

# Minimax rates for the covariance estimation of multi-dimensional Lévy processes with high-frequency data

Katerina Papagiannouli

*Institut für Mathematik  
Humboldt-Universität zu Berlin  
Unter den Linden 6  
10099 Berlin  
Germany  
e-mail: [papagai@hu-berlin.de](mailto:papagai@hu-berlin.de)*

**Abstract:** This article studies nonparametric methods to estimate the co-integrated volatility of multi-dimensional Lévy processes with high frequency data. We construct a spectral estimator for the co-integrated volatility and prove minimax rates for an appropriate bounded nonparametric class of Lévy processes. Given  $n$  observations of increments over intervals of length  $1/n$ , the rates of convergence are  $1/\sqrt{n}$  if  $r \leq 1$  and  $(n \log n)^{(r-2)/2}$  if  $r > 1$ , where  $r$  is the co-jump activity index and corresponds to the intensity of dependent jumps. These rates are optimal in a minimax sense. We bound the co-jump activity index from below by the harmonic mean of the jump activity indices of the components. Finally, we assess the efficiency of our estimator by comparing it with estimators in the existing literature.

**MSC2020 subject classifications:** Primary 60G51, 62G05; secondary 62G10, 62C20, 60J75.

**Keywords and phrases:** Co-jumps, infinite variation, co-integrated volatility, high-frequency data.

Received September 2019.

## 1. Introduction

Lévy processes are the main building blocks for stochastic continuous-time jump models. Whenever the modeling of a stochastic process in finance requires the inclusion of jumps, Lévy processes are those to be considered. They play an instrumental role, for example, in the modeling of financial data, see Carr et al. (2002); Barndorff-Nielsen and Shephard (2004, 2006); Wu (2007); Eberlein and Papapantoleon (2005); Geman (2002).

Consequently, the large amount of applications has given rise to a great demand for statistical methods in the study of Lévy processes, especially nonparametric methods. Using nonparametric methods relaxes any dependency on the model. The problem of estimating the characteristics of a Lévy process has received considerable attention over the past decade. Starting with the work by Belomestny and Reiß (2006), a number of articles have considered nonparametric estimation methods for Lévy processes. Therefore, one important task is to provide estimation methods for the characteristics of a Lévy process.

Moreover, statistical methods require the nature of the observation schemes to be classified as high frequency or low frequency; here, we focus on a high frequency setting. If we can assume high-frequency observations for a Lévy process, we can discretize a natural estimator based on continuous-time observations, where the jumps and the diffusion part are observed directly. In recent years, the literature on this subject has grown extensively, see Figueroa-Lopez and Houdré (2004); Todorov and Tauchen (2011); Coca (2018); Comte and Genon-Catalot (2009); Neumann and Reiß (2009). We now have vast amounts of data on the prices of various assets, exchange rates, and so on, typically *tick data* which are recorded at every transaction time.

Much has been written on the estimation of Lévy density using nonparametric techniques, for instance Nickl et al. (2016), Duval and Mariucci (2017), Comte et al. (2014) and the references therein. However, the seminal work of Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shephard (2002), proposed realized variance as an estimator for quadratic variation. In the presence of jumps, a well-known theoretical result proves that the realized variation converges in probability to the global quadratic variation as the time between two consecutive observations tends towards zero. This result motivated estimators that filter out the jumps, like Bipower Variation by Barndorff-Nielsen and Shephard (2004) and Truncated Realized Variation by Mancini (2009).

In the multivariate context, the recovery of co-integrated volatility (also known as covariance) becomes more complicated. Among various prominent works see Christensen et al. (2013), Bibinger and Vetter (2015), Bibinger and Winkelmann (2015). For models incorporating jumps, the realized covariation converges in probability to the global quadratic variation containing the co-jumps. Co-jumps refer to the case when the underlying processes jump at the same time with the same direction. This raises the question how we can assess the dependence structure among the jump components. We find the answer in the Lévy copula, a subject studied by Tankov (2004) in his PhD thesis. The interested reader should refer to Cont and Tankov (2004) and Kallsen and Tankov (2006). The Lévy Copula is the basic tool for the class of multidimensional Lévy processes. To mention only the few approaches which are close to our focus on Lévy processes we refer to Mancini (2017), Christensen et al. (2013), Martin and Vetter (2017), Bibinger et al. (2014), Jacod and Reiß (2014), Bücher and Vetter (2013) and Belomestny and Trabs (2018).

Our aim in the present work is to provide minimax rates of convergence for the estimation of co-integrated volatility when the underlying process belongs to a certain class of multi-dimensional Lévy processes. Many features of co-integrated volatility have already been studied, such as asynchronous observations, microstructure noise, and allowing for dependency among the jumps components. Whereas most of the aforementioned results prove central limit theorems for their estimators, at least to the best of our knowledge no work has dealt with optimal rates of convergence in the minimax sense. This work serves to fill this gap. Jacod and Reiß (2014) proposed a spectral estimator for integrated volatility achieving minimax rates. In the present work, we generalize their work on finite dimensions. By virtue of simplicity, we will concentrate

primarily on a two-dimensional regime, but extensions to the general multi-dimensional setting are straightforward to obtain as well.

For this purpose let us define, for a two-dimensional Lévy process  $X$ , the Blumenthal-Gettoor index  $r^*$ :

$$B(r) = \int_{\mathbb{R}^2} (1 \wedge \|\mathbf{x}\|^r) F(d\mathbf{x}), \quad I = \{0 < r \leq 2 : B(r) < \infty\}, \quad r^* = \inf I \tag{1.1}$$

where  $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$  is the size of the jump components,  $\|\cdot\|$  is the Euclidean norm in  $\mathbb{R}^2$  and  $F$  is the Lévy measure.  $B(r)$  is not specifically interesting, but the  $BG$ -index gives us the infimum number  $r$  for which  $B(r)$  is finite. This index is a very important number for Lévy processes, because using this index we can infer the behavior of small jump components around 0. When we have a two-dimensional Lévy process we have either independent jumps (i.e. disjoint or jumps in the axes) or dependent jumps(i.e. co-jumps or joint jumps). In the present work we focus on the case of co-jumps, when the two marginals jump at the same time in the same direction. The index  $r^*$  gives us information about the amount of disjoint and joint small jumps around 0. The behavior of co-jumps around 0 is described by

$$\int_{\mathbb{R}^2} (1 \wedge |x_1 x_2|^{r/2}) F(dx_1, dx_2) < \infty. \tag{1.2}$$

Here, we are interested in investigating the optimal rates for the estimation of co-integrated volatility when the model falls in a class of two-dimensional Lévy processes, in case the jump components are either of finite or infinite variation and satisfy (1.2). Let  $X = (X^{(1)}, X^{(2)})$  and  $r_1, r_2$  be the indices of jump activity for the small jump components of each process  $X^{(1)}, X^{(2)}$  respectively. We find that  $r$ , the activity index of co-jumps, is bounded from below by the harmonic mean of  $r_1, r_2$  even in the case of infinite variation jumps. This was not known up to now. Under this assumption for co-jumps we show that our spectral estimate for co-integrated volatility converges at a rate  $(n \log n)^{\frac{(r-2)}{2}}$  if  $r > 1$  and  $\frac{1}{\sqrt{n}}$  if  $r \leq 1$ .

Assuming a two-dimensional Itô semimartingale, Mancini (2017) proposed a truncated covariance estimator to estimate co-integrated volatility at the rate  $\frac{1}{\sqrt{n}}$  when  $r_1$  is small and  $r_2$  is close to 1,  $n^{-\frac{1}{2}(1+\frac{r_2}{r_1}-r_2)}$  when  $r_1, r_2$  are much bigger than 1 and close to 2,  $n^{\frac{(r_2}{2}-1)}$  when  $r_1$  is small and  $r_2$  is much bigger than  $r_1$  or in case of independent small jump components. However, these rates are sub-optimal when a two-dimensional Lévy process satisfies (1.2).

Let us describe the outline of this paper. In Section 2 we state the underlying model. In Section 3 we give the assumptions to be satisfied in order to prove the minimax rates. In Section 4 we construct our spectral estimator and state the results of this work. Section 5 gives the insight behind the co-jump activity index. In Section 6 we prove the upper bound for the family of our estimators. In Section 7 we present the proof of lower bound in a *minimax* sense. We provide some comparison of our estimator with existent estimators in the literature in Section 8. In the last section we provide a simulation study.

## 2. The underlying model

We assume equidistant discrete observation times  $i\Delta_n, i = 0, \dots, n$  for a mesh  $\Delta_n \rightarrow 0$ . Here, we use as a mesh  $\Delta_n = \frac{1}{n}$  and  $n \rightarrow \infty$ . Regarding the time horizon of the process, it is observed on a finite time span  $[0, 1]$ . Let  $\mathbf{X} = (X^{(1)}, X^{(2)})^\top$  be a two-dimensional Lévy process with Lévy-Itô decomposition as

$$\mathbf{X}_t = \mathbf{b}t + \mathbf{W}_t + \int_0^t \int_{\|\mathbf{x}\| \leq 1} \mathbf{x}(\mu - \tilde{\mu})(ds, d\mathbf{x}) + \int_0^t \int_{\|\mathbf{x}\| > 1} \mathbf{x}\mu(ds, d\mathbf{x}). \quad (2.1)$$

Unless stated otherwise, from now on  $\mathbf{b}$  is a drift vector in  $\mathbb{R}^2$ ,  $\mathbf{W} = (W^{(1)}, W^{(2)})$  denotes a bivariate Brownian motion with covariance matrix  $\Sigma\Sigma^\top$ , and  $\mu, \tilde{\mu}$  are the jump measure and its compensator, respectively. The compensator takes the form  $\tilde{\mu} = dsF(d\mathbf{x})$ , where  $F$  is the Lévy measure of  $\mathbf{X}$ .

Due to the independence of the continuous part and the discontinuous (jump part) of a Lévy process, the analysis of  $\mathbf{X}$  canonically splits into the inference on the covariance matrix and the inference on the jump measure  $F$ . Our focus on this paper is to investigate an estimator for the co-integrated volatility of  $\mathbf{X}$ .

We assume a filtered space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  supporting two independent standard Brownian motions  $W^{(1)}, W^{(3)}$  and two Poisson random measures  $\mu^{(j)}$  for  $j = 1, 2$  on  $\mathbb{R}^2 \times [0, 1]$ . Recall that  $W^{(1)}, W^{(2)}$  are correlated with  $d\langle W^{(1)}, W^{(2)} \rangle_t = \rho dt$ , where  $\rho$  is a constant on  $[-1, 1]$ . We construct  $W^{(2)}$  as a linear combination of the two independent Brownian motions so  $W_t^{(2)} = \rho W_t^{(1)} + \sqrt{1 - \rho^2} W_t^{(3)}$ . Next we calculate the variances and covariance of  $W^{(1)}, W^{(2)}$ , we see that the following holds

$$\begin{aligned} \text{Var}(W_t^{(1)}) &= \langle W_t^{(1)}, W_t^{(1)} \rangle = t \\ \text{Var}(W_t^{(2)}) &= \langle W_t^{(2)}, W_t^{(2)} \rangle = \rho^2 t + (1 - \rho^2)t = t. \end{aligned}$$

For the covariance we obtain

$$\text{Cov}(W_t^{(1)}, W_t^{(2)}) = \langle W_t^{(1)}, W_t^{(2)} \rangle = \rho \langle W_t^{(1)}, W_t^{(1)} \rangle + \sqrt{1 - \rho^2} \langle W_t^{(1)}, W_t^{(3)} \rangle = \rho t;$$

the last equality holds because of  $W^{(1)}, W^{(3)}$  being independent. So, without loss of generality we assume that

$$\Sigma = \begin{pmatrix} \sigma^{(1)} & 0 \\ \rho\sigma^{(2)} & \sqrt{1 - \rho^2}\sigma^{(2)} \end{pmatrix} \quad \text{so that} \quad \Sigma\Sigma^\top = \begin{pmatrix} (\sigma^{(1)})^2 & \rho\sigma^{(1)}\sigma^{(2)} \\ \rho\sigma^{(1)}\sigma^{(2)} & (\sigma^{(2)})^2 \end{pmatrix}$$

where  $\sigma^{(i)}, i = 1, 2$  are deterministic. Therefore, the global quadratic variation of  $\mathbf{X}$  is given by:

$$\langle X_t^{(1)}, X_t^{(2)} \rangle = \int_0^t \rho\sigma^{(1)}\sigma^{(2)} ds + \sum_{s \leq t} \Delta X_s^{(1)} \Delta X_s^{(2)} \quad (2.2)$$

where the first term is the co-integrated volatility and the second term is the sum of products of simultaneous jumps (called co-jumps). Our target of inference, the co-integrated volatility at time  $0 \leq t \leq 1$ , is

$$C_t^{12} = \int_0^t \rho \sigma^{(1)} \sigma^{(2)} ds. \tag{2.3}$$

### 3. Assumptions

To derive an estimator for co-integrated volatility and then prove a minimax bound for this estimator, we need to establish some assumptions regarding the behavior of small jumps and the class of our estimator. In particular, our setup is intrinsically nonparametric and related to the properties of the observed path. We use the following notation for a matrix:  $\|\cdot\|_\infty$  is the maximum absolute row sum of the matrix, (i.e. the  $\infty$ -norm).

*Assumption (1-M).* The  $\infty$ -norm of the covariance matrix is assumed to be bounded, i.e.  $\|\Sigma \Sigma^\top\|_\infty \leq M$ .

*Assumption (2-M).*  $\int_{\mathbb{R}^2} (1 \wedge |x_1 x_2|^{r/2}) F(dx_1, dx_2) \leq M$ , where  $r \in [0, 2)$ .

Notice that Assumption (2-M) follows from the classical condition to control the activity of small jumps in two dimensions. Through a trivial calculation

$$\begin{aligned} \int_{\mathbb{R}^2} (1 \wedge |x_1 x_2|^{r/2}) F(dx_1, dx_2) &\leq \int_{\mathbb{R}^2} (1 \wedge |x_1^2 + x_2^2|^{r/2}) F(dx_1, dx_2) \\ &= \int_{\mathbb{R}^2} (1 \wedge \|\mathbf{x}\|^r) F(d\mathbf{x}). \end{aligned}$$

By using this unconventional Assumption (2-M), we relax the classical condition for small jumps in two dimensions and strengthen our results, since we considered the case of dependent jumps.

Assumption (2-M) concerns the behavior of jump components with size smaller or equal to one. By this assumption we consider the problem of controlling the activity of *co-jumps*, i.e. joint jumps. Below, a co-jump indicates that both components jump at a time  $t$  but their jump sizes may *not* be the same. Ultimately, we are asking whether the small jump components are of finite or infinite variation. This question concerns the behavior of the compensator  $F$ , the Lévy measure near 0. The major difficulties here come from the possibly erratic behavior of  $F$  near 0 and the possible dependence between the jump components. In Section 5 we describe in more detail the dependence structure of the jump components and the co-jump activity index  $r$ .

The Blumenthal-Gettoor (BG) index allows us to classify the processes from least active to most active, according to the above description. We denote by  $r^*$  the BG index for a two-dimensional Lévy process which satisfies:

$$r^* = \inf \left\{ r \in (0, 2] : \int_{\mathbb{R}^2} (1 \wedge \|\mathbf{x}\|^r) F(d\mathbf{x}) < \infty \right\}. \tag{3.1}$$

Note that a stable Lévy process of index  $\beta \in (0, 2)$  satisfies

$$\int_{\mathbb{R}^2} (1 \wedge \|\mathbf{x}\|^r) F(d\mathbf{x}) < \infty$$

for all  $r > \beta$ , but not for  $r \leq \beta$ . The BG index of a  $\beta$ -stable is exactly  $\beta$ .

It should be noted that the problem of BG index estimation from discrete observations of a Lévy process has drawn much attention in the literature. In the case of high-frequency data, Aït-Sahalia and Jacod (2011) studied the problem of estimating the jump activity index that is defined for any Itô semimartingale. A consistent estimator for the BG index based on one-dimensional Lévy processes with low-frequency data was obtained in Belomestny (2010). The interested reader may refer to Belomestny and Reiß (2015), Section 7, for a detailed review of these results. An extension to time-changed Lévy processes can be found in Belomestny and Panov (2013a,b).

