Statistical inferences for price staleness

Document Version
Accepted author manuscript

Statistical inferences for price staleness[*]

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

January 16, 2020

Abstract

This paper proposes a nonparametric theory for statistical inferences on zero returns of high-frequency asset prices. Using an infill asymptotic design, we derive limit theorems for the percentage of zero returns observed on a finite time interval and for other related quantities. Within this framework, we develop two nonparametric tests. First, we test whether intra-day zero returns are independent and identically distributed. Second, we test whether intra-day variation of the likelihood of occurrence of zero returns can be solely explained by a deterministic diurnal pattern. In an empirical application to ten representative stocks of the NYSE, we provide evidence that the null of independent and identically distributed intra-day zero returns can be conclusively rejected. We further find that a deterministic diurnal pattern is

not sufficient to explain the intra-day variability of the distribution of zero returns.

1   Introduction

Traditional models in continuous-time finance entail that the price of a financial asset, traded in a frictionless market, evolves as a semimartingale. Bandi et al. (2017) provide empirical evidence against this hypothesis by 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 is pivotal from both an economic and an econometric point of view. Bandi et al. (2017) provide a microstructural model of price formation (following the spirit of Kyle, 1985; Hasbrouck and Ho, 1987; Glosten and Milgrom, 1985) where the lack of price updates is determined by the joint effect of asymmetric information, transaction costs and delays in the incorporation of the information flow into the assets' prices. 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., 2017), which results in distorting traditional jump tests towards false jump detection. Even though one may claim that such sluggish dynamics are only the spurious consequence of price discreteness, the empirical analysis in Bandi et al. (2019) shows that this argument is false. 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.

The occurrence of zero returns is thus an economically meaningful feature of the data-generating process of financial asset prices. 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) focusing on the erratic behavior of price paths, here we look at the other side of the coin and we answer the following research questions: does the price staleness vary empirically on an intra-daily basis? In the affirmative case, what is an appropriate model for such variability?

We assume the existence of an efficient price process Y, which we define as the asset price that would be observed if the market was perfectly liquid. Hereafter, we will assume that the process Y has not "genuine" zeros, i.e. that the increments of

Y follow a continuous distribution: $\forall t \neq s$, $\mathbb{P}[Y_t - Y_s = 0] = 0$.[1] In the presence of illiquidity frictions (such as trading costs), the trading activity is inhibited. We model this frictional dynamics assuming that, for any partition $0 = t_{0,n} < t_{1,n} < \dots < t_{n,n} = T$ of the interval $[0,T]$ (e.g., one trading day) there is a non-zero probability that, in each element of the partition, the efficient price is not observed. Whenever this event occurs, we substitute the missing price with the price observed in the previous point of the partition.[2] Assuming that the efficient price is observed with probability one in $t = 0$, the observed price process $X^{(n)}$ at the frequency[3] n is generated by the following recursive scheme

$$
\begin{aligned}
X_0^{(n)} &= Y_0, \\
X_{t_{j,n}}^{(n)} &= Y_{t_{j,n}}\left(1 - \mathbb{B}_{j,n}\right) + X_{t_{j-1,n}}^{(n)}\mathbb{B}_{j,n}, \quad j = 1, \dots, n,
\end{aligned}
\tag{1}
$$

where $Y_{t_{j,n}}$ indicates the efficient price sampled in the j-th element of the partition and where $(\mathbb{B}_{j,n})_{j=1,\dots,n}$ is a triangular array of Bernoulli random variates such that

$$
\frac{1}{T}\sum_{j=1}^{n}\left(t_{j,n} - t_{j-1,n}\right)\mathbb{B}_{j,n} \xrightarrow{p} \bar{\mathfrak{p}}_T, \quad \text{as} \quad n \to \infty,
\tag{2}
$$

for some (random) $\bar{\mathfrak{p}}_T \in (0,1)$, which can be interpreted as the average staleness on the time interval $[0,T]$. The recursive equation (1) implies that at each instant $t_{j,n}$ the observed price $X_{t_{j,n}}^{(n)}$ may either coincide with the latent efficient price (if $\mathbb{B}_{j,n} = 0$) or may not update and stay constant (if $\mathbb{B}_{j,n} = 1$), thus leading to a stale price. The event $\{\mathbb{B}_{j,n} = 1\}$ has, in this setting, a twofold interpretation: it may indicate the impossibility of observing the efficient price in the point $t_{j,n}$, or the occurrence of a zero return in the time interval $[t_{j-1,n}, t_{j,n}]$. For equi-spaced observations, i.e. imposing $t_{j,n} - t_{j-1,n} = \frac{T}{n}$, the assumption in equation (2) implies that the average probability, over the time interval $[0,T]$, of the event $\{\mathbb{B}_{j,n} = 1\}$ converges, asymptotically, to $\mathbb{E}[\bar{\mathfrak{p}}_T]$. This setting is in line with the data generating process

---

[1]This assumption is motivated by the ubiquitous semimartingale model for asset prices in continuous time. Nonetheless, it holds for a much larger class of stochastic processes. For instance, it holds for infinite activity pure-jump processes, such as α-stable motion (Cont and Tankov, 2004), or for non-semimartingales processes, such as fractional Brownian motion (Mandelbrot and Van Ness, 1968).

[2]Here, we work under the assumption of no multiple observations, i.e. at each instant of the partition $t_{j,n}$ we have at most one observation of the price process. However, recent literature (Liu et al., 2018) suggests that the presence of multiple observations is a common feature in (high-frequency) datasets. The inclusion of this feature in the proposed inferential theory is outside of the scope of this paper.

[3]The superscript (n) in the notation $X_{t_{j,n}}^{(n)}$ indicates that even if $t_{j,n} = t_{j,n'} = t$ for $n \neq n'$ it typically occurs that $X_t^{(n)} \neq X_t^{(n')}$. That is, we model the observed price process as a realization of a frequency-dependent process.

studied in Phillips and Yu (2007) and with the alternative (to a semimartingale null) hypothesis proposed by Bandi et al. (2017). Here, we build on this assumption by representing $\bar{\mathfrak{p}}_T$ as

$$\bar{\mathfrak{p}}_T = \int_0^T \mathfrak{p}_t \, dt,$$

where $\mathfrak{p}$ is a suitable stochastic process whose values fall within the open interval $(0, 1)$. In this paper, we derive suitable conditions for the triangular array $\mathbb{B}_{j,n}$ to guarantee the existence of a feasible local estimator of the process $\mathfrak{p}$. The econometric interpretation of $\mathfrak{p}$ is straightforward: as the spot volatility is the time derivative (or the instantaneous variation) of the integrated volatility (which is consistently estimated, in absence of micro-structural noise and jumps, by the realized volatility), the process $\mathfrak{p}$ represents the instantaneous variation[4] of $\bar{\mathfrak{p}}$, which is consistently estimated by the idle time defined in Bandi et al. (2017) or, as we will prove in this paper, by the *realized staleness* (defined below). For this reason, having in mind that $\bar{\mathfrak{p}}$ represents integrated price staleness, we call $\mathfrak{p}$ *instantaneous (price) staleness*. In other words, as $\bar{\mathfrak{p}}$ represents the average staleness on a time interval of finite length, $\mathfrak{p}$ represents the same average, but on an infinitesimal time interval.

In this paper, we develop an inferential theory for the dynamics of $\mathfrak{p}$ and, consequently, for the intra-day dynamics of instantaneous price staleness. Our first result is to show that, under a suitable assumption on the triangular array $\mathbb{B}_{j,n}$, the intra-day fraction of zero returns, which we address as *realized staleness*, is a consistent estimator of $\int_0^T \mathfrak{p}_t \, dt$. Then, under the assumption that the process $(\mathfrak{p}_t)_{t \in [0,T]}$ evolves as a Brownian semimartingale, we derive a (stable) central limit theorem (CLT) for the realized staleness. To set up a feasible confidence interval, we introduce a new economic indicator, called the *m-realized staleness*, and we derive its limiting properties when $n \to \infty$. Next, we introduce a consistent local estimator of $\mathfrak{p}$ and prove that 1) it allows the construction of a non-parametric test capable of distinguishing between a constant and a time-varying $\mathfrak{p}$ and 2) under the null of a Brownian semimartingale for $\mathfrak{p}$, it allows the definition of a consistent estimator of its integrated volatility. Empirically, using the transactions prices of ten representative stocks of the NYSE for the period 2006-2014, we prove that the null of a constant $\mathfrak{p}$ is not compatible with the data. This intra-day variability could be, in principle, solely driven by a deterministic seasonal effect or, in con-

---

[4]To clarify the intuition, note that, for a generic time t such that $0 < t < T$, one can define $\bar{\mathfrak{p}}_t$ as the average staleness on the time interval $[0, t]$, following the same construction used to define $\bar{\mathfrak{p}}_T$ as the average staleness on $[0, T]$. The process $\mathfrak{p}_t$ could be thus alternatively defined as

$$\mathfrak{p}_t = \lim_{h \to 0} \frac{\bar{\mathfrak{p}}_{t+h} - \bar{\mathfrak{p}}_t}{h} = \lim_{h \to 0} \frac{\int_t^{t+h} \mathfrak{p}_s \, ds}{h}$$

provided that the limit exists in an appropriate sense.

trast, it could be the result of stochastic shocks superimposed to a deterministic pattern. To test whether a recurrent deterministic pattern is sufficient to explain the time-dependence of instantaneous staleness, following Christensen et al. (2018), we derive a feasible test, based on transaction prices, that is (stably) distributed as a standard normal random variable if $\mathfrak{p}$ coincides, up to a multiplicative constant, with a deterministic recurrent pattern, and diverges (in probability) otherwise.

The test, when run on our dataset, reveals that the intra-day dynamics of instantaneous price staleness is richer than a simple deterministic seasonal effect, although the latter constitutes the leading determinant of its variability.

The remainder of the paper is organized as follows. Section 2 introduces the mathematical setting. Section 3 contains the limit results. The nonparametric tests are derived in Section 4. Section 5 shows the finite sample accuracy of our asymptotic theory using a Monte Carlo exercise and Section 6 presents the empirical results. Section 7 contains the conclusions. All technical proofs are confined to Appendix A. Finally, Appendix B contains an extension of the Monte Carlo exercize presented in Section 5.

2   The mathematical setting

We work on a filtered probability space $\left(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$ that supports all the 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. From now on let $T = 1$ for simplicity of notation. We consider refining equispaced partitions of the unit time interval $[0, 1]$, $\Pi_n = \{t_{0,n}, \ldots, t_{n,n}\}$, with $0 = t_{0,n} < t_{1,n} < \ldots < t_{n,n} = 1$, where n belongs to an increasing subsequence of $\mathbb{N}$, such that $\Pi_n \subseteq \Pi_{n'}$ for all $n' \geq n$ in the subsequence[5]. 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 $j = 0, ..., n$, and we indicate the distance between two consecutive points of the partition with $\Delta_n = 1/n$.

As anticipated in the introduction, our inferential theory develops around the concept of instantaneous staleness $\mathfrak{p}$. What follows formalizes this idea.

Assumption 1. *There exists an adapted a.s. Riemann-integrable continuous-time stochastic process* $(\mathfrak{p}_t)_{t\in[0,1]}$, *taking the values in* $(0, 1)$, *such that the triangular array* $(\mathbb{B}_{j,n})_{j=1,...,n}$ *in equation* (1) *consists of Bernoulli random variables* $\mathbb{B}_{j,n} = \mathbb{B}_{t_{j,n}}$

---

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

defined as:

$$\forall t, \quad \mathbb{B}_t \doteq \mathbb{1}_{\{\mathfrak{u}_t \leq \mathfrak{p}_t\}}, \tag{3}$$

where $\mathbb{1}_{\{\cdot\}}$ is the indicator function and where $(\mathfrak{u}_t)_{t \in [0,1]}$ is a collection of uniformly distributed random variables (independent of $\mathfrak{p}_t$) satisfying $\mathfrak{u}_t \perp \mathfrak{u}_{t'}$, $\forall t \neq t'$, and $\mathfrak{u}_t \in \mathcal{F}_t \, \forall t$.

Assumption 1 has two important implications and deserves some additional comments.

First, it preserves the *compatibility relationship* (see Aït-Sahalia and Jacod, 2014, p. 211) over different sampling frequencies for the array $\mathbb{B}_{j,n}$. 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 $j = 1, \ldots, n$ and $j' = 1, \ldots, n'$, then $\mathbb{B}_{j,n} = \mathbb{B}_{j',n'}$ whenever $j/n = j'/n'$. Since, as anticipated in the introduction, the event $\{\mathbb{B}_{j,n} = 1\}$ represents the impossibility of observing the efficient price in $t_{j,n}$, this means that if the efficient price is not observed at a given point in time, its value will not be revealed by increasing the frequency. On the other hand, if we do observe the efficient price at a given point, this observation will not be lost as n increases. To clarify, suppose that $\mathbb{B}_{t_{1,1}} = \mathbb{B}_1 = 1$, i.e. the efficient price process Y is not observed at $t_{1,1} = 1$. Then, the observed price at the frequency $n = 1$ in the point $t_{1,1}$ is set, according to the recursive scheme in equation (1), equal to the observed price in the previous point of the grid, i.e. $X_{t_{1,1}}^{(1)} = X_{t_{0,1}}^{(1)}$. As a consequence, at the frequency $n = 1$, a zero return over the time interval $[0, 1]$ is observed. Doubling the sampling frequency, i.e. going to $n = 2$, equation (3) implies that $\mathbb{B}_{t_{2,2}} = \mathbb{B}_{t_{1,1}} = \mathbb{B}_1 = 1$, but nothing more. That is, having observed at the frequency $n = 1$ a zero return over $[0, 1]$ does not imply that also the return in $[0, 1/2]$ is zero. In fact, the price process could have been observed, if sampled, at time $t = 1/2$, but since for $n = 1$ the partition contains only the points $\{t_{0,1} = 0, t_{1,1} = 1\}$, this event is ignored, because the instant $t = 1/2$ is not part of it.

Second, it implies that the probability of a zero return on any time interval $[t_{j-1,n}, t_{j,n}]$ depends only on the terminal point $t_{j,n}$, and not on the distance between the two points. This assumption could, in principle, be relaxed, replacing the $\mathfrak{p}$ in equation (1) with a sequence of stochastic processes $\mathfrak{p}^{(n)}$ converging, in some sense, to $\mathfrak{p}$. Here we abide by Assumption 1, avoiding further complications, since the asymptotic distributions of all the quantities of interest would be driven, in any case, by $\mathfrak{p}$. Frequency-dependent returns could be accommodated following other modelling choices. For example, one could assume that, for all $j = 0, ..., n$, the

probability of the event $\{\mathbb{B}_{j,n} = 1\}$ is given by

$$\mathfrak{p}_{t_{j,n}, \Delta_n} = \mathbb{E}\left[\mathbb{B}_{t_{j,n}}\right] = \mathbb{E}\left[\exp\left(-\int_{t_{j-1,n}}^{t_{j,n}} \lambda_t \, dt\right)\right], \tag{4}$$

for a suitable positive stochastic process $\lambda_t \geq 0$, which would play the role of a time-varying staleness intensity. It is immediately clear that equation (4) implies that

$$\frac{1}{n}\sum_{j=0}^{n} \mathbb{B}_{t_{j,n}} \xrightarrow{p} 1,$$

forcing, accordingly, the $\overline{\mathfrak{p}}$ in equation (2) to be exactly equal to one. Although it would be interesting to develop the intensity-based approach, we leave it for future projects, since in the present paper we focus on instantaneous price staleness.

Assumption 1 encompasses different specifications of $(\mathbb{B}_{j,n})_{j=1,\dots,n}$. If $\mathfrak{p}_t = \mathfrak{p}_0$ $\forall t \in [0,1]$, then the Bernoulli variates are *i.i.d.* with $\mathbb{P}[\mathbb{B}_{j,n} = 1] = \mathbb{E}[\mathfrak{p}_0]$. Another (more sophisticated) specification is obtained when $(\mathfrak{p}_t)_{t \in [0,1]}$ 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 $\mathfrak{p}$ is either constant (left panel) or a Brownian semimartingale (right panel). Although the number of zero returns (whose location is indicated 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.
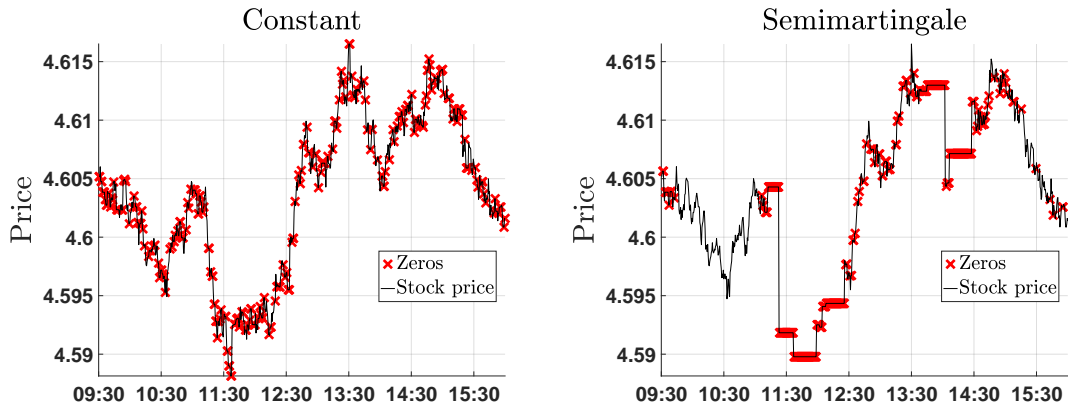
Figure 1: Shows two examples of price staleness. Zero returns are indicated by red crosses. Instantaneous staleness is either constant (left panel) or follows a semimartingale model (right panel). The total number of zeros is the same in both cases.

3 Asymptotic results

In this section, we derive a (stable) CLT for the fraction of zero returns within one day. This quantity is the natural estimator of the integrated (over one day) instantaneous staleness $\int_0^1 \mathfrak{p}_s \, ds$. As we will discuss below, the derivation of CLTs for the estimators of random variables of the form $\int_0^1 \mathfrak{p}_s^m \, ds$, with m any strictly positive integer, is required to define a test capable of distinguishing between a constant and a time-varying $\mathfrak{p}$. Nevertheless, as it happens for stable CLTs in the literature of integrated volatility estimators (see, among others, Kinnebrock and Podolskij, 2008; Podolskij and Vetter, 2009, 2010), an additional assumption on the dynamics of $\mathfrak{p}$ is required. In this paper, we stay with semimartingale dynamics for $\mathfrak{p}$.

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

$$\mathfrak{p}_t = \mathfrak{p}_0 + \int_0^t \mu_s \, ds + \int_0^t \nu_s \, dW_s, \tag{5}$$

where W_t is a standard Brownian motion, and μ_t and ν_t are adapted càdlàg processes, such that $\forall t$, $\mathfrak{p}_t \in (0,1)$ almost surely.

The following remark clarifies that the class of semimartingales whose values are confined within the open interval $(0,1)$ is not empty.

Remark 1. *Let $\mathsf{G} : \mathbb{R} \to (0,1)$ be a twice differentiable function and $(\chi_t)_{t \in [0,1]}$ be a Brownian semimartingale described by the following SDE*

$$\chi_t = \chi_0 + \int_0^t a_s \, ds + \int_0^t b_s \, dZ_s,$$

where Z is a $\mathcal{F}$-Brownian motion and the processes a and b are cádlág and $\mathcal{F}$-adapted. Consider the process defined as $\mathfrak{p}_t \doteq \mathsf{G}(\chi_t)$, for each $t \in [0,1]$. Hence, by construction, $\mathfrak{p}_t \in (0,1)$ for all t. By Itô lemma $(\mathfrak{p}_t)_{t \in [0,1]}$ is itself a Brownian semimartingale. In fact

$$\mathfrak{p}_t = \mathfrak{p}_0 + \int_0^t \left(a_s \, \mathsf{G}'(\chi_s) + \frac{1}{2} b_s^2 \, \mathsf{G}''(\chi_s) \right) ds + \int_0^t b_s \, \mathsf{G}'(\chi_s) \, dZ_s.$$

Remark 2. *Assumption 2 is not exhaustive for the asymptotic theory derived. For instance, one may allow the dynamics of the process $(\mathfrak{p}_t)_{t \in [0,1]}$ to include a jump component or even to be a pure-jump process. Investigating the jump dynamics of $(\mathfrak{p}_t)_{t \in [0,1]}$ would provide a number of interesting research questions. We defer the generalization of Assumption 2 to future projects.*

Having, as anticipated, the necessity to define an estimator of functionals of the form $\int_0^1 \mathfrak{p}_s^m \, ds$, in what follows we develop a theory of estimation of all integrals of the type $\int_0^1 f(\mathfrak{p}_s) \, ds$, 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 $\mathfrak{p}$ and an alternative in which the instantaneous staleness varies during the day. Finally, under Assumption 2, we derive a consistent estimator of the integrated volatility of $\mathfrak{p}$, that is, $\int_0^1 \nu_s \, ds$.

