LIMIT THEOREMS FOR INTEGRATED LOCAL EMPIRICAL CHARACTERISTIC EXPONENTS FROM NOISY HIGH-FREQUENCY DATA WITH APPLICATION TO VOLATILITY AND JUMP ACTIVITY ESTIMATION

By Jean Jacod and Viktor Todorov*

UPMC (Université Paris-6) and Northwestern University

We derive limit theorems for functionals of local empirical characteristic functions constructed from high-frequency observations of Itô semimartingales contaminated with noise. In a first step, we average locally the data to mitigate the effect of the noise and then in a second step we form local empirical characteristic functions from the pre-averaged data. The final statistics are formed by summing the local empirical characteristic exponents over the observation interval. The limit behavior of the statistics is governed by the observation noise, the diffusion coefficient of the Itô semimartingale and the behavior of its jump compensator around zero. Different choices for the block sizes for pre-averaging and formation of the local empirical characteristic function as well as for the argument of the characteristic function make the asymptotic role of the diffusion, the jumps and the noise differ. The derived limit results can be used in a wide range of applications and in particular for doing the following in a noisy setting: (1) efficient estimation of the time-integrated diffusion coefficient in presence of jumps of arbitrary activity, and (2) efficient estimation of the jump activity (Blumenthal-Getoor) index.

1. Introduction. In this paper we study the limit behavior of statistics based on empirical characteristic functions formed from discrete irregularly-sampled observations of an Itô semimartingale contaminated with observation noise. The asymptotic setting of the paper is one of fixed time span and mesh of the observation grid going to zero. The limit results derived here are rather general and can be applied for making inference regarding various quantities associated with the diffusion coefficient of the semimartingale and its jump component.

The statistics of interest are constructed as follows. We first average locally the data in order to mitigate the effect of the observation noise. This is done using the so-called pre-averaging technique of [10] and [19]. We then construct the real part of the empirical characteristic function (ecf) of the first-difference of the pre-averaged increments on local windows of asymptotically shrinking length. In a final step, we aggregate the local characteristic exponents over the observation interval.

By constructing the ecf over blocks of increments with sufficiently fast shrinking time span, the time variation of the characteristics of the semimartingale and that of the variance of the noise has an asymptotically negligible effect on our statistics. Therefore, the analysis of the real part of the ecf over the blocks can be performed as if the observed process is Lévy (i.e., one with constant characteristics) plus i.i.d. noise. Using the Lévy-Khintchine formula, the ecf of the increments of a Lévy process

*Todorov's work was partially supported by NSF grant SES-1530748.

AMS 2000 subject classifications: Primary 60F05, 60F17; secondary 60G51, 60G07.

Keywords and phrases: Blumenthal-Getoor Index, Central Limit Theorem, Empirical Characteristic Function, Integrated Volatility, Irregular Sampling, Itô Semimartingale, Jumps, Jump Activity, Microstructure Noise, Quadratic Variation, Stable Process.

observed with i.i.d. noise over a shrinking time interval is determined by the diffusion coefficient, the behavior of the Lévy measure of the jumps around zero as well as the variance of the noise.

By deriving the limit in probability of our statistics for different values of the argument of the characteristic function used in their construction, we can separately identify the diffusive and the jump part of the semimartingale as well as the variance of the observation noise. In particular, we can consistently estimate the diffusion coefficient and quantities pertaining to the jump compensator. We further derive associated Central Limit Theorem (CLT) results. By varying the asymptotic order of the argument of the characteristic function as well as that of the two local windows, used for pre-averaging and calculation of the ecf, we get a wide range of limit results depending on whether the diffusion, the jumps and/or noise dominate the asymptotic variance of the limit. In addition, we derive higher-order CLT when analyzing the limit behavior of differences of the statistics for different values of the argument of the characteristic function used in their construction.

Our results have a wide range of applications. First, using the proposed statistics we can construct efficient estimates of the integrated diffusion coefficient in the simultaneous presence of jumps of arbitrary activity (but of locally mixture-of-stable type), observation noise and irregular sampling. The problem of estimating the diffusion coefficient from high-frequency data has received a lot of attention. [3], [4] and [18] propose jump-robust estimators in a setting of no noise. The rate of convergence of these estimators, however, drops when the jumps are of infinite variation. [12, 13] propose an alternative estimator, again in the no-noise setting, which allows for efficient estimation even when jumps are of infinite variation (provided they are of locally mixture-of-stable type). Another strand of the literature has developed noise-robust estimators of the diffusion coefficient. Examples include [5], [6], [10], [17], [19] and [24]. These papers either do not allow for jumps in the underlying process or restrict its activity. The proposed estimator in this paper can work in the simultaneous presence of noise and jumps of arbitrary activity and it remains rate-efficient even when the latter are of infinite variation. To the best of our knowledge this has not been done in prior work.

Second, using the limit results of the current paper, we can develop estimates of the jump activity (Blumenthal-Getoor) index of the semimartingale in a noisy setting, both when the diffusive part of the semimartingale is present or not. Estimation of the jump activity index has been studied extensively in earlier work, with different methods of estimation and different setups affecting the rates of convergence of the estimation. In the no-noise setting and when the diffusion coefficient is present, [1], [7] and [15] use truncation based estimators. In the no-noise and no-diffusion setting, [21] and [23] use power variations and [20] uses local empirical characteristic functions. Finally, [14] adopt some of the above-mentioned no-noise estimators to noisy setting by doing initial pre-averaging of the raw data. The estimators that we propose here are more efficient than the ones based on truncated power variations considered in [14] and unlike [14] we derive the rate of convergence and a CLT for our estimators when the underlying process contains a diffusion.

The paper is organized as follows. In Section 2 we present our setting regarding the underlying process, the observation scheme and the noise. In Section 3 we construct our statistics from the high frequency data. Section 4 contains our limit results. All proofs are given in Section 5.

2. The setting.

Our setting contains three basic ingredients:

1. an underlying one-dimensional process X;
2. at each stage n, a strictly increasing sequence of observation times $0 = T(n,0) < T(n,1) < \cdots$;
3. at each stage n, a sequence of variables $(\chi_i^n : i \geq 0)$ which represents the observation noise; that is, at time $T(n,i)$ one does not observe directly $X_{T(n,i)}$, but instead $X_{T(n,i)} + \chi_i^n$, to account for the so-called *microstructure noise* in financial data.

All these objects are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and we now describe the assumptions for each of the three ingredients.

2.1. *The underlying process.* We start with the underlying process. The process X is a one-dimensional Itô semimartingale, relative to some càdlàg filtration $(\mathcal{F}_t)_{t\geq 0}$, and it takes the form

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s + \delta * (\underline{p} - \underline{q})_t + \delta' * \underline{p}_t, \tag{2.1}$$

where W is a Brownian motion, $\underline{p}$ is a Poisson measure on $\mathbb{R}_+ \times E$ with deterministic compensator $\underline{q}(dt, dz) = dt \otimes \eta(dz)$. Here E is a Polish space and η is a σ-finite measure on E. For a function ϕ on $\Omega \times \mathbb{R}_+ \times E$ and a random measure ν on $\mathbb{R}_+ \times E$, the notation $\phi * \nu_t$ stands for the double (ordinary or stochastic) integral $\int_0^t \int_E \phi(s, z)\, \nu(ds, dz)$. The process b is optional, the process σ is càdlàg adapted, the functions δ and δ' on $\Omega \times \mathbb{R}_+ \times E$ are predictable and such that the integrals in (2.1) make sense (this will be implied by our assumptions below).

We also assume that the volatility process σ is itself an Itô semimartingale, which can thus be written as

$$\sigma_t = \sigma_0 + \int_0^t b^\sigma_s\, ds + \int_0^t H^\sigma_s\, dW_s + \int_0^t H'^\sigma_s\, dW'_s + \delta^\sigma * (\underline{p} - \underline{q})_t + \delta'^\sigma * \underline{p}_t, \tag{2.2}$$

with $b^\sigma, H^\sigma, H'^\sigma$ optional and $\delta^\sigma, \delta'^\sigma$ predictable. Choosing the same Poisson measure $\underline{p}$ to drive both X and σ is not a restriction, and we use W and W' in (2.2) to allow for general dependence between the diffusion components of X and σ.

The Itô semimartingale assumption for σ is satisfied in many applications, e.g., when σ is modeled as Lévy-driven SDE. Such an assumption, however, rules out models in which σ is driven by a fractional Brownian motion, see e.g., [8] and [9]. We conjecture that our results can be extended to such settings (with possibly different - worse or better depending on the Hurst parameter - rates of convergence) but we leave such an extension for future work.

For the assumptions, we need to introduce two properties relative to a generic $(\mathcal{F}_t)$-optional process V, with some $q, q' > 0$ and where K is a constant depending on V:

$$\mathbb{E}\big(\sup_{s\in[T,S]} |V_s - V_T|^{q'}\big) \;\leq\; K\mathbb{E}\big((S-T)^q\big) \qquad \text{for any two finite stopping times } T \leq S, \tag{2.3}$$

$$\big|\,\mathbb{E}(V_S - V_T)\big| \;\leq\; K\mathbb{E}\big((S-T)^q\big) \qquad \text{for any two finite stopping times } T \leq S. \tag{2.4}$$

We will denote the first property above as $(2.3)_{q,q'}$, and the second one as $(2.4)_q$. If V is a bounded Itô semimartingale with bounded characteristics, then it satisfies $(2.3)_{2,1}$ and $(2.4)_1$. If in addition V is of finite variation $(2.3)_{1,1}$ holds as well.

We also recall that the process $\overline{X} = \delta * (\underline{p} - \underline{q})$ has a jump measure μ whose $(\mathcal{F}_t)$-predictable compensator ν factorizes as $\nu(dt, dx) = dt \otimes F_t(dx)$, where $F_t = F_{t,\omega}$, called the "spot Lévy measure", is the restriction to $\mathbb{R}\backslash\{0\}$ of the image of the measure λ by the map $z \mapsto \delta(\omega, t, z)$.

For an arbitrary positive measure F on $\mathbb{R}$ we consider the "tail function" defined for $x > 0$ by $\overline{F}(x) = F((-\infty, -x)) + F((x, \infty))$. Further, $\breve{F}$ denotes the measure defined by $\breve{F}(A) = \frac{1}{2}\big(F(A) + F(-A)\big)$ and where $-A = \{x : -x \in A\}$. When F is a signed measure, we denote by $|F|$ its "absolute value" (the smallest positive measure such that $|F| - F$ is also a positive measure), and by $\overline{|F|}$ the tail function of $|F|$.

With this notation, we next state our assumption for X.

Assumption (H). *We have (2.1) and (2.2), an integer $M \geq 1$, numbers $r, r', \beta_1, \dots, \beta_M$ such that $0 \leq r < \beta_M < \cdots < \beta_2 < \beta_1 \leq 2$ and nonnegative adapted càdlàg processes $a^1, \dots, a^M$, with the following properties: for each $r' \in (\beta_1, 2)$ we have a sequence τ_n of stopping times increasing to infinity, a sequence J_n of $[0,1]$-valued Borel functions on E with $\int J_n(z)\,\eta(dz) < \infty$ and a sequence Γ_n of numbers such that, with the notation*

$$F'_t(dx) = \breve{F}_t(dx) - \sum_{m=1}^{M} \frac{\beta_m a_t^m}{|x|^{1+\beta_m}}\, 1_{\{0<|x|\leq 1\}}\, dx, \tag{2.5}$$

we have

$$t < \tau_n \implies \left\{ \begin{array}{l} |b_t|,\, |\sigma_t|,\, |b_t^\sigma|,\, |H_t^\sigma|,\, |H_t'^\sigma|,\, a_t^m \leq \Gamma_n, \\ |\delta(t,z)|^{r'},\, |\delta^\sigma(t,z)|^2,\, 1_{\{\delta'^\sigma(t,z)\neq 0\}},\, 1_{\{\delta'(t,z)\neq 0\}} \leq J_n(z), \\ x > 0 \ \Rightarrow\ \overline{|F'_t|}(x) \leq \frac{\Gamma_n}{x^r}. \end{array} \right. \tag{2.6}$$

Moreover,
(i) the processes $b_{t\wedge\tau_n}$, $H^\sigma_{t\wedge\tau_n}$ and $\delta(t\wedge\tau_n, z)/J_n(z)^{1/r'}$ for all z satisfy $(2.3)_{2,1}$ with $K = \Gamma_n$,
(ii) the processes $(a^m_{t\wedge\tau_n})^{1/\beta_m}$ for $m = 1, \dots, M$ satisfy $(2.3)_{2,1}$ and $(2.4)_1$ with $K = \Gamma_n$.

We note that if (H) holds with some r it also holds for any $\overline{r} \in (r, \beta_M)$. Since we allow the processes a_t^m to be identically 0, we can always add another index β in $(r, 2)$ with the associated process a_t identically 0: this is of course immaterial and it looks a priori strange. However, we use the above formulation for a unified representation, which nests also the case where $M = 1$ and $a_t^1(\omega) = 0$ identically. In this case (2.5) reduces to $F'_t = \breve{F}_t$, and the last condition in (2.6) becomes $\overline{F}_t(x) \leq \Gamma_n/x^r$.

Assumption (H) restricts the behavior of the Lévy measure of X, but only around zero (and only when at least one of a_t^m is not identically zero). We note in that respect that the measures F'_t in (2.5) are *a priori* signed measures. The restriction of the Lévy measure around zero is to be like that of a sum of time-changed stable processes. The parameter r controls the degree of deviation of F_t from that of the mixture of time-changed stable processes. In many parametric jump specifications, (H) will be satisfied with $M = 1$ and $r < \beta_1/2$. In the Lévy case, for example, this will hold when $F_t(dx) = f(x)dx$ and $f(x)/x^{-1-\beta_1}$ converges to a positive constant as $x \to 0$ and has a bounded first derivative in a neighborhood of zero. This is the case for many parametric jump models, e.g., the tempered stable and the generalized hyperbolic.

Assumption (H) is satisfied for the class of time-changed Lévy processes with absolute continuous time change (the drift, diffusion and jumps can have separate time changes), provided the jump part of the Lévy process behaves around zero like that of a sum of stable processes. Although less obvious, (H) is also satisfied for the class of Lévy-driven stochastic differential equations (provided (H') below holds), i.e., when X takes the form

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s + \sum_{m=1}^{M} \int_0^t \overline{\sigma}^m_{s-}\, dZ^m_s + \delta' * \underline{p}_t, \tag{2.7}$$

with b, σ, δ' and $W, \underline{p}$ as above, and where the processes $\overline{\sigma}_t^m$ are càdlàg adapted and the processes Z^m are independent Lévy processes with no drift and no Gaussian part, and also independent of W and $\underline{p}$. We denote by F^m the Lévy measure of Z^m, so the "no drift and no Gaussian part" means that the characteristic function of Z_t^m is

$$\mathbb{E}\Big(e^{iuZ_t^m}\Big) = \exp\Big(t \int_{\mathbb{R}} \big(e^{iux} - 1 - iux\, 1_{\{|x|\leq 1\}}\big)\, F^m(dx)\Big). \tag{2.8}$$

Accordingly, we replace (2.2) by

$$\sigma_t = \sigma_0 + \int_0^t b_s^\sigma\, ds + \int_0^t H_s^\sigma\, dW_s + \int_0^t H_s'^\sigma\, dW_s' + \sum_{m=1}^M \int_0^1 H_{s-}^{\sigma,m}\, dZ_s^m + \delta^\sigma * (\underline{p} - \underline{q})_t + \delta'^\sigma * \underline{p}_t, \tag{2.9}$$

where each $H^{\sigma,m}$ is càdlàg adapted. The next assumption implies in particular that the Blumenthal-Getoor index of each Z^m is β_m.

Assumption (H'). *We have (2.7) and (2.9), and also numbers $2 > \beta_1 > \cdots > \beta_M > r \geq 0$, and $b_t, \sigma_t, b_t^\sigma, H_t^\sigma, H_t'^\sigma, \delta'(t,z), \delta^\sigma(t,z), \delta'^\sigma(t,z)$ satisfy the same conditions as in (H). Moreover, we have a sequence τ_n of stopping times increasing to infinity and a sequence Γ_n of numbers such that, for all $m = 1, \ldots, M$:*

(i) if $t < \tau_n$ then $|\overline{\sigma}_t^m| \leq \Gamma_n$ and $|H_t^{\sigma,m}| \leq \Gamma_n$;

(ii) the processes $|\overline{\sigma}_{t\wedge\tau_n}^m|$ satisfy $(2.3)_{2,1}$ and $(2.4)_1$ with $K = \Gamma_n$;

(iii) for some constant $\alpha_m, \alpha_m' > 0$ the signed measures $F'^m(dx) = \breve{F}^m(dx) - \frac{\alpha_m}{|x|^{1+\beta_m}} 1_{\{0<|x|\leq 1\}}\, dx$ (where $\breve{F}^m(A) = \frac{1}{2}\left(F^m(A) + F^m(-A)\right)$, as in (2.5)) satisfy $\overline{|F_m'|}(x) \leq \alpha_m'/x^r$ for all $x > 0$.

As proved below, see Lemma 8, any process satisfying (H') also satisfies (H). Therefore, henceforth, we always use the more general formulation (2.1).

2.2. *The observation scheme.* We next describe how observations take place. At stage n, that is, for a given frequency of observations, the successive observations occur at times $0 = T(n,0) < T(n,1) < \cdots$ for a sequence $T(n,i)$ of (possibly random) finite times. We will assume a rather special form for the sampling scheme, which involves a positive process λ_t and a double sequence $(\Phi_i^n : i, n \geq 1)$ of positive random variables, and at each stage n, the sampling times $T(n,i)$ are defined recursively in i, starting with $T(n,0) = 0$ and using the formulae

$$\Delta(n, i+1) = \Delta_n \lambda_{T(n,i)} \Phi_{i+1}^n, \qquad T(n, i+1) = T(n,i) + \Delta(n, i+1). \tag{2.10}$$

Here, $\Delta_n \to 0$ is a non-random sequence which plays the role of an "average mesh size" of the observation grid at stage n. Note, however, that the sampling times $T(n,i)$ and the inter-observation lags $\Delta(n,i)$ are observed up to the time horizon t, whereas Δ_n is a *non-observable* mathematical abstraction, which should not enter the various statistics constructed by the statistician.

We assume the following for the process λ_t and the sequence Φ_i^n.

Assumption (O). *There are a sequence (τ_m) of $(\mathcal{F}_t)$-stopping times increasing to ∞ and constants Γ_m, $\Gamma(p)$ such that*

(i) the process λ_t is càdlàg adapted with $1/\Gamma_m \leq \lambda_t \leq \Gamma_m$ for all $t < \tau_m$;

(ii) the stopped processes $\lambda_{t\wedge\tau_m}$ satisfy $(2.3)_{1,1}$ with $K = \Gamma_m$;

(iii) for each n the variables $(\Phi_i^n : i = 0, 1, \ldots)$ are mutually independent and independent of $\mathcal{F}_\infty$ and, for all $p > 0$:

$$\mathbb{E}(\Phi_{i+1}^n) = 1, \qquad \mathbb{E}((\Phi_{i+1}^n)^p) \leq \Gamma(p). \tag{2.11}$$

The number of observations up to time t is $N_t^n + 1$, where

$$N_t^n = \sum_{i\geq 1} 1_{\{T(n,i)\leq t\}}, \tag{2.12}$$

and we prove in Section 5 that the process $1/\lambda_t$ is like the "density" of observations, in the sense that

$$\Delta_n N^n_t \stackrel{\text{u.c.p.}}{\Longrightarrow} \Lambda_t := \int_0^t \frac{1}{\lambda_s}\, ds. \tag{2.13}$$

Part (ii) of (O) is somewhat restrictive. It is possible to assume only that each $V = \lambda_{t\wedge\tau_m}$ satisfies $(2.3)_{q,1}$ for some $q \in (1,2)$. This is at the expense of more complicated proofs and slower rates of convergence in some of the cases below, and the case $q = 2$ is unfortunately excluded.

A regular scheme $\Delta(n,i) = \Delta_n$ obviously satisfies (O) with $\lambda_t = 1$. A Poisson scheme, for which the counting process N^n is Poisson with parameter $1/\Delta_n$ and independent of X, also satisfies (O) with $\lambda_t = 1$. This happens when we take Φ^n_i to be exponential with parameter 1. More generally, assumption (O) allows the $\mathcal{F}_{T(n,i)}$-conditional law of $\Delta(n,i+1)$ to vary over time.

2.3. *The observation noise.* At stage n we do not observe $X_{T(n,i)}$ for $i = 0,1,\dots$, but rather $Y^n_i = X_{T(n,i)} + \chi^n_i$, where χ^n_i is "noise". The typical situation considered in the literature is when, for each n, the $(\chi^n_i)_{i\geq 0}$ are i.i.d. centered and independent of X and of the sampling times. Here we want to relax this assumption significantly, while keeping the property that the variables χ^n_i are centered and mutually independent as i varies, *conditionally on* the σ-field $\mathcal{H}^n_\infty = \mathcal{F}_\infty \bigvee \mathcal{K}^n_\infty$ with $\mathcal{K}^n_i = \sigma(\Phi^n_j : 0 \leq j \leq i)$. We also denote by $(\mathcal{H}^n_t)$ the smallest filtration containing $(\mathcal{F}_t)$ and with respect to which $T(n,i)$ is a stopping time for all $i \geq 0$.

It is obviously no restriction to "standardize" the noise by singling out a possible modulation via an $(\mathcal{F}_t)$-adapted process, times a new noise which has $\mathcal{H}^n_\infty$-conditional variance of 1. Therefore, we assume that, for a suitable $(\mathcal{F}_t)$-adapted process γ'_t, the i-th observation at stage n is

$$Y^n_i = X_{T(n,i)} + \chi^n_i = X_{T(n,i)} + \gamma'_{T(n,i)}\varepsilon^n_i. \tag{2.14}$$

We will use two different assumptions for the noise, which we state next.

Assumption (N-1). *We have (2.14), a sequence τ_m of $(\mathcal{F}_t)$-stopping times increasing to ∞, and for each integer $p \geq 1$ a càdlàg $(\mathcal{F}_t)$-adapted process $\gamma^{(p)}_t$ and constants $\Gamma(p)_m$ such that:*

(i) The stopped processes $\gamma'_{t\wedge\tau_m}$ and $\gamma^{(p)}_{t\wedge\tau_m}$ satisfy $(2.3)_{2,1}$ with $K = \Gamma_m$.

(ii) We have $\gamma^{(1)}_t = 0$ and $\gamma^{(2)}_t = 1$.

(iii) For all n the variables ε^n_i are independent as i varies, conditionally on $\mathcal{H}^n_\infty$, and satisfy for all integers $p \geq 1$ and all Borel subset B of $\mathbb{R}$:

$$\mathbb{E}((\varepsilon^n_i)^p \mid \mathcal{H}^n_\infty) = \gamma^{(p)}_{T(n,i)}, \qquad \mathbb{P}(\varepsilon^n_i \in B \mid \mathcal{H}^n_\infty) = \mathbb{P}(\varepsilon^n_i \in B \mid \mathcal{H}^n_{T(n,i)}). \tag{2.15}$$

Assumption (N-0). *We have (2.14), a sequence τ_m of $(\mathcal{F}_t)$-stopping times increasing to ∞, and constants $\Gamma(p)$ and Γ_m, such that:*

(i) The stopped processes $\gamma'_{t\wedge\tau_m}$ satisfy $(2.3)_{2,1}$ and $(2.4)_1$ with $K = \Gamma_m$.

(ii) For each n the sequence $(\varepsilon^n_i)_{i\geq 0}$ is independent of the σ-field $\mathcal{H}^n_\infty$ and i.i.d. as i varies and satisfies for all $p > 0$:

$$\mathbb{E}(\varepsilon^n_i) = 0, \qquad \mathbb{E}((\varepsilon^n_i)^2) = 1, \qquad \mathbb{E}(|\varepsilon^n_i|^p) < \infty. \tag{2.16}$$

Henceforth, we will use the notation $\gamma_t = (\gamma'_t)^2$ for the ($\mathcal{H}^n_\infty$-conditional) variance process of the noise. We note that the last part of (2.15) is equivalent to saying that $\varepsilon^n_i = f^n_i(\omega, \overline{\varepsilon}^n_i)$, where f^n_i is an $\mathcal{H}^n_{T(n,i)} \otimes \mathcal{R}$-measurable function on $\Omega \times \mathbb{R}$ and $\overline{\varepsilon}^n_i$ is a variable which is independent of $\mathcal{H}^n_\infty$.

Assumption (N-0) implies (N-1) with $\gamma^{(p)}_t$ identically equal to a constant for all $p \in \mathbb{N}$. (N-0) is satisfied in the case of a white noise independent of X. (N-0) also holds in the case of a "modulated" white noise, i.e., when the $\mathcal{H}^n_\infty$-conditional moments of the noise are time-varying. In particular, this allows for dependence between the observation noise and the unobservable X.

For financial applications where X is an asset price, there is typically a *rounding* effect, i.e., the observed price is integer-valued (prices can move only by multiples of ticks), and this effect is non-negligible if one is sampling the price very finely. The presence of rounding is basically incompatible with an Itô semimartingale plus a white noise (even a modulated one), i.e., assumption (N-0).

This is why we introduce the weaker assumption (N-1), which accommodates some kind of "additive noise plus rounding". Many versions are possible, the simplest one being as follows. For any $x \in \mathbb{R}$ we denote by $[x] = \max(n \in \mathbb{N} : n \leq x)$ its integer part and by $\{x\} = x - [x]$ its fractional part. At each stage n we have an i.i.d. sequence $(\zeta^n_i : i \geq 0)$, independent of $\mathcal{H}^n_\infty$, with the density $\frac{\alpha}{2}\, 1_{(-1,0)} + \frac{\alpha}{2}\, 1_{(1,2)} + (1-\alpha)\, 1_{(0,1)}$ for some $\alpha \in [0,1)$. The observation at time $T(n,i)$ is

$$Y^n_i = [X_{T(n,i)} + \zeta^n_i].$$

With the notation $Z_t = \{X_t\}(1 - \{X_t\})$ and $Z(p)_t = (1 - \{X_t\})^p - (-\{X_t\})^p$, a computation shows us that (2.14) and (2.15) are satisfied with, for each integer $p \geq 1$,

$$\gamma'_t = \sqrt{\alpha + Z_t},$$
$$\gamma^{(p)}_t = \tfrac{1}{(\alpha+Z_t)^{p/2}} \Big(\alpha\, 1_{\{p \geq 2\}} + Z_t Z(p-1)_t + \alpha \sum_{j=1}^{[(p-1)/2]} \tfrac{p!}{(2j)!(p-2j)!}\, Z_t Z(p-2j-1)_t \Big),$$

hence (ii,iii) of (N-1) holds. Moreover, (H) implies that, up to a localization, X satisfies $(2.3)_{2,1}$, which in turn implies that (again up to a localization) Z_t and $Z(p)_t$ satisfy $(2.3)_{2,1}$ as well. Then, as soon as $\alpha > 0$, one obtains that (i) of (N-1) holds. So, we have (N-1). However, although X also satisfies $(2.4)_1$ (up to a localization once more), the process Z_t does *not* satisfy $(2.4)_1$, so (i) of (N-0) cannot be true and this example cannot satisfy (N-0), even if we were to appropriately weaken (ii) of (N-O).

3. Construction of the statistics. In what follows it is convenient to single out some special cases, and towards this aim we introduce the following additional notation:

$$\kappa_1 = \begin{cases} 0 & \text{if } \sigma_t \equiv 0 \\ 1 & \text{otherwise} \end{cases}, \qquad \beta = \begin{cases} \beta_1 & \text{if } \sigma_t \equiv 0 \\ 2 & \text{otherwise} \end{cases}, \qquad \kappa_2 = \begin{cases} 0 & \text{under (N-0)} \\ 1 & \text{under (N-1)} \end{cases}. \tag{3.1}$$

Our ecf-based statistics are constructed in two steps. We first "de-noise" the observations, and then we compute local empirical characteristic functions. The first step needs a window of size h_n while the second step needs another window of size k_n and a sequence $u_n > 0$ of reals (both h_n and k_n are positive integers). We will specify later the conditions on these tuning parameters, but in any case they should always satisfy the following, for some $\varepsilon > 0$:

$$\begin{array}{l} k_n \Delta_n^\varepsilon, \;\; h_n \Delta_n^\varepsilon, \;\; u_n \Delta_n^\varepsilon, \;\; k_n^2 h_n^2 \Delta_n \;\to\; \infty, \\ k_n^2 h_n \Delta_n, \;\; h_n^3 \Delta_n^2, \;\; \frac{u_n^2}{h_n}, \;\; u_n^\beta h_n \Delta_n, \;\; u_n^2 (h_n \Delta_n)^3 \Delta_n^{-\varepsilon} \;\to\; 0. \end{array} \tag{3.2}$$

3.1. *Pre-averaging.* The first step in the construction of our statistics is to effectively "de-noise" the data which we do via pre-averaging ([10], [19]). The pre-averaging method amounts to average the data over a window of h_n successive increments, with the help of a weight (or, kernel) function g on $\mathbb{R}$, which satisfies

$$\begin{array}{l} g \text{ is continuous, piecewise } C^1 \text{ with a piecewise Lipschitz derivative } g', \\ s \notin (0,1) \;\Rightarrow\; g(s)=0, \qquad \int_0^1 g(s)^2 ds \;>\; 0. \end{array}$$

With g and the sequence h_n and $\beta \in [1,2)$ we associate the numbers (indexed by $n \geq 1$ and $i,j \in \mathbb{Z}$)

$$\begin{array}{lll} g_i^n = g(i/h_n), & \overline{g}_i^n \;=\; g_{i+1}^n - g_i^n, & \\ \phi_n = \frac{1}{h_n}\sum_{i\in\mathbb{Z}} (g_i^n)^2, & \overline{\phi}_n = h_n \sum_{i\in\mathbb{Z}} (\overline{g}_i^n)^2, & \widetilde{\phi}_n^{(\beta)} = \frac{1}{h_n}\sum_{i\in\mathbb{Z}} |g_i^n|^\beta, \end{array}$$

which satisfy as $n \to \infty$:

$$\phi_n \to \phi := \int g(u)^2\,du, \qquad \overline{\phi}_n \to \overline{\phi} := \int g'(u)^2\,du, \qquad \widetilde{\phi}_n^{(\beta)} \to \widetilde{\phi}^{(\beta)} := \int |g(u)|^\beta\,du. \tag{3.3}$$

Recall that we observe Y_i^n, as given by (2.14). More generally, for any process V we write $V_i^n = V_{T(n,i)}$, and also $\Delta_i^n V = V_i^n - V_{i-1}^n$, e.g., $\Delta_i^n X$ is the i-th increment and $\Delta_i^n Y = Y_i^n - Y_{i-1}^n$ is the i-th observed (noisy) increment. If V_i^n is any array of variables, we set

$$\widetilde{V}_i^n = \sum_{j=1}^{h_n-1} g_j^n\,\Delta_{i+j}^n V = -\sum_{j=0}^{h_n-1} \overline{g}_j^n\,V_{i+j}^n.$$

We note that the variable $\widetilde{V}_i^n$ implicitly depends on h_n and g. When $h_n = 2$, we simply have $\widetilde{V}_i^n = g(1/2)\Delta_{i+1}^n V$. The effect of the pre-averaging on the noise is to reduce its asymptotic order of magnitude by a factor of $\sqrt{h_n}$ while at the same time the order of magnitude of the pre-averaged Itô semimartingale remains unchanged. Thus, the asymptotic size of the noise relative to the Itô semimartingale after pre-averaging shrinks.

