HOUSE OF FINANCE

WORKING PAPER SERIES

Aleksey Kolokolov – Giulia Livieri – Davide Pirino

Statistical Inferences for Price Staleness

SAFE Working Paper No. 236

Statistical inferences for price staleness*

Aleksey Kolokolov[†] Giulia Livieri[‡] Davide Pirino[§]

November 6, 2018

Abstract

Asset transaction prices sampled at high frequency are much staler than one might expect in the sense that they frequently lack new updates showing zero returns. In this paper, we propose a theoretical framework for formalizing this phenomenon. It hinges on the existence of a latent continuous-time stochastic process p_t valued in the open interval $(0, 1)$, which represents at any point in time the probability of the occurrence of a zero return. Using a standard infill asymptotic design, we develop an inferential theory for non-parametrically testing, the null hypothesis that p_t is constant over one day. Under the alternative, which encompasses a semimartingale model for p_t, we develop non-parametric inferential theory for the probability of staleness that includes the estimation of various integrated functionals of p_t and its quadratic variation. Using a large dataset of stocks, we provide empirical evidence that the null of the constant probability of staleness is fairly rejected. We then show that the variability of p_t is mainly driven by transaction volume and is almost unaffected by bid-ask spread and realized volatility.

Keywords: staleness, idle time, liquidity, zero returns, stable convergence

*Davide Pirino acknowledges partial support via the RBSI14DDNN, "*A new measure of liquidity,*" financed within the program "Scientific Independence of Young Researchers" of the Italian Ministry of Education and Research. Aleksey Kolokolov acknowledges the support in the framework of the Trans-Atlantis Platform from the DFG under PE 2574/1-1 and the research support from the Research Center SAFE, funded by the State of Hessen initiative for research LOEWE. All errors are our own.

[†]Alliance Manchester Business School, The University of Manchester. E-mail: alexeiuo@gmail.com.

[‡]Scuola Normale Superiore, Pisa. E-mail: giulia.livieri@sns.it

[§]Università degli Studi di Roma "Tor Vergata", Dipartimento di Economia e Finanza, Via Columbia 2, 00173, Roma and Scuola Normale Superiore, Piazza dei Cavalieri 7, 56123, Pisa, Italy. E-mail: davide.pirino@gmail.com.

1 Introduction

Traditional models in continuous-time finance entail that the price of an asset traded in a frictionless market evolves as a semimartingale. Bandi et al. (2017) provide empirical evidence against this hypothesis showing that even at moderately high frequency asset prices do not update as frequently as expected under the semimartingale assumption. Indeed, while under the standard semimartingale hypothesis high-frequency returns should exceed an appropriately defined threshold with large probability, often the converse is true; asset prices are stale in the sense that they show a large incidence of zero or, more generally, "small" returns. The inclusion of price staleness in the data-generating process results is pivotal from both an economic and an econometric point of view. Bandi et al. (2017) provide a micro-structural model of price formation (following the spirit of Kyle, 1985; Hasbrouck and Ho, 1987; Glosten and Milgrom, 1985) where the emergence of zero returns is determined by the joint effect of asymmetric information, transaction costs, and delays in the incorporation of the information flow into the assets' prices. However, being agnostic about the sources of zero returns, Kolokolov and Renò (2017) show that neglecting price staleness leads to severe distortions of the widely used power and multi-power estimators (Woerner, 2006; Barndorff-Nielsen et al., 2006; Barndorff-Nielsen and Shephard, 2004; Lee and Mykland, 2008; Caporin et al., 2014), which results in distorting traditional jump tests toward false jump detection. Even though one may claim that such sluggish dynamics are the spurious consequence of price discreteness, the empirical analysis in Bandi et al. (2018) shows that this argument is falsified by data. On a large dataset of New York Stock Exchange (NYSE)-listed stocks, they document that high-frequency transaction prices show an excess of zero returns with respect to what would be expected from price rounding alone. Most importantly, they prove that this excess of staleness, being strictly related to transaction volumes, bid-ask spreads, and volatility, brings insightful economic information.

Therefore, the occurrence of zero returns is an economically meaningful feature of the data-generating process of financial asset prices, which deserves detailed investigation. As the past financial econometric literature has successfully investigated stochastic volatility (see, among many others, Hull and White, 1987; Scott, 1987; Heston, 1993; Bates, 1996) by focusing on the erratic behavior of price paths, here we look at the other side of the coin and ask the following research questions: Does the probability of the occurrence of zero returns vary empirically on an intradaily basis? In the affirmative case, what is an appropriate model for such variability and what are the economic variates that mostly determine it?

As a starting point, we assume the existence of an efficient price process Y, which we define as the asset price that would have been observed if the market was perfectly liquid. In the presence of illiquidity frictions (such as trading costs or asymmetric information), the trading activity is inhibited. The result is the random occurrence of periods in which the observed price process stays constant, a situation that in this paper we call

price staleness, following the nomenclature of Bandi et al. (2018). The greater the magnitude of these frictions, the more probable and the more persistent the staleness of the observed price. We model this frictional price dynamic following the formalism introduced by Bandi et al. (2017) and Bandi et al. (2018). In contrast to the standard assumption, for any partition $0 = t_{0,n} < t_{1,n} < \ldots < t_{n,n} = 1$ of the interval $[0, 1]$ (e.g., a trading day), the efficient price process Y is not observed in every point of the grid. Instead, the data are assumed to be generated by the following recursive equation[1]:

$$X^{(n)}_{t_{j,n}} = Y_{t_{j,n}} (1 - \mathbb{B}_{j,n}) + X^{(n)}_{t_{j-1,n}} \mathbb{B}_{j,n}, \quad j = 1, \ldots, n, \tag{1}$$

with the initial condition $X^{(n)}_0 = Y^{(n)}_0$, where $Y_{t_{j,n}}$ is the efficient price sampled in the j-th element of the partition and where $(\mathbb{B}_{j,n})_{j=1,\ldots,n}$ is a triangular array of Bernoulli random variates such that for some (random) $p_\infty \in (0, 1)$.

$$\frac{1}{n} \sum_{j=1}^{n} \mathbb{B}_{j,n} \xrightarrow{p} p_\infty, \quad \text{as} \quad n \to \infty.$$

The recursive equation (1) implies that at each instant $t_{j,n}$ the observed price $X^{(n)}_{j\Delta_n}$ may either coincide with the latent efficient price ($\mathbb{B}_{j,n} = 0$) or may not update and stay constant ($\mathbb{B}_{j,n} = 1$), thus leading to a stale price. Our theoretical framework hinges on the key assumption that this latter event occurs with probability $\mathbb{E}\left[p_{t_{j,n}}\right]$, where $(p_t)_{t\in[0,1]}$ is a latent stochastic process valued in $(0, 1)$. In this paper, we develop an inferential theory for the dynamics of p_t and consequently for the intraday dynamics of price staleness.

Our first result is to show that the intraday fraction of zeros, dubbed *idle time*[2] in Bandi et al. (2017), is a consistent estimator of the integrated probability of

[1] The superscript (n) in the notation $X^{(n)}_{t_{j,n}}$ intends to highlight that the observed price process is frequency-dependent. This means that even if $t_{j,n} = t_{j,n'} = t$ for $n \neq n'$ it typically occurs that:

$$X^{(n)}_t \neq X^{(n')}_t.$$

[2] Bandi et al. (2017) define *idle time* as the daily percentage of log-returns that in absolute value are smaller than a (asymptotically vanishing) threshold. Here, with a slight abuse of nomenclature, we call *idle time* the same percentage with the threshold set exactly to zero.

price staleness. Then, under the assumption that the process $(p_t)_{t\in[0,1]}$ evolves as a Brownian semimartingale, we derive a (stable) central limit theorem (CLT) for *idle time*. To set-up a feasible confidence interval, we introduce a new economic indicator, called *(m)-idle time*, and we derive its limiting properties when $n \to \infty$. Next, we introduce the notion of *local idle time* and prove that 1) it is a local estimator of p_t, the stochastic probability of price staleness; 2) it allows the construction of a non-parametric test able to distinguish between a constant and a time-varying p_t and 3) under the null of a Brownian semimartingale for p_t, it allows defining a consistent estimator of its integrated volatility.

We conclude the paper with an empirical application of the asymptotic theory that sheds light new features of the price formation mechanisms and answer our research questions on the dynamics of price staleness. First, using a large dataset of NYSE-listed stocks, we prove that the non-parametric test largely rejects the null of a constant probability of staleness. This means that zero returns are typically not uniformly distributed during the day, and therefore periods of no trading activity tend to cluster. Second, on the same dataset, we derive estimates of the integrated volatility of p_t for the sample stocks and with simple regressions prove how this newly defined realized measure conditionally driven on the average p_t by transaction volumes and is almost unaffected by bid-ask spreads and price volatility.

The remainder of the paper is organized as follows. Section 2 introduces the setting. Section 3 contains the limit results. Section 4 shows the finite sample accuracy of our asymptotic theory using a Monte Carlo exercise and Section 5 presents the empirical results. Section 6 consists of the conclusion. All technical proofs are confined to the Appendix.

2 The model

We work on a filtered probability space $\left(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P}\right)$ that supports all stochastic elements defined below. The structure of the filtration $(\mathcal{F}_t)_{t\geq 0}$ is quite technical and is reported in Appendix A.1. We consider refining partitions of the time interval $[0,1]$, $\Pi_n = \{t_{0,n}, \ldots, t_{n,n}\}$, with $0 = t_{0,n} < t_{1,n} < \ldots < t_{n,n} = 1$, such that $\Pi_n \subseteq \Pi_{n+1}$[3] where n is an increasing subsequence of $\mathbb{N}$, such that the partitions are equispaced. The value of a generic stochastic process X at a point $t_{j,n}$ of a partition Π_n are denoted with $X_{t_{j,n}}$ or, to avoid excessive subscripts, simply with $X_{j,n}$. Because the partitions are equispaced, we have $t_{j,n} = j/n$ for $j = 0, ..., n$,

[3]The requirement that $\Pi_n \subseteq \Pi_{n+1}$ allows us to significantly reduce the proofs, and it is natural for financial applications. For instance, one-minute partitioning of a trading day contains five-minute partitioning.

and we indicate the distance between two consecutive points of the partition with $\Delta_n = 1/n$.

As anticipated in the introduction, the key assumption of our theoretical framework consists in the existence of a latent stochastic process that plays the role of an "instantaneous probability of staleness". What follows formalizes this idea.

Assumption 1. *There exists an adapted Riemann-integrable continuous-time stochastic process* $(p_t)_{t\in[0,1]}$*, taking the values in* $(0,1)$*, such that the triangular array* $(\mathbb{B}_{j,n})_{j=1,\dots,n}$ *of Bernoulli random variables that appears in equation* (1) *is given by:*

$$\mathbb{B}_{j,n} \doteq \mathbb{I}_{\left\{U_{t_{j,n}} \leq p_{t_{j,n}}\right\}}, \quad j = 0, \dots, n, \tag{2}$$

where $\mathbb{I}_{\{\cdot\}}$ *is the indicator function and where* $(U_t)_{t\in[0,1]}$ *is a collection of uniformly distributed random variables (independent of* p_t*) satisfying* $U_t \perp U_{t'}$, $\forall t \neq t'$, *and* $U_t \in \mathcal{F}_t$ *for all* t*.*

Intuitively, the process p_t determines at any point in time the probability of occurrence of a zero return in the sense that $\mathbb{P}\left[\mathbb{B}_{j,n} = 1\right] = \mathbb{E}\left[p_{j,n}\right]$, that is $(p_t)_{t\geq 0}$ plays the role of an instantaneous stochastic probability of price staleness in continuous-time.

Notice that Assumption 1 preserves the *compatibility relationship* (cfr. Aït-Sahalia and Jacod, 2014, p. 211) over different sampling frequencies. Formally, this property guarantees that if $t_{j,n} = j/n$ and $t_{j',n} = j'/n$ are two equally spaced partitions of $[0,1]$, with $j = 1, \dots, n$ and $j' = 1, \dots, n'$, then $\mathbb{B}_{j,n} = \mathbb{B}_{j',n'}$ whenever $j/n = j'/n'$.

Assumption 1 encompasses different specifications of $(\mathbb{B}_{j,n})_{j=1,\dots,n}$. If $p_t = p^F$ $\forall t \in [0,1]$, then the Bernoulli variates are *i.i.d.* with the probability of staleness given by p^F. Another (more sophisticated) specification is obtained when $(p_t)_{t\in[0,1]}$ is described by a Brownian semimartingale. As an illustrative example, Figure 1 plots a simulated path of the observed price process in equation (1) in which p_t is either constant (right panel) or a Brownian semimartingale (left panel). Although the number of zero returns (signalled by a red cross) is the same, the two graphs look rather different. In the *i.i.d* scenario, stale prices are uniformly distributed over the trading day. However, in the semimartingale case, there is clustering of lack of price adjustments.

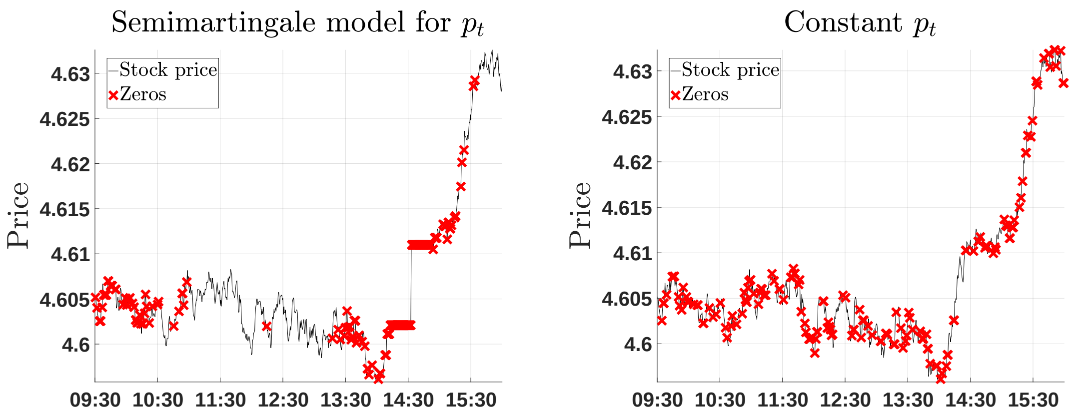

Figure 1: We give an example of a stale stock price where zero returns are signaled by a red cross. The probability of observing a zero return either follows a semimartingale dynamic (left panel) or it is equal to a constant (right panel). In both cases, the number of zeros is the same.

3 Asymptotic results

Here, we derive a (stable) CLT for the estimator of the fraction of zero returns within one day. For a feasible CLT, we need first to add a technical assumption regarding the dynamics of p_t.

Assumption 2. *The process $(p_t)_{t\in[0,1]}$ is described by the following stochastic differential equation (SDE):*

$$p_t = p_0 + \int_0^t \mu_s \, ds + \int_0^t \nu_s \, dW_s, \tag{3}$$

where W_t is a standard Brownian motion, and μ_t and ν_t are adapted cádlág processes, such that $\forall t$, $p_t \in (0,1)$ almost surely.

Second, we need a consistent estimator of functionals of the form $\int_0^1 p_s^m \, ds$, where $m \geq 1$ is an integer number. For this reason, below we develop a theory of estimation of all integrals of the type $\int_0^1 f(p_s) \, ds$ for a smooth enough test function $f(\cdot)$. We then derive a non-parametric test designed to asymptotically discriminate between the null of a time-independent p_t and an alternative in which the probability of staleness varies during the day. Finally, under the (alternative) hypothesis that p_t is driven by a semimartingale (i.e., under Assumption 2), we derive a consistent estimator of its integrated volatility.

3.1 Idle and multi-idle time

We borrow the definition[4] of *idle time* from Bandi et al. (2017)

$$\mathrm{IT}_n \doteq \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{\left\{X^{(n)}_{j,n} - X^{(n)}_{j-1,n} = 0\right\}}, \tag{4}$$

where $X^{(n)}$ is the observed price process defined in equation (1). The idle time, computed over a trading day, yields the fraction of a day for which the price adjustments are zero. Despite its simplicity, idle time encompasses economically meaningful features of the data-generating process of financial asset prices. However, we are not going to discuss this point further and we refer readers to the paper by Bandi et al. (2018) and the references therein for additional discussions. We focus instead on the limiting properties of IT_n, which are exposed in the following theorem.

Theorem 3.1. *Under Assumption 1, as $n \to \infty$, we have that:*

$$\mathrm{IT}_n \xrightarrow{u.c.p} \int_0^1 p_s \, ds.$$

In addition, if both Assumptions 1 and 2 hold, as $n \to \infty$,

$$\sqrt{n}\left(\mathrm{IT}_n - \int_0^1 p_s \, ds\right) \xrightarrow{stably} \mathcal{MN}\left(0, \Sigma_{IT}\right), \tag{5}$$

where $\mathcal{MN}(0, \Sigma_{IT})$ denotes the mixed-normal distribution with a stochastic variance Σ_{IT} and where Σ_{IT} is defined as:

$$\Sigma_{IT} = \int_0^1 p_s \,(1 - p_s)\, ds. \tag{6}$$

Proof. See Appendix A.2. □

The $u.c.p$ limit implies that IT_n is a consistent estimator of the integrated probability of price staleness over one trading day, and this result holds under very general assumptions on the dynamics of the process p_t.
Under Assumption 2, the difference $\mathrm{IT}_n - \int_0^1 p_s \, ds$ converges stably at rate $n^{1/2}$ to a zero-mean (mixed) normal distribution whose variability has an intuitive expression. Indeed, in the case of constant probability of price staleness, for example $p_t = p_0 \, \forall \, t \in [0, 1]$, the asymptotic variance coincides (given the independence of

[4]There is a subtle difference between the definition in equation (4) and the idle time as introduced by Bandi et al. (2017), as in this latter case idle time indicates the percentage of log returns that in absolute value are below an asymptotically vanishing threshold ξ_n. We set $\xi_n = 0$ because in our theoretical framework the introduction of a threshold is unnecessary.

the driving Bernoulli variates) with the variance of a Bernoulli random variable with mean p_0, that is $p_0\,(1-p_0)$. In the non-constant case, the expression of the asymptotic variance naturally generalizes to its integral version.

Because IT_n is a consistent estimator of $\int_0^1 p_s\,ds$, a feasible confidence interval for IT_n can be readily defined once a consistent estimator of $\int_0^1 p_s^2\,ds$ is available. Actually, we consider the more general problem of developing a consistent estimator of $\int_0^1 (p_s)^m\,ds$ for some integers $m \geq 2$. For this purpose, we introduce the (m)-*idle time* defined as:

$$\mathrm{IT}_n^{(m)} \stackrel{\text{def}}{=} \frac{1}{n-m}\sum_{j=1}^{n-m}\prod_{q=0}^{m}\mathbb{1}_{\left\{X_{j+q,n}^{(n)}-X_{j+q-1,n}^{(n)}=0\right\}}.$$

The rationale of the estimator is the following. Consider $\mathrm{IT}_n^{(m)}$ for a fixed $m \geq 2$ and $j \in \{1,\ldots,n-m\}$. If all the m consecutive price adjustments are zero, the product of the indicator functions is equal to one and contributes to the summation. Conversely, if at least one among the m price adjustments is different from zero, the product of the indicator functions is equal to zero and does not contribute to $\mathrm{IT}_n^{(m)}$. When Bernoulli variates are $i.i.d$, $\mathrm{IT}_n^{(m)}$ estimates the joint probability of m consecutive zeros. In the most general case, we have the following result.

Theorem 3.2. *Under Assumption 1 and Assumption 2, as $n \to \infty$, we have:*

$$\mathrm{IT}_n^{(m)} \xrightarrow{u.c.p} \int_0^1 (p_s)^m\,ds.$$

Moreover, as $n \to \infty$:

$$\sqrt{n}\begin{bmatrix}\mathrm{IT}_n - \int_0^1 p_s\,ds\\ \mathrm{IT}_n^{(m)} - \int_0^1 (p_s)^m\,ds\end{bmatrix} \stackrel{stably}{\Longrightarrow} \mathcal{MN}\left(0,\Sigma^{(m)}\right),$$

where $\mathcal{MN}(0,\Sigma^{(m)})$ denotes the mixed-normal distribution with covariance matrix:

$$\Sigma^{(m)} = \begin{bmatrix}\int_0^1 p_s\,(1-p_s)\,ds & \int_0^1 m\,p_s^m\,(1-p_s)\,ds\\ \int_0^1 m\,p_s^m\,(1-p_s)\,ds & \int_0^1 p_s^m\,\frac{p_s^m\,(2\,m+1)-p_s^{m+1}\,(2\,m-1)-(1+p_s)}{1-p_s}\,ds\end{bmatrix}.$$

Proof. See Appendix A.2, Lemma 5. □

The estimation of the entries of the matrix $\Sigma^{(m)}$ requires a consistent estimator of functionals of the form:

$$U\,(f) = \int_0^1 f\,(p_s)\,ds, \tag{7}$$

with $f(\cdot)$ being a (sufficiently regular) deterministic function. We discuss this point below.

3.2 Local estimation of the probability of staleness

The estimation of functionals of the type (7) is feasible once a local estimator of p_t is available. Therefore, we first choose a sequence $k_n \geq 2$ of integers that satisfies $k_n \to \infty$ and $k_n \Delta_n \to 0$, and then we define the *local idle time* as:

$$\widehat{p}_i(k_n) = \frac{1}{k_n} \sum_{j=0}^{k_n-1} \mathbb{1}_{\left\{X^{(n)}_{i+j+1,n} - X^{(n)}_{i+j,n} = 0\right\}} \qquad i \in \{1, \ldots, n-k_n\}. \tag{8}$$

Note that the condition $k_n \Delta_n \to 0$ ensures that we are taking local averages. The functional $U(f)$ can then be estimated via standard Riemann sums, in which the instantaneous probability of flatness is replaced by the local idle time in (8). For this reason, we define the discretized version of $U(f)$ as:

$$U(\Delta_n, f)^n = \Delta_n \sum_{i=1}^{n-k_n+1} f(\widehat{p}_i(k_n)),$$

and we derive its asymptotic properties in the following theorem.

Theorem 3.3. *Let $f(\cdot)$ be a locally bounded function. Under Assumption 1 and Assumption 2, as $n \to \infty$, it holds that:*

$$U(\Delta_n, f)^n \overset{u.c.p.}{\longrightarrow} \int_0^1 f(p_s)\, ds. \tag{9}$$

Proof. See Appendix A.3. □

The idea of estimating the functionals $U(f)$ through $U(\Delta_n, f)^n$ follows the same logic as in Jacod and Rosenbaum (2013, 2015) for the estimation of volatility functionals. As in their case, the $U(\Delta_n, f)^n$ in (9) admits a stable CLT with an $\mathcal{F}$-conditional Gaussian limit, which is, however, not centered. If $k_n \sim \theta/\sqrt{\Delta_n}$ for some constant θ, the $\mathcal{F}$-conditional mean of the limit consists of several bias terms depending on end effects, the second derivative of f, and the quadratic variation of p_t. If k_n diverges slower than $1/\sqrt{\Delta_n}$, the $\mathcal{F}$-conditional mean of the limit depends only on the second derivative of f, while the other bias terms are asymptotically immaterial. Because the estimation of the quadratic variation of p_t carries some complications (in particular, the convergence rate of the estimator is small, see Jacod and Rosenbaum, 2015), in what follows we will assume that $k_n\sqrt{\Delta_n} \to 0$.

Under these settings, the bias-corrected[5] version of $U(\Delta_n, f)^n$ takes the form:

$$U'(\Delta_n, f)^n = \Delta_n \sum_{i=1}^{n-k_n+1} \left(f(\widehat{p}_i(k_n)) - \frac{1}{2\,k_n} f''(\widehat{p}_i(k_n))\,\widehat{p}_i(k_n)(1-\widehat{p}_i(k_n)) \right) \quad (10)$$

and delivers the following stable CLT.

Theorem 3.4. *As $n \to \infty$, let k_n be a sequence of integers such that $k_n^2\Delta_n \to 0$ and $k_n^3\Delta_n \to \infty$. In addition, let f be a test function satisfying the following condition:*

$$\left|f^{(j)}(p)\right| \le K\left(1 + |p|^{m-j}\right), \quad j = 0, 1$$

for suitable positive constants K and m. As $n \to \infty$, under Assumption 1 and Assumption 2, we have that:

$$\frac{1}{\sqrt{\Delta_n}}\left(U'(\Delta_n, f)^n - \int_0^1 f(p_s)\,ds\right) \overset{stably}{\Longrightarrow} \mathcal{MN}(0, \Sigma_{\mathrm{U}}),$$

where $\mathcal{MN}(0, \Sigma_{\mathrm{U}})$ denotes the mixed-normal distribution with covariance matrix:

$$\Sigma_U = \int_0^1 f'(p_s)^2\, p_s(1-p_s)\,ds.$$

Proof. See Appendix A.3. □

Finally, in all the empirical applications we will adopt the finite-sample correction:

$$U''(\Delta_n, f)^n = \frac{(n-k_n+1)^{-1}}{\Delta_n} U'(\Delta_n, f)^n, \quad (11)$$

which mitigates the impact of the end effects in short samples and is asymptotically irrelevant. In Section 4, we compare the performances of $U''(\Delta_n, f)^n$ and $\mathrm{IT}_n^{(m)}$ in estimating the integrated powers of p_t.

3.3 A test for constant against time-varying probability of staleness

We now describe how the inferential theory discussed thus far allows deriving a statistical test able to discriminate, over one day of observation, between a constant and a time-varying probability of staleness, over one day of observation. Consider,

[5]Notice that the form of the bias is analogous to that of Jacod and Rosenbaum (2013), Equation (3.8). As in the latter, this bias is due to the local estimation of the probability of staleness.

therefore, the following partition of the sample space Ω:

$$\begin{aligned} \Omega_0 &= \left(\omega \in \Omega \,\middle|\, \int_0^1 (p_t(\omega))^m \, dt = \left(\int_0^1 p_t(\omega)\, dt\right)^m\right), \\ \Omega_1 &= \left(\omega \in \Omega \,\middle|\, \int_0^1 (p_t(\omega))^m \, dt \neq \left(\int_0^1 p_t(\omega)\, dt\right)^m\right), \end{aligned} \tag{12}$$

where m is an arbitrary integer. If a path $p_t(\omega)$ is constant on $[0,1]$ then $\omega \in \Omega_0$. On the contrary, if $\omega \in \Omega_1$ then the corresponding trajectory of the probability of staleness must, at least, depart from a constant path in a sub-set of the interval $[0,1]$ with a non-zero Lebesgue measure. By virtue of Theorem 3.1, Theorem 3.2, and the delta method, the random variable defined as:

$$\Psi_{n,m} \stackrel{\text{def}}{=} \frac{\sqrt{n}\left(\mathrm{IT}_n^{(m)} - (\mathrm{IT}_n)^m\right)}{\sqrt{\frac{(\mathrm{IT}_n)^{2m+1}(m^2+2m-1)-(\mathrm{IT}_n)^{2m}(2m^2+2m+1)+(\mathrm{IT}_n)^{m+1}+(\mathrm{IT}_n)^m}{\mathrm{IT}_n-1}}} \tag{13}$$

is the natural candidate for test statistics that may asymptotically distinguish whether the observed price staleness, as defined through Assumption 1, stems from a p_t in Ω_0 or in Ω_1. The asymptotic limits of the $\Psi_{n,m}$ test statistics are discussed in the following corollary.

Corollary 1. *As $n \to \infty$:*

$$\begin{cases} \Psi_{n,m} \stackrel{stably}{\Longrightarrow} \mathrm{N}(0,1) & on\ \Omega_0, \\ \Psi_{n,m} \stackrel{p}{\longrightarrow} +\infty & on\ \Omega_1. \end{cases}$$

Proof. See Appendix A.2. □

The limiting null distribution of $\Psi_{n,m}$ coincides with that of the zero-mean normal random variable with unit variance while, on the alterative, the test statistics diverges in probability, thus delivering a unit power. Note that the asymptotic properties of $\Psi_{n,m}$ are independent from the value of m. In the finite sample, however, m can trade off the size and power of the test. We will discuss this point in Section 4, which is dedicated to the Monte Carlo simulations.

3.4 On the estimation of the volatility of staleness

In this section, we prove that under the semimartingale model of Assumption 2, which falls within the class of processes compatible with Ω_1, it is possible to define a feasible and consistent estimator of the quadratic variation of the probability of staleness, that is $\overline{\nu} \doteq \int_0^1 \nu_s^2 \, ds$. From an economic point of view, interpreting

p_t as an illiquidity proxy[6], $\overline{\nu}$ is readily interpretable as a measure of volatility of illiquidity. If p_t were observed, the natural estimator for the quadratic variation of the semimartingale in (3) would be:

$$\sum_{i=1}^{n} \left(\Delta_i^n p\right)^2 ,$$

where $\Delta_i^n p = p_{i,n} - p_{i-1,n}$. However, the process p is latent, and therefore a proxy for the discrete increments $\Delta_i^n p$ is needed. The local idle time in equation (8) can be adopted for this purpose, as stated in the following theorem.

