



HAL
open science

Total variation distance for discretely observed Lévy processes: a Gaussian approximation of the small jumps

Alexandra Carpentier, Céline Duval, Ester Mariucci

► To cite this version:

Alexandra Carpentier, Céline Duval, Ester Mariucci. Total variation distance for discretely observed Lévy processes: a Gaussian approximation of the small jumps. 2018. hal-01891247v1

HAL Id: hal-01891247

<https://hal.science/hal-01891247v1>

Preprint submitted on 9 Oct 2018 (v1), last revised 28 Nov 2018 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Total variation distance for discretely observed Lévy processes: a Gaussian approximation of the small jumps

Alexandra Carpentier*
Céline Duval[†] and Ester Mariucci[‡]

Abstract

It is common practice to treat small jumps of Lévy processes as Wiener noise and thus to approximate its marginals by a Gaussian distribution. However, results that allow to quantify the goodness of this approximation according to a given metric are rare. In this paper, we clarify what happens when the chosen metric is the total variation distance. Such a choice is motivated by its statistical interpretation. If the total variation distance between two statistical models converges to zero, then no tests can be constructed to distinguish the two models which are therefore equivalent, statistically speaking. We elaborate a fine analysis of a Gaussian approximation for the small jumps of Lévy processes with infinite Lévy measure in total variation distance. Non asymptotic bounds for the total variation distance between n discrete observations of small jumps of a Lévy process and the corresponding Gaussian distribution is presented and extensively discussed. As a byproduct, new upper bounds for the total variation distance between discrete observations of Lévy processes are provided. The theory is illustrated by concrete examples.

Keywords. Lévy processes, Total variation distance, Small jumps, Gaussian approximation, Statistical test.
AMS Classification. Primary: 60G51, 62M99. Secondary: 60E99.

1 Introduction

Although Lévy processes, or equivalently infinite divisible distributions, are mathematical objects introduced almost a century ago and even though a good knowledge of their basic properties has since long been achieved, they have recently enjoyed renewed interest. This is mainly due to the numerous applications in various fields. To name some examples, Lévy processes or Lévy-type processes (time changed Lévy processes, Lévy driven SDE, etc...) play a central role in mathematical finance, insurance, telecommunications, biology, neurology, telecommunications, seismology, meteorology and extreme value theory. Examples of applications may be found in the textbooks [2] and [6] whereas the manuscripts [4] and [17] provide a comprehensive presentation of the properties of these processes.

The transition from the purely theoretical study of Lévy processes to the need to understand Lévy driven models used in real life applications has led to new challenges. For instance, the questions of how to simulate the trajectories of Lévy processes and how to make inference (prediction, testing, estimation, etc...) for this class of stochastic processes have become a key issue. The literature concerning these two aspects is already quite large; without any claim of completeness, we quote [1], Chapter VI in [2], [3], [5] and Part II in [6].

We specifically focus on statistics and simulation for Lévy processes, because our paper aims to give an exhaustive answer to a recurring question in these areas: When are the small jumps of Lévy

*Otto von Guericke Universität Magdeburg, Germany. E-mail: alexandra.carpentier@ovgu.de.

†Université Paris Descartes, Laboratoire MAP5, UMR CNRS 8145, France. E-mail: celine.duval@parisdescartes.fr.

‡Universität Potsdam, Germany. E-mail: mariucci@uni-potsdam.de.

processes morally Gaussian?

Before entering into the details, let us take a step back and see where this question comes from. Thanks to the Lévy-Itô decomposition, the structure of the paths of any Lévy process is well understood and it is well known that any Lévy process X can be decomposed into the sum of three independent Lévy processes: a Brownian motion with drift, a centered martingale M associated with the small jumps of X and a compound Poisson process describing the big jumps of X (see the decomposition (2) in Section 1.1 below). If the properties of continuously or discretely observed compound Poisson processes and of Gaussian processes are well understood, the same cannot be said for the small jumps M . As usual in mathematics, when one faces a complex object a natural reflection is whether the problem can be simplified by replacing the difficult part with an easier but, in a sense, equivalent one. There are of course various notions of equivalence ranging from the weakest, convergence in law, to the stronger convergence in total variation.

For some time now, many authors have noticed that marginal laws of small jumps of Lévy processes with infinite Lévy measures resemble to Gaussian random variables, see Figure 1 and 2. This remark

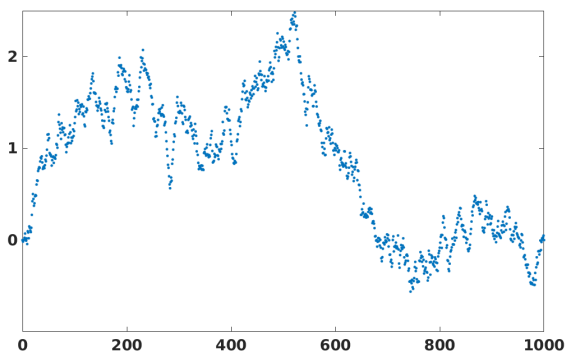


Figure 1: Discretized trajectory of a Lévy process $(0, 0, \frac{1_{(0,\varepsilon]}(x)}{x^{1+\beta}})$ for $(n = 10^3, \Delta = 1, \varepsilon = 10^{-2}, \beta = 0.9)$.

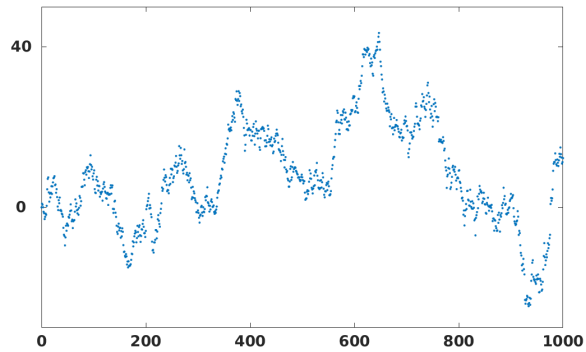


Figure 2: Discretized trajectory of a Lévy process $(0, 0, \frac{1_{[-\varepsilon,\varepsilon] \setminus \{0\}}(x)}{x^{1+\beta}})$ for $(n = 10^3, \Delta = 1, \varepsilon = 0.5, \beta = 1.8)$.

has led to propose algorithms of simulation of trajectories of Lévy processes based on a Gaussian approximation of the small jumps, see e.g. [5] or [6], Chapter 6. Regarding estimation procedures, a Gaussian approximation of the small jumps has, to the best of our knowledge, not been exploited yet. A fine control of the total variation distance between these two quantities could open the way of new statistical procedures.

In this paper we would like to shed some light on when a Gaussian approximation of the small jumps is adequate. To do so we carefully measure the distance in total variation between the small jumps and their corresponding Gaussian distribution. The choice of this distance is justified by its statistical interpretation: if the total variation distance between the law of the small jumps and the corresponding Gaussian component converges to zero then no statistical test can be built to distinguish between the two models. Rephrased in terms of information theory, this means that the two models are asymptotically equally informative.

Clearly, investigating the goodness of a Gaussian approximation of the small jumps of a Lévy process in total variation distance makes sense only if one deals with the discrete observations of the process. Indeed, if one disposes of continuous observations of two Lévy processes, the problem of separating the continuous part from the jumping part does not arise: the jumps are observed. It has

long been known that the measure corresponding to the continuous observation of a continuous Lévy process is orthogonal with respect to the measure corresponding to the continuous observation of a Lévy process with non trivial jump part, see e.g. [11]. However, the situation changes when dealing with discrete observations. In this case, the matter of disentangling continuous and discontinuous part of the processes is much more complex: intuitively, fine techniques are needed to understand whether, between two observations X_{t_0} and X_{t_1} , there has been a chaotic continuous behavior, many small jumps, one single bigger jump, or a mixture of these.

A criterion for the weak convergence for marginals of Lévy processes is given by Gnedenko and Kolmogorov, [9]:

Theorem 1 (Gnedenko, Kolmogorov). *Marginals of Lévy processes $X^n = (X_t^n)_{t \geq 0}$ with Lévy triplets (b_n, σ_n^2, ν_n) converge weakly to marginals of a Lévy process $X = (X_t)_{t \geq 0}$ with Lévy triplet (b, σ^2, ν) if and only if*

$$b_n \rightarrow b \text{ and } \sigma_n^2 \delta_0 + (x^2 \wedge 1) \nu_n(dx) \xrightarrow{w} \sigma^2 \delta_0 + (x^2 \wedge 1) \nu(dx),$$

where δ_0 is the Dirac measure in 0 and \xrightarrow{w} denotes weak convergence of finite measures.

A remarkable fact in the previous statement is the non-separation between the continuous and discontinuous parts of the processes: the law at time t of a pure jumps Lévy process can weakly converge to that of a continuous Lévy process. In particular, if X is a Lévy process with Lévy measure ν then, for any $\varepsilon \in (0, 1)$ and $t > 0$, the law of the centered jumps of X_t with magnitude less than ε converges weakly to a centered Gaussian distribution with variance $t\sigma^2(\varepsilon) := t \int_{|x| < \varepsilon} x^2 \nu(dx)$ as $\varepsilon \rightarrow 0$. We aim to understand this phenomenon in depth, using a notion of closeness stronger than the weak convergence, providing a quantitative translation of the result of Gnedenko and Kolmogorov in total variation distance.

There exist already several results for distances between Lévy processes. Most of them (see for example [8], [11] and [12]) are distances on the Skorohod space, that is distances between the continuous observation of the processes, and thus out of the scope of this paper. Concerning discretely observed Lévy processes we mention the results in [13] and [14]. Liese [13] proved the following upper bound in total variation for marginals of Lévy processes $X^j \sim (b_j, \Sigma_j^2, \nu_j)$, $j = 1, 2$: for any $t > 0$

$$\|\mathcal{L}(X_t^1) - \mathcal{L}(X_t^2)\|_{TV} \leq 2 \sqrt{1 - \left(1 - \frac{H^2(\mathcal{N}(\tilde{b}_1, t\Sigma_1^2), \mathcal{N}(\tilde{b}_2, t\Sigma_2^2))}{2}\right)^2} \exp(-tH^2(\nu_1, \nu_2))$$

with $\tilde{b}_1 = b_1 - \int_{-1}^1 x \nu_1(dx)$, $\tilde{b}_2 = b_2 - \int_{-1}^1 x \nu_2(dx)$ and H denotes the Hellinger distance.

This result is the analogous in discrete time of the result of Mémin and Shiryayev for continuously observed Lévy processes, see [15]. This is reflected in the fact that there is a clear separation between the continuous and discontinuous parts of the processes, which is unavoidable on the Skorohod space but that can be relaxed when dealing with the marginals. Clearly, from this kind of upper bounds it is not possible to deduce a Gaussian approximation for the small jumps in total variation: the bound is actually trivial whenever $tH^2(\nu_1, \nu_2) > 1$. However, such an approximation may hold even in total variation as proved in [14].

More precisely, in [14] the convolutional structure of Lévy processes with non-zero diffusion coefficients is exploited to transfer results from Wasserstein distances to total variation distance. However, it turns out that the Gaussian approximation of the small jumps thus obtained is suboptimal for large classes of Lévy processes, for instance for Lévy processes with infinite Lévy measures that are absolutely continuous with respect to the Lebesgue measure, see Section 2.1.3.

In the present paper we complete the work started in [14], providing a comprehensive answer to the following question: Under which asymptotics does a Gaussian approximation capture the behavior

of the small jumps of a discretely observed Lévy process adequately so that the two corresponding statistical models cannot be distinguished, they are equally informative?

Our analysis is developed for Lévy processes X with Lévy measures ν which are supposed to be infinite and absolutely continuous with respect to the Lebesgue measure. If ν is finite and non-zero, X is the convolution between a compound Poisson process and a Brownian motion plus drift. Several tools are already present in the literature to deal with this kind of processes, hence we do not focus on them. The absolute continuity of ν with respect to the Lebesgue measure is one way to avoid the orthogonality between the law of the small jumps and a Gaussian distribution. Indeed, if ν is concentrated on a finite set, then the total variation distance between X_t and any Gaussian distribution is always 1.

Differently from what was done in [14], we do not only consider the total variation distance between marginals of Lévy processes but we also establish sharp bounds for the distance between n given increments of the small jumps. This is an essential novelty. If it is true that from a bound in total variation between marginals one can always deduce a bound for the n sample by $\|P^{\otimes n} - Q^{\otimes n}\|_{TV} \leq \sqrt{2n}\|P - Q\|_{TV}$, this kind of control is in general sub-optimal. In many situations indeed, $\|P^{\otimes n} - Q^{\otimes n}\|_{TV}$ is of order of $\sqrt{n}\|P - Q\|_{TV}$, as n goes to infinity. This faster rate can indeed be obtained using the methods developed in the present work.

More concretely, our main result about the Gaussian approximation of the small jumps in total variation may be stated as follows:

Main Result 1. *Let $M(\varepsilon)$ be a Lévy process with Lévy triplet $(0, 0, \nu_\varepsilon)$ where ν_ε is an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure with support in $[-\varepsilon, \varepsilon]$. Introducing $\sigma^2(\varepsilon) = \int_{-\varepsilon}^{\varepsilon} x^2 \nu_\varepsilon(dx)$, $\mu_3(\varepsilon) = \int_{-\varepsilon}^{\varepsilon} x^3 \nu_\varepsilon(dx)$ and $\mu_4(\varepsilon) = \int_{-\varepsilon}^{\varepsilon} x^4 \nu_\varepsilon(dx)$, for any $0 < \varepsilon < 1$, $\Delta > 0$, $n \geq 1$ it holds*

$$\begin{aligned} & \|\mathcal{L}((M_{i\Delta}(\varepsilon) - M_{(i-1)\Delta}(\varepsilon)), i = 1 \dots, n) - \mathcal{N}(0, \Delta\sigma^2(\varepsilon))^{\otimes n}\|_{TV} \\ & \leq 3\sqrt{\left(1 + \frac{\mu_4^2(\varepsilon)}{\Delta^2(\sigma^2(\varepsilon))^4} + \frac{7}{6} \frac{\mu_3^2(\varepsilon)}{\Delta(\sigma^2(\varepsilon))^3}\right)^n - 1}. \end{aligned}$$

The non-asymptotic nature of the result above allows to quantify just how “small” must the small jumps be, in terms of the number of observations n and their frequency Δ , in order for it to be close in total variation to the corresponding Gaussian distribution. More precisely, provided $\frac{\mu_4^2(\varepsilon)}{(\sigma^2(\varepsilon))^4} = o(1)$ and $\frac{\mu_3^2(\varepsilon)}{(\sigma^2(\varepsilon))^3} = o(1)$ as $\varepsilon \rightarrow 0$, the rate of convergence in Main Result 1 is of order

$$r := \sqrt{n} \left(\frac{\mu_4(\varepsilon)}{\Delta(\sigma^2(\varepsilon))^2} + \frac{7}{6} \frac{\mu_3(\varepsilon)}{\sqrt{\Delta}(\sigma^2(\varepsilon))^{3/2}} \right). \quad (1)$$

A sufficient condition to ensure $\frac{\mu_4^2(\varepsilon)}{(\sigma^2(\varepsilon))^4} = o(1)$ and $\frac{\mu_3^2(\varepsilon)}{(\sigma^2(\varepsilon))^3} = o(1)$ is that $\frac{\varepsilon}{\sigma(\varepsilon)}$ as $\varepsilon \rightarrow 0$. In that case it is straightforward to see that if $\mu_3(\varepsilon) \neq 0$ then the rate of convergence of r is of order $\frac{\sqrt{n}\varepsilon}{\sqrt{\Delta}\sigma(\varepsilon)}$; when $\mu_3(\varepsilon) = 0$, this can be further improved to $\frac{\sqrt{n}\varepsilon^2}{\Delta\sigma^2(\varepsilon)}$.

To exemplify, consider the small jumps of symmetric β -stable processes, that is Lévy processes with Lévy measure given by $\nu_\varepsilon = \frac{1}{|x|^{1+\beta}} \mathbf{1}_{|x| \leq \varepsilon}$, $\beta \in (0, 2)$. In this case, the rate of convergence of the distance in total variation between the law of n given increments of the small jumps collected at a sampling rate Δ and the corresponding Gaussian distribution is of order $\sqrt{n} \frac{\varepsilon^\beta}{\Delta}$, as $\varepsilon \rightarrow 0$.

The idea behind the proof of Main Result 1 is the following. For any $\eta \in (0, \varepsilon)$ and any $t > 0$, we decompose $M_t(\varepsilon)$ into the sum of $M_t(\eta)$, the marginal of a Lévy process with jumps of size smaller than η , and $R_t(\eta, \varepsilon)$, the marginal of a centered compound Poisson process with jumps of size in $[\eta, \varepsilon]$.

For η small enough $M_t(\eta)$ is negligible (see Proposition 1) and $R_t(\eta, \varepsilon)$ can be approximated by a Gaussian distribution as follows. Firstly we write

$$R_t(\eta, \varepsilon) = \sum_{i=1}^{N_t(\eta, \varepsilon)} Y_i - \mathbb{E} \left[\sum_{i=1}^{N_t(\eta, \varepsilon)} Y_i \right] \approx \sum_{i=1}^{N_t(\eta, \varepsilon)} (Y_i - \mathbb{E}[Y_i]),$$

where $N(\eta, \varepsilon)$ is a Poisson process with intensity $\lambda(\eta, \varepsilon) = \int_{\eta \leq |x| \leq \varepsilon} \nu(dx)$ and the Y_i 's are i.i.d. bounded random variables. The approximation \approx is justified by Proposition 2. To deal with this random sum we separate the cases where $N_t(\eta, \varepsilon)$ is large and where it is not, which occur with very small probability, given that $\lambda_{\eta, \varepsilon} \rightarrow \infty$ as $\eta \rightarrow 0$ by the assumption $\nu(\mathbb{R}) = \infty$. Conditional on $N_t(\eta, \varepsilon)$ being close enough to its mean $t\lambda_{\eta, \varepsilon}$ (and thus large enough), the sum can be proven to be approximately Gaussian by means of an Edgeworth expansion. The main technical step in our proof of Main Result 1, however, is to derive an unconditional bound with explicit constants (see Proposition 3), to obtain a Gaussian approximation of the all of $R_t(\eta, \varepsilon)$. This involves very precise computations in order to get cancellation of the moments and thus the optimal rate of convergence in n .

Main Result 1 can be sharpened for Lévy processes with non-zero Gaussian components obtaining a faster rate of convergence, as follows.

Main Result 2. *With the same notation as in Main Result 1 the following holds. For any $0 < \varepsilon < 1$, $\Delta > 0$, $n \geq 1$, $b \in \mathbb{R}$ and $\Sigma \geq 0$, we have*

$$\begin{aligned} & \left\| (\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(\varepsilon))^{\otimes n} - \mathcal{N}(b\Delta, \Delta(\Sigma^2 + \sigma^2(\varepsilon)))^{\otimes n} \right\|_{TV} \leq \\ & 3 \sqrt{\left(1 + \frac{\mu_4^2(\varepsilon)}{\Delta^2(\Sigma^2 + \sigma^2(\varepsilon))^4} + \frac{7}{6} \frac{\mu_3^2(\varepsilon)}{\Delta(\Sigma^2 + \sigma^2(\varepsilon))^3} \right)^n - 1}. \end{aligned}$$

It can additionally be sharpened by considering separately large and rare jumps, see Corollary 2 - this has an impact in the case where the jumps of size of order ε are very rare.

Main Result 1 is optimal, in the sense that whenever the jumps of size of order ε are not “too rare” and whenever the above quantity r in (1) is larger than a constant, then the upper bound in Main Result 1 is trivial, but the total variation distance can be bounded away from 0 as shown in Main Result 3 below, which is in fact more general as it considers also processes with a drift and a Brownian part.

Main Result 3. *Let ν be an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure, let $b \in \mathbb{R}$ and $\Sigma \geq 0$. For any $n\Delta \geq 1$ and $\varepsilon \in (0, 1)$ such that $\lambda_{\varepsilon/2, \varepsilon} \geq \log(n)/(n\Delta)$, there exists an absolute sequence $\alpha_n \rightarrow 0$ (independent of $b, \nu, \varepsilon, \Sigma, \Delta$) and an absolute constant $C > 0$ (independent of $b, \nu, \varepsilon, \Sigma, \Delta$) such that the following holds:*

$$\begin{aligned} & \left\| (\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(\varepsilon))^{\otimes n} - \mathcal{N}(b\Delta, \Delta(\Sigma^2 + \sigma^2(\varepsilon)))^{\otimes n} \right\|_{TV} \geq \\ & \left\{ 1 - C \left[\frac{\Delta(\Sigma^2 + \sigma^2(\varepsilon))^3}{n\mu_3(\varepsilon)^2} \wedge \frac{\Delta^2(\Sigma^2 + \sigma^2(\varepsilon))^4}{n\mu_4(\varepsilon)^2} \right] - \alpha_n \right\}. \end{aligned}$$

A more technical and general lower bound, that holds without condition on the rarity of the jumps of size of order ε , is also available, see Theorem 5. This lower bound matches in order the general upper bound of Corollary 2 - implying optimality without conditions of our results. The proof of the lower bound for the total variation is based on the construction of a sharp Gaussian test for Lévy processes. This test combines three ideas, (i) the detection of extreme values that are ‘too large’ for being produced by a Brownian motion, (ii) the detection of asymmetries around the drift in the third

moment, and (iii) the detection of too heavy tails in the fourth moment for a Brownian motion. It can be of independent interest as it does not take knowledge on the tested process as parameter and detects optimally the presence of jumps. It uses classical ideas from testing through moments and extreme values [10], and it adapts them to the specific structure of Lévy processes. The closest related work is [16]. We improve on the test proposed there as we go beyond testing based on the fourth moment only, and we tighten the results regarding rare and large jumps.

To sum up, this paper provides a comprehensive answer to the question of when a Gaussian approximation of the small jumps of discretely observed Lévy processes is appropriate, in the following sense. Given the observation of the increments of a Lévy process X , if the number of observations n , the sampling rate $\Delta = \Delta_n$ and the size $\varepsilon = \varepsilon_n$ of the small jumps of X are such that the upper bounds for the total variation distance we give in Theorems 2 and 3 go to 0 as n goes to infinity, then a Gaussian approximation of the small jumps is statistically justified. Indeed, under these circumstances no statistical test can be constructed to distinguish between realizations of increments of the small jumps of X (potentially convoluted with the Gaussian part of X) and realizations of a Gaussian vector with matching mean and variance. Conversely, if n , Δ and ε are such that the upper bounds we provide in Theorems 2 and 3 do not vanish, then the Gaussian approximation of the small jumps is not statistically justifiable. In this case in fact the total variation distance is bounded from below and a statistical test can be constructed to distinguish the small jumps from Gaussian distributions, see Theorems 4 and 5.

An interesting byproduct of Main Results 1 and 2 is Theorem 6 in Section 3 which provides a new upper bound for the total variation distance between n given increments of each pairs of Lévy processes with infinite Lévy measures. A peculiarity of the result is the non-separation between the continuous and discontinuous part of the processes. In this sense Theorem 6 is close in spirit to Theorem 1 although it is worth noting that it gives a concrete measure of the closeness between discretely observed Lévy processes in total variation, that is using a much stronger notion of convergence than the weak convergence.

The paper is organized as follows. In Section 1.1 we fix the notation. Section 2 is devoted to the analysis of the Gaussian approximation of the small jumps of Lévy processes. More precisely, in Section 2.1 we present upper bounds in total variation distance whereas in Section 2.2 we provide lower bounds proving the optimality of our findings. In Section 3 new upper bounds for the total variation distance between n given increments of two general Lévy processes are derived. Most of the proofs are postponed to Section 4. The paper also contains two Appendices: in Appendix A technical results related to the proof of the lower bounds can be found whereas in Appendix B we recall some results about total variation distance and present some general, and probably not new, results about the total variation distance between Gaussian distributions and discrete observations of compound Poisson processes.

1.1 Statistical setting and notation

For X a Lévy process with Lévy triplet (b, Σ^2, ν) (we write $X \sim (b, \Sigma^2, \nu)$), where $b \in \mathbb{R}$, $\Sigma \geq 0$ and ν satisfies $\int (x^2 \wedge 1) \nu(dx) < \infty$, the Lévy-Itô decomposition gives a decomposition of X into the sum of three independent Lévy processes: a Brownian motion with drift b , a centered martingale associated with the small jumps of X and a compound Poisson process associated with the large jumps of X .

More precisely, for any $\varepsilon \in (0, 1)$, $X \sim (b, \Sigma^2, \nu)$ can be decomposed as

$$\begin{aligned} X_t &= b(\varepsilon)t + \Sigma W_t + \lim_{\eta \rightarrow 0} \left(\sum_{s \leq t} \Delta X_s \mathbf{1}_{\eta < |\Delta X_s| \leq \varepsilon} - t \int_{\eta < |x| \leq \varepsilon} x \nu(dx) \right) + \sum_{s \leq t} \Delta X_s \mathbf{1}_{|\Delta X_s| > \varepsilon}, \\ &:= b(\varepsilon)t + \Sigma W_t + M_t(\varepsilon) + Z_t(\varepsilon), \quad \forall t \geq 0 \end{aligned} \quad (2)$$

where $\Delta X_t = X_t - \lim_{s \uparrow t} X_s$ denotes the jump at time t of X and

- $b(\varepsilon) := b - \int_{1 \leq |x| \leq \varepsilon} x \nu(dx)$;
- $W = (W_t)_{t \geq 0}$ is a standard Brownian motion;
- $M(\varepsilon) = (M_t(\varepsilon))_{t \geq 0}$ is a centered Lévy process (and a martingale) with a Lévy measure $\nu_\varepsilon := \nu \mathbf{1}_{[-\varepsilon, \varepsilon]}$ with support in $[-\varepsilon, \varepsilon]$ *i.e.* it is the Lévy process associated to the jumps of X smaller than ε ;
- $Z(\varepsilon) = (Z_t(\varepsilon))_{t \geq 0}$ is a Lévy process with a Lévy measure concentrated on $\mathbb{R} \setminus [-\varepsilon, \varepsilon]$ *i.e.* it is a compound Poisson process of the form $Z_t(\varepsilon) := \sum_{i=1}^{P_t} Z_i$ with intensity $\Lambda_\varepsilon := \nu(\mathbb{R} \setminus [-\varepsilon, \varepsilon])$ and jump measure $\mathbb{P}(Z_1 \in B) = \frac{1}{\Lambda_\varepsilon} \int_B \mathbf{1}_{\mathbb{R} \setminus [-\varepsilon, \varepsilon]}(x) \nu(dx)$;
- W , $M(\varepsilon)$ and $Z(\varepsilon)$ are independent.