Now we will test the Assumption (2-M) about the boundedness of small co-jumps with some trivial examples. Despite its simple nature, the following example offers significant insight and intuitive understanding into co-jumps with infinite variation.

**Example 3.1.** Suppose we have independent jumps in the coordinates and  $F(d\mathbf{x})$  is a Lévy measure on  $\mathbb{R}^2$  and  $\mathbf{x}$  is a vector in  $\mathbb{R}^2$ . Then,  $\text{supp}(F) \subseteq \{\mathbb{R} \times \{0\} \cup \{0\} \times \mathbb{R}\}$  which means that

$$\begin{aligned} \int_{\mathbb{R}^2} (1 \wedge \|\mathbf{x}\|^r) F(d\mathbf{x}) &= \int_{\mathbb{R}^2} \left(1 \wedge |x_1^2 + x_2^2|^{r/2}\right) F(dx_1, dx_2) \\ &= \int_{\mathbb{R}} (1 \wedge |x_1|^r) F_1(dx_1) \\ &\quad + \int_{\mathbb{R}} (1 \wedge |x_2|^r) F_2(dx_2) < \infty, \end{aligned}$$

if the marginals of a two-dimensional Lévy process are finite in the one dimensional case. See the assumption section in Jacod and Reiß (2014).

In this example, we notice that

$$\int_{\mathbb{R}^2} \left(1 \wedge |x_1 x_2|^{r/2}\right) (\mathbb{1}_{\{x_1=0\}} + \mathbb{1}_{\{x_2=0\}}) F(dx_1, dx_2) = 0, \quad (3.2)$$

since the integrand is always equal to zero. This example shows us something more: Whenever we have independent jumps, no matter the choice of  $F$ , we can always find a control for the activity of small jumps even if we have jumps of infinite variation.

**Example 3.2.** Suppose we have independent jump size distributions and  $x_1, x_2 \in B_1(0)$ , where  $B_1(0)$  is the unit ball in  $\mathbb{R}^2$ . Therefore, for  $F_1$  and  $F_2$  are not

of zero measure,

$$\begin{aligned} \int_{B_1(0)} |x_1 x_2|^{r/2} F(dx_1, dx_2) &= \int_{B_1(0)} |x_1 x_2|^{r/2} F_1(dx_1) F_2(dx_2) \\ &\leq \int_{-1}^1 \left( \int_{-1}^1 |x_1|^{r/2} F_1(dx_1) \right) |x_2|^{r/2} F_2(dx_2) < \infty \end{aligned}$$

if and only if  $\int_{-1}^1 |x_i|^{r/2} F_i(dx_i) < \infty$  for  $i = 1, 2$ . However, we know that this inequality, for the marginal distributions, holds because of the results in Jacod and Reiß (2014). As a result, the Assumption (2-M) holds.

#### 4. Theoretical results

We use standard notation for asymptotic quantities like  $X_n = O_{\mathbb{P}}(w_n)$  if  $(X_n/w_n)_{n \geq 1}$  is stochastically bounded (i.e. bounded in probability or tight). We are in a nonparametric setting in which the process  $\mathbf{X}$  belongs to the class  $\mathcal{L}_M^r$ . Let us now define this class.

**Definition 4.1.** For  $M > 0$  and  $r \in [0, 2)$ , we define the class  $\mathcal{L}_M^r$ , the set of all Lévy processes, satisfying

$$\|C\|_{\infty} + \int_{\mathbb{R}^2} (1 \wedge |x_1 x_2|^{r/2}) F(dx_1, dx_2) \leq M, \tag{4.1}$$

where  $C = \begin{pmatrix} C^{11} & C^{12} \\ C^{21} & C^{22} \end{pmatrix}$  is the covariance matrix and  $\|C\|_{\infty} = \max\{C^{11} + C^{12}, C^{21} + C^{22}\}$ .

We adapt an estimator proposed by Jacod and Reiß (2014). Specifically, we let  $\mathbf{X}$  be a two-dimensional Lévy process with characteristic triplet  $(\mathbf{b}, C, F)$ . Let us remember that we are in a high-frequency setting and the consecutive time between two observations is  $\frac{1}{n}$ . The characteristic function of  $\mathbf{X}_{1/n}$  is given by:

$$\begin{aligned} \phi_n(\mathbf{u}_n) &= \exp \left\{ \frac{1}{n} \left( i \langle \mathbf{u}_n, \mathbf{b} \rangle - \frac{\langle C \mathbf{u}_n, \mathbf{u}_n \rangle}{2} + \int_{\mathbb{R}^2} (\exp(i \langle \mathbf{u}_n, \mathbf{x} \rangle) \right. \right. \\ &\quad \left. \left. - 1 - i \langle \mathbf{u}_n, \mathbf{x} \rangle \mathbb{1}_{\{\|\mathbf{x}\|_{\mathbb{R}^2} \leq 1\}} \right) F(d\mathbf{x}) \right\}, \end{aligned} \tag{4.2}$$

where  $\mathbf{u}_n = (U_n, U_n)$ . In the same vein, we define the characteristic function  $\phi_n(\tilde{\mathbf{u}}_n)$  where  $\tilde{\mathbf{u}}_n = (U_n, -U_n)$ . Here, we focus on estimating the characteristic function on the diagonal of first and fourth quadrant for sake of simplicity to our calculations. The results still hold even when we move away from the diagonal. Following a trivial calculation we get that

$$\begin{aligned} \langle C \mathbf{u}_n, \mathbf{u}_n \rangle &= C^{11} U_n^2 + C^{22} U_n^2 + 2C^{12} U_n^2 \\ \langle C \tilde{\mathbf{u}}_n, \tilde{\mathbf{u}}_n \rangle &= C^{11} U_n^2 + C^{22} U_n^2 - 2C^{12} U_n^2. \end{aligned}$$

So, the covariance is given by

$$C^{12} = \frac{\langle C\mathbf{u}_n, \mathbf{u}_n \rangle - \langle C\tilde{\mathbf{u}}_n, \tilde{\mathbf{u}}_n \rangle}{4U_n^2}. \quad (4.3)$$

We consider, based on the observations, the empirical characteristic function of the increments, at each stage  $n$ :

$$\hat{\phi}_n(\mathbf{u}_n) = \frac{1}{n} \sum_{j=1}^n e^{i\langle \mathbf{u}_n, \Delta_j^n \mathbf{X} \rangle} \quad \mathbf{u}_n \in \mathbb{R}^2. \quad (4.4)$$

Similarly, we consider the empirical characteristic function  $\hat{\phi}_n(\tilde{\mathbf{u}}_n)$ . Based on the trivial calculation (4.3), we now define the *spectral estimator*

$$\hat{C}_n^{12}(U_n) = \frac{n}{2U_n^2} \left( \log \left| \frac{\hat{\phi}_n(\tilde{\mathbf{u}}_n)}{\hat{\phi}_n(\mathbf{u}_n)} \right| \right) \mathbb{1}_{\{\hat{\phi}_n(\tilde{\mathbf{u}}_n) \neq 0 \text{ and } \hat{\phi}_n(\mathbf{u}_n) \neq 0\}}. \quad (4.5)$$

The first result of this paper is the following theorem.

**Theorem 4.2.** *Let  $\mathbf{X}$  belong to the class  $\mathcal{L}_M^r$ . Assume  $M > 0$  and  $r \in [0, 2)$ , then as  $n \rightarrow \infty$  the family of estimators  $\hat{C}_n^{12}(U_n)$  with*

$$U_n = \begin{cases} \sqrt{n} & \text{if } r \leq 1 \\ \sqrt{(r-1)n \log n / \sqrt{4M}} & \text{if } r > 1 \end{cases}$$

satisfies  $|\hat{C}_n^{12}(U_n) - C^{12}| = O_{\mathbb{P}}(w_n)$  within the class  $\mathcal{L}_M^r$  where

$$w_n = \begin{cases} 1/\sqrt{n} & \text{if } r \leq 1 \\ (n \log n)^{\frac{r-2}{2}} & \text{if } r > 1. \end{cases} \quad (4.6)$$

Particularly, we have that the family of estimators  $\hat{C}_n^{12}$  is consistent with the theoretical co-integrated volatility  $C^{12}$  with the exact rates of convergence  $w_n$ .

Theorem 4.2 gives us an upper bound for the family of our estimators  $\hat{C}_n^{12}$ . In Section 6, we give a proof of the upper bound for the family of our estimators  $\hat{C}_n^{12}$ . Let us finally show that on the class  $\mathcal{L}_M^r$  the rate  $w_n$  (4.6) can be achieved exactly and thus constitutes the exact minimax optimal rate.

**Theorem 4.3.** *There are constants  $A, B > 0$  such that*

$$\liminf_{n \rightarrow \infty} \inf_{\hat{C}_n^{12}} \sup_{\mathbf{X} \in \mathcal{L}_M^r} \mathbb{P}[d(\hat{C}_n^{12}, C^{12}) > Aw_n] \geq B > 0,$$

where  $\hat{C}_n^{12}$  is any estimator for the co-integrated volatility, and  $w_n$  is defined as in (4.6).

Theorem 4.3 gives us a lower bound for the family of our estimators  $\hat{C}_n^{12}$  within the class  $\mathcal{L}_M^r$ . The rates  $w_n$  (4.6) for estimating  $C^{12}$ , namely the co-integrated volatility at time  $t = 1$  are optimal in a minimax sense. In Section 7 we prove this result.



### 5. Co-jump activity index

Practically all financial applications require a multivariate model with dependence between the components. Examples are basket option pricing, portfolio risk and hedge management. In most of these applications, jumps in the price process must be taken into account. In particular, measuring the activity of co-jumps up to time  $t$  provides a tool for measuring the propagation among assets or effects due to important positive or negative economic developments. For portfolio and risk management, it is essential to understand co-jumps in order to distinguish between idiosyncratic and systemic risk. Knowledge of co-jump activity index helps us from a modeling viewpoint to judge whether the jump part of a multidimensional Lévy process is of finite or infinite variation.

Although we can always bound the co-jump activity index from above with the component of highest activity index, we are interested in bounding the co-jump activity index from below in case that at least one of the jump components is of infinite variation.

Each component  $X^{(i)}$  of a two-dimensional Lévy process has its own activity index  $r_i$  for  $i = 1, 2$ . In the following we will describe the method for bounding from below the activity index of co-jumps. The BG index of a Lévy process depends only on the Lévy measure  $F$ .  $r$  is an index taking care of positive and negative jumps, for simplicity's sake but without loss of generality we develop our method for the case in which the Lévy measure is one-sided, i.e.  $X^{(i)}$  only makes positive jumps.

*Assumption ( $r$ -stable).* Each  $X^{(i)}$  is following a  $r_i$ -stable Lévy process. Take  $r_2 \geq 1$ , and  $r_1 \in (0, r_2]$  with  $c_i > 0$ ,  $i = 1, 2$ . The jumps of each  $X^{(i)}$  have Lévy measure

$$F^{(i)}(dx_i) = c_i x_i^{-1-r_i} \mathbf{1}(x_i > 0) dx_i.$$

We denote, for each  $i = 1, 2$ , that

$$U_i(x_i) := F^{(i)}([x_i, +\infty)) = c_i \frac{x_i^{-r_i}}{r_i} \quad x_i \in [0, \infty) \tag{5.1}$$

be the tail integral of the marginal Lévy measure  $F^{(i)}$ . Note that  $r_i$  is the BG index of  $X^{(i)}$ .

The co-jump activity index  $r$  will be influenced by the dependence structure between the jump components. We will use a Lévy copula to describe this dependency. The concept of Lévy copula allows us to characterize in a time-independent scheme the dependence structure of the pure jump part of a Lévy process. Here, we use the Lévy copula, which permits a range from a dependent to a total independent framework. For the definition and concepts of independence and total positive dependence copula we refer to Kallsen and Tankov (2006). The next definition for joint jumps is taken from Mancini (2017).

**Definition 5.1.** *The occurrence of joint jumps in  $(X^{(1)}, X^{(2)})$  is described by the following tail integrals*

$$U(x_1, x_2) = F_\gamma([x_1, +\infty) \times [x_2, +\infty)) = C_\gamma(U_1(x_1), U_2(x_2)), \quad x_1, x_2 \in [0, \infty)$$

where  $C_\gamma : [0, \infty]^2 \rightarrow [0, \infty]$  is a Lévy copula of the form

$$C_\gamma(u_1, u_2) = \gamma C_\perp(u_1, u_2) + (1 - \gamma)C_\parallel(u_1, u_2),$$

where  $C_\perp(u_1, u_2) = u_2 \mathbf{1}(u_1 = \infty) + u_1 \mathbf{1}(u_2 = \infty)$  is the independence copula,  $C_\parallel(u_1, u_2) = u_1 \wedge u_2$  is the total positive dependence copula and  $\gamma$  varies in  $[0, 1]$ .

The following remark gives us some clarifications on the definition of the above Lévy copula.

*Remark 5.2.* The marginal tails  $U_i$  are defined on  $[0, \infty]$ , the joint tail is defined on  $[0, \infty]^2$ .  $u_1, u_2$  stands for  $U_1(x_1), U_2(x_2)$ , and  $(u_1, u_2) = (+\infty, +\infty)$  is allowed: both  $U_i(x_i)$  could be  $\infty$ , namely when both  $x_i = 0$ . In that case  $U(x_1, x_2) = C_\gamma(U_1(x_1), U_2(x_2))$  is  $+\infty$ , and  $C_\gamma(\infty, \infty) = 0$ .  $C_\gamma$  is a Lévy copula because it is a convex combination of two Lévy copulas, i.e.  $C_\gamma$  is a 2-increasing, grounded and with uniform margins, because  $C_\perp$  and  $C_\parallel$  are such.  $C_\gamma(u_1, u_2)$  is not a tail integral, it has different properties, for instance  $C_\gamma(u_1, +\infty) = u_1$ , while for any tail  $U$  we have  $U(x_1, +\infty) = 0$ . Finally, when  $\gamma = 0$  jump components are totally dependent while when  $\gamma = 1$  the opposite.

We observe that the activity index of co-jumps is bounded from below by the harmonic mean.

**Lemma 5.3.** *Suppose that Assumptions (2-M) and (r-stable) hold. Let  $X^{(i)}$  be an one-sided  $r_i$ -stable Lévy process for  $i = 1, 2$  with positive jumps  $r_2 \geq 1$ ,  $0 \leq r_1 \leq r_2$  and  $0 < c_1 \leq c_2$ . We assume either complete dependent or independent jumps using Lévy copulas defined as in Definition 5.1. Then, we have that*

$$r > \frac{2r_1r_2}{r_1 + r_2} \geq r_1 \wedge r_2$$

where  $r$  is the activity index of co-jumps.

*Proof.* The independent jumps have sizes of either  $(x_1, 0)$  or  $(0, x_2)$ . This means that we have jumps only on the Cartesian axes. The independent copula regulates such jumps. On the other hand, the complete dependent jumps are regulated by the dependent copula; their size falls into the point  $(x_1, x_2)$ . The complete dependent jumps are completely positively monotonic, i.e. there exists a strictly increasing and positive function  $f$  such that  $\forall t > 0, \Delta X_t^{(2)} = f(\Delta X_t^{(1)})$ . This means that when  $x_1$  is a jump realization so there is a realization  $x_2$  such as  $x_2 = f(x_1)$ , then  $x_1$  is interpreted as the first component of the joint jump. In fact, the sizes  $(x_1, x_2)$  are supported by the graph  $x_2 = f(x_1)$ . For the dependent copula we need the minimum between  $U_1(x_1)$  and  $U_2(x_2)$ , which is attained when  $U_1(x_1) = U_2(x_2)$ . Hence, the graph  $x_2 = U_2^{-1}(U_1(x_1))$  supports the joint jumps.

