# Statistical Analysis for Jumps in certain Semimartingale Models

# Dissertation

Zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt

der Fakultät für Mathematik
der Technischen Universität Dortmund

von

Dipl.-Math. Christian Palmes

Dortmund, März 2013

To my nieces and my nephew
Johanna, Julia, Claudia and Maximilian.

# Contents

# Introduction

This thesis has two objectives; the first (Chapter 1), *jump detection* in a *high frequency* setting and, second (Chapter 2 and 3), the analysis of certain *jump dependence structures* in a *low frequency* setting. All used models are semimartingales that are of special interest in finance since they provide a natural class for price processes, cf. Delbaen and Schachermayer [12, 13] and Shiryaev [41].

The three chapters are summarised below in order to give the reader a quick overview. More detailed introductions to the respective topics are given at the beginning of the respective chapters.

### Itô semimartingales and jumps

The classical Black-Scholes model is a time continuous model with continuous sample paths, i.e. it is not capable of modeling abrupt movements (jumps) in the financial market. Various investigations confirm that there is an essential difference whether one works with a continuous model or if one allows for jumps. It is therefore important to assess whether, for example, high frequency data (e.g. returns of a stock-price) should be modeled with a continuous or a jump process. Barndorff-Nielsen and Shephard [3] propose a nonparametric test to decide whether it suffices to use a continuous stochastic volatility model or if an additional jump term is required.

We investigate and develop a test based on classical extreme value theory for the same purpose. If there are no jumps and if the number of our observations tends to infinity on a fixed time interval, our test converges weakly to the Gumbel distribution. If there are jumps, the test converges to infinity. Simulation studies show that this technique results in a test with greater power than the Barndorff-Nielsen and Shephard test.

### Lévy processes and dependences

Let $X$ be a $d$-dimensional Lévy process with Lévy triplet $(\Sigma, \nu, \alpha)$ and $d \geqslant 2$. Given the low frequency observations $(X_t)_{t=1,\dots,n}$, the dependence structure of the jumps of $X$ is estimated. The Lévy measure $\nu$ describes the average jump behaviour in a time unit. Thus, the aim is to estimate the dependence structure of $\nu$ by estimating the Lévy copula $\mathfrak{C}$ of $\nu$, cf. Kallsen and Tankov [21].

We use the low frequency techniques presented in a one dimensional setting in Neumann and Reiß [29] and Nickl and Reiß [30] to construct a Lévy copula estimator $\widehat{\mathfrak{C}}_n$ based on the above $n$ observations. In doing so we prove

$$\widehat{\mathfrak{C}}_n \to \mathfrak{C}, \quad n \to \infty$$

uniformly on compact sets bounded away from zero with the convergence rate $(\log n)^{-\frac{1}{2}}$. This convergence holds under quite general assumptions, which also include Lévy triplets with $\Sigma \neq 0$ and $\nu$ of arbitrary Blumenthal-Getoor index $0 \leqslant \beta \leqslant 2$. Note that in a low frequency observation scheme, it is statistically difficult to distinguish between infinitely many small jumps and a Brownian motion part. Hence, the rather slow convergence rate $(\log n)^{-\frac{1}{2}}$ is not surprising.

In the complementary case of a compound Poisson process (CPP), an estimator $\widehat{C}_n$ for the copula $C$ of the jump distribution of the CPP is constructed under the same observation scheme. This copula $C$ is the analogue to the Lévy copula $\mathfrak{C}$ in the finite jump activity case, i.e. the CPP case. Here we establish

$$\widehat{C}_n \to C, \quad n \to \infty$$

with the convergence rate $n^{-\frac{1}{2}}$ uniformly on compact sets bounded away from zero.

Both convergence rates are optimal in the sense of Neumann and Reiß [29].

### Copula relations in compound Poisson processes

We investigate in multidimensional compound Poisson processes (CPP) the relation between the dependence structure of the jump distribution and the dependence structure of the respective components of the CPP itself. For this purpose the asymptotic $\lambda t \to \infty$ is considered, where $\lambda$ denotes the intensity and $t$ the time point of the CPP. For modeling the dependence structures we are using the concept of copulas. We prove that the copula of a CPP converges under quite general assumptions to a specific Gaussian copula, depending on the underlying jump distribution.

Let $F$ be a $d$-dimensional jump distribution ($d \geqslant 2$), $\lambda > 0$ and let $\Psi(\lambda, F)$ be the distribution of the corresponding CPP with intensity $\lambda$ at the time point 1 . Next denote the operator which maps a $d$-dimensional distribution on its copula as $\mathcal{T}$. The starting point for our investigation was the validity of the equation

$$\mathcal{T}(\Psi(\lambda, F)) = \mathcal{T}(\Psi(\lambda, \mathcal{T}F)).$$

Our asymptotic theory implies that this equation is, in general, not true.

### Danksagungen

Die Anfertigung dieser Arbeit wäre ohne die ideelle und fachliche Unterstützung vieler Freunde und Kollegen zumindest in dieser Form nicht möglich gewesen. Es ist mir daher ein besonderes Anliegen, ihnen an dieser Stelle meinen Dank kund zu tun.

An erster Stelle möchte ich mich bei meiner Betreuerin Frau Prof. Dr. Jeannette H.C. Woerner für die Überlassung des Themas und ausgezeichnete Betreuung meiner Doktorarbeit bedanken. Neben ihrer kontinuierlichen Unterstützung und Motivation habe ich sehr unsere konstruktiven Diskussionen geschätzt, aus denen für diese Arbeit entscheidende Anregungen entstanden sind. Zusätzlich ermutigte sie mich stets, an unterschiedlichen Konferenzen teilzunehmen, woraus viele gewinnbringende Diskussionen mit einem weit gefächerten Publikum entstanden.

Die Unterstützung durch die DFG-SNF Forschergruppe 916 *Statistical Regularization and Qualitative Constraints* schuf eine dreijährige exzellente Forschungsumgebung, wofür ich mich ebenfalls an dieser Stelle bedanken möchte.

Herrn Prof. Dr. Christoph Kühn danke ich, da er mich auf diese Stelle erst aufmerksam gemacht hat. Ferner bedanke ich mich bei Herrn Dr. Johannes Schmidt-Hieber, Herrn Dr. Mathias Vetter, Herrn Prof. Dr. Markus Reiß und Herrn Prof. Dr. Alexander Lindner für viele anregende und motivierende Fachgespräche hinsichtlich dieser Forschungsarbeit.

Des Weiteren ist es mir ein besonderes Anliegen, mich bei allen meinen Kollegen von der TU Dortmund für eine sehr schöne harmonische und freundschaftliche Zeit zu bedanken. Auf diese Weise enstand eine angenehme inspirierende Umgebung, von der ich sehr profitieren konnte.

Ebenso möchte ich meiner Familie danken. Meine Eltern haben durch ihre Zuverlässigkeit und Beständigkeit meinen gesamten akademischen Werdegang auf ein solides Fundament gestellt und somit erst ermöglicht. Meinen Geschwistern danke ich für ihre vielfältigen Ratschläge bei der Fertigstellung der Promotionsarbeit. Frau Proletina Girova danke ich für ihre Unterstützung und wünsche ihr für die Zukunft alles Gute.

Frau Dr. Elisabeth Justicz danke ich für die englische Korrektur dieser Arbeit.

# Chapter 1

# Itô semimartingales and jumps

In this chapter we have given high frequency observations from an underlying Itô semimartingale process with a possible additional jump component. Hence, the model

$$dY_t = \sigma_t \, dW_t + b_t \, dt + dJ_t, \quad t \geqslant 0$$

is used, where $W$ denotes a Brownian motion, $\sigma > 0$ a volatility process, $b$ a drift coefficient and $J$ a possible additional (external) jump process. The goal of this chapter is the development and investigation of a statistical test which, based on given high frequency observations, decides whether $Y$ possesses jumps ($J \neq 0$) or not ($J = 0$). $J$ is, for example, a compound Poisson process. $dW_t$ denotes the differential of an Itô integral. Korn and Korn [23] provide an introduction to the Itô integral from a financial point of view. Further literature related to this subject can be found in the following publications [20, 22, 32, 35].

*Section 1* outlines some technical preparations. Note that this section uses the rather restrictive Assumptions 1.1.1. These restrictions have the advantage of allowing us that we are able to derive some useful inequalities in Proposition 1.1.4, by using a certain moment estimation technique. Compare the proof of Proposition 1.1.4 for details. The restrictive assumptions are finally weakened to the more natural Assumptions 1.2.3 in the second section. They basically claim that each path $\omega$ of the volatility $t \mapsto \sigma_t(\omega)$ has to be Hölder continuous and that the drift coefficient has to be pathwise bounded.

*Section 2* contains the first important result of this chapter: Theorem 1.2.5 states the convergence to the Gumbel distribution under the weakened Assumptions 1.2.3 of the statistic

$$a_{n^2}(T_n - b_{n^2}), \quad n \to \infty$$

in the case $J = 0$, i.e. absence of external jumps. Compare Corollary 1.2.2 for the notations.

We show in *Section 6* that this statistic converges to infinity, if there are any external jumps, i.e. $J \neq 0$. Thus, this statistic can be used as a test to distinguish between the existence and non-existence of external jumps. The divergence behaviour of this statistic in the case of existing external jumps is also investigated in Section 6. Here the two main results are Theorems 1.6.1 and 1.6.3. The former covers infinite activity jumps ($J$ is a general semimartingale), while the latter, Theorem 1.6.3, investigates the case of finite activity jumps ($J$ is, for example, a compound Poisson process). Of course, different divergence types and rates are proven, depending on the jump activity. In particular, two different spot volatility estimators (see Definition 1.1.3) are investigated as to their nature in the presence of external jumps. For their finite sample behaviour, Proposition

1.6.5 should also be noted.

*Section 7* explores the finite sample behaviour of the above mentioned jump test using numerical Matlab computer simulations. A comparison between this test and the test of Barndorff-Nielsen, Shephard [3] is made.

A further interesting question is the behaviour of the test statistic, if the volatility process $\sigma$ itself possesses any jumps. *Section 5* is devoted to this point with the main results captured in Theorem 1.5.3 and Corollary 1.5.7. If $\sigma$ possesses only small jumps that are not larger than a certain boundary and there are no external jumps, Theorem 1.5.3 states that the statistic still converges to the Gumbel distribution. On the other hand, Corollary 1.5.7 provides an counter example of where the statistic diverges to infinity in the presence of an oversized jump. As a consequence, the above mentioned boundary is sharp. In this section, for technical simplification we assume that $\sigma$ is independent of $W$ and $b = 0$. The technical benefits of these simplifications are discussed in detail in *Section 4*.

As a by-product of the techniques developed in the Sections 1 and 2, we get a spot volatility estimator, which converges uniformly and pathwise to the underlying true volatility process $\sigma$. This spot volatility estimator is investigated in *Section 3*. The optimal convergence rate of this estimator is derived in Theorem 1.3.1 and a simple relationship between the convergence rate and the smoothness of the volatility process $\sigma$ is proven, cf. (1.42) and (1.43).

After completion of our research on this chapter, we became aware of a paper by Lee and Mykland [26] that investigates a similar test statistic for jump detection as we found. The separate research of an independent team with the same objective, emphasizes the importance and relevance of this research topic. Lee and Mykland [26], however, approached the topic from an economic point of view. We complement these investigations by an mathematically point of view with a much more subtle analysis. In order to give a transparent overview, a concise comparison between their results in [26] and ours in Chapter 1 follows:

The convergence to the Gumbel distribution, i.e. the case of no external jumps, is covered by Lee and Mykland [26] under the following assumptions.

**Assumptions of Lee and Mykland [26].** *It holds for every $\epsilon > 0$ the asymptotic*

*(i)*

$$\sup_{0 \leqslant i < n} \sup_{\frac{i}{n} \leqslant t \leqslant \frac{i+1}{n}} \left| b_t - b_{\frac{i}{n}} \right| = O_P\left( n^{-(\frac{1}{2}-\epsilon)} \right), \tag{1.1}$$

*(ii)*

$$\sup_{0 \leqslant i < n} \sup_{\frac{i}{n} \leqslant t \leqslant \frac{i+1}{n}} \left| \sigma_t - \sigma_{\frac{i}{n}} \right| = O_P\left( n^{-(\frac{1}{2}-\epsilon)} \right)$$

*for $n \to \infty$.*

Recall for two families of random variables $(X_n)_{n \in \mathbb{N}}$ and $(Y_n)_{n \in \mathbb{N}}$ with $P(X_n = 0) = 0$, $n \in \mathbb{N}$ the notation

$$Y_n = O_P(X_n) \quad :\iff \quad \forall \delta > 0 \ \exists K > 0 : \sup_{n \in \mathbb{N}} P\left( \left| \frac{Y_n}{X_n} \right| \geqslant K \right) \leqslant \delta.$$

This is basically the same as claiming that the volatility process $\sigma$ and the drift coefficient $b$ have to be $\alpha$-Hölder continuous in a stochastic sense for every $\alpha < \frac{1}{2}$. Lemma 1 in [26] corresponds to Theorem 1.2.5 of this thesis. Here we prove the convergence in distribution to the Gumbel distribution under Assumptions 1.2.3 which we restate below to facilitate a direct comparison:

**Assumptions 1.2.3.** *Let the volatility $\sigma$ be pathwise Hölder continuous, strictly positive and let the drift $b$ be pathwise bounded. This means that there are two functions*

$$\alpha : \Omega \to (0,1] \quad \textit{and} \quad K : \Omega \to (0,\infty)$$

*such that*

$$|\sigma_t(\omega) - \sigma_s(\omega)| \leqslant K(\omega)|t-s|^{\alpha(\omega)}, \quad 0 \leqslant s,t \leqslant 1, \quad \omega \in \Omega$$

*and*

$$|\sigma_t(\omega)| \vee |b_t(\omega)| \leqslant K(\omega), \quad 0 \leqslant t \leqslant 1, \quad \omega \in \Omega.$$

Note that our assumptions are much weaker since we do not have any Hölder continuity restriction as in (1.1) concerning the drift coefficient, and $t \mapsto \sigma_t(\omega)$ simply has to be $\alpha(\omega)$-Hölder continuous for every path. This means that $\alpha = \alpha(\omega)$ does not have to be arbitrarily close to $\frac{1}{2}$ and even $\inf_{\omega\in\Omega} \alpha(\omega) = 0$ is possible.

Concerning external jumps $J$, [26] requires $J$ to have the special shape

$$dJ_t = R_t\, dQ_t, \quad t \geqslant 0$$

where $Q$ is a counting process and $R$ is the jump size. Furthermore, $Q$ has to be independent of $W$, and the jump sizes $(R_t)_t$ have to be i.i.d. and independent of $W$ and $Q$, compare the beginning of the first paragraph in [26]. Concerning external jumps, we have proven more general results: Theorem 1.6.1 is a statement regarding general semimartingales. In particular, external jumps of infinite activity are possible. Moreover Theorem 1.6.3 treats finite activity jumps, but without any dependency and distributive restrictions as in [26].

Finally, the interesting case where the volatility process $\sigma$ itself possesses any jumps is not dealt with by [26]. This is covered in Section 1.5 in our investigations. As a by-product of our analysis, we also get optimal convergence rates of an interpolation based spot volatility estimator in Section 1.3, cf. Theorem 1.3.1.

## 1.1 Technical preparations

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, P)$ be a filtered stochastic space and let $(W_t)$,$(\sigma_t)$ and $(b_t)$ respectively be $(\mathcal{F}_t)$-adapted processes on this space. We assume here and throughout the chapter that the usual hypotheses are fulfilled. $W$ denotes a standard Brownian motion, $\sigma$ the volatility process, and $b$ the drift coefficient of the Itô semimartingales

$$Y_t \stackrel{\text{def}}{=} \int_0^t \sigma_s\, dW_s, \quad \check{Y}_t \stackrel{\text{def}}{=} \int_0^t \sigma_s\, dW_s + B_t \stackrel{\text{def}}{=} \int_0^t \sigma_s\, dW_s + \int_0^t b_s\, ds, \quad 0 \leqslant t \leqslant 1.$$

Throughout the chapter, we use the check notation $\check{Y}$, if we want to emphasize that this process has a possible non-vanishing drift term. Note further that without loss of generality we always consider the unit interval $[0,1]$ instead of an interval $[0,T]$ for some $T > 0$. This is only due to a simpler notation.

In this section, the following assumptions (weakened to more natural ones in the following sections of this chapter) are made:

**Assumptions 1.1.1.** *There are three global constants* $0 < V \leqslant K < \infty$ *and* $0 < \alpha \leqslant 1$*, such that we have for every* $\omega \in \Omega$

*(i)* $\sigma$ *is pathwise bounded, i.e.*

$$V \leqslant \sigma_t(\omega) \leqslant K, \quad 0 \leqslant t \leqslant 1,$$

*(ii)* $\sigma$ *is Hölder continuous, i.e.*

$$|\sigma_s(\omega) - \sigma_t(\omega)| \leqslant K|t-s|^\alpha, \quad 0 \leqslant s,\, t \leqslant 1,$$

*(iii)* $t \mapsto b_t(\omega)$ *is Lebesgue measurable and*

$$|b_t(\omega)| \leqslant K, \quad 0 \leqslant t \leqslant 1.$$

*We remark that any dependence structure between* $W$*,* $\sigma$ *and* $b$ *is allowed.*

In this chapter we are concerned with high-frequency statistical inference. To be more precise we are working with the observation vector

$$\left(Y_{\frac{0}{N}}(\omega_0), Y_{\frac{1}{N}}(\omega_0), \ldots, Y_{\frac{N}{N}}(\omega_0)\right) \in \mathbb{R}^{N+1} \tag{1.2}$$

for large $N \in \mathbb{N}$ and $\omega_0 \in \Omega$. If there is a drift term, we observe of course the respective variant with a check in (1.2). For our statistical experiment, this means, that we only observe one possible realisation, i.e. we see only one trajectory $\omega \in \Omega$ which we denote in (1.2) with $\omega_0$. Furthermore, we are not able to see the full trajectory, but only finite many dates at equidistant distance on the timeline. In (1.2) the sampling positions are $0, \frac{1}{N}, \ldots, 1$. This means that we observe exactly $N+1$ dates on an equidistant grid and want to infer using this information.

In our approach we set $N = n^2$ for $n \in \mathbb{N}$ and interpret the above described grid of observations as a double grid, in the sense that every observation position $\frac{l}{n^2}$, $l = 0, \ldots, n^2 - 1$ is uniquely represented as

$$\frac{l}{n^2} = \frac{kn+j}{n^2}, \quad 0 \leqslant k, j < n. \tag{1.3}$$

So the grid on the unit interval separates to two scales: the rough one, which is indexed in (1.3) by $k$, and the finer one, which is indexed in (1.3) by $j$. With this two-scale grid separation in mind, the following abbreviations become natural:

**Abbreviations 1.1.2.** Let $0 \leqslant k, j < n$ and $0 \leqslant t \leqslant 1$. We denote with

$$t_{k,j} = \frac{kn+j}{n^2}$$

a point on our equidistant two-scale grid. The index $k$ stands for the rough scale and $j$ for the fine one. Next, we approximate the volatility $\sigma$ with a step function via

$$\overline{\sigma}_t(\omega) = \sum_{k=0}^{n^2-1} \mathbb{1}_{(\frac{k}{n^2}, \frac{k+1}{n^2}]}(t)\sigma_{\frac{k}{n^2}}(\omega).$$

In doing so we can write the Itô integral

$$\overline{\mathrm{Y}}_t = \int_0^t \overline{\sigma}_s \, dW_s \tag{1.4}$$

in the closed-form expression

$$\overline{\mathrm{Y}}_{t_{k,j}} = \int_0^{t_{k,j}} \overline{\sigma}_s \, dW_s = \sum_{l=0}^{kn+j-1} \sigma_{\frac{l}{n^2}} (W_{\frac{l+1}{n^2}} - W_{\frac{l}{n^2}}).$$

Next, regarding the increments of the finer scale, we set

$$\Delta W_{k,j} = W_{t_{k,j}+\frac{1}{n^2}} - W_{t_{k,j}}, \quad \Delta Y_{k,j} = Y_{t_{k,j}+\frac{1}{n^2}} - Y_{t_{k,j}}, \quad \Delta \overline{\mathrm{Y}}_{k,j} = \overline{\mathrm{Y}}_{t_{k,j}+\frac{1}{n^2}} - \overline{\mathrm{Y}}_{t_{k,j}},$$
$$\Delta B_{k,j} = B_{t_{k,j}+\frac{1}{n^2}} - B_{t_{k,j}}$$

and note that (1.4) yields

$$\Delta \overline{\mathrm{Y}}_{k,j} = \sigma_{t_{k,j}} \Delta W_{k,j}.$$

We further set

$$Z_{k,j} = n \Delta W_{k,j}$$

and observe that $(Z_{k,j})_{0 \leqslant k,j < n}$ is a family of i.i.d. $\mathrm{N}(0,1)$ distributed random variables, since $W$ is a Brownian motion. Finally, we set some abbreviations concerning the volatility:

$$\sigma_{k,j} = \sigma_{t_{k,j}}, \quad \sigma_k = \sigma_{k,0}, \quad \epsilon_{k,j} = \sigma_{k,j} - \sigma_k.$$

All abbreviations up to here are also defined with a check in an analogue manner.

Crucial in what follows are the following two high-frequency spot-volatility estimators:

**Definition 1.1.3.** *Set for* $0 \leqslant k < n$

$$\widehat{\sigma}_k^2 \stackrel{\text{def}}{=} \frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} |\Delta Y_{k,j}| \Delta Y_{k,j+1}| \quad \text{(without drift)},$$
$$\check{\sigma}_k^2 \stackrel{\text{def}}{=} \frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} |\Delta \check{Y}_{k,j}||\Delta \check{Y}_{k,j+1}| \quad \text{(with drift)}.$$

Note that the factor $\frac{\pi}{2}$ in the above Definition 1.1.3 results from $E|\mathrm{N}(0,1)| = \sqrt{\frac{2}{\pi}}$. Now we can formulate our first essential proposition. In essence, this proposition makes the Itô integral for our purpose mathematically feasible. The technique found in the proof is based on moment estimates which follow from the Itô formula. This technique is motivated by the martingale moment inequalities of Millar [27] and Novikov [31]. (See also Karatzas and Shreve [22][Chapter 3, Proposition 3.26]). Finally, the Markov inequality and elementary considerations yield the result:

**Proposition 1.1.4.** *Set*

$$c_{k,j} \stackrel{\text{def}}{=} \sigma_{k,j}\sigma_{k,j+1} - \sigma_k^2, \quad 0 \leqslant k,j < n, \quad j < n-1,$$
$$F_k \stackrel{\text{def}}{=} \widehat{\sigma}_k^2 - \frac{\pi}{2(n-1)} \sum_{j=0}^{n-2} (\sigma_k^2 + c_{k,j}) |Z_{k,j}||Z_{k,j+1}|, \quad 0 \leqslant k < n,$$
$$H_{k,j} \stackrel{\text{def}}{=} n\Delta Y_{k,j} - (\sigma_k + \epsilon_{k,j}) Z_{k,j}, \quad 0 \leqslant k,j < n.$$

*For every fixed $m \in \mathbb{N}$, there are two constants $d_1, d_2 > 0$, such that the inequalities*

$$P(|F_k| \geqslant \epsilon) \leqslant \frac{d_1}{n^{2\alpha m-1}\epsilon^m}, \quad 0 \leqslant k < n, \tag{1.5}$$

$$P(|H_{k,j}| \geqslant \epsilon) \leqslant \frac{d_2}{n^{4\alpha m}\epsilon^{2m}}, \quad 0 \leqslant k, j < n. \tag{1.6}$$

*are true for every $\epsilon > 0$ and $n \in \mathbb{N}$. Furthermore, we have the trivial relations*

$$|\epsilon_{k,j}| \leqslant \frac{K}{n^\alpha}, \quad |c_{k,j}| \leqslant \frac{3K^2}{n^\alpha}, \quad 0 \leqslant k, j < n. \tag{1.7}$$

*Proof.* We separate the proof into four steps. In the first step, we establish the already mentioned moment inequalities. We prove (1.5) and (1.6) in the second and third step respectively. Finally, (1.7) is verified in the fourth step.

STEP 1. *Establishing the moment inequalities.*

$$E|\Delta Y_{k,j} - \Delta\, \overline{\mathrm{Y}}_{k,j}\,|^{2m} = E\left|\int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}} (\sigma_s - \overline{\sigma}_s)\, dW_s\right|^{2m} \leqslant \frac{v_m K^{2m}}{n^{4\alpha m+2m}}, \quad m \geqslant 1 \tag{1.8}$$

with

$$v_m \overset{\text{def}}{=} m^m(2m-1)^m.$$

To prove this, we apply Itô's formula to $f : x \mapsto x^{2m} \in \mathcal{C}^2(\mathbb{R})$ and

$$M_t \overset{\text{def}}{=} \int_{t_{k,j}}^t (\sigma_s - \overline{\sigma}_s)\, dW_s, \quad t_{k,j} \leqslant t \leqslant 1.$$

$M_t$ is obviously a continuous martingale and $f''(x) = 2m(2m-1)x^{2m-2}$. This yields

$$M^{2m}_{t_{k,j}+\frac{1}{n^2}} = f\left(M_{t_{k,j}+\frac{1}{n^2}}\right) = \int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}} f'(M_s)\, dM_s + m(2m-1)\int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}} M_s^{2m-2}\, d\langle M\rangle_s$$

where $\langle . \rangle$ denotes the angle bracket, c.f. Jacod Shiryaev [20][p. 38 ff.]. Taking the expectation on each side and using

$$\langle M\rangle_s = \int_{t_{k,j}}^s |\sigma_r - \overline{\sigma}_r|^2\, dr, \quad t_{k,j} \leqslant s \leqslant 1$$

results in

$$\begin{aligned}
EM^{2m}_{t_{k,j}+\frac{1}{n^2}} &= m(2m-1)\int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}} EM_s^{2m-2}|\sigma_s - \overline{\sigma}_s|^2\, ds \\
&\leqslant m(2m-1)\frac{K^2}{n^{4\alpha}}\int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}} EM_s^{2m-2}\, ds \\
&\leqslant m(2m-1)\frac{K^2}{n^{4\alpha+2}} EM^{2m-2}_{t_{k,j}+\frac{1}{n^2}}, \quad m \geqslant 1.
\end{aligned}$$

Note for the last inequality that $(M_s^{2m-2}, \mathcal{F}_s)_s$ is a submartingale. So we get (1.8) by iteration.

Consider further

$$E|\Delta Y_{k,j}|^{2m} = E\left|\int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}} \sigma_s \, dW_s\right|^{2m} \leqslant \frac{v_m K^{2m}}{n^{2m}}, \quad m \geqslant 1 \tag{1.9}$$

with the same argumentation as in (1.8). So the Cauchy-Schwarz inequality together with (1.8) and (1.9) yields

$$E\left(|\Delta Y_{k,j}|^m |\Delta Y_{k,j+1} - \Delta \overline{\mathrm{Y}}_{k,j+1}|^m\right) \leqslant \frac{v_m K^{2m}}{n^{2\alpha m+2m}}, \quad m \geqslant 1.$$

Due to

$$E|\Delta \overline{\mathrm{Y}}_{k,j+1}|^{2m} = \sigma_{k,j+1}^{2m} E|\Delta W_{k,j+1}|^{2m} \leqslant \frac{K^{2m}\mu_{2m}}{n^{2m}}, \quad \mu_{2m} \overset{\text{def}}{=} E|\mathrm{N}(0,1)|^{2m}, \quad m \geqslant 1,$$

we also get

$$E\left(|\Delta \overline{\mathrm{Y}}_{k,j+1}|^m |\Delta Y_{k,j} - \Delta \overline{\mathrm{Y}}_{k,j}|^m\right) \leqslant \frac{(v_m\mu_{2m})^{\frac{1}{2}} K^{2m}}{n^{2\alpha m+2m}}, \quad m \geqslant 1.$$

STEP 2. *Proof of (1.5).* Note in this regard

$$\begin{aligned} &\left|\widehat{\sigma}_k^2 - \frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} |\Delta \overline{\mathrm{Y}}_{k,j}||\Delta \overline{\mathrm{Y}}_{k,j+1}|\right| \qquad (1.10)\\ \leqslant\ & \frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} \big||\Delta Y_{k,j}\Delta Y_{k,j+1}| - |\Delta \overline{\mathrm{Y}}_{k,j}\,\Delta \overline{\mathrm{Y}}_{k,j+1}|\big| \\ \leqslant\ & \frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} |\Delta Y_{k,j}\Delta Y_{k,j+1} - \Delta \overline{\mathrm{Y}}_{k,j}\,\Delta \overline{\mathrm{Y}}_{k,j+1}| \\ \leqslant\ & \frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} |\Delta Y_{k,j}||\Delta Y_{k,j+1} - \Delta \overline{\mathrm{Y}}_{k,j+1}| + |\Delta \overline{\mathrm{Y}}_{k,j+1}||\Delta Y_{k,j} - \Delta \overline{\mathrm{Y}}_{k,j}|. \end{aligned}$$

Moreover, we have together with the results of the first step and the Markov inequality

$$\begin{aligned} & P\left(\frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} |\Delta Y_{k,j}||\Delta Y_{k,j+1} - \Delta \overline{\mathrm{Y}}_{k,j+1}| + |\Delta \overline{\mathrm{Y}}_{k,j+1}||\Delta Y_{k,j} - \Delta \overline{\mathrm{Y}}_{k,j}| \geqslant \epsilon\right) \qquad (1.11)\\ \leqslant\ & \sum_{j=0}^{n-2}\left[P\left(|\Delta Y_{k,j}||\Delta Y_{k,j+1} - \Delta \overline{\mathrm{Y}}_{k,j+1}| \geqslant \frac{1}{2(n-1)}\frac{2(n-1)}{\pi n^2}\epsilon\right)\right. \\ & \left. + P\left(|\Delta \overline{\mathrm{Y}}_{k,j+1}||\Delta Y_{k,j} - \Delta \overline{\mathrm{Y}}_{k,j}| \geqslant \frac{1}{2(n-1)}\frac{2(n-1)}{\pi n^2}\epsilon\right)\right] \\ \leqslant\ & n\frac{v_m K^{2m}\pi^m n^{2m}}{n^{2\alpha m+2m}\epsilon^m} + n\frac{(v_m\mu_{2m})^{\frac{1}{2}} K^{2m}\pi^m n^{2m}}{n^{2\alpha m+2m}\epsilon^m} \\ =\ & \frac{d_1}{n^{2\alpha m-1}\epsilon^m}, \end{aligned}$$

and $d_1$ is defined over the last equality. Now (1.5) follows from the inequalities in (1.10) and (1.11) together with

$$\begin{aligned}\frac{\pi n^2}{2(n-1)}\sum_{j=0}^{n-2}|\Delta\,\overline{\mathrm{Y}}_{k,j}\,||\Delta\,\overline{\mathrm{Y}}_{k,j+1}\,| &= \frac{\pi n^2}{2(n-1)}\sum_{j=0}^{n-2}\sigma_{k,j}\sigma_{k,j+1}|\Delta W_{k,j}||\Delta W_{k,j+1}| \\ &= \frac{\pi}{2(n-1)}\sum_{j=0}^{n-2}(\sigma_k^2+c_{k,j})|Z_{k,j}||Z_{k,j+1}|\end{aligned}$$

and step 2 is accomplished.

STEP 3. *Proof of (1.6).* Consider for this purpose the decomposition

$$\begin{aligned}n\Delta Y_{k,j} &= n\Delta\,\overline{\mathrm{Y}}_{k,j} + n(\Delta Y_{k,j}-\Delta\,\overline{\mathrm{Y}}_{k,j}) \\ &= \sigma_{k,j}Z_{k,j}+n(\Delta Y_{k,j}-\Delta\,\overline{\mathrm{Y}}_{k,j}) \\ &= (\sigma_k+\epsilon_{k,j})Z_{k,j}+n(\Delta Y_{k,j}-\Delta\,\overline{\mathrm{Y}}_{k,j})\end{aligned} \tag{1.12}$$

and write, using the results in step 1 and the Markov inequality:

$$P(|n(\Delta Y_{k,j}-\Delta\,\overline{\mathrm{Y}}_{k,j})|\geqslant\epsilon)\leqslant\frac{n^{2m}v_mK^{2m}}{\epsilon^{2m}n^{4\alpha m+2m}}=\frac{d_2}{\epsilon^{2m}n^{4\alpha m}}$$

where $d_2$ is defined over the last equality. Together with (1.12) this proves step 3.

STEP 4. *Proof of (1.7).* The first inequality in (1.7) follows directly from our Assumptions 1.1.1. The second inequality follows from the consideration

$$|c_{k,j}|=|\sigma_k(\epsilon_{k,j}+\epsilon_{k,j+1})+\epsilon_{k,j}\epsilon_{k,j+1}|\leqslant\frac{2K^2}{n^\alpha}+\frac{K^2}{n^{2\alpha}}\leqslant\frac{3K^2}{n^\alpha},\quad 0\leqslant k,j<n,$$

and the proof is complete. □

Next, we prove two $P$-a.s. convergences. These follow from the above Proposition 1.1.4, Corollary A.2 and Proposition A.3 in the appendix, and several Borel-Cantelli arguments. We use the abbreviations

$$\max_k f_k \stackrel{\text{def}}{=} \max_{0\leqslant k<n} f_k,\quad \max_{k,j} g_{k,j}\stackrel{\text{def}}{=}\max_{0\leqslant k,j<n} g_{k,j},\quad n\in\mathbb{N}$$

for functions $f$ resp. $g$ with the domains $\{0\leqslant k<n\}$ resp. $\{0\leqslant k,j<n\}$. In addition $f$ and $g$ may depend on $\omega\in\Omega$.

**Proposition 1.1.5.** *It holds for* $\gamma<\alpha$

$$n^\gamma\max_{k,j}|n\Delta Y_{k,j}-\sigma_kZ_{k,j}|\to 0,\quad P\text{-a.s.} \tag{1.13}$$

*and if additionally* $\gamma<\frac{1}{2}$, *i.e.* $\gamma<\alpha\wedge\frac{1}{2}$, *we have, furthermore,*

$$n^\gamma\max_k|\widehat{\sigma}_k^2-\sigma_k^2|\to 0,\quad P\text{-a.s.} \tag{1.14}$$

*Proof.* We separate the proof into two steps one for each equation.

STEP 1. *Proof of (1.13).* We write, using the notation in Proposition 1.1.4,

$$n^\gamma \max_{k,j} |n\Delta Y_{k,j} - \sigma_k Z_{k,j}| = n^\gamma \max_{k,j} |\epsilon_{k,j} Z_{k,j} + H_{k,j}| \leqslant K n^{\gamma-\alpha} \max_{k,j} |Z_{k,j}| + n^\gamma \max_{k,j} |H_{k,j}|.$$

Using Proposition A.3, (1.6) and the Markov inequality, this yields for $n$ sufficiently large

$$\begin{aligned}
& P(n^\gamma \max_{k,j} |n\Delta Y_{k,j} - \sigma_k Z_{k,j}| \geqslant \epsilon) \\
\leqslant\ & P\left(K n^{\gamma-\alpha} \max_{k,j} |Z_{k,j}| \geqslant \frac{\epsilon}{2}\right) + P\left(n^\gamma \max_{k,j} |H_{k,j}| \geqslant \frac{\epsilon}{2}\right) \\
\leqslant\ & \frac{2^m K^m n^{m(\gamma-\alpha)}}{\epsilon^m} E(\max_{k,j} |Z_{k,j}|)^m + \sum_{0 \leqslant k,j < n} P\left(|H_{k,j}| \geqslant \frac{\epsilon n^{-\gamma}}{2}\right) \\
\leqslant\ & n^{m(\gamma-\alpha)} \frac{2^m K^m}{\epsilon^m} \left(2^m (\log n^2)^{\frac{m}{2}} + 2m!\right) + n^2 d_2 \frac{2^{2m} n^{2\gamma m}}{\epsilon^{2m}} n^{-4\alpha m} \\
=\ & O\left(\frac{1}{n^2}\right),
\end{aligned} \tag{1.15}$$

for any fixed $\epsilon > 0$ and $m$ large enough. The last equation holds because of $\gamma < \alpha$ since

$$m(\gamma - \alpha) \downarrow -\infty, \quad 2\gamma m - 4\alpha m \downarrow -\infty, \quad m \uparrow \infty.$$

Next we set

$$A_l \stackrel{\text{def}}{=} \limsup_n \left\{ n^\gamma \max_{k,j} |n\Delta Y_{k,j} - \sigma_k Z_{k,j}| \geqslant \frac{1}{l} \right\}, \quad l \in \mathbb{N}.$$

Then, because of $\sum_{n>0} n^{-2} < \infty$ and (1.15), the Borel-Cantelli lemma yields

$$P(A_l) = 0, \quad l \in \mathbb{N}.$$

Since

$$\left(\bigcup_{l=1}^\infty A_l\right)^c = \{n^\gamma \max_{k,j} |n\Delta Y_{k,j} - \sigma_k Z_{k,j}| \to 0\},$$

(1.13) is proven.

STEP 2. *Proof of (1.14).* In view of Proposition 1.1.4 we write

$$\begin{aligned}
& \max_k |\hat{\sigma}_k^2 - \sigma_k^2| \\
=\ & \max_k \left| \left( \frac{\pi}{2(n-1)} \sum_{j=0}^{n-2} |Z_{k,j}||Z_{k,j+1}| - 1 \right) \sigma_k^2 + \frac{\pi}{2(n-1)} \sum_{j=0}^{n-2} c_{k,j} |Z_{k,j}||Z_{k,j+1}| + F_k \right| \\
\leqslant\ & K^2 \max_k \left| \frac{\pi}{2(n-1)} \sum_{j=0}^{n-2} |Z_{k,j}||Z_{k,j+1}| - 1 \right| + \frac{3\pi K^2}{2(n-1)n^\alpha} \sum_{j=0}^{n-2} \max_k |Z_{k,j}| \max_k |Z_{k,j+1}| \\
& + \max_k |F_k| \\
\stackrel{\text{def}}{=}\ & K^2 \max_k \eta_1^{(k)} + \eta_2 + \max_k |F_k|.
\end{aligned} \tag{1.16}$$

Corollary A.2 and the Markov inequality imply

$$P\left(n^\gamma \max_k \eta_1^{(k)} > \epsilon\right) \leqslant \sum_{k=0}^{n} P\left(\eta_1^{(k)} > \epsilon n^{-\gamma}\right) \leqslant nC\epsilon^{-2m} n^{2m\gamma} n^{-m}.$$

Using $\gamma < \frac{1}{2}$ we have

$$2m\gamma - m \downarrow -\infty, \quad m \uparrow \infty,$$

so that the same Borel-Cantelli argument as in step 1 yields, for $\gamma < \frac{1}{2}$, the convergence

$$n^\gamma K^2 \max_k \eta_1^{(k)} \to 0 \quad P\text{-a.s.} \tag{1.17}$$

In order to apply the same argument to $\eta_2$, we use Proposition A.3 and note that the random variables $\max_k |Z_{k,j}|$ and $\max_k |Z_{k,j+1}|$ are independent for every $j = 0, \ldots, n-2$. This yields

$$\begin{aligned} & P\left(n^{\gamma-\alpha} \sum_{j=0}^{n-2} \max_k |Z_{k,j}| \max_k |Z_{k,j+1}| \geqslant \epsilon(n-1)\right) \\ \leqslant \quad & \sum_{j=0}^{n-2} P\left(\max_k |Z_{k,j}| \max_k |Z_{k,j+1}| \geqslant \epsilon n^{\alpha-\gamma}\right) \\ \leqslant \quad & (n-1)\left(2^m (\log n)^{\frac{m}{2}} + 2m!\right)^2 \epsilon^{-m} n^{m(\gamma-\alpha)}. \end{aligned}$$

Since $\gamma < \alpha$, we have

$$1 + m(\gamma - \alpha) \downarrow -\infty, \quad m \uparrow \infty$$

and Borel-Cantelli yields, for $\gamma < \alpha$, the convergence

$$n^\gamma \eta_2 \to 0 \quad P\text{-a.s.} \tag{1.18}$$

Finally, we consider (1.5) and estimate

$$P\left(n^\gamma \max_k |F_k| > \epsilon\right) \leqslant \sum_{k=0}^{n-1} P\left(|F_k| > \epsilon n^{-\gamma}\right) \leqslant n\epsilon^{-m} n^{\gamma m} d_1 n^{1-2\alpha m}.$$

Again, since $\gamma < \alpha$

$$2 + \gamma m - 2\alpha m \downarrow -\infty, \quad m \uparrow \infty,$$

so that Borel-Cantelli once more yields, for $\gamma < \alpha$, the convergence

$$n^\gamma \max_k |F_k| \to 0 \quad P\text{-a.s.} \tag{1.19}$$

Lastly, an application of (1.17), (1.18) and (1.19) to (1.16) proves the claim (1.14). □

Our next proposition is the first one which acts with the drift term. Note that this proposition is itself interesting because it provides a possibility to estimate the spot volatility pathwise and uniform on a grid. We provide a refinement of this result in Section 1.3 where we investigate a spot volatility estimator, constructed via linear interpolation from the gridpoints.

**Proposition 1.1.6.** *We have for all* $\gamma < \alpha \wedge \frac{1}{2}$ *the convergence*

$$n^\gamma \max_{k,j} |\breve{\sigma}_k - \sigma_k| \to 0 \quad P\text{-a.s.} \tag{1.20}$$

*Proof.* First, we estimate

$$\begin{aligned}
|\breve{\sigma}_k^2 - \hat{\sigma}_k^2| &= \left| \frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} |\Delta Y_{k,j}||\Delta Y_{k,j+1}| - |\Delta \breve{Y}_{k,j}||\Delta \breve{Y}_{k,j+1}| \right| \\
&\leqslant \frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} |\Delta Y_{k,j}(\Delta Y_{k,j+1} - \Delta \breve{Y}_{k,j+1}) - \Delta \breve{Y}_{k,j+1}(\Delta \breve{Y}_{k,j} - \Delta Y_{k,j})| \\
&\leqslant \frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} |\Delta Y_{k,j}||\Delta B_{k,j+1}| + (|\Delta Y_{k,j+1}| + |\Delta B_{k,j+1}|)|\Delta B_{k,j}| \\
&\leqslant \frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} (|\Delta Y_{k,j} - \Delta \overline{\mathrm{Y}}_{k,j}| + |\Delta B_{k,j}|)|\Delta B_{k,j+1}| + |\Delta Y_{k,j+1} - \Delta \overline{\mathrm{Y}}_{k,j+1}||\Delta B_{k,j}| \\
&\quad + \frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} |\Delta \overline{\mathrm{Y}}_{k,j}||\Delta B_{k,j+1}| + |\Delta \overline{\mathrm{Y}}_{k,j+1}||\Delta B_{k,j}| \\
&\leqslant \frac{\pi K}{2(n-1)} \sum_{j=0}^{n-2} |\Delta Y_{k,j} - \Delta \overline{\mathrm{Y}}_{k,j}| + \frac{K}{n^2} + |\Delta Y_{k,j+1} - \Delta \overline{\mathrm{Y}}_{k,j+1}| \qquad (1.21) \\
&\quad + \frac{\pi K^2}{2(n-1)} \sum_{j=0}^{n-2} |\Delta W_{k,j}| + |\Delta W_{k,j+1}|
\end{aligned}$$

and write using (1.8)

$$\begin{aligned}
& P\left(n^\gamma \max_k \left(\frac{1}{n-1} \sum_{j=0}^{n-2} |\Delta Y_{k,j} - \Delta \overline{\mathrm{Y}}_{k,j}| + |\Delta Y_{k,j+1} - \Delta \overline{\mathrm{Y}}_{k,j+1}|\right) \geqslant \epsilon\right) \qquad (1.22) \\
\leqslant\ & \sum_{k=0}^{n-1} \sum_{j=0}^{n-1} P\left(|\Delta Y_{k,j} - \Delta \overline{\mathrm{Y}}_{k,j}| \geqslant \frac{\epsilon}{2} n^{-\gamma}\right) \\
\leqslant\ & n^2 \frac{2^{2m} v_m K^{2m} n^{2m\gamma}}{n^{4\alpha m + 2m} \epsilon^{2m}}.
\end{aligned}$$

Further, we write

$$\begin{aligned}
n^\gamma \max_k \frac{1}{2(n-1)} \sum_{j=0}^{n-2} (|\Delta W_{k,j}| + |\Delta W_{k,j+1}|) &= n^{\gamma-1} \max_k \frac{1}{2(n-1)} \sum_{j=0}^{n-2} (|Z_{k,j}| + |Z_{k,j+1}|) \\
&\leqslant n^{\gamma-1} \max_{k,j} |Z_{k,j}|.
\end{aligned}$$

This implies together with Proposition A.3 the inequalities

$$\begin{aligned}
P\left(n^\gamma \max_k \frac{1}{2(n-1)} \sum_{j=0}^{n-2} (|\Delta W_{k,j}| + |\Delta W_{k,j+1}|) \geqslant \epsilon\right) &\leqslant P(\max_{k,j} |Z_{k,j}| \geqslant \epsilon n^{1-\gamma}) \qquad (1.23) \\
&\leqslant \frac{n^{m(\gamma-1)}}{\epsilon^m} \left(2^m (\log n^2)^{\frac{m}{2}} + 2m!\right).
\end{aligned}$$

Finally, (1.22) and (1.23) yield, together with (1.21) and the same Borel-Cantelli argument as in Proposition 1.1.5, for all $\gamma < 1$ the convergence

$$n^\gamma \max_k |\breve{\sigma}_k^2 - \widehat{\sigma}_k^2| \to 0 \quad P\text{-a.s.} \tag{1.24}$$

Now (1.20) follows from (1.24) together with (1.14) and

$$n^\gamma \max_k |\breve{\sigma}_k - \sigma_k| = n^\gamma \max_k \frac{|\breve{\sigma}_k^2 - \sigma_k^2|}{\breve{\sigma}_k + \sigma_k} \leqslant V^{-1} n^\gamma \left( \max_k |\breve{\sigma}_k^2 - \widehat{\sigma}_k^2| + \max_k |\widehat{\sigma}_k^2 - \sigma_k^2| \right).$$

Recall the definition of $V$ in Assumption 1.1.1 (i). □

## 1.2 Extreme value theory and jumps

**Proposition 1.2.1.** *Let Assumptions 1.1.1 hold and let $\gamma < \alpha \wedge \frac{1}{2}$ be constant. Then, it holds*

$$n^\gamma \max_{k,j} \left| \frac{n\Delta \breve{Y}_{k,j}}{\breve{\sigma}_k} - Z_{k,j} \right| \to 0 \quad P\text{-a.s.} \tag{1.25}$$

*Proof.* Write

$$\max_{k,j} \left| \frac{n\Delta \breve{Y}_{k,j}}{\breve{\sigma}_k} - Z_{k,j} \right| \quad \leqslant \quad \max_{k,j} \left| \frac{n\Delta \breve{Y}_{k,j}}{\breve{\sigma}_k} - \frac{n\Delta Y_{k,j}}{\sigma_k} \right| + V^{-1} \max_{k,j} |n\Delta Y_{k,j} - \sigma_k Z_{k,j}|.$$

Based on (1.13) it suffices to prove for $\gamma < \alpha \wedge \frac{1}{2}$ the convergence

$$n^\gamma \max_{k,j} \left| \frac{n\Delta \breve{Y}_{k,j}}{\breve{\sigma}_k} - \frac{n\Delta Y_{k,j}}{\sigma_k} \right| \to 0 \quad P\text{-a.s.}$$

In order to prove this, note that

$$n^{1-\delta} \max_{k,j} |\Delta Y_{k,j}| \to 0 \quad P\text{-a.s.} \tag{1.26}$$

holds for every $\delta > 0$. Consider for this purpose

$$n^{1-\delta} \max_{k,j} |\Delta Y_{k,j}| \leqslant n^{-\delta} \max_{k,j} |n\Delta Y_{k,j} - \sigma_k Z_{k,j}| + K n^{-\delta} \max_{k,j} |Z_{k,j}| \tag{1.27}$$

together with (1.13), Proposition A.3 and our standard Borel-Cantelli argument. Next, write

$$\begin{aligned} \left| \frac{n\Delta \breve{Y}_{k,j}}{\breve{\sigma}_k} - \frac{n\Delta Y_{k,j}}{\sigma_k} \right| &= n \frac{|\Delta \breve{Y}_{k,j} \sigma_k - \Delta Y_{k,j} \breve{\sigma}_k|}{\breve{\sigma}_k \sigma_k} \\ &= n \frac{|(\Delta \breve{Y}_{k,j} - \Delta Y_{k,j})\sigma_k + \Delta Y_{k,j}(\sigma_k - \breve{\sigma}_k)|}{\breve{\sigma}_k \sigma_k} \\ &\leqslant \frac{\frac{K^2}{n} + n|\Delta Y_{k,j}||\sigma_k - \breve{\sigma}_k|}{\breve{\sigma}_k \sigma_k} \end{aligned} \tag{1.28}$$

Note further that we have because of (1.20) the convergence

$$\max_k |\breve{\sigma}_k \sigma_k - \sigma_k^2| \leqslant K \max_k |\breve{\sigma}_k - \sigma_k| \to 0 \quad \text{P-a.s.}$$

This yields

$$\min_k |\breve{\sigma}_k \sigma_k| \geqslant \min_k (\sigma_k^2 - |\breve{\sigma}_k \sigma_k - \sigma_k^2|) \geqslant V^2 - \max_k |\breve{\sigma}_k \sigma_k - \sigma_k^2| \to V^2 > 0 \quad \text{P-a.s.} \tag{1.29}$$

Now choose $\delta > 0$ such that $\gamma + \delta < \alpha \wedge \frac{1}{2}$. Then observe

$$\begin{aligned} n^\gamma \max_{k,j} \left( \frac{K^2}{n} + n|\Delta Y_{k,j}||\sigma_k - \breve{\sigma}_k| \right) &\leqslant K^2 n^{\gamma-1} + n^{1-\delta} \max_{k,j} |\Delta Y_{k,j}| n^{\gamma+\delta} \max_k |\sigma_k - \breve{\sigma}_k| \\ &\to 0 \quad P\text{-a.s.} \end{aligned} \tag{1.30}$$

where we get the convergence from (1.20) and (1.26). Finally, (1.28), (1.29) and (1.30) prove this proposition. □

**Corollary 1.2.2.** *Set*

$$a_N = \sqrt{2 \log N}, \quad b_N \overset{\text{def}}{=} a_N - \frac{\log(\log N) + \log(4\pi)}{2\sqrt{2 \log N}}, \quad N \in \mathbb{N} \tag{1.31}$$

*and define a statistic $T_n$ via*

$$T_n = n \max_{k,j} \left( \frac{\Delta \breve{Y}_{k,j}}{\breve{\sigma}_k} \right), \quad n \in \mathbb{N}. \tag{1.32}$$

*Then, under the Assumptions 1.1.1 it holds the weak convergence*

$$a_{n^2}(T_n - b_{n^2}) \overset{d}{\to} \mathcal{G}, \quad n \to \infty$$

*where $\mathcal{G}$ denotes the Gumbel distribution, i.e. the unique distribution with the cumulative distribution function $x \mapsto e^{-e^{-x}}$, $x \in \mathbb{R}$.*

*Proof.* It is known from extreme value theory that

$$a_{n^2}(\max_{k,j} Z_{k,j} - b_{n^2}) \overset{\text{d}}{\to} \mathcal{G}, \quad n \to \infty,$$

compare Example 1.1.7 in Haan, Ferreira [17]. This proves, together with Proposition 1.2.1, Slutsky's theorem and Lemma A.7, our claim. □

Note that the notation of $T_n$ does not distinguish between the presence or absence of a non-vanishing drift term, i.e. there is no check notation. This is only due to a simpler notation and should not cause any ambiguity.

Up to this point it was assumed that the Assumptions 1.1.1 hold. However, these assumptions have the disadvantage that the two constants $V$ and $K$ are chosen globally, i.e. they are independent of the path $\omega$. Particularly (ii) in Assumptions 1.1.1 is a huge restriction. A path dependent counterpart to the Assumptions 1.1.1 is, therefore, formulated which even allows $\alpha$ to depend on the path.

**Assumptions 1.2.3.** *Let the volatility $\sigma$ be pathwise Hölder continuous, strictly positive and let the drift $b$ be pathwise bounded. This means that there are two functions*

$$\alpha : \Omega \to (0,1] \quad \text{and} \quad K : \Omega \to (0,\infty),$$

*such that*

$$|\sigma_t(\omega) - \sigma_s(\omega)| \leqslant K(\omega)|t-s|^{\alpha(\omega)}, \quad 0 \leqslant s,t \leqslant 1, \quad \omega \in \Omega \tag{1.33}$$

*and*

$$|\sigma_t(\omega)| \vee |b_t(\omega)| \leqslant K(\omega), \quad 0 \leqslant t \leqslant 1, \quad \omega \in \Omega. \tag{1.34}$$

*Furthermore, we claim $t \mapsto b_t(\omega)$ to be Lebesgue measurable for all $\omega \in \Omega$.*

**Remark 1.2.4.** Because every path is assumed to be continuous and strictly positive, it follows from elementary analysis that

$$V(\omega) \stackrel{\text{def}}{=} \inf_{0 \leqslant t \leqslant 1} \sigma_t(\omega) > 0, \quad \omega \in \Omega.$$

We consider $V$ just like $K$ as a function $V : \Omega \to (0,\infty)$. Of course we still assume the processes $W$, $\sigma$ and $b$ to be resp. $(\mathcal{F}_t)$ adapted.

In the following we generalize some of the results proved under the Assumptions 1.1.1 to the weakened Assumptions 1.2.3. For this purpose we use stopping techniques and some properties of the Itô-integral. We start with the weak convergence to the Gumbel distribution as claimed in Corollary 1.2.2.

