Statistics and high frequency data

Jean Jacod*

1 Introduction

This short course is devoted to a few statistical problems related to the observation of a given process on a fixed time interval, when the observations occur at regularly spaced discrete times. This kind of observations may occur in many different contexts, but they are particularly relevant in finance: we do have now huge amounts of data on the prices of various assets, exchange rates, and so on, typically "tick data" which are recorded at every transaction time. So we are mainly concerned with the problems which arise in this context, and the concrete applications we will give are all pertaining to finance.

In some sense they are not "standard" statistical problems, for which we want to estimate some unknown parameter. We are rather concerned with the "estimation" of some random quantities. This means that we would like to have procedures that are as model-free as possible, and also that they are in some sense more akin to nonparametric statistics.

Let us describe the general setting in some more details. We have an underlying process $X = (X_t)_{t\geq 0}$, which may be multi-dimensional (its components are then denoted by $X^1, X^2, \cdots$). This process is defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We observe this process at discrete times, equally spaced, over some fixed finite interval $[0, T]$, and we are concerned with asymptotic properties as the time lag, denoted by Δ_n, goes to 0. In practice, this means that we are in the context of *high frequency data*.

The objects of interest are various quantities related to the particular outcome ω which is (partially) observed. The main object is the *volatility*, but other quantities or features are also of much interest for modeling purposes, for example whether the observed path has jumps and, when this is the case, whether several components may jump at the same times or not.

All these quantities are related in some way to the probabilistic model which is assumed for X: we do indeed need some model assumption, otherwise nothing can be said. In fact, any given set of observed values $X_0, X_{\Delta_n}, \cdots, X_{i\Delta_n}, \cdots$, with Δ_n fixed, is of course compatible with many different models for the continuous time process X: for example we can suppose that X is piecewise constant between the observation times, or that it

*Institut de mathématiques de Jussieu, Université Pierre et Marie Curie (Paris-6) and CNRS, UMR 7586, 4 place Jussieu, 75252 Paris, France

is piecewise linear between these times. Of course neither one of these two models is in general compatible with the observations if we modify the frequency of the observations.

So in the sequel we will always assume that X is an Itô semimartingale, that is a semimartingale whose characteristics are absolutely continuous with respect to Lebesgue measure. This is compatible with virtually all semimartingale models used for modeling quantities like asset prices or log-prices, although it rules out some non-semimartingale models sometimes used in this context, like the fractional Brownian motion.

Before stating more precisely the questions which we will consider, and in order to be able to formulate them in precise terms, we recall the structure of Itô semimartingales. We refer to [13], Chapter I, for more details.

Semimartingales: We start with a basic filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, the family of sub-$\sigma$-fields $(\mathcal{F}_t)$ of $\mathcal{F}$ being increasing and right-continuous in t. A semimartingale is simply the sum of a local martingale on this space, plus an adapted process of finite variation (meaning, its paths are right-continuous, with finite variation on any finite interval). In the multidimensional case it means that each component is a real-valued semimartingale.

Any multidimensional semimartingale can be written as

$$X_t = X_0 + B_t + X_t^c + \int_0^t \int_{\mathbb{R}^d} \kappa(x)(\mu - \nu)(ds, dx) + \int_0^t \int_{\mathbb{R}^d} \kappa'(x)\mu(ds, dx). \tag{1.1}$$

In this formula we use the following notation:

- μ is the "jump measure" of X: if we denote by $\Delta X_t = X_t - X_{t-}$ the size of the jump of X at time t (recall that X is right-continuous with left limits), then the set $\{t : \Delta X_t(\omega) \neq 0\}$ is at most countable for each ω, and μ is the random measure on $(0, \infty) \times \mathbb{R}^d$ defined by

$$\mu(\omega; dt, dx) = \sum_{s>0:\ \Delta X_s(\omega)\neq 0} \varepsilon_{(s, \Delta X_s(\omega))}(dt, dx), \quad \varepsilon_a \ = \ \text{the Dirac measure sitting at } a.$$

- ν is the "compensator" (or, predictable compensator) of μ. This is the unique random measure on $(0, \infty) \times \mathbb{R}^d$ such that, for any Borel subset A of $\mathbb{R}^d$ at a positive distance of 0, the process $\nu((0, t] \times A)$ is predictable and the difference $\mu((0, t] \times A) - \nu((0, t] \times A)$ is a local martingale.

- κ is a "truncation function", that is a function: $\mathbb{R}^d \to \mathbb{R}^d$, bounded with compact support, such that $\kappa(x) = x$ for all x in a neighborhood of 0. This function is fixed throughout, and we choose it to be *continuous* for convenience.

- κ' is the function $\kappa'(x) = x - \kappa(x)$.

- B is a predictable process of finite variation, with $B_0 = 0$.

- X^c is a continuous local martingale with $X_0^c = 0$, called the "continuous martingale part" of X.

With this notation, the decomposition (1.1) is unique (up to null sets), but the process B depends on the choice of the truncation function κ. The continuous martingale part does *not* depend on the choice of κ. Note that the first integral in (1.1) is a stochastic integral (in general), whereas the second one is a pathwise integral (in fact for any t is is simply the finite sum $\sum_{s\leq t} \kappa'(\Delta X_s)$). Of course (1.1) should be read "componentwise" in the multidimensional setting.

In the sequel we use the shorthand notation $\star$ to denote the (possibly stochastic) integral w.r.t. a random measure, and also $\bullet$ for the (possibly stochastic) integral of a process w.r.t. a semimartingale. For example, (1.1) may be written more shortly as

$$X \;=\; X_0 + B + X^c + \kappa \star (\mu - \nu) + \kappa' \star \mu. \tag{1.2}$$

The "*" symbol will also be used, as a superscript, to denote the transpose of a vector or matrix (no confusion may arise).

Another process is of great interest, namely the quadratic variation of the continuous martingale part X^c, which is the following $\mathbb{R}^d \otimes \mathbb{R}^d$-valued process:

$$C \;=\; \langle X^c, X^{c\star}\rangle, \quad \text{that is, componentwise,} \;\; C^{ij} = \langle X^{i,c}, X^{j,c}\rangle. \tag{1.3}$$

This is a continuous adapted process with $C_0 = 0$, which further is increasing in the set $\mathcal{M}_d^+$ of symmetric nonnegative matrices, that is $C_t - C_s$ belongs to $\mathcal{M}_d^+$ for all $t > s$.

The triple (B, C, ν) is called the *triple of characteristics* of X, this name coming from the fact that in "good cases" it completely determines the law of X.

The fundamental example of semimartingales is the case of Lévy processes. We say that X is a *Lévy process* if it is adapted to the filtration, with right-continuous and left-limited paths and $X_0 = 0$, and such that $X_{t+s} - X_t$ is independent of $\mathcal{F}_t$ and has the same law as X_s for all $s, t \geq 0$. Such a process is always a semimartingale, and its characteristics (B, C, ν) are of the form

$$B_t(\omega) = bt, \quad C_t = ct, \quad \nu(\omega; dt, dx) = dt \otimes F(dx). \tag{1.4}$$

Here $b \in \mathbb{R}^d$ and $c \in \mathcal{M}_d^+$ and F is a measure on $\mathbb{R}^d$ which does not charge 0 and integrates the function $x \mapsto \|x\|^2 \wedge 1$. The triple (b, c, F) is connected with the law of the variables X_t by the formula (for all $u \in \mathbb{R}^d$)

$$\mathbb{E}(e^{i\langle u, X_t\rangle}) = \exp t\Big(i\langle u, b\rangle - \frac{1}{2}\langle u, cu\rangle + \int F(dx)\Big(e^{i\langle u,x\rangle} - 1 - i\langle u, \kappa(x)\rangle\Big)\Big), \tag{1.5}$$

called Lévy-Khintchine's formula. So we sometimes call (b, c, F) the characteristics of X as well, and it is the Lévy-Khintchine characteristics of the law of X_1 in the context of infinitely divisible distributions. b is called the drift, c is the covariance matrix of the Gaussian part, and F is called the Lévy measure.

As seen above, for a Lévy process the characteristics (B, C, ν) are deterministic, and they do characterize the law of the process. Conversely, if the characteristics of a semimartingale X are deterministic one can show that X has independent increments, and if they are of the form (1.4) then X is a Lévy process.

Itô semimartingales. By definition, an *Itô semimartingale* is a semimartingale whose characteristics (B, C, ν) are absolutely continuous with respect to Lebesgue measure, in the following sense:

$$B_t(\omega) = \int_0^t b_s(\omega)ds, \quad C_t(\omega) = \int_0^t c_s(\omega)ds, \quad \nu(\omega; dt, dx) = dt\ F_{\omega,t}(dx). \tag{1.6}$$

here we can always choose a version of the processes b or c which is optional, or even predictable, and likewise choose F in such a way that $F_t(A)$ is optional, or even predictable, for all Borel subsets A of $\mathbb{R}^d$.

It turns out that Itô semimartingales have a nice representation in terms of a Wiener process and a Poisson random measure, and this representation will be very useful for us. Namely, it can be written as follows (where for example $\kappa'(\delta) * \underline{\mu}_t$ denotes the value at time t of the integral process $\kappa'(\delta) * \underline{\mu}$):

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \kappa(\delta) \star (\underline{\mu} - \underline{\nu})_t + \kappa'(\delta) \star \underline{\mu}_t. \tag{1.7}$$

In this formula W is a standard d'-dimensional Wiener process and $\underline{\mu}$ is a Poisson random measure on $(0,\infty) \times E$ with intensity measure $\underline{\nu}(dt, dx) = dt \otimes \lambda(dx)$, where λ is a σ-finite and infinite measure without atom on an auxiliary measurable set $(E, \mathcal{E})$.

Of course the process b_t is the same in (1.6) and in (1.7), and $\sigma = (\sigma^{ij})_{1\leq i\leq d, 1\leq j\leq d'}$ is an $\mathbb{R}^d \otimes \mathbb{R}^{d'}$-valued optional (or predictable, as one wishes to) process such that $c = \sigma\sigma^\star$, and $\delta = \delta(\omega, t, x)$ is a predictable function on $\Omega \times [0,\infty) \times E$ (that is, measurable with respect to $\mathcal{P} \otimes \mathcal{E}$, where $\mathcal{P}$ is the predictable σ-field of $\Omega \times [0,\infty)$). The connection between δ above and F in (1.6) is that $F_{t,\omega}$ is the image of the measure λ by the map $x \mapsto \delta(\omega, t, x)$, and restricted to $\mathbb{R}^d \backslash \{0\}$.

Remark 1.1 One should be a bit more precise in characterizing W and $\underline{\mu}$: W is an $(\mathcal{F}_t)$-Wiener process, meaning it is $\mathcal{F}_t$ adapted and $W_{t+s} - W_t$ is independent of $\mathcal{F}_t$ (on top of being Wiener, of course). Likewise, $\underline{\mu}$ is an $(\mathcal{F}_t)$-Poisson measure, meaning that $\underline{\mu}((0,t] \times A)$ is $\mathcal{F}_t$-measurable and $\underline{\mu}((t, t+s] \times A)$ is independent of $\mathcal{F}_t$, for all $A \in \mathcal{E}$. □

Remark 1.2 The original space $(\Omega, \mathcal{F}, \mathbb{P})$ on which X is defined may be too small to accommodate a Wiener process and a Poisson measure, so we may have to enlarge the space. Such an enlargement is always possible. □

Remark 1.3 When the matrix $c_t(\omega)$ is of full rank for all (ω, t) and $d' = d$, then it has a unique "square-root" $\sigma_t(\omega)$, which further is invertible. In this case we have $W = (\sigma)^{-1} \bullet X^c$. Otherwise, there are many ways of choosing σ such that $\sigma\sigma^\star = c$, hence many ways of choosing W and its dimension d' (which can always be taken such that $d' \leq d$).

In a similar way, we have a lot of freedom for the choice of $\underline{\mu}$. In particular we can choose at will the space $(E, \mathcal{E})$ and the measure λ, subject to the above conditions, and for example we can always take $E = \mathbb{R}$ with λ the Lebesgue measure, although in the d-dimensional case it is somewhat more intuitive to take $E = \mathbb{R}^d$. □

Of course a Lévy process is an Itô semimartingale (compare (1.2) and (1.6)). In this case the two representations (1.2) and (1.7) coincide if we take $E = \mathbb{R}^d$ and $\lambda = F$ (the

Lévy measure) and $\underline{\mu} = \mu$ (the jump measure of a Lévy process is a Poisson measure) and $\delta(\omega, t, x) = x$, and also if we recall that in this case the continuous martingale (or "Gaussian") part of X is always of the form $X^c = \sigma W$, with $\sigma\sigma^\star = c$.

The setting of Itô semimartingales encompasses most processes used for modeling purposes, at least in mathematical finance. For example, solutions of stochastic differential equations driven by a Wiener process, or a by a Lévy process, or by a Wiener process plus a Poisson random measure, are all Itô semimartingales. Such solutions are obtained directly in the form (1.7), which of course implies that X is an Itô semimartingale.

The volatility. In a financial context, the process c_t is called the volatility (sometimes it is σ_t which is thus called). This is by far the most important quantity which needs to be estimated, and there are many ways to do so. A very widely spread way of doing so consists in using the so-called "implied volatility", and it is performed by using the observed current prices of options drawn on the stock under consideration, by somehow inverting the Black-Scholes equation or extensions of it.

However, this way usually assumes a given type of models, for example that the stock prices is a diffusion process of a certain type, with unknown coefficients. Among the coefficients there is the volatility, which further may be "stochastic", meaning that it depends on some random inputs other than the Wiener process which drives the price itself. But then it is of primary importance to have a sound model, and this can be checked only by statistical means. That is, we have to make a statistical analysis, based on series of (necessarily discrete) observations of the prices.

In other words, there is a large body of work, essentially in the econometrical literature, about the (statistical) estimation of the volatility. This means finding good methods for estimating the path $t \mapsto c_t(\omega)$ for $t \in [0, T]$, on the basis of the observation of $X_{i\Delta_n}(\omega)$ for all $i = 0, 1, \cdots, [T/\Delta_n]$.

In a sense this is very similar to the non-parametric estimation of a function $c(t)$, say in the 1-dimensional case, when one observes the Gaussian process

$$Y_t = \int_0^t \sqrt{c(s)}\, dW_s$$

(here W is a standard 1-dimensional Wiener process) at the time $i\Delta_n$, and when Δ_n is "small" (that is, we consider the asymptotic $\Delta_n \to 0$). As is well known, this is possible only under some regularity assumptions on the function $c(t)$, whereas the "integrated" value $\int_0^t c(s)ds$ can be estimated as in parametric statistics, since it is just a number. On the other hand, if we know $\int_0^t c(s)ds$ for all t, then we also know the function $c(t)$, up to a Lebesgue-null set, of course: it should be emphasized that if we modify c on such a null set, we do not change the process Y itself; the same comment applies to the volatility process c_t in (1.6).

This is why we mainly consider, as in most of the literature, the problem of estimating the *integrated volatility*, which with our notation is the process C_t. One has to be aware of the fact that in the case of a general Itô semimartingale, this means "estimating" the random number or matrix $C_t(\omega)$, for the observed ω, although of course ω is indeed not "fully" observed.

Let us consider for simplicity the 1-dimensional case, when further X is continuous, that is

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s, \tag{1.8}$$

and σ_t (equivalently, $c_t = \sigma_t^2$) is random. It may be of the form $\sigma_t(\omega) = \sigma(X_t(\omega))$, it can also be by itself the solution of another stochastic differential equation, driven by W and perhaps another Wiener process W', and perhaps also some Poisson measures if it has jumps (even though X itself does not jump).

By far, the simplest thing to do is to consider the "realized" integrated volatility, or "approximate quadratic variation", that is the process

$$B(2, \Delta_n)_t = \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^2, \qquad \text{where } \Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}. \tag{1.9}$$

Then if (1.8) holds, well known results on the quadratic variation (going back to Itô in this case), we know that

$$B(2, \Delta_n)_t \stackrel{\mathbb{P}}{\longrightarrow} C_t \tag{1.10}$$

(convergence in probability), and this convergence is even uniform in t over finite intervals. Further, as we will see later, we have a rate of convergence (namely $1/\sqrt{\Delta_n}$) under some appropriate assumptions.

Now what happens when X is discontinuous ? We no longer have (1.10), but rather

$$B(2, \Delta_n)_t \stackrel{\mathbb{P}}{\longrightarrow} C_t + \sum_{s \le t} |\Delta X_s|^2 \tag{1.11}$$

(the right side above is always finite, and is the quadratic variation of the semimartingale X, also denoted $[X, X]_t$). Nevertheless we do want to estimate C_t: a good part of these notes is devoted to this problem. For example, we will show that both quantities

$$B(1, 1, \Delta_n)_t = \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X||\Delta_{i+1}^n X|, \qquad B(2, \varpi, \alpha)_t = \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^2 1_{\{|\Delta_i^n X| \le \alpha \Delta_n^{\varpi}\}} \tag{1.12}$$

converge in probability to $\frac{2}{\pi} C_t$ and C_t respectively, and as soon as $\varpi \in (0, 1/2)$ and $\alpha > 0$ for the second one.

Inference for jumps. Now, when X is discontinuous, there is also a lot of interest about jumps and, to begin with, are the observations compatible with a model without jumps, or should we use a model with jumps ? More complex questions may be posed: for a 2-dimensional process, do the jumps occur at the same times for the two components or not ? Is there infinitely many (small) jumps ? In this case, what is the "concentration" of the jumps near 0 ?

Here again, the analysis is based on the asymptotic behavior of quantities involving sums of functions of the increments $\Delta_i^n X$ of the observed process. So, before going to the main results in a general situation, we consider first two very simple cases: when $X = \sigma W$

for a constant $\sigma > 0$, and when $X = \sigma W + Y$ when Y is a compound Poisson process. It is also of primary importance to determine which quantities can be "consistently estimated" when $\Delta_n \to 0$, and which ones cannot be. We begin with the latter question.

2 What can be estimated ?

Recall that our underlying process X is observed at discrete times $0, \Delta_n, 2\Delta_n, \cdots$, up to some fixed time T. Obviously, we cannot have consistent estimators, as $\Delta_n \to 0$, for quantities which cannot be retrieved when we observe the whole path $t \mapsto X_t(\omega)$ for $t \in [0, T]$, a situation referred to below as the "complete observation scheme".

We begin with two simple observations:

1) The drift b_t can *never* be identified in the complete observation scheme, except in some very special cases, like when $X_t = X_0 + \int_0^t b_s ds$.

2) The quadratic variation of the process is fully known in the complete observation scheme, up to time T of course. This implies in particular that the integrated volatility C_t is known for all $t \leq T$, hence also the process c_t (this is of course up to a $\mathbb{P}$-null set for C_t, and a $\mathbb{P}(d\omega) \otimes dt$-null set for $c_t(\omega)$).

3) The jumps are fully known in the complete observation scheme, up to time T again.

Now, the jumps are not so interesting by themselves. More important is the "law" of the jumps in some sense. For Lévy processes the law of jumps is in fact determined by the Lévy measure. In a similar way, for a semimartingale the law of jumps can be considered as known if we know the measures $F_{t,\omega}$, since these measures specify the jump coefficient δ in (1.7). (Warning: this specification is in a "weak" sense, exactly as c specifies σ; we may have several square-root of c, as well as several δ such that F_t is the image of λ, but all choices of σ_t and δ which are compatible with a given c_t and F_t give rise to equations that have exactly the same weak solutions).

Consider Lévy processes first. Basically, the restriction of F to the complement of any neighborhood of 0, after normalization, is the law of the jumps of X lying outside this neighborhood. Hence to consistently estimate F we need potentially infinitely many jumps far from 0, and this possible only if $T \to \infty$. In our situation with T fixed there is no way of consistently estimating F.

We can still say something in the Lévy case: for the complete observation scheme, if there is a jump then F is not the zero measure; if we have infinitely many jumps in $[0, T]$ then F is an infinite measure; in this case, we can also determine for which $r > 0$ the sum $\sum_{s \leq T} |\Delta X_s|^r$ is finite, and this is also the set of r's such that $\int_{\{|x|<\infty\}} |x|^r F(dx) < \infty$.

The same statements also hold for more general semimartingales: we can decide for which r's the sum $\sum_{s \leq T} |\Delta X_s|^r$ is finite, and also if we have zero, or finitely many, or infinitely many jumps. Those are "characteristics" of the model which are of much interest for modelling purposes.

Hence we will be interested, when coming back to the actual discrete observation scheme, in estimating C_t for $t \leq T$, and whether there are zero or finitely many or infinitely many jumps in $[0,T]$.

3 Some simple limit theorems for Wiener plus compound Poisson processes

This section is about a very particular case: the underlying process is $X = \sigma X + Y$ for some $\sigma > 0$, and Y a compound Poisson process independent of W. And in the first subsection we even consider the most elementary case of $X = \sigma W$. In these two cases we state all limit theorems that are available about sums of a function of the increments. We do not give the full proofs, but heuristic reasons for the results to be true. The reason for devoting a special section to this simple case is to show the variety of results that can be obtained, whereas the full proofs can be easily reconstructed without annoying technical details.

Before getting started, we introduce some notation, to be used also for a general d-dimensional semimartingale X later on. Recall the increments $\Delta_i^n X$ in (1.9). First for any $p > 0$ and $j \leq d$ we set

$$B(p,j,\Delta_n)_t \;=\; \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X^j|^p. \tag{3.1}$$

In the 1-dimensional case this is written simply $B(p,\Delta_n)_t$. Next if f is a function on $\mathbb{R}^d$, the state space of X in general, we set

$$\left.\begin{array}{lcl} V(f,\Delta_n)_t &=& \sum_{i=1}^{[t/\Delta_n]} f(\Delta_i^n X), \\ V'(f,\Delta_n)_t &=& \sum_{i=1}^{[t/\Delta_n]} f(\Delta_i^n X/\sqrt{\Delta_n}). \end{array}\right\} \tag{3.2}$$

The reason for introducing the normalization $1/\sqrt{\Delta_n}$ will be clear below. These functionals are related one of the other by the trivial identity $V'(f,\Delta_n) = V(f_n,\Delta_n)$ with $f_n(x) = f(x/\sqrt{\Delta_n})$. Moreover, with the notation

$$y \in \mathbb{R} \;\mapsto\; h_p(y) \;=\; |y|^p, \qquad x = (x_j) \in \mathbb{R}^d \;\mapsto h_p^j(x) \;=\; |x_j|^p, \tag{3.3}$$

we also have $B(p,j,\Delta_n) = V(h_p^j,\Delta_n) = \Delta_n^{-p/2} V'(h_p^j,\Delta_n)$. Finally if we need to emphasize the dependency on the process X, we write these functionals as $B(X;p,j,\Delta_n)$ or $V(X;f,\Delta_n)$ or $V'(X;f,\Delta_n)$.

3.1 The Wiener case.

Here we suppose that $X = \sigma W$ for some constant $\sigma > 0$, so $d = 1$. Among all the previous functionals, the simplest ones to study are the functionals $V'(f,\Delta_n)$ with f a fixed function on $\mathbb{R}$. We need f to be Borel, of course, and "not too big", for example with polynomial growth, or even with exponential growth. In this case, the results are straightforward consequences of the usual law of large numbers (LNN) and central limit theorem (CLT).

Indeed, for any n the variables $(\Delta_i^n X/\sqrt{\Delta_n} : i \geq 1)$ are i.i.d. with law $\mathcal{N}(0,\sigma^2)$. In the formulas below we write ρ_σ for the law $\mathcal{N}(0,\sigma^2)$ and also $\rho_\sigma(g)$ the integral of a function g with respect to it. Therefore, with f as above, the variables $f(\Delta_i^n X/\sqrt{\Delta_n})$ when i varies are i.i.d. with moments of all orders, and their first and second moments equal $\rho_\sigma(f)$ and $\rho_\sigma(f^2)$ respectively. Then the classical LLN and CLT give us that

$$\left.\begin{array}{l} \Delta_n\ V'(f,\Delta_n)_t \ \stackrel{\mathbb{P}}{\longrightarrow}\ t\rho_\sigma(f) \\ \frac{1}{\sqrt{\Delta_n}}\Big(\Delta_n\ V'(f,\Delta_n)_t - t\rho_\sigma(g)\Big) \ \stackrel{\mathcal{L}}{\longrightarrow}\ \mathcal{N}\Big(0, t(\rho_\sigma(f^2)-\rho_\sigma(f)^2)\Big). \end{array}\right\} \tag{3.4}$$

We clearly see here why we have put the normalizing factor $1/\sqrt{\Delta_n}$ inside the function f.

The reader will observe that, contrary to the usual LNN, we get convergence in probability but *not* almost surely in the first part of (3.4). The reason is as follows: let ζ_i be a sequence of i.i.d. variables with the same law than $f(X_1)$. The LLN implies that $Z_n = \frac{t}{[t/\Delta_n]}\sum_{i=1}^{[t/\Delta_n]}\zeta_i$ converges a.s. to $t\rho_\sigma(f)$. Since $\Delta_n V'(f,\Delta_n)_t$ has the same law as Z_n we deduce the convergence in probability in (3.4) because, for a deterministic limit, convergence in probability and convergence in law are equivalent. However the variables $V'(f,\Delta_n)_t$ are connected one with the others in a way we do not really control when n varies, so we cannot conclude to $\Delta_n V'(f;\Delta_n)_t \to t\rho_\sigma(f)$ a.s.

(1.9) gives us the convergence for any time t, but we also have functional convergence:

1) First, recall that a sequence g_n of nonnegative increasing functions on $\mathbb{R}_+$ converging pointwise to a *continuous* function g also converges locally uniformly; then, from the first part of (1.9) applied separately for the positive and negative parts f^+ and f^- of f and using a "subsequence principle" for the convergence in probability, we obtain

$$\Delta_n\ V'(f,\Delta_n)_t \ \stackrel{\text{u.c.p.}}{\longrightarrow}\ t\rho_\sigma(f) \tag{3.5}$$

where $Z_t^n \stackrel{\text{u.c.p.}}{\longrightarrow} Z_t$ means "convergence in probability, locally uniformly in time": that is, $\sup_{s\leq t}|Z_s^n - Z_s| \stackrel{\mathbb{P}}{\longrightarrow} 0$ for all t finite.

2) Next, if instead of the 1-dimensional CLT we use the "functional CLT", or Donsker's Theorem, we obtain

$$\Big(\frac{1}{\sqrt{\Delta_n}}\Big(\Delta_n V'(f,\Delta_n)_t - t\rho_\sigma(f)\Big)\Big)_{t\geq 0} \ \stackrel{\mathcal{L}}{\Longrightarrow}\ \sqrt{\rho_\sigma(f^2)-\rho_s(f)^2}\ W' \tag{3.6}$$

where W' is another standard Wiener process, and $\stackrel{\mathcal{L}}{\Longrightarrow}$ stands for the convergence in law of processes (for the Skorokhod topology). Here we see a new Wiener process W' appear. What is its connection with the basic underlying Wiener process W ? To study that, one can try to prove the "joint convergence" of the processes on the left side of (3.6) together with W (or equivalently X) itself.

This is an easy task: consider the 2-dimensional process Z^n whose first component is the left side of (3.6) and second component is $X_{\Delta_n[t/\Delta_n]}$ (the discretized version of

X, which converges pointwise to X). Then Z^n takes the form $Z^n_t = \sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]} \zeta^n_i$, where the ζ^n_i are 2-dimensional i.i.d. variables as i varies, with the same distribution as $(g_1(X_1), g_2(X_1))$, where $g_1(x) = f(x) - \rho_\sigma(f)$ and $g_2(x) = x$. Then the 2-dimensional version of Donsker's Theorem gives us that

$$\Big(\frac{1}{\sqrt{\Delta_n}}\Big(\Delta_n V'(f;\Delta_n)_t - t\rho_\sigma(f)\Big), X_t\Big)_{t\geq 0} \stackrel{\mathcal{L}}{\Longrightarrow} \Big(B, X\Big) \tag{3.7}$$

and the pair (B, X) is a 2-dimensional (correlated) Wiener process, characterized by its variance-covariance at time 1, which is the following matrix:

$$\begin{pmatrix} \rho_\sigma(f^2) - \rho_s(f)^2 & \rho_\sigma(fg_2) \\ \rho_\sigma(fg_2) & \sigma^2 \end{pmatrix} \tag{3.8}$$

(note that $\sigma^2 = \rho_\sigma(g_2^2)$ and also $\rho_\sigma(g_2) = 0$, so the above matrix is semi-definite positive). Equivalently, we can write B as $B = \sqrt{\rho_\sigma(f^2) - \rho_s(f)^2}\; W'$ with W' a standard Brownian motion (as in (3.7))) which is correlated with W, the correlation coefficient being $\rho_\sigma(fg_2)/\sigma\sqrt{\rho_\sigma(f^2) - \rho_s(f)^2}$.

Now we turn to the processes $B(p,\Delta_n)$. Since $B(p,\Delta_n) = \Delta_n^{-p/2} V'(h_p,\Delta_n)$ this is just a particular case of (3.5) and (3.7), which we reformulate below (m_p denotes the pth absolute moment of the normal law $\mathcal{N}(0,1)$):

$$\Delta_n^{1-p/2}\; B(p,\Delta_n) \stackrel{\text{u.c.p.}}{\longrightarrow} t\sigma^p m_p, \tag{3.9}$$

$$\left.\begin{array}{l} \Big(\frac{1}{\sqrt{\Delta_n}}\Big(\Delta_n^{1-p/2} B(p,\Delta_n)_t - t\sigma^p m_p\Big), X_t\Big)_{t\geq 0} \stackrel{\mathcal{L}}{\Longrightarrow} \Big(B, X\Big), \\ \text{with } B \text{ a Wiener process unit variance } \sigma^{2p}(m_{2p} - m_p^2), \text{ independent of } X \end{array}\right\} \tag{3.10}$$

(the independence comes from that fact that $\rho_\sigma(g) = 0$, where $g(x) = x|x|^p$).

Finally for the functionals $V(f,\Delta_n)$, the important thing is the behavior of f near 0, since the increments $\Delta^n_i X$ are all going to 0 as $\Delta_n \to 0$. In fact, $\sup_{i\leq[t/\Delta_n]} |\Delta^n_i X| \to 0$ pointwise, so when the function f vanishes on a neighborhood of 0, for all n bigger than some (random) finite number N depending also on t we have

$$V(f,\Delta_n)_s = 0 \qquad \forall s \leq t. \tag{3.11}$$

For a "general" function f we can combine (3.9) with (3.11): we easily obtain that (3.9) holds with $V(f,\Delta_n)$ instead of $B(p,\Delta_n)$ as soon as $f(x) \sim |x|^p$ as $x \to 0$, and the same holds for (3.10) if we further have $f(x) = |x|^p$ on a neighborhood of 0.

Of course these results do not exhaust all possibilities for the convergence of $V(f;\Delta_n)$. For example on may prove the following:

$$f(x) = |x|^p \log|x| \quad \Rightarrow \quad \frac{\Delta_n^{1-p/2}}{\log(1/\Delta_n)}\; V(f,\Delta_n) \stackrel{\text{u.c.p.}}{\longrightarrow} -\frac{1}{2}\; t\sigma^p m_p, \tag{3.12}$$

and a CLT is also available in this situation. Or, we could consider functions f which behave like x^p as $x \downarrow\downarrow 0$ and like $(-x)^{p'}$ as $x \uparrow\uparrow 0$, with $p \neq p'$. However, we essentially restrict our attention to functions behaving like h_p: for simplicity first, and since more general functions do not really occur in the applications we have in mind, and also because the extension to processes X more general than the Brownian motion is not easy for other functions.

3.2 The Wiener plus compound Poisson case.

Our second example is when the underlying process X has the form $X = \sigma W + Y$, where as before $\sigma > 0$ and W is a Brownian motion, and Y is a compound Poisson process independent of W. We will write $X' = \sigma W$. Recall that Y has the form

$$Y_t \;=\; \sum_{p\geq 1} \Phi_p 1_{\{T_p\leq t\}}, \tag{3.13}$$

where the T_p's are the successive arrival times of a Poisson process, say with parameter 1 (they are finite stopping times, positive, strictly increasing with p and going to ∞), and the Φ_p's are i.i.d. variables, independent of the T_p's, and with some law G. Note that in (3.13) the sum, for any given t, is actually a finite sum.

The processes $V'(f,\Delta_n)$, which were particularly easy to study when X was a Wiener process, are not so simple to analyze now. This is easy to understand: let us fix t; at stage n, we have $\Delta_i^n X = \Delta_i^n X'$ for all $i \leq [t/\Delta_n]$, except for those finitely many i's corresponding to an interval $((i-1)\Delta_n, i\Delta_n]$ containing at least one of the T_p's. Furthermore, all those exceptional intervals contain exactly one T_p, as soon as n is large enough (depending on (ω,t)). Therefore for n large we have

$$\left.\begin{array}{l} V'(f,\Delta_n)_t = V'(X';f,\Delta_n)_t + A^n_t, \quad \text{where} \\ A'^n_t = \sum_{i=1}^{[t/\Delta_n]} \sum_{p\geq 1} 1_{\{(i-1)\Delta_n < T_p \leq i\Delta_n\}} \Big(f((\Phi_p + \Delta_i^n X')\sqrt{\Delta_n}) - f(\Delta_i^n X'/\sqrt{\Delta_n}) \Big). \end{array}\right\} \tag{3.14}$$

The double sum in A'^n_t is indeed a finite sum, with as many non-zero entries as the number of T_p's less than $\Delta_n[t/\Delta_n]$.

Therefore the behavior of $V'(f,\Delta_n)$ depends in an essential way on the behavior of f near infinity. There are essentially two possibilities:

1) The function f is bounded, or more generally satisfies $|f(x)| \leq K(1+|x|^p)$ for some $p < 2$. Then $|A'^n_t|$ above is "essentially" smaller than $K\sum_{q:T_q\leq t}(1+|\Phi_p|^r\Delta_n^{-p/2})$ for some constant K, and thus $\Delta_n A'^n_t \to 0$. So obviously the convergence (3.5) holds.

If further $p < 1$ we even have $\sqrt{\Delta_n}\; A'^n_t \to 0$. Therefore (3.7) holds. Observe that in this situation, the presence of the jumps does *not* modify the results that held for the Brownian case; this will be the rule for more general processes X as well.

2) The function f is equivalent to $|x|^p$ at infinity, for some $p > 2$. Then in (3.14) the leading term becomes A'^n_t, which is approximately equal to $\Delta_n^{-p/2}\sum_{s\leq t}|\Delta X_s|^p$. So $\Delta_n^{p/2} V'(f,\Delta_n)_t$ converges in probability to the variable

$$B(p)_t \;=\sum_{s\leq t} |\Delta X_s|^p \tag{3.15}$$

(we have just "proved" the convergence for any given t, but it is also a functional convergence, for the Skorokhod topology, in probability).

Again, these cases do not exhaust the possible behaviors of f, and further we have not given a CLT in the second situation above. But, when f is not bounded it looks a bit strange to impose a specific behavior at infinity, and without this there is simply *no convergence result* for $V'(f,\Delta_n)_t$, not to speak about CLTs.

Now we turn to the processes $V(f,\Delta_n)$. To begin with, we observe that, similar to (3.14), we have

$$\left.\begin{array}{l} V(f,\Delta_n)_t = V(X',f,\Delta_n)_t + A_t^n, \quad \text{where} \\ A_t^n = \sum_{i=1}^{[t/\Delta_n]} \sum_{p\geq 1} 1_{\{(i-1)\Delta_n < T_p \leq i\Delta_n\}} \Big(f(\Phi_p + \Delta_i^n X') - f(\Delta_i^n X')\Big). \end{array}\right\} \tag{3.16}$$

The first - fundamental - difference with the continuous case is that (3.11) fails now when f vanishes on a neighborhood of 0. In this case, though, for each given t and all n bigger than some number depending on (ω,t), we have $V(X';f,\Delta_n)_s = 0$ for all $s\leq t$ by (3.11), hence

$$V(f,\Delta_n)_s \;=\; \sum_{i=1}^{[s/\Delta_n]} \sum_{p\geq 1} 1_{\{(i-1)\Delta_n < T_p \leq i\Delta_n\}} f(\Phi_p + \Delta_i^n X'), \qquad \forall s\leq t. \tag{3.17}$$

Then, as soon as f is continuous and vanishes on a neighborhood of 0, we get

$$V(f,\Delta_n)_t \;\stackrel{\text{Sk}}{\longrightarrow}\; V(f)_t \;:=\; \sum_{s\leq t} f(\Delta X_s). \tag{3.18}$$

Here $\stackrel{\text{Sk}}{\longrightarrow}$ means "convergence for the Skorokhod topology", pointwise in ω (the reason for which we have convergence in the Skorokhod sense will be explained later; what is clear at this point is that we have the - pointwise in ω - convergence for all t such that X is continuous at t; we also have for each t an almost sure convergence above).

Next, we consider the case where f is still continuous and, say, coincides with h_p for some $p>0$ on a neighborhood of 0. For any given $\varepsilon>0$ we can write $f = f_\varepsilon + \widehat{f_\varepsilon}$ with f_ε and $\widehat{f_\varepsilon}$ continuous, and $f_\varepsilon(x) = h_p(x)$ if $|x|\leq \varepsilon/2$ and $f_\varepsilon(x)=0$ if $|x|\geq\varepsilon$ and $|f_\varepsilon|\leq h_p$ everywhere. Since $\widehat{f_\varepsilon}$ vanishes around 0, we have $V(\widehat{f_\varepsilon},\Delta_n)_t \to V(\widehat{f_\varepsilon})_t$ by (3.18), and $V(\widehat{f_\varepsilon})_t$ converges to $V(f)_t$ as $\varepsilon\to 0$. On the other hand the process A^n associated with f_ε by (3.16) is the sum of summands smaller than $2\varepsilon^p$, the number of them being bounded for each (ω,t) by a number independent of ε: hence A_t^n is negligible and $V(f_\varepsilon,\Delta_n)$ and $V(X';f_\varepsilon,\Delta_n)$ behave essentially in the same way. This means heuristically that, with the symbol $\asymp$ meaning "approximately equal to", we have

$$V(\widehat{f_\varepsilon},\Delta_n)_t \asymp V(f)_t, \qquad V(f_\varepsilon,\Delta_n)_t \asymp \Delta_n^{p/2-1} t\,\sigma^p\, m_p. \tag{3.19}$$

Adding these two expressions, we get

$$\left.\begin{array}{lll} V(f,\Delta_n)_t \stackrel{\mathbb{P}-\text{Sk}}{\longrightarrow} V(f)_t & & \text{if } \ p>2 \\ V(f,\Delta_n)_t \stackrel{\mathbb{P}-\text{Sk}}{\longrightarrow} V(f)_t + t\sigma^2 & & \text{if } \ p=2 \\ \Delta_n^{1-r/2} V(f,\Delta_n)_t \stackrel{\text{u.c.p.}}{\longrightarrow} t\sigma^{p/2} m_p & & \text{if } \ p<2. \end{array}\right\} \tag{3.20}$$

This type of LLN, which shows a variety of behaviors according to how f behaves near 0, will be found for much more general processes later, in (almost) exactly the same terms.

Now we turn to the CLT. Here again we single out first the case where f vanishes in a neighborhood of 0. We need to find out what happens to the difference $V(f,\Delta_n)-V(f)$. It is easier to evaluate is the difference $V(f,\Delta_n)_t-V(f)_{\Delta_n[t/\Delta_n]}$, since by (3.17) we have

$$V(f,\Delta_n)_s-V(f)_{\Delta_n[s/\Delta_n]} \;=\; \sum_{i=1}^{[s/\Delta_n]}\sum_{p\geq 1} 1_{\{(i-1)\Delta_n<T_p\leq i\Delta_n\}}\Big(f(\Phi_p+\Delta_i^n X')-f(\Phi_p)\Big) \tag{3.21}$$

for all $s\leq t$, as soon as n is large enough. Provided f is C^1, with derivative f', the pth summand above is approximately $f'(\Phi_p)\Delta_i^n X'$. Now the normalized increment $\Delta_i^n X'/\sqrt{\Delta_n}$, for the value of i such that $(i-1)\Delta_n<T_p\leq i\Delta_n$, has the law $\mathcal{N}(0,\sigma^2)$ (because X' and Y are independent), and it is asymptotically independent of the process X (more details are to be found later). Thus if $(U_p)_{p\geq 1}$ denotes a sequence of i.i.d. $\mathcal{N}(0,1)$ variables, independent of X, it is not difficult to see that

$$\frac{1}{\sqrt{\Delta_n}}\Big(V(f,\Delta_n)_t-V(f)_{\Delta_n[t/\Delta_n]}\Big) \;\stackrel{\mathcal{L}}{\Longrightarrow}\; \overline{B}(f)_t \;:=\; \sum_{p:T_p\leq t} f'(\Phi_p)\sigma U_p, \tag{3.22}$$

and in fact, this convergence in law (for the Skorokhod topology) is even stable (denoted $\stackrel{\mathcal{L}-\mathrm{s}}{\Longrightarrow}$), a stronger property than the mere convergence in law, which will be defined later only but nevertheless is used in the statements below.

When now f coincide with h_p for some $p>0$ on a neighborhood of 0 and is still C^1 outside 0, exactly as for (3.19) we obtain heuristically that

$$V(\widehat{f_\varepsilon},\Delta_n)_t\asymp V(f)_{\Delta_n[t/\Delta_n]}+\sqrt{\Delta_n}\;U_t^n,\quad V(f_\varepsilon,\Delta_n)_t\asymp \Delta_n^{p/2-1}t\sigma^p m_p+\Delta_n^{p/2-1/2}U_t'^n,$$

where U^n and U'^n converge stably in law to the right side of (3.22) and to the process B of (3.10), respectively. We then have two conflicting rates, and we can indeed prove that, with $\overline{B}(f)$ as in (3.22) and B as in (3.10) (thus depending on r):

$$\left.\begin{array}{lll}
\frac{1}{\sqrt{\Delta_n}}\Big(V(f,\Delta_n)_t-V(f)_{\Delta_n[t/\Delta_n]}\Big) \stackrel{\mathcal{L}-\mathrm{s}}{\Longrightarrow} \overline{B}(f)_t & \text{if} & p>3\\
\frac{1}{\sqrt{\Delta_n}}\Big(V(f,\Delta_n)_t-V(f)_{\Delta_n[t/\Delta_n]}\Big) \stackrel{\mathcal{L}-\mathrm{s}}{\Longrightarrow} t\sigma^3 m_3+\overline{B}(f)_t & \text{if} & p=3\\
\frac{1}{\Delta_n^{p/2-1}}\Big(V(f,\Delta_n)_t-V(f)_{\Delta_n[t/\Delta_n]}\Big) \stackrel{\mathrm{u.c.p.}}{\longrightarrow} t\sigma^p m_p & \text{if} & 2<p<3\\
\frac{1}{\sqrt{\Delta_n}}\Big(V(f,\Delta_n)_t-V(f)_{\Delta_n[t/\Delta_n]}-t\sigma^2\Big) \stackrel{\mathcal{L}-\mathrm{s}}{\Longrightarrow} B_t+\overline{B}(f)_t & \text{if} & p=2\\
\frac{1}{\Delta_n^{1-p/2}}\Big(\Delta_n^{1-p/2}V(f;\Delta_n)_t-t\sigma^p m_p\Big) \stackrel{\mathbb{P}-\mathrm{Sk}}{\longrightarrow} V(f)_t & \text{if} & 1<p<2\\
\frac{1}{\sqrt{\Delta_n}}\Big(\sqrt{\Delta_n}\;V(f,\Delta_n)_t-t\sigma m_1\Big) \stackrel{\mathcal{L}-\mathrm{s}}{\Longrightarrow} V(f)_t+B_t & \text{if} & p=1\\
\frac{1}{\sqrt{\Delta_n}}\Big(\Delta_n^{1-p/2}V(f,\Delta_n)_t-t\sigma^p m_p\Big) \stackrel{\mathcal{L}-\mathrm{s}}{\Longrightarrow} B_t & \text{if} & p<1.
\end{array}\right\} \tag{3.23}$$

Hence we obtain a genuine CLT, relative to the LLN (3.20), in the cases $p>3$, $p=2$ and $p<1$. When $p=3$ and $p=1$ we still have a CLT, with a bias. When $2<p<3$ or $1<p<2$ we have a "second order LNN", and the associated genuine CLTs run as follows:

$$\left.\begin{array}{lll}
\frac{1}{\sqrt{\Delta_n}}\Big(V(f,\Delta_n)_t-V(f)_{\Delta_n[t/\Delta_n]}-\Delta_n^{p/2-1}t\sigma^p m_p\Big) \stackrel{\mathcal{L}-\mathrm{s}}{\Longrightarrow} \overline{B}(f)_t & \text{if} & 2<p<3\\
\frac{1}{\Delta_n^{p/2-1/2}}\Big(V(f,\Delta_n)_t-V(f)_{\Delta_n[t/\Delta_n]}-\Delta_n^{p/2-1}t\sigma^p m_p\Big) \stackrel{\mathcal{L}-\mathrm{s}}{\Longrightarrow} B_t & \text{if} & 1<p<2
\end{array}\right\} \tag{3.24}$$

We see that these results exhibit again a large variety of behavior. This will be encountered also for more general underlying processes X, with of course more complicated statements and proofs (in the present situation we have not really given the complete proof, of course, but it is relatively easy along the lines outlined above). However, in the general situation we will not give such a complete picture, which is useless for practical applications. Only (3.20) and the cases $r > 2$ in (3.23) will be given.

4 Auxiliary limit theorems

The aims of this section are twofold: first we define the stable convergence in law, already mentioned in the previous section. Second, we recall a number of limit theorems for partial sums of triangular arrays of random variables.

1) Stable convergence in law. This notion has been introduced by Rényi in [22], for the very same reasons as we need it here. We refer to [4] for a very simple exposition and to [13] for more details.

It often happens that a sequence of statistics Z_n converges in law to a limit Z which has, say, a mixed centered normal distribution: that is, $Z = \Sigma U$ where U is an $\mathcal{N}(0,1)$ variable and Σ is a positive variable independent of U. This poses no problem other than computational when the law of Σ is known. However, in many instances the law of Σ is unknown, but we can find a sequence of statistics Σ_n such that the pair (Z_n, Σ_n) converges in law to (Z, Σ); so although the law of the pair (Z, Σ) is unknown, the variable Z_n/Σ_n converges in law to $\mathcal{N}(0,1)$ and we can base estimation or testing procedures on this new statistics Z_n/Σ_n. This is where the stable convergence in law comes into play.

The formal definition is a bit involved. It applies to a sequence of random variables Z_n, all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and taking their values in the same state space $(E, \mathcal{E})$, assumed to be Polish (= metric complete and separable). We say that Z_n *stably converges in law* if there is a probability measure η on the product $(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$, such that $\eta(A \times E) = \mathbb{P}(A)$ for all $A \in \mathcal{F}$ and

$$\mathbb{E}(Y f(Z_n)) \;\to\; \int Y(\omega) f(x) \eta(d\omega, dx) \tag{4.1}$$

for all bounded continuous functions f on E and bounded random variables Y on $(\Omega, \mathcal{F})$.

This is an "abstract" definition, similar to the definition of the convergence in law which says that $\mathbb{E}(f(Z_n)) \to \int f(x)\rho(dx)$ for some probability measure ρ. Now for the convergence in law we usually want a limit, that is we say $Z_n \xrightarrow{\mathcal{L}} Z$, and the variable Z is *any* variable with law ρ, of course. In a similar way it is convenient to "realize" the limit Z for the stable convergence in law.

We can always realize Z in the following way: take $\widetilde{\Omega} = \Omega \times E$ and $\widetilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{E}$ and endow $(\widetilde{\Omega}, \widetilde{\mathcal{F}})$ with the probability η, and put $Z(\omega, x) = x$. But, as for the simple convergence in law, we can also consider other extensions of $(\Omega, \mathcal{F}, \mathbb{P})$: that is, we have a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$, where $\widetilde{\Omega} = \Omega \times \widetilde{\Omega}'$ and $\widetilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}'$ for some auxiliary measurable space $(\Omega', \mathcal{F}')$ and $\widetilde{\mathbb{P}}$ is a probability measure on $(\widetilde{\Omega}, \widetilde{\mathcal{F}})$ whose first marginal is $\mathbb{P}$, and we also have a random variable Z on this extension. Then in this setting, (4.1) is equivalent to saying (with $\widetilde{\mathbb{E}}$ denoting the expectation w.r.t. $\widetilde{\mathbb{P}}$)

$$\mathbb{E}(Yf(Z_n)) \;\to\; \widetilde{\mathbb{E}}(Yf(Z)) \tag{4.2}$$

for all f and Y as above, as soon as $\widetilde{\mathbb{P}}(A \cap \{Z \in B\}) = \eta(A \times B)$ for all $A \in \mathcal{F}$ and $B \in \mathcal{E}$. We then say that Z_n converges stably to Z, and this convergence is denoted by $\stackrel{\mathcal{L}-\mathrm{s}}{\longrightarrow}$.

Clearly, when η is given, the property $\widetilde{\mathbb{P}}(A \cap \{Z \in B\}) = \eta(A \times B)$ for all $A \in \mathcal{F}$ and $B \in \mathcal{E}$ simply amounts to specifying the law of Z, conditionally on the σ-field $\mathcal{F}$. Therefore, saying $Z_n \stackrel{\mathcal{L}-\mathrm{s}}{\longrightarrow} Z$ amounts to saying that we have the stable convergence in law towards a variable Z, defined on any extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$, and with a specified conditional law knowing $\mathcal{F}$.