In our case we assume one-sided  $r_i$ -stable processes, which means that the union graph of the joint jumps is given by  $x_2 = \left(\frac{c_1 r_2}{r_1 c_2}\right)^{-1/r_2} \cdot x_1^{r_1/r_2}$ . We denote by  $F_\gamma$  the Lévy measure in terms of the Lévy copula, using the Definition 5.1. Therefore,

$$F_\gamma(d\mathbf{x}) = (1 - \gamma)C_\parallel(U_1(x_1), U_2(x_2)) + \gamma C_\perp(U_1(x_1), U_2(x_2)). \quad (5.2)$$

Observe that  $\int 1 \wedge (x_1 x_2)^{r/2} dC_{\perp}(U_1(x_1), U_2(x_2)) = 0$ , since the independent copula regulates the jumps on the axes. Inserting (5.2) into Assumption (2-M), it turns out that for  $\epsilon$  smaller than 1, we get

$$\begin{aligned} \int_{0 \leq x_1, x_2 \leq \epsilon} (x_1 x_2)^{r/2} F_{\gamma}(dx_1, dx_2) &= \int_{0 \leq x_1, x_2 \leq \epsilon} (x_1 x_2)^{r/2} dC_{\gamma}(U_1(x_1), U_2(x_2)) \\ &= (1 - \gamma) \int_{0 \leq x_1, x_2 \leq \epsilon} (x_1 x_2)^{r/2} dC_{\parallel}(U_1(x_1), U_2(x_2)). \end{aligned}$$

The second equality holds because of the fact that the integrand is always equally to zero in case of independent jumps. For sake of simplicity, we assume  $\gamma = 0$ , i.e. totally dependent jumps. We assume that the jump sizes  $(x_1, x_2)$  fall into the interval  $(0, \epsilon)$  for sufficiently small  $\epsilon > 0$ . By Assumption (r-stable) and  $c_1 \leq c_2$ , then we have  $U_1(\epsilon) \leq U_2(\epsilon)$ . Indeed, for the derivative of  $\frac{x^{-r}}{r}$  with respect to  $r$  to be positive on  $r \in (0, 2]$ , the condition  $r \log \epsilon < -1$  suffices, such that  $U_1(\epsilon) \leq U_2(\epsilon)$  holds. The last inequality implies  $\epsilon \geq U_2^{-1}(U_1(\epsilon)) = f(\epsilon)$ . Since we want to bound  $x_1 \leq \epsilon$  and  $x_2 = f(x_1) \leq \epsilon$ , this gives us  $x_1 \leq f^{-1}(\epsilon) \wedge \epsilon = \epsilon$ . Hence,

$$\begin{aligned} \int_{0 \leq x_1 \leq \epsilon} (x_1 \cdot f(x_1))^{r/2} dU_1(x_1) &= \int_{0 \leq x_1 \leq \epsilon} (x_1 \cdot f(x_1))^{r/2} c_1 x_1^{-1-r_1} dx_1 \\ &= C^{r/2} \cdot c_1 \int_{0 \leq x_1 \leq \epsilon} \left(x_1^{\frac{r_1}{r_2} + 1}\right)^{r/2} x_1^{-1-r_1} dx_1 \end{aligned} \tag{5.3}$$

where  $C = \left(\frac{c_1 r_2}{r_1 c_2}\right)^{-\frac{1}{r_2}}$ .

In light of the above calculations, in order for the integral in (5.3) not to be divergent we need  $\left(\frac{r_1}{r_2} + 1\right) \frac{r}{2} - 1 - r_1 > -1$ , which means that  $r > \frac{2r_1 r_2}{r_1 + r_2}$ . We observe that  $r$ , the activity index of co-jumps, is at least the harmonic mean of the indices  $r_1, r_2$ . In addition,  $\frac{2r_1 r_2}{r_1 + r_2} \geq r_1$ , since we assume  $r_1 \leq r_2$ . To conclude, the Blumenthal-Gettoor (BG) index of the co-jump activity will be bounded from below by

$$r > \frac{2r_1 r_2}{r_1 + r_2} \geq r_1 \wedge r_2. \tag{5.4}$$

The proof now is complete. □

We see here that the higher the activity of one jump component, the higher the activity of co-jumps.

Next we proceed to the proof of the upper bound Theorem 4.2 using a spectral estimate for the co-integrated volatility. Given the fact that we know an estimate for the *integrated volatility*  $\widehat{IV}$ , we should consider a straightforward estimate for *co-integrated volatility*. By polarization,  $\widehat{IV}(X^{(1)} + X^{(2)})/2 - \widehat{IV}(X^{(1)})/2 - \widehat{IV}(X^{(2)})/2$ , is a possible estimator for the co-integrated volatility. However, we refrain from using this estimate because the rates of convergence are slower than following the procedure as in Section 6. Let us illustrate this argument with an example.

**Example 5.4.** Let  $(X_t) \equiv (X_t^{(1)}, X_t^{(2)})$  be a Lévy process with characteristic triplet  $(0, 0, F(d\mathbf{x}))$ , i.e., without a Gaussian part. We assume its components are independent  $r_i$ -stable Lévy processes for  $i = 1, 2$  such that  $0 \leq r_1 \leq r_2 < 2$  and  $r_2 \geq 1$ . Using Lemma 4.1 from Kallsen and Tankov (2006)  $F$  is supported by the coordinates axes and it can be written as  $F(d\mathbf{x}) = F^{(1)}(dx_1) + F^{(2)}(dx_2)$ . The Lévy measures of the components are

$$F^{(1)}(dx_1) = \frac{1}{|x_1|^{1+r_1}} dx_1 \quad \text{and} \quad F^{(2)}(dx_2) = \frac{1}{|x_2|^{1+r_2}} dx_2.$$

More precisely,

$$\begin{aligned} \int 1 \wedge \|\mathbf{x}\|^r F(d\mathbf{x}) &\geq \int_{0 < x_1, x_2 < 1} |x_1^2 + x_2^2|^{r/2} F(dx_1, dx_2) \\ &= \int_{0 < x_1 < 1} |x_1|^r F^{(1)}(dx_1) + \int_{0 < x_2 < 1} |x_2|^r F^{(2)}(dx_2) \quad (5.5) \\ &= \int_{0 < x_1 < 1} |x_1|^{r-1-r_1} dx_1 + \int_{0 < x_2 < 1} |x_2|^{r-1-r_2} dx_2 \end{aligned}$$

In order the integrals in the last equality not to be divergent we need  $r > r_2$  and  $r > r_1$ . As a consequence, we find that  $r > \max(r_2, r_1)$ . Using (3.1) we find that the Blumenthal-Gettoor index  $r^* = r_2$ .

## 6. Upper Bound

In this section we prove Theorem 4.2. We say that a sequence of estimators  $\widehat{C}_n^{12}$  achieves the rate  $w_n$  on  $\mathcal{L}_M^r$ , for estimating  $C_{12}$ , if  $|\widehat{C}_n^{12} - C_{12}| = O_{\mathbb{P}}(w_n)$ . This means that the family  $\frac{1}{w_n} |\widehat{C}_n^{12} - C_{12}|$  is tight. Note that the argumentation is in line with the bias-variance decomposition.

### 6.1. The bias-variance decomposition

We start with deriving a bias-variance-type decomposition of the estimation error of the estimator for co-integrated volatility.

**Lemma 6.1.** *We have that*

$$\widehat{C}_n^{12}(U_n) - C_{12} = D_n + H_n.$$

The deterministic error given as

$$D_n = \frac{n}{2U_n^2} \left( \log |\phi_n(\tilde{\mathbf{u}}_n)| - \log |\phi_n(\mathbf{u}_n)| \right) - C_{12} \quad (6.1)$$

and the stochastic error as

$$H_n = -\frac{n}{2U_n^2} \left( \log \left| \frac{\phi_n(\tilde{\mathbf{u}}_n)}{\phi_n(\mathbf{u}_n)} \right| - \left( \log \left| \frac{\widehat{\phi}_n(\tilde{\mathbf{u}}_n)}{\widehat{\phi}_n(\mathbf{u}_n)} \right| \right) \mathbf{1}_{\{\widehat{\phi}_n(\tilde{\mathbf{u}}_n) \neq 0 \text{ and } \widehat{\phi}_n(\mathbf{u}_n) \neq 0\}} \right). \quad (6.2)$$

*Proof.* We set  $C_n^{12}(U_n) = \frac{n}{2U_n^2} \left( \log |\phi_n(\tilde{\mathbf{u}}_n)| - \log |\phi_n(\mathbf{u}_n)| \right)$ , recalling the form of the estimator (4.5). We get

$$\begin{aligned} \widehat{C}_n^{12}(U_n) - C_{12} &= \widehat{C}_n^{12}(U_n) + C_n^{12}(U_n) - C_n^{12}(U_n) - C_{12} \\ &= C_n^{12}(U_n) - C_{12} + \widehat{C}_n^{12}(U_n) - C_n^{12}(U_n). \end{aligned} \tag{6.3}$$

Inserting (4.5) into (6.3), we get that *estimation error* is given by  $\widehat{C}_n^{12} - C_{12} = D_n + H_n$ .  $\square$

Notice that the indicator of the event in (6.2) is crucial, since we need to guard large values for the empirical characteristic functions. We want to secure that both  $\widehat{\phi}_n(\mathbf{u}_n)$  and  $\widehat{\phi}_n(\tilde{\mathbf{u}}_n)$  are not equal to zero at the same time.

Our goal is to show that the estimation error is stochastically bounded, i.e.  $O_{\mathbb{P}}(w_n)$ . Firstly, we bound the deterministic error.

### 6.2. Bounding the deterministic error

**Lemma 6.2.** *Grant Assumption (2-M). The deterministic error satisfies*

$$|D_n| \leq 2^{r/2} M U_n^{r-2} + A U_n^{-2},$$

where  $A$  is a positive constant.

*Proof.* Recall the characteristic function of  $\mathbf{X}_{1/n}$  in (4.2). We define

$$d_n = 2 \int_{\mathbb{R}^2} \left( 1 - \cos(\langle \mathbf{u}_n, \mathbf{x} \rangle) \right) F(d\mathbf{x}) \tag{6.4}$$

$$\tilde{d}_n = 2 \int_{\mathbb{R}^2} \left( 1 - \cos(\langle \tilde{\mathbf{u}}_n, \mathbf{x} \rangle) \right) F(d\mathbf{x}). \tag{6.5}$$

Therefore,

$$|\phi_n(\mathbf{u}_n)| = \exp \left( -\frac{1}{2n} \left( \langle C\mathbf{u}_n, \mathbf{u}_n \rangle + d_n \right) \right)$$

and  $|\phi_n(\tilde{\mathbf{u}}_n)| = \exp \left( -\frac{1}{2n} \left( \langle C\tilde{\mathbf{u}}_n, \tilde{\mathbf{u}}_n \rangle + \tilde{d}_n \right) \right)$ . Notice that here we use an argument of complex analysis, if  $Z \in \mathbb{C}$  then  $|e^Z| = e^{\text{Re}(Z)}$ . Summing up,

$$\frac{n}{2U_n^2} \left( \log |\phi_n(\tilde{\mathbf{u}}_n)| - \log |\phi_n(\mathbf{u}_n)| \right) - C_{12} = \frac{1}{4U_n^2} \left( d_n - \tilde{d}_n \right). \tag{6.6}$$

By (6.6), we have

$$\begin{aligned} |D_n| &= \frac{1}{2U_n^2} \left| \int \left( 1 - \cos(\langle \mathbf{u}_n, \mathbf{x} \rangle) \right) F(d\mathbf{x}) - \int \left( 1 - \cos(\langle \tilde{\mathbf{u}}_n, \mathbf{x} \rangle) \right) F(d\mathbf{x}) \right| \\ &= \frac{1}{2U_n^2} \left| \int \left( \cos(\langle \tilde{\mathbf{u}}_n, \mathbf{x} \rangle) - \cos(\langle \mathbf{u}_n, \mathbf{x} \rangle) \right) F(d\mathbf{x}) \right| \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{2U_n^2} \int \left( 2 \wedge |\langle \mathbf{u}_n, \mathbf{x} \rangle^2 - \langle \tilde{\mathbf{u}}_n, \mathbf{x} \rangle^2| \right) F(d\mathbf{x}) \\ &= \frac{1}{2U_n^2} \int (2 \wedge |4U_n^2 x_1 x_2|) F(dx_1, dx_2), \end{aligned}$$

where we used the fact that  $|\cos x - \cos y| \leq 2 \wedge |x^2 - y^2|$ . Using the inequality  $a \wedge b \leq a^p b^{1-p}$  for  $p \in (0, 1)$ , the last term can be bounded as follows

$$\begin{aligned} |D_n| &\leq \frac{1}{U_n^2} \int \left( 1 \wedge 2U_n^2 |x_1 x_2| \right) F(dx_1, dx_2) \\ &\leq \frac{1}{U_n^2} \int_{B_1(0)} (2U_n^2 |x_1 x_2|)^{r/2} 1^{1-r/2} F(dx_1, dx_2) \\ &\quad + \frac{1}{U_n^2} \int_{\mathbb{R}^2 \setminus B_1(0)} 1 F(dx_1, dx_2) \\ &= U_n^{r-2} \int_{B_1(0)} (2|x_1 x_2|)^{r/2} F(dx_1, dx_2) + U_n^{-2} F(\mathbb{R}^2 \setminus B_1(0)), \end{aligned} \tag{6.7}$$

here  $\frac{r}{2} \in (0, 1)$  because  $r \in (0, 2)$  is the co-jump activity index. By Assumption (2-M) and for some constant  $A > 0$ ,

$$|D_n| \leq 2^{r/2} M U_n^{r-2} + A U_n^{-2} \tag{6.8}$$

as required.  $\square$

### 6.3. Bounding the stochastic error

We want to investigate how close the empirical characteristic function is to the characteristic function of a two-dimensional Lévy process. The variables  $e^{i\langle \mathbf{u}_n, \Delta_j^n \mathbf{X} \rangle}$  are i.i.d. as  $j$  varies, with expectation  $\phi_n(\mathbf{u}_n)$ . The respective statement holds true for  $e^{i\langle \tilde{\mathbf{u}}_n, \Delta_j^n \mathbf{X} \rangle}$  as well. So  $\hat{\phi}_n(\mathbf{u}_n)$  is an unbiased estimator because  $\mathbb{E}[\hat{\phi}_n(\mathbf{u}_n)] = \phi_n(\mathbf{u}_n)$ . Also, the variance of the empirical characteristic function is given by  $\text{Var}(\hat{\phi}_n(\mathbf{u}_n)) = \frac{1}{n} (1 - |\phi_n(\mathbf{u}_n)|^2)$ , where we use the following definition.

**Definition 6.3.** For a  $\mathbb{C}$ -valued random variable  $Z$  we define

$$\begin{aligned} \text{Var}(Z) &= \mathbb{E}[(Z - \mathbb{E}(Z))(\bar{Z} - \mathbb{E}(\bar{Z}))] \\ &= \mathbb{E}[\bar{Z}Z - Z\mathbb{E}(\bar{Z}) - \bar{Z}\mathbb{E}(Z) + \mathbb{E}(Z)\mathbb{E}(\bar{Z})] \\ &= \mathbb{E}(|Z|^2) - |\mathbb{E}(Z)|^2. \end{aligned}$$

**Lemma 6.4.** Let  $V_n = \hat{\phi}_n(\mathbf{u}_n) - \phi_n(\mathbf{u}_n)$  and  $\tilde{V}_n = \hat{\phi}_n(\tilde{\mathbf{u}}_n) - \phi_n(\tilde{\mathbf{u}}_n)$ , where  $V_n, \tilde{V}_n \in \mathbb{C}$ . Then,  $\mathbb{E}(|V_n|^2) \leq \frac{1}{n}$  and  $\mathbb{E}(|\tilde{V}_n|^2) \leq \frac{1}{n}$ .