3.1 Realized staleness

We define *realized staleness* as[6]

$$\mathsf{RZ}_n \doteq \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\left\{ X_{j,n}^{(n)} - X_{j-1,n}^{(n)} = 0 \right\}}, \tag{6}$$

where $X^{(n)}$ is the observed price process defined in equation (1). Despite its simplicity, RZ_n encompasses economically meaningful features of the data-generating process of financial asset prices. Here, we are not going to discuss this point further and we refer to the paper by Bandi et al. (2019) for additional discussions. We focus instead on the limiting properties of RZ_n, which are exposed in the following theorem.

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

$$\mathsf{RZ}_n \xrightarrow{p} \int_0^1 \mathfrak{p}_s \, ds.$$

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

$$\sqrt{n} \left(\mathsf{RZ}_n - \int_0^1 \mathfrak{p}_s \, ds \right) \overset{stably}{\Longrightarrow} \mathcal{MN} \left(0, \Sigma_{\mathsf{RZ}} \right), \tag{7}$$

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

$$\Sigma_{\mathsf{RZ}} = \int_0^1 \mathfrak{p}_s \left(1 - \mathfrak{p}_s \right) \, ds. \tag{8}$$

Proof. See Appendix A.2. □

---

[6]There is a subtle difference between the definition in equation (6) and the idle time 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 convergence in probability implies that RZ_n is a consistent estimator of the integrated instantaneous staleness over one trading day, and this result holds under very general assumptions on the dynamics of the process $\mathfrak{p}$. In fact, more general uniform convergence on compacts in probability of RZ_n can be established, as is done in Theorem 4.1 below.

Under Assumption 2, the difference $\mathsf{RZ}_n - \int_0^1 \mathfrak{p}_s \, ds$ converges stably, at rate $n^{1/2}$, to a zero-mean (mixed) normal distribution whose variability has an intuitive expression. Indeed, in the case $\mathfrak{p}_t = \mathfrak{p}_0 \, \forall \, t \in [0,1]$, the asymptotic variance coincides (given the independence of the driving Bernoulli variates) with the variance of a Bernoulli random variable with (random) mean $\mathfrak{p}_0$, that is, $\mathfrak{p}_0 \, (1 - \mathfrak{p}_0)$. In the non-constant case, the expression of the asymptotic variance naturally generalizes to its integral version.

Since RZ_n is a consistent estimator of $\int_0^1 \mathfrak{p}_s \, ds$, a feasible confidence interval for RZ_n can be defined once a consistent estimator of $\int_0^1 \mathfrak{p}_s^2 \, ds$ is available. Actually, we consider the more general problem of developing a consistent estimator of $\int_0^1 (\mathfrak{p}_s)^m \, ds$ for some integers $m \geq 2$. For this purpose, we introduce *m-realized staleness* as

$$\mathsf{RZ}_{m,n} \doteq \frac{1}{n-m} \sum_{j=1}^{n-m} \prod_{q=0}^{m-1} \mathbb{1}_{\left\{ X_{j+q,n}^{(n)} - X_{j+q-1,n}^{(n)} = 0 \right\}}. \tag{9}$$

The rationale of the estimator is the following. Consider, for a fixed $m \geq 2$ and $j \in \{1, \ldots, n-m\}$, the product of the indicator functions that appear in equation (9). 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 $\mathsf{RZ}_{m,n}$. When Bernoulli variates are *i.i.d.*, $\mathsf{RZ}_{m,n}$ 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*

$$\mathsf{RZ}_{m,n} \xrightarrow{p} \int_0^1 \mathfrak{p}_s^m \, ds.$$

Moreover, as $n \to \infty$

$$\sqrt{n} \begin{bmatrix} \mathsf{RZ}_n - \int_0^1 \mathfrak{p}_s \, ds \\ \mathsf{RZ}_{m,n} - \int_0^1 \mathfrak{p}_s^m \, ds \end{bmatrix} \xRightarrow{stably} \mathcal{MN}\left(0, \Sigma_m\right),$$

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

$$\Sigma_m = \begin{bmatrix} \int_0^1 \mathfrak{p}_s \left(1 - \mathfrak{p}_s\right) ds & \int_0^1 m\, \mathfrak{p}_s^m \left(1 - \mathfrak{p}_s\right) ds \\ \int_0^1 m\, \mathfrak{p}_s^m \left(1 - \mathfrak{p}_s\right) ds & \int_0^1 \mathfrak{p}_s^m \, \frac{\mathfrak{p}_s^m \left(2\,m+1\right) - \mathfrak{p}_s^{m+1} \left(2\,m-1\right) - \left(1+\mathfrak{p}_s\right)}{1-\mathfrak{p}_s} \, ds \end{bmatrix}.$$

Proof. See Appendix A.2, Lemma 6. □

The estimation of the entries of the matrix Σ_m requires a consistent estimator of functionals of the form

$$U\left(f\right) = \int_0^1 f\left(\mathfrak{p}_s\right) ds, \tag{10}$$

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

3.2 Local estimation of instantaneous staleness

The estimation of functionals of the type (10) is feasible once a local estimator of $\mathfrak{p}$ 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 estimator

$$\widehat{\mathfrak{p}}_j\left(k_n\right) = \frac{1}{k_n} \sum_{\ell=0}^{k_n-1} \mathbb{1}_{\left\{X_{j+\ell+1,n}^{(n)} - X_{j+\ell,n}^{(n)} = 0\right\}}, \quad j \in \{1, \ldots, n - k_n\}. \tag{11}$$

Note that the condition $k_n \Delta_n \to 0$ ensures that the number of observations used to perform the local average in (11) is an order of infinity smaller than the sampling frequency of the Bernoulli variates, which diverges as $1/\Delta_n$. The functional $U\left(f\right)$ can then be estimated via standard Riemann sums, in which instantaneous staleness $\mathfrak{p}$ is replaced by the estimator in (11). For this reason, we define the discretized version of $U\left(f\right)$ as

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

and we derive its asymptotic properties in the following theorem.

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

$$U\left(\Delta_n, f\right)^n \xrightarrow{p} \int_0^1 f\left(\mathfrak{p}_s\right) ds. \tag{12}$$

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 (12) 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 $\mathfrak{p}$. 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 $\mathfrak{p}$ 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[7] version of $U(\Delta_n, f)^n$ takes the form

$$
U'(\Delta_n, f)^n = \Delta_n \sum_{j=1}^{n-k_n+1} \left(f\left(\widehat{\mathfrak{p}}_j(k_n)\right) - \frac{1}{2\,k_n} f''\left(\widehat{\mathfrak{p}}_j(k_n)\right) \widehat{\mathfrak{p}}_j(k_n) \left(1 - \widehat{\mathfrak{p}}_j(k_n)\right) \right)
$$

(13)

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:*

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

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

$$
\frac{1}{\sqrt{\Delta_n}} \left(U'(\Delta_n, f)^n - \int_0^1 f(\mathfrak{p}_s)\, ds \right) \xRightarrow{stably} \mathcal{MN}(0, \Sigma_{\mathrm{U}}),
$$

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

$$
\Sigma_U = \int_0^1 f'(\mathfrak{p}_s)^2\, \mathfrak{p}_s\, (1 - \mathfrak{p}_s)\, ds.
$$

Proof. See Appendix A.3. □

---

[7] Notice that the form of the bias is analogous to that of Jacod and Rosenbaum (2013), equation (3.8). In both cases, this bias is due to the local estimation of the instantaneous staleness.

3.3 On the estimation of the volatility of instantaneous staleness

In this section, we prove that under the semimartingale model of Assumption 2, it is possible to define a feasible and consistent estimator of $\bar{\nu} \doteq \int_0^1 \nu_s^2 \, ds$, that is, the quadratic variation of $\mathfrak{p}$. From an economic point of view, interpreting $\mathfrak{p}$ as an illiquidity proxy,[8] $\bar{\nu}$ is readily interpretable as a measure of volatility of illiquidity. If $\mathfrak{p}$ was observed, the natural estimator for the quadratic variation of the semimartingale in (5) would be

$$\sum_{j=1}^n \left(\Delta_j^n \mathfrak{p} \right)^2,$$

where $\Delta_j^n \mathfrak{p} \doteq \mathfrak{p}_{j,n} - \mathfrak{p}_{j-1,n}$. However, the process $\mathfrak{p}$ is latent, and therefore a proxy for the discrete increments $\Delta_j^n \mathfrak{p}$ is needed. The estimator in equation (11) 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*

$$\widehat{\bar{\nu}}_n^\star(k_n) \doteq k_n^{-1} \sum_{j=1}^{n-2k_n+1} \left(\widehat{\mathfrak{p}}_{j+k_n}(k_n) - \widehat{\mathfrak{p}}_j(k_n) \right)^2 \xrightarrow{p} \frac{2}{3} \int_0^1 \nu_s^2 \, ds + \frac{2}{\theta^2} \int_0^1 \mathfrak{p}_s \left(1 - \mathfrak{p}_s \right) \, ds,$$

$$(14)$$

where $\widehat{\mathfrak{p}}_j(k_n)$ is the estimator defined in equation (11).

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 $\widehat{\mathfrak{p}}_j(k_n)$ converges in probability to two-thirds of the integrated volatility of $\mathfrak{p}$ plus a bias term that is proportional to the asymptotic variance of RZ_n (see Theorem 3.2). Nevertheless, this bias does not constitute an issue because it can be consistently estimated via $U''(\Delta_n, f)^n$ using a suitable f. Indeed, a consistent estimator of $\bar{\nu}$ can be defined as:

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

$$(15)$$

where $f(x) = x(1-x)$. Note that, by construction, it is not guaranteed that

---

[8]The interpretation of $\mathfrak{p}$ 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 micro-structural theories of price formation with transaction costs and asymmetries in information. In particular, the probability of the occurrence of a zero return is driven by $\mathfrak{p}$.

$\widehat{\overline{\nu}}_n \geq 0$. To circumvent this problem, one might use $\max\left(\widehat{\overline{\nu}}_n, 0\right)$ instead as a non-negative estimator.

4 Statistical tests

4.1 A test for time-varying versus constant instantaneous staleness

The inferential theory discussed thus far is the basis for a statistical test apt to discriminate, over one day of observation, between a constant and a time-varying instantaneous staleness. Consider, therefore, the following partition of the sample space Ω:

$$\begin{aligned}
\Omega_0 &= \left\{\omega \in \Omega \;\middle|\; \int_0^1 (\mathfrak{p}_t(\omega))^m \, dt = \left(\int_0^1 \mathfrak{p}_t(\omega) \, dt\right)^m\right\}, \\
\Omega_1 &= \left\{\omega \in \Omega \;\middle|\; \int_0^1 (\mathfrak{p}_t(\omega))^m \, dt \neq \left(\int_0^1 \mathfrak{p}_t(\omega) \, dt\right)^m\right\},
\end{aligned} \tag{16}$$

where m is an arbitrary integer. If a path $\mathfrak{p}_t(\omega)$ is constant on $[0,1]$ then $\omega \in \Omega_0$. On the other hand, if $\omega \in \Omega_1$ then the corresponding trajectory $\mathfrak{p}_t(\omega)$ must depart from a constant path in, at least, a subset of the interval $[0,1]$ with a non-zero Lebesgue measure. In particular, the null hypothesis is formally defined as $\mathcal{H}_0 : (\omega \in \Omega_0)$, whereas the alternative is $\mathcal{H}_1 : (\omega \in \Omega_1)$. By virtue of Theorem 3.1, Theorem 3.2, and the delta method, the random variable defined as

$$\Psi_{m,n} \doteq \frac{\sqrt{n}\left(\mathsf{RZ}_{m,n} - (\mathsf{RZ}_n)^m\right)}{\sqrt{\frac{(\mathsf{RZ}_n)^{2m+1}(m^2+2m-1) - (\mathsf{RZ}_n)^{2m}(2m^2+2m+1) + (\mathsf{RZ}_n)^{m+1} + (\mathsf{RZ}_n)^m}{\mathsf{RZ}_n - 1}}} \tag{17}$$

is the natural candidate for a test statistic that may asymptotically distinguish whether the observed instantaneous staleness stems from a $\mathfrak{p}$ in Ω_0 or in Ω_1. The asymptotic limits of the $\Psi_{m,n}$ test statistic are discussed in the following corollary.

Corollary 1. *As $n \to \infty$ it holds that*

$$\begin{cases}
\Psi_{m,n} \xRightarrow{stably} \mathcal{N}(0,1) & \text{on } \Omega_0, \\
\Psi_{m,n} \xrightarrow{p} +\infty & \text{on } \Omega_1.
\end{cases}$$

Proof. See Appendix A.5. $\qquad\qquad\square$

Under the null, the limiting null distribution of $\Psi_{m,n}$ coincides with that of the zero-mean normal random variable with unit variance while, under the alternative,

15

the test statistic diverges in probability, thus delivering a unit power. Note that the asymptotic properties of $\Psi_{m,n}$ 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 5, which is dedicated to the Monte Carlo simulations.

4.2 A test for deterministic instantaneous staleness

A rejection of the null hypothesis $\mathcal{H}_0 : (\omega \in \Omega_0)$ poses the question of which kind of intra-day variability is observed for the instantaneous staleness $\mathfrak{p}$. Since staleness is expected to be inversely related to volatility, an intra-day non-constant path for $\mathfrak{p}$ could be simply the result of an inverse U-shaped seasonal pattern. Alternatively, the intra-day variation could be the outcome of random shocks, superimposed on a deterministic seasonal pattern.

To discriminate between the two scenarios, we develop a non-parametric methodology which takes cues from the work of Christensen et al. (2018), in which the authors investigate the source of variation in intra-day volatility. The major challenge in designing the new test stems from the fact that $\mathfrak{p}$, as for the case of spot volatility, is not directly observed and both the (potential) stochastic and deterministic components of $\mathfrak{p}$ can change over the course of the day. To circumvent these problems, following Christensen et al. (2018), we resort to a long-span asymptotic theory, as opposed to the infill asymptotic framework used so far. Accordingly, some additional notation is needed. Let T be the number of days in the sample and let $[t-1, t]$ be the sub-interval indicating the t-th day. We assume that the processes are observed over $[0, T]$ with observations recorded on a refining partition $\Pi_n = \{t_{0,n}, \ldots, t_{nT,n}\}$, where $t_{j,n} = j/n$, $j = 0, \ldots, nT$. For a generic process X, let $X_{j,n}^{(t)}$ indicate the observation of X at point j/n of day t. Hereafter, $X_{j,n}^{(1)}$ is denoted as $X_{j,n}$.

The philosophy of the test reflects what it is usually done in the volatility literature (see, for instance, Engle and Sokalska, 2012). That is, we assume that at each time instant $s \in (t-1, t)$ the value of the process $\mathfrak{p}_s$ is equal to the product of a deterministic recurrent factor, say ϑ_s, and an intra-day stochastic part, say $\mathfrak{p}_s^\star$; see Assumption 3 below. We formulate a null hypothesis in which $\mathfrak{p}_s^\star$ is constant during the day. Under the alternative, instead, $\mathfrak{p}_s^\star = \mathfrak{p}_s / \vartheta_s$ is allowed to be time-varying and, since the deterministic intra-day variation of $\mathfrak{p}_s$ is assumed to be fully absorbed by ϑ_s, this scenario corresponds to an instantaneous staleness driven also by random shocks.

We put now some structure on the problem. We start from an additional assumption.

Assumption 3. *The process $\mathfrak{p}_t$ is equal to $\mathfrak{p}_t = \mathfrak{p}_t^\star \vartheta_t$, where $\mathfrak{p}_t^\star$ denotes the stochastic part of $\mathfrak{p}_t$ and ϑ_t the deterministic seasonal component. Moreover:*

(a) *The process $\mathfrak{p}_t^\star$ satisfies Assumption 2 and assumes values in the interval $\left(0, \bar{\vartheta}\right)$, with $\bar{\vartheta} = \left(\sup_{s \in [0,1]} \vartheta_s\right)^{-1}$. In addition, it is stationary and such that $\mathbb{E}\left[\mathfrak{p}_t^\star\right] = \pi^\star$ and $\sum_{k=1}^{\infty} \mathbb{C}ov\left[\mathfrak{p}_t^\star, \mathfrak{p}_{t+k}^\star\right] < \infty$.*

(b) *The diurnal component ϑ_t is a continuously differentiable one-periodic bounded function, bounded away from zero, with bounded derivative for $t \to 0$ and $t \to 1$. Moreover, it is normalized such that $\int_0^1 \vartheta_s \, ds = 1$.*

Some observations are in order. Point (a) ensures the stationarity of $\mathfrak{p}_t^\star$ as well as restricting its memory, which implies that the process is ergodic (Christensen et al., 2018). On the other hand, point (b) allows for the identification of both components of $\mathfrak{p}_t$ and, in particular, they allow the estimation of the diurnal component from a long sample of data. Indeed, let $s \in (0, 1)$ be a fixed time of the day. Then, we have

$$\frac{1}{T} \sum_{t=0}^{T} \mathfrak{p}_{t+s} = \frac{1}{T} \sum_{t=0}^{T} \mathfrak{p}_{t+s}^\star \vartheta_{t+s} = \vartheta_s \frac{1}{T} \sum_{t=0}^{T} \mathfrak{p}_{t+s}^\star \xrightarrow{p} \vartheta_s^\star \doteq \vartheta_s \pi^\star.$$

Then we define, for some integer $m \geq 2$, the *diurnal corrected m-realized staleness* (at day t) as:

$$\widetilde{\mathsf{RZ}}_{m,n}^{(t)} \doteq \frac{1}{n-m} \sum_{j=1}^{n-m+1} \prod_{q=0}^{m-1} \frac{\mathbb{1}_{\left\{X_{j+q,n}^{(t)} - X_{j+q-1,n}^{(t)} = 0\right\}}}{\widehat{\vartheta}_{j+q,n}^\star}, \tag{18}$$

where $\widehat{\vartheta}_{j,n}^\star$ is the long-scale estimator of $\vartheta_{j,n}^\star$ and it is defined as:

$$\widehat{\vartheta}_{j,n}^\star \doteq \frac{1}{T} \sum_{t=1}^{T} \mathbb{1}_{\left\{X_{j,n}^{(t)} - X_{j-1,n}^{(t)} = 0\right\}}. \tag{19}$$

We derive the limiting properties of $\widetilde{\mathsf{RZ}}_{m,n}^{(t)}$ in the following theorem.

Theorem 4.1. *Let Assumption 3 hold. Suppose that $n/T \longrightarrow 0$ as $n, T \to \infty$. Then we have*

$$\widetilde{\mathsf{RZ}}_{m,n}^{(t)} \xrightarrow{u.c.p.} \frac{1}{(\pi^\star)^m} \int_{t-1}^{t} (\mathfrak{p}_s^\star)^m \, ds,$$

and

$$\sqrt{n}\begin{bmatrix}\widetilde{\mathsf{RZ}}^{(t)}_{1,n} - \frac{1}{\pi^\star}\int_{t-1}^{t}\mathfrak{p}^\star_s\,ds\\[4pt]\widetilde{\mathsf{RZ}}^{(t)}_{m,n} - \frac{1}{(\pi^\star)^m}\int_{t-1}^{t}(\mathfrak{p}^\star_s)^m\,ds\end{bmatrix}\overset{stably}{\Longrightarrow}\mathcal{MN}\left(0,\widetilde{\Sigma}^{(m)}\right),$$

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

$$\widetilde{\Sigma}^{(m)}=\begin{bmatrix}\int_{t-1}^{t}\frac{\left(\mathfrak{p}^\star_s\vartheta_s^{-1}-(\mathfrak{p}^\star_s)^2\right)}{(\pi^\star)^2}\,ds & \int_{t-1}^{t}\frac{m\,(\mathfrak{p}^\star_s)^m\left(\vartheta_s^{-1}-\mathfrak{p}^\star_s\right)}{(\pi^\star)^{m+1}}\,ds\\[8pt]\int_{t-1}^{t}\frac{m\,(\mathfrak{p}^\star_s)^m\left(\vartheta_s^{-1}-\mathfrak{p}^\star_s\right)}{(\pi^\star)^{m+1}}\,ds & \int_{t-1}^{t}\sum_{\ell=0}^{m-1}\frac{(2m-2\ell-1)(\mathfrak{p}^\star_s)^{2m-\ell-1}\left(\vartheta_s^{-1}-\mathfrak{p}^\star_s\right)\vartheta_s^{-\ell}}{(\pi^\star)^{2m}}\,ds\end{bmatrix}.$$

Proof. See Appendix A.5. □

Theorem 4.1 shows that the m-realized staleness "scaled" by the diurnal effect converges, uniformly in probability, to a quantity that depends solely from the stochastic component of $\mathfrak{p}$. As for the limiting distribution, instead, note that the entries of $\widetilde{\Sigma}^{(m)}$ depend on ϑ_s. However, they can be estimated consistently by the following estimators

$$\widehat{\widetilde{\Sigma}}^{(m)}_{(1,1)} = \frac{1}{n}\sum_{j=1}^{n}\frac{\left(\widehat{\vartheta}^\star_n\,\widetilde{\mathsf{RZ}}^{(t)}_{1,n}\left(\frac{\widehat{\vartheta}^\star_{j,n}}{\widehat{\vartheta}^\star_n}\right)^{-1}-\left(\widehat{\vartheta}^\star_n\widetilde{\mathsf{RZ}}^{(t)}_{1,n}\right)^2\right)}{\left(\widehat{\vartheta}^\star_n\right)^2},$$

$$\widehat{\widetilde{\Sigma}}^{(m)}_{(1,2)} = \frac{1}{n}\sum_{j=1}^{n}\frac{m\left(\widehat{\vartheta}^\star_n\,\widetilde{\mathsf{RZ}}^{(t)}_{1,n}\right)^m\left(\left(\frac{\widehat{\vartheta}^\star_{j,n}}{\widehat{\vartheta}^\star_n}\right)^{-1}-\widehat{\vartheta}^\star_n\,\widetilde{\mathsf{RZ}}^{(t)}_{1,n}\right)}{\left(\widehat{\vartheta}^\star_n\right)^{m+1}},$$

$$\widehat{\widetilde{\Sigma}}^{(m)}_{(2,2)} = \frac{1}{n}\sum_{j=1}^{n}\sum_{\ell=0}^{m-1}\frac{(2\,m-2\,\ell-1)\left(\widehat{\vartheta}^\star_n\,\widetilde{\mathsf{RZ}}^{(t)}_{1,n}\right)^{2m-\ell-1}\left(\left(\frac{\widehat{\vartheta}^\star_{j,n}}{\widehat{\vartheta}^\star_n}\right)^{-1}-\widehat{\vartheta}^\star_n\,\widetilde{\mathsf{RZ}}^{(t)}_{1,n}\right)\left(\frac{\widehat{\vartheta}^\star_{j,n}}{\widehat{\vartheta}^\star_n}\right)^{-\ell}}{\left(\widehat{\vartheta}^\star_n\right)^{2m}},$$

$$(20)$$

where, to avoid clutter, we set $\widehat{\vartheta}^\star_n = \frac{1}{n}\sum_{j=1}^{n}\widehat{\vartheta}^\star_{j,n}$.