**Theorem 1.2.5.** *The statement of Corollary 1.2.2 still holds with the weakened Assumptions 1.2.3.*

*Proof.* Define for each $m \in \mathbb{N}$

$$S_m^{(1)} \stackrel{\text{def}}{=} \inf\left\{t \geqslant 0 : \sigma_t \notin \left[\frac{1}{m}, m\right]\right\}, \quad S_m^{(2)} \stackrel{\text{def}}{=} \inf\{t \geqslant 0 : |b_t| > m\}$$

and

$$S_m^{(3)} \stackrel{\text{def}}{=} \inf\left\{t \geqslant 0 : \exists s < t : |\sigma_t - \sigma_s| > m|t-s|^{\frac{1}{m}}\right\}$$

with $\inf \varnothing \stackrel{\text{def}}{=} 1$. Clearly $S_m^{(1)}$ and $S_m^{(2)}$ are stopping times as outlined by Protter [35][p. 4]. Furthermore, $S_m^{(3)}$ is also a stopping time because of

$$\{S_m^{(3)} < t\} = \bigcup_{\substack{0 \leqslant q_1 < q_2 < t, \\ q_1, q_2 \in \mathbb{Q}}} \left\{|\sigma_{q_2} - \sigma_{q_1}| > m|q_2 - q_1|^{\frac{1}{m}}\right\} \in \mathcal{F}_t,$$

taking into account that the Filtration $(\mathcal{F}_t)_t$ is right continuous. Set

$$A_m \stackrel{\text{def}}{=} \bigcap_{j=1}^{3} \left\{S_m^{(j)} = 1\right\} \cap \left\{\frac{1}{m} \leqslant \sigma_0 \leqslant m\right\}.$$

Then (1.33) and (1.34) yield $A_m \uparrow \Omega$. Note that $\sigma^{S_m^{(3)}}$ is Hölder continuous with exponent $\frac{1}{m}$ and coefficient $m$. We only need to consider the critical case $s < t = S_m^{(3)}$ to verify this:

$$|\sigma_{S_m^{(3)}} - \sigma_s| = \lim_k |\sigma_{S_m^{(3)}-\frac{1}{k}} - \sigma_s| \leqslant \lim_k m \left| S_m^{(3)} - \frac{1}{k} - s \right|^{\frac{1}{m}} = m|S_m^{(3)} - s|^{\frac{1}{m}}.$$

Next set $S_m \overset{\text{def}}{=} S_m^{(1)} \wedge S_m^{(2)} \wedge S_m^{(3)}$ and

$$\begin{aligned} \sigma_t^{(m)} &\overset{\text{def}}{=} \mathbb{1}_{[\frac{1}{m},m]}(\sigma_0)\sigma_t^{S_m} + \frac{1}{m}\mathbb{1}_{[\frac{1}{m},m]^c}(\sigma_0) \\ b_t^{(m)} &\overset{\text{def}}{=} \mathbb{1}_{[0,S_m)}(t) b_t, \quad 0 \leqslant t \leqslant 1, \quad m \in \mathbb{N}. \end{aligned}$$

Then $\sigma^{(m)}$ and $b^{(m)}$ fulfill the Assumptions 1.1.1 with

$$\alpha = V = \frac{1}{m}, \quad K = m. \tag{1.35}$$

As $\sigma^{(m)}$ is $(\mathcal{F}_t)_t$ adapted and $t \mapsto b_t^{(m)}(\omega)$ is measurable, we can define

$$\check{Y}_t^{(m)} \overset{\text{def}}{=} \int_0^t \sigma_s^{(m)}\, dW_s + \int_0^t b_s^{(m)}\, d\lambda(s), \quad m \in \mathbb{N}.$$

Using results of Jacod, Shiryaev [20][p. 46 ff.] and the fact that $S_m$ is a stopping time and $\mathbb{1}_{[\frac{1}{m},m]}(\sigma_0) \in \mathcal{F}_0$, we obtain for every $m \in \mathbb{N}$ and $0 \leqslant t \leqslant 1$ the $P$-a.s. equality

$$\int_0^t \sigma_s^{(m)}\, dW_s = \mathbb{1}_{[\frac{1}{m},m]}(\sigma_0)\left(\check{Y}_t^{S_m} + \mathbb{1}_{(S_m,1]}(t)(W_t - W_{S_m})\sigma_{S_m}\right) + \mathbb{1}_{[\frac{1}{m},m]^c}(\sigma_0)\frac{1}{m}W_t.$$

This yields

$$\check{Y}_t^{(m)}(\omega) = \check{Y}_t(\omega), \quad 0 \leqslant t \leqslant 1, \quad \omega \in A_m \cap N_m^c,$$

for a family of negligible sets $(N_m)$, so that we have for every $r \in \mathbb{R}$

$$P(a_n(T_n^{(m)} - b_n) \leqslant r) \leqslant P(a_n(T_n - b_n) \leqslant r) + P(A_m^c), \quad m, n \in \mathbb{N}. \tag{1.36}$$

Write $T_n^{(m)}$ for the statistic $T_n$ in (1.32) with $\check{Y}^{(m)}$ instead of $\check{Y}$. Then Corollary 1.2.2 yields

$$a_n(T_n^{(m)} - b_n) \xrightarrow{\text{d}} \mathcal{G}, \quad n \to \infty \tag{1.37}$$

for all $m \in \mathbb{N}$. From (1.36) and (1.37) follows

$$\liminf_n P(a_n(T_n - b_n) \leqslant r) \geqslant \mathcal{G}((-\infty, r]) - P(A_m^c), \quad m \in \mathbb{N},$$

so that we obtain because of $A_m^c \downarrow \varnothing$

$$\liminf_n P(a_n(T_n - b_n) \leqslant r) \geqslant \mathcal{G}((-\infty, r]), \quad r \in \mathbb{R}.$$

Similar considerations yield

$$\limsup_n P(a_n(T_n - b_n) \leqslant r) \leqslant \mathcal{G}((-\infty, r]), \quad r \in \mathbb{R},$$

providing the final prove for Theorem 1.2.5. □

Next we generalize some $P$-a.s. convergence results to the weakened Assumptions 1.2.3 with a fixed $0 < \alpha \leqslant 1$. To be more precise, we establish the following Corollary.

**Corollary 1.2.6.** *Assume that the weakened Assumptions 1.2.3 hold with a function $\alpha$ independent of the path $\omega$, i.e. $\alpha > 0$ is constant. Then the $P$-a.s. convergences (1.13), (1.14), (1.20) and (1.25) still hold.*

*Proof.* First, note that the requirement on $\alpha$, not to depend on the path $\omega$, is natural in view of what we are going to verify in this corollary. We use the same stopping techniques and notations as in Theorem 1.2.5 with the only difference that we replace $S_m^{(3)}$ with

$$S_m^{(3)} \stackrel{\text{def}}{=} \inf\{t \geqslant 0 \, : \, \exists s < t \, : \, |\sigma_t - \sigma_s| > m|t-s|^\alpha\} \tag{1.38}$$

because $\alpha$ is a constant in our actual setting. Here, we only prove (1.20), i.e.

$$n^\gamma \max_k |\breve{\sigma}_k - \sigma_k| \to 0 \quad P\text{-a.s.}$$

for all $\gamma < \alpha \wedge \frac{1}{2}$ because the proof of all assertions is based on the same idea.

Define $(\breve{\sigma}_k^{(m)})_{k=0,\dots,n-1}$ as a function of $(\Delta \breve{Y}_{k,j}^{(m)})_{0 \leqslant k,j<n}$ instead of $(\Delta \breve{Y}_{k,j})_{0 \leqslant k,j<n}$, i.e.

$$(\breve{\sigma}_k^{(m)})^2 = \frac{\pi n^2}{2(n-1)} \sum_{j=0}^{n-2} |\Delta \breve{Y}_{k,j}^{(m)}||\Delta \breve{Y}_{k,j+1}^{(m)}|, \quad 0 \leqslant k < n, \quad m \in \mathbb{N}.$$

Then, (1.20) states that we have for every $m \in \mathbb{N}$, the convergence

$$n^\gamma \max_{k,j} |\breve{\sigma}_k^{(m)} - \sigma_k^{(m)}|(\omega) \to 0, \quad n \to \infty, \quad \omega \in M_m^c \tag{1.39}$$

for a negligible set $M_m$. Crucial for this proof is, that we have

$$\breve{\sigma}_k^{(m)}(\omega) = \breve{\sigma}_k(\omega), \quad \sigma_k^{(m)}(\omega) = \sigma_k(\omega), \quad 0 \leqslant k < n, \quad \omega \in A_m \cap N_m^c, \quad m,n \in \mathbb{N}, \tag{1.40}$$

with $A_m$ as in Theorem 1.2.5 and $A_m \uparrow \Omega$. Now fix any

$$\omega \in A \stackrel{\text{def}}{=} \left( \bigcup_{m \in \mathbb{N}} A_m \cap N_m^c \right) \cap \left( \bigcup_{m \in \mathbb{N}} M_m \right)^c.$$

Then there is a number $m \in \mathbb{N}$ with $\omega \in A_m \cap N_m^c \cap M_m^c$. Hence, (1.39) and (1.40) yield

$$n^\gamma \max_{k,j} |\breve{\sigma}_k - \sigma_k|(\omega) \stackrel{(1.40)}{=} n^\gamma \max_{k,j} |\breve{\sigma}_k^{(m)} - \sigma_k^{(m)}|(\omega) \stackrel{(1.39)}{\to} 0, \quad n \to \infty.$$

Since $A_m \uparrow \Omega$, we have $P(A) = 1$ and our claim is proven. □

**Remark 1.2.7.** The rate $\alpha \wedge \frac{1}{2}$ in (1.20) is optimal as described in the next Section 1.3. Concerning the optimality of (1.13), choose the deterministic volatility

$$\sigma_s \stackrel{\text{def}}{=} 1 + s^\alpha, \quad 0 \leqslant s \leqslant 1.$$

Then we get using the Itô isometry

$$\limsup_n E\left(n^{\alpha}|n\Delta Y_{0,n-1} - \sigma_0 Z_{0,n-1}|\right)^2 \geqslant 1.$$

For this reason $n^{\alpha}\max_{k,j}|n\Delta Y_{k,j} - \sigma_k Z_{k,j}|$ does *not* converge in $L^2$ to zero. We obtain this result by considering the equations

$$E(n\Delta Y_{0,n-1} - \sigma_0 Z_{0,n-1})^2 = n^2 E\left(\int_{\frac{1}{n}-\frac{1}{n^2}}^{\frac{1}{n}} (\sigma_s - \sigma_0)\, dW_s\right)^2 = n^2 \int_{\frac{1}{n}-\frac{1}{n^2}}^{\frac{1}{n}} (\sigma_s - \sigma_0)^2\, ds$$

and

$$\int_{\frac{1}{n}-\frac{1}{n^2}}^{\frac{1}{n}} (\sigma_s - \sigma_0)^2\, ds = \int_{\frac{1}{n}-\frac{1}{n^2}}^{\frac{1}{n}} s^{2\alpha}\, ds \geqslant \frac{1}{n^2}\left(\frac{1}{n} - \frac{1}{n^2}\right)^{2\alpha} = \frac{1}{n^2}\left(\frac{1}{n}\right)^{2\alpha}\left(1 - \frac{1}{n}\right)^{2\alpha}.$$

Finally, we state (1.25) itself as a theorem under the weakened assumptions because this result is quite interesting in view of building a jump test in combination with Theorem 1.2.5.

**Theorem 1.2.8.** *Let the weakened Assumptions 1.2.3 hold with a constant $\alpha > 0$. Then we obtain for all $\gamma < \alpha \wedge \frac{1}{2}$ the convergence*

$$n^{\gamma}\left|n\max_{k,j}\frac{\Delta\breve{Y}_{k,j}}{\breve{\sigma}_k} - \max_{k,j} Z_{k,j}\right| \leqslant n^{\gamma}\max_{k,j}\left|\frac{n\Delta\breve{Y}_{k,j}}{\breve{\sigma}_k} - Z_{k,j}\right| \to 0 \quad P\text{-a.s.}$$

*Proof.* This is a direct consequence of Corollary 1.2.6 together with Lemma A.7. □

## 1.3 Uniform and pathwise estimate of the spot volatility

In this section, we are going to present an estimator for the spot volatility process $\sigma$, such that the estimated spot volatility converges uniformly and pathwise to the true spot volatility. The optimal convergence rate of this estimator is revealed. We emphasize that we only need very weak and natural assumptions to establish the convergence. To be more precise we assume the Assumptions 1.2.3 with a fixed $0 < \alpha \leqslant 1$. Our estimator is a linear interpolation of the realised bipower variation estimator in Definition 1.1.3 on the grid. Hence its calculation can be done very easily and quickly. A similar approach with a kernel type estimator was also taken by Fan and Wang [15]. They reach similar convergence rates as our test in the case $\alpha > \frac{1}{2}$. The case $\alpha < \frac{1}{2}$ is not covered by their results. An interesting result in [15] is the weak convergence of a suitably scaled supremum norm to the Gumbel distribution (Theorem 2 in their paper) which is valid in the case that $\sigma$ is stationary. Nevertheless, we are interested in pathwise, uniform convergences under consideration of the parameter $\alpha$. Note also an interesting alternative approach by Hoffman, Munk and Schmidt-Hieber [18]. They use a wavelet type estimator and consider the $L^p$ error, $p < \infty$.

In order avoid confusion with the notation, the time argument of all processes in this section is denoted in brackets and the grid fineness is denoted with an index. For (1.41), this means, for instance, $\breve{\sigma}_k$ is denoted as $\breve{\sigma}_n\left(\frac{k}{n}\right)$. Set

$$\breve{\sigma}_n\left(\frac{n}{n}\right) \stackrel{\text{def}}{=} \breve{\sigma}_n\left(\frac{n-1}{n}\right).$$

Then the statement (1.20) in Corollary 1.2.6 yields, together with

$$n^{\gamma}\left|\breve{\sigma}_n\left(\frac{n-1}{n}\right)-\sigma\left(\frac{n}{n}\right)\right| \leqslant n^{\gamma}\left|\breve{\sigma}_n\left(\frac{n-1}{n}\right)-\sigma\left(\frac{n-1}{n}\right)\right| + n^{\gamma-\alpha}K(\omega),$$

the convergence

$$n^{\gamma}\max_{0\leqslant k\leqslant n}\left|\breve{\sigma}_n\left(\frac{k}{n}\right)-\sigma\left(\frac{k}{n}\right)\right| \to 0, \quad P\text{-a.s.} \tag{1.41}$$

for all $\gamma < \alpha \wedge \frac{1}{2}$. Equation (1.41) is our starting point. Next, for every $n \in \mathbb{N}$, a natural estimator $\breve{\tau}_n$ for the spotvolatility $\sigma$ is defined which is based on $n^2+1$ equidistant high-frequency observations at the time points $0, \frac{1}{n^2}, \frac{2}{n^2}, \ldots, 1$ of the underlying process $\breve{Y}$, i.e. we have

$$\breve{\tau}_n : \Omega \times [0,1] \to \mathbb{R}_+, \quad n \in \mathbb{N}.$$

For this purpose $\breve{\tau}_n$ is defined to be the linear interpolation of the points $\left(\breve{\sigma}_n\left(\frac{k}{n}\right)\right)_{k=0,1,\ldots,n}$. This formally means

$$\breve{\tau}_n(t) \overset{\text{def}}{=} r(t)\breve{\sigma}_n(k(t)) + (1-r(t))\breve{\sigma}_n\left(k(t)+\frac{1}{n}\right), \quad 0 \leqslant t \leqslant 1, \quad n \in \mathbb{N}$$

with the two deterministic functions $k, r : [0,1] \mapsto \mathbb{R}_+$ defined via

$$k(t) \overset{\text{def}}{=} \frac{\lfloor nt \rfloor}{n}, \quad 0 \leqslant t < 1, \quad k(1) \overset{\text{def}}{=} \frac{n-1}{n}$$

and

$$r(t) \overset{\text{def}}{=} \begin{cases} 1, & nt \in \{0,1,\ldots,n-1\} \\ \lceil nt \rceil - nt, & \text{else} \end{cases}, \quad 0 \leqslant t \leqslant 1.$$

An estimator $\hat{\tau}_n$ in the case of a vanishing drift term is defined in an analogous manner. Note that our interpolation-based estimators coincide with the estimators in Definition 1.1.3 on the grid points. Next we define for every $0 < \alpha \leqslant 1$ the following sets

$$\mathcal{V}_\alpha \overset{\text{def}}{=} \{(\sigma,b) : (\sigma,b) \text{ satisfies Assumptions 1.2.3 with constant } \alpha\}$$

and denote for any continuous function $f : [0,1] \to \mathbb{R}$ with

$$\|f\|_\infty \overset{\text{def}}{=} \sup\{|f(t)| : 0 \leqslant t \leqslant 1\}$$

its supremum norm.

Now, a theorem concerning the convergence of the spot volatility estimator $\breve{\tau}_n$ can be formulated

**Theorem 1.3.1.** *Let the Assumptions 1.2.3 hold with a fixed $0 < \alpha \leqslant 1$. Then, we have for all $\gamma < \alpha \wedge \frac{1}{2}$ the convergence*

$$n^{\gamma}\|\breve{\tau}_n - \sigma\|_\infty \to 0, \quad n \to \infty \quad P\text{-a.s.} \tag{1.42}$$

*and the upper bound $\alpha \wedge \frac{1}{2}$ is sharp in the sense that*

$$\left\{\beta \in \mathbb{R} : n^{\beta}\|\breve{\tau}_n - \sigma\|_\infty \overset{P\text{-a.s.}}{\to} 0 \text{ for all } (\sigma,b) \in \mathcal{V}_\alpha\right\} \subseteq \left(-\infty, \frac{1}{2} \wedge \alpha\right]. \tag{1.43}$$

*Hence, the optimal convergence rate is* $n^{-(\alpha\wedge\frac{1}{2})}$.

*Proof.* The proof is divided into two steps. The first one is devoted to the convergence result (1.42) and the second one to the rate result (1.43).

STEP 1. *Proof of (1.42).* Using the linear interpolation approach, it holds for $0 \leqslant t \leqslant 1$

$$\begin{aligned} |\check{\tau}_n(t) - \sigma(t)| &\leqslant |\check{\tau}_n(t) - \check{\tau}_n(k(t))| + |\check{\tau}_n(k(t)) - \sigma(k(t))| + |\sigma(k(t)) - \sigma(t)| \\ &\leqslant \frac{1}{n}\max_{0\leqslant k\leqslant n}\check{\sigma}_n\left(\frac{k}{n}\right) + |\check{\sigma}_n(k(t)) - \sigma(k(t))| + \frac{K(\omega)}{n^\alpha} \\ &\leqslant \frac{1}{n}\left(\max_{0\leqslant k\leqslant n}\left|\check{\sigma}_n\left(\frac{k}{n}\right) - \sigma\left(\frac{k}{n}\right)\right| + K(\omega)\right) + |\check{\sigma}_n(k(t)) - \sigma(k(t))| + \frac{K(\omega)}{n^\alpha} \end{aligned}$$

which implies together with (1.41) the claim (1.42).

STEP 2. *Proof of (1.43).* Set for any $0 < \alpha < 1$

$$\sigma(s) = 1 + s^\alpha, \quad 0 \leqslant s \leqslant 1,$$

i.e. $\sigma$ is positive, deterministic and the prototype of an $\alpha$-Hölder continuous function. Furthermore, set $b = 0$ and note that $(\sigma, b) \in \mathcal{V}_\alpha$. Fix any $\beta > \alpha$. In what follows we are going to prove the pointwise divergence

$$n^\beta \|\hat{\tau}_n - \sigma\|_\infty(\omega) \to \infty, \quad n \to \infty, \quad \omega \in \Omega. \tag{1.44}$$

This divergence to infinity is due to our interpolation approach. To prove it, we write for $0 \leqslant t \leqslant 1$

$$\begin{aligned} \hat{\tau}_n(t) &= r(t)[\sigma(k(t)) - (\sigma(k(t)) - \hat{\sigma}_n(k(t)))] \\ &\quad + (1 - r(t))\left[\sigma\left(k(t) + \frac{1}{n}\right) - \left(\sigma\left(k(t) + \frac{1}{n}\right) - \hat{\sigma}_n\left(k(t) + \frac{1}{n}\right)\right)\right] \end{aligned}$$

which implies

$$\begin{aligned} |\hat{\tau}_n(t) - \sigma(t)| &= \Bigg| r(t)[\sigma(k(t)) - \sigma(t)] + (1 - r(t))\left[\sigma\left(k(t) + \frac{1}{n}\right) - \sigma(t)\right] \\ &\quad - r(t)[\sigma(k(t)) - \hat{\sigma}_n(k(t))] - (1 - r(t))\left[\sigma\left(k(t) + \frac{1}{n}\right) - \hat{\sigma}_n\left(k(t) + \frac{1}{n}\right)\right]\Bigg| \\ &\geqslant r(t)|\sigma(k(t)) - \sigma(t)| - (1 - r(t))\left|\sigma\left(k(t) + \frac{1}{n}\right) - \sigma(t)\right| \\ &\quad - |\sigma(k(t)) - \hat{\sigma}_n(k(t))| - \left|\sigma\left(k(t) + \frac{1}{n}\right) - \hat{\sigma}_n\left(k(t) + \frac{1}{n}\right)\right|, \end{aligned}$$

so that we obtain

$$\|\hat{\tau}_n - \sigma\|_\infty \geqslant \frac{1}{3}\sup_{0\leqslant t\leqslant 1}\left\{r(t)|\sigma(k(t)) - \sigma(t)| - (1 - r(t))\left|\sigma\left(k(t) + \frac{1}{n}\right) - \sigma(t)\right|\right\}. \tag{1.45}$$

Next, choose any $0 < \lambda < \frac{\beta}{\alpha} - 1$ and set $t_n \stackrel{\text{def}}{=} n^{-(1+\lambda)}$. Then we receive because of

$$t_n = (1 - n^{-\lambda}) \cdot 0 + n^{-\lambda} n^{-1}$$

the relations

$$k(t_n) = 0, \quad r(t_n) = 1 - n^{-\lambda}.$$

This yields

$$\begin{aligned} & n^{\beta}\left(r(t_n)|\sigma(k(t_n)) - \sigma(t_n)| - (1 - r(t_n))\left|\sigma\left(k(t_n) + \frac{1}{n}\right) - \sigma(t_n)\right|\right) \\ = \; & n^{\beta}\left(\left(1 - n^{-\lambda}\right)n^{-\alpha(1+\lambda)} - n^{-\lambda}\left(n^{-\alpha} - n^{-\alpha(1+\lambda)}\right)\right) \\ = \; & \left(1 - n^{-\lambda}\right)n^{-\alpha(1+\lambda)+\beta} - n^{\beta-\lambda-\alpha} + n^{\beta-\lambda-\alpha(1+\lambda)} \to \infty, \quad n \to \infty \end{aligned} \tag{1.46}$$

where the divergence to infinity holds since we have

$$-\alpha(1+\lambda) + \beta > \max(\beta - \lambda - \alpha, 0).$$

Finally, note that (1.45) and (1.46) establish (1.44). With (1.44) in mind, it obviously suffices to verify

$$P\left(\limsup_n n^{\frac{1}{2}}\|\check{\tau}_n - \sigma\|_\infty > 0\right) > 0 \tag{1.47}$$

for a pair $(\sigma, b) \in \mathcal{V}_1$ in order to establish our wanting claim (1.43). For this purpose, choose $\sigma = 1$ and $b = 0$, i.e. $Y$ is a standard Brownian motion and it holds obviously $(\sigma, b) \in \mathcal{V}_1$. Now, (1.47) holds because of the classical central limit theorem (CLT) for i.i.d random variables. Nevertheless, a rigorous proof of (1.47) is provided in the following: Note that

$$n^{\frac{1}{2}}\left(\frac{\pi n^2}{2(n-1)}\sum_{j=0}^{n-2}|\Delta Y_{0,j}||\Delta Y_{0,j+1}| - 1\right) = \frac{\pi n^{\frac{1}{2}}}{2(n-1)}\sum_{j=0}^{n-2}\left(|Z_{0,j}||Z_{0,j+1}| - \frac{2}{\pi}\right).$$

Write $Z_j \overset{\text{def}}{=} Z_{0,j}$, $j \in \mathbb{N}$ and consider a decomposition as in (A.9), i.e. write

$$\xi_n^{(1)} + \xi_n^{(2)} + \epsilon_n \overset{\text{def}}{=} n^{-\frac{1}{2}}\sum_{j=1}^{n}\left(|Z_j||Z_{j+1}| - \frac{2}{\pi}\right), \quad n \in \mathbb{N} \tag{1.48}$$

with

$$\xi_n^{(j)} \xrightarrow{d} \mathrm{N}(0, \tau), \quad \tau \overset{\text{def}}{=} \frac{1}{2}\mathrm{Var}|Z_1||Z_2|, \quad j = 1, 2$$

and

$$\epsilon_n \to 0, \quad P\text{-a.s. and in } L^2 \tag{1.49}$$

in the limit $n \to \infty$ respectively. Note next that

$$\left\{|\xi_n^{(1)} + \xi_n^{(2)}|^2\right\}_{n\in\mathbb{N}} \tag{1.50}$$

is a family of uniformly integrable (u.i.) random variables because of

$$E|\xi_n^{(1)} + \xi_n^{(2)}|^4 \leqslant 8\left(E|\xi_n^{(1)}|^4 + E|\xi_n^{(2)}|^4\right)$$

and Proposition A.1 together with (A.8). Now assume that (1.47) is not true. This and (1.49) imply

$$\xi_n^{(1)} + \xi_n^{(2)} \to 0 \quad P\text{-a.s.}$$

Since (1.50) is u.i. it necessarily follows

$$E|\xi_n^{(1)} + \xi_n^{(2)}|^2 \to 0, \quad n \to \infty.$$

But this is a contradiction to (1.48) because of (1.49) and the equations

$$\begin{aligned}\frac{1}{n}\left(\sum_{j=1}^{n}\left(|Z_j||Z_{j+1}| - \frac{2}{\pi}\right)\right)^2 &= \frac{1}{n}\sum_{0 \leqslant j,k < n} E\left[\left(|Z_j||Z_{j+1}| - \frac{2}{\pi}\right)\left(|Z_k||Z_{k+1}| - \frac{2}{\pi}\right)\right] \\ &= \frac{1}{n}\sum_{j=1}^{n} E\left(|Z_j||Z_{j+1}| - \frac{2}{\pi}\right)^2 \\ &= \operatorname{Var}|Z_1||Z_2| > 0.\end{aligned}$$

This proves (1.43). Thus, the proof of this theorem is completed under additional consideration of (1.44). □

## 1.4 Simplifications in the independent case

In this section it is assumed that the Brownian motion $W$, the volatility $\sigma$ and the drift coefficient $b$ are independent. A proof of the quadratic variation version of Theorem 1.2.5 is stated with the additional restriction that $0 < \alpha \leqslant 1$ is a constant, which, in this context, is the same as saying

$$\inf_{\omega \in \Omega} \alpha(\omega) > 0. \tag{1.51}$$

Note that we estimate the spot volatility in the following Theorem 1.4.2 with the realised quadratic variation, i.e. by using the squares of the increments and not the product of two neighbour increments (realised bipower variation) as in Theorem 1.2.5. Nevertheless, after a few minor changes, the proof of Theorem 1.4.2 also holds with the realised bipower variation estimator of the spot volatility. Vice versa, Theorem 1.2.5 is true with the realised quadratic variation estimator, i.e. in this sense both estimators are equivalent. Concerning external jumps, both estimators, however, are different. Their differences are investigated in more detail in Section 1.6. As already mentioned above, a redefinition is performed.

**Redefinition 1.4.1.** *Set for* $0 \leqslant k < n$

$$\begin{aligned}\widehat{\sigma}_k^2 &\stackrel{\text{def}}{=} n\sum_{j=0}^{n-1}(\Delta Y_{k,j})^2 \quad \text{(without drift)}, \\ \breve{\sigma}_k^2 &\stackrel{\text{def}}{=} n\sum_{j=0}^{n-1}(\Delta \check{Y}_{k,j})^2 \quad \text{(with drift)}.\end{aligned}$$

Due to the independence of $W$, $\sigma$ and $b$, we are capable to give a direct and simpler proof than the one of Theorem 1.2.5. The proof of Theorem 1.4.2 is based on the following representation (1.52) which yields an immense technical simplification. For example, Proposition 1.1.4 is not required. Representation (1.52) loosely states that we can assume w.l.o.g. that the volatility is deterministic. In view of this interpretation our restriction (1.51) becomes natural. It is also customary to state that we can condition on the volatility and drift processes.

**Theorem 1.4.2.** *The statement of Corollary 1.2.2 holds under the Assumptions 1.2.3 with constant $0 < \alpha \leqslant 1$ and the Redefinition 1.4.1.*

*Proof.* The proof is divided into three steps. The first one is based on the independence of $W$, $\sigma$ and $b$. Here, it is shown that we can work w.l.o.g. on another probability space which gives us a useful representation of the stochastic integral. The second step contains a short proof of Proposition 1.1.5 without using Proposition 1.1.4 due to the representation (1.52). Finally, the third step accomplishes the proof under citation of the proof of Proposition 1.2.1 from (1.28) downwards.

STEP 1. *A new probability space.* Let $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ be a probability space which possesses a sequence $(Z_{k,j})_{(k,j)\in\mathbb{N}_0 \times \mathbb{N}_0}$, $Z_{k,j} \sim \mathrm{N}(0,1)$ of i.i.d. random variables on it. Define

$$\bar{\Omega} = \Omega \times \widetilde{\Omega}, \quad \bar{\mathcal{F}} = \mathcal{F} \otimes \widetilde{\mathcal{F}}, \quad \bar{P} = P \otimes \widetilde{P}$$

and set

$$\bar{\sigma}_t(\bar{\omega}) = \sigma_t(\omega), \quad \bar{b}_t(\bar{\omega}) = b_t(\omega), \quad \bar{B}_t(\bar{\omega}) = \int_0^t \bar{b}_s(\bar{\omega})\lambda(ds), \quad \bar{Z}_{k,j}(\bar{\omega}) = Z_{k,j}(\widetilde{\omega})$$

where $\bar{\omega} \stackrel{\text{def}}{=} (\omega, \widetilde{\omega}) \in \bar{\Omega}$ and $t \in [0,1]$. Then $(\bar{\sigma}_t)$, $(\bar{Z}_{k,j})$ and $(\bar{B}_t)$ are $\bar{P}$-independent and the law of $(\sigma_t)$ under $P$ is equal to the law of $(\overline{\sigma}_t)$ under $\overline{P}$. The crucial fact is that we have

$$\left(\left(\int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}} \bar{\sigma}_s^2\, ds\right)^{\frac{1}{2}} \bar{Z}_{k,j}, \Delta\bar{B}_{k,j}\right)_{0\leqslant k,j<n} \stackrel{\mathrm{d}}{=} (\Delta Y_{k,j}, \Delta B_{k,j})_{0\leqslant k,j<n}. \tag{1.52}$$

See also (11) in Barndorff-Nielsen and Shephard [4]. This paper is based on the above representation. Since we intend to prove the weak convergence

$$a_{n^2}(T_n - b_{n^2}) \xrightarrow{d} \mathcal{G}, \tag{1.53}$$

we can consider w.l.o.g. the left hand side of (1.53) as a function of the left hand side of (1.52). However, for a more convenient notation $(\Omega, \mathcal{F}, P)$ is written instead of $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})$ and the overline notations of the processes are not used.

STEP 2. *An alternative proof of Proposition 1.1.5.* First the two convergences

$$n^{\delta} \max_k \left| \frac{1}{n} \sum_{j=0}^{n-1} Z_{k,j}^2 - 1 \right| \to 0, \quad P\text{-a.s.}, \quad \delta < \frac{1}{2}, \tag{1.54}$$

$$n^{-\delta} \max_{k,j} |Z_{k,j}| \to 0, \quad P\text{-a.s.}, \quad \delta > 0 \tag{1.55}$$

are established beginning with (1.54). Write for this purpose with $\delta < \frac{1}{2}$, $m \in \mathbb{N}$ and Corollary A.2, the inequalities

$$\begin{aligned} P\left(n^{\delta} \max_k \left| \frac{1}{n} \sum_{j=0}^{n-1} Z_{k,j}^2 - 1 \right| \geqslant \epsilon\right) &\leqslant nP\left(\left| \frac{1}{n} \sum_{j=0}^{n-1} Z_{0,j}^2 - 1 \right| \geqslant \epsilon n^{-\delta}\right) \\ &\leqslant nC\epsilon^{-2m} n^{2m\delta} n^{-m} \end{aligned}$$

and note that $2\delta - 1 < 0$ since $\delta < \frac{1}{2}$. Thus, Borel-Cantelli yields the desired convergence. For proving (1.55), use the same argument with Proposition A.3. For completeness, we write for $\delta > 0$ and $m \in \mathbb{N}$ the inequalities

$$\begin{aligned} P\left(n^{-\delta}\max_{k,j}|Z_{k,j}| \geqslant \epsilon\right) &\leqslant \epsilon^{-m}n^{-\delta m}\left(2^m\left(\log n^2\right)^{\frac{m}{2}} + 2m!\right) \\ &= \epsilon^{-m}o\left(n^{-\frac{\delta}{2}m}\right) \end{aligned}$$

and note $-\frac{\delta}{2} < 0$ since $\delta > 0$. Next it is proven

$$n^\gamma \max_{k,j}|n\Delta Y_{k,j} - \sigma_k Z_{k,j}| \to 0, \quad P\text{-a.s.}, \quad \gamma < \alpha, \tag{1.56}$$

$$n^\gamma \max_{k}|\hat{\sigma}_k^2 - \sigma_k^2| \to 0, \quad P\text{-a.s.}, \quad \gamma < \alpha \wedge \frac{1}{2}. \tag{1.57}$$

Starting with (1.56), we write using (1.52) the equality

$$n\Delta Y_{k,j} - \sigma_k Z_{k,j} = \left(\left(n^2\int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}}\sigma_s^2\,ds\right)^{\frac{1}{2}} - \sigma_k\right)Z_{k,j}, \quad 0 \leqslant k, j < n.$$

Note that we obtain from the mean value theorem for Riemann-integrals the estimates

$$\begin{aligned} \left|n^2\int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}}\sigma_s^2\,ds - \sigma_k^2\right| &= |\sigma_{\xi_{k,j}}^2 - \sigma_k^2| = |\sigma_{\xi_{k,j}} + \sigma_k||\sigma_{\xi_{k,j}} - \sigma_k| \leqslant 2K\cdot K\left|\xi_{k,j} - \frac{k}{n}\right|^\alpha \\ &\leqslant 2K^2\frac{1}{n^\alpha}, \quad 0 \leqslant k, j < n \end{aligned} \tag{1.58}$$

where $t_{k,j} \leqslant \xi_{k,j} \leqslant t_{k,j} + \frac{1}{n^2}$. Next fix any $\gamma < \alpha$ and $\delta > 0$ with $\gamma + \delta < \alpha$ and consider the inequality

$$n^\gamma\max_{k,j}|n\Delta Y_{k,j} - \sigma_k Z_{k,j}| \leqslant V^{-1}n^{\gamma+\delta}\max_{k,j}\left|n^2\int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}}\sigma_s^2\,ds - \sigma_k^2\right|n^{-\delta}\max_{k,j}|Z_{k,j}|.$$

This proves together with (1.55) and (1.58) the claim (1.56). Next, (1.57) is considered and for this purpose it is stated

$$\hat{\sigma}_k^2 = n\sum_{j=0}^{n-1}(\Delta Y_{k,j})^2 = \frac{1}{n}\sum_{j=0}^{n-1}\left(n^2\int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}}\sigma_s^2\,ds - \sigma_k^2 + \sigma_k^2\right)Z_{k,j}^2, \quad 0 \leqslant k < n$$

which implies

$$\hat{\sigma}_k^2 - \sigma_k^2 = \sigma_k^2\left(\frac{1}{n}\sum_{j=0}^{n-1}Z_{k,j}^2 - 1\right) + \frac{1}{n}\sum_{j=0}^{n-1}\left(n^2\int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}}\sigma_s^2\,ds - \sigma_k^2\right)Z_{k,j}^2.$$

So we can estimate

$$|\hat{\sigma}_k^2 - \sigma_k^2| \leqslant K^2\left|\frac{1}{n}\sum_{j=0}^{n-1}Z_{k,j}^2 - 1\right| + \max_{0\leqslant j<n}\left|n^2\int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}}\sigma_s^2\,ds - \sigma_k^2\right|\cdot\left(\left|\frac{1}{n}\sum_{i=0}^{n-1}Z_{k,i}^2 - 1\right| + 1\right).$$

This yields together with (1.54) and (1.58) for $\gamma < \alpha \wedge \frac{1}{2}$ the inequality (1.57).

STEP 3. *Drift analysis and handling the quotient.* With the same argumentation as in the proofs of Proposition 1.2.1 and Corollary 1.2.2, it suffices to verify

$$a_{n^2} \max_{k,j} \left| \frac{n\Delta \breve{Y}_{k,j}}{\breve{\sigma}_k} - \frac{n\Delta Y_{k,j}}{\sigma_k} \right| \to 0 \quad P\text{-a.s.} \tag{1.59}$$

To this end we first establish

$$n^{1-\delta} \max_{k,j} |\Delta Y_{k,j}| \;\to\; 0, \quad P\text{-a.s.}, \quad \delta > 0, \tag{1.60}$$

$$n^{\gamma} \max_{k} |\breve{\sigma}_k - \sigma_k| \;\to\; 0, \quad P\text{-a.s.}, \quad \gamma < \alpha \wedge \frac{1}{2}. \tag{1.61}$$

To prove (1.60) consider (1.27) and the statements (1.55) and (1.56). To prove (1.61) consider first the analogue quadratic variation calculation to (1.21), i.e.

$$\begin{aligned}
|\breve{\sigma}_k^2 - \widehat{\sigma}_k^2| &= n \left| \sum_{j=0}^{n-1} |\Delta \breve{Y}_{k,j}|^2 - |\Delta Y_{k,j}|^2 \right| \\
&\leqslant n \sum_{j=0}^{n-1} \Big| |\Delta Y_{k,j} + \Delta B_{k,j}| - |\Delta Y_{k,j}| \Big| \cdot \Big| |\Delta Y_{k,j} + \Delta B_{k,j}| + |\Delta Y_{k,j}| \Big| \\
&\leqslant n \sum_{j=0}^{n-1} |\Delta B_{k,j}| (2|\Delta Y_{k,j}| + |\Delta B_{k,j}|) \\
&\leqslant 2K \max_{k,j} |\Delta Y_{k,j}| + \frac{K^2}{n^2}.
\end{aligned} \tag{1.62}$$

This implies for any fixed $\gamma < 1$ because of (1.60) and (1.62) the convergence

$$n^{\gamma} \max_{k} |\breve{\sigma}_k^2 - \widehat{\sigma}_k^2| \to 0 \quad \text{P-a.s.}$$

Now, we arrive together with (1.57) for $\gamma < \alpha \wedge \frac{1}{2}$ at the inequality

$$n^{\gamma} \max_{k} |\breve{\sigma}_k - \sigma_k| = n^{\gamma} \max_{k} \frac{|\breve{\sigma}_k^2 - \sigma_k^2|}{\breve{\sigma}_k + \sigma_k} \leqslant V^{-1} n^{\gamma} \left( \max_{k} |\breve{\sigma}_k^2 - \widehat{\sigma}_k^2| + \max_{k} |\widehat{\sigma}_k^2 - \sigma_k^2| \right)$$

which yields (1.61).

Having proven (1.60) and (1.61), the estimation of the quotient is exactly the same as in the proof of Proposition 1.2.1 from (1.28) down to its end. □

## 1.5 Jumps in the volatility process

Jumps in the volatility process $\sigma$ cause some problems. Nevertheless, a positive result under appropriate strong assumptions is our starting point.

**Assumptions 1.5.1.** *Let the volatility $\sigma > 0$ and the Brownian motion $W$ be independent. Set* *further $b = 0$, i.e. absence of a drift, and assume*

$$V(\omega) \overset{\text{def}}{=} \inf_{0 \leqslant t \leqslant 1} \sigma_t(\omega) > 0, \quad \omega \in \Omega.$$

*Furthermore, fix $0 \leqslant \epsilon < 1$ and $0 < \alpha \leqslant 1$ and let $t \mapsto \sigma_t(\omega)$ be càglàd with, at most, finite many jumps of size not larger than $\epsilon(\sqrt{2}-1)V(\omega)$ and $\alpha$-Hölder continuous between the jumps for every path $\omega \in \Omega$. This means that there is a function $N : \Omega \to \mathbb{N}_0$ and a sequence $(S_l)_{l \geqslant 0}$ of stopping times with $S_0 \overset{\text{def}}{=} 0$,*

$$\begin{cases} S_l(\omega) < S_{l+1}(\omega), & 1 \leqslant l \leqslant N(\omega), \\ S_l(\omega) = \infty, & l > N(\omega), \end{cases}$$

*such that*

$$|\sigma_s(\omega) - \sigma_t(\omega)| \leqslant K(\omega)|t-s|^\alpha$$

*holds for all*

$$(s,t) \in \bigcup_{l=0}^{N(\omega)} (S_l(\omega), S_{l+1}(\omega) \wedge 1]^2, \quad \omega \in \Omega.$$

*Here, $N$ denotes the number of jumps in the respective path and $(S_l)_{1 \leqslant l \leqslant N}$ are the jump positions. Furthermore, we claim*

$$0 < |\Delta\sigma_{S_l}(\omega)| \leqslant \epsilon(\sqrt{2}-1)V(\omega), \quad 1 \leqslant l \leqslant N, \quad \omega \in \Omega$$

*and assume as usual*

$$|\sigma_t(\omega)| < K(\omega), \quad 0 \leqslant t \leqslant 1, \quad \omega \in \Omega.$$

**Remark 1.5.2.** The volatility $\sigma$ is required to be càglàd in the above Assumptions 1.5.1. This is due to the needed predictable integrands in the Itô-calculus. However, since the Brownian motion $W$ has continuous paths, there is in fact no difference to the corresponding càdlàg version, cf. Karatzas, Shreve [22] or Jacod, Shiryaev [20].

**Theorem 1.5.3.** *Corollary 1.2.2 holds under the Assumptions 1.5.1 with the spot volatility estimator in Redefinition 1.4.1.*

**Remark 1.5.4.** In general, Assumptions 1.5.1 say that sufficient small jumps with finite activity in the volatility process are allowed. It turns out that the jump size bound $(\sqrt{2}-1)V$ is sharp in the sense of Corollary 1.5.7 at the end of this section. We assume that $W$ and $\sigma$ are independent to keep the technical overhead as small as possible. The independence has the advantage that we can use the same technique as in the previous Section 1.4. In a way we generalize in what follows the proof of Theorem 1.4.2.

**Remark 1.5.5.** For completeness, a formal proof is provided that such a sequence of stopping times $(S_l)_{l \geqslant 0}$ as stated in the Assumptions 1.5.1 exists, $N$ is measurable and that $K$ can be chosen as a measurable function. This is also important for the proof of Theorem 1.5.3. Assume for this purpose that the Assumptions 1.5.1 hold and set for $l \geqslant 1$

$$S_l(\omega) \overset{\text{def}}{=} \begin{cases} \text{Position of the } l\text{-th jump in } \sigma(\omega), & \sigma(\omega) \text{ has at least } l \text{ jumps}, \\ \infty, & \text{else} \end{cases}$$

and set $S_0 \stackrel{\text{def}}{=} 0$. To understand that each $S_l$ is a stopping time, an inductive argument is provided. Firstly, define for this purpose with $r, s, u, v \in \mathbb{Q}$ and $m, n \in \mathbb{N}$, the sets

$$\begin{aligned} I_{u,v,n} &\stackrel{\text{def}}{=} \left\{(r,s) \in \mathbb{Q}^2 \,:\, u < r, s < v \text{ and } |r-s| < \frac{1}{n}\right\}, \\ A_{r,s,m} &\stackrel{\text{def}}{=} \left\{\omega \in \Omega \,:\, |\sigma_r(\omega) - \sigma_s(\omega)| > \frac{1}{m}\right\}. \end{aligned}$$

$S_0$ is obviously a stopping time. Assume for the induction step that $S_l$ is also a stopping time for some $l \in \mathbb{N}$. Observe, furthermore, for $t > 0$ and

$$C_{u,v} \stackrel{\text{def}}{=} \bigcup_{m\in\mathbb{N}} \bigcap_{n\in\mathbb{N}} \bigcup_{(r,s)\in I_{u,v,n}} A_{r,s,m} \in \mathcal{F}_v, \quad 0 < u, v < 1$$

the relation

$$\{S_{l+1} < t\} = \bigcup_{\substack{0<s<t,\\ s\in\mathbb{Q}}} \{S_l < s\} \cap C_{s,t} \in \mathcal{F}_t$$

which proves that $(S_l)$ are stopping times due to the right continuity of the filtration $(\mathcal{F}_t)$. Next

$$\{N = n\} = \bigcap_{m=1}^{n} \{S_m < \infty\} \cap \{S_{n+1} = \infty\} \in \mathcal{F}, \quad n \in \mathbb{N}_0$$

yields that $N$ is measurable. It remains to establish that $K$ can be chosen as a measurable function. To understand this, set

$$\psi_t \stackrel{\text{def}}{=} \sigma_t - \sum_{l=1}^{\infty} \Delta\sigma_{S_l} \, \mathbb{1}_{(S_l,1]}(t), \quad 0 \leqslant t \leqslant 1, \quad \Delta\sigma_\infty \stackrel{\text{def}}{=} 0$$

and observe that $\psi$ is $(\mathcal{F}_t)$ adapted since $\sigma$ is $(\mathcal{F}_t)$ adapted and $(S_l)$ are stopping times as proven previously. Note that $\psi$ is simply $\sigma$ without jumps. Since $\psi$ is pathwise $\alpha$-Hölder continuous, we can define

$$K(\omega) \stackrel{\text{def}}{=} \left( \sup_{\substack{0\leqslant s<t\leqslant 1,\\ s,t\in\mathbb{Q}}} \frac{|\psi_t(\omega) - \psi_s(\omega)|}{|t-s|^\alpha} \right) \vee \sup_{\substack{0\leqslant t\leqslant 1,\\ t\in\mathbb{Q}}} |\sigma_t(\omega)| < \infty, \quad \omega \in \Omega.$$

$K$ is obviously measurable and fulfills the requirements of the Assumptions 1.5.1. Compare for similar results in this context also Chapter I, Proposition 1.32 in Jacod, Shiryaev [20] or Chapter I in Protter [35].

We assume w.l.o.g. $P(S_1 = 0) = 0$ and set

$$\begin{aligned} K_{l,n} &\stackrel{\text{def}}{=} (\lceil nS_l \rceil - 1)\, \mathbb{1}_{\{l\leqslant N\}} + n\, \mathbb{1}_{\{l>N\}}, \\ G_{l,n} &\stackrel{\text{def}}{=} \{(g_1, \ldots, g_l) \in \{0, 1, \ldots, n-1\}^l \,:\, g_1 < g_2 < \ldots < g_l\}, \quad l, n \geqslant 1. \end{aligned}$$

Now we turn to the proof.

*Proof of Theorem 1.5.3.* Since the volatility $\sigma$ and the Brownian motion $W$ are independent according to the Assumptions 1.5.1, we consider the same probability space as the one in the proof of Theorem 1.4.2. This briefly means that we can assume w.l.o.g.

$$\Delta Y_{k,j} = \left( \int_{t_{k,j}}^{t_{k,j}+\frac{1}{n^2}} \sigma_s^2 \, ds \right)^{\frac{1}{2}} Z_{k,j}, \quad 0 \leqslant k, j < n, \quad n \in \mathbb{N}.$$

A precise consideration of the proof of Theorem 1.4.2 shows that we have proven

$$a_{n^2} \max_{k',j} \left| \frac{n \Delta Y_{k',j}}{\widehat{\sigma}_{k'}} - Z_{k',j} \right| \to 0, \quad n \to \infty, \quad P\text{-a.s.}$$

where

$$k' \in \{0, 1, \ldots, n-1\} - \bigcup_{l=1}^{N} \{K_{l,n}\}, \quad j \in \{0, 1, \ldots, n-1\}.$$

Note that $k'$ runs over all positions of the rough scaled grid in between the volatility does not possess a jump. This implies that for proving this theorem, it suffices to establish the following two convergences:

$$P\left( \max_{k,j} \frac{n \Delta Y_{k,j}}{\widehat{\sigma}_k} - \max_{k',j} \frac{n \Delta Y_{k',j}}{\widehat{\sigma}_{k'}} > 0 \right) \to 0, \quad n \to \infty \tag{1.63}$$

and

$$P\left( \max_{k,j} Z_{k,j} - \max_{k',j} Z_{k',j} > 0 \right) \to 0, \quad n \to \infty. \tag{1.64}$$

This is due to the inequality

$$\begin{aligned} &\left| \max_{k,j} \frac{n \Delta Y_{k,j}}{\widehat{\sigma}_k} - \max_{k,j} Z_{k,j} \right| \\ \leqslant \; & \max_{k',j} \left| \frac{n \Delta Y_{k',j}}{\widehat{\sigma}_{k'}} - Z_{k',j} \right| + \left( \max_{k,j} \frac{n \Delta Y_{k,j}}{\widehat{\sigma}_k} - \max_{k',j} \frac{n \Delta Y_{k',j}}{\widehat{\sigma}_{k'}} \right) + \left( \max_{k,j} Z_{k,j} - \max_{k',j} Z_{k',j} \right), \quad n \in \mathbb{N}. \end{aligned}$$

The following is divided into three steps. The first two steps prove (1.63). The first step simplifies the claim to a more elementary result which involves only the maximum of $\mathrm{N}(0,1)$ i.i.d. random variables. This simplified result is proven in the second step. Finally, in the third step, the first two steps are used in order to prove (1.64).

STEP 1. *Simplification of the claim.* Set

$$\eta_k \overset{\text{def}}{=} \frac{\widehat{\sigma}_k - \sigma_k}{\sigma_k}, \quad \zeta_{k,j} \overset{\text{def}}{=} n \Delta Y_{k,j} - \sigma_k Z_{k,j}, \quad 0 \leqslant k, j < n$$

and define for any fixed $0 < \gamma < \alpha \wedge \frac{1}{2}$

$$\begin{aligned} A_n &\overset{\text{def}}{=} \bigcap_{l=0}^{N-1} \left\{ S_{l+1} - S_l > \frac{1}{n} \right\} \cap \left\{ V > \left( 2Kn^{-\frac{\alpha}{2}} \right) \vee n^{-\frac{\gamma}{2}} \right\}, \\ B_n &\overset{\text{def}}{=} \left\{ \max_k \left| \frac{1}{n} \sum_{j=0}^{n-1} Z_{k,j}^2 - 1 \right| \leqslant \frac{1}{n^\gamma}, \; \max_{k'} |\eta_{k'}| \leqslant \frac{1}{\eta^\gamma}, \; \max_{k',j} |\zeta_{k',j}| \leqslant \frac{1}{n^\gamma}, \; \max_{k,j} Z_{k,j} > 0 \right\}, \end{aligned}$$

with the notation $\{M_1, M_2\} \stackrel{\text{def}}{=} M_1 \cap M_2$ for any sets $M_1, M_2$. Furthermore, we have with

$$\widetilde{k} \in \bigcup_{l=1}^{N} \{K_{l,n}\}$$

the inequality

$$\begin{aligned} & P\left(\max_{k,j} \frac{n\Delta Y_{k,j}}{\widehat{\sigma}_k} - \max_{k',j} \frac{n\Delta Y_{k',j}}{\widehat{\sigma}_{k'}} > 0\right) \\ \leqslant \; & P\left(\max_{\widetilde{k},j} \frac{n\Delta Y_{\widetilde{k},j}}{\widehat{\sigma}_{\widetilde{k}}} > \max_{k',j} \frac{n\Delta Y_{k',j}}{\widehat{\sigma}_{k'}},\, A_n,\, B_n\right) + P(A_n^c) + P(B_n^c). \end{aligned} \tag{1.65}$$

The proof of Theorem 1.4.2 shows $P(B_n^c) \to 0$, $n \to \infty$. Note also

$$\limsup P(A_n^c) \leqslant P(\limsup A_n^c) = P((\liminf A_n)^c) = P(\varnothing) = 0,$$

i.e. $P(A_n^c) \to 0$, $n \to \infty$. Since we aim to verify (1.63), the above yields that it is sufficient to prove that

$$\sum_{l=1}^{\infty} \sum_{g \in G_{l,n}} P\left(\max_{\widetilde{k},j} \frac{n\Delta Y_{\widetilde{k},j}}{\widehat{\sigma}_{\widetilde{k}}} - \max_{k',j} \frac{\Delta Y_{k',j}}{\widehat{\sigma}_{k'}} > 0,\ (K_{1,n}, \ldots, K_{l,n}) = g,\ N = l,\ A_n,\ B_n\right)$$

tends to zero, if $n$ tends to infinity. Define for this purpose

$$\begin{aligned} \lambda_k & \stackrel{\text{def}}{=} \inf_{\{\frac{k}{n} \leqslant s \leqslant \frac{k+1}{n}\}} \sigma_s, \quad 0 \leqslant k < n, \\ \delta & \stackrel{\text{def}}{=} \sup_{l \geqslant 1} |\Delta \sigma_{S_l}|\, \mathbb{1}_{\{S_l \leqslant 1\}} \leqslant \epsilon(\sqrt{2} - 1)V. \end{aligned}$$

Then we have for $\omega \in A_n$ and every $0 \leqslant k < n$, i.e. in particular for $\widetilde{k}$

$$\lambda_k(\omega) \leqslant \sigma_s(\omega) \leqslant \lambda_k(\omega) + 2K(\omega)n^{-\alpha} + \delta(\omega) \stackrel{\text{def}}{=} \lambda_k(\omega) + \delta_n(\omega), \quad \frac{k}{n} \leqslant s \leqslant \frac{k+1}{n}.$$

This yields the estimates

$$\lambda_k^2 \frac{1}{n} \sum_{j=0}^{n-1} Z_{k,j}^2 \leqslant \widehat{\sigma}_k^2 \leqslant (\lambda_k + \delta_n)^2 \frac{1}{n} \sum_{j=0}^{n-1} Z_{k,j}^2, \quad \omega \in A_n$$

because of

$$\widehat{\sigma}_k^2 = n \sum_{j=0}^{n-1} (\Delta Y_{k,j})^2 = n \sum_{j=0}^{n-1} \int_{t_{k,j}}^{t_{k,j} + \frac{1}{n^2}} \sigma_s^2 \, ds \, Z_{k,j}^2.$$

Using this we can write with $C_n \stackrel{\text{def}}{=} A_n \cap B_n$, $n \in \mathbb{N}$ the inequalities

$$P\left(\max_{\widetilde{k},j} \frac{n\Delta Y_{\widetilde{k},j}}{\widehat{\sigma}_{\widetilde{k}}} > \max_{k',j} \frac{n\Delta Y_{k',j}}{\widehat{\sigma}_{k'}},\, (K_{1,n}, \ldots, K_{l,n}) = g,\, N = l,\, C_n\right) \tag{1.66}$$

$$\leqslant \; P\left(\max_{\widetilde{k},j} \frac{n\left(\frac{1}{n^2}(\lambda_{\widetilde{k}}+\delta_n)^2\right)^{\frac{1}{2}} Z_{\widetilde{k},j}}{\left(\lambda_{\widetilde{k}}^2\left(1-\frac{1}{n^\gamma}\right)\right)^{\frac{1}{2}}} > \max_{k',j} \frac{\sigma_{k'} Z_{k',j}+\zeta_{k',j}}{\sigma_{k'}(1+\eta_{k'})}, (K_{1,n},\dots,K_{l,n})=g,\, N=l,\, C_n\right)$$

$$\leqslant \; P\left(\max_{\widetilde{k},j} \frac{\lambda_{\widetilde{k}}+\delta_n}{\lambda_{\widetilde{k}}} \frac{Z_{\widetilde{k},j}}{\left(1-\frac{1}{n^\gamma}\right)^{\frac{1}{2}}} > \max_{k',j} \frac{Z_{k',j}}{1+\frac{1}{n^\gamma}} - \frac{\frac{1}{n^\gamma}}{V\left(1-\frac{1}{n^\gamma}\right)}, (K_{1,n},\dots,K_{l,n})=g,\, N=l,\, C_n\right).$$

We have on $A_n$ for $n$ large enough

$$\frac{\lambda_{\widetilde{k}}+\delta_n}{\lambda_{\widetilde{k}}} \leqslant 1+\frac{n^{-\frac{\alpha}{2}}V+\epsilon(\sqrt{2}-1)V}{V} = \sqrt{2}+n^{-\frac{\alpha}{2}}-(\sqrt{2}-1)(1-\epsilon) \leqslant \sqrt{2-\kappa}$$

for some constant $\kappa > 0$. So if $g \in G_{l,n}$, (1.66) is not larger than

$$P\left(\frac{\sqrt{2-\kappa}}{\left(1-\frac{1}{n^\gamma}\right)^{\frac{1}{2}}} \max_{\widetilde{k},j} Z_{\widetilde{k},j} > \frac{1}{1+\frac{1}{n^\gamma}} \max_{k',j} Z_{k',j} - \frac{2}{n^{\frac{\gamma}{2}}}\right) P((K_{1,n},\dots,K_{l,n})=g,\; N=l)$$

where the latter maxima run over

$$\widetilde{k} \in \bigcup_{j=1}^{l}\{g_j\} \quad \text{resp.} \quad k' \in \{0,1,\dots,n-1\} - \bigcup_{j=1}^{l}\{g_j\}.$$

Here, we used the independence of $Z$ and $\sigma$. Define

$$D_{l,n}^g \overset{\text{def}}{=} \left\{\frac{\sqrt{2-\kappa}}{\left(1-\frac{1}{n^\gamma}\right)^{\frac{1}{2}}} \max_{\widetilde{k},j} Z_{\widetilde{k},j} > \frac{1}{1+\frac{1}{n^\gamma}} \max_{k',j} Z_{k',j} - \frac{2}{n^{\frac{\gamma}{2}}}\right\}, \quad l,n \geqslant 1,\, g \in G_{l,n}.$$

It suffices to prove $P(D_{l,n}^\pi) \to 0$, $n \to \infty$ with $\pi \overset{\text{def}}{=} (0,1,\dots,l-1) \in G_{l,n}$ for every fixed $l \geqslant 1$ because

$$\begin{aligned} &\sum_{l=1}^{\infty} \sum_{g\in G_{l,n}} P(D_{l,n}^g) P((K_{1,n},\dots,K_{l,n})=g,\, N=l) \\ =\; &\sum_{l=1}^{\infty} P(D_{l,n}^\pi) \sum_{g\in G_{l,n}} P((K_{1,n},\dots,K_{l,n})=g,\, N=l) \\ \leqslant\; &\sum_{l=1}^{\infty} P(D_{l,n}^\pi) P(N=l) \end{aligned}$$

and

$$P(D_{l,n}^\pi)P(N=l) \leqslant P(N=l), \quad \sum_{l=1}^{\infty} P(N=l) \leqslant 1 < \infty$$

holds. Finally, a dominated convergence argument, (1.65) and the results proven so far yields the desired convergence (1.63).

