Intraday Volatility Dynamics*

Carsten H. Chong† Marc Hoffmann‡ Mathieu Rosenbaum§
Grégoire Szymanski¶

Abstract

Return-based spot volatility estimates are noisy and prone to biases. This paper develops inference for the autocorrelation of intraday volatility changes that is robust to sampling errors in volatility estimates and biases due to jumps and microstructure noise. Our procedure builds on two results: First, we show that the limiting autocorrelation function of increments of a continuous-time process over shrinking horizons is parametrically determined and must coincide with that of fractional Gaussian noise. Second, we characterize the infill asymptotic distribution of realized autocovariances of spot variance estimators. The two results combine to yield a feasible generalized method of moments estimator of the autocorrelation of high-frequency volatility increments. In an empirical application to SPY transaction data, we document negative serial correlation in intraday movements of both latent and realized volatility.

Keywords: High-frequency data, realized autocovariance, regular variation, rough volatility, smoothness parameter.

*We thank Jia Li, Nour Meddahi, Viktor Todorov, Xinghua Zheng (discussant) and seminar and conference participants at Singapore Management University, the 17th Annual SoFiE Conference, the 13th Days of Econometrics for Finance in Rabat and the 2026 Financial Econometrics Conference at Toulouse School of Economics for insightful comments. We also thank Ferretti Berlincioni for help with preparing the empirical data. Carsten Chong acknowledges financial support from ECS Grant 26301724.

†Department of Information Systems, Business Statistics and Operations Management, The Hong Kong University of Science and Technology, e-mail: carstenchong@ust.hk.

‡CEREMADE, Université Paris-Dauphine-PSL, e-mails: hoffmann@ceremade.dauphine.fr

§CEREMADE, Université Paris-Dauphine-PSL, e-mails: mathieu.rosenbaum@dauphine.psl.eu

¶Department of Mathematics, University of Luxembourg, e-mail: gregoire.szymanski@uni.lu

1 Introduction

This paper conducts inference about the dynamical properties of intraday asset price volatility. We use high-frequency return data to construct an estimator of the autocorrelation of short-term volatility changes and derive its limiting distribution as the sampling frequency tends to infinity. Our estimator overcomes the latency of volatility and is robust to market microstructure noise and jumps in the underlying price and volatility dynamics.

Because spot volatility estimates are subject to sampling errors, the sample variance of their increments is a biased estimator of the population variance of volatility increments. In particular, the sample autocorrelation of first-differenced spot volatility estimates, which contains their sample variance in the denominator, is *inconsistent* for the autocorrelation of changes in true, unobserved, volatility. Key to the robustness of our estimator is an econometric procedure that allows us to estimate the autocorrelation of volatility increments at high frequency *without* estimating their variance.

In a first step, we show that for a continuous-time process satisfying mild regularity conditions, the convergence of the autocorrelation function (ACF) of its increments in the high-frequency limit implies that the variance of its increments, viewed as a function of the sampling frequency, must be regularly varying at the origin in the sense of Karamata (1930). We show that, as a consequence, the limiting ACF must coincide with that of fractional Gaussian noise.[1] We refer to the associated fractional parameter as the *smoothness parameter* of the process. Inference about the ACF of volatility increments, a nonparametric problem at any fixed frequency, thus reduces to a one-dimensional parametric problem under high-frequency asymptotics.

In a second step, we identify the asymptotic distribution of realized autocovariances of spot variance estimators. We start by estimating spot variance in the classical way, using truncated sums of squared returns over local blocks of high-frequency data. We then form first-order differences of these spot variance estimators and consider their sample autocovariances. Under mild structural assumptions on the data generating process, we show that any finite number of sample autocovariances at lags greater than or equal to two is consistent for their population counterparts with a mixed normal asymptotic distribution. We extend this central limit theorem (CLT) by including the sample variance plus twice the first-order sample autocovariance of increments of spot variance estimators. This particular

[1] Fractional Gaussian noise denotes the first-order difference of Mandelbrot and Van Ness (1968)'s fractional Brownian motion. The latter appears, for instance, in continuous-time limits of fractional time series models (see e.g., Sowell (1990), Comte and Renault (1996) and Müller and Watson (2008)).

linear combination ensures that sampling errors inherent to spot variance estimators cancel out. Crucially, we do not, and need not, consider autocovariances at lags zero and one separately, which individually contain biases due to sampling errors.

In a third step, we use a generalized method of moments (Hansen (1982), GMM) approach to derive consistent estimators and asymptotic confidence intervals for the smoothness parameter of volatility and, by an application of the delta method, for the limiting autocorrelation of volatility increments. This estimator minimizes the weighted sum of squared deviations between the sample autocovariances and their theoretical counterparts, with data-driven weights that minimize the asymptotic variance of the resulting estimator.

In an empirical application to transaction data of SPY, the exchange-traded fund tracking the S&P 500 index, we find strong evidence that short-term changes in spot volatility are negatively autocorrelated. Yearly estimates from 2011 to 2025 of the first-order autocorrelation coefficient are significant at a 5% level and negative throughout the sample, ranging from -0.42 to -0.21. This finding obtains in addition to the well-established diurnal pattern of intraday volatility. Our analysis further reveals that predictability, while attenuated, prevails in increments of short-horizon *realized* variance. We show that linear prediction of changes in 30-minute realized variance significantly improves both in sample and out of sample upon including lags of higher order.

What are the broader implications of this empirical result? As we show, classical statistics such as realized power variation of order different from two can have biases of the order Δ_n^{2e+1}, where Δ_n is the sampling frequency and e is the smoothness parameter of volatility. Therefore, if $e < -0.25$, these biases converge to zero at a slower rate than the standard error of the estimator, which is $O(\Delta_n^{1/2})$. This creates a trade-off for the econometrician between asymptotic efficiency in the mean square error sense, which improves with smaller values of Δ_n, and asymptotic inferential validity based on standard CLTs, which deteriorates as $\Delta_n \to 0$. A similar phenomenon occurs in high-frequency econometrics when asset price jumps are of infinite variation (Jacod and Todorov (2014)) or in other econometric contexts such as regression-discontinuity designs (Calonico et al. (2014)).

Related Literature. It has long been recognized that volatility is time-varying. Early models of stochastic volatility primarily describe daily or low-frequency dynamics of volatility. These include the GARCH models of Engle (1982) and Bollerslev (1986), the time-changed models studied by Clark (1973) and Tauchen and Pitts (1983), and continuous-time models like those of Hull and White (1987) and Heston (1993) which aim at option pricing

applications with horizons of more than one day.

Our paper, by contrast, focuses on the *intraday* dynamics of stochastic volatility and, in particular, on the autocorrelation of short-term volatility changes. It is well known since the early works on realized measures (Andersen et al. (2001, 2003), Barndorff-Nielsen and Shephard (2002)) that the econometric analysis of high-frequency intraday data involves infill asymptotics, where the time interval is fixed and the sampling frequency tends to infinity. A key insight of this paper is that for a broad class of continuous-time stochastic process, the ACF of its increments in the infill asymptotic limit, if it exists, is parametrically determined and coincides with that of fractional Gaussian noise.

Our paper is related to the literature on estimation of spot volatility and its properties. The idea of estimating spot volatility from local blocks of data appears in Foster and Nelson (1996) and has been extended in various directions; we refer to Chapter 8 of Aït-Sahalia and Jacod (2014) for an overview. The focus in this paper is not on estimating spot volatility itself but its dynamical properties within a trading day. In contrast to works that analyze deterministic patterns of volatility (e.g., Wood et al. (1985), Baillie and Bollerslev (1991) and Andersen et al. (2024)) or use spot volatility estimates for intraday event studies (e.g., Bollerslev et al. (2018, 2021, 2024)), we examine the second-order properties of intraday volatility on typical days without scheduled events.

Our paper contributes to a growing literature studying the dynamical properties of volatility using fractional instead of standard Brownian motion (Comte and Renault (1998), Gatheral et al. (2018)). While previous literature typically imposes either a semimartingale or a fractional volatility model as an assumption,[2] we provide a novel interpretation of this model choice: semimartingale models imply asymptotically uncorrelated volatility increments, while fractional models imply that volatility increments are correlated in the high-frequency limit. Furthermore, we derive asymptotic confidence intervals for the limiting autocorrelation of volatility increments that are valid in both semimartingale and fractional settings and are simultaneously robust to sampling errors, noise and jumps.

Our paper further shows that the smoothness parameter of a continuous-time stationary process is distinct from its memory parameter. While the latter concerns the low-frequency

[2]Papers investigating the dynamical properties of volatility include Bandi and Renò (2012), Aït-Sahalia et al. (2013, 2017), Wang and Mykland (2014), Kalnina and Xiu (2017), Li et al. (2022) and Chong and Todorov (2024) for volatility of volatility and the leverage effect and Jacod and Todorov (2010), Todorov and Tauchen (2011) and Bandi and Renò (2016) for volatility jumps, all in a semimartingale setting, as well as Bennedsen et al. (2022), Bolko et al. (2023), Chong et al. (2024), Wang et al. (2023) and Chong and Todorov (2025) in a fractional setting.

variation of the process (i.e., the decay of its ACF at infinity or, equivalently, the behavior of its spectral density at the origin), the former governs the high-frequency variation of its increments (i.e., the behavior of the ACF at the origin or, equivalently, the decay of the spectral density at infinity). Both smoothness and memory properties, however, are determined by a single Hurst exponent in standard stationary models based on fractional Brownian motion. Parametric inference in such models, if misspecified, can therefore draw on both high- and low-frequency information of the observations and lead to inconclusive findings about volatility smoothness and memory (see e.g., Shi and Yu (2023) and Li et al. (2025b)). The need to separate high- and low-frequency time series information has been observed earlier by Müller and Watson (2008) in a general context and by Bennedsen et al. (2022) for volatility. Our results contribute to this literature by enabling model-free inference about volatility smoothness while being silent on whether volatility has long or short memory. In particular, our empirical finding that volatility has negatively correlated increments complements, but neither supports nor contradicts, well-established results about long-memory volatility (see e.g., Andersen et al. (2001, 2003)).

Our paper studies a very particular instance of errors in variables. Because spot variance estimates contain sampling errors, the sample autocorrelation of their increments suffers from an attenuation bias (Hausman (2001)). While existing approaches to obtaining valid inferences in the presence of mismeasured variables often rely on instrumental variables (Hausman et al. (1991), Schennach (2007), Hu and Schennach (2008)), auxiliary information (Hausman et al. (1991), Schennach (2004), Chen et al. (2005)) or bias correction (Chesher (1991), Evdokimov and Zeleneev (2024)), our solution exploits the fact that the sampling errors inherent to spot variance estimates—while potentially correlated with the latter—are uncorrelated in the time series dimension. That is, the measurement errors in spot variance *increments* form an approximate MA(1) process with moving average parameter -1. As a consequence, we show that the sum of realized variance and twice the first-order realized autocovariance as well as all higher-order realized autocovariances of spot variance estimators are asymptotically unbiased. Together with the fractional form of the limiting ACF, this allows for asymptotically valid inference about the correlation of volatility changes without estimating their variance.

Outline. Section 2 introduces the smoothness parameter of a continuous-time process as a model-free way to characterize the ACF of its increments in the infill asymptotic limit. Section 3 presents a law of large numbers (LLN) and a CLT for the realized autocovari-

ances of increments of spot variance estimators. Section 4 constructs GMM estimators of the smoothness parameter of volatility and of the limiting autocorrelation of volatility increments. Section 5 evaluates the performance of this estimator in finite samples, while Section 6 contains an empirical application to SPY transaction data. Section 7 discusses implications of low volatility smoothness on estimators commonly used in high-frequency econometrics. Section 8 concludes. Appendices A and B contain the proofs for the main results of Section 2 and the assumptions behind the LLN and CLT of Section 3. Other proofs are deferred to the Supplementary Material.

2 The Autocorrelation of Increments of a Continuous-Time Process

In this section, we show that mild conditions guarantee that the limiting ACF of increments of a continuous-time process must be that of fractional Gaussian noise. Let $(X_t)_{t\geq 0}$ be a mean-square continuous stochastic process and $t > 0$ be a fixed time point.[3] Our main interest is when X is volatility, but X can be general for the purposes of this section.

2.1 A Characterization of the Limiting Autocorrelation Function

We define the local autocovariance function of the increments of X around t by[4]

$$\gamma_t(\tau;\Delta) = \mathrm{Cov}(X_{t+\Delta} - X_t, X_{t+(\tau+1)\Delta} - X_{t+\tau\Delta}), \quad \gamma_t(\Delta) = \gamma_t(0;\Delta), \quad \tau \geq 0. \tag{2.1}$$

In (2.1), $\Delta > 0$ denotes the step size of an increment. We are interested in $\Delta \to 0$ asymptotics for fixed t. Because $\gamma_t(\tau;\Delta)$ is allowed to depend on t, X can be nonstationary.

Assumption Γ. *We have $\gamma_t(\Delta) > 0$ for all $\Delta > 0$ and the following conditions:*

(i) There is a function $(\rho_t(\tau))_{\tau\geq 0}$ such that for all $\theta, \tau \geq 0$, we have

$$\left| \frac{\gamma_{t+\theta\Delta}(\tau;\Delta)}{\gamma_t(\Delta)} - \rho_t(\tau) \right| \longrightarrow 0 \quad \text{as } \Delta \to 0. \tag{2.2}$$

[3] It suffices to assume that X is mean-square continuous from the right at t. This would also accommodate situations in which X has fixed times of discontinuity (e.g., due to a scheduled announcement).

[4] The content of this section readily extends to *conditional* autocovariances/autocorrelation by considering the conditional probability measure $\mathbb{P}(\cdot \mid \mathcal{F}_t)$ and the process $(X_{t+s})_{s\geq 0}$.

(ii) We have the following regularity condition:

$$\liminf_{\lambda \downarrow 1} \liminf_{\Delta \to 0} \frac{\gamma_t(\lambda\Delta) - \gamma_t(\Delta)}{\gamma_t(\Delta)} \geq 0. \tag{2.3}$$

If we set $\theta = 0$ in (2.2), we see that $\rho_t(\tau)$ is precisely the infill asymptotic limit of the autocorrelation at lag τ of the increments of X. Assumption Γ (i) posits the existence of this limit. If X is a process with stationary increments, then condition (2.2) for $\theta = 0$ automatically implies the same condition for all $\theta > 0$. If X has nonstationary increments, condition (2.2) requires the dependence of $\gamma_t(\tau; \Delta)$ on t be of lower asymptotic order, so that we obtain the same limit $\rho_t(\tau)$ if time is shifted by a constant multiple of Δ. This condition is mild: for example, as we show in Propositions 1 and 2 below, if X is an Itô semimartingale or a moving-average process (including the price and volatility processes we consider in (3.1) and (3.13) below), then (2.2) is satisfied as soon as the coefficients of X are also mean-square continuous. Most continuous-time asset pricing models considered in prior literature satisfy this assumption. Exceptions include processes with explosive coefficients, examples of which we study in Proposition 3.

Condition (2.3) is a mild regularity assumption to rule out pathological cases. For example, if $\gamma_t(\Delta) = \mathrm{Var}(X_{t+\Delta} - X_t)$ is nondecreasing in Δ, then (2.3) is trivially satisfied. In general, for a continuous-time process, the limit in (2.3) is typically zero, which just means that the relative difference between $\mathrm{Var}(X_{t+\lambda\Delta} - X_t)$ and $\mathrm{Var}(X_{t+\Delta} - X_t)$ converges to zero as λ approaches unity.

Theorem 1. *If $(X_t)_{t\geq 0}$ is a mean-square continuous process satisfying Assumption Γ at time $t > 0$, then the following holds:*

(i) The variance function $\Delta \mapsto \gamma_t(\Delta)$ is regularly varying at zero with some index $\alpha \in [0, 2]$. That is, we have $\gamma_t(\lambda\Delta)/\gamma_t(\Delta) \to \lambda^\alpha$ as $\Delta \to 0$ for all $\lambda > 0$.

(ii) In (2.2), we must have $\rho_t(\tau) = \Psi_e(\tau)$ where $e = \frac{\alpha-1}{2}$ and, with the convention $0^0 = 0$,

$$\Psi_e(\tau) = \frac{1}{2}(|\tau + 1|^{2e+1} - 2|\tau|^{2e+1} + |\tau - 1|^{2e+1}), \quad \tau \geq 0. \tag{2.4}$$

In particular, if $\alpha \in (0, 2)$ (i.e., $e \in (-\frac{1}{2}, \frac{1}{2})$), $\rho_t(\tau)$ is the ACF of the increments of fractional Brownian motion with fractional parameter e.[5]

[5]Fractional Brownian motion is a centered Gaussian process $(B^e_t)_{t\geq 0}$ with covariance function $\mathbb{E}[B^e_t B^e_s] =$

Theorem 1 implies that for continuous-time processes satisfying Assumption Γ, the limiting ACF of its increments must be that of fractional Gaussian noise. In contrast to previous occurrences of fractional Brownian motion in the econometric literature (see e.g., Sowell (1990), Lo (1991), Comte and Renault (1996) and Müller and Watson (2008)), here it appears as a consequence of infill asymptotic theory, without assuming fractional differencing or fractional integration in the initial model.

We refer to $e = \frac{\alpha-1}{2}$ as the *smoothness parameter* of X at time t. If $e < 0$, non-overlapping increments of X are asymptotically negatively correlated and we say that X is *rough* at time t; if $e = 0$, they are asymptotically uncorrelated; and if $e > 0$, they are asymptotically positively correlated and we say that X is *smooth* at time t.[6]

2.2 Examples and Counterexamples

We first determine the smoothness parameter of continuous Itô semimartingales.

Proposition 1. *Let $X_t = Y_t + \int_0^t \sigma_s \, dW_s$ and $Y_t = Y_0 + \int_0^t b_s \, ds$ where W is a Brownian motion and b and σ are predictable and mean-square continuous. If $\mathbb{E}[\sigma_t^2] > 0$ and $\mathrm{Var}(b_t) > 0$ for all t, then X (Y) satisfies Assumption Γ with $e = \frac{1}{2}$ ($e = 0$) for all $t \geq 0$.*[7]

Thus, a pure drift has a maximal smoothness parameter of $e = \frac{1}{2}$, with asymptotically perfectly correlated increments ($\rho_t(\tau) = \Psi_{1/2}(\tau) = 1$). This is intuitive as such a process has differentiable sample paths, which by linear approximation behave locally like straight lines. By contrast, continuous Itô semimartingales with nonzero diffusion coefficients have a zero smoothness parameter, with asymptotically uncorrelated increments over disjoint intervals

$\frac{1}{2}(t^{2e+1} + s^{2e+1} - |t-s|^{2e+1})$ for $s, t \geq 0$. Its increments, $(B^e_{t+1} - B^e_t)_{t\geq 0}$, form *fractional Gaussian noise* and are stationary with ACF given by $\mathrm{Corr}(B^e_{t+1} - B^e_t, B^e_{t+\tau+1} - B^e_{t+\tau}) = \Psi_e(\tau)$ for all $t, \tau \geq 0$. The range of the fractional parameter is $e \in (-\frac{1}{2}, \frac{1}{2})$. Fractional Brownian motion and fractional Gaussian noise can also be parametrized by the Hurst index, which is given by $H = e + \frac{1}{2} = \frac{\alpha}{2}$.

[6] This terminology follows Gatheral et al. (2018) and Chong and Todorov (2025) and reflects the fact that processes with $e < 0$, $e = 0$ and $e > 0$ typically have paths of infinite, finite nonzero and zero quadratic variation, respectively. Moreover, if X is continuous, then the Hölder regularity of its sample paths is either lower or higher than that of Brownian motion paths depending on whether $e < 0$ or $e > 0$. We note that the boundary cases $e = -\frac{1}{2}$ and $e = \frac{1}{2}$ can also occur: in the first case, the increments of X have the same asymptotic autocorrelation as the first difference of white noise; in the second case, the increments of X are asymptotically perfectly correlated.

[7] While Theorem 1 allows for jumps in X, our main interest in the subsequent sections lies in the smoothness parameter of the *continuous* part of spot variance (after jumps have been removed). The smoothness parameter of an Itô semimartingale with square-integrable jumps and a nonzero diffusion part continues to be zero. This is because jumps are uncorrelated over disjoint intervals: they add variance, but not correlation, to the increments.

$(\rho_t(\tau) = \Psi_0(\tau) = (1-\tau)_+)$. This is because Itô semimartingales are locally dominated by their continuous martingale component, which has uncorrelated increments.

The next proposition shows that continuous-time moving average processes can have a smoothness parameter in the whole region $(-\frac{1}{2}, \frac{1}{2})$, with its value determined by the behavior of the moving-average kernel at the origin.

Proposition 2. *Suppose that $X_t = X_0 + \int_0^t g(t-s)\sigma_s \, dW_s$, where g has the form stated in (3.14) below (in particular, we have $g(t) \sim Ct^e$ as $t \to 0$ for some $e \in (-\frac{1}{2}, \frac{1}{2})$) and σ is predictable, mean-square continuous and satisfies $\mathbb{E}[\sigma_t^2] > 0$ for all t. Then X satisfies Assumption Γ and has smoothness parameter e for all $t > 0$.*

If $g(0)$ is a nonzero constant, X is in fact a semimartingale and we have $e = 0$, which is consistent with Proposition 1. If $g(0) = 0$, the process X has positively correlated increments and is smooth with $e > 0$. If $g(t) \to \infty$ as $t \to 0$, then X is asymptotically negatively correlated and is rough with $e < 0$. We emphasize that the value of e, in general, is unrelated to the memory properties of X. Section 2.3 discusses this relationship in greater detail.

We next show that Assumption Γ can fail if the coefficients of X are explosive (Laurent et al. (2024)). We define $r_t(\tau) = \lim_{\Delta \to 0} \operatorname{Corr}(X_{t+\Delta} - X_t, X_{t+(\tau+1)\Delta} - X_{t+\tau\Delta})$, if the limit exists. Note that we have $r_t(\tau) = \rho_t(\tau)$ automatically if (2.2) holds.

Proposition 3. *Let $X_t = X_0 + \int_0^t s^{-\alpha} b_s \, ds + \int_0^t s^{-\beta} \sigma_s \, dW_s$, where $\alpha \in [0, 1)$, $\beta \in [0, \frac{1}{2})$ and b, σ and W are as in Proposition 1 and $\operatorname{Var}(b_t) > 0$ and $\mathbb{E}[\sigma_t^2] > 0$ for all t.*[8]

(i) For any $t > 0$, Assumption Γ holds with $e = 0$ and $r_t(\tau) = (1-\tau)_+$.

(ii) If $\alpha < \beta + \frac{1}{2}$, Assumption Γ holds at $t = 0$, and we have $e = 0$ in this case, if and only if $\beta = 0$. Regardless of the value of $\beta \in [0, \frac{1}{2})$, we have $r_0(\tau) = 0$ for $\tau \geq 1$ and $r_0(\tau) = \frac{1 - \tau^{1-2\beta}}{\sqrt{(\tau+1)^{1-2\beta} - \tau^{1-2\beta}}}$ for $\tau \in [0, 1)$.

(iii) If $\alpha > \beta + \frac{1}{2}$, Assumption Γ fails at $t = 0$ and $r_0(\tau) = \Psi_e(\tau) = 1$ with $e = \frac{1}{2}$.

(iv) If $\alpha = \beta + \frac{1}{2}$, Assumption Γ fails at $t = 0$ and $r_0(\tau) = 1/\sqrt{q}$ for $\tau \geq 1$ and $r_0(\tau) = \frac{q[(\tau+1)^{1-\alpha} - \tau^{1-\alpha}] + (1-q)(1-\tau^{2-2\alpha})}{\sqrt{q[(\tau+1)^{1-\alpha} - \tau^{1-\alpha}]^2 + (1-q)[(\tau+1)^{2-2\alpha} - \tau^{2-2\alpha}]}}$ for $\tau \in [0, 1)$, where $q = \frac{\operatorname{Var}(b_0)}{\operatorname{Var}(b_0) + \frac{1-\alpha}{2}\mathbb{E}[\sigma_0^2]}$.

While X behaves like a continuous Itô semimartingale with bounded coefficients (and $e = 0$) for any $t > 0$, the smoothness properties of X at $t = 0$ depend on the strength of

[8] This is a variant of the drift and volatility burst model studied by Laurent et al. (2024). We consider a burst of the coefficients at time $t = 0$ for simplicity.

the drift and volatility burst. If $\alpha < \beta + \frac{1}{2}$, the drift burst is mild and negligible relative to the diffusion component. Therefore, X inherits the property of uncorrelated increments over disjoint intervals ($r_0(\tau) = 0$ for $\tau \geq 1$), which is not affected by the value of β. The value of β, however, does affect the correlation of overlapping increments ($\tau < 1$) and we have Assumption Γ with $e = 0$ and $r_0(\tau) = (1-\tau)_+$ precisely when there is no volatility burst ($\beta = 0$). If $\alpha > \beta + \frac{1}{2}$, the drift burst dominates the diffusion component around $t = 0$. As a consequence, $r_0(\tau) = 1$ as in the case of a pure drift studied in Proposition 1. Nevertheless, Assumption Γ fails because the limit $\rho_t(\tau)$ in (2.2) now depends on θ as the proof of the proposition shows. Finally, if $\alpha = \beta + \frac{1}{2}$, both drift and diffusion affect the limiting ACF $r_0(\tau)$, with the magnitude of their contributions depending on the ratio $q \in (0, 1)$.

Propositions 1 and 3 show that the asymptotic autocorrelation of nonoverlapping increments in continuous Itô semimartingale models, with or without explosive coefficients, is always nonnegative. It is equal to zero if the diffusion term dominates; it is one if the drift term dominates; and it takes a value in $(0, 1)$ in balanced cases (as in Proposition 2 (iv)). Explosive coefficients thus cannot generate negative autocorrelation in increments.

We next show that an infinite sum of Itô semimartingales with increasing mean-reversion speeds can lead to negative autocorrelation in its increments and a negative smoothness parameter. Multifactor volatility models (see e.g., Chernov et al. (2003)) involve a similar construction with finitely many factors. Suppose that

$$Z_t = X_t + \sum_{i=1}^{\infty} \rho_i X_t^i, \quad dX_t^i = -\lambda_i X_t^i \, dt + \sigma_t^i \, dW_t^i, \quad i \geq 1, \tag{2.5}$$

where X is the same process as in Proposition 1, $\rho_i, \lambda_i > 0$, $(X_0^i)_{i \geq 1}$ are independent and centered initial conditions, $(W_t^i)_{t \geq 0}$ are independent standard Brownian motions which are further independent of $(X_0^i)_{i \geq 1}$, and $(\sigma_t^i)_{t \geq 0}$ are predictable and mean-square continuous.

The process Z therefore has a main semimartingale component, X, as well as components X^i of Ornstein–Uhlenbeck type with increasing mean-reversion parameters. The smoothness parameter of the resulting process Z depends on the asymptotic behavior of the weights ρ_i and the mean-reversion parameters λ_i.

Proposition 4. *Suppose that* $2\mathbb{E}[(X_0^i)^2]\lambda_i \to v_0$ *and* $\mathbb{E}[(\sigma_t^i)^2] \to v_t$ *locally uniformly in* t *as* $i \to \infty$, *where* v *is a continuous positive function. If* $\rho_i \sim Ri^{-r}$ *and* $\lambda_i \sim Li^q$ *as* $i \to \infty$ *for some* $0 < r < \frac{1}{2}$, $q > 1 - 2r$ *and* $R, L > 0$, *then* Z *is well defined and satisfies Assumption* Γ *with smoothness parameter* $e = -\frac{1-2r}{2q} < 0$.

While the individual factor processes X^i, for $i \geq 1$, are continuous Itô semimartingales with $e = 0$ (see Proposition 1), their infinite linear combination can have a negative smoothness parameter, if the weights ρ_i decay slowly enough and the mean-reversion parameters λ_i diverge sufficiently fast. The exact value of e depends on the asymptotic rates of ρ_i and λ_i. The assumption in the proposition that $2\mathbb{E}[(X_0^i)^2]\lambda_i \to v_0$ reflects the fact that if σ^i is jointly covariance-stationary with dW^i, then the stationary solution X_t^i satisfies $\mathbb{E}[(X_t^i)^2] = \mathbb{E}[(\sigma_0^i)^2]/(2\lambda_i)$. Because we allow $\mathbb{E}[(\sigma_t^i)^2] \to v_t$ for a time-varying function v, the processes X_t^i can be nonstationary.

2.3 Smoothness and Memory

For a covariance-stationary process X, its smoothness parameter e is related to the decay of its spectral density at infinity.

Theorem 2. *Let $(X_t)_{t\in\mathbb{R}}$ be a continuous-time covariance-stationary process satisfying Assumption Γ with autocovariance function $C(r) = \mathrm{Cov}(X_0, X_r)$, $r \in \mathbb{R}$.*

(i) Let F be its spectral distribution function, that is, the unique distribution function with total mass $C(0)$ such that $C(r) = \int_{-\infty}^{\infty} \cos(\lambda r)\, dF(\lambda)$ for $r \in \mathbb{R}$. Then X has smoothness parameter $e \in (-\frac{1}{2}, \frac{1}{2})$ if and only if $C(0) - F$ is regularly varying at infinity with index $-1 - 2e$.

(ii) Consider the special case where $C(r)$ is integrable with spectral density f, that is, $f(\lambda) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \cos(\lambda r) C(r)\, dr$ for $\lambda \in \mathbb{R}$. Suppose that f is nonincreasing on (a, ∞) for large enough $a > 0$. Then X has smoothness parameter $e \in (-\frac{1}{2}, \frac{1}{2})$ if and only if f is regularly varying at infinity with index $-2 - 2e$.

The theorem highlights that the smoothness parameter e is distinct from the *memory parameter* d of X. A covariance-stationary process X has memory parameter $d \in (-\frac{1}{2}, \frac{1}{2})$ if its spectral density satisfies $f(\lambda) \asymp |\lambda|^{-2d}$ as $\lambda \to 0$ (Dittmann and Granger (2002)). If $d > 0$, X has *long memory*; if $d = 0$, X has *short memory*; and if $d < 0$, X is *antipersistent*. While d describes the behavior of the spectral density at the origin, e concerns its asymptotic decay at infinity.

Theorems 1 and 2 demonstrate that infill asymptotics separate high-frequency variation in X, encoded by e, from low-frequency variation in X, encoded by d. In our application to volatility, they allow us to conduct inference about the smoothness parameter of intraday

volatility independently of the well-established long memory property of volatility over daily or even longer time horizons (see e.g., Andersen et al. (2001, 2003)).

In many parametric fractional models, however, a single parameter controls the behavior of the spectral density at zero and infinity. As Müller and Watson (2008) point out, in such a case, even efficient statistical procedures may rely on both high- and low-frequency information, raising robustness concerns if the parametric model is misspecified. This is consistent with Shi and Yu (2023) and Li et al. (2025b) who find that roughness and long memory are only weakly identified in ARFIMA models.

The example of a fractional Ornstein–Uhlenbeck process driven by a fractional Brownian motion with Hurst index $H \in (0,1)$ illustrates the previous point. Its spectral density satisfies $f(\lambda) \asymp |\lambda|^{1-2H}$ as $\lambda \to 0$ and $f(\lambda) \asymp |\lambda|^{-1-2H}$ as $\lambda \to \infty$ (see e.g., Shi et al. (2024)). Hence, $e = d = H - \frac{1}{2}$ and the process is either smooth and has long memory (if $H > \frac{1}{2}$), rough and antipersistent (if $H < \frac{1}{2}$), or has the same smoothness as Brownian motion and short memory (if $H = \frac{1}{2}$). Importantly, a single parameter, H, controls both the smoothness and memory properties of the process.

This need not be the case for a general stationary process. For example, the sum of two independent fractional Ornstein–Uhlenbeck processes with Hurst indices H_1 and H_2, respectively, where $H_1 < H_2$, has a spectral density that satisfies $f(\lambda) \asymp |\lambda|^{1-2H_2}$ as $\lambda \to 0$ and $f(\lambda) \asymp |\lambda|^{-1-2H_1}$ as $\lambda \to \infty$. Hence, by Theorem 2, we have $e = H_1 - \frac{1}{2}$ and $d = H_2 - \frac{1}{2}$. If $H_1 < \frac{1}{2} < H_2$, the process will be rough *and* have long memory.

2.4 Localization

The next theorem shows that the definition of the smoothness parameter e can be extended to processes without a finite second moment by localization.

Theorem 3. *Suppose that $(X_t)_{t\geq 0}$ is a continuous-time process and there is a sequence of random times $(\tau_n)_{n\in\mathbb{N}}$ satisfying $\tau_n \to \infty$ almost surely such that the processes $X_t^{\tau_n} = X_{t\wedge\tau_n}$ have finite second moments, are mean-square continuous and satisfy Assumption Γ for all $n \in \mathbb{N}$, with limiting ACFs $\rho_t^{(n)}(\tau)$ that converge as $n \to \infty$ for $\tau = 1$. Then there exists $e \in [-\frac{1}{2}, \frac{1}{2}]$ such that*

$$\lim_{n\to\infty} \lim_{\Delta\to 0} \operatorname{Corr}(X_{t+\Delta}^{\tau_n} - X_t^{\tau_n}, X_{t+(\tau+1)\Delta}^{\tau_n} - X_{t+\tau\Delta}^{\tau_n}) = \Psi_e(\tau), \quad \tau \geq 0. \tag{2.6}$$

Examples covered by Theorem 3 include moving-average processes and, in particular, continuous Itô semimartingales (by taking $e = 0$ and $g \equiv 1$), as the next proposition shows.

Proposition 5. *Consider* $X_t = \int_0^t g(t-s)\eta_s\, dW_s$, *where* g *is a kernel as in* (3.14) *below for some* $e \in (-\frac{1}{2}, \frac{1}{2})$ *and* η *is a predictable process which is continuous in probability and locally bounded and locally bounded away from zero. If we define* $\tau_n = \inf\{t > 0 : |\eta_t| > n \text{ or } |\eta_t| < \frac{1}{n}\}$, *then* X^{τ_n} *is mean-square continuous for all* $n \in \mathbb{N}$ *and we have* (2.6).

3 Limit Theorems for Realized Autocovariances of Volatility Estimators

We move to inference about the smoothness parameter of volatility. The log-price x of the underlying asset is assumed to follow standard Itô semimartingale dynamics on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$:

$$x_t = x_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s + \int_0^t \int_E \delta^l(s,z)\, \pi(ds,dz) + \int_0^t \int_E \delta^s(s,z)\, \widetilde{\pi}(ds,dz). \quad (3.1)$$

In this expression, x_0 is an $\mathcal{F}_0$-measurable random variable, W is a standard Brownian motion, π is a Poisson random measure on $[0,\infty) \times E$ with compensator $dt\, \lambda(dz)$, $\widetilde{\pi}(dt, dz) = \pi(dt, dz) - dt\, \lambda(dz)$ and λ is a σ-finite measure on an auxiliary space E. The Itô semimartingale specification of the asset price is standard in continuous-time econometrics (Aït-Sahalia and Jacod (2014)). We specify further assumptions on the drift b, the volatility σ and the jump functions δ^l and δ^s later on.

3.1 Realized Autocovariances of Volatility Estimators

Our goal is to construct an estimator of the autocorrelation of volatility increments in the infill asymptotic limit. The estimator is based on high-frequency observations of asset returns on $[0, T]$, with step size Δ_n, where $T > 0$ is fixed and $\Delta_n \to 0$. Our inference procedure starts from jump-robust spot variance estimates constructed from truncated sums of squared returns over local blocks of return data (Aït-Sahalia and Jacod (2014)):

$$\widehat{c}_i^n = \frac{1}{k_n \Delta_n} \sum_{j=i+1}^{i+k_n} (\Delta_j^n x)^2 \mathbf{1}_{|\Delta_j^n x| \leq v_n}, \quad i = 0, \ldots, \lfloor T/\Delta_n \rfloor - k_n. \quad (3.2)$$

Here, $\Delta_i^n x = x_{i\Delta_n} - x_{(i-1)\Delta_n}$ is a high-frequency return on the asset, $(k_n)_{n\in\mathbb{N}}$ is a positive integer sequence increasing to infinity and $(v_n)_{n\in\mathbb{N}}$ is a sequence of positive truncation levels

shrinking to zero. Next, we compute truncated increments of these spot variance estimators and subsequently the *realized autocovariances* thereof. That is, we consider[9]

$$\Delta_{p_n}\widehat{c}_i^n = \widehat{c}_{i+p_n}^n - \widehat{c}_i^n, \quad \Delta_{p_n}^{\mathrm{tr}}\widehat{c}_i^n = \Delta_{p_n}\widehat{c}_i^n \mathbf{1}_{|\Delta_{p_n}\widehat{c}_i^n| \leq w_n} \tag{3.3}$$

and, for $\ell \geq 0$,

$$\widehat{V}_t^{n,\ell} = \frac{1}{p_n} \sum_{i=0}^{\lfloor t/\Delta_n \rfloor - (\ell+1)p_n - k_n} \Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{i+\ell p_n}^n \Delta_{p_n}^{\mathrm{tr}}\widehat{c}_i^n. \tag{3.4}$$

In (3.3) and (3.4), $(p_n)_{n\in\mathbb{N}}$ is an integer sequence increasing to infinity such that $p_n \geq k_n$. Choosing $p_n > k_n$ introduces gaps between $\widehat{c}_i^n$ and $\widehat{c}_{i+p_n}^n$, which helps mitigate the impact of potential microstructure noise on (3.4) in the case where $\ell = 1$. The sequence $(w_n)_{n\in\mathbb{N}}$, positive and convergent to zero, serves as truncation levels to remove volatility jumps from $\Delta_{p_n}\widehat{c}_i^n$. In one of the settings considered in the empirical application below, Δ_n is five seconds, $k_n\Delta_n$ is eight minutes and $p_n\Delta_n$ is ten minutes.

3.2 Microstructure Noise

If returns are sampled at high frequency, we may only observe

$$\Delta_i^n y = \Delta_i^n x + \varepsilon_i^n, \quad i = 1, \ldots, \lfloor T/\Delta_n \rfloor, \tag{3.5}$$

where ε_i^n are error variables due to market microstructure noise. It has long been recognized that observation errors in high-frequency asset prices can lead to biases in the spot variance estimates (3.2), which has prompted a large literature on noise-robust estimation of volatility.[10] As Bandi and Russell (2008) point out, there is an implicit bias–variance trade-off in any inferential procedure aimed at recovering volatility from noisy price observations. For example, under standard conditions, the most efficient noise-robust estimators of integrated volatility attain a rate of convergence of $\Delta_n^{-1/4}$, compared to $\Delta_n^{-1/2}$ in the noise-free case. For coefficients or parameters of volatility (e.g., volatility of volatility or the leverage effect), the rate drops from $\Delta_n^{-1/4}$ to $\Delta_n^{-1/8}$ for classical noise-robust estimators. Robustness to noise thus comes at the price of halving the rate exponent.

[9] Barndorff-Nielsen and Shephard (2004a), Barndorff-Nielsen et al. (2008), Laurent et al. (2024) and Li and Yang (2026) investigate realized (auto-)covariances for asset prices. Our statistics differ from these works in that we consider realized autocovariances of spot variance estimators.

[10] See e.g., Zhang et al. (2005), Bandi and Russell (2006), Barndorff-Nielsen et al. (2008), Jacod et al. (2009), Jacod et al. (2017), Da and Xiu (2021) and Li and Linton (2022).

For the smoothness parameter of volatility, the trade-off we face is different. If the observed returns follow (3.5), where ϵ_i^n is centered and stationary, then $\widehat{c}_i^n \sim \sigma^2_{(i-1)\Delta_n} + \rho_n^2/\Delta_n$ where $\rho_n^2 = \mathrm{Var}(\epsilon_i^n)$ is the noise variance that induces the asymptotic bias. The realized autocovariance statistics in (3.4), however, hinge on *increments* of volatility estimators, for which these biases cancel out such that $\widehat{c}_i^n$ remains centered at $\Delta_i^n \sigma^2$. Thus, to first order, noise does *not* bias the realized autocovariance statistics in (3.4).

Our approach in the following is in the same spirit as Andrews et al. (2026) and focuses on identifying assumptions under which noise contamination does *not* affect the large-sample properties of our estimators. This allows us to construct feasible estimators of the limiting autocorrelation of volatility increments and of the smoothness parameter of volatility with a rate of convergence of $\Delta_n^{-1/4}/\sqrt{\log \Delta_n^{-1}}$ in Section 4 and to understand the boundary conditions under which such a rate is attainable despite the presence of noise.

In the following, we assume that $\widehat{c}_i^n$ is constructed from the observed returns, $\Delta_i^n y$, instead of the efficient returns, $\Delta_i^n x$. We further assume that

$$\varepsilon_i^n = \Delta_n^{\alpha - 1/2} \Big(\int_{(i-1)\Delta_n}^{i\Delta_n} \rho_s^2 \, ds \Big)^{1/2} \chi_i, \tag{3.6}$$

where $(\chi_i)_{i \in \mathbb{N}}$ is a discrete stationary time series with zero mean and unit variance, ρ is an adapted noise intensity process and $\alpha \geq 0$ determines the strength of the noise.[11] If $\alpha = \frac{1}{2}$, the noise variables and the efficient returns have the same asymptotic order and both price variance, σ_t^2, and noise variance, ρ_t^2, contribute to the asymptotic limit of (3.4). If $\alpha < \frac{1}{2}$, microstructure noise dominates, while if $\alpha > \frac{1}{2}$, it is asymptotically smaller than the efficient returns. The case $\alpha = 0$ ("nonshrinking noise") is considered in Jacod et al. (2017) and Li and Linton (2022), while Aït-Sahalia and Jacod (2014) and Da and Xiu (2021) also consider the case $\alpha > 0$ ("shrinking noise"). Likewise, our theory allows for all $\alpha \geq 0$.[12]

We also assume that the noise variance process ρ_t^2 has the same smoothness parameter

[11] Note that the noise variables directly act on the returns. This specification implies that $\widehat{c}_i^n$ is an estimator of $(k_n \Delta_n)^{-1} \int_{i\Delta_n}^{(i+k_n)\Delta_n} (\sigma_s^2 + \Delta_n^{2\alpha-1} \rho_s^2) \, ds$, which allows for a unified treatment of price and noise volatility. If ρ is constant, (3.6) is equivalent to the more common specification $y_{i\Delta_n} = x_{i\Delta_n} + \Delta_n^\alpha \rho_{i\Delta_n} \chi_i$, which introduces errors on price levels. If ρ is time-varying, however, the latter implies that $\widehat{c}_i^n$ is an estimator of $(k_n \Delta_n)^{-1} \int_{i\Delta_n}^{(i+k_n)\Delta_n} (\sigma_s^2 + \Delta_n^{2\alpha-1} (\rho^2_{\lfloor s/\Delta_n \rfloor \Delta_n} + \rho^2_{\lfloor s/\Delta_n \rfloor \Delta_n + \Delta_n})) \, ds$. While the results remain unchanged, the proofs will be substantially longer without offering any additional insight, which is why we assume (3.6).

[12] While we do not need to know or estimate α, microstructure noise appears to be small in our empirical application in Section 6. The median ratio of five-second realized variance to five-minute realized variance is 1.0869 in our sample. This is in line with Aït-Sahalia and Xiu (2019) and Li and Yang (2026) who provide empirical evidence that the magnitude of microstructure noise has substantially decreased in recent years.

as σ_t^2. This assumption is satisfied if the noise intensity is a function of volatility, which is supported by empirical evidence of a strong relationship between noise intensity and volatility (Aït-Sahalia and Yu (2009)). Li et al. (2016) and Bollerslev et al. (2018) provide further evidence of this link by documenting a strong relationship between microstructure noise and trading volume and between trading volume and volatility, respectively. Remark 1 below discusses how our results generalize to cases in which volatility and noise intensity have different smoothness parameters.

3.3 Law of Large Numbers for Realized Autocovariances of Volatility Estimators

Because observed asset returns are contaminated by microstructure noise of the form (3.6), the spot variance estimators $\widehat{c}_i^n$ in (3.2), when multiplied by $\Delta_n^{(1-2\alpha)_+}$, approximate

$$c_t = \Delta_n^{(1-2\alpha)_+}\sigma_t^2 + \Delta_n^{(2\alpha-1)_+}\rho_t^2, \tag{3.7}$$

where $x_+ = \max(x, 0)$ denotes the positive part of x. We assume that

$$c_t = v_t + \int_0^t \int_E \delta_c^l(s,z)\,\pi(ds,dz) + \int_0^t \int_E \delta_c^s(s,z)\,\widetilde{\pi}(ds,dz), \tag{3.8}$$

where v, the continuous part of c, is an adapted process and δ_c^l and δ_c^s are the large and small jump parts of c, respectively. If $\alpha \neq \frac{1}{2}$, the process c as well as its components depend on n, which we do not reflect in the notation for simplicity. We prove an LLN for the realized autocovariances defined in (3.4) under Assumption LLN stated in Appendix B. The main structural assumptions on the continuous part of c are mild and comprise the following technical conditions. If v is a moving-average process as in Proposition 5, then (3.9)–(3.11) below are satisfied with $\gamma(\Delta_n) = \Delta_n^{2e+1}$ and $\zeta = \eta^2$.

Assumption V1. *The process v is adapted and satisfies the conditions of Theorem 3 for all $t > 0$, with a smoothness parameter $e \in (-\frac{1}{2}, 0]$ characterized by (2.6) and a sequence of stopping times $(\tau_m)_{m\in\mathbb{N}}$ increasing to infinity almost surely that does not depend on t and has the following properties: there exist a function γ regularly varying at zero with index $2e+1$ and a predictable positive process $(\zeta_t)_{t\geq 0}$ which is locally bounded and locally bounded away*

from zero such that for $m, p \in \mathbb{N}$ *and* $\iota > 0$*, we have*

$$\mathbb{E}\big[\big(\mathbb{E}[v^{\tau_m}_{s+\Delta_n} - v^{\tau_m}_s \mid \mathcal{F}_{s-r}]\big)^2\big] = o(\gamma(\Delta_n)), \tag{3.9}$$

$$\mathbb{E}\big[\big|\mathrm{Cov}\big(v^{\tau_m}_{s+\Delta_n} - v^{\tau_m}_s, v^{\tau_m}_{t+\Delta_n} - v^{\tau_m}_t \mid \mathcal{F}_{s-r}\big) - \zeta^{\tau_m}_{s-r}\gamma(\Delta_n)\Psi_e(\tfrac{t-s}{\Delta_n})\big|\big] = o(\gamma(\Delta_n)), \tag{3.10}$$

$$\mathbb{E}[|v^{\tau_m}_{s+\Delta_n} - v^{\tau_m}_s|^p]^{1/p} = O(\gamma(\Delta_n)^{1/2}), \quad \mathbb{E}[|\zeta^{\tau_m}_{s+\Delta_n} - \zeta^{\tau_m}_s|] = o(1) \tag{3.11}$$

as $\Delta_n \to 0$*, uniformly in* $0 \leq r \leq s \leq t \leq T$ *satisfying* $r \geq \Delta_n^{1-\iota}$*.*

Assumption V1 implies that v, the continuous part of c, has a smoothness parameter that does not depend on time. If e is time-varying, the periods in which e is the smallest (and thus v is the roughest) will dominate the asymptotic behavior of (3.4), see Remark 2 below. Our empirical results suggest that the time-series variation of e is mild in practice (see Figure 1).

