ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Specification Testing in Structural Nonparametric Cointegration

Chaohua Dong and Jiti Gao

January 2014

Working Paper 02/14
(revised 20/12)

Specification Testing in Structural Nonparametric Cointegration*

Chaohua Dong† and Jiti Gao†

Monash University†

January 13, 2014

Abstract

This paper proposes two simple and new specification tests based on the use of an orthogonal series for a considerable class of cointegrated time series models with endogeneity and nonstationarity. The paper then establishes an asymptotic theory for each of the proposed tests. The first test is initially proposed for the case where the regression function involved is integrable, which fills a gap in the literature, and the second test is an extended version of the first test for covering a class of non–integrable functions. Endogeneity in two general forms is allowed in the models to be tested. A potential global departure in the alternative hypothesis, which is being overlooked by the literature, is investigated. The finite sample performance of the proposed tests is examined through using several simulated examples. Meanwhile, the second test is naturally applicable to the case where there is a type of endogeneity inherited in the relationship between the United States aggregate consumers' consumption expenditure and disposable income over the period of 1960–2009. Our experience generally shows that the proposed tests are easily implementable and also have stable sizes and good power properties even when the 'distance' between the null hypothesis and a sequence of local alternatives is asymptotically negligible.

JEL classification: C12; C14

Keywords: Consumption–income model; Endogeneity; Integrated time series; Linear process; Orthogonal series estimation; Parametric specification

*The authors first acknowledge constructive comments and suggestions from Professor Peter Phillips and Professor Peter Robinson on earlier versions. Thanks from the authors go to the seminar participants at various seminars for their comments and suggestions. Thanks also go to the Australian Research Council Discovery Grants Program for its support under Grant number: DP1096374. *Corresponding author*: Jiti Gao, Department of Econometrics and Business Statistics, Monash University, Caulfield East, Victoria 3145, Australia. Email: jiti.gao@monash.edu

1 Introduction

Econometric model estimation for nonlinear structural cointegrating models is an increasingly active area in recent years. Recent literature includes Newey et al. (1999) for nonparametric estimation of triangular simultaneous equations models; Karlsen and Tjøstheim (2001) for a nonparametric kernel estimation for a class of null–recurrent time series; Ai and Chen (2003) for sieve estimation of models with unknown functions; Newey and Powell (2003) for instrumental variable estimation of nonparametric models; Karlsen et al. (2007) for nonparametric kernel estimation in a null–recurrent cointegration model; Wang and Phillips (2009a) for estimation in a nonparametric cointegration model; Wang and Phillips (2009b) for estimation in structural nonparametric cointegrating model with a type of endogeneity; Gao and Phillips (2013a) for estimation in a class of non– and semi–parametric cointegration models; and Gao and Phillips (2013b) for semiparametric estimation in triangular system equations of nonstationary time series. To avoid possible model misspecification issues, such nonparametric or semiparametric estimation method is normally used as the first step to suggest a possible parametric approximation when there is no prior information about which particular model should be used for a given set of data. Then, a rigorous specification procedure should be used to test whether the suggested parametric model may be accepted statistically. This therefore introduces the recent literature about nonparametric specification testing for parametric specification of nonlinear and nonstationary time series.

Closely related to the specification testing problem considered in this paper, there are some existing papers for studying specification testing issues in regression models with nonstationary regressors (mainly integrated processes). To name a few, see Gao et al. (2009a,b), Hong and Phillips (2010) and Wang and Phillips (2012a), among others. Note that Gao et al. (2009a) propose a nonparametric kernel–based test for parametric specification of nonstationary nonlinear models where the regressor that is an integrated time series is independent of the equation error term that is a martingale difference sequence; Gao et al. (2009b) consider a testing issue on a nonstationary nonlinear autoregression models where strict conditions are imposed on the density of the error term; Hong and Phillips (2010) test linearity of cointegrating relations with an application; and Wang and Phillips (2012a) investigate parametric specification of nonstationary nonlinear regression models where the regression function is imposed to have a certain growth rate when its variable tends to infinity such as polynomials, power functions, etc. (see Wang and Phillips, 2012a, Assumption 4, p. 730). In addition, Wang and Phillips (2012a, p. 731) point out that: "It seems that cases with integrable $f(x, \theta)$ require different techniques and these are left for future investigation". Wang and Phillips (2012a, p. 731) also point out that "It is unclear at the moment if the results of the present paper on testing extend to the more general error structure considered in Wang and Phillips (2009b), but simulation results suggest that this may be so."

Indeed, there is a gap in the literature that the specification testing for nonlinear integrable regression functions still remains unsolved. Meanwhile, existing tests mentioned above are not applicable to the case where there is a type of endogeneity. From a practical point of view, there is need for a proposal of such a test that accommodates a type of endogeneity, as it is inherited in the United States aggregate consumers' consumption expenditure and disposable income over the period of 1960–2009. This paper therefore aims at proposing a test statistic for the specification of integrable nonstationary regression models, in which two forms of endogeneity are allowed. Usually, an instrumental variable approach is needed to deal with such endogeneity issue, although it is difficult to find such instrumental variables in the nonlinear and nonstationary situation. As shown in Section 3 below, without introducing any instrumental variable, it is shown that each of the proposed tests has a simple and known limiting distribution under the null hypothesis, and is also asymptotically consistent under a sequence of local alternatives.

Consider a structural nonparametric cointegration model of the form

$$\begin{aligned} y_t =& m(x_t) + e_t, \\ x_t =& x_{t-1} + u_t, \\ E[u_t] =& E[e_t] = 0 \end{aligned} \tag{1.1}$$

for $t = 1, \cdots, n$, where n is the sample size, $m(\cdot)$ is an unknown function, $x_0 = O_P(1)$, and u_t and e_t are endogenous each other as defined in Assumption A below.

We are interested in testing the following hypotheses

$$H_0: \quad P(m(x_t) = g(x_t; \theta_0)) = 1 \quad \text{versus} \quad H_1: \quad P(m(x_t) = g(x_t; \theta_1) + \triangle_n(x_t)) = 1 \tag{1.2}$$

for all $t = 1, \cdots, n$, where $g(x, \cdot)$ is a known integrable function on $\mathbb{R}$, $\theta_0, \theta_1 \in \Theta$ and $\Theta \subset \mathbb{R}^d$ is a parameter space, and $\{\triangle_n(x)\}$, the so-called local departure, is a sequence of unknown functions satisfying $\lim_{n\to\infty} \triangle_n(x) = 0$ for every $x \in \mathbb{R}$.

To broaden our research, the models in the form of (1.1) are allowed to accommodate a certain endogeneity in two possible forms: Type (i) both $u_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ and $e_t = \sum_{j=0}^{\infty} \phi_j \epsilon_{t-j}$ are linear processes driven by the same innovations $\{\epsilon_i : -\infty < i < \infty\}$; and Type (ii) $u_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ is a linear process, while $e_t = \wp(\epsilon_t, \cdots, \epsilon_{t-m_0+1}; \eta_t, \cdots \eta_{t-m_1+1})$ is a functional of some innovations drawn from (ϵ_i, η_j). There are considerably possibilities of e_t and u_t within these settings, depending on the choice of the coefficients in the linear processes and functional forms. Two extreme cases are (a) $e_t = u_t$ and (b) e_t and u_t are independent. To the best of our knowledge, neither nonparametric estimation nor nonparametric specification testing have accommodated Type (i) endogeneity. Type (ii) endogeneity has been considered in Wang and Phillips (2009b) for nonparametric estimation purposes. To the best of our knowledge, the proposed tests of this paper are probably the first ones available in the nonparametric specification testing literature that take both endogeneity and nonstationarity into account. These settings considerably facilitate the applications of our specification testing method and theory in practice. As an application, we are able in

Section 5 below to naturally deal with model specification testing for the relationship between the consumption and income data, which exhibit a type of endogeneity.

Note that under the null of (1.2) we have $m(x) \in L^2(\mathbb{R})$. The proposed test statistic below basically measures the total difference of $y_t - g(x_t, \widehat{\theta})$ at each observation point with an orthonormal basis function in $L^2(\mathbb{R})$ being a weight, where $\widehat{\theta}$ is a consistent estimator of θ_0 under H_0. Then, the approach is extended to testing non-integrable models (i.e. the regression function is non-integrable), including linear model as a special case. The main idea on the second test is to exploit a simple transformation of the form $M(x) = m(x)\exp(-x^2/2)$ such that $M(x)$ is integrable in $\mathbb{R}$ even though $m(x)$ itself is not integrable. As a result, a large class of functions, such as polynomials and power functions, can be allowed for $m(x)$. These tests clearly cover the existing papers such as Gao et al. (2009a) and Wang and Phillips (2012a) as a subclass and fill the gap in the literature.

More interestingly, we notice that in H_1 of (1.2), if $\theta_1 \neq \theta_0$ there would be a global departure incurred, namely, $m(x_t) = g(x_t; \theta_0) + [g(x_t; \theta_1) - g(x_t; \theta_0)] + \triangle_n(x_t)$. The name 'global departure' for $g(x_t; \theta_1) - g(x_t; \theta_0)$ is coined as the function is free of the sample size. Given that the function $g(x; \theta_1) - g(x; \theta_0)$ has support with positive Lebesgue measure, the global departure $g(x_t; \theta_1) - g(x_t; \theta_0)$ will dominate the local one $\triangle_n(x_t)$, while when $g(x; \theta_1) - g(x; \theta_0)$ is a zero function in the space $L^2(\mathbb{R})$ or if $\theta_1 = \theta_0$, the local departure $\triangle_n(x_t)$ secures the proposed tests consistent. See the proof of Theorem 3.2 below. This fact has been overlooked by the literature.

Our experience shows that while each of the tests is of a simple quadratic form, it is not necessarily easy to establish and then prove an asymptotic distribution for each of the test statistics. This is mainly because existing central limit theorems available for standardised versions of quadratic forms (see, e.g., Theorem A.1 of Gao, 2007) are not applicable in our case where there is no martingale structure involved in model (1.1). For our test statistics, as shown in Section 3 below, asymptotic distributions are established without imposing any martingale structure and introducing an instrumental variable approach. In equations (3.4)–(3.7) in Section 3 below, we shall discuss this key feature and the relationship between the stationary case and the cointegrated time series case considered in this paper.

The main contributions of the paper are summarised as follows.

- This paper proposes a new test driven by an orthogonal series for specification testing of integrable functions of nonstationary time series. This test is then extended to non-integrable functions of nonstationary time series.

- The proposed tests are applicable to such models that encounter endogeneity. One of the proposed tests is then naturally applied in an empirical example.

- While the limiting distributions of the test statistics look simple, both the establishment and the proof are not trivial.

The rest of the paper is organised as follows. Section 2 gives some preliminaries about Hermite orthogonal polynomial system as well as assumptions for model (1.1). Our specification test for the case where $m(x)$ is integrable in $\mathbb{R}$ is proposed and then studied in Section 3.1. Section 3.2 discusses the non–integrable case. Section 4 investigates the finite–sample performance of the proposed tests. Section 5 analyses a set of data for the US aggregate consumption expenditure and disposable income and then examines possible empirical models. Section 6 concludes the main parts of this paper with some concluding remarks. Appendix A gives some useful technical lemmas. Appendix B gives the full proof of Theorem 3.2 and then outlines the proofs of Theorem 3.1 and 3.3–3.5 and Lemma 3.1. All detailed proofs of the lemmas and theorems (except Theorem 3.2) are available from Appendices C and D of the supplemental material.

Throughout the paper the following notation is used. $\|\cdot\|$ stands for Euclidean norm for vector and element-wise norm for matrix. For example, $\|A\| = (\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}^2)^{1/2}$ if $A = (a_{ij})_{nm}$. $(\mathcal{F}_j, j \in \mathbb{Z})$ is an information flow where $\mathcal{F}_j = \sigma(\cdots, \epsilon_{j-1}, \epsilon_j)$ and $\{\epsilon_j, j \in \mathbb{Z}\}$ is an independent and identically distributed sequence with mean zero and variance σ^2 (iid(0, σ^2)). $[a]$ is the maximum integer not exceeding a. $\mathbb{R}$ is the real line and $\mathbb{R}^+$ is the set of all positive real numbers. "$\to_P$" and "$\to_D$" signify convergence in probability and convergence in distribution, respectively.

2 Preliminaries and assumptions

In order to study unknown functions in $L^2(\mathbb{R})$, we first introduce an orthonormal basis in the space. Let $\{H_i(x)\}$ be the Hermite polynomial system orthogonal with respect to $\exp(-x^2)$. As a classical polynomial orthogonal system, $\{H_i(x)\}$ is a complete orthogonal system in the Hilbert space $L^2(\mathbb{R}, \exp(-x^2))$. In addition, the orthogonality of the system is expressed as the equation $\int H_i(x)H_j(x)\exp(-x^2)dx = \sqrt{\pi}2^i i!\delta_{ij}$, where δ_{ij} is the Kronecker delta. Denote $\varphi(x) = \exp(-\frac{1}{2}x^2)$ and define

$$\mathscr{H}_i(x) = \frac{1}{\sqrt[4]{\pi}\sqrt{2^i i!}} H_i(x)\varphi(x), \quad i \geq 0. \tag{2.1}$$

Note that $\mathscr{H}_i(x)$ are the so-called Hermite functions in the literature. Moreover, $\{\mathscr{H}_i(x)\}$ is uniformly bounded in both i and x. Consequently, $\sum_{i=0}^{k-1}\mathscr{H}_i^2(x) \leq O(1)k$ uniformly in x. In addition, the system is a complete orthonormal basis in $L^2(\mathbb{R})$ in which the inner product is defined as $(f, g) = \int f(x)g(x)dx$. Therefore, any function $g(x) \in L^2(\mathbb{R})$ admits an orthogonal expansion of the form $g(x) = \sum_{i=0}^{\infty}\beta_i\mathscr{H}_i(x)$ with $\beta_i = \int g(x)\mathscr{H}_i(x)dx$. Throughout, let k be a positive integer and define $g_k(x) = \sum_{i=0}^{k-1}\beta_i\mathscr{H}_i(x)$ to be the truncation series with truncation parameter k. It is known that the convergence of $g_k(x) \to g(x)$ on the real line relies on the property of $g(x)$.

The following assumptions are made for the subsequent theoretical development.

Assumption A Let $\{\epsilon_j, j \in \mathbb{Z}\}$ be a sequence of independent and identically distributed (iid) continuous random variables satisfying $E\epsilon_0 = 0$, $E\epsilon_0^2 = \sigma^2$ and $E\epsilon_0^4 < \infty$. Suppose further that the characteristic function of ϵ_0 is integrable, i.e. $\int |Ee^{it\epsilon_0}|dt < \infty$.

(a) Suppose that $\{u_t\}$ is a linear process defined by $u_t = \Psi(L;\rho_0)\epsilon_t = \sum_{j=0}^{\infty}\psi_j(\rho_0)\epsilon_{t-j}$, where L is the lag operator, $\psi_j(\rho)$ is known up to the unknown parameter vector $\rho_0 = (\alpha_0,\lambda_0)'$ with $\lambda_0 > 3/2$ and $\alpha_0 \neq 0$, the coefficient $\psi_0 = 1$ and $\lim_{j\to\infty}\psi_j(\rho_0)j^{\lambda_0}$ exists, and $\psi := \sum_{j=0}^{\infty}\psi_j(\rho_0) \neq 0$. In addition, when the parameter ρ varies, $\psi_j(\rho)$ is continuous at ρ for each j.

    For $t \geq 1$, let $x_t = x_{t-1} + u_t$ with $x_0 = O_P(1)$.

(b) Error process $\{e_t\}$ is generated either by

    (i) $e_t = \sum_{j=0}^{\infty}\phi_j\epsilon_{t-j}$, where $\phi_0 = 1$, $\lim_{j\to\infty}\phi_j j^{\gamma_0}$ exists with $\gamma_0 > 3/2$, and $\phi := \sum_{j=0}^{\infty}\phi_j \neq 0$, or by

    (ii) $e_t = \wp(\epsilon_t,\cdots,\epsilon_{t-m_0+1};\eta_t,\cdots\eta_{t-m_1+1})$ where $\min(m_0,m_1) \geq 1$, $\{\eta_t\}$ is a sequence of iid(0,1) continuous variables which is sequentially independent of $\{\epsilon_t\}$ and the function $\wp(\cdots)$ is a measurable mapping from $\mathbb{R}^{m_0+m_1} \mapsto \mathbb{R}$ such that $E[e_t] = 0$ and $E[e_t^4] < \infty$ for all $t > \max(m_0,m_1)$; we define $e_t = 0$ for $t \leq \max(m_0,m_1)$.

Remark 2.1. This assumption exhibits the structure of regressor and equilibrium error process in model (1.1) with building blocks ϵ_j and η_j. Here, x_t is integrated by u_i, viz, $\triangle x_t = u_t$, while u_t and e_t contain the same iid(0,1) sequence $\{\epsilon_j\}$. This setup gives the maximum possibility to accommodate both endogeneity and exogeneity in our model, depending on the choice of coefficients in linear processes and/or the form of $\wp$ function. For instance, if $\phi_j = \psi_j$ for each j, $e_t = u_t$ implying the highest endogeity in the context; when the functional $\wp$ does not include any ϵ_j, e_t is independent of u_s and therefore of x_t, giving the independence case.

In order to estimate the parameter in ψ_j later, we need to specify it in the assumption. However, this is not necessary for ϕ_j. The limits of $\psi_j(\rho_0)j^{\lambda_0}$ and $\phi_j j^{\gamma_0}$ could be zero, implying a faster decay rate of the coefficients, whereas nonzero limits guarantee the rates $O(j^{-\lambda_0})$ and $O(j^{-\gamma_0})$ for the coefficients in the processes, respectively. More importantly, such condition ensures that u_t is an invertible short-memory process so that ϵ_t may be represented via u_t for $t \geq 1$.

Under Assumption A we have $\sum_{j=0}^{\infty}|\psi_j| < \infty$ and $\sum_{j=0}^{\infty}|\phi_j| < \infty$. It follows from Wang et al. (2003) that, as $n \to \infty$

$$d_n^2 = E[x_n^2] = \psi^2 n\ (1 + o(1)). \tag{2.2}$$

In addition, from Wang and Phillips (2009a) we have

$$x_{nt} := \frac{1}{d_n}x_t \to_D B(r) \quad \text{for } r \in [0,1], \tag{2.3}$$

where $B(r)$ is a standard Brownian motion on [0,1]. Note also that $\frac{1}{d_t}x_t$ have densities $f_t(x)$ that are uniformly bounded in both t and x, and given $\mathcal{F}_s$, $\frac{1}{d_{ts}}(x_t - x_s)$ with $d_{ts} = |\psi|\sqrt{t-s}$ for $t > s$

have density functions $f_{ts}(x)$ that are uniformly bounded over all t, s and $x \in \mathbb{R}$. Such results can be found in Remarks 3.1 in Wang and Phillips (2009a, p. 321).

Assumption B

(a) Under H_0, there exists a consistent estimator $\widehat{\theta}$ of θ_0 such that $\|\widehat{\theta} - \theta_0\| = O_P(\zeta_n)$ where $\zeta_n n^{1/4} = O(1)$.

(b) Suppose that $g(x;\theta)$ is twicely differentiable with respect to θ and $g(x;\theta) \in L^2(\mathbb{R})$ for every fixed $\theta \in \Theta$. Denote that $l_1(x,\theta) := \frac{\partial}{\partial\theta} g(x;\theta)$ and $l_2(x,\theta) := \frac{\partial^2}{\partial\theta\partial\theta'} g(x;\theta)$. Suppose further that $\|l_1(x,\cdot)\|, \|l_2(x,\cdot)\| \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, and that there exists a positive function $l(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ such that $\|l_2(x,\cdot)\| \leq l(x)$.

Remark 2.2. Condition (a) stipulates a certain convergence rate on $\widehat{\theta} \to_P \theta_0$, where the convergence rate of $\zeta_n \to 0$ is at least and possibly faster than $O(n^{-1/4})$. It is known that (a) is achievable in particular cases such as (i) e_t is independent of u_t, and (ii) $(e_t, \mathcal{F}_t)$ is a martingale difference sequence for a given information flow $\mathcal{F}_t$ such that x_t is adapted with respect to $\mathcal{F}_{t-1}$, since there already exist such results in Park and Phillips (1999, 2001). Indeed, if $g(x,\theta_0)$ is an I–regular regression function (for more detail about I–regular functions, see Park and Phillips 2001), we then have $\sqrt[4]{n}(\widehat{\theta} - \theta) = O_P(1)$ in aforementioned two circumstances. We provide another simple example in the supplemental material for this paper.

Condition (b) is a set of standard requirements used in the similar situations (see, for example, Assumption 2.4 of Gao et al. (2011) and Assumption 4 of Wang and Phillips (2012a)). It is also easily seen that the following type of functions for $g(x;\theta)$ satisfy these conditions: (1) $g(x;\theta) = \theta g_1(x)$, where $g_1(x)$ is integrable and $\theta \in \mathbb{R}$ (so called linear-in-parameter regression function in Park and Phillips 2001); (2) $g(x;\theta) = \frac{1}{1+\theta x^2}$ where $\theta > c > 0$; (3) $g(x;\alpha,\beta) = \alpha \exp(-\beta x^2)$ where $\theta = (\alpha,\beta) \in \Theta \subset \mathbb{R} \times \mathbb{R}^+$.

Assumption C Let $\triangle_n(x) = \delta_n \triangle(x)$ satisfying the following conditions:

(a) $\triangle(x)$ is a bounded and integrable nonzero function.

(b) Let $\delta_n \to 0$ and $\delta_n^2 \sqrt{n}/k \to \infty$ as $n \to \infty$, where $k = [c \cdot n^\kappa]$ for some $\kappa : 0 < \kappa < \frac{1}{2}$ and constant $c > 0$.

Remark 2.3. Exclusion of $\triangle(x)$ to be zero function is trivial. This assumption on one hand guarantees our test statistic consistent, and on the other hand keeps the local departures as small as possible. Indeed, the convergence of $\delta_n \to 0$ and the boundedness and integrability of $\triangle(x)$ force the departures $\triangle_n(x) = \delta_n \triangle(x)$ smaller and smaller as the sample size increases, and $\delta_n^2 \sqrt{n}/k \to \infty$ is the only requirement that enables test statistic introduced later to detect whether there exists a departure from the null. To illustrate this idea, in the finite sample simulation below we choose that δ_n decays so fast that $\delta_n^2 \sqrt{n}/k = \frac{1}{4} \ln(n)$.

3 Specification testing

In this section, we first consider the case where $m(x)$ is integrable and then the case where $m(x)$ is non-integrable.

3.1 Integrable case

Suppose $m(\cdot) \in L^2(\mathbb{R})$. Then, $m(x)$ admits an orthogonal expansion $m(x) = \sum_{i=0}^{\infty} \beta_i \mathscr{H}_i(x)$. Let k be a positive integer and then define a truncated series of the form $m_k(x) = \sum_{i=0}^{k-1} \beta_i \mathscr{H}_i(x)$ and residue $\gamma_k(x) = \sum_{i=k}^{\infty} \beta_i \mathscr{H}_i(x)$ for any real x. Denote $Z(x) = (\mathscr{H}_0(x), \cdots, \mathscr{H}_{k-1}(x))'$ a vector of the first k functions in the basis and $\beta = (\beta_0, \cdots, \beta_{k-1})'$ a vector of the first k coefficients in the expansion. Then, $m_k(x) = Z(x)'\beta$ and model (1.1) under H_0 can be written as $y_t = Z(x_t)'\beta + \gamma_k(x_t) + e_t$, for $t = 1, \cdots, n$. Denote $Y = (y_1, \cdots, y_n)'$, $Z = (Z(x_1), \cdots, Z(x_n))'$ an $n \times k$ matrix, $e = (e_1, \cdots, e_n)'$ and $\gamma = (\gamma_k(x_1), \cdots, \gamma_k(x_n))'$. The equations are formulated into the following matrix form:

$$Y = Z\beta + \gamma + e. \tag{3.1}$$

Accordingly, we have the ordinary least squares (OLS) estimator of β given by $\widehat{\beta} = (Z'Z)^{-1}Z'Y$. Thereby, $\widehat{m}(x) = Z(x)'\widehat{\beta}$ is defined as an estimate for $m(x)$ at $\forall\ x \in \mathbb{R}$.