STEP 2. *Convergence of* $(D^{\pi}_{l,n})_n$. We know that

$$\alpha_{l,n} \overset{\text{def}}{=} \max_{\widetilde{k},j} Z_{\widetilde{k},j} - b_{ln} \to 0, \quad \beta_{l,n} \overset{\text{def}}{=} \max_{k',j} Z_{k',j} - b_{(n-l)n} \to 0, \quad n \to \infty \quad P\text{-stoch.}$$

Due to

$$\begin{aligned} & P(D^{\pi}_{l,n}) \\ = \;& P\left( \frac{\sqrt{2-\kappa}}{\left(1-\frac{1}{n^{\gamma}}\right)^{\frac{1}{2}}}(\alpha_{l,n}+b_{ln}) > \frac{1}{1+\frac{1}{n^{\gamma}}}(\beta_{l,n}+b_{(n-l)n}) - \frac{2}{n^{\frac{\gamma}{2}}} \right) \\ = \;& P\left( \frac{\sqrt{2-\kappa}}{\left(1-\frac{1}{n^{\gamma}}\right)^{\frac{1}{2}}}\alpha_{l,n} - \frac{1}{1+\frac{1}{n^{\gamma}}}\beta_{l,n} > \frac{1}{1+\frac{1}{n^{\gamma}}}b_{(n-l)n} - \frac{\sqrt{2-\kappa}}{\left(1-\frac{1}{n^{\gamma}}\right)^{\frac{1}{2}}}b_{ln} - \frac{2}{n^{\frac{\gamma}{2}}} \right) \end{aligned}$$

and the stochastic convergence of $(\alpha_{l,n})_n$ and $(\beta_{l,n})_n$ to zero, it suffices to validate

$$\frac{1}{1+\frac{1}{n^{\gamma}}}b_{(n-l)n} - \frac{\sqrt{2-\kappa}}{\left(1-\frac{1}{n^{\gamma}}\right)^{\frac{1}{2}}}b_{ln} \to \infty, \quad n \to \infty. \tag{1.67}$$

Substituting (1.31) in (1.67) yields

$$\begin{aligned} & \frac{1}{1+\frac{1}{n^{\gamma}}}\sqrt{2\log((n-l)n)} - \sqrt{\frac{2-\kappa}{1-\frac{1}{n^{\gamma}}}}\sqrt{2\log(ln)} + o(1), \quad n \to \infty \\ = \;& \sqrt{2\log\left(((n-l)n)^{\frac{1}{\left(1+\frac{1}{n^{\gamma}}\right)^2}}\right)} - \sqrt{2\log\left((ln)^{\frac{2-\kappa}{1-\frac{1}{n^{\gamma}}}}\right)} + o(1). \end{aligned}$$

Since

$$\frac{1}{\left(1+\frac{1}{n^{\gamma}}\right)^2} \to 1, \quad \frac{2-\kappa}{1-\frac{1}{n^{\gamma}}} \to 2-\kappa < 2, \quad n \to \infty,$$

step 2 is accomplished.

STEP 3. *Proof of (1.64).* Write as at the end of step 1

$$\begin{aligned} & P\left(\max_{k,j} Z_{k,j} - \max_{k',j} Z_{k',j} > 0,\, A_n\right) \\ \leqslant \;& \sum_{l=1}^{\infty}\sum_{g\in G_{l,n}} P\left(\max_{\widetilde{k},j} Z_{\widetilde{k},j} > \max_{k',j} Z_{k',j},\, (K_{1,n},\ldots,K_{l,n}) = g,\, N = l\right) \\ = \;& \sum_{l=1}^{\infty}\sum_{g\in G_{l,n}} P\left(\max_{\widetilde{k},j} Z_{\widetilde{k},j} > \max_{k',j} Z_{k',j}\right) P((K_{1,n},\ldots,K_{l,n}) = g,\, N = l). \end{aligned}$$

Again it suffices to establish

$$P\left(\max_{\widetilde{k},j} Z_{\widetilde{k},j} > \max_{k',j} Z_{k',j}\right) \to 0, \quad n \to \infty \tag{1.68}$$

for every fixed $l \in \mathbb{N}$ and $g \in G_{l,n}$. It appears that the proof of (1.68) is a simpler version of what was performed for the second step (set $\kappa = 1$). □

In the following, we demonstrate that the bound $(\sqrt{2}-1)V$ is sharp in a certain sense. Our main result in this context is Corollary 1.5.7 which is a stochastic generalization of the following Proposition 1.5.6. We illustrate that the convergence to the Gumbel distribution does not have to hold, if there is an oversized jump in the volatility process at some irrational position. Such an *irrational* jump position causes some problems since our grid consists of equidistant *rational* points. Compare the proof of the next Proposition 1.5.6 for an rigorous argumentation.

**Proposition 1.5.6.** *Let $h$ and $c$ be two numbers, such that $h > \sqrt{2}$ and $0 < c < 1$ is an irrational number. Set*

$$\sigma_t \stackrel{\text{def}}{=} h\,\mathbb{1}_{[0,c]}(t) + \mathbb{1}_{(c,1]}(t), \quad 0 \leqslant t \leqslant 1$$

*and $b = 0$. Then, there is a sequence $(n_l)_l$ of natural numbers, such that $n_l \uparrow \infty$ and*

$$P(T_{n_l} - b_{n_l^2} \geqslant \epsilon) \to 1, \quad l \to \infty \tag{1.69}$$

*for all $\epsilon > 0$. This implies in particular*

$$a_{n^2}(T_n - b_{n^2}) \not\to \mathcal{G}, \quad n \to \infty.$$

*Proof.* We use the fact that the spot volatility estimator $\widehat{\sigma}_k$ in Redefinition 1.4.1 estimates the average value of the spot volatility in the interval $\left[\frac{k}{n}, \frac{k+1}{n}\right]$. Thus, if the spot volatility jumps in this interval, we make obviously an error depending on the jump size. Our intention in the following is to make this error as large as possible to get the negative convergence result (1.69).

The proof is divided into two steps. Similar to the proof of Theorem 1.5.3, the first step simplifies our claim, so that it remains to prove a more elementary result which involves only the maximum of N(0, 1) i.i.d. random variables. We prove this result in the second step.

STEP 1. *Simplification of the claim.* We have

$$Y_t = \int_0^t \sigma_s\, dW_s = \begin{cases} hW_t, & t \leqslant c \\ W_t - W_c + hW_c, & t > c. \end{cases}$$

Let

$$f_n(t) = \lfloor n^2 t \rfloor - n\lfloor nt \rfloor, \quad 0 < t < 1$$

denote the fine scale position of $t$, compare with the beginning of Section 1.1, and choose $0 \leqslant k < n$, such that $c \in \left(\frac{k}{n}, \frac{k+1}{n}\right)$. Then we can write

$$\begin{aligned} \widehat{\sigma}_k^2 &= n \sum_{j=0}^{f_n(c)-1} (\Delta Y_{k,j})^2 + \epsilon_{k,n} + n \sum_{j=f_n(c)+1}^{n-1} (\Delta Y_{k,j})^2, \quad 0 \leqslant \epsilon_{k,n} \leqslant \frac{h^2}{n} Z^2_{k,f_n(c)} \\ &= \frac{h^2}{n} \sum_{j=0}^{f_n(c)-1} Z^2_{k,j} + \frac{1}{n} \sum_{j=f_n(c)+1}^{n-1} Z^2_{k,j} + \epsilon_{k,n}. \end{aligned}$$

Next set

$$r \stackrel{\text{def}}{=} \frac{h-\sqrt{2}}{4(h^2-1)} \in (0,1) \tag{1.70}$$

and note that Lemma A.5 implies that we have two sequences $(n_l)_l$ and $(k_l)_l$ of natural numbers, such that

$$\left\lfloor \frac{r}{2} n_l \right\rfloor \leqslant f_{n_l}(c) \leqslant \lfloor r n_l \rfloor, \quad c \in \left( \frac{k_l}{n_l}, \frac{k_l+1}{n_l} \right), \, l \in \mathbb{N}$$

and $n_l \uparrow \infty$. This implies together with the weak law of large numbers

$$\widehat{\sigma}_{k_l}^2 \leqslant \frac{h^2}{n_l} \sum_{j=0}^{\lfloor r n_l \rfloor} Z_{k_l,j}^2 + \frac{1}{n_l} \sum_{j=\lfloor r n_l \rfloor + 1}^{n_l - 1} Z_{k_l,j}^2 \to r h^2 + (1-r), \quad l \to \infty \quad (P\text{-stoch.})$$

which implicates

$$\begin{aligned} P\left(\widehat{\sigma}_{k_l}^2 \geqslant 1 + 2r(h^2-1)\right) &\leqslant P\left( \left| \frac{h^2}{n_l} \sum_{j=0}^{\lfloor r n_l \rfloor} Z_{k_l,j}^2 + \frac{1}{n_l} \sum_{j=\lfloor r n_l \rfloor + 1}^{n_l - 1} Z_{k_l,j}^2 - 1 - r(h^2-1) \right| \geqslant r(h^2-1) \right) \\ &\to 0, \quad l \to \infty. \end{aligned} \tag{1.71}$$

Next define

$$\lambda \overset{\text{def}}{=} \frac{h}{1 + 2r(h^2-1)} > \sqrt{2},$$

cf. (1.70) and

$$r_l \overset{\text{def}}{=} \frac{\lfloor \frac{r}{2} n_l \rfloor}{n_l} \geqslant \frac{r}{2} - \frac{1}{n_l}, \quad l \geqslant 1.$$

And note that we have for arbitrary $\epsilon > 0$ the inequalities

$$\begin{aligned} P\left(T_{n_l} - b_{n_l^2} \geqslant \epsilon\right) &\geqslant P\left( n_l \max_{0 \leqslant j < \lfloor \frac{r}{2} n_l \rfloor} \frac{\Delta Y_{k_l,j}}{\widehat{\sigma}_{k_l}} - b_{n_l^2} \geqslant \epsilon \right) \\ &\geqslant P\left( \lambda \max_{0 \leqslant j < r_l n_l} Z_{k_l,j} - b_{n_l^2} \geqslant \epsilon, \, \widehat{\sigma}_{k_l}^2 \leqslant 1 + 2r(h^2-1) \right). \end{aligned}$$

Thus regarding (1.71), it suffices to establish

$$P\left( \lambda \max_{0 \leqslant j < r_l n_l} Z_{k_l,j} - b_{n_l^2} \geqslant \epsilon \right) \to 1, \quad l \to \infty.$$

We complete this in the second step.

STEP 2. *Divergence to infinity of a $\lambda$-scaled partial-maximum.* Crucial in what follows is that we have the lower bound $\lambda > \sqrt{2}$, which is due to the choice of $r$. The notations

$$M_l \overset{\text{def}}{=} \max_{0 \leqslant j < r_l n_l} Z_{k_l,j}, \quad A_l \overset{\text{def}}{=} a_{r_l n_l} \left( M_l - b_{r_l n_l} \right), \quad l \geqslant 1$$

are used in the following. We know that $A_l \overset{\text{d}}{\to} \mathcal{G}$, $l \to \infty$, cf. Lemma 1.1.7 in Haan, Ferreira [17], and write

$$\begin{aligned} P(\lambda M_l - b_{n_l^2} \geqslant \epsilon) &= P(\lambda(A_l + a_{r_l n_l} b_{r_l n_l}) - a_{r_l n_l} b_{n_l^2} \geqslant a_{r_l n_l} \epsilon) \\ &= P\left( A_l \geqslant \frac{1}{\lambda} \left( a_{r_l n_l} b_{n_l^2} + a_{r_l n_l} \epsilon \right) - a_{r_l n_l} b_{r_l n_l} \right). \end{aligned}$$

Obviously it suffices to establish

$$\frac{1}{\lambda}b_{n_l^2} - b_{r_l n_l} \to -\infty, \quad l \to \infty.$$

This follows after a substitution of (1.31), i.e.

$$\begin{aligned} b_{n_l^2} - \lambda b_{r_l n_l} &= \sqrt{2\log n_l^2} - \lambda\sqrt{2\log(r_l n_l)} + o(1), \quad l \to \infty \\ &= 2\sqrt{\log n_l} - \sqrt{2}\lambda\sqrt{\log r_l + \log n_l} + o(1) \\ &= 2\sqrt{\log n_l}\left(1 - \underbrace{\frac{\lambda}{\sqrt{2}}}_{>1}\underbrace{\sqrt{\frac{\log r_l}{\log n_l} + 1}}_{\to 1}\right) + o(1) \\ &\to -\infty, \quad l \to \infty. \end{aligned}$$

□

**Corollary 1.5.7.** *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leqslant t \leqslant 1}, P)$ be a filtered probability space. Furthermore, assume that $W$ is a $(\mathcal{F}_t)$-adapted Brownian motion on this space and that there are two random variables $S, H : \Omega \to \mathbb{R}$, such that $(S, H)$ is $\mathcal{F}_0$ measurable and independent of $W$. Next assume that the distribution of $(S, H)$ has an atom at some point $(c, h)$. To be more precise, it is stated*

$$P((S, H) = (c, h)) > 0$$

*for some pair $(c, h)$ with*

$$0 < c < 1, \, c \notin \mathbb{Q}, \quad h > \sqrt{2}.$$

*Further set*

$$\begin{aligned} \sigma_t^{(S,H)} &\stackrel{\text{def}}{=} H\,\mathbb{1}_{[0,S]}(t) + \mathbb{1}_{(S,1]}(t), \\ Y_t^{(S,H)} &\stackrel{\text{def}}{=} \int_0^t \sigma_s^{(S,H)}\,dW_s, \quad 0 \leqslant t \leqslant 1 \end{aligned}$$

*and define $\left(T_n^{(S,H)}\right)_n$ analogue to $(T_n)_n$ as a function of $\left(Y_t^{(S,H)}\right)_t$. Then there is a sequence $(n_l)$ of natural numbers with $n_l \uparrow \infty$, such that*

$$P\left(T_{n_l}^{(S,H)} - b_{n_l^2} \geqslant \epsilon\right) \to 1, \quad l \to \infty$$

*for all $\epsilon > 0$. This implies in particular*

$$a_{n^2}(T_n^{(S,H)} - b_{n^2}) \not\to \mathcal{G}, \quad n \to \infty.$$

**Remark 1.5.8.** The assumptions of Corollary 1.5.7 basically say that the volatility jumps at the position $S$ with the jumpsize $H - 1$ if $0 < S \leqslant 1$ and $H \neq 1$. Furthermore, there is a positive probability that $\sigma$ jumps at some irrational position with a jumpsize larger than $\sqrt{2} - 1$. Note also that the existence of a filtration as stated in Corollary 1.5.7 does not cause any problems. This is due to the fact that if $\mathcal{H} \subseteq \mathcal{F}$ is a sub-$\sigma$-algebra, which is independent of $\mathcal{F}_1$, then

$$(W_t, \mathcal{H}_t)_{0 \leqslant t \leqslant 1}, \quad \mathcal{H}_t \stackrel{\text{def}}{=} \sigma(\mathcal{H} \cup \mathcal{F}_t), \quad 0 \leqslant t \leqslant 1$$

is also a Brownian motion. We have to consider such technical subtleties because the construction of the Itô integral needs $\sigma$ to be $(\mathcal{F}_t)_t$ adapted.

*Proof of Corollary 1.5.7.* Using the independence of $(S, H)$ and $W$, $P((S, H) = (c, h)) > 0$ and the statement of Proposition 1.5.6 we can write for any $\epsilon > 0$

$$\begin{aligned} P\left(T_n^{(S,H)} - b_{n^2} \geqslant \epsilon\right) &= \int_{\mathbb{R}^2} P\left(T_n^{(S,H)} - b_{n^2} \geqslant \epsilon \middle| (S,H) = (s,u)\right) dP^{(S,H)}(s,u) \\ &\geqslant P((S,H) = (c,h)) P\left(T_n^{(S,H)} - b_{n^2} \geqslant \epsilon \middle| (S,H) = (c,h)\right) \\ &= P((S,H) = (c,h)) P\left(T_n^{(c,h)} - b_{n^2} \geqslant \epsilon\right). \end{aligned}$$

This implies with the same sequence $(n_l)_l$ as in Proposition 1.5.6 the divergence

$$P\left(T_{n_l}^{(S,H)} - b_{n_l^2} \geqslant \epsilon\right) \to 1, \quad l \to \infty$$

and the statement of the non-holding weak convergence to the Gumbel distribution follows from Slutsky's theorem together with $a_{n^2} \uparrow \infty$ as $n \to \infty$. □

## 1.6 External jumps and divergence to infinity

In this section, the additional existence of external jumps is investigated. The behaviour of the statistic $a_{n^2}(T_n - b_{n^2})$ in the presence of external jumps is of particular interest in regard to Theorem 1.2.5 and Theorem 1.5.3. It turns out that this statistic converges to infinity under appropriate assumptions, compare Theorem 1.6.1 and Theorem 1.6.3 in this section. So we can use this statistic in order to distinguish between the jump case (convergence to infinity) and the non-jump case (convergence to the Gumbel distribution). We investigate by use of numerical simulations the finite sample behaviour of the resulting statistical test in the next Section 1.7. In order to get asymptotic correct confidence intervals, we use the quantiles of the Gumbel distribution.

We begin with a quite general result about semimartingales. Note that semimartingales have in general infinite activity jump paths. Observe for this that a Lévy process is some kind of a prototype semimartingale, cf. Jacod, Shiryaev [20][Chapter II, § 4c] and that every Lévy process which is not compound Poisson, has infinite activity jump paths, cf. Cont, Tankov [11][Proposition 3.3].

**Theorem 1.6.1.** *Let $(X_t, \mathcal{F}_t)_{t \in [0,1]}$ be a semimartingale with càdlàg paths and*

$$\Lambda \stackrel{\text{def}}{=} \{\omega \in \Omega : \exists t_0 \in (0, 1] : \Delta X_{t_0}(\omega) > 0\}. \tag{1.72}$$

*Let $\gamma < \frac{1}{2}$. We assert the convergence*

$$n^{-\gamma} a_{n^2}(T_n - b_{n^2}) \to \infty \quad P\text{-stoch. on } \Lambda \tag{1.73}$$

*where $a_n, b_n$ are defined as in (1.31), and $T_n$ is defined as in (1.32) with $(X_t)_t$ instead of $(\check{Y}_t)_t$.*

*Proof.* First, let us validate that $\Lambda \in \mathcal{F}$: Choose a sequence of stopping times $(T_n)_{n\in\mathbb{N}}$ that exhausts the jumps of $X$, cf. Chapter I, Proposition 1.32 in Jacod, Shiryaev [20]. Then,

$$\Lambda = \bigcup_{n,m\in\mathbb{N}} \left\{\Delta X_{T_n} > \frac{1}{m}\right\} \in \mathcal{F}_1 \subseteq \mathcal{F}.$$

Define two functions $k, j : (0,1] \to \{0,\dots,n-1\}$ by

$$k(t) = (\lceil nt\rceil - 1), \quad j(t) = (\lceil n^2 t\rceil - 1) - nM(t), \quad t \in (0,1]$$

and set as usual

$$\Delta X_{k,j} = X_{t_{k,j}+\frac{1}{n^2}} - X_{t_{k,j}}, \quad 0 \leqslant k, j < n.$$

Choose $\hat{\omega} \in \Lambda$ and let $\eta \in (0,1]$ be such that $\Delta X_\eta(\hat{\omega}) > 0$. Then we have

$$\liminf_n \max_{k,j} \Delta X_{k,j}(\hat{\omega}) \geqslant \liminf_n \Delta X_{k(\eta),j(\eta)}(\hat{\omega}) = \Delta X_\eta(\hat{\omega}) > 0 \tag{1.74}$$

where the equality sign above holds because the paths of $(X_t)_{t\in[0,1]}$ are càdlàg. We establish for arbitrary $L \in \mathbb{R}$ and $\gamma < \frac{1}{2}$ the convergence

$$P(\{n^{-\gamma} a_{n^2}(T_n - b_{n^2}) \leqslant L\} \cap \Lambda) \to 0, \quad n \to \infty.$$

Suppose the opposite, i.e. the existence of a subsequence $(n_r)_r$ is assumed, such that

$$P(\{n_r^{-\gamma} a_{n_r^2}(T_{n_r} - b_{n_r^2}) \leqslant L\} \cap \Lambda) \to \eta > 0, \quad r \to \infty \tag{1.75}$$

holds. Due to

$$\sum_{k=0}^{n-1}\sum_{j=0}^{n-1} (\Delta X_{k,j})^2 \to [X,X]_1, \quad n \to \infty \quad (P\text{-stoch.}),$$

cf. Chapter I, Theorem 4.47 in Jacod, Shiryaev [20], there exists a subsequence $(m_l)_l$ of $(n_r)_r$, such that we obtain

$$\sum_{k=0}^{m_l-1}\sum_{j=0}^{m_l-1} (\Delta X_{k,j})^2 \to [X,X]_1, \quad s \to \infty \quad (P\text{-a.s.}).$$

This yields together with

$$|\Delta X_{k,j}||\Delta X_{k,j+1}| \leqslant (\Delta X_{k,j})^2 + (\Delta X_{k,j+1})^2$$

the equation

$$P\left(\limsup_l \sum_{k=0}^{m_l-1}\sum_{j=0}^{m_l-2} |\Delta X_{k,j}||\Delta X_{k,j+1}| < \infty\right) = 1. \tag{1.76}$$

Next consider the inclusions

$$\begin{aligned} &\limsup_l \{m_l{}^{-\gamma} a_{m_l^2}(T_{m_l} - b_{m_l^2}) \leqslant L\} \cap \Lambda \\ = \;&\limsup_l \left\{T_{m_l} \leqslant \frac{L m_l{}^\gamma}{a_{m_l^2}} + b_{m_l^2}\right\} \cap \Lambda \end{aligned}$$

$$\begin{aligned}
&\subseteq \limsup_l \{T_{m_l} \leqslant m_l{}^{\gamma}\} \cap \Lambda \\
&= \limsup_l \left\{ \frac{\sqrt{2}\sqrt{m_l - 1}}{\sqrt{\pi}} \max_{k,j} \frac{\Delta X_{k,j}}{(\sum_{i=0}^{m_l-2} |\Delta X_{k,i}||\Delta X_{k,i+1}|)^{\frac{1}{2}}} \leqslant m_l{}^{\gamma} \right\} \cap \Lambda \\
&\subseteq \limsup_l \left\{ \max_{k,j} \frac{\Delta X_{k,j}}{(\sum_{i=0}^{m_l-2} |\Delta X_{k,i}||\Delta X_{k,i+1}|)^{\frac{1}{2}}} \leqslant 2 m_l{}^{\gamma - \frac{1}{2}} \right\} \cap \Lambda \\
&\subseteq \limsup_l \left\{ \max_{k,j} \Delta X_{k,j} \leqslant 2 m_l{}^{\gamma-\frac{1}{2}} \left( \sum_{k=0}^{m_l-1} \sum_{j=0}^{m_l-2} |\Delta X_{k,j}||\Delta X_{k,j+1}| \right)^{\frac{1}{2}} \right\} \cap \Lambda. \qquad (1.77)
\end{aligned}$$

The $P$-measure of the set in (1.77) is zero because of $\gamma - \frac{1}{2} < 0$, (1.74) and (1.76). Thus, the above inclusions yield together with Fatou's Lemma

$$\limsup_l P(\{m_l{}^{-\gamma} a_{m_l^2}(T_{m_l} - b_{m_l^2}) \leqslant L\} \cap \Lambda) \leqslant P(\limsup_l \{m_l{}^{-\gamma} a_{m_l}(T_{m_l} - b_{m_l}) \leqslant L\} \cap \Lambda) = 0,$$

which is a contradiction to (1.75) because of $(m_l)_l \subseteq (n_r)_r$. □

**Remark 1.6.2.** Note that (1.76) is the crucial property that we need for the proof of the above Theorem. This is a quite general assumption which is, for example, fulfilled by a semimartingal as proven above. Naturally, this raises the question, whether we can improve the convergence rate $n^{-\frac{1}{2}}$ in (1.73) under stronger assumptions. Assume for this purpose that we observe a process

$$\widetilde{Y} \stackrel{\text{def}}{=} \check{Y} + J$$

instead of a general semimartingale $X$. Here, $\check{Y}$ is as usual and $J$ denotes an additive, finite activity, external jump process. Then we obtain the heuristic

$$n \sum_{j=0}^{n-2} |\Delta \widetilde{Y}_{k,j}||\Delta \widetilde{Y}_{k,j+1}| = n \sum_{j=0}^{n-2} |\Delta \check{Y}_{k,j} + J_{k,j}||\Delta \check{Y}_{k,j+1} + J_{k,j+1}| = O_P(1). \qquad (1.78)$$

Note for this that every couple of neighbouring increments possesses at most one jump if the increment size is small enough, i.e. $n$ is large enough. This is due to the fact that there are only finite many jumps in each path. Using (1.78) we easily get the heuristic

$$n \max_{k,j} \frac{\Delta \widetilde{Y}_{k,j}}{\widetilde{\sigma}_k} = O_P(n) \qquad (1.79)$$

with the self-explanatory notation $(\widetilde{\sigma}_k)_{0 \leqslant k < n}$. Hence (1.73) should be true for each $\gamma < 1$ instead of $\gamma < \frac{1}{2}$, i.e. we have the convergence rate $n^{-1}$.

Note that, on the other hand, for the quadratic variation estimator in Redefinition 1.4.1 there is the heuristic

$$\widetilde{\sigma}_k^2 = n \sum_{j=0}^{n-1} (\Delta \widetilde{Y}_{k,j})^2 = n \sum_{j=0}^{n-1} (\Delta \check{Y}_{k,j} + \Delta J_{k,j})^2 = O_P(n),$$

so that only a convergence rate of $n^{-\frac{1}{2}}$ can be expected, compare (1.79). Observe that the proof of Theorem 1.6.1 also works with the quadratic variation estimator. Thus, the above heuristic implies that Theorem 1.6.1 with the quadratic variation estimator yields the optimal convergence rate $n^{-\frac{1}{2}}$, despite we have in general infinite activity jump paths.

The following Theorem 1.6.3 proves rigor to the above heuristic.

**Theorem 1.6.3.** *Let Assumptions 1.2.3 hold. Assume furthermore that there are two families* $(H_l)_{l\in\mathbb{N}}$ *and* $(R_l)_{l\in\mathbb{N}}$ *of random variables on the same probability space with*

$$R_l(\omega)\uparrow\infty,\quad 0<R_l(\omega),\quad R_l(\omega)<R_{l+1}(\omega),\quad l\in\mathbb{N},\quad \omega\in\Omega$$

*and* $(H_l)$ *arbitrary. Next, define a finite activity, pure jump process*

$$J_t\stackrel{\text{def}}{=}\sum_{l=1}^{\infty}H_l\,\mathbb{1}_{[R_l,\infty)}(t),\quad 0\leqslant t\leqslant 1,$$

*set*

$$\widetilde{Y}_t\stackrel{\text{def}}{=}\check{Y}_t+J_t,\quad 0\leqslant t\leqslant 1$$

*and define* $(T_n)$ *in the usual manner as a function of* $\widetilde{Y}$. *Fix any* $\gamma<1$. *Then it holds the divergence*

$$n^{-\gamma}a_{n^2}(T_n(\omega)-b_{n^2})\to\infty,\quad n\to\infty,\quad \omega\in\Lambda\cap N \tag{1.80}$$

*for a negligible set* $N$. *See (1.72) for the definition of* $\Lambda$.

*Proof.* The proof is divided into three steps. The first step proves two useful $P$-a.s. convergences under the Assumptions 1.1.1. The second step proves this theorem under these Assumptions by using the first step. Finally, the third step generalizes the results to the Assumptions 1.2.3.

STEP 1. *Two useful $P$-a.s. convergences.* Define two functions $k,j\,:\,(0,\infty)\times\mathbb{R}\to\{0,1,\ldots,n\}$ via

$$\begin{aligned}k(t,x)&\stackrel{\text{def}}{=}(\lceil nt\rceil-1)\,\mathbb{1}_{(0,1]\times(0,\infty)}(t,x)+n\,\mathbb{1}_{((0,1]\times(0,\infty))^c}(t,x),\\ j(t,x)&\stackrel{\text{def}}{=}((\lceil n^2t\rceil-1)-nk(t,x))\,\mathbb{1}_{(0,1]\times(0,\infty)}(t,x)+n\,\mathbb{1}_{((0,1]\times(0,\infty))^c}(t,x)\end{aligned}$$

and set

$$k_l^R\stackrel{\text{def}}{=}k(R_l,H_l),\quad j_l^R\stackrel{\text{def}}{=}j(R_l,H_l),\quad l\in\mathbb{N}$$

and

$$\Delta\check{Y}_{k,j}\stackrel{\text{def}}{=}\Delta Y_{k,j}\stackrel{\text{def}}{=}0,\quad \text{if } k\notin\{0,\ldots,n-1\}\text{ or } j\notin\{0,\ldots,n-1\}.$$

We assume first the stronger Assumptions 1.1.1 with vanishing drift (i.e. $b=0$) and show that there exists a negligible set $N\in\mathcal{F}$, such that

$$n^{\gamma}\sum_{j=0}^{n-2}|\Delta Y_{k_l^R,j}||\Delta Y_{k_l^R,j+1}|(\omega)\quad\to\quad 0,\quad n\to\infty \tag{1.81}$$

$$n^{\gamma}\Delta Y_{k_l^R,j_l^R+\rho}(\omega)\quad\to\quad 0,\quad n\to\infty \tag{1.82}$$

holds for every $l\in\mathbb{N}$, $\rho\in\{-1,1\}$ and $\omega\in N^c$. Starting with (1.81) and using the notation in Proposition 1.1.4, we write

$$n^{\gamma}\sum_{j=0}^{n-2}|\Delta Y_{k_l^R,j}||\Delta Y_{k_l^R,j+1}|\quad\leqslant\quad n^{\gamma+1}\max_{k,j}|\Delta Y_{k,j}|^2$$

$$
\begin{aligned}
&= n^{\gamma-1}\max_{k,j}|H_{k,j}+(\sigma_k+\epsilon_{k,j})Z_{k,j}|^2 \\
&\leqslant 2n^{\gamma-1}\left(\max_{k,j}|H_{k,j}|^2+K^2\max_{k,j}|Z_{k,j}|^2\right). \qquad (1.83)
\end{aligned}
$$

Note furthermore that we have, as stated in the Propositions 1.1.4 and A.3, for every $m \in \mathbb{N}$ the upper bounds

$$
P\left(\max_{k,j}|H_{k,j}|^2 \geqslant \epsilon\right) \leqslant n^2\frac{d_2}{n^{4\alpha m}\epsilon^m}, \qquad (1.84)
$$

$$
P\left(\max_{k,j}|Z_{k,j}|^2 \geqslant \epsilon\right) \leqslant \frac{2^m(\log n^2)^{\frac{m}{2}}+2m!}{\epsilon^{\frac{m}{2}}}. \qquad (1.85)
$$

Choosing $m$ large enough and observing $\gamma-1<0$, (1.83), (1.84) and (1.85) yield together with Borel-Cantelli (1.81). Concerning (1.82) we write

$$
n^\gamma|\Delta Y_{k_l^R,j_l^R+\rho}| \leqslant n^{\gamma-1}\max_{k,j}n|\Delta Y_{k,j}| \leqslant n^{\gamma-1}\left(\max_{k,j}|H_{k,j}|+K\max_{k,j}|Z_{k,j}|\right), \quad \rho\in\{-1,1\} \qquad (1.86)
$$

and the claim is proven analogous to (1.81). Next we assume that the drift does not necessarily vanish, and prove that also the check variants of (1.81) and (1.82) hold for some negligible set which is w.l.o.g. equal to $N \in \mathcal{F}$. For this purpose, consider for (1.81) the following estimation of the difference between the check and non-check variant:

$$
\begin{aligned}
&\left|n^\gamma\sum_{j=0}^{n-2}|\Delta\check{Y}_{k_l^R,j}||\Delta\check{Y}_{k_l^R,j+1}| - n^\gamma\sum_{j=0}^{n-2}|\Delta Y_{k_l^R,j}||\Delta Y_{k_l^R,j+1}|\right| \\
\leqslant\; & n^\gamma\sum_{j=0}^{n-2}\left|\Delta\check{Y}_{k_l^R,j}\Delta\check{Y}_{k_l^R,j+1}-\Delta Y_{k_l^R,j}\Delta Y_{k_l^R,j+1}\right| \\
\leqslant\; & n^\gamma\sum_{j=0}^{n-2}\frac{K}{n^2}\left(|\Delta Y_{k_l^R,j}|+|\Delta Y_{k_l^R,j+1}|+\frac{K}{n^2}\right) \\
\leqslant\; & \frac{K^2}{n^{3-\gamma}}+2Kn^{\gamma-1}\max_{k,j}|\Delta Y_{k,j}|.
\end{aligned}
$$

This means, using the proof of (1.81), that the difference tends $P$-a.s. to zero. Hence, the check variant of (1.81) is also established and the check variant of (1.82) is trivial because of the obvious convergence

$$
n^\gamma|\Delta B_{k_l^R,j_l^R+\rho}|(\omega) \leqslant Kn^{\gamma-2}\to 0, \quad n\to\infty, \quad \omega\in\Omega.
$$

STEP 2. *Divergence under the Assumptions 1.1.1.* Set

$$
\widehat{\Lambda} \stackrel{\text{def}}{=} \Lambda\cap N^c
$$

with $N$ as in the first step and set

$$
l = l(\widehat{\omega}) \stackrel{\text{def}}{=} \inf\{m\in\mathbb{N} : H_m(\widehat{\omega})>0\} = \inf\{m\in\mathbb{N} : k_m^R(\widehat{\omega})<n\}, \quad \widehat{\omega}\in\widehat{\Lambda}.
$$

Assume w.l.o.g. $\gamma > 0$. It obviously suffices to show that

$$\widehat{\Lambda} \subseteq \liminf_n \{n^{-\gamma} a_{n^2}(T_n - b_{n^2}) \geqslant L\}$$

holds for all $L > 0$. Fix for this purpose any $L > 0$. Then we have

$$\begin{aligned}
& \liminf_n \{n^{-\gamma} a_{n^2}(T_n - b_{n^2}) \geqslant L\} \cap \widehat{\Lambda} \\
= \ & \liminf_n \left\{ T_n \geqslant \frac{Ln^\gamma}{a_{n^2}} + b_{n^2} \right\} \cap \widehat{\Lambda} \\
\supseteq \ & \liminf_n \{T_n \geqslant n^\gamma\} \cap \widehat{\Lambda} \\
= \ & \liminf_n \left\{ \frac{\sqrt{2}\sqrt{n-1}}{\sqrt{\pi}} \max_{k,j} \frac{\Delta \widetilde{Y}_{k,j}}{(\sum_{i=0}^{n-2} |\Delta \widetilde{Y}_{k,i}||\Delta \widetilde{Y}_{k,i+1}|)^{\frac{1}{2}}} \geqslant n^\gamma \right\} \cap \widehat{\Lambda} \\
\supseteq \ & \liminf_n \left\{ \Delta \widetilde{Y}_{k_l^R, j_l^R} \geqslant 2n^{\gamma - \frac{1}{2}} \left( \sum_{j=0}^{n-2} |\Delta \widetilde{Y}_{k_l^R, j}||\Delta \widetilde{Y}_{k_l^R, j+1}| \right)^{\frac{1}{2}} \right\} \cap \widehat{\Lambda} \\
\supseteq \ & \liminf_n \left\{ \Delta \check{Y}_{k_l^R, j_l^R} + H_l \geqslant 2 \left( n^{2\gamma-1} \left( \sum_{j=0}^{n-2} |\Delta \check{Y}_{k_l^R, j}||\Delta \check{Y}_{k_l^R, j+1}| \right) \right.\right. \\
& \qquad \left.\left. + H_l n^{2\gamma - 1} (|\Delta \check{Y}_{k_l^R, j_l^R - 1}| + |\Delta \check{Y}_{k_l^R, j_l^R + 1}|) \right)^{\frac{1}{2}} \right\} \cap \widehat{\Lambda} \\
= \ & \widehat{\Lambda}.
\end{aligned} \tag{1.87}$$

Note that the last equality follows from $2\gamma - 1 < 1$ and from the check variants of (1.81) and (1.82). Consider for the last inclusion

$$\begin{aligned}
& |\Delta \widetilde{Y}_{k_l^R, j}||\Delta \widetilde{Y}_{k_l^R, j+1}| \\
\leqslant \ & (|\Delta \check{Y}_{k_l^R, j}| + |\Delta J_{k_l^R, j}|)(|\Delta \check{Y}_{k_l^R, j+1}| + |\Delta J_{k_l^R, j+1}|) \\
= \ & |\Delta \check{Y}_{k_l^R, j}||\Delta \check{Y}_{k_l^R, j+1}| + |\Delta \check{Y}_{k_l^R, j}||\Delta J_{k_l^R, j+1}| + |\Delta J_{k_l^R, j}||\Delta \check{Y}_{k_l^R, j+1}| + |\Delta J_{k_l^R, j}||\Delta J_{k_l^R, j+1}|,
\end{aligned}$$

and observe that we have for $\widehat{\omega} \in \widehat{\Lambda}$ and $n$ large enough

$$|\Delta J_{k_l^R, j_l^R}(\widehat{\omega})| = H_l(\widehat{\omega}) > 0, \quad |\Delta J_{k_l^R, j}(\widehat{\omega})| = 0, \quad j \in \{0, \ldots, n-1\} - \{j_l^R\}$$

because of $R_l(\widehat{\omega}) \uparrow \infty$.

STEP 3. *Generalizations to the Assumptions 1.2.3.* The same stopping techniques as in the proof of Theorem 1.2.5 respectively Corollary 1.2.6 are used, but without the replacement (1.38). Set

$$\widetilde{Y}_t^{(m)} \stackrel{\text{def}}{=} \check{Y}_t^{(m)} + J_t, \quad 0 \leqslant t \leqslant 1, \quad m \in \mathbb{N}.$$

We have proven so far that there exists a negligible set $M_m$ for each $m \in \mathbb{N}$, such that

$$n^{-\gamma} a_{n^2} \left( T_n^{(m)} - b_{n^2} \right)(\omega) \to \infty, \quad n \to \infty, \quad \omega \in \Lambda \cap M_m^c$$

where $T_n^{(m)}$ is in the usual manner a function of $\widetilde{Y}^{(m)}$. Note that in the notation of Theorem 1.2.5

$$\widetilde{Y}_t^{(m)}(\omega) = \widetilde{Y}_t(\omega), \quad 0 \leqslant t \leqslant 1, \quad \omega \in A_m \cap N_m^c, \quad m, n \in \mathbb{N}$$

holds. Set

$$B \overset{\text{def}}{=} \left( \bigcup_{m \in \mathbb{N}} A_m \right) \cap \left( \bigcup_{m \in \mathbb{N}} M_m \cup N_m \right)^c \cap \Lambda.$$

Then we have with the same argumentation as the one in the proof of Corollary 1.2.6 that the convergence

$$n^{-\gamma} a_{n^2}(T_n(\omega) - b_{n^2}) \to \infty, \quad n \to \infty, \quad \omega \in B$$

holds. Finally, the claim of this proposition is proven by setting

$$N \overset{\text{def}}{=} \bigcup_{m \in \mathbb{N}} M_m \cup N_m$$

for the negligible set $N$ in (1.80). □

As discussed above, we have shown that the bipower variation estimator is more suitable than the quadratic variation estimator in the case of external jumps. Another interesting and natural question is whether the bipower variation estimator converges even faster to the true volatility. It will turn out that this is in general not the case. We want to analyse in what follows the exact finite sample behaviour of both volatility estimators and assume for this that the volatility is the constant 1, i.e. $Y = W$ is a Brownian motion. Of course, this is a dramatical simplification, but, nevertheless, this approach also provide some insight in more general volatility processes $\sigma$: For example, assume that $\sigma$ is independent of the Brownian motion $W$ and that every path of $\sigma$ is a regulated function. Then, we can condition on $\sigma$ as described in Section 1.4 and, hence, assume that $\sigma$ is deterministic. Finally, after an approximation of $\sigma$ by deterministic step functions, we can use the results for the case $\sigma = 1$.

We calculate in the following the $L^2$ distance of the respective volatility estimators to the true volatility $\sigma = 1$. To not only have the two singular cases of the bipower and quadratic variation estimators, we define a family

$$\{\widehat{\sigma}_k(n,p) \,:\, 0 \leqslant p \leqslant 1,\, 0 \leqslant k < n,\, n \in \mathbb{N}\}$$

of volatility estimators which interpolate both cases in a natural way. Set for this purpose

$$\widehat{\sigma}_k^2(n,p) \overset{\text{def}}{=} \frac{\sum_{j=0}^{n-1} |\Delta Y_{k,j}|^{1+p} |\Delta Y_{k,j+1}|^{1-p}}{E\left(\sum_{j=0}^{n-1} |\Delta Y_{k,j}|^{1+p} |\Delta Y_{k,j+1}|^{1-p}\right)} \tag{1.88}$$

and note that $(\widehat{\sigma}_k^2(n-1,0))_{0 \leqslant k < n}$ yields the estimator in Definition 1.1.3 and $(\widehat{\sigma}_k^2(n,1))_{0 \leqslant k < n}$ the estimator in Redefinition 1.4.1.

**Remark 1.6.4.** To get a connection between the above discussed goal and the next Proposition 1.6.5, we have concerning the $L^2$ distance the equations (recall $\sigma = 1$)

$$E\left(\widehat{\sigma}_k^2(n-1,0) - 1\right)^2 = E\left( \frac{\pi}{2(n-1)} \sum_{j=0}^{n-2} |n\Delta Y_{k,j}||n\Delta Y_{k,j+1}| - 1 \right)^2 = h(n-1,0) \tag{1.89}$$

and

$$E\left(\widehat{\sigma}_k^2(n,1)-1\right)^2 = E\left(\frac{1}{n}\sum_{j=0}^{n-1}|n\Delta Y_{k,j}|^2-1\right)^2 = h(n,1) \tag{1.90}$$

with $h$ as in (1.91).

**Proposition 1.6.5.** *Let $(Z_j)_{0\leqslant j\leqslant n}$ be a family of i.i.d. standard normal distributed random variables. Set for $-\frac{3}{2}<p<\frac{3}{2}$ and $n\in\mathbb{N}$*

$$X(n,p) \stackrel{\text{def}}{=} \frac{1}{n}\sum_{j=0}^{n-1}|Z_j|^{1+p}|Z_{j+1}|^{1-p}$$

*and*

$$h(n,p) \stackrel{\text{def}}{=} E\left(\frac{X(n,p)}{EX(n,p)}-1\right)^2. \tag{1.91}$$

*Then we obtain*

$$h(n,p) = \frac{f_1(p)}{n}+\frac{f_2(p)}{n^2}, \quad -\frac{3}{2}<p<\frac{3}{2}, \quad n\in\mathbb{N}$$

*where*

$$f_1(p) = \left(\frac{1}{p^2}-4\right)\frac{\sin^2\left(\frac{p\pi}{2}\right)}{\cos(p\pi)}+\frac{2}{p}\sin\left(\frac{p\pi}{2}\right)-3, \quad f_2(p) = 2\left(1-\frac{\sin\left(\frac{p\pi}{2}\right)}{p}\right)$$

*and, in particular,*

$$(f_1(1),f_2(1)) = (2,0), \quad (f_1(0),f_2(0)) = \left(\frac{\pi^2}{4}+\pi-3, 2-\pi\right) \approx (2.609,-1.142).$$

*Proof.* Note first that in the strict sense, $f_1$ and $f_2$ are not defined on the whole interval $\left(-\frac{3}{2},\frac{3}{2}\right)$. However, this causes no problems since we always consider the continuous continuations which exist in this case. The proof is divided into two steps: first, some identities related to the Gamma function are calculated, second, in order to derive the claims of this proposition, these identities are used.

STEP 1. *Identities related to the Gamma function.* Denote with

$$\mu(r) \stackrel{\text{def}}{=} E|\mathrm{N}(0,1)|^r = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}|x|^r e^{-\frac{x^2}{2}}\,dx, \quad r>-1$$

the $r$-th moment of the standard normal distribution. Then we have using the substitution $y=\frac{x^2}{2}$

$$\mu(r) = \frac{2}{\sqrt{2\pi}}\int_0^\infty x^r e^{-\frac{x^2}{2}}\,dx = \frac{2^{\frac{r}{2}}}{\sqrt{\pi}}\int_0^\infty y^{\frac{r+1}{2}-1}e^{-y}\,dy = \frac{2^{\frac{r}{2}}}{\sqrt{\pi}}\Gamma\left(\frac{r+1}{2}\right), \quad r>-1. \tag{1.92}$$

Note that the Gamma function fulfills the Euler reflection identity

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}, \quad z\in\mathbb{C}-\mathbb{Z}, \tag{1.93}$$

cf. Theorem 1.2.1 in Andrews, Askey and Roy [1]. We require for the second step the functions

$$\nu_1(p) \stackrel{\text{def}}{=} \mu(1+p)\mu(1-p), \quad \nu_2(p) \stackrel{\text{def}}{=} \mu(2+2p)\mu(2-2p), \quad p\in\left(-\frac{3}{2},\frac{3}{2}\right).$$

Now we can write using (1.92), (1.93) and the functional equation of the Gamma function

$$\begin{aligned}\nu_1(p) &= \frac{2}{\pi}\Gamma\left(\frac{2+p}{2}\right)\Gamma\left(\frac{2-p}{2}\right) = \frac{2}{\pi}\Gamma\left(\frac{2+p}{2}\right)\Gamma\left(1+\left(1-\frac{2+p}{2}\right)\right) \\ &= \frac{2}{\pi}\left(1-\frac{2+p}{2}\right)\Gamma\left(\frac{2+p}{2}\right)\Gamma\left(1-\frac{2+p}{2}\right) \\ &= \frac{p}{\sin\left(\frac{p\pi}{2}\right)}, \quad p \in \left(-\frac{3}{2},\frac{3}{2}\right)\setminus\{0\} \end{aligned} \tag{1.94}$$

and

$$\begin{aligned}\nu_2(p) &= \frac{4}{\pi}\Gamma\left(\frac{3+2p}{2}\right)\Gamma\left(\frac{3-2p}{2}\right) = \frac{4}{\pi}\Gamma\left(\frac{3+2p}{2}\right)\Gamma\left(2+\left(1-\frac{3+2p}{2}\right)\right) \\ &= \frac{4}{\pi}\left(2-\frac{3+2p}{2}\right)\left(1-\frac{3+2p}{2}\right)\Gamma\left(\frac{3+2p}{2}\right)\Gamma\left(1-\frac{3+2p}{2}\right) \\ &= \frac{1-4p^2}{\cos(p\pi)}, \quad p \in \left(-\frac{3}{2},\frac{3}{2}\right)\setminus\left\{-\frac{1}{2},\frac{1}{2}\right\}. \end{aligned} \tag{1.95}$$

Note that since $\nu_1$, $\nu_2$ are continuous on $\left(-\frac{3}{2},\frac{3}{2}\right)$, the relations (1.94) and (1.95) hold on the whole interval $\left(-\frac{3}{2},\frac{3}{2}\right)$, if we consider as always the continuous continuation of the right-hand functions.

STEP 2. *Calculation of the crucial moments.* Starting with

$$h(n,p) = \frac{EX^2(n,p) - (EX(n,p))^2}{(EX(n,p))^2}, \quad p \in \left(-\frac{3}{2},\frac{3}{2}\right), n \in \mathbb{N} \tag{1.96}$$

the first and second moment of $X(n,p)$ are calculated. This yields

$$EX(n,p) = \frac{1}{n}\sum_{j=0}^{n-1} E\left(|Z_j|^{1+p}|Z_{j+1}|^{1-p}\right) = \nu_1(p).$$

For the calculation of the second moment, the random matrix

$$\left(|Z_j|^{1+p}|Z_{j+1}|^{1-p}|Z_k|^{1+p}|Z_{k+1}|^{1-p}\right)_{0\leqslant j,k<n}$$

is decomposed in its diagonal, two secondary diagonals and the remainder. This yields

$$\begin{aligned} EX^2(n,p) &= \frac{1}{n^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1} E\left(|Z_j|^{1+p}|Z_{j+1}|^{1-p}|Z_k|^{1+p}|Z_{k+1}|^{1-p}\right) \\ &= \frac{1}{n^2}\Big(nE\left(|Z_0|^{2+2p}|Z_1|^{2-2p}\right) + 2(n-1)E\left(|Z_0|^{1+p}|Z_1|^{(1-p)+(1+p)}|Z_2|^{1-p}\right) \\ &\quad +(n^2-n-2(n-1))E\left(|Z_0|^{1+p}|Z_1|^{1-p}|Z_2|^{1+p}|Z_3|^{1-p}\right)\Big) \\ &= \frac{1}{n^2}\left(n\nu_2(p) + 2(n-1)\nu_1(p) + (n^2-n-2(n-1))(\nu_1(p))^2\right). \end{aligned}$$

Thus, we can rewrite (1.96) as

$$h(n,p) = \frac{1}{n}\left(\frac{\nu_2(p)+2\nu_1(p)}{(\nu_1(p))^2} - 3\right) + \frac{2}{n^2}\left(1-\frac{1}{\nu_1(p)}\right), \quad p \in \left(-\frac{3}{2},\frac{3}{2}\right), n \in \mathbb{N}. \tag{1.97}$$

Now, (1.94) and (1.95) are substituted in (1.97) and it is calculated

$$\begin{aligned} f_1(p) &= \frac{\nu_2(p) + 2\nu_1(p)}{(\nu_1(p))^2} - 3 = \frac{1-4p^2}{\cos(p\pi)}\frac{\sin^2\left(\frac{p\pi}{2}\right)}{p^2} + 2\frac{\sin\left(\frac{p\pi}{2}\right)}{p} - 3 \\ &= \left(\frac{1}{p^2} - 4\right)\frac{\sin^2\left(\frac{p\pi}{2}\right)}{\cos(p\pi)} + \frac{2}{p}\sin\left(\frac{p\pi}{2}\right) - 3, \\ f_2(p) &= 2\left(1 - \frac{1}{\nu_1(p)}\right) = 2\left(1 - \frac{\sin\left(\frac{p\pi}{2}\right)}{p}\right). \end{aligned}$$

The remaining claims of this proposition concerning the values $p = 0, 1$ can be proven, for example, by means of L'Hôspital's rule and a numerical calculation. □

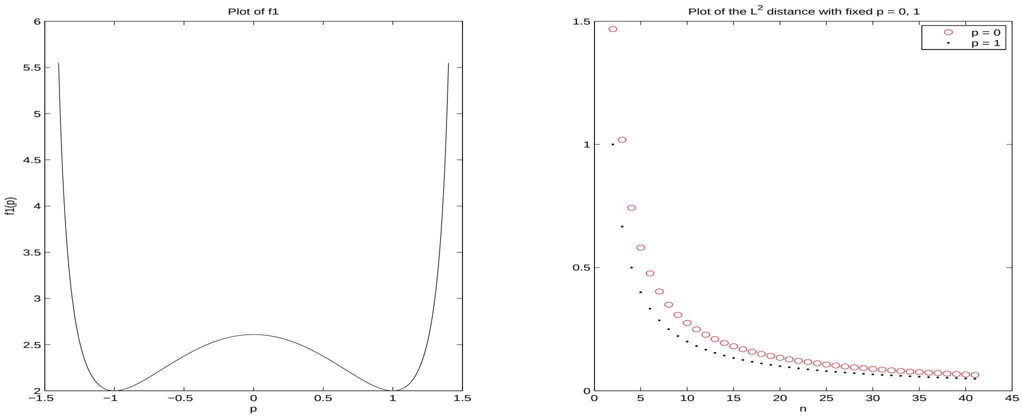


Figure 1.1: Interpolation between the bipower-estimator and the quadratic-estimator

**Remark 1.6.6.** The left-hand side in Figure 1.1 is a plot of the function $p \mapsto f_1(p)$. Note Remark 1.6.4 and the definition of the function $h$ in Proposition 1.6.5 to get the link to the finite sample behaviour of the volatility estimators in (1.88). The quadratic volatility estimator possesses the smallest $L^2$ distance to the true volatility $\sigma = 1$ because of

$$f_1(1) = \min_{-\frac{3}{2}<p<\frac{3}{2}} f_1(p).$$

The right-hand side in Figure 1.1 considers only the cases $p = 0, 1$, i.e. our cases of interest. Here, we have plotted the functions $n \mapsto h(n, 1)$ and $n \mapsto h(n-1, 0)$, compare with (1.89) and (1.90). Note that the quadratic term $\frac{f_2(p)}{n^2}$ in $h(n, p)$ particularly plays an important role for small $n$.