Contrary to the process $M(\varepsilon)$, the processes W and $Z(\varepsilon)$, $\varepsilon > 0$, are well understood and have been extensively studied in the literature. In the present paper we focus on the process $M(\varepsilon)$ and our objective is twofold. On the one hand, we would like to understand when a Gaussian approximation of the increments of the process $M(\varepsilon)$ is valid in total variation. On the other hand, we would like to deduce new results to control the total variation distance between n given increments of two Lévy processes. More precisely, in the first case, suppose one is given n increments of the process $M(\varepsilon)$, $\varepsilon \in (0, 1)$, collected at the sampling rate $\Delta > 0$ over the time interval $[0, n\Delta]$, *i.e.* one observes $(M_{i\Delta}(\varepsilon) - M_{(i-1)\Delta}(\varepsilon), i = 1 \dots, n)$. We propose a sharp control of

$$\left\| \mathcal{L}((M_{i\Delta}(\varepsilon) - M_{(i-1)\Delta}(\varepsilon), i = 1 \dots, n)) - \mathcal{N}\left(0, \Delta \int_{|x| \leq \varepsilon} x^2 \nu(dx)\right)^{\otimes n} \right\|_{TV}, \quad (3)$$

where $\mathcal{L}(\cdot)$ stands for “the law of .” and $\mathcal{N}(m, \Sigma^2)$ stands for “the law of a Gaussian random variable with mean m and variance Σ^2 ”. As $M(\varepsilon)$ is a centered Lévy process, its increments $(M_{i\Delta}(\varepsilon) - M_{(i-1)\Delta}(\varepsilon), i = 1 \dots, n)$ are i.i.d. centered random variables with variance given by $\Delta \int_{|x| \leq \varepsilon} x^2 \nu(dx) =: \Delta \sigma^2(\varepsilon)$, that is the reason why we compare $\mathcal{L}((M_{i\Delta}(\varepsilon) - M_{(i-1)\Delta}(\varepsilon), i = 1 \dots, n))$ with the product measure $\mathcal{N}\left(0, \Delta \int_{|x| \leq \varepsilon} x^2 \nu(dx)\right)^{\otimes n}$. In the second case, suppose one observes n increments of two distinct Lévy processes $X^i \sim (b_i, \Sigma_i^2, \nu_i)$, $i = 1, 2$, at the sampling rate Δ over the time interval $[0, n\Delta]$. We provide a control of

$$\left\| \mathcal{L}((X_{i\Delta}^1 - X_{(i-1)\Delta}^1, i = 1 \dots, n)) - \mathcal{L}((X_{i\Delta}^2 - X_{(i-1)\Delta}^2, i = 1 \dots, n)) \right\|_{TV}.$$

In the sequel, we exploit several approximations for different distances and objects. We introduce the following notations. For any $0 < \eta < \varepsilon$ define $M(\eta, \varepsilon)$, the centered compound Poisson process that approximates $M(\varepsilon)$ as $\eta \downarrow 0$, by

$$M_t(\eta, \varepsilon) = \sum_{s \leq t} \Delta X_s \mathbf{1}_{\eta < |\Delta X_s| \leq \varepsilon} - t \int_{\eta < |x| \leq \varepsilon} x \nu(dx) = \sum_{i=1}^{N_t} Y_i - t \int_{\eta < |x| \leq \varepsilon} x \nu(dx)$$

where N is a Poisson process with intensity $\lambda_{\eta,\varepsilon} = \int_{\eta < |x| \leq \varepsilon} \nu(dx)$ and the $(Y_i)_{i \geq 1}$ are i.i.d. with jump measure

$$\mathbb{P}(Y_1 \in B) = \frac{1}{\lambda_{\eta,\varepsilon}} \int_{B \cap \{\eta < |x| \leq \varepsilon\}} \nu(dx) \quad \forall B \in \mathcal{B}(\mathbb{R}).$$

Moreover, we denote the first moments of $M(\eta, \varepsilon)$ by

$$\sigma^2(\eta, \varepsilon) = \int_{\eta < |x| \leq \varepsilon} x^2 \nu(dx), \quad \mu_j(\eta, \varepsilon) = \int_{\eta < |x| \leq \varepsilon} x^j \nu(dx), \quad \text{and} \quad \mu_j(\varepsilon) = \mu_j(0, \varepsilon) \quad j = 3, 4.$$

Finally we recall that the total variation distance between two probability measures P_1 and P_2 defined on the same σ -field \mathcal{B} is defined as

$$\|P_1 - P_2\|_{TV} := \sup_{B \in \mathcal{B}} |P_1(B) - P_2(B)| = \frac{1}{2} \int \left| \frac{dP_1}{d\mu}(x) - \frac{dP_2}{d\mu}(x) \right| \mu(dx),$$

where μ is a common dominating measure for P_1 and P_2 . The χ^2 -distance is instead defined as follows:

$$\chi^2(P_1, P_2) = \begin{cases} \int \left(\frac{dP_1}{dP_2} - 1 \right)^2 dP_2 & \text{if } P_1 \ll P_2, \\ +\infty & \text{otherwise,} \end{cases}$$

where the symbol “ \ll ” stands for “absolutely continuous with respect to”.

In order to ease the reading, if X and Y are random variables with densities f_X and f_Y with respect to a same dominating measure, we sometimes write $\|X - Y\|_{TV}$ or $\|f_X - f_Y\|_{TV}$ instead of $\|\mathcal{L}(X) - \mathcal{L}(Y)\|_{TV}$. The same type of abbreviation is used for any other distance.

2 Gaussian approximation for $M_\Delta(\varepsilon)$ in total variation distance

Our main goal is to understand under which circumstances a Gaussian approximation for the small jumps of a Lévy process is adequate. Of course, what is meant by “adequate” must be clarified. As statisticians, such approximation is justified if we are not able to distinguish from n discrete observations of the increments of the small jumps and n independent Gaussian random variables. More concretely, we will be able to establish that a Gaussian approximation for $(M_{i\Delta}(\varepsilon) - M_{(i-1)\Delta}(\varepsilon), i = 1, \dots, n)$ is valid if we can provide a sharp control in total variation distance. This is the subject of this section. First, we prove an upper bound for (3) which is, using the independence and stationarity of the increments of $M(\varepsilon)$, equivalent to prove an upper bound for

$$\|M_\Delta(\varepsilon)^{\otimes n} - \mathcal{N}(0, \Delta\sigma^2(\varepsilon))^{\otimes n}\|_{TV}. \quad (4)$$

Several intermediate approximations given in Propositions 1, 2 and 3 are necessary to control (4). The final result is given in Theorem 2. At the beginning of Section 2.1, we outline the main steps needed to establish Theorem 2. Proofs are postponed to Section 4 whereas the optimality of Theorem 2 is confirmed in Section 2.2 where a lower bound result is presented.

2.1 Upper bound

2.1.1 Preliminary approximations

We establish an upper bound for (4) in Theorem 2 below, many auxiliary approximations are used to prove it, which are gathered in this Section. The starting point is that the process $M(\varepsilon)$ can be decomposed as follows. For any $t \geq 0$ and for any $0 < \eta < \varepsilon$,

$$M_t(\varepsilon) = M_t(\eta) + \sum_{i=1}^{N_t} Y_i - t \int_{\eta < |x| \leq \varepsilon} x \nu(dx) =: M_t(\eta) + M_t(\eta, \varepsilon), \quad (5)$$

where we recall that $(N_t)_{t \geq 0}$ is a Poisson process with intensity $\lambda_{\eta, \varepsilon} = \nu(\{\eta < |x| \leq \varepsilon\})$ and the Y_i 's are i.i.d. with density $\frac{1}{\lambda_{\eta, \varepsilon}} \frac{\nu(dx)}{dx} \mathbf{1}_{\eta < |x| \leq \varepsilon}$. The interest in writing $M_t(\varepsilon)$ as in (5) lies in the fact that the process $M(\eta, \varepsilon)$ is more manageable than $M(\varepsilon)$, being a Lévy process with finite Lévy measure. The spirit of the approximation of $M_\Delta(\varepsilon)$ by $\mathcal{N}(0, \Delta\sigma^2(\varepsilon))$ is to use the closeness, for η small enough, between $M_\Delta(\eta, \varepsilon)$ and $M_\Delta(\varepsilon)$ on the one hand and between $M_\Delta(\eta, \varepsilon)$ and $\mathcal{N}(0, \Delta\sigma^2(\varepsilon))$ on the other hand.

First, we try to understand if, for η small enough, working with discrete observations of $M(\varepsilon)$ is somehow equivalent to working with those of $M(\eta, \varepsilon)$. What is well known is that the law of $M_t(\eta, \varepsilon)$ converges almost surely and in L_2 to the law of $M_t(\varepsilon)$ as η goes to 0, but this is not strong enough for our purpose. In Proposition 1, we establish a control in total variation distance between n i.i.d. increments of these quantities, under the assumption that ν is an infinite Lévy measure.

Proposition 1. *Let ν be an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure. For any fixed $\varepsilon \in (0, 1)$, $n \geq 1$ and $\Delta > 0$, it holds,*

$$\limsup_{\eta \rightarrow 0} \|M_\Delta(\eta, \varepsilon)^{\otimes n} - M_\Delta(\varepsilon)^{\otimes n}\|_{TV} \leq 2 \liminf_{\varepsilon' \rightarrow 0} \|M_\Delta(\varepsilon', \varepsilon)^{\otimes n} - \mathcal{N}(0, \Delta\sigma^2(\varepsilon', \varepsilon))^{\otimes n}\|_{TV}.$$

Roughly speaking, Proposition 1 tells us that the distance between n i.i.d. observations of $M_\Delta(\varepsilon)$ and n i.i.d. observations of $M_\Delta(\eta, \varepsilon)$ is bounded by the approximation error one would make approximating the discrete observation of $M(\eta, \varepsilon)$ by the corresponding Gaussian observations.

Second, always under the assumption that ν is infinite, we control the distance between $M_\Delta(\eta, \varepsilon)$ and $\mathcal{N}(0, \Delta\sigma^2(\varepsilon))$, for that we use an Edgeworth expansion for the distribution of $M_\Delta(\eta, \varepsilon)$. However, this expansion will be sharp only when N_Δ , appearing in the definition of $M_\Delta(\eta, \varepsilon)$, is large. Therefore, we need to localize the values taken by the Poisson process N_Δ ; when η goes to 0, its intensity $\lambda_{\eta, \varepsilon}$ increases and N_Δ gets closer to its mean value $\Delta\lambda_{\eta, \varepsilon}$. This is why we introduce the following events, for a fixed $\psi \in (0, 1)$,

$$A_i := \{\omega : \lambda_{\eta, \varepsilon} \Delta(1 - \psi) < N_{\Delta i}(\omega) - N_{\Delta(i-1)}(\omega) < \lambda_{\eta, \varepsilon} \Delta(1 + \psi)\}, \quad i = 1, \dots, n. \quad (6)$$

This events are equiprobable, set $A := A_1$. For technical reasons, it is easier to work with the compound Poisson process with centered jumps $\widetilde{M}_\Delta(\eta, \varepsilon) := \sum_{i=1}^{N_\Delta} (Y_i - \mathbb{E}[Y_i])$ than with the centered compound Poisson process $M(\eta, \varepsilon)$. These two processes are equal if ν is symmetric and they are close to each other for η small enough. Denote by $f_{\eta, \varepsilon, A}$ (resp. $\widetilde{f}_{\eta, \varepsilon, A}$) the density of

$$\frac{1}{\mathbb{P}(A)} \left(\sum_{\substack{k \in \mathbb{N} \\ |k - \lambda_{\eta, \varepsilon} \Delta| \leq \psi \lambda_{\eta, \varepsilon} \Delta}} \mathbb{P}(N_\Delta = n) \sum_{i=1}^n Y_i - \lambda_{\eta, \varepsilon} \Delta \mathbb{E}[Y_1] \right) \quad \left(\text{resp.} \quad \sum_{\substack{k \in \mathbb{N} \\ |k - \lambda_{\eta, \varepsilon} \Delta| \leq \psi \lambda_{\eta, \varepsilon} \Delta}} \frac{\mathbb{P}(N_\Delta = n)}{\mathbb{P}(A)} \sum_{i=1}^n (Y_i - \mathbb{E}[Y_1]) \right) \quad (7)$$

and by $\varphi_{\eta, \varepsilon}$ (resp. $\widetilde{\varphi}_{\eta, \varepsilon}$) the density of a centered Gaussian distribution with variance $\Delta\sigma^2(\eta, \varepsilon)$ (resp. $\Delta\widetilde{\sigma}^2(\eta, \varepsilon) = \Delta\sigma^2(\eta, \varepsilon) - \frac{\Delta}{\lambda_{\eta, \varepsilon}} (\int_{\eta < |x| \leq \varepsilon} x\nu(dx))^2$). The densities $f_{\eta, \varepsilon, A}$ (resp. $\widetilde{f}_{\eta, \varepsilon, A}$) can be interpreted as the densities of $M_\Delta(\eta, \varepsilon)$ (resp. $\widetilde{M}_\Delta(\eta, \varepsilon)$) conditioned by the event $N_\Delta \in A$. Given that N_Δ is a Poisson random variable with mean $\Delta\lambda_{\eta, \varepsilon}$ and that $\lambda_{\eta, \varepsilon} \rightarrow \infty$ as $\eta \rightarrow 0$, the event A^c has exponentially small probability, that is why we focus on A .

Proposition 2 below justifies the substitution of $M_\Delta(\eta, \varepsilon)$ with $\widetilde{M}_\Delta(\eta, \varepsilon)$. We prove the convergence in L_1 , as $\eta \rightarrow 0$, between the laws of $M_\Delta(\eta, \varepsilon)$ and $\widetilde{M}_\Delta(\eta, \varepsilon)$ restricted to the event A , *i.e.* conditioned on the fact that $|N_\Delta - \Delta\lambda_{\eta, \varepsilon}| < \Delta\lambda_{\eta, \varepsilon}\psi$, for some $\psi \in (0, 1)$.

Proposition 2. Fix $0 < \varepsilon < 1$ and let ν be an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure. Then,

$$\lim_{\eta \rightarrow 0} \int |f_{\eta, \varepsilon, A}(x) - \tilde{f}_{\eta, \varepsilon, A}(x)| dx = 0.$$

Next, we approximate $\tilde{f}_{\eta, \varepsilon, A}$, the density of the random variable $\sum_{i=1}^{N_\Delta} (Y_i - \mathbb{E}[Y_i])$ conditional on the event $|N_\Delta - \Delta \lambda_{\eta, \varepsilon}| < \Delta \lambda_{\eta, \varepsilon} \psi$, when η is small with respect to ε . Under the assumption ν is infinite, $\lambda_{\eta, \varepsilon}$ is large and roughly speaking we can think about $\tilde{f}_{\eta, \varepsilon, A}$ as the density of the sum of $\Delta \lambda_{\eta, \varepsilon}$ i.i.d bounded and centered random variables. In view of the central limit theorem, $\tilde{f}_{\eta, \varepsilon, A}$ is therefore close to a Gaussian density. However, quantifying this proximity requires to choose a suitable distance between $\tilde{f}_{\eta, \varepsilon, A}$ and $\tilde{\varphi}_{\eta, \varepsilon}$. Since our final goal is to control product measures in total variation, we select a metric that behaves well with respect to product measures and that is connected with total variation, such as, the Hellinger distance, the Kullback-Leibler divergence or the χ^2 distance. The latter was easier to work with in this context and lead to the following Proposition 3.

Proposition 3. Fix $0 < \varepsilon < 1$, $\Delta > 0$ and let ν be an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure. Then, for any $C > \frac{7}{6}$ and $C' > 1$ there exists $\eta_0 > 0$ independent of Δ and ε , such that for any $\eta < \eta_0$

$$\chi^2(\tilde{f}_{\eta, \varepsilon, A}, \tilde{\varphi}_{\eta, \varepsilon}) \leq C \frac{\mu_3^2(\eta, \varepsilon)}{\Delta(\sigma^2(\eta, \varepsilon))^3} + C' \frac{\mu_4^2(\eta, \varepsilon)}{\Delta^2(\sigma^2(\eta, \varepsilon))^4}. \quad (8)$$

2.1.2 Upper bound result : Gaussian approximation

Collecting all the pieces together, it is finally possible to state under which conditions on Δ , n and ε , a Gaussian approximation of the small jumps is valid, provided that ν is sufficiently active.

Theorem 2. Let ν be an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure. Then, for any $0 < \varepsilon < 1$, $\Delta > 0$ and $n \geq 1$ we have

$$\|M_\Delta(\varepsilon)^{\otimes n} - \mathcal{N}(0, \Delta \sigma^2(\varepsilon))^{\otimes n}\|_{TV} \leq 3 \sqrt{\left(1 + \frac{\mu_4^2(\varepsilon)}{\Delta^2(\sigma^2(\varepsilon))^4} + \frac{7}{6} \frac{\mu_3^2(\varepsilon)}{\Delta(\sigma^2(\varepsilon))^3}\right)^n - 1}. \quad (9)$$

Proof of Theorem 2. Let $0 < \eta < \varepsilon$, by the triangle inequality it holds

$$\begin{aligned} \|M_\Delta^{\otimes n}(\varepsilon) - \mathcal{N}(0, \Delta \sigma^2(\varepsilon))^{\otimes n}\|_{TV} &\leq \|M_\Delta^{\otimes n}(\varepsilon) - M_\Delta(\eta, \varepsilon)^{\otimes n}\|_{TV} + \|M_\Delta(\eta, \varepsilon)^{\otimes n} - \mathcal{N}(0, \Delta \sigma^2(\eta, \varepsilon))^{\otimes n}\|_{TV} \\ &\quad + \|\mathcal{N}(0, \Delta \sigma^2(\eta, \varepsilon))^{\otimes n} - \mathcal{N}(0, \Delta \sigma^2(\varepsilon))^{\otimes n}\|_{TV} \\ &= T_1 + T_2 + T_3. \end{aligned}$$

The first term T_1 , is controlled by Proposition 1. For the term T_2 , recall that $M_\Delta(\eta, \varepsilon)$ is of the form $\sum_{i=1}^{N_\Delta} Y_i - \Delta \int_{\eta < |x| \leq \varepsilon} x \nu(dx)$; together with the definition of the sets A_i given in (6), it follows that

$$T_2 \leq \mathbb{P}\left((N_\Delta, N_{2\Delta} - N_\Delta, \dots, N_{n\Delta} - N_{(n-1)\Delta}) \in \left(\bigotimes_{i=1}^n A_i\right)^c\right) + \|f_{\eta, \varepsilon, A}^{\otimes n} - \varphi_{\eta, \varepsilon}^{\otimes n}\|_{TV}. \quad (10)$$

The first term in (10) is bounded using the Chernov inequality and $(1+\psi) \log(1+\psi) - \psi > \frac{\psi^2}{2}$, $\psi \in (0, 1)$,

$$n\mathbb{P}(|N_\Delta - \lambda_{\eta, \varepsilon} \Delta| > \psi \lambda_{\eta, \varepsilon} \Delta) \leq 2ne^{-\frac{\psi^2 \lambda_{\eta, \varepsilon} \Delta}{2}}. \quad (11)$$

For the second term in (10) observe that by triangular inequality joined with Lemma 13, we have

$$\begin{aligned} \|f_{\eta,\varepsilon,A}^{\otimes n} - \varphi_{\eta,\varepsilon}^{\otimes n}\|_{TV} &\leq \|f_{\eta,\varepsilon,A}^{\otimes n} - \tilde{f}_{\eta,\varepsilon,A}^{\otimes n}\|_{TV} + \|\tilde{f}_{\eta,\varepsilon,A}^{\otimes n} - \varphi_{\eta,\varepsilon}^{\otimes n}\|_{TV} \\ &\leq \sqrt{2n\|f_{\eta,\varepsilon,A} - \tilde{f}_{\eta,\varepsilon,A}\|_{TV}} + \|\tilde{f}_{\eta,\varepsilon,A}^{\otimes n} - \tilde{\varphi}_{\eta,\varepsilon}^{\otimes n}\|_{TV} + \|\varphi_{\eta,\varepsilon}^{\otimes n} - \tilde{\varphi}_{\eta,\varepsilon}^{\otimes n}\|_{TV}. \end{aligned} \quad (12)$$

The first term in (12) is controlled by means of Proposition 2, whereas for the second term $\|\tilde{f}_{\eta,\varepsilon,A}^{\otimes n} - \tilde{\varphi}_{\eta,\varepsilon}^{\otimes n}\|_{TV}$ we use the following well known relation on total variation and χ^2 distances (see *e.g.* Tsybakov [18])

$$\chi^2(P_1^{\otimes n}, P_2^{\otimes n}) = (1 + \chi^2(P_1, P_2))^n - 1 \quad \text{and} \quad \|P_1 - P_2\|_{TV} \leq \sqrt{\chi^2(P_1, P_2)}.$$

In particular we get

$$\|\tilde{f}_{\eta,\varepsilon,A}^{\otimes n} - \tilde{\varphi}_{\eta,\varepsilon}^{\otimes n}\|_{TV} \leq \sqrt{(1 + \chi^2(\tilde{f}_{\eta,\varepsilon,A}, \tilde{\varphi}_{\eta,\varepsilon}))^n - 1}.$$

Applying Proposition 3 we therefore deduce that for any $C > \frac{7}{6}$ and $C' > 1$ there exists $\eta_0 > 0$ such that

$$\|\tilde{f}_{\eta,\varepsilon,A}^{\otimes n} - \tilde{\varphi}_{\eta,\varepsilon}^{\otimes n}\|_{TV} \leq \sqrt{\left(1 + \frac{C\mu_3^2(\eta, \varepsilon)}{\Delta(\sigma^2(\eta, \varepsilon))^3} + C' \frac{\mu_4^2(\eta, \varepsilon)}{\Delta^2(\sigma^2(\eta, \varepsilon))^4}\right)^n - 1}. \quad (13)$$

The fact that T_3 and $T_4 = \|\varphi_{\eta,\varepsilon}^{\otimes n} - \tilde{\varphi}_{\eta,\varepsilon}^{\otimes n}\|_{TV}$ vanish as $\eta \rightarrow 0$ is a consequence of Lemma 10 joined with the fact that if $\nu(\mathbb{R}) = \infty$ then

$$\lim_{\eta \rightarrow 0} \frac{1}{\lambda_{\eta,\varepsilon}} \left(\int_{\eta < |x| < \varepsilon} x\nu(dx) \right)^2 = 0, \quad (14)$$

as proved in Lemma 2. Finally, for any $\psi \in (0, 1)$, having $\eta \rightarrow 0$ (11) vanishes and we have:

$$\begin{aligned} \|M_{\Delta}^{\otimes n}(\varepsilon) - \mathcal{N}(0, \Delta\sigma^2(\varepsilon))^{\otimes n}\|_{TV} &\leq \liminf_{\eta \rightarrow 0} \|M_{\Delta}^{\otimes n}(\varepsilon) - M_{\Delta}(\eta, \varepsilon)^{\otimes n}\|_{TV} \\ &\quad + \liminf_{\eta \rightarrow 0} \|M_{\Delta}(\eta, \varepsilon)^{\otimes n} - \mathcal{N}(0, \Delta\sigma^2(\eta, \varepsilon))^{\otimes n}\|_{TV} \\ &\quad + \liminf_{\eta \rightarrow 0} \|\mathcal{N}(0, \Delta\sigma^2(\eta, \varepsilon))^{\otimes n} - \mathcal{N}(0, \Delta\sigma^2(\varepsilon))^{\otimes n}\|_{TV} \\ &\leq 2 \liminf_{\varepsilon' \rightarrow 0} \|M_{\Delta}(\varepsilon', \varepsilon)^{\otimes n} - \mathcal{N}(0, \Delta\sigma^2(\varepsilon', \varepsilon))^{\otimes n}\|_{TV} \\ &\quad + \liminf_{\eta \rightarrow 0} \|M_{\Delta}(\eta, \varepsilon)^{\otimes n} - \mathcal{N}(0, \Delta\sigma^2(\eta, \varepsilon))^{\otimes n}\|_{TV} \\ &= 3 \liminf_{\eta \rightarrow 0} \|f_{\eta,\varepsilon,A}^{\otimes n} - \varphi_{\eta,\varepsilon}^{\otimes n}\|_{TV} \end{aligned}$$

The statement of Theorem 2 follows from (13) and the fact that $T_4 \rightarrow 0$ as $\eta \rightarrow 0$, using (14). \square

Theorem 3. *Let ν be an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure. Then, for any $0 < \varepsilon < 1$, $\Delta > 0$, $n \geq 1$, $b \in \mathbb{R}$ and $\Sigma \geq 0$, we have*

$$\begin{aligned} \|(\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_{\Delta}(\varepsilon))^{\otimes n} - \mathcal{N}(b\Delta, \Delta(\Sigma^2 + \sigma^2(\varepsilon)))^{\otimes n}\|_{TV} &\leq \\ &3\sqrt{\left(1 + \frac{\mu_4^2(\varepsilon)}{\Delta^2(\Sigma^2 + \sigma^2(\varepsilon))^4} + \frac{7}{6} \frac{\mu_3^2(\varepsilon)}{\Delta(\Sigma^2 + \sigma^2(\varepsilon))^3}\right)^n - 1}. \end{aligned} \quad (15)$$

Proof. The proof is analogous to that of Theorem 2, we only need to adequately modify Propositions 1–3. In this regard, observe that Proposition 1 becomes

$$\begin{aligned} \limsup_{\eta \rightarrow 0} \left\| \left(\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(\varepsilon) \right)^{\otimes n} - \mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(\eta, \varepsilon) \right\|_{TV} \\ \leq 2 \liminf_{\varepsilon' \rightarrow 0} \left\| \left(\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(\varepsilon', \varepsilon) \right)^{\otimes n} - \mathcal{N}(b\Delta, \Delta(\sigma^2(\varepsilon', \varepsilon) + \Sigma^2))^{\otimes n} \right\|_{TV}. \end{aligned}$$

Let $g_{\eta, \varepsilon, A}$ (resp. $\tilde{g}_{\eta, \varepsilon, A}$) be the convolution between $f_{\eta, \varepsilon, A}$ (resp. $\tilde{f}_{\eta, \varepsilon, A}$) and the density of $\mathcal{N}(b\Delta, \Delta\Sigma^2)$. Then, by means of Proposition 2 and Lemma 12 we get

$$\lim_{\eta \rightarrow 0} \int |g_{\eta, \varepsilon, A}(x) - \tilde{g}_{\eta, \varepsilon, A}(x)| dx \leq \lim_{\eta \rightarrow 0} \int |f_{\eta, \varepsilon, A}(x) - \tilde{f}_{\eta, \varepsilon, A}(x)| dx = 0.$$

Finally, tedious computations allow to rewrite Proposition 3 as: for any $C > \frac{7}{6}$ and $C' > 1$ there exists $\eta_0 > 0$ independent of Δ and ε , such that for any $\eta < \eta_0$

$$\chi^2(\tilde{f}_{\eta, \varepsilon, A}, \tilde{\varphi}_{\eta, \varepsilon}) \leq C \frac{\mu_3^2(\eta, \varepsilon)}{\Delta(\sigma^2(\eta, \varepsilon) + \Sigma^2)^3} + C' \frac{\mu_4^2(\eta, \varepsilon)}{\Delta^2(\sigma^2(\eta, \varepsilon) + \Sigma^2)^4}.$$

□

Corollary 1. *Let ν be an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure. Then, for all $\varepsilon \in (0, 1)$, for all $n \geq 1$ and $\Delta > 0$ it holds.*

- If ν is symmetric, then,

$$\|M_\Delta(\varepsilon)^{\otimes n} - \mathcal{N}(0, \Delta\sigma^2(\varepsilon))^{\otimes n}\|_{TV} \leq 3\sqrt{\left(1 + \left(\frac{\varepsilon}{\sqrt{\Delta}\sigma(\varepsilon)}\right)^4\right)^n - 1}. \quad (16)$$

- If ν is not symmetric, then,

$$\|M_\Delta(\varepsilon)^{\otimes n} - \mathcal{N}(0, \Delta\sigma^2(\varepsilon))^{\otimes n}\|_{TV} \leq 3\sqrt{\left(1 + \frac{7}{6}\left(\frac{\varepsilon}{\sqrt{\Delta}\sigma(\varepsilon)}\right)^2\right)^n - 1}. \quad (17)$$

Proof. It is a straightforward consequence of Theorem 2 together with

$$\mu_4(\varepsilon) \leq \varepsilon^2\sigma^2(\varepsilon), \quad \mu_3(\varepsilon) \leq \varepsilon\sigma^2(\varepsilon). \quad (18)$$

□

2.1.3 Comments

The results are non asymptotic and we stress here that (9), (15), (16) and (17) hold without any assumption on Δ : we are not assuming to work with high or low frequency observations. An explicit relation between ε , n and Δ depending on ν via $\sigma^2(\varepsilon)$, $\mu_3(\varepsilon)$ and $\mu_4(\varepsilon)$ is given. More precisely, provided that $\frac{\mu_4^2(\varepsilon)}{(\sigma^2(\varepsilon))^4} = o(1)$ and $\frac{\mu_3^2(\varepsilon)}{(\sigma^2(\varepsilon))^3} = o(1)$ as $\varepsilon \rightarrow 0$, the rate of convergence in (9) is of order

$$\sqrt{n} \left(\frac{\mu_4(\varepsilon)}{\Delta(\sigma^2(\varepsilon))^2} + \frac{7}{6} \frac{\mu_3(\varepsilon)}{\sqrt{\Delta}(\sigma^2(\varepsilon))^{3/2}} \right). \quad (19)$$

As highlighted in Corollary 1, a sufficient condition for the convergence to zero of (19) is given by $\frac{\varepsilon}{\sigma(\varepsilon)} \rightarrow 0$ as $\varepsilon \rightarrow 0$. Unsurprisingly, we observe that the rate of convergence for a Gaussian approximation of the small jumps is faster when the Lévy measure is symmetric.

