Statistical estimation of multivariate Ornstein-Uhlenbeck processes and applications to co-integration

Vicky Fasen[*]

September 5, 2012

Abstract

Ornstein-Uhlenbeck models are continuous-time processes which have broad applications in finance as, e.g., volatility processes in stochastic volatility models or spread models in spread options and pairs trading. The paper presents a least squares estimator for the model parameter in a multivariate Ornstein-Uhlenbeck model driven by a multivariate regularly varying Lévy process with infinite variance. We show that the estimator is consistent. Moreover, we derive its asymptotic behavior and test statistics. The results are compared to the finite variance case. For the proof we require some new results on multivariate regular variation of products of random vectors and central limit theorems. Furthermore, we embed this model in the setup of a co-integrated model in continuous time.

[*]ETH Zürich, Rämistrasse 101, 8092 Zürich Switzerland, Tel.: +41-44-632 6820, e-mail: vicky.fasen@math.ethz.ch

1 Introduction

In this paper we investigate the asymptotic properties of the least squares estimator for the model parameter in a multivariate Ornstein-Uhlenbeck model. Ornstein-Uhlenbeck processes are natural extensions of autoregressive processes of order one in discrete time to continuous time. Hence, they belong to the class of continuous-time autoregressive moving average (CARMA) processes. Not only that the Ornstein-Uhlenbeck process itself is a CARMA process but also the state space representation of a CARMA process leads to a multivariate Ornstein-Uhlenbeck process. Applications of CARMA processes include econometrics (see Bergstrom (1990), Phillips (1974)), high-frequency financial econometrics (see Todorov (2009)), and financial mathematics (see Benth et al. (2010)). However, CARMA processes clearly have potential applications in all areas involving time series data, e.g. social sciences, medicine, biology or physics.

Typical stylized facts of high frequency financial time series as asset returns and exchange rates are jumps and a heavy tailed distribution which is peaked around zero. These characteristics were already noticed in the 60's by the influential works of Mandelbrot (1963) and Fama (1965). Thus, α-stable distributions as generalization of a Gaussian distribution have often been discussed as more realistic models for asset returns than the usual normal distribution; see Rachev et al. (1999). More applications of heavy tailed distributions in economics and finance can be found in Adler et al. (1998), Rachev and Mittnik (2000) and Rachev (2003). Processes exhibiting infinite second variance have not only been appeared in finance but also, e.g., in insurance, signal processing and teletraffic data. For an overview on the topic of heavy tailed distributions and their applications we refer to the excellent monograph of Resnick (2007). It is well known that for heavy tailed distributions standard statistical techniques do not apply in the usual way.

Further common features of high frequency financial times series are non-stationarity and time-varying volatility. A famous stochastic volatility model is the Ornstein-Uhlenbeck model propagated by Barndorff-Nielsen and Shephard (2001). They start from the classical Black-Scholes model for the log asset price

$$\mathrm{d}X(t) = \left(\mu - \sigma^2\delta\right)\,dt + \sigma\,\mathrm{d}B(t), \tag{1.1}$$

where $\mu, \delta \in \mathbb{R}$ are the instantaneous drift and the premium parameter, $\sigma > 0$ is the constant volatility and $(B(t))_{t \geq 0}$ is a Brownian motion, and plug in (1.1) the stochastic Ornstein-Uhlenbeck process $(\sigma^2(t))_{t \geq 0}$ as volatility process instead of the constant volatility σ^2. This leads to

$$\mathrm{d}X(t) = \left(\mu - \sigma^2(t)\delta\right)\,dt + \sigma(t)\,\mathrm{d}B(t), \tag{1.2}$$

where $(\sigma(t))_{t \geq 0}$ has the representation

$$\sigma^2(t) = \mathrm{e}^{-\lambda t}\sigma^2(0) + \mathrm{e}^{-\lambda t}\int_0^t \mathrm{e}^{\lambda s}L(\mathrm{d}s) \quad \text{for } t \geq 0, \tag{1.3}$$

with $\lambda > 0$ and $(L(t))_{t \geq 0}$ a positive Lévy process also known as subordinator. This model is capable to show most of the stylized facts, e.g. volatility jumps, clustering and heavy tails (cf. Fasen et al.

(2006)). It was used and studied in detail in finance, see e.g., Barndorff-Nielsen et al. (2002), Griffin and Steel (2006), Roberts et al. (2004) and extended to the multivariate case with a multivariate Ornstein-Uhlenbeck type process by Pigorsch and Stelzer (2009).

Let $(\mathbf{L}(t))_{t \geq 0}$ be a p-dimensional Lévy process and $\Sigma \in \mathbb{R}^{d \times p}$, $\mathbf{\Lambda} \in \mathbb{R}^{d \times d}$, $d, p \in \mathbb{N}$, where the eigenvalues of $\mathbf{\Lambda}$ have strictly positive real parts. Then a multivariate Ornstein-Uhlenbeck process $(\mathbf{Z}(t))_{t \geq 0}$ in $\mathbb{R}^d$ is defined as

$$\mathbf{Z}(t) = \mathrm{e}^{-\mathbf{\Lambda} t} \mathbf{Z}(0) + \int_0^t \mathrm{e}^{-\mathbf{\Lambda}(t-s)} \Sigma \, \mathbf{L}(\mathrm{d}s) \quad \text{for } t \geq 0. \tag{1.4}$$

Under our assumptions on $\mathbf{L}$, which will follow below, we can choose a stationary version of $(\mathbf{Z}(t))_{t \geq 0}$ since the eigenvalues of $\mathbf{\Lambda}$ have strictly positive real parts (see Sato and Yamazato (1984), Theorem 4.1). For given observations

$$\mathbb{Z}_n' = (\mathbf{Z}(1), \dots, \mathbf{Z}(n)) \in \mathbb{R}^{d \times n}, \quad \text{where we write} \quad \mathbb{Z}_{i,j}' = (\mathbf{Z}(i), \dots, \mathbf{Z}(j)) \in \mathbb{R}^{d \times (j-i+1)},$$

at a discrete-time grid, we study the properties of the least squares estimator for $\mathrm{e}^{-\mathbf{\Lambda}}$, since the estimation of $\mathbf{\Lambda}$ itself is not identifiable (cf. Lemma 3.1 below). An Ornstein-Uhlenbeck process observed at discrete time points is a multivariate AR(1) process with representation

$$\mathbf{Z}(k) = \mathrm{e}^{-\mathbf{\Lambda}} \mathbf{Z}(k-1) + \boldsymbol{\xi}_k \quad \text{where} \quad \boldsymbol{\xi}_k = \int_{k-1}^k \mathrm{e}^{-\mathbf{\Lambda}(k-s)} \Sigma \, \mathbf{L}(\mathrm{d}s) \quad \text{for } k \in \mathbb{N}. \tag{1.5}$$

The usual least squares estimator is then

$$\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}} = \mathbb{Z}_{2,n}' \mathbb{Z}_{1,n-1} (\mathbb{Z}_{1,n-1}' \mathbb{Z}_{1,n-1})^{-1}. \tag{1.6}$$

If the second moment of $\|\mathbf{L}(1)\|$ exists, then it is well known from the statistical inference of multivariate ARMA models that the estimator (1.6) is asymptotically normal and unbiased (see Hannan (1970), Chapter 6, or Proposition 3.2 below). In the finite variance case the least squares estimator is inefficient. More efficient estimators for the mean reversion parameter in one-dimensional Ornstein-Uhlenbeck models with finite variance are presented in Jongbloed et al. (2005), Brockwell et al. (2007), Taufer and Leonenko (2009) and Spiliopoulos (2010).

The main focus of this paper is to derive the asymptotic distribution of the least squares estimator (1.6), if $\|\mathbf{L}(1)\|$ has infinite variance. To be precise we study in detail the case where $\mathbf{L}(1)$ is multivariate regularly varying of an index less than 2, derive test statistics and compare our results to the case with finite second moments. The least squares estimator has the advantage that it is easy to implement and it performs much better in the heavy tailed model than in the finite variance case. In the limit the least squares estimator converges to a heavy tailed distribution which is a functional of stable random variables. However, one can not calculate the distribution analytically and it still depends on the unknown tail index of the underlying process. This makes it difficult to develop asymptotic approximations for the purpose of statistical inference.

Our result extends those of Davis and Resnick (1986) for the estimation of the autocorrelation function of a discrete-time AR(1) process from the one-dimensional case to the multivariate case.

3

To the best of our knowledge statistical inference for multivariate linear processes with infinite variance has not been well explored in the literature yet, apart from the work of Davis et al. (1985) and Meerschaert and Scheffler (2000, 2001) regarding only the convergence in probability of the normalized autocovariance function and the cross-correlation. The estimation of heavy tailed continuous-time AR(1) models was only considered in Hu and Long (2007, 2009) paying attention only at one-dimensional stable Ornstein-Uhlenbeck processes. In contrast to us, their observation grid is getting finer if the time scale increases or they observe the process on a hole time interval.

It is worth noting that in the heavy tailed case of one-dimensional linear models in discrete time M-estimators, in particular, the least absolute deviation estimator, can be more efficient than least squares estimators because outliers do not dominant as in the case of least squares estimators (cf. Calder and Davis (1998), Davis et al. (1992)). There the least squares estimators give too much influence to outliers. Thus, the extension to M-estimators in the multivariate setup of continuous-time linear models will be considered in some future work and compared to the least squares estimator of this paper.

The second part of this paper is devoted to the application of Ornstein-Uhlenbeck processes and its statistical inference in the context of co-integration. Co-integration is a well known phenomenon in economic time series as e.g., interest rates on assets of different maturities, prices of commodities in different parts of the world, income and expenditure by the local government, the value of sales and production costs of an industry, and spot and future prices in commodity markets (see Engle and Granger (1991), Engle and White (1999), Lütkepohl and Krätzig (2004)). This means that even though time series are non-stationary there exist linear combinations of them that render stationarity. Typical models for asset prices are exponential Lévy models (cf. Mandelbrot and Taylor (1967) and Eberlein (2009)). Although, exponential Lévy models are not able to capture stochastic volatility, they are the straightforward extension of the geometric Brownian motion in the Black-Scholes model modeling jumps and going away from the Gaussian assumption. The analytic form of the exponential Lévy model is simple, easier to handle and to fit to data than stochastic volatility models. In spread options and pairs trading, which is a popular investment strategy among hedge funds and investment banks, the concept is to find some pairs of assets which tend to move together in the long-run, i.e., they are co-integrated. There the logarithmic asset prices of two assets are modeled as

$$
\begin{aligned}
X(t) &= AY(t) + Z(t) && \text{for } t \geq 0, \\
Y(t) &= L(t) && \text{for } t \geq 0,
\end{aligned}
\tag{1.7}
$$

where $(Z(t))_{t \geq 0}$ is a stationary Ornstein-Uhlenbeck process, $(L(t))_{t \geq 0}$ is a Lévy process and $A \in \mathbb{R}$ is a constant (see Benth and Benth (2006), Duan and Pliska (2004), Ekström et al. (2009), Elliott et al. (2005)). Models of this type are also applied in electricity spot price dynamics (see Benth et al. (2008)). In the long-term the first asset behaves like a multiple of the second asset only in the short-term there are some deviations modeled by $(Z(t))_{t \geq 0}$. The Ornstein-Uhlenbeck parameter λ of $(Z(t))_{t \geq 0}$ reflects the speed of mean reversion to the equilibrium and hence, this parameter is important to know and to estimate for the optimal strategy in a pairs trade. It is also possible to allow some short-term deviations of $(Y(t))_{t \geq 0}$ from $(L(t))_{t \geq 0}$ by adding a noise term (cf. Fasen

4

(2010)). For the ease of notation we neglect this here. The linear regression model (1.7) is commonly used and basic in econometrics. In a two-step procedure we will estimate A and $\mathrm{e}^{-\lambda}$. In this paper we investigate a multiple version of (1.7) and its statistical inference where the noise is modeled by a multivariate Ornstein-Uhlenbeck model.

The paper is structured in the following way. Section 2 starts with preliminaries on multivariate regular variation of products and central limit results essential for the derivation of the asymptotic distribution of the least squares estimator. The results on multivariate regular variation of products are interesting in its own, however we use them to prove the central limit theorems. Section 3 contains the main results of the paper including the consistency and the asymptotic distribution of the least squares estimator but also test statistics. After all we apply the results to a multiple regression model in Section 4 and embed it into the context of co-integration. Finally, Section 5 contains the proofs.

We use the notation $\Longrightarrow$ for weak convergence, $\overset{\mathbb{P}}{\longrightarrow}$ for convergence in probability, and $\overset{\nu}{\Longrightarrow}$ for vague convergence. Let $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ be the compactification of $\mathbb{R}$ and let $\mathcal{B}(\cdot)$ be the Borel-σ-algebra. Further, we denote by $\mathbf{e}_{ij} \in \mathbb{R}^{d \times d}$ a matrix containing 1 at place (i,j) and 0 otherwise. Similarly, we define $\mathbf{e}_i \in \mathbb{R}^d$ as unit vector having 1 in row i and 0 otherwise. For a vector $\mathbf{x} \in \mathbb{R}^d$ we write $\mathbf{x}'$ for its transposed. 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 $\mathbb{R}^{m_1 \times m_1}$, $m_1, m_2 \in \mathbb{N}_0$. The symbol $\otimes$ denotes the Kronecker product and we use as norms the Euclidean norm $\|\cdot\|$ in $\mathbb{R}^d$ and the corresponding operator norm $\|\cdot\|$ for matrices. Then $\lambda := \|\mathbf{\Lambda}\|$ is the spectral norm of $\mathbf{\Lambda}$. The operator vec transforms a matrix $\mathbf{A} \in \mathbb{R}^{p \times d}$ into an $\mathbb{R}^{pd}$-vector by stacking the columns of $\mathbf{A}$. An $S_\alpha(1, \beta, 0)$-stable distribution will be an α-stable distribution with scale parameter 1, skewness parameter β and shift parameter 0 in the sense of Samorodnitsky and Taqqu (1994). Finally, let $[\mathbf{X}, \mathbf{Y}]_t = ([\mathbf{X}_i, \mathbf{Y}_j]_t)_{i=1,\ldots,m_1, j=1,\ldots,m_2}$ for the semimartinagles $\mathbf{X} = ((\mathbf{X}_1(t), \ldots, \mathbf{X}_{m_1}(t))')_{t \geq 0}$ in $\mathbb{R}^{m_1}$, $\mathbf{Y} = ((\mathbf{Y}_1(t), \ldots, \mathbf{Y}_{m_2}(t))')_{t \geq 0}$ in $\mathbb{R}^{m_2}$, $m_1, m_2 \in \mathbb{N}$, denote the quadratic covariation of $\mathbf{X}$ and $\mathbf{Y}$.

2 Preliminaries

2.1 Regular variation

In this paper we investigate the Ornstein-Uhlenbeck model (1.4) where $\mathbf{L}(1)$ is a multivariate regularly varying random vector. We recall the definition.

Definition 2.1

(a) *A sequence $(c_n)_{n \in \mathbb{N}}$ of positive constants (a measurable function $f : (0, \infty) \to (0, \infty)$, respectively) is called regularly varying of index $-\alpha$, $\alpha \in \mathbb{R}$, if $\lim_{n \to \infty} c_{\lfloor nu \rfloor}/c_n = u^{-\alpha}$ ($\lim_{x \to \infty} f(xu)/f(x) = u^{-\alpha}$, respectively) for any $u > 0$, where $\lfloor nu \rfloor = \sup\{k \in \mathbb{N} : k \leq nu\}$. In that case we also write $(c_n)_{n \in \mathbb{N}} \in \mathcal{R}_{-\alpha}$ ($f \in \mathcal{R}_{-\alpha}$, respectively).*

(b) *A random matrix $\mathbf{U} \in \mathbb{R}^{p \times d}$ is multivariate regularly varying with index $-\alpha < 0$ if and only if there exists a non-zero Radon measure μ on $\overline{\mathbb{R}}^{p \times d} \setminus \{\mathbf{0}_{p \times d}\}$ with $\mu(\overline{\mathbb{R}}^{p \times d} \setminus \mathbb{R}^{p \times d}) = 0$ and a sequence*

$(a_n)_{n\in\mathbb{N}}$ of positive numbers increasing to ∞ such that

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

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

The sequence $(a_n)_{n\in\mathbb{N}}$ in (b) is in $\mathcal{R}_{1/\alpha}$ as well. If the representation of the limit measure μ or the norming sequence $(a_n)_{n\in\mathbb{N}}$ does not matter we also write $\mathcal{R}_{-\alpha}(a_n)$ and $\mathcal{R}_{-\alpha}$, respectively. For further information regarding multivariate regular variation of random vectors we refer to Resnick (2007). However, we can transfer the results to random matrices in $\mathbb{R}^{p\times d}$ by rewriting the random matrix as a random vector in $\mathbb{R}^{pd}$ with the vec operator.

