infty} e^{-\Lambda h_n j} \sum_{u=1}^{\infty} e^{-\Lambda h_n u} \xi_{n,-j} \xi_{n,u}^\top \xrightarrow{p} \mathbf{0}_{pd \times pd}. \tag{5.33}$$

By the uncorrelation of the components of $((e^{-\Lambda h_n j} e^{-\Lambda h_n u} \xi_{n,-j} \xi_{n,u}^\top)_{(l,m)})_{j,u \in \mathbb{N}}$ we obtain similarly as above

$$\begin{aligned} & \mathbb{E} \left(\left(\sum_{j=0}^{\infty} e^{-\Lambda h_n j} \sum_{u=1}^{\infty} e^{-\Lambda h_n u} \xi_{n,-j} \xi_{n,u}^\top \right)_{(l,m)}^2 \right) \\ & \leq C_{12} \sum_{j=0}^{\infty} e^{-2\lambda h_n j} \sum_{u=0}^{\infty} e^{-2\lambda h_n u} \mathbb{E} \|\xi_{n,-j}\|^2 \mathbb{E} \|\xi_{n,u}\|^2 \\ & \leq C_{13}. \end{aligned}$$

Putting all these elements together we obtain (5.33) and $h_n a_n^{-2} I_{n,2} \xrightarrow{p} \mathbf{0}_{pd \times pd}$ as $n \rightarrow \infty$.

Finally, we are able to prove the main statement in Theorem 3.3.

Proof of Theorem 3.3. (a) The observation equation (2.3) and Lemma 5.5 yield

$$\begin{aligned} h_n a_n^{-2} \sum_{k=1}^n \mathbf{V}(kh_n) \mathbf{V}(kh_n)^\top & \Rightarrow \int_0^\infty \mathbf{E} e^{-\Lambda s} \mathbf{B} [\mathbf{S}_1, \mathbf{S}_1]_1 \mathbf{B}^\top e^{-\Lambda^\top s} \mathbf{E}^\top ds \\ & = \int_0^\infty \mathbf{f}(s) [\mathbf{S}_1, \mathbf{S}_1]_1 \mathbf{f}(s)^\top ds \quad \text{as } n \rightarrow \infty. \end{aligned}$$

(b) An application of [13, Proposition 2.1] gives, with $\mathbf{g}(s) = e^{-\Lambda s} \mathbf{B} \mathbf{1}_{(0,\infty)}(s)$,

$$a_n^{-2} \sum_{k=1}^n \mathbf{Z}(kh) \mathbf{Z}(kh)^\top \Rightarrow [\mathbf{S}_{\mathbf{g},h}, \mathbf{S}_{\mathbf{g},h}]_1 \quad \text{as } n \rightarrow \infty,$$

resulting in

$$a_n^{-2} \sum_{k=1}^n \mathbf{V}(kh) \mathbf{V}(kh)^\top \Rightarrow \mathbf{E} [\mathbf{S}_{\mathbf{g},h}, \mathbf{S}_{\mathbf{g},h}]_1 \mathbf{E}^\top = [\mathbf{E} \mathbf{S}_{\mathbf{g},h}, \mathbf{E} \mathbf{S}_{\mathbf{g},h}]_1 \stackrel{D}{=} [\mathbf{S}_{\mathbf{f},h}, \mathbf{S}_{\mathbf{f},h}]_1.$$

Appendix A. Asymptotic behavior of stochastic integrals

In this appendix we present the tail behavior and extensions of Karamata’s theorem to stochastic integrals of the form $\int_0^{h_n} \mathbf{f}(s) d\mathbf{L}(s)$ where $h_n \downarrow 0$ as $n \rightarrow \infty$. First, we start with a driving Lévy process which has a finite second moment. In the subsequent subsection the driving Lévy process has a regularly varying tail.

A.1. Finite second moments

Proposition A.1. *Let $(\mathbf{L}(t))_{t \geq 0}$ be an \mathbb{R}^d -valued Lévy process with $\mathbb{E} \|\mathbf{L}(1)\|^2 < \infty$ and $\mathbb{E}(\mathbf{L}(1)\mathbf{L}(1)^\top) = \Sigma$. Suppose that $(h_n)_{n \in \mathbb{N}}$ is a sequence of positive constants such that $h_n \downarrow 0$ and $\lim_{n \rightarrow \infty} nh_n = \infty$. Moreover, let $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^{m \times d}$ be a measurable and bounded function with $\lim_{x \rightarrow 0} \mathbf{f}(x) = \mathbf{f}(0)$. Define $\xi_n = \int_0^{h_n} \mathbf{f}(s) d\mathbf{L}(s)$ for $n \in \mathbb{N}$. Finally, let $\delta \in (0, 2]$ and $x > 0$.*

- (a) *There exists a finite positive constant K such that $h_n^{-1} \mathbb{E} \|\xi_n\|^2 \leq K$ for all $n \in \mathbb{N}$. If $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_d$ then $\lim_{n \rightarrow \infty} h_n^{-1} \mathbb{E} \|\xi_n\|^2 = \mathbb{E} \|\mathbf{f}(0)\mathbf{L}(1)\|^2$.*

- (b) If $\mathbb{E}\|\mathbf{L}(1)\|^4 < \infty$ then there exists a finite positive constant K such that $\mathbb{E}\|\xi_n\|^4 \leq Kh_n$ for all $n \in \mathbb{N}$.
- (c) $n\mathbb{P}((nh_n)^{-1/2}\xi_n \in \cdot) \xrightarrow{v} 0$ as $n \rightarrow \infty$ on $\mathcal{B}(\overline{\mathbb{R}}^m \setminus \{\mathbf{0}_m\})$.
- (d) $\lim_{n \rightarrow \infty} n(nh_n)^{-\delta/2} \mathbb{E}(\|\xi_n\|^\delta \mathbf{1}_{\{\|\xi_n\| > \sqrt{nh_n x}\}}) = 0$.
- (e) There exists a finite positive constant K such that

$$h_n^{-1} \mathbb{E}(\|\xi_n\|^2 \mathbf{1}_{\{\|\xi_n\| \leq \sqrt{nh_n x}\}}) \leq K \quad \text{for all } n \in \mathbb{N}.$$

If $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_d$ then $\lim_{n \rightarrow \infty} h_n^{-1} \mathbb{E}(\|\xi_n\|^2 \mathbf{1}_{\{\|\xi_n\| \leq \sqrt{nh_n x}\}}) = \mathbb{E}\|\mathbf{f}(0)\mathbf{L}(1)\|^2$.