Theorem 3.5. *Let $k_n = \lfloor \theta \sqrt{n} \rfloor$ be a sequence of integers for some constant $\theta > 0$. As $n \to \infty$, under Assumption 1 and Assumption 2, it holds that:*

$$\widehat{\overline{\nu}}_n^{\star}(k_n) \doteq k_n^{-1} \sum_{i=1}^{n-2k_n+1} \left(\widehat{p}_{i+k_n}(k_n) - \widehat{p}_i(k_n)\right)^2 \xrightarrow{p} \frac{2}{3} \int_0^1 \nu_s^2 \, ds + \frac{2}{\theta^2} \int_0^1 p_s \left(1 - p_s\right) \, ds \quad ,(14)$$

where $\widehat{p}_i(k_n)$ is the local idle time defined in equation (8).

Proof. See Appendix A.4. □

Several remarks are needed at this point. First, in contrast to the assumption in Theorem 3.4, now we must assume $k_n \sim \theta/\sqrt{\Delta_n}$. Second, the sum of the squared increments of local idle time converge in probability to two-thirds of the integrated volatility of p_t plus a bias term that is proportional to the asymptotic variance of IT_n (cfr. Theorem 3.2). Nevertheless, this bias does not constitute an issue because it can be consistently estimated via $U''\left(\Delta_n, f\right)^n$ using a suitable f. Indeed, a consistent estimator of $\overline{\nu}$ can be defined as:

$$\widehat{\overline{\nu}}_n \doteq \frac{3}{2} \left(\widehat{\overline{\nu}}_n^{\star}(k_n) - \frac{2}{\theta^2} U''(\Delta_n, f)^n \right), \tag{15}$$

where $f\left(x\right) = x\left(1 - x\right)$. Note that, by construction, it is not guaranteed that $\widehat{\overline{\nu}}_n \geq 0$. To circumvent this problem, one might use $\max\left(\widehat{\overline{\nu}}_n, 0\right)$ instead as a non-negative estimator.

[6]The interpretation of p_t as an illiquidity proxy is (mainly) motivated by the work of Bandi et al. (2017), where the authors provide an economic rationale for zeros that hinges on microstructural theories of price formation with transaction costs and asymmetries in information. In particular, the probability of the occurrence of a zero return depends on $(p_t)_{t\geq 0}$.

4 Monte Carlo simulations

In the absence of finite-sample distortions, the implementation of the asymptotic theory developed in Section 3 would require the adoption of the highest frequency available for the data; the greater the frequency the closer the random quantities are to their limits (either in probability or stably in law). Nevertheless, price discreteness may affect these limits, producing unwanted spurious effects. More precisely, in the presence of rounding, there could be some extra zero returns not generated by "genuine" staleness, that is not driven by the stochastic process p_t defined in Assumption 1. In this section, we explore the finite sample contaminations of the asymptotic theory by means of Monte Carlo simulations. In particular, we want to assess both the accuracy of $\mathrm{IT}_n^{(m)}$ and $U''(\Delta_n, f)^n$ in estimating the integrated volatility functional of p_t and the sizes and the powers of the test $\Psi_{n,m}$ defined in (13). For this purpose, we generate a large artificial dataset of efficient price paths contaminated by staleness and rounded at one cent (as imposed by the actual settings of electronic financial markets). For each replication we simulate a trading day of 6.5 hours on a gird of one second for a total of $6.5 \times 60 \times 60$ steps. To begin, we create the path of an efficient log-price process $Y_t = \log(P_t)$ driven by a one-factor stochastic volatility model, the dynamics of which are described by the SDE:

$$\begin{aligned} d\log\sigma_t^2 &= \left(\alpha - \beta\,\log\sigma_t^2\right)\,dt + \eta\,dW_{\sigma,t}, \\ dY_t &= \mu\,dt + c_\sigma\,\sigma_t\,dW_{Y,t}, \end{aligned} \tag{16}$$

where $W_{\sigma,t}$ and $W_{Y,t}$ are two Brownian motions with $\mathrm{corr}(dW_{\sigma,t}, dW_{Y,t}) = \rho\,dt$. We adopt the values for the parameters α, β, η, μ, and ρ estimated by Andersen et al. (2002) on S&P500. The volatility factor c_σ can be tuned to generate different volatility scenarios. It will be equal to $c_\sigma = 2$, unless otherwise specified. Numerical integration of the SDE in (16) is performed on a one-second time grid via a standard Euler scheme and with the initial conditions $Y_0 = \log(P_0)$, with $P_0 = 100$, and $\log\sigma_0^2 = \alpha/\beta$. Once simulated, the efficient prices are sampled every 30 seconds. For each replication this sub-sample produces log prices $Y_{j,n}$ with $j = 1, ..., n$ and $n = 780$. Then, on the time grid of 30 seconds we construct the staleness-contaminated log-price process $X_{j,n}$ following the recursive equation:

$$\begin{cases} X_{0,n} = Y_{0,n} = \log(P_0) \\ X_{j,n} = (1 - \mathbb{B}_{j,n})\,Y_{j,n} + \mathbb{B}_{j,n}\,X_{j-1,n}, \end{cases} \tag{17}$$

where $\mathbb{B}_{j,n}$ are Bernoulli random variables generated as described below, according to either the null Ω_0 or the alternative Ω_1 (see equation (12)). Finally, the prices $\exp(X_{j,n})$ are rounded at one cent. The rounding is the only reason that prevents taking the highest frequency available.

The null Ω_0: constant probability of staleness. In this specification, the $\mathbb{B}_{j,n}$s are i.i.d. Bernoulli random variables with constant expected value $\mathbb{E}[\mathbb{B}_{j,n}] = p_F$ for all j. We put[7] $p_F = 0.5$.

The alternative Ω_1: semimartingale-driven probability of staleness. This specification corresponds to Assumption 2. For each replication, we generate a path of a latent stochastic process u with the following (discrete-time) integration scheme:

$$\begin{cases} u_{0,n} = F^{-1}(p_F) \\ u_{j,n} = u_{j-1,n} + \left(F^{-1}(p_F) - u_{j-1,n}\right)\Delta_n + \sigma_u\,\varepsilon_{j,n}\sqrt{\Delta_n}, \end{cases} \tag{18}$$

with $j = 1, ..., n$, $\Delta_n = 1/n$, $n = 780$, $p_F = 0.5$, and where $F^{-1}(x)$ is the inverse of the cumulative distribution function of a standard Gaussian variable. The $\varepsilon_{j,n}$s are i.i.d. standard Gaussian shocks, and σ_u is a tuning parameter that we set to $\sigma_u = 1.5$. Next, a path of the stochastic probability p_t, defined in equation (2) of Assumption 1, is generated as:

$$p_{j,n} = \int_{-\infty}^{u_{j,n}} \frac{1}{\sqrt{2\,\pi}}\, e^{-z^2/2}\, dz = F(u_{j,n}). \tag{19}$$

Note that since because u is a mean-reverting around $F^{-1}(p_F)$ by construction, then p_t is mean-reverting around p_F. Therefore, on average, the probability of zeros is similar to the value used in the constant probability case.

Figure 2 displays an example of observed log price under Assumption 2 for p_t. In the figure, observed stock price dynamics are indicated in black, red crosses represent zero return, and the dotted line with blue circles represents the probability of staleness. Looking at the first part of the (fictitious) trading day (before 10:30 EST) we see that the price process is considerably sticky with a high probability of flat trading, as indicated by the circles. As time flows, the price process experiences different regimes (in terms of staleness), changing from a quite erratic behaviour (from around 10:30 EST to 13:00 EST) to quite sluggish behaviour (from 13:00 EST until the end of the day).

[7] With this numerical choice we are assuming that, at the frequency of 30 seconds, fifty percent of the log-returns are zeros. This corresponds to a moderately high level of illiquidity for the asset.

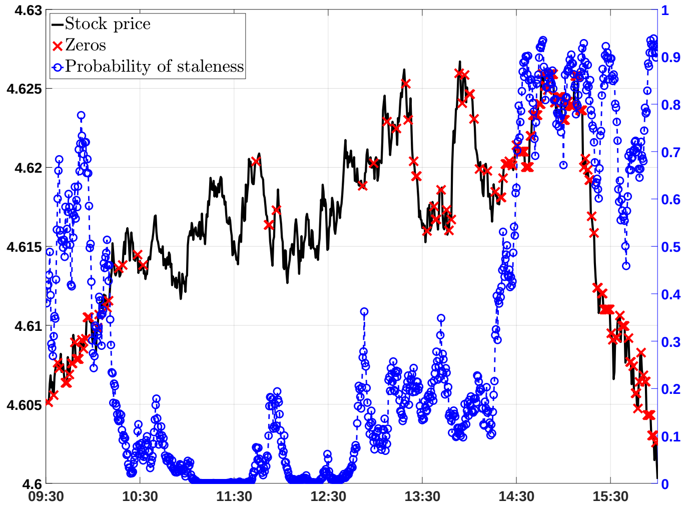

Figure 2: Example of a simulated log-price path of the observed stock price under Assumption 2 for p_t. The path of the stock price is indicated in black (with values reported on the left vertical axis) while red crosses signal the position of zero returns. The dotted line with blue circles represents the probability of staleness (with values reported on the right vertical axis).

4.1 Sizes and powers of $\Psi_{n,m}$

The test statistic $\Psi_{n,m}$ is characterized by a choice variable; more precisely, it depends on the number m of factors in the multi-idle time $\mathrm{IT}_n^{(m)}$ defined in (7). Asymptotically, the distribution of $\Psi_{n,m}$ is unaffected by the value of m, as well as its divergence toward $+\infty$ under the correspondent alternative hypothesis. Nevertheless, in finite sample m can be chosen to trade off the size and power of the test. Following the procedures described in Section 4, we generate 10^4 replications of (rounded) price paths under the null Ω_0 and the alternative Ω_1, and, for different choices of m, we evaluate the size and power of the test $\Psi_{n,m}$ by computing its rejection rates under the proper set of artificial data. Figure 3 summarizes the results of this numerical experiment; in particular, we report 5% rejection rates of the test under the null and the alternative. A reasonable trade-off between size and power is attained taking m around 5, a choice that maximizes power and gives a conservative (less than the theoretical 5%) size. Of course, the specific power of the test depends on how the alternative is formulated and, in our case, on the tuning parameter σ_u. For example, an higher value for the parameter σ_u in (18) would deliver a more powerful $\Psi_{n,m}$.

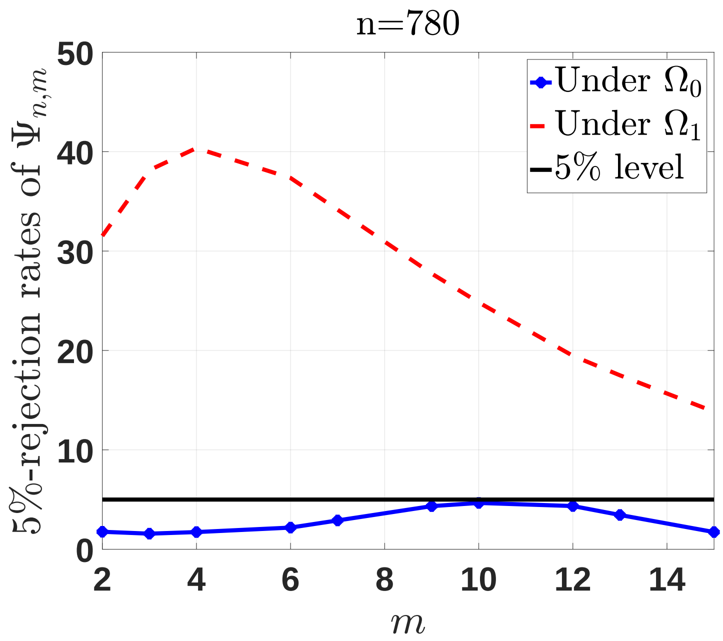

Figure 3: This plot reports 5% rejection rates of the test based on the asymptotic limit of the random variable $\Psi_{n,m}$ defined in equation (13). We set $n = 780$ (which corresponds to sampling prices every 30 seconds, assuming a day of 6.5 hours), and we use m as an independent variable. The blue line with circles and the red dotted line correspond, respectively, to the rejection rates under the null Ω_0 (i.e., the size of the test) and the alternative Ω_1 (i.e., the power of the test). Prices are rounded at one cent.

4.2 Estimation of $\int_0^1 p_s (1 - p_s)\, ds$ and of $\int_0^1 \nu_s^2\, ds$

This section is dedicated to the assessing the finite sample accuracy in the estimation of functionals of p_t, that is $U(f) = \int_0^1 f(p_s)\, ds$, and the integrated volatility of p_t, namely $\int_0^1 \nu_s^2\, ds$. Concerning the functionals, we focus on estimating the asymptotic variance of idle time that requires selecting $f(x) = x(1-x)$ (see Theorem 3.1) in $U(f)$. We generate sample paths of the probability of staleness on Ω_1, constructing the paths of p_t according to equations (18) and (19), with $\sigma_u = 2$. First, we consider estimation without rounding. Second, we analyze estimation when the price process generated according to equation (17), is rounded at one cent.

We start with estimating Σ_{IT}, defined in equation (6). According to Theorem 3.2 and Theorem 3.3, this random variable is estimated either by the difference $\mathrm{IT}_n - \mathrm{MIT}_n^{(2)}$ or by $U''(\Delta_n, f)^n$ with $f(x) = x - x^2$. Figure 4 summarizes the results of this numerical experiment for $\int_0^1 p_s (1 - p_s)\, ds$ ranging between 0 and 0.3. Both estimators are remarkably precise. However, the dispersion of $U''(\Delta_n, f)^n$, which is computed with the block length $k_n = 13$, is considerably smaller than the variance of $\mathrm{IT}_n - \mathrm{MIT}_n^{(2)}$. This is not totally surprising, as the former estimator constitutes a localized maximum likelihood estimator. This result is robust across different choices of the tuning parameter k_n as illustrated by Figure 5, which reports the root mean square error[8] (RMSE) of $U''(\Delta_n, f)^n$ as a function of k_n and the RMSE of the difference estimator, $\mathrm{IT}_n - \mathrm{MIT}_n^{(2)}$, which does not depend on k_n. The RMSE of $U''(\Delta_n, f)^n$ has a U-shaped pattern with the minimum at around $k_n \approx 13$, which roughly corresponds to $k_n \approx n^{2/5}$. However, the variation of RMSE of $U''(\Delta_n, f)^n$ across k_n is smaller than the reduction in the RMSE in comparison to the difference estimator.

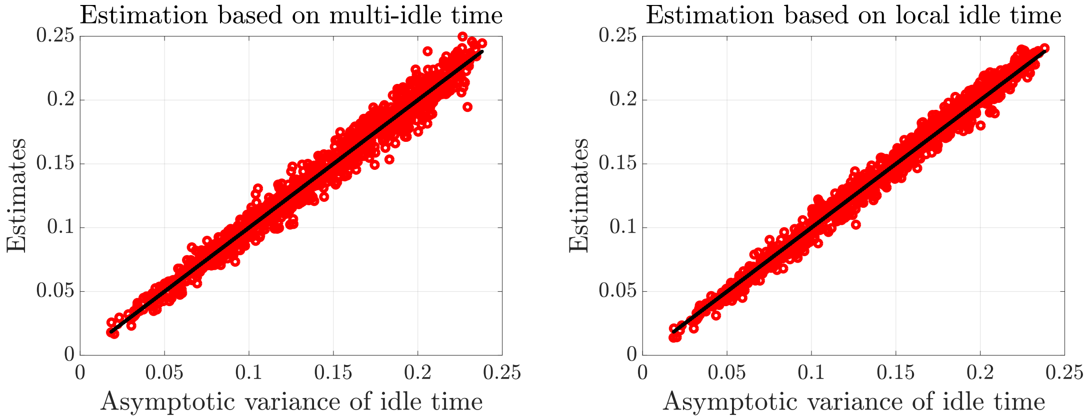

Figure 4: Scatter plot of the asymptotic variance of idle time, that is, $\int_0^1 p_s(1-p_s)\,ds$ and its estimated values based on multi-idle time (left panel) and local idle time (right panel) with $k_n = 13$. The black line represents the true value.

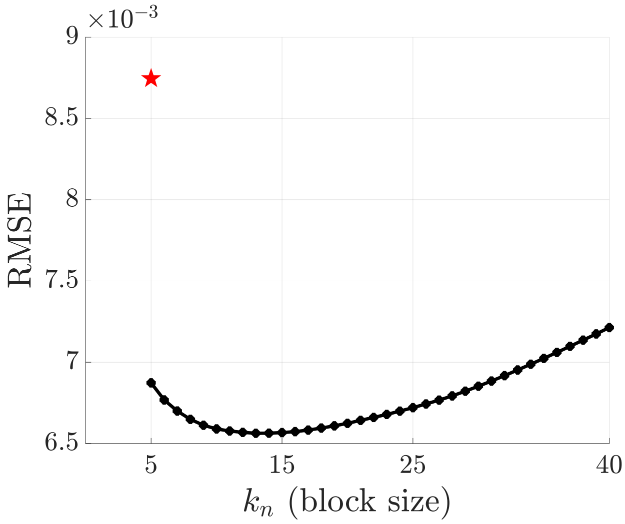

Figure 5: Root Mean Squared Error (RMSE) of $\mathrm{IT}_n - \mathrm{MIT}_n^{(2)}$ (red star) and of $U''(\Delta_n, f)^n$ (black line with circles) in estimating $\int_0^1 p_s(1-p_s)\,ds$.

[8]For both the estimators, the bias is an order of magnitude smaller than the standard deviation. Therefore, in practice the RMSEs coincides with the standard deviations. For this reason, in Figure 5 we report only the RMSE.

Next, we turn to estimating the integrated variance of p_t. In finite samples, the performance of our estimator, defined by equation (15), crucially depends on the implementation details. The estimator consists of the two parts:

$$\widehat{\overline{\nu}}_n = \frac{3}{2}\,\widehat{\overline{\nu}}^{\star}_n(k_n) - \frac{3}{\theta^2}U''(\Delta_n, f)^n, \tag{20}$$

where $\widehat{\overline{\nu}}^{\star}_n(k_n)$ is defined in equation (14).

Theorem 3.5 requires that the block length chosen for the computation of $\widehat{\overline{\nu}}^{\star}_n$ is $k_n \sim \theta/\sqrt{\Delta_n}$ for some $\theta > 0$. In contrast, based on Theorem 3.4, we need a different block length, say k'_n, satisfying $(k'_n)^2\,\Delta_n \to 0$ and $(k'_n)^3\,\Delta_n \to \infty$ for computing $U''(\Delta_n, f)^n$, the second part of the estimator. In finite samples, this means that the block length used for estimating the "bias" term in $\widehat{\overline{\nu}}_n$ ought to be smaller than the block length used for computing the leading part. Extensive Monte Carlo experiments suggest that the best finite-sample performance of $\widehat{\overline{\nu}}_n$ is achieved for $k_n = \lfloor 1/\sqrt{n} \rfloor$ and $k'_n = \lfloor k_n/5 \rfloor$.

With this heuristic rule of thumb, we estimate daily and monthly integrated variance of p_t. Figure 6 displays the histograms of the difference between true and estimated values. The left panel corresponds to daily estimates. The figure shows that the estimator is nearly unbiased, as the difference is centred at zero. However, it is highly volatile because the convergence rate of $\widehat{\overline{\nu}}_n$ is $n^{1/4}$ and not $n^{1/2}$ as, for example, for the estimator of integrated functionals of p_t. Estimating $\int_0^1 \nu_s^2\,ds$, monthly instead of daily allows reducing the variance, as illustrated in the right panel of Figure 6. Theoretically, the performance of the estimator can be improved by moving to the higher frequencies. However, in practice this is hardly possible due to the adverse effect of price discreteness discussed below.

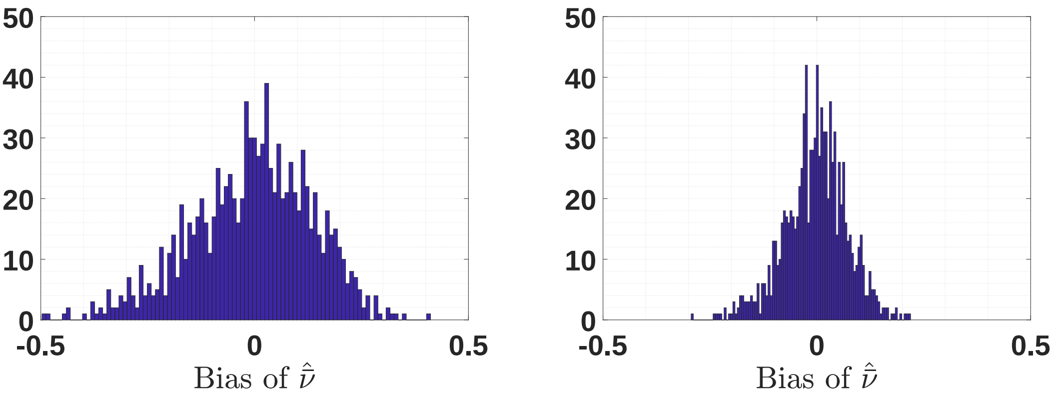

Figure 6: Histogram of the relative bias of the estimator $\hat{\bar{\nu}}_n$ defined in equation (15) in estimating the integrated volatility $\int_0^1 \nu_s^2\,ds$.

As discussed above, the rounding of prices may produce a number of zero returns, that are not associated with the Bernoulli random variables generated in equation (17). Therefore, it disrupts the theoretical limiting behavior of the estimators. To mitigate this contamination effect, the estimators ought to be implemented at sampling frequencies for which the number of zeros due to the rounding is relatively small. It can be easily verified that for a given sampling frequency the number of zeros produced by rounding is inversely related to both the (average) price level and its (average) volatility. To illustrate this effect, we report in Figure 7 the relative errors (in percentage and averaged across one thousand simulations) of the estimators $U'\left(\Delta_n, x(1-x)\right)^n$ (left panel) and $\widehat{\widehat{\nu}}_n$ (right panel) for different values of the initial price P_0 and of the average daily realized volatility[9]. In both cases, considering the area defined by initial prices larger than (roughly) 75 and daily realized volatilities larger than (roughly) 1%, the biases of the two estimators are negligible. Therefore, the estimation theory is meant to work for such assets. For assets with either lower prices of less volatility, the estimators ought to be implemented at lower frequencies.

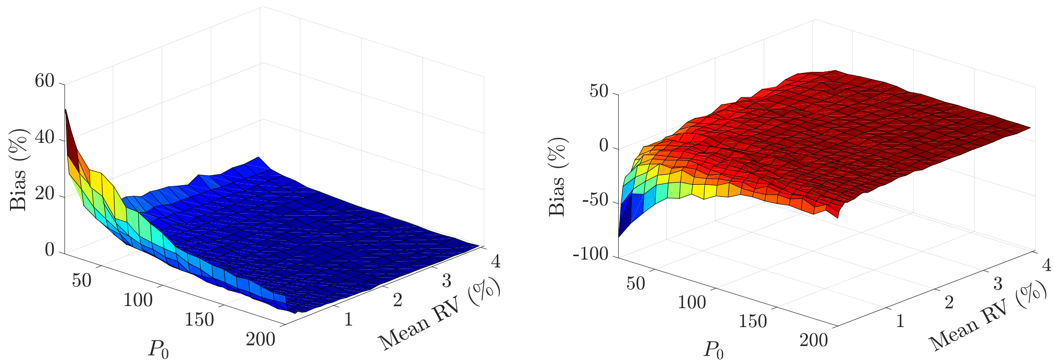

Figure 7: Relative bias (in percentage) as a function of the initial price P_0 and of the (average) daily volatility (in percentage) of (left panel) the estimator $U'\left(\Delta_n, f\right)^n$ defined in equation (10), with $f(x) = x(1-x)$, in estimating the stochastic integral $\int_0^1 p_s (1-p_s)\, ds$ and of (right panel) the estimator $\widehat{\widehat{\nu}}_n$ defined in equation (15), in estimating the integrated volatility $\int_0^1 \nu_s^2\, ds$.

5 Empirical application

As a first empirical application of our theoretical framework, we test whether the paths of p_t are constant or time-varying using a large dataset of transaction prices of NYSE-listed stocks. Following the limiting results discussed in Section 3.3, this amounts to discriminating between the two hypotheses:

$$\mathcal{H}_0 : \Big(\left(p_t(\omega)\right)_{t\in[0,1]} \in \Omega_0 \Big) \quad \text{against} \quad \mathcal{H}_1 : \Big(\left(p_t(\omega)\right)_{t\in[0,1]} \in \Omega_1 \Big) .$$

[9]We generate different levels of average volatility by varying the factor c_σ in equation (16).

We implement the test for every stock and day in our sample. We employ a data set whose constituents are the most 250 liquid[10] NYSE-listed stocks. The data range from January 3, 2006 to December 31, 2014, covering a time span of 2246 trading days. We rank the stocks in deciles according to total volume traded, and within each decile we pick up the stock with the highest average price. This choice is dictated by the necessity of mitigating, the impact of rounding as much as possible. The data filtering described so far produces the 10 tickers APA, BA, CVX, DE, EOG, GS, MCD, MMM, UNP, and XOM. Transaction prices are sampled with previous-tick interpolation, every 30 seconds producing 780 observations from 09 : 30 EST to 15 : 30 EST for each trading day.

Figure 8 shows the kernel density of the test statistics $\Psi_{n,m}$, computed pooling across the 10 stocks and all the days in the sample. The distribution of $\Psi_{n,m}$ significantly differs from that of a standard normal, indicating that for the majority of days and stocks the null hypothesis $\mathcal{H}_0$ is rejected. This constitutes clear empirical evidence that points toward a time-varying model for $(p_t)_{t\geq 0}$.

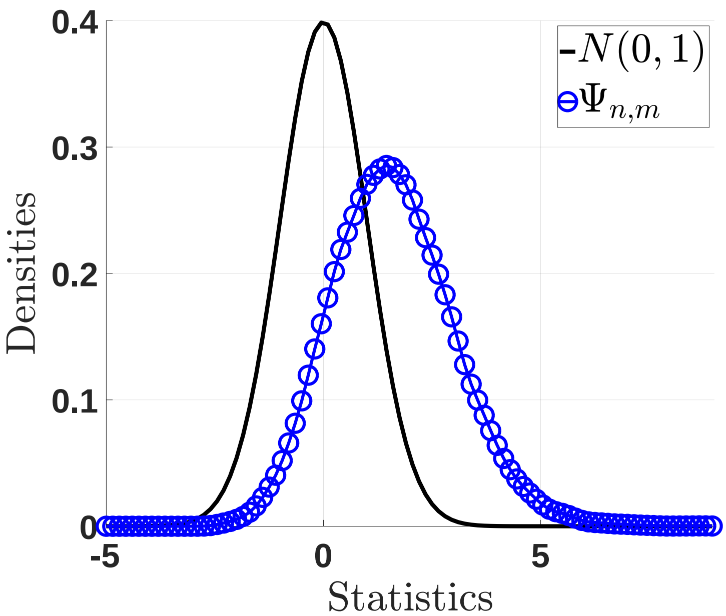

Figure 8: Kernel smooth density estimates of the daily test statistics of $\Psi_{n,m}$ computed pooling across the 10 selected stocks and the 2246 trading days.

This feature is further confirmed by looking at historical time series of local idle time. As an example, consider the plots in Figure 9 in which we consider the case of the ticker XOM, the most liquid of the 10 selected stocks. The left panel reports two paths of the estimator defined in equation (8), for the day with the smallest (black continuous line) and largest (blue lines with empty circles) value of

[10]In terms of average transaction volumes during the period considered.

$\hat{\tilde{\nu}}_n$ (cfr. equation (15)) . In the morning, the level of staleness is around 20% for both days with low and high volatility of p_t. While for the first day the process p_t hovers around an average value, for the other day it increases to around 60% at 10 : 30 EST and 14 : 30 EST. The right panel of Figure 9 shows the local idle time estimates, and 95% confidence bands averaged across the whole sample. Notice that local idle time exhibits a pronounced inverse U-shape pattern mirroring the U-shape of local volatility. In addition, the occurrence of zeros is almost two times less probable in the morning with respect to the noonday; the average local idle time is equal to 0.12 at 09 : 30 EST, and increases up to 0.24 at 12 : 30 EST. Analyzed together, the information provided by the left and right plots of Figure 9, indicates that both deterministic and stochastic components significantly contribute to the intraday variation of price staleness.

