Supplementary Materials for "Multi-Horizon Test for Market Frictions"

Z. Merrick Li* Xiye Yang†

September 30, 2025

Abstract

This note serves as the online supplementary material for Li and Yang (2025). Section S.1 presents additional simulation results. Section S.2 presents and proves several lemmas, which will be used to prove the main results in Li and Yang (2025).

S.1 Additional Simulation Results

S.1.1 Endogenous deviations

This section presents additional simulation results under the same conditions as Section 5 in Li and Yang (2025), with the explicit inclusion of cross-covariance between the deviations and the efficient price. The cross-covariance is modeled as a correlation between the deviation's innovation and the Brownian increments at a displacement of 1, denoted by $\nu := \mathbf{corr}(\mathrm{d}W_{i+1}, e_i)$. We evaluate two levels of correlation strength: $\nu = 0.3$ and $\nu = 0.6$. Results for negative ν are omitted, as the rejection rates for our left-sided tests are nearly always 100%.

Table S.1 reports the results. The top panel displays results for $\nu = 0.3$, and the bottom for $\nu = 0.6$. Consistent with our analysis in Section 4.4 of Li and Yang (2025), positive cross-covariance reduces the rejection rates for left-sided tests, with more pronounced reductions for larger cross-covariance values. For a fixed ν, the power reduction is more significant for smaller errors. Nonetheless, left-sided tests remain effective for large errors, even with positive cross-covariance.

When positive cross-covariance at displacement 1 is present, our *right-sided* test with a single horizon $\mathbf{K}_0 = \{1\}$ is expected to effectively detect the error terms. Table S.1 confirms this, showing that the right-sided test using returns over a single horizon (denoted as $\mathbf{K}_0^+$)

*Department of Economics, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China. Email: merrickli@cuhk.edu.hk.

†Department of Economics, Rutgers University, 75 Hamilton Street, New Jersey Hall, New Brunswick, NJ 08901, USA. Email: xiyeyang@economics.rutgers.edu.

achieves exceptionally high rejection rates across nearly all models. However, caution is warranted: part of this strong performance stems from the numerical specification of the cross-covariance at displacement 1. In practice, such prior knowledge about the cross-covariance structure is rarely available. Under these circumstances, two-sided tests offer a more robust and reliable alternative.

The power gains of two-sided tests—denoted as $|\mathbf{K}_\ell|$—are particularly significant when using a single horizon with $\ell = 0$, and this aligns with our earlier analysis. Compared to left-sided tests, two-sided tests with $\ell = 1, 2$ exhibit pronounced power gains across several models. Moreover, these gains grow as cross-covariance values increase.

This numerical study shows that the presence of cross-covariance between the deviations and the efficient price can reduce the power of our tests. However, a quick remedy is to use the two-sided tests, or a right-sided test if one has prior knowledge of the cross-covariance structure. These tests are quite robust to the presence of cross-covariance and maintain higher rejection rates even when such dependencies exist.

S.1.2 Other alternatives

We conduct simulation studies to evaluate our tests against various alternative models beyond the conventional signal-plus-noise frameworks discussed in the main text Li and Yang (2025). Although these alternative models are not the primary focus of this paper or the mainstream literature, we believe that the findings warrant a brief discussion to illustrate the usefulness and robustness of our method.

We evaluate three classes of alternative models: fractional Brownian motions with varying Hurst parameters, purely stationary ARMA processes, and drift bursts discussed in recent literature (Christensen et al., 2014; Andersen et al., 2023; Laurent et al., 2024). While the models under each alternative hypothesis can be identified using either left or right-sided tests with greater power, such testing strategies necessitate prior parametric knowledge, which is unrealistic in practice. Therefore, we advocate for the two-sided test outlined in Section 4.5 of Li and Yang (2025) to achieve robustness against a broader class of alternative models.

Table S.2 reports the rejection rates under various alternative models of market inefficiency. As expected, the fractional Brownian motion (fBm) with Hurst parameter $H = 0.55$, which is close to standard Brownian motion ($H = 0.5$), is difficult to detect, as evidenced by the relatively low rejection rate of approximately 16% across all test statistics.

In contrast, the stationary alternatives—represented by AR(1), MA(1), and ARMA(1,1) processes—are more readily detected, particularly when employing multi-horizon tests. The performance improves notably from $|\mathbf{K}_0|$ to $|\mathbf{K}_1|$ and further to $|\mathbf{K}_2|$, suggesting that incorporating multiple time horizons enhances the power of the tests against persistent but stationary deviations from efficiency.

For the drift burst process, we observe strong detection power even in the case of mild explosiveness, specifically when $\alpha = 0.55$. Notably, the multi-horizon test statistic $|\mathbf{K}_2|$ achieves a rejection rate of nearly 89%, indicating its sensitivity to subtle departures from

Statistics \ Model	i.i.d.			AR(1), $\varrho = 0.5$			AR(1), $\varrho = 0.9$			MA(1), $\vartheta = 0.5$			MA(1), $\vartheta = 0.9$		
	K_γ^2 =2.7e-10	5.5e-10	1.1e-09	5.5e-10	1.1e-09	2.2e-09	2.7e-09	5.5e-09	1.1e-08	4.6e-10	9.1e-10	1.8e-09	5.5e-10	1.1e-09	2.2e-09
$\mathbf{K}_0$	5.5	19.2	90.0	0.4	2.1	7.1	0.0	0.0	0.0	0.4	1.9	6.8	0.3	0.0	0.0
$\mathbf{K}_1$	5.0	19.0	86.0	12.0	41.1	97.2	0.0	0.0	1.0	27.6	83.1	100.0	74.4	100.0	100.0
$\mathbf{K}_2$	4.9	18.6	84.0	16.9	54.0	98.3	7.9	21.4	58.6	23.9	79.8	100.0	70.1	100.0	100.0
$\mathbf{K}_0^+$	37.4	3.8	0.0	91.0	86.3	58.9	100.0	100.0	100.0	46.8	44.0	17.9	4.2	5.4	6.9
$\|\mathbf{K}_0\|$	33.5	20.5	85.2	86.4	83.8	59.4	100.0	100.0	100.0	36.2	36.1	18.0	2.3	3.0	4.0
$\|\mathbf{K}_1\|$	18.1	19.5	82.8	11.3	34.4	95.0	48.0	53.0	48.3	22.8	79.1	100.0	68.3	100.0	100.0
$\|\mathbf{K}_2\|$	10.8	18.9	79.0	14.8	47.6	97.6	7.2	17.8	50.5	20.9	76.1	100.0	64.9	99.6	100.0

Statistics \ Model	AR(2), $\boldsymbol{\varrho} = (0.7, -0.2)$			AR(2), $\boldsymbol{\varrho} = (0.5, 0.2)$			MA(2), $\boldsymbol{\vartheta} = (0.5, 0.4)$			MA(2), $\boldsymbol{\vartheta} = (0.5, -0.2)$			ARMA(1,1), $\varrho = 0.6, \vartheta = 0.4$		
	K_γ^2 =6.6e-10	1.3e-09	2.6e-09	7.3e-10	1.5e-09	2.9e-09	5.4e-10	1.1e-09	2.2e-09	4.0e-10	7.9e-10	1.6e-09	1.1e-09	2.2e-09	4.5e-09
$\mathbf{K}_0$	0.0	0.0	0.0	0.6	3.1	10.3	0.4	2.2	7.4	2.9	9.4	29.9	0.0	0.0	0.0
$\mathbf{K}_1$	45.7	93.7	100.0	1.5	5.7	19.3	19.0	57.0	98.5	27.9	86.5	100.0	35.6	80.4	100.0
$\mathbf{K}_2$	54.1	96.5	100.0	10.1	27.9	76.1	31.8	82.8	100.0	24.4	83.2	100.0	72.7	99.2	100.0
$\mathbf{K}_0^+$	100.0	100.0	100.0	100.0	99.6	98.5	100.0	99.6	99.0	67.1	69.0	46.9	100.0	100.0	100.0
$\|\mathbf{K}_0\|$	96.8	99.7	99.8	95.9	89.2	62.1	96.0	91.1	71.6	4.0	8.2	25.7	99.8	100.0	100.0
$\|\mathbf{K}_1\|$	39.9	90.1	100.0	16.4	11.9	16.8	16.1	49.5	97.4	23.5	82.4	100.0	29.5	76.3	99.9
$\|\mathbf{K}_2\|$	48.4	95.2	100.0	8.3	24.0	70.3	26.9	77.4	100.0	21.7	80.1	100.0	67.0	98.2	100.0

Statistics \ Model	i.i.d.			AR(1), $\varrho = 0.5$			AR(1), $\varrho = 0.9$			MA(1), $\vartheta = 0.5$			MA(1), $\vartheta = 0.9$		
	K_γ^2 =2.7e-10	5.5e-10	1.1e-09	5.5e-10	1.1e-09	2.2e-09	2.7e-09	5.5e-09	1.1e-08	4.6e-10	9.1e-10	1.8e-09	5.5e-10	1.1e-09	2.2e-09
$\mathbf{K}_0$	0.4	2.2	7.7	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0
$\mathbf{K}_1$	0.4	2.2	7.6	3.2	12.3	49.5	0.0	0.0	0.0	10.0	36.0	99.1	67.5	99.8	100.0
$\mathbf{K}_2$	0.4	2.3	7.4	14.5	42.4	92.9	9.7	24.5	60.9	9.9	33.6	98.8	63.5	99.6	100.0
$\mathbf{K}_0^+$	100.0	99.6	97.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.6	40.8	64.3	87.7
$\|\mathbf{K}_0\|$	99.5	98.3	95.3	100.0	99.6	99.0	100.0	100.0	100.0	100.0	99.6	98.3	10.6	20.4	31.1
$\|\mathbf{K}_1\|$	97.1	96.2	88.9	18.5	15.6	43.7	98.2	99.9	99.4	13.4	32.1	98.8	61.8	99.6	100.0
$\|\mathbf{K}_2\|$	49.1	55.3	48.3	12.9	36.6	90.7	8.8	21.0	54.2	12.0	29.6	97.8	57.0	99.3	100.0