Besides, let us underline that the assumptions made to establish Theorems 2 and 3 are extremely mild. We only require to work with Lévy processes with infinite Lévy measures which are absolutely continuous with respect to the Lebesgue measure. The absolute continuity assumption permits to avoid degenerate cases such as finite support Lévy measure, where clearly a Gaussian approximation in total variation does not hold. We restrict our analysis to the case $\nu(\mathbb{R}) = \infty$ because the case $\nu(\mathbb{R}) < \infty$ is not relevant in this context. Having a finite Lévy measure means that we are considering a process which is the convolution between a Gaussian process and a compound Poisson process. Several tools have already been developed to deal with this kind of processes. Moreover, in this context the small jumps are summable and do not present any theoretical difficulty.

Concerning the literature, to the best of our knowledge, the only work in which non-asymptotic results are found for a Gaussian approximation of the small jumps of Lévy processes is [14]. Proposition 8 in [14] states that, for any $\Sigma > 0$:

$$\|\mathcal{N}(0, t\Sigma^2) * M_t(\varepsilon) - \mathcal{N}(0, t(\sigma^2(\varepsilon) + \Sigma^2))\|_{TV} \leq \frac{1}{\sqrt{2\pi t\Sigma^2}} \min\left(2\sqrt{t\sigma^2(\varepsilon)}, \frac{\varepsilon}{2}\right). \quad (20)$$

The assumption $\Sigma > 0$ is an artifact of the proof and it arises from the need of a convolutional structure which enables to transfer results in Wasserstein distance of order 1 to results in total variation distance, see Proposition 4 in [14]. However, this type of approach leads to suboptimal results even for large classes of Lévy processes as the present paper shows. Indeed, if we specialize our Theorem 3 to the case of $n = 1$ and we suppose $\Sigma > 0$ in order to be able to compare with (20), we have:

$$\|\mathcal{N}(0, t\Sigma^2) * M_t(\varepsilon) - \mathcal{N}(0, t(\sigma^2(\varepsilon) + \Sigma^2))\|_{TV} \leq 3\left(\frac{\mu_4(\varepsilon)}{t(\Sigma^2 + \sigma^2(\varepsilon))^2} + \sqrt{\frac{7}{6}} \frac{\mu_3(\varepsilon)}{t(\Sigma^2 + \sigma^2(\varepsilon))^{3/2}}\right)$$

which is faster than (20) thanks to (18). Moreover, let us stress that Proposition 8 in [14] establishes a Gaussian approximation for the small jumps of Lévy processes with non-zero diffusion coefficients only for the marginal laws. Having sharp results for the n -sample is a much more demanding task.

Finally, an interesting question is how to choose the size of the small jumps. Suppose we are given a Lévy process $X \sim (b, \Sigma^2, \nu)$ with $\nu(\mathbb{R}) = \infty$ and for some reasons (e.g. we would like to simulate a trajectory, make some inference, etc...) we would like to replace the small jumps, a complex object, with a Gaussian distribution, an easier object. The question is then what is the optimal choice for ε , the size of the small jumps, *i.e.* the largest ε such that we cannot distinguish between n realizations of $M_\Delta(\varepsilon) * \mathcal{N}(b\Delta, \Delta\Sigma^2)$ and some Gaussian random variable. As we will see in Section 2.2, in many situations this corresponds to taking the largest ε such that it holds

$$\frac{\sqrt{n}\mu_4(\varepsilon)}{\Delta(\Sigma^2 + \sigma^2(\varepsilon))^2} + \frac{\sqrt{n}\mu_3(\varepsilon)}{\Delta(\Sigma^2 + \sigma^2(\varepsilon))^{3/2}} \rightarrow 0.$$

An example is detailed in Section 2.3.

2.2 Lower bound

Theorems 2 and 3 are optimal in the following sense. If the upper bounds in Theorems 2 and 3 do not converge, then the total variation distance between the random vector associated with the increments of the small jumps -eventually convoluted with Gaussian distributions- and the corresponding Gaussian random vector does not converge to 0. This can be seen by means of the following result.

Theorem 4. *Let ν be an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure, let $b \in \mathbb{R}$ and $\Sigma \geq 0$. For any $n\Delta \geq 1$ and $\varepsilon \in (0, 1)$ such that $\lambda_{\varepsilon/2, \varepsilon} \geq \log(n)/(n\Delta)$, there exists an absolute sequence $\alpha_n \rightarrow 0$ (independent of $b, \nu, \varepsilon, \Sigma, \Delta$) and an absolute constant $C > 0$*

(independent of $b, \nu, \varepsilon, \Sigma, \Delta$) such that the following holds:

$$\begin{aligned} & \left\| (\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(\varepsilon))^{\otimes n} - \mathcal{N}(b\Delta, \Delta(\Sigma^2 + \sigma^2(\varepsilon)))^{\otimes n} \right\|_{TV} \geq \\ & \left\{ 1 - C \left[\frac{\Delta(\Sigma^2 + \sigma^2(\varepsilon))^3}{n\mu_3(\varepsilon)^2} \wedge \frac{\Delta^2(\Sigma^2 + \sigma^2(\varepsilon))^4}{n\mu_4(\varepsilon)^2} \right] - \alpha_n \right\}. \end{aligned}$$

To establish Theorem 4 we construct an appropriate statistical test and use the following fact.

Lemma 1. *Let \mathbb{P} and \mathbb{Q} be two probability distributions and Φ a test of level $\alpha_0 \in (0, 1)$ that separates \mathbb{P} from \mathbb{Q} from n i.i.d. observation with power larger than $1 - \alpha_1$. Then, $\|\mathbb{P}^{\otimes n} - \mathbb{Q}^{\otimes n}\|_{TV} \geq 1 - \alpha_0 - \alpha_1$.*

Let $X \sim (b, \Sigma^2, \nu)$, with b, Σ^2 and ν possibly unknown and consider the problem of testing whether the process contains jumps of size smaller than ε or not, i.e. whether $\nu_\varepsilon = 0$ or not, recall that $\nu_\varepsilon = \nu \mathbf{1}_{[-\varepsilon, \varepsilon]}$. Set $u^+ := u_{\nu_\varepsilon}^+$ for the positive real number such that

$$\lambda_{u^+, \varepsilon} n \Delta = \log(n).$$

Note that such a u^+ exists and is unique whenever ν_ε is an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure. Write now u^* for the largest $u \in [u^+, \varepsilon]$ such that

$$u^* = \sup\{u, u \in [u^+, \varepsilon], u \leq \sqrt{\Delta(\Sigma^2 + \sigma^2(u)) \log(n)^2}\} \vee u^+,$$

where $\sup \emptyset = -\infty$.

We prove the following result, of which Theorem 4 is an immediate corollary.

Theorem 5. *Suppose that $\varepsilon \in (0, 1)$, $n\Delta \geq 1$, ν is an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure, $b \in \mathbb{R}$ and $\Sigma \geq 0$. There exists an absolute sequence $\alpha_n \rightarrow 0$ (independent of $b, \nu, \varepsilon, \Sigma, \Delta$) and an absolute constant $C > 0$ (independent of $b, \nu, \varepsilon, \Sigma, \Delta$) such that the following holds:*

$$\begin{aligned} & \min_{B \in \mathbb{R}, S^2 \geq 0} \left\| (\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(\varepsilon))^{\otimes n} - \mathcal{N}(B\Delta, \Delta S^2)^{\otimes n} \right\|_{TV} \geq \\ & \left\{ 1 - C \left[\frac{\Delta(\Sigma^2 + \sigma^2(u^*))^3}{n\mu_3(u^*)^2} \wedge \frac{\Delta^2(\Sigma^2 + \sigma^2(u^*))^4}{n\mu_4(u^*)^2} \right] - \alpha_n \right\} \vee \left(1 - \exp(-\lambda_{u^*, \varepsilon} n \Delta) - \alpha_n \right). \end{aligned}$$

The construction of the test we use to derive Theorem 5 is actually quite involved, we refer to Section 4.2.1 for more details. Here, we only illustrate the main ideas. First, the intuitions behind the quantities u^+ and u^* are the following. The quantity u^+ is chosen such that with probability going to 1, there is (i) at least one jump larger than u^+ but (ii) not too many of such jumps, i.e. less than $2 \log(n)$, and finally (iii) at most one jump larger than u^+ per time increment Δ . Therefore, the discretized process of jumps larger than u^+ , $M(u^+, \varepsilon)$, does not look Gaussian at all. It is composed of many null entries and a few larger than u^+ . Now u^* is the largest quantity (larger than u^+) such that u^* is smaller than a number slightly larger than the standard deviation of the increment $X_{i\Delta} - X_{(i-1)\Delta}$ conditional to the event there are no jumps larger than u^* in $((i-1)\Delta, i\Delta]$. In other words, any increment having a jump larger than u^* is going to be quite visible.

Second, the idea behind the test, is to build an event that occurs with high probability if $\nu_\varepsilon = 0$ and with small probability otherwise. This would then allow us to bound the total variation distance between the two discretized processes using Lemma 1. The sketch of the proof is the following (all results mentioned below are stated and proved in Section 4 and Appendix A):

- First, we show that u^+ defined as above satisfies (i)-(ii)-(iii) with probability going to 1 and we bound the deviations of the difference of some¹ of the increments $X_{i\Delta} - X_{(i-1)\Delta} - (X_{(i-1)\Delta} - X_{(i-2)\Delta})$ (see Lemma 3).
- Second, we build an estimator of the standard deviation of the increments of the Lévy process (b, Σ, ν_{u^+}) . In order to do so, we use a robust estimator of the mean which drops large increments, and thus the ones larger than u^+ (see Lemma 4).
- From these preliminary steps, we prove that a test comparing the largest entry and the expected standard deviation if $\nu_\varepsilon = 0$ detects if there is a jump larger than u^* in the sample (see Proposition 4). In the following steps, we focus on tests *conditional to the event there is no jump larger than u^* in the sample* - otherwise they are eliminated by the present test. Two cases remain to be studied.
- If the dominant quantity in Theorem 3 is $\Delta\mu_3(u^*)$: we first construct a test for detecting if $\Delta\mu_6(u^*)$ is larger than a constant times $[\Delta(\Sigma^2 + \sigma^2(u^*))]^3$, to remove distributions that are too skewed (Proposition 5). Then, we build a test comparing the (estimated) third moment of the increments to the expected behavior if $\nu_\varepsilon = 0$ (see Proposition 6).
- If the dominant quantity in Theorem 3 is $\Delta\mu_4(u^*)$: we build a test comparing the (estimated) fourth moment of the increments to the expected behavior if $\nu_\varepsilon = 0$ (see Proposition 7).

2.2.1 Comments

Tightness of the lower bound on the total variation distance. The lower bound from Theorem 5 is tight as shown in the Corollary of Theorem 3 below.

Corollary 2. *Let ν be an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure and fix $0 < \varepsilon < 1$, $\Delta > 0$, $n \geq 1$, $b \in \mathbb{R}$ and $\Sigma \geq 0$. Then, there exists an absolute constant $C > 0$ such that*

$$\min_{B \in \mathbb{R}, S^2 \geq 0} \left\| (\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(\varepsilon))^{\otimes n} - \mathcal{N}(B\Delta, \Delta S^2)^{\otimes n} \right\|_{TV} \leq \min_{u \leq \varepsilon} \left[C \exp(-\lambda_{u,\varepsilon} n \Delta) \times \sqrt{\frac{n\mu_4^2(u)}{\Delta^2(\Sigma^2 + \sigma^2(u))^4} + \frac{n\mu_3^2(u)}{\Delta(\Sigma^2 + \sigma^2(u))^3}} + [1 - \exp(-\lambda_{u,\varepsilon} \Delta n)] \right]. \quad (21)$$

The bounds on the total variation we establish in Theorem 5 and Corollary 2 are tight in the following sense. Whenever $(\lambda_{u^*,\varepsilon} n \Delta) \vee \frac{n\mu_4^2(u^*)}{\Delta^2(\Sigma^2 + \sigma^2(u^*))^4} \vee \frac{n\mu_3^2(u^*)}{\Delta(\Sigma^2 + \sigma^2(u^*))^3}$ does not converge to 0, the total variation distance does not converge to 0. Note that if it converges to $+\infty$, the total variation converges to 1. Moreover, if $(\lambda_{u^*,\varepsilon} n \Delta) \vee \frac{n\mu_4^2(u^*)}{\Delta^2(\Sigma^2 + \sigma^2(u^*))^4} \vee \frac{n\mu_3^2(u^*)}{\Delta(\Sigma^2 + \sigma^2(u^*))^3}$ converges to 0 with n , then the total variation converges to 0 by Theorem 3. Another implication of these bounds is that the Gaussian random variable closest to $(\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(\varepsilon))^{\otimes n}$ is not necessary $\mathcal{N}(b\Delta, \Delta(\Sigma^2 + \sigma^2(\varepsilon)))^{\otimes n}$, in particular when rare and large jumps are present, a tighter Gaussian approximation is provided by $\mathcal{N}(b\Delta, \Delta(\Sigma^2 + \sigma^2(u^*)))^{\otimes n}$.

The lower bound on the total variation distance is a jump detection test. The proof of Theorem 5 is based on the construction of a test of Gaussianity, adapted to Lévy processes, that detects whether the discrete observations we have at our disposal are purely Gaussian, or whether they are realizations of a Lévy process with non trivial Lévy measure. More precisely, (see the proof of Theorem 5 for details) we build a uniformly consistent test for the testing problem

$$H_0 : \nu_\varepsilon = 0, \quad \text{against} \quad H_1 : \lim_{\eta \rightarrow 0} \lambda_{\eta,\varepsilon} = +\infty \quad \text{and} \quad E$$

¹Of all increments when $\nu_\varepsilon = 0$ and of those where a jump larger than u^* occurs otherwise.

where

$$E = \left[\mu_3(u^*)^2 \geq \frac{C\Delta(\Sigma^2 + \sigma^2(u^*))^3}{n} \text{ or } \mu_4(u^*)^2 \geq \frac{C\Delta^2(\Sigma^2 + \sigma^2(u^*))^4}{n} \right. \\ \left. \text{or a jump larger than } u^* \text{ occurs} \right].$$

This test can be of interest in itself: it does not need the knowledge of the Lévy triplet. At the expense of additional technicalities, one could extend this test - and therefore the lower bound on the total variation distance - to Lévy measures that are not absolutely continuous with respect to the Lebesgue measure.

Improvement of Theorem 4 for mixtures. An immediate corollary of Theorem 4 following from its proof is a lower bound on the total variation distance between any two mixture of Gaussian random variables and mixture of infinite Lévy measures concentrated in $[-\varepsilon, \varepsilon]$. More precisely, let $d\Lambda(b, \Sigma, \nu)$ and $d\Lambda'(b, \Sigma)$ be two priors on Lévy processes and Brownian motions, respectively. Assume that the support of $d\Lambda(b, \Sigma, \nu)$ is included in a set \mathcal{A} , and that for any $(b, \Sigma, \nu) \in \mathcal{A}$, we have $\lambda_{\varepsilon/2, \varepsilon} \geq \log(n)/(n\Delta)$. Then, it holds that

$$\left\| \int (\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_{\Delta}^{(\nu)}(\varepsilon))^{\otimes n} d\Lambda(b, \Sigma, \nu) - \int \mathcal{N}(b\Delta, \Delta\Sigma^2)^{\otimes n} d\Lambda'(b, \Sigma) \right\|_{TV} \geq \\ \min_{(b, \Sigma, \nu) \in \mathcal{A}} \left[\left\{ 1 - C \left[\frac{\Delta(\Sigma^2 + \sigma_{\nu}^2(\varepsilon))^3}{n\mu_3^{(\nu)}(\varepsilon)^2} \wedge \frac{\Delta^2(\Sigma^2 + \sigma_{\nu}^2(\varepsilon))^4}{n\mu_4^{(\nu)}(\varepsilon)^2} \right] - \alpha_n \right\} \right],$$

whereby $M_{\Delta}^{(\nu)}$, $\sigma_{\nu}^2(\varepsilon)$, $\mu_3^{(\nu)}(\varepsilon)$ and $\mu_4^{(\nu)}(\varepsilon)$ correspond to M_{Δ} , $\sigma^2(\varepsilon)$, $\mu_3(\varepsilon)$ and $\mu_4(\varepsilon)$ for the Lévy measure ν . A related result can be achieved for Theorem 5. Note that the corresponding lower bound on the total variation distance is a direct corollary of Theorem 3, but the lower bound displayed above it is not trivial and it holds because the test that we construct in the proof of Theorem 5 does not depend on the parameters of the Gaussian random variable or on the Lévy triplet.

Case of a finite Lévy measure. Let us emphasize that the assumption $n\Delta \geq 1$ is not too restrictive: it is implied by the fact that the observation interval is at least 1. But it is not necessary, and neither is the assumption that the Lévy measure is infinite. What one really needs for establishing Theorem 5 is only that $\lambda_{0, \varepsilon} \geq \Delta^{-1} \vee (\log(n)/(n\Delta))$.

If this is not the case, the asymptotic total variation is either 0 or 1, as discussed below.

- **Case 1 :** $a_n \Delta^{-1} n^{-1} \geq \lambda_{0, \varepsilon}$ **where** $a_n \rightarrow 0$. In this case one does never observe any jump with probability going to 1. So the total variation distance goes to 0.
- **Case 2 (only in the noiseless case $b = 0, \Sigma = 0$):** $a_n \Delta^{-1} n^{-1} \leq \lambda_{0, \varepsilon} \leq A_n \log(n) \Delta^{-1}$ **where** $a_n \rightarrow \infty$ **and** $A_n \rightarrow 0$. In this case the probability of observing at least a time step where one, and only one, jump occurs goes to 1, as well as the probability of having at least one time step where no jump occur goes to 1. We conclude therefore that the total variation distance goes to 1: such a process is very different from a Gaussian process.

2.3 Examples

In this Section we illustrate the implications of Theorems 2, 3 and Corollary 2 on the class of infinite stable processes. It is easy to extend the results valid for this example to other types of Lévy processes (e.g. inverse Gaussian processes, tempered stable distributions, etc..) as, around 0, stable measures well approximate many Lévy measures. Let $\beta \in (0, 2)$, $c_+, c_- \geq 0$, ($c_+ = c_-$ if $\beta = 1$) and set

$$\nu(dx) = \frac{c_+}{x^{1+\beta}} \mathbf{1}_{(0, +\infty)}(x) dx + \frac{c_-}{|x|^{1+\beta}} \mathbf{1}_{(-\infty, 0)}(x) dx.$$

In particular, following the notation used in the paper, $M(\varepsilon)$ will be a Lévy process with Lévy triplet $(0, 0, \nu_\varepsilon)$ where $\nu_\varepsilon := \nu \mathbf{1}_{|x| \leq \varepsilon}$.

Suppose we are interested in the question: “Given n and Δ , what is the largest ε such that it is not possible to distinguish between n independent realizations of $\mathcal{N}(0, \Delta \Sigma^2) * M_\Delta(\varepsilon)$ and the Gaussian vector $\mathcal{N}(0, \Delta(\sigma^2(\varepsilon) + \Sigma^2))^{\otimes n}$?” The answer to this question is provided by Theorems 3 and 4 which prescribe to choose the largest $\varepsilon \in (0, 1]$ such that

$$\frac{\sqrt{n}\mu_3(\varepsilon)}{\sqrt{\Delta(\Sigma^2 + \sigma^2(\varepsilon))^3}} \rightarrow 0 \quad \text{and} \quad \frac{\sqrt{n}\mu_4(\varepsilon)}{\Delta(\Sigma^2 + \sigma^2(\varepsilon))^2} \rightarrow 0.$$

Indeed, with such a choice of (n, Δ, ε) from Theorem 3 we get

$$\|(\mathcal{N}(0, \Delta \Sigma^2) * M_\Delta(\varepsilon))^{\otimes n} - \mathcal{N}(0, \Delta(\sigma^2(\varepsilon) + \Sigma^2))^{\otimes n}\|_{TV} \rightarrow 0, \quad (22)$$

while for any $(n, \Delta, \varepsilon^*)$ with $\varepsilon \ll \varepsilon^*$, Theorem 4 ensures that

$$\|(\mathcal{N}(0, \Delta \Sigma^2) * M_\Delta(\varepsilon^*))^{\otimes n} - \mathcal{N}(0, \Delta(\sigma^2(\varepsilon^*) + \Sigma^2))^{\otimes n}\|_{TV} \not\rightarrow 0.$$

In the following Table, each cell displays the order of the maximal value ε^* for $\varepsilon \in (0, 1]$ such that Equation (22) holds.

ν β -stable law, $\beta \in (0, 2)$	ν symmetric		ν non-symmetric	
Regime	$\frac{\sqrt{n}}{\Delta} = o(1)$	$\frac{\Delta}{\sqrt{n}} = o(1)$	$\frac{\sqrt{n}}{\sqrt{\Delta}} = o(1)$	$\frac{\sqrt{\Delta}}{\sqrt{n}} = o(1)$
$\Sigma > 0$	$\varepsilon^* \approx 1$	$\varepsilon^* \approx \left(\frac{\Delta}{\sqrt{n}}\right)^{\frac{1}{4-\beta}}$	$\varepsilon^* \approx 1$	$\varepsilon^* \approx \left(\frac{\sqrt{\Delta}}{\sqrt{n}}\right)^{\frac{1}{3-\beta}}$
$\Sigma = 0$	$\varepsilon^* \approx 1$	$\varepsilon^* \approx \left(\frac{\Delta}{\sqrt{n}}\right)^{\frac{1}{\beta}}$	$\varepsilon^* \approx 1$	$\varepsilon^* \approx \left(\frac{\sqrt{\Delta}}{\sqrt{n}}\right)^{\frac{1}{\beta}}$

If, instead, we were interested in the question: “For a given pair (n, Δ) and a precision $p = o(1)$ as n goes to infinity, how large can ε be so that the total variation distance (15) remains of order p ?”, the answer would be slightly different. Again, we could apply Theorem 3 and thus choose the largest $\varepsilon \in (0, 1]$ such that

$$\frac{\sqrt{n}\mu_3(\varepsilon)}{\sqrt{\Delta(\Sigma^2 + \sigma^2(\varepsilon))^3}} \lesssim p \quad \text{and} \quad \frac{\sqrt{n}\mu_4(\varepsilon)}{\Delta(\Sigma^2 + \sigma^2(\varepsilon))^2} \lesssim p. \quad (23)$$

However, observe that in this case Theorem 4 does not provide a meaningful lower bound for the total variation, so there is no guarantee that the largest ε such that (23) holds is the optimal choice for the size of the small jumps. Finally, if one is also interested in constants then we suggest to apply Corollary 2 rather than Theorem 3. More precisely, the idea is to choose the largest $u > 0$ such that

$$\sqrt{\left(1 + \frac{\mu_4^2(u)}{\Delta^2(\Sigma^2 + \sigma^2(u))^4} + \frac{7}{6} \frac{\mu_3^2(u)}{\Delta(\Sigma^2 + \sigma^2(u))^3}\right)^n - 1} \leq \frac{p}{6}$$

and then take $\varepsilon > u$ such that $\lambda_{u, \varepsilon} \leq \frac{p}{2n\Delta}$. In this way, by means of Corollary 2 and its proof, we get

$$\|(\mathcal{N}(0, \Delta \Sigma^2) * M_\Delta(\varepsilon))^{\otimes n} - (\mathcal{N}(0, \Delta(\sigma^2(u) + \Sigma^2)))^{\otimes n}\|_{TV} \leq p.$$

3 Total variation distance between Lévy processes

In this Section, let $X^i \sim (b_i, \Sigma_i^2, \nu_i)$, $i = 1, 2$, be two distinct Lévy processes with infinite Lévy measures. We shall use the notation introduced in Section 1.1 properly modified to take into account

the dependencies on X^1 and X^2 . For instance, $\mu_3(\varepsilon)$ and $\mu_4(\varepsilon)$ become

$$\mu_{j,i}(\varepsilon) = \int_{|x| \leq \varepsilon} x^j \nu_i(dx), \quad i = 1, 2, \quad j = 3, 4,$$

where $\mu_{j,1}(\varepsilon)$ (resp. $\mu_{j,2}(\varepsilon)$), $j = 3, 4$, denote the 3rd and 4th moment of ν_1 (resp. ν_2) restricted on $\{x : |x| \leq \varepsilon\}$.