Obviously, the stable convergence in law implies the convergence in law. But it implies much more, and in particular the following crucial result: if $Z_n \stackrel{\mathcal{L}-\mathrm{s}}{\longrightarrow} Z$ and if Y_n and Y are variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and with values in the same Polish space F, then

$$Y_n \stackrel{\mathbb{P}}{\longrightarrow} Y \quad \Rightarrow \quad (Y_n, Z_n) \stackrel{\mathcal{L}-\mathrm{s}}{\longrightarrow} (Y, Z). \tag{4.3}$$

On the other hand, there are criteria for stable convergence in law of a given sequence Z_n. The σ-field generated by all Z_n is necessarily separable, that is generated by a countable algebra, say $\mathcal{G}$. Then if for any finite family $(A_p : 1 \leq p \leq q)$ in $\mathcal{G}$, the sequence $(Z_n, (1_{A_p})_{1 \leq p \leq q})$ of $E \times \mathbb{R}^q$-valued variables converges in law as $n \to \infty$, then necessarily Z_n converges stably in law.

2) Convergence of triangular arrays. Our aim is to prove the convergence of functionals like in (3.1) and (3.2), which appear in a natural way as partial sums of triangular arrays. We really need the convergence for the terminal time T, but in most cases the available convergence criteria also give the convergence as processes, for the Skorokhod topology. So now we provide a set of conditions implying the convergence of partial sums of triangular arrays, all results being in [13].

We are not looking for the most general situation here, and we restrict our attention to the case where the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ is fixed. For each n we have a sequence of $\mathbb{R}^d$-valued variables $(\zeta_i^n : i \geq 1)$, the components being denoted by $\zeta_i^{n,j}$ for $j = 1, \cdots, d$. The *key assumption* is that for all n, i the variable ζ_i^n is $\mathcal{F}_{i\Delta_n}$-measurable, and this assumption is in force in the remainder of this section.

Conditional expectations w.r.t. $\mathcal{F}_{(i-1)\Delta_n}$ will play a crucial role, and to simplify notation we write it $\mathbb{E}_{i-1}^n$ instead of $\mathbb{E}(. \mid \mathcal{F}_{(i-1)\Delta_n})$, and likewise $\mathbb{P}_{i-1}^n$ is the conditional probability.

Lemma 4.1 *If we have*

$$\sum_{i=1}^{[t/\Delta_n]} \mathbb{E}_{i-1}^n(\|\zeta_i^n\|) \;\stackrel{\mathbb{P}}{\longrightarrow}\; 0 \qquad \forall t > 0, \tag{4.4}$$

then $\sum_{i=1}^{[t/\Delta_n]} \zeta_i^n \stackrel{u.c.p.}{\longrightarrow} 0$. *The same conclusion holds under the following two conditions:*

$$\sum_{i=1}^{[t/\Delta_n]} \mathbb{E}_{i-1}^n(\zeta_i^n) \stackrel{u.c.p.}{\longrightarrow} 0, \tag{4.5}$$

$$\sum_{i=1}^{[t/\Delta_n]} \mathbb{E}_{i-1}^n(\|\zeta_i^n\|^2) \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \forall t > 0. \tag{4.6}$$

In particular when ζ_i^n *is a martingale difference, that is* $\mathbb{E}_{i-1}^n(\zeta_i^n) = 0$, *then (4.6) is enough to imply* $\sum_{i=1}^{[t/\Delta_n]} \zeta_i^n \stackrel{u.c.p.}{\longrightarrow} 0$.

Lemma 4.2 *If we have*

$$\sum_{i=1}^{[t/\Delta_n]} \mathbb{E}_{i-1}^n(\zeta_i^n) \stackrel{u.c.p.}{\longrightarrow} A_t \tag{4.7}$$

for some continuous adapted $\mathbb{R}^d$*-valued process of finite variation* A, *and if further (4.6) holds, then we have* $\sum_{i=1}^{[t/\Delta_n]} \zeta_i^n \stackrel{u.c.p.}{\longrightarrow} A_t$.

Lemma 4.3 *If we have (4.7) for some (deterministic) continuous* $\mathbb{R}^d$*-valued function of finite variation* A, *and also the following two conditions:*

$$\sum_{i=1}^{[t/\Delta_n]} \Big(\mathbb{E}_{i-1}^n(\zeta_i^{n,j}\zeta_i^{n,k}) - \mathbb{E}_{i-1}^n(\zeta_i^{n,j})\mathbb{E}_{i-1}^n(\zeta_i^{n,k}) \Big) \stackrel{\mathbb{P}}{\longrightarrow} C_t'^{jk} \qquad \forall t > 0,\ j,k = 1,\cdots,d, \tag{4.8}$$

$$\sum_{i=1}^{[t/\Delta_n]} \mathbb{E}_{i-1}^n(\|\zeta_i^n\|^4) \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \forall t > 0, \tag{4.9}$$

where $C' = (C'^{jk})$ *is a (deterministic) function, continuous and increasing in* $\mathcal{M}_d^+$, *then the processes* $\sum_{i=1}^{[t/\Delta_n]} \zeta_i^n$ *converge in law to* $A + B$, *where* B *is a continuous centered Gaussian* $\mathbb{R}^d$*-valued process with independent increments with* $\mathbb{E}(B_t^j B_t^k) = C_t'^{jk}$.

(4.9) is a conditional *Lindeberg condition*, whose aims is to ensure that the limiting process is continuous; other, weaker, conditions of the same type are available, but not needed here. The conditions given above completely characterize, of course, the law of the process B. Equivalently we could say that B is a Gaussian martingale (relative to the filtration it generates), starting from 0, and with quadratic variation process C'.

3) Stable convergence of triangular arrays. The reader will have observed that the conditions (4.7) and (4.8) in Lemma 4.3 are very restrictive, because the limits are non-random. In the sequel, such a situation rarely occurs, and typically these conditions are satisfied with A and C' random. But then we need an additional condition, under which it turns out that the convergence holds not only in law, but even stably in law.

Note that the stable convergence in law has been defined for variables taking values in a Polish space, so it also applies to right-continuous and left limited d-dimensional processes: such a process can be viewed as a variable taking its values in the Skorokhod space $\mathbb{D}(\mathbb{R}^d)$ of all functions from $\mathbb{R}_+$ into $\mathbb{R}^d$ which are right-continuous with left limits, provided we endow this space with the Skorokhod topology which makes it a Polish space. See [10] or Chapter VI of [13] for details on this topology. In fact, in Lemma 4.3 the convergence in law is also relative to this Skorokhod topology. The stable convergence in law for processes is denoted as $\stackrel{\mathcal{L}-s}{\Longrightarrow}$ below.

In the previous results the fact that all variables were defined on the same space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and the ζ_i^n's were $\mathcal{F}_{i\Delta_n}$-measurable was essentially irrelevant. This is no longer the case for the next result, for which this setting is fundamental.

Below we single out, among all martingales on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, a possibly multidimensional Wiener process W. The following lemma holds for any choice of W, and even with no W at all (in which case a martingale "orthogonal to W" below means any martingale) but we will use it mainly with the process W showing in (1.7). The following is a particular case of Theorem IX.7.28 of [13].

Lemma 4.4 *Assume (4.7) for some continuous adapted $\mathbb{R}^d$-valued process of finite variation A, and (4.8) with some continuous adapted process $C' = (C'^{jk})$ with values in $\mathcal{M}_d^+$ and increasing in this set, and also (4.9). Assume also*

$$\sum_{i=1}^{[t/\Delta_n]} \mathbb{E}_{i-1}^n(\zeta_i^n\, \Delta_i^n N) \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \forall t > 0 \tag{4.10}$$

whenever N is one of the components of W or is a bounded martingale orthogonal to W. Then the processes $\sum_{i=1}^{[t/\Delta_n]} \zeta_i^n$ converge stably in law to $A + B$, where B is a continuous process defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of the space $(\Omega, \mathcal{F}, \mathbb{P})$ and which, conditionally on the σ-field $\mathcal{F}$, is a centered Gaussian $\mathbb{R}^d$-valued process with independent increments satisfying $\widetilde{\mathbb{E}}(B_t^j B_t^k \mid \mathcal{F}) = C_t'^{jk}$.

The conditions stated above completely specify the conditional law of B, knowing $\mathcal{F}$, so we are exactly in the setting explained in §1 above and the stable convergence in law is well defined. However one can say even more: letting $(\widetilde{\mathcal{F}}_t)$ be the smallest filtration on $\widetilde{\Omega}$ which make B adapted and which contains $(\mathcal{F}_t)$ (that is, $A \times \Omega' \in \widetilde{\mathcal{F}}_t$ whenever $A \in \mathcal{F}_t$), then B is a continuous local martingale on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\geq 0}, \widetilde{\mathbb{P}})$ which is orthogonal in the martingale sense to any martingale on the space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, and whose quadratic variation process is C'. Of course, on the extended space B is no longer Gaussian.

The condition (4.10) could be substituted with weaker ones. For example if it holds when N is orthogonal to W, whereas $\sum_{i=1}^{[t/\Delta_n]} \mathbb{E}_{i-1}^n(\zeta_i^n\, \Delta_i^n W^j)$ converges in probability to a continuous process for all indices j, we still have the stable convergence in law of $\sum_{i=1}^{[t/\Delta_n]} \zeta_i^n$, but the limit has the form $A + B + M$, where the process M is a stochastic integral with respect to W. Se [13] for more details.

5 A first LNN (Law of Large Numbers)

At this stage we start giving the basic limit theorems which are used later for statistical applications. Perhaps giving first all limit theorems in a purely probabilistic setting is not the most pedagogical way of proceeding, but it is the most economical in terms of space...

We are in fact going to provide a version of the results of Section 3, and other connected results, when the basic process X is an Itô semimartingale. There are two kinds of results: first some LNNs similar to (3.5), (3.9), (3.18) or (3.20); second, some "central limit theorems" (CLT) similar to (3.10) or (3.23). We will not give a complete picture, and rather restrict ourselves to those results which are used in the statistical applications.

Warning: Below, and in all these notes, the proofs are often sketchy and sometimes absent; for the full proofs, which are sometimes a bit complicated, we refer essentially to [15] (which is restricted to the 1-dimensional case for X, but the multidimensional extension is straightforward).

In this section, we provide some general results, valid for any d-dimensional semimartingale $X = (X^j)_{1\leq j\leq d}$, not necessarily Itô. We also use the notation (3.1) and (3.2). We start by recalling the fundamental result about quadratic variation, which says that for any indices j, k, and as $n \to \infty$ (recall $\Delta_n \to 0$):

$$\sum_{i=1}^{[t/\Delta_n]} \Delta_i^n X^j \Delta_i^n X^k \stackrel{\mathbb{P}-\mathrm{Sk}}{\longrightarrow} [X^j, X^k]_t = C_t^{jk} + \sum_{s\leq t} \Delta X_s^j \Delta X_s^k. \tag{5.1}$$

This is the convergence in probability, for the Skorokhod topology, and we even have the joint convergence for the Skorokhod topology for the d^2-dimensional processes, when $1 \leq j, k \leq d$. When further X has no fixed times of discontinuity, for example when it is an Itô semimartingale, we also have the convergence in probability for any fixed t.

Theorem 5.1 *Let f be a continuous function from $\mathbb{R}^d$ into $\mathbb{R}^{d'}$.*

a) If $f(x) = \mathrm{o}(\|x\|^2)$ as $x \to 0$, then

$$V(f, \Delta_n)_t \stackrel{\mathbb{P}-Sk}{\longrightarrow} f \star \mu_t = \sum_{s\leq t} f(\Delta X_s). \tag{5.2}$$

b) If f coincide on a neighborhood of 0 *with the function $g(x) = \sum_{j,k=1}^d \gamma_{jk} x_j x_k$ (here each γ_{jk} is a vector in $\mathbb{R}^{d'}$), then*

$$V(f, \Delta_n)_t \stackrel{\mathbb{P}-Sk}{\longrightarrow} \sum_{j,k=1}^d \gamma_{jk} C_t^{jk} + f \star \mu_t. \tag{5.3}$$

Moreover both convergences above also hold in probability for any fixed t such that $\mathbb{P}(\Delta X_t = 0) = 1$ (hence for all t when X is an Itô semimartingale).

Proof. 1) Suppose first that $f(x) = 0$ when $\|x\| \leq \varepsilon$, for some $\varepsilon > 0$. Denote by $S_1, S_2, \cdots$ the successive jump times of X corresponding to jumps of norm bigger than

$\varepsilon/2$, so $S_p \to \infty$. Fix $T > 0$. For each $\omega \in \Omega$ there are two integers $Q = Q(T,\omega)$ and $N = N(T,\omega)$ such that $S_Q(\omega) \le T < S_{Q+1}(\omega)$ and for all $n \ge N$ and for any interval $(i-1)\Delta_n, i\Delta_n]$ in $[0,T]$ then either there is no S_q in this interval and $\|\Delta_i^n X\| \le \varepsilon$, or there is exactly one S_q in it and then we set $\alpha_q^n = \Delta_i^n X - \Delta X_{S_q}$. Since $f(x) = 0$ when $\|x\| \le \varepsilon$ we clearly have for all $t \le T$ and $n \ge N$:

$$\Big\|V(f,\Delta_n)_t - \sum_{q:\ S_q \le \Delta_n[t/\Delta_n]} f(\Delta X_{S_q})\Big\| \ \le\ \sum_{q=1}^{Q} |f(\Delta X_{S_q} + \alpha_q^n) - f(\Delta X_{S_q})|.$$

Then the continuity of f yields (5.2), because $\alpha_q^n \to 0$ for all q.

2) We now turn to the general case in (a). For any $\eta > 0$ there is $\varepsilon > 0$ such that we can write $f = f_\varepsilon + f'_\varepsilon$, where f_ε is continuous and vanishes for $\|x\| \le \varepsilon$, and where $\|f'_\varepsilon(x)\| \le \eta\|x\|^2$. By virtue of (5.1) and the first part of the proof, we have

$$\begin{cases} \|V(f'_\varepsilon,\Delta_n)\| \ \le\ \eta \sum_{i=1}^{[t/\Delta_n]} \|\Delta_i^n X\|^2 \ \stackrel{\mathbb{P}-\mathrm{Sk}}{\longrightarrow}\ \eta \sum_{j=1}^{d} [X^j,X^j], \\ V(f_\varepsilon,\Delta_n) \ \stackrel{\mathbb{P}-\mathrm{Sk}}{\longrightarrow}\ f_\varepsilon \star \mu \end{cases}$$

Moreover, $f_\varepsilon \star \mu \ \stackrel{\mathrm{u.c.p.}}{\longrightarrow} f \star \mu$ as $\varepsilon \to 0$ follows easily from Lebesgue convergence theorem and the property $f(x) = \mathrm{o}(\|x\|^2)$ as $x \to 0$, because $\|x\|^2 \star \mu_t < \infty$ for all t. Since $\eta > 0$ and $\varepsilon > 0$ are arbitrarily small, we deduce (5.2) from $V(f,\Delta_n) = V(f_\varepsilon,\Delta_n) + V(f'_\varepsilon,\Delta_n)$.

3) Now we prove (b). Let $f' = f - g$, which vanishes on a neighborhood of 0. Then if we combine (5.1) and (5.2), plus a classical property of the Skorokhod convergence, we obtain that the pair $(V(g,\Delta_n), V(f',\Delta_n))$ converges (for the $2d'$-dimensional Skorokhod topology, in probability) to the pair $\Big(\sum_{j,k=1}^d \gamma_{jk} C^{jk} + g \star \mu, f' \star \mu\Big)$, and by adding the two components we obtain (5.3).

Finally the last claim comes from a classical property of the Skorokhod convergence, plus the fact that an Itô semimartingale has no fixed time of discontinuity. □

In particular, in the 1-dimensional case we obtain (recall (3.1)):

$$p > 2 \quad \Rightarrow \quad B(pr,\Delta_n) \ \stackrel{\mathbb{P}-\mathrm{Sk}}{\longrightarrow}\ B(p)_t \ :=\ \sum_{s \le t} |\Delta X_s|^p. \tag{5.4}$$

This result is due to Lépingle [18], who even proved the almost sure convergence. It completely fails when $r \le 2$ except under some special circumstances.

6 Some other LNNs

6.1 Hypotheses.

So far we have generalized (3.18) to any semimartingale, under appropriate conditions on f. If we want to generalize (3.5) or (3.14) we need X to be an Itô semimartingale, plus the fact that the processes (b_t) and (σ_t) and the function δ in (1.7) are locally bounded and (σ_t) is either right-continuous or left-continuous.

When it comes to the CLTs we need even more. So for a clearer exposition we gather all hypotheses needed in the sequel, either for LNNs or CLTs, in a single assumption.

Assumption (H): The process X has the form (1.7), and the volatility process σ_t is also an Itô semimartingale of the form

$$\sigma_t \;=\; \sigma_0 + \int_0^t \widetilde{b}_s\,ds + \int_0^t \widetilde{\sigma}\,dW_s + \widetilde{\kappa}(\widetilde{\delta}) \star (\underline{\mu} - \underline{\nu})_t + \widetilde{\kappa}'(\widetilde{\delta}) \star \underline{\mu}_t. \tag{6.1}$$

In this formula, σ_t (a $d \times d'$ matrix) is considered as an $\mathbb{R}^{dd'}$-valued process; $\widetilde{b}_t(\omega)$ and $\widetilde{\sigma}_t(\omega)$ are optional processes, respectively dd' and dd'^2-dimensional, and $\widetilde{\delta}(\omega,t,x)$ id a dd'-dimensional predictable function on $\Omega \times \mathbb{R}_+ \times E$; finally $\widetilde{\kappa}$ is a truncation function on $\mathbb{R}^{dd'}$ and $\widetilde{\kappa}'(x) = x - \widetilde{\kappa}(x)$.

Moreover, we have:

(a) The processes $\widetilde{b}_t(\omega)$ and $\sup_{x\in E} \frac{\|\delta(\omega,t,x)\|}{\gamma(x)}$ and $\sup_{x\in E} \frac{\|\widetilde{\delta}(\omega,t,x)\|}{\widetilde{\gamma}(x)}$ are locally bounded, where γ and $\widetilde{\gamma}$ are (non-random) nonnegative functions satisfying $\int_E (\gamma(x)^2 \wedge 1)\,\lambda(dx) < \infty$ and $\int_E (\widetilde{\gamma}(x)^2 \wedge 1)\,\lambda(dx) < \infty$.

(b) All paths $t \mapsto b_t(\omega)$, $t \mapsto \widetilde{\sigma}_t(\omega)$, $t \mapsto \delta(\omega,t,x)$ and $t \mapsto \widetilde{\delta}(\omega,t,x)$ are left-continuous with right limits. □

Recall that "$(\widetilde{b}_t)$ is locally bounded", for example, means that there exists an increasing sequence (T_n) of stopping times, with $T_n \to \infty$, and such that each stopped process $\widetilde{b}_t^{T_n} = \widetilde{b}_{t\wedge T_n}$ is bounded by a constant (depending on n, but not on (ω,t)).

Remark 6.1 For the LNNs, and also for the CLTs in which there is a discontinuous limit below, we need a weaker form of this assumption, namely *Assumption (H')*: this is as (H), except that we do not require σ_t to be an Itô semimartingale but only to be càdlàg (then of course $\widetilde{b}$, $\widetilde{\sigma}$, $\widetilde{\delta}$ are not present), and b_t is only locally bounded.

As a rule, we will state the results with the mention of this assumption (H'), when the full force of (H) is not needed. However, all proofs will be made assuming (H), because it simplifies the exposition, and because the most useful results need it anyway. □

Apart from the regularity and growth conditions (a) and (b), this assumption amounts to saying that both X and the process σ in (1.7) are Itô semimartingales: since the dimension d' is arbitrary large (and in particular may be bigger than d), this accommodates the case where in (1.7) only the first d components of W occur (by taking $\sigma_t^{ij} = 0$ when $j > d$), whereas in (6.1) other components of W come in, thus allowing σ_t to be driven by the same Wiener process than X, plus an additional multidimensional process. In the same way, it is no restriction to assume that both X and σ are driven by the same Poisson measure $\underline{\mu}$.

So in fact this hypothesis accommodates virtually all models of stock prices or exchange rates or interest rates, with stochastic volatility, including those with jumps, and allows for correlation between the volatility and the asset price processes. For example if we consider a q-dimensional equation

$$dY_t = f(Y_{t-})dZ_t \tag{6.2}$$

where Z is a multi-dimensional Lévy process, and f is a C^2 function with at most linear growth, then if X consists in a subset of the components of Y, it satisfies Assumption (H). The same holds for more general equations driven by a Wiener process and a Poisson random measure.

6.2 The results.

Now we turn to the results. The first, and most essential, result is the following; recall that we use the notation ρ_σ for the law $\mathcal{N}(0,\sigma\sigma^\star)$, and $\rho_\sigma^{\otimes k}$ denotes the k-fold tensor product. We also write $\rho_\sigma^{\otimes k}(f) = \int f(x)\rho_\sigma^{\otimes k}(dx)$ if f is a (Borel) function on $(\mathbb{R}^d)^k$. With such a function f we also associate the following processes

$$V'(f,k,\Delta_n)_t \;=\; \sum_{i=1}^{[t/\Delta_n]} f\Big(\Delta_i^n X/\sqrt{\Delta_n},\cdots,\Delta_{i+k-1}^n X/\sqrt{\Delta_n}\Big). \tag{6.3}$$

Of course when f is a function on $\mathbb{R}^d$, then $V'(f,1,\Delta_n) = V(f,\Delta_n)$, as defined by (3.2).

Theorem 6.2 *Assume (H) (or (H') only, see Remark 6.1)), and let f be a continuous function on $(\mathbb{R}^d)^k$ for some $k \geq 1$, which satisfies*

$$|f(x_1,\cdots,x_k))| \;\leq\; K_0 \prod_{j=1}^k (1+\|x_j\|^p) \tag{6.4}$$

for some $p \geq 0$ and K_0. If either X is continuous, or if $p<2$, we have $\Delta_n V'^n(f,k,\Delta_n)_t \stackrel{u.c.p.}{\longrightarrow} \int_0^t \rho_{\sigma_u}^{\otimes k}(f)du$.

In particular, if X is continuous and the function f on $\mathbb{R}^d$ satisfies $f(\lambda x) = \lambda^p f(x)$ for all $x \in \mathbb{R}^d$ and $\lambda \geq 0$, then

$$\Delta_n^{1-p/2}\, V(f,\Delta_n)_t \quad \stackrel{u.c.p.}{\longrightarrow} \quad \int_0^t \rho_{\sigma_u}(f)du. \tag{6.5}$$

The last claim above may be viewed as an extension of Theorem 5.1 to the case when the limit in (5.2) vanishes. The continuity of f can be somehow relaxed. The proof will be given later, after we state some other LLNs, of two kinds, to be proved later also.

Recalling that one of our main objective is to estimate the integrated volatility C_t^{jk}, we observe that Theorem 5.1 does not provide "consistent estimators" for C_t when X is discontinuous. There are two ways to solve this problem, and the first one is as follows: when X has jumps, (5.1) does not give information on C_t because of the jumps, essentially the "big" ones. However a big jump gives rise to a big increment $\Delta_i^n X$. So an idea, following Mancini [19], [20], consists in throwing away the big increments. The cutoff level has to be chosen carefully, so as to eliminate the jumps but keeping the increments which are "mainly" due to the continuous martingale part X^c, and those are of order $\sqrt{\Delta_n}$. So we choose two numbers $\varpi \in (0,1/2)$ and $\alpha > 0$, and for all indices $j,k \leq d$ we set

$$V^{jk}(\varpi,\alpha,\Delta_n)_t \;=\; \sum_{i=1}^{[t/\Delta_n]} (\Delta_i^n X^j \Delta_i^n X^k) 1_{\{\|\Delta_i^n X\| \leq \alpha\Delta_n^\varpi\}}. \tag{6.6}$$

More generally one can consider the truncated analogue of $V'(f,k,\Delta_n)$ of (6.3). With ϖ and α as above, and if f is a function on $(\mathbb{R}^d)^k$, we set

$$V'(\varpi,\alpha;f,k,\Delta_n)_t = \sum_{i=1}^{[t/\Delta_n]} f\Big(\Delta_i^n X/\sqrt{\Delta_n},\cdots,\Delta_{i+k-1}^n X/\sqrt{\Delta_n}\Big) 1_{\cap_{j=1,\cdots,k}\{\|\Delta_{i+j-1}^n X\|\le \alpha\Delta_n^{\varpi}\}}. \tag{6.7}$$

Theorem 6.3 *Assume (H) (or (H') only), and let f be a continuous function on $(\mathbb{R}^d)^k$ for some $k\ge 1$, which satisfies (6.4) for some $p\ge 0$ and some $K_0>0$. Let also $\varpi\in(0,\frac12)$ and $\alpha>0$. If either X is continuous, or X is discontinuous and $p\le 2$ we have $\Delta_n V'^n(\varpi,\alpha;f,k,\Delta_n)_t \stackrel{u.c.p.}{\longrightarrow} \int_0^t \rho_{\sigma_u}^{\otimes k}(f)du$.*

In particular, $V^{jk}(\varpi,\alpha,\Delta_n) \stackrel{u.c.p.}{\longrightarrow} C_t^{jk}$.

This result has no real interest when X is continuous. When X jumps, and at the expense of a more complicated proof, one could show that the result holds when $p\le 4$, and also when $p>4$ and $\varpi\ge\frac{p-4}{2p-2r-4}$ when additionally we have $\int(\gamma(x)^r\wedge 1)\lambda(dz)<\infty$ for some $r\in[0,2)$ (where γ is the function occurring in (H)).

The (slight) improvement on the condition on p, upon the previous theorem, allows to easily estimate not only C_t, but also the integral $\int_0^t g(c_s)ds$ for any polynomial g on the set of $d\times d$ matrices. For example if we take

$$f(x_1,\cdots,x_k) = \prod_{j=1}^k (x_j^{m_j} x_j^{n_j}), \tag{6.8}$$

for arbitrary indices m_j and n_j in $\{1,\cdots,d\}$, then we get

$$\Delta_n V'^n(\varpi,\alpha;f,k,\Delta_n)_t \stackrel{\text{u.c.p.}}{\longrightarrow} \int_0^t \prod_{j=1}^k c_s^{m_j n_j}\, ds. \tag{6.9}$$

The problem with this method is that we do not really know how to choose ϖ and α a priori: empirical evidence from simulation studies leads to choose ϖ to be very close to $1/2$, like $\varpi=0.47$ or 0.48, whereas α for estimating C_t^{jj}, say, should be chosen between 2 and 5 times the "average $\sqrt{c^{jj}}$" (recall $c=\sigma\sigma^\star$). So this requires a preliminary rough estimate of the order of magnitude of c^{jj}: of course for financial data this order of magnitude is usually pretty much well known.

Another way, initiated by Barndorff-Nielsen and Shephard (see [6] and [7]) consists in using the so-called bipower, or more generally multipower, variations. This is in fact a particular case of the Theorem 6.2. Indeed, recalling that m_r is the rth absolute moment of $\mathcal{N}(0,1)$, we set for any $r_1,\cdots,r_l\in(0,2)$ with $r_1+\cdots+r_l=2$ (hence $l\ge 2$):

$$V^{jk}(r_1,\cdots,r_l,\Delta_n)_t = \frac{1}{4m_{r_1}\cdots m_{r_l}} \sum_{i=1}^{[t/\Delta_n]} \left(\prod_{v=1}^l |\Delta_{i+v-1}^n(X^j+X^k)|^{r_v} - \prod_{v=1}^l |\Delta_{i+v-1}^n(X^j-X^k)|^{r_v}\right) \tag{6.10}$$

Then obviously this is equal to $\frac{1}{\Delta_n} V'(f,l,\Delta_n)$, where

$$f(x_1,\cdots,x_l) \;=\; \frac{1}{4m_{r_1}\cdots m_{r_l}}\left(\prod_{v=1}^{l}|x_v^j+x_v^k|^{r_v}-\prod_{v=1}^{l}|x_v^j-x_v^k|^{r_v}\right),$$

and $\rho_\sigma^{\otimes l}(f) = (\sigma\sigma^\star)^{jk}$ by a simple calculation. Then we deduce from Theorem 6.2 the following result:

Theorem 6.4 *Assume (H) (or (H') only)), and let $r_1,\cdots,r_l\in(0,2)$ be such that $r_1+\cdots+r_l=2$. Then $V^{jk}(r_1,\cdots,r_l,\Delta_n)\ \stackrel{u.c.p.}{\longrightarrow} C_t^{jk}$.*

Now, the previous LNNs are not enough for the statistical applications we have in mind. Indeed, we need consistent estimators for a few other processes than C_t, and in particular for the following one which appears as a conditional variance in some of the forthcoming CLTs:

$$D^{jk}(f)_t \;=\; \sum_{s\le t} f(\Delta X_s)(c_{s-}^{jk}+c_s^{jk}) \tag{6.11}$$

for indices $j,k\le d$ and a function f on $\mathbb{R}^d$ with $|f(x)|\le K\|x\|^2$ for $\|x\|\le 1$, so the summands above are non-vanishing only when $\Delta X_s\neq 0$ and the process $D^{jk}(f)$ is finite-valued.

To do this we take any sequence k_n of integers satisfying

$$k_n\to\infty,\qquad k_n\Delta_n\to 0, \tag{6.12}$$

and we let $I_{n,t}(i)=\{j\in N:\ j\neq i: 1\le j\le[t/\Delta_n],\ |i-j|\le k_n\}$ define a local window in time of length $k_n\Delta_n$ around time $i\Delta_n$. We also choose $\varpi\in(0,1/2)$ and $\alpha>0$ as in (6.6). We will consider two distinct cases for f and associate with it the functions f_n:

$$\left.\begin{array}{lll}\bullet\ f(x)=\mathrm{o}(\|x\|^2)\ \text{ as } x\to 0, & & f_n(x)=f(x)\\ \bullet\ f(x)=\sum_{v,w=1}^d\gamma_{vw}x_vx_w\ \text{ on a neighborhood of } 0, & & f_n(x)=f(x)1_{\{\|x\|>\alpha\Delta_n^\varpi\}}.\end{array}\right\} \tag{6.13}$$

Finally, we set

$$D^{jk}(f,\varpi,\alpha,\Delta_n)_t=\frac{1}{k_n\Delta_n}\sum_{i=1+k_n}^{[t/\Delta_n]-k_n} f_n(\Delta_i^nX)\sum_{l\in I_{n,t}(i)}(\Delta_l^nX^j\ \Delta_l^nX^k)1_{\{\|\Delta_l^nX\|\le\alpha\Delta_n^\varpi\}}. \tag{6.14}$$

Theorem 6.5 *Assume (H) (or (H') only), and let f be a continuous function on $\mathbb{R}^d$ satisfying (6.13), and $j,k\le d$ and $\varpi\in(0,1/2)$ and $\alpha>0$. Then*

$$D^{jk}(f,\varpi,\alpha,\Delta_n)\ \stackrel{\mathbb{P}-Sk}{\longrightarrow}\ D^{jk}(f). \tag{6.15}$$

If further X is continuous and $f(\lambda x)=\lambda^pf(x)$ for all $\lambda>0$ and $x\in\mathbb{R}^d$, for some $p>2$ (hence we are in the first case of (6.13)), then

$$\Delta_n^{1-p/2}D^{jk}(f,\varpi,\alpha,\Delta_n)\ \ \stackrel{u.c.p.}{\longrightarrow}\ \ 2\int_0^t\rho_{\sigma_u}(f)c_u^{jk}ds. \tag{6.16}$$

Before proceeding to the proof of all those results, we give some preliminaries.

6.3 A localization procedure.

The localization is a simple but very important tool for proving limit theorems for discretized processes, over a finite time interval. We describe it in details in the setting of the previous theorems, but it will also be used later for the CLTs.

The idea is that, for those theorems, we can replace the local boundedness assumptions in (H-r) for example by boundedness (by a constant), which is a much stronger assumption. More precisely, we set

Assumption (SH): We have (H) and also, for some constant Λ and all (ω, t, x):

$$\left.\begin{array}{l} \|b_t(\omega)\| \leq \Lambda, \quad \|\sigma_t(\omega)\| \leq \Lambda, \quad \|X_t(\omega)\| \leq \Lambda, \quad \|\widetilde{b}_t(\omega)\| \leq \Lambda, \quad \|\widetilde{\sigma}_t(\omega)\| \leq \Lambda \\ \|\delta(\omega,t,x)\| \leq \Lambda(\gamma(x) \wedge 1), \qquad \|\widetilde{\delta}(\omega,t,x)\| \leq \Lambda(\widetilde{\gamma}(x) \wedge 1) \end{array}\right\} \tag{6.17}$$

If these are satisfied, we can of course choose γ and $\widetilde{\gamma}$ smaller than 1.

Lemma 6.6 *If X satisfies (H) we can find a sequence of stopping times R_p increasing to $+\infty$ and a sequence of processes $X(p)$ satisfying (SH) and with volatility process $\sigma(p)$, such that*

$$t < R_p \quad \Rightarrow \quad X(p)_t = X_t, \quad \sigma(p)_t = \sigma_t. \tag{6.18}$$

Proof. Let X satisfy (H). The processes b_t, $\widetilde{b}_t$, $\widetilde{\sigma}_t$, $\sup_{x\in E} \frac{\|\delta(t,x)\|}{\gamma(x)}$ and $\sup_{x\in E} \frac{\|\widetilde{\delta}(t,x)\|}{\widetilde{\gamma}(x)}$ are locally bounded, so we can assume the existence of a "localizing sequence" of stopping times T_p (i.e. this sequence is increasing, with infinite limit) such that for $p \geq 1$:

$$t \leq T_p(\omega) \quad \Rightarrow \quad \left\{ \begin{array}{l} \|b_t(\omega)\| \leq p, \quad \|\widetilde{b}_t(\omega)\| \leq p, \quad \|\widetilde{\sigma}_t(\omega)\| \leq p, \\ \|\delta(\omega,t,x)\| \leq p\gamma(x), \quad \|\widetilde{\delta}(\omega,t,x)\| \leq p\widetilde{\gamma}(x). \end{array}\right. \tag{6.19}$$

We also set $S_p = \inf(t : \|X_t\| \geq p \text{ or } \|\sigma_t\| \geq p)$, so $R_p = T_p \wedge S_p$ is again a localizing sequence, and we have (6.19) for $t \leq R_p$ and also $\|X_t\| \leq p$ and $\|\sigma_t\| \leq p$ for $t < R_p$. Then we set

$$b(p)_t = \left\{ \begin{array}{ll} b_t & \text{if } t \leq R_p \\ 0 & \text{otherwise,} \end{array}\right. \qquad \widetilde{b}(p)_t = \left\{ \begin{array}{ll} \widetilde{b}_t & \text{if } t \leq R_p \\ 0 & \text{otherwise,} \end{array}\right. \qquad \widetilde{\sigma}(p)_t = \left\{ \begin{array}{ll} \widetilde{\sigma}_t & \text{if } t \leq R_p \\ 0 & \text{otherwise,} \end{array}\right.$$

$$\delta(p)(\omega,t,x) = \left\{ \begin{array}{ll} \delta(\omega,t,x) & \text{if } \|\delta(\omega,t,x)\| \leq 2p \text{ and } t \leq R_p \\ 0 & \text{otherwise,} \end{array}\right.$$

$$\widetilde{\delta}(p)(\omega,t,x) = \left\{ \begin{array}{ll} \widetilde{\delta}(\omega,t,x) & \text{if } \|\widetilde{\delta}(\omega,t,x)\| \leq 2p \text{ and } t \leq R_p \\ 0 & \text{otherwise,} \end{array}\right.$$

At this stage we define the process $\sigma(p)$ by (6.1) with the starting point $\sigma(p)_0 = \sigma_0$ if $\|\sigma_0\| < p$ and $\sigma(p)_0 = 0$ otherwise, and the coefficients $\widetilde{b}(p)$ and $\widetilde{\sigma}(p)$ and $\widetilde{\delta}(p)$, and then the process $X(p)$ by (1.7) with the starting point $X_0 = X_0$ if $\|X_0\| < p$ and $X(p)_0 = 0$ otherwise, and the coefficients $b(p)$ and $\sigma(p)$ (as defined just above) and $\delta(p)$.

We can write $\underline{\mu}$ as $\underline{\mu} = \sum_{t>0} 1_D(t)\, \varepsilon_{(t,\beta_t)}$ where D is the countable (random) support of $\underline{\mu}$ and β_t is E-valued. Outside a $\mathbb{P}$-null set N we have $\Delta X_t = 1_D(t)\,\delta(t,\beta_t)$ and $\Delta X(p)_t = 1_D(t)\,\delta(p)(t,\beta_t)$, and since $\|\Delta X_t\| \leq 2p$ when $t < R_p$ we deduce $\Delta X_t = \Delta X(p)_t$ if $t < R_p$, which implies that $\kappa'(\delta) * \underline{\mu}_t = \kappa'(\delta(p)) * \underline{\mu}_t$ for $t < R_p$. As for the two local martingales $\kappa(\delta) * (\underline{\mu} - \underline{\nu})$ and $\kappa(\delta(p)) * (\underline{\mu} - \underline{\nu})$, they have (a.s.) the same jumps on the predictable interval $[0, R_p]$ as soon as $\kappa(x) = 0$ when $\|x\| > 2p$ (this readily follows from the definition of $\delta(p)$, so they coincide a.s. on $[0, R_p]$.

The same argument shows that $\widetilde{\kappa}'(\widetilde{\delta}) * \underline{\mu}_t = \widetilde{\kappa}'(\widetilde{\delta}(p)) * \underline{\mu}_t$ for $t < R_p$, and $\widetilde{\kappa}(\widetilde{\delta}) * (\underline{\mu} - \underline{\nu})_t = \widetilde{\kappa}(\widetilde{\delta}(p)) * (\underline{\mu} - \underline{\nu})_t$ for $t \leq R_p$. It first follows in an obvious way that $\sigma(p)_t = \sigma_t$ for all $t < R_p$, and then $X(p)_t = X_t$ for all $t < R_p$, that is (6.18) holds.

Finally ny definition the coefficients $b(p)$, $\widetilde{b}(p)$, $\widetilde{\sigma}(p)$, $\delta(p)$ and $\widetilde{\delta}(p)$ satisfy (6.17) with $\Lambda = 2p$. Moreover the processes $\widetilde{\sigma}(p)$ and $X(p)$ are constant after time R_p, and they have jumps bounded by $2P$, so they satisfy (6.17) with $\Lambda = 3p$, and thus (SH) holds for $X(p)$.□

Now, suppose that, for example, Theorem 6.2 has been proved when X satisfies (SH). Let X satisfy (H) only, and $(X(p), R_p)$ be as above. We then know that, for all p, T and all appropriate functions f,

$$\sup_{t \leq T} \Big| \Delta_n V'^n(X(p); f, k, \Delta_n)_t - \int_0^t \rho^{\otimes k}_{\sigma(p)_u}(f)\,du \Big| \stackrel{\mathbb{P}}{\longrightarrow} 0. \tag{6.20}$$

On the set $\{R_p > T+1\}$, and if $k\Delta_n \leq 1$, we have $V'^n(X(p); f, k, \Delta_n)_t = V'^n(X; f, k, \Delta_n)_t$ and $\sigma(p)_t = \sigma_t$ for all $t \leq T$, by (6.18). Since $\mathbb{P}(R_p > T+1\} \to 1$ as $p \to \infty$, it readily follows that $\Delta_n V'^n(X; f, k, \Delta_n)_t \stackrel{\text{u.c.p.}}{\longrightarrow} \int_0^t \rho^{\otimes k}_{\sigma_u}(f)\,du$. This proves Theorem 6.2 under (H).

This procedure works in exactly the same way for all the theorems below, LNNs or CLTs, and we will call this the "localization procedure" without further comment.

Remark 6.7 If we assume (SH), and if we choose the truncation functions κ and $\widetilde{\kappa}$ in such a way that they coincide with the identity on the balls centered at 0 and with radius 2Λ, in $\mathbb{R}^d$ and $\mathbb{R}^{dd'}$ respectively, then clearly (1.7) and (6.1) can be rewritten as follows:

$$\left.\begin{array}{rcl} X_t &=& X_0 + \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s + \delta \star (\underline{\mu} - \underline{\nu})_t, \\ \sigma_t &=& \sigma_0 + \int_0^t \widetilde{b}_s\,ds + \int_0^t \widetilde{\sigma}\,dW_s + \widetilde{\delta} \star (\underline{\mu} - \underline{\nu})_t. \end{array}\right\} \tag{6.21}$$

6.4 Some estimates.

Below, we assume (SH), and we use the form (6.21) for X and σ. We will give a number of estimates, to be used for the LLNs and also for the CLTs, and we start with some notation. We set

$$\left.\begin{array}{l} \chi'^n_{i,l} = \frac{1}{\sqrt{\Delta_n}} \int_{(i+l-1)\Delta_n}^{(i+l)\Delta_n} \Big(b_s\,ds + (\sigma_s - \sigma_{(i-1)\Delta_n})\,dW_s \Big) \\ \beta^n_{i,l} = \sigma_{(i-1)\Delta_n} \Delta^n_{i+l} W / \sqrt{\Delta_n}, \qquad \chi''^n_{i,l} = \frac{1}{\sqrt{\Delta_n}}\,\Delta^n_{i+l}(\delta \star (\underline{\mu} - \underline{\nu})), \\ \chi^n_{i,l} = \chi'^n_{i,l} + \chi''^n_{i,l}, \qquad \beta^n_i = \beta^n_{i,0}, \qquad \chi^n_i = \chi^n_{i,0}, \qquad \chi'^n_i = \chi'^n_{i,0}. \end{array}\right\} \tag{6.22}$$

In particular, $\Delta_{i+l}^n X = \sqrt{\Delta_n}\,(\chi_{i,l}^n + \beta_{i,l}^n)$. It is well known that the boundedness of the coefficients in (SH) yields, through a repeated use of Doob and Davis-Burkholder-Gundy inequalities, for all $q > 0$ (below, K denotes a constant which varies from line to line and may depend on the constants occurring in (SH); we write it K_p if we want to emphasize its dependency on another parameter p):

$$\left.\begin{array}{ll} \mathbb{E}_{i-1}^n(\|\Delta_i^n X^c\|^q) \leq K_q \Delta_n^{q/2}, & \mathbb{E}(\|\sigma_{t+s} - \sigma_t\|^q \mid \mathcal{F}_t) \leq K_q s^{1\wedge(q/2)}, \\ \mathbb{E}_{i+l-1}^n(\|\beta_{i,l}^n\|^q) \leq K_q, & \mathbb{E}_{i+l-1}^n(\|\chi_{i,l}'^n\|^q) \leq K_{q,l}\Delta_n^{1\wedge(q/2)}, \\ \mathbb{E}_{i+l-1}^n(\|\chi_{i,l}^n\|^q + \|\chi_{i,l}''^n\|^q) \leq \left\{\begin{array}{ll} K_{q,l}\Delta_n^{-(1-q/2)^-} & \text{in general} \\ K_{q,l}\Delta_n^{1\wedge(q/2)} & \text{if } X \text{ is continuous} \end{array}\right. & \end{array}\right\} \tag{6.23}$$

We also use the following notation, for $\eta > 0$:

$$\psi_\eta(x) = \psi(x/\eta), \quad \psi \text{ a } C^\infty \text{ function on } \mathbb{R}^d \text{ with } 1_{\{\|x\|\leq 1\}} \leq \psi(x) \leq 1_{\{\|x\|\leq 2\}}. \tag{6.24}$$

Lemma 6.8 *Assume (SH) and let $r \in [0,2]$ be such that $\int(\gamma(x)^r \wedge 1)\lambda(dx) < \infty$, and α_n a sequence of numbers with $\alpha_n \geq 1$ and $\alpha_n\sqrt{\Delta_n} \to 0$. Then*

$$\lim_{n\to\infty} \sup_{i\geq 1,\ \omega\in\Omega} \Delta_n^{r/2-1}\alpha_n^{r-2}\ \mathbb{E}_{i+l-1}^n(\|\chi_{i,l}''^n\|^2 \wedge \alpha_n^2) \ = \ 0, \tag{6.25}$$

$$r \leq 1 \quad \Rightarrow \quad \lim_{n\to\infty} \sup_{i\geq 1,\ \omega\in\Omega} \Delta_n^{r/2-1}\alpha_n^{r-1}\ \mathbb{E}_{i+l-1}^n\left(\Big|\frac{1}{\sqrt{\Delta_n}}\ \Delta_{i+l}^n(\delta \star \underline{\mu})\Big| \wedge \alpha_n\right) \ = \ 0, \tag{6.26}$$

$$\lim_{\eta\to 0} \limsup_{n\to\infty} \sup_{i\geq 1,\ \omega\in\Omega} \frac{1}{\Delta_n}\ \mathbb{E}_{i+l-1}^n(\|\sqrt{\Delta_n}\ \chi_{i,l}^n\|^2 \wedge \eta^2) \ = \ 0. \tag{6.27}$$

(When $r \leq 1$ above, the two integral processes $\delta \star \underline{\mu}$ and $\delta \star \underline{\nu}$ are well defined, and of finite variation).

Proof. It is enough to consider the 1-dimensional case. For any $\varepsilon \in (0,1]$ we have $\delta \star (\underline{\mu} - \underline{\nu}) = N(\varepsilon) + M(\varepsilon) + B(\varepsilon)$, where ($\kappa$ is the truncation function in (1.1)).

$$N(\varepsilon) = (\delta 1_{\{|\delta|>\varepsilon\}}) \star \underline{\mu}, \quad M(\varepsilon) = (\delta 1_{\{|\delta|\leq\varepsilon\}}) \star (\underline{\mu} - \underline{\nu}), \quad B(\varepsilon) = -(\delta 1_{\{|\delta|>\varepsilon\}}) \star \underline{\nu}\,.$$

Then if $\gamma_\varepsilon = \int_{\{\gamma(x)\leq\varepsilon\}} \gamma(x)^r\lambda(dx)$, we have by (SH):

$$\left.\begin{array}{l} \mathbb{P}_{i+l-1}^n(\Delta_{i+l}^n N(\varepsilon) \neq 0) \ \leq \ \mathbb{E}_{i+l-1}^n(\Delta_{i+l}^n(1_{\{\gamma>\varepsilon\}} \star \underline{\mu})) \ = \ \Delta_n\lambda(\{\gamma > \varepsilon\}) \ \leq \ K\Delta_n\varepsilon^{-r} \\ \mathbb{E}_{i+l-1}^n((\Delta_{i+l}^n M(\varepsilon))^2) \ \leq \ \Delta_n \int_{\{\gamma(x)\leq\varepsilon\}} \gamma(x)^2\lambda(dx) \ \leq \ \Delta_n\varepsilon^{2-r}\gamma_\varepsilon, \\ |\Delta_{i+l}^n B(\varepsilon)| \ \leq \ K\Delta_n\Big(1 + \int_{\{\gamma(x)>\varepsilon\}}(\gamma(x) \wedge 1)\lambda(dx)\Big) \ \leq \ K\Delta_n\varepsilon^{-(r-1)^+}. \end{array}\right\}$$

We also trivially have

$$|\chi_{i,l}^n|^2 \wedge \alpha_n \leq \alpha_n^2 1_{\{\Delta_{i+l}^n N(\varepsilon)\neq 0\}} + 3|\chi_{i,l}'^n|^2 + 3\Delta_n^{-1}|\Delta_{i+l}^n M(\varepsilon)|^2 + 3\Delta_n^{-1}|\Delta_{i+l}^n B(\varepsilon)|^2.$$

Therefore, using (6.23), we get

$$\mathbb{E}^n_{i+l-1}(\|\chi^n_{i,l}\|^2\wedge\alpha_n^2)\leq K\Big(\frac{\alpha_n^2\Delta_n}{\varepsilon^r}+\Delta_n+\varepsilon^{2-r}\gamma_\varepsilon+\Delta_n\varepsilon^{-(r-1)^+}\Big).$$

Then since $\gamma_\varepsilon\to0$ as $\varepsilon\to0$, (6.25) follows by taking $\varepsilon=\varepsilon_n=u_n^2(u_n^{-1}\wedge(\gamma_{u_n})^{-1/4})$, where $u_n=\alpha_n^{1/2}\Delta_n^{1/4}\to0$ (note that $\varepsilon_n\leq u_n$, hence $\gamma_{\varepsilon_n}\leq\gamma_{u_n}$).

Next, suppose $r\leq1$. Then $\delta\star\underline{\mu}=N(\varepsilon)+A(\varepsilon)$, where $A(\varepsilon)=(\delta 1_{\{|\delta|\leq\varepsilon\}})\star\underline{\mu}$, and obviously $\mathbb{E}^n_{i+l-1}(|\Delta^n_{i+l}A(\varepsilon)|\leq K\Delta_n\varepsilon^{1-r}\gamma_\varepsilon$. Moreover

$$\Big|\frac{1}{\sqrt{\Delta_n}}\;\Delta^n_{i+l}(\delta\star\underline{\mu})\Big|\wedge\alpha_n\;\leq\;\alpha_n 1_{\{\Delta^n_{i+l}N(\varepsilon)\neq0\}}+\frac{1}{\sqrt{\Delta_n}}\;|\Delta^n_{i+l}A(\varepsilon)|.$$

Therefore

$$\mathbb{E}^n_{i+l-1}\left(\Big|\frac{1}{\sqrt{\Delta_n}}\;\Delta^n_{i+l}(\delta\star\underline{\mu})\Big|\wedge1\right)\;\leq\;K\left(\frac{\alpha_n\Delta_n}{\varepsilon^r}+\sqrt{\Delta_n}\;\varepsilon^{1-r}\gamma_\varepsilon\right),$$

and the same choice as above for $\varepsilon=\varepsilon_n$ gives (6.26).

