Time series regression on integrated continuous-time processes with heavy and light tails

Vicky Fasen*

Abstract

The paper presents a cointegration model in continuous time, where the linear combinations of the integrated processes are modeled by a multivariate Ornstein-Uhlenbeck process. The integrated processes are defined as vector-valued Lévy processes with an additional noise term. Hence, if we observe the process at discrete time points, we obtain a multiple regression model. As an estimator for the regression parameter we use the least squares estimator. We show that it is a consistent estimator and derive its asymptotic behavior. The limit distribution is a ratio of functionals of Brownian motions and stable Lévy processes, whose characteristic triplets have an explicit analytic representation. In particular, we present the Wald and the t-ratio statistic and simulate asymptotic confidence intervals. For the proofs we derive some central limit theorems for multivariate Ornstein-Uhlenbeck processes.

AMS 2010 Subject Classifications: primary: 91G70, 60F05
secondary: 62M10, 62H12

Keywords: Brownian motion, central limit theorem, cointegration, continuous time, Lévy process, multivariate regular variation, Ornstein-Uhlenbeck process, point process, t-ratio statistic, Wald statistic.

1 Introduction

Empirical studies of financial time series, as exchange rates, foreign currency spot and futures/forwards rates, stock prices within an industry and interest rates in different countries, show that they are cointegrated (cf. Brenner and Kroner [9] and references therein). Cointegrated originally means that even though time series are not stationary there exist linear combinations of them that render stationarity. This concept goes back to the seminal work of Granger [27] and Engle and Granger [21], and is well understood in discrete time if second moments exists, see e.g., the monographs of Johansen [33] and Lütkepohl [38].

The motivation for this paper comes from *pairs trading*, which is a popular investment strategy among hedge funds and investment banks, and involves trading of securities in pairs. The basic

*ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland, email: vicky.fasen@math.ethz.ch

concept is to find pairs of assets which tend to move together in the long-run, i.e. they are cointegrated, so that the difference of the log assets, called *spread*, is mean-reverting. Hence, if the spread is large then the trader sells the higher-priced asset and buys the lower-priced asset with the knowledge that in the long-term the prices will converge together again. Thus, if the spread tends to its mean value, the trader will sell the assets and realizes a profit (cf. Gatev et al. [26]). Pairs trading is a form of statistical arbitrage. This concept can be applied to any equilibrium relationship in financial markets. Besides pairs trading where it is essential to find the optimal investment strategy (cf. Ekström et al. [19]) there exist also *spread options*, used in fixed income markets, currency and foreign exchange markets, commodity future markets and energy markets, which take into account that the underlying assets are cointegrated if they price options.

In this paper we consider a cointegrated model in continuous time. Continuous-time models provide the basis of option pricing, asset allocation and term structure theory. The underlying observations of asset prices, exchange rates, and interest rates are often irregularly spaced, in particular, in the context of high frequently data. Consequently, one often works with continuous-time models which infer the implied dynamics and properties of the estimated model at different frequencies from the one used in the estimation.

Typical for high frequency financial time series as asset returns and exchange rates are jumps and a distribution which is peaked around zero with a tail distribution decreasing slower to zero than any exponential function. Empirical studies show that these distributions have often a power law tail with an index in $(2, 4)$ (cf. Adler [2], Rachev [57]), which implies that they have finite variances but infinite fourth moments. Already Mandelbrot [40] and Fama [22] noticed in the 60's that the Gaussian distribution is not the appropriate model and suggested to use α-stable distributions with $\alpha \in (0, 2)$ as natural generalization of the Gaussian distribution. Although α-stable distributions have an infinite second moment, tail index estimation are not sufficient to reject stable distributions; see McCulloch [43]. There is, e.g., empirical evidence that electricity prices and the daily trading volume of stocks have a power law tail with an index in $(1, 2)$ (cf. Weron [68] and Aban and Meerschaert [1]). More about α-stable distributions in financial modelling can be found in Mittnik and Rachev [56].

The model

A model for the price of an asset $S = (S(t))_{t\geq 0}$ is

$$S(t) = \exp(L(t) + \zeta(t)), \quad t \geq 0, \tag{1.1}$$

where $L = (L(t))_{t\geq 0}$ is a Lévy process, a stochastic process with independent and stationary increments, and $\zeta = (\zeta(t))_{t\geq 0}$ is a stationary process. This model extends the classical exponential Lévy model for stock prices which can be found, e.g., in the standard textbook of Jeanblanc et al. [31]. Models of the form (1.1) are popular for describing spot prices of commodities whose logarithmic prices are mean reverting. Our model (1.1) extends the model of Lucia and Schwartz [39]; see Benth et al. [7] and references therein for further examples. One interpretation is that L is a Brownian motion which reflects the *long-term equilibrium* and accounts the small variations in the spot prices when normal trading takes place, and the arrival of information, transaction and storages costs,

which causes large fluctuations, are modeled as jump process in the *short-term* behavior ζ. In this paper L and ζ will be very flexible. They may be jump processes or Gaussian processes; they may have a power like tail or a finite second moment. In an exponential Lévy model with finite time horizon (where $\zeta(t) = 0$ for $t \geq 0$) there is no arbitrage except in the case where L is increasing or decreasing, and the market is incomplete, except in the case where L is a Brownian motion and a Poisson process; see Selivanov [65]. Incomplete markets are, for example, typical for commodity markets.

Here, we suppose that we have two prices S_1, S_2, and

$$Y(t) := \log S_1(t) = L_1(t) + \zeta(t), \quad t \geq 0, \tag{1.2}$$

where $L_1 = (L_1(t))_{t\geq 0}$ is a Lévy process and $\zeta = (\zeta(t))_{t\geq 0}$ is some stationary process. The spread of the log prices

$$Z(t) = \log S_2(t) - a \log S_1(t), \quad t \geq 0$$

for some $a \in \mathbb{R}\backslash\{0\}$ is modeled by a mean reverting *Ornstein-Uhlenbeck process*, i.e.,

$$Z(t) = \mathrm{e}^{-\lambda t} Z(0) + \int_0^t \mathrm{e}^{-\lambda(t-s)} \, \mathrm{d}L_2(s), \quad t \geq 0,$$

for some $\lambda > 0$ and some Lévy process $L_2 = (L_2(t))_{t\geq 0}$; this is a common model, see [6, 18, 19, 20]. The parameter λ reflects the speed of mean reversion.

Finally, we suppose that $(Z(t), \zeta(t))_{t\geq 0}$ are jointly stationary, which holds obviously, if ζ and L_2 are independent. Then the price S_2 is also in the class (1.1) since

$$X(t) := \log S_2(t) = aY(t) + Z(t) = \widetilde{L}(t) + \widetilde{\zeta}(t), \quad t \geq 0, \tag{1.3}$$

where $\widetilde{L}(t) = aL_1(t)$ is a Lévy process and $\widetilde{\zeta}(t) = a\zeta(t) + Z(t)$ is a stationary process. In the case where (L_1, L_2) is a bivariate Brownian motion and $\zeta(t) = 0$ for $t \geq 0$ (which means that S_1 is a geometric Brownian motion), Duan and Pliska [18] showed that the model is complete and the price of any option is not affected by the cointegration, and remains as in the standard Black-Scholes framework.

We consider a *multivariate* version of such a cointegrated *regression model* (1.2) and (1.3) in continuous time. Extensions of discrete-time cointegrated autoregressive models to continuous time can be found in Comte [14], Phillips [50, 52] and Stockmarr and Jakobsen [67]. Let $\mathbf{L}_1 = (\mathbf{L}_1(t))_{t\geq 0}$ and $\mathbf{L}_2 = (\mathbf{L}_2(t))_{t\geq 0}$ be q-dimensional and d-dimensional Lévy processes. A multivariate Lévy process $(\mathbf{L}(t))_{t\geq 0}$ in $\mathbb{R}^m$ is characterized by its *Lévy-Khintchine representation* $\mathbb{E}(\mathrm{e}^{i\Theta'\mathbf{L}(t)}) = \exp(-t\Psi(\Theta))$ for $\Theta \in \mathbb{R}^m$ (for a vector $\mathbf{x} \in \mathbb{R}^m$ we write $\mathbf{x}'$ for the transpose of $\mathbf{x}$ and $\|\mathbf{x}\|$ for the Euclidean norm), where

$$\Psi(\Theta) = -i\gamma'\Theta + \frac{1}{2}\Theta'\Sigma\,\Theta + \int_{\mathbb{R}^m} \left(1 - \mathrm{e}^{i\Theta'\mathbf{x}} + i\mathbf{x}'\Theta \mathbb{1}_{\{\|\mathbf{x}\|\leq 1\}}\right) \nu_{\mathbf{L}}(\mathrm{d}\mathbf{x})$$

with $\gamma \in \mathbb{R}^m$, Σ a positive semi-definite matrix in $\mathbb{R}^{m\times m}$ and $\nu_{\mathbf{L}}$ a measure on $\mathbb{R}^m$, called *Lévy measure*, which satisfies $\int_{\mathbb{R}^m} \min\{\|\mathbf{x}\|^2, 1\}\, \nu_{\mathbf{L}}(\mathrm{d}\mathbf{x}) < \infty$ and $\nu_{\mathbf{L}}(\{\mathbf{0}_m\}) = 0$ (where $\mathbf{0}_m$ is the zero-vector in $\mathbb{R}^m$). The triplet $(\gamma, \Sigma, \nu_{\mathbf{L}})$ is also called *characteristic triplet*, because it characterizes

completely the distribution of the Lévy process (cf. the excellent monograph of Sato [63] for more details on Lévy processes). Typical examples for Lévy processes are the Brownian motion, whose increments are multivariate normal distributed with covariance matrix Σ and $\nu_{\mathbf{L}} = 0$, and α-stable Lévy processes, $\alpha \in (0, 2)$, where the Gaussian part represented by Σ is zero and the Lévy measure has the representation

$$\nu_L(\mathrm{d}\mathbf{x}) = r^{-1-\alpha}\mathrm{d}r\sigma(\mathrm{d}\mathbf{s}), \tag{1.4}$$

where $r = \|\mathbf{x}\|$, $\mathbf{s} = \frac{\mathbf{x}}{\|\mathbf{x}\|}$ and σ is a measure on the unit sphere of $\mathbb{R}^m$. An α-stable Lévy process with $\alpha = 2$ is defined to be a Brownian motion.

Moreover, let $\mathbf{A} \in \mathbb{R}^{d\times q}$ and $\mathbf{\Lambda} \in \mathbb{R}^{d\times d}$, where the eigenvalues of $\mathbf{\Lambda}$ have strictly positive real parts. The *multivariate cointegration model* is

$$\mathbf{X}(t) = \mathbf{A}\mathbf{Y}(t) + \mathbf{Z}(t), \quad t \geq 0, \quad \text{in } \mathbb{R}^d, \tag{1.5}$$

$$\mathbf{Y}(t) = \mathbf{L}_1(t) + \boldsymbol{\zeta}(t), \quad t \geq 0, \quad \text{in } \mathbb{R}^q, \tag{1.6}$$

where $\boldsymbol{\zeta} = (\boldsymbol{\zeta}(t))_{t\geq 0}$ is some stationary process in $\mathbb{R}^q$, and $\mathbf{Z} = (\mathbf{Z}(t))_{t\geq 0}$ is a stationary Ornstein-Uhlenbeck process in $\mathbb{R}^d$ with representation

$$\mathbf{Z}(t) = \mathrm{e}^{-\mathbf{\Lambda}t}\mathbf{Z}(0) + \int_0^t \mathrm{e}^{-\mathbf{\Lambda}(t-s)}\mathrm{d}\mathbf{L}_2(s), \quad t \geq 0. \tag{1.7}$$

We can take a stationary version of $\mathbf{Z}$ by defining $\mathbf{Z}(0) = \int_{-\infty}^0 \mathrm{e}^{\mathbf{\Lambda}s}\mathrm{d}\mathbf{L}_2(s)$ with $(\mathbf{L}_2(t))_{t\leq 0}$ an independent copy of $(\mathbf{L}_2(t-))_{t\geq 0}$, since the eigenvalues of $\mathbf{\Lambda}$ have strictly positive real parts and the logarithmic moments of the Lévy measure are finite under Assumption 3.1-3.3 (cf. Sato and Yamazato [64], Theorem 4.1). Thus, in the following $\mathbf{Z}$ will be a stationary Ornstein-Uhlenbeck process. We will furthermore extend the model in (1.5), and allow the short-run equilibrium $\mathbf{Z}$ to be more general than an Ornstein-Uhlenbeck model. If $\mathbf{\Lambda}$ is a diagonal matrix in $\mathbb{R}^{d\times d}$ then any component of $\mathbf{Z}$ is an one-dimensional Ornstein-Uhlenbeck process.

By definition (1.6), $\mathbf{Y}$ is integrated since it is non stationary but it has stationary increments. Moreover, it is not cointegrated if $\mathbf{L}_1$ has independent components, i.e. there exists no linear combinations of $\mathbf{Y}$ which are stationary. It is obvious that $(\mathbf{X}', \mathbf{Y}')'$ is cointegrated with cointegrating matrix $(\mathbf{I}_{d\times d}, -\mathbf{A})$ if $\mathbf{A} \neq \mathbf{0}_{d\times q}$ (where $\mathbf{I}_{d\times d}$ denotes the identity matrix in $\mathbb{R}^{d\times d}$ and $\mathbf{0}_{d\times q}$ the zero-matrix in $\mathbb{R}^{d\times q}$). Furthermore, if $\mathbf{L}_1$ has independent components, then the rank of $\mathbf{A}$ is equal to the rank of cointegration.

The estimation problem

Our aim is to present for the multiple cointegration model (1.5)-(1.7), an estimator for the regression parameter $\mathbf{A}$. Assume the following observation scheme

$$\mathbb{X}_n' = (\mathbf{X}(h), \ldots, \mathbf{X}(nh)) \in \mathbb{R}^{d\times n}, \qquad \mathbb{Y}_n' = (\mathbf{Y}(h), \ldots, \mathbf{Y}(nh)) \in \mathbb{R}^{q\times n}$$

with grid distance $h > 0$ (see Remark 3.11 *(iii)* if h depends also on n and tends to 0 as $n \to \infty$). We use as estimator for $\mathbf{A}$ the least squares estimator

$$\widehat{\mathbf{A}}_n = \mathbb{X}_n'\mathbb{Y}_n(\mathbb{Y}_n'\mathbb{Y}_n)^{-1}. \tag{1.8}$$

We will show that under general assumptions the least squares estimator is a consistent estimator and we will derive its asymptotic behavior when $\|\mathbf{L}_1(1)\|$ and $\|\mathbf{L}_2(1)\|$, respectively, has either a heavy tail in the sense that it is regularly varying of some index in $(0,2)$, or a finite second moment. We cover the possibility that one has a finite second moment and the other is heavy tailed, and that $\mathbf{L}_1$ and $\mathbf{L}_2$ are dependent. We obtain an explicit representation of the limit distribution of the estimation error, which allows us to present asymptotic confidence intervals for parameter tests on components of $\mathbf{A}$. The limit distribution is a functional of stable Lévy processes and Brownian motions depending on the tail behavior of $\|\mathbf{L}_i(1)\|$, $i = 1, 2$. Moreover, we derive the t-ratio statistic for $\mathbf{A}$. Simulation studies suggest that the asymptotic confidence intervals of that statistic do not depend on the tail behavior of $\|\mathbf{L}_i(1)\|$, $i = 1, 2$. Hence, the performance of the least squares estimator can be tested without knowing anything about $\mathbf{L}_1$ and $\mathbf{L}_2$, which is in the case of heavy tailed distributions unusual and valuable for statistical purpose.

The paper is structured in the following way. In Section 2 we present central limit results, which we need in order to derive the asymptotic behavior of our estimator in Section 3. In Section 3 we state that the least squares estimator is a consistent estimator and present its asymptotic behavior. We show that the results not only hold for a multivariate Ornstein-Uhlenbeck process $\mathbf{Z}$ but also for the much larger class of multivariate continuous-time ARMA (CARMA) processes, which, in particular, include Ornstein-Uhlenbeck processes. Examples of simulated confidence intervals are also presented in that section. Section 4 contains test statistics as the t-ratio and the Wald statistic for our setup. Finally, in Section 5 we derive the proofs of the results.

We will continue using the notation $\Longrightarrow$ for weak convergence, $\xrightarrow{\mathbb{P}}$ for convergence in probability, and $\overset{\nu}{\Longrightarrow}$ for vague convergence. Let $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ and let $\mathcal{B}(\cdot)$ be the Borel-σ-algebra. For $x \in \mathbb{R}$ we write $\lfloor x \rfloor = \sup\{k \in \mathbb{N} : k \leq x\}$. The expression $\operatorname{diag}(\mathbf{B}_1, \mathbf{B}_2)$ for $\mathbf{B}_1 \in \mathbb{R}^{m_1 \times m_1}$, $\mathbf{B}_2 \in \mathbb{R}^{m_2 \times m_2}$, $m_1, m_2 \in \mathbb{N}$, stands for a block diagonal matrix with block $\mathbf{B}_1$ and $\mathbf{B}_2$, respectively. For a vector $\mathbf{x} \in \mathbb{R}^m$ we write $\mathbf{x}'$ for the transpose of $\mathbf{x}$. The matrix $\mathbf{0}_{m_1 \times m_2}$ is the zero matrix in $\mathbb{R}^{m_1 \times m_2}$ and $\mathbf{I}_{m_1 \times m_1}$ is the identity matrix in $\mathbb{R}^{m_1 \times m_1}$. The symbol $\otimes$ denotes the Kronecker product and we use as norms the Euclidean norm $\|\cdot\|$ in $\mathbb{R}^m$ and the corresponding operator norm $\|\cdot\|$ for matrices. Then $\lambda := \|\mathbf{\Lambda}\|$ is the spectral norm of $\mathbf{\Lambda}$. An $S_\alpha(1, 0, 0)$-distribution will be an α-stable distribution with scale parameter 1, skewness and shift parameter 0 in the sense of Samorodnitsky and Taqqu [62]. Particularly, for $\alpha = 2$ this is a Gaussian distribution. Finally, for a metric space E we write $(\mathbb{D}[0,1], E)$ for the space of all càdlàg (continue à droite et limitée à gauche = right continuous, with left limits) functions on $[0,1]$ with values in E induced with the Skorokhod J_1 topology. The quadratic covariation process of two semimartingales $\mathbf{W}_1 = (\mathbf{W}_{1,1}(t), \ldots, \mathbf{W}_{1,m_1}(t))_{t \geq 0}$ in $\mathbb{R}^{m_1}$ and $\mathbf{W}_2 = (\mathbf{W}_{2,1}(t), \ldots, \mathbf{W}_{2,m_2}(t))_{t \geq 0}$ in $\mathbb{R}^{m_2}$ is denoted by $[\mathbf{W}_1, \mathbf{W}_2]_t = ([\mathbf{W}_{1,i}, \mathbf{W}_{2,j}]_t)_{i=1,\ldots,m_1, j=1,\ldots,m_2}$ for $t \geq 0$.

2 Limit results

2.1 Domain of attraction

The asymptotic behavior of our estimator in Section 3 is based on central limit results. Therefore, we have to distinguish the different domains of attractions of $\mathbf{L}_1(1)$ and $\mathbf{L}_2(1)$. We say that a random vector $\mathbf{U}$ in $\mathbb{R}^m$ belongs to the *domain of attraction* of an α-stable distribution $\mathbf{S}$ with $\alpha \in (0,2]$ (shortly DN(α)), where $\alpha = 2$ reflects the multivariate normal distribution, if there exists a sequence $(a_n)_{n\in\mathbb{N}}$ of positive numbers and (d_n) of real numbers such that for an iid (independently and identically distributed) sequence $(\mathbf{U}_n)_{n\in\mathbb{N}}$ with distribution $\mathbf{U}$ the asymptotic behavior

$$a_n^{-1}\sum_{k=1}^{n}\mathbf{U}_k - d_n \Longrightarrow \mathbf{S} \quad \text{as } n\to\infty \tag{2.1}$$

holds. The left hand side is only capable to converge to an α-stable distribution with $\alpha \in (0,2]$. Other limit distributions are not possible (cf. Rvačeva [61]). In particular, every α-stable distribution is in its own domain of attraction. A sufficient condition to be in the domain of attraction of a multivariate normal distribution is $\mathbb{E}\|\mathbf{U}\|^2 < \infty$. However, this is not a necessary assumption. In contrast, $\mathbf{U}$ is in the domain of attraction of an α-stable distribution with $\alpha \in (0,2)$ if and only if $\mathbf{U}$ is multivariate regularly varying of index $-\alpha$. Then $\mathbb{E}\|\mathbf{U}\|^2 = \infty$.

Recall that a random matrix $\mathbf{U}$ in $\mathbb{R}^{m\times d}$ is *multivariate regularly varying with index* $-\alpha < 0$ if and only if there exists a non-zero Radon measure μ on $\overline{\mathbb{R}}^{m\times d}\setminus\{\mathbf{0}_{m\times d}\}$ with $\mu(\overline{\mathbb{R}}^{m\times d}\setminus\mathbb{R}^{m\times d}) = 0$ and a sequence $(a_n)_{n\in\mathbb{N}}$ of positive numbers increasing to ∞ such that

$$n\mathbb{P}(a_n^{-1}\mathbf{U}\in\cdot) \overset{v}{\Longrightarrow} \mu(\cdot) \quad \text{as } n\to\infty \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}^{m\times d}\setminus\{\mathbf{0}_{m\times d}\}). \tag{2.2}$$

The limit measure μ is homogenous of order $-\alpha$, i.e., $\mu(uB) = u^{-\alpha}\mu(B)$ for $u > 0$ and $B \in \mathcal{B}(\overline{\mathbb{R}}^{m\times d}\setminus\{\mathbf{0}_{m\times d}\})$. The sequence $(a_n)_{n\in\mathbb{N}}$ in (2.1) can be chosen as in (2.2). We shortly 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. In particular, $\mathbb{E}\|\mathbf{U}\|^r < \infty$ for $r < \alpha$ and $\mathbb{E}\|\mathbf{U}\|^r = \infty$ for $r > \alpha$. For further information regarding multivariate regular variation of random vectors we refer to Resnick [60]. However, we can transfer the results to random matrices in $\mathbb{R}^{m\times d}$ by rewriting the random matrix as a random vector in $\mathbb{R}^{md}$. A typical example for a multivariate regularly random vector are multivariate α-stable distributions. If $\mathbf{L}$ is a multivariate α-stable Lévy process with Lévy measure $\nu_{\mathbf{L}}$ as given in (1.4), then $\mathbf{L}(1) \in \mathcal{R}_{-\alpha}(n^{1/\alpha},\nu_{\mathbf{L}})$.