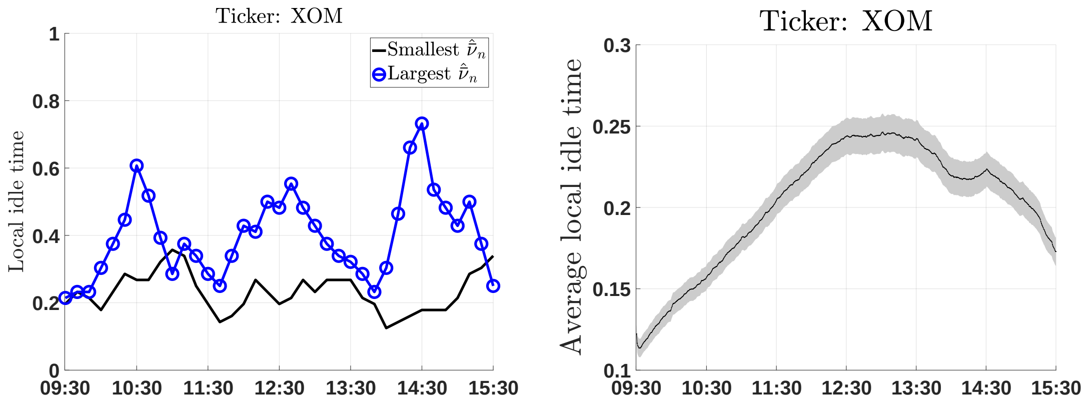

Figure 9: (Right panel). Local idle time estimates for days corresponding to the lowest (black line) and the highest (blue line with empty circles) volatility of $(p_t)_{t\geq 0}$ for XOM. (Left panel). Average, of the intraday local idle time estimates for XOM computed over the whole sample.

We conclude the empirical investigation with an econometric analysis of the economic determinants of the integrated volatility of p_t. More precisely, assuming that p_t is generated by the semimartingale in equation (3), we investigate which dimensions of illiquidity among transaction volume, bid-ask spread, and price volatility, mostly affect its integrated volatility $\int_0^1 \nu_s\, ds$ over a given period. In particular, for the purpose of minimizing finite sample distortions we focus on weekly integrated volatility. To carry out our analysis, for each stock in the sample we construct weekly estimates of:

1. integrated volatility of p_t, without bias correction, computed as $3/2 \cdot \widehat{\overline{\nu}}_n^*(k_n)$, where $\widehat{\overline{\nu}}_n^*(k_n)$ is defined as equation (17). Henceforth, we indicate this variate with $\tilde{\bar{\nu}}_t^{(w)}$.

2. Integrated p_t and p_t^2, henceforth indicated as $\text{Int}_t^{(w)}$ and $\text{Int2}_t^{(w)}$, respectively.

3. Average transaction volume, henceforth indicated as $\text{Vol}_t^{(w)}$.

4. Average bid-ask spread, henceforth indicated as $\text{Spread}_t^{(w)}$.

5. Five-minute realized volatility, henceforth indicated as $\sqrt{\text{RV}_t^{(w)}}$.

The value of k_n is set as in the Monte Carlo study. To construct the integrated quantities defined in point 2 above, we implement the estimator in equation (11), as suggested by the the numerical simulation. Moreover, to achieve a suitable scaling in the regressions, transaction volume is measured in thousands of USD. According to Theorem 3.5, $\tilde{\bar{\nu}}_t^{(w)}$ is the biased measure of the integrated volatility of p_t, with the bias proportional to the difference $\text{Int}_t^{(w)} - \text{Int2}_t^{(w)}$. Therefore, to isolate the effect of the explanatory variables on the volatility of p_t we include the variates $\text{Int}_t^{(w)}$ and $\text{Int2}_t^{(w)}$ in the set of regressors.

In light of these considerations, we run the following regression:

$$\tilde{\bar{\nu}}_t^{(w)} = a_0 + a_1\text{Int}_t^{(w)} + a_2\text{Int2}_t^{(w)} + a_3\text{Vol}_t^{(w)} + a_4\text{Spread}_t^{(w)} + a_5\sqrt{\text{RV}_t^{(w)}} + \varepsilon_t, \quad (21)$$

where ε_t denotes random errors, and the index t runs across the 451 weeks of our sample.

Table 1 reports the ordinary least squares (OLS) estimates of the regression (21). As, by construction, the dependent variable $\tilde{\bar{\nu}}_t^{(w)}$ includes a bias term of the form $\frac{3}{\theta}\left(\text{Int}_t^{(w)} - \text{Int2}_t^{(w)}\right)$, under the null of an integrated variance of p_t independent from p_t itself the coefficients a_1 and a_2 should sum up to zero and be both equal in absolute value to $\frac{3}{\theta}$. The estimates in Table 1 suggest that this is not the case, hence pointing toward a dependence of the instantaneous volatility process ν_t in (3) from the process p_t. The coefficients a_3, a_4, and a_5 capture the effect of volume, bid-ask spread, and price volatility on the weekly integrated variance of p_t. Table 1 shows that only the effect of volume (a_3) is quite significant and negative, whereas the spread and the realized variance (a_4 and a_5) are mostly not (or mildly) significant (apart for the notable exception of MCD). For all considered stocks, the adjusted coefficients of determination are quite high, ranging from 0.614 for MCD to 0.895 for GS. This suggests that over a given period, most of the quadratic variation of price staleness can be explained by the average level of staleness and the average transaction volume. Because the level of the staleness is itself affected by the transaction volume, this result points toward a non-linear dependence of the volatility of staleness on transaction volumes, something that deserves a deeper investigation, especially in terms of the implications for micro-structural models of price formation.

	a_0	a_1	a_2	a_3	a_4	a_5	R^2
EOG	0.033	2.970***	2.250**	−0.043***	0.586	−0.003	0.794
	(0.070)	(0.442)	(0.683)	(0.009)	(0.414)	(0.009)	
UNP	0.144**	2.010***	2.974***	−0.072***	−0.217	0.007	0.799
	(0.068)	(0.352)	(0.507)	(0.014)	(0.362)	(0.009)	
APA	0.009	2.920***	1.550*	−0.013**	0.433	−0.004	0.775
	(0.052)	(0.358)	(0.604)	(0.006)	(0.341)	(0.006)	
DE	0.005	3.004***	1.511*	−0.021**	−0.190	0.008	0.743
	(0.063)	(0.384)	(0.589)	(0.010)	(0.516)	(0.010)	
MMM	0.192***	0.558	5.960***	−0.034***	−1.143**	0.007	0.782
	(0.062)	(0.376)	(0.600)	(0.006)	(0.561)	(0.009)	
BA	0.055	1.689***	4.742***	−0.030***	−0.234	0.014	0.807
	(0.068)	(0.383)	(0.588)	(0.006)	(0.532)	(0.011)	
GS	0.041	2.171***	5.366***	−0.023***	−0.287	0.006***	0.895
	(0.030)	(0.271)	(0.545)	(0.004)	(0.192)	(0.002)	
MCD	0.307***	1.203	3.495***	−0.040***	−5.968***	0.026	0.614
	(0.117)	(0.588)	(0.794)	(0.004)	(1.504)	(0.021)	
CVX	0.204***	−0.842	10.374***	−0.007***	−0.144	−0.014**	0.809
	(0.046)	(0.377)	(0.753)	(0.002)	(0.442)	(0.006)	
XOM	0.114***	0.668	6.294***	−0.003***	−0.515	−0.006	0.826
	(0.043)	(0.317)	(0.575)	(0.001)	(0.667)	(0.007)	

Table 1: Table reports OLS estimates of the coefficients of the linear regression in equation (21) along with the correspondent standard errors (between brackets) and the adjusted coefficients of determination (R^2). Coefficients which result to be significant at 10%, 5% and 1% confidence levels are marked, respectively, with one, two and three stars.

6 Conclusions

In this paper, we provide a general econometric framework where the statistical properties of the likelihood of observing zero returns is driven by a stochastic process p_t, which we call the (instantaneous) probability of staleness, whose unconditional expected value determines the probability of observing a repeated price.
Zero returns are naturally linked to the lack of liquidity. The proposed framework allows us to make statistical inferences about liquidity-related economic variates in a way that is analogous to the analysis of integrated volatility. In particular, we develop an asymptotic theory designed to test whether, the process p_t follows a constant or a time-varying trajectory on an intradaily basis. In the first scenario, zero returns are uniformly distributed during the day, while the second scenario involves more complex dynamics, such as the clustering of zero returns.

Using the suitably defined statistical test, we show, on a large dataset of NYSE-listed stock that the second scenario is largely the most recurrent. Supported by such strong empirical evidence, we formulate the (alternative) hypothesis that p_t follows a Brownian semimartingale dynamic. Under this (new) null hypothesis, we derive CLT for the *idle time* and the (m)-*idle time*, two functionals of the observed price paths defined as the percentage (at a given sampling frequency) of zero returns and of m consecutive zero returns, respectively. Still under the null of a Brownian semimartingale for p_t, we prove how its integrated (over a given time horizon) volatility (indicated as $\bar{\nu}$) can be consistently estimated with a methodology analogous to that used for the estimation of integrated log-price volatility. High values of $\bar{\nu}$ are associated with high variability in the occurrence of zero returns during the period of estimation.

We conclude our analysis with an econometric exercise designed to identify which of the three important dimensions of illiquidity (i.e., transaction volume, bid-ask spread, and price volatility) are the main determinants of the volatility of p_t. Using the same dataset used for the empirical test, we provide evidence that the (weekly) integrated volatility of p_t is, conditionally on the (weekly) average p_t, negatively determined by (average weekly) transaction volume, and is almost unaffected by (average weekly) bid-ask spread and (weekly) realized volatility. In summary, while p_t is, as the idle time of Bandi et al. (2017), influenced simultaneously by all the three aforementioned dimensions of illiquidity, its volatility is, conditionally on the average p_t, almost exclusively driven by transaction volumes. This empirical evidence constitutes an additional feature of the price-formation dynamics useful for the formulation of realistic market micro-structure models.

References

Aït-Sahalia, Y. and J. Jacod (2012). Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data. *Journal of Economic Literature 50*(4), 1007–1050.

Aït-Sahalia, Y. and J. Jacod (2014). *High-Frequency Financial Econometrics*. Princeton University Press.

Andersen, T., L. Benzoni, and J. Lund (2002). An empirical investigation of continuous-time equity return models. *Journal of Finance 57*, 1239–1284.

Bandi, F. M., D. Pirino, and R. Renó (2017). Excess idle time. *Econometrica 85*(6), 1793–1846.

Bandi, F. M., D. Pirino, and R. Reno (2017). EXcess Idle Time. *Econometrica 85*(6), 1793–1846.

Bandi, F. M., D. Pirino, and R. Renó (2018). Zeros. *Working Paper*.

Barndorff-Nielsen, O. E., E. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2006). A central limit theorem for realised power and bipower variations of continuous semimartingales. In *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev*, pp. 33–68.

Barndorff-Nielsen, O. E. and N. Shephard (2004). Econometric analysis of realised covariation: high frequency based covariance, regression and correlation in financial economics. *Econometrica 72*, 885–925.

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*, 798–806.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies 9*(1), 69–107.

Caporin, M., A. Kolokolov, and R. Renò (2014). Multi-jumps. Working Paper.

Glosten, L. and P. Milgrom (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*(1), 71–100.

Hall, P. and C. Heyde (1980). *Martingale Limit Theory and Its Applications*. Academic Press, New York.

Hasbrouck, J. and T. Ho (1987). Order arrival, quote behavior, and the return-generating process. *The Journal of Finance 42*(4), 1035–1048.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies 6*(2), 327–343.

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *The journal of finance 42*(2), 281–300.

Jacod, J. (2008). Asymptotic properties of realized power variations and associated functionals of semimartingales. *Stochastic processes and their applications 118*, 517–559.

Jacod, J. (2012). Statistics and high frequency data. Lectures.

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

Jacod, J. and M. Rosenbaum (2013). Quarticity and other functionals of volatility: efficient estimation. *Annals of Statistics 41*, 1462–1484.

Jacod, J. and M. Rosenbaum (2015). Estimation of volatility functionals: the case of a $\sqrt{n}$ window. In P. Friz, J. Gatheral, A. Gulisashvili, A. Jacquier, and J. Teichmann (Eds.), *Large Deviations and Asymptotic Methods in Finance*, pp. 559–590. Springer.

Kolokolov, A. and R. Renò (2017). Jumps or flatness? *Working Paper*.

Kyle, P. (1985). Continuous auctions and insider trading. *Econometrica 43*, 1315–1335.

Lee, S. and P. Mykland (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial studies 21*(6), 2535.

Podolskij, M. and M. Vetter (2010). Understanding limit theorems for semimartingales: a short survey. *Statistica Neerlandica 64*(3), 329–351.

Scott, L. O. (1987). Option pricing when the variance changes randomly: Theory, estimation, and an application. *Journal of Financial and Quantitative analysis 22*(4), 419–438.

Woerner, J. (2006). Power and multipower variation: inference for high frequency data. In A. N. Shiryaev, M. d. R. Grossinho, P. Oliveira, and M. Esquivel (Eds.), *Stochastic Finance*, pp. 343–364. Springer.

A Appendix: Proofs

The appendix is divided into four parts. Section A.1 introduces the notation and collects auxiliary results on the convergence of triangular arrays. Section A.2 is dedicated to the proofs of limiting results from Sections 3.1, 3.2 and 3.3. Section A.3 presents the proofs of Theorems 3.3 and 3.4. Finally, the proof of Theorem 3.5 is presented in Section A.4.

A.1 Notations and auxiliary results

In what follows, we indicate with $t_{j,n} = j/n$, $j \in \{0, \ldots, n\}$ the deterministic equispaced partition of the interval $[0,1]$ and with $N_n(s) = \max\{j \mid t_{j,n} \leq s\}$. Trivially, $N_n(1) = n$. We use the symbol $\xrightarrow{p}$ for the convergence in probability, and $\xrightarrow{u.c.p}$ for the uniform convergence in probability.

Now, we specify the structure of the σ-field $\mathcal{F}$. We have the following flows of information on $\mathcal{F}$: i) $(\mathcal{F}_t^{(p)})_{t\in[0,1]}$ is the natural filtration associated with the process p_t, ii) $\mathcal{U}_{t_{j,n}}$ is the σ-algebra generated by random variables $U_{0,n}, \ldots, U_{j,n}$, and iii) $\mathcal{F}_{t_{j,n}} = \mathcal{F}_{t_{j,n}}^{(p)} \vee \mathcal{U}_{j,n}$ is a discrete time filtration associated with partitioning the interval $[0,1]$ with a descretization step $\Delta_n = 1/n$. Let $\mathcal{F}_\infty^{(p)} = \vee_{t\in[0,1]}\mathcal{F}_t^{(p)}$ be the smallest σ-algebra, that contains $\cup_{t\in[0,1]}\mathcal{F}_t^{(p)}$, $\mathcal{U}_\infty = \vee_{n=2}^{\infty}\mathcal{U}_{n,n}$, and $\mathcal{F}_{t_{n,n}} = \mathcal{F}_\infty^{(p)} \vee \mathcal{U}_{n,n}$. We than have $\mathcal{F} = \mathcal{F}_\infty^{(p)} \vee \mathcal{U}_\infty$.

For the sake of readability, we denote, for a generic index $j \in \{1, \ldots, n\}$, by $\mathbb{P}_j[\cdot]$, $\mathbb{E}_j[\cdot]$, $\mathbb{V}_j[\cdot]$ the conditional probability, the conditional expectation, and the conditional variance with respect to the filtration $\mathcal{F}_{t_{j,n}}$.

In what follows, our proofs and formalism will be inspired by those of Jacod (2012), Jacod and Protter (2012), and Aït-Sahalia and Jacod (2014). We say that a triangular array of random variables ξ_j^n, $j \in \{0, \ldots, n\}$ is asymptotically negligible (AN) if:

$$\sum_{j=1}^{n} \xi_j^n \xrightarrow{u.c.p} 0,$$

that is,

$$\sup_{s\in[0,1]} \left| \sum_{j=1}^{N_n(s)} \xi_j^n \right| \xrightarrow{p} 0. \tag{22}$$

The following two remarks state simple properties that will be invoked repeatedly during the proofs.

Remark 1. *Suppose that $\sum_{j=1}^n \left|\xi_j^n\right|$ converges to zero in $\mathbb{L}^1$, i.e.:*

$$\mathbb{E}\left[\sum_{j=1}^{n} \left|\xi_j^n\right|\right] \to 0. \tag{23}$$

By standard argument, this implies that $\sum_{j=1}^n \left|\xi_j^n\right| \xrightarrow{p} 0$ and so it is sufficient to note that

$$\sup_{s\in[0,1]} \left| \sum_{j=1}^{N_n(s)} \xi_j^n \right| \leq \sup_{s\in[0,1]} \sum_{j=1}^{N_n(s)} \left|\xi_j^n\right| = \sum_{j=1}^{n} \left|\xi_j^n\right| \xrightarrow{p} 0$$

to conclude that condition (23) is enough to guarantee that ξ_j^n is AN.

Remark 2. *Throughout the paper, we will implicitly use this simple fact. If $g(s)$ is a Riemann-integrable function then on $[0,1]$*

$$\sup_{t\in[0,1]} \int_0^t |g(s)|\, ds = \int_0^1 |g(s)|\, ds,$$

where for any sequence of function $g_n(s)$, uniform convergence on $[0,1]$ of the integral of $|g_n(s)|$ is equivalent to the convergence of $\int_0^1 |g_n(s)|\, ds$.

Finally, we remind remind readers of the following two lemmas that give us a simple criterion to conclude that a triangular array is AN; these are used repeatedly in the rest of the appendix. The first one is Lemma 4.1 in Jacod (2012) and the second is Lemma B.8 in Aït-Sahalia and Jacod (2014).

Lemma 1. *Let ξ_j^n be a triangular array of $\mathcal{F}_{t_{j,n}}$-measurable random variables. If the following condition is satisfied:*

$$\sum_{j=1}^{n} \mathbb{E}_{j-1}\left[\left|\xi_j^n\right|\right] \xrightarrow{p} 0,$$

then $\sum_{j=1}^{n} \xi_j^n \xrightarrow{u.c.p} 0$, i.e. ξ_j^n is AN. Moreover, the same conclusion holds under the following two conditions:

$$\sum_{j=1}^{n} \mathbb{E}_{j-1}\left[\xi_j^n\right] \xrightarrow{u.c.p} 0, \tag{24}$$

$$\sum_{j=1}^{n} \mathbb{E}_{j-1}\left[\left(\xi_j^n\right)^2\right] \xrightarrow{p} 0. \tag{25}$$

As a consequence, if $\mathbb{E}_{j-1}\left[\xi_j^n\right] = 0$ then condition (25) is sufficient to guarantee that $\sum_{j=1}^{n} \xi_j^n \xrightarrow{u.c.p} 0$.

Lemma 2. *If $m_n, \ell_n \geq 1$ are arbitrary integers, and if for all $n \geq 1$ and $1 \leq i \leq m_n$ the variable ξ_j^n is $\mathcal{F}_{t_{j+\ell,n}}$-measurable, and if*

$$\sum_{j=1}^{m_n} \left|\mathbb{E}_{j-1}\left[\xi_j^n\right]\right| \xrightarrow{p} 0, \quad \ell_n \sum_{j=1}^{m_n} \mathbb{E}\left[\left|\xi_j^n\right|^2\right] \to 0,$$

then

$$\sup_{i \leq m_n} \left|\sum_{j=1}^{i} \xi_j^n\right| \xrightarrow{p} 0,$$

that is $\sum_{j=1}^{n} \xi_j^n \xrightarrow{u.c.p} 0$.

We now turn to characterizing the *stable convergence* of triangular arrays (cfr. Podolskij and Vetter, 2010, Definition 1). For a sequence of random variables Y_n (representing the sequence of partial sums of a triangular array), the stable convergence is defined as follows.

Definition 1. *A sequence of random variables Y_n defined on $(\Omega, \mathcal{F}, \mathcal{P})$ is said to converge $\mathcal{G}$-stably with limit Y defined on an extension of the original probability space $(\Omega', \mathcal{F}', \mathcal{P}')$ if and only if for any bounded continuous function g and any bounded $\mathcal{G}$-measurable random variable Z it holds that:*

$$\mathbb{E}\left[g(Y_n)Z\right] \to \mathbb{E}\left[g(Y)Z\right].$$

In what follows, by stable convergence we mean $\mathcal{F}_\infty^{(p)}$-stable convergence (denoted simply $\overset{stably}{\Longrightarrow}$), unless otherwise stated. The classical stable Central Limit Theorem of Hall and Heyde (1980) is not valid for the triangular arrays considered in our paper. Indeed, by construction, we have that $\mathcal{F}_{t_{j,m}} \nsubseteq \mathcal{F}_{t_{j,n}}$ whenever $n > m$. As a consequence, the nesting assumption on the filtrations as in Theorem 3.2 of Hall and Heyde (1980) fails. However, a similar stable Central Limit Theorem holds.

Theorem A.1. *For any given integer ℓ consider the triangular array random variables:*

$$\gamma_{j,n}^{(\ell)} = \varphi\left(\mathbb{B}_{j-\ell,n}, \ldots, \mathbb{B}_{j,n}, \mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right], \ldots, \mathbb{E}_{j-1}\left[\mathbb{B}_{j+\ell,n}\right]\right),$$

where $\varphi : \mathbb{R}^{2\,\ell+1} \to \mathbb{R}$ *is a locally bounded function of a finite number of variables. Define the centred triangular array* $X^{(\ell)}_{j,n}$ *as:*

$$X^{(\ell)}_{j,n} = \frac{1}{\sqrt{n}} \left(\gamma^{(\ell)}_{j,n} - \mathbb{E}_{j-1}\left[\gamma^{(\ell)}_{j,n} \right] \right)$$

and assume that:

$$\sum_{j=1}^{n} \left(X^{(\ell)}_{j,n} \right)^2 \xrightarrow{p} \sigma^2, \tag{26}$$

for an a.s. finite random variable σ. *Then, as* $n \to \infty$*:*

$$\sum_{j=1}^{n} X^{(\ell)}_{j,n} \overset{stably}{\Longrightarrow} Z, \tag{27}$$

where Z *is a random variable with* $\mathcal{F}^{(p)}_\infty$*-conditional Gaussian distribution with variance* σ^2*, defined on an extension of the original probability space.*

Proof. The technicalities of the proof largely follow the results in Hall and Heyde (1980), Lemma 3.1, and Theorem 3.2. Because of the locally boundedness of φ and the distributional assumptions on random variables $\mathbb{B}_{j-\ell,n}, \ldots, \mathbb{B}_{j+\ell,n}$, it is easy to check that $\max_{1 \le j \le n} \left| X^{(\ell)}_{j,n} \right| \xrightarrow{p} 0$. Moreover, by hypothesis $\sum_{j=1}^{n} \left(X^{(\ell)}_{j,n} \right)^2 \xrightarrow{p} \sigma^2$ for an a.s. finite random variable σ. As a consequence (cfr. Lemma 3.1 in Hall and Heyde, 1980), to prove the statement above it is sufficient to prove that for all real t the random variable $T_n(t)$ defined as ($\imath = \sqrt{-1}$)

$$T_n(t) \doteq \prod_{j=1}^{n} \left(1 + \imath\, t\, X^{(\ell)}_{j,n} \right)$$

converges to 1 as $n \to \infty$ *weakly in* $\mathbb{L}^1$. By definition, this is equivalent proving that for all $E \in \mathcal{F}$, $\mathbb{E}\left[T_n(t)\, \mathbb{I}(E) \right] \to \mathbb{P}[E]$, where $\mathbb{I}(E)$ is the indicator function of the event E. For a fixed $2 \le m \le n$, let $E_m \in \mathcal{F}_{t_{m,m}}$. We compute:

$$\begin{aligned}
\mathbb{E}\left[T_n(t)\, \mathbb{I}(E_m) \right] &= \mathbb{E}\left[\mathbb{E}\left[T_n(t)\, \mathbb{I}(E_m) \,|\, \mathcal{F}_{t_{m,m}} \right] \right] = \mathbb{E}\left[\mathbb{E}\left[\prod_{j=1}^{n} \left(1 + \imath\, t\, X^{(\ell)}_{j,n} \right) \mathbb{I}(E_m) \,\Big|\, \mathcal{F}_{t_{m,m}} \right] \right] \\
&= \mathbb{E}\left[\prod_{j \in \mathrm{I}_1} \left(1 + \imath\, t\, X^{(\ell)}_{j,n} \right) \mathbb{I}(E_m)\, \mathbb{E}\left[\prod_{j \in \mathrm{I}_2 \cup \mathrm{I}_3} \left(1 + \imath\, t\, X^{(\ell)}_{j,n} \right) \Big|\, \mathcal{F}_{t_{m,m}} \right] \right] \\
&= \mathbb{E}\left[\prod_{j \in \mathrm{I}_1} \left(1 + \imath\, t\, X^{(\ell)}_{j,n} \right) \mathbb{I}(E_m)\, \mathbb{E}\left[\mathbb{E}\left[\prod_{j \in \mathrm{I}_2 \cup \mathrm{I}_3} \left(1 + \imath\, t\, X^{(\ell)}_{j,n} \right) \Big|\, \mathcal{F}^{(p)}_\infty \right] \Big|\, \mathcal{F}_{t_{m,m}} \right] \right] \\
&= \mathbb{E}\left[\prod_{j \in \mathrm{I}_1} \left(1 + \imath\, t\, X^{(\ell)}_{j,n} \right) \mathbb{I}(E_m)\, \mathbb{E}\left[\prod_{j \in \mathrm{I}_2} \left(1 + \imath\, t\, X^{(\ell)}_{j,n} \right) \Big|\, \mathcal{F}_{t_{m,m}} \right] \mathbb{E}\left[\prod_{j \in \mathrm{I}_3} \left(1 + \imath\, t\, X^{(\ell)}_{j,n} \right) \Big|\, \mathcal{F}^{(p)}_\infty \right] \right],
\end{aligned} \tag{28}$$

where I_1, I_2, and I_3 are three sets of indexes such that $X^{(\ell)}_{j,n} \in \mathcal{F}_{t_{m,m}}$ for $j \in \mathrm{I}_1$, $X^{(\ell)}_{j,n} \in \mathcal{F}_{t_{m+\ell,m+\ell}}$ for $j \in \mathrm{I}_2$, and $X^{(\ell)}_{j,n} \in \left(\mathcal{F}_{t_{n,n}} \diagdown \mathcal{F}_{t_{m+l,m+l}} \right)$ for $j \in \mathrm{I}_3$. In particular, $\left(\mathcal{F}_{t_{n,n}} \diagdown \mathcal{F}_{t_{m+l,m+l}} \right)$ denotes the smallest σ-algebra containing all the events of $\mathcal{F}_{t_n,t_n}$ that are not included in $\mathcal{F}_{t_{m+\ell},t_{m+\ell}}$. First, we note that I_1 and I_2 includes at most a finite number of terms and that:

$$\mathbb{E}\left[\prod_{j \in \mathrm{I}_3} \left(1 + \imath\, t\, X^{(\ell)}_{j,n} \right) \Big|\, \mathcal{F}^{(p)}_\infty \right] = \prod_{j \in \mathrm{I}_3} \mathbb{E}\left[\left(1 + \imath\, t\, X^{(\ell)}_{j,n} \right) \Big|\, \mathcal{F}^{(p)}_\infty \right] = 1,$$

because of the independence of the factors conditionally on $\mathcal{F}^{(p)}_\infty$ and the fact that for each $j \in \{1, \ldots, n\}$, $X^{(l)}_{j,n}$ has expected value equal to one. Equation (28) then become

$$\mathbb{E}\left[T_n(t)\,\mathbb{I}(E_m)\right] = \mathbb{E}\left[\mathbb{I}(E_m)\prod_{j\in \mathrm{I}_1\cup\mathrm{I}_2}\left(1+\imath\, t\, X_{j,n}^{(\ell)}\right)\right] = \mathbb{P}\left[E_m\right] + R_n,$$

where the remainder term R_n consists of at most $2^{2|\mathrm{I}_1\cup\mathrm{I}_2|}-1$ terms of the form $\mathbb{E}\left[\mathbb{I}(E_m)(it)^r X_{j_1,n}^{(\ell)}\dots X_{j_r,n}^{(\ell)}\right]$, with $1\le r\le |\mathrm{I}_1\cup\mathrm{I}_2|$ and $j_1,\dots j_r\in \mathrm{I}_1\cup\mathrm{I}_2$. Note that R_n converges to zero as $n\to\infty$. Consequently

$$\mathbb{E}\left[T_n(t)\,\mathbb{I}(E_m)\right] \xrightarrow{p} \mathbb{P}\left[E_m\right].$$

Finally, let $\triangle$ denote the symmetric difference. For any $E\in\mathcal{F}$ and any $\varepsilon>0$ there exists an m and an $E_m\in\mathcal{F}_{t_{m,m}}$, such that $\mathbb{P}\left[E\triangle E_m\right]\le\varepsilon$. Because T_n is uniformly integrable by assumption,

$$\left|\mathbb{E}\left[T_n(t)\,\mathbb{I}(E_m)\right]-\mathbb{E}\left[T_n(t)\,\mathbb{I}(E)\right]\right| \le \mathbb{E}\left[\left|T_n(t)\right|\mathbb{I}(E\triangle E_m)\right],$$

and $\sup_n\left|\mathbb{E}\left[T_n(t)\,\mathbb{I}(E_m)\right]-\mathbb{E}\left[T_n(t)\,\mathbb{I}(E)\right]\right|$ can be made arbitrarily small by choosing sufficiently small ε, and hence the thesis. □

We conclude this section with the following corollary, which will be used in the subsequent sections.