3.2. *Local empirical characteristic functions.* Below, we use the pre-averaged variables $\widetilde{Y}_i^n$, and we set $w_n = 2h_nk_n$. For any $y \in \mathbb{R}\backslash\{0\}$ we denote

$$L(y)_i^n = \frac{1}{k_n}\sum_{l=0}^{k_n-1} \cos\big(u_n y(\widetilde{Y}_{i+2lh_n}^n - \widetilde{Y}_{i+(2l+1)h_n}^n)\big), \tag{3.4}$$

which is the real part of the empirical characteristic function from a block of pre-averaged increments. In the no-noise setting, integrals over time of this statistic have been used by [22] for recovering the Laplace transform of the volatility process. Here, we will work with the characteristic exponent, i.e., we transform $L(y)_i^n$ as follows:

$$\widehat{c}(y)_i^n = -\log\Big(L(y)_i^n \bigvee \frac{1}{h_n}\Big).$$

For bias correction we will need further an estimate for the (locally integrated) variance of the noise, and for this we set

$$\widehat{Y}_i^n = \frac{1}{w_n}\sum_{l=1}^{w_n} (\Delta_{i+l}^n Y^n)^2, \tag{3.5}$$

and, with $f(x,y) = \frac{1}{2}\left(e^{2x-y} + e^{2x} - 2\right)$, we denote:

$$\widehat{C}(y)^n_t = \sum_{j=0}^{[N^n_t/w_n]-1} \Big(\widehat{c}(y)^n_{jw_n} - \frac{1}{2k_n} f(\widehat{c}(y)^n_{jw_n}, \widehat{c}(2y)^n_{jw_n}) - \frac{1}{2h_n} \overline{\phi}_n y^2 u_n^2 \widehat{Y}^n_{jw_n}\Big). \tag{3.6}$$

The last two terms on the right-hand side of (3.6) are bias corrections which are needed because of the nonlinear transformation of the local ecf and the presence of observation noise.

The above statistic can be viewed as the noise-robust analogue of the statistic proposed by [12, 13] for efficient volatility estimation ([16] use also the latter statistic for the purposes of testing for presence of diffusion in a no-noise setting). As we show later, $\widehat{C}(y)^n_t$ can be used not only for efficient estimation of the diffusion coefficient but also for estimating quantities associated with the jumps of X.

We conclude this section with introducing some additional notation. For $\gamma \in (0,2)$ we set

$$\chi(\gamma) = \int_0^\infty \frac{\sin y}{y^\gamma}\, dy,$$

which is a convergent integral for all $\gamma > 0$, but absolutely convergent when $\gamma > 1$ only. We also set

$$\begin{gathered} c_t = (\sigma_t)^2, \qquad C_t = \textstyle\int_0^t c_s\, ds, \qquad A^m_t = \int_0^t a^m_s\, ds, \\ \psi_\beta(y,y') = 2|y|^\beta + 2|y'|^\beta - |y+y'|^\beta - |y-y'|^\beta, \\ \overline{\psi}_\beta(y,y') = \psi_\beta(y,y') + y^2 y'^2 \psi_\beta(1,1) - y^2\psi_\beta(1,y') - y'^2\psi_\beta(y,1). \end{gathered} \tag{3.7}$$

Our key theorems in the next section describe the behavior of the centered processes

$$Z(y)^n_t = \widehat{C}(y)^n_t - \frac{y^2 u_n^2 \phi_n}{2k_n}\, C_t - \frac{2}{k_n}\sum_{m=1}^M |y|^{\beta_m} u_n^{\beta_m} \widetilde{\phi}_n^{\beta_m} \chi(\beta_m)\, A^m_t. \tag{3.8}$$

The centering terms in $Z(y)^n_t$ are scaled versions of C_t and $\{A^m_t\}_{m\geq 1}$, with the asymptotic magnitude of the scales depending on the order of magnitude of u_n and k_n. Since $u_n \to \infty$, the centering term involving C_t is asymptotically the largest, followed by the term involving A^1_t, etc. We note that in the centering of $\widehat{C}(y)^n_t$ above there is no term due to the noise. This is because we have already performed bias correction for the noise in the construction of $\widehat{C}(y)^n_t$ (the last summand in (3.6)).

4. Limit behavior of the statistics.

4.1. *Convergence in probability.* We start first with establishing convergence in probability of $v_n Z(y)^n_t$ towards 0 for an appropriate normalizing sequence v_n. The next theorem states the general result.

THEOREM 1. *Assume (H), (O), and (N-0) or (N-1). For any $t \geq 0$ and $y \neq 0$ we have $v_n Z(y)^n_t \stackrel{\mathbb{P}}{\longrightarrow} 0$ if the tuning parameters h_n, k_n, u_n and the sequence v_n of positive numbers satisfy* (3.2) *and, for some $\varepsilon > 0$ (as small as wanted) and all $m = 1,\ldots,M$,*

$$\begin{gathered} \tfrac{v_n^2}{k_n^2}\left(k_n^2 + u_n^{2r} + u_n^{\beta_1+\varepsilon} + \tfrac{u_n^4}{h_n^3\Delta_n} + \tfrac{u_n^8}{k_n^3 h_n^6 \Delta_n^2} + u_n^2 h_n\Delta_n + \tfrac{u_n^5}{h_n^2} + \tfrac{u_n^{2\beta_1}}{k_n^3} + u_n^{2\beta_m}(k_n h_n \Delta_n)^{\beta_m}\right) \to 0, \\ \kappa_1 = 1 \;\Rightarrow\; \tfrac{v_n^2}{k_n^2}\left(u_n^4 h_n\Delta_n + \tfrac{u_n^8 (h_n\Delta_n)^2}{k_n^3}\right) \to 0, \qquad \kappa_2 = 1 \;\Rightarrow\; \tfrac{v_n^2 u_n^4}{k_n h_n^3 \Delta_n} \to 0. \end{gathered} \tag{4.1}$$

This is a general "abstract" type of consistency result. It will allow us to estimate in a consistent way the integrated volatility C_t, the biggest index β_1 and the associated A^1_t. We will illustrate this in Section 4.3. In addition, it should be also possible to use the above result to estimate the next indices $\beta_2, \beta_3, \ldots$ (and the associated $A^2_t, A^3_t, \ldots$), but for simplicity we will not discuss this in this paper.

4.2. *Central Limit Theorems.* We continue with a CLT associated with the convergence in probability result in Theorem 1. By this we mean a result stating that, for a suitable sequence v_n, the variables $v_n Z(y)^n_t$ do not go to 0 but converge in law to a non-trivial limit. Depending on the choice of the tuning parameters and whether the underlying process X contains a diffusion, the CLT can be determined by the diffusion component of X, the jumps, the noise, or any combination of them. In addition, in some of the cases, the CLT for the difference $Z(y)^n_t - y^2 Z(1)^n_t$ becomes degenerate and we derive a higher-order CLT (joint with the CLT for $Z(y)^n_t$). We summarize these limit results in two different theorems, corresponding to the cases $\kappa_1 = 0$ and $\kappa_1 = 1$.

Before stating them, let us recall that a sequence U_n of R^q-valued variables on $(\Omega, \mathcal{F}, \mathbb{P})$ converges $\mathcal{F}_\infty$-stably in law to a limit U if the variable U is defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$ (that is, $\widetilde{\Omega} = \Omega \times \Omega'$ and $\widetilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}'$ for some extra measurable space $(\Omega', \mathcal{F}')$, and $\widetilde{\mathbb{P}}$ is a probability measure on $(\widetilde{\Omega}, \widetilde{\mathcal{F}})$ such that $\widetilde{\mathbb{P}}(A \times \Omega') = \mathbb{P}(A)$ for all $A \in \mathcal{F}$), and if $\mathbb{E}(f(U_n)Y) \to \widetilde{\mathbb{E}}(f(U)Y)$ for any bounded continuous function f on $\mathbb{R}^q$ and any bounded $\mathcal{F}_\infty$-measurable Y.

Concerning the tuning parameters, we need a set of conditions in the spirit of (4.1), which has the following form, for a sequence $\overline{v}_n$ as described in the theorems below, and for some $\varepsilon > 0$ (arbitrarily small), some integer P (arbitrarily large) and all $m = 1, \dots, M$:

$$
\begin{array}{l}
\frac{\overline{v}_n^2}{k_n^2} \Big(k_n^2 + u_n^{2r} + u_n^2 (h_n \Delta_n)^{\frac{2-\beta_1-\varepsilon}{\beta_1} \bigwedge 1} + \frac{u_n^{(4+\beta_1+\varepsilon) \bigvee 5}}{h_n^2}\Big) \to 0, \\
\frac{\overline{v}_n^2}{k_n^2} \Big(\frac{u_n^8}{h_n^5 \Delta_n} + \frac{u_n^8}{k_n^3 h_n^6 \Delta_n^2} + \frac{u_n^{2\beta_1}}{k_n^3} + u_n^{2\beta_m} (k_n h_n \Delta_n)^{\beta_m} + \frac{u_n^{2P}}{h_n^{P+1} \Delta_n}\Big) \to 0, \\
\kappa_1 = 1 \ \Rightarrow \ \frac{\overline{v}_n^2}{k_n^2} \Big(u_n^8 (h_n \Delta_n)^3 + \frac{u_n^8 (h_n \Delta_n)^2}{k_n^3} + \frac{u_n^8 \Delta_n}{h_n} + u_n^{(2+\beta_1+\varepsilon) \bigvee 3} h_n \Delta_n + \frac{u_n^{12} \Delta_n^2}{h_n}\Big) \to 0, \\
\kappa_2 = 1 \ \Rightarrow \ \frac{\overline{v}_n^2 u_n^4}{k_n h_n^3 \Delta_n} \to 0.
\end{array}
\tag{4.2}
$$

Finally, $\mathcal{Y}$ below is a fixed *finite* subset of $(0, \infty)$ with cardinal q. We start with a CLT for the case when X does not contain a diffusion.

THEOREM 2. *Assume (H) with $\kappa_1 = 0$, (O) and (N-0) or (N-1), and also (3.2),*

$$
\frac{u_n^{\beta_1} h_n^3 \Delta_n}{u_n^{\beta_1} h_n^3 \Delta_n + u_n^4} \to \eta,
\tag{4.3}
$$

and (4.2) with $\overline{v}_n = v_n$ given by

$$
v_n = k_n \sqrt{\frac{h_n^3 \Delta_n}{u_n^4 + u_n^{\beta_1} h_n^3 \Delta_n}}.
\tag{4.4}
$$

Then for any $t > 0$ the q-dimensional variables $(v_n Z(y)^n_t))_{y \in \mathcal{Y}}$ converge $\mathcal{F}_\infty$-stably in law to a variable $(Z(y)_t)_{y \in \mathcal{Y}}$ defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$, which conditionally on $\mathcal{F}$ is centered Gaussian with variance-covariance given by (recall (3.7) for $\psi_\beta(y, y')$):

$$
\widetilde{\mathbb{E}}(Z(y)_t Z(y')_t \mid \mathcal{F}) = \int_0^t \big(\eta \psi_{\beta_1}(y, y') \widetilde{\phi}^{(\beta_1)} \chi(\beta_1) a_s^1 \lambda_s + (1 - \eta) y^2 y'^2 \overline{\psi}^2 \gamma_s^2\big) \frac{1}{\lambda_s} \, ds.
\tag{4.5}
$$

When $\eta = 0$ the above CLT is driven by the noise, when $\eta = 1$ the CLT is determined by the jump component of X, and when $\eta \in (0, 1)$ both the jumps and the noise drive the limit.

The above theorem gives a CLT for the differences $Z(y)^n_t - y^2 Z(1)^n_t$, with a non-degenerate limit as soon as $\eta > 0$ in (4.3). On the other hand, if $\eta = 0$ the limit for this differences is degenerate and the proper rate should be $\overline{v}_n = k_n \sqrt{\frac{h_n^5 \Delta_n}{u_n^8}}$. However, in this case (4.2) cannot be fulfilled. Therefore, it is not clear whether in this case a genuine CLT for the differences $Z(y)^n_t - y^2 Z(1)^n_t$ does exist.

We next state a CLT for the case when X can contain a diffusion. To state the result, we introduce the two rates:

$$v_n = k_n \sqrt{\frac{h_n^3 \Delta_n}{u_n^4 (1 + h_n^2 \Delta_n)^2 + u_n^{\beta_1} h_n^3 \Delta_n}}, \qquad v'_n = \frac{k_n}{u_n^{\beta_1/2}}, \tag{4.6}$$

Clearly $v_n / v'_n \le K$, so (4.2) with $\overline{v}_n = v'_n$ implies (4.2) with $\overline{v}_n = v_n$.

Theorem 3. *Assume (H) with $\kappa_1 = 1$, (O) and (N-0) or (N-1), and also (3.2) and*

$$\frac{u_n^{\beta_1} h_n^3 \Delta_n}{u_n^{\beta_1} h_n^3 \Delta_n + u_n^4 (1 + h_n^2 \Delta_n)^2} \to \eta, \qquad \frac{h_n^2 \Delta_n}{1 + h_n^2 \Delta_n} \to \eta'. \tag{4.7}$$

a) Under (4.2) with $\overline{v}_n = v_n$, for any $t > 0$ the q-dimensional variables $(v_n Z(y)^n_t)_{y \in \mathcal{Y}}$ converge $\mathcal{F}_\infty$-stably in law to a variable $(Z(y)_t)_{y \in \mathcal{Y}}$ defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$, which conditionally on $\mathcal{F}$ is centered Gaussian with variance-covariance given by

$$\begin{aligned} \widetilde{\mathbb{E}}(Z(y)_t Z(y')_t \mid \mathcal{F}) = \textstyle\int_0^t \big(& \eta \psi_{\beta_1}(y, y') \widetilde{\phi}^{(\beta_1)} \chi(\beta_1) a_s^1 \lambda_s \\ & + (1 - \eta) y^2 y'^2 \big(\eta' \phi c_s \lambda_s + (1 - \eta') \overline{\phi} \gamma_s \big)^2 \big) \tfrac{1}{\lambda_s} \, ds. \end{aligned} \tag{4.8}$$

b) Under (4.2) with $\overline{v}_n = v'_n$ plus

$$u_n^{8 - \beta_1} \Big(\frac{1}{h_n^5 \Delta_n} + (h_n \Delta_n)^3 \Big) \to 0, \tag{4.9}$$

for any $t > 0$ the $q+1$-dimensional variables $\big(v_n Z(1)^n_t, \big(v'_n (Z(y)^n_t - y^2 Z(1)^n_t)\big)_{y \in \mathcal{Y}}\big)$ converge $\mathcal{F}_\infty$-stably in law to $\big(Z(1)_t, (Z(y)'_t)_{y \in \mathcal{Y}}\big)$, where $Z(1)_t$ and $(Z'(y)_t)_{y \in \mathcal{Y}}$ are defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ and are, conditionally on $\mathcal{F}$, two independent centered Gaussian variables with variances given by (4.8) for $Z(1)_t$ and by

$$\widetilde{\mathbb{E}}(Z'(y)_t \, Z'(y')_t \mid \mathcal{F}) = \overline{\psi}_{\beta_1}(y, y') \widetilde{\phi}^{(\beta_1)} \chi(\beta_1) \, A_t^1. \tag{4.10}$$

Part (a) of the theorem shows that the CLT for $\big(v_n Z(y)^n_t\big)_{y \in \mathcal{Y}}$ can be determined by the diffusion component of X, the jumps and the observation noise (their role in the asymptotic variance is controlled by η and η'). In part (b) of Theorem 3 we present a joint limit result for $\big(v_n Z(1)^n_t, \big(v'_n (Z(y)^n_t - y^2 Z(1)^n_t)\big)_{y \in \mathcal{Y}}\big)$. For this result we need the condition in (4.9) which guarantees that the jump component of X is the leading term for the difference $\big(v'_n (Z(y)^n_t - y^2 Z(1)^n_t)\big)$, i.e., the diffusion component of X and the noise play only an asymptotically negligible role. When this is not the case, similar to Theorem 2, there is no choice of the tuning parameters satisfying (4.2) with $\overline{v}_n = v'_n$ and a sequence v'_n ensuring a nontrivial limit in part (b) of Theorem 3.

Finally, as for the convergence in probability result in Theorem 1, we have a wide range of choices for our tuning parameter that satisfy Theorems 2 and 3. The choice of the tuning parameters can be optimized according to the specific application in mind as we will show in the next section.

4.3. *Applications.* We now illustrate some applications of the developed limit theory. We will focus attention on the estimation of the integrated volatility C_t and the leading jump activity index β_1. These problems have received a lot of attention in recent work. Our theory will allow estimation of C_t and β_1 in more general settings than previously considered and in many of the cases we will be also able to achieve faster rates of convergence than those of existing estimators and even rate efficiency.

We will develop the estimators, derive their rate of convergence, and provide a CLT for them. To make the inference feasible, one will need consistent estimates of the asymptotic variances of the estimators. Such estimates are relatively easy to derive using Theorem 1 (and consistent estimators for the variance of the noise), and for brevity we will not provide explicit expressions for them. In addition, the optimal choice for the tuning parameters u_n, h_n, k_n in many cases will depend on the unknown jump activity index β_1. Therefore, for a feasible estimation, one will need a preliminary estimator of β_1 based on an initial part of the sample of shrinking time span. Again, for brevity we will not further discuss this here, leaving instead the details pertaining to these issues for future applied work based on the theoretical results of the current paper.

4.3.1. *Estimation of β_1.* We start with the estimation of β_1. For the general case when X can contain a diffusion, we first set for $y > 0$:

$$\widehat{C}'(y)^n_t = \widehat{C}(y)^n_t - y^2\widehat{C}(1)^n_t = Z(y)^n_t - y^2 Z(1)^n_t + \frac{2}{k_n}\sum_{m=1}^M (y^{\beta_m} - y^2) u_n^{\beta_m} \widetilde{\phi}_n^{\beta_m} \chi(\beta_m)\, A^m_t. \tag{4.11}$$

Then, observing that the function $f(x) = \frac{4^x - 16}{2^x - 4}$ is C^∞ on the interval $(0,2)$, with a C^∞ reciprocal function f^{-1}, a natural estimator for β_1 is, for example,

$$\widehat{\beta}^{n,1}_t = f^{-1}\Big(\frac{\widehat{C}'(4)^n_t}{\widehat{C}'(2)^n_t}\Big). \tag{4.12}$$

An easy computation shows the consistency $\widehat{\beta}^{n,1}_t \stackrel{\mathbb{P}}{\longrightarrow} \beta_1$ in restriction to the set $\{A^1_t > 0\}$ on which the "component" with index β_1 is present, as soon as we have (4.1) with the sequence $v_n = k_n/u_n^{\beta_1}$. Therefore, we obtain consistent estimators for β_1 on the set $\{A^1_t > 0\}$ as soon as the tuning parameters h_n, k_n, u_n satisfy (3.2) and

$$\begin{array}{l} \frac{k_n}{u_n^{\beta_1}} + \frac{u_n^{4-2\beta_1}}{h_n^3\Delta_n} + \frac{u_n^{8-2\beta_1}}{k_n^3 h_n^6\Delta_n^2} + \frac{u_n^{8-2\beta_1}}{h_n^8\Delta_n^2} + u_n^{2-2\beta_1} h_n\Delta_n \to 0, \\ \kappa_1 = 1 \;\Rightarrow\; u_n^{4-2\beta_1} h_n\Delta_n + \frac{u_n^{8-2\beta_1}(h_n\Delta_n)^2}{k_n^3} \to 0, \qquad\qquad \kappa_2 = 1 \;\Rightarrow\; \frac{k_n u_n^{4-2\beta_1}}{h_n^3\Delta_n} \to 0. \end{array} \tag{4.13}$$

There is a wide range of tuning parameters achieving the above condition, provided we know that β_1 is strictly bigger than some known number $\alpha \in (0,1]$. For example, one may choose the integers h_n in such a way that $h_n^{11}\Delta_n^8 \to 0$ and $\inf_n\, h_n^2\Delta_n > 0$, and then $u_n = (h_n^3\Delta_n)^{1/8}$ and $k_n = [u_n^\alpha]$. With this choice we have consistency, and a single tuning parameter, regardless of whether κ_1 and κ_2 equal 0 or 1.

If we further know that $\kappa_1 = 0$, i.e., that X does not contain a diffusion, we do not need to use the differences $\widehat{C}'(y)^n_t$ but rather we can use directly $\widehat{C}(y)^n_t$. In particular, in this case, another sequence of estimators, which are consistent on the set $\{A^1_t > 0\}$, is naturally given by

$$\widehat{\beta}'^{n,1}_t = \frac{1}{\log 2}\,\log\Big(\frac{\widehat{C}(2)^n_t}{\widehat{C}(1)^n_t}\Big). \tag{4.14}$$

We note that given the above estimates of β_1, we readily get an estimate of A_t^1 using $\widehat{C}'(y)_t^n$ or $\widehat{C}(y)_t^n$. Hence the analysis of the estimation of A_t^1 is similar to that of the estimation of β_1 and for brevity is not discussed further.

We turn next to the rate of convergence of the estimation of β_1 and an associated CLT that can allow quantifying estimation uncertainty. For simplicity, we restrict attention to the typical case of $M = 1$ and $r < \beta_1/2$ (or, equivalently for what follows, $M \geq 2$ and $\beta_2 < \beta_1/2$).

Concerning the estimator in (4.14), which works only when $\kappa_1 = 0$, the joint convergence of $v_n Z(y)_t^n$ for $y = 1, 2$ to a non-degenerate limit is enough: we can apply Theorem 2 and the rate of convergence is $u_n^{\beta_1/2}$ (this explains the restrictions $r < \beta_1/2$ or $\beta_2 < \beta_1/2$), so we need to "maximize" u_n, of course within the constraints (3.2), (4.3) and (4.2). Actually, we cannot really achieve "best" rate, but only the best one up to some arbitrarily small $\varepsilon > 0$, as exhibited in the next result, which readily follows from Theorem 2 upon using the Delta method, and in which (4.3) gives us $\eta = 1$. Below, $a_n \asymp b_n$ means that both sequence a_n/b_n and b_n/a_n are bounded.

Theorem 4. *Assume (O) and (H) with $\kappa_1 = 0$ and either $M = 1$ and $r < \beta_1/2$ or $M \geq 2$ and $\beta_2 < \beta_1/2$. Let $t > 0$ and also $\varepsilon \in (0, 1/5)$ be arbitrarily small. Assume also either one of the following two hypotheses:*

(i) we have (N-0) (so $\kappa_2 = 0$) and the tuning parameters satisfy

$$(4.15)\qquad \begin{array}{lllll} \beta_1 \geq \frac{3}{2} & \Rightarrow & h_n \asymp \Delta_n^{-\frac{2}{2+\beta_1}}, & u_n \asymp \Delta_n^{-\frac{1-\varepsilon}{2+\beta_1}}, & k_n \asymp \Delta_n^{-\frac{\beta_1}{6+3\beta_1}}, \\ \frac{3}{4} \leq \beta_1 \leq \frac{3}{2} & \Rightarrow & h_n \asymp \Delta_n^{-\frac{15-2\beta_1}{21}}, & u_n \asymp \Delta_n^{-\frac{2(1-\varepsilon)}{7}}, & k_n \asymp \Delta_n^{-\frac{2\beta_1}{21}}, \\ \beta_1 \leq \frac{3}{4} & \Rightarrow & h_n \asymp \Delta_n^{-\frac{3-\beta_1}{5-2\beta_1}}, & u_n \asymp \Delta_n^{-\frac{1-\varepsilon}{5-2\beta_1}}, & k_n \asymp \Delta_n^{-\frac{\beta_1}{15-6\beta_1}}. \end{array}$$

(ii) we have (N-1) (so $\kappa_2 = 1$) and the tuning parameters satisfy

$$(4.16)\qquad \begin{array}{lllll} \beta_1 \geq \frac{3}{2} & \Rightarrow & h_n \asymp \Delta_n^{-\frac{12+\beta_1}{12+7\beta_1}}, & u_n \asymp \Delta_n^{-\frac{6(1-\varepsilon)}{12+7\beta_1}}, & k_n \asymp \Delta_n^{-\frac{2\beta_1}{12+7\beta_1}}, \\ \frac{3}{4} \leq \beta_1 \leq \frac{3}{2} & \Rightarrow & h_n \asymp \Delta_n^{-\frac{15-\beta_1}{21+\beta_1}}, & u_n \asymp \Delta_n^{-\frac{6(1-\varepsilon)}{21+\beta_1}}, & k_n \asymp \Delta_n^{-\frac{2\beta_1}{21+\beta_1}}, \\ \beta_1 \leq \frac{3}{4} & \Rightarrow & h_n \asymp \Delta_n^{-\frac{18-5\beta_1}{30-11\beta_1}}, & u_n \asymp \Delta_n^{-\frac{6(1-\varepsilon)}{30-11\beta_1}}, & k_n \asymp \Delta_n^{-\frac{2\beta_1}{30-11\beta_1}}. \end{array}$$

Then the sequence $u_n^{\beta_1/2}(\widehat{\beta}_t^{\prime n,1} - \beta_1)$ converges stably in law, in restriction to the set $\{A_t^1 > 0\}$, to a variable which is defined on an extension of the probability space and which, conditionally on $\mathcal{F}$, is centered Gaussian with variance

$$(4.17)\qquad \frac{16 + 28 \cdot 2^{\beta_1} + 16 \cdot 3^{\beta_1} - 17 \cdot 4^{\beta_1}}{4(\log 2)^2\, \widetilde{\phi}^{(\beta_1)}\, \chi(\beta_1)\, A_t^1}.$$

We turn next to the case when X can contain a diffusion, so we use the estimator (4.12) and Theorem 3, and the rate is again $u_n^{\beta_1/2}$. Exactly as before we cannot fully achieve the best possible rate. Using again the Delta method, we arrive at the following result (in case (i) below we have $\eta' \in (0, 1)$ and $\eta = 0$, and we use part (b) of Theorem 3; in case (ii) we have $\eta' = \eta = 1$, and therefore we use part (a) of Theorem 3).

Theorem 5. *Assume (O) and (H) with $\kappa_1 = 1$ and either $M = 1$ and $r < \beta_1/2$ or $M \geq 2$ and $\beta_2 < \beta_1/2$. Let $t > 0$. Assume also either one of the following two hypotheses:*

(i) we have (N-0) (so $\kappa_2 = 0$) and for some $\varepsilon \in \big(0, \frac{1}{8} - \beta_1 \wedge (2-\beta_1)\big)$ the tuning parameters satisfy

$$\begin{array}{lll} \beta_1 \geq \frac{16}{11} \quad \Rightarrow \quad h_n \asymp \frac{1}{\sqrt{\Delta_n}}, & u_n \asymp \Delta_n^{-\frac{3(1-\varepsilon)}{16-2\beta_1}}, & k_n \asymp \Delta_n^{-\frac{3\beta_1(1-2\varepsilon)}{32-4\beta_1}}, \\ \beta_1 \leq \frac{16}{11} \quad \Rightarrow \quad h_n \asymp \frac{1}{\sqrt{\Delta_n}}, & u_n \asymp \Delta_n^{-\frac{2(1-\varepsilon)}{16-5\beta_1}}, & k_n \asymp \Delta_n^{-\frac{\beta_1(1-2\varepsilon)}{16-5\beta_1}}. \end{array} \tag{4.18}$$

(ii) we have (N-1) (so $\kappa_2 = 1$) and for some $\varepsilon \in (0, 1/5)$ the tuning parameters satisfy

$$h_n \asymp \Delta_n^{-\frac{24-5\beta_1}{48-11\beta_1}}, \qquad u_n \asymp \Delta_n^{-\frac{6(1-\varepsilon)}{48-11\beta_1}}, \qquad k_n \asymp \Delta_n^{-\frac{2\beta_1}{48-11\beta_1}}. \tag{4.19}$$

Then the sequence $u_n^{\beta_1/2}(\widehat{\beta}_t^{n,1} - \beta_1)$ converges stably in law, in restriction to the set $\{A_t^1 > 0\}$, to a variable which is defined on an extension of the probability space and which, conditionally on $\mathcal{F}$, is centered Gaussian with variance

$$\frac{1}{4^{1+\beta_1}(\log 2)^2\, \widetilde{\phi}^{(\beta_1)}\, \chi(\beta_1)\, A_t^1} \Big(\frac{\overline{\psi}_{\beta_1}(4,4)}{(16-4^{\beta_1})^2} + \frac{\overline{\psi}_{\beta_1}(2,2)}{(4-2^{\beta_1})^2} - \frac{2\overline{\psi}_{\beta_1}(2,4)}{(16-4^{\beta_1})(4-2^{\beta_1})} \Big). \tag{4.20}$$

These two theorems are not directly applicable for three reasons. One is that we need consistent estimators for the conditional variances, and this could easily be taken care of. The second reason is somewhat more important: the choice of the tuning parameters in the various sets of conditions above depends on Δ_n which is not observable, so we will need to replace Δ_n by $1/N_t^n$ (making use of (2.13)). The third reason is that those conditions also depend on the unknown value β_1 and hence a preliminary estimate for it is needed. As mentioned before, we leave these practical considerations for a follow-up work.

We finish this section with a brief discussion of the achievable rates of convergence for estimating β_1. We start with the case of no diffusion ($\kappa_1 = 0$, so Theorem 4 applies) and the stronger assumption (N-0) for the noise. As a benchmark, we note that in a parametric model where X is a β_1-symmetric stable process and the noise is i.i.d. Gaussian, using empirical characteristic function, we can estimate β_1 at the rate $\Delta_n^{-\frac{\beta_1(1-\epsilon)}{4+2\beta_1}}$. Our estimator can achieve this parametric rate when $\beta_1 \geq 3/2$. For lower values of β_1, the achievable rate in our nonparametric setting drops. This is due to the effect from the presence of the drift term in X, the variation of the characteristics of X as well as the generality of our sampling scheme. Comparing the cases $\kappa_2 = 1$ and $\kappa_2 = 0$, when there is no diffusion, we notice that the weaker assumption for the noise slows down the rate of convergence. This effect is pretty small for high levels of β_1 (less than 10% loss in rate of convergence for $\beta_1 \geq 3/2$) and more significant for low values of β_1. Finally, we can compare the rate of convergence of our estimator of β_1 in the no diffusion setting with the one based on power variations in [14]. The rate of convergence for the latter is derived for $\beta_1 > \sqrt{2}$ and the best possible is $\Delta_n^{-\frac{\beta_1(1-\epsilon)}{(2\beta_1+8)}}$. This is much slower than the one achievable for our estimator $\widehat{\beta}_t'^{n,1}$.

Turning to the case when X can contain a diffusion, we can see that, as expected, the rate of convergence of the estimator drops. Focusing on the case of $\kappa_2 = 0$, we note that the loss of rate efficiency compared to the no diffusion case is relatively small for high levels of β_1: it is 19% for $\beta_1 = 3/2$ and it approaches 0% for β_1 approaching 2. To the best of our knowledge, the rate of convergence of estimators of β_1 in the simultaneous presence of diffusion and noise have not been analyzed thus far.

4.3.2. *Estimation of C_t.* We continue with the estimation of C_t and of course we assume $\kappa_1 = 1$. Consistent estimators of C_t are easy to construct. We take for example $y = 1$, and rewrite (3.8) as

$$\widehat{C}^n_t = C_t + R^n_t + S^n_t, \qquad \text{where}$$
$$\widehat{C}^n_t = \tfrac{2k_n}{u_n^2\phi_n}\,\widehat{C}(1)_t, \quad R^n_t = \textstyle\sum_{m=1}^M u_n^{\beta_m-2}\,\frac{4\widetilde{\phi}_n^{\beta_m}\chi(\beta_m)}{\phi_n}\,A^m_t, \quad S^n_t = \tfrac{2k_n}{u_n^2\phi_n}\,Z(1)^n_t. \tag{4.21}$$

Since $u_n \to \infty$ we have $R^n_t \xrightarrow{\mathbb{P}} 0$, hence as soon as $S^n_t \xrightarrow{\mathbb{P}} 0$ the statistics $\widehat{C}^n_t$ above are consistent estimators for C_t. In view of Theorem 1, this holds as soon as the sequence $v_n = k_n/u_n^2$ satisfies (4.1). Therefore, we have consistency as soon as the tuning parameters h_n, k_n, u_n satisfy (3.2) and

$$\frac{k_n}{u_n^2} + \frac{1}{h_n^3\Delta_n} + \frac{u_n^4}{k_n^3h_n^6\Delta_n^2} \to 0, \qquad \kappa_2 = 1 \;\Rightarrow\; \frac{k_n}{h_n^3\Delta_n} \to 0. \tag{4.22}$$

There is a wide range of tuning parameters achieving this. For example, we may choose the integers h_n in such a way that $h_n^3\Delta_n^2 \to 0$ and $h_n^5\Delta_n^2 \to \infty$, and then $u_n = h_n^{1/4}$ and $k_n = [u_n]$. This way we have consistency while using a single tuning parameter.