As a special case we remark that a measurable function $f : (0,\infty) \to (0,\infty)$ is called *regularly varying* of index $-\alpha$, $\alpha \in \mathbb{R}$, if $\lim_{x\to\infty} f(xu)/f(x) = u^{-\alpha}$ for any $u > 0$. In that case we also write $f \in \mathcal{R}_{-\alpha}$. If the random matrix $\mathbf{U} \in \mathcal{R}_{-\alpha}(a_n,\mu)$ then there exists an $\ell_1,\ell_2 \in \mathcal{R}_0$ such that $a_n = n^{1/\alpha}\ell_1(n)$ for $n \in \mathbb{N}$ and $\mathbb{P}(\|\mathbf{U}\| > x) = x^{-\alpha}\ell_2(x)$ for $x > 0$ as well.

In this paper we distinguish

(i) $\mathbf{L}_1(1) \in \text{DN}(\alpha)$, $\mathbf{L}_2(1) \in \text{DN}(\beta)$, $\alpha,\beta \in (0,2]$, and $\mathbf{L}_1$, $\mathbf{L}_2$ are independent Lévy processes;

(ii) $(\mathbf{L}_1(1)',\mathbf{L}_2(1)')' \in \text{DN}(\alpha)$, $\alpha \in (0,2]$.

However, if $\mathbf{L}_1$ and $\mathbf{L}_2$ are dependent and $\mathbf{L}_1(1) \in \mathrm{DN}(\alpha), \mathbf{L}_2(1) \in \mathrm{DN}(\beta)$ and $\beta > \alpha$, then $(\mathbf{L}_1(1)', \mathbf{L}_2(1)')'$ is also in $\mathrm{DN}(\alpha)$.

2.2 Central limit results in DN(α) for $\alpha \in (0,2)$

For the proof of the asymptotic behavior of the least squares estimator we require more general limit results than (2.1), where our increments are an iid sequence. If $(\mathbf{L}_1(1)', \mathbf{L}_2(1)')' \in \mathrm{DN}(\alpha)$, $\alpha \in (0,2)$, the proofs of our limit results rely on point process techniques (cf. Section 5.1). We follow Resnick's [58] notation of point processes. Let S denote the locally compact and separable Hausdorff space $[0,\infty) \times \overline{\mathbb{R}}^m \backslash \{\mathbf{0}_m\}$ with the Borel σ-field $\mathcal{B}(S)$, and let $M_P(S)$ denote the class of point measures (integer-valued Radon measures) on S provided with a metric that generates the topology of vague convergence. A measure of the form $\sum_{k\in I} \varepsilon_{x_k}$, where $x_k \in S$, I is at most countable and ε_{x_k} denotes the Dirac measure in x_k, is a point measure. A point process is a measurable map from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ into $(M_P(S), \mathcal{B}(M_P(S)))$. A famous point process is the Poisson random measure denoted by PRM(ϑ). A point process N is PRM(ϑ) if

(a) $N(A)$ is Poisson distributed with mean $\vartheta(A)$ for every $A \in \mathcal{B}(\mathcal{S})$,

(b) for all mutually disjoint sets $A_1, \ldots, A_n \in \mathcal{B}(\mathcal{S})$, $N(A_1), \ldots, N(A_n)$ are independent.

More about point processes can be found in Daley and Vere-Jones [15] and Kallenberg [34].

Proposition 2.1 *Let the multivariate cointegration model* (1.5)-(1.7) *be given. Suppose that* $(\mathbf{L}(t))_{t\geq 0}$ *is a* w*-dimensional Lévy process with* $\mathbf{L}(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$ *for* $0 < \alpha < 2$,

$$\mathbf{L}_1(t) := \Sigma_1 \mathbf{L}(t) \quad \text{and} \quad \mathbf{L}_2(t) := \Sigma_2 \mathbf{L}(t), \quad t \geq 0,$$

where $\Sigma_1 \in \mathbb{R}^{q\times w}$ *and* $\Sigma_2 \in \mathbb{R}^{d\times w}$. *Define for* $t \geq 0$ *and* $n \in \mathbb{N}$:

$$\begin{aligned}
\mathbf{S}_n^{(1)}(t) &= a_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \left(\Delta \mathbf{L}_1(kh) - \mathbb{E}(\Delta \mathbf{L}_1(h) \mathbb{1}_{\{\|\Delta \mathbf{L}_1(h)\| \leq a_n\}}) \right), \\
\mathbf{S}_n^{(2)}(t) &= a_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \left(\mathbf{Z}(kh) - \sum_{i=0}^{\infty} \mathrm{e}^{-h\mathbf{\Lambda} i} \mathbb{E} \left(\int_0^h \mathrm{e}^{-\mathbf{\Lambda} s} \, \mathrm{d}\mathbf{L}_2(s) \mathbb{1}_{\{\| \int_0^h \mathrm{e}^{-\mathbf{\Lambda} s} \, \mathrm{d}\mathbf{L}_2(s) \| \leq a_n\}} \right) \right), \\
\mathbf{S}_n^{(3)}(t) &= a_n^{-2} \sum_{k=1}^{\lfloor nt \rfloor} \mathbf{Z}(kh) \mathbf{Z}(kh)', \qquad \mathbf{S}_n^{(4)}(t) = a_n^{-2} \sum_{k=1}^{\lfloor nt \rfloor} \mathbf{Z}((k+1)h) \mathbf{Z}(kh)',
\end{aligned}$$

where $\Delta \mathbf{L}_1(kh) = \mathbf{L}_1(kh) - \mathbf{L}_1((k-1)h)$. *Let Leb be the Lebesgue measure and*

$$\sum_{k=1}^{\infty} \varepsilon_{(t_k, j_k^{(1)}, j_k^{(2)})} \sim \mathrm{PRM}(Leb \times \tilde{\mu})$$

be a point process on $S = [0,\infty) \times (\overline{\mathbb{R}}^{q+d} \backslash \{\mathbf{0}_{q+d}\})$, *where*

$$\widetilde{\mu}(B) = h\mathbb{E} \left(\mu \left(\left\{ \mathbf{x} \in \overline{\mathbb{R}}^w \backslash \{\mathbf{0}_w\} : \left(\Sigma_1 \mathbf{x}, \mathrm{e}^{-h\mathbf{\Lambda} U} \Sigma_2 \mathbf{x} \right) \in B \right\} \right) \right) \tag{2.3}$$

for $B \in \mathcal{B}(\overline{\mathbb{R}}^{q+d}\backslash\{\mathbf{0}_{q+d}\})$ *and* U *is a uniform random variable on* $(0,1)$ *and, similarly,*

$$\mu^*(B) = h\mathbb{E}\left(\mu\left\{\mathbf{x} \in \overline{\mathbb{R}}^w\backslash\{\mathbf{0}_w\} : \left(\ \Sigma_1\mathbf{x}, \textstyle\sum_{i=0}^{\infty} \mathrm{e}^{-h\mathbf{\Lambda}i}\mathrm{e}^{-h\mathbf{\Lambda}U}\Sigma_2\mathbf{x}\ \right) \in B\right\}\right). \tag{2.4}$$

Furthermore, let

$$\mu_1^*(\cdot) := \widetilde{\mu}_1(\cdot) := \mu^*(\cdot \times \overline{\mathbb{R}}^d), \quad \mu_2^*(\cdot) := \mu^*(\overline{\mathbb{R}}^q \times \cdot) \quad \text{and} \quad \widetilde{\mu}_2(\cdot) := \widetilde{\mu}(\overline{\mathbb{R}}^q \times \cdot).$$

Finally, define

$$\begin{aligned}
\mathbf{S}^{(1)}(t) &:= \sum_{t_k \le t} j_k^{(1)} \mathbb{1}_{\{\|j_k^{(1)}\|>1\}} + \lim_{\gamma\to 0}\left(\sum_{t_k\le t} j_k^{(1)}\mathbb{1}_{\{\gamma<\|j_k^{(1)}\|\le 1\}} - t\int_{\gamma<\|\mathbf{x}\|\le 1}\mathbf{x}\,\widetilde{\mu}_1(\mathrm{d}\mathbf{x})\right)\\
&=: \mathbf{S}_1(t) + t\begin{cases} \int_{\|\mathbf{x}\|>1}\mathbf{x}\,\mu_1^*(d\mathbf{x}) & \text{if } \alpha>1,\\ \mathbf{0}_q & \text{if } \alpha=1,\\ -\int_{\|\mathbf{x}\|\le 1}\mathbf{x}\,\mu_1^*(d\mathbf{x}) & \text{if } \alpha<1,\end{cases}\\
\mathbf{S}^{(2)}(t) &:= \sum_{i=0}^{\infty}\mathrm{e}^{-h\mathbf{\Lambda}i}\left(\sum_{t_k\le t} j_k^{(2)}\mathbb{1}_{\{\|j_k^{(2)}\|>1\}} + \lim_{\gamma\to 0}\left(\sum_{t_k\le t} j_k^{(2)}\mathbb{1}_{\{\gamma<\|j_k^{(2)}\|\le 1\}} - t\int_{\gamma<\|\mathbf{x}\|\le 1}\mathbf{x}\,\widetilde{\mu}_2(\mathrm{d}\mathbf{x})\right)\right)\\
&=: \mathbf{S}_2(t) + t\begin{cases} \int_{\|\mathbf{x}\|>1}\mathbf{x}\,\mu_2^*(d\mathbf{x}) & \text{if } \alpha>1,\\ \mathbf{0}_d & \text{if } \alpha=1,\\ -\int_{\|\mathbf{x}\|\le 1}\mathbf{x}\,\mu_2^*(d\mathbf{x}) & \text{if } \alpha<1,\end{cases}\\
\mathbf{S}^{(3)}(t) &:= \sum_{t_k\le t}\sum_{i=0}^{\infty}\mathrm{e}^{-h\mathbf{\Lambda}i} j_k^{(2)} j_k^{(2)'}\mathrm{e}^{-h\mathbf{\Lambda}'i} =: \mathbf{S}_3(t),\\
\mathbf{S}^{(4)}(t) &:= \sum_{t_k\le t}\sum_{i=0}^{\infty}\mathrm{e}^{-h\mathbf{\Lambda}(i+1)} j_k^{(2)} j_k^{(2)'}\mathrm{e}^{-h\mathbf{\Lambda}'i} = \mathrm{e}^{-h\mathbf{\Lambda}}\mathbf{S}^{(3)}(t) =: \mathbf{S}_4(t).
\end{aligned}$$

Let $0 \le t_1 \le \ldots \le t_l \le 1$ *and* $\mathbf{t} := (t_1,\ldots,t_l)$*. For a function* g *we write* $g(\mathbf{t}) := (g(t_1),\ldots,g(t_l))$*. Then we have as* $n\to\infty$*,*

$$\left((\mathbf{S}_n^{(1)}(t))_{t\ge 0}, \mathbf{S}_n^{(2)}(\mathbf{t}), \mathbf{S}_n^{(3)}(\mathbf{t}), \mathbf{S}_n^{(4)}(\mathbf{t})\right) \Longrightarrow \left((\mathbf{S}^{(1)}(t))_{t\ge 0}, \mathbf{S}^{(2)}(\mathbf{t}), \mathbf{S}^{(3)}(\mathbf{t}), \mathbf{S}^{(4)}(\mathbf{t})\right)$$

in $(\mathbb{D}[0,1],\mathbb{R}^q)\times\mathbb{R}^{d\times l}\times\mathbb{R}^{(d\times d)\times l}\times\mathbb{R}^{(d\times d)\times l}$ *equipped with the product topology.*

Remark 2.2

(i) The limit processes $\mathbf{S}^{(1)},\mathbf{S}^{(2)},\mathbf{S}^{(3)}$ and $\mathbf{S}^{(4)}$ are stable Lévy processes, where $\mathbf{S}^{(1)}$ has Lévy measure μ_1^*, $\mathbf{S}^{(2)}$ has Lévy measure μ_2^*, $\mathbf{S}^{(3)}$ has Lévy measure $\mu_2^*(\{\mathbf{x}\in\overline{\mathbb{R}}^d\backslash\{\mathbf{0}_d\} : \mathbf{x}\mathbf{x}'\in\cdot\})$, and $\mathbf{S}^{(4)}$ has Lévy measure $\mu_2^*(\{\mathbf{x}\in\overline{\mathbb{R}}^d\backslash\{\mathbf{0}_d\} : \mathrm{e}^{-h\mathbf{\Lambda}}\mathbf{x}\mathbf{x}'\in\cdot\})$. They are also jointly an α-stable Lévy process. If $\alpha>1$, $\mathbf{S}_1$ and $\mathbf{S}_2$ are centered stable Lévy processes with Lévy measures μ_1^* and μ_2^*, respectively. In the following we will work with the centered Lévy processes $\mathbf{S}_1$ and $\mathbf{S}_2$ and thus, we have defined for the ease of notation also $\mathbf{S}_3$ and $\mathbf{S}_4$ (whose means are infinite).

(ii) The limit result of Proposition 2.1 can also be used to derive estimators for $\mathbf{\Lambda}$ as in Fasen [24].

(iii) The convergence $\mathbf{S}_n^{(1)} \Longrightarrow \mathbf{S}^{(1)}$ as $n \to \infty$ holds in the Skorokhod J_1 topology. But $(\mathbf{S}_n^{(2)})_{n\in\mathbb{N}}$ does not converge in the Skorokhod J_1 topology; the proof is presented in Avram and Taqqu [3] and Remark 5.3 below; see also Remark 3.20 *(iii)* in Phillips and Solo [54] for a motivation in terms of the Beveridge-Nelson decomposition. Moreover, Avram and Taqqu [3] showed that in the one-dimensional case ($d = 1$) $(\mathbf{S}_n^{(2)})_{n\in\mathbb{N}}$ converges at least in the Skorokhod M_1 topology. However, in the multidimensional case, $d > 1$, the assumptions of Proposition 2.1 are not sufficient to obtain the convergence in the Skorokhod M_1 topology.

(iv) Similar results were proven in Meerschaert and Scheffler [45]. However, from Meerschaert and Scheffler [45] we can only follow under some slightly different assumptions the convergence of $(\mathbf{S}_n^{(i)}(t))_{n\in\mathbb{N}}$ for $i = 1, \ldots, 4$, $t \geq 0$, but not the joint convergence in the finite dimensional distributions, which we require for the forthcoming results of this paper. □

We continue with a corollary which gives, under some stronger assumptions, simple representations of $\mathbf{S}_i$, $i = 1, \ldots, 4$.

Corollary 2.3 *Let the assumption of Proposition 2.1 hold, and let* $0 < \alpha < 1$, *or* $1 < \alpha < 2$ *with* $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_w$. *Suppose that the components of the* $(q + d)$*-dimensional Lévy process* $((\mathbf{L}_1(t)', \mathbf{L}_2(t)')')_{t\geq 0}$ *are iid Lévy processes with distribution of* $(L(t))_{t\geq 0}$. *Furthermore, we assume that the tail balance condition*

$$\lim_{u\to\infty} \frac{\mathbb{P}(L(1) > u)}{\mathbb{P}(|L(1)| > u)} = \lim_{u\to\infty} \frac{\mathbb{P}(-L(1) > u)}{\mathbb{P}(|L(1)| > u)} = \frac{1}{2} \tag{2.5}$$

and

$$\lim_{n\to\infty} n\mathbb{P}(|L(1)| > a_n) = K_\alpha^{-1} \tag{2.6}$$

holds, where

$$K_\alpha = \begin{cases} \Gamma(1-\alpha)\cos\left(\frac{\pi\alpha}{2}\right) & \text{if } 0 < \alpha < 1, \\ \frac{\Gamma(2-\alpha)}{\alpha-1}|\cos(\frac{\pi\alpha}{2})| & \text{if } 1 < \alpha < 2. \end{cases}$$

Finally, suppose that $\mathbf{\Lambda} = diag(\lambda_1, \ldots, \lambda_d)$. *Define*

$$\mathbf{S}_{1,n}(t) = a_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \Delta\mathbf{L}_1(kh), \qquad \mathbf{S}_{2,n}(t) = a_n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \mathbf{Z}(kh),$$
$$\mathbf{S}_{3,n}(t) = a_n^{-2} \sum_{k=1}^{\lfloor nt \rfloor} \mathbf{Z}(kh)\mathbf{Z}(kh)', \qquad \mathbf{S}_{4,n}(t) = a_n^{-2} \sum_{k=1}^{\lfloor nt \rfloor} \mathbf{Z}((k+1)h)\mathbf{Z}(kh)'.$$

Then as $n \to \infty$,

$$((\mathbf{S}_{1,n}(t))_{t\geq 0}, \mathbf{S}_{2,n}(\mathbf{t}), \mathbf{S}_{3,n}(\mathbf{t}), \mathbf{S}_{4,n}(\mathbf{t})) \Longrightarrow ((\mathbf{S}_1(t))_{t\geq 0}, \mathbf{S}_2(\mathbf{t}), \mathbf{S}_3(\mathbf{t}), \mathbf{S}_4(\mathbf{t}))$$

in $(\mathbb{D}[0,1], \mathbb{R}^q) \times \mathbb{R}^{d\times l} \times \mathbb{R}^{(d\times d)\times l} \times \mathbb{R}^{(d\times d)\times l}$ *equipped with the product topology, where for* $t \geq 0$,

$$\mathbf{S}_1(t) = h^{\frac{1}{\alpha}}\mathbf{L}_1^*(t), \qquad \mathbf{S}_2(t) = E_{h\mathbf{\Lambda},1}^{-1} E_{h\mathbf{\Lambda},\alpha} D_{\mathbf{\Lambda},\alpha}\mathbf{L}_2^*(t),$$
$$\mathbf{S}_3(t) = E_{h\mathbf{\Lambda},2}^{-2} E_{h\mathbf{\Lambda},\alpha}^2 D_{\mathbf{\Lambda},\alpha}^2 [\mathbf{L}_2^*, \mathbf{L}_2^*]_t, \qquad \mathbf{S}_4(t) = \mathrm{e}^{-h\mathbf{\Lambda}}\mathbf{S}_3(t),$$

and $(\mathbf{L}_1^*(t))_{t\geq 0}$ *is a* q*-dimensional Lévy process independent of the* d*-dimensional Lévy process* $(\mathbf{L}_2^*(t))_{t\geq 0}$*. In both cases the components are iid* $S_\alpha(1,0,0)$*-stable Lévy motions. Finally,*

$$E_{h\mathbf{\Lambda},\alpha} = (\mathbf{I}_{d\times d} - \mathrm{e}^{-\alpha h\mathbf{\Lambda}})^{1/\alpha} \quad \text{and} \quad D_{\mathbf{\Lambda},\alpha} = diag((\alpha\lambda_1)^{-1/\alpha},\ldots,(\alpha\lambda_d)^{-1/\alpha}). \tag{2.7}$$

The Corollary is a direct conclusion of Proposition 2.1 and Samorodnitsky and Taqqu [62], Theorem 1.8.1, since for an $S_\alpha(1,0,0)$-stable random variable S_α the tail behavior

$$\lim_{n\to\infty} n\mathbb{P}(S_\alpha > n^{1/\alpha}) = \lim_{n\to\infty} n\mathbb{P}(-S_\alpha > n^{1/\alpha}) = K_\alpha^{-1}$$

holds.

2.3 Central limit results in DN(2)

On the other hand, if $\mathbf{L}_1(1)$ and $\mathbf{L}_2(1)$, respectively has at least a finite second moment and hence, $\mathbf{L}_1(1) \in \mathrm{DN}(2)$ and $\mathbf{L}_2(1) \in \mathrm{DN}(2)$, respectively, then we have the following result.

Proposition 2.4 *Let the multivariate cointegration model* (1.5)-(1.7) *be given.*

(a) Suppose $\mathbb{E}\|\mathbf{L}_1(1)\|^2 < \infty$ *and* $\mathbb{E}(\mathbf{L}_1(1)) = \mathbf{0}_q$*. Define* $\mathbf{\Omega}_1 = h\mathbb{E}(\mathbf{L}_1(1)\mathbf{L}_1(1)')$ *and* $a_n := n^{1/2}$*. Then as* $n\to\infty$*,*

$$\left(\mathbf{S}_{1,n}(t) := a_n^{-1}\sum_{k=1}^{\lfloor nt\rfloor}\Delta\mathbf{L}_1(kh)\right)_{t\geq 0} \Longrightarrow \left(\mathbf{\Omega}_1^{1/2}\mathbf{B}_1(t) =: \mathbf{S}_1(t)\right)_{t\geq 0} \tag{2.8}$$

in the Skorokhod J_1 *topology on* $(\mathbb{D}[0,1],\mathbb{R}^q)$*, where* $(\mathbf{B}_1(t))_{t\geq 0}$ *is a* q*-dimensional standard-Brownian motion.*

(b) Suppose $\mathbb{E}\|\mathbf{L}_2(1)\|^r < \infty$ *for some* $r > 2$ *and* $\mathbb{E}(\mathbf{L}_2(1)) = \mathbf{0}_d$*. Define*

$$\begin{aligned}\widetilde{\mathbf{\Omega}}_2 &:= \mathbb{E}(\mathbf{Z}(0)\mathbf{Z}(0)') = \int_0^\infty \mathrm{e}^{-\mathbf{\Lambda}s}\mathbb{E}(\mathbf{L}_2(1)\mathbf{L}_2(1)')\mathrm{e}^{-\mathbf{\Lambda}'s}\,\mathrm{d}s,\\ \mathbf{\Omega}_2 &:= \widetilde{\mathbf{\Omega}}_2 + \sum_{k=1}^\infty\left(\mathrm{e}^{-h\mathbf{\Lambda}k}\widetilde{\mathbf{\Omega}}_2 + \widetilde{\mathbf{\Omega}}_2\mathrm{e}^{-h\mathbf{\Lambda}'k}\right),\end{aligned}$$

and $b_n := n^{1/2}$*. If* $\mathbf{\Omega}_2$ *is invertible, then as* $n\to\infty$*,*

$$\left(\mathbf{S}_{2,n}(t) := b_n^{-1}\sum_{k=1}^{\lfloor nt\rfloor}\mathbf{Z}(kh)\right)_{t\geq 0} \Longrightarrow \left(\mathbf{\Omega}_2^{1/2}\mathbf{B}_2(t) =: \mathbf{S}_2(t)\right)_{t\geq 0} \tag{2.9}$$

in the Skorokhod J_1 *topology on* $(\mathbb{D}[0,1],\mathbb{R}^d)$*, where* $(\mathbf{B}_2(t))_{t\geq 0}$ *is a* d*-dimensional standard-Brownian motion. Furthermore,*

$$\mathbf{S}_{3,n}(1) := b_n^{-2}\sum_{k=1}^n \mathbf{Z}(kh)\mathbf{Z}(kh)' \overset{n\to\infty}{\longrightarrow} \widetilde{\mathbf{\Omega}}_2 =: \mathbf{S}_3(1) \quad \mathbb{P}\text{-a.s.} \tag{2.10}$$

on $\mathbb{R}^{d\times d}$*. If* $\mathbb{E}(\mathbf{L}_2(1)\mathbf{L}_2(1)') = \mathbf{I}_{d\times d}$ *and* $\mathbf{\Lambda} = diag(\lambda_1,\ldots,\lambda_d)$ *then* $\widetilde{\mathbf{\Omega}}_2^{1/2} = D_{\mathbf{\Lambda},2}$ *and* $\mathbf{\Omega}_2^{1/2} = E_{h\mathbf{\Lambda},1}^{-1}E_{h\mathbf{\Lambda},2}D_{\mathbf{\Lambda},2}$ *with the notation in (2.7).*

(c) *Let* $\mathbf{L} = ((\mathbf{L}_1(t)', \mathbf{L}_2(t)')')_{t\geq 0}$ *be a* $(q+d)$*-dimensional Lévy process. Suppose* $\mathbb{E}\|\mathbf{L}(1)\|^r < \infty$ *for some* $r > 2$ *and* $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_{q+d}$*. Define*

$$\begin{aligned} \boldsymbol{\Omega}^* &:= \int_0^h diag(\mathbf{I}_{q\times q}, \mathrm{e}^{-\boldsymbol{\Lambda} s})\mathbb{E}(\mathbf{L}(1)\mathbf{L}(1)')diag(\mathbf{I}_{q\times q}, \mathrm{e}^{-\boldsymbol{\Lambda}' s})\,\mathrm{d}s, \\ \boldsymbol{\Omega}^{1/2} &:= \boldsymbol{\Omega}^{*1/2} + \sum_{k=1}^{\infty} diag(\mathbf{0}_{q\times q}, \mathrm{e}^{-h\boldsymbol{\Lambda} k})\boldsymbol{\Omega}^{*1/2}. \end{aligned}$$

If $\boldsymbol{\Omega}$ *is invertible, then as* $n \to \infty$*,*

$$\left((\mathbf{S}_{1,n}(t)', \mathbf{S}_{2,n}(t)')'\right)_{t\geq 0} \Longrightarrow \left(\boldsymbol{\Omega}^{1/2}\mathbf{B}(t)\right)_{t\geq 0} \tag{2.11}$$

in the Skorokhod J_1 *topology on* $(\mathbb{D}[0,1], \mathbb{R}^{q+d})$*, where* $(\mathbf{B}(t))_{t\geq 0}$ *is a* $(q+d)$*-dimensional standard-Brownian motion. Furthermore,* $(\boldsymbol{\Omega}^{1/2}\mathbf{B}(t))_{t\geq 0} = ((\mathbf{S}_1(t)', \mathbf{S}_2(t)')')_{t\geq 0}$ *with* $\mathbf{S}_1$ *and* $\mathbf{S}_2$ *from (a) and (b), respectively.*

3 Consistency and asymptotic behavior of the estimators

The main results of this paper satisfy either the next Assumption 3.1 allowing that $\mathbf{L}_1$ and $\mathbf{L}_2$ are in different domains of attractions but are independent, or Assumption 3.2 and 3.3, respectively, where $(\mathbf{L}_1(1)', \mathbf{L}_2(1)')' \in \mathrm{DN}(\alpha)$, $\alpha \in (0, 2]$.