To test wether the diurnal pattern is sufficient to explain the observed intra-day variation in $\mathfrak{p}$, we partition the sample space Ω into

$$\Omega^\star_{t,0}\doteq\left\{\omega\in\Omega\,\middle|\,\mathfrak{p}^\star_s(\omega)\text{ is constant for }s\in[t-1,t]\right\}$$

and $\Omega^\star_{t,1}$, defined as the complement of $\Omega^\star_{t,0}$. Therefore, the null hypothesis is formally defined as $\mathcal{H}^\star_0:\left(\omega\in\Omega^\star_{t,0}\right)$, whereas the alternative is $\mathcal{H}^\star_1:\left(\omega\in\Omega^\star_{t,1}\right)$.

An implication of Theorem 4.1 is that for any integer $m \geq 2$ as $n \to \infty$

$$\widetilde{\mathsf{RZ}}_{m,n}^{(t)} - \left(\widetilde{\mathsf{RZ}}_{1,n}^{(t)} \right)^m \xrightarrow{u.c.p.} \frac{1}{(\pi^\star)^m} \left(\int_{t-1}^t (\mathfrak{p}_s^\star)^m \, ds - \left(\int_{t-1}^t \mathfrak{p}_s^\star \, ds \right)^m \right),$$

which means that, under the null $\mathcal{H}_0^\star$, the limit of the difference $\widetilde{\mathsf{RZ}}_{m,n}^{(t)} - \left(\widetilde{\mathsf{RZ}}_{1,n}^{(t)} \right)^m$ is zero while, under the alternative $\mathcal{H}_1^\star$, it is a strictly positive quantity. We thus build a test for $\mathcal{H}_0^\star$ against $\mathcal{H}_1^\star$ via the following statistic:

$$\Upsilon_{m,n}^{(t)} \doteq \frac{\sqrt{n} \left(\widetilde{\mathsf{RZ}}_{m,n}^{(t)} - \left(\widetilde{\mathsf{RZ}}_{1,n}^{(t)} \right)^m \right)}{\sqrt{m^2 \left(\widetilde{\mathsf{RZ}}_{1,n}^{(t)} \right)^{2(m-1)} \widehat{\widehat{\Sigma}}_{(1,1)}^{(m)} + \widehat{\widehat{\Sigma}}_{(2,2)}^{(m)} - 2m \left(\widetilde{\mathsf{RZ}}_{1,n}^{(t)} \right)^{m-1} \widehat{\widehat{\Sigma}}_{(1,2)}^{(m)}}},$$

where $\widehat{\widehat{\Sigma}}_{(1,1)}^{(m)}$, $\widehat{\widehat{\Sigma}}_{(2,2)}^{(m)}$ and $\widehat{\widehat{\Sigma}}_{(1,2)}^{(m)}$ are defined in equation (20). The asymptotic limits of $\Upsilon_{m,n}^{(t)}$ are described in the following corollary.

Corollary 2. *As $n \to \infty$, $T \to \infty$ and $\frac{n}{T} \to 0$,*

$$\begin{cases} \Upsilon_{m,n}^{(t)} \xoverset{stably}{\Longrightarrow} \mathcal{N}(0,1) & on \quad \Omega_{t,0}^\star, \\ \Upsilon_{m,n}^{(t)} \xrightarrow{p} +\infty & on \quad \Omega_{t,1}^\star. \end{cases}$$

Proof. See Appendix A.5. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Therefore, the limiting null distribution of $\Upsilon_{m,n}^{(t)}$ coincides with that of a standard normal while, under the alternative hypothesis, the test statistic diverges in probability. In the next section we will discuss how, in finite sample, the parameter m affects the size and power of both $\Psi_{m,n}$ and $\Upsilon_{m,n}^{(t)}$.

5 Monte Carlo simulations

The settings in force of electronic financial markets force transaction prices to round off at one cent. Hence, in real data, price discreteness may affect the limits derived in Section 3 and Section 4, producing unwanted spurious effects.

More precisely, in the presence of rounding, there could be some extra zero returns not generated by the stochastic process $\mathfrak{p}$, defined in Assumption 1. In this section, we explore the finite sample contaminations of the asymptotic theory by means of Monte Carlo simulations. For this purpose, we generate an artificial dataset of efficient price paths contaminated by staleness and rounded at one cent. For each replication we simulate a trading day of 6.5 hours on a grid 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 is driven by the SDE

$$
\begin{aligned}
d \log \sigma_t^2 &= \left(\alpha - \beta \log \sigma_t^2 \right) dt + \eta \, dW_{\sigma,t}, \\
dY_t &= \mu \, dt + c_\sigma \, \sigma_t \, dW_{Y,t},
\end{aligned}
\tag{21}
$$

where $W_{\sigma,t}$ and $W_{Y,t}$ are two Brownian motions with $\mathrm{corr}\left(dW_{\sigma,t}, dW_{Y,t}\right) = \rho \, dt$. 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 $c_\sigma = 3$, unless otherwise specified. Numerical integration of the SDE in (21) 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 $\log \sigma_0^2 = \alpha/\beta$ (different simulation scenarios are considered in Appendix B). Once simulated, the efficient prices are sampled every 30 seconds. Then, on the time grid of 30 seconds we construct the staleness-contaminated log-price process $X_{j,n}$ following the recursive equation

$$
\begin{cases}
X_{0,n} = Y_{0,n} = \log(P_0), \\
X_{j,n} = (1 - \mathbb{B}_{j,n}) \, Y_{j,n} + \mathbb{B}_{j,n} \, X_{j-1,n},
\end{cases}
\tag{22}
$$

where $\mathbb{B}_{j,n}$ are Bernoulli random variables generated as described below. Finally, the prices $\exp(X_{j,n})$ are rounded at one cent. The rounding is the only factor that prevents taking the highest frequency available. In the numerical experiments that follow we will consider different dynamics for $\mathfrak{p}$.

$\Omega_{0,\Psi}$: **constant instantaneous staleness.** In this specification we assume that $\mathfrak{p}$ is constant. Therefore, this choice corresponds to the null hypothesis of the test $\Psi_{m,n}$. In this specification, the $\mathbb{B}_{j,n}$'s are *i.i.d.* Bernoulli random variables with constant expected value $\mathbb{E}\left[\mathbb{B}_{j,n}\right] = \mathfrak{p}_F$ for all j. We put $\mathfrak{p}_F = 0.5$. Hence, 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.

$\Omega_{0,\Upsilon}$: **deterministic time-varying instantaneous staleness.** We input a $\mathfrak{p}$ solely driven by a deterministic diurnal pattern. In this scenario, the simulated instantaneous staleness $\mathfrak{p}$, once standardized by the seasonal component, is constant and equal to $\mathfrak{p}_F = 0.5$. The deterministic pattern is estimated from real data employing the methodology described in Section 4.2. This hypothesis is, simultaneously, the null hypothesis for the test statistic $\Upsilon_{m,n}^{(t)}$ and one possible alternative

20

for the test statistic $\Psi_{m,n}$.

$\Omega_{1,\Upsilon}$: stochastic plus deterministic time-varying instantaneous staleness.
We input an instantaneous staleness expressed as the product of two factors, as
described in Assumption 3 of Section 4.2. The diurnal deterministic component ϑ_t
is inputted as in the the hypothesis $\Omega_{0,\Upsilon}$. The stochastic component $\mathfrak{p}^\star$ is simulated
with the following scheme. For each replication, we generate a path of a latent
stochastic process u with the following (discrete-time) integration scheme

$$\begin{cases} u_{0,n} = F^{-1}(\mathfrak{p}_F) \\ u_{j,n} = u_{j-1,n} + (F^{-1}(\mathfrak{p}_F) - u_{j-1,n})/n + \sigma_u \varepsilon_{j,n}/\sqrt{n}, \end{cases} \tag{23}$$

with $j = 1, ..., n$, $n = 780$, $\mathfrak{p}_F = 0.5$, 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 $\sigma_u = 1.5$.
The path of $\mathfrak{p}^\star$ is generated as

$$\mathfrak{p}_{j,n}^\star = \int_{-\infty}^{u_{j,n}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \, dz = F(u_{j,n}). \tag{24}$$

Note that since the process u is mean-reverting around $F^{-1}(\mathfrak{p}_F)$ then, by construc-
tion, $\mathfrak{p}^\star$ is mean-reverting around $\mathfrak{p}_F$. This hypothesis is an alternative for both
$\Upsilon_{m,n}^{(t)}$ and $\Psi_{m,n}$.

5.1 Sizes and powers of $\Psi_{m,n}$ and $\Upsilon_{m,n}^{(t)}$

The test statistics $\Psi_{m,n}$ and $\Upsilon_{m,n}^{(t)}$ are characterized, for a given frequency of observa-
tion n, by the choice variable m. Asymptotically, their distributions are unaffected
by the value of m, as well as their divergence toward $+\infty$ under the corresponding
alternative hypothesis. Nevertheless, in a finite sample, m can be chosen to trade-
off the size and power of the tests. Our Monte Carlo simulations[9] shows that, at the
frequency of one minute, $m = 4$ maximizes the rejection powers and simultaneously
preserves the size of the tests under the null. Hence, in what follows, we will always
adopt $m = 4$.

Figure 2 shows the kernel densities of $\Psi_{m,n}$ over 10^4 replications of prices under
the null $\Omega_{0,\Psi}$ and the two alternatives $\Omega_{0,\Upsilon}$ and $\Omega_{1,\Upsilon}$, described in the previous
section. As expected, the highest rejection rates are found in the case of $\Omega_{1,\Upsilon}$.
Price rounding, which is already included in the simulations of Figure 2, has the

---

[9]All details and routines are available upon request. They are omitted here for the sake of
conciseness.

effect of slightly distorting the distribution under the null (toward negative values) and to reduce the rejection power of the tests.[10]

Similarly, Figure 3, shows the kernel densities of $\Upsilon_{m,n}^{(t)}$ over 10^4 replications of prices under the null $\Omega_{0,\Upsilon}$ and the alternative $\Omega_{1,\Upsilon}$. As for the case of the $\Psi_{m,n}$ statistic, rounding has the effect of slightly distorting the distribution of the test under $\Omega_{0,\Upsilon}$ and to reduce the rejection power under $\Omega_{1,\Upsilon}$.
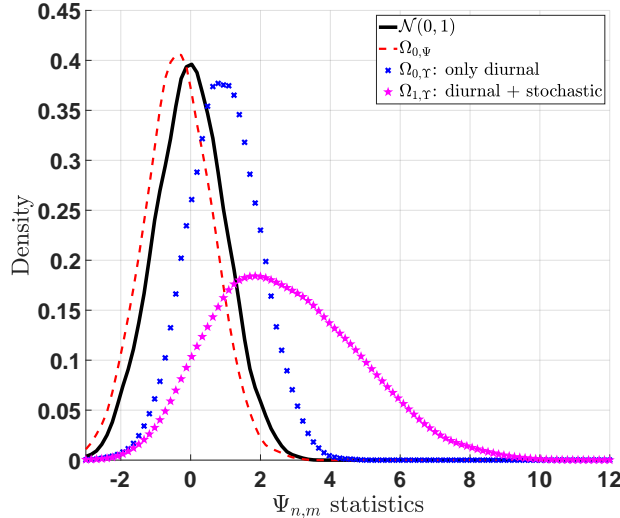
Figure 2: Kernel density of the test statistic $\Psi_{m,n}$ under the null $\Omega_{0,\Psi}$ (dotted red line) and the two alternatives $\Omega_{0,\Upsilon}$ (blue crosses) and $\Omega_{1,\Upsilon}$ (magenta stars). $n = 780$ and $m = 4$. Prices are rounded at one cent.

6   Empirical application

We employ a dataset whose constituents are the most 250 liquid[11] NYSE-listed stocks. The data range from January 3, 2006 to December 31, 2014, covering a time span of 2263 trading days. We rank the stocks in deciles according to the total volume traded (during the whole period available) and, within each decile, we select the stock with the highest average price. This choice is dictated by the necessity of mitigating the impact of rounding. This data filtering returns the ten tickers APA, BA, CVX, DE, EOG, GS, MCD, MMM, UNP, and XOM. Transaction prices are sampled every thirty seconds with previous-tick interpolation, producing 780 observations from 09:30 EST to 15:30 EST for each trading day.

---

[10]This claim is corroborated by comparing the results in Figure 2 with the corresponding plots obtained simulating prices without rounding. For conciseness, the results of these numerical experiments are omitted but are available upon request.

[11]In terms of average transaction volumes during the period considered.
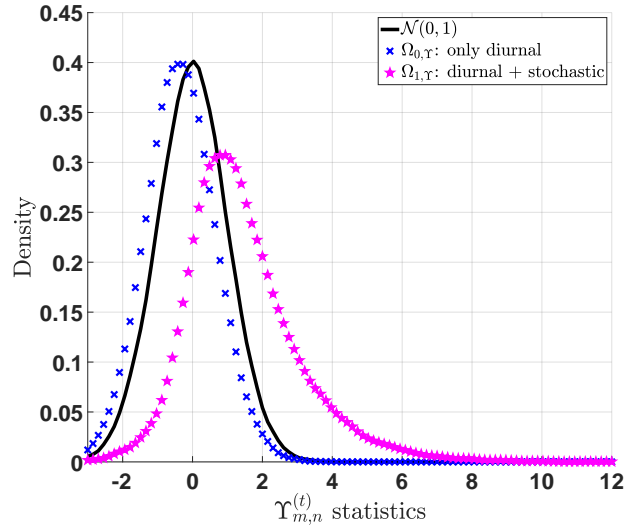
Figure 3: Kernel density of the test statistic $\Upsilon_{m,n}^{(t)}$ under the null $\Omega_{0,\Upsilon}$ (blue crosses) and the alternative $\Omega_{1,\Upsilon}$ (magenta stars). $n = 780$ and $m = 4$. Prices are rounded at one cent.
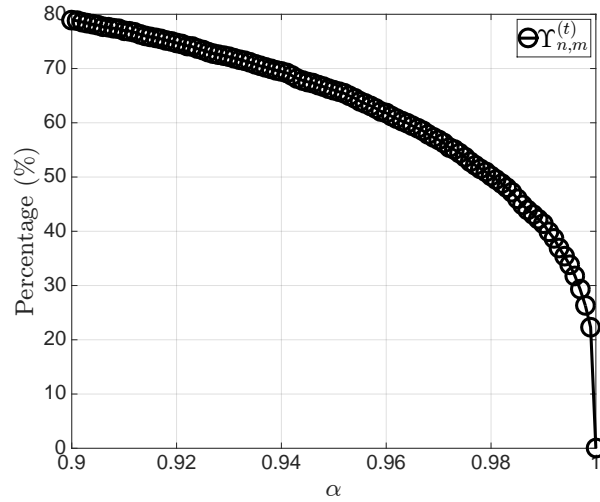
Figure 4: The black line reports the percentage of days (across the ten stocks) for which the null $\mathcal{H}_0$ is rejected through $\Psi_{m,n}$ at confidence level α. The red dotted line reports the percentage of days, among those selected rejecting the null $\mathcal{H}_0$ through $\Psi_{m,n}$ with 99% confidence level, for which the null $\mathcal{H}_0^\star$ is rejected through $\Upsilon_{m,n}^{(t)}$ at confidence level α.

We apply two quality cuts. In order to avoid spurious results due to scheduled market interruptions, we first remove all days with ten consecutive minutes of no trading activity. Then, we remove days with a realized staleness smaller than 20%. After these quality cuts, the final dataset is made of 11689 days out of $2263 \times 10 = 22630$ of the original sample.

23
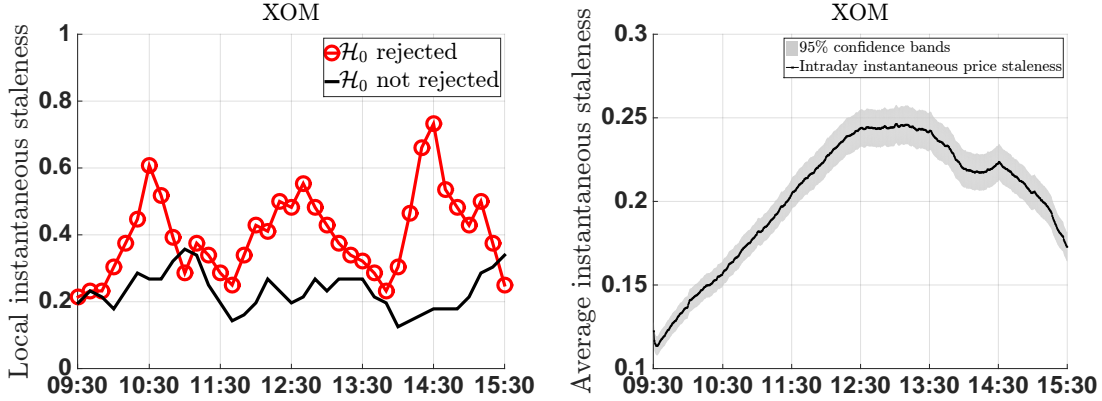
Figure 5: (Left panel) The red line with circles and the black continuous line represent, respectively, instantaneous price staleness of XOM during a day in which the null $\mathcal{H}_0$ of constant instantaneous price staleness is and is not rejected at 99% confidence level. (Right panel). Intra-day instantaneous price staleness estimates of XOM averaged over the whole sample available (black continuous line) and corresponding 95% confidence bands (grey shaded area).

First we test, for each day in the sample, for constancy of instantaneous price staleness using the test statistic $\Psi_{m,n}$. Then, on the sub-sample composed by the days for which $\mathcal{H}_0$ is rejected at 99% confidence level, we run the $\Upsilon_{m,n}^{(t)}$ test. Figure 4, black continuous line, shows the percentage of days, across the ten stocks, in which the null $\mathcal{H}_0$ is rejected at the confidence level α. At the 99% confidence level, the rejection rate is a remarkable 25%. This constitutes empirical evidence in favour of a time-varying model for $\mathfrak{p}$. The left panel of Figure 5 provides two examples of estimated (using the estimator defined in equation (11)) trajectories of the instantaneous staleness for a day in which the null $\mathcal{H}_0$ is rejected (red line with circles) and not rejected (black solid line), at the 99% confidence level. The difference in the dynamics of the two trajectories is striking.

Being related to trading activity, staleness is expected to be low at the beginning and the end of the trading day, and high in the middle. Indeed, right panel of Figure 5 illustrates the presence of a pronounced inverse U-shape pattern in the estimates of instantaneous staleness of XOM, averaged across the sample. Zero returns are mostly concentrated around midday: the average instantaneous staleness is equal to 0.12 at 09:30 EST, and increases up to 0.24 at 12:30 EST. Hence, we examine whether a deterministic recurrent pattern alone is sufficient to explain the intra-day variability of price staleness.