Concerning rates of convergence and an associated CLT, things are different. Let us first mention that, when X is continuous and the noise is an additive Gaussian white noise and sampling is regular, we know that the optimal rate for estimating C_t is $1/\Delta_n^{1/4}$: so this rate is a natural benchmark.

This optimal rate is achieved by the estimator $\widehat{C}^n_t$ only when $\beta_1 < 1$ (which implies that the bias term R^n_t in (4.21) is negligible at this rate) and $\kappa_2 = 0$. In the case when $\beta_1 < 1$ but $\kappa_2 = 1$, i.e., when the weaker assumption for the noise holds only, the rate of convergence of $\widehat{C}^n_t$ drops slightly. This result is a trivial application of Theorem 3-(a), with $\eta = 0$ and $\eta' = 1/2$ in case (a) and $\eta = 0$ and $\eta' = 1$ in case (b) and is given in the following theorem.

THEOREM 6. *Assume (O) and (H) with $\kappa_1 = 1$ and $\beta_1 < 1$.*

a) If (N-0) holds (so $\kappa_2 = 0$) and if the tuning parameters satisfy for some $\varepsilon \in (0, 1/12]$ and all $\varepsilon' > 0$:

$$h_n \asymp \frac{1}{\sqrt{\Delta_n}}, \qquad u_n = \frac{u'_n}{\Delta_n^{1/4}} \;\; \text{with} \;\; u'_n \to 0,\; u'_n\Delta_n^{-\varepsilon'} \to \infty, \qquad \Delta_n^{-\frac{1}{6}-\varepsilon} \le k_n \le \Delta_n^{-\frac{1}{4}+\varepsilon}, \tag{4.23}$$

the sequence $\Delta_n^{-1/4}(\widehat{C}^n_t - C_t)$ converges stably in law to a variable which is defined on an extension of the probability space and which, conditionally on $\mathcal{F}$, is centered Gaussian with variance

$$4\int_0^t \big(c_s\lambda_s + \frac{\overline{\phi}}{\phi}\gamma_s\big)^2\,\frac{1}{\lambda_s}\,ds. \tag{4.24}$$

b) If (N-1) holds (so $\kappa_2 = 1$) and if the tuning parameters satisfy for some $\varepsilon \in (0, 2(\beta_1 \wedge (1-\beta_1))$

$$h_n \asymp \Delta_n^{-\frac{12-5\beta_1+\varepsilon}{24-11\beta_1}}, \qquad u_n = \Delta_n^{-\frac{3}{24-11\beta_1}}, \qquad k_n \asymp \Delta_n^{-\frac{2\beta_1+\varepsilon}{24-11\beta_1}}, \tag{4.25}$$

the sequence $\Delta_n^{-\frac{12-6\beta_1-\varepsilon}{48-22\beta_1}}(\widehat{C}^n_t - C_t)$ converges stably in law, in restriction to the set $\{A^1_t > 0\}$, to a variable which is defined on an extension of the probability space and which, conditionally on $\mathcal{F}$, is centered Gaussian with variance

$$\frac{4\overline{\phi}^2}{\phi^2}\int_0^t \gamma_s^2\,\frac{1}{\lambda_s}\,ds. \tag{4.26}$$

In part (b) above, the rate is always faster than $\Delta_n^{-3/13}$, and approaches the optimal rate $\Delta_n^{-1/4}$ as β_1 becomes close to 0. Therefore, the loss of efficiency due to the weaker assumption for the noise is at most 8% in terms of rate of convergence. We can also observe that we have exactly the rate $\Delta_n^{-3/13}$, irrespective of the value of β_1 in $(0,1)$, if instead of (4.25) we take

$$h_n \asymp \Delta_n^{-\frac{7}{13}}, \qquad u_n = \Delta_n^{-\frac{3}{13}}, \qquad k_n \asymp \Delta_n^{-\frac{2}{13}}. \tag{4.27}$$

Now we turn to the case $\beta_1 \geq 1$. In this situation the bias term R_t^n in (4.21) is no longer negligible and we need to de-bias our estimators. We will restrict our attention to the case $M = 1$ and, similar to [12], we can use

$$\widehat{C}_t'^n = \frac{2k_n}{u_n^2 \phi_n} \Big(\widehat{C}(1)_t^n - \frac{(\widehat{C}(2)_t^n - 4\widehat{C}(1)_t^n)^2}{\widehat{C}(4)_t^n - 8\widehat{C}(2)_t^n + 16\widehat{C}(1)_t^n} \Big). \tag{4.28}$$

Then we need to use part (b) of Theorem 3, with $\eta' = 1/2$ in case (a) and $\eta' = 0$ in case (b) below, and always $\eta = 0$.

THEOREM 7. *Assume (O) and (H) with $\kappa_1 = 1$ and $r < \beta_1/2$ and either $M = 1$ or $\beta_2 < \beta_1/2$.*

a) Under the assumption (i) of Theorem 5, the sequence $\Delta_n^{-1/4}(\widehat{C}_t'^n - C_t)$ converges stably in law to a variable which is defined on an extension of the probability space and which, conditionally on $\mathcal{F}$, is centered Gaussian with variance given by (4.24).

b) Under the assumption (ii) of Theorem 5, the sequence $\Delta_n^{-\frac{24-6\beta_1}{96-22\beta_1}}(\widehat{C}_t^n - C_t)$ converges stably in law to a variable which is defined on an extension of the probability space and which, conditionally on $\mathcal{F}$, is centered Gaussian with variance given by (4.26).

The results of the above theorem hold irrespective of whether β_1 is smaller or bigger than 1, and the rate in case (b) of Theorem 7 is faster than the rate for $\widehat{C}_t^n$ in part (b) of Theorem 6 when $\beta_1 < 1$, but of course we need the additional assumptions $r < \beta_1/2$ and either $M = 1$ or $\beta_2 < \beta_1/2$ for Theorem 7. Note also that under (N-1) and upon making the choice (4.27) for the tuning parameter, we also have the convergence of $\Delta_n^{-3/13}(\widehat{C}_t'^n - C_t)$ to exactly the same limit as above.

5. Proofs.

We begin with the following lemma:

LEMMA 8. *If X satisfies (H'), it also satisfies (H).*

Proof. We assume (H'). Observe that (H'-iii) implies that β_m is the Blumenthal-Getoor index of F^m, and $\int (|x|^{r'} \wedge 1)\, F^m(dx) < \infty$ for any $r' > \beta_m$.

Let $\underline{p}^m$ be the jump measure of the Lévy process Z^m. This is a Poisson random measure with compensator $\underline{q}^m(dt, dx) = dt \otimes F^m(dx)$, and by hypothesis the $\underline{p}^m$ are independent when m varies, and also independent of W and $\underline{p}$. We aggregate the measures $\underline{p}$ and the $\underline{p}^m$'s as follows: we replace the space E by the union $\overline{E}$ of E and M copies $E_1, \ldots, E_M$ of $\mathbb{R}\backslash\{0\}$ (another Polish space), and set $\overline{\underline{p}}(A) = \underline{p}(A \cap E) + \sum_{m=1}^m \underline{p}^m(A \cap E_m)$ for any Borel subset A of $\overline{E}$. This is a new Poisson random measure, with compensator $\overline{\underline{q}}(dt, dz) = dt \otimes \overline{\eta}(dz)$, where $\overline{\eta}(A) = \eta(A)$ when $A \subset E$ and $\overline{\eta}(A) = F_m(A)$ when $A \subset E_m$ for some m.

We consider the functions f_m and f'_m on $\overline{E}$ defined by

$$f_m(\overline{z}) = \begin{cases} 0 & \text{if } \overline{z} \notin E_m \\ x\, 1_{\{|x| \leq 1\}} & \text{if } \overline{z} = x \in E_m \end{cases}, \qquad f'_m(\overline{z}) = \begin{cases} 0 & \text{if } \overline{z} \notin E_m \\ x\, 1_{\{|x| > 1\}} & \text{if } \overline{z} = x \in E_m \end{cases}.$$

By virtue of (2.8), each Z^m has the representation

$$Z^m_t = (x\,1_{\{|x|\le 1\}}) * (\underline{p}^m - \underline{q}^{\,m})_t + (x\,1_{\{|x|>1\}}) * \underline{p}^m_t = f_m * (\overline{\underline{p}} - \overline{\underline{q}})_t + f'_m * \overline{\underline{p}}_t.$$

Therefore, the processes X and σ of (2.7) and (2.9) can also be written as

$$\begin{aligned} X_t &= X_0 + \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s + \overline{\delta} * (\overline{\underline{p}} - \overline{\underline{q}})_t + \overline{\delta}' * \overline{\underline{p}}_t, \\ \sigma_t &= \sigma_0 + \int_0^t b^\sigma_s\,ds + \int_0^t H^\sigma_s\,dW_s + \int_0^t H'^\sigma_s\,dW'_s + \overline{\delta}^\sigma * (\overline{\underline{p}} - \overline{\underline{q}})_t + \overline{\delta}^\sigma * \overline{\underline{p}}_t, \end{aligned} \tag{5.1}$$

where

$$\overline{\delta}(t,\overline{z}) = \begin{cases} 0 & \text{if } \overline{z} = z \in E \\ \overline{\sigma}^m_{t-}\, f_m(\overline{z}) & \text{if } \overline{z} \in E_m \end{cases}, \qquad \overline{\delta}'(t,\overline{z}) = \begin{cases} \delta'(t,z) & \text{if } \overline{z} = z \in E \\ \overline{\sigma}^m_{t-}\, f'_m(\overline{z}) & \text{if } \overline{z} \in E_m \end{cases},$$
$$\overline{\delta}^\sigma(t,\overline{z}) = \begin{cases} \delta^\sigma(t,z) & \text{if } \overline{z} = z \in E \\ H^{\sigma,m}_{t-}\, f_m(\overline{z}) & \text{if } \overline{z} \in E_m \end{cases}, \qquad \overline{\delta}'^\sigma(t,\overline{z}) = \begin{cases} \delta'^\sigma(t,z) & \text{if } \overline{z} = z \in E \\ H^{\sigma,m}_{t-}\, f'_m(\overline{z}) & \text{if } \overline{z} \in E_m \end{cases}.$$

Letting J_n be an η-integrable bounded function such that (2.6) holds for $\delta', \delta^\sigma, \delta'^\sigma$, it is clear that $\overline{\delta}, \overline{\delta}', \overline{\delta}^\sigma, \overline{\delta}'^\sigma$ satisfy the same for the $\overline{\eta}$-integrable function $\overline{J}_n$ defined by

$$\overline{J}_n(\overline{z}) = \begin{cases} J_n(z) & \text{if } \overline{z} = z \in E, \\ \overline{\alpha}_n \| x|^{r'} \wedge 1) & \text{if } \overline{z} = x \in E_m. \end{cases}$$

for a constant $\overline{\alpha}_n$ depending on the bounds on $\overline{\sigma}^m_t, H^{\sigma,m}_t$ when $t < \tau_n$ (recall that here r' is arbitrary in $(\beta_1, 2)$, implying $\int(|x|^{r'} \wedge 1)\,F^m(dx) < \infty$ for all m). It is also obvious that $\overline{\delta}$ satisfies (H-i).

It remains to prove the existence of a decomposition (2.5), such that F'_t and a^m satisfy (2.6) and (H-ii). The spot Lévy measure F_t of $\overline{\delta} * (\overline{\underline{p}} - \overline{\underline{q}})$ is given by, for any Borel subset A of $\mathbb{R}\backslash\{0\}$:

$$F_t(A) = \sum_{m=1}^M \int_{\{|x|\le 1\}} 1_A(\overline{\sigma}^m_t\, x)\, F^m(dx),$$

hence the symmetrized measures $\breve{F}_t$ and $\breve{F}_m$ satisfy the same relationship. Then (2.5) holds with

$$\begin{aligned} a^m_t &= \tfrac{\alpha_m}{\beta_m}\, |\overline{\sigma}^m_t|^{\beta_m}, \\ F'_t(dx) &= \sum_{m=1}^M \Big(\int 1_A(\overline{\sigma}^m_t x)\, F'^m(dx) + \int_{\{1<|x|\le 1/|\overline{\sigma}^m_t|\}} \tfrac{\beta_m a^m_t}{|x|^{1+\beta_m}}\,dx - \int_{\{1/|\overline{\sigma}^m_t|<|x|\le 1\}} \tfrac{\beta_m a^m_t}{|x|^{1+\beta_m}}\,dx \Big). \end{aligned}$$

Our hypothesis on $\overline{\sigma}^m_t$ implies that each a^m_t is càdlàg adapted satisfying (2.6) and (H-ii). Moreover, when $t < \tau_n$ we have $|\overline{\sigma}^m_t| \le \Gamma_n$ and $\overline{|F'^m|}(x) \le \Gamma_n/|x|^r$ for $x > 0$, hence after a simple calculation

$$\overline{|F'_t|}(x) \le \sum_{m=1}^M \Big(\Gamma_n \frac{|\overline{\sigma}^m_t|^r}{x^r} + 2\big(1 + |\overline{\sigma}^m_t|^{\beta_m}\big)\big(1_{\{x\le 1\}} + \frac{1}{x^{\beta_m}}\, 1_{\{x>1\}}\big)\Big) \le \frac{\Gamma'_n}{x^r},$$

for a suitable constant Γ'_n depending on r, Γ_n, M only. So, we have the last part of (2.6). This completes the proof. □

5.1. *Strengthening the assumptions.* Below we take $\widetilde{r}'=1$ when $\beta_1<1$, and $\widetilde{r}'=r'$ otherwise, so in all cases $\widetilde{r}'$ can be used in place of r' in (H), and can be chosen arbitrarily close to β_1 when $\beta_1\geq 1$. The finite set $\mathcal{Y}$ is fixed throughout, and y and y' are always in $\mathcal{Y}$. It is also not a restriction to assume that $\Delta_n\leq \frac12$ for all n.

We introduce the following strengthened assumption (recall κ_2 as defined in (3.1))):

Assumption (SHON). *There is a constant* Γ *such that*

(i) We have (H), (O) and (N-κ_2*) with* $\tau_1\equiv\infty$*, we write* $J=J_1$*, and moreover we have* $|\delta'(.,z)|\leq \Gamma J(z)$ *and* $|\delta^\sigma(.,z)|^2\leq \Gamma J(z)$ *(so* δ',δ^σ *are bounded) and* $\delta'^\sigma\equiv 0$*.*

(ii) We have for all t*:*

$$\Delta_n N^n_t\leq 1+\Gamma t. \tag{5.2}$$

Lemma 9. *If Theorems 1 or 2 or 3 hold under (SHON), they also hold under (H), (O) and (N-*κ_2*).*

Proof. 1) According to the classical localization procedure, based on stopping $(\mathcal{F}_t)$-adapted processes such as X,σ,λ,γ at or strictly before $(\mathcal{F}_t)$-stopping times, it is enough to prove Theorems 1 or 2 or 3 under the Assumption (i) of (SHON). So below we assume (SHON)-(i), which in particular implies $\gamma_t\geq 1/\Gamma$ for a constant Γ.

2) In this step we construct another sequence Φ'^n_i, with the associated sampling scheme $(T'(n,i))$ and counting processes N'^n_t by (2.10) and (2.12), in such a way that we have:

$$\begin{array}{ll}\text{(i)} & \text{this scheme satisfies (O) and (5.2)},\\ \text{(ii)} & \mathbb{P}(B^n_t)\to 1, \quad \text{for all } t, \text{ where } \ B^n_t=\{T'(n,i)=T(n,i) \text{ for all } i \text{ with } T(n,i)\leq t\}.\end{array} \tag{5.3}$$

The construction of Φ'^n_i is as follows:

$$\Phi'^n_i=\begin{cases}\Phi^n_i & \text{if } i\leq \ell_n\\ 1 & \text{if } i>\ell_n,\end{cases} \quad \text{where } \ \ell_n=\inf\big(j\geq 1:\ S^n_j\leq j\Delta_n-1\big), \qquad S^n_j=\Delta_n\sum_{i=1}^j \Phi^n_i.$$

Observing that ℓ_n is a stopping time for the discrete-time filtration $(\mathcal{K}^n_i)_{i\geq 0}$, it is clear that Φ'^n_i satisfies (2.11), with the same constants $\Gamma(p)$, relative to $\mathcal{K}^n_i$, hence *a fortiori* relative to $\mathcal{K}'^n_i=\sigma(\Phi'^n_j:\, j\leq i)$.

Recall that $\lambda_t\geq 1/\Gamma$. Then $T(n,i)=T'(n,i)\geq (i\Delta_n-1)/\Gamma$ for $i<\ell_n$, hence $T'(n,\ell_n)=T(n,\ell_n)\geq ((\ell_n-1)\Delta_n-1)/\Gamma$, whereas $T'(n,\ell_n+j)-T'(n,\ell_n)\geq j\Delta_n/\Gamma$. We then deduce that $T'(n,j)\geq ((j-1)\Delta_n-1)/\Gamma$ for all $j\geq 0$. Since $T'(n,k+1)>t$ implies $N'^n_t\leq k$, we deduce that indeed $N'^n_t\leq k$ as soon as $k\Delta_n>\Gamma t+1$: so indeed (5.3)-(i) holds.

We now turn to (5.3)-(ii). Since $T'(n,i)=T(n,i)$ when $i\leq \ell_n$, whereas $T(n,i)\geq S^n_i/\Gamma$, this is implied by the property

$$t>0 \quad\Rightarrow\quad \mathbb{P}(S^n_{\ell_n}\geq t)\to 1. \tag{5.4}$$

Observe that $S^n_{[t/\Delta_n]}-t=\Delta_n\sum_{i=1}^{[t/\Delta_n]}(\Phi^n_i-1)+(\Delta_n[t/\Delta_n]-t)$. (2.11) implies $\mathbb{E}(\Phi^n_i-1\mid \mathcal{K}^n_{i-1})=0$ and $\mathbb{E}((\Phi^n_i-1)^2\mid \mathcal{K}^n_{i-1})\leq K$ for some constant K, whereas Φ^n_i is $\mathcal{K}^n_i$-measurable, hence a classical argument yields

$$\mathbb{E}\big((S^n_{[t/\Delta_n]}-t)^2\big)\leq 2\Delta_n+\Delta_n^2\mathbb{E}\Big(\sum_{i=1}^{[t/\Delta_n]}(\Phi^n_i-1)^2\Big)\leq 2\Delta_n^2+Kt\Delta_n.$$

We deduce that $S^n_{[t/\Delta_n]} \stackrel{\mathbb{P}}{\longrightarrow} t$, hence $S^n_{[t/\Delta_n]} \stackrel{\text{u.c.p.}}{\Longrightarrow} t$ as well, and thus $\theta_n = \inf\big(t : S^n_{[t/\Delta_n]} \leq t-1\big)$ satisfies $\mathbb{P}(\theta_n \leq t) \to 0$ for any $t>0$. Clearly, $\theta_n = \ell_n\Delta_n$, whereas $S^n_{\ell_n} \geq S^n_{\ell_n-1} > \theta_n - 1 - \Delta_n > \theta_n - 2$. Then $\mathbb{P}(S^n_{\ell_n} \leq t) \leq \mathbb{P}(\theta_n \leq t+2) \to 0$. We thus have (5.4), and (5.3) is proved.

3) In the previous setting we also construct a new noise process as follows. Let $(\rho_i)_{i\geq 1}$ be a sequence of i.i.d. $\mathcal{N}(0,1)$ variables, independent of all $T(n,i)$, $T'(n,i)$, ε^n_i and of $\mathcal{F}_\infty$. Then we set $\varepsilon'^n_i = \varepsilon^n_i$ if $T'(n,i) = T(n,i)$ and $\varepsilon'^n_i = \rho_i$ otherwise, and also $Y'^n_i = X_{T'(n,i)} + \gamma'_{T'(n,i)}\varepsilon'^n_i$. We have $Y'^n_i = Y^n_i$ if $T'(n,i) = T(n,i)$, and otherwise Y'^n_i is a fictitious observation. However, the family (Y'^n_i) satisfies (SHON)-(i) and (5.3), hence also (SHON). Thus, by our hypothesis the variables $Z'(y)^n_t$ constructed in the same way as $Z(y)^n_t$, on the basis of the sequence Y'^n_i and the sampling scheme $T'(n,i)$, satisfy the claims of Theorems 1 or 2 or 3 for any given t.

Since obviously $Z(y)^n_t = Z'(y)^n_t$ for all $y \in \mathcal{Y}$, in restriction to the set B^n_t, whereas $\mathbb{P}(B^n_t) \to 1$, we readily deduce that indeed the variables $Z(y)^n_t$ also satisfy these claims: this completes the proof. □

Below, (SHON) is in force. Recalling $\gamma_t = \gamma'^2_t$, this implies for some constants $\Gamma \geq 1$ (big enough to have (5.2)) and $\Gamma(p)$ and all p we have (recall that if a process V satisfies $(2.3)_{q,q'}$ or $(2.4)_q$ then $1/V$ satisfies the same, as soon as both V and $1/V$ are bounded):

$$
(5.5)\qquad
\begin{array}{l}
|b_t|,\ |\sigma_t|,\ a^m_t,\ |b^\sigma_t|,\ |H^\sigma_t|,\ |H'^\sigma_t|,\ \gamma_t,\ \gamma'_t,\ \gamma''_t,\ \lambda_t,\ 1/\lambda_t \leq \Gamma, \qquad |\gamma^{(p)}_t| \leq \Gamma(p),\\
|\delta(t,z)|^{r'} \leq J(z), \qquad |\delta^\sigma(t,z)|^2 \leq J(z), \qquad |\delta'(t,z)| \leq \Gamma J(z), \qquad 1_{\{\delta'(t,z)\neq 0\}} \leq \Gamma J(z),\\
X_t,\ \sigma_t,\ \lambda_t,\ 1/\lambda_t,\ (a^m_t)^{1/\beta_m},\ \text{and } \gamma_t,\ \gamma'_t \text{ if } \kappa_2 = 0, \quad \text{satisfy } (2.4)_1,\\
X_t,\ b_t,\ \sigma_t,\ H^\sigma_t,\ \frac{\delta(t,z)}{J(z)^{1/r'}},\ \gamma_t,\ \gamma'_t,\ \gamma^{(3)}_t,\ \lambda_t,\ 1/\lambda_t,\ (a^m_t)^{1/\beta_m} \ \text{ satisfy } (2.3)_{2,1},\\
\lambda_t \text{ and } 1/\lambda_t \ \text{ satisfy } (2.3)_{1,1},\\
\overline{F}_t(x) \leq \frac{\Gamma}{x^{\beta_1}}, \qquad \overline{|F'_t|}(x) \leq \frac{\Gamma}{x^r}.
\end{array}
$$

Below, K is a generic constant, changing from line to line, and possibly depending on $r, r', M, \beta_m, \Gamma$, and sometimes on some extra parameter q such as a power or on the set $\mathcal{Y}$, but never on n and the various indices $i, j, \dots$ or variables $u, y, \dots$ which may occur. Analogously, if $U_n = U_n(i, y, \dots)$ and $U'_n(i, y, \dots)$ are two sequences of variables possibly depending on $y \in \mathcal{Y}'$ and on indices $i, \dots$, we write $U_n = \mathrm{O}(U'_n)$, resp. $= \mathrm{o}(U'_n)$, if $U'_n = 0$ implies $U_n = 0$ and U_n/U'_n (with the convention $0/0 = 0$) is bounded uniformly in $n, i, y, \dots$, resp. goes to 0 uniformly in $i, y, \dots$ as $n \to \infty$.

We end this subsection with a general consequence of the properties (2.3) and (2.4) relative to an arbitrary filtration $(\mathcal{L}_t)$:

Lemma 10. *Suppose that a $(\mathcal{L}_t)$-adapted càdlàg process V satisfies $(2.3)_{q,q'}$, resp. $(2.4)_q$, with some constant K for all finite $(\mathcal{L}_t)$-stopping times $T \leq S$. Then we also have $(5.6)_{q,q'}$, resp. $(5.7)_q$, below, with the same constant K, for any pair S, T as above:*

$$
(5.6)\qquad \mathbb{E}\Big(\sup_{s\in[T,S]} |V_S - V_T|^{q'} \mid \mathcal{L}_T\Big) \leq K\,\mathbb{E}\big((S-T)^q \mid \mathcal{L}_T\big),
$$

$$
(5.7)\qquad \big|\mathbb{E}(V_S - V_T \mid \mathcal{L}_T)\big| \leq K\,\mathbb{E}\big((S-T)^q \mid \mathcal{L}_T\big).
$$

Proof. We prove that $(2.3)_{q,q'}$ implies $(5.6)_{q,q'}$ only, the other case being analogous. We fix two $(\mathcal{L}_t)$-stopping times $T \leq S$ and let $Y = \mathbb{E}(\sup_{s\in[T,S]} |V_S - V_T|^{q'} \mid \mathcal{L}_T)$ and $U = \mathbb{E}((S-T)^q \mid \mathcal{L}_T)$. We need to prove that the two $\mathcal{L}_T$-measurable sets $B_+ = \{Y > K\,U\}$ and $B_- = \{-Y > K\,U\}$ have a vanishing probability. Define another stopping time $T'_+ \leq S$ by setting $T'_+ = T$ on B_+ and $T'_+ = S$ on the complement B^c_+. Observe that $\mathbb{E}(\sup_{\in[T'_+,S]} |V_S - V_{T'_+}|^{q'} \mid \mathcal{L}_T)$ vanishes on B^c_+ and equals Y on B_+, hence as soon as $\mathbb{P}(B_+) > 0$ we have

$$\mathbb{E}(\sup_{s\in[T'_+,S]} |V_S - V_{T'_+}|^{q'}) = \mathbb{E}\big(Y\,1_{B_+}\big) > K\,\mathbb{E}\big(U\,1_{B_+}\big) = K\,\mathbb{E}\big((S-T)^q\,1_{B_+}\big) = K\,\mathbb{E}((S-T'_+)^q),$$

which contradicts $(2.3)_{q',q}$. Therefore $\mathbb{P}(B_+) = 0$, and $\mathbb{P}(B_-) = 0$ is proved analogously. □

5.2. *Properties of the sampling scheme.* We first prove (2.13):

LEMMA 11. *We have the convergence (2.13).*

Proof. We use the variables S^n_j of the proof of Lemma 9, in which $\overline{S}^n_t := S^n_{[t\Delta_n]} \stackrel{\text{u.c.p.}}{\Longrightarrow} t$ was proved, and we set

$$\Lambda^n_t = \Delta_n N^n_t, \qquad H^n_t = T(n,[t/\Delta_n]) = \Delta_n \sum_{i=1}^{[t/\Delta_n]} \lambda^n_{i-1}\Phi^n_i.$$

By the subsequence principle, it is enough to prove that any infinite sequence n_k contains a subsequence n'_k such that $\Lambda^{n'_k}_t \to \Lambda_t$ for all t, for all ω outside a null set. Note also that, from any subsequence one can extract a further subsequence such that the convergence $\overline{S}^n_t \to t$ holds, outside a null set again, locally uniformly in time. In other words, it is enough to show that if $\overline{S}^n_t(\omega) \to t$ locally uniformly in t for some given ω, then we have $\Lambda^n_t(\omega) \to \Lambda_t(\omega)$ for all t (then the convergence is automatically locally uniform).

Therefore, below we assume $\overline{S}^n_t(\omega) \to t$ locally uniformly, and omit to mention ω in S^n and also in λ and H^n. The definitions of H^n and $\overline{S}^n$ imply $H^n_t = \int_0^t \lambda_{H^n_{s-}}\,d\overline{S}^n_s$. (5.5) yields $H^n_{t+s} - H^n_t \leq K(\overline{S}^n_{t+s} - \overline{S}^n_t)$, hence by Ascoli's theorem, from any subsequence we can extract a further subsequence n' such that $H^{n'}$ converges locally uniformly to a continuous nondecreasing limit H. Picking any $\varepsilon > 0$, we denote by $t_1 < t_2 < \cdots$ the times at which $t \mapsto \lambda_t$ has a jump of size bigger than ε, and set $A_t = [0,t]\backslash\big(\cup_{i\geq 1}(t_i - \varepsilon, t_i + \varepsilon]\big)$. The modulus of continuity $w_t(\rho)$ of $\lambda^\varepsilon_s = \lambda_s - \sum_{i\geq 1} \Delta\lambda_{t_i}\,1_{\{t_i\leq s\}}$ on $[0,t]$ satisfies $\limsup_{\rho\to 0} w_t(\rho) \leq \varepsilon$, whereas $H^{n'}_s \to H_s$ locally uniformly, so $\limsup_{n'} \sup_{s\in A_t} |\lambda_{H^{n'}_{s-}} - \lambda_{H_s}| \leq \varepsilon$. Thus, for n' large enough, $\big|H^{n'}_t - \int_0^t \lambda_{H_s}\,d\overline{S}^{n'}_s\big| \leq 2\varepsilon\overline{S}^{n'}_t + K\int_{A_t} d\overline{S}^{n'}_s$, which in turn goes to $2\varepsilon\,t + K\int_{A_t} ds \leq K\varepsilon$. Since ε is arbitrarily small, we get $H^{n'}_t - \int_0^t \lambda_{H_s}\,d\overline{S}^{n'}_s \to 0$. Another application of $\overline{S}^n_s \to s$ for all s yields $\int_0^t \lambda_{H_s}\,d\overline{S}^{n'}_s \to \int_0^t \lambda_{H_s}\,ds$. Thus $H_t = \int_0^t \lambda_{H_s}\,ds$, so H is continuous strictly increasing and its inverse H^{-1} is Λ, as defined by (2.13). Therefore H is uniquely determined and the original sequence H^n converges to $H = \Lambda^{-1}$.

Now, the definitions of Λ^n_t and H^n_t imply that they are right-continuous inverses one from the other, hence $\Lambda^n_t \to H^{-1}_t = \Lambda_t$, and the proof is complete. □

We already introduced $(\mathcal{H}^n_t)$, the smallest filtration containing $(\mathcal{F}_t)$ and with respect to which $T(n,i)$ is a stopping time for all $i \geq 0$, and the σ-field $\mathcal{K}^n_\infty$ generated by the variables $(\Phi^n_i : i \geq 1)$. We will also need the filtration $(\overline{\mathcal{H}}^n_t)$ which is the smallest one containing $(\mathcal{F}_t)$ and such that $\mathcal{K}^n_\infty \subset \overline{\mathcal{H}}^n_0$ (below we prove the intuitively obvious fact that $\overline{\mathcal{H}}^n_t$ is bigger than $\mathcal{H}^n_t$).

Unless it vanishes identically, the noise is not measurable with respect to the previous filtration. To accommodate the noise, we define the following σ-fields:

(5.8)

$$\mathcal{G}^n_i = \mathcal{H}^n_{T(n,i)} \bigvee \sigma(\varepsilon^n_j : j < i), \qquad \overline{\mathcal{G}}^n_i = \overline{\mathcal{H}}^n_{T(n,i)} \bigvee \sigma(\varepsilon^n_j : j < i), \qquad \widehat{\mathcal{G}}^n_i = \mathcal{H}^n_\infty \bigvee \sigma(\varepsilon^n_j : j < i),$$

with the conventions $\mathcal{G}_0^n = \mathcal{F}_0$ and $\overline{\mathcal{G}}_0^n = \overline{\mathcal{H}}_0^n$ and $\widehat{\mathcal{G}}_0^n = \overline{\mathcal{H}}_\infty^n$. Note that the pre-averaged variable $\widetilde{Y}_i^n$ is $\mathcal{G}_{i+h_n}^n$-measurable.

LEMMA 12. *a) We have $\mathcal{H}_t^n \subset \overline{\mathcal{H}}_t^n$ and $\mathcal{H}_\infty^n = \overline{\mathcal{H}}_\infty^n$.*

b) Any càdlàg $(\mathcal{F}_t)$-adapted process satisfying $(2.3)_{q,q'}$, or $(2.4)_q$ for all finite $(\mathcal{F}_t)$-stopping times $T \leq S$ satisfies the same for all finite $(\overline{\mathcal{H}}_t^n)$-stopping times $T \leq S$.

c) Any $(\mathcal{F}_t)$-martingale is a $(\overline{\mathcal{H}}_t^n)$-martingale, hence a $(\mathcal{H}_t^n)$-martingale as well.

d) Any integrable $\mathcal{H}_\infty^n$-measurable variable Y satisfies

$$\mathbb{E}(Y \mid \mathcal{G}_i^n) = \mathbb{E}(Y \mid \mathcal{H}_{T(n,i)}^n). \tag{5.9}$$

Proof. First, (c) is a well known result because $(\overline{\mathcal{H}}_t^n)$ is the initial enlargement of $(\mathcal{F}_t)$ by the independent σ-field $\mathcal{K}_\infty^n$, and it is also a trivial consequence of (b).