Assumption 3.1
Let $\mathbf{L}_1$ *and* $\mathbf{L}_2$ *be independent Lévy processes of dimension* q *and* d*, respectively. Furthermore, suppose the following:*

(a) *Either*

$\mathbf{L}_1(1) \in \mathcal{R}_{-\alpha}(a_n, \mu_1)$, $0 < \alpha < 2$, $\mathbb{E}(\mathbf{L}_1(1)) = \mathbf{0}_q$ *if* $1 < \alpha < 2$ *and* $\mathbf{L}_1(1)$ *is symmetric if* $\alpha = 1$*. Furthermore,* $\mathbf{S}_1$ *is defined as in Proposition 2.1, i.e.* $\mathbf{S}_1$ *is an* α*-stable Lévy process with Lévy measure* μ_1*, and if* $1 < \alpha < 2$ *then* $\mathbb{E}(\mathbf{S}_1(1)) = \mathbf{0}_q$*. We suppose that* $\mu_1(\{\mathbf{0}_{i-1}\} \times \overline{\mathbb{R}}\backslash\{0\} \times \{\mathbf{0}_{q-i}\}) > 0$ *for* $i = 1, \ldots, q$*.*

Or

$\mathbb{E}\|\mathbf{L}_1(1)\|^2 < \infty$ *and* $\mathbb{E}(\mathbf{L}_1(1)) = \mathbf{0}_q$*. Define* $a_n := n^{1/2}$ *and* $\alpha := 2$*. Furthermore,* $\mathbf{S}_1$ *is defined as in Proposition 2.4 (a), i.e.* $\mathbf{S}_1$ *is a Brownian motion with covariance matrix* $\boldsymbol{\Omega}_1$*, and we suppose that* $\boldsymbol{\Omega}_1$ *is invertible.*

(b) *Either*

$\mathbf{L}_2(1) \in \mathcal{R}_{-\beta}(b_n, \mu_2)$, $0 < \beta < 2$, $\mathbb{E}(\mathbf{L}_2(1)) = \mathbf{0}_d$ *if* $1 < \beta < 2$ *and* $\mathbf{L}_2(1)$ *is symmetric if* $\beta = 1$*. Furthermore,* $\mathbf{S}_2, \mathbf{S}_3$ *are defined as in Proposition 2.1, i.e.* $\mathbf{S}_2$ *is a* β*-stable Lévy process with Lévy measure*

$$\mu_2^*(\cdot) = h\mathbb{E}\left(\mu_2\left\{\mathbf{x} \in \overline{\mathbb{R}}^d\backslash\{\mathbf{0}_d\} : \left(\sum_{i=0}^{\infty} \mathrm{e}^{-h\boldsymbol{\Lambda} i}\mathrm{e}^{-h\boldsymbol{\Lambda} U}\right)\mathbf{x} \in \cdot\right\}\right)$$

where U is a uniform random variable on $(0,1)$, and if $1 < \beta < 2$ then $\mathbb{E}(\mathbf{S}_2(1)) = \mathbf{0}_d$.

Or

$\mathbb{E}\|\mathbf{L}_2(1)\|^r < \infty$ for some $r > 2$ and $\mathbb{E}(\mathbf{L}_2(1)) = \mathbf{0}_d$. Define $b_n := n^{1/2}$ and $\beta := 2$. Furthermore, $\mathbf{S}_2, \mathbf{S}_3(1)$ are defined as in Proposition 2.4 (b), i.e. $\mathbf{S}_2$ is a Brownian motion with covariance matrix $\mathbf{\Omega}_2$ and $\mathbf{S}_3(1) = \mathbb{E}(\mathbf{Z}(0)\mathbf{Z}(0)')$.

Finally, $\mathbf{S}_1$ and $(\mathbf{S}_2, \mathbf{S}_3(1))$ are independent.

If $(\mathbf{L}_1, \boldsymbol{\zeta})$ and $\mathbf{L}_2$ are independent then the long-run equilibrium $\mathbf{AY}$ and the short-run equilibrium $\mathbf{Z}$ are independent which is a somewhat natural assumption. However, the next assumptions show that dependence between $\mathbf{L}_1$ and $\mathbf{L}_2$ is also allowed.

Assumption 3.2
Let $\mathbf{L} = ((\mathbf{L}_1(t)', \mathbf{L}_2(t)')')_{t\geq 0}$ be a $(q+d)$-dimensional Lévy process. Suppose $\mathbb{E}\|\mathbf{L}(1)\|^r < \infty$ for some $r > 2$ and $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_{q+d}$. Furthermore, $a_n := b_n := n^{1/2}$ and $\alpha := \beta := 2$. Finally, $\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3$ are given as in Proposition 2.4, i.e. $(\mathbf{S}_1', \mathbf{S}_2')'$ is a Brownian motion with covariance matrix $\mathbf{\Omega}$ and $\mathbf{S}_3(1) = \mathbb{E}(\mathbf{Z}(0)\mathbf{Z}(0)')$. We suppose that the covariance matrix $\mathbf{\Omega}_1$ of $\mathbf{S}_1$ is invertible.

Assumption 3.3
Let $\mathbf{L} = ((\mathbf{L}_1(t)', \mathbf{L}_2(t)')')_{t\geq 0}$ be a $(q+d)$-dimensional Lévy process. Suppose $(\mathbf{L}_1(1)', \mathbf{L}_2(1)')' \in \mathcal{R}_{-\alpha}(a_n, \mu)$, $0 < \alpha < 2$, $\mathbb{E}(\mathbf{L}_1(1)', \mathbf{L}_2(1)')' = \mathbf{0}_{q+d}$ if $1 < \alpha < 2$ and $(\mathbf{L}_1(1)', \mathbf{L}_2(1)')$ is symmetric if $\alpha = 1$. Furthermore, $b_n := a_n$ and $\beta := \alpha$. Finally, $\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3$ are given as in Proposition 2.1 with $\Sigma_1 := (\mathbf{I}_{q\times q}, \mathbf{0}_{q\times d})$ and $\Sigma_2 := (\mathbf{0}_{d\times q}, \mathbf{I}_{d\times d})$. We suppose that $\mu_1(\{\mathbf{0}_{i-1}\} \times \overline{\mathbb{R}}\backslash\{0\} \times \{\mathbf{0}_{q-i}\}) > 0$ for $i = 1, \ldots, q$.

These assumptions lead to the following asymptotic behavior of $\widehat{\mathbf{A}}_n$.

Theorem 3.4 *Let the multivariate cointegration model (1.5)-(1.7) be given, and let either of the Assumptions 3.1 – 3.3 hold. Furthermore, let the following conditions be satisfied:*

(i) $\sum_{k=1}^n \mathbf{Z}(kh)\boldsymbol{\zeta}(kh)' = o_p(a_n b_n)$ *as* $n \to \infty$.

(ii) $\sum_{k=1}^n \mathbf{L}_1(kh)\boldsymbol{\zeta}(kh)' = o_p(na_n^2)$ *as* $n \to \infty$.

(iii) $\sum_{k=1}^n \boldsymbol{\zeta}(kh)\boldsymbol{\zeta}(kh)' = o_p(na_n^2)$ *as* $n \to \infty$.

Then $\widehat{\mathbf{A}}_n$ as given in (1.8) satisfies as $n \to \infty$,

$$na_n b_n^{-1}(\widehat{\mathbf{A}}_n - \mathbf{A}) \implies \left(\mathbf{S}_2(1)\mathbf{S}_1(1)' - \int_0^1 \mathbf{S}_2(s-)\mathrm{d}\mathbf{S}_1(s)'\right)\left(\int_0^1 \mathbf{S}_1(s)\mathbf{S}_1(s)'\mathrm{d}s\right)^{-1} =: \mathbf{G}.$$

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

The conditions *(i)-(iii)* are very general in the sense that we do not require $\mathbf{L}_1, \mathbf{Z}$ and $\boldsymbol{\zeta}$ to be independent.

Example 3.5

(a) Let $\mathbf{Y}(t) = \mathbf{L}_1(t)$, $t \geq 0$, be a classical q-dimensional Lévy process, i.e. $\boldsymbol{\zeta}(t) = \mathbf{0}_q$ for $t \geq 0$. Then *(i)-(iii)* are satisfied. However, $\mathbf{Y}$ and $\mathbf{Z}$ can still be dependent, if Assumption 3.2 and 3.3, respectively hold. In the finite second moment case of Assumption 3.2 the dependence is measured by the covariance matrix of $\mathbf{L}(1)$. If Assumtion 3.3 holds, the dependence in extremes of $\mathbf{L}_1$ and $\mathbf{L}_2$ is measured by μ.

(b) Let either of Assumption 3.1 – 3.3 hold, and let

$$\boldsymbol{\zeta}(t) = \mathbf{B}\mathbf{O}(t), \quad t \geq 0,$$

where $\mathbf{B}$ is a random matrix in $\mathbb{R}^{q\times m}$ independent of the stationary multivariate Ornstein-Uhlenbeck process $\mathbf{O} = (\mathbf{O}(t))_{t\geq 0}$ in $\mathbb{R}^m$ and the Lévy process $\mathbf{L} = (\mathbf{L}_1', \mathbf{L}_2')'$. Suppose that the Ornstein-Uhlenbeck process $\mathbf{O}$ is driven by the Lévy process $\mathbf{L}_3$, which is independent of $\mathbf{L}$. If $\mathbf{L}_1(1) \in \mathcal{R}_{-\alpha}(a_n)$, we assume

$$\lim_{x\to\infty} \frac{\mathbb{P}(\|\mathbf{L}_3(1)\| > x)}{\mathbb{P}(\|\mathbf{L}_1(1)\| > x)} = C \in [0, \infty)\,,$$

and if $\mathbb{E}\|\mathbf{L}_1(1)\|^2 < \infty$, we assume $\mathbb{E}\|\mathbf{L}_3(1)\|^r < \infty$ for some $r > 2$. Then *(i)-(iii)* also hold. This structure of the noise term $(\boldsymbol{\zeta}(t))_{t\geq 0}$ is flexible and captures, in particular, multivariate continuous-time ARMA (CARMA) processes (cf. Marquardt and Stelzer [41]). □

Remark 3.6

(a) Since the limit distribution $\mathbf{G}$ depends on $\mathbf{S}_2$ it depends, in particularly, on the nuisance parameter $\boldsymbol{\Lambda}$, which is plugged in to the characteristic triplet of $\mathbf{S}_2$ (the Lévy measure μ_2^* and the covariance matrix $\boldsymbol{\Omega}_2$, respectively).

(b) The norming sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$, respectively, depend on the domain of attraction of $\mathbf{L}_1(1)$ and $\mathbf{L}_2(1)$, respectively, which are in general not known, and determine the convergence rate of the least squares estimator. Therefore we will introduce the t-statistic which is independent of $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$. However, if $\mathbf{L}_1(1) \in \mathrm{DN}(\alpha)$ and $\mathbf{L}_2(1) \in \mathrm{DN}(\alpha)$ for some $\alpha \in (0, 2]$, then $a_n = b_n$. In this case we have $n(\widehat{\mathbf{A}}_n - \mathbf{A}) \Longrightarrow \mathbf{G}$ and $\widehat{\mathbf{A}}_n \xrightarrow{\mathbb{P}} \mathbf{A}$ as $n \to \infty$.

(c) Suppose $\mathbf{L}_1(1) \in \mathcal{R}_{-\alpha}(a_n)$ and $\mathbf{L}_2(1) \in \mathcal{R}_{-\beta}(b_n)$ with $\beta > \alpha$ or $\mathbb{E}\|\mathbf{L}_2(1)\|^r < \infty$ for some $r > 2$. Then $(\mathbf{L}_1(1)', \mathbf{L}_2'(1))' \in \mathrm{DN}\ (\alpha)$. A conclusion of Theorem 3.4 is that $n(\widehat{\mathbf{A}}_n - \mathbf{A}) \Longrightarrow \mathbf{0}_{d\times q}$ as $n \to \infty$, which is unsatisfactory if one wants to compute asymptotic confidence intervals for the components of $\mathbf{A}$. If additionally $\mathbf{L}_1$ and $\mathbf{L}_2$ are independent, then Assumption 3.1 is satisfied and $na_nb_n^{-1}(\widehat{\mathbf{A}}_n - \mathbf{A}) \Longrightarrow \mathbf{G}$ as $n \to \infty$, where $\mathbf{G} \neq \mathbf{0}_{d\times q}$ $\mathbb{P}$-a.s. Thus, we are able to compute asymptotic confidence intervals for the components of $\mathbf{A}$. We conjecture that this result (in particular $\mathbf{S}_1$ and $\mathbf{S}_2$ independent) also holds under some mild technical assumptions for dependent $\mathbf{L}_1$ and $\mathbf{L}_2$. The distribution of $(\mathbf{L}_1(1)', \mathbf{L}_2(1)')$ has at least to be operator-stable (cf. Paulauskas and Rachev [48]).

(d) The integral $\int_0^t \mathbf{S}_2(s)\mathrm{d}\mathbf{S}_1(s)'$ in the representation of $\mathbf{G}$ already suggests that we show in the proof the convergence of stochastic integrals. However, the well known results of Kurtz and Protter [35] require that the integrand and integrator converge weakly in the Skorokhod J_1 topology.

As mentioned in Remark 2.2 the sequence $(\mathbf{S}_{2,n})_{n\in\mathbb{N}}$ does not converge in the Skorokhod J_1 topology such that we can not use these results directly. □

In the following we comment on relations of our results to those in the literature.

Remark 3.7

(i) If Assumption 3.2 holds, then our model is a special case of Phillips and Durlauf [53]. Phillips and Durlauf [53] investigate the discrete time model

$$\mathbf{X}_n = \mathbf{A}\mathbf{Y}_n + \boldsymbol{\varepsilon}_n^{(1)} \quad \text{and} \quad \mathbf{Y}_n = \mathbf{Y}_{n-1} + \boldsymbol{\varepsilon}_n^{(2)}, \quad n \in \mathbb{N}, \tag{3.1}$$

where the strongly mixing noise sequence $\{(\boldsymbol{\varepsilon}_n^{(1)}, \boldsymbol{\varepsilon}_n^{(2)})\}_{n\in\mathbb{N}}$ has at least finite r-moment for some $r > 2$.

An extension of these results to infinite second moments was given in Paulauskas and Rachev [48]. However, they restricted themselves to the case that $\{(\boldsymbol{\varepsilon}_n^{(1)}, \boldsymbol{\varepsilon}_n^{(2)})\}_{n\in\mathbb{N}}$ forms an iid sequence and derived only the asymptotic behavior of the least squares estimator $\widehat{\mathbf{A}}_n$ without going into detail into the structure of the limit distribution and to test statistics. A detailed analysis of the one-dimensional model of Paulauskas and Rachev [48] was done by Mittnik et al. [46].

(ii) Other models, which also allow an infinite variance of the noise term $(\mathbf{Z}_n)_{n\in\mathbb{N}}$ are, e.g., the regression model of Caner [10] of the form

$$\mathbf{X}_n = \mathbf{A}\mathbf{X}_{n-1} + \mathbf{Z}_n, \qquad \text{where} \qquad \mathbf{Z}_n = \sum_{k=0}^{\infty} \mathbf{C}_k \boldsymbol{\varepsilon}_{n-k}, \quad n \in \mathbb{N}, \tag{3.2}$$

is a stationary MA-process with $\mathbf{C}_k \in \mathbb{R}^{d\times d}$, and $(\boldsymbol{\varepsilon}_k)_{k\in\mathbb{Z}}$ is a sequence of iid symmetric d-dimensional random vectors with independent components and $\boldsymbol{\varepsilon}_1 \in \mathcal{R}_{-\alpha}(a_n)$. However, in that model they used as hypothesis only $\mathbf{A} = \mathbf{I}_{d\times d}$ which means that the model is not cointegrated, and they test for unit roots. The techniques of our paper can straightforwardly be applied to Caner's [10] model to avoid the assumption of iid symmetric components and will be presented in some future work. Note that Paulauskas et al. [47] pointed out a gap in Caner's proof of Theorem 2, although the other results of that paper are still valid. The one-dimensional case of (3.2) was already studied in Phillips [51] and Chan and Tran [12].

In the unit root model (3.2) different results apply than in our cointegrated model. For a survey on unit root models we refer to Chan [11]. □

Corollary 3.8 *Let the assumptions of Theorem 3.4 be satisfied, and let* $\mathbf{X}^*(t) = \mathbf{A}^*\mathbf{Y}(t) + \mathbf{B}^*\mathbf{Z}(t)$, $t \geq 0$, *where* $\mathbf{A}^* \in \mathbb{R}^{m\times q}$ *is deterministic and* $\mathbf{B}^*$ *in* $\mathbb{R}^{m\times d}$ *is a random matrix. Then the least squares estimator* $\widehat{\mathbf{A}}_n^* = \mathbb{X}_n^{*\prime}\mathbb{Y}_n(\mathbb{Y}_n'\mathbb{Y}_n)^{-1}$ *of* $\mathbf{A}^*$ *satisfies as* $n \to \infty$,

$$na_nb_n^{-1}(\widehat{\mathbf{A}}_n^* - \mathbf{A}^*) \implies \mathbf{B}^* \left(\mathbf{S}_2(1)\mathbf{S}_1(1)' - \int_0^1 \mathbf{S}_2(s-)\mathrm{d}\mathbf{S}_1(s)' \right) \left(\int_0^1 \mathbf{S}_1(s)\mathbf{S}_1(s)'\mathrm{d}s \right)^{-1}.$$

In particular, $\widehat{\mathbf{A}}_n^* \xrightarrow{\mathbb{P}} \mathbf{A}^*$ *as* $n \to \infty$ *if* $\beta > \alpha/(\alpha+1)$.

Remark 3.9 The model shows that we can take the short-term equilibrium more general than an Ornstein-Uhlenbeck process namely as $(\mathbf{B}^*\mathbf{Z}(t))_{t\geq 0}$ in our multivariate cointegration model, and still obtain consistency if $\beta > \alpha/(\alpha+1)$, and the asymptotic convergence of the least squares estimator $\widehat{\mathbf{A}}_n^*$ to a functional of stable Lévy processes and Brownian motions, whose characteristic triplets are known.

If we furthermore choose $\boldsymbol{\zeta}(t) = \mathbf{B}\mathbf{O}(t)$, $t \geq 0$, as in Example 3.5, then both $\mathbf{X}$ and $\mathbf{Y}$ are of the form Lévy process plus an additional Ornstein-Uhlenbeck noise which is multiplied by a random matrix. This means that $\mathbf{X}$ and $\mathbf{Y}$ are in the same class of processes.

The class of multivariate continuous-time processes with a representation $(\mathbf{B}^*\mathbf{Z}(t))_{t\geq 0}$ is huge and includes, in particularly, multivariate CARMA models. Moreover, the components of $(\mathbf{B}^*\mathbf{Z}(t))$ can be sums of dependent or independent Ornstein-Uhlenbeck processes, or CARMA processes as well. Furthermore, if $\mathbf{L}_2$ is a multivariate Brownian motion, then the distribution of $\mathbf{B}^*\mathbf{Z}(t)$ is a scaled mixtures of normals.

More complex noise terms than $(\mathbf{B}^*\mathbf{Z}(t))_{t\geq 0}$ will raise the problem that the characteristic triplet of the β-stable Lévy motion $(\mathbf{B}^*\mathbf{S}_2(t))_{t\geq 0}$ becomes analytically complex and, hence, the simulation of asymptotic confidence intervals for the components of $\mathbf{A}$ will be involved. □

Particularly, useful for the practical simulation of asymptotic confidence intervals is the next result.