For the purpose of constructing our test, we may avoid involving the inverse of $Z'Z$ in $\widehat{\beta}$ by defining $\widetilde{\beta} = Z'Z\widehat{\beta}$. Correspondingly, the estimator $\widehat{m}(x) = Z(x)'\widehat{\beta}$ can be replaced by $\widetilde{m}(x) = Z(x)'\widetilde{\beta}$ for testing purposes and, invoking the expression of $\widehat{\beta}$,

$$\widetilde{m}(x) = \sum_{t=1}^{n} [Z(x_t)'Z(x)] y_t. \tag{3.2}$$

Analogously, we may also have a similar version, $\widetilde{g}(x;\theta)$, for $g(x;\theta)$ of the form:

$$\widetilde{g}(x;\theta) = \sum_{t=1}^{n} [Z(x_t)'Z(x)] g(x_t;\theta). \tag{3.3}$$

Let $\widehat{\theta}$ be a consistent estimator of θ_0 under $H_0:\ y_t = g(x_t;\theta_0) + e_t$. Instead of comparing $\widehat{m}(x)$ with $g(x, \widehat{\theta})$, we measure the distance between $\widetilde{m}(x)$ and $\widetilde{g}(x;\widehat{\theta})$ and then propose a test statistic of the form:

$$L_n = \int_{-\infty}^{\infty} \left(\widetilde{m}(x) - \widetilde{g}(x;\widehat{\theta})\right)^2 dx = \int_{-\infty}^{\infty} \left(\sum_{t=1}^{n} [Z(x_t)'Z(x)](y_t - g(x_t;\widehat{\theta}))\right)^2 dx, \tag{3.4}$$

which, by virtue of the orthogonality of the basis, simplifies to

$$L_n = \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\widehat{e}_t\widehat{e}_s, \tag{3.5}$$

where $\widehat{e}_t := y_t - g(x_t;\widehat{\theta})$.

Clearly, L_n is interpreted as the distance in the space $L^2(\mathbb{R})$ between the proxies of $m(x)$ and $g(x, \widehat{\theta})$ based on the observation (y_t, x_t). Moreover, under the null $\widehat{e}_t = e_t + [g(x_t; \theta_0) - g(x_t; \widehat{\theta})]$ consists of two errors such that, as shown in Theorem 3.1 below, a suitably normalised version of L_n converges to a local–time random variable, while under a class of local alternatives, the suitably normalised version of L_n diverges to ∞, because of the existence of the local departure.

Meanwhile, we have

$$L_n = \sum_{t=1}^{n} \|Z(x_t)\|^2 \widehat{e}_t^2 + 2\sum_{t=2}^{n}\sum_{s=1}^{t-1} Z(x_t)'Z(x_s)\widehat{e}_t\widehat{e}_s \equiv L_{an} + L_{bn}. \tag{3.6}$$

In the case where x_t is stationary, as shown in Gao et al. (2002), a suitably standardised version of L_{an} converges in probability to a non–stochastic quantity. Thus, one has to use a standardised version of the form

$$\widehat{L}_{bn} = \frac{\sum_{t=2}^{n}\sum_{s=1}^{t-1} \widehat{e}_s Z(x_t)'Z(x_s)\widehat{e}_t}{\sqrt{\sum_{t=2}^{n}\sum_{s=1}^{t-1} \widehat{e}_s^2 \left(Z(x_t)'Z(x_s)\right)^2 \widehat{e}_t^2}} \tag{3.7}$$

to derive an asymptotically consistent test. As shown in the proof of Theorem 3.1 in Appendix B below for the nonstationary case, by comparison, L_{an} is the leading term that already converges in distribution to a local–time random variable while L_{bn} becomes asymptotically negligible compared with L_{an}. Therefore, there is no need to standardise L_{bn} as in the stationary case. As a consequence, both the size and power properties of L_n become stable and robust. These, along with the large and small–sample properties discussed in Sections 3–5 below, are the key features of L_n.

Another comparison with a nonparametric kernel test of the form

$$\widetilde{L}_n = \frac{\sum_{t=2}^{n}\sum_{s=1}^{t-1} \widehat{e}_s K\left(\frac{x_t - x_s}{h}\right)\widehat{e}_t}{\sqrt{\sum_{t=2}^{n}\sum_{s=1}^{t-1} \widehat{e}_s^2 K^2\left(\frac{x_t - x_s}{h}\right)\widehat{e}_t^2}} \tag{3.8}$$

proposed originally for the stationary time series case and then adopted in Gao et al. (2009a), Gao et al. (2009b) and Wang and Phillips (2012a) (with $K(\cdot)$ being a probability kernel function and h being a bandwidth parameter) for the nonstationary time series case, a normalised version of L_n of the form $\frac{\widehat{d}_n}{nk\,\widehat{\sigma}_e^2} L_n$ (as defined in Theorem 3.3 below) does not involve any random denominator of a quadratic form of nonstationary time series. As a consequence, the simple form of L_n of (3.6) makes it possible to establish Theorems 3.1–3.3 below under two types of endogeneity.

Theorem 3.1. *Suppose that Assumptions A and B hold. Under H_0, we have as $n \to \infty$*

$$\frac{d_n}{nk\,\sigma_e^2} L_n \to_D L_B(1, 0), \tag{3.9}$$

where d_n is defined by (2.2), $\sigma_e^2 = E[e_1^2]$ and $L_B(1, 0)$ is the local–time random variable of the standard Brownian motion $B(r)$ in (2.3).

Remark 3.1. It follows from Theorem 3.1 that, at the significance level α, we may reject H_0 if $\frac{d_n}{nk\sigma_e^2} L_n > l_{1-\alpha}$ with $l_{1-\alpha}$ satisfying $P(L_B(1, 0) > l_{1-\alpha}) = \alpha$, which means that the proposed test statistic has size α. For the definition and properties of the local–time processes, Revuz and Yor (1999) is a standard reference.

Theorem 3.2. *Suppose that Assumptions A–C hold. Under H_1, we have as $n \to \infty$*

$$\frac{d_n}{nk\ \sigma_e^2} L_n \to_P \infty. \tag{3.10}$$

Remark 3.2. Theorem 3.2 shows that the test statistic has nontrivial power against a sequence of local alternatives that may have a distance from the null approaching zero at a rate slower than $(\sqrt{n}/k)^{-1/2}$, as required in Assumption C for δ_n. This is similar to the stationary case discussed in Gao (Chapter 3, 2007). With the results in Theorem 3.1 and 3.2, we can differentiate the models underpinning by H_0 and H_1, respectively. We shall reject H_0 in favour of H_1 if $\frac{d_n}{nk} L_n$ is large enough.

It is noteworthy that there are two possible expressions for $\widehat{e}_t$ under H_1. If $\theta_1 = \theta_0$, $\widehat{e}_t = e_t + [g(x_t;\theta_0) - g(x_t;\widehat{\theta})] + \triangle_n(x_t)$. The divergence of the normalisation of L_n is due to the inclusion of the local departure $\triangle_n(x_t)$, although the function $\triangle_n(x)$ attenuates to zero uniformly in x when the sample size increases. If $\theta_1 \neq \theta_0$, we have $\widehat{e}_t = e_t + [g(x_t;\theta_0) - g(x_t;\widehat{\theta})] + \triangle_n(x_t) + [g(x_t;\theta_1) - g(x_t;\theta_0)]$. This means, except for the local depature $\triangle_n(x_t)$, we also have a global departure $g(x_t;\theta_1) - g(x_t;\theta_0)$. Intuitively, the global departure would dominate the local one (this will be shown in the proof of the theorem). Thus, the normalized L_n also diverges to infinity with probability approaching 1.

Note also that there are two nuisance parameters involved in the large sample theory of the proposed test, namely, ψ in d_n and σ_e^2, which should be replaced by their consistent estimates. Towards this end, write $\varepsilon_t = \Psi^{-1}(L;\rho_0)u_t = \Psi^{-1}(L;\rho_0)\triangle x_t$ for $t \geq 1$ (see Assumption A). For any admissible $\rho = (\alpha, \lambda)'$, define

$$\varepsilon_t(\rho) = \Psi^{-1}(L;\rho)u_t. \tag{3.11}$$

For a given user-chosen optimizing compact set $\mathbb{S}$, define

$$\widehat{\rho} = \operatorname*{argmin}_{\rho\in\mathbb{S}} \frac{1}{n}\sum_{t=1}^{n} \varepsilon_t^2(\rho) \tag{3.12}$$

as an estimator of ρ_0.

Lemma 3.1. *Define $\widehat{\psi} = \psi(\widehat{\rho}) = \sum_{j=0}^{\infty} \psi_j(\widehat{\rho})$, $\widehat{d}_n = |\widehat{\psi}|\sqrt{n}$ and $\widehat{\sigma}_e^2 = \frac{1}{n}\sum_{t=1}^{n} \widehat{e}_t^2$ with $\widehat{e}_t = y_t - g(x_t, \widehat{\theta})$. Then, under Assumption A and H_0, as $n \to \infty$*

$$\widehat{\psi} \to_P \psi, \quad \frac{\widehat{d}_n}{d_n} \to_P 1, \quad \textit{and} \quad \widehat{\sigma}_e^2 \to_P \sigma_e^2. \tag{3.13}$$

The proofs of Lemma 3.1 and Theorem 3.3 below are given in Appendix B.

Theorem 3.3. *When ψ and σ_e^2 are both unknown, under H_0 and Assumptions A and B, we have*

$$\frac{\widehat{d}_n}{nk\ \widehat{\sigma}_e^2} L_n \to_D L_B(1,0), \tag{3.14}$$

where $\widehat{d}_n$ and $\widehat{\sigma}_e^2$ are defined in Lemma 3.1. Under H_1 and Assumptions A–C, as $n \to \infty$,

$$\frac{\widehat{d}_n}{nk\ \widehat{\sigma}_e^2} L_n \to_P \infty. \tag{3.15}$$

Note that the quantity $\widehat{M_n} \equiv \frac{\widehat{d}_n}{nk\widehat{\sigma}_e^2} L_n$ not only has nice large sample properties as stated in Theorem 3.3, but also $\widehat{M_n}$ is now completely computable in practice.

3.2 Beyond integrability

We are about to relax the restriction for $m(x)$ in the last subsection. Our motivation is on the following observation, if $m(x)$ satisfies $\int m^2(x)e^{-x^2}dx < \infty$, then $m(x)\exp(-x^2/2) \in L^2(\mathbb{R})$. As a result, after a simple transformation we may borrow L_n to test models that have non-integrable regression functions.

Given observations $\{(x_t, y_t), t = 1, 2, \cdots, n\}$ from model (1.1), we are interested in testing in model (1.1) that

$$H_{10}:\quad P(m(x_t) = g(x_t;\theta_0)) = 1 \quad \text{versus} \quad H_{11}:\quad P(m(x_t) = g(x_t;\theta_1) + \triangle_n(x_t)) = 1 \tag{3.16}$$

for all $t = 1, \cdots, n$, where for any fixed $\theta \in \Theta \subset \mathbb{R}^d$, $g(x,\theta) \in L^2(\mathbb{R}, \exp(-x^2))$ is a known function, and $\triangle_n(x)$ is the same as in Assumption C.

Note that under H_{10}, model $y_t = g(x_t;\theta_0)+e_t$ has regression function $g(x;\theta) \in L^2(\mathbb{R}, \exp(-x^2))$. In order to make use of L_n, multiplying both sides by $\varphi(x_t) := \exp(-\frac{1}{2}x_t^2)$ will give us an integrable model. Specifically, by denoting $Y_t = y_t\varphi(x_t)$, $M(x_t) = \varphi(x_t)m(x_t)$ and $\varepsilon_t = \varphi(x_t)e_t$, we have

$$Y_t = M(x_t) + \varepsilon_t, \qquad t = 1, \cdots, n. \tag{3.17}$$

Moreover, Hypothesis H_{10} is completely equivalent to $H_{10}' :\ P(M(x_t) = G(x_t;\theta_0)) = 1$ in which $G(x_t;\theta) = g(x_t,\theta)\varphi(x_t)$ and Hypothesis H_{11} is completely equivalent with $H_{11}' :\ P(M(x_t) = G(x_t;\theta_1) + \triangle_n(x_t)\varphi(x_t)) = 1$. Thereinto, $M(x)$ is an unknown integrable function, while $G(x,\theta)$ is a known integrable function. Roughly speaking, we have transformed the dataset from 'non–integrable' (y_t, x_t) to 'integrable' (Y_t, x_t). Notice also that the deviation function $\triangle_n(x)\varphi(x)$ in H_{11}' has the same property in the purpose of testing as $\triangle_n(x)$ in H_{11}, that is, boundedness and integrability. After transformation, we are facing the same specification issue as in the last subsection.

Therefore, we propose the following test statistic for testing (3.16):

$$\Pi_n = \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\widehat{\varepsilon}_t\widehat{\varepsilon}_s, \tag{3.18}$$

where $\widehat{\varepsilon}_t := \widehat{e}_t\varphi(x_t) = (y_t - g(x_t;\widehat{\theta}))\varphi(x_t) = Y_t - G(x_t;\widehat{\theta})$ and $\widehat{\theta}$ is a consistent estimator of θ_0 under $H_{10}:\ y_t = g(x_t;\theta_0) + e_t$.

It is clear that $\Pi_n = \sum_{t=1}^n \sum_{s=1}^n Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)\widehat{e}_t\widehat{e}_s$ under H_{10}, while under H_{11}, $\Pi_n = \sum_{t=1}^n \sum_{s=1}^n Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)(\widehat{e}_t + \triangle_n(x_t))(\widehat{e}_s + \triangle_n(x_s))$, where $\widehat{e}_t = y_t - g(x_t;\widehat{\theta})$. Comparing with L_n, only factors $\varphi(x_t)\varphi(x_s)$ are adhered in each term. This is just the consequence of the transformation from $m(x_t)$ to $M(x_t)$.

Before stating the limit theory for Π_n, we need to give the following assumption which is the counterpart of Assumptions B for the case where m belongs to $L^2(\mathbb{R}, \exp(-x^2))$.

Assumption B*

(a) Under the null in (3.16), there exists a consistent estimator, $\widehat{\theta}$, of θ_0 such that $\|\widehat{\theta}-\theta_0\| = O_P(\zeta_n)$ as $n \to \infty$, where $\zeta_n n^{1/4} = O(1)$.

(b) Suppose that $g(x;\theta)$ is twicely differentiable with respect to θ. Let $l_1(x,\theta) := \frac{\partial}{\partial\theta} g(x;\theta)$ and $l_2(x,\theta) := \frac{\partial^2}{\partial\theta\partial\theta'} g(x;\theta)$. Suppose further that $\|l_1(x,\cdot)\|, \|l_2(x,\cdot)\| \in L^2(\mathbb{R}, e^{-x^2})$, while $\|l_2(x,\theta)\| \le l(x)$ uniformly over θ with positive function $l(x) \in L^2(\mathbb{R}, e^{-x^2})$.

Remark 3.3. There are two differences between Assumptions B* and B. One is the estimation of θ_0 is under H_{10} in (3.16). However, a faster convergence rate is not required in (a), although in existing literature the convergence rate for non-integrable regression function is much higher than $1/\sqrt[4]{n}$. This is an advantage of the proposed method, because non-integrable $m(x)$ has been transformed to an integrable function $M(x)$, so that Π_n works in the same environment as L_n does. Thus, all regression functions studied in the literature easily satisfy this assumption, for example, for functions like polynomials, power functions, $\theta \log|x|$, $\theta e^x/(1+e^x)$ and $x(1+\theta x)^{-1}1(x \ge 0)$, the rates of convergence are at least equal to even faster than $1/\sqrt{n}$. A detailed discussion can be found in Theorems 5.2 and 5.3 of Park and Phillips (2001, p. 135).

Another difference is the function space. As the regression function $m(x)$ is now in the class of $L^2(\mathbb{R}, e^{-x^2})$, we basically stipulate that its derivatives with respect to the parameter vector are also in $L^2(\mathbb{R}, e^{-x^2})$. Clearly, the examples given above satisfy this condition.

Theorem 3.4. *Suppose Assumptions A and B* hold. Under H_{10} of* (3.16)*, we have as $n \to \infty$*

$$\frac{d_n}{nk\ \sigma_e^2}\Pi_n \to_D \int \mathscr{T}(x)\varphi^2(x)dx \cdot L_B(1,0), \tag{3.19}$$

where d_n is given by (2.2)*, $\sigma_e^2 = E[e_1^2]$ and $\mathscr{T}(x) = \frac{1}{2\pi}\sqrt{4-x^2}$ on $x \in [-2,2]$ and 0 elsewhere, $\varphi(x) = \exp(-\frac{1}{2}x^2)$ and $L_B(1,0)$ is the local time process of Brownian motion $B(r)$.*

Theorem 3.5. *Let Assumptions A, B* and C hold. Under H_{11}, we have as $n \to \infty$*

$$\frac{d_n}{nk\ \sigma_e^2}\Pi_n \to_P \infty. \tag{3.20}$$

The proofs of Theorems 3.4 and 3.5 are given in Appendix B.

Remark 3.4. If ψ and σ_e^2 are unknown, we may have a similar corollary to Theorem 3.3. Due to the similarity, we omit it.

Both the proposed tests and the related theory will be evaluated using simulated and real data examples in Sections 4 and 5, respectively.

4 Finite sample experiment

Monte Carlo simulations are conducted in this section to investigate the performance of the proposed statistics L_n and Π_n in the finite-sample situations. To be in concert with Assumption A, two circumstances of generating data are considered:

I) Suppose $\{\epsilon_j, j \in \mathbb{Z}\}$ is a sequence of iid $N(0, \sigma^2)$. Let $e_t = \beta_0 e_{t-1} + \epsilon_t$ with $|\beta_0| < 1$, $u_t = \alpha_0 u_{t-1} + \epsilon_t$ with $|\alpha_0| < 1$ and $x_t = x_{t-1} + u_t$ with $x_0 = O_P(1)$.

II) Suppose $\{(\epsilon_j, \eta_j), j \in \mathbb{Z}\}$ is a sequence of iid $N(0, \sigma^2 I_2)$. Let $e_t = a\epsilon_t + b\eta_t$, $u_t = \alpha_0 u_{t-1} + \epsilon_t$ with $|\alpha_0| < 1$ and $x_t = x_{t-1} + u_t$ with $x_0 = O_P(1)$.

All constants in the above will be specified in the following tables. Note that in the first case if $\alpha_0 = \beta_0$, $e_t = u_t$ that reflects the highest endogeneity; in the second case when $a = 0$, $e_t = b\eta_t$ which is independent of any ϵ_j and hence of u_t, implying the independence between e_t and x_t.

Bootstrap Simulation Procedure: A bootstrap schedule is proposed below to generate critical values l_α^* for them where $\alpha = 1\%$ and 10%. The number of replications for Monte Carlo simulation is $M = 5000$ and that for the bootstrap procedure is $M_b = 250$.

Step 1 Let $\widehat{e}_t = y_t - g(x_t, \widehat{\theta})$, where $\widehat{\theta}$ is a consistent estimator of θ_0 based on the original sample (x_t, y_t) under the null hypothesis. Generate e_t^* as follows.

- For Case I) above, generate e_t^* by the conventional block bootstrap method (see, for example, Hall et al. (1995)) given as follows: Estimate β_0 by $\widehat{\beta}_0$ based on $\widehat{e}_t = \beta_0 \widehat{e}_{t-1} + \epsilon_t$ and then estimate σ^2 by $\widehat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^n \left(\widehat{e}_t - \widehat{\beta}_0 \widehat{e}_{t-1}\right)^2$.
  Let $l = \lfloor n^{\frac{1}{3}} \rfloor$ and choose integer λ such that $\lambda l = n$. Generate $e_{1l}^*(j) = [\widehat{e}_1(j), \cdots, \widehat{e}_l(j)]$, $\cdots$, $e_{Nl}^*(j) = \left[\widehat{e}_{(\lambda-1)l+1}(j), \cdots, \widehat{e}_{\lambda l}(j)\right]$ in step j for $N = n - l + 1$. Replicate the resample $J = 250$ times and obtain J bootstrap resamples $\{e_{sl}^*(j) : 1 \le s \le N; 1 \le j \le J\}$ and then take the average $e_{sl}^* = \frac{1}{J}\sum_{j=1}^J e_{sl}^*(j)$ to obtain a block bootstrap version of e_t of the form: $(e_1^*, \cdots, e_n^*) = (e_{1l}^*, \cdots, e_{Nl}^*)$.
- For Case II) above, generate e_t^* by a regression bootstrap method of the form $e_t^* = \widehat{e}_t \tau_t$ with τ_t being iid sequence possessing a two-point distribution of the form $P(\tau_1 = \mp(\sqrt{5} \mp 1)/2) = (\sqrt{5} \pm 1)/2\sqrt{5}$.

Step 2 Generate y_t^* by $y_t^* = g(x_t, \widehat{\theta}) + e_t^*$. Use the dateset (x_t, y_t^*) to re-estimate θ_0 and denote by $\widehat{\theta}^*$ the resulting estimate. Compute L_n^* that is a corresponding version of L_n with $\widehat{\theta}$ and $\{(x_t, y_t), 1 \le t \le n\}$ being replaced by $\widehat{\theta}^*$ and $\{(x_t, y_t^*), 1 \le t \le n\}$ in the expression of L_n.

Step 3 Repeat the above steps M_b times and thereby produce M_b versions of L_n^*, signified by $L_{n\ell}^*$ for $\ell = 1, \cdots, M_b$. Use the M_b values of $L_{n\ell}^*$ to construct their empirical bootstrap distribution function. The bootstrap distribution of $\frac{d_n}{nk}L_n^*$ given $\mathcal{W}_n = \{(x_t, y_t), 1 \le t \le n\}$ is defined by $P^*(\frac{d_n}{nk}L_n^* < x) = P(\frac{d_n}{nk}L_n^* < x|\mathcal{W}_n)$. Let l_α^* be the quantile such that $P^*(\frac{d_n}{nk}L_n^* \ge l_\alpha^*) = \alpha$ and then l_α^* is used to approximate l_α.

Step 4 Define the size and power functions by

$$\alpha = P\left(\frac{d_n}{nk}L_n \ge l_\alpha^*|H_0\right) \quad \text{and} \quad \beta = P\left(\frac{d_n}{nk}L_n \ge l_\alpha^*|H_1\right).$$

Sample sizes in the experiments are taken as $n = 200, 500$ and 1200, respectively. In addition, the truncation parameter is selected by $k = [c \cdot n^\kappa]$ with $\kappa = 1/5, 1/4.5, 1/4$ and $c = 2.2$, respectively. We choose $\triangle_n(x) = \delta_n \frac{1}{1+x^2}$, in which $\delta_n = \frac{1}{2}\sqrt{k \log(n)/\sqrt{n}}$. It is clear that we are about to keep both δ_n and $\triangle(x)$ very 'small' by choosing integrable function $\triangle(x) = \frac{1}{1+x^2}$ and the fast decay δ_n such that $\delta_n^2\sqrt{n}/k = \frac{1}{4}\log(n)$ diverges slowly. Note that the divergence of $\delta_n^2\sqrt{n}/k$ is a necessary condition for the proposed test statistics to be consistent, but we keep it very slow. We shall examine the finite–sample performance of L_n and Π_n with unknown ψ and σ_e^2 (see Theorem 3.3). That is, $\widehat{\sigma}_e^2$ is used in each experiment in the place of σ_e^2, but because ψ is only involved in x_t and we use bootstrap schedule to simulate the critical value, $\widehat{\psi}$ does not show up in the experiments.

Example 4.1. This example examines the test L_n using integrable regression models. The model for simulation is $y_t = m(x_t) + e_t$.

In situation I), the null hypothesis is H_0: $P(m(x_t) = 10\exp(-\theta_0 x_t^2)) = 1$ with $\theta_0 = 1$; and the alternative is H_1: $P(m(x_t) = 10\exp(-\theta_1 x_t^2) + \triangle_n(x_t)) = 1$ with $\theta_1 = 1$.