For (a) we first prove by induction on i that each $T(n,i)$ is a $(\overline{\mathcal{H}}_t^n)$-stopping time. This is obvious when $i=0$, and we have (recalling $\alpha_t \geq 1/\Gamma$)

$$\{T(n,i+1) \leq t\} = \{T(n,i) \leq t\} \cap A, \quad \text{with} \;\; A = \big\{\Phi_{i+1}^n \leq (t - T(n,i))/(\Delta_n \lambda_{T(n,i)})\big\}.$$

If $T(n,i)$ is a $(\overline{\mathcal{H}}_t^n)$-stopping time, and since Φ_{i+1}^n is $\overline{\mathcal{H}}_0^n$-measurable, we have $A \in \overline{\mathcal{H}}_{T(n,i)}^n$ and thus $\{T(n,i+1) \leq t\} \in \overline{\mathcal{H}}_t^n$. This, being true for all t, implies that $T(n,i+1)$ is also a $(\overline{\mathcal{H}}_t^n)$-stopping time. Therefore $\mathcal{H}_t^n \subset \overline{\mathcal{H}}_t^n$ for all t, including $t=\infty$. On the other hand, $\overline{\mathcal{H}}_\infty^n = \mathcal{F}_\infty \bigvee \mathcal{K}_\infty^n$ is obvious, and Φ_i^n is $\mathcal{H}_{T(n,i)}^n$-measurable by (2.10), so $\mathcal{K}_\infty^n \subset \mathcal{H}_\infty^n$. This yields $\mathcal{H}_\infty^n = \overline{\mathcal{H}}_\infty^n$, and (a) is proved.

Before showing (b) we give a description of the $(\overline{\mathcal{H}}_t^n)$-stopping times S. We consider $\overline{\Phi}^n = (\Phi_i^n)_{i\geq 0}$ as an E-valued random variable, with the Polish space $E = \mathbb{R}_+^{\mathbb{N}^*}$ and its Borel σ-field $\mathcal{E}$. Since $\overline{\mathcal{H}}_t^n = \mathcal{F}_t \vee \mathcal{K}_\infty^n$ we have $\{S > t\} = \{(\omega, \overline{\Phi}^n(\omega)) \in \overline{B}_t\}$ for some $\mathcal{F}_t \otimes \mathcal{E}$-measurable subset $\overline{B}_t$ of $\Omega \times E$. Setting $S'(\omega, \overline{\phi}) = \inf(s \in \mathbb{Q}^+ : (\omega, \overline{\phi}) \notin \overline{B}_s)$, so $\{\omega : S'(\omega, \overline{\phi}) \geq t\} = \cap_{s \in \mathbb{Q}\cap[0,t)} \{\omega : (\omega, \overline{\phi}) \in \overline{B}_s\}$ belongs to $\mathcal{F}_t$ for all t, and we readily deduce that

$$S(\omega) = S'(\omega, \overline{\Phi}^n(\omega)), \quad \text{where} \begin{cases} (i) & S' \text{ is } \mathcal{F} \otimes \mathcal{E}\text{-measurable on } \Omega \times E, \\ (ii) & S'(.,\overline{\phi}) \text{ is an } (\mathcal{F}_t)\text{-stopping time for each } \overline{\phi} \in E. \end{cases} \tag{5.10}$$

At this stage, we can prove (b), say in the case of $(2.3)_{q,q'}$, the other case being analogous. Let $T \leq S$ be two finite $(\overline{\mathcal{H}}_t^n)$-stopping times, with which we associate S' and T' as in (5.10). Upon replacing T' by $T' \wedge S'$, we can assume $T' \leq S'$ identically. Let μ be the law of $\overline{\Phi}^n$ (a probability measure on $(E, \mathcal{E})$). By the independence in Assumption (O), we have

$$\begin{aligned} \mathbb{E}\big(\sup_{s\in[T,S]} |V_S - V_T|^{q'}\big) &= \int \mathbb{E}\big(\sup_{s\mathcal{N}[T'(.,\overline{\phi}),S'(.,\overline{\phi})]} |V_s - V_{T'(.,\overline{\phi})}|^{q'}\big)\, \mu(d\overline{\phi})\Big| \\ &\leq K \int \mathbb{E}\big(|S'(.,\overline{\phi}) - T'(.,\overline{\phi})|^q\big)\, \mu(d\overline{\phi}) \;=\; K\mathbb{E}\big(|S-T|^q\big), \end{aligned}$$

where the inequality above follows from $(2.3)_{q,q'}$, applied with the $(\mathcal{F}_t)$-stopping times $S'(.,\overline{\phi})$ and $T'(.,\overline{\phi})$. This proves the claim.

For (d), let Y' and Y'' be the left and right hand sides of (5.9). It is enough to prove that $\mathbb{E}(Y'ZZ') = \mathbb{E}(Y''ZZ')$ for any bounded $\mathcal{H}_{T(n,i)}^n$-measurable Z and $\sigma(\varepsilon_j^n : j < i)$-measurable Z'. When $Z' = \prod_{j=1}^{i-1} f_j(\varepsilon_j^n)$ for bounded Borel functions f_j, we have $\mathbb{E}(Z' \mid \mathcal{H}_\infty^n) = \prod_{j=1}^{i-1} \mathbb{E}(f_j(\varepsilon_i^n) \mid \mathcal{H}_\infty^n)$ by (N), and each $\mathbb{E}(f_j(\varepsilon_i^n) \mid \mathcal{H}_\infty^n)$ is $\mathcal{H}_{T(n,i)}^n$-measurable (use the last part of (2.15)), hence $\mathbb{E}(Z' \mid \mathcal{H}_\infty^n)$ as well. By a density argument it follows that $\mathbb{E}(Z' \mid \mathcal{H}^n_\infty)$ is $\mathcal{H}^n_{T(n,i)}$-measurable for any $\sigma(\varepsilon^n_j : j < i)$-measurable Z'. Therefore

$$\mathbb{E}(Y'ZZ') = \mathbb{E}(YZZ') = \mathbb{E}\big(YZ\,\mathbb{E}(Z' \mid \mathcal{F}_\infty)\big) = \mathbb{E}\big(Y'Z\,\mathbb{E}(Z' \mid \mathcal{H}^n_\infty)\big) = \mathbb{E}(Y'ZZ'),$$

and the claim follows. □

This lemma will be used very often, typically without special mention. Its claim (c), for example, implies that X and σ are semimartingales satisfying (2.1) and (2.2), relative to the filtration $(\overline{\mathcal{H}}^n_t)$, with W, W' being $(\overline{\mathcal{H}}^n_t)$-Brownian motion and $\underline{q}$ still being the $(\overline{\mathcal{H}}^n_t)$-compensator of $\underline{p}$, and the same if $\overline{\mathcal{H}}^n_t$ is substituted with $\mathcal{H}^n_t$. Another application is the following estimate, easily deduced from (2.11) if we condition with respect to $\mathcal{H}^n_{T(n,i)}$, hence true as well if we condition with respect to $\mathcal{G}^n_i$: for all integers $j \geq 1$ and all $p > 0$ we have

$$\mathbb{E}\big((T(n,i+j) - T(n,i))^p \mid \mathcal{G}^n_i\big) \leq K_p (j\Delta_n)^p. \tag{5.11}$$

In particular, in combination with Lemma 10, this yields that, for any càdlàg $(\mathcal{F}_t)$-adapted process V, we have for all $j = 1, \ldots, 2k_nh_n$ (so $j\Delta_n \leq K$) and $p \geq q'$

$$\begin{array}{ll} \mathbb{E}\big(\sup_{s\in[T(n,i),T(n,i+j)]} |V_s - V_{T(n,i)}|^p \mid \mathcal{G}^n_i\big) \leq K(j\Delta_n)^q & \text{if } V \text{ satisfies } (2.3)_{q',q}, \\ \big|\,\mathbb{E}(V_{(T(n,i)+s)\,\bigwedge T(n,i+j)} - V_{T(n,i)} \mid \mathcal{G}^n_i)\big| \leq K(j\Delta_n)^q & \text{if } V \text{ satisfies } (2.4)_q. \end{array} \tag{5.12}$$

This and (5.5) imply the following estimate, uniform in $z \in E$:

$$\mathbb{E}\big(\sup_{s\in[T(n,i),T(n,i+j)]} |\delta(z,s) - \delta(z,T(n,i))|^2 \mid \mathcal{G}^n_i\big) \leq K J(z)^{2/r'} j\Delta_n. \tag{5.13}$$

Moreover, when $x, y \geq 0$, we have $x^{\beta_m} - y^{\beta_m} = \beta_m y^{\beta_m - 1}(x-y) + \mathrm{O}(|x-y|^{\beta_m})$ if $\beta_m \geq 1$ and $x^{\beta_m} - y^{\beta_m} = \mathrm{O}(|x-y| + |x-y|^{\beta_m})$ if $\beta_m < 1$. Therefore, using (5.5) again, we deduce for $p \geq 2$:

$$\begin{array}{l} \mathbb{E}\big(\sup_{s\in[T(n,i),T(n,i+j)]} |a^m_s - a^m_{T(n,i)}|^p \mid \mathcal{G}^n_i\big) \leq K(j\Delta_n)^{(p\beta_m/2)\wedge 1}, \\ \big|\,\mathbb{E}(a^m_{(T(n,i)+s)\wedge T(n,i+j)} - a^m_{T(n,i)} \mid \mathcal{G}^n_i)\big| \leq K(j\Delta_n)^{\beta_m/2}. \end{array} \tag{5.14}$$

5.3. *Estimates - 1.* The estimates (5.12) will not be enough for our purposes, and we proceed to complement them. The setting is somewhat complicated (because of our future needs), and to obtain notation and statements as simple as possible we fix n and i, but it is important to keep in mind that the (varying) constants K or K_p below do *not* depend on n, i.

We have a bounded sequence θ^n_j of numbers, with which we associate the process

$$\Theta_t = \sum_{j=1}^{2h_n-1} \theta^n_j \, 1_{(T(n,i+j-1)),T(n,i+j)]}(t).$$

We denote by $\mathcal{A}^n_i$ the set of all càdlàg $(\mathcal{H}^n_t)$-adapted processes V satisfying $V_t = 0$ for $t \leq T(n,i)$. If $V \in \mathcal{A}^n_i$ and U is a $(\mathcal{H}^n_t)$-local martingale we define the processes (all in $\mathcal{A}^n_i$):

$$L(V)_t = \int_0^t \Theta_s V_s\, ds, \qquad L'(V,U)_t = \int_0^t \Theta_s V_s\, dU_s, \qquad L''(V)_t = \int_0^t \Theta_s\Big(\int_0^s \Theta_v V_v\, dv\Big)\, ds, \tag{5.15}$$

and also (for $0 \leq j \leq \ell$) the variables

$$\overline{V}_{j,\ell} = \overline{V}^{n,i}_{j,\ell} = \sup_{s\in[T(n,i+j),T(n,i+j+\ell)]} |V_s - V_{T(n,i+j)}|.$$

Suppose that we are given nonnegative $\mathcal{H}^n_{T(n,i)}$-measurable variables $\widehat{\Psi}$ and $\overline{\Psi}$ and $\mathcal{H}^n_{T(n,i+j)}$-measurable variables Ψ_j for $j = 1, 2, \ldots$. We let $\overline{\mathcal{P}}(\overline{\Psi})$, $\widehat{\mathcal{P}}(\widehat{\Psi})$, $\mathcal{P}(\Psi_j)$ be the sets of all $V \in \mathcal{A}^n_i$ such that, for all $0 \leq j \leq 2h_n$ and all reals $z \geq 0$ and constants $\overline{K}_z$ with $\overline{K}_1 = 1$, we have:

$$(5.16)\qquad \begin{array}{ll} \text{for } \overline{\mathcal{P}}(\overline{\Psi}): & \big|\mathbb{E}(V_{T(n,i+j)} - V_{T(n,i)} \mid \mathcal{H}^n_{T(n,i)})\big| \leq \overline{\Psi}, \quad j = 1, \ldots, 2h_n, \\ \text{for } \widehat{\mathcal{P}}(\widehat{\Psi}): & \mathbb{E}((\overline{V}_{0,2h_n})^2 \mid \mathcal{H}^n_{T(n,i)}) \leq \widehat{\Psi}, \\ \text{for } \mathcal{P}(\Psi_j): & \mathbb{E}\big(\Delta(n,i+j+1)^z |\overline{V}_{j,j+1}|^2 \mid \mathcal{H}^n_{T(n,i+j)}\big) \leq \overline{K}_z \Psi_j \Delta^z_n, \quad j = 1, \ldots, 2h_n. \end{array}$$

These classes should indeed be indexed by i and n, as well as the process defined just below, but as already written we omit these indices.

There is of course a connection between $\widehat{\Psi}$ and the Ψ_j's, expressed in the following lemma:

LEMMA 13. *If $V \in \mathcal{A}^n_i$ we have for all $p \geq 2$:*

(5.17)

$$\mathbb{E}((\overline{V}_{0,2h_n})^p \mid \mathcal{H}^n_{T(n,i)}) \leq \begin{cases} K_p\, h_n^{p/2-1} \sum_{j=0}^{2h_n-1} \mathbb{E}(|\overline{V}_{j,j+1}|^p \mid \mathcal{H}^n_{T(n,i)}) & \text{if } V \text{ is a } (\mathcal{H}^n_t)\text{-local martingale,} \\ h_n^{p-1} \sum_{j=0}^{2h_n-1} \mathbb{E}(|\overline{V}_{j,j+1}|^p \mid \mathcal{H}^n_{T(n,i)}) & \text{otherwise.} \end{cases}$$

Hence if $V \in \mathcal{P}(\Psi_j)$, we have $V \in \widehat{\mathcal{P}}(\widehat{\Psi})$ with

$$(5.18)\qquad \begin{array}{l} \bullet\ \ \widehat{\Psi} = \sum_{j=0}^{2h_n-1} \mathbb{E}(\Psi_j \mid \mathcal{H}^n_{T(n,i)}), \\ \bullet\ \ \widehat{\Psi} = K \sum_{j=0}^{2h_n-1} \mathbb{E}(\Psi_j \mid \mathcal{H}^n_{T(n,i)}) \quad \text{if } V \text{ is a } (\mathcal{H}^n_t)\text{-local martingale.} \end{array}$$

Proof. The second part of (5.17) follows from $\overline{V}_{0,2h_n} \leq \sum_{j=0}^{2h_n-1} \overline{V}_{j,j+1}$ and Hölder's inequality. When V is a $(\mathcal{H}^n_t)$-martingale, the Burkholder-Gundy inequality for the discrete-time local martingale $(V_{T(n,i+j)})_{j\geq 0}$ and Hölder's inequality imply

$$\begin{aligned} \mathbb{E}(\sup_{0\leq j\leq 2h_n-1} |V_{T(n,i+j)}|^p \mid \mathcal{H}^n_{T(n,i)}) &\leq \mathbb{E}\Big(\Big(\sum_{j=0}^{2h_n-1} (\overline{V}_{j,j+1})^2\Big)^{p/2} \mid \mathcal{H}^n_{T(n,i)}\Big) \\ &\leq K_p\, h_n^{p/2-1} \sum_{j=0}^{2h_n-1} \mathbb{E}\big(|\overline{V}_{j,j+1}|^p \mid \mathcal{H}^n_{T(n,i)}\big), \end{aligned}$$

whereas $\overline{V}_{0,2h_n} \leq \sup_{0\leq j\leq 2h_n-1} \big(|V_{T(n,i+j)}| + \overline{V}_{j,j+1}\big)$, hence

$$(\overline{V}_{0,2h_n})^p \leq 2^{p-1} \sup_{0\leq j\leq 2h_n-1} |V_{T(n,i+j)}|^p + 2^{p-1} \sum_{j=0}^{2h_n-1} (\overline{V}_{j,j+1})^p,$$

and the first part of (5.17) follows. The last claim is obvious (take $p = 2$ above). □

Next, we give some criteria for a process $V \in \mathcal{A}^n_i$ to belong to these classes.

LEMMA 14. *a) If V' is càdlàg $(\mathcal{F}_t)$-adapted and satisfies (2.3)$_{2,q'}$, then $V_t = V'_t - V'_{t\wedge T(n,i)}$ belongs to $\widehat{\mathcal{P}}(\widehat{\Psi}) \cap \mathcal{P}(\Psi_j)$, with $\widehat{\Psi} = K(h_n\Delta_n)^q$ and $\Psi_j = K(\Delta_n)^q$.*

b) If V' is càdlàg $(\mathcal{F}_t)$-adapted and satisfies (2.4)$_q$, then $V_t = V'_t - V'_{t\wedge T(n,i)}$ belongs to $\overline{\mathcal{P}}(\overline{\Psi})$, with $\overline{\Psi} = K(h_n\Delta_n)^q$.

c) For $w = 1,2$ let $Y^w \in \mathcal{A}^n_i$ be a square-integrable martingale for the filtration $(\mathcal{H}^n_{T(n,i)} \bigvee \mathcal{F}_t)$, with predictable brackets $\langle Y^w, Y^w\rangle_t = \int_0^t \alpha^w_s\,ds$ with α^w_s bounded. Then if $M^w_t = \int_0^t \Theta^w_s\,dY^w_s$, where $\Theta^1_t = \Theta_t$ and $\Theta^2_t = \Theta'_t$, the product $V = M^1M^2$ belongs to $\widehat{\mathcal{P}}(\widehat{\Psi}) \cap \mathcal{P}(\Psi_j)$ with

$$\widehat{\Psi} = K(h_n\Delta_n)^2, \qquad \Psi_j = K\Delta_n\big(|\overline{M}^1_{0,j}|^2 + |\overline{M}^2_{0,j}|^2 + \Delta_n\big). \tag{5.19}$$

Proof. (b) and the claim $V \in \widehat{\mathcal{P}}(\widehat{\Psi})$ in (a) readily follow from (5.12) for V'. In view of (2.10), $\Delta(n,i+j+1)$ is $\overline{\mathcal{H}}^n_{T(n,i+j)}$-measurable, so by Lemmas 10 and 12 (5.12) for V' implies $\mathbb{E}(\overline{V}^2_{j,j+1} \mid \overline{\mathcal{H}}^n_{T(n,i+j)}) \leq K\Delta(n,i+j+1)^q$ in the case of (a). It follows that $V \in \mathcal{P}(\Psi_j)$ with $\Psi_j = K(\Delta_n)^q$.

Now we start the proof of (c). Observe that, under a regular version of the $\mathcal{H}^n_{T(n,i)}$-conditional probability, the new sampling scheme $T'(n,i) = T(n,i+j)$ for $j \geq 0$ satisfies (O) for the filtration $\mathcal{F}'_t = \mathcal{H}^n_{T(n,i)} \bigvee \mathcal{F}_{T(n,i)+t}$. Thus Lemma 12 implies that $Y^w_{T(n,i)+t}$ is a square-integrable martingale for $(\mathcal{H}^n_{T(n,i)+t})$ and for $(\overline{\mathcal{H}}^n_{T(n,i)+t})$. Since $Y^w_t = 0$ when $t \leq T(n,i)$, it follows that Y^w, hence M^w as well, are square-integrable martingales for $(\mathcal{H}^n_t)$ and for $(\overline{\mathcal{H}}^n_t)$. By Itô's formula,

$$V_t = V(1)_t + V(2)_t + V(3)_t, \qquad V(1)_t = \int_0^t M^1_s\,dM^2_s, \quad V(2)_t = \int_0^t M^2_s\,dM^1_s, \quad V(3)_t = \langle M^1, M^2\rangle_t,$$

and it suffices to prove the result for each $V(k)$. This is obvious for $V(3)$, because this process is absolutely continuous with a bounded density, so $\overline{V(3)}_{j,k} \leq K\big(T(n,i+k) - T(n,i+j)\big)$ for $j < k$ and because of (5.11).

Next, Doob's inequality and the boundedness of α^w_t and Θ^w_t and (5.11) imply first that $\mathbb{E}((\overline{M}^w_{j,k})^2 \mid \mathcal{H}^n_{T(n,i+j)}) \leq K(k-j)\Delta_n$ for any $j < k$, and also that $\mathbb{E}((\overline{M}^w_{j,j+1})^2 \mid \overline{\mathcal{H}}^n_{T(n,i+j)}) \leq K\Delta(n,i+j+1)$. The same arguments also yield

$$\begin{array}{rcl}\mathbb{E}(\overline{V(1)}^2_{j,j+1} \mid \overline{\mathcal{H}}^n_{T(n,i+j)}) & \leq & 4\theta'^2_{j+1}\mathbb{E}\big(\int_{T(n,i+j)}^{T(n,i+j+1)} (M^1_s)^2\alpha^2_s\,ds \mid \overline{\mathcal{H}}^n_{T(n,i+j)}\big) \\ & \leq & K\Delta(n,i+j+1)\big((\overline{M}^1_{0,j})^2 + \mathbb{E}((\overline{M}^1_{j,j+1})^2 \mid \overline{\mathcal{H}}^n_{T(n,i+j)})\big) \\ & \leq & K\Delta(n,i+j+1)(\overline{M}^1_{0,j})^2 + K\Delta(n,i+j+1)^2.\end{array}$$

Then, by conditioning on $\mathcal{H}^n_{T(n,i+j)}$, we see that $V(1) \in \mathcal{P}(\Psi_j)$ with Ψ_j given by (5.19) and, since $V(1)$ is an $(\mathcal{H}^n_t)$-martingale, it also belongs to $\widehat{\mathcal{P}}(\widehat{\Psi})$ with $\widehat{\Psi} = K(h_n\Delta_n)^2$ by the second part of (5.18). The same obviously holds for $V(2)$, and the proof is complete. □

LEMMA 15. *Let $V \in \mathcal{A}^n_i$ and U be a square-integrable martingale for the filtration $(\mathcal{H}^n_{T(n,i)} \bigvee \mathcal{F}_t)$ with predictable bracket $\langle U, U\rangle_t = \int_0^t \alpha_s\,ds$ with α_s bounded (note that $U = W$ satisfies this).*

a) If $V \in \mathcal{P}(\Psi_j) \cap \widehat{\mathcal{P}}(\widehat{\Psi})$ we have $L(V) \in \mathcal{P}(\Psi'_j) \cap \widehat{\mathcal{P}}(\widehat{\Psi}')$, where

$$\Psi'_j = K\big(\Psi_j + (\overline{V}_{0,j})^2\big)\Delta_n^2, \qquad \widehat{\Psi}' = K(h_n\Delta_n)^2\Big(\widehat{\Psi} + \frac{1}{h_n}\sum_{j=0}^{2h_n-1} \mathbb{E}(\Psi_j \mid \mathcal{H}^n_{T(n,i)})\Big). \tag{5.20}$$

b) If $V\in\mathcal{P}(\Psi_j)\cap\widehat{\mathcal{P}}(\widehat{\Psi})$ the process $L'(V,U)$ is a local martingale relative to $(\mathcal{H}^n_t)$ and $(\overline{\mathcal{H}}^n_t)$, and $L'(V,U)\in\mathcal{P}(\Psi'_j)\cap\widehat{\mathcal{P}}(\widehat{\Psi}')$, where

$$\Psi'_j=K\big(\Psi_j+(\overline{V}_{0,j})^2\big)\Delta_n,\qquad \widehat{\Psi}'=Kh_n\Delta_n\Big(\widehat{\Psi}+\frac{1}{h_n}\sum_{j=0}^{2h_n-1}\mathbb{E}(\Psi_j\mid\mathcal{H}^n_{T(n,i)})\Big). \tag{5.21}$$

If further V is bounded and $U=W$, we also have for all $p\geq 2$:

$$\mathbb{E}\big(|L'(V,W)_{T(i+2h_n)}|^p\mid\mathcal{H}^n_{T(n,i)}\big)\leq K_p(h_n\Delta_n)^{p/2}\Big(\widehat{\Psi}+\frac{1}{h_n}\sum_{j=0}^{2h_n-1}\mathbb{E}(\Psi_j\mid\mathcal{H}^n_{T(n,i)})\Big). \tag{5.22}$$

c) If $V\in\mathcal{P}(\Psi_j)\cap\overline{\mathcal{P}}(\overline{\Psi})\cap\widehat{\mathcal{P}}(\widehat{\Psi})$ we have

$$\begin{aligned}&\big|\mathbb{E}(L(V)_{T(n,i+2h_n)}\mid\mathcal{H}^n_{T(n,i)})\big|\\&\quad\leq K\Big(h_n\Delta_n\overline{\Psi}+(h_n\Delta_n)^{3/2}\sqrt{\widehat{\Psi}}+\Delta_n\sum_{j=0}^{2h_n-1}\mathbb{E}(\sqrt{\Psi_j}\mid\mathcal{H}^n_{T(n,i)})\Big),\\&\big|\mathbb{E}(L''(V)_{T(n,i+2h_n)}\mid\mathcal{H}^n_{T(n,i)})\big|\\&\quad\leq Kh_n\Delta_n\Big(h_n\Delta_n\overline{\Psi}+\big((h_n\Delta_n)^{3/2}+\Delta_n\big)\sqrt{\widehat{\Psi}}+\Delta_n\sum_{j=0}^{2h_n-1}\mathbb{E}(\sqrt{\Psi_j}\mid\mathcal{H}^n_{T(n,i)})\Big).\end{aligned} \tag{5.23}$$

Proof. a) The process $V'=L(V)$ is continuous and belongs to $\mathcal{A}^n_i$. We have, uniformly in $t\in[T(n,i+j),T(n,i+j+1)]$,

$$V'_t-V'_{T(n,i+j)}=\theta_{j+1}\int_{T(n,i+j)}^t V_s\,ds=\mathrm{O}\big(\overline{V}_{0,j+1}\,\Delta(n,i+j+1)\big),$$

hence $\overline{V}'_{j,j+1}\leq K(\overline{V}_{0,j}+\overline{V}_{j,j+1})\Delta(n,i+j+1)$. Thus, since $\overline{V}_{0,j}$ is $\mathcal{H}^n_{T(n,i+j)}$-measurable, we have $V'\in\mathcal{P}(\Psi'_j)$ with Ψ'_j given by (5.20). Then (5.18) applied to V' yields $V'\in\widehat{\mathcal{P}}(\widehat{\Psi}'')$ with $\widehat{\Psi}''=h_n\sum_{j=0}^{2h_n-1}\mathbb{E}(\Psi'_j\mid\mathcal{H}^n_{T(n,i)})$, which is smaller than Ψ' as given by (5.20), and the proof is complete.

b) Exactly as in (b) of the previous proof, U is a square-integrable martingale for the two filtrations $(\mathcal{H}^n_t)$ and $(\overline{\mathcal{H}}^n_t)$, so the process $V'=L'(V,U)$ is a local martingale for these two filtrations as well, and it vanishes for $t\leq T(n,i)$. The same argument as in (b) of the previous lemma again yields for all $p\geq 2$ if $U=W$ and for $p=2$ otherwise, and upon using the Burkholder-Gundy inequality:

$$\begin{aligned}\mathbb{E}\big((\overline{V}'_{j,j+1})^p\mid\overline{\mathcal{H}}^n_{T(n,i+j)}\big)&\leq K_p\,\theta^p_{j+1}\mathbb{E}\big(\big(\textstyle\int_{T(n,i+j)}^{T(n,i+j+1)}\alpha_s(V_s)^2\,ds\big)^{p/2}\mid\overline{\mathcal{H}}^n_{T(n,i+j)}\big)\\&\leq K_p\Delta(n,i+j+1)^{p/2}\big((\overline{V}_{0,j})^p+\mathbb{E}\big((\overline{V}_{j,j+1})^p\mid\overline{\mathcal{H}}^n_{T(n,i+j)}\big)\big).\end{aligned}$$

Using this with $p=2$ gives us $V'\in\mathcal{P}(\Psi'_j)$ for Ψ'_j as stated. The proof that $V'\in\widehat{\mathcal{P}}(\widehat{\Psi}')$ with $\widehat{\Psi}'$ as stated is the same as in Step (a), upon using now the second part of (5.18) for V'.

Assume further V bounded and $U=W$. Then obviously $(\overline{V}_{j,j+1})^p\leq K_p\,(\overline{V}_{j,j+1})^2$ and $(\overline{V}_{0,j})^p\leq K_p\,(\overline{V}_{0,j})^2$, hence $\mathbb{E}\big((\overline{V}'_{j,j+1})^p\mid\mathcal{H}^n_{T(n,i)}\big)\leq K_p\,\Delta_n^{p/2}\big(\mathbb{E}(\Psi_j\mid\mathcal{H}^n_{T(n,i)})+\widehat{\Psi}\big)$. Applying the first part of (5.17) to the $(\mathcal{H}^n_t)$-martingale V', we readily get (5.22).

c) (2.10) yields the decomposition

$$L(V)_{T(n,2h_n)}=\textstyle\sum_{j=0}^{2h_n-1}\theta^n_{j+1}(\zeta^n_j+\zeta'^n_j+\zeta''^n_j),\qquad\text{where}\quad \zeta^n_j=\Delta_n\lambda_{T(n,i)}V_{T(n,i+j)}\Phi^n_{i+j+1},$$
$$\zeta'^n_j=\Delta_nV_{T(n,i+j)}(\lambda_{T(n,i+j)}-\lambda_{T(n,i)})\Phi^n_{i+j+1},\qquad \zeta''^n_j=\textstyle\int_{T(n,i+j)}^{T(n,i+j+1)}(V_s-V_{T(n,i+j)})\,ds.$$

(O)-(ii) yields $\mathbb{E}(\zeta^n_j \mid \mathcal{H}^n_{T(n,i+j)}) = \Delta_n V_{T(n,i+j)}\lambda_{T(n,i)}$ and a similar property for ζ'^n_j. Then, $|\mathbb{E}(\zeta^n_j \mid \mathcal{H}^n_{T(n,i)})| \leq K\Delta_n\overline{\Psi}$, and (5.5) applied to λ_t and the Cauchy-Schwarz inequality yield $|\mathbb{E}(\zeta'^n_j \mid \mathcal{H}^n_{T(n,i)})| \leq Kh_n^{1/2}\Delta_n^{3/2}\widehat{\Psi}^{1/2}$, whereas $|\mathbb{E}(\zeta''^n_j \mid \mathcal{H}^n_{T(n,i)})| \leq K\Delta_n\mathbb{E}((\Psi_j)^{1/2} \mid \mathcal{H}^n_{T(n,i)})$ is obvious. The first part of (5.23) follows.