Corollary 3.10 *Let the assumptions of Theorem 3.4 be satisfied. Furthermore, suppose that* $\mathbf{L}_1$ *and* $\mathbf{L}_2$ *are independent, and* $\boldsymbol{\Lambda} = diag(\lambda_1, \ldots, \lambda_d)$. *Assume that either* $\mathbf{L}_1(1) \in \mathcal{R}_{-\alpha}(a_n, \mu_1)$, $0 < \alpha < 2$, $\alpha \neq 1$, *with iid components satisfying (2.5) and (2.6), or* $\mathbb{E}(\mathbf{L}_1(1)\mathbf{L}_1(1)') = \mathbf{I}_{q\times q}$. *Similarly assume that either* $\mathbf{L}_2(1) \in \mathcal{R}_{-\beta}(b_n, \mu_2)$, $0 < \beta < 2$, $\beta \neq 1$, *with iid components satisfying (2.5) and (2.6) (with* b_n *instead of* a_n, *and* β *instead of* α*), or* $\mathbb{E}(\mathbf{L}_2(1)\mathbf{L}_2(1)') = \mathbf{I}_{d\times d}$. *Then as* $n \to \infty$,

$$na_nb_n^{-1}h^{\frac{1}{\alpha}}E_{h\boldsymbol{\Lambda},1}E_{h\boldsymbol{\Lambda},\beta}^{-1}D_{\boldsymbol{\Lambda},\beta}^{-1}(\widehat{\mathbf{A}}_n - \mathbf{A}) \Longrightarrow \left(\int_0^1 \mathbf{L}_1^*(s-)\mathrm{d}\mathbf{L}_2^*(s)'\right)' \left(\int_0^1 \mathbf{L}_1^*(s)\mathbf{L}_1^*(s)'\mathrm{d}s\right)^{-1} =: \mathbf{G}^*$$

where $\mathbf{L}_1^*$ *is a q-dimensional Lévy process with components which are iid* $S_\alpha(1,0,0)$*-stable Lévy motions, independent of* $\mathbf{L}_2^*$ *a d-dimensional Lévy process with components which are iid* $S_\beta(1,0,0)$*-stable Lévy motions. Finally,*

$$E_{h\boldsymbol{\Lambda},\beta} = (\mathbf{I}_{d\times d} - \mathrm{e}^{-\beta h\boldsymbol{\Lambda}})^{1/\beta} \quad \text{and} \quad D_{\boldsymbol{\Lambda},\beta} = diag((\beta\lambda_1)^{-1/\beta}, \ldots, (\beta\lambda_d)^{-1/\beta}).$$

Remark 3.11

(i) In the model of Corollary 3.10 the components of $\mathbf{Z}$ are independent (uncorrelated, respectively), one-dimensional Ornstein-Uhlenbeck processes. Moreover, if $\boldsymbol{\zeta}(t) = \mathbf{0}_q$ for $t \geq 0$, then $\mathbf{Z}$ and $\mathbf{Y}$ are independent, and $\mathbf{Y}$ consists of independent and identically distributed components.

(ii) The result of Corollary 3.10 shows very nicely the influence of the nuisance parameter $\boldsymbol{\Lambda}$ on the limit result. The limit distribution $\mathbf{G}^*$ depends only on α and β. The parameter $\boldsymbol{\Lambda}$ influences the deterministic matrix $E_{h\boldsymbol{\Lambda},1}E_{h\boldsymbol{\Lambda},\beta}^{-1}D_{\boldsymbol{\Lambda},\beta}^{-1}$ and can be estimated as in Fasen [24].

(iii) Suppose we investigate the observation scheme

$$\mathbb{X}'_n = (\mathbf{X}(h_n), \dots, \mathbf{X}(nh_n)) \in \mathbb{R}^{d\times nh_n}, \qquad \mathbb{Y}'_n = (\mathbf{Y}(h_n), \dots, \mathbf{Y}(nh_n)) \in \mathbb{R}^{q\times nh_n}$$

with grid distance $h_n \longrightarrow 0$ and $nh_n \longrightarrow \infty$ as $n \to \infty$. Since

$$(na_n b_n^{-1} h_n^{\frac{1}{\alpha}} E_{h_n\mathbf{\Lambda},1} E^{-1}_{h_n\mathbf{\Lambda},\beta} D^{-1}_{\mathbf{\Lambda},\beta})(nh_n a_{nh_n} b^{-1}_{nh_n} \mathbf{\Lambda})^{-1} = \mathbf{I}_{d\times d},$$

this suggests that as $n \to \infty$,

$$nh_n a_{nh_n} b^{-1}_{nh_n} \mathbf{\Lambda}(\widehat{\mathbf{A}}_n - \mathbf{A}) \Longrightarrow \mathbf{G}^*.$$

That case is studied in detail in Fasen [23]. □

Example 3.12 To obtain the asymptotic $1-p-$confidence intervals of the components of $\mathbf{A}$ for some $p \in (0,1)$, we have simulated the $1-\frac{p}{2}-$quantiles of the components of $\mathbf{G}^*$ from Corollary 3.10 in Table 1 by 100.000 Monte Carlo simulations using the toolbox STABLE of Robust Analysis Inc, where $\mathbf{L}_1^*$ and $\mathbf{L}_2^*$ are both multivariate $S_\alpha(1,0,0)$-stable Lévy motions $(0 < \alpha \leq 2)$ of dimension q and d, respectively (i.e. $\alpha = \beta$). Then $\mathbf{G}^*$ is a random matrix whose components are identically distributed, since $(\mathbf{L}_1^*(t)'(\int_0^1 \mathbf{L}_1^*(s)\mathbf{L}_1^*(s)'\mathrm{d}s)^{-1})_{t\geq 0}$ has identically distributed components. Hence, the $1-\frac{p}{2}-$quantiles given here are the quantiles of any component of $\mathbf{G}^*$.

Let $x_p(\alpha)$ denote the $1-\frac{p}{2}$ quantile of a component of $\mathbf{G}^*$ which depends on the dimension d. Moreover, we see that if α decreases $x_p(\alpha)$ is increasing which reflects similarly to the sample paths behavior that the heavy tails of $\mathbf{S}_1$ and $\mathbf{S}_2$ are transferred to $\mathbf{G}^*$ if $\alpha < 2$. This is also confirmed by the huge quantiles for small α's.

Next, we compute the asymptotic $1-p-$confidence interval of $\mathbf{A}_{ij}$, the (i,j) component of $\mathbf{A} = (\mathbf{A}_{ij})_{i,j=1,\dots,d}$ if $a_n = b_n$, which is

$$\Big[(\widehat{\mathbf{A}}_n)_{ij} - \frac{x_p(\alpha)}{n} h^{-\frac{1}{\alpha}} (\alpha\lambda_i)^{-\frac{1}{\alpha}} (1-\mathrm{e}^{-h\alpha\lambda_i})^{\frac{1}{\alpha}} (1-\mathrm{e}^{-h\lambda_i})^{-1},$$
$$(\widehat{\mathbf{A}}_n)_{ij} + \frac{x_p(\alpha)}{n} h^{-\frac{1}{\alpha}} (\alpha\lambda_i)^{-\frac{1}{\alpha}} (1-\mathrm{e}^{-h\alpha\lambda_i})^{\frac{1}{\alpha}} (1-\mathrm{e}^{-h\lambda_i})^{-1}\Big].$$

If h is small then $(h\alpha\lambda_i)^{-1/\alpha}(1-\mathrm{e}^{-h\alpha\lambda_i})^{1/\alpha} \approx 1$ such that for decreasing α, the confidence intervals are getting larger, which results in larger statistical uncertainty.

4 Hypothesis testing

4.1 *t*-ratio statistic

In the following we define

$$\widehat{\mathbf{\Omega}}_n \;=\; n^{-1}(\mathbb{X}'_n - \widehat{\mathbf{A}}_n \mathbb{Y}'_n)(\mathbb{X}'_n - \widehat{\mathbf{A}}_n \mathbb{Y}'_n)'$$

which is under the assumption $\mathbb{E}\|\mathbf{L}_2(1)\|^2 < \infty$ an estimator for the covariance matrix $\widetilde{\mathbf{\Omega}}_2 = \mathbb{E}(\mathbf{Z}(0)\mathbf{Z}(0)')$. Furthermore, in a classical linear model with iid standard-normal noise $(\mathbf{Z}(k))_{k\in\mathbb{N}}$

$d=1$	$\alpha=0.9$	$\alpha=1.2$	$\alpha=1.4$	$\alpha=1.5$	$\alpha=1.6$	$\alpha=1.8$	$\alpha=2$
$p=0.1$	12.5	7.0	5.5	5.1	4.8	4.2	3.8
$p=0.05$	27.9	12.5	8.9	7.8	7.1	5.7	5.0
$p=0.025$	60.8	22.1	14.3	11.9	10.2	7.6	6.2
$p=0.01$	169.5	45.0	26.0	20.0	16.6	11.0	7.9
$d=2$	$\alpha=0.9$	$\alpha=1.2$	$\alpha=1.4$	$\alpha=1.5$	$\alpha=1.6$	$\alpha=1.8$	$\alpha=2$
$p=0.1$	17.2	9.6	7.7	6.9	6.4	5.5	5.0
$p=0.05$	39.0	17.1	12.1	10.4	9.3	7.5	6.5
$p=0.025$	88.6	30.1	19.0	15.6	13.1	9.8	7.9
$p=0.01$	255.0	64.5	34.8	27.2	21.6	13.7	9.8

Table 1: Simulated $1-\frac{p}{2}$-quantiles of the components of $\mathbf{G}^*$ for any $q \in \mathbb{N}$ and for different choices of α and d.

the matrix $\left(n^{-1}\mathbb{Y}_n'\mathbb{Y}_n\right)^{-1} \otimes \widehat{\mathbf{\Omega}}_n$ is an estimator for the covariance of the least squares estimator and is used as standardization in the t-ratio statistic to get an estimator with a covariance matrix which is the identity matrix. In the setup of a model with infinite variance the sequence of random matrices $\left(n^{-1}\mathbb{Y}_n'\mathbb{Y}_n\right)^{-1} \otimes \widehat{\mathbf{\Omega}}_n$ do not converge to a finite random matrix. However, as usual we use this sequence as standardization in the t-ratio statistic.

Theorem 4.1 *Let the assumptions of Theorem 3.4 hold and suppose* $\mathbb{P}(\det(\mathbf{S}_3(1)) = 0) = 0$. *Then as* $n \to \infty$,

$$t_{\widehat{\mathbf{A}}_n} = (\widehat{\mathbf{\Omega}}_n)^{-1/2}(\widehat{\mathbf{A}}_n - \mathbf{A})(\mathbb{Y}_n'\mathbb{Y}_n)^{1/2} \Longrightarrow \mathbf{S}_3(1)^{-1/2}\mathbf{G}\left(\int_0^1 \mathbf{S}_1(s)\mathbf{S}_1(s)'\mathrm{d}s\right)^{1/2}.$$

Let the stronger assumptions of Corollary 3.10 hold. Then as $n \to \infty$,

$$E_{h\mathbf{\Lambda},1}E_{h\mathbf{\Lambda},2}^{-1}t_{\widehat{\mathbf{A}}_n} \Longrightarrow [\mathbf{L}_2^*, \mathbf{L}_2^*]_1^{-1/2}\left(\int_0^1 \mathbf{L}_1^*(s-)\mathrm{d}\mathbf{L}_2^*(s)'\right)'\left(\int_0^1 \mathbf{L}_1^*(s)\mathbf{L}_1^*(s)'\mathrm{d}s\right)^{-1/2} =: \mathbf{G}^{**}.$$

Remark 4.2

(i) The t-ratio statistic has the advantage that the limit distribution does not depend on the regression order if $\mathbf{A} \neq \mathbf{0}_{d\times q}$ and $(d,q) = (1,1)$, i.e. if we regress $\mathbf{X}$ after $\mathbf{Y}$ or $\mathbf{Y}$ after $\mathbf{X}$ the asymptotic error distribution is the same. Hence, it makes no difference, if we test if $\mathbf{X}$ depends on $\mathbf{Y}$, or vice versa, which is natural.

(ii) Moreover, a goal of the t-ratio statistic is that we do not need the norming sequence $(na_nb_n^{-1})_{n\in\mathbb{N}}$, which is unknown anyway, and which depends on α and β.

(iii) For $\alpha = 2$ we know that $[\mathbf{L}_2^*, \mathbf{L}_2^*]_1 = \mathbf{I}_{d\times d}$. Thus, if $\alpha = \beta = 2$ then the conditional distribution of $\int_0^1 \mathbf{L}_1^*(s-)\mathrm{d}\mathbf{L}_2^*(s)'$ under $\mathbf{L}_1^*$ is a multivariate normal distribution with covariance matrix $\int_0^1 \mathbf{L}_1^*(s-)\mathbf{L}_1^*(s-)'\mathrm{d}s$ such that $\mathbf{G}^{**}$ is a multivariate standard normal distribution. □

Example 4.3 In Table 2 we present the simulated $1-\frac{p}{2}-$quantiles of $\mathbf{G}^{**}$ of Theorem 4.1, where both $\mathbf{L}_1^*$ and $\mathbf{L}_2^*$ are $S_\alpha(1,0,0)$-Lévy processes $(0<\alpha\leq 2)$, i.e. $\alpha=\beta$, based again on 100.000 Monte-Carlo simulations. The distribution of the components of $\mathbf{G}^{**}$ is independent of the dimension d and q.

The simulations suggest that high level quantiles of the components of $\mathbf{G}^{**}$ do not depend on α. In particular, we obtain the same high level quantiles of $\mathbf{G}^{**}$ for $\alpha=2$ (the Gaussian case) and $0<\alpha<2$ (the usual stable case). As noted in Remark 4.2 *(iii)*, if $\alpha=2$, then $\mathbf{G}^{**}$ is a multivariate standard normal distribution. Hence, in Table 2 we see the quantiles of a standard normal distribution. Already Mittnik et al. [46] observed a similar phenomena. For $\alpha<2$ the explanation of this result is not obvious and outside the scope of the present paper.

However, the independence of high level quantiles of $\mathbf{G}^{**}$ from α is very useful for statistical purpose because it shows that the confidence intervals do not depend on the model parameter α. Let $x_p(\alpha)$ denote the $1-\frac{p}{2}-$quantile of $\mathbf{G}^{**}$ and suppose $(d,q)=(1,1)$ with $\mathbf{A}=A$ and $\alpha=\beta$. Then A has the $1-p-$confidence interval

$$\left[\widehat{A}_n - x_p(\alpha)(1-\mathrm{e}^{-h\lambda})^{-1}(1-\mathrm{e}^{-2h\lambda})^{\frac{1}{2}}\widehat{\mathbf{\Omega}}_n^{\frac{1}{2}}(\mathbb{Y}_n'\mathbb{Y}_n)^{-\frac{1}{2}},\right.$$
$$\left.\widehat{A}_n + x_p(\alpha)(1-\mathrm{e}^{-h\lambda})^{-1}(1-\mathrm{e}^{-2h\lambda})^{\frac{1}{2}}\widehat{\mathbf{\Omega}}_n^{\frac{1}{2}}(\mathbb{Y}_n'\mathbb{Y}_n)^{-\frac{1}{2}}\right],$$

which is independent from α since $x_p(\alpha)$ is close to the $1-\frac{p}{2}-$quantile of the normal distribution.

$p=0.1$	1.6
$p=0.05$	1.9
$p=0.025$	2.2
$p=0.01$	2.5

Table 2: Simulated $1-\frac{p}{2}-$quantiles of the components of $\mathbf{G}^{**}$ for any $d,q\in\mathbb{N}$, $\alpha\in\{0.5,0.7,0.9,1.2,1.4,1.5,1.6,1.8,2\}$.

Remark 4.4 By $(X_k)_{k\in\mathbb{N}}$ we denote a sequence of iid $S_\alpha(1,0,0)$-distributed random variables and $L^*=L_1^*$. Then as $n\to\infty$,

$$\left(\frac{\sum_{k=1}^{\lfloor nt\rfloor}X_k}{\left(\sum_{k=1}^n X_k^2\right)^{\frac{1}{2}}}\right)_{t\in[0,1]} \Longrightarrow \left(\frac{L^*(t)}{[L^*,L^*]_1^{\frac{1}{2}}}\right)_{t\in[0,1]} \quad \text{in } (\mathbb{D}[0,1],\mathbb{R}).$$

De la Peña et al. [49], Theorem 4.4, states that the asymptotic behavior of the density of $L^*(1)[L^*,L^*]_1^{-\frac{1}{2}}$ is $c_1(\alpha)\exp(-x^2c_2(\alpha))$ for some constants $c_1(\alpha),c_2(\alpha)$ depending on α. These constants can unfortunately only be calculated numerically. However, Loretan and Phillips [37] simulated the quantiles of $L^*(1)[L^*,L^*]_1^{-\frac{1}{2}}$ for $1<\alpha<2$. They realized that for $p\leq 0.1$, $x_p(\alpha)$ is indeed smaller than the corresponding quantile $x_p(2)$ of the normal case. A similar phenomena was observed for the Student's t-statistic of the symmetric α-stable sequence $X_1,X_2,\ldots$ with $1<\alpha<2$, whose distribution

belongs to the larger class Gaussian scale mixtures; see Bakirov and Székely [4]. There

$$\lim_{n\to\infty} \mathbb{P}\left(\sqrt{n}\frac{\overline{X}_n}{S_{X,n}} > x\right) \leq 1-\Phi(x) \quad \text{for } x \geq \sqrt{3},$$

where $\overline{X}_n = n^{-1}\sum_{k=1}^n X_k$, $S_{X,n} = (n-1)^{-1}\sum_{k=1}^n (X_k - \overline{X}_n)^2$ and Φ is the standard normal distribution function. An open question is if these results are correlated to the properties of $\mathbf{G}^{**}$. A method to get asymptotically correct confidence intervals for our model could be subsampling as presented in Chapter 11 of Politis et al. [55], and was done in McElroy and Politis [44] for the selfnormalized-sum of heavy tailed moving averages. □

4.2 Wald statistic

Next we use the Wald-statistic to test the significance of subvectors and components of $\mathbf{A}$. It can be applied to test which components of $\mathbf{Y}$ have statistically significant information about future values of $\mathbf{X}$. For example, we divide the process $\mathbf{Y}' = (\mathbf{Y}^{(1)'}, Y^{(2)})$ into subprocesses $\mathbf{Y}^{(1)}$ in $\mathbb{R}^{q-1}$ and $Y^{(2)}$ in $\mathbb{R}$, and test if $Y^{(2)}$ not *Granger causes* $\mathbf{X}$. This means that past and present values of $Y^{(2)}$ can not be used to forecast $\mathbf{X}$. For more information we refer to the monograph of Lütkepohl [38]. In our model this is equivalent to $\mathbf{A}_{iq} = 0$ for $i = 1, \ldots, d$, if $\mathbf{A} = (\mathbf{A}_{ij})_{i=1,\ldots,d, j=1,\ldots,q}$. Thus, if we define $\mathbf{R} = (\mathbf{0}_{d\times(q-1)d}, \mathbf{I}_{d\times d}) \in \mathbb{R}^{d\times dq}$, then $\mathbf{A}_{iq} = 0$ for $i = 1, \ldots, d$, if and only if $\mathbf{R}\,\mathrm{vec}(\mathbf{A}) = \mathbf{0}_d$. The null hypothesis of non-Granger causality from $Y^{(2)}$ to $\mathbf{X}$ is then

$$H_0: \quad \mathbf{R}\,\mathrm{vec}(\mathbf{A}) = \mathbf{0}_d\,.$$

In general we obtain the following.

Theorem 4.5 *Let the assumptions and notation of Theorem 3.4 hold,* $\mathbb{P}(\det(\mathbf{S}_3(1)) = 0) = 0$ *and suppose that the null hypothesis* H_0*:* $\mathbf{R}\,\mathrm{vec}(\mathbf{A}) = \mathbf{r}$ *is true where* $\mathbf{R} \in \mathbb{R}^{l\times dq}$, $\mathbf{r} \in \mathbb{R}^l$ *and* $\mathrm{rank}(\mathbf{R}) = q$*. Then as* $n \to \infty$*,*

$$\begin{aligned} F_{\widehat{\mathbf{A}}_n} &= (\mathbf{R}\,\mathrm{vec}(\widehat{\mathbf{A}}_n) - \mathbf{r})' \left(\mathbf{R}((\mathbb{Y}_n'\mathbb{Y}_n)^{-1} \otimes \widehat{\mathbf{\Omega}}_n)\mathbf{R}'\right)^{-1} (\mathbf{R}\,\mathrm{vec}(\widehat{\mathbf{A}}_n) - \mathbf{r}) \\ &\Longrightarrow (\mathbf{R}\,\mathrm{vec}(\mathbf{G}))' \left(\mathbf{R}\left(\int_0^1 \mathbf{S}_1(s)\mathbf{S}_1(s)'\mathrm{d}s \otimes \mathbf{S}_3(1)\right)\mathbf{R}'\right)^{-1} (\mathbf{R}\,\mathrm{vec}(\mathbf{G})) =: G_{\mathbf{R},\mathbf{r}}. \end{aligned}$$

Remark 4.6

(i) Note that $F_{\widehat{\mathbf{A}}_n}$ and $G_{\mathbf{R},\mathbf{r}}$ are real-valued random variables. Let $x_p(\alpha, \beta, \mathbf{r}, \mathbf{R})$ denote the $1-p-$quantile of $G_{\mathbf{R},\mathbf{r}}$. Then the null hypothesis $\mathbf{R}\,\mathrm{vec}(\mathbf{A}) = \mathbf{r}$ is not rejected at significance level p if $F_{\widehat{\mathbf{A}}_n} \leq x_p(\alpha, \beta, \mathbf{r}, \mathbf{R})$.