By means of the Lévy-Itô decomposition recalled in Section 1.1, for any $t > 0$ and $\varepsilon > 0$ we have that the law of X_t^i , $i = 1, 2$, is the convolution between a Gaussian distribution and the law of the marginal at time t of the processes $M^i(\varepsilon)$ and $Z^i(\varepsilon)$, i.e.

$$X_t^i(\varepsilon) = \mathcal{N}(b_i(\varepsilon)t, t\Sigma_i^2) * M_t^i(\varepsilon) * Z_t^i(\varepsilon), \quad i = 1, 2.$$

By subadditivity of the total variation distance, see Lemma 12 in Appendix B, we have:

$$\|X_t^1 - X_t^2\|_{TV} \leq \|\mathcal{N}(b_1(\varepsilon)t, t\Sigma_1^2) * M_t^1(\varepsilon) - \mathcal{N}(b_2(\varepsilon)t, t\Sigma_2^2) * M_t^2(\varepsilon)\|_{TV} + \|Z_t^1(\varepsilon) - Z_t^2(\varepsilon)\|_{TV}.$$

By triangular inequality first and subadditivity of the total variation distance together with Lemma 10 in Appendix B then, we obtain

$$\begin{aligned} & \|\mathcal{N}(b_1(\varepsilon)t, t\Sigma_1^2) * M_t^1(\varepsilon) - \mathcal{N}(b_2(\varepsilon)t, t\Sigma_2^2) * M_t^2(\varepsilon)\|_{TV} \\ & \leq \|\mathcal{N}(b_1(\varepsilon)t, t\Sigma_1^2) * M_t^1(\varepsilon) - \mathcal{N}(b_1(\varepsilon)t, t\Sigma_1^2) * \mathcal{N}(0, t\sigma_1^2(\varepsilon))\|_{TV} \\ & \quad + \|\mathcal{N}(b_2(\varepsilon)t, t\Sigma_2^2) * M_t^2(\varepsilon) - \mathcal{N}(b_2(\varepsilon)t, t\Sigma_2^2) * \mathcal{N}(0, t\sigma_2^2(\varepsilon))\|_{TV} \\ & \quad + \|\mathcal{N}(b_1(\varepsilon)t, t\Sigma_1^2) * \mathcal{N}(0, t\sigma_1^2(\varepsilon)) - \mathcal{N}(b_2(\varepsilon)t, t\Sigma_2^2) * \mathcal{N}(0, t\sigma_2^2(\varepsilon))\|_{TV} \\ & \leq \sum_{i=1}^2 \|\mathcal{N}(b_i(\varepsilon)t, t\Sigma_i^2) * M_t^i(\varepsilon) - \mathcal{N}(b_i(\varepsilon)t, t(\sigma_i^2(\varepsilon) + \Sigma_i^2))\|_{TV} \\ & \quad + \frac{\sqrt{\frac{t}{2\pi}}|b_1(\varepsilon) - b_2(\varepsilon)| + \left| \sqrt{\Sigma_1^2 + \sigma_1^2(\varepsilon)} - \sqrt{\Sigma_2^2 + \sigma_2^2(\varepsilon)} \right|}{\sqrt{\Sigma_1^2 + \sigma_1^2(\varepsilon)} \vee \sqrt{\Sigma_2^2 + \sigma_2^2(\varepsilon)}}. \end{aligned}$$

The terms $\|\mathcal{N}(b_i(\varepsilon)t, t\Sigma_i^2) * M_t^i(\varepsilon) - \mathcal{N}(b_i(\varepsilon)t, t(\Sigma_i^2 + \sigma_i^2(\varepsilon)))\|_{TV}$, $i = 1, 2$ can be bounded by means of Theorem 3 whereas for $\|Z_t^1(\varepsilon) - Z_t^2(\varepsilon)\|_{TV}$ we can use Lemma 11 taking $n = 1$:

$$\|Z_t^1(\varepsilon) - Z_t^2(\varepsilon)\|_{TV} \leq t|\Lambda_1(\varepsilon) - \Lambda_2(\varepsilon)| + t(\Lambda_1(\varepsilon) \wedge \Lambda_2(\varepsilon)) \left\| \frac{\nu_1^\varepsilon}{\Lambda_1(\varepsilon)} - \frac{\nu_2^\varepsilon}{\Lambda_2(\varepsilon)} \right\|_{TV},$$

with $\nu_j^\varepsilon = \nu_j(\cdot \cap (\mathbb{R} \setminus (-\varepsilon, \varepsilon)))$ and $\Lambda_j(\varepsilon) = \nu_j^\varepsilon(\mathbb{R})$. Applying Lemma 10 for $n = 1$, we thus obtain the following upper bounds for the total variation distance between marginals of Lévy processes:

$$\begin{aligned} \|X_t^1 - X_t^2\|_{TV} & \leq \frac{\sqrt{\frac{t}{2\pi}}|b_1(\varepsilon) - b_2(\varepsilon)| + \left| \sqrt{\Sigma_1^2 + \sigma_1^2(\varepsilon)} - \sqrt{\Sigma_2^2 + \sigma_2^2(\varepsilon)} \right|}{\sqrt{\Sigma_1^2 + \sigma_1^2(\varepsilon)} \vee \sqrt{\Sigma_2^2 + \sigma_2^2(\varepsilon)}} \\ & \quad + 3 \sum_{i=1}^2 \sqrt{\frac{\mu_{4,i}^2(\varepsilon)}{t^2(\sigma_i^2(\varepsilon) + \Sigma_i^2)^4} + \frac{7}{6} \frac{\mu_{3,i}^2(\varepsilon)}{t(\Sigma_i^2 + \sigma_i^2(\varepsilon))^3}} \\ & \quad + t|\Lambda_1(\varepsilon) - \Lambda_2(\varepsilon)| + t(\Lambda_1(\varepsilon) \wedge \Lambda_2(\varepsilon)) \left\| \frac{\nu_1^\varepsilon}{\Lambda_1(\varepsilon)} - \frac{\nu_2^\varepsilon}{\Lambda_2(\varepsilon)} \right\|_{TV}. \end{aligned}$$

Extending these arguments to the case of discrete observations of Lévy processes we obtain the following result.

Theorem 6. Let $X^i \sim (b_i, \Sigma_i^2, \nu_i)$ be any Lévy process with $b_i \in \mathbb{R}$, $\Sigma_i \geq 0$ and ν_i infinite Lévy measure absolutely continuous with respect to the Lebesgue measure, $i = 1, 2$. For all $\Delta > 0$, $\varepsilon \in (0, 1)$ and $n \geq 1$:

$$\begin{aligned} \|(X_{k\Delta}^1 - X_{(k-1)\Delta}^1)_{k=1}^n - (X_{k\Delta}^2 - X_{(k-1)\Delta}^2)_{k=1}^n\|_{TV} &\leq \frac{\sqrt{n\Delta}}{\sqrt{2\pi}} \frac{|b_1(\varepsilon) - b_2(\varepsilon)|}{\max(\sqrt{\Sigma_1^2 + \sigma_1^2(\varepsilon)}, \sqrt{\Sigma_2^2 + \sigma_2^2(\varepsilon)})} \\ &+ 1 - \left(\frac{\min(\sqrt{\Sigma_1^2 + \sigma_1^2(\varepsilon)}, \sqrt{\Sigma_2^2 + \sigma_2^2(\varepsilon)})}{\max(\sqrt{\Sigma_1^2 + \sigma_1^2(\varepsilon)}, \sqrt{\Sigma_2^2 + \sigma_2^2(\varepsilon)})} \right)^n \\ &+ 3 \sum_{i=1}^2 \sqrt{\left(1 + \frac{\mu_{4,i}^2(\varepsilon)}{\Delta^2(\sigma_i^2(\varepsilon) + \Sigma_i^2)^4} + \frac{7}{6} \frac{\mu_{3,i}^2(\varepsilon)}{\Delta(\sigma_i^2(\varepsilon) + \Sigma_i^2)^3} \right)^n - 1}. \\ &+ 1 - \exp(-n\Delta|\Lambda_1(\varepsilon) - \Lambda_2(\varepsilon)|) + n\Delta(\Lambda_1(\varepsilon) \wedge \Lambda_2(\varepsilon)) \left\| \frac{\nu_1^\varepsilon}{\Lambda_1(\varepsilon)} - \frac{\nu_2^\varepsilon}{\Lambda_2(\varepsilon)} \right\|_{TV}. \end{aligned}$$

Proof. It directly follows from the Lévy-Itô decomposition together with Lemmas 12, 10, 11 and Theorem 3. \square

4 Proofs

Notations.

$$\begin{aligned} \bar{\mu}_{\eta,\varepsilon} &= \mathbb{E}[Y_1] = \frac{\int_{\eta < |x| \leq \varepsilon} x\nu(dx)}{\lambda_{\eta,\varepsilon}}, \quad \tilde{\sigma}^2(\eta, \varepsilon) = \lambda_{\eta,\varepsilon} \mathbb{V}(Y_1) = \sigma^2(\eta, \varepsilon) - \bar{\mu}_{\eta,\varepsilon}^2 \lambda_{\eta,\varepsilon}; \\ \gamma_{\eta,\varepsilon} &= \frac{\mathbb{E}[(Y_1 - \mathbb{E}[Y_1])^3]}{\mathbb{V}(Y_1)^{3/2}} = \frac{\sqrt{\lambda_{\eta,\varepsilon}} \left(\int_{\eta < |x| < \varepsilon} x^3 \nu(dx) - 3\bar{\mu}_{\eta,\varepsilon} \sigma^2(\eta, \varepsilon) + 2\lambda_{\eta,\varepsilon} \bar{\mu}_{\eta,\varepsilon}^3 \right)}{\tilde{\sigma}^3(\eta, \varepsilon)}, \quad \tau_{\eta,\varepsilon} = \frac{\mathbb{E}[(Y_1 - \bar{\mu}_{\eta,\varepsilon})^4]}{\mathbb{V}(Y_1)^2}. \end{aligned}$$

In the sequel, the dependency in η and ε of $\bar{\mu}$, λ , σ^2 , $\tilde{\sigma}^2$, γ and τ is not always made explicit to ease the reading. We define the set $K_\psi := \{k \in \mathbb{N}, |k - \lambda\Delta| \leq \psi\lambda\Delta\}$ where $\psi \in (0, 1)$ and we denote by H_n , $n \in \mathbb{N}$, the Hermite polynomials. We recall that:

$$H_3(x) = x^3 - 3x, \quad H_4(x) = x^4 - 6x^2 + 3, \quad H_6(x) = x^6 - 15x^4 + 45x^2 - 15.$$

Moreover $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ denotes the density function of the standard Gaussian random variable. Finally, we denote by $*$ the convolution product and thus the writing $X * Y$ will stand for $\mathcal{L}(X + Y)$, for independent random variables X and Y .

4.1 Proofs of the Propositions used in the proof of Theorem 2

4.1.1 Proof of Proposition 1

To lighten the notation we shall sometimes write $N(\varepsilon', \varepsilon)$ instead of $\mathcal{N}(0, \Delta\sigma^2(\varepsilon', \varepsilon))$. Let $\eta < \varepsilon' < \varepsilon^2$, by triangular inequality and Lemma 12, we have

$$\begin{aligned} \|M_\Delta(\varepsilon)^{\otimes n} - M_\Delta(\eta, \varepsilon)^{\otimes n}\|_{TV} &= \|(M_\Delta(\varepsilon') * M_\Delta(\varepsilon', \varepsilon))^{\otimes n} - (M_\Delta(\eta, \varepsilon') * M_\Delta(\varepsilon', \varepsilon))^{\otimes n}\|_{TV} \\ &\leq \|(M_\Delta(\varepsilon') * M_\Delta(\varepsilon', \varepsilon))^{\otimes n} - (M_\Delta(\varepsilon') * N(\varepsilon', \varepsilon))^{\otimes n}\|_{TV} \\ &\quad + \|(M_\Delta(\varepsilon') * N(\varepsilon', \varepsilon))^{\otimes n} - (M_\Delta(\eta, \varepsilon') * N(\varepsilon', \varepsilon))^{\otimes n}\|_{TV} \\ &\quad + \|(M_\Delta(\eta, \varepsilon') * N(\varepsilon', \varepsilon))^{\otimes n} - (M_\Delta(\eta, \varepsilon') * M_\Delta(\varepsilon', \varepsilon))^{\otimes n}\|_{TV} \\ &\leq 2\|(M_\Delta(\varepsilon', \varepsilon) - N(\varepsilon', \varepsilon))^{\otimes n}\|_{TV} + \|(M_\Delta(\varepsilon') * N(\varepsilon', \varepsilon))^{\otimes n} - (M_\Delta(\eta, \varepsilon') * N(\varepsilon', \varepsilon))^{\otimes n}\|_{TV} \end{aligned}$$

We are left to prove that

$$\lim_{\eta \rightarrow 0} \|(M_\Delta(\varepsilon') * N(\varepsilon', \varepsilon))^{\otimes n} - (M_\Delta(\eta, \varepsilon') * N(\varepsilon', \varepsilon))^{\otimes n}\|_{TV} = 0. \quad (24)$$

From Lemma 13, it follows that

$$\|(M_\Delta(\varepsilon') * N(\varepsilon', \varepsilon))^{\otimes n} - (M_\Delta(\eta, \varepsilon') * N(\varepsilon', \varepsilon))^{\otimes n}\|_{TV} \leq \sqrt{2n} \|M_\Delta(\varepsilon') * N(\varepsilon', \varepsilon) - M_\Delta(\eta, \varepsilon') * N(\varepsilon', \varepsilon)\|_{TV}.$$

Therefore, in order to prove (24), n being fixed, it is enough to prove that

$$\lim_{\eta \rightarrow 0} \|M_\Delta(\varepsilon') * N(\varepsilon', \varepsilon) - M_\Delta(\eta, \varepsilon') * N(\varepsilon', \varepsilon)\|_{TV} = 0. \quad (25)$$

Let us denote by $h_{\varepsilon'}$ (resp. $h_{\eta, \varepsilon'}$) the density with respect to the Lebesgue measure of $M_\Delta(\varepsilon') * N(\varepsilon', \varepsilon)$ (resp. $M_\Delta(\eta, \varepsilon') * N(\varepsilon', \varepsilon)$). By Scheffé Lemma, (25) is equivalent to

$$\lim_{\eta \rightarrow 0} |h_{\varepsilon'}(x) - h_{\eta, \varepsilon'}(x)| = 0, \quad \text{Lebesgue a.e.}$$

Observe that the characteristics functions of $M_\Delta(\varepsilon') * N(\varepsilon', \varepsilon)$ and $M_\Delta(\eta, \varepsilon') * N(\varepsilon', \varepsilon)$ are in $L_1(\text{Leb})$, therefore, by the inversion theorem we have that:

$$\begin{aligned} h_{\varepsilon'}(x) &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iux} \mathbb{E}[e^{iu(M_\Delta(\varepsilon') * N(\varepsilon', \varepsilon))}] du, \\ h_{\eta, \varepsilon'}(x) &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iux} \mathbb{E}[e^{iu(M_\Delta(\eta, \varepsilon') * N(\varepsilon', \varepsilon))}] du. \end{aligned}$$

Thus, for any $x \in \mathbb{R}$, we have

$$\begin{aligned} |h_{\varepsilon'}(x) - h_{\eta, \varepsilon'}(x)| &\leq \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\frac{\Delta u^2 \sigma^2(\varepsilon', \varepsilon)}{2}} |\mathbb{E}[e^{iuM_\Delta(\varepsilon')}] - \mathbb{E}[e^{iuM_\Delta(\eta, \varepsilon')}]| du \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\frac{\Delta u^2 \sigma^2(\varepsilon', \varepsilon)}{2}} |\mathbb{E}[e^{iuM_\Delta(\eta, \varepsilon')}]| |\mathbb{E}[e^{iuM_\Delta(\eta)}] - 1| du \\ &\leq \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\frac{\Delta u^2 \sigma^2(\varepsilon', \varepsilon)}{2}} |\mathbb{E}[e^{iuM_\Delta(\eta)}] - 1| du. \end{aligned}$$

Clearly, $\lim_{\eta \rightarrow 0} |\mathbb{E}[e^{iuM_\Delta(\eta)}] - 1| = 0$. Indeed, using that $|1 + iux - e^{iux}| \leq \frac{u^2 x^2}{2}$, we immediately obtain:

$$|\mathbb{E}[e^{iuM_\Delta(\eta)}] - 1| \leq \left| \Delta \int_{|x| < \eta} (1 + iux - e^{iux}) \nu(dx) \right| \leq \frac{\Delta}{2} \int_{|x| < \eta} u^2 x^2 \nu(dx) \rightarrow 0.$$

Moreover, since $\varepsilon' < \varepsilon^2$

$$e^{-\frac{\Delta u^2 \sigma^2(\varepsilon', \varepsilon)}{2}} |\mathbb{E}[e^{iuM_\Delta(\eta)}] - 1| \leq 2e^{-\frac{\Delta u^2 \sigma^2(\varepsilon^2, \varepsilon)}{2}} \in L_1(\text{Leb}),$$

an application of the dominated convergence theorem yields $\lim_{\eta \rightarrow 0} |h_{\varepsilon'}(x) - h_{\eta, \varepsilon'}(x)| = 0$, which completes the proof.

4.1.2 Proof of Proposition 2

We first establish a technical Lemma.

Lemma 2. *For any infinite Lévy measure ν it holds*

$$\lim_{\eta \rightarrow 0} \frac{1}{\lambda_{\eta, \varepsilon}} \left(\int_{\eta < |x| < \varepsilon} x \nu(dx) \right)^2 = 0. \quad (26)$$

Proof of Lemma 2. Let $0 < u < \varepsilon$ be arbitrarily. Then, using that $\lim_{\eta \rightarrow 0} \lambda_{\eta, \varepsilon} = +\infty$ and $\int_{u < |x| < \varepsilon} x\nu(dx)$ is finite, one has

$$\begin{aligned} \lim_{\eta \rightarrow 0} \frac{1}{\lambda_{\eta, \varepsilon}} \left(\int_{\eta < |x| < \varepsilon} x\nu(dx) \right)^2 &\leq \lim_{\eta \rightarrow 0} \frac{2}{\lambda_{\eta, \varepsilon}} \left(\left(\int_{\eta < |x| < u} x\nu(dx) \right)^2 + \left(\int_{u < |x| < \varepsilon} x\nu(dx) \right)^2 \right) \\ &= \lim_{\eta \rightarrow 0} \frac{2}{\lambda_{\eta, \varepsilon}} \left(\int_{\eta < |x| < u} x\nu(dx) \right)^2. \end{aligned}$$

For any $u \geq \eta > 0$, it holds by Cauchy-Schwartz that

$$\frac{1}{\lambda_{\eta, u}} \left(\int_{\eta < |x| < u} x\nu(dx) \right)^2 \leq \int_{\eta < |x| < u} x^2\nu(dx).$$

From $\lambda_{\eta, u} \leq \lambda_{\eta, \varepsilon}$ we deduce

$$\frac{1}{\lambda_{\eta, \varepsilon}} \left(\int_{\eta < |x| < u} x\nu(dx) \right)^2 \leq \int_{\eta < |x| < u} x^2\nu(dx)$$

and therefore

$$\lim_{\eta \rightarrow 0} \frac{1}{\lambda_{\eta, \varepsilon}} \left(\int_{\eta < |x| < \varepsilon} x\nu(dx) \right)^2 \leq \lim_{\eta \rightarrow 0} 2 \int_{\eta < |x| < u} x^2\nu(dx), \quad \forall 0 < u < \varepsilon.$$

As this holds for any $u \in (0, \varepsilon)$ and as $\int_{\eta < |x| < \varepsilon} x^2\nu(dx) < +\infty$, we conclude

$$\lim_{\eta \rightarrow 0} \frac{1}{\lambda_{\eta, \varepsilon}} \left(\int_{\eta < |x| < \varepsilon} x\nu(dx) \right)^2 = 0.$$

□

Proof of Proposition 2. Set $\tilde{Y}_1 = Y_1 - \mathbb{E}[Y_1]$ and let g be the density of \tilde{Y}_1 with respect to the Lebesgue measure and write $\lambda := \lambda_{\eta, \varepsilon}$. Using (7), we derive that

$$\begin{aligned} f_{\eta, \varepsilon, A}(x) &= \frac{1}{\mathbb{P}(A)} \sum_{k \in K_\psi} \mathbb{P}(N_\Delta = k) g^{*k}(x + (\lambda\Delta - k)\bar{\mu}) \\ \tilde{f}_{\eta, \varepsilon, A}(x) &= \frac{1}{\mathbb{P}(A)} \sum_{k \in K_\psi} \mathbb{P}(N_\Delta = k) g^{*k}(x). \end{aligned} \tag{27}$$

The triangle inequality leads to

$$\int |f_{\eta, \varepsilon, A}(x) - \tilde{f}_{\eta, \varepsilon, A}(x)| dx \leq \frac{1}{\mathbb{P}(A)} \sum_{k \in K_\psi} \mathbb{P}(N_\Delta = k) \int |g^{*k}(x + (\lambda\Delta - k)\bar{\mu}) - g^{*k}(x)| dx.$$

By the assumption $\nu(\mathbb{R}) = \infty$, it follows that $\lambda := \lambda_{\eta, \varepsilon} \rightarrow \infty$ as $\eta \rightarrow 0$ and therefore a Gaussian approximation for g^{*k} is appropriate for $k \in K_\psi$. We have:

$$\begin{aligned} g^{*k}(x + (\lambda\Delta - k)\bar{\mu}) &= \frac{1}{\sqrt{k\mathbb{V}(\tilde{Y}_1)}} \varphi \left(\frac{x + (\lambda\Delta - k)\bar{\mu}}{\sqrt{k\mathbb{V}(\tilde{Y}_1)}} \right) (1 + o(1)), \\ g^{*k}(x) &= \frac{1}{\sqrt{k\mathbb{V}(\tilde{Y}_1)}} \varphi \left(\frac{x}{\sqrt{k\mathbb{V}(\tilde{Y}_1)}} \right) (1 + o(1)). \end{aligned}$$

Using Lemma 10, we get

$$\begin{aligned} \int |g^{*k}(x + (\lambda\Delta - k)\bar{\mu}) - g^{*k}(x)| dx &= 2\|\mathcal{N}((\lambda\Delta - k)\mu, k\mathbb{V}(\tilde{Y}_1)) - \mathcal{N}(0, k\mathbb{V}(\tilde{Y}_1))\|_{TV} + o(1) \\ &\leq \frac{2}{\sqrt{2\pi}} \frac{|(\lambda\Delta - k)\bar{\mu}|}{\sqrt{k\mathbb{V}(\tilde{Y}_1)}} + o(1). \end{aligned}$$

We are left to control the following sum

$$\begin{aligned} \frac{2}{\mathbb{P}(A)} \sum_{k \in K_\psi} \frac{\mathbb{P}(N_\Delta = k)}{\sqrt{2\pi}} \frac{|(\lambda\Delta - k)\bar{\mu}|}{\sqrt{k\mathbb{V}(\tilde{Y}_1)}} &= \frac{2\sqrt{\lambda\bar{\mu}^2}}{\sqrt{2\pi}\mathbb{P}(A)} \sum_{k \in K_\psi} \mathbb{P}(N_\Delta = k) \frac{|\lambda\Delta - k|}{\sqrt{k(\sigma^2 - \lambda\bar{\mu}^2)}} \\ &\leq \frac{2\sqrt{\lambda\bar{\mu}^2}}{\mathbb{P}(A)} \frac{\mathbb{E}[|N_\Delta - \lambda\Delta|]}{\sqrt{2\pi(\sigma^2 - \lambda\bar{\mu}^2)\lambda\Delta(1 - \psi)}} \\ &\leq \frac{\sqrt{\lambda\bar{\mu}^2}}{\sqrt{2\pi(\sigma^2 - \lambda\bar{\mu}^2)(1 - \psi)}}, \end{aligned}$$

where we used that for η large enough $\mathbb{P}(A) \geq \frac{1}{2}$ since from (11) we have $\mathbb{P}(A^c) \leq e^{-\frac{\psi^2\lambda\Delta}{2}}$, $\mathbb{V}(\tilde{Y}_1) = \frac{\sigma^2 - \lambda\bar{\mu}^2}{\lambda}$ and $\mathbb{E}[|N_\Delta - \lambda\Delta|] \leq \sqrt{\lambda\Delta}$. Finally, using (26) which ensures that $\lambda\bar{\mu}^2 \rightarrow 0$ as $\eta \rightarrow 0$ together with $\sigma^2(\eta, \varepsilon) \rightarrow \sigma^2(\varepsilon) > 0$, as $\eta \rightarrow 0$, we get $\frac{\bar{\mu}^2\lambda_{\eta,\varepsilon}}{\sigma^2(\eta,\varepsilon) - \lambda_{\eta,\varepsilon}\bar{\mu}^2} \rightarrow 0$ as $\eta \rightarrow 0$; the proof is thus complete.

4.1.3 Proof of Proposition 3

From (27), the χ^2 distance between $\tilde{f}_{\eta,\varepsilon,A}$ and the corresponding Gaussian density $\tilde{\varphi}_{\eta,\varepsilon}$ is given by:

$$\chi^2(\tilde{f}_{\eta,\varepsilon,A}, \tilde{\varphi}_{\eta,\varepsilon}) + 1 = \int \frac{\tilde{f}_{\eta,\varepsilon,A}(x)^2}{\tilde{\varphi}_{\eta,\varepsilon}(x)} dx = \mathbb{P}(A)^{-2} e^{-2\lambda\Delta} \sum_{j,k \in K_\psi} \frac{(\lambda\Delta)^{j+k}}{j!k!} \int \frac{g^{*k}(x)g^{*j}(x)}{\tilde{\varphi}_{\eta,\varepsilon}(x)} dx, \quad (28)$$

where g denotes the density of $Y_1 - \mathbb{E}[Y_1]$ as in the proof of Proposition 2. A first difficulty in treating (28) comes from the integral. As in the proof of Proposition 2, we use that g^{*k} and g^{*j} are well approximated by Gaussian densities. However, here we want a sharp rate of convergence to thoroughly quantify the proximity between g^{*k} and g^{*j} and their corresponding Gaussian counterparts. We first normalize the densities and denote by $\tilde{g}^{*k}(x)$ the density of $\frac{1}{\sqrt{k\mathbb{V}(Y_1)}} \sum_{i=1}^k \tilde{Y}_i$. In particular, we have that

$$g^{*k}(x) = \frac{1}{\sqrt{k\mathbb{V}(Y_1)}} \tilde{g}^{*k}\left(\frac{x}{\sqrt{k\mathbb{V}(Y_1)}}\right).$$

The key point is to conduct meticulous computations using an Edgeworth expansion for \tilde{g}^{*k} (and \tilde{g}^{*j}):

$$\tilde{g}^{*k}(y) = \varphi(y) \left(1 + \frac{\gamma H_3(y)}{6\sqrt{k}} + \frac{(\tau - 3)H_4(y)}{24k} + \frac{\gamma^2 H_6(y)}{72k} + o(y^6 k^{-1} \vee k) \right), \quad (29)$$

where H_i are Hermite polynomials, γ the skewness and τ the kurtosis coefficients recalled at the beginning of this Section.