- (f) Let $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_d$. Then $\lim_{n \rightarrow \infty} n(nh_n)^{-1/2} \mathbb{E}(\xi_n \mathbf{1}_{\{\|\xi_n\| \leq \sqrt{nh_n x}\}}) = \mathbf{0}_{m \times m}$.
- (g) Let $(\xi_{n,k})_{k \in \mathbb{N}}$ be an i.i.d. sequence with $\xi_{n,1} \stackrel{d}{=} \xi_n$ for any $n \in \mathbb{N}$ and $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_d$. Then

$$(nh_n)^{-1} \sum_{k=1}^n \xi_{n,k} \xi_{n,k}^\top \xrightarrow{\mathbb{P}} \mathbf{f}(0) \Sigma \mathbf{f}(0)^\top \quad \text{as } n \rightarrow \infty.$$

Proof. (a) Suppose that $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_d$. Owing to Equation (2.10) of [25], the covariance matrix of ξ_n is $\int_0^{h_n} \mathbf{f}(s) \Sigma \mathbf{f}(s)^\top ds$. Hence, we obtain, as $n \rightarrow \infty$,

$$\begin{aligned} \mathbb{E}\|\xi_n\|^2 &= \int_0^{h_n} \|\text{diag}(\mathbf{f}(s) \Sigma \mathbf{f}(s)^\top)\|^2 ds \\ &\sim h_n \|\text{diag}(\mathbf{f}(0) \Sigma \mathbf{f}(0)^\top)\|^2 \\ &= h_n \mathbb{E}\|\mathbf{f}(0)\mathbf{L}(1)\|^2, \end{aligned} \tag{A.1}$$

where $\text{diag}(B)$ denotes the vector containing the diagonal elements of B .

Suppose that $\mathbb{E}(\mathbf{L}(1)) \neq \mathbf{0}_d$. Then $\mathbb{E}\|\xi_n\|^2 = \text{Var}(\|\xi_n\|) + (\mathbb{E}\|\xi_n\|)^2$. The proof of (a) then follows from (A.1).

(b) Suppose that $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_d$. The characteristic function of $\xi(t) = \int_0^t \mathbf{f}(s) ds =: (\xi_1(t), \dots, \xi_m(t))^\top$ is $\mathbb{E}(e^{i\Theta^\top \xi(t)}) = \exp(-\Psi_{f,t}(\Theta))$ for $\Theta \in \mathbb{R}^m$, where $\Psi_{f,t}(\Theta) = \int_0^t \Psi(\Theta^\top \mathbf{f}(s)) ds$ (cf. [33, Proposition 2.6]). Hence, for $k = 1, \dots, m$ and $\mathbf{e}_k = (0, \dots, 0, 1, 0, \dots, 0)^\top \in \mathbb{R}^m$,

$$\begin{aligned} \mathbb{E}|\xi_k(t)|^4 &= \frac{d}{d^4 \theta_k} \mathbb{E}(e^{i\theta_k^\top \xi(t)}) \Big|_{\theta_k=0} \\ &= 3 \left(\frac{d}{d^2 \theta_k} \Psi_{f,t}(\theta \mathbf{e}_k) \Big|_{\theta_k=0} \right)^2 - \left(\frac{d}{d^4 \theta_k} \Psi_{f,t}(\theta \mathbf{e}_k) \Big|_{\theta_k=0} \right) \\ &= 3 \left(\int_0^t \left(\frac{d}{d^2 \theta_k} \Psi(\theta \mathbf{e}_k^\top \mathbf{f}(s)) \right) \Big|_{\theta_k=0} ds \right)^2 - \int_0^t \left(\frac{d}{d^4 \theta_k} \Psi(\theta \mathbf{e}_k^\top \mathbf{f}(s)) \right) \Big|_{\theta_k=0} ds \\ &\sim 3t^2 C_2 + t C_3 \quad \text{as } t \rightarrow 0. \end{aligned}$$

Finally,

$$\mathbb{E}\|\xi_n\|^4 \leq C_4 \sum_{k=1}^m \mathbb{E}|\xi_k(h_n)|^4 \leq C_5 h_n \quad \text{for all } n \in \mathbb{N}. \tag{A.2}$$

Suppose that $\mathbb{E}(\mathbf{L}(1)) \neq \mathbf{0}_d$. Then, by (A.2), $\mathbb{E}\|\xi_n\|^4 \leq 8\mathbb{E}\|\int_0^{h_n} \mathbf{f}(s) d\tilde{\mathbf{L}}(s)\|^4 + C_6 h_n^4 \leq C_7 h_n$ where $\tilde{\mathbf{L}}(t) := \mathbf{L}(t) - t\mathbb{E}(\mathbf{L}(1))$ for $t \geq 0$.