For the proofs of our central limit results in Section 2.2 we have to understand multivariate regular variation of products of random vectors, where contrary to the common approaches both factors are multivariate regularly varying of the same index. Since there is no dominant factor in that model, the calculations are more involved. The one-dimensional version can be found in Cline (1986), who applies in his proof that for a positive random variable X we have $X \in \mathcal{R}_{-\alpha}$ iff $\lim_{x\to\infty} \mathbb{P}(X > e^{x-y})/\mathbb{P}(X > e^x) = e^{\alpha y}$ for any $y \in \mathbb{R}$. Unfortunately this can not be used in the multivariate setup.

Lemma 2.2

Let $\boldsymbol{\zeta}^{(1)} = (\zeta_1^{(1)}, \ldots, \zeta_p^{(1)})', \boldsymbol{\zeta}^{(2)} = (\zeta_1^{(2)}, \ldots, \zeta_p^{(2)})' \in \mathcal{R}_{-\alpha}(a_n, \mu)$ be iid random vectors in $\mathbb{R}^p$ with independent components such that

$$\lim_{x\to\infty} \frac{\mathbb{P}(\zeta_i^{(1)} > x)}{\mathbb{P}(\|\boldsymbol{\zeta}^{(1)}\| > x)} = p_i < \infty \quad \text{and} \quad \lim_{x\to\infty} \frac{\mathbb{P}(-\zeta_i^{(1)} > x)}{\mathbb{P}(\|\boldsymbol{\zeta}^{(1)}\| > x)} = q_i < \infty, \quad i = 1, \ldots, p,$$

and $\mathbb{E}\|\boldsymbol{\zeta}^{(1)}\|^\alpha = \infty$. Let $\widetilde{a}_n > 0$ be an increasing sequence of constants such that

$$0 < \lim_{n\to\infty} n\mathbb{P}(\|\boldsymbol{\zeta}^{(1)}\|\|\boldsymbol{\zeta}^{(2)}\| > \widetilde{a}_n) = C < \infty.$$

Then $\boldsymbol{\zeta}^{(1)}\boldsymbol{\zeta}^{(2)'} \in \mathcal{R}_{-\alpha}(\widetilde{a}_n, C\widetilde{\mu}(\cdot))$, where

$$\widetilde{\mu}(\cdot) = \sum_{i,j=1}^p (\widetilde{p}_{ij} + \widetilde{q}_{ij})\mathbb{E}\left(\mu_\alpha(\{x \in \mathbb{R}_+ : xR_{ij}\mathbf{e}_{ij} \in \cdot\})\right),$$

$\mu_\alpha(\mathrm{d}x) = \alpha x^{-\alpha-1}\mathbb{1}_{(0,\infty)}(x)\,\mathrm{d}x$ and $(R_{ij})_{i,j=1,\ldots,p}$ is an independent sequence of Bernoulli random variables with

$$\mathbb{P}(R_{ij} = 1) = \frac{\widetilde{p}_{ij}}{\widetilde{p}_{ij} + \widetilde{q}_{ij}} \quad \text{and} \quad \mathbb{P}(R_{ij} = -1) = \frac{\widetilde{q}_{ij}}{\widetilde{p}_{ij} + \widetilde{q}_{ij}},$$

$\widetilde{p}_{ij} = p_i p_j + q_i q_j$ and $\widetilde{q}_{ij} = p_i q_j + q_i p_j$.

With the above notation we use either of the following assumptions to prove the consistency of the least squares estimator (1.6).

Assumption A

Let $\mathbf{L}(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$, and let $(\mathbf{L}(t))_{t\geq 0} = ((L_1(t), \ldots, L_p(t))')_{t\geq 0}$ have independent components, where $\mathbb{E}\|\mathbf{L}(1)\|^\alpha = \infty$,

$$\lim_{x\to\infty} \frac{\mathbb{P}(L_i(1) > x)}{\mathbb{P}(\|\mathbf{L}(1)\| > x)} = p_i < \infty \quad and \quad \lim_{x\to\infty} \frac{\mathbb{P}(-L_i(1) > x)}{\mathbb{P}(\|\mathbf{L}(1)\| > x)} = q_i < \infty, \quad i = 1, \ldots, p.$$

Finally, let $a_n, \widetilde{a}_n > 0$ be increasing sequences of constants such that

$$\lim_{n\to\infty} n\mathbb{P}(\|\mathbf{L}(1)\| > a_n) = 1 \quad and \quad \lim_{n\to\infty} n\mathbb{P}(\|\mathbf{L}(1)\|\|\mathbf{L}(2) - \mathbf{L}(1)\| > \widetilde{a}_n) = 1.$$

Note that by Cline (1986), Proposition II, $(\widetilde{a}_n)_{n\in\mathbb{N}} \in \mathcal{R}_{1/\alpha}$ and by (3.5) in Davis and Resnick (1986), $\lim_{n\to\infty} \widetilde{a}_n/a_n = \infty$.

The next alternative and stronger Assumption B results in simpler limit results.

Assumption B

Let $(L(t))_{t\geq 0}$ be a one-dimensional Lévy process. Suppose that the components $(L_1(t))_{t\geq 0}, \ldots, (L_p(t))_{t\geq 0}$ of the p-dimensional Lévy process $(\mathbf{L}(t))_{t\geq 0} = ((L_1(t), \ldots, L_p(t))')_{t\geq 0}$ are independent Lévy processes with $(L_i(t))_{t\geq 0} \stackrel{d}{=} (L(t))_{t\geq 0}$ for $i = 1, \ldots, p$. Furthermore, assume that

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

and

$$\mathbb{P}(|L(1)| > u) = b(\log u)u^{-\alpha},$$

where $0 < \alpha < 2$ and $b \in \mathcal{R}_\gamma$ with $\gamma > -1$. If $1 < \alpha < 2$ then $\mathbb{E}(L(1)) = 0$. Let $a_n, \widetilde{a}_n > 0$ be increasing sequences of constants such that as $n \to \infty$,

$$a_n \sim (K_\alpha b(\log n)n)^{1/\alpha} \quad and \quad \widetilde{a}_n \sim (\alpha K_\alpha C_\gamma b^2(\log n)\log(n)n)^{1/\alpha}$$

with $K_\alpha = \frac{\Gamma(2-\alpha)}{\alpha-1}|\cos(\frac{\pi\alpha}{2})|$ and $C_\gamma = \Gamma(1+\gamma)^2/\Gamma(2+2\gamma)$. Finally, suppose that $\mathbf{\Lambda} = diag(\lambda_1, \ldots, \lambda_d)$ and $\Sigma = \mathbf{I}_{d\times d}$.

Under Assumption B the components of $(\mathbf{Z}(t))_{t\geq 0}$ driven by $(\mathbf{L}(t))_{t\geq 0}$ are independent one-dimensional Ornstein-Uhlenbeck processes driven by the Lévy processes $(L_i(t))_{t\geq 0}$ and with Ornstein-Uhlenbeck parameter $\lambda_i > 0$. Then the estimation can also be reduced to the estimation of λ_i in the one-dimensional model instead of looking at the multivariate model. However, in particular in the context of the co-integrated model in Section 4 it makes sense to investigate that special model of Assumption B.

Example 2.3

Let $(L(t))_{t\geq 0}$ be an α-stable Lévy motion. Then there exists a constant $C > 0$ such that

$$\mathbb{P}(|L(1)| > u) \sim Cu^{-\alpha} \quad as \ u \to \infty.$$

In Assumption B we can take $(a_n)_{n\in\mathbb{N}}$ and $(\widetilde{a}_n)_{n\in\mathbb{N}}$, respectively as

$$a_n = (K_\alpha Cn)^{1/\alpha} \quad and \quad \widetilde{a}_n = (\alpha K_\alpha C^2 \log(n)n)^{1/\alpha} \quad for \ n \in \mathbb{N}. \qquad \square$$

The next theorem extends Lemma 2.2 to the context of an Ornstein-Uhlenbeck model.

Theorem 2.4

Let $(\mathbf{Z}(t))_{t\geq 0}$ be given as in (1.4) and $(\boldsymbol{\xi}_k)_{k\in\mathbb{N}}$ as in (1.5). Furthermore, suppose that Assumption A holds. Define

$$\mu_{\boldsymbol{\xi}\boldsymbol{\xi}'}(\cdot) = \sum_{i,j=1}^{p} (\widetilde{p}_{ij} + \widetilde{q}_{ij})\mathbb{E}\left(\mu_\alpha(\{x \in \mathbb{R}_+ : xR_{ij}\mathrm{e}^{-\boldsymbol{\Lambda}U^{(1)}}\Sigma\mathbf{e}_{ij}\Sigma'\mathrm{e}^{-\boldsymbol{\Lambda}'U^{(2)}} \in \cdot\})\right), \tag{2.1}$$

where $\mu_\alpha(\mathrm{d}x) = \alpha x^{-\alpha-1}\mathbb{1}_{(0,\infty)}(x)\,\mathrm{d}x$ and $U^{(1)}, U^{(2)}$ are iid uniformly distributed random variables on $(0,1)$ independent of $(R_{ij})_{i,j=1,\ldots,p}$, an independent sequence of Bernoulli random variables with

$$\mathbb{P}(R_{ij} = 1) = \frac{\widetilde{p}_{ij}}{\widetilde{p}_{ij} + \widetilde{q}_{ij}} \quad \text{and} \quad \mathbb{P}(R_{ij} = -1) = \frac{\widetilde{q}_{ij}}{\widetilde{p}_{ij} + \widetilde{q}_{ij}},$$

$\widetilde{p}_{ij} = p_i p_j + q_i q_j$ and $\widetilde{q}_{ij} = p_i q_j + q_i p_j$. Then

$$\boldsymbol{\xi}_1\boldsymbol{\xi}_2' \in \mathcal{R}_{-\alpha}(\widetilde{a}_n, \mu_{\boldsymbol{\xi}\boldsymbol{\xi}'}), \tag{2.2}$$

and

$$(\boldsymbol{\xi}_{m+1}\boldsymbol{\xi}_m', \ldots, \boldsymbol{\xi}_{m+1}\boldsymbol{\xi}_1') \in \mathcal{R}_{-\alpha}\left(\widetilde{a}_n, \mu_{\boldsymbol{\xi}\boldsymbol{\xi}_m'}\right) \quad \text{with} \quad \mu_{\boldsymbol{\xi}\boldsymbol{\xi}_m'}(\cdot) = \sum_{k=1}^{m}\mu_{\boldsymbol{\xi}\boldsymbol{\xi}'} \circ \mathrm{pr}_k^{(d,m)}(E_k^{(d,m)} \cap \cdot), \tag{2.3}$$

where $\mathrm{pr}_k^{(d,m)} : \mathbb{R}^{d\times dm} \to \mathbb{R}^{d\times d}$ is the projection $\mathbf{A} = (A_{l,s})_{l=1,\ldots,d,s=1,\ldots,dm} \mapsto (A_{u,(k-1)d+v})_{u,v=1,\ldots,d}$ and $E_k^{(d,m)} = \{\mathbf{A} \in \mathbb{R}^{d\times dm} : A_{l,s} = 0 \text{ for } (l,s) \notin \{(u,(k-1)d+v) : u,v = 1,\ldots,d\}\}$.

2.2 Central limit results

We use the following special case of Fasen (2010), Proposition 2.1 (cf. Davis et al. (1985)) for the proof of the consistency of our estimator (1.6).

Proposition 2.5

Let $(\mathbf{Z}(t))_{t\geq 0}$ be given as in (1.4) and $(\boldsymbol{\xi}_k)_{k\in\mathbb{N}}$ as in (1.5). Furthermore, let $0 < \alpha < 2$. Suppose that $\mathbf{L}(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$ and $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_p$ if $1 < \alpha < 2$. Define for $n \in \mathbb{N}$:

$$\mathbf{S}_{1,n} = a_n^{-1}\sum_{k=1}^{n}\boldsymbol{\xi}_k, \qquad \mathbf{S}_{2,n} = a_n^{-2}\sum_{k=1}^{n}\boldsymbol{\xi}_k\boldsymbol{\xi}_k',$$

$$\mathbf{S}_{3,n} = a_n^{-2}\sum_{k=1}^{n}\mathbf{Z}(k)\mathbf{Z}(k)', \qquad \mathbf{S}_{4,n} = a_n^{-2}\sum_{k=1}^{n}\mathbf{Z}(k+1)\mathbf{Z}(k)'.$$

Let $(\mathbf{S}_1(t))_{t\geq 0}$ be an α-stable Lévy process in $\mathbb{R}^d$ with Lévy measure

$$\mu_{\boldsymbol{\xi}}(B) = \mathbb{E}\left(\mu\left(\{\mathbf{x} \in \overline{\mathbb{R}}^p\backslash\{\mathbf{0}_p\} : \mathrm{e}^{-\boldsymbol{\Lambda}U}\Sigma\mathbf{x} \in B\}\right)\right) \quad \text{for } B \in \mathcal{B}(\mathbb{R}^d\backslash\{\mathbf{0}_d\}), \tag{2.4}$$

where U is a uniformly distributed random variable on $(0,1)$, and if $1 < \alpha < 2$ then $\mathbb{E}(\mathbf{S}_1(t)) = \mathbf{0}_d$. Finally, define the random matrices

$$\mathbf{S}_1 := \mathbf{S}_1(1), \qquad \mathbf{S}_2 := [\mathbf{S}_1, \mathbf{S}_1]_1, \qquad \mathbf{S}_3 := \sum_{i=0}^{\infty}\mathrm{e}^{-\boldsymbol{\Lambda}i}[\mathbf{S}_1, \mathbf{S}_1]_1\mathrm{e}^{-\boldsymbol{\Lambda}'i}, \qquad \mathbf{S}_4 := \mathrm{e}^{-\boldsymbol{\Lambda}}\mathbf{S}_3.$$

Then as $n \to \infty$,

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

However, for the asymptotic behavior of our estimator we require a stronger result.

Theorem 2.6

Let $(\mathbf{Z}(t))_{t \geq 0}$ be given as in (1.4) and $(\boldsymbol{\xi}_k)_{k \in \mathbb{N}}$ as in (1.5). Furthermore, let $0 < \alpha < 2$. Suppose that $\mathbf{L}(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$, where $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_p$ if $1 < \alpha < 2$, and Assumption A holds. Define

$$\mathbf{S}_{5,n} := \widetilde{a}_n^{-1} \sum_{j=0}^{\infty} \sum_{k=1}^{n-1} \boldsymbol{\xi}_{k+1} \boldsymbol{\xi}'_{k-j} \mathrm{e}^{-\mathbf{\Lambda}'j} \quad \text{for } n \in \mathbb{N}, \qquad \text{and} \qquad \mathbf{S}_5 := \sum_{j=0}^{\infty} \mathbf{S}_j^* \mathrm{e}^{-\mathbf{\Lambda}'j},$$

where $(\mathbf{S}_j^*)_{j \in \mathbb{N}_0}$ are iid $\mathbb{R}^{d \times d}$-valued α-stable Lévy processes independent of $(\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, \mathbf{S}_4)$ with Lévy measure $\mu_{\boldsymbol{\xi}\boldsymbol{\xi}'}$ as given in (2.1) and $\mathbb{E}(\mathbf{S}_1^*) = \mathbf{0}_{d \times d}$ if $1 < \alpha < 2$. Then as $n \to \infty$,

$$(\mathbf{S}_{1,n}, \mathbf{S}_{2,n}, \mathbf{S}_{3,n}, \mathbf{S}_{4,n}, \mathbf{S}_{5,n}) \Longrightarrow (\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, \mathbf{S}_4, \mathbf{S}_5).$$

We conclude the Section with a corollary which gives under the stronger Assumption B simple representations of $\mathbf{S}_i$ for $i = 1, \ldots, 5$.

Corollary 2.7

Suppose Assumption B holds. Then as $n \to \infty$,

$$(\mathbf{S}_{1,n}, \mathbf{S}_{2,n}, \mathbf{S}_{3,n}, \mathbf{S}_{4,n}, \mathbf{S}_{5,n}) \Longrightarrow (\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, \mathbf{S}_4, \mathbf{S}_5),$$

where

$$\mathbf{S}_1 = E_{\mathbf{\Lambda},\alpha} D_{\mathbf{\Lambda},\alpha} \mathbf{S}_1^*(1), \qquad \mathbf{S}_2 = E_{\mathbf{\Lambda},\alpha}^2 D_{\mathbf{\Lambda},\alpha}^2 [\mathbf{S}_1^*, \mathbf{S}_1^*]_1, \qquad \mathbf{S}_3 = E_{\mathbf{\Lambda},2}^{-2} E_{\mathbf{\Lambda},\alpha}^2 D_{\mathbf{\Lambda},\alpha}^2 [\mathbf{S}_1^*, \mathbf{S}_1^*]_1,$$

$$\mathbf{S}_4 = \mathrm{e}^{-\mathbf{\Lambda}} \mathbf{S}_3, \qquad \qquad \mathbf{S}_5 = E_{\mathbf{\Lambda},\alpha} D_{\mathbf{\Lambda},\alpha} \mathbf{S}_2^*(1) D_{\mathbf{\Lambda},\alpha},$$

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

$$E_{\mathbf{\Lambda},\alpha} = \mathrm{diag}((1 - \mathrm{e}^{-\alpha \lambda_1})^{1/\alpha}, \ldots, (1 - \mathrm{e}^{-\alpha \lambda_d})^{1/\alpha}) \quad \text{and} \quad D_{\mathbf{\Lambda},\alpha} = \mathrm{diag}((\alpha \lambda_1)^{-1/\alpha}, \ldots, (\alpha \lambda_d)^{-1/\alpha}).$$

3 Least squares estimator

3.1 Statistical estimation

In this paper we estimate $\mathrm{e}^{-\mathbf{\Lambda}}$ instead of $\mathbf{\Lambda}$, since classical results of algebra going back to Cluver (1966) show that under general assumptions the estimation of $\mathbf{\Lambda}$ by the discrete-time model is not identifiable.