(ii) In a classical linear model with iid standard-normal noise $(\mathbf{Z}(k))_{k\in\mathbb{N}}$ and $\mathbf{R}\,\mathrm{vec}(\mathbf{A}) = \mathbf{r}$, the sequence of random matrices $\mathbf{R}((\mathbb{Y}_n'\mathbb{Y}_n)^{-1} \otimes \widehat{\mathbf{\Omega}}_n)\mathbf{R}'$ are estimators of the covariance matrix of $\mathbf{R}\,\mathrm{vec}(\widehat{\mathbf{A}}_n) - \mathbf{r}$. □

5 Proofs

5.1 Proofs of Section 2

Before we start with the proof of Proposition 2.1 we require some preliminary results. Without loss of generality we assume $h = 1$ in the following. Note that $(\mathbf{Z}(k))_{k\in\mathbb{N}}$ has the AR(1) representation

$$\mathbf{Z}(k) = \mathrm{e}^{-\mathbf{\Lambda}}\mathbf{Z}(k-1) + \boldsymbol{\xi}_k \quad \text{for } k \in \mathbb{N}, \tag{5.1}$$

where $\boldsymbol{\xi}_k = \int_{k-1}^{k} \mathrm{e}^{-\mathbf{\Lambda}(k-s)}\mathrm{d}\mathbf{L}_2(s)$ for $k \in \mathbb{Z}$. Then also the MA representation

$$\mathbf{Z}(k) = \sum_{j=-\infty}^{k} \mathrm{e}^{-\mathbf{\Lambda}(k-j)}\boldsymbol{\xi}_j = \sum_{i=0}^{\infty} \mathrm{e}^{-\mathbf{\Lambda}i}\boldsymbol{\xi}_{k-i} \tag{5.2}$$

holds and $(\mathbf{Z}(k))_{k\in\mathbb{N}}$ is stationary as well. Furthermore, we define for $\gamma > 0, m \in \mathbb{N}$,

$$\mathbf{Z}_{n,\gamma}(k) = \sum_{i=0}^{\infty} \mathrm{e}^{-\mathbf{\Lambda}i}\boldsymbol{\xi}_{k-i}\mathbb{1}_{\{\|\boldsymbol{\xi}_{k-i}\|>\gamma a_n\}}$$

and the truncated sums

$$\mathbf{Z}^{(m)}(k) = \sum_{i=0}^{m} \mathrm{e}^{-\mathbf{\Lambda}i}\boldsymbol{\xi}_{k-i}, \qquad \mathbf{Z}^{(m)}_{n,\gamma}(k) = \sum_{i=0}^{m} \mathrm{e}^{-\mathbf{\Lambda}i}\boldsymbol{\xi}_{k-i}\mathbb{1}_{\{\|\boldsymbol{\xi}_{k-i}\|>\gamma a_n\}}. \tag{5.3}$$

Finally,

$$\mathbb{Z}'_n = (\mathbf{Z}(1), \ldots, \mathbf{Z}(n)) \in \mathbb{R}^{d\times n}. \tag{5.4}$$

First of all we require some results on multivariate regular variation which we need for the explicit representation of the Lévy measure of $((\mathbf{S}^{(1)}(t)', \mathbf{S}^{(2)}(t)')')_{t\geq 0}$ in Proposition 2.1.

Proposition 5.1 *Let the assumptions of Proposition 2.1 hold. Then*

$$\mathbf{J} = \left(\mathbf{L}_1(1), \int_0^1 \mathrm{e}^{-\mathbf{\Lambda}(1-s)}\mathrm{d}\mathbf{L}_2(s)\right) \in \mathcal{R}_{-\alpha}(a_n, \tilde{\mu}) \qquad \textit{and} \qquad \mathbf{Z}(1) \in \mathcal{R}_{-\alpha}\left(a_n, \sum_{j=0}^{\infty} \tilde{\mu}_2 \circ \mathrm{e}^{\mathbf{\Lambda}j}\right).$$

Proof. Let $(\gamma, \Sigma, \nu_{\mathbf{L}})$ be the characteristic triplet of $\mathbf{L}$ and $\mathbb{S}^{w-1} = \{\mathbf{x} \in \mathbb{R}^w : \|\mathbf{x}\| \leq 1\}$ be the unit ball in $\mathbb{R}^w$. We factorize the Lévy measure $\nu_{\mathbf{L}}$ into two Lévy measures

$$\nu_{\mathbf{L},1}(B) = \nu_{\mathbf{L}}(B\backslash\mathbb{S}^{w-1}) \qquad \text{and} \qquad \nu_{\mathbf{L},2}(B) = \nu_{\mathbf{L}}(B \cap \mathbb{S}^{w-1})$$

such that $\nu_{\mathbf{L}} = \nu_{\mathbf{L},1} + \nu_{\mathbf{L},2}$. Then we can decompose $\mathbf{L}$ in two independent Lévy processes

$$\mathbf{L}(t) = \mathbf{L}^{(1)}(t) + \mathbf{L}^{(2)}(t), \quad t \geq 0,$$

where $\mathbf{L}^{(1)} = (\mathbf{L}^{(1)}(t))_{t\geq 0}$ has the characteristic triplet $(\mathbf{0}_d, \mathbf{0}_{d\times d}, \nu_{\mathbf{L},1})$ and $\mathbf{L}^{(2)} = (\mathbf{L}^{(2)}(t))_{t\geq 0}$ has the characteristic triplet $(\gamma, \Sigma, \nu_{\mathbf{L},2})$. Hence, $\mathbf{J}$ can be written as the sum of two independent random vectors

$$\mathbf{J} = \left(\ \Sigma_1\mathbf{L}^{(1)}(1), \textstyle\int_0^1 \mathrm{e}^{-\mathbf{\Lambda}(1-s)}\Sigma_2\mathrm{d}\mathbf{L}^{(1)}(s)\ \right) + \left(\ \Sigma_1\mathbf{L}^{(2)}(1), \textstyle\int_0^1 \mathrm{e}^{-\mathbf{\Lambda}(1-s)}\Sigma_2\mathrm{d}\mathbf{L}^{(2)}(s)\ \right) =: \mathbf{J}_1 + \mathbf{J}_2.$$

First, we will show that $\mathbf{J}_1$ is regularly varying with the limit measure $\tilde{\mu}$ as stated in (2.3), and secondly that all moments of $\|\mathbf{J}_2\|$ exists. Thus, we have by Lemma 3.12 in Jessen and Mikosch [32] that $\mathbf{J}$ is regularly varying with limit measure $\widetilde{\mu}$.

First, the Lévy measure of $\mathbf{J}_2$ has compact support. Thus, Sato [63], Corollary 25.8, gives that all moments of $\|\mathbf{J}_2\|$ exist.

Next, we prove the multivariate regular variation of $\mathbf{J}_1$. For this, let $(\boldsymbol{\zeta}_k)_{k\in\mathbb{N}}$ be a sequence of w-dimensional iid random vectors with common distribution $\frac{\nu_{\mathbf{L},1}(\cdot)}{\nu_{\mathbf{L},1}(\mathbb{R}^w)}$ and let N be a Poisson process independent of $(\boldsymbol{\zeta}_k)_{k\in\mathbb{N}}$, with intensity $\nu_{\mathbf{L},1}(\mathbb{R}^w)$ and jumping times $(\Gamma_k)_{k\in\mathbb{N}}$. Then $\mathbf{L}^{(1)}$ can be written as a compound Poisson process $\mathbf{L}^{(1)}(t) = \sum_{k=1}^{N(t)} \boldsymbol{\zeta}_k$. Hence,

$$\mathbf{J}_1 = \left(\Sigma_1 \mathbf{L}^{(1)}(1), \int_0^1 \mathrm{e}^{-\boldsymbol{\Lambda}(1-s)} \mathrm{d}\mathbf{L}^{(1)}(s) \right) \stackrel{d}{=} \sum_{k=1}^{N(1)} \left(\Sigma_1 \boldsymbol{\zeta}_k, \mathrm{e}^{-\boldsymbol{\Lambda} U_k} \Sigma_2 \boldsymbol{\zeta}_k \right) =: \sum_{k=1}^{N(1)} \widetilde{\boldsymbol{\zeta}}_k, \tag{5.5}$$

where $(U_k)_{k\in\mathbb{N}}$ is a sequence of iid uniform distributed random variables on $(0,1)$ independent of $(\boldsymbol{\zeta}_k)_{k\in\mathbb{N}}$ and N (cf. Resnick [59], Theorem 4.5.2). By a generalization of Breiman's result in Basrak et al. [5], Proposition 5.1, we obtain $\widetilde{\boldsymbol{\zeta}}_k \in \mathcal{R}_{-\alpha}(a_n, \frac{1}{\nu_{\mathbf{L},1}(\mathbb{R}^w)}\tilde{\mu})$. Finally, Theorem 1.30 in Lindskog [36] gives $\mathbf{J}_1 \in \mathcal{R}_{-\alpha}(a_n, \widetilde{\mu})$.

The multivariate regular variation of $\mathbf{Z}(1)$ follows then by (5.2) and Hult and Samorodnitsky [28], Theorem 2.1. □

A very similar point process result as the following was derived in Davis et al. [16].

Proposition 5.2 *Let the assumptions of Proposition 2.1 hold, and let* $(\mathbf{Z}(k))_{k\in\mathbb{N}}$ *and* $(\mathbf{Z}^{(m)}_{n,\gamma}(k))_{k\in\mathbb{N}}$ *be given as in* (5.2) *and* (5.3)*, respectively. Then as* $n \to \infty$*,*

$$\begin{aligned} N^{(m)}_{n,\gamma} &:= \sum_{k=1}^{\infty} \varepsilon_{(k/n, a_n^{-1}\Delta\mathbf{L}_1(k), a_n^{-1}\boldsymbol{\xi}_k, a_n^{-1}\mathbf{Z}^{(m)}_{n,\gamma}(k), a_n^{-1}\mathbf{Z}(k), a_n^{-1}\mathbf{Z}(k+1))} \\ &\Longrightarrow \sum_{k=1}^{\infty}\sum_{i=0}^{\infty} \varepsilon_{(t_k, j_k^{(1)}\mathbb{1}_{\{i=0\}}, j_k^{(2)}\mathbb{1}_{\{i=0\}}, \mathrm{e}^{-\boldsymbol{\Lambda} i}\mathbb{1}_{\{i\le m\}} j_k^{(2)} \mathbb{1}_{\{\|j_k^{(2)}\|>\gamma\}}, \mathrm{e}^{-\boldsymbol{\Lambda} i} j_k^{(2)}, \mathrm{e}^{-\boldsymbol{\Lambda}(i+1)} j_k^{(2)})} =: N^{(m)}_{\gamma} \end{aligned} \tag{5.6}$$

in $M_p([0,\infty)\times(\overline{\mathbb{R}}^q \times \overline{\mathbb{R}}^{d\times 4})\backslash\{\mathbf{0}_q \times \mathbf{0}_{d\times 4}\})$ *for* $m \in \mathbb{N}\cup\{\infty\}$ *where* $\sum_{k=1}^{\infty} \varepsilon_{(t_k, j_k^{(1)}, j_k^{(2)})} \sim \mathrm{PRM}(Leb\times\widetilde{\mu})$.

Proof. The sequence $(\Delta\mathbf{L}_1(k), \boldsymbol{\xi}_k)_{k\in\mathbb{N}}$ is a sequence of iid random vectors which are in $\mathcal{R}_{-\alpha}(a_n, \tilde{\mu})$ by Proposition 5.1. Therefore we have by Resnick [58], Proposition 3.21, that as $n \to \infty$,

$$\sum_{k=1}^{\infty} \varepsilon_{(k/n, a_n^{-1}\Delta\mathbf{L}_1(k), a_n^{-1}\boldsymbol{\xi}_k)} \Longrightarrow \sum_{k=1}^{\infty} \varepsilon_{(t_k, j_k^{(1)}, j_k^{(2)})} \tag{5.7}$$

in $M_p([0,\infty)\times\overline{\mathbb{R}}^{q+d}\backslash\{\mathbf{0}_{q+d}\})$. Now fix some integer $l \in \mathbb{N}$ and define the $d\times l$-dimensional random matrices

$$\boldsymbol{\xi}_{(k,l)} = (\boldsymbol{\xi}_k, \ldots, \boldsymbol{\xi}_{k-l+1}) \quad \text{and} \quad \boldsymbol{\xi}^n_{(k,l)} = (\boldsymbol{\xi}_k \mathbb{1}_{\{\|\boldsymbol{\xi}_k\|>a_n\gamma\}}, \ldots, \boldsymbol{\xi}_{k-l+1}\mathbb{1}_{\{\|\boldsymbol{\xi}_{k-l+1}\|>a_n\gamma\}}).$$

Further, we denote by $\mathbf{e}_i$ the unit vector in $\mathbb{R}^l$ having 1 in row i and 0 otherwise. Then, as in Theorem 2.2 in Davis and Resnick [17], we obtain with (5.7) that as $n \to \infty$,

$$\sum_{k=1}^{\infty} \varepsilon_{(k/n, a_n^{-1}\Delta \mathbf{L}_1(k), a_n^{-1}\boldsymbol{\xi}_{(k,l)})} \Longrightarrow \sum_{k=1}^{\infty}\sum_{i=1}^{l} \varepsilon_{(t_k, j_k^{(1)}\mathbb{1}_{\{i=1\}}, \mathbf{e}_i' \otimes j_k^{(2)})}$$

in $M_p([0,\infty) \times (\overline{\mathbb{R}}^q \times \overline{\mathbb{R}}^{d\times l}) \backslash \{\mathbf{0}_q \times \mathbf{0}_{d\times l}\})$. An application of the continuous mapping theorem and Resnick [58], Proposition 3.18, result that as $n \to \infty$,

$$\sum_{k=1}^{\infty} \varepsilon_{(k/n, a_n^{-1}\Delta \mathbf{L}_1(k), a_n^{-1}\boldsymbol{\xi}_{(k,l)}, a_n^{-1}\boldsymbol{\xi}^n_{(k,l)})} \Longrightarrow \sum_{k=1}^{\infty}\sum_{i=1}^{l} \varepsilon_{(t_k, j_k^{(1)}\mathbb{1}_{\{i=1\}}, \mathbf{e}_i' \otimes j_k^{(2)}, \mathbf{e}_i' \otimes j_k^{(2)}\mathbb{1}_{\{\|j_k^{(2)}\|>\gamma\}})}$$

in $M_p([0,\infty) \times (\overline{\mathbb{R}}^q \times \overline{\mathbb{R}}^{d\times l} \times \overline{\mathbb{R}}^{d\times l}) \backslash \{\mathbf{0}_q \times \mathbf{0}_{d\times l} \times \mathbf{0}_{d\times l}\})$. Using the continuous mapping theorem a second time gives

$$\begin{aligned} N_{n,\gamma}^{(m,l)} = & \sum_{k=1}^{\infty} \varepsilon_{(k/n, a_n^{-1}\Delta \mathbf{L}_1(k), a_n^{-1}\boldsymbol{\xi}_k, a_n^{-1}\sum_{i=0}^{l-1} \mathrm{e}^{-\boldsymbol{\Lambda} i}\mathbb{1}_{\{i\le m\}}\boldsymbol{\xi}_{k-i}\mathbb{1}_{\{\|\boldsymbol{\xi}_{k-i}\|>a_n\gamma\}}, a_n^{-1}\sum_{i=0}^{l-1}\mathrm{e}^{-\boldsymbol{\Lambda} i}\boldsymbol{\xi}_{k-i}, a_n^{-1}\sum_{i=0}^{l-1}\mathrm{e}^{-\boldsymbol{\Lambda}(i+1)}\boldsymbol{\xi}_{k-i})} \\ \Longrightarrow & \sum_{k=1}^{\infty}\sum_{i=0}^{l-1} \varepsilon_{(t_k, j_k^{(1)}\mathbb{1}_{\{i=0\}}, j_k^{(2)}\mathbb{1}_{\{i=0\}}, \mathrm{e}^{-\boldsymbol{\Lambda} i}\mathbb{1}_{\{i\le m\}}j_k^{(2)}\mathbb{1}_{\{\|j_k^{(2)}\|>\gamma\}}, \mathrm{e}^{-\boldsymbol{\Lambda} i}j_k^{(2)}, \mathrm{e}^{-\boldsymbol{\Lambda}(i+1)}j_k^{(2)})} =: N_{\gamma}^{(m,l)}. \end{aligned}$$

Furthermore, $N_{\gamma}^{(m,l)} \Longrightarrow N_{\gamma}^{(m)}$ as $l \to \infty$. Thus, if

$$\lim_{l\to\infty} \limsup_{n\to\infty} \mathbb{P}(\rho(N_{n,\gamma}^{(m,l)}, N_{n,\gamma}^{(m)}) > \eta) = 0, \tag{5.8}$$

where ρ is the metric inducing the vague topology on $M_p([0,\infty) \times \overline{\mathbb{R}}^q \times \overline{\mathbb{R}}^{d\times 4} \backslash \{\mathbf{0}_q \times \mathbf{0}_{d\times 4}\})$, we can finish the proof by Billingsley [8], Theorem 4.2. However, (5.8) holds as (2.11) in Davis and Resnick [17] using Hult and Samorodnitsky [28], Theorem 2.1 (or Lemma 2.3 in Davis et al. [16]). □

Proof of Proposition 2.1.

Let us define for $m \in \mathbb{N}$ and for $0 < \gamma < 1$,

$$\begin{aligned} \mathbf{S}_{n,\gamma}^{(1)}(t) &= \frac{1}{a_n}\sum_{k=1}^{\lfloor nt\rfloor}\left(\Delta\mathbf{L}_1(k)\mathbb{1}_{\{\|\Delta\mathbf{L}_1(k)\|>\gamma a_n\}} - \mathbb{E}(\Delta\mathbf{L}_1(1)\mathbb{1}_{\{\gamma a_n<\|\Delta\mathbf{L}_1(1)\|\le a_n\}})\right), \\ \mathbf{S}_{n}^{(2,m)}(t) &= \frac{1}{a_n}\sum_{k=1}^{\lfloor nt\rfloor}\left(\mathbf{Z}^{(m)}(k) - \sum_{i=0}^{m}\mathrm{e}^{-\boldsymbol{\Lambda} i}\mathbb{E}(\boldsymbol{\xi}_1\mathbb{1}_{\{\|\boldsymbol{\xi}_1\|\le a_n\}})\right), \\ \mathbf{S}_{n,\gamma}^{(2,m)}(t) &= \frac{1}{a_n}\sum_{k=1}^{\lfloor nt\rfloor}\left(\mathbf{Z}_{n,\gamma}^{(m)}(k) - \sum_{i=0}^{m}\mathrm{e}^{-\boldsymbol{\Lambda} i}\mathbb{E}(\boldsymbol{\xi}_1\mathbb{1}_{\{\gamma a_n<\|\boldsymbol{\xi}_1\|\le a_n\}})\right), \\ \mathbf{S}_{n,\gamma}^{(3)}(t) &= \frac{1}{a_n^2}\sum_{k=1}^{\lfloor nt\rfloor}\mathbf{Z}(k)\mathbf{Z}(k)'\mathbb{1}_{\{\|\mathbf{Z}(k)\|>\gamma a_n\}}, \\ \mathbf{S}_{n,\gamma}^{(4)}(t) &= \frac{1}{a_n^2}\sum_{k=1}^{\lfloor nt\rfloor}\mathbf{Z}(k+1)\mathbf{Z}(k)'\mathbb{1}_{\{\|\mathbf{Z}(k)\|>\gamma a_n \text{ or } \|\mathbf{Z}(k+1)\|>\gamma a_n\}}, \end{aligned}$$

and

$$\mathbf{S}_{n,\gamma}^m = ((\mathbf{S}_{n,\gamma}^{(1)}(t))_{t\geq 0}, \mathbf{S}_{n,\gamma}^{(2,m)}(\mathbf{t}), \mathbf{S}_{n,\gamma}^{(3)}(\mathbf{t}), \mathbf{S}_{n,\gamma}^{(4)}(\mathbf{t}))$$

in $((\mathbb{D}[0,1], \mathbb{R}^q) \times \mathbb{R}^{d\times l} \times \mathbb{R}^{(d\times d)\times l} \times \mathbb{R}^{(d\times d)\times l})$. There exist a map

$$\Phi : M_p([0,\infty) \times (\overline{\mathbb{R}}^q \times \overline{\mathbb{R}}^{d\times 4}) \backslash \{\mathbf{0}_q \times \mathbf{0}_{d\times 4}\}) \to ((\mathbb{D}[0,1], \mathbb{R}^q) \times \mathbb{R}^{d\times l} \times \mathbb{R}^{(d\times d)\times l} \times \mathbb{R}^{(d\times d)\times l})$$

with

$$\Phi(N_{n,\gamma}^{(m)}) = \mathbf{S}_{n,\gamma}^m. \tag{5.9}$$

Since Φ is a.s. continuous with respect to $N_\gamma^{(m)}$ (cf. Resnick [58], Proposition 3.13 and Resnick [60], Section 7.2.3), we have by Proposition 5.2 and the continuous mapping theorem as $n \to \infty$,

$$\Phi(N_{n,\gamma}^{(m)}) \Longrightarrow \Phi(N_\gamma^{(m)}).$$

Let us define $\mathbf{S}_\gamma^m := \Phi(N_\gamma^{(m)})$, i.e. $\mathbf{S}_{n,\gamma}^m \Longrightarrow \mathbf{S}_\gamma^m$ as $n \to \infty$. Furthermore, if $\gamma \to 0$ then $\mathbf{S}_\gamma^m$ converges weakly. The limit we denote by $\mathbf{S}^m$ and it is equal to $((\mathbf{S}^{(1)}(t))_{t\geq 0}, \mathbf{S}^{(2,m)}(\mathbf{t}), \mathbf{S}^{(3)}(\mathbf{t}), \mathbf{S}^{(4)}(\mathbf{t}))$ where for $t \geq 0$,

$$\mathbf{S}^{(2,m)}(t) = \sum_{i=0}^{m} \mathrm{e}^{-\mathbf{\Lambda} i} \left(\sum_{t_k \leq t} j_k^{(2)} \mathbb{1}_{\{\|j_k^{(2)}\|>1\}} + \lim_{\gamma\to 0} \left(\sum_{t_k \leq t} j_k^{(2)} \mathbb{1}_{\{\gamma<\|j_k^{(2)}\|\leq 1\}} - t \int_{\gamma<\|\mathbf{x}\|\leq 1} \mathbf{x}\, \widetilde{\mu}_2(\mathrm{d}\mathbf{x}) \right) \right).$$

We will now divide the proof in several steps. Therefore we will show that

$$\lim_{\gamma\to 0} \lim_{n\to\infty} \mathbb{P}(\sup_{t\in[0,1]} \|\mathbf{S}_{n,\gamma}^{(1)}(t) - \mathbf{S}_n^{(1)}(t)\| > \eta) = 0, \tag{5.10}$$

$$\lim_{\gamma\to 0} \lim_{n\to\infty} \mathbb{P}(\sup_{t\in[0,1]} \|\mathbf{S}_{n,\gamma}^{(2,m)}(t) - \mathbf{S}_n^{(2,m)}(t)\| > \eta) = 0, \tag{5.11}$$

$$\lim_{\gamma\to 0} \lim_{n\to\infty} \mathbb{P}(\sup_{t\in[0,1]} \|\mathbf{S}_{n,\gamma}^{(3)}(t) - \mathbf{S}_n^{(3)}(t)\| > \eta) = 0, \tag{5.12}$$

$$\lim_{\gamma\to 0} \lim_{n\to\infty} \mathbb{P}(\sup_{t\in[0,1]} \|\mathbf{S}_{n,\gamma}^{(4)}(t) - \mathbf{S}_n^{(4)}(t)\| > \eta) = 0 \tag{5.13}$$

for any $\eta > 0$. Because Billingsley [8], Theorem 4.2, then gives $\mathbf{S}_n^m \Longrightarrow \mathbf{S}^m$ as $n \to \infty$, where $\mathbf{S}^m \Longrightarrow ((\mathbf{S}^{(1)}(t))_{t\geq 0}, \mathbf{S}^{(2)}(\mathbf{t}), \mathbf{S}^{(3)}(\mathbf{t}), \mathbf{S}^{(4)}(\mathbf{t}))$ as $m \to \infty$. Since we want to show that

$$\mathbf{S}_n := ((\mathbf{S}_n^{(1)}(t))_{t\geq 0}, \mathbf{S}_n^{(2)}(\mathbf{t}), \mathbf{S}_n^{(3)}(\mathbf{t}), \mathbf{S}_n^{(4)}(\mathbf{t})) \Longrightarrow ((\mathbf{S}^{(1)}(t))_{t\geq 0}, \mathbf{S}^{(2)}(\mathbf{t}), \mathbf{S}^{(3)}(\mathbf{t}), \mathbf{S}^{(4)}(\mathbf{t})) \quad \text{as } n \to \infty,$$

it is then sufficient to prove that

$$\lim_{m\to\infty} \lim_{n\to\infty} \mathbb{P}(\sup_{t\in[0,1]} \|\mathbf{S}_n^m(t) - \mathbf{S}_n(t)\| > \eta) = 0 \tag{5.14}$$

for any $\eta > 0$. Then $((\mathbf{S}_n^{(1)}(t))_{t\geq 0}, \mathbf{S}_n^{(2)}(\mathbf{t}), \mathbf{S}_n^{(3)}(\mathbf{t}), \mathbf{S}_n^{(4)}(\mathbf{t})) \Longrightarrow ((\mathbf{S}^{(1)}(t))_{t\geq 0}, \mathbf{S}^{(2)}(\mathbf{t}), \mathbf{S}^{(3)}(\mathbf{t}), \mathbf{S}^{(4)}(\mathbf{t}))$ as $n \to \infty$ follows by Billingsley [8], Theorem 4.2, as well, and we can conclude the proof.