Statistics \ Model	AR(2), $\boldsymbol{\varrho} = (0.7, -0.2)$			AR(2), $\boldsymbol{\varrho} = (0.5, 0.2)$			MA(2), $\boldsymbol{\vartheta} = (0.5, 0.4)$			MA(2), $\boldsymbol{\vartheta} = (0.5, -0.2)$			ARMA(1,1), $\varrho = 0.6, \vartheta = 0.4$		
	K_γ^2 =6.6e-10	1.3e-09	2.6e-09	7.3e-10	1.5e-09	2.9e-09	5.4e-10	1.1e-09	2.2e-09	4.0e-10	7.9e-10	1.6e-09	1.1e-09	2.2e-09	4.5e-09
$\mathbf{K}_0$	0.0	0.0	0.0	0.0	0.1	0.9	0.0	0.0	0.4	0.4	1.4	5.5	0.0	0.0	0.0
$\mathbf{K}_1$	41.9	92.1	100.0	0.1	0.7	2.8	15.6	49.3	96.8	8.4	28.4	98.8	38.1	81.7	100.0
$\mathbf{K}_2$	60.3	98.0	100.0	12.2	30.2	79.5	30.1	80.5	100.0	7.7	26.3	98.3	91.0	100.0	100.0
$\mathbf{K}_0^+$	100.0	100.0	100.0	100.0	100.0	99.6	100.0	100.0	100.0	99.8	99.2	95.6	100.0	100.0	100.0
$\|\mathbf{K}_0\|$	100.0	100.0	100.0	100.0	99.7	99.4	100.0	99.6	99.3	56.5	61.3	42.8	100.0	100.0	100.0
$\|\mathbf{K}_1\|$	36.8	88.2	100.0	84.8	88.5	79.2	13.7	42.4	95.2	17.1	25.8	98.1	31.6	76.9	99.8
$\|\mathbf{K}_2\|$	53.6	96.6	100.0	10.4	26.6	74.8	26.0	75.5	100.0	13.9	25.1	97.3	88.2	100.0	100.0

Table S.1: Rejection percentage of $\mathbb{H}_0$ under $\mathbb{H}_1$ based on 1,000 simulations at 1% significance level. The numerical settings are the same as in Section 5 of Li and Yang (2025), except that $\mathbf{corr}(\mathrm{d}W_{i+1}, e_i) = 0.3$ for the results in the upper panel and $\mathbf{corr}(\mathrm{d}W_{i+1}, e_i) = 0.6$ for the results in the lower panel. $\mathbf{K}_\ell$, $\mathbf{K}_\ell^+$ and $|\mathbf{K}_\ell|$ represent the left-sided, right-sided, and two-sided tests statistics for $\ell = 0, 1, 2$.

efficiency induced by locally explosive trends. As α increases, reflecting stronger drift bursts, the power rapidly approaches 100%, demonstrating the effectiveness of our testing procedure in identifying increasingly pronounced inefficiencies.

The robustness of these results is supported by the consistent pattern across different alternative models. For instance, the superior performance of $|\mathbf{K}_2|$ is not limited to one specific data-generating process but holds across both stationary linear models and path-dependent drift burst dynamics. Moreover, the use of two-sided tests ensures that deviations from the null hypothesis are captured regardless of their directional bias, further reinforcing the reliability of our inference framework.

Model / Statistics	fBm H=0.2	fBm 0.55	fBm 0.8	AR(1) $\varrho_1 = 0.7$	MA(1) $\vartheta_1 = 0.7$	ARMA(1,1) (0.7, 0.2)	Drift burst $\alpha = 0.55$	0.65	0.75
$\lvert\mathbf{K}_0\rvert$	100.0	15.9	100.0	58.9	1.4	2.6	4.3	24.3	82.8
$\lvert\mathbf{K}_1\rvert$	100.0	19.6	100.0	84.1	100.0	62.8	45.8	90.2	99.9
$\lvert\mathbf{K}_2\rvert$	100.0	18.9	100.0	88.9	100.0	91.0	88.9	100.0	100.0

Table S.2: Rejection rates of the null hypothesis of an efficient price process ($\mathbb{H}_0$) under alternative models of market inefficiency, including fractional Brownian motion (fBm), stationary linear processes (AR(1), MA(1), ARMA(1,1)), and the drift burst hypothesis. Results are reported at the 1% significance level. The statistics $|\mathbf{K}_\ell|, \ell = 0, 1, 2$ denote two-sided tests derived from the methodology in Section 4.5 of the main text Li and Yang (2025). For the drift burst model, $\alpha \in (0.5, 1)$ governs the intensity of the locally explosive trend (higher α implies stronger drift bursts), as defined in Christensen et al. (2022). Simulations use 1,000 replications. The parameter ϱ_1 denotes the autoregressive coefficient in AR(1), ϑ_1 the moving average coefficient in MA(1), and (ϱ_1, ϑ_1) for ARMA(1,1). We set the innovations in the ARMA processes to have unit variance.

S.2 Mathematical Proofs

In this section, we present and prove several technical lemmas required for the main results in Li and Yang (2025). These results are derived under the general framework outlined in Section 4 and the Appendix of Li and Yang (2025), where the efficient price and its various components follow a semimartingale with a jump component, the observation scheme $\{T_i^n\}_i$ is random, and the deviations exhibit stochastic scaling. In the sequel, we denote $\mathcal{F}_i^n := \mathcal{F}_{T_i^n}$, and $V_i^n := V_{T_i^n}$ for any process V.

For clarity, equations, theorems, and other numbered items from the main text (Li and Yang, 2025) keep their original Arabic numerals, such as (1), (2), and (3). In this supplement, new equations, theorems, lemmas, and similar items are labeled with an "S." prefix to differentiate them from those in the main text. For example, new equations in the supplement are labeled as (S.1), (S.2), and (S.3).

In the sequel, K will denote a constant that may change from line to line, or even within one line; when it depends on some parameters *par*, we write it K_{par}; but it never depends on n or other indices such as i, j.

S.2.1 Preliminaries and Notations

We will use several classic estimates for Itô semimartingales (Jacod et al., 2017). Let V be any of the processes $X, b, \sigma, \alpha, 1/\alpha, \gamma$ (or any power of them), we have for any two finite stopping times $S \le S'$, and any $w \ge 2$ the following

$$|\mathbb{E}\left(V_{S'} - V_S \middle| \mathcal{F}_S\right)| \le K\mathbb{E}\left(S' - S \middle| \mathcal{F}_S\right), \tag{S.1}$$

$$\mathbb{E}\left(\sup_{S \le s \le S'} |V_s - V_S|^w \middle| \mathcal{F}_S\right) \le K_w\left(\mathbb{E}\left(S' - S \middle| \mathcal{F}_S\right) + \mathbb{E}\left((S' - S)^w \middle| \mathcal{F}_S\right)\right). \tag{S.2}$$

Under Assumption (H-X), we have the following decomposition $X_t = X'_t + J_t$, where

$$X'_t = X_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s; \quad J_t = \int_{[0,t]\times\mathbb{R}} \delta(s,x)\mu(\mathrm{d}s, \mathrm{d}x).$$

Let $q \in [0,1)$, and k be a positive integer. For any integers $i \in \mathbb{N}^*$, let

$$m(k)_i^q := (2i-1)k + q_k; \quad m(k)_{i\pm}^q := m(k)_i^q \pm k; \tag{S.3}$$

where $q_k := \lfloor 2qk \rfloor$ and $\lfloor\cdot\rfloor$ is the floor function.

We introduce the following notations for any two processes V and U:

$$F(V,U;k,q)_t^n := \sum_{j=0}^{n(k)_t^q - 1} f(V,U;k)_{q_k+(2j+1)k}^n, \; F(V,U;k)_t^n := \int_0^1 F(V,U;k,q)_t^n \mathrm{d}q.$$

where $f(V,U;k)_i^n := (V_i^n - V_{i-k}^n)(U_{i+k}^n - U_i^n)$ for any processes V and U and a positive window size k, and $n(k)_t^q := \lfloor (n_t - q_k)/(2k) \rfloor$. When $V = U$, the notation $f(V,U;k)_i^n$ is identical to $f(V;k)_i^n$ in Section 3 of Li and Yang (2025).

Note that $F(V,U;k,q)_t^n$ is the sum of the products of the increments of V and U over *non-overlapping* blocks of size $2k$, with initial points determined by q (or q_k), and $n(k)_t^q$ is the number of such blocks. The integral $F(V,U;k)_t^n$ is the average of $F(V,U;k,q)_t^n$ over $q \in [0,1)$, or equivalently, q_k from 0 to $2k-1$. Thus, one can also write

$$F(V,U;k)_t^n = \frac{1}{2k}\sum_{i=k}^{n_t - k} f(V,U;k)_i^n.$$

When $V = U$, we write $F(V;k)_t^n$ and $F(V;k,q)_t^n$ for brevity.