*Proof.* Set  $V_n = Z \in \mathbb{C}$  such that  $V_n = \hat{\phi}_n(\mathbf{u}_n) - \phi_n(\mathbf{u}_n)$ . Remember that  $\hat{\phi}_n$  is unbiased due to the fact that  $\mathbb{E}[\hat{\phi}_n(\mathbf{u}_n)] = \phi_n(\mathbf{u}_n)$ , thus  $|\mathbb{E}(Z)|^2 = 0$ . Taking this into consideration with the previous Definition 6.3, we obtain that

$\mathbb{E}[|Z|^2] = \text{Var}[Z]$ . Therefore,

$$\mathbb{E}(|V_n|^2) = \text{Var}(V_n) = \text{Var}(\widehat{\phi}_n(\mathbf{u}_n)) = \frac{1}{n} (1 - |\phi_n(\mathbf{u}_n)|^2) \leq \frac{1}{n}. \tag{6.9}$$

The same argument holds also for  $\mathbb{E}(|\tilde{V}_n|^2)$ . This completes the proof.  $\square$

We choose  $\mathbf{u}_n = (U_n, U_n)$  and  $\tilde{\mathbf{u}}_n = (U_n, -U_n)$ . Recall that we estimate the characteristic function on the diagonal of first and fourth quadrant for calculation simplicity. Particularly, we choose for  $M > 0$ ,  $r \in [0, 2)$  and  $n$  large enough

$$U_n = \begin{cases} \sqrt{n} & \text{if } r \leq 1 \\ \sqrt{(r-1)n \log n} / \sqrt{4M} & \text{if } r > 1. \end{cases} \tag{6.10}$$

**Lemma 6.5.** *Grant Assumption (1-M). For some positive constants  $B, \Gamma, M$  and on the event  $\left\{ |V_n| \vee |\tilde{V}_n| \leq \frac{1}{n^{r/4}} \right\}$  the stochastic error satisfies:*

$$\mathbb{E} \left[ |H_n| \mathbb{1}_{\left\{ |V_n| \vee |\tilde{V}_n| \leq \frac{1}{n^{r/4}} \right\}} \right] \leq \begin{cases} \frac{B\Gamma}{\sqrt{n}} & \text{if } r \leq 1 \\ \frac{BM}{(r-1)n^{\frac{2-r}{2}} \log n} & \text{if } r > 1. \end{cases} \tag{6.11}$$

*Proof.* Recalling the form of stochastic error (6.2), the first quantity we need to bound is:

$$\begin{aligned} \frac{1}{|\phi_n(\mathbf{u}_n)|} &= \exp \left( \frac{1}{2n} (\langle C\mathbf{u}_n, \mathbf{u}_n \rangle + d_n) \right) \\ &= \exp \left( \frac{1}{2n} (C^{11}U_n^2 + C^{22}U_n^2 + 2C^{12}U_n^2 + d_n) \right) \\ &\leq \exp \left( \frac{1}{2n} U_n^2 \left( C^{11} + C^{22} + 2C^{12} + 4 \int_{\mathbb{R}^2} (1 \wedge \|\mathbf{x}\|^2) F(d\mathbf{x}) \right) \right) \\ &\leq \exp \left( \frac{1}{2n} U_n^2 \left( 4\|C\|_\infty + 4 \int_{\mathbb{R}^2} (1 \wedge \|\mathbf{x}\|^2) F(d\mathbf{x}) \right) \right) \\ &\leq \exp \left( \frac{4M}{2n} U_n^2 \right). \end{aligned} \tag{6.12}$$

The first inequality holds because

$$\begin{aligned} d_n &= 2 \int (1 - \cos(\langle \mathbf{u}_n, \mathbf{x} \rangle)) F(d\mathbf{x}) \leq 2 \int (1 \wedge |\langle \mathbf{u}_n, \mathbf{x} \rangle|^2) F(d\mathbf{x}) \\ &\leq 4U_n^2 \int (1 \wedge \|\mathbf{x}\|^2) F(d\mathbf{x}), \end{aligned}$$

using the Cauchy-Schwarz inequality for  $|\langle \mathbf{u}_n, \mathbf{x} \rangle|^2 \leq \|\mathbf{u}_n\|^2 \|\mathbf{x}\|^2$  and the fact that  $U_n \geq 1$ . The last inequality in (6.12) derives from Assumption (1-M) and the fact that we always have  $\int_{\mathbb{R}^2} (1 \wedge \|\mathbf{x}\|^2) F(d\mathbf{x}) < \infty$ . Next, the form of  $U_n$  (6.10) implies that

$$\frac{1}{|\phi_n(\mathbf{u}_n)|} \leq \begin{cases} \Gamma & \text{if } r \leq 1 \\ n^{\frac{r-1}{2}} & \text{if } r > 1, \end{cases} \tag{6.13}$$

where  $\Gamma = e^{2M}$ . The equivalent statement holds for  $\phi_n(\tilde{\mathbf{u}}_n)$ . Let us now argue that

$$\frac{\widehat{\phi}_n(\tilde{\mathbf{u}}_n)}{\widehat{\phi}_n(\mathbf{u}_n)} = \frac{\tilde{V}_n + \phi_n(\tilde{\mathbf{u}}_n)}{V_n + \phi_n(\mathbf{u}_n)} = \frac{\phi_n(\tilde{\mathbf{u}}_n) \left(1 + \frac{\tilde{V}_n}{\phi_n(\tilde{\mathbf{u}}_n)}\right)}{\phi_n(\mathbf{u}_n) \left(1 + \frac{V_n}{\phi_n(\mathbf{u}_n)}\right)} \neq 0, \infty$$

as soon as  $n \geq n_0 = (2\Gamma)^{\frac{4}{(2-r)\wedge r}}$  by (6.13) and on the set  $\left\{|V_n| \vee |\tilde{V}_n| \leq \frac{1}{n^{r/4}}\right\}$ . Therefore,  $\left|\frac{V_n}{\phi_n(\mathbf{u}_n)}\right| \leq \frac{1}{2}$  and  $\left|\frac{\tilde{V}_n}{\phi_n(\tilde{\mathbf{u}}_n)}\right| \leq \frac{1}{2}$ . Accordingly, for the stochastic error on the event  $\left\{|V_n| \vee |\tilde{V}_n| \leq \frac{1}{n^{r/4}}\right\}$  we obtain for some deterministic constant  $B$ :

$$\begin{aligned} |H_n| &\leq \frac{n}{2U_n^2} \left| \log \left| \frac{\widehat{\phi}_n(\tilde{\mathbf{u}}_n)}{\widehat{\phi}_n(\mathbf{u}_n)} \right| - \log \left| \frac{\phi_n(\tilde{\mathbf{u}}_n)}{\phi_n(\mathbf{u}_n)} \right| \right| \\ &= \frac{n}{2U_n^2} \left| \log \left| 1 + \frac{\widehat{\phi}_n(\tilde{\mathbf{u}}_n) - \phi_n(\tilde{\mathbf{u}}_n)}{\phi_n(\tilde{\mathbf{u}}_n)} \right| - \log \left| 1 + \frac{\widehat{\phi}_n(\mathbf{u}_n) - \phi_n(\mathbf{u}_n)}{\phi_n(\mathbf{u}_n)} \right| \right| \\ &= \frac{n}{2U_n^2} \left| \log \left| 1 + \frac{\tilde{V}_n}{\phi_n(\tilde{\mathbf{u}}_n)} \right| - \log \left| 1 + \frac{V_n}{\phi_n(\mathbf{u}_n)} \right| \right| \\ &\leq \frac{Bn}{2U_n^2} \left| \left| \frac{\tilde{V}_n}{\phi_n(\tilde{\mathbf{u}}_n)} \right| - \left| \frac{V_n}{\phi_n(\mathbf{u}_n)} \right| \right|. \end{aligned}$$

In the last inequality, we use the linearized stochastic errors for

$$\log \left| 1 + \frac{\tilde{V}_n}{\phi_n(\tilde{\mathbf{u}}_n)} \right| \approx \left| \frac{\tilde{V}_n}{\phi_n(\tilde{\mathbf{u}}_n)} \right|$$

because of the fact that  $\frac{\tilde{V}_n}{\phi_n(\tilde{\mathbf{u}}_n)}$  and  $\frac{V_n}{\phi_n(\mathbf{u}_n)}$  are small enough. So there is a positive constant  $B$  such that  $\log \left| 1 + \frac{\tilde{V}_n}{\phi_n(\tilde{\mathbf{u}}_n)} \right| \leq B \left| \frac{\tilde{V}_n}{\phi_n(\tilde{\mathbf{u}}_n)} \right|$ . Therefore,

$$|H_n| \leq \frac{Bn}{U_n^2} \max \left( \left| \frac{\tilde{V}_n}{\phi_n(\tilde{\mathbf{u}}_n)} \right|, \left| \frac{V_n}{\phi_n(\mathbf{u}_n)} \right| \right). \tag{6.14}$$

Henceforth, combining  $\mathbb{E}[|V_n|] \leq \frac{1}{\sqrt{n}}$  and  $\mathbb{E}[|\tilde{V}_n|] \leq \frac{1}{\sqrt{n}}$  with (6.14), for  $n \geq n_0$ , and for some constant  $B > 0$ , we have

$$\begin{aligned} &\mathbb{E} \left[ |H_n| \mathbf{1}_{\{|V_n| \vee |\tilde{V}_n| \leq \frac{1}{n^{r/4}}\}} \right] \\ &\leq \mathbb{E} \left[ \frac{Bn}{U_n^2} \max \left( \left| \frac{\tilde{V}_n}{\phi_n(\tilde{\mathbf{u}}_n)} \right|, \left| \frac{V_n}{\phi_n(\mathbf{u}_n)} \right| \right) \mathbf{1}_{\{|V_n| \vee |\tilde{V}_n| \leq \frac{1}{n^{r/4}}\}} \right] \\ &\leq \frac{Bn}{U_n^2} \max \left( \left| \frac{1}{\phi_n(\mathbf{u}_n)} \right|, \left| \frac{1}{\phi_n(\tilde{\mathbf{u}}_n)} \right| \right) \frac{1}{\sqrt{n}}. \end{aligned}$$



The last inequality holds because we applied the Cauchy-Schwarz inequality and by Lemma 6.4. To sum up, by (6.13) we get that

$$\mathbb{E} \left[ |H_n| \mathbb{1}_{\{|V_n| \vee |\tilde{V}_n| \leq \frac{1}{n^{r/4}}\}} \right] \leq \begin{cases} \frac{B\Gamma}{\sqrt{n}} & \text{if } r \leq 1 \\ \frac{BM}{(r-1)n^{\frac{2-r}{2}} \log n} & \text{if } r > 1 \end{cases} \quad (6.15)$$

as required.  $\square$

*Remark 6.6.* Here, we are interested in the events  $\{|V_n| \leq \frac{1}{n^{r/4}}\}$  and  $\{|\tilde{V}_n| \leq \frac{1}{n^{r/4}}\}$  because the probabilities of the above events are negligible. Indeed, applying the Chebyshev inequality and by Lemma 6.4 we get

$$\mathbb{P} \left( |V_n| > \frac{1}{n^{r/4}} \right) \leq n^{r/2} \mathbb{E} (|V_n|^2) \leq n^{r/2} \frac{1}{n} = n^{\frac{r-2}{2}},$$

which tends towards zero as  $n \rightarrow \infty$ . Likewise, the probability of the event  $\{|\tilde{V}_n| > \frac{1}{n^{r/4}}\}$  tends towards zero as  $n \rightarrow \infty$ .

Until now we bound from above the deterministic and stochastic errors. We are now ready to prove that the family  $\frac{1}{w_n} |\hat{C}_n^{12} - C^{12}|$  is tight in  $\mathcal{L}_M^r$  and thus establish an upper bound for our estimator.

#### **End proof of Theorem 4.2**

Applying the Markov inequality, we get for any  $\epsilon > 0$ , there exists a finite  $L > 0$  such that,

$$\begin{aligned} \mathbb{P} \left( \frac{1}{w_n} |\hat{C}_n^{12}(U_n) - C^{12}| \geq L \right) &\leq \frac{1}{L} \mathbb{E} \left[ \frac{1}{w_n} |\hat{C}_n^{12}(U_n) - C^{12}| \mathbb{1}_{\{|V_n| \vee |\tilde{V}_n| \leq \frac{1}{n^{r/4}}\}} \right] \\ &\quad + \mathbb{P} \left( |V_n| > \frac{1}{n^{r/4}} \right) + \mathbb{P} \left( |\tilde{V}_n| > \frac{1}{n^{r/4}} \right) \\ &\leq \frac{1}{L} \mathbb{E} \left[ \frac{1}{w_n} |\hat{C}_n^{12}(U_n) - C^{12}| \mathbb{1}_{\{|V_n| \vee |\tilde{V}_n| \leq \frac{1}{n^{r/4}}\}} \right] \\ &\quad + 2n^{\frac{r-2}{2}} \\ &\leq \frac{1}{L} \mathbb{E} \left[ \frac{1}{w_n} |H_n| \mathbb{1}_{\{|V_n| \vee |\tilde{V}_n| \leq \frac{1}{n^{r/4}}\}} \right] + \frac{1}{L} \frac{1}{w_n} |D_n| \\ &\quad + 2n^{\frac{r-2}{2}}. \end{aligned}$$

Further applying Lemmas 6.5 and 6.2, we deduce that, as  $n \rightarrow \infty$ , there is a constant  $C_{L,r,M}$  which depends on  $L, r, M$  such that

$$\mathbb{P} \left( \frac{1}{w_n} |\hat{C}_n^{12}(U_n) - C^{12}| \geq L \right) \leq C_{L,r,M}, \quad (6.16)$$

for  $C_{L,r,M} \leq \epsilon$ , which proves that the family  $\frac{1}{w_n} |\hat{C}_n^{12}(U_n) - C^{12}|$  is tight in  $\mathbf{X} \in \mathcal{L}_M^r$ . The proof is complete.

## 7. Lower bound

In nonparametric statistics it is common to use a minimax approach in order to prove optimality of estimators. In the previous section, we proved Theorem 4.2, which gave us an upper bound for the estimation of co-integrated volatility using a spectral approach and establishing the rates (4.6) on the class  $\mathcal{L}_M^r$ .

In this section, we want to prove Theorem 4.3. The existence of a lower bound on the class  $\mathcal{L}_M^r$  constitutes the exact minimax rates for the estimation of co-integrated volatility. Indeed, we have something more for the lower bound, namely that any estimator on a general class of Itô semimartingales satisfying a condition similar to Definition 4.1 achieves a lower bound with rates (4.6). In view of the formulation of (4.1), we can consider, for any  $M > 0$ , the class

$$\begin{aligned} \mathcal{S}_M^r &= \text{the set of all Itô semimartingales with} \\ \|C_t\|_\infty + \int_{\mathbb{R}^2} (1 \wedge |x_1 x_2|^{r/2}) F_t(dx_1, dx_2) &\leq M \quad \text{for all } t. \end{aligned}$$

Observe that  $\mathcal{L}_M^r \subset \mathcal{S}_M^r$ . The lower bound for the class  $\mathcal{L}_M^r$  will hold for the class  $\mathcal{S}_M^r$  as well. So far, we do not know whether the spectral approach for the upper bound yields the same optimal rate on the larger class of Itô semimartingale.

We refer to Chapter 2 in Tsybakov (2009) for the techniques to prove the lower bounds. We establish the lower bound following the argumentation in line with a two-hypothesis test. Next, we introduce a distance between probability measures that will be useful for the lower bound.

**Definition 7.1.** Let  $(\mathcal{X}, \mathcal{A})$  be a measurable space and let  $\mathbb{P}_\mathbf{X}$  and  $\mathbb{P}_\mathbf{Y}$  be two probability measures on  $(\mathcal{X}, \mathcal{A})$ . The **total variation distance** between  $\mathbb{P}_\mathbf{X}, \mathbb{P}_\mathbf{Y}$  is defined as follows:

$$\text{TV}(\mathbb{P}_\mathbf{X}, \mathbb{P}_\mathbf{Y}) := \sup_{A \in \mathcal{A}} \left| \int (p_0 - p_1) \nu(dx) \right|,$$

where  $p_0 = d\mathbb{P}_\mathbf{X}/d\nu$ ,  $p_1 = d\mathbb{P}_\mathbf{Y}/d\nu$  and  $\nu = \mathbb{P}_\mathbf{X} + \mathbb{P}_\mathbf{Y}$  a  $\sigma$ -finite measure.

To sum up, in order to prove a lower bound on the minimax probability of error for hypotheses we use Theorem 2.2 in Tsybakov (2009). The lower bound is obtained when the following two properties are satisfied. First, we choose the appropriate parameters for the co-integrated volatility to be close enough but distinguished. Second, we bound from below the total variation distance between the two probability densities of our parameters.

Let us illustrate the above procedure giving a trivial lemma and proving that the above arguments are adequate, so as to obtain the lower bound corresponding to our family of estimators  $\mathcal{L}_M^r$ . The interested reader may refer to Lehmann and Romano (2006) who explore a lot of examples for hypothesis testing and distances between Gaussian random variables.

