



SCHOOL OF ECONOMICS AND MANAGEMENT
FACULTY OF SOCIAL SCIENCES
AARHUS UNIVERSITY



CREATES
Center for Research in Econometric
Analysis of Time Series

CREATES Research Paper 2009-47

Understanding limit theorems for semimartingales: a short survey

Mark Podolskij and Mathias Vetter

School of Economics and Management
Aarhus University
Bartholins Allé 10, Building 1322, DK-8000 Aarhus C
Denmark

Understanding limit theorems for semimartingales: a short survey ^{*}

Mark Podolskij
ETH Zürich and CREATES [†]

Mathias Vetter
Ruhr-University of Bochum [‡]

October 5, 2009

Abstract

This paper presents a short survey on limit theorems for certain functionals of semimartingales, which are observed at high frequency. Our aim is to explain the main ideas of the theory to a broader audience. We introduce the concept of stable convergence, which is crucial for our purpose. We show some laws of large numbers (for the continuous and the discontinuous case) that are the most interesting from a practical point of view, and demonstrate the associated stable central limit theorems. Moreover, we state a simple sketch of the proofs and give some examples.

Keywords: central limit theorem, high frequency observations, semimartingale, stable convergence.

JEL Classification: C10, C13, C14.

1 Introduction

In the last decade there has been a considerable development of the asymptotic theory for processes observed at a high frequency. This was mainly motivated by financial applications, where the data, such as stock prices or currencies, are observed very frequently. As under the no-arbitrage assumptions price processes must follow a semimartingale (see e.g. [7]), there was a need for probabilistic tools for functionals of semimartingales based on high frequency observations.

Inspired by potential applications, probabilists started to develop limit theorems for semimartingales. An important starting point was the unpublished work of Jacod [10], who

^{*}Mark Podolskij gratefully acknowledges financial support from CREATES funded by the Danish National Research Foundation.

[†]Department of Mathematics, ETH Zürich, HG G32.2, 8092 Zürich, Switzerland, Email: mark.podolskij@math.ethz.ch.

[‡]Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany, Email: mathias.vetter@rub.de

developed a first general (stable) central limit theorem for high frequency observations; the crucial part of this work was later published in [11] (see also Chapter IX in [14] for a detailed study of the asymptotic results). Later on those results were used to derive limit theorems for various functionals of semimartingales; we refer to [3], [4], [12], [13], [15] among many others. Statisticians applied the asymptotic theory to analyze the path properties of discretely observed semimartingales: for the estimation of certain volatility functionals and realised jumps (see e.g. Theorem 3.1, Example 3.2 and Theorem 3.6 of this paper or [4], [17]), or for performing various test procedures (see e.g. [1], [5], [8]).

The aim of this paper is to present a short survey of these theoretical results and to carefully explain the main concepts and ideas of the proofs. We remark that the formal proofs of various limit theorems are usually long and pretty complicated; however, we try to give the reader a simple and clear intuition behind the theory, making those limit theorems more accessible for non-specialists in the field of semimartingales and stochastic processes.

Throughout this paper we are in a framework of a one-dimensional Itô semimartingale, i.e.

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s + \delta 1_{\{|\delta| \leq 1\}} \star (\mu - \nu)_t + \delta 1_{\{|\delta| > 1\}} \star \mu_t, \quad (1.1)$$

defined on the filtered probability space $(\Omega, \mathbb{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. In (1.1) $(a_s)_{s \geq 0}$ is a stochastic drift process, $(\sigma_s)_{s \geq 0}$ is a stochastic volatility, W denotes a standard Brownian motion, δ is a predictable function, μ a Poisson random measure and ν its predictable compensator (the precise definition of μ , ν and $f \star \mu$ will be given later). The last two summands of (1.1) stand for the (compensated) small jumps and the large jumps, respectively.

Typically, the stochastic process X is observed at high frequency, i.e. the data points $X_{i\Delta_n}$, $i = 0, \dots, [t/\Delta_n]$ are given, and we are in the framework of *infill asymptotics*, that is $\Delta_n \rightarrow 0$. When X is a *continuous* process (i.e. the last two terms of (1.1) are 0 identically) we are interested in the behaviour of the functionals

$$V(f)_t^n = \Delta_n \sum_{i=1}^{[t/\Delta_n]} f\left(\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right), \quad t > 0, \quad (1.2)$$

where $\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$ and $f : \mathbb{R} \rightarrow \mathbb{R}$ is a smooth function. The scaling $\Delta_n^{-1/2}$ in the argument is explained by the selfsimilarity of the Brownian motion W .

When the process X contains jumps it is more appropriate to consider functionals of the type

$$\bar{V}(f)_t^n = \sum_{i=1}^{[t/\Delta_n]} f(\Delta_i^n X). \quad (1.3)$$

In contrast to $V(f)_t^n$, the asymptotic theory for $\bar{V}(f)_t^n$ crucially depends on the behaviour of the function f near 0. When $f(x) \sim x^p$ at 0 we observe the following: if $p > 2$ the limit of $\bar{V}(f)_t^n$ is driven by the jump part of X , if $0 < p < 2$ the limit of the normalized version of $\bar{V}(f)_t^n$ is driven by the continuous part of X , and if $p = 2$ both parts contribute to the limit. Finally, we remark that almost all high frequency statistics used for practical

applications are of the form (1.2), (1.3) or of related type (the two most well-known generalizations are *multipower variation* (see e.g. [6]), *truncated power variation* (see e.g. [17]) or combinations thereof (see e.g. [20])). Thus, it is absolutely crucial to understand the asymptotic theory for the functionals $V(f)_t^n$ and $\bar{V}(f)_t^n$. We will derive the law of large numbers for $V(f)_t^n$ and $\bar{V}(f)_t^n$, and prove the associated *stable* central limit theorems.

This paper is organized as follows: in Section 2 we introduce the concept of stable convergence and demonstrate Jacod's central limit theorem for semimartingales. We explain the intuition behind Jacod's theorem and give some examples to illustrate its application. Section 3 is devoted to the asymptotic results for functionals $V(f)_t^n$ and $\bar{V}(f)_t^n$. We state the theoretical results and present an intuitive (and rather precise) sketch of the proofs.

2 The mathematical background

We start this section by introducing the notion of *stable convergence* of random variables (or processes). As we will see in Section 3, we typically deal with *mixed normal* limits in the framework of semimartingales. More precisely, we have that $Y_n \xrightarrow{d} VU$, where $V > 0$, $U \sim N(0, 1)$ and the random variables V and U are independent (we write $Y_n \xrightarrow{d} MN(0, V^2)$), and the latter is called a mixed normal distribution with random variance V^2). Usually, the distribution of V is unknown and thus the weak convergence $Y_n \xrightarrow{d} MN(0, V^2)$ is useless for statistical purposes, since confidence intervals are unavailable. The problem can be explained as follows: as for the case of a normal distribution with *deterministic* variance V^2 , we would try to estimate V^2 , say by V_n^2 , and hope that

$$Y_n/V_n \xrightarrow{d} N(0, 1).$$

However, the weak convergence $Y_n \xrightarrow{d} VU$ does not imply $(Y_n, V_n) \xrightarrow{d} (VU, V)$ for a *random* variable V (which is required to conclude that $Y_n/V_n \xrightarrow{d} N(0, 1)$). For this reason we need a stronger mode of convergence that would imply the joint weak convergence of (Y_n, V) for any \mathbb{F} -measurable variable V .

Stable convergence is exactly the right type of convergence to guarantee this property. In the following subsection we give a formal definition of stable convergence and derive its most useful properties (in fact, all properties statisticians should know).

2.1 A crash course on stable convergence

In this subsection all random variables or processes are defined on some probability space $(\Omega, \mathbb{F}, \mathbb{P})$. We start with a definition of stable convergence.

Definition 2.1 *Let Y_n be a sequence of random variables with values in a Polish space (E, \mathcal{E}) . We say that Y_n converges stably with limit Y , written $Y_n \xrightarrow{st} Y$, where Y is defined on an extension $(\Omega', \mathbb{F}', \mathbb{P}')$, iff for any bounded, continuous function g and any bounded \mathbb{F} -measurable random variable Z it holds that*

$$\mathbb{E}(g(Y_n)Z) \rightarrow \mathbb{E}'(g(Y)Z) \quad (2.4)$$

as $n \rightarrow \infty$.

First of all, we remark that random variables Y_n in the above definition can be also random processes. We immediately see that stable convergence is a stronger mode of convergence than weak convergence, but weaker than convergence in probability.

For the sake of simplicity we will only deal with stable convergence of \mathbb{R}^d -valued random variables in this subsection (many of the results below transfer directly to stable convergence of processes). The next proposition gives a much simpler characterization of stable convergence which is closer to the original definition of Renyi [18] (see also [2]).