We start with (5.10). First,

$$\begin{aligned}
&\mathbb{P}(\sup_{t\in[0,1]}\|\mathbf{S}^{(1)}_{n,\gamma}(t)-\mathbf{S}^{(1)}_n(t)\|>\eta) \qquad (5.15)\\
&\quad=\mathbb{P}\left(\sup_{t\in[0,1]}\left\|\frac{1}{a_n}\sum_{k=1}^{\lfloor nt\rfloor}\left(\Delta\mathbf{L}_1(k)\mathbb{1}_{\{\|\Delta\mathbf{L}_1(k)\|\le\gamma a_n\}}-\mathbb{E}(\Delta\mathbf{L}_1(1)\mathbb{1}_{\{\|\Delta\mathbf{L}_1(1)\|\le\gamma a_n\}})\right)\right\|>\eta\right)\\
&\quad\le\sum_{j=1}^{q}\mathbb{P}\left(\sup_{0\le i\le n}\left|\frac{1}{a_n}\sum_{k=1}^{i}\left(\Delta L_{1,j}(k)\mathbb{1}_{\{\|\Delta\mathbf{L}_1(k)\|\le\gamma a_n\}}-\mathbb{E}(\Delta L_{1,j}(1)\mathbb{1}_{\{\|\Delta\mathbf{L}_1(1)\|\le\gamma a_n\}})\right)\right|>C_1\eta\right)
\end{aligned}$$

where $\Delta L_{1,j}(k)$ is the j-th component of $\Delta\mathbf{L}_1(k)=(\Delta L_{1,1}(k),\dots,\Delta L_{1,q}(k))$. Now, note that the terms in the sum are an iid sequence with mean zero and finite variance. Hence, we can apply Kolmogorov's inequality (cf. Kallenberg [34], Lemma 4.15) and obtain

$$\mathbb{P}(\sup_{t\in[0,1]}\|\mathbf{S}^{(1)}_{n,\gamma}(t)-\mathbf{S}^{(1)}_n(t)\|>\eta)\le\frac{1}{C_1^2\eta^2a_n^2}\sum_{j=1}^{q}n\mathbb{E}\left((\Delta L_{1,j}(1))^2\mathbb{1}_{\{\|\Delta\mathbf{L}_1(1)\|\le\gamma a_n\}}\right).$$

Since $\|\Delta\mathbf{L}_1(1)\|\in\mathcal{R}_{-\alpha}(a_n)$ or lighter tailed, Karamata's and Potter's Theorem (cf. Resnick [60], 2.5 on p. 36 and Proposition 2.6) result for some $\epsilon>0$ small in

$$\frac{n}{a_n^2}\mathbb{E}\left((\Delta L_{1,j}(1))^2\mathbb{1}_{\{\|\Delta\mathbf{L}_1(1)\|\le\gamma a_n\}}\right)\le C_2\frac{n}{a_n^2}\mathbb{E}\left(\|\Delta\mathbf{L}_1(1)\|^2\mathbb{1}_{\{\|\Delta\mathbf{L}_1(1)\|\le\gamma a_n\}}\right)\le C_3\gamma^{2-\alpha-\epsilon}\longrightarrow 0$$

as $\gamma\to 0$. Hence, we have (5.10). Next we prove (5.12). Here, applying Markov's inequality gives

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\in[0,1]}\|\mathbf{S}^{(3)}_{n,\gamma}(t)-\mathbf{S}^{(3)}_n(t)\|>\eta\right)&\le\mathbb{P}\left(\sup_{t\in[0,1]}C_4\sum_{k=1}^{\lfloor nt\rfloor}\|\mathbf{Z}(k)\|^2\mathbb{1}_{\{\|\mathbf{Z}(k)\|\le\gamma a_n\}}>\eta a_n^2\right)\\
&\le\frac{C_4}{\eta a_n^2}\sum_{k=1}^{n}\mathbb{E}\left(\|\mathbf{Z}(k)\|^2\mathbb{1}_{\{\|\mathbf{Z}(k)\|\le\gamma a_n\}}\right).
\end{aligned}$$

Again by Karamata's and Potter's Theorem, and $\|\mathbf{Z}(0)\|\le\sum_{k=0}^{\infty}\mathrm{e}^{-\lambda k}\|\boldsymbol{\xi}_k\|\in\mathcal{R}_{-\alpha}(a_n)$ or lighter tailed we can conclude that

$$\frac{n}{a_n^2}\mathbb{E}\left(\|\mathbf{Z}(0)\|^2\mathbb{1}_{\{\|\mathbf{Z}(0)\|\le\gamma a_n\}}\right)\le C_5\gamma^{2-\alpha-\epsilon}\longrightarrow 0\quad\text{ as }\gamma\to 0,$$

such that (5.12) holds. Analogously we derive (5.13).

Finally, we turn our attention to (5.11). We use the following decomposition:

$$\begin{aligned}
\mathbf{S}^{(2,m)}_n(t)-\mathbf{S}^{(2,m)}_{n,\gamma}(t)&=\frac{1}{a_n}\sum_{i=1-m}^{0}\left(\sum_{k=1}^{i+m}\mathrm{e}^{-\mathbf{\Lambda}(k-i)}\right)\left(\boldsymbol{\xi}_i\mathbb{1}_{\{\|\boldsymbol{\xi}_i\|\le\gamma a_n\}}-\mathbb{E}(\boldsymbol{\xi}_1\mathbb{1}_{\{\|\boldsymbol{\xi}_1\|\le\gamma a_n\}})\right)\\
&\quad+\frac{1}{a_n}\sum_{i=1}^{\lfloor nt\rfloor-m}\left(\sum_{k=i}^{i+m}\mathrm{e}^{-\mathbf{\Lambda}(k-i)}\right)\left(\boldsymbol{\xi}_i\mathbb{1}_{\{\|\boldsymbol{\xi}_i\|\le\gamma a_n\}}-\mathbb{E}(\boldsymbol{\xi}_1\mathbb{1}_{\{\|\boldsymbol{\xi}_1\|\le\gamma a_n\}})\right)\\
&\quad+\frac{1}{a_n}\sum_{i=\lfloor nt\rfloor-m+1}^{\lfloor nt\rfloor}\left(\sum_{k=i}^{\lfloor nt\rfloor}\mathrm{e}^{-\mathbf{\Lambda}(k-i)}\right)\left(\boldsymbol{\xi}_i\mathbb{1}_{\{\|\boldsymbol{\xi}_i\|\le\gamma a_n\}}-\mathbb{E}(\boldsymbol{\xi}_1\mathbb{1}_{\{\|\boldsymbol{\xi}_1\|\le\gamma a_n\}})\right)\\
&=:\ I^{(1)}_{n,\gamma}(t)+I^{(2)}_{n,\gamma}(t)+I^{(3)}_{n,\gamma}(t).
\end{aligned}$$

Note that $I^{(1)}_{n,\gamma}(t)$ is independent of t. Furthermore, $\|\sum_{k=1}^{i+m} \mathrm{e}^{-\mathbf{\Lambda}(k-i)}\| \le C_6 < \infty$ for $i = 1-m, \dots, 0$. Thus, we get for any $t \ge 0$,

$$\|I^{(1)}_{n,\gamma}(t)\| \le \frac{C_6}{a_n} \sum_{i=1-m}^{0} \|\boldsymbol{\xi}_i \mathbb{1}_{\{\|\boldsymbol{\xi}_i\| \le \gamma a_n\}} - \mathbb{E}(\boldsymbol{\xi}_1 \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \le \gamma a_n\}})\| \le C_7 \gamma m \longrightarrow 0 \quad \text{as } \gamma \to 0.$$

Next, we investigate $I^{(3)}_{n,\gamma}(t)$. Similarly as above we obtain

$$\left\| \sum_{k=i}^{\lfloor nt \rfloor} \mathrm{e}^{-\mathbf{\Lambda}(k-i)} \right\| \le C_8 \sum_{k=1}^{\infty} \mathrm{e}^{-\lambda k} \le C_9 < \infty.$$

This yields for any $t \ge 0$,

$$\|I^{(3)}_{n,\gamma}(t)\| \le \frac{C_9}{a_n} \sum_{i=\lfloor nt \rfloor - m+1}^{\lfloor nt \rfloor} \left\|\boldsymbol{\xi}_i \mathbb{1}_{\{\|\boldsymbol{\xi}_i\| \le \gamma a_n\}} - \mathbb{E}(\boldsymbol{\xi}_1 \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \le \gamma a_n\}})\right\| \le C_{10} m \gamma \longrightarrow 0 \quad \text{as } \gamma \to 0.$$

Finally, we treat the second term $I^{(2)}_{n,\gamma}$. As in the two cases before, we start with an upper bound $\|\sum_{k=0}^{m} \mathrm{e}^{-\mathbf{\Lambda} k}\| \le C_{11}$. We use Kolmogorov's inequality componentwise and afterward Karamata's and Potter's Theorem. Let $\xi_{i,j}$ be the j-th component of $\boldsymbol{\xi}_i = (\xi_{i,1}, \dots, \xi_{i,d})$. This gives

$$\begin{aligned}
\mathbb{P}(\sup_{t\in[0,1]} \|I^{(2)}_{n,\gamma}(t)\| > \eta) &\le \sum_{j=1}^{d} \mathbb{P}\left(\sup_{0\le k\le n} \left| \sum_{i=1}^{k-m} \left(\xi_{i,j} \mathbb{1}_{\{\|\boldsymbol{\xi}_i\| \le \gamma a_n\}} - \mathbb{E}(\xi_{1,j} \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \le \gamma a_n\}})\right) \right| > a_n \eta C_{12} \right) \\
&\le \frac{C_{13}}{a_n^2 \eta^2} n \mathbb{E}\left(\|\boldsymbol{\xi}_1\|^2 \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \le \gamma a_n\}} \right) \\
&\le \frac{C_{14}}{\eta^2} \gamma^{2-\alpha-\epsilon} \longrightarrow 0 \quad \text{as } \gamma \to 0.
\end{aligned}$$

Thus, we have shown not only (5.11) but also a stronger version

$$\lim_{\gamma \to 0} \sup_{n\in\mathbb{N}} \mathbb{P}(\sup_{t\in[0,1]} \|\mathbf{S}^{(2,m)}_{n,\gamma}(t) - \mathbf{S}^{(2,m)}_{n}(t)\| > \eta) = 0 \tag{5.16}$$

and hence, $\mathbf{S}^m_n \Longrightarrow \mathbf{S}^m$ as $n \to \infty$ holds.

Now we consider (5.14). Note that $\mathbf{S}_n$ and $\mathbf{S}^m_n$ differ only in the second component such that (5.14) is equivalent to

$$\lim_{m\to\infty} \lim_{n\to\infty} \mathbb{P}(\sup_{t\in[0,1]} \|\mathbf{S}^{(2)}_{n}(t) - \mathbf{S}^{(2,m)}_{n}(t)\| > \eta) = 0.$$

We use the following decomposition

$$\begin{aligned}
\mathbf{S}^{(2)}_n(t) - \mathbf{S}^{(2,m)}_n(t) &= \frac{1}{a_n} \sum_{i=-\infty}^{-m-1} \left(\sum_{k=1}^{\lfloor nt \rfloor} \mathrm{e}^{-\mathbf{\Lambda}(k-i)} \right) \left(\boldsymbol{\xi}_i \mathbb{1}_{\{\|\boldsymbol{\xi}_i\| \le a_n\}} - \mathbb{E}(\boldsymbol{\xi}_1 \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \le a_n\}}) \right) \\
&\quad + \frac{1}{a_n} \sum_{i=-m}^{\lfloor nt \rfloor - m - 1} \left(\sum_{k=i+m+1}^{\lfloor nt \rfloor} \mathrm{e}^{-\mathbf{\Lambda}(k-i)} \right) \left(\boldsymbol{\xi}_i \mathbb{1}_{\{\|\boldsymbol{\xi}_i\| \le a_n\}} - \mathbb{E}(\boldsymbol{\xi}_1 \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \le a_n\}}) \right) \\
&\quad + \frac{1}{a_n} \sum_{k=1}^{\lfloor nt \rfloor} \sum_{i=-\infty}^{k-m-1} \mathrm{e}^{-\mathbf{\Lambda}(k-i)} \boldsymbol{\xi}_i \mathbb{1}_{\{\|\boldsymbol{\xi}_i\| > a_n\}} \\
&=: J^{(1)}_n(t) + J^{(2)}_n(t) + J^{(3)}_n(t).
\end{aligned} \tag{5.17}$$

We will investigate all three terms. We start with $J_n^{(1)}$ by looking at it componentwise. Furthermore, we have $\sum_{j=1}^d \|\mathbf{e}_j' \mathrm{e}^{\mathbf{\Lambda} i}\|^2 \leq C_{15}\mathrm{e}^{2\lambda i}$ and $\|\sum_{k=1}^{\lfloor nt \rfloor} \mathrm{e}^{-\mathbf{\Lambda} k}\| \leq C_{16}$ for any $t \geq 0$. An application of Tschebyscheff's inequality results in

$$\begin{aligned}
\mathbb{P}(\sup_{t\in[0,1]} \|J_n^{(1)}(t)\| > \eta) &\leq \sum_{j=1}^d \mathbb{P}\left(\left|\sum_{i=-\infty}^{-m-1} \left(\mathbf{e}_j' \mathrm{e}^{\mathbf{\Lambda} i} \boldsymbol{\xi}_i \mathbb{1}_{\{\|\boldsymbol{\xi}_i\| \leq a_n\}} - \mathbb{E}(\mathbf{e}_j' \mathrm{e}^{\mathbf{\Lambda} i} \boldsymbol{\xi}_1 \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \leq a_n\}})\right)\right| > \eta a_n C_{17}\right) \\
&\leq \frac{C_{18}}{a_n^2 \eta^2} \sum_{i=-\infty}^{-m-1} \sum_{j=1}^d \|\mathbf{e}_j' \mathrm{e}^{\mathbf{\Lambda} i}\|^2 \mathbb{E}\left(\|\boldsymbol{\xi}_1\|^2 \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \leq a_n\}}\right) \leq \frac{C_{19}}{\eta^2} \sum_{i=m+1}^{\infty} \mathrm{e}^{-2\lambda i} \\
&\longrightarrow 0 \quad \text{as } m \to \infty. \qquad (5.18)
\end{aligned}$$

Next, we show that $\sup_{t\in[0,1]} \|J_n^{(2)}(t)\| \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$. Therefore we use the decomposition

$$\begin{aligned}
J_n^{(2)}(t) &= \frac{1}{a_n} \sum_{i=-m}^{\lfloor nt \rfloor - m - 1} \left(\sum_{k=m+1}^{\infty} \mathrm{e}^{-\mathbf{\Lambda} k}\right) \left(\boldsymbol{\xi}_i \mathbb{1}_{\{\|\boldsymbol{\xi}_i\| \leq a_n\}} - \mathbb{E}(\boldsymbol{\xi}_1 \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \leq a_n\}})\right) \\
&\quad - \frac{1}{a_n} \sum_{i=-m}^{\lfloor nt \rfloor - m - 1} \left(\sum_{k=\lfloor nt \rfloor - i + 1}^{\infty} \mathrm{e}^{-\mathbf{\Lambda} k}\right) \left(\boldsymbol{\xi}_i \mathbb{1}_{\{\|\boldsymbol{\xi}_i\| \leq a_n\}} - \mathbb{E}(\boldsymbol{\xi}_1 \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \leq a_n\}})\right) \\
&=: J_n^{(2,1)}(t) + J_n^{(2,2)}(t). \qquad (5.19)
\end{aligned}$$

First, we investigate $J_n^{(2,1)}$. Therefore we take $\left\|\sum_{k=m+1}^{\infty} \mathrm{e}^{-\mathbf{\Lambda} k}\right\| \leq C_{20}\mathrm{e}^{-\lambda m}$ into account. Similar calculations as above yield

$$\begin{aligned}
&\mathbb{P}(\sup_{t\in[0,1]} \|J_n^{(2,1)}(t)\| > \eta) \\
&\leq \sum_{j=1}^d \mathbb{P}\left(\sup_{t\in[0,1]} \left|\sum_{i=-m}^{\lfloor nt \rfloor - m - 1} \left(\xi_{i,j} \mathbb{1}_{\{\|\boldsymbol{\xi}_i\| \leq a_n\}} - \mathbb{E}(\xi_{1,j} \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \leq a_n\}})\right)\right| > C_{21}\mathrm{e}^{\lambda m} a_n \eta\right) \\
&\leq C_{22} \frac{\mathrm{e}^{-2\lambda m}}{\eta^2 a_n^2} n \mathbb{E}(\|\boldsymbol{\xi}_1\|^2 \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \leq a_n\}}) \\
&\leq C_{23}\mathrm{e}^{-2\lambda m} \longrightarrow 0 \quad \text{as } m \to \infty. \qquad (5.20)
\end{aligned}$$

Moreover, $\|\sum_{k=l-i+1}^{\infty} \mathrm{e}^{-\mathbf{\Lambda} k}\| \leq C_{24}\mathrm{e}^{-\lambda(l-i)}$ and $\sup_{1\leq l\leq n} \sum_{i=-m}^{l-m-1} \mathrm{e}^{-\lambda(l-i)} \leq C_{25}\mathrm{e}^{-\lambda m}$ give for the second term:

$$\begin{aligned}
&\mathbb{P}(\sup_{t\in[0,1]} \|J_n^{(2,2)}(t)\| > \eta) \\
&\leq \mathbb{P}\left(\sup_{1\leq l\leq n} \sum_{i=-m}^{l-m-1} \mathrm{e}^{-\lambda(l-i)} \left\|\boldsymbol{\xi}_i \mathbb{1}_{\{\|\boldsymbol{\xi}_i\| \leq a_n\}} - \mathbb{E}(\boldsymbol{\xi}_1 \mathbb{1}_{\{\|\boldsymbol{\xi}_1\| \leq a_n\}})\right\| > a_n \eta C_{26}\right) \\
&\leq \mathbb{P}\left(\mathrm{e}^{-\lambda m} a_n > a_n \eta C_{27}\right) \longrightarrow 0 \quad \text{as } m \to \infty. \qquad (5.21)
\end{aligned}$$

Thus, we have established the convergence of $\sup_{t\in[0,1]} \|J_n^{(2)}(t)\| \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$. Finally, we investigate $J_n^{(3)}$ where the upper bound

$$\sup_{t\in[0,1]} \|J_n^{(3)}(t)\| \leq \frac{C_{28}}{a_n} \sum_{k=1}^{n} \sum_{i=-\infty}^{k-m+1} \mathrm{e}^{-\lambda(k-i)} \|\boldsymbol{\xi}_i\| \mathbb{1}_{\{\|\boldsymbol{\xi}_i\| > a_n\}}$$

holds. In order to be able to apply Karamata's Theorem we have to treat two cases. First, let $\alpha > 1$. Markov's inequality gives

$$\mathbb{P}(\sup_{t\in[0,1]} \|J_n^{(3)}(t)\| > \eta) \leq \frac{C_{29}}{a_n \eta} \sum_{k=1}^{n} \sum_{i=-\infty}^{k-m+1} \mathrm{e}^{-\lambda(k-i)} \mathbb{E}\left(\|\boldsymbol{\xi}_1\| \mathbb{1}_{\{\|\boldsymbol{\xi}_1\|>a_n\}}\right) \leq C_{30} \left(\sum_{i=m+1}^{\infty} \mathrm{e}^{-\lambda i} \right) \longrightarrow 0$$

as $m \to \infty$. Now, let $\alpha \leq 1$ and $0 < \delta < \alpha \leq 1$. Then

$$\left(\frac{1}{a_n} \sum_{k=1}^{n} \sum_{i=-\infty}^{k-m-1} \mathrm{e}^{-\lambda(k-i)} \|\boldsymbol{\xi}_i\| \mathbb{1}_{\{\|\boldsymbol{\xi}_i\|>a_n\}} \right)^{\delta} \leq a_n^{-\delta} \sum_{k=1}^{n} \sum_{i=-\infty}^{k-m-1} \mathrm{e}^{-\lambda\delta(k-i)} \|\boldsymbol{\xi}_i\|^{\delta} \mathbb{1}_{\{\|\boldsymbol{\xi}_i\|>a_n\}},$$

and by Markov's inequality and Karamata's Theorem we obtain

$$\begin{aligned} \mathbb{P}(\sup_{t\in[0,1]} \|J_n^{(3)}(t)\| > \eta) &\leq \eta^{-\delta} a_n^{-\delta} \sum_{k=1}^{n} \sum_{i=-\infty}^{k-m-1} \mathrm{e}^{-\lambda\delta(k-i)} \mathbb{E}(\|\boldsymbol{\xi}_i\|^{\delta} \mathbb{1}_{\{\|\boldsymbol{\xi}_i\|>a_n\}}) \\ &\leq C_{31} \left(\sum_{i=m+1}^{\infty} \mathrm{e}^{-\lambda\delta i} \right) \longrightarrow 0 \end{aligned}$$

as $m \to \infty$ which proves statement (5.14) together with (5.17)-(5.36). Again we proved a stronger version namely

$$\lim_{m\to\infty} \sup_{n\in\mathbb{N}} \mathbb{P}(\sup_{t\in[0,1]} \|\mathbf{S}_n^{(2,m)}(t) - \mathbf{S}_n^{(2)}(t)\| > \eta) = 0 \tag{5.22}$$

for any $\eta > 0$. □

Remark 5.3 (Continuation of Remark 2.2) As noted in Remark 2.2 the process $(\mathbf{S}_n^{(2)})_{n\in\mathbb{N}}$ does not converge in the Skorokhod J_1 topology. The only part where the proof in Proposition 2.1 fails is that the adapted definition of Φ in (5.9) would not be a.s. continuous in the J_1 topology with respect to $N_\gamma^{(m)}$ anymore and thus, we are not allowed to apply the continuous mapping theorem. □

Proof of Proposition 2.4.