**Lemma 7.2.** Assume  $\mathbf{X}$  belongs to the class  $\mathcal{L}_M^r$  with Lévy-Khintchine triplet  $(0, \Sigma \Sigma^\top, 0)$ . Then there are constants  $A, K$  such that

$$\liminf_{n \rightarrow \infty} \inf_{\widehat{C}_n^{12}} \sup_{\mathbf{X} \in \mathcal{L}_M^r} \mathbb{P}[d(\widehat{C}_n^{12}, C^{12}) > Aw_n] \geq K,$$

where  $w_n = \frac{1}{\sqrt{n}}$ , and  $\widehat{C}_n^{12}$  is any estimator for the co-integrated volatility within the class  $\mathcal{L}_M^r$ .

*Proof.* Consider  $\mathbf{X}$  and  $\mathbf{Y}$  belong to  $\mathcal{L}_M^r$ . Also, we assume that no jumps are occurred so the Lévy Khintchine triplets for each process will satisfy  $(0, \Sigma_{\mathbf{X}} \Sigma_{\mathbf{X}}^\top, 0)$  and  $(0, \Sigma_{\mathbf{Y}} \Sigma_{\mathbf{Y}}^\top, 0)$  respectively. As a result,  $\mathbf{X}$  will evolve as follows:

$$\begin{aligned} dX^{(1)}(t) &= \sigma_t^{(1)} dW_t^{(1)} \\ dX^{(2)}(t) &= \sigma_t^{(2)} dW_t^{(2)} \end{aligned}$$

similarly for  $\mathbf{Y}$ . We know that the Itô integral  $dX^{(1)}(t) = \sigma_t^{(1)} dW_t^{(1)}$  is normally distributed with mean 0 and its variance is given by Itô isometry which translates to

$$X^{(1)} \sim \mathcal{N}\left(0, \int_0^1 (\sigma_t^{(1)})^2 dt\right).$$

Therefore,  $\mathbf{X}$  follows the parametric model

$$\mathbf{X} = \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma_{\mathbf{X}} \Sigma_{\mathbf{X}}^\top\right)$$

similarly for  $\mathbf{Y}$ .

We will prove the lower bound using the two-hypothesis test, as mentioned in the beginning of this section. We observe that

$$\mathcal{L}_M^r \supset \mathcal{B}_M$$

where  $\mathcal{B}_M$  is the class of all Brownian motions where the covariance matrix is bounded component-wise by a constant  $M$ . As a consequence,

$$\sup_{\mathbf{X} \in \mathcal{B}_M} \mathbb{P}[d(\widehat{C}_n^{12}, C^{12}) \geq Aw_n] \leq \sup_{\mathbf{X} \in \mathcal{L}_M^r} \mathbb{P}[d(C^{12}, \widehat{C}_n^{12}) \geq Aw_n].$$

It is enough to prove a lower bound for the rate  $w_n = \frac{1}{\sqrt{n}}$  for the class of all Brownian motions.

The two-hypothesis test is the following

$$\mathbb{P}_{\mathbf{X}} = \mathcal{N}(0, \Sigma_{\mathbf{X}} \Sigma_{\mathbf{X}}^\top) \quad \text{vs} \quad \mathbb{P}_{\mathbf{Y}} = \mathcal{N}(0, \Sigma_{\mathbf{Y}} \Sigma_{\mathbf{Y}}^\top),$$

where the covariance matrices are given by  $\Sigma_{\mathbf{X}} \Sigma_{\mathbf{X}}^\top = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$  and  $\Sigma_{\mathbf{Y}} \Sigma_{\mathbf{Y}}^\top = \begin{pmatrix} 2 & 1 + \frac{1}{\sqrt{n}} \\ 1 + \frac{1}{\sqrt{n}} & 1 \end{pmatrix}$ . Intuitively, we perturb the off-diagonal elements, namely the covariance, by the rate we want to achieve. Following the argumentation of the two-hypothesis test, it is sufficient to prove that the total variation distance is bounded. To do so, we use the Pinsker inequality. By Pinsker inequality we have that

$$\text{TV}(\mathbb{P}_{\mathbf{Y}}, \mathbb{P}_{\mathbf{X}}) \leq \sqrt{KL(\mathbb{P}_{\mathbf{Y}}, \mathbb{P}_{\mathbf{X}})/2},$$

where  $KL(\mathbb{P}_{\mathbf{Y}}, \mathbb{P}_{\mathbf{X}})$  is the Kullback-Leibler divergence. Next, we show that the Kullback-Leibler distance is bounded. We define the Kullback-Leibler divergence

between two multivariate normal distributions. Here, we denote by  $\Sigma_1 = \Sigma_{\mathbf{X}}\Sigma_{\mathbf{X}}^\top$  and  $\Sigma_2 = \Sigma_{\mathbf{Y}}\Sigma_{\mathbf{Y}}^\top$ . Therefore,

$$KL(\mathbb{P}_{\mathbf{Y}}, \mathbb{P}_{\mathbf{X}}) = \frac{1}{2} \left( \log \frac{|\Sigma_1|}{|\Sigma_2|} - 2 + \text{tr}(\Sigma_1^{-1}\Sigma_2) \right),$$

where  $|\cdot|$  denotes the determinant of a matrix. Calculating the appropriate quantities, we obtain

$$\begin{aligned} |\Sigma_1| &= \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1, \\ |\Sigma_2| &= \begin{vmatrix} 2 & 1+\frac{1}{\sqrt{n}} \\ 1+\frac{1}{\sqrt{n}} & 1 \end{vmatrix} = 2 - \left(1 + \frac{1}{\sqrt{n}}\right)^2, \\ \text{tr}(\Sigma_1^{-1}\Sigma_2) &= 2 - \frac{2}{\sqrt{n}}. \end{aligned}$$

Therefore,

$$KL(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}}) = \frac{1}{2} \left( \log \left( \frac{1}{2 - \left(1 + \frac{1}{\sqrt{n}}\right)^2} \right) - 2 + 2 - \frac{2}{\sqrt{n}} \right).$$

Consequently, we obtain that the right hand side tends to zero as  $n \rightarrow \infty$ . By Pinsker inequality, the total variation distance tends to zero. Upon using the minimax probability of error is bounded from below by 1/2 and the claim follows.  $\square$

To prove Theorem 4.3 we need to construct the two-hypothesis test in order to bound from below the minimax probability error as we described previously in Lemma 7.2.

### 7.1. Two-hypothesis test

Let  $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^2$  be Lévy processes with respective triplets  $(0, \Sigma_{\mathbf{X}}\Sigma_{\mathbf{X}}^\top, F_n)$ ,  $(0, \Sigma_{\mathbf{Y}}\Sigma_{\mathbf{Y}}^\top, G_n)$ , where  $F_n, G_n$  are Lévy measures in  $\mathbb{R}^2$  satisfying

$$\int_{\mathbb{R}^2} (1 \wedge |x_1 x_2|^{r/2}) F_n(dx_1, dx_2) \leq M \tag{7.1}$$

where  $\mathbf{x} = (x_1, x_2)$  is a vector in  $\mathbb{R}^2$  representing the size of small jumps for each process and  $M$  is a constant (below  $M$  changes from line to line and may depend on  $r$ , but all constants are denoted as  $M$ ). The same applies for the measure  $G_n$ . We set  $\Sigma_{\mathbf{X}}\Sigma_{\mathbf{X}}^\top = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$  and  $\Sigma_{\mathbf{Y}}\Sigma_{\mathbf{Y}}^\top = \begin{pmatrix} 2 & 1+\pi w_n \\ 1+\pi w_n & 1 \end{pmatrix}$  to be the parameters of our two-hypothesis test. Under this setting, we perturb the off-diagonal elements with the rate with which we want to achieve the lower bound. The quantity which we want to recover is the co-integrated volatility, so we need the off-diagonal elements. We use these forms of matrices in order for the Gaussian part to be non-degenerated, namely the eigenvalues of the matrices

to be positive. As we discussed in the beginning of this section, it is sufficient to construct two sequences  $\mathbf{X}_n, \mathbf{Y}_n$  which belong to the class  $\mathcal{L}_M^r$ , with the following two properties:

**Property 1.** *The two parameters, namely the two covariance matrices are close enough but distinguished.*

Note that for this property the object of our study is the distance between matrices. In this case we consider as a distance the Frobenius norm, and everything still holds. By construction and Frobenius norm

$$\begin{aligned} \|\Sigma_{\mathbf{Y}}\Sigma_{\mathbf{Y}}^\top - \Sigma_{\mathbf{X}}\Sigma_{\mathbf{X}}^\top\|_F &= \sqrt{\text{tr}(\Sigma_{\mathbf{Y}}\Sigma_{\mathbf{Y}}^\top - \Sigma_{\mathbf{X}}\Sigma_{\mathbf{X}}^\top)(\Sigma_{\mathbf{Y}}\Sigma_{\mathbf{Y}}^\top - \Sigma_{\mathbf{X}}\Sigma_{\mathbf{X}}^\top)^\top} \\ &= \sqrt{2}\pi w_n, \end{aligned}$$

which means that the parameters are close enough but distinguished.

**Property 2.** *The total variation distance between  $\mathbb{P}_{\mathbf{X}}$  and  $\mathbb{P}_{\mathbf{Y}}$  tends towards zero.*

As far as the second property is concerned, the total variation distance tends towards zero is not trivial. In fact, achieving the second property is quite demanding and we prove several lemmas to conclude this property.

### 7.2. Construction of the co-jump measure in $\mathbb{R}^2$

First, we have to construct a measure to satisfy property (7.1). Before we proceed with the technical part of this construction, let us highlight the idea behind it. Note that we are studying a two-dimensional Lévy process, so it is reasonable to include the possibility of dependence between the two jump components, more specifically the common jumps, i.e. the co-jumps.

Observe here that co-jumps are one-dimensional objects. Co-jumps are the jumps on the diagonal, due to the fact that the two processes jump at the same time with the same jump size. Mathematically speaking, this can be formalized as follows:

**Definition 7.3** (Co-jump measure). Let  $\mathbf{X} = (X^{(1)}, X^{(2)})$  be a Lévy processes, with  $\Delta X_t^{(j)} \neq 0$  for  $j = 1, 2$ . Here,  $\Delta X_t^{(j)} = X_t^{(j)} - X_{t^-}^{(j)}$  denotes the possible jump at time  $t$ . The measure on  $\mathbb{R}^2$  is defined by:

$$F_n(\omega; t, B) = \mathbb{E} \left[ \# \left\{ t \in [0, 1] : (\Delta X_t^{(1)}, \Delta X_t^{(2)}) \in B \right\} \right] = \mathbb{E} \left[ \mu^{X^{(1)}X^{(2)}}(\omega; t, B) \right],$$

where  $B = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = x_2\}$ . This is called the Lévy measure in  $\mathbb{R}^2$  of co-jumps for  $\mathbf{X}$ .  $F_n(B)$  is the expected number of joint jumps, per unite of time, whose size falls into  $B$ , and  $\mu$  is the Poisson random measure of co-jumps where,  $\mu^X(\omega; t, B) = \sum_{s \leq t} \mathbb{1}_B(\Delta X_s^{(1)}, \Delta X_s^{(2)})$ .

We assume that the co-jump measure  $F_n$  has a density  $f_n$  with respect to Lebesgue measure on  $B$ . Because the jump dynamics of the co-jump measure

is dictated by its density, say  $f_n$ , we can define the measure of the co-jumps as following.

**Definition 7.4.** For  $A \subset B \subset \mathbb{R}^2$ , we write the measure of co-jumps as

$$F_n(A) := \int f_n(x) \mathbb{1}_A(x, x) dx := \int f_n(x) \mathbb{1}_{\tilde{A}}(x) dx, \quad (7.2)$$

where  $\tilde{A} = \{x : (x, x) \in A\}$ .

The support of the co-jump measure is on  $B$  but the co-jump measure lives on  $\mathbb{R}^2$ . We focus on the case of co-jumps, i.e., when  $X^{(1)}$  and  $X^{(2)}$  jump at the same time with the same jump size. We are interested in the jumps on the diagonal.

Further, we do not integrate with respect to the Lebesgue measure, since it is equal to zero on the diagonal. In this case we integrate with respect to a measure that is not absolutely continuous with the Lebesgue measure, which we call *co-jump measure*. We assume that  $F_n, G_n$  have densities  $f_n$  and  $g_n$  respectively. By (7.2) we want to show that

$$\int_{\mathbb{R}^2} g(x_1, x_2) dF_n(x_1, x_2) = \int_{\mathbb{R}} g(x, x) f_n(x) dx \quad (7.3)$$

without being equal to zero. Being interested in the set of co-jumps, we pass from two dimensions to one dimension. Co-jumps are the concept of total dependency between the small jump components. Indeed, we use the argument of dependency in order to reduce dimensionality. In order to prove this argument, we need the following lemma.

**Lemma 7.5.** Let  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a measurable function and  $F_n$  be the co-jump measure on  $\mathbb{R}^2$ . Then

$$\int_{\mathbb{R}^2} g(x_1, x_2) dF_n(x_1, x_2) = \int_B g(x_1, x_2) dF_n(x_1, x_2) = \int f_n(x) g(x, x) dx,$$

where  $F_n(A)$  is defined as in Definition 7.4.

*Proof.* First we use the step functions to prove the lemma. This extends by linearity and by taking limits for all measurable functions  $g$ . Indeed, we only need to show the lemma for the case of step functions. Let

$$g(x_1, x_2) = \sum_{k=1}^m a_k \mathbb{1}_{A_k}(x_1, x_2)$$

where  $A_k \subset A$  and  $\cup_{k=1}^m A_k = A$ . Therefore,

$$\begin{aligned} \int g(x_1, x_2) dF_n(x_1, x_2) &= \int \sum_{k=1}^m a_k \mathbb{1}_{A_k}(x_1, x_2) dF_n(x_1, x_2) \\ &= \sum_{k=1}^m a_k \int \mathbb{1}_{A_k}(x_1, x_2) dF_n(x_1, x_2) \end{aligned}$$

$$\begin{aligned}
 &= \sum_{k=1}^m a_k \int_{A_k} dF_n(x_1, x_2) \\
 &= \sum_{k=1}^m a_k \int \mathbb{1}_{A_k}(x, x) f_n(x) dx \\
 &= \int \sum_{k=1}^m a_k \mathbb{1}_{A_k}(x, x) f_n(x) dx \\
 &= \int g(x, x) f_n(x) dx,
 \end{aligned}$$

and the claim follows. □

Furthermore, we need to find a measure whose mass is bounded away from the origin but may explode around 0 and integrates  $\|x\|^2$ . In order to construct the co-jump measure with the above properties, we need to find an appropriate density function for the co-jumps measure  $F_n(A)$  so as to satisfy the following condition for  $r \in (1, 2)$  and  $\mathbf{x} = (x, x)$ :

$$\int_A \left(1 \wedge |x_1 x_2|^{r/2}\right) F_n(dx_1, dx_2) = \int (1 \wedge |x|^r) f_n(x) \mathbb{1}_A(x, x) dx < \infty.$$

Indeed, the following lemma implies condition (7.1) by choosing properly the density function of the co-jumps.

**Proposition 7.6.** *Let  $w_n$  be defined by (4.6) and  $r \in (1, 2)$ . Let there be the even functions  $h_n : \mathbb{R}^2 \rightarrow \mathbb{R}$  such that  $h_n(U) = \tilde{h}_n(U) \cdot \hat{h}_n(U)$ , where*

$$\tilde{h}_n(U) = \begin{cases} \sqrt{w_n} & |U| \leq U_n \\ \sqrt{w_n} e^{-|U-U_n|^3} & |U| > U_n, \end{cases}$$

and  $U_n = 2w_n^{-\frac{1}{2-r}}$ . Then, for any  $A$  from Definition 7.4, we have

$$\int (1 \wedge |x|^r) \mathbb{1}_A(x, x) F_n(dx) < \infty, \tag{7.4}$$

where  $F_n(A) = \int_{\mathbb{R}} \frac{|H_n(x)|}{x^2} \mathbb{1}_A(x, x) dx$  and  $H_n$  is the Fourier transform of  $h_n$ .

*Proof.* The mathematical tool used for the formation of the density function is the Fourier transform. Intuitively, we use the function  $h_n$  as a constant inside a fixed interval and which decays exponentially outside this interval. Notice that in the exponential we used the power of 3 because we need to differentiate two times, as we shall see later. Also,  $U_n$  is different from the one used in Theorem 4.2.