Proposition 2.2 *The following properties are equivalent:*

- (i) $Y_n \xrightarrow{st} Y$
- (ii) $(Y_n, Z) \xrightarrow{d} (Y, Z)$ for any \mathbb{F} -measurable variable Z
- (iii) $(Y_n, Z) \xrightarrow{st} (Y, Z)$ for any \mathbb{F} -measurable variable Z

The assertion of Proposition 2.2 is easily shown and we leave the details to the reader.

For the moment it is not quite clear why an extension of the original probability space $(\Omega, \mathbb{F}, \mathbb{P})$ in Definition 2.1 is required. The next lemma gives the answer.

Lemma 2.3 *Assume that $Y_n \xrightarrow{st} Y$ and Y is \mathbb{F} -measurable. Then*

$$Y_n \xrightarrow{\mathbb{P}} Y.$$

Proof: As $Y_n \xrightarrow{st} Y$ and Y is \mathbb{F} -measurable, we deduce by Proposition 2.2(ii) that $(Y_n, Y) \xrightarrow{d} (Y, Y)$. Hence, $Y_n - Y \xrightarrow{d} 0$, and $Y_n \xrightarrow{\mathbb{P}} Y$ readily follows. \square

Lemma 2.3 tells us that the extension of the original probability space is not required iff we have $Y_n \xrightarrow{\mathbb{P}} Y$. But if we have "real" stable convergence $Y_n \xrightarrow{st} Y$, what type of extension usually appears? A partial answer is given in the following example.

Example 2.4 Let $(X_i)_{i \geq 1}$ be a sequence of i.i.d. random variables with $\mathbb{E}X_1 = 0$ and $\mathbb{E}X_1^2 = 1$, defined on $(\Omega, \mathbb{F}, \mathbb{P})$. Assume that $\mathbb{F} = \sigma(X_1, X_2, \dots)$. Setting $Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$ we obtain that

$$Y_n \xrightarrow{d} Y \sim N(0, 1),$$

which is of course a well-known result. Is there a stable version of this weak convergence? The answer is yes. Let $Y \sim N(0, 1)$ be independent of \mathbb{F} (thus it has to be defined on an orthogonal extension of $(\Omega, \mathbb{F}, \mathbb{P})$!). Then, for any collection $t_1, \dots, t_k \in \mathbb{N}$, we deduce that

$$(Y_n, X_{t_1}, \dots, X_{t_k}) \xrightarrow{d} (Y, X_{t_1}, \dots, X_{t_k})$$

as Y_n is asymptotically independent of $(X_{t_1}, \dots, X_{t_k})$. Thus, $(Y_n, Z) \xrightarrow{d} (Y, Z)$ for any \mathbb{F} -measurable variable, which implies that $Y_n \xrightarrow{st} Y$.

In fact, the described situation is pretty typical. Usually, we only require a new standard normal variable, independent of \mathbb{F} , to define the limiting variable Y (the canonical extension is simply the product space). We will see later that, when dealing with processes, we typically require a new Brownian motion, independent of \mathbb{F} , to define the limiting process. However, more complicated extensions may appear (see e.g. Section 3.2). \square

The last proposition of this subsection gives the answer to our original question and presents the Δ -method for stable convergence, which is quite often used in statistical applications.

Proposition 2.5 *Let Y_n, V_n, Y, X, V be \mathbb{R}^d -valued, \mathbb{F} -measurable random variables and let $g : \mathbb{R}^d \rightarrow \mathbb{R}$ be a C^1 -function.*

- (i) *If $Y_n \xrightarrow{st} Y$ and $V_n \xrightarrow{\mathbb{P}} V$ then $(Y_n, V_n) \xrightarrow{st} (Y, V)$.*
- (ii) *Let $d = 1$ and $Y_n \xrightarrow{st} Y \sim MN(0, V^2)$ with V being \mathbb{F} -measurable. Assume that $V_n \xrightarrow{\mathbb{P}} V$ and $V_n, V > 0$. Then $\frac{Y_n}{V_n} \xrightarrow{d} N(0, 1)$ (and there is also a stable version of this convergence).*
- (iii) *Let $\sqrt{n}(Y_n - Y) \xrightarrow{st} X$. Then $\sqrt{n}(g(Y_n) - g(Y)) \xrightarrow{st} \nabla g(Y)X$.*

Proof: Assertion (i) is trivial, since $Y_n \xrightarrow{st} Y$ implies $(Y_n, V) \xrightarrow{d} (Y, V)$ and we have $V_n - V \xrightarrow{\mathbb{P}} 0$ by assumption. Part (ii) follows by part (i) and the continuous mapping theorem, since $(Y_n, V_n) \xrightarrow{st} (Y, V)$. Finally, let us show part (iii). Since $\sqrt{n}(Y_n - Y) \xrightarrow{st} X$ we have $|Y_n - Y| \xrightarrow{\mathbb{P}} 0$. The mean value theorem implies that

$$\sqrt{n}(g(Y_n) - g(Y)) = \sqrt{n}\nabla g(\xi_n)(Y_n - Y)$$

for some ξ_n with $|\xi_n - Y| \leq |Y_n - Y|$. Clearly, $\xi_n \xrightarrow{\mathbb{P}} Y$. Thus, by part (i) we obtain $(\xi_n, \sqrt{n}(Y_n - Y)) \xrightarrow{st} (Y, X)$, which implies part (iii) because ∇g is continuous. \square

The Δ -method presented in Proposition 2.5 again demonstrates the importance of stable convergence. We would like to emphasize that such a result does not hold for the usual weak convergence when Y is random, which is a typical situation in a semimartingale framework (see Section 3).

2.2 Jacod's stable central limit theorem

In practice it is a difficult task to prove stable convergence, especially for processes. As for weak convergence, it is sufficient to show stable convergence of the finite dimensional distributions and tightness. However, proving stable convergence of the finite dimensional distributions is by far not easy, because the structure of the σ -algebra \mathbb{F} can be rather complicated (note that the σ -algebra \mathbb{F} from Example 2.4 has a pretty simple form).

Jacod [11] has derived a general stable central limit theorem for partial sums of triangular arrays. Below we assume that all processes are defined on the filtered probability space $(\Omega, \mathbb{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. We consider functionals of the form

$$Y_t^n = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} X_{in}, \quad (2.5)$$

where the X_{in} 's are $\mathcal{F}_{i\Delta_n}$ -measurable and square integrable random variables. Moreover, we assume that X_{in} 's are "fully generated" by a Brownian motion W^1 . Recall that the functionals $V(f)_t^n$ and $\bar{V}(f)_t^n$ are of the type (2.5).

Before we present the main theorem of this subsection, we need to introduce some notations. Below, $([M, N]_s)_{s \geq 0}$ denotes the covariation process of two (one-dimensional) semimartingales $(M_s)_{s \geq 0}$ and $(N_s)_{s \geq 0}$. We write $V^n \xrightarrow{u.c.p.} V$ whenever $\sup_{t \in [0, T]} |V_t^n - V_t| \xrightarrow{\mathbb{P}} 0$.

Theorem 2.6 (Jacod's Theorem [11])

Assume there exist absolutely continuous processes F , G , and a continuous process B with finite variation such that the following conditions are satisfied:

$$\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}(X_{in} | \mathcal{F}_{(i-1)\Delta_n}) \xrightarrow{u.c.p.} B_t, \quad (2.6)$$

$$\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \left(\mathbb{E}(X_{in}^2 | \mathcal{F}_{(i-1)\Delta_n}) - \mathbb{E}^2(X_{in} | \mathcal{F}_{(i-1)\Delta_n}) \right) \xrightarrow{\mathbb{P}} F_t = \int_0^t (v_s^2 + w_s^2) ds, \quad (2.7)$$

$$\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}(X_{in} \Delta_i^n W | \mathcal{F}_{(i-1)\Delta_n}) \xrightarrow{\mathbb{P}} G_t = \int_0^t v_s ds, \quad (2.8)$$

$$\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}(X_{in}^2 1_{\{|X_{in}| > \varepsilon\}} | \mathcal{F}_{(i-1)\Delta_n}) \xrightarrow{\mathbb{P}} 0 \quad \forall \varepsilon > 0, \quad (2.9)$$

$$\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}(X_{in} \Delta_i^n N | \mathcal{F}_{(i-1)\Delta_n}) \xrightarrow{\mathbb{P}} 0, \quad (2.10)$$

where $(v_s)_{s \geq 0}$ and $(w_s)_{s \geq 0}$ are predictable processes and condition (2.10) holds for all bounded \mathcal{F}_t -martingales with $N_0 = 0$ and $[W, N] \equiv 0$. Then we obtain the stable convergence of processes:

$$Y_t^n \xrightarrow{st} Y_t = B_t + \int_0^t v_s dW_s + \int_0^t w_s dW'_s, \quad (2.11)$$

where W' is a Brownian motion defined on an extension of the original probability space $(\Omega, \mathbb{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and independent of the original σ -algebra \mathbb{F} .