Lemma 3.1

Let $\mathbf{\Lambda} \in \mathbb{R}^{d \times d}$. Then the following are equivalent:

(a) *$x = \mathbf{\Lambda}$ is the unique solution to $\exp(-x) = \exp(-\mathbf{\Lambda})$ in $\mathbb{R}^{d \times d}$.*

(b) All eigenvalues of $\mathbf{\Lambda}$ are real and no elementary divisor (Jordan block) of $\mathbf{\Lambda}$ occurs more than once.

For completeness we start with the estimation of $\mathrm{e}^{-\mathbf{\Lambda}}$ in the case where second moments exists, which follows directly from Hannan (1970), Chapter 6 or Lütkepohl (2007), Section 3.3.

Proposition 3.2

Let $\mathbb{E}\|\mathbf{L}(1)\|^2 < \infty$ and $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_p$. Define

$$\Sigma_{\mathbf{Z}} = \int_0^\infty \mathrm{e}^{-\mathbf{\Lambda}s}\Sigma\mathbb{E}(\mathbf{L}(1)\mathbf{L}(1)')\Sigma'\mathrm{e}^{-\mathbf{\Lambda}'s}\,\mathrm{d}s \quad and \quad \Sigma_{\boldsymbol{\xi}} = \int_0^1 \mathrm{e}^{-\mathbf{\Lambda}s}\Sigma\mathbb{E}(\mathbf{L}(1)\mathbf{L}(1)')\Sigma'\mathrm{e}^{-\mathbf{\Lambda}'s}\,\mathrm{d}s. \quad (3.1)$$

Suppose $\Sigma_{\mathbf{Z}}$ is invertible. Then as $n \to \infty$,

$$\sqrt{n}\left(vec(\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}}) - vec(\mathrm{e}^{-\mathbf{\Lambda}})\right) \Longrightarrow \mathcal{N}(\mathbf{0}_{d^2}, \Sigma_{\mathbf{Z}}^{-1} \otimes \Sigma_{\boldsymbol{\xi}}).$$

In particular, $\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}} \xrightarrow{\mathbb{P}} \mathrm{e}^{-\mathbf{\Lambda}}$ as $n \to \infty$.

Remark 3.3

Particularly, $\Sigma_{\mathbf{Z}}$ is invertible if and only if $\Sigma_{\boldsymbol{\xi}}$ is invertible. If $\Sigma\mathbb{E}(\mathbf{L}(1)\mathbf{L}(1)')\Sigma'$ is invertible then also $\Sigma_{\boldsymbol{\xi}}$ and $\Sigma_{\mathbf{Z}}$ are invertible. A sufficient condition for both is that Σ has full rank and $\mathbf{L}$ has uncorrelated components. $\qquad\square$

Under very general assumptions we obtain also the consistency of the least squares estimator, if the second moment does not exist.

Theorem 3.4

Let $0 < \alpha < 2$ and $\mathbf{L}(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$, where $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_p$ if $1 < \alpha < 2$. Suppose $\mathbb{P}(\det(\mathbf{S}_3) = 0) = 0$. Then the estimator $\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}}$ as given in (1.6) is consistent, i.e. $\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}} \xrightarrow{\mathbb{P}} \mathrm{e}^{-\mathbf{\Lambda}}$ as $n \to \infty$.

This result can be obtained by a straightforward application of Davis et al. (1985), Theorem 3.1, in the bivariate case and under slightly different assumptions by Meerschaert and Scheffler (2000) for the multivariate case. However, it follows also directly from Proposition 2.5, since

$$\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}} = \mathbb{Z}_{2,n}'\mathbb{Z}_{1,n-1}(\mathbb{Z}_{1,n-1}'\mathbb{Z}_{1,n-1})^{-1} = \mathbf{S}_{4,n}\mathbf{S}_{3,n}^{-1} \Longrightarrow \mathrm{e}^{-\mathbf{\Lambda}}\mathbf{S}_3\mathbf{S}_3^{-1} = \mathrm{e}^{-\mathbf{\Lambda}} \quad as \ n \to \infty.$$

Remark 3.5

The claim $\mathbb{P}(\det(\mathbf{S}_3) = 0) = 0$ holds if and only if $\mathbb{P}(\det(\mathbf{S}_2) = 0) = 0$. If Assumption A is satisfied, a sufficient condition is $p_i + q_i > 0$ for $i = 1, \ldots, p$ and Σ has full rank. $\qquad\square$

For the asymptotic behavior of the least square estimator we require the slightly stronger Assumption A.

Theorem 3.6

Let $0 < \alpha < 2$. Suppose that $\mathbf{L}(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$ and $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_p$ if $1 < \alpha < 2$.

(a) Let Assumption A hold and suppose $\mathbb{P}(\det(\mathbf{S}_3) = 0) = 0$. Then as $n \to \infty$,

$$\frac{a_n^2}{\widetilde{a}_n}\left(\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}} - \mathrm{e}^{-\mathbf{\Lambda}}\right) \Longrightarrow \mathbf{S}_5\mathbf{S}_3^{-1} =: \mathbf{G}.$$

(b) *Let Assumption B hold. Then as $n \to \infty$,*

$$\frac{a_n^2}{\widetilde{a}_n} D_{\mathbf{\Lambda},\alpha}^{-1} E_{\mathbf{\Lambda},\alpha}^{-1} \left(\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}} - \mathrm{e}^{-\mathbf{\Lambda}} \right) D_{\mathbf{\Lambda},\alpha} E_{\mathbf{\Lambda},\alpha}^2 E_{\mathbf{\Lambda},2}^{-2} \Longrightarrow \mathbf{S}_2^*(1)([\mathbf{S}_1^*, \mathbf{S}_1^*]_1)^{-1} =: \mathbf{G}^*.$$

The Theorem generalizes the limit results of Davis and Resnick (1986) to the multivariate case. Note, that *(b)* is in particularly useful for the practical simulation of confidence intervals as we see below.

Remark 3.7
(a) Since the limit distribution $\mathbf{G}$ depends on $\mathbf{S}_3$ and on $\mathbf{S}_5$ it depends in particular on $\mathbf{\Lambda}$, which is plugged in to the Lévy measures of $\mathbf{S}_3$ and $\mathbf{S}_5$ which we see nicely in *(b)*.
(b) For statistical applications we keep in mind that the unknown parameter α is going into $a_n^2 \widetilde{a}_n^{-1}$, which is regularly varying sequence of index $1/\alpha$. Furthermore, the precise form of the slowly varying function in a_n and in $\widetilde{a}_n$, respectively is not known as well. The same problem arises in the t- and the Wald-statistic below and makes the statistical inference for such models difficult. $\qquad \square$

Next we compare the result in the finite variance and the infinite variance setting.

Remark 3.8
(a) Since $(a_n^2 \widetilde{a}_n^{-1})_{n \in \mathbb{N}} \in \mathcal{R}_{1/\alpha}$ and $1/\alpha > 1/2$, the convergence rate $a_n^2/\widetilde{a}_n$ of $\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}}$ in the heavy tailed case (Theorem 3.6) is faster than the usual convergence rate $n^{1/2}$ in the case where second moments exists (Proposition 3.2).
(b) For $\alpha = 2$ the random matrix $\mathbf{S}_2^*(1)([\mathbf{S}_1^*, \mathbf{S}_1^*]_1)^{-1}$ has iid standard normal distributed components. Hence, the finite variance case of Proposition 3.2 has for $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$ and $\Sigma \mathbb{E}(\mathbf{L}(1)\mathbf{L}(1)')\Sigma' = \mathbf{I}_{d \times d}$ the same representation as the infinite variance case of Theorem 3.6. $\square$

Finally, we comment what happens if we have a different observation grid.

Remark 3.9
Suppose we observe $(\mathbf{Z}(t))_{t \geq 0}$ in Theorem 3.6 *(b)* at the discrete time-grid kh for $k = 1, \ldots, n$, where $h > 0$ is fixed. Then we have the observations

$$\mathbb{Z}'_{h,n} = (\mathbf{Z}(h), \ldots, \mathbf{Z}(nh)) \in \mathbb{R}^{d \times n} \text{ where we write } \mathbb{Z}'_{h,i,j} = (\mathbf{Z}(ih), \ldots, \mathbf{Z}(jh)) \in \mathbb{R}^{d \times (j-i+1)},$$

and define the least squares estimator

$$\widehat{\mathrm{e}^{-\mathbf{\Lambda}_{h,n}}} = \mathbb{Z}'_{h,2,n} \mathbb{Z}_{h,1,n-1} (\mathbb{Z}'_{h,1,n-1} \mathbb{Z}_{h,1,n-1})^{-1}.$$

We obtain as $n \to \infty$,

$$\frac{a_n^2}{\widetilde{a}_n} D_{h\mathbf{\Lambda},\alpha}^{-1} E_{h\mathbf{\Lambda},\alpha}^{-1} \left(\widehat{\mathrm{e}^{-\mathbf{\Lambda}_{h,n}}} - \mathrm{e}^{-h\mathbf{\Lambda}} \right) D_{h\mathbf{\Lambda},\alpha} E_{h\mathbf{\Lambda},\alpha}^2 E_{h\mathbf{\Lambda},2}^{-2} \Longrightarrow \mathbf{G}^*,$$

which means for the one-dimensional case $\mathbf{\Lambda} = \lambda$ that as $n \to \infty$,

$$\frac{a_n^2}{\widetilde{a}_n} (1 - \mathrm{e}^{-\alpha h \lambda})^{1/\alpha} (1 - \mathrm{e}^{-2h\lambda})^{-1} \left(\widehat{\mathrm{e}^{-\lambda_{h,n}}} - \mathrm{e}^{-h\lambda} \right) \Longrightarrow \frac{S_2^*}{S_1^*},$$

11

where S_2^* is an $S_\alpha(1,0,0)$-stable random variable independent of the positive $S_{\alpha/2}(1,1,0)$-stable random variable S_1^*. On the other hand,

$$\lim_{h \to 0} \frac{(1 - \mathrm{e}^{-\alpha h\lambda})^{1/\alpha}(1 - \mathrm{e}^{-2h\lambda})^{-1}}{(\alpha h\lambda)^{1/\alpha}(2h\lambda)^{-1}} = 1$$

suggests that for a observation scheme with grid distance h_n, which is getting finer as n increases, i.e., $h_n \downarrow 0$ and $nh_n \longrightarrow \infty$ as $n \to \infty$, we have as $n \to \infty$,

$$\frac{a_n^2}{\widetilde{a}_n}(\alpha\lambda h_n)^{\frac{1}{\alpha}}(2\lambda h_n)^{-1}\left(\widehat{\mathrm{e}^{-\lambda_{h_n,n}}} - \mathrm{e}^{-\lambda h_n}\right) \Longrightarrow \frac{S_2^*}{S_1^*}.$$

If additionally $\lim_{n \to \infty} h_n^{\frac{\alpha+1}{\alpha}} a_n^2 \widetilde{a}_n^{-1} = 0$ then we will obtain as $n \to \infty$,

$$h_n^{\frac{1}{\alpha}}\frac{a_n^2}{\widetilde{a}_n}(\alpha\lambda)^{\frac{1}{\alpha}}(2\lambda)^{-1}\left(\frac{1 - \widehat{\mathrm{e}^{-\lambda_{h_n,n}}}}{h_n} - \lambda\right) \Longrightarrow \frac{S_2^*}{S_1^*}.$$

This was shown in Hu and Long (2009) for the least squares estimator of an α-stable Ornstein-Uhlenbeck model. In some future work we will show that this result holds for general regularly varying Ornstein-Uhlenbeck models. $\qquad\square$

One can use the above results to derive confidence intervals for the components of $\mathrm{e}^{-\mathbf{\Lambda}}$. In the definition of $\mathbf{G}^*$ in Theorem 3.6 the random matrix $\mathbf{S}_2^*(1)$ has iid components and $[\mathbf{S}_1^*, \mathbf{S}_1^*]_1$ is a diagonal matrix whose diagonal consists of iid components. Thus, $\mathbf{G}^*$ is a random matrix with identically distributed components which do not depend on d, i.e. every component has the distribution of S_2^*/S_1^* of Remark 3.9. Furthermore, the distribution of S_2^*/S_1^* is heavy tailed and becomes heavier for smaller α. Let $x_p(\alpha)$ denote the $1 - \frac{p}{2}$ quantile of S_2^*/S_1^*. An $1 - p$-confidence interval of $(\mathrm{e}^{-\mathbf{\Lambda}})_{ik}$ (the (i,k)-component of $\mathrm{e}^{-\mathbf{\Lambda}}$) has then the representation

$$\left[(\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}})_{ik} \pm x_p(\alpha)\frac{\widetilde{a}_n}{a_n^2}\left(\frac{\lambda_k}{\lambda_i}\right)^{1/\alpha}(1 - \mathrm{e}^{-\alpha\lambda_i})^{1/\alpha}(1 - \mathrm{e}^{-\alpha\lambda_k})^{-2/\alpha}(1 - \mathrm{e}^{-2\lambda_k})\right].$$

In practise the distribution of S_2^*/S_1^* can not be computed theoretically but can be derived by simulation or numerical approximation, since there is no analytic representation of the density. In Table 1 we have simulated the $1 - \frac{p}{2}$-quantiles of S_2^*/S_1^* by 100,000 Monte Carlo simulations using the toolbox STABLE of Robust Analysis Inc.

	$\alpha = 1.2$	$\alpha = 1.4$	$\alpha = 1.5$	$\alpha = 1.6$	$\alpha = 1.8$
$p = 0.1$	4.9	3.1	2.5	2.1	1.5
$p = 0.05$	9.0	4.9	3.8	3.0	2.0
$p = 0.025$	16.0	7.8	5.6	4.2	2.5
$p = 0.01$	33.2	14.3	9.6	6.9	3.5

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

3.2 Asymptotic tests

In this section we want to test the significance of subvectors and components of $e^{-\boldsymbol{\Lambda}}$. Usual tests on the linear restriction $\mathbf{R}\operatorname{vec}(e^{-\boldsymbol{\Lambda}}) = \mathbf{r}$ for some $\mathbf{R} \in \mathbb{R}^{q \times d^2}$, $\mathbf{r} \in \mathbb{R}^q$ are the t- and the Wald-statistic. The idea in both tests is that if the restriction is true then the difference $\mathbf{R}\operatorname{vec}(\widehat{e^{-\boldsymbol{\Lambda}_n}}) - \mathbf{r}$ is close to $\mathbf{0}_q$.

A straightforward conclusion of Proposition 3.2 and Hannan (1970), Chapter 4, Theorem 6, is the next proposition for the finite variance case, where in *(a)* we investigate the t- and in *(b)* the Wald statistic.

Proposition 3.10
Let $\mathbb{E}\|\mathbf{L}(1)\|^2 < \infty$ and $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_p$. Define

$$\widehat{\boldsymbol{\Omega}}_n := n^{-1}(\mathbb{Z}'_{2,n} - \widehat{e^{-\boldsymbol{\Lambda}_n}}\mathbb{Z}'_{1,n-1})(\mathbb{Z}'_{2,n} - \widehat{e^{-\boldsymbol{\Lambda}_n}}\mathbb{Z}'_{1,n-1})'. \tag{3.2}$$

Suppose $\Sigma_{\mathbf{Z}}$ as given in (3.1) is invertible.