Notice that  $h_n$  has a range on  $\mathbb{R}$ , which is why we use the Fourier transform on  $\mathbb{R}$ . The pair of Fourier transform takes the following form:

$$H_n(x) = \mathcal{F}[h_n(U)](x) = \int_{-\infty}^{\infty} e^{iUx} h_n(U) dU$$

$$h_n(U) = \mathcal{F}^{-1}[H_n(x)](U) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iUx} H_n(x) dx$$

and the respective first derivatives will have the form

$$\begin{aligned} \partial_1 H_n(x) &= i\mathcal{F}[U h_n(U)](x) \\ \partial_1 h_n(U) &= (-i)\mathcal{F}^{-1}[x H_n(x)](U). \end{aligned}$$

For a thorough analysis of the Fourier transform the interested reader should refer to Bracewell (1986).

In the next step, the pair of Fourier transform will provide us with a proper and well-defined density function for the co-jump measure. First, we note that the  $\mathbb{L}^2$ -norm of  $h_n$  is bounded. Indeed,

$$\begin{aligned} \|h_n\|_{L^2(B)} &= \|\tilde{h}_n^2\|_{L^2(\mathbb{R})} = \left( \int_{-\infty}^{\infty} |\tilde{h}_n(U)|^4 dU \right)^{1/2} \\ &= w_n \left( \int_{|U| \leq U_n} dU + \int_{|U| > U_n} e^{-4|U-U_n|^3} dU \right)^{1/2} \\ &= \sqrt{2}w_n \left( U_n + \int_{U_n}^{\infty} e^{-4|U-U_n|^3} dU \right)^{1/2} \quad (7.5) \\ &= \sqrt{2}w_n \left( U_n + \int_0^{\infty} e^{-4K^3} dK \right)^{1/2} \\ &\leq Cw_n \sqrt{U_n} \leq Cw_n^{\frac{3-2r}{4-2r}}, \end{aligned}$$

for  $U_n \geq 1$ . In the last equality, substituting  $K = U - U_n$ , we get that  $\int_0^{\infty} e^{-4K^3} dK$  is bounded by a constant  $C$ . In addition,  $h_n$  is an  $\mathbb{L}^2$ -function. Applying the Plancherel theorem we deduce that

$$\|H_n\|_{\mathbb{L}^2(\mathbb{R})} = \sqrt{2\pi} \|h_n\|_{\mathbb{L}^2(\mathbb{R})} \leq Cw_n^{\frac{3-2r}{4-2r}}. \quad (7.6)$$

Similarly, we get a bound for the  $L^2$ -norm of the first derivative of  $h_n$

$$\|\partial_1 h_n\|_{\mathbb{L}^2(\mathbb{R})} \leq \frac{C}{\sqrt{2\pi}} w_n. \quad (7.7)$$

Moreover,  $\|H_n\|_{\mathbb{L}^1}$  is also bounded

$$\begin{aligned} \int_{-\infty}^{+\infty} |H_n(x)| dx &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{1+x^2}} \sqrt{1+x^2} |H_n(x)| dx \\ &\leq \left( \int_{-\infty}^{+\infty} \frac{1}{1+x^2} dx \right)^{1/2} \left( \int_{-\infty}^{+\infty} H_n^2(x)(1+x^2) dx \right)^{1/2} \\ &\leq \sqrt{\pi} \left( \int_{-\infty}^{+\infty} H_n^2(x) + x^2 H_n^2(x) dx \right)^{1/2} \quad (7.8) \\ &\leq \sqrt{\pi} (\|H_n\|_{\mathbb{L}^2} + \|\partial_1 h_n\|_{\mathbb{L}^2}) \\ &\leq \sqrt{\pi} C (1 + w_n^{\frac{3-2r}{4-2r}}) < \infty. \end{aligned}$$



We get the first inequality because of the Cauchy-Schwarz inequality. By means of (7.6) and (7.7), we get the last inequality. The  $\mathbb{L}^1$ -norm of  $H_n$  is bounded in  $n$  when  $r \leq 3/2$ .

At this point we are ready to define the co-jumps measures  $F_n(A)$  and  $G_n(A)$  in terms of the Fourier transform  $H_n(x)$ .

$$F_n(A) = \int_{\mathbb{R}} \frac{|H_n(x)|}{x^2} \mathbb{1}_A(x, x) dx \quad G_n(A) = F_n(A) + \int_{\mathbb{R}} \frac{H_n(x)}{x^2} \mathbb{1}_A(x, x) dx. \tag{7.9}$$

for any  $A \in \mathcal{B}(\mathbb{R}^2)$ . These measures satisfy the basic properties of Lévy measures. They are non-negative, integrate  $x^2$ , and may explode around zero since  $\frac{H_n(x)}{x^2} \rightarrow \infty$  as  $x \rightarrow 0$ .

It remains to prove (7.1) under this argumentation. Based on the above construction and  $A = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = x_2 = x\}$ , (7.1) transforms into:

$$\int (1 \wedge |x|^r) \frac{|H_n(x)|}{|x|^2} dx, \tag{7.10}$$

so we need to show that (7.10) is finite. Next we show how to bound from above  $|H_n(x)|$ .

$$\begin{aligned} H_n(x) &= \int_{-\infty}^{+\infty} h_n(U) \cos(Ux) dU + i \int_{-\infty}^{+\infty} h_n(U) \sin(Ux) dU \\ &= \int_{-\infty}^{\infty} h_n(U) \cos(Ux) dU. \end{aligned} \tag{7.11}$$

In the first line the second term is equal to zero since it is the integral of the product of an even and an odd function.

$$\begin{aligned} H_n(x) &= 2w_n \int_0^{U_n} \cos(Ux) dU + 2w_n \int_{U_n}^{\infty} e^{-2(U-U_n)^3} \cos(Ux) dU \\ &= 2w_n \left( \frac{\sin(U_n x)}{x} + \int_0^{\infty} e^{-2K^3} \cos((K + U_n)x) dK \right) \end{aligned} \tag{7.12}$$

Note that in the second line we substitute  $K = U - U_n$  and the integral is always bounded from above by a constant  $C$ . So,

$$|H_n(x)| \leq Cw_n \left( \frac{|\sin(U_n x)|}{|x|} + 1 \right). \tag{7.13}$$

On the sets  $\{|x| \leq \frac{1}{U_n}\}$ ,  $\{\frac{1}{U_n} < |x| \leq 1\}$ ,  $\{|x| > 1\}$  we deduce that

1.  $|x| \leq \frac{1}{U_n} \Rightarrow |U_n x| \leq 1 \Rightarrow |\sin(U_n x)| \leq |U_n x| \Rightarrow \frac{|\sin(U_n x)|}{|x|} + 1 \leq CU_n$ .
2.  $\frac{1}{U_n} < |x| \leq 1 \Rightarrow \frac{|\sin(U_n x)|}{|x|} + 1 = \frac{|\sin(U_n x)| + |x|}{|x|} \leq \frac{2}{|x|}$ .
3.  $|x| > 1 \Rightarrow \frac{|\sin(U_n x)|}{|x|} + 1 \leq 2$ .

In turn, we get that

$$|H_n(x)| \leq Cw_n \left( U_n \mathbf{1} \left( |x| \leq \frac{1}{U_n} \right) + \frac{1}{|x|} \mathbf{1} \left( \frac{1}{U_n} < |x| \leq 1 \right) + \mathbf{1}(|x| > 1) \right). \tag{7.14}$$

By splitting the integration domain into the sets  $\{|x| \leq \frac{1}{U_n}\}$ ,  $\{\frac{1}{U_n} < |x| \leq 1\}$ ,  $\{|x| > 1\}$  and recalling that  $r \in (1, 2)$ , condition (7.1) will take the form:

$$\begin{aligned} & \int (1 \wedge |x|^r) \frac{|H_n(x)|}{|x|^2} dx \\ & \leq Cw_n \int \frac{1 \wedge |x|^r}{|x|^2} U_n \mathbf{1} \left( |x| \leq \frac{1}{U_n} \right) dx \\ & + \int \frac{1 \wedge |x|^r}{|x|^2} \frac{1}{|x|} \mathbf{1} \left( \frac{1}{U_n} < |x| \leq 1 \right) dx + \int \frac{1 \wedge |x|^r}{|x|^2} \mathbf{1}(|x| > 1) dx \\ & \leq Cw_n(U_n^{2-r} + 1) \leq C. \end{aligned}$$

In light of the form of  $U_n$  the last inequality holds. Recall that  $U_n = 2/w_n^{1/(2-r)}$ . Therefore, (7.4) is satisfied, which implies condition (7.1), by which the proof is complete.  $\square$

Till now we constructed the co-jump measure, which satisfies (7.1), and the covariance matrices for the hypothesis test. So the triplets for the hypothesis test are now defined. Next step, we study the characteristic functions of the two processes, which will be useful later on in the proof of Property 2.

### 7.3. Characteristic functions of $\mathbf{X}_{1/n}$ and $\mathbf{Y}_{1/n}$

At this point, we study the processes  $\mathbf{X}, \mathbf{Y}$  for one observation at the moment  $t = \frac{1}{n}$ . We denote by  $\psi_n(\mathbf{u}), \phi_n(\mathbf{u})$  the characteristic functions of  $\mathbf{X}_{1/n}, \mathbf{Y}_{1/n}$  respectively, and by  $\eta_n(\mathbf{u}) = \psi_n(\mathbf{u}) - \phi_n(\mathbf{u})$  their difference. The characteristic triplet for each process is

$$\begin{aligned} \mathbf{X}_{1/n} & \sim (\mathbf{0}, \Sigma_{\mathbf{X}} \Sigma_{\mathbf{X}}^\top, G_n(d\mathbf{x})) \\ \mathbf{Y}_{1/n} & \sim (\mathbf{0}, \Sigma_{\mathbf{Y}} \Sigma_{\mathbf{Y}}^\top, F_n(d\mathbf{x})), \end{aligned}$$

where  $\Sigma_{\mathbf{X}} \Sigma_{\mathbf{X}}^\top = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$  and  $\Sigma_{\mathbf{Y}} \Sigma_{\mathbf{Y}}^\top = \begin{pmatrix} 2 & 1 + \pi w_n \\ 1 + \pi w_n & 1 \end{pmatrix}$ . Denote by  $C_{\mathbf{X}} = \Sigma_{\mathbf{X}} \Sigma_{\mathbf{X}}^\top$  and  $C_{\mathbf{Y}} = \Sigma_{\mathbf{Y}} \Sigma_{\mathbf{Y}}^\top$ . The characteristic functions will be defined as follows

$$\phi_n(\mathbf{u}) = \exp \left\{ -\frac{1}{2n} \left( \langle C_{\mathbf{Y}} \mathbf{u}, \mathbf{u} \rangle + 2\tilde{\phi}_n(\mathbf{u}) \right) \right\}$$

and

$$\psi_n(\mathbf{u}) = \exp \left\{ -\frac{1}{2n} \left( \langle C_{\mathbf{X}} \mathbf{u}, \mathbf{u} \rangle + 2\tilde{\phi}_n(\mathbf{u}) + 2\tilde{\psi}_n(\mathbf{u}) \right) \right\}.$$

Thus, we get

$$\psi_n(\mathbf{u}) = \psi_n(U) = \exp \left\{ -\frac{1}{2n} \left( 5U^2 + 2\tilde{\phi}_n(U) + 2\tilde{\psi}_n(U) \right) \right\} \tag{7.15}$$

$$\phi_n(\mathbf{u}) = \phi_n(U) = \exp \left\{ -\frac{1}{2n} \left( 5U^2 + 2\pi w_n U^2 + 2\tilde{\phi}_n(U) \right) \right\}. \tag{7.16}$$

We denote by

$$\tilde{\phi}_n(\mathbf{u}) = \tilde{\phi}_n(U) = \int_A (1 - \cos(Ux)) \frac{|H_n(x)|}{x^2} dx \tag{7.17}$$

and

$$\tilde{\psi}_n(\mathbf{u}) = \tilde{\psi}_n(U) = \int_A (1 - \cos(Ux)) \frac{H_n(x)}{x^2} dx \tag{7.18}$$

because of the form of co-jump measure (7.9) and the fact that  $H_n$  is an even function, its Fourier transform will be a real function. Also, recall the Fourier transform of the co-jump measure has support on  $\mathbb{R}$  and  $A$  is a subset of the diagonal. Moreover,

$$\langle C_{\mathbf{X}} \mathbf{u}, \mathbf{u} \rangle = 5U^2 \quad \text{and} \quad \langle C_{\mathbf{Y}} \mathbf{u}, \mathbf{u} \rangle = 5U^2 + 2\pi w_n U^2.$$

Next we bound from above (7.17) and (7.18). First, observe that

$$\tilde{\psi}_n''(U) = \int \cos(Ux) H_n(x) dx = 2\pi h_n(U),$$

since  $H_n$  is an even function. We consider the following two cases:

$$|U| \leq U_n : \tilde{\psi}_n''(U) = 2\pi w_n \Rightarrow \tilde{\psi}_n'(U) = 2\pi w_n U \Rightarrow \tilde{\psi}_n(U) = 2\pi w_n \frac{U^2}{2} \tag{7.19}$$

$$|U| \geq U_n : |\tilde{\psi}_n'(U)| \leq 2\pi w_n |U| \Rightarrow |\tilde{\psi}_n(U)| \leq 2\pi w_n \frac{U^2}{2}.$$

Now, concerning the  $\tilde{\phi}_n(U)$  we exploit the same arguments as before and by (7.8) we obtain

$$|\tilde{\phi}_n'(U)| \leq \int \frac{|x \sin(Ux)|}{x^2} |H_n(x)| dx \leq \int \frac{|U|x|}{x} |H_n(x)| dx \leq C|U| \left( 1 + w_n^{\frac{3-2r}{4-2r}} \right) \tag{7.20}$$

$$\tilde{\phi}_n(U) = \int \frac{1 - \cos(Ux)}{x^2} |H_n(x)| dx \leq \int \frac{(Ux)^2}{x^2} |H_n(x)| dx \leq CU^2 \left( 1 + w_n^{\frac{3-2r}{4-2r}} \right). \tag{7.21}$$

#### 7.4. Total variation distance

As we discussed in the first step,  $\mathbf{X}$  and  $\mathbf{Y}$  have a nonvanishing Gaussian part so that the variables  $\mathbf{X}_{1/n}$  and  $\mathbf{Y}_{1/n}$  have densities. We denote by  $f_{1/n}$  and

$g_{1/n}$  the densities of  $\mathbf{X}_{1/n}$  and  $\mathbf{Y}_{1/n}$  respectively. In order to establish a lower bound for our class with the rates (4.6), the last ingredient to be shown is Property 2, that the total variation distance between  $\mathbb{P}_{\mathbf{X}}$  and  $\mathbb{P}_{\mathbf{Y}}$  goes towards zero. Mathematically speaking, this formulates as

$$\text{TV}(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}}) = 2n \int |f_{1/n}(x) - g_{1/n}(x)| dx \rightarrow 0. \tag{7.22}$$

We denote as  $k_n = f_{1/n} - g_{1/n}$  the difference between the densities,  $\eta_n = \mathcal{F}k_n$  the Fourier transform of  $k_n$ , and  $\partial_1 \eta_n$  the first derivative of  $\eta_n$ . One would be tempted to use the following

$$\begin{aligned} \text{TV}(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}}) &= 2n \int |f_{1/n}(x) - g_{1/n}(x)| dx \\ &= 2n \int \left| \int e^{-iUx} (\phi_n(U) - \psi_n(U)) dU \right| dx \\ &\leq 2n \int \left( \int |e^{-iUx}| |\phi_n(U) - \psi_n(U)| dU \right) dx \\ &\leq 2n \int \left( \int |\phi_n(U) - \psi_n(U)| dU \right) dx. \end{aligned} \tag{7.23}$$

In the second equality we wrote the density function as the Inverse Fourier transform of its characteristic function. But the last integral is infinite. Hence, this procedure is not working for our goal. Since we want to prove that

$$2n \int |f_{1/n}(x) - g_{1/n}(x)| dx \rightarrow 0, \tag{7.24}$$

we know that the total variation distance between  $\mathbb{P}_{\mathbf{X}}$  and  $\mathbb{P}_{\mathbf{Y}}$  is not more than  $2n$  times  $\int |k_n(x)| dx$ . By using the same argument as for the Jacod and Reiß (2014) Theorem 3.1, by Cauchy-Schwarz inequality and Plancherel theorem, we obtain

$$\begin{aligned} \int |k_n(x)| dx &= \int \frac{1}{\sqrt{1+x^2}} (\sqrt{1+x^2}) |k_n(x)| dx \\ &\leq \sqrt{\pi} \left( \int (k_n^2(x) + x^2 k_n^2(x)) dx \right)^{1/2} \\ &\leq \sqrt{\pi} (\|\eta_n\|_{\mathbb{L}^2} + \|\partial_1 \eta_n\|_{\mathbb{L}^2})^{1/2}. \end{aligned}$$

In the second inequality we used the Cauchy-Schwarz inequality and in the last one we used Plancherel identity.