¹Roughly speaking, this means that there is no martingale N with $[W, N] \equiv 0$ that has a substantial contribution to X_{in} (otherwise condition (2.10) of Theorem 2.6 would be violated). We also remark that the central limit theorem in [11] is formulated with respect to a reference continuous (local) martingale M , which is supposed to generate the X_{in} 's (and has to be chosen by the user). However, for continuous Itô semimartingale models we can always choose $M = W$.

Remark 2.7 To the best of our knowledge, Theorem 2.6 is the only (general) stable central limit theorem for the case of infill asymptotics! Another stable central limit theorem (for random variables) can be found in [9] (see Theorem 3.2 therein), but it requires a certain *nesting condition* for the sequence of filtrations, which is not satisfied by $\mathcal{F}_{i\Delta_n}$. This underlines the huge importance of Jacod's theorem.

Furthermore, Theorem 2.6 is optimal in the following sense: there are no extra conditions among (2.6) - (2.10) that guarantee the stability of the central limit theorem. Even weak convergence $Y^n \Rightarrow Y$ would not hold under less conditions. \square

Remark 2.8 First of all, Theorem 2.6 is a probabilistic result that has no statistical applications in general, because there is no way to access the distribution of Y . However, when $B \equiv 0$ and $v \equiv 0$, which is the case for the most interesting situations, things become different! We remark that, for any fixed $t > 0$,

$$\int_0^t w_s dW'_s \sim MN\left(0, \int_0^t w_s^2 ds\right),$$

since W' is independent of \mathbb{F} . Hence

$$\frac{Y_t^n}{\sqrt{\int_0^t w_s^2 ds}} \xrightarrow{d} N(0, 1),$$

and the convergence still holds true if we replace the denominator by a consistent estimator. The latter can be applied to obtain confidence bands or to solve other statistical problems. \square

Although the formal proof of Theorem 2.6 is rather complicated, it is worthwhile to explain the meaning of the conditions (2.6) - (2.10). First of all, we observe the decomposition

$$Y_t^n = \underbrace{\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \left(X_{in} - \mathbb{E}(X_{in} | \mathcal{F}_{(i-1)\Delta_n}) \right)}_{\text{martingale part}} + \underbrace{\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}(X_{in} | \mathcal{F}_{(i-1)\Delta_n})}_{\text{drift part}},$$

where the first summand is a $\mathcal{F}_{i\Delta_n}$ -martingale. By (2.6), $\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}(X_{in} | \mathcal{F}_{(i-1)\Delta_n}) \xrightarrow{\text{u.c.p.}} B_t$, and consequently it is sufficient to assume that Y_t^n is a $\mathcal{F}_{i\Delta_n}$ -martingale and to show that

$$Y_t^n \xrightarrow{st} Y_t = \int_0^t v_s dW_s + \int_0^t w_s dW'_s.$$

Next, we observe that (2.9) is a classical (conditional) Lindeberg condition that ensures that the limiting process Y_t has no jumps. Now, let us analyze the quadratic variation structure of Y_t^n . Setting $W_t^n = W_{\Delta_n \lfloor t/\Delta_n \rfloor}$ and $N_t^n = N_{\Delta_n \lfloor t/\Delta_n \rfloor}$ we deduce from conditions

(2.7), (2.8) and (2.10) that

$$\begin{aligned} [Y^n, Y^n]_t &\xrightarrow{\mathbb{P}} [Y, Y]_t = F_t = \int_0^t (v_s^2 + w_s^2) ds, \\ [Y^n, W^n]_t &\xrightarrow{\mathbb{P}} [Y, W]_t = G_t = \int_0^t v_s ds, \\ [Y^n, N^n]_t &\xrightarrow{\mathbb{P}} [Y, N]_t = 0, \end{aligned}$$

for some predictable processes $(v_s)_{s \geq 0}$ and $(w_s)_{s \geq 0}$. The second convergence suggests that the process $\int_0^t v_s dW_s$ must be a part of Y_t . But, since $[Y, N]_t = 0$ and $w \not\equiv 0$ in general, the continuous \mathcal{F}_t -martingales cannot fully explain the quadratic variation of Y , and thus another martingale, which lives on the extension of $(\Omega, \mathbb{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, is required in the representation of Y . But why must this term be of the form $\int_0^t w_s dW'_s$? The reason is the Dambis-Dubins-Schwarz theorem (see e.g. Theorem V.1.6 in [19]): conditions (2.7), (2.8) and (2.10) imply that, conditionally on \mathbb{F} , the quadratic variation of this martingale is absolutely continuous. Thus, it must be a time-changed Brownian motion; hence, it must be of the form $\int_0^t w_s dW'_s$.

Finally, let us present a simple but important example to illustrate how Theorem 2.6 is applied in practice.

Example 2.9 Let σ be a càdlàg, \mathcal{F}_t -adapted and bounded process and let $g, h : \mathbb{R} \rightarrow \mathbb{R}$ be continuous functions with h being of polynomial growth. Define

$$Y_t^n = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} X_{in}, \quad X_{in} = \Delta_n^{1/2} g(\sigma_{(i-1)\Delta_n}) \left(h\left(\frac{\Delta_i^n W}{\sqrt{\Delta_n}}\right) - \mathbb{E}h\left(\frac{\Delta_i^n W}{\sqrt{\Delta_n}}\right) \right). \quad (2.12)$$

Note that the X_{in} 's have a pretty simple structure, since $\Delta_i^n W$ is independent of $\mathcal{F}_{(i-1)\Delta_n}$, and thus of $\sigma_{(i-1)\Delta_n}$, and $\Delta_i^n W / \sqrt{\Delta_n} \sim N(0, 1)$. Now we need to check the conditions (2.6) - (2.10) of Theorem 2.6. As $\mathbb{E}(X_{in} | \mathcal{F}_{(i-1)\Delta_n}) = 0$ we can set $B \equiv 0$. A simple calculation shows that

$$F_t = a^2 \int_0^t g^2(\sigma_s) ds, \quad G_t = b \int_0^t g(\sigma_s) ds,$$

where $a^2 = \text{var}(h(U))$, $b = \mathbb{E}(h(U)U)$ and $U \sim N(0, 1)$. Thus, we can set

$$w_s = \sqrt{a^2 - b^2} g(\sigma_s), \quad v_s = b g(\sigma_s)$$

in (2.7) and (2.8). On the other hand, it holds that

$$\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}(X_{in}^2 1_{\{|X_{in}| > \varepsilon\}} | \mathcal{F}_{(i-1)\Delta_n}) \leq \varepsilon^{-2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}(X_{in}^4 | \mathcal{F}_{(i-1)\Delta_n}) \leq C \frac{\Delta_n}{\varepsilon^2}$$

for some $C > 0$, because σ is a bounded process. Hence, condition (2.9) holds. The key to prove (2.10) is the Itô-Clark representation theorem (see Proposition V.3.2 in [19]). It says that there exists a process η^n such that

$$h\left(\frac{\Delta_i^n W}{\sqrt{\Delta_n}}\right) - \mathbb{E}h\left(\frac{\Delta_i^n W}{\sqrt{\Delta_n}}\right) = \int_{(i-1)\Delta_n}^{i\Delta_n} \eta_s^n dW_s.$$

From the Itô isometry we deduce that

$$\begin{aligned}\mathbb{E}(X_{in}\Delta_i^n N|\mathcal{F}_{(i-1)\Delta_n}) &= \Delta_n^{1/2}g(\sigma_{(i-1)\Delta_n}) \mathbb{E}\left(\int_{(i-1)\Delta_n}^{i\Delta_n} \eta_s^n dW_s \int_{(i-1)\Delta_n}^{i\Delta_n} dN_s\right) \\ &= \Delta_n^{1/2}g(\sigma_{(i-1)\Delta_n}) \mathbb{E}\left(\int_{(i-1)\Delta_n}^{i\Delta_n} \eta_s^n d[W, N]_s\right) = 0\end{aligned}$$

as $[W, N] \equiv 0$. This implies (2.10) and we obtain that

$$Y_t^n \xrightarrow{st} Y_t = b \int_0^t g(\sigma_s) dW_s + \sqrt{a^2 - b^2} \int_0^t g(\sigma_s) dW'_s.$$

Furthermore, when h is an even function then $b = 0$ and we have

$$Y_t^n \xrightarrow{st} Y_t = a \int_0^t g(\sigma_s) dW'_s,$$

and the limiting process Y is mixed normal. \square

3 Asymptotic results

As we mentioned above we need to distinguish between the continuous and the discontinuous case to derive the asymptotic results for $V(f)_t^n$ and $\bar{V}(f)_t^n$. We start with the continuous case. Below, for any process V , we define $V_{t-} = \lim_{s \nearrow t} V_s$ and $\Delta V_t = V_t - V_{t-}$.