In situation II), the null hypothesis is H_0: $P(m(x_t) = (1 + x_t^2)\exp(-\theta_0 x_t^2)) = 1$ with $\theta_0 = 1$; and the alternative is H_1: $P(m(x_t) = (1 + x_t^2)\exp(-\theta_1 x_t^2) + \triangle_n(x_t)) = 1$ with $\theta_1 = 1$.

All results are reported in Tables 1-4. Thereinto, Tables 1 and 3 contain all sizes. It is readily seen that almost all sizes fluctuate reasonably around the given significance levels, although they are not perfectly performed. In Table 3 a slight less than half experiments are undersized, while most sizes in Table 1 are oversized. Note also that in Table 3, in spite of the independence between e_t and x_t due to $a = 0$, the sizes in the test perform similarly to that of the test with $a = 0.2$. Overall, the sizes become stable when the sample size increases.

Tables 2 and 4 accommodate all powers. Although the local departure function $\triangle_n(x) = \delta_n\triangle(x)$ is asymptotically negligible as δ_n approaches zero, the power values are very strong. It is clear that in most cases the power increases as either the sample size or κ involved in the truncation parameter increases. This is because the increase of the sample size or κ would bring more terms of the orthogonal basis into L_n and, as can be seen in the proof of Theorem 3.2, each extra term may make contribution to the power. Meanwhile, comparing the independent and dependent cases in Table 4, the endogeity does not attenuate the powers at all and, conversely, the powers still perform strong and robust in the experiments.

Table 1: Size: $m(x) = 10\exp(-\theta_0 x^2)$, $\theta_0 = 1$

n		Nominal size 1%			Nominal size 10%		
	$\kappa =$	1/5	1/4.5	1/4	1/5	1/4.5	1/4
				$\alpha_0 = -0.01$,	$\beta_0 = -0.01$		
200		0.0160	0.0126	0.0110	0.1222	0.1162	0.1090
500		0.0112	0.0140	0.0120	0.1162	0.1156	0.1046
1200		0.0110	0.0120	0.0106	0.1006	0.1002	0.0918
				$\alpha_0 = 0.01$,	$\beta_0 = 0.05$		
200		0.0162	0.0130	0.0140	0.1236	0.1176	0.1172
500		0.0136	0.0130	0.0130	0.1122	0.1052	0.1080
1200		0.0110	0.0098	0.0094	0.1030	0.1050	0.1010

$u_t = \alpha_0 u_{t-1} + \epsilon_t$, $e_t = \beta_0 e_{t-1} + \epsilon_t$ and $\sigma = 0.7$.

Table 2: Power: $m(x) = 10\exp(-\theta_1 x^2) + \triangle_n(x)$, $\theta_1 = 1$

n		Nominal size 1%			Nominal size 10%		
	$\kappa =$	1/5	1/4.5	1/4	1/5	1/4.5	1/4
				$\alpha_0 = -0.01$,	$\beta_0 = -0.01$		
200		0.7810	0.7906	0.8016	0.9040	0.9070	0.9126
500		0.8196	0.8300	0.8410	0.9156	0.9176	0.9210
1200		0.8544	0.8580	0.8660	0.9328	0.9320	0.9344
				$\alpha_0 = 0.01$,	$\beta_0 = 0.05$		
200		0.7840	0.7956	0.8070	0.9030	0.9047	0.9092
500		0.8112	0.8252	0.8326	0.9150	0.9170	0.9220
1200		0.8528	0.8588	0.8694	0.9318	0.9330	0.9380

$u_t = \alpha_0 u_{t-1} + \epsilon_t$, $e_t = \beta_0 e_{t-1} + \epsilon_t$ and $\sigma = 0.7$.

Table 3: Size: $m(x) = (1 + x^2)\exp(-\theta_0 x^2)$, $\theta_0 = 1$

n		Nominal size 1%			Nominal size 10%		
	$\kappa =$	1/5	1/4.5	1/4	1/5	1/4.5	1/4
		$\alpha_0 = 0.05$,	$a = 0.2$,	$b = 0.9$			
200		0.0146	0.0136	0.0130	0.1092	0.1052	0.1052
500		0.0124	0.0116	0.0118	0.0958	0.0956	0.0958
1200		0.0110	0.0104	0.0104	0.1022	0.1004	0.1008
		$\alpha_0 = 0.2$,	$a = 0$,	$b = 1$			
200		0.0094	0.0090	0.0080	0.1040	0.1058	0.1024
500		0.0114	0.0102	0.0082	0.0950	0.1012	0.0980
1200		0.0102	0.0098	0.0114	0.0988	0.0992	0.0984

$u_t = \alpha_0 u_{t-1} + \epsilon_t$, $e_t = a\epsilon_t + b\eta_t$, $(\epsilon_t, \eta_t) \sim iiN(0, \sigma^2 I_2)$ and $\sigma = 0.5$.

Table 4: Power: $m(x) = (1 + x^2)\exp(-\theta_1 x^2) + \triangle_n(x)$, $\theta_1 = 1$

n		Nominal size 1%			Nominal size 10%		
	$\kappa =$	1/5	1/4.5	1/4	1/5	1/4.5	1/4
		$\alpha_0 = 0.1$,	$a = 0.02$,	$b = -0.3$			
200		0.9968	0.9974	0.9978	0.9980	0.9988	0.9990
500		0.9972	0.9972	0.9974	0.9984	0.9992	0.9994
1200		0.9978	0.9976	0.9986	0.9990	0.9990	0.9996
		$\alpha_0 = 0.2$,	$a = 0$,	$b = 0.5$			
200		0.9846	0.9856	0.9858	0.9954	0.9964	0.9958
500		0.9860	0.9868	0.9870	0.9956	0.9958	0.9962
1200		0.9864	0.9860	0.9872	0.9986	0.9968	0.9968

$u_t = \alpha_0 u_{t-1} + \epsilon_t$, $e_t = a\epsilon_t + b\eta_t$, $(\epsilon_t, \eta_t) \sim iiN(0, \sigma^2 I_2)$ and $\sigma = 0.5$.

Example 4.2. This example examines the test Π_n using non-integrable regression models. The model for simulation is $y_t = m(x_t) + e_t$.

In situation I), the null hypothesis is H_{10}: $P(m(x_t) = \theta_0 x_t^2(x_t + \sin(x_t))) = 1$ with $\theta_0 = 1$; and the alternative is H_{11}: $P(m(x_t) = \theta_1 x_t^2(x_t + \sin(x_t)) + \triangle_n(x_t)) = 1$ with $\theta_1 = 1$.

In situation II), the null hypothesis is H_{10}: $P(m(x_t) = \theta_0 x_t^2(2 + x_t + \sin(x_t))) = 1$ with $\theta_0 = 1$; and the alternative is H_{11}: $P(m(x_t) = \theta_1 x_t^2(2 + x_t + \sin(x_t)) + \triangle_n(x_t)) = 1$ with $\theta_1 = 1$.

Table 5: Size: $m(x) = \theta_0 x^2(x + \sin(x))$, $\theta_0 = 1$

n		Nominal size 1%			Nominal size 10%		
	$\kappa =$	1/5	1/4.5	1/4	1/5	1/4.5	1/4
			$\alpha_0 = -0.05$,	$\beta_0 = -0.05$			
200		0.0068	0.0074	0.0070	0.0824	0.0868	0.0854
500		0.0104	0.0100	0.0094	0.0908	0.0994	0.0920
1200		0.0100	0.0104	0.0102	0.0994	0.0990	0.0988
			$\alpha_0 = -0.01$,	$\beta_0 = 0.04$			
200		0.0040	0.0044	0.0044	0.0814	0.0818	0.0844
500		0.0084	0.0080	0.0094	0.0898	0.0918	0.0954
1200		0.0098	0.0106	0.0112	0.1060	0.1046	0.0984

$u_t = \alpha_0 u_{t-1} + \epsilon_t$, $e_t = \beta_0 e_{t-1} + \epsilon_t$, $\sigma = 0.6$.

Tables 5 and 7 record the sizes of the test according to the two circumstances of data generation, respectively. As can be seen, undersized experiments account for a larger percentage, especially in Table 7. Note also in Table 7 that, at 1% level, when $\kappa = 1/4.5$ the sizes are not close enough to the significant level even when $n = 1200$. However, almost all sizes approach the significant lever stably as the sample size increase. By large, the performance of all sizes is satisfactory.

Tables 6 and 8 give all the power values of the experiments. Power values are still quite strong and robust in this example, in the sense that they are increasing with the increase of sample size and κ in the truncation parameter. The reason is the same as in Example 4.1. Also, it is noteworthy that in Table 8 there is a clear difference in the powers between two experiments, possibly due to the choice of the parameters α_0, a and b, or some unclear reason.

5 Empirical analysis

Since existing literature does not provide such a test that accommodates both endogeneity and nonstationarity, as pointed out before, one of the main motivations for the proposal of Π_n is to

Table 6: Power: $m(x) = \theta_0 x^2(x + \sin(x)) + \triangle_n(x)$, $\theta_0 = 1$

n		Nominal size 1%			Nominal size 10%		
	$\kappa =$	1/5	1/4.5	1/4	1/5	1/4.5	1/4
		$\alpha_0 = -0.05$,	$\beta_0 = -0.05$				
200		0.9174	0.9294	0.9386	0.9562	0.9632	0.9686
500		0.9270	0.9350	0.9496	0.9590	0.9660	0.9730
1200		0.9474	0.9520	0.9638	0.9756	0.9776	0.9824
		$\alpha_0 = -0.01$,	$\beta_0 = 0.04$				
200		0.9050	0.9184	0.9324	0.9496	0.9564	0.9626
500		0.9280	0.9388	0.9536	0.9630	0.9694	0.9736
1200		0.9526	0.9582	0.9650	0.9758	0.9788	0.9840

$u_t = \alpha_0 u_{t-1} + \epsilon_t$, $e_t = \beta_0 e_{t-1} + \epsilon_t$, $\sigma = 0.6$.

Table 7: Size: $m(x) = \theta_0 x^2(2 + x + \sin(x))$, $\theta_0 = 1$

n		Nominal size 1%			Nominal size 10%		
	$\kappa =$	1/5	1/4.5	1/4	1/5	1/4.5	1/4
		$\alpha_0 = 0.01$,	$a = -0.05$,	$b = -0.1$			
200		0.0058	0.0054	0.0050	0.0968	0.0900	0.0920
500		0.0068	0.0070	0.0060	0.0944	0.0924	0.0972
1200		0.0104	0.0082	0.0104	0.0980	0.0926	0.0986
		$\alpha_0 = 0.05$,	$a = 0.2$,	$b = 0.9$			
200		0.0060	0.0052	0.0054	0.0920	0.0872	0.0868
500		0.0070	0.0074	0.0076	0.1062	0.0950	0.0940
1200		0.0086	0.0078	0.0096	0.1004	0.0970	0.0990

$u_t = \alpha_0 u_{t-1} + \epsilon_t$, $e_t = a\epsilon_t + b\eta_t$ and $\sigma = 0.8$.

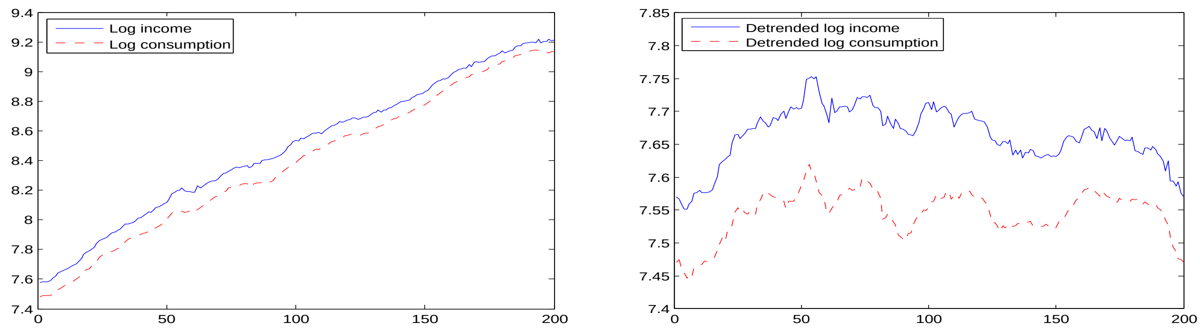
Table 8: Power: $m(x) = \theta_0 x^2(2 + x + \sin(x)) + \triangle_n(x)$, $\theta_0 = 1$

n		Nominal size 1%			Nominal size 10%		
	$\kappa =$	1/5	1/4.5	1/4	1/5	1/4.5	1/4
				$\alpha_0 = 0.01, \quad a = -0.05, \quad b = -0.1$			
200		0.9978	0.9984	0.9988	0.9982	0.9988	0.9990
500		0.9988	0.9990	0.9992	0.9992	0.9990	0.9996
1200		0.9990	0.9988	0.9990	0.9992	0.9992	0.9996
				$\alpha_0 = 0.05, \quad a = 0.2, \quad b = 0.9$			
200		0.7478	0.7934	0.8172	0.8234	0.8960	0.9014
500		0.8174	0.8260	0.8474	0.8778	0.9052	0.9182
1200		0.8632	0.8844	0.8928	0.9114	0.9418	0.9416

$u_t = \alpha_0 u_{t-1} + \epsilon_t$, $e_t = a\epsilon_t + b\eta_t$ and $\sigma = 0.8$.

naturally specify the relationship between the United States aggregate consumers' consumption expenditure and disposable income, in which a type of endogenity is inherited and the data sets are both nonstationary.

Example 5.1 Let us investigate the relationship between the United States aggregate consumers' consumption expenditure and disposable income over the time span 1960-2009. The data set is quarterly data from the Bureau of Economic Analysis at the website http://www.bea.gov. Let $c_t = \log(C_t)$ and $i_t = \log(I_t)$, where C_t and I_t are the aggregate consumption expenditures and the disposable incomes for $t = 1, \cdots, n$ with $n = 200$.



(a) Log income i_t and Log consumption c_t

(b) Detrended i_t and c_t

Figure 1: The real data and the detrended data

A visual inspection at the plot of i_t and c_t in Figure 1a suggests that there is a trend for each series. In order to obtain data satisfying the theoretical requirement, we need to remove the trend.

Towards this end, suppose

$$i_t = \mu_1 + i_{t-1} + u_t, \quad \text{and} \quad c_t = \mu_2 + c_{t-1} + v_t,$$

for $t = 2, \cdots, 200$ where u_t and v_t are two stationary mean zero processes. Thus,

$$\widehat{\mu}_1 = \frac{1}{n-1}\sum_{t=2}^{n}(i_t - i_{t-1}) = \frac{1}{199}(i_{200} - i_1) = 0.0082,$$
$$\widehat{\mu}_2 = \frac{1}{n-1}\sum_{t=2}^{n}(c_t - c_{t-1}) = \frac{1}{199}(c_{200} - c_1) = 0.0083.$$

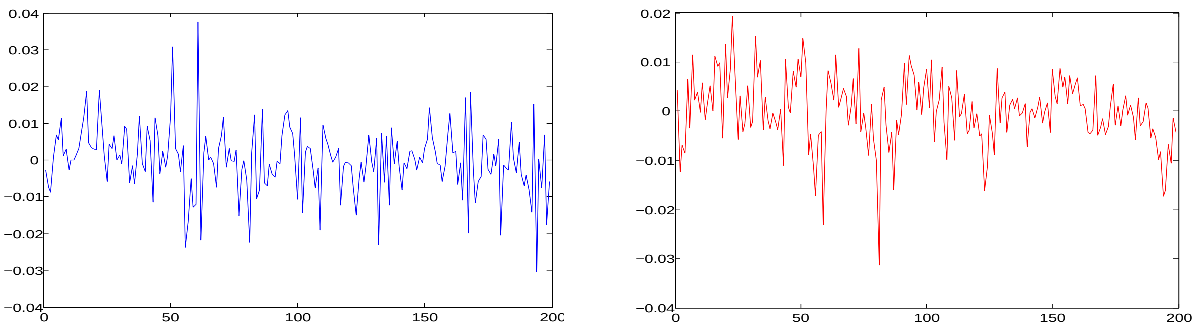
Hence, we may define

$$x_t = i_t - \widehat{\mu}_1 t, \quad \text{and} \quad y_t = c_t - \widehat{\mu}_2 t,$$

as detrended log income and log consumption, respectively. See Figure 1b. It follows that

$$x_t = x_{t-1} + u_t, \quad \text{and} \quad y_t = y_{t-1} + v_t.$$

Unit root tests are conducted for $\{x_t\}$ and $\{y_t\}$, respectively, using the augmented Dickey Fuller test with nonzero mean. According to the reports, all tests fail to reject the null of possessing unit root with p-values 0.5207 and 0.5619, respectively. These also are verified by the plots for the differences $\triangle x_t$ and $\triangle y_t$ in Figure 2.



(a) Difference of x_t (b) Difference of y_t

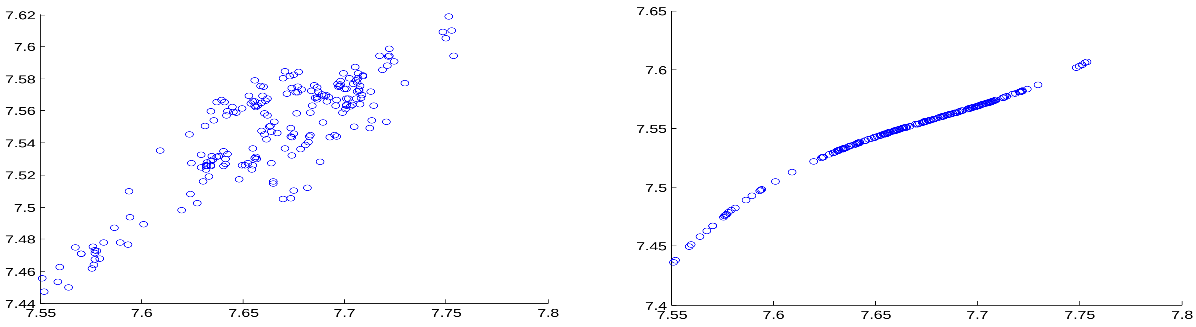
Figure 2: Difference of the detrended data

Consider a structural time series relationship between y_t and x_t of the form

$$y_t = m(x_t) + e_t, \tag{5.1}$$

where $m(\cdot)$ is an unknown function defined on $\mathbb{R}$.

The scatter of (x_t, y_t) in Figure 3a suggests that $m(x)$ is unlikely integrable on $\mathbb{R}$. Thus, we will approximate $m(x)$ in the space $L^2(\mathbb{R}, \exp(-x^2))$ by $m_k(x) = \sum_{i=0}^{k-1} b_i h_i(x)$, where $h_i(x) = \frac{1}{\sqrt{\sqrt{\pi}2^i i!}} H_i(x)$ and $H_i(x)$ is the Hermite polynomial of order i. Then $\beta = (b_0, \cdots, b_{k-1})'$ is estimated by (5.1).



(a) Scatter of (x_t, y_t) (b) Scatter of $(x_t, g(x_t; \widehat{\theta}))$

Figure 3: Scatters of the real data and fitted data

In doing so, we first have to choose a suitable truncation parameter k such that $m(x)$ can be better approximated by $m_k(x)$ based on the observations. We propose using the Generalised Cross–Validation (GCV) method (see Gao et al., 2002) to select an optimal value for k. While the GCV method is originally proposed for the stationary case, the finite sample analysis in this example shows that the GCV method works well numerically for the nonstationary case. Let $\widehat{k}$ denote the optimal value such that

$$\widehat{k} = \underset{k \in K_n}{\operatorname{argmin}} \left(1 - \frac{k}{n}\right)^{-2} \widehat{\sigma}^2(k), \tag{5.2}$$

where $K_n = \{[c \cdot n^{1/(2(q+1)+1)-\epsilon}], \cdots, [d \cdot n^{1/(2(q+1)+1)+\epsilon}]\}$, in which $0 < c < d < \infty$, q is the smoothness order of $m(x)$, $0 < \epsilon < 1/[2(q+1)(2(q+1)+1)]$, and $\widehat{\sigma}^2(k) = \frac{1}{n}\sum_{t=1}^{n}\left(y_t - \widehat{m}_k(x_t)\right)^2$, in which $\widehat{m}_k(x) = \sum_{i=0}^{k-1} \widehat{b}_i h_i(x)$ with $\widehat{b}_i$ being the i–th component of $\widehat{\beta}$.

In this example, we choose $c = 0.8$, $d = 2.5$, $q = 1$, so that $K_n = [2, \cdots, 7]$. We then have $\widehat{k} = 5$ by GCV and $\widehat{\beta} = (-1689.2, -799.8, 428.6, -74.6, 4.8)$ is the resulting OLS estimate of β, that is, GCV suggests $\widehat{m}_{\widehat{k}}(x) = \sum_{i=0}^{4} \widehat{b}_i h_i(x)$ where $\widehat{b}_i$ are the elements of $\widehat{\beta}$. This motivates us to test the following null hypothesis:

$$H_{10}: \ P(m(x_t) = g(x_t, \theta_0)) = 1. \tag{5.3}$$

where $g(x, \theta_0) = \theta_{00} + \theta_{01}x + \theta_{02}x^2 + \theta_{03}x^3 + \theta_{04}x^4$ with $\theta_0 = (\theta_{00}, \theta_{01}, \cdots, \theta_{04})'$. Under H_{10}, we have the following parametric model

$$y_t = g(x_t, \theta_0) + e_t. \tag{5.4}$$

Thus, θ_0 is estimated as $\widehat{\theta} = (-442.3, -1410.8, 596.7, -79.7, 3.5)$ by OLS. The proposed test Π_n is then applicable to deal with H_{10}. We employ a bootstrap scheme to generate a simulated P–value for Π_n under H_{10}. A commonly used bootstrap method is described as follows.

Step 1 Let $\{e_t\}$ of (5.1) follow an autoregressive time series of order one. Generate $\{e_t^*\}$ in the same way as for Case I) in Step 1 of the Bootstrap Simulation Procedure proposed in Section 4.

Step 2 Define $y_t^* = g(x_t; \widehat{\theta}) + e_t^*$. Use the data set $\{(x_t, y_t^*), 1 \le t \le n\}$ to re-estimate θ_0 and denote its estimate by $\widehat{\theta}^*$. Then calculate the test Π_n^*, which is the corresponding version of Π_n with $\{(x_t, y_t)\}$ and $\widehat{\theta}$ being replaced by $\{(x_t, y_t^*)\}$ and $\widehat{\theta}^*$, respectively.

Step 3 Repeat Steps 1–2 $M_b = 250$ times. We then have $\Pi_{n\ell}^*$, $\ell = 1, \cdots, 250$ and compute the proportion of $\Pi_n < \Pi_{n\ell}^*$ for model (5.4). This proportion is an approximate P–value of Π_n.

According to the simulated P–value, the statistic Π_n fails to reject the null H_{10} in (5.3) with 14.8% of P–value. The fitted curve is $y = g(x; \widehat{\theta}) = \widehat{\theta}_{00} + \widehat{\theta}_{01}x + \widehat{\theta}_{02}x^2 + \widehat{\theta}_{03}x^3 + \widehat{\theta}_{04}x^4$. Figure 3 gives both the scatter of the real data (x_t, y_t) and that of $(x_t, g(x_t; \widehat{\theta}))$.

To investigate further why the linear, quadratic, cubic, the fifth-order and the sixth-order polynomial relations between y_t and x_t are not favourable, we proceed in the following three aspects.

First, since y_t and x_t are I(1) processes, their differences $\triangle y_t$ and $\triangle x_t$ are stationary. If y_t and x_t follow a linear relationship, then $\triangle y_t$ and $\triangle x_t$ retain the linearity. Thus, we would like to test if there exists a linear relationship between $\triangle y_t$ and $\triangle x_t$ using $\widehat{L}_{bn}$ as given in equation (3.6). The simulated P–value of the test statistic $\widehat{L}_{bn}$ is 0.5%, implying that there is no evidence to support the linearity between $\triangle y_t$ and $\triangle x_t$. This outcome coincides with the intuition of the plot ($\triangle y_t$, $\triangle x_t$) in Figure 4.