Thus, the only ingredient which remains to be shown is the following lemma in order to satisfy Property 2.

**Lemma 7.7.** *We show that*

$$4n^2 \int |\eta_n(U)|^2 dU \rightarrow 0 \quad \text{and} \quad 4n^2 \int |\partial_1 \eta_n(U)|^2 dU \rightarrow 0 \tag{7.25}$$

as  $n \rightarrow \infty$ .

*Proof.* First, we study the convergence of  $\eta_n(U)$ :

$$\begin{aligned} |\eta_n(U)| &= |\phi_n(U) - \psi_n(U)| = |\psi_n(U)| \left| 1 - \frac{\phi_n(U)}{\psi_n(U)} \right| \\ &= |\psi_n(U)| \left| 1 - \exp\left(\frac{1}{2n}(2\tilde{\psi}_n(U) - 2\pi w_n U^2)\right) \right| \\ &\leq |\psi_n(U)| \left| \frac{1}{2n}(2\tilde{\psi}_n(U) - 2\pi w_n U^2) \right|. \end{aligned}$$

The last inequality holds due to the fact that  $1 - e^{-x} \leq x$ . The difference of the characteristic functions vanishes for  $|U| \leq U_n$  because of  $2\tilde{\psi}_n(U) = 2\pi w_n U^2$  by (7.19).

By means of  $\tilde{\phi}_n(U) + \tilde{\psi}_n(U) \geq 0$ , we get that  $|\psi_n(U)| \leq e^{-\frac{U^2}{2n}}$  and  $|\phi_n(U)| \leq e^{-\frac{U^2}{2n}}$ . Thus, for some multiplicative constant  $C$

$$|\eta_n(U)| \leq C \frac{U^2 w_n}{2n} e^{-\frac{U^2}{2n}} \mathbf{1}_{\{|U| \geq U_n\}}. \tag{7.26}$$

We define by

$$A := \int_{\{|U| \geq U_n\}} U^4 e^{-\frac{U^2}{n}} dU. \tag{7.27}$$

Using Cauchy-Schwarz inequality we bound the integral (7.27) from above by

$$\begin{aligned} A &\leq 2 \left( \int_{U_n}^{\infty} U^5 e^{-\frac{U^2}{n}} dU \right)^{1/2} \cdot \left( \int_{U_n}^{\infty} U^3 e^{-\frac{U^2}{n}} dU \right)^{1/2} \\ &\leq C \left( n U_n^4 e^{-\frac{U_n^2}{n}} \right)^{1/2} \left( n U_n^2 e^{-\frac{U_n^2}{n}} \right)^{1/2}. \end{aligned} \tag{7.28}$$

The integrals in the first line on the r.h.s can be calculated exactly by calculus methods, and recalling  $U_n = 2/w_n^{1/(2-r)} = 2\sqrt{n \log n}$  we get

$$4n^2 \int |\eta_n(U)|^2 dU \leq C w_n^2 \int_{\{|U| \geq U_n\}} U^4 e^{-\frac{U^2}{n}} dU \leq C \frac{(\log n)^{3/2}}{n^{3/2}}.$$

The first part of the (7.25) follows. Now recall the form of the characteristic functions and their difference  $\eta_n = \psi_n(U) - \phi_n(U)$ . Therefore by (7.19), (7.20), and the fact that  $|\psi_n(U)| \leq e^{-\frac{U^2}{2n}}$ ,  $|\phi_n(U)| \leq e^{-\frac{U^2}{2n}}$  we get that

$$\begin{aligned} |\partial_1 \eta_n(U)| &\leq \frac{1}{n} \left( 2\pi w_n |U| |\phi_n(U)| + |\tilde{\phi}'_n(U) + 5U| |\eta_n(U)| + |\tilde{\psi}'_n(U)| |\psi_n(U)| \right) \\ &\leq \frac{1}{n} \left( 2\pi w_n |U| e^{-\frac{U^2}{2n}} + |\tilde{\phi}'_n(U) + 5U| e^{-\frac{U^2}{2n}} \frac{C w_n U^2}{2n} + 2\pi w_n |U| e^{-\frac{U^2}{2n}} \right) \\ &\leq \frac{1}{n} \left( 4\pi w_n |U| e^{-\frac{U^2}{2n}} + \left( 5|U| + C|U| (1 + w_n^{\frac{3-2r}{4-2r}}) \right) e^{-\frac{U^2}{2n}} \frac{C w_n U^2}{2n} \right) \\ &\leq C \frac{w_n |U|}{n} e^{-\frac{U^2}{2n}} \left( 1 + \left( w_n^{\frac{3-2r}{4-2r}} + 1 \right) \frac{U^2}{2n} \right). \end{aligned}$$

Therefore,

$$4n^2 \int_{U_n}^{\infty} |\partial_1 \eta_n(U)|^2 \leq C w_n^2 \int_{U_n}^{\infty} U^2 e^{-\frac{U^2}{n}} dU \\ + C \frac{w_n^2}{n^2} \left( 1 + w_n^{\frac{3-2r}{4-2r}} \right)^2 \int_{U_n}^{\infty} U^6 e^{-\frac{U^2}{n}} dU.$$

Now,  $\frac{1+w_n^{\frac{3-2r}{4-2r}}}{n}$  tends towards zero. Additionally, the integrals on the right hand side can be bounded again using Cauchy-Schwarz inequality, like integral A in (7.27). Following these, we can calculate the integrals through basic calculus methods exactly. As a result,

$$4n^2 \int_{U_n}^{\infty} |\partial_1 \eta_n(U)|^2 dU \leq C \frac{(\log n)^{5/2}}{n^{1/2}},$$

which also goes to zero as  $n \rightarrow \infty$  and the proof is completed.  $\square$

### *End proof of Theorem 4.3*

**Lower bound for the rate  $w_n = \frac{1}{\sqrt{n}}$  when  $r \in (0, 2)$ .** To prove this bound, it is enough to show that it holds on the subclass of all Brownian motions since  $\mathcal{B}_M \subset \mathcal{L}_M^r$ . Taken together with Lemma 7.2, this bound is achieved.

**Lower bound for the rate  $w_n = 1/(n \log n)^{\frac{2-r}{2}}$  when  $r \in (1, 2)$ .** The main steps of this proof are to show that Property 1 and Property 2 are satisfied. Now, with reference to Lemma 7.5 for the construction of co-jump measure, Proposition 7.6 and Lemma 7.7 we conclude the proof of Theorem 4.3.

## 8. Discussion

In this section we make some important remarks concerning the upper bound and the rates of convergence. First, we want to compare the efficiency of our estimator with the work of Mancini (2017) in which she considered at least one jump component of a two-dimensional Itô semimartingale with infinity variation. Mancini (2017) introduced the truncated realized covariance (TRC) as an estimator for co-integrated volatility. The proposed estimator is

$$\widehat{IC} = \sum_{i=1}^n \Delta_i X^{(1)} \mathbf{1}_{\{(\Delta_i X^{(1)})^2 \leq r_h\}} \Delta_i X^{(2)} \mathbf{1}_{\{(\Delta_i X^{(2)})^2 \leq r_h\}},$$

where  $r_h = h^{2u}$  is the truncation level with  $h = 1/n$ ,  $u \in (0, \frac{1}{2})$  and  $n \rightarrow \infty$ . It is clear that, when  $r_h \rightarrow 0$ , all jumps are asymptotically excluded. It is assumed that the two jump components have an activity index  $r_1, r_2$  where  $0 \leq r_1 \leq r_2 < 2$  and  $r_2 \geq 1$ . Notice by recalling Lemma 5.3 that in our case

we used the index  $r$  for the activity of co-jumps. In the following we use the notation “ $a \gg b$ ” when  $a$  is much greater than  $b$ .

The truncated estimator achieves the rate  $(1-\gamma)\sqrt{r_h}^{(1+\frac{r_2}{r_1}-r_2)}$  when  $r_1, r_2 \gg 1$  and the jump components are dependent. The TRC estimator reaches the rate  $hr_h^{-\frac{r_2}{2}}$  when the two jump components are independent and  $r_2 \gg r_1$  and the rate  $\sqrt{h}$  when  $r_1$  is small and  $r_2$  is close to 1. The parameter  $\gamma$  describes the dependence structure of the two jumps with  $\gamma \in [0, 1]$ . When  $\gamma = 0$  we have full dependency between the jump components, while  $\gamma = 1$  means independence between the jump components.

Finally, for a fair comparison with the spectral estimator, the TRC rates read in our notation using the following modifications. We assume  $\sqrt{r_h}$  to be approximately  $\frac{1}{\sqrt{n}}$  as truncation level, since  $u \in (0, \frac{1}{2})$ . The truncation level is not optimal, but the work of Figueroa-López and Mancini (2017) proposed an optimal way for the truncation level using mean and conditional mean square error for the case of a one-dimensional Itô semimartingale.

For reasons of simplicity we take into consideration only the two extreme cases of dependency. In the first two rows we assume  $\gamma = 0$ , i.e. fully dependency between the jump components, while in the last row we consider  $\gamma = 1$ , i.e. the jumps are totally independent. Further, we assume that the observations are drawn from a two-dimensional stable Lévy process. In the second row, by Lemma 5.3, we get the rate for the spectral estimator in case of dependent jumps. In the last row, using Example 5.4 we conclude the rate for the spectral estimator, since the last row is referring to independent jumps.

Rates of convergence		
$r, r_1, r_2 \in (0, 2], \gamma \in [0, 1]$	TRC estimator	Spectral estimator
$r_2$ close to 1, $r_1$ small, $r = 1, \gamma = 0$	$n^{-\frac{1}{2}}$	$n^{-\frac{1}{2}}$
$r_1 \leq r_2$ close to 2, $r = r_2, \gamma = 0$	$n^{-\frac{1}{2}(1+\frac{r_2}{r_1}-r_2)}$	$(n \log n)^{\frac{r-2}{2}}$
$r_1$ small, $r_2$ close to 2, $r = r_2, \gamma = 1$	$n^{\frac{r_2-2}{2}}$	$(n \log n)^{\frac{r-2}{2}}$

In the first row we observe that the truncated estimator establishes the same rates as the spectral estimator when both jump components have activity indices close to 1. Notwithstanding, the second row refers to the case of both jump components being of infinite variation. Then by Lemma 5.3, the spectral estimator will achieve the rate  $(n \log n)^{\frac{r-2}{2}}$ . Here, we notice that the truncated estimator achieves slower rates than the spectral estimator. The same happens in the last row, when we assume independent jumps and one component is of infinite variation. Then by Example 5.4 the spectral estimator will achieve the rate  $(n \log n)^{\frac{r-2}{2}}$ , which is faster than the rate of TRC estimator. However, the above comparison refers to the special case of a two-dimensional Lévy process. The TRC estimator still achieves its rates of convergence in case of a two-dimensional Itô semimartingale. In contrary, the spectral estimator reaches the optimal rates only in case of a two-dimensional Lévy process.

### 8.1. Numerical experiments

In this section we test our estimates with Monte-Carlo experiments.<sup>1</sup> This means that we first have to simulate the sample paths of a bivariate Lévy process on  $[0, 1]$ .

Section 6 of Cont and Tankov (2004) suggested various simulation algorithms for Lévy processes. We extend here Algorithms 6.6, 6.5, 6.3 to a bivariate setting. In addition, we use the generalized shot noise method for series representation of a two-dimensional Lévy process of infinite variation introduced by Rosinski (1990).

We now perform Monte-Carlo tests of our spectral estimate  $\widehat{C}_n^{12}(U_n)$ , comparing it to the Truncated Realized Covariance (TRC) estimate  $\widehat{IC}_T$  of Mancini (2017) for a two-dimensional Itô semimartingale. To provide a balanced comparison, we will draw our observations from a process  $X_t = B_t + J_t$ , where  $B_t$  is a two-dimensional Brownian motion and  $J_t$  is a two-dimensional jump process. Its jumps are driven by a two-dimensional  $r$ -stable process.  $X_t$  thus models a process with both diffusion and jump components. For Figure 1, in each run of our simulation, we will generate  $n = 1,000$  observations, corresponding to observations taken every  $1/1,000$  over a time interval  $[0, 1]$ .

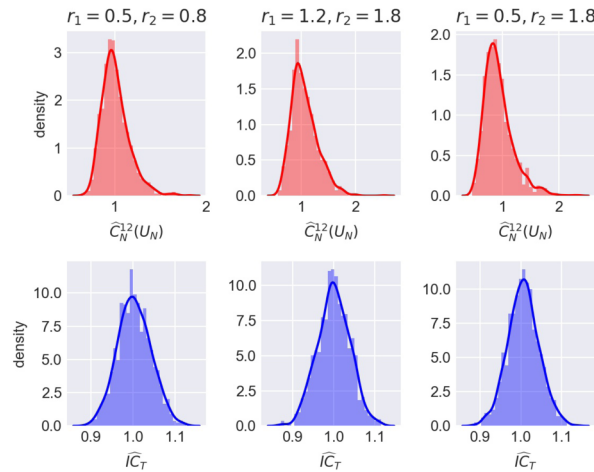


FIG 1. Simulated distributions of the estimates.

The estimates  $\widehat{C}_n^{12}(U_n)$  and  $\widehat{IC}_T$  depend on a number of parameters. We begin by considering the covariance matrix  $C = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$  for two correlated Brownian motions. In our simulations, the cointegrated volatility of  $X_t$  is equal to 1, and so we may choose the parameters accordingly. In our tests, we found the value  $M = 4.229$  worked well for bounding from above the jump activity in the case

<sup>1</sup>The interested reader can view the code at <https://github.com/KarinaPapayia/Co-integrated-volatility-multidimensional-Levy-processes>



of infinite variation jumps. In the case of  $\widehat{IC}_T$  we chose  $h = 1/1,000$ ,  $u = 0.387$ , and as truncation level  $r_h = (\frac{1}{1,000})^{2*0.387}$ . We found that this truncation level cuts the jumps bigger than  $(\frac{1}{1,000})^{2*0.387}$ , which means that almost all jumps were eliminated.

Figure 1 plots the simulated distributions of the estimates  $\widehat{C}_n^{12}(U_n)$  and  $\widehat{IC}_T$  together with a nonparametric fit, shown as a solid line. We can see that in every choice for  $r_1, r_2$ , the estimates are centered around 1, which is the expected theoretical cointegrated volatility. Figure 1 shows that the spectral estimator is more skewed, while the spread of the TRC estimator is much smaller.

Figure 2 plots the RMSEs of the estimates  $\widehat{C}_n^{12}(U_n)$ ,  $\widehat{IC}_T$  against different choices for the activity index of the co-jumps. We study the performance of the estimates under finite, moderate, and infinite activity of co-jumps. We can see that, as  $n$  grows, the RMSE of the spectral estimate  $\widehat{C}_n^{12}(U_n)$  is getting slightly smaller compared with the truncated estimate. However, we observe that the RMSEs of the truncated estimate  $\widehat{IC}_T$  are smaller compared with the spectral estimate when  $n = 1,000$ .