By a classic localization procedure, we can replace the three assumptions (H-X), (O-ρ) and (N-θ-v) by the following stronger one:

Assumption (S-HON). *Assume Assumptions (H-X), (O-ρ) (with $\tau_1 = \infty$) and (N-θ-v) hold. Assumption (H) hold for $b, \sigma, \alpha, \gamma$. The function δ and the processes $b, \sigma, \alpha, 1/\alpha, \gamma, X$ are bounded,* *and there exists a nonnegative function Γ on $\mathbb{R}$, satisfying*

$$
\begin{aligned}
&|\delta(\omega,t,x)| \leq \Gamma(x),\ \int_{\mathbb{R}} (\Gamma(x)\wedge 1)\,\lambda(\mathrm{d}x) < \infty,\ n_t \leq Kt\Delta_n^{-1};\\
&\left|\mathbb{E}\left(\Delta(n,i) - \Delta_n/\alpha_{i-1}^n \middle| \mathcal{F}_{i-1}^n\right)\right| \leq K\Delta_n^{1+\rho},\ \mathbb{E}\left(\Delta(n,i)^\kappa \middle| \mathcal{F}_{i-1}^n\right) \leq K\Delta_n^\kappa,\ \kappa \geq 2.
\end{aligned}
\tag{S.4}
$$

S.2.2 Technical Lemmas

We first present a useful lemma concerning the random observation schemes $\{T_i^n\}_i$. For detailed assumptions and descriptions of this scheme, see Section 4 and Appendix A.2 in Li and Yang (2025).

Lemma S.1. For any $j' > j \geq 1$, $\kappa \geq 2$,

$$\mathbb{E}\left(T_{j'}^n - T_j^n \middle| \mathcal{F}_j^n\right) \leq K(j'-j)\Delta_n, \quad \mathbb{E}\left((T_{j'}^n - T_j^n)^\kappa \middle| \mathcal{F}_j^n\right) \leq K((j'-j)\Delta_n)^\kappa; \tag{S.5}$$

$$\left|\mathbb{E}\left(T_{j'}^n - T_j^n - (j'-j)\Delta_n/\alpha_j^n \middle| \mathcal{F}_j^n\right)\right| \leq K(j'-j)\Delta_n\left(\Delta_n^\rho \vee (j'-j)\Delta_n\right); \tag{S.6}$$

$$\mathbb{E}\left(\left(T_{j'}^n - T_j^n - (j'-j)\Delta_n/\alpha_j^n\right)^2 \middle| \mathcal{F}_j^n\right) \leq K((j'-j)\Delta_n)^2\left(\Delta_n^\rho \vee (j'-j)^{-1} \vee (j'-j)\Delta_n\right). \tag{S.7}$$

Proof. (S.5) follows directly from (S.4), the boundedness of $1/\alpha$, and Hölder's inequality. Now let $d_k^n := \Delta(n,k) - \frac{\Delta_n}{\alpha_{k-1}^n}$. Note that

$$T_{j'}^n - T_j^n - (j'-j)\Delta_n/\alpha_j^n =: \mathfrak{A}(1)_{j,j'}^n + \mathfrak{A}(2)_{j,j'}^n,$$

where $\mathfrak{A}(1)_{j,j'}^n := \Delta_n \sum_{k=j+1}^{j'}\left(\frac{1}{\alpha_{k-1}^n} - \frac{1}{\alpha_j^n}\right)$, $\mathfrak{A}(2)_{j,j'}^n := \sum_{k=j+1}^{j'} d_k^n$. (S.4) and (S.1) imply $\left|\mathbb{E}\left(\mathfrak{A}(1)_{j,j'}^n \middle| \mathcal{F}_j^n\right)\right| \leq K(\Delta_n(j'-j))^2$; (S.4) also implies $\left|\mathbb{E}\left(\mathfrak{A}(2)_{j,j'}^n \middle| \mathcal{F}_j^n\right)\right| \leq K\Delta_n^{1+\rho}(j'-j)$, and this proves (S.6).

Apply (S.2) (for $1/\alpha$) and (S.5), we have

$$\mathbb{E}\left((\mathfrak{A}(1)_{j,j'}^n)^2 \middle| \mathcal{F}_j^n\right) \leq K(j'-j)\Delta_n^2 \sum_{k=j+1}^{j'} \mathbb{E}\left(\left(\frac{1}{\alpha_{k-1}^n} - \frac{1}{\alpha_j^n}\right)^2 \middle| \mathcal{F}_j^n\right) \leq K((j'-j)\Delta_n)^3.$$

By applying (S.4), successive conditioning, and Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
\mathbb{E}\left((\mathfrak{A}(2)_{j,j'}^n)^2 \middle| \mathcal{F}_j^n\right) &\leq K\left(\Delta_n^2(j'-j) + \sum\nolimits_{k,k':k'>k} \mathbb{E}\left(|d_k^n|\left|\mathbb{E}\left(d_{k'}^n \middle| \mathcal{F}_{k'-1}^n\right)\right| \middle| \mathcal{F}_j^n\right)\right)\\
&\leq K((j'-j)\Delta_n)^2(\Delta_n^\rho \vee (j'-j)^{-1}).
\end{aligned}
$$

This completes the proof of (S.7). □

Lemma S.2. Let S, S' be two finite stopping times, and let V be a local martingale. $U_{S'}$ is a bounded variable that is measurable with respect to $\mathcal{F}_{S'}$. Then, we have

$$\left|\mathbb{E}\left(U_{S'}(V_{S'} - V_S) \middle| \mathcal{F}_{S'\wedge S}\right)\right| \leq K\left|\mathbb{E}\left(S' - S \middle| \mathcal{F}_{S'\wedge S}\right)\right|.$$

Proof. (i) Assume $S' \leq S$, then the result follows immediately from (S.1) and the fact that U is bounded. (ii) Now assume $S' > S$, we define a (bounded) martingale $M_t := \mathbb{E}\left(U_{S'} \middle| \mathcal{F}_t\right)$. Thus, we have $M_{S'} = U_{S'}$. By the analysis in (i), we have

$$|\mathbb{E}\left(M_S(V_{S'} - V_S) \middle| \mathcal{F}_{S'\wedge S}\right)| \leq K\mathbb{E}\left(S' - S \middle| \mathcal{F}_S\right). \tag{S.8}$$

Next, by the representation theorems for martingale, see, e.g., Theorem III 4.34 of Jacod and Shiryaev (2003), both M and V have bounded quadratic variation whence bounded co-variation as well. It is easy to see that

$$|\mathbb{E}\left((M_{S'} - M_S)(V_{S'} - V_S) \middle| \mathcal{F}_{S'\wedge S}\right)| \leq K\left|\mathbb{E}\left(\int_S^{S'} \mathrm{d}[V, M]_s \,\middle|\, \mathcal{F}_S\right)\right| \leq K\mathbb{E}\left(S' - S \middle| \mathcal{F}_S\right),$$

which, when combined with the earlier result in (S.8), completes the proof. □

Recall X' below is the continuous part of the efficient price X.

Lemma S.3. Let k be a positive integer. For any $i \geq k$, we have

$$\left|\mathbb{E}\left(f(X';k)_i^n \middle| \mathcal{F}_{i-k}^n\right)\right| \leq K(k\Delta_n)^2; \tag{S.9}$$

$$\left|\mathbb{E}\left(\left(f(X';k)_i^n/k\Delta_n\right)^2 \middle| \mathcal{F}_{i-k}^n\right) - (\sigma_{i-k}^n)^4/(\alpha_{i-k}^n)^2\right| \leq K(k\Delta_n \vee \Delta_n^\rho); \tag{S.10}$$

$$\left|\mathbb{E}\left(f(X';k)_i^n f(X';k)_{i+k}^n \middle| \mathcal{F}_{i-k}^n\right)\right| \leq K(k\Delta_n)^3. \tag{S.11}$$

Proof of (S.9). First of all, we rewrite $\mathbb{E}\left(X'^n_{i+k} - X'^n_i \mid \mathcal{F}_i^n\right)$ as

$$\frac{b_i^n}{\alpha_i^n}k\Delta_n + b_i^n\mathbb{E}\left(T_{i+k}^n - T_i^n - \frac{k\Delta_n}{\alpha_i^n} \,\middle|\, \mathcal{F}_i^n\right) + \mathbb{E}\left(\int_{T_i^n}^{T_{i+k}^n} (b_s - b_i^n)\mathrm{d}s \,\middle|\, \mathcal{F}_i^n\right).$$

Recall that b/α is a bounded stochastic process, we have, by Lemma S.1 and Lemma S.2 that $\left|\mathbb{E}\left((X'^n_i - X'^n_{i-k})b_i^n/\alpha_i^n \mid \mathcal{F}_{i-k}^n\right)\right| \leq Kk\Delta_n$. Now we have by the boundedness of b, (S.6) and Lemma S.2 that

$$\left|\mathbb{E}\left(b_i^n\mathbb{E}\left(\frac{T_{i+k}^n - T_i^n}{k\Delta_n} - \frac{1}{\alpha_i^n} \mid \mathcal{F}_i^n\right)(X'^n_i - X'^n_{i-k}) \mid \mathcal{F}_{i-k}^n\right)\right| \leq K(k\Delta_n)(k\Delta_n \vee \Delta_n^\rho).$$

Finally, by first conditioning on $\bigvee_{j=i}^{i+k}\boldsymbol{\sigma}(\Delta(n,j)) \bigvee \mathcal{F}_i^n$ and then applying (S.1) (for b), we have $\left|\mathbb{E}\left(\int_{T_i^n}^{T_{i+k}^n}(b_s - b_i^n)\mathrm{d}s \mid \mathcal{F}_i^n\right)\right| \leq K(k\Delta_n)^2$, which yields the following after another application of Lemma S.2 :

$$\left|\mathbb{E}\left[\mathbb{E}\left(\int_{T_i^n}^{T_{i+k}^n}(b_s - b_i^n)\mathrm{d}s \,\middle|\, \mathcal{F}_i^n\right)(X'^n_i - X'^n_{i-k}) \,\middle|\, \mathcal{F}_{i-k}^n\right]\right| \leq K(k\Delta_n)^3.$$