3.1 The continuous case

In this subsection we present the asymptotic results for the functional $V(f)_t^n$ for continuous Itô semimartingales X . More precisely, we consider a continuous semimartingale X of the form

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s, \quad (3.13)$$

where $(a_s)_{s \geq 0}$ is a càglàd process and $(\sigma_s)_{s \geq 0}$ is a càdlàg, adapted process.

We start with the law of large numbers for $V(f)_t^n$. For any function $f : \mathbb{R} \rightarrow \mathbb{R}$, we define

$$\rho_x(f) = \mathbb{E}f(xU), \quad (3.14)$$

for $x \in \mathbb{R}$ and $U \sim N(0, 1)$.

Theorem 3.1 *Assume that the function f is continuous and has polynomial growth. Then*

$$V(f)_t^n \xrightarrow{u.c.p.} V(f)_t = \int_0^t \rho_{\sigma_s}(f) ds. \quad (3.15)$$

We remark that the drift process $(a_s)_{s \geq 0}$ does not influence the limit $V(f)_t$; we will see later why. Next, we present Theorem 3.1 for an important subclass of $V(f)_t^n$.

Example 3.2 (Realised power variation)

The class of statistics $V(f)_t^n$ with $f(x) = |x|^p$ ($p > 0$) is called *realised power variation*. It has some important applications in high frequency econometrics; see e.g. [4]. For $f(x) = |x|^p$, Theorem 3.1 translates to

$$V(f)_t^n \xrightarrow{\text{u.c.p.}} V(f)_t = m_p \int_0^t |\sigma_s|^p ds$$

with $m_p = \mathbb{E}(|U|^p)$, $U \sim N(0, 1)$. For $f(x) = x^2$ we rediscover a well-known result

$$V(f)_t^n \xrightarrow{\text{u.c.p.}} [X, X]_t = \int_0^t \sigma_s^2 ds.$$

□

Now, let us give a sketch of the proof of Theorem 3.1.

- *From local boundedness to boundedness:* Our assumptions imply that the processes $(a_s)_{s \geq 0}$ and $(\sigma_{s-})_{s \geq 0}$ are locally bounded, i.e. there exists an increasing sequence of stopping times T_k with $T_k \xrightarrow{\text{a.s.}} \infty$ such that the stopped processes are bounded:

$$|a_s| + |\sigma_{s-}| \leq C_k, \quad \forall s \leq T_k$$

for all $k \geq 1$. Indeed, it is possible to assume w.l.o.g. that $(a_s)_{s \geq 0}$, $(\sigma_{s-})_{s \geq 0}$ are bounded, because Theorem 3.1 is *stable under stopping*. To illustrate these ideas set $a_s^{(k)} = a_s 1_{\{s \leq T_k\}}$, $\sigma_s^{(k)} = \sigma_s 1_{\{s \leq T_k\}}$. Note that the processes $a^{(k)}$, $\sigma^{(k)}$ are bounded for all $k \geq 1$. Associate $X^{(k)}$ with $a^{(k)}$, $\sigma^{(k)}$ by (3.13), $V^{(k)}(f)_t^n$ with $X^{(k)}$ by (1.2) and $V^{(k)}(f)_t$ with $\sigma^{(k)}$ by (3.15). Now, notice that

$$V^{(k)}(f)_t^n = V(f)_t^n, \quad V^{(k)}(f)_t = V(f)_t, \quad \forall t \leq T_k.$$

As $T_k \xrightarrow{\text{a.s.}} \infty$ it is sufficient to prove $V^{(k)}(f)_t^n \xrightarrow{\text{u.c.p.}} V^{(k)}(f)_t$ for each $k \geq 1$. Thus, we can assume w.l.o.g. that the process $(a_s)_{s \geq 0}$, $(\sigma_{s-})_{s \geq 0}$ are bounded. □

- *The crucial approximation:* First of all, observe that

$$\Delta_i^n X = \underbrace{\int_{(i-1)\Delta_n}^{i\Delta_n} a_s ds}_{=O_p(\Delta_n)} + \underbrace{\int_{(i-1)\Delta_n}^{i\Delta_n} \sigma_s dW_s}_{=O_p(\Delta_n^{1/2})},$$

where the second approximation follows by *Burkholder's inequality* (see e.g. Theorem IV.4.1 in [19]). Thus, the influence of the drift process $(a_s)_{s \geq 0}$ is negligible for the first order asymptotics. Indeed, we have

$$\frac{\Delta_i^n X}{\sqrt{\Delta_n}} \approx \alpha_i^n = \Delta_n^{-1/2} \sigma_{(i-1)\Delta_n} \Delta_i^n W, \quad (3.16)$$

which is the crucial approximation for proving all asymptotic results. Note that the α_i^n 's have a very simple structure: they are uncorrelated and $\alpha_i^n \sim MN(0, \sigma_{(i-1)\Delta_n}^2)$. As f is continuous and σ is càdlàg, it is relatively easy to show that

$$V(f)_t^n - \Delta_n \sum_{i=1}^{[t/\Delta_n]} f(\alpha_i^n) \xrightarrow{\text{u.c.p.}} 0. \quad (3.17)$$

On the other hand, it holds that

$$\Delta_n \sum_{i=1}^{[t/\Delta_n]} \mathbb{E}(f(\alpha_i^n) | \mathcal{F}_{(i-1)\Delta_n}) = \Delta_n \sum_{i=1}^{[t/\Delta_n]} \rho_{\sigma_{(i-1)\Delta_n}}(f) \xrightarrow{\text{u.c.p.}} V(f)_t$$

and $\Delta_n^2 \sum_{i=1}^{[t/\Delta_n]} \mathbb{E}(f^2(\alpha_i^n) | \mathcal{F}_{(i-1)\Delta_n}) \xrightarrow{\text{u.c.p.}} 0$. Hence

$$\Delta_n \sum_{i=1}^{[t/\Delta_n]} f(\alpha_i^n) \xrightarrow{\text{u.c.p.}} V(f)_t,$$

which implies $V(f)_t^n \xrightarrow{\text{u.c.p.}} V(f)_t$. \square

Now we turn our attention to the stable central limit theorem associated with Theorem 3.1. Here we require a stronger assumption on the volatility process σ to be able to deal with the approximation error induced by (3.16). More precisely, the process σ is assumed to be a continuous Itô semimartingale:

$$\sigma_t = \sigma_0 + \int_0^t \tilde{a}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \tilde{\tau}_s dV_s, \quad (3.18)$$

where the processes $(\tilde{a}_s)_{s \geq 0}$, $(\tilde{\sigma}_s)_{s \geq 0}$, $(\tilde{\tau}_s)_{s \geq 0}$ are càdlàg adapted and V is a Brownian motion independent of W .

In fact, the condition (3.18) is motivated by potential applications, as it is satisfied for many stochastic volatility models. Next, for any function $f : \mathbb{R} \rightarrow \mathbb{R}$ and $k \in \mathbb{N}$, we define

$$\rho_x(f, k) = \mathbb{E}(f(xU)U^k), \quad U \sim N(0, 1). \quad (3.19)$$

Note that $\rho_x(f) = \rho_x(f, 0)$.

Theorem 3.3 *Assume that $f \in C^1(\mathbb{R})$ with f, f' having polynomial growth and that condition (3.18) is satisfied. Then the stable convergence of processes*

$$\Delta_n^{-1/2} \left(V(f)_t^n - V(f)_t \right) \xrightarrow{st} L(f)_t = \int_0^t b_s ds + \int_0^t v_s dW_s + \int_0^t w_s dW'_s, \quad (3.20)$$

holds, where

$$\begin{aligned} b_s &= a_s \rho_{\sigma_s}(f') + \frac{1}{2} \tilde{\sigma}_s (\rho_{\sigma_s}(f', 2) - \rho_{\sigma_s}(f')), \\ v_s &= \rho_{\sigma_s}(f, 1), \\ w_s &= \sqrt{\rho_{\sigma_s}(f^2) - \rho_{\sigma_s}^2(f) - \rho_{\sigma_s}^2(f, 1)} \end{aligned}$$

and W' is a Brownian motion defined on an extension of the original probability space $(\Omega, \mathbb{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and independent of the original σ -algebra \mathbb{F} .

As a consequence of Theorem 3.3 we obtain a simple but very important lemma.

Lemma 3.4 *Assume that $f : \mathbb{R} \rightarrow \mathbb{R}$ is an even function and that the conditions of Theorem 3.3 hold. Then $\rho_x(f') = \rho_x(f', 2) = \rho_x(f, 1) = 0$, and we deduce that*

$$\Delta_n^{-1/2} \left(V(f)_t^n - V(f)_t \right) \xrightarrow{st} L(f)_t = \int_0^t w_s dW'_s$$

with $w_s = \sqrt{\rho_{\sigma_s}(f^2) - \rho_{\sigma_s}^2(f)}$.