We only consider $e \in (-\frac{1}{2}, 0]$ in Assumption V1 to simplify the exposition. Theorem 4 below can be extended to the whole range $e \in (-\frac{1}{2}, \frac{1}{2})$ if one replaces $k_n \asymp \Delta_n^{-1/2} \log \Delta_n^{-1}$ (see Assumption LLN (v)) by a larger window of size $k_n \asymp \Delta_n^{-2/3}$ and imposes stronger conditions on the price and volatility jumps in Assumption LLN (iii) and (iv). As a larger window size leads to a loss of efficiency of the estimators we consider in Section 4 if the true value of e is negative, and Chong and Todorov (2025) reject the case $e \geq 0$ in favor of $e < 0$ for almost all considered assets and time periods, we do not detail such an extension.

The conditions (3.9)–(3.11) are mild. The first two concern the conditional expectation and autocovariances of volatility changes over short intervals of length Δ_n. Condition (3.9) assumes that predicted changes of volatility over periods of length Δ_n become smoother as the forecast horizons (when measured in units of Δ_n) increases. If v^{τ_m} has stationary increments, this is satisfied if the ACF of its increments decays to zero at infinity, however slowly. Equation (3.10) is a conditional version of (2.6) and states that the conditional autocovariances of short-term volatility changes are asymptotically equivalent to the those of fractional Gaussian noise scaled by a volatility-of-volatility process ζ. If $\alpha > \frac{1}{2}$ ($\alpha < \frac{1}{2}$) in (3.7), ζ only depends on σ^2 (ρ^2); if $\alpha = \frac{1}{2}$, it depends on both. The next theorem determines the large-sample limit of the realized autocovariances $V^{n,\ell}_t$. We use $\overset{L^1}{\Longrightarrow}$ to denote uniform convergence in L^1 on $[0, T]$ for any fixed $T > 0$. The constants $\Phi_{\kappa,e}(\ell)$ are defined in (B.2) of Appendix B.

Theorem 4. *Under Assumption LLN, we have* $V^{n,\ell} \xrightarrow{L^1} V^\ell$ *for all* $\ell \geq 0$*, where*

$$V_t^{n,\ell} = \frac{p_n \Delta_n}{\gamma(p_n \Delta_n)} \Delta_n^{2(1-2\alpha)_+} \widehat{V}_t^{n,\ell}, \qquad V_t^\ell = \Phi_{\kappa,e}(\ell) \int_0^t \zeta_s \, ds \tag{3.12}$$

The large-sample limit of $V_t^{n,\ell}$ is thus given by the product of a limiting ACF $(\Phi_{\kappa,e}(\ell))_{\ell \geq 0}$ and integrated volatility of volatility, $\int_0^t \zeta_s \, ds$. We note that $\Phi_{\kappa,e}$ is *not* the ACF of fractional Gaussian noise defined in (2.4) but the ACF of averages of fractional Gaussian noise (see (B.2) in Appendix B). This is because $\widehat{c}_i^n$ is primarily an estimator of local averages of spot variance.

Remark 1. *What if the noise variance process,* ρ_t^2*, has a smoothness parameter* $e' \leq 0$ *(together with a variance function* γ'*) which is different from the smoothness parameter* e *of spot variance,* σ_t^2*? In this case, Theorem 4 continues to hold with* $c = \sigma^2$ *(*$c = \rho^2$*) and* $V_t^{n,\ell} = \frac{p_n \Delta_n}{\gamma(p_n \Delta_n)} \widehat{V}_t^{n,\ell}$ *(*$V_t^{n,\ell} = \frac{p_n \Delta_n}{\gamma'(p_n \Delta_n)} \Delta_n^{1-2\alpha} \widehat{V}_t^{n,\ell}$*) if* $\alpha > \frac{1}{2} + \frac{e-e'}{4}$ *(*$\alpha < \frac{1}{2} + \frac{e-e'}{4}$*). Therefore, microstructure noise does not affect the estimation of* e *unless it is strong* and *its intensity has a smoothness parameter which is largely different from that of volatility.*[13]

Remark 2. *What if the smoothness parameter of* c *is time-varying? In this case, the realized autocovariance statistics will be dominated by the periods in which* e *is the smallest. Consider, for example, the case where the smoothness parameter is a piecewise constant function of time:* $e(t) = \sum_{j=1}^k e_j \mathbf{1}_{I_j}(t)$ *where* $e_j < 0$ *and* $(I_j)_{j=1,\ldots,k}$ *forms a partition of* $[0, T]$ *into intervals of positive length. In that case, we still have Theorem 4 with* $e = e_1 \wedge \cdots \wedge e_k$ *and* $V_t^\ell = \Phi_{\kappa,e}(\ell) \sum_{j:e_j=e} \int_{I_j} \zeta_s \mathbf{1}_{s \leq t} \, ds$*.*

3.4 Central Limit Theorem for Realized Autocovariances of Volatility Estimators

In order to obtain a CLT for the realized autocovariances of spot variance estimators, we need stronger structural assumptions on c. We present the main assumption here and gather additional regularity conditions in Assumption CLT of Appendix B.

[13] One can use classical pre-averaging techniques (see e.g., Jacod et al. (2017), Li and Linton (2022)) to construct an estimator of the volatility smoothness parameter which is robust to strong microstructure noise with highly irregular intensity. Such an estimator, however, has a rate of convergence of $\Delta_n^{-1/8}$ and therefore requires observations at much higher frequencies in practice than the five-second data we consider in Section 6. Given empirical evidence that the magnitude of microstructure noise has decreased in recent years (see e.g., Aït-Sahalia and Xiu (2019) and Li and Yang (2026)) and the strong relationship between volatility and noise (Aït-Sahalia and Yu (2009)), we do not pursue such an extension here.

Assumption V2. *We have (3.8) where v has the form*

$$v_t = v_0 + \int_0^t a_s\,ds + \int_0^t h(t-s)\theta_s\,ds + \int_0^t g(t-s)(\eta_s\,dW_s + \overline{\eta}_s\,d\overline{W}_s) \tag{3.13}$$

for an $\mathcal{F}_0$-measurable initial condition v_0. The processes a, θ, η and $\overline{\eta}$ are predictable and locally bounded, and $\overline{W}$ is a standard Brownian motion independent of W. Moreover,

$$g(t) = g_e(t) + g_0(t), \quad h(t) = g_{e'}(t) + h_0(t), \tag{3.14}$$

where $g_e(t) = K_e^{-1}t_+^e$, $K_e = \Gamma(1+e)/\sqrt{\cos(\pi e)\Gamma(2+2e)}$, $t_+^e = t^e\mathbf{1}_{t>0}$, $e, e' \in (-\frac{1}{2}, 0]$ and $g_0, h_0 \in \mathcal{A}$. Here $\mathcal{A}$ is the class of all absolutely continuous functions $f : [0,\infty) \to \mathbb{R}$ which vanish at the origin, are twice continuously differentiable on $(0,\infty)$ and satisfy $|f'(t)| \leq Ct^{\alpha-1}$ and $|f''(t)| \leq Ct^{\alpha-2}$ for $0 < t < 1$ and some $\alpha > \frac{1}{2}$.

Under Assumption V2, the continuous part of c_t is a continuous-time moving average process driven by Brownian motion. By the Wold representation theorem (see e.g., Theorem 2' in Section 4 of Hannan (1970)), stationary processes also have them as their main building block. Assumption V2 is a direct generalization of the semimartingale assumption on v_t underlying many CLTs in high-frequency econometrics (see e.g., Assumption (K) on page 116 in Jacod and Protter (2012)). The semimartingale case corresponds to $e = e' = 0$.

Because $g(t) \asymp t^e$ as $t \downarrow 0$, the smoothness parameter of the continuous part of c is e (see Proposition 5). The constant K_e is chosen such that $\mathbb{E}[(v_{t+\Delta} - v_t)^2] \sim \Delta^{2e+1}$ if $\eta_t^2 + \overline{\eta}_t^2 \equiv 1$. The behavior of g as $t \to \infty$ depends on g_0 and is not further specified. In particular, if v is stationary, the smoothness parameter e has no implications for the memory properties of v as these depend on the decay of the autocovariance function at infinity, which in turn is governed by the behavior of g at infinity. Thus, the setting in (3.13) and (3.14) allows for inference about the smoothness parameter e without making any assumptions on or drawing any conclusions about the memory parameter d.[14]

The process v has a differentiable drift, $\int_0^t a_s\,ds$, and a nondifferentiable drift, $\int_0^t h(t-s)\theta_s\,ds$. We include the latter in (3.13) as it is present, for example, in the rough Heston

[14] Comte and Renault (1996) study a variant of (3.13) in which the behavior of g at zero is coupled with the decay of g at infinity (see their Proposition 5 and condition (8)). Hence, the value d in Comte and Renault (1996) determines both the smoothness and the memory properties of the resulting process. By contrast, Assumption V2 specifies the behavior of g at the origin but leaves its decay at infinity unspecified. The memory properties of v in (3.13) are thus unrelated to e. For example, the kernel $g(t) = t_+^e(1+t)^{d-1-e}$ with $d, e \in (-\frac{1}{2}, \frac{1}{2})$ implies a smoothness parameter of e and a memory parameter of d (see Bennedsen et al. (2022) for further examples).

model (El Euch and Rosenbaum (2019)) considered in the Monte Carlo study of Section 5.

The next result determines the joint asymptotic distribution of $\widehat{V}_t^{n,\ell}$, $\ell = 2, \dots, L$ for some $L \geq 2$, together with $\widehat{V}_t^{n,0} + 2\widehat{V}_t^{n,1}$. We consider this particular linear combination because sampling errors in $\widehat{c}_i^n$ lead to bias terms in $\widehat{V}_t^{n,0}$ and $\widehat{V}_t^{n,1}$. While they converge to zero in probability and therefore do not affect the LLN limit of $\widehat{V}_t^{n,0}$ and $\widehat{V}_t^{n,1}$, they affect the CLT as higher-order bias terms. However, as we show, these bias terms cancel out up to an asymptotically negligible contribution when linearly combined in the above fashion. This in turn is because the sampling errors in $\widehat{c}_i^n$ and $\widehat{c}_{i+p_n}^n$ are uncorrelated in the limit.

Theorem 5. *Let $V_t^{n,\ell}$ and V_t^ℓ be given by (3.12) with $\gamma(\Delta) = \Delta^{2e+1}$ and $\zeta_t = \eta_t^2 + \overline{\eta}_t^2$. For fixed $m \geq 1$ and $\ell = 1, \dots, m$, define $\mathcal{Z}_t^{n,\ell} = (p_n\Delta_n)^{-1/2}(V_t^{n,\ell} - V_t^\ell)$ if $\ell \geq 2$ and $\mathcal{Z}_t^{n,1} = (p_n\Delta_n)^{-1/2}(V_t^{n,0} + 2V_t^{n,1} - V_t^0 - 2V_t^1)$ if $\ell = 1$. Under Assumption CLT, the process $\mathcal{Z}_t^n = (\mathcal{Z}_t^{n,1}, \dots, \mathcal{Z}_t^{n,m})^T$ satisfies the joint CLT[15] $\mathcal{Z}^n \overset{\mathcal{L}-s}{\Longrightarrow} \mathcal{Z}$, where $\mathcal{Z} = ((\mathcal{Z}_t^1, \dots, \mathcal{Z}_t^m)^T)_{t\geq 0}$ is defined on $(\overline{\Omega}, \overline{\mathcal{F}}, (\overline{\mathcal{F}}_t)_{t\geq 0}, \overline{\mathbb{P}})$ and is $\mathcal{F}$-conditionally Gaussian with independent increments, mean zero and covariance matrix*

$$\mathcal{C}_t^{\ell\ell'} = \overline{\mathbb{E}}[\mathcal{Z}_t^\ell \mathcal{Z}_t^{\ell'} \mid \mathcal{F}] = \mathcal{R}_{\kappa,e}(\ell, \ell') \int_0^t (\eta_s^2 + \overline{\eta}_s^2)^2 \, ds. \tag{3.15}$$

The constants $\mathcal{R}_{\kappa,e}(\ell, \ell')$ are defined in (B.3) of Appendix B.

The asymptotic distribution of the realized autocovariances of spot variance estimates is mixed normal, with the asymptotic conditional covariance matrix given by (3.15). As (B.3) shows, $\mathcal{R}_{\kappa,e}(\ell, \ell')$ is related to the autocovariance of products of integrated fractional Gaussian noise. Naturally, the integrated quarticity of spot variance, $\int_0^t (\eta_s^2 + \overline{\eta}_s^2)^2 \, ds$, enters (3.15) as well. Just like the LLN (see the discussion after Assumption V1), Theorem 5 can be extended to $e \in (0, \frac{1}{4})$ with minor modifications of the proofs if one imposes more restrictive assumptions on the jumps and chooses a window size which is asymptotically of the order $\Delta_n^{-3/5}$. If one knows that $e \leq 0$ (see e.g., Chong and Todorov (2025) for the case of volatility), the resulting estimator would have a slower rate of convergence than the current one, which is based on $k_n \asymp \Delta_n^{-1/2} \log \Delta_n^{-1}$. Theorem 5, however, does not extend to $e \in (\frac{1}{4}, \frac{1}{2})$. In this case, the CLT can be recovered by considering second- instead of first-order increments of volatility (cf. Corcuera et al. (2013)). Because the confidence intervals we construct in the empirical application are strictly contained in $(-\frac{1}{2}, 0)$, we do not detail such extensions here.

[15] Here $\overset{\mathcal{L}-s}{\Longrightarrow}$ denotes functional stable convergence in the local uniform topology and $(\overline{\Omega}, \overline{\mathcal{F}}, (\overline{\mathcal{F}}_t)_{t\geq 0}, \overline{\mathbb{P}})$ is a very good filtered extension of the original probability space (see e.g., Chapter 2 of Jacod and Protter (2012)).

4 Inference About the Asymptotic Autocorrelation of Volatility Increments

This section proposes estimators of the asymptotic autocorrelation of volatility increments based on the limit theorems from Section 3. For brevity, we focus on deriving feasible asymptotic confidence intervals from Theorem 5; the same method allows the construction of consistent estimators under the weaker assumptions of Theorem 4.

Because volatility is unobserved, direct inference on the asymptotic ACF using the realized autocovariance statistics from (3.4) incurs attenuation biases of the same kind as in classical errors-in-variables models (Hausman (2001)): sampling errors in spot variance estimators inflate estimated variances but not autocovariances, shrinking the empirical autocorrelation of volatility increments towards zero. The present problem, however, differs from standard mismeasured econometric models in that the latent variables, increments in latent volatility, and the error terms have a time series structure. As Theorem 5 shows, an important implication is that $V^{n,\ell}$ for $\ell \geq 2$ as well as the linear combination $V^{n,0} + 2V^{n,1}$ are asymptotically unbiased and stably converge to a mixed normal distribution centered at V^{ℓ} and $V^0 + 2V^1$, respectively.

This fact alone is not yet sufficient for feasible estimation of the asymptotic ACF of volatility increments, which requires division by V^0 and not $V^0 + 2V^1$. Theorem 1 provides a remedy by establishing the equivalence between estimating the asymptotic ACF of volatility increments and estimating the smoothness parameter e of volatility.

Our approach, therefore, is to first construct a feasible GMM estimator of e and then apply the delta method to obtain an estimator of $(\Psi_e(\tau))_{\tau \geq 0}$. For the first step, we choose $m \geq 2$ lags $\ell = 1, \dots, m$ as well as symmetric positive definite (and possibly random) weight matrices $\mathcal{W}_n \in \mathbb{R}^{m \times m}$. We obtain an estimator of $\Theta_t = (e, R_t)^T$, where $R_t = \int_0^t \zeta_s \, ds$, by minimizing the weighted distance between the realized autocovariances of volatility increments and their large-sample limit under a generic parameter vector $\Theta = (E, R)^T$:

$$\widehat{\Theta}_t^n = (\widehat{e}_n, \widehat{R}_n)^T = \underset{(E,R)}{\arg\min} \, F_t^n(E, R) \tag{4.1}$$

where $F_t^n(E,R) = \|\mathcal{W}_n^{1/2}(\widetilde{V}_t^n - \widetilde{\Phi}_{\kappa,E}^n R)\|^2$, $\widetilde{V}_t^n = (\widehat{V}_t^{n,0} + 2\widehat{V}_t^{n,1}, \widehat{V}_t^{n,2}, \dots, \widehat{V}_t^{n,m})^T$, $\widetilde{\Phi}_{\kappa,E}^n = (p_n \Delta_n)^{2E} \overline{\Phi}_{\kappa,E}$ and $\overline{\Phi}_{\kappa,E} = (\Phi_{\kappa,E}(0) + 2\Phi_{\kappa,E}(1), \Phi_{\kappa,E}(2), \dots, \Phi_{\kappa,E}(m))^T$.[16]

[16]Definition (4.1) is to be understood as follows: for every $E \in (-\frac{1}{2}, \frac{1}{2})$, because $\mathcal{W}_n$ is positive definite, $\widehat{R}_n(E) = (\widetilde{\Phi}_{\kappa,E}^n)^T \mathcal{W}_n \widetilde{V}_t^n / (\widetilde{\Phi}_{\kappa,E}^n)^T \mathcal{W}_n \widetilde{\Phi}_{\kappa,E}^n$ is the unique minimizer of the function $R \mapsto F_t^n(E,R)$. We

Theorem 6. *Let $t \in (0,T]$ and suppose that $\mathcal{W}_n \xrightarrow{\mathbb{P}} \mathcal{W}$, where $\mathcal{W} \in \mathbb{R}^{m\times m}$ is a deterministic symmetric positive definite matrix. Under Assumption CLT, the sequence (4.1) is unique with probability converging to one and*

$$D_n^{-1}(\check{\Theta}_t^n - \Theta_t) \xrightarrow{\mathcal{L}-s} \left(G_t^T \mathcal{W} G_t\right)^{-1} G_t^T \mathcal{W} \mathcal{Z}_t, \tag{4.2}$$

where $\check{\Theta}_t^n = (\widehat{e}_n, \Delta_n^{2(1-2\alpha)_+}\widehat{R}_t^n)^T$, $\mathcal{Z}$ is defined in Theorem 5, $G_t = (\overline{\Phi}'_{\kappa,e}R_t, \overline{\Phi}_{\kappa,e}) \in \mathbb{R}^{m\times 2}$ has full rank, $\overline{\Phi}'_{\kappa,e} = \partial_e \overline{\Phi}_{\kappa,e}$ and $D_n = (p_n\Delta_n)^{1/2}\left(\begin{smallmatrix} 1 & 0 \\ -2\log(p_n\Delta_n)R_t & 1\end{smallmatrix}\right)$.

Theorem 6 implies that the rate of convergence of $\widehat{e}_n$ is $(p_n\Delta_n)^{-1/2}$. We note that $\Delta_n^{2(1-2\alpha)_+}\widehat{R}_t^n \xrightarrow{\mathbb{P}} R_t$. Thus, $\widehat{R}_t^n$ consistently estimates R_t only if $\alpha \geq \frac{1}{2}$. If $\alpha < \frac{1}{2}$, R_t can be estimated consistently if α is known. As we show below, one does not need to know α in order to construct feasible confidence intervals for e.

The asymptotic variance in (4.2) has the usual sandwich form. It is minimized if the weight matrix $\mathcal{W}_n$ is chosen such that $\mathcal{W}_n \xrightarrow{\mathbb{P}} \mathcal{C}_t^{-1}$ from (3.15). We now construct a consistent estimator of the asymptotic covariance matrix by adopting Li and Xiu (2016)'s approach. For fixed $t \in (0,T]$, define centered products of volatility increments by

$$\psi_i^{n,\ell} = \begin{cases} (\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_i^n)^2 + 2\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{i+p_n}^n\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_i^n - m_i^{n,0} - 2m_i^{n,1} & \text{if } \ell = 1, \\ \Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{i+\ell p_n}^n\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_i^n - m_i^{n,\ell} & \text{if } \ell \geq 2, \end{cases}$$

where $m_i^{n,\ell} = \frac{1}{K_n}\sum_{j=i}^{i+K_n-1}\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{j+\ell p_n}^n\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_j^n$ for a bandwidth parameter $K_n \geq 1$. We estimate their cross-covariances at lag L by $\widehat{\Sigma}_t^{n,L} = (\widehat{\Sigma}_{\ell,\ell',t}^{n,L})_{1\leq\ell,\ell'\leq m}$, where

$$\widehat{\Sigma}_{\ell,\ell',t}^{n,L} = \Delta_n \sum_{i=0}^{N_{\ell,\ell'}^{n,L}(t)} \psi_i^{n,\ell}\psi_{i+L}^{n,\ell'}, \tag{4.3}$$

and $N_{\ell,\ell'}^{n,L}(t) = \lfloor t/\Delta_n\rfloor - K_n - k_n - [(\ell+1)p_n \vee ((\ell'+1)p_n + L)] + 1$. Finally, we define

$$\widehat{\Sigma}_t^n = \widehat{\Sigma}_t^{n,0} + \sum_{L=1}^{L_n} W(L,L_n)(\widehat{\Sigma}_t^{n,L} + (\widehat{\Sigma}_t^{n,L})^T), \tag{4.4}$$

where $W(\cdot,\cdot)$ is a deterministic weight function and $L_n \geq 1$ is the number of lags used in

then define $\widehat{e}_n = \arg\min_{E\in[-1/2,1/2]} F_t^n(E, \widehat{R}_n(E))$, where the minimum exists because $F_t^n(E,\widehat{R}_n(E))$ is continuous on $(-\frac{1}{2},\frac{1}{2}]$ and can be continuously extended to $E = -\frac{1}{2}$. Finally, we define $\widehat{R}_n = 0$ if $\widehat{e}_n = -\frac{1}{2}$ and $\widehat{R}_n = \widehat{R}_n(\widehat{e}_n)$ otherwise. We show that $\mathbb{P}(\widehat{e}_n = -\frac{1}{2}) \to 0$.

the asymptotic covariance estimation.

Theorem 7. *Grant the assumptions of Theorem 6 and suppose in addition that* $\widetilde{\varpi} > e + \frac{1}{4}$, $K_n \sim \vartheta \Delta_n^{-u'}$ *and* $L_n \sim \vartheta' \Delta_n^{-u}$ *for some* $\vartheta, \vartheta' > 0$ *and* $u, u' > \frac{1}{2}$ *satisfying* $\frac{1}{2} < u < \frac{1}{2}(1-e) \wedge (\widetilde{\varpi} - e + \frac{1}{4}) \wedge (\frac{1}{4} + \frac{1}{2}u') \wedge (\frac{1}{2} + (\frac{3}{8} \wedge (1-u'))(e^{\eta} \wedge e^{\overline{\eta}} + \frac{1}{2}))$. *Further assume that* W *is uniformly bounded and satisfies* $W(L, L_n) \to 1$ *for every* $L \geq 0$. *Then* $(p_n \Delta_n)^{-2-4e} p_n^{-1} \Delta_n^{4(1-2\alpha)_+} \widehat{\Sigma}_t^n \xrightarrow{\mathbb{P}} \mathcal{C}_t$.

We obtain interval estimates for e and, by the delta method, of the asymptotic ACF of volatility increments as an immediate consequence.

Corollary 1. *Assume that the conditions of Theorem 7 hold and let* $e_1 = (1,0)^T$ *and* $\widehat{G}_t = (\overline{\Phi}'_{\kappa,\widehat{e}_n} \widehat{R}_t^n, \overline{\Phi}_{\kappa,\widehat{e}_n}) \in \mathbb{R}^{m \times 2}$. *If*

$$\mathbb{V}_n^e = (p_n \Delta_n)^{-1-4\widehat{e}_n} p_n^{-1} e_1^T (\widehat{G}_t^T \mathcal{W}_n \widehat{G}_t)^{-1} \widehat{G}_t^T \mathcal{W}_n \widehat{\Sigma}_t^n \mathcal{W}_n \widehat{G}_t (\widehat{G}_t^T \mathcal{W}_n \widehat{G}_t)^{-1} e_1, \tag{4.5}$$

which for $\mathcal{W}_n = (\widehat{\Sigma}_t^n)^{-1}$ *simplifies to*

$$\mathbb{V}_n^e = (p_n \Delta_n)^{-1-4\widehat{e}_n} p_n^{-1} e_1^T (\widehat{G}_t^T (\widehat{\Sigma}_t^n)^{-1} \widehat{G}_t)^{-1} e_1, \tag{4.6}$$

then asymptotic $(1-\alpha)$*-confidence intervals for* e *and* $\Psi_e(\tau)$, *where* $\alpha \in (0,1)$ *and* $\tau \geq 0$, *are given by*

$$[\widehat{e}_n \pm \Phi^{-1}(1 - \tfrac{\alpha}{2})(\mathbb{V}_n^e)^{1/2}] \quad \text{and} \quad [\Psi_{\widehat{e}_n}(\tau) \pm \Phi^{-1}(1 - \tfrac{\alpha}{2}) |\partial_e \Psi_{\widehat{e}_n}(\tau)| (\mathbb{V}_n^e)^{1/2}], \tag{4.7}$$

respectively, where Φ^{-1} *is the quantile function of the standard normal distribution.*

5 Monte Carlo Study

This section studies the finite-sample behavior of the GMM estimator from Section 4.

5.1 Asset Price Dynamics and Noise

The efficient log-price satisfies

$$dx_t = (b - \tfrac{1}{2}\sigma_t^2)\, dt + \sigma_t\, dW_t + \int_{\mathbb{R}} z\, (\mu(dt, dz) - \nu_t(dz)\, dt), \tag{5.1}$$

where W is a standard Brownian motion, the equity risk premium is $b = 0.1$ in annualized terms and μ is an integer-valued random measure generating jumps in the log-price. Its compensator is $\nu_t(dz)\,dt$, where $\nu_t(dz) = c\sigma_t^2 \frac{e^{-\lambda|z|}}{|z|^{\beta+1}}\,dz$. Such a specification implies that jumps are from a time-changed tempered stable process, with the time change given by spot variance (cf. Duffie et al. (2000)). The parameter β is the Blumenthal–Getoor index of x and controls the activity of small jumps. We experiment with $\beta = 0.5$ (Case J1), which leads to price jumps of infinite activity but finite variation, and $\beta = 1.5$ (Case J2), which leads to infinite variation jumps. The latter case also allows us to examine the robustness of our estimators to high jump activity as $\beta = 1.5$ is not covered by Assumption CLT. We set $\lambda = 500$ and $c = 0.1\lambda^{2-\beta}/\Gamma(2-\beta)$. This implies $\int_{\mathbb{R}} z^2\,\nu_t(dz) = 0.2\sigma_t^2$, which is consistent with previous empirical results about the percentage of return variance due to jumps.

We consider Heston (1993)'s model for spot variance in which

$$\sigma_t^2 = \sigma_0^2 + \int_0^t \kappa(\theta - \sigma_s^2)\,ds + \int_0^t \nu\sigma_s\,dB_s \tag{5.2}$$

and the smoothness parameter is given by $e = 0$. The standard Brownian motion B satisfies $\mathrm{Corr}(dW_t, dB_t) = \rho\,dt$ with $\rho = -0.7$, which generates a negative leverage effect. We further choose $\theta = 0.02$, $\kappa = 8.32$ and $\nu = 0.55$ in annualized terms. This corresponds to an unconditional mean of spot variance of 0.02 and a half-life of a shock of one month. The parameter ν is chosen close to the upper bound imposed by Feller's condition.

We further consider cases in which spot variance has rough Heston dynamics (El Euch and Rosenbaum (2019)) with $e \in \{-0.4, -0.3, -0.2, -0.1\}$. Then

$$\sigma_t^2 = \sigma_0^2 + \frac{1}{\Gamma(e+1)}\left(\int_0^t (t-s)^e\kappa(\theta - \sigma_s^2)\,ds + \int_0^t (t-s)^e\nu\sigma_s\,dB_s\right),$$

where ν is chosen in such a way that σ_t^2 has the same unconditional variance as in the standard Heston model from (5.2). That is, we have $\nu = 0.16, 0.29, 0.39, 0.48$ if $e = -0.4, -0.3, -0.2, -0.1$.

In addition, we also consider cases in which volatility has jumps:

$$\sigma_t^2 = \begin{cases} \sigma_0^2 + \displaystyle\int_0^t \kappa(\theta - \sigma_s^2)\,ds + \int_0^t \nu\sigma_s\,dB_s + L_t & \text{if } e = 0, \\ \sigma_0^2 + \dfrac{1}{\Gamma(e+1)}\left(\displaystyle\int_0^t (t-s)^e\kappa(\theta - \sigma_s^2)\,ds + \int_0^t (t-s)^e\nu\sigma_s\,dB_s\right) + L_t & \text{otherwise.} \end{cases}$$

Here L is an independent gamma process with Lévy density $c_v e^{-\lambda_v z} z^{-1} \mathbf{1}_{z>0}$, where $c_v = 3.45$ (in annualized terms) and $\lambda_v = 17$ are the same volatility jump parameters as in Chong and Todorov (2025). We label the cases without (with) volatility jumps as V1 (V2).

Observed returns include additive noise of the form $\varepsilon_i^n = \sigma_{\text{noise}}(\int_{(i-1)\Delta_n}^{i\Delta_n} \sigma_s^2 \, ds)^{1/2} \chi_i$. We experiment with the case where $\chi_i = (\zeta_i - \zeta_{i-1})/\sqrt{2}$ for independent standard normal variables ζ_i and $\sigma^2_{\text{noise}} = 1.09 - 1 = 0.09$ (N1). The value 1.09 is approximately the median empirical ratio of daily realized variance computed from five-second versus five-minute returns in our SPY sample. To check the robustness of our estimator, we consider a second configuration (N2) in which σ_{noise} increases by a factor of 10 and χ_i exhibits time series dependence and is the first-order difference of $(\sum_{j=1}^{10} \rho_{\text{noise}}^{2j})^{-1} \sum_{j=1}^{10} \rho_{\text{noise}}^{j} \zeta_{i-j}$. We follow Li and Linton (2022) and choose $\rho_{\text{noise}} = 0.7$. This specification is consistent with Li and Linton (2022)'s observation that market microstructure noise decorrelates at four to five lags in calendar time.

5.2 Implementation and Choice of Tuning Parameters

The efficient price process is generated by an Euler scheme with a step size of one second and subsampling every five seconds, yielding 4,680 returns per trading day. We simulate 300 replications per configuration, each of which consists of 252 independently simulated trading days with variance initialized at θ.

We choose $k_n = 96$ and $p_n = 120$, which corresponds to estimating spot variance every eight minutes and leaving a two-minute gap in between. We use lags $\ell = 1, \ldots, 6$, which correspond to considering autocorrelation of volatility changes up to one hour apart, and the optimal weight matrix $\mathcal{W}_n = (\widehat{\Sigma}_t^n)^{-1}$ (see (4.4)) in our GMM procedure. We choose $K_n = 720$, $L_n = 180$ as well as Parzen weights $W(L, L_n)$ for estimating the asymptotic variance. The threshold for price jump truncation in (3.2) is set to $v_n = 3\sqrt{\frac{\Delta_n}{T} BV^n_{t,t+T}}$ on day t, where $BV^n_{t,t+T}$ is the bipower variation estimator of Barndorff-Nielsen and Shephard (2004b) based on five-second returns on day t. The threshold w_n for volatility jumps on day t is set to three empirical standard deviations of all volatility increments $\Delta_{p_n} \widehat{c}_i^n$ on that day, with the largest 5% of increments (in absolute value) removed. To improve the finite-sample performance of our estimator, we replace volatility increments exceeding w_n in absolute value by w_n or $-w_n$ (depending on the sign) instead of zero. In order to mimic the empirical application, we further normalize spot variance increments on each day by five-minute realized variance over that day.

5.3 Results

Table 1 reports the bias, standard error (SE) and root mean square error (RMSE) of $\widehat{e}_n$ in absolute terms. We observe that $\widehat{e}_n$ has a small upward bias of up to 0.03 for all considered values of e, except for the case $e = 0$, in which it is essentially unbiased. While our theory shows that this bias vanishes asymptotically, it is nevertheless present in finite samples. Standard errors tend to decrease as e changes from -0.4 to -0.3 and increase again as e increases further, attaining the largest values when $e = 0$.

Consistent with our theory, price and volatility jumps as well as noise do not meaningfully impact the bias or standard error of $\widehat{e}_n$. Overall, the simulation confirms a good finite-sample behavior of our GMM estimator.

Table 1: Monte Carlo Results with Jumps

	e	J1-V1-N1	J1-V1-N2	J1-V2-N1	J1-V2-N2	J2-V1-N1	J2-V1-N2	J2-V2-N1	J2-V2-N2
Bias		0.020	0.021	0.022	0.023	0.020	0.020	0.023	0.024
SE	-0.4	0.031	0.030	0.031	0.030	0.031	0.030	0.029	0.029
RMSE		0.037	0.036	0.038	0.038	0.037	0.037	0.037	0.037
Bias		0.013	0.013	0.015	0.015	0.013	0.013	0.016	0.016
SE	-0.3	0.027	0.026	0.027	0.026	0.026	0.025	0.026	0.026
RMSE		0.030	0.029	0.031	0.030	0.029	0.029	0.031	0.031
Bias		0.022	0.023	0.025	0.025	0.024	0.023	0.026	0.025
SE	-0.2	0.027	0.028	0.026	0.028	0.026	0.027	0.027	0.028
RMSE		0.035	0.036	0.036	0.038	0.035	0.035	0.037	0.038
Bias		0.026	0.027	0.029	0.029	0.027	0.028	0.030	0.029
SE	-0.1	0.034	0.038	0.034	0.037	0.034	0.038	0.035	0.036
RMSE		0.043	0.046	0.044	0.047	0.043	0.047	0.046	0.046
Bias		-0.003	0.005	-0.004	0.004	-0.002	0.004	-0.001	0.004
SE	0	0.062	0.063	0.058	0.060	0.065	0.065	0.060	0.060
RMSE		0.062	0.063	0.058	0.060	0.065	0.065	0.060	0.060

Note: The table reports the bias, SE and RMSE of $\widehat{e}_n$ in absolute terms in the presence of jumps and truncation. The numbers are based on 300 Monte Carlo replications.

6 Empirical Application

We estimate the smoothness parameter and the asymptotic autocorrelation of intraday volatility changes using SPY transaction data from 2010 to 2025. We sample prices every five seconds from 9:30 a.m. to 4:00 p.m. ET after applying standard filters to the tick-by-tick data. We remove trading days with halts of longer than five minutes, half-trading days and FOMC announcement days.

6.1 The Smoothness Parameter of Volatility

We implement the GMM estimator $\widehat{e}_n$ of the smoothness parameter of intraday volatility from Section 4. We choose tuning parameters exactly as in the Monte Carlo study. As a robustness check, we also report results for $k_n = 48$ and $p_n = 60$, which correspond to estimating volatility on four-minute blocks with a one-minute gap in between.

We take measures to mitigate the effects of heteroscedasticity in daily volatility and diurnal variation in intraday volatility. We first divide spot variance estimates by five-minute truncated realized variance on the same day. As a consequence, when we form increments of these estimates, we consider changes in volatility relative to the overall level of volatility on that day. As we are interested in estimating autocorrelation, this normalization has no effect on the asymptotic behavior of our estimators but improves their finite-sample performance by reducing interdaily heteroscedasticity in the spot variance estimates.[17] Next, we estimate the diurnal pattern of volatility by computing averages of normalized spot variance estimates at fixed times of the day over the previous 252 trading days. We divide the normalized spot variance estimates by this estimated pattern before applying the GMM procedure from Section 4 to their realized autocovariances.

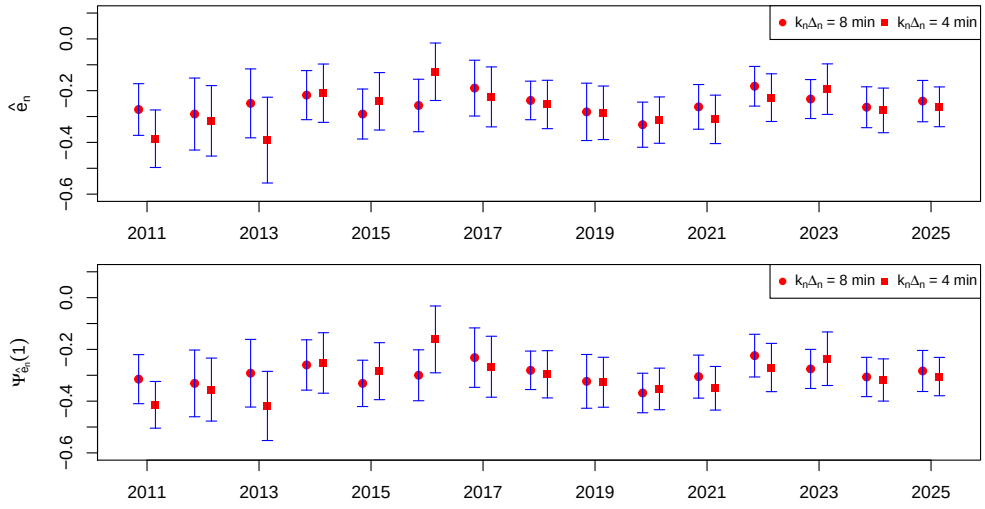


Figure 1: 95% confidence intervals of the smoothness parameter of volatility (top panel) and of the first-order autocorrelation of volatility increments (bottom panel).

The top panel of Figure 1 shows yearly estimates of the smoothness parameter of volatility

[17] We obtain very similar results when we take logarithm of spot variance estimates instead.

from 2011 to 2025. (The data from 2010 is only used to estimate the diurnal pattern of volatility.) The point estimates are all negative, with 95% confidence intervals not covering zero in all years. This provides significant evidence of negative serial dependence in intraday volatility changes. At the same time, the 95% confidence intervals also lie above $e = -0.4$ in most years. In models driven by fractional Brownian motion, $e = -0.4$ corresponds to a Hurst index of $H = 0.1$, which is reported in previous works on the roughness of daily (realized or integrated) volatility (see e.g., Gatheral et al. (2018), Bolko et al. (2023) and Wang et al. (2023)). The bottom panel of Figure 1 shows point and interval estimates based on (4.7) for the first-order autocorrelation of intraday volatility changes.

The yearly estimates are relatively stable throughout the considered sample. Using all data from 2011 to 2025, we obtain an aggregate estimate of $\widehat{e}_n = -0.250$ ($\widehat{e}_n = -0.257$) when $k_n = 96$ ($k_n = 48$). In models based on fractional Brownian motion, this corresponds to a Hurst index of $H = 0.250$ ($H = 0.243$), which is consistent with the findings of Muhle-Karbe et al. (2026). Figure 2 shows the estimated limiting ACF $(\Psi_{\widehat{e}_n}(\tau))_{\tau=0,\dots,10}$ of volatility increments based on $\widehat{e}_n = -0.250$. As the ACF corresponding to $\widehat{e}_n = -0.257$ is visually almost indistinguishable, we do not plot it in Figure 2.

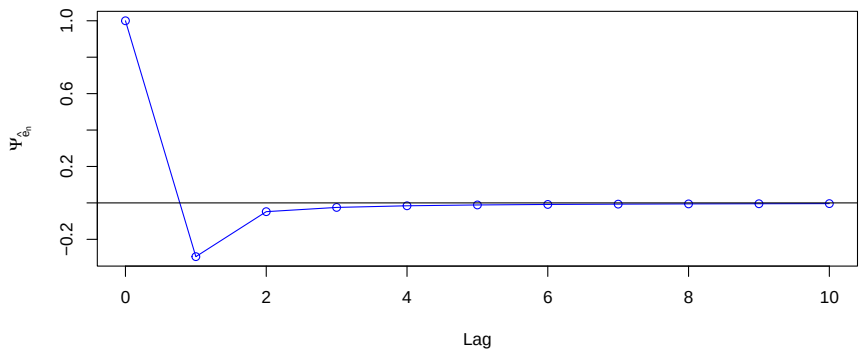


Figure 2: Estimated limiting ACF of changes in latent spot variance. The graph shows $(\Psi_{\widehat{e}_n}(\tau))_{\tau=0,\dots,10}$ as defined in (2.4), where $\widehat{e}_n = -0.250$ is the estimate of e based on all data from 2011 to 2025.

6.2 Predictability of Intraday Realized Variance

While the previous results show significant predictability in changes of latent spot volatility, the effect size will be attenuated for intraday realized variance due to the presence of sampling

errors.[18] We now evaluate the magnitude of this attenuation bias and both the in-sample and out-of-sample predictability of short-horizon realized variance.

We start by investigating to what extent the nonrobust empirical autocorrelation estimator $\widehat{r}_2 = \widehat{V}_t^{n,2}/\widehat{V}_t^{n,0}$ is affected by attenuation bias due to sampling errors in spot variance. The first row in Table 2 shows $\widehat{r}_2$, with $\widehat{V}_t^{n,2}$ and $\widehat{V}_t^{n,0}$ computed over the whole sample from 2011 to 2025, for different choices of k_n. The values of k_n correspond to estimating volatility on intervals of 5, 10, 20, 30, 60 and 90 minutes, respectively. We choose $p_n = k_n$ for simplicity; leaving a one-minute gap (i.e., taking $k_n = p_n - 12$) does not meaningfully change the results.

Table 2: Empirical Autocorrelation and Sign Persistence of Spot Variance Increments

	$k_n\Delta_n$						
	5 minutes	10 minutes	20 minutes	30 minutes	60 minutes	90 minutes	**Robust Estimate**
$\widehat{r}_2$	-0.032	-0.045	-0.091	-0.137	-0.152	-0.121	**-0.125**
$\widehat{p}$	49.1%	48.8%	48.0%	47.2%	45.6%	45.2%	**46.0%**

Note: The table reports $\widehat{r}_2 = \widehat{V}_t^{n,2}/\widehat{V}_t^{n,0}$ as well as the proportion of consecutive increments of spot variance estimates with the same sign, $\widehat{p}$, for different window sizes $k_n\Delta_n$. The last column contains robust estimates given by $\Phi_{1,\widehat{e}_n}(2)/\Phi_{1,\widehat{e}_n}(0)$ and $1/2 + \arcsin(\Phi_{1,\widehat{e}_n}(2)/\Phi_{1,\widehat{e}_n}(0))/\pi$, where $\Phi_{\kappa,e}(\ell)$ is defined in (B.2) and $\widehat{e}_n = 0.250$ is the full-sample estimate of the smoothness parameter. We use the fact that $\mathbb{P}(XY > 0) = 1/2 + \arcsin(\rho)/\pi$ for two mean-zero normal variables X and Y with correlation ρ.

As we can see from the first row of the table, sampling errors of spot variance at five-minute frequency are large and shrink the empirical second-order autocorrelation of volatility increments towards zero, while our robust estimator $\widehat{e}_n$ estimates the population value at -0.125. As the interval length for volatility estimation increases, the attenuation bias declines and becomes small for window sizes of 30 minutes or more.

The second row of Table 2 reports the average proportion of consecutive increments of spot variance estimates with the same sign, that is, $\widehat{p}$ is the mean of $\mathbf{1}_{\Delta^{\mathrm{tr}}_{p_n}\widehat{c}^n_{i+2p_n}\Delta^{\mathrm{tr}}_{p_n}\widehat{c}^n_i>0}$. In the absence of sampling errors, one would expect this estimate to be above, around or below 50%, depending on whether $e > 0$, $e = 0$ or $e < 0$. The exact limiting probability is given by $1/2 + \arcsin(\Phi_{1,e}(2)/\Phi_{1,e}(0))/\pi$. Consistent with our previous observations, sampling errors lead to estimates $\widehat{p}$ which are close to 50% for small values of k_n. As k_n increases, $\widehat{p}$ becomes smaller and approximately equals the robust estimate of 46.0% at 60-minute frequency. The

[18] Andersen et al. (2005) emphasize this point for daily realized variance.

overall good match between $\widehat{p}$ for larger values of k_n and the robust estimate implied by $\widehat{e}_n$ shows that the negative serial correlation in volatility changes is *not* merely caused by short extreme periods of low volatility smoothness, which could be a concern a priori due to Remark 2.

Next, we consider prediction of intraday movements of 30-minute truncated realized variance, that is, of $\Delta_{k_n}\widehat{c}^n_{jk_n}$ from (3.3) constructed from five-second returns with $p_n = k_n = 360$. As Table 2 shows, sampling errors do not have a large impact on the second-order autocorrelation estimate $\widehat{r}_2$ for this value of k_n. We employ the linear predictor[19]

$$\Delta_{k_n}\widehat{c}^{n,\mathrm{pred}}_{jk_n} = \beta_{L,0} + \sum_{\ell=1}^{L} \beta_{L,\ell}\Delta_{k_n}\widehat{c}^n_{(j-2\ell)k_n}, \tag{6.1}$$

for $L = 0, \ldots, 3$. This corresponds to predicting the change of volatility (after removing the intraday pattern) between 12:30pm and 1:30pm, between 1:30pm and 2:30pm, and between 2:30pm and 3:30pm using volatility increments up to three hours before.

If volatility is an Itô semimartingale ($e = 0$), one expects nonoverlapping increments of volatility estimates to be asymptotically uncorrelated. Therefore, models with $L \geq 1$ should do no better than the constant model with $L = 0$ after accounting for model complexity. If volatility has a smoothness parameter $e \approx -0.5$, volatility increments behave like MA(1) noise (see footnote 6), in which case one would expect the model with $L = 1$ to perform best. If e is strictly between -0.5 and 0, one expects (6.1) with $L = 2$ and $L = 3$ to outperform the $L = 1$ model on a complexity-adjusted basis.

Table 3: Least Squares Estimates and HAC Standard Errors

L	$\widehat{\beta}_{L,0}$	$\widehat{\beta}_{L,1}$	$\widehat{\beta}_{L,2}$	$\widehat{\beta}_{L,3}$
0	-0.008 (0.004)			
1	-0.009 (0.003)	-0.160 (0.013)		
2	-0.009 (0.003)	-0.168 (0.013)	-0.065 (0.013)	
3	-0.009 (0.003)	-0.169 (0.014)	-0.069 (0.013)	-0.048 (0.013)

[19] In contrast to the existing literature on realized variance forecasting (see e.g., Andersen et al. (2003), Andersen et al. (2005), Corsi (2009) and Wang et al. (2023) for daily realized variance and Bennedsen et al. (2022) and Li et al. (2025a) for intraday spot variance), we focus on predictability of *changes* in short-horizon realized variance. We consider nonoverlapping volatility increments in (6.1) in order to avoid capturing predictability due to sampling errors only. Moreover, we remove the intraday volatility pattern and normalize volatility changes by daily truncated realized variance. We further trim outliers exceeding in absolute value the truncation threshold w_n as determined in Section 5.2.