Hence, the conclusion readily follows. □

Proof of (S.10). Denote

$$\mathfrak{B}(1)_i^n := \int_{T_{i-k}^n}^{T_i^n} b_s \mathrm{d}s, \qquad \mathfrak{C}(1)_i^n := \int_{T_{i-k}^n}^{T_i^n} \sigma_s^2 \mathrm{d}s, \qquad \mathfrak{D}(1)_i^n := \int_{T_{i-k}^n}^{T_i^n} \sigma_s \mathrm{d}W_s,$$
$$\mathfrak{B}(2)_i^n := \int_{T_i^n}^{T_{i+k}^n} b_s \mathrm{d}s, \qquad \mathfrak{C}(2)_i^n := \int_{T_i^n}^{T_{i+k}^n} \sigma_s^2 \mathrm{d}s, \qquad \mathfrak{D}(2)_i^n := \int_{T_i^n}^{T_{i+k}^n} \sigma_s \mathrm{d}W_s.$$

Thus, we have $(f(X';k)_i^n)^2 = (\mathfrak{B}(1)_i^n + \mathfrak{D}(1)_i^n)^2(\mathfrak{B}(2)_i^n + \mathfrak{D}(2)_i^n)^2$. An immediate observation is that that the leading term in $\mathbb{E}\left((f(X';k)_i^n)^2 \middle| \mathcal{F}_{i-k}^n\right)$ is given by $\mathbb{E}\left((\mathfrak{D}(1)_i^n)^2(\mathfrak{D}(2)_i^n)^2 \middle| \mathcal{F}_{i-k}^n\right)$, which is equal to $\mathbb{E}\left(\mathfrak{C}(1)_i^n \mathfrak{C}(2)_i^n \middle| \mathcal{F}_{i-k}^n\right)$ by successive conditioning and Itô 's isometry.

Rewrite $\mathfrak{C}(1)_i^n = \mathscr{C}(1)_i^n + \mathscr{D}(1)_i^n$, $\mathfrak{C}(2)_i^n = \mathscr{C}(2)_i^n + \mathscr{D}(2)_i^n$, where

$$\mathscr{C}(1)_i^n := \int_{T_{i-k}^n}^{T_i^n} (\sigma_s^2 - (\sigma_{i-k}^n)^2)\mathrm{d}s, \quad \mathscr{D}(1)_i^n := (\sigma_{i-k}^n)^2(T_i^n - T_{i-k}^n),$$
$$\mathscr{C}(2)_i^n := \int_{T_i^n}^{T_{i+k}^n} (\sigma_s^2 - (\sigma_i^n)^2)\mathrm{d}s, \quad \mathscr{D}(2)_i^n := (\sigma_i^n)^2(T_{i+k}^n - T_i^n).$$

Cauchy-Schwarz inequality and the estimate (S.2) (for σ) yield

$$\mathbb{E}(|\mathscr{C}(1)_i^n| \mid \mathcal{F}_{i-k}^n) \vee \mathbb{E}(|\mathscr{C}(2)_i^n| \mid \mathcal{F}_i^n) \le K(k\Delta_n)^{3/2}.$$

Thus, the leading term in $\mathbb{E}\left(\mathfrak{C}(1)_i^n \mathfrak{C}(2)_i^n \middle| \mathcal{F}_{i-k}^n\right)$ is $\mathbb{E}\left(\mathscr{D}(1)_i^n \mathscr{D}(2)_i^n \middle| \mathcal{F}_{i-k}^n\right)$. Now apply (S.6), we have

$$\left|\mathbb{E}\left(\frac{\mathscr{D}(1)_i^n \mathscr{D}(2)_i^n}{(k\Delta_n)^2} \middle| \mathcal{F}_{i-k}^n\right) - \frac{(\sigma_{i-k}^n)^2}{\alpha_{i-k}^n} \frac{(\sigma_i^n)^2}{\alpha_i^n}\right| \le K(k\Delta_n \vee \Delta_n^\rho).$$

By (S.1), we have

$$\left|\mathbb{E}\left((\sigma_{i-k}^n)^2/\alpha_{i-k}^n \left((\sigma_i^n)^2/\alpha_i^n - (\sigma_{i-k}^n)^2/\alpha_{i-k}^n\right) \middle| \mathcal{F}_{i-k}^n\right)\right| \le K\Delta_n.$$

Now the proof is complete. □

Proof of (S.11). Now let $\mathfrak{C}_i^n := \frac{1}{(k\Delta_n)^2}\mathbb{E}\big((X_i'^n - X_{i+k}'^n)^2(X_{i+2k}'^n - X_{i+k}'^n) \mid \mathcal{F}_i^n\big)$. By successive conditioning and (S.1) and (S.2), we have the boundedness of $\mathfrak{C}_i^n$. Then Lemma S.1 and Lemma S.2 imply that $\left|\mathbb{E}\left((X_i'^n - X_{i-k}'^n)\mathfrak{C}_i^n \middle| \mathcal{F}_{i-k}^n\right)\right| \le K(k\Delta_n)^3$. This proves (S.11).

□

Lemma S.4. We have for $\theta \in [0,1]$ and any positive integer k,

$$\mathbb{E}\big(\big|F(Y;k)_t^n - F'(\chi;k)_t^n\big|\big) \le K\Delta_n^{\frac{1}{2}\wedge\theta}, \tag{S.12}$$

where $F'(\chi;k)_t^n := \frac{1}{2k}\sum_{i=k}^{n_t-k}(\Delta_n^\theta \gamma_i^n)^2 f(\chi;k)_i^n$.

Proof. We make the following decomposition

$$2k\left(F(Y;k)_t^n - F'(\chi;k)_t^n\right) = \sum_{i=k}^{n_t-k}\sum_{\ell=1}^{8}\mathfrak{E}(\ell)_i^n, \tag{S.13}$$

where

$$\mathfrak{E}(1)_i^n := f(X;k)_i^n,\ \ \mathfrak{E}(2)_i^n := \Delta_n^{2\theta}\chi_{i-k}\chi_{i+k}f(\gamma;k)_i^n,\ \ \mathfrak{E}(3)_i^n := \Delta_n^{\theta}\chi_{i+k}f(X,\gamma;k)_i^n,$$
$$\mathfrak{E}(4)_i^n := \Delta_n^{\theta}\chi_{i-k}f(\gamma,X;k)_i^n,\ \ \mathfrak{E}(5)_i^n := \Delta_n^{\theta}\gamma_i^n f(\chi,X;k)_i^n,\ \ \mathfrak{E}(6)_i^n := \Delta_n^{\theta}\gamma_i^n f(X,\chi;k)_i^n,$$
$$\mathfrak{E}(7)_i^n := \Delta_n^{2\theta}\gamma_i^n\chi_{i+k}f(\chi,\gamma;k)_i^n,\ \ \mathfrak{E}(8)_i^n := \Delta_n^{2\theta}\gamma_i^n\chi_{i-k}f(\gamma,\chi;k)_i^n.$$

It's trivial to show that

$$\mathbb{E}\Big((\mathfrak{E}(\ell)_i^n)^2\Big) \le \begin{cases} K\Delta_n^{2+4\theta} & \ell = 2; \\ K\Delta_n^{2+2\theta} & \ell = 3, 4. \end{cases}$$

Whence, by Cauchy-Schwarz inequality, we have

$$\sum\nolimits_{i=k}^{n_t-k}\mathbb{E}(|\mathfrak{E}(\ell)_i^n|) \le K\Delta_n^{\theta},\ \ell = 2,3,4. \tag{S.14}$$

Based on successive conditioning and the standard results provided in Chapter 2 of Jacod and Protter (2011), it is not hard to show

$$\mathbb{E}(|\mathbb{E}\left(f(X;k)_i^n|\,\mathcal{F}_i^n\right)|) + \mathbb{E}\big((f(X;k)_i^n)^2\big) \le K\Delta_n^2,$$

which in turn imply

$$\mathbb{E}\left(\left|\sum\nolimits_{i=k}^{n_t-k}\mathfrak{E}(1)_i^n\right|\right) \le K\sqrt{\Delta_n}. \tag{S.15}$$

Now let $\mathcal{H}_i^n := \mathcal{F}_{i-k}^n \otimes \mathcal{G}$. Then, we have

$$\mathbb{E}(|\mathbb{E}\left(f(X,\chi)_i^n|\,\mathcal{H}_i^n\right)|) \le K\mathbb{E}\big(\left|\mathbb{E}\left(X_i - X_{i-k}|\,\mathcal{F}_{i-k}^n\right)\right|\big) \le K\Delta_n.$$

On the other hand, we also have $\mathbb{E}\big((f(X,\chi;k)_i^n)^2\big) \le K\Delta_n$. Another application of Lemma A.6 of Jacod et al. (2017) leads to

$$\mathbb{E}\left(\sum\nolimits_{i=k}^{n_t-k}|f(X,\chi;k)_i^n|\right) \le K.$$

Similarly, we can prove $\mathbb{E}\left(\sum_{i=k}^{n_t-k}|f(\chi,X;k)_i^n|\right) \le K$. Using the same technique, we can show

$$\mathbb{E}\left(\sum\nolimits_{i=k}^{n_t-k}\left(|\chi_{i+k}f(\chi,\gamma;k)_i^n| + |\chi_{i-k}f(\gamma,\chi;k)_i^n|\right)\right) \le K.$$

By the boundedness of γ, we have

$$\sum_{i=k}^{n_t-k} \mathbb{E}(|\mathfrak{E}(\ell)|) \leq \begin{cases} K\Delta_n^{\theta}, & \ell = 5, 6; \\ K\Delta_n^{2\theta}, & \ell = 7, 8. \end{cases} \tag{S.16}$$