As we mentioned in Remark 2.8, $L(f)_t$ has obviously a mixed normal distribution (for any $t > 0$) when f is an even function. Indeed, this is the case for almost all statistics used in practice. Let us now return to Example 3.2.

Example 3.5 (Realised power variation)

We consider again the class of functions $f(x) = |x|^p$ ($p > 0$), which are obviously even. By Lemma 3.4 we deduce that

$$\Delta_n^{-1/2} \left(V(f)_t^n - m_p \int_0^t |\sigma_s|^p \right) \xrightarrow{st} L(f)_t = \sqrt{m_{2p} - m_p^2} \int_0^t |\sigma_s|^p dW'_s. \quad (3.21)$$

(In fact, the above convergence can be deduced from Lemma 3.4 only for $p > 1$, since otherwise $f(x) = |x|^p$ is not differentiable at 0. However, it is possible to extend the theory to the case $0 < p \leq 1$ under a further condition on σ ; see [3]). By Theorem 3.1 and Proposition 2.5 we are able to derive a *feasible* version of Lemma 3.4 associated with $f(x) = |x|^p$:

$$\frac{\Delta_n^{-1/2} \left(V(f)_t^n - m_p \int_0^t |\sigma_s|^p \right)}{\sqrt{\frac{m_{2p} - m_p^2}{m_{2p}} V(f^2)_t^n}} \xrightarrow{d} N(0, 1),$$

which can be used for statistical purposes. For the case of quadratic variation, i.e. $f(x) = x^2$, this translates to

$$\frac{\Delta_n^{-1/2} \left(\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} |\Delta_i^n X|^2 - \int_0^t \sigma_s^2 \right)}{\sqrt{\frac{2}{3} \Delta_n^{-1} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} |\Delta_i^n X|^4}} \xrightarrow{d} N(0, 1).$$

Quite surprisingly, the stable convergence for the case of quadratic variation can be proved without assuming the condition (3.18) (thus under very weak assumptions on the process X); this is not possible anymore for other powers p . \square

We present the main ideas behind the proof of Theorem 3.3, which ends this subsection.

- *CLT for the approximation (3.16)*: First of all, we observe that Theorem 3.3 is also

stable under stopping. Thus, we can assume w.l.o.g. that the processes $(a_s)_{s \geq 0}$, $(\sigma_s)_{s \geq 0}$, $(\tilde{a}_s)_{s \geq 0}$, $(\tilde{\sigma}_s)_{s \geq 0}$, $(\tilde{\tau}_s)_{s \geq 0}$ are bounded. In a first step, we show the central limit theorem for the approximation α_i^n . More precisely, we want to prove that

$$\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} X_{in} \xrightarrow{st} \int_0^t v_s dW_s + \int_0^t w_s dW'_s, \quad X_{in} = \Delta_n^{1/2} \left(f(\alpha_i^n) - \mathbb{E}(f(\alpha_i^n) | \mathcal{F}_{(i-1)\Delta_n}) \right),$$

where the process $(v_s)_{s \geq 0}$ and $(w_s)_{s \geq 0}$ are defined in Theorem 3.3. In principle, we can follow the ideas of Example 2.9: we immediately deduce the convergence

$$\begin{aligned} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}(X_{in}^2 | \mathcal{F}_{(i-1)\Delta_n}) &\xrightarrow{\mathbb{P}} F_t = \int_0^t (\rho_{\sigma_s}(f^2) - \rho_{\sigma_s}^2(f)) ds, \\ \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}(X_{in} \Delta_i^n W | \mathcal{F}_{(i-1)\Delta_n}) &\xrightarrow{\mathbb{P}} G_t = \int_0^t \rho_{\sigma_s}(f, 1) ds. \end{aligned}$$

On the other hand, conditions (2.6) with $B \equiv 0$, (2.9) and (2.10) of Theorem 2.6 are shown as in Example 2.9 (in fact, the proof of (2.10) is a bit more complicated here). Consequently, we deduce that $\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} X_{in} \xrightarrow{st} \int_0^t v_s dW_s + \int_0^t w_s dW'_s$. \square

• *CLT for the canonical process:* Before we proceed with the proof of Theorem 3.3 we need to present a further intermediate step. In fact, it is much more natural to consider a central limit theorem for the "canonical process"

$$L(f)_t^n = \Delta_n^{1/2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \left\{ f\left(\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right) - \mathbb{E}\left(f\left(\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right) \middle| \mathcal{F}_{(i-1)\Delta_n}\right) \right\}$$

since the latter is a martingale. Since f is continuous and σ is càdlàg, it is easy to see that

$$L(f)_t^n - \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} X_{in} \xrightarrow{\text{u.c.p.}} 0,$$

where the X_{in} 's are defined as in the previous step, because the above expression is a sum of martingale differences whose quadratic variation is shown to converge to 0 in probability as in (3.17). Hence, $L(f)_t^n \xrightarrow{st} \int_0^t v_s dW_s + \int_0^t w_s dW'_s$. \square

• *Putting things together:* Now, we are left to proving

$$\Delta_n^{-1/2} \left(V(f)_t^n - V(f)_t \right) - L(f)_t^n \xrightarrow{\text{u.c.p.}} \int_0^t b_s ds,$$

where the process $(b_s)_{s \geq 0}$ is given in Theorem 3.3. In view of the previous step, it is sufficient to show that

$$\Delta_n^{-1/2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \int_{(i-1)\Delta_n}^{i\Delta_n} (\rho_{\sigma_s}(f) - \rho_{\sigma_{(i-1)\Delta_n}}(f)) ds \xrightarrow{\text{u.c.p.}} 0, \quad (3.22)$$

$$\Delta_n^{1/2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}\left(f\left(\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right) - f(\alpha_i^n) \middle| \mathcal{F}_{(i-1)\Delta_n}\right) \xrightarrow{\text{u.c.p.}} \int_0^t b_s ds. \quad (3.23)$$

We remark that $\rho_{\sigma_s}(f) - \rho_{\sigma_{(i-1)\Delta_n}}(f) \approx \rho'_{\sigma_{(i-1)\Delta_n}}(f)(\sigma_s - \sigma_{(i-1)\Delta_n})$. By assumption (3.18) the left-hand side of (3.22) becomes asymptotically equivalent to a sum of martingale differences and the convergence in (3.22) readily follows. Finally, let us highlight the proof of (3.23) which is the crucial step. Assume for simplicity that

$$\sigma_t = \int_0^t \tilde{\sigma}_s dW_s$$

instead of (3.18), as the other components in (3.18) do not contribute to the limit process. In the following we write $Y^n \asymp X^n$ whenever $Y^n - X^n \xrightarrow{\text{u.c.p.}} 0$. The most important idea in the whole proof is the following approximation step

$$\begin{aligned} & \Delta_n^{1/2} \sum_{i=1}^{[t/\Delta_n]} \mathbb{E} \left(f \left(\frac{\Delta_i^n X}{\sqrt{\Delta_n}} \right) - f(\alpha_i^n) \middle| \mathcal{F}_{(i-1)\Delta_n} \right) \\ & \asymp \Delta_n^{1/2} \sum_{i=1}^{[t/\Delta_n]} \mathbb{E} \left(f'(\alpha_i^n) \left(\frac{\Delta_i^n X}{\sqrt{\Delta_n}} - \alpha_i^n \right) \middle| \mathcal{F}_{(i-1)\Delta_n} \right) \\ & \asymp \sum_{i=1}^{[t/\Delta_n]} \mathbb{E} \left(f'(\alpha_i^n) \left(\Delta_n a_{(i-1)\Delta_n} + \int_{(i-1)\Delta_n}^{i\Delta_n} (\sigma_s - \sigma_{(i-1)\Delta_n}) dW_s \right) \middle| \mathcal{F}_{(i-1)\Delta_n} \right) \\ & \asymp \sum_{i=1}^{[t/\Delta_n]} \mathbb{E} \left(f'(\alpha_i^n) \left(\Delta_n a_{(i-1)\Delta_n} + \tilde{\sigma}_{(i-1)\Delta_n} \int_{(i-1)\Delta_n}^{i\Delta_n} (W_s - W_{(i-1)\Delta_n}) dW_s \right) \middle| \mathcal{F}_{(i-1)\Delta_n} \right). \end{aligned}$$

By an application of Itô's formula and Riemann integrability we obtain

$$\begin{aligned} & \sum_{i=1}^{[t/\Delta_n]} \mathbb{E} \left(f'(\alpha_i^n) \left(\Delta_n a_{(i-1)\Delta_n} + \tilde{\sigma}_{(i-1)\Delta_n} \int_{(i-1)\Delta_n}^{i\Delta_n} (W_s - W_{(i-1)\Delta_n}) dW_s \right) \middle| \mathcal{F}_{(i-1)\Delta_n} \right) \\ & \xrightarrow{\text{u.c.p.}} \int_0^t b_s ds, \end{aligned}$$

which completes the proof of Theorem 3.3. \square

3.2 The discontinuous case

This subsection is devoted to the analysis of $\bar{V}(f)_t^n$ in the framework of an Itô semimartingale exhibiting jumps, and we start with a discussion of the representation

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s + \delta 1_{\{|\delta| \leq 1\}} \star (\mu - \nu)_t + \delta 1_{\{|\delta| > 1\}} \star \mu_t$$

from (1.1). Again, $(a_s)_{s \geq 0}$ is càglàd and $(\sigma_s)_{s \geq 0}$ is càdlàg adapted.