For this purpose, we apply the $\Upsilon_{m,n}^{(t)}$ test to a Ψ-*filtered sample* defined by selecting the days for which the null hypothesis of constant instantaneous staleness $\mathcal{H}_0$ is rejected at 99% confidence level. The sample consists of 3102 days out of the 11689 of the original sample. Figure 4, red dotted line, displays the percentage

of days, in the Ψ-*filtered sample*, for which the null hypothesis $\mathcal{H}_0^\star$ is rejected at the confidence level α. With a rejection rate that at the 99% confidence level is still 40%, this plot provides empirical evidence that the diurnal pattern alone is not sufficient to explain the intra-day variation of $\mathfrak{p}$. The left panel of Figure 6 reports an example of estimated instantaneous staleness $\mathfrak{p}$ (red markers) for a (randomly selected) day in which the null $\mathcal{H}_0^\star$ is not rejected with 99% confidence level. The black continuous line shows the (non-standardized) intra-day pattern of $\mathfrak{p}$. The right panel of Figure 6 illustrates a (randomly selected) day in which the null $\mathcal{H}_0^\star$ is rejected with the same confidence level. When $\mathcal{H}_0^\star$ is not rejected, the estimated instantaneous staleness almost coincides with the intra-day seasonal pattern. On the other hand, should the null $\mathcal{H}_0^\star$ be rejected as in the case of the right panel, the intra-day path of $\mathfrak{p}$ significantly deviates from the seasonal pattern. Overall, the analysis suggests that, although the diurnal pattern certainly constitutes an important driving factor, it is not sufficient to fully explain the intra-day variation of the staleness. A more complex modelling is required.
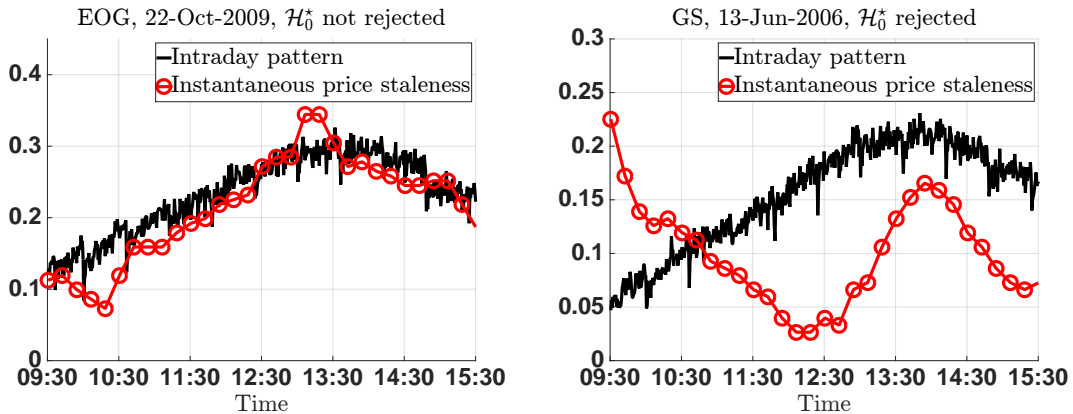
Figure 6: Examples of the estimated trajectories of $\mathfrak{p}_t$ (red markers) and the (non-standardized) intra-day pattern, ϑ_t (black continuous lines). Left panel: the null $\mathcal{H}_0^\star$ is not rejected at 99% confidence level. Right panel: the null $\mathcal{H}_0^\star$ is rejected at 99% confidence level.

7 Conclusions

Transaction prices of financial assets show, from moderate to high-frequency (e.g. from five-minute to higher frequency), a non-null probability of being repeated. This feature translates in an excess of zero (or small) returns with respect to what would be expected under semimartingale dynamics (Bandi et al., 2017).

Building on this point, we provided a new econometric framework, where the statistical properties of the likelihood of observing zero returns are driven by a

stochastic process $\mathfrak{p}$, which we refer to as instantaneous staleness. Once integrated over a given time horizon, the expected value of instantaneous staleness delivers the asymptotic average probability of observing a zero return, where the average is relative to the time window considered. Hence, instantaneous staleness can be interpreted as the localized version of average price staleness which, in turn, is consistently estimated by the realized staleness or, also, by the idle time, defined by Bandi et al. (2017).

Zero returns are naturally linked to lack of liquidity. The proposed framework thus allows to derive statistical inferences on liquidity-related variates. In particular, we derived a test statistic capable of distinguishing between a null in which the process $\mathfrak{p}$ is constant and an alternative in which it is time-varying. We show, on ten representative stocks of the NYSE, that the 99%-rejection rate of the statistic is a remarkable 25%. We investigated further the dynamics of instantaneous staleness, developing an asymptotic theory to test whether a deterministic recurrent pattern is sufficient to explain the time-variation in $\mathfrak{p}$. For this purpose, following Christensen et al. (2018), we derived a test statistic that is (stably) distributed as a standard normal random variable if $\mathfrak{p}$ coincides, up to a multiplicative constant, with a periodic deterministic function of time. Under the alternative, which encompasses all models in which the de-seasonalized instantaneous staleness is time-varying, the test diverges (in probability) to infinity. Running this newly defined test we found that the intra-day seasonal pattern is the leading effect in determining the time-variation of $\mathfrak{p}$. Nevertheless, it cannot explain all the variability in $\mathfrak{p}$ since, for example, at 99% confidence level the null is still rejected in 40% of the days considered. In summary, our empirical results point toward the necessity of investigating which dynamical model is more appropriate to explain the intra-day pattern of instantaneous staleness.

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. G., L. Benzoni, and J. Lund (2002). An empirical investigation of continuous-time equity return models. *The Journal of Finance 57*(3), 1239–1284.

Bandi, F. M., A. Kolokolov, D. Pirino, and R. Renò (2019). Zeros. *Forthcoming on Management Science*.

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

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 Realized Covariation: High Frequency Based Covariance, Regression and Correlation in Financial Economics. *Econometrica 72*(3), 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*(5), 796–806.

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 (2017). Systemic co-jumps. *Journal of Financial Economics 126*(3), 563–591.

Christensen, K., U. Hounyo, and M. Podolskij (2018). Is the diurnal pattern sufficient to explain intraday variation in volatility? A nonparametric assessment. *Journal of Econometrics 205*(2), 336–362.

Cont, R. and P. Tankov (2004). *Financial Modelling with Jump Processes*. Boca Raton, FL:Chapman & Hall/CRC.

Engle, R. F. and M. E. Sokalska (2012). Forecasting intraday volatility in the US equity market. Multiplicative component GARCH. *Journal of Financial Econometrics 10*(1), 54–83.

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 application.* 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. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *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 related functionals of semimartingales. *Stochastic Processes and their Applications 118*(4), 517–559.

Jacod, J. (2012). Statistics and high-frequency data. *Statistical Methods for Stochastic Differential Equations*, 191.

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. *The Annals of Statistics 41*(3), 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.

Kinnebrock, S. and M. Podolskij (2008). A note on the central limit theorem for bipower variation of general functions. *Stochastic Processes and Their Applications 118*(6), 1056–1070.

Kolokolov, A. and R. Renò (2017). Jumps or Staleness? *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.

Liu, Z., X. Kong, and J. B. (2018). Estimating the Integrated Volatility Using High Frequency Data with Zero Durations. *Journal of Econometrics* (204), 18–32.

Mandelbrot, B. B. and J. W. Van Ness (1968). Fractional brownian motions, fractional noises and applications. *SIAM review 10*(4), 422–437.

Phillips, P. C. B. and J. Yu (2007). Information loss in volatility measurement with flat price trading. Cowles Foundation Discussion Paper No. 1598.

Podolskij, M. and M. Vetter (2009). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli 15*(3), 634–668.

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

This appendix is divided into five 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 and 3.2. Section A.3 presents the proofs of Theorems 3.3 and 3.4 in the main text. Section A.4 presents the proof of Theorem 3.5 in the main text. Finally, the proofs of the results in Section 4 of the main text are reported in Section A.5.

A.1   Notations and auxiliary results

In what follows, we indicate with $t_{j,n} = j/n$, $j \in \{0, \dots, n\}$, the deterministic equispaced partition of the interval $[0,1]$ and with $N_n(s) \doteq \max\{j \mid t_{j,n} \le s\}$. Trivially, $N_n(1) = n$. When necessary, we assume the existence of finitely many out-of-sample values (i.e. the variables defined outside the interval $[0,1]$) in order to simplify the notations (this allows us, for example, considering quantities like $\sum_{i=1}^{n} \mathbb{B}_{i,n} \mathbb{B}_{i+1,n}$ instead of $\sum_{i=1}^{n-1} \mathbb{B}_{i,n} \mathbb{B}_{i+1,n}$, which have identical asymptotic properties). We use the symbol $\xrightarrow{p}$ for convergence in probability, and $\xrightarrow{u.c.p.}$ for 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 $\mathfrak{p}_t$, ii) $\mathcal{U}_{t_{j,n}}$ is the σ-algebra generated by random variables $\mathfrak{u}_{0,n}, \dots, \mathfrak{u}_{j,n}$, and iii) $\mathcal{F}_{t_{j,n}} = \mathcal{F}_{t_{j,n}}^{(p)} \vee \mathcal{U}_{t_{j,n}}$ 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 $\cup_{t \in [0,1]} \mathcal{F}_t^{(p)}$, $\mathcal{U}_\infty = \vee_{n=1}^{\infty} \mathcal{U}_{n,n}$, and $\mathcal{F}_{t_{n,n}} = \mathcal{F}_\infty^{(p)} \vee \mathcal{U}_{n,n}$.

For the sake of readability, we denote, for a generic index $j \in \{1, \dots, n\}$, 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, $j \in \{0, \dots, n\}$, is asymptotically negligible (AN) if:

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

that is,

$$\sup_{s \in [0,1]} \left| \sum_{j=1}^{N_n(s)} \xi_j^n \right| \xrightarrow{p} 0. \tag{25}$$

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

Remark 3. *Suppose that $\sum_{j=1}^{n} \left| \xi_j^n \right|$ converges to zero in $\mathbb{L}^1$, i.e.:*

$$\mathbb{E}\left[\sum_{j=1}^{n} \left| \xi_j^n \right| \right] \to 0. \tag{26}$$

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

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

to conclude that condition (26) is enough to guarantee that ξ_j^n is AN.

Remark 4. *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 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:

$$\sum_{j=1}^n \mathbb{E}_{j-1}\left[\xi_j^n\right] \xrightarrow{u.c.p.} 0, \tag{27}$$

$$\sum_{j=1}^n \mathbb{E}_{j-1}\left[\left(\xi_j^n\right)^2\right] \xrightarrow{p} 0. \tag{28}$$

As a consequence, if $\mathbb{E}_{j-1}\left[\xi_j^n\right] = 0$ then condition (28) 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 (see 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 $(\Omega, \mathcal{F}, \mathbb{P})$ is said to converge $\mathcal{G}$-stably with limit Y defined on an extension of the original probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ if and only if for any bounded continuous function g and any bounded $\mathcal{G}$-measurable random variable Z it holds that:*

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

In what follows, by stable convergence we mean $\mathcal{F}_\infty^{(p)}$-stable convergence (denoted simply $\xrightarrow{stably}$), unless otherwise stated. The classical stable CLT of Hall and Heyde (1980) is not valid for the triangular arrays considered in the present paper. Indeed, by construction, we have that $\mathcal{F}_{t_{j,m}} \not\subseteq \mathcal{F}_{t_{j,n}}$ 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 CLT 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 $X_{j,n}^{(\ell)}$ as:

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

and assume that:

$$\sum_{j=1}^{n}\left(X_{j,n}^{(\ell)}\right)^2 \overset{p}{\longrightarrow} \sigma^2, \tag{29}$$

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

$$\sum_{j=1}^{n} X_{j,n}^{(\ell)} \overset{stably}{\Longrightarrow} Z, \tag{30}$$

where Z is a random variable with $\mathcal{F}_{\infty}^{(p)}$-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 $\max_{1 \leq j \leq n}\left|X_{j,n}^{(\ell)}\right| \overset{p}{\longrightarrow} 0$. Moreover, by hypothesis $\sum_{j=1}^{n}\left(X_{j,n}^{(\ell)}\right)^2 \overset{p}{\longrightarrow} \sigma^2$ for an a.s. finite random variable σ. As a consequence (see 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 ($\iota = \sqrt{-1}$)

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

converges to 1 as $n \to \infty$ *weakly in* $\mathbb{L}^1$. By definition, this is equivalent to prove 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 $2 \leq m \leq n$, let $E_m \in \mathcal{F}_{t_{m,m}}$. We compute

$$
\begin{aligned}
\mathbb{E}\left[T_n(t)\,\mathbb{I}(E_m)\right] &= \mathbb{E}\left[\mathbb{E}\left[T_n(t)\,\mathbb{I}(E_m)\,|\mathcal{F}_{t_{m,m}}\right]\right] = \mathbb{E}\left[\mathbb{E}\left[\prod_{j=1}^{n}\left(1 + \iota\, t\, X_{j,n}^{(\ell)}\right)\mathbb{I}(E_m)\,\bigg|\mathcal{F}_{t_{m,m}}\right]\right] \\[2mm]
&= \mathbb{E}\left[\prod_{j\in I_1}\left(1 + \iota\, t\, X_{j,n}^{(\ell)}\right)\mathbb{I}(E_m)\,\mathbb{E}\left[\prod_{j\in I_2 \cup I_3}\left(1 + \iota\, t\, X_{j,n}^{(\ell)}\right)\,\bigg|\mathcal{F}_{t_{m,m}}\right]\right] \\[2mm]
&= \mathbb{E}\left[\prod_{j\in I_1}\left(1 + \iota\, t\, X_{j,n}^{(\ell)}\right)\mathbb{I}(E_m)\,\mathbb{E}\left[\mathbb{E}\left[\prod_{j\in I_2 \cup I_3}\left(1 + \iota\, t\, X_{j,n}^{(\ell)}\right)\,\bigg|\mathcal{F}_{\infty}^{(p)}\right]\,\bigg|\mathcal{F}_{t_{m,m}}\right]\right] \\[2mm]
&= \mathbb{E}\left[\prod_{j\in I_1}\left(1 + \iota\, t\, X_{j,n}^{(\ell)}\right)\mathbb{I}(E_m)\,\mathbb{E}\left[\prod_{j\in I_2}\left(1 + \iota\, t\, X_{j,n}^{(\ell)}\right)\,\bigg|\mathcal{F}_{t_{m,m}}\right]\mathbb{E}\left[\prod_{j\in I_3}\left(1 + \iota\, t\, X_{j,n}^{(\ell)}\right)\,\bigg|\mathcal{F}_{\infty}^{(p)}\right]\right], \tag{31}
\end{aligned}
$$

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

and that

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

because of the independence of the factors conditionally on $\mathcal{F}_\infty^{(p)}$ and the fact that for each $j\in\{1,\dots,n\}$ $X_{j,n}^{(\ell)}$ has expected value equal to one. Equation (31) then becomes

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

where the remainder term R_n consists of at most $2^{2|\mathrm{I}_1\cup \mathrm{I}_2|}-1$ terms of the form $\mathbb{E}\left[\mathbb{I}\left(E_m\right)\left(it\right)^r X_{j_1,n}^{(\ell)}\dots X_{j_r,n}^{(\ell)}\right]$, 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,

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

Finally, let $\triangle$ denote the symmetric difference. For any $E\in\mathcal{F}_\infty^{(p)}$ 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,

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

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

We state now the following corollary, which will be used in the subsequent sections.

Corollary 3. *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*

$$\sum_{j=1}^n \mathbf{X}_{j,n}^{(\ell)}\left(\mathbf{X}_{j,n}^{(\ell)}\right)'\xrightarrow{p}\Sigma,\tag{32}$$

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

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

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

Proof. The condition (32) 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 $\mathbf{c}=\left(c_1,\dots,c_q\right)'$. Consequently, by Theorem A.1, we have

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

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

Remark 5. *The statement of Theorem A.1 remains true if the condition (29) is replaced by the analogous condition for conditional variances:*

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

We conclude the present section with the following useful lemma.

Lemma 3. *For any sequence of bivariate random vectors $\mathcal{B}_n$, and any differentiable function $\nu(X,Y) : \mathbb{R}^2 \to \mathbb{R}$, the stable convergence*

$$\sqrt{n}\left(\mathcal{B}_n - \mathcal{B}\right) \overset{stably}{\Longrightarrow} \mathcal{N}\left(0, \Sigma\right),$$

implies that

$$\frac{\sqrt{n}\left(\nu\left(\mathcal{B}_n\right) - \nu\left(\mathcal{B}\right)\right)}{\sqrt{\nu_1\left(\mathcal{B}\right)^2 \sigma_{1,1} + \nu_2\left(\mathcal{B}\right)^2 \sigma_{2,2} + 2\nu_1\left(\mathcal{B}\right)\nu_2\left(\mathcal{B}\right)\sigma_{1,2}}} \overset{stably}{\Longrightarrow} \mathcal{N}\left(0, 1\right),$$

where $\mathcal{B}$ is a bivariate random vector, $\Sigma = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_{2,2} \end{bmatrix}$, and $\nu_1\left(\mathcal{B}\right)$ and $\nu_2\left(\mathcal{B}\right)$ denote, respectively, the derivatives of the function $\nu(X,Y)$ with respect to the first and second argument.

Proof. The proof of the result above can be found in Podolskij and Vetter (2010). □

A.2 Proofs of limit theorems of Section 3.1 and 3.2

The proofs of the limiting results from Sections 3.1 follow directly from several auxiliary 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 6. *Under Assumption 2,*

$$\mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right] = \mathfrak{p}_{j-1,n} + O_p\left(\Delta_n^{1/2}\right). \tag{33}$$

Indeed,

$$\mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbb{B}_{j,n} \,\Big|\, \mathcal{F}_{t_{j-1,n}} \vee \mathcal{F}_{t_{j,n}}^{(p)}\right]\right] = \mathbb{E}_{j-1}\left[\mathfrak{p}_{j,n}\right] = \mathfrak{p}_{j-1,n} + \mathbb{E}_{j-1}\left[\mathfrak{p}_{j,n} - \mathfrak{p}_{j-1,n}\right], \tag{34}$$

where

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

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

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

which implies that for every finite integer k

$$\mathfrak{p}_{j+k} = \mathfrak{p}_{j-1} + O_p\left(k\left(\Delta_n \left|\log \Delta_n\right|\right)^{1/2}\right). \tag{35}$$

Lemma 4. *Under Assumption 2, as $n \to \infty$,*

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

Proof. Consider the following quantity:

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

34

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:

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

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

$$A_n = \frac{1}{n} \sum_{j=m}^{n} \sum_{\ell=0}^{m-1} \varsigma_{j-\ell,\ell}^{(m)} + \frac{\mathcal{R}_n}{n}, \tag{36}$$

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

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

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

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

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 (36) 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:

$$\sum_{j=1}^{n} \mathbb{E}_{j-1} \left[\zeta_j^n \right] = \sum_{j=1}^{n} \frac{1}{n} \mathbb{E}_{j-1} \left[\varsigma_{j-\ell,\ell}^{(m)} \right] \xrightarrow{u.c.p.} 0 \tag{37}$$

35

and

$$\sum_{j=1}^{n} \mathbb{E}_{j-1}\left[\left(\varsigma_j^n\right)^2\right] \xrightarrow{p} 0. \tag{38}$$

In particular, to prove equation (37) 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

$$\sum_{j=1}^{n} \mathbb{E}_{j-1}\left[|\xi_j^n|\right] \xrightarrow{p} 0. \tag{39}$$

Therefore, we start from the assertion in (39) and we prove:

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

At this point, it is enough to prove the convergence in equation (38). This is an easy check because of the boundedness of the Bernoulli variates, that is:

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

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

$$\frac{1}{n}\sum_{j=1}^{n}\left(\mathfrak{p}_{j-1,n}\right)^m \longrightarrow \int_0^1 \left(\mathfrak{p}_s\right)^m ds,$$

which completes the proof. □

Before proceeding, we state and prove another useful lemma.