Finally, we have for any $\eta>0$:

$$|\sqrt{\Delta_n}\;\chi^n_{i,l}|^2\wedge\eta^2\leq\eta^2 1_{\{\Delta^n_{i+l}N(\varepsilon)\neq0\}}+3\Delta_n|\chi'^n_{i,l}|^2+3|\Delta^n_{i+l}M(\varepsilon)|^2+3|\Delta^n_{i+l}B(\varepsilon)|^2,$$

hence if we take $\varepsilon=\sqrt{\eta}$ above we get $\mathbb{E}^n_{i+l-1}\left(|\sqrt{\Delta_n}\;\chi^n_{i,l}|^2\wedge\eta^2\right)\leq K\Delta_n g'_n(\eta)$, where

$$g'_n(\eta)\;=\;\eta^{2-r/2}+\Delta_n+\eta^{1-r/2}\gamma_{\sqrt{\eta}}+\Delta_n\eta^{-(r-1)^+}.$$

Since $g'_n(\eta)\to g'(\eta):=\eta^{2-r/2}+\eta^{1-r/2}\gamma_{\sqrt{\eta}}$ and $\gamma_\varepsilon\to0$ as $\varepsilon\to0$, we readily get (6.27). $\Box$

Lemma 6.9 *Assume (SH). Let $k\geq1$ and $l\geq0$ be integers and let $q>0$. Let f be a continuous function on $(\mathbb{R}^d)^k$, satisfying (6.4) for some $p\geq0$ and $K_0>0$.*

a) If either X is continuous or if $qp<2$, we have as $n\to\infty$:

$$\sup_{i\geq l,\;\omega\in\Omega}\;\mathbb{E}^n_{i-l-1}\left(\left|f\left(\frac{\Delta^n_iX}{\sqrt{\Delta_n}},\cdots,\frac{\Delta^n_{i+k-1}X}{\sqrt{\Delta_n}}\right)-f\left(\beta^n_{i-l,l},\cdots,\beta^n_{i-l,l+k-1}\right)\right|^q\right)\to0.\quad(6.28)$$

b) If $qp\leq2$, and if α_n is like in the previous lemma, we have as $n\to\infty$:

$$\sup_{i\geq l,\;\omega\in\Omega}\;\mathbb{E}^n_{i-l-1}\left(\left|f\left(\frac{\Delta^n_iX}{\sqrt{\Delta_n}},\cdots,\frac{\Delta^n_{i+k-1}X}{\sqrt{\Delta_n}}\right)1_{\cap_{1\leq j\leq k}\{\|\Delta^n_{i+j-1}X\|\leq\alpha_n\}}\right.\right.$$
$$\left.\left.-f\left(\beta^n_{i-l,l},\cdots,\beta^n_{i-l,l+k-1}\right)\right|^q\right)\to0.\qquad(6.29)$$

Proof. For any $A>0$, the supremum $G_A(\varepsilon)$ of $|f(x_1+y_1,\cdots,x_k+y_k)-f(x_1,\cdots,x_k)|$ over all $\|x_j\|\leq A$ and $\|y_j\|\leq\varepsilon$ goes to 0 as $\varepsilon\to0$. We set $g(x,y)=1+\|x\|^{qp}+\|y\|^{qp}$. If we want to prove (6.29) the sequence α_n is of course as above, whereas if we want to prove

(6.28) we put $\alpha_n = \infty$ for all n. Then for all $A > 1$ and $s \geq 0$ and $\varepsilon > 0$ and $\alpha \in [1, \infty]$ we have, by a (tedious) calculation using (6.4), the constant K depending on K_0, q, k:

$$\begin{aligned} &|f(x_1 + y_1, \cdots, x_k + y_k) 1_{\cap_{j=1,\cdots,k}\{\|x_j\| \leq \alpha_n\}} - f(x_1, \cdots, x_k)|^q \\ &\quad \leq \; G_A(\varepsilon)^q + K \sum_{m=1}^{k} \left(h_{\varepsilon,s,A,n}(x_m, y_m) \prod_{j=1,\cdots,k,\; j \neq m} g(x_j, y_j) \right), \end{aligned} \tag{6.30}$$

where

$$h_{\varepsilon,s,A,n}(x, y) \;=\; \frac{\|x\|^{pq+1}}{A} + \|x\|^{pq}(\|y\| \wedge 1) + A^{pq} \frac{\|y\|^2 \wedge 1}{\varepsilon^2} + \frac{\|y\|^{pq+s} \wedge \alpha_n^{pq+s}}{A^s}.$$

We apply these estimates with $x_j = \beta^n_{i-l,l+j-1}$ and $y_j = \chi^n_{i-l,l+j-1}$. In view of (6.23) we have if X is continuous or if $pq \leq 2$:

$$\mathbb{E}^n_{i+j-2}(g(\beta^n_{i-l,l+j-1}, \chi^n_{i-l,l+j-1})) \leq K. \tag{6.31}$$

Next consider $\zeta^n_{i,j,\varepsilon,A} = \mathbb{E}^n_{i+j-2}(h_{\varepsilon,s,A,n}(\beta^n_{i-l,l+j-1}, \chi^n_{i-l,l+j-1}))$ for an adequate choice of s, to be done below. When X is continuous we take $s = 1$, and (6.23) and Cauchy-Schwarz inequality yield $\zeta^n_{i,j,\varepsilon,A} \leq K(1/A + \sqrt{\Delta_n} + \Delta_n A^{pq}/\varepsilon^2)$. In the discontinuous case when $pq < 2$ and $\alpha_n = \infty$ we take $s = 2 - pq > 0$ and by (6.23) and Cauchy-Schwarz again, plus (6.25) with $r = 2$, we get the existence on a sequence $\delta_n \to 0$ such that $\zeta^n_{i,j,\varepsilon,A} \leq K(1/A + 1/A^s + A^{pq}\delta_n/\varepsilon^2)$. Finally in the discontinuous case when $\alpha_n < \infty$ we have $pq \leq 2$ and we take $s = 0$ and we still obtain $\zeta^n_{i,j,\varepsilon,A} \leq K(1/A + A^{pq}\delta_n/\varepsilon^2)$ by the same argument. To summarize, in all cases we have for all $\varepsilon > 0$:

$$\sup_{\omega,i,j} \zeta^n_{i,j,\varepsilon,A}(\omega) \;\leq\; \psi_n(A, \varepsilon), \qquad \text{where} \quad \lim_{A\to\infty} \limsup_{n\to\infty} \psi_n(A, \varepsilon) \;=\; 0. \tag{6.32}$$

At this stage, we make use of (6.30) and use the two estimates (6.31) and (6.32) and take successive downward conditional expectations to get the left sides of (6.28) and (6.29) are smaller than $G_A(\varepsilon)^q + K\psi_n(A, \varepsilon)$. This hold for all $A > 1$ and $\varepsilon > 0$. Then by using $G_A(\varepsilon) \to 0$ as $\varepsilon \to 0$ and the last part of (6.32), we readily get the results. $\square$

Lemma 6.10 *Under (SH), for any function $(\omega, x) \mapsto g(\omega, x)$ on $\Omega \times \mathbb{R}^d$ which is $\mathcal{F}_{(i-1)\Delta_n} \otimes \mathcal{R}^d$-measurable, and even and with polynomial growth in x, we have*

$$\mathbb{E}^n_{i-1}\left(\Delta^n_i N \; g(., \beta^n_i)\right) \;=\; 0 \tag{6.33}$$

for N being any component of W, or being any bounded martingale orthogonal to W.

Proof. When $N = W^j$ we have $\Delta^n_i N g(\beta^n_i)(\omega) = h(\sigma_{(i-1)\Delta_n}, \Delta^n_i W)(\omega)$ for a function $h(\omega, x, y)$ which is odd and with polynomial growth in y, so obviously (6.33) holds.

Next assume that N is bounded and orthogonal to W. We consider the martingale $M_t = \mathbb{E}(g(., \beta^n_i)|\mathcal{F}_t)$, for $t \geq (i-1)\Delta_n$. Since W is an $(\mathcal{F}_t)$-Brownian motion, and since β^n_i is a function of $\sigma_{(i-1)\Delta_n}$ and of $\Delta^n_i W$, we see that $(M_t)_{t \geq (i-1)\Delta_n}$ is also, conditionally on $\mathcal{F}_{i-1)\Delta_n}$, a martingale w.r.t. the filtration which is generated by the process $W_t -$

$W_{(i-1)\Delta_n}$. By the martingale representation theorem the process M is thus of the form $M_t = M_{(i-1)\Delta_n} + \int_{(i-1)\Delta_n}^t \eta_s dW_s$ for an appropriate predictable process η. It follows that M is orthogonal to the process $N'_t = N_t - N_{(i-1)\Delta_n}$ (for $t \geq (i-1)\Delta_n$), or in other words the product MN' is an $(\mathcal{F}_t)_{t\geq(i-1)\Delta_n}$–martingale. Hence

$$\mathbb{E}^n_{i-1}(\Delta^n_i N\ g(.,\sqrt{\Delta_n}\ \sigma_{(i-1)\Delta_n}\Delta^n_i W)) = \mathbb{E}^n_{i-1}(\Delta^n_i N' M_{i\Delta_n}) = \mathbb{E}^n_{i-1}\Delta^n_i N'\Delta^n_i M) = 0,$$

and thus we get (6.33). □

6.5 Proof of Theorem 6.2.

When $f(\lambda x) = \lambda^p f(x)$ we have $V(f,\Delta_n) = \Delta_n^{p/2} V'(f,\Delta_n)$, hence (6.5) readily follows from the first claim. For this first claim, and as seen above, it is enough to prove it under the stronger assumption (SH).

If we set

$$V''(f,k,\Delta_n)_t \;=\; \sum_{i=1}^{[t/\Delta_n]} f(\beta^n_{i,0},\cdots,\beta^n_{i,k-1}),$$

we have $\Delta_n(V'(f,k,\Delta_n) - V''(f,k,\Delta_n)) \stackrel{\text{u.c.p.}}{\longrightarrow} 0$ by Lemma 6.9-(a) applied with $l=0$ and $q=1$. Therefore it is enough to prove that $\Delta_n V''(f,k,\Delta_n)_t \stackrel{\text{u.c.p.}}{\longrightarrow} \int_0^t \rho^{\otimes k}_{\sigma_v}(f_v)dv$. For this, with $I(n,t,l)$ denoting the set of all $i \in \{1,\cdots,[t/\Delta_n]\}$ which are equal to l modulo k, it is obviously enough to show that for $l = 0,1,\cdots,k-1$:

$$\sum_{i\in I(n,t,l)} \eta^n_i \;\stackrel{\text{u.c.p.}}{\longrightarrow}\; \frac{1}{k}\int_0^t \rho^{\otimes k}_{\sigma_v}(f_v)dv, \qquad \text{where } \eta^n_i = \Delta_n f(\beta^n_{i,0},\cdots,\beta^n_{i,k-1}). \tag{6.34}$$

Observe that η^n_i is $\mathcal{F}_{(i+k-1)\Delta_n}$-measurable, and obviously

$$\mathbb{E}^n_{i-1}(\eta^n_i) \;=\; \Delta_n \rho^{\otimes k}_{\sigma_{(i-1)\Delta_n}}(f), \qquad \mathbb{E}^n_{i-1}(|\eta^n_i|^2) \;\leq\; K\Delta_n^2.$$

By Riemann integration, we have $\sum_{i\in I(n,t,l)} \mathbb{E}^n_{i-1}(\eta^n_i) \stackrel{\text{u.c.p.}}{\longrightarrow} \frac{1}{k}\int_0^t \rho^{\otimes k}_{\sigma_v}(f_v)dv$, because $t \mapsto \rho^{\otimes k}_{\sigma_t}(f)$ is right-continuous with left limits. Hence (6.34) follows from Lemma 4.1.

6.6 Proof of Theorem 6.3.

The proof is exactly the same as for Theorem 6.2, once noticed that in view of Lemma 6.9-(b) applied with $\alpha_n = \alpha\Delta_n^{\varpi-1/2}$ we have

$$\Delta_n(V'(\varpi,\alpha;f,k,\Delta_n)_t - V''(f,k,\Delta_n)_t)) \;\stackrel{\text{u.c.p.}}{\longrightarrow}\; 0.$$

6.7 Proof of Theorem 6.5.

Once more we may assume (SH). Below, j,k are fixed, as well as ϖ and α and the function f, satisfying (6.13), and for simplicity we write $D = D^{jk}(f)$ and $D^n = D^{jk}(f,\varpi,\alpha,\Delta_n)$. Set also

$$\left.\begin{array}{l} \widehat{D}^n_t = \frac{1}{k_n}\sum_{i=1+k_n}^{[t/\Delta_n]-k_n} f_n(\Delta^n_i X)\sum_{l\in I_{n,t}(i)} \beta^{n,j}_l \beta^{n,k}_l, \\ \widehat{D}'^n_t = \frac{1}{k_n}\sum_{i=1+k_n}^{[t/\Delta_n]-k_n} f(\sqrt{\Delta_n}\ \beta^n_i)\sum_{l\in I_{n,t}(i)} \beta^{n,j}_l \beta^{n,k}_l. \end{array}\right\} \tag{6.35}$$

Lemma 6.11 *We have* $\widehat{D}^n \stackrel{\mathbb{P}-Sk}{\longrightarrow} D$.

Proof. a) Let ψ_ε be as in (6.24) and

$$Y(\varepsilon)^n_t = \frac{1}{k_n} \sum_{i=1+k_n}^{[t/\Delta_n]-k_n} (f_n\psi_\varepsilon)(\Delta_i^n X) \sum_{l\in I_{n,t}(i)} \beta_l^{n,j}\beta_l^{n,k}, \qquad Z(\varepsilon)^n_t = \widehat{D}^n_t - Y(\varepsilon)^n_t.$$

It is obviously enough to show the following three properties, for some suitable processes $Z(\varepsilon)$:

$$\lim_{\varepsilon\to 0}\ \limsup_n\ \mathbb{E}(\sup_{s\le t}|Y(\varepsilon)^n_s|) \ = \ 0, \tag{6.36}$$

$$\varepsilon\in(0,1),\ n\to\infty \quad\Rightarrow\quad Z(\varepsilon)^n \stackrel{\mathbb{P}-\mathrm{Sk}}{\longrightarrow} Z(\varepsilon), \tag{6.37}$$

$$\varepsilon\to 0 \quad\Rightarrow\quad Z(\varepsilon) \stackrel{\mathrm{u.c.p.}}{\longrightarrow} D. \tag{6.38}$$

b) Let us prove (6.36) in the first case of (6.13). We have $|(f\psi_\varepsilon)(x)|\le \phi(\varepsilon)\|x\|^2$ for some function ϕ such that $\phi(\varepsilon)\to 0$ as $\varepsilon\to 0$. Hence (6.23) yields $\mathbb{E}^n_{i-1}(|(f\psi_\varepsilon)(\Delta_i^n X)|)\le K\phi(\varepsilon)\Delta_n$. Now, $Y(\varepsilon)^n_t$ is the sum of less than $2k_n[t/\Delta_n]$ terms, all smaller in absolute value than $\frac{1}{k_n}\ |(f\psi_\varepsilon)(\Delta_i^n X)|\|\beta_j^n\|^2$ for some $i\neq j$. By taking two successive conditional expectations and by using again (6.23) the expectation of such a term is smaller than $K\phi(\varepsilon)\Delta_n/k_n$, hence the expectation in (6.36) is smaller than $Kt\phi(\varepsilon)$ and we obtain (6.36).

Next, consider the second case of (6.13). Then $(f_n\psi_\varepsilon)(x) = g(x)1_{\{\alpha\Delta_n^{\varpi}<\|x\|\le\varepsilon\}}$, where g is an homogeneous polynomial of degree 2. Then if $\alpha\Delta_n^{\varpi}<\varepsilon<1/2$ we have

$$|(f_n\psi_\varepsilon)(x+y)| \ \le\ K\Big(\|x\|^4\Delta_n^{-2\varpi} + \|y\|^2\wedge\varepsilon^2\Big).$$

Using this with $x=\sqrt{\Delta_n}\ \beta_i^n$ and $y=\Delta_i^n X'$, we deduce from (6.23) and (6.27) that

$$\mathbb{E}^n_{i-1}(|(f_n\psi_\varepsilon)(\Delta_i^n X)|) \ \le\ K\Delta_n\Big(\Delta_n^{1-2\varpi} + \alpha(n,\varepsilon)\Big),$$

where $\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\alpha(n,\varepsilon)=0$. Then exactly as for the first case, we deduce that the expectation in (6.36) is smaller than $Kt(\Delta_n^{1-2\varpi}+\alpha(n,\varepsilon))$, and we obtain again (6.36).

c) Now we define $Z(\varepsilon)$. Let us call $T_q(\varepsilon)$ for $q=1,2,\cdots$ the successive jump times of the Poisson process $\underline{\mu}([0,t]\times\{x:\gamma(x)>\varepsilon/2\})$, and set

$$Z(\varepsilon)_t = \sum_{q: T_q(\varepsilon)\le t} (f(1-\psi_\varepsilon))(\Delta X_{T_q(\varepsilon)})\ (c^{jk}_{T_q(\varepsilon)-} + c^{jk}_{T_q(\varepsilon)}).$$

For all $\omega\in\Omega$, $q\ge 1$, $\varepsilon'\in(0,\varepsilon)$ there is q' such that $T_q(\varepsilon)(\omega)=T_{q'}(\varepsilon')(\omega)$, whereas $1-\psi_\varepsilon$ increases to the indicator of $\mathbb{R}^d\backslash\{0\}$. Thus we obviously have (6.38).

d) It remains to prove (6.37). Fix $\varepsilon\in(0,1)$ and write $T_q=T_q(\varepsilon)$. Recall that for u different from all T_q's, we have $\|\Delta X_u\|\le\varepsilon/2$. Hence, for each ω and each $t>0$, we have the following properties for all n large enough: there is no T_q in $(0,k_n\Delta_n]$, nor in $(t-(k_n+1)\Delta_n,t]$; there is at most one T_q in an interval $((i-1)\Delta_n,i\Delta_n]$ with $i\Delta_n\le t$, and if this is not the case we have $\psi_\varepsilon(\Delta_i^n X)=1$. Hence for n large enough we have

$$Z(\varepsilon)_t = \sum_{q:\ k_n\Delta_n<T_q\le t-(k_n+1)\Delta_n} \zeta_q^n,$$

where

$$\zeta_q^n = \frac{1}{k_n}\,(f(1-\psi_\varepsilon))(\Delta_{i(n,q)}^n X) \sum_{l\in I'(n,q)} \beta_l^{n,j}\beta_l^{n,k},$$

and $i(n,q) = \inf(i : i\Delta_n \geq T_q)$ and $I'(n,q) = \{l : l \neq i(n,q), |l - i(n,q)| \leq k_n\}$.

To get (6.37) it is enough that $\zeta_q^n \stackrel{\mathbb{P}}{\longrightarrow} (f(1-\psi_\varepsilon))(\Delta X_{T_q})\,(c_{T_q-}^{jk} + c_{T_q}^{jk})$ for any q. Since $(f(1-\psi_\varepsilon))(\Delta_{i(n,q)}^n X) \to (f(1-\psi_\varepsilon))(\Delta X_{T_q})$ pointwise, it remains to prove that

$$\frac{1}{k_n}\sum_{l\in I'_-(n,q)} \beta_l^{n,j}\beta_l^{n,k} \stackrel{\mathbb{P}}{\longrightarrow} c_{T_q-}^{jk}, \qquad \frac{1}{k_n}\sum_{l\in I'_+(n,q)} \beta_l^{n,j}\beta_l^{n,k} \stackrel{\mathbb{P}}{\longrightarrow} c_{T_q}^{jk}. \tag{6.39}$$

where $I'_-(n,q)$ and $I'_+(n,q)$ are the subsets of $I'(n,q)$ consisting in those l smaller, respectively bigger, than $i(n,q)$. Letting $l(n,q)$ be the smallest l in $I'_-(n,q)$, we see that the left side of the first expression in (6.39) is $U_q^n + U_q'^n$, where

$$U_q^n = \sum_{r,s=1}^{d'} \sigma_{l(n,q)\Delta_n}^{jr}\sigma_{l(n,q)\Delta_n}^{ks}\overline{U}_n^q(r,s), \quad \overline{U}_q^n(r,s) = \frac{1}{k_n\Delta_n}\sum_{l\in I'_-(n,q)} \Delta_l^n W^r \Delta_l^n W^s,$$

$$U_q'^n = \sum_{r,s=1}^{d'} \frac{1}{k_n\Delta_n}\sum_{l\in I'_-(n,q)} (\sigma_{(l-1)\Delta_n}^{jr} - \sigma_{l(n,q)\Delta_n}^{jr})(\sigma_{(l-1)\Delta_n}^{ks} - \sigma_{l(n,q)\Delta_n}^{ks})\Delta_l^n W^r \Delta_l^n W^s.$$

On the one hand, the variables $\Delta_i^n W$ are i.i.d. $N(0,\Delta_n I_{d'})$, so $\overline{U}_q^n(r,s)$ is distributed as $1/k_n$ times the sum of k_n i.i.d. variables with the same law as $W_1^r W_1^s$, hence obviously $\overline{U}_q^n(r,s)$ converges in probability to 1 if $r=s$ and to 0 otherwise. Since $\sigma_{l(n,q)\Delta_n} \to \sigma_{T_q-}$, we deduce that $U_q^n \stackrel{\mathbb{P}}{\longrightarrow} c_{T_q-}^{jk}$.

On the other hand, due to (6.23) and by successive integrations we obtain

$$\mathbb{E}(|U_q'^n|) \;\leq\; \frac{1}{k_n}\sum_{l\in I'_-(n,q)} \mathbb{E}(\|\sigma_{(l-1)\Delta_n} - \sigma_{l(n,q)\Delta_n}\|^2) \leq K k_n \Delta_n$$

which goes to 0 by virtue of (6.12). Therefore we have proved the first part of (6.39), and the second part is proved in a similar way. □

Lemma 6.12 *If f is continuous and $f(\lambda x) = \lambda^p f(x)$ for all $\lambda > 0$, $x\in\mathbb{R}^d$ and some $p \geq 2$, we have $\Delta_n^{1-p/2}\widehat{D}_t'^n \stackrel{u.c.p.}{\longrightarrow} 2\int_0^t \rho_{\sigma_u}(f)c_u^{jk}\,du$.*

Proof. First we observe that by polarization, and exactly as in the proof of Theorem 6.3, it is enough to show the result when $j=k$, and of course when $f\geq 0$: then $\widehat{D}_t'^n$ is increasing in t, and $\int_0^t \rho_{\sigma_u}(f)c_u^{jk}\,du$ is also increasing and continuous. Then instead of proving the local uniform convergence it is enough to prove the convergence (in probability) for any given t.

With our assumptions on f, we have

$$\Delta_n^{1-p/2}\widehat{D}_t'^n \;=\; \frac{\Delta_n}{k_n}\sum_{i=1+k_n}^{[t/\Delta_n]-k_n}\ \sum_{l\in I_{n,t}(i)} f(\beta_i^n)\beta_i^{n,j}\beta_i^{n,k}.$$

Moreover, $\Delta_n \sum_{i=1+k_n}^{[t/\Delta_n]-k_n} \rho_{\sigma_{(i-1)\Delta_n}}(f) c^{jk}_{(i-1-k_n)\Delta_n} \stackrel{\mathbb{P}}{\longrightarrow} \int_0^t \rho_{\sigma_u}(f) c^{jk}_u \, du$ by Riemann integration. Therefore, it is enough to prove the following two properties:

$$\sum_{i=1+k_n}^{[t/\Delta_n]-k_n} \Delta_n (f(\beta_i^n) - \rho_{\sigma_{(i-1)\Delta_n}}(f)) \; c^{jk}_{(i-1-k_n)\Delta_n} \;\stackrel{\mathbb{P}}{\longrightarrow}\; 0, \tag{6.40}$$

$$Y_t^n := \frac{\Delta_n}{k_n} \sum_{i=1+k_n}^{[t/\Delta_n]-k_n} \sum_{l \in I_{n,t}(i)} \zeta^n_{i,l} \stackrel{\mathbb{P}}{\longrightarrow} 0, \quad \text{where} \;\; \zeta^n_{i,l} = f(\beta_i^n)(\beta_l^{n,j}\beta_l^{n,k} - c^{jk}_{(i-1-k_n)\Delta_n}). \tag{6.41}$$

Each summand, say ζ^n_i, in the left side of (6.40) is $\mathcal{F}_{i\Delta_n}$-measurable with $\mathbb{E}^n_{i-1}(\zeta^n_i) = 0$ and $\mathbb{E}^n_{i-1}((\zeta^n_i)^2) \leq K\Delta_n^2$ (apply (6.23) and recall that $|f(x)| \leq K\|x\|^r$ with our assumptions on f), so (6.40) follows from Lemma 4.1.

Proving (6.41) is a bit more involved. We set

$$\zeta'^n_{i,l} = f(\sigma_{(i-1-k_n)\Delta_n} \Delta_i^n W/\sqrt{\Delta_n}\,)\; (\beta_l^{n,j}\beta_l^{n,k} - c^{jk}_{(i-1-k_n)\Delta_n}), \quad Y'^n_t = \frac{\Delta_n}{k_n} \sum_{i=1+k_n}^{[t/\Delta_n]-k_n} \sum_{l \in I_{n,t}(i)} \zeta'^n_{i,l}$$

On the one hand, for any $l \in I_{n,t}(i)$ (hence either $l < i$ or $l > i$) and by successive integration we have

$$|\mathbb{E}^n_{i-1-k_n}(\zeta'^n_{i,l})| = |\rho_{\sigma_{(i-1-k_n)\Delta_n}}(f) \mathbb{E}^n_{i-1-k_n}(c^{jk}_{(l-1)\Delta_n} - c^{jk}_{(i-1-k_n)\Delta_n})| \leq K\sqrt{k_n\Delta_n}$$

by (6.23), the boundedness of σ and $|f(x)| \leq K\|x\|^r$. Moreover $\mathbb{E}^n_{i-1-k_n}((\zeta'^n_{i,l})^2) \leq K$ is obvious. Therefore, since $\mathbb{E}((Y'^n_t)^2)$ is Δ^2/k_n^2 times the sum of all $\mathbb{E}(\zeta'^n_{i,l}\; \zeta'^n_{i',l'})$ for all $1+k_n \leq i, i' \leq [t/\Delta_n] - k_n$ and $l \in I_{n,t}(i)$ and $l' \in I_{n,t}(i')$, by singling out the cases where $|i - i'| > 2k_n$ and $|i - i'| \leq 2k_n$ and in the first case by taking two successive conditional expectations, and in the second case by using Cauchy-Schwarz inequality, we obtain that

$$\mathbb{E}((Y'^n_t)^2) \;\leq\; K\frac{\Delta_n^2}{k_n^2}\left(4k_n^2[t/\Delta_n]^2(k_n\Delta_n) + 4k_n^2[t/\Delta_n]\right) \;\leq\; K(t^2 k_n\Delta_n + t\Delta_n) \;\to\; 0.$$

In order to get (6.41) it remains to prove that $Y^n_t - Y'^n_t \stackrel{\mathbb{P}}{\longrightarrow} 0$. By Cauchy-Schwarz inequality and (6.23), we have

$$\mathbb{E}(|\zeta^n_{i,l} - \zeta'^n_{i,l}|) \leq K\left(\mathbb{E}(|\rho_{\sigma_{(i-1)\Delta_n}}(f) - \rho_{\sigma_{(i-1-k_n)\Delta_n}}(f)|^2)\right)^{1/2}.$$

Then another application of Cauchy-Schwarz yields $\mathbb{E}(|Y^n_t - Y'^n_t|) \leq K_t\sqrt{\alpha_n(t)}$, where

$$\begin{aligned}
\alpha_n(t) &= \frac{\Delta_n}{k_n} \sum_{i=1+k_n}^{[t/\Delta_n]-k_n} \sum_{l \in I_{n,t}(i)} \mathbb{E}(|\rho_{\sigma_{(i-1)\Delta_n}}(f) - \rho_{\sigma_{(i-1-k_n)\Delta_n}}(f)|^2) \\
&= 2\Delta_n \sum_{i=1+k_n}^{[t/\Delta_n]-k_n} \mathbb{E}(|\rho_{\sigma_{(i-1)\Delta_n}}(f) - \rho_{\sigma_{(i-1-k_n)\Delta_n}}(f)|^2) \;\leq\; 2\int_0^t g_n(s)ds,
\end{aligned}$$

with the notation $g_n(s) = \mathbb{E}((\rho_{\sigma_{\Delta_n(k_n+[s\Delta_n])}}(f) - \rho_{\sigma_{\Delta_n[s/\Delta_n]}}(f))^2)$. Since c_t is bounded and f is with polynomial growth, we first have $g_n(s) \leq K$. Since further $t \mapsto \sigma_t$ has no fixed time of discontinuity and f is continuous and $\Delta_n k_n \to 0$, we next have $g_n(s) \to 0$ pointwise: hence $\alpha_n(t) \to 0$ and we have the result. □

Proof of (6.15). In view of Lemma 6.11 it is enough to prove that $\widehat{D}^n - D^n \stackrel{\text{u.c.p.}}{\longrightarrow} 0$, and this will obviously follow if we prove that

$$\sup_{i\neq l} \frac{1}{\Delta_n^2}\, \mathbb{E}(|f_n(\Delta_i^n X)\zeta_l^n|) \;\to\; 0 \qquad \text{as } \; n\to\infty, \tag{6.42}$$

where $\zeta_i^n = \Delta_l^n X^j \Delta_l^n X^k 1_{\{\|\Delta_l^n X\|\leq \alpha\Delta_n^{\varpi}\}} - \Delta_n \beta_l^{n,j}\beta_l^{n,k}$.

A simple computation shows that for $x,y\in\mathbb{R}^d$ and $\varepsilon>0$, we have

$$|(x_j+y_j)(x_k+y_k)1_{\{\|x+y\|\leq\varepsilon\}} - x_jx_k| \leq K\Big(\frac{1}{\varepsilon}\|x\|^3 + \|x\|(\|y\|\wedge\varepsilon) + \|y\|^2\wedge\varepsilon^2\Big).$$

We apply this to $x=\sqrt{\Delta_n}\,\beta_l^n$ and $y=\sqrt{\Delta_n}\,\chi_l^n$ and $\varepsilon = \alpha\Delta_n^{\varpi}$, and (6.23) and (6.27) with $\eta=\varepsilon$ and Cauchy-Schwarz inequality, to get

$$\mathbb{E}_{l-1}^n(|\zeta_l^n|) \leq K\Delta_n(\Delta_n^{1/2-\varpi} + \alpha_n)$$

for some α_n going to 0. On the other hand, (SH) implies that $\Delta_i^n X$ is bounded by a constant, hence (6.13) yields $|f_n(\Delta_i^n X)| \leq K\|\Delta_i^n X\|^2$ and (6.23) again gives $\mathbb{E}_{i-1}^n(|f_n(\Delta_i^n X)|) \leq K\Delta_n$. Then, by taking two successive conditional expectations, we get $\mathbb{E}(|f_n(\Delta_i^n X)\zeta_l^n|) \leq K\Delta_n^2(\Delta_n^{1/2-\varpi}+\alpha_n)$ as soon as $l\neq i$, and (6.42) follows. □

Proof of (6.16). In view of Lemma 6.12 it is enough to prove that $\Delta_n^{1-r/2}(\widehat{D}'^n - D^n) \stackrel{\text{u.c.p.}}{\longrightarrow} 0$, when X is continuous and $f(\lambda x) = \lambda^r f(x)$ for some $r>2$. With the notation ζ_i^n of the previous proof, this amounts to prove the following two properties:

$$\sup_{i\neq l}\frac{1}{\Delta_n^{1+r/2}}\,\mathbb{E}(|f(\Delta_i^n X)\zeta_l^n|) \;\to\; 0 \qquad \text{as } \; n\to\infty, \tag{6.43}$$

$$\sup_{i\neq l}\frac{1}{\Delta_n^{r/2}}\,\mathbb{E}\Big(|f(\Delta_i^n X)|1_{\{\|\Delta_i^n X\|>\alpha\Delta_n^{\varpi}\}}\|\beta_l^n\|\Big) \;\to\; 0 \qquad \text{as } \; n\to\infty. \tag{6.44}$$

Since X is continuous and $|f(x)|\leq K\|x\|^r$, we have $\mathbb{E}_{i-1}^n(|f_n(\Delta_i^n X)|) \leq K\Delta_n^{r/2}$, hence the proof of (6.43) is like in the previous proof. By Bienaymé-Tchebycheff inequality and (6.23) we also have $\mathbb{E}_{i-1}^n(|f(\Delta_i^n X)|1_{\{\|\Delta_i^n X\|>\alpha\Delta_n^{\varpi}\}}) \leq K_q\Delta_n^q$ for any $q>0$, hence (6.44) follows. □

7 A first CLT

As we have seen after (3.14), we have the CLT (3.7) when X is the sum of a Wiener process and a compound Poisson process, as soon as the function f in $V'(f,\Delta_n)$ satisfies

$f(x)/|x|^p \to 0$ as $|x| \to \infty$, for some $p < 1$. In this section we prove the same result, and even a bit more (the stable convergence in law) when X satisfies (H).

In other words, we are concerned with the CLT associated with Theorem 6.2. For statistical purposes we need a CLT when the function $f = (f_1, \cdots, f_q)$ is multidimensional: in this case, $V'(f, k, \Delta_n)$ is also multidimensional, with components $V'(f_j, k, \Delta_n)$. On the other hand, we will strongly restrict the class of functions f for which we give a CLT: although much more general situations are available, they also are much more complicated and will not be used in the sequel. Let us however mention that the present setting does not allow to consider the CLT for multipower variations in the interesting cases like in (6.10): for this, we refer to [8] when X is continuous, and to [9] when X is a discontinuous Lévy process. For discontinuous semimartingales which are not Lévy processes, essentially nothing is known as far as CLTs are concerned, for processes like (6.10).

One of the difficulties of this question is to characterize the limit, and more specifically the quadratic variation of the limiting process. To do this, we consider a sequence $(U_i)_{i \geq 1}$ of independent $\mathcal{N}(0, I_d)$ variables (they take values in $\mathbb{R}^d$, and I_d is the unit $d \times d$ matrix). Recall that ρ_σ, defined before (6.3), is also the law of σU_1, and so $\rho_\sigma(g) = \mathbb{E}(g(\sigma U_1))$. In a similar way, for any q-dimensional function $f = (f_1, \cdots, f_q)$ on $(\mathbb{R}^d)^k$, say with polynomial growth, we set for $i, j = 1, \cdots, q$:

$$\begin{aligned} R_\sigma^{ij}(f,k) \;=\; & \sum_{l=-k+1}^{k-1} \mathbb{E}\Big(f_i(\sigma U_k, \cdots, \sigma U_{2k-1}) f_j(\sigma U_{l+k}, \cdots, \sigma U_{l+2k-1}) \Big) \\ & - (2k-1)\mathbb{E}(f_i(\sigma U_1, \cdots, \sigma U_k))\mathbb{E}(f_j(\sigma U_1, \cdots, \sigma U_k)). \end{aligned} \tag{7.1}$$

One can of course express this in terms of integrals of f with respect to the measures ρ_σ and their tensor powers, but this is very complicated. Let us just mention the special case where $k = 1$:

$$R_\sigma^{ij}(f,1) \;=\; \rho_\sigma(f_i f_j) - \rho_\sigma(f_i)\rho_\sigma(f_j). \tag{7.2}$$

The main result goes as follows:

Theorem 7.1 *Assume (H). Let f be a q-dimensional function on $(\mathbb{R}^d)^k$ for some $k \geq 1$, which is even in each argument, that is*

$$f(x_1, \cdots, x_{l-1}, -x_l, x_{l+1}, \cdots, x_k) = f(x_1, \cdots, x_{l-1}, x_l, x_{l+1}, \cdots, x_k)$$

identically for all l. In the following two cases:

a) X is continuous, and f is C^1 with derivatives having polynomial growth,

b) f is C_b^1 (bounded with first derivatives bounded), and $\int(\gamma(x) \wedge 1)\lambda(dx) < \infty$ (hence the jumps of X are summable over each finite interval),

the q-dimensional processes

$$\frac{1}{\sqrt{\Delta_n}} \Big(\Delta_n V'(f, k, \Delta_n)_t - \int_0^t \rho_{\sigma_u}^{\otimes k}(f) du \Big)$$

converge stably in law to a continuous process $V'(f, k)$ defined defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of the space $(\Omega, \mathcal{F}, \mathbb{P})$, which conditionally on the σ-field $\mathcal{F}$ is a centered Gaussian

$\mathbb{R}^q$*-valued process with independent increments, satisfying*

$$\widetilde{\mathbb{E}}(V'(f_i,k)_t V'(f_j,k)_t) \;=\; \int_0^t R^{ij}_{\sigma_u}(f,k)du. \tag{7.3}$$

Another, equivalent, way to characterize the limiting process $V'(f,k)$ is as follows, see [13]: for each σ, the matrix $R_\sigma(f,k)$ is symmetric nonnegative, so we can find a square-root $S_\sigma(f,k)$, that is $S_\sigma(f,k)S_\sigma(f,k)^\star = R_\sigma(f,k)$, which as a function of σ is measurable. Then there exists a q-dimensional Brownian motion $B = (B^i)_{i\leq q}$ on an extension of the space $(\Omega,\mathcal{F},\mathbb{P})$, independent of $\mathcal{F}$, and $V'(f,k)$ is given componentwise by

$$V'(f_i,k)_t \;=\; \sum_{j=1}^q \int_0^t S^{ij}_{\sigma_u}(f,k)dB^j_u. \tag{7.4}$$

As a consequence we obtain a CLT for estimating C^{jk}_t when X is continuous. It suffices to apply the theorem with $k=1$ and the $d\times d$-dimensional function f with components $f_{jl}(x) = x^j x^k$. Upon a simple calculation using (7.2) in this case, we obtain:

Corollary 7.2 *Assume (H)* (or (H') only,although it is not then a consequence of the previous theorem) *and that X is continuous. Then the $d\times d$-dimensional process with components*

$$\frac{1}{\sqrt{\Delta_n}}\Big(\sum_{i=1}^{[t/\Delta_n]} \Delta^n_i X^j \Delta^n_i X^k - C^{jk}_t\Big)$$

converge stably in law to a continuous process $(V^{jk})_{1\leq j,k\leq d}$ defined defined on an extension $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}})$ of the space $(\Omega,\mathcal{F},\mathbb{P})$, which conditionally on the σ-field $\mathcal{F}$ is a centered Gaussian $\mathbb{R}^q$-valued process with independent increments, satisfying

$$\widetilde{\mathbb{E}}(V^{jk}_t V^{j'k'}_t) \;=\; \int_0^t (c^{jk'}_u c^{j'k}_u + c^{jj'}_u c^{kk'}_u)du. \tag{7.5}$$

It turns out that this result is very special: Assumption (H) is required for Theorem 7.1, essentially because one needs that $\Delta_n \sum_{i=1}^{[t/\Delta_n]} \rho_{\sigma_{(i-1)\Delta_n}}(g)$ converges to $\int_0^t \rho_{\sigma_s}(g)ds$ at a rate faster than $1/\sqrt{\Delta_n}$, and this necessitates strong assumptions on σ (instead of assuming that it is an Itô semimartingale, as in (H), one could require some Hölder continuity of its paths, with index bigger than 1/2). However, for the corollary, and due to the quadratic form of the test function, some cancelations occur which allow to obtain the result under the weaker assumption (H') only. Although this is a theoretically important point, it is not proved here.

There is a variant of Theorem 7.1 which concerns the case where in (6.3) on take the sum over the i's that are multiple of k. More precisely we set

$$V''(f,k,\Delta_n)_t \;=\; \sum_{i=1}^{[t/k\Delta_n]} f\Big(\Delta^n_{(i-1)k+1}X/\sqrt{\Delta_n},\cdots,\Delta^n_{ik}X/\sqrt{\Delta_n}\Big). \tag{7.6}$$

The LLN is of course exactly the same as Theorem 6.2, except that the limit should be divided by k in (6.5). As for the CLT, it runs as follows (and although similar to Theorem 7.1 i is not a direct consequence):

Theorem 7.3 *Under he same assumptions than in Theorem 7.1, the q-dimensional processes*

$$\frac{1}{\sqrt{\Delta_n}}\,\Big(\Delta_n V''(f,k,\Delta_n)_t-\frac{1}{k}\int_0^t\rho_{\sigma_u}^{\otimes k}(f)du\Big)$$

converge stably in law to a continuous process $V''(f,k)$ defined defined on an extension $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}})$ of the space $(\Omega,\mathcal{F},\mathbb{P})$, which conditionally on the σ-field $\mathcal{F}$ is a centered Gaussian $\mathbb{R}^q$-valued process with independent increments, satisfying

$$\widetilde{\mathbb{E}}(V''(f_i,k)_tV''(f_j,k)_t\mid\mathcal{F})\;=\;\frac{1}{k}\int_0^t\Big(\rho_{\sigma_u}^{\otimes k}(f_if_j)-\rho_{\sigma_u}^{\otimes k}(f_i)\rho_{\sigma_u}^{\otimes k}(f_j)\Big)du. \tag{7.7}$$

Theorem 7.1 does not allow to deduce a CLT associated with Theorem 6.4, since the function f which is used in (6.10) cannot meet the assumptions above. Nevertheless such a CLT is available when X is continuous: see [8], under the (weak) additional assumption that $\sigma_t\sigma_t^\star$ is everywhere invertible. When X is discontinuous and with the additional assumption that $\int(\gamma(x)\wedge 1)\lambda(dx)<\infty$, it is also available, see [9] for the case when in addition X is a Lévy process.

We do however give the CLT associated with Theorem 6.3, although it is not a direct consequence of the previous one.

Theorem 7.4 *Assume (H), and also that X is continuous or that $\int(\gamma(x)^r\wedge 1)\lambda(dx)<\infty$ for some $r\in[0,1)$. Then for all $\varpi\in[\frac{1}{2(2-r)},\frac{1}{2})$ and $\alpha>0$ the $d\times d$-dimensional process with components*

$$\frac{1}{\sqrt{\Delta_n}}\,\Big(V^{jk}(\varpi,\alpha,\Delta_n)_t-C_t^{jk}\Big)$$

converge stably in law to the continuous process $(V^{jk})_{1\le j,k\le d}$ defined in Corollary 7.2.

In the discontinuous case, this is not fully satisfactory since we need the assumption about $r<1$, which we a priori do not know to hold, and further ϖ has to be bigger than $\frac{1}{2(2-r)}$. In the continuous case for X the assumption is simply (H), but of course in this case there is no reason to prefer the estimators $V^{jk}(\varpi,\alpha,\Delta_n)_t$ to $\sum_{i=1}^{[t/\Delta_n]}\Delta_i^nX^j\Delta_i^nX^k$.

7.1 The scheme of the proof of Theorem 7.1.

This theorem is rather long to prove, and quite technical. We first describe here the main steps. Note that the localization argument expounded earlier works here as well, so we can and will assume (SH) instead of (H), without special mention. Also, the multidimensional case for f reduces to the 1-dimensional one by polarization, as in the proof of Theorem 6.3, so below we suppose that f is 1-dimensional (that is, $q=1$). These assumptions are in force through the remainder of this section. We also denote by $\mathcal{M}'$ the set of all $d'\times d'$ matrices bounded by K where K is a bound for the process $\|\sigma_t\|$.

We use the notation

$$\zeta_i^n=f(\Delta_i^nX/\sqrt{\Delta_n},\cdots,\Delta_{i-k}^nX/\sqrt{\Delta_n}),\quad \zeta_i'^n=f(\beta_{i,0}^n,\beta_{i,1}^n,\cdots,\beta_{i,k-1}^n),\quad \zeta_i''^n=\zeta_i^n-\zeta_i'^n.$$

First, we replace each normalized increment $\Delta_{i+l}^nX/\sqrt{\Delta_n}$ in (6.3) by $\beta_{i,l}^n$ (notation (6.22)): this is of course much simpler, and we have the following:

Proposition 7.5 *The processes*

$$\overline{V}^n_t \;=\; \sqrt{\Delta_n}\,\sum_{i=1}^{[t/\Delta_n]}\Big(\zeta'^n_i - \rho^{\otimes k}_{\sigma_{(i-1)\Delta_n}}(f)\Big) \tag{7.8}$$

converge stably in law to the process $V'(f,k)$, as defined in Theorem 7.1.

Next, we successively prove the following three properties:

$$\sqrt{\Delta_n}\,\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}^n_{i-1}(\zeta''^n_i)\;\stackrel{\text{u.c.p.}}{\longrightarrow}\;0, \tag{7.9}$$

$$\sqrt{\Delta_n}\,\sum_{i=1}^{[t/\Delta_n]}\Big(\zeta''^n_i-\mathbb{E}^n_{i-1}(\zeta''^n_i)\Big)\;\stackrel{\text{u.c.p.}}{\longrightarrow}\;0, \tag{7.10}$$

$$\frac{1}{\sqrt{\Delta_n}}\Big(\Delta_n\sum_{i=1}^{[t/\Delta_n]}\rho^{\otimes k}_{\sigma_{(i-1)\Delta_n}}(f)-\int_0^t\rho^{\otimes k}_{\sigma_u}(f)du\Big)\;\stackrel{\text{u.c.p.}}{\longrightarrow}\;0. \tag{7.11}$$

Obviously our theorem is a consequence of these three properties and of Proposition 7.5. Apart from (7.10), which is a simple consequence of Lemma 6.9, all these steps are non trivial, and the most difficult is (7.9).

7.2 Proof of (7.10).

We use the notation $I(n,t,l)$ of the proof of Theorem 6.2, and it is of course enough to prove

$$\sqrt{\Delta_n}\sum_{i\in I(n,t,l)}\Big(\zeta''^n_i-\mathbb{E}^n_{i-1}(\zeta''^n_i)\Big)\;\stackrel{\text{u.c.p.}}{\longrightarrow}\;0$$

for each $l=0,\cdots,k-1$. Since each ζ''^n_i is $\mathcal{F}_{(i+k-1)\Delta_n}$-measurable, by Lemma 4.1 it is even enough to prove that

$$\Delta_n\sum_{i\in I(n,t,l)}\mathbb{E}^n_{i-1}((\zeta''^n_i)^2)\;\stackrel{\text{u.c.p.}}{\longrightarrow}\;0.$$

But this is a trivial consequence of Lemma 6.9 applied with $q=2$ and $l=0$: in case (a) the function f obviously satisfies (6.4) for some $r\geq 0$ and X is continuous, whereas in case (b) it satisfies (6.4) with $r=0$.

7.3 Proof of (7.11).

Let us consider the function $g(\sigma)=\rho^{\otimes k}_\sigma(f)$, defined on the set $\mathcal{M}'$. (7.11) amounts to

$$\sum_{i=1}^{[t/\Delta_n]}\eta^n_i\;\stackrel{\text{u.c.p.}}{\longrightarrow}\;0,\qquad\text{where}\quad\eta^n_i=\frac{1}{\sqrt{\Delta_n}}\int_{(i-1)\Delta_n}^{i\Delta_n}(g(\sigma_u)-g(\sigma_{(i-1)\Delta_n}))du. \tag{7.12}$$

Since f is at least C^1 with derivatives having polynomial growth, the function g is C^1_b on $\mathcal{M}$. However, the problem here is that σ may have jumps, and even when it is continuous its paths are typically Hölder with index $\alpha > 1/2$, but nor $\alpha = 1/2$: so (7.12) is not trivial.

With ∇g denoting the gradient of g (a $d \times d'$-dimensional function), we may write $\eta^n_i = \eta'^n_i + \eta''^n_i$ where (with matrix notation)

$$\eta'^n_i = \frac{1}{\sqrt{\Delta_n}}\ \nabla g(\sigma_{(i-1)\Delta_n}) \int_{(i-1)\Delta_n}^{i\Delta_n} (\sigma_u - \sigma_{(i-1)\Delta_n})\ du,$$

$$\eta''^n_i = \frac{1}{\sqrt{\Delta_n}} \int_{(i-1)\Delta_n}^{i\Delta_n} \big(g(\sigma_u) - g(\sigma_{(i-1)\Delta_n}) - \nabla g(\sigma_{(i-1)\Delta_n})(\sigma_u - \sigma_{(i-1)\Delta_n})\big)\ du.$$

In view of (6.21) we can decompose further η'^n_i as $\eta'^n_i = \mu^n_i + \mu'^n_i$, where

$$\mu^n_i = \frac{1}{\sqrt{\Delta_n}}\ \nabla g(\sigma_{(i-1)\Delta_n}) \int_{(i-1)\Delta_n}^{i\Delta_n} du \int_{(i-1)\Delta_n}^{u} \widetilde{b}_s ds,$$

$$\begin{aligned}\mu'^n_i \ &= \ \frac{1}{\sqrt{\Delta_n}}\ \nabla g(\sigma_{(i-1)\Delta_n}) \int_{(i-1)\Delta_n}^{i\Delta_n} du \left(\int_{(i-1)\Delta_n}^{u} \widetilde{\sigma}_s dW_s \right. \\ &\qquad \left. + \int_{(i-1)\Delta_n}^{u} \int \widetilde{\delta}(s,x)(\underline{\mu} - \underline{\nu})(ds,dx)\right).\end{aligned}$$

On the one hand, we have $|\mu^n_i| \leq K\Delta_n^{3/2}$ (recall that g is C^1_b and $\widetilde{b}$ is bounded), so $\sum_{i=1}^{[t/\Delta_n]} \mu^n_i \stackrel{\text{u.c.p.}}{\longrightarrow} 0$. On the other hand, we have $\mathbb{E}^n_{i-1}(\mu'^n_i) = 0$ and $\mathbb{E}^n_{i-1}((\mu'^n_i)^2) \leq K\Delta_n^2$ by Doob and Cauchy–Schwarz inequalities, hence $\sum_{i=1}^{[t/\Delta_n]} \mu'^n_i \stackrel{\text{u.c.p.}}{\longrightarrow} 0$ by Lemma 4.1.