In conclusion, if external jumps are not of concern, the quadratic-volatility estimator is preferable.

## 1.7 Simulation results

In this section, the efficiency of the test statistic (1.32) concerning a jump detection test is investigated using numerical simulations. The resulting test is called the *Gumbel test*. Furthermore, the latter test is compared with a test developed by Barndorff-Nielsen and Shephard in [3], and both tests are applied to a real dataset. For this purpose, some MATLAB functions were written.

First, consider the approximation to the Gumbel distribution in Theorem 1.2.5. For this purpose the drift term is set zero and an Ornstein-Uhlenbeck process is chosen as the volatility process with initial value $a = 1$, mean reversion $\mu = 1$, mean reversion speed $\theta = 0.5$ and diffusion $\overline{\sigma} = 0.2$ which is bounded away from zero, i.e.

$$dZ_t = \theta(\mu - Z_t)\,dt + \overline{\sigma}\,d\overline{W}_t, \quad Z_0 = a, \quad 0 \leqslant t \leqslant 1, \tag{1.98}$$

$$\sigma_t = \max(0.1, Z_t). \tag{1.99}$$

Here, $\overline{W}$ denotes a Brownian motion that is independent of the integrator $W$ in $Y_t = \int_0^t \sigma_s\,dW_s$, i.e. $\sigma$ and $W$ are independent. In Figure 1.2, the grid sizes $\frac{1}{n^2}$, $n = 20, 50, 200, 1000$ are used and the respective empirical distribution functions using 10000 paths for each $n$ are calculated.

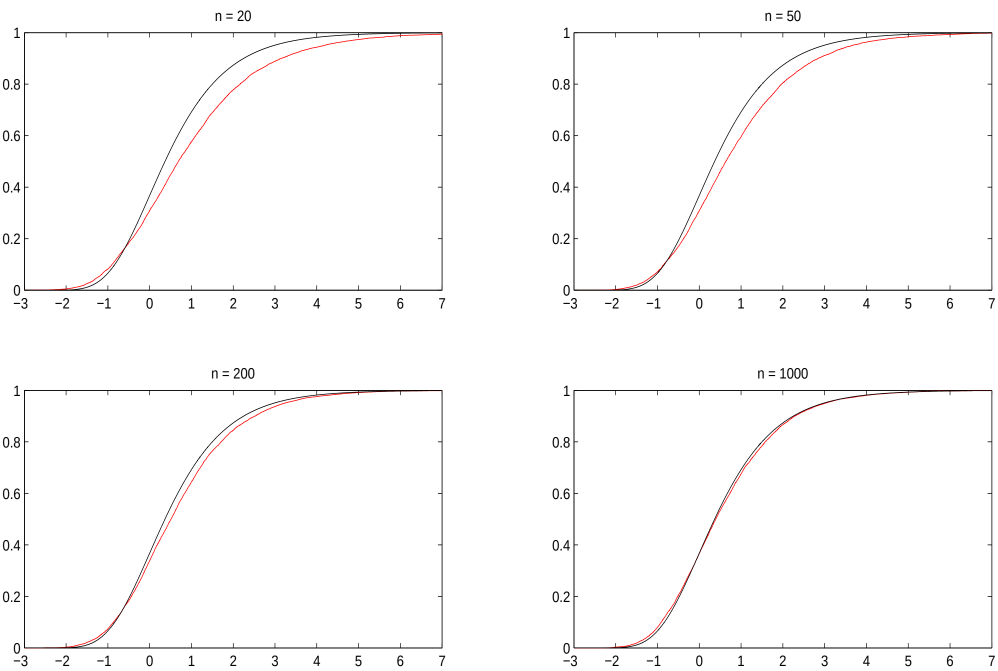


Figure 1.2: Approximation to the Gumbel distribution

Next, set $n = 50$ and consider the additional jump process

$$J_t = \sum_{i=0}^{N_t} U_i, \quad 0 \leqslant t \leqslant 1 \tag{1.100}$$

where $N, U, \sigma, W$ are independent, $N$ is a Poisson process with intensity $\lambda$ and $(U_i)_i$ are i.i.d. $\Gamma$ distributed random variables with shape parameter $k$ and scale parameter $\theta$, i.e. $J$ is a compound Poisson process. Figure 1.3 presents four simulation plots with the respective settings $\lambda = 5$,

$k = 10L^2$ and $\theta = \frac{1}{500L}$, $L = 1, 2, 3, 4$. This results in

$$E(U_i) = \frac{L}{50}, \quad \mathrm{Var}(U_i) = \frac{1}{10 \cdot 50^2}.$$

Note that it holds by the Chebyshev inequality

$$P\left(\left|U_i - \frac{L}{50}\right| \geqslant \frac{1}{50}\right) \leqslant \frac{1}{10}$$

and that, in the case of a constant, deterministic volatility $\sigma = \sigma_0 > 0$ and no external jumps ($J = 0$), the process $Y_t = \sigma_0 W_t$ is obtained. In this case, it follows that

$$E|Y_{\frac{i+1}{n^2}} - Y_{\frac{i}{n^2}}| = \sigma_0 E|W_{\frac{1}{n^2}}| = \frac{\sigma_0\sqrt{2}}{n\sqrt{\pi}} = \frac{\sigma_0\sqrt{2}}{\sqrt{\pi}} \cdot \frac{1}{50}.$$

Hence, the jumps have a critical size in the sense that it is not clear whether they can be detected or not. (recall $a = 1$ in (1.98))

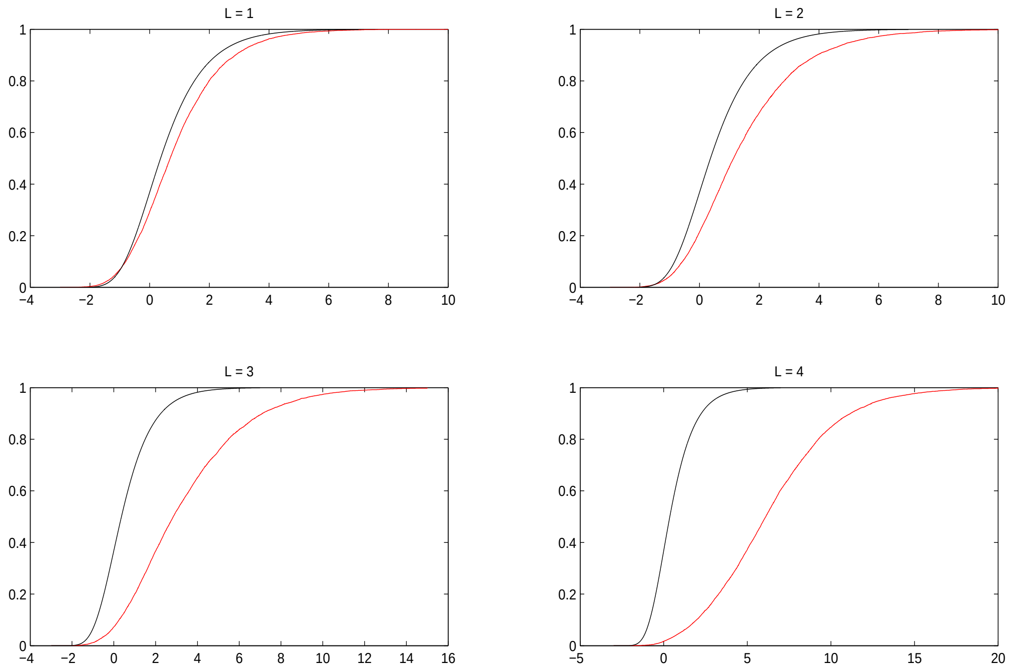


Figure 1.3: Divergence from the Gumbel distribution

Next, the test proposed by Barndorff-Nielsen and Shephard in [3] is compared with the Gumbel test. Barndorff-Nielsen and Shephard use the statistic

$$S_n = \frac{\mu^{-2} Y_n^{[1,1]} - Y_n^{[2]}}{\sqrt{\vartheta \mu^{-4} Y_n^{[1,1,1,1]}}}$$

with

$$\mu = \sqrt{\frac{2}{\pi}}, \quad \vartheta = \frac{\pi^2}{4} + \pi - 5$$

and

$$\begin{aligned} Y_n^{[2]} &= \sum_{i=0}^{n^2-1} |\Delta Y_i|^2, \quad \Delta Y_i = Y_{\frac{i+1}{n^2}} - Y_{\frac{i}{n}}, \\ Y_n^{[1,1]} &= \sum_{i=0}^{n^2-2} |\Delta Y_i||\Delta Y_{i+1}|, \\ Y_n^{[1,1,1,1]} &= \sum_{i=0}^{n^2-4} |\Delta Y_i||\Delta Y_{i+1}||\Delta Y_{i+2}||\Delta Y_{i+3}|. \end{aligned}$$

It holds according to Theorem 1 in [3]

$$S_n \xrightarrow{\mathrm{d}} \mathrm{N}(0,1), \quad n \to \infty$$

under certain conditions on $\sigma$ which are fulfilled by the choice in (1.99). Furthermore, [3] states

$$S_n \to -\infty, \quad n \to \infty \quad P\text{-stoch.},$$

if there is an additional external jump term $J$ as in (1.100). Define the null hypothesis

$$H_0 \ : \text{ there are no jumps}$$

and the alternative hypothesis

$$H_1 \ : \text{ there are jumps (i.e. J in (1.100) is added).}$$

We have the two errors types

$$\begin{aligned} &\text{Type I error} &&: \text{ rejecting a true null hypothesis,} \\ &\text{Type II error} &&: \text{ failing to reject a false null hypothesis} \end{aligned}$$

and decide that a path $\widehat{\omega}$ possesses a jump on the significance level $\alpha \in (0,1)$, iff

$$G(a_{n^2}(T_n(\widehat{\omega}) - b_{n^2})) \geqslant 1 - \alpha \quad \text{resp.} \quad F(S_n(\widehat{\omega})) \leqslant \alpha \tag{1.101}$$

with

$$G(x) = \mathcal{G}((-\infty, x)), \quad F(x) = \mathrm{N}(0,1)((-\infty, x)), \quad x \in \mathbb{R}.$$

Figure 1.4 is based on a setting with the gridsize $\frac{1}{n^2}, n = 50$, the volatility process in (1.99) and $\lambda = 10$, $L = 4$ for the jump process. Here, 2000 sample paths for each $\alpha = \frac{i}{1000}$, $i = 0, 1, \ldots, 100$ were simulated and it is verified that the Gumbel test is more sensitive than the test proposed in [3].

In order to check whether the Gumbel test has a larger power, both tests were recalibrated so that the type I error is exactly $\alpha$ for both tests. Hence $G$ and $F$ are replaced by $G_n$ and $F_n$ in

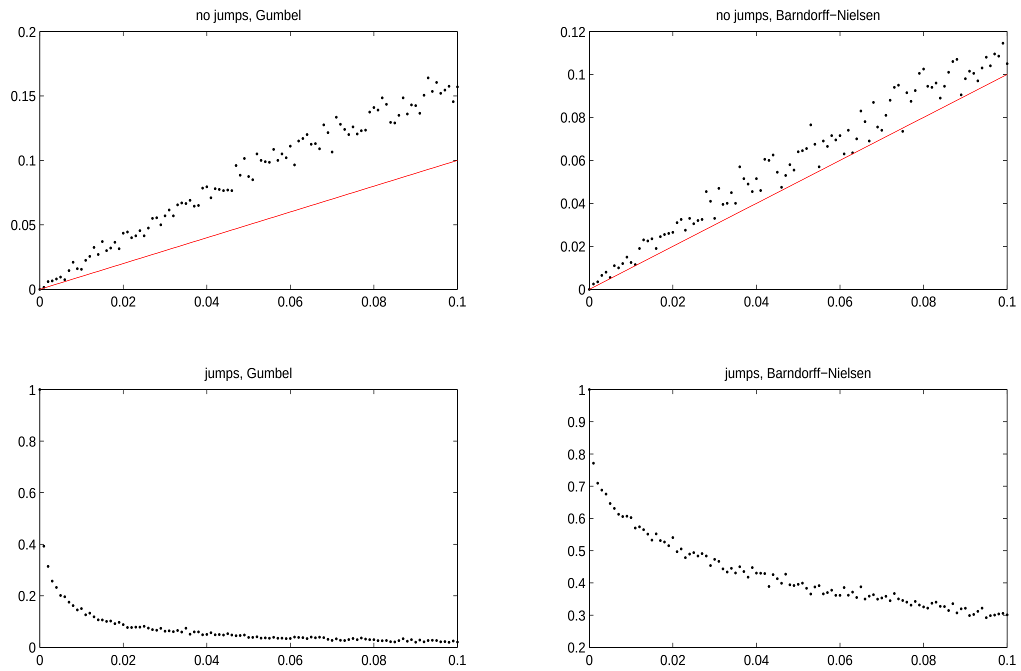


Figure 1.4: Error types

(1.101) with

$$G_n(x) \overset{\text{def}}{=} P(a_{n^2}(T_n - b_{n^2}) \leqslant x) = G_n(x), \quad F_n(x) \overset{\text{def}}{=} P(S_n \leqslant x), \quad x \in \mathbb{R} \quad (n = 50).$$

$G_n$ resp. $F_n$ were approximated with the resp. empiric distribution functions based on 10000 paths. Figure 1.5 shows the type I error calculated with 2000 paths per $\alpha$. Observe the desired approximation to the diagonal.

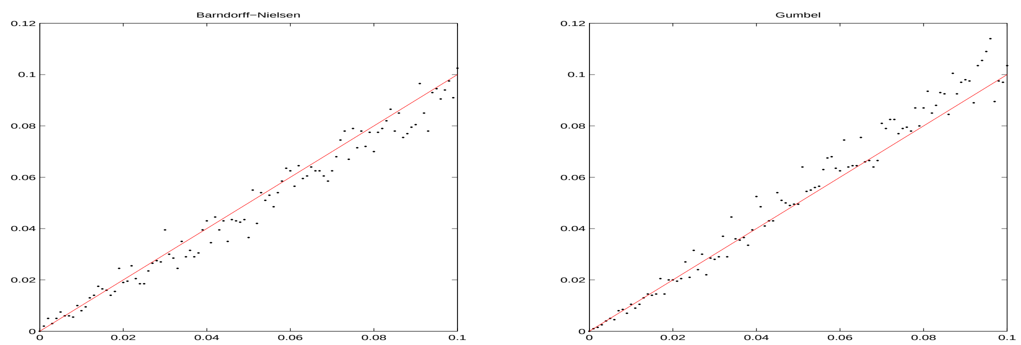


Figure 1.5: Recalibrated type I error

As mentioned above, the power of the tests was calculated. It is defined by

$$\text{power}_\alpha = 1 - \text{type II error}_\alpha$$

with $G_n$ resp. $F_n$. For this purpose, we set $\lambda = 2, 15$ and $L = 4$.

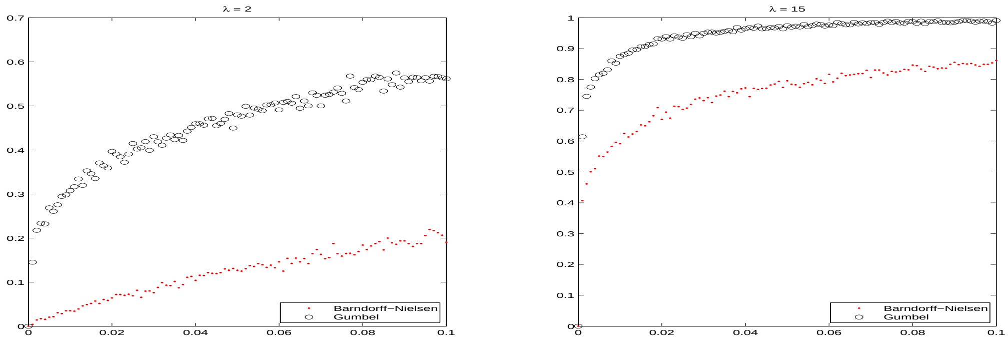


Figure 1.6: Power of tests

In this sense, the Gumbel test clearly has more power than the test proposed by Barndorff-Nielsen and Shephard in [3]. Nevertheless, note that only a special setting (here an OU-process as volatility process and a $\Gamma$ distributed compound Poisson process as external jump process) can be simulated and that the test statistic (1.32) has difficulties with jumps in the volatility process $\sigma$, see Section 1.5.

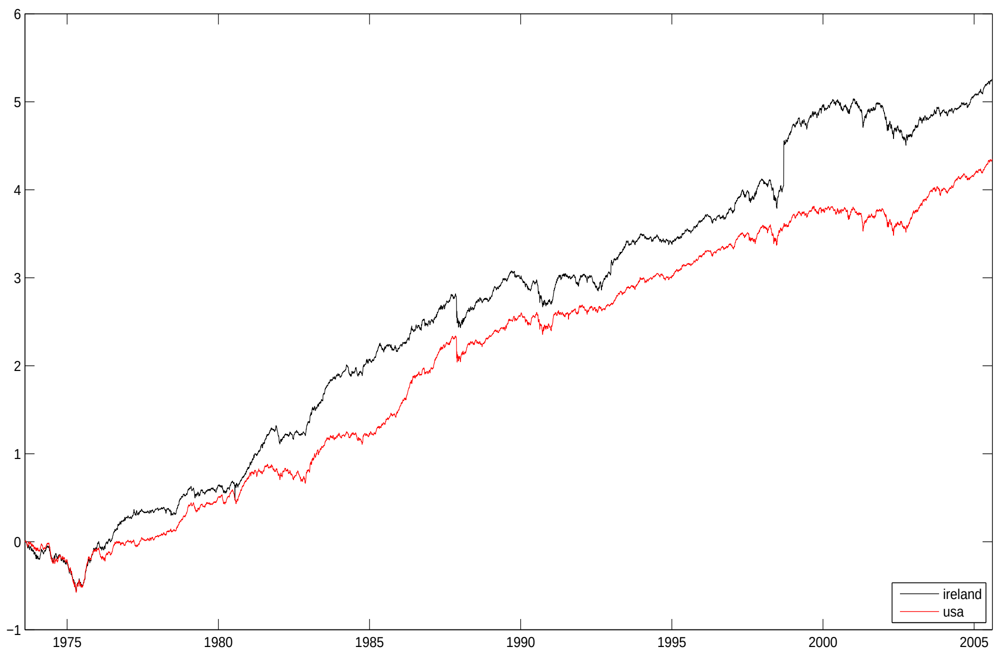


Figure 1.7: Worldstock index

Finally, both tests are applied to a real dataset, i.e. the world stock indices of the USA and Ireland. With

$$\widetilde{T} = \widetilde{T}_n = a_{n^2}(T_n - b_{n^2})$$

the results

$$G(\widetilde{T}^{\text{up}}_{\text{Ireland}}) = 1,\; G(\widetilde{T}^{\text{down}}_{\text{Ireland}}) = 1, \quad G(\widetilde{T}^{\text{up}}_{\text{USA}}) = 0.99268,\; G(\widetilde{T}^{\text{down}}_{\text{USA}}) = 1$$

and

$$F(S_{\text{Ireland}}) = 0, \quad F(S_{\text{USA}}) = 0.81401$$

are achieved. Note that the Gumbel test not only indicates a jump but also states the position and direction of the jump, i.e. whether we have an upwards or downwards jump. In order to detect downwards jumps, we simply have to switch from $Y$ to $-Y$ and the maximum in the Gumbel test statistic turns to a minimum . This, of course, is an enormous advantage over the Barndorff-Nielsen and Shephard test.

The latter results indicate the following: Using the Barndorff-Nielsen and Shephard test the world stock index for Ireland possesses a jump but that of the USA does not. Using the Gumbel test, the index for Ireland possesses an upwards and downwards jump and that for the USA has a downwards jump for sure and an upwards jump with high probability. The different results for the indices of the two countries are not surprising considering that the Gumbel test is more sensitive as discussed above.

We are grateful to Prof. Dr. Eckhard Platen for providing us the world stock index data set. See in this context also the publications [19, 33] by Platen et al.

# Chapter 2

# Lévy processes and dependences

In this chapter, we analyse the dependence structure of jumps in a multidimensional Lévy process. This Lévy process is discretely observed in a *low frequency* scheme. To be more precise, let $(X_t)_{t\geqslant 0}$ be a $d$-dimensional Lévy process on a probability space $(\Omega, \mathcal{F}, P_{\Sigma,\nu,\alpha})$ with the Lévy triplet $(\Sigma, \nu, \alpha)$. Based on the equidistant observations $(X_t(\widehat{\omega}))_{t=0,1,\ldots,n}$ for some fixed path $\widehat{\omega} \in \Omega$, we intend to estimate the dependence structure of the jumps between the coordinates of $X$. Hence, we have a statistical problem. Next, a rigorous formulation of what is meant by this dependence structure is given:

First, we state a well-known concept proposed by Sklar [42]. In this context, see also the monograph of Nelsen [28]. Given $d$ random variables $Y_1, \ldots, Y_d : (\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P}) \to (\mathbb{R}, \mathcal{B})$, a well-known concept to describe the dependence structure within $(Y_1, \ldots, Y_d)$ is provided by its copula $C_{Y_1,\ldots,Y_d}$. This is a $d$ dimensional distribution function with uniform margins, such that we have

$$\widetilde{P}(Y_1 \leqslant y_1, \ldots Y_d \leqslant y_d) = C_{Y_1,\ldots,Y_d}(\widetilde{P}(Y_1 \leqslant y_1), \ldots, \widetilde{P}(Y_d \leqslant y_d)), \quad y_1, \ldots y_d \in \mathbb{R}.$$

Thus, a copula provides in a certain sense the additional information that is needed to obtain the vector distribution from the marginal distributions. However, in this chapter, we deal with a stochastic process

$$X : \mathbb{R}_+ \times \Omega \to \mathbb{R}^d$$

and not with only finite many real valued random variables. Observe the notations

$$\mathbb{R}_+ \stackrel{\text{def}}{=} [0, \infty), \quad \mathbb{R}^* \stackrel{\text{def}}{=} \mathbb{R}\backslash\{0\}, \quad \mathbb{R}_+^* \stackrel{\text{def}}{=} \mathbb{R}_+ \cap \mathbb{R}^*.$$

Generally, it is problematic to determine the meaning of the dependence structure of $X$ or even the dependence structure between the coordinates of the jumps of $X$. Nevertheless, in the case of a *Lévy process* $X$, a natural approach is reported by Kallsen and Tankov [21] which uses the fact that $X$ is characterized by its Lévy triplet $(\Sigma, \nu, \alpha)$. Here, $\nu$ describes the jumps of $X$ in the sense that

$$\nu(A) = E|\{t \in [0, 1] : \Delta X_t \in A\}|, \quad A \in \mathcal{B}(\mathbb{R}^d),$$

cf. Sato [38][Theorem 19.2]. With the jump structure of $X$, we mean the dependence structure of $\nu$. Observe the problem that $\nu$ is in general not a probability measure, so that the copula concept cannot be applied to $\nu$. However, it is at least known that

$$\int_{\mathbb{R}^d} |x|^2 \wedge 1 \, \nu(dx) < \infty$$

holds which implies $\nu(A) < \infty$ for all $A \in \mathcal{B}(\mathbb{R}^d)$ with $0 \notin \overline{A}$. Thus, $0 \in \mathbb{R}^d$ is the only possible singular point. Based on these facts, Kallsen and Tankov [21] introduced the concept of *Lévy-copulas*. We concisely state in what follows the definition of a Lévy copula and the analogous statement to Sklars theorem. We quote for this purpose the respective issues in [21]. See also this paper for a detailed discussion on this topic.

Set $\overline{\mathbb{R}} \overset{\text{def}}{=} \mathbb{R} \cup \{\infty\}$ and

$$\operatorname{sgn}(x) \overset{\text{def}}{=} \begin{cases} 1, & x \geqslant 0, \\ -1, & x < 0, \end{cases} \qquad \mathcal{I}(x) \overset{\text{def}}{=} \begin{cases} (x, \infty), & x \geqslant 0, \\ (-\infty, x], & x < 0, \end{cases} \qquad x \in \mathbb{R}$$

and write

$$(a, b] \overset{\text{def}}{=} (a_1, b_1] \times \ldots \times (a_d, b_d], \quad a, b \in \overline{\mathbb{R}}^d.$$

**Definition 2.1. in [21].** *Let $F : S \to \overline{\mathbb{R}}$ for some subset $S \subseteq \overline{\mathbb{R}}^d$. For $a, b \in S$ with $a \leqslant b$ and $\overline{(a, b]} \subseteq S$, the $F$-volume of $(a, b]$ is defined by*

$$V_F((a, b]) \overset{\text{def}}{=} \sum_{u \in \{a_1, b_1\} \times \ldots \times \{a_d, b_d\}} (-1)^{N(u)} F(u),$$

*where* $N(u) \overset{\text{def}}{=} |\{k : u_k = a_k\}|$.

**Definition 2.2. in [21].** *A function $F : S \to \overline{\mathbb{R}}$ for some subset $S \subseteq \overline{\mathbb{R}}^d$ is called d-increasing if $V_F((a, b]) \geqslant 0$ for all $a, b \in S$ with $a \leqslant b$ and $\overline{(a, b]} \subseteq S$.*

**Definition 2.4. in [21].** *Let $F : \overline{\mathbb{R}}^d \to \overline{\mathbb{R}}$ be a d-increasing function such that $F(u_1, \ldots, u_d) = 0$ if $u_i = 0$ for at least one $i \in \{1, \ldots, d\}$. For any non-empty index set $I \subseteq \{1, \ldots, d\}$, the $I$-margin of $F$ is the function $F^I : \overline{\mathbb{R}}^{|I|} \to \overline{\mathbb{R}}$, defined by*

$$F^I((u_i)_{i \in I}) \overset{\text{def}}{=} \lim_{a \to \infty} \sum_{(u_i)_{i \in I^c} \in \{-a, \infty\}^{|I^c|}} F(u_1, \ldots, u_d) \prod_{i \in I^c} \operatorname{sgn}(u_i),$$

*where* $I^c \overset{\text{def}}{=} \{1, \ldots, d\} \backslash I$.

**Definition 3.1. in [21].** *A function $\mathfrak{C} : \overline{\mathbb{R}}^d \to \overline{\mathbb{R}}$ is called a Lévy copula if*

*(i)* $\mathfrak{C}(u_1, \ldots, u_d) \neq \infty$ *for* $(u_1, \ldots, u_d) \neq (\infty, \ldots, \infty)$,

*(ii)* $\mathfrak{C}(u_1, \ldots, u_d) = 0$ *if* $u_i = 0$ *for at least one* $i \in \{1, \ldots, d\}$,

*(iii)* $\mathfrak{C}$ *is d-increasing,*

*(iv)* $\mathfrak{C}^{\{i\}}(u) = u$ *for every* $i \in \{1, \ldots, d\}$, $u \in \mathbb{R}$.

**Definition 3.3. in [21].** *Let $X$ be a $\mathbb{R}^d$-valued Lévy process with Lévy measure $\nu$. The tail integral of $X$ is the function $U : (\mathbb{R} \backslash \{0\})^d \to \mathbb{R}$ defined by*

$$U(x_1, \ldots, x_d) \overset{\text{def}}{=} \prod_{i=1}^{d} \operatorname{sgn}(x_i) \nu \left( \prod_{j=1}^{d} \mathcal{I}(x_j) \right).$$

**Definition 3.4. in [21].** *Let $X$ be a $\mathbb{R}^d$-valued Lévy process and let $I \subseteq \{1, \ldots, d\}$ be non-empty. The $I$-marginal tail integral $U^I$ of $X$ is the tail integral of the process $X^I \stackrel{\text{def}}{=} (X^i)_{i\in I}$. To simplify notation, we denote one-dimensional margins by $U_i \stackrel{\text{def}}{=} U^{\{i\}}$.*

**Theorem 3.6. in [21].** *(i) Let $X = (X^1, \ldots, X^d)$ be a $\mathbb{R}^d$-valued Lévy process. Then there exists a Lévy copula $\mathfrak{C}$ such that the tail integrals of $X$ satisfy*

$$U^I((x_i)_{i\in I}) = \mathfrak{C}^I((U_i(x_i))_{i\in I}) \tag{2.1}$$

*for any non-empty $I \subseteq \{1, \ldots, d\}$ and any $(x_i)_{i\in I} \in (\mathbb{R}\backslash\{0\})^{|I|}$. The Lévy copula $\mathfrak{C}$ is unique on $\prod_{i=1}^d \overline{\operatorname{Ran} U_i}$.*

*(ii) Let $\mathfrak{C}$ be a $d$-dimensional Lévy copula and $U_i$, $i = 1, \ldots, d$ tail integrals of real-valued Lévy processes. Then there exists a $\mathbb{R}^d$-valued Lévy process $X$ whose components have tail integrals $U_1, \ldots, U_d$ and whose marginal tail integrals satisfy (2.1) for any non-empty $I \subseteq \{1, \ldots, d\}$ and any $(x_i)_{i\in I} \in (\mathbb{R}\backslash\{0\})^{|I|}$. The Lévy measure $\nu$ of $X$ is uniquely determined by $\mathfrak{C}$ and $U_i$, $i = 1, \ldots, d$.*

*Proof.* For the proof, refer also the paper of Kallsen, Tankov [21]. □

Our future assumptions in this chapter are going to ensure that the Lévy copula $\mathfrak{C}$ has always the special shape

$$\mathfrak{C}(u, v) = \begin{cases} U(U_1^{-1}(u), U_2^{-1}(v)), & u, v > 0, \\ 0, & u \leqslant 0 \text{ or } v \leqslant 0 \end{cases} \tag{2.2}$$

where

$$U(x, y) \stackrel{\text{def}}{=} \nu([x, \infty) \times [y, \infty)), \quad U_1(x) \stackrel{\text{def}}{=} \nu([x, \infty) \times \mathbb{R}_+), \quad U_2(y) \stackrel{\text{def}}{=} \nu(\mathbb{R}_+ \times [y, \infty)), \quad x, y \in \mathbb{R}_+.$$

Regarding this issue, also compare the remark in Kallsen, Tankov [21] below Theorem 3.6.

At this point we can specify the goal of this chapter: Our Aim is to construct and investigate an estimator for the Lévy copula of $\nu$ based on low frequency observations. The only existing reference in this context is, to our best knowledge, the unpublished paper of Schicks [39]. This paper, however, only deals with the compound Poisson case and will be discussed later in this thesis in the next chapter. Note that also Bücher, Vetter [8] and Laeven [25] and Krajina, Laeven [24] have published relevant information close to this subject. Nevertheless, all their approaches work with the following high frequency observation scheme:

$$(X_t(\widehat{\omega}))_{t=0,\Delta_n,2\Delta_n,\ldots,n\Delta_n}, \quad \Delta_n \to 0, \quad n\Delta_n \to \infty$$

which results in a completely different analysis than our low frequency observation scheme. Our approach is mostly motivated by Neumann, Reiß [29] and Nickl, Reiß [30] which provide the required low frequency techniques for our needs.

This chapter is divided into three sections. In *Section 1*, we state an estimator $\widehat{\nu}_n$ for the Lévy measure $\nu$, which is based on the low frequency observations $(X_t(\widehat{\omega}))_{t=0,1,\ldots,n}$. Our assumptions in this section imply that the second moment of $(\widehat{\nu}_n)$ and $\nu$ exist, i.e.

$$\int_{\mathbb{R}^d} |x|^2 \, \nu(dx) < \infty, \quad \int_{\mathbb{R}^d} |x|^2 \, \widehat{\nu}_n(dx) < \infty, \quad n \in \mathbb{N}, \quad \omega \in \Omega.$$

We prove in Theorem 2.1.5 the weak convergence of Borel measures

$$|x|^2\widehat{\nu}_n(dx)\overset{w}{\to}|x|^2\nu(dx), \quad n\to\infty.$$

This means that we have $P_{\Sigma,\nu,\alpha}$-a.s. the convergence

$$\int_{\mathbb{R}^d} f(x)|x|^2\,\widehat{\nu}_n(dx)\to\int_{\mathbb{R}^d} f(x)|x|^2\,\nu(dx), \quad n\to\infty$$

for all bounded, continuous functions $f$, i.e. $f\in\mathcal{C}_b(\mathbb{R}^d)$. Our prove works only under the assumption $\Sigma=0$, i.e. with vanishing Brownian motion part. This is due to the fact that it is statistically hard to distinguish between the small jumps of infinite activity and the Brownian motion part. This issue is stated more precisely in Lemma 2.1.1. Note that Neumann and Reiß [29] solve this problem in the one dimensional case by estimating

$$\nu_\sigma(dx)\overset{\text{def}}{=}\sigma^2\delta_0(dx)+x^2\nu(dx), \quad x\in\mathbb{R}, \quad \sigma^2\overset{\text{def}}{=}\Sigma$$

instead of $\nu$. As a result, the exponent of the characteristic function gets the shape

$$\Psi_{\nu_\sigma,\alpha}(u)=iu\alpha+\int_{\mathbb{R}}\frac{e^{iux}-1-iux}{x^2}\,\nu_\sigma(dx), \quad u\in\mathbb{R},$$

assuming that the second moment of $\nu$ is finite, compare Section 4 in [29]. Unfortunately, such a transition from $\nu$ to $\nu_\sigma$ does not work in the multidimensional case $d\geqslant 2$. Nevertheless, we aim to estimate the Lévy copula of $\nu$ under assumptions that do *not* exclude the existence of a Brownian motion part, i.e. $\Sigma\neq 0$. Thus, we have to deal somehow with the small jumps of $X$.

This is described in *Section 2*. Here, everything is developed for the case $d=2$. This is only due to a simpler notation of the anyway high technical approach. We first construct an estimator $\widehat{N}_n$ based on the same $n+1$ equidistant observations as $\widehat{\nu}_n$, such that it holds under certain smoothness and decay conditions on $\nu$

$$\sup_{(a,b)\in\mathfrak{R}}\eta(a,b)\left|\nu([a,\infty)\times[b,\infty))-\widehat{N}_n(a,b)\right|=O_{P_{\Sigma,\nu,\alpha}}\left(\frac{(\log\log n)^2}{\sqrt{\log n}}\right), \quad n\to\infty \tag{2.3}$$

with

$$\mathfrak{R}\overset{\text{def}}{=}[0,\infty)^2\backslash\{(0,0)\}, \quad \eta(a,b)\overset{\text{def}}{=}|(a,b)|^2\wedge|(a,b)|^4.$$

This is proven in Theorem 2.2.11. Note that the right hand side of (2.3) is independent of $(a,b)\in\mathfrak{R}$. If $|(a,b)|\to 0$, $\eta(a,b)\asymp|(a,b)|^4\to 0$ slows down the convergence speed in (2.3). Vice versa $|(a,b)|\to\infty$ implies $\eta(a,b)\asymp|(a,b)|^2\to\infty$ which accelerates the convergence speed. This way of treating the small jumps is sufficient of getting satisfying results concerning the estimation of the Lévy copula. Note that (2.3) also yields estimations for $\nu([a,\infty)\times\mathbb{R}_+)$ resp. $\nu(\mathbb{R}_+\times[b,\infty))$ by setting $a>0$, $b=0$ resp. $a=0$, $b>0$. Our assumptions in this section ensure that the Lévy copula of $\nu$ can be written in the form (2.2). We are capable to estimate $U,U_1,U_2$ with the use of (2.3). Our intention is to create a plug-in estimator for (2.2), i.e. we also need an estimator for $U_k^{-1}$, $k=1,2$. This basically works by building the pseudo inverse of the estimator of $U_k$. At this point, we again have to pay attention to the small jumps. This inverting procedure is performed by Corollary 2.2.13 which is the stochastic counterpart of Proposition 2.2.12. This proposition contains the analysis needed for the inversion operation. Finally, Theorem 2.2.14 states that the resulting plug-in estimator $\widehat{\mathfrak{C}}_n$ uniformly converges on compact sets bounded away from zero with a convergence rate $(\log n)^{-\frac{1}{2}}$, i.e.: It holds for two arbitrary and fixed numbers $0 < a < b < \infty$ the asymptotic

$$\sup_{a \leqslant u,v \leqslant b} |\mathfrak{C}(u,v) - \widehat{\mathfrak{C}}_n(u,v)| = O_{P_{\Sigma,\nu,\alpha}}\left(\frac{(\log\log n)^9}{\sqrt{\log n}}\right). \tag{2.4}$$

The term $(\log\log n)^9$ in (2.4) is not relevant in the sense that we have

$$\frac{(\log\log n)^9}{(\log n)^\epsilon} \to 0, \quad n \to \infty$$

for all $\epsilon > 0$. Note that the convergence in (2.4) holds in a wide, non-pathologic class of Lévy triplets which contains Lévy measures of every Blumenthal Getoor index $0 \leqslant \beta \leqslant 2$, i.e. the test can separate the small jumps from the Brownian motion part in a low frequency setting even in the case $\beta = 2$. This is proven in Corollary 2.2.6. Furthermore, observe that $[a,b] \subseteq [0,\infty)^2$ is bounded away from zero. However, we have to treat the small jumps tending to zero in order to estimate, for example, $U_k^{-1}$, $k = 1, 2$, compare the proof of Theorem 2.2.14. Apart from that, our technical approach would easily yield a similar treatment of $[a,b]^2$ in the case $a \downarrow 0$, $b \uparrow \infty$ as in (2.3). The respective $\eta$ would then, however, depend on $U$, i.e. on $\nu$ which is the unknown estimating entity. Hence, such convergence rates are not statistically feasible and thus, we have not calculated them.

Finally, in *Section 3*, we apply the techniques developed in Section 2 to the compound Poisson process (CPP) case. For simplicity we assume that the intensity $\Lambda = \nu(\mathbb{R}^2)$ is known. We propose an estimator $\widehat{C}_n$ for the copula $C$ of the probability measure $\Lambda^{-1}\nu$, which is based on $n+1$ low frequency observations. For this purpose, we show that everything developed in the previous Section 2 also works in this case. Here, we obtain the better and natural convergence rate $n^{-\frac{1}{2}}$ as expected. Namely, we show in Theorem 2.3.7 that

$$\sup_{a \leqslant u,v \leqslant b} |C(u,v) - \widehat{C}_n(u,v)| = O_{P_{\nu,\alpha}}\left(\frac{(\log n)^{10}}{\sqrt{n}}\right), \quad n \to \infty$$

holds under certain assumptions in the CPP case.

Neumann and Reiß [29][Theorem 4.4] prove in a one dimensional setting that, in the case of a non-vanishing Brownian motion part, a logarithmic convergence rate for estimating $\nu_\sigma$ is optimal. Furthermore, $n^{-\frac{1}{2}}$ is the optimal rate in the CPP case. Hence, the convergence rates of our Lévy copula estimators can be considered to be optimal in the sense that the optimal rates in the one dimensional setting still hold in the multidimensional setting *and* after an inversion operation.

## 2.1 Estimating the Lévy measure

Let $(X_t)_{t \geqslant 0}$ be a $d$-dimensional Lévy process on a probability space $(\Omega, \mathcal{F}, P_{\Sigma,\nu,\alpha})$ with the Lévy triplet $(\Sigma, \nu, \alpha)$. Based on the equidistant observations $(X_t(\widehat{\omega}))_{t=0,1,\ldots,n}$ for some fixed path $\widehat{\omega} \in \Omega$, we intend to estimate the Lévy triplet $(\Sigma, \nu, \alpha)$. First of all note that it is statistical not possible to distinguish between the existence of a Brownian motion part and an accumulation of infinitely many jumps in a uniform consistent way. This is explained by the following lemma which is a generalization of Remark 3.2 in Neumann, Reiß [29].

**Lemma 2.1.1.** *Set $d = 1$ and write $\sigma = \Sigma$. Then we have*

$$\sup_{\sigma,\nu,\alpha} P_{\sigma,\nu,\alpha}\left(|\widehat{\sigma} - \sigma| \geqslant \frac{1}{2}\right) \geqslant \frac{1}{2}$$

*where $\widehat{\sigma}$ is any real valued random variable. For example $\widehat{\sigma}$ can be any estimator based on the above low frequency observations.*

*Proof.* Denote with $P_m$, $m \geqslant 1$ the $2 - \frac{1}{m}$ symmetric stable law, i.e. the law with the characteristic function

$$\varphi_m(u) = e^{-\frac{|u|^{2-\frac{1}{m}}}{2}}.$$

As $\varphi_m$ is Lebesgue integrable, $P_m$ has the Lebesgue density

$$f_m(x) = \frac{1}{2\pi} \int e^{-iux} \varphi_m(u)\, du.$$

Now consider the total variation (TV) between $P_m$ and $P_\infty \overset{\text{def}}{=} \mathrm{N}(0,1)$. Scheffé's Lemma yields

$$\|P_m - P_\infty\|_{TV} = \frac{1}{2} \int |f_m(x) - f_\infty(x)|\, dx. \tag{2.5}$$

We have $f_m \to f_\infty$ pointwise because of

$$|f_m(x) - f_\infty(x)| \leqslant \int |\varphi_m(u) - \varphi_\infty(u)|\, du$$

and the integrable $L^1(\lambda^1)$ majorant

$$|\varphi_m(u) - \varphi_\infty(u)| \leqslant 2e^{-\frac{u^2 \wedge |u|}{2}}, \quad m \in \mathbb{N}.$$

This implies together with

$$\int f_m(x)\, dx = \int f_\infty(x)\, dx = 1$$

and a theorem of Riesz, cf. [5][Theorem 15.4] that $f_m \to f_\infty$ in $L^1(\lambda)$, i.e. with (2.5)

$$\|P_m - P_\infty\|_{TV} \to 0, \quad m \to \infty.$$

Now consider the two sets

$$A_0 \overset{\text{def}}{=} \left\{|\widehat{\sigma}| \geqslant \frac{1}{2}\right\}, \quad A_1 \overset{\text{def}}{=} \left\{|\widehat{\sigma} - 1| \geqslant \frac{1}{2}\right\}.$$

Fix $\epsilon > 0$ and choose $m$ large enough, such that $\|P_m - P_\infty\|_{TV} < \epsilon$. Note that the Brownian part $\sigma$ of $P_m$ is zero and the Brownian part of $P_\infty$ is one. Assume that

$$P_m\left(|\widehat{\sigma} - \sigma| \geqslant \frac{1}{2}\right) = P_m(A_0) < \frac{1}{2}.$$

Then we have $P_\infty(A_0) < \frac{1}{2} + \epsilon$ which yields because of $A_0 \cup A_1 = \Omega$

$$P_\infty(A_1) + \frac{1}{2} + \epsilon > P_\infty(A_0) + P_\infty(A_1) \geqslant P_\infty(A_0 \cup A_1) = 1.$$

This results in

$$P_\infty\left(|\widehat{\sigma}-\sigma| \geqslant \frac{1}{2}\right) = P_\infty(A_1) \geqslant \frac{1}{2}-\epsilon.$$

The lemma is proven since $\epsilon > 0$ was chosen arbitrarily. □

We denote by

$$\widehat{\varphi}_n(u) \stackrel{\text{def}}{=} \frac{1}{n}\sum_{t=1}^{n} e^{i\langle u, X_t - X_{t-1}\rangle}, \quad u \in \mathbb{R}^d$$

the empirical characteristic function of the increments and write, furthermore,

$$\varphi_{\Sigma,\nu,\alpha}(u) \stackrel{\text{def}}{=} E_{\Sigma,\nu,\alpha}\left(e^{i\langle u, X_1\rangle}\right), \quad u \in \mathbb{R}^d.$$

Next let $w : \mathbb{R}^d \to \mathbb{R}^{>0}$ denote a weight function which is specified later. For the following we only require that $w$ is bounded and vanishes at infinity. Define the weighted supremum

$$\|\psi\|_{L^\infty(w)} \stackrel{\text{def}}{=} \sup_{u\in\mathbb{R}^d}\{w(u)|\psi(u)|\}$$

for mappings $\psi : \mathbb{R}^d \to \mathbb{C}$. The following proposition is needed in order to prove Theorem 2.1.5. This theorem yields a consistent estimator $\widehat{\nu}_n$ for the Lévy measure $\nu$ as described in the introduction of this chapter.

**Proposition 2.1.2.** *For every $(\Sigma, \nu, \alpha)$ with $E_{\Sigma,\nu,\alpha}|X_1|^4 < \infty$, there exists a $P_{\Sigma,\nu,\alpha}$ negligible set $N$, such that*

$$\left\|\frac{\partial^l}{\partial u_k^l}\left(\widehat{\varphi}_n(u) - \varphi_{\Sigma,\nu,\alpha}(u)\right)\right\|_{L^\infty(w)} \to 0, \quad n \to \infty, \quad l = 0,1,2, \quad 1 \leqslant k \leqslant d \tag{2.6}$$

*holds on $N^c$.*

*Proof.* The proof is divided into two steps. We have to address the problem that the negligible set $N$ has to be independent of $u \in \mathbb{R}^d$ in (2.6). $\mathbb{R}^d$ is uncountable and unbounded. Step 1 yields two continuity inequalities, (2.7) and (2.8). These inequalities enable us to replace $\mathbb{R}^d$ in the second step by certain countable sets. The use of a weight function $w$ solves the problem that $\mathbb{R}^d$ is unbounded. Finally, in step 2, all necessary analyses are performed to complete the proof.

We set $k = 1$, $l = 2$ to simplify the notation. Then (2.6) claims that

$$\sup_{u\in\mathbb{R}^d}\left\{w(u)\left|\frac{1}{n}\sum_{t=1}^{n}(X_{t,1}(\omega) - X_{t-1,1}(\omega))^2 e^{i\langle u, X_t(\omega)-X_{t-1}(\omega)\rangle} - E_{\Sigma,\nu,\alpha}X_{1,1}^2 e^{i\langle u, X_1\rangle}\right|\right\} \to 0, \quad n \to \infty$$

holds for all $\omega \in N^c$ where $X_{t,1} \in \mathbb{R}$ denotes the first component of $X_t \in \mathbb{R}^d$.

STEP 1. We show that there exists a constant $C > 0$, such that we have for all $u, v \in \mathbb{R}^d$ the inequalities

$$\left|\frac{\partial^2\varphi_{\Sigma,\nu,\alpha}(u)}{\partial u_1^2} - \frac{\partial^2\varphi_{\Sigma,\nu,\alpha}(v)}{\partial v_1^2}\right| \leqslant C|u-v|(1+|u|+|v|), \tag{2.7}$$

$$\left|\frac{\partial^2\widehat{\varphi}_n(u)}{\partial u_1^2} - \frac{\partial^2\widehat{\varphi}_n(v)}{\partial v_1^2}\right| \leqslant (C+Y_n)|u-v|(1+|u|+|v|), \quad n \in \mathbb{N} \tag{2.8}$$

for some random variables $(Y_n)_n$ with

$$Y_n \to 0, \quad n \to \infty \quad (P_{\Sigma,\nu,\alpha}\text{-a.s.}).$$

Consider with $u, v \in \mathbb{R}^d$

$$\left|\frac{\partial^2 \widehat{\varphi}_n(u)}{\partial u_1^2} - \frac{\partial^2 \widehat{\varphi}_n(v)}{\partial v_1^2}\right| \leqslant \frac{1}{n}\sum_{t=1}^n (X_{t,1} - X_{t-1,1})^2 \left|e^{i\langle u, X_t - X_{t-1}\rangle} - e^{i\langle v, X_t - X_{t-1}\rangle}\right|.$$

Using the identity

$$e^{ix} = 1 + ix - \int_0^x (x-y)e^{iy}\,dy, \quad x \in \mathbb{R},$$

cf. Sato [38][Lemma 8.6], we can write

$$\begin{aligned}
& e^{i\langle u, X_t - X_{t-1}\rangle} - e^{i\langle v, X_t - X_{t-1}\rangle} \\
= \;& i\,\langle u - v, X_t - X_{t-1}\rangle - \int_0^{\langle u, X_t - X_{t-1}\rangle} (\langle u, X_t - X_{t-1}\rangle - y)e^{iy}\,dy \\
& + \int_0^{\langle v, X_t - X_{t-1}\rangle} (\langle v, X_t - X_{t-1}\rangle - y)e^{iy}\,dy \\
= \;& i\,\langle u - v, X_t - X_{t-1}\rangle + \int_0^{\langle v, X_t - X_{t-1}\rangle} \langle v - u, X_t - X_{t-1}\rangle\, e^{iy}\,dy \\
& - \int_{\langle v, X_t - X_{t-1}\rangle}^{\langle u, X_t - X_{t-1}\rangle} (\langle u, X_t - X_{t-1}\rangle - y)e^{iy}\,dy.
\end{aligned}$$

This yields together with the Cauchy-Schwarz inequality

$$\begin{aligned}
& |e^{i\langle u, X_t - X_{t-1}\rangle} - e^{i\langle v, X_t - X_{t-1}\rangle}| \\
\leqslant \;& |\,\langle u - v, X_t - X_{t-1}\rangle\,| + |\,\langle u - v, X_t - X_{t-1}\rangle\,||\,\langle v, X_t - X_{t-1}\rangle\,| \\
& + |\,\langle u - v, X_t - X_{t-1}\rangle\,|(|\,\langle u, X_t - X_{t-1}\rangle\,| + |\,\langle v, X_t - X_{t-1}\rangle\,|) \\
\leqslant \;& |u - v|[|X_t - X_{t-1}| + |v||X_t - X_{t-1}|^2 + |X_t - X_{t-1}|^2(|u| + |v|)] \\
= \;& |u - v||X_t - X_{t-1}| + |u - v|(|u| + 2|v|)|X_t - X_{t-1}|^2.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
& \left|\frac{\partial^2 \widehat{\varphi}_n(u)}{\partial u_1^2} - \frac{\partial^2 \widehat{\varphi}_n(v)}{\partial v_1^2}\right| \\
\leqslant \;& |u - v|\frac{1}{n}\sum_{t=1}^n (X_{t,1} - X_{t-1,1})^2 |X_t - X_{t-1}| \\
& + |u - v|(|u| + 2|v|)\frac{1}{n}\sum_{t=1}^n (X_{t,1} - X_{t-1,1})^2 |X_t - X_{t-1}|^2
\end{aligned}$$

and the strong law of large numbers yields (2.8) which in turn implies (2.7). Set $N_1 \overset{\text{def}}{=} \{Y_n \nrightarrow 0\}$.

STEP 2. We split the supremum in (2.6) as follows

$$\sup_{u \in \mathbb{R}^d} w(u)\left|\frac{\partial^2}{\partial u_1^2}(\widehat{\varphi}_n(u) - \varphi_{\Sigma,\nu,\alpha}(u))\right| \leqslant \sup_{u \in K_n} w(u)\left|\frac{\partial^2}{\partial u_1^2}(\widehat{\varphi}_n(u) - \varphi_{\Sigma,\nu,\alpha}(u))\right| \tag{2.9}$$

$$+ \sup_{u \in K_n^c} w(u) \left| \frac{\partial^2}{\partial u_1^2} (\widehat{\varphi}_n(u) - \varphi_{\Sigma,\nu,\alpha}(u)) \right| \tag{2.10}$$

where $K_n \subseteq \mathbb{R}^d$ may depend on $\omega$, i.e. $K_n = K_n(\omega)$. We choose $K_n$ in such a manner that

$$R_n \stackrel{\text{def}}{=} \inf\{|u| \,:\, u \in K_n^c\} \to \infty, \quad n \to \infty$$

for all $\omega \in \Omega$. Due to

$$\left| \frac{\partial^2}{\partial u_1^2} (\widehat{\varphi}_n(u) - \varphi_{\Sigma,\nu,\alpha}(u)) \right| \leqslant \frac{1}{n} \sum_{t=1}^{n} (X_{t,1} - X_{t-1,1})^2 + E_{\Sigma,\nu,\alpha} X_{1,1}^2,$$

the strong law of large numbers and

$$R_n(\omega) \to \infty, \quad n \to \infty, \quad \omega \in \Omega$$

together with

$$w(u) \to 0, \quad |u| \to \infty,$$

(2.10) converges $P_{\Sigma,\nu,\alpha}$-a.s. to zero, i.e. for all $\omega \in N_2^c$ for some negligible set $N_2$.