(c) In the following $f^* := \sup_{s \in \mathbb{R}} \|f(s)\|$. Let $(\gamma_L, \Sigma_L, \nu_L)$ be the characteristic triplet of $(L(t))_{t \geq 0}$ and $\mathbb{B}^{d-1} = \{x \in \mathbb{R}^d : \|x\| \leq 1\}$ be the unit ball in \mathbb{R}^d . We factorize the Lévy measure ν_L into the two Lévy measures

$$\nu_{L_1}(A) := \nu_L(A \setminus \mathbb{B}^{d-1}) \quad \text{and} \quad \nu_{L_2}(A) := \nu_L(A \cap \mathbb{B}^{d-1}) \quad \text{for } A \in \mathcal{B}(\mathbb{R}^d \setminus \{0_d\})$$

such that $\nu_L = \nu_{L_1} + \nu_{L_2}$. Then we can decompose $(L(t))_{t \geq 0}$ into two independent Lévy processes,

$$L(t) = L^{(1)}(t) + L^{(2)}(t) \quad \text{for } t \geq 0, \tag{A.3}$$

where $L^{(1)} = (L^{(1)}(t))_{t \geq 0}$ has the characteristic triplet $(0_d, 0_{d \times d}, \nu_{L_1})$ and $L^{(2)} = (L^{(2)}(t))_{t \geq 0}$ has the characteristic triplet $(\gamma_L, \Sigma_L, \nu_{L_2})$. Then

$$\xi_n = \int_0^{h_n} f(s) dL^{(1)}(s) + \int_0^{h_n} f(s) dL^{(2)}(s) =: \xi_n^{(1)} + \xi_n^{(2)}, \tag{A.4}$$

and $\xi_n^{(1)}$ and $\xi_n^{(2)}$ are independent. Since the Lévy measure of $L^{(1)}$ is finite and $L^{(1)}$ is without Gaussian part and drift, $L^{(1)}$ has the compound Poisson process representation

$$L^{(1)}(t) = \sum_{k=1}^{N(t)} J_k, \quad t \geq 0, \quad \text{and} \quad \xi_n^{(1)} = \sum_{k=1}^{N(h_n)} f(\Gamma_k) J_k, \tag{A.5}$$

where $(J_k)_{k \in \mathbb{N}}$ is a sequence of i.i.d. random vectors independent of the Poisson process $(N(t))_{t \geq 0}$ with intensity $\lambda = \nu_{L_1}(\mathbb{R}^d)$ and jump times $(\Gamma_k)_{k \in \mathbb{N}}$. Now, let B be a relatively compact set in $\mathcal{B}(\mathbb{R}^m \setminus \{0_m\})$ with $\mu(\partial B) = 0$ and $\gamma_B = \inf_{x \in B} \|x\|$, which is larger than 0. Then

$$n\mathbb{P}((nh_n)^{-1/2}\xi_n \in B) \leq n\mathbb{P}\left(\|\xi_n^{(1)}\| > \frac{\gamma_B \sqrt{nh_n}}{2}\right) + n\mathbb{P}\left(\|\xi_n^{(2)}\| > \frac{\gamma_B \sqrt{nh_n}}{2}\right).$$

First, we will show that the first summand with $\xi_n^{(1)}$ converges to 0, using the following conclusions. On the one hand, for $l \geq 1$,

$$\frac{\mathbb{P}(N(h_n) = l)}{h_n} = e^{-\lambda h_n} \frac{(\lambda h_n)^l}{h_n!} \leq C_8 \mathbb{P}(N(1) = l). \tag{A.6}$$

On the other hand,

$$\lim_{n \rightarrow \infty} \frac{\mathbb{P}(N(h_n) = l)}{h_n} = \begin{cases} \lim_{n \rightarrow \infty} e^{-\lambda h_n} \lambda = \lambda & \text{for } l = 1, \\ \lim_{n \rightarrow \infty} e^{-\lambda h_n} \frac{\lambda^l h_n^{l-1}}{l!} = 0 & \text{for } l \geq 2. \end{cases} \tag{A.7}$$

If $U_{l,1} < U_{l,2} < \dots < U_{l,l}$ denotes the order statistic of l i.i.d. uniform random variables on $(0, 1)$ then

$$n\mathbb{P}\left(\|\xi_n^{(1)}\| > \frac{\gamma_B \sqrt{nh_n}}{2}\right) = n \sum_{l=1}^{\infty} \mathbb{P}\left(\left\| \sum_{k=1}^l f(h_n U_{l,k}) J_k \right\| > \frac{\gamma_B \sqrt{nh_n}}{2}\right) \mathbb{P}(N(h_n) = l)$$

(see [34, Theorem 4.5.2]). On the one hand, by (A.6),

$$n\mathbb{P}\left(\|\xi_n^{(1)}\| > \frac{\gamma_B \sqrt{nh_n}}{2}\right) \leq nh_n C_9 \mathbb{P}\left(f^* \sum_{k=1}^{N(1)} \|J_k\| > \frac{\gamma_B \sqrt{nh_n}}{2}\right) \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

since $\mathbb{E}((\sum_{k=1}^{N(1)} \|J_k\|)^2) < \infty$ by [38, Corollary 25.8]. On the other hand, since the Lévy measure of $L^{(2)}$ has compact support, all moments of $L^{(2)}(1)$ are finite (cf. [38, Corollary 25.8]). We conclude from (b) that

$$n\mathbb{P}\left(\|\xi_n^{(2)}\| > \frac{\gamma_B\sqrt{nh_n}}{2}\right) \leq n\left(\frac{\gamma_B\sqrt{nh_n}}{2}\right)^{-4} \mathbb{E}\|\xi_n^{(2)}\|^4 \leq C_{10}n(nh_n)^{-2}h_n \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

(d) Note that, for any random variable X with $\mathbb{E}|X|^2 < \infty$, the limits $\lim_{y \rightarrow \infty} y^2\mathbb{P}(|X| > y) = 0$ and $\lim_{y \rightarrow \infty} y^{2-\delta}\mathbb{E}(|X|^\delta \mathbf{1}_{\{|X|>y\}}) = 0$ (apply Hölder inequality) hold. Then

$$\begin{aligned} & n(nh_n)^{-\delta/2}\mathbb{E}(\|\xi_n^{(1)}\|^\delta \mathbf{1}_{\{\|\xi_n^{(1)}\| > \sqrt{nh_n x}\}}) \\ & \leq C_{11}(nh_n)^{(2-\delta)/2}\mathbb{E}\left(\left(f^* \sum_{k=1}^{N(1)} \|J_k\|\right)^\delta \mathbf{1}_{\{f^* \sum_{k=1}^{N(1)} \|J_k\| > \sqrt{nh_n x}\}}\right) \\ & \rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned} \tag{A.8}$$