Regarding the latter two terms, recall that for some optional function $W(\omega, s, x)$ and some random measure κ on $\mathbb{R}_+ \times \mathbb{R}$ the notation $W \star \kappa_t$ is an abbreviation for the stochastic integral process

$$W \star \kappa_t(\omega) = \int_{[0, t] \times \mathbb{R}} W(\omega, s, x) \kappa(\omega; ds, dx),$$

as long as it exists. These processes are typically used to represent the jump part of a semimartingale, since $x \star \mu^X$ with $(\varepsilon$ is the Dirac measure)

$$\mu^X(\omega; dt, dx) = \sum_s 1_{\{\Delta X_s(\omega) \neq 0\}} \varepsilon_{(s, \Delta X_s(\omega))}(dt, dx)$$

is the sum of the jumps of X . In general, those jumps must not be summable, and thus compensating the small jumps (X is càdlàg, so there are only finitely many jumps larger than any given η) with ν^X becomes necessary. This random measure is the unique predictable one such that $W \star (\mu^X - \nu^X)_t$ is a local martingale for all optional W . Assume for example that we are given a Poisson process N_t with parameter λ : In this case, the compensator becomes $\nu^N(\omega; dt, dx) = \lambda dt \otimes \varepsilon_1(dx)$, and $x \star (\mu^N - \nu^N)$ takes the well-known form $N_t - \lambda t$.

For technical reasons we use a slightly different approach, as for Itô semimartingales it is always possible to choose μ as the specific Poisson random measure, whose compensator is given by $\nu(\omega; ds, dx) = ds \otimes dx$. This happens at the cost of a change in the integrator: x is replaced by some predictable function δ on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$.

Throughout this section we restrict ourselves to the two choices of f , which are the most interesting for applications, namely power variations with the respective cases $p > 2$ and $p = 2$. The same result for arbitrary semimartingales is proven in Lepingle [16].

Theorem 3.6 *Let $f(x) = |x|^p$ for a non-negative exponent p . For any $t \geq 0$ we have*

$$\overline{V}(f)_t^n \xrightarrow{\mathbb{P}} \overline{V}(f)_t = \begin{cases} \sum_{s \leq t} |\Delta X_s|^p, & p > 2, \\ [X, X]_t, & p = 2. \end{cases} \quad (3.24)$$

Remark 3.7 Recall that

$$[X, X]_t = \int_0^t \sigma_s^2 ds + \sum_{s \leq t} |\Delta X_s|^2$$

is almost surely finite for any (Itô) semimartingale. This implies in particular that $\sum_{s \leq t} |\Delta X_s|^p$ is finite for any $p > 2$ as well. \square

Remark 3.8 Following Jacod [12] there is a similar result for more general functions of polynomial growth, but the limiting behaviour of $\overline{V}(f)_t^n$ depends heavily on additional properties of the function f and the semimartingale X . In particular, assuming that f is continuous with $f(x) \sim |x|^p$ around zero, we have a more general version of Theorem 3.6:

For $p > 2$ the limit is always $\sum_{s \leq t} f(\Delta X_s)$, whereas for $p = 2$ it is $\int_0^t \rho_{\sigma_s}(f) ds + \sum_{s \leq t} f(\Delta X_s)$. For $p < 2$, the conditions on X come into play: If the Wiener part is non-vanishing, it dominates $\overline{V}(f)_t^n$, which in turn converges to infinity. However, for the standardised version $V(f)_t^n$ we have the same limiting behaviour as in Theorem 3.1, no matter what the jumps of X look like. If $1 < p < 2$ and there is no Wiener part, we have the limit $\sum_{s \leq t} f(\Delta X_s)$ again, provided that the jumps of power p are summable. A similar result holds for $0 < p \leq 1$, if the (genuine) drift part is zero as well. \square

Before we come to a sketch of the proof of Theorem 3.6, we state a local boundedness condition on the jumps, which is assumed to be satisfied for the rest of this section: δ is locally bounded by a family (γ_k) of deterministic functions with $\int (1 \wedge \gamma_k^2(x)) dx < \infty$. Though not necessary for the LLN, this assumption simplifies the proof and it is crucial for the CLT to hold. As Theorem 3.6 is also stable under stopping, we may assume again that a and σ are actually bounded and that all γ_k can be replaced by a bounded function γ satisfying $\int (1 \wedge \gamma^2(x)) dx < \infty$.

• *A fundamental decomposition:* The basic idea in essentially all of the proofs on discontinuous semimartingales is to fix an integer q first (which eventually tends to infinity) and to decompose X into the sum of the jumps larger than $1/q$ and the remaining terms, including the compensated jumps smaller than $1/q$. Precisely, we have for any q :

$$X_t = X(q)_t + X(q)'_t \quad \text{with} \quad X(q)'_t := X_0 + Q_t + M(q)_t + B(q)_t, \quad (3.25)$$

where

$$\left. \begin{aligned} X(q)_t &= \delta 1_{\{|\gamma| > 1/q\}} \star \mu_t, & Q_t &= \int_0^t a_s ds + \int_0^t \sigma_s dW_s, \\ M(q)_t &= \delta 1_{\{|\gamma| \leq 1/q\}} \star (\mu - \nu)_t, & B(q)_t &= \delta 1_{\{|\delta| \leq 1, \gamma > 1/q\}} \star \nu_t. \end{aligned} \right\} \quad (3.26)$$

If X exhibits only finitely many jumps, the decomposition becomes much simpler: $X(q)$ can be interpreted as the pure jump part of the semimartingale, whereas $X'(q)$ denotes its continuous part, and in this case one does not need the additional parameter q . Keeping this intuition in mind, it might be easier to follow the proofs.

It is crucial that $X(q)_t$ has only finitely many jumps, as this makes its contribution to $\overline{V}(f)_t^n$ rather simple to analyze. Setting

$$\overline{V}(R, p)_t^n = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} |\Delta_i^n R|^p$$

for any càdlàg process R and using

$$\overline{V}(Q, p)_t^n \xrightarrow{\mathbb{P}} \begin{cases} 0, & p > 2, \\ \int_0^t \sigma_s^2 ds, & p = 2 \end{cases}$$

from Theorem 3.1, the proof essentially reduces to showing that both $V(B(q), p)_t^n$ and $V(M(q), p)_t^n$ are small and that $V(X(q), p)_t^n$ converges to $\sum_{s \leq t} |\Delta X_s|^p$. One has to be careful here, as all quantities above depend both on n and q . Formally, this means proving

$$\left. \begin{aligned} \lim_{q \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left(|\overline{V}(B(q), p)_t^n| + |\overline{V}(M(q), p)_t^n| > \eta \right) &= 0, \\ \lim_{q \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \overline{V}(X(q), p)_t^n - \sum_{s \leq t} |\Delta X_s|^p \right| > \eta \right) &= 0 \end{aligned} \right\} \quad (3.27)$$

for all $\eta > 0$.

• *Some basic computations:* For the first claim in (3.27), a simple calculation shows that $B(q)$ behaves in a similar way as the drift term in Q ; precisely, we have $|\Delta_i^n B(q)| < C_q \Delta_n$.