Lemma 5. *Under Assumption 2, for any finite numbers $\ell, d \geq 0$ and powers $q_1, \ldots, q_d \geq 0$, as $n \to \infty$,*

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

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

Proof. First, by Remark 33:

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

36

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

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

Therefore, by Theorem 2.13 in Hall and Heyde (1980)[12] applied to the martingale difference $X^{(\ell)}_{j,n} = \mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j,n}\mathfrak{p}^v_{j-1,n} - \mathfrak{p}_{j-\ell,n}\cdots\mathfrak{p}_{j,n}\mathfrak{p}^v_{j-1,n}$:

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

Using Remark (33) again:

$$\frac{1}{n}\sum_{j=1}^{n}\mathfrak{p}_{j-\ell,n}\cdots\mathfrak{p}_{j,n}\,\mathfrak{p}^v_{j-1,n} = \frac{1}{n}\sum_{j=1}^{n}\mathfrak{p}^{\ell+v}_{j-1,n} + O_p\left(\Delta^{1/2}\right).$$

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

$$\frac{1}{n}\sum_{j=1}^{n}\mathfrak{p}^{\ell+v}_{j-1,n} \longrightarrow \int_0^1 \mathfrak{p}^{\ell+v}_s\,ds,$$

which completes the proof. $\qquad\square$

Lemma 6. *Let $m \geq 2$ be a given integer number. Under Assumption 2, as $n \to \infty$*

$$\sqrt{n}\begin{bmatrix} \mathsf{RZ}_n - \int_0^1 \mathfrak{p}_s\,ds \\ \mathsf{RZ}_{m,n} - \int_0^1 (\mathfrak{p}_s)^m\,ds \end{bmatrix} \stackrel{stably}{\Longrightarrow} \mathcal{MN}\left(0,\Sigma_m\right), \tag{40}$$

where

$$\mathsf{RZ}_n = \frac{1}{n}\sum_{j=1}^{n}\mathbb{B}_{j,n} \qquad \mathsf{RZ}_{m,n} = \frac{1}{n}\sum_{j=1}^{n}\prod_{i=0}^{m-1}\mathbb{B}_{j+i,n},$$

and $\mathcal{MN}(0,\Sigma_m)$ denotes the mixed-normal distribution with covariance matrix Σ_m:

$$\Sigma_m = \begin{bmatrix} \int_0^1 \mathfrak{p}_s\left(1-\mathfrak{p}_s\right)ds & \int_0^1 m\,\mathfrak{p}^m_s\left(1-\mathfrak{p}_s\right)ds \\ \int_0^1 m\,\mathfrak{p}^m_s\left(1-\mathfrak{p}_s\right)ds & \int_0^1 \mathfrak{p}^m_s\,\frac{\mathfrak{p}^m_s\,(2\,m+1)-\mathfrak{p}^{m+1}_s\,(2\,m-1)-(1+\mathfrak{p}_s)}{1-\mathfrak{p}_s}\,ds \end{bmatrix}.$$

Proof. We consider the following decomposition:

$$\sqrt{n}\begin{bmatrix} \mathsf{RZ}_n - \int_0^1 \mathfrak{p}_s\,ds \\ \mathsf{RZ}_{m,n} - \int_0^1 (\mathfrak{p}_s)^m\,ds \end{bmatrix} = A_1 + A_2,$$

where

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

A_2 is AN. Therefore, it is enough to prove that $A_1 \stackrel{stably}{\Longrightarrow} \mathcal{MN}\left(0,\Sigma_m\right)$. 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:

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

---

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

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] \\
&\quad - \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, i.e. $A_1(2) = n^{-1/2}\sum_{j=1}^{n}\varphi_{j,m}$, can be rewritten as:

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

Reasoning as in Lemma 4, 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

$$
\begin{aligned}
\mathcal{R}_3 &= \sum_{j=m}^{n} \sum_{\ell=0}^{m-1} \left(\zeta_{j-\ell,\ell}^{(m)} - \widetilde{\zeta}_{j-\ell,\ell}^{(m)} \right) \\
&= \sum_{j=m}^{n} \sum_{\ell=0}^{m-1} \mathbb{B}_{j-\ell,n} \, \mathbb{B}_{j-\ell+1,n} \cdots \mathbb{B}_{j-1,n} \left(\mathbb{B}_{j,n} - \mathbb{E}_{j-1} \left[\mathbb{B}_{j,n} \right] \right) \times \\
&\quad \times \left(\mathbb{E}_j \left[\mathbb{B}_{j+1,n} \right] \cdots \mathbb{E}_{j-\ell+m-2} \left[\mathbb{B}_{j-\ell+m-1,n} \right] - \mathbb{E}_{j-1} \left[\mathbb{B}_{j+1,n} \right] \cdots \mathbb{E}_{j-1} \left[\mathbb{B}_{j+m-\ell-1,n} \right] \right) \\
&= \sum_{j=m}^{n} \sum_{\ell=0}^{m-1} \mathcal{R}_3(\ell) = \sum_{\ell=0}^{m-1} \sum_{j=m}^{n} \mathcal{R}_3(\ell),
\end{aligned}
\tag{41}
$$

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

$$
\begin{aligned}
\mathcal{R}_3(\ell) &= \sum_{j=m}^{n} \mathbb{B}_{j-\ell,n} \, \mathbb{B}_{j-\ell+1,n} \cdots \mathbb{B}_{j-1,n} \left(\mathbb{B}_{j,n} - \mathbb{E}_{j-1} \left[\mathbb{B}_{j,n} \right] \right) \times \\
&\quad \times \left(\mathbb{E}_j \left[\mathbb{B}_{j+1,n} \right] \cdots \mathbb{E}_{j-\ell+m-2} \left[\mathbb{B}_{j-\ell+m-1,n} \right] - \mathbb{E}_{j-1} \left[\mathbb{B}_{j+1,n} \right] \cdots \mathbb{E}_{j-1} \left[\mathbb{B}_{j+m-\ell-1,n} \right] \right) \\
&= \sum_{j=m}^{n} r_j(\ell).
\end{aligned}
\tag{42}
$$

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:

$$
\mathbb{E}_{j-1} \left[\frac{1}{\sqrt{n}} r_j(\ell) \right] = 0.
\tag{43}
$$

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 $M = m - \ell - 1$)

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

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

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

where the last inequality follows from Remark 6. Then, because m and ℓ are finite, we have:

$$
\sum_{j=m}^{n} \mathbb{E}_{j-1} \left[\left(\frac{1}{\sqrt{n}} r_j(\ell) \right)^2 \right] \leq K \sum_{j=m}^{n} \frac{1}{n} \mathbb{E}_{j-1} \left[\left(\sum_{i=1}^{m-\ell-1} |\mathbb{E}_{j-1} \left[\mathbb{B}_{j+i,n} \right] - \mathbb{E}_{j+i-1} \left[\mathbb{B}_{j+i,n} \right]| \right)^2 \right] \to 0.
\tag{44}
$$

Therefore, by Lemma 2, $\frac{1}{\sqrt{n}}\mathcal{R}_3(\ell)$ are AN $\forall \ell \in \{0, 1, \dots, m-1\}$, 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

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

and where the reminders are given by

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

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:

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

To establish the previous convergence, we use Corollary 3. We have to find two functions $\varphi^{(1)}$ and $\varphi^{(2)}$ such that:

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

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

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

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

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

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

$$= \sum_{\ell=0}^{m-1} \mathbb{B}_{j-\ell,n}\,\mathbb{B}_{j-\ell+1,n}\cdots\mathbb{B}_{j-1,n}\underbrace{\mathbb{E}_{j-1}\left[\left(\mathbb{B}_{j,n}-\mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\right)\right]}_{=0}\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\mathbb{E}_{j-1}\left[\mathbb{B}_{j+2,n}\right]\cdots\mathbb{E}_{j-1}\left[\mathbb{B}_{j-\ell+m-1,n}\right].$$

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

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

By Lemma 5,

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

Now consider the following product:

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

We note that

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

and

$$\widetilde{\zeta}_{j-\ell,\ell}^{(m)}\widetilde{\zeta}_{j-\ell-k,\ell+k}^{(m)}$$

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

$$\times\mathbb{B}_{j-\ell-k,n}\,\mathbb{B}_{j-\ell-k+1,n}\cdots\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j-1,n}\left(\mathbb{B}_{j,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\right)\mathbb{E}_{j-1}\left[\mathbb{B}_{j+1,n}\right]\cdots\mathbb{E}_{j-1}\left[\mathbb{B}_{j+m-\ell-k-1,n}\right]$$

$$= \underbrace{\mathbb{B}_{j-\ell-k,n}...\mathbb{B}_{j-1,n}}_{\ell+k \text{ factors}}\left(\mathbb{B}_{j,n} - \mathbb{E}_{j-1}\left[\mathbb{B}_{j,n}\right]\right)^2\times$$

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

Consequently, using Lemma 5,

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

which, after some standard algebra becomes:

$$
\begin{aligned}
\Sigma_m(2,2) &= \int_0^1 \mathfrak{p}_s^{2m-1}\,(1-\mathfrak{p}_s)\left(\sum_{\ell=0}^{m-1}\mathfrak{p}_s^{-\ell}+2\sum_{\ell=0}^{m-1}(m-\ell-1)\,\mathfrak{p}_s^{-\ell}\right)ds \\
&= \int_0^1 \frac{\mathfrak{p}_s^m\,(1+\mathfrak{p}_s-(2m(1-\mathfrak{p}_s)+1+\mathfrak{p}_s)\mathfrak{p}_s^m)}{1-\mathfrak{p}_s}\,ds \\
&= \int_0^1 \mathfrak{p}_s^m\,\frac{\mathfrak{p}_s^m\,(2\,m+1)-\mathfrak{p}_s^{m+1}\,(2\,m-1)-(1+\mathfrak{p}_s)}{1-\mathfrak{p}_s}\,ds.
\end{aligned}
\tag{45}
$$

Finally:

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

Applying Lemma 5 again:

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

which completes the proof. $\qquad\square$

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

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

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

$$
\widehat{\mathfrak{p}}_j\left(k_n\right)-\mathfrak{p}_{j-1,n}=\alpha_j^n+\beta_j^n,\quad j\in\{1,\dots,h_n+1\}.
$$

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

Lemma 7. *Under Assumptions 1, 2, and 3, for $C > 0$ and for all $q \geq 2$, we have:*

$$\mathbb{E}_{j-1}\left[\sup_{s\in[0,\Delta_n]}|\mathfrak{p}_{j-1+s,n} - \mathfrak{p}_{j-1,n}|^q\right] \leq C \cdot \Delta_n^{1\wedge(q/2)} \tag{46}$$

$$|\mathbb{E}_{j-1}\left[\mathfrak{p}_{j,n} - \mathfrak{p}_{j-1,n}\right]| \leq C \cdot \Delta_n \tag{47}$$

$$|\mathbb{E}_{j-1}\left[\beta_j^n\right]| \leq C \cdot k_n\Delta_n \tag{48}$$

$$\mathbb{E}_{j-1}\left[|\beta_j^n|^q\right] \leq C \cdot (k_n\Delta_n)^{q/2} \tag{49}$$

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

$$\mathbb{E}_{j-1}\left[|\alpha_j^n|^q\right] \leq Ck_n^{-q/2} \tag{51}$$

$$\left|\mathbb{E}_{j-1}\left[(\alpha_i^n)^2 - \frac{1}{k_n}\mathfrak{p}_{i-1,n}(1-\mathfrak{p}_{j-1,n})\right]\right| \leq C \cdot \Delta_n \tag{52}$$

$$|\mathbb{E}_{j-1}\left[\alpha_j^n\beta_j^n\right]| = 0 \tag{53}$$

Proof. The proof of (46)–(50) 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 (50)–(53). Equality (50) easily follows by conditioning on the path of the process $\mathfrak{p}_t$.

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

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

$$\mathbb{E}_{i-1}\left[(\alpha_j^n)^2\right] = \frac{1}{k_n^2}\mathbb{E}_{j-1}\left[\sum_{i=0}^{k_n-1}(\mathbb{B}_{j+i,n} - \mathfrak{p}_{j+i,n})^2\right] + \frac{2}{k_n}\mathbb{E}_{j-1}\left[\sum_{i=0}^{k_n-2}\sum_{m=1}^{k_n-1-i}(\mathbb{B}_{j+i,n} - \mathfrak{p}_{j+i,n})(\mathbb{B}_{j+i+m,n} - \mathfrak{p}_{j+i+m,n})\right]$$
$$= \frac{1}{k_n^2}\sum_{i=0}^{k_n-1}\mathbb{E}_{j-1}\left[(\mathbb{B}_{j+i,n} - \mathfrak{p}_{j+i,n})^2\right] = \frac{1}{k_n^2}\sum_{i=0}^{k_n-1}\mathbb{E}_{j-1}\left[\mathfrak{p}_{j+i,n}(1-\mathfrak{p}_{j+i,n})\right] \leq \frac{C}{k_n}, \tag{54}$$

where the last inequality is due to the fact that $\mathfrak{p}_t \in (0,1)$. Moreover, we have:

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

By applying triangular inequality, we obtain:

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

Therefore, (52) follows from (47), whereas (51) follows from Hölder's inequality and (54). Finally, (53) is obtained by conditioning on the path $(\mathfrak{p}_t)_{t\in[0,1]}$ and by using equation (50). $\square$

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

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

First, we prove that $\widehat{\mathfrak{p}}(k_n,t)$ converges in probability to $\mathfrak{p}_t$ 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:

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

Second, we have:

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

The first of the two terms converges to zero by the boundedness of $\mathbb{B}_{j_t+i,n}$ and $\mathfrak{p}_t$. Concerning the second, by conditioning on $(\mathfrak{p}_t)_{t \in [0,1]}$ and (47) we have that:

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

Therefore,

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

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

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

and we compute:

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

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

$$f(\widehat{\mathfrak{p}}(k_n, s)) \xrightarrow{p} f(\mathfrak{p}_s). \tag{55}$$

Nonetheless, because the sequence of random variables $f(\widehat{\mathfrak{p}}(k_n, s))$ is uniformly integrable (again using the locally boundedness of $f(\cdot)$ and the boundedness of $\widehat{\mathfrak{p}}(k_n, s)$), the convergence in equation (55) 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), $U(\Delta_n, f)^n \xrightarrow{u.c.p.} \int_0^1 f(\mathfrak{p}_s)\, ds$ by the dominated convergence theorem (see Jacod and Protter, 2012, Theorem 9.4.1). $\qquad\square$

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

$$\frac{1}{\sqrt{\Delta_n}} \left(U'\left(\Delta_n, f\right)^n - U\left(f\right) \right) = \sqrt{\Delta_n} \sum_{j=1}^{h_n+1} \left(f\left(\widehat{\mathfrak{p}}_j\left(k_n\right)\right) - \frac{1}{2\,k_n} f''\left(\widehat{\mathfrak{p}}_j\left(k_n\right)\right) \widehat{\mathfrak{p}}_j\left(k_n\right)\left(1 - \widehat{\mathfrak{p}}_j\left(k_n\right)\right) \right) - \frac{1}{\sqrt{\Delta_n}} \int_0^1 f\left(\mathfrak{p}_s\right)\,ds$$

$$= \sum_{r=1}^4 U\left(r\right)^n,$$

with

$$U\left(1\right)^n = \frac{1}{\sqrt{\Delta_n}} \sum_{j=1}^{h_n+1} \int_{(j-1)\Delta_n}^{j\Delta_n} \left(f\left(\mathfrak{p}_{j-1,n}\right) - f\left(\mathfrak{p}_s\right)\right)\,ds - \frac{1}{\sqrt{\Delta_n}} \int_{(h_n+1)\,\Delta_n}^1 f\left(\mathfrak{p}_s\right)\,ds$$

$$U\left(2\right)^n = \sqrt{\Delta_n} \sum_{j=1}^{h_n+1} f'\left(\mathfrak{p}_{j-1,n}\right) \beta_j^n$$

$$U\left(3\right)^n = \sqrt{\Delta_n} \sum_{j=1}^{h_n+1} \left(f\left(\widehat{\mathfrak{p}}_j\left(k_n\right)\right) - f\left(\mathfrak{p}_{j-1,n}\right) - f'\left(\mathfrak{p}_{j-1,n}\right)\left(\alpha_j^n + \beta_j^n\right) - \frac{1}{2\,k_n} f''\left(\widehat{\mathfrak{p}}_j\left(k_n\right)\right) \widehat{\mathfrak{p}}_j\left(k_n\right)\left(1 - \widehat{\mathfrak{p}}_j\left(k_n\right)\right) \right)$$

$$U\left(4\right)^n = \sqrt{\Delta_n} \sum_{j=1}^{h_n+1} f'\left(\mathfrak{p}_{j-1,n}\right) \alpha_j^n.$$

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 $U(4)^n \overset{stably}{\Longrightarrow} \mathcal{MN}(0, \Sigma_{\mathrm{U}})$.

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 $1 - (h_n + 1)\,\Delta_n = 1 - (n - k_n + 1)\,\Delta_n = k_n\,\Delta_n - \Delta_n$. Because $f(\mathfrak{p}_s)$ is bounded, for the second term of $U(1)^n$ we have:

$$\left| \frac{1}{\sqrt{\Delta_n}} \int_{(h_n+1)\,\Delta_n}^1 f\left(\mathfrak{p}_s\right)\,ds \right| \leq C k_n \sqrt{\Delta_n} \longrightarrow 0.$$

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

$$\xi_j^n = \frac{1}{\sqrt{\Delta_n}} \int_{(j-1)\Delta_n}^{j\Delta_n} \left(f\left(\mathfrak{p}_{j-1,n}\right) - f\left(\mathfrak{p}_s\right)\right)\,ds.$$

Because the process $f(\mathfrak{p}_t)$ is a bounded semimartingale, by using inequality (47) we get:

$$\left| \mathbb{E}\left[\xi_j^n\right] \right| = \frac{1}{\sqrt{\Delta_n}} \left| \mathbb{E}\left[\int_{(j-1)\Delta_n}^{j\Delta_n} \left(f\left(\mathfrak{p}_{j-1,n}\right) - f\left(\mathfrak{p}_s\right)\right)\,ds \right] \right| = \frac{1}{\sqrt{\Delta_n}} \left| \int_{(j-1)\Delta_n}^{j\Delta_n} \mathbb{E}\left[\mathbb{E}_{j-1}\left[\left(f\left(\mathfrak{p}_{j-1,n}\right) - f\left(\mathfrak{p}_s\right)\right)\right]\,ds \right| $$

$$\leq \frac{1}{\sqrt{\Delta_n}} \int_{(j-1)\Delta_n}^{j\Delta_n} \left| \mathbb{E}\left[\mathbb{E}_{j-1}\left[\left(f\left(\mathfrak{p}_{j-1,n}\right) - f\left(\mathfrak{p}_s\right)\right)\right]\right] \right|\,ds \leq \frac{1}{\sqrt{\Delta_n}} \int_{(j-1)\Delta_n}^{j\Delta_n} \mathbb{E}\left[\left|\mathbb{E}_{j-1}\left[\left(f\left(\mathfrak{p}_{j-1,n}\right) - f\left(\mathfrak{p}_s\right)\right)\right]\right|\right]\,ds$$

$$\leq \frac{C}{\sqrt{\Delta_n}} \Delta_n^2 = C\left(\Delta_n\right)^{3/2} \longrightarrow 0,$$

while using inequality (46) and Holder's inequality, we obtain

$$
\begin{aligned}
\mathbb{E}\left[\left|\xi_j^n\right|^2\right] &= \frac{1}{\Delta_n}\mathbb{E}\left[\left(\int_{(j-1)\Delta_n}^{j\Delta_n}\left(f\left(\mathfrak{p}_{j-1,n}\right)-f\left(\mathfrak{p}_s\right)\right)ds\right)^2\right] \\
&= \frac{1}{\Delta_n}\mathbb{E}\left[\int_{(j-1)\Delta_n}^{j\Delta_n}\int_{(j-1)\Delta_n}^{j\Delta_n}\left(f(\mathfrak{p}_{j-1,n})-f(\mathfrak{p}_q)\right)\left(f(\mathfrak{p}_{j-1,n})-f(\mathfrak{p}_s)\right)ds\,dq\right] \\
&= \frac{1}{\Delta_n}\int_{(j-1)\Delta_n}^{j\Delta_n}\int_{(j-1)\Delta_n}^{j\Delta_n}\mathbb{E}\left[\left(f(\mathfrak{p}_{j-1,n})-f(\mathfrak{p}_q)\right)\left(f(\mathfrak{p}_{j-1,n})-f(\mathfrak{p}_s)\right)\right]ds\,dq \\
&\leq \frac{1}{\Delta_n}\int_{(j-1)\Delta_n}^{j\Delta_n}\int_{(j-1)\Delta_n}^{j\Delta_n}\sqrt{\mathbb{E}\left[\left|f(\mathfrak{p}_{j-1,n})-f(\mathfrak{p}_q)\right|^2\right]\mathbb{E}\left[\left|f(\mathfrak{p}_{j-1,n})-f(\mathfrak{p}_s)\right|^2\right]}\,ds\,dq \\
&\leq \frac{1}{\Delta_n}\int_{(j-1)\Delta_n}^{j\Delta_n}\int_{(j-1)\Delta_n}^{j\Delta_n}C\,\Delta_n\,ds\,dq \leq C\,\Delta_n^2 \longrightarrow 0.
\end{aligned}
$$

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

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

Using Lemma 7 and the boundedness of $f'\left(\mathfrak{p}_{j-1,n}\right)$, we obtain

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

and

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

and so

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

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

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

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

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

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

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

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

$$\frac{1}{2} f''(\mathfrak{p}_{j-1,n})(\alpha_j^n + \beta_j^n)^2 = \frac{1}{2} f''(\mathfrak{p}_{j-1,n})\left((\alpha_j^n)^2 + 2\alpha_j^n\beta_j^n - \frac{1}{k_n}\mathfrak{p}_{j-1,n}(1 - \mathfrak{p}_{j-1,n})\right)$$
$$+ \frac{1}{2k_n} f''(\mathfrak{p}_{j-1,n})\,\mathfrak{p}_{j-1,n}(1 - \mathfrak{p}_{j-1,n}) + \frac{1}{2} f''(\mathfrak{p}_{j-1,n})(\beta_j^n)^2.$$

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:

$$v_j^n(1) = \frac{\sqrt{\Delta_n}}{2} f''(\mathfrak{p}_{j-1,n})\left((\alpha_j^n)^2 + 2\alpha_j^n\beta_j^n - \frac{1}{k_n}\mathfrak{p}_{j-1,n}(1 - \mathfrak{p}_{j-1,n})\right),$$

$$v_j^n(2) = \frac{\sqrt{\Delta_n}}{2k_n} f''(\mathfrak{p}_{j-1,n})\,\mathfrak{p}_{j-1,n}(1 - \mathfrak{p}_{j-1,n}) - \frac{\sqrt{\Delta_n}}{2k_n} f''(\widehat{\mathfrak{p}}_j(k_n))\,\widehat{\mathfrak{p}}_j(k_n)\left(1 - \widehat{\mathfrak{p}}_j(k_n)\right),$$

$$v_j^n(3) = \frac{\sqrt{\Delta_n}}{2} f''(\mathfrak{p}_{j-1,n})(\beta_j^n)^2,$$

$$v_j^n(4) = \frac{\sqrt{\Delta_n}}{6} f'''(p_j^\star)(\alpha_j^n + \beta_j^n)^3.$$