(a) Then as $n \to \infty$,

$$((\mathbb{Z}'_{1,n-1}\mathbb{Z}_{1,n-1})^{1/2} \otimes \widehat{\boldsymbol{\Omega}}_n^{-1/2})\left(\operatorname{vec}(\widehat{e^{-\boldsymbol{\Lambda}_n}}) - \operatorname{vec}(e^{-\boldsymbol{\Lambda}})\right) \Longrightarrow \mathcal{N}(\mathbf{0}_{d^2}, \mathbf{I}_{d^2 \times d^2}).$$

Let $\mathbf{R} \in \mathbb{R}^{1 \times d^2}$ and $r \in \mathbb{R}$. If the null hypothesis H_0: $\mathbf{R}\operatorname{vec}(e^{-\boldsymbol{\Lambda}}) = r$ is true, then as $n \to \infty$,

$$(\mathbf{R}((\mathbb{Z}'_{1,n-1}\mathbb{Z}_{1,n-1}) \otimes \widehat{\boldsymbol{\Omega}}_n^{-1})\mathbf{R}')^{1/2}\left(\mathbf{R}\operatorname{vec}(\widehat{e^{-\boldsymbol{\Lambda}_n}}) - r\right) \Longrightarrow \mathcal{N}(0,1).$$

(b) Let $\mathbf{R} \in \mathbb{R}^{q \times d^2}$, $\mathbf{r} \in \mathbb{R}^q$ and $\operatorname{rank}(\mathbf{R}) = q$. If the null hypothesis H_0: $\mathbf{R}\operatorname{vec}(e^{-\boldsymbol{\Lambda}}) = \mathbf{r}$ is true, then as $n \to \infty$,

$$(\mathbf{R}\operatorname{vec}(\widehat{e^{-\boldsymbol{\Lambda}_n}}) - \mathbf{r})'\left(\mathbf{R}((\mathbb{Z}'_{1,n-1}\mathbb{Z}_{1,n-1})^{-1} \otimes \widehat{\boldsymbol{\Omega}}_n)\mathbf{R}'\right)^{-1}(\mathbf{R}\operatorname{vec}(\widehat{e^{-\boldsymbol{\Lambda}_n}}) - \mathbf{r}) \Longrightarrow \chi_q^2.$$

In contrast to the asymptotic limit of the least squares estimator in Proposition 3.2, in the t- and the Wald-statistic the limit distributions are independent of $\boldsymbol{\Lambda}$.

The sequence of random matrices $n(\mathbb{Z}'_{1,n-1}\mathbb{Z}_{1,n-1})^{-1} \otimes \widehat{\boldsymbol{\Omega}}_n$ are estimators for the covariance matrix of $\operatorname{vec}(\widehat{e^{-\boldsymbol{\Lambda}_n}})$. However, in the infinite variance case the covariance does not exist and hence, the sequence does not converge.

Theorem 3.11
Let $0 < \alpha < 2$ and $\mathbf{L}(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$, where $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_p$ if $1 < \alpha < 2$. Suppose Assumption A holds and $\mathbb{P}(\det(\mathbf{S}_3) = 0) = 0$. Let $\widehat{\boldsymbol{\Omega}}_n$ be given as in (3.2).

(a) Then as $n \to \infty$,

$$t_{\widehat{e^{-\boldsymbol{\Lambda}_n}}} := \frac{a_n^2}{\widetilde{a}_n}n^{-1/2}\widehat{\boldsymbol{\Omega}}_n^{-1/2}\left(\widehat{e^{-\boldsymbol{\Lambda}_n}} - e^{-\boldsymbol{\Lambda}}\right)(\mathbb{Z}'_{1,n-1}\mathbb{Z}_{1,n-1})^{1/2} \Longrightarrow \mathbf{S}_2^{-1/2}\mathbf{S}_5\mathbf{S}_3^{-1/2}.$$

Let $\mathbf{R} \in \mathbb{R}^{1 \times d^2}$ and $r \in \mathbb{R}$. If the null hypothesis H_0: $\mathbf{R}\operatorname{vec}(e^{-\boldsymbol{\Lambda}}) = r$ is true, then as $n \to \infty$,

$$\frac{a_n^2}{\widetilde{a}_n}n^{-1/2}(\mathbf{R}((\mathbb{Z}'_{1,n-1}\mathbb{Z}_{1,n-1}) \otimes \widehat{\boldsymbol{\Omega}}_n^{-1})\mathbf{R}')^{1/2}\left(\mathbf{R}\operatorname{vec}(\widehat{e^{-\boldsymbol{\Lambda}_n}}) - r\right) \Longrightarrow (\mathbf{R}(\mathbf{S}_3 \otimes \mathbf{S}_2^{-1})\mathbf{R}')^{1/2}\mathbf{R}\operatorname{vec}(\mathbf{S}_5\mathbf{S}_3^{-1}).$$

(b) *Suppose the stronger Assumption B holds. Then as $n \to \infty$,*

$$t_{\widehat{\mathrm{e}^{-\Lambda_n}}} E_{\Lambda,\alpha} E_{\Lambda,2}^{-1} \Longrightarrow [\mathbf{S}_1^*, \mathbf{S}_1^*]_1^{-1/2} \mathbf{S}_2^*(1) [\mathbf{S}_1^*, \mathbf{S}_1^*]_1^{-1/2}.$$

(c) *Let $\mathbf{R} \in \mathbb{R}^{q \times d^2}$, $\mathbf{r} \in \mathbb{R}^q$ and $\mathrm{rank}(\mathbf{R}) = q$. If the null hypothesis $H_0 \colon \mathbf{R} \, \mathrm{vec}(\mathrm{e}^{-\Lambda}) = \mathbf{r}$ is true, then as $n \to \infty$,*

$$\frac{a_n^4}{\widetilde{a}_n^2} n^{-1} (\mathbf{R} \, \mathrm{vec}(\widehat{\mathrm{e}^{-\Lambda_n}}) - \mathbf{r})' \left(\mathbf{R}((\mathbb{Z}_{1,n-1}' \mathbb{Z}_{1,n-1})^{-1} \otimes \widehat{\mathbf{\Omega}}_n) \mathbf{R}' \right)^{-1} (\mathbf{R} \, \mathrm{vec}(\widehat{\mathrm{e}^{-\Lambda_n}}) - \mathbf{r})$$

$$\Longrightarrow (\mathbf{R} \, \mathrm{vec}(\mathbf{S}_5 \mathbf{S}_3^{-1}))' \left(\mathbf{R} \left(\mathbf{S}_3^{-1} \otimes \mathbf{S}_2 \right) \mathbf{R}' \right)^{-1} (\mathbf{R} \, \mathrm{vec}(\mathbf{S}_5 \mathbf{S}_3^{-1})).$$

Remark 3.12

To understand the limit results better, we look at the one-dimensional case $d = 1$ with $\Lambda = \lambda > 0$. Then for the t- and the Wald-statistic we obtain as $n \to \infty$,

$$
\begin{aligned}
t_{\widehat{\mathrm{e}^{-\lambda_n}}} &\Longrightarrow (1 - \mathrm{e}^{-\alpha\lambda})^{-1/\alpha}(1 - \mathrm{e}^{-2\lambda})^{1/2} \frac{S_2^*}{S_1^*}, \\
t_{\widehat{\mathrm{e}^{-\lambda_n}}}^2 &\Longrightarrow (1 - \mathrm{e}^{-\alpha\lambda})^{-2/\alpha}(1 - \mathrm{e}^{-2\lambda}) \frac{S_2^{*2}}{S_1^{*2}},
\end{aligned}
\tag{3.3}
$$

where S_2^* is an $S_\alpha(1, 0, 0)$-stable random variable independent of the positive $S_{\alpha/2}(1, 1, 0)$-stable random variable S_1^*. On the one hand, we see that the t- and the Wald-statistic still depend on λ. On the other hand, the limit distributions are again heavy tailed in contrast to the normal and the χ^2-distribution in the finite variance setting of Proposition 3.10, which are the usual limit distributions of the t- and the Wald-statistic. Furthermore, we still get not rid of the norming constants a_n and $\widetilde{a}_n$ which are in general not known. To conclude, this makes it difficult to use the t- and the Wald statistic for the statistical analysis of heavy tailed models. However, we have again in the case with infinite variance a faster convergence rate of the t- and the Wald-statistic than in the finite variance setting of Proposition 3.10.

Again Proposition 3.10 can be interpreted as a special case of Theorem 3.11 by taking $\alpha = 2$. If $\alpha = 2$ then $S_2^* \sim \mathcal{N}(0, 1)$, $S_1^* = 1$ and $(1 - \mathrm{e}^{-\alpha\lambda})^{-1/\alpha}(1 - \mathrm{e}^{-2\lambda})^{1/2} = 1$ in (3.3). $\qquad \square$

4 Multiple time series regression with Ornstein-Uhlenbeck processes

In this section we investigate a multiple regression model used in co-integration. Therefore, let $((\mathbf{L}_1(t)', \mathbf{L}_2(t)')')_{t \geq 0}$ be an $h + d$-dimensional Lévy process and $\mathbf{A} \in \mathbb{R}^{d \times h}$. Then the regression model is defined as

$$
\begin{aligned}
\mathbf{X}(t) &= \mathbf{A}\mathbf{Y}(t) + \mathbf{Z}(t) &\text{for } t \geq 0, &\quad \text{in } \mathbb{R}^d, \\
\mathbf{Y}(t) &= \mathbf{L}_1(t) &\text{for } t \geq 0, &\quad \text{in } \mathbb{R}^h,
\end{aligned}
\tag{4.1}
$$

where $(\mathbf{Z}(t))_{t \geq 0}$ is a stationary Ornstein-Uhlenbeck process in $\mathbb{R}^d$ with representation

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

as in (1.4). The system of simultaneous equations (4.1) is co-integrated and in particulary, $\mathbf{X}$ and $\mathbf{Y}$ are non-stationary. Models of this kind can be found in any equilibrium relationship which are common in financial and econometric time series; see Fasen (2010). The equidistant observations are

$$
\begin{aligned}
\mathbb{X}'_n &= (\mathbf{X}(1), \ldots, \mathbf{X}(n)) &&\in \mathbb{R}^{d \times n}, & \mathbb{X}'_{i,j} &= (\mathbf{X}(i), \ldots, \mathbf{X}(j)) &&\in \mathbb{R}^{d \times (j-i+1)}, \\
\mathbb{Y}'_n &= (\mathbf{Y}(1), \ldots, \mathbf{Y}(n)) &&\in \mathbb{R}^{h \times n}, & \mathbb{Y}'_{i,j} &= (\mathbf{Y}(i), \ldots, \mathbf{Y}(j)) &&\in \mathbb{R}^{h \times (j-i+1)}.
\end{aligned}
$$

Note that here we do not observe $\mathbb{Z}'_n$. The matrices $\mathbf{A}$ and $\mathrm{e}^{-\mathbf{\Lambda}}$ are then estimated in a two-stage procedure. First, we will estimate $\mathbf{A}$ and then we use this result to estimate $\mathrm{e}^{-\mathbf{\Lambda}}$. As estimator for $\mathbf{A}$ we take the least squares estimator

$$
\widehat{\mathbf{A}}_n = \mathbb{X}'_n \mathbb{Y}_n (\mathbb{Y}'_n \mathbb{Y}_n)^{-1}. \tag{4.2}
$$

The asymptotic properties of $\widehat{\mathbf{A}}_n$ are studied in detail in Fasen (2010). However, the objective is now to estimate $\mathrm{e}^{-\mathbf{\Lambda}}$ in a second stage. Thus, define

$$
\widehat{\mathbb{Z}}'_{1,n-1} := \mathbb{X}'_{1,n-1} - \widehat{\mathbf{A}}_n \mathbb{Y}_{1,n-1} \qquad \text{and} \qquad \widehat{\mathbb{Z}}'_{2,n} := \mathbb{X}'_{2,n} - \widehat{\mathbf{A}}_n \mathbb{Y}'_{2,n}
$$

as estimators for the noise $\mathbb{Z}'_{1,n-1}$ and $\mathbb{Z}'_{2,n}$, respectively. Then we use as estimator for $\mathrm{e}^{-\mathbf{\Lambda}}$ the least squares estimator (1.6) where we plug in the estimated noise resulting in

$$
\widetilde{\mathrm{e}^{-\mathbf{\Lambda}_n}} = \widehat{\mathbb{Z}}'_{2,n} \widehat{\mathbb{Z}}_{1,n-1} (\widehat{\mathbb{Z}}'_{1,n-1} \widehat{\mathbb{Z}}_{1,n-1})^{-1}. \tag{4.3}
$$

In this section we will see that the estimator (4.3) has the same asymptotic properties as the estimator (1.6) in Section 3 where the noise is known. However, we suppose the following assumption on the Lévy process $(\mathbf{L}_1(t))_{t \geq 0}$ in (4.1).

Assumption C

Suppose that $(\mathbf{L}_1(t))_{t \geq 0} = ((L_{1,1}(t), \ldots, L_{1,h}(t))')_{t \geq 0}$ satisfies either condition (a) or (b) below.

(a) Let $0 < \alpha^* < 2$, $\mathbf{L}_1(1) \in \mathcal{R}_{-\alpha^*}(b_n, \mu_1)$ and $\mu_1(\{\mathbf{0}_{i-1}\} \times \overline{\mathbb{R}} \backslash \{0\} \times \{\mathbf{0}_{h-i}\}) > 0$ for $i = 1, \ldots, h$. If $1 < \alpha^* < 2$ then $\mathbb{E}(\mathbf{L}_1(1)) = \mathbf{0}_h$.

(b) Let $\mathbb{E}\|\mathbf{L}_1(1)\|^2 < \infty$, $\Sigma_{\mathbf{L}_1(1)}$ be invertible and $\mathbb{E}(\mathbf{L}_1(1)) = \mathbf{0}_h$. Then define $b_n := \sqrt{n}$ and $\alpha^* := 2$.

4.1 Statistical estimation

Again we start with the case where second moments exists.

Theorem 4.1

Let $(\mathbf{L}_1(t))_{t \geq 0}$ and $(\mathbf{L}_2(t))_{t \geq 0}$ be independent Lévy processes, where $(\mathbf{L}_1(t))_{t \geq 0}$ satisfies Assumption C, $\mathbb{E}\|\mathbf{L}_2(1)\|^r < \infty$ for some $r > 2$ and $\mathbb{E}(\mathbf{L}_2(1)) = \mathbf{0}_d$. Define $\Sigma_{\mathbf{Z}}, \Sigma_{\boldsymbol{\xi}}$ as in (3.1) with $\Sigma = \mathbf{I}_{d \times d}$ and $\mathbf{L} = \mathbf{L}_2$, and suppose $\Sigma_{\mathbf{Z}}$ is invertible. Then as $n \to \infty$,

$$
\sqrt{n}\left(vec(\widetilde{\mathrm{e}^{-\mathbf{\Lambda}_n}}) - vec(\mathrm{e}^{-\mathbf{\Lambda}})\right) \Longrightarrow \mathcal{N}(\mathbf{0}_{d^2}, \Sigma_{\mathbf{Z}}^{-1} \otimes \Sigma_{\boldsymbol{\xi}}).
$$

In particular, $\widetilde{\mathrm{e}^{-\mathbf{\Lambda}_n}} \overset{\mathbb{P}}{\longrightarrow} \mathrm{e}^{-\mathbf{\Lambda}}$ as $n \to \infty$.

We conjecture that it is also sufficient to assume a finite second moment of $\|\mathbf{L}_2(1)\|$. However, in the proof we apply results of Fasen (2010) which are based on slightly higher moments.

In the infinite variance case we first present the result for the consistency and then the asymptotic distribution. The consistency holds under more general assumptions than the asymptotic limit result.

Theorem 4.2

Let $0 < \alpha < 2$ and let $(\mathbf{L}(t))_{t \geq 0}$ be a p-dimensional Lévy process satisfying $\mathbf{L}(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$ with $\mathbb{E}(\mathbf{L}(1)) = \mathbf{0}_p$ if $1 < \alpha < 2$. Suppose that $(\mathbf{L}_2(t))_{t \geq 0} = (\Sigma \mathbf{L}(t))_{t \geq 0}$ for $\Sigma \in \mathbb{R}^{d \times p}$, $\mathbb{P}(\det(\mathbf{S}_3) = 0) = 0$ and that either

(a) *$(\mathbf{L}_1(1), \mathbf{L}_2(1)) \in \mathcal{R}_{-\alpha}(a_n, \widetilde{\mu})$ and $\widetilde{\mu}(\{\mathbf{0}_{i-1}\} \times \overline{\mathbb{R}} \backslash \{0\} \times \{\mathbf{0}_{h+d-i}\}) > 0$ for $i = 1, \ldots, h$, or*

(b) *$(\mathbf{L}_1(t))_{t \geq 0}$ and $(\mathbf{L}(t))_{t \geq 0}$ are independent and $(\mathbf{L}_1(t))_{t \geq 0}$ satisfies Assumption C.*

Then the estimator $\mathrm{e}^{\widetilde{-\mathbf{\Lambda}_n}}$ as given in (4.3) is consistent, i.e. $\mathrm{e}^{\widetilde{-\mathbf{\Lambda}_n}} \overset{\mathbb{P}}{\longrightarrow} \mathrm{e}^{-\mathbf{\Lambda}}$ as $n \to \infty$.

The asymptotic behavior of the estimator requires inter alia the additional assumption $\alpha > 1$, i.e. we require a finite first moment. This reflects also the feature of most empirical data, where α is larger than 1.5.

Theorem 4.3

Let the assumptions of Theorem 4.2 hold. Suppose additionally $1 < \alpha < 2$.