Next, we estimate (2.9). Set

$$\overline{K}_m \stackrel{\text{def}}{=} [-m,m]^d, \quad \Delta_m \stackrel{\text{def}}{=} \frac{1}{m^2}, \quad m \in \mathbb{N}.$$

Define furthermore

$$Q_m \stackrel{\text{def}}{=} \left\{ \frac{z}{m^2} \,:\, z \in \mathbb{Z} \right\}^d, \quad \overline{K}_m^Q \stackrel{\text{def}}{=} \overline{K}_m \cap Q_m.$$

Then we have for any $x \in \overline{K}_m$

$$\inf\{|x-y| \,:\, y \in \overline{K}_m^Q\} \leqslant \left( \left(\frac{1}{m^2}\right)^2 + \ldots + \left(\frac{1}{m^2}\right)^2 \right)^{\frac{1}{2}} = \frac{\sqrt{d}}{m^2}.$$

As $\bigcup_{m=1}^{\infty} \overline{K}_m^Q \subseteq \mathbb{R}^d$ is countable, there exists because of the strong law of large numbers a negligible set $N_3$, such that

$$\frac{\partial^2}{\partial u_1^2} (\widehat{\varphi}_n(u) - \varphi_{\Sigma,\nu,\alpha}(u)) \to 0, \quad n \to \infty, \quad \omega \in N_3^c, \quad u \in \bigcup_{m=1}^{\infty} \overline{K}_m^Q.$$

Now, (2.7) and (2.8) yield

$$\begin{aligned} & \sup_{u \in \overline{K}_m} w(u) \left| \frac{\partial^2}{\partial u_1^2} (\widehat{\varphi}_n(u) - \varphi_{\Sigma,\nu,\alpha}(u)) \right| \qquad (2.11) \\ \leqslant \; & \|w\|_\infty \left( \max_{u \in \overline{K}_m^Q} \left| \frac{\partial^2}{\partial u_1^2} (\widehat{\varphi}_n(u) - \varphi_{\Sigma,\nu,\alpha}(u)) \right| + 2(C + |Y_n|) \frac{\sqrt{d}}{m^2} (1 + 2\sqrt{d} m) \right). \end{aligned}$$

We obtain

$$\max_{u \in \overline{K}_m^Q} \left| \frac{\partial^2}{\partial u_1^2} (\widehat{\varphi}_n(u) - \varphi_{\Sigma,\nu,\alpha}(u)) \right| (\omega) \leqslant \frac{1}{m}, \quad \omega \in N_3^c \tag{2.12}$$

for all $n \geqslant N(m,\omega)$. We assume w.l.o.g. that $N(\cdot,\omega) \uparrow \infty$ for every $\omega \in N_3^c$. Set

$$N^{-1}(n,\omega) \stackrel{\text{def}}{=} \sup\{m \in \mathbb{N} : N(m,\omega) \leqslant n\} \uparrow \infty, \quad n \to \infty, \quad \sup \varnothing \stackrel{\text{def}}{=} 1, \quad \omega \in N_3^c$$

and

$$K_n = K_n(\omega) \stackrel{\text{def}}{=} \begin{cases} \overline{K}_{N^{-1}(n,\omega)}, & \omega \in N_3^c \\ \overline{K}_n, & \omega \in N_3. \end{cases} \tag{2.13}$$

Now, under consideration of (2.11) and (2.12), we obviously obtain

$$\max_{u \in K_n} w(u) \left| \frac{\partial^2}{\partial u_1^2} (\widehat{\varphi}_n(u) - \varphi_{\Sigma,\nu,\alpha}(u)) \right| \to 0, \quad n \to \infty, \quad \omega \in (N_1 \cup N_2 \cup N_3)^c,$$

ensuring together with (2.10) that everything is proven. □

**Remark 2.1.3.** Note that this proof is an illustrative straightforward proof in contrast to the proof of Theorem 2.1.6 which is much more involved. With the latter, we prove a similar statement using a deep result from empirical process theory.

**Assumptions 2.1.4.** *The Lévy triplet $(\Sigma,\nu,\alpha)$ has a vanishing Brownian part and possesses a finite fourth moment, i.e.*

$$\int |x|^4 \nu(dx) < \infty, \quad \Sigma = 0.$$

The above Assumptions 2.1.4 are motivated by Lemma 2.1.1 and Proposition 2.1.2. Note that given the Assumptions 2.1.4 we do not need a truncation function in the representation of the characteristic function of the Lévy process, i.e. we are going to use the representation

$$\varphi_{\nu,\alpha}(u) = \exp\left( i\langle u,\alpha\rangle + \int_{\mathbb{R}^d} (e^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle)\,\nu(dx) \right). \tag{2.14}$$

Define the following metric on $\mathcal{C}^2(\mathbb{R}^d)$ :

$$d^{(2)}(\varphi_1,\varphi_2) \stackrel{\text{def}}{=} \|\varphi_1 - \varphi_2\|_{L^\infty(w)} + \sum_{k=1}^d \left\| \frac{\partial}{\partial u_k} (\varphi_1 - \varphi_2) \right\|_{L^\infty(w)} + \sum_{k=1}^d \left\| \frac{\partial^2}{\partial u_k^2} (\varphi_1 - \varphi_2) \right\|_{L^\infty(w)}.$$

Let Assumptions 2.1.4 hold. Given the equidistant observations as described at the beginning of this section, we introduce the minimum distance estimator $(\widehat{\nu}_n, \widehat{\alpha}_n)$ via

$$d^{(2)}(\widehat{\varphi}_n, \varphi_{\widehat{\nu}_n,\widehat{\alpha}_n}) \leqslant \inf_{\nu,\alpha} d^{(2)}(\widehat{\varphi}_n, \varphi_{\nu,\alpha}) + \delta_n, \quad n \in \mathbb{N}, \quad \omega \in \Omega \tag{2.15}$$

for a given sequence $\delta_n \downarrow 0$. This means that $(\widehat{\nu}_n, \widehat{\alpha}_n)$ are chosen in such a way that (2.15) and the Assumptions 2.1.4 are fulfilled. This is exactly the multidimensional variant of (2.3) in Neumann, Reiss [29].

**Theorem 2.1.5.** *Let Assumptions 2.1.4 hold and let $w$ be continuous and vanishing at infinity. Then we have $P_{\nu,\alpha}$-a.s. the weak convergence*

$$|x|^2 \widehat{\nu}_n(dx) \stackrel{w}{\to} |x|^2 \nu(dx).$$

*Proof.* Proposition 2.1.2 yields

$$\begin{aligned} d^{(2)}(\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n},\varphi_{\nu,\alpha}) &\leqslant d^{(2)}(\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n},\widehat{\varphi}_n)+d^{(2)}(\widehat{\varphi}_n,\varphi_{\nu,\alpha}) \\ &\leqslant 2d^{(2)}(\widehat{\varphi}_n,\varphi_{\nu,\alpha})+\delta_n \\ &\to 0, \quad P_{\nu,\alpha}\text{-a.s.} \end{aligned} \tag{2.16}$$

It follows for any compact $K \subseteq \mathbb{R}^d$

$$\int_K |\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n}(u)-\varphi_{\nu,\alpha}(u)|\,\lambda^d(du) \leqslant \left(\inf_{u\in K} w(u)\right)^{-1} \lambda^d(K)\|\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n}-\varphi_{\nu,\alpha}\|_{L^\infty(w)}, \quad \omega\in\Omega. \tag{2.17}$$

Note that

$$\|\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n}-\varphi_{\nu,\alpha}\|_{L^\infty(w)} \leqslant d^{(2)}(\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n},\varphi_{\nu,\alpha}) \to 0, \quad P_{\nu,\alpha}\text{-a.s.}$$

Fix any $v=(v_1,\dots,v_d)\in\mathbb{R}^d$ and define

$$K_n = K_n(\omega) \overset{\text{def}}{=} [v_1, v_1+\epsilon_n]\times\ldots\times[v_d,v_d+\epsilon_n] \subseteq \mathbb{R}^d, \quad n\in\mathbb{N}$$

for a sequence $(\epsilon_n)_n$ with $\epsilon_n \downarrow 0$, which may depend on $\omega\in\Omega$. Then we have because of the strict positivity and continuity of $w$

$$C \overset{\text{def}}{=} \sup_{n\in\mathbb{N}} \left(\inf_{u\in K_n} w(u)\right)^{-1} = \left(\inf_{u\in K_1} w(u)\right)^{-1} < \infty.$$

From (2.17), it follows that

$$\frac{1}{\epsilon_n^d}\int_{K_n} |\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n}(u)-\varphi_{\nu,\alpha}(u)|\,\lambda^d(du) \leqslant C\|\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n}-\varphi_{\nu,\alpha}\|_{L^\infty(w)} \to 0, \quad P_{\nu,\alpha}\text{-a.s.} \tag{2.18}$$

As $u\mapsto |\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n}(u)-\varphi_{\nu,\alpha}(u)|$ is continuous, we can choose $\epsilon_n=\epsilon_n(\omega)$ small enough that

$$\left|\frac{1}{\epsilon_n^d}\int_{K_n} |\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n}(u)-\varphi_{\nu,\alpha}(u)|\,\lambda^d(du) - |\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n}(v)-\varphi_{\nu,\alpha}(v)|\right| \leqslant \frac{1}{n}, \quad \omega\in\Omega.$$

This yields together with (2.18)

$$\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n}(v) \to \varphi_{\nu,\alpha}(v) \quad P_{\nu,\alpha}\text{-a.s.}$$

and the negligible set is independent of $v\in\mathbb{R}^d$. As a result, Lévy-Cramér yields that we have $P_{\nu,\alpha}$-a.s. the weak convergence

$$P^{X_1}_{\widehat{\nu}_n,\widehat{\alpha}_n} \overset{w}{\to} P^{X_1}_{\nu,\alpha}.$$

With (2.14), we have

$$\begin{aligned} \int_{\mathbb{R}^d} |x|^2\,\widehat{\nu}_n(dx) &= \sum_{k=1}^d \left(\frac{\partial\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n}}{\partial u_k}(0)\right)^2 - \frac{\partial^2\varphi_{\widehat{\nu}_n,\widehat{\alpha}_n}}{\partial u_k^2}(0) \\ &\to \sum_{k=1}^d \left(\frac{\partial\varphi_{\nu,\alpha}}{\partial u_k}(0)\right)^2 - \frac{\partial^2\varphi_{\nu,\alpha}}{\partial u_k^2}(0) = \int_{\mathbb{R}^d} |x|^2\,\nu(dx), \quad P_{\nu,\alpha}\text{-a.s.} \end{aligned}$$

because of (2.16). It therefore suffices to prove the vague convergence, cf. Chung [10].

$$|x|^2\widehat{\nu}_n(dx) \xrightarrow{v} |x|^2\nu(dx), \quad P_{\nu,\alpha}\text{-a.s.}$$

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a continuous function with compact support. It remains to verify

$$\int_{\mathbb{R}^d} f(x)|x|^2\widehat{\nu}_n(dx) \to \int_{\mathbb{R}^d} f(x)|x|^2\,\nu(dx), \quad P_{\nu,\alpha}\text{-a.s.} \tag{2.19}$$

For this purpose set

$$h_\epsilon(r) \stackrel{\text{def}}{=} \begin{cases} 0, & 0 \leqslant r < \frac{\epsilon}{2} \\ \frac{2}{\epsilon}\left(r - \frac{\epsilon}{2}\right), & \frac{\epsilon}{2} \leqslant r < \epsilon \\ 1, & r \geqslant \epsilon \end{cases}, \quad r \geqslant 0$$

and

$$g_\epsilon(x) \stackrel{\text{def}}{=} h_\epsilon(|x|), \quad x \in \mathbb{R}^d.$$

Fix any $\delta > 0$ and choose $\epsilon = \epsilon(\omega) > 0$ small enough that

$$\|f\|_\infty \sup_{n\in\mathbb{N}} \int_{\{|x|\leqslant\epsilon\}} |x|^2\,\widehat{\nu}_n(dx) < \delta, \quad \|f\|_\infty \int_{\{|x|\leqslant\epsilon\}} |x|^2\,\nu(dx) < \delta, \quad P_{\nu,\alpha}\text{-a.s.}$$

This is possible because of Sato [38] Theorem 8.7.(2) and our assumption $\Sigma = \widehat{\Sigma}_n = 0$ on $\Omega$. Based on this, we obtain

$$\int_{\mathbb{R}^d} f(x)(1-g_\epsilon(x))|x|^2\widehat{\nu}_n(dx) < \delta, \quad \int_{\mathbb{R}^d} f(x)(1-g_\epsilon(x))|x|^2\,\nu(dx) < \delta, \quad P_{\nu,\alpha}\text{-a.s.} \tag{2.20}$$

On the other hand, $x \mapsto f(x)|x|^2 g_\epsilon(x)$, $x \in \mathbb{R}^d$ is a bounded, continuous function which vanishes on a neighborhood of zero. Thus, Sato [38][Theorem 8.7.(1)] yields

$$\int_{\mathbb{R}^d} f(x)g_\epsilon(x)|x|^2\widehat{\nu}_n(dx) \to \int_{\mathbb{R}^d} f(x)g_\epsilon(x)|x|^2\,\nu(dx), \quad P_{\nu,\alpha}\text{-a.s.}$$

Hence, together with (2.20), we obtain

$$\limsup_n \left| \int_{\mathbb{R}^d} f(x)|x|^2\widehat{\nu}_n(dx) - \int_{\mathbb{R}^d} f(x)|x|^2\,\nu(dx) \right| \leqslant 2\delta, \quad P_{\nu,\alpha}\text{-a.s.}$$

This proves (2.19) because $\delta > 0$ was chosen arbitrarily. □

Set $Z_t = X_t - X_{t-1}$, $t \in \mathbb{N}$ and define

$$A_n(u) \stackrel{\text{def}}{=} n^{-\frac{1}{2}} \sum_{t=1}^{n} \left( e^{i\langle u, Z_t\rangle} - E\left(e^{i\langle u, Z_1\rangle}\right)\right), \quad u \in \mathbb{R}^d, \quad n \in \mathbb{N}.$$

Next, we aim to establish a similar result as in Proposition 2.1.2 which will be useful for the next Section 2.2. The difference to the previous result is that we are not interested in a $P_{\Sigma,\nu,\alpha}$-a.s. result, but in finding an upper bound for the expectation values as stated in the next theorem. Furthermore, $A_n(u)$ is scaled with $n^{-\frac{1}{2}}$ and not with $n^{-1}$ as in Proposition 2.1.2. Therefore it is not surprising that we have to make some further restrictions to the weight function $w$. To be more precise, we choose $w$ as

$$w(u) = (\log(e + |u|))^{-\frac{1}{2}-\delta}, \quad u \in \mathbb{R}^d$$

for some fixed $\delta > 0$. This is the natural generalization to $d$ dimensions of the weight function in [29].

The proof of the next Theorem 2.1.6 uses a result from empirical process theory. We briefly repeat in the following some definitions needed for this result. The respective notations in van der Vaart [43] are used: Let $(\mathcal{X}, \mathcal{A}, P)$ be a probability space and let $\mathcal{F}$ be a class of measurable functions $f : \mathcal{X} \to \mathbb{R}$ in $L^2(P)$. Fix any $\epsilon > 0$ and let $l, u : \mathcal{X} \to \mathbb{R}$ be two functions in $L^2(P)$ with $\int (l-u)^2 \, dP < \epsilon^2$. Then,

$$[l, u] \stackrel{\text{def}}{=} \{f : \mathcal{X} \to \mathbb{R}, \quad \text{measurable}, \quad l \leqslant f \leqslant u\}$$

is called an $\epsilon$-bracket. Denote further with $N_{[]}(\epsilon, \mathcal{F})$ the minimum number of such $\epsilon$-brackets needed to cover $\mathcal{F}$. Note that $l$ and $u$ are not required to belong to $\mathcal{F}$. Next

$$J_{[]}(\delta, \mathcal{F}) \stackrel{\text{def}}{=} \int_0^\delta \sqrt{\log N_{[]}(\epsilon, \mathcal{F})} \, d\epsilon, \quad \delta > 0$$

is called the bracketing integral. Let $(X_i)_{i \in \mathbb{N}}$ be a sequence of i.i.d. $\mathcal{X}$-valued and $P$ distributed random variables and set for $n \in \mathbb{N}$

$$G_n f \stackrel{\text{def}}{=} \frac{1}{\sqrt{n}} \sum_{i=1}^n (f(X_i) - Ef(X_i)), \quad \|G_n\|_{\mathcal{F}} \stackrel{\text{def}}{=} \sup_{f \in \mathcal{F}} |G_n f|.$$

Then we have

**Corollary 19.35 in van der Vaart [43].** *For any class $\mathcal{F}$ of measurable functions with envelope function $F$, it holds*

$$E^* \|G_n\|_{\mathcal{F}} \lesssim J_{[]}\left(\sqrt{\int |F|^2 \, dP}, \mathcal{F}\right). \tag{2.21}$$

Note that $\lesssim$ means not larger up to a constant which does not depend on $n \in \mathbb{N}$. An envelope function $F$ is any $L^2(P)$ function, such that $|f|(x) \leqslant F(x)$ holds for all $x \in \mathcal{X}$ and $f \in \mathcal{F}$. Finally, observe the star notation $E^*$ instead of $E$. This is due to certain measurability problems which are typical in empirical process theory, compare for this the first chapter in van der Vaart and Wellner [44]. Fortunately, we are not concerned with such measurability problems in our case and, thus, can simply write $E$ instead of $E^*$ in (2.21). In general, observe also the helpful monographs of Pollard [34] and Dudley [14].

Now, we can prove the next theorem which is a generalization of Theorem 4.1 in Neumann, Reiß [29] to the multidimensional case:

**Theorem 2.1.6.** *Let $(\Sigma, \nu, \alpha)$ be a Lévy triplet, such that $E|X_1|^{8+\gamma} < \infty$ holds for some $\gamma > 0$. Then we have*

$$\sup_{n \geqslant 1} E_{\Sigma,\nu,\alpha} \left\| \frac{\partial^l}{\partial u_k^l} A_n(u) \right\|_{L^\infty(w)} < \infty$$

*for all $l = 0, 1, 2, 3, 4$ and $1 \leqslant k \leqslant d$.*

*Proof.* Write $\Re(z)$ for the real part of a complex number $z$ and $\Im(z)$ for its imaginary part. It suffices to prove the Theorem separately for the real and imaginary part because of

$$\left\| \frac{\partial^l}{\partial u_k^l} A_n(u) \right\|_{L^\infty(w)} \leqslant \left\| \frac{\partial^l}{\partial u_k^l} \Re(A_n(u)) \right\|_{L^\infty(w)} + \left\| \frac{\partial^l}{\partial u_k^l} \Im(A_n(u)) \right\|_{L^\infty(w)} .$$

Here, we only treat the real part because the imaginary part can be proven in exactly the same way. We have

$$\Re(A_n(u)) = n^{-\frac{1}{2}} \sum_{t=1}^{n} (\cos(\langle u, Z_t \rangle) - E \cos(\langle u, Z_1 \rangle)), \quad u \in \mathbb{R}^d.$$

Set

$$G_{l,k} \stackrel{\text{def}}{=} \left\{ z \mapsto w(u) \frac{\partial^l}{\partial u_k^l} \cos(\langle u, z \rangle) \; : \; u \in \mathbb{R}^d \right\}$$

with $l = 0, 1, 2, 3, 4,\ 1 \leqslant k \leqslant d$. Next, it is the crucial idea to apply the above Corollary 19.35 in [43], i.e. empirical process theory. Here, we choose

$$f_{l,k}(z) = |z_k|^l, \quad z \in \mathbb{R}^d$$

as envelope function for the set $G_{l,k}$. Then the cited corollary implies

$$E \left\| \frac{\partial^l}{\partial u_k^l} \Re(A_n(u)) \right\|_{L^\infty(w)} \lesssim J_{[]} \left( \sqrt{E Z_{1,k}^{2l}}, G_{l,k} \right) \tag{2.22}$$

in the above Notation. With

$$M \stackrel{\text{def}}{=} M(\epsilon, l, k) = \inf \left\{ m > 0 \; : \; E(Z_{1,k}^{2l} \, \mathbb{1}_{(m,\infty)}(|Z_1|)) \leqslant \epsilon^2 \right\}$$

for $\epsilon > 0$, define

$$g_j^{\pm}(z) \stackrel{\text{def}}{=} \left( w(u^{(j)}) \frac{\partial^l}{\partial u_k^l} \cos\left( \left\langle u^{(j)}, z \right\rangle \right) \pm \epsilon |z_k|^l \right) \mathbb{1}_{[0,M]}(|z|) \pm |z_k|^l \, \mathbb{1}_{(M,\infty)}(|z|)$$

for later defined fixed points $u^{(j)} \in \mathbb{R}^d$. This yields

$$\begin{aligned} E(g_j^+(Z_1) - g_j^-(Z_1))^2 &\leqslant E\left( 4\epsilon^2 Z_{1,k}^{2l} \, \mathbb{1}_{[0,M]}(|Z_1|) + 4 Z_{1,k}^{2l} \, \mathbb{1}_{(M,\infty)}(|Z_1|) \right) \\ &\leqslant 4\epsilon^2 (E Z_{1,k}^{2l} + 1). \end{aligned}$$

Set $C \stackrel{\text{def}}{=} 2\sqrt{E Z_{1,k}^{2l} + 1}$. Then $[g_j^-, g_j^+]$ is a $C \cdot \epsilon$-bracket. Since we are only interested in the finiteness of the right-hand side of (2.22), we can assume w.l.o.g. $C = 1$. Hence, $[g_j^-, g_j^+]$ is an $\epsilon$-bracket. Next we perform some calculations in order to determine the points $u^{(j)}$, such that the upper bound in (2.22) is finite. Obviously,

$$w^1(r) \stackrel{\text{def}}{=} (\log(e + r))^{-\frac{1}{2} - \delta}, \quad r \geqslant 0$$

is Lipschitz continuous, so that we have

$$|w(u) - w(v)| = |w^1(|u|) - w^1(|v|)| \leqslant L ||u| - |v|| \leqslant L|u - v|, \quad u, v \in \mathbb{R}^d$$

for some $L > 0$. With $u, z \in \mathbb{R}^d$ and $|z| \leqslant M$, we obtain the inequalities

$$\begin{aligned} & \left| w(u) \frac{\partial^l}{\partial u_k^l} \cos(\langle u, z \rangle) - w(u^{(j)}) \frac{\partial^l}{\partial u_k^l} \cos\left(\left\langle u^{(j)}, z \right\rangle\right) \right| \\ \leqslant\ & |w(u) - w(u^{(j)})| \left| \frac{\partial^l}{\partial u_k^l} \cos(\langle u, z \rangle) \right| \\ & + |w(u^{(j)})| \left| \frac{\partial^l}{\partial u_k^l} \cos(\langle u, z \rangle) - \frac{\partial^l}{\partial u_k^l} \cos\left(\left\langle u^{(j)}, z \right\rangle\right) \right| \\ \leqslant\ & L|u - u^{(j)}||z_k|^l + \left| z_k^l \left( \langle u, z \rangle - \left\langle u^{(j)}, z \right\rangle \right) \right| \\ \leqslant\ & L|u - u^{(j)}||z_k|^l + |u - u^{(j)}||z_k|^l |z| \\ \leqslant\ & |z_k|^l |u - u^{(j)}|(L + M). \end{aligned} \tag{2.23}$$

This yields that (2.23) is not larger than

$$|z_k|^l \min\left\{ \sqrt{d}|u - u^{(j)}|_\infty (L + M), w(u) + w(u^{(j)}) \right\}, \quad z \in \mathbb{R}^d \ :\ |z| \leqslant M. \tag{2.24}$$

Set

$$U(\epsilon) \stackrel{\text{def}}{=} \inf\left\{ u > 0 \ :\ \sup_{v \in \mathbb{R}^d \,:\, |v| \geqslant u} w(v) \leqslant \frac{\epsilon}{2} \right\}$$

and

$$J(\epsilon) \stackrel{\text{def}}{=} \inf\left\{ l \in \mathbb{N} \ :\ \frac{l\epsilon}{\sqrt{d}(L + M)} \geqslant U(\epsilon) \right\}. \tag{2.25}$$

Now, we specify the points $u^{(j)}$ as

$$u^{(j)} \stackrel{\text{def}}{=} \frac{j\epsilon}{\sqrt{d}(L + M)}, \quad j \in \mathbb{Z}^d \ :\ |j|_\infty \leqslant J(\epsilon). \tag{2.26}$$

This choice guarantees that

$$G_{l,k} \subseteq \bigcup_{j \in \mathbb{Z}^d \,:\, |j|_\infty \leqslant J(\epsilon)} [g_j^-, g_j^+].$$

To understand this, fix any $u \in \mathbb{R}^d$. If

$$|u|_\infty \leqslant \frac{J(\epsilon)\epsilon}{\sqrt{d}(L + M)},$$

set

$$j_u \stackrel{\text{def}}{=} \left( \left\lfloor \frac{\sqrt{d}(L + M)}{\epsilon} u_1 \right\rfloor, \ldots, \left\lfloor \frac{\sqrt{d}(L + M)}{\epsilon} u_d \right\rfloor \right) \in \mathbb{Z}^d$$

with $\lfloor x \rfloor = -\lfloor |x| \rfloor$, $x < 0$. Note that $|j_u|_\infty \leqslant J(\epsilon)$. Then

$$\sqrt{d}|u - u^{(j_u)}|_\infty (L + M) \leqslant \epsilon$$

and the corresponding function belongs to $[g_{j_u}^-, g_{j_u}^+]$. If

$$|u|_\infty > \frac{J(\epsilon)\epsilon}{\sqrt{d}(L + M)} \geqslant U(\epsilon),$$

the corresponding function belongs to

$$j_u \stackrel{\text{def}}{=} (J(\epsilon), 0, \ldots, 0) \in \mathbb{Z}^d$$

because of (2.24) and

$$w(|u|) \leqslant w(|u|_\infty) \leqslant \frac{\epsilon}{2}, \quad w(|u^{(j_u)}|) = w^1\left(\frac{J(\epsilon)\epsilon}{\sqrt{d}(L+M)}\right) \leqslant \frac{\epsilon}{2}.$$

Note that (2.26) implies

$$N_{[]}(\epsilon, G_{l,k}) \leqslant (2J(\epsilon)+1)^d. \tag{2.27}$$

Next we establish

$$M \leqslant \left(\frac{E|Z_1|^{2l+\gamma}}{\epsilon^2}\right)^{\frac{1}{\gamma}}. \tag{2.28}$$

This is easily done by setting

$$m^\gamma \stackrel{\text{def}}{=} \frac{E|Z_1|^{2l+\gamma}}{\epsilon^2}$$

and considering the inequalities

$$\begin{aligned} E|Z_1|^{2l+\gamma} &\geqslant E\left(|Z_1|^{2l+\gamma}\, \mathbb{1}_{(m,\infty)}(|Z_1|)\right) \\ &\geqslant E\left(|Z_{1,k}|^{2l} m^\gamma\, \mathbb{1}_{(m,\infty)}(|Z_1|)\right) \\ &= m^\gamma E\left(|Z_{1,k}|^{2l}\, \mathbb{1}_{(m,\infty)}(|Z_1|)\right). \end{aligned}$$

(2.25) yields

$$J(\epsilon) \leqslant \frac{U(\epsilon)\sqrt{d}(L+M)}{\epsilon} + 1. \tag{2.29}$$

The special shape of $w$, furthermore, yields

$$\log(U(\epsilon)) = \left(\frac{\epsilon}{2}\right)^{-(\delta+\frac{1}{2})^{-1}} + o(1), \quad \epsilon \to 0,$$

so that we have together with (2.27), (2.28) and (2.29)

$$\begin{aligned} \log(N_{[]}(\epsilon, G_{l,k})) &\leqslant d\log(2J(\epsilon)+1) \\ &= O\left(\epsilon^{-\left(\delta+\frac{1}{2}\right)^{-1}} + \log\left(\epsilon^{-1-\frac{2}{\gamma}}\right)\right), \quad \epsilon \to 0. \end{aligned}$$

As $\left(\delta + \frac{1}{2}\right)^{-1} < 2$, we have established

$$\int_0^{\sqrt{EZ_{1,k}^{2l}}} \sqrt{\log(N_{[]}(\epsilon, G_{l,k}))}\, d\epsilon < \infty$$

and (2.22) is finite. □

Next, we define a $\mathcal{C}^4(\mathbb{R}^d)$ metric

$$d(\varphi_1,\varphi_2) \stackrel{\text{def}}{=} d^{(4)}(\varphi_1,\varphi_2) \stackrel{\text{def}}{=} d^{(2)}(\varphi_1,\varphi_2) + \sum_{k=1}^d \left\| \frac{\partial^3}{\partial u_k^3}(\varphi_1-\varphi_2)\right\|_{L^\infty(w)} + \sum_{k=1}^d \left\| \frac{\partial^4}{\partial u_k^4}(\varphi_1-\varphi_2)\right\|_{L^\infty(w)}.$$

Then, as a direct consequence of Theorem 2.1.6, it holds the following statement:

**Corollary 2.1.7.** *Let $(\Sigma, \nu, \alpha)$ be a Lévy triplet such that $E|X_1|^{8+\gamma} < \infty$ holds for some $\gamma > 0$. Then we have*

$$E_{\Sigma,\nu,\alpha} d(\widehat{\varphi}_n, \varphi_{\Sigma,\nu,\alpha}) = O\left(n^{-\frac{1}{2}}\right), \quad n \to \infty.$$

## 2.2 Nonparametric low frequency Lévy copula estimation

We denote with $\mathcal{F}$ the Fourier transform of a function or a finite measure. To be more precise, we set for $u \in \mathbb{R}^d$

$$(\mathcal{F}f)(u) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} e^{i\langle u,x\rangle} f(x)\, \lambda^2(dx), \quad f \in L^1(\lambda^d)$$

and

$$(\mathcal{F}\mu)(u) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} e^{i\langle u,x\rangle}\, \mu(dx),$$

where $\mu$ denotes a finite positive measure on the space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

As described in the introduction of this chapter, we aim to estimate the Lévy measure $\nu$ in order to construct a Lévy copula estimator. Motivated by Nickl and Reiß [30], we do not estimate directly $\nu$, but a smoothed version of $\nu$. The statistical estimation of this smoothed version is investigated in the proof of Theorem 2.2.11. An upper bound of the error which we make by using a smoothed version of $\nu$ instead of $\nu$ itself, is calculated in Lemma 2.2.7. We consider the convolution of $\nu$ with a Kernel $K$ in order to get such a smoothed version of $\nu$, cf. Lemma 2.2.7. Such a Kernel, of course, has to fulfill some assumptions which are stated next:

**Assumptions 2.2.1.** Let $K : \mathbb{R}^2 \to \mathbb{R}_+$ be a kernel function with the properties

(i) $K \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$, $\int_{\mathbb{R}} K(x)\, \lambda^2(dx) = 1$

(ii) $\operatorname{supp}(\mathcal{F}K) \subseteq [-1,1]^2$

(iii) $u \mapsto (\mathcal{F}K)(u)$ is Lipschitz continuous.

It is natural to consider Lévy processes in the Fourier space because of the Lévy-Khintchine formula. From this point of view, Assumption 2.2.1 (ii) is particularly useful because it provides compact support for many important integrands we use.

**Example 2.2.2.** In the following example, we state a kernel function $K$ which fulfills the Assumptions 2.2.1. First, set

$$\begin{array}{rccl} K_1 : & \mathbb{R} & \to & \mathbb{R}_+ \\ & x_1 & \mapsto & \frac{2}{\pi}\left(\frac{\sin\left(\frac{x_1}{2}\right)}{x_1}\right)^2. \end{array}$$

A straightforward calculation yields

$$(\mathcal{F}K_1)(u_1) = (1 - |u_1|)\, \mathbb{1}_{(-1,1)}(u_1), \quad u_1 \in \mathbb{R}.$$

Set

$$K(x) \stackrel{\text{def}}{=} K_1(x_1) \cdot K_1(x_2), \quad x = (x_1, x_2) \in \mathbb{R}^2.$$

This implies

$$(\mathcal{F}K)(u) = (\mathcal{F}K_1)(u_1) \cdot (\mathcal{F}K_1)(u_2), \quad u = (u_1, u_2) \in \mathbb{R}^2.$$

Note that $K$ fulfills the desired conditions since

$$\int_{\mathbb{R}^2} K(x)\lambda^2(x) = (\mathcal{F}K)(0) = (\mathcal{F}K_1)(0)\cdot(\mathcal{F}K_1)(0) = 1$$

and we have for $u, v \in \mathbb{R}^2$

$$\begin{aligned} |\mathcal{F}K(u) - \mathcal{F}K(v)| &\leqslant |\mathcal{F}K_1(u_1)(\mathcal{F}K_1(u_2) - \mathcal{F}K_1(v_2))| + |\mathcal{F}K_1(v_2)(\mathcal{F}K_1(u_1) - \mathcal{F}K_1(v_1))| \\ &\leqslant |\mathcal{F}K_1(u_2) - \mathcal{F}K_1(v_2)| + |\mathcal{F}K_1(u_1) - \mathcal{F}K_1(v_1)| \\ &\leqslant |u_2 - v_2| + |u_1 - v_1| \\ &\leqslant \sqrt{2}|u-v|. \end{aligned}$$

The remaining conditions are obviously true.

Next set for $h > 0$

$$K_h(x) \overset{\text{def}}{=} h^{-2}K(h^{-1}x) = (h^{-1}K_1(h^{-1}x_1))\cdot(h^{-1}K_1(h^{-1}x_2)), \quad x \in \mathbb{R}^2$$

and observe that standard results from Fourier analysis yield

$$(\mathcal{F}K_h)(u) = (\mathcal{F}K)(hu), \quad u \in \mathbb{R}^2.$$

Recall the definitions

$$U(x,y) = \nu([x,\infty)\times[y,\infty)), \quad U_1(x) = \nu([x,\infty)\times\mathbb{R}_+), \quad U_2(y) = \nu(\mathbb{R}_+\times[y,\infty)), \quad x,y \in \mathbb{R}_+$$

and $\mathfrak{R} = \mathbb{R}_+^2\backslash\{(0,0)\}$ from the introduction of this chapter. Define, furthermore,

$$g_{a,b}(x) \overset{\text{def}}{=} \frac{1}{x_1^4 + x_2^4}\,\mathbb{1}_{[a,\infty)\times[b,\infty)}(x_1,x_2), \quad (a,b)\in\mathfrak{R}, \quad x \in \mathbb{R}^2.$$

**Assumptions 2.2.3.** *Next we state some assumptions concerning the Lévy measure $\nu$:*

*(i)* $\nu\left(\mathbb{R}^2\backslash[0,\infty)^2\right) = 0$, *i.e. only positive jumps,*

*(ii)* $\exists\gamma > 0 : \int |x|^{8+\gamma}\,\nu(dx) < \infty$, *i.e. finite $8+\gamma$-th moment,*

*(iii)* $\mathcal{F}((x_1^4 + x_2^4)\nu)(u) \lesssim (1+|u_1|)^{-1}(1+|u_2|)^{-1}$, $u \in \mathbb{R}^2$,

*(iv)* $U_k : (0,\infty) \to (0,\infty)$ *is a $\mathcal{C}^1$-bijection with $U_k' < 0$ and*

$$\inf_{0<x_k\leqslant 1} |U_k'(x_k)| > 0, \quad \sup_{x_k>0}(1\wedge x_k^3)|U_k'(x_k)| < \infty, \quad k = 1,2. \tag{2.30}$$

**Remark 2.2.4.** Assumption 2.2.3 (i) assures that there are no negative jumps. This simplifies the shape of the Lévy copula of $\nu$, cf. (2.2) and serves to keep the technical overhead as small as possible. (ii) required to use the statement of Corollary 2.1.7. (iii) is perhaps the most non-transparent assumption. It guarantees a certain decay behaviour of some integrands in the Fourier space. Finally, (iv) is needed to build a pseudo inverse in order to estimate the Lévy copula of $\nu$ which is our final goal.

**Proposition 2.2.5.** *Let $f : \mathbb{R}_+^2 \to \mathbb{R}_+$ be a continuous function with the properties* *(i)* $f(x) > 0, \quad x \in (\{0\} \times \mathbb{R}_+^*) \times (\mathbb{R}_+^* \times \{0\}),$

*(ii)* $|x|^2 \lesssim f(x) \lesssim (\log|x|)^{-2}, \quad x \in \mathbb{R}_+^2 \,:\, |x| \leqslant \frac{1}{2},$

*(iii)* $f(x) \lesssim (1+|x|)^{-(6+\epsilon)}, \quad x \in \mathbb{R}_+^2$

*for some* $\epsilon > 0$. *Then,*

$$\nu(dx) \overset{\text{def}}{=} \mathbb{1}_{\mathfrak{R}}(x)(x_1^4 + x_2^4)^{-1} f(x) \lambda^2(dx)$$

*is a Lévy measure and fulfills the Assumptions 2.2.3 (i), (ii) and (iv).*

*Proof.* First, observe that $\nu$ is a Lévy measure since

$$\int_{\mathbb{R}_+^2} |x|^2 \,\nu(dx) \lesssim \int_0^{\frac{1}{2}} r^2 r^{-4} (\log r)^{-2} r\,dr + \int_{\frac{1}{2}}^{\infty} r^2 r^{-4} (1+r)^{-(6+\epsilon)} r\,dr < \infty$$

holds. Next, we turn to the claimed Assumptions 2.2.3 (i), (ii) and (iv).

(i) This is obviously true due to $\mathfrak{R} \subseteq \mathbb{R}_+^2$.

(ii) Note that we have

$$\int_{\mathbb{R}_+^2} |x|^{8+\frac{\epsilon}{2}} \,\nu(dx) \lesssim \int_{\mathbb{R}_+^2} |x|^{4+\frac{\epsilon}{2}} f(x) \,\lambda^2(dx) \lesssim \int_0^\infty r^{4+\frac{\epsilon}{2}} (1+r)^{-(6+\epsilon)} r\,dr < \infty.$$

Hence, Sato [38][Theorem 25.3] yields that the $(8+\gamma)$-th moment with $\gamma \overset{\text{def}}{=} \frac{\epsilon}{2} > 0$ of the corresponding Lévy process exists.

(iv) First, observe

$$0 = \lim_{x_1 \uparrow \infty} U_1(x_1) \leqslant \lim_{x_1 \downarrow 0} U_1(x_1) = \infty \tag{2.31}$$

because of

$$[x_1, \infty) \times \mathbb{R}_+ \downarrow \varnothing, \quad x_1 \uparrow \infty$$

and

$$\nu(\mathbb{R}_+^2) = \int_{\mathbb{R}_+^2} (x_1^4 + x_2^4)^{-1} f(x) \,\lambda^2(dx) \gtrsim \int_0^{\frac{1}{2}} r^{-4} \cdot r^2 r\,dr = \infty.$$

Next, $U_1(x_1) > 0$ follows from $f(\cdot, 0) > 0$ on $(0, \infty)$ and the continuity of $f$. Hence (2.31) yields that $U_1 : (0,\infty) \to (0,\infty)$ is a surjection. We, furthermore, have for $x_1 > 0$

$$U_1'(x_1) = \frac{\partial}{\partial x_1} \int_{x_1}^\infty \int_0^\infty (y_1^4 + y_2^4)^{-1} f(y) \,dy_2 \,dy_1 = -\int_0^\infty (x_1^4 + y_2^4)^{-1} f(x_1, y_2) \,dy_2. \tag{2.32}$$

Again due to the continuity of $f$ and $f(\cdot, 0) > 0$, this implies $U_1' < 0$ on $(0, \infty)$. Hence, $U_1$ is also injective, i.e. a bijection. Finally, (2.32) also implies (2.30) for $k = 1$. Observe for this purpose

$$|U_1'(x_1)| \leqslant \|f\|_\infty \int_0^\infty (x_1^4 + y_2^4)^{-1} \,dy_2 = \frac{\|f\|_\infty \sqrt{2}\pi}{4x_1^3} \lesssim x_1^{-3}, \quad x_1 > 0$$

and, with the use of Fatou's Lemma,

$$\liminf_{x_1 \to 0} |U_1'(x_1)| \geqslant \int_0^\infty \liminf_{x_1 \to 0} [(x_1^4 + y_2^4)^{-1} f(x_1, y_2)] \,dy_2 = \int_0^\infty y_2^{-4} f(0, y_2) \,dy_2 > 0.$$

Now, (iv) is verified since everything works equally with $U_2$ instead of $U_1$.

□

**Corollary 2.2.6.** *It exists, for every $0 \leqslant \beta \leqslant 2$, a Lévy measure $\nu_\beta$ with Blumenthal Getoor index (BGi) $\beta$, such that $\nu_\beta$ fulfills the Assumptions 2.2.3.*

*Proof.* We treat the cases $0 \leqslant \beta < 2$ and $\beta = 2$ separately in two steps:

STEP 1. The case $0 \leqslant \beta < 2$. Set

$$f_\beta(x) \stackrel{\text{def}}{=} r^{2-\beta} e^{-r}, \quad r \stackrel{\text{def}}{=} |x|, \quad x \in \mathbb{R}_+^2$$

and

$$\nu_\beta(dx) \stackrel{\text{def}}{=} \mathbb{1}_{\mathfrak{R}}(x)(x_1^4 + x_2^4)^{-1} f_\beta(x)\, \lambda^2(dx).$$

Then $\nu_\beta$ is a Lévy measure of BGi $\beta$ because it holds for $\gamma > \beta$

$$\int_{\{|x| \leqslant 1\}} |x|^\gamma \, \nu(dx) \lesssim \int_0^1 r^\gamma r^{-4} r^{2-\beta} r \, dr = \int_0^1 r^{\gamma-\beta-1} \, dr < \infty$$

and for $\gamma = \beta$

$$\int_{\{|x| \leqslant 1\}} |x|^\beta \, \nu(dx) \gtrsim \int_0^1 r^\beta r^{-4} r^{2-\beta} r \, dr = \int_0^1 r^{-1} \, dr = \infty.$$

Furthermore, the Assumptions 2.2.3 (i), (ii) and (iv) are fulfilled because of Proposition 2.2.5 and

$$|x|^2 \lesssim |x|^{2-\beta} e^{-|x|} \lesssim (\log |x|)^{-2}, \quad x \in \mathbb{R}_+^2 \, : \, |x| \leqslant \frac{1}{2}, \quad 0 \leqslant \beta < 2.$$

Next, we show that $\nu_\beta$ fulfills Assumption 2.2.3 (iii). Note for this after a straightforward calculation the equations

$$\begin{aligned} \frac{\partial f_\beta}{\partial x_k}(x) &= x_k \left( (2-\beta) r^{-\beta} - r^{1-\beta} \right) e^{-r}, \quad k = 1, 2, \\ \frac{\partial^2 f_\beta}{\partial x_1 \partial x_2}(x) &= x_1 x_2 \left( \beta(\beta-2) r^{-\beta-2} + (2\beta-3) r^{-\beta-1} + r^{-\beta} \right) e^{-r}. \end{aligned}$$

Hence, it holds for $x \in \mathfrak{R}$, $r = |x| > 0$ and $k = 1, 2$

$$\left| \frac{\partial f_\beta}{\partial x_k} \right| (x) \lesssim \mathbb{1}_{(0,1)}(r) r^{1-\beta} + r^2 e^{-r}, \quad \left| \frac{\partial^2 f_\beta}{\partial x_1 \partial x_2} \right| (x) \lesssim \mathbb{1}_{(0,1)}(r) r^{-\beta} + r^2 e^{-r}.$$

Now, fix any $0 < \delta < 1$. Then, Proposition B.2 yields

$$\mathcal{F}(f_\beta \cdot \mathbb{1}_{[\delta,\infty)^2})(u) \leqslant \frac{\Lambda_{\beta,\delta}}{|u_1 u_2|}, \quad u \in (\mathbb{R}^*)^2 \tag{2.33}$$

with

$$\begin{aligned} \Lambda_{\beta,\delta} &= |f_\beta(\delta,\delta)| + \int_\delta^\infty \frac{\partial f_\beta}{\partial x_1}(x_1, \delta) \, dx_1 + \int_\delta^\infty \frac{\partial f_\beta}{\partial x_2}(\delta, x_2) \, dx_2 + \int_{[\delta,\infty)^2} \left| \frac{\partial^2 f_\beta}{\partial x_1 \partial x_2} \right| (x) \, dx \\ &\lesssim 1 + \int_\delta^1 (x_1^2 + \delta^2)^{\frac{1-\beta}{2}} \, dx_1 + \int_\delta^1 (\delta^2 + x_2^2)^{\frac{1-\beta}{2}} \, dx_2 + \int_\delta^1 r^{-\beta} r \, dr. \end{aligned}$$

If $0 \leqslant \beta \leqslant 1$, we have

$$\Lambda_{\beta,\delta} \lesssim 1 + 2\int_0^1 (x_1^2+1)^{\frac{1-\beta}{2}}\,dx_1 + \int_0^1 r^{1-\beta}\,dr < \infty. \tag{2.34}$$

If $1 < \beta < 2$, it holds

$$\Lambda_{\beta,\delta} \lesssim 1 + 2\int_0^1 x_1^{1-\beta}\,dx_1 + \int_0^1 r^{1-\beta}\,dr < \infty. \tag{2.35}$$

Note that the constants in the $\lesssim$ sign are independent of $0 < \delta < 1$ and that we obtain by dominated convergence

$$\mathcal{F}(f_\beta \cdot \mathbb{1}_{[\delta,\infty)^2})(u) \to \mathcal{F}(f_\beta)(u), \quad \delta \downarrow 0 \tag{2.36}$$

pointwise for all $u \in \mathbb{R}^2$. Thus, (2.33) - (2.36) yield together for fixed $\beta$

$$\mathcal{F}(f_\beta)(u) \lesssim \frac{1}{|u_1||u_2|}, \quad u \in (\mathbb{R}^*)^2. \tag{2.37}$$

Hence, $\nu_\beta$ satisfies

$$\mathcal{F}((x_1^4+x_2^4)\nu_\beta)(u) \lesssim \frac{1}{|u_1||u_2|}, \quad u \in (\mathbb{R}^*)^2$$

This proves together with Lemma B.5 (i) and the continuity of $u \mapsto \mathcal{F}((x_1^4+x_2^4)\nu)(u)$ the Assumption 2.2.3 (iii) and, thus, the first step is accomplished.

STEP 2. The case $\beta = 2$. Let $\phi : \mathbb{R}_+^2 \to [0,1]$ be a $\mathcal{C}^\infty$ function with

$$\phi(x) = \begin{cases} 1, & |x| \leqslant \frac{1}{2} \\ 0, & |x| > \frac{3}{4} \end{cases}, \quad x \in \mathbb{R}_+^2.$$

A detailed construction of such a function is given in Rudin [36][§1.46]. Set

$$f_2(x) \overset{\text{def}}{=} \phi(r)(\log r)^{-2} + (1-\phi(r))e^{-r}, \quad r = |x|, \quad x \in \mathfrak{R}$$

and observe that

$$\nu_2(dx) \overset{\text{def}}{=} \mathbb{1}_{\mathfrak{R}}(x)(x_1^4+x_2^4)^{-1} f_2(x)\,\lambda^2(dx)$$

is a Lévy measure of BGi 2 because we have

$$\int_{\{|x|\leqslant 1\}} |x|^2\,\nu(dx) \lesssim \int_0^{\frac{3}{4}} r^2 r^{-4}(\log r)^{-2} r\,dr = \left(\log\left(\frac{4}{3}\right)\right)^{-1} < \infty$$

and

$$\int_{\{|x|\leqslant 1\}} |x|^{2-\gamma}\,\nu(dx) \gtrsim \int_0^{\frac{1}{2}} r^{2-\gamma} r^{-4}(\log r)^{-2} r\,dr \gtrsim \int_0^{\frac{1}{2}} r^{-1-\frac{\gamma}{2}}\,dr = \infty$$

for every $\gamma > 0$. Note further that the Assumptions 2.2.3 (i), (ii) and (iv) hold because of Proposition 2.2.5 and

$$|x|^2 \lesssim (\log|x|)^{-2} + e^{-|x|}(1-\phi(|x|)) \lesssim (\log|x|)^{-2}, \quad x \in \mathbb{R}_+^2 : |x| \leqslant \frac{1}{2}.$$

Next, we establish the Assumption 2.2.3 (iii). For this purpose, set $Lr \overset{\text{def}}{=} r\log r$, $r > 0$ and note that it holds for $0 < r < \frac{1}{2}$ and $k = 1, 2$

$$\begin{aligned}
\frac{\partial f_2}{\partial x_k}(x) &= x_k(\phi'(r)(Lr)^{-2}r - 2\phi(r)(Lr)^{-3}r), \quad k = 1, 2, \\
\frac{\partial^2 f_2}{\partial x_1 \partial x_2}(x) &= x_1 x_2(\phi''(r)(Lr)^{-2} - 4\phi'(r)(Lr)^{-3} - \phi'(r)(Lr)^{-2}r^{-1} + 6\phi(r)(Lr)^{-4} \\
&\quad +4\phi(r)(Lr)^{-3}r^{-1}).
\end{aligned}$$

This implies for $0 < r < \frac{1}{2}$ the asymptotics

$$\left|\frac{\partial f_2}{\partial x_k}\right|(x) \lesssim r^{-1}(\log r)^{-2}, \quad \left|\frac{\partial^2 f_2}{\partial x_1 \partial x_2}\right|(x) \lesssim r^{-2}(\log r)^{-2}.$$

Observe that we have in the complementary case $r > 1$

$$\begin{aligned}
\frac{\partial f_2}{\partial x_k}(x) &= -x_k r^{-1} e^{-r}, \\
\frac{\partial^2 f_2}{\partial x_1 \partial x_2}(x) &= x_1 x_2 (r^{-2} + r^{-3}) e^{-r}.
\end{aligned}$$

This yields for $r > 1$ the asymptotics

$$\left|\frac{\partial f_2}{\partial x_k}\right|(x) \lesssim e^{-r}, \quad \left|\frac{\partial^2 f_2}{\partial x_1 \partial x_2}\right|(x) \lesssim e^{-r}.$$

Now, we get the claim of Assumption 2.2.3 (iii) with the same procedure as in the first step. □

Denote with $(K_h\lambda^2) * ((x_1^4 + x_2^4)\nu)$ in the following the convolution of the two finite Borel measures $d(K_h\lambda^2) \overset{\text{def}}{=} K_h \, d\lambda^2$ and $d((x_1^4 + x_2^4)\nu) \overset{\text{def}}{=} (x_1^4 + x_2^4)\, d\nu$.

**Lemma 2.2.7.** *Let the above Assumptions 2.2.1 and 2.2.3 (iii) hold. Then we have*

$$\left|\nu([a,\infty) \times [b,\infty)) - \int_{\mathbb{R}^2} g_{a,b}(x)[(K_h\lambda^2) * (x_1^4 + x_2^4)\nu)](dx)\right| \lesssim |h \log h|(|(a,b)|^{-2} \vee |(a,b)|^{-4})$$

*for all $(a,b) \in \mathfrak{R}$ and $0 < h < \frac{1}{2}$.*

*Proof.* Write

$$\begin{aligned}
&\left|\nu([a,\infty) \times [b,\infty)) - \int_{\mathbb{R}^2} g_{a,b}(x)[(K_h\lambda^2) * ((x_1^4 + x_2^4)\nu)](dx)\right| \\
= &\left|\int_{\mathbb{R}^2} g_{a,b}(x) \left\{[(x_1^4 + x_2^4)\nu](dx) - [(K_h\lambda^2) * ((x_1^4 + x_2^4)\nu)](dx)\right\}\right| \\
= &\frac{1}{4\pi}\left|\int_{\mathbb{R}^2} (\mathcal{F}g_{a,b})(-u)(1 - (\mathcal{F}K_h)(u))\mathcal{F}((x_1^4 + x_2^4)\nu)(u)\, \lambda^2(du)\right| \qquad (2.38)
\end{aligned}$$

where we use Lemma B.6 and the fact that a convolution becomes a simple multiplication in the Fourier space for the last inequality. Note further

$$|1 - (\mathcal{F}K_h)(u)| = |(\mathcal{F}K)(0) - (\mathcal{F}K)(hu)| \lesssim \min(h|u|, 1), \quad u \in \mathbb{R}^2,$$

due to the Lipschitz continuity and our assumption that $K$ is normalized. Hence, using Corollary B.4 together with Lemma B.5 (i), (2.38) is up to a constant not larger than $I_1 + I_2$ with

$$I_1 \overset{\text{def}}{=} h|(a,b)|^{-2} \int_{[-1,1]^2} |u|(1+|u_1|)^{-1}(1+|u_2|)^{-1}\, \lambda^2(du)$$

and

$$I_2 \overset{\text{def}}{=} |(a,b)|^{-4} \int_{\mathbb{R}^2} \min(h|u|,1)(1+|u_1|)^{-2}(1+|u_2|)^{-2}\, \lambda^2(du).$$

This finally proves under consideration of Lemma B.5 (ii), Fubinis theorem and

$$\min(h|u|,1) \leqslant \min(h|u_1|,1) + \min(h|u_2|,1), \qquad \int_{\mathbb{R}} (1+|z|)^{-2}\, \lambda^1(dz) < \infty$$

this lemma. □

**Lemma 2.2.8.** *Let $\varphi_{\Sigma,\nu,\alpha}$ be the characteristic function of an infinitesimal divisible two dimensional distribution with Lévy triplet $(\Sigma,\nu,\alpha)$ and finite second moment. Then we have*

$$|\varphi_{\Sigma,\nu,\alpha}(u)| \geqslant e^{-C(1+|u|)^2}, \quad u \in \mathbb{R}^2 \tag{2.39}$$

*with a constant $C$ depending only on the triplet $(\Sigma,\nu,\alpha)$.*

**Remark 2.2.9.** Note, that the fast exponential decay to zero in (2.39) as $|u|$ tends to infinity results from a possible non-vanishing $\Sigma$. Otherwise $\varphi_{\Sigma,\nu,\alpha}(u)$ may possibly have a slower convergence rate to zero. In this context, review the results in Neumann, Reiß [29]. In the case of a compound Poisson process, it is even bounded away from zero, cf. Lemma 2.3.4.