This allows to focus on the local martingale $M(q)$ only. Following Proposition II.2.17 in [14] its quadratic variation process is given by

$$N(q)_t = \langle M(q), M(q) \rangle_t = |\delta|^2 1_{\{|\gamma| \leq 1/q\}} \star \nu_t,$$

and we have

$$|\Delta_i^n N(q)| = \left| \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{\{|\gamma(x)| \leq 1/q\}} |\delta(\omega, s, x)|^2 dx ds \right| \leq \Delta_n \int_{\{|\gamma(x)| \leq 1/q\}} |\gamma(x)|^2 dx = e_q \Delta_n,$$

and $e_q \rightarrow 0$ for $q \rightarrow \infty$ by assumption on γ . Thus the first part of (3.27) follows from *Burkholder's inequality* again, since

$$\mathbb{E}(|\Delta_i^n M(q)|^p) \leq \mathbb{E}(|\Delta_i^n N(q)|^{p/2}) \leq e_q^{p/2} \Delta_n^{p/2}$$

holds and $p \geq 2$. Finally, we know from the structure of the compensator $\nu(\omega; ds, dx) = ds \otimes dx$ that the finitely many (say: $K_q(t)$) jump times of $X(q)$ within $[0, t]$ have the same distribution (conditionally on $K_q(t)$) as a sample of $K_q(t)$ independent uniformly distributed variables on the same interval. Thus, for growing n it becomes less likely that two or more jump times are within the same interval $[(i-1)\Delta_n, i\Delta_n]$, and precisely we have $\Omega_n(t, q) \rightarrow \Omega$ almost surely, if we denote by $\Omega_n(t, q)$ the set of those ω for which all jump times of $X(q)$ are at least $2\Delta_n$ apart and none occurs in the interval $[\frac{t\Delta_n}{\Delta_n}, t]$. So w.l.o.g. we are on $\Omega_n(t, q)$, where we have

$$\bar{V}(X(q), p)_t^n = \sum_{s \leq t} |\Delta X(q)_s|^p$$

identically. Thus the last step of (3.27) follows from Lebesgue's Theorem, namely

$$\mathbb{E} \left(\left| \sum_{s \leq t} |\Delta X(q)_s|^p - \sum_{s \leq t} |\Delta X_s|^p \right| \right) \leq \mathbb{E} \left(\sum_{s \leq t} |\Delta X_s|^p 1_{\{|\Delta X_s| \leq 1/q\}} \right) \rightarrow 0$$

for $q \rightarrow \infty$. □

We have central limit theorems associated with any of the two types of convergence in Theorem 3.6, and it is no surprise that both limiting processes are fundamentally different from the one in (3.21).

Before we state the result, we have to introduce some further quantities. First, we need an extension of the original probability space, which supports a Brownian motion W' , two sequences (U_n) and (U'_n) of independent $N(0, 1)$ variables and a sequence (κ_n) of independent $U(0, 1)$ variables, all being mutually independent and independent of \mathbb{F} . Let further be (T_m) any choice of stopping times with disjoint graphs that exhausts the jumps of X , which means that $\Delta X_t \neq 0$ implies $t = T_m$ for some m and that $T_m \neq T_{m'}$ for $m \neq m'$. Then we set for $p = 2$ and $p > 3$ (there is no CLT for $2 < p \leq 3$, since the Brownian part within $\bar{V}(f)_t^n$ is not negligible at the rate of convergence $\sqrt{\Delta_n}$):

$$\bar{L}(f)_t = \sum_{m: T_m \leq t} f'(\Delta X_{T_m}) \left(\sqrt{\kappa_m} U_m \sigma_{T_m-} + \sqrt{1 - \kappa_m} U'_m \sigma_{T_m} \right).$$

The theorem goes then as follows.

Theorem 3.9 *Let $f(x) = |x|^p$ for a non-negative exponent p . For any $t \geq 0$ we have*

$$\Delta_n^{-1/2} \left(\bar{V}(f)_t^n - \bar{V}(f)_t \right) \xrightarrow{st} \begin{cases} \bar{L}(f)_t & p > 3, \\ \bar{L}(f)_t + L(f)_t, & p = 2. \end{cases} \quad (3.28)$$

Remark 3.10 Note that $\bar{L}(f)_t$ for $p = 2$ does not converge in general (because it is not absolutely summable), and thus it depends on the specific choice of the stopping times (T_m) . However, it can be shown easily that its \mathbb{F} -conditional law does *not*, since conditionally on \mathbb{F} the summands

$$\alpha_m = f'(\Delta X_{T_m}) \left(\sqrt{\kappa_m} U_m \sigma_{T_m-} + \sqrt{1 - \kappa_m} U'_m \sigma_{T_m} \right)$$

are independent, mean zero variables with

$$\mathbb{E}(\alpha_m^2 | \mathbb{F}) = \frac{1}{2} f'(\Delta X_{T_m})^2 (\sigma_{T_m-}^2 + \sigma_{T_m}^2).$$

By definition $f(x) = |x|^2$, and so

$$\sum_{m: T_m \leq t} f'(\Delta X_{T_m})^2 (\sigma_{T_m-}^2 + \sigma_{T_m}^2) < C \sum_{m: T_m \leq t} |\Delta X_{T_m}|^2 < \infty,$$

showing that the \mathbb{F} -conditional variance of $\bar{L}(f)_t$ is absolutely summable.

As we are interested in proving \mathbb{F} -stable convergence towards $\bar{L}(f)_t$, it is by definition only its \mathbb{F} -conditional law that matters. The previous claim thus gives us the freedom to work with any choice of (T_m) for the rest of this section, and we choose a convenient one as follows: Consider for any $q \geq 1$ the finitely many jump times $(T(m, q))$ of the Poisson process $\mu([0, t] \times \{1/q < \gamma(z) \leq 1/(q-1)\})$. Then we denote with $(T_m)_{m \geq 1}$ any reordering of the double sequence $(T(m, q) : m, q \geq 1)$, and we set $P_q = \{m : T_m = T(m', q') \text{ with } q' \leq q\}$. \square

Remark 3.11 In contrast to the continuous case, we have stated both the LLN and the CLT pointwise in t , but not in a functional sense. It is possible to do so, but one has to be careful then: If T is a specific jump time of X , then the jump ΔX_T is typically not included in any of the discretized processes $\bar{V}(f)_T^n$ (whose corresponding jump times T_n are not the same as T), but it obviously occurs in the limit. This prevents $\bar{V}(f)_T^n$ from converging uniformly in probability in the LLN (one has convergence in probability for the Skorokhod topology, however), and we only have a functional CLT for a discretized version, namely for $\Delta_n^{-1/2} \left(\bar{V}(f)_t^n - \sum_{s \leq \lfloor \frac{t \Delta_n}{\Delta_n} \rfloor} |\Delta X_s|^p \right)$. See [12] for details. \square

We conclude with a sketch of the proof of Theorem 3.3.

• *The case $p > 3$:* Recall the notation in (3.25) and set $I(m, n) = \min\{i : T_m \leq i \Delta_n\}$, so T_m is in $[(I(m, n) - 1) \Delta_n, I(m, n) \Delta_n]$. The main idea of the proof is again to separate the (finitely many) large jumps from the other terms in X . Precisely, on $\Omega_n(t, q)$ the identity

$$\Delta_n^{-1/2} \left(\bar{V}(f)_t^n - \sum_{s \leq t} f(\Delta X_s) \right) = \Delta_n^{-1/2} \left(V(X'(q), p)_t^n - \sum_{m \notin P_q} f(\Delta X_{T_m}) \right) + \sum_{m \in P_q} \zeta_m^n \quad (3.29)$$

holds, where we have defined

$$\zeta_m^n = \Delta_n^{-1/2} \left\{ f(\Delta_{I(m,n)}^n X) - f(\Delta X_{T_m}) - f(\Delta_{I(m,n)}^n X'(q)) \right\}.$$

The first term in (3.29) comprises the contributions of the continuous part of X plus the small jumps, and from a simple but tedious application of Itô's formula one gets

$$\lim_{q \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left(\Delta_n^{-1/2} \left| V(X'(q), p)_t^n - \sum_{m \notin P_q} f(\Delta X_{T_m}) \right| > \eta \right) = 0.$$

Following (3.16) this result is not surprising for Q and $B(q)$ within $X'(q)$ (recall $p > 3$), and the main part is to prove that the contributions of the small jumps cancel out.

We may thus focus on $\sum_{m \in P_q} \zeta_m^n$ only, and a Taylor expansion around ΔX_{T_m} gives for the dominating terms within each ζ_m^n :

$$f(\Delta_{I(m,n)}^n X) - f(\Delta X_{T_m}) = f'(\Delta X_{T_m}) \Delta_{I(m,n)}^n X'(q) + f''(\xi_m^n) (\Delta_{I(m,n)}^n X'(q))^2,$$

where ξ_m^n lies between $\Delta_{I(m,n)}^n X$ and ΔX_{T_m} . As f is a power function, a simple calculation gives

$$\begin{aligned} & \sum_{m \in P_q} |\zeta_m^n - \Delta_n^{-1/2} f'(\Delta X_{T_m}) \Delta_{I(m,n)}^n X'(q)| \\ & \leq C_p \sum_{m \in P_q} \Delta_n^{-1/2} (|\Delta_{I(m,n)}^n X'(q)|^p + |\Delta X_{T_m}|^{p-2} |\Delta_{I(m,n)}^n X'(q)|^2), \end{aligned}$$

