Estimating smooth structural change in cointegrated system with nearly or mildly integrated regressors

Yu Bai* Hsein Kew

City University of Macau *Monash University*

February 2026

In this paper, we derive the asymptotic properties of the Nadaraya-Watson type local level regression estimator of time–varying cointegrating coefficients when the regressors are either nearly or mildly integrated with autoregressive roots of the form $\rho_{ni} = 1 + c_i/n^{\alpha}$, where $\alpha \in (0, 1]$ and $c_i < 0$ are constant parameters. In nearly integrated case, it is shown that weighted signal matrix becomes asymptotically singular and a rotation decomposition is developed to restore the limit theory. In mildly integrated case, it is shown that the signal matrix is asymptotically well-behaved the estimator is consistent, so long as $n^{1-\alpha}h \to \infty$, where $h := h_n$ is the bandwidth parameter. Achieving standard asymptotic normality, however, requires bias correction and stronger rate conditions. The theoretical findings are illustrated via an extensive Monte Carlo study.

Keywords: Time–varying cointegration; Near integration; Mild integration; Local level regression

JEL-Classification: C13; C14; C32

*Corresponding to: Yu Bai, Faculty of Finance, City University of Macau. Avenida Xian Xing Hai - No. 81-121, Macau, China. Email: yubai@cityu.edu.mo.

1 — Introduction

There has been growing interest in the limit theory of nonparametric time-varying coefficient cointegration regressions. This literature dates back to Park and Hahn 1999, who develop a series estimator for time-varying cointegrating vectors in a single-equation framework. Within a kernel estimation setting, Phillips, Li, and Gao 2017 show that conventional asymptotic theory under stationarity breaks down when regressors are integrated, due to a kernel-induced degeneracy in the weighted signal matrix. They introduce a novel rotational decomposition to resolve this issue and establish the limit theory. Building on this work, Li, Phillips, and Gao 2020 extend the framework to allow for cointegrated regressors and mixtures of stochastic and deterministic trend components, while Kapetanios et al. 2020 develop inference procedures for time-varying cointegration and apply them to long-run macroeconomic ratios for the UK (UK great ratios).

A common feature of this literature is the assumption that regressors are exactly integrated, with autoregressive roots fixed at unity. However, a large body of empirical evidence suggests that economic and financial time series often exhibit strong persistence without being precisely unit root processes, with autoregressive roots lying in a neighborhood of unity. For example, Kostakis, Magdalinos, and Stamatogiannis 2015 study stock return predictors that may be stationary, mildly integrated, nearly integrated, or integrated, while Shi and Phillips 2023 model housing market fundamentals using regressors that are assumed to be mildly or nearly integrated.

As in the related literature, consider the autoregressive coefficient matrix for the regressors $\mathbf{X}_t$:

$$\mathbf{R}_n = \mathbf{I}_k + \frac{\mathbf{C}}{n^{\alpha}},$$

where $\mathbf{C} = \mathrm{diag}(c_1, \ldots, c_k)$ contains non-positive diagonal elements and $\alpha \in [0, 1]$. This specification nests stationary regressors ($\alpha = 0$), local-to-unity processes ($\alpha = 1$; see Chan and Wei 1987; Phillips 1987), and the exact unit root case as special cases. When $c_i < 0$ and $0 < \alpha < 1$, we follow Phillips and Magdalinos 2007a, Phillips and Magdalinos 2007b, and Magdalinos and Phillips 2009 in referring to such processes as *mildly integrated*. Following Phillips 1988, the local-to-unity case is also referred to as *nearly integrated*.

Despite their empirical relevance, allowing for mildly or nearly integrated regressors substantially complicates nonparametric estimation. In particular, kernel-based methods face additional challenges due to the interaction between persistent regressors and local weighting schemes. To address robustness concerns for instability in predictive regressions, Liu, Phillips, and Zhang 2025 propose a series-based approach for time-varying stock return predictability that extends the framework of Kostakis, Magdalinos, and Stamatogiannis 2015. However, the exact large-sample properties of kernel estimators in time-varying coefficient cointegration models, especially in the presence of mildly or nearly integrated regressors, have not yet been fully worked out. This paper aims to fill this gap.

In this paper, we establish limit theory for Nadaraya-Watson local level estimation of time-varying cointegrating coefficients in a cointegrated system when regressors are either nearly or mildly integrated. In the nearly integrated case, we show that, as in the exactly integrated setting, conventional kernel methods encounter a degeneracy problem due to asymptotic singularity of the weighted signal matrix. Following Phillips, Li, and Gao 2017, we employ a rotational decomposition to develop the limit theory. The resulting asymptotics exhibit two distinct convergence rates in the different directions associated with the rotations: a *type I* super-consistency rate of $n\sqrt{h}$ and a *type II* super-consistency rate of nh, where $h := h_n$ denotes the bandwidth parameter.

The results differ fundamentally in the mildly integrated case. We show that the weighted signal matrix is asymptotically well behaved and that the estimator is consistent, provided that $n^{1-\alpha}h \to \infty$. This condition is stronger than the standard requirement $nh \to \infty$ under stationarity and reflects the reduction in effective local sample size induced by regressor persistence. Achieving standard asymptotic normality, however, requires bias correction and stronger rate conditions. In particular, the bandwidth must satisfy $n^{(1+\alpha)/2}h^{5/2} \to c \in (0, \infty)$ together with the additional conditions $n^{1-\alpha}h^2 \to \infty$ and $n^{1/2-\alpha}\sqrt{h} \to 0$.

The theoretical findings are illustrated via an extensive Monte Carlo study. In the nearly integrated case, the simulation results show that the bias decays at a faster rate in one rotation direction than in the orthogonal direction. In mildly integrated case, we examine both mean squared error and 95% coverage probability of the estimator and its bias-corrected counterpart across a range of persistence parameters α and dependence structures. The results show that the bias-corrected estimator delivers substantial efficiency gains for relatively small values of α. As α increases, however, the impact of bias correction diminishes and the two estimators exhibit nearly identical behavior. In addition, the coverage probabilities of confidence intervals display a hump-shaped pattern as α varies, reflecting the finite-sample trade-off between eliminating simultaneous-equation bias and satisfying the rate conditions required for valid variance approximation. Overall, the Monte Carlo evidence closely aligns with the theoretical analysis and highlights the practical relevance of the proposed bias correction in empirically relevant settings.

The rest of the paper is organised as follows. Section 2 introduces the model, lays out the assumptions, and presents the estimator. Section 3 develops the asymptotic theory for nearly integrated system and Section 4 establishes the asymptotic theory for mildly integrated system. Section 5 presents an extensive Monte Carlo study and Section 6 concludes. All mathematical proofs and additional technical details are collected in the Appendix, and the Online Supplement contains additional lemmas needed for the theoretical development.

2 — Model and the estimator

We consider the triangular system (cf. Phillips 1991) with time-varying coefficients

$$\mathbf{y}_t = \mathbf{A}_t \mathbf{X}_t + \mathbf{u}_{0t}, \tag{1}$$

$$\mathbf{X}_t = \mathbf{R}_n \mathbf{X}_{t-1} + \mathbf{u}_{xt}, \tag{2}$$

$$\mathbf{R}_n = \mathbf{I}_k + \frac{\mathbf{C}}{n^{\alpha}}, \quad \alpha \in (0,1], \ \ \mathbf{C} = \text{diag}\left(c_1, \cdots, c_k\right), \tag{3}$$

for $t = 1, 2, \cdots, n$. The system is initialized at some $\mathbf{X}_0 = O_p(1)$. $\mathbf{A}_t := \mathbf{A}(t/n)$ is an $d \times k$ matrix of time-varying "cointegrating" coefficients, with elements measured as functions of scaled time point t/n. c_i is assumed to be negative for all i, where $i = 1, 2, \cdots, k$.

For any fixed $z_0 \in (0,1)$, under a suitable smoothness condition on $\mathbf{A}(\cdot)$ (Assumption 2), we have

$$\mathbf{A}(t/n) = \mathbf{A}(z_0) + O\left(\left|\frac{t}{n} - z_0\right|\right) \approx \mathbf{A}(z_0), \tag{4}$$

where t/n lies in a small neighbourhood of z_0. The Nadaraya-Watson type local level regression estimator of $\mathbf{A}(z_0)$ has the usual form given by

$$\widehat{\mathbf{A}}_n(z_0) = \left[\sum_{t=1}^{n} K\left(\frac{t - nz_0}{nh}\right)\mathbf{y}_t \mathbf{X}_t^{\top}\right]\left[\sum_{t=1}^{n} K\left(\frac{t - nz_0}{nh}\right)\mathbf{X}_t \mathbf{X}_t^{\top}\right]^{-1}, \tag{5}$$

where $K(\cdot)$ is some kernel function, and h is a bandwidth parameter. We make the following assumptions on the kernel function $K(\cdot)$ and bandwidth h.

Assumption 1. (i) The kernel function $K(\cdot)$ is Lipschitz-continuous, positive, symmetric, and has compact support $[-1,1]$ with $\int_{-1}^{1} K(u)du = 1$.
(ii) The bandwidth $h := h_n$ satisfies $h \to 0$ and $nh \to \infty$, as $n \to \infty$.

The smoothness condition on $\mathbf{A}(\cdot)$ is given below.

Assumption 2. Let $a^{(i_1,i_2)}(t/n)$ be the (i_1, i_2)th elements in $\mathbf{A}(t/n)$, where $i_1 = 1, 2, \cdots, d$, $i_2 = 1, 2, \cdots, k$. $\mathbf{A}^{(i_1,i_2)}(\cdot)$ is a real function and is twice continuously differentiable on $(0, 1)$.

We assume that the innovations $\mathbf{u}_t = \left(\mathbf{u}_{0t}^\top, \mathbf{u}_{xt}^\top\right)^\top$ are determined according to the linear process

$$\mathbf{u}_t = \mathbf{\Phi}(\mathcal{L})\boldsymbol{\varepsilon}_t = \sum_{j=0}^{\infty} \mathbf{\Phi}_j \boldsymbol{\varepsilon}_{t-j}, \tag{6}$$

where $\mathbf{\Phi}(\mathcal{L}) = \sum_{j=0}^{\infty} \mathbf{\Phi}_j \mathcal{L}^j$, $\mathbf{\Phi}_j$ is a sequence of $(d+k) \times (d+k)$ matrices, $\mathcal{L}$ is the lag operator, and $(\boldsymbol{\varepsilon}_t)_t$ is a sequence of independent and identically distributed $(i.i.d.)$ random vectors with dimension $d + k$. We impose the following conditions.

Assumption 3. (i) Let $(\boldsymbol{\varepsilon}_t)_t$ be $i.i.d.$ $(d+k)$-dimensional random vectors with $\mathbb{E}(\boldsymbol{\varepsilon}_t) = \mathbf{0}$, $\mathbf{\Sigma}_\varepsilon := \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t^\top) > 0$, and $\mathbb{E}\|\boldsymbol{\varepsilon}_t\|^p < \infty$, for some $p > 4$. (ii) The coefficient matrices in (6) satisfy $\sum_{j=0}^{\infty} j\|\mathbf{\Phi}_j\| < \infty$ and $|\mathbf{\Phi}(1)| \neq 0$.

Under Assumption 3, $\mathbf{u}_t$ has covariance matrix $\mathbf{\Sigma} = \sum_{j=0}^{\infty} \mathbf{\Phi}_j \mathbf{\Sigma}_\varepsilon \mathbf{\Phi}_j^\top$, with $\mathbb{E}\|\mathbf{u}_t\|^p < \infty$, for some $p > 4$. By functional central limit theory (FCLT) for linear process (cf. Phillips and Solo 1992), we have

$$n^{-1/2} \sum_{t=1}^{n} u_t \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathbf{\Omega}\right), \tag{7}$$

with covariance matrix $\mathbf{\Omega} = \mathbf{\Phi}(1)\mathbf{\Sigma}_\varepsilon \mathbf{\Phi}(1)^\top > 0$ and $\mathbf{\Phi}(1) = \sum_{j=0}^{\infty} \mathbf{\Phi}_j$. For the ease of theoretical development, we partition $\mathbf{\Sigma}$, $\mathbf{\Phi}(1)$, and $\mathbf{\Omega}$ comfortably with $\mathbf{u}_t$ as

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{00} & \mathbf{\Sigma}_{0x} \\ \mathbf{\Sigma}_{x0} & \mathbf{\Sigma}_{xx} \end{bmatrix}, \quad \mathbf{\Phi}(1) = \begin{bmatrix} \mathbf{\Phi}_0(1) \\ \mathbf{\Phi}_x(1) \end{bmatrix}, \quad \mathbf{\Omega} = \begin{bmatrix} \mathbf{\Omega}_{00} & \mathbf{\Omega}_{0x} \\ \mathbf{\Omega}_{x0} & \mathbf{\Omega}_{xx} \end{bmatrix},$$

where $\mathbf{\Phi}_0(1)$ and $\mathbf{\Phi}_x(1)$ are of dimension $d \times (d+k)$ and $k \times (d+k)$, respectively. Using the Beveridge–Nelson (BN) decomposition, we have the following decomposition for $\mathbf{u}_t$:

$$\mathbf{u}_t = \mathbf{\Phi}(1)\boldsymbol{\varepsilon}_t - \Delta\widetilde{\boldsymbol{u}}_t, \text{ for } \widetilde{\mathbf{u}}_t = \sum_{j=0}^{\infty} \widetilde{\mathbf{\Phi}}_j \boldsymbol{\varepsilon}_{t-j}, \quad \widetilde{\mathbf{\Phi}}_j = \sum_{k=j+1}^{\infty} \mathbf{\Phi}_j, \tag{8}$$

where . Corresponding to the partition of $\mathbf{u}_t$, we write

$$\mathbf{u}_{0t} = \sum_{j=0}^{\infty} \mathbf{\Phi}_{0j}\boldsymbol{\varepsilon}_{t-j}, \quad \widetilde{\mathbf{u}}_{0t} = \sum_{j=0}^{\infty} \widetilde{\mathbf{\Phi}}_{0j}\boldsymbol{\varepsilon}_{t-j}, \tag{9}$$

and also

$$\mathbf{u}_{xt} = \sum_{j=0}^{\infty} \mathbf{\Phi}_{xj}\boldsymbol{\varepsilon}_{t-j}, \quad \widetilde{\mathbf{u}}_{xt} = \sum_{j=0}^{\infty} \widetilde{\mathbf{\Phi}}_{xj}\boldsymbol{\varepsilon}_{t-j}, \tag{10}$$

where $\mathbf{\Phi}_{0j}$, $\widetilde{\mathbf{\Phi}}_{0j}$ are of dimension $d \times (d+k)$ and $\mathbf{\Phi}_{xj}$, $\widetilde{\mathbf{\Phi}}_{xj}$ are of dimension $K \times (d+k)$, respectively.

In addition, the limit theory also involves the partitioned components of the one-sided long-run covariance matrix

$$\mathbf{\Lambda} := \begin{bmatrix} \mathbf{\Lambda}_{00} & \mathbf{\Lambda}_{0x} \\ \mathbf{\Lambda}_{x0} & \mathbf{\Lambda}_{xx} \end{bmatrix} = \sum_{j=1}^{\infty} \mathbb{E}\left(\mathbf{u}_t \mathbf{u}_{t-j}^{\top}\right).$$

As usual, we have $\mathbf{\Omega} = \mathbf{\Sigma} + \mathbf{\Lambda} + \mathbf{\Lambda}^{\top}$ and

$$\mathbf{\Lambda} = \sum_{j=1}^{\infty}\sum_{k=0}^{\infty} \mathbf{\Phi}_{k+j}\mathbf{\Sigma}_{\varepsilon}\mathbf{\Phi}_k^{\top} = \sum_{k=0}^{\infty}\left(\sum_{j=1}^{\infty} \mathbf{\Phi}_{k+j}\right)\mathbf{\Sigma}_{\varepsilon}\mathbf{\Phi}_k^{\top} = \sum_{k=0}^{\infty} \widetilde{\mathbf{\Phi}}_k \mathbf{\Sigma}_{\varepsilon}\mathbf{\Phi}_k^{\top} = \mathbb{E}\left(\widetilde{\mathbf{u}}_t \mathbf{u}_t^{\top}\right).$$

3 — Asymptotic theory for nearly integrated system

In this section[1], we derive the asymptotic properties of the local level regression estimator (5) when regressors $\mathbf{X}_t$ are nearly integrated. We shall first consider the asymptotic behavior of the

1. Throughout Section 3 and the subsequent proofs in Sections A.1, A.2, and B, matrices involving the inverse of the weighted signal matrix are written using the Moore-Penrose generalized inverse $(\cdot)^{+}$ rather than the usual inverse $(\cdot)^{-1}$, to accommodate the possible degeneracy of the limit matrix. Once nonsingularity is established $a.s.$, the generalized inverse coincides with the ordinary inverse, and the distinction between $(\cdot)^{+}$ and $(\cdot)^{-1}$ becomes immaterial.

weighted signal matrix

$$\mathbf{\Delta}_n(z_0) := \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathbf{X}_t\mathbf{X}_t^\top.$$

This is characterised in the next lemma.

Lemma 1. If Assumptions 1-3 are satisfied. Then, for any fixed $z_0 \in (0,1)$,

$$\frac{1}{n^2h}\mathbf{\Delta}_n(z_0) \xrightarrow{d} \mathbf{K}_C(z_0)\mathbf{K}_C^\top(z_0).$$

Observe that $\mathbf{K}_C(z_0)$ is a Gaussian process which, for fixed $z_0 \in (0,1)$, has the distribution

$$\mathbf{K}_C(z_0) \equiv \mathcal{N}\left(\mathbf{0}_{k\times 1}, \mathbf{\Sigma}_K(z_0)\right),$$

where $\mathbf{\Sigma}_K(z_0) := \int_0^{z_0} e^{(z_0-s)\mathbf{C}}\mathbf{\Omega}_{xx}e^{(z_0-s)\mathbf{C}'}ds$. Consequently, $\mathbf{K}_C(z_0)\mathbf{K}_C^\top(z_0)$ is a (possibly singular) Wishart variate with one degree of freedom and scale matrix $\mathbf{\Sigma}_K(z_0)$, and hence has rank one almost surely whenever $\mathbf{\Sigma}_K(z_0)$ is nonsingular. Therefore, the limiting matrix in Lemma 1 is rank one, implying that $\frac{1}{n^2h}\mathbf{\Lambda}_{n1}(z_0)$ is asymptotically singular whenever $\dim(\mathbf{X}_t) = k > 1$.