*Proof of Lemma 2.2.8.* We have $\varphi(u) = \exp(\Psi(u))$ with

$$\Psi(u) = -\frac{1}{2}\langle u, \Sigma u\rangle + i\langle u,\alpha\rangle + \int_{\mathbb{R}^2} \left(e^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle\right) \nu(dx), \quad u \in \mathbb{R}^2. \tag{2.40}$$

Next, we estimate each summand of $\Psi$ separately:

$$\begin{aligned} |\langle u,\Sigma u\rangle| &\leqslant |u||\Sigma u| \leqslant |\Sigma||u|^2, \\ |\langle u,\alpha\rangle| &\leqslant |\alpha||u|, \quad u \in \mathbb{R}^2. \end{aligned}$$

Furthermore, Sato [38][Lemma 8.6.] yields

$$e^{i\langle u,x\rangle} = 1 + i\langle u,x\rangle + \theta_{u,x}\frac{|\langle u,x\rangle|^2}{2}, \quad u,x \in \mathbb{R}^2, \quad \theta_{u,x} \in \mathbb{C}, \quad |\theta_{u,x}| \leqslant 1$$

which implies

$$\left|\int_{\mathbb{R}^2} (e^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle)\, \nu(dx)\right| \leqslant \int_{\mathbb{R}^2} |\langle u,x\rangle|^2\, \nu(dx) \leqslant |u|^2 \int_{\mathbb{R}^2} |x|^2\, \nu(dx).$$

This yields with $C \overset{\text{def}}{=} |\alpha| + |\Sigma| + \int_{\mathbb{R}^2} |x|^2\, \nu(dx)$ the estimate

$$|\varphi(u)| = |e^{\Psi(u)}| \geqslant e^{-|\Psi(u)|} \geqslant e^{-C(1+|u|)^2}, \quad u \in \mathbb{R}^2$$

and this Lemma is proven. □

In the following we construct, based on low frequency observations, a uniform estimator for the values

$$\{\nu([a,\infty)\times[b,\infty))\,:\,(a,b)\in\mathfrak{R}\}.$$

Let us assume that the corresponding Lévy process has a finite fourth moment. Our motivation is the following fact:

$$\begin{aligned}\left(\frac{\partial^4\Psi_{\Sigma,\nu,\alpha}}{\partial u_1^4}+\frac{\partial^4\Psi_{\Sigma,\nu,\alpha}}{\partial u_2^4}\right)(u) &= \int_{\mathbb{R}^2} e^{i\langle u,x\rangle}(x_1^4+x_2^4)\,\nu(dx)\\ &= \mathcal{F}((x_1^4+x_2^4)\nu)(u),\quad u\in\mathbb{R}^2. \end{aligned}\tag{2.41}$$

**Remark 2.2.10.** Note that we have to take at least the third derivations of $\Psi$ in order to dump the Brownian motion part. However, this seems not to be sufficient to deal with Lévy measures with Blumenthal Getoor indices greater than one. That is why we take the fourth derivations of $\Psi$. Doing so, we are, for example, capable to prove Corollary 2.2.6.

A simple calculation yields for $k=1,2$

$$\begin{aligned}\frac{\partial^4\Psi_{\Sigma,\nu,\alpha}}{\partial u_k^4} &= \frac{\partial^4\varphi_{\Sigma,\nu,\alpha}}{\partial u_k^4}\varphi_{\Sigma,\nu,\alpha}^{-1}-4\frac{\partial\varphi_{\Sigma,\nu,\alpha}}{\partial u_k}\frac{\partial^3\varphi_{\Sigma,\nu,\alpha}}{\partial u_k^3}\varphi_{\Sigma,\nu,\alpha}^{-2}-3\left(\frac{\partial^2\varphi_{\Sigma,\nu,\alpha}}{\partial u_k^2}\varphi_{\Sigma,\nu,\alpha}^{-1}\right)^2\\ &\quad+12\left(\frac{\partial\varphi_{\Sigma,\nu,\alpha}}{\partial u_k}\right)^2\frac{\partial^2\varphi_{\Sigma,\nu,\alpha}}{\partial u_k^2}\varphi_{\Sigma,\nu,\alpha}^{-3}-6\left(\frac{\partial\varphi_{\Sigma,\nu,\alpha}}{\partial u_k}\varphi_{\Sigma,\nu,\alpha}^{-1}\right)^4.\end{aligned}\tag{2.42}$$

Note that we are going to estimate $\varphi_{\Sigma,\nu,\alpha}$ by

$$\widehat{\varphi}_n(u)=\frac{1}{n}\sum_{t=1}^n e^{i\langle u,X_t-X_{t-1}\rangle},\quad u\in\mathbb{R}^2.$$

Hence, we set for $k=1,2$

$$\begin{aligned}\frac{\partial\widehat{\Psi}_n}{\partial u_k^4} &\overset{\text{def}}{=} \frac{\partial^4\widehat{\varphi}_n}{\partial u_k^4}\widehat{\varphi}_n^{-1}-4\frac{\partial\widehat{\varphi}_n}{\partial u_k}\frac{\partial^3\widehat{\varphi}_n}{\partial u_k^3}\widehat{\varphi}_n^{-2}-3\left(\frac{\partial^2\widehat{\varphi}_n}{\partial u_k^2}\widehat{\varphi}_n^{-1}\right)^2+12\left(\frac{\partial\widehat{\varphi}_n}{\partial u_k}\right)^2\frac{\partial^2\widehat{\varphi}_n}{\partial u_k^2}\widehat{\varphi}_n^{-3}\\ &\quad-6\left(\frac{\partial\widehat{\varphi}_n}{\partial u_k}\widehat{\varphi}_n^{-1}\right)^4,\end{aligned}\tag{2.43}$$

i.e. $\frac{\partial^4\widehat{\Psi}_n}{\partial u_k^4}$ is a function of derivatives of $\widehat{\varphi}_n$ in exactly the same manner as $\frac{\partial^4\Psi_{\Sigma,\nu,\alpha}}{\partial u_k^4}$ is a function of the derivatives of $\varphi_{\Sigma,\nu,\alpha}$, compare (2.42). Of course we cannot write $\widehat{\varphi}_n(u)=e^{\widehat{\Psi}_n(u)}$ since $\widehat{\varphi}_n$ need not be a characteristic function of an infinitesimal divisible measure for each $\omega\in\Omega$.

Considering (2.41), we set

$$\widetilde{N}_n(a,b)\overset{\text{def}}{=}\int_{\mathbb{R}^2} g_{a,b}(x)\mathcal{F}^{-1}\left(\left(\frac{\partial^4\widehat{\Psi}_n}{\partial u_1^4}+\frac{\partial^4\widehat{\Psi}_n}{\partial u_2^4}\right)\mathcal{F}K_h\right)(x)\,\lambda^2(dx),\quad (a,b)\in\mathfrak{R}\tag{2.44}$$

for an estimator of $\nu([a,\infty)\times[b,\infty))$. Note furthermore that (2.44) is only well-defined on

$$\widetilde{A}_{h,n}\overset{\text{def}}{=}\left\{\omega\in\Omega\,:\,\widehat{\varphi}_n(u)\neq 0,\quad\text{for all } u\in\left[-\frac{1}{h},\frac{1}{h}\right]^2\right\},\quad h>0,\quad n\in\mathbb{N}$$

because $\operatorname{supp}(\mathcal{F}K_h) \subseteq \left[-\frac{1}{h}, \frac{1}{h}\right]^2$ and $\widehat{\varphi}_n$ has to be non-zero on $\operatorname{supp}(\mathcal{F}K_h)$. At the same time, we have for $\omega \in \widetilde{A}_{h,n}$

$$\left(\frac{\partial^4 \widehat{\Psi}_n}{\partial u_1^4} + \frac{\partial^4 \widehat{\Psi}_n}{\partial u_2^4}\right)(\omega) \cdot \mathcal{F}K_h \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$$

since the left-hand side is a continuous function with compact support. Thus, the inverse Fourier transform in (2.44) is well-defined on $\widetilde{A}_{h,n}$. Set

$$A_{h,n} \stackrel{\text{def}}{=} \left\{ |\widehat{\varphi}_n(u)| > \frac{1}{2}|\varphi(u)|, \quad u \in \left[-\frac{1}{h}, \frac{1}{h}\right]^2 \right\} \subseteq \widetilde{A}_{h,n}, \quad h > 0, \quad n \in \mathbb{N}.$$

Based on the above discussion, finally set

$$\widehat{N}_n(a,b) \stackrel{\text{def}}{=} \begin{cases} \int_{\mathbb{R}^2} g_{a,b}(x) \mathcal{F}^{-1}\left(\left(\frac{\partial^4 \widehat{\Psi}_n}{\partial u_1^4} + \frac{\partial^4 \widehat{\Psi}_n}{\partial u_2^4}\right) \mathcal{F}K_h\right)(x)\, \lambda^2(dx), & \omega \in A_{h,n}, \\ 0, & \omega \in A_{h,n}^c, \end{cases}$$

for all $(a,b) \in \mathfrak{R}$, $h > 0$, $n \in \mathbb{N}$. Of course, the bandwidth $h = h_n$ has to be chosen in an optimal manner. It turns out that

$$h_n \stackrel{\text{def}}{=} \frac{\log\log n}{\sqrt{\log n}}, \quad n \in \mathbb{N}$$

yields a satisfying result:

**Theorem 2.2.11.** *It holds, under the Assumptions 2.2.1 and 2.2.3 (i)-(iii), the asymptotic*

$$\sup_{(a,b)\in\mathfrak{R}} |(a,b)|^2 \wedge |(a,b)|^4 \left|\nu([a,\infty) \times [b,\infty)) - \widehat{N}_n(a,b)\right| = O_{P_{\Sigma,\nu,\alpha}}\left(\frac{(\log\log n)^2}{\sqrt{\log n}}\right), \quad n \to \infty. \tag{2.45}$$

*Proof.* The proof is divided into three steps. The probability that the inverse Fourier transform is well defined tends to one. This is shown in the first step. The second step estimates the approximation error between $\widehat{\Psi}_n$ and $\Psi$. Finally, the third step uses these estimations together with the statement of Lemma 2.2.7 to prove the desired convergence rate.

STEP 1. First, we establish $P(A_n^c) \to 0$, $n \to \infty$ with $A_n \stackrel{\text{def}}{=} A_{h_n,n}$, $n \in \mathbb{N}$ and $P \stackrel{\text{def}}{=} P_{\Sigma,\nu,\alpha}$. Note for this that we have with $B_{\frac{1}{h}} \stackrel{\text{def}}{=} \left[-\frac{1}{h}, \frac{1}{h}\right]^2$ and $\varphi \stackrel{\text{def}}{=} \varphi_{\Sigma,\nu,\alpha}$ the inclusions

$$\begin{aligned} A_{h,n}^c = \left\{\exists u \in B_{\frac{1}{h}} : |\widehat{\varphi}_n(u)| \leqslant \frac{1}{2}|\varphi(u)|\right\} &\subseteq \left\{\exists u \in B_{\frac{1}{h}} : \frac{|\widehat{\varphi}_n(u) - \varphi(u)|}{|\varphi(u)|} \geqslant \frac{1}{2}\right\} \\ &\subseteq \left\{\exists u \in B_{\frac{1}{h}} : \frac{d(\widehat{\varphi}_n, \varphi)}{|\varphi(u)||w(u)|} \geqslant \frac{1}{2}\right\}. \end{aligned}$$

Observe

$$w(u) = (\log(e + |u|))^{-\frac{1}{2}-\delta} \geqslant e^{-\left(\frac{1}{2}+\delta\right)|u|}, \quad u \in \mathbb{R}^2,$$

so that together with Lemma 2.2.8 we obtain

$$|\varphi(u)||w(u)| \geqslant e^{-C(1+|u|)^2}, \quad u \in \mathbb{R}^2 \tag{2.46}$$

with a constant $C > 0$. This, finally, implies

$$\left\{\exists u \in B_{\frac{1}{h}} \ : \ \frac{d(\widehat{\varphi}_n, \varphi)}{|\varphi(u)||w(u)|} \geqslant \frac{1}{2}\right\} \subseteq \left\{d(\widehat{\varphi}_n, \varphi) e^{C\left(1+\frac{\sqrt{2}}{h}\right)^2} \geqslant \frac{1}{2}\right\},$$

and, the Markov inequality yields together with Corollary 2.1.7

$$P(A^c_{h,n}) \leqslant n^{-\frac{1}{2}} e^{C\left(1+\frac{\sqrt{2}}{h}\right)^2} O(1) \lesssim n^{-\frac{1}{2}} e^{\frac{4C}{h^2}} = e^{-\frac{1}{2}\log n + \frac{4C}{h^2}}, \quad n \in \mathbb{N}. \tag{2.47}$$

A substitution with $h_n = \frac{\log\log n}{\sqrt{\log n}}$ yields

$$-\frac{1}{2}\log n + \frac{4C}{h_n^2} = -\frac{1}{2}\log n + 4C\frac{\log n}{(\log\log n)^2} \to -\infty, \quad n \to \infty,$$

so that (2.47) implies $P(A^c_n) \to 0$, $n \to \infty$.

STEP 2. Next, we consider the difference

$$\frac{\partial^4 \widehat{\Psi}_n}{\partial u_k^4} - \frac{\partial^4 \Psi}{\partial u_k^4}, \quad k = 1, 2, \quad n \in \mathbb{N}.$$

(2.42) and (2.43) consist of respectively five terms. Subtracting (2.42) from (2.43), results in five difference terms. We rearrange for $k = 1, 2$ these terms in (2.48) - (2.52) for our needs:

$$\begin{aligned}&\frac{\partial^l \varphi}{\partial u_k^l}\varphi^{-1} - \frac{\partial^l \widehat{\varphi}_n}{\partial u_k^l}\widehat{\varphi}_n^{-1}, \quad l = 1,2,3,4 \\ = \ &\frac{\partial^l(\varphi - \widehat{\varphi}_n)}{\partial u_k^l}\widehat{\varphi}_n^{-1} + \frac{\partial^l \varphi}{\partial u_k^l}\varphi^{-1}(\widehat{\varphi}_n - \varphi)\widehat{\varphi}_n^{-1},\end{aligned} \tag{2.48}$$

$$\begin{aligned}&\frac{\partial \varphi}{\partial u_k}\frac{\partial^3 \varphi}{\partial u_k^3}\varphi^{-2} - \frac{\partial \widehat{\varphi}_n}{\partial u_k}\frac{\partial^3 \widehat{\varphi}_n}{\partial u_k^3}\widehat{\varphi}_n^{-2} \\ = \ &\frac{\partial \varphi}{\partial u_k}\varphi^{-1}\left(\frac{\partial^3 \varphi}{\partial u_k^3}\varphi^{-1} - \frac{\partial^3 \widehat{\varphi}_n}{\partial u_k^3}\widehat{\varphi}_n^{-1}\right) + \frac{\partial^3 \widehat{\varphi}_n}{\partial u_k^3}\widehat{\varphi}_n^{-1}\left(\frac{\partial \varphi}{\partial u_k}\varphi^{-1} - \frac{\widehat{\varphi}_n}{\partial u_k}\widehat{\varphi}_n^{-1}\right),\end{aligned} \tag{2.49}$$

$$\begin{aligned}&\left(\frac{\partial^2 \varphi}{\partial u_k^2}\varphi^{-1}\right)^2 - \left(\frac{\partial^2 \widehat{\varphi}_n}{\partial u_k^2}\widehat{\varphi}_n^{-1}\right)^2 \\ = \ &\left(\frac{\partial^2 \varphi}{\partial u_k^2}\varphi^{-1} + \frac{\partial^2 \widehat{\varphi}_n}{\partial u_k^2}\widehat{\varphi}_n^{-1}\right)\left(\frac{\partial^2 \varphi}{\partial u_k^2}\varphi^{-1} - \frac{\partial^2 \widehat{\varphi}_n}{\partial u_k^2}\widehat{\varphi}_n^{-1}\right),\end{aligned} \tag{2.50}$$

$$\begin{aligned}&\left(\frac{\partial \varphi}{\partial u_k}\right)^2 \frac{\partial^2 \varphi}{\partial u_k^2}\varphi^{-3} - \left(\frac{\partial \widehat{\varphi}_n}{\partial u_k}\right)^2 \frac{\partial^2 \widehat{\varphi}_n}{\partial u_k^2}\widehat{\varphi}_n^{-3} \\ = \ &\left(\frac{\partial \varphi}{\partial u_k}\right)^2 \varphi^{-2}\left(\frac{\partial^2 \varphi}{\partial u_k^2}\varphi^{-1} - \frac{\partial^2 \widehat{\varphi}_n}{\partial u_k^2}\widehat{\varphi}_n^{-1}\right) + \frac{\partial^2 \widehat{\varphi}_n}{\partial u_k^2}\widehat{\varphi}_n^{-1}\left(\left(\frac{\partial \varphi}{\partial u_k}\varphi^{-1}\right)^2 - \left(\frac{\partial \widehat{\varphi}_n}{\partial u_k}\widehat{\varphi}_n^{-1}\right)^2\right),\end{aligned} \tag{2.51}$$

$$
\begin{aligned}
&\left(\frac{\partial \varphi}{\partial u_k}\varphi^{-1}\right)^4 - \left(\frac{\partial \widehat{\varphi}_n}{\partial u_k}\widehat{\varphi}_n^{-1}\right)^4 \\
= \ & \left(\left(\frac{\partial \varphi}{\partial u_k}\varphi^{-1}\right)^2 + \left(\frac{\partial \widehat{\varphi}_n}{\partial u_k}\widehat{\varphi}_n^{-1}\right)^2\right)\left(\left(\frac{\partial \varphi}{\partial u_k}\varphi^{-1}\right)^2 - \left(\frac{\partial \widehat{\varphi}_n}{\partial u_k}\widehat{\varphi}_n^{-1}\right)^2\right).
\end{aligned} \tag{2.52}
$$

Next, after some straightforward calculations, observe

$$
\frac{\partial \varphi}{\partial u_k}\varphi^{-1} = \frac{\partial \Psi}{\partial u_k}, \tag{2.53}
$$

$$
\frac{\partial^2 \varphi}{\partial u_k^2}\varphi^{-1} = \left(\frac{\partial \Psi}{\partial u_k}\right)^2 + \frac{\partial^2 \Psi}{\partial u_k^2}, \tag{2.54}
$$

$$
\frac{\partial^3 \varphi}{\partial u_k^3}\varphi^{-1} = \left(\frac{\partial \Psi}{\partial u_k}\right)^3 + \frac{\partial^3 \Psi}{\partial u_k^3} + 3\frac{\partial \Psi}{\partial u_k}\frac{\partial^2 \Psi}{\partial u_k^2}, \tag{2.55}
$$

$$
\frac{\partial^4 \varphi}{\partial u_k^4}\varphi^{-1} = \left(\frac{\partial \Psi}{\partial u_k}\right)^4 + 3\left(\frac{\partial^2 \Psi}{\partial u_k^2}\right)^2 + 6\left(\frac{\partial \Psi}{\partial u_k}\right)^2\frac{\partial^2 \Psi}{\partial u_k^2} + 4\frac{\partial \Psi}{\partial u_k}\frac{\partial^3 \Psi}{\partial u_k^3} + \frac{\partial^4 \Psi}{\partial u_k^4}. \tag{2.56}
$$

In what follows, we estimate the derivatives of $\Psi$:

$$
\langle u, \Sigma u \rangle = \sigma_{11}u_1^2 + 2\sigma_{12}u_1u_2 + \sigma_{22}u_2^2, \quad u \in \mathbb{R}^2
$$

and the representation (2.40) yields

$$
\frac{\partial \Psi}{\partial u_1}(u) = -\frac{1}{2}(2u_1\sigma_{11} + 2u_2\sigma_{12}) + i\alpha_1 + \int_{\mathbb{R}^2}\left(ix_1e^{i\langle u,x\rangle} - ix_1\right)\nu(dx)
$$

which yields together with

$$
\left|e^{i\langle u,x\rangle} - 1\right| \leqslant |\langle u,x\rangle| \leqslant |u||x|, \quad u,x \in \mathbb{R}^2
$$

and $\int_{\mathbb{R}^2}|x|^2\,\nu(dx) < \infty$ the inequality

$$
\left|\frac{\partial \Psi}{\partial u_1}\right|(u) \lesssim 1 + |u|, \quad u \in \mathbb{R}^2
$$

where the constant in the $\lesssim$ sign depends only on the Lévy triplet $(\Sigma, \nu, \alpha)$. Similarly we get

$$
\left|\frac{\partial^2 \Psi}{\partial u_1^2}\right|(u) = \left|-\sigma_{11} - \int_{\mathbb{R}^2}x_1^2e^{i\langle u,x\rangle}\,\nu(dx)\right| \leqslant \sigma_{11} + \int_{\mathbb{R}^2}x_1^2\,\nu(dx) < \infty
$$

and

$$
\begin{aligned}
\left|\frac{\partial^3 \Psi}{\partial u_1^3}\right|(u) &= \left|-i\int_{\mathbb{R}^2}x_1^3e^{i\langle u,x\rangle}\,\nu(dx)\right| \leqslant \int_{\mathbb{R}^2}|x_1|^3\,\nu(dx) < \infty, \\
\left|\frac{\partial^4 \Psi}{\partial u_1^4}\right|(u) &= \left|\int_{\mathbb{R}^2}x_1^4e^{i\langle u,x\rangle}\,\nu(dx)\right| \leqslant \int_{\mathbb{R}^2}x_1^4\,\nu(dx) < \infty.
\end{aligned}
$$

The derivatives $\frac{\partial^l}{\partial u_2^l}$, $l = 1, 2, 3, 4$ yield analogous estimates. Hence, (2.53) - (2.56) imply

$$\left|\frac{\partial^l \varphi}{\partial u_k^l}\varphi^{-1}\right|(u) \lesssim (1+|u|)^l, \quad u \in \mathbb{R}^2, \quad l = 1, 2, 3, 4. \tag{2.57}$$

Additionally (2.57) yields together with (2.48) on $A_n$

$$\begin{aligned}\left|\frac{\partial^l \widehat{\varphi}_n}{\partial u_k^l}\widehat{\varphi}_n^{-1}\right|(u) &\leqslant \left|\frac{\partial^l \varphi}{\partial u_k^l}\varphi^{-1}\right|(u) + \left|\frac{\partial^l \varphi}{\partial u_k^l}\varphi^{-1} - \frac{\partial^l \widehat{\varphi}_n}{\partial u_k^l}\widehat{\varphi}_n^{-1}\right|(u) \\ &\lesssim \left(1 + \frac{d(\varphi, \widehat{\varphi}_n)}{w(u)|\varphi(u)|}\right)(1+|u|)^l, \quad u \in \left[-\frac{1}{h_n}, \frac{1}{h_n}\right]^2, \quad l = 1, 2, 3, 4.\end{aligned} \tag{2.58}$$

Hence, (2.42), (2.43) and (2.48) - (2.52) and (2.57), (2.58) finally yield on $A_n$

$$\left|\frac{\partial^4 \widehat{\Psi}_n}{\partial u_k^4} - \frac{\partial^4 \Psi}{\partial u_k^4}\right|(u) \lesssim \sum_{j=1}^4 \left(\frac{d(\varphi, \widehat{\varphi}_n)}{w(u)|\varphi(u)|}\right)^j (1+|u|)^4, \quad u \in \left[-\frac{1}{h_n}, \frac{1}{h_n}\right]^2.$$

STEP 3. Next, observe with Sato [38][Proposition 2.5 (xii)]

$$\mathcal{F}^{-1}\left(\left(\frac{\partial^4 \Psi}{\partial u_1^4} + \frac{\partial^4 \Psi}{\partial u_2^4}\right)\mathcal{F}K_{h_n}\right)\lambda^2 = (K_{h_n}\lambda^2) * ((x_1^4 + x_2^4)\nu).$$

Using the Plancherel identity, we get the following essential estimates:

$$\begin{aligned}&\mathbb{1}_{A_n}\left|\int_{\mathbb{R}^2} g_{a,b}(x)\mathcal{F}^{-1}\left(\left(\frac{\partial^4 \widehat{\Psi}_n}{\partial u_1^4} + \frac{\partial^4 \widehat{\Psi}_n}{\partial u_2^4}\right)\mathcal{F}K_{h_n}\right)(x)\,\lambda^2(dx)\right. \\ &\left.\quad - \int_{\mathbb{R}^2} g_{a,b}(x)\mathcal{F}^{-1}\left(\left(\frac{\partial^4 \Psi}{\partial u_1^4} + \frac{\partial^4 \Psi}{\partial u_2^4}\right)\mathcal{F}K_{h_n}\right)(x)\,\lambda^2(dx)\right| \\ &= \mathbb{1}_{A_n}\frac{1}{4\pi^2}\left|\int_{\mathbb{R}^2}(\mathcal{F}g_{a,b})(-u)\sum_{k=1}^2 \frac{\partial^4(\widehat{\Psi}_n - \Psi)}{\partial u_k^4}(u)\mathcal{F}K_{h_n}(u)\,\lambda^2(du)\right| \\ &\lesssim \mathbb{1}_{A_n}|(a,b)|^{-2}\int_{\left[-\frac{1}{h_n}, \frac{1}{h_n}\right]^2}\sum_{j=1}^4\left(\frac{d(\varphi, \widehat{\varphi}_n)}{w(u)|\varphi(u)|}\right)^j (1+|u|)^4\,\lambda^2(du).\end{aligned} \tag{2.59}$$

Note that

$$d(\varphi, \widehat{\varphi}_n) = O_P(n^{-\frac{1}{2}})$$

and

$$(w(u)|\varphi(u)|)^{-j}(1+|u|)^4 \lesssim e^{C(1+|u|)^2}, \quad j = 1, 2, 3, 4, \quad u \in \mathbb{R}^2$$

hold for suitable $C > 0$, compare (2.46). Hence, (2.59) is not larger than

$$\begin{aligned}&|(a,b)|^{-2}\int_{\left[-\frac{1}{h_n}, \frac{1}{h_n}\right]^2} e^{C(1+|u|)^2}\,\lambda^2(du)\cdot O_P(n^{-\frac{1}{2}}) \\ &= |(a,b)|^{-2}\int_0^{\frac{\sqrt{2}}{h_n}} e^{C(1+r)^2}(1+r)\,dr \cdot O_P(n^{-\frac{1}{2}}) = |(a,b)|^{-2}\left.\frac{e^{C(1+r)^2}}{2C}\right|_0^{\frac{\sqrt{2}}{h_n}} \cdot O_P(n^{-\frac{1}{2}})\end{aligned}$$

$$= \quad |(a,b)|^{-2} n^{-\frac{1}{2}} e^{\frac{4C}{h_n^2}} \cdot O_P(1) = |(a,b)|^{-2} e^{\left(\frac{4C}{(\log\log n)^2} - \frac{1}{2}\right)\log n} O_P(1) = |(a,b)|^{-2} O_P(n^{\epsilon - \frac{1}{2}})$$

for every $\epsilon > 0$. Together with Lemma 2.2.7 and

$$|h_n \log h_n| = -\frac{\log\log n}{\sqrt{\log n}}(\log\log\log n - \frac{1}{2}\log\log n) \lesssim \frac{(\log\log n)^2}{\sqrt{\log n}}$$

this proves this theorem. □

### The inverting operation

Considering the Lévy copula (2.2), our next goal is to establish an inversion operation. For this purpose, we first define some function spaces and an inversion operation $\mathcal{I}$ on those spaces: Set

$$\begin{aligned}
\widehat{\mathcal{C}} &\stackrel{\text{def}}{=} \{g : (0,\infty) \to \mathbb{R}_+, \quad g \in \mathcal{C}, \quad \lim_{x\to\infty} g(x) = 0\},\\
\mathcal{D}_\delta &\stackrel{\text{def}}{=} \{h : (0,\infty) \to \mathbb{R}_+, \quad h \text{ is càdlàg, decreasing and } \lim_{x\to\infty} h(x) = \delta\}, \quad \delta > 0,\\
\mathcal{D} &\stackrel{\text{def}}{=} \bigcup_{\delta>0} \mathcal{D}_\delta
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{I} : \quad \widehat{\mathcal{C}} \times (0,\infty) &\to \mathcal{D}\\
(g,\delta) &\mapsto (z \mapsto \inf\{x \geqslant \delta : \inf_{\delta \leqslant y \leqslant x} g(y) \leqslant z\}).
\end{aligned} \tag{2.60}$$

Furthermore, let $R : (0,\infty) \to (0,\infty)$ be a function and $(\epsilon_n)$, $(\delta_n)$ be two sequences of positive numbers, such that

$$\gamma_n \stackrel{\text{def}}{=} R(\delta_n)\epsilon_n \to 0, \quad \epsilon_n \downarrow 0, \quad \delta_n \downarrow 0, \quad n \to \infty$$

hold. Note that we have

$$\mathcal{I}(g,\delta) \in \mathcal{D}_\delta, \quad g \in \widehat{\mathcal{C}}, \quad \delta > 0.$$

The next Proposition 2.2.12 investigates the behaviour of an approximation error under the inversion operation $\mathcal{I}$. Note that $\mathcal{I}$ is the pseudo inverse with the starting position $\delta > 0$, cf. (2.60). The introduction of such an offset $\delta > 0$ is required for the subsequent treatment of the small jumps. Note that we have $\Lambda = \infty$ ($\Lambda$ in Proposition 2.2.12) in this section. The case $\Lambda < \infty$ is important for the investigations in the next Section 2.3 of the compound Poisson process case.

**Proposition 2.2.12.** *Let $f : (0,\infty) \to (0,\Lambda)$, $\Lambda \in (0,\infty]$ be a $\mathcal{C}^1$-bijection with*

$$f' < 0, \quad \inf_{0<x\leqslant 1} |f'(x)| > 0$$

*and let $(f_n)_n \subseteq \widehat{\mathcal{C}}$ be a family of functions, such that*

$$\sup_{x \geqslant \delta_n} |f_n(x) - f(x)| \leqslant \gamma_n, \quad n \in \mathbb{N}$$

*holds. Fix any $0 < a < b < \Lambda$. Then it holds also for each $n \in \mathbb{N}$ with $2\gamma_n < a \wedge (\Lambda - b)$ and $\delta_n < f^{-1}(b + 2\gamma_n)$ the inequality*

$$\sup_{a \leqslant z \leqslant b} |\mathcal{I}(f_n,\delta_n)(z) - f^{-1}(z)| \leqslant 2\gamma_n \left( \inf_{0<x\leqslant f^{-1}\left(\frac{a}{2}\right)} |f'(x)| \right)^{-1}.$$

*Proof.* Set

$$F_n(x) \overset{\text{def}}{=} \inf_{\delta_n \leqslant y \leqslant x} f_n(y), \quad h_n(z) \overset{\text{def}}{=} \mathcal{I}(f_n, \delta_n)(z), \quad x \geqslant \delta_n, \quad z > 0.$$

Note that $F_n : [\delta_n, \infty) \to \mathbb{R}_+$ is a decreasing, continuous function with $F_n(x) \to 0$, $x \to \infty$ for each $n \in \mathbb{N}$. First, we show the inequality

$$\sup_{x \geqslant \delta_n} |f(x) - F_n(x)| \leqslant \gamma_n, \quad n \in \mathbb{N}. \tag{2.61}$$

Note for this

$$F_n(x) = f_n(c_x) \leqslant f_n(x), \quad \delta_n \leqslant c_x \leqslant x$$

and

$$\begin{aligned} f(x) - F_n(x) &= f(x) - f_n(c_x) \leqslant f(c_x) - f_n(c_x) \leqslant \gamma_n, \\ f(x) - F_n(x) &\geqslant f(x) - f_n(x) \geqslant -\gamma_n \end{aligned}$$

for all $n \in \mathbb{N}$. Next, fix any $0 < a \leqslant z \leqslant b < \Lambda$ and $n \in \mathbb{N}$ with $2\gamma_n < a \wedge (\Lambda - b)$, $\delta_n < f^{-1}(b + 2\gamma_n)$. Set

$$y \overset{\text{def}}{=} z - \gamma_n > \frac{a}{2}, \quad y' \overset{\text{def}}{=} z + 2\gamma_n < \Lambda$$

and

$$x \overset{\text{def}}{=} f^{-1}(y) \geqslant x' \overset{\text{def}}{=} f^{-1}(y') \geqslant f^{-1}(b + 2\gamma_n) \geqslant \delta_n.$$

(2.61) implies $F_n(x) \leqslant f(x) + \gamma_n$ which yields, since $h_n$ is the pseudo-inverse of $F_n$,

$$h_n(z) = h_n(f(x) + \gamma_n) \leqslant x = f^{-1}(z - \gamma_n).$$

Equally, we have $F_n(x') \geqslant f(x') - \gamma_n > f(x') - 2\gamma_n$ which implies

$$h_n(z) = h_n(f(x') - 2\gamma_n) \geqslant x' = f^{-1}(z + 2\gamma_n),$$

so that, altogether we have

$$f^{-1}(z + 2\gamma_n) \leqslant h_n(z) \leqslant f^{-1}(z - \gamma_n).$$

Using the mean value theorem, this yields, on the one hand,

$$h_n(z) - f^{-1}(z) \leqslant f^{-1}(z - \gamma_n) - f^{-1}(z) = -\gamma_n (f^{-1})'(\xi_1)$$

and on the other hand

$$f^{-1}(z) - h_n(z) \leqslant f^{-1}(z) - f^{-1}(z + 2\gamma_n) = -2\gamma_n (f^{-1})'(\xi_2)$$

with

$$\xi_1, \xi_2 \in [z - \gamma_n, z + 2\gamma_n] \subseteq \left[\frac{a}{2}, \Lambda\right).$$

Thus, we finally obtain

$$|h_n(z) - f^{-1}(z)| \leqslant 2\gamma_n \sup_{\frac{a}{2} \leqslant y < \Lambda} |(f^{-1})'(y)| = 2\gamma_n \left( \inf_{0 < x \leqslant f^{-1}\left(\frac{a}{2}\right)} |f'(x)| \right)^{-1}.$$

□

Next, we state a stochastic version of Proposition 2.2.12, which is adapted to our later needs.

**Corollary 2.2.13.** *Given a probability space $(\Omega, \mathcal{F}, P)$ and a family of functions*

$$\widehat{Z}_n : \Omega \to \widehat{\mathcal{C}}, \quad n \in \mathbb{N},$$

*such that $\omega \mapsto [\widehat{Z}_n(\omega)](x)$ is $\mathcal{F}$-measurable for every $n \in \mathbb{N}$, $x > 0$ and such that*

$$\sup_{x \geqslant \delta_n} |\widehat{Z}_n(x) - f(x)| = O_P(\gamma_n), \quad n \to \infty \tag{2.62}$$

*holds with a function $f$ as in Proposition 2.2.12. Then it also holds for any fixed $0 < a < b < \Lambda$*

$$\sup_{a \leqslant z \leqslant b} |\mathcal{I}(\widehat{Z}_n, \delta_n)(z) - f^{-1}(z)| = O_P(\gamma_n), \quad n \to \infty.$$

*Proof.* Write $(\gamma_n \cdot X_n)_{n \in \mathbb{N}}$ instead of $O_P(\gamma_n)$ in (2.62), i.e. $(X_n)_n$ is a family of random variables, which are uniformly bounded in probability. Set, furthermore,

$$A_n \stackrel{\text{def}}{=} \left\{ 2\gamma_n X_n < a \wedge (\Lambda - b), \quad \delta_n < f^{-1}(b + 2\gamma_n X_n) \right\}, \quad n \in \mathbb{N}.$$

Then, Proposition 2.2.12 states that we have for $\omega \in A_n$

$$\sup_{a \leqslant z \leqslant b} |\mathcal{I}(\widehat{Z}_n(\omega), \delta_n)(z) - f^{-1}(z)| \leqslant 2\gamma_n X_n(\omega) \left( \inf_{0 < x \leqslant f^{-1}\left(\frac{a}{2}\right)} |f'(x)| \right)^{-1} \lesssim \gamma_n X_n(\omega).$$

This proves this Corollary since $P(A_n^c) \to 0$ for $n \to \infty$. □

Denote with $\Re_+(c) \stackrel{\text{def}}{=} \Re(c) \vee 0$ the positive real part of a complex number $c \in \mathbb{C}$. Finally, we combine the statements developed so far and get the following main result:

**Theorem 2.2.14.** *Let the Assumptions 2.2.1 and 2.2.3 hold and $0 < a < b < \infty$ be two fixed numbers. Set $\delta_n \stackrel{\text{def}}{=} (\log\log n)^{-1}$ and*

$$\widehat{U}_{1,n}^{-1} \stackrel{\text{def}}{=} \mathcal{I}(\Re_+ \widehat{N}_n(\cdot, 0), \delta_n), \quad \widehat{U}_{2,n}^{-1} \stackrel{\text{def}}{=} \mathcal{I}(\Re_+ \widehat{N}_n(0, \cdot), \delta_n), \quad n \in \mathbb{N}.$$

*Then, it holds with the plug-in estimator*

$$\widehat{\mathfrak{C}}_n(u, v) = \widehat{N}_n(\widehat{U}_{1,n}^{-1}(u), \widehat{U}_{2,n}^{-1}(v)), \quad u, v > 0$$

*the asymptotic*

$$\sup_{a \leqslant u, v \leqslant b} |\mathfrak{C}(u, v) - \widehat{\mathfrak{C}}_n(u, v)| = O_{P_{\Sigma,\nu,\alpha}}\left( \frac{(\log\log n)^9}{\sqrt{\log n}} \right), \quad n \to \infty.$$

*Proof.* First, note that we can replace $\widehat{N}_n$ by $\Re_+ \widehat{N}_n$ and (2.45) is still valid. This is due to the fact that we have for all $c \in \mathbb{C}$ and $r \in \mathbb{R}_+$ the inequality

$$|c - r| = \sqrt{(\Re(c - r))^2 + (\Im(c - r))^2} \geqslant |\Re(c - r)| = |\Re(c) - r| \geqslant |\Re_+(c) - r|.$$

Observe, furthermore,

$$\Re_+\widehat{N}_n(\cdot,0),\ \Re_+\widehat{N}_n(0,\cdot)\in\widehat{\mathcal{C}},\quad n\in\mathbb{N}.$$

Set

$$\epsilon_n \stackrel{\text{def}}{=} \frac{(\log\log n)^2}{\sqrt{\log n}},\quad R(x)\stackrel{\text{def}}{=} x^{-4},\quad x>0$$

and note that

$$\gamma_n = R(\delta_n)\epsilon_n = \frac{(\log\log n)^6}{\sqrt{\log n}}\to 0,\quad n\to\infty.$$

Theorem 2.2.11 states

$$\sup_{x\geqslant\delta_n}|\Re_+\widehat{N}_n(x,0)-U_1(x)| = O_{P_{\Sigma,\nu,\alpha}}(\gamma_n),\quad n\to\infty,$$

so that Corollary 2.2.13 implies with $\Lambda\stackrel{\text{def}}{=}\infty$, $\widehat{Z}_n\stackrel{\text{def}}{=}\Re_+\widehat{N}_n(\cdot,0)$ and Assumption 2.2.3 (iv)

$$\sup_{a\leqslant u\leqslant b}|\widehat{U}_{1,n}^{-1}(u)-U_1^{-1}(u)| = O_{P_{\Sigma,\nu,\alpha}}(\gamma_n),\quad n\to\infty. \tag{2.63}$$

Of course, exactly the same considerations yield the $U_2$ analogue of (2.63). Next, write

$$u_n\stackrel{\text{def}}{=} U_1\circ\widehat{U}_{1,n}^{-1}(u),\quad v_n\stackrel{\text{def}}{=} U_2\circ\widehat{U}_{2,n}^{-1}(v),\quad a\leqslant u,v\leqslant b$$

and note that the Lévy-copula $\mathfrak{C}$ is Lipschitz continuous, cf. Kallsen, Tankov [21][Lemma 3.2]. More precise, we have

$$|\mathfrak{C}(u,v)-\mathfrak{C}(u',v')|\leqslant|u-u'|+|v-v'|,\quad u,u',v,v'>0. \tag{2.64}$$

Together with the mean value theorem and $\widehat{U}_{1,n}^{-1}\in\mathcal{D}_{\delta_n}$ we have for $a\leqslant u\leqslant b$

$$|u-u_n| = |U_1\circ U_1^{-1}(u)-U_1\circ\widehat{U}_{1,n}^{-1}(u)| = |U_1^{-1}(u)-\widehat{U}_{1,n}^{-1}(u)||U_1'(x)|,\quad x\in[U_1^{-1}(b)\wedge\delta_n,\infty).$$

Thus, (2.63) yields together with Assumption 2.2.3 (iv)

$$\sup_{a\leqslant u\leqslant b}|u-u_n| = \delta_n^{-3}O_{P_{\Sigma,\nu,\alpha}}(\gamma_n),\quad n\to\infty. \tag{2.65}$$

Hence, (2.64) and (2.65) imply

$$\sup_{a\leqslant u,v\leqslant b}|\mathfrak{C}(u,v)-\mathfrak{C}(u_n,v_n)|\leqslant\sup_{a\leqslant u\leqslant b}|u-u_n|+\sup_{a\leqslant v\leqslant b}|v-v_n| = \delta_n^{-3}O_{P_{\Sigma,\nu,\alpha}}(\gamma_n),\quad n\to\infty. \tag{2.66}$$

Theorem 2.2.11 yields because of $\widehat{U}_{j,n}^{-1}\in\mathcal{D}_{\delta_n}$, $j=1,2$ the asymptotic

$$\begin{aligned}\sup_{a\leqslant u,v\leqslant b}|\mathfrak{C}(u_n,v_n)-\widehat{\mathfrak{C}}_n(u,v)| &= \sup_{a\leqslant u,v\leqslant b}|U(\widehat{U}_{1,n}^{-1}(u),\widehat{U}_{2,n}^{-1}(v))-\widehat{N}_n(\widehat{U}_{1,n}^{-1}(u),\widehat{U}_{2,n}^{-1}(v))|\\ &= O_{P_{\Sigma,\nu,\alpha}}(\gamma_n),\quad n\to\infty. \end{aligned}\tag{2.67}$$

Finally, (2.66) and (2.67) prove this theorem. □

## 2.3 The Compound Poisson Process Case

Note that we do not use the special shape of the weight function

$$w(u) = (\log(e + |u|))^{-\frac{1}{2}-\delta}, \quad \delta > 0, \quad u \in \mathbb{R}^2$$

in the previous Section 2.2. Neither do we use the convergence rate $n^{-\frac{1}{2}}$ obtained in Theorem 2.1.6. In fact, the proofs in the previous section also work if we only had

$$e^{-C(1+|u|)^2} \lesssim w(u), \quad u \in \mathbb{R}^2$$

for some constant $C > 0$ and, concerning Theorem 2.1.6,

$$n^{-\frac{1}{2}+\epsilon} \sup_{n \geqslant 1} E_{\Sigma,\nu,\alpha} \left\| \frac{\partial^l}{\partial u_k^l} A_n(u) \right\|_{L^\infty(w)} < \infty, \quad k = 1, 2, \quad l = 0, 1, 2, 3, 4$$

for some $\epsilon > 0$. This is due to the fast decay behaviour of $\varphi_{\Sigma,\nu,\alpha}$ if $\Sigma \neq 0$, cf. Lemma 2.2.8. Therefore we cannot derive any benefit from these stronger results. However, if $\varphi_{\Sigma,\nu,\alpha}$ decays more slowly, we can benefit from these stronger results as we shall demonstrate in the case of a compound Poisson process with drift. This is, in some sense, the complementary case of the one we investigated in the previous section.

**Assumptions 2.3.1.** *We state here the assumptions concerning $\nu$ in the compound Poisson case:*

*(i) The corresponding Lévy process is a compound Poisson process with intensity $0 < \Lambda < \infty$ and has only positive jumps, i.e.*

$$0 < \Lambda \overset{\text{def}}{=} \nu(\mathbb{R}^2) = \nu(\mathbb{R}_+^2) < \infty,$$

*(ii) $\exists \gamma > 0 : \int |x|^{8+\gamma} \nu(dx) < \infty$, i.e. finite $8+\gamma$-th moment,*

*(iii) $\mathcal{F}((x_1^4 + x_2^4)\nu)(u) \lesssim (1 + |u_1|)^{-1}(1 + |u_2|)^{-1}, \quad u \in \mathbb{R}^2$,*

*(iv) $U_k : (0, \infty) \to (0, \Lambda)$ is a $\mathcal{C}^1$-bijection with $U_k' < 0$ and*

$$\inf_{0 < x_k \leqslant 1} |U_k'(x_k)| > 0, \quad \sup_{x_k > 0} (1 \wedge x_k)|U_k'(x_k)| < \infty, \quad k = 1, 2.$$

**Proposition 2.3.2.** *Let $f : \mathbb{R}_+^2 \to \mathbb{R}_+$ be a continuous function with the properties*

*(i) $f(x) > 0, \quad x \in (\{0\} \times \mathbb{R}_+^*) \times (\mathbb{R}_+^* \times \{0\})$,*

*(ii) $f(x) \lesssim |x|^2(\log|x|)^{-2}, \quad x \in \mathbb{R}_+^2 : |x| \leqslant \frac{1}{2}$,*

*(iii) $f(x) \lesssim (1 + |x|)^{-(6+\epsilon)}, \quad x \in \mathbb{R}_+^2$*

*for some $\epsilon > 0$. Then*

$$\nu(dx) \overset{\text{def}}{=} \mathbb{1}_{\mathfrak{R}}(x)(x_1^4 + x_2^4)^{-1} f(x) \lambda^2(dx)$$

*is a Lévy measure and fulfills the Assumptions 2.3.1 (i), (ii) and (iv).*

*Proof.* We only highlight the deviations from the proof in Proposition 2.2.5:

(i) We have

$$0 < \Lambda = \nu(\mathbb{R}_+^2) \lesssim \int_{\mathbb{R}_+^2} |x|^{-4} f(x)\,\lambda^2(dx) \lesssim \int_0^{\frac{1}{2}} r^{-4} r^2 (\log r)^{-2} r\,dr + \int_{\frac{1}{2}}^{\infty} r^{-4} r\,dr < \infty.$$

(iv) Observe first

$$[x_1, \infty) \times \mathbb{R}_+ \uparrow \mathbb{R}_+^* \times \mathbb{R}_+, \quad x_1 \downarrow 0,$$

so that we obtain

$$U_1(x_1) = \nu([x_1, \infty) \times \mathbb{R}_+) \to \nu(\mathbb{R}_+^* \times \mathbb{R}_+) = \nu(\mathbb{R}_+^2) = \Lambda.$$

This yields that $U_1 : (0, \infty) \to (0, \Lambda)$ is a surjection. Compare for this the argumentation in Proposition 2.2.5. Finally, note that it holds for $x_1 > 0$

$$|U_1'(x_1)| = \int_0^\infty (x_1^4 + y_2^4)^{-1} f(x_1, y_2)\,dy_2 \lesssim \int_0^1 (x_1 + y_2)^{-2}\,dy_2 + \int_1^\infty y_2^{-4}\,dy_2 = \frac{1}{x_1} - \frac{1}{x_1 + 1} + \frac{1}{3}$$

which implies

$$\sup_{x_1 > 0} (1 \wedge x_1)|U_1'(x_1)| < \infty.$$

□

**Corollary 2.3.3.** *It exists a Lévy measure $\nu_0$ that fulfills the Assumptions 2.3.1 with the property*

$$\int_{\mathbb{R}^2} |x|^{-\epsilon} \nu_0(dx) = \infty$$

*for all $\epsilon > 0$.*

*Proof.* We imitate the proof of step 2 in Corollary 2.2.6. For this purpose, set

$$f_0(x) \stackrel{\text{def}}{=} \psi(r)(\log r)^{-2} + (1 - \phi(r))e^{-r}, \quad r \stackrel{\text{def}}{=} |x|, \quad x \in \mathfrak{R}$$

and

$$\nu_0(dx) \stackrel{\text{def}}{=} \mathbb{1}_{\mathfrak{R}}(x)(x_1^4 + x_2^4)^{-1} f_0(x)\,\lambda^2(dx)$$

with $\phi$ as in the proof of Corollary 2.2.6 and

$$\psi(x) \stackrel{\text{def}}{=} |x|^2 \phi(x) = r^2 \phi(r), \quad x \in \mathfrak{R}.$$

Then Assumptions 2.3.1 (i), (ii) and (iv) are fulfilled because of Proposition 2.3.2. Furthermore, we have

$$\int_{\mathbb{R}^2} |x|^{-\epsilon}\,\nu_0(dx) \geqslant \int_0^{\frac{1}{2}} r^{-\epsilon} r^{-4} r^2 (\log r)^{-2} r\,dr \gtrsim \int_0^{\frac{1}{2}} r^{-1-\frac{\epsilon}{2}}\,dr = \infty$$

for all $\epsilon > 0$. Concerning Assumption 2.3.1 (iii), note that it holds for $r \geqslant 0$

$$\begin{aligned} \psi'(r) &= 2r\phi(r) + r^2\phi'(r), \\ \psi''(r) &= 2\phi(r) + 4r\phi'(r) + r^2\phi''(r), \end{aligned}$$

i.e. $\|\psi\|_\infty < \infty$, $\|\psi'\|_\infty < \infty$ and $\|\psi''\|_\infty < \infty$. The remaining proof works exactly as step 2 in the proof of Corollary 2.2.6. □

Note that Assumption 2.3.1 (iv) implies that the one-dimensional compound Poisson coordinate processes also have the intensity $\Lambda$. Furthermore, we have the representation

$$\varphi_{\nu,\alpha}(u) = \exp\left(i\,\langle u,\alpha\rangle + \int_{\mathbb{R}^2} (e^{i\langle u,x\rangle} - 1)\,\nu(dx)\right), \quad u \in \mathbb{R}^2$$

with a finite measure $\nu$ and $\alpha \in \mathbb{R}^2$. Lemma 2.2.8 now turns into the following statement:

**Lemma 2.3.4.** *Given a Lévy triplet $(\Sigma,\nu,\alpha)$ with $\Sigma = 0$ and $\nu(\mathbb{R}^2) < \infty$. Then it holds*

$$\inf_{u\in\mathbb{R}^2} |\varphi_{\nu,\alpha}(u)| > 0.$$

*Proof.* We have

$$|\varphi_{\nu,\alpha}(u)| = \left|\exp\left(\int_{\mathbb{R}^2}\left(e^{i\langle u,x\rangle} - 1\right)\,\nu(dx)\right)\right| \geqslant \exp(-2\nu(\mathbb{R}^2)) > 0.$$

□

Using this, we can prove the following theorem with the same technique as Theorem 2.2.11. Set $h_n \overset{\text{def}}{=} n^{-\frac{1}{2}}$.

**Theorem 2.3.5.** *It holds under the Assumptions 2.2.1 and 2.3.1 (i)-(iii) the asymptotic*

$$\sup_{(a,b)\in\mathfrak{R}} |(a,b)|^2 \wedge |(a,b)|^4 |\nu([a,\infty)\times[b,\infty)) - \widehat{N}_n(a,b)| = O_{P_{\nu,\alpha}}\left(\frac{(\log n)^5}{\sqrt{n}}\right), \quad n\to\infty.$$

*Proof.* We only note the changes in the proof of Theorem 2.2.11 :

First, we establish

$$P(A_n^c) = P(A_{h_n,n}^c) = P\left(A^c_{\frac{1}{\sqrt{n}},n}\right) \to 0, \quad n\to\infty.$$

Note

$$A_n^c \subseteq \left\{\exists u \in B_{\frac{1}{h_n}} \,:\, \frac{d(\widehat{\varphi}_n,\varphi)}{|\varphi(u)||w(u)|} \geqslant \frac{1}{2}\right\} \subseteq \left\{n^{-\frac{1}{2}}(\log(e+\sqrt{2}n^{\frac{1}{2}}))^{\frac{1}{2}+\delta} O_P(1) \geqslant 1\right\}, \quad n\in\mathbb{N},$$

so that $n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}+\delta} \to 0$ yields $P(A_n^c) \to 0$. Note furthermore

$$\left|\frac{\partial\Psi}{\partial u_k}\right|(u) \lesssim 1, \quad u\in\mathbb{R}^2, \quad k = 1,2,$$

since

$$\int_{\mathbb{R}^2} |x_k|\,\nu(dx) < \infty, \quad k=1,2, \quad \text{cf. Assumptions 2.3.1 (i), (ii).}$$

This is why (2.59) can be replaced by

$$\begin{aligned} & \mathbb{1}_{A_n} \int_{\left[-\frac{1}{h_n}, \frac{1}{h_n}\right]^2} (\log(e + |u|))^{2+4\delta} |\mathcal{F} g_{a,b}(-u)| \, \lambda^2(u) O_P(n^{-\frac{1}{2}}) \\ = \; & \mathbb{1}_{A_n} \, |(a,b)|^{-2} \vee |(a,b)|^{-4} \\ & \left( 1 + \int_{\left[-\frac{1}{h_n}, \frac{1}{h_n}\right]^2} (\log(e+|u|))^{2+4\delta} (1+|u_1|)^{-1} (1+|u_2|)^{-1} \, \lambda^2(du) \right) O_P(n^{-\frac{1}{2}}). \quad (2.68) \end{aligned}$$

Set w.l.o.g. $\delta \stackrel{\text{def}}{=} \frac{1}{4}$. Then, (2.68) is not larger than

$$\begin{aligned} & \mathbb{1}_{A_n} \, |(a,b)|^{-2} \vee |(a,b)|^{-4} (\log n)^3 \left( \int_0^{\frac{1}{h_n}} (1+x)^{-1} \, \lambda^1(dx) \right)^2 O_P(n^{-\frac{1}{2}}) \\ = \; & \mathbb{1}_{A_n} \, |(a,b)|^{-2} \vee |(a,b)|^{-4} (\log n)^5 O_P(n^{-\frac{1}{2}}). \end{aligned}$$

Again this proves together with Lemma 2.2.7 and $|h_n \log h_n| = \frac{\log n}{2\sqrt{n}}$ this theorem. □

Note that we are interested in estimating the copula $C$ of $(\nu(\mathbb{R}^2))^{-1}\nu$ instead of the Lévy copula $\mathfrak{C}$ of $\nu$. Here, we do not need the principle of a Lévy copula because $\nu$ has no singularity in the origin. Nevertheless, we still treat the origin with our technique as a singularity point. This is due to the fact that we have originally developed this technique for the setting of the previous Section 2.2. However, it is also possible to get some considerable results in case of the compound Poisson with this technique without much extra effort.