Moreover, by Markov’s inequality,

$$n(nh_n)^{-\delta/2}\mathbb{E}(\|\xi_n^{(2)}\|^\delta \mathbf{1}_{\{\|\xi_n^{(2)}\| > \sqrt{nh_n x}\}}) \leq C_{12}(nh_n)^{-1} \rightarrow 0 \quad \text{as } n \rightarrow \infty. \tag{A.9}$$

Taking $\mathbb{E}\|\xi_n\|^\delta \leq (\mathbb{E}\|\xi_n\|^2)^{\delta/2} \leq C_{13}h_n^{\delta/2}$ into account, the inequality

$$\begin{aligned} & n(nh_n)^{-\delta/2}\mathbb{E}\|\xi_n^{(2)}\|^\delta \mathbb{P}\left(\|\xi_n^{(1)}\| > \frac{\sqrt{nh_n x}}{2}\right) + n(nh_n)^{-\delta/2}\mathbb{E}\|\xi_n^{(1)}\|^\delta \mathbb{P}\left(\|\xi_n^{(2)}\| > \frac{\sqrt{nh_n x}}{2}\right) \\ & \leq C_{14}n(nh_n)^{-\delta/2}h_n^{\delta/2}h_n(nh_n)^{-1} \rightarrow 0 \quad \text{as } n \rightarrow \infty \end{aligned} \tag{A.10}$$

is valid. Finally, applying (A.8)–(A.10) yields $n(nh_n)^{-\delta/2}\mathbb{E}(\|\xi_n\|^\delta \mathbf{1}_{\{\|\xi_n\| > \sqrt{nh_n x}\}}) \rightarrow 0$ as $n \rightarrow \infty$.

(f) Since $\mathbb{E}(\xi_n) = \mathbf{0}_m$, an application of (d) results in

$$\lim_{n \rightarrow \infty} n(nh_n)^{-1/2}\|\mathbb{E}(\xi_n \mathbf{1}_{\{\|\xi_n\| \leq \sqrt{nh_n x}\}})\| = \lim_{n \rightarrow \infty} n(nh_n)^{-1/2}\|\mathbb{E}(\xi_n \mathbf{1}_{\{\|\xi_n\| > \sqrt{nh_n x}\}})\| = 0.$$

(g) [18, Theorem 3.1] and $\lim_{n \rightarrow \infty} h_n^{-1}\mathbb{E}(\xi_n \xi_n^\top) = \lim_{n \rightarrow \infty} h_n^{-1} \int_0^{h_n} f(s) \Sigma f(s)^\top ds = f(0) \Sigma f(0)^\top$ gives $(nh_n)^{-1} \sum_{k=1}^n \xi_{n,k} \xi_{n,k}^\top \xrightarrow{\mathbb{P}} f(0) \Sigma f(0)^\top$ as $n \rightarrow \infty$.

A.2. Infinite second moments

Here we present some asymptotic results for the integral if $L(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$, $\alpha \in (0, 2)$.

Proposition A.2. *Let $(L(t))_{t \geq 0}$ be an \mathbb{R}^d -valued Lévy process with $L(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$, $0 < \alpha < 2$. Suppose that $(h_n)_{n \in \mathbb{N}}$ is a sequence of positive constants such that $h_n \downarrow 0$ and $\lim_{n \rightarrow \infty} nh_n = \infty$. Set $a_t := a_{\lfloor t \rfloor}$ for $t \geq 0$. Let $f : \mathbb{R} \rightarrow \mathbb{R}^{m \times d}$ be a measurable and bounded function with $\lim_{x \rightarrow 0} f(x) = f(0)$. Define $\xi_n = \int_0^{h_n} f(s) dL(s)$ for $n \in \mathbb{N}$.*

(a) $n\mathbb{P}(a_{nh_n}^{-1} \xi_n \in \cdot) \xrightarrow{\nu} \mu \circ f(0)^{-1}(\cdot)$ on $\mathcal{B}(\mathbb{R}^m \setminus \{\mathbf{0}_m\})$.

(b) There exists a finite positive constant K such that

$$\lim_{n \rightarrow \infty} n\mathbb{P}(\|\xi_n\| > a_{nh_n x}) = Kx^{-\alpha} \quad \text{for } x > 0.$$

(c) Let either $\delta \geq 2$ or $\delta > \alpha$, and let $(L(t))_{t \geq 0}$ be a compound Poisson process. Then, for any $x > 0$, there exists a finite positive constant K_δ such that

$$na_{nh_n}^{-\delta} \mathbb{E}(\|\xi_n\|^\delta \mathbf{1}_{\{\|\xi_n\| \leq a_{nh_n}x\}}) \leq K_\delta x^{\delta-\alpha} \quad \text{for all } n \in \mathbb{N}.$$

(d) Let $\delta \in (0, \alpha)$. Then, for any $x > 0$, there exists a finite positive constant K_δ such that

$$na_{nh_n}^{-\delta} \mathbb{E}(\|\xi_n\|^\delta \mathbf{1}_{\{\|\xi_n\| > a_{nh_n}x\}}) \leq K_\delta x^{\delta-\alpha} \quad \text{for all } n \in \mathbb{N}.$$

(e) Suppose that $\alpha \neq 1$ and $\mathbb{E}(L(1)) = \mathbf{0}_d$ if $1 < \alpha < 2$. Then for any $x > 0$, there exists a finite positive constant K such that

$$na_{nh_n}^{-1} \|\mathbb{E}(\xi_n \mathbf{1}_{\{\|\xi_n\| \leq a_{nh_n}x\}})\| \leq Kx^{1-\alpha} \quad \text{for all } n \in \mathbb{N}.$$

The proof of Proposition A.2 uses the next two lemmas, which can be derived using arguments similar to those given in Section A.1.