Finally since g is C^1_b on the compact set $\mathcal{M}$ we have $|g(\sigma') - g(\sigma) - \nabla g(\sigma)(\sigma' - \sigma)| \leq K\|\sigma' - \sigma\| h(\|\sigma' - \sigma\|)$ for all $\sigma, \sigma' \in \mathcal{M}$, where $h(\varepsilon) \to 0$ as $\varepsilon \to 0$. Therefore

$$\begin{aligned}|\eta''^n_i| \ &\leq \ \frac{1}{\sqrt{\Delta_n}} \int_{(i-1)\Delta_n}^{i\Delta_n} h(\|\sigma_u - \sigma_{(i-1)\Delta_n}\|)\ \|\sigma_u - \sigma_{(i-1)\Delta_n}\|\ du \\ &\leq \ \frac{1}{\sqrt{\Delta_n}}\ h(\varepsilon) \int_{(i-1)\Delta_n}^{i\Delta_n} \|\sigma_u - \sigma_{(i-1)\Delta_n}\|\ du + \frac{K}{\varepsilon\sqrt{\Delta_n}} \int_{(i-1)\Delta_n}^{i\Delta_n} \|\sigma_u - \sigma_{(i-1)\Delta_n}\|^2\ du.\end{aligned}$$

Since $h(\varepsilon)$ is arbitrarily small we deduce from the above and from (6.23) that $\sum_{i=1}^{[t/\Delta_n]} \mathbb{E}(|\eta''^n_i|) \to 0$. This clearly finishes to prove (7.12).

7.4 Proof of Proposition 7.5.

We prove the result when $k = 2$ only. The case $k \geq 3$ is more tedious but similar.

Letting $g_t(x) = \int \rho_{\sigma_t}(dy) f(x,y)$, we have $\overline{V}^n(f)_t = \sum_{i=2}^{[t/\Delta_n]+1} \eta^n_i + \gamma'^n_1 - \gamma'^n_{[t/\Delta_n]+1}$, where $\eta^n_i = \gamma^n_i + \gamma'^n_i$ and

$$\begin{aligned}\gamma^n_i \ &= \ \sqrt{\Delta_n}\ \Big(f(\beta^n_{i-1,0}, \beta^n_{i-1,1}) - \int \rho_{\sigma_{(i-2)\Delta_n}}(dx) f(\beta^n_{i-1,0}, x)\Big), \\ \gamma'^n_i \ &= \ \sqrt{\Delta_n}\ \Big(\int \rho_{\sigma_{(i-1)\Delta_n}}(dx) f(\beta^n_{i,0}, x) - \rho^{\otimes 2}_{\sigma_{(i-1)\Delta_n}}(f)\Big).\end{aligned}$$

Since obviously $\mathbb{E}(|\gamma_i'^n|) \leq K\sqrt{\Delta_n}$, it is enough to prove that $\overline{V}'^n(f)_t = \sum_{i=2}^{[t/\Delta_n]+1} \eta_i^n$ converges stably in law to the process $V'(f,2)$.

Note that η_i^n is $\mathcal{F}_{i\Delta_n}$-measurable, and a (tedious) calculation yields

$$\mathbb{E}_{i-1}^n(\eta_i^n) = 0, \qquad \mathbb{E}_{i-1}^n((\eta_i^n)^2) = \Delta_n \phi_i^n, \qquad \mathbb{E}_{i-1}^n(|\eta_i^n|^4) \leq K\Delta_n^2, \tag{7.13}$$

where $\phi_i^n = g((i-2)\Delta_n, (i-1)\Delta_n, \beta_{i-1}^n)$ and

$$\begin{aligned} g(s,t,x) &= \int \rho_{\sigma_t}(dy)\Big(f(x,y)^2 + \Big(\int \rho_{\sigma_t}(dz) f(y,z)\Big)^2\Big) - \Big(\int \rho_{\sigma_t}(dy) f(x,y)\Big)^2 \\ &\quad -\Big(\rho_{\sigma_t}^{\otimes 2}(f)\Big)^2 - 2\rho_{\sigma_t}^{\otimes 2}(f)\int \rho_{\sigma_s}(dy) f(x,y) + 2\int \rho(dy)\rho(dz) f(x,\sigma_s y) f(\sigma_t y, \sigma_t z) \end{aligned}$$

(here, ρ is the law $\mathcal{N}(0, I_{d'})$). Then if we can prove the following two properties:

$$\sum_{i=2}^{[t/\Delta_n]+1} \mathbb{E}_{i-1}^n(\Delta_i^n N\ \eta_i^n) \stackrel{\mathbb{P}}{\longrightarrow} 0 \tag{7.14}$$

for any N which is a component of W or is a bounded martingale orthogonal to W, and

$$\Delta_n \sum_{i=2}^{[t/\Delta_n]+1} \phi_i^n \stackrel{\mathbb{P}}{\longrightarrow} \int_0^t R_{\sigma_u}(f,2)du \tag{7.15}$$

(with the notation (7.1); here f is 1-dimensional, so $R_\sigma(f,2)$ is also 1-dimensional), then Lemma 4.4 will yield the stable convergence in law of $\overline{V}'^n$ to $V'(f,2)$.

Let us prove first (7.14). Recall $\eta_i^n = \gamma_i^n + \gamma_i'^n$, and observe that

$$\begin{aligned} \gamma_i^n &= \sqrt{\Delta_n}\ h(\sigma_{(i-2)\Delta_n}, \Delta_{i-1}^n W/\sqrt{\Delta_n}, \Delta_i^n W/\sqrt{\Delta_n}) \\ \gamma_i'^n &= \sqrt{\Delta_n}\ h'(\sigma_{(i-1)\Delta_n}, \Delta_i^n W/\sqrt{\Delta_n}), \end{aligned}$$

where $h(\sigma,x,y)$ and $h'(\sigma,x)$ are continuous functions with polynomial growth in x an y, uniform in $\sigma \in \mathcal{M}'$. Then (7.14) when N is a bounded martingale orthogonal to W readily follows from Lemma 6.10.

Next, suppose that N is a component of W, say W^1. Since f is globally even and ρ_{σ_s} is a measure symmetric about the origin, the function $h'(\sigma,x)$ is even in x, so $h'(\sigma,x)x^1$ is odd in x and obviously $\mathbb{E}_{i-1}^n(\gamma_i'^n \Delta_i^n W^1) = 0$. So it remains to prove that

$$\sum_{i=2}^{[t/\Delta_n]+1} \zeta_i^n \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad \text{where } \zeta_i^n = \mathbb{E}_{i-1}^n(\gamma_i^n \Delta_i^n W^1). \tag{7.16}$$

An argument similar to the previous one shows that $h(\sigma,x,y)$ is globally even in (x,y), so ζ_i^n has the form $\Delta_n\ k(\sigma_{(i-2)\Delta_n}, \Delta_{i-1}^n W/\sqrt{\Delta_n})$ where $k(\sigma,x) = \int \rho_\sigma(dy) h(\sigma,x,y) y^1$ is odd in x, and also C^1 in x with derivatives with polynomial growth, uniformly in $\sigma \in \mathcal{M}'$. Then $\mathbb{E}_{i-2}^n(\zeta_i^n) = 0$ and $\mathbb{E}_{i-2}^n(|\zeta_i^n|^2) \leq K\Delta_n^2$. Since ζ_i^n is also $\mathcal{F}_{(i-1)\Delta_n}$-measurable, we deduce (7.16) from Lemma 4.1, and we have finished the proof of (7.14).

Now we prove (7.15). Observe that ϕ_i^n is $\mathcal{F}_{(i-1)\Delta_n}$-measurable and

$$\mathbb{E}_{i-2}^n(\phi_i^n) \ = \ h((i-2)\Delta_n,(i-1)\Delta_n), \qquad \mathbb{E}_{i-2}^n(|\phi_i^n|^2) \ \leq \ K,$$

where $h(s,t)=\int \rho_{\sigma_s}(dx)g(s,t,x)$. Then, by Lemma 4.1,the property (7.15) follows from

$$\Delta_n \sum_{i=1}^{[t/\Delta_n]} h((i-1)\Delta_n, i\Delta_n) \ \stackrel{\mathbb{P}}{\longrightarrow} \ \int_0^t R_{\sigma_u}(f,2)du, \tag{7.17}$$

so it remains to show (7.17). On the one hand we have $|h(s,t)| \leq K$. On the other hand, since f is continuous with polynomial growth and σ_t is bounded we clearly have $h(s_n,t_n)\to h(t,t)$ for any sequences $s_n,t_n\to t$ which are such that σ_{s_n} and σ_{t_n} converge to σ_t: since the later property holds, for $\mathbb{P}$-almost all ω and Lebesgue-almost all t, for all sequences $s_n,t_n\to t$, we deduce that

$$\Delta_n \sum_{i=1}^{[t/\Delta_n]} h((i-1)\Delta_n, i\Delta_n) \ \stackrel{\mathbb{P}}{\longrightarrow} \ \int_0^t h(u,u)du.$$

Since

$$h(t,t) \ = \ \rho_{\sigma_t}^{\otimes 2}(f^2) - 3\Big(\rho_{\sigma_t}^{\otimes 2}(f)\Big)^2 + 2\int \rho_{\sigma_t}(dx)\rho_{\sigma_t}(dy)\rho_{\sigma_t}(dz)f(x,y)f(y,z),$$

is trivially equal to $R_{\sigma_t}(f,2)$, as given by (7.1). Hence we have (7.17).

7.5 Proof of (7.9).

As said before, this is the hard part, and it is divided into a number of steps.

Step 1. For $l=0,\cdots,k-1$ we define the following (random) functions on $\mathbb{R}^d$:

$$g_{i,l}^n(x) \ = \ \int f\Big(\frac{\Delta_i^n X}{\sqrt{\Delta_n}},\cdots,\frac{\Delta_{i+l-1}^n X}{\sqrt{\Delta_n}},x,x_{l+1},\cdots,x_{k-1}\Big) \ \rho_{\sigma_{(i-1)\Delta_n}}^{\otimes(k-l-1)}(dx_{l+1},\cdots,x_{k-1})$$

(for $l=0$ we simply integrate $f(x,x_{l+1},\cdots,x_{k-1})$, whereas for $l=k-1$ we have no integration). As a function of ω this is $\mathcal{F}_{(i+l-1)\Delta_n}$-measurable. As a function of x it is C^1, and further it has the following properties, according to the case (a) or (b) of Theorem 7.1 (we heavily use the fact that σ_t is bounded, and also (6.23)):

$$\left.\begin{array}{l} |g_{i,l}^n(x)|+\|\nabla g_{i,l}^n(x)\| \ \leq \ KZ_{i,l}^n(1+\|x\|^r) \qquad \text{where} \\ \text{in case (a):} \ \ r\geq 0, \ \ \mathbb{E}_{i-1}^n(|Z_{i,l}^n|^p)\leq K_p \ \ \forall p>0, \ \ Z_{i,l}^n \ \text{ is } \mathcal{F}_{(i+l-2)\Delta_n}\text{-measurable} \\ \text{in case (b):} \ \ r=0, \ \ Z_{i,l}^n=1. \end{array}\right\} \tag{7.18}$$

For all $A\geq 1$ there is also a positive function $G_A(\varepsilon)$ tending to 0 as $\varepsilon\to 0$, such that with $Z_{i,l}^n$ as above:

$$\|x\|\leq A, \ Z_{i,l}^n\leq A, \ \|y\|\leq\varepsilon \quad\Rightarrow\quad \|\nabla g_{i,l}^n(x+y)-\nabla g_{i,l}^n(x)\| \ \leq \ G_A(\varepsilon). \tag{7.19}$$

Observing that $\zeta_i''^n$ is the sum over l from 0 to $k-1$ of

$$f\Big(\frac{\Delta_i^n X}{\sqrt{\Delta_n}},\cdots,\frac{\Delta_{i+l}^n X}{\sqrt{\Delta_n}},\beta_{i,l+1}^n,\cdots,\beta_{i,k-1}^n\Big)-f\Big(\frac{\Delta_i^n X}{\sqrt{\Delta_n}},\cdots,\frac{\Delta_{i+l-1}^n X}{\sqrt{\Delta_n}},\beta_{i,l}^n,\cdots,\beta_{i,k-1}^n\Big),$$

we have

$$\mathbb{E}_{i-1}^n(\zeta_i''^n) \;=\; \sum_{l=0}^{k-1}\mathbb{E}_{i-1}^n\Big(g_{i,l}^n(\Delta_{i+l}^n X/\sqrt{\Delta_n})-g_{i,l}^n(\beta_{i,l}^n)\Big).$$

Therefore it is enough to prove that for any $l\geq 0$ we have

$$\sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n\Big(g_{i,l}^n(\Delta_{i+l}^n X/\sqrt{\Delta_n})-g_{i,l}^n(\beta_{i,l}^n)\Big)\;\stackrel{\text{u.c.p.}}{\longrightarrow}\;0. \tag{7.20}$$

Step 2. In case (b) the process X has jumps, but we assume that $\int(\gamma(x)\wedge 1)\lambda(dx)<\infty$, hence the two processes $\delta\star\underline{\mu}$ and $\delta\star\underline{\nu}$ are well defined. Moreover (7.18) readily gives $|g_{i,l}^n(x+y)-g_{i,l}^n(x)|\leq K(\|y\|\wedge 1)$. Hence it follows from (6.26) with $\alpha_n=1$ that

$$\sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n\Big(g_{i,l}^n(\Delta_{i+l}^n X/\sqrt{\Delta_n})-g_{i,l}^n(\Delta_{i+l}^n X/\sqrt{\Delta_n}-\Delta_{i+l}^n(\delta\star\underline{\mu})/\sqrt{\Delta_n})\Big)\;\stackrel{\text{u.c.p.}}{\longrightarrow}\;0.$$

Therefore if we put

$$\xi_{i,l}^n \;=\; \begin{cases}\Delta_{i+l}^n X/\sqrt{\Delta_n}-\beta_{i,l}^n & \text{in case (a)}\\ \Delta_{i+l}^n X/\sqrt{\Delta_n}-\Delta_{i+l}^n(\delta\star\underline{\mu})/\sqrt{\Delta_n}-\beta_{i,l}^n & \text{in case (b),}\end{cases} \tag{7.21}$$

(7.20) amounts to

$$\sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n\Big(g_{i,l}^n(\beta_{i,l}^n+\xi_{i,l}^n)-g_{i,l}^n(\beta_{i,l}^n)\Big)\;\stackrel{\text{u.c.p.}}{\longrightarrow}\;0. \tag{7.22}$$

Step 3. At this stage, we set (for simplicity, in the forthcoming formulas we write $S=S(i,l,n)=(i+l-1)\Delta_n$ and $T=T(i,l,n)=(i+l)\Delta_n$; recall that $x\mapsto\delta(s,x)$ is λ-integrable (in case (a) because then $\delta\equiv 0$, in case (b) because $|\delta(s,.)|\leq K(\gamma\wedge 1)$):

$$\begin{aligned}\widehat{\xi}_{i,l}^n &= \int_S^T\Big(b_s-b_S+\int_E(\delta(s,x)-\delta(S,x))\lambda(dx)\Big)ds\\ &\quad+\int_S^T\Big(\int_S^s(\widetilde{b}_u du+(\widetilde{\sigma}_u-\widetilde{\sigma}_S)dW_u)+\int_S^s\int_E(\widetilde{\delta}(u,x)-\widetilde{\delta}(S,x))(\underline{\mu}-\underline{\nu})(du,dx)\Big)dW_s\\ \widetilde{\xi}_{i,l}^n &= \Big(b_S+\int_E\delta(S,x)\lambda(dx)\Big)\Delta_n+\int_S^T\Big(\widetilde{\sigma}_S\int_S^s dW_u+\int_S^s\int\widetilde{\delta}(S,x)(\underline{\mu}-\underline{\nu})(du,dx)\Big)dW_s\end{aligned}$$

In view of (7.21), we obviously have $\xi_{i,l}^n=\Big(\widehat{\xi}_{i,l}^n+\widetilde{\xi}_{i,l}^n\Big)/\sqrt{\Delta_n}$.

Consider the process $Y=(\widetilde{\gamma}^2\wedge 1)\star\underline{\mu}$. This is an increasing pure jump Lévy process, whose Laplace transform is

$$u\;\mapsto\;\mathbb{E}(e^{-u(Y_{s+t}-Y_s)})\;=\;\exp\,t\int\Big(e^{-u(\widetilde{\gamma}(x)^2\wedge 1}-1\Big)\;\lambda(dx).$$

If q is a non zero integer, we compute the qth moment of $Y_{s+t}-Y_s$ by differentiating q times its Laplace transform at 0: this is the sum, over all choices $p_1,\dots,p_k$ of positive integers with $\sum_{i=1}^k p_i = q$, of suitable constants times the product for all $i=1,\dots,k$ of the terms $t\int(\widetilde{\gamma}(x)^{2p_i}\wedge 1)\lambda(dx)$, each one being smaller than Kt. Then we deduce that $\mathbb{E}((Y_{s+t}-Y_s)^q \mid \mathcal{F}_s)\leq K_q t$, and by interpolation this also holds for any real $q\geq 1$.

Then, coming back to the definition of $\widehat{\xi}^n_{i,l}$ and $\widetilde{\xi}^n_{i,l}$, and using the properties $\|\delta(t,x)\|\leq K(\gamma(x)\wedge 1)$ and $\|\widetilde{\delta}(t,x)\|\leq K(\widetilde{\gamma}(x)\wedge 1)$, plus the fact that $\int(\gamma(x)\wedge 1)\lambda(dx)<\infty$ when δ is not identically 0, and the boundedness of $b,\widetilde{b},\sigma,\widetilde{\sigma}$, we deduce from Burkholder-Davis-Gundy and Hölder inequalities that

$$q\geq 2 \quad\Rightarrow\quad \mathbb{E}^n_{i+l-1}(|\widehat{\xi}^n_{i,l}|^q)+\mathbb{E}^n_{i+l-1}(|\widetilde{\xi}^n_{i,l}|^q) \;\leq\; K\Delta_n^{1+q/2}. \tag{7.23}$$

The same arguments, plus Cauchy-Schwarz inequality, yield that with notation

$$\begin{aligned}\alpha^n_{i,l} \;=\; \mathbb{E}^n_{i+l-1}\Big(\int_S^T \Big(&\|b_s-b_S\|^2+\|\widetilde{\sigma}_s-\widetilde{\sigma}_S\|^2+\int\|\widetilde{\delta}(s,x)-\widetilde{\delta}(S,x)\|^2\lambda(dx)\\ &+\int\|\delta(s,x)-\delta(S,x)\|\lambda(dx)\Big)ds\Big),\end{aligned}$$

then

$$\mathbb{E}^n_{i+l-1}(|\widehat{\xi}^n_{i,l}|^2) \;\leq\; K\Delta_n\Big(\Delta_n^2+\alpha^n_{i,l}\Big). \tag{7.24}$$

Next, since the restriction of $\underline{\mu}$ to $(S,\infty)\times E$ and the increments of W after time S are independent, conditionally on $\mathcal{F}'_S=\mathcal{F}_S\vee\sigma(W_t: t\geq 0)$, we get

$$\mathbb{E}(\widetilde{\xi}^n_{i,l}\mid\mathcal{F}'_S)=\Big(b_S+\int_E\delta(S,x)\lambda(dx)\Big)\Delta_n+\widetilde{\sigma}_S\int_S^T\Big(\int_S^s dW_u\Big)dW_s.$$

Hence the product of the right side above with $h(\beta^n_{i,l})$, where h is an odd function on $\mathbb{R}^d$ with polynomial growth, is a function of the form $Y(\omega,(W_{S+t}-W_S)_{t\geq 0})$ on $\Omega\times C(\mathbb{R}_+,\mathbb{R}^{d'})$ which is $\mathcal{F}_S\otimes\mathcal{C}$-measurable ($\mathcal{C}$ is the Borel σ-field on $C(\mathbb{R}_+,\mathbb{R}^{d'})$), and such that $Y(\omega,w)=Y(\omega,-w)$. Therefore we deduce

$$\mathbb{E}^n_{i+l-1}(\widetilde{\xi}^n_{i,l}\;h(\beta^n_{i,l})) \;=\; 0. \tag{7.25}$$

Step 4. Here we prove the following auxiliary result:

$$\sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\sqrt{\mathbb{E}(\alpha^n_{i,l})} \;\to\; 0. \tag{7.26}$$

Indeed, by Cauchy-Schwarz inequality the square of the left side of (7.26) is smaller than

$$\begin{aligned}t\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}(\alpha^n_{i,l}) \;=\; t\mathbb{E}\Bigg(\int_{l\Delta_n}^{\Delta_n(l+[t/\Delta_n]}&\Big(\|b_s-b_{\Delta_n[s/\Delta_n]}\|^2+\|\widetilde{\sigma}_s-\widetilde{\sigma}_{\Delta_n[s/\Delta_n]}\|^2\\ +\int\|\widetilde{\delta}(s,x)-\widetilde{\delta}(\Delta_n[s/\Delta_n],x)\|^2\lambda(dx)&+\int\|\delta(s,x)-\delta(\Delta_n[s/\Delta_n],x)\|\lambda(dx)\Big)ds\Bigg),\end{aligned}$$

which goes to 0 by the dominated convergence theorem and the bounds given in (SH) and $\int(\gamma(x)\wedge 1)\lambda(dx)<\infty$.

Step 5. By a Taylor expansion we can write

$$g^n_{i,l}(\beta^n_{i,l}+\xi^n_{i,l})-g^n_{i,l}(\beta^n_{i,l}) \;=\; \nabla g^n_{i,l}(\beta^n_{i,l})\xi^n_{i,l}+(\nabla g^n_{i,l}(\beta'^n_{i,l})-\nabla g^n_{i,l}(\beta^n_{i,l}))\xi^n_{i,l},$$

where $\beta'^n_{i,l}$ is some (random) vector lying on the segment between $\beta^n_{i,l}$ and $\beta^n_{i,l}+\xi^n_{i,l}$. Therefore we can write $g^n_{i,l}(\beta^n_{i,l}+\xi^n_{i,l})-g^n_{i,l}(\beta^n_{i,l})=\sum_{j=1}^3\zeta^n_{i,l}(j)$, where

$$\zeta^n_{i,l}(1)=\frac{1}{\sqrt{\Delta_n}}\,\nabla g^n_{i,l}(\beta^n_{i,l})\widetilde{\xi}^n_{i,l},\qquad \zeta^n_{i,l}(2)=\frac{1}{\sqrt{\Delta_n}}\,\nabla g^n_{i,l}(\beta^n_{i,l})\widehat{\xi}^n_{i,l},$$
$$\zeta^n_{i,l}(3)=(\nabla g^n_{i,l}(\beta'^n_{i,l})-\nabla g^n_{i,l}(\beta^n_{i,l}))\xi^n_{i,l}.$$

Then at this point it remains to prove that we have, for $j=1,2,3$:

$$\sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}^n_{i-1}(\zeta^n_{i,l}(j)) \;\stackrel{\text{u.c.p.}}{\longrightarrow}\; 0. \tag{7.27}$$

For $j=1$ this is obvious: indeed f is even in each of its (d-dimensional) arguments, so the functions $g^n_{i,l}$ are even as well, hence $\nabla g^n_{i,l}$ is odd and by (7.25) the left side of (7.27) is equal to 0.

Step 6) Now we prove (7.27) for $j=3$. By (7.18) and (7.19) we have for all $A\geq 1$ and $\varepsilon>0$:

$$\begin{aligned}|\zeta^n_{i,l}(3)| \;&\leq\; G_A(\varepsilon)\|\xi^n_{i,l}\|+KZ^n_{i,l}(1+\|\beta^n_{i,l}\|^r+\|\xi^n_{i,l}\|^r)\|\xi^n_{i,l}\| \\ &\qquad\qquad (1_{\{Z^n_{i,l}>A\}}+1_{\{\|\beta^n_{i,l}\|>A\}}+1_{\{\|\xi^n_{i,l}\|>\varepsilon\}}) \\ &\leq\; G_A(\varepsilon)\|\xi^n_{i,l}\|+KZ^n_{i,l}(1+Z^n_{i,l}) \\ &\qquad \Big(\frac{1+\|\beta^n_{i,l}\|)^{r+1}}{A}+\frac{(1+\|\beta^n_{i,l}\|)^r\|\xi^n_{i,l}\|}{\varepsilon}+\frac{(1+\|\beta^n_{i,l}\|)\|\xi^n_{i,l}\|^{r+1}}{A}+\frac{\|\xi^n_{i,l}\|^{r+2}}{\varepsilon}\Big).\end{aligned}$$

By (7.23) we have $\mathbb{E}^n_{i+l-1}(\|\xi^n_{i,l}|^q)\leq K_q\Delta_n$ if $q\geq 2$. Then in view of (6.23) we get by Hölder inequality:

$$\mathbb{E}^n_{i+l-1}(|\zeta^n_{i,l}(3)|)\;\leq\; K\sqrt{\Delta_n}\,\Big(G_A(\varepsilon)+Z^n_{i,l}(1+Z^n_{i,l})\Big(\frac{1}{A}+\frac{\Delta_n^{1/6}}{\varepsilon}\Big)\Big).$$

Then since $\mathbb{E}((Z^n_{i,l})^q)\leq K_q$ for all $q>0$ we have

$$\mathbb{E}\Big(\sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\Big|\mathbb{E}^n_{i+l-1}(\zeta^n_{i,l}(3))\Big|\Big)\;\leq\; Kt\Big(G_A(\varepsilon)+\frac{1}{A}+\frac{\Delta_n^{1/6}}{\varepsilon}\Big),$$

and (7.27) for $j=3$ follows (choose A big and then ε small).

Step 7) It remains to prove (7.27) for $j=2$. By (7.18) we have

$$|\zeta^n_{i,l}(2)|\;\leq\;\frac{K}{\sqrt{\Delta_n}}\,Z^n_{i,l}(1+\|\beta^n_{i,l}\|^r)\|\widehat{\xi}^n_{i,l}\|.$$

Hence by Cauchy-Schwarz inequality and (6.23),

$$\mathbb{E}\Big(\Big|\mathbb{E}^n_{i+l-1}(\zeta^n_{i,l}(2))\Big|\Big) \;\leq\; K\mathbb{E}\Big(Z^n_{i,l}(\Delta_n+\sqrt{\alpha^n_{i,l}})\Big) \;\leq\; K\Big(\Delta_n+\sqrt{\mathbb{E}(\alpha^n_{i,l})}\;\Big).$$

Then, in view of (7.26), the result is obvious.

7.6 Proof of Theorem 7.3.

The proof is exactly the same as above, with the following changes:

1) In Proposition 7.5 we substitute $\overline{V}^n$ and $V'(f,k)$ with

$$\overline{V}''^n_t \;=\; \sqrt{\Delta_n}\sum_{i=1}^{[t/k\Delta_n]}\Big(\zeta'^n_{(i-1)k+1}-\rho^{\otimes k}_{\sigma_{(i-1)k\Delta_n}}(f)\Big)$$

and $V''(f,k)$ respectively. The proof is then much shorter, because $\eta^n_i=\sqrt{\Delta_n}(\zeta'^n_{(i-1)k+1}-\rho^{\otimes k}_{\sigma_{(i-1)k\Delta_n}}(f))$ is $\mathcal{F}_{ik\Delta_n}$-measurable. We have

$$\mathbb{E}^n_{(i-1)k}(\eta^n_i))=0,\qquad \mathbb{E}^n_{(i-1)k}((\eta^n_i)^2)=\Delta_n\phi^n_i,\qquad \mathbb{E}^n_{(i-1)k}((\eta^n_i)^4)\leq K\Delta_n^2,$$

with $\phi^n_i=\rho^{\otimes k}_{\sigma_{(i-1)k\Delta_n}}(f^2)-\rho^{\otimes k}_{\sigma_{(i-1)k\Delta_n}}(f)^2$, and (7.17) is replaced by the obvious convergence of $\sum_{i=1}^{[t/k\Delta_n]}\mathbb{E}^n_{(i-1)k}((\eta^n_i)^2)$ to the right side of (7.7) (recall that we assumed $q=1$ here). We also have $\mathbb{E}^n_{(i-1)k}((N_{ik\Delta_n}-N_{(i-1)k\Delta_n})\zeta'^n_i)=0$ when N is a bounded martingale orthogonal to W by Lemma 6.10, and if N is one of the components of W because then this conditional expectation is the integral of a globally odd function, with respect to a measure on $(\mathbb{R}^{d'})^k$ which is symmetric about 0. So Lemma 4.4 readily applies directly, and the proposition is proved.

2) Next, we have to prove the analogues of (7.9), (7.10) and (7.11), where we only take the sum for those i of the form $i=(j-1)k+1$, and where in (7.11) we divide the integral by k. Proving the new version of (7.10) is of course simpler than the old one; the new version of (7.11) is the old one for Δ_n, whereas for (7.9) absolutely nothing is changed. So we are done.

7.7 Proof of Theorem 7.4.

For this theorem again we can essentially reproduce the previous proof, with $k=1$, and with the function f with components $f_{jm}(x)=x^jx^m$ (here m replaces the index k in the theorem). Again it suffices by polarization to prove the result for a single pair (j,m).

Below we set $\alpha_n=\alpha\Delta_n^{\varpi-1/2}$, which goes to ∞. Introduce the function on $\mathbb{R}^d$ defined by $g_n(x)=x^jx^m\psi_{\alpha_n}(x)$ (recall (6.24)), and set

$$\eta^n_i=\frac{\Delta^n_iX^j\Delta^n_iX^k}{\Delta_n}\;1_{\{\|\Delta^n_iX\|\leq\alpha\Delta_n^{\varpi}\}}-g_n(\Delta^n_iX/\sqrt{\Delta_n}),\quad \eta'^n_i=g_n(\Delta^n_iX/\sqrt{\Delta_n})-\beta^{n,m}_i\beta^{n,k}_i.$$

Proposition 7.5 implies that the processes

$$\sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\Big(\beta_i^{n,j}\beta_i^{n,m}-c_{(i-1)\Delta_n}^{jm}\Big)$$

converges stably in law to V^{jm}, and we also have (7.11), which here reads as

$$\frac{1}{\sqrt{\Delta_n}}\Big(\Delta_n\sum_{i=1}^{[t/\Delta_n]}c_{(i-1)\Delta_n}^{jj}-\int_0^t((c_u^{jm})^2+c_u^{jj}c_u^{mm})\,du\Big)\ \stackrel{\text{u.c.p.}}{\longrightarrow}\ 0.$$

Therefore it remains to prove the following three properties:

$$\sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\eta_i^n\ \stackrel{\text{u.c.p.}}{\longrightarrow}\ 0, \tag{7.28}$$

$$\sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n(\eta_i'^n)\ \stackrel{\text{u.c.p.}}{\longrightarrow}\ 0, \tag{7.29}$$

$$\sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\Big(\eta_i'^n-\mathbb{E}_{i-1}^n(\eta_i'^n)\Big)\ \stackrel{\text{u.c.p.}}{\longrightarrow}\ 0. \tag{7.30}$$

Proof of (7.28). Observe that $|\eta_i^n|\leq(\|\Delta_i^nX\|^2/\Delta_n)1_{\{\alpha\Delta_n^{\varpi}<\|\Delta_i^nX\|\leq 2\alpha\Delta_n^{\varpi}\}}$, hence

$$\begin{aligned}|\eta_i^n|&\leq 2\alpha_n1_{\{\|\beta_i^n\|>\alpha_n/2\}}+4\|\chi_i^n\|^21_{\{\alpha_n/2<\|\chi_i^n\|\leq 3\alpha_n\}}\\&\leq 2\alpha_n^{1-q}\|\beta_i^n\|^q++36(\|\chi_i^n\|^2\wedge\alpha_n^2)\end{aligned}$$

for any $q>0$ (recall that $\Delta_i^nX/\sqrt{\Delta_n}=\beta_i^n+\chi_i^n$). Then we take q such that $(q-1)(1/2-\varpi)>2\varpi-\varpi r$, and we apply (6.23) and (6.25) with α_n as above (so $\alpha_n\geq 1$ for n large enough, and $\alpha_n\sqrt{\Delta_n}\to 0$), to get $\mathbb{E}(|\eta_i^n|)\leq K\Delta_n^{2\varpi-\varpi r}u_n$, where $u_n\to 0$. Hence

$$\sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}(|\eta_i^n|)\ \leq\ Kt\Delta_n^{2\varpi-\varpi r-1/2}u_n,$$

which goes to 0 if $\varpi\geq\frac{1}{2(2-r)}$. Hence we have (7.28).

Proof of (7.29). From the properties of ψ, the function ψ_{α_n} is differentiable and $\|\nabla\psi_{\alpha_n}(x)\|\leq(K/\alpha_n)1_{\{\|x\|\leq 2\alpha_n\}}$. Hence we clearly have $\|\nabla g_n(x)\|\leq K(\|x\|\wedge\alpha_n)$, and thus

$$\left.\begin{aligned}|g_n(x+y)-g_n(x))|&\leq K\alpha_n(\|y\|\wedge\alpha_n),\\|g_n(x+y)-g_n(x)-\nabla g_n(x)y|&\leq K\|y\|^2.\end{aligned}\right\} \tag{7.31}$$

If we use the first estimate above and (6.26) we obtain, as in Step 2 of the previous proof (we use again $\varpi\geq\frac{1}{2(2-r)}$ here), that

$$\sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n\Big(g_n(\Delta_{i+l}^nX/\sqrt{\Delta_n})-g_n(\Delta_{i+l}^nX/\sqrt{\Delta_n}-\Delta_{i+l}^n(\delta\star\underline{\mu})/\sqrt{\Delta_n})\Big)\ \stackrel{\text{u.c.p.}}{\longrightarrow}\ 0.$$

Then with the notation of (7.21), in order to prove (7.29) it is enough to prove (7.12) with g_n instead of $g^n_{i,l}$, and $l = 0$. Then the second estimate in (7.31) allows to write $g_n(\beta^n_{i,l} + \xi^n_{i,l}) - g_n(\beta^n_{i,l}) = \sum_{j=1}^3 \zeta^n_i(j)$, where

$$\zeta^n_i(1) = \frac{1}{\sqrt{\Delta_n}}\, \nabla g_n(\beta^n_i)\widetilde{\xi}^n_{i,0}, \qquad \zeta^n_i(2) = \frac{1}{\sqrt{\Delta_n}}\, \nabla g^n_{i,l}(\beta^n_i)\widehat{\xi}^n_{i,0}, \qquad |\zeta^n_i(3)| \leq K\|\xi^n_{i,0}\|^2.$$

Then it remains to prove (7.27) for $j = 1,2,3$, with $\zeta^n_i(j)$ instead of $\zeta^n_{i,l}(j)$.

Since g_n is even, this property for $j = 1$ follows from (7.25) exactly as in the previous proof. The proof for $j = 2$ is the same as in Step 7 of the previous proof (here $Z^n_{l,i} = 1$ and $r = 2$). Finally by (7.23) we have $\mathbb{E}(\|\xi^n_{i,0}\|^2) \leq K\Delta_n$, so the result for $j = 3$ is immediate.

Proof of (7.30). Exactly as for (7.10) it is enough to prove that

$$\Delta_n \sum_{i=1}^{[t/\Delta_n]} \mathbb{E}^n_{i-1}((\eta'^n_i)^2) \ \stackrel{\text{u.c.p.}}{\longrightarrow}\ 0. \tag{7.32}$$

First, we have

$$\begin{aligned}\sum_{i=1}^{[t/\Delta_n]} \mathbb{E}^n_{i-1}\left(\left(\beta^{n,j}_i\beta^{n,m}_i - g_n(\beta^n_i)\right)^2\right) \ &\leq\ \sum_{i=1}^{[t/\Delta_n]} \mathbb{E}^n_{i-1}\left(\|\beta^n_i\|^4 1_{\{\|\beta^n_i\|>\alpha_n\}}\right) \\ &\leq\ K\Delta_n^{q((1/2-\varpi)} \sum_{i=1}^{[t/\Delta_n]} \mathbb{E}^n_{i-1}(\|\beta^n_i\|^{4+q}) \ \leq\ Kt\Delta_n,\end{aligned}$$

by choosing appropriately q for the last inequality.

Second, since $\Delta^n_i X = \sqrt{\Delta_n}\,(\beta^n_i + \xi^n_{i,0}) + \Delta^n_i(\delta \star \underline{\mu})$, we deduce from (7.31) that

$$\left(g_n(\Delta^n_i X/\sqrt{\Delta_n}) - g_n(\beta^n_i)\right)^2 \leq K\alpha_n^2\|\xi'^n_{i,0}\|^2 + K\alpha_n^3\Big((|\Delta^n_i(\delta\star\underline{\mu})|/\sqrt{\Delta_n}) \bigwedge \alpha_n\Big).$$

Then by (6.26) and (7.23) again, we get

$$\mathbb{E}\left(\left(g_n(\Delta^n_i X/\sqrt{\Delta_n}) - g_n(\beta^n_i)\right)^2\right) \leq K(\alpha_n^2\Delta_n + \alpha_n^{4-r}\Delta_n^{1-r/2}) \leq K\Delta_n^{\varpi(4-r)-1}.$$

If we put together these estimates, we find that

$$\Delta_n \sum_{i=1}^{[t/\Delta_n]} \mathbb{E}((\eta'^n_i)^2) \ \leq\ Kt(\Delta_n + \Delta_n^{\varpi(4-r)-1}),$$

which goes to 0 because $\varpi \geq \frac{1}{2(2-r)}$. Hence we have (7.32).

8 CLT with discontinuous limits

So far we have been concerned with CLTs associated with Theorems 6.2, 6.3 and 6.4, in which the limiting processes are always continuous. Now, as seen in the case $r = 3$ of (3.23) there are cases where the limit is a sum of jumps, and we are looking at this kind of question here. In case $r = 2$ of (3.23) we even have a "mixed" limit with a continuous and a purely discontinuous parts: this has less statistical interest, and we will state the result without proof.

Here, more than in the continuous case even, it is important and not completely trivial to define the limiting processes. This is the aim of the first subsection below. Throughout, we assume (H), and we also fix an integer $k \geq 2$.

8.1 The limiting processes.

As for the case of continuous limits, we will have stable convergence in law, and the limiting processes will be defined on an extension of the space $(\Omega, \mathcal{F}, \mathbb{P})$. To do this, it is convenient to introduce another probability space $(\Omega', \mathcal{F}', \mathbb{P}')$. We assume that this space supports the following variables:

- four sequences (U_p), (U'_p), $(\overline{U}_p)$, $(\overline{U}'_p)$ of d'-dimensional $N(0, I_{d'})$ variables;
- a sequence (κ_p) of uniform variables on $[0, 1]$;
- a sequence (L_p) of uniform variables on the finite set $\{0, 1, \cdots, k-1\}$, where $k \geq 2$ is some fixed integer;

and all these variables are mutually independent. Then we put

$$\widetilde{\Omega} = \Omega \times \Omega', \qquad \widetilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}', \qquad \widetilde{\mathbb{P}} = \mathbb{P} \otimes \mathbb{P}'. \tag{8.1}$$

We extend the variables X_t, b_t, ... defined on Ω and U_p, κ_p,... defined on Ω' to the product $\widetilde{\Omega}$ in the obvious way, without changing the notation. We write $\widetilde{\mathbb{E}}$ for the expectation with respect to $\widetilde{\mathbb{P}}$.

Next, we need a filtration $(\widetilde{\mathcal{F}}_t)_{t \geq 0}$ on our extension. To this effect, we first denote by $(S_p)_{p \geq 1}$ a sequence of stopping times which exhausts the "jumps" of the Poisson measure $\underline{\mu}$: this means that for each ω we have $S_p(\omega) \neq S_q(\omega)$ if $p \neq q$, and that $\underline{\mu}(\omega, \{t\} \times E) = 1$ if and only if $t = S_p(\omega)$ for some p. There are many ways of constructing those stopping times, but it turns out that what follows does not depend on the specific description of them.

With a given choice of the above stopping times S_p, we let $(\widetilde{\mathcal{F}}_t)$ be the smallest (right-continuous) filtration of $\widetilde{\mathcal{F}}$ containing the filtration $(\mathcal{F}_t)$ and such that U_p, U'_p, $\overline{U}_p$, $\overline{U}'_p$, κ_p and L_p are $\widetilde{\mathcal{F}}_{S_p}$-measurable for all p. Obviously, $\underline{\mu}$ is still a Poisson measure with compensator $\underline{\nu}$, and W and W' is still a Wiener process on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t \geq 0}, \widetilde{\mathbb{P}})$. Finally we define the q-dimensional variables

$$\left.\begin{array}{l} R_p = \sqrt{\kappa_p}\ \sigma_{S_p-} U_p + \sqrt{1-\kappa_p}\ \sigma_{S_p} \overline{U}_p \\ R''_p = \sqrt{L_p}\ \sigma_{S_p-} U'_p + \sqrt{k-1-L_p}\ \sigma_{S_p} \overline{U}'_p \\ R'_p = R_p + R''_p. \end{array}\right\} \tag{8.2}$$

If f is a C^q function on $\mathbb{R}^d$ we denote by $\nabla^r f$ for $r \leq q$ the tensor of its rth derivatives, and if we want to be more specific, we write $\partial^r_{i_1,\cdots,i_r} f$ the rh partial derivative with respect to the components $x^{i_1}, \cdots, x^{j_r}$, and simply ∇f and $\partial_i f$ when $r = 1$. If f and g are two C^1 functions we set

$$C(f,g)_t \ := \ \sum_{s \leq t} \ \sum_{i,j=1}^d \ (\partial_i f \partial_j g)(\Delta X_s)\ (c^{ij}_{s-} + c^{ij}_s). \tag{8.3}$$

This makes sense (that is, the series above converges for all t) as soon as $f(0) = 0$, because then $\|\nabla f(x)\| \leq K\|x\|$ for $\|x\| \leq 1$ and the process c is locally bounded and $\sum_{s \leq t} \|\Delta X_s\|^2 < \infty$; the process $C(f,g)$ is then of finite variation, and even increasing when $g = f$. In the same way, if f is C^2 and $\|\nabla f(x)\| \leq K\|x\|^2$ when $\|x\| \leq 1$, the following defines a process of finite variation:

$$\overline{C}(f)_t \ := \ \sum_{s \leq t} \ \sum_{i,j=1}^d \ \partial^2_{ij} f(\Delta X_s)\ (c^{ij}_{s-} + c^{ij}_s). \tag{8.4}$$

In the following lemma we define and prove the existence of our limiting processes, at the same time. We do it for a q-dimensional function $f = (f_1, \cdots, f_q)$, since it costs us nothing.

Lemma 8.1 *a) Let f be a q-dimensional C^1 function on $\mathbb{R}^d$, vanishing at 0. The formulas*

$$Z(f_l)_t = \sum_{p:\ S_p \leq t} \ \sum_{i=1}^d \ \partial_i f_l(\Delta X_{S_p})\, R^i_p, \quad Z'(f_l)_t = \sum_{p:\ S_p \leq t} \ \sum_{i=1}^d \ \partial_i f_l(\Delta X_{S_p})\, R'^i_p \tag{8.5}$$

define two q-dimensional processes $Z(f) = (Z(f_l))_{l \leq q}$ and $Z'(f) = (Z'(f_l))_{l \leq q}$, and conditionally on $\mathcal{F}$ the pair $(Z(f), Z'(f))$ is a square-integrable martingale with independent increments, zero mean and variance-covariance given by

$$\left.\begin{array}{l} \widetilde{\mathbb{E}}(Z(f_l)_t Z(f_{l'})_t \mid \mathcal{F}) \ = \ \widetilde{\mathbb{E}}(Z(f_l)_t Z'(f_{l'})_t \mid \mathcal{F}) \ = \ \frac{1}{2}\ C(f_l, f_{l'})_t, \\ \widetilde{\mathbb{E}}(Z'(f_l)_t Z'(f_{l'})_t \mid \mathcal{F}) \ = \ \frac{k}{2}\ C(f_l, f_{l'})_t. \end{array}\right\} \tag{8.6}$$

Moreover, if X and c have no common jumps, conditionally on $\mathcal{F}$ the process $(Z(f), Z'(f))$ is a Gaussian martingale.

b) Let f be a q-dimensional C^2 function on $\mathbb{R}^d$, with $\|\nabla^2 f(x)\| \leq \|x|^2$ for $\|x\| \leq 1$. The formulas

$$\left.\begin{array}{l} \overline{Z}(f_l)_t = \sum_{p:\ S_p \leq t}\ \sum_{i,j=1}^d\ \partial^2_{ij} f_l(\Delta X_{S_p})\, R^i_p, R^j_p \\ \overline{Z}'(f_l)_t = \sum_{p:\ S_p \leq t}\ \sum_{i,j=1}^d\ \partial^2_{ij} f_l(\Delta X_{S_p})\, R'^i_p R'^j_p \end{array}\right\} \tag{8.7}$$

define two q-dimensional processes $\overline{Z}(f) = (\overline{Z}(f_l))_{l\leq q}$ *and* $\overline{Z}'(f) = (\overline{Z}'(f_l))_{l\leq q}$ *of finite variation, and with* $\mathcal{F}$*-conditional expectations given by*

$$\left.\begin{array}{rcl} \widetilde{\mathbb{E}}(\overline{Z}(f_l)_t \mid \mathcal{F}) &=& \frac{1}{2}\,\overline{C}(f_l)_t, \\ \widetilde{\mathbb{E}}(\overline{Z}'(f_l)_t \mid \mathcal{F}) &=& \frac{k}{2}\,\overline{C}(f_l)_t. \end{array}\right\} \tag{8.8}$$

c) The processes $(Z(f), Z'(f))$ *and* $(\overline{Z}(f), \overline{Z}'(f))$ *above depend on the choice of he sequence* (S_p) *of stopping times exhausting the jumps of* $\underline{\mu}$*, but their* $\mathcal{F}$*-conditional laws do not.*

Proof. a) Among several natural proofs, here is an "elementary" one. We set $\alpha_p(l,l') = \sum_{i,j=1}^d (\partial_i f_l \partial_j f_{l'})(\Delta X_{S_p})(c^{ij}_{S_p-} + c^{ij}_{S_p})$, so $C(f_l, f_{l'})_t = \sum_{p:\ S_p\leq t} \alpha_p(l,l')$. We fix $\omega\in\Omega$, and we consider the q-dimensional variables $\Phi_p(\omega,.)$ and $\Phi'_p(\omega,.)$ on $(\Omega',\mathcal{F}')$ with components

$$\Phi^l_p(\omega,\omega') = \sum_{i=1}^d \partial_i f_l(\Delta X_{S_p}(\omega)) R^i_p(\omega,\omega'), \quad \Phi'^l_p(\omega,\omega') = \sum_{i=1}^d \partial_i f_l(\Delta X_{S_p}(\omega)) R'^i_p(\omega,\omega').$$

The variables $(\Phi_p(\omega,.), \Phi'_p(\omega,.))$ on $(\Omega',\mathcal{F}',\mathbb{P}')$ are independent as p varies, and a simple calculation shows that they have zero mean and variance-covariance given by

$$\left.\begin{array}{l} \mathbb{E}'(\Phi^l_p(\omega,.)\Phi^{l'}_p(\omega,.)) \;=\; \mathbb{E}'(\Phi^l_p(\omega,.)\Phi'^{l'}_p(\omega,.)) \;=\; \frac{1}{2}\,\alpha_p(l,l';\omega) \\ \mathbb{E}'(\Phi'^l_p(\omega,.)\Phi'^{l'}_p(\omega,.)) \;=\; \frac{k}{2}\,\alpha_p(l,l';\omega) \end{array}\right\} \tag{8.9}$$

Since $\sum_{p:\ S_p(\omega)\leq t} \alpha_p(l,l';\omega) < \infty$, a standard criterion for convergence of series of independent variables yields that the formulas

$$Z(f_l)_t(\omega,.) = \sum_{p:\ S_p(\omega)\leq t} \Phi^l_p(\omega,.), \qquad Z'(f_l)_t(\omega,.) = \sum_{p:\ S_p(\omega)\leq t} \Phi'^l_p(\omega)) R'_p(\omega,.)$$

define a $2q$-dimensional process $(\omega',t)\mapsto (Z(f)(\omega,\omega')_t, Z'(f)(\omega,\omega')_t)$, which obviously is a martingale with independent increments, and with $((2q)\times 2$-dimensional) predictable bracket being deterministic (that is, it does not depend on ω') and equal at time t to the sum over all p with $S_p(\omega)\leq t$ of the right sides of (8.9). That is, we can consider $(Z(f), Z'(f))$ as a process on the extended space, and it satisfies (8.6). Since the law of a centered martingale with independent increments depends only on its predictable bracket we see that the law of $(Z(f), Z'(f))$, conditional on $\mathcal{F}$, only depends on the processes $C(f_l, f_{l'})$ and thus does not depend on the particular choice of the sequence (S_p).