To deal with this limiting degeneracy issue, we follow the rotation decomposition approach introduced in Phillips, Li, and Gao 2017 for the exactly integrated case. Define $\mathbf{k} \equiv \mathbf{k}_{z_0} = \mathbf{K}_C(z_0)$ and set

$$\mathbf{q} = \frac{\mathbf{k}}{(\mathbf{k}^\top\mathbf{k})^{1/2}} = \frac{\mathbf{k}}{\|\mathbf{k}\|}.$$

Let $\mathbf{q}_n^\perp$ be the $k \times (k-1)$ be the orthogonal component matrix such that $\mathbf{Q} = \left(\mathbf{q}, \mathbf{q}^\perp\right)$, $\mathbf{Q}^\top\mathbf{Q} = \mathbf{I}_k$. The sample version of these quantities is

$$\mathbf{q}_n = \frac{\mathbf{k}_n}{(\mathbf{k}_n^\top\mathbf{k}_n)^{1/2}} = \frac{\mathbf{k}_n}{\|\mathbf{k}_n\|}, \quad \mathbf{k}_n \equiv \mathbf{k}_n(z_0) = \frac{1}{\sqrt{n}}\mathbf{X}_{z_n}.$$

Moreover, let $\mathbf{Q}_n = \left(\mathbf{q}_n, \mathbf{q}_n^\perp\right)$, $\mathbf{Q}_n^\top\mathbf{Q}_n = \mathbf{I}_k$ and defined the standardization matrix

$$\mathbf{D}_n = \text{diag}\left\{n\sqrt{h}, (nh)\mathbf{I}_{k-1}\right\}.$$

Note that both $\mathbf{Q}$ and $\mathbf{Q}_n$ are random, path dependent, and localized to the point z_0 and $z_n = \lfloor (z_0 - h)n \rfloor$, respectively.

The following theorem gives the asymptotic distribution of $\mathbf{A}_n(z_0)$.

Theorem 1. If Assumptions 1-3 are satisfied and $nh^{3/2} \to c \in (0, \infty)$. For any fixed $z_0 \in (0, 1)$, we have

$$\left(\widehat{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0)\right) \mathbf{Q}_n \mathbf{D}_n - \mathbb{B}_n(z_0) \xrightarrow{d} \mathbf{\Gamma}(z_0)\mathbf{\Delta}^{+}(z_0). \tag{11}$$

$\mathbb{B}_n(z_0)$ is a bias process of order $nh^{3/2}$ in probability. $\Delta(z_0)$ is defined as

$$\mathbf{\Delta}(z_0) := \begin{bmatrix} \Delta(z_0,1) & \mathbf{\Delta}(z_0,2) \\ \mathbf{\Delta}(z_0,2)^\top & \mathbf{\Delta}(z_0,3) \end{bmatrix}$$

with

$$\begin{aligned}
\mathbf{\Delta}(z_0,1) &= \mathbf{k}^\top \mathbf{k}, \\
\mathbf{\Delta}(z_0,2) &= \sqrt{2}\,\|\mathbf{k}\| \left[\int_{-1}^{1} K(r)\mathbf{K}_C^{*\top}\left(\frac{r+1}{2}\right) dr\right] \mathbf{q}^\perp, \\
\mathbf{\Delta}(z_0,3) &= 2\left(\mathbf{q}^\perp\right)^\top \left[\int_{-1}^{1} K(r)\mathbf{K}_C^{*}\left(\frac{r+1}{2}\right)\mathbf{K}_C^{*\top}\left(\frac{r+1}{2}\right) dr\right] \mathbf{q}^\perp,
\end{aligned}$$

and is nonsingular $a.s.$. $\Gamma(z_0)$ is defined as

$$\mathbf{\Gamma}(z_0) = \begin{bmatrix} \mathbf{\Gamma}(z_0,1) & \mathbf{\Gamma}(z_0,2) \end{bmatrix},$$

with

$$\begin{aligned}
\mathbf{\Gamma}(z_0,1) &= \sqrt{2}\,\|\mathbf{k}\| \int_{-1}^{1} K(r) d\mathbf{B}_0^{*}\left(\frac{r+1}{2}\right), \\
\mathbf{\Gamma}(z_0,2) &= 2\left[\int_{-1}^{1} K(r) d\mathbf{B}_0^{*}\left(\frac{r+1}{2}\right)\mathbf{K}_C^{*\top}\left(\frac{r+1}{2}\right) + \frac{1}{2}\left(\mathbf{\Sigma}_{0x} + \mathbf{\Lambda}_{0x}\right)\right] \mathbf{q}^\perp.
\end{aligned}$$

and $\left(\mathbf{K}_C^{\top *}(r), \mathbf{B}_0^{\top *}(r)\right)^\top$ is an independent copy of $\left(\mathbf{K}_C^{\top}(r), \mathbf{B}_0^{\top}(r)\right)^\top$.

Several comments are in order. First, as we expect, two types of convergence rates are involved. In the direction of $\mathbf{q}_n$, we have a faster convergence rate given by

$$\left(\widehat{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0)\right) \mathbf{q}_n = O_p\left(\frac{1}{n\sqrt{h}}\right),$$

which is called *type I* super-consistency as in Phillips, Li, and Gao 2017. In the direction of $\mathbf{q}_n^{\perp}$, (11) gives

$$\left(\widehat{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0)\right)\mathbf{q}_n^{\perp} = O_p\left(\frac{1}{nh}\right),$$

which is slower than the $n\sqrt{h}$ rate in the direction of $\mathbf{q}_n$. We also name it *type II* super-consistency.

Second, while in the direction of $\mathbf{q}_n$, $\int_{-1}^{1} K(r)d\mathbf{B}_0^*\left(\frac{r+1}{2}\right)$ is distributed as $\mathcal{N}\left(0, \frac{1}{2}v_0\mathbf{\Omega}_{00}\right)$ with $v_0 = \int_{-1}^{1} K^2(u)du$. This is identical to the unit root case considered in Phillips, Li, and Gao 2017. In the direction of $\mathbf{q}_n^{\perp}$, $\int_{-1}^{1} K(r)\mathbf{K}_C^*\left(\frac{r+1}{2}\right) d\mathbf{B}_0^*\left(\frac{r+1}{2}\right)$ has a nonstandard limiting distribution, which is also different from the unit root case.

Third, in analogy with standard kernel estimation under stationarity, one might expect that a condition such as $nh^{5/2} \to c$ for some finite constant c would be sufficient. In the present setting, as detailed in the proof of Theorem 1, $\mathbb{B}_n(z_0)$ inherits two distinct stochastic rates associated with the rotated directions $\mathbf{q}_n$ and $\mathbf{q}_n^{\perp}$, so that to ensure that this random bias does not dominate the leading stochastic term in the asymptotic distribution, we have to impose the condition $nh^{3/2} \to c$.

Finally, it is worth noting that, unlike the stationary case, the bias process $\mathbb{B}_n(z_0)$ is random and path dependent. Under the bandwidth condition $nh^{3/2} \to c \in (0, \infty)$, this term contributes a nondegenerate component to the limit distribution. If instead the bandwidth is chosen so that $nh^{3/2} \to 0$, as in Phillips, Li, and Gao 2017 (corresponding to $\gamma_1 = 1$ in their notation), then the estimator operates in an undersmoothing regime and $\mathbb{B}_n(z_0) = o_p(1)$. In that case, the smoothing bias vanishes asymptotically.

4 — Asymptotic theory for mildly integrated system

In this section, we consider the case when regressors $\mathbf{X}_t$ are mildly integrated. In view of Lemma 1, it is natural to ask whether the weighted signal matrix $\mathbf{\Delta}_n(z_0)$ admits a nondegenerate probability limit. We show below that this difficulty does not arise under mild integration, provided that certain condition is satisfied.

Lemma 2. Under Assumptions 1-3, we have, for any fixed $z_0 \in (0, 1)$

$$\frac{1}{n^{1+\alpha}h}\mathbf{\Delta}_n(z_0) = \mathbf{V}_{xx} + O_p\left(\frac{1}{n^{1-\alpha}h}\right), \tag{12}$$

where $\mathbf{V}_{xx} := \int_0^\infty e^{r\mathbf{C}}\mathbf{\Omega}_{xx}e^{r\mathbf{C}}dr$.

When the condition $n^{1-\alpha}h \to \infty$ is satisfied, $\frac{1}{n^{1+\alpha}h}\mathbf{\Delta}_n(z_0)$ remains well behaved asymptotically in the sense that

$$\frac{1}{n^{1+\alpha}h}\mathbf{\Delta}_n(z_0) \xrightarrow{p} \mathbf{V}_{xx}.$$

This local information condition $n^{1-\alpha}h \to \infty$ reflects the reduction in the effective local sample size induced by regressor persistence. When $\alpha \to 0$, this condition reduces to the standard requirement $nh \to \infty$ under stationarity. When $\alpha \to 1$, corresponding to the case considered in the previous section, the condition $n^{1-\alpha}h \to \infty$ no longer holds, since $h \to 0$.

We now have our theorem regarding to the consistency of our estimator (5).

Theorem 2. If Assumptions 1-3 are satisfied. Then, for any fixed $z_0 \in (0,1)$ and any $\alpha \in (0,1)$ such that the condition $n^{1-\alpha}h \to \infty$ holds. We have

$$\left\|\widehat{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0)\right\| = O_p\left(h^2 + \frac{1}{n^{1-\alpha}} + \frac{1}{n^\alpha} + \frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\right). \tag{13}$$

Several comments are in order. First, Theorem 2 implies that $\widehat{\mathbf{A}}_n(z_0) \xrightarrow{p} \mathbf{A}(z_0)$ for any fixed $z_0 \in (0,1)$, as long as $n^{1-\alpha}h \to \infty$. Of course, when $\alpha \to 0$, it becomes $\left\|\widehat{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0)\right\| = O_p(1)$, so that $\widehat{\mathbf{A}}_n(z_0)$ is inconsistent due to the usual simultaneous equation bias problem. Second, if we write $\left(\mathrm{vec}\left(\widehat{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0)\right)\right)^\top \left(\mathrm{vec}\left(\widehat{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0)\right)\right) = r_{n,h}q_n$, where q_n is a scalar $O_p(1)$ random variable and $r_{n,h} = h^2 + \frac{1}{n^{1-\alpha}} + \frac{1}{n^\alpha} + \frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}$. The first-order condition of $r_{n,h}$ with respect to h gives the desired result that the MSE-optimal bandwidth is given by $h \asymp n^{-(1+\alpha)/5}$, so that $n^{\frac{1+\alpha}{2}}h^{5/2} \to c \in (0,\infty)$.[2] Finally, in contrast to the stationary case, the smoothing bias is not $O_p(h^2)$ alone. Instead, it is $O_p\left(\frac{1}{n^{1-\alpha}} + h^2\right)$. The extra $n^{-(1-\alpha)}$ term reflects the remainder induced by persistent regressors. Under the MSE-optimal bandwidth $h \asymp n^{-(1+\alpha)/5}$, the two bias components h^2 and $n^{-(1-\alpha)}$ do not admit a uniform ordering. A clear classical ordering is recovered only under the stronger condition $n^{1-\alpha}h^2 \to \infty$.

To establish a central limit theorem (C.L.T.) requires the knowledge of dominating term in the expansion of $\widehat{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0)$. However, based on Theorem 2, there is no clear ordering of these terms, unless further restrictions are imposed. Then, we have the following theorem.

2. $h \asymp n^{-(1+\alpha)/5}$ entails under smoothing relative to the stationary case, where it is $h \asymp n^{-1/5}$ (Cai 2007) for any fixed $z_0 \in (0,1)$.

Theorem 3. If Assumptions 1-3 are satisfied. Then, for any fixed $z_0 \in (0,1)$ and any $\alpha \in (0,1)$ such that the conditions $n^{\frac{1+\alpha}{2}} h^{5/2} \to c \in (0,\infty)$, $n^{1-\alpha}h^2 \to \infty$, and $n^{\frac{1}{2}-\alpha}\sqrt{h} \to 0$ hold. We have

$$n^{\alpha} \cdot \left(\widehat{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0)\right) \xrightarrow{p} (\mathbf{\Lambda}_{0x} + \mathbf{\Sigma}_{0x})\ \mathbf{V}_{xx}^{-1}, \tag{14}$$

The limiting distribution degenerates. As detailed in the proof of Lemma C.4, the condition $n^{\frac{1}{2}-\alpha}\sqrt{h} \to 0$ only eliminates one source of bias but terms related to simultaneous equation bias still dominates. To restore the standard asymptotic Normality, we propose a bias-corrected local level regression estimator $\widetilde{\mathbf{A}}_n(z_0)$ defined by

$$\widetilde{\mathbf{A}}_n(z_0) = \left[\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathbf{y}_t\mathbf{X}_t' - K_{1n}\cdot(\mathbf{\Lambda}_{0x} + \mathbf{\Sigma}_{0x})\right]\left[\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathbf{X}_t\mathbf{X}_t^{\top}\right]^{-1}, \tag{15}$$

where $K_{1n} = \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)$. In practice, one may estimate $\mathbf{\Lambda}_{0x} + \mathbf{\Sigma}_{0x}$ using HAC-type long-run covariance estimator or by exploiting the finite order VAR automated regression method by Phillips 2005.[3] The asymptotic distribution of the estimator $\widetilde{\mathbf{A}}_n(z_0)$ is given in the next theorem.

Theorem 4. If Assumptions 1-3 are satisfied. Then, for any fixed $z_0 \in (0,1)$ and any $\alpha \in (0,1)$ such that the conditions $n^{\frac{1+\alpha}{2}} h^{5/2} \to c \in (0,\infty)$, $n^{1-\alpha}h^2 \to \infty$, and $n^{\frac{1}{2}-\alpha}\sqrt{h} \to 0$ hold. We have

$$n^{\frac{1+\alpha}{2}}\sqrt{h}\cdot \mathrm{vec}\left(\widetilde{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0) - \frac{h^2}{2}\,\mu_2\,\nabla^2\mathbf{A}(z_0)\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}_{dk\times 1}, \mathbb{V}\right), \tag{16}$$

where $\mathbb{V} := v_0\mathbf{V}_{xx}^{-1} \otimes \mathbf{\Omega}_{00}$, $v_0 = \int_{-1}^{1} K^2(u)du$, and $\mathbf{V}_{xx}$ is defined as in Lemma 2.

From Theorem 4, we see that the convergence rate of the estimator is $n^{\frac{1+\alpha}{2}}\sqrt{h}$, which nests the rates in the existing literature. When $\alpha \to 0$, the rate reduces to $\sqrt{nh}$, coinciding with the stationary case considered in Cai (2007). When $\alpha \to 1$, the rate becomes $n\sqrt{h}$, which corresponds to the *type I super-consistency* as in Theorem 1 for the nearly integrated system.

Remark 1. It is worth noting that, both Theorem 3 and Theorem 4 are derived under a set of additional conditions. Given the MSE-optimal bandwidth choice $h = c \cdot n^{-(1+\alpha)/5}$, the con-

3. The feasible version of the estimator ($\mathbf{\Lambda}_{0x}+\mathbf{\Sigma}_{0x}$ replaced by $\widehat{\mathbf{\Lambda}}_{0x}+\widehat{\mathbf{\Sigma}}_{0x}$ in (15)) is analogous to the nonparametric version of the fully modified OLS estimator (FM-OLS) (Phillips and Hansen 1990) proposed in Phillips, Li, and Gao 2017 for the integrated case. The role of this adjustment is to remove the leading simultaneous-equation bias. No additional endogeneity correction is required beyond this adjustment.

dition $n^{1-\alpha}h^2 \to \infty$ requires $\alpha < 3/7$. On the other hand, eliminating an additional source of simultaneous-equation bias (Lemma C.4) requires $n^{\frac{1}{2}-\alpha}\sqrt{h} \to 0$, which implies that $\alpha > 4/11$. Therefore, the allowable range of α for Normal inference is $(4/11, 3/7)$. We shall investigate more on finite sample implications in Section 5.2.

5 — Monte Carlo study

5.1. Nearly integrated system

We use the Data Generating Process (D.G.P.) as in Phillips, Li, and Gao 2017:

$$y_t = \mathbf{A}_t^\top \mathbf{X}_t + u_{0t},$$

where $\mathbf{A}_t = (a_{1t}, a_{2t})^\top$ has the following functional forms

- M1: $a_{1t} = 1 + \frac{t}{n}$, $a_{2t} = e^{-\frac{t}{n}}$;
- M2: $a_{1t} = \cos(2\pi t/n)$, $a_{2t} = \sin(2\pi t/n)$.

The regressor $\mathbf{X}_t = (x_{1t}, x_{2t})^\top$ is generated according to $x_{i,t} = \left(1 - \frac{1}{n}\right) x_{i,t-1} + u_{i,xt}$ for $i = 1, 2$. We generate $\mathbf{u}_t = (u_{0t}, u_{1,xt}, u_{2,xt})^\top$ from a VAR(1) process

$$\mathbf{u}_t = \mathbf{\Pi}\mathbf{u}_{t-1} + \boldsymbol{\varepsilon}_t,$$

where $\mathbf{\Pi} = \text{diag}\left(\rho, \rho_1, \rho_2\right)$ and the innovation vector $\boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(\mathbf{0}_{3\times 1}, (1-\lambda)\mathbf{I}_3 + \mathbf{J}_3\right)$. We consider seven cases:

- Case 1: $\rho = \rho_1 = \rho_2 = 0$, $\lambda = 0$;
- Case 2: $\rho = \rho_1 = \rho_2 = 0.5$, $\lambda = 0.5$;
- Case 3: $\rho = \rho_1 = 0.5$, $\rho_2 = -0.5$, $\lambda = 0.5$;
- Case 4: $\rho = \rho_2 = 0.5$, $\rho_1 = -0.5$, $\lambda = 0.5$;
- Case 5: $\rho = \rho_1 = \rho_2 = 0.5$, $\lambda = -0.5$;
- Case 6: $\rho = \rho_1 = 0.5$, $\rho_2 = -0.5$, $\lambda = -0.5$;
- Case 7: $\rho = \rho_2 = 0.5$, $\rho_1 = -0.5$, $\lambda = -0.5$.

To compute the estimator (5), we employ the uniform kernel, $K(u) = \frac{1}{2}\mathbf{1}(|u| \leq 1)$. The bandwidth parameter h is set to $h = c \cdot n^{-(2/3}$ with c ranging from 0.01 to 1.5 with step size

0.02. Then, c is selected by leave-one-out cross-validation method. We consider two sample sizes: $n = 500, 1000$. The Monte Carlo analysis is based on $R = 10,000$ replications.

As in Section 3, we use the following notations: $z_n = \lfloor (z_0 - h)n \rfloor$, $\mathbf{X}_{z_n} = (x_{1,z_n}, x_{2,z_n})^\top$, $\mathbf{k}_n(z_0) = \frac{1}{\sqrt{n}}\mathbf{X}_{z_n}$, $\mathbf{q}_n(z_0) = \frac{\mathbf{k}_n}{\|\mathbf{k}_n\|} = \left(\frac{x_{1,z_n}}{\|\mathbf{X}_{z_n}\|}, \frac{x_{2,z_n}}{\|\mathbf{X}_{z_n}\|}\right)^\top \equiv (q_{1n}(z_0), q_{2n}(z_0))^\top$, and $\mathbf{q}_n^\perp(z_0) = (p_{1n}(z_0), p_{2n}(z_0))^\top$. Given that $\mathbf{Q}_n(z_0) = \left[\mathbf{q}_n(z_0), \mathbf{q}_n^\perp(z_0)\right]$ has to be orthonormal, we set $p_{1n}(z_0) = q_{2n}(z_0)$ and $p_{2n}(z_0) = -q_{1n}(z_0)$. The performance of the estimators is evaluated based on the transformed quantities

$$\tilde{a}_{1n}(z_0) = q_{1n}(z_0)\left[\hat{a}_{1n}(z_0) - a_1(z_0)\right] + q_{2n}(z_0)\left[\hat{a}_{2n}(z_0) - a_2(z_0)\right],$$
$$\tilde{a}_{2n}(z_0) = p_{1n}(z_0)\left[\hat{a}_{1n}(z_0) - a_1(z_0)\right] + p_{2n}(z_0)\left[\hat{a}_{2n}(z_0) - a_2(z_0)\right],$$

where $\hat{\mathbf{A}}_t = (\hat{a}_{1n}(z_0), \hat{a}_{2n}(z_0))^\top$. We report $\frac{1}{R}\sum_{j=1}^{R} \tilde{a}_{in,j}(z_0)$ for $i = 1, 2$ at the middle point $z_0 = 0.5$, where $\tilde{a}_{in,j}(z_0)$ is the value of $\tilde{a}_{in}(z_0)$ at the jth iteration.