In the same way, we have $L''(V)_{T(n,2h_n)} = \sum_{j=0}^{2h_n-1}\theta^n_{j+1}\sum_{w=1}^5\zeta^{n,w}_j$, where

$$\begin{aligned}
\zeta^{n,1}_j &= \Delta_n^2\textstyle\sum_{k=j+1}^{2h_n-1}\theta^n_{k+1}\lambda^2_{T(n,i)}V_{T(n,i+j)}\Phi^n_{i+j+1}\Phi^n_{i+k+1},\\
\zeta^{n,2}_j &= \Delta_n^2\textstyle\sum_{k=j+1}^{2h_n-1}\theta^n_{k+1}\big(\lambda^2_{T(n,i+j)}-\lambda^2_{T(n,i)}\big)V_{T(n,i+j)}\Phi^n_{i+j+1}\Phi^n_{i+k+1},\\
\zeta^{n,3}_j &= \Delta_n\textstyle\sum_{k=j+1}^{2h_n-1}\theta^n_{k+1}\big(\lambda_{T(n,i+k)}-\lambda_{T(n,i+j)}\big)V_{T(n,i+j)}\Delta(n,i+j+1)\Phi^n_{i+k+1},\\
\zeta^{n,4}_j &= \theta^n_{j+1}V_{T(n,i+j)}\textstyle\int_{T(n,i+j)}^{T(n,i+j+1)}(T(n,i+j+1)-s)\,ds,\\
\zeta^{n,5}_j &= \theta^n_{j+1}\textstyle\int_{T(n,i+j)}^{T(n,i+j+1)}(V_s-V_{T(n,i+j)})(T(n,i+j+1)-s)\,ds\\
&\quad+\textstyle\sum_{k=j+1}^{2h_n-1}\theta^n_{k+1}\Delta(n,i+k+1)\int_{T(n,i+j)}^{T(n,i+j+1)}(V_s-V_{T(n,i+j)})\,ds.
\end{aligned}$$

By successive conditioning and the same arguments as above, we see that $|\mathbb{E}(\zeta^{n,1}_j \mid \mathcal{H}^n_{T(n,i)})| \leq Kh_n\Delta_n^2\overline{\Psi}$, and also that $\big|\mathbb{E}(\zeta^{n,w}_j \mid \mathcal{H}^n_{T(n,i)})\big|$ is smaller than $Kh_n^{3/2}\Delta_n^{5/2}\widehat{\Psi}^{1/2}$ if $w=2,3$, than $K\Delta_n^2\widehat{\Psi}^{1/2}$ if $w=4$, and than $Kh_n\Delta_n^2\mathbb{E}((\Psi_j)^{1/2} \mid \mathcal{H}^n_{T(n,i)})$ if $w=5$. Since $h_n\Delta_n\to 0$, All these estimates give us the second part of (5.23). $\square$

After these general technical results we introduce some processes more specifically related to our problem. For any $y>0$ we set

$$\begin{aligned}
&U(y)^n_t = e^{-\Upsilon(y)^n_t}, \qquad \Upsilon(y)^n_t = c(y)^n_t\lambda_t + a(y)^n_t\lambda_t + \gamma(y)^n_t \quad \text{with}\\
&c(y)^n_t = y^2u_n^2h_n\Delta_n\phi_n\,c_t, \qquad \gamma(y)^n_t = y^2u_n^2h_n^{-1}\,\overline{\phi}_n\,\gamma_t,\\
&a(y)^n_t = 4\textstyle\sum_{m=1}^M|y|^{\beta_m}u_n^{\beta_m}h_n\Delta_n\,\widetilde{\phi}_n^{\beta_m}\chi(\beta_m)\,a^m_t.
\end{aligned} \tag{5.24}$$

and also

$$\begin{aligned}
V^\#(y,y')^n_t &= U(y+y')^n_t + U(|y-y'|)^n_t - 2U(y)^n_t\,U(y')^n_t,\\
V(y,y')^n_t &= \frac{V^\#(y,y')^n_t}{2\,U(y)^n_t\,U(y')^n_t}, \qquad \widetilde{V}(y)^n_t = V(y,1)^n_t - y^2V(1,1)^n_t,\\
\overline{V}(y,y')^n_t &= V(y,y')^n_t + y^2y'^2V(1,1)^n_t - y'^2V(y,1)^n_t - y^2V(y',1)^n_t.
\end{aligned}$$

Upon increasing Γ if necessary, we have

$$\frac{1}{\Gamma} \leq U(y)^n_t \leq 1. \tag{5.25}$$

In view of (5.5), (5.12) and (5.14), for $U^n = U(y)^n$ hence for $U^n = V^\#(y,y')^n$, $U^n = V(y,y')^n$ and $U^n = \overline{V}(y,y)^n$ as well (upon using (5.25)), we have for $j=1,\dots,w_n=2h_nk_n$ and $p\geq 2$:

$$\begin{gathered}
\big|\mathbb{E}(U^n_{T(n,i+j)} - U^n_{T(n,i)} \mid \mathcal{G}^n_i)\big| \leq K\chi_{n,j}, \quad \mathbb{E}(|U(y)^n_{T(n,i+j)} - U(y)^n_{T(n,i)}|^p \mid \mathcal{G}^n_i) \leq K\chi(p)_{n,j},\\
\chi_{n,j} = u_n^2h_n^{-1}j\Delta_n + \kappa_2u_n^2h_n^{-1}(j\Delta_n)^{1/2} + h_n\Delta_n\Big(\kappa_1u_n^2j\Delta_n + \sum_{m=1}^M u_n^{\beta_m}(j\Delta_n)^{\beta_m/2}\Big),\\
\chi(p)_{n,j} = u_n^{2p}h_n^{-p}j\Delta_n + (h_n\Delta_n)^p\Big(\kappa_1u_n^{2p}j\Delta_n + \sum_{m=1}^M u_n^{p\beta_m}(j\Delta_n)^{1\wedge(p\beta_m/2)}\Big).
\end{gathered} \tag{5.26}$$

Moreover, an expansion of the exponential function gives us

$$
\begin{array}{lcl}
V^{\#}(y,y')^n_t &=& \mathrm{O}\big(u_n^{\beta_1}h_n\Delta_n + \frac{u_n^4}{h_n^2} + \kappa_1 u_n^4(h_n\Delta_n)^2\big),\\
V(y,y')^n_t &=& 2u_n^{\beta_1}h_n\Delta_n\psi_{\beta_1}(y,y')\widetilde{\phi}_n^{(\beta_1)}\chi(\beta_1)\,a^1_t\lambda_t + 2y^2y'^2u_n^4\big(h_n\Delta_n\phi_n\,c_t\lambda_t + \frac{1}{h_n}\,\overline{\phi}_n\,\gamma_t\big)^2\\
&& +\frac{2}{3}\,y^4y'^4u_n^8\big(h_n\Delta_n\phi_n\,c_t\lambda_t + \frac{1}{h_n}\,\overline{\phi}_n\,\gamma_t\big)^4\\
&& +\mathrm{o}\big(u_n^{\beta_1}h_n\Delta_n + \frac{u_n^8}{h_n^4} + \kappa_1 u_n^8(h_n\Delta_n)^4\big),\\
\widetilde{V}(y)^n_t &=& \mathrm{O}\big(u_n^{\beta_1}h_n\Delta_n + \frac{u_n^8}{h_n^4} + \kappa_1 u_n^8(h_n\Delta_n)^4\big),\\
\overline{V}(y,y')^n_t &=& 2h_n\Delta_n u_n^{\beta_1}\overline{\psi}_{\beta_1}(y,y')\widetilde{\phi}_n^{(\beta_1)}\chi(\beta_1)\,a^1_t\lambda_t\\
&& +\frac{2}{3}\,y^2y'^2(y^2-1)(y'^2-1)u_n^8\big(h_n\Delta_n\phi_n\,c_t\lambda_t + \frac{1}{h_n}\,\overline{\phi}_n\,\gamma_t\big)^4\\
&& +\mathrm{o}\big(u_n^{\beta_1}h_n\Delta_n + \frac{u_n^8}{h_n^4} + \kappa_1 u_n^8(h_n\Delta_n)^4\big).
\end{array} \tag{5.27}
$$

5.4. *Estimates - 2.* In this subsection we prove various estimates for a number of arrays of variables, which we presently define. Since we take differences of two successive pre-averaged values, it is convenient to introduce the following:

$$
\begin{array}{lll}
g'^n_j = -g^n_j, & \overline{g}'^n_j = -\overline{g}^n_j, & \text{if } 1\le j\le h_n-1,\\
g'^n_j = g^n_{j-h_n}, & \overline{g}'^n_j = \overline{g}^n_{j-h_n}, & \text{if } h_n\le j\le 2h_n-1,
\end{array} \tag{5.28}
$$

so that $\widetilde{V}^n_{i+h_n} - \widetilde{V}^n_i = \sum_{j=1}^{2h_n-1} g'^n_j\Delta^n_{i+j}V$. Recalling $\gamma^n_i = \gamma_{T(n,i)}$ and $\sigma^n_i = \sigma_{T(n,i)}$ and writing $\delta^n_i(z) = \delta(T(n,i),z)$, we set

$$
\begin{array}{ll}
\psi^{n,i}_t = \sum\limits_{j=1}^{2h_n-1} g'^n_j\,1_{(T(n,i+j-1)),T(n,i+j)]}(t), & \rho^{n,1}_i = u_n\sigma^n_i\int_0^\infty \psi^{n,i}_s\,dW_s,\\
\rho^{n,2}_i = u_n\int_0^\infty\int_E \delta^n_i(z)\psi^{n,i}_s\,(\underline{p}-\underline{q})(dt,dz), & \rho^{n,3}_i = u_n\gamma'^n_i\sum\limits_{j=1}^{2h_n-1}\overline{g}'^n_j\varepsilon^n_{i+j},\\
\rho^n_i = \rho^{n,1}_i + \rho^{n,2}_i + \rho^{n,3}_i, & \overline{\rho}^n_i = u_n(\widetilde{Y}^n_i - \widetilde{Y}^n_{i+h_n}),
\end{array}
$$

and

$$
\xi(y)^{w,n}_j = \begin{cases}
\frac{1}{k_n}\sum_{l=0}^{k_n-1}\big(\cos(y\rho^n_{2h_n(jk_n+l)}) - U(y)^n_{T(n,2h_n(jk_n+l))}\big) & \text{if } w=1,\\
\frac{1}{k_n}\sum_{l=0}^{k_n-1}\big(\cos(y\overline{\rho}^n_{2h_n(jk_n+l)}) - \cos(y\rho^n_{2h_n(jk_n+l)})\big) & \text{if } w=2,\\
\frac{1}{k_n}\sum_{l=0}^{k_n-1}(U(y)^n_{T(n,2h_n(jk_n+l))} - U(y)^n_{T(n,jw_n)}) & \text{if } w=3,
\end{cases}
$$

$$
\xi(y)^n_j = \frac{1}{U(y)^n_{T(n,jw_n)}}\sum_{w=1}^3 \xi(y)^{w,n}_j,
$$

$$
\Omega(y)_{n,t} = \bigcap\nolimits_{0\le j<[N^n_t/w_n]}\{|\xi(y)^n_j|\le\tfrac{1}{2}\},\qquad \Omega_{n,t} = \bigcap\nolimits_{y\in\mathcal{Y}}\big(\Omega(y)_{n,t}\cap\Omega(2y)_{n,t}\big).
$$

We also introduce a (long) list of numerical sequences, which all go to 0 by (3.2), and where P is an arbitrarily large integer:

$$\begin{aligned}
\alpha_n^1 &= h_n\Delta_n + u_n(h_n\Delta_n)^{1/2+1/\widetilde{r}'} + u_n^{2+\widetilde{r}'/2}\Delta_n + u_n^{P+1}h_n^{-P/2}\Delta_n^{1/2} + \kappa_2 u_n^2 h_n^{-1/2}\Delta_n^{1/2}, \\
&\quad + \kappa_1\big(u_n^{1+\widetilde{r}'/2}(h_n\Delta_n)^{3/2} + u_n^5 h_n^{-1}\Delta_n + u_n^6 h_n^{1/2}\Delta_n^2 + u_n^6(h_n\Delta_n)^{7/2} + u_n^4 h_n^2\Delta_n^{5/2}\big), \\
\alpha_n^2 &= h_n\Delta_n + u_n^2(h_n\Delta_n)^3 + u_n^{\widetilde{r}'}(h_n\Delta_n)^{1+\widetilde{r}'/2}, \\
\alpha_n^3 &= u_n^r h_n\Delta_n + \kappa_2 u_n^3 h_n^{-3/2}\Delta_n^{1/2} + u_n^4 h_n^{-3}, \\
\alpha_n^4 &= u_n^{\beta_1} h_n\Delta_n + \kappa_2 u_n^3 h_n^{-3/2}\Delta_n^{1/2} + u_n^4 h_n^{-2} + \kappa_1 u_n^4(h_n\Delta_n)^2, \\
\alpha_n^5 &= u_n^2 k_n\Delta_n + \kappa_2 u_n^2 k_n^{1/2} h_n^{-1/2}\Delta_n^{1/2} + \textstyle\sum_{m=1}^M u_n^{\beta_m} k_n^{\beta_m/2}(h_n\Delta_n)^{1+\beta_m/2} + \kappa_1 u_n^2 k_n(h_n\Delta_n)^2, \\
\overline{\alpha}(p)_n &= (\alpha_n^3)^p + \alpha_n^4 k_n^{-p/2}, \\
\widehat{\alpha}(p)_n &= (\alpha_n^1)^p + \alpha_n^2 k_n^{-p/2} + u_n^{2p} k_n h_n^{1-p}\Delta_n + \kappa_1 u_n^{2p} k_n(h_n\Delta_n)^{p+1} \\
&\qquad + \textstyle\sum_{m=1}^M u_n^{p\beta_m} k_n^{1\wedge(p\beta_m/2)}(h_n\Delta_n)^{p+1\wedge(p\beta_m/2)}.
\end{aligned}$$

LEMMA 16. *For all $p\geq 2$ and $y\in\mathcal{Y}$, we have*

$$\big|\mathbb{E}(\cos(y\overline{\rho}_i^n) - \cos(y\rho_i^n) \mid \mathcal{G}_i^n\big| \leq K\alpha_n^1, \qquad \mathbb{E}(|\cos(y\overline{\rho}_i^n) - \cos(y\rho_i^n)|^p \mid \mathcal{G}_i^n\big| \leq K\alpha_n^2. \tag{5.29}$$

Proof. 1) Since $\int_0^\infty \psi_t^{n,i}\,dt = 0$, we have $y\overline{\rho}_i^n = \overline{\theta}(6)_i^n$, where $\overline{\theta}(k)_i^n = \sum_{j=1}^k \theta(j)_i^n$ and

$$\begin{aligned}
&\theta(1)_i^n = y\rho_i^n, \qquad\qquad \theta(2)_i^n = yu_n\textstyle\int_0^\infty(\sigma_s - \sigma_i^n)\psi_s^{n,i}\,dW_s, \\
&\theta(3)_i^n = yu_n\textstyle\sum_{j=0}^{2h_n-1}\overline{g}_j'^n(\gamma_{i+j}'^n - \gamma_i'^n)\varepsilon_{i+j}^n, \\
&\theta(4)_i^n = yu_n\textstyle\int_0^\infty\int_E(\delta(s,z) - \delta_i^n(z))\psi_s^{n,i}(\underline{p} - \underline{q})(ds,dz), \\
&\theta(5)_i^n = yu_n\textstyle\int_0^\infty(b_s - b_i^n)\psi_s^{n,i}\,ds, \\
&\theta(6)_i^n = yu_n\textstyle\int_0^\infty\int_E\delta'(s,z)\psi_s^{n,i}\,\underline{p}(ds,dz).
\end{aligned}$$

2) This step is devoted to proving the following estimates, for any $a\geq\widetilde{r}'$ and $p\geq 2$:

$$\begin{aligned}
&\mathbb{E}(|\rho_i^{n,1}|^p \mid \mathcal{G}_i^n) \leq K_p\kappa_1 u_n^p(h_n\Delta_n)^{p/2}, &&\mathbb{E}(|\rho_i^{n,2}|^a \mid \mathcal{G}_i^n) \leq K_a u_n^a h_n\Delta_n, \\
&\mathbb{E}(|\theta(2)_i^n|^p \mid \mathcal{G}_i^n) \leq K_p\kappa_1 u_n^p(h_n\Delta_n)^{1+p/2}, &&\mathbb{E}(|\theta(4)_i^n|^{\widetilde{r}'} \mid \mathcal{G}_i^n) \leq K u_n^{\widetilde{r}'}(h_n\Delta_n)^{1+\widetilde{r}'/2}, \\
&\mathbb{E}(|\theta(5)_i^n|^2 \mid \mathcal{G}_i^n) \leq K u_n^2(h_n\Delta_n)^3, &&\mathbb{P}(\theta(6)_i^n \neq 0 \mid \mathcal{G}_i^n) \leq K h_n\Delta_n.
\end{aligned} \tag{5.30}$$

By virtue of Lemma 12-(d) we can always condition on $\mathcal{H}_{T(n,i)}^n$ instead of $\mathcal{G}_i^n$. The claim for $\rho_i^{n,1}$ follows from Burkholder-Gundy inequality and $|\sigma_i^n\psi_t^{n,i}|\leq K$, plus (5.11). A trivial reformulation of Lemma 2.1.5 of [11] entails that, for any predictable function on $\Omega\times\mathbb{R}_+\times E$ with $|\delta''(t,z)|^{r'}\leq KJ(z)$ and any two $(\mathcal{H}_t^n)$-stopping times $T\leq S$ we have for $a\geq\widetilde{r}'$ and $Z = \delta''*(\underline{p} - \underline{q})$:

$$\mathbb{E}\big(|Z_S - Z_T|^a \mid \mathcal{H}_T^n\big) \leq K\mathbb{E}\Big(\int_{(S,T]\times E}|\delta''(t,z)|^a\,dt\,\eta(dz) + \Big(\int_{(S,T]\times E}|\delta''(t,z)|^{a\wedge 2}\,dt\,\eta(dz)\Big)^{(a\vee 2)/2} \mid \mathcal{H}_T^n\Big),$$

since J is bounded and η-integrable, $J^{a/r'}$ is also η-integrable. This with $\delta''(t,z) = \delta_i^n(z)\psi_t^{n,i}$ and (5.11) yield (5.30) for $\rho_i^{n,2}$. If $\delta''(t,z) = (\delta(t,z) - \delta_i^n(z))\psi_t^{n,i}$ and $V_t^z = \delta(t,z)/J(z)^{1/r'}$ it also implies with $a = \widetilde{r}'$ and $T = T(n,i)$ and $S = T(n,i+2h_n)$ and η' is the finite measure $\eta'(dz)J(z)^{2/r'}\eta(dz)$:

$$\mathbb{E}\big(|Z_S - Z_T|^p \mid \mathcal{H}_T^n\big) \leq K\int_E\mathbb{E}\Big(\int_S^T|V_t^z - V_T^z|^{\widetilde{r}'}\,dt\,\eta'(dz)\mid \mathcal{H}_T^n\Big)\,\eta'(dz).$$

Then (5.30) for $\theta(4)_i^n$ readily follows from (5.11) and (5.12) (and the sentence which follows it) applied with each V^z, plus Hölder inequality. Next, (5.5) and $\int J(z)\,\eta(dz)<\infty$ yield

$$\mathbb{P}(\theta(6)_i^n\neq 0\mid \mathcal{G}_i^n)\leq \frac{1}{\Gamma}\,\mathbb{E}\Big(\int_{R_+\times E} J(z)\,|\psi_s^{n,i}|\underline{p}(ds,dz)\mid \mathcal{G}_i^n\Big)\leq K\mathbb{E}(T(n,i+2h_n)-T(n,i)\mid \mathcal{G}_i^n),$$

and hence (5.30) for $\theta(6)_i^n$ by (5.11).

Next, we apply (a) of Lemma 14 with $V'=b$, hence $V_t=b_t-b_{t\wedge T(n,i)}$, and (5.20) to obtain the claim for $\theta(5)_i^n$ which is equal to $yu_nL(V)_{2h_n}$ (with the notation (5.15) and $\theta_j^n=g_j'^n$).

Analogously, Lemma 14-(a) with $V'=\sigma$, hence $V_t=\sigma_t-\sigma_{t\wedge T(n,i)}$, and (5.22) yields the claim for $\theta(2)_i^n$, which is equal to $yu_nL'(V,W)_{2h_n}$.

3) We turn to estimates for $\rho_i^{n,3}$ and $\theta(3)_i^n$. First, we have for all $p\geq 2$:

$$\mathbb{E}(|\rho_i^{n,3}|^p\mid \widehat{\mathcal{G}}_i^n)\leq K_p\,u_n^p h_n^{-p/2},\qquad \mathbb{E}(|\theta(3)_i^n|^p\mid \mathcal{G}_i^n)\leq K_p\,u_n^p h_n^{1-p/2}\Delta_n. \tag{5.31}$$

The first part above follows from Burkholder-Gundy inequality and $|\overline{g}_j'^n|\leq K/h_n$, because the $(\varepsilon_j^n: j\geq i)$ are independent and centered with bounded moments, conditionally on $\mathcal{H}_\infty^n$. Analogously,

$$\mathbb{E}(|\theta(3)_i^n|^p\mid \widehat{\mathcal{G}}_i^n)\leq K\,\frac{u_n^p}{h_n^{1+p/2}}\sum_{j=1}^{2h_n-1}|\gamma_{i+j}'^n-\gamma_i'^n|^{p/2}.$$

Then the second part of (5.31) follows from the last part of (5.12) applied with $V=\gamma'$.

However, (5.31) is not quite enough for us, and we need some further estimates, here and later on. For any integer $w\geq 1$ we denote by J_w the family of all w-uplet $\mathbf{j}=(j_1,\dots,j_w)$ of integer between 1 and $2h_n-1$. Within J_w we single out the subset J_w' of those $\mathbf{j}$'s for which at least one j_m is different from all others, and $J_w''=J_w\backslash J_w'$. When $\mathbf{j}\in J_w$ the integers j_m for $m=1,\dots,w$ take $\ell=\ell(\mathbf{j})$ distinct values $\overline{j}_1,\dots,\overline{j}_\ell$ and for each m there are $s_m\geq 1$ integers j_k equal to $\overline{j}_m$, and further $s_m\geq 2$ and $l\leq w/2$ when $\mathbf{j}\in J_w''$, whereas $s_m=1$ for at least one m when $\mathbf{j}\in J_w'$.

With this notation, we set

$$\begin{aligned} D_i^{n,w,y}&=\mathbb{E}\big((y\rho_i^{n,3})^w\mid \mathcal{H}_\infty^n\big),\qquad D_i'^{n,w,y}=\mathbb{E}\big(\theta(3)_i^n(y\rho_i^{n,3})^w\mid \mathcal{H}_\infty^n\big),\\ \overline{D}_i^{n,w,y}&=y^w u_n^w(\gamma_i'^n)^w\sum_{\mathbf{j}\in J_w''}\prod_{m=1}^{l}(\overline{g}_{\overline{j}_m}'^n)^{s_m}(\gamma^{(s_m)})_i^n. \end{aligned} \tag{5.32}$$

Recalling the properties of the noise, and in particular $\gamma_t^{(0)}=0$, we see that

$$\begin{aligned} D_i^{n,w,y}&=y^w u_n^w(\gamma_i'^n)^w\sum_{\mathbf{j}\in J_w''}\prod_{m=1}^{l}(\overline{g}_{\overline{j}_m}'^n)^{s_m}(\gamma^{(s_m)})_{i+\overline{j}_m}^n,\\ D_i'^{n,w,y}&=y^{w+1}u_n^{w+1}(\gamma_i'^n)^w\sum_{\mathbf{j}\in J_{w+1}''}\Big((\gamma_{i+j_{w+1}}'^n-\gamma_i'^n)\prod_{m=1}^{l}(\overline{g}_{\overline{j}_m}'^n)^{s_m}(\gamma^{(s_m)})_{i+\overline{j}_m}^n\Big). \end{aligned} \tag{5.33}$$

Recall $|\overline{g}_j'^n|\leq K/h_n$, whereas $\#J_w''\leq K_w h_n^{w-[(w+1)/2]}$. We have $\gamma_t^{(2)}=1$ and, for $q\geq 3$, the process $\gamma_t^{(q)}$ equals a constant when $\kappa_2=0$ and satisfies the last part of (5.12) and is bounded when $\kappa_2=1$, and γ_t' satisfies the same in all cases, plus the first part of (5.12) when $\kappa_2=0$. Then a simple calculation shows us that, for $p=2,4$,

$$
\begin{array}{l}
|D_i^{n,w,y}|+|\overline{D}_i^{n,w,y}|\leq K_w\,1_{\{w\geq 2\}}\,\frac{u_n^w}{h_n^{[(w+1)/2]}},\\
\mathbb{E}\big(|D_i^{n,w,y}-\overline{D}_i^{n,w,y}|^p\mid \mathcal{G}_i^n\big)\leq K_w\kappa_2\,1_{\{w\geq 3\}}\,\frac{u_n^{pw}h_n\Delta_n}{h_n^{p[(w+1)/2]}},\\
\big|\,\mathbb{E}(D_i'^{n,w,y}\mid \mathcal{G}_i^n)\big|\leq K_w\,\frac{u_n^{w+1}h_n\Delta_n}{h_n^{[w/2]+1}}+\kappa_2K_w\,\frac{u_n^{w+1}(h_n\Delta_n)^{1/2}}{h_n^{[w/2]+1}},\\
\mathbb{E}\big(|D_i'^{n,w,y}|^2\mid \mathcal{G}_i^n\big)\leq K_w\,\frac{u_n^{2w+2}h_n\Delta_n}{h_n^{2[w/2]+2}}.
\end{array}
\tag{5.34}
$$

4) Since $|\cos(u+v)-\cos(u)|\leq K(1\wedge|v|)$ and $|\cos(u+v)-\cos(u)-v\sin(u)|\leq Kv^2$, we deduce from (3.2) and (5.30) and $1/\widetilde{r}'<1$ that

$$
\begin{array}{l}
\mathbb{E}\big(|\cos(y\overline{\rho}_i^n)-\cos(y\rho_i^n)|^2\mid \mathcal{G}_i^n\big)\leq K\alpha_n^2,\\
\mathbb{E}\big(\cos(y\overline{\rho}_i^n)-\cos(\overline{\theta}(3)_i^n)\mid \mathcal{G}_i^n\big)\leq K\big(h_n\Delta_n+u_n(h_n\Delta_n)^{1/2+1/\widetilde{r}'}\big),\\
\mathbb{E}\big(\cos(\overline{\theta}(3)_i^n)-\cos(\overline{\theta}(2)_i^n)-\theta(3)_i^n\,\sin(\overline{\theta}(2)_i^n)\mid \mathcal{G}_i^n\big)\leq Ku_n^2\Delta_n\leq Kh_n\Delta_n,\\
\mathbb{E}\big(\cos(\overline{\theta}(2)_i^n)-\cos(y\rho_i^n)-\theta(2)_i^n\,\sin(y\rho_i^n)\mid \mathcal{G}_i^n\big)\leq K\kappa_1u_n^2(h_n\Delta_n)^2\leq Kh_n\Delta_n.
\end{array}
\tag{5.35}
$$

The first estimate above yields the second part of (5.29) for $p=2$, hence for all $p\geq 2$ as well.

Next, we evaluate $\mathbb{E}(\theta(3)_i^n\,\sin(\overline{\theta}(2)_i^n)\mid \mathcal{G}_i^n)$. Set $\widehat{\theta}_i^n=\overline{\theta}(2)_i^n-y\rho_i^{n,3}$. A Taylor expansion of the function $f(x)=\sin x$ around $\widehat{\theta}_i^n$ and the fact that the derivatives $f^{(w)}$ of f are all bounded by 1 yield, for any even integer $P\geq 2$,

$$
\theta(3)_i^n\,\sin(\overline{\theta}(2)_i^n)=\sum_{w=0}^{P-1}\frac{1}{w!}\,f^{(w)}(\widehat{\theta}_i^n)(y\rho_i^{n,3})^w\theta(3)_i^n+\mathrm{O}\big(|y\rho_i^{n,3}|^P|\theta(3)_i^n|\big).
$$

Since $\mathbb{E}(\theta(3)_i^n\mid \mathcal{H}_\infty^n)=0$ we have $\mathbb{E}(\theta(3)_i^n\,\sin(\overline{\theta}(2)_i^n)\mid \mathcal{G}_i^n)=\sum_{w=1}^P\frac{1}{w!}\,\eta_i^{n,w}$, where (with $a_w=(-1)^{[w/2]}$)

$$
\eta_i^{n,w}=\begin{cases}a_w\,\mathbb{E}(D_i'^{n,w,y}\cos(y\widehat{\theta}_i^n)\mid \mathcal{G}_i^n)=\mathrm{O}\big(\mathbb{E}(D_i'^{n,w,y}+|D_i'^{n,w,y}|((\widehat{\theta}_i^n)^2\wedge 1)\mid \mathcal{G}_i^n)\big) & \text{if } w<P \text{ is odd},\\ a_w\,\mathbb{E}(D_i'^{n,w,y}\sin(y\widehat{\theta}_i^n)\mid \mathcal{G}_i^n)=\mathrm{O}\big(\mathbb{E}(|D_i'^{n,w,y}\widehat{\theta}_i^n|\mid \mathcal{G}_i^n)\big) & \text{if } w<P \text{ is even},\\ \mathrm{O}\big(\mathbb{E}(|D_i'^{n,P,y}|\mid \mathcal{G}_i^n)\big) & \text{if } w=P\end{cases}
$$

We have $\widehat{\theta}_i^n=\theta(2)_i^n+y\rho_i^{n,1}+y\rho_i^{n,2}$, hence if we combine (5.30) and (5.34) plus the Cauchy-Schwarz inequality and $u_n^2\leq Kh_n$ and (5.31) for the last estimate below, we see that

$$
\eta_i^{n,w}\leq\begin{cases}K_w\big(h_n\Delta_n+u_n^{2+\widetilde{r}'/2}\Delta_n+\kappa_1u_n^4h_n^{1/2}\Delta_n^{3/2}+\kappa_2u_n^2h_n^{-1/2}\Delta_n^{1/2}\big) & \text{if } w<P \text{ is odd},\\ K_wh_n\Delta_n & \text{if } w<P \text{ is even},\\ Ku_n^{P+1}h_n^{-P/2}\Delta_n^{1/2} & \text{if } w=P.\end{cases}
$$

Then we end up with

$$
\big|\,\mathbb{E}(\theta(3)_i^n\,\sin(\overline{\theta}(2)_i^n)\mid \mathcal{G}_i^n)\big|\leq K\Big(h_n\Delta_n+u_n^{2+\widetilde{r}'/2}\Delta_n+\kappa_1u_n^4h_n^{1/2}\Delta_n^{3/2}+\frac{u_n^{P+1}\Delta_n^{1/2}}{h_n^{P/2}}+\kappa_2\frac{u_n^2\Delta_n^{1/2}}{h_n^{1/2}}\Big).
\tag{5.36}
$$

5) Now we estimate $\mathbb{E}(\theta(2)_i^n\,\sin(y\rho_i^n)\mid \mathcal{G}_i^n)$, assuming $\kappa_1=1$, otherwise this vanishes identically. First, $|\sin(y\rho_i^n)-\sin(y\rho_i^{n,1}+y\rho_i^{n,3})|^2\leq K|\rho_i^{n,2}|^{\widetilde{r}'}$ and (5.30) and (5.31) yield

$$
\big|\,\mathbb{E}\big(\theta(2)_i^n\,(\sin(y\rho_i^n)-\sin(y\rho_i^{n,1}+y\rho_i^{n,3}))\mid \mathcal{G}_i^n\big)\big|\leq Ku_n^{1+\widetilde{r}'/2}(h_n\Delta_n)^{3/2}.
\tag{5.37}
$$

Next, expand $f(x)=\sin(x)$ around $y\rho_i^{n,1}$ and use (5.30) and (5.31) to get

$$\mathbb{E}\big(\theta(2)_i^n\,\sin(y\rho_i^{n,1}+y\rho_i^{n,3})\mid\mathcal{G}_i^n\big)=\sum_{w=0}^{3}\tfrac{1}{w!}\,v_i^{n,w}+\mathrm{O}(u_n^5\Delta_n/h_n),\quad\text{where}$$
$$v_i^{n,w}=\begin{cases}\mathbb{E}\big(\theta(2)_i^n\,\sin(y\rho_i^{n,1})\mid\mathcal{G}_i^n\big) & \text{if } w=0,\\ \mathbb{E}\big(\theta(2)_i^n\,f^{(w)}(y\rho_i^{n,1})\,(y\rho_i^{n,3})^w\mid\mathcal{G}_i^n\big)=\mathbb{E}\big(\theta(2)_i^n\,f^{(w)}(y\rho_i^{n,1})\,D_i^{n,w,y}\mid\mathcal{G}_i^n\big) & \text{if } w\geq 1.\end{cases}$$

Then (5.30) and (5.34) yield

$$v_i^{n,1}=0,\qquad w=2,3\;\Rightarrow\;|v_i^{n,w}|\leq Ku_n^w h_n^{-[(w+1)/2]}\,\big|\,\mathbb{E}(\theta(2)_i^n\,f^{(w)}(y\rho_i^{n,1})\mid\mathcal{G}_i^n))\big|.$$

Thus $|v_i^{n,3}|\leq Ku_n^4\Delta_n/h_n\leq Kh_n\Delta_n$ by (5.30), whereas $f^{(2)}=-f$ yields $|v_i^{n,2}|\leq K|v_i^{n,0}|$, hence

$$\big|\,\mathbb{E}\big(\theta(2)_i^n\,\sin(y\rho_i^{n,1}+y\rho_i^{n,3})\mid\mathcal{G}_i^n\big)\big|\leq K|v_i^{n,0}|+K\big(h_n\Delta_n+\frac{u_n^5\Delta_n}{h_n}\big).$$