Lemma A.1. Let $(L(t))_{t \geq 0}$ be an \mathbb{R}^d -valued Lévy process with $\mathbb{E}\|L(1)\|^2 < \infty$, let $(a_t)_{t \geq 0}$ be an increasing sequence of positive constants in $\mathcal{R}_{1/\alpha}$, $0 < \alpha < 2$, and let $(h_n)_{n \in \mathbb{N}}$ be a sequence of positive constants such that $h_n \downarrow 0$ as $n \rightarrow \infty$ and $\lim_{n \rightarrow \infty} nh_n = \infty$. Moreover, let $f: \mathbb{R} \rightarrow \mathbb{R}^{m \times d}$ be a measurable and bounded function with $\lim_{x \rightarrow 0} f(x) = f(0)$. Define $\xi_n = \int_0^{h_n} f(s) dL(s)$ for $n \in \mathbb{N}$. Finally, let $(\alpha - 1)_+ < \delta < 2$.

(a) $\lim_{n \rightarrow \infty} n\mathbb{P}(a_{nh_n}^{-1} \xi_n \in B) = 0$ for any relatively compact set $B \in \mathcal{B}(\overline{\mathbb{R}}^m \setminus \{\mathbf{0}_m\})$.

(b) $\lim_{n \rightarrow \infty} na_{nh_n}^{-\delta} \mathbb{E}(\|\xi_n\|^\delta \mathbf{1}_{\{\|\xi_n\| > a_{nh_n}x\}}) = 0$ for $x > 0$.

Lemma A.2. Let $L_1 = (\sum_{k=1}^{N(t)} J_k)_{t \geq 0}$ be an \mathbb{R}^d -valued compound Poisson process, and let $f: \mathbb{R} \rightarrow \mathbb{R}^{m \times d}$ be a measurable and bounded function with $\lim_{x \rightarrow 0} f(x) = f(0)$. Define $\xi_n = \int_0^{h_n} f(s) dL_1(s)$ for $n \in \mathbb{N}$. Then, for any $0 < \delta \leq 1$ with $\mathbb{E}\|L(1)\|^\delta < \infty$, there exists a finite positive constant K such that

$$\mathbb{E}\|\xi_n\|^\delta \leq Kh_n.$$

Note that, for an arbitrary driving Lévy process, the result is not valid, e.g. Brownian motion. In general, we only have $\mathbb{E}\|\xi_n\|^\delta \leq Ch_n^{\delta/2}$.

Proof of Proposition A.2. (a) We use the decomposition $\xi_n = \xi_n^{(1)} + \xi_n^{(2)}$ as given in the proof of Proposition A.1 and the notation there. Moreover, $f(0)J_1 \in \mathcal{R}_{-\alpha}(a_n, \lambda^{-1}\mu \circ f(0)^{-1}(\cdot))$ due to [19, Lemma 2.1] and $\|J_1\| \in \mathcal{R}_{-\alpha}(a_n)$ as well. First, we will show that $\xi_n^{(1)}$ satisfies the statement. Now, let B be a relatively compact set in $\mathcal{B}(\overline{\mathbb{R}}^m \setminus \{\mathbf{0}_m\})$ with $\mu(\partial B) = 0$ and $\gamma_B = \inf_{x \in B} \|x\|$, which is larger than 0. We define

$$\begin{aligned} n\mathbb{P}(a_{nh_n}^{-1} \xi_n^{(1)} \in B) &= \sum_{l=1}^{\infty} n\mathbb{P}\left(a_{nh_n}^{-1} \sum_{k=1}^l f(h_n U_{l,k}) J_k \in B\right) \mathbb{P}(N(h_n) = l) \\ &=: \sum_{l=1}^{\infty} a_{n,l}^*. \end{aligned} \tag{A.11}$$

Furthermore, (A.6) gives, for any $l \geq 1$,

$$0 \leq a_{n,l}^* \leq C_1 nh_n \mathbb{P}\left(a_{nh_n}^{-1} f^* \sum_{k=1}^l \|J_k\| > \gamma_B\right) \mathbb{P}(N(1) = l) =: b_{n,l}^*,$$

and, for some finite constants $C_2, C_3, C_4 > 0$,

$$\begin{aligned} \lim_{n \rightarrow \infty} b_{n,l}^* &= C_2 l f^{*\alpha} \gamma_B^{-\alpha} \mathbb{P}(N(1) = l), \\ \lim_{n \rightarrow \infty} \sum_{l=1}^{\infty} b_{n,l}^* &= C_3 \lim_{n \rightarrow \infty} nh_n \mathbb{P}\left(a_{nh_n}^{-1} \sum_{k=1}^{N(1)} \|J_k\| > f^{*-1} \gamma_B\right) = C_4 f^{*\alpha} \gamma_B^{-\alpha}, \end{aligned}$$

where we have used the fact that $\sum_{k=1}^l \|J_k\|$ and $\sum_{k=1}^{N(1)} \|J_k\|$ are in $\mathcal{R}_{-\alpha}(a_n)$ by [36, Theorem 6.1 and Proposition 7.4] and [19, Lemma 2.1], respectively. Since (A.7), (A.11), and $\lim_{n \rightarrow \infty} f(h_n U_{1,1}) = f(0)$, \mathbb{P} -almost surely yield

$$\lim_{n \rightarrow \infty} a_{n,1}^* = \lim_{n \rightarrow \infty} nh_n \mathbb{P}(a_{nh_n}^{-1} f(0) J_1 \in B) \lambda = \mu \circ f(0)^{-1}(B),$$

and, moreover, (A.7) results in $\lim_{n \rightarrow \infty} a_{n,l}^* = 0$ for $l \geq 2$, we conclude from Pratt’s theorem (see [32]) that

$$\lim_{n \rightarrow \infty} n \mathbb{P}(a_{nh_n}^{-1} \xi_n^{(1)} \in B) = \sum_{l=1}^{\infty} \lim_{n \rightarrow \infty} a_{n,l}^* = \mu \circ f(0)^{-1}(B). \tag{A.12}$$

Furthermore, the Lévy measure of $L^{(2)}$ has compact support. Thus, it follows from [38, Corollary 25.8] that all moments of $\|L^{(2)}(1)\|$ exist. The statement then follows from Lemma A.1(a), (A.4), and (A.12).

Part (b) follows from (a) and [35, Proposition 3.12].