(a) *Let further Assumption A hold. Then as $n \to \infty$,*

$$\frac{a_n^2}{\widetilde{a}_n} \left(\mathrm{e}^{\widetilde{-\mathbf{\Lambda}_n}} - \mathrm{e}^{-\mathbf{\Lambda}} \right) \Longrightarrow \mathbf{S}_5 \mathbf{S}_3^{-1}.$$

(b) *Let further Assumption B hold. Then as $n \to \infty$,*

$$\frac{a_n^2}{\widetilde{a}_n} D_{\mathbf{\Lambda}, \alpha}^{-1} E_{\mathbf{\Lambda}, \alpha}^{-1} \left(\mathrm{e}^{\widetilde{-\mathbf{\Lambda}_n}} - \mathrm{e}^{-\mathbf{\Lambda}} \right) D_{\mathbf{\Lambda}, \alpha} E_{\mathbf{\Lambda}, \alpha}^2 E_{\mathbf{\Lambda}, 2}^{-2} \Longrightarrow \mathbf{S}_2^*(1)([\mathbf{S}_1^*, \mathbf{S}_1^*]_1)^{-1}.$$

To conclude, both in the infinite variance and the finite variance case the classical least squares estimator (1.6) based on the observed Ornstein-Uhlenbeck process and the estimator (4.3) based on no observations of the Ornstein-Uhlenbeck process have the same properties.

4.2 Asymptotic tests

For the t- and the Wald statistic we obtain in the multiple regression model (4.1) the same results as in Section 3.2.

Theorem 4.4

Let the assumptions of Theorem 4.1 hold, and define

$$\widetilde{\mathbf{\Omega}}_n := n^{-1}(\widehat{\mathbb{Z}}_{2,n}' - \mathrm{e}^{\widetilde{-\mathbf{\Lambda}_n}} \widehat{\mathbb{Z}}_{1,n-1}')(\widehat{\mathbb{Z}}_{2,n}' - \mathrm{e}^{\widetilde{-\mathbf{\Lambda}_n}} \widehat{\mathbb{Z}}_{1,n-1}')'. \tag{4.4}$$

(a) Then as $n \to \infty$,

$$((\widehat{\mathbb{Z}}'_{1,n-1}\widehat{\mathbb{Z}}_{1,n-1})^{1/2} \otimes \widetilde{\mathbf{\Omega}}_n^{-1/2}) \left(vec(\widetilde{e^{-\mathbf{\Lambda}_n}}) - vec(e^{-\mathbf{\Lambda}})\right) \implies \mathcal{N}(\mathbf{0}_{d^2}, \mathbf{I}_{d^2 \times d^2}).$$

Let $\mathbf{R} \in \mathbb{R}^{1 \times d^2}$ and $r \in \mathbb{R}$. If the null hypothesis $H_0: \mathbf{R}\,vec(e^{-\mathbf{\Lambda}}) = r$ is true, then as $n \to \infty$,

$$(\mathbf{R}((\widehat{\mathbb{Z}}'_{1,n-1}\widehat{\mathbb{Z}}_{1,n-1}) \otimes \widetilde{\mathbf{\Omega}}_n^{-1})\mathbf{R}')^{1/2} \left(\mathbf{R}vec(\widetilde{e^{-\mathbf{\Lambda}_n}}) - r\right) \implies \mathcal{N}(0,1).$$

(b) Let $\mathbf{R} \in \mathbb{R}^{q \times d^2}$, $\mathbf{r} \in \mathbb{R}^q$ and $\mathrm{rank}(\mathbf{R}) = q$. If the null hypothesis $H_0: \mathbf{R}\,vec(e^{-\mathbf{\Lambda}}) = \mathbf{r}$ is true, then as $n \to \infty$,

$$(\mathbf{R}\,vec(\widetilde{e^{-\mathbf{\Lambda}_n}}) - \mathbf{r})' \left(\mathbf{R}((\widehat{\mathbb{Z}}'_{1,n-1}\widehat{\mathbb{Z}}_{1,n-1})^{-1} \otimes \widetilde{\mathbf{\Omega}}_n)\mathbf{R}'\right)^{-1} \left(\mathbf{R}\,vec(\widetilde{e^{-\mathbf{\Lambda}_n}}) - \mathbf{r}\right) \implies \chi_q^2.$$

Theorem 4.5

Let the assumptions of Theorem 4.2 and Assumption A hold, and define $\widetilde{\mathbf{\Omega}}_n$ as in (4.4). Suppose additionally $1 < \alpha < 2$.

(a) Then as $n \to \infty$,

$$t_{\widetilde{e^{-\mathbf{\Lambda}_n}}} := \frac{a_n^2}{\widetilde{a}_n} n^{-1/2} \widetilde{\mathbf{\Omega}}_n^{-1/2} \left(\widetilde{e^{-\mathbf{\Lambda}_n}} - e^{-\mathbf{\Lambda}}\right) (\widehat{\mathbb{Z}}'_{1,n-1}\widehat{\mathbb{Z}}_{1,n-1})^{1/2} \implies \mathbf{S}_2^{-1/2}\mathbf{S}_5\mathbf{S}_3^{-1/2}.$$

Let $\mathbf{R} \in \mathbb{R}^{1 \times d^2}$ and $r \in \mathbb{R}$. If the null hypothesis $H_0: \mathbf{R}\,vec(e^{-\mathbf{\Lambda}}) = r$ is true, then as $n \to \infty$,

$$\frac{a_n^2}{\widetilde{a}_n} n^{-1/2}(\mathbf{R}((\widehat{\mathbb{Z}}'_{1,n-1}\widehat{\mathbb{Z}}_{1,n-1}) \otimes \widetilde{\mathbf{\Omega}}_n^{-1})\mathbf{R}')^{1/2}(\mathbf{R}vec(\widetilde{e^{-\mathbf{\Lambda}_n}}) - r) \implies (\mathbf{R}(\mathbf{S}_3 \otimes \mathbf{S}_2^{-1})\mathbf{R}')^{1/2}\mathbf{R}vec(\mathbf{S}_5\mathbf{S}_3^{-1}).$$

(b) Suppose the stronger Assumption B holds. Then as $n \to \infty$,

$$t_{\widetilde{e^{-\mathbf{\Lambda}_n}}} E_{\mathbf{\Lambda},\alpha} E_{\mathbf{\Lambda},2}^{-1} \implies [\mathbf{S}_1^*, \mathbf{S}_1^*]_1^{-1/2} \mathbf{S}_2^*(1)[\mathbf{S}_1^*, \mathbf{S}_1^*]_1^{-1/2}.$$

(c) Let $\mathbf{R} \in \mathbb{R}^{q \times d^2}$, $\mathbf{r} \in \mathbb{R}^q$ and $\mathrm{rank}(\mathbf{R}) = q$. If the null hypothesis $H_0: \mathbf{R}\,vec(e^{-\mathbf{\Lambda}}) = \mathbf{r}$ is true, then as $n \to \infty$,

$$\frac{a_n^4}{\widetilde{a}_n^2} n^{-1}(\mathbf{R}\,vec(\widetilde{e^{-\mathbf{\Lambda}_n}}) - \mathbf{r})' \left(\mathbf{R}((\widehat{\mathbb{Z}}'_{1,n-1}\widehat{\mathbb{Z}}_{1,n-1})^{-1} \otimes \widetilde{\mathbf{\Omega}}_n)\mathbf{R}'\right)^{-1} (\mathbf{R}\,vec(\widetilde{e^{-\mathbf{\Lambda}_n}}) - \mathbf{r})$$

$$\implies (\mathbf{R}\,vec(\mathbf{S}_5\mathbf{S}_3^{-1}))' \left(\mathbf{R}\left(\mathbf{S}_3^{-1} \otimes \mathbf{S}_2\right)\mathbf{R}'\right)^{-1} (\mathbf{R}\,vec(\mathbf{S}_5\mathbf{S}_3^{-1})).$$

5 Proofs

5.1 Proofs of Section 2

Proof of Lemma 2.2. Let us define

$$\mu_j(\mathrm{d}x) := p_j \alpha x^{-\alpha-1} \mathbb{1}_{(0,\infty)}(x)\,\mathrm{d}x + q_j \alpha(-x)^{-\alpha-1} \mathbb{1}_{(-\infty,0)}(x)\,\mathrm{d}x$$

and $K := \lim_{n\to\infty} n\mathbb{P}(\|\boldsymbol{\zeta}^{(1)}\| > a_n)$ such that $\zeta_j^{(1)} \in \mathcal{R}_{-\alpha}(a_n, K\mu_j(\cdot))$, $j = 1, \ldots, p$. Then by the independence of the components $\zeta_1^{(1)}, \ldots, \zeta_p^{(1)}$ of $\boldsymbol{\zeta}^{(1)}$ we have for μ the representation

$$\mu(\cdot) = K \sum_{j=1}^p \mu_j \circ \mathrm{pr}_j^{(p)}(E_j^{(p)} \cap \cdot), \tag{5.1}$$

where $E_j^{(p)} = \{y\mathbf{e}_j : y \in \mathbb{R}\} \subseteq \mathbb{R}^p$ and $\mathrm{pr}_j^{(p)} : \mathbb{R}^p \to \mathbb{R}$ is the projection on the j-th coordinate. Furthermore, we define

$$\mu_{ij}(\mathrm{d}x) := \widetilde{p}_{ij}\alpha x^{-\alpha-1}\mathbb{1}_{(0,\infty)}(x)\,\mathrm{d}x + \widetilde{q}_{ij}\alpha(-x)^{-\alpha-1}\mathbb{1}_{(-\infty,0)}(x)\,\mathrm{d}x.$$

Then Cline (1986), Proposition II, results in $\zeta_i^{(1)}\zeta_j^{(2)} \in \mathcal{R}_{-\alpha}(\widetilde{a}_n, C\mu_{ij}(\cdot))$.

Let $\mathcal{S}$ be the collection of all sets B of the form $B = \{(x_{ij})_{i,j=1,\ldots,p} \in \mathbb{R}^{p\times p} : x_{ij} \in (b_{ij}, c_{ij}], i, j = 1, \ldots, p\} \subseteq \mathbb{R}^{p\times p}$, which are bounded away from $\mathbf{0}_{p\times p}$ and $b_{ij} < c_{ij}$ $i, j = 1, \ldots, p$. It is clear that $\mathcal{S}$ is a DC-semiring (cf. Lindskog (2004), Lemma 1.10). W.l.o.g. $b_{ij} \neq 0$, $c_{ij} \neq 0$ for $i, j = 1, \ldots, p$. Moreover, since $B \in \mathcal{S}$ is bounded away from $\mathbf{0}_{p\times p}$, either

$$B \cap E_{ij}^{(p)} = \emptyset \quad \text{for } i, j = 1, \ldots, p, \tag{C1}$$

where $E_{ij}^{(p)} = \{y\mathbf{e}_{ij} : y \in \mathbb{R}\} \subseteq \mathbb{R}^{p\times p}$, or

$$B \cap E_{ij}^{(p)} = (b_{ij}, c_{ij}]\,\mathbf{e}_{ij} \quad \text{and} \quad B \cap E_{kl}^{(p)} = \emptyset \quad \text{for } (k,l) \neq (i,j), k, l = 1, \ldots, p. \tag{C2}$$

That is, B has either empty intersection with all coordinate axes or intersects exactly one. Note that in (C2),

$$b_{kl} < 0 < c_{kl} \quad \text{for } k \neq i \text{ or } l \neq j \qquad \text{and} \qquad 0 \notin (b_{ij}, c_{ij}].$$

Case (C1): Then $b_{ij}, c_{ij} > 0$ or $b_{ij}, c_{ij} < 0$ for $i, j = 1, \ldots, p$. Define $x^* := \min(|b_{11}|, |c_{11}|) > 0$ and $y^* := \min(|b_{pp}|, |c_{pp}|) > 0$. Thus,

$$n\mathbb{P}(\widetilde{a}_n^{-1}\boldsymbol{\zeta}^{(1)}\boldsymbol{\zeta}^{(2)'} \in B) \leq n\mathbb{P}(|\zeta_1^{(1)}\zeta_1^{(2)}| > \widetilde{a}_n x^*)\mathbb{P}(|\zeta_p^{(1)}\zeta_p^{(2)}| > \widetilde{a}_n y^*) \longrightarrow 0 \quad \text{as } n \to \infty.$$

Case (C2): W.l.o.g. $B \cap E_{11}^{(p)} = (b_{11}, c_{11}]\,\mathbf{e}_{11}$. Define $x^* := \min(|b_{11}|, |c_{11}|) > 0$ and $y^* := \min_{\substack{k,l=1,\ldots,p \\ (k,l)\neq(1,1)}}(|b_{kl}|, |c_{kl}|) > 0$. Then

$$\begin{aligned}
&|n\mathbb{P}(\widetilde{a}_n^{-1}\boldsymbol{\zeta}^{(1)}\boldsymbol{\zeta}^{(2)'} \in B) - n\mathbb{P}(b_{11}\widetilde{a}_n < \zeta_1^{(1)}\zeta_1^{(2)} < c_{11}\widetilde{a}_n)| \\
&\quad \leq 2n\sum_{k=2}^p \mathbb{P}(|\zeta_1^{(1)}\zeta_1^{(2)}| > \widetilde{a}_n x^*, |\zeta_1^{(1)}\zeta_k^{(2)}| > \widetilde{a}_n y^*) + n\sum_{k,l=2}^p \mathbb{P}(|\zeta_1^{(1)}\zeta_1^{(2)}| > \widetilde{a}_n x^*, |\zeta_k^{(1)}\zeta_l^{(2)}| > \widetilde{a}_n y^*) \\
&\quad \leq 2np\mathbb{P}(|\zeta_1^{(1)}\zeta_1^{(2)}| > \widetilde{a}_n x^*, |\zeta_1^{(1)}|\max_{k=2,\ldots,p}|\zeta_k^{(2)}| > \widetilde{a}_n y^*) \\
&\qquad + np^2\mathbb{P}(|\zeta_1^{(1)}\zeta_1^{(2)}| > \widetilde{a}_n x^*)\mathbb{P}(\max_{k=2,\ldots,p}|\zeta_k^{(1)}|\max_{l=2,\ldots,p}|\zeta_l^{(2)}| > \widetilde{a}_n y^*) \longrightarrow 0 \quad \text{as } n \to \infty,
\end{aligned}$$

where the first term tends to 0, since $\widetilde{a}_n a_n^{-1} \to \infty$ as $n \to \infty$ and

$$n\mathbb{P}(|\zeta_1^{(1)}\zeta_1^{(2)}| > \widetilde{a}_n x^*, |\zeta_1^{(1)}| \max_{k=2,\ldots,p} |\zeta_k^{(2)}| > \widetilde{a}_n y^*)$$

$$\leq n\mathbb{P}(|\zeta_1^{(1)}| > a_n M) + n\mathbb{P}(|\zeta_1^{(1)}\zeta_1^{(2)}| > \widetilde{a}_n x^*)\mathbb{P}(\max_{k=2,\ldots,p} |\zeta_k^{(2)}| > \widetilde{a}_n y^*(a_n M)^{-1})$$

$$\overset{n\to\infty}{\longrightarrow} (p_1 + q_1)M^{-\alpha} \overset{M\to\infty}{\longrightarrow} 0.$$

Thus, we obtain

$$\lim_{n\to\infty} n\mathbb{P}(\widetilde{a}_n^{-1}\boldsymbol{\zeta}^{(1)}\boldsymbol{\zeta}^{(2)'} \in B) = C\sum_{i,j=1}^{p} \mu_{ij} \circ \mathrm{pr}_{ij}^{(p)}(E_{ij}^{(p)} \cap B) = C\widetilde{\mu}(B),$$

where $\mathrm{pr}_{ij}^{(p)} : \mathbb{R}^{p\times p} \to \mathbb{R}$ is the projection on the (i,j)-th coordinate. Finally, we can conclude $\boldsymbol{\zeta}^{(1)}\boldsymbol{\zeta}^{(2)'} \in \mathcal{R}_{-\alpha}(a_n, C\widetilde{\mu}(\cdot))$. $\qquad\square$

Proof of Theorem 2.4. Let $(\mathbf{b}, \mathbf{C}, \nu)$ denote the characteristic triplet of $(\mathbf{L}(t))_{t\geq 0}$.