The simulation results are reported in Table 1. In most of the cases, $|\tilde{a}_{1n}(z_0)|$ is consistently smaller than $|\tilde{a}_{2n}(z_0)|$, which is broadly consistent with the asymptotic theory developed in Section 3 that $|\tilde{a}_{1n}(z_0)|$ converges to zero at a faster rate than $|\tilde{a}_{2n}(z_0)|$.

Table 1—Small sample properties of the estimator (5) for near integrated regressors: absolute average of $\tilde{a}_{in}(z_0)$ at $z_0 = 0.5$

		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
M1				$n = 500$				
	$\lvert\tilde{a}_{1n}(z_0)\rvert$	0.0009	0.0013	0.0023	0.0021	0.0009	0.0030	0.0018
	$\lvert\tilde{a}_{2n}(z_0)\rvert$	0.0059	0.0064	0.0021	0.0055	0.0059	0.0058	0.0076
				$n = 1000$				
	$\lvert\tilde{a}_{1n}(z_0)\rvert$	0.0001	0.0004	0.0013	0.0008	0.0004	0.0004	0.0001
	$\lvert\tilde{a}_{2n}(z_0)\rvert$	0.0015	0.0056	0.0030	0.0067	0.0039	0.0038	0.0023
M2				$n = 500$				
	$\lvert\tilde{a}_{1n}(z_0)\rvert$	0.0010	0.0019	0.0013	0.0029	0.0014	0.0032	0.0032
	$\lvert\tilde{a}_{2n}(z_0)\rvert$	0.0016	0.0035	0.0024	0.0099	0.0086	0.0035	0.0005
				$n = 1000$				
	$\lvert\tilde{a}_{1n}(z_0)\rvert$	0.0000	0.0002	0.0015	0.0009	0.0006	0.0001	0.0008
	$\lvert\tilde{a}_{2n}(z_0)\rvert$	0.0040	0.0013	0.0042	0.0029	0.0047	0.0043	0.0047

5.2. Mildly integrated system

We generate data using the model:

$$y_t = A_t X_t + u_{0t},$$
$$X_t = R_n X_{t-1} + u_{xt}, \quad t = 1, 2, \cdots, n,$$

where $R_n = 1 - 1/n^\alpha, \alpha \in (0,1)$, $A_t = A(t/n) = 1 + t/n$. As in Phillips and Hansen 1990, $\mathbf{u}_t = (u_{0t}, u_{xt})'$ is generated according to a bivariate MA(1) process

$$\mathbf{u}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\Theta}\boldsymbol{\varepsilon}_{t-1}, \quad \boldsymbol{\varepsilon}_t \overset{iid}{\sim} \mathcal{N}(0, \Sigma_\varepsilon)$$

with

$$\Theta = \begin{pmatrix} 0.3 & -0.4 \\ \theta_{21} & 0.6 \end{pmatrix}, \qquad \Sigma_\varepsilon = \begin{pmatrix} 1 & \sigma_{21} \\ \sigma_{21} & 1 \end{pmatrix},$$

and θ_{21}, σ_{21} allowed to vary. We evaluate the estimators over the combinations $\theta_{21} \in 0.8, 0.4, 0.0$ and $\sigma_{21} \in -0.8, -0.4, 0.4, 0.8$.

To compute the estimators (5) and (15), we employ the uniform kernel, $K(u) = \frac{1}{2}\mathbf{1}(|u| \leq 1)$. The bandwidth parameter h is set to $h = c \cdot n^{-(1+\alpha)/5}$ with c ranging from 0.01 to 1.5 with step size 0.02. Then, c is selected by leave-one-out cross-validation method. We consider two sample sizes: $n = 500, 1000$. The Monte Carlo analysis is based on 10,000 replications.

The performance of the estimators is evaluated by the mean squared error (MSE). For the estimator (15), we also report the 95% coverage probability (CP), which is the estimated probability that the true $A(z_0)$ lies in the interval

$$\left(\hat{A}(z_0) - 1.96\,\text{s.d.}(\hat{A}(z_0)), \hat{A}(z_0) + 1.96\,\text{s.d.}(\hat{A}(z_0))\right),$$

where s.d.$(\hat{A}(z_0))$ is the asymptotic variance of the estimator obtained from Theorem 1. For both MSE and CP, we report the results at the middle point $z_0 = 0.5$.

Tables 2-3 report average MSE for both estimators (5) and (15) for $\alpha \in \{0.1, 0.2, 0.3\}$ and $\alpha \in \{0.5, 0.7, 0.9\}$, respectively. Let us first comment on the results for $\alpha \in \{0.1, 0.2, 0.3\}$. First, MSE decreases with the sample size for both estimators. Second, MSE reduction from (15) are substantial when σ_{21} is negative. This is not surprising. Under our data generating process (D.G.P.), straightforward calculation shows that $\Lambda_{0x} = 0.3\sigma_{21} - 0.4$, $\Sigma_{0x} = -0.24 + 1.18\sigma_{21} + 0.3\theta_{21} - 0.4\sigma_{21}\theta_{21}$. By Lemma C.4, under the condition $n^{\frac{1}{2}-\alpha}\sqrt{h} \to 0$, the amount of remaining simultaneous equation bias is given by $\Lambda_{0x} + \Sigma_{0x}$, which is substantially larger when σ_{21} is negative.

Turning to the case $\alpha \in \{0.5, 0.7, 0.9\}$, the performances of the two estimators are generally indistinguishable. By Theorem 2 and the construction of (15), the bias-correction term removes a component of order $O(n^{-\alpha})$. As α increases, this term vanishes at a faster rate, so that the two estimators exhibit nearly identical finite-sample behavior.

Figures 1–2 report the 95% coverage probability (CP) for the case $\sigma_{21} = -0.8$ over the grid $\alpha \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$. The CP displays a pronounced hump-shaped pattern, increasing for small values of α before declining as α increases. While consistency (Theorem 2) only requires minimal additional restriction $n^{1-\alpha}h \to \infty$, asymptotic Normality does require more restrictions (Theorem 4). As explained in Remark 1, the admissible range of α for Normal inference is $(4/11, 3/7)$. As $4/11 \approx 0.364$ and $3/7 \approx 0.429$, this theoretical threshold is broadly consistent with the coverage plots, where CP peaks at $\alpha = 0.3$ and gradually declines afterwards.

6 — Conclusion

This paper establishes the limit theory for Nadaraya–Watson type local level estimation of time-varying cointegrating coefficients in systems with nearly or mildly integrated regressors. In the nearly integrated case, we show that, conventional kernel methods encounter a degeneracy problem due to asymptotic singularity of the weighted signal matrix. We employ a rotational decomposition to develop the limit theory. In the mildly integrated case, We show that, the signal matrix remains asymptotically well behaved, and the estimator is consistent under a strengthened local information condition that reflects the reduced effective sample size induced by regressor persistence. At the same time, the resulting asymptotic behavior differs markedly from the stationary benchmark, with convergence rates and bias properties that depend explicitly on the degree of persistence. A bias-corrected local level estimator is proposed to restore standard

Table 2 — Small sample properties of the estimators (5) and (15) for mildly integrated regressor: MSE at $z_0 = 0.5$ for $\alpha \in \{0.1, 0.2, 0.3\}$

	$\alpha = 0.1$			$\alpha = 0.2$			$\alpha = 0.3$		
	$\theta_{21} = 0.8$	0.4	0.0	$\theta_{21} = 0.8$	0.4	0.0	$\theta_{21} = 0.8$	0.4	0.0
$n = 500$									
$\sigma_{21} = -0.8$									
$\widehat{A}_n(z_0)$	0.435	0.502	0.378	0.277	0.230	0.142	0.145	0.092	0.049
$\widetilde{A}_n(z_0)$	0.112	0.140	0.070	0.084	0.075	0.019	0.066	0.033	0.010
$\sigma_{21} = -0.4$									
$\widehat{A}_n(z_0)$	0.076	0.140	0.164	0.052	0.071	0.068	0.031	0.034	0.026
$\widetilde{A}_n(z_0)$	0.034	0.053	0.047	0.021	0.027	0.018	0.015	0.016	0.009
$\sigma_{21} = 0.4$									
$\widehat{A}_n(z_0)$	0.005	0.006	0.008	0.002	0.003	0.005	0.002	0.002	0.003
$\widetilde{A}_n(z_0)$	0.004	0.006	0.010	0.002	0.003	0.005	0.002	0.002	0.003
$\sigma_{21} = 0.8$									
$\widehat{A}_n(z_0)$	0.016	0.032	0.069	0.004	0.009	0.019	0.002	0.003	0.006
$\widetilde{A}_n(z_0)$	0.002	0.002	0.004	0.001	0.001	0.002	0.001	0.001	0.002
$n = 1000$									
$\sigma_{21} = -0.8$									
$\widehat{A}_n(z_0)$	0.413	0.455	0.337	0.219	0.175	0.108	0.092	0.058	0.032
$\widetilde{A}_n(z_0)$	0.074	0.078	0.038	0.037	0.025	0.008	0.022	0.011	0.003
$\sigma_{21} = -0.4$									
$\widehat{A}_n(z_0)$	0.065	0.122	0.144	0.036	0.051	0.051	0.017	0.019	0.016
$\widetilde{A}_n(z_0)$	0.022	0.029	0.025	0.010	0.011	0.007	0.006	0.006	0.004
$\sigma_{21} = 0.4$									
$\widehat{A}_n(z_0)$	0.003	0.004	0.004	0.001	0.002	0.002	0.001	0.001	0.002
$\widetilde{A}_n(z_0)$	0.002	0.004	0.006	0.001	0.002	0.002	0.001	0.001	0.002
$\sigma_{21} = 0.8$									
$\widehat{A}_n(z_0)$	0.013	0.027	0.059	0.003	0.006	0.014	0.001	0.002	0.004
$\widetilde{A}_n(z_0)$	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001

Table 3—Small sample properties of the estimators (5) and (15) for mildly integrated regressor: MSE at $z_0 = 0.5$ for $\alpha \in \{0.1, 0.2, 0.3\}$

	$\alpha = 0.5$			$\alpha = 0.7$			$\alpha = 0.9$		
	$\theta_{21} = 0.8$	0.4	0.0	$\theta_{21} = 0.8$	0.4	0.0	$\theta_{21} = 0.8$	0.4	0.0
				$n = 500$					
$\sigma_{21} = -0.8$									
$\widehat{A}_n(z_0)$	0.043	0.020	0.008	0.019	0.009	0.004	0.011	0.005	0.002
$\widetilde{A}_n(z_0)$	0.032	0.022	0.009	0.017	0.017	0.011	0.011	0.010	0.008
$\sigma_{21} = -0.4$									
$\widehat{A}_n(z_0)$	0.012	0.010	0.006	0.006	0.005	0.004	0.004	0.003	0.002
$\widetilde{A}_n(z_0)$	0.008	0.009	0.007	0.004	0.005	0.007	0.003	0.004	0.004
$\sigma_{21} = 0.4$									
$\widehat{A}_n(z_0)$	0.002	0.002	0.002	0.001	0.001	0.002	0.001	0.001	0.001
$\widetilde{A}_n(z_0)$	0.002	0.002	0.002	0.001	0.001	0.002	0.001	0.001	0.001
$\sigma_{21} = 0.8$									
$\widehat{A}_n(z_0)$	0.001	0.001	0.002	0.001	0.001	0.001	0.000	0.001	0.001
$\widetilde{A}_n(z_0)$	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001
				$n = 1000$					
$\sigma_{21} = -0.8$									
$\widehat{A}_n(z_0)$	0.019	0.009	0.004	0.007	0.003	0.001	0.004	0.002	0.001
$\widetilde{A}_n(z_0)$	0.009	0.005	0.002	0.005	0.004	0.003	0.003	0.003	0.002
$\sigma_{21} = -0.4$									
$\widehat{A}_n(z_0)$	0.005	0.004	0.003	0.002	0.002	0.001	0.001	0.001	0.001
$\widetilde{A}_n(z_0)$	0.003	0.003	0.002	0.002	0.002	0.002	0.001	0.001	0.001
$\sigma_{21} = 0.4$									
$\widehat{A}_n(z_0)$	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000	0.000
$\widetilde{A}_n(z_0)$	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000	0.000
$\sigma_{21} = 0.8$									
$\widehat{A}_n(z_0)$	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000
$\widetilde{A}_n(z_0)$	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000	0.000

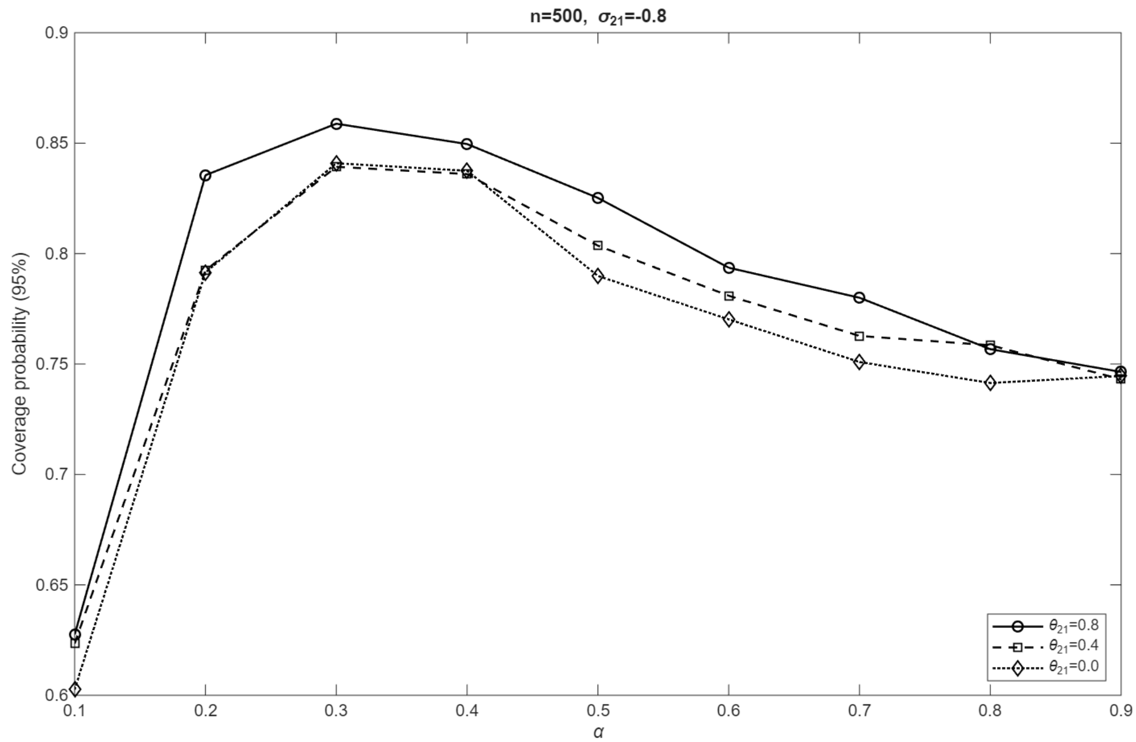

Figure 1 — 95% coverage probabilities: $n = 500$, $\sigma_{21} = -0.8$.

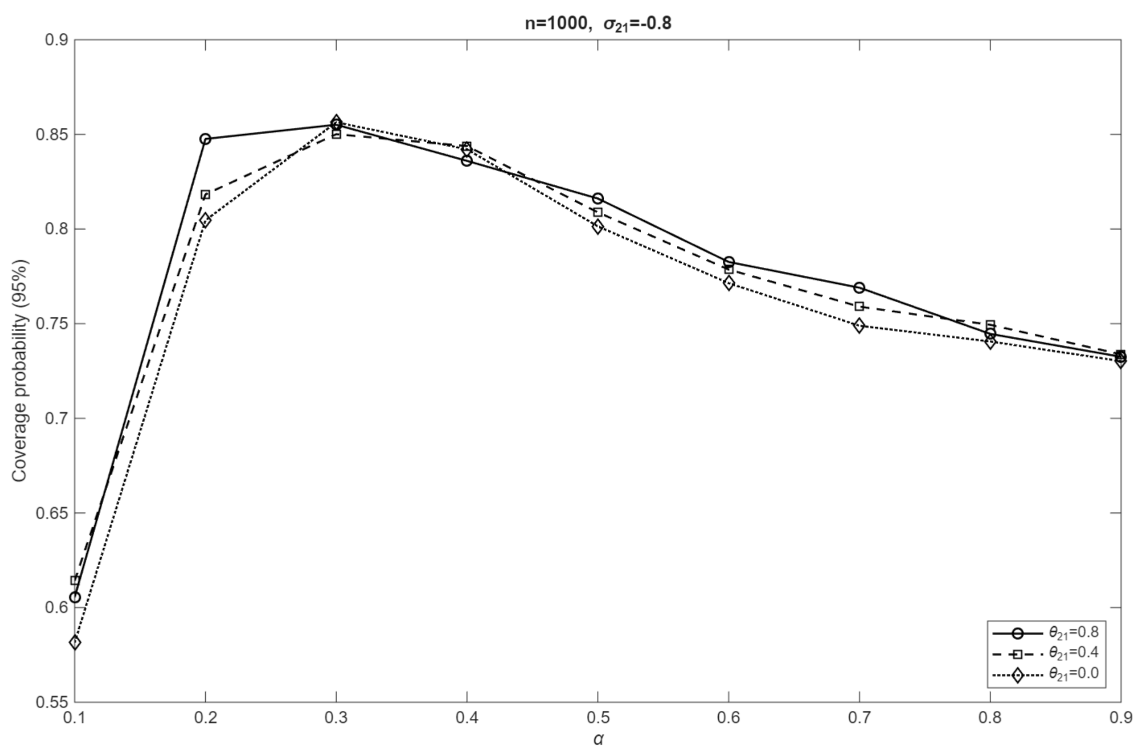

Figure 2 — 95% coverage probabilities: $n = 1000$, $\sigma_{21} = -0.8$.

asymptotic normality under additional rate conditions.

Monte Carlo simulations illustrate the theoretical results and show that a clear pattern for different directional convergence rates for the nearly integrated case. In the mildly integrated case, the bias correction is particularly effective when persistence is moderate, while its impact diminishes as the persistence parameter increases. Overall, the analysis clarifies how kernel smoothing interacts with near and mild integration and highlights the finite-sample implications from the asymptotic theory.