Proof of Equation (8): case ν symmetric. We will first prove Proposition 3 in the easier setting where ν is symmetric. In this case $\tilde{f}_{\eta,\varepsilon,A} = f_{\eta,\varepsilon,A}$, $\bar{\mu} = 0$, $\gamma = 0$, $\mu_3(\eta, \varepsilon) = 0$ and $\mathbb{V}(Y_1) = \frac{\sigma^2}{\lambda}$. To show

(8) we compute a Taylor expansion of (28) as $\eta \rightarrow 0$ and thus $\lambda = \lambda_{\eta, \varepsilon} \rightarrow \infty$. From (29) it holds:

$$\int \frac{g^{*k}(x)g^{*j}(x)}{\varphi_{\eta, \varepsilon}(x)} dx = \frac{\sqrt{\Delta\sigma^2}}{\sqrt{2\pi k j \mathbb{V}(Y_1)^2}} \int e^{\frac{x^2}{2\Delta\sigma^2}} e^{-\frac{x^2}{2k\mathbb{V}(Y_1)}} \left(1 + \frac{(\tau-3)H_4(x/\sqrt{k\mathbb{V}(Y_1)})}{24k} + o(\tau x^4 k^{-3}) \right) \times \\ e^{-\frac{x^2}{2j\mathbb{V}(Y_1)}} \left(1 + \frac{(\tau-3)H_4(x/\sqrt{j\mathbb{V}(Y_1)})}{24j} + o(\tau x^4 j^{-3}) \right) dx.$$

Denote by $V = \frac{\lambda}{k} + \frac{\lambda}{j} - \frac{1}{\Delta}$ and observe that $V > 0$, for all $k, j \in K_\psi$ with $0 < \psi < 1$. Therefore, for η small enough, using that $k, j \in K_\psi$ and that a Gaussian random variable has finite moment of order 4, we have

$$\int \frac{g^{*k}(x)g^{*j}(x)}{\varphi_{\eta, \varepsilon}(x)} dx = \frac{\sqrt{\Delta\sigma^2}}{\sqrt{2\pi k j \mathbb{V}(Y_1)^2}} \int e^{-\frac{x^2}{2}V} \left(1 + \frac{(\tau-3)H_4(x/\sqrt{k\mathbb{V}(Y_1)})}{24k} \right) \\ \left(1 + \frac{(\tau-3)H_4(x/\sqrt{j\mathbb{V}(Y_1)})}{24j} \right) dx + o\left(\frac{\tau}{\lambda^2\Delta^2}\right).$$

Replacing H_4 and $\mathbb{V}(Y_1)$ by their expressions and expanding the product we get

$$\left(1 + \frac{(\tau-3)H_4(x/\sqrt{k\mathbb{V}(Y_1)})}{24k} \right) \left(1 + \frac{(\tau-3)H_4(x/\sqrt{j\mathbb{V}(Y_1)})}{24j} \right) = T_0 + T_2x^2 + T_4x^4 + T_6x^6 + T_8x^8,$$

where

$$T_0 = 1 + \frac{\tau-3}{8} \left(\frac{1}{k} + \frac{1}{j} + \frac{\tau-3}{8kj} \right), \quad T_2 = \frac{(3-\tau)\lambda}{4\sigma^2} \left(\frac{1}{k^2} + \frac{1}{j^2} + \frac{\tau-3}{8} \left(\frac{1}{j^2k} + \frac{1}{k^2j} \right) \right), \\ T_4 = \frac{(\tau-3)\lambda^2}{24\sigma^4} \left(\frac{1}{k^3} + \frac{1}{j^3} + \frac{\tau-3}{8jk} \left(\frac{1}{k^2} + \frac{1}{j^2} \right) + \frac{3}{2} \frac{\tau-3}{j^2k^2} \right), \\ T_6 = -\frac{(\tau-3)^2\lambda^3}{96j^2k^2\sigma^6} \left(\frac{1}{k} + \frac{1}{j} \right) \quad \text{and} \quad T_8 = \frac{(\tau-3)^2\lambda^4}{(24)^2j^3k^3\sigma^8}.$$

Therefore, we have

$$\int e^{-\frac{x^2}{2\sigma^2}V} \left(1 + \frac{(\tau-3)H_4(x/\sqrt{k\mathbb{V}(Y_1)})}{24k} \right) \left(1 + \frac{(\tau-3)H_4(x/\sqrt{j\mathbb{V}(Y_1)})}{24j} \right) dx \\ = \int e^{-\frac{x^2}{2\sigma^2}V} (T_0 + T_2x^2 + T_4x^4 + T_6x^6 + T_8x^8) dx \\ = \sqrt{2\pi\frac{\sigma^2}{V}} \left(T_0 + T_2\frac{\sigma^2}{V} + 3T_4\frac{\sigma^4}{V^2} + 15T_6\frac{\sigma^6}{V^3} + 105T_8\frac{\sigma^8}{V^4} \right).$$

Set $\delta_k = k - \lambda\Delta$ and $\delta_j = j - \lambda\Delta$, we have:

$$\int \frac{g^{*k}(x)g^{*j}(x)}{\varphi_{\eta, \varepsilon}(x)} dx = \frac{\lambda\sqrt{\Delta}}{\sqrt{2\pi\sigma^2kj}} \sqrt{2\pi\frac{\sigma^2}{V}} \left(T_0 + T_2\frac{\sigma^2}{V} + 3T_4\frac{\sigma^4}{V^2} + 15T_6\frac{\sigma^6}{V^3} + 105T_8\frac{\sigma^8}{V^4} \right) + o\left(\frac{\tau}{\lambda^2\Delta^2}\right) \\ = \frac{\lambda\Delta}{\sqrt{\lambda\Delta j + \lambda\Delta k - kj}} \left(T_0 + T_2\frac{\sigma^2}{V} + 3T_4\frac{\sigma^4}{V^2} + 15T_6\frac{\sigma^6}{V^3} + 105T_8\frac{\sigma^8}{V^4} \right) + o\left(\frac{\tau}{\lambda^2\Delta^2}\right) \\ = \frac{1}{\sqrt{1 - \frac{\delta_k\delta_j}{\lambda^2\Delta^2}}} (1 + (\tau-3)I_1) + o\left(\frac{\tau}{\lambda^2\Delta^2}\right), \quad (30)$$

with

$$\begin{aligned}
I_1 &= \frac{1}{8k} + \frac{1}{8j} + \frac{\tau-3}{64kj} - \frac{\lambda\Delta kj}{4(\lambda^2\Delta^2 - \delta_k\delta_j)} \left(\frac{1}{k^2} + \frac{1}{j^2} - \frac{3-\tau}{8jk} \left(\frac{1}{j} + \frac{1}{k} \right) \right) \\
&+ \frac{\lambda^2\Delta^2 kj}{8(\lambda^2\Delta^2 - \delta_k\delta_j)^2} \left(\frac{k^3 + j^3}{k^2j^2} + (\tau-3) \left(\frac{1}{8j^2} + \frac{1}{8k^2} + \frac{3}{2jk} \right) \right) \\
&+ \frac{5}{32} \frac{(3-\tau)\lambda^3\Delta^3 kj}{(\lambda^2\Delta^2 - \delta_k\delta_j)^3} \left(\frac{1}{k} + \frac{1}{j} \right) + \frac{35(\tau-3)kj\Delta^4\lambda^4}{192(\lambda^2\Delta^2 - \delta_k\delta_j)^4}.
\end{aligned}$$

Observe that

$$\frac{1}{8k} + \frac{1}{8j} - \frac{\lambda\Delta kj}{4(\lambda^2\Delta^2 - \delta_k\delta_j)} \left(\frac{1}{k^2} + \frac{1}{j^2} \right) + \frac{\lambda^2\Delta^2}{8(\lambda^2\Delta^2 - \delta_k\delta_j)^2} \left(\frac{k^3 + j^3}{kj} \right) = \frac{\lambda\Delta\delta_j^2 + \lambda\Delta\delta_k^2 + \delta_j^2\delta_k + \delta_j\delta_k^2}{8(\lambda^2\Delta^2 - \delta_j\delta_k)^2},$$

that is

$$\begin{aligned}
1 + (\tau-3)I_1 &= 1 + (\tau-3) \frac{\lambda\Delta\delta_j^2 + \lambda\Delta\delta_k^2 + \delta_j^2\delta_k + \delta_j\delta_k^2}{8(\lambda^2\Delta^2 - \delta_j\delta_k)^2} \\
&+ (\tau-3)^2 \left(\frac{1}{64kj} - \frac{\lambda\Delta(k+j)}{32kj(\lambda^2\Delta^2 - \delta_k\delta_j)} + \frac{\lambda^2\Delta^2 kj}{8(\lambda^2\Delta^2 - \delta_k\delta_j)^2} \left(\frac{1}{8j^2} + \frac{1}{8k^2} + \frac{3}{2jk} \right) \right. \\
&\quad \left. - \frac{5}{32} \frac{\lambda^3\Delta^3(k+j)}{(\lambda^2\Delta^2 - \delta_k\delta_j)^3} + \frac{35}{192} \frac{kj\lambda^4\Delta^4}{(\lambda^2\Delta^2 - \delta_k\delta_j)^4} \right).
\end{aligned}$$

In particular, for $j, k \in [\lambda\Delta(1-\psi), \lambda\Delta(1+\psi)]$ and $\delta_k\delta_j \in [-\psi^2\lambda^2\Delta^2, \psi^2\lambda^2\Delta^2]$, we have

$$1 + (\tau-3)I_1 = 1 + (\tau-3) \frac{\lambda\Delta\delta_j^2 + \lambda\Delta\delta_k^2 + \delta_j^2\delta_k + \delta_j\delta_k^2}{8(\lambda^2\Delta^2 - \delta_j\delta_k)^2} + c \frac{(\tau-3)^2}{\lambda^2\Delta^2} + o\left(\frac{\tau^2}{\lambda^2\Delta^2}\right), \quad (31)$$

for a positive constant $0 < c < (8(1-\psi))^{-2}$. Let us now compute

$$\mathbb{P}(A)^{-2} e^{-2\lambda\Delta} \sum_{j,k \in K_\psi} \frac{(\lambda\Delta)^{j+k}}{j!k!} \frac{1}{\sqrt{1 - \frac{\delta_k\delta_j}{\lambda^2\Delta^2}}} (1 + (\tau-3)I_1).$$

By means of (31) we thus need to control these two quantities

$$I_2 = \mathbb{P}(A)^{-2} e^{-2\lambda\Delta} \sum_{j,k \in K_\psi} \frac{(\lambda\Delta)^{j+k}}{j!k!} \frac{1}{\sqrt{1 - \frac{\delta_k\delta_j}{\lambda^2\Delta^2}}} \left(1 + c \frac{(\tau-3)^2}{\lambda^2\Delta^2} \right) \quad (32)$$

$$I_3 = \mathbb{P}(A)^{-2} e^{-2\lambda\Delta} \sum_{j,k \in K_\psi} \frac{(\lambda\Delta)^{j+k}}{j!k!} \frac{\tau-3}{\sqrt{1 - \frac{\delta_k\delta_j}{\lambda^2\Delta^2}}} \frac{\lambda\Delta\delta_j^2 + \lambda\Delta\delta_k^2 + \delta_j^2\delta_k + \delta_j\delta_k^2}{8(\lambda^2\Delta^2 - \delta_j\delta_k)^2}. \quad (33)$$

Using that for $|x| < 1$, $a > 0$, $(1-x)^{-a} = 1 + ax + \frac{a(a+1)x^2}{2} + O(x^3)$ applied to (32) and (33) with $x = \frac{\delta_k\delta_j}{\lambda^2\Delta^2} \in [-\psi^2, \psi^2]$ and $a = \frac{1}{2}, \frac{5}{2}$, respectively, we are left to control $\sum_{j,k \in K_\psi} \frac{(\lambda\Delta)^{j+k}}{j!k!} \delta_k^l \delta_j^r$, for all

$l, r \in \{1, 2, 3, 4\}$. To that aim set $k_m = \min K_\psi$ and $k_M = \max K_\psi$. We have

$$\begin{aligned} \sum_{k \in K_\psi} e^{-\lambda\Delta} \frac{(\lambda\Delta)^k}{k!} \delta_k &= \lambda\Delta \sum_{k=k_m-1}^{k_M-1} e^{-\lambda\Delta} \frac{(\lambda\Delta)^k}{k!} - \lambda\Delta \sum_{k=k_m}^{k_M} e^{-\lambda\Delta} \frac{(\lambda\Delta)^k}{k!} \\ &= \lambda\Delta e^{-\lambda\Delta} \left(\frac{(\lambda\Delta)^{k_m-1}}{(k_m-1)!} - \frac{(\lambda\Delta)^{k_M}}{k_M!} \right), \end{aligned} \quad (34)$$

$$\sum_{k \in K_\psi} e^{-\lambda\Delta} \frac{(\lambda\Delta)^k}{k!} \delta_k^2 \leq \sum_{k \in \mathbb{N}} e^{-\lambda\Delta} \frac{(\lambda\Delta)^k}{k!} \delta_k^2 = \lambda\Delta, \quad (35)$$

$$\begin{aligned} \sum_{k \in K_\psi} e^{-\lambda\Delta} \frac{(\lambda\Delta)^k}{k!} \delta_k^3 &= \sum_{k=k_m}^{k_M} e^{-\lambda\Delta} \frac{(\lambda\Delta)^k}{k!} (k(k-1)(k-2) + 3k^2 - 2k - 3k^2\lambda\Delta + 3k\lambda^2\Delta^2 - \lambda^3\Delta^3) \\ &= e^{-\lambda\Delta} \left(\lambda^3\Delta^3 \sum_{k=k_m-3}^{k_M-3} \frac{(\lambda\Delta)^k}{k!} + 3\lambda^2\Delta^2(1-\lambda\Delta) \sum_{k=k_m-2}^{k_M-2} \frac{(\lambda\Delta)^k}{k!} \right. \\ &\quad \left. + \lambda\Delta(1-3\lambda\Delta+3\lambda^2\Delta^2) \sum_{k=k_m-1}^{k_M-1} \frac{(\lambda\Delta)^k}{k!} - \lambda^3\Delta^3 \sum_{k=k_m}^{k_M} \frac{(\lambda\Delta)^k}{k!} \right) \\ &= \lambda\Delta e^{-\lambda\Delta} \sum_{k=k_m}^{k_M-3} \frac{(\lambda\Delta)^k}{k!} + e^{-\lambda\Delta} \left(\frac{(\lambda\Delta)^{k_m-3}}{(k_m-3)!} (\lambda\Delta)^3 + \frac{(\lambda\Delta)^{k_m-2}}{(k_m-2)!} (3\lambda^2\Delta^2 - 2\lambda^3\Delta^3) \right. \\ &\quad \left. + \frac{(\lambda\Delta)^{k_m-1}}{(k_m-1)!} (3\lambda^2\Delta^2 - 2\lambda^3\Delta^3) + \frac{(\lambda\Delta)^{k_M-2}}{(k_M-2)!} (3\lambda^2\Delta^2 - 4\lambda^3\Delta^3) \right. \\ &\quad \left. + \frac{(\lambda\Delta)^{k_M-1}}{(k_M-1)!} (3\lambda^2\Delta^2 - 4\lambda^3\Delta^3) - \lambda^3\Delta^3 \frac{(\lambda\Delta)^{k_M}}{(k_M)!} \right) \\ &= \lambda\Delta e^{-\lambda\Delta} \sum_{k=k_m}^{k_M-3} \frac{(\lambda\Delta)^k}{k!} + R_1, \end{aligned} \quad (36)$$

$$\sum_{k \in K_\psi} e^{-\lambda\Delta} \frac{(\lambda\Delta)^k}{k!} \delta_k^4 \leq \sum_{k \in \mathbb{N}} e^{-\lambda\Delta} \frac{(\lambda\Delta)^k}{k!} \delta_k^4 = \lambda\Delta(1+3\lambda\Delta).$$

We will now prove that (34) and R_1 in (36) converge exponentially to 0 as $\eta \rightarrow 0$. For this reason, we will freely ignore polynomial terms in $\lambda\Delta$ whenever this helps with readability. In particular, we can approximate $k_m - 1 \approx \lambda\Delta(1 - \psi)$ and $k_M \approx \lambda\Delta(1 + \psi)$ and $(k_m - 1)!$ (resp. $k_M!$) with $\Gamma(\lambda\Delta(1 - \psi))$ (resp. $\Gamma(\lambda\Delta(1 + \psi))$). Using Stirling approximation,

$$\lambda\Delta e^{-\lambda\Delta} \frac{(\lambda\Delta)^{k_m-1}}{(k_m-1)!} \approx \frac{\Gamma(\lambda\Delta)(\lambda\Delta)^{-\lambda\Delta\psi}}{\Gamma(\lambda\Delta(1-\psi))} = \frac{\lambda\Delta}{\lambda\Delta} \frac{\lambda\Delta-1}{\lambda\Delta} \cdots \frac{\lambda\Delta(1-\psi)+1}{\lambda\Delta} \leq \left(1 - \frac{\psi}{2}\right)^{\lambda\Delta\frac{\psi}{2}} \quad (37)$$

where the last inequality in (37) holds true because at least half of the terms of $\prod_{k=0}^{\lambda\Delta\psi-1} \frac{\lambda\Delta-k}{\lambda\Delta}$ are smaller than $1 - \frac{\psi}{2}$. This shows that $\lambda_{\eta,\varepsilon} \Delta e^{-\lambda_{\eta,\varepsilon} \Delta} \frac{(\lambda_{\eta,\varepsilon} \Delta)^{k_m-1}}{(k_m-1)!} \rightarrow 0$ exponentially in $\lambda_{\eta,\varepsilon}$ as $\eta \rightarrow 0$. Analogously we have

$$\lambda\Delta e^{-\lambda\Delta} \frac{(\lambda\Delta)^{k_M}}{(k_M)!} \approx \frac{\Gamma(\lambda\Delta)(\lambda\Delta)^{\lambda\Delta\psi}}{\Gamma(\lambda\Delta(1+\psi))} = \frac{\lambda\Delta}{\lambda\Delta+1} \frac{\lambda\Delta}{\lambda\Delta+2} \cdots \frac{\lambda\Delta}{\lambda\Delta(1+\psi)} \leq \left(\frac{1}{1+\frac{\psi}{2}}\right)^{\lambda\Delta\frac{\psi}{2}}$$

which goes to 0 exponentially in $\lambda_{\eta,\varepsilon}$ as $\eta \rightarrow 0$. Therefore, we deduce that (34) converges exponentially to 0 as $\eta \rightarrow 0$. The exponential convergence to 0 of R_1 in (36) is proven in a similar way.

Thanks to (35), (36) and the exponential convergence to 0 of (34) and R_1 in (36), when computing (32) and (33) we can neglect the terms in δ_k , the terms in δ_k^2 will count as $\lambda\Delta$ as well as those in δ_k^3 and finally the terms in δ_k^4 as $\lambda\Delta + \lambda^2\Delta^2$.

We can thus deduce that (32) and (33) are of order

$$\begin{aligned} I_2 &= 1 + \frac{3}{8\lambda^2\Delta^2} + \frac{c(\tau-3)^2}{\lambda^2\Delta^2} + O\left(\frac{(\tau-3)^2}{\lambda^4\Delta^4}\right), \\ I_3 &= \frac{\tau-3}{\lambda^4\Delta^4} \left(1 + \frac{e^{-2\lambda\Delta}}{\mathbb{P}(A)^2} \sum_{j,k \in K_\psi} \frac{(\lambda\Delta)^{j+k}}{8j!k!} \left(\frac{5\delta_k\delta_j}{2\lambda^2\Delta^2} + \frac{35\delta_k^2\delta_j^2}{8\lambda^4\Delta^4}\right) (\lambda\Delta\delta_j^2 + \lambda\Delta\delta_k^2 + \delta_j^2\delta_k + \delta_j\delta_k^2)\right) + o\left(\frac{1}{\lambda^4\Delta^4}\right) \\ &\leq 2\frac{\tau-3}{\lambda^4\Delta^4} + o(1). \end{aligned}$$

Gathering (30), (31), and the controls for I_2 and I_3 we derive

$$\chi^2(f_{\eta,\varepsilon,A}, \varphi_{\eta,\varepsilon}) + 1 = 1 + c\frac{(\tau-3)^2}{\lambda^2\Delta^2} + o\left(\frac{\tau}{\lambda^2\Delta^2}\right).$$

For $\psi \leq \frac{7}{8}$ we get $c < 1$. This together with (26) and recalling that $\tau = \lambda\frac{\mu_4(\eta,\varepsilon)}{\sigma^4(\eta,\varepsilon)}$, complete the proof of (8) when ν is symmetric.

Proof of Equation (8): the general case. In the general case, we will also have to keep track of the terms depending on γ . For the remaining ones, the computation in the symmetric case applies.

As above, we begin by applying (29) and we obtain

$$\begin{aligned} \int \frac{g^{*k}(x)g^{*j}(x)}{\tilde{\varphi}_{\eta,\varepsilon}(x)} dx &= \frac{\sqrt{2\pi\Delta\tilde{\sigma}^2}}{2\pi\mathbb{V}(Y_1)\sqrt{kj}} \int e^{-\frac{x^2}{2\Delta\tilde{\sigma}^2}} e^{-\frac{x^2}{2k\mathbb{V}(Y_1)}} \left(1 + \frac{\gamma H_3(x/\sqrt{k\mathbb{V}(Y_1)})}{6\sqrt{k}} + \frac{(\tau-3)H_4(x/\sqrt{k\mathbb{V}(Y_1)})}{24k}\right. \\ &\quad \left.+ o(\tau x^4 k^{-3}) + o(\gamma x^3 k^{-2})\right) \times \\ &\quad e^{-\frac{x^2}{2j\mathbb{V}(Y_1)}} \left(1 + \frac{\gamma H_3(x/\sqrt{j\mathbb{V}(Y_1)})}{6\sqrt{j}} + \frac{(\tau-3)H_4(x/\sqrt{j\mathbb{V}(Y_1)})}{24j} + o(\tau x^4 j^{-3}) + o(\gamma x^3 j^{-2})\right) dx. \end{aligned}$$

We therefore only need to focus on

$$\frac{\lambda\sqrt{\Delta}}{\sqrt{2\pi\tilde{\sigma}^2kj}} \int e^{-\frac{x^2}{2\tilde{\sigma}^2}\left(\frac{\lambda}{k} + \frac{\lambda}{j} - \frac{1}{\Delta}\right)} \left(1 + \frac{\gamma H_3(x/\sqrt{k\mathbb{V}(Y_1)})}{6\sqrt{k}}\right) \left(1 + \frac{\gamma H_3(x/\sqrt{j\mathbb{V}(Y_1)})}{6\sqrt{j}}\right) dx + o\left(\frac{\gamma}{\lambda\Delta}\right). \quad (38)$$

Expanding the product in (38) and replacing H_3 by its expression we get

$$\begin{aligned} &\left(1 + \frac{\gamma H_3(x/\sqrt{k\mathbb{V}(Y_1)})}{6\sqrt{k}}\right) \left(1 + \frac{\gamma H_3(x/\sqrt{j\mathbb{V}(Y_1)})}{6\sqrt{j}}\right) \\ &= 1 - \frac{\gamma}{2\sqrt{\mathbb{V}(Y_1)}} \left(\frac{1}{k} + \frac{1}{j}\right)x + \frac{\gamma^2 x^2}{4\mathbb{V}(Y_1)kj} + \frac{\gamma(k^2 + j^2)x^3}{6(\mathbb{V}(Y_1))^{3/2}k^2j^2} - \frac{\gamma^2 x^4}{12\mathbb{V}(Y_1)^2kj} \left(\frac{1}{k} + \frac{1}{j}\right) + \frac{\gamma^2 x^6}{36\mathbb{V}(Y_1)^3k^2j^2} \\ &:= 1 + T_1x + T_2x^2 + T_3x^3 + T_4x^4 + T_6x^6. \end{aligned}$$

Therefore, denoting by $V = \frac{\lambda}{k} + \frac{\lambda}{j} - \frac{1}{\Delta}$, we obtain

$$\begin{aligned} &\int e^{-\frac{x^2}{2\tilde{\sigma}^2}V} \left(1 + \frac{\gamma H_3(x/\sqrt{k\mathbb{V}(Y_1)})}{6\sqrt{k}}\right) \left(1 + \frac{\gamma H_3(x/\sqrt{j\mathbb{V}(Y_1)})}{6\sqrt{j}}\right) dx \\ &= \sqrt{\frac{2\pi\tilde{\sigma}^2}{V}} \left(1 + T_2\frac{\tilde{\sigma}^2}{V} + 3T_4\frac{\tilde{\sigma}^4}{V^2} + 15T_6\frac{\tilde{\sigma}^6}{V^3}\right). \end{aligned}$$

Replacing $\mathbb{V}(Y_1) = \frac{\sigma^2 - \lambda \bar{\mu}^2}{\lambda} = \frac{\tilde{\sigma}^2}{\lambda}$, it follows that (38) is of order

$$\frac{\lambda \sqrt{\Delta}}{\sqrt{kjV}} \left(1 + \frac{\lambda \gamma^2}{4Vkj} - \frac{\lambda^2 \gamma^2 (k+j)}{4V^2 k^2 j^2} + \frac{5\lambda^3 \gamma^2}{12V^3 k^2 j^2} \right) + o\left(\frac{\gamma}{\lambda \Delta}\right).$$

Set $\delta_k = k - \lambda \Delta$, $\delta_j = j - \lambda \Delta$ and notice that $kjV = \lambda^2 \Delta \left(1 - \frac{\delta_k \delta_j}{\lambda^2 \Delta^2}\right)$. The following equalities hold:

$$\begin{aligned} \frac{\lambda \sqrt{\Delta}}{\sqrt{kjV}} &= \frac{1}{\sqrt{1 - \frac{\delta_k \delta_j}{\lambda^2 \Delta^2}}}, & \frac{\lambda \sqrt{\Delta}}{\sqrt{kjV}} \frac{\lambda \gamma^2}{4Vkj} &= \frac{\gamma^2}{4\lambda \Delta \left(1 - \frac{\delta_k \delta_j}{\lambda^2 \Delta^2}\right)^{3/2}} \\ \frac{\lambda \sqrt{\Delta}}{\sqrt{kjV}} \frac{\lambda^2 \gamma^2 (k+j)}{4V^2 k^2 j^2} &= \frac{\gamma^2 (\delta_k + \delta_j + 2\lambda \Delta)}{4\lambda^2 \Delta^2 \left(1 - \frac{\delta_k \delta_j}{\lambda^2 \Delta^2}\right)^{5/2}} \\ \frac{\lambda \sqrt{\Delta}}{\sqrt{kjV}} \frac{5\lambda^3 \gamma^2}{12V^3 k^2 j^2} &= \frac{5\gamma^2 (\delta_k \delta_j + \lambda \Delta (\delta_k + \delta_j) + \lambda^2 \Delta^2)}{12\lambda^3 \Delta^3 \left(1 - \frac{\delta_k \delta_j}{\lambda^2 \Delta^2}\right)^{7/2}}. \end{aligned}$$

Consequently, (38) can be expanded as follows:

$$\begin{aligned} \int \frac{g^{*k}(x) g^{*j}(x)}{\tilde{\varphi}_{\eta, \varepsilon}(x)} dx &= 1 + \frac{\delta_k \delta_j}{2\lambda^2 \Delta^2} + \frac{3}{8} \frac{\delta_k^2 \delta_j^2}{\lambda^4 \Delta^4} + \frac{\gamma^2}{\lambda \Delta} \left(\frac{1}{4} + \frac{3}{8} \frac{\delta_k \delta_j}{\lambda^2 \Delta^2} + \frac{15}{32} \frac{\delta_k^2 \delta_j^2}{\lambda^4 \Delta^4} \right) \\ &+ \gamma^2 \frac{\delta_k + \delta_j + 2\lambda \Delta}{4\lambda^2 \Delta^2} \left(1 + \frac{5}{2} \frac{\delta_k \delta_j}{\lambda^2 \Delta^2} + \frac{35}{8} \frac{\delta_k^2 \delta_j^2}{\lambda^4 \Delta^4} \right) \\ &+ \gamma^2 \frac{5}{12} \frac{\delta_k \delta_j + \lambda \Delta (\delta_k + \delta_j) + \lambda^2 \Delta^2}{\Delta^3 \lambda^3} \left(1 + \frac{7}{2} \frac{\delta_k \delta_j}{\lambda^2 \Delta^2} + \frac{63}{8} \frac{\delta_k^2 \delta_j^2}{\lambda^4 \Delta^4} \right) + O\left(\frac{\delta_k^3 \delta_j^3}{\lambda^6 \Delta^6}\right) + o\left(\frac{\gamma}{\lambda \Delta}\right). \end{aligned} \quad (39)$$

Thanks to (28), in order to compute $\chi^2(\tilde{f}_{\eta, \varepsilon, A}, \tilde{\varphi}_{\eta, \varepsilon})$ we are only left to take the sum over k and j in (39) weighted with $\frac{(\lambda \Delta)^{j+k}}{j! k!}$. Thanks to (35), (36) and the exponential convergence to 0 of (34) and R_1 in (36), when passing to the weighted sum in (39) we can neglect the terms in δ_k and the terms in δ_k^2 will count as $\lambda \Delta$ as well as those in δ_k^3 .