Moreover this martingale is purely discontinuous and jumps at times $S_p(\omega)$, and if X and c have no common jumps, the jump of $(Z(f)(\omega,.), Z'(f)(\omega,.))$ at $S_p(\omega)$ equals

$$\Big(\nabla f(\Delta X_{S_p})\sigma_{S_p}(\omega)\Big(\sqrt{\kappa_p}\,U_p + \sqrt{1-\kappa_p}\,\overline{U}_p\Big), \\ \nabla f(\Delta X_{S_p})\sigma_{S_p}(\omega)\Big(\sqrt{\kappa_p}\,U_p + \sqrt{1-\kappa_p}\,\overline{U}_p + \sqrt{L_p}\,U'_p + \sqrt{k-1-L_p}\,\overline{U}'_p\Big)\Big)$$

(we use here product matrix notation); this 2-dimensional variable is $\mathcal{F}$-conditionally Gaussian and centered, so in this case the pair $(Z(f), Z'(f))$ is $\mathcal{F}$-conditionally a Gaussian process.

b) Since $\widetilde{\mathbb{E}}(|R_p^i R_p^j| \mid \mathcal{F}) \leq K(\|c_{S_p-}\| + \|c_{S_p}\|)$, and the same with R'_p, it is obvious in view of our assumption on f that the $\mathcal{F}$-conditional expectation of the two variables

$$\sum_{p:\ S_p \leq t} \Big| \sum_{i,j=1}^d \partial^2_{ij} f_l(\Delta X_{S_p})\, R_p^i, R_p^j \Big|, \qquad \sum_{p:\ S_p \leq t} \Big| \sum_{i,j=1}^d \partial^2_{ij} f_l(\Delta X_{S_p})\, R'^i_p R'^j_p \Big|$$

is finite for all t. Then all claims are obvious.

It remains to prove (c) for the process $(\overline{Z}(f), \overline{Z}'(f))$. For this, we observe that conditionally on $\mathcal{F}$ this process is the sum of its jump and it has independent increments. Moreover it jumps only when X jumps, and if T is a finite $(\mathcal{F}_t)$-stopping time such that $\Delta X_T \neq 0$, then its jump at time T is

$$\Big(\sum_{i,j=1}^d \partial^2_{ij} f(\Delta X_T) \widetilde{R}^{ij}, \sum_{i,j=1}^d \partial^2_{ij} f(\Delta X_T) \widetilde{R}'^{ij} \Big),$$

where $WR^{ij} = \sum_{p \geq 1} R_p^i R_p^j 1_{\{S_p = T\}}$ and a similar expression for $\widetilde{R}'^{ij}$. But the $\mathcal{F}$-conditional law of $(\widetilde{R}^{ij}, WR'^{ij})$ clearly depends only on σ_{T-} and σ_T, but not on the particular choice of the sequence (S_p). This proves the result. □

8.2 The results.

Now we proceed to giving a CLT associated with the convergence in (5.2), and as seen already in (3.23) we need some smoothness for the test function f, and also that $f(x)$ goes to 0 faster than $\|x\|^3$ instead of $\|x\|^2$ as $x \to 0$. As in Theorem 7.1 we also consider a q-dimensional function $f = (f_1, \cdots, f_q)$.

Theorem 8.2 *Assume (H) (or (H') only), and let f be a q-dimensional C^2 function on $\mathbb{R}^d$ satisfying $f(0) = 0$ and $\nabla f(0) = 0$ and $\nabla^2 f(x) = \mathrm{o}(\|x\|)$ as $x \to 0$. The pair of q-dimensional processes*

$$\Big(\frac{1}{\sqrt{\Delta_n}}\ (V(f, \Delta_n)_t - f \star \mu_{\Delta_n[t/\Delta_n]}), \frac{1}{\sqrt{\Delta_n}}\ (V(f, k\Delta_n)_t - f \star \mu_{k\Delta_n[t/k\Delta_n]}) \Big) \tag{8.10}$$

converges stably in law, on the product $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^q) \times \mathbb{D}(\mathbb{R}_+, \mathbb{R}^q)$ of the Skorokhod spaces, to the process $(Z(f), Z'(f))$.

We have the (stable) convergence in law of the above processes, as elements of the product functional space $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^q)^2$, but usually not as elements of the space $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^{2q})$ with the ($2q$-dimensional) Skorokhod topology, because a jump of X at time S, say, entails a jump for both components above at two times S_n and S'_n which both converge to S but are in general different (with a probability close to $(k-1)/k$, in fact): this prevents the $2q$-dimensional Skorokhod convergence. In the same way, although $S_n \to S$, we have $S_n \neq S$ and $V(f, \Delta_n)$ jumps at S_n whereas $f \star \mu$ jumps at S: this is why, if we want Skorokhod convergence, we have to center $V(f, \Delta_n)$ around the discretized version of $f \star \mu$.

However, in most applications we are interested in the convergence at a given fixed time t. Since $\mathbb{P}(\Delta X_t \neq 0) = 0$ for all t, in view of the properties of the Skorokhod convergence we immediately get the following corollary:

Corollary 8.3 *Under the assumptions of the previous theorem, for any fixed $t > 0$ the $2q$-dimensional variables*

$$\Big(\frac{1}{\sqrt{\Delta_n}}\,(V(f,\Delta_n)_t - f\star\mu_t),\frac{1}{\sqrt{\Delta_n}}\,(V(f,k\Delta_n)_t - f\star\mu_t)\Big)$$

converges stably in law to the variable $(Z(f)_t, Z'(f)_t)$.

Now, it may happen that f is such that $f\star\mu = 0$, and also $(\nabla f)\star\mu = 0$: this is the case when X is continuous, of course, but it may also happen when X is discontinuous, as we will see in some statistical applications later. Then the above result degenerates, and does not give much insight. So we need a further CLT, which goes as follows. There is a general result in the same spirit as Theorem 8.2, but here we consider a very special situation, which is enough for the applications we have in mind:

Theorem 8.4 *Assume (H) (or (H') only), and suppose that the two components X^1 and X^2, say, never jump at the same times. Let f be the function $f(x) = (x^1x^2)^2$. Then the 2-dimensional processes*

$$\Big(\frac{1}{\Delta_n}\;V(f,\Delta_n),\frac{1}{\Delta_n}\;V(f,k\Delta_n)\Big) \tag{8.11}$$

converge stably in law, on the product $\mathbb{D}(\mathbb{R}_+,\mathbb{R})\times\mathbb{D}(\mathbb{R}_+,\mathbb{R})$ of the Skorokhod spaces, to the process

$$\Big(\frac{1}{2}\;\overline{Z}(f)_t + \int_0^t(c_u^{ii}c_u^{jj} + 2(c_u^{ij})^2)du,\;\frac{1}{2}\;\overline{Z}'(f)_t + k\int_0^t(c_u^{ii}c_u^{jj} + 2(c_u^{ij})^2)du\Big) \tag{8.12}$$

Of course the same result holds for any two other components. More generally a similar result holds when f is an homogeneous polynomial of degree 4, which satisfies outside a $\mathbb{P}$-null set:

$$f\star\mu \;=\; 0, \qquad (\nabla f)\star\mu \;=\; 0\,. \tag{8.13}$$

Finally as said before, we also state, without proof, the result about the quadratic variation itself. Although not so important for statistical applications, it is of great theoretical significance. Exactly as in Corollary 7.2, only the weak assumption (H') is required here (see [15] for a proof, and [14] for an early version stated somewhat differently).

Theorem 8.5 *Assume (H'), Then the $d\times d$-dimensional process with components*

$$\frac{1}{\sqrt{\Delta_n}}\;\Big(\sum_{i=1}^{[t/\Delta_n]}\Delta_i^nX^j\Delta_i^nX^k - [X^j,X^k]_{[t/\Delta_n]\Delta_n]}\Big)$$

converge stably in law to $V + Z(f)$, where V is a s described in Corollary 7.2 and $Z(f)$ is as above with $f_{jk}(x) = x_jx_k$, and conditionally on $\mathcal{F}$ the processes V and $Z(f)$ are independent.

8.3 Some preliminary on stable convergence.

Once more, for the above results it is enough to prove them under (SH), which we assume henceforth. The basis of the proof is a rather general result of stable convergence about discontinuous processes, which cannot be found in a book form so far.

Although what follows does not depend on the choice of the sequence (S_p), for convenience we make a specific choice. For any $m \geq 1$ we denote by $(T(m,r) : r \geq 1)$ the successive jump times of the process $N^m = 1_{\{1/m < \gamma \leq 1/(m-1)\}} \star \underline{\mu}$ (note that N^m is an homogeneous Poisson process with intensity $\lambda(\{z : \frac{1}{m} < \gamma(z) < \frac{1}{m-1}\})$). Then (S_p) is a reordering of the double sequence $(T(m,r) : r, m \geq 1)$ into a single sequence.

Next we introduce some notation. For any $p \geq 1$ the time S_p is in one and only one interval $((ik+j)\Delta_n, (ik+j+1)\Delta_n]$, for some $i \geq 0$ and $j = 0, \cdots, k-1$. So, we can define a number of quantities by setting their values on each set $\{(ik+j)\Delta_n < S_p \leq (ik+j+1)\Delta_n\}$:

$$\left.\begin{array}{l}
L(n,p) = j, \qquad\qquad K(n,p) = \frac{S_p}{\Delta_n} - (ik+j) \\
\alpha_-(n,p) = \frac{1}{\sqrt{\Delta_n}}\,(W_{S_p} - W_{(ik+j)\Delta_n}), \quad \alpha_+(n,p) = \frac{1}{\sqrt{\Delta_n}}\,(W_{(ik+j+1)\Delta_n} - W_{S_p}) \\
\beta_-(n,p) = \frac{1}{\sqrt{\Delta_n}}\,(W_{(ik+j)\Delta_n} - W_{ik\Delta_n}) \\
\beta_+(n,p) = \frac{1}{\sqrt{\Delta_n}}\,(W_{(i+1)k\Delta_n} - W_{(ik+j+1)\Delta_n}) \\
A(n,p) = (\alpha_-(n,p), \alpha_+(n,p), \beta_-(n,p), \beta_+(n,p)) \\
\widehat{R}^n_p = \sigma_{(ik+j)\Delta_n}\alpha_-(n,p) + \sigma_{T_p}\alpha_+(n,p), \quad \widehat{R}'^n_p = \sigma_{ik\Delta_n}\beta_-(n,p) + \sigma_{T_p}\beta_+(n,p) \\
R^n_p = \frac{1}{\sqrt{\Delta_n}}\,(X_{(ik+j+1)\Delta_n} - X_{(ik+j)\Delta_n} - \Delta X_{S_p}) \\
R'^n_p = \frac{1}{\sqrt{\Delta_n}}\,(X_{(i+1)k\Delta_n} - X_{(ik+j+1)\Delta_n} + X_{(ik+j)\Delta_n} - X_{ik\Delta_n}).
\end{array}\right\} \quad (8.14)$$

In the next lemma, we consider the variables $\Theta_n = A(n,p)_{p\geq 1}$ taking values in the Polish space $F = (\mathbb{R}^4)^{\mathbb{N}^\star}$, and also the variable $\Theta = (A_p)_{p\geq 1}$ taking values in F as well, where $A_p = (\sqrt{\kappa_p}\,U_p, \sqrt{1-\kappa_p}\,\overline{U}_p, \sqrt{L_p}\,U'_p, \sqrt{k-1-L_p}\,\overline{U}'_p)$ uses the variables introduced at the beginning of this section.

Lemma 8.6 *The sequence* (Θ_n) *of variables stably converges in law to* Θ*.*

Proof. We need to prove that

$$\mathbb{E}(Zh(\Theta_n)) \to \widetilde{\mathbb{E}}(Zh(\Theta)) \qquad (8.15)$$

for any bounded $\mathcal{F}$-measurable variable Z and any bounded continuous function h on F.

Let $\mathcal{G}$ be the σ-field of Ω generated by the process W and the random measure $\underline{\mu}$. Each Θ_n is $\mathcal{G}$-measurable and Θ is $\mathcal{G} \otimes \mathcal{F}'$-measurable, so $\mathbb{E}(Zh(\Theta_n)) = \mathbb{E}(Z'h(\Theta_n))$ and $\widetilde{\mathbb{E}}(Zh(\Theta)) = \widetilde{\mathbb{E}}(Z'h(\Theta))$, where $Z' = \mathbb{E}(Z \mid \mathcal{G})$. Hence it suffices to prove (8.15) when Z is $\mathcal{G}$-measurable.

We can go further: recalling that μ has the form $\underline{\mu} = \sum_{p\geq 1} \varepsilon_{(S_p,V_p)}$ for suitable E-valued variables V_p (ε_a = Dirac mass at a), then $\mathcal{G}$ generated by W and the variables

(S_p, V_p). Then by a density argument it is enough to prove (8.15) when

$$Z = f(W) \prod_{p=1}^{P} g_p(S_p) g'_p(V_p), \qquad h((z_p)_{p\geq 1}) = \prod_{p=1}^{P} h_p(z_p)$$

where f is continuous and bounded on the space of all continuous $\mathbb{R}^{d'}$-valued functions, and the g_p's are continuous and bounded on R_+ and the g'_p's are continuous and bounded on E, and the h_p's are continuous and bounded on $\mathbb{R}^4$, and P is an integer.

Let $W^n_t = W_t - \sum_{p=1}^{P} (W_{S_p+2k\Delta_n} - W_{(S_p-2k\Delta_n)^+})$. Clearly $W^n \to W$ uniformly (for each ω), hence $f(W^n) \to f(W)$. If $\Omega(n,P) = \cap_{p,p'\in\{1,\cdots,P\},\ p\neq p'} \{|S_p - S_{p'}| > k\Delta_n\}$, we also have $\Omega(n,P) \to \Omega$ as $n \to \infty$. Therefore by Lebesgue theorem,

$$\mathbb{E}\Big(f(W) \prod_{p=1}^{P} g_p(S_q) g'_p(V_p) h_p(A(n,p))\ 1_{\Omega(n,P)}\Big) - \mathbb{E}\Big(f(W^n) \prod_{p=1}^{P} g_p(S_p) g'_p(V_p) h_p(A(n,p))\Big)$$

goes to 0, and we are left to prove that

$$\mathbb{E}\Big(f(W^n) \prod_{p=1}^{P} g_p(S_p) g'_p(V_p) h_p(A(n,p))\ 1_{\Omega(n,P)}\Big) \ \to\ \widetilde{\mathbb{E}}\Big(f(W) \prod_{p=1}^{P} g_p(S_p) g'_q(V_p) h_p(A_p)\Big).$$

Now, W and $\underline{\mu}$ are independent, and with our choice of the sequence (S_p) the two sequences (S_p) and (V_p) are also independent. This implies that W^n, the family (V_p) and the family $(A(n,p))_{p\leq P}$ are independent as well. Therefore the left side above equals the product of $\mathbb{E}\Big(f(W^n) \prod_{p=1}^{P} g'_p(V_p)\Big)$ with $\mathbb{E}\Big(\prod_{p=1}^{P} g_p(T_p) h_p(A(n,p))\ 1_{\Omega(n,P)}\Big)$, and likewise for the right side. So finally it remains to prove that

$$\mathbb{E}\Big(\prod_{p=1}^{P} g_p(S_p) h_p(A(n,p))\ 1_{\Omega(n,P)}\Big) \ \to\ \widetilde{\mathbb{E}}\Big(\prod_{p=1}^{P} g_p(S_p) h_p(A_p)\Big). \tag{8.16}$$

At this stage, and by another application of the independence between W and $\underline{\mu}$, we observe that in restriction on the set $\Omega(n,P)$, the sequence $(A(n,p) : p = 1,\cdots,P)$ has the same law than the sequence $(A'(n,p) : p = 1,\cdots,P)$, where

$$A'(n,p) = (\sqrt{K(n,p)}\ U_p, \sqrt{1-K(n,p)}\ \overline{U}_q, \sqrt{L(n,p)}\ U'_p, \sqrt{k-1-L(n,p)}\ \overline{U}'_p).$$

Therefore (8.16) amounts to proving that the sequence $((S_p, K(n,p), L(n,p)) : p = 1\cdots,P)$ converges in law to $((S_p, \kappa_p, L_p) : p = 1\cdots,P)$.

To see this, on may introduce the fractional part $G(n,p)$ of $[S_p/k\Delta_n]$, which equals $\frac{1}{\Delta_n}(S_p - ik\Delta_n)$ on the set $\{ik\Delta_n \leq S_p < (i+1)k\Delta_n\}$. Since the family $(S_p : p = 1,\cdots,P)$ admits a smooth density on its support in $\mathbb{R}^P_+$ (again because of our choice of (S_p)), an old result of Tukey in [23] shows that the sequences $((S_p, G(n,p)) : p = 1,\cdots,P)$ converge in law, as $n \to \infty$, to $((S_p, G_p) : p = 1,\cdots,P)$ where the G_p's are independent one from the other and from the S_p's and uniformly distributed on $[0,1]$ (Tukey's result deals with 1-dimensional variables, but the multidimensional extension is straightforward). Since $K(n,p)$ and $L(n,p)$ are respectively the fractional part and the integer part of $G(n,p)/k$, and since the fractional part and the integer part of G_p/k are independent and respectively uniform on $[0,1]$ and uniform on $\{0,\cdots,k-1\}$, the desired result is now obvious. □

Lemma 8.7 *The sequence of $(\mathbb{R}^{2d})^{\mathbb{N}^\star}$-valued variables $((R_p^n, R_p''^n) : p \geq 1)$ stably converges in law to $((R_p, R_p'') : p \geq 1)$ (see (8.2)).*

Proof. This result is a consequence of one of the basic properties of the stable convergence in law. Namely, if a sequence Y_n of E-variables defined on the space $(\Omega, \mathcal{F}, \mathbb{P})$ stably converges in law to Y (defined on an extension), and if a sequence Z_n of F-variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$ again converges in probability to Z, then for any continuous function f on $E \times F$ the variables $f(Y_n, Z_n)$ stably converge in law to $f(Y, Z)$.

A first application of this property allows to deduce from the previous lemma and from the fact that σ_t is right continuous with left limits is that $((\widehat{R}_p^n, \widehat{R}_p''^n) : p \geq 1)$ stably converges in law to $((R_p, R_p'') : p \geq 1)$. A second application of the same shows that, in order to get our result, it is enough to prove that for each $p \geq 1$ we have

$$R_p^n - \widehat{R}_p^n \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad R_p''^n - \widehat{R}_p''^n \stackrel{\mathbb{P}}{\longrightarrow} 0. \tag{8.17}$$

We will prove the first part of (8.17), the proof of the second part being similar. Recall that $S_p = T(m, r)$ for some $r, m \geq 1$, and set $X' = X - X^c - (\delta 1_{\{\gamma > 1/m\}}) \star \underline{\mu}$ and

$$\begin{cases} \zeta_i^n(t) = \frac{1}{\sqrt{\Delta_n}} \Big(\int_{(i-1)\Delta_n}^t (\sigma_u - \sigma_{(i-1)\Delta_n}) dW_u + \int_t^{i\Delta_n} (\sigma_u - \sigma_t) dW_u \Big) 1_{((i-1)\Delta_n, i\Delta_n]}(t) \\ \zeta_i'^n = \frac{1}{\sqrt{\Delta_n}} \, \Delta_i^n X', \end{cases}$$

and observe that

$$R_p^n - \widehat{R}_p^n \; = \; \sum_{i \geq 1} \Big(\zeta_i^n(S_p) + \zeta_i'^n \Big) 1_{D_i^n}, \quad \text{where} \;\; D_i^n = \{(i-1)\Delta_n < S_p \leq i\Delta_n\}. \tag{8.18}$$

There is a problem here: it is easy to evaluate the conditional expectations of $|\zeta_i^n(t)|$ and $|\zeta_i'^n|$ w.r.t. $\mathcal{F}_{(i-1)\Delta_n}$ and to check that they go to 0, uniformly in i, but the set D_i^n is not $\mathcal{F}_{(i-1)\Delta_n}$-measurable. To overcome this difficulty we denote by $(\mathcal{G}_t)_{t \geq 0}$ the smallest filtration such that $\mathcal{G}_t$ contains $\mathcal{F}_t$ and $\sigma(S_p)$. Then W and the restriction $\underline{\mu}'$ of $\underline{\mu}$ to the set $\mathbb{R}_+ \times \{z : \gamma(z) \leq 1/m\}$ are still a Wiener process and a Poison random measure relative to this bigger filtration $(\mathcal{G}_t)$, and X' is driven by $\underline{\mu}'$.

Therefore applying (6.25) with $\alpha_n = 1$ and $r = 2$ and to the process X' instead of X, we get $\mathbb{E}(|\Delta_i^n X'| \wedge 1 \mid \mathcal{G}'_{(i-1)\Delta_n}) \leq \varepsilon_n$, where $\varepsilon_n \to 0$. Since $D_i^n \in \mathcal{G}'_{(i-1)\Delta_n}$ we then have

$$\mathbb{E}\Big(\sum_{i \geq 1} (|\zeta_i'^n| \wedge 1) 1_{D_i^n} \Big) = \mathbb{E}\Big(\sum_{i \geq 1} 1_{D_i^n} \, \mathbb{E}(|\zeta_i'^n| \wedge 1 \mid \mathcal{G}'_{(i-1)\Delta_n}) \Big) \leq \varepsilon_n \mathbb{E}\Big(\sum_{i \geq 1} 1_{D_i^n} \Big) = \varepsilon_n. \tag{8.19}$$

By Doob inequality and the fact that W is an $(\mathcal{F}'_t)$-Wiener process, for any $t \in ((i-1)\Delta_n, i\Delta_n]$, the conditional expectation $\mathbb{E}(|\zeta_i'^n(t)|^2 \mid \mathcal{G}'_{(i-1)\Delta_n}))$ is smaller than

$$K \mathbb{E}\Big(\int_{(i-1)\Delta_n}^t \|\sigma_u - \sigma_{(i-1)\Delta_n}\|^2 du + \int_t^{i\Delta_n} \|\sigma_u - \sigma_t\|^2 du \mid \mathcal{G}'_{(i-1)\Delta_n} \Big).$$

Then the same argument as above yields

$$\mathbb{E}\Big(\sum_{i\geq 1}|\zeta_i^n(T_p)|^2 1_{D_i^n}\Big) \;\leq$$
$$K\mathbb{E}\Big(\int_{\Delta_n[T_p/\Delta_n]-\Delta_n}^{T_p}\|\sigma_u-\sigma_{(i-1)\Delta_n}\|^2du+\int_{T_p}^{\Delta_n[T_p/\Delta_n]}\|\sigma_u-\sigma_t\|^2du\mid\mathcal{F}'_{(i-1)\Delta_n})\Big).$$

This quantity goes to 0 by Lebesgue theorem, because σ is right continuous with left limit, so this together with (8.18) and (8.19) gives us the first part of (8.17). □

8.4 Proof of Theorem 8.2.

Step 1) We begin with some preliminaries, to be used also for the next theorem. We fix $m\geq 1$ and let P_m be the set of all p such that $S_p=T(m',r)$ for some $r\geq 1$ and some $m'\leq m$ (see the previous subsection). We also set

$$X(m)_t \;=\; X_t-\sum_{p\in P_m:\ T_p\leq t}\Delta X_{S_p} \;=\; X_t-(\delta 1_{\{\gamma>1/m\}})\star\underline{\mu}_t. \tag{8.20}$$

Observe that, due to (6.21), and with the notation $b(m)_t=b_t-\int_{\{z:\gamma(z)>1/m\}}\delta(t,z)\lambda(dz)$, we have $X(m)=X'(m)+X''(m)$, where

$$X'(m)_t \;=\; X_0+\int_0^t b(m)_s ds+\int_0^t\sigma_s dW_s,\qquad X''(m)\;=\;(\delta 1_{\{\gamma\leq 1/m\}})\star(\underline{\mu}-\underline{\nu}). \tag{8.21}$$

Then we denote by $\Omega_n(t,m)$ the set of all ω satisfying the following for all $p\geq 1$:

$$\left.\begin{array}{ll} p,p'\in P_m,\ S_p(\omega)\leq t & \Rightarrow\quad |S_p(\omega)-S_{p'}(\omega)|>k\Delta_n,\\ 0\leq s\leq t,\ 0\leq u\leq k\Delta_n & \Rightarrow\quad \|X(m)_{s+u}(\omega)-X(m)_s(\omega)\|\leq 2/m.\end{array}\right\} \tag{8.22}$$

Since $\|\delta\|\leq\gamma$, implying $\|\Delta X(m)_s\|\leq 1/m$, we deduce that for all $t>0$ and $m\geq 1$:

$$\Omega_n(t,m)\;\to\;\Omega\quad\text{a.s. as}\quad n\to\infty. \tag{8.23}$$

If g is C^2 with $g(0)=0$ and $\nabla g(0)=0$, for any integer $l\geq 1$ and any d-dimensional semimartingale Z, we write $G^n(Z,g,l)_t \;=\; V(Z,g,l\Delta_n)_t-\sum_{s\leq l\Delta_n[t/l\Delta_n]}g(\Delta Z_s)$. Observe that on the set $\Omega_n(t,m)$ we have for all $s\leq t$ and $l=1$ or $l=k$:

$$G^n(X,g,l)_t \;=\; G^n(X(m),g,l)+Y^n(m,g,l), \tag{8.24}$$

where

$$\left.\begin{array}{rcl} Y^n(m,g,l)_t &=& \sum_{p\in P_m:\ S_p\leq l\Delta_n[t/l\Delta_n]}\zeta(g,l)_p^n,\\ \zeta(g,1)_p^n &=& g(\Delta X_{S_p}+\sqrt{\Delta_n}\,R_p^n)-g(\Delta X_{S_p})-g(\sqrt{\Delta_n}\,R_p^n)\\ \zeta(g,k)_p^n &=& g(\Delta X_{S_p}+\sqrt{\Delta_n}\,(R_p^n+R_p^{\prime\prime n}))-g(\Delta X_{S_p})-g(\sqrt{\Delta_n}\,(R_p^n+R_p^{\prime\prime n})).\end{array}\right\} \tag{8.25}$$

Step 2) Now we turn to the proof itself, with a function f satisfying the relevant assumptions. Recall in particular that $f(x)/\|x\| \to 0$ as $x \to 0$. A Taylor expansion in the expressions giving $\zeta(m,f,l)^n_q$ and Lemma 8.7 readily gives

$$\Big(\frac{1}{\sqrt{\Delta_n}}\,\zeta(f,1)^n_p, \frac{1}{\sqrt{\Delta_n}}\,\zeta(f,k)^n_p\Big)_{p\geq 1} \stackrel{\mathcal{L}-s}{\longrightarrow} \Big(\nabla f(\Delta X_{S_p})R_p, \nabla f(\Delta X_{S_p})R'_p\Big)_{p\geq 1}$$

(here, $\nabla f(\Delta X_{S_p})R_p$ for example stands for the q-dimensional vector with components $\sum_{i=1}^d \partial_i f_l(\Delta X_{S_p})R^i_p$). Since the sum giving $Y^n(m,f,l)_t$ has in fact finitely many entries, we deduce from well known properties of the Skorokhod topology that, as $n\to\infty$:

$$\left.\begin{array}{l}\text{the processes } \Big(\frac{1}{\sqrt{\Delta_n}}\,Y^n(m,f,1), \frac{1}{\sqrt{\Delta_n}}\,Y^n(m,f,k)\Big) \text{ converges}\\ \text{stably in law, in } \mathbb{D}(\mathbb{R}_+,\mathbb{R}^q)\times\mathbb{D}(\mathbb{R}_+,\mathbb{R}^q)\text{, to the process } (Z^m(f), Z'^m(f))\end{array}\right\} \tag{8.26}$$

where $(Z^m(f), Z'^m(f))$ is defined componentwise by (8.5), except that the sum is taken over all $p\in P_m$ only.

If we consider, say, the first component, we have by (8.6) and Doob's inequality:

$$\begin{aligned}\widetilde{\mathbb{E}}\Big(\sup_{s\leq t}|Z^m(f_1)_s - Z(f_1)_s|^2\Big) &= \widetilde{\mathbb{E}}\Big(\widetilde{\mathbb{E}}\Big(\sup_{s\leq t}|Z^m(f_1)_s - Z(f_1)_s|^2 \mid \mathcal{F}\Big)\Big)\\ &\leq 4\mathbb{E}\Big(\sum_{p\notin P_m,\ S_p\leq t}\ \sum_{i,j=1}^d (\partial_i f_1\partial_j f_1)(\Delta X_{S_p})(c^{ij}_{S_p-} + c^{ij}_{S_p})\Big).\end{aligned}$$

The variable of which the expectation is taken in the right side above is smaller than $K\sum_{s\leq t}\|\Delta X_s\|^2 1_{\{\|\Delta X_s\|\leq 1/m\}}$ (because c_t is bounded and if $p\notin P_m$ then $\|\Delta X_s\|\leq 1/m$), so by Lebesgue theorem this expectation goes to 0 as $m\to\infty$. The same argument works for the other components, and thus we have proved that

$$(Z^m(f), Z'^m(f)) \stackrel{\text{u.c.p.}}{\longrightarrow} (Z(f), Z'(f)). \tag{8.27}$$

Hence, in view of (8.26) and (8.27), and also of (8.23) and (8.24), it remains to prove

$$\lim_{m\to\infty}\ \limsup_{n\to\infty}\ \mathbb{P}\Big(\Omega_n(t,m)\cap\Big\{\sup_{s\leq t}\frac{1}{\sqrt{\Delta_n}}\,|G^n(X(m),f_r,l)_s| > \eta\Big\}\Big) = 0. \tag{8.28}$$

Step 3) Now we proceed to proving (8.28), and we drop the index r, pretending that f is 1-dimensional. It is also enough to consider the case $l=1$ (the case $l=k$ is the same, upon replacing everywhere Δ_n by $k\Delta_n$). We set

$$k(x,y) = f(x+y) - f(x) - f(y), \qquad g(x,y) = k(x,y) - \nabla f(x)y. \tag{8.29}$$

Recall that f is C^2 and that (6.21) and (8.20) hold. Then we apply Itô's formula to the process $X(m)_s - X(m)_{i\Delta_n}$ and the function f, for $t > i\Delta_n$ to get

$$\frac{1}{\sqrt{\Delta_n}}\,\Big(G^n(X(m),f,1)_t - G^n(X(m),f,1)_{i\Delta_n}\Big) = A(n,m,i)_t + M(n,m,i)_t, \tag{8.30}$$

where $M(n,m,i)$ is a locally square-integrable martingale with predictable bracket $A'(n,m,i)$, and with

$$A(n,m,i)_t = \int_{i\Delta_n}^t a(n,m,i)_u \, du, \qquad A'(n,m,i)_t = \int_{i\Delta_n}^t a'(n,m,i)_u \, du, \tag{8.31}$$

and

$$\left\{ \begin{array}{lll} a(n,m,i)_t &=& \frac{1}{\sqrt{\Delta_n}} \Big(\sum_{j=1}^d \partial_j f(X(m)_t - X(m)_{i\Delta_n}) b(m)_t^j \\ && + \frac{1}{2} \sum_{j,l=1}^d \partial^2_{jl} f(X(m)_t - X(m)_{i\Delta_n}) c_t^{jl} \\ && + \int_{\{z:\gamma(z)\leq 1/m\}} g(X(m)_t - X(m)_{i\Delta_n}, \delta(t,z)) \ \lambda(dz) \Big) \\ a'(n,m,i)_t &=& \frac{1}{\Delta_n} \Big(\sum_{j,l=1}^d (\partial_j f \ \partial_l f)(X(m)_t - X(m)_{i\Delta_n}) c_t^{jl} \\ && + \int_{\{z:\gamma(z)\leq 1/m\}} k(X(m)_t - X(m)_{i\Delta_n}, \delta(t,z))^2 \ \lambda(dz) \Big). \end{array} \right.$$

Now we set $T(n,m,i) = \inf(s > i\Delta_n : \ \|X(m)_s - X(m)_{i\Delta_n}\| > 2/m)$. On the set $\Omega_n(t,m)$ we have by construction $T(n,m,i) > (i+1)\Delta_n$ for all $i < [t/\Delta_n]$. Therefore in view of (8.30) we have on this set:

$$\begin{array}{rcl} \frac{1}{\sqrt{\Delta_n}} \sup_{s\leq t} |G^n(X(m),f,1)_s| &\leq& \sum_{i=1}^{[t/\Delta_n]} |A(n,m,i-1)_{(i\Delta_n)\wedge T(n,m,i-1)}| \\ && + \left| \sum_{i=1}^{[t/\Delta_n]} M(n,m,i-1)_{(i\Delta_n)\wedge T(n,m,i-1)} \right|. \end{array}$$

Henceforth in order to get (8.28), it is enough to prove the following:

$$\left. \begin{array}{lll} \lim_{m\to\infty} \ \limsup_n \ \mathbb{E}\Big(\sum_{i=1}^{[t/\Delta_n]} |A(n,m,i-1)_{(i\Delta_n)\wedge T(n,m,i-1)}| \Big) &=& 0, \\ \lim_{m\to\infty} \ \limsup_n \ \mathbb{E}\Big(\sum_{i=1}^{[t/\Delta_n]} A'(n,m,i-1)_{(i\Delta_n)\wedge T(n,m,i-1)} \Big) &=& 0. \end{array} \right\} \tag{8.32}$$

Recall that $f(0) = 0$ and $\nabla f(0) = 0$ and $\|\nabla^2 f(x)\| = o(\|x\|)$ as $x \to 0$, so we have

$$j = 0,1,2, \ \ \|x\| \leq \frac{3}{m} \quad \Rightarrow \quad \|\nabla^j f(x)| \ \leq \ \alpha_m \ \|x\|^{3-j} \tag{8.33}$$

for some α_m going to 0 as $m \to \infty$, which implies

$$\|x\| \leq \frac{3}{m}, \ \|y\| \leq \frac{1}{m} \quad \Rightarrow \quad |k(x,y)| \leq K\alpha_m \|x\| \ \|y\|, \quad |g(x,y)| \leq K\alpha_m \|x\| \ \|y\|^2. \tag{8.34}$$

Observe that $\|X(m)_{s\wedge T(n,m,i)} - X(m)_{i\Delta_n}\| \leq 3/m$ for $s \geq i\Delta_n$ (because the jumps of $X(m)$ are smaller than $1/m$). Then in view of (SH) and (8.34) and of the fact that $\|b(m)_t\| \leq Km$ we obtain for $i\Delta_n \leq t \leq T(n,m,i)$:

$$\left\{ \begin{array}{l} |a(n,m,i)_t| \leq \frac{K\alpha_m}{\sqrt{\Delta_n}} \ (\|X(m)_t - X(m)_{i\Delta_n}\| + m\|X(m)_t - X(m)_{i\Delta_n}\|^2), \\ a'(n,m,i)_t \leq \frac{K\alpha_m^2}{\Delta_n} \ \|X(m)_t - X(m)_{i\Delta_n}\|^2. \end{array} \right.$$

Now, exactly as for (6.23), one has $\mathbb{E}(\|X(m)_{t+s} - X(m)_t\|^p) \leq K_p(s^{p/2} + m^p s^p)$ for all $p \in (0,2]$ and $s,t \geq 0$, under (SH). Applying this with $p = 1$ and $p = 2$, respectively, gives that the two "lim sup" in (8.32) are smaller than $Kt\alpha_m$ and $Kt\alpha_m^2$ respectively. Then (8.32) holds, and we are finished.

8.5 Proof of Theorem 8.4.

We essentially reproduce the previous proof, with the same notation. Recall that $f(x) = (x^1x^2)^2$.

Step1) The assumption that X^1 and X^2 have no common jumps implies that $f(\Delta X_{S_p}) = 0$ and $\nabla f(\Delta X_{S_p}) = 0$ for all $p \geq 1$, whereas $f(x)/\|x\|^2 \to 0$ as $x \to 0$. Then a second order Taylor expansion in the expressions giving $\zeta(m,f,l)_q^n$ and Lemma 8.7 gives

$$\Big(\frac{1}{\Delta_n}\ \zeta(f,1)_p^n, \frac{1}{\Delta_n}\ \zeta(f,k)_p^n\Big)_{p\geq 1} \stackrel{\mathcal{L}-\mathrm{s}}{\longrightarrow} \Big(\frac{1}{2}\sum_{i,j=1}^d \partial^2 f(\Delta X_{S_p})R_p^iR_p^j, \frac{1}{2}\sum_{i,j=1}^d \partial^2 f(\Delta X_{S_p})R_p'^iR_p'^j\Big)_{p\geq 1}.$$

From this we deduce that, instead of (8.26), and as $n\to\infty$:

$$\left.\begin{array}{l}\text{the processes } \Big(\frac{1}{\Delta_n}\ Y^n(m,f,1), \frac{1}{\Delta_n}\ Y^n(m,f,k)\Big) \text{ converges}\\ \text{stably in law, in } \mathbb{D}(\mathbb{R}_+,\mathbb{R})\times\mathbb{D}(\mathbb{R}_+,\mathbb{R}), \text{ to the process } \frac{1}{2}(\overline{Z}^m(f),\overline{Z}'^m(f))\end{array}\right\}$$

where $(\overline{Z}^m(f),\overline{Z}'^m(f))$ is defined componentwise by (8.7), except that the sum is taken over all $p\in P_m$ only. By Lebesgue theorem, we readily obtain

$$(\overline{Z}^m(f),\overline{Z}'^m(f)) \stackrel{\text{u.c.p.}}{\longrightarrow} (\overline{Z}(f),\overline{Z}'(f)).$$

Hence, in view of (8.23) and (8.24), and since here $G(X(m),f,l) = V(X(m),f,l\Delta_n)$, it remains to prove that with the notation $\overline{C}_t = \int_0^t (c_u^{ii}c_u^j + 2(c_u^{ij})^2)du$ we have for all $t,\eta>0$ and for $l=1$ and $l=k$:

$$\lim_{m\to\infty}\ \limsup_{n\to\infty}\ \mathbb{P}\Big(\sup_{s\leq t}\frac{1}{\Delta_n}\ |V(X(m),f,l\Delta_n)_s - l\overline{C}_{l\Delta_n[s/l\Delta_n]}| > \eta\Big) = 0. \tag{8.35}$$

Step 2) Recall (8.21), and set $g(x) = \|x\|^4$. By Theorem 6.2 applied to the process $X'(m)$ we have for each $m\geq 1$:

$$\frac{1}{l\Delta_n}\ V(X'(m),f,l\Delta_n) \stackrel{\text{u.c.p.}}{\longrightarrow} \overline{C}, \quad \frac{1}{l\Delta_n}\ V(X'(m),g,l\Delta_n)_t \stackrel{\text{u.c.p.}}{\longrightarrow} \int_0^t \rho_{\sigma_u}(g)du. \tag{8.36}$$

Therefore for getting (8.35) it is enough to prove that

$$\lim_{m\to\infty}\ \limsup_{n\to\infty}\ \mathbb{P}\Big(\sup_{s\leq t}\frac{1}{\Delta_n}\ |V(X(m),f,l\Delta_n)_s - V(X'(m),f,l\Delta_n)_s| > \eta\Big) = 0. \tag{8.37}$$

Here again, it is obviously enough to prove the result for $l=1$.

Now, the special form of f implies that for each $\varepsilon>0$ there is a constant K_ε with

$$|f(x+y)-f(x)| \ \leq\ \varepsilon\|x\|^4 + K_\varepsilon\|x\|^2\,\|y\|^2 + K_\varepsilon f(y),$$

hence

$$|V(X(m),f,\Delta_n) - V(X'(m),f,\Delta_n)| \ \leq\ \varepsilon V(X'(m),g,\Delta_n) + K_\varepsilon(U^n + U'^n),$$

where

$$U^n_t \;=\; \sum_{i=1}^{[t/\Delta_n]} (\Delta^n_i X(m)^1)^2\,(\Delta^n_i X'(m)^2)^2, \qquad U'^n_t \;=\; \sum_{i=1}^{[t/\Delta_n]} \|\Delta^n_i X(m)\|^2\,\|\Delta^n_i X'(m)\|^2.$$

$\varepsilon>0$ being arbitrarily small, by the second part of (8.36) it is then enough to prove

$$\lim_{m\to\infty}\ \limsup_{n\to\infty}\ \frac{1}{\Delta_n}\,\mathbb{E}(U^n_t) \;=\; 0, \qquad \lim_{m\to\infty}\ \limsup_{n\to\infty}\ \frac{1}{\Delta_n}\,\mathbb{E}(U'^n_t) \;=\; 0. \tag{8.38}$$

Step 3) Exactly as in the proof of Lemma 6.8, we have

$$\mathbb{E}(\|X''(m)_{t+s}-X''(m)_t\|^2) \;\le\; \alpha_m\, s, \quad \text{where} \quad \alpha_m=\int_{\{z:\gamma(z)\le 1/m\}} \gamma(z)^2\lambda(dz). \tag{8.39}$$

Now, as for (8.30), we deduce from Itô's formula that

$$(\Delta^n_i X''^1(m))^2\,(\Delta^n_i X''^2(m))^2 \;=\; M(n,m,i)_{i\Delta_n} + \int_{(i-1)\Delta_n}^{i\Delta_n} a(n,m,i-1)_s ds, \tag{8.40}$$

where $M(n,m,i)$ is a martingale and $a(n,m,i)_t=H_m(X''(m)_t-X''(m)_{(i-1)\Delta_n})$ and

$$H_m(x) \;=\; \int_{\{z:\gamma(z)\le 1/m\}} \Big(f(x+\delta(t,z))-f(x)-\nabla f(x)\delta(t,z)\Big)\ \lambda(dz).$$

Now, since X^1 and X^2 have no common jumps, we have $\delta(\omega,t,z)^1\delta(\omega,t,z)^2 = 0$ for λ-almost all z. Therefore a simple calculation shows that

$$H(x) \;=\; \int_{\{z:\gamma(z)\le 1/m\}} \Big((x^1)^2(\delta(t,z)^2)^2+(x^2)^2(\delta(t,z)^1)^2\Big)\ \lambda(dz),$$

and thus

$$0 \;\le\; H_m(x) \;\le \alpha_m\,\|x\|^2.$$

Recall also that $\mathbb{E}(\|X''(m)_{t+s}-X''(m)_t\|^2)\le Kt$. Then taking the expectation in (8.40) gives us, with

$$\mathbb{E}((\Delta^n_i X(m)^1)^2\,(\Delta^n_i X'(m)^2)^2) = \mathbb{E}\left(\int_{(i-1)\Delta_n}^{i\Delta_n} H_m(X''(m)_s-X''(m)_{(i-1)\Delta_n})\, ds\right) \le \alpha_m \Delta_n^2.$$

Then, since $\alpha_m\to 0$ as $m\to\infty$, we readily deduce the first part of (8.38).

It remains to prove the second part of (8.38). Itô's formula again yields

$$\|\Delta^n_i X''(m)\|^2\,\|\Delta^n_i X'(m)\|^2 \;=\; M'(n,m,i)_{i\Delta_n} + \int_{(i-1)\Delta_n}^{i\Delta_n} a'(n,m,i-1)_s ds, \tag{8.41}$$

where $M'(n,m,i)$ is a martingale and $a'(n,m,i)_t=H'_m(X'(m)_t-X'(m)_{(i-1)\Delta_n}, X''(m)_t-X''(m)_{(i-1)\Delta_n})$ and

$$H'_m(x,y) \;=\; 2\|y\|^2\sum_{i=1}^d b(m)^i_t\, x^i + \|y\|^2\sum_{i=1}^d c^{ii}_t + \|x\|^2\int_{\{z:\gamma(z)\le 1/m\}} \|\delta(t,z))\|^2\ \lambda(dz),$$

and thus

$$|H'_m(x,y)| \ \leq \ K\Big(\alpha_m\|x\|^2 + \|y\|^2(1+m\|x\|\Big)$$

because $\|b(m)_t\| \leq Km$ and $\|c\| \leq K$. Then using (8.39) and $\mathbb{E}(\|X'(m)_{t+s} - X'(m)_t\|^p) \leq K_p(s^{p/2} + m^p s^p)$ for all $p > 0$, we deduce from Cauchy-Schwarz inequality, and by taking the expectation in (8.41), that

$$\begin{aligned}
&\mathbb{E}(\|\Delta_i^n X''(m)\|^2 \, \|\Delta_i^n X'(m)\|^2) \\
&= \mathbb{E}\left(\int_{(i-1)\Delta_n}^{i\Delta_n} H'_m(X'(m)_s - X'(m)_{(i-1)\Delta_n}, X''(m)_s - X''(m)_{(i-1)\Delta_n})\, ds\right) \\
&\leq \ K\Delta_n^2\Big(\alpha_m(1+m^2\Delta_n) + m\sqrt{\alpha_m\Delta_n}\ (1+m\Delta_n)\Big).
\end{aligned}$$

Then again since $\alpha_m \to 0$ as $m \to \infty$, we deduce the second part of (8.38), and the proof is finished.

9 Estimation of the integrated volatility

At this point we have established the theoretical results which are needed for the statistical problems we have in mind, and we can turn to these problems. We start by a warning, which applies to all problems studied below:

The underlying process X is observed at times $0, \Delta_n, 2\Delta_n, \cdots$ without measurement errors.

This assumption is clearly *not satisfied* in general in the context of high-frequency data, at least in finance where there is an important microstructure noise. However, dealing with measurement errors involves a lot of complications which would go beyond the scope of this course.

As said before the first and probably the most important question is the estimation of the integrated volatility, at least when the underlying process is continuous. This is the object of this section.

9.1 The continuous case.

Here we assume that the underlying process X is a *continuous* Itô semimartingale, i.e. is of the form

$$X_t \ = \ X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s. \tag{9.1}$$

Most of the literature is concerned with the 1-dimensional case, but mathematically speaking there is no complication whatsoever in considering the d-dimensional case: so above W is a d'-dimensional Wiener process, and b_t and σ_t are d and $d \times d'$-dimensional (so implicitly in (9.1) the second integral is in fact a sum of stochastic integrals w.r.t. the various components W^j of W).

Our aim is to "estimate" the *integrated volatility*, that is the quadratic variation-covariation process of X:

$$C_t^{jk} \;=\; \int_0^t c_s^{jk} ds, \qquad \text{where} \quad c_t = \sigma_t \sigma_t^\star. \tag{9.2}$$

Recall that the process X is observed at the discrete times $0, \Delta_n, 2\Delta_n, \cdots$ over a finite interval $[0,T]$, and one wants to infer C_T, or sometimes the increments $C_t - C_s$ for some pairs (s,t) with $0 \leq s \leq t \leq T$. Each of these increments is a random variable taking values in the set of $d \times d$ symmetric nonnegative matrices.

One point should be mentioned right away, and is in force not only for the integrated volatility but for all quantities estimated in this course: although we speak about estimating the matrix C_T, it is *not* a statistical problem in the usual sense since the quantity to estimate is a *random variable*; so the "estimator", say $\widetilde{C}_T^n$ (the "n" is here to emphasize that it is a function of the observation $(X_{i\Delta_n} : 0 \leq i \leq [T/\Delta_n]))$ does not estimate a parameter, but a variable which depends on the outcome ω, and the quality of this estimator is something which fundamentally depends on ω as well.

Nevertheless we are looking for estimators which behave as in the classical case, asymptotically as $n \to \infty$ (that is, as $\Delta_n \to 0$). We say that $\widetilde{C}_T^n$ is *consistent* if $\widetilde{C}_T^n$ converges in probability to C_T (one should say "weakly" consistent; of course in the present setting, even more than in classical statistics, on would like to have estimators which converge *for all* ω, or at least almost surely, but this is in general impossible to achieve). Then we also aim to a *rate of convergence*, and if possible to a limit theorem so as to allow for quantitatively asserting the quality of the estimator and for constructing confidence intervals, for example.

Two consistent estimators can be compared on the basis of their rates of convergence and, if those are the same, on their asymptotic variances for example. However, unlike in classical statistics, we do not have a theory for asymptotic optimality, like the LAN or LAMN theory. The best one can do is to check whether our estimators are asymptotically optimal (in the usual sense) when the problem reduces to a classical parametric problem, that is when C_T is deterministic (this happens when for example the volatility σ_t is not random, like in the Black-Scholes model for the log-returns).

After these lengthy preliminaries we now introduce the estimator. Of course all authors use the approximated quadratic variation given in (1.9), and often called "realized volatility". Since we are in the d-dimensional case, we have a matrix $B(2,\Delta_n)_t$ with components

$$B(2,\Delta_n)_t^{jk} \;=\; \sum_{i=1}^{[t/\Delta_n]} \Delta_i^n X^j \Delta_i^n X^k. \tag{9.3}$$

These estimators have the following properties:

Property 9.1 (Consistency) $B(2,\Delta_n)_t \xrightarrow{\mathbb{P}} C_t$.

Property 9.2 (Asymptotic normality-1) $\frac{1}{\sqrt{\Delta_n}} \Big(B(2,\Delta_n)_t - C_t\Big)$ converges in law to a $d \times d$-dimensional variable which, conditionally on the path of X over $[0,t]$, is centered normal with variance-covariance (Γ_t^{jklm}) (the covariance of the (jk) and the (lm) components) given by

$$\Gamma_t^{jklm} = \int_0^t (c_s^{jl} c_s^{km} + c_s^{jm} c_s^{kl}) ds. \tag{9.4}$$

These are obvious consequences of Theorem 5.1-(b) and Corollary 7.2: for the consistency there is no assumption other than (9.1); for the asymptotic normality we need (H) in these notes, but in fact it is enough that $\int_0^t \|c_s\|^2 ds < \infty$ a.s. (see [14]).