We first consider the in-sample performance of the prediction model (6.1). Table 3 shows the ordinary least squares estimates of $\beta_{L,\ell}$ together with heteroscedasticity- and autocorrelation-consistent (HAC) standard errors. For all values of L, the estimated autoregressive coefficients are negative and significant at 5% level. This indicates that realized variance tends to reverse during the day and is consistent with a volatility smoothness parameter between, but bounded away from, -0.5 and 0.

In order to compare forecasting models with different lags while accounting for model complexity, we report the Alkaike information criterion (AIC) and the Bayesian information criterion (BIC) in the second and third columns of Table 4. Both AIC and BIC decrease monotonically in L. To evaluate the significance of these in-sample performance differences, we compute Hansen et al. (2011)'s model confidence sets at the 90% level. We observe that models with lags $L = 2$ and $L = 3$ lie within the model confidence set for both AIC and BIC, while the models with $L = 0$ and $L = 1$ do not for either criterion.

Table 4: Comparison of Forecasts of Changes in 30-Minute Realized Variance

L	AIC	BIC	MSE	MAE
0	-15320.98	-15306.40	1064.96	2211.25
1	-15596.80	-15574.93	**1043.02**	**2198.40**
2	**-15629.89**	**-15600.74**	**1043.10**	**2198.75**
3	**-15642.52**	**-15606.08**	**1043.42**	**2200.03**

Note: The loss functions associated with mean square error (MSE) and mean absolute error (MAE) are $L(y, \widehat{y}) = |y - \widehat{y}|^2$ and $L(y, \widehat{y}) = |y - \widehat{y}|$, respectively. Models within Hansen et al. (2011)'s 90% model confidence sets are highlighted in bold.

Next, we evaluate (pseudo) out-of-sample performance. For every value of $L = 0, \ldots, 3$, we compute the predictor in (6.1) using ordinary least squares estimates of $\beta_{L,\ell}$ based on data in a rolling window of the previous 252 days.[20] Columns 4 and 5 in Table 4 show the aggregate forecasting losses across the sample from 2012 to 2025 based on the mean square error (MSE) and the mean absolute error (MAE) loss functions. For both metrics, the model confidence set at 90% level contains the models corresponding to $L = 1, 2, 3$, but not the baseline model where $L = 0$.[21] This provides strong evidence for out-of-sample predictability

[20] For both training and prediction, we divide $\widehat{c}_i^n$ by the intraday volatility pattern estimated based on data in this 252-day window. In the training step, we further consider trimmed volatility increments, $\Delta_{k_n}^{\mathrm{tr}} \widehat{c}_{i-\ell}^n$, and normalize them by daily truncated realized variance to improve estimation accuracy. We do not, however, trim or normalize volatility increments in the prediction step. Hence, there is no look-ahead bias in the prediction step.

[21] As we estimate parameters on a rolling window of fixed length, we can use asymptotic inference based

of short-term volatility changes.

7 Implications for Large-Sample Properties of High-Frequency Estimators

Much of the econometric theory for high-frequency data hinges on the assumption that coefficients of the data-generating process such as volatility are sufficiently regular and approximately constant over short time intervals. For Itô semimartingale processes, such an assumption implies that the underlying model is locally Gaussian, which in turn enables nonparametric inference in an infill asymptotic setup (Mykland and Zhang (2009), Bollerslev et al. (2021)).

In this section, we illustrate the limits to such local-to-constant approximations in situations where volatility smoothness is low. It is well known that for an Itô semimartingale x as in (3.1) with locally bounded drift and volatility and no jumps, the realized variance estimator $RV^n(t) = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (\Delta_i^n x)^2$ satisfies the CLT

$$\Delta_n^{-1/2}(RV^n(t) - IV(t)) \xrightarrow{d} N\left(0, 2\int_0^t \sigma_s^4\, ds\right), \quad IV(t) = \int_0^t \sigma_s^2\, ds. \tag{7.1}$$

This CLT, which is the backbone of volatility estimation based on high-frequency data, extends to more general moment estimators. For $p \geq 2$, consider the *realized power variation* $RP_p^n(t) = \mu_p^{-1}\Delta_n^{1-p/2}\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} |\Delta_i^n x|^p$, where $\mu_p = \mathbb{E}[|N(0,1)|^p]$, as an estimator of $IP_p(t) = \int_0^t \sigma_s^p\, ds$. Then

$$\Delta_n^{-1/2}(RP_p^n(t) - IP_p(t)) \xrightarrow{d} N\big(0, \mu_p^{-2}\rho_p IP_{2p}(t)\big), \quad \rho_p = \mathrm{Var}(|N(0,1)|^p). \tag{7.2}$$

The two limit results differ in the required assumptions. While the CLT in (7.1) for realized variance holds without further assumptions about volatility, the CLT in (7.2) for $p \neq 2$ is proved under the additional hypothesis that volatility itself is an Itô semimartingale (see Theorems 5.3.6 and 5.4.2 of Jacod and Protter (2012)).

We now show that the CLT in (7.2) can break down if volatility has a low smoothness parameter.

on the normal distribution for the Diebold and Mariano (1995) and West (1996)-type tests in constructing the model confidence set (Giacomini and White (2006)).

Theorem 8. *Consider an Itô semimartingale process x as in (3.1), where b is locally bounded and càdlàg and $\sigma_t^2 = c_t$ satisfies Assumption V2 and Assumption CLT (i)–(iii), with $\delta^l \equiv \delta^s \equiv \delta_c^l \equiv \delta_c^s \equiv \delta_\eta^l \equiv \delta_{\overline{\eta}}^l \equiv \delta_\eta^s \equiv \delta_{\overline{\eta}}^s \equiv 0$ (no jumps in c_t, η_t^2 and $\overline{\eta}_t^2$) and $e = e^\eta = e^{\overline{\eta}}$ (i.e., c_t, η_t^2 and $\overline{\eta}_t^2$ have the same smoothness parameter). If $e > -0.25$, then (7.2) continues to hold, while if $e < -0.25$, we have*

$$\Delta_n^{-2e-1}(RP_p^n(t) - IP_p(t)) \xrightarrow{\mathbb{P}} B_p(t), \tag{7.3}$$

where

$$\begin{aligned} B_p(t) = \frac{p(p-2)}{4} \int_0^t c_s^{p/2-2} \Bigg[& \frac{2\cos(\pi e)(\eta_s^2 + c_s \eta_s^\eta / \eta_s) - (\eta_s^2 + \overline{\eta}_s^2)}{2(2e+2)(2e+3)} \\ & + \frac{(p-4)\cos(\pi e)\Gamma(2e+2)\eta_s^2}{2\Gamma(e+3)^2} \Bigg] ds. \end{aligned} \tag{7.4}$$

Note that $\Delta_n^{1/2} = o(\Delta_n^{2e+1})$ if $e < -0.25$. So in this case, nonquadratic realized power variation estimators have an asymptotic bias of intermediate order that tends to zero at the slower rate of Δ_n^{2e+1} and renders inference about $IP_p(t)$ based on the CLT in (7.2) invalid.

8 Conclusion

This paper studies the second-order properties of short-term volatility changes within a trading day. We develop point and interval estimators for the limiting ACF of increments of volatility over shrinking horizons that are robust to sampling errors of spot variance estimators, market microstructure noise and jumps in both asset price and volatility dynamics. We first use the theory of regular variation to show that the ACF of increments of a continuous-time process is determined by a single smoothness parameter in the infill asymptotic limit. We then derive GMM estimators of the smoothness parameter of volatility and, by extension, of the limiting ACF of volatility increments from CLTs for the realized autocovariances of spot variance estimators.

In an application to SPY transaction data from 2011 to 2025, we find significant negative autocorrelation in intraday volatility changes in each year of the sample. The confidence intervals for the smoothness parameter indicate that short-term volatility reversals occur more often than implied by classical Itô semimartingale volatility models but also less often than implied by rough volatility models with a near-zero Hurst exponent. This empirical

finding informs model choice in studies that focus on, or otherwise depend on, the dynamical properties of intraday volatility.

The empirical results carry over to changes in short-horizon *realized* variance. While the effect size is attenuated due to sampling errors, we find strong evidence of predictability in intraday volatility movements, both in sample and out of sample. This holds in addition to the pronounced diurnal pattern of intraday volatility.

We further show that a smoothness parameter of volatility below -0.25, as we find in about half of the years in our sample, can lead to biases of intermediate order that force a trade-off between asymptotic efficiency and inferential validity based on standard limit results. This occurs for nonquadratic estimators as simple as realized power variations. Determining how such biases affect more complex estimators and developing methods that restore valid asymptotic inference are promising avenues for future research.

A Proof of Theorems 1–3

Proof of Theorem 1. Using the identity $X_{t+\ell\Delta} - X_t = \sum_{j=0}^{\ell-1}(X_{t+(j+1)\Delta} - X_{t+j\Delta})$ and (2.2), we obtain

$$\begin{aligned}\frac{\gamma_t(\ell\Delta)}{\gamma_t(\Delta)} = \frac{\operatorname{Var}(\sum_{j=0}^{\ell-1}(X_{t+(j+1)\Delta} - X_{t+j\Delta}))}{\gamma_t(\Delta)} &= \sum_{j_1,j_2=0}^{\ell-1} \frac{\gamma_{t+(j_1\wedge j_2)\Delta}(|j_1 - j_2|;\Delta)}{\gamma_t(\Delta)} \\ &\longrightarrow \sum_{j_1,j_2=0}^{\ell-1} \rho_t(|j_1 - j_2|) \quad \text{for all integers } \ell \geq 1.\end{aligned} \tag{A.1}$$

By Karamata's characterization theorem, the convergence of $\gamma_t(\ell\Delta)/\gamma_t(\Delta)$ for integer $\ell \geq 1$ extends to all real-valued $\lambda > 0$, which means that γ_t is regularly varying at zero.[22] As a consequence, there is some $\alpha \in \mathbb{R}$ such that

$$\frac{\gamma_t(\lambda\Delta)}{\gamma_t(\Delta)} \longrightarrow \lambda^\alpha \quad \text{for all real-valued } \lambda > 0. \tag{A.2}$$

In order to identify $\rho_t(\tau)$, we use the polarization identity $2\operatorname{Cov}(Z_2 - Z_1, Z_4 - Z_3) = \operatorname{Var}(Z_4 - Z_1) + \operatorname{Var}(Z_3 - Z_2) - \operatorname{Var}(Z_3 - Z_1) - \operatorname{Var}(Z_4 - Z_2)$ to rewrite the covariance term

[22] More precisely, because (A.1) holds for $\ell = 2$ and $\ell = 3$ and $\log_2 3$ is irrational, we can use the equivalence of (i) and (iv) in Theorem 1.4.3 of Bingham et al. (1987) (and the subsequent remarks concerning a substitute for (1.4.4), which is (2.3) in our setting).

$\gamma_t(\tau;\Delta)$ as a linear combination of variance terms:

$$\begin{aligned} 2\gamma_t(\tau;\Delta) &= 2\operatorname{Cov}(X_{t+\Delta} - X_t, X_{t+(\tau+1)\Delta} - X_{t+\tau\Delta}) \\ &= \operatorname{Var}(X_{t+(\tau+1)\Delta} - X_t) + \operatorname{Var}(X_{t+\tau\Delta} - X_{t+\Delta}) - \operatorname{Var}(X_{t+\tau\Delta} - X_t) \\ &\quad - \operatorname{Var}(X_{t+(\tau+1)\Delta} - X_{t+\Delta}) \\ &= \gamma_t((\tau+1)\Delta) + \gamma_{t+(\tau\wedge 1)\Delta}(|\tau-1|\Delta) - \gamma_t(\tau\Delta) - \gamma_{t+\Delta}(\tau\Delta). \end{aligned} \tag{A.3}$$

As $\gamma_{t+\theta\Delta}(\lambda\Delta)/\gamma_t(\Delta) = \gamma_{t+\theta/\lambda(\lambda\Delta)}(\lambda\Delta)/\gamma_t(\lambda\Delta) \times \gamma_t(\lambda\Delta)/\gamma_t(\Delta) \to \rho_t(0)\lambda^\alpha = \lambda^\alpha$ for all $\lambda > 0$ and $\theta \geq 0$ by (2.2) and (A.2), dividing by $\gamma_t(\Delta)$ and letting $\Delta \to 0$ in (A.3) lead to $2\rho_t(\tau) = (\tau+1)^\alpha - 2\tau^\alpha + |\tau-1|^\alpha$, which shows (2.4) with $e = \frac{\alpha-1}{2}$. By the Cauchy–Schwarz inequality and (2.2), we have $|\gamma_t(\tau;\Delta)/\gamma_t(\Delta)| \leq \sqrt{\gamma_{t+\tau\Delta}(\Delta)/\gamma_t(\Delta)} \to \sqrt{\rho_t(0)} = 1$, which shows $|\rho_t(\tau)| \leq 1$ and thus $\alpha \in [0,2]$ (or $e \in [-\frac{1}{2}, \frac{1}{2}]$). □

Proof of Theorem 2. The proof is based on Abelian and Tauberian theorems for the cosine transform. Since X is covariance-stationary, the variance function $\gamma \equiv \gamma_t$ defined in (2.1) does not depend on t and $\gamma(r) = \operatorname{Var}(X_r - X_0) = 2(C(0) - C(r))$. By Theorem 1, X has smoothness parameter $e \in (-\frac{1}{2}, \frac{1}{2})$ if and only if γ, or $1 - C(r)/C(0)$, is regularly varying at zero with index $2e+1$. By Theorem 8.1.10 of Bingham et al. (1987), this in turn is equivalent to $1 - F(\lambda)/C(0) + F(-\lambda)/C(0)$ being regularly varying at infinity with index $-1-2e$. By symmetry, we have $F(-\lambda) = C(0) - F(\lambda)$ and therefore $1 - F(\lambda)/C(0) + F(-\lambda)/C(0) = 2(1 - F(\lambda)/C(0))$. This implies (i). In the special case where C is integrable, we have $C(0) - F(\lambda) = \int_\lambda^\infty f(x)\,dx$. Therefore, part (ii) of the theorem follows from Theorems 1.5.11 (ii) and 1.7.2 in Bingham et al. (1987). □

Proof of Theorem 3. We apply Theorem 1 to the processes X^{τ_n}, which shows that $\rho_t^{(n)}(k)$ is of the form (2.4) for some $e = e_n$. As the mapping $e \mapsto \Psi_e(1) = \frac{1}{2}(2^{2e+1} - 2)$ is one-to-one with a continuous inverse, the convergence of $\rho_t^{(n)}(1)$ implies that $e_n \to e$ for some $e \in [-\frac{1}{2}, \frac{1}{2}]$. This value of e therefore satisfies (2.6). □

B Assumptions and Limiting Constants in Section 3

We detail the assumptions for Theorem 4 and Theorem 5.

Assumption LLN. *We have* (3.1), (3.5)–(3.8) *together with Assumption V1 and the following properties:*

(i) *The initial condition x_0 is $\mathcal{F}_0$-measurable, W is a standard Brownian motion and π is a Poisson random measure on $[0,\infty)\times E$ with compensator $dt\,\lambda(dz)$, where λ is a σ-finite measure on an auxiliary Polish space $(E,\mathcal{E})$. The processes b, σ, ρ and v are predictable and locally bounded and v is locally bounded away from zero.*

(ii) *The sequence $(\chi_i)_{i\in\mathbb{N}}$ is stationary, independent of $\mathcal{F}_\infty = \bigvee_{t\geq 0}\mathcal{F}_t$, with finite moments of all orders, mean 0 and variance 1 and exponentially ρ-mixing: if $\mathcal{G}_i = \sigma(\chi_j,\, j\leq i)$, $\mathcal{G}^i = \sigma(\chi_j,\, j\geq i)$ and $\rho_k(\chi) = \sup\{|\mathbb{E}[UV]| : U$ is $\mathcal{G}_0$-measurable, V is $\mathcal{G}^k$-measurable, $\mathbb{E}[U]=\mathbb{E}[V]=0$, $\mathbb{E}[U^2]=\mathbb{E}[V^2]=1\}$, then there are $K,\rho>0$ such that $|\rho_k(\chi)|\leq K\exp(-\rho k)$ for all $k\geq 0$.*

(iii) *$\delta^s,\delta^l,\delta^s_c,\delta^l_c : \Omega\times\mathbb{R}_+\times E\to\mathbb{R}$ are predictable and there are deterministic measurable functions $\delta^*_m : E\to\mathbb{R}$, $0<r<\frac{2}{2+e}$ and $0<r_c<\frac{1}{e+1}$ such that $\int_E \delta^*_m(z)\,\lambda(dz)<\infty$ and $(|\delta^s(t,z)|^r + |\delta^s_c(t,z)|^{r_c} + |\delta^l(t,z)|^{r\wedge 1} + |\delta^l_c(t,z)|^{r_c\wedge 1})\wedge 1 \leq \delta^*_m(z)$ for all $m\in\mathbb{N}$, $t\leq\tau_m$ and $z\in E$.*

(iv) *We have $v_n\sim v\Delta_n^{\varpi}$ and $w_n\sim w\Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{\widetilde{\varpi}}$ for some $v,w>0$, $\varpi<\frac{1}{2}\wedge\alpha$ and $\frac{1+4e}{2(2-r_c)}<\widetilde{\varpi}<\frac{1}{2}+e$ satisfying $2(2-r)\varpi+\widetilde{\varpi}+2(1-2\alpha)_+>\frac{3}{2}+2e$ and, if $r>1$, also $\widetilde{\varpi}+\frac{2}{r}+(1-2\alpha)_+>\frac{5}{2}+2e$.*

(v) *We have $k_n\sim\theta\Delta_n^{-1/2}\log(\Delta_n^{-1})$ and $p_n/k_n=\kappa$ for some $\theta>0$ and $\kappa\in[1,\infty)$. If $e=0$, we further have $k_n^{-1}=o(\gamma(p_n\Delta_n))$.*

Condition (i) imposes standard measurability and boundedness assumptions on the coefficients in the asset and noise dynamics. Condition (ii) concerns the dependence structure of microstructure noise. The assumptions on χ_i are slightly weaker than those in Chong and Todorov (2025) (where m-dependence is assumed) and slightly stronger than those in Jacod et al. (2017), Li and Linton (2022) and Da and Xiu (2021) (where only a polynomial decay of the ρ-mixing coefficients is assumed). We conjecture that our results can be extended to situations in which χ_i is only polynomially ρ-mixing.

Conditions (iii) and (iv) impose restrictions on the jump activity of the price process x as well as c, the sum of spot variance and noise variance. For all $e<0$, we can allow both asset price jumps and volatility jumps to be of infinite variation, with the highest permissible jump activity index being larger for smaller values of e. This is intuitive as a rougher continuous part of the volatility process is more pronounced and therefore easier to separate from jumps than a smoother one. Remark 3 below explains how the condition on the price jump activity index can be further relaxed. Regarding the truncation thresholds ϖ and $\widetilde{\varpi}$, we note that

for all values of $e < 0$, $0 < r < \frac{2}{2+e}$ and $0 < r_c < \frac{1}{e+1}$, we can find permissible values of ϖ and $\widetilde{\varpi}$ by taking them close enough to their respective upper bounds. The restrictions (iii) and (iv) stem from the CLT result in Theorem 5. For the LLN result in Theorem 4, weaker assumptions would suffice. In order to avoid a separate treatment of jumps for the LLN and the CLT in the proofs, we do not pursue this direction here.

Condition (v) specifies a window size of $\Delta_n^{-1/2} \log \Delta_n^{-1}$ for spot variance estimation. This choice implies that the sampling errors are of a lower asymptotic order than the increments in latent volatility for all values of $e < 0$. The value of κ determines the size of the gap between volatility increments. For example, $k_n\Delta_n = 8\,\text{min}$ and $p_n\Delta_n = 10\,\text{min}$ (as in our empirical application) correspond to $\kappa = 1.25$.

Remark 3. *Extensions to cases where the asset price process has infinite variation jumps with an activity index $r \geq 1 - e$, which are ruled out by Assumption LLN, are possible. Consider, for example, the case where the jumps of x are given by those of a time-changed Lévy process $L_{\int_0^t \lambda_s\, ds}$, where L is a Lévy process and $(\lambda_t)_{t\geq 0}$ is a predictable and locally bounded process. If λ has the same smoothness parameter as c, then Theorem 5 can be extended to cases where L has infinite variation jumps with a Blumenthal–Getoor index exceeding $1 - e$. Just like microstructure noise, infinite variation jumps introduce biases in spot variance estimators but they cancel to first order when we form increments of spot variance estimators. Therefore, as long as the jump intensity λ has the same smoothness parameter as c, infinite variation jumps will not affect the validity of Theorem 5. We confirm this in the Monte Carlo study of Section 5 where we consider a configuration with $r = \frac{3}{2}$, which is formally excluded by Assumption LLN.*

Assumption CLT. *We have Assumptions LLN and V2 and the following properties:*

(i) The volatility-of-volatility processes have the following representations:

$$\begin{aligned}
\eta_t^2 &= \eta_0^2 + \int_0^t a_s^\eta\, ds + \int_0^t \int_E \delta_\eta^l(s,z)\, \pi(ds,dz) + \int_0^t \int_E \delta_\eta^s(s,z)\, \widetilde{\pi}(ds,dz) \\
&\quad + \int_0^t h^\eta(t-s)\theta_s^\eta\, ds + \int_0^t g^\eta(t-s)(\eta_s^\eta\, dW_s + \overline{\eta}_s^\eta\, d\overline{W}_s + \widetilde{\eta}_s^\eta\, d\widetilde{W}_s), \\
\overline{\eta}_t^2 &= \overline{\eta}_0^2 + \int_0^t a_s^{\overline{\eta}}\, ds + \int_0^t \int_E \delta_{\overline{\eta}}^l(s,z)\, \pi(ds,dz) + \int_0^t \int_E \delta_{\overline{\eta}}^s(s,z)\, \widetilde{\pi}(ds,dz) \\
&\quad + \int_0^t h^{\overline{\eta}}(t-s)\theta_s^{\overline{\eta}}\, ds + \int_0^t g^{\overline{\eta}}(t-s)(\eta_s^{\overline{\eta}}\, dW_s + \overline{\eta}_s^{\overline{\eta}}\, d\overline{W}_s + \widetilde{\eta}_s^{\overline{\eta}}\, d\widetilde{W}_s + \widehat{\eta}_s^{\overline{\eta}}\, d\widehat{W}_s),
\end{aligned} \tag{B.1}$$

where $(W, \overline{W}, \widetilde{W}, \widehat{W})^T$ forms a four-dimensional standard Brownian motion.

(ii) *The processes* a^η, $a^{\overline{\eta}}$, θ^η, $\theta^{\overline{\eta}}$, η^η, $\overline{\eta}^\eta$, $\widetilde{\eta}^\eta$, $\eta^{\overline{\eta}}$, $\overline{\eta}^{\overline{\eta}}$, $\widetilde{\eta}^{\overline{\eta}}$ *and* $\widehat{\eta}^{\overline{\eta}}$ *are predictable and locally bounded. The variables* η_0 *and* $\overline{\eta}_0$ *are* $\mathcal{F}_0$*-measurable. Both* c *and* η *are locally bounded away from zero. There exists* $\beta > e - e'$ *such that for all* $m, p \in \mathbb{N}$ *we have* $\mathbb{E}[|\theta^{\tau_m}_{s+\Delta} - \theta^{\tau_m}_s|^p]^{1/p} = O(\Delta^\beta)$ *as* $\Delta \to 0$ *uniformly in* $0 \le s \le T$.

(iii) *There exist* $e^\eta, e'^\eta, e^{\overline{\eta}}, e'^{\overline{\eta}} \in (-\frac{1}{2}, 0)$ *and continuously differentiable functions* g^η_0, h^η_0, $g^{\overline{\eta}}_0$ *and* $h^{\overline{\eta}}_0$ *vanishing at the origin such that* $g^\eta = K^{-1}_{e^\eta} t^{e^\eta}_+ + g^\eta_0$, $h^\eta = K^{-1}_{e'^\eta} t^{e'^\eta}_+ + h^\eta_0$, $g^{\overline{\eta}} = K^{-1}_{e^{\overline{\eta}}} t^{e^{\overline{\eta}}}_+ + g^{\overline{\eta}}_0$ *and* $h^{\overline{\eta}} = K^{-1}_{e'^{\overline{\eta}}} t^{e'^{\overline{\eta}}}_+ + h^{\overline{\eta}}_0$. *In addition, we have* $e^\eta \ge \frac{e}{2} - \frac{1}{4}$.

(iv) $\delta^s_\eta, \delta^l_\eta, \delta^s_{\overline{\eta}}, \delta^l_{\overline{\eta}} : \Omega \times \mathbb{R}_+ \times E \to \mathbb{R}$ *are predictable and satisfy* $(|\delta^s_\eta(t,z)|^2 + |\delta^l_\eta(t,z)| + |\delta^s_{\overline{\eta}}(t,z)|^2 + |\delta^l_{\overline{\eta}}(t,z)|) \wedge 1 \le \delta^*_n(z)$ *for all* $n \in \mathbb{N}$, $t \le \tau_n$ *and* $z \in E$.

Condition (i) (together with (iii) and (iv)) assumes that the volatility-of-volatility processes η_t^2 and $\overline{\eta}^2$ have same structural form as c in Assumption V2, with a moving-average representation for their continuous parts. Condition (ii) imposes a minimal Hölder regularity assumption on the nondifferentiable drift coefficient θ.

Finally, we specify the constants in (3.12) and (3.15). Define the central difference operator $\delta^n_h \varphi(x) = \sum_{k=0}^n (-1)^k \binom{n}{k} \varphi(x + (\frac{n}{2} - k)h)$ where $n \in \mathbb{N}$, φ is a real-valued function and $x, h \in \mathbb{R}$. We further define $\delta^n_{h_1,h_2}\varphi(x) = \delta^{n/2}_{h_1}[\delta^{n/2}_{h_2}\varphi](x)$ if n is even. Writing $H_\alpha(x) = |x|^\alpha$, we have

$$\begin{aligned}\Phi_{\kappa,e}(\ell) &= \mathrm{Cov}\left(\kappa \int_0^{\kappa^{-1}} (B_{1+s} - B_s)\, ds, \kappa \int_0^{\kappa^{-1}} (B_{1+s+\ell} - B_{s+\ell})\, ds\right) \\ &= \frac{\kappa^2}{8(e+1)(e+\frac{3}{2})} \delta^4_{1,1/\kappa} H_{2e+3}(\ell)\end{aligned} \tag{B.2}$$

and

$$\mathcal{R}_{\kappa,e}(\ell,\ell') = \begin{cases} R_{\kappa,e}(\ell,\ell') & \text{if } \ell,\ell' \ge 2, \\ R_{\kappa,e}(0,\ell') + 2R_{\kappa,e}(1,\ell') & \text{if } \ell = 1 \text{ and } \ell' \ge 2, \\ R_{\kappa,e}(\ell,0) + 2R_{\kappa,e}(\ell,1) & \text{if } \ell \ge 2 \text{ and } \ell' = 1, \\ R_{\kappa,e}(0,0) + 4R_{\kappa,e}(0,1) + 4R_{\kappa,e}(1,1) & \text{if } \ell = \ell' = 1, \end{cases} \tag{B.3}$$

where[23]

$$
\begin{aligned}
R_{\kappa,e}(\ell,\ell') = \int_{\mathbb{R}} \mathrm{Cov}\Bigg(& \kappa^2 \int_0^{\kappa^{-1}} (B_{1+s} - B_s)\, ds \int_0^{\kappa^{-1}} (B_{1+s+\ell} - B_{s+\ell})\, ds, \\
& \kappa^2 \int_0^{\kappa^{-1}} (B_{1+s+u} - B_{s+u})\, ds \int_0^{\kappa^{-1}} (B_{1+s+u+\ell'} - B_{s+u+\ell'})\, ds \Bigg)\, du \qquad \text{(B.4)} \\
= & \frac{\kappa^4 \Gamma(2e+2)^2 (1 + 1/\cos(2\pi e))}{4\Gamma(4e+8)} (\delta^8_{1,1/\kappa} H_{4e+7}(\ell+\ell') + \delta^8_{1,1/\kappa} H_{4e+7}(\ell-\ell')).
\end{aligned}
$$

While $\Phi_{\kappa,e}(\ell)$ is the autocovariance at lag ℓ of the averages of fractional Brownian motion increments, $R_{\kappa,e}(\ell,\ell')$ is related to the covariance of products of such averages.

References

Aït-Sahalia, Yacine, Jianqing Fan, Roger J. Laeven, Christina Dan Wang, and Xiye Yang (2017): "Estimation of the Continuous and Discontinuous Leverage Effects," *Journal of the American Statistical Association*, 112 (520), 1744–1758. 4

Aït-Sahalia, Yacine, Jianqing Fan, and Yingying Li (2013): "The Leverage Effect Puzzle: Disentangling Sources of Bias at High Frequency," *Journal of Financial Economics*, 109 (1), 224–249. 4

Aït-Sahalia, Yacine and Jean Jacod (2014): *High-Frequency Financial Econometrics*, Princeton: Princeton University Press. 4, 13, 15

Aït-Sahalia, Yacine and Dacheng Xiu (2019): "A Hausman Test for the Presence of Market Microstructure Noise in High Frequency Data," *Journal of Econometrics*, 211 (1), 176–205. 15, 18

Aït-Sahalia, Yacine and Jialin Yu (2009): "High Frequency Market Microstructure Noise Estimates and Liquidity Measures," *The Annals of Applied Statistics*, 3 (1), 422–457. 16, 18

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens

[23] The expression in (B.4) is formally defined only for $e \neq -\frac{1}{4}$. However, it admits a continuous extension at $e = -1/4$. Using L'Hôpital's rule, one can find that $R_{\kappa,-1/4}(\ell,\ell') = \frac{1}{5760}\kappa^4(\delta^8_{1,1/\kappa}\mathbb{H}_6(\ell+\ell') + \delta^8_{1,1/\kappa}\mathbb{H}_6(\ell-\ell'))$ where $\mathbb{H}_\alpha(x) = |x|^\alpha \log|x|$ for $x \neq 0$ and $\mathbb{H}_\alpha(0) = 0$.

(2001): "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61 (1), 43–76. 4, 5, 12

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71 (2), 579–625. 4, 5, 12, 30

Andersen, Torben G., Tim Bollerslev, and Nour Meddahi (2005): "Correcting the Errors: Volatility Forecast Evaluation Using High-Frequency Data and Realized Volatilities," *Econometrica*, 73 (1), 279–296. 29, 30

Andersen, Torben G., Tao Su, Viktor Todorov, and Zhiyuan Zhang (2024): "Intraday Periodic Volatility Curves," *Journal of the American Statistical Association*, 119 (546), 1181–1191. 4

Andrews, Isaiah, Jiafeng Chen, and Otavio Tecchio (2026): "The Purpose of an Estimator is What It Does: Misspecification, Estimands, and Over-Identification," *Preprint*. 15

Baillie, Richard T. and Tim Bollerslev (1991): "Intra-Day and Inter-Market Volatility in Foreign Exchange Rates," *The Review of Economic Studies*, 58 (3), 565–585. 4

Bandi, Federico M. and Roberto Renò (2012): "Time-Varying Leverage Effects," *Journal of Econometrics*, 169 (1), 94–113. 4

——— (2016): "Price and Volatility Co-Jumps," *Journal of Financial Economics*, 119 (1), 107–146. 4

Bandi, Federico M. and Jeffrey R. Russell (2006): "Separating Microstructure Noise from Volatility," *Journal of Financial Economics*, 79 (3), 655–692. 14

——— (2008): "Microstructure Noise, Realized Variance, and Optimal Sampling," *The Review of Economic Studies*, 75 (2), 339–369. 14

Barndorff-Nielsen, Ole E., Peter Reinhard Hansen, Asger Lunde, and Neil Shephard (2008): "Designing Realized Kernels to Measure the ex post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76 (6), 1481–1536. 14

Barndorff-Nielsen, Ole E. and Neil Shephard (2002): "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 64 (2), 253–280. 4

——— (2004a): "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics," *Econometrica*, 72 (3), 885–925. 14

——— (2004b): "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2 (1), 1–37. 25

Bennedsen, Mikkel, Asger Lunde, and Mikko S. Pakkanen (2022): "Decoupling the Short- and Long-Term Behavior of Stochastic Volatility," *Journal of Financial Econometrics*, 20 (5), 961–1006. 4, 5, 19, 30

Bingham, N. H., C. M. Goldie, and J. L. Teugels (1987): *Regular Variation*, vol. 27 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge. 34, 35, 50

Bolko, Anine E., Kim Christensen, Mikko S. Pakkanen, and Bezirgen Veliyev (2023): "A GMM Approach to Estimate the Roughness of Stochastic Volatility," *Journal of Econometrics*, 235 (2), 745–778. 4, 28

Bollerslev, Tim (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31 (3), 307–327. 3

Bollerslev, Tim, Jia Li, and Zhipeng Liao (2021): "Fixed-k Inference for Volatility," *Quantitative Economics*, 12 (4), 1053–1084. 4, 32

Bollerslev, Tim, Jia Li, and Yuexuan Ren (2024): "Optimal Inference for Spot Regressions," *American Economic Review*, 114 (3), 678–708. 4

Bollerslev, Tim, Jia Li, and Yuan Xue (2018): "Volume, Volatility, and Public News Announcements," *The Review of Economic Studies*, 85 (4), 2005–2041. 4, 16

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik (2014): "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 82 (6), 2295–2326. 3

Chen, Xiaohong, Han Hong, and Elie Tamer (2005): "Measurement Error Models with Auxiliary Data," *The Review of Economic Studies*, 72 (2), 343–366. 5

Chernov, Mikhail, A. Ronald Gallant, Eric Ghysels, and George Tauchen (2003): "Alternative Models for Stock Price Dynamics," *Journal of Econometrics*, 116 (1), 225–257. 10

Chesher, Andrew (1991): "The Effect of Measurement Error," *Biometrika*, 78 (3), 451–462. 5

Chong, Carsten H., Marc Hoffmann, Yanghui Liu, Mathieu Rosenbaum, and Grégoire Szymanski (2024): "Statistical Inference for Rough Volatility: Central Limit Theorems," *The Annals of Applied Probability*, 34 (3), 2600–2649. 4, 64, 65, 69

Chong, Carsten H. and Viktor Todorov (2024): "Volatility of Volatility and Leverage Effect from Options," *Journal of Econometrics*, 240 (1), Paper No. 105669, 21. 4

——— (2025): "Nonparametric Test for Rough Volatility," *Journal of the American Statistical Association*, 120 (552), 2772–2783. 4, 8, 17, 20, 25, 36

Clark, Peter K. (1973): "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41 (1), 135–155. 3

Comte, Fabienne and Eric Renault (1996): "Long Memory Continuous Time Models," *Journal of Econometrics*, 73 (1), 101–149. 2, 8, 19

——— (1998): "Long Memory in Continuous-Time Stochastic Volatility Models," *Mathematical Finance*, 8 (4), 291–323. 4

Corcuera, José Manuel, Emil Hedevang, Mikko S. Pakkanen, and Mark Podolskij (2013): "Asymptotic Theory for Brownian Semi-Stationary Processes with Application to Turbulence," *Stochastic Processes and their Applications*, 123 (7), 2552–2574. 20

Corsi, Fulvio (2009): "A Simple Approximate Long-Memory Model of Realized Volatility," *Journal of Financial Econometrics*, 7 (2), 174–196. 30

Da, Rui and Dacheng Xiu (2021): "When Moving-Average Models Meet High-Frequency Data: Uniform Inference on Volatility," *Econometrica*, 89 (6), 2787–2825. 14, 15, 36

Diebold, Francis X. and Roberto S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13 (3), 253–263. 32

Dittmann, Ingolf and Clive W. J. Granger (2002): "Properties of Nonlinear Transformations of Fractionally Integrated Processes," *Journal of Econometrics*, 110 (2), 113–133. 11

Duffie, Darrell, Jun Pan, and Kenneth Singleton (2000): "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68 (6), 1343–1376. 24

El Euch, Omar and Mathieu Rosenbaum (2019): "The Characteristic Function of Rough Heston Models," *Mathematical Finance*, 29 (1), 3–38. 20, 24

Engle, Robert F. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50 (4), 987–1007. 3

Evdokimov, Kirill S. and Andrei Zeleneev (2024): "Nonparametric Identification and Estimation with Non-Classical Errors-in-Variables," *Preprint*. 5

Foster, Dean P. and Dan B. Nelson (1996): "Continuous Record Asymptotics for Rolling Sample Variance Estimators," *Econometrica*, 64 (1), 139–174. 4

Gatheral, Jim, Thibault Jaisson, and Mathieu Rosenbaum (2018): "Volatility Is Rough," *Quantitative Finance*, 18 (6), 933–949. 4, 8, 28

Giacomini, Raffaella and Halbert White (2006): "Tests of Conditional Predictive Ability," *Econometrica*, 74 (6), 1545–1578. 32

Hannan, Edward J. (1970): *Multiple Time Series*, John Wiley & Sons, Inc., New York-London-Sydney. 19

Hansen, Lars Peter (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50 (4), 1029–1054. 3

Hansen, Peter R., Asger Lunde, and James M. Nason (2011): "The Model Confidence Set," *Econometrica*, 79 (2), 453–497. 31

Hausman, Jerry (2001): "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left," *Journal of Economic Perspectives*, 15 (4), 57–67. 5, 21

Hausman, Jerry A., Whitney K. Newey, Hidehiko Ichimura, and James L. Powell (1991): "Identification and Estimation of Polynomial Errors-in-Variables Models," *Journal of Econometrics*, 50 (3), 273–295. 5

Heston, Steven L. (1993): "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *The Review of Financial Studies*, 6 (2), 327–343. 3, 24

Hörmander, Lars (1990): *The Analysis of Linear Partial Differential Operators. I*, vol. 256 of *Grundlehren der mathematischen Wissenschaften*, Springer, Berlin, second ed. 65

Hu, Yingyao and Susanne M. Schennach (2008): "Instrumental Variable Treatment of Nonclassical Measurement Error Models," *Econometrica*, 76 (1), 195–216. 5

Hull, John and Alan White (1987): "The Pricing of Options on Assets with Stochastic Volatilities," *The Journal of Finance*, 42 (2), 281–300. 3

Jacod, Jean, Yingying Li, Per A. Mykland, Mark Podolskij, and Mathias Vetter (2009): "Microstructure Noise in the Continuous Case: The Pre-Averaging Approach," *Stochastic Processes and their Applications*, 119 (7), 2249–2276. 14

Jacod, Jean, Yingying Li, and Xinghua Zheng (2017): "Statistical Properties of Microstructure Noise," *Econometrica*, 85 (4), 1133–1174. 14, 15, 18, 36

Jacod, Jean and Philip Protter (2012): *Discretization of Processes*, vol. 67 of *Stochastic Modelling and Applied Probability*, Springer, Heidelberg. 19, 20, 32, 52, 58, 59, 81

Jacod, Jean and Viktor Todorov (2010): "Do Price and Volatility Jump Together?" *The Annals of Applied Probability*, 20 (4), 1425–1469. 4

——— (2014): "Efficient Estimation of Integrated Volatility in Presence of Infinite Variation Jumps," *The Annals of Statistics*, 42 (3), 1029–1069. 3

Kalnina, Ilze and Dacheng Xiu (2017): "Nonparametric Estimation of the Leverage Effect: A Trade-Off Between Robustness and Efficiency," *Journal of the American Statistical Association*, 112 (517), 384–396. 4

Karamata, Jovan (1930): "Sur un mode de croissance régulière des fonctions," *Mathematica (Cluj)*, 4, 38–53. 2

Laurent, Sébastien, Roberto Renò, and Shuping Shi (2024): "Realized Drift," *Journal of Econometrics*, 105813. 9, 14

Li, Jia, Yifan Li, Andrew J. Patton, and Tian Xie (2025a): "Online Forecast Evaluation with High-Frequency Proxies," *SSRN Preprint*. 30

Li, Jia, Peter C. B. Phillips, Shuping Shi, and Jun Yu (2025b): "Weak Identification of Long Memory with Implications for Inference," *The Review of Financial Studies*, 38 (10), 3117–3148. 5, 12

Li, Jia and Dacheng Xiu (2016): "Generalized Method of Integrated Moments for High-Frequency Data," *Econometrica*, 84 (4), 1613–1633. 22

Li, Yingying, Guangying Liu, and Zhiyuan Zhang (2022): "Volatility of Volatility: Estimation and Tests Based on Noisy High Frequency Data with Jumps," *Journal of Econometrics*, 229 (2), 422–451. 4

Li, Yingying, Shangyu Xie, and Xinghua Zheng (2016): "Efficient Estimation of Integrated Volatility Incorporating Trading Information," *Journal of Econometrics*, 195 (1), 33–50. 16

Li, Z. Merrick and Oliver Linton (2022): "A ReMeDI for Microstructure Noise," *Econometrica*, 90 (1), 367–389. 14, 15, 18, 25, 36

Li, Z. Merrick and Xiye Yang (2026): "Multi-Horizon Test for Market Frictions," *Journal of Econometrics*, 253, 106171. 14, 15, 18

Lo, Andrew W. (1991): "Long-Term Memory in Stock Market Prices," *Econometrica*, 59 (5), 1279–1313. 8

Mandelbrot, Benoit B. and John W. Van Ness (1968): "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Review*, 10, 422–437. 2

Mies, Fabian and Mark Podolskij (2023): "Estimation of Mixed Fractional Stable Processes Using High-Frequency Data," *The Annals of Statistics*, 51 (5), 1946–1964. 72, 73

Muhle-Karbe, Johannes, Youssef Ouazzani Chahdi, Mathieu Rosenbaum, and Grégoire Szymanski (2026): "A Unified Theory of Order Flow, Market Impact, and Volatility," *arXiv*. 28

Mykland, Per A. and Lan Zhang (2009): "Inference for Continuous Semimartingales Observed at High Frequency," *Econometrica*, 77 (5), 1403–1445. 32

Müller, Ulrich K. and Mark W. Watson (2008): "Testing Models of Low-Frequency Variability," *Econometrica*, 76 (5), 979–1016. 2, 5, 8, 12

Nualart, David (2006): *The Malliavin Calculus and Related Topics*, Springer, Berlin, second ed. 83, 85

Račkauskas, Alfredas and Charles Suquet (2005): "Central Limit Theorems in Hölder Topologies for Banach Space Valued Random Fields," *Theory of Probability & Its Applications*, 49 (1), 77–92. 56

Schennach, Susanne M. (2004): "Estimation of Nonlinear Models with Measurement Error," *Econometrica*, 72 (1), 33–75. 5

——— (2007): "Instrumental Variable Estimation of Nonlinear Errors-in-Variables Models," *Econometrica*, 75 (1), 201–239. 5

Shi, Shuping and Jun Yu (2023): "Volatility Puzzle: Long Memory or Antipersistency," *Management Science*, 69 (7), 3861–3883. 5, 12

Shi, Shuping, Jun Yu, and Chen Zhang (2024): "On the Spectral Density of Fractional Ornstein-Uhlenbeck Processes," *Journal of Econometrics*, 245 (1-2), Paper No. 105872, 28. 12

Sowell, Fallaw (1990): "The Fractional Unit Root Distribution," *Econometrica*, 58 (2), 495–505. 2, 8

Tauchen, George E. and Mark Pitts (1983): "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica*, 51 (2), 485–505. 3

Todorov, Viktor and George Tauchen (2011): "Volatility Jumps," *Journal of Business & Economic Statistics*, 29 (3), 356–371. 4

Wang, Christina D. and Per A. Mykland (2014): "The Estimation of Leverage Effect with High-Frequency Data," *Journal of the American Statistical Association*, 109 (505), 197–215. 4

Wang, Xiaohu, Weilin Xiao, and Jun Yu (2023): "Modeling and Forecasting Realized Volatility with the Fractional Ornstein–Uhlenbeck Process," *Journal of Econometrics*, 232 (2), 389–415. 4, 28, 30

West, Kenneth D. (1996): "Asymptotic Inference about Predictive Ability," *Econometrica*, 64 (5), 1067–1084. 32

Wood, Robert A., Thomas H. McInish, and J. Keith Ord (1985): "An Investigation of Transactions Data for NYSE Stocks," *The Journal of Finance*, 40 (3), 723–739. 4

Yokoyama, Ryozo (1980): "Moment Bounds for Stationary Mixing Sequences," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 52 (1), 45–57. 59

Zhang, Lan, Per A. Mykland, and Yacine Aït-Sahalia (2005): "A Tale of Two Tme Scales: Determining Integrated Volatility with Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100 (472), 1394–1411. 14

Internet Appendix for
"Supplement to "Intraday Volatility Dynamics"'

C Proof of Propositions 1–5

For a continuous-time process X, we write $\Delta_t X = X_{t+\Delta} - X_t$.