Now the result follows immediately from (S.13), (S.14), (S.15), and (S.16). □

Lemma S.5. Let $\theta \in \left[0, \frac{3}{4}\right)$. For any given positive integer k, we have

$$\mathbb{E}\left(\left|2k\Delta_n^{1-2\theta} F(Y;k)_t^n + g(k,r)\int_0^t \gamma_s^2 \mathrm{d}A_s\right|\right) \leq K\Delta_n^{\frac{1}{2}\wedge\frac{v}{1+v}\wedge(1-\theta)\wedge\left(\frac{3}{2}-2\theta\right)}. \tag{S.17}$$

Proof. We introduce the following decomposition

$$2\Delta_n k F(Y;k)_t^n + g(k;r)\Delta_n^{2\theta}\int_0^t \gamma_s^2 \mathrm{d}A_s = \sum_{\ell=1}^{4} \mathfrak{D}(\ell)_t^n,$$

with

$$\begin{aligned}
\mathfrak{D}(1)_t^n &:= \Delta_n \sum_{i=k}^{n_t-k} (\Delta_n^{\theta}\gamma_i^n)^2 \left(f(\chi;k)_i^n + g(k;r)\right); \\
\mathfrak{D}(2)_t^n &:= g(k;r) \sum_{i=k}^{n_t-k} (\Delta_n^{\theta}\gamma_i^n)^2 \left(\alpha_i^n \Delta(n,i+1) - \Delta_n\right); \\
\mathfrak{D}(3)_t^n &:= g(k;r)\Delta_n^{2\theta}\left(\int_0^t \gamma_s^2 \mathrm{d}A_s - \sum_{i=k}^{n_t-k} (\gamma_i^n)^2 \alpha_i^n \Delta(n,i+1)\right); \\
\mathfrak{D}(4)_t^n &:= 2\Delta_n k \left(F(Y;k)_t^n - F'(\chi;k)_t^n\right).
\end{aligned}$$

Let $k_n \geq 2k$, and define

$$\widetilde{\mathcal{H}}_i^n := \mathcal{F}_{i-k}^n \otimes \mathcal{G}_{i-k_n}, \quad \delta_i^n := (\gamma_i^n)^2 \left(f(\chi;k)_i^n + g(k;r)\right).$$

Thus, δ_i^n is $\widetilde{\mathcal{H}}_{i+k+k_n}^n$- measurable. We have

$$\mathbb{E}\left(\left|\mathbb{E}\left(\delta_i^n | \widetilde{\mathcal{H}}_i^n\right)\right|\right) \leq K k_n^{-v}, \; \mathbb{E}\left((\delta_i^n)^2\right) \leq K.$$

Lemma A.6 in Jacod et al. (2017) yields $\mathbb{E}(|\mathfrak{D}(1)_t^n|) \leq K\Delta_n^{2\theta}\left(k_n^{-v} \vee \sqrt{\Delta_n k_n}\right)$. Now let $k_n \asymp \Delta_n^{-1/(1+v)}$, we have $\mathbb{E}(|\mathfrak{D}(1)_t^n|) \leq K\Delta_n^{2\theta+\frac{v}{1+v}}$. Next, by Lemma A.2 of Jacod et al. (2017), we have $\mathbb{E}(|\mathfrak{D}(\ell)_t^n|) \leq K\Delta_n^{\frac{1}{2}+2\theta}$ for $\ell = 2, 3$. Lemma S.4 implies $\mathbb{E}(|\mathfrak{D}(4)_t^n|) \leq K\Delta_n^{(1+\theta)\wedge\frac{3}{2}}$. This completes the proof. □

Now we state and prove a key theorem.

Theorem S.6. Suppose that Assumption (S-HON) holds, we have the following functional stable convergence in law:

$$\left(\frac{1}{\sqrt{\Delta_n}} F(X';k,q)_t^n\right)_{q\in[0,1)} \xrightarrow{\mathcal{L}_s} (\mathcal{Z}_t^q)_{q\in[0,1)},$$

where the limits are defined on an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$. Conditional on $\mathcal{F}$, each Z_t^q is Gaussian, and for any $q, q' \in [0, 1)$, we have

$$\tilde{\mathbb{E}}\big(\mathcal{Z}_t^q \mathcal{Z}_t^{q'} \mid \mathcal{F}\big) = \Phi_k^{q,q'} \int_0^t \frac{\sigma_s^4}{\alpha_s^2} \mathrm{d}A_s,$$

where $\tilde{\mathbb{E}}$ is the expectation with respect to $\tilde{\mathbb{P}}$ and $\Phi_k^{q,q'} := \frac{(k - |q_k - q_k'|)^2}{2k}$.

Proof. (i) We first prove the convergence result for any $q \in [0, 1)$. According to (S.9), and the fact that $n(k)_t^q \leq \frac{Kt}{\Delta_n k}$, we readily get that

$$\frac{1}{\sqrt{k\Delta_n}} \sum_{i=1}^{n(k)_t^q} \Big| \mathbb{E}\big(f(X'; k)_{m(k)_i^q}^n \mid \mathcal{F}_{m(k)_{i-}^q}^n\big) \Big| \leq K\sqrt{k\Delta_n}.$$

According to (S.10) and Lemma A.11 of Jacod et al. (2019), we can derive that

$$\frac{1}{k\Delta_n} \sum_{i=1}^{n(k)_t^q} \mathbb{E}\big(f(X'; k)_{m(k)_i^q}^2 \mid \mathcal{F}_{m(k)_{i-}^q}^n\big) = \sum_{i=1}^{n(k)_t^q} \frac{(\sigma_{m(k)_i^q}^n)^4 k\Delta_n}{(\alpha_{m(k)_i^q}^n)^2} + O_p(k\Delta_n \vee \Delta_n^\rho) \xrightarrow{\mathbb{P}} \int_0^t \frac{\sigma_s^4}{2\alpha_s^2} \mathrm{d}A_s.$$

Using (S.2), and by successive conditioning, we have

$$\frac{1}{(k\Delta_n)^2} \sum_{i=1}^{n(k)_t^q} \mathbb{E}\big(f(X'; k)_{m(k)_i^q}^4 \mid \mathcal{F}_{m(k)_{i-}^q}^n\big) \leq Kk\Delta_n \longrightarrow 0.$$

Now assume M is a bounded martingale that is orthogonal (in the martingale sense) to W. Then, by the orthogonality and the martingale properties of W and M, it is easy to see that

$$\frac{1}{\sqrt{k\Delta_n}} \sum_{i=1}^{n(k)_t^q} \mathbb{E}\big(f(X'; k)_{m(k)_i^q}(M_{m(k)_{i+}^q}^n - M_{m(k)_{i-}^q}^n) \mid \mathcal{F}_{m(k)_{i-}^q}^n\big) \xrightarrow{\mathbb{P}} 0.$$

When $M = W$, one can use Lemma S.2 and the same techniques used in the proofs of the previous lemmas to show that the LHS above converges to zero in probability. Now Theorem 2.2.15 in Jacod and Protter (2011) yields the following functional stable convergence in law:

$$\frac{1}{\sqrt{k\Delta_n}} F(X'; k, q)_t^n \xrightarrow{\mathcal{L}_s} \int_0^t \beta_s \mathrm{d}B_s^q,$$

with $\beta_s := \frac{\sigma_s^2}{\sqrt{2\alpha_s}}$ for a Brownian motion B^q that is independent of $\mathcal{F}$. It is obvious that the role of q is trivial in this part. Hence, the same conclusion holds for q'. (ii) Now we turn to the covariation/correlation between B^q and $B^{q'}$ here for $q \neq q'$ (the case $q = q'$ reduces to variance). According to (S.9), and upon successive conditioning, one can readily get that

$$\frac{1}{k\Delta_n} \sum_{i=1}^{n(k)_t^q} \sum_{\substack{j=1 \\ |j-i|\geq 2}}^{n(k)_t^{q'}} \Big| \mathbb{E}\big(f(X'; k)_{m(k)_i^q}^n f(X'; k)_{m(k)_j^{q'}}^n\big) \Big| \leq \frac{K}{k\Delta_n} \sum_{i=1}^{n(k)_t^q} \sum_{\substack{j=1 \\ |j-i|\geq 2}}^{n(k)_t^{q'}} (k\Delta_n)^4 \leq Kk\Delta_n.$$

Hence, we only need to consider the cases where $i = j$ or $|i - j| = 1$.