**Definition 2.3.6.** *Let Assumptions 2.2.1 and 2.3.1 hold. Set with the same notation as in Theorem 2.2.14, but* $\delta_n \stackrel{\text{def}}{=} (\log n)^{-1}$

$$\begin{aligned} V_1(x_1) &\stackrel{\text{def}}{=} \Lambda^{-1}\nu([0,x_1] \times \mathbb{R}_+), & V_2(x_2) &\stackrel{\text{def}}{=} \Lambda^{-1}\nu(\mathbb{R}_+ \times [0,x_2]), & x_1, x_2 > 0, \\ \widehat{V}_{1,n}^{-1}(u) &\stackrel{\text{def}}{=} \widehat{U}_{1,n}^{-1}(\Lambda(1-u)), & \widehat{V}_{2,n}^{-1}(v) &\stackrel{\text{def}}{=} \widehat{U}_{2,n}^{-1}(\Lambda(1-v)), & 0 < u, v < 1 \end{aligned}$$

*and*

$$\begin{aligned} \widehat{M}_n(a,b) &\stackrel{\text{def}}{=} 1 + \Lambda^{-1}(\widehat{N}_n(a,b) - \widehat{N}_n(a,0) - \widehat{N}_n(0,b)), \quad (a,b) \in \mathfrak{R}, \\ \widehat{C}_n(u,v) &\stackrel{\text{def}}{=} \widehat{M}_n(\widehat{V}_{1,n}^{-1}(u), \widehat{V}_{2,n}^{-1}(v)), \quad 0 < u, v < 1. \end{aligned}$$

*Let furthermore $C$ denote the unique copula of the probability measure $\Lambda^{-1}\nu$, i.e.*

$$C(u,v) = M(V_1^{-1}(u), V_2^{-1}(v)), \quad 0 < u, v < 1$$

*with*

$$M(a,b) \stackrel{\text{def}}{=} \Lambda^{-1}\nu([0,a] \times [0,b]) = 1 + \Lambda^{-1}(U(a,b) - U(a,0) - U(0,b)), \quad (a,b) \in \mathfrak{R}.$$

Note that $V_k : (0,\infty) \to (0,1)$, $k = 1,2$ is a bijection and that its inverse is

$$V_k^{-1}(u) = U_k^{-1}(\Lambda(1-u)), \quad u \in (0,1).$$

**Theorem 2.3.7.** *Set* $\delta_n \stackrel{\text{def}}{=} (\log n)^{-1}$ *and let the Assumptions 2.2.1 and 2.3.1 hold. Then we have* *for arbitrary and fixed* $0 < a < b < 1$ *the asymptotic*

$$\sup_{a\leqslant u,v\leqslant b} |C(u,v) - \widehat{C}_n(u,v)| = O_{P_{\nu,\alpha}}\left(\frac{(\log n)^{10}}{\sqrt{n}}\right), \quad n \to \infty.$$

*Proof.* We imitate in the following the proof of Theorem 2.2.14. First Theorem 2.3.5 yields with

$$\epsilon_n \overset{\text{def}}{=} \frac{(\log n)^5}{\sqrt{n}}, \quad R(x) \overset{\text{def}}{=} x^{-4}, \quad x > 0$$

and the notation

$$\gamma_n \overset{\text{def}}{=} R(\delta_n)\epsilon_n = \frac{(\log n)^9}{\sqrt{n}} \to 0, \quad n \to \infty$$

the asymptotic

$$\sup_{x\geqslant\delta_n} |\Re_+ \widehat{N}_n(x,0) - U_1(x)| = O_{P_{\nu,\alpha}}(\gamma_n), \quad n \to \infty.$$

Next Corollary 2.2.13 implies, since the Assumption 2.3.1 (iv) holds

$$\sup_{a\leqslant u\leqslant b} |\widehat{V}_{1,n}^{-1}(u) - V_1^{-1}(u)| = \sup_{\Lambda(1-b)\leqslant u\leqslant\Lambda(1-a)} |\widehat{U}_{1,n}^{-1}(u) - U_1^{-1}(u)| = O_{P_{\nu,\alpha}}(\gamma_n), \quad n \to \infty. \tag{2.69}$$

Again as in Theorem 2.2.14, set

$$u_n \overset{\text{def}}{=} V_1 \circ \widehat{V}_{1,n}^{-1}(u), \quad v_n \overset{\text{def}}{=} V_2 \circ \widehat{V}_{2,n}^{-1}(v), \quad a \leqslant u, v \leqslant b$$

and note that a copula also is Lipschitz continuous, cf. Nelsen [28][Theorem 2.2.4.]. In particular, we have

$$|C(u,v) - C(u',v')| \leqslant |u - u'| + |v - v'|, \quad 0 < u, u', v, v' < 1. \tag{2.70}$$

Together with the mean value theorem and

$$\widehat{U}_{1,n}^{-1} \in \mathcal{D}_{\delta_n} \implies \inf_{0<u<1} \widehat{V}_{1,n}^{-1}(u) \geqslant \inf_{u>0} \widehat{U}_{1,n}^{-1}(u) = \delta_n, \tag{2.71}$$

we have for $a \leqslant u \leqslant b$

$$|u - u_n| = |V_1 \circ V_1^{-1}(u) - V_1 \circ \widehat{V}_{1,n}^{-1}(u)| = |V_1^{-1}(u) - \widehat{V}_{1,n}^{-1}(u)||V_1'(x)|, \quad x \in [V_1^{-1}(a) \wedge \delta_n, \infty).$$

So (2.69) yields together with the Assumption 2.3.1 (iv)

$$\sup_{a\leqslant u\leqslant b} |u - u_n| = \delta_n^{-1} O_{P_{\nu,\alpha}}(\gamma_n), \quad n \to \infty. \tag{2.72}$$

Hence (2.70) and (2.72) imply

$$\sup_{a\leqslant u,v\leqslant b} |C(u,v) - C(u_n,v_n)| \leqslant \sup_{a\leqslant u\leqslant b} |u - u_n| + \sup_{a\leqslant v\leqslant b} |v - v_n| = \delta_n^{-1} O_{P_{\nu,\alpha}}(\gamma_n), \quad n \to \infty. \tag{2.73}$$

Furthermore, Theorem 2.3.5 yields together with (2.71)

$$\begin{aligned} \sup_{a\leqslant u,v\leqslant b} |C(u_n,v_n) - \widehat{C}_n(u,v)| &= \sup_{a\leqslant u,v\leqslant b} |M(\widehat{V}_{1,n}^{-1}(u), \widehat{V}_{2,n}^{-1}(v)) - \widehat{M}_n(\widehat{V}_{1,n}^{-1}(u), \widehat{V}_{2,n}^{-1}(v))| \\ &= O_{P_{\nu,\alpha}}(\gamma_n), \quad n \to \infty. \end{aligned} \tag{2.74}$$

Finally, (2.73) and (2.74) prove this theorem. □

# Chapter 3

# Copula relations in compound Poisson processes (CPP)

Let $(N_t)_{t \geqslant 0}$ be a Poisson process on a probability space $(\Omega, \mathcal{F}, P)$ with intensity $\lambda > 0$ and let

$$X_j : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}^d)), \quad j \in \mathbb{N}$$

be a sequence of i.i.d. random variables, such that $(X_j)_{j \in \mathbb{N}}$ and $(N_t)_{t \geqslant 0}$ are independent. Set $F \overset{\text{def}}{=} P^{X_1}$ and

$$Y_t \overset{\text{def}}{=} \sum_{j=1}^{N_t} X_j, \quad t \geqslant 0. \tag{3.1}$$

Then $Y$ is a Lévy process with Lévy triplet $(\Sigma, \nu, \alpha) = (0, \lambda F, 0)$, i.e. a compound Poisson process (CPP). In Chapter 2, Section 3, we constructed for $d = 2$ an estimator $\widehat{C}_n$ for the jump distribution $F$ was based on the $n+1$ low frequency observations $(Y_t(\widehat{\omega}))_{t=0,1,\ldots,n}$, where $\widehat{\omega}$ denotes our observed path. Our approach in Chapter 2 based on the Lévy triplet estimation techniques by Neumann, Reiß [29] and further techniques motivated by Nickl, Reiß [30]. In the current Chapter 3, we explicitly treat compound Poisson processes (CPP), i.e. $(\Sigma, \nu, \alpha) = (0, \nu, 0)$, $\lambda = \nu(\mathbb{R}^d) < \infty$. In this case, the paper of Buchmann, Grübel [7] also offers a possibility of estimating the Lévy triplet under the same low frequency observation scheme for $d = 1$. The techniques in [7] do not, in contrast to [29, 30], use any Fourier inversion operations, but are based on a direct deconvolution approach. To get more acquainted with their approach, we quote Lemma 7 in [7]:

**Lemma 7 of Buchmann, Grübel [7].** *Let $F$ and $G$ be probability distributions on $\mathbb{R}_+$ with $\int_{(0,\infty)} e^{-\tau y}\, G(dy) < e^{-\lambda}$ for some $\lambda, \tau > 0$ and*

$$G \overset{\text{def}}{=} \Psi(\lambda, F) \overset{\text{def}}{=} e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} F^{*j}.$$

*Then it holds*

$$F = \Phi(\lambda, G) \overset{\text{def}}{=} \sum_{j=1}^{\infty} \frac{(-1)^{j+1} e^{\lambda j}}{\lambda j} G^{*j}.$$

The convergence of the right-hand sum holds in some suitable Banach space $D(\tau)$ introduced in detail in Buchmann, Grübel [7].

Note that (3.1) implies the relation $P^{Y_1} = \Psi(\lambda, F)$. Under this point of view, $\Phi$ in the quoted

Lemma 7 regains the jump distribution $F$ out of $\Psi(\lambda, F)$. Schicks [39] generalizes the paper of Buchmann and Grübel [7] to the multidimensional case and estimates the copula $C_F$ of the jump distribution $F$ with this approach. As before, everything in this chapter is stated in the two dimensional case $d = 2$ which again is only due to a simpler notation.

Given the assumption that the intensity $\lambda > 0$ is known, Schicks proposes in [39] for a jump distribution $F$ with no negative jumps, i.e. $F([0,\infty)^2) = 1$, the following estimator for the copula $C_F$:

$$\widehat{C}_n^{BG}(u,v) \stackrel{\text{def}}{=} \Phi[\lambda, G_n(G_{1,n}^{-1} \circ E_\lambda, G_{2,n}^{-1} \circ E_\lambda)](u,v), \quad 0 \leqslant u, v \leqslant 1, \quad n \in \mathbb{N}, \tag{3.2}$$

where the index $BG$ stands for Buchmann, Grübel. Note that $\Phi$ in (3.2) is applied to a two dimensional distribution. It can be shown in a straightforward consideration that a multidimensional analogue of Lemma 7 also holds, cf. Schicks [39]. Nevertheless, it is unnecessary to indicate this in our notation, i.e. we again simply write $\Phi$. We proceed in the same way with $\Psi$. Further, we use the notations

$$\begin{aligned} G_{j,n} &= \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{[0,x]}(Y_{j,k} - Y_{j,k-1}), \quad x \geqslant 0, \quad j = 1,2, \\ G_n(x,y) &= \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{[0,x]\times[0,y]}(Y_k - Y_{k-1}), \quad x, y \geqslant 0, \\ E_\lambda &= \Psi(\lambda, U[0,1]), \end{aligned}$$

where $Y_{j,k}$, $j = 1,2$ denotes the coordinates of $Y_k$ and $U[0,1]$ denotes the uniform distribution on $[0,1]$. Finally, $G_{j,n}^{-1}$, $j = 1,2$ denote the respective pseudo inverse. (3.2) implies that it holds in the limit $n \to \infty$ the identity

$$C_F = \Phi(\lambda, C_G(E_\lambda, E_\lambda)) \iff \Psi(\lambda, C_F) = C_G(E_\lambda, E_\lambda),$$

if $\widehat{C}_n^{BG}$ is close to $C_F$ in some sense. This is equivalent to

$$C_{\Psi(\lambda, C_F)} = C_{\Psi(\lambda, F)}. \tag{3.3}$$

Thus, the estimator in (3.2) would only be correct under this heuristic point of view, if (3.3) were true. For a better point of view, (3.3) is re-formulated by using the underlying compound Poisson process: Let

$$Y_t = \sum_{j=1}^{N_t} X_j, \quad B_t = \sum_{j=1}^{N_t} A_j$$

be two compound Poisson processes. Then (3.3) is true if, and only if,

$$C_{X_1} = C_{A_1} \implies C_{Y_1} = C_{B_1}$$

holds. As we shall see from of the results in this chapter, (3.3) is in general false.

Nevertheless, this provides an interesting starting point to investigate the correlation of the dependence structure (via copulas) between the components of $(Y_t)$ and the dependence structure of the underlying jump distribution, i.e. $C_{X_1}$. Concerning this, we prove some asymptotic results. This chapter is divided into three sections:

*Section 1* simply states some useful definitions for our needs. If $F$ is a two dimensional distribution with continuous margins, i.e. $F \in \mathcal{M}^c$, we denote with $\mathcal{T}F$ its unique copula, cf. Proposition 3.1.1. Note furthermore that we use, in this chapter, the notation $\mathcal{C}$ for the set of all copulas and *not* for the set of all continuous functions as in Chapter 2.

In *Section 2*, we consider the copula of a compound Poisson process $Y$ under the asymptotic $\lambda t \to \infty$, i.e. we consider the limit behaviour of

$$\mathcal{T}P^{Y_t} = \mathcal{T}\Psi(\lambda t, P^{X_1}), \quad \lambda t \to \infty. \tag{3.4}$$

Obviously, (3.4) implies that we can fix w.l.o.g. $t \overset{\text{def}}{=} 1$ and consider only the intensity limit $\lambda \to \infty$. In this context, Theorem 3.2.5 yields the convergence

$$\mathcal{T}\Psi(\lambda, F) \to \mathcal{T}N(0, \Sigma), \quad \lambda \to \infty$$

which is uniform on $[0,1]^2$. Here, $\Sigma \in \mathbb{R}^{2\times 2}$ denotes a positive definite matrix defined in Theorem 3.2.5. Hence, it follows that the copula of a CPP uniformly converges under this asymptotic to the Gaussian copula $\mathcal{T}N(0,\Sigma)$. To distinguish between the Gaussian limit copulas, we have to investigate whether $\mathcal{T}N(0,\Sigma) = \mathcal{T}N(0,\Sigma')$ holds for two positive definite matrices $\Sigma, \Sigma' \in \mathbb{R}^{2\times 2}$. This is done in Proposition 3.2.4: With the notations in Section 1 the entity

$$\rho(P^{X_1}) = \rho(F) = \frac{\int xy \, dF(x,y)}{\sqrt{\int x^2 \, dF(x,y) \int y^2 \, dF(x,y)}}$$

determines the limit copula of the CPP with jump distribution $F = P^{X_1}$. Using this asymptotic approach, the statement of Corollary 3.2.6 implies that (3.3) is in general not true, compare Remark 3.2.7.

Finally, in *Section 3*, we analyse the resulting limit copulas of all compound Poisson processes in a certain way. For this purpose, we investigate the map $F \mapsto \rho(F)$. First, note that Cauchy Schwarz yields for every $F$ the inequalities $-1 \leqslant \rho(F) \leqslant 1$. Note that $\rho$ can be geometrically interpreted as the cosinus of the two coordinates of $X_1$ in the Hilbert space $L^2(\Omega, \mathcal{F}, P)$. The first interesting question is whether

$$\rho(\mathcal{C}) \overset{?}{=} \rho(\mathcal{M}^c_+) = (0,1]$$

holds where $\mathcal{M}^c_+$ denotes all two dimensional distributions $F$ with $F([0,\infty)^2) = 1$ and continuous margins. To say it in prose: The question is, whether the set of jump distributions which consists of the set of copulas $\mathcal{C}$, can generate every limit copula, which belongs to a CPP with positive jumps. Proposition 3.3.1 states that this is not the case because it holds

$$\rho(\mathcal{C}) \subseteq \left[\frac{1}{2}, 1\right].$$

Thus, from the above geometric point of view, copulas always span an angle between 0 and 60 degrees. Additionally, Example 3.3.2 states that all limit copulas that are reachable by a copula jump distribution are even obtained by a Clayton copula, i.e.

$$\rho(\{C_\theta : \theta \in [-1,\infty)\backslash\{0\}\} \cup \{\Pi\} \cup \{M\}) = \left[\frac{1}{2}, 1\right],$$

see Section 1 for the notation. Finally, Example 3.3.3 provides the answer to the question of how to obtain the remaining limit copulas which belong to $\rho \in \left(0, \frac{1}{2}\right)$. In this example, we describe a constructive procedure how to construct such jump distributions: Fix any $0 < \epsilon < 1$. Simulate two independent $U[0,1]$ distributed random variables $U$ and $V$. If $|U - V| \geqslant \epsilon$, make the jump $(U, V) \in \mathbb{R}^2$. Else repeat this procedure until the difference between $U$ and $V$ is not less that $\epsilon$, and make afterwards the jump $(U, V) \in \mathbb{R}^2$. All necessary repetitions are performed independently from each other. Then, if $\epsilon$ runs through the interval $(0, 1)$, we get a set of corresponding $\rho$ values that includes $\left(0, \frac{1}{2}\right)$. Note that the resulting jumps of the above procedure are all positive.

## 3.1 Basic definitions

We denote with $\mathcal{M}$ the set of all probability measures on $(\mathbb{R}^2, \mathcal{B}^2)$. Here $\mathcal{B}^2$ are the Borel sets of $\mathbb{R}^2$. Furthermore write

$$F \in \mathcal{M}^c :\Longleftrightarrow F \in \mathcal{M} \text{ and } F(\{x\} \times \mathbb{R}) = F(\mathbb{R} \times \{x\}) = 0, \quad x \in \mathbb{R},$$

i.e. the case that the margins of $F$ are both continuous. If we have even

$$F([0, x] \times \mathbb{R}) = F(\mathbb{R} \times [0, x]) = x, \quad 0 \leqslant x \leqslant 1,$$

we write $F \in \mathcal{C}$ and call it a copula. Hence, we have defined a further subclass and have altogether the inclusions

$$\mathcal{C} \subseteq \mathcal{M}^c \subseteq \mathcal{M}.$$

Define next

$$\mathcal{M}_+ \stackrel{\text{def}}{=} \left\{F \in \mathcal{M} : F\left((\mathbb{R}_+^2)^c\right) = 0\right\}, \quad \mathcal{M}_+^c \stackrel{\text{def}}{=} \mathcal{M}_+ \cap \mathcal{M}^c$$

and note $\mathcal{C} \subseteq \mathcal{M}_+^c$.

For a more convenient notation, we do not distinguish between a probability measure and its distribution function, e.g. we shall write without confusion

$$F((-\infty, x] \times (-\infty, y]) = F(x, y), \quad x, y \in \mathbb{R},\ F \in \mathcal{M}.$$

The definition of the map $\mathcal{T}$ in the following proposition is crucial for our purpose.

**Proposition 3.1.1.** *There exists a unique map*

$$\mathcal{T} : \mathcal{M}^c \to \mathcal{C}$$

*with the property*

$$F(x, y) = (\mathcal{T}F)(F_1(x), F_2(y)), \quad x, y \in \mathbb{R}.$$

*Here, $F_k$, $k = 1, 2$ denotes the $k$-th marginal distribution of $F \in \mathcal{M}^c$.*

*Proof.* This is a consequence of Theorem 2.3.3. (Sklar) in Nelsen [28]. □

$\mathcal{T}$ is a map that transforms a probability measure in its copula. We require the margins to be continuous in order to get a unique map. Note that of course $\mathcal{T}_{|\mathcal{C}} = \mathrm{id}_{|\mathcal{C}}$. We shall deal with the following concrete copulas. Let $0 \leqslant u, v \leqslant 1$.

| | |
|---|---|
| $\Pi(u,v) = uv$ | independence copula, |
| $W(u,v) = \max(u+v-1,0)$ | Fréchet-Hoeffding lower bound, |
| $M(u,v) = \min(u,v)$ | Fréchet-Hoeffding upper bound, |
| $C_\theta(u,v) = \left\{\left(\max\{u^{-\theta}+v^{-\theta}-1,0\}\right)^{-\frac{1}{\theta}}\right\}_{\theta\in[-1,\infty)\setminus\{0\}}$ | family of Clayton copulas. |

Furthermore, we define a function $\rho$ via

$$\rho(F) = \frac{\int xy\, dF(x,y)}{\sqrt{\int x^2\, dF(x,y) \int y^2\, dF(x,y)}}$$

on the domain of all $F \in \mathcal{M}^c$ which possess square integrable margins. We further write as in Buchmann and Grübel [7]

$$\Psi(\lambda, F) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} F^{*k}, \quad F \in \mathcal{M}.$$

## 3.2 Asymptotic results

**Lemma 3.2.1.** *Given $F$, $G \in \mathcal{M}^c$ and $\lambda > 0$. Then $F * G \in \mathcal{M}^c$ and $\Psi(\lambda, F) \in \mathcal{M}^c$ holds.*

*Proof.* Fix any $r \in \mathbb{R}$. Then, Fubinis theorem yields

$$\begin{aligned} (F * G)_j(\{r\}) &= (F_j \otimes G_j)(\{(x,y) \in \mathbb{R}^2 : x + y = r\}), \quad j = 1,2 \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_{\{r-y\}}(x)\, dF_j(x)\, dG_j(y) \\ &= \int_{\mathbb{R}} F_j(\{r-y\})\, dG_j(y) \\ &= 0. \end{aligned}$$

In order to prove the second assertion note that we have

$$\begin{aligned} \Psi(\lambda, F)_j(\{r\}) &= \left(e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} F^{*k}\right)_j (\{r\}), \quad j = 1,2 \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} (F^{*k})_j(\{r\}) \\ &= 0 \end{aligned}$$

where the last expression equals zero because of what we have proven at the beginning. □

**Proposition 3.2.2.** *Let $(F_n)_{n\in\mathbb{N}_0} \subseteq \mathcal{M}^c$ and*

$$F_n \xrightarrow{d} F_0, \quad n \to \infty.$$

*Then we have*

$$\sup_{u,v\in[0,1]} |(\mathcal{T}F_n)(u,v) - (\mathcal{T}F_0)(u,v)| \to 0, \quad n \to \infty. \tag{3.5}$$

*Proof.* Set $C_n \overset{\text{def}}{=} \mathcal{T} F_n$ and let $F_n^j$, $j = 1, 2$ denote the two marginal distributions of $F_n$. Fix $(u, v) \in (0, 1)^2$. Then, there exist $x, y \in \mathbb{R}$ with

$$F_0^1(x) = u, \quad F_0^2(y) = v.$$

We have

$$F_0(x, y) = C_0(F_0^1(x), F_0^2(y)) = C_0(u, v).$$

Since, with the margins of $F_0$, $F_0$ itself is also continuous, the assumption $F_n \xrightarrow{\text{d}} F_0$ yields

$$C_n(F_n^1(x), F_n^2(y)) = F_n(x, y) \to F_0(x, y) = C_0(u, v), \quad n \to \infty. \tag{3.6}$$

Next, it holds

$$|C_n(u, v) - C_n(F_n^1(x), F_n^2(y))| \leqslant |u - F_n^1(x)| + |v - F_n^2(y)| \to 0, \quad n \to \infty \tag{3.7}$$

because every copula is Lipschitz continuous, cf. Nelsen [28][Theorem 2.2.4]. The latter convergence to zero results from $F_n^j \xrightarrow{\text{d}} F_0^j$, $j = 1, 2$ which is a direct consequence of the Cramér-Wold Theorem. The pointwise convergence in (3.5) follows from (3.6) and (3.7).

For the uniform convergence, fix $\epsilon > 0$ and choose any $m > \frac{1}{\epsilon}$, $m \in \mathbb{N}$. Then, we have for all $0 \leqslant u, v \leqslant 1$ and

$$u_m \overset{\text{def}}{=} \frac{\lfloor um \rfloor}{m}, \quad v_m \overset{\text{def}}{=} \frac{\lfloor vm \rfloor}{m}$$

$$\begin{aligned} |C_n(u, v) - C(u, v)| &\leqslant |C_n(u, v) - C_n(u_m, v_m)| + |C_n(u_m, v_m) - C(u_m, v_m)| \\ &\quad + |C(u_m, v_m) - C(u, v)| \\ &\leqslant \frac{4}{m} + \max_{0 \leqslant j, k \leqslant m} \left| C_n\left(\frac{j}{m}, \frac{k}{m}\right) - C\left(\frac{j}{m}, \frac{k}{m}\right) \right| \\ &\leqslant 5\epsilon \end{aligned}$$

for all $n \in \mathbb{N}$ large enough. □

Observe also the paper of Sempi [40] for further results in this area.

**Remark 3.2.3.** Let $\Sigma \in \mathbb{R}^{2\times 2}$ be a positive-semidefinite matrix. Then, we obviously have

$$N(0, \Sigma) \in \mathcal{M}^c \iff \sigma_{11}\sigma_{22} > 0.$$

Assume this is the case, i.e. $\sigma_{11}\sigma_{22} > 0$. Then

(i) $\Sigma$ is strictly positive definite, iff $\frac{|\sigma_{12}|}{\sqrt{\sigma_{11}\sigma_{22}}} < 1$.

(ii) $\mathcal{T} N(0, \Sigma) = W$, iff $\frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}} = -1$.

(iii) $\mathcal{T} N(0, \Sigma) = M$, iff $\frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}} = 1$.

**Proposition 3.2.4.** *Let $\Sigma$, $\Sigma' \in \mathbb{R}^{2\times 2}$ be two positive-definite matrices with $\sigma_{11}\sigma_{22} > 0$. Then we have*

$$\mathcal{T} N(0, \Sigma) = \mathcal{T} N(0, \Sigma') \quad \iff \quad \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}} = \frac{\sigma'_{12}}{\sqrt{\sigma'_{11}\sigma'_{22}}}. \tag{3.8}$$

*Proof.* First, we can assume because of the previous Remark 3.2.3 w.l.o.g. that $\Sigma$ and $\Sigma'$ are strictly positive definite. We have

$$\mathcal{T}N(0,\Sigma)(u,v) = \Phi_\Sigma(\phi_{\sigma_{11}}^{-1}(u), \phi_{\sigma_{22}}^{-1}(v)), \quad u, v \in [0,1]$$

where $\phi_\Sigma$ resp. $\phi_{\sigma_{jj}}$ denotes the cumulative distribution function of $\mathrm{N}(0,\Sigma)$ resp. $\mathrm{N}(0,\sigma_{jj})$, $j = 1,2$. Set further $\phi \stackrel{\text{def}}{=} \phi_1$. Considering the respective densities, the equality of the left hand side in (3.8) is equivalent to

$$\frac{\partial^2}{\partial u \partial v}\mathcal{T}N(0,\Sigma)(u,v) = \frac{\partial^2}{\partial u \partial v}\mathcal{T}N(0,\Sigma')(u,v), \quad u, v \in [0,1].$$

Note that

$$\begin{aligned}
\frac{\partial^2}{\partial u \partial v}\Phi_\Sigma(\phi_{\sigma_{11}}^{-1}(u), \phi_{\sigma_{22}}^{-1}(v)) &= \frac{\partial \phi_{\sigma_{11}}^{-1}}{\partial u}\frac{\partial \phi_{\sigma_{22}}^{-1}}{\partial v}\frac{\partial^2 \Phi_\Sigma}{\partial x \partial y} \\
&= \left( \frac{e^{-\frac{x^2}{2\sigma_{11}} - \frac{y^2}{2\sigma_{22}}}}{2\pi\sqrt{\sigma_{11}\sigma_{22}}} \right)^{-1} \frac{e^{-\frac{1}{2}(x,y)\Sigma^{-1}\binom{x}{y}}}{2\pi(\det\Sigma)^{\frac{1}{2}}} \Bigg|_{\binom{x}{y} = \binom{\sqrt{\sigma_{11}}\phi^{-1}(u)}{\sqrt{\sigma_{22}}\phi^{-1}(v)}} \\
&= \sqrt{\frac{\sigma_{11}\sigma_{22}}{\det\Sigma}} e^{-\frac{1}{2}z^t(\Sigma^{-1} - D^{-2})z} \Bigg|_{z = \binom{x}{y}},
\end{aligned}$$

with

$$D \stackrel{\text{def}}{=} \begin{pmatrix} \sqrt{\sigma_{11}} & 0 \\ 0 & \sqrt{\sigma_{22}} \end{pmatrix}.$$

We have

$$\begin{aligned}
z^t(\Sigma^{-1} - D^{-2})z &= \left( D \begin{pmatrix} \phi^{-1}(u) \\ \phi^{-1}(v) \end{pmatrix} \right)^t (\Sigma^{-1} - D^{-2}) D \begin{pmatrix} \phi^{-1}(u) \\ \phi^{-1}(v) \end{pmatrix} \\
&= \begin{pmatrix} \phi^{-1}(u) \\ \phi^{-1}(v) \end{pmatrix}^t D^t(\Sigma^{-1} - D^{-2}) D \begin{pmatrix} \phi^{-1}(u) \\ \phi^{-1}(v) \end{pmatrix}.
\end{aligned}$$

Furthermore, note

$$\begin{aligned}
D^t(\Sigma^{-1} - D^{-2})D = D\Sigma^{-1}D - I = (D^{-1}\Sigma D^{-1})^{-1} - I \\
= \frac{1}{1-a^2}\begin{pmatrix} 1 & -a \\ -a & 1 \end{pmatrix} - I, \quad a \stackrel{\text{def}}{=} \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}
\end{aligned}$$

and

$$\frac{\det\Sigma}{\sigma_{11}\sigma_{22}} = 1 - a^2.$$

This proves, together with the fact that

$$\begin{array}{rcl} (0,1)^2 & \to & \mathbb{R}^2 \\ (u,v) & \mapsto & (\phi^{-1}(u), \phi^{-1}(v)) \end{array}$$

is a surjection, our claim. □

**Theorem 3.2.5.** *Let $F \in \mathcal{M}^c$ be a distribution with square integrable margins, i.e.*

$$\int (x^2 + y^2)\, dF(x,y) < \infty.$$

*Then,*

$$\sup_{0 \leqslant u,v \leqslant 1} |\mathcal{T}\Psi(\lambda, F) - \mathcal{T}N(0,\Sigma)| \to 0, \quad \lambda \to \infty$$

*with*

$$\Sigma \stackrel{\text{def}}{=} \begin{pmatrix} \int x^2\, dF(x,y) & \int xy\, dF(x,y) \\ \int xy\, dF(x,y) & \int y^2\, dF(x,y) \end{pmatrix}.$$

*Proof.* Denote with $\sum_{j=1}^{N_t(\lambda)} X_j$ the corresponding compound Poisson process to $\Psi(\lambda, F)$, i.e.

$$P^{Z_\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k F^{*k}}{k!} = \Psi(\lambda, F)$$

with

$$Z_\lambda \stackrel{\text{def}}{=} \sum_{j=1}^{N_1(\lambda)} X_j, \quad \lambda > 0.$$

Next, define a map

$$\begin{array}{rccl} \varphi_\lambda : & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (x,y) & \mapsto & \frac{1}{\sqrt{\lambda}}((x,y) - EZ_\lambda). \end{array}$$

It suffices to show that

$$\varphi_\lambda(Z_\lambda) \xrightarrow{d} N(0,\Sigma), \quad \lambda \to \infty \tag{3.9}$$

because Theorem 2.4.3 in Nelsen [28] yields

$$\mathcal{T}P^{\varphi_\lambda(Z_\lambda)} = \mathcal{T}P^{Z_\lambda} = \mathcal{T}\Psi(\lambda, F),$$

so that Proposition 3.2.2 proves this theorem. Note that the latter application of $\mathcal{T}$ is allowed because of Lemma 3.2.1 and the fact that $\varphi_\lambda$ is injective.

We verify (3.9) by use of Lévy's continuity theorem: For a convenient notation, first set

$$Y_v \stackrel{\text{def}}{=} \langle v, X_1 \rangle, \quad v \in \mathbb{R}^2.$$

Next, observe that we have

$$\mathcal{F}N(0,\Sigma)(v) = \exp\left(-\frac{v^t \Sigma v}{2}\right) = \exp\left(-\frac{EY_v^2}{2}\right).$$

Hence, it suffices to establish for every $v \in \mathbb{R}^2$ the convergence

$$\mathcal{F}[Z_\lambda](v) \to \exp\left(-\frac{EY_v^2}{2}\right), \quad \lambda \to \infty.$$

Write for this purpose

$$E\left(\exp\left(i\left\langle v, \varphi_\lambda(Z_\lambda)\right\rangle\right)\right) = E\left(\exp\left(\frac{i}{\sqrt{\lambda}}(\langle v, Z_\lambda\rangle - \lambda E\langle v, X_1\rangle)\right)\right)$$

$$
\begin{aligned}
&= E\left(\exp\left(\frac{i}{\sqrt{\lambda}}\sum_{j=1}^{N_1(\lambda)}\langle v, X_j\rangle\right)\right)\exp(-i\sqrt{\lambda}E\langle v, X_1\rangle) \\
&= \exp\left(\lambda(\mathcal{F}[Y_v](\lambda^{-\frac{1}{2}})-1)\right)\exp(-i\sqrt{\lambda}EY_v) \\
&= \exp\left(\lambda\left(\frac{i}{\sqrt{\lambda}}EY_v - \frac{1}{2\lambda}EY_v^2 + o(\lambda^{-1})\right)\right)\exp(-i\sqrt{\lambda}EY_v) \\
&= \exp\left(-\frac{EY_v^2}{2}+o(1)\right), \quad \lambda\to\infty.
\end{aligned}
$$

This proves the desired convergence. Note that we used Sato [38][Theorem 4.3] for the third equal sign and Chow, Teicher [9][8.4 Theorem 1] for the fourth equal sign.

□

**Corollary 3.2.6.** *Let $F \in \mathcal{M}^c$ be a distribution with square integrable margins. Then there exists another such distribution $G$ and a number $\Lambda > 0$, such that $\mathcal{T}F = \mathcal{T}G$, but*

$$\mathcal{T}\Psi(\lambda, F) \neq \mathcal{T}\Psi(\lambda, G), \quad \lambda \geqslant \Lambda.$$

*To be more precise, there exists $u_0, v_0 \in (0,1)$ such that*

$$\lim_{\lambda\uparrow\infty}|\mathcal{T}\Psi(\lambda, F) - \mathcal{T}\Psi(\lambda, G)|(u_0, v_0) > 0.$$

*Additionally, we can choose $G \in \mathcal{M}^c_+$ if $F \in \mathcal{M}^c_+$.*

*Proof.* For any $c, d \geqslant 0$ set $G \stackrel{\text{def}}{=} \delta_{(c,d)} * F \in \mathcal{M}^c$ and observe $G \in \mathcal{M}^c_+$ if $F \in \mathcal{M}^c_+$. Note that because of Theorem 2.4.3 in Nelsen [28] we have $\mathcal{T}F = \mathcal{T}G$. Next, let $X$ be a random variable with $X \sim F$, so that $(X_1 + c, X_2 + d) \sim G$. Due to Theorem 3.2.5 together with Proposition 3.2.4, we only have to show that we can choose $c$ and $d$ such that $\rho(F) \neq \rho(G)$. For this purpose, note that

$$
\begin{aligned}
E(X_1+c)^2 &= EX_1^2 + 2cEX_1 + c^2 \\
E(X_2+d)^2 &= EX_2^2 + 2dEX_2 + d^2 \\
E(X_1+c)(X_2+d) &= EX_1X_2 + cEX_2 + dEX_1 + cd.
\end{aligned}
$$

Set

$$
\begin{aligned}
\widetilde{\rho}: \quad \mathbb{R}^2_+ &\to [0,1] \\
(c,d) &\mapsto \left(\frac{E(X_1+c)(X_2+d)}{\sqrt{E(X_1+c)^2E(X_2+d)^2}}\right)^2
\end{aligned}
$$

and observe

$$\widetilde{\rho}(c,d) = \frac{(EX_1X_2 + cEX_2 + dEX_1 + cd)^2}{(EX_1^2 + 2cEX_1 + c^2)(EX_2^2 + 2dEX_2 + d^2)}.$$

Assume $(c, d) \mapsto \widetilde{\rho}(c, d)$ is constant. Then

$$d \mapsto \breve{\rho}(d) \stackrel{\text{def}}{=} \lim_{c\uparrow\infty}\widetilde{\rho}(c,d) = \frac{(EX_2+d)^2}{E(X_2+d)^2}$$

is also constant. This implies

$$\frac{(EX_2)^2}{EX_2^2} = \breve{\rho}(0) = \lim_{d\uparrow\infty}\breve{\rho}(d) = 1$$

which is only possible if $\mathrm{Var}X_2 = 0$, i.e. $X_2$ is a.s. constant which is a contradiction to the assumed continuity of the second marginal distribution of $F$. □

**Remark 3.2.7.** (3.3) is equivalent to (12) in [39]. Unfortunately, (12) in [39] does not hold in view of Corollary 3.2.6. The equality in (12) is an important assumption of the test constructed in that paper, i.e. (3.2). Nevertheless, we think that the basic ideas in [39] can still yield a useful Lévy copula estimator. However, the limit distribution will probably depend on the margins of the jump distribution in contrast to what is claimed in [39].

## 3.3 Two examples

Consider the introduction of this chapter for the motivation of the following two examples.

**Proposition 3.3.1.** *We have*

$$\frac{1}{2} \leqslant \rho(C) \leqslant 1, \quad C \in \mathcal{C}.$$

*Proof.* The upper bound is a direct consequence of the Cauchy-Schwarz inequality. For the lower bound suppose $(U, V) \sim C$, i.e. in particular $U, V \sim U[0,1]$. We then have to show that

$$EUV \geqslant \frac{1}{6}.$$

Cauchy-Schwarz yields

$$E(1-V)U \leqslant (E(1-V)^2 EU^2)^{\frac{1}{2}} = EU^2.$$

This implies

$$EUV \geqslant EU - EU^2 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

which proves the claim. □

**Example 3.3.2.** *Let $(C_\theta)_{\theta \in [-1,\infty)\backslash\{0\}}$ be the family of Clayton copulas. Then we have*

$$\rho(\{C_\theta \,:\, \theta \in [-1,\infty)\backslash\{0\}\} \cup \{\Pi\} \cup \{M\}) = \left[\frac{1}{2}, 1\right].$$

*Proof.* First, we show the continuity of the map

$$\theta \mapsto \rho(C_\theta), \quad \theta \in [-1,\infty)\backslash\{0\}. \tag{3.10}$$

For this purpose, choose a sequence $(\theta_n)_{n \in \mathbb{N}_0} \subseteq [-1,\infty)\backslash\{0\}$ with $\theta_n \to \theta_0$. The pointwise convergence

$$C_{\theta_n}(u,v) \to C_{\theta_0}(u,v), \quad 0 \leqslant u, v \leqslant 1$$

yields the convergence of measures $C_{\theta_n} \xrightarrow{d} C_{\theta_0}$. Define the product function

$$H \,:\, [0,1]^2 \to [0,1], \quad (u,v) \mapsto uv.$$

Since $H$ is continuous, we have $C^H_{\theta_n} \xrightarrow{d} C^H_{\theta_0}$, which implies

$$C_{\theta_n}(H \leqslant t) \to C_\theta(H \leqslant t), \quad \theta_n \to \theta \quad (t\text{-a.e.}). \tag{3.11}$$

Finally we can write

$$\begin{aligned}\left|\int H\,dC_{\theta_n} - \int H\,dC_\theta\right| &= \left|\int_0^1 C_{\theta_n}(H > t)\,dt - \int_0^1 C_\theta(H > t)\,dt\right| \\ &\leqslant \int_0^1 |C_{\theta_n}(H \leqslant t) - C_\theta(H \leqslant t)|\,dt \to 0, \quad n \to \infty\end{aligned}$$

where the last convergence holds because of (3.11) and dominated convergence. This proves the claimed continuity. Next, observe the pointwise convergences

$$\begin{aligned}&C_{\theta_n} \to \Pi, \quad \theta_n \to 0, \quad \theta_n \in [-1,\infty)\backslash\{0\},\\ &C_{\theta_n} \to M, \quad \theta_n \to \infty\end{aligned}$$

and $C_{-1} = W$, cf. Nelsen [28] (4.2.1). This completes, together with the continuity of (3.10),

$$\rho(W) = \frac{1}{2}, \quad \rho(\Pi) = \frac{3}{4}, \quad \rho(M) = 1$$

and Proposition 3.3.1, the proof. □

**Example 3.3.3.** *Let* $\{U_k : k \in \mathbb{N}\} \cup \{V_k : k \in \mathbb{N}\}$ *be a family of i.i.d* $U[0,1]$ *distributed random variables and fix any* $0 < \epsilon < 1$. *Set*

$$T_\epsilon \stackrel{\text{def}}{=} \inf\{k \in \mathbb{N} : |U_k - V_k| \geqslant \epsilon\}.$$

*Then the following two statements are true:*

*(i)* $(U_{T_\epsilon}, V_{T_\epsilon}) \sim U(I_\epsilon)$ *with* $I_\epsilon = \{(u,v) \in [0,1]^2 : |u-v| \geqslant \epsilon\}$.

*(ii) Set*

$$\begin{aligned}\varphi:\ (0,1) &\to [0,1]\\ \epsilon &\mapsto \rho(P^{(U_{T_\epsilon},V_{T_\epsilon})}).\end{aligned}$$

*It holds* $(0, \frac{3}{4}) \subseteq \varphi((0,1))$.

*Proof.* To have an unambiguous notation in this proof, the two dimensional Lebesgue measure is denoted in the following with $\mathfrak{l}^2$ instead of $\lambda^2$.

(i): Let $A \in \mathcal{B}^2$. Then we have

$$\begin{aligned}&P((U_{T_\epsilon}, V_{T_\epsilon}) \in A)\\ =\ &\sum_{k=1}^{\infty} P((U_{T_\epsilon}, V_{T_\epsilon}) \in A | T_\epsilon = k) P(T_\epsilon = k)\\ =\ &\sum_{k=1}^{\infty} P((U_k, V_k) \in A | (U_1, V_1) \in I_\epsilon^c, \ldots, (U_{k-1}, V_{k-1}) \in I_\epsilon^c, (U_k, V_k) \in I_\epsilon) P(T_\epsilon = k)\\ =\ &\sum_{k=1}^{\infty} P((U_k, V_k) \in A | (U_k, V_k) \in I_\epsilon) P(T_\epsilon = k)\\ =\ &\frac{P((U_1, V_1) \in A \cap I_\epsilon)}{P((U_1, V_1) \in I_\epsilon)}.\end{aligned}$$

Note that $P(T_\epsilon = \infty) = 0$.
(ii): Obviously $\mathfrak{l}^2(I_\epsilon) = (1-\epsilon)^2$ holds. Next, we obtain

$$\begin{aligned} EU_{T_\epsilon}^2 &= (\mathfrak{l}^2(I_\epsilon))^{-1} \int_{[0,1]^2} u^2\, \mathbb{1}_{I_\epsilon}(u,v) d\mathfrak{l}^2(u,v) \\ &= (1-\epsilon)^{-2} \left( \int_\epsilon^1 \int_0^{u-\epsilon} u^2\, dv\, du + \int_\epsilon^1 \int_0^{v-\epsilon} u^2\, du\, dv \right) \\ &= \frac{\epsilon^2}{6} + \frac{1}{3} \end{aligned}$$

and

$$\begin{aligned} E(U_{T_\epsilon} V_{T_\epsilon}) &= (\mathfrak{l}^2(I_\epsilon))^{-1} \int_{[0,1]^2} uv\, \mathbb{1}_{I_\epsilon}(u,v)\, d\mathfrak{l}^2(u,v) \\ &= 2(1-\epsilon)^{-2} \int_\epsilon^1 \int_0^{u-\epsilon} uv\, dv\, du \\ &= \frac{(1-\epsilon)(3+\epsilon)}{12}. \end{aligned}$$

A symmetry argument yields $E(V_{T_\epsilon}^2) = E(U_{T_\epsilon}^2)$, so that we have

$$\varphi(\epsilon) = \frac{(1-\epsilon)(3+\epsilon)}{2(\epsilon^2+2)}.$$

Continuity of $\varphi$ and

$$0 = \lim_{\epsilon \uparrow 1} \varphi(\epsilon) < \lim_{\epsilon \downarrow 0} \varphi(\epsilon) = \frac{3}{4}$$

proves (ii). □

# Appendix A

# Auxiliary results for Chapter 1

**Proposition A.1.** *Let $(H_i)_{i\in\mathbb{N}}$ be a family of i.i.d. random variables with the property*

$$\exists r > 0 \,:\, Ee^{tH_1} < \infty, \quad t \in (-r, r).$$

*Then we have for every $N \in \mathbb{N}$*

$$E\left[\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(H_i - EH_1)\right)^{2N}\right] \to E[N(0, \mathrm{Var}\ H_1)]^{2N}, \quad n \to \infty, \tag{A.1}$$

*which in turn implies*

$$E\left(\frac{1}{n}\sum_{i=1}^{n} H_i - EH_1\right)^{2N} = O\left(\frac{1}{n^N}\right), \quad n \to \infty. \tag{A.2}$$

*Proof.* First, we establish for all $N \in \mathbb{N}$

$$\sup_{n\in\mathbb{N}} E\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(H_i - EH_1)\right)^{2N} < \infty. \tag{A.3}$$

Fix for this purpose any $t \in (0, r)$ and note that because of

$$\frac{t^{2N}}{(2N)!}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(H_i - EH_1)\right)^{2N} \leqslant e^{\frac{t}{\sqrt{n}}\sum_{i=1}^{n}(H_i - EH_1)}, \quad n, N \in \mathbb{N},$$

it suffices to verify

$$\sup_{n\in\mathbb{N}} E\left(e^{\frac{t}{\sqrt{n}}\sum_{i=1}^{n}(H_i - EH_1)}\right) < \infty.$$

Further, since

$$E\left(e^{\frac{t}{\sqrt{n}}\sum_{i=1}^{n}(H_i - EH_1)}\right) = \left[E\left(e^{\frac{t}{\sqrt{n}}(H_1 - EH_1)}\right)\right]^n, \quad n \in \mathbb{N},$$

it is enough to prove

$$E\left(e^{\frac{t}{\sqrt{n}}(H_1 - EH_1)}\right) \leqslant 1 + \frac{C}{n}, \quad n \in \mathbb{N}$$

for some constant $C > 0$. Consider for this

$$E\left(e^{\frac{t}{\sqrt{n}}(H_1-EH_1)}\right) \leqslant 1+\sum_{k=2}^{\infty}\frac{1}{k!}\left(\frac{t}{\sqrt{n}}\right)^k E(|H_1-EH_1|^k) \leqslant 1+\frac{1}{n}\underbrace{\sum_{k=2}^{\infty}\frac{t^k}{k!}E|H_1-EH_1|^k}_{\stackrel{\text{def}}{=}C}, \quad n\in\mathbb{N}$$

and

$$C \leqslant E\left(e^{t|H_1-EH_1|}\right) \leqslant e^{t|EH_1|}\left(E\left(e^{tH_1}\right)+E\left(e^{-tH_1}\right)\right) < \infty.$$

Next, Bauer [5][§21, Exercise 5] implies together with (A.3) the uniform integrability of

$$\left\{\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(H_i-EH_1)\right)^{2N}\right\}_{n\in\mathbb{N}}. \tag{A.4}$$

Moreover, the central limit theorem states

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(H_i-EH_1) \xrightarrow{\text{d}} N(0,\text{Var } H_1). \tag{A.5}$$

Finally Billingsley [6][Theorem 3.5.] proves together with the uniform integrability of (A.4) and the weak convergence (A.5) the claim (A.1). □

**Corollary A.2.** *Let $(Z_i)_{i\in\mathbb{N}}$ be a family of $N(0,1)$ i.i.d. random variables and $N\in\mathbb{N}$. Then there is a constant $C = C(N) > 0$ such that*

$$P\left(\left|\frac{1}{n}\sum_{i=0}^{n-1}Z_i^2-1\right| \geqslant \epsilon\right) \leqslant \frac{C}{\epsilon^{2N}n^N} \tag{A.6}$$

*and*

$$P\left(\left|\frac{\pi}{2(n-1)}\sum_{i=0}^{n-2}|Z_i||Z_{i+1}|-1\right| \geqslant \epsilon\right) \leqslant \frac{C}{\epsilon^{2N}n^N} \tag{A.7}$$

*for every $n\in\mathbb{N}$ and $\epsilon > 0$.*

*Proof.* Note that

$$Ee^{tZ_1^2} = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{tx^2}e^{-\frac{x^2}{2}}\,dx < \infty, \quad t < \frac{1}{2}.$$

Hence, (A.6) follows from the Markov inequality and Proposition A.1.

To prove (A.7), observe

$$E|Z_1||Z_2| = \frac{2}{\pi}, \quad Ee^{t|Z_1||Z_2|} \leqslant Ee^{\frac{t}{2}(Z_1^2+Z_2^2)} = \left(Ee^{\frac{t}{2}Z_1^2}\right)^2 < \infty, \quad t < 1 \tag{A.8}$$

and decompose

$$\frac{\pi}{2(n-1)}\sum_{i=0}^{n-2}|Z_i||Z_{i+1}|-1$$

$$
\begin{aligned}
&= \frac{\pi}{2(n-1)} \sum_{i=0}^{n-2} \left( |Z_i||Z_{i+1}| - \frac{2}{\pi} \right) \\
&= \frac{\pi}{4} \left( \frac{2}{n-1} \sum_{i=0}^{\lceil \frac{n}{2} \rceil - 2} \left( |Z_{2i}||Z_{2i+1}| - \frac{2}{\pi} \right) + \frac{2}{n-1} \sum_{i=0}^{\lceil \frac{n}{2} \rceil - 2} \left( |Z_{2i+1}||Z_{2i+2}| - \frac{2}{\pi} \right) \right. \qquad \text{(A.9)} \\
&\qquad \left. + \frac{2}{n-1} |Z_{n-2}||Z_{n-1}| \, \mathbb{1}_{2\mathbb{N}}(n) - \frac{4}{\pi(n-1)} \, \mathbb{1}_{2\mathbb{N}}(n) \right).
\end{aligned}
$$

Once again, (A.7) follows from the Markov inequality and Proposition A.1. □

**Proposition A.3.** *Let $(Z_i)_{i \in \mathbb{N}}$ be a family of i.i.d. random variables with $Z_i \sim N(0,1)$ and $N \in \mathbb{N}$. Then we have*

$$
E \left( \max_{1 \leqslant i \leqslant n} |Z_i| \right)^N \leqslant 2^N (\log n)^{\frac{N}{2}} + 2N!
$$

*for all sufficient large $n \in \mathbb{N}$. To put a finer point on it, the above inequality holds at least for $n \geqslant 55$.*

*Proof.* The proof is divided into two steps. The first step derives some elementary inequalities. Finally, the second step estimates the desired $N$-th moment of the maximum using the inequalities in the first step.

STEP 1. *Compilation of some helpful inequalities.* First, we note the best-known inequality

$$
1 + y \leqslant \exp(y), \quad y \in \mathbb{R}. \qquad \text{(A.10)}
$$

For the opposite direction, an analysis of the extrema of the function

$$
y \mapsto \log(1+y) - y + y^2, \quad y > -1
$$

implies

$$
\log(1+y) \geqslant y - y^2, \quad -\frac{1}{2} < y < \infty. \qquad \text{(A.11)}
$$

Further, we have

$$
P(Z_1 \geqslant t) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-\frac{x^2}{2}} \, dx \leqslant \frac{1}{\sqrt{2\pi}} \int_t^\infty \frac{x}{t} e^{-\frac{x^2}{2}} \, dx = \frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{t^2}{2}}}{t}, \quad t > 0,
$$

which yields

$$
2P(Z_1 \geqslant t) \leqslant \sqrt{\frac{2}{\pi}} \frac{1}{t} e^{-\frac{t^2}{2}} \leqslant e^{-\frac{t^2}{2}}, \quad t \geqslant 1. \qquad \text{(A.12)}
$$

If $n \geqslant \lceil e^4 \rceil = 55$, it follows $2\sqrt{\log n} \geqslant 2\sqrt{\log e^4} = 4$. Hence, $t \geqslant 2\sqrt{\log n}$ obviously implies $\frac{t^2}{4} \geqslant t$. Due to this, we get for $n \geqslant 55$ the implications

$$
t \geqslant 2\sqrt{\log n} \implies t^2 \geqslant 4 \log n, \; \frac{t^2}{4} \geqslant t \implies e^{\frac{t^2}{2} - \frac{t^2}{4}} \geqslant n, \; \frac{t^2}{4} \geqslant t \implies e^{\frac{t^2}{2} - t} \geqslant n,
$$

so that we have

$$
t \geqslant 2\sqrt{\log n}, \; n \geqslant 55 \implies \frac{1}{n} e^{-t} \geqslant e^{-\frac{t^2}{2}}. \qquad \text{(A.13)}
$$

STEP 2. *Estimate of the expected value.* We have

$$\begin{aligned}
E\big(\max_{1\leqslant i\leqslant n}|Z_i|^N\big) &= N\int_0^\infty t^{N-1}P\left(\max_{1\leqslant i\leqslant n}|Z_i|>t\right)dt\\
&= N\int_0^\infty t^{N-1}(1-[P(|Z_1|\leqslant t)]^n)\,dt\\
&= N\int_0^\infty t^{N-1}(1-[1-2P(Z_1\geqslant t)]^n)\,dt\\
&\overset{(A.12)}{\leqslant} N\int_0^{2\sqrt{\log n}} t^{N-1}\,dt + N\int_{2\sqrt{\log n}}^\infty t^{N-1}\left(1-\left(1-e^{-\frac{t^2}{2}}\right)^n\right)dt\\
&\overset{(A.13)}{\leqslant} 2^N(\log n)^{\frac{N}{2}} + N\int_{2\sqrt{\log n}}^\infty t^{N-1}\left(1-\left(1-\frac{e^{-t}}{n}\right)^n\right)dt.
\end{aligned}$$

With $y \overset{\text{def}}{=} -\frac{e^{-t}}{n}$ inequality (A.11) yields

$$n\log\left(1-\frac{e^{-t}}{n}\right)\geqslant n\left(-\frac{e^{-t}}{n}-\frac{e^{-2t}}{n^2}\right),$$

so that

$$\left(1-\frac{e^{-t}}{n}\right)^n \geqslant e^{-e^{-t}-\frac{e^{-2t}}{n}} \overset{(A.10)}{\geqslant} 1-e^{-t}-\frac{e^{-2t}}{n}.$$

This yields, using the Gamma function,

$$\begin{aligned}
N\int_{2\sqrt{\log n}}^\infty t^{N-1}\left(1-\left(1-\frac{e^{-t}}{n}\right)^n\right)dt &\leqslant N\int_0^\infty t^{N-1}\left(e^{-t}+\frac{e^{-2t}}{n}\right)dt\\
&= N\int_0^\infty t^{N-1}e^{-t}\,dt + \frac{N}{n2^N}\int_0^\infty t^{N-1}e^{-t}\,dt\\
&= N\left(1+\frac{1}{n2^N}\right)(N-1)!\\
&\leqslant 2N!
\end{aligned}$$

which proves the claim. □

**Remark A.4.** In view of the scaling constants that are needed to make the maximum of i.i.d. normal distributed random variables converge to the Gumbel distribution, the upper bound in the above Proposition A.3 seems natural.