Step 1. First, we will assume that $(\mathbf{L}(t))_{t\geq 0}$ is a compound Poisson process such that

$$\boldsymbol{\xi}_1 = \sum_{j=1}^{N_1(1)} \mathrm{e}^{-\boldsymbol{\Lambda} U_j^{(1)}} \Sigma \boldsymbol{\zeta}_j^{(1)} \qquad \text{and} \qquad \boldsymbol{\xi}_2 = \sum_{k=1}^{N_2(1)} \mathrm{e}^{-\boldsymbol{\Lambda} U_k^{(2)}} \Sigma \boldsymbol{\zeta}_k^{(2)},$$

where $\boldsymbol{\zeta}_1^{(1)}, \boldsymbol{\zeta}_1^{(2)}, \boldsymbol{\zeta}_2^{(1)}, \boldsymbol{\zeta}_2^{(2)}, \ldots$ is an iid sequence with distribution $\nu(\cdot)/\nu(\mathbb{R}^p)$ where the components of any vector are independent and $(N_1(t))_{t\geq 0}, (N_2(t))_{t\geq 0}$ are iid Poisson processes with intensity $\nu(\mathbb{R}^p)$. Furthermore, $U_1^{(1)}, U_1^{(2)}, U_2^{(1)}, U_2^{(2)}, \ldots$ is a sequence of iid random variables uniformly distributed on $(0,1)$. Finally, all random variables are independent. Then

$$\boldsymbol{\xi}_1\boldsymbol{\xi}_2' = \sum_{j=1}^{N_1(1)}\sum_{k=1}^{N_2(1)} \mathrm{e}^{-\boldsymbol{\Lambda} U_j^{(1)}} \Sigma \boldsymbol{\zeta}_j^{(1)}\boldsymbol{\zeta}_k^{(2)'} \Sigma' \mathrm{e}^{-\boldsymbol{\Lambda}' U_k^{(2)}}.$$

Now, let $B = \{(x_{ij})_{i,j=1,\ldots,d} \in \mathbb{R}^{d\times d} : x_{ij} \in (b_{ij}, c_{ij}], i, j = 1, \ldots, d\} \subseteq \mathbb{R}^{d\times d}$ with $b_{ij} < c_{ij}$, $b_{ij} \neq 0$, $c_{ij} \neq 0$, $i, j = 1, \ldots, d$, be bounded away from $\mathbf{0}_{d\times d}$ as in Lemma 2.2. We conclude that

$$\lim_{n\to\infty} n\mathbb{P}(\widetilde{a}_n^{-1}\boldsymbol{\xi}_1\boldsymbol{\xi}_2' \in B) \tag{5.2}$$

$$= \lim_{n\to\infty} \sum_{l=0}^{\infty}\sum_{m=0}^{\infty} n\mathbb{P}\left(\widetilde{a}_n^{-1}\sum_{j=1}^{l}\sum_{k=1}^{m} \mathrm{e}^{-\boldsymbol{\Lambda} U_j^{(1)}} \Sigma \boldsymbol{\zeta}_j^{(1)}\boldsymbol{\zeta}_k^{(2)'} \Sigma' \mathrm{e}^{-\boldsymbol{\Lambda}' U_k^{(2)}} \in B\right) \mathbb{P}(N_1(1) = l)\mathbb{P}(N_2(1) = m).$$

Next, since $\boldsymbol{\zeta}_1^{(1)}$ has distribution $\nu(\cdot)/\nu(\mathbb{R}^p)$, which is the normalized Lévy measure of $\mathbf{L}(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$ with μ as given in (5.1), we have by Hult and Lindskog (2007), Lemma 2.1, that $\boldsymbol{\zeta}_1^{(1)} \in \mathcal{R}_{-\alpha}(a_n, \mu(\cdot)/\nu(\mathbb{R}^p))$ and $\lim_{n\to\infty} n\mathbb{P}(\|\boldsymbol{\zeta}_1^{(1)}\| > a_n) = \nu(\mathbb{R}^p)^{-1}$. Thus, by Cline (1986), Proposition II,

$$\lim_{n\to\infty} n\mathbb{P}(\|\boldsymbol{\zeta}_1^{(1)}\|\|\boldsymbol{\zeta}_1^{(2)}\| > \widetilde{a}_n) = \nu(\mathbb{R}^p)^{-2}$$

such that Lemma 2.2 results in $\boldsymbol{\zeta}_1^{(1)}\boldsymbol{\zeta}_1^{(2)'} \in \mathcal{R}_{-\alpha}(\widetilde{a}_n, \widetilde{\mu}(\cdot)/\nu(\mathbb{R}^p)^2)$. Since $\mathbb{E}\|\mathrm{e}^{-\boldsymbol{\Lambda}U_1^{(i)}}\|^\delta < \infty$ for any $\delta > 0$, a multivariate version of Breiman's result (cf. Basrak et al. (2002), Proposition A.1) gives then

$$\mathrm{e}^{-\boldsymbol{\Lambda}U_1^{(1)}}\Sigma\boldsymbol{\zeta}_1^{(1)}\boldsymbol{\zeta}_1^{(2)'}\Sigma'\mathrm{e}^{-\boldsymbol{\Lambda}'U_1^{(2)}} \in \mathcal{R}_{-\alpha}\left(\widetilde{a}_n, \mu_{\boldsymbol{\xi}\boldsymbol{\xi}'}(\cdot)/\nu(\mathbb{R}^p)^2\right).$$

We interpret for a moment $(\mathrm{e}^{-\boldsymbol{\Lambda}U_j^{(1)}}\Sigma\boldsymbol{\zeta}_j^{(1)}\boldsymbol{\zeta}_k^{(2)'}\Sigma'\mathrm{e}^{-\boldsymbol{\Lambda}'U_k^{(2)}})_{j,k=1,\ldots,m}$ as $(dm \times dm)$ block matrix and write it as

$$(\mathrm{e}^{-\boldsymbol{\Lambda}U_j^{(1)}}\Sigma\boldsymbol{\zeta}_j^{(1)}\boldsymbol{\zeta}_k^{(2)'}\Sigma'\mathrm{e}^{-\boldsymbol{\Lambda}'U_k^{(2)}})_{j,k=1,\ldots,m} = \mathrm{diag}(\mathrm{e}^{-\boldsymbol{\Lambda}U_1^{(1)}}\Sigma,\ldots,\mathrm{e}^{-\boldsymbol{\Lambda}U_m^{(1)}}\Sigma)\mathrm{vec}(\boldsymbol{\zeta}_1^{(1)},\ldots,\boldsymbol{\zeta}_m^{(1)})$$
$$\mathrm{vec}(\boldsymbol{\zeta}_1^{(2)},\ldots,\boldsymbol{\zeta}_m^{(2)})'\mathrm{diag}(\mathrm{e}^{-\boldsymbol{\Lambda}U_1^{(2)}}\Sigma,\ldots,\mathrm{e}^{-\boldsymbol{\Lambda}U_m^{(2)}}\Sigma)',$$

where diag is a block diagonal matrix in $\mathbb{R}^{dm \times pm}$ and

$$\mathrm{vec}(\boldsymbol{\zeta}_1^{(1)},\ldots,\boldsymbol{\zeta}_m^{(1)})\mathrm{vec}(\boldsymbol{\zeta}_1^{(2)},\ldots,\boldsymbol{\zeta}_m^{(2)})' \in \mathcal{R}_{-\alpha}\left(\widetilde{a}_n, \sum_{j,k=1}^m \widetilde{\mu}\circ\mathrm{pr}_{j,k}^{(p,m)}(E_{j,k}^{(p,m)}\cap\cdot)/\nu(\mathbb{R}^p)^2\right)$$

is a random matrix in $\mathbb{R}^{pm \times pm}$ with independent components on an extremal level by Lemma 2.2, where $\mathrm{pr}_{j,k}^{(p,m)} : \mathbb{R}^{pm \times pm} \to \mathbb{R}^{p \times p}$ is the projection $\mathbf{A} = (A_{l,s})_{l,s=1,\ldots,pm} \mapsto (A_{(j-1)p+u,(k-1)p+v})_{u,v=1,\ldots,p}$ and $E_{j,k}^{(p,m)} = \{\mathbf{A} \in \mathbb{R}^{pm \times pm} : A_{l,s} = 0 \text{ for } (l,s) \notin \{((j-1)p+u,(k-1)p+v) : u,v = 1,\ldots,p\}\}$. Thus, again by Basrak et al. (2002), Proposition A.1,

$$(\mathrm{e}^{-\boldsymbol{\Lambda}U_j^{(1)}}\Sigma\boldsymbol{\zeta}_j^{(1)}\boldsymbol{\zeta}_k^{(2)'}\Sigma'\mathrm{e}^{-\boldsymbol{\Lambda}'U_k^{(2)}})_{j,k=1,\ldots,m} \in \mathcal{R}_{-\alpha}\left(\widetilde{a}_n, \sum_{j,k=1}^m \mu_{\boldsymbol{\xi}\boldsymbol{\xi}'}\circ\mathrm{pr}_{j,k}^{(d,m)}(E_{j,k}^{(d,m)}\cap\cdot)/\nu(\mathbb{R}^p)^2\right),$$

which means that the blocks of this random matrix are independent on an extremal level. A final application of Basrak et al. (2002), Proposition A.1, gives

$$\lim_{n\to\infty} n\mathbb{P}\left(\widetilde{a}_n^{-1}\sum_{k=1}^l\sum_{j=1}^m \mathrm{e}^{-\boldsymbol{\Lambda}U_j^{(1)}}\Sigma\boldsymbol{\zeta}_j^{(1)}\boldsymbol{\zeta}_k^{(2)'}\Sigma'\mathrm{e}^{-\boldsymbol{\Lambda}'U_k^{(2)}} \in B\right)$$
$$= \lim_{n\to\infty} lmn\mathbb{P}\left(\widetilde{a}_n^{-1}\mathrm{e}^{-\boldsymbol{\Lambda}U_1^{(1)}}\Sigma\boldsymbol{\zeta}_1^{(1)}\boldsymbol{\zeta}_1^{(2)'}\Sigma'\mathrm{e}^{-\boldsymbol{\Lambda}'U_1^{(2)}} \in B\right) = lm\frac{\mu_{\boldsymbol{\xi}\boldsymbol{\xi}'}(B)}{\nu(\mathbb{R}^p)^2}. \quad (5.3)$$

If we suppose that it is allowed to exchange the limit and the infinite sums in (5.2) then we obtain with

$$\lim_{n\to\infty} n\mathbb{P}(\widetilde{a}_n^{-1}\boldsymbol{\xi}_1\boldsymbol{\xi}_2' \in B) = \mathbb{E}(N_1(1))^2 \lim_{n\to\infty} n\mathbb{P}\left(\widetilde{a}_n^{-1}\left(\mathrm{e}^{-\boldsymbol{\Lambda}U^{(1)}}\Sigma\boldsymbol{\zeta}_1^{(1)}\boldsymbol{\zeta}_1^{(2)'}\Sigma'\mathrm{e}^{-\boldsymbol{\Lambda}'U^{(2)}}\right) \in B\right) = \mu_{\boldsymbol{\xi}\boldsymbol{\xi}'}(B)$$

the result. Hence, it remains to show that the assumptions of Pratt's Theorem (see Pratt (1960)) are satisfied which allows us to exchange the limit and the infinite sums in (5.2).

Therefore, we define $x^* = \inf_{\mathbf{x}\in B}\|\mathbf{x}\|$, which is larger than 0 by the construction of B. Then

$$\mathbb{P}\left(\widetilde{a}_n^{-1}\sum_{k=1}^l\sum_{j=1}^m \mathrm{e}^{-\boldsymbol{\Lambda}U_j^{(1)}}\Sigma\boldsymbol{\zeta}_j^{(1)}\boldsymbol{\zeta}_k^{(2)'}\Sigma'\mathrm{e}^{-\boldsymbol{\Lambda}'U_k^{(2)}} \in B\right) \le \mathbb{P}\left(\widetilde{a}_n^{-1}\sum_{k=1}^l\sum_{j=1}^m \|\boldsymbol{\zeta}_j^{(1)}\|\|\boldsymbol{\zeta}_k^{(2)}\| > C_1 x^*\right). \quad (5.4)$$

20

If we define $\widetilde{L}^{(1)} := \sum_{k=1}^{N_1(1)} \|\boldsymbol{\zeta}_k^{(1)}\|$ and $\widetilde{L}^{(2)} := \sum_{k=1}^{N_2(1)} \|\boldsymbol{\zeta}_k^{(2)}\|$ then $\widetilde{L}^{(i)} \in \mathcal{R}_{-\alpha}(a_n)$ with $\mathbb{E}|\widetilde{L}^{(i)}|^\alpha = \infty$, $i = 1, 2$. By Cline (1986), Proposition II, we know that $\widetilde{L}^{(1)}\widetilde{L}^{(2)} \in \mathcal{R}_{-\alpha}(\widetilde{a}_n)$. Hence,

$$\lim_{n\to\infty} \sum_{l=0}^\infty \sum_{m=0}^\infty n\mathbb{P}\left(\widetilde{a}_n^{-1} \sum_{k=1}^l \sum_{j=1}^m \|\boldsymbol{\zeta}_j^{(1)}\|\|\boldsymbol{\zeta}_k^{(2)}\| > C_1 x^*\right) \mathbb{P}(N_1(1) = l)\mathbb{P}(N_2(1) = m)$$
$$= \lim_{n\to\infty} n\mathbb{P}(\widetilde{a}_n^{-1}\widetilde{L}^{(1)}\widetilde{L}^{(2)} > C_1 x^*) = C_2 < \infty,$$

such that the assumptions of Pratt's Theorem are satisfied.

Step 2. Finally, let $(\mathbf{L}(t))_{t\geq0}$ be a general Lévy process, and $(\mathbf{L}^{(1)}(t))_{t\geq0}$ and $(\mathbf{L}^{(2)}(t))_{t\geq0}$ denote two independent Lévy processes with characteristic triplets $(\mathbf{0}_p, \mathbf{0}_{p\times p}, \nu(\cdot \cap E^c))$ and $(\mathbf{b}, \mathbf{C}, \nu(\cdot \cap E))$, respectively, where $E = \{\mathbf{x} \in \mathbb{R}^p\backslash\{\mathbf{0}_p\} : \|\mathbf{x}\| \leq 1\}$, such that $(\mathbf{L}(t))_{t\geq0} = (\mathbf{L}^{(1)}(t))_{t\geq0} + (\mathbf{L}^{(2)}(t))_{t\geq0}$. Then

$$\boldsymbol{\xi}_1 = \int_0^1 e^{-\mathbf{\Lambda}(1-s)}\Sigma\,\mathbf{L}^{(1)}(ds) + \int_0^1 e^{-\mathbf{\Lambda}(1-s)}\Sigma\,\mathbf{L}^{(2)}(ds) =: \boldsymbol{\xi}_1^{(1)} + \boldsymbol{\xi}_1^{(2)}.$$

The Lévy measure of $\boldsymbol{\xi}_1^{(2)}$ is by Masuda (2004), Proposition 2.1, given by

$$\nu_{\boldsymbol{\xi}^{(2)}}(B) = \int_0^1 \nu(\{\mathbf{x} \in E : e^{-\mathbf{\Lambda}s}\Sigma\mathbf{x} \in B\})\,ds \quad \text{for } B \in \mathcal{B}(\mathbb{R}^d\backslash\{\mathbf{0}_d\}),$$

which has a bounded support. Thus, a conclusion of Sato (2002), Corollary 25.8, is that all moments of $\|\boldsymbol{\xi}_1^{(2)}\|$ exist. We have an analog decomposition of $\boldsymbol{\xi}_2$ as $\boldsymbol{\xi}_2^{(1)} + \boldsymbol{\xi}_2^{(2)}$. Then

$$\boldsymbol{\xi}_1\boldsymbol{\xi}_2' = \boldsymbol{\xi}_1^{(1)}\boldsymbol{\xi}_2^{(1)'} + \boldsymbol{\xi}_1^{(1)}\boldsymbol{\xi}_2^{(2)'} + \boldsymbol{\xi}_1^{(2)}\boldsymbol{\xi}_2^{(1)'} + \boldsymbol{\xi}_1^{(2)}\boldsymbol{\xi}_2^{(2)'}. \tag{5.5}$$

Since by Hult and Lindskog (2007), Lemma 2.1, $\mathbf{L}^{(1)}(1) \in \mathcal{R}_{-\alpha}(a_n, \mu)$ and $(\mathbf{L}^{(1)}(t))_{t\geq0}$ is a compound Poisson process, we have by Step 1 that the first summand $\boldsymbol{\xi}_1^{(1)}\boldsymbol{\xi}_2^{(1)'} \in \mathcal{R}_{-\alpha}(\widetilde{a}_n, \mu_{\boldsymbol{\xi}\boldsymbol{\xi}'})$. On the other hand, $\boldsymbol{\xi}_1^{(1)}, \boldsymbol{\xi}_2^{(1)} \in \mathcal{R}_{-\alpha}(a_n)$ and $\mathbb{E}\|\boldsymbol{\xi}_1^{(2)}\|^{\alpha+1}, \mathbb{E}\|\boldsymbol{\xi}_2^{(2)}\|^{\alpha+1} < \infty$ such that $\boldsymbol{\xi}_1^{(1)}\boldsymbol{\xi}_2^{(2)'}, \boldsymbol{\xi}_1^{(2)}\boldsymbol{\xi}_2^{(1)'} \in \mathcal{R}_{-\alpha}(a_n)$ and $\mathbb{E}\|\boldsymbol{\xi}_1^{(2)}\boldsymbol{\xi}_2^{(2)'}\|^{\alpha+1} < \infty$. Hence,

$$\lim_{x\to\infty} \frac{\mathbb{P}(\|\boldsymbol{\xi}_1^{(1)}\boldsymbol{\xi}_2^{(2)'} + \boldsymbol{\xi}_1^{(2)}\boldsymbol{\xi}_2^{(1)'} + \boldsymbol{\xi}_1^{(2)}\boldsymbol{\xi}_2^{(2)'}\| > x)}{\mathbb{P}(\|\boldsymbol{\xi}_1^{(1)}\boldsymbol{\xi}_2^{(1)'}\| > x)} = 0.$$

The statement (2.2) follows then by (5.5) and Jessen and Mikosch (2006), Lemma 3.3.