Finally, collecting all the pieces together and using the computations already done for the symmetric case, we conclude that

$$\chi^2(\tilde{f}_{\eta, \varepsilon, A}, \tilde{\varphi}_{\eta, \varepsilon}) + 1 \leq 1 + \frac{7}{6} \frac{\gamma^2}{\lambda \Delta} + c \frac{(\tau - 3)^2}{\lambda^2 \Delta^2} + o\left(\frac{\gamma}{\lambda \Delta}\right) + o\left(\frac{\tau}{\lambda^2 \Delta^2}\right), \quad \text{as } \eta \rightarrow 0,$$

for $c < 1$. Recalling that

$$\begin{aligned} \bar{\mu} &= \frac{1}{\lambda} \int_{\eta < |x| < \varepsilon} x \nu(dx), & \tilde{\sigma}^2 &= \sigma^2(\eta, \varepsilon) - \bar{\mu}^2 \lambda, \\ \gamma &= \frac{\sqrt{\lambda} (\mu_3(\eta, \varepsilon) - 3\bar{\mu} \sigma^2 + 2\lambda \bar{\mu}^3)}{\tilde{\sigma}^3}, & \tau &= \lambda \frac{\mu_4(\eta, \varepsilon)}{\sigma^4(\eta, \varepsilon)}, \end{aligned}$$

and observing that thanks to (26) the terms $\frac{\gamma}{\lambda \Delta}$ and $\bar{\mu}^2 \lambda$ vanish as $\eta \rightarrow 0$, we finally established

$$\chi^2(\tilde{f}_{\eta, \varepsilon, A}, \tilde{\varphi}_{\eta, \varepsilon}) \leq \frac{7}{6} \frac{\mu_3^2(\eta, \varepsilon)}{\Delta \sigma^6(\eta, \varepsilon)} + \frac{\mu_4^2(\eta, \varepsilon)}{\Delta^2 (\sigma^2(\eta, \varepsilon))^4} + o(1), \quad \text{as } \eta \rightarrow 0.$$

The proof of Proposition 3 is now complete.

4.2 Proof of Theorem 5

4.2.1 Preliminary work: Four statistical tests

In this Section we will work with $X \sim (b, \Sigma^2, \nu_\varepsilon)$; in particular this means that the increments $X_{i\Delta} - X_{(i-1)\Delta}$ are i.i.d. realizations of

$$X_\Delta = \Delta b + \Sigma W_\Delta + M_\Delta(\varepsilon), \quad \text{where } W_\Delta \sim \mathcal{N}(0, \Delta).$$

For any $n \in \mathbb{N}$, set $\tilde{n} = \lfloor n/2 \rfloor$ and define

$$Z_i := |(X_{2i\Delta} - X_{(2i-1)\Delta}) - (X_{(2i-1)\Delta} - X_{(2i-2)\Delta})|, \quad i = 1, \dots, \tilde{n} = \lfloor n/2 \rfloor$$

$$S_n = \frac{1}{\tilde{n}} \sum_{i=1}^{\tilde{n}-2\log(n)} Z_{(i)}, \quad \text{and} \quad Z_{(\tilde{n})} = \max\{Z_i, 1 \leq i \leq \tilde{n}\},$$

where for any sequence a_\cdot , the sequence $a_{(\cdot)}$ is a reordering of a in increasing order.

For any $0 < u \leq \varepsilon$, we write X as $X = \bar{X}(u) + M(u, \varepsilon)$, where

$$\bar{X}(u)_t = bt + \Sigma W_t + M_t(u)$$

is a Lévy process whose jumps are smaller (or equal) than u and $M(u, \varepsilon) = M_t(\varepsilon) - M_t(u)$ is a pure jumps Lévy process with jumps of size between u and ε . We write $N(u)$ for the number of jumps larger than u , that is, for any $t > 0$, $N_t(u)$ is a Poisson random variable with mean $t\lambda_{u,\varepsilon}$.

Furthermore, in order to present the test needed to prove Theorem 5, we introduce the following notations:

$$\bar{X}_{\Delta,n} := \frac{1}{n - \lfloor n/2 \rfloor} \sum_{i=\lfloor n/2 \rfloor+1}^n (X_{i\Delta} - X_{(i-1)\Delta})$$

$$\bar{Y}_{n,3} := \frac{1}{\lfloor n/2 \rfloor} \sum_{i=1}^{\lfloor n/2 \rfloor} ((X_{i\Delta} - X_{(i-1)\Delta}) - \bar{X}_{\Delta,n})^3$$

$$\bar{Y}_{n,2} := \frac{1}{\lfloor n/4 \rfloor} \sum_{i=1}^{\lfloor n/4 \rfloor} Z_i^2, \quad \bar{Y}'_{n,2} := \frac{1}{\lfloor n/4 \rfloor} \sum_{i=\lfloor n/4 \rfloor+1}^{\lfloor n/2 \rfloor} Z_i^2$$

$$\bar{Y}_{n,4} := \frac{1}{\lfloor n/2 \rfloor} \sum_{i=1}^{\lfloor n/2 \rfloor} Z_i^4, \quad \bar{Y}_{n,6} := \frac{1}{\lfloor n/2 \rfloor} \sum_{i=1}^{\lfloor n/2 \rfloor} Z_i^6$$

$$T_n^{(3)} := \frac{1}{1 - (n - \lfloor n/2 \rfloor)^{-2}} \bar{Y}_{n,3}, \quad T_n^{(4)} := 4^{-1} (\bar{Y}_{n,4} - 3\bar{Y}_{n,2}\bar{Y}'_{n,2}).$$

By definition, $\bar{X}_{\Delta,n}$ is the empirical version of $\mathbb{E}[X_\Delta]$ computed on the second half of the sample only and $\bar{Y}_{n,3}$ (resp. $\bar{Y}_{n,6}$) is an estimator of $\mathbb{E}[(X_\Delta - \Delta b)^3]$ (resp. of $8\mathbb{E}[(X_\Delta - \Delta b)^6]$) computed on the first half of the sample. Moreover, since $\mathbb{E}[(X_\Delta - b\Delta)^3] = \Delta\mu_3(\varepsilon)$, using Corollary 5 joined with the independence of X_Δ and $\bar{X}_{\Delta,n}$, we have that $T_n^{(3)}$ is an unbiased estimator of $\Delta\mu_3(\varepsilon)$. $\bar{Y}_{n,4}$ is instead an estimator of $4\mathbb{E}[(X_\Delta - \Delta b)^4]$ while $\bar{Y}_{n,2}$ and $\bar{Y}'_{n,2}$ are two independent estimators of $2\mathbb{E}[(X_\Delta - \Delta b)^2]$ (see also Corollary 5). Using that $\mathbb{E}[Z_1^4] - 3(\mathbb{E}[Z_1^2])^2 = 4\Delta\mu_4(\varepsilon)$, it is easy to prove that $T_n^{(4)}$ is an unbiased estimator of $\Delta\mu_4(\varepsilon)$ (see, e.g. [7]).

Finally, let $C > 0$ be the absolute constant introduced in Lemma 5 below and consider the following events:

- If $\nu_\varepsilon = 0$, set

$$\xi_n := \{\forall i, Z_i \leq 4\sqrt{\Delta\Sigma^2 \log(n)}\}, \quad (40)$$

$$\xi'_n := \left\{ \frac{\sqrt{\Delta\Sigma^2}}{\sqrt{\pi}} \leq S_n \right\}, \quad (41)$$

$$\xi''_n := \{C(\sqrt{\Delta\Sigma^2})^6 \geq \bar{Y}_{n,6}\}. \quad (42)$$

- If $\nu_\varepsilon \neq 0$, set

$$\begin{aligned} \xi_n &:= \{1 \leq N_{n\Delta}(u^+) \leq 2 \log(n)\} \cap \{\forall i \leq n, N_{i\Delta}(u^+) - N_{(i-2)\Delta}(u^+) \leq 1\} \\ &\quad \cap \{\forall i \text{ s.t. } N_{i\Delta}(u^+) - N_{(i-2)\Delta}(u^+) \neq 0, \\ &\quad |\bar{X}_{i\Delta}(u^*) - \bar{X}_{(i-1)\Delta}(u^*) - (\bar{X}_{(i-1)\Delta}(u^*) - \bar{X}_{(i-2)\Delta}(u^*))| \leq 2\sqrt{(\Sigma^2 + \sigma^2(u^*))\Delta} \log(n)\}, \quad (43) \\ \xi'_n &:= \{S_n \leq 2\sqrt{2\Delta(\Sigma^2 + \sigma^2(u^*))}\}, \quad (44) \\ \xi''_n &:= \left\{ \bar{Y}_{n,6} \geq \frac{\Delta\mu_6(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^3}{2} \right\}. \end{aligned}$$

Lemma 3. *There exists a universal sequence $\alpha_n \rightarrow 0$ such that $\mathbb{P}(\xi_n) \geq 1 - \alpha_n$.*

Lemma 4. *There exists a universal sequence $\alpha_n \rightarrow 0$ such that $\mathbb{P}(\xi'_n) \geq 1 - \alpha_n$.*

Lemma 5. *There exist a universal sequence $\alpha_n \rightarrow 0$ and a universal constant $C > 0$ such that the following holds.*

Whenever $\nu_\varepsilon = 0$, with probability larger than $1 - \alpha_n$ we have

$$C(\sqrt{\Delta\Sigma^2})^6 \geq \bar{Y}_{n,6}.$$

In any case, with probability larger than $1 - \alpha_n$ and conditional on $N_{n\Delta}(u^) = 0$, it holds*

$$\bar{Y}_{n,6} \geq \frac{\Delta\mu_6(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^3}{2}.$$

Observe that Lemmas 3, 4 and Lemma 5 joined with Equation (45), imply the existence of two absolute sequences $\alpha_n \rightarrow 0$ and $\beta_n \rightarrow 0$ such that

$$\begin{aligned} \mathbb{P}(\xi_n \cap \xi'_n) &\geq 1 - \alpha_n, \\ \mathbb{P}(\xi_n \cap \xi'_n \cap \xi''_n) &\geq 1 - \beta_n. \end{aligned} \quad (45)$$

We are finally ready to introduce the four tests we use to establish Theorem 5:

$$\begin{aligned} \Phi_n^{(\max)} &= \mathbf{1}\{Z_{(\bar{n})} \geq \log(n)^{3/2} S_n\}, & \Phi_{n,c}^{(6)} &= \mathbf{1}\{\bar{Y}_{n,6} \geq cS_n^6\}, \\ \Phi_{n,c,\alpha}^{(3)} &= \mathbf{1}\left\{|T_n^{(3)}| \geq \frac{c}{\sqrt{\alpha}} \sqrt{\frac{S_n^6}{n}}\right\}, & \Phi_{n,c,\alpha}^{(4)} &= \mathbf{1}\left\{|T_n^{(4)}| \geq \frac{c}{\sqrt{\alpha}} S_n^4 \sqrt{\frac{1}{n}}\right\}. \end{aligned}$$

Their properties are investigated in Propositions 4, 5, 6 and 7 below.

Proposition 4. *Under H_0 , for any $n > e^{4\sqrt{\pi}}$, it holds that $\xi_n \cap \xi'_n \subset \{\Phi_n^{(\max)} = 0\}$. Moreover, for any $n > e^2$, it holds $\xi_n \cap \xi'_n \cap \{N_{n\Delta}(u^*) \geq 1\} \subset \{N_{n\Delta}(u^*) \geq 1\} \cap \{\Phi_n^{(\max)} = 1\}$.*

Proposition 5. *There exist $c > 0$ a universal constant and C_c depending only on c such that the following holds, for n larger than an universal constant. Under H_0 , it holds that $\xi'_n \cap \xi''_n \subset \{\Phi_{n,c}^{(6)} = 0\}$. Moreover if $\Delta\mu_6(u^*) \geq C_c \Delta^3 (\Sigma^2 + \sigma^2(u^*))^3$, then $\xi'_n \cap \xi''_n \cap \{N_{n\Delta}(u^*) = 0\} \subset \{N_{n\Delta}(u^*) = 0\} \cap \{\Phi_{n,c}^{(6)} = 1\}$.*

Proposition 6. Let $\alpha > 2\log(n)^{-1}$. Let $c > 0$ and $c' > 0$ be large enough absolute constant and let $C_{c,c'} > 0$ be a large enough absolute constant depending only on c and c' . Then, the following holds.

Under H_0 , $\Phi_{n,c,\alpha}^{(3)} = 0$ with probability larger than $1 - \alpha - \mathbb{P}(\xi'_n < c)$.

Under the hypothesis $H_{1,\rho_n}^{(3)} : \mu_3(u^*) > \rho_n^{(3)}$, $\lim_{\eta \rightarrow 0} \lambda_{\eta,\varepsilon} = +\infty$, and conditional to the event $N_{n\Delta}(u^*) = 0$, if

$$u^* > u^+, \quad \Delta\mu_6(u^*) \leq c' \Delta^3 (\Sigma^2 + \sigma^2(u^*))^3, \quad \rho_n^{(3)} \geq \frac{C_{c,c'}}{\sqrt{\alpha}} \left\{ \frac{\sqrt{\Delta(\Sigma^2 + \sigma^2(u^*))^3}}{\sqrt{n}} \right\}, \quad (46)$$

it holds that $\Phi_{n,c,\alpha}^{(3)} = 1$ with probability larger than $1 - \alpha - \mathbb{P}(\xi'_n < c)$.

Proposition 7. Let $\alpha > 2\log(n)^{-1}$. Let $c > 0$ and $c' > 0$ be large enough absolute constant and let $C_{c,c'} > 0$ be a large enough absolute constant depending only on c and c' . Then, the following holds.

Under H_0 , it holds that $\Phi_{n,c,\alpha}^{(4)} = 0$ with probability larger than $1 - \alpha - \mathbb{P}(\xi'_n < c)$.

Under the hypothesis $H_{1,\rho_n}^{(4)} : \mu_4(u^*) > \rho_n^{(4)}$, $\lim_{\eta \rightarrow 0} \lambda_{\eta,\varepsilon} = +\infty$, and conditional to the event $N_{n\Delta}(u^*) = 0$, if

$$u^* > u^+, \quad \rho_n^{(4)} \geq \frac{C_c}{\sqrt{\alpha}} \left\{ \frac{\Delta(\Sigma^2 + \sigma^2(u^*))^2}{\sqrt{n}} \right\},$$

it holds that $\Phi_{n,c,\alpha}^{(4)} = 1$ with probability larger than $1 - \alpha - \mathbb{P}(\xi'_n < c)$.

4.2.2 Completion on the proof of Theorem 5

Let $(\tilde{b}, \tilde{\Sigma}, \tilde{\nu}_\varepsilon)$ be a Lévy triplet where $\tilde{\nu}_\varepsilon$ is an infinite Lévy measure absolutely continuous with respect to the Lebesgue measure with support in $[-\varepsilon, \varepsilon]$. Assume that we want to test

$$H_0 : \nu_\varepsilon = 0, \quad \text{against} \quad H_1 : (b, \Sigma, \nu_\varepsilon) = (\tilde{b}, \tilde{\Sigma}, \tilde{\nu}_\varepsilon).$$

We write $\tilde{\mu}_\cdot, \tilde{\lambda}_\cdot$ and \tilde{u}^* for all the quantities related to $(\tilde{b}, \tilde{\Sigma}, \tilde{\nu}_\varepsilon)$.

We can choose $c^{(3)}, c^{(4)}, c^{(6)} > 0$ large enough universal constants and $C^{(3)}, C^{(4)}, C^{(6)} > 0$ large enough depending only on $c^{(3)}, c^{(4)}, c^{(6)}$, and an absolute sequence α_n that converges to 0 such that Propositions 4, 5, 6 and 7 hold. Set $\alpha = \left\{ \left(\frac{C^{(3)}}{\tilde{\mu}_3(\tilde{u}^*)} \frac{\sqrt{\Delta(\Sigma^2 + \sigma^2(\tilde{u}^*))^3}}{\sqrt{n}} \right)^2 \wedge \left(\frac{C^{(4)}}{\tilde{\mu}_4(u^*)} \frac{\Delta(\Sigma^2 + \sigma^2(\tilde{u}^*))^2}{\sqrt{n}} \right)^2 \right\} \vee \alpha_n$. Write $i = 3$ if $\left(\frac{C^{(3)}}{\tilde{\mu}_3(\tilde{u}^*)} \frac{\sqrt{\Delta(\Sigma^2 + \sigma^2(\tilde{u}^*))^3}}{\sqrt{n}} \right)^2 \leq \left(\frac{C^{(4)}}{\tilde{\mu}_4(u^*)} \frac{\Delta(\Sigma^2 + \sigma^2(\tilde{u}^*))^2}{\sqrt{n}} \right)^2$ and $i = 4$ otherwise. In the remaining of the proof α_n denotes a vanishing sequence whose value may change from line to line.

Case 1 : $1 - \exp(-\tilde{\lambda}_{u^*,\varepsilon} n \Delta) \geq 1 - \alpha$. In this case, consider the test $\Phi_n = \Phi_n^{(\max)}$. If $X \sim (b, \Sigma^2, \nu_\varepsilon)$ is in H_0 (i.e. $\nu_\varepsilon = 0$), an application of Proposition 4 and Lemmas 3 and 4 yields $\mathbb{P}(\Phi_n = 0) \geq 1 - \alpha_n$. If, instead, X is such that $(b, \Sigma, \nu_\varepsilon) = (\tilde{b}, \tilde{\Sigma}, \tilde{\nu}_\varepsilon)$, by means of Proposition 4 and Lemmas 3, 4 we get

$$\mathbb{P}(\Phi_n = 1) \geq \mathbb{P}(\{N_{n\Delta}(\tilde{u}^*) \neq 0\} \cap \xi_n \cap \xi'_n) \geq 1 - \exp(-\tilde{\lambda}_{\tilde{u}^*,\varepsilon} n \Delta) - \alpha_n.$$

So by Lemma 1 it follows that the total variation between the observations of n increments of X at the sampling rate Δ and the closest Gaussian random variable is larger than

$$1 - \exp(-\tilde{\lambda}_{\tilde{u}^*,\varepsilon} n \Delta) - \alpha_n.$$

Case 2 : $1 - \exp(-\tilde{\lambda}_{\tilde{u}^*, \varepsilon} n \Delta) \leq 1 - \alpha$. In this case consider the test

$$\Phi_{n, c^{(i)}, c^{(6)}, \alpha} = \Phi_n^{(\max)} \vee \Phi_{n, c^{(i)}, \alpha}^{(i)} \vee \Phi_{n, c^{(6)}}^{(6)}.$$

If X is in H_0 (i.e. $\nu_\varepsilon = 0$), by Propositions 4, 5, 6 and 7 we have that

$$\mathbb{P}(\Phi_{n, c^{(i)}, c^{(6)}, \alpha} = 0) \geq 1 - \alpha - \alpha_n.$$

If X is such that $(b, \Sigma, \nu_\varepsilon) = (\tilde{b}, \tilde{\Sigma}, \tilde{\nu}_\varepsilon)$, we distinguish two cases.

- If $\Delta \mu_6(u^*) \geq C^{(6)}(\Delta(\Sigma^2 + \sigma^2(u^*)))^3$: Propositions 4, 5 yield

$$\mathbb{P}(\Phi_{n, c_\alpha^{(i)}, c^{(6)}, \alpha} = 1) \geq \mathbb{P}(\{N_{n\Delta}(\tilde{u}^*) \neq 0\} \cap \xi_n \cap \xi'_n) + \mathbb{P}(\{N_{n\Delta}(\tilde{u}^*) = 0\})(1 - \alpha_n).$$

- If $\Delta \mu_6(u^*) < C^{(6)}(\Delta(\Sigma^2 + \sigma^2(u^*)))^3$: Propositions 4, 6, 7 joined with $\{u^* > u^+\}$ yield

$$\mathbb{P}(\Phi_{n, c_\alpha^{(i)}, c^{(6)}, \alpha} = 1) \geq \mathbb{P}(\{N_{n\Delta}(\tilde{u}^*) \neq 0\} \cap \xi_n \cap \xi'_n) + \mathbb{P}(\{N_{n\Delta}(\tilde{u}^*) = 0\})(1 - \alpha - \alpha_n).$$

In both cases we conclude that,

$$\begin{aligned} \mathbb{P}(\Phi_{n, c^{(i)}, c^{(6)}, \alpha} = 1) &\geq \mathbb{P}(\{N_{n\Delta}(\tilde{u}^*) \neq 0\}) + \mathbb{P}(\{N_{n\Delta}(\tilde{u}^*) = 0\})(1 - \alpha) - \alpha_n \\ &\geq 1 - \alpha \exp(-\tilde{\lambda}_{\tilde{u}^*, \varepsilon} n \Delta) - \alpha_n. \end{aligned}$$

By Lemma 1 we deduce that the total variation distance between the observations of n increments of X at the sampling rate Δ and the closest Gaussian random variable is larger than $1 - 2\alpha - \alpha_n$.

4.2.3 Proof of Corollary 2

Fix $0 < u \leq \varepsilon$, using (5) it holds $M(\varepsilon) = M(u) + M(u, \varepsilon)$ where $M(u, \varepsilon)$ is a compound Poisson process with intensity $\lambda_{u, \varepsilon}$ independent of $M(u)$. Decomposing on the values of the Poisson process induced by $M(u, \varepsilon)$ at time $n\Delta$ and using that the total variation distance is bounded by 1, we have

$$\begin{aligned} &\|(\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(\varepsilon))^{\otimes n} - \mathcal{N}(b\Delta, \Delta(\Sigma^2 + \sigma^2(u)))^{\otimes n}\|_{TV} \\ &\leq \exp(-\lambda_{u, \varepsilon} n \Delta) \|(\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(u))^{\otimes n} - \mathcal{N}(b\Delta, \Delta(\Sigma^2 + \sigma^2(u)))^{\otimes n}\|_{TV} + [1 - \exp(-\lambda_{u, \varepsilon} \Delta n)] \\ &\leq 3 \exp(-\lambda_{u, \varepsilon} n \Delta) \sqrt{\left(1 + \frac{\mu_4^2(u)}{\Delta^2(\Sigma^2 + \sigma^2(u))^4} + \frac{7}{6} \frac{\mu_3^2(u)}{\Delta(\Sigma^2 + \sigma^2(u))^3}\right)^n - 1} + [1 - \exp(-\lambda_{u, \varepsilon} \Delta n)], \quad (47) \end{aligned}$$

applying Theorem 3 to get the last inequality. Finally, the result is obtained using that

$$\begin{aligned} \min_{B \in \mathbb{R}, S^2 \geq 0} \|(\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(\varepsilon))^{\otimes n} - \mathcal{N}(B\Delta, \Delta S^2)^{\otimes n}\|_{TV} &\leq \\ \|(\mathcal{N}(b\Delta, \Delta\Sigma^2) * M_\Delta(\varepsilon))^{\otimes n} - \mathcal{N}(b\Delta, \Delta(\Sigma^2 + \sigma^2(u)))^{\otimes n}\|_{TV} & \end{aligned}$$

and taking the infimum in $u \in (0, \varepsilon]$ in (47).

Acknowledgements. The work of A. Carpentier is partially supported by the Deutsche Forschungsgemeinschaft (DFG) Emmy Noether grant MuSyAD (CA 1488/1-1), by the DFG - 314838170, GRK 2297 MathCoRe, by the DFG GRK 2433 DAEDALUS, by the DFG CRC 1294 'Data Assimilation', Project A03, and by the UFA-DFH through the French-German Doktorandenkolleg CDFA 01-18.

The work of E. Mariucci has been partially funded by the Federal Ministry for Education and Research through the Sponsorship provided by the Alexander von Humboldt Foundation, by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 314838170, GRK 2297 MathCoRe, and by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1294 'Data Assimilation'.

References

- [1] S. Asmussen and J. Rosiński. “Approximations of small jumps of Lévy processes with a view towards simulation”. In: *Journal of Applied Probability* 38.2 (2001), pp. 482–493.
- [2] O. E. Barndorff-Nielsen, T. Mikosch, and S. I. Resnick. *Lévy processes: theory and applications*. Springer Science & Business Media, 2012.
- [3] D. Belomestny, F. Comte, V. Genon-Catalot, H. Masuda, and M. Reiß. *Lévy Matters IV - Estimation for discretely observed Lévy processes*. Vol. 2128. Springer International Publishing, 2015.
- [4] J. Bertoin. *Lévy processes*. Vol. 121. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1996, pp. x+265.
- [5] S. Cohen and J. Rosinski. “Gaussian approximation of multivariate Lévy processes with applications to simulation of tempered stable processes”. In: *Bernoulli* 13.1 (2007), pp. 195–210.
- [6] R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004, xvi+535.
- [7] Y. Dodge and V. Rousson. “The complications of the fourth central moment”. In: *The American Statistician* 53.3 (1999), pp. 267–269.
- [8] P. Étoré and E. Mariucci. “ L_1 -distance for additive processes with time-homogeneous Lévy measures”. In: *Electronic Communications in Probability* 19 (2014).
- [9] B. Gnedenko and A. Kolmogorov. *Limit distributions for sums of independent random variables*. Cambridge, Massachusetts: Addison-Wesley, 1954.
- [10] Y. Ingster and I. A. Suslina. *Nonparametric goodness-of-fit testing under Gaussian models*. Vol. 169. Springer Science & Business Media, 2012.
- [11] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*. Second. Vol. 288. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 2003, pp. xx+661.
- [12] Y. A. Kutoyants. *Statistical inference for spatial Poisson processes*. Vol. 134. Springer Science & Business Media, 2012.
- [13] F. Liese. “Estimates of Hellinger integrals of infinitely divisible distributions”. In: *Kybernetika* 23.3 (1987), pp. 227–238.
- [14] E. Mariucci and M. Reiß. “Wasserstein and total variation distance between marginals of Lévy processes”. In: *Electronic Journal of Statistics* 12.2 (2018), pp. 2482–2514.
- [15] J. Mémin and A. Shiryaev. “Distance de Hellinger-Kakutani des lois correspondant à deux processus à accroissements indépendants”. In: *Probability Theory and Related Fields* 70.1 (1985), pp. 67–89.
- [16] M. Reiß. “Testing the characteristics of a Lévy process”. In: *Stochastic Processes and their Applications* 123.7 (2013), pp. 2808–2828.
- [17] K.-i. Sato. *Lévy processes and infinitely divisible distributions*. Vol. 68. Cambridge Studies in Advanced Mathematics. Translated from the 1990 Japanese original, Revised by the author. Cambridge: Cambridge University Press, 1999, pp. xii+486.
- [18] A. B. Tsybakov. *Introduction to nonparametric estimation. Revised and extended from the 2004 French original. Translated by Vladimir Zaiats*. 2009.

A Technical results

Since it will be used several times in the rest of the paper, we write BCI for the Bienaymé-Chebyshev inequality which states that, if Z is a random variable with finite variance, then with probability larger than $1-\alpha$, it holds

$$\mathbb{E}[Z] - \sqrt{\mathbb{V}(Z)/\alpha} \leq Z \leq \mathbb{E}[Z] + \sqrt{\mathbb{V}(Z)/\alpha}.$$

Finally, in several occasions we will use that $\sigma(u^+) \leq \sigma(u^*)$.

A.1 Proofs of the Propositions involved in the proof of Theorem 5

A.1.1 Proof of Proposition 4

Under H_0 : By means of Equations (40) and (41) we have that

$$\max_i Z_i \leq 4\sqrt{\Delta\Sigma^2 \log(n)}, \quad \text{on } \xi_n \quad \text{and} \quad S_n \geq \frac{\sqrt{\Delta\Sigma^2}}{\pi}, \quad \text{on } \xi'_n.$$

Therefore, for n strictly larger than $e^{4\sqrt{\pi}}$, on the event $\xi_n \cap \xi'_n$ we have that $Z_{(\tilde{n})} < \log(n)^{3/2} S_n$, and thus $\Phi_n^{(\max)} = 0$ as desired.