Another expansion of the function f, around 0 this time, yields $v_i^{n,0}=\overline{v}_i^{n,1}+\overline{v}_i^{n,3}+\overline{v}_i^n$, where

$$\overline{v}_i^{n,w}=\frac{y^w}{w!}\,\mathbb{E}\big(\theta(2)_i^n\,(\rho_i^{n,1})^w\mid\mathcal{G}_i^n\big),\qquad|\overline{v}_i^n|\leq\mathbb{E}\big(|\theta(2)_i^n|\,|\rho_i^{n,1}|^5\mid\mathcal{G}_i^n\big)\leq Ku_n^6(h_n\Delta_n)^{7/2},$$

(use (5.30) again), so we deduce

(5.38)
$$\big|\,\mathbb{E}\big(\theta(2)_i^n\,\sin(y\rho_i^{n,1}+y\rho_i^{n,3})\mid\mathcal{G}_i^n\big)\big|\leq K\Big(|\overline{v}_i^{n,1}|+|\overline{v}_i^{n,3}|+h_n\Delta_n+u_n^6(h_n\Delta_n)^{7/2}+\frac{u_n^5\Delta_n}{h_n}+u_n^6h_n^{1/2}\Delta_n^2\Big).$$

6) It remains to evaluate $\overline{v}_i^{n,w}$ for $w=1,3$. Omitting the indices n,i we write

$$S=T(n,i),\quad T=T(n,i+2h_n),\qquad M_t=\textstyle\int_{t\wedge S}^t\psi_s^{n,i}\,dW_s,$$
$$Y_t=H_S^\sigma(W_t-W_{t\wedge S})+\textstyle\int_{t\wedge S}^t H_s'^\sigma\,dW_s'+\int_{t\wedge S}^t\int_E\delta^\sigma(s,z)\,(\underline{p}-\underline{q}\,)(ds,dz),\qquad M_t'=\int_{t\wedge S}^t\psi_s^{n,i}Y_s\,dW_s.$$

Observe that

$$\rho_i^{n,1}=u_n\sigma_i^nM_T,\qquad\theta(2)_i^n=yu_n\big(\mu(1)_i^n+\mu(2)_i^n+\mu(3)_i^n\big),\qquad\text{where}$$
$$\mu(1)_i^n=M_T',\qquad\mu(2)_i^n=\textstyle\int_S^T\psi_s^{n,i}\Big(\int_S^s b_t^\sigma\,dt\Big)\,dW_s,\qquad\mu(3)_i^n=\int_S^T\psi_s^{n,i}\Big(\int_S^s(H_t^\sigma-H_S^\sigma)\,dW_t\Big)\,dW_s.$$

Then

$$|\overline{v}_i^{n,w}|\leq Ku_n^{w+1}\sum_{k=1}^3|v_{w,k}|,\quad\text{where}\quad v_{w,k}=\mathbb{E}\big(\mu(k)_i^n(M_T)^w\mid\mathcal{G}_i^n\big).\tag{5.39}$$

Since $\big|\int_S^s b_t^\sigma\,dt\big|\leq K(s-S)$, Doob's inequality and (5.11) yield $\mathbb{E}((\mu(2)_i^n)^2\mid\mathcal{G}_i^n)\leq K(h_n\Delta_n)^3$. For the case $k=3$ we first apply Lemma 14 to $V'=H^\sigma$, then (5.22) with V, which gives estimates for the process $V''=L'(V,W)$ (with $\theta_j^n=1$), and finally (5.22) again to the process V'' (with $\theta_j^n=g_i'^n$). Upon observing that $\mu(3)_i^n=L'(V'')_T$, and after some calculations, we end up with $\mathbb{E}((\mu(3)_i^n)^2\mid\mathcal{G}_i^n)\leq K(h_n\Delta_n)^3$. Using the first estimate in (5.30) and $u_n^2h_n\Delta_n\to 0$ (since $\kappa_1=1$ here), we thus get

$$k=2,3\quad\Rightarrow\quad|u_n^{w+1}v_{w,k}|\leq Kh_n\Delta_n.\tag{5.40}$$

Since Y,M,M' are martingales for $(\mathcal{H}^n_t)$ with integrable powers of any order and moreover $\langle M,Y\rangle_t = H^g_S\int_0^t \psi^{n,i}_s\,ds$ (recall that W and W' are orthogonal), we have by a repeated use of Itô's formula:

$$v_{1,1}=\mathbb{E}(\zeta_1\mid\mathcal{G}^n_i),\qquad v_{3,1}=\sum_{k=2}^6\mathbb{E}(\zeta_k\mid\mathcal{G}^n_i),\qquad\text{where}$$

$$\zeta_1=\int_S^T(\psi^{n,i}_s)^2Y_s\,ds,\qquad \zeta_2=(3+12H^g_S)\int_S^T(\psi^{n,i}_s)^2\Big(\int_S^s\psi^{n,i}_tM_tY_t\,dW_t\Big)\,ds,$$
$$\zeta_3=3\int_S^T(\psi^{n,i}_s)^2\Big(\int_S^s\psi^{n,i}_tM'_t\,dW_t\Big)\,ds,\qquad \zeta_4=6H^g_S\int_S^T(\psi^{n,i}_s)^2\Big(\int_S^sM_t^2\,dY_t\Big)\,ds,$$
$$\zeta_5=(3+6H^g_S)\int_S^T(\psi^{n,i}_s)^2\Big(\int_S^s(\psi^{n,i}_t)^2Y_t\,dt\Big)\,ds,\quad \zeta_6=12H^g_S\int_S^T(\psi^{n,i}_s)^2\Big(\int_S^s(\psi^{n,i}_t)^2M_t\,dt\Big)\,ds.$$

First, Lemma 14-(a) implies $Y\in\mathcal{P}(\Psi_j)\cap\widehat{\mathcal{P}}(\widehat{\Psi})$ with $\psi_j=K\Delta_n$ and $\widehat{\Psi}=Kh_n\Delta_n$. With the notation (5.15) we have $\zeta_1=L(Y)$ (with $\theta^n_i=(g'^n_i)^2$, hence (5.23) yields

$$|u_n^2v_{1,1}|\le Ku_n^2((h_n\Delta_n)^2+h_n\Delta_n^{3/2})\le K(h_n\Delta_n+u_n^2h_n\Delta_n^{3/2}). \tag{5.41}$$

Next, Lemma 14-(c) implies $MY\in\mathcal{P}(\Psi_j)\cap\widehat{\mathcal{P}}(\widehat{\Psi})$ with $\psi_j=K(\Delta_n^2+\Delta(\overline{M}_{0,j})^2+\Delta(\overline{Y}_{0,j})^2)$ and $\widehat{\Psi}=K(h_n\Delta_n)^2$; then (5.21) with $\theta^n_j=g'^n_j$ yields that $V=L'(MY,W)$ belong to $\mathcal{P}(\Psi_j)\cap\widehat{\mathcal{P}}(\widehat{\Psi})$ with $\widehat{\Psi}=K(h_n\Delta_n)^3$ and $\Psi_j=K\Delta_n(\Delta_n^2+\Delta_n\overline{M}^2_{0,j}+\Delta_n\overline{M}^2_{0,,j}+\overline{M}^2_{0,j}\overline{M}^2_{0,,j})$, whereas $\zeta_2=L(V)_T$ (with $\theta^n_j=(g'^n_j)^2$), hence (5.23) yields for $k=2$:

$$|\mathbb{E}(\zeta_k\mid\mathcal{G}^n_i)|\le K(h_n\Delta_n)^3. \tag{5.42}$$

Note that $U=Y$ satisfies the assumptions of Lemma 15 and M' (resp. M^2) belongs to $MY\in\mathcal{P}(\Psi_j)\cap\widehat{\mathcal{P}}(\widehat{\Psi})$ with $\widehat{\Psi}=K(h_n\Delta_n)^2$ and with $\psi_j=K(\Delta_n^2+\Delta(\overline{Y}_{0,j})^2)$ (resp. $\psi_j=K(\Delta_n^2+\Delta(\overline{M}_{0,j})^2)$) (for the case of M' we first use (5.21) and the fact that $M'=L(Y,W)$ with $\theta^n_j=g'^n_j$). Then the same argument shows that (5.42) holds for $k=3,4$. Furthermore, $\zeta_5=L''(Y)$ and $\zeta_6=L''(M)$ with $\theta^n_j=(g'^n_j)^2$, so upon using (5.23) we get $|\mathbb{E}(\zeta_k\mid\mathcal{G}^n_i)|\le K(h_n\Delta_n)^2(h_n\Delta_n+\sqrt{\Delta_n})$ for $k=5,6$. Summarizing, we deduce

$$|u_n^4v_{3,1}|\le K\big(h_n\Delta_n+u_n^4h_n^2\Delta_n^{5/2}\big).$$

and upon using (5.37), (5.38), (5.39), (5.40) and (5.41), we end up with

$$\begin{aligned}\big|\,\mathbb{E}\big(\theta(2)^n_i\,\sin(y\rho^n_i)\mid\mathcal{G}^n_i\big)\big|&\le K\kappa_1\big(h_n\Delta_n+u_n^2h_n\Delta_n^{3/2}\\&\quad+u_n^4h_n^2\Delta_n^{5/2}+u_n^6h_n^{1/2}\Delta_n^2+u_n^6(h_n\Delta_n)^{7/2}+u_n^5h_n^{-1}\Delta_n+u_n^{1+\widetilde{r}'/2}(h_n\Delta_n)^{3/2}\big).\end{aligned}$$

In turn, this combined with (5.35) and (5.36) gives us the first part of (5.29). $\square$

We can in fact cut the $\rho^{n,w}_i$ for $w=1,2$ into pieces corresponding to sub-intervals $[T(n,i+j),T(n,i+j+l)]$ of $[T(n,i),T(n,i+2h_n)]$ when $0\le j<l\le 2h_n$, as follows:

$$\rho^{n,1}_{i,j,l}=u_n\sigma^n_i\int_{T(n,i+j)}^{T(n,i+l)}\psi^{n,i}_s\,dW_s,\qquad \rho^{n,2}_{i,j,l}=u_n\int_{T(n,i+j)}^{T(n,i+l)}\int_E\delta^n_i(z)\psi^{n,i}_s(\underline{p}-\underline{q})(ds,dz),$$
$$\widehat{\rho}^n_{i,j,l}=\rho^{n,1}_{i,j,l}+\rho^{n,2}_{i,j,l}, \tag{5.43}$$

so $\widehat{\rho}^n_{i,0,2h_n}=\rho^{n,1}_i+\rho^{n,2}_i$. In all the sequel, since i is an index, we write $\iota=\sqrt{-1}$.

Lemma 17. *There are $\mathcal{G}^n_i$-measurable real-valued variables $B(y)^n_{i,q}$ and $B'(y)^n_{i,q}$ satisfying*

$$\begin{aligned}&\Big|\,B(y)^n_{i,q}-\tfrac12\,y^2(g'^n_q)^2u_n^2\Delta_nc^n_i-2\sum_{m=1}^M|y|^{\beta_m}u_n^{\beta_m}\Delta_n|g'^n_j|^{\beta_m}\chi(\beta_m)\,a^m_{T(n,i)}\Big|\le Ku_n^r\Delta_n,\\&|B'(y)^n_{i,q}|\le Ku_n^{\beta_1}\Delta_n,\qquad 1\le q\le h_n\ \Rightarrow\ B'(y)^n_{i,q}=-B'(y)^n_{i,q+h_n},\end{aligned} \tag{5.44}$$

for $0\leq g\leq 2h_n$*, and such that, if* $\overline{B}(y)^n_{i,j,l}=\sum_{q=j+1}^l B(y)^n_{i,q}$ *and* $\overline{B}'(y)^n_{i,j,l}=\sum_{q=j+1}^l B'(y)^n_{i,q}$*, we have for all* $0\leq j<l\leq 2h_n$*:*

$$\mathbb{E}\big(e^{\iota y\widehat{\rho}^n_{i,j,l}}\mid \mathcal{G}^n_{i+j}\big)=e^{-\lambda^n_{i+j}(\overline{B}(y)^n_{i,j,l}+\iota\overline{B}'(y)^n_{i,j,l})}+\mathrm{O}(u_n^\beta(l-j)\Delta_n^2). \tag{5.45}$$

Proof. 1) Set $\overline{v}_{n,q}=yu_n g'^n_q$ (so $|\overline{v}_{n,q}|\leq Ku_n$) and $V^n=(\delta^n_i\,1_{(T(n,i),\infty)})*(\underline{p}-\underline{q}\,)$. We have

$$y\widehat{\rho}^n_{i,j,l}=\sum_{q=j+1}^l\mu^n_q,\qquad \mu^n_q=\overline{v}_{n,q}\big(\sigma^n_i\Delta^n_{i+q}W+\Delta^n_{i+q}V^n\big).$$

We also consider the two functions of $v\in\mathbb{R}$:

$$G^n_i(v)=\int_{\mathbb{R}}(1-\cos(vx))\,F_{T(n,i)}(dx),\qquad H^n_i(u)=\int_{\mathbb{R}}(ux-\sin(vx))\,F_{T(n,i)}(dx),$$

which satisfy $G^n_i(v)+|H^n_i(v)|\leq K(v^2\wedge|v|^{\beta_1})$ because $\overline{F}_t(1)=0$ and $\overline{F}_t(z)\leq K/z^{\beta_1}$. Observe that, conditionally on $\mathcal{G}^n_{i+q-1}$, the process $\sigma^n_i(W_{T(n,i+q-1)+t}-W_{T(n,i+q-1)})+V^n_{T(n,i+q-1)+t}-V^n_{T(n,i+q-1)}$ for $t\geq 0$ is a Lévy process with Lévy measure $F_{T(n,i)}$ and variance c^n_i for the Gaussian part, independent of the variable Φ^n_{i+j}. Then,

$$\begin{aligned}\zeta^n_q:=\mathbb{E}\big(e^{\iota\mu^n_q}\mid\mathcal{G}^n_{i+q-1}\big)&=\mathbb{E}\Big(e^{-\lambda^n_{i+q-1}\Phi^n_{i+q}\Theta^n_q}\mid\mathcal{G}^n_{i+q-1}\Big),\\ \text{where}\quad \Theta^n_q&=\Delta_n\big(\tfrac12\,\overline{v}^2_{n,q}c^n_i+G^n_i(\overline{v}_{n,q})+\iota H^n_i(\overline{v}_{n,q})\big).\end{aligned}\tag{5.46}$$

Note that $|\Theta^n_q|\leq Ku_n^\beta\Delta_n$ and Θ^n_q is $\mathcal{G}^n_i$-measurable. The moment properties of Φ^n_{i+q} imply $\big|\mathbb{E}(e^{z\Phi^n_{i+q}}\mid\mathcal{G}^n_{i+q-1})-e^z\big|\leq K|z|^2$, uniformly in $z\in\mathbb{C}$ with $\mathcal{R}e(z)\leq 0$, yielding

$$\big|\,\zeta^n_q-e^{-\lambda^n_{i+q-1}\Theta^n_q}\big|\leq K(\Theta^n_q)^2\leq Ku_n^{2\beta}\Delta_n^2. \tag{5.47}$$

The variables $\Theta'^n_q=\sum_{m=q}^l\Theta^n_m$ satisfy $|\Theta'^n_q|\leq Ku_n^\beta(l-q+1)\Delta_n$ for $q=1,\dots,l$, so (5.12) yields

$$\mathbb{E}\Big(\big|\,e^{-\lambda^n_{i+q}\Theta'^n_{q+1}}-e^{-\lambda^n_{i+q-1}\Theta'^n_{q+1}}\big|\mid\mathcal{G}^n_{i+q-1}\Big)\leq K|\Theta'^n_{q+1}|\,\mathbb{E}(|\lambda^n_{i+q}-\lambda^n_{i+q-1}|\mid\mathcal{G}^n_{i+q-1})\leq Ku_n^\beta\Delta_n^2. \tag{5.48}$$

Now, with the notation $\Gamma^n_q=\exp\big(\iota\sum_{m=q}^l\mu^n_m\big)$, we will prove that

$$\big|\,\mathbb{E}(\Gamma^n_q\mid\mathcal{G}^n_{i+q-1})-e^{-\lambda^n_{i+q-1}\Theta'^n_q}\big|\leq 2Cu_n^\beta(l-q+1)\Delta_n^2, \tag{5.49}$$

for $q=1,\dots,l$, by downward induction on q, and where C is a constant ay least as big as the constants K showing in (5.47) and (5.48). When $q=l$ this readily follows from (5.46) and (5.47). Now, applying successively (5.49) for $q+1$, (5.48) and (5.47), and using also the $\mathcal{G}^n_{i+q}$-measurability of μ^n_q and the $\mathcal{G}^n_{i+q-1}$-measurability of $\lambda^n_{i+q-1}\Theta'^n_{q+1}$, we get with $a_{n,\dots}=\mathrm{O}_1(b_{n,\dots})$ meaning $|a_{n,\dots}|\leq[b_{n,\dots}|$:

$$\begin{aligned}\mathbb{E}(\Gamma^n_q\mid\mathcal{G}^n_{i+q-1})\ &=\mathbb{E}\big(\Gamma^n_{q+1}\,e^{\iota\mu^n_q}\mid\mathcal{G}^n_{i+q-1}\big)\\ &=\mathbb{E}\big(e^{\iota\mu^n_q}\,e^{-\lambda^n_{i+q}\Theta'^n_{q+1}}\mid\mathcal{G}^n_{i+j-1}\big)+2C\,\mathrm{O}_1((l-q)u_n^\beta\Delta_n^2)\\ &=\mathbb{E}\big(e^{\iota\mu^n_q}\,e^{-\lambda^n_{i+q-1}\Theta'^n_{q+1}}\mid\mathcal{G}^n_{i+q-1}\big)+2C\,\mathrm{O}_1((l-q+\tfrac12)u_n^\beta\Delta_n^2)\\ &=\zeta^n_q\,e^{-\lambda^n_{i+q-1}\Theta'^n_{q+1}}+2C\,\mathrm{O}_1((l-q+\tfrac12)u_n^\beta\Delta_n^2)\\ &=e^{-\lambda^n_{i+q-1}\Theta'^n_q}+2C\,\mathrm{O}_1((l-q+1)u_n^\beta\Delta_n^2).\end{aligned}$$

Then (5.49) is proved and, applied with $q=j+1$, it yields

$$\mathbb{E}\big(e^{\iota y\widehat{\rho}^n_{i,j,l}}\mid\mathcal{G}^n_{i+j}\big)=e^{-\lambda^n_i\Theta'^n_{j+1}}+\mathrm{O}(u_n^\beta(l-j)\Delta_n^2).$$

This is (5.45), upon taking

$$B(y)^n_{i,q}=\Delta_n\big(\frac{1}{2}\,\overline{v}^2_{n,q}c^n_i+G^n_i(\overline{v}_{n,q})\big),\qquad B'(y)^n_{i,q}=H^n_i(\overline{v}_{n,q}). \tag{5.50}$$

2) It remains to prove (5.43). The function H^n_i is odd with $|H^n_i(u)|\leq Ku^{\beta_1}$, whereas $\overline{v}_{n,j}=-\overline{v}_{n,h_n+j}$ when $0\leq j<h_n$, hence the second part of the claim. For the first part we first observe that, recalling that $\breve{F}_t$ is the symmetrized version of the measure F_t and since the cosine function is even, we have

$$G^n_i(u)=\int_{\mathbb{R}}(1-\cos(ux))\,\breve{F}_{T(n,i)}(dx).$$

Then, recalling (2.5), for simplicity, we write $M=|F'_{T(n,i)}|$, with its tail function $\overline{M}$, and also $T=T(n,i)$. The last part of (5.5) yields $G^n_i(v)=\sum_{m=1}^M A_m(v)+\mathrm{O}(A'(v))$, where

$$A_m(v)=\int_{\{|x|\leq 1\}}\beta_m\,a^m_T\,\frac{1-\cos(vx)}{|x|^{1+\beta_m}}\,dx,\qquad A'(v)=\int_{\{|x|\leq 1\}}\big(1-\cos(vx)\big)\,M(dx).$$

First, by symmetry, change of variable and integration by parts, and for $v>1$,

$$A_m(v)=2\beta_m v^{\beta_m}\,a^m_T\int_0^v\frac{1-\cos x}{x^{1+\beta_m}}\,dx=2\beta_m v^{\beta_m}\,a^m_T\int_0^\infty\frac{1-\cos x}{x^{1+\beta_m}}\,dx+\mathrm{O}(1)=2a^m_T v^{\beta_m}\chi(\beta_m)+\mathrm{O}(1).$$

Second, $1-\cos(vx)\leq|vx|^2\wedge 2$ and Fubini's theorem yield

$$A'(v)=2\overline{M}(1/v)+v^2\int_0^{1/v}x\,\overline{M}(x)\,dx\leq K\Big(v^r+v^2\int_0^{1/v}x^{1-r}\,dx\Big)\leq Kv^r.$$

Therefore we get $G^n_i(v)=2\sum_{m=1}^M a^m_T v^{\beta_m}\chi(\beta_m)+\mathrm{O}(v^r)$ as $v\to\infty$, hence by substituting v with $\overline{v}_{n,j}$ and using (5.50), we have the first part of (5.43). □

LEMMA 18. *We have, for any $p\geq 2$:*

$$\begin{aligned}&\big|\mathbb{E}(\cos(y\rho^n_i)-U(y)^n_{T(n,i)}\mid\mathcal{G}^n_i)\big|\leq K\alpha_n^3,\\&\mathbb{E}\big((\cos(y\rho^n_i)-U(y)^n_{T(n,i)})(\cos(y'\rho^n_i)-U(y')^n_{T(n,i)}\mid\mathcal{G}^n_i\big)=\tfrac{1}{2}\,V^{\#}(y,y')_{T(n,i)}+\mathrm{O}(\alpha_n^3),\\&\mathbb{E}(|\cos(y\rho^n_i)-U(y)^n_{T(n,i)}|^p\mid\mathcal{G}^n_i)\leq K_p\,\alpha_n^4.\end{aligned} \tag{5.51}$$

Proof. 1) In a first step, we compute the variable $\mathbb{E}\big(e^{\iota y\rho^{n,3}_i}\mid\widehat{\mathcal{G}}^n_i\big)$, and for this we set for $v\in\mathbb{R}$:

$$\Psi^n_j(v)=\mathbb{E}\big(e^{\iota v\varepsilon^n_j}\mid\widehat{\mathcal{G}}^n_i\big).$$

The properties (2.15) or (2.16) yield $|\Psi^n_j(v)-1+v^2/2+\iota\,v^3(\gamma^{(3)})^n_j/6|\leq Kv^4$, hence also, with the notation $\overline{w}_{n,j}=yu_n\overline{g}'^n_j\gamma'^n_i$ (so $|\overline{w}_{n,j}|\leq Ku_n/h_n$),

$$\mathbb{E}\big(e^{\iota y\rho^{n,3}_i}\mid\widehat{\mathcal{G}}^n_i\big)=\prod_{j=0}^{2h_n-1}\Psi^n_{(i+j)\Delta_n}=A_n+\mathrm{O}(u_n^4/h_n^3),\quad A_n=\prod_{j=0}^{2h_n-1}\Big(1-\frac{\overline{w}^2_{n,j}}{2}-\frac{\iota}{6}\,\overline{w}^3_{n,j}(\gamma^{(3)})^n_{i+j}\Big).$$

We can go further, and compare A_n with the variable

$$A'_n=\prod_{j=0}^{2h_n-1}\Big(1-\frac{\overline{w}_{n,j}^2}{2}-\frac{\iota\,\overline{w}_{n,j}^3}{6}\,(\gamma^{(3)})_i^n\Big)=\prod_{j=0}^{h_n-1}\Big(\big(1-\frac{\overline{w}_{n,j}^2}{2}\big)^2+\frac{\overline{w}_{n,j}^6}{36}\,((\gamma^{(3)})_i^n)^2\Big),$$

where the last equality comes from $\overline{w}_{n,j+h_n}=-\overline{w}_{n,j}$ for $0\leq j<h_n$. Since each factor in the definition of A_n has an absolute value smaller than 1 for all n large enough, we have $|A_n-A'_n|\leq K\sum_{j=1}^{2h_n-1}\frac{u_n^3}{h_n^3}\,|(\gamma^{(3)})_{i+j}^n-(\gamma^{(3)})_i^n|$. Since $\mathbb{E}(|(\gamma^{(3)})_{i+j}^n-(\gamma^{(3)})_i^n|\mid\mathcal{G}_i^n)\leq K\kappa_2\sqrt{j\Delta_n}$ by (5.5), we deduce

$$\mathbb{E}\big(|A_n-A'_n|\mid\mathcal{G}_i^n\big)\leq K\kappa_2u_n^3\Delta_n^{1/2}/h_n^{3/2}.$$

Moreover, $\big|\,A'_n-\exp\big(-\sum_{j=0}^{h_n-1}\overline{w}_{n,j}^2\big)\big|\leq Ku_n^4/h_n^3$ and $\sum_{j=0}^{h_n-1}\overline{w}_{n,j}^2=u_n^2\overline{\phi}_n/h_n$, hence

$$\mathbb{E}\Big(\big|\,\mathbb{E}\big(e^{\iota y\rho_i^{n,3}}\mid\widehat{\mathcal{G}}_i^n\big)-e^{-\gamma(y)_{T(n,i)}^n}\big|\mid\mathcal{G}_i^n\Big)\leq K\big(\kappa_2u_n^3h_n^{-3/2}\Delta_n^{1/2}+u_n^4h_n^{-3}\big).\tag{5.52}$$

2) We complement (5.24) with (for some fixed y):

$$U_t'^n=e^{-c(y)_t^n\lambda_t-a(y)_t^n\lambda_t},\qquad U_t''^n=e^{-\gamma(y)_t^n},$$

so $U(y)_t^n=U_t'^nU_t''^n$. Observe that

$$\mathbb{E}\big(e^{\iota y\rho_i^n}\mid\mathcal{G}_i^n\big)=\mathbb{E}\big(e^{\iota y\rho_i^{n,1}+\iota y\rho_i^{n,2}}\mathbb{E}\big(e^{\iota y\rho_i^{n,3}}\mid\widehat{\mathcal{G}}_i^n\big)\mid\mathcal{G}_i^n\big),$$

so (5.52) gives us

$$\big|\,\mathbb{E}\big(e^{\iota y\rho_i^n}-U(y)_{T(n,i)}^n\mid\mathcal{G}_i^n\big)\big|\leq\big|\,\mathbb{E}\big(e^{\iota y\widehat{\rho}_{i,0,2h_n}^n}-U_{T(n,i)}'^n\mid\mathcal{G}_i^n\big)\big|+K\big(\kappa_2u_n^3h_n^{-3/2}\Delta_n^{1/2}+u_n^4h_n^{-3}\big).$$

(5.44) yields $|\overline{B}(y)_{i,0,2h_n}^n-c(y)_{T(n,i)}^n-a(y)_{T(n,i)}^n|\leq Ku_n^rh_n\Delta_n$ and $\overline{B}'(y)_{i,0,2h_n}^n=0$, so (5.45) implies

$$\big|\,\mathbb{E}\big(e^{\iota y\widehat{\rho}_{i,0,2h_n}^n}-U_{T(n,i)}'^n\mid\mathcal{G}_i^n\big)\big|\leq K(u_n^rh_n\Delta_n+u_n^\beta h_n\Delta_n^2)\leq Ku_n^rh_n\Delta_n.$$

Taking the real part above, we deduce the first part of (5.51). Upon using $\cos(x)\cos(x')=\frac12\big(\cos(x+x')+\cos(x-x')\big)$, the second part of (5.51) is a trivial consequence of the first part, plus the definition of $V^\#(y,y')_t^n$. For the last part, since the integrand is bounded, it suffices to show it when $p=2$, in which case it follows from the second part with $y'=y$ and (5.27). □

LEMMA 19. *For all $p\geq 2$ we have*

$$\begin{aligned}&\mathbb{E}(\xi(y)_j^{1,n}\mid\mathcal{G}_{jw_n}^n)=\mathrm{O}(\alpha_n^3),\\&\mathbb{E}(\xi(y)_j^{1,n}\,\xi(y')_j^{1,n}\mid\mathcal{G}_{jw_n}^n)=\tfrac{1}{2k_n}\,V^\#(y,y')_{T(n,jw_n)}^n+\mathrm{O}\big(\alpha_n^3+\tfrac{\alpha_n^5}{k_n}\big),\\&\mathbb{E}(|\xi(y)_j^{1,n}|^p\mid\mathcal{G}_{jw_n}^n)=\mathrm{O}(\overline{\alpha}(p)_n),\end{aligned}\tag{5.53}$$

$$\mathbb{E}(\xi(y)_j^{2,n}\mid\mathcal{G}_{jw_n}^n)=\mathrm{O}(\alpha_n^1),\qquad\mathbb{E}(|\xi(y)_j^{2,n}|^p\mid\mathcal{G}_{jw_n}^n)=\mathrm{O}(\widehat{\alpha}(p)_n),\tag{5.54}$$

$$\mathbb{E}(\xi(y)_j^{3,n}\mid\mathcal{G}_{jw_n}^n)=\mathrm{O}(\alpha_n^5),\qquad\mathbb{E}(|\xi(y)_j^{3,n}|^p\mid\mathcal{G}_{jw_n}^n)=\mathrm{O}(\widehat{\alpha}(p)_n).\tag{5.55}$$

Proof. We have $\xi(y)^{w,n}_j = \frac{1}{k_n}\sum_{l=0}^{k_n-1}\zeta(y)^w_l$, where $\zeta(y)^w_l$ is the l-th summand in the definition of $\xi(y)^{w,n}_j$. When $w=1,2$ we also set $\zeta(y)'^w_l = \mathbb{E}(\zeta(y)^w_l \mid \mathcal{G}_{2h_n(jk_n+l)})$ and $\zeta(y)''^w_l = \zeta(y)^w_l - \zeta(y)'^w_l$.

1) (5.51) yields $|\zeta(y)'^1_k| \le K\alpha^3_n$, and the first part of (5.53) follows. Next, $\xi(y)^{1,n}_j\,\xi(y')^{1,n}_j$ is the sum of the k_n^2 terms $a_{l,l'} = \zeta(y)^1_l\zeta(y')^1_{l'}/k_n^2$. For the off-diagonal terms, say when $l<l'$, we have $\mathbb{E}(a_{l,l'} \mid \mathcal{G}^n_{2h_n(jk_n+l')}) = \zeta(y)^1_l\zeta(y')'^1_{l'}/k_n^2$, whereas $|\zeta(y)^1_l| \le K$: hence the $\mathcal{G}^n_{jw_n}$-conditional expectation of the total contribution of those off-diagonal terms is $\mathrm{O}(\alpha^3_n)$. (5.51) again gives us $\mathbb{E}(a_{l,l} \mid \mathcal{G}^n_{2h_n(jk_n+l)}) = \frac{1}{2k_n^2}\,V^{\#}(y,y')^n_{T(n,2h_n(jk_n+l))} + \mathrm{O}(\alpha^3_n/k_n^2)$. In view of (5.26) for $V^{\#}(y,y')^n$, we deduce $\mathbb{E}(a_{l,l} \mid \mathcal{G}^n_{jw_n}) = \frac{1}{2k_n^2}\big(V^{\#}(y,y')^n_{T(n,jw_n)} + \mathrm{O}(\alpha^3_n+\alpha^5_n)\big)$, and the second part of (5.53) follows.

Finally, $\xi(y)^{1,n}_j = A'+A''$, where $A' = \frac{1}{k_n}\sum_{l=0}^{k_n-1}\zeta(y)'^1_l$ and $A'' = \frac{1}{k_n}\sum_{l=0}^{k_n-1}\zeta(y)''^1_l$. We have seen $|A'| \le K\alpha^3_n$ and, by the Burkholder-Gundy and Hölder's inequalities, we have for all $p\ge 2$:

$$\mathbb{E}(|A''|^p \mid \mathcal{G}^n_{jw_n}) \le K_p\,\frac{1}{k_n^p}\,\mathbb{E}\Big(\Big(\sum_{l=0}^{k_n-1}(\zeta(y)''^1_l)^2\Big)^{p/2} \mid \mathcal{G}^n_{jw_n}\Big) \le K_p\,\frac{1}{k_n^{p/2}}\sum_{l=0}^{k_n-1}\mathbb{E}\big((\zeta(y)^1_l)^p \mid \mathcal{G}^n_{jw_n}\big),$$

which is smaller than $K_p\alpha^4_n/k_n^{p/2}$ by (5.26). The third estimate in (5.53) follows.