(c) *Step 1.* Let $(L(t))_{t \geq 0}$ be a compound Poisson process as given in (A.5), let $f(s) = I_{d \times d} \mathbf{1}_{[0, \infty)}(s)$, and let $\delta > \alpha$ (if $\delta \geq 2$ then, in particular, $\delta > \alpha$). Keep in mind that $L(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$ and $J_1 \in \mathcal{R}_{-\alpha}(a_n, \mu/\lambda)$ by [19, Lemma 2.1]. Then

$$\begin{aligned} &\mathbb{E}(\|L(h_n)\|^\delta \mathbf{1}_{\{\|L(h_n)\| \leq a_{nh_n} x\}}) \\ &= h_n \mathbb{E}(\|J_1\|^\delta \mathbf{1}_{\{\|J_1\| \leq a_{nh_n} x\}}) \frac{\mathbb{P}(N(h_n) = 1)}{h_n} \\ &\quad + \sum_{l=2}^{\infty} h_n \mathbb{E}\left(\left\|\sum_{k=1}^l J_k\right\|^\delta \mathbf{1}_{\{\|\sum_{k=1}^l J_k\| \leq a_{nh_n} x\}}\right) \frac{\mathbb{P}(N(h_n) = l)}{h_n}. \end{aligned} \tag{A.13}$$

By [36, Theorem 6.1 and Proposition 7.4] and $\|\sum_{k=1}^l J_k\| \in \mathcal{R}_{-\alpha}(a_n)$, it follows from Karamata’s theorem that, for any $l \geq 1$,

$$\lim_{n \rightarrow \infty} nh_n a_{nh_n}^{-\delta} \mathbb{E}\left(\left\|\sum_{k=1}^l J_k\right\|^\delta \mathbf{1}_{\{\|\sum_{k=1}^l J_k\| \leq a_{nh_n} x\}}\right) = l C_5 x^{\delta-\alpha}. \tag{A.14}$$

As in (a), we apply Pratt’s theorem such that (A.7), (A.13), and (A.14) result in

$$\lim_{n \rightarrow \infty} na_{nh_n}^{-\delta} \mathbb{E}(\|L(h_n)\|^\delta \mathbf{1}_{\{\|L(h_n)\| \leq a_{nh_n} x\}}) = \lambda C_5 x^{\delta-\alpha}.$$

Step 2. Let $(L(t))_{t \geq 0}$ be the compound Poisson process given in (A.5), let f be arbitrary, and let $\delta > \alpha$. Since

$$\mathbb{P}(\|\xi_n\| > y) \leq \mathbb{P}\left(f^* \sum_{k=1}^{N(h_n)} \|J_k\| > y\right) \quad \text{for any } y > 0$$

and $L^*(t) := f^* \sum_{k=1}^{N(t)} \|J_k\|$ for $t \geq 0$ is a compound Poisson process with $L^*(1) \in \mathcal{R}_{-\alpha}(a_n)$, we have

$$\begin{aligned} & n a_{nh_n}^{-\delta} \mathbb{E}(\|\xi_n\|^\delta \mathbf{1}_{\{\|\xi_n\| \leq a_{nh_n} x\}}) \\ & \leq n x^\delta \mathbb{P}(L^*(h_n) > a_{nh_n} x) + n a_{nh_n}^{-\delta} \mathbb{E}(L^*(h_n)^\delta \mathbf{1}_{\{L^*(h_n) \leq a_{nh_n} x\}}), \end{aligned}$$

which converges to $C_6 x^{\delta-\alpha}$ due to (b) and step 1.

Step 3. Let $(L(t))_{t \geq 0}$ be a Lévy process, let f be arbitrary and $\delta \geq 2$, and let $\xi_n = \xi_n^{(1)} + \xi_n^{(2)}$ as given in (A.4). Furthermore, let $\epsilon > 0$. Then the decomposition

$$\begin{aligned} n a_{nh_n}^{-\delta} \mathbb{E}(\|\xi_n\|^\delta \mathbf{1}_{\{\|\xi_n\| \leq a_{nh_n} x\}}) &= n a_{nh_n}^{-\delta} \mathbb{E}(\|\xi_n\|^\delta \mathbf{1}_{\{\|\xi_n\| \leq a_{nh_n} x\}} \mathbf{1}_{\{\|\xi_n^{(1)}\| \leq a_{nh_n}(x+\epsilon)\}}) \\ & \quad + n a_{nh_n}^{-\delta} \mathbb{E}(\|\xi_n\|^\delta \mathbf{1}_{\{\|\xi_n\| \leq a_{nh_n} x\}} \mathbf{1}_{\{\|\xi_n^{(1)}\| > a_{nh_n}(x+\epsilon)\}}) \\ & =: I_{n,1} + I_{n,2} \end{aligned}$$

holds. Furthermore,

$$\begin{aligned} I_{n,1} &\leq n a_{nh_n}^{-\delta} 2^\delta \mathbb{E}(\|\xi_n^{(1)}\|^\delta \mathbf{1}_{\{\|\xi_n^{(1)}\| \leq a_{nh_n}(x+\epsilon)\}}) \\ & \quad + n 2^\delta (2x + \epsilon)^\delta \mathbb{E}\left(\left\|\frac{\xi_n^{(2)}}{a_n(2x + \epsilon)}\right\|^\delta \mathbf{1}_{\{\|\xi_n^{(2)}\| \leq a_n(2x+\epsilon)\}}\right) \\ &\leq n a_{nh_n}^{-\delta} 2^\delta \mathbb{E}(\|\xi_n^{(1)}\|^\delta \mathbf{1}_{\{\|\xi_n^{(1)}\| \leq a_{nh_n}(x+\epsilon)\}}) + n C_7 a_{nh_n}^{-2} \mathbb{E}\|\xi_n^{(2)}\|^2 \\ &\rightarrow C_8(x + \epsilon)^{\delta-\alpha} \quad \text{as } n \rightarrow \infty \end{aligned}$$

by step 2 and Proposition A.1(a). In the last inequality we required that $\delta \geq 2$. Moreover, applying (b) and Proposition A.1(a) results in

$$\begin{aligned} I_{n,2} &\leq n \mathbb{P}(\|\xi_n^{(2)}\| > a_{nh_n} \epsilon) \mathbb{P}(\|\xi_n^{(1)}\| > a_{nh_n}(x + \epsilon)) \\ &\leq C_9 \epsilon^{-2} h_n a_{nh_n}^{-2} n \mathbb{P}(\|\xi_n^{(1)}\| > a_{nh_n}(x + \epsilon)), \end{aligned}$$

which tends to 0 as $n \rightarrow \infty$. Thus, (c) follows.