Without loss of generality, assume $q < q'$. Let us examine the case where $q' - q < 1/2$ first. Consider the case $j = i$. For notation simplicity, let

$$
\begin{aligned}
t(0)_i^n &= T^n_{m(k)^q_{i-}}, \quad t(1)_i^n = T^n_{m(k)^{q'}_{i-}}, \quad t(2)_i^n = T^n_{m(k)^q_i}; \\
t(3)_i^n &= T^n_{m(k)^{q'}_i}, \quad t(4)_i^n = T^n_{m(k)^q_{i+}}, \quad t(5)_i^n = T^n_{m(k)^{q'}_{i+}}.
\end{aligned}
$$

Obviously, we have $t(0)_i^n < t(1)_i^n < t(2)_i^n < t(3)_i^n < t(4)_i^n < t(5)_i^n$. Then we can rewrite

$$
\begin{aligned}
f(X';k)^n_{m(k)^q_i} &= \Big((X'_{t(2)_i^n} - X'_{t(1)_i^n}) + (X'_{t(1)_i^n} - X'_{t(0)_i^n})\Big)\Big((X'_{t(4)_i^n} - X'_{t(3)_i^n}) + (X'_{t(3)_i^n} - X'_{t(2)_i^n})\Big); \\
f(X';k)^n_{m(k)^{q'}_i} &= \Big((X'_{t(3)_i^n} - X'_{t(2)_i^n}) + (X'_{t(2)_i^n} - X'_{t(1)_i^n})\Big)\Big((X'_{t(5)_i^n} - X'_{t(4)_i^n}) + (X'_{t(4)_i^n} - X'_{t(3)_i^n})\Big).
\end{aligned}
$$

One observes that the only common term of $f(X';k)^n_{m(k)^q_i}$ and $f(X';k)^n_{m(k)^{q'}_i}$ after expansion is $(X'_{t(2)_i^n} - X'_{t(1)_i^n})(X'_{t(4)_i^n} - X'_{t(3)_i^n})$, as illustrated in Figure S.1. Upon using Lemma S.2 in the

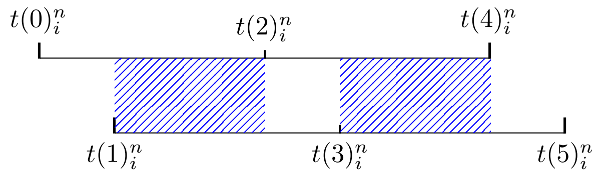

Figure S.1: Illustration of the case $q < q' < q + 1/2$.

way we prove, e.g., (S.9), one can show that the following after some elementary calculations:

$$
\Big|\mathbb{E}\big(f(X';k)^n_{m(k)^q_i} f(X';k)^n_{m(k)^{q'}_i} - (X'_{t(2)_i^n} - X'_{t(1)_i^n})^2 (X'_{t(3)_i^n} - X'_{t(4)_i^n})^2 \mid \mathcal{F}_{t(0)_i^n}\big)\Big| \le K(k\Delta_n)^3.
$$

Similar to (S.10), one can prove that

$$
\Big|\mathbb{E}\Big[\Big(\frac{X'_{t(2)_i^n} - X'_{t(1)_i^n}}{\sqrt{k\Delta_n}}\Big)^2 \Big(\frac{X'_{t(3)_i^n} - X'_{t(4)_i^n}}{\sqrt{k\Delta_n}}\Big)^2 - \frac{\sigma^4_{t(0)_i^n}}{\alpha^2_{t(0)_i^n}}\Big(1 - \frac{|q_k - q'_{k'}|}{k}\Big) \,\Big|\, \mathcal{F}_{t(0)_i^n}\Big]\Big| \le K(k\Delta_n \vee \Delta_n^\rho).
$$

In the case $|j - i| = 1$, there are no common terms. Hence, the conditional expectation of the product is of order $O_p(k\Delta_n)^3$. It then follows that, when $q < q' < q + 1/2$, the main term of the covariance

$$
\begin{aligned}
&\frac{1}{k\Delta_n} \sum_{i=1}^{n(k)_t^{q'}} \mathbb{E}\big(f(X';k)^n_{m(k)^q_i} f(X';k)^n_{m(k)^{q'}_i} \mid \mathcal{F}^n_{m(k)^q_i}\big) \\
&= \frac{\Phi_k^{q,q'}}{k} \sum_{i=1}^{n(k)_t^{q'}} \Big(2k\Delta_n \frac{\sigma^4_{t(0)_i^n}}{\alpha^2_{t(0)_i^n}} + o_p(k\Delta_n)\Big) \xrightarrow{\mathbb{P}} \frac{\Phi_k^{q,q'}}{k} \int_0^t \frac{\sigma_s^4}{\alpha_s^2} \mathrm{d}A_s.
\end{aligned}
$$

For the case $q + 1/2 \le q' < q + 1$, there are no common terms when $j = i$. Instead, there is a common term when $j = i + 1$.

In this case, one can define $t(6)_i^n := T^n_{m(k)^q_{i+1}}$ and $t(8)_i^n := T^n_{m(k)^q_{(i+1)+}}$ following the above manner. This scenario is illustrated in Figure S.2. Note that there is no common term when $q' = q + 1/2$ ($t(3)_i^n = t(4)_i^n$ and $t(5)_i^n = t(6)_i^n$). Following a similar analysis as above, one can

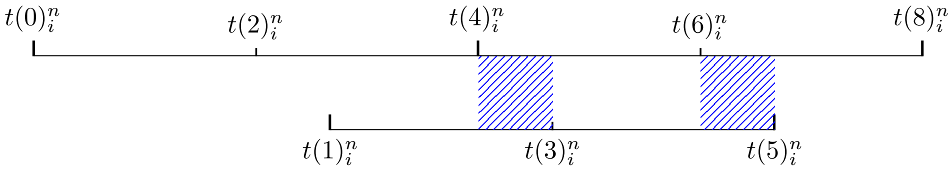

Figure S.2: Illustration of the case $q + 1/2 \le q' < q + 1$

derive the same covariance. The uniform convergence in q is trivial, as there are only finite q_k and uniformity reduces to finite dimensional convergence. Now the proof is complete. □

The next two lemmas will deal with the truncations used to remove jumps. For any process V, we define the truncated process $\overline{V}_t^n := \sum_{i=2}^{n_t} (\Delta_i^n V) \mathbf{1}_{\{|\Delta_i^n V| \le u_n\}}$.

Lemma S.7. Under the assumptions of Theorem S.6, we have

$$\frac{1}{\sqrt{k\Delta_n}} \sum_{i=1}^{n(k)_t^q} \mathbb{E}\Big(\Big| f(\overline{X}^n; k)_{m(k)_i^q}^n - f(X'; k)_{m(k)_i^q}^n \Big| \;\Big|\; \mathcal{F}_{i-k}^n \Big) \xrightarrow{\mathbb{P}} 0. \tag{S.18}$$

Proof. For any $j = 1, \dots, 2k$, define

$$x_j := \frac{\Delta^n_{m(k)_i^q - k + j} X}{\sqrt{\Delta_n}} \quad \text{and} \quad \tilde{x}_j := x_j \cdot 1_{\{|x_j| \le u_n \Delta_n^{-1/2}\}}.$$

Similarly, one can define x_j', $\tilde{x}_j'$, x_j'', and $\tilde{x}_j''$ by replacing X with X' or J accordingly. Note that $u_n \Delta_n^{-1/2} = a\Delta_n^{\varpi - 1/2}$ diverges to infinity.

For any integer l, let $A_{n,l} := \{|\Delta_l^n X'| \ge u_n\}$. The increment $\Delta_l^n X'$ is dominated by the Brownian motion part. Hence, the term $\Delta_l^n X' / \sqrt{\Delta_n}$ is (approximately) normally distributed conditional on $\mathcal{F}_{l-1}^n$, which implies that the probability $\mathbb{P}(A_{n,l}^c)$ decays exponentially to zero. Note that the impact of the assumed random sampling interval is negligible here. We can obtain $\sum_{l=1}^{n_t} \mathbb{P}(A_{n,l}^c) \longrightarrow 0, \text{ as } n \to 0, \forall t \in (0, \infty)$.

Following a similar argument as Jacod and Protter (2011), it suffices to prove the result on the set $\cap_{l=1}^{n_t} A_{n,l}$, on which we always have $\tilde{x}_j' = x_j'$. We introduce a new function $\mathbb{R}^{2k} \mapsto \mathbb{R}$: $h(x_1, \dots, x_{2k}) = (x_1 + \dots + x_k)(x_{k+1} + \dots + x_{2k})$. On the refined set, we have

$$\begin{aligned}
&\frac{1}{\Delta_n} \Big(f(X_{u_n}; k)_{m(q)_i^n}^n - f(X'; k)_{m(q)_i^n}^n \Big) = h(\tilde{x}_1, \dots, \tilde{x}_{2k}) - h(\tilde{x}_1', \dots, \tilde{x}_{2k}') \\
&= \sum_{j=1}^{2k} \Big(h(\tilde{x}_1, \dots, \tilde{x}_{j-1}, \tilde{x}_j, \tilde{x}_{j+1}', \dots, \tilde{x}_{2k}') - h(\tilde{x}_1, \dots, \tilde{x}_{j-1}, \tilde{x}_j', \tilde{x}_{j+1}', \dots, \tilde{x}_{2k}') \Big) \\
&= \sum_{j=k+1}^{2k} (\tilde{x}_1 + \dots + \tilde{x}_k)(\tilde{x}_j - \tilde{x}_j') + \sum_{j=1}^{k} (\tilde{x}_j - \tilde{x}_j')(\tilde{x}_{k+1}' + \dots + \tilde{x}_{2k}')
\end{aligned}$$

When $|x_j| \le u_n \Delta_n^{-1/2}$, we have $\tilde{x}_j - \tilde{x}_j' = x_j''$, where $|x_j''| \le 2u_n \Delta_n^{-1/2}$. Otherwise, we get $\tilde{x}_j - \tilde{x}_j' = -x_j'$, where $|x_j'| \le u_n$. It then follows that $|\tilde{x}_j - \tilde{x}_j'| \le K|\tilde{x}_j'' \wedge u_n \Delta_n^{-1/2}|$. A careful examine of the proof of Corollary 2.1.9 in Jacod and Protter (2011) indicates that our assumed random observation scheme does not make a real difference. Whenever a is a constant or measurable to $\mathcal{F}^n_{m(k)^q_i-k+j-1}$ (e.g., the estimated spot volatility prior to T^n_{j-1}), part (c) of the corollary implies that

$$\mathbb{E}\big(|\tilde{x}''_j \wedge u_n\Delta_n^{-1/2}| \mid \mathcal{F}^n_{j-1}\big) \le K u_n \Delta_n^{-1/2}\Delta_n^{1-\varpi}\phi_n \le K\Delta_n^{1/2}\phi_n,$$

where $\phi_n \to 0$ as $n \to \infty$. Even when a is not adaptive to $\mathcal{F}^n_{m(k)^q_i-k+j-1}$ (e.g., average integrated volatility multiplied by some constant), the above bound still holds as long as a_{j-1}, a_j, and $\tilde{x}''_j \wedge (a_j - a_{j-1})$ are both $O_p(1)$, where $a_j = \mathbb{E}(a|\mathcal{F}^n_{m(k)^q_i-k+j-1})$. It then follows that the LHS of (S.18) is bounded by $\frac{1}{\sqrt{k\Delta_n}}\sum_{i=1}^{n(k)^q_t} Kk^2\Delta_n^{3/2}\phi_n \le K\sqrt{k}\phi_n \to 0$. This completes the proof. □

Lemma S.8. Assume $\theta \in [0, \frac{3}{4})$, $\varpi \in (0, \frac{1}{2})$. Then, we have for any positive integer k,

$$\mathbb{E}\Big(k\Delta_n^{1-2\theta}\,\big|\overline{F}(Y;k)^n_t - F(Y;k)^n_t\big|\Big) \le K\Delta_n^{\frac{1}{2}-\varpi},$$
$$\mathbb{E}\Big(\Delta_n^{2-4\theta}|\overline{Q}(Y)^n_t - Q(Y)^n_t|\Big) \le K\Delta_n^{\frac{3}{2}-\varpi},$$

where $\overline{Q}(Y)^n_t$ is the Q statistic applied to the truncated returns.