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 (52) and (53) imply that $\left|\mathbb{E}_{j-1}\left[v_j^n(1)\right]\right| \le C\,\Delta_n^{3/2}$, and so:

$$\sum_{j=1}^{h_n+1} \left|\mathbb{E}_{j-1}\left[v_j^n(1)\right]\right| \le C\,\Delta_n^{1/2} \xrightarrow{p} 0. \tag{56}$$

In addition,

$$v_j^n(1)^2 = \frac{\Delta_n}{4} f''(\mathfrak{p}_{j-1,n})^2 \left((\alpha_j^n)^4 + 4(\alpha_j^n\beta_j^n)^2 + \frac{1}{k_n^2}\mathfrak{p}_{j-1,n}^2(1 - \mathfrak{p}_{j-1,n})^2 + \right.$$
$$\left. + 4(\alpha_j^n)^3\beta_j^n - 2\frac{(\alpha_j^n)^2}{k_n}\mathfrak{p}_{j-1,n}(1 - \mathfrak{p}_{j-1,n}) - \frac{4\alpha_j^n\beta_j^n}{k_n}\mathfrak{p}_{j-1,n}(1 - \mathfrak{p}_{j-1,n})\right)$$
$$\le \frac{\Delta_n}{4} f''(\mathfrak{p}_{j-1,n})^2 \left((\alpha_j^n)^4 + 4(\alpha_j^n\beta_j^n)^2 + \frac{1}{k_n^2}\mathfrak{p}_{j-1,n}^2(1 - \mathfrak{p}_{j-1,n})^2 + \right.$$
$$\left. + 4\left|(\alpha_j^n)^3\beta_j^n\right| + 2\frac{(\alpha_j^n)^2}{k_n}\mathfrak{p}_{j-1,n}(1 - \mathfrak{p}_{j-1,n}) - \frac{4\alpha_j^n\beta_j^n}{k_n}\mathfrak{p}_{j-1,n}(1 - \mathfrak{p}_{j-1,n})\right).$$

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

- Inequality (51) implies that
$$\mathbb{E}_{j-1}\left[(\alpha_j^n)^4\right] \le C\,k_n^{-2},$$
and that
$$\mathbb{E}_{j-1}\left[\frac{(\alpha_j^n)^2}{k_n}\mathfrak{p}_{j-1,n}(1 - \mathfrak{p}_{j-1,n})\right] \le C\,k_n^{-2}.$$

- Cauchy-Schwartz inequality plus (51) and (49) imply that
$$\mathbb{E}_{j-1}\left[(\alpha_j^n\beta_j^n)^2\right] \le \left(\mathbb{E}_{j-1}\left[(\alpha_j^n)^4\right]\right)^{1/2}\left(\mathbb{E}_{j-1}\left[(\beta_j^n)^4\right]\right)^{1/2} \le C\,\Delta_n$$
and that
$$\left|\mathbb{E}_{j-1}\left[(\alpha_j^n)^3\beta_j^n\right]\right| \le \left(\mathbb{E}_{j-1}\left[(\alpha_j^n)^6\right]\right)^{1/2}\left(\mathbb{E}_{j-1}\left[(\beta_j^n)^2\right]\right)^{1/2} \le C\,k_n^{-1}\Delta_n^{1/2}.$$

- Equation (53) implies that $\mathbb{E}_{j-1}\left[\alpha_j^n\beta_j^n k_n\mathfrak{p}_{j-1,n}(1 - \mathfrak{p}_{j-1,n})\right] = 0$.

Summing up:

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

where

$$k_n\sum_{j=1}^{h_n}\mathbb{E}\left[v_j^n(1)^2\right] \longrightarrow 0. \tag{57}$$

Therefore, the limits in (56) and (57) imply, through Lemma 2, that $v_j^n(1)$ is AN. Now, consider $v_j^n(4)$. Because both $\mathfrak{p}_t$ and $\widehat{\mathfrak{p}}_i(k_n)$ are in $[0,1]$, $|f'''(\mathfrak{p}_i^\star)| \le C$ for some constant $C > 0$ we therefore have:

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

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

- inequality (51) implies

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

- Inequalities (51) and (49), and Cauchy-Schwartz, imply

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

and

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

- Inequality (49) implies

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

Therefore:

$$\sum_{j=1}^{h_n+1}\left|\mathbb{E}_{j-1}\left[v_j^n(4)\right]\right|\xrightarrow{p} 0. \tag{58}$$

Now consider

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

inequalities (51) and (49), respectively, imply:

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

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

and, using also Cauchy-Schwartz, they imply:

$$k_n \sum_{j=1}^{h_n+1} \Delta_n \mathbb{E}\left[\left|\alpha_j^n\right|^2 \left|\beta_j^n\right|^4\right] \le C \left(k_n \Delta_n\right)^2 \longrightarrow 0,$$

$$k_n \sum_{j=1}^{h_n+1} \Delta_n \mathbb{E}\left[\left|\alpha_j^n\right|^4 \left|\beta_j^n\right|^2\right] \le C \, k_n^{-2} \Delta_n \longrightarrow 0,$$

$$k_n \sum_{j=1}^{h_n+1} \Delta_n \mathbb{E}\left[\left|\alpha_j^n\right|^5 \left|\beta_j^n\right|\right] \le C \, k_n^{-1} \Delta_n^{1/2} \longrightarrow 0,$$

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

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

Consequently:

$$k_n \sum_{j=1}^{h_n+1} \mathbb{E}\left[v_j^n(4)^2\right] \longrightarrow 0. \tag{59}$$

As before, the limits in (58) and (59) imply, through Lemma 2, that $v_j^n(4)$ is AN. Similarly, for $v_j^n(3)$ we have:

$$\sum_{j=1}^{h_n+1} \mathbb{E}_{j-1}\left[\left|\frac{\sqrt{\Delta_n}}{2} f''\left(\mathfrak{p}_{j-1,n}\right)\left(\beta_j^n\right)^2\right|\right] \le C \cdot k_n \sqrt{\Delta_n} \xrightarrow{p} 0. \tag{60}$$

In addition:

$$k_n \sum_{j=1}^{h_n+1} \mathbb{E}\left[\left|\frac{\Delta_n}{4}\left(f''\left(\mathfrak{p}_{j-1,n}\right)\right)^2 \left(\beta_j^n\right)^4\right|\right] \le C \cdot \left(k_n^{3/2} \Delta_n\right)^2 \longrightarrow 0. \tag{61}$$

Therefore, the limits in (60) and (61) 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 $\widehat{\mathfrak{p}}_j(k_n) - \mathfrak{p}_{j-1,n} = \alpha_j^n + \beta_j^n$):

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

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

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

Using Lemma 7, we have:

$$
\sum_{j=1}^{h_n}\left|\mathbb{E}_{j-1}\left[\frac{\sqrt{\Delta_n}}{k_n} f''\left(\mathfrak{p}_{(j-1)\Delta_n}\right)\alpha_j^n\right]\right| = 0,
$$

$$
k_n \sum_{j=1}^{h_n}\mathbb{E}\left[\frac{\Delta_n}{k_n^2}\left(f''\left(\mathfrak{p}_{(j-1)\Delta_n}\right)\right)^2\left|\alpha_j^n\right|^2\right] \leq C\,k_n^{-2},
$$

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

$$
k_n \sum_{j=1}^{h_n}\mathbb{E}\left[\frac{\Delta_n}{k_n^2}\left(f''\left(\mathfrak{p}_{(j-1)\Delta_n}\right)\right)^2\left|\beta_j^n\right|^2\right] \leq C\,\Delta_n,
$$

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

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

and being $\left(\widehat{\mathfrak{p}}_j(k_n)+\mathfrak{p}_{j-1,n}\right)$ 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 $U^n(4)\overset{stably}{\Longrightarrow}\mathcal{MN}\left(0,\Sigma_{\mathrm{U}}\right)$
Recall that $U(4)^n$ is defined as:

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

For the sake of readability, we temporarily define the variables

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

50

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:

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

Hence,

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

where

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

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

$$\left|w_j^n\right| \leq C \sup_{s\in[(j-1)\Delta_n,(j+k_n-1)\Delta_n]} \left|\mathfrak{p}_s - \mathfrak{p}_{j-1,n}\right|.$$

Hence, inequality (46) implies that $\mathbb{E}\left[\left|w_j^n\right|^2\right] \leq C\sqrt{\Delta_n}$ when $k_n \leq j \leq \lfloor 1/\Delta_n\rfloor - k_n$ and $\left|w_j^n\right| \leq C$ always. Therefore, since both $\mathbb{B}_{j,n}$ and $\mathfrak{p}_t$ are bounded,

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

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

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

Because $\left(f' \left(\mathfrak{p}_{j-1,n} \right) \right)^2$ is bounded, using (47) we have:

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

Therefore,

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

Consequently,

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

which completes the proof. $\qquad \square$

A.4 Proof of Theorem 3.5 from Section 3.3

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

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

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

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

We have:

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

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

$$\widehat{\mathfrak{p}}_{j+k_n} \left(k_n \right) - \widehat{\mathfrak{p}}_j \left(k_n \right) = \frac{1}{k_n} \sum_{i=0}^{2k_n-1} \epsilon(1)_i \left(\mathbb{B}_{j+i,n} - \mathfrak{p}_{j+i,n} \right) + \frac{1}{k_n} \sum_{i=0}^{k_n-1} \left(\mathfrak{p}_{j+i+k_n,n} - \mathfrak{p}_{j+i,n} \right), \tag{62}$$

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:

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

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

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

which can be rewritten as:

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

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 (62) becomes:

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

Therefore:

$$
\begin{aligned}
\left(\widehat{\mathfrak{p}}_{j+k_n}(k_n) - \widehat{\mathfrak{p}}_j(k_n) \right)^2 &= \frac{1}{k_n^2} \sum_{i=0}^{2k_n-1} \Bigg(\epsilon(2)_i^2 \, (\mathbb{B}_{j+i,n} - \mathfrak{p}_{j+i,n})^2 + \epsilon(2)_i^2 \, (\mathfrak{p}_{j+i+1,n} - \mathfrak{p}_{j+i,n})^2 \\
&\quad + 2\,\epsilon(2)_i \, \epsilon(2)_i \, (\mathbb{B}_{j+i,n} - \mathfrak{p}_{j+i,n}) \, (\mathfrak{p}_{j+i+1,n} - \mathfrak{p}_{j+i,n}) \Bigg) \\
&\quad + 2 \sum_{j=0}^{2\,k_n-2} \sum_{\ell'=j+1}^{2\,k_n-1} \Bigg(\epsilon(2)_i \epsilon(1)_\ell \, (\mathbb{B}_{j+i,n} - \mathfrak{p}_{j+i,n}) \, (\mathbb{B}_{j+\ell,n} - \mathfrak{p}_{j+\ell,n}) \\
&\quad + \epsilon(2)_i \, \epsilon(2)_\ell \, (\mathbb{B}_{j+i,n} - \mathfrak{p}_{j+i,n}) \, (\mathfrak{p}_{j+\ell+1,n} - \mathfrak{p}_{j+\ell,n}) \\
&\quad + \epsilon(1)_\ell \, \epsilon(2)_i \, (\mathbb{B}_{j+\ell,n} - \mathfrak{p}_{j+\ell,n}) \, (\mathfrak{p}_{j+i+1,n} - \mathfrak{p}_{j+i,n}) \\
&\quad + \epsilon(2)_i \, \epsilon(2)_\ell \, (\mathfrak{p}_{j+i+1,n} - \mathfrak{p}_{j+i,n}) \, (\mathfrak{p}_{j+\ell+1,n} - \mathfrak{p}_{j+\ell,n}) \Bigg).
\end{aligned} \tag{63}
$$

So, setting

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

we have the following more compact expression:

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

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

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

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

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

$$v_i^n(5) = \frac{2}{k_n^2} \sum_{i=0}^{2k_n-2} \sum_{\ell=j+1}^{2k_n-1} \epsilon(2)_i \epsilon(2)_\ell \left(\mathfrak{p}_{j+i+1,n} - \mathfrak{p}_{j+i,n}\right) \left(\mathfrak{p}_{j+\ell+1,n} - \mathfrak{p}_{j+\ell,n}\right),$$

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

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

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 $s = 1, \ldots 7$. 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 six 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

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

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

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

where we use conditioning on $(\mathfrak{p}_t)_{t \in [0,1]}$, triangular inequality, and Lemma 7. Next, using the boundedness of $\mathfrak{p}_t$, we

obtain:

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

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

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

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

$$\frac{1}{k_n}\sum_{j=1}^{g_n} v_j^n(1) \overset{u.c.p.}{\longrightarrow} \frac{2}{\theta^2}\int_0^1 \mathfrak{p}_s(1-\mathfrak{p}_s)\,ds. \tag{64}$$

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

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

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

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

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,

$$\mathfrak{p}_t = \mathfrak{p}_0 + \int_0^t \mu_s\,ds + \int_0^t \nu_s\,dW_s$$

now implies that

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

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

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

and, by Markov inequality, the Itô isometry and the boundedness of[13] ν:

$$\int_0^\Delta \nu_s \, dW_s = \nu_0 \left(W_\Delta - W_0 \right) + O_p(\Delta^{1/2}), \tag{65}$$

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

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

The square of the increment $\Delta_j^n \mathfrak{p} = \left(\mathfrak{p}_{j+1,n} - \mathfrak{p}_{j,n} \right)$ then becomes

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

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

$$\left(\Delta_j^n \mathfrak{p}\right)^2 = \nu_{j\,\Delta_n}^2 \left(\Delta_j^n W\right)^2 + O_p\left(\Delta_n^{1/2}\right) \left(\Delta_j^n W\right)^2 + \nu_{j,n} \left(\Delta_j^n W\right) O_p\left(\Delta_n\right), \tag{66}$$

so that:

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

Now consider the same increment shifted by i units:

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

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

$$\left(\Delta_{j+i}^n \mathfrak{p}\right)^2 = \nu_{j,n}^2 \left(\Delta_{j+i}^n W\right)^2 + \left(\Delta_{j+i}^n W\right)^2 O_p\left(\sqrt{i\,\Delta_n}\right), \tag{67}$$

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 $\left(\Delta_{j+i}^n \mathfrak{p}\right)^2$ and $\left(\Delta_j^n \mathfrak{p}\right)^2$ has the following order in probability:

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

implying that

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

---

[13] 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 (65) 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:

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

From (66) we get

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

and therefore:

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

Similarly, from (67) we get

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

and hence:

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

Summing up the two fourth powers so obtained:

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

Finally consider that

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

where

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

Therefore

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

which implies:

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

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

$$
\begin{aligned}
\mathbb{E}\left[|\mathcal{D}_{j,n}|\right] &= \frac{2}{k_n^5}\sum_{i=0}^{2k_n-2}\sum_{\ell=i+1}^{2k_n-1}\varepsilon_{i,\ell}^2\,\mathbb{E}\left[\left|\left(\Delta_{j+i}^n\mathfrak{p}\right)^2 - \left(\Delta_j^n\mathfrak{p}\right)^2\right|\,\left|\left(\Delta_{j+\ell}^n\mathfrak{p}\right)^2 - \left(\Delta_j^n\mathfrak{p}\right)^2\right|\right] \\[2mm]
&\leq \frac{2}{k_n^5}\sum_{i=0}^{2k_n-2}\sum_{\ell=i+1}^{2k_n-1}\varepsilon_{i,\ell}^2\left(\mathbb{E}\left[\left(\left(\Delta_{j+i}^n\mathfrak{p}\right)^2 - \left(\Delta_j^n\mathfrak{p}\right)^2\right)^2\right]\right)^{\frac12}\left(\mathbb{E}\left[\left(\left(\Delta_{j+\ell}^n\mathfrak{p}\right)^2 - \left(\Delta_j^n\mathfrak{p}\right)^2\right)^2\right]\right)^{\frac12} \\[2mm]
&= \frac{2}{k_n^5}\sum_{i=0}^{2k_n-2}\sum_{\ell=i+1}^{2k_n-1}\varepsilon_{i,\ell}^2\left(\mathbb{E}\left[6\,\nu_{j,n}^4\,\Delta_n^2 + O_p\left(j^{\frac12}\,\Delta_n^{\frac52}\right)\right]\right)^{\frac12}\left(\mathbb{E}\left[6\,\nu_{j,n}^4\,\Delta_n^2 + O_p\left(\ell^{\frac12}\,\Delta_n^{\frac52}\right)\right]\right)^{\frac12}.
\end{aligned}
$$

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

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

so that

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

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:

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

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:

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

First, we observe that, conditionally on $(\mathfrak{p}_t)$, 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:

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

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

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

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

Therefore:

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

Therefore:

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

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

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

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

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

where

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

and

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

By conditioning on $(\mathfrak{p}_t)$ again, we have $\mathbb{E}\left[\zeta(1)_{j+i}\zeta(1)_{j+\ell}\zeta(1)_{j+u}\zeta(1)_{j+m}\right] = 0$ 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 $|\zeta(1)_{j+\ell}| \leq C$, for some constant $C > 0$,

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

Therefore:

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

Consequently, by Lemma 2, $\frac{1}{k_n}v_j^n(4)$ is AN.

Part 5: Proof of the convergence in probability of $v_i^n(5)$

By successive conditioning and using Lemma 7, we obtain

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

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

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

Next, we have

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

where:

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

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

Furthermore, we have

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

where:

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

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

Using the estimate (46) of Lemma 7, and the fact that $\sum_{i=0}^{2k_n-2}\sum_{\ell=i+1}^{2k_n-1} \left(\epsilon(2)_i \epsilon(2)_\ell\right)^2 \sim k_n^6$, we obtain

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

which implies that $k_n \sum_{i=1}^{g_n} \mathbb{E}\left[\mathcal{A}_{1,1,n}\right] \le k_n \Delta_n \to 0$. Next, using the estimates (46) and (47) of Lemma 7, we have

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

Therefore, we have

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

where:

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

$$S_2 = \sum_{i=0}^{2k_n-3} \sum_{\ell=i+1}^{2k_n-2} \sum_{m=i+2}^{2k_n-1} \left(\epsilon(2)_i\right)^2 \epsilon(2)_\ell \epsilon(2)_m \mathbb{I}\left(\{\ell \ne m\}\right) = \sum_{i=0}^{2k_n-3} \sum_{\ell=i+1}^{2k_n-2} \sum_{m=i+2}^{2k_n-1} \left(\epsilon(2)_i\right)^2 \epsilon(2)_\ell \epsilon(2)_m - S_1 \sim k_n^7.$$

Consequently,

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

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

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

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 $(\mathfrak{p}_t)$ 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

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

and

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

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

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

Thus:

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

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.

A.5 Proof of the results in Section 4

Proof of Corollary 1. By Theorem 3.2, as $n \to \infty$,

$$\sqrt{n}\begin{bmatrix} \mathsf{RZ}_n - \int_0^1 \mathfrak{p}_s \, ds \\ \mathsf{RZ}_{m,n} - \int_0^1 \left(\mathfrak{p}_s\right)^m \, ds \end{bmatrix} \overset{stably}{\Longrightarrow} \mathcal{MN}\left(0, \Sigma_m\right),$$

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

$$\Sigma_m = \begin{bmatrix} \int_0^1 \mathfrak{p}_s \left(1 - \mathfrak{p}_s\right) ds & \int_0^1 m \, \mathfrak{p}_s^m \left(1 - \mathfrak{p}_s\right) ds \\ \int_0^1 m \, \mathfrak{p}_s^m \left(1 - \mathfrak{p}_s\right) ds & \int_0^1 \mathfrak{p}_s^m \frac{\mathfrak{p}_s^m \, (2\,m+1) - \mathfrak{p}_s^{m+1} \, (2\,m-1) - (1+\mathfrak{p}_s)}{1-\mathfrak{p}_s} \, ds \end{bmatrix}.$$

Therefore, applying Lemma 3 with $\nu(X, Y) = Y - (X)^m$, we obtain:

$$\frac{\sqrt{n}\left(\mathsf{RZ}_{m,n} - (\mathsf{RZ}_n)^m - \left(\int_0^1 (\mathfrak{p}_s)^m \, ds - \left(\int_0^1 \mathfrak{p}_s \, ds\right)^m\right)\right)}{\sqrt{m^2 (\mathsf{RZ}_n)^{2(m-1)} \int_0^1 \mathfrak{p}_s (1-\mathfrak{p}_s) \, ds + \int_0^1 \mathfrak{p}_s^m \frac{\mathfrak{p}_s^m (2m+1) - \mathfrak{p}_s^{m+1} (2m-1) - (1+\mathfrak{p}_s)}{1-\mathfrak{p}_s} \, ds + 2m (\mathsf{RZ}_n)^{m-1} \int_0^1 m \mathfrak{p}_s^m (1-\mathfrak{p}_s) \, ds}} \overset{stably}{\Longrightarrow} \mathcal{N}(0, 1).$$

On Ω_0, $\left(\int_0^1 (\mathfrak{p}_s)^m \, ds - \left(\int_0^1 \mathfrak{p}_s \, ds\right)^m\right) = 0$ and the integrals $\int_0^1 (\mathfrak{p}_s)^m \, ds$ in the denominator of the test statistic can be consistently estimated by $(\mathsf{RZ}_n)^m$. Consequently, plugging in the estimates of the integrals $\int_0^1 (\mathfrak{p}_s)^m \, ds$ and simplifying, we obtain:

$$\Psi_{m,n} = \frac{\sqrt{n}\left(\mathsf{RZ}_{m,n} - (\mathsf{RZ}_n)^m\right)}{\sqrt{\frac{(\mathsf{RZ}_n)^{2m+1}(m^2+2m-1) - (\mathsf{RZ}_n)^{2m}(2m^2+2m+1) + (\mathsf{RZ}_n)^{m+1} + (\mathsf{RZ}_n)^m}{\mathsf{RZ}_n - 1}}} \overset{stably}{\Longrightarrow} \mathcal{N}(0, 1), \quad \text{on } \Omega_0. \tag{68}$$

On Ω_1,

$$\mathsf{RZ}_{m,n} - (\mathsf{RZ}_n)^m \overset{p}{\longrightarrow} \left(\int_0^1 (\mathfrak{p}_s)^m \, ds - \left(\int_0^1 \mathfrak{p}_s \, ds\right)^m\right) > 0, \tag{69}$$

which implies that :

$$\sqrt{n}\left(\mathsf{RZ}_{m,n} - (\mathsf{RZ}_n)^m\right) \overset{p}{\longrightarrow} +\infty. \tag{70}$$

Due to the boundedness of the Bernoulli random variables, the denominator of $\Psi_{m,n}$ is bounded as well. Hence, $\Psi_{m,n} \overset{p}{\longrightarrow} +\infty$, which completes the proof.