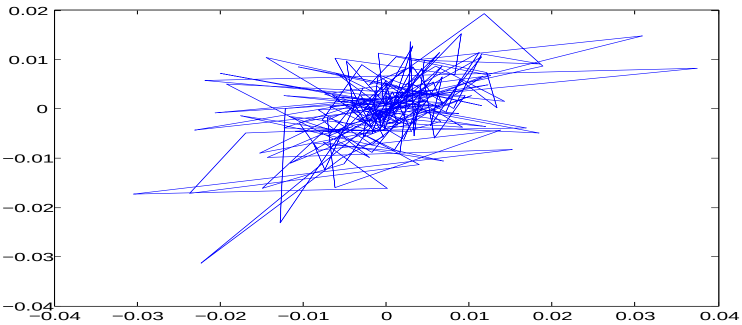


Figure 4: Plot of $(\triangle x_t, \triangle y_t)$

Second, in order to see how our proposed test statistic Π_n performs on the other regression forms that the GCV does not suggest, we conduct testings for $\mathcal{H}_{i0} : P(m(x_t) = P_i(x_t; a_i)) = 1$ with $P_i(x, a_i)$ $(i = 1, 2, 3, 5, 6)$ being linear, quadratic, cubic, fifth-order and sixth-order polynomials, respectively, where $a_i = (a_{i0}, \cdots, a_{ii})$ are the vector of the coefficients in the polynomials. The statistic Π_n reject all nulls because all P-values are less than 6%. This is understandable because we may write $P_i(x, a_i) = g(x, \theta) + \triangle_{in}(x)$ $(i = 1, 2, 3, 5, 6)$. Thus, the departure term $\triangle_{in}(x)$ that is a polynomial in x is easily detected by our proposed statistic as it is significantly 'larger' than $\triangle_n(x)$ in Assumption C.

Third, to compare further among the six polynomials, we regress y_t on $P_i(x_t; a_i)$, respectively, to obtain $\widehat{a}_i$, where $P_i(x; a_i) = a_{i0} + \cdots + a_{ii}x^i$ with $i = 1, \cdots, 6$. Then, we calculate the so-called in-sample mean square error (MSE)

$$\text{MSE}_{in} = \frac{1}{n}\sum_{t=1}^{n}(y_t - P_i(x_t, \widehat{a}_i))^2,$$

for $i = 1, \cdots, 6$.

Meanwhile, the so-called out-of-sample MSEs are also investigated. More precisely, based on the data sets $D_j = \{(y_t, x_t), 1 \le t \le 159 + 2j\}$ where $j = 1, 2, \cdots, 20$, we regress $\{y_t, t \le 159 + 2j\}$ on $\{P_i(x_t, a_i), t \le 159 + 2j\}$ for $i = 1, \cdots, 6$. Then the next value y_{160+2j} are estimated by $\widehat{y}^i_{160+2j} = P_i(x_{160+2j}, \widehat{a}_{i;j})$ where $\widehat{a}_{i;j} = (\widehat{a}_{i0;j}, \cdots, \widehat{a}_{ii;j})$ are the estimation of the parameters in each polynomial based on D_j by OLS. Hence, the out-of-sample MSEs are given by

$$\text{MSE}_{out} = \frac{1}{20}\sum_{j=1}^{20}(y_{160+2j} - \widehat{y}^i_{160+2j})^2,$$

for $i = 1, \cdots, 6$.

All the in-sample and out-of-sample MSEs are reported in Table 9. As can be seen, the fourth-order polynomial has the minimum MSEs in both the in-sample and the out-of-sample experiments among all six polynomials. This shows more evidence and support to the choice of the proposed statistic.

Table 9: The in-sample and out-of-sample MSEs

$i =$	1	2	3	4	5	6
MSE_{in}	$3.071e^{-4}$	$2.782e^{-4}$	$2.661e^{-4}$	$2.607e^{-4}$	$2.659e^{-4}$	$2.608e^{-4}$
MSE_{out}	$8.031e^{-4}$	$6.644e^{-4}$	$6.257e^{-4}$	$5.874e^{-4}$	$6.233e^{-4}$	$5.882e^{-4}$

Particularly, three fitted curves of the linear, the quadratic and the fourth-order are plotted in one figure to illustrate visually their difference. See Figure 5.

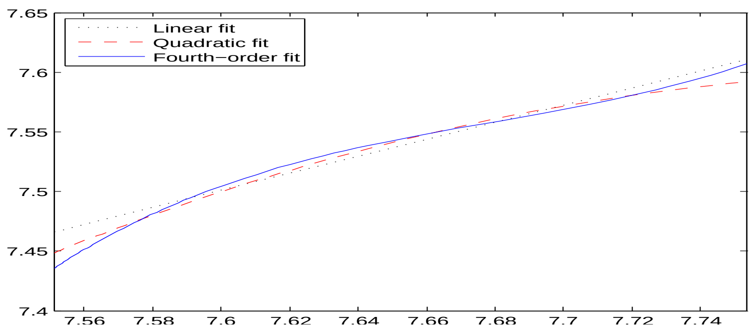


Figure 5: Plot of the linear, quadratic and fourth-order fitted curves

To conclude, the relationship between the detrended aggregate consumptions and incomes during 1960–2009 is tested by Π_n proposed in Section 3 and the fact that the detrended log consumption is the fourth-order polynomial of the detrended log income fails to reject. Other evidences also are provided to support the recommendation of the specification testing.

6 Conclusions and discussions

We have proposed two concise and computationally simple tests for parametric specification of structural time series models with endogeity and nonstationarity in both cases of integrable and non-integrable regression functions. An asymptotic theory for each of the proposed test statistics has been established. Several Monte Carlo simulation experiments have been used to evaluate the finite–sample performance of the proposed tests. Overall, the sizes and power values are found to be satisfactory. In addition, an empirical analysis has been provided to conclude that a simple linear relationship between the United States consumers' consumption expenditure and disposable income may not be justifiable, while a fourth-order polynomial relationship is not rejected.

As pointed out in the introductory section, the proposed tests have been developed for the univariate time series regressor case. In the case where one will need to model the relationship between y_t and a vector of integrated time series regressors x_t, one may consider using a varying–coefficient model of the form (see, e.g., Gao and Phillips, 2013a)

$$y_t = u_t'\beta(v_t, r_t) + e_t, \tag{6.1}$$

where $x_t = (u_t', v_t)'$ with u_t being a vector of integrated time series regressors and v_t being a univariate integrated time series regressor, $\{r_t\}$ is a vector of stationary regressors, e_t is an error time series and $\beta(\cdot)$ is a vector of unknown functions defined on $\mathbb{R}$.

As discussed in Gao and Phillips (2013b), one may also consider a semiparametric cointegration model of the form

$$y_t = u_t'\alpha + g(v_t) + e_t, \tag{6.2}$$

where α is a vector of unknown parameters and $g(\cdot)$ is an unknown function.

In this case, there is need in empirical applications to check whether the functional form of $\beta(\cdot)$ may be parametrically specified through testing $H_{20}: \ P\left(\beta(v_t, r_t) = \beta(v_t, r_t; \theta_0)\right) = 1$. If H_{20} is true, then model (6.1) will become a parametric time series model of the form

$$y_t = u_t'\beta(v_t, r_t; \theta_0) + e_t. \tag{6.3}$$

Similarly, for model (6.2), it is interested to test $H_{30}: \ P\left(g(v_t) = g(v_t; \theta_0)\right) = 1$. If H_{30} is true, model (6.2) becomes a partially nonlinear model of the form

$$y_t = u_t'\alpha + g(v_t; \theta_0) + e_t. \tag{6.4}$$

While the constructions of two corresponding tests for H_{20} and H_{30} follow similarly from that of L_n, the establishments of the resulting limiting distributions require some additional conditions and proofs. Letting y_t and $u_t = x_t$ be defined in the same way as in Example 5.1 above and v_t be the real interest rate, an important empirical application is to examine whether model (6.3) is more appropriate for the data than what has been discussed in the literature. A recent paper by Gao and Phillips (2013a) considers such an empirical problem, but it treats the real interest rate, which probably should be considered as nonstationary, as a stationary time series. We wish to leave such further discussion for future research.

It is of interest that models (6.1) and (6.2) allow for endogenity in some form that is similar to what has been considered in Assumption A of Section 2. There may also be some other extensions. One of them is whether the proposed test is extendable to accommodate the case where x_t is a vector of multivariate regressors in a nonparametric multivariate time series case. Our experience suggests that it may be possible to construct a multivariate series expansion and then a multivariate version of the proposed test before a resulting theory may be established. Another issue is how to develop a data–driven method for the choice of the truncated parameter k such that it is optimal for testing purposes. In the stationary time series case, Chapter 3 of Gao (2007) shows that one may choose a suitable truncated parameter such that the power function of the test under consideration is maximised while the size function is under control by a significance level. Since such extensions require developing new techniques, we also wish to leave them for future research.

A Lemmas and justification

The first section in Appendix is to present a few technical lemmas which are crucial for the proofs of the theorems, as well as a justification for Assumption B. Let us first introduce function sequence $\{\mathscr{T}_k(x)\}$ and function $\mathscr{T}(x)$:

$$\mathscr{T}_k(x) := \frac{1}{k}\|Z(x)\|^2 = \frac{1}{k}\sum_{i=0}^{k-1}\mathscr{H}_i^2(x) \quad \text{and} \quad \mathscr{T}(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2}, & \text{if } |x| \le 2, \\ 0, & \text{if } |x| > 2. \end{cases}$$

Lemma A.1. *(1) The sequence $\mathscr{T}_k(x)$ converges to $\mathscr{T}(x)$ for any $x \in \mathbb{R}$ as $k \to \infty$. Moreover, as $k \to \infty$*

$$\int_{-\infty}^{\infty} |\mathscr{T}_k(x) - \mathscr{T}(x)|\, dx \to 0. \tag{A.1}$$

(2) Furthermore, $\lim_{k\to\infty}\iint_{x\neq y}|Z(x)'Z(y)|dxdy = 0$.

Proof. (1) The convergence of $\mathscr{T}_k(x) \to \mathscr{T}(x)$ is the consequence of the celebrated Wigner's semicircle law (Wigner, 1958), which appears in Edelman and Rao (2005, p. 29) but is shown rigorously in Tao (2012, Section 2.6).

Since $\mathscr{T}_k(x)$ for each $k \ge 1$ can be viewed as a probability density and $\mathscr{T}(x)$ is actually the so-called Wigner semicircle law with radius 2, namely, a probability density as well, it follows from Scheffe's Theorem of Billingsley (1968, p. 224) that (A.1) holds.

(2) We will use Mehler's formula (see Szego, 1975, p. 380), for any $|r| < 1$,

$$\sum_{i=0}^{\infty} \mathscr{H}_i(x)\mathscr{H}_i(y) r^i = \frac{1}{\sqrt{1-r^2}} \exp\left(\frac{2r}{1+r}xy - \frac{r^2}{1-r^2}(x-y)^2\right),$$

and

$$\lim_{r\to 1^-} \frac{1}{\sqrt{1-r^2}} \exp\left(\frac{2r}{1+r}xy - \frac{r^2}{1-r^2}(x-y)^2\right) = \delta(x-y)$$

where $\delta(\cdot)$ is the Deric delta function , which only has point 0 as its support. See Example 2 of Gelf́and and Shilov (1964, P. 36).

Observe that

$$\begin{aligned}
&\lim_{k\to\infty} \iint_{x\neq y} |Z(x)'Z(y)| dxdy = \lim_{k\to\infty} \iint_{x\neq y} \left|\sum_{i=0}^{k-1} \mathscr{H}_i(x)\mathscr{H}_i(y)\right| dxdy \\
=& \lim_{k\to\infty} \iint_{x\neq y} \lim_{r\to 1^-} \left|\sum_{i=0}^{k-1} \mathscr{H}_i(x)\mathscr{H}_i(y) r^i\right| dxdy = \iint_{x\neq y} \lim_{r\to 1^-} \left|\sum_{i=0}^{\infty} \mathscr{H}_i(x)\mathscr{H}_i(y) r^i\right| dxdy \\
=& \iint_{x\neq y} \lim_{r\to 1^-} \frac{1}{\sqrt{1-r^2}} \exp\left(\frac{2r}{1+r}xy - \frac{r^2}{1-r^2}(x-y)^2\right) dxdy \\
=& \iint_{x\neq y} \delta(x-y) dxdy = 0.
\end{aligned}$$

Taking limit under the integration for $k \to \infty$ is due to the dominated convergence theorem for the series is dominated by an integrable function $\sum_{i=0}^{\infty} |\mathscr{H}_i(y)\mathscr{H}_i(x)| r^i$ with $0 < r < 1$. Exchange of two limits for k and r is because the series is convergent uniformly in r in any bounded interval. In fact, noting that Hermite polynomials have an integral expression $H_j(x) = (-1)^j \frac{1}{2\sqrt{\pi}} e^{x^2} \int (iu)^j e^{iux-u^2/4} du$ where i is imaginary unit,

$$\begin{aligned}
&\sum_{j=0}^{k-1} \mathscr{H}_j(y)\mathscr{H}_j(x) r^j = \sum_{j=0}^{k-1} \frac{r^j}{2^j j! \sqrt{\pi}} H_j(y) H_j(x) e^{-(x^2+y^2)/2} \\
=& \frac{e^{(x^2+y^2)/2}}{4\pi\sqrt{\pi}} \iint \left(\sum_{j=0}^{k-1} \frac{1}{2^j j!} (-ruv)^j\right) e^{i(ux+vy)-u^2/4-v^2/4} dudv \\
\to& \frac{e^{(x^2+y^2)/2}}{4\pi\sqrt{\pi}} \iint e^{-ruv/2} e^{i(ux+vy)-u^2/4-v^2/4} dudv
\end{aligned}$$

where the sum converges to $e^{-ruv/2}$ uniformly in r in any fixed finite interval, giving the Mehler's formula by calculating the double integral. $\square$

We shall consider several versions of decomposition for x_t. Without loss of generality, in what follows let $x_0 = 0$ almost surely. Let $j \le t$ be fixed. Observe that, when $j > 0$,

$$\begin{aligned}
x_t =& \sum_{i=1}^{t} u_i = \sum_{i=1}^{t}\sum_{a=0}^{\infty} \psi_a \epsilon_{i-a} = \sum_{i=1}^{t}\sum_{a=-\infty}^{i} \psi_{i-a}\epsilon_a = \sum_{i=1}^{j-1}\sum_{a=-\infty}^{i} \psi_{i-a}\epsilon_a + \sum_{i=j}^{t}\sum_{a=-\infty}^{i} \psi_{i-a}\epsilon_a \\
=& \sum_{i=1}^{j-1}\sum_{a=-\infty}^{i} \psi_{i-a}\epsilon_a + \sum_{i=j}^{t}\sum_{a=-\infty,\neq j}^{i} \psi_{i-a}\epsilon_a + \epsilon_j \sum_{i=j}^{t} \psi_{i-j} := \epsilon_j \sum_{i=j}^{t} \psi_{i-j} + W_{1t,j},
\end{aligned}$$

and when $j \leq 0$,

$$
\begin{aligned}
x_t =& \sum_{i=1}^{t} u_i = \sum_{i=1}^{t}\sum_{a=0}^{\infty} \psi_a \epsilon_{i-a} = \sum_{i=1}^{t}\sum_{a=-\infty}^{i} \psi_{i-a}\epsilon_a \\
=& \epsilon_j \sum_{i=1}^{t} \psi_{i-j} + \sum_{i=1}^{t} \sum_{a=-\infty, \neq j}^{i} \psi_{i-a}\epsilon_a := \epsilon_j \sum_{i=1}^{t} \psi_{i-j} + W_{2t,j}.
\end{aligned}
$$

For the sake of convenience and due to unambiguity, denote by $x_{t/j}$ for $W_{1t,j}$ or $W_{2t,j}$, which in fact stands for the remaining variable after extracting ϵ_j from x_t. Thus we always have

$$
x_t = b\epsilon_j + x_{t/j} \tag{A.2}
$$

where $b = \sum_{i=\max(1,j)}^{t} \psi_{i-j}$ for any $j \leq t$.

Additionally, letting $1 \leq s < j \leq t$, x_t also has the following decomposition.

$$
\begin{aligned}
x_t =& \sum_{i=1}^{t} u_i = x_s + \sum_{i=s+1}^{t}\sum_{a=0}^{\infty} \psi_a \epsilon_{i-a} = x_s + \sum_{i=s+1}^{t}\sum_{a=-\infty}^{i} \psi_{i-a}\epsilon_a \\
=& x_s + \sum_{i=s+1}^{t}\sum_{a=-\infty}^{s} \psi_{i-a}\epsilon_a + \sum_{i=s+1}^{t}\sum_{a=s+1}^{i} \psi_{i-a}\epsilon_a = x_s + \sum_{i=s+1}^{t}\sum_{a=-\infty}^{s} \psi_{i-a}\epsilon_a + \sum_{a=s+1}^{t} \epsilon_a \sum_{i=0}^{t-a} \psi_i \\
=& x_s + \sum_{i=s+1}^{t}\sum_{a=-\infty}^{s} \psi_{i-a}\epsilon_a + \epsilon_j \sum_{i=0}^{t-j} \psi_i + \sum_{a=s+1, \neq j}^{t} \epsilon_a \sum_{i=0}^{t-a} \psi_i := x_s^* + \epsilon_j \sum_{i=j}^{t} \psi_{i-j} + x_{ts/j},
\end{aligned} \tag{A.3}
$$

where $x_s^* = x_s + \bar{x}_s$ with $\bar{x}_s = \sum_{i=s+1}^{t}\sum_{a=-\infty}^{s} \psi_{i-a}\epsilon_a$ containing all information available up to s and $x_{ts/j} = \sum_{a=s+1,\neq j}^{t} \epsilon_a \sum_{i=0}^{t-a} \psi_i$. Evidently, $x_{ts/j}$ captures all information containing in x_t on the time periods $(s, j) \cup (j, t]$. Note that $\bar{x}_s = O_P(1)$, since

$$
\begin{aligned}
E[\bar{x}_s]^2 =& E\left(\sum_{i=s+1}^{t}\sum_{a=-\infty}^{s} \psi_{i-a}\epsilon_a\right)^2 = \sum_{a=-\infty}^{s}\left(\sum_{i=s+1}^{t} \psi_{i-a}\right)^2 \\
=& \sum_{a=-\infty}^{s}\sum_{i=s+1}^{t} \psi_{i-a}^2 + 2\sum_{a=-\infty}^{s}\sum_{i=s+2}^{t}\sum_{\ell=s+1}^{i-1} \psi_{i-a}\psi_{\ell-a} \\
=& O(1)\sum_{i=1}^{t-s}\sum_{a=0}^{\infty}(i+a)^{-2\lambda} + O(1)\sum_{i=2}^{t-s}\sum_{\ell=1}^{i-1}\sum_{a=0}^{\infty}(i+a)^{-\lambda}(\ell+a)^{-\lambda} \\
\leq& O(1)\sum_{i=1}^{t-s} i^{-2\lambda+1} + O(1)\sum_{i=2}^{t-s}\sum_{\ell=1}^{i-1} i^{-\lambda+1/2}\ell^{-\lambda+1/2} \\
=& O(1)(1-(t-s)^{-2\lambda+2}) + O(1)(2^{-\lambda+3/2} - (t-s)^{-\lambda+3/2}) = O(1).
\end{aligned}
$$

Lemma A.2. *Assume Assumption A holds.*

(1) Let $j \leq t$. $\frac{1}{\sqrt{t}}x_{t/j}$ where $x_{t/j}$ is given by (A.2) *has a uniformly bounded density $f_{t/j}(x)$ over all t, j and x satisfying uniform Lipschitz condition $\sup_x |f_t(x+y) - f_t(x)| \leq C|y|$ for any y and some constant $C > 0$.*

(2) Let $1 \leq s < j \leq t$. $\frac{1}{\sqrt{t-s}} x_{ts/j}$ where $x_{ts/j}$ is given by (A.3) *has a uniformly bounded density $f_{ts/j}(x)$ over all t, j and s, x satisfying the above uniform Lipschitz condition as well.*

Proof. Since $Ex_t^2 = O(1)t$ by (2.2) and $b = O(1)$ defined by (A.2), it follows from the orthogonality between ϵ_j and $x_{t/j}$ that $Ex_{t/j}^2 = Ex_t^2 - b^2 E[\epsilon_j^2] = O(1)t - b^2 = O(1)t$. Similarly, $Ex_{ts/j}^2 = O(1)(t-s)$. Therefore, the existence of the densities in (1) and (2) follows from the proof of Corollary 2.2 in Wang and Phillips (2009a) and (8.2) in Wang and Phillips (2012b, p. 13), respectively. The uniform Lipschitz condition can be shown by the routine argument of the relationship between density and characteristic function. See also Wang and Phillips (2012b, p. 13). □

Lemma A.3. *Assume Assumption A holds. Let j be a fixed integer such that $j \leq t$. For any functions U and $f : \mathbb{R} \mapsto \mathbb{R}$ such that $\int |U(w)|dw < \infty$ and $E|\epsilon_j f(\epsilon_j)| < \infty$, we have*

(1) $E[U(x_t) f(\epsilon_j)] = E[U(x_{t/j})]E[f(\epsilon_j)] + c_U \frac{1}{t}$ where $x_{t/j}$ defined by (A.2) *is independent of ϵ_j, c_U is a quantity such that $|c_U| \leq O(1)E|\epsilon_j f(\epsilon_j)| \int |U(w)|dw$. In particular, if $Ef(\epsilon_j) = 0$, $E[U(x_t)f(\epsilon_j)] = c_U \frac{1}{t}$;*

(2) $E|U(x_t)f(\epsilon_j)| \leq O(1)\frac{1}{\sqrt{t}} E|f(\epsilon_j)| \int |U(w)|dw$;

(3) For any ℓ: $j \neq \ell \leq t$, $E[U(x_t)f(\epsilon_j)|\epsilon_\ell] = E[U(x_{t/j})|\epsilon_\ell]E[f(\epsilon_j)] + \frac{1}{t}\eta_\ell$ where $x_{t/j}$ is defined by (A.2), *η_ℓ is a random variable depending on ϵ_ℓ such that $|\eta_\ell| \leq O(1) \int |U(w)|dw$ almost surely.*

When $E[f(\epsilon_j)] = 0$, $E[U(x_t)f(\epsilon_j)|\epsilon_\ell] = \frac{1}{t}\eta_\ell$.

Meanwhile, $E[|U(x_t)f(\epsilon_j)||\epsilon_\ell] \leq O(1)\frac{1}{\sqrt{t}} E|f(\epsilon_j)| \int |U(w)|dw$ almost surely.

(4) For $1 \leq s < j \leq t$, $E[U(x_t)f(\epsilon_j)|\mathcal{F}_s] = E[f(\epsilon_j)]E[U(x_{t/j})|\mathcal{F}_s] + \frac{1}{t-s}\xi_s$ where $|\xi_s| \leq O(1)E|\epsilon_j f(\epsilon_j)| \int |U(x)|dx$ almost surely; meanwhile, $E[|U(x_t)f(\epsilon_j)||\mathcal{F}_s] \leq O(1)\frac{1}{\sqrt{t-s}} E[|f(\epsilon_j)|] \int |U(w)|dw$ a.s.

(5) Similarly, letting $1 \leq s < j \neq \ell \leq t$ and $g(x)$ be a function such that $E|\epsilon_\ell g(\epsilon_\ell)| < \infty$, we have $E[U(x_t)f(\epsilon_j)g(\epsilon_\ell)|\mathcal{F}_s] = E[f(\epsilon_j)g(\epsilon_\ell)]E[U(x_{t/j\ell})|\mathcal{F}_s] + \frac{1}{t-s}\xi_s'$ where $x_{t/j\ell}$ is defined in the same way as $x_{t/j}$ and ξ_s' satisfies a similar condition as ξ_s, that is, $|\xi_s'| \leq O(1) \int |U(x)|dx$ almost surely. Moreover, similar to (2), we have $E|U(x_t)f(\epsilon_j)g(\epsilon_\ell)| \leq O(1)\frac{1}{\sqrt{t}} E|f(\epsilon_j)|E|g(\epsilon_\ell)| \int |U(w)|dw$.

(6) For $s \leq t$, $E[U(x_t)e_s] = c_U' \frac{1}{t}$ where c_U' satisfies a similar condition as the c_U in (1) and $E|U(x_t)e_s^2| \leq O(1)\frac{1}{\sqrt{t}} \int |U(w)|dw$ for e_t defined by a linear process in b(i) of Assumption A.