Corollary 2. *Let $\mathbf{X}_{j,n}^{(\ell)}$ be a q-dimensional random vector with each component defined as $X_{j,n}^{(\ell)}$ in Theorem A.1, such that:*

$$\sum_{j=1}^{n}\mathbf{X}_{j,n}^{(\ell)}\left(\mathbf{X}_{j,n}^{(\ell)}\right)' \xrightarrow{p} \Sigma, \tag{29}$$

for an a.s. finite positive definite random matrix $\Sigma=\{\sigma_{i,j}\}$. Then,

$$\sum_{j=1}^{n}\mathbf{X}_{j,n}^{(\ell)} \overset{stably}{\Longrightarrow} \mathcal{MN}(0,\Sigma),$$

where $\mathcal{MN}(0,\Sigma)$ is a q-dimensional mixed-normal random variable.

Proof. The condition (29) implies that:

$$\sum_{j=1}^{n}\left(\mathbf{c}'\mathbf{X}_{j,n}^{(\ell)}\right)^2 \xrightarrow{p} \mathbf{c}'\Sigma\mathbf{c}$$

for an arbitrary real valued vector $\mathbf{c}=(c_1,...,c_q)'$. Consequently, by Theorem A.1, we have:

$$\sum_{j=1}^{n}\mathbf{c}'\mathbf{X}_{j,n}^{(\ell)} \overset{stably}{\Longrightarrow} \mathcal{MN}(0,\mathbf{c}'\Sigma\mathbf{c}),$$

where $\mathcal{MN}(0,\mathbf{c}'\Sigma\mathbf{c})$ denotes a mixed-normal random variable. Because $\mathbf{c}$ is arbitrary, the later convergence implies the statement of the Corollary. □

Remark 3. *The statement of Theorem A.1 remains true if the condition (26) is replaced by the analogous condition for conditional variances:*

$$\sum_{j=1}^{n}\mathbb{E}\left[\left(X_{j,n}^{(\ell)}\right)^2\,\middle|\,\mathcal{F}_{t_{j,n}}\right] \xrightarrow{p} \sigma^2.$$

A.2 Proofs of limit theorems from Section 3.1, 3.2, and 3.3

The proofs of the limiting results from Sections 3.1, 3.1, and 3.3 follow directly from several auxillary lemmas on the limiting behaviour of triangular arrays of Bernoulli random variables presented below.
We start with a remark about Assumption 2, which is repeatedly used in the subsequent proofs.

Remark 4. *Under Assumption 2,*

$$\mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right] = p_{j-1,n} + O_p\left(\Delta_n^{1/2}\right). \tag{30}$$

Indeed,

$$\mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbb{B}_{j,n} \,\middle|\, \mathcal{F}_{t_{j-1,n}} \vee \mathcal{F}^{(p)}_{t_{j,n}}\right]\right] = \mathbb{E}_{j-1}\left[p_{j,n}\right] = p_{j-1,n} + \mathbb{E}_{j-1}\left[p_{j,n} - p_{j-1,n}\right], \tag{31}$$

where

$$\left|\mathbb{E}_{j-1}\left[p_{j,n} - p_{j-1,n}\right]\right| \leq \mathbb{E}_{j-1}\left[\left|p_{j,n} - p_{j-1,n}\right|\right] \leq C\left(\Delta_n\right)^{1/2},$$

where the last inequality follows from standard estimates for semimartingales (Jacod, 2008). Moreover, by Proposition 1 of Barndorff-Nielsen et al. (2006),

$$\left|p_{j,n} - p_{j-1,n}\right| = O_p\left(\left(\Delta_n \left|\log \Delta_n\right|\right)^{1/2}\right),$$

which implies that for every finite integer k

$$p_{j+k} = p_{j-1} + O_p\left(k\,\left(\Delta_n \left|\log \Delta_n\right|\right)^{1/2}\right). \tag{32}$$

Lemma 3. *Under Assumption 2, as $n \to \infty$,*

$$\frac{1}{n}\sum_{j=1}^{n}\prod_{i=0}^{m-1}\mathbb{B}_{i+j,n} \xrightarrow{u.c.p} \int_0^1 \left(p_s\right)^m ds.$$

Proof. Consider the following quantity:

$$A_n = \frac{1}{n}\sum_{j=1}^{n}\prod_{i=0}^{m-1}\mathbb{B}_{i+j,n} - \frac{1}{n}\sum_{j=1}^{n}\left(p_{j-1,n}\right)^m = \frac{1}{n}\sum_{j=1}^{n}\left[\mathbb{B}_{j,n}\mathbb{B}_{j+1,n}\cdots\mathbb{B}_{j+(m-1),n} - \left(p_{j-1,n}\right)^m\right],$$

We show that $A_n \xrightarrow{u.c.p} 0$. To do so, we rewrite the quantity A_n as a sum of a $\mathcal{F}_{t_{j,n}}$-measurable quantity and a negligible term. We introduce the following quantity:

$$\varsigma^{(m)}_{j,\ell} = \mathbb{B}_{j,n}\mathbb{B}_{j+1,n}\cdots\mathbb{B}_{j+\ell-1,n}\left(\mathbb{B}_{j+\ell,n} - p_{j-1,n}\right)\left(p_{j-1,n}\right)^{m-\ell-1}$$

and we show that A_n can be rewritten in the following way:

$$A_n = \frac{1}{n}\sum_{j=m}^{n}\sum_{\ell=0}^{m-1}\varsigma^{(m)}_{j-\ell,\ell} + \frac{\mathcal{R}_n}{n} \tag{33}$$

where $\mathcal{R}_n/n$ is ANe. Let us consider the following expressions:

$$\begin{aligned}
\psi_{j,1} &= \mathbb{B}_{j,n} - p_{j-1,n} \doteq \varsigma^{(1)}_{j,0} \\
\psi_{j,2} &= \mathbb{B}_{j,n}\mathbb{B}_{j+1,n} - p^2_{j-1,n} = \mathbb{B}_{j,n}\left(\mathbb{B}_{j+1,n} - p_{j-1,n}\right) + \left(\mathbb{B}_{j,n} - p_{j-1,n}\right)p_{j-1,n} \doteq \varsigma^{(2)}_{j,1} + \varsigma^{(2)}_{j,0} \\
\psi_{j,3} &= \mathbb{B}_{j,n}\mathbb{B}_{j+1,n}\left(\mathbb{B}_{j+2,n} - p_{j-1,n}\right) + \mathbb{B}_{j,n}\left(\mathbb{B}_{j+1,n} - p_{j-1,n}\right)p_{j-1,n} + \left(\mathbb{B}_{j,n} - p_{j-1,n}\right)p^2_{j-1,n} \doteq \varsigma^{(3)}_{j,2} + \varsigma^{(3)}_{j,1} + \varsigma^{(3)}_{j,0},
\end{aligned}$$

and similarly for each fixed m. Then $A_n = n^{-1}\sum_{j=1}^{n}\psi_{j,m}$ becomes:

$$A_n = \frac{1}{n}\sum_{j=1}^{n}\sum_{\ell=0}^{m-1}\varsigma_{j,\ell}^{(m)} = \frac{1}{n}\sum_{j=m}^{n}\sum_{\ell=0}^{m-1}\varsigma_{j,\ell}^{(m)} + \frac{1}{n}\sum_{j=1}^{m-1}\sum_{\ell=0}^{m-1}\varsigma_{j,\ell}^{(m)} = \frac{1}{n}\sum_{j=m}^{n}\sum_{\ell=0}^{m-1}\varsigma_{j-\ell,\ell}^{(m)} + \frac{1}{n}\underbrace{\sum_{j=m}^{n}\sum_{\ell=0}^{m-1}\left(\varsigma_{j,\ell}^{(m)} - \varsigma_{j-\ell,\ell}^{(m)}\right)}_{\mathcal{R}_1} + \frac{1}{n}\underbrace{\sum_{j=1}^{m-1}\sum_{\ell=0}^{m-1}\varsigma_{j,\ell}^{(m)}}_{\mathcal{R}_2}.$$

We show now that both $\mathcal{R}_1/n$ and $\mathcal{R}_2/n$ are $o_p(1)$. Because m is fixed, by the boundedness of the Bernoulli variables we have $\mathcal{R}_2/n = o_p(1)$. Now, considering that all the terms with $\ell = 0$ in $\mathcal{R}_{1,n}$ are identically zero, we get:

$$\begin{aligned}
\mathcal{R}_1 &= \sum_{\ell=1}^{m-1}\sum_{j=m}^{n}\left(\varsigma_{j,\ell}^{(m)} - \varsigma_{j-\ell,\ell}^{(m)}\right) = \sum_{\ell=1}^{m-1}\left(\sum_{j=m}^{n}\varsigma_{j,\ell}^{(m)} - \sum_{j=m}^{n}\varsigma_{j-\ell,\ell}^{(m)}\right) = \sum_{\ell=1}^{m-1}\left(\sum_{j=m}^{n}\varsigma_{j,\ell}^{(m)} - \sum_{j=m-\ell}^{n-\ell}\varsigma_{j,\ell}^{(m)}\right) \\
&= \sum_{\ell=1}^{m-1}\left(\underbrace{\sum_{j=n-\ell+1}^{n}\varsigma_{j,\ell}^{(m)}}_{\ell \text{ addends}} - \underbrace{\sum_{j=m-\ell}^{m-1}\varsigma_{j,\ell}^{(m)}}_{\ell\text{addends}}\right).
\end{aligned}$$

Therefore, as for $\mathcal{R}_2$, for given m the number of addends in $\mathcal{R}_1$ is independent of n (and bounded) so that $\mathcal{R}_1/n = o_p(1)$. Thus, by setting $\mathcal{R}_n \doteq \mathcal{R}_1 + \mathcal{R}_2$ the decomposition in (33) hold; that is

$$A_n = \frac{1}{n}\sum_{j=m}^{n}\sum_{\ell=0}^{m-1}\varsigma_{j-\ell,\ell}^{(m)} + o_p(1)$$

To conclude, we have to show that A_n is AN. Before proceeding, for the sake of clarity, we briefly describe how we achieve this result. Let us set $\zeta_j^n = \frac{1}{n}\varsigma_{j-\ell,\ell}^{(m)}$, for fixed ℓ and m. We note that to prove the asymptotic negligibility of A_n, it is sufficient to prove that ζ_j^n is AN. By Lemma 1, this amounts showing that the following two conditions are satisfied

$$\sum_{j=1}^{n}\mathbb{E}_{j-1}\left[\zeta_j^n\right] = \sum_{j=1}^{n}\frac{1}{n}\mathbb{E}_{j-1}\left[\varsigma_{j-\ell,\ell}^{(m)}\right] \xrightarrow{u.c.p} 0 \tag{34}$$

and

$$\sum_{j=1}^{n}\mathbb{E}_{j-1}\left[\left(\zeta_j^n\right)^2\right] \xrightarrow{p} 0. \tag{35}$$

In particular, to prove equation (34) we set $\xi_j^n = n^{-1}\mathbb{E}_{j-1}\left[\varsigma_{j-\ell,\ell}^{(m)}\right]$ and by using Lemma 1 again, we show that

$$\sum_{j=1}^{n}\mathbb{E}_{j-1}\left[\left|\xi_j^n\right|\right] \xrightarrow{p} 0. \tag{36}$$

Therefore, we start from the assertion in (36) and we prove:

$$\begin{aligned}
\sum_{j=1}^{n}\mathbb{E}_{j-1}\left[\left|\xi_j^n\right|\right] &= \sum_{j=1}^{n}\mathbb{E}_{j-1}\left[\left|\frac{1}{n}\mathbb{E}_{j-1}\left[\varsigma_{j-\ell,\ell}^{(m)}\right]\right|\right] = \sum_{j=1}^{n}\frac{1}{n}\left|\mathbb{E}_{j-1}\left[\varsigma_{j-\ell,\ell}^{(m)}\right]\right| \\
&= \sum_{j=1}^{n}\frac{1}{n}\left|\mathbb{E}_{j-1}\left[\mathbb{B}_{j-\ell,n}\cdot\ldots\cdot\mathbb{B}_{j-1,n}\left(p_{j-\ell-1,n}\right)^{m-\ell-1}\left(\mathbb{B}_{j,n}-p_{j-\ell-1,n}\right)\right]\right| \\
&= \sum_{j=1}^{n}\frac{1}{n}\left|\mathbb{B}_{j-\ell,n}\cdot\ldots\cdot\mathbb{B}_{j-1,n}\left(p_{j-\ell-1,n}\right)^{m-\ell-1}\mathbb{E}_{j-1}\left[\left(\mathbb{B}_{j,n}-p_{j-\ell-1,n}\right)\right]\right| \\
&= \sum_{j=1}^{n}\frac{1}{n}\left|\mathbb{B}_{j-\ell,n}\cdot\ldots\cdot\mathbb{B}_{j-1,n}\left(p_{j-1,n}\right)^{m-\ell-1}\mathbb{E}_{j-1}\left[\left(p_{j,n}-p_{j-l-1,n}\right)\right]\right| \\
&\leq \sum_{j=1}^{n}\frac{1}{n}\mathbb{E}_{j-1}\left[\left|p_{j,n}-p_{j-\ell-1,n}\right|\right] \leq \sum_{j=1}^{n}\frac{1}{n}C\Delta_n^{1/2} \leq C\Delta_n^{1/2}.
\end{aligned}$$

At this point, it is enough to prove the convergence in equation (35). This is an easy check because of the boundedness of the Bernoulli variates, that is:

$$\sum_{i=1}^{n}\mathbb{E}_{j-1}\left[\left(\zeta_j^n\right)^2\right] = \frac{1}{n^2}\mathbb{E}_{j-1}\left[\left(\varsigma_{j-\ell,\ell}^{(m)}\right)^2\right] \leq K\Delta_n \longrightarrow 0,$$

which implies the asymptotic negligibility of A_n. Finally, by Riemann integrability:

$$\frac{1}{n}\sum_{j=1}^{n}\left(p_{j-1,n}\right)^m \longrightarrow \int_0^1 \left(p_s\right)^m ds,$$

which completes the proof. □

Before proceeding, we state and prove another useful lemma.

Lemma 4. *Under Assumption 2, for any finite numbers $\ell, d \geq 0$ and powers $q_1, \ldots, q_d \geq 0$, as $n \to \infty$,*

$$\frac{1}{n}\sum_{j=1}^{n}\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j,n}\left(\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\right)^{q_1}\cdots\left(\mathbb{E}_{j-1}\left[\mathbb{B}_{j+d,n}\right]\right)^{q_d} \xrightarrow{p} \int_0^1 p_s^{\ell+v}ds,$$

where $v = q_1 + \ldots + q_d$.

Proof. First, by Remark 30:

$$\frac{1}{n}\sum_{j=1}^{n}\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j,n}\left(\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\right)^{q_1}\cdots\left(\mathbb{E}_{j-1}\left[\mathbb{B}_{j+d,n}\right]\right)^{q_d} = \frac{1}{n}\sum_{j=1}^{n}\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j,n}\,p_{j-1,n}^{v} + O_p\left(\Delta^{1/2}\right).$$

Next, by conditioning on $\mathcal{F}_\infty^{(p)}$ and using the law of iterated expectations:

$$\mathbb{E}\left[\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j,n}\,p_{j-1,n}^{v} - p_{j-\ell,n}\cdots p_{j,n}\,p_{j-1,n}^{v}\right] = 0.$$

Therefore, by Theorem 2.13 in Hall and Heyde (1980)[11] applied to the martingale difference $X_{j,n}^{(\ell)} = \mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j,n}p_{j-1,n}^{v} -$

[11]The hypothesis of the Theorem are readily satisfied because of the boundedness of the Bernoulli random variables with $\mathbb{B}_n = n$.

$p_{j-\ell,n} \cdots p_{j,n} p_{j-1,n}^{v}$:

$$\frac{1}{n}\sum_{j=1}^{n}\left(\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j,n}\,p_{j-1,n}^{v}-p_{j-\ell,n}\cdots p_{j,n}\,p_{j-1,n}^{v}\right)\xrightarrow{p}0.$$

Using Remark (30) again:

$$\frac{1}{n}\sum_{j=1}^{n}p_{j-\ell,n}\dots p_{j,n}\,p_{j-1,n}^{v}=\frac{1}{n}\sum_{j=1}^{n}p_{j-1,n}^{\ell+v}+O_p\left(\Delta^{1/2}\right).$$

Finally, by Riemann integrability we have, path-wise on Ω:

$$\frac{1}{n}\sum_{j=1}^{n}p_{j-1,n}^{\ell+v}\longrightarrow\int_0^1 p_s^{\ell+v}ds,$$

which completes the proof. □

Lemma 5. *Let $m \geq 2$ be a given integer number. Under Assumption 2, as $n \to \infty$:*

$$\sqrt{n}\begin{bmatrix}\mathrm{IT}_n-\int_0^1 p_s\,ds\\ \mathrm{IT}_n^{(m)}-\int_0^1 (p_s)^m\,ds\end{bmatrix}\overset{stably}{\Longrightarrow}\mathcal{MN}\left(0,\Sigma^{(m)}\right), \tag{37}$$

where

$$\mathrm{IT}_n=\frac{1}{n}\sum_{j=1}^{n}\mathbb{B}_{j,n}\qquad \mathrm{IT}_n^{(m)}=\frac{1}{n}\sum_{j=1}^{n}\prod_{i=0}^{m-1}\mathbb{B}_{j+i,n},$$

and $\mathcal{MN}(0,\Sigma^{(m)})$ denotes the mixed-normal distribution with covariance matrix $\Sigma^{(m)}$

$$\Sigma^{(m)}=\begin{bmatrix}\int_0^1 p_s\,(1-p_s)\,ds & \int_0^1 m\,p_s^m\,(1-p_s)\,ds\\ \int_0^1 m\,p_s^m\,(1-p_s)\,ds & \int_0^1 p_s^m\,\frac{p_s^m\,(2\,m+1)-p_s^{m+1}\,(2\,m-1)-(1+p_s)}{1-p_s}\,ds\end{bmatrix}.$$

Proof. We consider the following decomposition:

$$\sqrt{n}\begin{bmatrix}\mathrm{IT}_n-\int_0^1 p_s\,ds\\ \mathrm{IT}_n^{(m)}-\int_0^1 (p_s)^m\,ds\end{bmatrix}=A_1+A_2,$$

where

$$A_1=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\begin{bmatrix}\mathbb{B}_{j,n}-\mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\\ \prod_{i=0}^{m-1}\mathbb{B}_{j+i,n}-\prod_{i=0}^{m-1}\mathbb{E}_{j+i-1}\left[\mathbb{B}_{j+i,n}\right]\end{bmatrix},\quad A_2=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\begin{bmatrix}\mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]-\int_0^1 p_s\,ds\\ \prod_{i=0}^{m-1}\mathbb{E}_{j+i-1}\left[\mathbb{B}_{j+i,n}\right]-\int_0^1 (p_s)^m\,ds\end{bmatrix}.$$

A_2 is AN. Therefore, it is enough to prove that $A_1 \overset{stably}{\Longrightarrow} \mathcal{MN}(0,\Sigma)$. To do so, we rewrite the quantity A_1 as a sum of a $\mathcal{F}_{t_{j,n}}$-measurable quantity and a negligible term. We introduce the following quantity:

$$\zeta_{j,\ell}^{(m)}=\mathbb{B}_{j,n}\,\mathbb{B}_{j+1,n}\cdots\mathbb{B}_{j+\ell-1,n}\,\left(\mathbb{B}_{j+\ell,n}-\mathbb{E}_{j+\ell-1}\left[\mathbb{B}_{j+\ell,n}\right]\right)\,\mathbb{E}_{j+\ell}\left[\mathbb{B}_{j+\ell+1,n}\right]\cdots\mathbb{E}_{j+m-2}\left[\mathbb{B}_{j+m-1,n}\right],$$

and we consider the following expression:

$$\varphi_{j,m}=\prod_{i=0}^{m-1}\mathbb{B}_{j+i,n}-\prod_{i=0}^{m-1}\mathbb{E}_{j+i-1}\left[\mathbb{B}_{j+i,n}\right]$$

for a generic m. Note that $\varphi_{j,m} = \sum_{\ell=0}^{m-1} \zeta_{j,\ell}^{(m)}$. Indeed:

$$\begin{aligned}
\varphi_{j,1} &= \mathbb{B}_{j,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right] \equiv \zeta_{j,0}^{(1)} \\
\varphi_{j,2} &= \mathbb{B}_{j,n}\,\mathbb{B}_{j+1,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\,\mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right] \\
&= \mathbb{B}_{j,n}\,\mathbb{B}_{j+1,n} - \mathbb{B}_{j,n}\,\mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right] + \mathbb{B}_{j,n}\,\mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right] - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\,\mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right] \\
&= \mathbb{B}_{j,n}\,\left(\mathbb{B}_{j+1,n} - \mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right]\right) + \mathbb{B}_{j,n}\,\mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right] - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\,\mathbb{E}_{i}\left[\mathbb{B}_{j+1,n}\right] \\
&= \underbrace{\mathbb{B}_{j,n}\,\left(\mathbb{B}_{j+1,n} - \mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right]\right)}_{\zeta_{j,1}^{(2)}} + \underbrace{\left(\mathbb{B}_{j,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\right)\,\mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right]}_{\zeta_{j,0}^{(2)}} \\
\varphi_{j,3} &= \mathbb{B}_{j,n}\,\mathbb{B}_{j+1,n}\,\mathbb{B}_{j+2,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\,\mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right]\,\mathbb{E}_{j+1}\left[\mathbb{B}_{j+2,n}\right] \\
&= \mathbb{B}_{j,n}\,\mathbb{B}_{j+1,n}\,\mathbb{B}_{j+2,n} - \mathbb{B}_{j,n}\,\mathbb{B}_{j+1,n}\,\mathbb{E}_{j+1}\left[\mathbb{B}_{j+2,n}\right] + \mathbb{B}_{j,n}\,\mathbb{B}_{j+1,n}\,\mathbb{E}_{j+1}\left[\mathbb{B}_{j+2,n}\right] \\
&- \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\,\mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right]\,\mathbb{E}_{j+1}\left[\mathbb{B}_{j+2,n}\right] \\
&= \mathbb{B}_{j,n}\,\mathbb{B}_{j+1,n}\,\left(\mathbb{B}_{j+2,n} - \mathbb{E}_{j+1}\left[\mathbb{B}_{j+2,n}\right]\right) + \\
&\quad + \mathbb{B}_{j,n}\,\mathbb{B}_{j+1,n}\,\mathbb{E}_{j+1}\left[\mathbb{B}_{j+2,n}\right] - \mathbb{B}_{j,n}\,\mathbb{E}_{i}\left[\mathbb{B}_{j+1,n}\right]\,\mathbb{E}_{j+1}\left[\mathbb{B}_{j+2,n}\right] + \\
&\quad + \mathbb{B}_{j,n}\,\mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right]\,\mathbb{E}_{j+1}\left[\mathbb{B}_{j+2,n}\right] - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\,\mathbb{E}_{i}\left[\mathbb{B}_{j+1,n}\right]\,\mathbb{E}_{j+1}\left[\mathbb{B}_{j+2,n}\right] \\
&= \underbrace{\mathbb{B}_{j,n}\,\mathbb{B}_{j+1,n}\,\left(\mathbb{B}_{j+2,n} - \mathbb{E}_{j+1}\left[\mathbb{B}_{j+2,n}\right]\right)}_{\zeta_{i,2}^{(3)}} + \underbrace{\mathbb{B}_{j,n}\,\left(\mathbb{B}_{j+1,n} - \mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right]\right)\,\mathbb{E}_{j+1}\left[\mathbb{B}_{j+2,n}\right]}_{\zeta_{j,1}^{(3)}} + \\
&\quad + \underbrace{\left(\mathbb{B}_{j,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\right)\,\mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right]\,\mathbb{E}_{j+1}\left[\mathbb{B}_{j+2,n}\right]}_{\zeta_{j,0}^{(3)}},
\end{aligned}$$

and so on for every m. Therefore, the second component of A_1, $A_1(2) = n^{-1/2}\sum_{j=1}^{n}\varphi_{j,m}$, can be rewritten as:

$$\begin{aligned}
A_1(2) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\sum_{\ell=0}^{m-1}\zeta_{j,\ell}^{(m)} &= \frac{1}{\sqrt{n}}\sum_{j=m}^{n}\sum_{\ell=0}^{m-1}\zeta_{j,\ell}^{(m)} + \sum_{j=1}^{m-1}\sum_{\ell=0}^{m-1}\zeta_{j,\ell}^{(m)} \\
&= \frac{1}{\sqrt{n}}\sum_{j=m}^{n}\sum_{\ell=0}^{m-1}\zeta_{j-\ell,\ell}^{(m)} + \frac{1}{\sqrt{n}}\underbrace{\sum_{j=m}^{n}\sum_{\ell=0}^{m-1}\left(\zeta_{j,\ell}^{(m)} - \zeta_{j-\ell,\ell}^{(m)}\right)}_{\mathcal{R}_1} + \frac{1}{\sqrt{n}}\underbrace{\sum_{j=1}^{m-1}\sum_{\ell=0}^{m-1}\zeta_{j,\ell}^{(m)}}_{\mathcal{R}_2}.
\end{aligned}$$

Reasoning as in Lemma 3, one can prove that both $\mathcal{R}_1/\sqrt{n}$ and $\mathcal{R}_2/\sqrt{n}$ are $o_p(1)$. To render $A_1(2)$ $\mathcal{F}_{t_{j,n}}$-measurable, a further step is necessary. We define:

$$\widetilde{\zeta}_{j-\ell,\ell}^{(m)} = \mathbb{B}_{j-\ell,n}\,\mathbb{B}_{j-\ell+1,n}\cdots\mathbb{B}_{j-1,n}\,\left(\mathbb{B}_{j,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\right)\,\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\cdots\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-\ell-1,n}\right],$$

and consider:

$$\begin{aligned}
\mathcal{R}_3 &= \sum_{j=m}^{n}\sum_{\ell=0}^{m-1}\left(\zeta_{j-\ell,\ell}^{(m)}-\widetilde{\zeta}_{j-\ell,\ell}^{(m)}\right)\\
&= \sum_{j=m}^{n}\sum_{\ell=0}^{m-1}\mathbb{B}_{j-\ell,n}\,\mathbb{B}_{j-\ell+1,n}\cdots\mathbb{B}_{j-1,n}\left(\mathbb{B}_{j,n}-\mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\right)\times\\
&\times\ \left(\mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right]\cdots\mathbb{E}_{j-\ell+m-2}\left[\mathbb{B}_{j-\ell+m-1,n}\right]-\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\cdots\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-\ell-1,n}\right]\right)\\
&= \sum_{j=m}^{n}\sum_{\ell=0}^{m-1}\mathcal{R}_3(\ell)=\sum_{\ell=0}^{m-1}\sum_{j=m}^{n}\mathcal{R}_3(\ell), \qquad (38)
\end{aligned}$$

where for every $\ell\in\{0,1,\ldots,m-1\}$ we have:

$$\begin{aligned}
\mathcal{R}_3(\ell) &= \sum_{j=m}^{n}\mathbb{B}_{j-\ell,n}\,\mathbb{B}_{j-\ell+1,n}\cdots\mathbb{B}_{j-1,n}\left(\mathbb{B}_{j,n}-\mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\right)\times\\
&\times\ \left(\mathbb{E}_{j}\left[\mathbb{B}_{j+1,n}\right]\cdots\mathbb{E}_{j-\ell+m-2}\left[\mathbb{B}_{j-\ell+m-1,n}\right]-\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\cdots\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-\ell-1,n}\right]\right)\\
&= \sum_{j=m}^{n}r_j(\ell). \qquad (39)
\end{aligned}$$

Using Lemma 2, we show that $\frac{1}{\sqrt{n}}\mathcal{R}_3(\ell)$ are AN $\forall\ell\in\{0,1,\ldots,m-1\}$. Notice that $r_k(\ell)$ is $\mathcal{F}_{t_{j+m-\ell-2,n}}$-measurable. Using the law of iterated expectations, we obtain:

$$\mathbb{E}_{j-1}\left[\frac{1}{\sqrt{n}}r_j(\ell)\right]=0. \qquad (40)$$

Now note that using the triangular inequality and a recursive decomposition for any set of bounded random variables $x_1,\ldots,x_{m-\ell-1}$, $y_1,\ldots,y_{m-\ell-1}$ we obtain (to reduce notation we put $M=m-\ell-1$):

$$\begin{aligned}
|x_1\cdots x_M-y_1\cdots y_M| &= |x_1\cdots x_{M-1}\left(x_M-y_M\right)+\left(x_1\cdots x_{M-1}-y_1\ldots y_{M-1}\right)y_M|\\
&\le |x_1\cdots x_{M-1}\left(x_M-y_M\right)|+|\left(x_1\cdots x_{M-1}-y_1\cdots y_{M-1}\right)y_M|\\
&\le K\ |(x_M-y_M)|+K\ |(x_1\cdots x_{M-1}-y_1\cdots y_{M-1})|\\
&\le \ldots\\
&\le K\sum_{k=1}^{M}|x_k-y_k|\,,
\end{aligned}$$

where the constant K changes from line to line. Applying this inequality to $|r_j(\ell)|$, we obtain:

$$|r_j(\ell)|\le K\sum_{i=1}^{m-\ell-1}|\mathbb{E}_{j-1}\left[\mathbb{B}_{j+i},n\right]-\mathbb{E}_{j+i-1}\left[\mathbb{B}_{j+i},n\right]|\le K\Delta_n^{1/2},$$

where the last inequality follows from Remark 4. Then, because m and l are finite, we have:

$$\sum_{j=m}^{n}\mathbb{E}_{j-1}\left[\left(\frac{1}{\sqrt{n}}r_j(\ell)\right)^2\right]\le K\sum_{j=m}^{n}\frac{1}{n}\mathbb{E}_{j-1}\left[\left(\sum_{i=1}^{m-\ell-1}|\mathbb{E}_{j-1}\left[\mathbb{B}_{j+i},n\right]-\mathbb{E}_{j+i-1}\left[\mathbb{B}_{j+i},n\right]|\right)^2\right]\to 0. \qquad (41)$$

Therefore, by Lemma 2, $\frac{1}{\sqrt{n}}\mathcal{R}_3(\ell)$ are AN $\forall \ell \in \{0,1,\ldots,m-1\}$, which implies that $\mathcal{R}_3$ AN as well.