2) For (5.54) we argue in exactly the same way, except that we now use (5.29) (the proof is in fact quite simpler). For (5.55) the first estimate directly follows from (5.26), and the second one from the same and Hölder's inequality. □

Lemma 20. *For all $p\ge 2$ and $j<[t/w_n]$, we have*

$$\begin{aligned}
&\mathbb{E}(\xi(y)^n_j \mid \mathcal{G}^n_{jw_n}) = \mathrm{O}(\alpha^1_n+\alpha^3_n+\alpha^5_n),\\
&\mathbb{E}(\xi(y)^n_j\,\xi(y')^n_j \mid \mathcal{G}^n_{jw_n}) = \tfrac{1}{k_n}\,V(y,y')^n_{jw_n} + \mathrm{O}\big(\alpha^3_n + \tfrac{\alpha^5_n}{k_n} + \widehat{\alpha}(2)_n + \sqrt{\overline{\alpha}(2)_n\widehat{\alpha}(2)_n}\,\big),\\
&\mathbb{E}(|\xi(y)^n_j|^p \mid \mathcal{G}^n_{jw_n}) = \mathrm{O}\big(\overline{\alpha}(p)_n+\widehat{\alpha}(p)_n\big).
\end{aligned}\tag{5.56}$$

Proof. In view of (5.25) and of the previous lemma, the first and last parts of (5.56) are obvious. For the second part, in view of (5.53) it is enough to prove that

$$|\mathbb{E}(\xi(y)^{z,n}_j\,\xi(y')^{w,n}_j \mid \mathcal{G}^n_{jw_n})| \le K\big(\alpha^3_n + \frac{\alpha^5_n}{k_n} + \widehat{\alpha}(2)_n + \sqrt{\overline{\alpha}(2)_n\widehat{\alpha}(2)_n}\,\big),$$

for all $z,w=1,2,3$ but $z=w=1$. These properties follow from the Cauchy-Schwarz inequality and (5.55) with $q=2$. □

We also need some estimates on the variables $\widehat{Y}^n_j$ of (3.5):

Lemma 21. *For all $p\ge 2$ and all integers $j<[t/w_n]$ and $k\ge 2$, we have*

$$\begin{aligned}
&\mathbb{E}(\widehat{Y}^n_j \mid \mathcal{G}^n_{jw_n}) = 2\gamma_{T(n,jw_n)} + \mathrm{O}(k_nh_n\Delta_n + \kappa_2\sqrt{k_nh_n\Delta_n}\,),\\
&\mathbb{E}((\widehat{Y}^n_j)^k \mid \mathcal{G}^n_{jw_n}) = (2\gamma_{T(n,jw_n)})^k + \mathrm{O}\big(k_nh_n\Delta_n + \kappa_2\sqrt{k_nh_n\Delta_n} + \tfrac{1}{k_nh_n}\big),\\
&\mathbb{E}\big(((\widehat{Y}^n_j)^k - (2\gamma_{T(n,jw_n)})^k)^2 \mid \mathcal{G}^n_{jw_n}\big) = \mathrm{O}\big(k_nh_n\Delta_n + \kappa_2\sqrt{k_nh_n\Delta_n} + \tfrac{1}{k_nh_n}\big).
\end{aligned}\tag{5.57}$$

Proof. The properties of the noise and (5.12) with $V=X$ and $V=\gamma$ imply for all $p>0$:

$$\begin{aligned}
&\mathbb{E}((\Delta^n_iY^n)^2 \mid \mathcal{G}^n_i) = \mathbb{E}((\Delta^n_iX)^2 + \gamma^n_i + \gamma^n_{i+1} \mid \mathcal{G}^n_i) = 2\gamma^n_i + \mathrm{O}(\Delta_n + \kappa_2\sqrt{\Delta_n}\,),\\
&\mathbb{E}(|\Delta^n_iY^n|^p \mid \mathcal{G}^n_i) \le K.
\end{aligned}$$

Then (5.12) with $V=\gamma$ again yields the first part of (5.57).

Next, let $i\leq j_1<\cdots<j_k\leq i+w_n\leq [t/w_n]$ with $j_l>j_{l-1}+1$. The properties of the noise and successive conditioning allow us to write

$$\begin{aligned}\mathbb{E}\big(\prod_{l=1}^k(\Delta_{j_l}^nY^n)^2\mid\mathcal{G}_i^n\big) &= \mathbb{E}\Big(2\gamma_{j_k}^n\prod_{l=1}^{k-1}(\Delta_{j_l}^nY^n)^2\mid\mathcal{G}_i^n\Big)+\mathrm{O}(\Delta_n+\kappa_2\sqrt{\Delta_n}\,)\\ &= \mathbb{E}\Big(2\gamma_{j_{k-1}+2}^n\prod_{l=1}^{k-1}(\Delta_{j_l}^nY^n)^2\mid\mathcal{G}_i^n\Big)+\mathrm{O}(k_nh_n\Delta_n+\kappa_2\sqrt{k_nh_n\Delta_n}\,)\\ &= \mathbb{E}\Big(2\gamma_{j_{k-1}}^n\prod_{l=1}^{k-1}(\Delta_{j_l}^nY^n)^2\mid\mathcal{G}_i^n\Big)+\mathrm{O}(k_nh_n\Delta_n+\kappa_2\sqrt{k_nh_n\Delta_n}+\sqrt{\Delta_n}\,),\end{aligned}$$

and we deduce by induction that

$$\mathbb{E}\big(\prod_{l=1}^k(\Delta_{j_l}^nY^n)^2\mid\mathcal{G}_i^n\big)=(2\gamma_i^n)^k+\mathrm{O}(k_nh_n\Delta_n+\kappa_2\sqrt{k_nh_n\Delta_n}+\sqrt{\Delta_n}\,).$$

In the expansion of $(\widehat{Y}_i^n)^k$ as a sum of w_n^k terms of the form $\prod_{l=1}^k(\Delta_{j_l}^nY^n)^2$, the number of terms for which $|j_l-j_r|\leq 1$ for at least one pair (l,r) is less than Kw_n^{k-1}, so the second part of (5.57) follows (notice that $\sqrt{\Delta_n}\leq k_nh_n\Delta_n+1/k_nh_n$ always). This, upon expanding the square $((\widehat{Y}_j^n)^k-(2\gamma_{T(n,jw_n)})^k)^2$, yields the third part. □

5.5. *Reducing the problem.* Below, we basically reproduce Subsection 6.2 of [12], with a few changes. Observe that

$$f(\Upsilon(y)_t^n,\Upsilon(2y)_t^n)=V(y,y)_t^n,$$

and the two arguments of f above go to 0 as $n\to\infty$, uniformly in $y\in\mathcal{Y}$ and $t\geq 0$. We have $\log U(y)_t^n=-\Upsilon(y)_t^n$ and, by construction, $L(y)_j^n=U(y)_{T(n,jw_n)}^n(1+\xi(y)_j^n)$. Moreover, $U(y)_t^n\geq 1/\Gamma$ by (5.25) and there is a non random integer n_0 such that $h_n\geq 2\Gamma$ for $n\geq n_0$, implying $L(y)_j^n\geq 1/h_n$ whenever $1+\xi(y)_j^n\geq\frac12$. Hence we deduce that, if $j\leq[N_t^n/w_n]$,

$$n\geq n_0,\ \omega\in\Omega_{n,t}\quad\Longrightarrow\quad \widehat{c}(y)_j^n=\Upsilon(y)_{T(n,jw_n)}^n-\log(1+\xi(y)_j^n),$$

and in particular $|\widehat{c}(y)_j^n|\leq K$. Again on the set $\Omega_{n,t}$ and for $n\geq n_0$, we can expand $\log(1+x)$ around 0 and f around the pair $(\Upsilon(y)_{T(n,jw_n)}^n,\Upsilon(2y)_{T(n,jw_n)}^n)$ to obtain (since $|\xi(y)_i^n|\leq 1/2$, and with $\overline{\rho}(y)_j^n$ being suitable $\mathcal{G}_{jw_n}^n$-measurable variables with $|\overline{\rho}(y)_j^n|\leq K$:

$$\begin{aligned}&\widehat{c}(y)_j^n=\Upsilon(y)_{T(n,jw_n)}^n-\xi(y)_j^n+\tfrac12\,|\xi(y)_j^n|^2+\mathrm{O}\big(|\xi(y)_j^n|^3\big),\\ &f(\widehat{c}(y)_j^n,\widehat{c}(2y)_j^n)=V(y,y)_{T(n,jw_n)}^n+\overline{\rho}(y)_j^n\xi(y)_j^n+\overline{\rho}(2y)_j^n\xi(2y)_j^n+\mathrm{O}(|\xi(y)_j^n|^2+|\xi(2y)_j^n|^2).\end{aligned}$$

In turn, this yields on the set $\Omega_{n,t}$ and for $n\geq n_0$ again:

$$\begin{aligned}&\widehat{c}(y)_j^n-\Upsilon(y)_{T(n,jw_n)}^n-\tfrac{1}{2k_n}\,f(\widehat{c}(y)_j^n,\widehat{c}(2y)_j^n)\\ &\qquad=-\xi(y)_j^n-\tfrac{1}{2k_n}\,\big(\overline{\rho}(y)_j^n\xi(y)_j^n+\overline{\rho}(2y)_j^n\xi(2y)_j^n\big)\\ &\qquad\quad+\tfrac12\,|\xi(y)_j^n|^2-\tfrac{1}{2k_n}\,V(y,y)_{T(n,jw_n)}^n+\mathrm{O}\Big(\tfrac{|\xi(y)_j^n|^2+|\xi(2y)_j^n|^2}{k_n}+|\xi(y)_j^n|^3\Big).\end{aligned}\tag{5.58}$$

Recall $w_n=2k_nh_n$ and set $N_t'^n=[N_t^n/w_n]$. Observe that $Z(y)_t^n=\widetilde{V}_t^{n,y}+\sum_{l=1}^2V_t^{n,y,l}$, where

$$\begin{aligned}V_t^{n,y,1}&=-\Big(\widehat{c}(y)_{N_t'^n}^n-\tfrac{1}{2k_n}\,f(\widehat{c}(y)_{N_t'^n}^n,\widehat{c}(2y)_{N_t'^n}^n)-\tfrac{1}{2h_n}\,y^2\overline{\phi}_nu_n^2\widehat{Y}_{w_nN_t'^n}^n\Big),\\ V_t^{n,y,2}&=\sum_{j=0}^{N_t'^n}\Big(\Upsilon(y)_{T(n,jw_n)}^n-\tfrac{1}{2h_n}\,\overline{\phi}_ny^2u_n^2\widehat{Y}_{jw_n}^n\Big)-\tfrac{y^2u_n^2\phi_n}{2k_n}\,C_t-\tfrac{2}{k_n}\sum_{m=1}^M|y|^{\beta_m}u_n^{\beta_m}\widetilde{\phi}_n^{\beta_m}\chi(\beta_m)\,A_t^m,\\ \widetilde{V}_t^{n,y}&=\sum_{j=0}^{N_t'^n}\Big(\widehat{c}(y)_j^n-\Upsilon(y)_{T(n,jw_n)}^n-\tfrac{1}{2k_n}\,f(\widehat{c}(y)_j^n,\widehat{c}(2y)_j^n)\Big).\end{aligned}$$

The reason for summing up to N'^n_t, instead of $N'^n_t - 1$, is that the j-th summands above are measurable with respect to $\mathcal{G}^n_{(j+1)w_n}$ and, in order to apply classical results for triangular arrays we need that, for any n, the sum over j is taken up to a stopping time for the discrete-time filtration $(\mathcal{G}^n_{(j+1)w_n})_{j\geq 0}$: this is true of N'^n_t, but not of $N'^n_t - 1$ in general.

We also introduce the following processes:

$$V^{n,y,3}_t = \sum_{j=0}^{N'^n_t} \big(\tfrac12 \,|\xi(y)^n_j|^2 - \tfrac{1}{2k_n}\, V(y,y)^n_{T(n,jw_n)}\big), \quad V^{n,y,4}_t = -\sum_{j=0}^{N'^n_t} \tfrac{1}{2k_n}\, \big(\overline{\rho}(y)^n_j \xi(y)^n_j + \overline{\rho}(2y)^n_j \xi(2y)^n_j\big),$$
$$V^{n,y}_t = -\sum_{j=0}^{N'^n_t} \xi(y)^n_j, \qquad R^{n,y}_t = \sum_{j=0}^{N'^n_t} \Big(\tfrac{|\xi(y)^n_j|^2 + |\xi(2y)^n_j|^2}{k_n} + |\xi(y)^n_j|^3\Big).$$

By virtue of (3.8) and (5.58) we then obtain

$$\Big| Z(y)^n_s - V^{n,y}_s - \sum_{l=1}^4 V^{n,y,l}_s \Big| \leq K R^{n,y}_t \qquad \text{on } \Omega_{n,t}, \ \text{ for all } s \leq t.$$

Therefore, for Theorems 1–3 it is enough to prove (i) below, and either (ii) or (iii) or (iv), for appropriate rates $v_n, \overline{v}_n$ with $v_n/\overline{v}_n \to 0$ in (iv), depending on the case:

$$(5.59)\qquad \begin{array}{ll} \text{(i)} & \mathbb{P}(\Omega_{n,t}) \to 1, \qquad \overline{v}_n V^{n,y,l}_t \stackrel{\mathbb{P}}{\longrightarrow} 0 \ \text{ for } l=1,2,3,4, \qquad \overline{v}_n R^{n,y}_t \stackrel{\mathbb{P}}{\longrightarrow} 0, \\ \text{(ii)} & \overline{v}_n V^{n,y}_t \stackrel{\mathbb{P}}{\longrightarrow} 0, \\ \text{(iii)} & (\overline{v}_n V^{n,y}_t)_{y\in\mathcal{Y}} \ \text{ converges } \mathcal{F}_\infty\text{-stably in law}, \\ \text{(iv)} & \big(v_n V^{n,1}_t, (\overline{v}_n \overline{V}^{n,y}_t)_{y\in\mathcal{Y}}\big) \ \text{ converges } \mathcal{F}_\infty\text{-stably in law.} \end{array}$$

We prove (i) in the forthcoming subsection, and (ii)–(iv) in the next one.

5.6. *Technical lemmas.*

LEMMA 22. *We have $\mathbb{P}(\Omega_{n,t}) \to 1$, as soon as*

$$(5.60)\qquad \frac{1}{k_n h_n \Delta_n}\,\big(\overline{\alpha}(p)_n + \widehat{\alpha}(p)_n\big) \to 0 \quad \textit{for } p \textit{ large enough.}$$

Proof. Since $N'^n_t \leq K_t/w_n\Delta_n$ by (5.2), the claim is implied by the following consequence of (5.56):

$$\mathbb{P}((\Omega(y)_{n,t})^c) \leq 2^p\, \mathbb{E}\Big(\sum_{j=0}^{N'^n_t} |\xi(y)^n_j|^p\Big) \leq 2^p \sum_{j=0}^{[K_t/w_n\Delta_n]} \mathbb{E}(|\xi(y)^n_j|^p) \leq K_{t,p} \frac{\overline{\alpha}(p)_n + \widehat{\alpha}(p)_n}{k_n h_n \Delta_n}.$$

□

LEMMA 23. *Let G be a càdlàg bounded $(\mathcal{F}_t)$-adapted satisfying $(2.3)_{2,q}$ and $(2.4)_{\overline{q}}$ with $\overline{q} \leq 1$. As soon as $v'_n\big((k_n h_n \Delta_n)^{\overline{q}\wedge \frac{1+q}{2}} + \Delta_n^{\frac{q}{2}}\big) \to 0$, we have*

$$G^n_t := v'_n\Big(w_n\Delta_n \sum_{j=0}^{N'^n_t} (G\lambda)_{T(n,jw_n)} - \int_0^t G_s\, ds\Big) \stackrel{\mathbb{P}}{\longrightarrow} 0.$$

Proof. Recalling (2.10), we have $G_t^n = \sum_{l=1}^4 G_t^{n,l}$, where

$$\begin{aligned}
G_t^{n,1} &= v'_n \int_t^{T(n,w_n(N'^n_t+1))} G_s\,ds,\\
G_t^{n,2} &= -\sum_{j=0}^{N'^n_t} v'_n \int_{T(n,jw_n)}^{T(n,(j+1)w_n)} \big(G_s - G_{T(n,jw_n)}\big)\,ds,\\
G_t^{n,3} &= -\sum_{j=0}^{N'^n_t} v'_n \Delta_n\, G_{T(n,jw_n)} \sum_{m=1}^{w_n} \lambda_{T(n,jw_n+m-1)}(\Phi^n_{jw_n+m}-1),\\
G_t^{n,4} &= -\sum_{j=0}^{N'^n_t} v'_n \Delta_n\, G_{T(n,jw_n)} \sum_{m=1}^{w_n} (\lambda_{T(n,jw_n+m-1)} - \lambda_{T(n,jw_n)}),
\end{aligned}$$

and we will show $G_t^{n,l} \stackrel{\mathbb{P}}{\longrightarrow} 0$ for $l = 1, 2, 3, 4$.

The first part of (5.12) with $p = q = 1$ for $V_t = \lambda_t$ plus $N^n_t \leq (1+\Gamma t)/\Delta_n$ by (5.2), hence $N'^n_t \leq K_t/k_nh_n\Delta_n$, yield $\mathbb{E}(|G_t^{n,4}|) \leq K_t v'_n k_n h_n \Delta_n$, which goes to 0 by hypothesis (recall $\overline{q} \leq 1$).

For the case $l = 1$ we need a preliminary result. By (3.2), there is some $\varepsilon > 0$ such that $k_nh_n\Delta_n^{\varepsilon} \to \infty$. Set $B_n = \{\Delta(n,i) \leq \Delta_n^{1-\varepsilon} : i = 1, \ldots, N^n_t\}$. (2.10) and $\lambda_t \leq \Gamma$ imply $\Delta(n,i) \leq \Gamma\Delta_n\Phi^n_i$. Then, upon using (2.11) with $p = 2/\varepsilon$ and again $N^n_t \leq (1+\Gamma t)/\Delta_n$, we get

$$\mathbb{P}(B_n^c) \leq \sum_{i=1}^{[(1+\Gamma t)/\Delta_n]} \mathbb{P}(\Gamma\Phi^n_i\Delta_n^{\varepsilon} > 1) \leq (\Gamma\Delta_n^{\varepsilon})^p \sum_{i=1}^{[(1+\Gamma t)/\Delta_n]} \mathbb{E}((\Phi^n_i)^p) \leq K_t\Delta_n.$$

Since $T(n, w_n(N'^n_t+1)) \leq T(n, N^n_t + w_n)$ and $|G_s| \leq K$, we see that for all $\varepsilon' > 0$ we have

$$\mathbb{P}(|G_t^{n,1}| > \varepsilon') \leq \mathbb{P}(B_n^c) + \frac{Kv'_n}{\varepsilon'}\,\mathbb{E}\big(\Delta_n^{1-\varepsilon} + (T(n, N^n_t + w_n) - T(n, N^n_t+1)\big).$$

Hence, since the set $\{T(n,i-1) \leq t < T(n,i)\}$ belongs to $\mathcal{G}^n_i$, (5.11) yields

$$\begin{aligned}
\mathbb{P}(|G_t^{n,1}| > \varepsilon') &\leq K_t\Delta_n + \tfrac{Kv'_n}{\varepsilon'}\Big(\Delta_n^{1-\varepsilon} + \sum_{i=0}^{\infty} \mathbb{E}\big(\big(T(n,i+w_n) - T(n,i)\big)\,1_{\{T(n,i-1)\leq t < T(n,i)\}}\big)\Big)\\
&\leq K_t\Delta_n + \tfrac{Kv'_n}{\varepsilon'}\Big(\Delta_n^{1-\varepsilon} + w_n\Delta_n\sum_{i=0}^{\infty}\mathbb{P}(T(n,i-1) \leq t < T(n,i))\Big)\\
&\leq K_{t,\varepsilon'}\big(\Delta_n + v'_n\Delta_n^{1-\varepsilon} + v'_nh_nk_n\Delta_n\big) \;\leq\; K_{t,\varepsilon'}\big(\Delta_n + v'_nh_nk_n\Delta_n\big),
\end{aligned}$$

where the last inequality comes from our choice of ε. The claim for $l = 1$ follows.

For the cases $l = 2, 3$, we use a martingale-type argument. We denote by ζ^n_j the j-th summand in $G_t^{n,l}$, and use the property that for each n the sequence $(\zeta^n_j : j \geq 0)$ is adapted to the discrete-time filtration $(\mathcal{G}^n_{(j+1)w_n})_{j\geq -1}$, whereas N'^n_t is a stopping time for this filtration, Then, with the notation $\zeta'^n_j = \mathbb{E}(\zeta^n_j \mid \mathcal{G}^n_{jw_n})$, the claim is implied by the convergences $\sum_{i=1}^{N'^n_t} |\zeta'^n_j| \stackrel{\mathbb{P}}{\longrightarrow} 0$ and $\sum_{j=1}^{N'^n_t} (\zeta^n_j)^2 \stackrel{\mathbb{P}}{\longrightarrow} 0$. In view of (5.2), so $N'^n_t \leq K_t/w_n\Delta_n$, it is thus enough to show that

$$\theta_n := \sup_j \mathbb{E}(|\zeta'^n_j|) = \mathrm{o}(k_nh_n\Delta_n), \qquad \theta'_n := \sup_j \mathbb{E}(\zeta^n_j)^2) = \mathrm{o}(k_nh_n\Delta_n). \tag{5.61}$$

Note that the results of Subsection 5.3 apply for any sequence h_n of integers with $h_n\Delta_n \leq K$, and below we use them with k_nh_n instead of h_n, and also with $\theta^n_j = 1$, hence $\Theta_t = 1$, in (5.15), and with $i = jw_n$. When $l = 2$ we have $\zeta^n_j = v'_nL(V)_{T(n,(j+1)w_n)}$ with the process $V_t = G_{t\bigvee T(n,jw_n)} - G_{T(n,jw_n)}$, which belongs to $\mathcal{P}(\Psi_j) \cap \overline{\mathcal{P}}(\overline{\Psi}) \cap \widehat{\mathcal{P}}(\widehat{\Psi})$ with $\Psi_j = K\Delta_n^q$ and $\widehat{\Psi} = K(k_nh_n\Delta_n)^q$ and $\overline{\Psi} = K(k_nh_n\Delta_n))^{\overline{q}}$.

(5.20) and (5.23) yield $\theta_n\leq Kv'_nk_nh_n\Delta_n\big((k_nh_n\Delta_n)^{\overline{q}\wedge\frac{1+q}{2}}+\Delta_n^{q/2}\big)$ and $\theta'_n\leq Kv'^2_n(k_nh_n\Delta_n)^{2+q}$, hence (5.61) holds for $l=2$.

When $l=3$ we use (2.10) and $\mathbb{E}(\Phi^n_{i+1}\mid\mathcal{G}^n_i)=1$ and $\mathbb{E}((\Phi^n_{i+1})^2\mid\mathcal{G}^n_i)\leq K$ to get $\theta_n=0$ and $\theta'_n\leq Kv^2_nk_nh_n\Delta_n^2$, hence (5.61). □

LEMMA 24. *We have* $\overline{v}_nV^{n,y,l}_t\stackrel{\mathbb{P}}{\longrightarrow}0$ *for* $l=1,2,3,4$ *as soon as*

$$\begin{array}{l}\overline{v}_n\to 0,\qquad \frac{\overline{v}_n}{k_nh_n\Delta_n}\big(\alpha^3_n+\frac{\alpha^5_n}{k_n}+\widehat{\alpha}(2)_n+\frac{\alpha^1_n}{k_n}+\sqrt{\widehat{\alpha}(2)_n\overline{\alpha}(2)_n}\big)\to 0,\\ \frac{\overline{v}^2_n}{k_nh_n\Delta_n}\big(\frac{\overline{\alpha}(2)_n+\widehat{\alpha}(2)_n}{k^2_n}+\overline{\alpha}(4)_n+\widehat{\alpha}(4)_n\big)\to 0,\\ \frac{\overline{v}^2_n}{k_nh_n\Delta_n}\big(\frac{\kappa_1u^8_n(h_n\Delta_n)^4}{k^2_n}+\frac{u^8_n}{k^2_nh^4_n}+\frac{u^{2\beta_1}_n(h_n\Delta_n)^2}{k^2_n}\big)\to 0.\end{array}\tag{5.62}$$

Proof. 1) If $l=1$, on the set $\Omega_{n,t}$ we have $|\widehat{c}(y)^n_j|\leq K$ and thus $f(\widehat{c}(y)^n_j,\widehat{c}(2y)^n_j)\leq K$ for all $j\leq N'^n_t+1$, whereas (5.57) implies $\mathbb{E}(|\widehat{Y}^n_{N'^n_t}|)\leq K$ because N'^n_t is a stopping time for the filtration $(\mathcal{G}^n_{jw_n})_{j\geq 0}$, so $\mathbb{E}(|V^{n,y,2}_t|)\leq K$ and the claim follows from $\overline{v}_n\to 0$.

2) If $l=2$, and in view of (5.24) plus the convergences $\phi_n\to\phi$ and $\overline{\phi}_n\to\overline{\phi}$ and $\widetilde{\phi}^\beta_n\to\widetilde{\phi}^\beta$, it suffices to prove the following properties:

$$\begin{array}{ll}\frac{u^2_n\overline{v}_n}{k_n}\Big(2k_nh_n\Delta_n\sum_{j=0}^{N'^n_t}(c\lambda)_{T(n,jw_n)}-C_t\Big)\stackrel{\mathbb{P}}{\longrightarrow}0 & \text{when } \kappa_1=1,\\ \frac{u^{\beta_m}_n\overline{v}_n}{k_n}\Big(2k_nh_n\Delta_n\sum_{j=0}^{N'^n_t}(a^m\lambda)_{T(n,jw_n)}-A^m_t\Big)\stackrel{\mathbb{P}}{\longrightarrow}0 & \text{for } m=1,\dots,M,\\ \frac{u^2_n\overline{v}_n}{h_n}\sum_{j=1}^{N'^n_t}\big(2\gamma_{T(n,jw_n)}-\widehat{Y}^n_{jw_n}\big). & \end{array}$$

The first two properties follow from the previous lemma applied with $G=c\lambda$ (so $q=\overline{q}=1$ by (5.5)) and $G=a^m\lambda$ (so $\overline{q}=\beta_m/2$ and $q=1\wedge\beta_m$ by (5.14)), because $k^2_nh^2_n\Delta_n\to\infty$ and respectively $\overline{v}_n\to 0$ and $u^2_nh_n\delta_n\to 0$ when $\kappa_1=1$, and $\overline{v}_nu^{\beta_m}_n(k_nh_n\Delta_n)^{\beta_m/2}/k_n\to 0$ under (5.62). The last property follows from the same martingale argument as in the proof of Lemma 23, upon using Lemma 21 and (3.2) and $\overline{v}_n\to 0$.

3) In the cases $l=3,4$ we again use a martingale argument, with θ_n and θ'_n as in the proof of Lemma 23, relative to $\overline{v}_nV^{n,y,l}_t$ instead of $G^{n,l}_t$. When $l=3$, by (5.56) and $|V(y,y)^n_t|\leq K(\kappa_1u^4_n(h_n\Delta_n)^2+u^4_nh^{-2}_n+u^{\beta_1}_nh_n\Delta_n)$ we get

$$\begin{array}{l}\theta_n\leq K\overline{v}_n\big(\alpha^3_n+\frac{\alpha^5_n}{k_n}+\widehat{\alpha}(2)_n+\sqrt{\overline{\alpha}(2)_n\widehat{\alpha}(2)_n}\,\big),\\ \theta'_n\leq K\overline{v}^2_n\Big(\overline{\alpha}(4)_n+\widehat{\alpha}(4)_n+\frac{\kappa_1u^8_nh^4_n\Delta^4_n}{k^2_n}+\frac{u^8_n}{k^2_nh^4_n}+\frac{u^{2\beta_1}_nh^2_n\Delta^2_n}{k^2_n}\Big).\end{array}$$

When $l=4$ (5.56) gives us $\theta_n\leq K\overline{v}_n(\alpha^1_n+\alpha^3_n+\alpha^5_n)/k_n$ and $\theta'_n\leq K\overline{v}^2_n(\overline{\alpha}(2)_n+\widehat{\alpha}(2)_n)/k^2_n$. In both cases, (5.62) implies (5.61), and the proof is complete. □

LEMMA 25. *We have* $\overline{v}_nR^{n,y}_t\stackrel{\mathbb{P}}{\longrightarrow}0$ *if*

$$\frac{\overline{v}_n}{k_nh_n\Delta_n}\big(\frac{\overline{\alpha}(2)_n+\widehat{\alpha}(2)_n}{k_n}+\overline{\alpha}(3)_n+\widehat{\alpha}(3)_n\big)\to 0.\tag{5.63}$$

Proof. The claim readily follows from $N_t'^n \leq K_t/w_n\Delta_n$ and (5.56). □

We need another auxiliary result:

LEMMA 26. *Under* (5.62) *and* $\overline{v}_n\alpha_n^1/k_nh_n\Delta_n \to 0$, *for any continuous square-integrable* $(\mathcal{F}_t)$-*martingale* M *we have*

$$\overline{v}_n \sum_{j=0}^{N_t'^n} \mathbb{E}\big((M_{T(n,(j+1)w_n)} - M_{T(n,jw_n)})\,\xi(y)_j^n \mid \mathcal{G}_{jw_n}^n\big) \stackrel{\mathbb{P}}{\longrightarrow} 0. \tag{5.64}$$

Proof. In the whole proof t is fixed, and the (varying) constant K may depend on t. Recalling (5.2), we have $N_t'^n \leq m_n$, where $m_n = 1 + [(1+\Gamma t)/w_n\Delta_n]$, and we have $m_n \leq K/k_nh_n\Delta_n$. By a classical result, it suffices to prove the claim when the continuous martingale M is either orthogonal to W and bounded, or is W itself.

1) We begin with some preliminaries. First, we use the following notation:

$$\widehat{M}_j^n = M_{T(n,(j+1)w_n)} - M_{T(n,jw_n)}, \qquad \widehat{M}_i'^n = M_{T(n,i+2h_n)} - M_{T(n,i)}.$$

Using Doob's inequality and the properties of the approximate quadratic variation, plus the fact that M is a $(\mathcal{H}_t^n)$-martingale by Lemma 12, we see that

$$\begin{aligned}\mathbb{E}\Big(\sum_{j=0}^{m_n}(\widehat{M}_j^n)^2\Big) &= \mathbb{E}\Big(\sum_{i=0}^{(m_n+1)w_n-2h_n}(\widehat{M}_i'^n)^2\Big) = \mathbb{E}\Big(\sum_{i=1}^{(m_n+1)w_n}(\Delta_i^n M)^2\Big)\\ &= \mathbb{E}\big(M_{T(n,(m_n+1)w_n)} - M_0)^2\big) \leq K.\end{aligned} \tag{5.65}$$

(the last inequality is obvious when M is bounded; when $M = W$ it comes from $\mathbb{E}(T(n,(m_n+1)w_n)) \leq K(m_n+1)w_n\Delta_n \leq K$.)