Proof of Proposition 1. Consider the process Y. By the linearity of covariance, $\gamma_{t+\theta\Delta}(\tau;\Delta) = \int_{t+\theta\Delta}^{t+(\theta+1)\Delta}\int_{t+(\theta+\tau)\Delta}^{t+(\theta+\tau+1)\Delta}\mathrm{Cov}(b_s,b_r)\,dr\,ds$. Because b is mean-square continuous, we have $\mathrm{Cov}(b_s,b_r) = \mathrm{Cov}(b_t,b_t)+o(1) = \mathrm{Var}(b_t)+o(1)$, with $\mathrm{Var}(b_t)>0$, uniformly in $r,s\in[t,t+(\theta+\tau+1)\Delta]$. This shows that $\gamma_{t+\theta\Delta}(\tau;\Delta) = \mathrm{Var}(b_t)\Delta^2+o(\Delta^2)$ and therefore $\gamma_{t+\theta\Delta}(\tau;\Delta)/\gamma_t(\Delta)\to 1$ and $(\gamma_t(\lambda\Delta)-\gamma_t(\Delta))/\gamma_t(\Delta)\to\lambda^2-1$ as $\Delta\to 0$, for any $\theta,\tau\geq 0$ and $\lambda>1$. This shows Assumption Γ with $\rho_t(\tau)=\Psi_{1/2}(\tau)=1$. For X_t, we apply Proposition 3 with $\alpha=\beta=0$. □

Proof of Proposition 2. We show in Section D.1 that g_0 only introduces a negligible drift term. We can thus assume $g_0\equiv 0$ and $g\equiv g_e$ without loss of generality. Writing $\Delta_t g_e(s) = g_e(t+\Delta-s)-g_e(t-s)$ and using the fact that $g_e(s)=0$ for $s\leq 0$, we have $\gamma_{\theta\Delta}(\tau;\Delta) = \int_0^\infty \Delta_{t+\theta\Delta}g_e(s)\Delta_{t+(\theta+\tau)\Delta}g_e(s)\mathbb{E}[\sigma_s^2]\,ds$. A change of variables $(t+\theta\Delta-s = r\Delta)$ yields $\gamma_{\theta\Delta}(\tau;\Delta) = [\int_{-\infty}^{t/\Delta+\theta}(g_e(r+1)-g_e(r))(g_e(r+\tau+1)-g_e(r+\tau))\mathbb{E}[\sigma^2_{t+(\theta-r)\Delta}]\,dr]\Delta^{2e+1}$. Because $(g_e(r+1)-g_e(r))(g_e(r+\tau+1)-g_e(r+\tau))\sim r^{2e-2}$, which is integrable, we can truncate the tail of the integral at $+\infty$. We obtain $\gamma_{\theta\Delta}(\tau;\Delta)\sim[\int_{-\infty}^{1/\sqrt{\Delta}}(g_e(r+1)-g_e(r))(g_e(r+\tau+1)-g_e(r+\tau))\mathbb{E}[\sigma^2_{t+(\theta-r)\Delta}]\,dr]\Delta^{2e+1}$. Note that the integrand is nonzero only on $r\in[-1,1/\sqrt{\Delta}]$. On this range, we have $\mathbb{E}[\sigma^2_{t+(\theta-r)\Delta}]\sim\mathbb{E}[\sigma_t^2]+o(1)$, uniformly in r. As a result, $\gamma_{\theta\Delta}(\tau;\Delta)\sim[\int_{-\infty}^{1/\sqrt{\Delta}}(g_e(r+1)-g_e(r))(g_e(r+\tau+1)-g_e(r+\tau))\,dr]\mathbb{E}[\sigma_t^2]\Delta^{2e+1}\sim\int_{\mathbb{R}}(g_e(r+1)-g_e(r))(g_e(r+\tau+1)-g_e(r+\tau))\,dr\,\mathbb{E}[\sigma_t^2]\Delta^{2e+1}$. Since the last integral is the lag-τ autocovariance of fractional Gaussian noise, we obtain $\gamma_{\theta\Delta}(\tau;\Delta)\sim\Psi_e(\tau)\mathbb{E}[\sigma_t^2]\Delta^{2e+1}$. From here, we readily deduce (2.2) and (2.3) with $\rho_t(\tau)=\Psi_e(\tau)$. □

Proof of Proposition 3. Note that X is a continuous Itô semimartingale with mean-square continuous coefficients on $[t,\infty)$, for any $t>0$. If $\alpha=\beta=0$, this extends to $t=0$. Therefore, the proof of the proposition for $t>0$ is exactly the same as for $t=0$ in the case $\alpha=\beta=0$. We thus focus on $t=0$ in the following. Using the linearity of covariance and Itô's isometry, we have $\gamma_{\theta\Delta}(\tau;\Delta) = \mathrm{Cov}(X_{(\theta+1)\Delta}-X_{\theta\Delta}, X_{(\theta+\tau+1)\Delta}-X_{(\theta+\tau)\Delta}) = \sum_{i=1}^4 A_\theta^i(\tau;\Delta)$ where $A_\theta^1(\tau;\Delta) = \int_{\theta\Delta}^{(\theta+1)\Delta}\int_{(\theta+\tau)\Delta}^{(\theta+\tau+1)\Delta} s^{-\alpha}r^{-\alpha}\,\mathrm{Cov}(b_s,b_r)\,dr\,ds\sim\mathrm{Var}(b_0)(1-\alpha)^{-2}[(\theta+1)^{1-\alpha}-\theta^{1-\alpha}][(\theta+\tau+1)^{1-\alpha}-(\theta+\tau)^{1-\alpha}]\Delta^{2-2\alpha}$, $A_\theta^2(\tau;\Delta) = \mathbb{E}[\int_{\theta\Delta}^{(\theta+1)\Delta}s^{-\beta}\sigma_s\,dW_s\int_{(\theta+\tau)\Delta}^{(\theta+\tau+1)\Delta}r^{-\beta}\sigma_r\,dW_r] =$

$\int_{(\theta+\tau)\Delta}^{(\theta+1)\Delta} s^{-2\beta}\mathbb{E}[\sigma_s^2]\,ds\mathbf{1}_{\tau<1} \sim \mathbb{E}[\sigma_0^2](1-2\beta)^{-1}[(\theta+1)^{1-2\beta}-(\theta+\tau)^{1-2\beta}]_+\Delta^{1-2\beta}$ and, by the Cauchy–Schwarz inequality, $A_\theta^3(\tau;\Delta) = \mathbb{E}[\int_{\theta\Delta}^{(\theta+1)\Delta} s^{-\alpha}b_s ds \int_{(\theta+\tau)\Delta}^{(\theta+\tau+1)\Delta} r^{-\beta}\sigma_r dW_r] = \mathbb{E}[\int_{\theta\Delta}^{(\theta+1)\Delta} s^{-\alpha}[b_s-b_0]ds \int_{(\theta+\tau)\Delta}^{(\theta+\tau+1)\Delta} r^{-\beta}\sigma_r dW_r] = o(\Delta^{3/2-\alpha-\beta})$ and similarly, $A_\theta^4(\tau;\Delta) = \mathbb{E}[\int_{\theta\Delta}^{(\theta+1)\Delta} s^{-\beta}\sigma_s\,dW_s \int_{(\theta+\tau)\Delta}^{(\theta+\tau+1)\Delta} r^{-\alpha}b_r\,dr] = o(\Delta^{3/2-\alpha-\beta})$. This shows that $\gamma_{\theta\Delta}(\tau;\Delta) \sim A_\theta^1(\tau;\Delta)+A_\theta^2(\tau;\Delta)$.

If $\alpha < \beta+\frac{1}{2}$, we have $\gamma_{\theta\Delta}(\tau;\Delta) \sim A_\theta^2(\tau;\Delta) \sim \frac{\mathbb{E}[\sigma_0^2]}{1-2\beta}[(\theta+1)^{1-2\beta}-(\theta+\tau)^{1-2\beta}]_+\Delta^{1-2\beta}$ and therefore, $\gamma_{\theta\Delta}(\tau;\Delta)/\gamma_0(\Delta) \to [(\theta+1)^{1-2\beta}-(\theta+\tau)^{1-2\beta}]_+$, which is independent of θ (as required by Assumption Γ (i)) if and only if $\beta=0$. Nonetheless, we always have $r_0(\tau) = \lim_{\Delta\to0}\gamma_0(\tau;\Delta)/\sqrt{\gamma_0(\Delta)\gamma_{\tau\Delta}(\Delta)} = (1-\tau^{1-2\beta})_+/\sqrt{(\tau+1)^{1-2\beta}-\tau^{1-2\beta}}$. In the case $\beta=0$, we also have $\gamma_0(\lambda\Delta)/\gamma_0(\Delta)-1 \to \lambda^{1-2\beta}-1$ for all $\lambda>1$, which proves (2.3) in this case.

If $\alpha > \beta+\frac{1}{2}$, we have $\gamma_{\theta\Delta}(\tau;\Delta) \sim A_\theta^1(\tau;\Delta) \sim \frac{\mathrm{Var}(b_0)}{(1-\alpha)^2}[(\theta+1)^{1-\alpha}-\theta^{1-\alpha}][(\theta+\tau+1)^{1-\alpha}-(\theta+\tau)^{1-\alpha}]\Delta^{2-2\alpha}$. Therefore, $\gamma_{\theta\Delta}(\tau;\Delta)/\gamma_0(\Delta) \to [(\theta+1)^{1-\alpha}-\theta^{1-\alpha}][(\theta+\tau+1)^{1-\alpha}-(\theta+\tau)^{1-\alpha}]$. As this depends on θ, Assumption Γ is not satisfied at $t=0$. But we can still compute the asymptotic autocorrelation of increments: we have $r_0(\tau) = \lim_{\Delta\to0}\gamma_0(\tau;\Delta)/\sqrt{\gamma_0(\Delta)\gamma_{\tau\Delta}(\Delta)} = [(\tau+1)^{1-\alpha}-\tau^{1-\alpha}]/\sqrt{[(\tau+1)^{1-\alpha}-\tau^{1-\alpha}]^2} = 1$.

If $\alpha = \beta+\frac{1}{2}$, a similar calculation shows that the limit of $\gamma_{\theta\Delta}(\tau;\Delta)/\gamma_0(\Delta)$ depends on θ, which is why Assumption Γ fails. Since the derivation of $r_0(\tau)$ only requires combining the two cases analyzed above, we omit the details. □

Proof of Proposition 4. Let $Y = Z - X = \sum_{i=1}^\infty \rho_i X^i$. Since $X_s^i = e^{-\lambda_i(s-r)}X_r^i + \int_r^s e^{-\lambda_i(s-v)}\sigma_v^i\,dW_v^i$ for $i\ge1$ and $r\le s$, Itô's isometry, the independence properties of W^i and (X_0^i) and the assumption $q>1-2r$ imply that $\mathbb{E}[Y_t^2] = \sum_{i=1}^\infty \rho_i^2(\mathbb{E}[(X_0^i)^2]e^{-2\lambda_i t} + \int_0^t e^{-2\lambda_i(t-r)}\mathbb{E}[(\sigma_r^i)^2]\,dr) \le C\sum_{i=1}^\infty i^{-2r-q} < \infty$, which shows that Y_t is well defined.

We next prove that $2\lambda_i\mathbb{E}[(X_{t+\theta\Delta}^i)^2] = v_t + o(1)$ as $i\to\infty$ and $\Delta\to0$, for all $t\ge0$. For $t=0$, the continuity of v and the conditions $\mathbb{E}[(X_0^i)^2] \sim v_0/(2\lambda_i)$ and $\mathbb{E}[(\sigma_s^i)^2] \sim v_s$ yield $2\lambda_i\mathbb{E}[(X_{\theta\Delta}^i)^2] - v_0 = e^{-2\lambda_i\theta\Delta}(2\lambda_i\mathbb{E}[(X_0^i)^2]-v_0) + \int_0^{\theta\Delta} 2\lambda_i e^{-2\lambda_i(\theta\Delta-s)}(\mathbb{E}[(\sigma_s^i)^2]-v_0)\,ds = o(1)$. For $t>0$, a change of variables shows that $2\lambda_i\mathbb{E}[(X_{t+\theta\Delta}^i)^2] = e^{-2\lambda_i(t+\theta\Delta)}2\lambda_i\mathbb{E}[(X_0^i)^2] + \int_0^\infty e^{-r}\mathbb{E}[(\sigma_{t+\theta\Delta-r/(2\lambda_i)}^i)^2]\mathbf{1}_{\{r\le2\lambda_i(t+\theta\Delta)\}}\,dr$. As $i\to\infty$ and $\Delta\to0$, the first term converges to 0, because $t>0$. In the second term, the integrand converges pointwise to $e^{-r}v_t$ for all $r>0$. By dominated convergence, we obtain $2\lambda_i\mathbb{E}[(X_{t+\theta\Delta}^i)^2] \to v_t$, as claimed.

For any $t,\theta\ge0$ and $u>0$, we thus have $\sum_{i=1}^\infty \rho_i^2(1-e^{-\lambda_i u\Delta})^2(\mathbb{E}[(X_{t+\theta\Delta}^i)^2] - \frac{v_t}{2\lambda_i}) = \frac{1}{2}\sum_{i=1}^\infty \frac{\rho_i^2}{\lambda_i}(1-e^{-\lambda_i u\Delta})^2(2\lambda_i\mathbb{E}[(X_{t+\theta\Delta}^i)^2]-v_t) = o(\sum_{i=1}^\infty \frac{\rho_i^2}{\lambda_i}(1-e^{-\lambda_i u\Delta})) = o(\Delta^\alpha)$, where $\alpha =$

$2e+1$ and the last step follows from Karamata's Tauberian theorem for Laplace transforms.[24] Similarly, we have $\sum_{i=1}^\infty \rho_i^2 \int_{t+\theta\Delta}^{t+(\theta+u)\Delta} e^{-2\lambda_i(t+(\theta+u)\Delta-s)}(\mathbb{E}[(\sigma_s^i)^2]-v_t)ds = o(\sum_{i=1}^\infty \frac{\rho_i^2}{\lambda_i}(1-e^{-2u\Delta}))$ $= o(\Delta^\alpha)$.

The previous calculations imply that $\gamma^Y_{t+\theta\Delta}(u\Delta) = \mathrm{Cov}(Y_{t+(\theta+u)\Delta} - Y_{t+\theta\Delta}, Y_{t+(\theta+\tau+u)\Delta} - Y_{t+(\theta+\tau)\Delta}) = \sum_{i=1}^\infty \rho_i^2 \mathrm{Var}(X^i_{t+(\theta+u)\Delta} - X^i_{t+\theta\Delta}) = \sum_{i=1}^\infty \rho_i^2 \big[(1-e^{-\lambda_i u\Delta})^2\mathbb{E}[(X^i_{t+\theta\Delta})^2] + \int_{t+\theta\Delta}^{t+(\theta+u)\Delta} e^{-2\lambda_i(t+(\theta+u)\Delta-s)}\mathbb{E}[(\sigma_s^i)^2]\,ds\big] \sim \frac{1}{2}v_t\sum_{i=1}^\infty \frac{\rho_i^2}{\lambda_i}[(1-e^{-\lambda_i u\Delta})^2 + (1-e^{-2\lambda_i u\Delta})] = v_t\sum_{i=1}^\infty \frac{\rho_i^2}{\lambda_i}(1-e^{-\lambda_i u\Delta}) \sim \frac{R^2L^{\alpha-1}u^\alpha}{q\alpha}\Gamma(1-\alpha)\Delta^\alpha$. Since $\mathrm{Var}(X_{t+u\Delta}-X_t) = O(\Delta)$, we conclude that $\gamma^Z_{t+\theta\Delta}(u\Delta) \sim \gamma^Y_{t+\theta\Delta}(u\Delta)$ and therefore, $\gamma^Z_{t+\theta\Delta}(u\Delta)/\gamma^Z_t(\Delta) \to u^\alpha = u^{2e+1}$ for all $t,\theta \geq 0$ and $u > 0$. The arguments after (A.3) prove (2.2) with $\rho_t(\tau) = \Psi_e(\tau)$. Since $(\gamma^Z_t(u\Delta) - \gamma^Z_t(\Delta))/\gamma^Z_t(\Delta) \to u^\alpha - 1$, we also have (2.3). □

Proof of Proposition 5. As η_t is bounded on $[0,\tau_n]$, the process X^{τ_n} is mean-square continuous for all $n \in \mathbb{N}$. As we show in Section D.1, the component g_0 of g only introduces a drift term, which is asymptotically negligible. We can thus assume $g_0 \equiv 0$ and $g \equiv g_e$ for simplicity. Write $t_n = t \wedge \tau_n$, $\Delta_t X = X_{t+\Delta} - X_t$ and $\Delta_t^{(n)} g_e(s) = g_e((t+\Delta)_n - s) - g_e(t_n - s)$. By Itô's isometry, $\gamma_t^{(n)}(k;\Delta) = \mathbb{E}[\int_0^{(t+\Delta)_n} \Delta_t^{(n)} g_e(s)\Delta^{(n)}_{t+k\Delta} g_e(s)\eta_s^2\,ds] = I_1^n + I_2^n$, where $I_1^n = \mathbb{E}[\int_0^{(t+\Delta)_n} \Delta_t^{(n)} g_e(s)\Delta^{(n)}_{t+k\Delta} g_e(s)\eta_s^2\,ds\,\mathbf{1}_{t+(k+1)\Delta\leq\tau_n}]$ and $I_2^n = \mathbb{E}[\int_0^{(t+\Delta)_n} \Delta_t^{(n)} g_e(s) \times \Delta^{(n)}_{t+k\Delta} g_e(s)\eta_s^2\,ds\,\mathbf{1}_{t+k\Delta<\tau_n<t+(k+1)\Delta}]$.

Note that $|\Delta_t^{(n)} g_e(s)| \leq |\Delta_{t_n} g_e(s)|$ where $\Delta_t g_e(s) = g_e(t+\Delta-s) - g_e(t-s)$. Moreover, for any $t > 0$, we have $\int_{\mathbb{R}} \Delta_t g_e(s)^2\,ds = \Delta^{2e+1}$ (see the discussion regarding K_e after (3.14)). Hence, $|I_2^n| \leq n^2(\int_{\mathbb{R}} \Delta_t^{(n)} g_e(s)^2\,ds\ \int_{\mathbb{R}} \Delta^{(n)}_{t+k\Delta} g_e(s)^2\,ds)^{1/2}\mathbb{P}(\tau_n - t \in (k\Delta,(k+1)\Delta)) \leq n^2\Delta^{2e+1}\mathbb{P}(\tau_n - t \in (k\Delta,(k+1)\Delta))$. As τ_n increases to infinity, we have that $\mathbb{P}(\tau_n - t \in (k\Delta,(k+1)\Delta)) = o(1)$ as $\Delta \to 0$, uniformly in n. As a result, $|I_2^n| \leq n^2\Delta^{2e+1}w(\Delta)$ where w is a function which does not depend on n and is continuous and equal to zero at the origin.

Next, we decompose $I_1^n = \int_0^{t+\Delta} \Delta_t g_e(s)\Delta_{t+k\Delta} g_e(s)\mathbb{E}[\eta_s^2\mathbf{1}_{t+(k+1)\Delta\leq\tau_n}]\,ds = I_{11}^n + I_{12}^n$, where I_{11}^n and I_{12}^n are the restrictions of the integral to $[0, t-\sqrt{\Delta}]$ and $[t-\sqrt{\Delta}, t+\Delta]$, respectively. We can bound $|I_{11}^n| \leq n^2(\int_0^{t-\sqrt{\Delta}} \Delta_t g_e(s)^2\,ds\ \int_0^{t-\sqrt{\Delta}} \Delta_{t+k\Delta} g_e(s)^2\,ds)^{1/2}$. We have $\int_0^{t-\sqrt{\Delta}} \Delta_t g_e(s)^2\,ds = \int_{\sqrt{\Delta}+\Delta}^{t+\Delta} (g_e(s) - g_e(s-\Delta))^2\,ds = \Delta^{2e+1}\int_{1+1/\sqrt{\Delta}}^{1+t/\Delta} (g_e(r) - g_e(r-1))^2\,dr \leq \Delta^{2e+1}\int_{1+1/\sqrt{\Delta}}^{\infty} (g_e(r) - g_e(r-1))^2\,dr = O(\Delta^{e+3/2}) = o(\Delta^{2e+1})$. The same bound applies to

[24] Define the measure $\mu = \sum_{i=1}^\infty \frac{\rho_i^2}{\lambda_i}\delta_{\lambda_i}$. Then $\mu((x,\infty)) = \sum_{\lambda_i > x} \frac{\rho_i^2}{\lambda_i} \sim \frac{R^2}{L}\sum_{i\geq(x/L)^{1/q}} i^{-2r-q} \sim \frac{R^2/L}{2r+q-1}(x/L)^{(1-2r-q)/q} = \frac{R^2L^{\alpha-1}}{q\alpha}x^{-\alpha}$ and $\int_0^x \mu((y,\infty))\,dy \sim \frac{R^2L^{\alpha-1}}{q\alpha(1-\alpha)}x^{1-\alpha}$ as $x \to \infty$. Theorem 1.7.1 in Bingham et al. (1987) implies that $\int_0^\infty e^{-sx}\mu((x,\infty))\,dx \sim \frac{R^2L^{\alpha-1}}{q\alpha(1-\alpha)}\Gamma(2-\alpha)s^{\alpha-1} = \frac{R^2L^{\alpha-1}}{q\alpha}\Gamma(1-\alpha)s^{\alpha-1}$ as $s \to 0$. Therefore, $\sum_{i=1}^\infty \frac{\rho_i^2}{\lambda_i}(1-e^{-\lambda_i u\Delta}) = \int_0^\infty (1-e^{-u\Delta y})\,\mu(dy) = u\Delta\int_0^\infty \int_0^y e^{-u\Delta x}\,dx\,\mu(dy) = u\Delta\int_0^\infty e^{-u\Delta x}\mu((x,\infty))\,dx \sim \frac{R^2L^{\alpha-1}u^\alpha}{q\alpha}\Gamma(1-\alpha)\Delta^\alpha$ as $\Delta \to 0$.

$\int_0^{t-\sqrt{\Delta}} \Delta_{t+k\Delta} g_e(s)^2\, ds$, whence $|I_{11}^n| \leq Cn^2\Delta^{2e+1}w(\Delta)$ for a suitable w as before.

Since $\eta = \eta^{\tau_n}$ on $\{t+(k+1)\Delta \leq \tau_n\}$ and η^{τ_n} is mean-square continuous (as η is continuous in probability), there exists w_n as before (which may now depend on n) such that

$$\begin{aligned} I_{12}^n &= \mathbb{E}[\eta_t^2 \mathbf{1}_{t+(k+1)\Delta\leq\tau_n}] \int_{t-\sqrt{\Delta}}^{t+\Delta} \Delta_t g_e(s) \Delta_{t+k\Delta} g_e(s)\, ds + O(\Delta^{2e+1} w_n(\Delta)) \\ &= \mathbb{E}[\eta_t^2 \mathbf{1}_{t+(k+1)\Delta\leq\tau_n}] \Psi_e(k) \Delta^{2e+1} + O(\Delta^{2e+1} w_n(\Delta)) \\ &= \mathbb{E}[\eta_t^2 \mathbf{1}_{t\leq\tau_n}] \Psi_e(k) \Delta^{2e+1} + O(\Delta^{2e+1} w_n(\Delta)). \end{aligned}$$

The second equality holds because we can add back the portion of the integral from 0 to $t-\sqrt{\Delta}$ (see the analysis of I_{11}^n) and the resulting integral is asymptotically equivalent to the autocovariance at lag k of standard fractional Brownian motion increments over intervals of length Δ. For the last equality, we can argue as in the analysis of I_2^n. Altogether, we have

$$\frac{\gamma_{t+j\Delta}^{(n)}(k;\Delta)}{\gamma_t^{(n)}(\Delta)} = \frac{\mathbb{E}[\eta_t^2 \mathbf{1}_{t\leq\tau_n}] \Psi_e(k) \Delta^{2e+1} + O(\Delta^{2e+1} w_n(\Delta))}{\mathbb{E}[\eta_t^2 \mathbf{1}_{t\leq\tau_n}] \Delta^{2e+1} + O(\Delta^{2e+1} w_n(\Delta))} \longrightarrow \Psi_e(k) \quad \text{as } \Delta \to 0,$$

which shows that X satisfies the conditions in Theorem 3 with $\rho_t^{(n)}(k) \equiv \Psi_e(k)$. □

D Proof of Theorems 4 and 5

This section provides an overview of the proof of Theorems 4 and 5.

D.1 Notation and Localization Procedure

We write $A \lesssim B$ if $A \leq CB$ for some constant C that does not depend on any important variable and $\widetilde{\mathbb{E}} = \mathbb{E}[\,\cdot\,|\mathcal{F}_\infty]$. Given a (deterministic or random) function $f(t)$ (or of the form f_t) and $h, h' > 0$, we define $\Delta_h f(t) = f(t+h) - f(t)$ and $\Delta_{h,h'}^2 f(t) = \Delta_h f(t+h') - \Delta_h f(t) = f(t+h+h') - f(t+h') - f(t+h) + f(t)$. Also, if $h = 1$ ($h = h'$), we simply write $\Delta f(t)$ ($\Delta_h^2 f(t)$). If $f = (f_i)_i$ is a sequence, we write $\Delta_h f_i = f_{i+h} - f_i$.

We denote the continuous part of x by $x_t^c = x_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s$ be and the jump part by $x_t^j = \int_0^t \int_E \delta^l(s,z)\, \pi(ds,dz) + \int_0^t \int_E \delta^s(s,z)\, \widetilde{\pi}(ds,dz)$. Similarly, we decompose $c = v + c^j$, where v is the continuous part of c and $c_t^j = \int_0^t \int_E \delta_c^l(s,z)\, \pi(ds,dz) + \int_0^t \int_E \delta_c^s(s,z)\, \widetilde{\pi}(ds,dz)$. We also use the abbreviation $\boldsymbol{\eta}_s\, d\mathbf{W}_s = \eta_s\, dW_s + \overline{\eta}_s\, d\overline{W}_s$.

If Assumption V2 holds, the nondifferentiable drift and moving average components in c

can be simplified. Because $g_0, h_0 \in \mathcal{A}$, we have $\int_0^t g_0(t-s)\boldsymbol{\eta}_s\, d\mathbf{W}_s = \int_0^t \int_0^r g_0'(r-s)\boldsymbol{\eta}_s\, d\mathbf{W}_s\, dr$ and $\int_0^t h_0(t-s)\theta_s\, ds = \int_0^t \int_s^t h_0'(r-s)dr\, \theta_s\, ds = \int_0^t \int_0^r h_0'(r-s)\theta_s\, ds\, dr$, which shows that both terms are differentiable drifts and can be absorbed into $\int_0^t a_s ds$.[25] The same argument applies to the representations in (B.1). We may therefore assume

$$g_0 \equiv h_0 \equiv g_0^{\eta} \equiv h_0^{\eta} \equiv g_0^{\overline{\eta}} \equiv h_0^{\overline{\eta}} \equiv 0 \tag{D.1}$$

without loss of generality. In the proofs we show that both the differentiable drift, $\int_0^t a_s\, ds$, and the nondifferentiable one, $\int_0^t h(t-s)\theta_s\, ds$, are negligible. Because the former is a special case of the latter obtained by setting $h \equiv 1$, we can also assume

$$a \equiv 0. \tag{D.2}$$

As usual when infill asymptotics are considered (cf. Lemma 4.4.9 in Jacod and Protter (2012)), we can assume the following strengthened hypotheses without loss of generality.

Assumption LLN'. *In addition to Assumption LLN, there is a deterministic constant $K \in (0,\infty)$ such that x, b, σ^2, ρ^2, v, v^{-1}, ζ, ζ^{-1} and $(|\delta^s(t,z)|^r \mathbf{1}_{r\geq 1} + |\delta^l(t,z)|^r \mathbf{1}_{r<1} + |\delta_c^s(t,z)|^{r_c}\mathbf{1}_{r_c\geq 1} + |\delta_c^l(t,z)|^{r_c}\mathbf{1}_{r_c<1})/\delta^*(z)$ are uniformly bounded by K, where $\delta^*(z)$ is a deterministic nonnegative measurable function satisfying $\delta^*(z) \leq K$ for all $z \in E$ and $\int_E \delta^*(z)\,\lambda(dz) < \infty$. Moreover, $\delta^l(t,z)\mathbf{1}_{r\geq 1} \equiv \delta^s(t,z)\mathbf{1}_{r<1} \equiv \delta_c^l(t,z)\mathbf{1}_{r_c\geq 1} \equiv \delta_c^s(t,z)\mathbf{1}_{r_c<1} \equiv 0$ and for any $\iota > 0$ and $p \in \mathbb{N}$, we have*

$$\mathbb{E}\big[\big(\mathbb{E}[v_{s+\Delta_n} - v_s \mid \mathcal{F}_{s-r}]\big)^2\big] = o(\gamma(\Delta_n)), \tag{D.3}$$

$$\mathbb{E}\big[\big|\mathrm{Cov}\big(v_{s+\Delta_n} - v_s, v_{t+\Delta_n} - v_t \mid \mathcal{F}_{s-r}\big) - \zeta_{s-r}\gamma(\Delta_n)\Psi_e(\tfrac{t-s}{\Delta_n})\big|\big] = o(\gamma(\Delta_n)), \tag{D.4}$$

$$\mathbb{E}[|v_{s+\Delta_n} - v_s|^p]^{1/p} = O(\gamma(\Delta_n)^{1/2}), \quad \mathbb{E}[|\zeta_{s+\Delta_n} - \zeta_s|] = o(1) \tag{D.5}$$

as $\Delta_n \to 0$, uniformly in $0 \leq r \leq s \leq t \leq T$ satisfying $r \geq \Delta_n^{1-\iota}$.

Assumption CLT'. *In addition to Assumption CLT, we have (D.1) and (D.2) and there is a deterministic constant $K \in (0,\infty)$ such that the processes listed in Assumption CLT (ii) as well as x, σ^2, ρ^2, a, θ, c^{-1}, η^{-1} and $(|\delta^s(t,z)|^r \mathbf{1}_{r\geq 1} + |\delta^l(t,z)|^r \mathbf{1}_{r<1} + |\delta_c^s(t,z)|^{r_c}\mathbf{1}_{r_c\geq 1} +$*

[25] Since $\alpha > \frac{1}{2}$, the existence of a continuous version of the process $r \mapsto \int_0^r g_0'(r-s)\boldsymbol{\eta}_s\, d\mathbf{W}_s$ follows from the bound $\int_0^{r+\tau} |g_0'(r+\tau-s) - g_0'(r-s)|^2\, ds = \int_0^\tau |g_0'(s)|^2\, ds + \int_0^r |g_0'(s+\tau) - g_0'(s)|^2\, ds \leq C\tau^{2\alpha-1}$ and Kolmogorov's continuity theorem. The continuity of $\int_0^r h_0'(r-s)\theta_s\, ds$ follows from the bound $\int_0^{r+\tau} |h_0'(r+\tau-s) - h_0'(r-s)|\, ds = \int_0^\tau |h_0'(s)|\, ds + \int_0^r |h_0'(s+\tau) - h_0'(s)|\, ds \leq C\tau^{2\alpha}$.

$|\delta_c^l(t,z)|^{r_c}\mathbf{1}_{r_c<1}+|\delta_\eta^s(t,z)|^2+|\delta_{\overline{\eta}}^s(t,z)|^2)/\delta^*(z)$ *are uniformly bounded by* K*, where* $\delta^*(z)$ *is the same function as in Assumption LLN'. Moreover,* $\delta^l(t,z)\mathbf{1}_{r\geq 1}\equiv\delta^s(t,z)\mathbf{1}_{r<1}\equiv\delta_c^l(t,z)\mathbf{1}_{r_c\geq 1}\equiv\delta_c^s(t,z)\mathbf{1}_{r_c<1}\equiv\delta_\eta^l(t,z)\equiv\delta_{\overline{\eta}}^l(t,z)\equiv 0$.

D.2 Main Steps

Proof of Theorem 4 and Theorem 5. The theorems follow by combining (D.6), (D.8) as well as (D.9) (for Theorem 4) and (D.10) (for Theorem 5), obtained in three steps in the remainder of this section. □

Step 1: Removal of Price Jumps

We first show that $V_t^{n,1,\ell}=\frac{\Delta_n}{\Gamma_n}\Delta_n^{2(1-2\alpha)_+}\sum_{i=0}^{\lfloor t/\Delta_n\rfloor-(\ell+1)p_n-k_n}\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{i+\ell p_n}^{n,1}\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_i^{n,1}$ can be substituted for $V_t^{n,\ell}$ under Assumption LLN', where $\Gamma_n=\gamma(p_n\Delta_n)$ ($\Gamma_n=(p_n\Delta_n)^{2e+1}$) for the proof of Theorem 4 (Theorem 5), $\widehat{c}_i^{n,1}=(k_n\Delta_n)^{-1}\sum_{j=i+1}^{i+k_n}(\Delta_j^n x^c+\varepsilon_j^n)^2$, $\Delta_{p_n}\widehat{c}_i^{n,1}=\widehat{c}_{i+p_n}^{n,1}-\widehat{c}_i^{n,1}$ and $\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_i^{n,1}=\Delta_{p_n}\widehat{c}_i^{n,1}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_i^{n,1}|\leq w_n}$.

Proposition D.1. *We have* $\mathbb{E}[|\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{i+\ell p_n}^n\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_i^n-\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{i+\ell p_n}^{n,1}\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_i^{n,1}|]\lesssim\Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{\widetilde{\varpi}}\times(\Delta_n^{(2-r)\varpi}+\Delta_n^{(1/2\wedge\alpha)-1+1/r}\mathbf{1}_{r>1}+\Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{1-r_c\widetilde{\varpi}+\widetilde{\varpi}})$, *uniformly in* i.

Since we have $L(\Delta_n)=o(\Delta_n^{-\epsilon})$ for any $\epsilon>0$ and slowly varying function L, we obtain the bound $(p_n\Delta_n)^{-1/2}\mathbb{E}[\sup_{t\leq T}|V_t^{n,\ell}-V_t^{n,1,\ell}|]\lesssim(p_n\Delta_n)^{-3/2-2e+\widetilde{\varpi}-\epsilon}(\Delta_n^{(1-2\alpha)_++(2-r)\varpi}+\Delta_n^{(1-2\alpha)_++(1/2\wedge\alpha)-1+1/r}\mathbf{1}_{r>1}+(p_n\Delta_n)^{1-r_c\widetilde{\varpi}+\widetilde{\varpi}})$, which tends to zero for sufficiently small ϵ by Assumption LLN (iv) and (v). We have shown that

$$(p_n\Delta_n)^{-1/2}(V_t^{n,\ell}-V_t^{n,1,\ell})\overset{L^1}{\Longrightarrow}0. \tag{D.6}$$

Step 2: Removal of Volatility Jumps

Let $J_{1,i}^n=\frac{1}{k_n\Delta_n}\int_{i\Delta_n}^{(i+k_n)\Delta_n}\Delta_n^{-(1-2\alpha)_+}v_u du$, $J_{2,i}^n=\frac{1}{k_n\Delta_n}\sum_{j=i+1}^{i+k_n}[(\Delta_j^n x^c)^2-\int_{(j-1)\Delta_n}^{j\Delta_n}\sigma_u^2 du]$, $J_{3,i}^n=\frac{1}{k_n\Delta_n}\sum_{j=i+1}^{i+k_n}[(\varepsilon_j^n)^2-\Delta_n^{2\alpha-1}\int_{(j-1)\Delta_n}^{j\Delta_n}\rho_u^2\,du]$, $J_{4,i}^n=\frac{2}{k_n\Delta_n}\sum_{j=i+1}^{i+k_n}\Delta_j^n x^c\varepsilon_j^n$ and $J_{5,i}^n=\frac{1}{k_n\Delta_n}\int_{i\Delta_n}^{(i+k_n)\Delta_n}\Delta_n^{-(1-2\alpha)_+}c_u^j\,du$. Then $\widehat{c}_i^{n,1}=\sum_{k=1}^5 J_{k,i}^n$. Our goal is to remove $J_{5,i}^n$ from $\widehat{c}_i^{n,1}$ in order to obtain $\widehat{c}_i^{n,2}=\sum_{k=1}^4 J_{k,i}^n$ and to replace $V_t^{n,1,\ell}$ by

$$V_t^{n,2,\ell}=\frac{\Delta_n}{\Gamma_n}\Delta_n^{2(1-2\alpha)_+}\sum_{i=0}^{\lfloor t/\Delta_n\rfloor-(\ell+1)p_n-k_n}\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2}\Delta_{p_n}\widehat{c}_i^{n,2}. \tag{D.7}$$

Proposition D.2. *For $\epsilon > 0$, we have $\mathbb{E}[|\Delta_{p_n}^{\mathrm{tr}} \widehat{c}_{i+\ell p_n}^{n,1} \Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^{n,1} - \Delta_{p_n} \widehat{c}_{i+\ell p_n}^{n,2} \Delta_{p_n} \widehat{c}_i^{n,2}|] \lesssim \Delta_n^{-2(1-2\alpha)_+}$ $\times [(p_n\Delta_n)^{e+3/2+\widetilde{\varpi}(1-r_c)-\epsilon} + (p_n\Delta_n)^{e+1/2+1/r_c-\epsilon}]$ uniformly in i.*

The proposition implies that $(p_n\Delta_n)^{-1/2}(V_t^{n,1,\ell} - V_t^{n,2,\ell})$ is of the order $(p_n\Delta_n)^{-3/2-2e-\epsilon} \times ((p_n\Delta_n)^{e+3/2+(1-r_c)\widetilde{\varpi}} + (p_n\Delta_n)^{e+1/2+1/r_c}) \lesssim (p_n\Delta_n)^{-e+(1-r_c)\widetilde{\varpi}-\epsilon} + (p_n\Delta_n)^{-1-e+1/r_c-\epsilon}$ for any $\epsilon > 0$ and $\ell \geq 0$. Since $r_c < \frac{1}{e+1}$ and $\widetilde{\varpi} < \frac{1}{2} + e$, we have $-e + \widetilde{\varpi} - r_c\widetilde{\varpi} > 0$, which implies

$$(p_n\Delta_n)^{-1/2}(V_t^{n,1,\ell} - V_t^{n,2,\ell}) \overset{L^1}{\Longrightarrow} 0. \tag{D.8}$$

Step 3: Limit Theorems in the Absence of Price and Volatility Jumps

The last part of the proof consists in showing that

$$\left(V_t^{n,2,0} - V_t^0, \ldots, V_t^{n,2,m} - V_t^m\right)^T \overset{\mathbb{P}}{\Longrightarrow} 0 \tag{D.9}$$

under the assumptions of Theorem 4 and

$$(p_n\Delta_n)^{-1/2}\left(V_t^{n,2,0} + 2V_t^{n,2,1} - (V_t^0 + 2V_t^1), V_t^{n,2,2} - V_t^2, \ldots, V_t^{n,2,m} - V_t^m\right)^T \overset{\mathcal{L}-s}{\Longrightarrow} \mathcal{Z} \tag{D.10}$$

under the assumptions of Theorem 5. We decompose $V_t^{n,2,\ell} = \sum_{j_1,j_2=1}^4 Z_{j_1,j_2}^{n,\ell}(t)$ where

$$Z_{j_1,j_2}^{n,\ell}(t) = \frac{\Delta_n}{\Gamma_n}\Delta_n^{2(1-2\alpha)_+} \sum_{i=0}^{\lfloor t/\Delta_n \rfloor - (\ell+1)p_n - k_n} \Delta_{p_n} J_{j_1,i}^n \Delta_{p_n} J_{j_2,i+\ell p_n}^n, \quad j_1, j_2 = 1, \ldots, 4.$$

We show (D.9) and (D.10) in two propositions.

Proposition D.3.

(i) Under Assumption LLN', we have $Z_{1,1}^{n,\ell} - V^\ell \overset{L^1}{\Longrightarrow} 0$ for all $\ell \geq 0$.

(ii) Under Assumption CLT', we have $(p_n\Delta_n)^{-1/2}(Z_{1,1}^{n,0} + 2Z_{1,1}^{n,1} - (V^0 + 2V^1), Z_{1,1}^{n,2} - V^2, \ldots, Z_{1,1}^{n,m} - V^m)^T \overset{\mathcal{L}-s}{\Longrightarrow} \mathcal{Z}$.

Proposition D.4. *Consider $j_1, j_2 = 1, \ldots, 4$ such that $(j_1, j_2) \neq (1, 1)$.*

(i) Under Assumption LLN', we have $Z_{j_1,j_2}^{n,\ell} \overset{L^1}{\Longrightarrow} 0$ for $\ell \geq 0$.

(ii) Under Assumption CLT', we have $(p_n\Delta_n)^{-1/2}(Z_{j_1,j_2}^{n,0} + 2Z_{j_1,j_2}^{n,1}) \overset{L^1}{\Longrightarrow} 0$ as well as $(p_n\Delta_n)^{-1/2} Z_{j_1,j_2}^{n,\ell} \overset{L^1}{\Longrightarrow} 0$ for $\ell \geq 2$.

E Preliminary Bounds

The assumptions of Theorem 4 and Assumption LLN' are in force throughout this section.

E.1 Statement of the Preliminary Bounds

We start with bounds on convolutions and recall the notation in (3.14).

Lemma E.1. *For $e \in (-1,0]$, consider $A_t = \int_0^t g_e(t-s)a_s\,ds$ where the process a is uniformly bounded by a constant K. For any $p \geq 1$, there exists $C > 0$ depending only on e, K and p such that $\mathbb{E}[|A_{t+h}-A_t|^p]^{1/p} \leq Ch^{e+1}$ for all $t, h \geq 0$. If in addition $\mathbb{E}[|a_t-a_s|^p]^{1/p} \leq K(t-s)^\beta$ for some $0 < \beta < -e$ and all $s, t \geq 0$, then $\mathbb{E}[|A_{t+h} - A_t|^p]^{1/p} \leq C(1+t^{-\beta})h^{\beta+e+1}$.*

Next, we gather estimates for $J^n_{k,i}$, $k = 1, \dots, 5$, defined before (D.7). To this end, we first decompose $J^n_{1,i}$ and $J^n_{2,i}$ further and recall the notation from Section D.1 as well as (D.1) and (D.2). The decomposition of $J^n_{1,i}$ is only needed for the proof of Theorem 5 and assumes that v satisfies (3.13). Defining $G_e(t) = [(e+1)K_e]^{-1}t_+^{e+1}$, $\Delta^2_{n,i}G_e(r) = \Delta^2_{k_n\Delta_n, p_n\Delta_n}G_e(i\Delta_n - r)$, we can write $\int_{i\Delta_n}^{(i+k_n)\Delta_n}\int_0^u g_e(u-r)\boldsymbol{\eta}_r\,d\mathbf{W}_r\,du = \int_0^{(i+k_n)\Delta_n}\Delta_{k_n\Delta_n}G_e(i\Delta_n - r)\boldsymbol{\eta}_r\,d\mathbf{W}_r$ and therefore, $\Delta_{p_n}J^n_{1,i} = \Delta_{p_n}J^n_{11,i} + \Delta_{p_n}J^n_{12,i}$ where

$$\begin{aligned}
\Delta_{p_n}J^n_{11,i} &= \frac{\Delta_n^{-(1-2\alpha)_+}}{k_n\Delta_n}\int_0^{(i+p_n+k_n)\Delta_n}\Delta^2_{n,i}G_e(r)\boldsymbol{\eta}_r\,d\mathbf{W}_r,\\
\Delta_{p_n}J^n_{12,i} &= \frac{\Delta_n^{-(1-2\alpha)_+}}{k_n\Delta_n}\int_{i\Delta_n}^{(i+k_n)\Delta_n}(A_{s+p_n\Delta_n} - A_s)\,ds, \quad A_t = \int_0^t g_{e'}(t-s)\theta_s\,ds.
\end{aligned} \tag{E.1}$$

Itô's formula and the notation $z_t = \int_0^t \sigma_s\,dW_s$ allow us to write $J^n_{2,i} = J^n_{21,i} + J^n_{22,i}$, where

$$\begin{aligned}
J^n_{21,i} &= \frac{2}{k_n\Delta_n}\int_{i\Delta_n}^{(i+k_n)\Delta_n}(z_s - z_{\lfloor s/\Delta_n\rfloor\Delta_n})\,dz_s,\\
J^n_{22,i} &= \frac{2}{k_n\Delta_n}\int_{i\Delta_n}^{(i+k_n)\Delta_n}\Big((z_s - z_{\lfloor s/\Delta_n\rfloor\Delta_n})b_s + \int_{\lfloor s/\Delta_n\rfloor\Delta_n}^s b_r\,dr\,b_s\Big)\,ds\\
&\quad + \frac{2}{k_n\Delta_n}\int_{i\Delta_n}^{(i+k_n)\Delta_n}\int_{\lfloor s/\Delta_n\rfloor\Delta_n}^s b_r\,dr\,dz_s.
\end{aligned} \tag{E.2}$$

Lemma E.2. *Under the assumptions of Theorem 5 and Assumption CLT', we have for $p \geq 1$ that $\mathbb{E}[|\Delta_{p_n}J^n_{11,i}|^p]^{1/p} \lesssim \Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{e+1/2}$ and $\mathbb{E}[|\Delta_{p_n}J^n_{12,i}|^p]^{1/p} \lesssim \Delta_n^{-(1-2\alpha)_+} \times [(p_n\Delta_n)^{e'+1} \wedge ((p_n\Delta_n)^{\beta+e'+1} + (p_n\Delta_n)^{\beta+e'}\int_{i\Delta_n}^{(i+k_n)\Delta_n} s^{-\beta}\,ds)]$, uniformly in i. In particular, we have $\mathbb{E}[|\Delta_{p_n}J^n_{1,i}|^p]^{1/p} \lesssim \Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{e+1/2}$.*

Lemma E.3. *For any $p \geq 1$, $\mathbb{E}[|J^n_{21,i}|^p \mid \mathcal{F}_{i\Delta_n}]^{1/p} \lesssim k_n^{-1/2}$ and $\mathbb{E}[|J^n_{22,i}|^p \mid \mathcal{F}_{i\Delta_n}]^{1/p} \lesssim \Delta_n^{1/2}$ uniformly in i and ω.*

Lemma E.4. *For any $p \geq 1$, $\mathbb{E}[|J^n_{3,i}|^p \mid \mathcal{F}_\infty]^{1/p} \lesssim \Delta_n^{2\alpha-1}k_n^{-1/2}$ and $|\mathbb{E}[J^n_{3,i_1}J^n_{3,i_2} \mid \mathcal{F}_\infty]| \lesssim \Delta_n^{4\alpha-2}k_n^{-2}$, uniformly in i as well as i_1 and i_2 with $|i_1 - i_2| \geq k_n$.*

Lemma E.5. *For any $p \geq 1$, $\mathbb{E}[|J^n_{4,i}|^p \mid \mathcal{F}_{i\Delta_n}]^{1/p} \lesssim \Delta_n^{\alpha-1/2}k_n^{-1/2}$ and $\mathbb{E}[\widetilde{\mathbb{E}}[J^n_{4,i_1}J^n_{4,i_2}]^2 \mid \mathcal{F}_{(i_1\wedge i_2)\Delta_n}]^{1/2} \lesssim \Delta_n^{2\alpha-1}k_n^{-2}$, uniformly in i as well as i_1 and i_2 with $|i_1 - i_2| \geq k_n$.*

Lemma E.6. *If $r_c \geq 1$ and $p \in [r_c, 2]$ or if $r_c < 1$ and $p \in [r_c, 1]$, we have $\mathbb{E}[|\Delta_{p_n}J^n_{5,i}|^p \mid \mathcal{F}_{i\Delta_n}]^{1/p} \lesssim \Delta_n^{-(1-2\alpha)_+}(k_n\Delta_n)^{1/p}$ and $\mathbb{E}[(|\Delta_{p_n}J^n_{5,i}| \wedge (2w_n))^p \mid \mathcal{F}_{i\Delta_n}]^{1/p} \lesssim \Delta_n^{-(1-2\alpha)_+} \times (k_n\Delta_n)^{(1-\widetilde{\varpi}r_c)/p+\widetilde{\varpi}}$, uniformly in i.*

Finally, we estimate the truncation probabilities for returns and volatility increments.