$\square$

Proof of Theorem 4.1 We start from the consistency result and we state the following lemma.

Lemma 8. *Under Assumption 3, as $n \to \infty$,*

$$\frac{1}{n} \sum_{j=1}^n \prod_{i=0}^{m-1} \frac{\mathbb{B}_{i+j,n}^{(t)}}{\vartheta_{i+j,n}^\star} \overset{u.c.p.}{\longrightarrow} \frac{1}{(\pi^\star)^m} \int_{t-1}^t (\mathfrak{p}_s^\star)^m \, ds.$$

Proof. In order to simplify notation, we present the proof for the case $t = 1$. The proof follows the lines of that of Lemma 4 and therefore we omit some details. Consider the following quantity:

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

We show that $A_n \overset{u.c.p.}{\longrightarrow} 0$. In particular, the latter can be rewritten as

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

where

$$\varsigma_{j,\ell}^{(m)} = \frac{\mathbb{B}_{j,n}\mathbb{B}_{j+1,n} \cdots \mathbb{B}_{j+\ell-1,n}}{\vartheta_{j,n}^\star \vartheta_{j+1,n}^\star \cdots \vartheta_{j+\ell-1,n}^\star} \left(\frac{\mathbb{B}_{j+\ell,n}}{\vartheta_{j+\ell,n}^\star} - \frac{\mathfrak{p}_{j-1,n}^\star}{\pi^\star}\right) \left(\frac{\mathfrak{p}_{j-1,n}^\star}{\pi^\star}\right)^{m-\ell-1}.$$

For fixed ℓ and m let $\zeta_j^n = \frac{1}{n} \varsigma_{j-\ell,\ell}^{(m)}$. Thus, to prove that A_n is AN, it is sufficient to prove that $\sum_{j=1}^n \zeta_j^n$ is AN, i.e., to show that

$$\sum_{j=1}^n \frac{1}{n} \left|\mathbb{E}_{j-1}\left[\varsigma_{j-\ell,\ell}^{(m)}\right]\right| \overset{p}{\longrightarrow} 0$$

and

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

Notice that, by conditioning on $\mathfrak{p}_{j,n}$ we have:

$$\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j,n}}{\vartheta_{j,n}^\star}\right] = \mathbb{E}_{j-1}\left[\frac{\mathfrak{p}_{j,n}}{\vartheta_{j,n}^\star}\right] = \mathbb{E}_{j-1}\left[\frac{\mathfrak{p}_{j,n}^\star}{\pi^\star}\right].$$

Therefore:

$$\begin{aligned}
\sum_{j=1}^{n}\frac{1}{n}\left|\mathbb{E}_{j-1}\left[\varsigma_{j-\ell,\ell}^{(m)}\right]\right| &= \sum_{j=1}^{n}\frac{1}{n}\left|\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j-1,n}}{\vartheta_{j-\ell,n}^\star\cdots\vartheta_{j-\ell,n}^\star}\left(\frac{\mathfrak{p}_{j-\ell-1,n}^\star}{\pi^\star}\right)^{m-\ell-1}\left(\frac{\mathbb{B}_{j,n}}{\vartheta_{j,n}^\star}-\frac{\mathfrak{p}_{j-\ell-1,n}^\star}{\pi^\star}\right)\right]\right| \\
&= \sum_{j=1}^{n}\frac{1}{n}\left|\frac{\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j-1,n}}{\vartheta_{j-\ell,n}^\star\cdots\vartheta_{j-\ell,n}^\star}\left(\frac{\mathfrak{p}_{j-\ell-1,n}^\star}{\pi^\star}\right)^{m-\ell-1}\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j,n}}{\vartheta_{j,n}^\star}-\frac{\mathfrak{p}_{j-\ell-1,n}^\star}{\pi^\star}\right]\right| \\
&= \sum_{j=1}^{n}\frac{1}{n}\left|\frac{\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j-1,n}}{\vartheta_{j-\ell,n}^\star\cdots\vartheta_{j-\ell,n}^\star}\left(\frac{\mathfrak{p}_{j-\ell-1,n}^\star}{\pi^\star}\right)^{m-\ell-1}\mathbb{E}_{j-1}\left[\frac{\mathfrak{p}_{j,n}^\star-\mathfrak{p}_{j-\ell-1,n}^\star}{\pi^\star}\right]\right| \\
&\leq \sum_{j=1}^{n}\frac{1}{n}\mathbb{E}_{j-1}\left[\frac{\left|\mathfrak{p}_{j,n}^\star-\mathfrak{p}_{j-\ell-1,n}^\star\right|}{\pi^\star}\right] \leq C\Delta_n^{1/2} \longrightarrow 0.
\end{aligned}$$

Moreover, by using the boundedness of the Bernoulli variates,

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

which implies the AN of A_n. Finally, by Riemann integrability,

$$\sum_{j=1}^{n}\left(\frac{\mathfrak{p}_{j-1,n}^\star}{\pi^\star}\right)^m\frac{1}{n} \longrightarrow \frac{1}{(\pi^\star)^m}\int_0^1\left(\mathfrak{p}_s^\star\right)^m ds,$$

which completes the proof. □

As for the limiting distribution, the proof is divided into several steps. We first determine the joint limiting distribution for the following two quantities (see Lemma 9):

$$\widehat{\mathsf{RZ}}_{t,n} = \frac{1}{n}\sum_{j=1}^{n}\frac{\mathbb{B}_{j,n}^{(t)}}{\vartheta_{j,n}^\star} \qquad \widehat{\mathsf{RZ}}_{m,n}^{(t)} = \frac{1}{n}\sum_{j=1}^{n}\prod_{i=0}^{m-1}\frac{\mathbb{B}_{j+i,n}^{(t)}}{\vartheta_{j+i,n}^\star}.$$

However, in the previous definitions we pretend that the diurnal components $\vartheta_{j,n}^\star$ and $\vartheta_{j+1,n}^\star$ of $\mathfrak{p}_t$ are available to deflate the series of Bernoulli variates. In practice, these components are unobserved and estimated by $\widehat{\vartheta}_{j,n}^\star$ and $\widehat{\vartheta}_{j+1,n}^\star$. Lemma 10 enables us to prove the result of Theorem 4.1 and shows that the error induced from estimating $\vartheta_{j,n}^\star$ does not alter the analysis in Lemma 9. That is, the limiting distribution of $\widetilde{\mathsf{RZ}}_{m,n}^{(t)}$ is not altered.

Step I

Lemma 9. *Let $m \geq 2$ be a given integer number. Under Assumption 3, as $n \to \infty$,*

$$\sqrt{n}\begin{bmatrix}\widehat{\mathsf{RZ}}_{t,n} - \frac{1}{\pi^\star}\int_{t-1}^{t}\mathfrak{p}_s^\star ds \\ \widehat{\mathsf{RZ}}_{m,n}^{(t)} - \frac{1}{(\pi^\star)^m}\int_{t-1}^{t}\left(\mathfrak{p}_s^\star\right)^m ds\end{bmatrix} \xRightarrow{stably} \mathcal{MN}\left(0, \widetilde{\Sigma}^{(m)}\right), \tag{71}$$

where

$$\widehat{\mathsf{RZ}}_{t,n} = \frac{1}{n} \sum_{j=1}^{n} \frac{\mathbb{B}_{j,n}^{(t)}}{\vartheta_{j,n}^{\star}} \qquad \widehat{\mathsf{RZ}}_{m,n}^{(t)} = \frac{1}{n} \sum_{j=1}^{n} \prod_{i=0}^{m-1} \frac{\mathbb{B}_{j+i,n}^{(t)}}{\vartheta_{j+i,n}^{\star}}$$

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

$$\widetilde{\Sigma}^{(m)} = \begin{bmatrix} \int_0^1 \frac{\left(\mathfrak{p}_s^{\star} \vartheta_s^{-1} - (\mathfrak{p}_s^{\star})^2\right)}{(\pi^{\star})^2} \, ds & \int_0^1 \frac{m(\mathfrak{p}_s^{\star})^m \left(\vartheta_s^{-1} - \mathfrak{p}_s^{\star}\right)}{(\pi^{\star})^{m+1}} \, ds \\ \int_0^1 \frac{m(\mathfrak{p}_s^{\star})^m \left(\vartheta_s^{-1} - \mathfrak{p}_s^{\star}\right)}{(\pi^{\star})^{m+1}} \, ds & \int_0^1 \sum_{\ell=0}^{m-1} \frac{(2m-2\ell-1)(\mathfrak{p}_s^{\star})^{2m-\ell-1}\left(\vartheta_s^{-1} - \mathfrak{p}_s^{\star}\right)\vartheta_s^{-\ell}}{(\pi^{\star})^{2m}} \, ds \end{bmatrix}.$$

Proof. Again, in order simplify the notation, we present the proof for the case $t = 1$. Moreover, the proof is similar to that of Lemma 6 and therefore we omit some details. We consider the following decomposition:

$$\sqrt{n} \begin{bmatrix} \widehat{\mathsf{RZ}}_{1,n} - \frac{1}{\pi^{\star}} \int_0^1 p_s^{\star} \, ds \\ \widehat{\mathsf{RZ}}_{1,n}^{(m)} - \frac{1}{(\pi^{\star})^m} \int_0^1 (p_s^{\star})^m \, ds \end{bmatrix} = A_1 + A_2,$$

where

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

It is easy to see that A_2 is AN. Thus, it is enough to prove that $A_1 \overset{stably}{\Longrightarrow} \mathcal{MN}\left(0, \widetilde{\Sigma}^{(m)}\right)$. In particular, it is sufficient to establish the convergence

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

where $\eta_j = [\eta_j(1) \ \eta_j(2)]'$, with

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

and

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

At this point we note that if for any finite numbers $k, \ell, d \geq 0$ and powers $q_1, \ldots, q_d \geq 0$, we set

$$\xi_{j,n} = \underbrace{\frac{\mathbb{B}_{j-\ell-k+1,n} \cdots \mathbb{B}_{j-\ell,n}}{\vartheta_{j-\ell-k+1,n}^{\star} \cdots \vartheta_{j-\ell,n}^{\star}}}_{k \text{ factors}} \times \underbrace{\frac{\mathbb{B}_{j-\ell+1,n} \cdots \mathbb{B}_{j,n}}{\left(\vartheta_{j-\ell+1,n}^{\star} \cdots \vartheta_{j,n}^{\star}\right)^2}}_{\ell \text{ factors}} \times \underbrace{\left(\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j+1,n}}{\vartheta_{j+1,n}^{\star}}\right]\right)^{q_1} \cdots \left(\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j+d,n}}{\vartheta_{j+d,n}^{\star}}\right]\right)^{q_d}}_{d \text{ factors}},$$

then, as $n \to \infty$,

$$\frac{1}{n} \sum_{j=m}^{n} \xi_{j,n} \overset{p}{\to} \int_0^1 \frac{(\mathfrak{p}_s^{\star})^{k+\ell+v} \vartheta_s^{-\ell}}{(\pi^{\star})^{k+2\ell+v}} \, ds,$$

where $v = q_1 + \cdots + q_d$. This result is analogous to Lemma 5 and can be proven in a similar way.

Now, consider each component of the matrix $\eta_j \eta_j'$ separately:

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

64

Hence, by the convergence result above,

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

Now consider the product:

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

Note that

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

and

$$\widetilde{\zeta}_{j-\ell,\ell}^{(m)}\,\widetilde{\zeta}_{j-\ell-k,\ell+k}^{(m)}$$

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

$$\times\frac{\mathbb{B}_{j-\ell-k,n}\,\mathbb{B}_{j-\ell-k+1,n}\cdots\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j-1,n}}{\vartheta_{j-\ell-k,n}^\star\,\vartheta_{j-\ell-k+1,n}^\star\cdots\vartheta_{j-\ell,n}^\star\cdots\vartheta_{j-1,n}^\star}\left(\frac{\mathbb{B}_{j,n}}{\vartheta_{j,n}^\star}-\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j,n}}{\vartheta_{j,n}^\star}\right]\right)\left(\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j+1,n}}{\vartheta_{j+1,n}^\star}\right]\cdots\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j+m-\ell-k-1,n}}{\vartheta_{j+m-\ell-k-1,n}^\star}\right]\right)$$

$$=\underbrace{\frac{\mathbb{B}_{j-\ell-k,n}\cdots\mathbb{B}_{j-\ell-1,n}}{\vartheta_{j-\ell-k,n}^\star\cdots\vartheta_{j-\ell-1,n}^\star}}_{k\text{ factors}}\underbrace{\frac{\mathbb{B}_{j-\ell,n}\cdots\mathbb{B}_{j-1,n}}{\left(\vartheta_{j-\ell,n}^\star\cdots\vartheta_{j-1,n}^\star\right)^2}}_{\ell\text{ factors}}\left(\frac{\mathbb{B}_{j,n}}{\vartheta_{j,n}^\star}-\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j,n}}{\vartheta_{j,n}^\star}\right]\right)^2\times$$

$$\underbrace{\left(\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j+1,n}}{\vartheta_{j+1,n}^\star}\right]\cdots\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j+1,n}}{\vartheta_{j+m-\ell-k-1,n}^\star}\right]\right)^2}_{m-(\ell+k)-1\text{ factors}}\underbrace{\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j+m-\ell-k,n}}{\vartheta_{j+m-\ell-k,n}^\star}\right]\cdots\mathbb{E}_{j-1}\left[\frac{\mathbb{B}_{j+m-\ell-1,n}}{\vartheta_{j+m-\ell-1,n}^\star}\right]}_{k\text{ factors}}.$$

Hence

$$\frac{1}{n}\sum_{j=m}^{n}\left(\widetilde{\zeta}_{j-\ell,\ell}^{(m)}\right)^2\xrightarrow{p}\int_0^1\frac{(\mathfrak{p}_s^\star)^{2m-\ell-1}\left(\vartheta_s^{-1}-\mathfrak{p}_s^\star\right)\vartheta_s^{-\ell}}{\left(\pi^\star\right)^{2m}}\,ds,$$

and

$$\frac{1}{n}\sum_{j=m}^{n}\widetilde{\zeta}_{j-\ell,\ell}^{(m)}\,\widetilde{\zeta}_{j-\ell-k,\ell+k}^{(m)}\xrightarrow{p}\int_0^1\frac{(\mathfrak{p}_s^\star)^{2m-\ell-1}\left(\vartheta_s^{-1}-\mathfrak{p}_s^\star\right)\vartheta_s^{-\ell}}{\left(\pi^\star\right)^{2m}}\,ds,$$

Consequently,

$$\frac{1}{n}\sum_{j=m}^{n}\mathbb{E}_{j-1}\left[\eta_j(2)\eta_j(2)\right]\xrightarrow{p}\int_0^1\sum_{\ell=0}^{m-1}\frac{(2m-2\ell-1)(\mathfrak{p}_s^\star)^{2m-\ell-1}\left(\vartheta_s^{-1}-\mathfrak{p}_s^\star\right)\vartheta_s^{-\ell}}{\left(\pi^\star\right)^{2m}}\,ds$$

Finally, for the covariance term, we have:

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

Consequently,

$$
\frac{1}{n}\sum_{j=m}^n \mathbb{E}_{j-1}\left[\eta_j(1)\eta_j(2)\right] \xrightarrow{p} \int_0^1 \frac{m\,(\mathfrak{p}_s^\star)^m\,(\vartheta_s^{-1} - \mathfrak{p}_s^\star)}{(\pi^\star)^{m+1}}\,ds,
$$

which completes the proof. $\qquad\square$

Step II

Lemma 10. *Let $\widehat{\vartheta}_{j,n}^\star$ be the long-scale estimator of $\vartheta_{j,n}^\star$. Then, as $T \to \infty$, under Assumption 3 we have*

$$
\widehat{\vartheta}_{j,n}^\star = \vartheta_{j,n}^\star + O_p\left(T^{-1/2}\right).
$$

Proof. First, we consider the following decomposition:

$$
\widehat{\vartheta}_{j,n}^\star - \vartheta_{j,n}^\star = \frac{1}{T}\sum_{t=1}^T \left(\mathbb{B}_{j,n}^{(t)} - \mathfrak{p}_{j,n}^{(t)}\right) + \frac{1}{T}\sum_{t=1}^T \mathfrak{p}_{j,n}^{(t)} - \vartheta_{j,n}^\star.
$$

Now, the proof is complete if we can show that:

$$
\frac{1}{T}\sum_{t=1}^T \left(\mathbb{B}_{j,n}^{(t)} - \mathfrak{p}_{j,n}^{(t)}\right) = O_{L^2}\left(T^{-1/2}\right) \tag{72}
$$

$$
\frac{1}{T}\sum_{t=1}^T \mathfrak{p}_{j,n}^{(t)} - \vartheta_{j,n}^\star = O_{L^2}\left(T^{-1/2}\right) \tag{73}
$$

As for (72), note that by conditioning on $\mathfrak{p}_{j,n}^{(t)}$'s we obtain:

$$
\mathbb{V}\mathrm{ar}\left[\frac{1}{T}\sum_{t=1}^T \left(\mathbb{B}_{j,n}^{(t)} - \mathfrak{p}_{j,n}^{(t)}\right)\right] = \frac{1}{T^2}\sum_{t=1}^T \mathbb{E}\left[\mathfrak{p}_{j,n}^{(t)} - \left(\mathfrak{p}_{j,n}^{(t)}\right)^2\right] \leq \frac{K}{T},
$$

where $K > 0$ is some constant and the last inequality follows from the boundedness of $\mathfrak{p}_t$. Therefore, we have:

$$
\frac{1}{T}\sum_{t=1}^T \left(\mathbb{B}_{j,n}^{(t)} - \mathfrak{p}_{j,n}^{(t)}\right) = O_{L^2}\left(T^{-1/2}\right).
$$

Hence (72) follows.

Next, (73) can be verified as follows. First, we write:

$$
\frac{1}{T}\sum_{t=1}^T \mathfrak{p}_{j,n}^{(t)} - \vartheta_{j,n}^\star = \vartheta_{j,n}\left(\frac{1}{T}\sum_{t=1}^T \mathfrak{p}_{j,n}^{\star(t)} - \pi^\star\right).
$$

66

Then, due to Assumption 3 for some constant K, we have

$$\mathbb{V}\mathrm{ar}\left[\frac{1}{T}\sum_{t=1}^{T}p_{j,n}^{\star(t)}-\pi^{\star}\right]\leq\frac{K}{T}\left(1+2\sum_{k=1}^{\infty}\mathbb{C}\mathrm{ov}\left[\mathfrak{p}_{t+j/n}^{\star},\mathfrak{p}_{t+k+j/n}^{\star}\right]\right),$$

which implies that $\frac{1}{T}\sum_{t=1}^{T}\mathfrak{p}_{j,n}^{\star(t)}-\pi^{\star}=O_{L^{2}}\left(T^{-1/2}\right)$. Since $\vartheta_{j,n}$ is bounded, this also implies that:

$$\frac{1}{T}\sum_{t=1}^{T}\mathfrak{p}_{j,n}^{(t)}-\vartheta_{j,n}^{\star}=O_{L^{2}}\left(T^{-1/2}\right).$$

$\square$

Step III We prove now the second part of Theorem 4.1.