Property 9.2 gives a rate of convergence equal to $1/\sqrt{\Delta_n}$, but the name "asymptotic normality" is not really adequate since the limiting variable after centering and normalization is not unconditionally normal, and indeed it has a law which is essentially unknown. So it is useless in practice. But fortunately we not only have the convergence in law, but also the *stable* convergence in law. That is, as soon as one can find a sequence Γ_t^n of variables, depending on the observations at stage n only, and which converge in probability to the variance given by (9.4), then by normalizing once more by the square-root of the inverse of Γ_t^n (supposed to be invertible), we get a limit which is standard normal.

The "complete" result involving all components of C_t at once is a bit messy to state. In practice one is interested in the estimation of a particular component C_t^{jk} (often with $k = j$ even). So for simplicity we consider below the estimation of a given component C_t^{jk}. The asymptotic variance is $\Gamma_t^{jkjk} = \int_0^t (c_s^{jj} c_s^{kk} + (c_s^{jk})^2) ds$ and we need an estimator for Γ_t^{jkjk}, which is provided by Theorem 6.2. More specifically, this theorem implies that

$$\Gamma(\Delta_n)_t^{jkjk} = \frac{1}{\Delta_n} \sum_{i=1}^{[t/\Delta_n]} \Big((\Delta_i^n X^j)^2 (\Delta_{i+1}^n X^k)^2 + \Delta_i^n X^j \Delta_i^n X^k \Delta_{i+1}^n X^j \Delta_{i+1}^n X^k \Big) \tag{9.5}$$

converges in probability to Γ_t^{klkl}. Therefore we have the following *standardized CLT* :

Theorem 9.3 (Asymptotic normality-2) *Assume (H). With the previous notation, and in restriction to the set $\{\Gamma_t^{jkjk} > 0\}$, the variables*

$$\frac{1}{\sqrt{\Delta_n \, \Gamma(\Delta_n)_t^{jkjk}}} \Big(B(2, \Delta_n)_t^{jk} - C_t^{jk} \Big) \tag{9.6}$$

converge stably in law to an $\mathcal{N}(0,1)$ random variable independent of $\mathcal{F}$.

The reader will notice the proviso "in restriction to the set $A := \{\Gamma_t^{jkjk} > 0\}$". This set is in fact equal to the set where $s \mapsto c_s^{jj}$ and $s \mapsto c_s^{kk}$ are not Lebesgue-almost surely vanishing on $[0,t]$, and also $\mathbb{P}$-a.s. to the set where neither one of the two paths $s \mapsto X_s^j$ and $s \mapsto X_s^k$ is of finite variation over $[0,t]$. So in practice $A = \Omega$ and the above is the mere (stable) convergence in law.

When $A \neq \Omega$, the stable convergence in law in restriction to A means that $\mathbb{E}(f(T_n)Y) \to \mathbb{E}(Y)\widetilde{\mathbb{E}}(f(U))$ for all bounded continuous functions f and all $\mathcal{F}$-measurable bounded variables Y vanishing outside A, and where T_n is the statistics in (9.6) and U is $\mathcal{N}(0,1)$.

This result is immediately applicable in practice, in contrast to Property 9.2: it may be used to derive confidence intervals for example, in the customary way.

Proof. As above, T_n is the variable (9.6), and we also set $S_n = \frac{1}{\sqrt{\Delta_n}} \left(B(2,\Delta_n)_t^{jk} - C_t^{jk}\right)$. We know that S_n converges stably in law to a variable which can be expressed as the product $\sqrt{\Gamma_t^{jkjk}}\, U$, where U is $\mathcal{N}(0,1)$ and independent of $\mathcal{F}$. By the properties of the stable convergence in law, and since $\Gamma(\Delta_n)_t^{jkjk} \stackrel{\mathbb{P}}{\longrightarrow} \Gamma_t^{jkjk}$, we also have stable convergence of the pair $(S_n, \Gamma(\Delta_n)_t^{jkjk})$ towards $\left(\sqrt{\Gamma_t^{jkjk}}\, U, \Gamma(\Delta_n)_t^{jkjk}\right)$, Obviously this also holds in restriction to the set A described above. Since $T_n = S_n/\sqrt{\Gamma(\Delta_n)_t^{jkjk}}$ and $\Gamma(\Delta_n)_t^{jkjk} \stackrel{\mathbb{P}}{\longrightarrow} \Gamma_t^{jkjk} > 0$ on A, the result follows from the continuous mapping theorem. □

Remark 9.4 When $\sigma_t(\omega) = \sigma$ is a constant matrix, so up to the drift the process X is a Wiener process, then we are in the classical setting of estimation of a matrix-valued parameter $c = \sigma\sigma^*$. In this case we have the LAN property, and it is well known that the estimators $B(2,\Delta_n)_t$ are asymptotically efficient for estimating c in this setting (and when the drift vanishes, it is even the MLE). Note that c is identifiable, but usually not σ itself since there might be many square-roots σ for the matrix c. □

Remark 9.5 There are many ways, indeed, to find consistent estimators for Γ_t^{jkjk}, and (9.5) is just possibility. A full set of consistent estimators is provided by the formulas below, where q is a non-zero integer (recall that m_r is the rth absolute moment of $\mathcal{N}(0,1)$):

$$\Gamma(q,\Delta_n)_t^{jkjk} = \frac{1}{8m_{2/q}^{2q}\Delta_n} \sum_{i=1}^{[t/\Delta_n]} g_q^{jk}(\Delta_i^n X, \Delta_{i+1}^n X, \cdots, \Delta_{i+2q-1}^n X), \tag{9.7}$$

where

$$\begin{aligned} g_q^{jk}(x_1,\cdots,x_{2q}) &= \prod_{i=1}^{2q} |x_i^j + x_i^k|^{2/q} + \prod_{i=1}^{2q} |x_i^j - x_i^k|^{2/q} - 2\prod_{i=1}^{2q} |x_i^j|^{2/q} \\ &\quad -2\prod_{i=1}^{2q} |x_i^k|^{2/q} + 4\prod_{i=1}^{q} |x_i^j|^{2/q} \prod_{i=q+1}^{2q} |x_i^k|^{2/q}. \end{aligned} \tag{9.8}$$

Indeed a simple computation yields that $\rho_t(g_q^{jk}) = 8m_{2/q}^{2q}\left(c_t^{jj}c_t^{kk} + (c_t^{jk})^2\right)$, so the property $\Gamma(q,\Delta_n)_t^{jkjk} \stackrel{\mathbb{P}}{\longrightarrow} \Gamma_t^{jkjk}$ again follows from Theorem 6.2. And of course one could make variations on this formula, like taking various powers summing up to 4 instead of the uniform power $2/q$, or varying the order in which the components x_i^j and x_i^k are taken in the last term of (9.8): for example one could take $2\prod_{i=1}^q |x_i^j|^{2/q} \prod_{i=q+1}^{2q} |x_i^k|^{2/q} + 2\prod_{i=1}^q |x_i^k|^{2/q} \prod_{i=q+1}^{2q} |x_i^j|^{2/q}$ instead of the last term in (9.8): then, with this substitution, we have in fact $\Gamma(q,\Delta_n)^{jkjk} = \Gamma(\Delta_n)^{jkjk}$ when $q = 2$.

The important fact is that Theorem 9.3 is unchanged, if $\Gamma(\Delta_n)_t^{jkjk}$ is substituted with $\Gamma(q,\Delta_n)_t^{jkjk}$. □

9.2 The discontinuous case.

Now we come back to the general situation, where X is an Itô semimartingale satisfying (H). In this situation the integrated volatility is probably of less importance than in the continuous case because it captures only a part of the behavior of X and says nothing about jumps, but still many people wish to estimate it.

In this case things are more complicated. For example $B(2,\Delta_n)_t$ is no longer a consistent estimator for C_t, as seen in (5.3). However we have constructed in Section 6 some consistent estimators:

Property 9.6 (Consistency) Assuming (H), the truncated variation $V^{jk}(\varpi,\alpha,\Delta_n)_t$ of (6.6), and the multipower variation $V^{jk}(r_1,\cdots,r_l,\Delta_n)_t$ of (6.10) converge in probability to C_t^{jk}, for all $\alpha>0$ and $\varpi\in(0,\frac{1}{2})$ for the first one, and for all integer $l\geq 2$ and all $r_1,\cdots,r_l>0$ with $r_1+\cdots+r_l=2$ for the second one.

This is nice enough, but the associated CLTs need some more assumption, as seen in Theorems 7.1 and 7.4. In Theorem 7.1 we need the test function f to be bounded when X jumps, and this precludes the use of multipower variations; hence in these notes we actually have a CLT for truncated powers only, as a consequence of Theorem 7.4 (we do have a CLT for multipower variations as well, under the same assumption $r<1$ as below, but it is slightly too complicated to prove here; see however [9] for the Lévy case).

Property 9.7 (Asymptotic normality-1) Assume (H) and that $\int(\gamma(z)^r\wedge 1)\lambda(dz)<\infty$ for some $r\in[0,1)$. If $\alpha>0$ and $\varpi\in[\frac{1}{2(2-r)},\frac{1}{2})$ then the $d\times d$-dimensional processes with components $\frac{1}{\sqrt{\Delta_n}}\left(V^{jk}(\varpi,\alpha,\Delta_n)_t-C_t^{jk}\right)$ converge in law to a $d\times d$-dimensional variable which, conditionally on the path of X over $[0,t]$, is centered normal with variance-covariance (Γ_t^{jkjk}) given by (9.4).

The comments made after property 9.2, about the need for a standardized version of the CLT, are in order here. We need a consistent estimator for Γ_t^{jkjk}. Of course (9.5) does not any longer provide us with such an estimate, but we can use the "truncated" version

$$\begin{aligned}\Gamma'(\varpi,\alpha;\Delta_n)_t^{jkjk} \;=\; & \frac{1}{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\Big((\Delta_i^nX^j)^2(\Delta_{i+1}^nX^k)^2\\ & +\Delta_i^nX^j\Delta_i^nX^k\Delta_{i+1}^nX^j\Delta_{i+1}^nX^k\Big)1_{\{\|\Delta_i^nX\|\leq\alpha\Delta_n^{\varpi},\|\Delta_{i+1}^nX\|\leq\alpha\Delta_n^{\varpi}\}}.\end{aligned}\tag{9.9}$$

By virtue of (6.9), we have $\Gamma'(\varpi,\alpha;\Delta_n)_t^{jkjk}\stackrel{\mathbb{P}}{\longrightarrow}\Gamma_t^{jkjk}$. Then the same proof as for Theorem 9.3 gives:

Theorem 9.8 (Asymptotic normality-2) *Assume (H) and that* $\int(\gamma(z)^r\wedge 1)\lambda(dz)<\infty$ *for some* $r\in[0,1)$. *If* $\alpha>0$ *and* $\varpi\in[\frac{1}{2(2-r)},\frac{1}{2})$. *Then and in restriction to the set* $\{\Gamma_t^{jkjk}>0\}$, *the variables*

$$\frac{1}{\sqrt{\Delta_n\,\Gamma'(\varpi,\alpha;\Delta_n)_t^{jkjk}}}\left(V^{jk}(\varpi,\alpha;\Delta_n)_t^{jk}-C_t^{jk}\right)\tag{9.10}$$

converge stably in law to an $\mathcal{N}(0,1)$ *random variable independent of* $\mathcal{F}$.

One could also use multipower variations to estimate Γ_t^{jkjk}.

Remark 9.9 The assumption $\int(\gamma(z)^r \wedge 1)\lambda(dz) < \infty$ for some $r \in [0,1)$ is quite restrictive, but so far there is no known estimator for C_t with a rate $1/\sqrt{\Delta_n}$, if this fails. However we do have a (worse) rate in almost every situation. Namely, if $\int(\gamma(z)^r \wedge 1)\lambda(dz) < \infty$ for some $r \in [0,2)$ then the sequence $\frac{1}{\Delta_n^{(2-r)\varpi}}\left(V^{jk}(\varpi,\alpha,\Delta_n)_t - C_t^{jk}\right)$ is tight (or, bounded in probability), see [15]. This does not give a limit theorem, which we do not know to exist, but it is a bound for the rate.

Note that the rate gets worse when r approaches 2, and does not exist when $r = 2$ (that is, with no special assumption on the jumps). This is because, when $r \to 2$, the discontinuous part $\kappa(\delta) \star (\underline{\mu} - \underline{\nu})$ of the process X gets closer to a Brownian motion in some sense. To take a more specific example, the symmetric stable processes of index $\alpha \in (0,2)$ (which satisfy the above assumption for $r > \alpha$ and not for $r \leq \alpha$) converge to the Brownian motion as $\alpha \to 0$. The fact that the rate worsens when r increases is not surprising: it is more and more difficult to distinguish between the continuous part X^c and the discontinuous part when r approaches 2.

9.3 Estimation of the spot volatility.

If one is so much interested in the integrated volatility it is probably because one does not really know how to estimate the volatility c_t itself. In principle the knowledge of the process C_t entails the knowledge of its derivative c_t as well. But practically speaking, with discrete observations, the estimation of c_t is quite another matter, and we are not going to give here a serious account on the subject, which still features many open problems.

Let us just say a few words. This is very much like a non-parametric problem for which one wants to estimate an unknown function f, for example the density of a sequence of n i.i.d. variables. In this case, and depending of course of the kind of criterion one chooses (one can consider the estimation error pointwise, or in some $\mathbb{L}^p$), the rate of convergence of the best estimators strongly depends on the smoothness of the estimated function f, although this smoothness is usually not known beforehand. More precisely, if f is "r-Hölder" (that is, Hölder with index r when $r \in (0,1]$, and if $r > 1$ it means that f is $[r]$ times differentiable and its $[r]$th derivative is $(r - [r])$-Hölder), typically the rate of convergence of the best non-parametric estimators is $n^{r/(1+2r)}$, always smaller than $n^{1/2}$.

Here, the unknown function is $t \mapsto c_t(\omega)$, for a given ω. If it were not dependent of ω and if X were simply (say, in the 1-dimensional case) $X_t = \int_0^t \sqrt{c_s}\, dW_s$, the observed increments $\Delta_i^n X$ would be independent, centered, with variances $\int_{(i-1)\Delta_n}^{i\Delta_n} c_s ds$. That is, we would have a genuine non-parametric problem and the rate of convergence of "good" estimators would indeed be $\Delta_n^{-r/(1+2r)}$ with r being the smoothness of the function c_t in the above sense. Now of course c_t is random, and possibly discontinuous, and X has also a drift and possibly jumps.

When σ_t is an Itô semimartingale (hypothesis (H)) and is further continuous, then the path of $t \mapsto c_t$ are a.s. Hölder with any index $r < 1/2$, and not Hölder with index 1/2. And worse, σ_t can be discontinuous. Nevertheless, one expects estimators which converge at the rate $\Delta_n^{-1/4}$ (the rate when $r = 1/2$). This is what happens for the most elementary kernel estimators which are

$$U_t^{n,jk} \;=\; \frac{1}{k_n\Delta_n}\sum_{i\in I_n(t)} \Delta_i^n X^j \Delta_i^n X^k \, 1_{\{\|\Delta_i^n X\|\leq \alpha\Delta_n^{\varpi}\}}, \tag{9.11}$$

where α and ϖ are as before, and the sequence k_n of integers goes to ∞ with $\Delta_n k_n \to 0$ (as in (6.12)), and $I_n(t)$ is a set of k_n consecutive integers containing $[t/\Delta_n]$. This formula should of course be compared with (6.14). The "optimal" choice, as far as rates are concerned, consists in taking $k_n \sim 1/\sqrt{\Delta_n}$, and it is even possible to prove that the variables $\frac{1}{\Delta_n^{1/4}}\left(U_t^{n,jk} - c_t^{jk}\right)$ converge in law under appropriate conditions (this is *not* a functional CLT, and the limit behaves, as t varies, as a white noise) .

10 Testing for jumps

This section is about testing for jumps. As before we observe the process X at discrete times $0, \Delta_n, \cdots$ over a finite interval, and on the basis of these observations we want to decide whether the process has jumps or not. This is a crucial point for modeling purposes, and assuming that there are jumps brings out has important mathematical and financial consequences (option pricing and hedging, portfolio optimization).

It would seem that a simple glance at the dataset should be sufficient to decide this issue, and this is correct when a "big" jump occurs. Such big jumps usually do not belong to the model itself, and either they are considered as breakdowns in the homogeneity of the model, or they are dealt with using different methods like risk management. On the other hand, a visual inspection of most time series in finance does not provide a clear evidence for either the presence or the absence of small or medium sized jumps.

Determining whether a process has jumps has been considered by a number of authors. Let us quote for example [1], [11], [7], [16], [12] and [17]. Here we closely follow the approach initiated in [3].

10.1 Preliminary remarks.

The present problem is 1-dimensional: if X jumps then at least one of its components jumps, so we can and will assume below that X is 1-dimensional (in the multidimensional case one can apply the forthcoming procedure to each of the components successively). We will also strengthen Hypothesis (H) in a rather innocuous way:

Assumption (K): We have (H); furthermore with the notation $S = \inf(t : \Delta X_t \neq 0)$, we have:

(a) $C_t > 0$ when $t > 0$,

(b) $t \mapsto \int \kappa(\delta(\omega, t, z))\lambda(dz)$ is left-continuous with right limits on the set $(0, S(\omega)]$. $\square$

(a) above is a non-degeneracy condition for the continuous martingale part X^c. As for (b), it may appear as a strong assumption because it supposes that $z \mapsto \kappa(\delta(\omega, t, z))$ is λ-

integrable if $t < S(\omega)$. However one may remark that is "empty" on the set where $S(\omega) = 0$, that is where X has infinitely many jumps near the origin. It is also automatically implied by (H) when $\int(\gamma(z) \wedge 1)\lambda(dz) < \infty$. Moreover, if $F = \{(\omega, t, z) : \ \delta(\omega, t, z) \neq 0\}$, the variable $1_F \star \underline{\mu}_t$ is the number of jumps of X on the interval $(0, t]$, so by the very definition of S we have $1_F \star \underline{\mu}_S \leq 1$. Since F is predictable and $\underline{\nu}$ is the predictable compensator of $\underline{\mu}$, we have

$$\mathbb{E}\left(\int_0^S ds \int 1_F(s, z)\lambda(dz)\right) \; = \; \mathbb{E}(1_F \star \underline{\nu}_S) \; = \; \mathbb{E}(1_F \star \underline{\mu}_S) \; \leq \; 1.$$

Therefore outside a $\mathbb{P}$-null set we have $\int_0^S ds \int 1_F(s, z)\lambda(dz) < \infty$ and thus, upon modifying δ on a $\mathbb{P}$-null set, $z \mapsto \kappa(\delta(\omega, t, z))$ is λ-integrable if $t < S(\omega)$. So the condition (b) is really a very mild additional smoothness assumption, of the same nature as (b) of (H).

Before getting started we begin with a very important remark: Suppose that we are in the ideal situation where the path of $t \mapsto X_t(\omega)$ is fully observed over the time interval $[0, T]$. Then we know whether the path jumps or not, but we know nothing about other paths; so, exactly as for the integrated volatility in the previous section we can at the best make an inference about the outcome ω which is (partially) observed. But here there is even more: if we find that there are jumps we should conclude to a model with jumps, of course. But if we find no jump it does not really mean that the model should not have jumps, only that our particular observed path is continuous (and, if jumps occur like for a compound Poisson process, for instance, although the model should include jumps we always have a positive probability that a path does not jump over $[0, T]$).

Therefore, the problem which we really try to solve here is to decide, on the basis of the observations $X_{i\Delta_n}$, in which of the following two complementary sets the path which we have discretely observed falls:

$$\left.\begin{array}{l} \Omega_T^j = \{\omega : \ s \mapsto X_s(\omega) \text{ is discontinuous on } \ [0, T]\} \\ \Omega_T^c = \{\omega : \ s \mapsto X_s(\omega) \text{ is continuous on } \ [0, T]\}. \end{array}\right\} \tag{10.1}$$

10.2 The level and the power function of a test.

In view of (10.1) we have two possibilities for the "null hypothesis", namely "there are no jumps" (that is, we are in Ω_T^c), and "there are jumps" (that is, we are in Ω_T^c).

Consider for example the first case where the null hypothesis is "no jump". We are thus going to construct a critical (rejection) region $C_{T,n}^c$ at stage n, which should depend only on the observations $X_0, X_{\Delta_n}, \cdots, X_{\Delta_n[t/\Delta_n]}$. We are not here in a completely standard situation: the problem is asymptotic, and the hypothesis involves the outcome ω.

In a classical asymptotic test problem, the unknown probability measure $\mathbb{P}_\theta$ depends on a parameter $\theta \in \Theta$ (Θ can be a functional space), and the null hypothesis corresponds to θ belonging to some subset Θ_0 of Θ. At stage n one constructs a critical region C_n. The asymptotic level is

$$\alpha \; = \; \sup_{\theta \in \Theta_0} \; \limsup_n \; \mathbb{P}_\theta(C_n), \tag{10.2}$$

whereas the asymptotic power function is defined on $\Theta_1 = \Theta \backslash \Theta_0$ as

$$\beta(\theta) \;=\; \liminf_n \; \mathbb{P}_\theta(C_n).$$

Sometimes one exchanges the supremum and the lim sup in (10.2), which is probably more sensible but in general impossible to achieve, in the sense that often $\sup_{\theta\in\Theta_0} \mathbb{P}_\theta(C_n) = 1$. Moreover, usually a prescribed level α_0 is given, and the aim is to construct C_n so that (10.2) holds with $\alpha \leq \alpha_0$ (and if possible even, $\alpha = \alpha_0$, which generally increases the power function). Finally a "good" asymptotic critical region satisfies $\beta(\theta) = 1$ for all $\theta \in \Theta_1$ (we cannot hope for $\mathbb{P}_\theta(C_n) = 1$ if $\theta \in \Theta_1$ at any stage n).

In the present situation we have no genuine parameter (although the law of X itself can in a sense be considered as a parameter, or perhaps its characteristics (B, C, ν) can). Rather, the outcome ω, or at least the fact that it lies in Ω_t^c or not, can be considered as a kind of parameter. So, keeping the analogy with (10.2), we are led to consider the following definition for the asymptotic level of our critical region $C_{t,n}^c$:

$$\alpha_t^c \;=\; \sup\Big(\limsup_{n\to\infty} \; \mathbb{P}(C_{t,n}^c \mid A) : \; A \in \mathcal{F}, \; A \subset \Omega_t^c\Big). \tag{10.3}$$

Here $\mathbb{P}(C_{t,n}^c \mid A)$ is the usual conditional probability with respect to the set A, *with the convention that it vanishes if* $\mathbb{P}(A) = 0$. If $\mathbb{P}(\Omega_t^c) = 0$ then $\alpha_t^c = 0$, which is a rather natural convention. It would seem better to define the level as the essential supremum $\alpha_t'^c$ (in ω) over Ω_t^c of $\limsup_n \; \mathbb{P}(C_{t,n}^c \mid \mathcal{F})$; the two notions are closely related and $\alpha_t'^c \geq \alpha_t^c$, but we cannot exclude a strict inequality here, whereas we have no way (so far) to handle $\alpha_t'^c$. Note that α_t^c features some kind of "uniformity" over all subsets $A \subset \Omega_t^c$, in the spirit of the uniformity in $\theta \in \Theta_0$ in (10.2).

As for the asymptotic power function, we define it as

$$\beta_t^c \;=\; \liminf_n \; \mathbb{P}(C_{t,n}^c \mid \mathcal{F}) \tag{10.4}$$

and of course only the restriction of this "power function" (a random variable, indeed) to the alternative set Ω_t^j imports.

When on the opposite we take "there are jumps" as our null hypothesis, that is Ω_t^j, in a similar way we associate to the critical region $C_{t,n}^j$ the asymptotic level α_t^j and the power function β_t^j (simply exchange everywhere Ω_t^c and Ω_t^j).

10.3 The test statistics.

First we recall the processes (3.1), except that here we do not specify the component since X is 1-dimensional:

$$B(p, \Delta_n)_p \;=\; \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p. \tag{10.5}$$

The test statistics we will use to construct the critical regions, for both null hypotheses, are the following ones:

$$\widehat{S}(p, k, \Delta_n)_t \;=\; \frac{B(p, k\Delta_n)_t}{B(p, \Delta_n)_t}, \tag{10.6}$$

where $k \geq 2$ is an integer, and $p > 3$. Note that the numerator is obtained by considering only the increments of X between successive intervals of length $k\Delta_n$. Then we have (and the assumption (K) is unnecessarily strong for this):

Theorem 10.1 Assume (K). For all $t > 0$ we have the following convergence:

$$\widehat{S}(p,k,\Delta_n)_t \;\stackrel{\mathbb{P}}{\longrightarrow}\; \begin{cases} 1 & \text{on the set } \Omega_t^j \\ k^{p/2-1} & \text{on the set } \Omega_t^c. \end{cases} \tag{10.7}$$

Proof. By Theorem 5.1 the two variables $B(p,\Delta_n)_t$ and $B(p,k\Delta_n)_t$ both converge in probability to $\sum_{s\leq t}|\Delta X_s|^p$ (this is true as soon as $p > 2$, indeed), and the latter variable is strictly positive on the set Ω_t^j: hence the convergence on the set Ω_t^j is obvious.

When X has no jump, we can apply (6.5) to obtain that $\Delta_n^{1-p/2}B(p,\Delta_n)_t$, and of course $(k\Delta_n)^{1-p/2}B(p,k\Delta_n)_t$ as well, converge to $m_p\int_0^t c_s^{p/2}ds$, which by (H')-(a) is not 0. Then obviously we have the second limit in (10.7) when X is continuous.

This does not end the proof, however, except in the case $\Omega_t^c = \Omega$. It may happen that $0 < \mathbb{P}(\Omega_t^c) < 1$, so X is not (a.s.) continuous even on $[0,t]$, but some of its path are. However, suppose that we have proved the following:

$$X_s = X'_s \;\text{ for all } s \leq t, \text{ on the set } \Omega_t^c, \text{ where } X' \text{ satisfies (K) and is continuous.} \tag{10.8}$$

Then obviously $B(X,p,\Delta_n)_t = B(X',p,\Delta_n)_t$ and $B(X,p,k\Delta_n)_t = B(X',p,k\Delta_n)_t$ on the set Ω_t^c, and we get the result by applying (6.5) to X' instead of X.

The construction of X' involves the assumption (K)-(b). In fact we set

$$X'_t \;=\; X_0 + \int_0^t b'_s ds + \int_0^t \sigma_s dW_s \tag{10.9}$$

where $b'_t = b_t - b''_t$ and $b''_t = \Big(\int \kappa(\delta(t,z))\lambda(dz)\Big)1_{\{t<S\}}$. Then b'_t is adapted, with left-continuous and right limited paths, so X' satisfies (K), and it is continuous. Now suppose that we are in Ω_t^c. Then $t < S$, hence $\kappa'(\delta)\star\underline{\mu}_s = 0$ for all $s \leq t$. As for the stochastic integral $\kappa(\delta)\star(\underline{\mu}-\underline{\nu})_s$ for $s \leq t$, we observe that in fact $\kappa(\delta)\star\underline{\nu}_s$ is well-defined as an ordinary integral and equals $\int_0^s b''_u du$; hence $\kappa(\delta)\star\underline{\mu}_s$ is also an ordinary integral, and since $s < S$ it actually vanishes: therefore we deduce that $X_s = X'_s$ if $s \leq t$ and we are done. $\Box$

We now turn to the central limit theorem. We introduce two processes, with $q > 0$ and $q \geq 2$ respectively:

$$A(q)_t \;=\; \int_0^t c_u^{q/2}du, \qquad D(q)_t \;=\; \sum_{s\leq t}|\Delta X_s|^q(c_{s-}+c_s). \tag{10.10}$$

Recalling that $d = 1$ here, these two processes are respectively the right side of (6.5) and the process $D^{11}(f)$ of (6.11), when we take the function $f(x) = |x|^q$. For this function we also write $|x|^q\star\mu$ instead of $f\star\mu$. In addition to the absolute moments m_p used before, we also set

$$m_{2p}(k) \;=\; \mathbb{E}\Big(|\sqrt{k-1}\,U+V|^p\;|V|^p\Big), \tag{10.11}$$

where U and V are two independent $\mathcal{N}(0,1)$ variables. Finally we set

$$M(p,k) \;=\; \frac{1}{m_p^2}\,\Big(k^{p-2}(1+k)(m_{2p}-m_p^2)-2k^{p/2-1}(m_{2p}(k)-k^{p/2}m_p^2)\Big). \tag{10.12}$$

When $p=4$ we get $M(p,k)=16k(2k^2-k-1)/35$, and in particular $M(4,2)=\frac{32}{7}$.

Theorem 10.2 *Assume (K), and let $t>0$, $p>3$ and $k\leq 2$.*

(a) In restriction to the set Ω_t^j, the variables $\frac{1}{\sqrt{\Delta_n}}\,(\widehat{S}(p,k,\Delta_n)_t-1)$ converge stably in law to a variable $S(p,k)_t^j$ which, conditionally on $\mathcal{F}$, is centered with variance

$$\widetilde{\mathbb{E}}\Big((S(p,k)_t^j)^2\mid\mathcal{F}\Big) \;=\; \frac{(k-1)p^2}{2}\,\frac{D(2p-2)_t}{(|x|^p\star\mu)_t^2}. \tag{10.13}$$

Moreover if the processes σ and X have no common jumps, the variable $S(p,k)_t^j$ is $\mathcal{F}$-conditionally Gaussian.

(b) In restriction to the set Ω_t^c, the variables $\frac{1}{\sqrt{\Delta_n}}\,(\widehat{S}(p,k,\Delta_n)_t-2)$ converge stably in law to a variable $S(p,k)_t^c$ which, conditionally on $\mathcal{F}$, is centered Gaussian with variance

$$\widetilde{\mathbb{E}}\Big((S(p,k)_t^c)^2\mid\mathcal{F}\Big) \;=\; M(p,k)\,\frac{A(2p)_t}{(A(p)_t)^2}. \tag{10.14}$$

We have already encountered in and explained after Theorem 9.3 the notion of stable convergence in law in restriction to a subset of Ω. It is also worth noticing that the conditional variances (10.13) and (10.15), although of course random, are more or less behaving in time like $1/t$.

Proof. a) Write $U_n=\frac{1}{\sqrt{\Delta_n}}\,(B(p,\Delta_n)_t-|x|^p\star\mu_t)$ and $V_n=\frac{1}{\sqrt{\Delta_n}}\,(B(p,k\Delta_n)_t-|x|^p\star\mu_t)$. Then

$$\widehat{S}(p,k,\Delta_n)_t-1=\frac{B(p,k\Delta_n)_t}{B(p,\Delta_n)_t}-1=\sqrt{\Delta_n}\,\frac{V_n-U_n}{B(p,\Delta_n)_t}.$$

Since $p>3$, Corollary 8.3 yields that V_n-U_n converges stably in law to $Z'(f)_t-Z(f)_t$, and the result readily follows from (8.6), from the fact that $B(p,\Delta_n)_t\stackrel{\mathbb{P}}{\longrightarrow}|x|^p\star\mu_t$, and from the last claim in Lemma 8.1.

b) Exactly as for Theorem 10.1 it is enough to prove the result for the process X' of (10.9). This amounts to assume that the process X itself is continuous, so $\Omega_t^c=\Omega$. Write $U_n'=\frac{1}{\sqrt{\Delta_n}}\,(\Delta_n^{1-p/2}B(p,\Delta_n)-A(p)_t)$ and $V_n'=\frac{1}{\sqrt{\Delta_n}}\,(\Delta_n^{1-p/2}B(p,k\Delta_n)-k^{p/2-1}A(p)_t)$. Then

$$\widehat{S}(p,k,\Delta_n)_t-k^{p/2-1}=\frac{B(p,k\Delta_n)_t}{B(p,\Delta_n)_t}-k^{p/2-1}=\sqrt{\Delta_n}\;\frac{V_n'-k^{p/2-1}U_n'}{\Delta_n^{1-p/2}B(p,\Delta_n)_t}.$$

Now we consider the 2-dimensional function f whose components are $|x_1|^p+\cdots+|x_k|^p$ and $|x_1+\cdots+x_k|^p$. Recalling (7.6), the two components of $\frac{1}{\sqrt{\Delta_n}}\Big(V''(f,k,\Delta_n)_t-\frac{1}{k}\int_0^t\rho_{\sigma_u}^{\otimes k}(f)\Big)$ are respectively $U_n'+U_n''$ and V_n', where

$$U_n'' \;=\; \sqrt{\Delta_n}\sum_{k[t/k\Delta_n]<i\leq[t/\Delta_n]}|\Delta_i^nX|^p.$$

Obviously $U''_n \to 0$, hence Theorem 7.3 implies that the pair (U'_n, V'_n) converges stably in law to a vector which is $\mathcal{F}$-conditionally centered Gaussian, with $\mathcal{F}$-conditional covariance $MA(2p)_t$, where the entries of M are $M_{11} = 1 - m_p^2/m_{2p}$ and $M_{12} = M_{21} = (m_{2p}(k) - m_p^2)/m_{2p}$ and $M_{22} = k^{p-1}(1 - m_p^2/m_{2p})$. Therefore, using also the fact that $\Delta_n^{1-p/2} B(p, \Delta_n)_t \xrightarrow{\mathbb{P}} A(p)_t$, we readily deduce the result. □

Exactly as for estimating the volatility, see Theorem 9.3, this CLT is useless in practice and one has to standardize the test statistics so as to obtain a usable result. As usual, the standardization is done by dividing by the square-root of any consistent estimators for the conditional variances in (10.13) and (10.14). For the first one we can use again the fact that $B(p, \Delta_n)_t \xrightarrow{\mathbb{P}} |x|^p \star \mu_t$, plus the following version of (6.14), which by Theorem 6.5 converges to $D(q)_t$ if $q > 2$:

$$D(q, \varpi, \alpha, \Delta_n)_t = \frac{1}{k_n \Delta_n} \sum_{i=1+k_n}^{[t/\Delta_n]-k_n} |\Delta_i^n X|^q \sum_{j:\, j\neq i, |j-i|\leq k_n} |\Delta_j^n X|^2 1_{\{|\Delta_j^n X| \leq \alpha \Delta_n^{\varpi}\}}. \qquad (10.15)$$

where $\alpha > 0$ and $\varpi \in (0, \frac{1}{2})$, and k_n satisfies (6.12).

For the right side of (10.14) we can use estimators of $A(p)_t$, as provided in Theorem 6.3; for example, with ϖ and α as above, we can take

$$A(p, \varpi, \alpha, \Delta_n)_t \;=\; \Delta_n^{1-p/2} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p 1_{\{|\Delta_i^n X| \leq \alpha \Delta_n^{\varpi}\}}, \qquad (10.16)$$

which converges to $A(p)_t$ when X is continuous (and also when X has jumps, in restriction to Ω_t^c, as in the proof of Theorem 10.1). The variables $\Delta_n^{1-p/2} B(p, \Delta_n)_t$ also converge to $A(p)_t$ on Ω_t^c. Hence the next result follows from Theorem 10.2, with exactly the same proof than for Theorem 9.3:

Theorem 10.3 *Assume (K) and let $t > 0$, $p > 3$ and $k \geq 2$.*

(a) In restriction to the set Ω_t^j, the variables $\frac{1}{\sqrt{\Gamma^j(t,n)}}\,(\widehat{S}(p, k, \Delta_n)_t - 1)$, where

$$\Gamma^j(t,n) \;=\; \frac{\Delta_n (k-1) p^2 \; D(2p-2, \varpi, \alpha, \Delta_n)_t}{(B(p, \Delta_n)_t)^2} \qquad (10.17)$$

converge stably in law to a variable which, conditionally on $\mathcal{F}$, is centered with variance 1, and which additionally is $\mathcal{F}$-conditionally normal if the processes σ and X have no common jumps.

(b) In restriction to the set Ω_t^c, the variables $\frac{1}{\sqrt{\Gamma^c(t,n)}}\,(\widehat{S}(p, k, \Delta_n)_t - k^{p/2-1})$, where either

$$\Gamma^c(t,n) \;=\; \frac{\Delta_n M(p,k) \; A(2p, \varpi, \alpha, \Delta_n)_t}{(A(p, \varpi, \alpha, \Delta_n)_t)^2}, \qquad (10.18)$$

$$\Gamma^c(t,n) \;=\; \frac{M(p,k) \; B(2p, \Delta_n)_t}{(B(p, \Delta_n)_t)^2}, \qquad (10.19)$$

converge stably in law to a variable which, conditionally on $\mathcal{F}$, is $\mathcal{N}(0,1)$.

We will see later that, although both choice of $\Gamma^c(t,n)$ are asymptotically equivalent for determining the level of our tests, it is no longer the case for the power function: the second choice (10.19) should *never* prevail.

10.4 Null hypothesis = no jump.

We now use the preceding results to construct actual tests, either for the null hypothesis that there are no jumps, or for the null hypothesis that jumps are present. We start with the first one here. The null hypothesis is then "Ω_t^c", and we are going to construct a critical (rejection) region $C_{t,n}^c$ for it. In view of Theorem 10.1 it is natural to take a region of the form

$$C_{t,n}^c = \{\widehat{S}(p,k,\Delta_n)_t < \gamma_{t,n}^c\} \tag{10.20}$$

for some sequence $\gamma_{t,n}^c > 0$, possibly even a random sequence. What we want, though, is to achieve an asymptotic level α prescribed in advance. For this we need to introduce the α-quantile of $N(0,1)$, that is $\mathbb{P}(U > z_\alpha) = \alpha$ where U is $N(0,1)$.

Theorem 10.4 *Assume (K), and let $t > 0$, $p > 3$ and $k \geq 2$. For any prescribed level $\alpha \in (0,1)$ we define the critical region $C_{t,n}^j$ by (10.20), with*

$$\gamma_{t,n}^c \;=\; k^{p/2-1} - z_\alpha \sqrt{\Gamma_{t,n}^c}, \tag{10.21}$$

where $\Gamma^c(t,n)$ is given either by (10.18) or by (10.19).

(a) The asymptotic level α_t^c for testing the null hypothesis of "no jump" is not bigger than α and equal to α when $\mathbb{P}(\Omega_t^c) > 0$; we even have $\mathbb{P}(C_{t,n}^c \mid A) \to \alpha$ for all $A \subset \Omega_t^c$ with $\mathbb{P}(A) > 0$.

(b) The asymptotic power function β_t^c is a.s. equal to 1 on the complement Ω_t^j if we use (10.18) for $\Gamma^c(t,n)$, with $\varpi \in (\frac{1}{2} - \frac{1}{p}, \frac{1}{2})$, but this fails in general if we use (10.19).

Proof. For (a) it is enough to prove that if $A \in \Omega_t^c$ has $\mathbb{P}(A) > 0$, then $\mathbb{P}(C_{t,n}^c \mid A) \to \alpha$. Let $U_n = \frac{1}{\sqrt{\Gamma^c(t,n)}} (\widehat{S}(p,k,\Delta_n)_t - k^{p/2-1})$. We know that this variable converges stably in law, as $n \to \infty$, and in restriction to Ω_t^c, to an $\mathcal{N}(0,1)$ variable U independent of $\mathcal{F}$. Therefore for A as above we have

$$\mathbb{P}(C_{t,n}^c \cap A) \;=\; \mathbb{P}(\{U_n \leq -z_\alpha\} \cap A) \;\to\; \mathbb{P}(A)\mathbb{P}(U \leq -z_\alpha) \;=\; \alpha\,\mathbb{P}(A),$$

and the result follows.

For (b) we can assume $\mathbb{P}(\Omega_t^j) > 0$, otherwise there is nothing to prove. Theorem 10.1 implies that $\widehat{S}(p,k,\Delta_n)_t \stackrel{\mathbb{P}}{\longrightarrow} 1$ on Ω_t^j. If we use the version (10.19) for $\Gamma^c(t,n)$, then Theorem 5.1 implies that $\Gamma^c(t,n)$ converges in probability to a positive finite variable, on Ω_t^j again. Hence on this set the variable U_n converges in probability to a limiting variable U (equal in fact to $(1-k^{p/2})|x|^p \star \mu_t / \sqrt{M(p,k)|x|^{2p} \star \mu_t}$). In general this variable is not a.s. smaller than $-z_\alpha$ on Ω_t^j, and thus the power function is not equal to 1 on this set.

On the opposite, suppose that we have chosen the version (10.18), with $\varpi \in (\frac{1}{2} - \frac{1}{p}, \frac{1}{2})$. Suppose also that

$$\frac{\Delta_n A(2p, \varpi, \alpha, \Delta_n)t}{(A(p, \varpi, \alpha, \Delta_n)_t)^2} \stackrel{\mathbb{P}}{\longrightarrow} 0. \tag{10.22}$$

This means that $\Gamma^c(t,n) \stackrel{\mathbb{P}}{\longrightarrow} 0$. Since $1 - k^{p/2-1} < 0$ we deduce that $U_n \stackrel{\mathbb{P}}{\longrightarrow} -\infty$ on the set Ω_t^k. Then

$$\mathbb{P}(C_{t,n}^c \cap \Omega_t^j) \;=\; \mathbb{P}(\{U_n \leq -z_\alpha\} \cap \Omega_t^j) \;\to\; \mathbb{P}(\Omega_t^j).$$

This trivially implies $\mathbb{P}(C_{t,n}^c \mid \mathcal{F}) \stackrel{\mathbb{P}}{\longrightarrow} 1$ on the set Ω_t^j.

It remains to prove (10.22), and for this it is no restriction to assume (SH). The reader will observe that when X is continuous this trivially follows from Theorem 6.3, but unfortunately we need this property on Ω_t^j. With the notation of (6.22), one easily check that for all $B > 0$:

$$\left| \Delta_n^{1-p/2} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p 1_{\{|\Delta_i^n X| \leq \sqrt{B\Delta_n}\}} - \Delta_n \sum_{i=1}^{[t/\Delta_n]} |\beta_i^n|^p \right| \;\leq\; K Z_n(B), \tag{10.23}$$

where

$$Z_n(B) = \Delta_n \sum_{i=1}^{[t/\Delta_n]} \Big(|\beta_i^n|^p 1_{\{|\beta_i^n| > \sqrt{B}/2\}} B^{p/2-1}(|\chi_i^n|^2 \wedge B) + |\beta_i^n|^{p-1}(|\chi_i^n| \wedge \sqrt{B}) \Big).$$

(6.23) and Bienaymé-Tchebycheff, plus (6.25) and Cauchy-Schwarz give us

$$\limsup_n \; \mathbb{E}(Z_n(B)) \;\leq\; \frac{Kt}{B} \tag{10.24}$$

On the other hand we know that $\Delta_n^{1-p/2} \sum_{i=1}^{[t/\Delta_n]} |\beta_i^n|^p \stackrel{\mathbb{P}}{\longrightarrow} A(p)_t$. Combining this with the above estimates and (10.23), we obtain for all $\eta, B > 0$:

$$\mathbb{P}\left(\Delta_n^{1-p/2} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p 1_{\{|\Delta_i^n X| \leq \sqrt{B\Delta_n}\}} < A(p)_t - Z_n(B) - \eta \right) \;\to\; 0.$$

Now, for any $B \geq 1$ we have $\alpha \Delta_n^{\varpi} > \sqrt{B\Delta_n}$ for all n large enough because $\varpi < 1/2$. Therefore we a fortiori have

$$\mathbb{P}\left(A(p, \varpi, \alpha, \Delta_n)_t < A(p)_t - Z_n(B) - \eta\right) \;\to\; 0.$$

Now (10.24) imply that $\lim_{B\to\infty} \; \limsup_n \; \mathbb{P}(Z_n(B) > \eta) = 0$, hence

$$\mathbb{P}\left(A(p, \varpi, \alpha, \Delta_n)_t < A(p)_t - 2\eta\right) \;\to\; 0.$$

Since $A(p)_t > 0$ a.s., we finally deduce

$$\mathbb{P}\left(A(p, \varpi, \alpha, \Delta_n)_t < \frac{A(p)_t}{2} \right) \;\to\; 0. \tag{10.25}$$

At this stage, the proof of (10.22) is straightforward: since $|\Delta_i^n X|^{2p} \leq \alpha^p \Delta_n^{p\varpi} |\Delta_i^n X|^p$ when $|\Delta_i^n X| \leq \alpha \Delta_n^{\varpi}$, one deduces from (10.16) that

$$\frac{\Delta_n\, A(2p,\varpi,\alpha,\Delta_n)_t}{A(p,\varpi,\alpha,\Delta_n)_t^2} \;\leq\; \frac{K\Delta_n^{p\varpi+1-p/2}}{A(p,\varpi,\alpha,\Delta_n)_t}.$$

Since $p\varpi + 1 - p/2 > 0$, the result readily follows from (10.25). □

10.5 Null hypothesis = there are jumps.

In a second case, we set the null hypothesis to be that there are jumps, that is "Ω_t^j". Then we take a critical region of the form

$$C_{t,n}^j = \{\widehat{S}(p,k,\Delta_n)_t > \gamma_{t,n}^j\}. \tag{10.26}$$

for some sequence $\gamma_{t,n}^j > 0$. As in (10.3) and (10.4), the asymptotic level and power functions are

$$\alpha_t^j = \sup\Big(\limsup_n\ \mathbb{P}(C_{t,n}^j \mid A):\ A \in \mathcal{F}, A \subset \Omega_t^j\Big), \qquad \beta_t^d = \liminf_n \mathbb{P}(C_{t,n}^j \mid \mathcal{F}).$$

Theorem 10.5 *Assume (K), and let $t > 0$, $p > 3$ and $k \geq 2$. Define $\Gamma^j(t,n)$ by (10.17), and let $\alpha \in (0,1)$ be a prescribed level.*

(i) With the critical region $C_{t,n}^j$ given by (10.26), with

$$\gamma_{t,n}^j \;=\; 1 + \frac{1}{\sqrt{\alpha}}\,\sqrt{\Gamma^j(t,n)}, \tag{10.27}$$

the asymptotic level α_t^j for testing the null hypothesis of "jumps" is not bigger than α.

(ii) With the critical region $C_{t,n}^j$ given by (10.26), with

$$\gamma_{t,n}^j \;=\; 1 + z_\alpha\,\sqrt{\Gamma^j(t,n)}, \tag{10.28}$$

and if further the two processes X and σ do not jump at the same times, the asymptotic level α_t^j for testing the null hypothesis of "jumps" is not bigger than α, and equals to α when $\mathbb{P}(\Omega_t^j) > 0$; we even have $\mathbb{P}(C_{t,n}^j \mid A) \to \alpha$ for all $A \subset \Omega_t^j$ with $\mathbb{P}(A) > 0$.

(iii) In both cases the asymptotic power function β_t^j is a.s. equal to 1 on the complement Ω_t^c of Ω_t^j.

Since $z_\alpha < 1/\sqrt{\alpha}$ the critical region is larger with the version (10.28) than with the version (10.27). Hence, even though asymptotically the two power functions are equal, at any stage n the power is bigger with (10.28) than with (10.27), so one should use (10.28) whenever possible (however, when there are jumps, it is usually the case that the volatility jumps together with X).

Proof. We know that the variables $U_n = \frac{1}{\sqrt{\Gamma^c(t,n)}}\ (\widehat{S}(p,k,\Delta_n)_t - 1)$ converges stably in law, as $n \to \infty$, and in restriction to Ω_t^j, to a variable which conditionally on $\mathcal{F}$ is centered with variance 1, and is further $\mathcal{N}(0,1)$ if X and σ do not jump at the same times. Then $\mathbb{P}(U > 1/\sqrt{\alpha}) \leq \alpha$, and also $\mathbb{P}(U > z_\alpha) = \alpha$ in the latter case, the two statements (i) and (ii) follow exactly as in Theorem 10.4.

For (iii) we can assume $\mathbb{P}(\Omega_t^c) > 0$, otherwise there is nothing to prove. Then in restriction to Ω_t^c the statistics $\widehat{S}(p,k,\Delta_n)_t$ converge in probability to $k^{p/2-1} > 1$. Moreover, on this set again, both $D(2p-2,\varpi,\alpha,\Delta_n)_t$ and $B(p,\Delta_n)_t$ are the same as if they were computed on the basis of the continuous process X' of (10.9). Therefore, by virtue of Theorems 6.2 and 6.5 we have that $\Delta_n^{1-p/2} B(p,\Delta_n)_t \stackrel{\mathbb{P}}{\longrightarrow} A(p)_t$ and $\Delta_n^{2-p} D(2p-2,\varpi,\alpha,\Delta_n)_t \stackrel{\mathbb{P}}{\longrightarrow} \frac{2m_{2p-2}}{m_{2p}} A(2p)_t$ on Ω_t^c. Since by (H') we have $A(p)_t > 0$ it follows that $\Gamma^j(t,n) \stackrel{\mathbb{P}}{\longrightarrow} 0$ on Ω_t^j. Therefore we have $U_n \stackrel{\mathbb{P}}{\longrightarrow} +\infty$ on Ω_t^c, and as in Theorem 10.4 we conclude that $\mathbb{P}(C_{n,t}^j \cap \Omega_t^c) \to \mathbb{P}(\Omega_t^c)$, hence $\mathbb{P}(C_{t,n}^c \mid \mathcal{F}) \stackrel{\mathbb{P}}{\longrightarrow} 1$ on the set Ω_t^c. □

11 Testing for common jumps

This section is again about jumps. We suppose here that our underlying process is multidimensional, and that it has jumps, and we want to check whether any two components have jumps occurring at the same time.