Proof. (1) Recalling that $x_t = b\epsilon_j + x_{t/j}$ with $b = \sum_{i=1\vee j}^t \psi_{i-j} = O(1)$ by the definition of the ψ_j, $x_{t/j}$ has a uniformly bounded density $f_{t/j}(x)$ satisfying Lipschitz condition and supposing that ϵ_j has density $h_\epsilon(v)$, we have

$$E[U(x_t)f(\epsilon_j)] = E[U(b\epsilon_j + x_{t/j})f(\epsilon_j)] = \iint U(bv + \sqrt{t}w)f(v)h_\epsilon(v)f_{t/j}(w)dvdw$$

$$= \iint U(w)f(v)h_\epsilon(v)\frac{1}{\sqrt{t}} f_{t/j}\left(\frac{w - bv}{\sqrt{t}}\right) dvdw = \iint U(w)f(v)h_\epsilon(v)\frac{1}{\sqrt{t}} f_{t/j}\left(\frac{w}{\sqrt{t}}\right) dvdw$$

$$
\begin{aligned}
&+\iint U(w)f(v)h_\epsilon(v)\frac{1}{\sqrt{t}}\left[f_{t/j}\left(\frac{w-bv}{\sqrt{t}}\right)-f_{t/j}\left(\frac{w}{\sqrt{t}}\right)\right]dvdw\\
=&\int f(v)h_\epsilon(v)dv\int U(w)\frac{1}{\sqrt{t}}f_{t/j}\left(\frac{w}{\sqrt{t}}\right)dw+c_U\frac{1}{t}\\
=&E[f(\epsilon_j)]\int U(\sqrt{t}w)f_{t/j}(w)dw+c_U\frac{1}{t}=E[U(x_{t/j})]E[f(\epsilon_j)]+c_U\frac{1}{t},
\end{aligned}
$$

where $c_U := \sqrt{t}\iint U(w)f(v)h_\epsilon(v)\left[f_{t/j}\left(\frac{w-bv}{\sqrt{t}}\right)-f_{t/j}\left(\frac{w}{\sqrt{t}}\right)\right]dvdw$ satisfies

$$
\begin{aligned}
|c_U|\le&\sqrt{t}\int|f(v)|h_\epsilon(v)\int|U(w)|\left|f_{t/j}\left(\frac{w-bv}{\sqrt{t}}\right)-f_{t/j}\left(\frac{w}{\sqrt{t}}\right)\right|dwdv\\
\le&O(1)\int|f(v)|h_\epsilon(v)\int|U(w)||bv|dwdv=O(1)E|\epsilon_jf(\epsilon_j)|\int|U(w)|dw,
\end{aligned}
$$

using Lipschitz condition for $f_{t/j}$ in Lemma A.2. Clearly, if $Ef(\epsilon_j)=0$, $E[U(x_t)f(\epsilon_j)]=c_U\frac{1}{t}$.

(2) It follows that

$$
\begin{aligned}
E|U(x_t)f(\epsilon_j)|=&E|U(b\epsilon_j+x_{t/j})f(\epsilon_j)|=\iint|U(bv+\sqrt{t}w)f(v)|h_\epsilon(v)f_{t/j}(w)dvdw\\
=&\frac{1}{\sqrt{t}}\iint|U(w)f(v)|h_\epsilon(v)f_{t/j}\left(\frac{w-bv}{\sqrt{t}}\right)dvdw\le O(1)\frac{1}{\sqrt{t}}\iint|U(w)f(v)|h_\epsilon(v)dvdw\\
=&O(1)\frac{1}{\sqrt{t}}\int|f(v)|h_\epsilon(v)dv\int|U(w)|dw=O(1)\frac{1}{\sqrt{t}}E|f(\epsilon_j)|\int|U(w)|dw.
\end{aligned}
$$

(3) Because of similarity we only consider here $\ell>j>0$. In this case, similar to (A.2) we have

$$
\begin{aligned}
x_t=&\sum_{i=1}^t u_i=\sum_{i=1}^t\sum_{a=0}^\infty\psi_a\epsilon_{i-a}=\sum_{i=1}^t\sum_{a=-\infty}^i\psi_{i-a}\epsilon_a\\
=&\sum_{i=1}^{j-1}\sum_{a=-\infty}^i\psi_{i-a}\epsilon_a+\sum_{i=j}^t\sum_{a=-\infty}^{j-1}\psi_{i-a}\epsilon_a+\epsilon_j\sum_{i=j}^t\psi_{i-j}+\sum_{i=j}^t\sum_{a=j+1}^{\ell-1}\psi_{i-a}\epsilon_a\\
&+\epsilon_\ell\sum_{i=\ell}^t\psi_{i-\ell}+\sum_{i=\ell}^t\sum_{a=\ell+1}^i\psi_{i-a}\epsilon_a:=b_1\epsilon_j+b_2\epsilon_\ell+x_{t/j\ell},
\end{aligned}
$$

where $x_{t/j\ell}$ includes all terms except ϵ_ℓ and ϵ_j, and by convention $\sum_{a=a_1}^{a_2}=0$ for any $a_2<a_1$.

Moreover, $Ex_{t/j\ell}^2=Ex_t^2-b_1^2-b_2^2=O(1)t$ and, similar to Lemma A.2, we may show that $\frac{1}{\sqrt{t}}x_{t/j\ell}$ has a density $f_{t/j\ell}(x)$ and $f_{t/j\ell}(x)$ satisfies Lipschtiz condition uniformly on $\mathbb{R}$. Recalling that ϵ_j has density $h_\epsilon(v)$,

$$
\begin{aligned}
E[U(x_t)f(\epsilon_j)|\epsilon_\ell]=&E[U(b_1\epsilon_j+b_2\epsilon_\ell+x_{t/j\ell})f(\epsilon_j)|\epsilon_\ell]\\
=&\iint U(b_1v+b_2\epsilon_\ell+\sqrt{t}w)f(v)h_\epsilon(v)f_{t/j\ell}(w)dvdw\\
=&\iint U(w)f(v)h_\epsilon(v)\frac{1}{\sqrt{t}}f_{t/j\ell}\left(\frac{w-b_1v-b_2\epsilon_\ell}{\sqrt{t}}\right)dvdw\\
=&\iint U(w)f(v)h_\epsilon(v)\frac{1}{\sqrt{t}}f_{t/j\ell}\left(\frac{w-b_2\epsilon_\ell}{\sqrt{t}}\right)dvdw\\
&+\iint U(w)f(v)h_\epsilon(v)\frac{1}{\sqrt{t}}\left[f_{t/j\ell}\left(\frac{w-b_1v-b_2\epsilon_\ell}{\sqrt{t}}\right)-f_{t/j\ell}\left(\frac{w-b_2\epsilon_\ell}{\sqrt{t}}\right)\right]dvdw
\end{aligned}
$$

$$=E[U(x_{t/j\ell}+b_2\epsilon_\ell)|\epsilon_\ell]E[f(\epsilon_j)]+\frac{1}{t}\eta_\ell=E[U(x_{t/j})|\epsilon_\ell]E[f(\epsilon_j)]+\frac{1}{t}\eta_\ell$$

and using Lipschitz condition

$$|\eta_\ell|\leq O(1)\int|f(v)|h_\epsilon(v)\int|U(w)||b_1v|dwdv=O(1)\int|vf(v)|h_\epsilon(v)dv\int|U(w)|dw,\quad a.s.$$

When $E[f(\epsilon_j)]=0$ we shall have $E[U(x_t)f(\epsilon_j)|\epsilon_\ell]=\frac{1}{t}\eta_\ell$. Additionally,

$$\begin{aligned}E[|U(x_t)f(\epsilon_j)||\epsilon_\ell]&=\iint|U(b_1v+b_2\epsilon_\ell+\sqrt{t}w)f(v)|h_\epsilon(v)f_{t/j\ell}(w)dvdw\\ =&\iint|U(w)f(v)h_\epsilon(v)|\frac{1}{\sqrt{t}}f_{t/j\ell}\left(\frac{w-b_1v-b_2\epsilon_\ell}{\sqrt{t}}\right)dvdw\\ \leq&O(1)\frac{1}{\sqrt{t}}\iint|U(w)f(v)h_\epsilon(v)|dvdw=O(1)\frac{1}{\sqrt{t}}E|f(\epsilon_j)|\int|U(w)|dw,\end{aligned}$$

almost surely.

(4) Recalling that $x_t=x_s^*+b\epsilon_j+x_{ts/j}$ and $\frac{1}{\sqrt{t-s}}x_{ts/j}$ has a uniformly bounded density $f_{ts/j}(x)$ satisfying uniform Lipschitz condition,

$$\begin{aligned}&E[U(x_t)f(\epsilon_j)|\mathcal{F}_s]=E[U(x_s^*+b\epsilon_j+x_{ts/j})f(\epsilon_j)|\mathcal{F}_s]=\iint U(x_s^*+bv+\sqrt{t-s}x)f(v)h_\epsilon(v)f_{ts/j}(x)dxdv\\ =&\iint U(x)f(v)h_\epsilon(v)\frac{1}{\sqrt{t-s}}f_{ts/j}\left(\frac{x-bv-x_s^*}{\sqrt{t-s}}\right)dvdx=\iint U(x)f(v)h_\epsilon(v)\frac{1}{\sqrt{t-s}}f_{ts/j}\left(\frac{x-x_s^*}{\sqrt{t-s}}\right)dvdx\\ &+\iint U(x)f(v)h_\epsilon(v)\frac{1}{\sqrt{t-s}}\left[f_{ts/j}\left(\frac{x-bv-x_s^*}{\sqrt{t-s}}\right)-f_{ts/j}\left(\frac{x-x_s^*}{\sqrt{t-s}}\right)\right]dvdx\\ =&\int f(v)h_\epsilon(v)dv\int U(x)\frac{1}{\sqrt{t-s}}f_{ts/j}\left(\frac{x-x_s^*}{\sqrt{t-s}}\right)dvdx+\frac{1}{t-s}\xi_s=E[f(\epsilon_j)]E[U(x_{ts/j}+x_s^*)|\mathcal{F}_s]+\frac{1}{t-s}\xi_s,\end{aligned}$$

where

$$\xi_s=\sqrt{t-s}\iint U(x)f(v)h_\epsilon(v)\left[f_{ts/j}\left(\frac{x-bv-x_s^*}{\sqrt{t-s}}\right)-f_{ts/j}\left(\frac{x-x_s^*}{\sqrt{t-s}}\right)\right]dvdx,\tag{A.4}$$

and using Lipschitz condition,

$$|\xi_s|\leq C\iint|U(x)f(v)|h_\epsilon(v)|bv|dvdx=O(1)E|\epsilon_jf(\epsilon_j)|\int|U(x)|dx\quad a.s.$$

Consequently, when $E[f(\epsilon_j)]=0$, $E[U(x_t)f(\epsilon_j)|\mathcal{F}_s]=\frac{1}{t-s}\xi_s$. Meanwhile,

$$\begin{aligned}E[|U(x_t)f(\epsilon_j)||\mathcal{F}_s]&=\iint|U(x_s^*+bv+\sqrt{t-s}x)f(v)|h_\epsilon(v)f_{ts/j}(x)dxdv\\ =&\iint|U(x)f(v)|h_\epsilon(v)\frac{1}{\sqrt{t-s}}f_{ts/j}\left(\frac{x-bv-x_s^*}{\sqrt{t-s}}\right)dvdx\\ \leq&O(1)\frac{1}{\sqrt{t-s}}\iint|U(x)f(v)|h_\epsilon(v)dxdv=O(1)\frac{1}{\sqrt{t-s}}E|f(\epsilon_j)|\int|U(x)|dx.\end{aligned}$$

(5) As we may decompose $x_t=x_s^*+b_1\epsilon_j+b_2\epsilon_\ell+x_{ts/j\ell}$ where x_s^* contains all information available up to s and $x_{ts/j\ell}$ is the remaining term deducting the first three from x_t, and we may show that $\frac{1}{\sqrt{t-s}}x_{ts/j\ell}$ has a uniformly bounded density which satisfies the uniform Lipschitz condition, the assertion follows similarly.

(6) $E[U(x_t)e_s] = \sum_{j=0}^{\infty}\phi_j E[U(x_t)\epsilon_{s-j}] = \frac{1}{t}\sum_{j=0}^{\infty}\phi_j c_{U,j} := c_U' \frac{1}{t}$ where $c_{U,j}$ are the counterparts of c_U in (1), and

$$E|U(x_t)e_s^2| = E[|U(x_t)|e_s^2] = \sum_{j=-\infty}^{s}\phi_{s-j}^2 E[|U(x_t)|\epsilon_j^2] + 2\sum_{j=-\infty}^{s-1}\sum_{\ell=j+1}^{s}\phi_{s-j}\phi_{s-\ell}E[|U(x_t)|\epsilon_j\epsilon_\ell)]$$
$$\le \frac{1}{\sqrt{t}}\int|U(w)|dw\sum_{j=0}^{\infty}\phi_j^2 + O(1)\frac{1}{t}\sum_{j=1}^{\infty}\sum_{\ell=0}^{j-1}\phi_j\phi_\ell\eta_\ell = O(1)\frac{1}{\sqrt{t}}\int|U(w)|dw,$$

where η_ℓ is generated by (3) and $|\eta_\ell| \le O(1)\int|U(w)|dw$ a.s. uniformly in ℓ. □

Corresponding to e_t defined by b(ii) in Assumption A, the following shall develop a similar version of the preceding lemma. Observe that, if $t > m_0$,

$$\begin{aligned} x_t =& x_{t-m_0} + \sum_{i=t-m_0+1}^{t} u_i = x_{t-m_0} + \sum_{i=t-m_0+1}^{t}\sum_{a=-\infty}^{i}\psi_{i-a}\epsilon_a \\ =& x_{t-m_0} + \sum_{i=t-m_0+1}^{t}\sum_{a=-\infty}^{t-m_0}\psi_{i-a}\epsilon_a + \sum_{i=t-m_0+1}^{t}\sum_{a=t-m_0+1}^{i}\psi_{i-a}\epsilon_a \\ =& x_{t-m_0} + \sum_{i=t-m_0+1}^{t}\sum_{a=-\infty}^{t-m_0}\psi_{i-a}\epsilon_a + \sum_{a=t-m_0+1}^{t}\epsilon_a\sum_{i=0}^{t-a}\psi_i \\ :=& x_{t-m_0}^* + b(\epsilon_t, \cdots, \epsilon_{t-m_0+1}), \end{aligned} \tag{A.5}$$

where $x_{t-m_0}^* = x_{t-m_0} + \sum_{i=t-m_0+1}^{t}\sum_{a=-\infty}^{t-m_0}\psi_{i-a}\epsilon_a$ and $b(\epsilon_t, \cdots, \epsilon_{t-m_0+1}) := \sum_{a=t-m_0+1}^{t}\epsilon_a\psi(t-a)$, a linear function with $\psi(t-a) := \sum_{i=0}^{t-a}\psi_i$, a partial sum of $\sum_{i=0}^{\infty}\psi_i$.

Then, noting $\sum_{i=t-m_0+1}^{t}\sum_{a=-\infty}^{t-m_0}\psi_{i-a}\epsilon_a = O_P(1)$ and thus $E[x_{t-m_0}^*]^2 = O(1)(t-m_0)$, it follows similarly as Lemma A.2 that $\frac{1}{\sqrt{t-m_0}}x_{t-m_0}^*$ has a density $f_{t-m_0}(w)$ uniformly bounded and satisfying uniform Lipschitz condition.

If $t - s > m_0$,

$$\begin{aligned} x_t =& x_s + \sum_{i=s+1}^{t-m_0} u_i + \sum_{i=t-m_0+1}^{t} u_i = x_s + \sum_{i=s+1}^{t-m_0}\sum_{a=-\infty}^{i}\psi_{i-a}\epsilon_a + \sum_{i=t-m_0+1}^{t}\sum_{a=-\infty}^{i}\psi_{i-a}\epsilon_a \\ =& x_s + \sum_{i=s+1}^{t}\sum_{a=-\infty}^{s}\psi_{i-a}\epsilon_a + \sum_{i=s+1}^{t-m_0}\sum_{a=s+1}^{i}\psi_{i-a}\epsilon_a + \sum_{i=t-m_0+1}^{t}\sum_{a=s+1}^{t-m_0}\psi_{i-a}\epsilon_a + \sum_{i=t-m_0+1}^{t}\sum_{a=t-m_0+1}^{i}\psi_{i-a}\epsilon_a \\ =& x_s + \sum_{i=s+1}^{t}\sum_{a=-\infty}^{s}\psi_{i-a}\epsilon_a + \sum_{a=s+1}^{t-m_0}\epsilon_a[2\psi(t-a) - \psi(t-a-m_0+1)] + \sum_{a=t-m_0+1}^{t}\epsilon_a\psi(t-a) \\ =& x_s^* + x_{s+1,t-m_0}' + b(\epsilon_t, \cdots, \epsilon_{t-m_0+1}), \end{aligned} \tag{A.6}$$

where $x_s^* = x_s + \sum_{i=s+1}^{t}\sum_{a=-\infty}^{s}\psi_{i-a}\epsilon_a$, $x_{s+1,t-m_0}' = \sum_{a=s+1}^{t-m_0}\epsilon_a\Psi(t-a)$ with $\Psi(t-a) := [2\psi(t-a) - \psi(t-a-m_0+1)]$ for convenience, $b(\cdots)$ is defined as before. By virtue of the structure of $x_{s+1,t-m_0}'$, $\frac{1}{\sqrt{t-s-m_0}}x_{s+1,t-m_0}'$ has a density $f_{s+1,t-m_0}(w)$ uniformly bounded and satisfying uniform Lipschitz condition.

Lemma A.4. *Let e_t be defined by b(ii) in Assumption A. Let $U(\cdot)$ be an integrable function on $\mathbb{R}$, i.e. $\int |U(w)|dw < \infty$, $f(\cdot)$ be a function such that $E|b(\epsilon_t, \cdots, \epsilon_{t-m_0+1})f(e_t)| < \infty$.*

(1) For $t > m_0$, $E[U(x_t)f(e_t)] = E[U(x^*_{t-m_0})]E[f(e_t)] + \frac{1}{t-m_0}c_U$, where $x^*_{t-m_0}$ is given by (A.5) *and c_U satisfies $|c_U| \le O(1)\int |U(w)|dw$ almost surely. In particular, if $E[f(e_t)] = 0$, $E[U(x_t)f(e_t)] = \frac{1}{t-m_0}c_U$; meanwhile, $|E[U(x_t)f(e_t)]| \le C\frac{1}{\sqrt{t-m_0}}E|f(e_t)|\int |U(w)|dw$, where C is an absolutely constant.*

(2) For $t - s > m_0$, $E[U(x_t)f(e_t)|\mathcal{F}_s] = E[U(x^*_s + x'_{s+1,t-m_0})|\mathcal{F}_s]E[f(e_t)] + \frac{1}{t-s-m_0}\xi_s$ with $|\xi_s| \le CE|b(\epsilon_t \cdots \epsilon_{t-m_0+1})f(e_t)|\int |U(w)|dw$ almost surely, where x^*_s and $x'_{s+1,t-m_0}$ are defined by (A.6)*. If $E[f(e_t)] = 0$, $E[U(x_t)f(e_t)|\mathcal{F}_s] = \frac{1}{t-s-m_0}\xi_s$; otherwise, we have $|E[U(x_t)f(e_t)|\mathcal{F}_s]| \le C\frac{1}{\sqrt{t-s-m_0}}\int |U(w)|dwE|f(e_t)|$ where C is an absolutely constant.*

(3) For $0 < t - s \le m_0, s > m_0$ and any function $g(\cdot)$ satisfying the similar condition as f, $|E[U(x_s)f(e_t)g(e_s)]| \le O(1)\frac{1}{\sqrt{s-m_0}}\int |U(x)|dxE|f(e_t)g(e_s)|$.

Proof. (1). Suppose that ϵ_0 and η_0 have densities $h_\epsilon(x)$ and $h_\eta(x)$, respectively.

$$
\begin{aligned}
E[U(x_t)f(e_t)] =& E[U(x^*_{t-m_0} + b(\epsilon_t, \cdots, \epsilon_{t-m_0+1}))f(\wp(\epsilon_t, \cdots, \epsilon_{t-m_0+1}; \eta_t, \cdots \eta_{t-m_1+1}))] \\
=& \int \cdots \int U(\sqrt{t-m_0}w + b(v_1, \cdots, v_{m_0}))f(\wp(v_1, \cdots, v_{m_0}; w_1, \cdots, w_{m_1})) \\
& \times f_{t-m_0}(w)\prod_{i=1}^{m_0} h_\epsilon(v_i)\prod_{i=1}^{m_1} h_\eta(w_i)dwdv_1\cdots dv_{m_0}dw_1\cdots dw_{m_1} \\
=& \frac{1}{\sqrt{t-m_0}}\int \cdots \int U(w)f(\wp(v_1, \cdots, v_{m_0}; w_1, \cdots, w_{m_1})) \\
& \times f_{t-m_0}\left(\frac{w - b(v_1, \cdots, v_{m_0})}{\sqrt{t-m_0}}\right)\prod_{i=1}^{m_0} h_\epsilon(v_i)\prod_{i=1}^{m_1} h_\eta(w_i)dwdv_1\cdots dw_{m_1} \\
=& \frac{1}{\sqrt{t-m_0}}\int \cdots \int U(w)f(\wp(v_1, \cdots, v_{m_0}; w_1, \cdots, w_{m_1})) \\
& \times f_{t-m_0}\left(\frac{w}{\sqrt{t-m_0}}\right)\prod_{i=1}^{m_0} h_\epsilon(v_i)\prod_{i=1}^{m_1} h_\eta(w_i)dwdv_1\cdots dw_{m_1} \\
& + \frac{1}{\sqrt{t-m_0}}\int \cdots \int U(w)f(\wp(v_1, \cdots, v_{m_0}; w_1, \cdots, w_{m_1}))\prod_{i=1}^{m_0} h_\epsilon(v_i)\prod_{i=1}^{m_1} h_\eta(w_i) \\
& \times \left[f_{t-m_0}\left(\frac{w - b(v_1, \cdots, v_{m_0})}{\sqrt{t-m_0}}\right) - f_{t-m_0}\left(\frac{w}{\sqrt{t-m_0}}\right)\right]dwdv_1\cdots dw_{m_1} \\
:=& E[U(x^*_{t-m_0})]E[f(e_t)] + \frac{1}{t-m_0}c_U.
\end{aligned}
$$

Thus, using Lipschitz condition for $f_{t-m_0}(\cdot)$ yields

$$
|c_U| \le C\int |U(w)|dwE|b(\epsilon_t \cdots \epsilon_{t-m_0+1})f(e_t)|.
$$

In particular, if $Ef(e_t) = 0$, $E[U(x_t)f(e_t)] = \frac{1}{t-m_0}c_U$; meanwhile, we may derive that $|E[U(x_t)f(e_t)]| \le \frac{1}{\sqrt{t-m_0}}E|f(e_t)|\int |U(w)|dw$ from the above.