Now, decompose A_1 as:

$$A_1 = \frac{1}{\sqrt{n}}\sum_{j=m}^{n}\eta_j + \frac{1}{\sqrt{n}}R_n = \frac{1}{\sqrt{n}}\sum_{j=m}^{n}\begin{bmatrix}\eta_j(1)\\ \eta_j(2)\end{bmatrix} + \frac{1}{\sqrt{n}}\begin{bmatrix}R_n(1)\\ R_n(2)\end{bmatrix},$$

with

$$\eta_j(1) \doteq \mathbb{B}_{j,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right], \quad \eta_j(2) \doteq \sum_{\ell=0}^{m-1}\widetilde{\zeta}^{(m)}_{j-\ell,\ell},$$

and where the reminders are given by:

$$R_n(1) = \sum_{j=1}^{m-1}\left(\mathbb{B}_{j\Delta,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j\Delta,n}\right]\right), \quad R_n(2) = \mathcal{R}_1 + \mathcal{R}_2 + \mathcal{R}_3.$$

Since the first component of R_n consists of a finite number of bounded terms and the second component of R_n is the sum of AN terms, $R_n/\sqrt{n}$ is AN. Therefore, it is enough to establish the following convergence:

$$\frac{1}{\sqrt{n}}\sum_{j=m}^{n}\eta_j \overset{stably}{\Longrightarrow} \mathcal{MN}\left(0,\Sigma_{\text{MIT}}\right).$$

To establish the previous convergence, we use Corollary 2. We have to find two functions $\varphi^{(1)}$ and $\varphi^{(2)}$ such that:

$$\eta_j\left(1\right) = \varphi^{(1)}\left(\mathbb{B}_{j-m+1,n},\ldots,\mathbb{B}_{j,n},\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right],\ldots,\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-1,n}\right]\right) - \mathbb{E}_{j-1}\left[\varphi^{(1)}\left(\mathbb{B}_{j-m+1,n},\ldots,\mathbb{B}_{j,n},\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right],\ldots,\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-1,n}\right]\right)\right]$$

and the same is necessary for $\eta_j\left(2\right)$. The case of $\eta_j\left(1\right)$ is trivial because it is enough to define $\varphi^{(1)}\left(x_1\right) \doteq x_1$ to have the identity $\eta_j\left(1\right) = \varphi^{(1)}\left(\mathbb{B}_j\right) - \mathbb{E}_{j-1}\left[\varphi^{(1)}\left(\mathbb{B}_j\right)\right]$. Concerning, $\eta_j\left(2\right)$ note that:

$$\begin{aligned}
\eta_j\left(2\right) &= \sum_{\ell=0}^{m-1}\widetilde{\zeta}^{(m)}_{j-\ell,\ell} = \sum_{\ell=0}^{m-1}\mathbb{B}_{j-\ell,n}\,\mathbb{B}_{j-\ell+1,n}\cdots\mathbb{B}_{j-1,n}\,\left(\mathbb{B}_{j,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\right)\,\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\cdots\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-\ell-1,n}\right]\\
&= \left(\mathbb{B}_{j,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\right)\,\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\cdots\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-1,n}\right] +\\
&\quad +\mathbb{B}_{j-1,n}\,\left(\mathbb{B}_{j,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\right)\,\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\cdots\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-2,n}\right] + \ldots\\
&\quad +\mathbb{B}_{j-m+1,n}\,\mathbb{B}_{j-m+2,n}\cdots\mathbb{B}_{j-1,n}\,\left(\mathbb{B}_{j,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\right)\\
&= \mathbb{B}_{j,n}\,\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\cdots\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-1,n}\right] + \mathbb{B}_{j-1,n}\,\mathbb{B}_{j,n}\,\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\cdots\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-2,n}\right] + \ldots\\
&\quad +\mathbb{B}_{j-m+1,n}\,\mathbb{B}_{j-m+2,n}\cdots\mathbb{B}_{j-1,n}\,\mathbb{B}_{j,n} - \left(\mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\,\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\cdots\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-1,n}\right]\right) + \ldots\\
&\quad +\mathbb{B}_{j-m+1,n}\,\mathbb{B}_{j-m+2,n}\cdots\mathbb{B}_{j-1,n}\,\mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\\
&= \varphi^{(2)}\left(\mathbb{B}_{j-m+1,n},\ldots,\mathbb{B}_{j,n},\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right],\ldots,\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-1,n}\right]\right)\\
&- \mathbb{E}_{j-1}\left[\varphi^{(2)}\left(\mathbb{B}_{j-m+1,n},\ldots,\mathbb{B}_{j,n},\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right],\ldots,\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-1,n}\right]\right)\right]
\end{aligned}$$

where $\varphi^{(2)}:\mathbb{R}^{2\,(m-1)+1} \to \mathbb{R}$ takes the following form:

$$\varphi^{(2)}\left(x_1,\cdots,x_m,\cdots,x_{2(m-1)+1}\right) \doteq x_m\,x_{m+1}\cdots x_{2(m-1)+1} + x_{m-1}\,x_m\cdots x_{2(m-1)} + \ldots + x_1\,x_2\cdots x_m.$$

We now proceed by noticing that for all j the vector η_j is $\mathcal{F}_{t_{j,n}}$-measurable and bounded, where:

$$\sum_{j=m}^{n} \mathbb{E}_{j-1}\left[\left\|\frac{1}{\sqrt{n}}\eta_j\right\|^4\right] \xrightarrow{p} 0,$$

and $\mathbb{E}_{j-1}[\eta_j(1)] = 0$. To also see that $\mathbb{E}_{j-1}[\eta_j(2)] = 0$, it is better to write down $\mathbb{E}_{j-1}[\eta_j(2)]$ explicitly:

$$\begin{aligned}
&\mathbb{E}_{j-1}[\eta_j(2)] = \sum_{\ell=0}^{m-1} \mathbb{E}_{j-1}\left[\widetilde{\zeta}^{(m)}_{j-\ell,\ell}\right] \\
= \;& \sum_{\ell=0}^{m-1} \mathbb{B}_{j-\ell,n}\,\mathbb{B}_{j-\ell+1,n}\cdots\mathbb{B}_{j-1,n}\,\underbrace{\mathbb{E}_{j-1}\left[(\mathbb{B}_{j,n}-\mathbb{E}_{j-1}[\mathbb{B}_{j,n}])\right]}_{=0}\,\mathbb{E}_{j-1}[\mathbb{B}_{j+1,n}]\,\mathbb{E}_{j-1}[\mathbb{B}_{j+2,n}]\cdots\mathbb{E}_{j-1}[\mathbb{B}_{j-\ell+m-1,n}].
\end{aligned}$$

Consequently, it is enough to show that $n^{-1}\sum_{i=m}^{n}\mathbb{E}_{j-1}\left[\eta_j\eta_j'\right] \xrightarrow{p} \Sigma$. Consider each component of the matrix $\eta_j\eta_j'$ separately.

$$\eta_j(1)\eta_j(1) = \mathbb{B}_{j,n} - 2\mathbb{B}_{j,n}\mathbb{E}_{j-1}[\mathbb{B}_{j,n}] + (\mathbb{E}_{j-1}[\mathbb{B}_{j,n}])^2.$$

By Lemma 4,

$$\frac{1}{n}\sum_{i=m}^{n}\mathbb{E}_{i-1}[\eta_i(1)\eta_i(1)] \xrightarrow{p} \int_0^1 \left(p_s - p_s^2\right) ds.$$

Now consider the product:

$$\eta_j(2)\,\eta_j(2) = \sum_{\ell=0}^{m-1}\left(\widetilde{\zeta}^{(m)}_{j-\ell,\ell}\right)^2 + 2\sum_{\ell=0}^{m-1}\sum_{\ell'=\ell+1}^{m-1}\widetilde{\zeta}^{(m)}_{j-\ell,\ell}\,\widetilde{\zeta}^{(m)}_{j-\ell',\ell'} = \sum_{\ell=0}^{m-1}\left(\widetilde{\zeta}^{(m)}_{j-\ell,\ell}\right)^2 + 2\sum_{\ell=0}^{m-1}\sum_{k=1}^{m-\ell-1}\widetilde{\zeta}^{(m)}_{j-\ell,\ell}\,\widetilde{\zeta}^{(m)}_{j-\ell-k,\ell+k}.$$

We note that :

$$\left(\widetilde{\zeta}^{(m)}_{j-\ell,\ell}\right)^2 = \underbrace{\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j-1,n}}_{\ell\text{ factors}}\,(\mathbb{B}_{j,n}-\mathbb{E}_{j-1}[\mathbb{B}_{j,n}])^2\,\underbrace{(\mathbb{E}_{j-1}[\mathbb{B}_{j+1,n}]\cdots\mathbb{E}_{j-1}[\mathbb{B}_{j+m-\ell-1,n}])^2}_{m-\ell-1\text{ factors}}$$

and

$$\begin{aligned}
&\widetilde{\zeta}^{(m)}_{j-\ell,\ell}\,\widetilde{\zeta}^{(m)}_{j-\ell-k,\ell+k} \\
= \;& \mathbb{B}_{j-\ell,n}\,\mathbb{B}_{j-\ell+1,n}\cdots\mathbb{B}_{j-1,n}\,(\mathbb{B}_{j,n}-\mathbb{E}_{j-1}[\mathbb{B}_{j,n}])\,\mathbb{E}_{j-1}[\mathbb{B}_{j+1,n}]\cdots\mathbb{E}_{j-1}[\mathbb{B}_{j+m-\ell-1,n}]\times \\
& \times\mathbb{B}_{j-\ell-k,n}\,\mathbb{B}_{j-\ell-k+1,n}\cdots\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j-1,n}\,(\mathbb{B}_{j,n}-\mathbb{E}_{j-1}[\mathbb{B}_{j,n}])\,\mathbb{E}_{j-1}[\mathbb{B}_{j+1,n}]\cdots\mathbb{E}_{j-1}[\mathbb{B}_{j+m-\ell-k-1,n}] \\
= \;& \underbrace{\mathbb{B}_{j-\ell-k,n}...\mathbb{B}_{j-1,n}}_{\ell+k\text{ factors}}\,(\mathbb{B}_{j,n}-\mathbb{E}_{j-1}[\mathbb{B}_{j,n}])^2\times \\
& \underbrace{(\mathbb{E}_{j-1}[\mathbb{B}_{j+1,n}]\cdots\mathbb{E}_{j-1}[\mathbb{B}_{j+m-\ell-k-1,n}])^2}_{m-(\ell+k)-1\text{ factors}}\,\underbrace{\mathbb{E}_{j-1}[\mathbb{B}_{j+m-\ell-k,n}]\cdots\mathbb{E}_{j-1}[\mathbb{B}_{j+m-\ell-1,n}]}_{k\text{ factors}}.
\end{aligned}$$

Consequently, using Lemma 4:

$$\frac{1}{n}\sum_{j=m}^{n}\mathbb{E}_{j-1}[\eta_j(2)\eta_j(2)] \xrightarrow{p} \Sigma_{22} \doteq \int_0^1\left(\sum_{\ell=0}^{m-1}p_s^{2m-\ell-1}(1-p_s) + 2\sum_{\ell=0}^{m-1}(m-\ell-1)\,p_s^{2m-\ell-1}(1-p_s)\right) ds,$$

which, after some standard algebra becomes:

$$\begin{aligned}\Sigma_{22} &= \int_0^1 p_s^{2m-1}\,(1-p_s)\left(\sum_{\ell=0}^{m-1} p_s^{-\ell} + 2\sum_{\ell=0}^{m-1}(m-\ell-1)\,p_s^{-\ell}\right) ds \\ &= \int_0^1 \frac{p_s^m\,(1+p_s-(2m(1-p_s)+1+p_s)p_s^m)}{1-p_s}\,ds \\ &= \int_0^1 p_s^m\,\frac{p_s^m\,(2\,m+1)-p_s^{m+1}\,(2\,m-1)-(1+p_s)}{1-p_s}\,ds. \end{aligned} \tag{42}$$

Finally:

$$\begin{aligned}\eta_j(1)\eta_j(2) &= (\mathbb{B}_{j,n}-\mathbb{E}_{j-1}[\mathbb{B}_{j,n}])^2\,\mathbb{E}_{j-1}[\mathbb{B}_{j+1,n}]\ldots\mathbb{E}_{j-1}[\mathbb{B}_{j+m-1,n}] \\ &+ \mathbb{B}_{j-1,n}\,(\mathbb{B}_{j,n}-\mathbb{E}_{j-1}[\mathbb{B}_{j,n}])^2\,\mathbb{E}_{j-1}[\mathbb{B}_{j+1,n}]\ldots\mathbb{E}_{j-1}[\mathbb{B}_{j+m-2,n}] \\ &+ \ldots \\ &+ \mathbb{B}_{j-m-1,n}\ldots\mathbb{B}_{j-1,n}\,(\mathbb{B}_{j,n}-\mathbb{E}_{j-1}[\mathbb{B}_{j,n}])^2. \end{aligned}$$

Applying Lemma 4 again:

$$\frac{1}{n}\sum_{j=m}^{n}\mathbb{E}_{j-1}[\eta_j(1)\eta_j(2)] \xrightarrow{p} \int_0^1 m\,p_s^m\,(1-p_s)\,ds,$$

which completes the proof. □

A.3 Proofs of Theorems 3.3 and 3.4 from Section 3.2

For an arbitrary sequence of integers k_n such that $k_n \to \infty$ and $k_n\Delta_n = \frac{k_n}{n} \to 0$, let:

$$\alpha_j^n \doteq \frac{1}{k_n}\sum_{i=0}^{k_n-1}(\mathbb{B}_{j+i,n}-p_{j+i,n}), \qquad \beta_j^n \doteq \frac{1}{k_n}\sum_{i=0}^{k_n-1}(p_{j+i,n}-p_{j-1,n}),$$

and set $h_n = n - k_n$. Note that:

$$\widehat{p}_j(k_n) - p_{j-1,n} = \alpha_j^n + \beta_j^n, \quad j \in \{1,\ldots,h_n+1\}.$$

The auxiliary results for the proofs of Theorems 3.3 and 3.4 are summarized by the following Lemma.

Lemma 6. *Under Assumptions 1, 2, and 3, for $C > 0$ and for all $q \geq 2$, we have:*

$$\mathbb{E}_{j-1}\left[\sup_{s\in[0,\Delta_n]}|p_{j-1+s,n}-p_{j-1,n}|^q\right] \leq C\cdot\Delta_n^{1\wedge(q/2)} \tag{43}$$

$$\left|\mathbb{E}_{j-1}\left[p_{j,n}-p_{j-1,n}\right]\right| \leq C\cdot\Delta_n \tag{44}$$

$$\left|\mathbb{E}_{j-1}\left[\beta_j^n\right]\right| \leq C\cdot k_n\Delta_n \tag{45}$$

$$\mathbb{E}_{j-1}\left[\left|\beta_j^n\right|^q\right] \leq C\cdot(k_n\Delta_n)^{q/2} \tag{46}$$

$$\left|\mathbb{E}_{j-1}\left[\alpha_j^n\right]\right| = 0 \tag{47}$$

$$\mathbb{E}_{j-1}\left[\left|\alpha_j^n\right|^q\right] \leq Ck_n^{-q/2} \tag{48}$$

$$\left|\mathbb{E}_{j-1}\left[(\alpha_i^n)^2-\frac{1}{k_n}p_{i-1,n}\left(1-p_{j-1,n}\right)\right]\right| \leq C\cdot\Delta_n \tag{49}$$

$$\left|\mathbb{E}_{j-1}\left[\alpha_j^n\beta_j^n\right]\right| = 0 \tag{50}$$

Proof. The proof of (43)-(47) follows the same arguments as in the proof of the results of Appendix A and Lemma B-4 of Aït-Sahalia and Jacod (2012). To complete the proof of the Lemma, we need to prove (47)-(50). Equality (47) easily follows by conditioning on the path of the process p_t.

$$\left|\mathbb{E}_{j-1}\left[\alpha_j^n\right]\right| = \left|\frac{1}{k_n}\sum_{j=0}^{k_n-1}\mathbb{E}_{j-1}\left[\mathbb{B}_{j+i,n}-p_{j+i,n}\right]\right| = 0.$$

To prove the other relationships, we first observe that conditioning on the path $(p_t)_{t\in[0,1]}$ we have:

$$\begin{aligned}\mathbb{E}_{i-1}\left[\left(\alpha_j^n\right)^2\right] &= \frac{1}{k_n^2}\mathbb{E}_{j-1}\left[\sum_{i=0}^{k_n-1}\left(\mathbb{B}_{j+i,n}-p_{j+i,n}\right)^2\right]+\frac{2}{k_n}\mathbb{E}_{j-1}\left[\sum_{i=0}^{k_n-2}\sum_{m=1}^{k_n-1-i}\left(\mathbb{B}_{j+i,n}-p_{j+i,n}\right)\left(\mathbb{B}_{j+i+m,n}-p_{j+i+m,n}\right)\right]\\ &= \frac{1}{k_n^2}\sum_{i=0}^{k_n-1}\mathbb{E}_{j-1}\left[\left(\mathbb{B}_{j+i,n}-p_{j+i,n}\right)^2\right] = \frac{1}{k_n^2}\sum_{i=0}^{k_n-1}\mathbb{E}_{j-1}\left[p_{j+i,n}\left(1-p_{j+i,n}\right)\right] \leq \frac{C}{k_n},\end{aligned} \tag{51}$$

where the last inequality is due to the fact that $p_t \in (0,1)$. Moreover, we have:

$$\mathbb{E}_{j-1}\left[\left(\alpha_j^n\right)^2-\frac{1}{k_n}p_{j-1,n}(1-p_{j-1,n})\right] = \frac{1}{k_n^2}\sum_{i=0}^{k_n-1}\mathbb{E}_{j-1}\left[p_{j+i,n}-p_{j-1,n}\right]-\frac{1}{k_n^2}\sum_{i=0}^{k_n-1}\mathbb{E}_{j-1}\left[p_{j+i,n}^2-p_{j-1,n}^2\right].$$

By applying triangular inequality, we obtain:

$$\left|\mathbb{E}_{j-1}\left[\left(\alpha_j^n\right)^2-\frac{1}{k_n}p_{j-1,n}(1-p_{j-1,n})\right]\right| \leq \frac{1}{k_n^2}\sum_{i=0}^{k_n-1}\left|\mathbb{E}_{j-1}\left[p_{j+i,n}-p_{j-1,n}\right]\right|+\frac{1}{k_n^2}\sum_{i=0}^{k_n-1}\left|\mathbb{E}_{j-1}\left[p_{j+i,n}^2-p_{j-1,n}^2\right]\right|.$$

Therefore, (49) follows from (44), whereas (48) follows from Hölder's inequality and (51). Finally, (50) is obtained by conditioning on the path $(p_t)_{t\in[0,1]}$ and by using equation (47). □

Proof of Theorem 3.3. For any $t > 0$, define a function of t, $\widehat{p}(k_n, t)$, as:

$$\widehat{p}(k_n,t) \doteq \widehat{p}_j(k_n), \quad t \in ((j-2)\Delta_n, (j-1)\Delta_n].$$

First, we prove that $\widehat{p}(k_n, t)$ converges in probability to p_t for every $t \in [0, 1]$. For any $t \in [0, 1]$ and j_t such that

$t \in ((j_t-2)\Delta_n, (j_t-1)\Delta_n]$, we have:

$$(j+1)\Delta_n \leq (j_t+j)\Delta_n - t \leq (j+2)\Delta_n.$$

Second, we have:

$$\begin{aligned}
\mathbb{E}\left[(\widehat{p}(k_n,t)-p_t)^2\right] &= \mathbb{E}\left[\left(\frac{1}{k_n}\sum_{i=0}^{k_n-1}(\mathbb{B}_{j_t+i,n}-p_t)\right)^2\right] = \mathbb{E}\left[\frac{1}{k_n^2}\sum_{i=0}^{k_n-1}(\mathbb{B}_{j_t+i,n}-p_t)^2 + \frac{1}{k_n^2}\sum_{i\neq i'}(\mathbb{B}_{j_t+i,n}-p_t)\,(\mathbb{B}_{j_t+i',n}-p_t)\right] \\
&= \mathbb{E}\left[\frac{1}{k_n^2}\sum_{i=0}^{k_n-1}(\mathbb{B}_{j_t+i,n}-p_t)^2\right] + \mathbb{E}\left[\frac{1}{k_n^2}\sum_{i\neq i'}(\mathbb{B}_{j_t+i,n}-p_t)\,(\mathbb{B}_{j_t+i',n}-p_t)\right].
\end{aligned}$$

The first of the two terms converges to zero by the boundedness of $\mathbb{B}_{j_t+i,n}$ and p_t. Concerning the second by conditioning on $(p_t)_{t\in[0,1]}$ and (44) we have that

$$\left|\mathbb{E}\left[(\mathbb{B}_{j_t+i,n}-p_t)\,(\mathbb{B}_{j_t+i',n}-p_t)\right]\right| = \left|\mathbb{E}\left[p_{(j_t+i)\Delta_n}-p_t\right]\mathbb{E}\left[p_{(j_t+i')\Delta_n}-p_t\right]\right| \leq C(k_n\Delta_n)^2.$$

Therefore,

$$\left|\mathbb{E}\left[\frac{1}{k_n^2}\sum_{j\neq j'}(\mathbb{B}_{i_t+j,n}-p_t)\,(\mathbb{B}_{i_t+j',n}-p_t)\right]\right| \leq C(k_n\Delta_n)^2 \longrightarrow 0.$$

Thus, $\widehat{p}(k_n,t) \xrightarrow{p} p_t$ for each $t\in[0,1]$. Now, we write $U(\Delta_n,f)^n$ as:

$$U(\Delta_n,f)^n = \Delta_n f(\widehat{p}_1(k_n)) + \int_0^{h_n\Delta_n} f(\widehat{p}(k_n,t))\,ds.$$

and we compute:

$$\begin{aligned}
\mathbb{E}\left[\left|U(\Delta_n,f)^n - \int_0^1 f(p_s)\,ds\right|\right] &= \Delta_n\,\mathbb{E}\left[\left|f(\widehat{p}_1(k_n)) - \int_0^1 f(p_s)\,ds\right|\right] + \int_0^{h_n\Delta_n} a_s\,ds \\
&= \Delta_n\,\mathbb{E}\left[\left|\int_0^1 \left(f(\widehat{p}_1(k_n)) - f(p_s)\right)\,ds\right|\right] + \int_0^{h_n\Delta_n} a_s\,ds \\
&\leq \Delta_n\,\mathbb{E}\left[\int_0^1 \left|\left(f\left(\widehat{p}_1\left(k_n\right)\right) - f(p_s)\right)\right|\,ds\right] + \int_0^{h_n\Delta_n} a_s\,ds \\
&\leq C\,\Delta_n + \int_0^{h_n\Delta_n} a_n(s)\,ds,
\end{aligned}$$

where $a_n(s) \doteq \mathbb{E}\left[|f(\widehat{p}(k_n,s)) - f(p_s)|\right]$, C is a suitable constant, and we used the locally boundedness of $f(\cdot)$ and the boundedness of p_s and $\widehat{p}(k_n,s)$. By the continuous mapping theorem, condition $\widehat{p}(k_n,t) \xrightarrow{p} p_t$ implies that for a given $s\in[0,1]$:

$$f(\widehat{p}(k_n,s)) \xrightarrow{p} f(p_s). \tag{52}$$

Nonetheless, because the sequence of random variables $f(\widehat{p}(k_n,s))$ is uniformly integrable (again using the locally boundedness of $f(\cdot)$ and the boundedness of $\widehat{p}(k_n,s)$) the convergence in equation (52) is also in $\mathbb{L}^1$ norm and therefore $a_n(s) \longrightarrow 0$ for each s. In addition, because $a_n(s)$ is uniformly bounded in (n,s), $U(\Delta_n,f)^n \stackrel{u.c.p.}{\longrightarrow} \int_0^1 f(p_s)\,ds$ by the dominated convergence theorem (cfr. Jacod and Protter, 2012, Theorem 9.4.1). □

Proof of Theorem 3.4. First, consider the following decomposition:

$$\begin{aligned}\frac{1}{\sqrt{\Delta_n}}\left(U'\left(\Delta_n,f\right)^n-U\left(f\right)\right) &= \sqrt{\Delta_n}\sum_{j=1}^{h_n+1}\left(f\left(\widehat{p}_j\left(k_n\right)\right)-\frac{1}{2\,k_n}f''\left(\widehat{p}_j\left(k_n\right)\right)\widehat{p}_j\left(k_n\right)\left(1-\widehat{p}_j\left(k_n\right)\right)\right)-\frac{1}{\sqrt{\Delta_n}}\int_0^1 f\left(p_s\right)\,ds \\ &= \sum_{r=1}^{4}U\left(r\right)^n,\end{aligned}$$

with

$$\begin{aligned} U\left(1\right)^n &= \frac{1}{\sqrt{\Delta_n}}\sum_{j=1}^{h_n+1}\int_{(j-1)\Delta_n}^{j\Delta_n}\left(f\left(p_{j-1,n}\right)-f\left(p_s\right)\right)\,ds-\frac{1}{\sqrt{\Delta_n}}\int_{(h_n+1)\,\Delta_n}^{1}f\left(p_s\right)\,ds \\ U\left(2\right)^n &= \sqrt{\Delta_n}\sum_{j=1}^{h_n+1}f'\left(p_{j-1,n}\right)\beta_j^n \\ U\left(3\right)^n &= \sqrt{\Delta_n}\sum_{j=1}^{h_n+1}\left(f\left(\widehat{p}_j\left(k_n\right)\right)-f\left(p_{j-1,n}\right)-f'\left(p_{j-1,n}\right)\left(\alpha_j^n+\beta_j^n\right)-\frac{1}{2\,k_n}f''\left(\widehat{p}_j\left(k_n\right)\right)\widehat{p}_j\left(k_n\right)\left(1-\widehat{p}_j\left(k_n\right)\right)\right) \\ U\left(4\right)^n &= \sqrt{\Delta_n}\sum_{j=1}^{h_n+1}f'\left(p_{j-1,n}\right)\alpha_j^n.\end{aligned}$$

At this point, the rest of the proof is divided into four parts. In the first three, we prove that $U\left(k\right)^n$, $k=1,2,3$, is AN, whereas in the last part we show that $U(4)^n \overset{stably}{\Longrightarrow} \mathcal{MN}(0,\Sigma)$.