11.1 Preliminary remarks.

Clearly the problem at hand is 2-dimensional, since in the multidimensional situation one can perform the tests below for any pair of components. So below we assume that $X = (X^1, X^2)$ is 2-dimensional. Exactly as in the previous section, we need a slightly stronger assumption than (H):

Assumption (K'): We have (H); furthermore with the notation $\tau = \inf(t : \Delta X_t^1 \Delta X_t^2 \neq 0)$ (the infimum of all common jump times) and $\Gamma = \{(\omega,t,x) : \delta^1(\omega,t,x)\delta^2(\omega,t,x) \neq 0\}$, we have

(a) $C_t \neq 0$ when $t > 0$

(b) $t \mapsto \int \kappa(\delta(\omega,t,z)) 1_\Gamma(\omega,t,z)\lambda(dz)$ is left-continuous with right limits on the interval $(0,\tau(\omega)]$. □

(a) above is again a non-degeneracy assumption for X^c, similar in the 2-dimensional case to (a) of (K). As for (b) here, we can state the same remarks as for (b) of (K): it is "empty" on the set $\{\tau = 0\}$, that is where X^1 and X^2 have infinitely many common jumps near the origin. It is implied by (H) when $\int (\gamma(z) \wedge 1)\lambda(dz) < \infty$. Moreover in all generality, and outside a $\mathbb{P}$-null set, $z \mapsto \kappa(\delta(\omega,t,z))$ is λ-integrable if $t < \tau(\omega)$. So again (b) is a very mild additional smoothness assumption, of the same nature as (b) of (H).

Next, and again like in the previous section, what we can really test on the basis of discrete observations of X over a finite time interval $[0,T]$ is whether the two paths $t \mapsto X_t^1(\omega)$ and $t \mapsto X_t^2(\omega)$ have common jump times or not. That is, we can (hopefully) decide in which one of the following two disjoint subsets of Ω we are:

$$\left.\begin{array}{lcl}\Omega_T^{cj} & = & \{\omega: \ s\mapsto X_s^1(\omega) \text{ and } s\mapsto X_s^2(\omega) \text{ have common jumps on } [0,T]\} \\ \Omega_T^{dj} & = & \{\omega: \text{ both } s\mapsto X_s^1(\omega) \text{ and } s\mapsto X_s^2(\omega) \text{ have jumps, but they have} \\ & & \quad \text{no common jump, on } [0,T]\}.\end{array}\right\} \tag{11.1}$$

The union of these two sets is not Ω, but their global complement is

$$\Omega_T^{cc} \ = \ \{\omega: \text{ both } s\mapsto X_s^1(\omega) \text{ and } s\mapsto X_s^2(\omega) \text{ are continuous on } [0,T]\}. \tag{11.2}$$

All three sets above may have a positive probability. However, we can first perform the tests developed in the previous section, separately on both components, to decide whether both of them jump. Then in this case only it makes sense to test for joint jumps. That is, we suppose that this preliminary testing has been done and that we have decided that we are *not* in Ω_T^{cc}.

At this point we again have two possible null hypotheses, namely "common jumps" (we are in Ω_T^{cj}) and "disjoint jumps" (we are in Ω_T^{dj}). Exactly as in the previous section we construct at stage n a critical region $C_{T,n}^{cj}$ for the null Ω_T^{cj}, and a critical region $C_{T,n}^{dj}$ for the null Ω_T^{dj}. In the first case the asymptotic level and power function are respectively

$$\alpha_T^{cj} = \sup\Big(\limsup_n \ \mathbb{P}(C_{T,n}^{cj} \mid A): \ A\in\mathcal{F}, A\subset\Omega_T^{cj}\Big), \quad \beta_T^{cj} = \liminf_n \ \mathbb{P}(C_{T,n}^{cj} \mid \mathcal{F}). \tag{11.3}$$

In the second case, they are

$$\alpha_T^{dj} = \sup\Big(\limsup_n \ \mathbb{P}(C_{T,n}^{dj} \mid A): \ A\in\mathcal{F}, A\subset\Omega_T^{dj}\Big), \quad \beta_T^{dj} = \liminf_n \ \mathbb{P}(C_{T,n}^{dj} \mid \mathcal{F}). \tag{11.4}$$

11.2 The test statistics.

Three functions will be used in the construction of our test statistics (here $x=(x^1,x^2)\in\mathbb{R}^2$):

$$f(x) \ = \ (x^1x^2)^2, \qquad g_1(x) \ = \ (x^1)^4, \qquad g_2(x) \ = \ (x^2)^4. \tag{11.5}$$

Then, with $k\geq 2$ being an integer fixed throughout, we put

$$\widehat{T}^{cj}(k,\Delta_n)_t \ = \ \frac{V(f,k\Delta_n)_t}{V(f,\Delta_n)_t}, \qquad \widehat{T}^{dj}(\Delta_n)_t \ = \ \frac{V(f,\Delta_n)_t}{\sqrt{V(g_1,\Delta_n)_t\,V(g_2,\Delta_n)_t}}. \tag{11.6}$$

These statistics will we used to construct respectively, the two critical regions $C_{t,n}^{cj}$ and $C_{t,n}^{dj}$. Unlike for simply testing jumps, we have to resort to two different statistics to deal with our two cases.

We have now to determine the asymptotic behavior of these statistics, deriving an LLN and a CLT for each one. To prepare for this we need to introduce a number of processes to come in in the limiting variables. First we set

$$F_t \ = \ \int_0^t (c_s^{11}c_s^{22} + 2(c_s^{12})^2)\,ds. \tag{11.7}$$

Second, on the extended space described in Subsection 8.1 and with the notation S_p, R_p and R'_p of this subsection (recall (8.2), here R_p and R'_p are 2-dimensional), we set

$$\left.\begin{array}{lcl} D_t &=& \sum_{p:S_p\le t}\Big((\Delta X^1_{S_p}R^2_p))^2+(\Delta X^2_{S_p}R^1_p)^2\Big) \\ D'_t &=& \sum_{p:S_p\le t}\Big((\Delta X^1_{S_p}R'^2_p))^2+(\Delta X^2_{S_p}R'^1_p)^2\Big). \end{array}\right\} \tag{11.8}$$

If we are on the set Ω^{dj}_T it turns out (via an elementary calculation) that in fact $D_t = \overline{Z}(f)_t/2$ and $D'_t = \overline{Z}'(f)_t/2$ for all $t \le T$.

Theorem 11.1 *Assume (K').*

(a) We have

$$\widehat{T}^{cj}(k,\Delta_n)_t \xrightarrow{\mathbb{P}} 1 \qquad \text{on the set } \Omega^{cj}_t, \tag{11.9}$$

and $\widehat{T}^{cj}(k,\Delta_n)_t$ converges stably in law, in restriction to the set Ω^{dj}_t, to

$$T^{cj}(k) = \frac{D'_t + kF_t}{D_t + F_t} \tag{11.10}$$

which is a.s. different from 1.

(b) We have

$$\widehat{T}^{dj}(\Delta_n)_t \xrightarrow{\mathbb{P}} \begin{cases} f\star\mu_t/\sqrt{(g_1\star\mu_t)(g_2\star\mu_t)} > 0 & \text{on the set } \Omega^{cj}_t \\ 0 & \text{on the set } \Omega^{dj}_t. \end{cases} \tag{11.11}$$

The second part of (a) is a kind of LLN because it concerns the behavior of $\widehat{T}^{cj}(k,\Delta_n)_t$ without centering or normalization, but it is also a kind of CLT.

Proof. On both sets Ω^{cj}_t and Ω^{dj}_t both components of X jumps before t, so $g_1\star\mu_t > 0$ and $g_2\star\mu_t > 0$, and also $f\star\mu_t > 0$ on Ω^{cj}_t. Then all claims except the second one in (a) are trivial consequences of Theorem 5.1.

Let us now turn to the behavior of $\widehat{T}^{cj}(k,\Delta_n)_t$ on Ω^{dj}_t. If we make the additional assumption that X^1 and X^2 *never* jump at the same time, then the stable convergence in law towards $T^{cj}(k)$, as defined by (11.10), is a trivial consequence of Theorem 8.4 and of the remark which follows (11.8). Moreover, the $\mathcal{F}$-conditional law of the pair of variable (D_t, D'_t), in restriction to Ω^{dj}_t, clearly admits a density, hence $\mathbb{P}(\Omega^{dj}_t \cap \{T^{cj} = 1\}) = 0$ and we have the last claim of (a).

Now, exactly as in Theorem 10.1, this is not quite enough for proving our claim, since it may happen that both Ω^{dj}_t and Ω^{cj}_t have positive probability. However, suppose that

$$\left.\begin{array}{l} X_s = X'_s \text{ for all } s\le t, \text{ on the set } \Omega^{dj}_t, \text{ where } X' \text{ satisfies (K')}, \\ \text{and the two components } X'^1 \text{ and } X'^2 \text{ never jump at the same times.} \end{array}\right\} \tag{11.12}$$

Then the above argument applied for X' instead of X yields the result.

The construction of X' involves (K')-(b). We set $b'_t = b_t - b''_t$, where the process $b''_t = \Big(\int \kappa(\delta(t,z))\, 1_\Gamma(t,z)\lambda(dz)\Big) 1_{\{t<\tau\}}$ is well-defined and left-continuous with right limits everywhere. Set also $\delta' = \delta 1_{\Gamma^c}$. Then the process

$$X'_t \;=\; X_0 + \int_0^t b''_s de + \int_0^t \sigma_s dW_s + \kappa(\delta') \star (\underline{\mu} - \underline{\nu})_t - \kappa'(\delta') \star \underline{\mu}_t$$

satisfies all requirements in (11.12) (we should more careful here; it satisfies (H"), except for one fact, namely we do not know whether $t \mapsto \delta'(\omega, t, z)$ is left-continuous with right limits; however, this particular property plays no role in the proof of Theorem 8.4, so the proof is nevertheless complete.) $\square$

Now we turn to the associated CLTs. Here again we need to complement the notation. Set

$$\overline{D}_t \;=\; \frac{1}{2} \sum_{s \le t} \Big((\Delta X^1_s)^2 (c^{22}_{s-} + c^{22}_s) + (\Delta X^2_s)^2 (c^{11}_{s-} + c^{11}_s) \Big), \tag{11.13}$$

$$\overline{D}'_t = 2 \sum_{s \le t} (\Delta X^1_s \Delta X^2_s)^2 \Big((\Delta X^2_s)^2 (c^{11}_{s-} + c^{11}_s) + (\Delta X^1_s)^2 (c^{22}_{s-} + c^{22}_s) + 2 \Delta X^1_s \Delta X^2_s (c^{12}_{s-} + c^{12}_s) \Big), \tag{11.14}$$

In other words, with the notation (8.3) and (8.4), we have $\overline{D}_s = \frac{1}{2}\, \overline{C}(f)_s$ for all $s \le t$ on the set Ω^{dj}_t, and $\overline{D}' = \frac{1}{2}\, C(f,f)$ everywhere.

Theorem 11.2 *Assume (K').*

(a) In restriction to the set Ω^{cj}_T the sequence $\frac{1}{\sqrt{\Delta_n}} (\widehat{T}^{cj}(k, \Delta_n)_t - 1)$ converges stably in law to a variable $T'^{cj}(k)$ which, conditionally on $\mathcal{F}$, is centered with variance

$$\widetilde{\mathbb{E}}\Big((T'^{cj}(k))^2 \mid \mathcal{F} \Big) \;=\; (k-1) \frac{\overline{D}'_t}{(f \star \mu_t)^2}, \tag{11.15}$$

and is even Gaussian conditionally on $\mathcal{F}$ if the processes X and σ have no common jumps.

(b) In restriction to the set Ω^{dj}_T the sequences $\frac{1}{\Delta_n} \widehat{T}^{dj}(D_n)$ converges stably in law to the positive variable $T^{dj} = (D_t + F_t)/\sqrt{(g_1 \star \mu_t)(g_2 \star \mu_t)}$ which, conditionally on $\mathcal{F}$, satisfies

$$\widetilde{\mathbb{E}}(T^{dj} \mid \mathcal{F}) \;=\; \frac{\overline{D}_t + F_t}{\sqrt{(g_1 \star \mu_t)(g_2 \star \mu_t)}}. \tag{11.16}$$

Proof. a) This is the very same proof as for (a) of Theorem 10.2: we write $U_n = \frac{1}{\sqrt{\Delta_n}}\, (V(f, \Delta_n)_t - f \star \mu_t)$ and $V_n = \frac{1}{\sqrt{\Delta_n}}\, (V(f, k\Delta_n)_t - f \star \mu_t)$ and observe that

$$\widehat{T}^{cj}(k, \Delta_n)_t - 1 \;=\; \sqrt{\Delta_n}\, \frac{V_n - U_n}{V(f, \Delta_n)_t}.$$

Then we conclude using Corollary 8.3, plus (11.14) and the remark that follows, in exactly the same way.

b) Exactly as in the previous theorem, we can replace X by a process X' satisfying (11.12), or equivalently we can assume that the two X^1 and X^2 never jump at the same times. Then the result immediately derives from Theorem 8.4. □

Finally we need to standardize our statistics, and thus to find consistent estimators for the conditional variance in (11.15), and conditional first moment in (11.16). For the variables $f \star \mu_t$, $g_1 \star \mu_t$ and $\mathcal{G}_2 \star \mu_t$ we can use $V(f,\Delta_n)_t$, $V(g_1,\Delta_n)_t$ and $V(g_2,\Delta_n)_t$ respectively. For F_t we can use the truncated powers (see Theorem 6.3; we have to be careful here, because X is discontinuous, whereas f is a polynomial of degree 4; so we choose the version given by (6.8)-(6.9)): we choose $\varpi \in (0,\frac{1}{2})$ and $\alpha > 0$, and we set

$$\begin{aligned} A(\varpi,\alpha,\Delta_n)_t &= \frac{1}{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\Big(|\Delta_i^n X^1|^2|\Delta_{i+1}^n X^2|^2 \\ &\quad +2\Delta_i^n X^1\Delta_i^n X^2\Delta_{i+1}^n X^1\Delta_{i+1}^n X^2\Big)1_{\{\|\Delta_i^n X\|\le\alpha\Delta_n^\varpi\},\ \|\Delta_{i+1}^n X\|\le\alpha\Delta_n^\varpi\}}. \qquad (11.17)\end{aligned}$$

Finally, by virtue of Theorem 6.5, we can estimate $\overline{D}_t$ and $\overline{D}'_t$ by the following variables, where in addition to ϖ and α we have chosen a sequence k_n of integers satisfying (6.12):

$$\begin{aligned} \overline{D}(\varpi,\alpha,\Delta_n)_t &= \frac{1}{2k_n\Delta_n}\sum_{i=1+k_n}^{[t/\Delta_n]-k_n} 1_{\{\|\Delta_i^n X\|>\alpha\Delta_n^\varpi\}} \\ &\quad \sum_{j\in I_n(i)}\Big((\Delta_i^n X^1)^2(\Delta_j^n X^2)^2+(\Delta_i^n X^2)^2(\Delta_j^n X^1)^2\Big)1_{\{\|\Delta_j^n X\|\le\alpha\Delta_n^\varpi\}} \qquad (11.18)\end{aligned}$$

$$\begin{aligned} \overline{D}'(\varpi,\alpha,\Delta_n)_t &= \frac{2}{k_n\Delta_n}\sum_{i=1+k_n}^{[t/\Delta_n]-k_n}\sum_{j\in I_n(i)}(\Delta_i^n X^1)^2(\Delta_i^n X^2)^2 \\ &\quad \Big(\Delta_i^n X^1\Delta_j^n X^2+\Delta_i^n X^2\Delta_j^n X^1\Big)^2 1_{\{\|\Delta_j^n X\|\le\alpha\Delta_n^\varpi\}}. \qquad (11.19)\end{aligned}$$

Then we have the following trivial consequence of Theorem 11.2:

Theorem 11.3 *Assume (K').*

(a) In restriction to the set Ω_t^{cj}*, the variables* $\frac{1}{\sqrt{\Gamma^{cj}(t,n)}}\,(\widehat{T}^{cj}(k,\Delta_n)_t-1)$ *, where*

$$\Gamma^{cj}(n,t) = \frac{\Delta_n(k-1)\overline{D}'(\varpi,\alpha,\Delta_n)_t}{(V(f,\Delta_n)_t)^2}, \qquad (11.20)$$

converge stably in law to a variable which, conditionally on $\mathcal{F}$*, is centered with variance 1, and which additionally is* $\mathcal{F}$*-conditionally Gaussian if the processes* X *and* σ *have no common jumps.*

(b) In restriction to the set Ω_t^{dj}*, the variables* $\frac{1}{\Gamma^{dj}(t,n)}\,\widehat{T}^{dj}(\Delta_n)_t$ *, where*

$$\Gamma^{dj}(t,n) = \frac{\Delta_n(\overline{D}(\varpi,\alpha,\Delta_n)_t+A(\varpi,\alpha,\Delta_n)_t)}{\sqrt{V(g_1,\Delta_n)_t\,V(g_2,\Delta_n)_t}}, \qquad (11.21)$$

converge stably in law, in restriction to the set Ω_T^{dj}, to a positive variable which, conditionally on $\mathcal{F}$, has expectation 1.

11.3 Null hypothesis = common jumps.

Now we are in a position to construct the critical regions we are looking for. We start with the null hypothesis being "there are common jumps", that is we are in Ω_t^{cj}. In view of Theorem 11.1 it is natural to take a critical region of the form

$$C_{t,n}^{cj} \;=\; \{|\widehat{T}^{cj}(k,\Delta_n)-1| \geq \gamma_{t,n}^{cj}\}. \tag{11.22}$$

For $\alpha \in (0,1)$ we denote by z'_α the symmetric α-quantile of an $\mathcal{N}(0,1)$ variable U, that is $\mathbb{P}(|U| \geq z_\alpha) = \alpha$.

Theorem 11.4 *Assume (K'), and let $t>0$ and $k \geq 2$. Define $\Gamma^{cj}(t,n)$ by (11.20), and let $\alpha \in (0,1)$ be a prescribed level.*

(i) With the critical region $C_{t,n}^{cj}$ given by (11.22), with

$$\gamma_{t,n}^{cj} \;=\; 1+\frac{1}{\sqrt{\alpha}}\,\sqrt{\Gamma^{cj}(t,n)}, \tag{11.23}$$

the asymptotic level α_t^{cj} for testing the null hypothesis of "common jumps" is not bigger than α.

(ii) With the critical region $C_{t,n}^{cj}$ given by (11.22), with

$$\gamma_{t,n}^{cj} \;=\; 1+z'_\alpha\,\sqrt{\Gamma^{cj}(t,n)}, \tag{11.24}$$

and if further the two processes X and σ do not jump at the same times, the asymptotic level α_t^j for testing the null hypothesis of "common jumps" is not bigger than α, and equals to α when $\mathbb{P}(\Omega_t^{cj})>0$; we even have $\mathbb{P}(C_{t,n}^{cj} \mid A) \to \alpha$ for all $A \subset \Omega_t^{cj}$ with $\mathbb{P}(A)>0$.

(iii) In both cases the asymptotic power function β_t^{cj} is a.s. equal to 1 on the set Ω_t^{dj}.

Again $z'_\alpha < 1/\sqrt{\alpha}$, so whenever possible one should choose the critical region defined by (11.24).

Proof. In view of the previous theorem, (i) and (ii) are proved exactly as in Theorem 10.5 for example. For (iii), we observe first that, in view of Theorem 11.1(-a), the variable $\widehat{T}^{cj}(k,\Delta_n)_t$ converges stably in law to $T^j(k)-1$, which a.s. noon vanishing. On the other hand we have $\overline{D}'(\varpi,\alpha,\Delta_n)_t \stackrel{\mathbb{P}}{\longrightarrow} \overline{D}_t$ everywhere and $V(f,\Delta_n)_t \stackrel{\mathbb{P}}{\longrightarrow} f\star\mu_t > 0$ on Ω_t^{dj}, hence $\Gamma^{cj}(t,n) \stackrel{\mathbb{P}}{\longrightarrow} 0$ on Ω_t^{dj}. That is, $\gamma_{t,n}^{dj} \stackrel{\mathbb{P}}{\longrightarrow} 1$ on this set, and this implies the result. □

11.4 Null hypothesis = no common jumps.

In a second case, we set the null hypothesis to be that "no common jumps", that is we are in Ω_t^{dj}. We take a critical region of the form

$$C_{t,n}^{dj} = \{\widehat{T}^{dj}(\delta_n)_t \geq \gamma_{t,n}^{dj}\}. \tag{11.25}$$

Theorem 11.5 *Assume (K'), and let $t > 0$. Define $\Gamma^{dj}(t,n)$ by (11.21), and let $\alpha \in (0,1)$ be a prescribed level.*

(a) With the critical region $C^{dj}_{t,n}$ given by (11.25), with

$$\gamma^{dj}_{t,n} \;=\; \frac{\Gamma^{dj}(t,n)}{\alpha}, \tag{11.26}$$

the asymptotic level α^{dj}_t for testing the null hypothesis of "common jumps" is not bigger than α.

(b) The asymptotic power function β^{dj}_t is a.s. equal to 1 on the set Ω^{cj}_t.

Proof. The variables $U_n = \widehat{T}^{dj}(\Delta_n)_t/\Gamma^{dj}(t,n)$ converge stably in law to a limit $U > 0$ having $\widetilde{\mathbb{E}}(U \mid \mathcal{F}) = 1$, in restriction to Ω^{dj}_t. Hence if $A \in \mathcal{F}$ is included into Ω^{dj}_t we have

$$\alpha \mathbb{P}(A) \;\geq\; \mathbb{P}(A \cap \{U \geq \frac{1}{\alpha}\}) \;\leq\; \limsup_n \; \mathbb{P}(A \cap \{U_n \geq \frac{1}{\alpha}\}) \;=\; \mathbb{P}(C^{dj}_{t,n} \cap A).$$

and (a) readily follows.

For (b) one observes that $\Gamma^{dj}(t,n) \stackrel{\mathbb{P}}{\longrightarrow} 0$ o Ω^{cj}_t, whereas on this set $\widehat{T}^{dj}(\Delta_n)_t$ converge to a positive variable by Theorem 11.1, hence $U_n \stackrel{\mathbb{P}}{\longrightarrow} +\infty$ on Ω^{cj}_t and the result becomes obvious. □

12 The Blumenthal-Getoor index

In the last section of these notes we wish to use the already made observation that if the path $s \mapsto X_s(\omega)$ is fully observed on $[0,t]$, then one also know the processes

$$H(r)_t \;=\; \sum_{s \leq T} \|\Delta X_s\|^r \tag{12.1}$$

for any $r \geq 0$ (with the convention $0^0 = 0$). This is not especially interesting, and it has no predictive value about the laws of the jumps, *but for one point*: we know for which r's we have $H(r)_t < \infty$. We will call *Blumenthal-Getoor index up to time T* the following random number

$$R_T \;=\; \inf(r: \; H(r)_T < \infty). \tag{12.2}$$

This is increasing with T, and $0 \leq R_T \leq 2$ always, and we have $H(r)_T = \infty$ for all $r < R_T$, and $H(r)_T < \infty$ for all $r > R_T$, whereas $H(R_T)_T$ may be finite or infinite (except that $H(2)_T < \infty$ always again). We will consider in this section the "estimation" of $R_T(\omega)$, in the same sense as we estimated the integrated volatility above. Clearly, R_T is the maximum of the Blumenthal-Getoor indices $R^i_t(\omega)$ for all components X^i, so this problem is essentially 1-dimensional, and in the sequel we assume X to be 1-dimensional.

To understand why this index is important let us consider the special situation where $X = X' + Y$, where X' is a *continuous* Itô semimartingale and Y is a Lévy process.

Of course $H(r)_t = \sum_{s\leq t} |\Delta Y_s|^r$, and the Lévy property yields the following equivalence, which holds for all $t > 0$:

$$H(r)_t < \infty \ \text{ a.s.} \quad \Longleftrightarrow \quad \int (|x|^r \wedge 1) F(dx) < \infty, \tag{12.3}$$

where F is the Lévy measure of Y. It is also characterized in the following way: writing

$$x > 0 \quad \mapsto \quad \overline{F}(x) \ = \ F([-x,x]^c), \tag{12.4}$$

for its (symmetrical) tail function (more generally, $\overline{H}(x) = H([-x,x]^c)$ for any measure H on $\mathbb{R}$), then the Blumenthal-Getoor index β is the unique number in $[0,2]$ such that for all $\varepsilon > 0$ we have

$$\lim_{x\to 0} \ x^{\beta+\varepsilon}\overline{F}(x) \ = \ 0, \qquad \limsup_{x\to 0} \ x^{\beta-\varepsilon}\overline{F}(x) \ = \ \infty. \tag{12.5}$$

Unfortunately, the "limsup" above is usually not a limit.

If Y is a stable process, its Blumenthal-Getoor index is the stability index, which is probably the most important parameter in the law of Y (the other three, a scaling constant and a drift and a skewness parameter are also of course important but no as much; note that here the scaling and skewness parameters can also be in principle estimated exactly, but the drift cannot). More generally, for a Lévy process the observation over $[0,t]$ does not allow to infer the Lévy measure, but one can infer in principle the Blumenthal-Getoor index, which indeed is about the only information which is known about F: this is an essential characteristic of the process, for modeling purposes for example.

So we are going to estimate R_T. Unfortunately, to do this we need some very restrictive assumptions. We start with the simple case when X is a symmetric stable process plus possibly a Brownian motion, then we state the results when X is a "general" Itô semimartingale, and we come back to Lévy process with a slightly different problem. The proofs are mainly given at the end.

To end these introductory remarks, let us introduce the processes which we will use here. The Blumenthal-Getoor index is related with the behavior of "small jumps", which correspond in our discrete observation scheme to the increments $\Delta_i^n X$ that are "small"; however we also have the continuous part X', which plays a preponderant role in those small increments. So we need to "truncate" from below the increments to get rid of the process X'. This leads us to take, as in the previous sections, two numbers $\varpi \in (0, \frac{1}{2})$ and $\alpha > 0$ and, this time, to consider increments bigger than $\alpha\Delta_n^\varpi$ only. We could a priori take a "general" test function, but it turns out that simply counting those not too small increments is enough. Hence we set for $u > 0$

$$U(u, \Delta_n)_t \ = \ \sum_{i=1}^{[t/\Delta_n]} 1_{\{|\Delta_i^n X| > u\}}, \tag{12.6}$$

and use in fact the processes $U(\alpha\Delta_n^\varpi, \Delta_n)$ or $U(\alpha\Delta_n^\varpi, 2\Delta_n)$. On the basis of these we introduce two different statistics, which will be in fact our estimators. Below, we choose $\varpi \in (0, \frac{1}{2})$ and two numbers $\alpha' > \alpha > 0$, and we set

$$\widehat{\beta}_n(t,\varpi,\alpha,\alpha') \ = \ \frac{\log(U(\alpha\Delta_n^\varpi, \Delta_n)_t / U(\alpha'\Delta_n^\varpi, \Delta_n)_t)}{\log(\alpha'/\alpha)}. \tag{12.7}$$

Other estimators of the same kind, but involving increments of sizes Δ_n and $k\Delta_n$ and the same cut-off level $\alpha\Delta_n^{\varpi}$ are possible, in the spirit of the previous two sections, but the results are essentially the same, and in particular the rates.

12.1 The stable process case.

In this subsection Y denotes a symmetric stable process with index $\beta \in (0,2)$. This is a Lévy process whose characteristic function is of the form $\mathbb{E}(e^{iuY_t}) = \exp(-ct|u|^\beta)$ for some constant c, and the Lévy measure is of the form

$$F(dx) \;=\; \frac{A\beta}{2|x|^{1+\beta}}\,dx, \qquad \text{hence}\quad \overline{F}(x) \;:=\; \frac{A}{x^\beta}\quad \text{for } x>0 \tag{12.8}$$

for some scale parameter $A>0$, related of course with the c above. The law of Y_1 has an even density g and a tail function $\overline{G}(x) = \mathbb{P}(|Y_1|>x)$ satisfying, as $x\to\infty$ (see [26], Theorems 2.4.2 and Corollary 2 of Theorem 2.5.1):

$$g(x) \;=\; \frac{A\beta}{2|x|^{1+\beta}} + \mathrm{O}\Big(\frac{1}{x^{1+2\beta}}\Big), \qquad \overline{G}(x) \;=\; \frac{A}{x^\beta} + \mathrm{O}\Big(\frac{1}{x^{2\beta}}\Big). \tag{12.9}$$

Let us begin with the case $X=Y$. In this case, $U(\alpha\Delta_n^{\varpi},\Delta_n)_t$ is the sum of $[t/\Delta_n]$ i.i.d. $\{0,1\}$-valued variables which, by the scaling property of Y (namely, Y_t has the same law as $t^{1/\beta}Y_1$) have the probability $\overline{G}(\alpha\Delta_n^{\varpi-1/\beta})$ of taking the value 1. Then the following result is completely elementary to prove (it will follows from the more general results proved later):

Theorem 12.1 *Assume that $X=Y$. Let $0<\alpha<\alpha'$ and $\varpi>0$ and $t>0$.*

a) If $\varpi<\frac{1}{\beta}$, the estimators $\widehat{\beta}_n(t,\varpi,\alpha,\alpha')$ converge in probability to β.

b) If $\varpi<\frac{2}{3\beta}$, we have

$$\frac{1}{\Delta_n^{\varpi\beta/2}}\,\big(\widehat{\beta}_n(t,\varpi,\alpha,\alpha')-\beta\big) \;\stackrel{\mathcal{L}}{\longrightarrow}\; \mathcal{N}\Big(0,\frac{\alpha'^\beta-\alpha^\beta}{At(\log(\alpha'/\alpha))^2}\Big), \tag{12.10}$$

The reader will observe that we do not necessarily assume $\varpi<\frac12$, because there is no Brownian part, and the restriction over ϖ will be explained later.

These estimators are not rate-efficient. To see that, one can recall from [2] that the model in which one observes the values $X_{i\Delta_n}$ for $i\Delta_n\le t$ is regular, and its Fisher information (for estimating β) is asymptotically of the form

$$I_n \;\sim\; \frac{\log(1/\Delta_n)}{\Delta_n}\,C_\beta\,t \tag{12.11}$$

for some constant C_β. So rate-efficient estimators would be such the rate of convergence is $\Delta_n^{-1/2}\sqrt{\log(1/\Delta_n)}$, instead of $\Delta_n^{-\varpi\beta/2}$ found here. With the "optimal" choice of ϖ, namely smaller than but as close as possible to $2/3\beta$, we get a rate which is "almost"

$\Delta_n^{-1/3}$ only. In addition β is unknown, so a conservative choice of ϖ is $\varpi = 1/3$ and the rate in (12.10) become $\Delta_n^{-\beta/6}$, quite far from the optimal rate.

The reason for this (huge) lack of optimality is that our method results in discarding a large part of the data. In the absence of a Brownian component this is of course unnecessary, but as seen immediately below the situation is different is a Brownian motion is present.

Now we turn to the situation where $X_t = bt + \sigma W_t + Y_t$, with Y as above.

Theorem 12.2 *Assume that $X_t = bt + \sigma W_t + Y_t$. Let $0 < \alpha < \alpha'$ and $\varpi > 0$ and $t > 0$.*

a) If $\varpi < \frac{1}{2}$, the estimators $\widehat{\beta}_n(t, \varpi, \alpha, \alpha')$ converge in probability to β.

b) If $\varpi < \frac{1}{2+\beta}$, we have (12.10).

These estimators are again not rate-efficient. In fact, one can extends [2] to obtain that in the present situation the Fisher information for estimating β, at stage n, satisfies

$$I_n \sim \frac{A\ (\log(1/\Delta_n))^{2-\beta/2}}{\sigma^\beta\ \Delta_n^{\beta/2}}\ C'_\beta\ t \tag{12.12}$$

for another constant C'_β. The discrepancy here comes from the fact that we have absolutely not used the fact that we exactly know the law of X. If one consider the (partial) statistical model where we observe only the increments bigger than $\alpha\Delta_n^\varpi$, the Fisher information becomes

$$I_n \sim \frac{A(1-\varpi)^2\ (\log(1/\Delta))^2}{\alpha^\beta\ \Delta_n^{\varpi\beta}}\ C'''_\beta\ t. \tag{12.13}$$

This still gives a faster rate than in the theorem, but by a (negligible) factor of $\log(1/\Delta_n)$. There is however the restriction $\varpi < \frac{1}{2+\beta}$, which does not appear in (12.13).

12.2 The general result.

The title of this subsection is rather misleading, since the solution of the problem requires quite strong assumptions. Unfortunately, this seems consubstantial to this problem, as one can see in the next subsection in a much simpler situation. We will assume that X is an Itô semimartingale, with conditions on σ_t even weaker than in (H) or (H'), but the assumptions on the Lévy measures $F_t = F_{\omega,t}(dx)$ of (1.6) are rather strong:

Assumption (L) : The process X is a 1-dimensional Itô semimartingale, with b_t and σ_t locally bounded. There are three (non-random) numbers $\beta \in (0, 2)$ and $\beta' \in [0, \beta/2)$ and $\gamma > 0$, and a locally bounded process $L_t \geq 1$, such that we have for all (ω, t):

$$F_t = F'_t + F''_t, \tag{12.14}$$

where

a) F'_t has the form

$$F'_t(dx) = \frac{1 + |x|^\gamma f(t,x)}{|x|^{1+\beta}} \left(a_t^{(+)} 1_{\{0<x\leq z_t\}} + a_t^{(-)} 1_{\{-z_t\leq x<0\}}\right) dx, \tag{12.15}$$

for some predictable non-negative processes $a_t^{(+)}, a_t^{(-)}, z_t$ and some predictable function $f(\omega, t, x)$, satisfying:

$$\frac{1}{L_t} \leq z_t \leq 1, \quad a_t^{(+)} + a_t^{(-)} \leq L_t, \quad 1 + |x| f(t,x) \geq 0, \quad |f(t,x)| \leq L_t. \tag{12.16}$$

b) F_t'' is a measure which is singular with respect to F_t' and satisfies

$$\int_{\mathbb{R}} (|x|^{\beta'} \wedge 1)\, F_t''(dx) \ \leq \ L_t. \tag{12.17}$$

□

This assumption implies in particular that $(|x|^r \wedge 1) * \nu_T$ is finite for all $r > \beta$, and infinite for all $r < \beta$ on the set $\{\overline{A}_T > 0\}$, where we have put

$$A_t = \frac{a_t^{(+)} + a_t^{(-)}}{\beta}, \qquad \overline{A}_t = \int_0^t A_s ds. \tag{12.18}$$

Therefore the Blumenthal-Getoor index R_T satisfies

$$R_T \leq \beta, \qquad A_T > 0 \ \Rightarrow \ R_T = \beta. \tag{12.19}$$

A stable process with index β satisfies (L), and this assumption really means that the small jumps of X behave like the small jumps of such a stable process, on the time set $\{t : A_t > 0\}$, whereas on the complement of this set they are "negligible" in comparison with the small jumps of the stable process. The solution of an equation like (6.2) satisfies (L) when Z is a stable process, and (much) more generally when Z is a Lévy process which itself satisfies (L) (like for example the sum of two stable processes plus a Wiener process, or of a stable process plus a compound Poisson process plus a Wiener process).

Theorem 12.3 *Let $0 < \alpha < \alpha'$ and $0 < \varpi < \frac{1}{2}$ and $t > 0$. Assume (L).*

a) We have $\widehat{\beta}_n'(t, \varpi, \alpha, \varpi') \stackrel{\mathbb{P}}{\longrightarrow} \beta$ on the set $\{\overline{A}_t > 0\}$.

b) If further $\beta' \in [0, \frac{\beta}{2+\beta})$ and $\gamma > \beta/2$, and if $\varpi < \frac{1}{2+\beta} \bigwedge \frac{1}{3\beta}$, in restriction to the set $\{\overline{A}_t > 0\}$ we have

$$\frac{1}{\Delta_n^{\varpi\beta/2}} \ (\widehat{\beta}_n(t, \varpi, \alpha, \alpha') - \beta) \ \stackrel{\mathcal{L}-s}{\longrightarrow} \ U, \tag{12.20}$$

where U is defined on an extension of the original space and is $\mathcal{F}$-conditionally centered Gaussian, with variance:

$$\widetilde{\mathbb{E}}(U^2 \mid \mathcal{F}) \ = \ \frac{\alpha'^{\beta} - \alpha^{\beta}}{\overline{A}_t (\log(\alpha'/\alpha))^2}. \tag{12.21}$$

At this point, we can replace the variances in (12.21) by estimators for them, to get a standardized CLT:

Theorem 12.4 *Under (L) and the assumptions of (b) of the previous theorem, the variables*

$$\frac{\log(\alpha'/\alpha)}{\sqrt{\frac{1}{U(\alpha'\Delta_n^{\varpi}, \Delta_n)_t} - \frac{1}{U(\alpha\Delta_n^{\varpi}, \Delta_n)_t}}} \ \left(\widehat{\beta}_n(t, \varpi, \alpha, \alpha') - \beta\right) \tag{12.22}$$

converge stably in law, in restriction to the set $\{\overline{A}_t > 0\}$, to a standard normal variable independent of $\mathcal{F}$.

Despite the strong assumptions, these estimators are thus reasonably good for estimating β on the (random) set $\{\overline{A}_t > 0\}$ on which the Blumenthal-Getoor index is actually β; unfortunately we do not know how they behave on the complement of this set.

12.3 Coming back to Lévy processes.

Let us restrict the setting of the previous subsection by assuming that X is a Lévy process, that is an Itô semimartingale with characteristics of the form (1.4). (L) may hold or not, but when it does we have $A_t = at$ for some constant $a > 0$, and so the two theorems 12.3 and 12.4 hold on the whole of Ω.

What is important here, though, is that those results probably fail, even in this simple setting, when (L) fails. We cannot really show this in a serious mathematical way, but we can see on a closely related and even simpler problem why strong assumptions are needed on the Lévy measure. This is what we are going to explain now.

The model is as follows: instead of observing the increments of X, we observe all its jumps (between 0 and t) whose sizes are bigger than $\alpha\Delta_n^\varpi$. A priori, this should give us more information on the Lévy measure than the original observation scheme.

In this setting the estimators (12.7) have no meaning, but may be replaced by

$$\overline{\beta}_n(t,\varpi,\alpha,\alpha') \;=\; \frac{\log(\overline{U}(\alpha\Delta_n^\varpi)_t/\overline{U}(\alpha'\Delta_n^\varpi)_t)}{\log(\alpha'/\alpha)}, \quad \text{where } \overline{U}(u)_t \;=\; \sum_{s\leq t} 1_{\{|\Delta X_s|>u\}}. \tag{12.23}$$

Lemma 12.5 *Let $\gamma_n(\alpha) \;=\; \overline{F}(\alpha\Delta_n^\varpi)$ and*

$$M^n(\alpha)_t \;=\; \frac{1}{\sqrt{\gamma_n(\alpha)}}\Big(\overline{U}(\alpha\Delta_n^\varpi)_t - \gamma_n(\alpha)\; t\Big). \tag{12.24}$$

a) The processes $M^n(\alpha)$ converge stably in law to a standard Wiener process, independent of $\mathcal{F}$.

b) If $\alpha < \alpha'$ all limit points of the sequence $\frac{\gamma_n(\alpha')}{\gamma_n(\alpha)}$ are in $[0,1]$. If further this sequence converges to γ then the pairs $(M^n(\alpha), M^n(\alpha'))$ of processes converge stably in law to a process $(\overline{W}, \overline{W}')$, independent of X, where $\overline{W}$ and $\overline{W}'$ are correlated standard Wiener processes with correlation $\sqrt{\gamma}$.

Proof. The processes $M^n = M^n(\alpha)$ and $M'^n = M^n(\alpha')$ are Lévy processes and martingales, with jumps going uniformly to 0, and with predictable brackets

$$\langle M^n, M^n\rangle_t \;=\; \langle M'^n, M'^n\rangle_t \;=\; t, \qquad \langle M^n, M'^n\rangle_t \;=\; \frac{\sqrt{\gamma_n(\alpha')}}{\sqrt{\gamma_n(\alpha)}}\; t.$$

Observe also that $\alpha'\Delta_n^\varpi \geq \alpha\Delta_n^\varpi$, hence $\gamma_n(\alpha') \leq \gamma_n(\alpha)$. All results are then obvious (see [13], Chapter VII). □

Theorem 12.6 *If $\alpha' > \alpha$ and if $\frac{\gamma_n(\alpha')}{\gamma_n(\alpha)} \to \gamma \in [0,1]$, then the sequence*

$$\sqrt{\gamma_n(\alpha')}\,\left(\overline{\beta}_n(t,\varpi,\alpha,\alpha') - \frac{\log(\gamma_n(\alpha)/\gamma_n(\alpha'))}{\log(\alpha'/\alpha)}\right) \tag{12.25}$$

converges stably in law to a variable, independent of $\mathcal{F}$ and with the law $\mathcal{N}\Big(0, \frac{1-\gamma}{t(\log(\alpha'/\alpha))^2}\Big)$.

This result is a simple consequence of the previous lemma, and its proof is the same as for Theorem 12.3 and is thus omitted.

This result shows that in general, that is without specific assumptions on F, the situation is hopeless. These estimators are not even consistent for estimating the Blumenthal-Getoor index β of F, because of a bias, and to remove the bias we have to know the ratio $\gamma_n(\alpha')/\gamma_n(\alpha)$ (or at least its asymptotic behavior in a precise way), and further there is no CLT if this ratio does not converge (a fact which we a priori do not know, of course).

The major difficulty comes from the possible erratic behavior of $\overline{F}$ near 0. Indeed, we have (12.5), but there are Lévy measures F satisfying this, and such that for any $r \in (0,\beta)$ we have $x_n^r\overline{F}(x_n) \to 0$ for a sequence $x_n \to 0$ (depending on r, of course). If F is such, the sequence $\gamma_n(\varpi,\alpha')/\gamma_n(\varpi,\alpha)$ may have the whole of $[0,1]$ as limit points, depending on the parameter values ϖ,α,α', and in a completely uncontrolled way for the statistician.

So we need some additional assumption on F. For the consistency a relatively weak assumption is enough, for the asymptotic normality, we need in fact (L). Recall that under (L) we have necessarily $\overline{A}_t = at$ for some $a \geq 0$, in the Lévy case.

Theorem 12.7 *a) If the tail function $\overline{F}$ is regularly varying at 0, with index $\beta \in (0,2)$ we have $\overline{\beta}_n(t,\varpi,\alpha,\alpha') \stackrel{\mathbb{P}}{\longrightarrow} \beta$.*

b) Under (L) with $a >$, the sequence $\frac{1}{\Delta_n^{\varpi\beta}}\Big(\overline{\beta}_n(t,\varpi,\alpha,\alpha') - \beta\Big)$ converges stably in law to a variable, independent of $\mathcal{F}$ and with law $\mathcal{N}\Big(0, \frac{\alpha'^\beta - \alpha^\beta}{t\,\alpha'^\beta\,(\log(\alpha'/\alpha))^2}\Big)$.

Proof. The regular variation implies $\gamma_n(\alpha) \to \infty$ and $\gamma_n(\alpha')/\gamma_n(\alpha) \to (\alpha/\alpha')^\beta$, so the previous theorem yields (a). (L) clearly implies

$$\sqrt{\gamma_n(\alpha)}\;\frac{\log(\gamma_n(\alpha)/\gamma_n(\alpha'))}{\log(\alpha'/\alpha)} \to \beta,$$

and also $\gamma_n(\varpi,\alpha) \sim a/\alpha^\beta \Delta_n^{\varpi\beta}$, so (b) follows again from the previous theorem. □

It may of course happen that the regular variation or (L) fail and nevertheless the conclusions of the previous theorem hold for a particular choice of the parameters ϖ,α,α'. But in view of Theorem 12.6 and of the previous proof these assumptions are *necessary* if we want those conclusions to hold *for all choices* of ϖ,α,α'.

Now if we come back to the original problem, for which only increments of X are observed. We have Theorem 12.3 whose part (b) looks like (b) above; however there are restrictions on ϖ, unlike in Theorem 12.7. This is because an increment $\Delta_i^n X$ with size bigger than $\alpha\Delta_n^\varpi$ is, with a high probability, almost equal to a "large" jump only when the cutoff level is higher than a typical Brownian increments, implying at least $\varpi < 1/2$.

12.4 Estimates.

As all the results in these notes, Theorem 12.3 is "local" in time. So by our usual localization procedure we may assume that (L) is replaced by the stronger assumption below:

Assumption (SL): We have (L), and the process L_t is in fact a constant L, and further $|b_t| \leq L$ and $|\sigma_t| \leq L$ and $|X_t| \leq L$. □

Before proceeding, we mention a number of elementary consequences of (SL), to be used many times. First, F_t is supported by the interval $[-2L, 2L]$. This and (12.15) and (12.17) imply that for all $u, v, x, y > 0$ we have

$$\left.\begin{array}{l} \overline{F}''_t(x) \;\leq\; \frac{K}{x^{\beta'}}, \qquad \left|\overline{F}_t(x) - \frac{A_t}{x^\beta}\right| \;\leq\; \frac{K}{x^{(\beta-\gamma)\vee\beta'}}, \qquad \overline{F}_t(x) \;\leq\; \frac{K}{x^\beta}, \\ \int_{\{|x|\leq u\}} x^2 F_t(dx) \;\leq\; K u^{2-\beta}, \qquad \int |x| F''_t(dx) \;\leq\; K \\ \int_{\{|x|>u\}} (|x|^v \wedge 1) F_t(dx) \;\leq\; \begin{cases} K_v & \text{if } v > \beta \\ K_v \log(1/u) & \text{if } v = \beta \\ K_v u^{v-\beta} & \text{if } v < \beta, \end{cases} \\ \overline{F}_t(x) - \overline{F}_t(x+y) \;\leq\; \frac{K}{x^\beta}\left(1 \wedge \frac{y}{x} + x^{(\beta-\beta')\wedge\gamma}\right). \end{array}\right\} \tag{12.26}$$

In the next lemma, Y is a symmetric stable process with Lévy measure (12.8), and for $\eta \in (0,1)$ we set

$$Y(\eta)_t \;=\; \sum_{s\leq t} \Delta Y_s 1_{\{|\Delta Y_s|>\eta)\}}, \qquad Y'(\eta) \;=\; Y - Y(\eta). \tag{12.27}$$

Lemma 12.8 *There is a constant K depending on (A,β), such that for all $s, \eta \in (0,1)$,*

$$\mathbb{P}(|Y'(\eta)_s| > \eta/2) \;\leq\; K\, s^{4/3}/\eta^{4\beta/3}. \tag{12.28}$$

Proof. We use the notation (12.8) and (12.9). Set $\eta' = \eta/2$ and $\theta = s\overline{F}(\eta') = sA/\eta'^\beta$, and consider the processes $Y' = Y'(\eta')$ and $Z_t = \sum_{r\leq t} 1_{\{|\Delta Y_r|>\eta'\}}$. Introduce also the sets

$$D = \{|Y_s| > \eta'\}, \quad D' = \{|Y'_s| > \eta'\}, \quad B = \{Z_s = 1\}, \quad B' = \{Z_s = 0\}.$$

It is of course enough to prove the result for s/η^β small, so below we assume $\theta \leq 1/2$.

By scaling, $\mathbb{P}(D) = \overline{G}(\eta' s^{-1/\beta})$, so (12.9) yields

$$|\mathbb{P}(D) - \theta| \;\leq\; K\theta^2. \tag{12.29}$$

On the other hand Z_s is a Poisson variable with parameter $\theta \leq 1/2$, hence

$$|\mathbb{P}(B) - \theta| \;\leq\; K\theta^2. \tag{12.30}$$

Since Y' is a purely discontinuous Lévy process without drift and whose Lévy measure is the restriction of F to $[-\eta', \eta']$, we deduce from (12.8) that

$$\mathbb{E}((Y'_s)^2) \;=\; s\int_{\{|x|\leq\eta'\}} x^2 F(dx) \;\leq\; K\theta\eta^2. \tag{12.31}$$

The two processes Y' and Z are independent, and conditionally on B the law of the variable $Y_s - Y'_s$ is the restriction of the measure $\frac{s}{\theta}F$ to $[-\eta',\eta']^c$, and $\mathbb{P}(B) = \theta e^{-\theta}$. Thus

$$\begin{aligned}\mathbb{P}(B \cap D^c) &= e^{-\theta}\ s \int_{\{|x|>\eta'\}} F(dx)\ \mathbb{P}(|Y'_s + x| \leq \eta') \\ &\leq\ s\Big(F(\{\eta' < |x| \leq \eta'(1+\theta^{1/3})/2\}) + F(\{|x|>\eta'\})\ \mathbb{P}(|Y'_s| > \eta'\theta^{1/3})\Big) \\ &\leq\ \theta\Big(1-(1+\theta^{1/3})^{-\beta} + \frac{4}{\eta'^2\theta^{2/3}}\ \mathbb{E}((Y'_s)^2)\Big)\ \leq\ K\,\theta^{4/3}, \end{aligned} \tag{12.32}$$

where we have used (12.31) for the last inequality.