(a) is a multivariate version of the invariance principle of Donsker (cf. Phillips and Durlauf [53], Corollary 2.2).

(b) By Masuda [42], Theorem 4.3, the multivariate Ornstein-Uhlenbeck process $(\mathbf{Z}(t))_{t\geq 0}$ is ergodic and strongly mixing with geometric rate. Hence, the same holds for $(\mathbf{Z}(k))_{k\in\mathbb{N}}$. Thus, (2.11) is a conclusion of Phillips and Durlauf [53], Corollary 2.2.

Since the components $((\mathbf{Z}(k)\mathbf{Z}(k)')_{ij})_{k\in\mathbb{N}}$, $i, j = 1, \ldots, d$ of $(\mathbf{Z}(k)\mathbf{Z}(k)')_{k\in\mathbb{N}}$ are also ergodic and strongly mixing, the statement (2.10) follows from the ergodic theorem (cf. Shiryaev [66], Theorem 3 on p.413).

(c) follows with the same arguments as in *(b)*. □

5.2 Proofs of Section 3

Proof of Theorem 3.4.
(a) Since $\mathbb{X}_n' = \mathbf{A}\mathbb{Y}_n' + \mathbb{Z}_n'$ we have

$$\widehat{\mathbf{A}}_n - \mathbf{A} = \mathbf{A}\mathbb{Y}_n'\mathbb{Y}_n(\mathbb{Y}_n'\mathbb{Y}_n)^{-1} + \mathbb{Z}_n'\mathbb{Y}_n(\mathbb{Y}_n'\mathbb{Y}_n)^{-1} - \mathbf{A} = \mathbb{Z}_n'\mathbb{Y}_n(\mathbb{Y}_n'\mathbb{Y}_n)^{-1}. \tag{5.23}$$

This gives

$$na_nb_n^{-1}\left(\widehat{\mathbf{A}}_n - \mathbf{A}\right) = na_nb_n^{-1}(\mathbb{Z}_n'\mathbb{Y}_n)(\mathbb{Y}_n'\mathbb{Y}_n)^{-1} = \left(b_n^{-1}\mathbb{Z}_n'\mathbb{Y}_na_n^{-1}\right)\left(n^{-1}(a_n^{-1}\mathbb{Y}_n'\mathbb{Y}_na_n^{-1})\right)^{-1}.$$

Now we will prove the convergence

$$\begin{aligned}&\left(b_n^{-1}\mathbb{Z}_n'\mathbb{Y}_na_n^{-1}, n^{-1}(a_n^{-1}\mathbb{Y}_n'\mathbb{Y}_na_n^{-1})\right)\\ &\quad\Longrightarrow \left(\mathbf{S}_2(1)\mathbf{S}_1(1)' - \int_0^1 \mathbf{S}_2(s-)\mathrm{d}\mathbf{S}_1(s)', \int_0^1 \mathbf{S}_1(s)\mathbf{S}_1(s)'\mathrm{d}s\right)\end{aligned} \tag{5.24}$$

in $\mathbb{R}^{d\times q}\times\mathbb{R}^{q\times q}$ as $n\to\infty$, giving us the claim by a continuous mapping theorem, since

$$\mathbb{P}\left(\det\left(\int_0^1 \mathbf{S}_1(s)\mathbf{S}_1(s)'\mathrm{d}s\right) = 0\right) = 0$$

by $\mu_1(\{\mathbf{0}_{i-1}\}\times\overline{\mathbb{R}}\backslash\{0\}\times\{\mathbf{0}_{q-i}\}) > 0$ and $\mathbf{\Omega}_1$ invertible, respectively. We define the processes

$$\mathbf{S}_{1,n}(t) := a_n^{-1}\sum_{k=1}^{\lfloor nt\rfloor}\Delta\mathbf{L}_1(k) \qquad \text{and} \qquad \mathbf{S}_{2,n}(t) := b_n^{-1}\sum_{k=1}^{\lfloor nt\rfloor}\mathbf{Z}(k) \quad \text{for } t\geq 0.$$

We get for the first term on the left-hand side of (5.24), by assumption *(i)-(iii)*

$$\begin{aligned}a_n^{-1}b_n^{-1}\mathbb{Z}_n'\mathbb{Y}_n &= \sum_{k=1}^n b_n^{-1}\mathbf{Z}(k)\left(a_n^{-1}\sum_{j=1}^k\Delta\mathbf{L}_1(j)\right)' + o_p(1)\\ &= \left(b_n^{-1}\sum_{k=1}^n\mathbf{Z}(k)\right)\left(a_n^{-1}\sum_{j=1}^n\Delta\mathbf{L}_1(j)\right)' - \sum_{j=1}^n\left(b_n^{-1}\sum_{k=1}^{j-1}\mathbf{Z}(k)\right)\left(a_n^{-1}\Delta\mathbf{L}_1(j)\right)' + o_p(1)\\ &= \mathbf{S}_{2,n}(1)\mathbf{S}_{1,n}(1)' - \int_0^1\mathbf{S}_{2,n}(s-)\,\mathrm{d}\mathbf{S}_{1,n}(s)' + o_p(1),\end{aligned} \tag{5.25}$$

and for the second term,

$$\begin{aligned}n^{-1}a_n^{-2}\mathbb{Y}_n'\mathbb{Y}_n &= n^{-1}\sum_{k=1}^n\left(a_n^{-1}\sum_{j=1}^k\Delta\mathbf{L}_1(j)\right)\left(a_n^{-1}\sum_{j=1}^k\Delta\mathbf{L}_1(j)\right)' + o_p(1) \qquad (5.26)\\ &= n^{-1}\sum_{k=1}^n\mathbf{S}_{1,n}\left(\frac{k}{n}\right)\mathbf{S}_{1,n}\left(\frac{k}{n}\right)' + o_p(1)\\ &= \int_0^1\mathbf{S}_{1,n}(s)\mathbf{S}_{1,n}(s)'\mathrm{d}s + o_p(1).\end{aligned}$$

Since $(\mathbf{S}_1(t))_{t\geq 0}$ and $(\mathbf{S}_2(t))_{t\geq 0}$ are both not necessarily of unbounded variation (depending on α and β, respectively), and $(\mathbf{S}_{2,n}(t))_{t\geq 0}$ does not necessarily converges in the Skorokhod J_1 topology (only for $\beta = 2$), we can not apply the continuous mapping theorem in (5.25), in contrast to (5.26). However, we will show that in $\mathbb{R}^q \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathbb{R}^{d\times d} \times \mathbb{R}^{q\times q} \times \mathbb{R}^{d\times q}$ as $n \to \infty$,

$$\begin{aligned}\Big(\mathbf{S}_{1,n}(1), \mathbf{S}_{2,n}(1), \mathbf{S}_n^{(3)}(1), \mathbf{S}_n^{(4)}(1), \int_0^1 \mathbf{S}_{1,n}(s)\mathbf{S}_{1,n}(s)'\mathrm{d}s, \int_0^1 \mathbf{S}_{2,n}(s-)\,\mathrm{d}\mathbf{S}_{1,n}(s)'\Big) \\ \Longrightarrow \quad \Big(\mathbf{S}_1(1), \mathbf{S}_2(1), \mathbf{S}_3(1), \mathbf{S}_4(1), \int_0^1 \mathbf{S}_1(s)\mathbf{S}_1(s)'\mathrm{d}s, \int_0^1 \mathbf{S}_2(s-)\,\mathrm{d}\mathbf{S}_1(s)'\Big), \end{aligned} \tag{5.27}$$

where the main part is to prove

$$\int_0^1 \mathbf{S}_{2,n}(s-)\,\mathrm{d}\mathbf{S}_{1,n}(s)' \Longrightarrow \int_0^1 \mathbf{S}_2(s-)\,\mathrm{d}\mathbf{S}_1(s)' \quad \text{as } n \to \infty. \tag{5.28}$$

We define

$$\widetilde{\mathbf{S}}_{2,n}(t) = b_n^{-1} \sum_{j=0}^{\infty} \mathrm{e}^{-\mathbf{\Lambda} j} \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{\xi}_k \quad \text{for } t \geq 0. \tag{5.29}$$

By a special case of Proposition 2.1 (replacing $(\Delta \mathbf{L}_1(k))_{k\in\mathbb{N}}$ by $((\Delta \mathbf{L}_1(k)', \boldsymbol{\xi}_k' \sum_{j=0}^{\infty} \mathrm{e}^{-\mathbf{\Lambda}' j})')_{k\in\mathbb{N}}$ whose distribution is in $\mathcal{R}_{-\alpha}(a_n, \mu^*)$) we get

$$\begin{aligned}\Big(\mathbf{S}_{1,n}(1), \mathbf{S}_{2,n}(1), \mathbf{S}_n^{(3)}(1), \mathbf{S}_n^{(4)}(1), ((\mathbf{S}_{1,n}(t)', \widetilde{\mathbf{S}}_{2,n}(t)')')_{t\geq 0}\Big) \\ \Longrightarrow \quad \big(\big(\mathbf{S}_1(1), \mathbf{S}_2(1), \mathbf{S}_3(1), \mathbf{S}_4(1), ((\mathbf{S}_1(t)', \mathbf{S}_2(t)')')_{t\geq 0}\big)\big)\end{aligned}$$

in $\mathbb{R}^q \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathbb{R}^{d\times d} \times (\mathbb{D}[0,1], \mathbb{R}^{q+d})$. A straightforward conclusion of the continuous mapping theorem is then

$$\begin{aligned}\Big(\mathbf{S}_{1,n}(1), \mathbf{S}_{2,n}(1), \mathbf{S}_n^{(3)}(1), \mathbf{S}_n^{(4)}(1), \int_0^1 \mathbf{S}_{1,n}(s)\mathbf{S}_{1,n}(s)'\mathrm{d}s, ((\mathbf{S}_{1,n}(t)', \widetilde{\mathbf{S}}_{2,n}(t)')')_{t\geq 0}\Big) \\ \Longrightarrow \quad \Big(\mathbf{S}_1(1), \mathbf{S}_2(1), \mathbf{S}_3(1), \mathbf{S}_4(1), \int_0^1 \mathbf{S}_1(s)\mathbf{S}_1(s)'\mathrm{d}s, ((\mathbf{S}_1(t)', \mathbf{S}_2(t)')')_{t\geq 0}\Big).\end{aligned}$$

Since $(\mathbf{S}_{1,n}(t))_{t\geq 0}$ is *P-UT* (Predictably Uniformly Tight; see Jacod and Shiryaev [30]) by Lemma 5.5, a result of Jacod and Shiryaev [30], Theorem VI.6.22, is that as $n \to \infty$,

$$\begin{aligned}\Big(\mathbf{S}_{1,n}(1), \mathbf{S}_{2,n}(1), \mathbf{S}_n^{(3)}(1), \mathbf{S}_n^{(4)}(1), \int_0^1 \mathbf{S}_{1,n}(s)\mathbf{S}_{1,n}(s)'\mathrm{d}s, \int_0^1 \widetilde{\mathbf{S}}_{2,n}(s-)\,\mathrm{d}\mathbf{S}_{1,n}(s)'\Big) \\ \Longrightarrow \quad \Big(\mathbf{S}_1(1), \mathbf{S}_2(1), \mathbf{S}_3(1), \mathbf{S}_4(1), \int_0^1 \mathbf{S}_1(s)\mathbf{S}_1(s)'\mathrm{d}s, \int_0^1 \mathbf{S}_2(s-)\,\mathrm{d}\mathbf{S}_1(s)'\Big)\end{aligned} \tag{5.30}$$

in $\mathbb{R}^q \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathbb{R}^{d\times d} \times \mathbb{R}^{q\times q} \times \mathbb{R}^{d\times q}$. Furthermore, we have a kind of Beveridge-Nelsen decomposition

$$\mathbf{S}_{2,n}(t) = \widetilde{\mathbf{S}}_{2,n}(t) + b_n^{-1} \sum_{j=1}^{\infty} \mathrm{e}^{-\mathbf{\Lambda} j} \left[\mathbf{Z}(0) - \mathbf{Z}(\lfloor nt \rfloor)\right]$$

giving

$$\int_0^1 \mathbf{S}_{2,n}(s-)\,\mathrm{d}\mathbf{S}_{1,n}(s)' \tag{5.31}$$
$$= \int_0^1 \widetilde{\mathbf{S}}_{2,n}(s-)\,\mathrm{d}\mathbf{S}_{1,n}(s)' + \sum_{j=1}^{\infty} \mathrm{e}^{-\mathbf{\Lambda} j}\left[b_n^{-1}\mathbf{Z}(0)\mathbf{S}_{1,n}(1) - \sum_{k=1}^{n} b_n^{-1}\mathbf{Z}(k-1)a_n^{-1}\Delta\mathbf{L}_1(k)'\right].$$

Applying Lemma 5.6 from below and $\mathbf{S}_{1,n}(1) \Longrightarrow \mathbf{S}_1(1)$ as $n \to \infty$ gives

$$\sum_{j=1}^{\infty} \mathrm{e}^{-\mathbf{\Lambda} j}\left[b_n^{-1}\mathbf{Z}(0)\mathbf{S}_{1,n}(1) - \sum_{k=1}^{n} b_n^{-1}\mathbf{Z}(k-1)a_n^{-1}\Delta\mathbf{L}_1(k)'\right] \xrightarrow{\mathbb{P}} \mathbf{0}_{d\times q} \quad \text{as } n\to\infty. \tag{5.32}$$

Finally, from (5.30)-(5.32) the statement (5.27) follows. Then the theorem is a conclusion of (5.27) and the continuous mapping theorem. □

Remark 5.4 Under the stronger assumption $\mathbb{E}\|\mathbf{L}(1)\|^r < \infty$ for some $r > 4$, the convergence of (5.28) in the Skorokhod J_1 topology follows directly from Ibragimov and Phillips [29], Theorem 4.3. In particular, in that case $(\mathbf{S}_{2,n})_{n\in\mathbb{N}}$ converges in the Skorokhod J_1 topology. □

Lemma 5.5 *Let either of the Assumptions 3.1 – 3.3 hold. Then for any $t > 0$ the sequence of random vectors $(\mathbf{S}_{1,n}(t))_{n\in\mathbb{N}}$ and $(\widetilde{\mathbf{S}}_{2,n}(t))_{n\in\mathbb{N}}$ are P-UT.*

Proof. We show that $(\mathbf{S}_{1,n}(t))_{n\in\mathbb{N}}$ is *P-UT* for some $t > 0$. The proof of the *P-UT*ness of $(\widetilde{\mathbf{S}}_{2,n}(t))_{n\in\mathbb{N}}$ is analogous. Thus, we define for $s \geq 0$

$$\begin{aligned}
\mathbf{M}_n(s) &= a_n^{-1}\sum_{k=1}^{\lfloor ns\rfloor}\left(\Delta\mathbf{L}_1(k)\mathbb{1}_{\{\|\Delta\mathbf{L}_1(k)\|\leq a_n\}} - \mathbb{E}(\Delta\mathbf{L}_1(1)\mathbb{1}_{\{\|\Delta\mathbf{L}_1(1)\|\leq a_n\}})\right),\\
\mathbf{D}_n(s) &= \lfloor ns\rfloor a_n^{-1}\mathbb{E}\left(\Delta\mathbf{L}_1(1)\mathbb{1}_{\{\|\Delta\mathbf{L}_1(1)\|\leq a_n\}}\right),\\
\mathbf{V}_n(s) &= a_n^{-1}\sum_{k=1}^{\lfloor ns\rfloor}\Delta\mathbf{L}_1(k)\mathbb{1}_{\{\|\Delta\mathbf{L}_1(k)\|> a_n\}},
\end{aligned}$$

and the filtration $(\mathcal{F}_s^n)_{s\geq 0} = (\sigma(\Delta\mathbf{L}_1(k) : k \leq \lfloor ns\rfloor))_{s\geq 0}$, $n \in \mathbb{N}$. It is obvious that $(\mathbf{M}_n(s))_{s\geq 0}$ is a $(\mathcal{F}_s^n)_{s\geq 0}$ martingale for any $n \in \mathbb{N}$ and in particular, a local martingale. All three processes are adapted with respect to $(\mathcal{F}_s^n)_{s\geq 0}$ and we have the semimartingale decomposition

$$\mathbf{S}_{1,n}(s) = \mathbf{M}_n(s) + \mathbf{D}_n(s) + \mathbf{V}_n(s) \quad \text{for } s \geq 0.$$

If $(\mathbf{M}_n(t))_{n\in\mathbb{N}}$, $(\mathbf{D}_n(t))_{n\in\mathbb{N}}$ and $(\mathbf{V}_n(t))_{n\in\mathbb{N}}$ are *P-UT* then VI.6.4 in Jacod and Shiryaev [30] gives that the sum $(\mathbf{S}_{1,n}(t))_{n\in\mathbb{N}}$ is also *P-UT*.

Let $\mathrm{VT}_t(\mathbf{W}) = \sup_{i=1,\ldots,d}\mathrm{VT}_t(\mathbf{W}_i)$ for $t \geq 0$ denote the variation process of the càdlàg stochastic process $\mathbf{W} = (\mathbf{W}_1(t),\ldots,\mathbf{W}_d(t))_{t\geq 0}$. To prove the uniform tightness of $(\mathbf{D}_n(t))_{n\in\mathbb{N}}$ and $(\mathbf{V}_n(t))_{n\in\mathbb{N}}$ it is sufficient to show that $(\mathrm{VT}_t(\mathbf{D}_n))_{n\in\mathbb{N}}$ and $(\mathrm{VT}_t(\mathbf{V}_n))_{n\in\mathbb{N}}$ are tight; see Jacod and Shiryaev [30], VI.6.6. We start with the verification of the tightness of $(\mathrm{VT}_t(\mathbf{D}_n))_{n\in\mathbb{N}}$ by showing

that it is uniformly bounded. In the following we will use the next inequality for random variables, e.g, W, with finite second moments. That is Markov's inequality gives for any $\delta \in (0,2)$,

$$\mathbb{E}(|W|^\delta \mathbb{1}_{\{|W|>|x|\}}) \quad = \quad |x|^\delta \mathbb{P}(|W|>x) + \int_{|x|^\delta}^\infty \mathbb{P}(|W|^\delta > y)\, dy \leq C_1 \mathbb{E}|W|^2 |x|^{-2+\delta}. \tag{5.33}$$

If $0<\alpha<1$, Proposition 5.1, Karamata's and Potter's Theorem give the uniform bound

$$\sup_{n\in\mathbb{N}} \mathrm{VT}_t(\mathbf{D}_n) \leq C_2 \sup_{n\in\mathbb{N}} nta_n^{-1} \mathbb{E}(\|\Delta \mathbf{L}_1(1)\|\, \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\|\leq a_n\}}) \leq C_3 t. \tag{5.34}$$

If $\alpha = 1$, then by symmetry $\mathbf{D}_n = \mathbf{0}_q$. If $1<\alpha<2$ and $\mathbb{E}(\mathbf{L}_2(1)) = \mathbf{0}_d$, Karamata's and Potter's Theorem result in the uniform bound

$$\sup_{n\in\mathbb{N}} \mathrm{VT}_t(\mathbf{D}_n) \leq C_4 \sup_{n\in\mathbb{N}} nta_n^{-1} \mathbb{E}(\|\Delta \mathbf{L}_1(1)\|\, \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\|> a_n\}}) \leq C_5 t. \tag{5.35}$$

Finally, for $\alpha = 2$ the conclusion (5.35) follows from (5.33) with $\delta = 1$. To conclude, for $\eta \geq \max(C_3, C_5)t$, we have

$$\sup_{n\in\mathbb{N}} \mathbb{P}(\mathrm{VT}_t(\mathbf{D}_n) > \eta) = 0,$$

which results in the tightness of $(\mathrm{VT}_t(\mathbf{D}_n))_{n\in\mathbb{N}}$.

For the proof of the tightness of $(\mathrm{VT}_t(\mathbf{V}_n))_{n\in\mathbb{N}}$ we distinguish the cases $0<\alpha\leq 1$ and $\alpha>1$. Let $0<\alpha\leq 1$ and $0<\delta<\alpha$. Then

$$(\mathrm{VT}_t(\mathbf{V}_n))^\delta \leq C_6 a_n^{-\delta} \sum_{k=1}^{\lfloor nt\rfloor} \|\Delta \mathbf{L}_1(k)\|^\delta \mathbb{1}_{\{\|\Delta \mathbf{L}_1(k)\|>a_n\}},$$

and by Markov's inequality and Karamata's Theorem we obtain

$$\sup_{n\in\mathbb{N}} \mathbb{P}(\mathrm{VT}_t(\mathbf{V}_n) > \eta) \leq C_7 \eta^{-\delta} \sup_{n\in\mathbb{N}} a_n^{-\delta} \sum_{k=1}^{\lfloor nt\rfloor} \mathbb{E}(\|\Delta \mathbf{L}_1(k)\|^\delta \mathbb{1}_{\{\|\Delta \mathbf{L}_1(k)\|>a_n\}}) \leq C_8 \eta^{-\delta} t \overset{\eta\to\infty}{\longrightarrow} 0. \tag{5.36}$$

If $\alpha>1$ then Markov's inequality and (5.35) give

$$\sup_{n\in\mathbb{N}} \mathbb{P}(\mathrm{VT}_t(\mathbf{V}_n) > \eta) \leq C_9 \eta^{-1} \sup_{n\in\mathbb{N}} a_n^{-1} \sum_{k=1}^{\lfloor nt\rfloor} \mathbb{E}(\|\Delta \mathbf{L}_1(k)\| \mathbb{1}_{\{\|\Delta \mathbf{L}_1(k)\|>a_n\}}) \leq C_{10} \eta^{-1} t \overset{\eta\to\infty}{\longrightarrow} 0. \tag{5.37}$$

Hence, $(\mathrm{VT}_t(\mathbf{V}_n))_{n\in\mathbb{N}}$ is also tight.