Part 1: Proof of the AN of $U\left(1\right)^n$

Remember that $h_n=n-k_n$ and that $n=1/\Delta_n$, where $1-\left(h_n+1\right)\,\Delta_n=1-\left(n-k_n+1\right)\,\Delta_n=k_n\,\Delta_n-\Delta_n$. Because $f(p_s)$ is bounded, for the second term of $U(1)^n$ we have:

$$\left|\frac{1}{\sqrt{\Delta_n}}\int\limits_{(h_n+1)\,\Delta_n}^{1}f\left(p_s\right)\,ds\right|\leq Ck_n\sqrt{\Delta_n}\longrightarrow 0.$$

The first term of $U(1)_1^n$ can be expressed as $\sum\limits_{j=1}^{h_n+1}\xi_j^n$, where:

$$\xi_j^n=\frac{1}{\sqrt{\Delta_n}}\int_{(j-1)\Delta_n}^{j\Delta_n}\left(f\left(p_{j-1,n}\right)-f\left(p_s\right)\right)\,ds.$$

Because the process $f(p_t)$ is a bounded semimartingale, by using inequality (44) we get:

$$\begin{aligned}\left|\mathbb{E}\left[\xi_j^n\right]\right| &= \frac{1}{\sqrt{\Delta_n}}\left|\mathbb{E}\left[\int_{(j-1)\Delta_n}^{j\Delta_n}\left(f\left(p_{j-1,n}\right)-f\left(p_s\right)\right)\,ds\right]\right|=\frac{1}{\sqrt{\Delta_n}}\left|\int_{(j-1)\Delta_n}^{j\Delta_n}\mathbb{E}\left[\mathbb{E}_{j-1}\left[\left(f\left(p_{j-1,n}\right)-f\left(p_s\right)\right)\right]\,ds\right]\right| \\ &\leq \frac{1}{\sqrt{\Delta_n}}\int_{(j-1)\Delta_n}^{j\Delta_n}\left|\mathbb{E}\left[\mathbb{E}_{j-1}\left[\left(f\left(p_{j-1,n}\right)-f\left(p_s\right)\right)\right]\right]\right|ds\leq\frac{1}{\sqrt{\Delta_n}}\int_{(j-1)\Delta_n}^{j\Delta_n}\mathbb{E}\left[\left|\mathbb{E}_{j-1}\left[\left(f\left(p_{j-1,n}\right)-f\left(p_s\right)\right)\right]\right|\right]ds \\ &\leq \frac{C}{\sqrt{\Delta_n}}\Delta_n^2=C\left(\Delta_n\right)^{3/2}\longrightarrow 0,\end{aligned}$$

while using inequality (43) and Holder's inequality, we obtain:

$$\begin{aligned}
\mathbb{E}\left[\left|\xi_j^n\right|^2\right] &= \frac{1}{\Delta_n}\mathbb{E}\left[\left(\int_{(j-1)\Delta_n}^{j\Delta_n} \left(f\left(p_{j-1,n}\right) - f\left(p_s\right)\right) ds\right)^2\right] \\
&= \frac{1}{\Delta_n}\mathbb{E}\left[\int\limits_{(j-1)\Delta_n}^{j\Delta_n}\int\limits_{(j-1)\Delta_n}^{j\Delta_n} \left(f(p_{j-1,n}) - f(p_q)\right)\left(f(p_{j-1,n}) - f(p_s)\right) ds\, dq\right] \\
&= \frac{1}{\Delta_n}\int\limits_{(j-1)\Delta_n}^{j\Delta_n}\int\limits_{(j-1)\Delta_n}^{j\Delta_n} \mathbb{E}\left[\left(f(p_{j-1,n}) - f(p_q)\right)\left(f(p_{j-1,n}) - f(p_s)\right)\right] ds\, dq \\
&\leq \frac{1}{\Delta_n}\int\limits_{(j-1)\Delta_n}^{j\Delta_n}\int\limits_{(j-1)\Delta_n}^{j\Delta_n} \sqrt{\mathbb{E}\left[\left|f(p_{j-1,n}) - f(p_q)\right|^2\right]\mathbb{E}\left[\left|f(p_{j-1,n}) - f(p_s)\right|^2\right]}\, ds\, dq \\
&\leq \frac{1}{\Delta_n}\int\limits_{(j-1)\Delta_n}^{j\Delta_n}\int\limits_{(j-1)\Delta_n}^{j\Delta_n} C\,\Delta_n\, ds\, dq \leq C\,\Delta_n^2 \longrightarrow 0.
\end{aligned}$$

Consequently, by Lemma 2, $U(1)^n$ is AN.

Part 2: Proof of the AN of $U\left(2\right)^n$

Using Lemma 6 and the boundedness of $f'\left(p_{j-1,n}\right)$, we obtain:

$$\sum_{j=1}^{h_n+1}\left|\mathbb{E}_{j-1}\left[\sqrt{\Delta_n}f'\left(p_{j-1,n}\right)\beta_j^n\right]\right| \leq C\sum_{j=1}^{h_n+1}\sqrt{\Delta_n}\left|\mathbb{E}_{j-1}\left[\beta_j^n\right]\right| \leq C\sum_{j=1}^{h_n+1}k_n\left(\Delta_n\right)^{3/2} \longrightarrow 0$$

and

$$\sum_{j=1}^{h_n+1}\mathbb{E}_{j-1}\left[\left|\sqrt{\Delta_n}f'\left(p_{j-1,n}\right)\beta_j^n\right|^2\right] \leq C\sum_{j=1}^{h_n+1}\mathbb{E}_{j-1}\left[\Delta_n\left|\beta_j^n\right|^2\right] \leq C\sum_{j=1}^{h_n+1}k_n\left(\Delta_n\right)^2 = C\,\left(n-k_n\right)\,k_n\,\Delta_n^2 \leq C\,k_n\,\Delta_n \longrightarrow 0,$$

and so

$$k_n\sum_{j=1}^{h_n+1}\mathbb{E}\left[\left|\sqrt{\Delta_n}f'\left(p_{j-1,n}\right)\beta_j^n\right|^2\right] \leq C\,k_n^2\,\Delta_n \longrightarrow 0.$$

Consequently, by applying Lemma 2 we get that $U(2)^n$ is AN.

Part 3: Proof of the AN of $U\left(3\right)^n$

As a first step, we rewrite $U\left(3\right)^n$ as $U(3)^n = \sum\limits_{j=1}^{h_n+1}\sum\limits_{k=1}^{4} v_j^n(k)$ with $v_n^j\left(k\right)$, $k = 1, \ldots, 4$, suitably defined triangular arrays. To do so, we remind readers that:

$$\alpha_j^n + \beta_j^n = \frac{1}{k_n}\sum_{i=0}^{k_n-1}\left(\mathbb{B}_{j+i,n} - p_{j-1,n}\right) = \widehat{p}_j\left(k_n\right) - p_{j-1,n}.$$

Using Taylor expansion of $f\left(p\right)$ around $p_0 = p_{j-1,n}$ and computing the expansion in $p = \widehat{p}_j\left(k_n\right)$, we obtain:

$$f\left(\widehat{p}_j\left(k_n\right)\right) - f\left(p_{j-1,n}\right) - f'\left(p_{j-1,n}\right)\left(\alpha_j^n + \beta_j^n\right) = \frac{1}{2}f''\left(p_{j-1,n}\right)\left(\alpha_j^n + \beta_j^n\right)^2 + \frac{1}{6}f'''\left(p_j^\star\right)\left(\alpha_j^n + \beta_j^n\right)^3,$$

where $p_j^\star$ is a point between $p_{j-1,n}$ and $p_{j-1,n} + \alpha_j^n + \beta_j^n$. We then have:

$$\begin{aligned}\frac{1}{2} f''(p_{j-1,n}) \left(\alpha_j^n + \beta_j^n\right)^2 &= \frac{1}{2} f''(p_{j-1,n}) \left((\alpha_j^n)^2 + 2\alpha_j^n \beta_j^n - \frac{1}{k_n} p_{j-1,n}(1-p_{j-1,n}) \right) \\ &+ \frac{1}{2\,k_n} f''(p_{j-1,n})\, p_{j-1,n}(1-p_{j-1,n}) + \frac{1}{2} f''(p_{j-1,n}) \left(\beta_j^n\right)^2.\end{aligned}$$

Consequently, $U(3)^n$ can be represented as $U(3)^n = \sum_{j=1}^{h_n+1} \sum_{k=1}^{4} v_j^n(k)$, where:

$$\begin{aligned}v_j^n(1) &= \frac{\sqrt{\Delta_n}}{2} f''(p_{j-1,n}) \left((\alpha_j^n)^2 + 2\alpha_j^n \beta_j^n - \frac{1}{k_n} p_{j-1,n}(1-p_{j-1,n}) \right), \\ v_j^n(2) &= \frac{\sqrt{\Delta_n}}{2\,k_n} f''(p_{j-1,n})\, p_{j-1,n}(1-p_{j-1,n}) - \frac{\sqrt{\Delta_n}}{2\,k_n} f''(\widehat{p}_j(k_n))\, \widehat{p}_j(k_n)\,(1-\widehat{p}_j(k_n)), \\ v_j^n(3) &= \frac{\sqrt{\Delta_n}}{2} f''(p_{j-1,n}) \left(\beta_j^n\right)^2, \\ v_j^n(4) &= \frac{\sqrt{\Delta_n}}{6} f'''(p_j^\star) \left(\alpha_j^n + \beta_j^n\right)^3.\end{aligned}$$

We have to prove that all the triangular arrays $v_j^n(k)$ are AN for $k = 1, 2, 3, 4$. First, consider $v_i^n(1)$. Inequalities (49) and (50) imply that $\left|\mathbb{E}_{j-1}\left[v_j^n(1)\right]\right| \leq C\,\Delta_n^{3/2}$, and so:

$$\sum_{j=1}^{h_n+1} \left|\mathbb{E}_{j-1}\left[v_j^n(1)\right]\right| \leq C\,\Delta_n^{1/2} \xrightarrow{p} 0. \tag{53}$$

In addition,

$$\begin{aligned}v_j^n(1)^2 &= \frac{\Delta_n}{4} f''(p_{j-1,n})^2 \Big((\alpha_j^n)^4 + 4\left(\alpha_j^n \beta_j^n\right)^2 + \frac{1}{k_n^2} p_{j-1,n}^2 (1-p_{j-1,n})^2 + \\ &\quad + 4\left(\alpha_j^n\right)^3 \beta_j^n - 2\,\frac{(\alpha_j^n)^2}{k_n} p_{j-1,n}(1-p_{j-1,n}) - \frac{4\,\alpha_j^n\,\beta_j^n}{k_n} p_{j-1,n}(1-p_{j-1,n}) \Big) \\ &\leq \frac{\Delta_n}{4} f''(p_{j-1,n})^2 \Big((\alpha_j^n)^4 + 4\left(\alpha_j^n \beta_j^n\right)^2 + \frac{1}{k_n^2} p_{j-1,n}^2 (1-p_{j-1,n})^2 + \\ &\quad + 4\left|\left(\alpha_j^n\right)^3 \beta_j^n\right| + 2\,\frac{(\alpha_j^n)^2}{k_n} p_{j-1,n}(1-p_{j-1,n}) - \frac{4\,\alpha_j^n\,\beta_j^n}{k_n} p_{j-1,n}(1-p_{j-1,n}) \Big).\end{aligned}$$

Now, in computing $\mathbb{E}\left[v_j^n(1)^2\right]$ we consider that:

- Inequality (48) implies that:

$$\mathbb{E}_{j-1}\left[(\alpha_j^n)^4\right] \leq C\,k_n^{-2},$$

  and that

$$\mathbb{E}_{j-1}\left[\frac{(\alpha_j^n)^2}{k_n} p_{j-1,n}(1-p_{j-1,n})\right] \leq C\,k_n^{-2}.$$

- Cauchy-Schwartz inequality plus (48) and (46) imply that:

$$\mathbb{E}_{j-1}\left[\left(\alpha_j^n \beta_j^n\right)^2\right] \leq \left(\mathbb{E}_{j-1}\left[\left(\alpha_j^n\right)^4\right]\right)^{1/2} \left(\mathbb{E}_{j-1}\left[\left(\beta_j^n\right)^4\right]\right)^{1/2} \leq C\,\Delta_n$$

  and that

$$\left|\mathbb{E}_{j-1}\left[\left(\alpha_j^n\right)^3 \beta_j^n\right]\right| \leq \left(\mathbb{E}_{j-1}\left[\left(\alpha_j^n\right)^6\right]\right)^{1/2} \left(\mathbb{E}_{j-1}\left[\left(\beta_j^n\right)^2\right]\right)^{1/2} \leq C\,k_n^{-1}\,\Delta_n^{1/2}.$$

- Equation (50) implies $\mathbb{E}_{j-1}\left[\alpha_j^n\,\beta_j^n k_n p_{j-1,n}(1-p_{j-1,n})\right] = 0$.

Summing up:

$$\mathbb{E}_{j-1}\left[v_j^n(1)^2\right] \leq C\,\Delta_n\left(\frac{1}{k_n^2}+\Delta_n+\frac{\sqrt{\Delta_n}}{k_n}\right)$$

where:

$$k_n\sum_{j=1}^{h_n}\mathbb{E}\left[v_j^n(1)^2\right] \longrightarrow 0. \tag{54}$$

Summing up, the limits in (53) and (54) imply, through Lemma 2, that $v_j^n(1)$ is AN. Now, consider $v_j^n(4)$. Because both p_t and $\widehat{p}_i(k_n)$ are in $[0,1]$, $|f'''(p_i^\star)| \leq C$ for some constant $C>0$ we therefore have:

$$\sum_{j=1}^{h_n+1}\left|\frac{\sqrt{\Delta_n}}{6}f'''\left(p_j^\star\right)\left(\alpha_j^n+\beta_j^n\right)^3\right| \leq C\sum_{j=1}^{h_n+1}\sqrt{\Delta_n}\left|\left(\alpha_j^n+\beta_j^n\right)^3\right| = C\sum_{j=1}^{h_n+1}\sqrt{\Delta_n}\left(\left|\alpha_j^n\right|^3+3\left|\alpha_j^n\right|\left|\beta_j^n\right|^2+3\left|\alpha_j^n\right|^2\left|\beta_j^n\right|+\left|\beta_j^n\right|^3\right).$$

Using estimates from the preliminary results and Cauchy-Schwartz inequality, we have the following implications.

- inequality (48) implies:

$$\sum_{j=1}^{h_n+1}\sqrt{\Delta_n}\mathbb{E}_{j-1}\left[\left|\alpha_j^n\right|^3\right] \leq C\cdot k_n^{-3/2}(\Delta_n)^{-1/2} \xrightarrow{p} 0,$$

- inequalities (48) and (46), and Cauchy-Schwartz, imply:

$$\sum_{j=1}^{h_n+1}\sqrt{\Delta_n}\mathbb{E}_{j-1}\left[\left|\alpha_j^n\right|^2\left|\beta_j^n\right|\right] \leq C\sum_{j=1}^{h_n+1}\sqrt{\Delta_n}\sqrt{\mathbb{E}_{j-1}\left[\left|\alpha_j^n\right|^4\right]\mathbb{E}_{j-1}\left[\left|\beta_j^n\right|^2\right]} \leq C\,k_n^{-1/2} \xrightarrow{p} 0$$

and

$$\sum_{j=1}^{h_n+1}\sqrt{\Delta_n}\mathbb{E}_{j-1}\left[\left|\alpha_j^n\right|\left|\beta_j^n\right|^2\right] \leq C\sum_{j=1}^{h_n+1}\sqrt{\Delta_n}\sqrt{\mathbb{E}_{j-1}\left[\left|\alpha_j^n\right|^2\right]\mathbb{E}_{j-1}\left[\left|\beta_j^n\right|^4\right]} \leq C\cdot(k_n\,\Delta_n)^{1/2} \xrightarrow{p} 0.$$

- Inequality (46) implies:

$$\sum_{j=1}^{h_n+1}\sqrt{\Delta_n}\mathbb{E}_{j-1}\left[\left|\beta_j^n\right|^3\right] \leq C\cdot k_n^{3/2}\Delta_n \xrightarrow{p} 0.$$

Therefore:

$$\sum_{j=1}^{h_n+1}\left|\mathbb{E}_{j-1}\left[v_j^n(4)\right]\right| \xrightarrow{p} 0. \tag{55}$$

Now consider:

$$\begin{aligned}
\sum_{j=1}^{h_n+1}v_j^n(4)^2 &\leq C\sum_{j=1}^{h_n+1}\Delta_n\Big(\left|\alpha_j^n\right|^6+9\left|\alpha_j^n\right|^2\left|\beta_j^n\right|^4+9\left|\alpha_j^n\right|^4\left|\beta_j^n\right|^2+\left|\beta_j^n\right|^6+6\left|\alpha_j^n\right|^4\left|\beta_j^n\right|^2+\\
&+6\left|\alpha_j^n\right|^5\left|\beta_j^n\right|+2\left|\alpha_j^n\right|^3\left|\beta_j^n\right|^3+18\left|\alpha_j^n\right|^3\left|\beta_j^n\right|^3+6\left|\alpha_j^n\right|\left|\beta_j^n\right|^5+6\left|\alpha_j^n\right|^2\left|\beta_j^n\right|^4\Big).\\
&= C\sum_{j=1}^{h_n+1}\Delta_n\Big(\left|\alpha_j^n\right|^6+15\left|\alpha_j^n\right|^2\left|\beta_j^n\right|^4+15\left|\alpha_j^n\right|^4\left|\beta_j^n\right|^2+\left|\beta_j^n\right|^6+6\left|\alpha_j^n\right|^5\left|\beta_j^n\right|+20\left|\alpha_j^n\right|^3\left|\beta_j^n\right|^3+6\left|\alpha_j^n\right|\left|\beta_j^n\right|^5\Big).
\end{aligned}$$

inequalities (48) and (46), respectively, imply:

$$k_n\sum_{j=1}^{h_n+1}\Delta_n\mathbb{E}\left[\left|\alpha_j^n\right|^6\right] \leq C\,k_n^{-2} \longrightarrow 0,$$

$$k_n \sum_{j=1}^{h_n+1} \Delta_n \mathbb{E}\left[\left|\beta_j^n\right|^6\right] \leq C\,\left(k_n^{4/3}\,\Delta_n\right)^3 \longrightarrow 0,$$

and, using also Cauchy-Schwartz, they imply:

$$k_n \sum_{j=1}^{h_n+1} \Delta_n \mathbb{E}\left[\left|\alpha_j^n\right|^2\,\left|\beta_j^n\right|^4\right] \leq C\,\left(k_n\,\Delta_n\right)^2 \longrightarrow 0$$
$$k_n \sum_{j=1}^{h_n+1} \Delta_n \mathbb{E}\left[\left|\alpha_j^n\right|^4\,\left|\beta_j^n\right|^2\right] \leq C\,k_n^{-2}\,\Delta_n \longrightarrow 0$$
$$k_n \sum_{j=1}^{h_n+1} \Delta_n \mathbb{E}\left[\left|\alpha_j^n\right|^5\,\left|\beta_j^n\right|\right] \leq C\,k_n^{-1}\,\Delta_n^{1/2} \longrightarrow 0$$
$$k_n \sum_{j=1}^{h_n+1} \Delta_n \mathbb{E}\left[\left|\alpha_j^n\right|\,\left|\beta_j^n\right|^5\right] \leq C\,\left(k_n^{6/5}\,\Delta_n\right)^{5/2} \longrightarrow 0$$
$$k_n \sum_{j=1}^{h_n+1} \Delta_n \mathbb{E}\left[\left|\alpha_j^n\right|^3\,\left|\beta_j^n\right|^3\right] \leq C\,\left(k_n^{2/3}\,\Delta_n\right)^{3/2} \longrightarrow 0.$$

Consequently:

$$k_n \sum_{j=1}^{h_n+1} \mathbb{E}\left[v_j^n\,(4)^2\right] \longrightarrow 0. \tag{56}$$

As before, the limits in (55) and (56) imply, through Lemma 2, that $v_j^n(4)$ is AN. Similarly, for $v_j^n(3)$ we have:

$$\sum_{j=1}^{h_n+1} \mathbb{E}_{j-1}\left[\left|\frac{\sqrt{\Delta_n}}{2} f''\left(p_{j-1,n}\right)\left(\beta_j^n\right)^2\right|\right] \leq C \cdot k_n \sqrt{\Delta_n} \xrightarrow{p} 0, \tag{57}$$

In addition:

$$k_n \sum_{j=1}^{h_n+1} \mathbb{E}\left[\left|\frac{\Delta_n}{4}\left(f''\left(p_{j-1,n}\right)\right)^2\left(\beta_j^n\right)^4\right|\right] \leq C \cdot \left(k_n^{3/2}\,\Delta_n\right)^2 \longrightarrow 0, \tag{58}$$

Therefore, the limits in (57) and (58) imply, through Lemma 2, that $v_j^n(3)$ is AN. Finally, consider $v_j^n(2)$. Using Taylor's expansion, we have (remember that $\widehat{p}_j(k_n) - p_{j-1,n} = \alpha_j^n + \beta_j^n$):

$$f''\left(\widehat{p}_j(k_n)\right) = f''\left(p_{j-1,n}\right) + f'''\left(p_j^\star\right)\left(\alpha_j^n + \beta_j^n\right).$$

Consequently, $v_j^n(2)$ takes the form:

$$\begin{aligned}
v_j^n(2) &= \frac{\sqrt{\Delta_n}}{2\,k_n} f''\left(p_{j-1,n}\right) p_{(j-1)\Delta_n}(1-p_{(j-1)\Delta_n}) - \frac{\sqrt{\Delta_n}}{2\,k_n} f''\left(\widehat{p}_j\left(k_n\right)\right) \widehat{p}_j\left(k_n\right)\left(1-\widehat{p}_j\left(k_n\right)\right) \\
&= \frac{\sqrt{\Delta_n}}{2\,k_n} f''\left(p_{j-1,n}\right) p_{(j-1)\Delta_n}(1-p_{(j-1)\Delta_n}) - \frac{\sqrt{\Delta_n}}{2\,k_n}\left(f''\left(p_{j-1,n}\right) + f'''\left(p_j^\star\right)\left(\alpha_j^n+\beta_j^n\right)\right) \widehat{p}_j\left(k_n\right)\left(1-\widehat{p}_j\left(k_n\right)\right) \\
&= \frac{\sqrt{\Delta_n}}{2\,k_n} f''\left(p_{j-1,n}\right) p_{(j-1)\Delta_n} - \frac{\sqrt{\Delta_n}}{2\,k_n} f''\left(p_{j-1,n}\right) p^2_{(j-1)\Delta_n} \\
&- \frac{\sqrt{\Delta_n}}{2\,k_n} f''\left(p_{j-1,n}\right) \widehat{p}_j\left(k_n\right) + \frac{\sqrt{\Delta_n}}{2\,k_n} f''\left(p_{j-1,n}\right) \widehat{p}_j\left(k_n\right)^2 - \frac{\sqrt{\Delta_n}}{2\,k_n} f'''\left(p_j^\star\right)\left(\alpha_j^n+\beta_j^n\right) \widehat{p}_j\left(k_n\right)\left(1-\widehat{p}_j\left(k_n\right)\right) \\
&= -\frac{\sqrt{\Delta_n}}{2\,k_n} f''\left(p_{j-1,n}\right)\left(\widehat{p}_j(k_n) - p_{j-1,n}\right) + \frac{\sqrt{\Delta_n}}{2\,k_n} f''\left(p_{j-1,n}\right)\left(\widehat{p}_j(k_n)^2 - p^2_{j-1,n}\right) \\
&- \frac{\sqrt{\Delta_n}}{2\,k_n} f'''\left(p_j^\star\right)\left(\alpha_j^n+\beta_j^n\right) \widehat{p}_j\left(k_n\right)\left(1-\widehat{p}_j\left(k_n\right)\right) \\
&= -\underbrace{\frac{\sqrt{\Delta_n}}{2\,k_n} f''\left(p_{j-1,n}\right)\left(\alpha_j^n+\beta_j^n\right)}_{\mathcal{A}_{j,n}} + \underbrace{\frac{\sqrt{\Delta_n}}{2\,k_n} f''\left(p_{j-1,n}\right)\left(\widehat{p}_j(k_n)^2 - p^2_{j-1,n}\right)}_{\mathcal{B}_{j,n}} \\
&- \underbrace{\frac{\sqrt{\Delta_n}}{2\,k_n} f'''\left(p_j^\star\right)\left(\alpha_j^n+\beta_j^n\right) \widehat{p}_j\left(k_n\right)\left(1-\widehat{p}_j\left(k_n\right)\right)}_{\mathcal{C}_{j,n}}.
\end{aligned}$$

Using Lemma 6, we have:

$$\begin{aligned}
&\sum_{j=1}^{h_n}\left|\mathbb{E}_{j-1}\left[\frac{\sqrt{\Delta_n}}{k_n} f''\left(p_{(j-1)\Delta_n}\right)\alpha_j^n\right]\right| = 0, \\
&k_n \sum_{j=1}^{h_n}\mathbb{E}\left[\frac{\Delta_n}{k_n^2}\left(f''\left(p_{(j-1)\Delta_n}\right)\right)^2\left|\alpha_j^n\right|^2\right] \le C\,k_n^{-2}, \\
&\sum_{j=1}^{h_n}\left|\mathbb{E}_{j-1}\left[\frac{\sqrt{\Delta_n}}{k_n} f''\left(p_{(j-1)\Delta_n}\right)\beta_j^n\right]\right| \le C\,\Delta_n^{1/2}, \\
&k_n \sum_{j=1}^{h_n}\mathbb{E}\left[\frac{\Delta_n}{k_n^2}\left(f''\left(p_{(j-1)\Delta_n}\right)\right)^2\left|\beta_j^n\right|^2\right] \le C\,\Delta_n,
\end{aligned}$$

which imply, through Lemma 2, that $\mathcal{A}_{j,n}$ is AN. Now because

$$\mathcal{B}_{j,n} = \frac{\sqrt{\Delta_n}}{2\,k_n} f''\left(p_{j-1,n}\right)\left(\alpha_j^n+\beta_j^n\right)\left(\widehat{p}_j(k_n)+p_{j-1,n}\right) = \mathcal{A}_{j,n}\ \left(\widehat{p}_j(k_n)+p_{j-1,n}\right)$$

and being $(\widehat{p}_j(k_n)+p_{j-1,n})$ bounded, we can apply to $\mathcal{B}_{j,n}$ the same reasoning used for $\mathcal{A}_{j,n}$, ; therefore, $\mathcal{B}_{j,n}$ is AN. An identical reasoning applies to $\mathcal{C}_{j,n}$, which is then AN as well.