(2). Similarly,

$$
\begin{aligned}
&E[U(x_t)f(e_t)|\mathcal{F}_s] = E[U(x_s^* + x'_{s+1,t-m_0} + b(\epsilon_t, \cdots, \epsilon_{t-m_0+1}))f(e_t)|\mathcal{F}_s] \\
=&\int \cdots \int U(x_s^* + \sqrt{t-s-m_0}w + b(v_1, \cdots, v_{m_0}))f(\wp(v_1, \cdots, v_{m_0}; w_1, \cdots, w_{m_1})) \\
&\times f_{s+1,t-m_0}(w) \prod_{i=1}^{m_0} h_\epsilon(v_i) \prod_{i=1}^{m_1} h_\eta(w_i) dw dv_1 \cdots dw_{m_1} \\
=&\frac{1}{\sqrt{t-s-m_0}} \int \cdots \int U(w) f(\wp(v_1, \cdots, v_{m_0}; w_1, \cdots, w_{m_1})) \prod_{i=1}^{m_0} h_\epsilon(v_i) \prod_{i=1}^{m_1} h_\eta(w_i) \\
&\times f_{s+1,t-m_0}\left(\frac{w - x_s^* - b(v_1, \cdots, v_{t-s-m_0})}{\sqrt{t-s-m_0}}\right) dw dv_1 \cdots dw_{m_1} \\
=&\frac{1}{\sqrt{t-s-m_0}} \int \cdots \int U(w) f(\wp(v_1, \cdots, v_{m_0}; w_1, \cdots, w_{m_1})) \prod_{i=1}^{m_0} h_\epsilon(v_i) \prod_{i=1}^{m_1} h_\eta(w_i) \\
&\times f_{s+1,t-m_0}\left(\frac{w - x_s^*}{\sqrt{t-s-m_0}}\right) dw dv_1 \cdots dw_{m_1} \\
&+ \frac{1}{\sqrt{t-s-m_0}} \int \cdots \int U(w) f(\wp(v_1, \cdots, v_{m_0}; w_1, \cdots, w_{m_1})) \prod_{i=1}^{m_0} h_\epsilon(v_i) \prod_{i=1}^{m_1} h_\eta(w_i) \\
&\times \left[f_{s+1,t-m_0}\left(\frac{w - x_s^* - b(v_1, \cdots, v_{m_0})}{\sqrt{t-s-m_0}}\right) - f_{s+1,t-m_0}\left(\frac{w - x_s^*}{\sqrt{t-s-m_0}}\right) \right] dw dv_1 \cdots dw_{m_1} \\
=&E[U(x_s^* + x'_{s+1,t-m_0})|\mathcal{F}_s] E[f(e_t)] + \frac{1}{t-s-m_0}\xi_s,
\end{aligned}
$$

where using Lipschitz condition for $f_{s+1,t-m_0}(\cdot)$ gives

$$
|\xi_s| \le CE|b(\epsilon_t \cdots \epsilon_{t-m_0+1})f(e_t)| \int |U(w)|dw, \quad a.s.
$$

in which C stands for the constant in Lipschitz condition.

Clearly, if $E[f(e_t)] = 0$, $E[U(x_t)f(e_t)|\mathcal{F}_s] = \frac{1}{t-s-m_0}\xi_s$ almost surely, while $E[f(e_t)] \neq 0$, $|E[U(x_t)f(e_t)|\mathcal{F}_s]| \le C\frac{1}{\sqrt{t-s-m_0}} \int |U(w)|dwE|f(e_t)|$.

(3) Note that $m_0 < s < t < s + m_0$, implying $1 \le s - m_0 < t - m_0 < s$. Notice also that $x_s = x_{s-m_0}^* + b(\epsilon_s, \cdots, \epsilon_{s-m_0+1})$, and $\frac{1}{\sqrt{s-m_0}} x_{s-m_0}^*$ has a density $f_{s-m_0}(x)$. Define $\mathcal{G}_{s-m_0}^t = \sigma(\epsilon_t, \cdots, \epsilon_s, \cdots, \epsilon_{s-m_0}; \eta_j, \forall j)$, so that both e_t and e_s are adapted with $\mathcal{G}_{s-m_0}^t$. Then

$$
\begin{aligned}
&|EU(x_s)f(e_t)g(e_s)| = |E[E(U(x_{s-m_0}^* + b(\epsilon_s, \cdots, \epsilon_{s-m_0+1}))f(e_t)g(e_s)|\mathcal{G}_{s-m_0}^t)]| \\
=&\left| E \int U(\sqrt{s-m_0}w + b(\epsilon_s, \cdots, \epsilon_{s-m_0+1}))f(e_t)g(e_s)f_{s-m_0}(w)dw \right| \\
=&\frac{1}{\sqrt{s-m_0}} \left| E \int U(w)f(e_t)g(e_s)f_{s-m_0}\left(\frac{w - b(\epsilon_s, \cdots, \epsilon_{s-m_0+1})}{\sqrt{s-m_0}}\right) dw \right| \\
\le& O(1)\frac{1}{\sqrt{s-m_0}} \int |U(w)|dwE|f(e_t)g(e_s)|.
\end{aligned}
$$

□

Justification of Assumption B. This is to justify that under the null, $\|\widehat{\theta} - \theta\| = O_P(\zeta_n)$ with $\zeta_n n^{-1/4} = O(1)$ is achievable in the setting of this paper. Consider the estimate of a scalar parameter θ from regression equation

$$y_t = \theta g(x_t) + e_t, \quad t = 1, \cdots, n,$$

where x_t and e_t verify Assumption A, and function $g(\cdot)$ is integrable. Then, we will show

$$\sqrt[4]{n}(\widehat{\theta} - \theta) = \frac{n^{-1/4}\sum_{t=1}^{n} g(x_t)e_t}{n^{-1/2}\sum_{t=1}^{n} g^2(x_t)} = O_P(1).$$

For the denominator, by Theorem 3.2 of Park and Phillips (2001), we have $n^{-1/2}\sum_{t=1}^{n} g^2(x_t) \to_P \int g^2(x)dxL_B(1,0)$ where $L_B(1,0)$ is the local time process of Brownian motion $B(r)$ generated by (2.3). For the nominator, due to integrability of g function, $E|n^{-1/4}\sum_{t=1}^{m_n} g(x_t)e_t| = o(1)$ if $m_n \to \infty$ but $m_n n^{-1/4} \to 0$. In what follows we only consider large t. If e_t is a linear process stipulated in b(i) of Assumption A, using (6), (4) and (3) in Lemma A.3,

$$\begin{aligned}
E\left(n^{-1/4}\sum_{t=m_n}^{n} g(x_t)e_t\right)^2 =& \frac{1}{\sqrt{n}}\sum_{t=m_n}^{n} E[g^2(x_t)e_t^2] + \frac{2}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=m_n}^{t-1} E[g(x_t)e_t g(x_s)e_s] \\
\le& O(1)\frac{1}{\sqrt{n}}\sum_{t=m_n}^{n}\frac{1}{\sqrt{t}}\int g^2(x)dx + \frac{2}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=m_n}^{t-1}\sum_{j=s+1}^{t}\phi_{t-j}E[E(g(x_t)\epsilon_j|\mathcal{F}_s)g(x_s)e_s] \\
&+ \frac{2}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=m_n}^{t-1}\sum_{j=-\infty}^{s}\phi_{t-j}E[E(g(x_t)|\mathcal{F}_s)\epsilon_j g(x_s)e_s] \\
\le& O(1) + \frac{2}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=m_n}^{t-1}\sum_{j=s+1}^{t}\phi_{t-j}\frac{1}{t-s}E[\xi_s g(x_s)e_s] \\
&+ \frac{2}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=m_n}^{t-1}\sum_{j=-\infty}^{s}|\phi_{t-j}|\frac{1}{\sqrt{t-s}}E|\epsilon_j g(x_s)e_s|\int|g(x)|dx \\
=& O(1) + 2T_1 + 2T_2, \quad \text{say.}
\end{aligned}$$

Note that ξ_s is given by (A.4) with replacement U by g and $f(v) = v$, that is,

$$\xi_s = \sqrt{t-s}\iint g(x)vh_\epsilon(v)\left[f_{ts}\left(\frac{x - bv - x_s^*}{\sqrt{t-s}}\right) - f_{ts}\left(\frac{x - x_s^*}{\sqrt{t-s}}\right)\right]dvdx,$$

implying $|\xi_s| \le \int|g(x)|dx$ almost surely by the Lipschitz condition on f_{ts}.

To tackle $E[\xi_s g(x_s)e_s]$ in T_1, noting $x_s^* = x_s + \bar{x}_s$ and $\bar{x}_s = O_P(1)$, it follows from the fact $s \ge m_n$ and Lemma 2 in Renyi (1958, p. 223) that $\frac{1}{\sqrt{s}}x_s$ and $\frac{1}{\sqrt{s}}x_s^*$ have the same density $f_s(x)$ asymptotically. Because of this, in the calculation of $E[\xi_s g(x_s)e_s]$ we treat x_s and x_s^* the same and emphasize this by denoting $\xi_s = \xi(x_s)$. Hence, by (6) of Lemma A.3,

$$\begin{aligned}
|T_1| =& \frac{1}{\sqrt{n}}\left|\sum_{t=m_n+1}^{n}\sum_{s=m_n}^{t-1}\frac{1}{t-s}\sum_{j=s+1}^{t}\phi_{t-j}E[\xi(x_s)g(x_s)e_s]\right| = \frac{1}{\sqrt{n}}\left|\sum_{t=m_n+1}^{n}\sum_{s=m_n}^{t-1}\sum_{j=s+1}^{t}\phi_{t-j}\frac{1}{t-s}\frac{1}{s}c_g\right| \\
\le& O(1)\frac{1}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=m_n}^{t-1}\frac{1}{t-s}\frac{1}{s} = O(1)\frac{1}{\sqrt{n}}\ln^2(n) = o(1),
\end{aligned}$$

where c_g is such that $|c_g| \le O(1)\int |\xi(x)g(x)|dx \le O(1)\left(\int |g(x)|dx\right)^2$. Meanwhile, in view of $\phi_j = O(1)j^{-\gamma_0}$ stipulated in Assumption A,

$$
\begin{aligned}
|T_2| =& \frac{1}{\sqrt{n}} \sum_{t=m_n+1}^{n} \sum_{s=m_n}^{t-1} \sum_{j=-\infty}^{s} |\phi_{t-j}| \frac{1}{\sqrt{t-s}} E|\epsilon_j g(x_s)e_s| \int |g(x)|dx \\
\le& O(1)\frac{1}{\sqrt{n}} \sum_{t=m_n+1}^{n} \sum_{s=m_n}^{t-1} (t-s)^{-\gamma_0+1} \frac{1}{\sqrt{t-s}} \frac{1}{\sqrt{s}} = O(1)n^{-\gamma_0+3/2}\ln(n) = o(1),
\end{aligned}
$$

due to $\gamma_0 > 3/2$, where $E|\epsilon_j g(x_s)e_s| \le \sum_{\ell=-\infty}^{s} |\phi_{s-\ell}| E|g(x_s)\epsilon_j\epsilon_\ell| \le O(1)\frac{1}{\sqrt{s}}$ by (2) and (3) of Lemma A.3.

If e_t has the function form in Assumption A, by Lemma A.4,

$$
\begin{aligned}
&E\left(n^{-1/4}\sum_{t=m_n}^{n} g(x_t)e_t\right)^2 = \frac{1}{\sqrt{n}}\sum_{t=m_n}^{n} E[g^2(x_t)e_t^2] + \frac{2}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=m_n}^{t-1} E[g(x_t)e_t g(x_s)e_s] \\
\le& \frac{1}{\sqrt{n}}\sum_{t=m_n}^{n} \frac{1}{\sqrt{t-m_0}} E[e_t^2]\int g^2(x)dx + \frac{2}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=t-m_0}^{t-1} E[g(x_t)e_t g(x_s)e_s] \\
&+ \frac{2}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=m_n}^{t-m_0-1} E[E(g(x_t)e_t|\mathcal{F}_s)g(x_s)e_s] \\
\le& O(1) + O(1)\frac{1}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=t-m_0}^{t-1} E[|e_t g(x_s)e_s|] + \frac{2}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=m_n}^{t-m_0-1} \frac{1}{t-s} E[\xi_s g(x_s)e_s] \\
\le& O(1) + O(1)\frac{1}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=t-m_0}^{t-1} \frac{1}{\sqrt{s}} + O(1)\frac{1}{\sqrt{n}}\sum_{t=m_n+1}^{n}\sum_{s=m_n}^{t-m_0-1} \frac{1}{t-s}\frac{1}{s} \\
=& O(1) + O(1)\frac{1}{\sqrt{n}}\sum_{t=m_n+1}^{n} (\sqrt{t-1} - \sqrt{t-m_0}) + O(1)\frac{1}{\sqrt{n}}\ln^2(n) \\
\le& O(1) + O(1)\frac{1}{\sqrt{n}}\sum_{t=m_n+2}^{n} \frac{1}{\sqrt{t-m_0}} = O(1),
\end{aligned}
$$

where we have used the same argument as in the first part to deal with $E[\xi_s g(x_s)e_s]$, the boundedness of g function and $E[|e_t g(x_s)e_s|] \le O(1)\frac{1}{\sqrt{s}}$ by (3) of Lemma A.4. □

B Proofs of the main results

Proof of Theorem 3.1. This proof is mainly for the case where e_t is a linear process stipulated by b(i) in Assumption A as an exemplar. Due to Lemma A.4, the proof for the case where e_t takes a functional form stipulated by b(ii) in Assumption A can be shown similarly. Under the hypothesis H_0, $y_t = g(x_t;\theta_0) + e_t$ for all $t = 1, \cdots, n$. Hence, it is easy to rewrite $L_n = L_{1n} + L_{2n} + L_{3n}$, where

$$L_{1n} = \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)e_t e_s, \tag{B.1}$$

$$L_{2n} = 2\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)e_t\widehat{g}(x_s), \tag{B.2}$$

$$L_{3n} \quad = \quad \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\widehat{g}(x_t)\widehat{g}(x_s), \tag{B.3}$$

and $\widehat{g}(x) := g(x;\theta_0) - g(x;\widehat{\theta})$ for any real x.

We first deal with L_{1n}. Observe that

$$\begin{aligned} L_{1n} =& \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)e_s e_t = \sum_{t=1}^{n} \|Z(x_t)\|^2 e_t^2 + 2\sum_{t=2}^{n}\sum_{s=1}^{t-1} Z(x_t)'Z(x_s)e_s e_t \\ =&\sigma_e^2 \sum_{t=1}^{n} \|Z(x_t)\|^2 + \sum_{t=1}^{n} \|Z(x_t)\|^2 (e_t^2 - \sigma_e^2) + 2\sum_{t=2}^{n}\sum_{s=1}^{t-1} Z(x_t)'Z(x_s)e_s e_t \\ =&\sigma_e^2 L'_{1n} + L''_{1n} + 2L'''_{1n} \end{aligned}$$

where $\sigma_e^2 = Ee_t^2$. Using the notations in Lemma A.1,

$$\begin{aligned} \frac{d_n}{nk}L'_{1n} =& \frac{d_n}{nk}\sum_{t=1}^{n} \|Z(x_t)\|^2 = \frac{d_n}{n}\sum_{t=1}^{n} \frac{1}{k}\|Z(x_t)\|^2 = \frac{d_n}{n}\sum_{t=1}^{n} \mathscr{T}_k(x_t) \\ =& \frac{d_n}{n}\sum_{t=1}^{n} \mathscr{T}(x_t) + \frac{d_n}{n}\sum_{t=1}^{n} \left[\mathscr{T}_k(x_t) - \mathscr{T}(x_t)\right]. \end{aligned}$$

It follows from Wang and Phillips (2009a) that, since $\int \mathscr{T}(x)dx = 1$,

$$\frac{d_n}{n}\sum_{t=1}^{n} \mathscr{T}(x_t) = \frac{d_n}{n}\sum_{t=1}^{n} \mathscr{T}(d_n x_{nt}) \to_D L_B(1,0).$$

Additionally, $\frac{d_n}{n}\sum_{t=1}^{n}[\mathscr{T}_k(x_t) - \mathscr{T}(x_t)] \to_P 0$. In fact, noting that x_t/d_t has density function $f_t(x)$ that has a uniform bound, say C, over all x and t, by Lemma A.1 we have

$$\begin{aligned} &E\left|\frac{d_n}{n}\sum_{t=1}^{n}[\mathscr{T}_k(x_t) - \mathscr{T}(x_t)]\right| \le \frac{d_n}{n}\sum_{t=1}^{n} E\left|\mathscr{T}_k(x_t) - \mathscr{T}(x_t)\right| \\ =& \frac{d_n}{n}\sum_{t=1}^{n} \int \left|\mathscr{T}_k(d_t x) - \mathscr{T}(d_t x)\right| \; f_t(x)dx = \frac{d_n}{n}\sum_{t=1}^{n} \frac{1}{d_t}\int \left|\mathscr{T}_k(x) - \mathscr{T}(x)\right| \; f_t(x/d_t)dx \\ \le& C\frac{d_n}{n}\sum_{t=1}^{n} \frac{1}{d_t}\int \left|\mathscr{T}_k(x) - \mathscr{T}(x)\right| dx = O(1)\int \left|\mathscr{T}_k(x) - \mathscr{T}(x)\right| dx \to 0. \end{aligned}$$

Hence, $\frac{d_n}{nk}L'_{1n} \to_D L_B(1,0)$. To complete the proof, it suffices to show that L''_{1n}, L'''_{1n}, L_{2n} and L_{3n} after normalisation all are $o_P(1)$. We shall tackle them one by one.

For universal convenience, let m_n be a sequence such that $m_n^4/n \to 0$ and $m_n \to \infty$ as $n \to \infty$.

(I) For $\frac{d_n}{nk}L''_{1n}$, split the sum into two parts, $t \le m_n$ and $t \ge m_n + 1$. The first part is $o_P(1)$ due to $\|Z(x)\|^2 \le O(1)k$ uniformly in x and the stationarity of e_t. We thus only consider the second part in what follows. Note that

$$\begin{aligned} &E\left(\frac{d_n}{nk}\sum_{t=m_n+1}^{n} \|Z(x_t)\|^2(e_t^2 - \sigma_e^2)\right)^2 \\ =&E\frac{d_n^2}{n^2k^2}\sum_{t=m_n+1}^{n} \|Z(x_t)\|^4(e_t^2 - \sigma_e^2)^2 + 2\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1} \|Z(x_t)\|^2(e_t^2 - \sigma_e^2)\|Z(x_s)\|^2(e_s^2 - \sigma_e^2) \end{aligned}$$

$$=\frac{d_n^2}{n^2k^2}\sum_{t=m_n+1}^{n}E[\|Z(x_t)\|^4e_t^4]-2\sigma_e^2\frac{d_n^2}{n^2k^2}\sum_{t=m_n+1}^{n}E[\|Z(x_t)\|^4e_t^2]+\sigma_e^4\frac{d_n^2}{n^2k^2}\sum_{t=m_n+1}^{n}E\|Z(x_t)\|^4$$
$$+2\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2(e_t^2-\sigma_e^2)\|Z(x_s)\|^2(e_s^2-\sigma_e^2).$$

To begin with the easiest one, noting that x_t/d_t has a uniformly bounded density $f_t(x)$, $\|Z(x)\|^2 \le O(1)k$ uniformly in x and $\int \|Z(x)\|^2 dx = k$ by orthogonality,

$$\frac{d_n^2}{n^2k^2}\sum_{t=m_n+1}^{n}E\|Z(x_t)\|^4 \le O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}E\|Z(x_t)\|^2$$
$$=O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}\int\|Z(d_tx)\|^2f_t(x)dx=O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}\frac{1}{d_t}\int\|Z(x)\|^2f_t\left(\frac{x}{d_t}\right)dx$$
$$\le O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}\frac{1}{d_t}\int\|Z(x)\|^2dx=O(1)\frac{1}{\sqrt{n}}=o(1).$$

Moreover, by virtue of (6) of Lemma A.3,

$$\frac{d_n^2}{n^2k^2}\sum_{t=m_n+1}^{n}E[\|Z(x_t)\|^4e_t^2]\le O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}E[\|Z(x_t)\|^2e_t^2]$$
$$\le O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}\frac{1}{\sqrt{t}}\int\|Z(x)\|^2dx=O(1)\frac{1}{\sqrt{n}}=o(1).$$

Next, to compute $E[\|Z(x_t)\|^4e_t^4]$ in the first term, notice that

$$e_t^4=\left(\sum_{j=-\infty}^{t}\phi_{t-j}\epsilon_j\right)^4=\sum_{j_1,j_2,j_3,j_4=-\infty}^{t}\phi_{t-j_1}\phi_{t-j_2}\phi_{t-j_3}\phi_{t-j_4}\epsilon_{j_1}\epsilon_{j_2}\epsilon_{j_3}\epsilon_{j_4}$$
$$=\sum_{j=-\infty}^{t}\phi_{t-j}^4\epsilon_j^4+6\sum_{j=-\infty}^{t}\sum_{\ell=-\infty,\neq j}^{t}\phi_{t-\ell}^2\epsilon_j^2\phi_{t-j}^2\epsilon_\ell^2+4\sum_{j=-\infty}^{t}\sum_{\ell=-\infty,\neq j}^{t}\phi_{t-j}\epsilon_j\phi_{t-\ell}^3\epsilon_\ell^3$$
$$+\sum_{\substack{j_1,j_2,j_3,j_4=-\infty\\ j_1\neq j_2\neq j_3\neq j_4}}^{t}\phi_{t-j_1}\phi_{t-j_2}\phi_{t-j_3}\phi_{t-j_4}\epsilon_{j_1}\epsilon_{j_2}\epsilon_{j_3}\epsilon_{j_4}.$$

Hence, by (2) of Lemma A.3, $E[\|Z(x_t)\|^4\epsilon_j^4] \le O(1)E[\epsilon_j^4]\frac{1}{\sqrt{t}}\int\|Z(x)\|^4dx = O(1)\frac{1}{\sqrt{t}}k^2$, by (3) of Lemma A.3, $E[\|Z(x_t)\|^4\epsilon_j^2\epsilon_j^2] \le O(1)\frac{1}{\sqrt{t}}k^2$, $|E[\|Z(x_t)\|^4\epsilon_j\epsilon_\ell^3]| \le O(1)\frac{1}{\sqrt{t}}k^2$, and with a similar derivation, $|E[\|Z(x_t)\|^4\epsilon_{j_1}\epsilon_{j_2}\epsilon_{j_3}\epsilon_{j_4}]| \le O(1)\frac{1}{\sqrt{t}}k^2$. It follows that

$$\frac{d_n^2}{n^2k^2}\sum_{t=m_n+1}^{n}E[\|Z(x_t)\|^4e_t^4]\le O(1)\frac{1}{nk^2}\sum_{t=m_n+1}^{n}\frac{1}{\sqrt{t}}k^2=O(1)\frac{1}{\sqrt{n}}=o(1).$$

Since $e_t^2-\sigma_e^2=\sum_{j=0}^{\infty}\phi_j^2(\epsilon_{t-j}^2-1)+\sum_{j=0}^{\infty}\sum_{j_1=0,\neq j}^{\infty}\phi_j\phi_{j_1}\epsilon_{t-j}\epsilon_{t-j_1}=\sum_{j=-\infty}^{t}\phi_{t-j}^2(\epsilon_j^2-1)+2\sum_{j=-\infty}^{t-1}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}$,

$$\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2(e_t^2-\sigma_e^2)\|Z(x_s)\|^2(e_s^2-\sigma_e^2)$$

$$
\begin{aligned}
=&\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\left(\sum_{j=-\infty}^{t}\phi_{t-j}^2(\epsilon_j^2-1)+2\sum_{j=-\infty}^{t-1}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}\right)\\
&\times\left(\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2(\epsilon_\ell^2-1)+2\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}\phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1}\right)\\
=&\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{j=-\infty}^{t}\phi_{t-j}^2(\epsilon_j^2-1)\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2(\epsilon_\ell^2-1)\\
&+2\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{j=-\infty}^{t}\phi_{t-j}^2(\epsilon_j^2-1)\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}\phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1}\\
&+2\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2(\epsilon_\ell^2-1)\sum_{j=-\infty}^{t-1}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}\\
&+4\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{j=-\infty}^{t-1}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}\phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1}.
\end{aligned}
$$

We shall deal with each term in the following four parts.