If we show that $([\mathbf{M}_n, \mathbf{M}_n]_t)_{n\in\mathbb{N}}$ is tight, then the *P-UT*ness of $(\mathbf{M}_n(t))_{t\geq 0}$ follows by Jacod and Shiryaev [30], Proposition VI.6.13. Here, we use again Markov's inequality, and Karamata's Theorem if $\alpha<2$ and $\mathbb{E}(\|\Delta \mathbf{L}_1(1)\|^2 \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\|\leq a_n\}}) \leq \mathbb{E}\|\Delta \mathbf{L}_1(1)\|^2$ if $\alpha = 2$, which results in

$$\sup_{n\in\mathbb{N}} \mathbb{P}(\|[\mathbf{M}_n, \mathbf{M}_n]_t\| > \eta) \leq \eta^{-1} \sup_{n\in\mathbb{N}} a_n^{-2} n \mathbb{E}(\|\Delta \mathbf{L}_1(1)\|^2 \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\|\leq a_n\}}) \leq C_{11} \eta^{-1} \longrightarrow 0 \quad \text{ as } \eta\to\infty.$$

Finally, $([\mathbf{M}_n, \mathbf{M}_n]_t)_{n\in\mathbb{N}}$ is tight as well. □

Lemma 5.6 *Let the assumptions of Theorem 3.4 hold. Then as $n \to \infty$,*

$$a_n^{-1} b_n^{-1} \sum_{k=1}^{n} \mathbf{Z}(k-1) \Delta \mathbf{L}_1(k)' \xrightarrow{\mathbb{P}} \mathbf{0}_{d \times q} \,.$$

Proof. We divide the sum in four parts, namely

$$\begin{aligned}
& a_n^{-1} b_n^{-1} \sum_{k=1}^{n} \mathbf{Z}(k-1) \Delta \mathbf{L}_1(k)' \\
&= a_n^{-1} b_n^{-1} \sum_{k=1}^{n} \mathbf{Z}(k-1) \mathbb{1}_{\{\|\mathbf{Z}(k-1)\| \le b_n\}} (\Delta \mathbf{L}_1(k)' \mathbb{1}_{\{\|\Delta \mathbf{L}_1(k)\| \le a_n\}} - \mathbb{E}(\Delta \mathbf{L}_1(1)' \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\| \le a_n\}})) \\
&\quad + a_n^{-1} b_n^{-1} \sum_{k=1}^{n} \mathbf{Z}(k-1) \mathbb{1}_{\{\|\mathbf{Z}(k-1)\| \le b_n\}} \mathbb{E}(\Delta \mathbf{L}_1(1)' \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\| \le a_n\}}) \\
&\quad + a_n^{-1} b_n^{-1} \sum_{k=1}^{n} \mathbf{Z}(k-1) \mathbb{1}_{\{\|\mathbf{Z}(k-1)\| > b_n\}} \Delta \mathbf{L}_1(k)' \mathbb{1}_{\{\|\Delta \mathbf{L}_1(k)\| \le a_n\}} \\
&\quad + a_n^{-1} b_n^{-1} \sum_{k=1}^{n} \mathbf{Z}(k-1) \Delta \mathbf{L}_1(k)' \mathbb{1}_{\{\|\Delta \mathbf{L}_1(k)\| > a_n\}} \\
&=: I^{(n,1)} + I^{(n,2)} + I^{(n,3)} + I^{(n,4)}. \qquad (5.38)
\end{aligned}$$

Now we will investigate the four terms in (5.38). The sequence of random matrices

$$(\mathbf{Z}(k-1) \mathbb{1}_{\{\|\mathbf{Z}(k-1)\| \le b_n\}} [\Delta \mathbf{L}_1(k)' \mathbb{1}_{\{\|\Delta \mathbf{L}_1(k)\| \le a_n\}} - \mathbb{E}(\Delta \mathbf{L}_1(1)' \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\| \le a_n\}})])_{k \in \mathbb{N}}$$

is uncorrelated. Thus, for any (i,j)-component $I_{ij}^{(n,1)}$ of $I^{(n,1)}$ there exists a constant $C_{ij} > 0$ such that

$$\mathbb{E}(I_{ij}^{(n,1)})^2 \le n b_n^{-2} C_{ij} \mathbb{E}(\|\mathbf{Z}(1)\|^2 \mathbb{1}_{\{\|\mathbf{Z}(1)\| \le b_n\}}) a_n^{-2} \mathbb{E}(\|\Delta \mathbf{L}_1(1)\|^2 \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\| \le a_n\}}),$$

which tends to 0 as $n \to \infty$ by Karamata's Theorem if $\alpha < 2$ and $\beta < 2$, respectively. If $\alpha = 2$ then $\mathbb{E}(\|\Delta \mathbf{L}_1(1)\|^2 \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\| \le a_n\}}) \le \mathbb{E}\|\Delta \mathbf{L}_1(1)\|^2$ and similarly for $\beta = 2$. This results in $\lim_{n \to \infty} \mathbb{E}(\|I^{(n,1)}\|^2) = 0$.

If $0 < \alpha < 1$, then

$$\mathbb{E}\|I^{(n,2)}\| \le b_n^{-1} \mathbb{E}(\|\mathbf{Z}(1)\| \mathbb{1}_{\{\|\mathbf{Z}(1)\| \le b_n\}}) n a_n^{-1} \mathbb{E}(\|\Delta \mathbf{L}_1(1)\| \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\| \le a_n\}}) \overset{n \to \infty}{\longrightarrow} 0$$

by Karamata's Theorem, where $\mathbb{E}(\|\mathbf{Z}(1)\| \mathbb{1}_{\{\|\mathbf{Z}(1)\| \le b_n\}}) \le \mathbb{E}\|\mathbf{Z}(1)\|$ if $\beta > 1$. Moreover, $I^{(n,2)} = \mathbf{0}_{d \times q}$ if $\alpha = 1$. Let $1 < \alpha \le 2$. Since $\mathbb{E}(\Delta \mathbf{L}_1(1)) = \mathbf{0}_q$ and hence, $\mathbb{E}(\Delta \mathbf{L}_1(1) \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\| \le a_n\}}) = \mathbb{E}(\Delta \mathbf{L}_1(1) \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\| > a_n\}})$ we have

$$\mathbb{E}\|I^{(n,2)}\| \le b_n^{-1} \mathbb{E}(\|\mathbf{Z}(1)\| \mathbb{1}_{\{\|\mathbf{Z}(1)\| \le b_n\}}) n a_n^{-1} \mathbb{E}(\|\Delta \mathbf{L}_1(1)\| \mathbb{1}_{\{\|\Delta \mathbf{L}_1(1)\| > a_n\}}).$$

Then $\mathbb{E}\|I^{(n,2)}\|$ tends to 0 by Karamata's Theorem if $\alpha < 2$ and (5.33) if $\alpha = 2$, respectively.

In the following, if $0 < \min(\alpha, \beta) \le 1$, we choose some $\delta \in (0, \min(\alpha, \beta))$ and define $\delta' := \delta$. $\delta' := 1$ if $\min(\alpha, \beta) \in (1, 2]$. Next,

$$\mathbb{E}\|I^{(n,3)}\|^{\delta'} \le n b_n^{-\delta'} \mathbb{E}(\|\mathbf{Z}(1)\|^{\delta'} \mathbb{1}_{\{\|\mathbf{Z}(1)\| > b_n\}}) a_n^{-\delta'} \mathbb{E}(\|\Delta \mathbf{L}_1(1)\|^{\delta'}),$$

which tends to 0 by Karamata's Theorem if $\beta < 2$ and (5.33) if $\beta = 2$, respectively.

Finally,

$$\mathbb{E}\|I^{(n,4)}\|^{\delta'} \le b_n^{-\delta'}\mathbb{E}\|\mathbf{Z}(1)\|^{\delta'} n a_n^{-\delta'}\mathbb{E}(\|\Delta\mathbf{L}_1(1)\|^{\delta'}\mathbb{1}_{\{\|\Delta\mathbf{L}_1(1)\|>a_n\}}),$$

which tends again to 0 by Karamata's Theorem if $\alpha < 2$ and (5.33) if $\alpha = 2$, respectively. □

Proof of Example 3.5.

(a) is clear.

(b) We give only a sketch of the proof. It is sufficient to show *(i)-(iii)* with $\mathbf{O}$ instead of $\boldsymbol{\zeta}$ since $\mathbf{B}$ is independent of $\mathbf{L}_1, \mathbf{L}_2$ and $\mathbf{L}_3$.

We distinguish three different cases to show *(i)*.

Case 1: If $\mathbb{E}\|\mathbf{L}_1(1)\|^2 < \infty$ and $\mathbb{E}\|\mathbf{L}_2(1)\|^2 < \infty$. Then $a_n = b_n = \sqrt{n}$, $\mathbb{E}(\mathbf{Z}(1)\boldsymbol{\zeta}(1)') = \mathbf{0}_{d\times q}$ and $\mathbb{E}(\|\mathbf{Z}(1)\|^2\|\boldsymbol{\zeta}(1)\|^2) < \infty$. Hence, *(i)* follows by the ergodic theorem (cf. proof of Proposition 2.4).

Case 2: $\mathbf{L}_1(1) \in \mathcal{R}_{-\alpha}(a_n)$ and $\mathbf{L}_2(1) \in \mathcal{R}_{-\alpha}(b_n)$. Then $\|\mathbf{O}(1)\| \in \mathcal{R}_{-\alpha}(a_n)$ and hence,

$$\lim_{n\to\infty} n\mathbb{P}(\|\mathbf{Z}(1)\|\|\mathbf{O}(1)\| > a_n b_n) = 0$$

by Cline [13], and *(i)* follows (cf. Fasen [24]).

Case 3: W.l.o.g. $\mathbf{L}_2(1) \in \mathcal{R}_{-\beta}(b_n)$, and $\mathbf{L}_1(1) \in \mathcal{R}_{-\alpha}(a_n)$ with $\alpha > \beta$ or $\mathbb{E}\|\mathbf{L}_1(1)\|^2 < \infty$. Then $\mathbb{E}\|\boldsymbol{\zeta}(1)\|^\beta < \infty$ such that $\|\mathbf{O}(1)\|\|\mathbf{Z}(1)\| \in \mathcal{R}_{-\beta}(b_n)$. Thus, we obtain *(i)* again.

Next, *(ii)* is a conclusion of Lemma 5.6 since $\mathbf{L}_1$ and $\mathbf{O}$ are independent. Finally, *(iii)* follows directly from Proposition 2.1 and Proposition 2.4. □

5.3 Proofs of Section 4

Proof of Theorem 4.1.

First,

$$\begin{aligned} n\widehat{\boldsymbol{\Omega}}_n &= ((\mathbf{A}-\widehat{\mathbf{A}}_n)\mathbb{Y}_n' + \mathbb{Z}_n')((\mathbf{A}-\widehat{\mathbf{A}}_n)\mathbb{Y}_n' + \mathbb{Z}_n')' \\ &= (\mathbf{A}-\widehat{\mathbf{A}}_n)\mathbb{Y}_n'\mathbb{Y}_n(\mathbf{A}-\widehat{\mathbf{A}}_n)' + \mathbb{Z}_n'\mathbb{Y}_n(\mathbf{A}-\widehat{\mathbf{A}}_n)' + (\mathbf{A}-\widehat{\mathbf{A}}_n)\mathbb{Y}_n'\mathbb{Z}_n + \mathbb{Z}_n'\mathbb{Z}_n \\ &=: \boldsymbol{\varepsilon}_n^{(1)} + \mathbb{Z}_n'\mathbb{Z}_n. \end{aligned} \tag{5.39}$$

On the one hand, (5.23) and (5.24) result in

$$n b_n^{-2}\boldsymbol{\varepsilon}_n^{(1)} \Longrightarrow \boldsymbol{\varepsilon}^{(1)} \quad \text{as } n\to\infty, \tag{5.40}$$

where $\boldsymbol{\varepsilon}^{(1)}$ is some a.s. finite random matrix. On the other hand, by (5.27) we have as $n \to \infty$,

$$b_n^{-2}\mathbb{Z}_n'\mathbb{Z}_n \Longrightarrow \mathbf{S}_3(1). \tag{5.41}$$

Thus, (5.39)-(5.41) result in

$$n b_n^{-2}\widehat{\boldsymbol{\Omega}}_n \Longrightarrow \mathbf{S}_3(1). \tag{5.42}$$

Moreover, by (5.24) as $n \to \infty$,

$$n^{-1}a_n^{-2}(\mathbb{Y}_n'\mathbb{Y}_n) \Longrightarrow \int_0^1 \mathbf{S}_1(s)\mathbf{S}_1(s)'\mathrm{d}s.$$

Hence, applying the continuous mapping theorem, (5.27) and Theorem 3.4, respectively we obtain the result. □

Proof of Theorem 4.5. It is again an application of Theorem 3.4, Proposition 2.1, (5.42), (5.27) and the continuous mapping theorem. □

Acknowledgement

My special thanks go to Holger Rootzén, Carl Lindberg and their colleagues at the Department of Mathematical Statistics at Chalmers University of Technology for their hospitality during my visit in fall 2008 and, in particular, for raising my attention to the topic of cointegration. I take pleasure in thanking Christoph Ferstl, who has written his diploma thesis [25] at the Technische Universität München using some preliminary notes of some earlier results for having patience with my research progress. Furthermore, I am deeply grateful to some anonymous referees and to Claudia Klüppelberg for some useful comments. Finally, I would like to thank John Nolan for providing me with the toolbox STABLE for Matlab.

References

[1] Aban, B. and Meerschaert, M. M. (2004). Generalized least-squares estimators for the thickness of heavy tails. *J. Statist. Plann. Inference* **119**, 341–152.

[2] Adler, R. J., Feldman, R. E., and Taqqu, M. S. (Eds.) (1998). *A Practical Guide to Heavy Tails*. Birkhäuser, Boston.

[3] Avram, F. and Taqqu, M. (1991). Weak convergence of sums of moving average processes in the α-stable domain of attraction. *Ann. Probab.* **20**, 483–503.

[4] Bakirov, N. K. and Székely, G. J. (2006). Students t-test for Gaussian scale mixtures. *J. Math. Sci.* **139**, 6497–6505.

[5] Basrak, B., Davis, R. A., and Mikosch, T. (2002). Regular variation of GARCH processes. *Stoch. Proc. Appl.* **99**, 95–116.

[6] Benth, F. E. and Benth, J. S. (2006). Analytic approximatins for the price dynamics of spark spread options. *Stud. Nonlinear Dyn. Econom.* **10** (2). Article 8.

[7] Benth, F. E., Benth, J. S., and Koekebakker, S. (2008). *Stochastic Modelling of Electricity and Related Markets*. World Scientific, Singapore.

[8] Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[9] Brenner, R. J. and Kroner, K. F. (1995). Arbitrage, cointegration, and testing the unbiasedness hypothesis in financial markets. *J. Financ. Quant. Anal.* **10**, 23–42.

[10] Caner, M. (1997). Weak convergence of matrix stochastic integral with stable processes. *Econometric Theory* **13**, 506–528.

[11] Chan, N. H. (2009). Time series with roots on or near the unit circle. In: T. G. Andersen, R. A. Davis, J.-P. Kreiß, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp. 695–707. Springer, Berlin.

[12] Chan, N. H. and Tran, L. T. (1989). On the first order autoregressive process with infinite variance. *Econom. Theory* **5**, 354–362.

[13] Cline, D. B. H. (1986). Convolution tails, product tails and domains of attraction. *Probab. Theory Related Fields* **72**, 529–557.

[14] Comte, F. (1999). Discrete and continuous time cointegration. *J. Econometrics* **88**, 207–222.

[15] Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes. Vol I: Elementary Theory and Methods*. 2nd edn. Springer, New York.

[16] Davis, R., Marengo, J., and Resnick, S. (1985). Extremal properties of a class of multivariate moving averages. *Bull. Internat. Statist. Institute* **51**, 26.1–26.14.

[17] Davis, R. and Resnick, S. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13**, 179–195.

[18] Duan, J.-C. and Pliska, S. R. (2004). Option valuation with co-inetgrated assets. *J. Econom. Dynam. Control* **28**, 727–754.

[19] Ekström, E., Lindberg, C., and Tysk, J. (2009). Optimal liquidation of pairs trade. Submitted for Publication.

[20] Elliott, R. J., Von der Hoeck, J., and Malcom, W. P. (2005). Pairs trading. *Quant. Finance* **5**, 271–276.

[21] Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica* **55**, 251–276.

[22] Fama, E. F. (1965). The behavior of stock market prices. *J. Business* **38**, 34–105.

[23] Fasen, V. (2011). Limit theory for high frequency sampled MCARMA models. Submitted for publication.

[24] Fasen, V. (2011). Statistical estimation of multivariate Ornstein-Uhlenbeck processes and applications to co-integration. *J. Econometrics*. To appear.

[25] Ferstl, C. (2009). *Cointegration in Discrete and Continuous time*. Diploma Thesis, Technische Universität München.

[26] Gatev, E., Goetzmann, W. N., and Rouwenhorst, K. G. (2006). Pairs trading: Performance of a relative-value arbitrage rule. *Rev. Financ. Stud.* **19**, 797–827.

[27] Granger, C. W. J. (1981). Some properties of time series data and their use in econometric model specification. *J. Econometrics* **16**, 121–130.

[28] Hult, H. and Samorodnitsky, G. (2008). Tail probabilities for infinite series of regularly varying random vectors. *Bernoulli* **14**, 838–864.

[29] Ibragimov, R. and Phillips, P. C. B. (2008). Regression asymptotics using martingale convergence methods. *Econom. Theory* **24**, 888–947.

[30] Jacod, J. and Shiryaev, A. N. (2002). *Limit Theorems for Stochastic Processes*. 2nd edn. Springer, Heidelberg.

[31] Jeanblanc, M., Yor, M., and Chesney, M. (2009). *Mathematical Methods for Financial Markets*. Springer.

[32] Jessen, A. and Mikosch, T. (2006). Regularly varying functions. *Publications de l'Institut Mathématique* **80**, 171–192.

[33] Johansen, S. (1996). *Likelihood-Based Inference on Cointegration in the Vector Autoregressive Model*. Oxford University Press, Oxford.

[34] Kallenberg, O. (1997). *Foundations of Modern Probability*. Springer, New York.

[35] Kurtz, T. and Protter, P. (1991). Characterizing the weak convergence of stochastic integrals. *Stochastic Analysis* **167**, 255–259.

[36] Lindskog, F. (2004). *Multivariate Extremes and Regular Variation for Stochastic Processes*. Ph.D. thesis, ETH Zürich.

[37] Loretan, M. and Phillips, P. C. B. (1994). Testing the covariance stationarity of heavy tailed time series. *J. Emp. Finance* **1**, 211–248.

[38] Lütkepohl, H. (2007). *New Introduction to Multiple Time Series Analysis*. 2nd edn. Springer, Berlin.

[39] Lucia, J. J. and Schwartz, E. (2002). Electricity prices and power derivatives: Evidence from the nordic power exchange. *Rev. Derivatives Res.* **5**, 5–50.

[40] Mandelbrot, B. B. (1963). The variation of certain speculative prices. *J. Business* **36**, 394–419.

[41] Marquardt, T. and Stelzer, R. (2007). Multivariate CARMA processes. *Stoch. Proc. Appl.* **117**, 96–120.

[42] Masuda, H. (2004). On multidimensional Ornstein-Uhlenbeck processes driven by a general Lévy process. *Bernoulli* **10** (1), 97–120.

[43] McCulloch, J. H. (1997). Measuring tail thickness to estimate the stable index α: a critique. *J. Bus. Econom. Statist.* **15**, 74–81.

[44] McElroy, T. and Politis, D. N. (2002). Robust inference for the mean in the presence of serial correlation and heavy tailed distributions. *Econom. Theory* **18**, 1019–1039.

[45] Meerschaert, M. M. and Scheffler, H.-P. (2000). Moving averages of random vectors with regularly varying tails. *J. Time Ser. Anal.* **21**, 297–328.

[46] Mittnik, S., Paulauskas, V., and Rachev, S. T. (2001). Statistical inference in regression with heavy-tailed integrated variables. *Math Comput. Modelling* **34**, 1145–1158.

[47] Paulauskas, A., Rachev, S., and Fabozzi, F. J. (2011). Comment on "weak convergence to a matrix stochastic integral with stable processes". *Econom. Theory.* To appear.

[48] Paulauskas, V. and Rachev, S. (1998). Cointegrated processes with infinite variance innovations. *Ann. Appl. Probab.* **8**, 775–792.

[49] de la Peña, V. H., Shao, Q., and Lai, T. L. (2009). *Self-Normalized Processes: Limit Theory and Statistical Applications*. Springer, Berlin.

[50] Phillips, P. C. B. (1987). Towards a unified asymptotic theory for autoregression. *Biometrika* **74**, 535–547.

[51] Phillips, P. C. B. (1990). Time series regression with a unit root and infinite-variance errors. *Econom. Theory* **6**, 44–62.

[52] Phillips, P. C. B. (1991). Error correction and long-run equilibrium in continuous time. *Econometrica* **59**, 967–980.

[53] Phillips, P. C. B. and Durlauf, S. N. (1986). Multiple time series regression with integrated processes. *Rev. Econom. Stud.* **LIII**, 473–495.

[54] Phillips, P. C. B. and Solo, V. (1992). Asymptotics for linear processes. *Ann. Statist.* **20**, 971–1001.

[55] Politis, D. N., Romano, J. P., and Wolf, M. (1999). *Subsampling*. Springer.

[56] Rachev, S. and Mittnik, S. T. (2000). *Stable Paretian Models in Finance*. Wiley.

[57] Rachev, S. T. (Ed.) (2003). *Handbook of Heavy Tailed Distributions in Finance*, vol. 1 of *Handbooks in Finance*. Elsevier, North-Holland.

[58] Resnick, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.

[59] Resnick, S. I. (1992). *Adventures in Stochastic Processes*. Birkhäuser, Boston.

[60] Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York.

[61] Rvačeva, E. L. (1962). On domains of attraction of multidimensional distributions. *Selected Translations in Math. Statist. Prob. Theory* **2**, 183–205.

[62] Samorodnitsky, G. and Taqqu, M. S. (1994). *Stable Non-Gaussian Random Processes*. Chapman & Hall, New York.

[63] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge.

[64] Sato, K.-I. and Yamazato, M. (1984). Operator selfdecomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stoch. Proc. Appl.* **17**, 73–100.

[65] Selivanvov, A. V. (2005). On the martingale measures in exponential Lévy models. *Theory Probab. Appl.* **49**, 261–274.

[66] Shiryaev, A. (1995). *Probability*. 2nd edn. Springer.

[67] Stockmarr, A. and Jacobsen, M. (1994). Gaussian diffusions and autoregressive processes: weak convergence and statistical inference. *Scand. J. Statist.* **21**, 403–429.

[68] Weron, R. (2006). *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. John Wiley & Sons.