Lemma E.7. *If $\alpha > 0$, we have $\mathbb{P}(|\Delta^n_j y| > v_n) \lesssim \Delta_n^{1-r\varpi}$ and $\mathbb{P}(|\Delta^n_j x^c + \varepsilon^n_j| > v_n) \lesssim \Delta_n^\eta$ for any $\eta > 0$ uniformly in j.*

Lemma E.8. *We have $\mathbb{P}(|\Delta_{p_n}\widehat{c}^{n,1}_i| \geq w_n) \lesssim (p_n\Delta_n)^{1-r_c\widetilde{\varpi}}$ uniformly in i.*

E.2 Proof of the Preliminary Bounds

Proof of Lemma E.1. The case $e = 0$ is trivial. For $e < 0$, recall the forward difference notation from Section D.1. Since a is uniformly bounded and $A_{t+h} - A_t = \int_t^{t+h} g_e(t+h-s)a_s\,ds + \int_0^t \Delta_h g_e(t-s)a_s\,ds$, we have $|A_{t+h} - A_t| \lesssim \int_0^h g_e(s)\,ds + \int_0^t |\Delta_h g_e(s)|\,ds \lesssim h^{e+1}$. If a is β-Hölder continuous in L^p, the identity $\Delta_h^2 A_t = \int_0^t \Delta_h g_e(s)\Delta_h a_s\,ds + \int_0^h \Delta_h g_e(t+h-s)a_s\,ds + \int_0^h g_e(s)\Delta_h a_{t+h-s}\,ds$ implies that $\mathbb{E}[|\Delta_h^2 A_t|^p]^{1/p} \lesssim h^\beta(\int_0^t |\Delta_h g_e(s)|\,ds + \int_0^h g_e(s)\,ds) + \int_0^h |\Delta_h g_e(t+h-s)|\,ds \lesssim h^{\beta+e+1} + h^2 t^{e-1} \wedge h^{e+1} \leq (1+t^{-\beta})h^{\beta+e+1}$. We conclude using Proposition 2 in Račkauskas and Suquet (2005). □

Proof of Lemma E.2. It suffices to consider the case $\alpha = \frac{1}{2}$ and, by Jensen's inequality, $p \geq 2$. The bound on $\Delta_{p_n}J^n_{11,i}$ follows from standard semimartingale estimates for $e = 0$. If $e < 0$, applying the Burkholder–Davis–Gundy inequalities and using the assumption that $|\eta|, |\overline{\eta}| \leq K$, we obtain $\mathbb{E}[|\Delta_{p_n}J^n_{11,i}|^p]^{1/p} \lesssim (k_n\Delta_n)^{-1}(\int_0^{(i+p_n+k_n)\Delta_n} |\Delta^2_{n,i}G_e(r)|^2\,dr)^{1/2} = (k_n\Delta_n)^{e+1/2}(\int_0^{i/k_n+\kappa+1}[G_e(u) - G_e(u-1) - G_e(u-\kappa) + G_e(u-\kappa-1)]^2\,du)^{1/2}$ by changing variables $u = i/k_n + \kappa + 1 - r/(k_n\Delta_n)$. Because $|G_e(u) - G_e(u-1) - G_e(u-\kappa) + G_e(u-\kappa-1)| \lesssim 1 \wedge u^{e-1}$, the integral from 0 to ∞ is finite and we have shown $\mathbb{E}[|\Delta_{p_n}J^n_{11,i}|^p]^{1/p} \lesssim (k_n\Delta_n)^{e+1/2}$. The bound on $\Delta_{p_n}J^n_{12,i}$ follows directly from Lemma E.1 and Assumption CLT'. □

Proof of Lemma E.3. The lemma follows from classical martingale estimates. □

Proof of Lemma E.4. The general case follows from the special case $\alpha = \frac{1}{2}$, which we consider. Recall that $\widetilde{\mathbb{E}}$ stands for $\mathbb{E}[\cdot \mid \mathcal{F}_\infty]$ and that $\mathcal{G}_i = \sigma(\chi_j,\, j \le i)$ and $\mathcal{G}^i = \sigma(\chi_j,\, j \ge i)$ are the pre- and post-σ-fields at time i. In this proof, we use the notation $\Gamma_j^n = \int_{(j-1)\Delta_n}^{j\Delta_n} \rho_s^2\, ds$ for $j \ge 1$. By Jensen's inequality, it is enough to treat the case $p = 2m$, where $m \ge 1$ is an integer, for the first bound. In that case, we can explicitly develop

$$\widetilde{\mathbb{E}}[(J_{3,i}^n)^{2m}] = \frac{1}{(k_n\Delta_n)^{2m}} \sum_{j_1,\dots,j_{2m}=i+1}^{i+k_n} \Big(\prod_{k=1}^{2m} \Gamma_{j_k}^n\Big) \widetilde{\mathbb{E}}\Big[\prod_{k=1}^{2m} \big(\chi_{j_k}^2 - 1\big)\Big].$$

As ρ is bounded by Assumption LLN', we have $|\Gamma_j^n| \le K\Delta_n$ for all $j \ge 1$. Thus,

$$\widetilde{\mathbb{E}}[(J_{3,i}^n)^{2m}] \lesssim \frac{1}{k_n^{2m}} \sum_{j_1,\dots,j_{2m}=i+1}^{i+k_n} \Big|\widetilde{\mathbb{E}}\Big[\prod_{k=1}^{2m} \big(\chi_{j_k}^2 - 1\big)\Big]\Big|.$$

Since $(\chi_j)_j$ is centered, stationary, exponentially ρ-mixing, with unit variance and finite moments of all orders, the sequence $(Y_j)_j$ defined by $Y_j = \chi_j^2 - 1$ is also centered, stationary, exponentially ρ-mixing and has finite moments of all orders. Lemma E.9 yields $\widetilde{\mathbb{E}}[(J_{3,i}^n)^{2m}] \lesssim k_n^{-m}$, which is the first bound in the lemma.

For the second bound, we assume without loss of generality that $i_2 \ge i_1 + k_n$. Then

$$\widetilde{\mathbb{E}}[J_{3,i_1}^n J_{3,i_2}^n] = \frac{1}{k_n^2\Delta_n^2} \sum_{j_1=i_1+1}^{i_1+k_n} \sum_{j_2=i_2+1}^{i_2+k_n} \Big\{\Gamma_{j_1}^n \Gamma_{j_2}^n \widetilde{\mathbb{E}}\big[\big(\chi_{j_1}^2 - 1\big)\big(\chi_{j_2}^2 - 1\big)\big]\Big\},$$

where $\chi_{j_1}^2 - 1$ is $\mathcal{G}_{j_1}$-measurable and $\chi_{j_2}^2 - 1$ is $\mathcal{G}^{j_2}$-measurable. Both variables are centered and have a finite variance. Using Assumption LLN (ii), we know that $\rho_{|j_2-j_1|}(\chi) \le K\exp(-\rho|j_2 - j_1|)$ and therefore, $\widetilde{\mathbb{E}}[(\chi_{j_1}^2 - 1)(\chi_{j_2}^2 - 1)] \lesssim \exp(-\rho|j_2 - j_1|)$. Using again the bound $|\Gamma_j^n| \le K\Delta_n$, we obtain

$$\begin{aligned}\big|\widetilde{\mathbb{E}}[J_{3,i_1}^n J_{3,i_2}^n]\big| &\lesssim \frac{1}{k_n^2} \sum_{j_1=i_1+1}^{i_1+k_n} \sum_{j_2=i_2+1}^{i_2+k_n} e^{-\rho(j_2-j_1)} = \frac{1}{k_n^2} e^{-\rho(i_2-i_1)} \frac{(1-e^{-\rho k_n})(e^{\rho k_n}-1)}{(1-e^{-\rho})(e^{\rho}-1)} \\ &\lesssim k_n^{-2} \exp(-\rho(i_2 - i_1 - k_n)).\end{aligned}$$

Since $i_2 - i_1 \ge k_n$ we obtain the second bound in the lemma. □

Proof of Lemma E.5. As for the previous lemma, it is enough to consider $\alpha = \frac{1}{2}$ and prove

the first bound for $p = 2m$, where $m \geq 1$. In this case, we can develop explicitly

$$\widetilde{\mathbb{E}}[(J^n_{4,i})^{2m}] = \frac{4^m}{(k_n\Delta_n)^{2m}} \sum_{j_1,\dots,j_{2m}=i+1}^{i+k_n} \prod_{k=1}^{2m} \big(\Delta^n_{j_k} x^c (\Gamma^n_{j_k})^{1/2}\big) \widetilde{\mathbb{E}}\Big[\prod_{k=1}^{2m} \chi_{j_k}\Big].$$

Since $(\chi_j)_{j\geq 1}$ is independent of $\mathcal{F}_\infty$, we deduce that

$$\mathbb{E}\big[(J^n_{4,i})^{2m} \mid \mathcal{F}_{i\Delta_n}\big] = \frac{4^m}{(k_n\Delta_n)^{2m}} \sum_{j_1,\dots,j_{2m}=i+1}^{i+k_n} \mathbb{E}\Big[\prod_{k=1}^{2m} \big(\Delta^n_{j_k} x^c (\Gamma^n_{j_k})^{1/2}\big) \,\Big|\, \mathcal{F}_{i\Delta_n}\Big] \mathbb{E}\Big[\prod_{k=1}^{2m} \chi_{j_k}\Big].$$

Because $|\Gamma^n_j| \lesssim \Delta_n$ and $\mathbb{E}[|\Delta^n_j x^c|^p \mid \mathcal{F}_{i\Delta_n}] \lesssim \Delta_n^{p/2}$ for all p, Hölder's inequality yields $\mathbb{E}[\prod_{k=1}^{2m} \big(\Delta^n_{j_k} x^c (\Gamma^n_{j_k})^{1/2}\big) \mid \mathcal{F}_{i\Delta_n}] \lesssim \Delta_n^{2m}$. We obtain

$$\mathbb{E}\big[(J^n_{4,i})^{2m} \mid \mathcal{F}_{i\Delta_n}\big] \lesssim \frac{1}{k_n^{2m}} \sum_{j_1,\dots,j_{2m}=i+1}^{i+k_n} \mathbb{E}\Big[\prod_{k=1}^{2m} \chi_{j_k}\Big].$$

Lemma E.9 implies the first bound in Lemma E.5.

For the second bound, we assume without loss of generality that $i_2 \geq i_1 + k_n$. Then

$$\widetilde{\mathbb{E}}[J^n_{4,i_1} J^n_{4,i_2}] = \frac{4}{k_n^2\Delta_n^2} \sum_{j_1=i_1+1}^{i_1+k_n} \sum_{j_2=i_2+1}^{i_2+k_n} \Big\{\Delta^n_{j_1} x^c \Delta^n_{j_2} x^c (\Gamma^n_{j_1}\Gamma^n_{j_2})^{1/2} \widetilde{\mathbb{E}}[\chi_{j_1}\chi_{j_2}]\Big\}$$

As before, we obtain $\mathbb{E}[\widetilde{\mathbb{E}}[J^n_{4,i_1} J^n_{4,i_2}]^2 \mid \mathcal{F}_{(i_1\wedge i_2)\Delta_n}]^{1/2} \lesssim k_n^{-2} \sum_{j_1=i_1+1}^{i_1+k_n} \sum_{j_2=i_2+1}^{i_2+k_n} e^{-\rho(j_2-j_1)}$. We conclude using the same arguments as in Lemma E.4. □

Proof of Lemma E.6. It is enough to consider $\alpha = \frac{1}{2}$. If $r_c \geq 1$, we have $\delta^l_c(t,z) \equiv 0$ by Assumption LLN'. Therefore,

$$\begin{aligned}\Delta_{p_n} J^n_{5,i} &= \frac{1}{k_n\Delta_n} \int_{i\Delta_n}^{(i+k_n)\Delta_n} \int_u^{u+p_n\Delta_n} \int_E \delta^s_c(s,z)\, \widetilde{\pi}(ds,dz)\, du \qquad \text{(E.3)}\\ &= \frac{1}{k_n\Delta_n} \int_{i\Delta_n}^{(i+p_n+k_n)\Delta_n} \int_E \big((i+k_n)\Delta_n \wedge s - (s-p_n\Delta_n)\vee i\Delta_n\big)\delta^s_c(s,z)\, \widetilde{\pi}(ds,dz).\end{aligned}$$

Lemma 2.1.5 in Jacod and Protter (2012) yields $\mathbb{E}[|\Delta_{p_n} J^n_{5,i}|^p \mid \mathcal{F}_{i\Delta_n}]^{1/p} \lesssim (k_n\Delta_n)^{1/p}$ and $\mathbb{E}[|\Delta_{p_n} J^n_{5,i}|^p \wedge (2w_n)^p \mid \mathcal{F}_{i\Delta_n}]^{1/p} \lesssim \mathbb{E}[|\Delta_{p_n} J^n_{5,i}|^{r_c} \mid \mathcal{F}_{i\Delta_n}]^{1/p} w_n^{1-r_c/p} \lesssim (k_n\Delta_n)^{(1-\widetilde{\varpi}r_c)/p+\widetilde{\varpi}}$ for $p \in [r_c, 2]$. If $r_c < 1$, we have (E.3) with δ^l_c instead of δ^s_c and $\pi(ds,dz)$ instead of $\widetilde{\pi}(ds,dz)$. Lemma 2.1.7 in Jacod and Protter (2012) yields the same bounds for all $p \in [r_c, 1]$. □

Proof of Lemma E.7. Recall the decomposition $x = x^c + x^j$ from Appendix D.1, where by Assumption LLN' x^j only contains small (large) jumps if $r \geq 1$ ($r < 1$). For all $q \geq 1$,

$$\begin{aligned}\mathbb{P}(|\Delta_j^n y| > v_n) &\leq \mathbb{P}(|\Delta_j^n x^c| > v_n/3) + \mathbb{P}(|\varepsilon_j^n| > v_n/3) + \mathbb{P}(|\Delta_j^n x^j| > v_n/3) \\ &\leq 3^q \mathbb{E}[|\Delta_j^n x^c|^q] v_n^{-q} + 3^q \mathbb{E}[|\varepsilon_j^n|^q] v_n^{-q} + 3\mathbb{E}[|\Delta_j^n x^j|^r] v_n^{-r}.\end{aligned}$$

Using Assumption LLN', we have $\mathbb{E}[|\Delta_j^n x^c|^q] \lesssim \Delta_n^{q/2}$ and $\mathbb{E}[|\varepsilon_j^n|^q] \lesssim \Delta_n^{q\alpha}\mathbb{E}[|\chi_0|^q]$. Lemmas 2.1.5 (if $r \geq 1$) and 2.1.7 (if $r < 1$) in Jacod and Protter (2012) show that $\mathbb{E}[|\Delta_j^n x^j|^r] \lesssim \Delta_n$. Combining these bounds, we get $\mathbb{P}(|\Delta_j^n y| > v_n) \lesssim \Delta_n^{q(1/2\wedge\alpha-\varpi)} + \Delta_n^{1-r\varpi}$. The first bound follows by taking q large enough because $\varpi < \frac{1}{2} \wedge \alpha$. The second bound is shown similarly. □

Proof of Lemma E.8. We have $\Delta_{p_n}\widehat{c}_i^{n,1} = \sum_{k=1}^5 \Delta_{p_n} J_{k,i}^n$ and $\mathbb{P}(|\Delta_{p_n}\widehat{c}_i^{n,1}| \geq w_n) \leq \sum_{k=1}^5 \mathbb{P}(|\Delta_{p_n} J_{k,i}^n| \geq w_n/5) \leq 5^p w_n^{-p} \sum_{k=1}^4 \mathbb{E}[|\Delta_{p_n} J_{k,i}^n|^p] + w_n^{-r_c}\mathbb{E}[|\Delta_{p_n} J_{5,i}^n|^{r_c}]$ for $p > 1$. Using Lemmas E.3–E.6 and (D.5) (for the proof of Theorem 4) or Lemma E.2 (for the proof of Theorem 5), we obtain $\mathbb{P}(|\Delta_{p_n}\widehat{c}_i^{n,1}| \geq w_n) \lesssim w_n^{-p}[\Delta_n^{-(1-2\alpha)_+}(k_n\Delta_n)^{(e+1/2)}L(k_n\Delta_n)]^p + w_n^{-p}\Delta_n^{-p(1-2\alpha)_+}k_n^{-p/2} + w_n^{-r_c}\Delta_n^{-r_c(1-2\alpha)_+}k_n\Delta_n$ for some slowly varying function L. Because $k_n \sim \theta\Delta_n^{-1/2}\log\Delta_n^{-1}$ and $e < 0$ (or $k_n^{-1} = o(\gamma(p_n\Delta_n))$ if $e = 0$), we have $k_n^{-1/2} = o((k_n\Delta_n)^{e+1/2} \times L(k_n\Delta_n))$. Since $w_n \sim w\Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{\widetilde{\varpi}}$ and $p_n/k_n = \kappa$, we obtain $\mathbb{P}(|\Delta_{p_n}\widehat{c}_i^{n,1}| \geq w_n) \lesssim [(p_n\Delta_n)^{(e+1/2-\widetilde{\varpi})}L(k_n\Delta_n)]^p + (p_n\Delta_n)^{1-r_c\widetilde{\varpi}}$. As $\widetilde{\varpi} < \frac{1}{2} + e$, we can conclude by taking p large enough. □

E.3 A Bound on Mixing Sequences

We prove the following bound for exponentially mixing random variables, which is implicitly contained in Yokoyama (1980).

Lemma E.9. *Suppose that $(Y_j)_{j\geq 1}$ is centered, stationary, exponentially ρ-mixing in the sense of Assumption LLN (ii), with finite moments of all orders. Then for any $m \geq 1$, there exists a constant C_m such that for all $k \geq 1$,*

$$\sum_{j_1,\ldots,j_m=1}^{k} \left|\mathbb{E}\Big[\prod_{l=1}^{m} Y_{j_l}\Big]\right| \leq C_m k^{\lfloor m/2 \rfloor}. \tag{E.4}$$

Proof. By symmetry, (E.4) is bounded by $m!\sum_{\boldsymbol{j}} |\mathbb{E}[\prod_{l=1}^m Y_{j_l}]|$, where $\boldsymbol{j} = (j_1, \ldots, j_m)$ run through all indices satisfying $1 \leq j_1 \leq j_2 \leq \cdots \leq j_m \leq k$. Hence, it is enough to prove

that $\sum_{\boldsymbol{j}} |\mathbb{E}[\prod_{l=1}^m Y_{j_l}]| \le C_m k^{\lfloor m/2 \rfloor}$. We proceed by induction on m. For $m = 1$, the claim is immediate since Y_j is centered. Suppose now that $m \ge 2$, and that the result holds for all smaller m. For $\boldsymbol{j}$ as before, define $M(\boldsymbol{j}) = \max\{j_{i+1} - j_i : 1 \le i \le m-1\}$ and choose $1 \le h \le m-1$ (which depends on $\boldsymbol{j}$) such that $j_{h+1} - j_h = M(\boldsymbol{j})$. By the mixing property of $(Y_j)_{j \ge 1}$,

$$\begin{aligned}\Big|\mathbb{E}\Big[\prod_{l=1}^m Y_{j_l}\Big]\Big| &\le \Big|\mathbb{E}\Big[\prod_{l=1}^h Y_{j_l}\Big]\mathbb{E}\Big[\prod_{l=h+1}^m Y_{j_l}\Big]\Big| + K\exp(-M(\boldsymbol{j})\rho)\mathbb{E}\Big[\prod_{l=1}^h Y_{j_l}^2\Big]^{1/2}\mathbb{E}\Big[\prod_{l=h+1}^m Y_{j_l}^2\Big]^{1/2} \\ &\le \Big|\mathbb{E}\Big[\prod_{l=1}^h Y_{j_l}\Big]\mathbb{E}\Big[\prod_{l=h+1}^m Y_{j_l}\Big]\Big| + C_m \exp(-M(\boldsymbol{j})\rho),\end{aligned}$$

where C_m only depends on m and the first $2m$ moments of Y_j. If $h = 1$ or $h = m-1$, the first term on the right-hand side is zero because Y_j is centered. Only the cases $h \in \{2, \dots, m-2\}$ remain and we have

$$\sum_{\boldsymbol{j}} \Big|\mathbb{E}\Big[\prod_{l=1}^m Y_{j_l}\Big]\Big| \le \sum_{\boldsymbol{j}} \sum_{h=2}^{m-2} \Big|\mathbb{E}\Big[\prod_{l=1}^h Y_{j_l}\Big]\mathbb{E}\Big[\prod_{l=h+1}^m Y_{j_l}\Big]\Big| + C_m \sum_{M=0}^{k-1} \exp(-M\rho) \sum_{\boldsymbol{j}} \mathbf{1}_{M(\boldsymbol{j})=M}.$$

For the first sum, separating the first h from the last $m-h$ indices yields the upper bound

$$\sum_{h=2}^{m-2} \sum_{1 \le j_1 \le j_2 \le \cdots \le j_h \le k} \Big|\mathbb{E}\Big[\prod_{l=1}^h Y_{j_l}\Big]\Big| \sum_{1 \le j_{h+1} \le j_{h+2} \le \cdots \le j_m \le k} \Big|\mathbb{E}\Big[\prod_{l=h+1}^m Y_{j_l}\Big]\Big|,$$

which by the induction hypothesis is bounded by a constant times $\sum_{h=2}^{m-2} k^{\lfloor h/2 \rfloor + \lfloor (m-h)/2 \rfloor} \lesssim k^{\lfloor m/2 \rfloor}$. As the number of sequences $\boldsymbol{j}$ with $M(\boldsymbol{j}) = M$ is bounded by $k(M+1)^{m-1}$, the second sum is bounded by $kC_m \sum_{M=0}^{k-1} \exp(-M\rho)(M+1)^{m-1} \lesssim k \lesssim k^{\lfloor m/2 \rfloor}$, which completes the proof of the lemma. □

F Proof of Propositions D.1–D.4

The assumptions of Theorem 4 and Assumption LLN' are in force throughout this section. For brevity, we write $\sum_i$ instead of $\sum_{i=0}^{\lfloor t/\Delta_n \rfloor - (\ell+1)p_n - k_n}$.

F.1 Main Proofs

Proof of Proposition D.1. We express the difference on the left-hand side of the inequality as

$$\begin{aligned}
&\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{i+\ell p_n}^{n}\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{i}^{n}-\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{i+\ell p_n}^{n,1}\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{i}^{n,1}\\
&\quad=\Delta_{p_n}\widehat{c}_{i}^{n}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i}^{n}|\le w_n}\Big\{\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n}|\le w_n}-\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}|\le w_n}\Big\}\\
&\qquad+\Big\{\Delta_{p_n}\widehat{c}_{i}^{n}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i}^{n}|\le w_n}-\Delta_{p_n}\widehat{c}_{i}^{n,1}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i}^{n,1}|\le w_n}\Big\}\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}|\le w_n}.
\end{aligned}$$

Bounding $\Delta_{p_n}\widehat{c}_{i}^{n}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i}^{n}|\le w_n}$ and $\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}|\le w_n}$ by w_n, we obtain the assertion of the proposition by proving the following uniform bound in i:

$$\begin{aligned}
&\mathbb{E}[|\Delta_{p_n}\widehat{c}_{i}^{n}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i}^{n}|\le w_n}-\Delta_{p_n}\widehat{c}_{i}^{n,1}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i}^{n,1}|\le w_n}|]\\
&\quad\lesssim\Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{1-r_c\widetilde{\varpi}+\widetilde{\varpi}}+\Delta_n^{(2-r)\varpi}+\Delta_n^{(1/2\wedge\alpha)-1+1/r}\mathbf{1}_{r>1}.
\end{aligned}\tag{F.1}$$

The left-hand side of (F.1) is equal to $\mathbb{E}[|\Delta_{p_n}\widehat{c}_{i}^{n}-\Delta_{p_n}\widehat{c}_{i}^{n,1}|\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i}^{n}|\le w_n,\,|\Delta_{p_n}\widehat{c}_{i}^{n,1}|\le w_n}]+\mathbb{E}[|\Delta_{p_n}\widehat{c}_{i}^{n}|\times\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i}^{n}|\le w_n,\,|\Delta_{p_n}\widehat{c}_{i}^{n,1}|>w_n}]+\mathbb{E}[|\Delta_{p_n}\widehat{c}_{i}^{n,1}|\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i}^{n}|>w_n,\,|\Delta_{p_n}\widehat{c}_{i}^{n,1}|\le w_n}]$. Lemma F.1 shows that the first expectation is bounded by $\mathbb{E}[|\widehat{c}_{i}^{n}-\widehat{c}_{i}^{n,1}|]+\mathbb{E}[|\widehat{c}_{i+p_n}^{n}-\widehat{c}_{i+p_n}^{n,1}|]\lesssim\Delta_n^{(2-r)\varpi}+\Delta_n^{(1/2\wedge\alpha)-1+1/r}\mathbf{1}_{r>1}$. For the second and third terms, we bound $|\Delta_{p_n}\widehat{c}_{i}^{n}|,|\Delta_{p_n}\widehat{c}_{i}^{n,1}|\le w_n$, $\mathbb{P}(|\Delta_{p_n}\widehat{c}_{i}^{n,1}|>w_n)\lesssim(p_n\Delta_n)^{1-r_c\widetilde{\varpi}}$ (see Lemma E.8) and $\mathbb{P}(|\Delta_{p_n}\widehat{c}_{i}^{n}|\ge w_n)\lesssim\mathbb{P}(|\Delta_{p_n}\widehat{c}_{i}^{n,1}|\ge w_n/2)+\mathbb{P}(|\Delta_{p_n}\widehat{c}_{i}^{n}-\Delta_{p_n}\widehat{c}_{i}^{n,1}|\ge w_n/2)\lesssim(p_n\Delta_n)^{1-r_c\widetilde{\varpi}}+w_n^{-1}\Delta_n^{(2-r)\varpi}$ to obtain $\Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{1-r_c\widetilde{\varpi}+\widetilde{\varpi}}+\Delta_n^{(2-r)\varpi}+\Delta_n^{(1/2\wedge\alpha)-1+1/r}\mathbf{1}_{r>1}$ as the overall bound. □

Proof of Proposition D.2. Since $k_n\asymp\Delta_n^{-1/2}\log\Delta_n^{-1}$ if $e<0$ (or $k_n^{-1}=o(\gamma(p_n\Delta_n))$ if $e=0$), Lemmas E.3–E.5 together with either (D.5) or Lemma E.2 show that, for some slowly varying function L,

$$\begin{aligned}
\mathbb{E}[|\Delta_{p_n}\widehat{c}_{i}^{n,2}|^p]^{1/p}&\lesssim\Delta_n^{-(1-2\alpha)_+}\big[(p_n\Delta_n)^{e+1/2}L(p_n\Delta_n)+k_n^{-1/2}\big]\\
&\lesssim\Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{e+1/2}L(p_n\Delta_n)
\end{aligned}\tag{F.2}$$

uniformly in i. We decompose $\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{i+\ell p_n}^{n,1}\Delta_{p_n}^{\mathrm{tr}}\widehat{c}_{i}^{n,1}-\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2}\Delta_{p_n}\widehat{c}_{i}^{n,2}=\sum_{k=1}^{4}\delta_{k,i}^{n,\ell}$, where

$$\begin{aligned}
\delta_{1,i}^{n,\ell}&=-\Delta_{p_n}\widehat{c}_{i}^{n,1}\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i}^{n,1}|\le w_n,\,|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}|>w_n},\\
\delta_{2,i}^{n,\ell}&=\Delta_{p_n}\widehat{c}_{i}^{n,1}(\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}-\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2})\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i}^{n,1}|\le w_n,\,|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}|\le w_n},\\
\delta_{3,i}^{n,\ell}&=(\Delta_{p_n}\widehat{c}_{i}^{n,1}-\Delta_{p_n}\widehat{c}_{i}^{n,2})\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_{i}^{n,1}|\le w_n},
\end{aligned}$$

$$\delta_{4,i}^{n,\ell} = -\Delta_{p_n}\widehat{c}_i^{n,2}\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_i^{n,1}|>w_n}.$$

We have $\mathbb{E}[|\delta_{1,i}^{n,\ell}|] \leq w_n\mathbb{E}[|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2}|^q]^{1/q}\mathbb{P}[|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}| > w_n]^{1/q'}$ where $\frac{1}{q}+\frac{1}{q'}=1$. By (F.2) and Lemma E.8, we get $\mathbb{E}[|\delta_{1,i}^{n,\ell}|] \lesssim w_n\Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{e+1/2+(1-r_c\widetilde{\varpi})/q'}L(p_n\Delta_n)$. Taking q' close to 1, we obtain $\mathbb{E}[|\delta_{1,i}^{n,\ell}|] \lesssim \Delta_n^{-2(1-2\alpha)_+}(p_n\Delta_n)^{\widetilde{\varpi}+e+3/2-r_c\widetilde{\varpi}-\epsilon}$ for any $\epsilon>0$. If $r_c<1$, we bound $\mathbb{E}[|\delta_{2,i}^{n,\ell}|] \leq w_n\mathbb{E}[|\Delta_{p_n}J_{5,i+\ell p_n}^n|\mathbf{1}_{\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}|\leq w_n}] \leq w_n\mathbb{E}[|\Delta_{p_n}J_{5,i+\ell p_n}^n| \times \mathbf{1}_{|\Delta_{p_n}J_{5,i+\ell p_n}^n|\leq 2w_n}] + w_n\mathbb{E}[|\Delta_{p_n}J_{5,i+\ell p_n}^n|\mathbf{1}_{|\Delta_{p_n}J_{5,i+\ell p_n}^n|>2w_n,\,|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}|\leq w_n}]$. By Lemma E.6, the first term is $O(\Delta_n^{-2(1-2\alpha)_+}(p_n\Delta_n)^{1+(2-r_c)\widetilde{\varpi}})$. For the second term note that we have $|\Delta_{p_n}J_{k,i+\ell p_n}^n| > w_n/4$ for at least one $k=1,\dots,4$. The proof of Lemma E.8 shows that this probability is $o((p_n\Delta_n)^p)$ for any $p>0$. Therefore, the second term in negligible and we have $\mathbb{E}[|\delta_{2,i}^{n,\ell}|] \lesssim \Delta_n^{-2(1-2\alpha)_+}(p_n\Delta_n)^{1+(2-r_c)\widetilde{\varpi}}$.

If $r_c \geq 1$, we let r_c' satisfy $1/r_c+1/r_c'=1$ (with the following terms suitably interpreted for $r_c=1$ and $r_c'=\infty$). Then, by Hölder's inequality,

$$\begin{aligned}\mathbb{E}[|\delta_{2,i}^{n,\ell}|] \leq \mathbb{E}[|\Delta_{p_n}\widehat{c}_i^{n,1}|^{r_c'}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_i^{n,1}|\leq w_n}]^{1/r_c'}\Big(&\mathbb{E}[|\Delta_{p_n}J_{5,i+\ell p_n}^n|^{r_c}\mathbf{1}_{|\Delta_{p_n}J_{5,i+\ell p_n}^n|\leq 2w_n}]^{1/r_c}\\ &+\mathbb{E}[|\Delta_{p_n}J_{5,i+\ell p_n}^n|^{r_c}\mathbf{1}_{|\Delta_{p_n}J_{5,i+\ell p_n}^n|>2w_n,\,|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,1}|\leq w_n}]^{1/r_c}\Big).\end{aligned}$$

As $\mathbb{E}[|\Delta_{p_n}\widehat{c}_i^{n,1}|^{r_c'}\mathbf{1}_{|\Delta_{p_n}\widehat{c}_i^{n,1}|\leq w_n}]^{1/r_c'} \lesssim \Delta_n^{-(1-2\alpha)_+}[(p_n\Delta_n)^{e+1/2-\epsilon}+(p_n\Delta_n)^{1/r_c'+\widetilde{\varpi}(1-r_c/r_c')}]$ by (F.2) and Lemma E.6, the first term on the right-hand side of the previous display is $O(\Delta_n^{-2(1-2\alpha)_+}\times[(p_n\Delta_n)^{e+1/2-\epsilon+1/r_c}+(p_n\Delta_n)^{1+\widetilde{\varpi}(1-r_c/r_c')}])$. As in the case $r_c<1$, the second term is negligible, so $\mathbb{E}[|\delta_{2,i}^{n,\ell}|] \lesssim \Delta_n^{-2(1-2\alpha)_+}[(p_n\Delta_n)^{e+1/2-\epsilon+1/r_c}+(p_n\Delta_n)^{1+\widetilde{\varpi}(2-r_c)}]$.

We turn to $\delta_{3,i}^{n,\ell}$. If $r_c \geq 1$, Lemma E.6 and (F.2) imply that $\mathbb{E}[|\delta_{3,i}^{n,\ell}|] \leq \mathbb{E}[|\Delta_{p_n}J_{5,i}^n|^{r_c}]^{1/r_c}\times \mathbb{E}[|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2}|^{r_c'}]^{1/r_c'} \lesssim \Delta_n^{-2(1-2\alpha)_+}(p_n\Delta_n)^{e+1/2+1/r_c-\epsilon}$. If $r_c<1$, we choose $q,q'>1$ with $\frac{1}{q}+\frac{1}{q'}=1$ to bound $\mathbb{E}[|\delta_{3,i}^{n,\ell}|] \leq \mathbb{E}[|\Delta_{p_n}J_{5,i}^n|^q\mathbf{1}_{|\Delta_{p_n}J_{5,i}^n|\leq 2w_n}]^{1/q}\mathbb{E}[|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2}|^{q'}]^{1/q'}+\mathbb{E}[|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2}|^4]^{1/4}\times\mathbb{P}[|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2}|>w_n]^{1/4}\mathbb{E}[|\Delta_{p_n}J_{5,i}^n|^2]^{1/2}$. Because $\mathbb{P}[|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2}|>w_n]=o((p_n\Delta_n)^p)$ for any $p>0$ (see the proof of Lemma E.8), the second term is negligible. Taking q close to 1, we conclude that $\mathbb{E}[|\delta_{3,i}^{n,\ell}|] \lesssim \Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{e+3/2+(1-r_c)\widetilde{\varpi}-\epsilon}$.

Finally, $\mathbb{E}[|\delta_{4,i}^{n,\ell}|] \lesssim \mathbb{E}[|\Delta_{p_n}\widehat{c}_i^{n,2}|^{2q'}]^{1/(2q')}\mathbb{E}[|\Delta_{p_n}\widehat{c}_{i+\ell p_n}^{n,2}|^{2q'}]^{1/(2q')}\mathbb{P}[|\Delta_{p_n}\widehat{c}_i^{n,1}|>w_n]^{1/q}$. Applying Lemma E.8 and (F.2) and taking q' close to 1, we obtain $\mathbb{E}[|\delta_{4,i}^{n,\ell}|] \lesssim \Delta_n^{-(1-2\alpha)_+}\times(p_n\Delta_n)^{2e+2-r_c\widetilde{\varpi}-\epsilon}$ for any $\epsilon>0$. Observe now that the bounds for $\mathbb{E}[|\delta_{3,i}^{n,\ell}|]$ and $\mathbb{E}[|\delta_{4,i}^{n,\ell}|]$ are of lower order than the bounds for $\mathbb{E}[|\delta_{2,i}^{n,\ell}|]$ and $\mathbb{E}[|\delta_{1,i}^{n,\ell}|]$, respectively, because $\widetilde{\varpi}<e+\frac{1}{2}$. Also observe that $(p_n\Delta_n)^{1+\widetilde{\varpi}(2-r_c)}=o((p_n\Delta_n)^{e+3/2+(1-r_c\widetilde{\varpi})-\epsilon})$ for the same reason. □

Proof of Proposition D.3. We start with (i). For $\ell \geq 0$ and a sequence $\lambda_n \to \infty$, we de-

compose $Z^{n,\ell}_{1,1}(t) = \sum_{m_1,m_2=1}^{2} \widehat{Z}^{n,\ell}_{m_1,m_2}(t)$, where $\widehat{Z}^{n,\ell}_{m_1,m_2}(t) = \frac{\Delta_n}{\gamma(p_n\Delta_n)} \sum_i \xi^{n,\ell}_{m_1,i,i} \xi^{n,\ell}_{m_2,i+\ell p_n,i}$ for $m_1, m_2 = 1, 2$ and where $\xi^{n,\ell}_{1,i,j} = \mathbb{E}[\Delta_n^{(1-2\alpha)_+} \Delta_{p_n} J^n_{1,i} \mid \mathcal{F}_{(j-\lambda_n p_n)\Delta_n}]$ and $\xi^{n,\ell}_{2,i,j} = \Delta_n^{(1-2\alpha)_+} \times (\Delta_{p_n} J^n_{1,i} - \mathbb{E}[\Delta_{p_n} J^n_{1,i} \mid \mathcal{F}_{(j-\lambda_n p_n)\Delta_n}])$. By (D.5) and (D.3), we have $\mathbb{E}[|\xi^{n,\ell}_{2,i,i}|^2 + |\xi^{n,\ell}_{2,i+\ell p_n,i}|^2]^{1/2} \lesssim \gamma(p_n\Delta_n)^{1/2}$ and $\mathbb{E}[|\xi^{n,\ell}_{1,i,i}|^2 + |\xi^{n,\ell}_{1,i+\ell p_n,i}|^2]^{1/2} = o(\gamma(p_n\Delta_n)^{1/2})$ uniformly in n and i, so applying the Cauchy–Schwarz inequality we obtain $\widehat{Z}^{n,\ell}_{1,1} + \widehat{Z}^{n,\ell}_{1,2} + \widehat{Z}^{n,\ell}_{2,1} \overset{L^1}{\Longrightarrow} 0$.

Let $\widehat{Z}^{n,\ell}_{2,2} = \widehat{Z}^{\prime n,\ell}_{2,2} + \widehat{Z}^{\prime\prime n,\ell}_{2,2}$, where $\widehat{Z}^{\prime n,\ell}_{2,2}(t) = \frac{\Delta_n}{\gamma(p_n\Delta_n)} \sum_i \mathbb{E}[\xi^{n,\ell}_{2,i,i} \xi^{n,\ell}_{2,i+\ell p_n,i} \mid \mathcal{F}_{(i-\lambda_n p_n)\Delta_n}]$ and $\widehat{Z}^{\prime\prime n,\ell}_{2,2}(t) = \frac{\Delta_n}{\gamma(p_n\Delta_n)} \sum_i \{\xi^{n,\ell}_{2,i,i} \xi^{n,\ell}_{2,i+\ell p_n,i} - \mathbb{E}[\xi^{n,\ell}_{2,i,i} \xi^{n,\ell}_{2,i+\ell p_n,i} \mid \mathcal{F}_{(i-\lambda_n p_n)\Delta_n}]\}$. The ith term in $\widehat{Z}^{\prime n,\ell}_{2,2}(t)$ is $(k_n\Delta_n)^{-2} \int_{i\Delta_n}^{(i+k_n)\Delta_n} \int_{i\Delta_n}^{(i+k_n)\Delta_n} \mathrm{Cov}(v_{s+p_n\Delta_n} - v_s, v_{r+(\ell+1)p_n\Delta_n} - v_{r+\ell p_n\Delta_n} \mid \mathcal{F}_{(i-\lambda_n p_n)\Delta_n})\, ds\, dr$, which is $\frac{\gamma(p_n\Delta_n)}{(k_n\Delta_n)^2} \zeta_{(i-\lambda_n p_n)\Delta_n} \int_{i\Delta_n}^{(i+k_n)\Delta_n} \int_{i\Delta_n}^{(i+k_n)\Delta_n} \Psi_e(\frac{r-s}{p_n\Delta_n} + \ell)\, ds\, dr + o(\gamma(p_n\Delta_n)) = \zeta_{(i-\lambda_n p_n)\Delta_n} \times \gamma(p_n\Delta_n) \int_0^1 \int_0^1 \Psi_e(\frac{r-s}{\kappa} + \ell)\, ds\, dr + o(\gamma(p_n\Delta_n))$ by (D.4). Since $\Psi_e(r) = \mathrm{Cov}(B^e_1, B^e_{1+r} - B^e_r)$ for a standard fractional Brownian motion with fractional parameter e, $\Psi_e(r) = \int_{\mathbb{R}} (g_e(1-u) - g_e(-u))(g_e(1+r-u) - g_e(r-u))\, du$ where g_e is the kernel defined in (3.14). Recalling the notation $G_e = [(e+1)K_e]^{-1} t_+^{e+1}$ and $\Delta^2_{\kappa,1} G_e(v) = G_e(v+\kappa+1) - G_e(v+1) - G_e(v+\kappa) + G_e(v)$, we obtain

$$
\begin{aligned}
&\int_0^1 \int_0^1 \Psi_e(\tfrac{r-s}{\kappa} + \ell)\, ds\, dr \\
&= \int_0^1 \int_0^1 \int_{\mathbb{R}} (g_e(1-u) - g_e(-u))(g_e(1 + \tfrac{r-s}{\kappa} + \ell - u) - g_e(\tfrac{r-s}{\kappa} + \ell - u))\, du\, ds\, dr \\
&= \kappa^{-1} \int_{\mathbb{R}} \int_0^1 \int_0^1 (g_e(1 + \tfrac{s+v}{\kappa}) - g_e(\tfrac{s+v}{\kappa}))\, ds \int_0^1 (g_e(1 + \tfrac{r+v}{\kappa} + \ell) - g_e(\tfrac{r+v}{\kappa} + \ell))\, dr\, dv \\
&= \kappa^{-1-2e} \int_{\mathbb{R}} \Delta^2_{\kappa,1} G_e(v) \Delta^2_{\kappa,1} G_e(v + \kappa\ell)\, dv = \Phi_{\kappa,e}(\ell)
\end{aligned} \tag{F.3}
$$

by Lemma F.2. Thus, $\widehat{Z}^{\prime n,\ell}_{2,2}(t) = \Phi_{\kappa,e}(\ell) \Delta_n \sum_i \zeta_{(i-\lambda_n p_n)\Delta_n} + o_p(1) \overset{L^1}{\Longrightarrow} \Phi_{\kappa,e}(\ell) \int_0^t \zeta_s\, ds + o_p(1)$ by the second property in (D.5).

Next, we write $\chi^{n,\ell}_i$ for the ith term in the definition of $\widehat{Z}^{\prime\prime n,\ell}_{2,2}$ and decompose $\widehat{Z}^{\prime\prime n,\ell}_{2,2} = \sum_{j=1}^{2\lambda_n p_n} \widehat{Z}^{n,\ell}_j$, where $\widehat{Z}^{n,\ell}_j(t) = \frac{\Delta_n}{\gamma(p_n\Delta_n)} \sum_k \chi^{n,\ell}_{(j-1)+2k\lambda_n p_n}$ and the sum ranges over all $k \geq 0$ for which $(j-1) + 2k\lambda_n p_n \leq \lfloor t/\Delta_n \rfloor - (\ell+1)p_n - k_n$. For any j, $\widehat{Z}^{n,\ell}_j(t)$ is a martingale sum as $\chi^{n,\ell}_{(j-1)+2k\lambda_n p_n}$ is $\mathcal{F}_{((j-1)+(2k+1)\lambda_n p_n)\Delta_n}$-measurable with a vanishing $\mathcal{F}_{((j-1)+(2k-1)\lambda_n p_n)\Delta_n}$-conditional expectation. By Doob's maximal inequality and (D.5), $\mathbb{E}[\sup_{t\in[0,T]} |\widehat{Z}^{n,\ell}_j(t)|^2] \lesssim \frac{\Delta_n^2}{\gamma(p_n\Delta_n)^2} \sum_k \mathbb{E}[|\chi^{n,\ell}_{(j-1)+2k\lambda_n p_n}|^2] \lesssim \Delta_n/(\lambda_n p_n)$. As a result, $\mathbb{E}[\sup_{t\in[0,T]} |\widehat{Z}^{\prime\prime n,\ell}_{2,2}|^2]^{1/2} \lesssim \lambda_n p_n \times \sqrt{\Delta_n/(\lambda_n p_n)} = \sqrt{\lambda_n p_n \Delta_n}$, which converges to 0 if $\lambda_n \to \infty$ slowly enough. This completes part (i) of the proposition.

For part (ii), we work under the assumptions of Theorem 5 and Assumption CLT'. By

(E.1), $Z^{n,\ell}_{1,1}(t) = \overline{Z}^{n,\ell}_{11}(t) + \overline{Z}^{n,\ell}_{12}(t) + \overline{Z}^{n,\ell}_{21}(t) + \overline{Z}^{n,\ell}_{22}(t)$, where for $m_1, m_2 = 1, 2$ we define $\overline{Z}^{n,\ell}_{m_1 m_2}(t) = (p_n\Delta_n)^{-2e} p_n^{-1} \Delta_n^{2(1-2\alpha)_+} \sum_i \Delta_{p_n} J^n_{1m_1,i} \Delta_{p_n} J^n_{1m_2,i+\ell p_n}$. We show in Lemma F.3 below that $\overline{Z}^{n,\ell}_{12}(t)$, $\overline{Z}^{n,\ell}_{21}(t)$ and $\overline{Z}^{n,\ell}_{22}(t)$ are negligible. To study $\overline{Z}^{n,\ell}_{11}(t)$, we use (E.1) and integration by parts to write $\Delta_n^{2(1-2\alpha)_+} \Delta_{p_n} J^n_{11,i} \Delta_{p_n} J^n_{11,i+\ell p_n} = I^{n,\ell}_{1,i} + I^{n,\ell}_{2,i} + I^{n,\ell}_{3,i}$ where

$$
\begin{aligned}
I^{n,\ell}_{1,i} &= \frac{1}{(k_n\Delta_n)^2} \int_0^{(i+p_n+k_n)\Delta_n} \Delta^2_{n,i} G_e(r) \int_0^r \Delta^2_{n,i+\ell p_n} G_e(s) \boldsymbol{\eta}_s \, d\mathbf{W}_s \boldsymbol{\eta}_r \, d\mathbf{W}_r, \\
I^{n,\ell}_{2,i} &= \frac{1}{(k_n\Delta_n)^2} \int_0^{(i+(\ell+1)p_n+k_n)\Delta_n} \Delta^2_{n,i+\ell p_n} G_e(s) \int_0^s \Delta^2_{n,i} G_e(r) \boldsymbol{\eta}_r \, d\mathbf{W}_r \boldsymbol{\eta}_s \, d\mathbf{W}_s, \\
I^{n,\ell}_{3,i} &= \frac{1}{(k_n\Delta_n)^2} \int_0^{(i+p_n+k_n)\Delta_n} \Delta^2_{n,i} G_e(r) \Delta^2_{n,i+\ell p_n} G_e(r) |\boldsymbol{\eta}_r|^2 \, dr
\end{aligned} \tag{F.4}
$$

and $|\boldsymbol{\eta}_r|^2 = \eta_r^2 + \overline{\eta}_r^2$. We obtain $\overline{Z}^{n,\ell}_{11}(t) = I^{n,\ell}_1(t) + I^{n,\ell}_2(t) + I^{n,\ell}_3(t)$, where $I^{n,\ell}_j(t) = (p_n\Delta_n)^{-2e} p_n^{-1} \sum_i I^{n,\ell}_{j,i}$ for $j = 1, 2, 3$. The first two terms, $I^{n,\ell}_1(t)$ and $I^{n,\ell}_2(t)$, can be studied analogously to the terms $M^{n,\ell,k_n}_{21}(t)$ and $M^{n,\ell,k_n}_{22}(t)$ in Chong et al. (2024). Their sum converges stably in law to the $\mathcal{Z}$ in Theorem 5. Lemma F.4 analyzes $I^{n,\ell}_3$. □

Proof of Proposition D.4. Part (i) follows from Lemmas E.3–E.5 which show that $\Delta_{p_n} J^n_{m,i} = O(\Delta_n^{-(1-2\alpha)_+} k_n^{-1/2})$ for $m = 2, 3, 4$ uniformly in i. To prove (ii), we define $Y^{n,\ell}_{j_1,j_2}(t) = (p_n\Delta_n)^{-2e} p_n^{-1} \Delta_n^{2(1-2\alpha)_+} \sum_i J^n_{j_1,i} J^n_{j_2,i+\ell p_n}$ for $\ell \geq 0$. Using the results of Appendix E, we have that $Z^{n,\ell}_{j_1,j_2}(t) = 2Y^{n,\ell}_{j_1,j_2}(t) - Y^{n,\ell-1}_{j_1,j_2}(t) - Y^{n,\ell+1}_{j_1,j_2}(t) + o_p((p_n\Delta_n)^{1/2})$ for $\ell \geq 2$ and $Z^{n,0}_{j_1,j_2}(t) + 2Z^{n,1}_{j_1,j_2}(t) = 2Y^{n,1}_{j_1,j_2}(t) - 2Y^{n,2}_{j_1,j_2}(t) + o_p((p_n\Delta_n)^{1/2})$. We show in Lemma F.5 below that the terms corresponding to $(j_1, j_2) \in \{2, 3, 4\}$ are negligible. The proof for $j_1 = 1$ or $j_2 = 1$ is different because Lemma E.2 applies to $\Delta_{p_n} J^n_{1,i}$ and not to $J^n_{1,i}$. Lemma F.6 analyzes the cases where $(j_1, j_2) = (1, 2), (2, 1)$, while Lemma F.7 considers the remaining cases. □

F.2 Auxiliary Lemmas

Lemma F.1. *We have* $\mathbb{E}[|\widehat{c}^n_i - \widehat{c}^{n,1}_i|] \lesssim \Delta_n^{(2-r)\varpi} + \Delta_n^{(1/2\wedge\alpha)-1+1/r} \mathbf{1}_{r>1}$ *uniformly in* i.