(1a) The first part.

$$
\begin{aligned}
&\frac{d_n^2}{n^2k^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{t}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\phi_{t-j}^2E\|Z(x_t)\|^2\|Z(x_s)\|^2(\epsilon_j^2-1)(\epsilon_\ell^2-1)\\
=&\frac{d_n^2}{n^2k^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=s+1}^{t}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\phi_{t-j}^2E\|Z(x_t)\|^2\|Z(x_s)\|^2(\epsilon_j^2-1)(\epsilon_\ell^2-1)\\
&+\frac{d_n^2}{n^2k^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\phi_{t-j}^2E\|Z(x_t)\|^2\|Z(x_s)\|^2(\epsilon_j^2-1)(\epsilon_\ell^2-1)\\
=&\frac{d_n^2}{n^2k^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=s+1}^{t}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\phi_{t-j}^2E\|Z(x_t)\|^2\|Z(x_s)\|^2(\epsilon_j^2-1)(\epsilon_\ell^2-1)\\
&+\frac{d_n^2}{n^2k^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}\phi_{t-j}^2\phi_{s-j}^2E\|Z(x_t)\|^2\|Z(x_s)\|^2(\epsilon_j^2-1)^2\\
&+\frac{d_n^2}{n^2k^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}\sum_{\ell=-\infty,\neq j}^{s}\phi_{t-j}^2\phi_{s-\ell}^2E\|Z(x_t)\|^2\|Z(x_s)\|^2(\epsilon_j^2-1)(\epsilon_\ell^2-1)\\
:=&I_1+I_2+I_3.
\end{aligned}
$$

For I_1, note by (4) of Lemma A.3 that, $E[\|Z(x_t)\|^2(\epsilon_j^2-1)|\mathcal{F}_s]=\frac{1}{t-s}\xi_s$ where $|\xi_s|\le O(1)\int\|Z(x)\|^2dx=O(1)k$ a.s. Therefore,

$$
\begin{aligned}
|I_1|=&\frac{d_n^2}{n^2k^2}\left|\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=s+1}^{t}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\phi_{t-j}^2E\|Z(x_t)\|^2\|Z(x_s)\|^2(\epsilon_j^2-1)(\epsilon_\ell^2-1)\right|\\
=&\frac{d_n^2}{n^2k^2}\left|O(1)\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=s+1}^{t}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\phi_{t-j}^2\frac{1}{t-s}E[\xi_s\|Z(x_s)\|^2(\epsilon_\ell^2-1)]\right|
\end{aligned}
$$

$$\leq O(1)\frac{1}{nk^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=s+1}^{t}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\phi_{t-j}^2\frac{k}{t-s}E[\|Z(x_s)\|^2|\epsilon_\ell^2-1|]$$

$$\leq O(1)\frac{1}{nk^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=s+1}^{t}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\phi_{t-j}^2\frac{k}{t-s}\frac{1}{\sqrt{s}}\int\|Z(x)\|^2dx$$

$$=O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\frac{1}{\sqrt{s}}=O(1)\frac{1}{\sqrt{n}}\ln(n)=o(1),$$

using (2) of Lemma A.3.

For I_2, observe that

$$I_2=\frac{d_n^2}{n^2k^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}\phi_{t-j}^2\phi_{s-j}^2E\|Z(x_t)\|^2\|Z(x_s)\|^2(\epsilon_j^2-1)^2$$

$$\leq O(1)\frac{1}{nk}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}\phi_{t-j}^2\phi_{s-j}^2E\|Z(x_t)\|^2(\epsilon_j^2-1)^2$$

$$\leq O(1)\frac{1}{nk}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}\phi_{t-j}^2\phi_{s-j}^2\frac{1}{\sqrt{t}}\int\|Z(x)\|^2dx$$

$$\leq O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\frac{1}{\sqrt{t}}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}(t-j)^{-2\gamma_0}=O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\frac{1}{\sqrt{t}}\sum_{s=m_n+1}^{t-1}(t-s)^{-2\gamma_0+1}$$

$$=O(1)\frac{1}{n}\sum_{t=m_n+1}^{n}\frac{1}{\sqrt{t}}(1-(t-m_n)^{-2\gamma_0+2})=O(1)\frac{1}{\sqrt{n}}=o(1)$$

using (2) of Lemma A.3 again and the condition $\gamma_0>3/2$.

As for I_3, notice by (3) of Lemma A.3 that

$$|I_3|\leq\frac{1}{nk}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}\sum_{\ell=-\infty,\neq j}^{s}\phi_{t-j}^2\phi_{s-\ell}^2E\|Z(x_t)\|^2|(\epsilon_j^2-1)(\epsilon_\ell^2-1)|$$

$$\leq O(1)\frac{1}{nk}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}\phi_{t-j}^2\frac{1}{\sqrt{t}}\int\|Z(x)\|^2dx$$

$$=O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\frac{1}{\sqrt{t}}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}(t-j)^{-2\gamma_0}=O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\frac{1}{\sqrt{t}}\sum_{s=m_n+1}^{t-1}(t-s)^{-2\gamma_0+1}$$

$$=O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\frac{1}{\sqrt{t}}(1-(t-m_n)^{-2\gamma_0+2})=O(1)\frac{1}{\sqrt{n}}=o(1).$$

(1b) The second part. Notice that

$$\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{j=-\infty}^{t}\phi_{t-j}^2(\epsilon_j^2-1)\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}\phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1}$$

$$=\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{j=s+1}^{t}\phi_{t-j}^2(\epsilon_j^2-1)\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}\phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1}$$

$$+\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{j=-\infty}^{s}\phi_{t-j}^2(\epsilon_j^2-1)\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}\phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1}.$$

For $j > s$, once again $E[\|Z(x_t)\|^2(\epsilon_j^2-1)|\mathcal{F}_s] = \frac{1}{t-s}\xi_s$ where $|\xi_s| \le O(1)k$ almost surely. Hence,

$$\left|\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{j=s+1}^{t}\phi_{t-j}^2(\epsilon_j^2-1)\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}\phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1}\right|$$

$$=\left|O(1)\frac{1}{nk^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\sum_{j=s+1}^{t}\phi_{t-j}^2\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}\phi_{s-\ell}\phi_{s-\ell_1}E[\xi_s\|Z(x_s)\|^2\epsilon_\ell\epsilon_{\ell_1}]\right|$$

$$\le O(1)\frac{1}{nk}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}|\phi_{s-\ell}\phi_{s-\ell_1}|E[\|Z(x_s)\|^2|\epsilon_\ell\epsilon_{\ell_1}|]$$

$$\le O(1)\frac{1}{nk}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\frac{1}{\sqrt{s}}\int\|Z(x)\|^2dx \le O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\frac{1}{\sqrt{s}}$$

$$=O(1)\frac{1}{n}\sum_{t=m_n+1}^{n}\frac{1}{\sqrt{t}}\ln(t) \le O(1)\frac{1}{n}n^{1/2}\ln(n) = O(1)\frac{1}{n^{1/2}}\ln(n) = o(1),$$

where we have used (3) of Lemma A.3, as well as $\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}|\phi_{s-\ell}\phi_{s-\ell_1}| \le (\sum_{\ell=0}^{\infty}|\phi_\ell|)^2 < \infty$.

Additionally, a similar derivation as (3) of Lemma A.3 yields

$$E[\|Z(x_t)\|^2|\epsilon_j,\epsilon_\ell,\epsilon_{\ell_1}] \le O(1)\frac{1}{\sqrt{t}}\int\|Z(x)\|^2dx = O(1)k\frac{1}{\sqrt{t}}.$$

Therefore, we have

$$\frac{d_n^2}{n^2k^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}\phi_{t-j}^2\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}|\phi_{s-\ell}\phi_{s-\ell_1}|E[\|Z(x_t)\|^2\|Z(x_s)\|^2|(\epsilon_j^2-1)\epsilon_\ell\epsilon_{\ell_1}|]$$

$$\le O(1)\frac{1}{nk}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}\phi_{t-j}^2\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}|\phi_{s-\ell}\phi_{s-\ell_1}|E[\|Z(x_t)\|^2|(\epsilon_j^2-1)\epsilon_\ell\epsilon_{\ell_1}|]$$

$$\le O(1)\frac{1}{nk}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}\phi_{t-j}^2\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}|\phi_{s-\ell}\phi_{s-\ell_1}|\frac{1}{\sqrt{t}}E[|(\epsilon_j^2-1)\epsilon_\ell\epsilon_{\ell_1}|]\int\|Z(x)\|^2dx$$

$$\le O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s}\phi_{t-j}^2\frac{1}{\sqrt{t}} = O(1)\frac{1}{n}\sum_{t=m_n+1}^{n}\frac{1}{\sqrt{t}}\sum_{s=m_n}^{t-1}(t-s)^{-2\gamma_0+1}$$

$$=O(1)\frac{1}{n}\sum_{t=m_n+1}^{n}\frac{1}{\sqrt{t}}(1-(t-m_n)^{-2\gamma_0+2}) = O(1)\frac{1}{\sqrt{n}} = o(1).$$

(1c) We attempt to show the following expectation is $o(1)$,

$$\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2(\epsilon_\ell^2-1)\sum_{j=-\infty}^{t-1}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}$$

$$=\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2(\epsilon_\ell^2-1)$$

$$\times\left(\sum_{j=-\infty}^{s-1}\sum_{j_1=j+1}^{s}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}+\sum_{j=-\infty}^{s}\sum_{j_1=s+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}+\sum_{j=s+1}^{t-1}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}\right)$$

$$=\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2(\epsilon_\ell^2-1)\sum_{j=-\infty}^{s-1}\sum_{j_1=j+1}^{s}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}$$

$$+\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2(\epsilon_\ell^2-1)\sum_{j=-\infty}^{s}\sum_{j_1=s+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}$$

$$+\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2(\epsilon_\ell^2-1)\sum_{j=s+1}^{t-1}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}.$$

Once again, since $E[\|Z(x_t)\|^2|(\epsilon_\ell^2-1)\epsilon_j\epsilon_{j_1}|]\le O(1)\frac{1}{\sqrt{t}}\int\|Z(x)\|^2dx=O(1)\frac{1}{\sqrt{t}}k$, the first item is dealt with as follows.

$$\begin{aligned}
&\frac{d_n^2}{n^2k^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\sum_{j=-\infty}^{s-1}\sum_{j_1=j+1}^{s}|\phi_{t-j}\phi_{t-j_1}|E[\|Z(x_t)\|^2\|Z(x_s)\|^2|(\epsilon_\ell^2-1)\epsilon_j\epsilon_{j_1}|]\\
\le&O(1)\frac{1}{nk}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\sum_{j=-\infty}^{s-1}\sum_{j_1=j+1}^{s}|\phi_{t-j}\phi_{t-j_1}|E[\|Z(x_t)\|^2|(\epsilon_\ell^2-1)\epsilon_j\epsilon_{j_1}|]\\
\le&O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\sum_{j=-\infty}^{s-1}\sum_{j_1=j+1}^{s}|\phi_{t-j}\phi_{t-j_1}|\frac{1}{\sqrt{t}}=O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\frac{1}{\sqrt{t}}\sum_{s=m_n+1}^{t-1}(t-s)^{-2\gamma_0+2}\\
=&O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\frac{1}{\sqrt{t}}t^{-2\gamma_0+3}=O(1)\frac{1}{n}n^{-2\gamma_0+7/2}=O(1)n^{-2\gamma_0+5/2}=o(1)
\end{aligned}$$

for $\gamma_0>3/2$, where we specify the coefficients $\phi_j=O(1)j^{-\gamma_0}$.

By virtue of (4) in Lemma A.3, $E[\|Z(x_t)\|^2\epsilon_{j_1}|\mathcal{F}_s]=O(1)\frac{1}{t-s}\xi_s$ where $|\xi_s|\le O(1)k$ a.s. Thus, for the second item,

$$\begin{aligned}
&\left|\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2(\epsilon_\ell^2-1)\sum_{j=-\infty}^{s}\phi_{t-j}\epsilon_j\sum_{j_1=s+1}^{t}\phi_{t-j_1}\epsilon_{j_1}\right|\\
=&\left|\frac{d_n^2}{n^2k^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\sum_{j=-\infty}^{s}\phi_{t-j}\sum_{j_1=s+1}^{t}\phi_{t-j_1}E[\xi_s\|Z(x_s)\|^2\epsilon_j(\epsilon_\ell^2-1)]\right|\\
\le&O(1)\frac{1}{nk}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\sum_{j=-\infty}^{s}|\phi_{t-j}|\sum_{j_1=s+1}^{t}|\phi_{t-j_1}|E[\|Z(x_s)\|^2|\epsilon_j(\epsilon_\ell^2-1)|]\\
\le&O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\sum_{j=-\infty}^{s}|\phi_{t-j}|\frac{1}{\sqrt{s}}=O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\frac{1}{\sqrt{s}}(t-s)^{-\gamma_0+1}\\
\le&O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\frac{1}{t^{\gamma_0-1/2}}\ln(t)=O(1)\frac{1}{n}(m_n^{-\gamma_0+3/2}-n^{-\gamma_0+3/2})\ln(n)=o(1).
\end{aligned}$$

To calculate the third item, using (5) of Lemma A.3, for $j\ne j_1$ and $1\le s<j,j_1\le t$, we have $E[\|Z(x_t)\|^2\epsilon_j\epsilon_{j_1}|\mathcal{F}_s]=\frac{1}{t-s}\xi_s'$ where ξ_s' is a random variable verifying $|\xi_s'|\le O(1)k$ a.s. Therefore, the third item can be evaluated as

$$\begin{aligned}
&\left|\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2(\epsilon_\ell^2-1)\sum_{j=s+1}^{t}\phi_{t-j}\epsilon_j\sum_{j_1=j+1}^{t}\phi_{t-j_1}\epsilon_{j_1}\right|\\
=&\left|\frac{d_n^2}{n^2k^2}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2\sum_{j=s+1}^{t}\phi_{t-j}\sum_{j_1=j+1}^{t}\phi_{t-j_1}E[\xi_s'\|Z(x_s)\|^2(\epsilon_\ell^2-1)]\right|
\end{aligned}$$

$$
\begin{aligned}
&\leq O(1)\frac{1}{nk}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}^2 E|\|Z(x_s)\|^2(\epsilon_\ell^2-1)| \\
&\leq O(1)\frac{1}{nk}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\frac{1}{\sqrt{s}}\int\|Z(x)\|^2dx \leq O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{t-s}\frac{1}{\sqrt{s}} \\
&\leq O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\frac{1}{\sqrt{t}}\ln(t) = O(1)\frac{1}{n}\sqrt{n}\ln(n) = o(1).
\end{aligned}
$$

(1d) We start to calculate the last term. Because

$$
\begin{aligned}
\sum_{j=-\infty}^{t-1}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1} &= \sum_{j=-\infty}^{s}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1} + \sum_{j=s+1}^{t-1}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1} \\
&= \sum_{j=-\infty}^{s-1}\sum_{j_1=j+1}^{s}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1} + \sum_{j=-\infty}^{s}\sum_{j_1=s+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1} + \sum_{j=s+1}^{t-1}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1},
\end{aligned}
$$

we have

$$
\begin{aligned}
&\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{j=-\infty}^{t-1}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}\phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1} \\
=&\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{j=-\infty}^{s-1}\sum_{j_1=j+1}^{s}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}\phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1} \\
&+\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{j=-\infty}^{s}\sum_{j_1=s+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}\phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1} \\
&+\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\|Z(x_t)\|^2\|Z(x_s)\|^2\sum_{j=s+1}^{t-1}\sum_{j_1=j+1}^{t}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}\sum_{\ell=-\infty}^{s-1}\sum_{\ell_1=\ell+1}^{s}\phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1},
\end{aligned}
$$

which clearly can be shown to be $o(1)$ one by one similar to what we did in the first three parts, as $n\to\infty$. The following is an outline. In the first item $\|Z(x_s)\|^2 \leq O(1)k$, $E\|Z(x_t)\|^2|\epsilon_j\epsilon_{j_1}\epsilon_\ell\epsilon_{\ell_1}| \leq O(1)k\frac{1}{\sqrt{t}}$ and $\sum_{j=-\infty}^{s-1}\sum_{j_1=j+1}^{s}|\phi_{t-j}\phi_{t-j_1}| \leq \left(\sum_{j=-\infty}^{s-1}|\phi_{t-j}|\right)^2 = O(1)(t-s)^{-2\gamma_0+2}$ and all of these facts ensure that the first item is $o(1)$; while in the second item, we may use (4) of Lemma A.3 to derive $E[\|Z(x_t)\|^2\epsilon_{j_1}|\mathcal{F}_s] = \frac{1}{t-s}\xi_s$ and in the third item, we may use (5) of Lemma A.3 to derive $E[\|Z(x_t)\|^2\epsilon_{j_1}\epsilon_j|\mathcal{F}_s] = \frac{1}{t-s}\xi'_s$, which along with a routine calculation for the remaining term yield the desired results.

(II) For L'''_{1n}, we have

$$
\frac{d_n}{nk}L'''_{1n} = \frac{d_n}{nk}\sum_{t=2}^{m_n}\sum_{s=1}^{t-1}Z(x_t)'Z(x_s)e_se_t + \frac{d_n}{nk}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}Z(x_t)'Z(x_s)e_se_t.
$$

However, we may show the first term is $o_P(1)$. Indeed,

$$
\frac{d_n}{nk}\sum_{t=2}^{m_n}\sum_{s=1}^{t-1}E|Z(x_t)'Z(x_s)e_se_t| \leq \frac{d_n}{n}\sum_{t=2}^{m_n}\sum_{s=1}^{t-1}E|e_se_t| \leq O(1)\frac{1}{\sqrt{n}}m_n^2 = o(1).
$$

Hence, only large t is considered in what follows.

$$
\begin{aligned}
&E\left(\frac{d_n}{nk}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}Z(x_t)'Z(x_s)e_se_t\right)^2\\
=&\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\left(\sum_{s=1}^{t-1}Z(x_t)'Z(x_s)e_se_t\right)^2\\
&+2\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\left(\sum_{s=1}^{t_1-1}Z(x_{t_1})'Z(x_s)e_se_{t_1}\right)\left(\sum_{s=1}^{t_2-1}Z(x_{t_2})'Z(x_s)e_se_{t_2}\right)\\
=&\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}[Z(x_t)'Z(x_s)e_se_t]^2\\
&+2\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}Z(x_t)'Z(x_{s_1})e_{s_1}e_t^2Z(x_t)'Z(x_{s_2})e_{s_2}\\
&+2\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=1}^{t_1-1}Z(x_{t_1})'Z(x_{s_1})e_{s_1}e_{t_1}\sum_{s_2=1}^{t_2-1}Z(x_{t_2})'Z(x_{s_2})e_{s_2}e_{t_2}\\
:=&A_1+2A_2+2A_3,\quad \text{say.}
\end{aligned}
$$

Noting that $e_t=\sum_{j=-\infty}^{t}\phi_{t-j}\epsilon_j=\sum_{j=s+1}^{t}\phi_{t-j}\epsilon_j+\sum_{j=-\infty}^{s}\phi_{t-j}\epsilon_j$, we have

$$
\begin{aligned}
A_1=&\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}[Z(x_t)'Z(x_s)e_se_t]^2\\
\le&2\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}[Z(x_t)'Z(x_s)]^2e_s^2\left(\sum_{j=s+1}^{t}\phi_{t-j}\epsilon_j\right)^2\\
&+2\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}[Z(x_t)'Z(x_s)]^2e_s^2\left(\sum_{j=-\infty}^{s}\phi_{t-j}\epsilon_j\right)^2\\
=&2\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}[Z(x_t)'Z(x_s)]^2e_s^2\sum_{j=s+1}^{t}\phi_{t-j}^2\epsilon_j^2\\
&+4\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}[Z(x_t)'Z(x_s)]^2e_s^2\sum_{j=s+2}^{t}\sum_{j_1=s+1}^{j-1}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}\\
&+2\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}[Z(x_t)'Z(x_s)]^2e_s^2\left(\sum_{j=-\infty}^{s}\phi_{t-j}\epsilon_j\right)^2\\
:=&2A_{11}+4A_{12}+2A_{13}.
\end{aligned}
$$

To tackle A_{11}, notice that $E[(Z(x_t)'Z(x_s))^2\epsilon_j^2|\mathcal{F}_s]\le O(1)\frac{1}{\sqrt{t-s}}\int(Z(x)'Z(x_s))^2dx=O(1)\frac{1}{\sqrt{t-s}}\|Z(x_s)\|^2$ by (4) of Lemma A.3 and the orthogonality of the basis. Thus,

$$
A_{11}=\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}[Z(x_t)'Z(x_s)]^2\sum_{j=s+1}^{t}\phi_{t-j}^2\epsilon_j^2e_s^2\le O(1)\frac{1}{nk^2}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{t-s}}E[\|Z(x_s)\|^2e_s^2]
$$

$$\leq O(1)\frac{1}{nk^2}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{t-s}}\frac{1}{\sqrt{s}}\int\|Z(x)\|^2dx = O(1)\frac{1}{k} = o(1),$$

where because of (6) of Lemma A.3, we have $E[\|Z(x_s)\|^2 e_s^2] \leq O(1)\frac{1}{\sqrt{s}}\int\|Z(x)\|^2dx = O(1)\frac{1}{\sqrt{s}}k$.

Regarding of A_{12}, using (5) of Lemma A.3 $E([Z(x_t)'Z(x_s)]^2\epsilon_j\epsilon_{j_1}|\mathcal{F}_s) = \frac{1}{t-s}\xi_s'$ where $|\xi_s'| \leq O(1)\int[Z(x)'Z(x_s)]^2dx = O(1)\|Z(x_s)\|^2$ a.s. Then,

$$|A_{12}| = \left|\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}[Z(x_t)'Z(x_s)]^2\sum_{j=s+2}^{t}\sum_{j_1=s+1}^{j-1}\phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}e_s^2\right|$$

$$=\left|\frac{d_n^2}{n^2k^2}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\frac{1}{t-s}\sum_{j=s+2}^{t}\sum_{j_1=s+1}^{j-1}\phi_{t-j}\phi_{t-j_1}E\xi_s'e_s^2\right| \leq O(1)\frac{1}{nk^2}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\frac{1}{t-s}E[\|Z(x_s)\|^2e_s^2]$$

$$\leq O(1)\frac{1}{nk^2}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\frac{1}{t-s}\frac{1}{\sqrt{s}}\int\|Z(x)\|^2dx = O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}\frac{1}{\sqrt{t}}\ln(t) = O(1)\frac{1}{nk}\sqrt{n}\ln(n) = o(1),$$

where once again we have used (6) of Lemma A.3 for $E[\|Z(x_s)\|^2e_s^2]$.