If a jump larger than u^* occurs : If $u^* = \varepsilon$, then Proposition 4 is satisfied as $\lambda_{\varepsilon, \varepsilon} = 0$, i.e. no jumps larger than ε happen. So, without loss of generality, let us assume that $u^* < \varepsilon$. In particular, since ν_ε is absolutely continuous with respect to the Lebesgue measure, we have that $u^* = \sqrt{(\Sigma^2 + \sigma^2(u^*))\Delta \log(n)^2}$. Furthermore, let us assume that $N_{n\Delta}(u^*) \geq 1$, i.e. from now on we always condition by this event. This assumption, combined with (43), implies that on ξ_n there exists i such that $N_{i\Delta}(u^*) - N_{(i-2)\Delta}(u^*) = 1$, and therefore

$$|M_{i\Delta}(u^*, \varepsilon) - M_{(i-1)\Delta}(u^*, \varepsilon) - (M_{(i-1)\Delta}(u^*, \varepsilon) - M_{(i-2)\Delta}(u^*, \varepsilon))| \geq u^*.$$

In addition, by means of Equation (43), we also know that on ξ_n

$$|\bar{X}_{i\Delta}(u^*) - \bar{X}_{(i-1)\Delta}(u^*) - (\bar{X}_{(i-1)\Delta}(u^*) - \bar{X}_{(i-2)\Delta}(u^*))| \leq 2\sqrt{(\Sigma^2 + \sigma^2(u^*))\Delta \log(n)}.$$

Recalling the definition of u^* and taking $n > e^2$ we can conclude that, on ξ_n , it holds that $Z_i \geq u^*/2$.

Furthermore, by Equation (44) we know that on ξ'_n

$$S_n \leq 2\sqrt{2\Delta(\Sigma^2 + \sigma^2(u^+))} \leq 2\sqrt{2\Delta(\Sigma^2 + \sigma^2(u^*))},$$

which allows to conclude that, for $n > e^2$, on $\xi_n \cap \xi'_n$, it holds

$$Z_i \geq \frac{S_n \log(n)^2}{2\sqrt{2}} > S_n \log(n)^{3/2},$$

that is $\Phi_n^{(\max)} = 1$, as desired.

A.1.2 Proof of Proposition 5

Under H_0 : By means of Equations (41) and (42), for any $\omega' \in \xi'_n$ and $\omega'' \in \xi''_n$, we have

$$S_n(\omega') \geq \frac{\sqrt{\Delta\Sigma^2}}{\sqrt{\pi}} \quad \text{and} \quad \bar{Y}_{n,6}(\omega'') \leq C(\sqrt{\Delta\Sigma^2})^6.$$

Therefore, on $\xi''_n \cap \xi'_n$, we have $\bar{Y}_{n,6} < C\pi^3 S_n^6$, and thus $\Phi_{n,c}^{(6)} = 0$, as desired.

If $\Delta\mu_6(u^*)$ is large and no large jumps occur: On the one hand, by Equation (42) we know that, on $\xi_n'' \cap \{N_{n\Delta}(u^*) = 0\}$, it holds

$$\bar{Y}_{n,6} \geq \frac{1}{2} [2\Delta\mu_6(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^3].$$

On the other hand, on ξ_n' , by means of Equation (44), we have that $S_n \leq 2\sqrt{2\Delta(\Sigma^2 + \sigma^2(u^*))}$. Thus, denoting by C_c an absolute constant depending only on c , whenever

$$\Delta\mu_6(u^*) \geq C_c \Delta^3 (\Sigma^2 + \sigma^2(u^*))^3,$$

it holds that $\bar{Y}_{n,6} > cS_n^6$, on $\xi_n'' \cap \xi_n' \cap \{N_{n\Delta}(u^*) = 0\}$, for n larger than an absolute constant. We therefore conclude that $\Phi_{n,c}^{(6)} = 1$, as desired.

A.1.3 Proof of Proposition 6

We begin with some preliminary results.

Lemma 6. *For n larger than an universal constant, $\varepsilon \in (0, 1]$ and any $\log(n)^{-1} < \alpha \leq 1$ there exists an event ξ_n''' of probability larger than $1 - \alpha$ and two universal constants $c, C > 0$ such that the following holds:*

$$|\mathbb{E}[T_n^{(3)} | \xi_n'''] - \Delta\mu_3(\varepsilon)| \leq c \frac{\Delta^{3/2}(\Sigma^2 + \sigma^2(\varepsilon))^{3/2}}{\sqrt{n\alpha}},$$

and

$$\mathbb{V}(T_n^{(3)} | \xi_n''') \leq \frac{C}{n} [\Delta\mu_6(\varepsilon) + \Delta^2(\Sigma^2 + \sigma^2(\varepsilon))^3].$$

Corollary 3. *For any $\varepsilon \in (0, 1]$ and for any $\log(n)^{-1} < \alpha \leq 1$, there exists an event ξ_n''' of probability larger than $1 - \alpha$ and two universal constants $c, C > 0$ such that the following holds:*

$$|\mathbb{E}[T_n^{(3)} | \xi_n''', N_{n\Delta}(u^*) = 0] - \Delta\mu_3(u^*)| \leq c \frac{\Delta^{3/2}(\Sigma^2 + \sigma^2(u^*))^{3/2}}{\sqrt{n\alpha}},$$

and

$$\mathbb{V}(T_n^{(3)} | \xi_n''', N_{n\Delta}(u^*) = 0) \leq \frac{C}{n} [\Delta\mu_6(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^3].$$

Proof of Proposition 6. For some given α , let ξ_n''' be an event as in Corollary 3. If $3 \leq k \leq 6$, thanks to the hypothesis (46) on $\Delta\mu_6(u^*)$, there exists an universal constant $C > 0$ such that

$$\mathbb{V}(T_n^{(3)} | N_{n\Delta}(u^*) = 0, \xi_n''') \leq \frac{C}{n} [\Delta(\Sigma^2 + \sigma^2(u^*))]^3.$$

Therefore, using BCI, we have

$$\mathbb{P}\left(\left|\frac{T_n^{(3)} - \Delta\mu_3(u^*)}{\sqrt{\frac{C}{n} [\Delta(\Sigma^2 + \sigma^2(u^*))]^3}}\right| > \frac{c/C + 1}{\sqrt{\alpha}} | N_{n\Delta}(u^*) = 0\right) \leq 2\alpha. \quad (48)$$

Under H_0 : $\mu_3(u^*)$ and $\sigma^2(u^*)$ are zero and thus

$$\mathbb{P}\left(\left|\frac{T_n^{(3)}}{\sqrt{\frac{C}{n} \Delta^3 \Sigma^6}}\right| > \frac{1}{\sqrt{\alpha}}\right) \leq \alpha.$$

Therefore, recalling the definition of ξ_n' (see Equation (41)), we have that for $c > 0$ a large enough absolute constant, with probability larger than $1 - \alpha - \mathbb{P}((\xi_n')^c)$ it holds

$$|T_n^{(3)}| \leq \frac{c}{\sqrt{\alpha n}} S_n^3,$$

which means that the test is accepted with probability larger than $1 - 2\alpha$.

Under $H_{1,\rho}^{(3)}$ and conditional to $N_{\Delta n}(u^*) = 0$: Assume that for some $C_c > 0$ large enough absolute constant depending only on $c > 0$ we have by definition of u^* (since $u^* > u^+$)

$$\Delta|\mu_3(u^*)| \geq \rho \geq \frac{C_c}{\sqrt{n\alpha}} \sqrt{[\Delta(\Sigma^2 + \sigma^2(u^*))]^3}.$$

This implies by Equation (48) and for C_c large enough depending only on c that with probability larger than $1 - \alpha$

$$|T_n^{(3)}| \geq \frac{C_c}{2\sqrt{\alpha}} \left\{ \frac{\sqrt{\Delta^3(\Sigma^2 + \sigma^2(u^*))^3}}{\sqrt{n}} \right\}.$$

Therefore, for C_c large enough depending only on c , we have by definition of ξ'_n in Equation (41) that with probability larger than $1 - \alpha - \mathbb{P}(\xi'_n \leq c)$

$$|T_n^{(3)}| \geq \frac{c}{\sqrt{\alpha n}} S_n^3.$$

The test is thus rejected with probability larger than $1 - \alpha - \mathbb{P}(\xi'_n \leq c)$. \square

A.1.4 Proof of Proposition 7

Proposition 7 can be proved with arguments very similar to those used in the proof of Proposition 6.

Lemma 7. *For any $\varepsilon \in (0, 1]$ it holds $\mathbb{E}[T_n^{(4)}] = \Delta\mu_4(\varepsilon)$. For n larger than an absolute constant and for some universal constant $C > 0$, it holds*

$$\mathbb{V}(T_n^{(4)}) \leq \frac{C}{n} \left[\Delta\mu_8(\varepsilon) + [\Delta(\Sigma^2 + \sigma^2(\varepsilon))]^4 \right].$$

Corollary 4. *For any $\varepsilon \in (0, 1]$, it holds $\mathbb{E}[T_n^{(4)} | N_{n\Delta}(u^*) = 0] = \Delta\mu_4(u^*)$. Moreover, there exists a universal constant $C > 0$ such that*

$$\mathbb{V}(T_n^{(4)} | N_{n\Delta}(u^*) = 0) \leq \frac{C}{n} \left[\Delta\mu_8(u^*) + [\Delta(\Sigma^2 + \sigma^2(u^*))]^4 \right]. \quad (49)$$

Proof of Proposition 7. The proof follows the same scheme as the one in Lemma 6. Here we only remark that Equation (49) implies

$$\mathbb{V}(T_n^{(4)} | N_{n\Delta}(u^*) = 0) \leq \frac{C}{n} \left[\Delta\mu_8(u^*) + [\Delta(\Sigma^2 + \sigma^2(u^*))]^4 \right] \leq \frac{C}{n} \left[(u^*)^4 \Delta\mu_4(u^*) + [\Delta(\Sigma^2 + \sigma^2(u^*))]^4 \right],$$

since $\Delta\mu_8(u^*) \leq (u^*)^4 \Delta\mu_4(u^*)$. By means of BCI we thus deduce that

$$\mathbb{P} \left(\left| \frac{T_n^{(4)} - \Delta\mu_4(u^*)}{\sqrt{\frac{C}{n} \left[(u^*)^4 \Delta\mu_4(u^*) + [\Delta(\Sigma^2 + \sigma^2(u^*))]^4 \right]}} \right| > \frac{1}{\sqrt{\alpha}} | N_{n\Delta}(u^*) = 0 \right) \leq \alpha,$$

which, using that $\alpha \geq \log(n)^{-1}$ and the definition of u^* , implies

$$\mathbb{P} \left(\frac{T_n^{(4)} - \frac{1}{2} \Delta\mu_4(u^*)}{\sqrt{\frac{C'}{n} [\Delta(\Sigma^2 + \sigma^2(u^*))]^4}} < \frac{1}{\sqrt{\alpha}} | N_{n\Delta}(u^*) = 0 \right) \leq \alpha,$$

for some universal constant C' and for n larger than a universal constant. \square

A.2 Proofs of the Lemmas involved in the proof of Theorem 5

A.2.1 Proof of Lemma 1

Let Φ be such a test for $H_0 : \mathbb{P}$ against $H_1 : \mathbb{Q}$. The conditions on Φ lead to

$$\|\mathbb{P}^{\otimes n} - \mathbb{Q}^{\otimes n}\|_{TV} \geq |\mathbb{P}_{H_0}(\Phi = 0) - \mathbb{P}_{H_1}(\Phi = 0)| \geq 1 - \alpha_1 - \alpha_0.$$

A.2.2 Proof of Lemma 3

If $\nu_\varepsilon = 0$: Under H_0 we know that all Z_i are i.i.d. realizations of the absolute values of centered Gaussian random variables with variance $2\Delta\Sigma^2$. By Gaussian concentration we thus have that with probability larger than $1 - 1/n$, $\max_{i \leq \tilde{n}} Z_i \leq 4\sqrt{\Delta\Sigma^2 \log(n)}$, since $\tilde{n} = \lfloor n/2 \rfloor \leq n$.

If $\nu_\varepsilon \neq 0$ is an infinite Lévy measure : By BCI, with probability larger than $1 - \alpha$, it holds

$$|N_{n\Delta}(u^+) - \Delta n \lambda_{u^+, \varepsilon}| \leq \sqrt{\Delta n \lambda_{u^+, \varepsilon} / \alpha},$$

i.e. for $\alpha = 4 \log(n)^{-1}$ we have that with probability larger than $1 - 4 \log(n)^{-1}$

$$\log(n)/2 \leq N_{n\Delta}(u^+) \leq 2 \log(n).$$

Furthermore, we observe that

$$\begin{aligned} \mathbb{P}(N_{i\Delta}(u^+) - N_{(i-2)\Delta}(u^+) \leq 1) &= \exp(-2\Delta\lambda_{u^+, \varepsilon}) + 2\Delta\lambda_{u^+, \varepsilon} \exp(-2\Delta\lambda_{u^+, \varepsilon}) \\ &= \exp(-2 \log(n)/n) + 2 \frac{\log(n)}{n} \exp(-2 \log(n)/n). \end{aligned}$$

It thus follows

$$\begin{aligned} \mathbb{P}\left(\{\forall i \leq n, N_{i\Delta}(u^+) - N_{(i-2)\Delta}(u^+) \leq 1\}\right) &= \left(\exp(-2 \log(n)/n) + 2 \frac{\log(n)}{n} \exp(-2 \log(n)/n)\right)^n \\ &= \exp(-2 \log(n))(1 + 2 \log(n)/n)^n \rightarrow 1, \end{aligned}$$

at a rate which does not depend on $\nu, \varepsilon, b, \Sigma$.

Finally, by BCI, with probability larger than $1 - \alpha$, we have

$$|\bar{X}_{i\Delta}(u^*) - \bar{X}_{(i-1)\Delta}(u^*) - b\Delta| \leq \sqrt{(\Sigma^2 + \sigma^2(u^*))\Delta\alpha^{-1}}.$$

So, conditional on $\{1 \leq N_{n\Delta}(u^+) \leq 2 \log(n)\}$, with probability larger than $1 - \log(n)^{-1}$, we have that $\forall i$ s.t. $N_{i\Delta}(u^+) - N_{(i-1)\Delta}(u^+) \neq 0$

$$|\bar{X}_{i\Delta}(u^*) - \bar{X}_{(i-1)\Delta}(u^*) - b\Delta| \leq \sqrt{(\Sigma^2 + \sigma^2(u^*))\Delta} \log(n).$$

We conclude observing that, conditional on $\{1 \leq N_{n\Delta}(u^+) \leq 2 \log(n)\}$, with probability larger than $1 - \log(n)^{-1}$, we have $\forall i$ s.t. $N_{i\Delta}(u^+) - N_{(i-2)\Delta}(u^+) \neq 0$

$$|\bar{X}_{i\Delta}(u^*) - \bar{X}_{(i-1)\Delta}(u^*) - (\bar{X}_{(i-1)\Delta}(u^*) - \bar{X}_{(i-2)\Delta}(u^*))| \leq 2\sqrt{(\Sigma^2 + \sigma^2(u^*))\Delta} \log(n).$$

A.2.3 Proof of Lemma 4

Preliminary. Denote by $\tilde{Z}_i = |(X_{i\Delta}(u^+) - X_{(i-1)\Delta}(u^+)) - (X_{(i-1)\Delta}(u^+) - X_{(i-2)\Delta}(u^+))|$ and assume that ξ_n holds. We begin by observing that, for any $\omega \in \xi_n$, we have:

$$\frac{1}{\tilde{n}} \sum_{i=1}^{\tilde{n}-4\log(n)} \tilde{Z}_{(i)}(\omega) \leq S_n(\omega) \leq \frac{1}{\tilde{n}} \sum_{i=1}^{\tilde{n}} \tilde{Z}_i(\omega). \quad (50)$$

To show (50), let $I := \{i : \tilde{Z}_i = Z_i\}$, that is the set where no jumps of size larger than u^+ occur between $(i-2)\Delta$ and $i\Delta$. By means of the positivity of the variables \tilde{Z}_i and Z_i , we get

$$\frac{1}{\tilde{n}} \sum_{1 \leq i \leq \tilde{n}-2\log(n), i \in I} Z_{(i)} \leq S_n = \frac{1}{\tilde{n}} \sum_{i=1}^{\tilde{n}-2\log(n)} Z_{(i)}.$$

Moreover, since $\#I^c \leq 2\log(n)$ on ξ_n , we have

$$\frac{1}{\tilde{n}} \sum_{1 \leq i \leq \tilde{n}-2\log(n), i \in I} Z_{(i)} \leq S_n \leq \frac{1}{\tilde{n}} \sum_{i \in I} Z_i.$$

Using again that $\mathbb{P}\{\#I^c \leq 2\log(n) \cap \xi_n\} = 1$, the definition of I and the fact that \tilde{Z}_i, Z_i are positive, we obtain (50).

Control when ξ'_n is given by (44). Note that $\mathbb{E}\tilde{Z}_i^2 = 2\Delta(\Sigma^2 + \sigma^2(u^+))$, so by the Cauchy-Schwartz inequality $\mathbb{E}\tilde{Z}_i \leq \sqrt{2\Delta(\Sigma^2 + \sigma^2(u^+))}$. It follows by BCI, that with probability larger than $1 - 1/n$

$$\frac{1}{\tilde{n}} \sum_{i=1}^{\tilde{n}} \tilde{Z}_i \leq 2\sqrt{2\Delta(\Sigma^2 + \sigma^2(u^+))}.$$

Then, on ξ_n , by Equation (50), with probability larger than $1 - 1/n$ it holds $S_n \leq 2\sqrt{2\Delta(\Sigma^2 + \sigma^2(u^+))}$.

Control when ξ'_n is given by (41). In this case $\nu_\varepsilon = 0$, then $Z_i = \tilde{Z}_i$ which are i.i.d. and distributed like the absolute value of a centered Gaussian random variable with variance $2\Delta\Sigma^2$. By Gaussian concentration it then follows that with probability larger than $1 - \alpha$

$$\max_i |Z_i| \leq 2\sqrt{2\Delta\Sigma^2 \log(2/\alpha)}.$$

Using that $\mathbb{E}_{V \sim \mathcal{N}(0,1)} |V| = \sqrt{\frac{2}{\pi}}$ and BCI, we conclude that with probability larger than $1 - \alpha$

$$\frac{1}{\tilde{n}} \sum_{i=1}^{\tilde{n}-4\log(n)} \tilde{Z}_{(i)} \geq \sqrt{\frac{2}{\pi}} \sqrt{2\Delta\Sigma^2} - \sqrt{\frac{1}{n\alpha} 2\Delta\Sigma^2} - \frac{4\log(n)}{n} \sqrt{2\Delta\Sigma^2 \log(2/\alpha)}.$$

By Equation (50), with probability larger than $1 - 2^9\pi n^{-1}$, it thus holds

$$S_n \geq \frac{\sqrt{\Delta\Sigma^2}}{\sqrt{\pi}}.$$

A.2.4 Preliminaries for the proofs of Lemmas 5, 6 and 7

If $Y \sim \mathcal{N}(m, \Sigma^2)$ its moments can be computed through the recursive formula:

$$\mathbb{E}[Y^k] = (k-1)\mathbb{V}(Y)\mathbb{E}[Y^{k-2}] + \mathbb{E}[Y]\mathbb{E}[Y^{k-1}], \quad k \in \mathbb{N}.$$

Lemma 8. *Let $Y \sim \mathcal{N}(m, \Sigma^2)$, then it holds*

$$\begin{aligned} \mathbb{E}[Y^3] &= 3\Sigma^2 m + m^3, & \mathbb{E}[Y^4] &= 3\Sigma^4 + 6m^2\Sigma^2 + m^4, & \mathbb{E}[Y^5] &= 15m\Sigma^4 + 10m^3\Sigma^2 + m^5, \\ \mathbb{E}[Y^6] &= 15\Sigma^6 + 45m^2\Sigma^4 + 15m^4\Sigma^2 + m^6, & \mathbb{E}[Y^7] &= 105m^4\Sigma^6 + 105m^3\Sigma^4 + 21m^5\Sigma^2 + m^7, \\ \mathbb{E}[Y^8] &= 105\Sigma^8 + 420m^2\Sigma^6 + 210m^4\Sigma^4 + 28m^6\Sigma^2 + m^8. \end{aligned}$$

Similarly, using the series expansion of the characteristic function, together with the Lévy Kintchine formula and $\varepsilon \in (0, 1]$, we get

$$\mathbb{E}[M_\Delta(\varepsilon)^k] = \frac{d^k}{du^k} \exp\left(\Delta \int_{-\varepsilon}^{\varepsilon} (e^{iux} - 1 - iux)\nu_\varepsilon(dx)\right) \Big|_{u=0}. \quad (51)$$

Lemma 9. *For $\varepsilon \in (0, 1]$, set $\sigma^2(\varepsilon) = \int_{-\varepsilon}^{\varepsilon} y^2\nu(dy)$ and $\mu_k(\varepsilon) := \int_{-\varepsilon}^{\varepsilon} y^k\nu(dy)$, $k \geq 3$. We have*

$$\begin{aligned} \mathbb{E}[M_\Delta(\varepsilon)] &= 0, & \mathbb{E}[(M_\Delta(\varepsilon))^2] &= \Delta\sigma^2(\varepsilon), & \mathbb{E}[(M_\Delta(\varepsilon))^3] &= \Delta\mu_3(\varepsilon) \\ \mathbb{E}[(M_\Delta(\varepsilon))^4] &= \Delta\mu_4(\varepsilon) + 3\Delta^2\sigma^4(\varepsilon), & \mathbb{E}[(M_\Delta(\varepsilon))^5] &= \Delta\mu_5(\varepsilon) + 10\Delta^2\sigma^2(\varepsilon)\mu_3(\varepsilon), \\ \mathbb{E}[(M_\Delta(\varepsilon))^6] &= \Delta\mu_6(\varepsilon) + \Delta^2(10\mu_3(\varepsilon)^2 + 15\sigma^2(\varepsilon)\mu_4(\varepsilon)) + 15\Delta^3\sigma^6(\varepsilon), \\ \mathbb{E}[(M_\Delta(\varepsilon))^7] &= \Delta\mu_7(\varepsilon) + \Delta^2(21\sigma^2(\varepsilon)\mu_5(\varepsilon) + 35\mu_3(\varepsilon)\mu_4(\varepsilon)) + 105\Delta^3\sigma^4(\varepsilon)\mu_3(\varepsilon), \\ \mathbb{E}[(M_\Delta(\varepsilon))^8] &= \Delta\mu_8(\varepsilon) + \Delta^2(35\mu_4(\varepsilon)^2 + 56\mu_3(\varepsilon)\mu_5(\varepsilon) + 28\sigma^2(\varepsilon)\mu_6(\varepsilon)) \\ &\quad + \Delta^3(280\sigma^2(\varepsilon)\mu_3(\varepsilon)^2 + 210\sigma^4(\varepsilon)\mu_4(\varepsilon)) + 105\Delta^4\sigma^8(\varepsilon). \end{aligned}$$

More generally if ν_ε is an infinite Lévy measure, for any $k \geq 2$ even integer it holds that

$$\mathbb{E}[M_\Delta(\varepsilon)^k] \leq \tilde{C}_k \left(\Delta\mu_k(\varepsilon) + (\Delta\sigma^2(\varepsilon))^{k/2} \right),$$

where $\tilde{C}_k > 0$ is a constant that depends only on k .

Proof of Lemma 9. The explicit first 8 moments are computed using Equation (51). We prove now the last part of the Lemma. Denote $M(\eta, \varepsilon)$ the Lévy process with jump measure $\nu_\varepsilon \mathbf{1}_{[-\eta, \eta]^c}$, $N(\eta, \varepsilon)$ the corresponding Poisson process and let $\mathcal{L}(\eta, \varepsilon)$ be the law of the jumps given by $\frac{\nu_\varepsilon \mathbf{1}_{[-\eta, \eta]^c}}{\lambda_{\eta, \varepsilon}}$.

By Rosenthal's inequality there exists a constant \tilde{C}_k , depending on k only, such that for any $k \geq 2$, k even

$$\mathbb{E}[|M_\Delta(\eta, \varepsilon)|^k | N_\Delta(\eta, \varepsilon) = N] \leq \tilde{C}_k \max \left[N\mathbb{E}_{X \sim \mathcal{L}(\eta, \varepsilon)}[X^k], [N\mathbb{E}_{X \sim \mathcal{L}(\eta, \varepsilon)}[X^2]]^{k/2} \right].$$

Averaging over $N_\Delta(\eta, \varepsilon) \sim \mathcal{P}(\Delta\lambda_{\eta, \varepsilon})$, we have that

$$\begin{aligned} \mathbb{E}[|M_\Delta(\eta, \varepsilon)|^k] &\leq \tilde{C}_k \max \left[\Delta\lambda_{\eta, \varepsilon}\mathbb{E}_{X \sim \mathcal{L}(\eta, \varepsilon)}[X^k], \mathbb{E}[N_\Delta(\eta, \varepsilon)^{k/2}] [\mathbb{E}_{X \sim \mathcal{L}(\eta, \varepsilon)}[X^2]]^{k/2} \right] \\ &\leq \tilde{C}'_k \max \left[\Delta\lambda_{\eta, \varepsilon}\mathbb{E}_{X \sim \mathcal{L}(\eta, \varepsilon)}[X^k], [1 + \Delta\lambda_{\eta, \varepsilon}]^{k/2} [\mathbb{E}_{X \sim \mathcal{L}(\eta, \varepsilon)}[X^2]]^{k/2} \right], \end{aligned}$$

where we used that for a Poisson random variable it holds for any $k' \geq 1$ that $\mathbb{E}_{N \sim \mathcal{P}(\lambda)}X^{k'} \leq c_k(1+\lambda)^{k'}$ where c_k is a constant that depends on k only. Setting $\mu_k(\eta, \varepsilon) = \int_{\eta \leq |x| \leq \varepsilon} y^k\nu(dy)$ we have for η small

enough such that $\Delta\lambda_{\eta,\varepsilon} \geq 1$ (which always exists for infinite Lévy measure)

$$\begin{aligned} \mathbb{E}[|M_\Delta(\eta, \varepsilon)|^k] &\leq \tilde{C}_k'' \max \left[\Delta\mu_k(\eta, \varepsilon), [\Delta\lambda_{\eta,\varepsilon}]^{k/2} [\mathbb{E}_{X \sim \mathcal{P}(\eta,\varepsilon)} X^2]^{k/2} \right] \\ &\leq \tilde{C}_k'' \max \left[\Delta\mu_k(\eta, \varepsilon), [\Delta\mu_2(\eta, \varepsilon)]^{k/2} \right]. \end{aligned}$$

where \tilde{C}_k'' is some constant that only depends on k . Making η converge to 0 gives the result. \square

Corollary 5. For $\varepsilon \in (0, 1]$, set $\sigma^2(\varepsilon) = \int_{-\varepsilon}^{\varepsilon} y^2 \nu(dy)$ and $\mu_k(\varepsilon) := \int_{-\varepsilon}^{\varepsilon} y^k \nu(dy)$, $k \geq 2$. It holds that

$$\begin{aligned} \mathbb{E}(X_\Delta - b\Delta) &= 0, & \mathbb{E}(X_\Delta - b\Delta)^3 &= \Delta\mu_3(\varepsilon), & \mathbb{E}(X_\Delta - \bar{X}_{\Delta,n}) &= 0, \\ \mathbb{E}(X_\Delta - \bar{X}_{\Delta,n})^3 &= \left(1 - \frac{1}{(n - \lfloor n/2 \rfloor)^2}\right) \Delta\mu_3(\varepsilon), & \mathbb{E}Z_1^2 &= 2\Delta(\Sigma^2 + \sigma^2(\varepsilon)) \\ \mathbb{E}Z_1^4 &= 4\Delta\mu_4(\varepsilon) + 12(\Delta(\Sigma^2 + \sigma^2(\varepsilon)))^2, & \mathbb{E}Z_1^6 &\geq 2\Delta\mu_6(\varepsilon) + (\Delta(\Sigma^2 + \sigma^2(\varepsilon)))^3. \end{aligned}$$

Proof. The result is obtained combining Lemmas 8 and 9 with the independence between $(X_{2\Delta} - X_\Delta)$ and X_Δ . \square

Corollary 6. For $\varepsilon \in (0, 1]$, set $\sigma^2(\varepsilon) = \int_{-\varepsilon}^{\varepsilon} y^2 \nu(dy)$ and $\mu_k(\varepsilon) := \int_{-\varepsilon}^{\varepsilon} y^k \nu(dy)$, $k \geq 2$. It holds for any $k \geq 2$, k even integer that

$$\begin{aligned} \mathbb{E}(X_\Delta(\varepsilon) - b\Delta)^k &\leq \bar{C}_k(\Delta\mu_k(\varepsilon) + (\Delta(\Sigma^2 + \sigma^2(\varepsilon)))^{k/2}), \\ \mathbb{E}(Z_1^k) &\leq \bar{C}_k(\Delta\mu_k(\varepsilon) + (\Delta(\Sigma^2 + \sigma^2(\varepsilon)))^{k/2}), \end{aligned}$$

where \bar{C}_k is a constant that depends only on k .