Step 3. We define $\widetilde{\boldsymbol{\xi}}_m := \text{vec}(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{m+1})$, $\widetilde{\mathbf{L}}(t) = \text{vec}(\mathbf{L}(t), \mathbf{L}(t+1) - \mathbf{L}(t), \ldots, \mathbf{L}(t+m+1) - \mathbf{L}(t+m))$ for $t \in [0,1]$, $\widetilde{\mathbf{\Lambda}} = \text{diag}(\mathbf{\Lambda}, \ldots, \mathbf{\Lambda}) \in \mathbb{R}^{d(m+1)\times d(m+1)}$, $\widetilde{\Sigma} = \text{diag}(\Sigma, \ldots, \Sigma) \in \mathbb{R}^{d(m+1)\times p(m+1)}$ such that $\widetilde{\boldsymbol{\xi}}_m = \int_0^1 e^{-\widetilde{\mathbf{\Lambda}}(1-s)}\widetilde{\Sigma}\,\widetilde{\mathbf{L}}(ds)$. Then the conclusion (2.3) can be derived as special case of (2.2) and a continuous-mapping theorem. $\square$

Proof of Theorem 2.6. As in Davis and Resnick (1986), Theorem 3.3 (see also Resnick (2007), (9.91)), we have by (2.3) and the $m+1$-dependence of $(\text{vec}(\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_k', \ldots, \boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_{k-m}'))_{k\in\mathbb{N}}$ as $n \to \infty$,

$$\left(\widetilde{a}_n^{-1}\sum_{k=1}^{n-1}\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_k', \ldots, \widetilde{a}_n^{-1}\sum_{k=1}^{n-1}\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_{k-m}'\right) \Longrightarrow (\mathbf{S}_0^*, \ldots, \mathbf{S}_m^*).$$

Applying the continuous mapping theorem yields as $n \to \infty$,

$$\sum_{j=0}^{m}\left(\widetilde{a}_n^{-1}\sum_{k=1}^{n-1}\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_{k-j}'\right)\mathrm{e}^{-\boldsymbol{\Lambda}'j} \Longrightarrow \sum_{j=0}^{m}\mathbf{S}_j^*\mathrm{e}^{-\boldsymbol{\Lambda}'j}.$$

Since $\sum_{j=0}^{m}\mathbf{S}_j^*\mathrm{e}^{-\boldsymbol{\Lambda}'j} \Longrightarrow \sum_{j=0}^{\infty}\mathbf{S}_j^*\mathrm{e}^{-\boldsymbol{\Lambda}'j}$ as $m \to \infty$ it remains to show that for any $\eta > 0$

$$\lim_{m\to\infty}\lim_{n\to\infty}\mathbb{P}\left(\left\|\sum_{j>m}\left(\widetilde{a}_n^{-1}\sum_{k=1}^{n-1}\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_{k-j}'\right)\mathrm{e}^{-\boldsymbol{\Lambda}'j}\right\| > \eta\right) = 0, \tag{5.6}$$

because a conclusion of Billingsley (1968), Theorem 4.2, is then that as $n \to \infty$,

$$\sum_{j=0}^{\infty}\left(\widetilde{a}_n^{-1}\sum_{k=1}^{n-1}\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_{k-j}'\right)\mathrm{e}^{-\boldsymbol{\Lambda}'j} \Longrightarrow \sum_{j=0}^{\infty}\mathbf{S}_j^*\mathrm{e}^{-\boldsymbol{\Lambda}'j}. \tag{5.7}$$

For the proof of (5.6) let us define $\mathbf{K}_n := \mathbb{E}(\boldsymbol{\xi}_1\boldsymbol{\xi}_2'\mathbb{1}_{\{\|\boldsymbol{\xi}_1\boldsymbol{\xi}_2'\|\leq\widetilde{a}_n\}})$ and

$$
\begin{aligned}
J_1(m,n) &:= \sum_{j>m}\widetilde{a}_n^{-1}\sum_{k=1}^{n-1}\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_{k-j}'\mathbb{1}_{\{\|\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_{k-j}'\|>\widetilde{a}_n\}}\mathrm{e}^{-\boldsymbol{\Lambda}'j}, \\
J_2(m,n) &:= \sum_{j>m}\widetilde{a}_n^{-1}\sum_{k=1}^{n-1}(\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_{k-j}'\mathbb{1}_{\{\|\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_{k-j}'\|\leq\widetilde{a}_n\}} - \mathbf{K}_n)\mathrm{e}^{-\boldsymbol{\Lambda}'j}, \\
J_3(m,n) &:= \widetilde{a}_n^{-1}(n-1)\mathbf{K}_n\sum_{j>m}\mathrm{e}^{-\boldsymbol{\Lambda}'j}.
\end{aligned}
$$

Let $0 < \alpha \leq 1$ and $0 < \delta < \alpha$. Then

$$\|J_1(m,n)\|^\delta \leq C_1\widetilde{a}_n^{-\delta}\sum_{j>m}\sum_{k=1}^{n-1}\|\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_{k-j}'\|^\delta\mathbb{1}_{\{\|\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_{k-j}'\|>\widetilde{a}_n\}}\mathrm{e}^{-\lambda\delta j}$$

such that by Markov's inequality and Karamata's Theorem

$$\mathbb{P}(\|J_1(m,n)\| > \eta) \leq C_2 a_n^{-\delta}n\mathbb{E}(\|\boldsymbol{\xi}_1\boldsymbol{\xi}_2'\|^\delta\mathbb{1}_{\{\|\boldsymbol{\xi}_1\boldsymbol{\xi}_2'\|>\widetilde{a}_n\}})\sum_{j>m}\mathrm{e}^{-\lambda\delta j} \stackrel{n\to\infty}{\longrightarrow} C_3\sum_{j>m}\mathrm{e}^{-\lambda\delta j} \stackrel{m\to\infty}{\longrightarrow} 0. \tag{5.8}$$

Let $1 < \alpha < 2$. Then Markov's inequality and Karmata's Theorem give

$$\mathbb{P}(\|J_1(m,n)\| > \eta) \leq C_4\widetilde{a}_n^{-1}n\mathbb{E}(\|\boldsymbol{\xi}_1\boldsymbol{\xi}_2'\|\mathbb{1}_{\{\|\boldsymbol{\xi}_1\boldsymbol{\xi}_2'\|>\widetilde{a}_n\}})\sum_{j>m}\mathrm{e}^{-\lambda j} \stackrel{n\to\infty}{\longrightarrow} C_5\sum_{j>m}\mathrm{e}^{-\lambda j} \stackrel{m\to\infty}{\longrightarrow} 0. \tag{5.9}$$

Similarly we have also

$$\lim_{m\to\infty}\lim_{n\to\infty}\|J_3(m,n)\| = 0. \tag{5.10}$$

It remains to investigate $J_2(m,n)$. We assume now without loss of generality that $J_2(m,n)$ is one-dimensional otherwise we have to investigate $J_2(m,n)$ coordinatenwise. Note that by the Cauchy-Schwarz inequality

$$|\mathbb{E}((\xi_1\xi_2\mathbb{1}_{\{|\xi_1\xi_2|\leq\widetilde{a}_n\}} - K_n)(\xi_1\xi_3\mathbb{1}_{\{|\xi_1\xi_3|\leq\widetilde{a}_n\}} - K_n))| \leq \mathbb{E}(\xi_1^2\xi_2^2\mathbb{1}_{\{|\xi_1\xi_2|\leq\widetilde{a}_n\}}) =: \sigma_n^2.$$

Defining

$$\sigma_n(k+1, k-i, l+1, l-j) := \mathbb{E}((\xi_{k+1}\xi_{k-i}\mathbb{1}_{\{|\xi_{k+1}\xi_{k-i}|\leq\widetilde{a}_n\}} - K_n)(\xi_{l+1}\xi_{l-j}\mathbb{1}_{\{|\xi_{l+1}\xi_{l-j}|\leq\widetilde{a}_n\}} - K_n)),$$

we have by the independence of the sequence $(\xi_k)_{k\in\mathbb{N}}$ that

$$
\begin{aligned}
\mathrm{Var}(\widetilde{a}_n J_2(m,n)) &:= \sum_{k=1}^{n-1}\sum_{l=1}^{n-1}\sum_{i>m}\sum_{j>m}\sigma_n(k+1,k-i,l+1,l-j)\mathrm{e}^{-\lambda i}\mathrm{e}^{-\lambda j} \\
&= \sum_{k=1}^{n-1}\sum_{i>m}\sum_{j>m}\sigma_n(k+1,k-i,k+1,k-j)\mathrm{e}^{-\lambda i}\mathrm{e}^{-\lambda j} \\
&\quad + 2\sum_{k=m+3}^{n-1}\sum_{l=1}^{k-m-2}\sum_{j>m}\sigma_n(k+1,l+1,l+1,l-j)\mathrm{e}^{-\lambda(k-l-1)}\mathrm{e}^{-\lambda j} \\
&\quad + 2\sum_{k=1}^{n-1}\sum_{l=1}^{k-1}\sum_{i>m+k-l}\sigma_n(k+1,k-i,l+1,k-i)\mathrm{e}^{-\lambda i}\mathrm{e}^{-\lambda(l-k+i)} \\
&\leq n\sigma_n^2\left(\sum_{j>m}\mathrm{e}^{-\lambda j}\right)^2 + 2\sigma_n^2\left(\sum_{j>m}\mathrm{e}^{-\lambda j}\right)\sum_{k=1}^{n}\sum_{l=1}^{k}\mathrm{e}^{-\lambda l} + 2\sigma_n^2\left(\sum_{i>m}\mathrm{e}^{-\lambda i}\right)\sum_{k=1}^{n}\sum_{l=1}^{k-1}\mathrm{e}^{-\lambda l} \\
&\leq C_6 n\sigma_n^2\left(\sum_{j>m}\mathrm{e}^{-\lambda j}\right).
\end{aligned}
$$

Hence, Tschebyscheff's inequality and Karamata's Theorem result in

$$\lim_{m\to\infty}\lim_{n\to\infty}\mathbb{P}(\|J_2(m,n)\| > \eta) \leq \lim_{m\to\infty}\lim_{n\to\infty}C_6\widetilde{a}_n^{-2}n\sigma_n^2\left(\sum_{j>m}\mathrm{e}^{-\lambda j}\right) = 0. \tag{5.11}$$

Finally, (5.8)-(5.11) give (5.6) and thus, (5.7) holds, i.e. $\mathbf{S}_{5,n} \Longrightarrow \mathbf{S}_5$ as $n \to \infty$.

On the other hand, by Proposition 2.5 we have $(\mathbf{S}_{1,n}, \mathbf{S}_{2,n}, \mathbf{S}_{3,n}, \mathbf{S}_{4,n}) \Longrightarrow (\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, \mathbf{S}_4)$ as $n \to \infty$. The joint convergence and the independence of $(\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, \mathbf{S}_4)$ with $(\mathbf{S}_j^*)_{j\in\mathbb{N}_0}$ comes from the fact that we can straightforward extend Davis and Resnick (1986), Proposition 3.2, to the multivariate case

$$\sum_{k=1}^{\infty}\varepsilon_{(k/n, a_n^{-1}\boldsymbol{\xi}_k, \widetilde{a}_n^{-1}(\boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_k', \dots, \boldsymbol{\xi}_{k+1}\boldsymbol{\xi}_{k+1-m}'))} \Longrightarrow \sum_{k=1}^{\infty}\left[\varepsilon_{(t_k, j_k^{(1)}, \mathbf{0}_{d\times d})} + \varepsilon_{(t_k, \mathbf{0}_d, j_k^{(2)})}\right] \qquad \text{as } n \to \infty,$$

where $\sum_{k=1}^{\infty}\varepsilon_{(t_k, j_k^{(1)}, j_k^{(2)})} \sim \mathrm{PRM}(\mathrm{Leb} \times \mu_{\boldsymbol{\xi}} \times \mu_{\boldsymbol{\xi\xi}_m'})$ with $\mu_{\boldsymbol{\xi\xi}_m'}$ given in (2.3) and $\mu_{\boldsymbol{\xi}}$ in (2.4), and then follow the proof of Fasen (2010), Proposition 2.1, step by step. $\qquad\square$

5.2   Proofs of Section 3

Proof of Theorem 3.6. Note that $(\mathbf{Z}(k))_{k\in\mathbb{N}}$ has the MA representation

$$\mathbf{Z}(k) = \mathrm{e}^{-\boldsymbol{\Lambda}}\mathbf{Z}(k-1) + \boldsymbol{\xi}_k = \sum_{i=-\infty}^{k}\mathrm{e}^{-\boldsymbol{\Lambda}(k-i)}\boldsymbol{\xi}_i = \sum_{j=0}^{\infty}\mathrm{e}^{-\boldsymbol{\Lambda}j}\boldsymbol{\xi}_{k-j}$$

23

where

$$\boldsymbol{\xi}_{-k} = \int_{-k+1}^{-k} e^{-\boldsymbol{\Lambda}(-k-s)} \Sigma\, \mathbf{L}(ds) \quad \text{ for } k \in \mathbb{N}$$

with $(\mathbf{L}(-t))_{t\geq 0}$ an independent copy of $(\mathbf{L}(t))_{t\geq 0}$. Thus, we have

$$
\begin{aligned}
\mathbb{Z}'_{2,n}\mathbb{Z}_{1,n-1} &= \sum_{k=1}^{n-1} \mathbf{Z}(k+1)\mathbf{Z}(k)' = e^{-\boldsymbol{\Lambda}} \sum_{k=1}^{n-1} \mathbf{Z}(k)\mathbf{Z}(k)' + \sum_{k=1}^{n-1} \boldsymbol{\xi}_{k+1}\mathbf{Z}(k)' \\
&= e^{-\boldsymbol{\Lambda}} \mathbb{Z}'_{1,n-1}\mathbb{Z}_{1,n-1} + \sum_{j=0}^{\infty} \left[\sum_{k=1}^{n-1} \boldsymbol{\xi}_{k+1}\boldsymbol{\xi}'_{k-j} \right] e^{-\boldsymbol{\Lambda}'j}.
\end{aligned}
$$

Hence,

$$a_n^2 \widetilde{a}_n^{-1}(\widehat{e^{-\boldsymbol{\Lambda}_n}} - e^{-\boldsymbol{\Lambda}}) = \left(\sum_{j=0}^{\infty} \left(\widetilde{a}_n^{-1} \sum_{k=1}^{n-1} \boldsymbol{\xi}_{k+1}\boldsymbol{\xi}'_{k-j} \right) e^{-\boldsymbol{\Lambda}'j} \right) (a_n^{-2}\mathbb{Z}'_{1,n-1}\mathbb{Z}_{1,n-1})^{-1}.$$

An application of Theorem 2.6 and the continuous mapping theorem (accounting for $\mathbb{P}(\det(\mathbf{S}_3) = 0) = 0$) result in the proof. $\qquad\square$

Proof of Theorem 3.11. We will show that as $n \to \infty$,

$$na_n^{-2}\widehat{\boldsymbol{\Omega}}_n \Longrightarrow \mathbf{S}_2. \tag{5.12}$$

Then we obtain the statement again by an application of the continuous mapping theorem and Theorem 3.6.

To prove (5.12) note that as $n \to \infty$,

$$
\begin{aligned}
na_n^{-2}\left(\widehat{\boldsymbol{\Omega}}_n - n^{-1}\sum_{k=1}^{n} \boldsymbol{\xi}_k\boldsymbol{\xi}'_k \right) &= a_n^{-2}(\mathbb{Z}'_{2,n} - \widehat{e^{-\boldsymbol{\Lambda}_n}}\mathbb{Z}'_{1,n-1})(\mathbb{Z}'_{2,n} - \widehat{e^{-\boldsymbol{\Lambda}_n}}\mathbb{Z}'_{1,n-1})' \tag{5.13} \\
&\quad - a_n^{-2}(\mathbb{Z}'_{2,n} - e^{-\boldsymbol{\Lambda}}\mathbb{Z}'_{1,n-1})(\mathbb{Z}'_{2,n} - e^{-\boldsymbol{\Lambda}}\mathbb{Z}'_{1,n-1})' \Longrightarrow \mathbf{0}_{d\times d}
\end{aligned}
$$

by an expansion of the product, Theorem 2.6 and Theorem 3.4. Then the conclusion (5.12) follows from Proposition 2.5. $\qquad\square$

5.3 Proofs of Section 4

We start with the proofs in the infinite variance case of $\mathbf{L}_2$ in Theorem 4.2, Theorem 4.3 and Theorem 4.5, and afterwards we will present the proofs of Theorem 4.1 and Theorem 4.4 in the finite variance case of $\mathbf{L}_2$.