References

Abadir, Karim M, Walter Distaso, Liudas Giraitis, and Hira L Koul. 2014. "Asymptotic normality for weighted sums of linear processes." *Econometric Theory* 30 (1): 252–284. (Cited on pages 26, 38, 41, 43, 47).

Bernstein, Dennis. 2018. *Scalar, vector, and matrix mathematics: theory, facts, and formulas-revised and expanded edition.* Princeton university press. (Cited on page 35).

Cai, Zongwu. 2007. "Trending time-varying coefficient time series models with serially correlated errors." *Journal of Econometrics* 136 (1): 163–188. (Cited on pages 10, 11).

Chan, Ngai H, and Ching-Zong Wei. 1987. "Asymptotic inference for nearly nonstationary AR (1) processes." *The Annals of Statistics,* 1050–1063. (Cited on page 2).

Kapetanios, George, Stephen Millard, Katerina Petrova, and Simon Price. 2020. "Time-varying cointegration with an application to the UK Great Ratios." *Economics Letters* 193:109213. (Cited on page 2).

Kostakis, Alexandros, Tassos Magdalinos, and Michalis P Stamatogiannis. 2015. "Robust econometric inference for stock return predictability." *The Review of Financial Studies* 28 (5): 1506–1553. (Cited on pages 2, 3).

Li, Degui, Peter CB Phillips, and Jiti Gao. 2020. "Kernel-based inference in time-varying coefficient cointegrating regression." *Journal of Econometrics* 215 (2): 607–632. (Cited on page 2).

Liu, Nan, Peter CB Phillips, and Yajie Zhang. 2025. "Robust Inference for Time Varying Predictability: A Sieve-IVX Approach," (cited on page 3).

Magdalinos, Tassos, and Peter CB Phillips. 2009. "Limit theory for cointegrated systems with moderately integrated and moderately explosive regressors." *Econometric Theory* 25 (2): 482–526. (Cited on pages 2, 26, 44, 45, 47, 48).

Park, Joon Y, and Sang B Hahn. 1999. "Cointegrating regressions with time varying coefficients." *Econometric Theory* 15 (5): 664–703. (Cited on page 2).

Phillips, Peter CB. 1987. "Towards a unified asymptotic theory for autoregression." *Biometrika* 74 (3): 535–547. (Cited on page 2).

———. 1988. "Regression theory for near-integrated time series." *Econometrica: Journal of the Econometric Society,* 1021–1043. (Cited on page 2).

———. 1991. "Optimal inference in cointegrated systems." *Econometrica: Journal of the Econometric Society,* 283–306. (Cited on page 4).

———. 2005. "HAC estimation by automated regression." *Econometric Theory* 21 (1): 116–142. (Cited on page 11).

Phillips, Peter CB, and Bruce E Hansen. 1990. "Statistical inference in instrumental variables regression with I (1) processes." *The review of economic studies* 57 (1): 99–125. (Cited on pages 11, 14).

Phillips, Peter CB, Degui Li, and Jiti Gao. 2017. "Estimating smooth structural change in cointegration models." *Journal of Econometrics* 196 (1): 180–195. (Cited on pages 2, 3, 7, 9, 11, 12, 32).

Phillips, Peter CB, and Tassos Magdalinos. 2007a. "Limit theory for moderate deviations from a unit root." *Journal of Econometrics* 136 (1): 115–130. (Cited on page 2).

———. 2007b. "Limit Theory for Moderate Deviations From a Unit Root Under Weak Dependence." *The Refinement of Econometric Estimation and Test Procedures,* 123–162. (Cited on page 2).

Phillips, Peter CB, and Victor Solo. 1992. "Asymptotics for linear processes." *The Annals of Statistics,* 971–1001. (Cited on pages 5, 44).

Shi, Shuping, and Peter CB Phillips. 2023. "Diagnosing housing fever with an econometric thermometer." *Journal of Economic Surveys* 37 (1): 159–186. (Cited on page 2).

Yan, Yayi, Jiti Gao, and Bin Peng. 2025. "ASYMPTOTICS FOR TIME-VARYING VECTOR MA (∞) PROCESSES." *Econometric Theory* 41 (3): 584–616. (Cited on pages 26, 38, 41, 43, 47).

A — Mathematical proofs

A.1. Proof of Lemma 1

Take a neighborhood $N_{nz_0}(h) = [\lfloor (z_0 - h)n \rfloor, \lfloor (z_0 + h)n \rfloor]$ of $\lfloor z_0 n \rfloor$ and let $z_n = \lfloor (z_0 - h)n \rfloor$. We have the following representation for $\mathbf{\Lambda}_n(z_0)$:

$$
\begin{aligned}
\mathbf{\Delta}_n(z_0) &= \sum_{t=1}^{n} K\left(\frac{t - nz_0}{nh}\right) \left(\mathbf{X}_t - \mathbf{X}_{z_n} + \mathbf{X}_{z_n}\right) \left(\mathbf{X}_t - \mathbf{X}_{z_n} + \mathbf{X}_{z_n}\right)^\top \\
&= \sum_{t=1}^{n} K\left(\frac{t - nz_0}{nh}\right) \mathbf{X}_{z_n}\mathbf{X}_{z_n}^\top + \sum_{t=1}^{n} K\left(\frac{t - nz_0}{nh}\right) \left(\mathbf{X}_t - \mathbf{X}_{z_n}\right) \mathbf{X}_{z_n}^\top \\
&\quad + \sum_{t=1}^{n} K\left(\frac{t - nz_0}{nh}\right) \mathbf{X}_{z_n} \left(\mathbf{X}_t - \mathbf{X}_{z_n}\right)^\top + \sum_{t=1}^{n} K\left(\frac{t - nz_0}{nh}\right) \left(\mathbf{X}_t - \mathbf{X}_{z_n}\right) \left(\mathbf{X}_t - \mathbf{X}_{z_n}\right)^\top \\
&:= \mathbf{\Delta}_{n1}(z_0) + \mathbf{\Delta}_{n2}(z_0) + \mathbf{\Delta}_{n3}(z_0) + \mathbf{\Delta}_{n4}(z_0).
\end{aligned}
$$

By Assumption 1 and a Riemann-sum approximation, we have

$$
\frac{1}{nh} \sum_{t=1}^{n} K\left(\frac{t - nz_0}{nh}\right) \longrightarrow \int_{-1}^{1} K(u)\, du = 1, \qquad 0 < z_0 < 1.
$$

Together with Lemma S.1(i) and Slutsky's theorem, this implies

$$\frac{1}{n^2h}\mathbf{\Delta}_{n1}(z_0) = \left(\frac{\mathbf{X}_{z_n}\mathbf{X}_{z_n}^\top}{n}\right)\left(\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\right) \xrightarrow{d} \mathbf{K}_C(z_0)\mathbf{K}_C^\top(z_0). \tag{A.1}$$

For $\mathbf{\Lambda}_{n2}(z_0)$, Lemma S.1 implies $\|\mathbf{X}_{z_n}\| = O_p(\sqrt{n})$ and, uniformly for $t \in \overline{N}_{nz_0}(h)$, $\|\mathbf{X}_t - \mathbf{X}_{z_n}\| = O_p(\sqrt{nh})$, where $\overline{N}_{nz_0}(h)$ is a set of integers in $N_{nz_0}(h)$. We have

$$\begin{aligned}\|\mathbf{\Lambda}_{n2}(z_0)\| &\le \|\mathbf{X}_{z_n}\| \sum_{t=\lfloor (z_0-h)n\rfloor+1}^{\lfloor (z_0+h)n\rfloor} K\left(\frac{t-nz_0}{nh}\right)\|\mathbf{X}_t - \mathbf{X}_{z_n}\| \\ &= O_p\left(\sqrt{n}\right)\times O(nh)\times O_p\left(\sqrt{nh}\right) = O_p\left(n^2h^{3/2}\right).\end{aligned} \tag{A.2}$$

Similarly, we can show that

$$\|\mathbf{\Lambda}_{n3}(z_0)\| = O_p\left(n^2h^{3/2}\right), \quad \|\mathbf{\Lambda}_{n4}(z_0)\| = O_p\left(n^2h^2\right), \tag{A.3}$$

In view of (A.1), (A.2), and (A.3), we deduce that

$$\frac{1}{n^2h}\mathbf{\Lambda}_n(z_0) \xrightarrow{d} \mathbf{K}_C(z_0)\mathbf{K}_C^\top(z_0).$$

A.2. Proof of Theorem 1

We have the following decomposition

$$\begin{aligned}&\left(\hat{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0)\right)\mathbf{Q}_n\mathbf{D}_n \\ =&\left[\left(\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)(\mathbf{A}_t - \mathbf{A}(z_0))\mathbf{X}_t\mathbf{X}_t^\top\right)\mathbf{Q}_n\mathbf{D}_n^+\right]\left[\mathbf{D}_n^+\mathbf{Q}_n^\top\left(\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{X}_t\mathbf{X}_t^\top\right)\mathbf{Q}_n\mathbf{D}_n^+\right]^+ \\ &+\left[\left(\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\mathbf{X}_t^\top\right)\mathbf{Q}_n\mathbf{D}_n^+\right]\left[\mathbf{D}_n^+\mathbf{Q}_n^\top\left(\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{X}_t\mathbf{X}_t^\top\right)\mathbf{Q}_n\mathbf{D}_n^+\right]^+ \\ :=& \mathbf{B}_n(z_0)\mathbf{\Delta}_n^+(z_0) + \mathbf{\Gamma}_n(z_0)\mathbf{\Delta}_n^+(z_0).\end{aligned} \tag{A.4}$$

We shall analyze $\mathbf{B}_n(z_0)$. For $\mathbf{\Delta}_n(z_0)$ and $\mathbf{\Gamma}_n(z_0)$, they are given in Lemma B.1 and Lemma

B.2, respectively. Notice that

$$
\begin{aligned}
\mathbf{B}_n(z_0) &= \left(\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \left(\mathbf{A}_t - \mathbf{A}(z_0)\right) \mathbf{X}_t \mathbf{X}_t^\top \right) \mathbf{Q}_n \mathbf{D}_n^+ \\
&= \begin{bmatrix} \frac{1}{n\sqrt{h}} \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \left(\mathbf{A}_t - \mathbf{A}(z_0)\right) \mathbf{X}_t \mathbf{X}_t^\top \mathbf{q}_n & \frac{1}{nh} \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \left(\mathbf{A}_t - \mathbf{A}(z_0)\right) \mathbf{X}_t \mathbf{X}_t^\top \mathbf{q}_n^\perp \end{bmatrix} \\
&:= \begin{bmatrix} \mathbf{B}_n(z_0, 1) & \mathbf{B}_n(z_0, 2) \end{bmatrix}.
\end{aligned}
$$

Let us first consider $\mathbf{B}_n(z_0, 1)$. By a second-order Taylor expansion of $A(\cdot)$ around $z_0 \in (0, 1)$, we have

$$
\begin{aligned}
\frac{1}{n\sqrt{h}} \mathbf{B}_n(z_0, 1) = \nabla \mathbf{A}(z_0) & \left(\frac{1}{n^2 h} \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \left(\frac{t}{n} - z_0\right) \mathbf{X}_t \mathbf{X}_t^\top \right) \mathbf{q}_n \\
& + \frac{1}{2} \nabla^2 \mathbf{A}(z_0) \left(\frac{1}{n^2 h} \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \left(\frac{t}{n} - z_0\right)^2 \mathbf{X}_t \mathbf{X}_t^\top \right) \mathbf{q}_n + \text{s.o.},
\end{aligned}
$$

where s.o. denotes higher-order negligible terms. Observe that

$$
\begin{aligned}
\frac{1}{nh} \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \left(\frac{t}{n} - z_0\right) &= h \int_{-1}^{1} uK(u)du + O(1/n) = O(1/n), \\
\frac{1}{nh} \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \left(\frac{t}{n} - z_0\right)^2 &= h^2 \int_{-1}^{1} u^2 K(u)du + o(h^2) = O(h^2),
\end{aligned}
$$

where $O(1/n)$ in the first line comes from the Riemann-sum approximation of an integral. Using similar arguments to establish Lemma 1 and Lemma B.1, we can show that

$$
\begin{aligned}
&\frac{1}{n\sqrt{h}}\mathbf{B}_n(z_0,1)\\
&=\nabla\mathbf{A}(z_0)\left(\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\left(\frac{t}{n}-z_0\right)\right)\left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right)\left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right)^{\top}\mathbf{q}_n\\
&\quad+\frac{1}{2}\nabla^2\mathbf{A}(z_0)\left(\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\left(\frac{t}{n}-z_0\right)^2\right)\left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right)\left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right)^{\top}\mathbf{q}_n\\
&\quad+\nabla\mathbf{A}(z_0)\cdot\frac{1}{n\sqrt{h}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\left(\frac{t}{n}-z_0\right)\left(\frac{\mathbf{X}_t-\mathbf{X}_{z_n}}{\sqrt{nh}}\right)\left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right)^{\top}\mathbf{q}_n\\
&\quad+\nabla\mathbf{A}(z_0)\cdot\frac{1}{n\sqrt{h}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\left(\frac{t}{n}-z_0\right)\left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right)\left(\frac{\mathbf{X}_t-\mathbf{X}_{z_n}}{\sqrt{nh}}\right)^{\top}\mathbf{q}_n\\
&\quad+\nabla\mathbf{A}(z_0)\cdot\frac{1}{n}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\left(\frac{t}{n}-z_0\right)\left(\frac{\mathbf{X}_t-\mathbf{X}_{z_n}}{\sqrt{nh}}\right)\left(\frac{\mathbf{X}_t-\mathbf{X}_{z_n}}{\sqrt{nh}}\right)^{\top}\mathbf{q}_n+\text{s.o.}\\
&=O_p(h^{3/2})+O_p(h^2)+O_p(h^{3/2})+O_p(h^{3/2})+O_p(h^2)=O_p(h^{3/2}),
\end{aligned}
$$

which implies that $\|\mathbf{B}_n(z_0,1)\|=O_p\left(nh^2\right)$. An entirely analogous argument shows that $\|\mathbf{B}_n(z_0,2)\|=O_p\left(nh^{3/2}\right)$. Then,

$$
\begin{aligned}
\left\|\left(\hat{\mathbf{A}}_n(z_0)-\mathbf{A}(z_0)\right)\mathbf{Q}_n\mathbf{D}_n-\mathbf{\Gamma}_n(z_0)\mathbf{\Delta}_n^{+}(z_0)\right\| &\le \|\mathbf{B}_n(z_0)\|\ \left\|\mathbf{\Delta}_n^{+}(z_0)\right\|\\
&\le \left(\|\mathbf{B}_n(z_0,1)\|+\|\mathbf{B}_n(z_0,2)\|\right)\ \left\|\mathbf{\Delta}_n^{+}(z_0)\right\|\\
&=O_p\left(nh^{3/2}\right).
\end{aligned}
$$

Since $\mathbf{\Delta}_n(z_0)$ and $\mathbf{\Gamma}_n(z_0)$ are both functions of the same localized partial-sum process of the innovation vector $\mathbf{u}_t$ together with the same random rotation based on X_{z_n}, their joint convergence follows directly by a straightforward application of the continuous mapping theorem to a vector collecting the relevant block components. Theorem 1 follows by combining Lemma B.1 and Lemma B.2.

A.3. Proof of Lemma 2

By (2), we have

$$\mathbf{X}_t\mathbf{X}_t^\top = \mathbf{R}_n\mathbf{X}_{t-1}\mathbf{X}_{t-1}^\top\mathbf{R}_n + \mathbf{R}_n\mathbf{X}_{t-1}\mathbf{u}_{xt}^\top + \mathbf{u}_{xt}\mathbf{X}_{t-1}^\top\mathbf{R}_n + \mathbf{u}_{xt}\mathbf{u}_{xt}^\top.$$

Recall the identities $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^\top \otimes \mathbf{A})\,\mathrm{vec}(\mathbf{B})$ and $\mathcal{K}_k\,\mathrm{vec}(\mathbf{A}) = \mathrm{vec}(\mathbf{A}^\top)$. Applying vec on both sides gives

$$\mathrm{vec}(\mathbf{X}_t\mathbf{X}_t^\top) = (\mathbf{R}_n \otimes \mathbf{R}_n)\,\mathrm{vec}(\mathbf{X}_{t-1}\mathbf{X}_{t-1}^\top) + (\mathbf{I}_{k^2} + \mathcal{K}_k)\,(\mathbf{I}_k \otimes \mathbf{R}_n)\,\mathrm{vec}(\mathbf{X}_{t-1}\mathbf{u}_{xt}^\top) + \mathrm{vec}(\mathbf{u}_{xt}\mathbf{u}_{xt}^\top).$$

For each $z_0 \in (0,1)$, by first multiplying the kernel weights $K\left(\frac{t-nz_0}{nh}\right)$ and then summing over $t = 1, 2, \cdots, n$, we obtain

$$\begin{aligned}
&\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}\left(\mathbf{X}_t\mathbf{X}_t^\top\right)\\
&= (\mathbf{R}_n \otimes \mathbf{R}_n)\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}(\mathbf{X}_{t-1}\mathbf{X}_{t-1}^\top)\\
&\quad + (\mathbf{I}_{k^2} + \mathcal{K}_k)\,(\mathbf{I}_k \otimes \mathbf{R}_n)\,\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}\left(\mathbf{X}_{t-1}\mathbf{u}_{xt}^\top\right) + \frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}\left(\mathbf{u}_{xt}\mathbf{u}_{xt}^\top\right).
\end{aligned}$$

By rearranging terms and summation by parts, we have

$$\begin{aligned}
&\left[\mathbf{I}_{k^2} - \mathbf{R}_n \otimes \mathbf{R}_n\right]\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}\left(\mathbf{X}_{t-1}\mathbf{X}_{t-1}^\top\right)\\
&= \frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\left\{\mathrm{vec}\left(\mathbf{X}_{t-1}\mathbf{X}_{t-1}^\top\right) - \mathrm{vec}\left(\mathbf{X}_t\mathbf{X}_t^\top\right)\right\}\\
&\qquad + (\mathbf{I}_{k^2} + \mathcal{K}_k)\,(\mathbf{I}_k \otimes \mathbf{R}_n)\,\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}\left(\mathbf{X}_{t-1}\mathbf{u}_{xt}^\top\right) + \frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}\left(\mathbf{u}_{xt}\mathbf{u}_{xt}^\top\right)\\
&= (\mathbf{I}_{k^2} + \mathcal{K}_k)\,(\mathbf{I}_k \otimes \mathbf{R}_n)\,\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}\left(\mathbf{X}_{t-1}\mathbf{u}_{xt}^\top\right)\\
&\qquad\quad + \frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}\left(\mathbf{u}_{xt}\mathbf{u}_{xt}^\top\right) + O_p\left(\frac{1}{n^{1-\alpha}h}\right).
\end{aligned}$$

The second equality follows from the fact that summation by parts gives

$$
\begin{aligned}
&\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\left\{\mathrm{vec}\left(\mathbf{X}_{t-1}\mathbf{X}_{t-1}^{\top}\right)-\mathrm{vec}\left(\mathbf{X}_{t}\mathbf{X}_{t}^{\top}\right)\right\}\\
=&\frac{1}{nh}\left[-K\left(\frac{1-z_0}{h}\right)\mathrm{vec}\left(\mathbf{X}_{n}\mathbf{X}_{n}^{\top}\right)+K\left(\frac{1-nz_0}{nh}\right)\mathrm{vec}\left(\mathbf{X}_{0}\mathbf{X}_{0}^{\top}\right)\right]\\
&+\frac{1}{nh}\sum_{t=1}^{n-1}\left[K\left(\frac{t+1-nz_0}{nh}\right)-K\left(\frac{t-nz_0}{nh}\right)\right]\mathrm{vec}\left(\mathbf{X}_{t}\mathbf{X}_{t}^{\top}\right).
\end{aligned}
$$

Assumption 1(i) implies that $K\left(\frac{1-z_0}{h}\right)=K\left(\frac{1-nz_0}{nh}\right)=0$ for sufficiently large n. In addition,

$$
\begin{aligned}
&\left\|\frac{1}{nh}\sum_{t=1}^{n-1}\left[K\left(\frac{t+1-nz_0}{nh}\right)-K\left(\frac{t-nz_0}{nh}\right)\right]\mathrm{vec}\left(\mathbf{X}_{t}\mathbf{X}_{t}^{\top}\right)\right\|\\
\leq&\sup_{1\leq t\leq n-1}\left\|\mathbf{X}_{t}\mathbf{X}_{t}^{\top}\right\|\frac{1}{nh}\sum_{t=1}^{n-1}\left|K\left(\frac{t+1-nz_0}{nh}\right)-K\left(\frac{t-nz_0}{nh}\right)\right|=O_p\left(\frac{1}{n^{1-\alpha}h}\right),
\end{aligned}
$$

which follows by Lemma 3.1(i) in Magdalinos and Phillips 2009 and the Lipschitz continuity of $K(\cdot)$. Therefore, $\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\left\{\mathrm{vec}\left(\mathbf{X}_{t-1}\mathbf{X}_{t-1}^{\top}\right)-\mathrm{vec}\left(\mathbf{X}_{t}\mathbf{X}_{t}^{\top}\right)\right\}=o_p(1)$.