Proof of Corollary 6.

$$\begin{aligned} \mathbb{E}(X_\Delta(\varepsilon) - b\Delta)^k &= \mathbb{E}(M_\Delta(\varepsilon) + \Sigma W_\Delta)^k \leq 2^{k+1}(\mathbb{E}M_\Delta(\varepsilon)^k + \Sigma^k \mathbb{E}W_\Delta^k) \\ &\leq C_k(\Delta^{k/2}\Sigma^k + \Delta\mu_k(\varepsilon) + (\Delta\sigma^2(\varepsilon))^{k/2}), \end{aligned}$$

where C_k is a constant that depends on k only. \square

A.2.5 Proof of Lemma 5

If $\nu_\varepsilon = 0$. By Corollary 6, there exist universal constants \bar{C}_6 and \bar{C}_{12} such that $\mathbb{E}[Z_i^6] \leq \bar{C}_6\Delta^3\Sigma^6$ and $\mathbb{E}[Z_i^{12}] \leq \bar{C}_{12}\Delta^6\Sigma^{12}$. Using that Z_i are i.i.d. we get

$$\mathbb{E}[\bar{Y}_{n,6}] \leq \bar{C}_6\Delta^3\Sigma^6, \quad \mathbb{V}(\bar{Y}_{n,6}) \leq \frac{\bar{C}_{12}}{n - \lfloor n/2 \rfloor} \Delta^6\Sigma^{12}.$$

Therefore, by means of BCI, with probability larger than $1 - \log(n)^{-1}$ conditional to $N_{n\Delta}(u^*) = 0$ we have

$$\bar{Y}_{n,6} \leq \bar{C}_6\Delta^3\Sigma^6 + \sqrt{\frac{\bar{C}_{12} \log(n)}{(n - \lfloor n/2 \rfloor)}} (\Delta\Sigma^2)^6,$$

which allows to deduce that for n larger than a universal constant, with probability larger than $1 - \log(n)^{-1}$, we have $\bar{Y}_{n,6} \leq 2\bar{C}_6\Delta^3\Sigma^6$.

If $\nu_\varepsilon \neq 0$. By Corollaries 5 and 6, conditional to $N_{n\Delta}(u^*) = 0$, for any i it holds

$$\begin{aligned}\mathbb{E}[Z_i^6 | N_{n\Delta}(u^*) = 0] &\geq \Delta\mu_6(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^3 \\ \mathbb{E}[Z_i^{12} | N_{n\Delta}(u^*) = 0] &\leq \bar{C}_{12}(\Delta\mu_{12}(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^6),\end{aligned}$$

where \bar{C}_{12} is the universal constant from Corollary 6. Using that the Z_i are i.i.d. we obtain

$$\begin{aligned}\mathbb{E}[\bar{Y}_{n,6} | N_{n\Delta}(u^*) = 0] &\geq \Delta\mu_6(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^6 \\ \mathbb{V}(\bar{Y}_{n,6} | N_{n\Delta}(u^*) = 0) &\leq \frac{\bar{C}_{12}}{n - \lfloor n/2 \rfloor} (\Delta\mu_{12}(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^6).\end{aligned}$$

It follows by BCI that with probability larger than $1 - \log(n)^{-1}$, conditional to $N_{n\Delta}(u^*) = 0$, we have

$$\bar{Y}_{n,6} \geq \Delta\mu_6(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^3 - \sqrt{\frac{\bar{C}_{12} \log(n)}{(n - \lfloor n/2 \rfloor)} (\Delta\mu_{12}(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^6)}.$$

Since $\mu_{12}(u^*) \leq (u^*)^6 \mu_6(u^*)$, with probability larger than $1 - \log(n)^{-1}$ and conditional to $N_{n\Delta}(u^*) = 0$, we have

$$\bar{Y}_{n,6} \geq \Delta\mu_6(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^3 - \sqrt{\frac{\bar{C}_{12} \log(n)}{(n - \lfloor n/2 \rfloor)} ((u^*)^6 \Delta\mu_6(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^6)}.$$

Finally, by means of the definition of u^* and for n larger than a universal constant, with probability larger than $1 - \log(n)^{-1}$ conditional to $N_{n\Delta}(u^*) = 0$, it holds

$$\bar{Y}_{n,6} \geq \frac{\Delta\mu_6(u^*) + (\Delta(\Sigma^2 + \sigma^2(u^*)))^3}{2}.$$

A.2.6 Proof of Lemma 6

Note that $\mathbb{E}[\bar{X}_{\Delta,n}] = b\Delta$ and that $\mathbb{V}(\bar{X}_{\Delta,n}) = (n - \lfloor n/2 \rfloor)^{-1} \Delta(\Sigma^2 + \sigma^2(\varepsilon))$. Write

$$\xi_n''' = \{|\bar{X}_{\Delta,n} - b\Delta| \leq r_n\} \quad \text{where} \quad r_n := r_n(\alpha) = \sqrt{\frac{\Delta(\Sigma^2 + \sigma^2(\varepsilon))}{\alpha(n - \lfloor n/2 \rfloor)^{-1}}}.$$

By BCI we have for any $0 < \alpha \leq 1$

$$\mathbb{P}(\xi_n''') \geq 1 - \alpha. \tag{52}$$

Conditional on ξ_n''' we have by Corollary 5 and the definition of r_n

$$|\mathbb{E}[\bar{Y}_{n,3} | \xi_n'''] - \Delta\mu_3(\varepsilon)| \leq r_n^3 + 3\Delta(\Sigma^2 + \sigma^2(\varepsilon))r_n \leq 100 \frac{(\Delta(\Sigma^2 + \sigma^2(\varepsilon)))^{3/2}}{\sqrt{n\alpha}},$$

for $\alpha \geq \log(n)^{-1}$.

Now we compute $\mathbb{V}(T_n^{(3)})$. Since $\bar{X}_{\Delta,n}$ and $\bar{X}_{\Delta,n}^j$ are independent, as they are computed on two independent samples, the elements of the sum are independent of each other conditional on the second half of the sample. Then, conditional on the second half of the sample

$$\begin{aligned}\mathbb{V}(T_n^{(3)} | \xi_n''') &\leq \left(1 - \frac{1}{(n - \lfloor n/2 \rfloor)^2}\right)^2 \frac{1}{\lfloor n/2 \rfloor^2} \sum_{i=1}^{\lfloor n/2 \rfloor} \mathbb{E}\left[[(X_{i\Delta} - X_{(i-1)\Delta} - \bar{X}_{\Delta,n})^3]^2 | \xi_n'''\right] \\ &\leq \frac{16}{n} \mathbb{E}\left[[X_{i\Delta} - X_{(i-1)\Delta} - \Delta b]^6 + [\bar{X}_{\Delta,n} - \Delta b]^6 | \xi_n'''\right] \\ &\leq \frac{16}{n} \left[\bar{C}_6 \left(\Delta\mu_6(\varepsilon) + [\Delta(\Sigma^2 + \sigma^2(\varepsilon))]^3\right) + r_n^6\right],\end{aligned}$$

where \bar{C}_6 is the constant from Corollary 6. Hence, by Equation (52), there exists a universal constant C such that

$$\mathbb{V}(T_n^{(3)}|\xi_n''') \leq \frac{C}{n} \left[\Delta\mu_6(\varepsilon) + [\Delta(\Sigma^2 + \sigma^2(\varepsilon))]^3 \right].$$

A.2.7 Proof of Lemma 7

The main ingredient of the proof consists in establishing expansions of $\mathbb{V}(T_n^{(4)})$. Computations are cumbersome but not difficult, we only give the main tools here but we do not provide all computations. By Corollary 5 and since $\bar{Y}_{n,2}$ and $\bar{Y}'_{n,2}$ are independent, we have

$$\begin{aligned} \mathbb{E}[\bar{Y}_{n,2}] &= 2\Delta(\Sigma^2 + \sigma^2(\varepsilon)), & \mathbb{E}[\bar{Y}_{n,2}\bar{Y}'_{n,2}] &= 4\Delta^2(\Sigma^2 + \sigma^2(\varepsilon))^2, \\ \mathbb{E}[\bar{Y}_{n,4}] &= 3 \times 4\Delta^2(\Sigma^2 + \sigma^2(\varepsilon))^2 + 4\Delta\mu_4(\varepsilon). \end{aligned}$$

In particular, $\mathbb{E}[T_n^{(4)}] = \Delta\mu_4(\varepsilon)$. Next, we have $\mathbb{V}[T_n^{(4)}] \leq 9\mathbb{V}[\bar{Y}_{n,2}\bar{Y}'_{n,2}] + \mathbb{V}[\bar{Y}_{n,4}]$. We analyse these two terms separately.

Since the Z_i in the sum composing $\bar{Y}_{n,4}$ are i.i.d. we have

$$\begin{aligned} \mathbb{V}(\bar{Y}_{n,4}) &\leq \frac{1}{[n/2]^2} \sum_{i=1}^{[n/2]} \mathbb{E}(Z_i^8) \leq \frac{1}{[n/2]} 2^9 \bar{C}_8 \left(\Delta\mu_8(\varepsilon) + [\Delta(\Sigma^2 + \sigma^2(\varepsilon))]^4 \right) \\ &\leq \frac{C}{n} \left(\Delta\mu_8(\varepsilon) + [\Delta(\Sigma^2 + \sigma^2(\varepsilon))]^4 \right), \end{aligned}$$

where \bar{C}_8 is the constant from Corollary 6, and where C is a universal constant.

Similarly, as the Z_i in the sums composing $\bar{Y}_{n,2}$ and $\bar{Y}'_{n,2}$ are i.i.d. we have

$$\begin{aligned} \mathbb{V}(\bar{Y}_{n,2}\bar{Y}'_{n,2}) &\leq 4\mathbb{E}[\bar{Y}_{n,2}^2]\mathbb{V}(\bar{Y}'_{n,2}) \leq 4\frac{\bar{C}_4}{[n/4]} \left(\Delta\mu_4(\varepsilon) + [\Delta(\Sigma^2 + \sigma^2(\varepsilon))]^2 \right) \\ &\quad \times \left[\frac{\bar{C}_4}{[n/4]} \left(\Delta\mu_4(\varepsilon) + [\Delta(\Sigma^2 + \sigma^2(\varepsilon))]^2 \right) + 4\Delta^2(\Sigma^2 + \sigma^2(\varepsilon))^2 \right] \\ &\leq \frac{C'}{n} \left[[\Delta(\Sigma^2 + \sigma^2(\varepsilon))]^4 + \frac{\Delta^2\mu_4(\varepsilon)^2}{n} \right], \end{aligned}$$

where \bar{C}_4 is the constant from Corollary 6, and where C' is a universal constant. Combining this with the last displayed equation, completes the proof.

B Some bounds in total variation distance

Lemma 10. *If $\mu_1, \mu_2 \in \mathbb{R}$, $0 \leq \Sigma_1^2 \leq \Sigma_2^2$ and $\Sigma_2 \in \mathbb{R}_{>0}$, then*

$$\|\mathcal{N}(\mu_1, \Sigma_1^2)^{\otimes n} - \mathcal{N}(\mu_2, \Sigma_2^2)^{\otimes n}\|_{TV} \leq 1 - \left(\frac{\Sigma_1}{\Sigma_2} \right)^n + \sqrt{n} \frac{|\mu_1 - \mu_2|}{\sqrt{2\pi\Sigma_2^2}}, \quad \forall n \geq 1.$$

Proof. Since the Lemma is trivially true if $\Sigma_1 = 0$ we can assume that $\Sigma_1 > 0$ without loss of generality. By triangular inequality, for any $n \geq 1$ we have:

$$\|\mathcal{N}(\mu_1, \Sigma_1^2)^{\otimes n} - \mathcal{N}(\mu_2, \Sigma_2^2)^{\otimes n}\|_{TV} \leq \|\mathcal{N}(\mu_1, \Sigma_1^2)^{\otimes n} - \mathcal{N}(\mu_1, \Sigma_2^2)^{\otimes n}\|_{TV} + \|\mathcal{N}(\mu_1, \Sigma_2^2)^{\otimes n} - \mathcal{N}(\mu_2, \Sigma_2^2)^{\otimes n}\|_{TV}.$$

Denote by $\varphi_n(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{x_1^2 + \dots + x_n^2}{2}}$ the density of a standard normal distribution on \mathbb{R}^n and observe that

$$\begin{aligned}
\|\mathcal{N}(\mu_1, \Sigma_1^2)^{\otimes n} - \mathcal{N}(\mu_1, \Sigma_2^2)^{\otimes n}\|_{TV} &= \|\mathcal{N}(0, \Sigma_1^2)^{\otimes n} - \mathcal{N}(0, \Sigma_2^2)^{\otimes n}\|_{TV} \\
&= \frac{1}{2} \int_{\mathbb{R}^n} \left| \varphi_n(z_1, \dots, z_n) - \left(\frac{\Sigma_1}{\Sigma_2}\right)^n \varphi_n\left(\frac{\Sigma_1}{\Sigma_2} z_1, \dots, \frac{\Sigma_1}{\Sigma_2} z_n\right) \right| dz_1 \dots dz_n \\
&\leq \frac{1}{2} \int_{\mathbb{R}^n} \left| 1 - \left(\frac{\Sigma_1}{\Sigma_2}\right)^n \right| \varphi_n(z_1, \dots, z_n) dz_1 \dots dz_n \\
&\quad + \frac{1}{2} \int_{\mathbb{R}^n} \left(\frac{\Sigma_1}{\Sigma_2}\right)^n \left| \varphi_n(z_1, \dots, z_n) - \varphi_n\left(\frac{\Sigma_1}{\Sigma_2} z_1, \dots, \frac{\Sigma_1}{\Sigma_2} z_n\right) \right| dz_1 \dots dz_n \\
&= \frac{1 - \left(\frac{\Sigma_1}{\Sigma_2}\right)^n}{2} + \frac{1}{2} \int_{\mathbb{R}^n} \left(\frac{\Sigma_1}{\Sigma_2}\right)^n \left(\varphi_n\left(\frac{\Sigma_1}{\Sigma_2} z_1, \dots, \frac{\Sigma_1}{\Sigma_2} z_n\right) - \varphi_n(z_1, \dots, z_n) \right) dz_1 \dots dz_n \\
&= \frac{1 - \left(\frac{\Sigma_1}{\Sigma_2}\right)^n}{2} + \frac{1}{2} \left(\frac{\Sigma_1}{\Sigma_2}\right)^n \left(\left(\frac{\Sigma_2}{\Sigma_1}\right)^n - 1 \right) = 1 - \left(\frac{\Sigma_1}{\Sigma_2}\right)^n.
\end{aligned}$$

Let U be any $n \times n$ orthonormal matrix with first column given by $(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}})^t$, set $\alpha = \frac{\mu_1 - \mu_2}{\Sigma_2}$ and $\beta = U^t(\alpha, \dots, \alpha)^t = (\sqrt{n}\alpha, 0, \dots, 0)^t$. Observe that $\varphi_n(Ux) = \varphi_n(x)$ for any $x \in \mathbb{R}^n$. We have:

$$\begin{aligned}
\|\mathcal{N}(\mu_1, \Sigma_2^2)^{\otimes n} - \mathcal{N}(\mu_2, \Sigma_2^2)^{\otimes n}\|_{TV} &= \frac{1}{2\Sigma_2^2} \int_{\mathbb{R}^n} \left| \varphi_n\left(\frac{x_1 - \mu_1}{\Sigma_2}, \dots, \frac{x_n - \mu_1}{\Sigma_2}\right) - \varphi_n\left(\frac{x_1 - \mu_2}{\Sigma_2}, \dots, \frac{x_n - \mu_2}{\Sigma_2}\right) \right| dx_1 \dots dx_n \\
&= \frac{1}{2} \int_{\mathbb{R}^n} \left| \varphi_n(z_1 - \alpha, \dots, z_n - \alpha) - \varphi_n(z_1, \dots, z_n) \right| dz_1 \dots dz_n.
\end{aligned}$$

Let $z = (z_1, \dots, z_n)^t$. By the change of variable $y = U^t z$ we get

$$\begin{aligned}
&\int_{\mathbb{R}^n} \left| \varphi_n(z_1 - \alpha, \dots, z_n - \alpha) - \varphi_n(z_1, \dots, z_n) \right| dz_1 \dots dz_n \\
&= \int_{\mathbb{R}^n} |\varphi_n(Uy - (\alpha, \dots, \alpha)^t) - \varphi_n(Uy)| dy_1 \dots dy_n \\
&= \int_{\mathbb{R}^n} |\varphi_n(U(y - \beta)) - \varphi_n(Uy)| dy_1 \dots dy_n \\
&= \int_{\mathbb{R}^n} |\varphi_n(y_1 - \sqrt{n}\alpha, y_2, \dots, y_n) - \varphi_n(y_1, \dots, y_n)| dy_1 \dots dy_n \\
&= \int_{\mathbb{R}^n} \left| \varphi_1(y_1 - \sqrt{n}\alpha) \prod_{j=2}^n \varphi_1(y_j) - \prod_{j=1}^n \varphi_1(y_j) \right| dy_1 \dots dy_n \\
&= \int_{\mathbb{R}} |\varphi_1(y_1 - \sqrt{n}\alpha) - \varphi_1(y_1)| dy_1 = 2 \int_{-\frac{\sqrt{n}|\alpha|}{2}}^{\frac{\sqrt{n}|\alpha|}{2}} \varphi_1(y) dy \leq \frac{2\sqrt{n}|\alpha|}{\sqrt{2\pi}}.
\end{aligned}$$

We thus deduce that

$$\|\mathcal{N}(\mu_1, \Sigma_2^2)^{\otimes n} - \mathcal{N}(\mu_2, \Sigma_2^2)^{\otimes n}\|_{TV} \leq \frac{\sqrt{n}|\mu_1 - \mu_2|}{\sqrt{2\pi}\Sigma_2}.$$

□

Lemma 11. Let $(X_i)_{i \geq 1}$ and $(Y_i)_{i \geq 1}$ be sequences of i.i.d. random variables a.s. different from zero and N, N' be two Poisson random variables with N (resp. N') independent of $(X_i)_{i \geq 1}$ (resp. $(Y_i)_{i \geq 1}$).

Denote by λ (resp. λ') the mean of N (resp. N'). Then, for any $n \geq 1$

$$\left\| \mathcal{L} \left(\sum_{i=1}^N X_i \right)^{\otimes n} - \mathcal{L} \left(\sum_{i=1}^{N'} Y_i \right)^{\otimes n} \right\|_{TV} \leq n(\lambda \wedge \lambda') \|X_1 - Y_1\|_{TV} + 1 - e^{-n|\lambda - \lambda'|}.$$

Proof. The proof is a minor extension of the proof of Theorem 13 in [14]. For the ease of the reader we repeat here the essential steps. Without loss of generality, let us suppose that $\lambda \geq \lambda'$ and write $\lambda = \alpha + \lambda'$, $\alpha \geq 0$. By triangle inequality,

$$\begin{aligned} \left\| \mathcal{L} \left(\sum_{i=1}^N X_i \right)^{\otimes n} - \mathcal{L} \left(\sum_{i=1}^{N'} Y_i \right)^{\otimes n} \right\|_{TV} &\leq \left\| \mathcal{L} \left(\sum_{i=1}^N X_i \right)^{\otimes n} - \mathcal{L} \left(\sum_{i=1}^{N''} X_i \right)^{\otimes n} \right\|_{TV} \\ &\quad + \left\| \mathcal{L} \left(\sum_{i=1}^{N''} X_i \right)^{\otimes n} - \mathcal{L} \left(\sum_{i=1}^{N'} Y_i \right)^{\otimes n} \right\|_{TV}, \end{aligned} \quad (53)$$

where N'' is a random variable independent of $(X_i)_{i \geq 1}$ and with the same law as N' .

Let P be a Poisson random variable independent of N'' and $(X_i)_{i \geq 1}$ with mean α . Then,

$$\begin{aligned} \left\| \mathcal{L} \left(\sum_{i=1}^N X_i \right)^{\otimes n} - \mathcal{L} \left(\sum_{i=1}^{N''} X_i \right)^{\otimes n} \right\|_{TV} &= \left\| \mathcal{L} \left(\sum_{i=1}^{N''+P} X_i \right)^{\otimes n} - \mathcal{L} \left(\sum_{i=1}^{N''} X_i \right)^{\otimes n} \right\|_{TV} \\ &\leq \left\| \delta_0 - \mathcal{L} \left(\sum_{i=1}^P X_i \right)^{\otimes n} \right\|_{TV} \\ &= \mathbb{P} \left(\left(\sum_{i=1}^P X_i \right)^{\otimes n} \neq (0, \dots, 0) \right) \leq 1 - e^{-n\alpha}, \end{aligned} \quad (54)$$

where (54) follows by Lemma 12.

In order to bound the second addendum in (53) we condition on N' and use again Lemma 12 joined with the fact that $\mathcal{L}(N') = \mathcal{L}(N'')$. Denoting by N'_1, \dots, N'_n n independent copies of N' and by $X_{i,j}$ (resp. $Y_{i,j}$), $j = 1, \dots, n$ and $i \geq 1$, n independent copies of X_i (resp. Y_i), we have:

$$\begin{aligned} &\left\| \mathcal{L} \left(\sum_{i=1}^{N''} X_i \right)^{\otimes n} - \mathcal{L} \left(\sum_{i=1}^{N'} Y_i \right)^{\otimes n} \right\|_{TV} \\ &= \sum_{m_1 \geq 0, \dots, m_n \geq 0} \left\| \mathcal{L} \left(\sum_{i=1}^{m_1} X_{i,1}, \dots, \sum_{i=1}^{m_n} X_{i,n} \right) - \mathcal{L} \left(\sum_{i=1}^{m_1} Y_{i,1}, \dots, \sum_{i=1}^{m_n} Y_{i,n} \right) \right\|_{TV} \mathbb{P}(N'_1 = m_1) \dots \mathbb{P}(N'_n = m_n) \\ &\leq \sum_{m_1 \geq 0, \dots, m_n \geq 0} (m_1 \|X_{1,1} - Y_{1,1}\|_{TV} + \dots + m_n \|X_{1,n} - Y_{1,n}\|_{TV}) \mathbb{P}(N'_1 = m_1) \dots \mathbb{P}(N'_n = m_n) \\ &= \|X_1 - Y_1\|_{TV} (\mathbb{E}[N'_1] + \dots + \mathbb{E}[N'_n]) = n\lambda' \|X_1 - Y_1\|_{TV}. \end{aligned}$$

□

Lemma 12. Let X, Y and Z be three random variables on \mathbb{R}^N , $N \geq 1$, with Z independent of X and Y . Then,

$$\|\mathcal{L}(X + Z) - \mathcal{L}(Y + Z)\|_{TV} \leq \|X - Y\|_{TV}. \quad (55)$$

Let $(X_i)_{i=1}^m$ and $(Y_i)_{i=1}^m$ be independent random variables on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$, $N \geq 1$. Then, for any $m \geq 1$,

$$\left\| \sum_{i=1}^m X_i - \sum_{i=1}^m Y_i \right\|_{TV} \leq \sum_{i=1}^m \|X_i - Y_i\|_{TV}. \quad (56)$$

Proof. To prove (55) one can use the variational definition of the total variation. Denoting by P_A the law of a random variable A , it holds

$$\begin{aligned} 2\|\mathcal{L}(X+Z) - \mathcal{L}(Y+Z)\|_{TV} &= \sup_{\|f\|_\infty \leq 1} \int f(x)(P_{X+Z}(dx) - P_{Y+Z}(dx)) \\ &= \sup_{\|f\|_\infty \leq 1} \int f(x) \left(\int P_Z(dx-z)(P_X(dz) - P_Y(dz)) \right) \\ &= \sup_{\|f\|_\infty \leq 1} \int \int f(x)P_Z(dx-z)(P_X(dz) - P_Y(dz)). \end{aligned}$$

Denote by $g_f(z) = \int f(x)P_Z(dx-z)$ and observe that g_f is measurable with $\|g_f\|_\infty \leq \|f\|_\infty$. It then follows that

$$\sup_{\|f\|_\infty \leq 1} \int g_f(z)(P_X(dz) - P_Y(dz)) \leq \sup_{\|g\|_\infty \leq 1} \int g(z)(P_X(dz) - P_Y(dz)) = 2\|X - Y\|_{TV}.$$

Equation (56) is straightforward using (55). Indeed, by induction, it suffices to prove the case $n = 2$. Let \tilde{X}_2 be a random variable equal in law to X_2 and independent of Y_1 and of X_1 . By triangle inequality we deduce that

$$\|\mathcal{L}(X_1 + X_2) - \mathcal{L}(Y_1 + Y_2)\|_{TV} \leq \|\mathcal{L}(X_1 + X_2) - \mathcal{L}(Y_1 + \tilde{X}_2)\|_{TV} + \|\mathcal{L}(Y_1 + \tilde{X}_2) - \mathcal{L}(Y_1 + Y_2)\|_{TV}$$

and by means of (55) we conclude that

$$\begin{aligned} \|\mathcal{L}(Y_1 + \tilde{X}_2) - \mathcal{L}(Y_1 + Y_2)\|_{TV} &\leq \|\tilde{X}_2 - Y_2\|_{TV} = \|X_2 - Y_2\|_{TV} \\ \|\mathcal{L}(X_1 + X_2) - \mathcal{L}(Y_1 + \tilde{X}_2)\|_{TV} &\leq \|X_1 - Y_1\|_{TV}. \end{aligned}$$

□

Lemma 13. *Let P and Q be probability density. For any $n \geq 1$ it holds:*

$$\|P^{\otimes n} - Q^{\otimes n}\|_{TV} \leq \sqrt{2n\|P - Q\|_{TV}}.$$

Proof. Let $H(P, Q)$ denote the Hellinger distance between P and Q , i.e.

$$H(P, Q) = \sqrt{\int \left(\sqrt{\frac{dP}{d\mu}} - \sqrt{\frac{dQ}{d\mu}} \right)^2 d\mu}$$

where μ it is a common dominating measure for P and Q . It is well known (see e.g. Lemma 2.3 and Property (iv) page 83 in [18]) that

$$\frac{H^2(P, Q)}{2} \leq \|P - Q\|_{TV} \leq H(P, Q), \quad H^2(P^{\otimes n}, Q^{\otimes n}) = 2 \left(1 - \left(1 - \frac{H^2(P, Q)}{2} \right)^n \right).$$

In particular $H^2(P^{\otimes n}, Q^{\otimes n}) \leq nH^2(P, Q)$ and thus

$$\|P^{\otimes n} - Q^{\otimes n}\|_{TV} \leq \sqrt{nH^2(P, Q)} \leq \sqrt{2n\|P - Q\|_{TV}},$$

as desired.

□