Part 4: Proof of the convergence $U^n(4) \overset{stably}{\Longrightarrow} \mathcal{MN}(0,\Sigma)$
Recall that $U(4)^n$ is defined as:

$$U(4)^n = \frac{\sqrt{\Delta_n}}{k_n}\sum_{j=1}^{h_n+1} f'\left(p_{j-1,n}\right)\sum_{i=0}^{k_n-1}\mathbb{B}_{j+i,n}.$$

For the sake of readability define, , we temporarily define the variables:

$$a_{j-1} = f'\left(p_{j-1,n}\right), \mathbb{B}_{j+i} = \mathbb{B}_{j+i,n} - p_{j+i,n}$$

so that:

$$U(4)^n = \frac{\sqrt{\Delta_n}}{k_n} \sum_{j=1}^{n-k_n+1} a_{j-1} \sum_{i=0}^{k_n-1} \mathbb{B}_{j+i}.$$

The convolution of summation in $U(4)^n$ can be rewritten as:

$$\begin{aligned}
\sum_{j=1}^{n-k_n+1} a_{j-1} \sum_{i=0}^{k_n-1} \mathbb{B}_{j+i} &= a_0 \left(\mathbb{B}_1 + \mathbb{B}_2 + \cdots + \mathbb{B}_{k_n}\right) + a_1 \left(\mathbb{B}_2 + \mathbb{B}_3 + \cdots + \mathbb{B}_{k_n+1}\right) + \cdots \\
\cdots & + a_{k_n-1} \left(\mathbb{B}_{k_n} + \mathbb{B}_{k_n+1} + \cdots + \mathbb{B}_{2k_n-1}\right) + a_{k_n} \left(\mathbb{B}_{k_n+1} + \mathbb{B}_{k_n+2} + \cdots + \mathbb{B}_{2k_n}\right) + \cdots \\
\cdots & + a_{n-k_n-1} \left(\mathbb{B}_{n-k_n} + \mathbb{B}_{n-k_n+1} + \cdots + \mathbb{B}_{n-1}\right) + a_{n-k_n} \left(\mathbb{B}_{n-k_n+1} + \mathbb{B}_{n-k_n+1} + \cdots + \mathbb{B}_n\right) \\
&= \mathbb{B}_1 a_0 + \mathbb{B}_2 \left(a_0 + a_1\right) + \mathbb{B}_3 \left(a_0 + a_1 + a_2\right) + \cdots + \mathbb{B}_{k_n} \left(a_0 + a_1 + a_2 + \ldots + a_{k_n-1}\right) \\
&+ \mathbb{B}_{k_n+1} \left(a_1 + a_2 + a_3 + \ldots + a_{k_n}\right) + \mathbb{B}_{k_n+2} \left(a_2 + a_3 + a_4 + \ldots + a_{k_n+1}\right) + \cdots \\
&+ \mathbb{B}_{n-k_n+1} \left(a_{n-2k_n+1} + a_{n-2k_n+1} + \ldots + a_{n-k_n}\right) \\
&+ \mathbb{B}_{n-k_n+2} \left(a_{n-2k_n+2} + a_{n-2k_n+3} + \ldots + a_{n-k_n}\right) + \cdots + \mathbb{B}_{n-1} \left(a_{n-k_n-1} + a_{n-k_n}\right) + \mathbb{B}_n a_{n-k_n} \\
&= \sum_{j=1}^{k_n} \mathbb{B}_j \sum_{i=0}^{j-1} a_i + \sum_{j=k_n+1}^{n-k_n+1} \mathbb{B}_j \sum_{i=j-k}^{j-1} a_i + \sum_{j=n-k_n+1}^{n} \mathbb{B}_j \sum_{i=j-k_n}^{n-k_n} a_i \\
(i \to j-i-1) &= \sum_{j=1}^{k_n} \mathbb{B}_j \sum_{j=0}^{i-1} a_{j-i-1} + \sum_{j=k_n+1}^{n-k_n+1} \mathbb{B}_j \sum_{i=0}^{k_n-1} a_{j-i-1} + \sum_{j=n-k_n+1}^{n} \mathbb{B}_j \sum_{i=j-n+k_n-1}^{k_n-1} a_{j-i-1} \\
&= \sum_{j=1}^{n} \sum_{i=j-n+k_n-1 \vee 0}^{(j-1)\wedge(k_n-1)} a_{j-i-1} \mathbb{B}_j.
\end{aligned}$$

Hence,

$$\begin{aligned}
U(4)^n &= \sqrt{\Delta_n} \sum_{j=1}^{n} \frac{1}{k_n} \sum_{i=j-n+k_n-1 \vee 0}^{(j-1)\wedge(k_n-1)} f'\left(p_{j-i-1,n}\right) \left(\mathbb{B}_{j,n} - p_{j\Delta_n}\right) \\
&= \sqrt{\Delta_n} \sum_{j=1}^{n} \left(\left(\frac{1}{k_n} \sum_{i=j-n+k_n-1 \vee 0}^{(j-1)\wedge(k_n-1)} f'\left(p_{j-i-1,n}\right) \right) - f'\left(p_{j-1,n}\right) + f'\left(p_{j-1,n}\right) \right) \left(\mathbb{B}_{j,n} - p_{j\Delta_n}\right) \\
&= \sqrt{\Delta_n} \sum_{j=1}^{n} f'\left(p_{j-1,n}\right) \left(\mathbb{B}_{j,n} - p_{j\Delta_n}\right) + \sqrt{\Delta_n} \sum_{j=1}^{n} w_j^n \left(\mathbb{B}_{j,n} - p_{j\Delta_n}\right),
\end{aligned}$$

where

$$w_j^n = \frac{1}{k_n} \sum_{i=j-n+k_n-1 \vee 0}^{(j-1)\wedge(k_n-1)} f'\left(p_{j-i-1,n}\right) - f'\left(p_{j-1,n}\right).$$

By conditioning on $(p_t)_{t\in[0,1]}$, $\mathbb{E}\left[w_j^n \left(\mathbb{B}_{j,n} - p_{j\Delta_n}\right)\right] = 0$. Next, by the assumption about the derivative of f,

$$\left|w_j^n\right| \le C \sup_{s \in [(j-1)\Delta_n, (j+k_n-1)\Delta_n]} \left|p_s - p_{j-1,n}\right|.$$

Hence, inequality (43) implies that $\mathbb{E}\left[\left|w_j^n\right|^2\right] \le C\sqrt{\Delta_n}$ when $k_n \le j \le \lfloor 1/\Delta_n \rfloor - k_n$ and $\left|w_j^n\right| \le C$ always. Therefore, since both $\mathbb{B}_{j,n}$ and p_t are bounded,

$$\mathbb{E}_{j-1}\left[\left|\sqrt{\Delta_n} w_j^n \left(\mathbb{B}_{j,n} - p_{j\Delta_n}\right)\right|^2\right] \le \begin{cases} C\Delta_n^{3/2} & k_n \le j \le h_n, \\ C\Delta_n & \text{otherwise.} \end{cases}$$

Consequently, $\sum_{j=1}^{\lfloor 1/\Delta_n \rfloor} \mathbb{E}_{j-1}\left[\left|\sqrt{\Delta_n} w_j^n\left(\mathbb{B}_{j,n}-p_{j\Delta_n}\right)\right|^2\right] \longrightarrow 0$, which by Lemma 2 implies that $\sqrt{\Delta_n}\sum_{j=1}^{\lfloor 1/\Delta_n \rfloor} w_j^n\left(\mathbb{B}_{j,n}-p_{j\Delta_n}\right)$ is AN. Now, set $\xi_j^n = \sqrt{\Delta_n} f'\left(p_{j-1,n}\right)\left(\mathbb{B}_{j,n}-p_{j\Delta_n}\right)$. Clearly, $\mathbb{E}\left[\xi_j^n\right]=0$, and we have:

$$\mathbb{E}_{j-1}\left[\left(\xi_j^n\right)^2\right] = \Delta_n\left(f'\left(p_{j-1,n}\right)\right)^2 \mathbb{E}_{j-1}\left[p_{j\Delta_n} - \left(p_{j\Delta_n}\right)^2\right].$$

Because, $\left(f'\left(p_{j-1,n}\right)\right)^2$ is bounded, using (44) we have:

$$\left|\mathbb{E}_{j-1}\left[\left(\xi_j^n\right)^2\right] - \Delta_n\left(f'\left(p_{j-1,n}\right)\right)^2\left(p_{j-1,n} - \left(p_{j-1,n)^2}\right)\right| \leq C(\Delta_n)^2.$$

Therefore,

$$\sum_{j=1}^{\lfloor 1/\Delta_n \rfloor} \mathbb{E}_{j-1}\left[\left(\xi_j^n\right)^2\right] \xrightarrow{P} \int_0^1 f'\left(p_s\right)^2 p_s(1-p_s)\, ds.$$

Consequently:

$$\sum_{j=1}^{\lfloor 1/\Delta_n \rfloor} \xi_j^n \overset{stably}{\Longrightarrow} \mathcal{MN}(0,\Sigma),$$

which completes the proof. □

A.4 Proof of Theorem 3.5 from Section 3.4

For any process X, denote the increments by $\Delta_j^n X = X_{(j+1)\,\Delta_n} - X_{j\,\Delta_n}$. Set $k_n = \theta\left\lfloor\sqrt{n}\right\rfloor$ and define:

$$\widetilde{\nu}_n = \sum_{i=1}^{n-2k_n+1}\left(\widehat{p}_{i+k_n}\left(k_n\right) - \widehat{p}_i\left(k_n\right)\right)^2.$$

We then have to prove that, as $n \to \infty$:

$$k_n^{-1}\,\widetilde{\nu}_n \xrightarrow{p} \frac{2}{3}\int_0^1 \nu_s^2\, ds + \frac{2}{\theta^2}\int_0^1 p_s\left(1-p_s\right)\, ds.$$

We have:

$$\widehat{p}_j\left(k_n\right) = \frac{1}{k_n}\sum_{i=0}^{k_n-1}\left(\mathbb{B}_{j+i,n} - p_{j+i,n}\right) + \frac{1}{k_n}\sum_{i=0}^{k_n-1} p_{j+i,n}.$$

Consequently, the difference between $\widehat{p}_{j+k_n}\left(k_n\right)$ and $\widehat{p}_j\left(k_n\right)$ can be expressed as:

$$\widehat{p}_{j+k_n}\left(k_n\right) - \widehat{p}_j\left(k_n\right) = \frac{1}{k_n}\sum_{i=0}^{2k_n-1}\epsilon(1)_i\left(\mathbb{B}_{j+i,n} - p_{j+i,n}\right) + \frac{1}{k_n}\sum_{i=0}^{k_n-1}\left(p_{j+i+k_n,n} - p_{j+i,n}\right), \qquad (59)$$

where, for $m \in \{0,\ldots,2k_n-1\}$:

$$\epsilon(1)_m = \begin{cases} -1, & 0 \leq m < k_n, \\ +1, & k_n \leq m < 2k_n. \end{cases}$$

Then, using telescopic sums, notice that:

$$\left(p_{j+i+k_n,n} - p_{j+i,n}\right) = \sum_{\ell=0}^{k_n-1}\Delta_{j+i+\ell,n}p.$$

Now note that the sum $S_{j,n} = \sum_{i=0}^{k_n-1} (p_{j+i+k_n,n} - p_{j+i,n})$, collecting identical terms, becomes:

$$\begin{aligned}
S_{j,n} &= \Delta_j^n p + \Delta_{j+1}^n p + \Delta_{j+2}^n p + \ldots + \Delta_{j+k_n-1}^n p \\
&\quad + \Delta_{j+1}^n p + \Delta_{j+2}^n p + \ldots + \Delta_{j+k_n-1}^n p + \Delta_{j+k_n}^n p \\
&\quad\quad + \Delta_{j+2}^n p + \ldots + \Delta_{j+k_n-1}^n p + \Delta_{j+k_n}^n p + \Delta_{j+k_n+1}^n p \\
&\quad\quad\quad \vdots \\
&\quad\quad\quad + \Delta_{j+k_n-1}^n p + \Delta_{j+k_n}^n p + \Delta_{j+k_n+1}^n p + \ldots + \Delta_{j+2\,k_n-2}^n p \\
&= \underbrace{\Delta_j^n p + 2\,\Delta_{j+1}^n p + 3\,\Delta_{j+2}^n p + \ldots + k_n\,\Delta_{j+k_n-1}^n p}_{k_n \text{ terms}} + \underbrace{(k_n-1)\,\Delta_{j+k_n}^n p + \ldots + \Delta_{j+2\,k_n-2}^n p}_{k_n-1 \text{ terms}},
\end{aligned}$$

which can be rewritten as:

$$\frac{1}{k_n}\sum_{i=0}^{k_n-1} (p_{j+i+k_n,n} - p_{j+i,n}) = \frac{1}{k_n}\sum_{i=0}^{2k_n-1} \epsilon(2)_i\, (p_{j+i+1,n} - p_{j+i,n}),$$

where, for $i \in \{0, \ldots, 2k_n - 1\}$

$$\epsilon(2)_i = (i+1) \wedge (2k_n - i - 1),$$

and, in particular, $\epsilon(2)_{2k_n-1} = 0$. Now expression (59) become:

$$\widehat{p}_{j+k_n}(k_n) - \widehat{p}_j(k_n) = \frac{1}{k_n}\sum_{j=0}^{2k_n-1} \left(\epsilon(2)_i\,(\mathbb{B}_{i+j,n} - p_{j+i,n}) + \epsilon(2)_i\,(p_{j+i+1,n} - p_{j+i,n})\right).$$

Therefore:

$$\begin{aligned}
\left(\widehat{p}_{j+k_n}(k_n) - \widehat{p}_j(k_n)\right)^2 &= \frac{1}{k_n^2}\sum_{i=0}^{2k_n-1}\Bigg(\epsilon(2)_i^2\,(\mathbb{B}_{j+i,n} - p_{j+i,n})^2 + \epsilon(2)_i^2\,(p_{j+i+1,n} - p_{j+i,n})^2 \\
&+ 2\,\epsilon(2)_i\,\epsilon(2)_i\,(\mathbb{B}_{j+i,n} - p_{j+i,n})\,(p_{j+i+1,n} - p_{j+i,n})\Bigg) \\
&+ 2\sum_{j=0}^{2\,k_n-2}\sum_{\ell'=j+1}^{2\,k_n-1}\Bigg(\epsilon(2)_i\epsilon(1)_\ell\,(\mathbb{B}_{j+i,n} - p_{j+i,n})\,(\mathbb{B}_{j+\ell,n} - p_{j+\ell,n}) \\
&+ \epsilon(2)_i\,\epsilon(2)_\ell\,(\mathbb{B}_{j+i,n} - p_{j+i,n})\,(p_{j+\ell+1,n} - p_{j+\ell,n}) \\
&+ \epsilon(1)_\ell\,\epsilon(2)_i\,(\mathbb{B}_{j+\ell,n} - p_{j+\ell,n})\,(p_{j+i+1,n} - p_{j+i,n}) \\
&+ \epsilon(2)_i\,\epsilon(2)_\ell\,(p_{j+i+1,n} - p_{j+i,n})\,(p_{j+\ell+1,n} - p_{j+\ell,n})\Bigg).
\end{aligned} \tag{60}$$

So, setting;

$$\zeta(1)_j = \mathbb{B}_{j,n} - p_{j,n},\ \zeta(2)_j = p_{j+1,n} - p_{j,n},$$

we have the following more compact expression:

$$\left(\widehat{p}_{j+k_n}(k_n) - \widehat{p}_j(k_n)\right)^2 = \frac{1}{k_n^2}\sum_{u,v=1}^{2}\left(\sum_{i=0}^{2k_n-1}\epsilon(u)_i\epsilon(v)_i\zeta(u)_{j+i}\zeta(v)_{j+i} + 2\sum_{i=0}^{2k_n-2}\sum_{l=j+1}^{2k_n-1}\epsilon(u)_i\epsilon(v)_l\zeta(u)_{j+i}\zeta(v)_{i+l}\right).$$

Consequently, $\widetilde{\nu}_n$ can be expressed as:

$$\widetilde{\nu}_n = \sum_{s=1}^{7}\sum_{i=1}^{n-2k_n+1} v_i^n(s),$$

where

$$v_i^n(1) = \frac{1}{k_n^2} \sum_{i=0}^{2k_n-1} \left(\mathbb{B}_{j+i,n} - p_{j+i,n}\right)^2, \quad v_i^n(2) = \frac{1}{k_n^2} \sum_{i=0}^{2k_n-1} \epsilon(2)_i^2 \left(p_{j+i+1,n} - p_{j+i,n}\right)^2,$$

$$v_i^n(3) = \frac{2}{k_n^2} \sum_{i=0}^{2k_n-1} \epsilon(1)_i \epsilon(2)_i \left(\mathbb{B}_{j+i,n} - p_{j+i,n}\right) \left(p_{j+i+1,n} - p_{j+i,n}\right),$$

$$v_i^n(4) = \frac{2}{k_n^2} \sum_{i=0}^{2k_n-2} \sum_{l=i+1}^{2k_n-1} \epsilon(1)_i \epsilon(1)_l \left(\mathbb{B}_{j+i,n} - p_{j+i,n}\right) \left(\mathbb{B}_{j+l,n} - p_{j+l,n}\right),$$

$$v_i^n(5) = \frac{2}{k_n^2} \sum_{i=0}^{2k_n-2} \sum_{l=j+1}^{2k_n-1} \epsilon(2)_i \epsilon(2)_l \left(p_{j+i+1,n} - p_{j+i,n}\right) \left(p_{j+l+1,n} - p_{j+l,n}\right),$$

$$v_i^n(6) = \frac{2}{k_n^2} \sum_{i=0}^{2k_n-2} \sum_{l=j+1}^{2k_n-1} \epsilon(1)_i \epsilon(2)_l \left(\mathbb{B}_{j+i,n} - p_{j+i,n}\right) \left(p_{j+l+1,n} - p_{j+l,n}\right),$$

$$v_i^n(7) = \frac{2}{k_n^2} \sum_{i=0}^{2k_n-2} \sum_{l=j+1}^{2k_n-1} \epsilon(2)_i \epsilon(1)_l \left(p_{j+i+1,n} - p_{j+i,n}\right) \left(\mathbb{B}_{j+l+1,n} - p_{j+l,n}\right).$$

Consequently, to study the convergence of $\widetilde{\nu}_n$ in probability, we need to study the convergence of the sums $\sum_{j=1}^{n-2k_n+1} v_j^n(s)$ for $s = 1, \dots 7$. In what follows, we use the abbreviation $g_n = n - 2k_n + 1$. For the sake of readability, we divide the rest of the proof into seven parts.

Part 1: Proof of the convergence in probability of $v_i^n(1)$

The quantity $\frac{1}{k_n} \sum_{j=1}^{g_n} v_j^n(1)$ can be decomposed as:

$$\frac{1}{k_n} \sum_{j=1}^{g_n} v_j^n(1) = \sum_{j=1}^{g_n} d_{j,1}^{(n)} + \sum_{j=1}^{g_n} d_{j,2}^{(n)},$$

where

$$d_{j,1}^{(n)} = \frac{1}{k_n^3} \sum_{j=0}^{2k_n-1} \Big(\left(\mathbb{B}_{j+i,n} - p_{j+i,n}\right)^2 - p_{i-1,n} \left(1 - p_{i-1,n}\right) \Big), \quad d_{j,2}^{(n)} = \frac{1}{k_n^3} \sum_{j=0}^{2k_n-1} p_{i-1,n} \left(1 - p_{i-1,n}\right).$$

First, we show that $\sum_{j=1}^{g_n} d_{j,1}^{(n)}$ is AN. We have:

$$\begin{aligned}
\sum_{j=1}^{g_n} \left| \mathbb{E}_{j-1} \left[d_{j,1}^{(n)} \right] \right| &= \sum_{j=1}^{g_n} \frac{1}{k_n^3} \sum_{j=0}^{2k_n-1} \left| \mathbb{E}_{j-1} \left[p_{j+i,n} - p_{j-1,n} + p_{j-1,n}^2 - p_{j+i,n}^2 \right] \right| \\
&\leq \sum_{j=1}^{g_n} \frac{1}{k_n^3} \sum_{i=0}^{2k_n-1} \left(\left| \mathbb{E}_{j-1} \left[p_{j+i,n} - p_{j-1,n} \right] \right| + \left| \mathbb{E}_{j-1} \left[p_{j+i,n}^2 - p_{j-1,n}^2 \right] \right| \right) \\
&= \sum_{j=1}^{g_n} \frac{1}{k_n^3} \sum_{i=0}^{2k_n-1} \left(\left| \mathbb{E}_{j-1} \left[p_{j+i,n} - p_{j-1,n} \right] \right| + \left| \mathbb{E}_{j-1} \left[\left(p_{j+i,n} + p_{j-1,n}\right) \left(p_{j+i,n} - p_{j-1,n}\right) \right] \right| \right) \\
&\leq C \sum_{j=1}^{g_n} \frac{1}{k_n^3} \sum_{i=0}^{2k_n-1} k_n \Delta_n = C \frac{k_n \Delta_n (2k_n - 1) g_n}{k_n^3} \sim \frac{1}{k_n} \longrightarrow 0,
\end{aligned}$$

where we use conditioning on $(p_t)_{t\in[0,1]}$, triangular inequality, and Lemma 6. Next, using the boundedness of p_t, we obtain:

$$k_n \sum_{j=1}^{g_n} \mathbb{E}_{j-1}\left[\left|d_{j,1}^{(n)}\right|^2\right] \leq k_n \sum_{j=1}^{g_n} \frac{1}{k_n^6}\left(\sum_{i=0}^{2k_n-1} C\right)^2 = C\,\frac{(2k_n-1)^2\, g_n}{k_n^5} \sim \frac{1}{k_n^3\,\Delta_n} \longrightarrow 0.$$

Consequently, by Lemma 2, $\sum_{j=1}^{g_n} d_{j,1}^{(n)}$ is AN. Now, consider $\sum_{j=1}^{g_n} d_{j,2}^{(n)}$. We have:

$$\sum_{j=1}^{g_n} d_{j,2}^{(n)} = \frac{2}{k_n^2}\sum_{j=1}^{g_n}\frac{1}{2k_n}\sum_{j=0}^{2k_n-1} p_{j-1,n}(1-p_{j-1,n}) = \frac{2}{\theta^2}\sum_{j=1}^{g_n} p_{j-1,n}(1-p_{j-1,n})\frac{1}{\lfloor\sqrt{n}\rfloor^2} \longrightarrow \frac{2}{\theta^2}\int_0^1 p_s(1-p_s)\,ds,$$

where the convergence is point-wise, by Riemann integrability. Combining the two results, we obtain:

$$\frac{1}{k_n}\sum_{j=1}^{g_n} v_j^n(1) \xrightarrow{u.c.p.} \frac{2}{\theta^2}\int_0^1 p_s(1-p_s)\,ds. \tag{61}$$

Part 2: Proof of the convergence in probability of $v_i^n(2)$
To begin, note that $v_j^n(2)$ can be written as :

$$v_j^n(2) = \frac{1}{k_n^2}\sum_{i=0}^{2k_n-1}\epsilon(2)_i^2\left(\Delta_{j+i}^n p\right)^2 = \frac{1}{k_n^2}\sum_{i=0}^{2k_n-1}\epsilon(2)_i^2\left(\Delta_j^n p\right)^2 + \frac{1}{k_n^2}\sum_{i=1}^{2k_n-1}\epsilon(2)_i^2\left[\left(\Delta_{j+i}^n p\right)^2 - \left(\Delta_j^n p\right)^2\right],$$

so that the sum over the index i of all the $v_j^n(2)$ becomes:

$$\frac{1}{k_n}\sum_{j=0}^{g_n} v_j^n(2) = \underbrace{\frac{1}{k_n^3}\sum_{i=0}^{2k_n-1}\epsilon(2)_i^2\sum_{j=0}^{g_n}\left(\Delta_j^n p\right)^2}_{\mathcal{A}_n} + \underbrace{\frac{1}{k_n^3}\sum_{j=0}^{g_n}\sum_{i=1}^{2k_n-1}\epsilon(2)_i^2\left[\left(\Delta_{j+i}^n p\right)^2 - \left(\Delta_j^n p\right)^2\right]}_{\mathcal{B}_n}.$$

Now we want to prove that $\mathcal{A}_n$ converges in probability to a finite quantity, while $\mathcal{B}_n$ is AN. Using the definition of the integers' coefficients $\epsilon(2)_i$ it is easy to show that:

$$\frac{1}{k_n^3}\sum_{j=0}^{2\,k_n-1}\epsilon(2)_i^2 = \frac{1}{3\,k_n^3}\left(2k_n^3 + k_n\right) \longrightarrow \frac{2}{3}.$$

Hence, the standard theory of realized volatility for the semimartingale:

$$p_t = p_0 + \int_0^t \mu_s\,ds + \int_0^t \nu_s\,dW_s$$

now implies that:

$$\mathcal{A}_n \xrightarrow{p} \frac{2}{3}\int_0^1 \nu_s^2\,ds.$$

Concerning $\mathcal{B}_n$, we write it as:

$$\mathcal{B}_n = \sum_{j=0}^{g_n}\vartheta_{j+1,n} \text{ with } \vartheta_{j+1,n} = \frac{1}{k_n^3}\sum_{i=1}^{2k_n-1}\epsilon(2)_i^2\left[\left(\Delta_{j+i}^n p\right)^2 - \left(\Delta_j^n p\right)^2\right],$$

and, by Markov inequality, the Itô isometry and the boundedness of[12] ν

$$\int_0^{\Delta} \nu_s \, dW_s = \nu_0 \, (W_\Delta - W_0) + O_p(\Delta^{1/2}), \tag{62}$$

Considering also that $\int_0^t \mu_s \, ds$ is $O_p\left(\Delta_n\right)$ for bounded μ , we have

$$\begin{aligned} p_{j+1,n} - p_{j,n} &= \int_{j\Delta_n}^{(j+1)\,\Delta_n} \mu_s \, ds + \int_{j\Delta_n}^{(j+1)\,\Delta_n} \nu_s \, dW_s = \left(\nu_{j,n} + O_p\left(\sqrt{\Delta_n}\right)\right) \left(W_{j+1,n} - W_{j,n}\right) + O_p(\Delta_n) \\ &= \nu_{j,n} \, \left(W_{j+1,n} - W_{j,n}\right) + O_p\left(\Delta_n^{1/2}\right) \left(W_{j+1,n} - W_{j,n}\right) + O_p\left(\Delta_n\right). \end{aligned}$$

The square of the increment $\Delta_j^n p = (p_{j+1,n} - p_{j,n})$ then becomes:

$$\begin{aligned} \left(\Delta_j^n p\right)^2 &= \nu_{j,n}^2 \, \left(\Delta_j^n W\right)^2 + \left(\Delta_j^n W\right)^2 O_p(\Delta_n) + O_p\left(\Delta_n^2\right) + \left(\Delta_j^n W\right)^2 O_p\left(\Delta_n^{1/2}\right) + \left(\Delta_j^n W\right) O_p\left(\Delta_n\right) + \left(\Delta_j^n W\right) O_p\left(\Delta_n^{3/2}\right) \\ &= \nu_{j,n}^2 \, \left(\Delta_j^n W\right)^2 + O_p\left(\Delta_n^2\right) + \left(\Delta_j^n W\right)^2 O_p\left(\Delta_n^{1/2}\right) + \left(\Delta_j^n W\right) O_p\left(\Delta_n\right), \end{aligned}$$

which, by preserving only the leading terms can be further simplified into:

$$\left(\Delta_j^n p\right)^2 = \nu_{j\,\Delta_n}^2 \, \left(\Delta_j^n W\right)^2 + O_p\left(\Delta_n^{1/2}\right) \left(\Delta_j^n W\right)^2 + \nu_{j,n} \, \left(\Delta_j^n W\right) O_p\left(\Delta_n\right), \tag{63}$$

so that:

$$\mathbb{E}_j\left[\left(\Delta_j^n p\right)^2\right] = \nu_{j,n}^2 \, \Delta_n + O_p(\Delta_n^{3/2}).$$

Now consider the same increment shifted by i units:

$$\begin{aligned} \left(\Delta_{j+i}^n p\right)^2 &= \nu_{i+j,n}^2 \, \left(\Delta_{j+i}^n W\right)^2 + O_p\left(\Delta_n^2\right) + \left(\Delta_{j+i}^n W\right)^2 O_p\left(\Delta_n^{1/2}\right) + \left(\Delta_{j+i}^n W\right) O_p\left(\Delta_n\right) \\ &= \left(\nu_{j\,\Delta_n}^2 + O_p\left(j\,\Delta_n\right) + O_p\left(\sqrt{j\,\Delta_n}\right)\right) \left(\Delta_{j+i}^n W\right)^2 + O_p\left(\Delta_n^2\right) + \left(\Delta_{j+i}^n W\right)^2 O_p\left(\Delta_n^{1/2}\right) + \left(\Delta_{j+i}^n W\right) O_p\left(\Delta_n\right) \\ &= \nu_{j\,\Delta_n}^2 \, \left(\Delta_{j+i}^n W\right)^2 + \left(\Delta_{j+i}^n W\right)^2 O_p\left(\sqrt{j\,\Delta_n}\right) + O_p\left(\Delta_n^2\right) + \left(\Delta_{j+i}^n W\right)^2 O_p\left(\Delta_n^{1/2}\right) + \left(\Delta_{j+i}^n W\right) O_p\left(\Delta_n\right), \end{aligned}$$

which, by preserving only the leading terms, can be further simplified into:

$$\left(\Delta_{j+i}^n p\right)^2 = \nu_{j,n}^2 \, \left(\Delta_{j+i}^n W\right)^2 + \left(\Delta_{j+i}^n W\right)^2 O_p\left(\sqrt{i\,\Delta_n}\right) \tag{64}$$

and so:

$$\mathbb{E}_j\left[\left(\Delta_{j+i}^n p\right)^2\right] = \nu_{j,n}^2 \, \Delta_n + O_p\left(i^{1/2}\,\Delta_n^{3/2}\right).$$