By Lemma C.5(ii), we have $\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}\left(\mathbf{X}_{t-1}\mathbf{u}_{xt}^{\top}\right)=\mathbf{\Lambda}_{xx}+O_p\left(\frac{1}{n^{1-\alpha}h}\right)$. In addition, by the weak law of large numbers (WLLN) for weighted sums of linear processes (Abadir et al. 2014; Yan, Gao, and Peng 2025), we have $\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}\left(\mathbf{u}_{xt}\mathbf{u}_{xt}^{\top}\right)\xrightarrow{p}\mathbf{\Sigma}_{xx}$. Substituting all these above leads to

$$
\begin{aligned}
\left[\mathbf{I}_{k^2}-\mathbf{R}_n\otimes\mathbf{R}_n\right]\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}\big(\mathbf{X}_{t-1}\mathbf{X}_{t-1}^{\top}\big)&=\left(\mathbf{I}_{k^2}+\mathcal{K}_k\right)\mathrm{vec}(\mathbf{\Lambda}_{xx})+\mathrm{vec}(\mathbf{\Sigma}_{xx})+O_p\left(\frac{1}{n^{1-\alpha}h}\right)\\
&=\mathrm{vec}\big(\mathbf{\Lambda}_{xx}+\mathbf{\Lambda}_{xx}^{\top}+\mathbf{\Sigma}_{xx}\big)+O_p\left(\frac{1}{n^{1-\alpha}h}\right)\\
&=\mathrm{vec}(\mathbf{\Omega}_{xx})+O_p\left(\frac{1}{n^{1-\alpha}h}\right).
\end{aligned}
$$

In addition, since

$$
\begin{aligned}
\mathbf{I}_{k^2}-(\mathbf{R}_n\otimes\mathbf{R}_n)&=\mathbf{I}_{k^2}-\left(\mathbf{I}_k+\frac{\mathbf{C}}{n^{\alpha}}\right)\otimes\left(\mathbf{I}_k+\frac{\mathbf{C}}{n^{\alpha}}\right)\\
&=-\frac{1}{n^{\alpha}}\left(\mathbf{C}\otimes\mathbf{I}_k+\mathbf{I}_k\otimes\mathbf{C}+\frac{\mathbf{C}\otimes\mathbf{C}}{n^{\alpha}}\right),
\end{aligned}
$$

we conclude that

$$\frac{1}{n^{1+\alpha}h}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathrm{vec}\left(\mathbf{X}_{t-1}\mathbf{X}_{t-1}^{\top}\right)=-\left(\mathbf{C}\otimes\mathbf{I}_k+\mathbf{I}_k\otimes\mathbf{C}\right)^{-1}\mathrm{vec}\left(\mathbf{\Omega}_{xx}\right)+O_p\left(\frac{1}{n^{1-\alpha}h}\right).$$

Let $\mathbf{V}_{xx}:=\int_0^\infty e^{r\mathbf{C}}\mathbf{\Omega}_{xx}e^{r\mathbf{C}}dr$. Observe that

$$\begin{aligned}\mathbf{C}\mathbf{V}_{xx}+\mathbf{V}_{xx}\mathbf{C}&=\int_0^\infty\left(\mathbf{C}e^{r\mathbf{C}}\mathbf{\Omega}_{xx}e^{r\mathbf{C}}+e^{r\mathbf{C}}\mathbf{\Omega}_{xx}e^{r\mathbf{C}}\mathbf{C}\right)dr\\&=\int_0^\infty\left(\frac{d}{dr}e^{r\mathbf{C}}\mathbf{\Omega}_{xx}e^{r\mathbf{C}}\right)dr=-\mathbf{\Omega}_{xx}.\end{aligned}$$

Applying vec on both sides gives

$$\left(\mathbf{C}\otimes\mathbf{I}_k+\mathbf{I}_k\otimes\mathbf{C}\right)\mathrm{vec}\left(\mathbf{V}_{xx}\right)=-\,\mathrm{vec}(\mathbf{\Omega}_{xx}).$$

So, we can write

$$\frac{1}{n^{1+\alpha}h}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{X}_{t-1}\mathbf{X}_{t-1}^{\top}=\mathbf{V}_{xx}+O_p\left(\frac{1}{n^{1-\alpha}h}\right).$$

This completes the proof.

A.4. Proof of Theorem 2

Observe that, for any fixed $z_0\in(0,1)$, we have

$$\begin{aligned}\widehat{\mathbf{A}}_n(z_0)-\mathbf{A}(z_0)=&\left[\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\left(\mathbf{A}_t-\mathbf{A}(z_0)\right)\mathbf{X}_t\mathbf{X}_t^{\top}\right]\left[\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{X}_t\mathbf{X}_t^{\top}\right]^{-1}\\&+\left[\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\mathbf{X}_t^{\top}\right]\left[\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{X}_t\mathbf{X}_t^{\top}\right]^{-1}\\:=&\,\mathbf{B}_n(z_0)\mathbf{\Delta}_n^{-1}(z_0)+\mathbf{\Gamma}_n(z_0)\mathbf{\Delta}_n^{-1}(z_0).\end{aligned}\tag{A.5}$$

By (A.5) and triangular inequality, we have

$$\left\|\widehat{\mathbf{A}}_n(z_0)-\mathbf{A}(z_0)\right\|\leq\left(\left\|\frac{1}{n^{1+\alpha}h}\mathbf{B}_n(z_0)\right\|+\left\|\frac{1}{n^{1+\alpha}h}\mathbf{\Gamma}_n(z_0)\right\|\right)\left\|\left(\frac{1}{n^{1+\alpha}h}\mathbf{\Delta}_n(z_0)\right)^{-1}\right\|_{sp}.$$

First, it follows immediately from Lemma C.4 that

$$\left\| \frac{1}{n^{1+\alpha}h} \mathbf{\Gamma}_n(z_0) \right\| = O\left(\frac{1}{n^{\alpha}}\right) + O_p\left(\frac{1}{n^{\frac{3}{2}\alpha}}\right) + O_p\left(\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\right).$$

Second, note that

$$\begin{aligned}\left\| \left(\frac{1}{n^{1+\alpha}h} \mathbf{\Delta}_n(z_0)\right)^{-1} \right\|_{sp} &\leq \left\|(\mathbf{V}_{xx})^{-1}\right\|_{sp} \left\| \left(\mathbf{I}_k + (\mathbf{V}_{xx})^{-1}\left(\frac{1}{n^{1+\alpha}h}\mathbf{\Delta}_n(z_0) - \mathbf{V}_{xx}\right)\right)^{-1} \right\|_{sp} \\ &\leq \left\|(\mathbf{V}_{xx})^{-1}\right\|_{sp} \left(1 - \left\| \frac{1}{n^{1+\alpha}h}\mathbf{\Delta}_n(z_0) - \mathbf{V}_{xx} \right\|_{sp}\right)^{-1} = O_p(1),\end{aligned}$$

since $\lambda_{\min}(\mathbf{V}_{xx}) > 0$ and $\frac{1}{n^{1+\alpha}h}\mathbf{\Delta}_n(z_0) \xrightarrow{p} \mathbf{V}_{xx}$ as $n^{1-\alpha}h \to \infty$.

The proofs are completed if we could show that

$$\left\| \frac{1}{n^{1+\alpha}h} \mathbf{B}_n(z_0) \right\| = O_p\left(\frac{1}{n^{1-\alpha}} + h^2\right). \tag{A.6}$$

To see this, notice that, by a second-order Taylor expansion of $A(\cdot)$ around $z_0 \in (0,1)$, we have

$$\begin{aligned}\mathbf{B}_n(z_0) &= \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \nabla \mathbf{A}(z_0) \left(\frac{t}{n} - z_0\right) \mathbf{X}_{t-1}\mathbf{X}_{t-1}^{\top} \\ &\quad + \frac{1}{2}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \nabla^2 \mathbf{A}(z_0) \left(\frac{t}{n} - z_0\right)^2 \mathbf{X}_{t-1}\mathbf{X}_{t-1}^{\top} + \text{s.o.},\end{aligned}$$

where s.o. denotes higher-order negligible terms. By similar arguments as in the derivation for Lemma 2, we can show that

$$\begin{aligned}&\frac{1}{n^{1+\alpha}h^2}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \nabla \mathbf{A}(z_0) \left(\frac{t}{n} - z_0\right) \mathbf{X}_{t-1}\mathbf{X}_{t-1}^{\top} \\ &= \nabla \mathbf{A}(z_0) \frac{1}{n^{1+\alpha}h}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\left(\frac{t-nz_0}{nh}\right) \mathbf{X}_{t-1}\mathbf{X}_{t-1}^{\top} \\ &= \nabla \mathbf{A}(z_0)\, \mu_1\, \mathbf{V}_{xx} + O_p\left(\frac{1}{n^{1-\alpha}h}\right),\end{aligned}$$

and

$$
\begin{aligned}
&\frac{1}{n^{1+\alpha}h^3}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\nabla^2\mathbf{A}(z_0)\left(\frac{t}{n}-z_0\right)^2\mathbf{X}_{t-1}\mathbf{X}_{t-1}^\top \\
&=\nabla^2\mathbf{A}(z_0)\frac{1}{n^{1+\alpha}h}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\left(\frac{t-nz_0}{nh}\right)^2\mathbf{X}_{t-1}\mathbf{X}_{t-1}^\top \\
&=\nabla^2\mathbf{A}(z_0)\,\mu_2\,\mathbf{V}_{xx}+O_p\left(\frac{1}{n^{1-\alpha}h}\right),
\end{aligned}
$$

where $\mu_2=\int_{-1}^{1}u^2K(u)\,du$. Since $\mu_1=\int_{-1}^{1}uK(u)\,du=0$, we have

$$
\begin{aligned}
\left\|\frac{1}{n^{1+\alpha}h}\mathbf{B}_n(z_0)\right\| &\leq \left\|\frac{1}{n^{1+\alpha}h}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\nabla\mathbf{A}(z_0)\left(\frac{t}{n}-z_0\right)\mathbf{X}_{t-1}\mathbf{X}_{t-1}^\top\right\| \\
&\quad+\left\|\frac{1}{n^{1+\alpha}h}\frac{1}{2}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\nabla^2\mathbf{A}(z_0)\left(\frac{t}{n}-z_0\right)^2\mathbf{X}_{t-1}\mathbf{X}_{t-1}^\top\right\| \\
&=O_p\left(\frac{1}{n^{1-\alpha}}\right)+O_p(h^2)
\end{aligned}
$$

This establishes (A.6), which completes the proofs of Theorem 2.

A.5. Proof of Theorem 3

Note that

$$
\begin{aligned}
n^{\frac{1+\alpha}{2}}\sqrt{h}\cdot\left(\widehat{\mathbf{A}}_n(z_0)-\mathbf{A}(z_0)\right) &= \left(\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\,\mathbf{B}_n(z_0)\right)\left(\frac{1}{n^{1+\alpha}h}\,\mathbf{\Delta}_n(z_0)\right)^{-1} \\
&\quad+\left(\frac{K_{1n}}{n^{\frac{1+\alpha}{2}}\sqrt{h}}(\mathbf{\Lambda}_{0x}+\mathbf{\Sigma}_{0x})\right)\left(\frac{1}{n^{1+\alpha}h}\,\mathbf{\Delta}_n(z_0)\right)^{-1} \\
&\quad+\left(\mathbf{\Phi}_0(1)\,\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}'\right)\left(\frac{1}{n^{1+\alpha}h}\,\mathbf{\Delta}_n(z_0)\right)^{-1} \\
&:=\mathcal{B}_{1n}+\mathcal{B}_{2n}+\mathcal{V}_n+o_p(1).
\end{aligned}
$$

We have shown in the proof of Theorem 2 that $\left\|\left(\frac{1}{n^{1+\alpha}h}\mathbf{\Delta}_n(z_0)\right)^{-1}\right\|_{sp}=O_p(1)$. Moreover, it is straightforward to deduce that

$$\mathcal{B}_{1n} = O_p\left(n^{\frac{1+\alpha}{2}} h^{5/2} + n^{\frac{3}{2}\alpha - \frac{1}{2}}\sqrt{h}\right), \;\; \mathcal{B}_{2n} = O_p\left(n^{\frac{1-\alpha}{2}}\sqrt{h}\right), \;\; \mathcal{V}_n = O_p(1),$$

Under $n^{\frac{1+\alpha}{2}} h^{5/2} \to c \in (0,\infty)$ the first component of $\mathcal{B}_{1n}$ is $O_p(1)$. In addition, the restriction $n^{1-\alpha}h^2 \to \infty$ implies the second component of $\mathcal{B}_{1n}$ is of smaller order than the first one, so that $\mathcal{B}_{1n} = O_p(1)$. Since $n^{\frac{1-\alpha}{2}}\sqrt{h} \to \infty$, $\mathcal{B}_{2n}$ dominates $\mathcal{B}_{1n}$ and $\mathcal{V}_n$. The simultaneous–equation bias dominates and we have

$$n^{\alpha} \cdot \left(\widehat{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0)\right) \xrightarrow{p} (\boldsymbol{\Lambda}_{0x} + \boldsymbol{\Sigma}_{0x})\ \mathbf{V}_{xx}^{-1}.$$

A.6. Proof of Theorem 4

Recall that, since $n^{\frac{1}{2}-\alpha}\sqrt{h} \to 0$, Lemma C.4 yields

$$\begin{aligned} & n^{\frac{1+\alpha}{2}}\sqrt{h} \cdot \operatorname{vec}\left(\widetilde{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0)\right) \\ = & \left\{\left(\frac{1}{n^{1+\alpha}h}\,\boldsymbol{\Delta}_n(z_0)\right)^{-1} \otimes \mathbf{I}_d\right\} \operatorname{vec}\left(\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\,\mathbf{B}_n(z_0)\right) \\ & + \left\{\left(\frac{1}{n^{1+\alpha}h}\,\boldsymbol{\Delta}_n(z_0)\right)^{-1} \otimes \mathbf{I}_d\right\} \operatorname{vec}\left(\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\widetilde{\boldsymbol{\Gamma}}_n(z_0)\right) + o_p(1), \end{aligned}$$

where $\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\widetilde{\boldsymbol{\Gamma}}_n(z_0) := \boldsymbol{\Phi}_0(1)\,\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^{\top}$. First, under the condition $n^{1-\alpha}h^2 \to \infty$, Lemma 2 gives $\frac{1}{n^{1+\alpha}h}\boldsymbol{\Delta}_n(z_0) \xrightarrow{p} \mathbf{V}_{xx}$. In addition, the derivation of (A.6) for Theorem 2 shows that $\frac{1}{n^{1+\alpha}h}\mathbf{B}_n(z_0) = \frac{h^2}{2}\,\nabla^2\mathbf{A}(z_0)\,\mu_2\,\mathbf{V}_{xx} + o_p(1)$. This implies that

$$\begin{aligned}
&\left\|\left\{\left(\frac{1}{n^{1+\alpha}h}\,\boldsymbol{\Delta}_n(z_0)\right)^{-1}\otimes\mathbf{I}_d\right\}\mathrm{vec}\left(\frac{1}{n^{1+\alpha}h}\,\mathbf{B}_n(z_0)\right) - \left\{(\mathbf{V}_{xx})^{-1}\otimes\mathbf{I}_d\right\}\mathrm{vec}\left(\frac{h^2}{2}\,\nabla^2\mathbf{A}(z_0)\,\mu_2\,\mathbf{V}_{xx}\right)\right\| \\
&\leq \left\|\left\{\left(\frac{1}{n^{1+\alpha}h}\,\boldsymbol{\Delta}_n(z_0)\right)^{-1}\otimes\mathbf{I}_d\right\} - \left\{(\mathbf{V}_{xx})^{-1}\otimes\mathbf{I}_d\right\}\right\|\,\left\|\mathrm{vec}\left(\frac{1}{n^{1+\alpha}h}\,\mathbf{B}_n(z_0)\right) - \mathrm{vec}\left(\frac{h^2}{2}\,\nabla^2\mathbf{A}(z_0)\,\mu_2\,\mathbf{V}_{xx}\right)\right\| \\
&\quad + \left\|\left\{\left(\frac{1}{n^{1+\alpha}h}\,\boldsymbol{\Delta}_n(z_0)\right)^{-1}\otimes\mathbf{I}_d\right\} - \left\{(\mathbf{V}_{xx})^{-1}\otimes\mathbf{I}_d\right\}\right\|\,\left\|\mathrm{vec}\left(\frac{h^2}{2}\,\nabla^2\mathbf{A}(z_0)\,\mu_2\,\mathbf{V}_{xx}\right)\right\| \\
&\quad + \left\|(\mathbf{V}_{xx})^{-1}\otimes\mathbf{I}_d\right\|\,\left\|\mathrm{vec}\left(\frac{1}{n^{1+\alpha}h}\,\mathbf{B}_n(z_0)\right) - \mathrm{vec}\left(\frac{h^2}{2}\,\nabla^2\mathbf{A}(z_0)\,\mu_2\,\mathbf{V}_{xx}\right)\right\| \\
&= o_p(1).
\end{aligned}$$

Thus, we can write

$$n^{\frac{1+\alpha}{2}}\sqrt{h}\cdot\mathrm{vec}\left(\widetilde{\mathbf{A}}_n(z_0) - \mathbf{A}(z_0) - \frac{h^2}{2}\,\mu_2\,\nabla^2\mathbf{A}(z_0)\right) = \left\{\left(\frac{1}{n^{1+\alpha}h}\,\boldsymbol{\Delta}_n(z_0)\right)^{-1}\otimes\mathbf{I}_d\right\}\mathrm{vec}\left(\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\widetilde{\boldsymbol{\Gamma}}_n(z_0)\right).$$

Similarly, we have

$$\left\|\left\{\left(\frac{1}{n^{1+\alpha}h}\,\boldsymbol{\Delta}_n(z_0)\right)^{-1}\otimes\mathbf{I}_d\right\}\mathrm{vec}\left(\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\widetilde{\boldsymbol{\Gamma}}_n(z_0)\right) - \left\{(\mathbf{V}_{xx})^{-1}\otimes\mathbf{I}_d\right\}\mathrm{vec}\left(\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\widetilde{\boldsymbol{\Gamma}}_n(z_0)\right)\right\| = o_p(1).$$

Since

$$\begin{aligned}
\mathrm{vec}\left(\boldsymbol{\Phi}_0(1)\,\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^\top\right) &= (\mathbf{I}_k\otimes\boldsymbol{\Phi}_0(1))\,\mathrm{vec}\left(\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^\top\right) \\
&= (\mathbf{I}_k\otimes\boldsymbol{\Phi}_0(1))\,\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)(\mathbf{X}_{t-1}\otimes\boldsymbol{\varepsilon}_t)\,,
\end{aligned}$$

the result follows immediately by applying Lemma C.5(iii).