Proof of Theorem 4.2.
We write

$$
\begin{aligned}
\widehat{\mathbb{Z}}'_{2,n}\widehat{\mathbb{Z}}_{1,n-1} &= ((\mathbf{A} - \widehat{\mathbf{A}}_n)\mathbb{Y}'_{2,n} + \mathbb{Z}'_{2,n})((\mathbf{A} - \widehat{\mathbf{A}}_n)\mathbb{Y}'_{1,n-1} + \mathbb{Z}'_{1,n-1})' \tag{5.14} \\
&= (\mathbf{A} - \widehat{\mathbf{A}}_n)\mathbb{Y}'_{2,n}\mathbb{Y}_{1,n-1}(\mathbf{A} - \widehat{\mathbf{A}}_n)' + \mathbb{Z}'_{2,n}\mathbb{Y}_{1,n-1}(\mathbf{A} - \widehat{\mathbf{A}}_n)' \\
&\quad + (\mathbf{A} - \widehat{\mathbf{A}}_n)\mathbb{Y}'_{2,n}\mathbb{Z}_{1,n-1} + \mathbb{Z}'_{2,n}\mathbb{Z}_{1,n-1} \\
&=: \boldsymbol{\varepsilon}_n^{(1)} + \mathbb{Z}'_{2,n}\mathbb{Z}_{1,n-1}.
\end{aligned}
$$

From Fasen (2010) we know that

$$\left(a_n^{-1}b_n^{-1}\mathbb{Z}_n'\mathbb{Y}_n, n^{-1}b_n^{-2}\mathbb{Y}_n'\mathbb{Y}_n, a_n^{-2}\mathbb{Z}_{2,n}'\mathbb{Z}_{1,n-1}, nb_na_n^{-1}(\widehat{\mathbf{A}}_n - \mathbf{A}) \right)$$

$$\Longrightarrow \left(\widetilde{\mathbf{S}}_2(1)\widetilde{\mathbf{S}}_1(1)' - \int_0^1 \widetilde{\mathbf{S}}_2(s)\,\widetilde{\mathbf{S}}_1(\mathrm{d}s)', \int_0^1 \widetilde{\mathbf{S}}_1(s)\widetilde{\mathbf{S}}_1(s)'\mathrm{d}s, \mathrm{e}^{-\mathbf{\Lambda}}\mathbf{S}_3, \right.$$

$$\left. \widetilde{\mathbf{S}}_2(1)\widetilde{\mathbf{S}}_1(1)' - \int_0^1 \widetilde{\mathbf{S}}_2(s)\,\widetilde{\mathbf{S}}_1(\mathrm{d}s)' \left(\int_0^1 \widetilde{\mathbf{S}}_1(s)\widetilde{\mathbf{S}}_1(s)'\mathrm{d}s \right)^{-1} \right), \qquad (5.15)$$

where $(\widetilde{\mathbf{S}}_1(t))_{t\geq 0}$ is an h-dimensional α^*-stable Lévy process and $(\widetilde{\mathbf{S}}_2(t))_{t\geq 0}$ is an d-dimensional α-stable Lévy process, respectively. Thus,

$$na_n^{-2}\varepsilon_n^{(1)} \Longrightarrow \varepsilon^{(1)} \quad \text{as } n \to \infty, \qquad (5.16)$$

where $\varepsilon^{(1)}$ is some random matrix in $\mathbb{R}^{d\times d}$. On the other hand, by Proposition 2.5 we have as $n \to \infty$,

$$a_n^{-2}\mathbb{Z}_{2,n}'\mathbb{Z}_{1,n-1} \Longrightarrow \mathrm{e}^{-\mathbf{\Lambda}}\mathbf{S}_3 \qquad (5.17)$$

with $\mathbb{P}(\det(\mathrm{e}^{-\mathbf{\Lambda}}\mathbf{S}_3) = 0) = 0$ such that as $n \to \infty$,

$$(\widehat{\mathbb{Z}}_{2,n}'\widehat{\mathbb{Z}}_{1,n-1})(\mathbb{Z}_{2,n}'\mathbb{Z}_{1,n-1})^{-1} \xrightarrow{\mathbb{P}} \mathbf{I}_{d\times d}. \qquad (5.18)$$

The same arguments give as $n \to \infty$,

$$(\widehat{\mathbb{Z}}_{1,n-1}'\widehat{\mathbb{Z}}_{1,n-1})(\mathbb{Z}_{1,n-1}'\mathbb{Z}_{1,n-1})^{-1} \xrightarrow{\mathbb{P}} \mathbf{I}_{d\times d}. \qquad (5.19)$$

However, a conclusion of Proposition 2.5 is that as $n \to \infty$,

$$\mathbb{Z}_{2,n}'\mathbb{Z}_{1,n-1}(\mathbb{Z}_{1,n-1}'\mathbb{Z}_{1,n-1})^{-1} \Longrightarrow \mathrm{e}^{-\mathbf{\Lambda}}\mathbf{S}_3\mathbf{S}_3^{-1} = \mathrm{e}^{-\mathbf{\Lambda}} \qquad (5.20)$$

which finishes with (5.18) and (5.19) the proof. $\qquad\square$

Proof of Theorem 4.3.

We use the notation of the proof of Theorem 4.2 and write similarly $\widehat{\mathbb{Z}}_{1,n-1}'\widehat{\mathbb{Z}}_{1,n-1} - \mathbb{Z}_{1,n-1}'\mathbb{Z}_{1,n-1} =: \varepsilon_n^{(2)}$. Hence, we obtain

$$\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}} - \widetilde{\mathrm{e}^{-\mathbf{\Lambda}_n}} = (\mathbb{Z}_{2,n}'\mathbb{Z}_{1,n-1})(\mathbb{Z}_{1,n-1}'\mathbb{Z}_{1,n-1})^{-1}\varepsilon_n^{(2)}(\mathbb{Z}_{1,n-1}'\mathbb{Z}_{1,n-1} + \varepsilon_n^{(2)})^{-1}$$

$$- \varepsilon_n^{(1)}(\mathbb{Z}_{1,n-1}'\mathbb{Z}_{1,n-1} + \varepsilon_n^{(2)})^{-1} \qquad (5.21)$$

$$= (\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}} - \mathrm{e}^{-\mathbf{\Lambda}})\varepsilon_n^{(2)}(\mathbb{Z}_{1,n-1}'\mathbb{Z}_{1,n-1} + \varepsilon_n^{(2)})^{-1} + (\mathrm{e}^{-\mathbf{\Lambda}}\varepsilon_n^{(2)} - \varepsilon_n^{(1)})(\mathbb{Z}_{1,n-1}'\mathbb{Z}_{1,n-1} + \varepsilon_n^{(2)})^{-1}.$$

On the one hand, $na_n^{-2}\varepsilon_n^{(2)} \Longrightarrow \varepsilon^{(2)}$ as $n \to \infty$, (5.17) and Theorem 3.6 give as $n \to \infty$,

$$a_n^2\widetilde{a}_n^{-1}(\widehat{\mathrm{e}^{-\mathbf{\Lambda}_n}} - \mathrm{e}^{-\mathbf{\Lambda}})\varepsilon_n^{(2)}(\mathbb{Z}_{1,n-1}'\mathbb{Z}_{1,n-1} + \varepsilon_n^{(2)})^{-1} \Longrightarrow \mathbf{0}_{d\times d}. \qquad (5.22)$$

Finally, for $1 < \alpha < 2$ we have $a_n^2(\widetilde{a}_n n)^{-1} \longrightarrow 0$ and

$$\frac{a_n^2}{\widetilde{a}_n}(\mathrm{e}^{-\mathbf{\Lambda}}\varepsilon_n^{(2)} - \varepsilon_n^{(1)})(\mathbb{Z}_{1,n-1}'\mathbb{Z}_{1,n-1} + \varepsilon_n^{(2)})^{-1} \qquad (5.23)$$

$$= \frac{a_n^2}{\widetilde{a}_n n}\left(na_n^{-2}\mathrm{e}^{-\mathbf{\Lambda}}\varepsilon_n^{(2)} - na_n^{-2}\varepsilon_n^{(1)} \right)\left(a_n^{-2}\mathbb{Z}_{1,n-1}'\mathbb{Z}_{1,n-1} + a_n^{-2}\varepsilon_n^{(2)} \right)^{-1} \Longrightarrow \mathbf{0}_{d\times d}$$

as $n \to \infty$ which results with (5.21), (5.22) and Theorem 3.6 in the statement. $\square$

Proof of Theorem 4.5.

Note, that as $n \to \infty$,

$$
\begin{aligned}
na_n^{-2}(\widetilde{\boldsymbol{\Omega}}_n - \widehat{\boldsymbol{\Omega}}_n) = &\ a_n^{-2}(\widehat{\mathbb{Z}}'_{2,n} - \widetilde{\mathrm{e}^{-\boldsymbol{\Lambda}_n}}\widehat{\mathbb{Z}}'_{1,n-1})(\widehat{\mathbb{Z}}'_{2,n} - \widetilde{\mathrm{e}^{-\boldsymbol{\Lambda}_n}}\widehat{\mathbb{Z}}'_{1,n-1})' \\
&\ -a_n^{-2}(\mathbb{Z}'_{2,n} - \mathrm{e}^{-\boldsymbol{\Lambda}}\mathbb{Z}'_{1,n-1})(\mathbb{Z}'_{2,n} - \mathrm{e}^{-\boldsymbol{\Lambda}}\mathbb{Z}'_{1,n-1})' \Longrightarrow \mathbf{0}_{d\times d}
\end{aligned} \tag{5.24}
$$

by an expansion of the product, Proposition 2.5, Theorem 3.4, (5.18) and (5.19). Thus, (5.12) and (5.24) leads to

$$
na_n^{-2}\widetilde{\boldsymbol{\Omega}}_n \Longrightarrow \mathbf{S}_2 \quad \text{as } n \to \infty.
$$

Finally, the asymptotics of the t- and the Wald-statistic follow by Theorem 4.3 and the continuous mapping theorem. $\square$

Proof of Theorem 4.1 and Theorem 4.4.

The proof goes step by step as the proof of Theorem 4.3 using $a_n = \widetilde{a}_n = \sqrt{n}$, (5.15), Proposition 3.2 and the conclusions of Brockwell and Davis (2006), Theorem 11.2.1 that $\frac{1}{n}\sum_{k=1}^n \boldsymbol{\xi}_k\boldsymbol{\xi}'_k \overset{\mathbb{P}}{\longrightarrow} \Sigma_{\boldsymbol{\xi}}$, $\widehat{\boldsymbol{\Omega}}_n \overset{\mathbb{P}}{\longrightarrow} \Sigma_{\boldsymbol{\xi}}$, $\frac{1}{n}\mathbb{Z}'_{1,n-1}\mathbb{Z}_{1,n-1} \overset{\mathbb{P}}{\longrightarrow} \Sigma_{\mathbf{Z}}$ and $\frac{1}{n}\mathbb{Z}'_{2,n}\mathbb{Z}_{1,n-1} \overset{\mathbb{P}}{\longrightarrow} \mathrm{e}^{-\boldsymbol{\Lambda}}\Sigma_{\mathbf{Z}}$ as $n \to \infty$. $\square$

Acknowledgment

References

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

Barndorff-Nielsen, O.E., Nicolata, E. and Shephard, N. (2002). Some recent developments in stochastic volatility modelling. *Quant. Finance*, **2**, 11–23.

Barndorff-Nielsen, O.E. and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics (with discussion). *J. Roy. Statist. Soc. Ser. B*, **63**, 167–241.

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

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

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

Benth, F.E., Koekebakker, S. and Zakamouline, V. (2010). The CARMA interest rate model. Available at SSRN: http://ssrn.com/abstract=1138632.

Bergstrom, A.R. (1990). *Continuous Time Econometric Modelling*. Oxford University Press, Oxford.

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

Brockwell, P.J. and Davis, R.A. (2006). *Time Series: Theory and Method.* 2nd edn. Springer, New York.

Brockwell, P., Davis, R.A. and Yang, V. (2007). Estimation for non-negative Levy-driven Ornstein-Uhlenbeck processes. *J. Appl. Prob.*, **44**, 977–989.

Calder, M. and Davis, R.A. (1998). Inference for linear processes with stable noise. In: R. Adler, R. Feldman and M. Taqqu (Eds.), *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, pages 159–176. Birkhäuser, Boston.

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

Cluver, W.J. (1966). On the existence and uniqueness of the real logarithm of a matrix. *Proc. American Math. Soc.*, **17**, 1146–1151.

Davis, R., Knight, K. and Liu, J. (1992). *M*-estimation for autoregressions with infinite variance. *Stoch. Process. Appl.*, **40**, 145–180.

Davis, R., Marengo, J. and Resnick, S. (1985). Extremal properties of a class of multivariate moving averages. In: *Proceeding of the 45th session of the Internationa Statistical Institute*, vol. 51 of *Bull. Inst. Internat. Statist.*, pages 185–192.

Davis, R. and Resnick, S. (1986). Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist.*, **14**, 533–558.

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

Eberlein, E. (2009). Jump-type Lévy processes. In: T.G. Andersen, R.A. Davis, J.P. Kreiß and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pages 439–455. Springer, Berlin.

Ekström, E., Lindberg, C. and Tysk, J. (2009). Optimal liquidation of pairs trade. Available at http://www.math.uu.se/ johant/preprints.

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

Engle, R.F. and Granger, C.W.J. (1991). *Long-Run Economic Relations: Readings in Cointegration.* Oxford University Press.

Engle, R.F. and White, H. (1999). *Cointegration, Causality and Forecasting: A Festschrift in Honour of Clive W.J. Granger.* Oxford University Press.

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

Fasen, V. (2010). Time series regression on integrated continuous-time processes with heavy and light tails. Available at http://www-m4.ma.tum.de/Papers/.

Fasen, V., Klüppelberg, C. and Lindner, A. (2006). Extremal behavior of stochastic volatility models. In: A. Shiryaev, M.d.R. Grossinho, P.E. Oliviera and M.L. Esquivel (Eds.), *Stochastic Finance*, pages 107–155. Springer, New York.

Griffin, J.E. and Steel, M.F.J. (2006). Inference with non-Gaussian Ornstein-Uhlenbeck processes for stochastic volatility modelling. *J. Econometrics*, **134**, 606–644.

Hannan, E.J. (1970). *Multiple Time Series.* John Wiley & Sons Inc.

Hu, Y. and Long, H. (2007). Parameter estimation for Ornstein-Uhlenbeck processes driven by α-stable Lévy motions. *Commun. Stoch. Anal.*, **1**, 175–192.

Hu, Y. and Long, H. (2009). Least squares estimator for Ornstein-Uhlenbeck processes driven by α-stable motions. *Stoch. Proc. Appl.*, pages 2465–2480.

Hult, P. and Lindskog, H. (2007). Extremal behaviour of stochastic integrals driven by regularly varying Lévy processes. *Ann. Probab.*, **35**, 309–339.

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

Jongbloed, G., van der Meulen, F.H. and van der Vaart, A.W. (2005). Nonparametric inference for Lévy driven Ornstein-Uhlenbeck processes. *Bernoulli*, **11**, 759–791.

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

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

Lütkepohl, H. and Krätzig, M. (2004). *Applied Time Series Econometrics.* Cambridge University Press.

Mandelbrot, B. and Taylor, H.M. (1967). On the distribution of stock price differences. *Oper. Res.*, **15**, 1057–1062.

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

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

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

Meerschaert, M.M. and Scheffler, H.P. (2001). Sample cross-correlations for moving averages with regularly varying tails. *J. Time Ser. Anal.*, **22**, 481–492.

Phillips, P.C.B. (1974). The estimation of some continuous time models. *Econometrica*, **42**, 803–823.

Pigorsch, C. and Stelzer, R. (2009). A multivariate Ornstein-Uhlenbeck type stochastic volatility model. Available at http://www-m4.ma.tum.de/Papers/.

Pratt, J. (1960). On interchanging limits and integrals. *Ann. Math. Statist.*, **31**, 74–77.

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

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

Rachev, S.T., Kim, J. and Mittnik, S. (1999). Stable Paretian econometrics. *Math. Sci.*, **24**, 24–25 and 113–127.

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

Roberts, G.O., Papaspiliopoulos, O. and Dellaportas, P. (2004). Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatility processes. *J. Roy. Statist. Soc. Ser. B*, **66**, 369–393.

Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes*. Chapman and Hall, London.

Sato, K. (2002). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Oxford.

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

Spiliopoulos, K. (2010). Method of moment estimation of Ornstein-Uhlenbeck processes driven by general Lévy processes. Available at http://arxiv.org/abs/0807.2832.

Taufer, E. and Leonenko, N. (2009). Characteristic function estimation of non-Gaussian Ornstein-Uhlenbeck processes. *J. Statist. Plann. Inference*, **139**, 3050–3063.

Todorov, V. (2009). Estimation of continuous-time stochastic volatility models with jumps using high-frequency data. *J. Econometrics*, **148**, 131–148.