Proof. Let $u^n_j := K\sqrt{\hat{c}^n_j}\Delta_n^{\varpi}$, where $\hat{c}^n_j$ is a noisy (hence biased) estimate of volatility. For example, one can set $\hat{c}^n_j := \frac{1}{d_n\Delta_n}\sum_{i=j-d_n+1}^{j}(\Delta^n_i Y)^2$ to be pre-estimator for some $d_n \to \infty$. Alternatively, one can set it to be the average realized variance. Following a similar analysis as in the proof of Lemma S.4, one can show that $\hat{c}^n_j = O_p(\Delta_n^{2\theta-1} \vee 1)$. It is easy to see that

$$\begin{aligned}&\big|(\Delta^n_j Y)\mathbf{1}_{\{|\Delta^n_j Y|\le u^n_j\}}(\Delta^n_{j'}Y)\mathbf{1}_{\{\{|\Delta^n_{j'}Y|\le u^n_{j'}\}\}} - (\Delta^n_j Y)(\Delta^n_{j'}Y)\big|\\ &\le \big|(\Delta^n_j Y)(\Delta^n_{j'}Y)\big|\big(\mathbf{1}_{\{|\Delta^n_j Y|>u^n_j\}} + \mathbf{1}_{\{|\Delta^n_{j'}Y|>u^n_{j'}\}}\big).\end{aligned}$$

When $\theta \in (\frac{1}{2}, \frac{3}{4})$, the Brownian motion part dominates the weak noise in the return process. The proof in this case basically follows the same idea as in the proof of Lemma S.7, hence is omitted here. When $\theta \in [0, \frac{1}{2}]$, we have the following estimates

$$\mathbb{P}(|\Delta^n_j Y| > u^n_j) \le \mathbb{E}\big((\Delta^n_j Y)^2/(u^n_j)^2\big) \le K\Delta_n^{1-2\varpi}; \quad \mathbb{E}\big((\Delta^n_j Y)^{2\iota}\big) \le K\Delta_n^{2\theta\iota}, \iota = 1, 2. \tag{S.19}$$

The above estimates and an application of Cauchy-Schwarz inequality yield

$$\mathbb{E}\Big(\big|(\Delta^n_j Y)(\Delta^n_{j'}Y)\big|^{\iota}\big(\mathbf{1}_{\{|\Delta^n_j Y|>u^n_j\}} + \mathbf{1}_{\{|\Delta^n_{j'}Y|>u^n_{j'}\}}\big)\Big) \le K\Delta_n^{2\iota\theta+\frac{1}{2}-\varpi}.$$

The results now readily follow. □

For any process V, we denote $\Delta^n_i V := V^n_i - V^n_{i-1}$.

Lemma S.9. We have

$$\frac{1}{\Delta_n}\sum_{i=2}^{n_t-1}(\Delta_i^n\overline{X}^n)^2(\Delta_{i+1}^n\overline{X}^n)^2 \xrightarrow{\mathbb{P}} \int_0^t \frac{\sigma_s^4}{\alpha_s^2}\mathrm{d}A_s.$$

Proof. One can follow the same argument as in the proof of Lemma S.7, using another h function defined as $h(x_1,x_2)=x_1^2x_2^2$ to show that

$$\frac{1}{\Delta_n}\sum_{i=2}^{n_t-1}\left((\Delta_i^n\overline{X}^n)^2(\Delta_{i+1}^n\overline{X}^n)^2-(\Delta_i^nX')^2(\Delta_{i+1}^nX')^2\right) \xrightarrow{\mathbb{P}} 0.$$

Therefore, it suffices to prove

$$\frac{1}{\Delta_n}\sum_{i=1}^{n_t-1}(\Delta_i^nX')^2(\Delta_{i+1}^nX')^2 \xrightarrow{\mathbb{P}} \int_0^t \frac{\sigma_s^4}{\alpha_s^2}\mathrm{d}A_s. \tag{S.20}$$

Let $\mathscr{S}_i^n=(\Delta_i^nX')^2(\Delta_{i+1}^nX')^2$, and $\overline{\mathscr{S}}_i^n=\mathscr{S}_i^n-\mathbb{E}\left(\mathscr{S}_i^n|\mathcal{F}_{i-1}^n\right)$. An application of (S.10) (with $k=1$) yields

$$\left|\sum_{i=1}^{n_t-1}\frac{\mathbb{E}\left(\mathscr{S}_i^n|\mathcal{F}_{i-1}^n\right)}{\Delta_n}-\Delta_n\sum_{i=1}^{n_t-1}\frac{(\sigma_{i-1}^n)^4}{(\alpha_{i-1}^n)^2}\right|\le K\Delta_n^{1\wedge\rho}.$$

(S.10) also implies that $\mathbb{E}\Big((\overline{\mathscr{S}}_i^n)^2\Big)\le K\Delta_n^4$. Therefore, we have

$$\mathbb{E}\Big((\textstyle\sum_{i=1}^{n_t-1}\overline{\mathscr{S}}_i^n/\Delta_n)^2\Big)=\displaystyle\sum_{i=1}^{n_t-1}\mathbb{E}\Big((\overline{\mathscr{S}}_i^n)^2/\Delta_n^2\Big)\le K\Delta_n.$$

The above two estimates and Lemma A.11 of Jacod et al. (2019) lead to

$$\textstyle\sum_{i=1}^{n_t-1}\mathscr{S}_i^n/\Delta_n \xrightarrow{\mathbb{P}} \displaystyle\int_0^t \frac{\sigma_s^4}{\alpha_s^2}\mathrm{d}A_s.$$

The proof is complete. □

Lemma S.10. Given two positive integers $k\le k'$, let $L=\mathrm{lcm}(k,k')$ be the least common multiple of k and k'. For any positive integer $l\in\{1,\dots,L/k'\}$, and $q_k\in\{0,1,\dots,2k-1\}$, $q'_{k'}\in\{0,1,\dots,2k'-1\}$, define

$$c(q_k,q'_{k'})_l:=(m'_l\wedge m_l-(m'_l-k')\vee(m_l-k))((m'_l+k')\wedge(m_l+k)-m'_l\vee m_l),$$

where

$$m'_l:=(2l-1)k'+q'_{k'},\quad m_l:=\Big(2\Big\lfloor\frac{m'_l-q_k}{2k}\Big\rfloor+1\Big)k+q_k.$$

Then, we have

$$\sum_{q_k=0}^{2k-1}\sum_{q'_{k'}=0}^{2k'-1}\sum_{l=1}^{L/k'} c(q_k, q'_{k'})_l = 2L\overline{\Phi}_{k,k'}, \tag{S.21}$$

where $\overline{\Phi}_{k,k'} := k^3 - \frac{1}{3}\big[(2k-k')^3 - (2k-k')\big]\mathbf{1}_{\{k'\le 2k\}}$.

Proof. For simplicity, we denote c_l as a short form of $c(q_k, q'_{k'})_l$. We first define two remainders:

$$a \equiv k' \pmod{2k}, \quad b \equiv (2l-1)a + q'_k - q_k \pmod{2k}.$$

It is then trivial to obtain $m'_l = \mathscr{C}_l + b,\ m_l = \mathscr{C}_l + k$, where

$$\mathscr{C}_l := 2\Big\lfloor\frac{k'}{2k}\Big\rfloor(2l-1)k + 2\Big\lfloor\frac{(2l-1)a + q'_{k'} - q_k}{2k}\Big\rfloor k + q_k.$$

One observation is that c_l remains unchanged when m'_l and m_l change by the same amount, in particular, $\mathscr{C}_l$. Consequently, it follows that

$$c_l = [b\wedge k - (b-k')\vee 0][(b+k')\wedge 2k - b\vee k]. \tag{S.22}$$

Another key observation here is that the remainder b will take on all values in $0,\ldots,2k-1$ when q_k runs over the same set, regardless of the values of l and $q'_{k'}$. This implies that we can first sum c_l over q_k and the result will be invariant to l and $q'_{k'}$.