B — Auxiliary lemmas for Section 3

Lemma B.1. If Assumptions 1-3 are satisfied. Then, for any fixed $z_0\in(0,1)$,

$$\boldsymbol{\Delta}_n(z_0)\xrightarrow{d}\boldsymbol{\Delta}(z_0),$$

where $\mathbf{\Delta}(z_0)$ is defined as in Theorem 1 and is nonsingular $a.s.$.

Proof. We follow the steps as in Phillips, Li, and Gao 2017. Observe that

$$
\begin{aligned}
&\mathbf{D}_n^+\mathbf{Q}_n^\top\left(\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{X}_t\mathbf{X}_t^\top\right)\mathbf{Q}_n\mathbf{D}_n^+\\
&=\begin{bmatrix}\frac{1}{nh}\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{q}_n^\top\left(\frac{\mathbf{X}_t}{\sqrt{n}}\right)\left(\frac{\mathbf{X}_t}{\sqrt{n}}\right)^\top\mathbf{q}_n & \frac{1}{nh^{3/2}}\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{q}_n^\top\left(\frac{\mathbf{X}_t}{\sqrt{n}}\right)\left(\frac{\mathbf{X}_t}{\sqrt{n}}\right)^\top\mathbf{q}_n^\perp\\ \frac{1}{nh^{3/2}}\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)(\mathbf{q}_n^\perp)^\top\left(\frac{\mathbf{X}_t}{\sqrt{n}}\right)\left(\frac{\mathbf{X}_t}{\sqrt{n}}\right)^\top\mathbf{q}_n & \frac{1}{nh^2}\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)(\mathbf{q}_n^\perp)^\top\left(\frac{\mathbf{X}_t}{\sqrt{n}}\right)\left(\frac{\mathbf{X}_t}{\sqrt{n}}\right)^\top\mathbf{q}_n^\perp\end{bmatrix}\\
&:=\begin{bmatrix}\Delta_n(z_0,1) & \mathbf{\Delta}_n(z_0,2)\\ \mathbf{\Delta}_n(z_0,2)^\top & \mathbf{\Delta}_n(z_0,3)\end{bmatrix}.
\end{aligned}
$$

Let us first consider $\Delta_n(z_0,1)$. Following the proof of Lemma 2, we can show that

$$
\begin{aligned}
\mathbf{\Delta}_n(z_0,1)&=\frac{1}{nh}\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{q}_n^\top\left(\frac{\mathbf{X}_t-\mathbf{X}_{z_n}+\mathbf{X}_{z_n}}{\sqrt{n}}\right)\left(\frac{\mathbf{X}_t-\mathbf{X}_{z_n}+\mathbf{X}_{z_n}}{\sqrt{n}}\right)^\top\mathbf{q}_n\\
&=\mathbf{q}_n^\top\left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right)\left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right)^\top\mathbf{q}_n\left(\frac{1}{nh}\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\right)+o_p(1) \qquad \text{(B.1)}\\
&=\mathbf{q}_n^\top\mathbf{k}_n\mathbf{k}_n^\top\mathbf{q}_n+o_p(1)\\
&=\mathbf{k}_n^\top\mathbf{k}_n+o_p(1).
\end{aligned}
$$

Now we consider $\mathbf{\Delta}_n(z_0,2)$. Recall that $z_n=\lfloor n(z_0-h)\rfloor$, we have the decomposition

$$
\mathbf{X}_t=\mathbf{X}_{z_n}+\sum_{j=z_n+1}^t\mathbf{R}_n^{t-j}\mathbf{u}_{xj}+\sum_{j=1}^{z_n}\left(\mathbf{R}_n^{t-j}-\mathbf{R}_n^{z_n-j}\right)\mathbf{u}_{xj}+\left(\mathbf{R}_n^t-\mathbf{R}_n^{z_n}\right)\mathbf{X}_0,
$$

where

$$
\boldsymbol{\eta}_t:=\sum_{j=z_n+1}^t\mathbf{R}_n^{t-j}\mathbf{u}_{xj},\quad \boldsymbol{\xi}_t:=\sum_{j=1}^{z_n}\left(\mathbf{R}_n^{t-j}-\mathbf{R}_n^{z_n-j}\right)\mathbf{u}_{xj}+\left(\mathbf{R}_n^t-\mathbf{R}_n^{z_n}\right)\mathbf{X}_0.
$$

Since by construction $(\mathbf{q}_n^{\perp})^\top \mathbf{X}_{z_n} = 0\ a.s.$, we have

$$\begin{aligned}
\mathbf{\Lambda}_n(z_0,2) &= \frac{1}{nh^{3/2}} \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \mathbf{q}_n^\top \left(\frac{\mathbf{X}_t}{\sqrt{n}}\right) \left(\frac{\boldsymbol{\eta}_t + \boldsymbol{\xi}_t}{\sqrt{n}}\right)^\top \mathbf{q}_n^{\perp} + o_p(1) \\
&= \frac{1}{nh^{3/2}} \mathbf{q}_n^\top \left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right) \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \left(\frac{\boldsymbol{\eta}_t + \boldsymbol{\xi}_t}{\sqrt{n}}\right)^\top \mathbf{q}_n^{\perp} \\
&\quad + \frac{1}{nh^{3/2}} \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \mathbf{q}_n^\top \left(\frac{\boldsymbol{\eta}_t + \boldsymbol{\xi}_t}{\sqrt{n}}\right) \left(\frac{\boldsymbol{\eta}_t + \boldsymbol{\xi}_t}{\sqrt{n}}\right)^\top \mathbf{q}_n^{\perp} + o_p(1) \\
&:= \mathbf{\Lambda}_n(z_0,2,1) + \mathbf{\Lambda}_n(z_0,2,2) + o_p(1).
\end{aligned}$$

By Lemma S.1(ii), we have $\sup_{t \in \overline{N}_{nz_0}(h)} \|\boldsymbol{\eta}_t + \boldsymbol{\xi}_t\| = O_p\left(\sqrt{nh}\right)$, where $\overline{N}_{nz_0}(h)$ is a set of integers in $N_{nz_0}(h) = [\lfloor (z_0 - h)n \rfloor, \lfloor (z_0 + h)n \rfloor]$. This implies that

$$\begin{aligned}
\|\mathbf{\Lambda}_n(z_0,2,2)\| &\leq \frac{1}{nh^{3/2}} \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \left| \mathbf{q}_n^\top \left(\frac{\boldsymbol{\eta}_t + \boldsymbol{\xi}_t}{\sqrt{n}}\right) \left(\frac{\boldsymbol{\eta}_t + \boldsymbol{\xi}_t}{\sqrt{n}}\right)^\top \mathbf{q}_n^{\perp} \right| \\
&\leq \frac{1}{nh^{3/2}} \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \left\| \frac{\boldsymbol{\eta}_t + \boldsymbol{\xi}_t}{\sqrt{n}} \right\|^2 \\
&\leq \frac{1}{nh^{3/2}} \cdot \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \cdot \sup_{t \in \overline{N}_{nz_0}(h)} \left\| \frac{\boldsymbol{\eta}_t + \boldsymbol{\xi}_t}{\sqrt{n}} \right\|^2 \\
&= \frac{1}{nh^{3/2}} \cdot O(nh) \cdot O_p(h) = O_p(\sqrt{h}).
\end{aligned}$$

The dominating term is thus $\mathbf{\Lambda}_n(z_0,2,1)$. We have

$$\begin{aligned}
\mathbf{\Lambda}_n(z_0,2,1) &= \frac{1}{nh^{3/2}} \mathbf{q}_n^\top \left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right) \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \left(\frac{\boldsymbol{\xi}_t}{\sqrt{n}}\right)^\top \mathbf{q}_n^{\perp} \\
&\quad + \frac{1}{nh^{3/2}} \mathbf{q}_n^\top \left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right) \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \left(\frac{\boldsymbol{\eta}_t}{\sqrt{n}}\right)^\top \mathbf{q}_n^{\perp} + o_p(1).
\end{aligned}$$

Obviously

$$
\begin{aligned}
&\left\| \frac{1}{nh^{3/2}} \mathbf{q}_n^\top \left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}} \right) \sum_{t=1}^n K\left(\frac{t - nz_0}{nh} \right) \left(\frac{\boldsymbol{\xi}_t}{\sqrt{n}} \right)^\top \mathbf{q}_n^\perp \right\| \\
\leq{}& \frac{1}{\sqrt{nh}} \cdot \left\| \mathbf{q}_n^\top \left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}} \right) \right\| \cdot \left(\frac{1}{nh} \sum_{t=1}^n K\left(\frac{t - nz_0}{nh} \right) \| \boldsymbol{\xi}_t^\top \mathbf{q}_n^\perp \| \right) \\
\leq{}& \frac{1}{\sqrt{nh}} \cdot \left\| \mathbf{q}_n^\top \left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}} \right) \right\| \cdot \sup_{t \in \overline{N}_{nz_0}(h)} \| \boldsymbol{\xi}_t \| \cdot \left(\frac{1}{nh} \sum_{t=1}^n K\left(\frac{t - nz_0}{nh} \right) \right) = o_p(1).
\end{aligned}
$$

Then, we have

$$
\boldsymbol{\Delta}_n(z_0, 2) = \frac{\sqrt{2}}{nh} \mathbf{q}_n^\top \left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}} \right) \sum_{t=1}^n K\left(\frac{t - nz_0}{nh} \right) \left(\frac{\boldsymbol{\eta}_t}{\sqrt{2nh}} \right)^\top \mathbf{q}_n^\perp + o_p(1). \tag{B.2}
$$

Finally, following the proof of (B.2) and the fact that $(\mathbf{q}_n^\perp)^\top \mathbf{X}_{z_n} = 0 \; a.s.$, we can show that

$$
\begin{aligned}
\boldsymbol{\Lambda}_n(z_0, 3) &= \frac{1}{nh^2} \sum_{t=1}^n K\left(\frac{t - nz_0}{nh} \right) (\mathbf{q}_n^\perp)^\top \left(\frac{\boldsymbol{\eta}_t + \boldsymbol{\xi}_t}{\sqrt{n}} \right) \left(\frac{\boldsymbol{\eta}_t + \boldsymbol{\xi}_t}{\sqrt{n}} \right)^\top \mathbf{q}_n^\perp + o_p(1) \\
&= \frac{1}{nh^2} \sum_{t=1}^n K\left(\frac{t - nz_0}{nh} \right) (\mathbf{q}_n^\perp)^\top \left(\frac{\boldsymbol{\eta}_t}{\sqrt{n}} \right) \left(\frac{\boldsymbol{\eta}_t}{\sqrt{n}} \right)^\top \mathbf{q}_n^\perp + o_p(1) \\
&= (\mathbf{q}_n^\perp)^\top \left[\frac{2}{nh} \sum_{t=1}^n K\left(\frac{t - nz_0}{nh} \right) \left(\frac{\boldsymbol{\eta}_t}{\sqrt{2nh}} \right) \left(\frac{\boldsymbol{\eta}_t}{\sqrt{2nh}} \right)^\top \right] \mathbf{q}_n^\perp + o_p(1).
\end{aligned} \tag{B.3}
$$

Using the invariance principle for linear processes under Assumption **??** and Lemma S.1(i), there exist two independent Ornstein–Uhlenbeck processes $\mathbf{K}_C(\cdot)$ and $\mathbf{K}_C^*(\cdot)$ such that, for any fixed $z_0 \in (0, 1)$ and $0 < p \leq 1$,

$$
\begin{pmatrix} \mathbf{k}_n \\ \frac{1}{\sqrt{2nh}} \sum_{j=z_n+1}^{z_n(p)} \mathbf{R}_n^{z_n(p)-j} \mathbf{u}_{xj} \end{pmatrix} \Rightarrow \begin{pmatrix} \mathbf{k} \\ \mathbf{K}_C^*(p) \end{pmatrix}, \qquad \mathbf{k}_n := \frac{\mathbf{X}_{z_n}}{\sqrt{n}}, \; \mathbf{k} := \mathbf{K}_C(z_0).
$$

Since all block components in $\boldsymbol{\Delta}_n(z_0)$ are functions of the same localized partial-sum process of X_t around z_0 and the same random rotation based on X_{z_n}, their joint convergence follows directly, which may be viewed as a straightforward application of the continuous mapping theorem to a vector of block components. Lemma B.1 then follows by combining (B.1)–(B.3).

Now we show that $\mathbf{\Delta}(z_0)$ is nonsingular $a.s.$. In view of Proposition 3.9.7. in Bernstein 2018, we need to verify that $\Delta(z_0, 1) > 0$ $a.s.$ and $\mathbf{\Delta}(z_0, 3) - \frac{1}{\Delta(z_0,1)}\mathbf{\Delta}^\top(z_0, 2)\mathbf{\Delta}(z_0, 2)$ is positive definite $a.s.$. Given that $\mathbf{k} \equiv \mathbf{k}_{z_0} = \mathbf{K}_C(z_0)$ is Gaussian at fixed $z_0 \in (0, 1)$ with full rank covariance matrix $\mathbf{\Sigma}_K(z_0) := \int_0^{z_0} e^{(z_0-s)\mathbf{C}}\mathbf{\Omega}_{xx}e^{(z_0-s)\mathbf{C}'}ds$, $\mathbb{P}(\mathbf{k} = \mathbf{0}) = 0$ so that $\mathbb{P}\left(\mathbf{k}^\top\mathbf{k} = \mathbf{0}\right) = 0$.

Next, given that $\int_{-1}^1 K(r)dr = 1$, we can write (by ignoring normalizing constant)

$$\begin{aligned}
&\mathbf{\Delta}(z_0, 3) - \frac{1}{\Delta(z_0, 1)}\mathbf{\Delta}^\top(z_0, 2)\mathbf{\Delta}(z_0, 2)\\
&= \left(\mathbf{q}^\perp\right)^\top \Bigg[\int_{-1}^1 K(r)\mathbf{K}_C^*\left(\frac{r+1}{2}\right)\mathbf{K}_C^{*\top}\left(\frac{r+1}{2}\right)dr\\
&\qquad - \left(\int_{-1}^1 K(r)\mathbf{K}_C^*\left(\frac{r+1}{2}\right)dr\right)\left(\int_{-1}^1 K(r)\mathbf{K}_C^{*\top}\left(\frac{r+1}{2}\right)dr\right)\Bigg]\mathbf{q}^\perp\\
&= \left(\mathbf{q}^\perp\right)^\top \Bigg[\int_{-1}^1 K(r)\left(\mathbf{K}_C^*\left(\frac{r+1}{2}\right) - \int_{-1}^1 K(r)\mathbf{K}_C^{*\top}\left(\frac{r+1}{2}\right)dr\right)\\
&\qquad \cdot\left(\mathbf{K}_C^*\left(\frac{r+1}{2}\right) - \int_{-1}^1 K(r)\mathbf{K}_C^{*\top}\left(\frac{r+1}{2}\right)dr\right)^\top dr\Bigg]\mathbf{q}^\perp.
\end{aligned}$$

Take any $\mathbf{x} \in \mathbb{R}^{k-1} \neq \mathbf{0}$ and let $\omega := \mathbf{q}^\perp\mathbf{x}$. We can write the above as

$$\begin{aligned}
&\mathbf{x}^\top\left(\mathbf{q}^\perp\right)^\top \Bigg[\int_{-1}^1 K(r)\left(\mathbf{K}_C^*\left(\frac{r+1}{2}\right) - \int_{-1}^1 K(r)\mathbf{K}_C^*\left(\frac{r+1}{2}\right)dr\right)\\
&\qquad \cdot\left(\mathbf{K}_C^*\left(\frac{r+1}{2}\right) - \int_{-1}^1 K(r)\mathbf{K}_C^*\left(\frac{r+1}{2}\right)dr\right)^\top dr\Bigg]\mathbf{q}^\perp\\
&= \int_{-1}^1 K(r)\left(\omega^\top\left(\mathbf{K}_C^*\left(\frac{r+1}{2}\right) - \int_{-1}^1 K(r)\mathbf{K}_C^*\left(\frac{r+1}{2}\right)dr\right)\right)^2 dr \geq 0, \quad a.s.
\end{aligned}$$

What remains is to show that

$$\mathbb{P}\left\{\int_{-1}^1 K(r)\left(\omega^\top\left(\mathbf{K}_C^*\left(\frac{r+1}{2}\right) - \int_{-1}^1 K(r)\mathbf{K}_C^{*\top}\left(\frac{r+1}{2}\right)dr\right)\right)^2 dr = 0\right\} = 0. \tag{B.4}$$

Since Assumption **??** implies that $K(r) \geq 0$ and is positive on a set of nonzero Lebesgue measure in $[-1, 1]$, on the event that (B.4) holds we must have

$$\omega^\top\mathbf{K}_C^*\left(\frac{r+1}{2}\right) = \omega^\top\int_{-1}^1 K(s)\mathbf{K}_C^*\left(\frac{s+1}{2}\right)ds, \quad a.s.$$

for all $r \in [-1,1]$ with $K(r) > 0$ except on a set of Lebesgue measure zero. This can only happen if for almost every realization, the sample path $r \mapsto \boldsymbol{\omega}^\top \mathbf{K}_C^*\left(\frac{r+1}{2}\right)$ is constant in r, which has probability zero since $\mathbf{K}_C^*\left(\frac{r+1}{2}\right)$ is a nondegenerate continuous Gaussian process. □

Lemma B.2. If Assumptions 1-3 are satisfied. Then, for any fixed $z_0 \in (0,1)$,

$$\boldsymbol{\Gamma}_n(z_0) \xrightarrow{d} \boldsymbol{\Gamma}(z_0),$$

where $\boldsymbol{\Gamma}(z_0)$ is defined as in Theorem 1.