Proof. The lemma follows from the bound $\mathbb{E}[|(\Delta^n_j y)^2 \mathbf{1}_{|\Delta^n_j y| \leq v_n} - (\Delta^n_j x^c + \varepsilon^n_j)^2|] \lesssim \Delta_n^{1+(2-r)\varpi} + \Delta_n^{(1/2\wedge\alpha)+1/r} \mathbf{1}_{r>1}$. To show this, we write the left-hand side as

$$
\begin{aligned}
&\mathbb{E}[|\Delta^n_j x^c + \varepsilon^n_j|^2 \mathbf{1}_{|\Delta^n_j y| > v_n}] + \mathbb{E}[|(\Delta^n_j y)^2 - (\Delta^n_j x^c + \varepsilon^n_j)^2| \mathbf{1}_{|\Delta^n_j y| \leq v_n, |\Delta^n_j x^c + \varepsilon^n_j| \leq v_n}] \\
&\quad + \mathbb{E}[|(\Delta^n_j y)^2 - (\Delta^n_j x^c + \varepsilon^n_j)^2| \mathbf{1}_{|\Delta^n_j y| \leq v_n, |\Delta^n_j x^c + \varepsilon^n_j| > v_n}].
\end{aligned} \tag{F.5}
$$

We use Lemma E.7 and Hölder's inequality with $q, q' > 1$, $1/q + 1/q' = 1$, to bound

the first term by $\mathbb{E}[|\Delta_j^n x^c + \varepsilon_j^n|^{2q}]^{1/q}\mathbb{P}(|\Delta_j^n y| > v_n)^{1/q'} \lesssim \Delta_n^{(1\wedge 2\alpha)+(1-r\varpi)/q'}$. Taking q' close to 1 ensures that this bound is $O(\Delta_n^{(1\wedge 2\alpha)+1-r\varpi-\epsilon})$ for any small $\epsilon > 0$, which in turn is $o(\Delta_n^{1+(2-r)\varpi})$ if ϵ is small enough since $\varpi < 1/2 \wedge \alpha$.

For the second term in (F.5), note that when $|\Delta_j^n y| \le v_n$ and $|\Delta_j^n x^c + \varepsilon_j^n| \le v_n$, then we must have $|\Delta_j^n x^j| \le 2v_n$. Hence, $|(\Delta_j^n y)^2 - (\Delta_j^n x^c + \varepsilon_j^n)^2| = |(\Delta_j^n x^j)^2 + 2\Delta_j^n x^j(\Delta_j^n x^c + \varepsilon_j^n)| \lesssim (|\Delta_j^n x^j| \wedge (2v_n))^2 + (|\Delta_j^n x^j| \wedge (2v_n))|\Delta_j^n x^c + \varepsilon_j^n|$ in this case. We obtain

$$\begin{aligned}&\mathbb{E}[|(\Delta_j^n y)^2 - (\Delta_j^n x^c + \varepsilon_j^n)^2|\mathbf{1}_{|\Delta_j^n y|\le v_n,\, |\Delta_j^n x^c + \varepsilon_j^n| \le v_n}] \\ &\quad\lesssim \mathbb{E}[(|\Delta_j^n x^j| \wedge (2v_n))^2] + \mathbb{E}[(|\Delta_j^n x^j| \wedge (2v_n))^q]^{1/q}\mathbb{E}[|\Delta_j^n x^c + \varepsilon_j^n|^{q'}]^{1/q'}.\end{aligned} \tag{F.6}$$

We have $\mathbb{E}[(|\Delta_j^n x^j| \wedge (2v_n))^2] \lesssim \mathbb{E}[|\Delta_j^n x^j|^r]v_n^{2-r} \lesssim \Delta_n^{1+(2-r)\varpi}$ and $\mathbb{E}[(|\Delta_j^n x^j| \wedge (2v_n))^q]^{1/q} \lesssim \mathbb{E}[|\Delta_j^n x^j|^r]^{1/q} v_n^{(q-r)/q} \lesssim \Delta_n^{(1+\varpi(q-r))/q}$. Moreover, $\mathbb{E}[(\Delta_j^n x^c + \varepsilon_j^n)^{q'}]^{1/q'} \lesssim \Delta_n^{1/2\wedge\alpha}$, so (F.6) is $\lesssim \Delta_n^{1+(2-r)\varpi} + \Delta_n^{1/2\wedge\alpha+(1+\varpi(q-r))/q}$ Taking q is close enough to $r \vee 1$, we obtain the bound $O(\Delta_n^{1+(2-r)\varpi} + \Delta_n^{(1/2\wedge\alpha)+1+\varpi(1-r)-\epsilon})$ if $r \le 1$ and $O(\Delta_n^{1+(2-r)\varpi} + \Delta_n^{(1/2\wedge\alpha)+1/r})$ if $r > 1$. As $\varpi < 1/2 \wedge \alpha$, we obtain the bound $O(\Delta_n^{1+(2-r)\varpi})$ for $r \le 1$.

Lastly, we bound

$$\begin{aligned}&\mathbb{E}[|(\Delta_j^n y)^2 - (\Delta_j^n x^c + \varepsilon_j^n)^2|\mathbf{1}_{|\Delta_j^n y|\le v_n,\, |\Delta_j^n x^c + \varepsilon_j^n| > v_n}] \\ &\quad\lesssim \mathbb{E}[|(\Delta_j^n y)^2 - (\Delta_j^n x^c + \varepsilon_j^n)^2|^q\mathbf{1}_{|\Delta_j^n y|\le v_n}]^{1/q}\mathbb{P}(|\Delta_j^n x^c + \varepsilon_j^n| > v_n)^{1/q'}.\end{aligned} \tag{F.7}$$

Using $\mathbb{E}[(\Delta_j^n x^c + \varepsilon_j^n)^2|^q]^{1/q} \lesssim \Delta_n^{1\wedge 2\alpha}$ for the first term and Lemma E.7 for the second, we see that (F.7) is $\lesssim \Delta_n^\eta(v_n^2 + \Delta_n^{1\wedge 2\alpha}) = o(\Delta_n^{1+(2-r)\varpi})$ if η is large enough. □

Lemma F.2. *We have* $\Phi_{\kappa,e}(\ell) = \kappa^{-2e-1}\int_{-1-\kappa}^{\infty} \Delta_{\kappa,1}^2 G_e(v)\Delta_{\kappa,1}^2 G_e(v+\kappa\ell)\,dv$.

Proof. The proof uses Fourier analysis and generalizes Lemma 5.6 in Chong et al. (2024) to the case where $\kappa \neq 1$. Let $\mathcal{F}[\varphi](\xi) = \int_{\mathbb{R}} \varphi(x)e^{-ix\xi}\,d\xi$ be the Fourier transform of an L^2-function φ and extend this operator to the space of tempered distributions. Parseval's identity states that $\int_{\mathbb{R}} \varphi_1(x)\varphi_2(x)\,dx = \frac{1}{2\pi}\int_{\mathbb{R}} \mathcal{F}[\varphi_1](\xi)\overline{\mathcal{F}[\varphi_2](\xi)}\,d\xi$ for $\varphi_1, \varphi_2 \in L^2(\mathbb{R})$, where $\overline{z}$ is the complex conjugate of $z \in \mathbb{C}$. Recall that $H_\alpha(x) = |x|^\alpha$ and define $J_\alpha^\pm(x) = (\pm x)^\alpha \mathbf{1}_{\{\pm x>0\}}$ and $K_\alpha^\pm(x) = J_\alpha^+(x) + e^{\pm i\pi\alpha}J_\alpha^-(x)$. Then $z\overline{z} = |z|^2$, $H_\alpha(x) = J_\alpha^+(x) + J_\alpha^-(x)$, $J_\alpha^+(x)J_\alpha^-(x) = 0$, $\overline{K_\alpha^\pm(x)} = K_\alpha^\mp(x)$ and $|K_\alpha^\pm(x)| = H_\alpha(x)$ for all $x \in \mathbb{R}$ and $\alpha \in \mathbb{C} \setminus \{0, -1, -2, \ldots\}$. Moreover, by Section 3.2 and Example 7.1.17 in Hörmander (1990), we have

$$\mathcal{F}[J_\alpha^\pm] = \Gamma(\alpha+1)e^{\mp\frac{1}{2}i\pi(\alpha+1)}K_{-\alpha-1}^\mp, \quad \mathcal{F}[H_\alpha] = 2\Gamma(\alpha+1)\cos(\tfrac{\pi}{2}(\alpha+1))H_{-\alpha-1}. \tag{F.8}$$

The standard translation rule for Fourier transforms, $\mathcal{F}[\varphi(\cdot+a)](\xi)=e^{ia\xi}\mathcal{F}[\varphi](\xi)$, implies that $\mathcal{F}[\Delta^2_{\kappa,1}\varphi](\xi)=(e^{i\kappa\xi}-1)(e^{i\xi}-1)\mathcal{F}[\varphi](\xi)$. Similarly, the rule $\mathcal{F}[\varphi e^{ia(\cdot)}](\xi)=\mathcal{F}[\varphi](\xi-a)$ implies that $\mathcal{F}[(e^{i\kappa(\cdot)/2}-e^{-i\kappa(\cdot)/2})^2(e^{i(\cdot)/2}-e^{-i(\cdot)/2})^2\varphi](\xi)=\delta^4_{\kappa,1}\mathcal{F}[\varphi](\xi)$. Therefore,

$$\begin{aligned}
&\int_{-1-\kappa}^{\infty}\Delta^2_{\kappa,1}G_e(v)\Delta^2_{\kappa,1}G_e(v+\kappa\ell)\,dv=\int_{\mathbb{R}}\Delta^2_{\kappa,1}G_e(v)\Delta^2_{\kappa,1}G_e(v+\kappa\ell)\,dv\\
&=\frac{1}{2\pi}\int_{\mathbb{R}}(e^{i\kappa\xi}-1)(e^{i\xi}-1)\mathcal{F}[G_e](\xi)(e^{-i\kappa\xi}-1)(e^{-i\xi}-1)\overline{\mathcal{F}[G_e](\xi)}e^{-i\kappa\ell\xi}\,d\xi\\
&=\frac{K_e^{-2}}{2\pi(e+1)^2}\int_{\mathbb{R}}(e^{\frac{1}{2}i\kappa\xi}-e^{-\frac{1}{2}i\kappa\xi})^2(e^{\frac{1}{2}i\xi}-e^{-\frac{1}{2}i\xi})^2\mathcal{F}[J^+_{e+1}](\xi)\overline{\mathcal{F}[J^+_{e+1}](\xi)}e^{-i\kappa\ell\xi}\,d\xi\\
&=\frac{K_e^{-2}\Gamma(e+2)^2}{2\pi(e+1)^2}\int_{\mathbb{R}}(e^{\frac{1}{2}i\kappa\xi}-e^{-\frac{1}{2}i\kappa\xi})^2(e^{\frac{1}{2}i\xi}-e^{-\frac{1}{2}i\xi})^2H_{-2e-4}(\xi)e^{-i\kappa\ell\xi}\,d\xi\\
&=\frac{2K_e^{-2}\Gamma(e+2)^2\Gamma(-2e-3)\cos(\frac{\pi}{2}(2e+3))}{2\pi(e+1)^2}\delta^4_{\kappa,1}H_{2e+3}(\kappa\ell).
\end{aligned}$$

The definition of K_e (after (3.14)) and identities involving the gamma function imply that the fraction in the last line equals $1/[8(e+1)(e+\frac{3}{2})]$. As $\delta^4_{\kappa,1}H_{2e+3}(\kappa\ell)=\kappa^{2e+3}\delta^4_{1,1/\kappa}H_{2e+3}(\ell)$, the lemma is proved. □

For Lemmas F.3–F.7, we work under Assumption CLT'.

Lemma F.3. *We have* $(p_n\Delta_n)^{-1/2}(\overline{Z}^{n,\ell}_{12}+\overline{Z}^{n,\ell}_{21}+\overline{Z}^{n,\ell}_{22})\overset{L^1}{\Longrightarrow}0$ *for all* $\ell\geq 0$.

Proof. We first bound $\mathbb{E}[|\overline{Z}^{n,\ell}_{12}(t)|]\lesssim(p_n\Delta_n)^{-2e}p_n^{-1}\Delta_n^{2(1-2\alpha)_+}\sum_i\mathbb{E}[|\Delta_{p_n}J^n_{12,i}|^2]^{1/2}\times\mathbb{E}[|\Delta_{p_n}J^n_{11,i+\ell p_n}|^2]^{1/2}$. Then we use Lemma E.2 to bound $\mathbb{E}[|\Delta_n^{(1-2\alpha)_+}\Delta_{p_n}J^n_{12,i}|^2]^{1/2}$ and $\mathbb{E}[|\Delta_n^{(1-2\alpha)_+}\Delta_{p_n}J^n_{11,i+\ell p_n}|^2]^{1/2}$. For the former, we use the first estimate in Lemma E.2 if $i\leq p_n$ and the second if $i>p_n$. Assuming without loss of generality that $\beta<-e'$, we obtain $\mathbb{E}[|\overline{Z}^{n,\ell}_{12}(t)|]\lesssim(p_n\Delta_n)^{3/2-e+e'}+(p_n\Delta_n)^{-e+1/2+e'+\beta}+\frac{(p_n\Delta_n)^{-e+1/2+\beta+e'}}{p_n}\int_{p_n\Delta_n}^t s^{-\beta}\,ds\lesssim(p_n\Delta_n)^{3/2-e+e'}+(p_n\Delta_n)^{-e+1/2+e'+\beta}$. Since $k_n\asymp\Delta_n^{-1/2}\log(\Delta_n^{-1})$, $e'>-1$, $e<0$ and $\beta>e-e'$ by assumption, we conclude that $(p_n\Delta_n)^{-1/2}\mathbb{E}[|\overline{Z}^{n,\ell}_{12}(t)|]\to 0$. The same proof applies to $\overline{Z}^{n,\ell}_{21}(t)$. Similarly, $\mathbb{E}[|\overline{Z}^{n,\ell}_{22}(t)|]\lesssim(p_n\Delta_n)^{2-2e+2e'}+(p_n\Delta_n)^{-2e+2e'+2\beta+1}+\frac{(p_n\Delta_n)^{-2e+2e'+\beta+1}}{p_n}\times\int_{p_n\Delta_n}^t s^{-\beta}\,ds\lesssim(p_n\Delta_n)^{2-2e+2e'}+(p_n\Delta_n)^{-2e+2e'+2\beta+1}=o((p_n\Delta_n)^{1/2})$. As the previous estimates are uniform in t, the lemma is proved. □

Lemma F.4. *We have* $(p_n\Delta_n)^{-1/2}(I^{n,\ell}_3(t)-\Phi_{\kappa,e}(\ell)\int_0^t|\boldsymbol{\eta}_r|^2\,dr)\overset{L^1}{\Longrightarrow}0$ *for all* $\ell\geq 0$.

Proof. We first show that instead of $I^{n,\ell}_3(t)$, it suffices to study

$$\overline{I}^{n,\ell}_3(t)=\frac{\kappa^{-2e}}{p_n}\int_0^t\sum_{i=1-p_n-k_n}^{\infty}\Delta^2_{\kappa,1}G_e(\tfrac{i-\{r/\Delta_n\}}{k_n})\Delta^2_{\kappa,1}G_e(\tfrac{i+\ell p_n-\{r/\Delta_n\}}{k_n})|\boldsymbol{\eta}_r|^2\,dr,$$

where $\{x\}$ denotes the fractional part of x.

By definition, $\Delta_{n,i}^2 G_e(r) = (k_n\Delta_n)^{1+e}\Delta_{\kappa,1}^2 G_e(\frac{i-r/\Delta_n}{k_n})$. Hence, if we interchange integration and summation and decompose $r/\Delta_n = \lfloor r/\Delta_n\rfloor + \{r/\Delta_n\}$, we can write $I_3^{n,\ell}(t)$ as

$$\begin{aligned}
&\frac{\kappa^{-2e}}{p_n}\sum_{i=0}^{\lfloor t/\Delta_n\rfloor-(\ell+1)p_n-k_n}\int_0^{(i+p_n+k_n)\Delta_n}\Delta_{\kappa,1}^2 G_e(\tfrac{i-r/\Delta_n}{k_n})\Delta_{\kappa,1}^2 G_e(\tfrac{i+\ell p_n-r/\Delta_n}{k_n})|\boldsymbol{\eta}_r|^2\,dr\\
&=\frac{\kappa^{-2e}}{p_n}\int_0^{(\lfloor t/\Delta_n\rfloor-\ell p_n)\Delta_n}\sum_{i=(\lceil r/\Delta_n\rceil-p_n-k_n)_+}^{\lfloor t/\Delta_n\rfloor-(\ell+1)p_n-k_n}\Delta_{\kappa,1}^2 G_e(\tfrac{i-r/\Delta_n}{k_n})\Delta_{\kappa,1}^2 G_e(\tfrac{i+\ell p_n-r/\Delta_n}{k_n})|\boldsymbol{\eta}_r|^2\,dr\\
&=\frac{\kappa^{-2e}}{p_n}\int_0^{(\lfloor t/\Delta_n\rfloor-\ell p_n)\Delta_n}\sum_{i=(1-p_n-k_n)\vee(-\lfloor r/\Delta_n\rfloor)}^{\lfloor t/\Delta_n\rfloor-\lfloor r/\Delta_n\rfloor-(\ell+1)p_n-k_n}\Delta_{\kappa,1}^2 G_e(\tfrac{i-\{r/\Delta_n\}}{k_n})\\
&\quad\times\Delta_{\kappa,1}^2 G_e(\tfrac{i+\ell p_n-\{r/\Delta_n\}}{k_n})|\boldsymbol{\eta}_r|^2\,dr,
\end{aligned}\tag{F.9}$$

where the last step follows from changing the index $i-\lfloor r/\Delta_n\rfloor$ to i.

Note that $\overline{I}_3^{n,\ell}$ is obtained by removing $\vee(-\lfloor r/\Delta_n\rfloor)$ from the lower bound of i, replacing the upper bound by ∞ and extending the integral up to t. Note that the first change only applies if $r<(1+p_n+k_n)\Delta_n$ and in this case adds at most $1+p_n+k_n$ terms. As $\Delta_{\kappa,1}^2 G_e(t)$ is uniformly bounded in t, the total error is $O(p_n^{-1}(p_n\Delta_n)p_n = p_n\Delta_n = o((p_n\Delta_n)^{1/2})$. In order to replace the upper bound in the summation by ∞, we study the expression (F.9) when the sum is taken from $i=\lfloor t/\Delta_n\rfloor-\lfloor r/\Delta_n\rfloor-(\ell+1)p_n-k_n+1$ to ∞. We consider the dr-integral for $r\le t-p_n^\alpha\Delta_n$ and for $r>t-p_n^\alpha\Delta_n$ separately, where $(2e-\frac{3}{2})/(2e-1)<\alpha<\frac{3}{2}$ (such a choice is always possible for $e\in(-\frac{1}{2},0)$). In the former case, the resulting term has an absolute first moment which is $\lesssim p_n^{-1}\sum_{i=p_n^\alpha/2}^\infty(\frac{i}{k_n})^{2e-2} = O(p_n^{\alpha(2e-1)-2e+1}) = o((p_n\Delta_n)^{1/2})$ since $(2e-\frac{3}{2})/(2e-1)<\alpha$ and $p_n\asymp\Delta_n^{-1/2}\log\Delta_n^{-1}$. In the latter case, because $|\Delta_{\kappa,1}^2 G_e(\frac{i-\{r/\Delta_n\}}{k_n})\Delta_{\kappa,1}^2 G_e(\frac{i+\ell p_n-\{r/\Delta_n\}}{k_n})|\le(\frac{i}{k_n})^{2e-2}\wedge 1\le(\frac{i}{k_n})^{-1-\epsilon}\wedge 1$ for any small enough $\epsilon>0$, we obtain the bound $\lesssim p_n^{-1}k_n^{1+\epsilon}\int_{t-p_n^\alpha\Delta_n}^{(\lfloor t/\Delta_n\rfloor-\ell p_n)\Delta_n}\mathbb{E}[|\boldsymbol{\eta}_r|^2]\,dr = O(k_n^\epsilon p_n^\alpha\Delta_n)$, which is $o((p_n\Delta_n)^{1/2})$ for small enough ϵ because $\alpha<\frac{3}{2}$. Finally, by similar arguments, extending the integral to t incurs an error of the order $O(p_n^{-1}p_n\Delta_n k_n^{1+\epsilon}) = o((p_n\Delta_n)^{1/2})$. In summary, we have shown that $(p_n\Delta_n)^{-1/2}(I_3^{n,\ell}(t)-\overline{I}_3^{n,\ell}(t))\overset{L^1}{\Longrightarrow}0$.

It remains to show $(p_n\Delta_n)^{-1/2}(\overline{I}_3^{n,\ell}(t)-\Phi_{\kappa,e}(\ell)\int_0^t|\boldsymbol{\eta}_r|^2\,dr)\overset{L^1}{\Longrightarrow}0$. Lemma F.2 below shows that the constant from (B.2) satisfies $\Phi_{\kappa,e}(\ell) = \kappa^{-2e-1}\int_{-1-\kappa}^\infty\Delta_{\kappa,1}^2 G_e(v)\Delta_{\kappa,1}^2 G_e(v+\kappa\ell)\,dv$.

Hence, changing v to $u = i - vk_n$ and then u to $v = (u + \lfloor r/\Delta_n \rfloor \Delta_n)$, we obtain

$$\begin{aligned}
\Phi_{\kappa,e}(\ell) &= \kappa^{-2e-1} \sum_{i=1-p_n-k_n}^{\infty} \int_{\frac{i-1}{k_n}}^{\frac{i}{k_n}} \Delta^2_{\kappa,1} G_e(v) \Delta^2_{\kappa,1} G_e(v + \kappa\ell)\, dv \\
&= \frac{\kappa^{-2e}}{p_n} \sum_{i=1-p_n-k_n}^{\infty} \int_0^1 \Delta^2_{\kappa,1} G_e(\tfrac{i-u}{k_n}) \Delta^2_{\kappa,1} G_e(\tfrac{i+\ell p_n - u}{k_n})\, du \\
&= \frac{\kappa^{-2e}}{p_n \Delta_n} \sum_{i=1-p_n-k_n}^{\infty} \int_{\lfloor r/\Delta_n \rfloor \Delta_n}^{(\lfloor r/\Delta_n \rfloor + 1)\Delta_n} \Delta^2_{\kappa,1} G_e(\tfrac{i-\{v/\Delta_n\}}{k_n}) \Delta^2_{\kappa,1} G_e(\tfrac{i+\ell p_n - \{v/\Delta_n\}}{k_n})\, dv
\end{aligned}$$

for any $r > 0$. As a result, $\overline{I}_3^{n,\ell} - \Phi_{\kappa,e}(\ell) \int_0^t |\boldsymbol{\eta}_r|^2\, dr$ is equal to

$$\begin{aligned}
&\frac{\kappa^{-2e}}{p_n} \sum_{i=1-p_n-k_n}^{\infty} \int_0^t \Big\{ \Delta^2_{\kappa,1} G_e(\tfrac{i-\{r/\Delta_n\}}{k_n}) \Delta^2_{\kappa,1} G_e(\tfrac{i+\ell p_n - \{r/\Delta_n\}}{k_n}) \\
&\quad - \Delta_n^{-1} \int_{\lfloor r/\Delta_n \rfloor \Delta_n}^{(\lfloor r/\Delta_n \rfloor + 1)\Delta_n} \Delta^2_{\kappa,1} G_e(\tfrac{i-\{v/\Delta_n\}}{k_n}) \Delta^2_{\kappa,1}(\tfrac{i+\ell p_n - \{v/\Delta_n\}}{k_n})\, dv \Big\} |\boldsymbol{\eta}_r|^2\, dr \\
&= \frac{\kappa^{-2e}}{p_n \Delta_n} \sum_{j=1}^{\lfloor t/\Delta_n \rfloor} \sum_{i=1-p_n-k_n}^{\infty} \int_{(j-1)\Delta_n}^{j\Delta_n} \int_{(j-1)\Delta_n}^{j\Delta_n} \Big\{ \Delta^2_{\kappa,1} G_e(\tfrac{i-\{r/\Delta_n\}}{k_n}) \Delta^2_{\kappa,1} G_e(\tfrac{i+\ell p_n - \{r/\Delta_n\}}{k_n}) \\
&\quad - \Delta^2_{\kappa,1} G_e(\tfrac{i-\{v/\Delta_n\}}{k_n}) \Delta^2_{\kappa,1}(\tfrac{i+\ell p_n - \{v/\Delta_n\}}{k_n}) \Big\} |\boldsymbol{\eta}_r|^2\, dv\, dr + O_p(\Delta_n).
\end{aligned}$$

By symmetry, the $dv\, dr$-integral for $r \le v$ is the negative of the same integral for $v \le r$. Hence,

$$\begin{aligned}
&\overline{I}_3^{n,\ell} - \Phi_{\kappa,e}(\ell) \int_0^t |\boldsymbol{\eta}_r|^2\, dr \\
&= \frac{\kappa^{-2e}}{p_n \Delta_n} \sum_{j=1}^{\lfloor t/\Delta_n \rfloor} \sum_{i=1-p_n-k_n}^{\infty} \int_{(j-1)\Delta_n}^{j\Delta_n} \int_{(j-1)\Delta_n}^{r} \Big\{ \Delta^2_{\kappa,1} G_e(\tfrac{i-\{r/\Delta_n\}}{k_n}) \Delta^2_{\kappa,1} G_e(\tfrac{i+\ell p_n - \{r/\Delta_n\}}{k_n}) \\
&\quad - \Delta^2_{\kappa,1} G_e(\tfrac{i-\{v/\Delta_n\}}{k_n}) \Delta^2_{\kappa,1}(\tfrac{i+\ell p_n - \{v/\Delta_n\}}{k_n}) \Big\} (|\boldsymbol{\eta}_r|^2 - |\boldsymbol{\eta}_v|^2)\, dv\, dr + O_p(\Delta_n). \qquad \text{(F.10)}
\end{aligned}$$

The last line involves $\eta_r^2 - \eta_v^2$ and $\overline{\eta}_r^2 - \overline{\eta}_v^2$, which can be analyzed separately. We only do so for the former. Recalling (B.1), we note that the L^2-norm of the increment on $[v, r]$ of the components involving a^η and $\delta^{s,\eta}$ (recall that $\delta^{l,\eta} \equiv 0$ by Assumption CLT') is $O(\Delta_n^{1/2})$. By Lemma E.1, the same applies to $A^\eta = \int_0^t g_{e'^\eta}(t-s)\theta_s^\eta\, ds$ as $\mathbb{E}[(A_r^\eta - A_v^\eta)^2] \lesssim \Delta_n^{2e'^\eta + 2} = O(\Delta_n)$. Their contributions to (F.10) are thus $O_p((p_n\Delta_n)^{-1}\Delta_n^{-1}\Delta_n^2 k_n^{1+\epsilon}\Delta_n^{1/2}) =$

$O_p(k_n^{\epsilon}\Delta_n^{1/2}) = o_p((p_n\Delta_n)^{1/2})$. The fact that the fractional components of η are also negligible derives from the proof of Lemma 5.7 in Chong et al. (2024). □

Lemma F.5. *We have* $(p_n\Delta_n)^{-1/2}Y^{n,\ell}_{j_1,j_2} \xrightarrow{L^1} 0$ *for all* $\ell \geq 1$ *and* $(j_1,j_2) \in \{2,3,4\}$.

Proof. We start with $(j_1,j_2) = (2,2)$. By (E.2), we have $Y^{n,\ell}_{2,2}(t) = \overline{Y}^{n,\ell}_{11}(t) + \overline{Y}^{n,\ell}_{12}(t) + \overline{Y}^{n,\ell}_{21}(t) + \overline{Y}^{n,\ell}_{22}(t)$, where $\overline{Y}^{n,\ell}_{m_1m_2}(t) = (p_n\Delta_n)^{-2e}p_n^{-1}\Delta_n^{2(1-2\alpha)_+}\sum_i J^n_{2m_1,i}J^n_{2m_2,i+\ell p_n}$ for $m_1,m_2 = 1,2$. We first consider $\overline{Y}^{n,\ell}_{11}(t)$. As $J^n_{21,i}$ is $\mathcal{F}_{(i+k_n)\Delta_n}$-measurable and satisfies $\mathbb{E}[J^n_{21,i} \mid \mathcal{F}_{i\Delta_n}] = 0$ for all $i \geq 0$, we have $\mathbb{E}[J^n_{21,i}J^n_{21,i+\ell p_n} \mid \mathcal{F}_{i\Delta_n}] = 0$ for all $\ell \geq 1$. So if we decompose $\overline{Y}^{n,\ell}_{11}(t) = \sum_{j=0}^{(\ell+1)p_n-1}\overline{Y}^{n,\ell,j}_{11}(t)$ where $\overline{Y}^{n,\ell,j}_{11}(t) = (p_n\Delta_n)^{-2e}p_n^{-1}\Delta_n^{2(1-2\alpha)_+}\sum_{\iota\geq 0}J^n_{21,j+(\ell+1)p_n\iota} \times J^n_{21,j+(\ell+1)p_n\iota+\ell p_n}$ sums every $(\ell+1)p_n$th term starting from the jth one (with the convention that $J^n_{21,i} = 0$ for $i > \lfloor t/\Delta_n\rfloor - (\ell+1)p_n - k_n$), then $\overline{Y}^{n,\ell,j}_{11}(t)$ is a martingale sum. By Doob's inequality, we have $\mathbb{E}[|\overline{Y}^{n,\ell}_{11}(t)|^2]^{1/2} \leq \sum_{j=0}^{(\ell+1)p_n-1}\mathbb{E}[|\overline{Y}^{n,\ell,j}_{11}(t)|^2]^{1/2} \leq (p_n\Delta_n)^{-2e}p_n^{-1}\sum_{j=0}^{(\ell+1)p_n-1}(\sum_{\iota\geq 0}\mathbb{E}[|J^n_{21,j+(\ell+1)p_n\iota}J^n_{21,j+(\ell+1)p_n\iota+\ell p_n}|^2])^{1/2}$. By Lemma E.3, we have $\mathbb{E}[|J^n_{21,i}J^n_{21,i+\ell p_n}|^2] \lesssim k_n^{-2}$ for each $i \geq 0$. As the sum over ι has $O((p_n\Delta_n)^{-1})$ terms and $k_n, p_n \asymp \Delta_n^{-1/2}\log(\Delta_n^{-1})$, we get $\mathbb{E}[|\overline{Y}^{n,\ell}_{11}(t)|] \lesssim (p_n\Delta_n)^{-2e-1/2}k_n^{-1} = o((p_n\Delta_n)^{1/2})$. Similarly, the bound $\mathbb{E}[|\overline{Y}^{n,\ell}_{21}(t)|] \lesssim (p_n\Delta_n)^{-2e}p_n^{-1}\sum_i\mathbb{E}[(J^n_{22,i})^2]^{1/2}\mathbb{E}[(J^n_{21,i+\ell p_n})^2]^{1/2}$ and Lemma E.3 yield $\mathbb{E}[|\overline{Y}^{n,\ell}_{21}(t)|] \lesssim (p_n\Delta_n)^{-2e-1}k_n^{-1/2}\Delta_n^{1/2} = o((p_n\Delta_n)^{1/2})$ since $e < 0$. The same result for $\overline{Y}^{n,\ell}_{12}(t)$ and $\overline{Y}^{n,\ell}_{22}(t)$ derives from analogous arguments.

Next, we move to the case where $j_1, j_2 \in \{3,4\}$. We have

$$\mathbb{E}[Y^{n,\ell}_{j_1,j_2}(t)^2] \leq \frac{(p_n\Delta_n)^{-4e}}{p_n^2}\Delta_n^{4(1-2\alpha)_+}\Bigg(\sum_{|i_1-i_2|\leq(\ell+2)p_n}\mathbb{E}\Big[J^n_{j_1,i_1}J^n_{j_2,i_2}J^n_{j_1,i_1+\ell p_n}J^n_{j_2,i_2+\ell p_n}\Big] + \sum_{|i_1-i_2|>(\ell+2)p_n}\mathbb{E}\Big[\widetilde{\mathbb{E}}\Big[J^n_{j_1,i_1}J^n_{j_2,i_2}J^n_{j_1,i_1+\ell p_n}J^n_{j_2,i_2+\ell p_n}\Big]\Big]\Bigg). \tag{F.11}$$

Lemmas E.4 and E.5 show that $\mathbb{E}[J^n_{j_1,i_1}J^n_{j_2,i_2}J^n_{j_1,i_1+\ell p_n}J^n_{j_2,i_2+\ell p_n}] \leq \Delta_n^{4(2\alpha-1)}k_n^{-2}$. Hence, the first sum in (F.11) is $O((p_n\Delta_n)^{-4e-1}k_n^{-2}) = o(p_n\Delta_n)$ for $e \leq 0$. For the second sum, we decompose $\widetilde{\mathbb{E}}[J^n_{j_1,i_1}J^n_{j_2,i_2}J^n_{j_1,i_1+\ell p_n}J^n_{j_2,i_2+\ell p_n}] = \widetilde{\mathbb{E}}[J^n_{j_1,i_1}J^n_{j_1,i_1+\ell p_n}]\widetilde{\mathbb{E}}[J^n_{j_2,i_2}J^n_{j_2,i_2+\ell p_n}] + \widetilde{\mathbb{E}}[(J^n_{j_1,i_1}J^n_{j_1,i_1+\ell p_n} - \widetilde{\mathbb{E}}[J^n_{j_1,i_1}J^n_{j_1,i_1+\ell p_n}])(J^n_{j_2,i_2}J^n_{j_2,i_2+\ell p_n} - \widetilde{\mathbb{E}}[J^n_{j_2,i_2}J^n_{j_2,i_2+\ell p_n}])]$. Using Lemmas E.4 and E.5 and the fact that $\ell \geq 1$, we have $\mathbb{E}[\widetilde{\mathbb{E}}[J^n_{j,i}J^n_{j,i+\ell p_n}]^2]^{1/2} \lesssim \Delta_n^{2(2\alpha-1)}k_n^{-2}$ for any i and for $j \in \{3,4\}$. Therefore, $\mathbb{E}[|\widetilde{\mathbb{E}}[J^n_{j_1,i_1}J^n_{j_1,i_1+\ell p_n}]\widetilde{\mathbb{E}}[J^n_{j_2,i_2}J^n_{j_2,i_2+\ell p_n}]|] \lesssim \Delta_n^{4(2\alpha-1)}k_n^{-4}$.

We consider the covariance term and note that $\mathbb{E}[(J^n_{m,i}J^n_{m,i+\ell p_n} - \widetilde{\mathbb{E}}[J^n_{m,i}J^n_{m,i+\ell p_n}])^2]^{1/2} \lesssim \mathbb{E}[(J^n_{m,i}J^n_{m,i+\ell p_n})^2]^{1/2} \lesssim \mathbb{E}[(J^n_{m,i})^4]^{1/4}\mathbb{E}[(J^n_{m,i+\ell p_n})^4]^{1/4} \lesssim \Delta_n^{2(2\alpha-1)}k_n^{-1}$ for any i and $m \in \{3,4\}$. The same arguments as in the proof of Lemmas E.4 and E.5 and Assumption LLN (ii) imply $|\widetilde{\mathbb{E}}[(J^n_{j_1,i_1}J^n_{j_1,i_1+\ell p_n} - \widetilde{\mathbb{E}}[J^n_{j_1,i_1}J^n_{j_1,i_1+\ell p_n}])(J^n_{j_2,i_2}J^n_{j_2,i_2+\ell p_n} - \widetilde{\mathbb{E}}[J^n_{j_2,i_2}J^n_{j_2,i_2+\ell p_n}])]| \lesssim \exp(-\rho(|i_2 - i_1| -$

$(\ell+1)p_n))\Delta_n^{4(2\alpha-1)}k_n^{-2} \lesssim \exp(-\rho p_n)\Delta_n^{4(2\alpha-1)}k_n^{-2}$. The second term in (F.11) is therefore $O((p_n\Delta_n)^{-4e-2}k_n^{-4}) = o(p_n\Delta_n)$.

Finally, we study the cases where $j_1 = 2$ and $j_2 \in \{3,4\}$ and where $j_1 \in \{3,4\}$ and $j_2 = 2$. The two cases being similar, we only consider the former case. Since $\mathbb{E}[Y_{2,j}^{n,\ell}(t)^2] = (p_n\Delta_n)^{-4e}p_n^{-2}\Delta_n^{2(1-2\alpha)_+}\sum_{i_1,i_2}\mathbb{E}[J_{2,i_1}^n J_{2,i_2}^n J_{m,i_1+\ell p_n}^n J_{m,i_2+\ell p_n}^n]$ for $j \in \{3,4\}$, we can use Lemmas E.3–E.5 to bound $\mathbb{E}[J_{2,i_1}^n J_{2,i_2}^n J_{m,i_1+\ell p_n}^n J_{m,i_2+\ell p_n}^n]$ by $\mathbb{E}[(J_{2,i_1}^n)^4]^{1/4}\mathbb{E}[(J_{2,i_2}^n)^4]^{1/4}\mathbb{E}[(J_{m,i_1+\ell p_n}^n)^4]^{1/4}\times \mathbb{E}[(J_{m,i_2+\ell p_n}^n)^4]^{1/4} \lesssim \Delta_n^{2(2\alpha-1)}k_n^{-2}$ if $|i_1-i_2| \le 2p_n$ and by $\mathbb{E}[(J_{2,i_1}^n)^4]^{1/4}\mathbb{E}[(J_{2,i_1}^n)^4]^{1/4} \times \mathbb{E}[\widetilde{\mathbb{E}}[J_{m,i_1+\ell p_n}^n J_{m,i_2+\ell p_n}^n]^2]^{1/2} \lesssim \Delta_n^{2(2\alpha-1)}k_n^{-3}$ if $|i_1-i_2| > 2p_n$. Thus, we obtain $\mathbb{E}[Y_{2,j}^{n,\ell}(t)^2] \lesssim (p_n\Delta_n)^{-4e-1}k_n^{-2} + (p_n\Delta_n)^{-4e-2}k_n^{-3} = o(p_n\Delta_n)$. □

Lemma F.6. *We have* $(p_n\Delta_n)^{-1/2}(Z_{1,2}^{n,\ell} + Z_{2,1}^{n,\ell}) \overset{L^1}{\Longrightarrow} 0$ *for all* $\ell \ge 0$.

Proof. Analogously to Lemma F.3, one can show that $(p_n\Delta_n)^{-1/2}(Z_{1,2}^{n,\ell} - Z_{1,2}^{\prime n,\ell}) \overset{L^1}{\Longrightarrow} 0$ and $(p_n\Delta_n)^{-1/2}(Z_{2,1}^{n,\ell} - Z_{2,1}^{\prime n,\ell}) \overset{L^1}{\Longrightarrow} 0$, where for $j_1, j_2 \in \{1,2\}$, we write $Z_{j_1,j_2}^{\prime n,\ell}(t) = (p_n\Delta_n)^{-2e}p_n^{-1} \times \Delta_n^{2(1-2\alpha)_+}\sum_i \Delta_{p_n}J_{j_1 1,i}^n \Delta_{p_n}J_{j_2 1,i+\ell p_n}^n$.

We first prove $(p_n\Delta_n)^{-1/2}Z_{2,1}^{\prime n,\ell}(t) \overset{L^1}{\Longrightarrow} 0$. By the integration by parts formula for semi-martingales, we have $\Delta_{p_n}J_{21,i}^n\Delta_{p_n}J_{11,i+\ell p_n}^n = \Delta_n^{-(1-2\alpha)_+}(I_{1,i}^{\prime n} + I_{2,i}^{\prime n} + I_{3,i}^{\prime n} + I_{4,i}^{\prime n})$ where

$$
\begin{aligned}
I_{1,i}^{\prime n} &= \frac{2}{(k_n\Delta_n)^2}\int_{i\Delta_n}^{(i+p_n+k_n)\Delta_n} \zeta_i^n(s)(z_s - z_{\lfloor s/\Delta_n\rfloor\Delta_n})\int_0^s \Delta_{n,i+\ell p_n}^2 G_e(r)\boldsymbol{\eta}_r\, d\mathbf{W}_r\, \sigma_s\, dW_s,\\
I_{2,i}^{\prime n} &= \frac{2}{(k_n\Delta_n)^2}\int_{i\Delta_n}^{(i+(\ell+1)p_n+k_n)\Delta_n} \Delta_{n,i+\ell p_n}^2 G_e(r)\\
&\quad\times \int_{i\Delta_n}^{(i+p_n+k_n)\Delta_n\wedge r} \zeta_i^n(s)(z_s - z_{\lfloor s/\Delta_n\rfloor\Delta_n})\sigma_s\, dW_s\boldsymbol{\eta}_r\, d\mathbf{W}_r,\\
I_{3,i}^{\prime n} &= \frac{2}{(k_n\Delta_n)^2}\int_{i\Delta_n}^{(i+p_n+k_n)\Delta_n} \zeta_i^n(s)(z_s - z_{\lfloor s/\Delta_n\rfloor\Delta_n})\Delta_{n,i+\ell p_n}^2 G_e(s)(\eta\sigma)_{\lfloor s/\Delta_n\rfloor\Delta_n}\, ds,\\
I_{4,i}^{\prime n} &= \frac{2}{(k_n\Delta_n)^2}\int_{i\Delta_n}^{(i+p_n+k_n)\Delta_n} \zeta_i^n(s)(z_s - z_{\lfloor s/\Delta_n\rfloor\Delta_n})\\
&\quad\times \Delta_{n,i+\ell p_n}^2 G_e(s)[\eta_s\sigma_s - (\eta\sigma)_{\lfloor s/\Delta_n\rfloor\Delta_n}]\, ds
\end{aligned}
$$

and $\zeta_i^n(s) = 1$ if $(i+p_n)\Delta_n \le s < (i+p_n+k_n)\Delta_n$, $\zeta_i^n(s) = -1$ if $i\Delta_n \le s < (i+k_n)\Delta_n$ and $\zeta_i^n(s) = 0$ otherwise. Note that $I_{1,i}^{\prime n}$ and $I_{2,i}^{\prime n}$ are both $\mathcal{F}_{(i+(\ell+1)p_n+k_n)\Delta_n}$-measurable and satisfy $\mathbb{E}[I_{1,i}^{\prime n} \mid \mathcal{F}_{i\Delta_n}] = 0$ and $\mathbb{E}[I_{2,i}^{\prime n} \mid \mathcal{F}_{i\Delta_n}] = 0$. Moreover, proceeding as in Lemmas E.2 and E.3, we can show that $\mathbb{E}[(I_{1,i}^{\prime n})^2]^{1/2} + \mathbb{E}[(I_{2,i}^{\prime n})^2]^{1/2} \lesssim (k_n\Delta_n)^{e+1/2}k_n^{-1/2}$. Just as $\mathbb{E}[|\overline{Y}_{11}^{n,\ell}(t)|] = o((p_n\Delta_n)^{1/2})$ in the proof of Lemma F.5, these two terms are asymptotically negligible.

It remains to analyze $\overline{Z}_j'^{n,\ell}(t) = (p_n\Delta_n)^{-2e}p_n^{-1}\Delta_n^{(1-2\alpha)_+}\sum_i I_{j,i}'^n$ for $j=3,4$. For $j=3$, we rewrite $I_{3,i}'^n = 2(k_n\Delta_n)^{-2}\sum_{j=1}^{p_n+k_n}(\eta\sigma)_{(i+j-1)\Delta_n}\int_{(i+j-1)\Delta_n}^{(i+j)\Delta_n}\zeta_i^n(s)(z_s - z_{(i+j-1)\Delta_n})\Delta_{n,i+\ell p_n}^2 G_e(s)\,ds$ and

$$\overline{Z}_3'^{n,\ell}(t) = \frac{2(p_n\Delta_n)^{-2e}}{p_n(k_n\Delta_n)^2}\Delta_n^{(1-2\alpha)_+}\sum_{j=1}^{p_n+k_n}\Bigg(\sum_{i=0}^{\lfloor t/\Delta_n\rfloor-(\ell+1)p_n-k_n}(\eta\sigma)_{(i+j-1)\Delta_n} \times\int_{(i+j-1)\Delta_n}^{(i+j)\Delta_n}\zeta_i^n(s)(z_s - z_{(i+j-1)\Delta_n})\Delta_{n,i+\ell p_n}^2 G_e(s)\,ds\Bigg). \tag{F.12}$$

Consider for fixed j the sum over i. The ith term if $\mathcal{F}_{(i+j)\Delta_n}$-measurable with a vanishing $\mathcal{F}_{(i+j-1)\Delta_n}$-conditional expectation. Thus, the sum over i is a martingale sum. Because $\mathbb{E}[|z_s - z_{\lfloor s/\Delta_n\rfloor\Delta_n}|] = O(\Delta_n^{1/2})$ and $\int_{(i+j-1)\Delta_n}^{(i+j)\Delta_n}|\Delta_{n,i+\ell p_n}^2 G_e(s)|\,ds = O((p_n\Delta_n)^{e+2}/p_n)$, it follows that $\mathbb{E}[|\overline{Z}_3'^{n,\ell}(t)|] \lesssim (p_n\Delta_n)^{-2e-2}p_n^{-1}\sum_{j=1}^{p_n+k_n}\mathbb{E}[(\cdots)^2]^{1/2} \lesssim (p_n\Delta_n)^{-e}\Delta_n^{-1/2}\Delta_n^{1/2}/p_n = o((p_n\Delta_n)^{1/2})$, where $(\cdots)$ denotes what is inside the parentheses in (F.12).