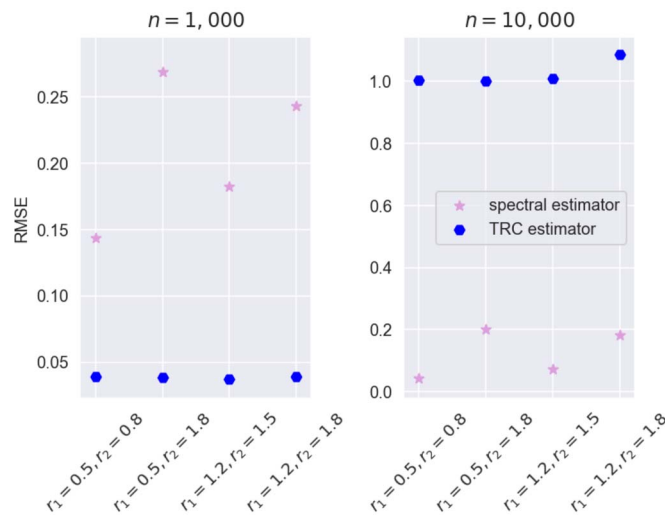


FIG 2. Simulated RMSEs of the estimates  $\widehat{C}_n^{12}(U_n)$  and  $\widehat{IC}_T$ .

We observe this behavior in Figure 2 for the truncated estimate  $\widehat{IC}_T$  because of our choice of truncation level, which is not an optimal. While the choice of  $u = 0.387$  for the truncation level  $r_h$  works well for  $n = 1,000$ , it does not work well when the number of observations is bigger, for example when  $n = 10,000$ . Figure 8 shows the sensitivity of the TRC estimator in case of IV jumps.

Figures 3, 4 give violin plots for the spectral estimate  $\widehat{C}_n^{12}(U_n)$  under a number of choices for the amount of observations  $n$  and the activity index for the co-jumps, whilst Figures 5, 6 show violin plots for the truncated estimate  $\widehat{IC}_T$  under the same settings. The number of observations varies from 1,000 to 10,000

by step 1,000.

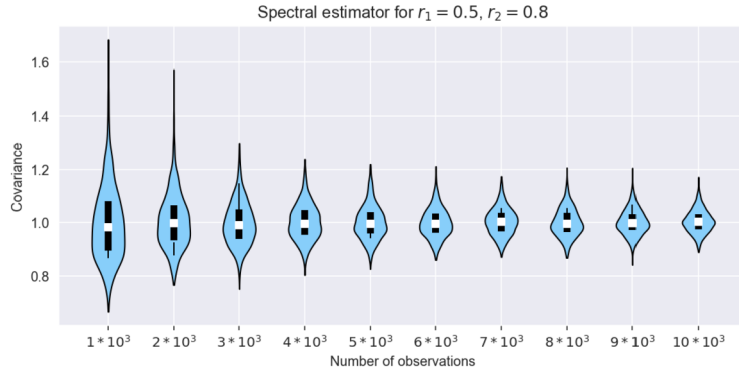


FIG 3. Violin plots for the estimates  $\hat{C}_n^{12}(U_n)$ .

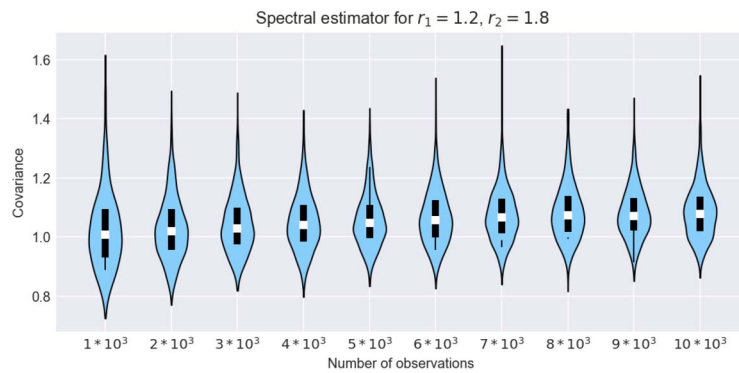


FIG 4. Violin plots for the estimates  $\hat{C}_n^{12}(U_n)$ .

In Figure 3, we used as an activity index for the jumps  $r_1 = 0.5$ ,  $r_2 = 0.8$ , while in Figure 4 we set  $r_1 = 1.2$  and  $r_2 = 1.8$ . In the case of  $r_1 = 1.2$  and  $r_2 = 1.8$ , we see that the estimation for the covariance slightly deviates from the center as  $n$  grows. We tuned the parameter  $M$  for  $n = 5,000$ .

Figures 5 and 6 show again that the truncated estimate  $\widehat{IC}_T$  slightly deviates from the center as  $n$  is larger than 6,000 or smaller than 5,000, an expected effect due to the choice of truncation level. We should mention here that the parameter  $u$  is tuned for  $n = 5,000$ . We expect this effect to disappear once the optimal choice for the threshold  $r_h$  is established.

Finally,  $U_n$  is the parameter which controls the frequency for our spectral estimate  $\hat{C}_n^{12}(U_n)$ .  $U_n$  depends on  $n, M, r$ . In view of the form (6.10) for  $U_n$  we can find a constant to multiply which will give us the optimal choice for  $U_n$ . The results will still hold. In fact Figure 7 shows that the spectral estimate for  $M > 3$ ,  $n = 5,000$  and  $r = 1.5$  is centered around the theoretical co-integrated

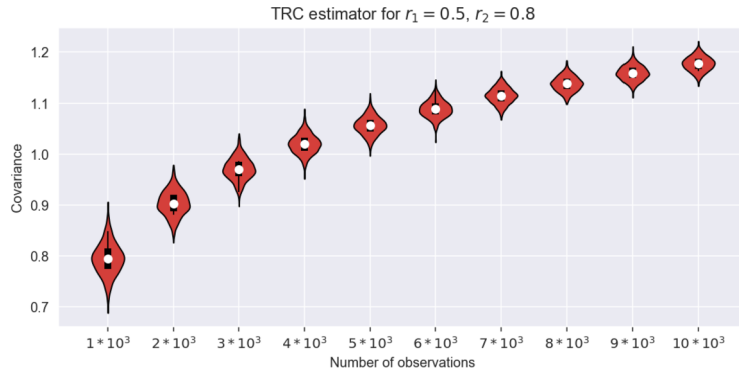


FIG 5. Violin plots for the estimates  $\widehat{IC}_T$ .

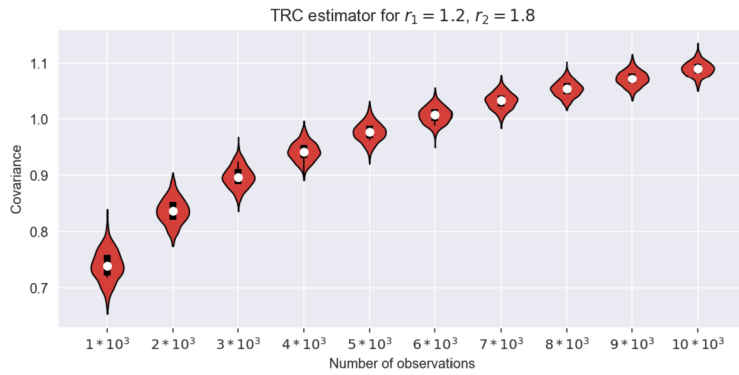


FIG 6. Violin plots for the estimates  $\widehat{IC}_T$ .

volatility  $C^{12}$ . Figure 7 shows a violin plot for the spectral estimate tuning up the parameter  $M$ .

Figure 8 shows that the truncated estimate  $\widehat{IC}_T$  is quite sensitive to the choice of threshold. Here, we used  $n = 5,000, r_1 = 1.2, r_2 = 1.5, h = 1/5,000$  and  $u$  varies from 0.410 to 0.419. Recall that  $r_h = h^{2u}$ . We notice that the threshold estimate deviates strongly from the theoretical co-integrated volatility. Figure 8 shows that the threshold estimate is centered around  $C^{12}$  only when  $u = 0,419$ . As a result, it is crucial to choose an optimal threshold for the truncated estimate.

**Acknowledgement**

The author is very grateful to an anonymous referee for many helpful questions and remarks that have led to considerable improvements, to Markus Reiß for stimulating comments and discussions, and to Cecilia Mancini for inspiring dis-

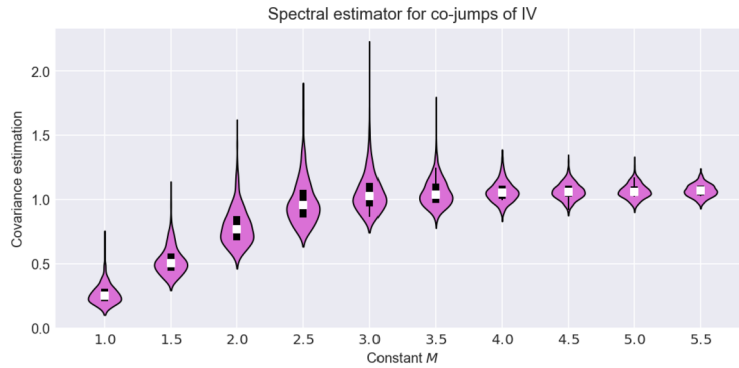


FIG 7. Tuning parameter  $M$  for the estimates  $\widehat{C}_n^{12}$ .

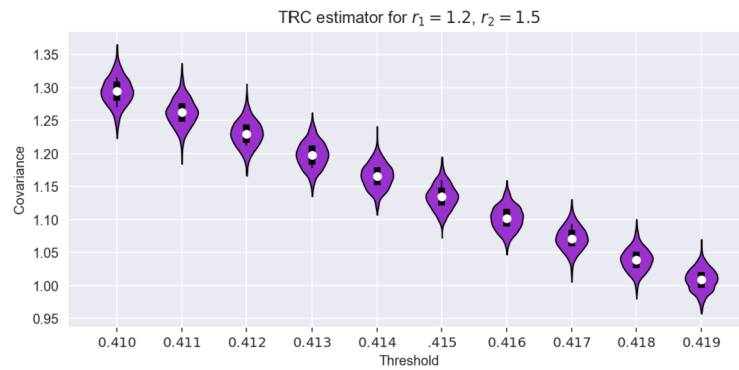


FIG 8. Tuning parameter  $r_h$  for the estimates  $\widehat{IC}_T$ .

cussion of co-jumps. This work was partially supported by Deutsche Forschungsgemeinschaft via IRTG 1792 *High Dimensional Nonstationary Time Series*.

## References

- Y. Aït-Sahalia and J. Jacod. Testing whether jumps have finite or infinite activity. *The Annals of Statistics*, 39(3):1689–1719, 2011. [MR2850217](#)
- T. Andersen and T. Bollerslev. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4):885–905, 1998.
- O. Barndorff-Nielsen and N. Shephard. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2):253–280, 2002. [MR1904704](#)
- O. Barndorff-Nielsen and N. Shephard. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37, 2004.

- O. Barndorff-Nielsen and N. Shephard. Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1):1–30, 2006. [MR2051439](#)
- D. Belomestny. Spectral estimation of the fractional order of a Lévy process. *The Annals of Statistics*, 38(1):317–351, 2010. [MR2589324](#)
- D. Belomestny and V. Panov. Abelian theorems for stochastic volatility models with application to the estimation of jump activity of volatility. *Stochastic Processes and their Applications*, 123(1):15–44, 2013a. [MR2988108](#)
- D. Belomestny and V. Panov. Estimation of the activity of jumps in time-changed Lévy models. *Electronic Journal of Statistics*, 7:2970–3003, 2013b. [MR3151759](#)
- D. Belomestny and M. Reiß. Spectral calibration of exponential Lévy models. *Finance and Stochastics*, 10(4):449–474, 2006. [MR2276314](#)
- D. Belomestny and M. Reiß. Estimation and Calibration of Lévy Models via Fourier Methods. *Lévy Matters IV*, 2015. [MR3364256](#)
- D. Belomestny and M. Trabs. Low-rank diffusion matrix estimation for high-dimensional time-changed Lévy processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 54(3):1583–1621, 2018. [MR3825892](#)
- M. Bibinger and M. Vetter. Estimating the quadratic covariation of an asynchronously observed semimartingale with jumps. *Annals of the Institute of Statistical Mathematics*, 67(4):707–743, 2015. [MR3357936](#)
- M. Bibinger and L. Winkelmann. Econometrics of co-jumps in high-frequency data with noise. *Journal of Econometrics*, 184(2):361–378, 2015. [MR3291008](#)
- M. Bibinger, N. Hautsch, P. Malec, and M. Reiß. Estimating the quadratic covariation matrix from noisy observations: Local method of moments and efficiency. *The Annals of Statistics*, 42(4):1312–1346, 2014. [MR3226158](#)
- R. Bracewell. *The Fourier transform and its applications*. McGraw-Hill, New York, 1986. [MR0924577](#)
- A. Bücher and M. Vetter. Nonparametric inference on Lévy measures and copulas. *The Annals of Statistics*, 41(3):1485–1515, 2013. [MR3113819](#)
- P. Carr, H. Geman, D. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. *The Journal of Business*, 75(2):305–332, 2002.
- K. Christensen, M. Podolskij, and M. Vetter. On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes. *Journal of Multivariate Analysis*, 120:59–84, 2013. [MR3072718](#)
- A. Coca. Efficient nonparametric inference for discretely observed compound poisson processes. *Probability Theory and Related Fields*, 170(1-2):475–523, 2018. [MR3748329](#)
- F. Comte and V. Genon-Catalot. Nonparametric estimation for pure jump Lévy processes based on high frequency data. *Stochastic Processes and their Applications*, 119(12):4088–4123, 2009. [MR2565560](#)
- F. Comte, C. Duval, and V. Genon-Catalot. Nonparametric density estimation in compound Poisson processes using convolution power estimators. *Metrika*, 77(1):163–183, 2014. [MR3152023](#)
- R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman &

- Hall/CRC, Boca Raton, 2004. [MR2042661](#)
- C. Duval and E. Mariucci. Compound Poisson approximation to estimate the Lévy density. *arXiv preprint* [arXiv:1702.08787](#), 2017.
- E. Eberlein and A. Papapantoleon. Symmetries and pricing of exotic options in Lévy models. *Exotic Option Pricing and Advanced Lévy Models*, 2005. [MR2343210](#)
- E. Figueroa-Lopez and C. Houdré. Nonparametric estimation for Lévy processes with a view towards mathematical finance. *arXiv preprint* [math/0412351](#), 2004. [MR2706761](#)
- J. Figueroa-López and C. Mancini. Optimum thresholding using mean and conditional mean square error. *arXiv preprint* [arXiv:1708.04339](#), 2017. [MR3906967](#)
- H. Geman. Pure jump Lévy processes for asset price modelling. *Journal of Banking & Finance*, 26(7):1297–1316, 2002.
- J. Jacod and M. Reiß. A remark on the rates of convergence for integrated volatility estimation in the presence of jumps. *The Annals of Statistics*, 42(3):1131–1144, 2014. [MR3224283](#)
- J. Kallsen and P. Tankov. Characterization of dependence of multidimensional Lévy processes using Lévy copulas. *Journal of Multivariate Analysis*, 97(7):1551–1572, 2006. [MR2275419](#)
- E. Lehmann and J. Romano. *Testing statistical hypotheses*. Springer Science & Business Media New York, 2006. [MR2135927](#)
- C. Mancini. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36(2):270–296, 2009. [MR2528985](#)
- C. Mancini. Truncated realized covariance when prices have infinite variation jumps. *Stochastic Processes and their Applications*, 127(6):1998–2035, 2017. [MR3646438](#)
- O. Martin and M. Vetter. The null hypothesis of common jumps in case of irregular and asynchronous observations. *arXiv preprint* [arXiv:1712.07159](#), 2017. [MR3788181](#)
- M. Neumann and M. Reiß. Nonparametric estimation for Lévy processes from low-frequency observations. *Bernoulli*, 15(1):223–248, 2009. [MR2546805](#)
- R. Nickl, M. Reiß, J. Söhl, and M. Trabs. High-frequency Donsker theorems for Lévy measures. *Probability Theory and Related Fields*, 164(1-2):61–108, 2016. [MR3449386](#)
- J. Rosinski. On series representations of infinitely divisible random vectors. *The Annals of Probability*, 18(1):405–430, 1990. [MR1043955](#)
- P. Tankov. Processus de Lévy en Finance: Problèmes Inverses et Modélisation de Dépendance. *Ph.D Thesis*, 2004.
- V. Todorov and G. Tauchen. Volatility jumps. *Journal of Business & Economic Statistics*, 29(3):356–371, 2011. [MR2848503](#)
- A. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Series in Statistics, New York, 2009. [MR2724359](#)
- L. Wu. Modeling financial security returns using Lévy processes. *Handbooks in operations research and management science*, 15:117–162, 2007.