Proof. We have

$$\begin{aligned}
\boldsymbol{\Gamma}_n(z_0) &= \left(\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\mathbf{X}_t^\top\right)\mathbf{Q}_n\mathbf{D}_n^+ \\
&= \left[\tfrac{1}{n\sqrt{h}}\sum_{t=1}^n K\left(\tfrac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\mathbf{X}_t^\top\mathbf{q}_n \quad \tfrac{1}{nh}\sum_{t=1}^n K\left(\tfrac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\mathbf{X}_t^\top\mathbf{q}_n^\perp\right] \\
&:= \left[\boldsymbol{\Gamma}_n(z_0,1) \quad \boldsymbol{\Gamma}_n(z_0,2)\right].
\end{aligned}$$

For $\boldsymbol{\Gamma}_n(z_0,1)$, notice that

$$\begin{aligned}
\boldsymbol{\Gamma}_n(z_0,1) &= \frac{1}{n\sqrt{h}}\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\left(\mathbf{X}_t - \mathbf{X}_{z_n} + \mathbf{X}_{z_n}\right)^\top\mathbf{q}_n \\
&= \frac{1}{\sqrt{2nh}}\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\cdot\sqrt{2}\left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right)^\top\mathbf{q}_n + \frac{1}{n\sqrt{h}}\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\left(\mathbf{X}_t - \mathbf{X}_{z_n}\right)^\top\mathbf{q}_n \\
&:= \boldsymbol{\Gamma}_n(z_0,1,1) + \boldsymbol{\Gamma}_n(z_0,1,2).
\end{aligned}$$

Using Lemma B.3, we obtain

$$\begin{aligned}
\|\boldsymbol{\Gamma}_n(z_0,1,2)\| &\leq \frac{1}{n\sqrt{h}}\cdot\left\|\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\left(\mathbf{X}_t - \mathbf{X}_{z_n}\right)^\top\right\|\cdot\|\mathbf{q}_n\| \\
&= \frac{1}{n\sqrt{h}}\cdot O_p(nh) = O_p\left(\sqrt{h}\right),
\end{aligned}$$

so that the dominating term is $\mathbf{\Gamma}_n(z_0,1,1)$. For $\mathbf{\Gamma}_n(z_0,2)$, given that $(\mathbf{q}_n^{\perp})^{\top}\mathbf{X}_{z_n}=0$ $a.s.$, we have

$$\begin{aligned}\mathbf{\Gamma}_n(z_0,2)&=\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\left(\mathbf{X}_t-\mathbf{X}_{z_n}+\mathbf{X}_{z_n}\right)^{\top}\mathbf{q}_n^{\perp}\\&=\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\left(\mathbf{X}_t-\mathbf{X}_{z_n}\right)^{\top}\mathbf{q}_n^{\perp}+o_p(1):=\mathbf{\Gamma}_n(z_0,2,1)+o_p(1).\end{aligned}$$

Then, using the weak convergence result of

$$\begin{pmatrix}\frac{1}{\sqrt{2nh}}\sum_{j=z_n+1}^{z_n(p)}\mathbf{u}_{0j}\\ \frac{1}{\sqrt{2nh}}\sum_{j=z_n+1}^{z_n(p)}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\end{pmatrix}\Rightarrow\begin{pmatrix}\mathbf{B}_0(p)\\ \mathbf{K}_C(p)\end{pmatrix},$$

where $z_n(p)=z_n+\lfloor 2pnh\rfloor+1$ for $0<p\le 1$, Lemma B.3, and the continuous mapping theorem, we can show that

$$\begin{aligned}&\begin{bmatrix}\mathbf{\Gamma}_n(z_0,1,1) & \mathbf{\Gamma}_n(z_0,2,1)\end{bmatrix}\\&=\begin{bmatrix}\frac{1}{\sqrt{2nh}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\cdot\sqrt{2}\left(\frac{\mathbf{X}_{z_n}}{\sqrt{n}}\right)^{\top}\mathbf{q}_n & \frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\left(\mathbf{X}_t-\mathbf{X}_{z_n}\right)^{\top}\mathbf{q}_n^{\perp}\end{bmatrix}\\&\xrightarrow{d}\begin{bmatrix}\sqrt{2}\,\|\mathbf{k}\|\int_{-1}^{1}K(r)d\mathbf{B}_0^*\left(\frac{r+1}{2}\right) & \left[\int_{-1}^{1}K(r)d\mathbf{B}_0^*\left(\frac{r+1}{2}\right)\mathbf{K}_C^{*\top}\left(\frac{r+1}{2}\right)+\frac{1}{2}\left(\mathbf{\Sigma}_{0x}+\mathbf{\Lambda}_{0x}\right)\right]\mathbf{q}^{\perp}\end{bmatrix}.\end{aligned}$$

□

Lemma B.3. If Assumptions 1-3 are satisfied. Then, for any fixed $z_0\in(0,1)$,

$$\frac{1}{2nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\left(\mathbf{X}_t-\mathbf{X}_{z_n}\right)^{\top}\xrightarrow{d}\int_{-1}^{1}d\mathbf{B}_0^*\left((r+1)/2\right)K(r)\mathbf{K}_C^{*\top}\left((r+1)/2\right)+\frac{1}{2}\left[\mathbf{\Sigma}_{0x}+\mathbf{\Lambda}_{0x}\right],$$

where $z_n=\lfloor n(z_0-h)\rfloor$ and $\left(\mathbf{K}_C^{\top *}(r),\mathbf{B}_0^{\top *}(r)\right)^{\top}$ is an independent copy of $\left(\mathbf{K}_C^{\top}(r),\mathbf{B}_0^{\top}(r)\right)^{\top}$.

Proof. Let $\overline{N}_{nz_0}(h)$ be a set of integers in $N_{nz_0}(h)=[\lfloor(z_0-h)n\rfloor,\lfloor(z_0+h)n\rfloor]$ and $z_n=\lfloor n(z_0-$

$h)\rfloor$. We have

$$
\begin{aligned}
&\frac{1}{2nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\left(\mathbf{X}_t-\mathbf{X}_{z_n}\right)^\top \\
&= \frac{1}{2nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\left\{\sum_{j=z_n+1}^{t}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}+\sum_{j=1}^{z_n}\left(\mathbf{R}_n^{t-j}-\mathbf{R}_n^{z_n-j}\right)\mathbf{u}_{xj}+\left(\mathbf{R}_n^{t}-\mathbf{R}_n^{z_n}\right)\mathbf{X}_0\right\}^\top \\
&= \frac{1}{2nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\left\{\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right\}^\top + \frac{1}{2nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\mathbf{u}_{xt}^\top \\
&\quad + \frac{1}{2nh}\sum_{t\in\overline{N}^+_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\left\{\sum_{j=1}^{z_n}\left(\mathbf{R}_n^{t-j}-\mathbf{R}_n^{z_n-j}\right)\mathbf{u}_{xj}\right\}^\top \\
&\quad + \frac{1}{2nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\left\{\left(\mathbf{R}_n^{t}-\mathbf{R}_n^{z_n}\right)\mathbf{X}_0\right\}^\top \\
&:= \sum_{i=1}^{4}\mathbf{S}_{in}.
\end{aligned}
$$

We first show that $\|\mathbf{S}_{3n}\| = o_p(1)$. To see this, note that

$$
\|\mathbf{S}_{3n}\| \le \sup_{t\in\overline{N}^+_{nz_0}(h)}\left\|\sum_{j=1}^{z_n}\left(\mathbf{R}_n^{t-j}-\mathbf{R}_n^{z_n-j}\right)\mathbf{u}_{xj}\right\|\left\|\frac{1}{2nh}\sum_{t\in\overline{N}^+_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\right\|.
$$

Following the proofs of Lemma S.1(ii) and the central limit theorem (C.L.T.) for weighted linear processes (Abadir et al. 2014; Yan, Gao, and Peng 2025), we have $\sup_{t\in\overline{N}_{nz_0}(h)}\left\|\sum_{j=1}^{z_n}\left(\mathbf{R}_n^{t-j}-\mathbf{R}_n^{z_n-j}\right)\mathbf{u}_{xj}\right\| = o_p(\sqrt{nh})$ and $\left\|\frac{1}{2nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\right\| = O_p\left(\frac{1}{\sqrt{nh}}\right)$. This implies that $\|\mathbf{S}_{3n}\| = o_p(1)$. Similar arguments also yields $\|\mathbf{S}_{4n}\| = o_p(1)$.

We now move on to $\mathbf{S}_{1n}$ and $\mathbf{S}_{2n}$. $\mathbf{S}_{2n}$ is straightforward, since by the weak law of large numbers (W.L.L.N.) for weighted linear processes (Abadir et al. 2014; Yan, Gao, and Peng 2025), we have

$$
2\mathbf{S}_{2n} \xrightarrow{p} \mathbf{\Sigma}_{0x}.
$$

What remains are $\mathbf{S}_{1n}$. BN decomposition on $\mathbf{u}_{0t}$ gives

$$\begin{aligned}\mathbf{S}_{1n} &= \mathbf{\Phi}_0(1)\frac{1}{2nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\left\{\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right\}^\top \\ &\quad - \frac{1}{2nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\Delta\tilde{\mathbf{u}}_{0t}\left\{\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right\}^\top \\ &:= \mathbf{S}_{1n}^{(1)} + \mathbf{S}_{1n}^{(2)}.\end{aligned}$$

For $\mathbf{S}_{1n}^{(1)}$, let

$$\mathbf{U}_n(r) := \mathbf{\Phi}_0(1)\frac{1}{\sqrt{2nh}}\sum_{j=z_n+1}^{z_n+\lfloor 2nhr\rfloor+1}\boldsymbol{\varepsilon}_j,\quad \mathbf{V}_n(r) := \frac{1}{\sqrt{2nh}}\sum_{j=z_n+1}^{z_n+\lfloor 2nhr\rfloor+1}\mathbf{R}_n^{z_n+\lfloor 2nhr\rfloor-j}\mathbf{u}_{xj}, \qquad r\in[0,1].$$

We note that

$$\begin{aligned}\mathbf{S}_{1n}^{(1)} &= \sum_{t\in\overline{N}_{nz_0}(h)}\left(\frac{1}{\sqrt{2nh}}\mathbf{\Phi}_0(1)\boldsymbol{\varepsilon}_t\right)\left\{\frac{1}{\sqrt{2nh}}K\left(\frac{t-nz_0}{nh}\right)\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right\}^\top \\ &= \int_0^1 K(r)d\mathbf{U}_n(r)\mathbf{V}_n^\top(r) + o_p(1).\end{aligned}$$

Using the weak convergence result of

$$\begin{pmatrix}\mathbf{U}_n(r)\\ \mathbf{V}_n(r)\end{pmatrix} \Rightarrow \begin{pmatrix}\mathbf{B}_0(r)\\ \mathbf{K}_C(r)\end{pmatrix}, \quad r\in[0,1],$$

we obtain

$$\mathbf{S}_{1n}^{(1)} \xrightarrow{d} \int_{-1}^1 K(r)d\mathbf{B}_0^*\left((r+1)/2\right)\mathbf{K}_C^{*\top}\left((r+1)/2\right),$$

where $\left(\mathbf{K}_C^{\top *}(r), \mathbf{B}_0^{\top *}(r)\right)^\top$ is an independent copy of $\left(\mathbf{K}_C^\top(r), \mathbf{B}_0^\top(r)\right)^\top$. Moving on to $\mathbf{S}_{1n}^{(2)}$, summation by parts yields

$$
\begin{aligned}
2\mathbf{S}_{1n}^{(2)} &= \frac{1}{nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\widetilde{\mathbf{u}}_{0t}\left\{\sum_{j=z_n+1}^{t}\mathbf{R}_n^{t+1-j}\mathbf{u}_{xj}-\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right\}^{\top} \\
&\quad+\frac{1}{nh}\sum_{t\in\overline{N}_{nz_0}(h)}\left\{K\left(\frac{t-nz_0}{nh}\right)-K\left(\frac{t-1-nz_0}{nh}\right)\right\}\tilde{\mathbf{u}}_{0,t-1}\left(\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right)^{\top} \\
&= \frac{1}{nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\widetilde{\mathbf{u}}_{0t}\left(\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right)^{\top}(\mathbf{R}_n-\mathbf{I}_k)^{\top}+\frac{1}{nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\tilde{\mathbf{u}}_{0t}\mathbf{u}_{xt}^{\top}\mathbf{R}_n^{\top} \\
&\quad+\frac{1}{nh}\sum_{t\in\overline{N}_{nz_0}(h)}\left\{K\left(\frac{t-nz_0}{nh}\right)-K\left(\frac{t-1-nz_0}{nh}\right)\right\}\tilde{\mathbf{u}}_{0,t-1}\left(\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right)^{\top}.
\end{aligned}
$$

Note that

$$
\begin{aligned}
&\left\|\frac{1}{nh}\sum_{t\in\overline{N}^{+}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\widetilde{\mathbf{u}}_{0t}\left(\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right)^{\top}(\mathbf{R}_n-\mathbf{I}_k)^{\top}\right\| \\
&\le \|\mathbf{R}_n-\mathbf{I}_k\|\sup_{t\in\overline{N}_{nz_0}(h)}\left\|\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right\|\left(\frac{1}{nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\|\widetilde{\mathbf{u}}_{0t}\|\right) \\
&= O(1/n)O_p\left(\sqrt{nh}\right)=o_p(1),
\end{aligned}
$$

which follows from the fact that by Lemma S.1(i) we have $\sup_{t\in\overline{N}_{nz_0}(h)}\left\|\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right\| = O_p\left(\sqrt{nh}\right)$ and $\|\widetilde{\mathbf{u}}_{0t}\|=O_p(1)$ holds uniformly over $t\in\overline{N}_{nz_0}(h)$. We also have

$$
\begin{aligned}
&\left\|\frac{1}{nh}\sum_{t\in\overline{N}_{nz_0}(h)}\left\{K\left(\frac{t-nz_0}{nh}\right)-K\left(\frac{t-1-nz_0}{nh}\right)\right\}\tilde{\mathbf{u}}_{0,t-1}\left(\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right)^{\top}\right\| \\
&\le \sup_{t\in\overline{N}_{nz_0}(h)}\left\|\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right\|\frac{1}{nh}\sum_{t\in\overline{N}_{nz_0}(h)}\left\|\left\{K\left(\frac{t-nz_0}{nh}\right)-K\left(\frac{t-1-nz_0}{nh}\right)\right\}\tilde{\mathbf{u}}_{0,t-1}\right\| \\
&\le \sup_{t\in\overline{N}_{nz_0}(h)}\left\|\sum_{j=z_n+1}^{t-1}\mathbf{R}_n^{t-j}\mathbf{u}_{xj}\right\|\left(\frac{O(1)}{nh}\sum_{t\in\overline{N}_{nz_0}(h)}(1/nh)\|\tilde{\mathbf{u}}_{0t}\|\right) \\
&= O_p\left(\sqrt{nh}\right)O(1/nh)=o_p(1),
\end{aligned}
$$

where we have used the fact $K(\cdot)$ is Lipschitz-continuous on $[-1,1]$. Finally, by the weak law of large numbers (W.L.L.N.) for weighted linear processes (Abadir et al. 2014; Yan, Gao, and Peng 2025), we have

$$\frac{1}{nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\tilde{\mathbf{u}}_{0t}\mathbf{u}_{xt}' \xrightarrow{p} \mathbf{\Lambda}_{0x}.$$

Given that

$$\begin{aligned}&\left\|\frac{1}{nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\widetilde{\mathbf{u}}_{0t}\mathbf{u}_{xt}^\top\mathbf{R}_n^\top - \frac{1}{nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\widetilde{\mathbf{u}}_{0t}\mathbf{u}_{xt}^\top\right\| \\ \leq\ & \|\mathbf{R}_n-\mathbf{I}_k\|\left\|\frac{1}{nh}\sum_{t\in\overline{N}_{nz_0}(h)} K\left(\frac{t-nz_0}{nh}\right)\widetilde{\mathbf{u}}_{0t}\mathbf{u}_{xt}^\top\right\| = O_p(1/n) = o_p(1),\end{aligned}$$

together with Slutsky's theorem, leads to

$$\mathbf{S}_{1n} \xrightarrow{d} \int_{-1}^{1} K(r)d\mathbf{B}_0^*\left((r+1)/2\right)\mathbf{K}_C^{*\top}\left((r+1)/2\right) + \frac{1}{2}\mathbf{\Lambda}_{0x}.$$

The proofs are completed by combining the results above. □

C — Auxiliary lemmas for Section 4

Lemma C.4. Under Assumptions 1-3, we have, for any fixed $z_0 \in (0,1)$

$$\begin{aligned}&\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\mathbf{\Gamma}_n(z_0) \\ =\ & \mathbf{\Phi}_0(1)\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^\top + \frac{K_{1n}}{n^{\frac{1+\alpha}{2}}\sqrt{h}}(\mathbf{\Lambda}_{0x}+\mathbf{\Sigma}_{0x}) + O_p\left(n^{\frac{1}{2}-\alpha}\sqrt{h}\right) + o_p(1),\end{aligned}$$

where $K_{1n} = \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)$ and $\mathbf{\Gamma}_n(z_0)$ is defined as in (A.5).

Proof. By BN decomposition on $\mathbf{u}_{0t}$ and summation by parts, we have the following decomposition:

$$\begin{aligned}
\boldsymbol{\Gamma}_n(z_0) &= \sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\mathbf{X}_{t-1}^\top\mathbf{R}_n + \sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\mathbf{u}_{xt}^\top \\
&= \boldsymbol{\Phi}_0(1)\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^\top\mathbf{R}_n - \sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\Delta\tilde{\boldsymbol{u}}_{0t}\mathbf{X}_{t-1}^\top\mathbf{R}_n + \sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\mathbf{u}_{xt}^\top \\
&= \boldsymbol{\Phi}_0(1)\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^\top\mathbf{R}_n + \sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_{0t}\Delta\mathbf{X}_t^\top\mathbf{R}_n \\
&\quad + K\left(\frac{1-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_{00}\mathbf{X}_0^\top\mathbf{R}_n - K\left(\frac{1-z_0}{h}\right)\tilde{\boldsymbol{u}}_{0n}\mathbf{X}_n^\top\mathbf{R}_n \\
&\quad + \sum_{t=1}^{n-1}\left\{K\left(\frac{t+1-nz_0}{nh}\right) - K\left(\frac{t-nz_0}{nh}\right)\right\}\tilde{\boldsymbol{u}}_{0t}\mathbf{X}_t^\top\mathbf{R}_n + \sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\mathbf{u}_{xt}^\top \\
&= \boldsymbol{\Phi}_0(1)\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^\top + \boldsymbol{\Phi}_0(1)\frac{1}{n^\alpha}\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^\top\mathbf{C} \\
&\quad + \frac{1}{n^\alpha}\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_{0t}\mathbf{X}_{t-1}^\top\mathbf{C} + \sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_{0t}\mathbf{u}_{xt}^\top \\
&\quad + \sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{0t}\mathbf{u}_{xt}^\top + O_p\left(n^{\alpha/2}\right),
\end{aligned}$$

where the final equality follows from Assumption 1(i), so that $K\left(\frac{1-nz_0}{nh}\right) = K\left(\frac{1-z_0}{h}\right) = 0$, for sufficiently large n, together with the fact that

$$\begin{aligned}
&\left\|\sum_{t=1}^{n-1}\left\{K\left(\frac{t+1-nz_0}{nh}\right) - K\left(\frac{t-nz_0}{nh}\right)\right\}\tilde{\boldsymbol{u}}_{0t}\mathbf{X}_t^\top\right\| \\
\leq& \max_{1\leq t\leq n-1}\|\mathbf{X}_t\|\sum_{t=1}^{n-1}\left\|\left\{K\left(\frac{t+1-nz_0}{nh}\right) - K\left(\frac{t-nz_0}{nh}\right)\right\}\tilde{\boldsymbol{u}}_{0t}\right\| = O_p\left(n^{\alpha/2}\right),
\end{aligned}$$

which follows from the Lipschitz-continuity of $K(\cdot)$.