Now, we have

$$\mathbb{P}(D \cap B^c)\ =\ \mathbb{P}(D) - \mathbb{P}(B) + \mathbb{P}(B \cap D^c).$$

Observe also that $D \cap B' = D' \cap B'$, and D' and B' are independent, hence

$$\mathbb{P}(D')\ =\ \frac{\mathbb{P}(D' \cap B')}{\mathbb{P}(B')}\ =\ \frac{\mathbb{P}(D \cap B')}{\mathbb{P}(B')}\ \leq\ \frac{\mathbb{P}(D \cap B^c)}{\mathbb{P}(B')}\ \leq\ K\mathbb{P}(D \cap B^c)$$

because $\mathbb{P}(B') = e^{-\theta} \geq e^{-1/2}$. The last two displays, plus (12.29), (12.30) and (12.32) give us $\mathbb{P}(D')\ \leq\ K\theta^{4/3}$, hence the result. □

Now we turn to semimartingales. We have (12.14) and there exists a predictable subset Φ of $\Omega \times (0,\infty) \times \mathbb{R}$ such that

$$\begin{array}{l} F''_t(\omega,.)\ \text{ is supported by the set }\ \{x : (\omega,t,x) \in \Phi\} \\ F'_t(\omega,.)\ \text{ is supported by the set }\ \{x : (\omega,t,x) \notin \Phi\}. \end{array} \tag{12.33}$$

Next we will derive a decomposition of X a bit similar to (8.20), but here we have a control on the Lévy measure of X itself, through (SL), so it is more convenient to truncate at the value taken by ΔX_t rather than by the function γ. Recall that the jumps of X are bounded, so we can write X in the form (6.21), with still b_t bounded. For any $\eta \in (0,1]$ we set

$$b(\eta)_t\ =\ b_t - \int_{\{|x|>\eta\}} F'_t(dx)x - \int F''_t(dx)x$$

By (12.26) and (SL) the process $b(\eta)_t$ is well defined and satisfies $|b(\eta)_t| \leq K/\eta$. Then by (6.21) we can write $X = X(\eta) + X'(\eta)$, where $X'(\eta) = \widehat{X}(\eta) + \widehat{X}'(\eta) + \widehat{X}''(\eta)$ and

$$\begin{array}{ll} X(\eta) = (x1_{\{|x|>\eta\}}) \star \mu, & \widehat{X}(\eta)_t = X_0 + \int_0^t b(\eta)_s ds + \int_0^t \sigma_s dW_s \\ \widehat{X}'(\eta) = (x1_{\{|x|\leq\eta\}}\, 1_{\Phi^c}) \star (\mu - \nu), & \widehat{X}''(\eta) = (x1_{\{|x|\leq\eta\}}\, 1_{\Phi}) \star \mu. \end{array}$$

Lemma 12.9 *Assume (SL). We have for all $p \geq 2$:*

$$\left.\begin{array}{l} \mathbb{E}^n_{i-1}(|\Delta^n_i \widehat{X}(\eta)|^p)\ \leq\ K_p\,(\Delta_n^{p/2} + \eta^{-p}\Delta_n^p) \\ \mathbb{E}^n_{i-1}(|\Delta^n_i \widehat{X}'(\eta)|^2)\ \leq\ K\,\Delta_n\,\eta^{2-\beta} \\ \mathbb{E}^n_{i-1}(|\Delta^n_i \widehat{X}''(\eta)|^{\beta'})\ \leq\ K\,\Delta_n. \end{array}\right\} \tag{12.34}$$

Proof. The first estimate is obvious (see after (8.34)), whereas the second one is obtained from the second line of (12.26). Since $\beta' < 1$, we have $|\sum_j x_j|^{\beta'} \leq \sum_j |x_j|^{\beta'}$ for any sequence (x_j), hence

$$\begin{aligned}\mathbb{E}^n_{i-1}(|\Delta^n_i \widehat{X}''(\eta)|^{\beta'}) &\leq \mathbb{E}^n_{i-1}\Big(\Delta^n_i((|x|^{\beta'} 1_{\{|x|\leq\eta\}} 1_\Phi 1_{(t,\infty)}) \star \mu)\Big) \\ &= \mathbb{E}^n_{i-1}\Big(\int_{(i-1)\Delta_n}^{i\Delta_n} dr \int_{\{|x|\leq\eta\}} |x|^{\beta'} F'_r(dx)\Big) \leq K\Delta_n. \qquad \square\end{aligned}$$

Next, we give a general result on counting processes. Let N be a counting process (that is, right continuous with $N_0 = 0$, piecewise constant, with jumps equal to 1) adapted to $(\mathcal{F}_t)$ and with predictable compensator of the form $G_t = \int_0^t g_s ds$.

Lemma 12.10 *With N and G as above, and if further $g_t \leq u$ for some constant $u > 0$, we have*

$$|\mathbb{P}^n_{i-1}(\Delta^n_i N = 1) - \mathbb{E}^n_{i-1}(\Delta^n_i G)| + \mathbb{P}^n_{i-1}(\Delta^n_i N \geq 2) \leq (u\Delta_n)^2. \tag{12.35}$$

Proof. Introduce the successive jump times $T_1, T_2, \cdots$ of N after time $(i-1)\Delta_n$, the sets $D = \{\Delta^n_i N = 1\}$ and $D' = \{\Delta^n_i N \geq 2\}$ and the variable $G'^n_i = \mathbb{E}^n_{i-1}(\Delta^n_i G)$. Then

$$\mathbb{P}^n_{i-1}(D) = \mathbb{E}^n_{i-1}(N_{(i\Delta_n)\wedge T_1} - N_{(i-1)\Delta_n}) = \mathbb{E}^n_{i-1}\Big(\int_{(i-1)\Delta_n}^{(i\Delta_n)\wedge T_1} g_r dr\Big) \leq G'^n_i \leq u\Delta_n,$$

$$G'^n_i - \mathbb{P}^n_{i-1}(D) = \mathbb{E}^n_{i-1}\Big(\int_{(i\Delta_n)\wedge T_1}^{i\Delta_n} g_r dr\Big) \leq u\Delta_n \, \mathbb{P}^n_{i-1}(D) \leq (u\Delta_n)^2$$

This gives us the first estimate. Next,

$$\begin{aligned}\mathbb{P}^n_{i-1}(D') &= \mathbb{P}^n_{i-1}(T_2 \leq i\Delta_n) = \mathbb{E}^n_{i-1}\Big(1_{\{T_1 < i\Delta_n\}} \, \mathbb{P}^n_{i-1}(T_2 \leq i\Delta_n \mid \mathcal{F}_{T_1})\Big) \\ &= \mathbb{E}^n_{i-1}\Big(1_{\{T_1 < i\Delta_n\}} \, \mathbb{E}\Big(\int_{T_1}^{(i\Delta_n)\wedge T_2} g_r dr \mid \mathcal{F}_{T_1}\Big)\Big) \leq u\Delta_n \, \mathbb{P}^n_{i-1}(D) \leq (u\Delta_n)^2,\end{aligned}$$

hence the second estimate. $\square$

Lemma 12.11 *With the notation $N(\eta)_t = \sum_{s\leq t} 1_{\{|\Delta X_s|>\eta\}}$, for all $\eta \in (0,1]$, $\zeta \in (0, \frac{1}{2})$ and $p \geq 2$ we have*

$$\mathbb{P}^n_{i-1}(\Delta^n_i N(\eta) \geq 1, \, |\Delta^n_i X'(\eta)| > \eta\zeta) \leq K_p\Big(\frac{\Delta_n^{p/2}}{\zeta^p \, \eta^p} + \frac{\Delta_n^p}{\zeta^p \eta^{2p}} + \frac{\Delta_n^2}{\zeta^2 \, \eta^{2\beta}} + \frac{\Delta_n}{\eta^{\beta'} \zeta^{\beta'}}\Big). \tag{12.36}$$

Proof. (12.34) and Bienaymé-Tchebycheff inequality yield

$$\mathbb{P}^n_{i-1}\Big(|\Delta^n_i \widehat{X}(\eta)| > \frac{\eta\zeta}{4}\Big) \leq K_p \, \Big(\frac{\Delta_n^{p/2}}{\eta^p \zeta^p} + \frac{\Delta_n^p}{\eta^{2p} \zeta^p}\Big), \qquad \mathbb{P}^n_{i-1}\Big(|\Delta^n_i \widehat{X}''(\eta)| > \frac{\eta\zeta}{4}\Big) \leq K \, \frac{\Delta_n}{\eta^{\beta'} \zeta^{\beta'}}.$$

Since $X'(\eta)=\widehat{X}(\eta)+\widehat{X}'(\eta)+\widehat{X}'(\eta)$ it remains to prove

$$\mathbb{P}^n_{i-1}\Big(\Delta^n_i N(\eta)\geq 1,\ |\Delta^n_i\widehat{X}'(\eta)|>\frac{\eta\zeta}{2}\Big)\ \leq\ K\frac{\Delta_n^2}{\zeta^2\,\eta^{2\beta}}. \tag{12.37}$$

For simplicity, write $N_s=N(\eta)_{(i-1)\Delta_n+s}-N(\eta)_{(i-1)\Delta_n}$ and $Y_s=\widehat{X}'(\eta)_{(i-1)\Delta_n+s}-\widehat{X}'(\eta)_{(i-1)\Delta_n}$. By Bienaymé-Tchebycheff inequality again the left side of (12.37) is not bigger than $4\mathbb{E}(N_{\Delta_n}Y^2_{\Delta_n})/\eta^2\zeta^2$. Now, N is a counting process and Y is a purely discontinuous square-integrable martingale, and they have no common jumps, so Itô's formula yields

$$N_sY_s^2\ =\ 2\int_0^s N_{r-}Y_{r-}dY_r+\int_0^s Y^2_{r-}dN_r+\sum_{r\leq s}N_{r-}(\Delta Y_r)^2.$$

Moreover, the compensator N is as in the previous lemma, with $g_s\leq K\eta^{-\beta}$, and the predictable quadratic variation of Y is $G'_s=\int_0^s g'_r dr$ with $g'_r\leq K\eta^{2-\beta}$ (see Lemma 12.9). Then taking expectations in the above display, and since the first term of the right side above is a martingale, we get

$$\begin{aligned}\mathbb{E}^n_{i-1}(N_{\Delta_n}Y^2_{\Delta_n})\ &=\ \mathbb{E}^n_{i-1}\Big(\int_0^{\Delta_n}Y_r^2dG_r+\int_0^{\Delta_n}N_rdG'_r\Big)\leq K\eta^{-\beta}\int_0^{\Delta_n}\mathbb{E}^n_{i-1}\Big(Y_r^2+\eta^2N_r\Big)\,dr\\ &=\ K\eta^{-\beta}\int_0^{\Delta_n}\mathbb{E}^n_{i-1}\Big(G'_r+\eta^2G_r\Big)\,dr\ \leq K\eta^{2(1-\beta)}\,\Delta_n^2.\end{aligned}$$

(12.37) is then obvious. □

The following lemma is key to the whole proof. We use the notation $u_n=\alpha\Delta_n^{\varpi}$.

Lemma 12.12 *Let $\alpha>0$, $\varpi\in(0,\frac12))$ and $\eta\in(0,\frac12-\varpi)$, and set*

$$\rho=\eta\wedge(\varpi(\beta-\beta')-\beta'\eta)\wedge(\varpi\gamma)\wedge(1-\varpi\beta-2\eta) \tag{12.38}$$

There is a constant K depending on (α,ϖ,η), and also on the characteristics of X, such that

$$\Big|\mathbb{P}^n_{i-1}(|\Delta^n_iX|>u_n)-\mathbb{E}^n_{i-1}\Big(\int_{(i-1)\Delta_n}^{i\Delta_n}\overline{F}_r(u_n)dr\Big)\Big|\leq K\Delta_n^{1-\varpi\beta+\rho} \tag{12.39}$$

$$\mathbb{P}^n_{i-1}(u_n<|\Delta^n_iX|\leq u_n(1+\Delta_n^{\eta}))\ \leq\ K\Delta_n^{1-\varpi\beta+\rho} \tag{12.40}$$

$$\mathbb{P}^n_{i-1}(|\Delta^n_iX|>u_n)\ \leq\ K\Delta_n^{1-\varpi\beta}. \tag{12.41}$$

Proof. 1) Observe that $\rho>0$, and it is clearly enough to prove the results when Δ_n is smaller than some number $\xi\in(0,1)$ to be chosen later, and independent of i and n.

We can apply (12.34) and Bienaymé-Tchebycheff inequality to obtain

$$\mathbb{P}^n_{i-1}(|\Delta^n_i\widehat{X}(u_n)|>u_n\Delta_n^{\eta}/2)\ \leq\ K_p\,\Delta_n^{p(1-2\varpi-2\eta)/2}$$

$$\mathbb{P}^n_{i-1}(|\Delta^n_i\widehat{X}''(u_n)|>u_n\Delta_n^{\eta}/2)\ \leq\ K\,\Delta_n^{1-\beta'(\varpi+2\eta)}.$$

Then by choosing p large enough and by (12.38), we see that $Y^n = \widehat{X}(u_n) + \widehat{X}''(u_n)$ satisfies

$$\mathbb{P}^n_{i-1}(|\Delta_i^n Y^n| > u_n \Delta_n^\eta) \ \leq \ K\,\Delta_n^{1-\varpi\beta+\rho}. \tag{12.42}$$

2) By (SL) we have $F'_r(dx) \leq (L'/|x|^{1+\beta})dx$ in restriction to $[-1,1]$, for some constant L'. We fix n. For each $\omega \in \Omega$ we endow the canonical (Skorokhod) space $(\Omega', \mathcal{F}', (\mathcal{F}'_t))$ of all càdlàg functions on $\mathbb{R}_+$ starting from 0 with the (unique) probability measure Q_ω under which the canonical process X' is a semimartingale with characteristics $(0,0,\nu'_\omega)$, where

$$\nu'_\omega(\omega', dr, dx) \ = \ dr\ 1_{\{|x|\leq u_n\}} \Big(\frac{L'}{|x|^{1+\beta}}\ dx - F'_r(\omega, dx)\Big). \tag{12.43}$$

This measure does not depend on ω', hence under Q_ω the process X' has independent increments; $\nu'_\omega(\omega', dr, dx)$ depends measurably on ω, hence $Q_\omega(d\omega')$ is a transition probability from $(\Omega, \mathcal{F})$ into $(\Omega', \mathcal{F}')$. Then we extend X, X' and other quantities defined on Ω or Ω' in the usual way (without changing the symbols) to the product $\widetilde{\Omega} = \Omega \times \Omega'$ endowed with the product σ-field $\widetilde{\mathcal{F}}$, the product filtration $(\widetilde{\mathcal{F}}_t)$, and the probability measure $\widetilde{\mathbb{P}}(d\omega, d\omega') = \mathbb{P}(d\omega)\ Q_\omega(d\omega')$.

Because of (12.26) and (12.43), and as in Lemma 12.9, $\mathbb{E}_{Q_\omega}(|\Delta_i^n X'|^2 \mid \widetilde{\mathcal{F}}_{(i-1)\Delta_n}) \leq K\Delta_n u_n^{2-\beta}$, so for some constant C depending on α and β but not on n and ω we have

$$Q_\omega(|\Delta_i^n X'| > u_n\Delta_n^\eta \mid \widetilde{\mathcal{F}}_{(i-1)\Delta_n}) \ \leq \ C\Delta_n^{1-\varpi\beta-2\eta} \ \leq \ C\Delta_n^\rho. \tag{12.44}$$

3) By well known results on extensions of spaces (see e.g. [13], Section II.7; note that the present extension of the original space is a "very good extension"), X' is a semimartingale on the extension with characteristics $(0,0,\nu')$, where $\nu'((\omega,\omega'), dr, dx) = \nu'_\omega(dr,dx)$, and any semimartingale on the original space is a semimartingale on the extension, with the same characteristics. Moreover X and X' have almost surely no common jump, so the sum $Y'(u_n) = \widehat{X}'(u_n) + X'$ is a semimartingale with characteristics $(0,0,\nu')$, where

$$\nu'(dr,dx) \ = \ dr\ 1_{\{|x|\leq u_n\}}\, F'_r(dx) + \nu_\omega(dr,dx) \ = \ 1_{\{|x|\leq u_n\}}\, \frac{L'}{|x|^{1+\beta}}\ dr\ dx,$$

where the last equality comes from (12.43). It follows that $Y'(u_n)$ is a Lévy process with Lévy measure given above, or in other words it is a version of the process $Y'(u_n)$ of (12.27) with $A = 2L'/\beta$. Hence, recalling (12.26), we deduce from (12.28) and from the Lévy property of $Y'(u_n)$ that, as soon as $\Delta_n^\eta \leq 1/4$, and if $A \in \mathcal{F}_{(i-1)\Delta_n}$:

$$\widetilde{\mathbb{P}}(A \cap \{|\Delta_i^n Y'(u_n)| > u_n(1-2\Delta_n^\eta)\}) \ \leq \ K\Delta_n^{4/3-4\varpi\beta/3}. \tag{12.45}$$

Next, let ξ be such that $C\xi^\rho \leq 1/2$. With A as above, and if $\Delta_n \leq \xi$, we can write

$$\begin{aligned}
&\widetilde{\mathbb{P}}(A \cap \{|\Delta_i^n Y'(u_n)| > u_n(1-2\Delta_n^\eta)\}) \\
&\quad \geq \ \widetilde{\mathbb{P}}\Big(A \cap \{|\Delta_i^n \widehat{X}'(u_n)| > u_n(1-\Delta_n^\eta)\} \cap \{|\Delta_i^n X'| \leq u_n\Delta_n^\eta\}\Big) \\
&\quad = \ \widetilde{\mathbb{E}}\Big(1_{\{A\cap\{|\Delta_i^n \widehat{X}'(u_n)|>u_n(1-\Delta_n^\eta)\}\}}\ Q_\cdot\Big(|X'_{t+s} - X'_t| \leq u_n\Delta_n^\eta\Big)\Big) \\
&\quad \geq \ \frac{1}{2}\ \mathbb{P}\Big(A \cap \{|\Delta_i^n \widehat{X}'(u_n)| > u_n(1-\Delta_n^\eta)\}\Big)
\end{aligned}$$

where the last inequality comes from (12.44). Then by (12.45) and the facts that A is arbitrary in $\mathcal{F}_{(i-1)\Delta_n}$ and that $\rho \leq \frac{1-\varpi\beta}{3}$ we deduce

$$\mathbb{P}^n_{i-1}\Big(|\Delta^n_i \widehat{X}'(u_n)| > u_n(1-\Delta^\eta_n)\Big) \;\leq\; K\Delta_n^{4/3-4\varpi\beta/3} \;\leq\; K\Delta_n^{1-\varpi\beta+\rho}.$$

In turn, combining this with (12.42), we readily obtain

$$\mathbb{P}^n_{i-1}\Big(|\Delta^n_i X'(u_n)| > u_n\Big) \;\leq\; K\,\Delta_n^{1-\varpi\beta-\rho}. \tag{12.46}$$

4) Now we write $u'_n = u_n(1+\Delta^\eta_n)$ and also

$$\theta^n_i \;=\; \mathbb{E}^n_{i-1}\Big(\int_{(i-1)\Delta_n}^{i\Delta_n} \overline{F}_r(u_n)dr\Big), \qquad \theta'^n_i \;=\; \mathbb{E}\Big(\int_{(i-1)\Delta_n}^{i\Delta_n} \overline{F}_r(u'_n)dr\Big),$$

and introduce the following two counting process

$$N^n_t \;=\; \sum_{s\leq t} 1_{\{|\Delta X_s|>u_n\}}, \qquad N'^n_t \;=\; \sum_{s\leq t} 1_{\{|\Delta X_s|>u'_n\}}.$$

Their predictable compensators are $\int_0^t \overline{F}_r(u_n)dr$ and $\int_0^t \overline{F}_r(u'_n)dr$, whereas both $\overline{F}_r(u_n)$ and $\overline{F}_r(u'_n)$ are smaller than $K/\Delta_n^{\varpi\beta}$. Hence (12.35) gives

$$\begin{array}{l} |\mathbb{P}^n_{i-1}(\Delta^n_i N^n = 1) - \theta^n_i| + \mathbb{P}^n_{i-1}(\Delta^n_i N^n \geq 2) \;\leq\; K\Delta_n^{2(1-\varpi\beta)}, \\ |\mathbb{P}^n_{i-1}(\Delta^n_i N'^n = 1) - \theta'^n_i| \;\leq\; K\Delta_n^{2(1-\varpi\beta)}. \end{array} \tag{12.47}$$

Since $N^n - N'^n$ is non-decreasing, we have

$$\begin{array}{l} \mathbb{P}^n_{i-1}(\Delta^n_i N^n = 1,\; \Delta^n_i N'^n = 0) \;=\; \mathbb{P}^n_{i-1}(\Delta^n_i N^n = 1) \\ \qquad -\mathbb{P}^n_{i-1}(\Delta^n_i N'^n = 1) + \mathbb{P}^n_{i-1}(\Delta^n_i N^n \geq 2,\; \Delta^n_i N'^n = 1). \end{array}$$

Then (12.47) yields

$$|\mathbb{P}^n_{i-1}(\Delta^n_i N^n = 1,\; \Delta^n_i N'^n = 0) - (\theta^n_i - \theta'^n_i)| \;\leq\; K\Delta_n^{2(1-\varpi\beta)}. \tag{12.48}$$

Moreover (12.26) clearly implies $\theta^n_i - \theta'^n_i \leq K\Delta_n^{1-\varpi\beta}(\Delta^\eta_n + \Delta_n^{\varpi(\gamma\wedge(\beta-\beta'))}) \leq K\Delta_n^{1-\varpi\beta+\rho}$. We then deduce from (12.48) that

$$\mathbb{P}^n_{i-1}(\Delta^n_i N^n = 1,\; \Delta^n_i N'^n = 0) \;\leq\; K\Delta_n^{1-\varpi\beta+\rho}. \tag{12.49}$$

5) If $\Delta^n_i N^n = \Delta^n_i N'^n = 1$ and $|\Delta^n_i X| \leq u_n$, then necessarily $|\Delta^n_i X(u'_n)| > u_n\Delta^\eta_n$. Hence

$$\begin{array}{l} \mathbb{P}^n_{i-1}(\Delta^n_i N^n = 1, |\Delta^n_i X| \leq u_n) \leq \mathbb{P}^n_{i-1}(\Delta^n_i N^n = 1,\; \Delta^n_i N'^n = 0) \\ \qquad +\mathbb{P}^n_{i-1}(\Delta^n_i N^n = 1,\; |\Delta^n_i X'(u_n)| > u_n\Delta^\eta_n). \end{array}$$

Then if we apply (12.36) with p large enough and $\eta = u_n$ and $\zeta = \Delta^\eta_n$, and (12.49), we deduce

$$\mathbb{P}^n_{i-1}(\Delta^n_i N^n = 1,\; |\Delta^n_i X| \leq u_n) \;\leq\; K\Delta_n^{1-\varpi\beta+\rho}. \tag{12.50}$$

Finally $\Delta_i^n X = \Delta_i^n X'(u_n)$ on the set $\{\Delta_i^n N^n = 0\}$, so

$$\begin{aligned}\mathbb{P}_{i-1}^n(|\Delta_i^n X| > u_n) \;&=\; \mathbb{P}_{i-1}^n(\Delta_i^n N^n = 1) - \mathbb{P}_{i-1}^n(\Delta_i^n N^n = 1,\ |\Delta_i^n X| \le u_n))\\ &\quad + \mathbb{P}_{i-1}^n(\Delta_i^n N^n = 0,\ |\Delta_i^n X'(u_n)| > u_n)\\ &\quad + \mathbb{P}_{i-1}^n(\Delta_i^n N^n \ge 2,\ |\Delta_i^n X| > u_n).\end{aligned}$$

Then if we combine (12.46), (12.47) and (12.50), if $\Delta_n \le \xi$h we readily obtain (12.39). We also trivially deduce (12.41) from (12.26) and (12.39),

6) Finally, a close look at the previous argument shows that (12.39) also holds with $\alpha\Delta_n^\varpi(1+\Delta_n^\eta)$ and $\theta_i'^n$ in place of $\alpha\Delta_n^\varpi$ and θ_i^n. Therefore (12.40) follows, upon using the property $\theta_i^n - \theta_i'^n \le K\Delta_n^{1-\varpi\beta+\rho}$ proved above. □

Lemma 12.13 *Under the assumption and with the notation of Lemma 12.12, and if M is a bounded continuous martingale, we have (with K depending also on M):*

$$\Big|\mathbb{E}_{i-1}^n\left(\Delta_i^n M\, 1_{\{|\Delta_i^n X| > u_n\}}\right)\Big| \le\; K\Delta_n^{1-\varpi\beta+\rho} + K\Delta_n^{1-(\varpi+\eta)\beta}\,\mathbb{E}_{i-1}^n(|\Delta_i^n M|). \tag{12.51}$$

Proof. 1) There exist C^2 functions f_n such that (with K independent of n):

$$\begin{aligned}&1_{\{|x|>u_n(1+2\Delta_n^\eta/3)\}} \le f_n(x) \le 1_{\{|x|>u_n(1+\Delta_n^\eta/3)\}}\\ &|f_n'(x)| \le \tfrac{K}{\Delta_n^{\varpi+\eta}}, \qquad |f_n''(x)| \le \tfrac{K}{\Delta_n^{2(\varpi+\eta)}}.\end{aligned} \tag{12.52}$$

With $\widehat{X}' = X - B - X^c$, and since M is bounded, we have

$$\begin{aligned}&\Big|\mathbb{E}_{i-1}^n(\Delta_i^n M\, 1_{\{|\Delta_i^n X|>u_n\}}) - \mathbb{E}_{i-1}^n(\Delta_i^n M\, f_n(\Delta_i^n \widehat{X}'))\Big|\\ &\le\; K\mathbb{P}_{i-1}^n(u_n < |\Delta_i^n X| \le u_n(1+\Delta_n^\eta)) + K\mathbb{E}_{i-1}^n(|f_n(\Delta_i^n X) - f_n(\Delta_i^n \widehat{X}')|).\end{aligned} \tag{12.53}$$

Now we have

$$|f_n(x+y) - f_n(x)| \;\le\; 1_{\{|y|>u_n\Delta_n^\eta/3)\}} + \frac{K}{\Delta_n^{\varpi+\eta}}|y| 1_{\{u_n<|x+y|\le\ u_n(1+\Delta_n^\eta)\}}.$$

If we apply this with $x = \Delta_i^n \widehat{X}'$ and $y = \Delta_i^n(B + X^c)$, plus (12.34) for p large enough and Bienaymé-Tchebycheff inequality and $1-2\varpi-2\eta > 0$, plus (12.40) and (12.34) again and Hölder's inequality, we obtain that the right side of (12.53) is smaller than $K\Delta_n^{1-\varpi\beta+\rho}$. Therefore it remains to prove that

$$\Big|\mathbb{E}_{i-1}^n(\Delta_i^n M\, f_n(\Delta_i^n \widehat{X}'))\Big| \;\le\; K\Delta_n^{1-\varpi\beta+\rho)} + K\Delta_n^{1-(\varpi+\eta)\beta}\,\mathbb{E}_{i-1}^n(|\Delta_i^n M|). \tag{12.54}$$

2) For simplicity we write $Y_t = \widehat{X}'_{(i-1)\Delta_n+t} - \widehat{X}'_{(i-1)\Delta_n}$ and $Z_r = M_{(i-1)\Delta_n+t} - M_{(i-1)\Delta_n}$. Since Z is a bounded continuous martingale and Y a semimartingale with vanishing continuous martingale part, and $f_n(Y)$ is bounded, we deduce from Itô's formula that the product $Z_t f_n(Y_t)$ is the sum of a martingale plus the process $\int_0^t \Gamma_u^n du$, where

$$\Gamma_t^n \;=\; Z_t \int F_{(i-1)\Delta_n+t}(dx)\, g_n(Y_t, x), \quad g_n(y,x) = f_n(y+x) - f_n(y) - f_n'(y)x1_{\{|x|\le 1\}}.$$

An easy computation allows to deduce of (12.52) that

$$|g_n(y,x)| \le 1_{\{|x|>u_n\Delta_n^\eta/3\}} + K1_{\{u_n<|y|\le u_n(1+\Delta_n^\eta)\}}\Big(\frac{x^2}{\Delta_n^{2\varpi+2\eta}}\,1_{\{|x|\le u_n\Delta_n^\eta\}} + \frac{|x|\wedge 1}{\Delta_n^{\varpi+\eta}}\,1_{\{|x|>u_n\Delta_n^\eta\}}\Big).$$

Now, we apply (12.26) to get for any $\varepsilon > 0$:

$$|\Gamma_t^n| \le K\,|Z_t|\,\Delta_n^{-(\varpi+\eta)\beta} + K_\varepsilon\,|Z_t|\,\Delta_n^{-(\beta+\varepsilon)(\varpi+\eta)}\,1_{\{u_n<|Y_t|\le u_n(1+\Delta_n^\eta)\}}.$$

Since $\eta < 1/2 - \varpi$ we have $\beta(\varpi + \eta) < 1$ and thus $(\beta + \varepsilon)(\varpi + \eta) = 1$ for a suitable $\varepsilon > 0$. Moreover $\mathbb{E}(|Z_u|) \le \mathbb{E}(|Z_s|)$ if $u \le s$ because Z is a martingale. Therefore, since Z is bounded we obtain

$$\begin{aligned}\Big|\mathbb{E}_{i-1}^n(\Delta_i^n M\, f_n(\Delta_i^n \widehat{X}'))\Big| \;&=\; \Big|\mathbb{E}_{i-1}^n\Big(\int_0^{\Delta_n}\Gamma_t^n dt\Big)\Big| \;\le\; \int_0^{\Delta_n}\mathbb{E}(|\Gamma_t^n|)\,dt\\ \le\; K\Delta_n^{1-(\varpi+\eta)\beta}\,\mathbb{E}_{i-1}^n(|\Delta_i^n M|) &+ K\Delta_n^{-1}\int_0^{\Delta_n}\mathbb{P}_{i-1}^n(u_n<|Y_t|\le u_n(1+\Delta_n^\eta))\,dt.\end{aligned}$$

By (12.40) for the process $\widehat{X}'$ instead of X, we readily deduce (12.54). □

12.5 Some auxiliary limit theorems.

Below, recall the process $\overline{A}$ of (12.18). We still assume (SL) and write $u_n = \alpha\Delta_n^\varpi$.

Lemma 12.14 *Let $\rho' < \frac{1}{2} \wedge (\varpi\gamma) \wedge (\varpi(\beta-\beta'))$. Then for all $t > 0$ we have*

$$\Delta_n^{-\rho'}\,\Big(\sum_{i=1}^{[t/\Delta_n]}\Delta_n^{\varpi\beta}\,\mathbb{E}_{i-1}^n\Big(\int_{(i-1)\Delta_n}^{i\Delta_n}\overline{F}_t(u_n)dt\Big) - \frac{\overline{A}_t}{\alpha^\beta}\Big) \;\stackrel{\mathbb{P}}{\longrightarrow}\; 0. \qquad (12.55)$$

Proof. Let $\theta_i^n = \int_{(i-1)\Delta_n}^{i\Delta_n}\overline{F}_t(u_n)dt$ and $\eta_i^n = \int_{(i-1)\Delta_n}^{i\Delta_n}A_t dt$. We deduce from (12.26) that

$$\Big|\Delta_n^{\varpi\beta}\theta_i^n - \frac{1}{\alpha^\beta}\,\eta_i^n\Big| \;\le\; K\Delta_n^{1+\varpi(\beta-(\beta-\gamma)\vee\beta')} \;\le\; K(\Delta_n^{1+\varpi\gamma} + \Delta_n^{1+\varpi(\beta-\beta')}).$$

Then obviously

$$\mathbb{E}\Big(\Delta_n^{-\rho'}\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n\Big(\Big|\Delta_n^{\varpi\beta}\,\theta_i^n - \frac{1}{\alpha^\beta}\,\eta_i^n\Big|\Big)\Big) \;\to\; 0,$$

and since A_t is bounded we have $\Big|\overline{A}_t - \sum_{i=1}^{[t/\Delta_n]}\eta_i^n\Big| \le Kt\Delta_n$, whereas $\rho' < 1$. It thus remains to prove that

$$\Delta_n^{-\rho'}\,\Big(\sum_{i=1}^{[t/\Delta_n]}\big(\eta_i^n - \mathbb{E}_{i-1}^n(\eta_i^n)\big)\Big) \;\stackrel{\mathbb{P}}{\longrightarrow}\; 0. \qquad (12.56)$$

Since $\zeta_i^n = \Delta_n^{-\rho'}\big(\eta_i^n - \mathbb{E}_{i-1}^n(\eta_i^n)\big)$ is a martingale increment, for (12.56) it is enough to check that $a_n(t) = \mathbb{E}\Big(\sum_{i=1}^{[t/\Delta_n]}(\zeta_i^n)^2\Big)$ goes to 0. However, since A_t is bounded, we have $|\zeta_i^n|^2 \le K\Delta_n^{2-2\rho'}$, so $a_n(t) \le Kt\Delta_n^{1-2\rho'} \to 0$ because $\rho' < 1/2$. □

Lemma 12.15 *a) Let* $\chi < (\varpi\gamma) \wedge \frac{1-\varpi}{3} \wedge \frac{\varpi(\beta-\beta')}{1+\beta'} \wedge \frac{1-2\varpi}{2}$*. Then for all* $t > 0$ *we have*

$$\Delta_n^{-\chi}\Big(\Delta_n^{\varpi\beta}\sum_{i=1}^{[t/\Delta_n]}\mathbb{P}_{i-1}^n(|\Delta_i^n X| > u_n) - \frac{\overline{A}_t}{\alpha^\beta}\Big) \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{12.57}$$

and in particular

$$\Delta_n^{\varpi\beta}\sum_{i=1}^{[t/\Delta_n]}\mathbb{P}_{i-1}^n(|\Delta_i^n X| > u_n) \stackrel{\mathbb{P}}{\longrightarrow} \frac{\overline{A}_t}{\alpha^\beta} \tag{12.58}$$

b) If further $\beta' < \frac{\beta}{2+\beta}$ *and* $\gamma > \frac{\beta}{2}$ *and* $\varpi < \frac{1}{2+\beta}\bigwedge\frac{1}{3\beta}$*, and if* M *is a bounded continuous martingale, we also have*

$$\Delta_n^{-\varpi\beta/2}\Big(\Delta_n^{\varpi\beta}\sum_{i=1}^{[t/\Delta_n]}\mathbb{P}_{i-1}^n(|\Delta_i^n X| > u_n) - \frac{\overline{A}_s}{\alpha^\beta}\Big) \stackrel{\mathbb{P}}{\longrightarrow} 0. \tag{12.59}$$

$$\Delta_n^{\varpi\beta/2}\sum_{i=1}^{[t/\Delta_n]}\Big|\mathbb{E}_{i-1}^n(\Delta_i^n M 1_{\{|\Delta_i^n X|>u_n\}})\Big| \stackrel{\mathbb{P}}{\longrightarrow} 0. \tag{12.60}$$

Proof. a) In Lemma 12.12 we can take $\eta = \frac{1-\varpi\beta}{3} \wedge \frac{\varpi(\beta-\beta')}{1+\beta'} \wedge \frac{1-2\varpi-\varepsilon}{2}$ for some $\varepsilon > 0$, and ρ is given by (12.38). Upon taking ε small enough, we then have $\chi < \rho$, and also $\chi \le \rho'$ for a ρ' satisfying the conditions of Lemma 12.14. Then (12.57) readily follows from (12.39) and (12.55).

b) Our conditions on γ, β' and ϖ imply (after some calculations) that one may take $\chi = \varpi\beta/2$ satisfying the condition in (a), so (12.59) follows from (12.57).

It remains to prove (12.60). By (12.51), the left side of (12.60) is smaller than

$$K\,t\,\Delta_n^{\rho-\varpi\beta/2} + K\Delta_n^{1-\eta\beta-\varpi\beta/2}\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n(|\Delta_i^n M|).$$

By the Cauchy-Schwarz inequality this is smaller than

$$K(t+\sqrt{t}\,)\Big(\Delta_n^{\rho-\varpi\beta/2} + \Delta_n^{1/2-\eta\beta-\varpi\beta/2}\Big(\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n(|\Delta_i^n M|^2)\Big)^{1/2}\Big).$$

A well known property of martingales yields

$$\mathbb{E}\Big(\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n(|\Delta_i^n M|^2)\Big) = \mathbb{E}\left((M_{\Delta_n[t/\Delta_n]} - M_0)^2\right),$$

which is bounded (in n). Therefore we deduce (12.60), provided we have $\rho > \varpi\beta/2$ and also $1 - 2\eta\beta > \varpi\beta$. The first condition has already been checked, but the second one may fail with our previous choice of η. However since $\varpi < 1/3\beta$ we have $\varpi\beta/2 < (1-\varpi\beta)/2\beta$, and we can find η' strictly between these two numbers. Then we replace ρ and η by $\overline{\rho} = \rho \wedge \eta$ and $\overline{\eta} = \eta \wedge \eta'$, which still satisfy (12.38), and now the required conditions are fulfilled by $\overline{\rho}$ and $\overline{\eta}$. This ends the proof. □

Proposition 12.16 *Assume (SL). For each $t > 0$ we have*

$$\Delta_n^{\varpi\beta}\, U(\alpha\Delta_n^{\varpi},\Delta_n)_t \;\stackrel{\mathbb{P}}{\longrightarrow}\; \frac{\overline{A}_t}{\alpha^\beta}. \tag{12.61}$$

Proof. Set

$$\zeta_i^n \;=\; \Delta_n^{\varpi\beta/2}\Big(1_{\{|\Delta_i^n X|>\alpha\Delta_n^{\varpi}\}} - \mathbb{P}_{i-1}^n(|\Delta_i^n X| > \alpha\Delta_n^{\varpi})\Big)\,. \tag{12.62}$$

By virtue of (12.58), it suffices to prove that the sequence $\sum_{i=1}^{[t/\Delta_n]}\zeta_i^n$ is tight. Since the ζ_i^n's are martingale increments, it is enough to show that the sequence $a_n(t) = \sum_{i=1}^{[t/\Delta_n]}\mathbb{E}((\zeta_i^n)^2)$ is bounded. But (12.41) yields $\mathbb{E}((\zeta_i^n)^2) \le K\Delta_n$, which in turn yields $a_n(t) \le Kt$. □

Proposition 12.17 *Assume (SL). Let $\alpha' > \alpha$. If we have $\beta' < \frac{\beta}{2+\beta}$ and $\gamma > \frac{\beta}{2}$ and $\varpi < \frac{1}{2+\beta}\bigwedge\frac{1}{3\beta}$, the pair of processes*

$$\Delta_n^{-\varpi\beta/2}\left(\Delta_n^{\varpi\beta}\; U(\alpha\Delta_n^{\varpi},\Delta_n)_t - \frac{\overline{A}_t}{\alpha^\beta},\; \Delta_n^{\varpi\beta}\; U(\alpha'\Delta_n^{\varpi},\Delta_n)_t - \frac{\overline{A}_t}{\alpha'^\beta}\right) \tag{12.63}$$

converges stably in law to a process $(\overline{W},\overline{W}')$ defined on an extension of $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge0},\mathbb{P})$, and with conditionally on $\mathcal{F}$ is a continuous Gaussian martingale with

$$\widetilde{\mathbb{E}}(\overline{W}_t^2 \mid \mathcal{F}) \;=\; \frac{\overline{A}_t}{\alpha^\beta}, \qquad \widetilde{\mathbb{E}}(\overline{W}_t'^2 \mid \mathcal{F}) \;=\; \frac{\overline{A}_t}{\alpha'^\beta}, \qquad \widetilde{\mathbb{E}}(\overline{W}_t\,\overline{W}_t' \mid \mathcal{F}) \;=\; \frac{\overline{A}_t}{\alpha'^\beta}. \tag{12.64}$$

Proof. Define ζ_i^n by (12.62), and associate $\zeta_i'^n$ with α' in the same way. In view of (12.59) the result amounts to proving the stable convergence in law of the pair of processes $\Big(\sum_{i=1}^{[t/\Delta_n]}\zeta_i^n,\sum_{i=1}^{[t/\Delta_n]}\zeta_i'^n\Big)$ to $(\overline{W},\overline{W}')$. The variables $\zeta^n i$ and $\zeta_i'^n$ are martingale increments and are smaller than $K\Delta_n^{\varpi\beta/2}$, so in view of Lemma 4.4 it is enough to prove the following

$$\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n((\zeta_i^n)^2) \stackrel{\mathbb{P}}{\longrightarrow} \frac{\overline{A}_t}{\alpha^\beta}, \quad \sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n((\zeta_i'^n)^2) \stackrel{\mathbb{P}}{\longrightarrow} \frac{\overline{A}_t}{\alpha'^\beta}, \quad \sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n(\zeta_i^n\,\zeta_1'^n) \stackrel{\mathbb{P}}{\longrightarrow} \frac{\overline{A}_t}{\alpha'^\beta}. \tag{12.65}$$

$$\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n(\zeta_i^n\;\Delta_i^n M) \;\stackrel{\mathbb{P}}{\longrightarrow}\; 0, \qquad \sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n(\zeta_i'^n\;\Delta_i^n M) \;\stackrel{\mathbb{P}}{\longrightarrow}\; 0, \tag{12.66}$$

where M is any bounded martingale.

Since $\alpha < \alpha'$, we have

$$\mathbb{E}_{i-1}^n(\zeta_i^n\,\zeta_1'^n) = \Delta_n^{\varpi\beta}\Big(\mathbb{P}_{i-1}^n(|\Delta_i^n X| > \alpha'\Delta_n^{\varpi}) - \mathbb{P}_{i-1}^n(|\Delta_i^n X| > \alpha\Delta_n^{\varpi})\,\mathbb{P}_{i-1}^n(|\Delta_i^n X| > \alpha'\Delta_n^{\varpi})\Big),$$

whereas $\mathbb{P}_{i-1}^n(|\Delta_i^n X| > \alpha\Delta_n^{\varpi}) \le K\Delta_n^{1-\varpi\beta}$ by (12.41). Therefore we deduce the last part of (12.65) from (12.61), and the first two parts are proved in the same way.

Now we turn to (12.66). Since $\mathbb{E}_{i-1}^n(\Delta_i^n M) = 0$, this follows from (12.60), which has been proved when M is continuous. Now, since any bounded martingale is the sum of a continuous martingale and a purely discontinuous martingale with bounded jumps, and up to a localization, it remains to prove (12.66) when M is a bounded purely discontinuous martingale.

In this case, we consider the discretized process $M_t^n = M_{\Delta_n[t/\Delta_n]}$, and we set $Z^n = \sum_{i=1}^{[t/\Delta_n]} \zeta_i^n$. We know by (12.65) that the sequence (of discrete-time martingales) Z^n is tight, whereas the convergence $M^n \to M$ (pathwise, in the Skorokhod sense) is a known fact. Since further any limiting process of Z^n is continuous, the pair (Z^n, M^n) is tight. From any subsequence of indices we pick a further subsequence, say (n_k), such that (Z^{n_k}, M^{n_k}) is tight, with the limit (Z, M). Another well known fact is that the quadratic covariation $[M^{n_k}, Z^{n_k}]$ converges to $[M, Z]$, and since M is purely discontinuous and Z is continuous it follows that $[M, Z] = 0$. Then by Lenglart inequality (since the jumps of the discrete processes $[M^n, Z^n]$ are bounded by a constant), the predictable compensators of $[M^n, Z^n]$ also go to 0 in probability. Now, those compensators are exactly $\sum_{i=1}^{[t/\Delta_n]} \mathbb{E}_{i-1}^n(\zeta_i^n \Delta_i^n M)$, which thus goes to 0 in probability along the subsequence n_k; it readily follows that the first part of (12.66) holds, and the second part is similarly analyzed. $\square$

12.6 Proof of Theorem 12.3.

At this point, the proof is nearly trivial. As said before, it is no restriction to assume (SL). Then in view of Proposition 12.16 the consistency result (a) is obvious.

As for (b), we apply Proposition 12.17, to obtain that

$$U(\alpha \Delta_n^{\varpi}, \Delta_n)_t = \frac{\overline{A}_t}{\Delta_n^{\varpi\beta} \alpha^\beta} + \Delta_n^{\varpi\beta/2} V_n, \qquad U(\alpha' \Delta_n^{\varpi}, \Delta_n)_t = \frac{\overline{A}_t}{\Delta_n^{\varpi\beta} \alpha'^\beta} + \Delta_n^{\varpi\beta/2} V_n',$$

where the pair (V_n, V_n') converge stably in law to a variable (V, V') which is $\mathcal{F}$-conditionally Gaussian centered with covariance matrix $\begin{pmatrix} \overline{A}_t/\alpha^\beta & \overline{A}_t/\alpha'^\beta \\ \overline{A}_t/\alpha'^\beta & \overline{A}_t/\alpha'^\beta \end{pmatrix}$. Then a simple computation shows that the variable $\frac{1}{\Delta_n^{\varpi\beta/2}} \left(\widehat{\beta}_n(t, \varpi, \alpha, \alpha') - \beta \right)$ is equivalent (in probability) to

$$\frac{\alpha^\beta V_n - \alpha'^\beta V_n'}{\overline{A}_t \log(\alpha'/\alpha)},$$

on the set $\{\overline{A}_t > 0\}$. The result readily follows . $\square$

References

[1] Aït Sahalia Y. (2004): Disentangling diffusion from jumps. *J. Financial Econometrics* **74**, 487-528.

[2] Aït Sahalia Y. and J. Jacod (2008): Fisher's information for discretely sampled Lévy processes. *Econometrica*, **76**, 727-761.

[3] Aït Sahalia Y. and J. Jacod (2009): Testing for jumps in a discretely observed process. *Annals of Statistics*, **37**, 1, 184-222.

[4] Aldous, D., and G. K. Eagleson (1978): On mixing and stability of limit theorems. *Ann. Probab.***6**, 325-331.

[5] Andersen, T. G., T. Bollerslev, and F. X. Diebold (2005): Parametric and nonparametric measurement of volatility. In Y. Aït-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland. Forthcoming.

[6] Barndorff-Nielsen, O. E. and N. Shephard (2003): Realised power variation and stochastic volatility. *Bernoulli* **9**, 243–265. Correction published in pages 1109–1111.

[7] Barndorff-Nielsen, O. E. and N. Shephard (2004): Power and bipower variation with stochastic volatility and jumps. *J. Financial Econometrics* **2**, 1–48.

[8] Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, M. Podolskij and N. Shephard (2006): A central limit theorem for realised bipower variations of continuous semimartingales. In *From Stochastic Calculus to Mathematical Finance*, the Shiryaev Festschrift, Yu. Kabanov, R. Liptser, J. Stoyanov eds, pp. 33-69, Springer Verlag: Berlin.

[9] Barndorff-Nielsen, O. E., N. Shephard, and M. Winkel (2006): Limit theorems for multipower variation in the presence of jumps. *Stochastic Processes and Appl.* **116**, 796-826.

[10] Billingsley, P. (1968): *Convergence of Probability Measures*, Wiley and sons: New York.

[11] Carr, P. and L. Wu (2003): What type of process underlies options ? A simple robust test. t . Finance **58**, 2581-2610.

[12] Huang, X. and G. Tauchen (2006): The relative contribution of jumps to total price variance. *J. Financial Econometrics* **4**, 456-499.

[13] Jacod, J. and A. Shiryaev (2003): *Limit Theorems for Stochastic Processes*, 2d ed., Springer-Verlag: Berlin.

[14] Jacod, J. and Protter, P. (1998): Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.*, **26**, 267-307.

[15] Jacod, J. (2008): Asymptotic properties of realized power variations and related functionals of semimartingales. *Stoch. Proc. Appl.*, **118**, 517-559.

[16] Jiang G.J. and R.C. Oomen (2008): A new test for jumps in asset prices. *Journal of Econometrics* **144**, 352-370.

[17] Lee, S. and P. Mykland (2008): Jumps in financial markets: a new nonparametric test and jump clustering. *Review of Financial Studies* **21**, 2535-2563.

[18] Lépingle, D. (1976): La variation d'ordre p des semimartingales. *Z. für Wahr. Th.*, **36**, 285–316.

[19] Mancini, C. (2001): Disentangling the jumps of the diffusion in a geometric jumping Brownian motion. *Giornale dell'Instituto Italiano degli Attuari* **LXIV** 19–47.

[20] Mancini, C. (2008): Non parametric threshold Estimation for models with stochastic diffusion coefficient and jumps. To appear in *Scandinavian J. of Statistics*.

[21] Mykland, P.A. and L. Zhang (2006): ANOVA for Diffusions and Itô Processes, *The Annals of Statistics*, **34**, 1931-1963.

[22] Rényi A. (1963): On Stable Sequences of Events, *Sankyā*, Series A, **25** 293-302.

[23] Tukey J.W. (1939): On the distribution of the fractional part of a statistical variable. *Math. Sb.*, **4** 561-562.

[24] Woerner,J. (2005): Estimation of integrated volatility in stochastic volatility models. *Appl. Stochastic Models Bus. Ind.*, **21**, 27-44.

[25] Woerner,J. (2006): Power and multipower variation: inference for high frequency data. In *Stochastic Finance*, A.N. Shiryaev, M. do Rosário Grosshino, P. Oliviera, M. Esquivel, eds., 343-354, Springer.

[26] Zolotarev, V.M. (1986): *One-dimensional stable distributions*, AMS, Translations of Mathematical Monographs **65**: Providence.