With any arrays (η_i^n, z_i^n) or $(\eta_{i,l}^n, z_i^n)$ of variables, with $|z_i^n| \leq K$, we associate the variables

$$\begin{aligned}
S(\{\eta_i^n, z_i^n\})_n &= \overline{v}_n \sum_{j=0}^{N_t'^n} z_j^n\, \mathbb{E}(\eta_j^n \widehat{M}_j^n \mid \mathcal{G}_{jw_n}^n),\\
S'(\{\eta_i^n, z_i^n\})_n &= \tfrac{\overline{v}_n}{k_n} \sum_{j=0}^{N_t'^n}\sum_{l=0}^{k_n-1} z_{jw_n+2lh_n}^n\, \mathbb{E}(\eta_{jw_n+2lh_n}^n \widehat{M}_{jw_n+2lh_n}'^n \mid \mathcal{G}_{jw_n}^n),\\
S''(\{\eta_{i,l}^n, z_i^n\})_n &= \tfrac{\overline{v}_n}{k_n} \sum_{j=0}^{N_t'^n}\sum_{l=0}^{k_n-1} z_{jw_n+2lh_n+m}^n \sum_{m=1}^{2h_n} \mathbb{E}(\eta_{jw_n+2lh_n,m}^n\, \Delta_{jw_n+2lh_n+m}^n M \mid \mathcal{G}_{jw_n}^n),
\end{aligned}$$

and consider the properties

$$\begin{array}{lll}
A_2: & \mathbb{E}((\eta_i^n)^2 \mid \mathcal{G}_{jw_n}^n) \leq a_n, & \text{and } \overline{v}_n^2 a_n/(k_nh_n\Delta_n) \to 0,\\
A_1': & \big|\mathbb{E}(\eta_i^n\, \widehat{M}_i^n \mid \mathcal{G}_i^n)\big| \leq a_n, & \text{and } \overline{v}_n a_n/(k_nh_n\Delta_n) \to 0,\\
A_2': & \mathbb{E}((\eta_i^n)^2 \mid \mathcal{G}_i^n) \leq a_n, & \text{and } \overline{v}_n^2 a_n/(k_n^2h_n\Delta_n) \to 0,\\
A_1'': & \big|\mathbb{E}(\eta_i^n\, \Delta_i^n M \mid \mathcal{G}_{i-1}^n)\big| \leq a_n, & \text{and } \overline{v}_n a_n/(k_n\Delta_n) \to 0,\\
A_2'': & \mathbb{E}((\eta_i^n)^2 \mid \mathcal{G}_{i-1}^n) \leq a_n, & \text{and } \overline{v}_n^2 a_n/(k_n^2\Delta_n) \to 0.
\end{array}$$

The last part of this step is devoted to proving the following:

$$\begin{array}{lll}
A_2 & \Rightarrow & S(\{\eta_i^n, z_i^n\})_n \stackrel{\mathbb{P}}{\longrightarrow} 0,\\
A_1' \text{ or } A_2' & \Rightarrow & S'(\{\eta_i^n, z_i^n\})_n \stackrel{\mathbb{P}}{\longrightarrow} 0,\\
A_1'' \text{ or } A_2'' & \Rightarrow & S''(\{\eta_{i,l}^n, z_i^n\})_n \stackrel{\mathbb{P}}{\longrightarrow} 0.
\end{array} \tag{5.66}$$

The claims under A_1' and A_1'' are obvious because $N_t'^n \leq m_n$ and $|z_j^n| \leq K$. Assuming A_2' for example, (5.65) and the Cauchy-Schwarz inequality yields

$$\begin{aligned}\mathbb{E}(|S'(\{\eta_i^n, z_i^n\})_n|^2) &\leq K\,\frac{\overline{v}_n^2}{k_n^2}\,\mathbb{E}\Big(\Big(\sum_{j=0}^{m_n}\sum_{l=0}^{k_n-1}|\mathbb{E}(\eta^n_{jw_n+2lh_n}\widehat{M}'^n_{jw_n+2lh_n}\mid\mathcal{G}^n_{jw_n})|\Big)^2\Big)\\ &\leq K\,\frac{\overline{v}_n^2 m_n}{k_n}\,\mathbb{E}\Big(\sum_{j=0}^{m_n}\sum_{l=0}^{k_n-1}\mathbb{E}((\eta^n_{jw_n+2lh_n})^2\mid\mathcal{G}^n_{jw_n})\,\mathbb{E}((\widehat{M}'^n_{jw_n+2lh_n})^2\mid\mathcal{G}^n_{jw_n})\Big)\\ &\leq K\,\frac{\overline{v}_n^2 m_n a_n}{k_n}\,\mathbb{E}\Big(\sum_{j=0}^{m_n}\sum_{l=0}^{k_n-1}(\widehat{M}'^n_{jw_n+2lh_n})^2\Big)\;\leq\;K\,\frac{\overline{v}_n^2 m_n a_n}{k_n}\;\leq\;K\,\frac{\overline{v}_n^2 a_n}{k_n^2 h_n\Delta_n},\end{aligned}$$

implying $S'(\{\eta_i'^n, z_i^n\})_n \stackrel{\mathbb{P}}{\longrightarrow} 0$. The proof of (5.66) under A_2 or A_2'' is analogous, hence the claim.

2) By the definition of $\xi(y)_i^n$, the left side of (5.64) is $\sum_{w=1}^3 S(\{\xi(y)_i^{n,w}, z_i^n\})_n$, where $z_i^n = 1/U(y)^n_{T(n,iw_n)}$. Lemma 19 yields that the arrays $(\xi(y)_i^{n,w})$ satisfy A_2 with $a_n = \overline{\alpha}(2)_n$ when $w = 2,3$, hence (5.62) and (5.66) yield $S(\{\xi(y)_i^{n,w}, z_i^n\})_n \stackrel{\mathbb{P}}{\longrightarrow} 0$ in these cases. Now, $S(\{\xi(y)_i^{n,1}, z_i^n\})_n = S'(\{\cos(y\rho_i^n), z_i'^n\})_n + S'(\{U(y)^n_{T(n,i)}, z_i'^n\})_n$ if $z_i'^n = z_j^n$ when $jw_n \leq i < (j+1)w_n$, and since M is a martingale the array $(U(y)^n_{T(n,i)})$ satisfies A_1' with $a_n = 0$. Thus we are left to show that $S'(\{\cos(y\rho_i^n), z_i'^n\})_n \stackrel{\mathbb{P}}{\longrightarrow} 0$.

Toward this aim we set $\widehat{\rho}_i^n = \rho_i^{n,1} + \rho_i^{n,2}$. Expanding the cosine function around $\widehat{\rho}_i^n$, for any integer $P \geq 2$ we have

$$\cos(y\rho_i^n) = \sum_{w=0}^{P+1}\frac{1}{w!}\,\eta_i^{n,w}, \qquad \eta_i^{n,w} = \begin{cases}(-1)^{w/2}(y\rho_i^{n,3})^w\,\cos(y\widehat{\rho}_i^n) & \text{if } w \text{ is even and } 0\leq w\leq P,\\ (-1)^{(w+1)/2}(y\rho_i^{n,3})^w\,\sin(y\widehat{\rho}_i^n) & \text{if } w \text{ is odd and } 1\leq w\leq P,\\ \mathrm{O}\big(|\rho_i^{n,3}|^{P+1}\big) & \text{if } w = P+1.\end{cases}$$

Note that $\mathbb{E}((\eta_i'^{n,P+1})^2 \mid \mathcal{G}_i^n) \leq K u_n^{2P+2}/h_n^{P+1}$ by (5.31), so the array $(\eta_i^{n,P+1})$ satisfies A_2' for P large enough because $\overline{v}_n\alpha_n^1/k_n h_n\Delta_n \to 0$ is assumed, so $S'(\{\eta_i^{n,P+1}, z_i'^n\})_n \stackrel{\mathbb{P}}{\longrightarrow} 0$. Observe also that if $\eta_i'^{n,w} = \mathbb{E}(\eta_i^{n,w} \mid \mathcal{H}_\infty^n)$, we have $S'(\{\eta_i^{n,w}, z_i'^n\})_n = S'(\{\eta_i'^{n,w}, z_i'^n\})_n$ because M and $z_i'^n$ are $\mathcal{H}_\infty^n$-measurable, whereas with the notation (5.32) we have for $1 \leq w \leq P$:

$$w \text{ even } \Rightarrow\; \eta_i'^{n,w} = (-1)^{w/2}D_i^{n,w,y}\,\cos(y\widehat{\rho}_i^n), \qquad w \text{ odd } \Rightarrow\; \eta_i'^{n,w} = (-1)^{(w+1)/2}D_i^{n,w,y}\,\sin(y\widehat{\rho}_i^n).$$

(5.34) yields that $\eta_i'^{n,1} = 0$ and, when $w \geq 3$ is odd and upon combining with (5.30) and $u_n^2 \leq Kh_n$, the array $(\eta_i'^{n,w})$ satisfies A_2' with $a_n = Kh_n\Delta_n$. Thus for all w odd we have $S'(\{\eta_i^{n,w}, z_i'^n\})_n \stackrel{\mathbb{P}}{\longrightarrow} 0$.

Now, suppose that w is even. Since $|\overline{D}_i^{n,w,y}| \leq K$ by (5.32), we see that with the new array $z_i''^n = z_i'^n\overline{D}_i^{n,w,y}$ which is bounded again and $\mathcal{G}_i^n$-measurable, we have

$$S'(\{\eta_i^{n,w}, z_i'^n\})_n = (-1)^{w/2}\,S'(\{\eta_i^{n,0}, z_i''^n\})_n + S'(\{\eta_i''^{n,w}, z_i'^n\})_n, \qquad |\eta_i''^{n,w}| \leq |D_i^{n,w,y} - \overline{D}_i^{n,w,y}|.$$

(5.34) again yields that $\eta_i''^{n,w}$ satisfies A_2' with $a_n = Kh_n\Delta_n$, hence $S'(\{\eta_i''^{n,w}, z_i'^n\})_n \stackrel{\mathbb{P}}{\longrightarrow} 0$.

Summarizing the previous partial results, we deduce that we are left to prove that, for any variables z_i^n uniformly bounded and $\mathcal{G}_i^n$-measurable, we have

$$S'(\{\cos(y\widehat{\rho}_i^n), z_i^n\})_n \stackrel{\mathbb{P}}{\longrightarrow} 0. \tag{5.67}$$

3) For simplicity of notation we argue with complex valued variables below, and with the notation (5.43) we have $\widehat{\rho}_i^n = \widehat{\rho}^n_{i,0,2h_n}$. We have $\widehat{M}_i'^n = \sum_{l=1}^{2h_n}\Delta^n_{i+l}M$, so by successive conditioning (recall (5.9)

and the $(\mathcal{H}^n_t)$-martingale property of M) we get

$$S'(\{\cos(y\widehat{\rho}^n_i),z^n_i\})_n=\mathcal{R}e\big(S''(\{\eta^n_{i,l},z^n_i\})_n\big),\quad \eta^n_{i,l}=\mathbb{E}\big(e^{\iota y\widehat{\rho}^n_{i,0,2h_n}}\mid\mathcal{G}^n_{i+l}\big).$$

With the notation of Lemma 17, if $B_l=B(y)^n_{i,l}+\iota B'(y)^n_{i,l}$ and $\overline{B}_l=\overline{B}(y)^n_{i,l,2h_n}+\iota\overline{B}'(y)^n_{i,l,2h_n}$, (5.45) yields $\eta^n_{i,l}=\sum_{w=1}^3\eta^{n,w}_{i,l}$, where

$$\eta^{n,1}_{i,l}=e^{\iota y\widehat{\rho}^n_{i,0,l-1}-\lambda^n_{i+l-1}\overline{B}_l}e^{-\iota y\widehat{\rho}^n_{i,l-1,l}},\quad \eta^{n,2}_{i,l}=e^{\iota y\widehat{\rho}^n_{i,0,l}}\big(e^{-\lambda^n_{i+l}\overline{B}_l}-e^{-\lambda^n_{i+l-1}\overline{B}_l}\big),\quad |\eta^{n,3}_{i,l}|\le Ku^\beta_nh_n\Delta^2_n.$$

The array $(\eta^{n,w}_{i,l})$ satisfies A''_2 with $a_n=u^{2\beta}_nh^2_n\Delta^4_n=\mathrm{O}(\Delta^2_n)$ in the case $w=3$, and with $a_n=u^{2\beta}_nh^2_n\Delta^3_n=\mathrm{O}(\Delta_n)$ in the case $w=2$ (use $|\overline{B}_l|\le Ku^\beta_nh_n\Delta_n$ and the third part of (5.12) with $V=\lambda$), so $S''(\{\eta^{n,w}_{i,l},z^n_i\})_n\xrightarrow{\mathbb{P}}0$ in these cases, and it remains to prove $S''(\{\eta^{n,1}_{i,l},z^n_i\})_n\xrightarrow{\mathbb{P}}0$.

4) We first assume that M is orthogonal to W. Let $W'_t=W_{T(n,i+l-1)+t}-W_{T(n,i+l-1)}$ and $\underline{p}'$ be the restriction of $\underline{p}$ to $[T(n,i+l-1),\infty)\times E$, shifted in time by $-T(n,i+l-1)$: so W' and $\underline{p}'$ are still a Brownian motion and a Poisson measure with compensator $\underline{q}$, relative to the smallest filtration $(\mathcal{L}^n_t)$ to which they are adapted and with $\mathcal{L}^n_0=\mathcal{K}^n_\infty\bigvee\mathcal{G}^n_{i+l-1}$. Moreover, by Lemma 12 they have the same properties for the filtration $(\overline{\mathcal{L}}^n_t)$ which is the smallest one containing $(\mathcal{L}^n_t)$ and such that $\mathcal{K}^n_\infty\subset\overline{\mathcal{L}}^n_0$, and the process $M'_t=M_{T(n,i+l-1)+t}-M_{T(n,i+l-1)}$ is a continuous bounded $(\overline{\mathcal{L}}^n_t)$-martingale orthogonal to W.

Now, $\Delta(n,i+l)$ is $\mathcal{L}^n_0$-measurable and $\widehat{\rho}^n_{i,l-1,l}$ is $\mathcal{L}^n_{\Delta(n,i+l)}$-measurable. By the representation property for $(\mathcal{L}^n_t)$-martingales, the bounded variable $e^{-\iota y\widehat{\rho}^n_{i,l-1,l}}$ is the sum of an $\mathcal{L}^n_0$-measurable variable, plus two stochastic integral with respect to W' and $\underline{p}'-\underline{q}$, on the time interval $[0,\Delta(n,i+l)]$, thus $\mathbb{E}(e^{-\iota y\widehat{\rho}^n_{i,l-1,l}}M_{\Delta(n,i+l)}\mid\mathcal{L}^n_0)=0$. We deduce $S''(\{\eta^{n,1}_{i,l},z^n_i\})_n=0$.

When $\kappa_1=0$ we drop W' from the definition of $(\mathcal{L}^n_t)$. Then the variable $e^{-\iota y\widehat{\rho}^n_{i,l-1,l}}$ is the sum of an $\mathcal{L}^n_0$-measurable variable plus a single stochastic integral with respect to $\underline{p}'-\underline{q}$, so the orthogonality argument above applies also when $M=W$, and $S''(\{\eta^{n,1}_{i,l},z^n_i\})_n=0$ again in this case.

5) Finally, suppose $M=W$ and $\kappa_1=1$. Analogous with (5.46), we have

$$\mathbb{E}\big(e^{\iota y\widehat{\rho}^n_{i,l-1,l}}W'_{\Delta(n,i+l)}\mid\mathcal{L}^n_0\big)=e^{-\lambda^n_{i+l-1}\Phi^n_{i+l}\Delta_n(G^n_i(\overline{v}_{n,k})+\iota H^n_i(\overline{v}_{n,k}))}\,\mathbb{E}\big(e^{\iota yu_n\sigma^n_ig'^n_lW'_{\Delta(n,i+l)}}W'_{\Delta(n,i+l)}\mid\mathcal{L}^n_0\big),$$

and since $W'_{\Delta(n,i+l)}$ is $\mathcal{N}(0,\lambda^n_{i+j-1}\Phi^n_{i+l}\Delta_n)$, conditionally on $\mathcal{L}^n_0$, we deduce that, with $A^n_{i,l}=y\sigma^n_ig'^n_l$:

$$\mathbb{E}\big(e^{\iota y\widehat{\rho}^n_{i,l-1,l}}W'_{\Delta(n,i+l)}\mid\mathcal{L}^n_0\big)=\iota A^n_{i,l}u_n\Delta_n\lambda^n_{i+l-1}\Phi^n_{i+l}\,e^{-\lambda^n_{i+l-1}\Phi^n_{i+l}B_l},$$

which in turn gives us

$$\mathbb{E}\big(\eta^{n,1}_{i,l}\,\Delta^n_{i+l}W\mid\mathcal{L}^n_0\big)=\iota A^n_{i,l}u_n\Delta_n\lambda^n_{i+l-1}\Phi^n_{i+l}\,e^{\iota\widehat{\rho}^n_{i,0,l-1}-\lambda^n_{i+l-1}\Phi^n_{i+l}B_l-\lambda^n_{i+l-1}\overline{B}_l}.$$

Exactly as for (5.47), we then deduce (since here $\beta=2$)

$$\mathbb{E}\big(\eta^{n,1}_{i,l}\,\Delta^n_{i+l}W\mid\mathcal{G}^n_{i+l}\big)=\iota A^n_{i,l}u_n\Delta_n\lambda^n_{i+l-1}\,e^{\iota\widehat{\rho}^n_{i,0,l-1}-\lambda^n_{i+l-1}\overline{B}_{l-1}}+\mathrm{O}(u^5_n\Delta^3_n).$$

Then, using the second part of (5.12) for $V=\lambda$ plus $|\overline{L}_{l-1}|\le K$, we get

$$\mathbb{E}\Big(\big|\,\mathbb{E}\big(\eta^{n,1}_{i,l}\,\Delta^n_{i+l}W\mid\mathcal{G}^n_{i+l-1}\big)-\iota A^n_{i,l}u_n\Delta_n\lambda^n_i\,e^{\iota\widehat{\rho}^n_{i,0,l-1}-\lambda^n_i\overline{B}_{l-1}}\big|\mid\mathcal{G}^n_i\Big)\le K(u^5_n\Delta^3_n+u_nh_n\Delta^2_n),$$

and thus (5.45) yields, because $u_n^5\Delta_n^3 + u_n^3 h_n\Delta_n^3 \leq u_n h_n\Delta_n^2$,

$$\mathbb{E}\big(\eta_{i,l}^{n,1}\,\Delta_{i+l}^n W \mid \mathcal{G}_i^n\big) = \iota A_{i,l}^n u_n\Delta_n\lambda_i^n\, e^{-\lambda_i^n\overline{B}_0} + \mathrm{O}(u_n h_n\Delta_n^2).$$

Now, recalling (5.45), the real part of the above is

$$\mathbb{E}\big(\mathcal{R}e(\eta_{i,l}^{n,1}\,\Delta_{i+l}^n W) \mid \mathcal{G}_i^n\big) = A_{i,l}^n u_n\Delta_n\lambda_i^n\,\sin(\lambda_i^n B'(y)_{i,0,2h_n}^n) + \mathrm{O}(u_n h_n\Delta_n^2) = \mathrm{O}(u_n^{1+\beta_1}h_n\Delta_n^2).$$

In other words, the array $(\mathcal{R}e(\eta_{i,l}^{n,1}))$ satisfies A_1'' with $a_n = K u_n^{1+\beta_1}h_n\Delta_n^2$. Since $\overline{v}_n\alpha_n^1/k_n h_n\Delta_n \to 0$, we have in particular $\overline{v}_n u_n^{1+\widetilde{r}'/2}(h_n\Delta_n)^{1/2}/k_n \to 0$ if $\kappa_1 = 1$, which implies $\overline{v}_n u_n^{1+\beta_1}h_n\Delta_n/k_n \to 0$ by (3.2) and $\widetilde{r}' \geq \beta_1$: therefore $S''(\{\eta_{i,l}^{n,1}, z_i^n\})_n \xrightarrow{\mathbb{P}} 0$, and the proof is complete. □

5.7. *Proof of the main results.* At this stage, and under (5.60), (5.62) and (5.63), we are left to proving (ii)–(iv) of (5.59), according to the case. Toward this aim, we introduce a series of conditions:

$$\overline{v}_n\sum_{j=0}^{N_t'^n}\big|\,\mathbb{E}(\xi(y)_j^n \mid \mathcal{G}_{jw_n}^n)\big| \xrightarrow{\mathbb{P}} 0, \tag{5.68}$$

$$\overline{v}_n^3\sum_{j=0}^{N_t'^n}\mathbb{E}(|\xi(y)_j^n|^3 \mid \mathcal{G}_{jw_n}^n) \xrightarrow{\mathbb{P}} 0, \tag{5.69}$$

$$\overline{v}_n^2\sum_{j=0}^{N_t'^n}\mathbb{E}(\xi(y)_j^n\,\xi(y')_j^n \mid \mathcal{G}_{jw_n}^n) \xrightarrow{\mathbb{P}} \Gamma(y,y')_t, \tag{5.70}$$

where $\Gamma(y,y')$ will be defined later, depending on the case at hand. Let us mention the following facts, based on (5.56) plus the consequence $N_t'^n \leq \frac{K(1+t)}{k_n h_n\Delta_n}$ of (5.2):

1. We have (5.68) as soon as

$$\frac{\overline{v}_n}{k_n h_n\Delta_n}\,(\alpha_n^1 + \alpha_n^3 + \alpha_n^5) \to 0. \tag{5.71}$$

2. We have (5.69) as soon as $\frac{\overline{v}_n^3}{h_n k_n\Delta_n}\,(\overline{\alpha}(3)_n + \widehat{\alpha}(3)_n) \to 0$, hence under (5.63) and $\overline{v}_n \to 0$.
3. Under (5.62), (5.70) amounts to having

$$\frac{\overline{v}_n^2}{k_n}\sum_{j=0}^{N_t'^n}V(y,y)_{jw_n}^n \xrightarrow{\mathbb{P}} \Gamma(y,y')_t. \tag{5.72}$$

A (very) tedious but elementary calculation shows us that, under (3.2), the set of conditions (5.60), (5.62), (5.63), (5.71) is indeed equivalent to (4.2), and thus by the previous lemmas

$$(4.2) \quad\Longrightarrow\quad (5.59)\text{-(i,ii)},\ (5.64),\ (5.68),\ (5.69). \tag{5.73}$$

In connection with (4.3), we also set

$$\eta_n = \frac{u_n^{\beta_1} h_n^3 \Delta_n}{u_n^{\beta_1} h_n^3 \Delta_n + u_n^4 (1 + \kappa_1 h_n^2 \Delta_n)^2}, \qquad \eta'_n = \frac{h_n^2 \Delta_n}{1 + h_n^2 \Delta_n}.$$

Proof of Theorem 1. We fix $t > 0$ and we need to prove (5.59)-(ii) with $\overline{v}_n = v_n$ satisfying (4.1). By a classical convergence result for sums of triangular arrays, since $N_t'^n$ is a stopping time for the discrete-time filtration $(\mathcal{G}^n_{jw_n})_{j\geq 0}$ and each $\xi(y)^n_j$ is $\mathcal{G}^n_{(j+1)w_n}$-measurable, the three conditions (5.68), (5.69) and (5.70) for $y' = y$ and $\Gamma(y,y)_t = 0$ imply $\overline{v}_n V_t^{n,y} \xrightarrow{\mathbb{P}} 0$. By (5.27) and $N_t'^n \leq \frac{K(1+t)}{k_n h_n \Delta_n}$ again, the third condition is implied by

$$\zeta_n := \frac{\overline{v}_n^2}{k_n^2} \Big(u_n^{\beta_1} + \frac{u_n^4}{h_n^3 \Delta_n} + \kappa_1 u_n^4 h_n \Delta_n \Big) \ \to \ 0.$$

The above plus (5.73) are indeed equivalent to (4.1) under (3.2), hence the result. □

Proof of Theorem 2. We have $\kappa_1 = 0$ here and $v_n^2/k_n^2 = h_n^3 \Delta_n/(u_n^4 + u_n^{\beta_1} h_n^3 \Delta_n) = \eta_n/u_n^{-\beta_1}$. If $\overline{v}_n = v_n$, a simple consequence of the estimate (5.27) is

$$\frac{\overline{v}_n^2}{k_n} V(y,y')^n_t = w_n \Delta_n \big(\eta_n \psi_{\beta_1}(y,y') \widetilde{\phi}_n^{(\beta_1)} \chi(\beta_1)\, a^1_t \lambda_t + (1-\eta_n) y^2 y'^2 \overline{\phi}_n\, \gamma_t^2 \big) + \mathrm{o}(w_n \Delta_n).$$

Lemma 23 applied with $\overline{v}'_n = 1$ and with G equal to a^1 or γ^2/λ or $1/\lambda$ (which all satisfy the assumptions of that lemma) plus (3.3) and $\eta_n \to \eta$ by (4.3) and (5.56) yield (5.72) with $\Gamma(y,y')_t$ given by the right-hand side of (4.5), for all t.

Now, by another classical result for the $\mathcal{F}_\infty$-stable convergence in law of triangular arrays, the property (iii) of (5.59) with the limit $(Z(y)_t)_{y\in\mathcal{Y}}$ holds as soon as, for all $t > 0$, we have (5.68), (5.69), (5.70), and also (5.64) for any continuous square integrable martingale M. Then the claim follows from (5.73). □

Proof of Theorem 3. We have $\kappa_1 = 1$ here and v_n and v'_n are given by (4.6).

a) Note that $v_n^2/k_n^2 = \eta_n/u_n^{\beta_1}$, so if (5.27) yields

$$\frac{v_n^2}{k_n} V(y,y')^n_t = w_n \Delta_n \big(\eta_n \psi_{\beta_1}(y,y') \widetilde{\phi}_n^{(\beta_1)} \chi(\beta_1)\, a^1_t + (1-\eta_n) y^2 y'^2 \big(\eta'_n \phi_n c_t \lambda_t + (1-\eta'_n) \overline{\phi}_n \gamma_t \big)^2 \big) + \mathrm{o}(w_n \Delta_n). \tag{5.74}$$

Then the results follows exactly as for the previous theorem, upon using $\overline{v}_n = v_n$.

b) We have $v_n'^2/k_n^2 = 1/u_n^{\beta_1}$ and we now set $\overline{v}_n = v'_n$. The property (iv) of (5.59) with the limit $(Z(1)_t, (Z'(y)_t)_{y\in\mathcal{Y}})$ holds as soon as we have (5.64), (5.68) and (5.69) for all t and all M, plus the following (again for all t) instead of (5.70):

$$\begin{aligned}
& v_n^2 \textstyle\sum_{j=0}^{N_t'^n} \mathbb{E}\big((\xi(1)^n_j)^2 \mid \mathcal{G}^n_{jw_n} \big) \xrightarrow{\mathbb{P}} \Gamma(1,1)_t, \\
& v_n'^2 \textstyle\sum_{j=0}^{N_t'^n} \mathbb{E}\big((\xi(y)^n_j - y^2 \xi(1)^n_i)(\xi(y')^n_j - y'^2 \xi(1)^n_i) \mid \mathcal{G}^n_{jw_n} \big) \xrightarrow{\mathbb{P}} \overline{\Gamma}(y,y')_t, \\
& v_n v'_n \textstyle\sum_{j=0}^{N_t'^n} \mathbb{E}\big((\xi(y)^n_j - y^2 \xi(1)^n_j) \xi(1)^n_j \mid \mathcal{G}^n_{jw_n} \big) \xrightarrow{\mathbb{P}} 0,
\end{aligned} \tag{5.75}$$

where $\Gamma(y,y')_t$ and $\overline{\Gamma}(y,y')_t$ are the right-hand sides in (4.8) and (4.10). Now, we have (5.74), and (5.27) also yields

$$\frac{v_n'^2}{k_n}\,\overline{V}(y,y')^n_t = w_n\Delta_n\overline{\psi}_{\beta_1}(y,y')\widetilde{\phi}_n^{(\beta_1)}\chi(\beta_1)\,a^1_t + \mathrm{o}(w_n\Delta_n),$$
$$\frac{v_nv_n'}{k_n}\,\widetilde{V}(y)^n_t = \mathrm{o}(w_n\Delta_n),$$

At this stage, again the same argument as previously gives us the result. □

Acknowledgements. Research partially supported by NSF Grant SES-1530748. We would like to thank the editor, an associate editor and an anonymous referee for many constructive comments and suggestions.

References.

[1] Ait-Sahalia, Y. and J. Jacod (2009). Estimating the Degree of Activity in High Frequency Data *Annals of Statistics*. **37**, 2202–2244.

[2] Ait-Sahalia, Y. and J. Jacod (2014). *High-Frequency Financial Econometrics*. Princeton University Press.

[3] Barndorff-Nielsen, O.E. and N. Shephard (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics* **2**, 1–48.

[4] Barndorff-Nielsen, O.E., N. Shephard and M. Winkel (2006). Limit Theorems for Multipower Variation in the Presence of Jumps. *Stochastic Processes and Their Applications* **116**, 796–806.

[5] Barndorff-Nielsen, O. E., P.R. Hansen, A. Lunde and N. Shephard. (2008). Designing Realised Kernels to Measure the Ex-post Variation of Equity Prices in the Presence of Noise. *Econometrica* **76**, 1481–1536.

[6] Bibinger, M., N. Hautsch, P. Malec and M. Reiss. (2014). Estimating the Quadratic Covariation Matrix from Noisy Observations: Local method of Moments and Efficiency. *Annals of Statistics* **42**, 1312–1346.

[7] Bull, A. (2016). Near-optimal Estimation of Jump Activity in Semimartingales. *Annals of Statistics* **44**, 58–86.

[8] Comte, F. and E. Renault. (1996). Long Memory Continuous Time Models. *Journal of Econometrics* **73**, 101–149.

[9] Gatheral, J., T. Jaisson and M. Rosenbaum. (2016). Volatility is Rough. *Preprint*.

[10] Jacod, J., Y. Li, P. Mykland, M. Podolskij and M. Vetter. (2009). Microstructure Noise in the Continuous Case: The Pre-averaging Approach. *Stochastic Processes and their Applications* **119**, 2249–2276.

[11] Jacod, J. and P. Protter (2012). *Discretization of Processes*, Springer-Verlag, Berlin.

[12] Jacod, J. and V. Todorov (2014). Efficient Estimation of Integrated Volatility in Presence of Infinite Variation Jumps. *Annals of Statistics* **42**, 1029-1069.

[13] Jacod, J. and V. Todorov (2015). Efficient Estimation of Integrated Volatility in Presence of Infinite Variation Jumps with Multiple Activity Indices. In *The fascination of Probability, Statistics and Their Applications – In honour of Ole E. Barndorff-Nielsen on his 80th birthday*, Springer-Verlag.

[14] Jing, B., X. Kong and Z. Liu. (2011). Estimating the Jump Activity Index under Noisy Observations using High-Frequency Data. *Journal of the American Statistical Association* **106**, 558–568.

[15] Jing, B., X. Kong, Z. Liu and P. Mykland. (2012). On the Jump Activity Index for Semimartingales. *Journal of Econometrics* **166**, 213–223.

[16] Kong, X., Z. Liu and B. Jing. (2015). Testing for Pure-Jump Processes for High-Frequency Data. *Annals of Statistics* **43**, 847–877.

[17] Koike, Y. (2016). Estimation of Integrated Covariances in the Simultaneous Presence of Nonsynchronicity, Microstructure Noise and Jumps. *Econometric Theory* **32**, 533–611.

[18] Mancini, C. (2001). Disentangling the Jumps of the Diffusion in a Geometric Brownian Motion. *Giornale dell'Istituto Italiano degli Attuari* **LXIV**, 19–47.

[19] Podolskij, M. and M. Vetter. (2009). Estimation of Volatility Functionals in the Simultaneous Presence of Microstructure Noise and Jumps. *Bernoulli* **15**, 634-658.

[20] Todorov, V. (2015). Jump Activity Estimation for Pure-Jump Semimartingales via Self-Normalized Statistics. *Annals of Statistics* **43**, 1831–1864.

[21] Todorov, V. and G. Tauchen (2011). Limit Theorems for Power Variations of Pure-Jump Processes with Application to Activity Estimation. *Annals of Applied Probability* **21**, 546–588.

[22] Todorov, V. and G. Tauchen (2012). The Realized Laplace Transform of Volatility. *Econometrica* **80**, 1105–1127.

[23] Woerner, J. (2003). Variational Sums and Power Variation: A Unifying Approach to Model Selection and Estimation in Semimartingale Models. *Statistics and Decisions* **21**, 47–68.

[24] Zhang, L. (2006). Efficient Estimation of Stochastic Volatility using Noisy Observations: A Multi-scale Approach. *Bernoulli* **12**, 1019–1043.

Institut de Mathématiques de Jussieu
CNRS – UMR 7586
Université Pierre et Marie Curie–P6
4 Place Jussieu, 75252 Paris-Cedex 05
E-mail: jean.jacod@gmail.com

Department of Finance
Northwestern University
Evanston, IL 60208-2001
E-mail: v-todorov@northwestern.edu