By Lemma C.5(iii), we have $\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^\top = O_p\left(n^{\frac{1+\alpha}{2}}\sqrt{h}\right)$ so that

$$\frac{1}{n^\alpha}\sum_{t=1}^n K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^\top\mathbf{C} = O_p\left(n^{\frac{1-\alpha}{2}}\sqrt{h}\right).$$

In addition, Lemma C.5(i) leads to

$$\frac{1}{n^{\alpha}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_{0t}\mathbf{X}_{t-1}^{\top}\mathbf{C}=O_p\left(n^{1-\frac{\alpha}{2}}h\right).$$

Finally, by the central limit theorem (C.L.T.) for weighted linear processes (Abadir et al. 2014; Yan, Gao, and Peng 2025), we obtain

$$\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\left(\tilde{\boldsymbol{u}}_{0t}\mathbf{u}_{xt}^{\top}-\boldsymbol{\Lambda}_{0x}\right)=O_p(\sqrt{nh}),\quad \sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\left(\mathbf{u}_{0t}\mathbf{u}_{xt}^{\top}-\boldsymbol{\Sigma}_{0x}\right)=O_p(\sqrt{nh}).$$

Substituting all results above yields

$$\begin{aligned}\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\boldsymbol{\Gamma}_n(z_0)&=\boldsymbol{\Phi}_0(1)\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^{\top}\\&\quad+\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)(\boldsymbol{\Lambda}_{0x}+\boldsymbol{\Sigma}_{0x})+O_p\left(n^{\frac{1}{2}-\alpha}\sqrt{h}\right)+O_p\left(\frac{1}{n^{\frac{\alpha}{2}}}\right)+O_p\left(\frac{1}{\sqrt{nh}}\right),\end{aligned}$$

which implies that

$$\begin{aligned}&\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\left\{\boldsymbol{\Gamma}_n(z_0)-\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)(\boldsymbol{\Lambda}_{0x}+\boldsymbol{\Sigma}_{0x})\right\}\\&=\boldsymbol{\Phi}_0(1)\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^{\top}+O_p\left(n^{\frac{1}{2}-\alpha}\sqrt{h}\right)+o_p(1).\end{aligned}$$

This completes the proof. □

Lemma C.5. For model (1)-(3), suppose that Assumptions 1, 2, and 3 are satisfied. Then, for any fixed $z_0\in(0,1)$, we have

(i) $\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_t\mathbf{X}_{t-1}^{\top}=O_p\left(n^{1+\alpha/2}h\right)$;

(ii) $\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{xt}\mathbf{X}_{t-1}^{\top}=\Lambda_{xx}+O_p\left(\frac{1}{n^{1-\alpha}h}\right)$;

(iii) for fixed $\alpha\in(0,1)$, $\frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)(\mathbf{X}_{t-1}\otimes\boldsymbol{\varepsilon}_t)\xrightarrow{d}\mathcal{N}\left(\mathbf{0},\nu_0\mathbf{V}_{xx}\otimes\boldsymbol{\Sigma}_{\varepsilon}\right)$, where $\nu_0:=\int_{-1}^{1}K^2(u)du$.

Proof. Let us start with (i). We first note that

$$
\begin{aligned}
&\left\|\text{vec}\left(\frac{1}{n^{1+\alpha/2}h}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_t\mathbf{X}_{t-1}^{\top}\right)\right\| \\
&=\frac{1}{n^{1+\alpha/2}h}\left\|\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)(\mathbf{X}_{t-1}\otimes\tilde{\boldsymbol{u}}_t)\right\| \\
&\leq\frac{1}{n^{1+\alpha/2}h}\sum_{t=1}^{n}\left\|K\left(\frac{t-nz_0}{nh}\right)\left[\left(\sum_{j=1}^{t-1}\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}+\mathbf{R}_n^{t-1}\mathbf{X}_0\right)\otimes\tilde{\boldsymbol{u}}_t\right]\right\| \\
&\leq\frac{1}{n^{1+\alpha/2}h}\sum_{t=1}^{n}\left\|K\left(\frac{t-nz_0}{nh}\right)\left[\left(\sum_{j=1}^{t-1}\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}\right)\otimes\tilde{\boldsymbol{u}}_t\right]\right\| \\
&\qquad+\frac{1}{n^{1+\alpha/2}h}\sum_{t=1}^{n}\left\|K\left(\frac{t-nz_0}{nh}\right)\left[\left(\mathbf{R}_n^{t-1}\mathbf{X}_0\right)\otimes\tilde{\boldsymbol{u}}_t\right]\right\| \\
&\equiv\frac{1}{n^{1+\alpha/2}h}\sum_{t=1}^{n}\left\|K\left(\frac{t-nz_0}{nh}\right)\left(\sum_{j=1}^{t-1}\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}\right)\right\|\|\tilde{\boldsymbol{u}}_t\|+o_p(1),
\end{aligned}
$$

where $\frac{1}{n^{1+\alpha/2}h}\sum_{t=1}^{n}\left\|K\left(\frac{t-nz_0}{nh}\right)\left[\left(\mathbf{R}_n^{t-1}\mathbf{X}_0\right)\otimes\tilde{\boldsymbol{u}}_t\right]\right\|=o_p(1)$ is shown in Lemma S.2(i).

By Lemma 3.1(i) in Magdalinos and Phillips 2009, we have $\max_{1\leq t\leq n}\ \mathbb{E}\left\|\frac{1}{n^{\alpha/2}}\sum_{j=1}^{t-1}\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}\right\|^2\leq C$. Then, Cauchy–Schwarz inequality gives

$$
\begin{aligned}
&\mathbb{E}\left(\frac{1}{n^{1+\alpha/2}h}\sum_{t=1}^{n}\left\|K\left(\frac{t-nz_0}{nh}\right)\left(\sum_{j=1}^{t-1}\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}\right)\right\|\|\tilde{\boldsymbol{u}}_t\|\right) \\
&\leq\frac{\left(\mathbb{E}\|\tilde{\boldsymbol{u}}_1\|^2\right)^{1/2}}{n^{1+\alpha/2}h}\sum_{t=1}^{n}\left(K^2\left(\frac{t-nz_0}{nh}\right)\mathbb{E}\left\|\sum_{j=1}^{t-1}\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}\right\|^2\right)^{1/2} \\
&\leq\ \left(\mathbb{E}\|\tilde{\boldsymbol{u}}_1\|^2\right)^{1/2}\left(\max_{1\leq t\leq n}\ \mathbb{E}\left\|\frac{1}{n^{\alpha/2}}\sum_{j=1}^{t-1}\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}\right\|^2\right)^{1/2}\left(\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\right) \\
&\leq O(1)\left(\mathbb{E}\|\tilde{\boldsymbol{u}}_1\|^2\right)^{1/2},
\end{aligned}
$$

which follows from the fact that Assumption 3 ensures $\sum_{j=0}^{\infty}\left\|\tilde{\boldsymbol{\Phi}}_j\right\|<\infty$ (Phillips and Solo 1992), the triangle inequality and Cauchy–Schwarz yield $\mathbb{E}\|\tilde{\boldsymbol{u}}_t\|^2<\infty$ for all $t=1,2,\cdots,n$. This establishes (i).

For (ii), we first show that

$$\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^\top = o_p(1). \tag{C.1}$$

To see this, note that

$$\begin{aligned}\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^\top &= \frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\left(\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}+\mathbf{R}_n^{t-1}\mathbf{X}_0\right)^\top \\ &= \frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\left(\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}\right)^\top + \frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\left(\mathbf{R}_n^{t-1}\mathbf{X}_0\right)^\top\end{aligned}$$

By Lemma S.2(ii), we have

$$\left\|\mathrm{vec}\left(\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\left(\mathbf{R}_n^{t-1}\mathbf{X}_0\right)^\top\right)\right\| \le \frac{1}{nh}\sum_{t=1}^{n}\left\|K\left(\frac{t-nz_0}{nh}\right)\left[\left(\mathbf{R}_n^{t-1}\mathbf{X}_0\right)\otimes\boldsymbol{\varepsilon}_t\right]\right\| = o_p(1),$$

so that the dominating term is the first one. Note that $\boldsymbol{\varepsilon}_t\left(\sum_{j=1}^{t-1}\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}\right)^\top$ forms a martingale difference sequence. Thus, by Lemma 3.1(i) in Magdalinos and Phillips 2009 and the fact that $\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)^2 \to \int_{-1}^{1} K^2(u)du$, we have

$$\begin{aligned}&\mathbb{E}\left\|\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\ \mathrm{vec}\left[\boldsymbol{\varepsilon}_t\left(\sum_{j=1}^{t-1}\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}\right)^\top\right]\right\|^2 \\ &= \frac{1}{(nh)^2}\,\mathbb{E}\left\|\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)\left[\left(\sum_{j=1}^{t-1}\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}\right)\otimes\boldsymbol{\varepsilon}_t\right]\right\|^2 \\ &= \frac{1}{(nh)^2}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)^2\mathbb{E}\left\|\sum_{j=1}^{t-1}\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}\right\|^2\mathbb{E}\left\|\boldsymbol{\varepsilon}_t\right\|^2 \\ &\le \frac{\mathbb{E}\left\|\boldsymbol{\varepsilon}_1\right\|^2}{n^{1-\alpha}h}\left(\max_{2\le t\le n}\mathbb{E}\left\|n^{-\alpha/2}\sum_{j=1}^{t-1}\mathbf{R}_n^{t-j-1}\mathbf{u}_{xj}\right\|^2\right)\left(\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right)^2\right) \\ &= O\left(\frac{1}{n^{1-\alpha}h}\right).\end{aligned}$$

Now, by BN decomposition on $\mathbf{u}_{xt}$, we obtain

$$\begin{aligned}\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{xt}\mathbf{X}_{t-1}^{\top} &= \mathbf{\Phi}_x(1)\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\boldsymbol{\varepsilon}_t\mathbf{X}_{t-1}^{\top} - \frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\Delta\tilde{\boldsymbol{u}}_{xt}\mathbf{X}_{t-1}^{\top}\\ &= O_p\left(\frac{1}{n^{1-\alpha}h}\right) - \frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\Delta\tilde{\boldsymbol{u}}_{xt}\mathbf{X}_{t-1}^{\top}.\end{aligned}$$

Summation by parts gives

$$\begin{aligned}&-\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\Delta\tilde{\boldsymbol{u}}_{xt}\mathbf{X}_{t-1}^{\top}\\ &=\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_{xt}\left(\Delta\mathbf{X}_t\right)^{\top} - \frac{1}{nh}K\left(\frac{1-z_0}{h}\right)\tilde{\boldsymbol{u}}_{xn}\mathbf{X}_n^{\top} + \frac{1}{nh}K\left(\frac{1-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_{x0}\mathbf{X}_0^{\top}\\ &\quad+\frac{1}{nh}\sum_{t=1}^{n-1}\left\{K\left(\frac{t+1-nz_0}{nh}\right)-K\left(\frac{t-nz_0}{nh}\right)\right\}\tilde{\boldsymbol{u}}_{xt}\mathbf{X}_t^{\top}\\ &=\frac{1}{(nh)^{1+\alpha}}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_{xt}\mathbf{X}_{t-1}^{\top}\mathbf{C} + \frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_{xt}\mathbf{u}_{xt}^{\top} + O_p\left(\frac{1}{n^{1-\alpha/2}h}\right)\\ &= O_p\left(\frac{1}{n^{\alpha/2}}\right) + \frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_{xt}\mathbf{u}_{xt}^{\top} + o_p(1)\\ &= \frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_{xt}\mathbf{u}_{xt}^{\top} + o_p(1),\end{aligned}$$

where the second equality follows from the fact that $K(\cdot)$ is defined on $[-1,1]$ so $K\left(\frac{1-z_0}{h}\right) = K\left(\frac{1-nz_0}{nh}\right) = 0$, for sufficient large n, and the fact that

$$\begin{aligned}&\left\|\sum_{t=1}^{n-1}\left\{K\left(\frac{t+1-nz_0}{nh}\right)-K\left(\frac{t-nz_0}{nh}\right)\right\}\tilde{\boldsymbol{u}}_{xt}\mathbf{X}_t^{\top}\right\|\\ &\leq \max_{1\leq t\leq n-1}\|\mathbf{X}_t\|\sum_{t=1}^{n-1}\left\|\left\{K\left(\frac{t+1-nz_0}{nh}\right)-K\left(\frac{t-nz_0}{nh}\right)\right\}\tilde{\boldsymbol{u}}_{xt}\right\| = O_p\left(n^{\alpha/2}\right),\end{aligned}$$

which follows from the Lipschitz-continuity of $K(\cdot)$. The third inequality also follows from Lemma C.5(i). This implies that

$$\left\|\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\mathbf{u}_{xt}\mathbf{X}_{t-1}^{\top} - \frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{t-nz_0}{nh}\right)\tilde{\boldsymbol{u}}_{xt}\mathbf{u}_{xt}^{\top}\right\| = O_p\left(\frac{1}{n^{1-\alpha}h}\right).$$

Now, by the weak law of large numbers (WLLN) for weighted sums of linear processes (Abadir et al. 2014; Yan, Gao, and Peng 2025), we have, as $n \to \infty$

$$\frac{1}{nh} \sum_{t=1}^{n} K\left(\frac{t-nz_0}{nh}\right) \tilde{\boldsymbol{u}}_{xt} \mathbf{u}_{xt}^{\top} \xrightarrow{p} \Lambda_{xx}.$$

This establishes (ii).

For (iii), let $\boldsymbol{\xi}_{nt} = \frac{1}{n^{\frac{1+\alpha}{2}}\sqrt{h}} K\left(\frac{t-nz_0}{nh}\right) \mathbf{X}_{t-1} \otimes \boldsymbol{\varepsilon}_t$ and the filtration $\mathcal{F}_{nt} = \sigma\left(\mathbf{X}_0, \boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_{t-1}, \cdots\right)$. Clearly, $\{\boldsymbol{\xi}_{nt}, \mathcal{F}_{nt}\}$ is a martingale difference array. In addition, we have

$$\begin{aligned}
\sum_{t=1}^{n} \mathbb{E}\left(\boldsymbol{\xi}_{nt}\boldsymbol{\xi}_{nt}^{\top} | \mathcal{F}_{n,t-1}\right) &= \frac{1}{n^{1+\alpha}h} \sum_{t=1}^{n} K^2\left(\frac{t-nz_0}{nh}\right) \left(\mathbf{X}_{t-1}\mathbf{X}_{t-1}^{\top}\right) \otimes \mathbb{E}\left(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t^{\top} | \mathcal{F}_{n,t-1}\right) \\
&= \left(\frac{1}{n^{1+\alpha}h} \sum_{t=1}^{n} K^2\left(\frac{t-nz_0}{nh}\right) \mathbf{X}_{t-1}\mathbf{X}_{t-1}^{\top}\right) \otimes \boldsymbol{\Sigma}_{\varepsilon} \\
&= \nu_0 \mathbf{V}_{xx} \otimes \boldsymbol{\Sigma}_{\varepsilon} + o_p(1),
\end{aligned}$$

where $\nu_0 = \int_{-1}^{1} K^2(u)du$ and the last equality follows from arguments analogous to those used to establish (12), with $K(\cdot)$ replaced by $K^2(\cdot)$. Thus, in view of Proposition A1 in Magdalinos and Phillips 2009, it suffices to verify the conditional Lindeberg condition. For any $\delta > 0$, we need to show that

$$\sum_{t=1}^{n} \mathbb{E}\left(\left\|\boldsymbol{\xi}_{nt}\right\|^2 \mathbb{1}_{\{\|\boldsymbol{\xi}_{nt}\|>\delta\}} \,\middle|\, \mathcal{F}_{n,t-1}\right) = o_p(1). \tag{C.2}$$

Using the inequality $\mathbb{1}_{\{\|\boldsymbol{\xi}_{nt}\|>\delta\}} \leq \left\|\boldsymbol{\xi}_{nt}\right\|^2 / \delta^2$, we obtain

$$\begin{aligned}
\sum_{t=1}^{n} \mathbb{E}\left(\left\|\boldsymbol{\xi}_{nt}\right\|^2 \mathbb{1}_{\{\|\boldsymbol{\xi}_{nt}\|>\delta\}} \,\middle|\, \mathcal{F}_{n,t-1}\right) &\leq \frac{1}{\delta^2} \sum_{t=1}^{n} \mathbb{E}\left(\left\|\boldsymbol{\xi}_{nt}\right\|^4 \,\middle|\, \mathcal{F}_{n,t-1}\right) \\
&= \frac{1}{\delta^2} \frac{1}{n^{2(1+\alpha)}h^2} \sum_{t=1}^{n} K^4\left(\frac{t-nz_0}{nh}\right) \mathbb{E}\left(\left\|\mathbf{X}_{t-1}\right\|^4 \left\|\boldsymbol{\varepsilon}_t\right\|^4 \,\middle|\, \mathcal{F}_{n,t-1}\right) \\
&= \frac{1}{\delta^2} \frac{1}{n^{2(1+\alpha)}h^2} \sum_{t=1}^{n} K^4\left(\frac{t-nz_0}{nh}\right) \left\|\mathbf{X}_{t-1}\right\|^4 \mathbb{E}\left\|\boldsymbol{\varepsilon}_t\right\|^4 \\
&\leq O(1) \max_{1\leq t\leq n} \left\|\mathbf{X}_{t-1}\right\|^4 \cdot \frac{1}{n^{2(1+\alpha)}h^2} \sum_{t=1}^{n} K^4\left(\frac{t-nz_0}{nh}\right) \\
&= O_p\left(\frac{1}{nh}\right) = o_p(1),
\end{aligned}$$

where the third equality follows from the fact that $\{\varepsilon_t\}$ is i.i.d. and the final line uses $\mathbb{E}\,\|\varepsilon_t\|^4 < \infty$ (assumed by Assumption 3(i)), $\max_{1\le t\le n} \|\mathbf{X}_{t-1}\|^4 = O_p(n^{2\alpha})$ (by Lemma 3.1(i) in Magdalinos and Phillips 2009), and $\sum_{t=1}^{n} K^4\left(\frac{t-nz_0}{nh}\right) = O(nh)$ (implied by Assumption 1(i)). This verifies the conditional Lindeberg condition and completes the proof of (iii). □