**Lemma A.5.** *Fix any $0 < c, r < 1$ and let $c$ be an irrational number. Then there are sequences $(n_l)$ and $(k_l)$ of natural numbers with the properties*

$$n_l \uparrow \infty, \quad 0 \leqslant k_l < n_l$$

*and*

$$\frac{k_l}{n_l}+\frac{r}{2n_l} < c < \frac{k_l}{n_l}+\frac{r}{n_l}, \quad l\geqslant 1. \tag{A.14}$$

**Remark A.6.** Note that the statement of Lemma A.5 is wrong if $c\in\mathbb{Q}$. To understand this, assume

$$c=\frac{p}{q}\in(0,1), \quad p,q\in\mathbb{N}$$

and note that (A.14) is equivalent to

$$qk_l + \frac{qr}{2} < pn_l < qk_l + qr, \quad l \geqslant 1. \tag{A.15}$$

Next, set

$$r \stackrel{\text{def}}{=} \frac{1}{2q} \in (0,1)$$

and observe that

$$\left(qk_l + \frac{1}{4}, qk_l + \frac{1}{2}\right) \cap \mathbb{N} = \varnothing$$

since $qk_l \in \mathbb{N}$. This is a contradiction to (A.15) because of $pn_l \in \mathbb{N}$. Nevertheless, we are going to prove that the lemma holds if $c \notin \mathbb{Q}$.

*Proof of Lemma A.5.* Consider the function

$$g : \mathbb{N} \to [0,1], \quad n \mapsto nc - \lfloor nc \rfloor.$$

$g(\mathbb{N})$ is a dense subset of $[0,1]$. This is due to the irrationality of $c$ and can be proven by the pigeon-hole principle, c.f. Arnold [2][§24, page 222]. Observe that (A.14) is the same as claiming

$$n_l c - k_l \in \left(\frac{r}{2}, r\right), \quad l \geqslant 1. \tag{A.16}$$

Since $\left(\frac{r}{2}, r\right) \subseteq [0,1]$ is open and $g(\mathbb{N})$ is dense in $[0,1]$, it follows that

$$g(\mathbb{N}) \cap \left(\frac{r}{2}, r\right) \subseteq [0,1]$$

consists of infinite many points. So we can choose a sequence $n_l \uparrow \infty$ of natural numbers such that

$$g(n_l) \in \left(\frac{r}{2}, r\right), \quad l \geqslant 1 \tag{A.17}$$

holds. Finally, we set

$$0 \leqslant k_l \stackrel{\text{def}}{=} \lfloor n_l c \rfloor < n_l, \quad l \geqslant 1$$

and observe that this choice yields the equivalence of (A.16) and (A.17) which proves this lemma. □

Eventually, let us state an easy, but useful lemma.

**Lemma A.7.** *Fix $a_1, \ldots, a_n$, $b_1, \ldots, b_n \in \mathbb{R}$. Then, the following inequalities hold:*

*(i)*

$$\min(b_1, \ldots, b_n) \leqslant \max(a_1 + b_1, \ldots, a_n + b_n) - \max(a_1, \ldots, a_n) \leqslant \max(b_1, \ldots, b_n)$$

*(ii)*

$$|\max(a_1 + b_1, \ldots, a_n + b_n) - \max(a_1, \ldots, a_n)| \leqslant \max(|b_1|, \ldots, |b_n|)$$

*Proof.* (i) implies (ii), because

$$\max(b_1, \ldots, b_n) \leqslant \max(|b_1|, \ldots, |b_n|)$$

and

$$\min(b_1, \ldots, b_n) = -\max(-b_1, \ldots, -b_n) \geqslant -\max(|b_1|, \ldots, |b_n|).$$

We therefore only have to show (i): Consider for this

$$\begin{aligned} \max(a_1 + b_1, \ldots, a_n + b_n) &\geqslant \max(a_1 + \min_k b_k, \ldots, a_n + \min_k b_k) \\ &= \max(a_1, \ldots, a_n) + \min(b_1, \ldots, b_n) \end{aligned}$$

and

$$\begin{aligned} \max(a_1 + b_1, \ldots, a_n + b_n) &\leqslant \max(\max_k a_k + b_1, \ldots, \max_k a_k + b_n) \\ &= \max(a_1, \ldots, a_n) + \max(b_1, \ldots, b_n). \end{aligned}$$

□

# Appendix B

# Auxiliary results for Chapter 2

Proposition B.2 is a generalization of the following Proposition B.1 which can be proven with standard results from Fourier analysis.

**Proposition B.1.** *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a Schwartz space function. Then it holds*

$$|\mathcal{F}f|(u) \leqslant \frac{1}{|u_1u_2|}\int_{\mathbb{R}^2}\left|\frac{\partial^2 f}{\partial x_1 \partial x_2}(x)\right| \lambda^2(dx), \quad u \in (\mathbb{R}^*)^2. \tag{B.1}$$

*Proof.* Due to Rudin [36][Theorem 7.4 (c)], it holds the equation

$$u_1u_2(\mathcal{F}f)(u) = -\mathcal{F}\left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)(u), \quad u \in \mathbb{R}^2$$

which implies

$$|\mathcal{F}f|(u) = \frac{1}{|u_1u_2|}\left|\mathcal{F}\left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)\right|(u) \leqslant \frac{1}{|u_1u_2|}\int_{\mathbb{R}^2}\left|\frac{\partial^2 f}{\partial x_1 \partial x_2}(x)\right| \lambda^2(dx), \quad u \in (\mathbb{R}^*)^2.$$

□

The situation is more involved in Proposition B.2 since $f$ does not need to be a Schwartz space function. Generally, it is not even a continuous function. The claim of Proposition B.2 states that, in this situation, a similar result as (B.1) also holds. We only have to take the boundaries into account. Here, we are going to provide, for completeness, an elementary proof of Proposition B.2, although the technique is straightforward.

**Proposition B.2.** *Let $g : \mathbb{R}_+^2 \to \mathbb{R}$ be a $\mathcal{C}^2$-function with*

$$g \in L^1(\mathbb{R}_+^2), \quad \left|\frac{\partial g}{\partial x_j}\right|(x) \lesssim (1+|x|)^{-(1+\epsilon)}, \quad \frac{\partial^2 g}{\partial x_1 \partial x_2} \in L^1(\mathbb{R}_+^2), \quad j = 1, 2$$

*for some $\epsilon > 0$ and define a function $f : \mathbb{R}^2 \to \mathbb{R}$ via*

$$f(x) = \begin{cases} g(x), & x \in (\mathbb{R}_+^*)^2, \\ 0, & \textit{else}. \end{cases} \tag{B.2}$$

*Set*

$$\Lambda_g \overset{\text{def}}{=} |g(0,0)| + \int_{\mathbb{R}_+} \left|\frac{\partial g}{\partial x_1}\right| (x_1,0)\, \lambda^1(dx_1) + \int_{\mathbb{R}_+} \left|\frac{\partial g}{\partial x_2}\right| (0,x_2)\, \lambda^1(dx_2) + \int_{\mathbb{R}_+^2} \left|\frac{\partial^2 g}{\partial x_1 \partial x_2}\right| (x)\, \lambda^2(dx).$$

*Then it holds*

$$|\mathcal{F}f|(u) \leqslant \frac{\Lambda_g}{|u_1 u_2|}, \quad u \in (\mathbb{R}^*)^2. \tag{B.3}$$

**Remark B.3.** Note first that $\mathcal{F}f$ is, of course, independent of the values of $f$ on the negligible set

$$\mathcal{N}_{(0,0)} \overset{\text{def}}{=} (\{0\} \times \mathbb{R}_+) \cup (\mathbb{R}_+ \times \{0\}).$$

We only choose the representation in (B.2) for a more comfortable approach in the subsequent proof.

Note furthermore that it holds for every $y \in \mathbb{R}$

$$\mathcal{F}f(u) = e^{i\langle u,y\rangle} \mathcal{F}(f(\cdot + y))(u), \quad u \in \mathbb{R}^2.$$

Hence, our discontinuity set could also have been

$$\mathcal{N}_{(y_1,y_2)} \overset{\text{def}}{=} (\{y_1\} \times [y_2,\infty)) \cup ([y_1,\infty) \times \{y_2\})$$

and the above Proposition B.2 remains true. The choice $y = 0$ is only due to a simpler notation.

*Proof of Proposition B.2.* The proof is divided into three steps. In the first step, we approximate $f$ with a sequence of step functions and prove the $L^1$ convergence of the sequence to the given function $f$. The second step calculates the Fourier transforms of those step functions in relation to the Fourier transform of $f$. Finally, the third step combines the results of the first two steps and proves the desired result.

STEP 1. Set for $j,k \in \mathbb{Z}$ and $n \in \mathbb{N}$

$$A^n_{j,k} \overset{\text{def}}{=} \left[\frac{j}{n}, \frac{j+1}{n}\right) \times \left[\frac{k}{n}, \frac{k+1}{n}\right).$$

Next, approximate $f$ by

$$r_{n,m}(x) \overset{\text{def}}{=} \sum_{j=0}^{m-1} \sum_{k=0}^{m-1} f\left(\frac{j}{n}, \frac{k}{n}\right) \mathbb{1}_{A^n_{j,k}}(x), \quad x \in \mathbb{R}^2, \quad m \in \mathbb{N}.$$

Fix any $N \in \mathbb{N}$. If $x \in (0,N]^2$ and $m \overset{\text{def}}{=} N \cdot n$, we have

$$\begin{aligned} |r_{n,m}(x) - f(x)| &= \left| f\left(\frac{j}{n}, \frac{k}{n}\right) - f(x) \right|, \quad \text{if } x \in A^n_{j,k} \\ &\to 0, \quad n \to \infty, \end{aligned}$$

because of

$$\left| \left(\frac{j}{n}, \frac{k}{n}\right) - x \right| \leqslant \frac{\sqrt{2}}{n}$$

and the continuity of $f$ on $(0,N]^2$. Using the dominated convergence theorem with the constant

$L^1((0,N]^2)$ majorant

$$\sup_{x\in(0,N]^2}|f(x)| = \sup_{x\in[0,N]^2}|g(x)| < \infty,$$

we get $r_{n,N\cdot n}\to f$, $n\to\infty$ in $L^1((0,N]^2)$, i.e.

$$\int_{[0,N]^2}|r_{n,N\cdot n}(x)-f(x)|\lambda^2(dx)\to 0,\quad n\to\infty.$$

Set

$$n_N \stackrel{\text{def}}{=} \inf\{l\in\mathbb{N} \,:\, \int_{[0,N]^2}|r_{n,N\cdot n}(x)-f(x)|\,\lambda^2(dx) < \frac{1}{N},\quad \forall n\geqslant l\}$$

and

$$r_N \stackrel{\text{def}}{=} r_{n_N,N\cdot n_N},\quad N\in\mathbb{N}.$$

Then, we have

$$\int_{\mathbb{R}_+^2}|r_N(x)-f(x)|\,\lambda^2(dx) = \int_{\mathbb{R}_+^2\setminus[0,N]^2}|f(x)|\,\lambda^2(dx) + \int_{[0,N]^2}|r_N(x)-f(x)|\,\lambda^2(dx)\to 0,\quad N\to\infty,$$

i.e. $r_N\to f$ in $L^1(\mathbb{R}_+^2)$.

STEP 2. Write with $u=(u_1,u_2)\in\mathbb{R}^2$

$$\mathcal{F}r_{n,m}(u) = -\frac{1}{u_1u_2}\sum_{j=0}^{m-1}\sum_{k=0}^{m-1} f\left(\frac{j}{n},\frac{k}{n}\right)\left(e^{i\frac{j+1}{n}u_1}-e^{i\frac{j}{n}u_1}\right)\left(e^{i\frac{k+1}{n}u_2}-e^{i\frac{k}{n}u_2}\right)$$

and set

$$f_{j,k}\stackrel{\text{def}}{=} f\left(\frac{j}{n},\frac{k}{n}\right),\quad a_j\stackrel{\text{def}}{=} e^{i\frac{j}{n}u_1},\quad b_k\stackrel{\text{def}}{=} e^{i\frac{k}{n}u_2},\quad j,k\in\mathbb{Z}.$$

This implies

$$\mathcal{F}r_{n,m}(u) = -\frac{1}{u_1u_2}\sum_{j=0}^{m-1}(a_{j+1}-a_j)\sum_{k=0}^{m-1} f_{j,k}(b_{k+1}-b_k). \tag{B.4}$$

Note next that it holds for arbitrary $s_0,\dots,s_{l-1},t_0,\dots,t_l\in\mathbb{C}$, $l\in\mathbb{N}$ the equality

$$\sum_{k=0}^{l-1}s_k(t_{k+1}-t_k) = \sum_{k=0}^{l-1}s_kt_{k+1} - \sum_{k=-1}^{l-2}s_{k+1}t_{k+1} = s_{l-1}t_l - s_0t_0 - \sum_{k=0}^{l-2}t_{k+1}(s_{k+1}-s_k). \tag{B.5}$$

This yields, in particular,

$$\sum_{k=0}^{m-1}f_{j,k}(b_{k+1}-b_k) = \overbrace{f_{j,m-1}b_m}^{\stackrel{\text{def}}{=}(I)} - \overbrace{f_{j,0}b_0}^{=0} - \overbrace{\sum_{k=0}^{m-2}b_{k+1}(f_{j,k+1}-f_{j,k})}^{\stackrel{\text{def}}{=}(II)}.$$

Regarding $(I)$ together with (B.4) we next consider

$$\sum_{j=0}^{m-1}f_{j,m-1}(a_{j+1}-a_j) \quad=\quad f_{m-1,m-1}a_m - \overbrace{f_{0,m-1}a_0}^{=0} - \sum_{j=0}^{m-2}a_{j+1}(f_{j+1,m-1}-f_{j,m-1})$$

$$= \quad a_m f_{m-1,m-1} - \overbrace{\sum_{j=1}^{m-2} a_{j+1}(f_{j+1,m-1} - f_{j,m-1})}^{(*)} - a_1 f_{1,m-1}.$$

Note that $f_{m-1,m-1}$ and $f_{1,m-1}$ tends to zero, if $\frac{m}{n}$ tends to infinity since $f$ vanishes at infinity. Next, there exists for each $1 \leqslant j \leqslant m-2$ a number $\eta_j \in \left[\frac{j}{n}, \frac{j+1}{n}\right]$ such that

$$f_{j+1,m-1} - f_{j,m-1} = \frac{1}{n}\frac{\partial f}{\partial x_1}\left(\eta_j, \frac{m-1}{n}\right).$$

This implies, together with our assumptions,

$$|f_{j+1,m-1} - f_{j,m-1}| \lesssim \frac{1}{n}\left(\frac{n}{m-1}\right)^{1+\epsilon},$$

so that we have

$$\left|\sum_{j=1}^{m-2} a_{j+1}(f_{j+1,m-1} - f_{j,m-1})\right| \leqslant \sum_{j=1}^{m-2} |f_{j+1,m-1} - f_{j,m-1}| \lesssim \frac{m}{n}\left(\frac{n}{m-1}\right)^{1+\epsilon}.$$

Regarding $(II)$ together with (B.4) and (B.5), we consider

$$\begin{aligned}
&\sum_{j=0}^{m-1}(a_{j+1} - a_j)\sum_{k=0}^{m-2} b_{k+1}(f_{j,k+1} - f_{j,k}) \\
= \quad & \overbrace{a_m \sum_{k=0}^{m-2} b_{k+1}(f_{m-1,k+1} - f_{m-1,k})}^{\text{cf. } (*)} - \overbrace{a_0 \sum_{k=0}^{m-2} b_{k+1}(f_{0,k+1} - f_{0,k})}^{=0} \\
& \underbrace{- \sum_{j=0}^{m-2} a_{j+1} \sum_{k=0}^{m-2} b_{k+1}(f_{j+1,k+1} - f_{j+1,k} - f_{j,k+1} + f_{j,k})}_{\stackrel{\text{def}}{=}(III)}.
\end{aligned}$$

So it remains to consider $(III)$. Decompose for this purpose

$$\begin{aligned}
&\sum_{j=0}^{m-2}\sum_{k=0}^{m-2} |f_{j+1,k+1} - f_{j+1,k} - f_{j,k+1} + f_{j,k}| \\
= \quad & |f_{1,1} - f_{1,0} - f_{0,1} + f_{0,0}| + \sum_{k=1}^{m-2} |f_{1,k+1} - f_{1,k} - f_{0,k+1} + f_{0,k}| \\
& + \sum_{j=1}^{m-2} |f_{j+1,1} - f_{j+1,0} - f_{j,1} + f_{j,0}| + \sum_{j=1}^{m-2}\sum_{k=1}^{m-2} |f_{j+1,k+1} - f_{j+1,k} - f_{j,k+1} + f_{j,k}| \\
= \quad & |f_{1,1}| + \sum_{k=1}^{m-2} |f_{1,k+1} - f_{1,k}| + \sum_{j=1}^{m-2} |f_{j+1,1} - f_{j,1}| + \sum_{j=1}^{m-2}\sum_{k=1}^{m-2} |f_{j+1,k+1} - f_{j+1,k} - f_{j,k+1} + f_{j,k}|.
\end{aligned}$$

Each addend of the above expression is considered separately. First, note

$$|f_{1,1}| = \left| g\left(\frac{1}{n}, \frac{1}{n}\right) \right|.$$

Next consider

$$\sum_{k=1}^{m-2} |f_{1,k+1} - f_{1,k}| \leqslant \sum_{k=1}^{m-2} \frac{1}{n} \sup_{\frac{k}{n} \leqslant x_2 \leqslant \frac{k+1}{n}} \left| \frac{\partial g}{\partial x_2} \right| \left(\frac{1}{n}, x_2\right) \tag{B.6}$$

and set

$$h_n(x_2) \stackrel{\text{def}}{=} \sum_{k=1}^{m-2} \sup_{\frac{k}{n} \leqslant x_2 \leqslant \frac{k+1}{n}} \left| \frac{\partial g}{\partial x_2} \right| \left(\frac{1}{n}, x_2\right) \mathbb{1}_{\left[\frac{k}{n}, \frac{k+1}{n}\right)}(x_2), \quad x_2 \in (0, N]$$

Set as in the first step $m \stackrel{\text{def}}{=} N \cdot n$ and note that we have because of the continuity of $\frac{\partial g}{\partial x_2}$ on $\mathbb{R}_+^2$ the pointwise convergence

$$h_n(x_2) \to \left| \frac{\partial g}{\partial x_2} \right| (0, x_2), \quad n \to \infty, \quad x_2 \in (0, N].$$

The dominated convergence theorem with the constant $L^1((0, N])$ majorant

$$\sup_{(x_1,x_2) \in [0,1] \times [0,N]} \left| \frac{\partial g}{\partial x_2} \right| (x_1, x_2) < \infty$$

yields

$$h_n \to \left| \frac{\partial g}{\partial x_2} \right| (0, \cdot), \quad n \to \infty \quad \text{in } L^1((0, N]).$$

This means that

$$\sum_{k=1}^{N \cdot n-2} \frac{1}{n} \sup_{\frac{k}{n} \leqslant x_2 \leqslant \frac{k+1}{n}} \left| \frac{\partial g}{\partial x_2} \right| \left(\frac{1}{n}, x_2\right) \to \int_{[0,N]} \left| \frac{\partial g}{\partial x_2} \right| (0, x_2)\, \lambda^1(dx_2), \quad n \to \infty.$$

So (B.6) implies

$$\limsup_n \sum_{k=1}^{n \cdot N-2} |f_{1,k+1} - f_{1,k}| \leqslant \int_{[0,N]} \left| \frac{\partial g}{\partial x_2} \right| (0, x_2)\, \lambda^1(dx_2). \tag{B.7}$$

Equally, we get

$$\limsup_n \sum_{j=1}^{n \cdot N-2} |f_{j+1,1} - f_{j,1}| \leqslant \int_{[0,N]} \left| \frac{\partial g}{\partial x_1} \right| (x_1, 0)\, \lambda^1(dx_1). \tag{B.8}$$

Note further that we have for $1 \leqslant j, k \leqslant m-2$ and $x = (x_1, x_2) \in \mathbb{R}^2$ the estimate

$$|f_{j+1,k+1} - f_{j+1,k} - f_{j,k+1} + f_{j,k}| = \left| \int_{A_{j,k}} \frac{\partial^2 f}{\partial x_1 \partial x_2}\, \lambda^2(dx) \right| \leqslant \frac{1}{n^2} \sup_{x \in A_{j,k}} \left| \frac{\partial^2 f}{\partial x_1 \partial x_2} \right| (x).$$

Hence, we get with a similar argumentation

$$\limsup_n \sum_{j=1}^{(n \cdot N-2)} \sum_{k=1}^{(n \cdot N-2)} |f_{j+1,k+1} - f_{j+1,k} - f_{j,k+1} + f_{j,k}| \leqslant \int_{[0,N]^2} \left| \frac{\partial^2 g}{\partial x_1 \partial x_2} \right| \lambda^2(dx). \tag{B.9}$$

STEP 3. In view of (B.7) set

$$n_N^{(1)} \overset{\text{def}}{=} \inf\{l \in \mathbb{N} : \sum_{k=1}^{n\cdot N-2} |f_{1,k+1} - f_{1,k}| \leqslant \int_{\mathbb{R}_+} \left|\frac{\partial g}{\partial x_2}\right|(0,x_2)\,\lambda^1(dx_2) + \frac{1}{N}, \quad \forall n \geqslant l\}.$$

Set analogously $n_N^{(2)}$ and $n_N^{(3)}$ regarding (B.8) and (B.9). Finally, define

$$m_N \overset{\text{def}}{=} n_N \vee n_N^{(1)} \vee n_N^{(2)} \vee n_N^{(3)}, \quad f_N \overset{\text{def}}{=} r_{m_N, N\cdot m_N}, \quad N \in \mathbb{N}.$$

Considering the results in the first step together with the estimates in the second step and (B.7), (B.8) and (B.9), we have for $u \in (\mathbb{R}^*)^2$

$$|\mathcal{F}f(u)| \leqslant |\mathcal{F}(f-f_N)(u)| + |\mathcal{F}f_N(u)| \leqslant \|f-f_N\|_1 + \frac{1}{|u_1u_2|}A_N, \quad N \in \mathbb{N} \tag{B.10}$$

with a sequence $(A_N)_N$ independent of $u \in (\mathbb{R}^*)^2$ and

$$\limsup_N A_N \leqslant \Lambda_g, \quad \lim_N \|f-f_N\|_1 = 0.$$

Taking the limes superior on both sides of (B.10), finally, yields (B.3). □

**Corollary B.4.** *Recall*

$$g_{a,b}(x) \overset{\text{def}}{=} \frac{1}{x_1^4+x_2^4}\,\mathbb{1}_{[a,\infty)\times[b,\infty)}(x_1,x_2), \quad (a,b) \in \mathfrak{R}, \quad x \in \mathbb{R}^2$$

*and*

$$\mathfrak{R} \overset{\text{def}}{=} [0,\infty)^2 \backslash \{(0,0)\}.$$

*It holds for all $(a,b) \in \mathfrak{R}$ and $u \in \mathbb{R}^2$ the inequality*

$$|\mathcal{F}g_{a,b}|(u) \lesssim \left(\frac{|(a,b)|^{-4}}{|u_1u_2|}\,\mathbb{1}_{(\mathbb{R}^*)^2}(u)\right) \wedge |(a,b)|^{-2}$$

*where the constant in the above $\lesssim$ is independent of $(a,b) \in \mathfrak{R}$.*

*Proof.* Note that we have

$$\begin{aligned} |\mathcal{F}g_{a,b}|(u) &\leqslant \int_{[a,\infty)\times[b,\infty)} \frac{1}{x_1^4+x_2^4}\,\lambda^2(dx) \lesssim \int_{[a,\infty)\times[b,\infty)} \frac{1}{|x|^4}\,\lambda^2(dx) \\ &\lesssim \int_{[|(a,b)|,\infty)} \frac{1}{r^4} r\,\lambda^1(dr) = 2^{-1}|(a,b)|^{-2} \end{aligned}$$

where we have used the norm equivalence in $\mathbb{R}^2$ and a polar coordinate transformation.

Next, fix any $u \in (\mathbb{R}^*)^2$ and apply Proposition B.2. Observe for this purpose that we have

$$\int_{[a,\infty)} \left|\frac{\partial}{\partial x_1}\frac{1}{x_1^4+x_2^4}\right|(x_1,b)\,\lambda^1(dx_1) = -\int_{[a,\infty)} \frac{\partial}{\partial x_1}\frac{1}{x_1^4+b^4}\,\lambda^1(dx_1) = \frac{1}{a^4+b^4}$$

and analogously

$$\int_{[b,\infty)} \left| \frac{\partial}{\partial x_2} \frac{1}{x_1^4 + x_2^4} \right| (a, x_2)\, \lambda^1(dx_2) = -\int_{[b,\infty)} \frac{\partial}{\partial x_2} \frac{1}{a^4 + x_2^4}\, \lambda^1(dx_2) = \frac{1}{a^4 + b^4}.$$

Finally, a similar consideration yields

$$\int_{[a,\infty)\times[b,\infty)} \left| \frac{\partial^2}{\partial x_1 \partial x_2} \frac{1}{x_1^4 + x_2^4} \right| (x)\, \lambda^2(dx) = \frac{1}{a^4 + b^4} \lesssim |(a,b)|^{-4}.$$

□

**Lemma B.5.** *This lemma states two simple but useful facts.*

*(i) We have with* $u \in \mathbb{R}^2$ *and* $|u_1|, |u_2| \geqslant 1$ *the inequality*

$$\frac{1}{|u_1 u_2|} \leqslant \frac{4}{(1 + |u_1|)(1 + |u_2|)}. \tag{B.11}$$

*(ii) It holds for every* $0 < h < \frac{1}{2}$

$$\int_{\mathbb{R}} \frac{\min(h|z|, 1)}{(1 + |z|)^2}\, dz \lesssim |h \log h|.$$

*Proof.* (i): Set

$$\epsilon_j \overset{\text{def}}{=} |u_j| - 1 \geqslant 0, \quad j = 1, 2.$$

(B.11) is equivalent to

$$(2 + \epsilon_1)(2 + \epsilon_2) \overset{!}{\leqslant} 4(1 + \epsilon_1)(1 + \epsilon_2) = (2 + 2\epsilon_1)(2 + 2\epsilon_2)$$

which is obviously true.

(ii): Write

$$\begin{aligned}
\int_{\mathbb{R}} \frac{\min(h|z|, 1)}{(1 + |z|)^2}\, dz &= 2\int_0^{\frac{1}{h}} \frac{hz}{(1 + z)^2}\, dz + 2\int_{\frac{1}{h}}^{\infty} \frac{dz}{(1 + z)^2} \\
&= 2h\left(\log\left(\frac{1 + h}{h}\right) - \frac{1}{1 + h}\right) + \frac{2h}{1 + h} \\
&= 2h(\log(1 + h) - \log h) \\
&\lesssim |h \log h|, \quad 0 < h < \frac{1}{2}.
\end{aligned}$$

□

Next, we state a version of the Plancherel theorem. The well-known classical Plancherel theorem as stated in Rudin [36][Theorem 7.9] yields the equality

$$\int_{\mathbb{R}^2} \overline{f(x)} g(x)\, \lambda^2(dx) = \frac{1}{4\pi} \int_{\mathbb{R}^2} \overline{(\mathcal{F}f)(u)} (\mathcal{F}g)(u)\, \lambda^2(du) \tag{B.12}$$

for all complex valued functions $f, g \in L^2(\lambda^2)$. Lemma B.6 establishes that an analogous version also holds for $g$ replaced by a measure $\mu$ under certain assumptions. The proof of Lemma B.6 demonstrates how to use the well-known result (B.12) in order to verify the claim (B.13).

**Lemma B.6.** *Let $\mu$ be a finite positive measure on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ and $f : \mathbb{R}^2 \to \mathbb{R}$ be a measurable bounded function, such that $\mathcal{F}\mu \in L^2(\lambda^2)$ and $f \in L^1(\lambda^2)$ hold. Then we have*

$$\int f \, d\mu = \frac{1}{4\pi} \int (\mathcal{F}f)(-u)(\mathcal{F}\mu)(u) \, \lambda^2(du). \tag{B.13}$$

*Proof.* First, note that the integral of the right-hand side in (B.13) exists:

$$|f|^2 \leqslant \left( \sup_{x \in \mathbb{R}^2} |f(x)| \right) \cdot |f| \in L^1(\lambda^2)$$

implies $f \in L^2(\lambda^2)$ and therefore $\mathcal{F}f \in L^2(\lambda^2)$. This yields together with the Hölder inequality

$$u \mapsto (\mathcal{F}f)(-u)(\mathcal{F}\mu)(u) \in L^1(\lambda^2).$$

The proof is divided into two steps. We assume in the first step that $f$ is continuous and, finally, drop this restriction in the second step.

STEP 1. Suppose, additionally, that $f$ is continuous. Set

$$\mu_n \overset{\text{def}}{=} \mu * \left( N\left(0, \frac{1}{n}\right) \otimes N\left(0, \frac{1}{n}\right) \right), \quad n \in \mathbb{N}$$

and note that we have

$$\int |\mathcal{F}\mu_n| \, d\lambda^2 = \int |\mathcal{F}\mu|(u) e^{-\frac{|u|^2}{2n}} \, \lambda^2(du) < \infty, \quad n \in \mathbb{N}.$$

Hence, Sato [38][Proposition 2.5 (xii)] yields the existence of a bounded and continuous $\lambda^2$ density

$$f_n \overset{\text{def}}{=} \frac{d\mu_n}{d\lambda^2} \in L^1(\lambda^2), \quad n \in \mathbb{N}.$$

The boundedness of $f$ and $f_n$ implies $f, f_n \in L^2(\lambda^2)$, so that the classical Plancherel theorem (B.12) yields

$$\int f \, d\mu_n = \int f f_n \, d\lambda^2 = \frac{1}{4\pi} \int \overline{\mathcal{F}f} \mathcal{F}f_n \, d\lambda^2, \quad n \in \mathbb{N}. \tag{B.14}$$

Furthermore, the pointwise convergence

$$\mathcal{F}\mu_n = \mathcal{F}\mu \cdot \mathcal{F}\left( N\left(0, \frac{1}{n}\right) \otimes N\left(0, \frac{1}{n}\right) \right) \to \mathcal{F}\mu, \quad n \to \infty$$

implies $\mu_n \overset{d}{\to} \mu$, i.e.

$$\int f \, d\mu_n \to \int f \, d\mu, \quad n \to \infty. \tag{B.15}$$

Finally, dominated convergence and $|\overline{\mathcal{F}f}\mathcal{F}\mu_n| \leqslant |\overline{\mathcal{F}f}\mathcal{F}\mu| \in L^1(\lambda^2)$ yield

$$\int \overline{\mathcal{F}f} \mathcal{F}f_n \, d\lambda^2 = \int \overline{\mathcal{F}f} \mathcal{F}\mu_n \, d\lambda^2 \to \int \overline{\mathcal{F}f} \mathcal{F}\mu \, d\lambda^2, \quad n \to \infty.$$

This accomplishes together with (B.14), (B.15) and $\overline{\mathcal{F}f}(\cdot) = \mathcal{F}f(-\cdot)$ step 1.

STEP 2. Next, we drop the restriction on $f$ to be continuous. Note that the space $\mathcal{C}_c$ of continuous functions with compact support is a dense subspace of $L^1(\lambda^2 + \mu)$, cf. Rudin [37][Theorem 3.14]. Hence, there is a sequence of functions $(g_n)_n \subseteq \mathcal{C}_c$ with $g_n \to f$ in $L^1(\lambda^2 + \mu)$. We can assume furthermore $\|g_n\|_\infty \leqslant \|f\|_\infty$, $n \in \mathbb{N}$ because $f$ is bounded. This yields in particular

$$g_n \to f \quad \text{in } L^1(\mu) \text{ and } L^2(\lambda^2). \tag{B.16}$$

Since $\mathcal{F}$ is an isometric $L^2(\lambda^2)$ isomorphism, we, in particular, have $\mathcal{F}g_n \to \mathcal{F}f$ in $L^2(\lambda^2)$ which implies by using of the Hölder inequality

$$\int \mathcal{F}g_n(-\cdot)\mathcal{F}\mu \, d\lambda^2 \to \int \mathcal{F}f(-\cdot)\mathcal{F}\mu \, d\lambda^2, \quad n \to \infty. \tag{B.17}$$

Finally, (B.16) and (B.17) prove together with step 1 the claim of this lemma. □

# Appendix C

# Matlab listings

In the following, the Matlab source code files for the numerical simulations in Chapter 1, Section 7 are provided. The files are listed in alphabetical order and separated into Matlab macros and Matlab function files. Further, a concise description of the meaning of the respective files is provided. Every figure in Chapter 1.7 is created using exactly one macro file which is based on the stated function files:

GApprox.m, GApproxJmp.m, errorTypes.m, GBNSCompRef.m, GBNSComp.m, realTest.m

create Figures 1.2 up to Figure 1.7, in this order. An execution of the file doAll.m creates all figures in Chapter 1.7. However, the needed calculations may take several hours on a customary personal computer with our parameter setting.

Explanations of the function files are as follows:

- `approxBNScdf.m` calculates the empirical distribution function of the test statistic by Barndorff -Nielsen and Shephard. For this purpose, we sample the underlying unit interval with the grid fineness $\frac{1}{n^2}$ and use the jump diffusion process in Chapter 1.7 with the parameters `lambda`, $k = 10L^2$ and $\theta = \frac{1}{10L \cdot n}$. Thus, the jump distribution has expectation $\frac{L}{n}$ and variance $\frac{1}{10n^2}$. Finally, to get the empirical distribution function, the test statistic is evaluated $m$ times, based on $m$ independent simulations of the underlying jump diffusion process. The result is exported in the vector $r$ which is, of course, discretized, i.e. we obtain $r$ as a function on $-I, -I + dt, -I + 2dt, \cdots, 2I$ with a suitable stepsize $dt$.

- `approxGcdf.m` proceeds in the same way as approxBNScdf.m with the Gumbel statistic instead of then Barndorff-Nielsen and Shephard statistic.

- `CGamPoiss.m` simulates a compound Poisson process with a Gamma distributed jump distribution. Here, `lambda` is the intensity of the process and $k$, `theta` are the parameters of the Gamma distribution. Again, $\frac{1}{n^2}$ describes the sampling step size.

- `getQuantile.m` calculates the quantile function, i.e. the pseudo-inverse of a given distribution function $F$ which is defined on the discretized interval $[-I, I]$ with step size $dt$.

- `OU.m` simulates an Ornstein-Uhlenbeck process with the parameters: `a` starting point; `mu` mean reversion level; `theta` mean reversion speed; `sigma` volatility. The parameter $n$ is again used for the grid fineness.

- `powerBNS.m` calculates the power of the Barndorff-Nielsen and Shephard test. For this purpose, it is assumed that the test statistic has the finite sample distribution given by the quantile function q, if there are no jumps. Based on $m$ independent simulations with the grid fineness $\frac{1}{n^2}$, the power concerning the significance levels $a, a+dt, a+2dt, \ldots, b$ is calculated. The underlying process is again the jump diffusion process in Chapter 1.7 with the jump parameters $L$ and `lambda`, cf. approxBNScdf.m.

- `powerG.m` is the same as powerBNS.m, but based on the Gumbel test.
- `testBNS.m` applies the Barndorf-Nielsen and Shephard test $m$-times on $m$ independently simulated jump diffusion processes or, alternatively, on a real dataset $Y$. Again, $n$ is used for the grid fineness and $L$, `lambda` is used for the compound Poisson jump process.
- `testG.m` is an analogue to testBNS.m, but based on the Gumbel test.
- `Yf.m` simulates the underlying jump diffusion processes.

# Matlab macros

## errorTypes.m

```
%m1 = 10000;
m2 = 2000;
n = 50;
I = 20;
z = (-I:0.0001:I);

alpha = (0:0.001:0.1);

G = exp(-exp(-z));
F = normcdf(z);
qG = getQuantile(G,I,0.0001);
qF = getQuantile(F,I,0.0001);

error1G = powerG(n,m2,qG,0,0.1,0.01,0,0);                   %Type I error
error1F = powerBNS(n,m2,qF,0,0.1,0.01,0,0);

error2G = 1 - powerG(n,m2,qG,0,0.1,0.01,4,10);              %Type II error
error2F = 1 - powerBNS(n,m2,qF,0,0.1,0.01,4,10);

figure(3);

subplot(2,2,1);
plot(alpha,error1G,'.','Color','k');
hold on;
plot(alpha,alpha,'Color','red');
title('no jumps, Gumbel');
subplot(2,2,2);
plot(alpha,error1F,'.','Color','k');
hold on;
plot(alpha,alpha,'Color','red');
title('no jumps, Barndorff-Nielsen');

subplot(2,2,3);
plot(alpha,error2G,'.','Color','k'); title('jumps, Gumbel');
subplot(2,2,4);
plot(alpha,error2F,'.','Color','k'); title('jumps, Barndorff-Nielsen');

set(gcf,'PaperPositionMode','Auto');
```

## GApprox.m

```
m = 10000;
I = -3;
J = 7;

x = (I:0.001:J);
y = exp(-exp(-x));
```

```
F20 = approxGcdf(20,m,0.001,I,J,0,0);
F50 = approxGcdf(50,m,0.001,I,J,0,0);
F200 = approxGcdf(200,m,0.001,I,J,0,0);
F1000 = approxGcdf(1000,m,0.001,I,J,0,0);

figure(1);

subplot(2,2,1); plot(x,F20,'r',x,y,'k'); title('n = 20');
subplot(2,2,2); plot(x,F50,'r',x,y,'k'); title('n = 50');
subplot(2,2,3); plot(x,F200,'r',x,y,'k'); title('n = 200');
subplot(2,2,4); plot(x,F1000,'r',x,y,'k'); title('n = 1000');

set(gcf,'PaperPositionMode','Auto');
```

## GApproxJmp.m

```
m = 10000;
I = -3;

x_1 = (I:0.001:10);
x_2 = (I:0.001:10);
x_3 = (I:0.001:15);
x_4 = (I:0.001:20);

y_1 = exp(-exp(-x_1));
y_2 = exp(-exp(-x_2));
y_3 = exp(-exp(-x_3));
y_4 = exp(-exp(-x_4));


F50_1 = approxGcdf(50,m,0.001,I,10,1,5);
F50_2 = approxGcdf(50,m,0.001,I,10,2,5);
F50_3 = approxGcdf(50,m,0.001,I,15,3,5);
F50_4 = approxGcdf(50,m,0.001,I,20,4,5);

figure(2);

subplot(2,2,1); plot(x_1,F50_1,'r',x_1,y_1,'k'); title('L = 1');
subplot(2,2,2); plot(x_2,F50_2,'r',x_2,y_2,'k'); title('L = 2');
subplot(2,2,3); plot(x_3,F50_3,'r',x_3,y_3,'k'); title('L = 3');
subplot(2,2,4); plot(x_4,F50_4,'r',x_4,y_4,'k'); title('L = 4');

set(gcf,'PaperPositionMode','Auto');
```

## GBNSComp.m

```
m1 = 10000;
m2 = 2000;
n = 50;
I = 20;

alpha = (0:0.001:0.1);

F = approxBNScdf(n,m1,0.0001,I,0,0);
qF = getQuantile(F,I,0.0001);

G = approxGcdf(n,m1,0.0001,-I,I,0,0);
qG = getQuantile(G,I,0.0001);

pF1 = powerBNS(n,m2,qF,0,0.1,0.01,4,2);
pF2 = powerBNS(n,m2,qF,0,0.1,0.01,4,15);

pG1 = powerG(n,m2,qG,0,0.1,0.01,4,2);
pG2 = powerG(n,m2,qG,0,0.1,0.01,4,15);

figure(5);

subplot(1,2,1);
plot(alpha,pF1,'.','Color','r');
hold on;
plot(alpha,pG1,'o','Color','k');
title('\lambda = 2');
leg = legend('Barndorff-Nielsen','Gumbel');
set(leg,'Location','SouthEast');

subplot(1,2,2);
plot(alpha,pF2,'.','Color','r');
hold on;
plot(alpha,pG2,'o','Color','k');
title('\lambda = 15');
leg = legend('Barndorff-Nielsen','Gumbel');
set(leg,'Location','SouthEast');

set(gcf,'PaperPositionMode','Auto');
```

## GBNSCompRef.m

```
m1 = 10000;
m2 = 2000;
n = 50;
I = 20;

alpha = (0:0.001:0.1);
x = (0:0.0001:0.1);

F = approxBNScdf(n,m1,0.0001,I,0,0);
qF = getQuantile(F,I,0.0001);

G = approxGcdf(n,m1,0.0001,-I,I,0,0);
```

```
qG = getQuantile(G,I,0.0001);

pF = powerBNS(n,m2,qF,0,0.1,0.01,0,0);
pG = powerG(n,m2,qG,0,0.1,0.01,0,0);

figure(4);

subplot(1,2,1);
plot(alpha,pF,'.','Color','k');
hold on;
plot(x,x,'Color','r');
title('Barndorff-Nielsen');

subplot(1,2,2);
plot(alpha,pG,'.','Color','k');
hold on;
plot(x,x,'Color','r');
title('Gumbel');

set(gcf,'PaperPositionMode','Auto');
```

## realTest.m

```
data = xlsread('LogEWI104s_nonstandardized.xls');
usa = data(:,1)';
ireland = data(:,14)';

usa = [0 cumsum(usa)];
ireland = [0 cumsum(ireland)];

N = 93^2;
usa = usa(1:N+1);
ireland = ireland(1:N+1);

uG = exp(-exp(-testG(93,1,-1,-1,usa)));
iG = exp(-exp(-testG(93,1,-1,-1,ireland)));

uGMin = exp(-exp(-testG(93,1,-1,-1,-usa)));
iGMin = exp(-exp(-testG(93,1,-1,-1,-ireland)));

uBNS = normcdf(testBNS(93,1,-1,-1,usa));
iBNS = normcdf(testBNS(93,1,-1,-1,ireland));

figure(6);

x = 1973.6+(0:1/N:1)*(2005.6-1973.6);
plot(x,ireland,'k',x,usa,'r');
leg = legend('ireland','usa');
set(leg,'Location','SouthEast');
axis([1973.6 2005.6 -1 6])


%---- Print the test results ---
```

```
iG
iGMin
uG
uGMin
iBNS
uBNS
%-------------------------------


set(gcf,'PaperPositionMode','Auto');
```

## doAll.m

```
GApprox;
GApproxJmp;
errorTypes;
GBNSCompRef;
GBNSComp;
realTest;
```

# Matlab functions

### approxBNScdf.m

```
function [r] = approxBNScdf(n,m,dt,I,L,lambda)

T = testBNS(n,m,L,lambda);

T = sort(T);                            %empirical distribution function
r = zeros(1,2*I/dt+1);

for j = (1:m)
    v = ceil((T(j)+I)/dt);
    if (v <= 2*I/dt)
        r((max(v+1,1):end)) = r((max(v+1,1):end)) + 1;
    end
end

r = r/m;
r(end) = 1;
```

### approxGcdf.m

```
function [r] = approxGcdf(n,m,dt,I,J,L,lambda)

T = testG(n,m,L,lambda);

T = sort(T);                            %empirical distribution function
r = zeros(1,(J-I)/dt+1);

for j = (1:m)
    v = ceil((T(j)-I)/dt);
    if (v <= (J-I)/dt)
        r((max(v+1,1):end)) = r((max(v+1,1):end)) + 1;
    end
end

r = r/m;
r(end) = 1;
```

### CGamPoiss.m

```
function [r] = CGamPoiss(n,lambda,k,theta)

N = n^2;
J = zeros(1,N+1);                                   %n^2+1 row vector
jQuant = poissrnd(lambda);                          %quantity
jPos = sort(rand(1,jQuant),2);                      %position
```

```
jSize = gamrnd(k,theta,1,jQuant);                    %size

jCumSize = cumsum(jSize);
pos = [ceil(jPos*N) N+1];
for k = (1:jQuant)
    for h = (pos(k)+1:pos(k+1))
        J(h) = jCumSize(k);
    end
end

r = J;
```

## getQuantile.m

```
function [r] = getQuantile(F,I,dt)

m = 1/dt;
q = zeros(1,m+1);
j = 0;
for k = (0:m)
    while ((j+1 <= size(F,2)) && (F(j+1) < k/m))
        j = j+1;
    end
    q(k+1) = j;
end

r = 2*I*q/(size(F,2)-1) - I;

r(1) = -inf;
r(end) = inf;
```

## OU.m

```
function [r] = OU(n,a,mu,theta,sigma)
N = n^2;
dt = 1/N;
dW = sqrt(dt)*randn(1,N);
z = (0:1/N:1);
x = a * exp(-theta*z) + mu*(1-exp(-theta*z));   %deterministic part


J = sigma*exp(theta*z(1:end-1)).*dW;            %stochastic integral
J = cumsum(J);

J = J.*exp(-theta*z(2:end));

r = x + [0 J];
```

## powerBNS.m

```
function [r] = powerBNS(n,m,q,a,b,dt,L,lambda)

A = 1/dt;
v = zeros(1,A+1);

for i = (0:A)

    alpha = i * dt * (b-a) + a;

    Q = q(floor(alpha*(size(q,2)-1) + 1));
    T = testBNS(n,m,L,lambda);

    v(i + 1) = sum(T >= Q) / m;

end

r = 1 - v;
```

## powerG.m

```
function [r] = powerG(n,m,q,a,b,dt,L,lambda)

A = 1/dt;
v = zeros(1,A+1);

for i = (0:A)

    alpha = i * dt * (b-a) + a;

    Q = q(floor((1-alpha)*(size(q,2)-1) + 1));
    T = testG(n,m,L,lambda);

    v(i + 1) = sum(T <= Q) / m;

end

r = 1 - v;
```

## testBNS.m

```
function [r] = testBNS(n,m,L,lambda,Y)

theta = pi^2/4 + pi - 5;
mu1 = sqrt(2/pi);

T = zeros(1,m);
for j = (1:m)                       %produces m values of the statistic
```

```
    if (lambda >= 0)
            Y = Yf(n,L,lambda);
    end
    dY = abs(Y(2:end)-Y(1:end-1));

    dY2 = dY*dY';
    dY11 = dY(1:end-1)*dY(2:end)';
    dY1111 = sum(dY(1:end-3).*dY(2:end-2).*dY(3:end-1).*dY(4:end));

    T(j) = (mu1^(-2)*dY11 - dY2)/sqrt(theta*mu1^(-4)*dY1111);
end

r = T;
```

## testG.m

```
function [r] = testG(n,m,L,lambda,Y)

aa = 2*sqrt(log(n));
bb = aa - (log(log(n)) + log(8*pi))/(4*sqrt(log(n)));

T = zeros(1,m);

for j = (1:m)                       %produces m values of the statistic
        if (lambda >= 0)
            Y = Yf(n,L,lambda);
        end
        dY = Y(2:end)-Y(1:end-1);
        sigma = zeros(1,n^2);
        for k = (1:n)
            dYpartial = abs(dY((k-1)*n + 1 : k*n));
            sigma(((k-1)*n + 1 : k*n)) = dYpartial(1:end-1)*dYpartial(2:end)';
        end
        sigma = sqrt(pi/(2*(n-1))*sigma);
        dY = dY./sigma;
        T(j) = aa*(max(dY) - bb);
end

r = T;
```

## Yf.m

```
function [r] = Yf(n,L,lambda)

N = n^2;
dt = 1/N;
dW = sqrt(dt)*randn(1,N);
```

```
sigma = max(OU(n,1,1,0.5,0.2),0.1);
y = cumsum(sigma(1:end-1).*dW);

J = 0;
if (L > 0)                        %mean = L/n , variance = 1/(10n^2)
    J = CGamPoiss(n,lambda,10*L^2,1/(10*L*n));
end


r = [0 y] + J;
```

# Bibliography


[1] Andrews, G. E., Askey, R., Roy, R. (2001) Special Functions. *Cambridge University Press.*

[2] Arnold, V. I. (2006) Ordinary Differential Equations. *Springer Press.*
Translated from the 3th Russian original edition.

[3] Barndorff-Nielsen, O. E., Shephard, N. (2006). Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation. *Journal of Financial Econometrics* **4**, 1-30.

[4] Barndorff-Nielsen, O. E., Shephard, N. (2003). Realized power variation and stochastic volatility models. *Bernoulli* **9 (2)**, 243-265.

[5] Bauer, H. (2001). Measure and Integration Theory. *Springer Press.*

[6] Billingsley, P. (1999). Convergence of Probability Measures, 2. Edition. *Wiley-Interscience Press.*

[7] Buchmann, B., Grübel, R. (2003). Decompounding: An estimation problem for Poisson random sums. *The Annals of Statistics* **31 (4)**, 1054-1074.

[8] Bücher Axel, Vetter, M. (2012). Nonparametric inference on Lévy measures and copulas. *Unpublished*, arXiv:1205.0417

[9] Chow, Y. S., Teicher, H. (2003). Probability Theory: Independence, Interchangeability, Martingales, 3. Edition. *Springer Texts in Statistics.*

[10] Chung, K. L. (2000). A Course in Probability Theory, 2. Edition. *Academic Press.*

[11] Cont, R., Tankov, P. (2003). Financial Modelling with Jump Processes. *Chapman & Hall/CRC Financial Mathematics Series.*

[12] Delbaen, F., Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Math. Ann.* **300**, 463-520.

[13] Delbaen, F., Schachermayer, W. (1998) The fundamental theorem of asset pricing for unbounded stochastic processes. *Math. Ann.* **312**, 215-250.

[14] Dudley, R. M. (1999). Uniform Central Limit Theorems. *Cambridge Studies in advanced mathematics*

[15] Fan, J., Wang, W. (2008). Spot volatility estimation for high-frequency data. *Statistics and its interface.* **1**, 279-288.

[16] FERMANIAN, J. D., RADULOVIĆ, D., WEGKAMP M. (2004) Weak convergence of empirical copula processes. *Bernoulli* **10 (5)**, 847-860.

[17] HAAN, L., FERREIRA, A. (2006). Extreme Value Theory: An Introduction. *Springer Press.*

[18] HOFFMANN, M., MUNK, A., SCHMIDT-HIEBER, J. Adaptive wavelet estimation of the diffusion coefficient under additive error measurements. *To appear in Annales de l'Institut Henri Poincaré*, arXiv:1007.4622

[19] IGNATIEVA, K., PLATEN, E. (2012). Estimating the diffusion coefficient function for a diversified world stock index. *Computational Statistics and Data Analysis.* **56 (6)**, 1333-1349.

[20] JACOD, J., SHIRYAEV, A. N. (2002). Limit theorems for stochastic processes, 2. Edition. *Springer Press.*

[21] KALLSEN, J., TANKOV P. (2006). Characterization of dependence of multidimensional Lévy processes using Lévy copulas. *Journal of Multivariate Analysis.* **97**, 1551-1572.

[22] KARATZAS, I., SHREVE, S. E. (1991). Brownian Motion and Stochastic Calculus, 2. Edition. *Springer Press.*

[23] KORN, R., KORN E. (2001). Options Pricing and Portfolio Optimization. Modern Methods of Financial Mathematics. *Oxford University Press.*

[24] KRAJINA, A., LAEVEN, R. Modeling Jump Dependence using Lévy copulas. *http://www.eurandom.nl/events/workshops/2011/ISI_MRM/Presentation/Krajina.pdf*

[25] LAEVEN, R. (2009). Non-parametric Estimation of the Lévy Copula. *http://www.eurandom.tue.nl/events/workshops/2009/Levy_processes/Laeven.pdf.*

[26] LEE, S. S., MYKLAND, P. A. (2008). Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics. *The Review of Financial Studies* **v 21 n 6**, 2535-2563.

[27] MILLAR, P. W. (1968). Martingale integrals. *Trans. Amer. Math. Soc.* **133**, 145-166.

[28] NELSEN, R. (2005). An Introduction to Copulas, 2. Edition. *Springer Press.*

[29] NEUMANN, M. H., REISS, M. (2009). Nonparametric estimation for Lévy processes from low-frequency observations. *Bernoulli* **15 (1)**, 223-248.

[30] NICKL, R., REISS, M. (2012). A Donsker Theorem for Lévy Measures. *To appear in Journal of Functional Analysis*, arXiv:1201.0590

[31] NOVIKOV, A. A. (1971). On moment inequalities for stochastic integrals. *Theory Probab. Appl.* **16** 538-541.

[32] ØKSENDAL, B. (2000). Stochastic Differential Equations, 6. Edition. *Springer Press.*

[33] PLATEN, E., RENDEK, R.J. (2008). Empirical evidence on Student-t log-returns of diversified world stock indices. *Journal of Statistical Theory and Practice.* **2 (2)** 233-251.

[34] POLLARD, D. (1984). Convergence of Stochastic Processes. *Springer Series in Statistics.*

[35] Protter, P. E. (2005). Stochastic Integration and Differential Equations, 2. Edition. *Springer Press.*

[36] Rudin, W. (1973). Functional Analysis. *McGraw-Hill series in higher mathematics.*

[37] Rudin, W. (1986). Real and Complex Analysis, 3. Edition. *McGraw-Hill Book Company.*

[38] Sato, K. I. (1999). Lévy processes and infinitely divisible distributions. *Cambridge studies in advances mathematics* **68**.

[39] Schicks, M. (2009). Non-Parametric Estimation for Multivariate Compound Poisson Processes. *Unpublished.*

[40] Sempi, C. (2004). Convergence of copulas: Critical remarks. *Radovi Matematički* **12**, 241-249.

[41] Shiryaev, A. N. (1999). Essentials of Stochastic Finance. *Advanced Series on Statistical Science & Applied Probability.* Vol. **3**.

[42] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Inst. Statist. Univ. Paris* **8**, 229-231.

[43] van der Vaart, A. W. (1997). Asymptotic statistics. *Cambridge series in statistical and probabilistic mathematics.*

[44] van der Vaart, A. W., Wellner, J. A. (2000). Weak Convergence and Empirical Processes. *Springer Series in Statistics.*