(d) Follows with Karamata’s and Pratt’s theorem and similar arguments as in (c).

(e) Step 1. Let $1 < \alpha < 2$. Then $\mathbb{E}(\xi_n) = \mathbf{0}_m$. Hence,

$$\begin{aligned} n a_{nh_n}^{-1} \|\mathbb{E}(\xi_n \mathbf{1}_{\{\|\xi_n\| \leq a_{nh_n} x\}})\| &= n a_{nh_n}^{-1} \|\mathbb{E}(\xi_n \mathbf{1}_{\{\|\xi_n\| > a_{nh_n} x\}})\| \\ &\leq n a_{nh_n}^{-1} \mathbb{E}(\|\xi_n\| \mathbf{1}_{\{\|\xi_n\| > a_{nh_n} x\}}), \end{aligned}$$

such that we can apply (d).

Step 2. Let $\alpha \in (0, 1)$. Again, we use the decomposition $\xi_n = \xi_n^{(1)} + \xi_n^{(2)}$ given in (A.4). Thus,

$$\mathbb{E}(\xi_n \mathbf{1}_{\{\|\xi_n\| \leq a_{nh_n} x\}}) = \mathbb{E}(\xi_n^{(1)} \mathbf{1}_{\{\|\xi_n\| \leq a_{nh_n} x\}}) + \mathbb{E}(\xi_n^{(2)} \mathbf{1}_{\{\|\xi_n\| \leq a_{nh_n} x\}}) =: I_{n,1} + I_{n,2}.$$

On the one hand, let, for some $\epsilon > 0$,

$$\|I_{n,1}\| \leq \left[\int_0^{a_{nh_n}(x+\epsilon)} + \int_{a_{nh_n}(x+\epsilon)}^\infty \right] \mathbb{P}(\|\xi_n^{(1)}\| > y, \|\xi_n\| \leq a_{nh_n} x) \, dy =: I_{n,1,1} + I_{n,1,2}.$$

Then

$$I_{n,1,1} \leq \mathbb{E}(\|\xi_n^{(1)}\| \mathbf{1}_{\{\|\xi_n^{(1)}\| \leq a_{nh_n}(x+\epsilon)\}}).$$

Hence, by (c) and Proposition A.1(a)

$$\limsup_{n \rightarrow \infty} na_{nh_n}^{-1} I_{n,1,1} \leq C_{13}x^{1-\alpha}.$$

Furthermore,

$$I_{n,1,2} \leq \mathbb{P}(\|\xi_n^{(1)}\| > a_{nh_n}(x + \epsilon))\mathbb{E}\|\xi_n^{(2)}\|,$$

such that, by (b) and Proposition A.1(a),

$$\limsup_{n \rightarrow \infty} na_{nh_n}^{-1} I_{n,1,2} = 0.$$

To conclude, $na_{nh_n}^{-1} I_{n,1} \leq C_{14}x^{1-\alpha}$ for all $n \in \mathbb{N}$. On the other hand, we have

$$\begin{aligned} \|I_{n,2}\| &\leq \|\mathbb{E}((\xi_n^{(2)} - \mathbb{E}(\xi_n^{(2)}))\mathbf{1}_{\{\|\xi_n\| > a_{nh_n}x\}})\| + \|\mathbb{E}(\mathbb{E}(\xi_n^{(2)})\mathbf{1}_{\{\|\xi_n\| \leq a_{nh_n}x\}})\| \\ &\leq \mathbb{E}(\|\xi_n^{(2)}\| \mathbf{1}_{\{\|\xi_n^{(1)}\| > a_{nh_n}x/2\}}) + \mathbb{E}(\|\xi_n^{(2)}\| \mathbf{1}_{\{\|\xi_n^{(2)}\| > a_{nh_n}x/2\}}) + 2\|\mathbb{E}(\xi_n^{(2)})\| \\ &=: I_{n,2,1} + I_{n,2,2} + I_{n,2,3}. \end{aligned}$$

Then, by (b), Proposition A.1(a), and the facts that $\|\mathbb{E}(\xi_n^{(2)})\| \leq C_{15}h_n$ and $\alpha \in (0, 1)$,

$$\begin{aligned} na_{nh_n}^{-1} I_{n,2,1} &= a_{nh_n}^{-1} \mathbb{E}\|\xi_n^{(2)}\| n\mathbb{P}(\|\xi_n^{(1)}\| > a_{nh_n}x/2) \rightarrow 0 \quad \text{as } n \rightarrow \infty, \\ na_{nh_n}^{-1} I_{n,2,3} &\leq C_{16}nh_n a_{nh_n}^{-1} \rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Finally, by Markov’s inequality,

$$\begin{aligned} na_{nh_n}^{-1} I_{n,2,2} &= \frac{nx}{2\mathbb{P}(\|\xi_n^{(2)}\| > a_{nh_n}x/2)} + na_{nh_n}^{-1} \int_{a_{nh_n}x/2}^{\infty} \mathbb{P}(\|\xi_n^{(2)}\| > y) dy \leq C_{17}nh_n a_{nh_n}^{-2} \\ &\rightarrow 0 \quad \text{as } n \rightarrow \infty; \end{aligned}$$

thus, $\lim_{n \rightarrow \infty} na_{nh_n}^{-1} I_{n,2} = 0$.

Acknowledgements

Financial support from the Deutsche Forschungsgemeinschaft through the research grants STE 2005/1-1 and FA 809/2-2, and RiskLab, ETH Zürich, is gratefully acknowledged.