For $j=4$, we have $|I_{4,i}'^n| \lesssim (k_n\Delta_n)^{-2}\int_{i\Delta_n}^{(i+p_n+k_n)\Delta_n}|z_s - z_{\lfloor s/\Delta_n\rfloor\Delta_n}||\Delta_{n,i+\ell p_n}^2 G_e(s)||\eta_s\sigma_s - (\eta\sigma)_{\lfloor s/\Delta_n\rfloor\Delta_n}|\,ds$. We now apply the bounds $\mathbb{E}[|z_s - z_{\lfloor s/\Delta_n\rfloor\Delta_n}|^2]^{1/2} = O(\Delta_n^{1/2})$, $\mathbb{E}[|\eta_s\sigma_s - (\eta\sigma)_{\lfloor s/\Delta_n\rfloor\Delta_n}|^2]^{1/2} = O(\Delta_n^{(e\wedge e^\eta)+1/2})$ (note (3.13), (B.1) and that σ and η are bounded away from 0 by assumption) and $\int_{i\Delta_n}^{(i+p_n+k_n)\Delta_n}|\Delta_{n,i+\ell p_n}^2 G_e(s)|\,ds = O((p_n\Delta_n)^{e+2})$ to obtain $\mathbb{E}[|\overline{Z}_4'^{n,\ell}(t)|] = O((p_n\Delta_n)^{-e-1})\Delta_n^{1+e\wedge e^\eta}$, which is $o((p_n\Delta_n)^{1/2})$ for our choice of p_n because $e^\eta \geq \frac{1}{2}e - \frac{1}{4}$ by assumption. We have shown that $(p_n\Delta_n)^{-1/2}Z_{2,1}'^{n,\ell}(t) \overset{L^1}{\Longrightarrow} 0$. A similar proof shows $(p_n\Delta_n)^{-1/2}Z_{1,2}'^{n,\ell}(t) \overset{L^1}{\Longrightarrow} 0$. □

Lemma F.7. *We have* $(p_n\Delta_n)^{-1/2}(Z_{1,3}^{n,\ell} + Z_{3,1}^{n,\ell} + Z_{1,4}^{n,\ell} + Z_{4,1}^{n,\ell}) \overset{L^1}{\Longrightarrow} 0$.

Proof. For $j\in\{3,4\}$ we have $Z_{1,j}^{n,\ell}(t) = \widetilde{Z}_1^{n,\ell,j}(t) + \widetilde{Z}_2^{n,\ell,j}(t)$ where for $m=1,2$, we define $\widetilde{Z}_m^{n,\ell,j}(t) = (p_n\Delta_n)^{-2e}p_n^{-1}\Delta_n^{2(1-2\alpha)_+}\sum_i \Delta_{p_n}J_{1m,i}^n\Delta_{p_n}J_{j,i+\ell p_n}^n$. By the same proof as for Lemma F.3, we have $\mathbb{E}[|\widetilde{Z}_2^{n,\ell,j}(t)|] \lesssim k_n^{-1/2}((p_n\Delta_n)^{1-2e+e'} + (p_n\Delta_n)^{-2e+e'+\beta})$. This is $o((p_n\Delta_n)^{1/2})$ because $k_n, p_n \asymp \Delta_n^{-1/2}\log\Delta_n^{-1}$, $e\leq 0$ and $\beta > e - e'$.

Next,

$$\begin{aligned}&\mathbb{E}[\widetilde{Z}_1^{n,\ell,j}(t)^2]\\ &= \frac{(p_n\Delta_n)^{-4e}\Delta_n^{4(1-2\alpha)_+}}{p_n^2}\Bigg(\sum_{|i_1-i_2|\leq 2p_n}\mathbb{E}\Big[\Delta_{p_n}J_{11,i_1}^n\Delta_{p_n}J_{11,i_2}^n\Delta_{p_n}J_{j,i_1+\ell p_n}^n\Delta_{p_n}J_{j,i_2+\ell p_n}^n\Big]\\ &\quad+\sum_{|i_1-i_2|>2p_n}\mathbb{E}\Big[\Delta_{p_n}J_{11,i_1}^n\Delta_{p_n}J_{11,i_2}^n\Delta_{p_n}J_{j,i_1+\ell p_n}^n\Delta_{p_n}J_{j,i_2+\ell p_n}^n\Big]\Bigg).\end{aligned}$$

By Lemmas E.2–E.5 we see that $\Delta_n^{4(1-2\alpha)_+}\mathbb{E}[\Delta_{p_n}J^n_{11,i_1}\Delta_{p_n}J^n_{11,i_2}\Delta_{p_n}J^n_{j,i_1+\ell p_n}\Delta_{p_n}J^n_{j,i_2+\ell p_n}]$ is $O((p_n\Delta_n)^{2e+1}k_n^{-1})$ if $|i_1-i_2|\leq 2p_n$ and $\leq \Delta_n^{4(1-2\alpha)_+}\mathbb{E}[|\Delta_{p_n}J^n_{11,i_1}|^4]^{1/4}\mathbb{E}[|\Delta_{p_n}J^n_{11,i_2}|^4]^{1/4}\times \mathbb{E}[\widetilde{\mathbb{E}}[\Delta_{p_n}J^n_{j,i_1+\ell p_n}\Delta_{p_n}J^n_{j,i_2+\ell p_n}]^2]^{1/2}\lesssim (p_n\Delta_n)^{2e+1}k_n^{-2}$ if $|i_1-i_2|>2p_n$. As a result, $\mathbb{E}[\widetilde{Z}_1^{n,\ell,j}(t)^2]\lesssim (p_n\Delta_n)^{-2e-1}p_n^{-2}=o((p_n\Delta_n))$. In summary, we have shown that $(p_n\Delta_n)^{-1/2}Z^{n,\ell}_{1,j}\overset{L^1}{\Longrightarrow}0$ for $j\in\{3,4\}$. The proof of $(p_n\Delta_n)^{-1/2}Z^{n,\ell}_{j,1}\overset{L^1}{\Longrightarrow}0$ is identical and therefore omitted. □

G Proof of Theorems 6 and 7

Proof of Theorem 6. Because $\mathcal{W}_n$ is positive definite, $\widehat{R}_n(E)=\frac{(\widetilde{\Phi}^n_{\kappa,E})^T\mathcal{W}_n\widetilde{V}^n_t}{(\widetilde{\Phi}^n_{\kappa,E})^T\mathcal{W}_n\widetilde{\Phi}^n_{\kappa,E}}$ is the unique minimizer of the function $R\mapsto F^n_t(E,R)$ for fixed $E\in(-\frac{1}{2},\frac{1}{2}]$. By L'Hôpital's rule, the function $E\mapsto F^n_t(E)=F^n_t(E,\widehat{R}_n(E))=(\widetilde{V}^n_t)^T\mathcal{W}_n\widetilde{V}^n_t-\frac{(\overline{\Phi}^T_{\kappa,E}\mathcal{W}_n\widetilde{V}^n_t)^2}{\overline{\Phi}^T_{\kappa,E}\mathcal{W}_n\overline{\Phi}_{\kappa,E}}$ continuously extends to $E=-\frac{1}{2}$ with $F^n_t(-\frac{1}{2})=(\widetilde{V}^n_t)^T\mathcal{W}_n\widetilde{V}^n_t-\frac{(\overline{\Phi}'^T_{\kappa,-1/2}\mathcal{W}_n\widetilde{V}^n_t)^2}{\overline{\Phi}'^T_{\kappa,-1/2}\mathcal{W}_n\overline{\Phi}'_{\kappa,-1/2}}$. We assume tacitly in the following that similar expressions are continuously extended to $E=-\frac{1}{2}$. Consider now the scaled function $\overline{F}^n_t(E)=(p_n\Delta_n)^{-4e}\Delta_n^{4(1-2\alpha)_+}F^n_t(E)$, whose minimizer is $\widehat{e}_n$ as well. By Theorem 5 and the assumption $\mathcal{W}_n\xrightarrow{\mathbb{P}}\mathcal{W}$, we have $\overline{F}^n_t(E)\xrightarrow{\mathbb{P}}\overline{F}_t(E)=\big(\overline{\Phi}^T_{\kappa,e}\mathcal{W}\overline{\Phi}_{\kappa,e}-\frac{(\overline{\Phi}^T_{\kappa,E}\mathcal{W}\overline{\Phi}_{\kappa,e})^2}{\overline{\Phi}^T_{\kappa,E}\mathcal{W}\overline{\Phi}_{\kappa,E}}\big)R_t^2$, uniformly in $E\in[-\frac{1}{2},\frac{1}{2}]$. By the Cauchy–Schwarz inequality for the inner product $(x,y)\mapsto x^T\mathcal{W}y$, we have $\overline{F}_t(E)\geq 0$. Moreover, $\overline{F}_t(E)=0$ if and only if $\overline{\Phi}_{\kappa,E}$ is a constant multiple of $\overline{\Phi}_{\kappa,e}$. Because $\overline{\Phi}_{\kappa,E}(2)/\overline{\Phi}_{\kappa,E}(1)=\Phi_{\kappa,E}(2)/(\Phi_{\kappa,E}(0)+2\Phi_{\kappa,E}(1))$ is increasing in E, the only zero of $\overline{F}_t$ is $E=e$. It follows that $\widehat{e}_n\xrightarrow{\mathbb{P}}e$. By similar arguments, we have $\overline{F}''_t(e)>0$ and $(\overline{F}^n_t)''(E)\xrightarrow{\mathbb{P}}\overline{F}''_t(E)$ uniformly in E. We conclude that the minimizer $\widehat{e}_n$ is unique with probability converging to one.

We now derive the asymptotic distribution of the GMM estimator from Theorem A.2 in Mies and Podolskij (2023).[26] Because e belongs to the interior of $[-\frac{1}{2},\frac{1}{2}]$, we have $\mathbb{P}(\nabla F^n_t(\widehat{e}_n,\widehat{R}_n)=0)\to 1$. Hence, $\check{\Theta}^n_t=(\widehat{e}_n,\Delta_n^{2(1-2\alpha)_+}\widehat{R}_n)^T$ satisfies $\mathbb{P}(\nabla\check{F}^n_t(\check{\Theta}^n_t)=0)\to 1$ where $\check{F}^n_t(E,R)=(U^n_t)^T\mathcal{W}_nU^n_t$, $U^n_t(\Theta)=\widetilde{V}^n_t-\check{\Phi}^n_{\kappa,E}R$ and $\check{\Phi}^n_{\kappa,E}=\widetilde{\Phi}^n_{\kappa,E}\Delta_n^{-2(1-2\alpha)_+}$. Explicit computations show that

$$\begin{aligned}
\partial_E U^n_t(\Theta)&=-(p_n\Delta_n)^{2E}\Delta_n^{-2(1-2\alpha)_+}\big(2\log(p_n\Delta_n)\overline{\Phi}_{\kappa,E}+\overline{\Phi}'_{\kappa,E}\big)R,\quad \partial_R U^n_t(\Theta)=-\check{\Phi}^n_{\kappa,E},\\
\partial^2_{EE}U^n_t(\Theta)&=-(p_n\Delta_n)^{2E}\Delta_n^{-2(1-2\alpha)_+}\big(4\log^2(p_n\Delta_n)\overline{\Phi}_{\kappa,E}+4\log(p_n\Delta_n)\overline{\Phi}'_{\kappa,E}+\overline{\Phi}''_{\kappa,E}\big)R,\\
\partial^2_{ER}U^n_t(\Theta)&=-(p_n\Delta_n)^{2E}\Delta_n^{-2(1-2\alpha)_+}\big(2\log(p_n\Delta_n)\overline{\Phi}_{\kappa,E}+\overline{\Phi}'_{\kappa,E}\big),\quad \partial^2_{RR}U^n_t(E,R)=0,
\end{aligned}$$

[26] Although the theorem in Mies and Podolskij (2023) only states convergence in distribution, the proof readily extends to stable convergence.

where $\overline{\Phi}''_{\kappa,E} = \partial^2_{EE}\overline{\Phi}_{\kappa,E}$. Also,

$$\begin{aligned}
\partial_E \check{F}^n_t(\Theta) &= 2\partial_E U^n_t(\Theta)^T \mathcal{W}_n U^n_t(\Theta), \quad \partial_R \check{F}^n_t(\Theta) = 2\partial_R U^n_t(\Theta)^T \mathcal{W}_n U^n_t(\Theta),\\
\partial^2_{EE} \check{F}^n_t(\Theta) &= 2\partial^2_{EE} U^n_t(\Theta)^T \mathcal{W}_n U^n_t(\Theta) + 2\partial_E U^n_t(\Theta)^T \mathcal{W}_n \partial_E U^n_t(\Theta),\\
\partial^2_{ER} \check{F}^n_t(\Theta) &= 2\partial^2_{ER} U^n_t(\Theta)^T \mathcal{W}_n U^n_t(\Theta) + 2\partial_E U^n_t(\Theta)^T \mathcal{W}_n \partial_R U^n_t(\Theta),\\
\partial^2_{RR} \check{F}^n_t(\Theta) &= 2\partial_R U^n_t(\Theta)^T \mathcal{W}_n \partial_R U^n_t(\Theta).
\end{aligned}$$

We now verify conditions (E.1) and (E.2)' in Theorem A.2 of Mies and Podolskij (2023), starting with the former. By Theorem 5, we have $(p_n\Delta_n)^{-1/2-2e}\Delta_n^{2(1-2\alpha)_+} U^n_t(e,R_t) \xrightarrow{\mathcal{L}-s} \mathcal{Z}$. Choosing the nondiagonal rate matrix $A_n = (p_n\Delta_n)^{-1/2-4e}\Delta_n^{4(1-2\alpha)_+}\left(\begin{smallmatrix} 1 & -2\log(p_n\Delta_n)R_t \\ 0 & 1 \end{smallmatrix}\right)$ transforms it into a CLT for $\nabla \check{F}^n_t(e,R_t)$:

$$A_n \nabla \check{F}^n_t(e,R_t) \xrightarrow{\mathcal{L}-s} (-2R_t(\overline{\Phi}'_{\kappa,e})^T \mathcal{W}\mathcal{Z}_t, \, -2\overline{\Phi}^T_{\kappa,e}\mathcal{W}\mathcal{Z}_t)^T = -2G_t^T \mathcal{W}\mathcal{Z}_t, \tag{G.1}$$

which verifies (E.1).

For (E.2)', we define $B_n = C_n^T = (p_n\Delta_n)^{2e+1/2}\Delta_n^{-2(1-2\alpha)_+}A_n$. Straightforward matrix algebra and Theorem 5 imply that $\sup\{\|B_n\nabla^2\check{F}_n(E,R)C_n - W\| : |E-e|+|R-R_t| \le u_n\} \xrightarrow{\mathbb{P}} 0$, where $u_n = |\log(p_n\Delta_n)|^{-2}$ and $W = 2\left(\begin{smallmatrix} (\overline{\Phi}'_{\kappa,e})^T\mathcal{W}\overline{\Phi}'_{\kappa,e}R_t^2 & (\overline{\Phi}'_{\kappa,e})^T\mathcal{W}\overline{\Phi}_{\kappa,e}R_t \\ (\overline{\Phi}'_{\kappa,e})^T\mathcal{W}\overline{\Phi}_{\kappa,e}R_t & \overline{\Phi}^T_{\kappa,e}\mathcal{W}\overline{\Phi}_{\kappa,e} \end{smallmatrix}\right) = 2G_t^T\mathcal{W}G_t$. Local uniformity follows from Theorem 5 and the fact that $(p_n\Delta_n)^{-2u_n}u_n \to 0$. We also have $\|C_n\|\|B_nA_n^{-1}\|/u_n \lesssim |\log(p_n\Delta_n)|(p_n\Delta_n)^{1/2}u_n^{-1} \lesssim |\log(p_n\Delta_n)|^3(p_n\Delta_n)^{1/2} \to 0$, which completes the proof of (E.2)'. As $D_n = (p_n\Delta_n)^{4e+1}\Delta_n^{-4(1-2\alpha)_+}A_n^T$ and G_t has full rank, Theorem 6 follows. □

Proof of Theorem 7. As we consider convergence in probability, it suffices to examine each coordinate separately. Recalling the notation from (B.3), we aim to prove for fixed $\ell,\ell' = 1,\dots,m$ that $\Sigma^{n,0}_{\ell,\ell',t} + \sum_{L=1}^{L_n} W(L,L_n)(\Sigma^{n,L}_{\ell,\ell',t} + \Sigma^{n,L}_{\ell',\ell,t}) \xrightarrow{\mathbb{P}} \mathcal{R}_{\kappa,e}(\ell,\ell')\int_0^t \zeta_s^2\,ds$, where $\Sigma^{n,L}_{\ell,\ell',t} = (p_n\Delta_n)^{-2-4e}p_n^{-1}\Delta_n^{4(1-2\alpha)_+}\widehat{\Sigma}^{n,L}_{\ell,\ell',t}$. To simplify notation, we set $W(0,L_n) = \frac{1}{2}$ and $W(-L,L_n) = W(L,L_n)$ for all $L > 0$. We also assume that $W(L,L_n) = 0$ whenever $L > L_n$. We then extend the definition of $\widehat{\Sigma}^{n,L}_{\ell,\ell',t}$ (and $\Sigma^{n,L}_{\ell,\ell',t}$) to $L \in \mathbb{Z}$ by imposing the lower bound $i \ge 0\vee(-L)$ on the index of summation in (4.3). With this notation and the bounds established in Appendix E, we can show that $\Sigma^{n,L}_{\ell,\ell',t}$ and $\Sigma^{n,-L}_{\ell,\ell',t}$ are equal up to negligible terms and therefore, $\Sigma^n_t = \sum_{L\in\mathbb{Z}} W(L,L_n)\Sigma^{n,L}_t + o_p(1)$. Finally, we restrict our attention to $\ell,\ell' \ge 2$ as the cases where $\ell = 1$ and/or $\ell' = 1$ can be handled analogously using the techniques developed in the proofs of Propositions D.3 and D.4. The remaining proof proceeds in three steps paralleling those in the proof of Theorem 5.

Step 1: Removal of Price Jumps

Recall the definitions preceding Proposition D.1 and let $\psi_i^{n,\ell,1} = \Delta_{p_n}^{\mathrm{tr}} \widehat{c}_{i+\ell p_n}^{n,1} \Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^{n,1} - m_i^{n,\ell,1}$, where $m_i^{n,\ell,1} = {K_n}^{-1} \sum_{j=i}^{i+K_n-1} \Delta_{p_n}^{\mathrm{tr}} \widehat{c}_{j+\ell p_n}^{n,1} \Delta_{p_n}^{\mathrm{tr}} \widehat{c}_j^{n,1}$. Define $\Sigma_t^{n,1} = \sum_{L\in\mathbb{Z}} W(L, L_n) \Sigma_t^{n,1,L}$, where $\Sigma_t^{n,1,L} = (\Sigma_{\ell,\ell',t}^{n,1,L})_{1\le\ell,\ell'\le m}$ and

$$\Sigma_{\ell,\ell',t}^{n,1,L} = \frac{(p_n\Delta_n)^{-1-4e}}{p_n^2} \Delta_n^{4(1-2\alpha)_+} \sum_{i=0\vee(-L)}^{N_{\ell,\ell'}^{n,L}(t)} \psi_i^{n,\ell,1} \psi_{i+L}^{n,\ell',1}. \tag{G.2}$$

Since $|\Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^n| + |\Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^{n,1}| \lesssim w_n \lesssim \Delta_n^{-(1-2\alpha)_+} (p_n\Delta_n)^{\widetilde{\varpi}}$ by Assumption LLN (iv) and $\mathbb{E}[|\Delta_{p_n}^{\mathrm{tr}} \widehat{c}_{i+\ell p_n}^n \Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^n - \Delta_{p_n}^{\mathrm{tr}} \widehat{c}_{i+\ell p_n}^{n,1} \Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^{n,1}|] \lesssim \Delta_n^{-(1-2\alpha)_+} (p_n\Delta_n)^{\widetilde{\varpi}} (\Delta_n^{(2-r)\varpi} + \Delta_n^{(1/2\wedge\alpha)-1+1/r} \mathbf{1}_{r>1} + \Delta_n^{-(1-2\alpha)_+} (p_n\Delta_n)^{1-r_c\widetilde{\varpi}+\widetilde{\varpi}})$ by Proposition D.1, we get $|\psi_i^{n,\ell}| + |\psi_i^{n,\ell,1}| \lesssim \Delta_n^{-2(1-2\alpha)_+} (p_n\Delta_n)^{2\widetilde{\varpi}}$ and $\mathbb{E}[|\psi_i^{n,\ell} - \psi_i^{n,\ell,1}|] \lesssim \Delta_n^{-(1-2\alpha)_+} (p_n\Delta_n)^{\widetilde{\varpi}} (\Delta_n^{(2-r)\varpi} + \Delta_n^{(1/2\wedge\alpha)-1+1/r} \mathbf{1}_{r>1} + \Delta_n^{-(1-2\alpha)_+} \times (p_n\Delta_n)^{1-r_c\widetilde{\varpi}+\widetilde{\varpi}})$. Because $\psi_i^{n,\ell}\psi_{i+L}^{n,\ell'} - \psi_i^{n,\ell,1}\psi_{i+L}^{n,\ell',1} = \psi_i^{n,\ell}(\psi_{i+L}^{n,\ell'} - \psi_{i+L}^{n,\ell',1}) + \psi_{i+L}^{n,\ell',1}(\psi_i^{n,\ell} - \psi_i^{n,\ell,1})$, we obtain the bound $\mathbb{E}[|\Sigma_t^n - \Sigma_t^{n,1}|] \lesssim L_n (p_n\Delta_n)^{-2-4e+3\widetilde{\varpi}} p_n^{-1} (\Delta_n^{(1-2\alpha)_+ + (2-r)\varpi} + (p_n\Delta_n)^{1-r_c\widetilde{\varpi}+\widetilde{\varpi}}) \lesssim L_n\Delta_n[(p_n\Delta_n)^{-3-4e+3\widetilde{\varpi}} \Delta_n^{(1-2\alpha)_+ + [(2-r)\varpi + ((1/2\wedge\alpha)-1+1/r)\mathbf{1}_{r>1}]} + (p_n\Delta_n)^{-2-4e+(4-r_c)\widetilde{\varpi}}]$. Since $L_n \asymp \Delta_n^{-u}$ and $p_n \asymp \Delta_n^{-1/2} \log \Delta_n^{-1}$ by assumption, we obtain $\mathbb{E}[|\Sigma_t^n - \Sigma_t^{n,1}|] \to 0$ provided that $u < [(2-r)\varpi + ((1/2\wedge\alpha)-1+1/r)\mathbf{1}_{r>1}] + (1-2\alpha)_+ + 3\widetilde{\varpi}/2 - 2e - 1/2$ and $u < (2-r_c/2)\widetilde{\varpi} - 2e$. These bounds follow from Assumption LLN (iv) and the condition $u < \widetilde{\varpi} - e + \frac{1}{4}$,

Step 2: Removal of Volatility Jumps

Recall the notation introduced before Proposition D.2. We introduce quantities that are analogous to those defined in the paragraph ending in (G.2) by using the superscript notation $(\cdot)^2$ (or $(\cdot)^{2,\mathrm{tr}}$) to indicate the replacement of $\Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^{n,1}$ by $\Delta_{p_n} \widehat{c}_i^{n,2}$ (or $\Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^{n,2} = \Delta_{p_n} \widehat{c}_i^{n,2} \mathbf{1}_{|\Delta_{p_n} \widehat{c}_i^{n,2}| \le w_n}$). We show that $\Sigma_t^{n,1} - \Sigma_t^{n,2,\mathrm{tr}}$ and $\Sigma_t^{n,2,\mathrm{tr}} - \Sigma_t^{n,2}$ are negligible.

Arguing as in the proof of Proposition D.2, one can show that $\mathbb{E}[|\Delta_{p_n}^{\mathrm{tr}} \widehat{c}_{i+\ell p_n}^{n,1} \Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^{n,1} - \Delta_{p_n}^{\mathrm{tr}} \widehat{c}_{i+\ell p_n}^{n,2} \Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^{n,2}|] \lesssim \Delta_n^{-2(1-2\alpha)_+} [(p_n\Delta_n)^{e+3/2+(1-r_c)\widetilde{\varpi}-\epsilon} + (p_n\Delta_n)^{e+1/2+1/r_c-\epsilon}]$. Because $|\Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^{n,1}| \le w_n$ and $|\Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^{n,2}| \le w_n$ by definition and $u < \widetilde{\varpi} - e + \frac{1}{4}$, $r_c < \frac{1}{e+1}$ and $\widetilde{\varpi} < \frac{1}{2} + e$ by assumption, we have $\Sigma_t^{n,1} - \Sigma_t^{n,2,\mathrm{tr}} = O_p(L_n\Delta_n((p_n\Delta_n)^{-3/2-3e+(3-r_c)\widetilde{\varpi}-\epsilon} + (p_n\Delta_n)^{-5/2-3e+2\widetilde{\varpi}+1/r_c-\epsilon}) = o_p(1)$. Next, as $\mathbb{E}[|\Delta_{p_n} \widehat{c}_i^{n,2}|^p]^{1/p} \lesssim \Delta_n^{-(1-2\alpha)_+} (p_n\Delta_n)^{e+1/2}$ for all $p \ge 1$ by (F.2) and $\widetilde{\varpi} < \frac{1}{2} + e$, we deduce that $\mathbb{P}(\Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^{n,2} \ne \Delta_{p_n} \widehat{c}_i^{n,2}) \lesssim \Delta_n^{\eta}$ for any $\eta > 0$. Choosing $\eta > 1$, we obtain $\sum_i \mathbb{P}(\Delta_{p_n}^{\mathrm{tr}} \widehat{c}_i^{n,2} \ne \Delta_{p_n} \widehat{c}_i^{n,2}) \lesssim \Delta_n^{\eta-1} \to 0$ and thus $\Sigma_t^{n,2,\mathrm{tr}} - \Sigma_t^{n,2} \xrightarrow{\mathbb{P}} 0$.

Step 3: Consistency in the Absence of Price and Volatility Jumps

We are left to prove $\Sigma_t^{n,2} \xrightarrow{\mathbb{P}} \mathcal{C}_t$. Recall that $\Delta_{p_n}\widehat{c}_i^{n,2} = \Delta_{p_n}J_{1,i}^n + \Delta_{p_n}J_{2,i}^n + \Delta_{p_n}J_{3,i}^n + \Delta_{p_n}J_{4,i}^n$ where $J_{k,i}^n$ is defined before (D.7). We decompose $\Sigma_t^{n,2} = \sum_{j_1,j_2,j_3,j_4=1}^4 \Sigma_{j_1,j_2,j_3,j_4,t}^{n,2}$, where $\Sigma_{j_1,j_2,j_3,j_4,t}^{n,2} = \sum_{L\in\mathbb{Z}} W(L,L_n)\Sigma_{j_1,j_2,j_3,j_4,t}^{n,2,L}$, $\Sigma_{j_1,j_2,j_3,j_4,t}^{n,2,L} = (\Sigma_{j_1,j_2,j_3,j_4,t}^{n,2,L,\ell,\ell'})_{1\le\ell,\ell'\le m}$,

$$\Sigma_{j_1,j_2,j_3,j_4,t}^{n,2,L,\ell,\ell'} = \frac{(p_n\Delta_n)^{-1-4e}}{p_n^2}\Delta_n^{4(1-2\alpha)_+}\sum_{i=0\vee(-L)}^{N_{\ell,\ell'}^{n,L}(t)}\psi_{j_1,j_2,i}^{n,\ell,2}\psi_{j_3,j_4,i+L}^{n,\ell',2} \tag{G.3}$$

and

$$\psi_{j,j',i}^{n,\ell,2} = \Delta_{p_n}J_{j,i}^n\Delta_{p_n}J_{j',i+\ell p_n}^n - m_{j,j',i}^{n,\ell,2}, \quad m_{j,j',i}^{n,\ell,2} = \frac{1}{K_n}\sum_{k=i}^{i+K_n-1}\Delta_{p_n}J_{j,k}^n\Delta_{p_n}J_{j',k+\ell p_n}^n. \tag{G.4}$$

The proof of Theorem 7 is completed by establishing the following two propositions. □

Proposition G.5. *If* $(j_1,j_2,j_3,j_4)\neq(1,1,1,1)$, *then* $\Sigma_{j_1,j_2,j_3,j_4,t}^{n,2}\xrightarrow{\mathbb{P}}0$.

Proposition G.6. $\Sigma_{1,1,1,1,t}^{n,2}\xrightarrow{\mathbb{P}}\mathcal{C}_t$.

Proof of Proposition G.5. First recall from Lemmas E.2–E.5 that

$$\mathbb{E}[|\Delta_{p_n}J_{j,i}^n|^p]^{1/p} \lesssim \begin{cases}\Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{e+1/2} & \text{if } j=1,\\ \Delta_n^{-(1-2\alpha)_+}k_n^{-1/2} & \text{if } j=2,3,4\end{cases}$$

for all $p\ge1$. Applying Hölder's inequality, we get $\mathbb{E}[|\Sigma_{j_1,j_2,j_3,j_4,t}^{n,2}|] \lesssim L_n\Delta_n(p_n\Delta_n)^{-3-4e}\times(p_n\Delta_n)^{q(e+1/2)}k_n^{-(4-q)/2}$, where $q=\mathbf{1}_{j_1=1}+\mathbf{1}_{j_2=1}+\mathbf{1}_{j_3=1}+\mathbf{1}_{j_4=1}$. As $L_n\asymp\Delta_n^{-u}$ and $k_n,p_n\asymp\Delta_n^{-1/2}\log\Delta_n^{-1}$, we obtain $\mathbb{E}[|\Sigma_{j_1,j_2,j_3,j_4,t}^{n,2}|]=O(\Delta_n^{-u+e(q/2-2)+1/2-\epsilon})$. By assumption, we know that $0\le q\le3$ and $u<(1-e)/2$, in which case the last term is $o(1)$. □

Proof of Proposition G.6. Recall from (E.1) that $\Delta_{p_n}J_{1,i}^n = \Delta_{p_n}J_{11,i}^n+\Delta_{p_n}J_{12,i}^n$, where $\Delta_{p_n}J_{1,i}^n = O_p(\Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{e+1/2})$ and $\Delta_{p_n}J_{12,i}^n = O_p(\Delta_n^{-(1-2\alpha)_+}(p_n\Delta_n)^{e'+1})$ by Lemma E.2. As a consequence, we have $\Delta_{p_n}J_{1,i}^n\Delta_{p_n}J_{1,i+\ell p_n}^n - \Delta_{p_n}J_{11,i}^n\Delta_{p_n}J_{11,i+\ell p_n}^n = \Delta_{p_n}J_{1,i}^n\Delta_{p_n}J_{12,i+\ell p_n}^n + \Delta_{p_n}J_{11,i+\ell p_n}^n\Delta_{p_n}J_{12,i}^n = O_p(\Delta_n^{-2(1-2\alpha)_+}(p_n\Delta_n)^{e+e'+3/2})$. Defining $\overline{\Sigma}_{1,1,1,1,t}^{n,2}$ in the same way as $\Sigma_{1,1,1,1,t}^{n,2}$ except that $\Delta_{p_n}J_{1,i}^n$ is replaced by $\Delta_{p_n}J_{11,i}^n$, we conclude that $\mathbb{E}[|\Sigma_{1,1,1,1,t}^{n,2}-\overline{\Sigma}_{1,1,1,1,t}^{n,2}|]\lesssim L_n\Delta_n(p_n\Delta_n)^{-3-4e}(p_n\Delta_n)^{2e+1}(p_n\Delta_n)^{e+e'+3/2}\lesssim L_n\Delta_n(p_n\Delta_n)^{-1/2-e+e'}$, which converges to 0 because $u<(1-e)/2$ and $p_n\asymp\Delta_n^{-1/2}\log\Delta_n^{-1}$ by assumption.

Using (F.4), we further decompose $\Delta_n^{2(1-2\alpha)_+}\Delta_{p_n}J^n_{11,i}\Delta_{p_n}J^n_{11,i+\ell p_n} = I^{n,\ell}_{12,i}+I^{n,\ell}_{3,i}$ with $I^{n,\ell}_{12,i} = I^{n,\ell}_{1,i} + I^{n,\ell}_{2,i}$. As a result, $\overline{\Sigma}^n_{1,1,1,1,t} = \overline{\Sigma}^n_{12,12,t} + \overline{\Sigma}^n_{12,3,t} + \overline{\Sigma}^n_{3,12,t} + \overline{\Sigma}^n_{3,3,t}$ where $\overline{\Sigma}^n_{k,k',t} = \sum_{L\in\mathbb{Z}} W(L,L_n)\overline{\Sigma}^{n,L}_{k,k',t}$, $\overline{\Sigma}^{n,L}_{k,k',t} = (\overline{\Sigma}^{n,L,\ell,\ell'}_{k,k',t})_{1\le\ell,\ell'\le m}$,

$$\overline{\Sigma}^{n,L,\ell,\ell'}_{k,k',t} = \frac{(p_n\Delta_n)^{-1-4e}}{p_n^2}\sum_{i=0\vee(-L)}^{N^{n,L}_{\ell,\ell'}(t)}(I^{n,\ell}_{k,i}-\overline{m}^{n,\ell}_{k,i})(I^{n,\ell'}_{k',i+L}-\overline{m}^{n,\ell'}_{k',i+L}) \tag{G.5}$$

and

$$\overline{m}^{n,\ell}_{k,i} = \frac{1}{K_n}\sum_{j=i}^{i+K_n-1} I^{n,\ell}_{k,j} \tag{G.6}$$

for $k,k'\in\{12,3\}$. Lemmas G.1 and G.2 below together with the Cauchy–Schwarz inequality imply that $\mathbb{E}[|\overline{\Sigma}^n_{12,3,t}|] \lesssim L_n(p_n\Delta_n)^{-1-4e}p_n^{-2}(k_n\Delta_n)^{4e+2}[\Delta_n^{-1}(K_n\Delta_n)^{e^\eta\wedge e^{\overline{\eta}}+1/2} + (K_n/k_n)k_n] = o(1)$ because $k_n,p_n \asymp \Delta_n^{-1/2}\log\Delta_n^{-1}$ and $u < \frac{1}{2}+(1-u')(e^\eta\wedge e^{\overline{\eta}}+1/2)$. The same conclusion holds for $\mathbb{E}[|\overline{\Sigma}^n_{3,12,t}|]$ and $\mathbb{E}[|\overline{\Sigma}^n_{3,3,t}|]$.

It remains to prove $\overline{\Sigma}^n_{12,12,t} \xrightarrow{\mathbb{P}} \mathcal{C}_t$. Using Lemma G.1 and $\sum_i$ to abbreviate the sum in (G.5), we have

$$\begin{aligned}\overline{\Sigma}^{n,L,\ell,\ell'}_{12,12,t} &= \frac{(p_n\Delta_n)^{-1-4e}}{p_n^2}\sum_i\Big(I^{n,\ell}_{12,i}I^{n,\ell'}_{12,i+L} + O((k_n\Delta_n)^{4e+2}(k_n/K_n)^{1/2})\Big)\\ &= \frac{(p_n\Delta_n)^{-1-4e}}{p_n^2}\sum_i I^{n,\ell}_{12,i}I^{n,\ell'}_{12,i+L} + O((k_nK_n)^{-1/2}).\end{aligned}$$

As $u < \frac{1}{4}+\frac{1}{2}u'$, it suffices to analyze the main term in the previous line, which we denote by $\widetilde{\Sigma}^{n,L,\ell,\ell'}_{12,12,t}$. By (F.4), $I^{n,\ell}_{12,i} = I^{n,\ell}_{1,i}+I^{n,\ell}_{2,i} = (k_n\Delta_n)^{-2}\int_0^t\int_0^s \mathcal{G}^{n,\ell}_i(s,r)\boldsymbol{\eta}_r\,d\mathbf{W}_r\boldsymbol{\eta}_s\,d\mathbf{W}_s$ where $\mathcal{G}^{n,\ell}_i(s,r) = \Delta^2_{n,i+\ell p_n}G_e(s)\Delta^2_{n,i}G_e(r)\mathbf{1}_{s\le(i+(\ell+1)p_n+k_n)\Delta_n} + \Delta^2_{n,i}G_e(s)\Delta^2_{n,i+\ell p_n}G_e(r)\mathbf{1}_{s\le(i+p_n+k_n)\Delta_n}$. Therefore, integration by parts yields

$$\begin{aligned}&(k_n\Delta_n)^4 I^{n,\ell}_{12,i}I^{n,\ell'}_{12,i+L}\\ &\quad= \int_0^t\bigg[\int_0^s \mathcal{G}^{n,\ell}_i(s,r)\boldsymbol{\eta}_r\,d\mathbf{W}_r\int_0^s\int_0^u \mathcal{G}^{n,\ell'}_{i+L}(u,v)\boldsymbol{\eta}_v\,d\mathbf{W}_v\boldsymbol{\eta}_u\,d\mathbf{W}_u\bigg]\boldsymbol{\eta}_s\,d\mathbf{W}_s\\ &\qquad+\int_0^t\bigg[\int_0^s \mathcal{G}^{n,\ell'}_{i+L}(s,r)\boldsymbol{\eta}_r\,d\mathbf{W}_r\int_0^s\int_0^u \mathcal{G}^{n,\ell}_{i}(u,v)\boldsymbol{\eta}_v\,d\mathbf{W}_v\boldsymbol{\eta}_u\,d\mathbf{W}_u\bigg]\boldsymbol{\eta}_s\,d\mathbf{W}_s\\ &\qquad+\int_0^t\int_0^s \mathcal{G}^{n,\ell}_i(s,r)\boldsymbol{\eta}_r\,d\mathbf{W}_r\int_0^s \mathcal{G}^{n,\ell'}_{i+L}(s,r)\boldsymbol{\eta}_r\,d\mathbf{W}_r|\boldsymbol{\eta}_s|^2\,ds.\end{aligned}$$

We note that the first two terms on the right-hand side are centered terms. As in the proof of Proposition D.3, one can show that they are negligible for the limit of $\widetilde{\Sigma}^{n,L,\ell,\ell'}_{12,12,t}$. The latter is determined by the quadratic variation term in the last line of the previous display. Integrating by parts one more time, we obtain

$$\int_0^s \mathcal{G}_i^{n,\ell}(s,r)\boldsymbol{\eta}_r\,d\mathbf{W}_r \int_0^s \mathcal{G}_{i+L}^{n,\ell'}(s,r)\boldsymbol{\eta}_r\,d\mathbf{W}_r = \int_0^s \mathcal{G}_i^{n,\ell}(s,r)\int_0^r \mathcal{G}_{i+L}^{n,\ell'}(s,u)\boldsymbol{\eta}_u\,d\mathbf{W}_u\boldsymbol{\eta}_r\,d\mathbf{W}_r$$
$$+\int_0^s \mathcal{G}_{i+L}^{n,\ell'}(s,r)\int_0^r \mathcal{G}_i^{n,\ell}(s,u)\boldsymbol{\eta}_u\,d\mathbf{W}_u\boldsymbol{\eta}_r\,d\mathbf{W}_r + \int_0^s \mathcal{G}_i^{n,\ell}(s,r)\mathcal{G}_{i+L}^{n,\ell'}(s,r)|\boldsymbol{\eta}_r|^2\,dr.$$

As the centered terms do not contribute to the limit, we can replace $\widetilde{\Sigma}^{n,L,\ell,\ell'}_{12,12,t}$ by

$$\Sigma'^{n,L,\ell,\ell'}_{12,12,t} = \frac{\kappa^4(p_n\Delta_n)^{-5-4e}}{p_n^2}\sum_i \int_0^t\int_0^s \mathcal{G}_i^{n,\ell}(s,r)\mathcal{G}_{i+L}^{n,\ell'}(s,r)|\boldsymbol{\eta}_r|^2\,dr|\boldsymbol{\eta}_s|^2\,ds.$$

We further simplify $\Sigma'^{n,L,\ell,\ell'}_{12,12,t}$. Let $\mathbb{G}^\ell_{\kappa,e}(s,r) = \Delta^2_{\kappa,1}G_e(s+\ell\kappa)\Delta^2_{\kappa,1}G_e(r)\mathbf{1}_{s\geq -1-(\ell+1)\kappa} + \Delta^2_{\kappa,1}G_e(s)\Delta^2_{\kappa,1}G_e(r+\ell\kappa)\mathbf{1}_{s\geq -1-\kappa}$. Since $\Delta^2_{n,i}G_e(r) = (k_n\Delta_n)^{1+e}\Delta^2_{\kappa,1}G_e(\frac{i-r/\Delta_n}{k_n})$, we can change variables to rewrite $\Sigma'^{n,L,\ell,\ell'}_{12,12,t}$ as

$$\frac{(p_n\Delta_n)^{-1}}{\kappa^{4e}p_n^2}\sum_i\int_0^t\int_0^s \mathbb{G}^\ell_{\kappa,e}(\tfrac{i-s/\Delta_n}{k_n},\tfrac{i-r/\Delta_n}{k_n})\mathbb{G}^{\ell'}_{\kappa,e}(\tfrac{i+L-s/\Delta_n}{k_n},\tfrac{i+L-r/\Delta_n}{k_n})|\boldsymbol{\eta}_r|^2\,dr|\boldsymbol{\eta}_s|^2\,ds$$
$$= \frac{\Delta_n}{\kappa^{4e+3}k_n}\sum_i\int_{-\infty}^{\frac{i}{k_n}}\int_s^{\frac{i}{k_n}} \mathbb{G}^\ell_{\kappa,e}(s,r)\mathbb{G}^{\ell'}_{\kappa,e}(s+\tfrac{L}{k_n},r+\tfrac{L}{k_n})|\boldsymbol{\eta}_{(i-rk_n)\Delta_n}|^2\,dr|\boldsymbol{\eta}_{(i-sk_n)\Delta_n}|^2\,ds.$$

Consider now the restriction of the $dr\,ds$-integral to the set where $r > \Delta_n^{-v}$ or $s > \Delta_n^{-v}$ for some small constant $v > 0$. Because $|\Delta^2_{\kappa,1}G_e(s)| \lesssim 1 \wedge |s|^{e-1}$, the restricted $dr\,ds$-integral is $O(\Delta_n^{-v(2e-1)})$. Because $\boldsymbol{\eta}$ is $(e^\eta \wedge e^{\overline{\eta}} + 1/2)$-Hölder continuous and the integrand is only nonzero if $s > -1 - [(\ell+1)\kappa \wedge (\frac{L}{k_n} + (\ell'+1)\kappa)]$, which is $O(1)$, it follows that

$$\begin{aligned}\Sigma'^{n,L,\ell,\ell'}_{12,12,t} &= \frac{\Delta_n}{\kappa^{4e+3}k_n}\sum_i |\boldsymbol{\eta}_{i\Delta_n}|^4\int_{-\infty}^{\Delta_n^{-v}}\int_s^{\Delta_n^{-v}} \mathbb{G}^\ell_{\kappa,e}(s,r)\mathbb{G}^{\ell'}_{\kappa,e}(s+\tfrac{L}{k_n},r+\tfrac{L}{k_n})\,dr\,ds\\ &\quad + O_p(\Delta_n^{-v(2e-1)}/k_n + (k_n\Delta_n^{1-v})^{e^\eta\wedge e^{\overline{\eta}}+1/2}/k_n)\\ &= \frac{\Gamma(t)}{\kappa^{4e+3}k_n}\int_{-\infty}^{\infty}\int_s^{\infty} \mathbb{G}^\ell_{\kappa,e}(s,r)\mathbb{G}^{\ell'}_{\kappa,e}(s+\tfrac{L}{k_n},r+\tfrac{L}{k_n})\,dr\,ds\\ &\quad + O_p(\Delta_n^{-v(2e-1)}/k_n + (k_n\Delta_n^{1-v})^{e^\eta\wedge e^{\overline{\eta}}+1/2}/k_n).\end{aligned}$$

Because $\Delta_n^{-u-v(2e-1)}/k_n + \Delta_n^{-u}(k_n\Delta_n^{1-v})^{e^\eta \wedge e^{\overline{\eta}}+1/2}/k_n = o(1)$ for $v = 1/8$ by our assumptions, the O_p-term in the last line of the previous display does not contribute asymptotically. Summing over L and recalling that $W(L, L_n) \to 1$ uniformly in $|L| \leq L_n$, we see that

$$\begin{aligned}
\Sigma'^{n,\ell,\ell'}_{12,12,t} &= \sum_{L\in\mathbb{Z}} W(L, L_n)\Sigma'^{n,L,\ell,\ell'}_{12,12,t} = \sum_{|L|\leq L_n} \Sigma'^{n,L,\ell,\ell'}_{12,12,t} + o_p(1) \\
&= \frac{\Gamma(t)}{\kappa^{4e+3}k_n} \sum_{|L|\leq L_n} \int_{-\infty}^{\infty}\int_{s}^{\infty} \mathbb{G}^{\ell}_{\kappa,e}(s, r)\mathbb{G}^{\ell'}_{\kappa,e}(s + \tfrac{L}{k_n}, r + \tfrac{L}{k_n})\, dr\, ds + o_p(1) \\
&= \frac{\Gamma(t)}{\kappa^{4e+3}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{\infty} \mathbb{G}^{\ell}_{\kappa,e}(s, s + u)\mathbb{G}^{\ell'}_{\kappa,e}(s + v, s + u + v)\, du\, ds\, dv + o_p(1).
\end{aligned}$$

The integral in the last line can be written as $\int_0^\infty (\int_{\mathbb{R}} \mathbb{G}^{\ell}_{\kappa,e}(s, s+u)\, ds \int_{\mathbb{R}} \mathbb{G}^{\ell'}_{\kappa,e}(v, u+v)\, dv)\, du$. As we show in Lemma G.3 below, it is equal to $R_{\kappa,e}(\ell, \ell')\kappa^{4e+3}$. □

Lemma G.1. *For all $\ell \geq 0$, $i \geq 0$ and $j \geq 4k_n$, we have $\mathbb{E}[|I^{n,\ell}_{12,i}|^2]^{1/2} \lesssim (k_n\Delta_n)^{2e+1}$, $|\mathbb{E}[I^{n,\ell}_{12,i}I^{n,\ell}_{12,i+j}]| \lesssim (k_n\Delta_n)^{4e+2}(j/k_n)^{e-1}$ and $\mathbb{E}[|\overline{m}^{n,\ell}_{12,i}|^2]^{1/2} \lesssim (k_n\Delta_n)^{2e+1}(k_n/K_n)^{1/2}$.*

Proof. By definition, $I^n_{12,i} = I^{n,\ell}_{1,i} + I^{n,\ell}_{2,i}$ where both $I^{n,\ell}_{1,i}$ and $I^{n,\ell}_{2,i}$ are similar martingale terms. We only prove the lemma with $I^{n,\ell}_{1,i}$ instead of $I^n_{12,i}$. The contributions from $I^{n,\ell}_{2,i}$ and cross-terms between $I^{n,\ell}_{1,i}$ and $I^{n,\ell}_{2,i}$ can be handled in the same way. Using the definition (F.4) and the assumption that $|\boldsymbol{\eta}|^2$ is uniformly bounded, we have $\mathbb{E}[|I^{n,\ell}_{1,i}|^2] \lesssim (k_n\Delta_n)^{-4}\int_0^{(i+p_n+k_n)\Delta_n}\int_0^r (\Delta^2_{n,i}G_e(r)\Delta^2_{n,i+\ell p_n}G_e(s))^2\, ds\, dr$. As $\Delta^2_{n,i}G_e(r) = (k_n\Delta_n)^{1+e} \times \Delta^2_{\kappa,1}G_e(\frac{i-r/\Delta_n}{k_n})$ and $|\Delta^2_{\kappa,1}G_e(r)| \lesssim 1 \wedge |r|^{e-1}$, we obtain

$$\begin{aligned}
\mathbb{E}[|I^{n,\ell}_{1,i}|^2] &\lesssim (k_n\Delta_n)^{4e}\int_0^{(i+p_n+k_n)\Delta_n}\int_0^r \Big(\Delta^2_{\kappa,1}G_e(\tfrac{i-r/\Delta_n}{k_n})\Delta^2_{\kappa,1}G_e(\tfrac{i+\ell p_n-s/\Delta_n}{k_n})\Big)^2\, ds\, dr \\
&\lesssim (k_n\Delta_n)^{4e+2}\int_{-1-\kappa}^{\infty}\int_r^{\infty} \Big(\Delta^2_{\kappa,1}G_e(r)\Delta^2_{\kappa,1}G_e(s + \ell\kappa)\Big)^2\, ds\, dr,
\end{aligned}$$

which proves the first claim of the lemma.