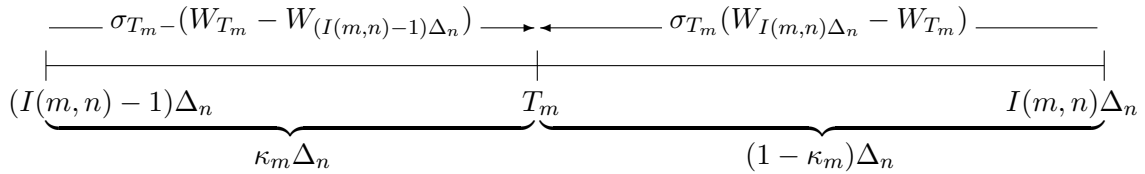
and by similar arguments as in the proof of Theorem 3.6 (note that P_q has only finitely many elements) this quantity converges in probability to zero for any q as $n \rightarrow \infty$. The proof of the first claim in Theorem 3.9 is finished, once one has shown

$$\sum_{m \in P_q} \Delta_n^{-1/2} f'(\Delta X_{T_m}) \Delta_{I(m,n)}^n X'(q) \xrightarrow{st} \bar{L}(f, q)_t, \quad (3.30)$$

where $\bar{L}(f, q)_t$ is the same quantity as $\bar{L}(f)_t$, but where the sum goes over the terms in P_q only (the convergence of $\bar{L}(f, q)_t$ towards $\bar{L}(f)_t$, as $q \rightarrow \infty$, is straight-forward). Proving (3.30) mainly amounts to showing the stable convergence

$$(\Delta_n^{-1/2} \Delta_{I(m,n)}^n X'(q))_{m \in P_q} \xrightarrow{st} (\sqrt{\kappa_m} U_m \sigma_{T_m} + \sqrt{1 - \kappa_m} U'_m \sigma_{T_m})_{m \in P_q}$$

for any fixed q . This result makes sense, as (3.16), the proof of Theorem 3.6 and Lemma 5.12 in [12] allow us to replace $\Delta_{I(m,n)}^n X'(q)$ by $\sigma_{T_m} - (W_{T_m} - W_{(I(m,n)-1)\Delta_n}) + \sigma_{T_m} (W_{I(m,n)\Delta_n} - W_{T_m})$. Since the jump time T_m within $[(I(m,n)-1)\Delta_n, I(m,n)\Delta_n]$ is uniformly distributed (see the graphic below), the additional factor $\kappa_m \sim U(0, 1)$ shows up.



Showing finally that the convergence is indeed a stable one is a bit tricky, since we cannot use Theorem 2.6 here. One works directly with Definition 2.1 instead, making extensive use of the choice of the stopping times (T_m) as well as of the fact that the homogeneous Poisson measure μ restricted to $\{\gamma > 1/q\}$ is independent of W . See again [12].

• *The case $p = 2$:* Let $f(x) = x^2$. The main idea is similar, as we have on $\Omega_n(t, q)$ the decomposition

$$\Delta_n^{-1/2} \left(\bar{V}(f)_t^n - [X, X]_t \right) = \Delta_n^{-1/2} \left(V(X'(q), 2)_t^n - \int_0^t \sigma_s^2 ds - \sum_{m \notin P_q} |\Delta X_{T_m}|^2 \right) + \sum_{m \in P_q} \zeta_m^n$$

with ζ_m^n as above, but with $p = 2$. In a same way as before, we have

$$\sum_{m \in P_q} \Delta_n^{-1/2} f'(\Delta X_{T_m}) \Delta_{I(m,n)}^n X'(q) \xrightarrow{st} \bar{L}(f, q)_t, \quad (3.31)$$

whereas we have

$$\Delta_n^{-1/2} \left(V(Q, 2)_t^n - \int_0^t \sigma_s^2 ds \right) \xrightarrow{st} L(f)_t \quad (3.32)$$

as in (3.21). Following Lemma 5.8 in [12] we also have the joint stable convergence in (3.31) and (3.32), and so one is left to show

$$\lim_{q \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left(\sup_{s \leq t} \left| \Delta_n^{-1/2} \left(V(X'(q), 2)_t^n - V(Q, 2)_t^n - \sum_{m \notin P_q} |\Delta X_{T_m}|^2 \right) \right| > \eta \right) \rightarrow 0,$$

which is again a consequence of Itô's formula. \square

References

- [1] Ait-Sahalia, Y., and J. Jacod (2009): Analyzing the spectrum of asset returns: jump and volatility components in high frequency data. Working paper. Available at <http://www.princeton.edu/~yacine/research.htm>
- [2] Aldous, D.J., and Eagleson, G.K. (1978): On mixing and stability of limit theorems. *Ann. of Prob.* 6(2), 325–331.
- [3] Barndorff-Nielsen, O.E., S.E. Graversen, J. Jacod, M. Podolskij, N. Shephard (2006): A central limit theorem for realised power and bipower variations of continuous semimartingales. In: Yu. Kabanov, R. Liptser and J. Stoyanov (Eds.), *From Stochastic Calculus to Mathematical Finance. Festschrift in Honour of A.N. Shiryaev*, Heidelberg: Springer, 2006, 33–68.
- [4] Barndorff-Nielsen, O.E., and N. Shephard (2004): Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics* 2, 1–48.

- [5] Barndorff-Nielsen, O.E., and N. Shephard, (2006): Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* 4, 1-30.
- [6] Barndorff-Nielsen, O.E., Shephard, N., Winkel, M. (2006): Limit theorems for multipower variation in the presence of jumps. *Stochastic Process. Appl.*, 116, 796-806.
- [7] Delbaen, F., and W. Schachermayer (1994): A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300, 463–520.
- [8] Dette, H., Podolskij, M., and M. Vetter (2006): Estimation of integrated volatility in continuous time financial models with applications to goodness-of-fit testing. *Scandinavian Journal of Statistics* 33, 259-278.
- [9] Hall, P. and C.C. Heyde (1980): *Martingale limit theory and its application*. Academic Press, New York.
- [10] Jacod, J. (1994): Limit of random measures associated with the increments of a Brownian semimartingale. *Preprint number 120, Laboratoire de Probabilités, Univ. P. et M. Curie*.
- [11] Jacod, J. (1997): On continuous conditional Gaussian martingales and stable convergence in law. *Seminaire de Probabilités XXXI*, 232-246.
- [12] Jacod, J. (2008): Asymptotic properties of realized power variations and related functionals of semimartingales. *Stoch. Proc. Appl.*, 118, 517-559.
- [13] Jacod, J. and Protter, P. (1998): Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.*, 26, 267-307.
- [14] Jacod, J. and A.N. Shiryaev (2003): *Limit Theorems for Stochastic Processes*, 2d ed., Springer-Verlag: Berlin.
- [15] Kinnebrock, S. and M. Podolskij (2008): A note on the central limit theorem for bipower variation of general functions. *Stochastic Processes and Their Applications* 118, 1056-1070.
- [16] Lépingle, D. (1976): La variation d'ordre p des semimartingales. *Z. für Wahr. Th.* 36, 285-316.
- [17] Mancini, C. (2001): Disentangling the jumps of the diffusion in a geometric jumping Brownian motion. *Giornale dell'Istituto Italiano degli Attuari LXIV* 1947.
- [18] Renyi, A. (1963): On stable sequences of events. *Sankhya A* 25, 293-302.
- [19] Revuz, D. and M. Yor (1998): *Continuous martingales and Brownian motion*, 3d ed., Springer, New York.
- [20] Vetter, M. (2009): Limit theorems for bipower variation of semimartingales. Working paper. Available at <http://www.rub.de/mathematik3/team/vetter.html>.

- 2009-34: Peter Christoffersen, Steven Heston and Kris Jacobs: The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well
- 2009-35: Peter Christoffersen, Jeremy Berkowitz and Denis Pelletier: Evaluating Value-at-Risk Models with Desk-Level Data
- 2009-36: Tom Engsted and Thomas Q. Pedersen: The dividend-price ratio does predict dividend growth: International evidence
- 2009-37: Michael Jansson and Morten Ørregaard Nielsen: Nearly Efficient Likelihood Ratio Tests of the Unit Root Hypothesis
- 2009-38: Frank S. Nielsen: Local Whittle estimation of multivariate fractionally integrated processes
- 2009-39: Borus Jungbacker, Siem Jan Koopman and Michel van der Wel: Dynamic Factor Models with Smooth Loadings for Analyzing the Term Structure of Interest Rates
- 2009-40: Niels Haldrup, Antonio Montañés and Andreu Sansó: Detection of additive outliers in seasonal time series
- 2009-41: Dennis Kristensen: Pseudo-Maximum Likelihood Estimation in Two Classes of Semiparametric Diffusion Models
- 2009-42: Ole Eiler Barndorff-Nielsen and Robert Stelzer: The multivariate supOU stochastic volatility model
- 2009-43: Lasse Bork, Hans Dewachter and Romain Houssa: Identification of Macroeconomic Factors in Large Panels
- 2009-44: Dennis Kristensen: Semiparametric Modelling and Estimation: A Selective Overview
- 2009-45: Kim Christensen, Silja Kinnebrock and Mark Podolskij: Pre-averaging estimators of the ex-post covariance matrix
- 2009-46: Matias D. Cattaneo, Richard K. Crump and Michael Jansson: Robust Data-Driven Inference for Density-Weighted Average Derivatives
- 2009-47: Mark Podolskij and Mathias Vetter: Understanding limit theorems for semimartingales: a short survey