Consider the decomposition:

$$\sqrt{n}\begin{bmatrix}\widetilde{\mathsf{RZ}}_{1,n}^{(t)}-\frac{1}{\pi^{\star}}\int_{t-1}^{t}\mathfrak{p}_{s}^{\star}\,ds\\ \widetilde{\mathsf{RZ}}_{m,n}^{(t)}-\frac{1}{\pi^{\star}}\int_{t-1}^{t}(\mathfrak{p}_{s}^{\star})^{m}\,ds\end{bmatrix}=A_{1}+A_{2},$$

where

$$A_{1}=\sqrt{n}\begin{bmatrix}\frac{1}{n}\sum_{j=1}^{n}\frac{\mathbb{B}_{j,n}^{(t)}}{\widehat{\vartheta}_{j,n}^{\star}}-\frac{1}{n}\sum_{j=1}^{n}\frac{\mathbb{B}_{j,n}^{(t)}}{\vartheta_{j,n}^{\star}}\\ \frac{1}{n-m}\sum_{j=1}^{n-m+1}\prod_{q=0}^{m-1}\frac{\mathbb{B}_{j+q,n}^{(t)}}{\widehat{\vartheta}_{j+q,n}^{\star}}-\frac{1}{n-m}\sum_{j=1}^{n-m+1}\prod_{q=0}^{m-1}\frac{\mathbb{B}_{j+q,n}^{(t)}}{\vartheta_{j+q,n}^{\star}}\end{bmatrix},\quad A_{2}=\sqrt{n}\begin{bmatrix}\frac{1}{n}\sum_{j=1}^{n}\frac{\mathbb{B}_{j,n}^{(t)}}{\vartheta_{j,n}^{\star}}-\frac{1}{\pi^{\star}}\int_{t-1}^{t}\mathfrak{p}_{s}^{\star}\,ds\\ \frac{1}{n-m}\sum_{j=1}^{n-m+1}\prod_{q=0}^{m-1}\frac{\mathbb{B}_{j+q,n}^{(t)}}{\vartheta_{j+q,n}^{\star}}-\frac{1}{\pi^{\star}}\int_{t-1}^{t}(\mathfrak{p}_{s}^{\star})^{m}\,ds\end{bmatrix}.$$

By Lemma 9,

$$A_{2}\overset{stably}{\Longrightarrow}\mathcal{MN}\left(0,\widetilde{\Sigma}^{(m)}\right).$$

Hence, it is sufficient to show that A_{1} becomes negligible as $T\to\infty$. For simplicity we consider only the first component of A_{1}, as the second component can be treated analogously. We have:

$$A_{1}(1)=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\frac{\mathbb{B}_{j,n}^{(t)}}{\widehat{\vartheta}_{j,n}^{\star}}-\frac{\mathbb{B}_{j,n}^{(t)}}{\vartheta_{j,n}^{\star}}\right)=\sum_{j=1}^{n}\frac{1}{\sqrt{n}}\cdot\frac{\vartheta_{j,n}^{\star}-\widehat{\vartheta}_{j,n}^{\star}}{\vartheta_{j,n}^{\star}\widehat{\vartheta}_{j,n}^{\star}}\mathbb{B}_{j,n}^{(t)}.$$

By Lemma 10, $\vartheta_{j,n}^{\star}-\widehat{\vartheta}_{j,n}^{\star}=O_{p}(T^{-1/2})$. Consequently, since $0\leq\frac{\mathbb{B}_{j,n}^{(t)}}{\vartheta_{j,n}^{\star}\widehat{\vartheta}_{j,n}^{\star}}\leq K$, for some constant K, we have:

$$A_{1}(1)=O_{p}(n^{1/2}T^{-1/2}),$$

which completes the proof.

Proof of Corollary 2. Combining Theorem 4.1 and Lemma 3 we deduce that on $\Omega_{t,0}^{\star}$,

$$\frac{\sqrt{n}\left(\widetilde{\mathsf{RZ}}_{m,n}^{(t)}-\left(\widetilde{\mathsf{RZ}}_{1,n}^{(t)}\right)^{m}\right)}{\sqrt{m^{2}\left(\widetilde{\mathsf{RZ}}_{1,n}^{(t)}\right)^{2(m-1)}\widetilde{\Sigma}_{(1,1)}^{(m)}+\widetilde{\Sigma}_{(2,2)}^{(m)}-2m\left(\widetilde{\mathsf{RZ}}_{1,n}^{(t)}\right)^{m-1}\widetilde{\Sigma}_{(1,2)}^{(m)}}}\overset{stably}{\Longrightarrow}\mathcal{N}\left(0,1\right),\tag{74}$$

where $\widetilde{\Sigma}_{(1,1)}^{(m)}$, $\widetilde{\Sigma}_{(2,2)}^{(m)}$ and $\widetilde{\Sigma}_{(1,2)}^{(m)}$ denote the elements of $\widetilde{\Sigma}^{(m)}$. Hence, in order to prove that $\Upsilon_{m,n}^{(t)}\overset{stably}{\Longrightarrow}\mathcal{N}(0,1)$ on $\Omega_{t,0}^{\star}$ it is enough to show that $\widehat{\widetilde{\Sigma}}_{(k,\ell)}^{(m)}\overset{p}{\longrightarrow}\widetilde{\Sigma}_{(k,\ell)}^{(m)}$ for $k,\ell=1,2$. Consider estimating $\widetilde{\Sigma}_{(1,1)}^{(m)}$. Notice that, on $\Omega_{t,0}^{\star}$, $\mathfrak{p}_{s}^{\star}=\mathfrak{p}_{t-1}^{\star}$,

$\forall s \in [t-1, t]$. Hence,

$$\widetilde{\Sigma}_{(1,1)}^{(m)} = \int_{t-1}^{t} \frac{\left(\mathfrak{p}_s^\star \vartheta_s^{-1} - (\mathfrak{p}_s^\star)^2\right)}{(\pi^\star)^2} \, ds = \frac{\mathfrak{p}_{t-1}^\star}{(\pi^\star)^2} \int_{t-1}^{t} \vartheta_s^{-1} \, ds - \left(\frac{\mathfrak{p}_{t-1}^\star}{\pi^\star}\right)^2.$$

By Lemma 10, $\widehat{\vartheta}_{j,n}^\star = \vartheta_{j,n}^\star + O_p\left(T^{-1/2}\right)$. Since $\vartheta_{j,n}^\star$ is bounded away from zero and bounded from above we also have

$$\left(\widehat{\vartheta}_{j,n}^\star\right)^{-\ell} = \left(\vartheta_{j,n}^\star\right)^{-\ell} + O_p\left(T^{-1/2}\right),$$

for every real ℓ. Next, we have:

$$\frac{1}{n} \sum_{j=1}^{n} \left(\widehat{\vartheta}_{j,n}^\star\right)^{-\ell} = \frac{1}{n} \sum_{j=1}^{n} \left(\vartheta_{j,n}^\star\right)^{-\ell} + O_p\left(T^{-1/2}\right).$$

By the Riemann integrability of ϑ_s,

$$\frac{1}{n} \sum_{j=1}^{n} \left(\vartheta_{j,n}^\star\right)^{-\ell} \longrightarrow \int_{t-1}^{t} (\vartheta_s^\star)^{-\ell} \, ds, \tag{75}$$

hence:

$$\frac{1}{n} \sum_{j=1}^{n} \left(\widehat{\vartheta}_{j,n}^\star\right)^{-\ell} \xrightarrow{p} \int_{t-1}^{t} (\vartheta_s^\star)^{-\ell} \, ds = (\pi^\star)^{-\ell} \int_{t-1}^{t} (\vartheta_s)^{-\ell} \, ds. \tag{76}$$

In particular, due to the fact that $\int_0^1 \vartheta_s \, ds = 1$ (by Assumption 3), setting $\ell = -1$ gives:

$$\frac{1}{n} \sum_{j=1}^{n} \widehat{\vartheta}_{j,n}^\star \xrightarrow{p} \pi^\star \int_{t-1}^{t} \vartheta_s \, ds = \pi^\star. \tag{77}$$

Consequently,

$$\frac{1}{n} \sum_{j=1}^{n} \left(\frac{\widehat{\vartheta}_{j,n}^\star}{\frac{1}{n} \sum_{j=1}^{n} \widehat{\vartheta}_{j,n}^\star}\right)^{-\ell} \xrightarrow{p} \int_{t-1}^{t} \vartheta_s^{-\ell} \, ds. \tag{78}$$

Finally, notice that on $\Omega_{t,0}^\star$,

$$\widetilde{\mathsf{RZ}}_{1,n}^{(t)} \xrightarrow{p} \frac{\mathfrak{p}_{t-1}^\star}{\pi^\star}. \tag{79}$$

Combining (79), (78), (77) and the expression for $\widetilde{\Sigma}_{(1,1)}^{(m)}$, we obtain

$$\widehat{\widetilde{\Sigma}}_{(1,1)}^{(m)} = \frac{1}{n} \sum_{j=1}^{n} \frac{\left(\widehat{\vartheta}_n^* \, \widetilde{\mathsf{RZ}}_{1,n}^{(t)} \left(\frac{\widehat{\vartheta}_{j,n}^*}{\widehat{\vartheta}_n^*}\right)^{-1} - \left(\widehat{\vartheta}_n^* \widetilde{\mathsf{RZ}}_{1,n}^{(t)}\right)^2\right)}{\left(\widehat{\vartheta}_n^*\right)^2} \xrightarrow{p} \widetilde{\Sigma}_{(1,1)}^{(m)},$$

where $\widehat{\theta}_n^* = \frac{1}{n} \sum_{j=1}^{n} \widehat{\theta}_{j,n}^*$. With a similar reasoning one can prove the remaining convergences, which implies that:

$$\Upsilon_{m,n}^{(t)} \stackrel{stably}{\Longrightarrow} \mathcal{N}(0,1), \quad \text{on } \Omega_{t,0}^\star.$$

Finally, we prove that $\Upsilon_{m,n}^{(t)}$ is unbounded in probability on $\Omega_{t,1}^\star$. By Theorem 4.1, for any integer $m \geq 2$,

$$\widetilde{\mathsf{RZ}}_{m,n}^{(t)} - \left(\widetilde{\mathsf{RZ}}_{1,n}^{(t)}\right)^m \stackrel{u.c.p.}{\longrightarrow} \frac{1}{(\pi^\star)^m} \left(\int_{t-1}^{t} (\mathfrak{p}_s^\star)^m \, ds - \left(\int_{t-1}^{t} \mathfrak{p}_s^\star \, ds\right)^m\right) > 0. \tag{80}$$

Consequently,

$$\sqrt{n} \left(\widetilde{\mathsf{RZ}}_{m,n}^{(t)} - \left(\widetilde{\mathsf{RZ}}_{1,n}^{(t)}\right)^m\right) \xrightarrow{p} \infty. \tag{81}$$

68

Due to the boundedness of the Bernoulli random variables, the denominator of $\Upsilon_{m,n}^{(t)}$ is bounded, and, therefore,

$$\Upsilon_{m,n}^{(t)} \xrightarrow{p} \infty, \quad \text{on } \Omega_{t,1}^{\star},$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

B   Appendix: Additional simulations

In the Monte Carlo simulations reported in Section 5, the log-price paths are generated with the initial condition $Y_0 = \log(P_0)$, $P_0 = 100$ and the volatility factor $c_\sigma = 3$, which corresponds to a daily volatility, $\overline{\Sigma}$, of (roughly) 2%. Here we conduct additional simulations varying both P_0 and c_σ, while keeping other parameters unchanged. In particular, we consider $P_0 \in \{5, 15, 30, 75, 100\}$ and we also consider four possible values for c_σ, chosen in such a way that the daily average realized volatility $\overline{\Sigma} \in \{2\%, 6\%, 10\%, 13\%\}$. Hence, in our simulations the percentage of zeros attributed to rounding ranges from roughly 95% (when $P_0 = 5$ and $\overline{\Sigma} \approx 2\%$) to roughly 5% (when $P_0 = 100$ and $\overline{\Sigma} \approx 13\%$). The number of zeros produced by rounding declines as either P_0 or c_σ increases.

Figure 7 presents the kernel densities of $\Psi_{m,n}$ computed on 10^4 replications of prices under the null ($\omega \in \Omega_{0,\Psi}$) or the two alternatives ($\omega \in \Omega_{0,\Upsilon}$ and $\omega \in \Omega_{1,\Upsilon}$). The figure shows that, if $P_0 > 15$ and $\overline{\Sigma} > 6\%$, $\Psi_{m,n}$ is correctly centred under the null and displays only a slight distortion due to price discreteness. For smaller values of P_0 and $\overline{\Sigma}$, the test statistic $\Psi_{m,n}$ is negatively biased and the power of the test is reduced accordingly. However, it improves steadily as P_0 and $\overline{\Sigma}$ grow large. For all values of P_0 and $\overline{\Sigma}$, the rejection rate is higher on $\Omega_{1,\Upsilon}$, as expected.

Figure 8 shows the kernel densities of $\Upsilon_{m,n}^{(t)}$ over 10^4 replications of price paths under the null ($\omega \in \Omega_{0,\Upsilon}$) and the alternative ($\omega \in \Omega_{1,\Upsilon}$). Similarly to the case of $\Psi_{m,n}$, the test statistic $\Upsilon_{m,n}^{(t)}$ shows good size and power for initial prices larger than 15 and daily average volatility greater than 2%. For smaller values of P_0 and $\overline{\Sigma}$, the test is negatively biased and shows a poor rejection power.
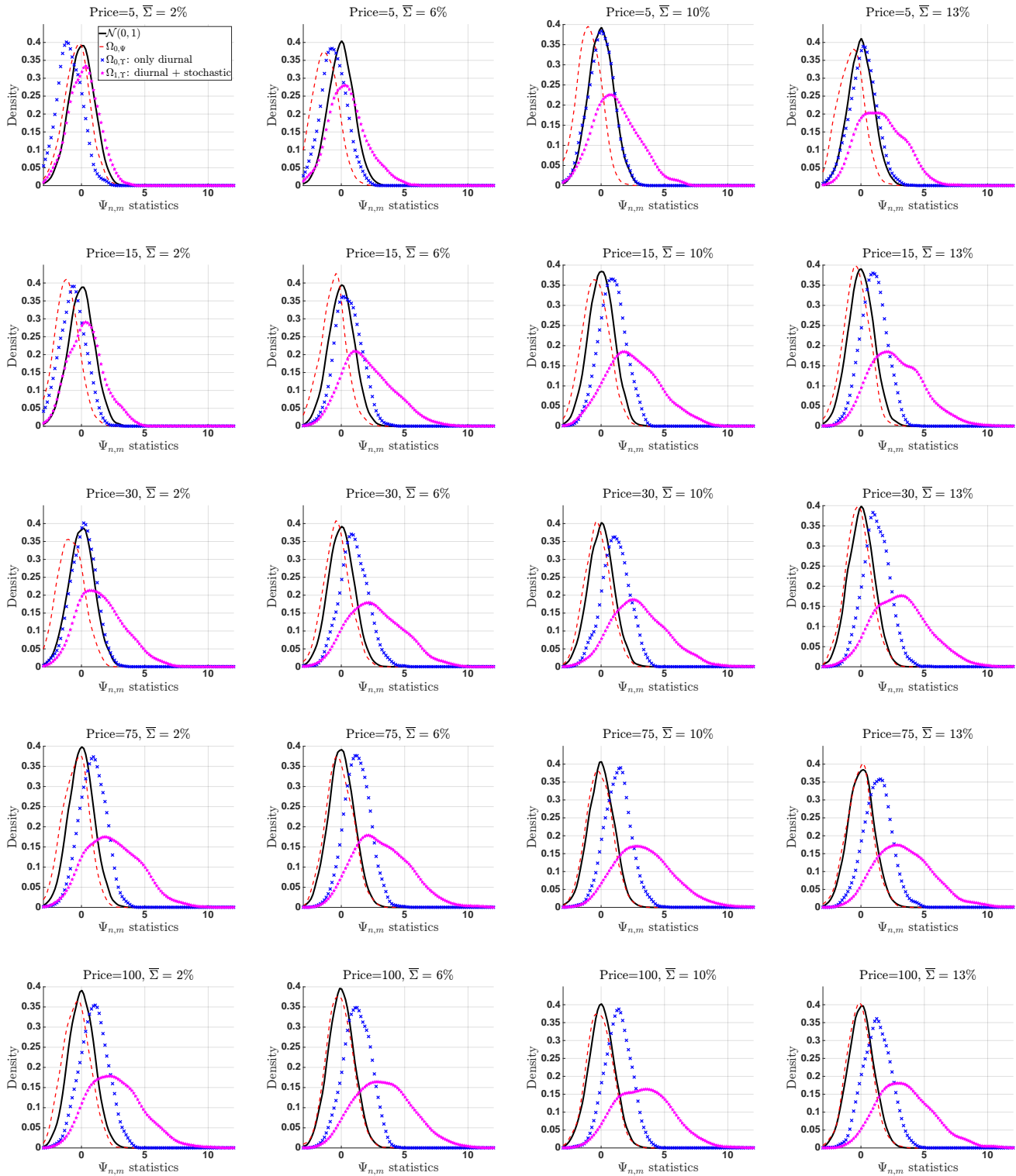
Figure 7: *From left to right, from top to bottom*: This plot reports the kernel density estimates of the test statistic $\Psi_{m,n}$ under the null $\Omega_{0,\Psi}$ (dotted red line) and the two alternatives $\Omega_{0,\Upsilon}$ (blue crosses) and $\Omega_{1,\Upsilon}$ (magenta star) for different values of the daily average volatility $\overline{\Sigma}$ and of the initial price P_0. We set $n = 780$. Prices are rounded at one cent.
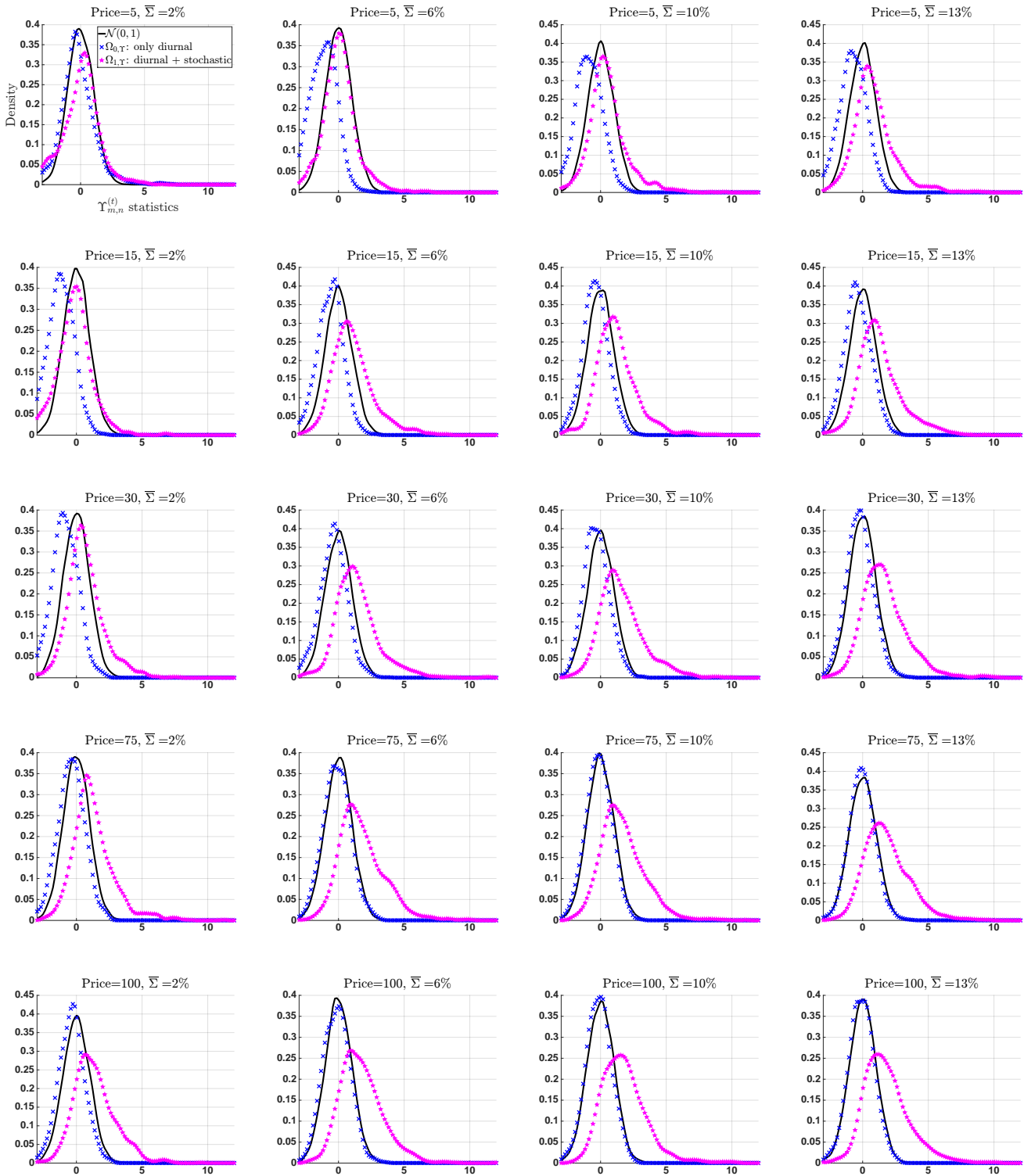
Figure 8: *From left to right, from top to bottom*: This plot reports the kernel density estimates of the test statistic $\Upsilon_{m,n}^{(t)}$ under the null $\Omega_{0,\Upsilon}$ (blue crosses) and the alternative $\Omega_{1,\Upsilon}$ (magenta stars) and $\Omega_{1,\Upsilon}$ (magenta star) for different values of the daily average volatility $\overline{\Sigma}$ and of the initial price P_0. We set $n = 780$. Prices are rounded at one cent.