When $k' \ge 2k$, it is easy to check that

$$c_l = (b\wedge k)[k - (b-k)\vee 0] = k[b\wedge(2k-b)].$$

Therefore, we obtain, by observing the dependence of c_l on q_k, that

$$\overline{\Phi}_{k,k'} := \sum_{q_k=0}^{2k-1} c_l = \sum_{b=0}^{2k-1} k[b\wedge(2k-b)] = k^3.$$

When $k \le k' < 2k$, there are four categories:

$$c_l = \begin{cases} b(b+k'-k) & \text{if } 0\le b< 2k-k'; \\ bk & \text{if } 2k-k'\le b\le k; \\ (2k-b)k & \text{if } k\le b\le k'; \\ (2k-b)(k'+k-b) & \text{if } k'< b< 2k. \end{cases}$$

Note that we consider the case $b = k$ in both the second and the third categories, for convenience. Also note that the above four categories are symmetric around $b = k$ (which is counted twice). Thus, we can get

$$
\begin{aligned}
\overline{\Phi}_{k,k'} &= \sum_{q_k=0}^{2k-1} c_l = 2\Big(\sum_{b=1}^{2k-k'-1} b(b+k'-k) + \sum_{b=2k-k'}^{k} bk \Big) - k^2 \\
&= \frac{1}{3}(2k-k')(2k-k'-1)(k-k'-1) + k(3k-k')(k'-k+1) - k^2 \\
&= \frac{1}{3}\Big(3k^3 + (k'-2k)^3 - (k'-2k)\Big).
\end{aligned}
$$

In particular, if we plug $k' = 2k$ into the last expression, we get $\overline{\Phi}_{k,2k} = k^3$, which is the same as the case $k' \geq 2k$. That said, the term $\overline{\Phi}_{k,2k}$ is "continuous" at the point $k' = 2k$. Hence, for $k' \geq k$, we can write

$$
\overline{\Phi}_{k,k'} = \frac{1}{3}\Big(3k^3 - \big[(2k-k')^3 - (2k-k')\big]1_{\{k' \leq 2k\}}\Big).
$$

Then, the triple summation of c_l is simply $2L\overline{\Phi}_{k,k'}$. □

Theorem S.11. For any two integers k, k', we have the jointly convergence in law

$$
\left(\frac{F(X';k)_t^n}{\sqrt{\Delta_n}}, \frac{F(X';k')_t^n}{\sqrt{\Delta_n}} \right) \xrightarrow{\mathcal{L}_s} (\overline{\mathcal{Z}}_t, \overline{\mathcal{Z}}'_t),
$$

where the limits are defined on an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$. Conditional on $\mathcal{F}$, $(\overline{\mathcal{Z}}_t, \overline{\mathcal{Z}}'_t)$ are jointly Gaussian, with (co)-variances $\tilde{\mathbb{E}}\big(\overline{\mathcal{Z}}_t \overline{\mathcal{Z}}'_t \mid \mathcal{F}\big) = \Phi_{k,k'} \int_0^t \frac{\sigma_s^4}{\alpha_s^2} \mathrm{d}A_s$.

Proof. We consider two distinct sampling mechanisms: one characterized by a window size k beginning at $q_k = \lfloor 2qk \rfloor$ for some $q \in [0,1)$, and the other by a window size k' starting at $q'_{k'} = \lfloor 2q'k' \rfloor$ for some $q' \in [0,1)$. Assuming $k \leq k'$ without loss of generality, we define $L = \mathrm{lcm}(k, k')$ to represent the *least common multiple* of the window sizes k and k'.

To determine the covariance between $F(X';k,q)_t^n$ and $F(X';k',q')_t^n$, we can leverage the two sampling schemes. It is evident that a repeating pattern emerges every $2L$ observations. This pattern corresponds to L/k' non-overlapping windows when using the (k',q')-scheme. The covariance structure is determined by these intervals. For simplicity, we will focus our analysis on the first interval that contains the first L observations, as the remaining intervals follow the same analysis.

For $l = 1, \ldots, L$, the middle point of the (k',q')-scheme is given by m'_l, defined in Lemma S.10. When considering the local windows from the (k,q)-scheme, if the middle point of a window is more than k observations away from m'_l, the contribution to the covariance from these two windows will be asymptotically negligible. Therefore, we only need to focus on the windows from the (k,q)-scheme whose middle point is within k observations from m'_l. The middle point of such a window is given by m_l, which is also defined in Lemma S.10.

In this case, there are only two possibilities: $m_l < m'_l$ and $m_l \geq m'_l$. Simple illustrations of these possibilities are provided in Figure S.3. It is important to note that in the first possibility, $m_l - k$ could be smaller than $m'_l - k'$, while in the second possibility, $m_l + k$ could be larger than $m'_l + k'$. This explains the two multiplicative factors of $c(q_k, q'_{k'})_l$ defined in Lemma S.10, which are represented by the shaded intervals in Figure S.3.

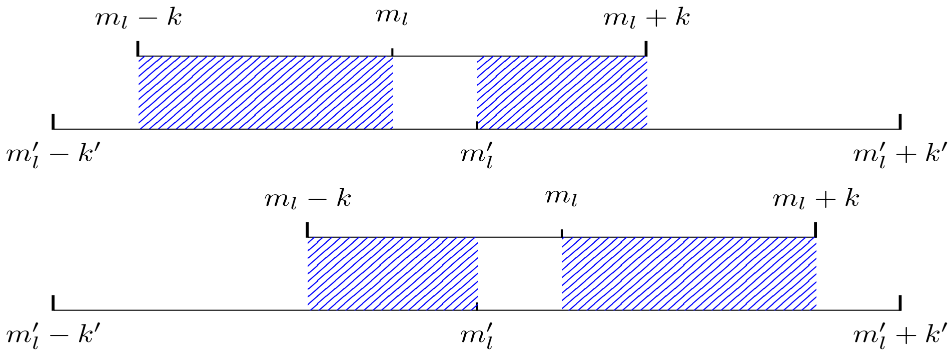

Figure S.3: Two cases illustrating the relationship between m_l and m'_l when $|m_l - m'_l| \leq k$.

Following Lemma A.11 of Jacod et al. (2019), the $\mathcal{F}$-conditional covariance of $F(X'; k, q)^n_t$ and $F(X'; k', q')^n_t$ can be simplified as:

$$\begin{aligned}
&\mathrm{Cov}\Big(\frac{1}{\sqrt{k\Delta_n}} F(X'; k, q)^n_t, \frac{1}{\sqrt{k'\Delta_n}} F(X'; k', q')^n_t \,|\, \mathcal{F}\Big) \\
&= \sum_{v=1}^{\lfloor \frac{n_t}{2L} \rfloor} \frac{\sigma^4_{T^n_{2(v-1)L}}}{2\alpha^2_{T^n_{2(v-1)L}}} \sum_{l=1}^{\frac{L}{k'}} \frac{c_l}{\sqrt{kk'}} \Delta_n + o_p(1) = \sum_{l=1}^{\frac{L}{k'}} \frac{c_l}{\sqrt{kk'}L} \int_0^t \frac{\sigma^4_s}{2\alpha^2_s} \mathrm{d}A_s + o_p(1).
\end{aligned}$$

The bilinear property of the covariance implies that

$$\begin{aligned}
&\mathrm{Cov}\Big(\frac{1}{\sqrt{k\Delta_n}} F(X'; k)^n_t, \frac{1}{\sqrt{k'\Delta_n}} F(X'; k')^n_t | \mathcal{F}\Big) \\
&= \frac{1}{2k} \frac{1}{2k'} \sum_{q_k=0}^{2k-1} \sum_{q'_{k'}=0}^{2k'-1} \mathrm{Cov}\Big(\frac{1}{\sqrt{k\Delta_n}} F(X'; k, q)^n_t, \frac{1}{\sqrt{k'\Delta_n}} F(X'; k', q')^n_t \,|\, \mathcal{F}\Big) \\
&= \frac{1}{4kk'} \sum_{q_k=0}^{2k-1} \sum_{q'_{k'}=0}^{2k'-1} \sum_{l=1}^{L/k'} \frac{c_l}{\sqrt{kk'}L} \int_0^t \frac{\sigma^4_s}{2\alpha^2_s} \mathrm{d}A_s + o_p(1).
\end{aligned}$$

Thus, by Lemma S.10, we have

$$\mathbf{Cov}\Big(\frac{1}{\sqrt{\Delta_n}} F(X'; k)^n_t, \frac{1}{\sqrt{\Delta_n}} F(X'; k')^n_t | \mathcal{F}\Big) \xrightarrow{\mathbb{P}} \frac{\overline{\Phi}_{k,k'}}{2kk'} \int_0^t \frac{\sigma^4_s}{2\alpha^2_s} \mathrm{d}s = \Phi_{k,k'} \int_0^t \frac{\sigma^4_s}{\alpha^2_s} \mathrm{d}s.$$

Now the convergence follows from Theorem S.6. □

References

Andersen, T. G., Li, Y., Todorov, V., and Zhou, B. (2023). Volatility measurement with pockets of extreme return persistence. *Journal of Econometrics*, 237(2):105048.

Christensen, K., Oomen, R., and Renò, R. (2022). The drift burst hypothesis. *Journal of Econometrics*, 227(2):461–497.

Christensen, K., Oomen, R. C., and Podolskij, M. (2014). Fact or friction: Jumps at ultra high frequency. *Journal of Financial Economics*, 114(3):576–599.

Jacod, J., Li, Y., and Zheng, X. (2017). Statistical properties of microstructure noise. *Econometrica*, 85(4):1133–1174.

Jacod, J., Li, Y., and Zheng, X. (2019). Estimating the integrated volatility with tick observations. *Journal of Econometrics*, 208(1):80–100.

Jacod, J. and Protter, P. E. (2011). *Discretization of Processes*, volume 67. Springer Science & Business Media.

Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes*, volume 288. Springer-Verlag Berlin.

Laurent, S., Renò, R., and Shi, S. (2024). Realized drift. *Journal of Econometrics*, page 105813.

Li, Z. M. and Yang, X. (2025). Multi-horizon test for market frictions. *working paper*.