References

- [1] ANDRESEN, A., BENTH, F. E., KOEKEBAKKER, S. AND ZAKAMULIN, V. (2014). The CARMA interest rate model. *Internat. J. Theoret. Appl. Finance* **17**, 1450008.
- [2] BERGSTROM, A. R. (1990). *Continuous Time Econometric Modelling*. Oxford University Press.
- [3] BEVERIDGE, S. AND NELSON, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the ‘business cycle’. *J. Monetary Econom.* **7**, 151–174.
- [4] BROCKWELL, P. J. (2001). Lévy-driven CARMA processes. *Ann. Inst. Statist. Math.* **53**, 113–124.
- [5] BROCKWELL, P. J. (2009). Lévy-driven continuous-time ARMA processes. In *Handbook of Financial Time Series*, Springer, Berlin, pp. 457–480.
- [6] BROCKWELL, P. J., FERRAZZANO, V. AND KLÜPPELBERG, C. (2013). High-frequency sampling and kernel estimation for continuous-time moving average processes. *J. Time Ser. Anal.* **34**, 385–404.
- [7] COMTE, F. (1999). Discrete and continuous time cointegration. *J. Econometrics* **88**, 207–226.
- [8] DAVIS, R. AND RESNICK, S. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Prob.* **13**, 179–195.

- [9] DAVIS, R., MARENGO, J. AND RESNICK, S. (1985). Extremal properties of a class of multivariate moving averages. *Bull. Inst. Internat. Statist.* **51**, 185–192.
- [10] DOOB, J. L. (1944). The elementary Gaussian processes. *Ann. Math. Statist.* **15**, 229–282.
- [11] ENGLE, R. F. AND GRANGER, C. W. J. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica* **55**, 251–276.
- [12] FASEN, V. (2013). Statistical estimation of multivariate Ornstein–Uhlenbeck processes and applications to co-integration. *J. Econometrics* **172**, 325–337.
- [13] FASEN, V. (2013). Time series regression on integrated continuous-time processes with heavy and light tails. *Econometric Theory* **29**, 28–67.
- [14] FASEN, V. AND FUCHS, F. (2013). On the limit behavior of the periodogram of high-frequency sampled stable CARMA processes. *Stoch. Process. Appl.* **123**, 229–273.
- [15] FASEN, V. AND FUCHS, F. (2013). Spectral estimates for high-frequency sampled continuous-time autoregressive moving average processes. *J. Time Series Anal.* **34**, 532–551.
- [16] GARCÍA, I., KLÜPPELBERG, C. AND MÜLLER, G. (2011). Estimation of stable CARMA models with an application to electricity spot prices. *Statist. Modelling* **11**, 447–470.
- [17] GARNIER, H. AND WANG, L. (eds) (2008). *Identification of Continuous-Time Models from Sampled Data*. Springer, London.
- [18] GUT, A. (1992). Complete convergence of arrays. *Period. Math. Hungar.* **25**, 51–75.
- [19] HULT, H. AND LINDSKOG, F. (2007). Extremal behavior of stochastic integrals driven by regularly varying Lévy processes. *Ann. Prob.* **35**, 309–339.
- [20] JACOD, J. AND SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd edn. Springer, Berlin.
- [21] JOHANSEN, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.
- [22] KALLENBERG, O. (1997). *Foundations of Modern Probability*. Springer, New York.
- [23] KESSLER, M. AND RAHBEK, A. (2001). Asymptotic likelihood based inference for co-integrated homogenous Gaussian diffusions. *Scand. J. Statist.* **28**, 455–470.
- [24] LARSSON, E. K., MOSSBERG, M. AND SÖDERSTRÖM, T. (2006). An overview of important practical aspects of continuous-time ARMA system identification. *Circuits Systems Signal Process.* **25**, 17–46.
- [25] MARQUARDT, T. AND STELZER, R. (2007). Multivariate CARMA processes. *Stoch. Process. Appl.* **117**, 96–120.
- [26] MEERSCHAERT, M. M. AND SCHEFFLER, H.-P. (2000). Moving averages of random vectors with regularly varying tails. *J. Time Ser. Anal.* **21**, 297–328.
- [27] MEERSCHAERT, M. M. AND SCHEFFLER, H.-P. (2001). *Limit Distributions for Sums of Independent Random Vectors*. John Wiley, New York.
- [28] MOSER, M. AND STELZER, R. (2011). Tail behavior of multivariate Lévy-driven mixed moving average processes and supOU stochastic volatility models. *Adv. Appl. Prob.* **43**, 1109–1135.
- [29] PAULASKAS, V. AND RACHEV, S. T. (1998). Cointegrated processes with infinite variance innovations. *Ann. Appl. Prob.* **8**, 775–792.
- [30] PHILLIPS, P. C. B. (1991). Error correction and long-run equilibrium in continuous time. *Econometrica* **59**, 967–980.
- [31] PHILLIPS, P. C. B. AND SOLO, V. (1992). Asymptotics for linear processes. *Ann. Statist.* **20**, 971–1001.
- [32] PRATT, J. W. (1960). On interchanging limits and integrals. *Ann. Math. Statist.* **31**, 74–77.
- [33] RAJPUT, B. S. AND ROSIŃSKI, J. (1989). Spectral representation of infinitely divisible processes. *Prob. Theory Relat. Fields* **82**, 451–487.
- [34] RESNICK, S. I. (1986). Point processes, regular variation and weak convergence. *Adv. Appl. Prob.* **18**, 66–138.
- [35] RESNICK, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.
- [36] RESNICK, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York.
- [37] RVAČEVA, E. L. (1962). On domains of attraction of multi-dimensional distributions. *Select. Transl. Math. Statist. Prob.* **2**, 183–205.
- [38] SATO, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.
- [39] 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.
- [40] STOCKMARR, A. AND JACOBSEN, M. (1994). Gaussian diffusions and autoregressive processes: weak convergence and statistical inference. *Scand. J. Statist.* **21**, 403–419.
- [41] TODOROV, V. (2009). Estimation of continuous-time stochastic volatility models with jumps using high-frequency data. *J. Econometrics* **148**, 131–148.