Therefore the $\mathcal{F}_{t_{i,n}}$-conditional expected value of the difference between $\left(\Delta_{j+i}^n p\right)^2$ and $\left(\Delta_j^n p\right)^2$ has the following order in probability:

$$\mathbb{E}_j\left[\left(\Delta_{j+i}^n p\right)^2 - \left(\Delta_j^n p\right)^2\right] = O_p\left(i^{1/2}\,\Delta_n^{3/2}\right),$$

implying that

$$\sum_{j=0}^{g_n} \mathbb{E}_j\left[\vartheta_{j+1,n}\right] = \frac{1}{k_n^3} \sum_{j=0}^{g_n} \sum_{i=0}^{2k_n-1} \epsilon(2)_i^2 O_p\left(i^{1/2}\,\Delta_n^{3/2}\right) = O_p\left(\left(k_n\,\Delta_n\right)^{1/2}\right) \xrightarrow{p} 0,$$

[12] Here we follow the standard approach

$$\mathbb{P}\left(\left|\frac{1}{\sqrt{\Delta}}\left[\int_0^{\Delta} \nu_s \, dW_s - \nu_0(W_\Delta - W_0)\right]\right| > M\right) \le \frac{1}{M^2\Delta}\mathbb{E}\left(\left|\int_0^{\Delta} (\nu_s - \nu_0)\, dW_s\right|^2\right) = \frac{1}{M^2\Delta}\mathbb{E}\left(\int_0^{\Delta} (\nu_s - \nu_0)^2 \, ds\right),$$

and then the identity (62) follows from the boundedness of ν.

which is the first of the two conditions in Lemma 2 that guarantee AN. To prove that also the second condition is satisfied consider:

$$k_n\,\vartheta^2_{j+1,n}=\underbrace{\frac{1}{k_n^5}\sum_{i=0}^{2k_n-1}\epsilon(2)^4_i\left[\left(\Delta^n_{j+i}p\right)^2-\left(\Delta^n_j p\right)^2\right]^2}_{\mathcal{C}_{i,n}}+\underbrace{\frac{2}{k_n^5}\sum_{i=0}^{2k_n-2}\sum_{\ell=j+1}^{2k_n-1}\epsilon(2)^2_i\,\epsilon(2)^2_\ell\,\left[\left(\Delta^n_{j+i}p\right)^2-\left(\Delta^n_j p\right)^2\right]\left[\left(\Delta^n_{j+\ell}p\right)^2-\left(\Delta^n_j p\right)^2\right]}_{\mathcal{D}_{i,n}}.$$

From (63) we get

$$\begin{aligned}\left(\Delta^n_j p\right)^4 &= \nu^4_{j,n}\,\left(\Delta^n_j W\right)^4+O_p(\Delta_n)\,\left(\Delta^n_j W\right)^4+\nu^2_{j,n}\,\left(\Delta^n_j W\right)^2 O_p\left(\Delta^2_n\right)+2\,\nu^2_{j,n}\,\left(\Delta^n_j W\right)^4 O_p\left(\Delta^{1/2}_n\right)\\ &+\;2\,\nu^3_{j,n}\,\left(\Delta^n_j W\right)^3 O_p\left(\Delta_n\right)+2\,\nu_{j,n}\,\left(\Delta^n_j W\right)^3 O_p\left(\Delta^{3/2}_n\right)\end{aligned}$$

and therefore:

$$\mathbb{E}_j\left[\left(\Delta^n_j p\right)^4\right]=3\,\nu^4_{j,n}\,\Delta^2_n+O_p(\Delta^3_n)+O_p(\Delta^3_n)+O_p(\Delta^{5/2}_n)=3\,\nu^4_{j,n}\,\Delta^2_n+O_p(\Delta^{5/2}_n).$$

Similarly, from (64) we get:

$$\left(\Delta^n_{j+i}p\right)^4=\nu^4_{j,n}\,\left(\Delta^n_{j+i}W\right)^4+\left(\Delta^n_{j+i}W\right)^4 O_p\left(j\,\Delta_n\right)+2\,\nu^2_{j,n}\,\left(\Delta^n_{j+i}W\right)^4 O_p\left(\sqrt{i\,\Delta_n}\right)$$

and hence

$$\mathbb{E}_j\left[\left(\Delta^n_{j+i}p\right)^4\right]=3\,\nu^4_{j,n}\,\Delta^2_n+O_p\left(i\,\Delta^3_n\right)+O_p\left(i^{1/2}\,\Delta^{5/2}_n\right)=3\,\nu^4_{j,n}\,\Delta^2_n+O_p\left(i^{1/2}\,\Delta^{5/2}_n\right).$$

Summing up the two fourth powers so obtained:

$$\mathbb{E}_j\left[\left(\Delta^n_{j+i}p\right)^4+\left(\Delta^n_{j+i}p\right)^4\right]=6\,\nu^4_{j,n}\,\Delta^2_n+O_p\left(i^{1/2}\,\Delta^{5/2}_n\right).$$

Finally consider that:

$$\begin{aligned}\left(\Delta^n_{j+i}p\right)^2\,\left(\Delta^n_j p\right)^2 &= \left(\nu^2_{j,n}\,\left(\Delta^n_{j+i}W\right)^2+\left(\Delta^n_{j+i}W\right)^2 O_p\left(\sqrt{i\,\Delta_n}\right)\right)\times\\ &\quad\left(\nu^2_{j,n}\,\left(\Delta^n_j W\right)^2+\left(\Delta^n_j W\right)^2 O_p\left(\Delta^{1/2}_n\right)+\nu_{j,n}\,\left(\Delta^n_j W\right)\,O_p\left(\Delta_n\right)\right)\\ &= \nu^4_{j,n}\,\left(\Delta^n_{j+i}W\right)^2\,\left(\Delta^n_j W\right)^2+\nu^2_{j,n}\,\left(\Delta^n_{j+i}W\right)^2\,\left(\Delta^n_j W\right)^2 O_p\left(\Delta^{1/2}_n\right)+\nu^3_{j,n}\,\left(\Delta^n_{j+i}W\right)^2\,\left(\Delta^n_j W\right)\,O_p\left(\Delta_n\right)\\ &+\;\nu^2_{j,n}\,\left(\Delta^n_{j+i}W\right)^2\,\left(\Delta^n_j W\right)^2 O_p\left(\sqrt{i\,\Delta_n}\right)+\left(\Delta^n_{j+i}W\right)^2\,\left(\Delta^n_j W\right)^2 O_p\left(i^{1/2}\,\Delta_n\right)\\ &+\;\nu_{j,n}\,\left(\Delta^n_{j+i}W\right)^2\,\left(\Delta^n_j W\right)\,O_p\left(i^{1/2}\,\Delta^{3/2}_n\right),\end{aligned}$$

where:

$$\mathbb{E}_j\left[\left(\Delta^n_{j+i}p\right)^2\,\left(\Delta^n_j p\right)^2\right]=\nu^4_{j,n}\,\Delta^2_n+O_p\left(\Delta^{5/2}_n\right)+O_p\left(i^{1/2}\,\Delta^{5/2}_n\right)+O_p\left(i^{1/2}\,\Delta^3_n\right)=\nu^4_{j,n}\,\Delta^2_n+O_p\left(i^{1/2}\,\Delta^{5/2}_n\right).$$

Therefore:

$$\mathbb{E}_j\left[\left(\Delta^n_{j+i}p\right)^4+\left(\Delta^n_{j+i}p\right)^4-2\,\left(\Delta^n_{j+i}p\right)^2\,\left(\Delta^n_j p\right)^2\right]=6\,\nu^4_{j,n}\,\Delta^2_n+O_p\left(i^{1/2}\,\Delta^{5/2}_n\right),$$

which implies:

$$\sum_{i=0}^{g_n}\mathbb{E}\left[\mathcal{C}_{j,n}\right]=\frac{1}{k_n^5}\sum_{i=0}^{g_n}\sum_{i=0}^{2k_n-1}\epsilon(2)^4_i\mathbb{E}\left[\left(\left(\Delta^n_{j+i}p\right)^2-\left(\Delta^n_j p\right)^2\right)^2\right]=\frac{1}{k_n^5}\sum_{i=0}^{g_n}\sum_{i=0}^{2k_n-1}\epsilon(2)^4_i\mathbb{E}\left[\nu^4_{j,n}\,\Delta^2_n+O_p\left(i^{1/2}\,\Delta^{5/2}_n\right)\right]=O\left(\Delta_n\right)\longrightarrow 0.$$

Concerning $\mathcal{C}_{j,n}$, first call $\varepsilon_{i,\ell}^2 = \epsilon(2)_i^2\,\epsilon(2)_\ell^2$ and then note that:

$$\begin{aligned}
\mathbb{E}\left[|\mathcal{D}_{j,n}|\right] &= \frac{2}{k_n^5}\sum_{i=0}^{2k_n-2}\sum_{\ell=i+1}^{2k_n-1}\varepsilon_{i,\ell}^2\,\mathbb{E}\left[\left|\left(\Delta_{j+i}^n p\right)^2-\left(\Delta_j^n p\right)^2\right|\,\left|\left(\Delta_{j+\ell}^n p\right)^2-\left(\Delta_j^n p\right)^2\right|\right]\\
&\le \frac{2}{k_n^5}\sum_{i=0}^{2k_n-2}\sum_{\ell=i+1}^{2k_n-1}\varepsilon_{j,\ell}^2\left(\mathbb{E}\left[\left(\left(\Delta_{j+i}^n p\right)^2-\left(\Delta_j^n p\right)^2\right)^2\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\left(\left(\Delta_{j+\ell}^n p\right)^2-\left(\Delta_j^n p\right)^2\right)^2\right]\right)^{\frac{1}{2}}\\
&= \frac{2}{k_n^5}\sum_{i=0}^{2k_n-2}\sum_{\ell=i+1}^{2k_n-1}\varepsilon_{j,\ell}^2\left(\mathbb{E}\left[6\,\nu_{j,n}^4\,\Delta_n^2+O_p\left(j^{\frac{1}{2}}\,\Delta_n^{\frac{5}{2}}\right)\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[6\,\nu_{j,n}^4\,\Delta_n^2+O_p\left(\ell^{\frac{1}{2}}\,\Delta_n^{\frac{5}{2}}\right)\right]\right)^{\frac{1}{2}}.
\end{aligned}$$

Because $\epsilon(2)_i^2\,\epsilon(2)_\ell^2 \le C\,k_n^4$, we get:

$$\mathbb{E}\left[|\mathcal{D}_{j,n}|\right] \le C\,k_n\,\Delta_n^2$$

so that:

$$\sum_{j=0}^{g_n}\mathbb{E}\left[|\mathcal{D}_{j,n}|\right] \le C\,k_n\,\Delta_n \to 0,$$

and hence, in conclusion, $\mathcal{B}_n$ is AN.

Part 3: Proof of the convergence in probability of $v_i^n(3)$

In what follows, we call:

$$\zeta(1)_j \doteq \mathbb{B}_{j,n}-p_{j,n} \text{ and } \zeta(2)_j \doteq p_{j+1,n}-p_{j,n}.$$

The quantity $v_j^n(3)$ can be rewritten as:

$$v_j^n(3) = \frac{2}{k_n^2}\sum_{i=0}^{2k_n-1}\epsilon(1)_i\,\epsilon(2)_i\,\zeta(1)_{j+i}\,\zeta(2)_{j+i}.$$

Therefore, the quantity $\frac{1}{k_n}\sum_{i=0}^{g_n}v_i^n(3)$ becomes:

$$\frac{1}{k_n}\sum_{i=0}^{g_n}v_j^n(3) = \sum_{i=0}^{g_n}\frac{2}{k_n^3}\sum_{i=0}^{2k_n-1}\epsilon(1)_i\,\epsilon(2)_i\,\zeta(1)_{j+i}\,\zeta(2)_{j+i}.$$

First, we observe that, conditionally on (p_t), we have that $\mathbb{E}\left[\zeta(1)_j\right]=0$ and so $\mathbb{E}_{j-1}\left[v_j^n(3)\right]=0$. Then, we note that term $k_n\left(\frac{v_j^n(3)}{k_n}\right)^2$ can be decomposed as:

$$\begin{aligned}
k_n\left(\frac{v_j^n(3)}{k_n}\right)^2 &= \frac{4}{k_n^5}\sum_{i=0}^{2k_n-1}\left(\epsilon(2)_i\right)^2\left(\zeta(1)_{j+i}\right)^2\left(\zeta(2)_{j+i}\right)^2+\frac{8}{k_n^5}\sum_{j=0}^{2k_n-2}\sum_{i=0}^{2k_n-1}\epsilon(1)_j\epsilon(2)_j\zeta(1)_{i+j}\zeta(2)_{i+j}\epsilon(1)_l\epsilon(2)_l\zeta(1)_{i+l}\zeta(2)_{i+l}\\
&\doteq \mathcal{A}_{1,n}+\mathcal{A}_{2,n}
\end{aligned}$$

Now, by conditioning on (p_t), we readily obtain that $\mathbb{E}\left[\mathcal{A}_{2,n}\right]=0$. Concerning $\mathcal{A}_{1,n}$, we have:

$$\mathbb{E}\left[|\mathcal{A}_{1,n}|\right] \le \mathbb{E}\left[\frac{4}{k_n^5}\sum_{i=0}^{2k_n-1}\left(\epsilon(2)_i\right)^2\left(\zeta(1)_{j+i}^2\left(\zeta(2)_{j+i}\right)^2\right)\right] \le \frac{C}{k_n^5}\Delta_n\sum_{i=0}^{2k_n-1}\left(\epsilon(2)_i\right)^2$$

By the boundedness of Bernoulli random variables and (p_t) we have that $\left(\zeta\left(1\right)_{j+i}\right)^2 \leq C$ for some positive constant C. Therefore:

$$\mathbb{E}\left[\left|\mathcal{A}_{1,n}\right|\right] \leq \frac{C}{k_n^5}\Delta_n \sum_{j=0}^{2k_n-1}\left(\epsilon\left(2\right)_i\right)^2 = \frac{C}{k_n^5}\Delta_n \frac{2k_n^3+k_n}{3} \sim \frac{\Delta_n}{k_n^2}.$$

Therefore:

$$\sum_{j=1}^{g_n}\mathbb{E}\left[\left|\mathcal{A}_{1,n}\right|\right] \leq \frac{C}{k_n} \to 0.$$

Consequently, by Lemma 2, $\frac{1}{k_n}v_j^n\left(3\right)$ is AN.

Part 4: Proof of the convergence in probability of $v_i^n\left(4\right)$

First, by conditioning on (p_t) we readily obtain $\mathbb{E}_{j-1}\left[v_j^n\left(4\right)\right] = 0$. Next, consider the decomposition:

$$\left(\frac{v_j^n\left(4\right)}{k_n}\right)^2 = \mathcal{A}_{1,n} + \mathcal{A}_{2,n},$$

where

$$\mathcal{A}_{1,n} = \frac{C}{k_n^6}\sum_{j=0}^{2k_n-2}\left(\sum_{l=i+1}^{2k_n-1}\epsilon(1)_i\epsilon(1)_l\zeta(1)_{j+i}\zeta(1)_{j+l}\right)^2,$$

and

$$\begin{aligned}\mathcal{A}_{2,n} &= \frac{C}{k_n^6}\sum_{i=0}^{2k_n-3}\sum_{m=i+1}^{2k_n-2}\left(\sum_{l=i+1}^{2k_n-1}\epsilon(1)_i\epsilon(1)_l\zeta(1)_{j+i}\zeta(1)_{j+l}\right)\left(\sum_{u=m+1}^{2k_n-1}\epsilon(1)_m\epsilon(1)_u\zeta(1)_{j+m}\zeta(1)_{j+u}\right),\\ &= \frac{C}{k_n^6}\sum_{i=0}^{2k_n-3}\sum_{m=i+1}^{2k_n-2}\sum_{l=i+1}^{2k_n-1}\sum_{u=i+2}^{2k_n-1}\epsilon(1)_i\epsilon(1)_l\epsilon(1)_m\epsilon(1)_u\zeta(1)_{j+i}\zeta(1)_{j+l}\zeta(1)_{j+u}\zeta(1)_{j+m}.\end{aligned}$$

By conditioning on (p_t) again, we have $\mathbb{E}\left[\zeta(1)_{j+i}\zeta(1)_{j+l}\zeta(1)_{j+u}\zeta(1)_{j+m}\right] = 0$ if at least two of the indexes i, l, u, m are different. Because in the sums that appear in $\mathcal{A}_{2,n}$ one among m, l, or u is different from i, we have $\mathbb{E}\left[\mathcal{A}_{2,n}\right] = 0$. Analogously: the expected value of the cross-product terms in $\mathcal{A}_{1,n}$ is zero. Next, because $|\zeta(1)_{j+l}| \leq C$, for some constant $C > 0$,

$$\mathbb{E}\left[\mathcal{A}_{1,n}\right] = \frac{C}{k_n^6}\sum_{i=0}^{2k_n-2}\sum_{l=i+1}^{2k_n-1}\mathbb{E}\left[\left(\zeta(1)_{j+i}\zeta(1)_{j+l}\right)^2\right] \leq \frac{C(2k_n-2)(2k_n-1)}{k_n^6} \sim \frac{1}{k_n^4}.$$

Therefore:

$$k_n\sum_{i=1}^{g_n}\mathbb{E}\left[\left(\frac{v_j^n\left(4\right)}{k_n}\right)^2\right] \leq \frac{C}{k_n^3\Delta_n} \to 0.$$

Consequently, by Lemma 2, $\frac{1}{k_n}v_j^n\left(4\right)$ is AN. *Part 5: Proof of the convergence in probability of* $v_i^n\left(5\right)$

By successive conditioning and using Lemma 6, we obtain

$$\left|\mathbb{E}_{j-1}\left[v_j^n\left(5\right)\right]\right| \leq \frac{C}{k_n^2}\sum_{i=0}^{2k_n-2}\sum_{l=i+1}^{2k_n-1}\epsilon(2)_i\epsilon(2)_l\Delta_n^2 = C\frac{\Delta_n^2}{k_n^2}\sum_{i=0}^{2k_n-2}\sum_{l=i+1}^{2k_n-1}\epsilon(2)_i\epsilon(2)_l \sim C\,\Delta_n^2 k_n^2,$$

where we use the fact that $\sum_{i=0}^{2k_n-2}\sum_{l=i+1}^{2k_n-1}\epsilon(2)_i\epsilon(2)_l \sim k_n^4$. Therefore, we have:

$$\sum_{j=1}^{g_n}\frac{1}{k_n}\left|\mathbb{E}_{j-1}\left[v_j^n\left(5\right)\right]\right| \sim \Delta_n k_n \to 0.$$

Next, we have:

$$\left(\frac{v_j^n\,(5)}{k_n}\right)^2 = \mathcal{A}_{1,n} + \mathcal{A}_{2,n},$$

where:

$$\mathcal{A}_{1,n} = \frac{C}{k_n^6}\sum_{i=0}^{2k_n-2}\left(\sum_{l=i+1}^{2k_n-1}\epsilon(2)_i\epsilon(2)_l\zeta(2)_{j+i}\zeta(2)_{j+l}\right)^2,$$
$$\mathcal{A}_{2,n} = \frac{C}{k_n^6}\sum_{j=0}^{2k_n-3}\sum_{m=j+1}^{2k_n-2}\sum_{l=i+1}^{2k_n-1}\sum_{u=j+2}^{2k_n-1}\epsilon(2)_i\epsilon(2)_l\epsilon(2)_m\epsilon(2)_u\zeta(2)_{j+i}\zeta(2)_{j+l}\zeta(2)_{i+u}\zeta(2)_{i+m}.$$

Furthermore, we have:

$$\mathcal{A}_{1,n} = \mathcal{A}_{1,1,n} + \mathcal{A}_{1,2,n},$$

where:

$$\mathcal{A}_{1,1,n} = \frac{C}{k_n^6}\sum_{i=0}^{2k_n-2}\sum_{l=i+1}^{2k_n-1}\left(\epsilon(2)_i\epsilon(2)_l\zeta(2)_{j+i}\zeta(2)_{j+l}\right)^2,$$
$$\mathcal{A}_{1,2,n} = \frac{C}{k_n^6}\sum_{i=0}^{2k_n-3}\sum_{l=i+1}^{2k_n-2}\sum_{m=i+2}^{2k_n-1}\left(\epsilon(2)_i\right)^2\epsilon(2)_l\epsilon(2)_m\left(\zeta(2)_{j+i}\right)^2\zeta(2)_{j+l}\zeta(2)_{j+m}.$$

Using the estimate (43) of Lemma 6, and the fact that $\sum_{i=0}^{2k_n-2}\sum_{l=i+1}^{2k_n-1}\left(\epsilon(2)_i\epsilon(2)_l\right)^2 \sim k_n^6$, we obtain:

$$\mathbb{E}\left[\mathcal{A}_{1,1,n}\right] \leq C\,\frac{\Delta_n^2}{k_n^6}\sum_{i=0}^{2k_n-2}\sum_{l=i+1}^{2k_n-1}\left(\epsilon(2)_i\epsilon(2)_l\right)^2 \sim \Delta_n^2,$$

which implies that $k_n\sum_{i=1}^{g_n}\mathbb{E}\left[\mathcal{A}_{1,1,n}\right] \leq k_n\,\Delta_n \to 0$. Next, using the estimates (43) and (44) of Lemma 6, we have

$$\mathbb{E}\left[\left(\zeta(2)_{j+i}\right)^2\zeta(2)_{j+l}\zeta(2)_{j+m}\right] \leq C\begin{cases}\Delta_n^2, & l=m,\\ \Delta_n^3, & i\neq l\neq m.\end{cases}$$

Therefore, we have:

$$\mathbb{E}\left[\mathcal{A}_{1,2,n}\right] \leq C\frac{\Delta_n^2}{k_n^6}S_1 + C\frac{\Delta_n^3}{k_n^6}S_2 \sim \Delta_n^2 \vee \Delta_n^3 k_n,$$

where:

$$S_1 = \sum_{j=0}^{2k_n-3}\sum_{l=i+1}^{2k_n-2}\sum_{m=i+2}^{2k_n-1}\left(\epsilon(2)_i\right)^2\epsilon(2)_l\epsilon(2)_m\mathbb{I}\left(\{l=m\}\right) = \sum_{i=0}^{2k_n-3}\sum_{l=i+2}^{2k_n-2}\left(\epsilon(2)_i\right)^2\left(\epsilon(2)_l\right)^2 \sim k_n^6,$$
$$S_2 = \sum_{i=0}^{2k_n-3}\sum_{l=i+1}^{2k_n-2}\sum_{m=i+2}^{2k_n-1}\left(\epsilon(2)_i\right)^2\epsilon(2)_l\epsilon(2)_m\mathbb{I}\left(\{l\neq m\}\right) = \sum_{i=0}^{2k_n-3}\sum_{l=i+1}^{2k_n-2}\sum_{m=i+2}^{2k_n-1}\left(\epsilon(2)_i\right)^2\epsilon(2)_l\epsilon(2)_m - S_1 \sim k_n^7.$$

Consequently,

$$k_n\sum_{j=1}^{g_n}\mathbb{E}\left[\mathcal{A}_{1,2,n}\right] \leq C\,\Delta_n k_n \to 0.$$

So, summing up $k_n \sum_{j=1}^{g_n} \mathbb{E}[\mathcal{A}_{1,n}] \to 0$. With a procedure similar to that used for $\mathcal{A}_{1,2,n}$, we obtain

$$k_n \sum_{j=1}^{g_n} \mathbb{E}[\mathcal{A}_{2,n}] \leq C\,\Delta_n k_n \to 0.$$

Thus, $\frac{1}{k_n} v_j^n(5)$ is AN by Lemma 2.

Part 6: Proof of the convergence in probability of $v_i^n(6)$ and $v_i^n(7)$. First, by conditioning on (p_t) we readily obtain $\mathbb{E}_{j-1}\left[v_j^n(6)\right] = 0$. Next, consider the decomposition:

$$\left(\frac{v_j^n(6)}{k_n}\right)^2 = \mathcal{A}_{1,n} + \mathcal{A}_{2,n},$$

where

$$\mathcal{A}_{1,n} = \frac{C}{k_n^6} \sum_{i=0}^{2k_n-2} \left(\sum_{l=i+1}^{2k_n-1} \epsilon(1)_i \epsilon(2)_l \zeta(1)_{j+i} \zeta(2)_{j+l} \right)^2$$

and

$$\begin{aligned}
\mathcal{A}_{2,n} &= \frac{C}{k_n^6} \sum_{i=0}^{2k_n-3} \sum_{m=i+1}^{2k_n-2} \left(\sum_{l=i+1}^{2k_n-1} \epsilon(1)_i \epsilon(2)_l \zeta(1)_{j+i} \zeta(2)_{j+l} \right) \left(\sum_{u=m+1}^{2k_n-1} \epsilon(1)_m \epsilon(2)_u \zeta(1)_{j+m} \zeta(2)_{j+u} \right), \\
&= \frac{C}{k_n^6} \sum_{i=0}^{2k_n-3} \sum_{m=i+1}^{2k_n-2} \sum_{l=i+1}^{2k_n-1} \sum_{u=i+2}^{2k_n-1} \epsilon(1)_i \epsilon(2)_l \epsilon(1)_m \epsilon(2)_u \zeta(1)_{j+i} \zeta(2)_{j+l} \zeta(1)_{j+u} \zeta(2)_{j+m}.
\end{aligned}$$

By conditioning on (p_t), $\mathbb{E}[\mathcal{A}_{2,n}] = 0$, because $\mathbb{E}[\zeta(1)_{j+i}\zeta(1)_{j+u}] = 0$ for $u > i$. Analogously, the expected value of the cross-product terms in $\mathcal{A}_{1,n}$ is zero. Therefore, we have:

$$\mathbb{E}[\mathcal{A}_{1,n}] = \frac{C}{k_n^6} \sum_{i=0}^{2k_n-2} \sum_{l=i+1}^{2k_n-1} \mathbb{E}\left[\left(\epsilon(1)_i \epsilon(2)_l \zeta(1)_{j+i} \zeta(2)_{j+l}\right)^2 \right] \leq C\,\frac{\Delta_n}{k_n^6} \sum_{i=0}^{2k_n-2} \sum_{l=i+1}^{2k_n-1} (\epsilon(2)_l)^2 \sim \frac{\Delta_n}{k_n^2}.$$

Thus:

$$k_n \sum_{j=1}^{g_n} \mathbb{E}[\mathcal{A}_{1,n}] \leq \frac{C}{k_n} \to 0.$$

Consequently, $\frac{1}{k_n} v_j^n(6)$ is AN by Lemma 2. Analogously, $\frac{1}{k_n} v_j^n(7)$ is AN as well.

Recent Issues

No. 235	Christian Kubitza, Loriana Pelizzon, Mila Getmansky Sherman	The Pitfalls of Central Clearing in the Presence of Systematic Risk
No. 234	Alejandro Bernales, Nicolás Garrido, Satchit Sagade, Marcela Valenzuela, Christian Westheide	A Tale of One Exchange and Two Order Books: Effects of Fragmentation in the Absence of Competition
No. 233	Baptiste Massenot, Yuri Pettinicchi	Can Households See into the Future? Survey Evidence from the Netherlands
No. 232	Jannic Alexander Cutura	Debt Holder Monitoring and Implicit Guarantees: Did the BRRD Improve Market Discipline?
No. 231	Benjamin Clapham, Peter Gomber, Jens Lausen, Sven Panz	Liquidity Provider Incentives in Fragmented Securities Markets
No. 230	Yalin Gündüz, Giorgio Ottonello, Loriana Pelizzon, Michael Schneider, Marti G. Subrahmanyam	Lighting up the Dark: Liquidity in the German Corporate Bond Market
No. 229	Daniel Harenberg	Asset Pricing in OLG Economies With Borrowing Constraints and Idiosyncratic Income Risk
No. 228	Roberto C. Panzica	Idiosyncratic Volatility Puzzle: The Role of Assets' Interconnections
No. 227	Mila Getmansky, Ravi Jagannathan, Loriana Pelizzon, Ernst Schaumburg, Darya Yuferova	Stock Price Crashes: Role of Slow-Moving Capital
No. 226	Loriana Pelizzon, Marti G. Subrahmanyam, Davide Tomio, Jun Uno	Central Bank–Driven Mispricing?
No. 225	Monica Billio, Massimiliano Caporin, Lorenzo Frattarolo, Loriana Pelizzon	Networks in risk spillovers: A multivariate GARCH perspective