Regarding the second bound, note that

$$\begin{aligned}
\mathbb{E}[I^{n,\ell}_{1,i}I^{n,\ell}_{1,i+j}] = \frac{1}{(k_n\Delta_n)^4}\int_0^{(i+p_n+k_n)\Delta_n} &\Delta^2_{n,i}G_e(r)\Delta^2_{n,i+j}G_e(r) \\
&\times \mathbb{E}\bigg[\int_0^r \Delta^2_{n,i+\ell p_n}G_e(s)\boldsymbol{\eta}_s\, d\mathbf{W}_s \int_0^r \Delta^2_{n,i+j+\ell p_n}G_e(s)\boldsymbol{\eta}_s\, d\mathbf{W}_s|\boldsymbol{\eta}_r|^2\bigg]\, dr.
\end{aligned}$$

Since $|\boldsymbol{\eta}|^2$ is bounded, we can argue as above to see that the absolute value of the expectation

in the second line is $O((k_n\Delta_n)^{2e+3})$. Therefore, we get

$$\begin{aligned}
|\mathbb{E}[I^{n,\ell}_{1,i} I^{n,\ell}_{1,i+j}]| &\lesssim (k_n\Delta_n)^{2e-1} \int_0^{(i+p_n+k_n)\Delta_n} \left|\Delta^2_{n,i} G_e(r) \Delta^2_{n,i+j} G_e(r)\right| dr \\
&= (k_n\Delta_n)^{4e+1} \int_0^{(i+p_n+k_n)\Delta_n} \left|\Delta^2_{\kappa,1} G_e(\tfrac{i-r/\Delta_n}{k_n}) \Delta^2_{\kappa,1} G_e(\tfrac{i+j-r/\Delta_n}{k_n})\right| dr \\
&= (k_n\Delta_n)^{4e+2} \int_{-1-\kappa}^{i/k_n} \left|\Delta^2_{\kappa,1} G_e(r) \Delta^2_{\kappa,1} G_e(r+\tfrac{j}{k_n})\right| dr.
\end{aligned}$$

Using that $\Delta^2_{\kappa,1} G_e(r) \lesssim 1 + |r|^{e-1}$, we obtain $|\mathbb{E}[I^{n,\ell}_{1,i} I^{n,\ell}_{1,i+j}]| \lesssim (k_n\Delta_n)^{4e+2}(\frac{j}{k_n} - 1 - \kappa)^{e-1} \lesssim (k_n\Delta_n)^{4e+2}(\frac{j}{k_n})^{e-1}$ for $j > 4k_n$, which proves the second bound of the lemma. For the final one, combine the previous two to get

$$\begin{aligned}
\mathbb{E}[(\overline{m}^{n,\ell}_{12})^2] &\lesssim \frac{1}{K_n^2} \sum_{j_1=i}^{i+K_n-1} \sum_{j_2=i}^{i+K_n-1} \mathbb{E}\Big[I^{n,\ell}_{12,j_1} I^{n,\ell}_{12,j_2}\Big] \\
&\lesssim \frac{(k_n\Delta_n)^{4e+2}}{K_n^2} \left(k_n K_n + \sum_{j_1=i}^{i+K_n-1} \sum_{j_2=i}^{i+K_n-1} \Big(\frac{|j_1-j_2|}{k_n}\Big)^{e-1} \mathbf{1}_{|j_1-j_2|\geq 4k_n} \right) \\
&\lesssim (k_n\Delta_n)^{4e+2}(k_n/K_n). \qquad \square
\end{aligned}$$

Lemma G.2. *For all $i \geq 0$ and $p \geq 1$, we have $\mathbb{E}[|I^{n,\ell}_{3,i} - \overline{m}^{n,\ell}_{3,i}|^p]^{1/p} \lesssim (k_n\Delta_n)^{2e+1} \times [(K_n\Delta_n)^{e^\eta \wedge e^{\overline{\eta}}+1/2} + \frac{K_n}{k_n}(1 \wedge (i/k_n)^{2e-2})]$.*

Proof. By (F.4), (G.6), the identity $\Delta^2_{n,i} G_e(r) = (k_n\Delta_n)^{1+e} \Delta^2_{\kappa,1} G_e(\frac{i-r/\Delta_n}{k_n})$ and a change of variables, we have

$$\begin{aligned}
&I^{n,\ell}_{3,i} - \overline{m}^{n,\ell}_{3,i} \\
&= \frac{1}{K_n(k_n\Delta_n)^2} \sum_{j=i}^{i+K_n-1} \left(\int_0^{(i+p_n+k_n)\Delta_n} \Delta^2_{n,i} G_e(r) \Delta^2_{n,i+\ell p_n} G_e(r) |\boldsymbol{\eta}_r|^2 \, dr \right. \\
&\quad \left. - \int_0^{(j+p_n+k_n)\Delta_n} \Delta^2_{n,j} G_e(r) \Delta^2_{n,j+\ell p_n} G_e(r) |\boldsymbol{\eta}_r|^2 \, dr \right) \\
&= \frac{(k_n\Delta_n)^{2e}}{K_n} \sum_{j=i}^{i+K_n-1} \left(\int_0^{(i+p_n+k_n)\Delta_n} \Delta^2_{\kappa,1} G_e(\tfrac{i-r/\Delta_n}{k_n}) \Delta^2_{\kappa,1} G_e(\tfrac{i+\ell p_n-r/\Delta_n}{k_n}) |\boldsymbol{\eta}_r|^2 \, dr \right. \\
&\quad \left. - \int_0^{(j+p_n+k_n)\Delta_n} \Delta^2_{\kappa,1} G_e(\tfrac{j-r/\Delta_n}{k_n}) \Delta^2_{\kappa,1} G_e(\tfrac{j+\ell p_n-r/\Delta_n}{k_n}) |\boldsymbol{\eta}_r|^2 \, dr \right)
\end{aligned}$$

$$= \frac{(k_n\Delta_n)^{2e+1}}{K_n} \sum_{j=i}^{i+K_n-1} \bigg(\int_{-1-\kappa}^{i/k_n} \Delta_{\kappa,1}^2 G_e(r) \Delta_{\kappa,1}^2 G_e(r+\ell\kappa) |\boldsymbol{\eta}_{i\Delta_n - rk_n\Delta_n}|^2 \, dr$$
$$- \int_{-1-\kappa}^{j/k_n} \Delta_{\kappa,1}^2 G_e(r) \Delta_{\kappa,1}^2 G_e(r+\ell\kappa) |\boldsymbol{\eta}_{j\Delta_n - rk_n\Delta_n}|^2 \, dr \bigg)$$
$$= \frac{(k_n\Delta_n)^{2e+1}}{K_n} \sum_{j=i}^{i+K_n-1} \bigg(- \int_{i/k_n}^{j/k_n} \Delta_{\kappa,1}^2 G_e(r) \Delta_{\kappa,1}^2 G_e(r+\ell\kappa) |\boldsymbol{\eta}_{j\Delta_n - rk_n\Delta_n}|^2 \, dr$$
$$+ \int_{-1-\kappa}^{i/k_n} \Delta_{\kappa,1}^2 G_e(r) \Delta_{\kappa,1}^2 G_e(r+\ell\kappa) (|\boldsymbol{\eta}_{i\Delta_n - rk_n\Delta_n}|^2 - |\boldsymbol{\eta}_{j\Delta_n - rk_n\Delta_n}|^2) \, dr \bigg).$$

For the first integral, note that $|\Delta_{\kappa,1}^2 G_e(r)| \lesssim 1 \wedge |r|^{e-1}$ and thus $|\Delta_{\kappa,1}^2 G_e(r) \Delta_{\kappa,1}^2 G_e(r+\ell\kappa)| \lesssim 1 \wedge |r|^{2e-2}$. As a result,

$$\bigg| \int_{i/k_n}^{j/k_n} \Delta_{\kappa,1}^2 G_e(r) \Delta_{\kappa,1}^2 G_e(r+\ell\kappa) |\boldsymbol{\eta}_{j\Delta_n - rk_n\Delta_n}|^2 \, dr \bigg| \lesssim \int_{i/k_n}^{j/k_n} (1 \wedge r^{2e-2}) \, dr$$
$$\lesssim (1 \wedge (i/k_n)^{2e-1}) - (1 \wedge (j/k_n)^{2e-1}) \lesssim (K_n/k_n)(1 \wedge (i/k_n)^{2e-2})$$

almost surely. For the second integral, we use the fact that η and $\overline{\eta}$ are $(e^\eta \wedge e^{\overline{\eta}} + 1/2)$-Hölder regular in L^p for all $p \geq 1$ to obtain

$$\mathbb{E}\bigg[\bigg| \int_{-1-\kappa}^{i/k_n} \Delta_{\kappa,1}^2 G_e(r) \Delta_{\kappa,1}^2 G_e(r+\ell\kappa) (|\boldsymbol{\eta}_{i\Delta_n - rk_n\Delta_n}|^2 - |\boldsymbol{\eta}_{j\Delta_n - rk_n\Delta_n}|^2) \, dr \bigg|^p \bigg]^{1/p}$$
$$\lesssim (K_n\Delta_n)^{e^\eta \wedge e^{\overline{\eta}} + 1/2} \int_{-1-\kappa}^{i/k_n} |\Delta_{\kappa,1}^2 G_e(r) \Delta_{\kappa,1}^2 G_e(r+\ell\kappa)| \, dr \lesssim (K_n\Delta_n)^{e^\eta \wedge e^{\overline{\eta}} + 1/2}.$$

These bounds combine to show Lemma G.2. □

Lemma G.3. *We have* $\int_0^\infty [\int_{\mathbb{R}} \mathbb{G}_{\kappa,e}^\ell(s, s+u) \, ds \int_{\mathbb{R}} \mathbb{G}_{\kappa,e}^{\ell'}(v, v+u) \, dv] \, du = R_{\kappa,e}(\ell, \ell') \kappa^{4e+3}$.

Proof. The calculations are similar those in the proof of Lemma F.2. First note that

$$\int_{\mathbb{R}} \mathbb{G}_{\kappa,e}^\ell(s, s+u) \, ds = \int_{\mathbb{R}} \Delta_{\kappa,1}^2 G_e(s) \Big[\Delta_{\kappa,1}^2 G_e(s+u-\ell\kappa) + \Delta_{\kappa,1}^2 G_e(s+u+\ell\kappa) \Big] \, ds$$
$$= \frac{\delta_{\kappa,1}^4 H_{2e+3}(u+\kappa\ell) + \delta_{\kappa,1}^4 H_{2e+3}(u-\kappa\ell)}{8(e+1)(e+\frac{3}{2})},$$

where the second step follows from the proof of Lemma F.2. Therefore, by symmetry,

$$64(e+1)^2(e+\tfrac{3}{2})^2\int_0^\infty\left[\int_{\mathbb{R}}\mathbb{G}^{\ell}_{\kappa,e}(s,s+u)\,ds\int_{\mathbb{R}}\mathbb{G}^{\ell'}_{\kappa,e}(v,v+u)\,dv\right]du$$
$$=\int_{\mathbb{R}}\delta^4_{\kappa,1}H_{2e+3}(u+\kappa\ell)(\delta^4_{\kappa,1}H_{2e+3}(u+\kappa\ell')+\delta^4_{\kappa,1}H_{2e+3}(u-\kappa\ell'))\,du.$$

For any $\ell'\in\mathbb{R}$, the Fourier analysis results reviewed in the proof of Lemma F.2 show that

$$\begin{aligned}
&\int_{\mathbb{R}}\delta^4_{\kappa,1}H_{2e+3}(u+\kappa\ell)\delta^4_{\kappa,1}H_{2e+3}(u+\kappa\ell')\,du\\
&=\frac{1}{2\pi}\int_{\mathbb{R}}(e^{\frac{1}{2}i\kappa\xi}-e^{-\frac{1}{2}i\kappa\xi})^4(e^{\frac{1}{2}i\xi}-e^{-\frac{1}{2}i\xi})^4e^{i\kappa(\ell-\ell')\xi}\mathcal{F}[H_{2e+3}](\xi)\overline{\mathcal{F}[H_{2e+3}](\xi)}\,d\xi\\
&=\frac{2\Gamma(2e+4)^2\cos^2(\pi(e+2))}{\pi}\int_{\mathbb{R}}(e^{\frac{1}{2}i\kappa\xi}-e^{-\frac{1}{2}i\kappa\xi})^4(e^{\frac{1}{2}i\xi}-e^{-\frac{1}{2}i\xi})^4e^{-i\kappa(\ell'-\ell)\xi}H_{-4e-8}(\xi)\,d\xi\\
&=\frac{4\Gamma(2e+4)^2\cos^2(\pi(e+2))\Gamma(-4e-7)\cos(\frac{\pi}{2}(4e+7))}{\pi}\delta^8_{\kappa,1}H_{4e+7}(\kappa(\ell'-\ell))
\end{aligned}$$

for all $e\in(-\frac{1}{2},0)\backslash\{-\frac{1}{4}\}$. The final formula admits a continuous extension to the case $e=-\frac{1}{4}$. As the fraction in the last line equals $16(e+1)^2(e+\frac{3}{2})^2\Gamma(2e+2)^2(1+1/\cos(2\pi e))/\Gamma(4e+8)$, the lemma follows by noting that $\delta^8_{\kappa,1}H_{4e+7}(\kappa(\ell'-\ell))=\kappa^{4e+7}\delta^8_{1,1/\kappa}H_{4e+7}(\ell'-\ell)$. □

H Proof of Theorem 8

Proof of Theorem 8. Writing $\mathbb{E}^n_{i-1}=\mathbb{E}[\cdot\mid\mathcal{F}_{(i-1)\Delta_n}]$, we can decompose $RP^n_p(t)-IP_p(t)=Z^n_p(t)+B^n_p(t)$ where $Z^n_p(t)=\Delta_n\mu_p^{-1}\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\{|\frac{\Delta^n_i x}{\sqrt{\Delta_n}}|^p-\mathbb{E}^n_{i-1}[|\frac{\Delta^n_i x}{\sqrt{\Delta_n}}|^p]\}$ and $B^n_p(t)=\Delta_n\mu_p^{-1}\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\mathbb{E}^n_{i-1}[|\frac{\Delta^n_i x}{\sqrt{\Delta_n}}|^p]-\int_0^t\sigma^p_s\,ds$. Since $\Delta_n^{-1/2}Z^n_p(t)$ converges stably to the right-hand side of (7.2) (see Theorem 5.5.2 in Jacod and Protter (2012)), regardless of the value of e, we only have to analyze $B^n_p(t)$, and we do so assuming the strengthened conditions of Assumption CLT'. Because the proof of Theorem 5.3.6 in Jacod and Protter (2012) shows that drifts are asymptotically negligible for $B^n_p(t)$, we can further assume $b\equiv a\equiv a^\eta\equiv a^{\overline{\eta}}\equiv 0$. We also assume $\theta\equiv\theta^\eta\equiv\theta^{\overline{\eta}}\equiv 0$ as they lead to fractional terms that are similar to but are of higher order than the main terms we analyze below. That is, we assume in the following that $c_t=c_0+\int_0^t g_e(t-s)\boldsymbol{\eta}_s\,d\boldsymbol{W}_s$, $\eta^2_t=\eta^2_0+\int_0^t g_e(t-s)\boldsymbol{\eta}^\eta_s\,d\boldsymbol{W}_s$ and $\overline{\eta}^2_t=\overline{\eta}^2_0+\int_0^t g_e(t-s)\boldsymbol{\eta}^{\overline{\eta}}_s\,d\boldsymbol{W}_s$, where $\boldsymbol{\eta}_t=(\eta_t,\overline{\eta}_t,0,0)$, $\boldsymbol{\eta}^\eta_t=(\eta^\eta_t,\overline{\eta}^\eta_t,\widetilde{\eta}^\eta_t,0)$, $\boldsymbol{\eta}^{\overline{\eta}}_t=(\eta^{\overline{\eta}}_t,\overline{\eta}^{\overline{\eta}}_t,\widetilde{\eta}^{\overline{\eta}}_t,\widehat{\eta}^{\overline{\eta}}_t)$ and $\boldsymbol{W}_t=(W_t,\overline{W}_t,\widetilde{W}_t,\widehat{W}_t)^T$.

We observe that $\int_0^t\sigma^p_s\,ds=\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\int_{(i-1)\Delta_n}^{i\Delta_n}\mathbb{E}^n_{i-1}[c^{p/2}_s]\,ds+o_p(\Delta_n^{1/2})$. This is because

$\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\int_{(i-1)\Delta_n}^{i\Delta_n}(c_s^{p/2}-\mathbb{E}_{i-1}^n[c_s^{p/2}])\,ds$ is a martingale sum in which the ith term is $O_p(\Delta_n^{2e+1}) = o_p(1)$. Next, for $s\in[(i-1)\Delta_n, i\Delta_n]$, let $c_s^{n,i}=\mathbb{E}_{i-1}^n[c_s]$ and $\sigma_s^{n,i}=\sqrt{c_s^{n,i}}$. Then $c_s - c_s^{n,i} = \int_{(i-1)\Delta_n}^s g_e(s-r)\boldsymbol{\eta}_r^\eta\,d\boldsymbol{W}_r$ and, by Taylor's theorem and the assumption that c is bounded away from zero, $\sigma_s-\sigma_s^{n,i}+\frac{1}{2\sigma_s^{n,i}}(c_s-c_s^{n,i})-\frac{1}{8(\sigma_s^{n,i})^3}(c_s-c_s^{n,i})^2+o_p(\Delta_n^{2e+1})$. Defining $\phi_p(x)=|x|^p$, $\mathcal{X}_{i,0}^n=\Delta_n^{-1/2}\int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_s^{n,i}\,dW_s$, $\mathcal{X}_{i,1}^n=\frac{1}{2\sqrt{\Delta_n}}\int_{(i-1)\Delta_n}^{i\Delta_n}\frac{c_s-c_s^{n,i}}{\sigma_s^{n,i}}\,dW_s$ and $\mathcal{X}_{i,2}^n=-\frac{1}{8\sqrt{\Delta_n}}\int_{(i-1)\Delta_n}^{i\Delta_n}\frac{(c_s-c_s^{n,i})^2}{(\sigma_s^{n,i})^3}\,dW_s$, we obtain the expansion $B_p^n(t)=\Delta_n\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\sum_{j=1}^5 Q_{i,j}^n + o_p(\Delta_n^{2e+1}+\Delta_n^{1/2})$ where $Q_{i,1}^n=(\Delta_n^{-1}\int_{(i-1)\Delta_n}^{i\Delta_n}c_s^{n,i}\,ds)^{p/2}$, $Q_{i,2}^n=\frac{1}{\mu_p}\mathbb{E}_{i-1}^n[\phi_p'(\mathcal{X}_{i,0}^n)(\mathcal{X}_{i,1}^n+\mathcal{X}_{i,2}^n)+\frac{1}{2}\phi_p''(\mathcal{X}_{i,0}^n)(\mathcal{X}_{i,1}^n)^2]-Q_{i,3}^n$, $Q_{i,3}^n=\frac{p(p-2)}{4\Delta_n^2}c_{(i-1)\Delta_n}^{p/2-2}|\boldsymbol{\eta}_{(i-1)\Delta_n}|^2\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{(i-1)\Delta_n}^s\int_{(i-1)\Delta_n}^r g_e(r-u)\times g_e(s-u)\,du\,dr\,ds$, $Q_{i,4}^n=-\Delta_n^{-1}\int_{(i-1)\Delta_n}^{i\Delta_n}(c_s^{n,i})^{p/2}\,ds$ and $Q_{i,5}^n=-\frac{p(p-2)}{8\Delta_n}c_{(i-1)\Delta_n}^{p/2-2}|\boldsymbol{\eta}_{(i-1)\Delta_n}|^2\times\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{(i-1)\Delta_n}^s g_e(s-r)^2\,dr\,ds$.

Applying a second-order Taylor expansion to $Q_{i,1}^n$ and $Q_{i,4}^n$, we note that zeroth- and first-order terms cancel out, such that $Q_{i,14}^n=Q_{i,1}^n+Q_{i,4}^n=\frac{p(p-2)}{8}c_{(i-1)\Delta_n}^{p/2-2}[(\Delta_n^{-1}\int_{(i-1)\Delta_n}^{i\Delta_n}(c_s^{n,i}-c_{(i-1)\Delta_n})\,ds)^2-\Delta_n^{-1}\int_{(i-1)\Delta_n}^{i\Delta_n}(c_s^{n,i}-c_{(i-1)\Delta_n})^2\,ds]+o_p(\Delta_n^{2e+1})$. Rearranging the terms in brackets, we can equivalently write $Q_{i,14}^n=\frac{p(p-2)}{8\Delta_n^2}c_{(i-1)\Delta_n}^{p/2-2}\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{(i-1)\Delta_n}^{i\Delta_n}(c_s^{n,i}-c_{(i-1)\Delta_n})\times(c_r^{n,i}-c_s^{n,i})\,dr\,ds+o_p(\Delta_n^{2e+1})$. Consider now a slowly diverging sequence λ_n. Using a martingale argument and the fact that $Q_{i,14}^n=O_p(\Delta_n^{e+1/2})$, we have $\Delta_n\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}(Q_{i,14}^n-\mathbb{E}_{i-\lambda_n}^n[Q_{i,14}^n])=O_p((\lambda_n\Delta_n)^{1/2}\Delta_n^{e+1/2})=o_p(\sqrt{\Delta_n})$ if $\lambda_n=o(\Delta_n^{-2e-1})$. Therefore, it suffices to consider $\Delta_n\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\mathbb{E}_{i-\lambda_n}^n[Q_{i,14}^n]$ in the following.

As $c_s^{n,i}-c_{(i-1)\Delta_n}=\int_0^{(i-1)\Delta_n}[g_e(s-u)-g_e((i-1)\Delta_n-u)]\boldsymbol{\eta}_u\,d\boldsymbol{W}_u$ and $c_r^{n,i}-c_s^{n,i}=\int_0^{(i-1)\Delta_n}[g_e(r-u)-g_e(s-u)]\boldsymbol{\eta}_u\,d\boldsymbol{W}_u$, we have $\mathbb{E}_{i-\lambda_n}^n[(c_s^{n,i}-c_{(i-1)\Delta_n})(c_r^{n,i}-c_s^{n,i})]=\int_0^{(i-\lambda_n)\Delta_n}[g_e(s-u)-g_e((i-1)\Delta_n-u)]\boldsymbol{\eta}_u\,d\boldsymbol{W}_u\int_0^{(i-\lambda_n)\Delta_n}[g_e(r-u)-g_e(s-u)]\boldsymbol{\eta}_u\,d\boldsymbol{W}_u+\int_{(i-\lambda_n)\Delta_n}^{(i-1)\Delta_n}[g_e(s-u)-g_e((i-1)\Delta_n-u)][g_e(r-u)-g_e(s-u)]\mathbb{E}_{i-\lambda_n}^n[|\boldsymbol{\eta}_u|^2]\,du$. The two $\int_0^{(i-\lambda_n)\Delta_n}$-integrals have L^2-norms of the order $\Delta_n^{e+1/2}(\int_{\lambda_n}^\infty(g_e(v)-g_e(v-1))^2\,dv)^{1/2}=o(\Delta_n^{e+1/2})$ and are therefore negligible. Because the $\int_{(i-\lambda_n)\Delta_n}^{(i-1)\Delta_n}$-integral is already $O_p(\Delta_n^{2e+1})$, we can replace $\mathbb{E}_{i-\lambda_n}^n[|\boldsymbol{\eta}_u|^2]$ by $|\boldsymbol{\eta}_{(i-1)\Delta_n}|^2$. Since $\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{(i-1)\Delta_n}^{i\Delta_n}g_e((i-1)\Delta_n-u)[g_e(r-u)-g_e(s-u)]\,dr\,ds=0$ by symmetry, it follows that $\mathbb{E}_{i-\lambda_n}^n[Q_{i,13}^n]=\frac{p(p-2)}{8}c_{(i-1)\Delta_n}^{p/2-2}|\boldsymbol{\eta}_{(i-1)\Delta_n}|^2\Delta_n^{2e+1}\int_0^1\int_0^1\int_0^{\lambda_n-1}g_e(s+u)[g_e(r+u)-g_e(s+u)]\,du\,dr\,ds+o_p(\Delta_n^{2e+1})$. The identity $\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{(i-1)\Delta_n}^s\int_{(i-1)\Delta_n}^r g_e(r-u)g_e(s-u)\,du\,dr\,ds=\frac{1}{2}\Delta_n^{2e+1}\int_0^1\int_0^1\int_{-1}^0 g_e(r+u)g_e(s+u)\,du\,dr\,ds$ shows that $Q_{i,3}^n+Q_{i,5}^n=\frac{p(p-2)}{8}c_{(i-1)\Delta_n}^{p/2-2}|\boldsymbol{\eta}_{(i-1)\Delta_n}|^2\Delta_n^{2e+1}\int_0^1\int_0^1\int_{-1}^0 g_e(s+u)[g_e(r+u)-g_e(s+u)]\,du\,dr\,ds+o_p(\Delta_n^{2e+1})$. As a result, $\Delta_n\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\sum_{j=2}^5 Q_{i,j}^n=\frac{p(p-2)}{8}\Delta_n^{2e+1}\int_0^t c_s^{p/2-2}|\boldsymbol{\eta}_s|^2\,ds\int_0^1\int_0^1\int_{\mathbb{R}}g_e(s+u)[g_e(r+u)-g_e(s+u)]\,du\,dr\,ds+o_p(\Delta_n^{2e+1})$. By symmetry and the fact that $\int_{\mathbb{R}}(g_e(r+u)-g_e(s+u))^2\,du=|s-r|^{2e+1}$, the triple integral equals $-\frac{1}{2}\int_0^1\int_0^1\int_{\mathbb{R}}(g_e(r+u)-g_e(s+u))^2\,du\,dr\,ds=-\frac{1}{(2e+2)(2e+3)}$.

We have shown that

$$\Delta_n \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \sum_{j=1,3,4,5} Q^n_{i,j} = -\frac{p(p-2)\Delta_n^{2e+1}}{8(2e+2)(2e+3)} \int_0^t c_s^{p/2-2} |\boldsymbol{\eta}_s|^2 \, ds + o_p(\Delta_n^{2e+1} + \Delta_n^{1/2}). \tag{H.1}$$

It remains to analyze $Q^n_{i,2}$. We define $\boldsymbol{B}_u = (B_u, \overline{B}_u)^T$, $B_u = \Delta_n^{-1/2}(W_{(i-1)\Delta_n+u\Delta_n} - W_{(i-1)\Delta_n})$ and $\overline{B}_u = \Delta_n^{-1/2}(\overline{W}_{(i-1)\Delta_n+u\Delta_n} - \overline{W}_{(i-1)\Delta_n})$ for $u \in [0,1]$. Conditionally on $\mathcal{F}_{(i-1)\Delta_n}$, $\boldsymbol{B}$ is a two-dimensional standard Brownian motion. With this notation, we have

$$\begin{aligned}
\mathcal{X}^n_{i,0} &= \sigma_{(i-1)\Delta_n} B_1 + O_p(\Delta_n^{e+1/2}), \\
\mathcal{X}^n_{i,1} &= \frac{\Delta_n^{e+1/2}}{2\sigma_{(i-1)\Delta_n}} \int_0^1 \boldsymbol{\eta}_{(i-1)\Delta_n} \int_0^s g_e(s-r) \, d\boldsymbol{B}_r \, dB_s + o_p(\Delta_n^{e+1/2}), \\
\mathcal{X}^n_{i,2} &= -\frac{\Delta_n^{2e+1}}{8\sigma^3_{(i-1)\Delta_n}} \int_0^1 \left(\boldsymbol{\eta}_{(i-1)\Delta_n} \int_0^s g_e(s-r) \, d\boldsymbol{B}_r \right)^2 dB_s + o_p(\Delta_n^{2e+1}).
\end{aligned} \tag{H.2}$$

Next, we need some notation for multiple Wiener integrals. For $m \geq 1$ and $f \in L^2([0,1]^m)$, define $I_m(f) = \sum_\pi \int_{t_{\pi(1)} < \cdots < t_{\pi(m)}} f(t_1, \ldots, t_m) \, dB_{t_{\pi(1)}} \cdots dB_{t_{\pi(m)}}$ as a sum of iterated Itô integrals with respect to B, where π ranges over all permutations of $1, \ldots, m$. Then $\mathcal{X}^n_{i,1} = \frac{\Delta_n^{e+1/2}}{2\sigma_{(i-1)\Delta_n}}(\eta_{(i-1)\Delta_n} I_2(f_2) + \overline{\eta}_{(i-1)\Delta_n} I_1(f_1^{\overline{B}})) + o_p(\Delta_n^{e+1/2})$ where $f_2(r,s) = \frac{1}{2} g_e(|s-r|)$ and $f_1^{\overline{B}}(s) = \int_0^s g_e(s-r) \, d\overline{B}_r$. Squaring $\mathcal{X}^n_{i,1}$ and using Proposition 1.1.3 in Nualart (2006), we obtain $(\mathcal{X}^n_{i,1})^2 = \frac{\Delta_n^{2e+1}}{4c_{(i-1)\Delta_n}} \big[\eta^2_{(i-1)\Delta_n} (I_4(h_4) + 4I_2(h_2) + 2\|f_2\|^2_{L^2}) + \overline{\eta}^2_{(i-1)\Delta_n} (I_2(h_2^{\overline{B}}) + \|f_1^{\overline{B}}\|^2_{L^2}) + 2\eta_{(i-1)\Delta_n} \overline{\eta}_{(i-1)\Delta_n} I_2(f_2) I_1(f_1^{\overline{B}}) \big] + o_p(\Delta_n^{2e+1})$ where $h_4(r,s,u,v) = \frac{1}{4} g_e(|s-r|) g_e(|u-v|)$, $h_2(r,s) = \int_0^1 f_2(r,u) f_2(s,u) \, du = \frac{1}{4} \int_0^1 (\int_0^1 g_e(|s-r|) \, dr)^2 \, ds$ and $h_2^{\overline{B}}(r,s) = f_1^{\overline{B}}(r) f_1^{\overline{B}}(s)$.

If we first condition on B in the analysis of $\mathbb{E}^n_{i-1}[\phi_p''(\mathcal{X}^n_{i,0})(\mathcal{X}^n_{i,1})^2]$, we can see that the mixed term, $2\eta_{(i-1)\Delta_n} \overline{\eta}_{(i-1)\Delta_n} I_2(f_2) I_1(f_1^{\overline{B}})$, has a zero conditional mean due to $I_1(f_1^{\overline{B}})$. Therefore, only the first two terms in the expansion of $(\mathcal{X}^n_{i,1})^2$ contribute. By Lemma H.1, $\mathbb{E}^n_{i-1}[\phi_p''(\mathcal{X}^n_{i,0}) I_4(h_4)] \sim p(p-1) c^{p/2-1}_{(i-1)\Delta_n} \mathbb{E}^n_{i-1}[|B_1|^{p-2} \mathbb{E}^n_{i-1}[I_4(h_4) \mid B_1]] = p(p-1) c^{p/2-1}_{(i-1)\Delta_n} \times \|h_4\|_{L^1} \mathbb{E}[|B_1|^{p-2}(B_1^4 - 6B_1^2 + 3)] = p(p-2)(p-4) \mu_p c^{p/2-1}_{(i-1)\Delta_n} \|h_4\|_{L^1}$. Similarly, we have $\mathbb{E}^n_{i-1}[\phi_p''(\mathcal{X}^n_{i,0}) 4I_2(h_2)] \sim 4p(p-1) c^{p/2-1}_{(i-1)\Delta_n} \|h_2\|_{L^1} \mathbb{E}[|B_1|^{p-2}(B_1^2 - 1)] = 4p(p-2) \mu_p c^{p/2-1}_{(i-1)\Delta_n} \|h_2\|_{L^1}$, $\mathbb{E}^n_{i-1}[\phi_p''(\mathcal{X}^n_{i,0}) 2\|f_2\|^2_{L^2}] \sim 2p\mu_p c^{p/2-1}_{(i-1)\Delta_n} \|f_2\|^2_{L^2}$, $\mathbb{E}^n_{i-1}[\phi_p''(\mathcal{X}^n_{i,0}) I_2(h_2^{\overline{B}})] \sim p(p-2) \mu_p c^{p/2-1}_{(i-1)\Delta_n} \times \mathbb{E}^n_{i-1}[(\int_0^1 \int_0^r g_e(r-u) \, d\overline{B}_u \, dr)^2] = p(p-2) \mu_p c^{p/2-1}_{(i-1)\Delta_n} \int_0^1 (\int_u^1 g_e(r-u) \, dr)^2 \, du$ as well as $\mathbb{E}^n_{i-1}[\phi_p''(\mathcal{X}^n_{i,0}) \|f_1^{\overline{B}}\|^2_{L^2}] \sim p\mu_p c^{p/2-1}_{(i-1)\Delta_n} \int_0^1 \mathbb{E}^n_{i-1}[(\int_0^s g_e(s-r) \, d\overline{B}_r)^2] \, ds \sim 2p\mu_p c^{p/2-1}_{(i-1)\Delta_n} \|f_2\|^2_{L^2}$. In the same way, one can show that $\mathbb{E}^n_{i-1}[\phi_p'(\mathcal{X}^n_{i,0}) \mathcal{X}^n_{i,2}] \sim -\frac{1}{8\sigma^3_{(i-1)\Delta_n}} \Delta_n^{2e+1} (\eta^2_{(i-1)\Delta_n} \mathbb{E}^n_{i-1}[\phi_p'(\mathcal{X}^n_{i,0}) \times [I_3(k_3) + I_1(k_1)]] + \overline{\eta}^2_{(i-1)\Delta_n} \mathbb{E}^n_{i-1}[\phi_p'(\mathcal{X}^n_{i,0}) I_1(k_1)])$, where $k_1(s) = \int_0^s g_e(s-r)^2 \, dr$, $k_3(r,s,u) =$

$g_e(s-r)g_e(s-u)$, $\mathbb{E}^n_{i-1}[\phi_p'(\mathcal{X}^n_{i,0})I_3(k_3)] \sim pc^{p/2-1/2}_{(i-1)\Delta_n}\mathbb{E}[|B_1|^{p-2}B_1(B_1^3-3B_1)]\|k_3\|_{L^1} = p(p-2)c^{p/2-1/2}_{(i-1)\Delta_n}\mu_p\|k_3\|_{L^1}$ and $\mathbb{E}^n_{i-1}[\phi_p'(\mathcal{X}^n_{i,0})I_1(k_1)] \sim pc^{p/2-1/2}_{(i-1)\Delta_n}\mu_p\|k_1\|_{L^1}$.

For the first part of $Q^n_{i,2}$, note that $\mathbb{E}^n_{i-1}[\phi_p'(\mathcal{X}^n_{i,0})\mathcal{X}^n_{i,1}] = pc^{p/2-1/2}_{(i-1)\Delta_n}\mathbb{E}^n_{i-1}[|B_1|^{p-2}B_1\mathcal{X}^n_{i,1}] + p(p-1)c^{p/2-1}_{(i-1)\Delta_n}\mathbb{E}^n_{i-1}[|B_1|^{p-2}\int_0^1(\sigma^{n,i}_s-\sigma_{(i-1)\Delta_n})\,dB_s\,\mathcal{X}^n_{i,1}] + o_p(\Delta_n^{2e+1})$. Conditionally on $\mathcal{F}_{(i-1)\Delta_n}$, the product $|B_1|^{p-2}\int_0^1(\sigma^{n,i}_s-\sigma_{(i-1)\Delta_n})\,dB_s$ belongs to the direct sum of Wiener chaoses of odd orders, while the main component of $\mathcal{X}^n_{i,1}$ in (H.2) consists of one term in the second Wiener chaos with respect to B and one term that involves a single integral with respect to $\overline{B}$. These terms yield a zero covariance with $|B_1|^{p-2}\int_0^1(\sigma^{n,i}_s-\sigma_{(i-1)\Delta_n})\,dB_s$, whence $\mathbb{E}^n_{i-1}[\phi_p'(\mathcal{X}^n_{i,0})\mathcal{X}^n_{i,1}] = pc^{p/2-1/2}_{(i-1)\Delta_n}\mathbb{E}^n_{i-1}[|B_1|^{p-2}B_1\mathcal{X}^n_{i,1}] + o_p(\Delta_n^{2e+1})$.

Expanding $\mathcal{X}^n_{i,1}$ in (H.2) one level further, we obtain $\mathbb{E}^n_{i-1}[\phi_p'(\mathcal{X}^n_{i,0})\mathcal{X}^n_{i,1}] = \frac{1}{2}pc^{p/2-1/2}_{(i-1)\Delta_n}\Delta_n^{e+1/2}\times \mathbb{E}^n_{i-1}[|B_1|^{p-2}B_1\int_0^1(\sigma^{n,i}_{(i-1+s)\Delta_n})^{-1}\int_0^s g_e(s-r)(\mathbb{E}^n_{i-1}[\eta_{(i-1+r)\Delta_n}]\,dB_r + \overline{\eta}_{(i-1+r)\Delta_n}\,d\overline{B}_r)\,dB_s] + \frac{1}{4}p\frac{\eta^\eta_{(i-1)\Delta_n}}{\eta_{(i-1)\Delta_n}}c^{p/2-1}_{(i-1)\Delta_n}\Delta_n^{2e+1}\mathbb{E}^n_{i-1}[|B_1|^{p-2}B_1\int_0^1\int_0^s\int_0^r g_e(s-r)g_e(r-u)\,dB_u\,dB_r\,dB_s] + o_p(\Delta_n^{2e+1})$. The first term on the right-hand side is identically zero. This is because conditioning on B removes the part involving $\overline{B}$ and what remains consists of $|B_1|^{p-2}B_1$, an odd function of B_1, and a double integral, the two of which are uncorrelated. Thus, by Lemma H.1, $\mathbb{E}^n_{i-1}[\phi_p'(\mathcal{X}^n_{i,0})\mathcal{X}^n_{i,1}] \sim \frac{1}{4}p\frac{\eta^\eta_{(i-1)\Delta_n}}{\eta_{(i-1)\Delta_n}}c^{p/2-1}_{(i-1)\Delta_n}\Delta_n^{2e+1}\mathbb{E}[|B_1|^{p-2}B_1(B_1^3-3B_1)]\|l\|_{L^1} = \frac{1}{4}p(p-2)\mu_p\times \frac{\eta^\eta_{(i-1)\Delta_n}}{\eta_{(i-1)\Delta_n}}c^{p/2-1}_{(i-1)\Delta_n}\Delta_n^{2e+1}\|l\|_{L^1}$, with $l(r,s,u) = g_e(s-r)g_e(r-u)$.

Since $Q^n_{i,3} = \frac{1}{8}p(p-2)c^{p/2-2}_{(i-1)\Delta_n}|\boldsymbol{\eta}_{(i-1)\Delta_n}|^2\Delta_n^{2e+1}\int_0^1(\int_0^s g_e(s-r)\,dr)^2\,ds + o_p(\Delta_n^{2e+1}) = \frac{1}{8}p(p-2)c^{p/2-2}_{(i-1)\Delta_n}|\boldsymbol{\eta}_{(i-1)\Delta_n}|^2\Delta_n^{2e+1}\|k_3\|_{L^1} + o_p(\Delta_n^{2e+1})$ and, by direct calculation, we have $\|k_1\|_{L^1} = 2\|f_2\|^2_{L^2}$, $2\|h_2\|_{L^1} - \|k_3\|_{L^1} = \|l\|_{L^1} = \int_0^1\int_0^s\int_0^r g_e(s-r)g_e(r-u)\,du\,dr\,ds = \frac{\cos(\pi e)}{(2e+2)(2e+3)}$ and $\|h_4\|_{L^1} = (\int_0^1\int_0^s g_e(s-r)\,dr\,ds)^2 = \frac{1}{K_e^2(e+1)^2(e+2)^2} = \frac{\cos(\pi e)\Gamma(2e+2)}{\Gamma(e+3)^2}$, we conclude that

$$\Delta_n\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}Q^n_{i,2} = \frac{p(p-2)}{4}\Delta_n^{2e+1}\int_0^t c_s^{p/2-2}\bigg[\frac{\cos(\pi e)(\eta_s^2+c_s\eta_s^\eta/\eta_s)}{(2e+2)(2e+3)} + \frac{(p-4)\cos(\pi e)\Gamma(2e+2)\eta_s^2}{2\Gamma(e+3)^2}\bigg]\,ds + o_p(\Delta_n^{2e+1}). \tag{H.3}$$

The theorem follows by combining (H.1) and (H.3). □

Lemma H.1. *Suppose that $f \in L^2([0,1]^m)$ for some $m \geq 1$ and $(B_t)_{t\in[0,1]}$ is a standard Brownian motion. Then $\mathbb{E}[I_m(f) \mid B_1] = m!H_m(B_1)\int_{[0,1]^m} f(t_1,\ldots,t_m)\,dt_m\cdots dt_1$, where $H_m(x) = \frac{(-1)^m}{m!}e^{x^2/2}\frac{d^m}{dx^m}e^{-x^2/2}$ is the mth Hermite polynomial.*

The first four Hermite polynomials are given by $H_1(x) = x$, $H_2(x) = \frac{1}{2}(x^2-1)$, $H_3(x) = \frac{1}{6}(x^3-3x)$ and $H_4(x) = \frac{1}{24}(x^4-6x^2+3)$.

Proof of Lemma H.1. Let $\widetilde{f}$ denote the the symmetrization of f (see page 9 in Nualart (2006)) and $h^{\otimes m}(t_1,\dots,t_m) = h(t_1)\cdots h(t_m)$. Because $I_m(f) = I_m(\widetilde{f})$ and the linear span of $\{h^{\otimes m} : h \in L^2([0,1])\}$ is dense in the space of symmetric functions in $L^2([0,1]^m)$, it suffices to show the lemma for $f = h^{\otimes m}$. Let $e = \mathbf{1}_{[0,1]}$ be the constant unit function on $[0,1]$ and $Ph = (\int_0^1 h(s)\,ds)e$ be the projection of h on e and $(I-P)h = h - Ph$. Then Ph and $(I-P)h$ are orthogonal in $L^2([0,1])$ and $\int_0^1 Ph(s)\,dB_s$ and $\int_0^1 (I-P)h(s)\,dB_s$ are independent. In particular, $h = Ph + (I-P)h$ and $\|h\|_{L^2}^2 = \|Ph\|_{L^2}^2 + \|(I-P)h\|_{L^2}^2$. The The Wiener–Itô chaos expansion of the stochastic exponential (see page 23 in Nualart (2006)) states that $\exp(t\int_0^1 h(s)\,dB_s - \frac{t^2}{2}\|h\|_{L^2}^2) = \sum_{m=0}^\infty \frac{t^m}{m!} I_m(h^{\otimes m})$ for all $t \in \mathbb{R}$.

Conditioning on B_1 yields $\sum_{m=0}^\infty \frac{t^m}{m!}\mathbb{E}[I_m(h^{\otimes m}) \mid B_1] = \exp(t\int_0^1 Ph(s)\,dB_s - \frac{t^2}{2}\|Ph\|_{L^2}^2)$ $\times\, \mathbb{E}[\exp(t\int_0^1 (I-P)h(s)\,dB_s - \frac{t^2}{2}\|(I-P)h\|_{L^2}^2) \mid B_1] = \exp(t\int_0^1 Ph(s)\,dB_s - \frac{t^2}{2}\|Ph\|_{L^2}^2)$ $\times\, \mathbb{E}[\exp(t\int_0^1 (I-P)h(s)\,dB_s - \frac{t^2}{2}\|(I-P)h\|_{L^2}^2)]$, where the first step follows from the fact that $\int_0^1 Ph(s)\,dB_s = (\int_0^1 h(s)\,ds)B_1$ and the second holds because of the independence between B_1 and $\int_0^1 Ph(s)\,dB_s$. Because the last expectation is one, we have shown that $\sum_{m=0}^\infty \frac{t^m}{m!}\mathbb{E}[I_m(h^{\otimes m}) \mid B_1] = \exp(t\int_0^1 Ph(s)\,dB_s - \frac{t^2}{2}\|Ph\|_{L^2}^2) = \sum_{m=0}^\infty \frac{t^m}{m!} I_m((Ph)^{\otimes m})$ for all $t \in \mathbb{R}$. Comparing coefficients, we conclude that $\mathbb{E}[I_m(h^{\otimes m}) \mid B_1] = I_m((Ph)^{\otimes m})$, which is equal to $m!(\int_0^1 h(s)\,ds)^m H_m(B_1)$ by Proposition 1.1.4 in Nualart (2006) and the fact that $Ph = (\int_0^1 h(s)\,ds)e$. This shows the lemma for $f = h^{\otimes m}$ and, by a density argument, for all $f \in L^2([0,1]^m)$. □