As for the last term A_{13}, notice that, given $\mathcal{F}_s$, $\frac{1}{\sqrt{t-s}}(x_t - x_s)$ has a uniformly bounded density $f_{ts}(w)$. Hence,

$$\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}[Z(x_t)'Z(x_s)]^2e_s^2\left(\sum_{j=-\infty}^{s}\phi_{t-j}\epsilon_j\right)^2$$

$$=O(1)\frac{1}{nk^2}E\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}E[(Z(x_t)'Z(x_s))^2|\mathcal{F}_s]\left(\sum_{j=-\infty}^{s}\phi_{t-j}\epsilon_j\right)^2e_s^2$$

$$=O(1)\frac{1}{nk^2}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}E\int[Z(\sqrt{t-s}w+x_s)'Z(x_s)]^2f_{ts}(w)dw\left(\sum_{j=-\infty}^{s}\phi_{t-j}\epsilon_j\right)^2e_s^2$$

$$\leq O(1)\frac{1}{nk^2}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{t-s}}E\|Z(x_s)\|^2\left(\sum_{j=-\infty}^{s}\phi_{t-j}\epsilon_j\right)^2e_s^2$$

$$\leq O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{t-s}}E\left(\sum_{j=-\infty}^{s}\phi_{t-j}\epsilon_j\right)^2e_s^2$$

$$\leq O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{t-s}}\sum_{j=-\infty}^{s}\phi_{t-j}^2E[\epsilon_j^2e_s^2]$$

$$+O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{t-s}}\sum_{j=-\infty}^{s}\sum_{\ell=-\infty,\neq j}^{s}\phi_{t-j}\phi_{t-\ell}E[\epsilon_j\epsilon_\ell e_s^2]$$

$$\leq O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{t-s}}\sum_{j=-\infty}^{s}\phi_{t-j}^2 + O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{t-s}}\sum_{j=-\infty}^{s}\sum_{\ell=-\infty,\neq j}^{s}|\phi_{t-j}\phi_{t-\ell}|$$

$$\leq O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{t-s}}(t-s)^{-2\gamma_0+2} = O(1)\frac{1}{nk}\sum_{t=m_n+1}^{n}[1-t^{-2\gamma_0+5/2}] = O(1)\frac{1}{k} = o(1).$$

Now we move on to A_2. Note that

$$e_t^2 = \left(\sum_{j=-\infty}^{t} \phi_{t-j}\epsilon_j\right)^2 = \left(\sum_{j=-\infty}^{s_1} \phi_{t-j}\epsilon_j + \sum_{j=s_1+1}^{t} \phi_{t-j}\epsilon_j\right)^2$$
$$= \left(\sum_{j=s_1+1}^{t} \phi_{t-j}\epsilon_j\right)^2 + \left(\sum_{j=-\infty}^{s_1} \phi_{t-j}\epsilon_j\right)^2 + 2\sum_{j=s_1+1}^{t} \phi_{t-j}\epsilon_j \sum_{j=-\infty}^{s_1} \phi_{t-j}\epsilon_j.$$

Thus,

$$\begin{aligned}
A_2 =& \frac{d_n^2}{n^2k^2} E \sum_{t=m_n+1}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2}) e_t^2 e_{s_1} e_{s_2} \\
=& \frac{d_n^2}{n^2k^2} E \sum_{t=m_n+1}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2}) \left(\sum_{j=s_1+1}^{t} \phi_{t-j}\epsilon_j\right)^2 e_{s_1} e_{s_2} \\
&+ \frac{d_n^2}{n^2k^2} E \sum_{t=m_n+1}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2}) \left(\sum_{j=-\infty}^{s_1} \phi_{t-j}\epsilon_j\right)^2 e_{s_1} e_{s_2} \\
&+ 2\frac{d_n^2}{n^2k^2} E \sum_{t=m_n+1}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2}) \sum_{j=s_1+1}^{t} \phi_{t-j}\epsilon_j \sum_{j=-\infty}^{s_1} \phi_{t-j}\epsilon_j e_{s_1} e_{s_2} \\
:=& A_{21} + A_{22} + A_{23}.
\end{aligned}$$

It follows that

$$\begin{aligned}
A_{21} =& \frac{d_n^2}{n^2k^2} E \sum_{t=m_n+1}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2}) \left(\sum_{j=s_1+1}^{t} \phi_{t-j}\epsilon_j\right)^2 e_{s_1} e_{s_2} \\
=& \frac{d_n^2}{n^2k^2} E \sum_{t=m_n+1}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2}) \sum_{j=s_1+1}^{t} \phi_{t-j}^2\epsilon_j^2 e_{s_1} e_{s_2} \\
&+ 2\frac{d_n^2}{n^2k^2} E \sum_{t=m_n+1}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2}) \sum_{j_1=s_1+2}^{t} \sum_{j_2=s_1+1}^{j_1-1} \phi_{t-j_1}\phi_{t-j_2}\epsilon_{j_1}\epsilon_{j_2} e_{s_1} e_{s_2} \\
=& \frac{d_n^2}{n^2k^2} E \sum_{t=m_n+1}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2}) \\
&\times \sum_{j=s_1+1}^{t} \phi_{t-j}^2\epsilon_j^2 \sum_{\ell_1=-\infty}^{s_1} \phi_{s_1-\ell_1}\epsilon_{\ell_1} \sum_{\ell_2=-\infty}^{s_2} \phi_{s_2-\ell_2}\epsilon_{\ell_2} \\
&+ 2\frac{d_n^2}{n^2k^2} E \sum_{t=m_n+1}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2}) \sum_{j_1=s_1+2}^{t} \sum_{j_2=s_1+1}^{j_1-1} \phi_{t-j_1}\phi_{t-j_2}\epsilon_{j_1}\epsilon_{j_2} e_{s_1} e_{s_2} \\
=& \frac{d_n^2}{n^2k^2} \sum_{t=m_n+1}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} \sum_{j=s_1+1}^{t} \phi_{t-j}^2 \sum_{\ell_1=s_2+1}^{s_1} \phi_{s_1-\ell_1} \sum_{\ell_2=-\infty}^{s_2} \phi_{s_2-\ell_2} \\
&\times E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\epsilon_j^2\epsilon_{\ell_1}\epsilon_{\ell_2}] \\
&+ \frac{d_n^2}{n^2k^2} \sum_{t=m_n+1}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} \sum_{j=s_1+1}^{t} \phi_{t-j}^2 \sum_{\ell_1=-\infty}^{s_2} \phi_{s_1-\ell_1} \sum_{\ell_2=-\infty}^{s_2} \phi_{s_2-\ell_2}
\end{aligned}$$

$$\times E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\epsilon_j^2\epsilon_{\ell_1}\epsilon_{\ell_2}]$$

$$+ 2\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\sum_{j_1=s_1+2}^{t}\sum_{j_2=s_1+1}^{j_1-1}\phi_{t-j_1}\phi_{t-j_2}\epsilon_{j_1}\epsilon_{j_2}e_{s_1}e_{s_2}$$

$$=\frac{d_n^2}{n^2k^2}\sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}\sum_{j=s_1+1}^{t}\phi_{t-j}^2\sum_{\ell_1=s_2+1}^{s_1}\phi_{s_1-\ell_1}\sum_{\ell_2=-\infty}^{s_2}\phi_{s_2-\ell_2}$$

$$\times E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\epsilon_j^2\epsilon_{\ell_1}\epsilon_{\ell_2}]$$

$$+ \frac{d_n^2}{n^2k^2}\sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}\sum_{j=s_1+1}^{t}\phi_{t-j}^2\sum_{\ell=-\infty}^{s_2}\phi_{s_1-\ell}\phi_{s_2-\ell}$$

$$\times E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\epsilon_j^2\epsilon_\ell^2]$$

$$+ \frac{d_n^2}{n^2k^2}\sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}\sum_{j=s_1+1}^{t}\phi_{t-j}^2\sum_{\ell_1=-\infty}^{s_2}\phi_{s_1-\ell_1}\sum_{\ell_2=-\infty,\neq\ell_1}^{s_2}\phi_{s_2-\ell_2}$$

$$\times E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\epsilon_j^2\epsilon_{\ell_1}\epsilon_{\ell_2}]$$

$$+ 2\frac{d_n^2}{n^2k^2}E\sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\sum_{j_1=s_1+2}^{t}\sum_{j_2=s_1+1}^{j_1-1}\phi_{t-j_1}\phi_{t-j_2}\epsilon_{j_1}\epsilon_{j_2}e_{s_1}e_{s_2}$$

$$:=\sum_{i=1}^{4}A_{21}(i), \quad \text{say.}$$

For the following development, we have to use the argument of Deric delta function, a generalized function. Notice that as the probabilities of $x_t = x_{s_1}$ and $x_t = x_{s_2}$ are all zero when $t \neq s_1$ and $t \neq s_2$, the expectations in the above can be computed excluding the regions $x_t = x_{s_1}$ and $x_t = x_{s_2}$. In a usual situation, such exclusion does not make any difference, however, this time it really matters. The detail is as follows.

$A_{21}(1)$ is dealt with first. We partition s_1 into two parts: (1) $2 \le s_1 \le \sqrt{m_n}$, and (2) $\sqrt{m_n}+1 \le s_1 \le t-1$. In situation (1), we repeatedly use (2) of Lemma A.3 to derive

$$\frac{d_n^2}{n^2k^2}\left|\sum_{t=m_n+1}^{n}\sum_{s_1=2}^{\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}\sum_{j=s_1+1}^{t}\phi_{t-j}^2\sum_{\ell_1=s_2+1}^{s_1}\phi_{s_1-\ell_1}\sum_{\ell_2=-\infty}^{s_2}\phi_{s_2-\ell_2}E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\epsilon_j^2\epsilon_{\ell_1}\epsilon_{\ell_2}]\right|$$

$$\le O(1)\frac{1}{nk^2}\sum_{t=m_n+1}^{n}\sum_{s_1=2}^{\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}\frac{1}{\sqrt{t-s_1}}\sum_{\ell_1=s_2+1}^{s_1}|\phi_{s_1-\ell_1}|\sum_{\ell_2=-\infty}^{s_2}|\phi_{s_2-\ell_2}|$$

$$\times\int E|Z(x)'Z(x_{s_1})Z(x)'Z(x_{s_2})\epsilon_{\ell_1}\epsilon_{\ell_2}|dx$$

$$\le O(1)\frac{1}{nk^2}\sum_{t=m_n+1}^{n}\sum_{s_1=2}^{\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}\frac{1}{\sqrt{t-s_1}}\frac{1}{\sqrt{s_1-s_2}}\frac{1}{\sqrt{s_2}}\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(x)'Z(z)|dxdydz$$

$$=O(1)\frac{\sqrt{m_n}}{\sqrt{n}k^2}\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(x)'Z(z)|dxdydz \to 0,$$

due to $\frac{\sqrt{m_n}}{\sqrt{n}k^2} = o(1)$ and (2) of Lemma A.1.

In situation (2), invoking (4) of Lemma A.3,

$$E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\epsilon_j^2|\mathcal{F}_{s_1}] = E[Z(x_{t/j})'Z(x_{s_1})Z(x_{t/j})'Z(x_{s_2})|\mathcal{F}_{s_1}] + \frac{1}{t-s_1}\xi_{s_1}$$

where $x_{t/j} = x_{s_1}^* + x_{ts_1/j}$ and

$$E[Z(x_{t/j})'Z(x_{s_1})Z(x_{t/j})'Z(x_{s_2})|\mathcal{F}_{s_1}] = \frac{1}{\sqrt{t-s_1}}\int Z(x)'Z(x_{s_1})Z(x)'Z(x_{s_2})f_{ts_1/j}\left(\frac{x-x_{s_1}^*}{\sqrt{t-s_1}}\right)dx,$$

$$\xi_{s_1} = \sqrt{t-s_1}\iint Z(x)'Z(x_{s_1})Z(x)'Z(x_{s_2})v^2 h_\epsilon(v)\left[f_{ts_1/j}\left(\frac{x-bv-x_{s_1}^*}{\sqrt{t-s_1}}\right) - f_{ts_1/j}\left(\frac{x-x_{s_1}^*}{\sqrt{t-s_1}}\right)\right]dxdv,$$

because $\frac{1}{\sqrt{t-s_1}}x_{ts_1/j}$ has density $f_{ts_1/j}(x)$. Thus, we have

$$\begin{aligned}
&\frac{d_n^2}{n^2k^2}\Bigg|\sum_{t=m_n+1}^{n}\sum_{s_1=\sqrt{m_n}+1}^{t}\sum_{s_2=1}^{s_1-1}\sum_{j=s_1+1}^{t}\phi_{t-j}^2\sum_{\ell_1=s_2+1}^{s_1}\phi_{s_1-\ell_1}\sum_{\ell_2=-\infty}^{s_2}\phi_{s_2-\ell_2}\\
&\times E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\epsilon_j^2\epsilon_{\ell_1}\epsilon_{\ell_2}]\Bigg|\\
=&\frac{d_n^2}{n^2k^2}\Bigg|\sum_{t=m_n+1}^{n}\sum_{s_1=\sqrt{m_n}+1}^{t}\sum_{s_2=1}^{s_1-1}\sum_{j=s_1+1}^{t}\phi_{t-j}^2\sum_{\ell_1=s_2+1}^{s_1}\phi_{s_1-\ell_1}\sum_{\ell_2=-\infty}^{s_2}\phi_{s_2-\ell_2}\\
&\times E\left\{E[Z(x_{t/j})'Z(x_{s_1})Z(x_{t/j})'Z(x_{s_2})|\mathcal{F}_{s_1}] + \frac{1}{t-s_1}\xi_{s_1}\right\}\Bigg|\\
\le&\frac{d_n^2}{n^2k^2}\Bigg|\sum_{t=m_n+1}^{n}\sum_{s_1=\sqrt{m_n}+1}^{t}\sum_{s_2=1}^{s_1-1}\sum_{j=s_1+1}^{t}\phi_{t-j}^2\sum_{\ell_1=s_2+1}^{s_1}\phi_{s_1-\ell_1}\sum_{\ell_2=-\infty}^{s_2}\phi_{s_2-\ell_2}\\
&\times E\left\{E[Z(x_{t/j})'Z(x_{s_1})Z(x_{t/j})'Z(x_{s_2})|\mathcal{F}_{s_1}]\epsilon_{\ell_1}\epsilon_{\ell_2}\right\}\Bigg|\\
&+\frac{d_n^2}{n^2k^2}\Bigg|\sum_{t=m_n+1}^{n}\sum_{s_1=\sqrt{m_n}+1}^{t}\sum_{s_2=1}^{s_1-1}\sum_{j=s_1+1}^{t}\phi_{t-j}^2\sum_{\ell_1=s_2+1}^{s_1}\phi_{s_1-\ell_1}\sum_{\ell_2=-\infty}^{s_2}\phi_{s_2-\ell_2}\frac{1}{t-s_1}E[\xi_{s_1}\epsilon_{\ell_1}\epsilon_{\ell_2}]\Bigg|
\end{aligned}$$

Since $s_1 > \sqrt{m_n}$ is large, it follows from Lemma 2 of Renyi (1958) that $\frac{1}{\sqrt{s_1}}x_{s_1}^*$ and $\frac{1}{\sqrt{s_1}}x_{s_1}$ have the same distribution function asymptotically. Thus, in this part we treat x_s and x_s^* the same, or simply replace x_s^* by x_s. Then we continue to compute the first part by (4) of Lemma A.3 as follows.

$$\begin{aligned}
&\frac{d_n^2}{n^2k^2}\Bigg|\sum_{t=m_n+1}^{n}\sum_{s_1=\sqrt{m_n}+1}^{t}\sum_{s_2=1}^{s_1-1}\sum_{j=s_1+1}^{t}\phi_{t-j}^2\sum_{\ell_1=s_2+1}^{s_1}\phi_{s_1-\ell_1}\sum_{\ell_2=-\infty}^{s_2}\phi_{s_2-\ell_2}\\
&\times E\left\{E[Z(x_{t/j})'Z(x_{s_1})Z(x_{t/j})'Z(x_{s_2})|\mathcal{F}_{s_1}]\epsilon_{\ell_1}\epsilon_{\ell_2}\right\}\Bigg|\\
=&\frac{d_n^2}{n^2k^2}\Bigg|\sum_{t=m_n+1}^{n}\sum_{s_1=\sqrt{m_n}+1}^{t}\sum_{s_2=1}^{s_1-1}\sum_{j=s_1+1}^{t}\phi_{t-j}^2\sum_{\ell_1=s_2+1}^{s_1}\frac{1}{\sqrt{t-s_1}}\phi_{s_1-\ell_1}\sum_{\ell_2=-\infty}^{s_2}\phi_{s_2-\ell_2}\\
&\times E\int Z(x)'Z(x_{s_1})Z(x)'Z(x_{s_2})f_{ts_1/j}\left(\frac{x-x_{s_1}}{\sqrt{t-s_1}}\right)\epsilon_{\ell_1}\epsilon_{\ell_2}dx\Bigg|\\
=&\frac{d_n^2}{n^2k^2}\Bigg|\sum_{t=m_n+1}^{n}\sum_{s_1=\sqrt{m_n}+1}^{t}\sum_{s_2=1}^{s_1-1}\sum_{j=s_1+1}^{t}\phi_{t-j}^2\sum_{\ell_1=s_2+1}^{s_1}\frac{1}{\sqrt{t-s_1}}\frac{1}{s_1-s_2}\phi_{s_1-\ell_1}\sum_{\ell_2=-\infty}^{s_2}\phi_{s_2-\ell_2}E\xi_{s_2}\epsilon_{\ell_2}\Bigg|
\end{aligned}$$

where

$$\xi_{s_2} = \sqrt{s_1 - s_2} \iiint_{x\neq y} Z(x)'Z(y)Z(x)'Z(x_{s_2}) v h_\epsilon(v) f_{ts_1/j}\left(\frac{x-y}{\sqrt{t-s_1}}\right) \times \left[f_{s_1s_2/\ell_1}\left(\frac{y-x^*_{s_2}-bv}{\sqrt{s_1-s_2}}\right) - f_{s_1s_2/\ell_1}\left(\frac{y-x^*_{s_2}}{\sqrt{s_1-s_2}}\right)\right] dxdydv$$

satisfying $|\xi_{s_2}| \le O(1)\iint_{x\neq y}|Z(x)'Z(y)Z(x)'Z(x_{s_2})|dxdy$ almost surely by Lipschitz condition. Moreover, $|E\xi_{s_2}\epsilon_{\ell_2}| \le O(1)\frac{1}{\sqrt{s_2}}\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(x)'Z(z)|dxdydz$ by (2) of Lemma A.3. Putting all of them together yields

$$\begin{aligned}
&\frac{d_n^2}{n^2k^2}\left|\sum_{t=m_n+1}^{n}\sum_{s_1=\sqrt{m_n}+1}^{t}\sum_{s_2=1}^{s_1-1}\sum_{j=s_1+1}^{t}\phi_{t-j}^2\sum_{\ell_1=s_2+1}^{s_1}\phi_{s_1-\ell_1}\sum_{\ell_2=-\infty}^{s_2}\phi_{s_2-\ell_2}\right.\\
&\left.\times E\left\{E[Z(x_{t/j})'Z(x_{s_1})Z(x_{t/j})'Z(x_{s_2})|\mathcal{F}_{s_1}]\epsilon_{\ell_1}\epsilon_{\ell_2}\right\}\right|\\
\le& O(1)\frac{1}{nk^2}\sum_{t=m_n+1}^{n}\sum_{s_1=\sqrt{m_n}+1}^{t}\sum_{s_2=1}^{s_1-1}\frac{1}{\sqrt{t-s_1}}\frac{1}{s_1-s_2}\frac{1}{\sqrt{s_2}}\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(x)'Z(z)|dxdydz\\
=& O(1)\frac{1}{k^2}\ln(n)\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(x)'Z(z)|dxdydz \to 0
\end{aligned}$$

because $\frac{\ln(n)}{k^2} = o(1)$ and (2) of Lemma A.1.

To finish the proof of $A_{21}(1) = o_P(1)$, we now consider the remaining part. Due to the same reason as before, we treat $x^*_{s_1}$ in ξ_{s_1} the same as x_{s_1}. Thus, by virtue of (1) of Lemma A.3, $E[\xi_{s_1}\epsilon_{\ell_1}|\mathcal{F}_{s_2}] = \frac{1}{s_1-s_2}\xi_{s_2}$ and here $|\xi_{s_2}| \le \iint_{x\neq y}|Z(x)'Z(y)Z(x)'Z(x_{s_2})|dxdy$ almost surely. With further using (2) of Lemma A.3, we have $|E[\xi_{s_2}\epsilon_{\ell_2}]| \le O(1)\frac{1}{\sqrt{s_2}}\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(x)'Z(z)|dxdydz$. Putting all ingredients together will result in the remaining term to be $o_P(1)$ easily by (2) of Lemma A.1.

In $A_{21}(2)$, by (2) of Lemma A.3,

$$|E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\epsilon_j^2\epsilon_\ell^2]| \le \frac{1}{\sqrt{s_1-s_2}}\frac{1}{\sqrt{t-s_1}}\frac{1}{\sqrt{s_2}}\int\left(\int_{y\neq x}|Z(y)'Z(x)|dx\right)^2 dy,$$

which together with $\sum_{\ell=-\infty}^{s_2}|\phi_{s_1-\ell}\phi_{s_2-\ell}| \le O(1)(s_1-s_2)^{-\gamma_0+1}$ and $\gamma_0 > 3/2$ as well as (2) of Lemma A.1 yields that $A_{21}(2) = o(1)$. For the same reason $A_{21}(3) = o(1)$.

As for $A_{21}(4)$, because of $j_1 \neq j_2$ and (5) of Lemma A.3,

$$E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\epsilon_{j_1}\epsilon_{j_2}|\mathcal{F}_{s_1}] = \frac{1}{t-s_1}\xi_{s_1}$$

where $|\xi_{s_1}| \le O(1)\int|Z(x)'Z(x_{s_1})Z(x)'Z(x_{s_2})|dx$ a.s. Then, following a similar fashion (that is, evaluate the expectation of $\xi_{s_1}e_{s_1}e_{s_2}$ by Lemma A.3 and then use (2) of Lemma A.1) we may show $A_{21}(4) = o(1)$.

Next, we shall compute A_3,

$$A_3 = \frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=1}^{t_1-1}Z(x_{t_1})'Z(x_{s_1})e_{s_1}e_{t_1}\sum_{s_2=1}^{t_2-1}Z(x_{t_2})'Z(x_{s_2})e_{s_2}e_{t_2}$$

$$
\begin{aligned}
=&\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}Z(x_{t_1})'Z(x_{t_2})e_{t_2}^2e_{t_1}\sum_{s_2=1}^{t_2-1}Z(x_{t_2})'Z(x_{s_2})e_{s_2}\\
&+\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=1}^{t_2-1}Z(x_{t_1})'Z(x_{s_1})e_{s_1}e_{t_1}\sum_{s_2=1}^{t_2-1}Z(x_{t_2})'Z(x_{s_2})e_{s_2}e_{t_2}\\
&+\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=t_2+1}^{t_1-1}Z(x_{t_1})'Z(x_{s_1})e_{s_1}e_{t_1}\sum_{s_2=1}^{t_2-1}Z(x_{t_2})'Z(x_{s_2})e_{s_2}e_{t_2}\\
=&\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}Z(x_{t_1})'Z(x_{t_2})e_{t_2}^2e_{t_1}\sum_{s_2=1}^{t_2-1}Z(x_{t_2})'Z(x_{s_2})e_{s_2}\\
&+\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=1}^{t_2-1}Z(x_{t_1})'Z(x_{s_1})e_{t_1}Z(x_{t_2})'Z(x_{s_1})e_{s_1}^2e_{t_2}\\
&+2\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=2}^{t_2-1}Z(x_{t_1})'Z(x_{s_1})e_{s_1}e_{t_1}\sum_{s_2=1}^{s_1-1}Z(x_{t_2})'Z(x_{s_2})e_{s_2}e_{t_2}\\
&+\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=t_2+1}^{t_1-1}Z(x_{t_1})'Z(x_{s_1})e_{s_1}e_{t_1}\sum_{s_2=1}^{t_2-1}Z(x_{t_2})'Z(x_{s_2})e_{s_2}e_{t_2}\\
:=&A_{31}+A_{32}+A_{33}+A_{34},\quad \text{say.}
\end{aligned}
$$

Thereinto, A_{31}–A_{34} are partitioned according as $t_2 = s_1$, $t_2 > s_1 = s_2$, $t_2 > s_1 > s_2$ and $s_1 > t_2 > s_2$. To begin with,

$$
\begin{aligned}
A_{31}=&\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_2=1}^{t_2-1}Z(x_{t_1})'Z(x_{t_2})Z(x_{t_2})'Z(x_{s_2})e_{t_1}e_{t_2}^2e_{s_2}\\
=&\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_2=1}^{t_2-1}Z(x_{t_1})'Z(x_{t_2})Z(x_{t_2})'Z(x_{s_2})\sum_{j_1=-\infty}^{t_1}\phi_{t_1-j_1}\epsilon_{j_1}e_{t_2}^2e_{s_2}\\
=&\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_2=1}^{t_2-1}Z(x_{t_1})'Z(x_{t_2})Z(x_{t_2})'Z(x_{s_2})\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\epsilon_{j_1}e_{t_2}^2e_{s_2}\\
&+\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_2=1}^{t_2-1}Z(x_{t_1})'Z(x_{t_2})Z(x_{t_2})'Z(x_{s_2})\sum_{j_1=-\infty}^{t_2}\phi_{t_1-j_1}\epsilon_{j_1}e_{t_2}^2e_{s_2}.
\end{aligned}
$$

Because for $j_1 > t_2$, $E[Z(x_{t_1})'Z(x_{t_2})Z(x_{t_2})'Z(x_{s_2})\epsilon_{j_1}|\mathcal{F}_{t_2}] = \frac{1}{t_1-t_2}\xi_{t_2}$ and

$$
|\xi_{t_2}| \le O(1)\int |Z(x)'Z(x_{t_2})Z(x_{t_2})'Z(x_{s_2})|dx, a.s.
$$

it follows from (6) of Lemma A.3 that

$$
\begin{aligned}
&\left|\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_2=1}^{t_2-1}Z(x_{t_1})'Z(x_{t_2})Z(x_{t_2})'Z(x_{s_2})\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\epsilon_{j_1}e_{t_2}^2e_{s_2}\right|\\
=&\left|\frac{d_n^2}{n^2k^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_2=1}^{t_2-1}\frac{1}{t_1-t_2}\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}E[\xi_{t_2}e_{t_2}^2e_{s_2}]\right|\\
\le&O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_2=1}^{t_2-1}\frac{1}{t_1-t_2}\int E|Z(x)'Z(x_{t_2})Z(x_{t_2})'Z(x_{s_2})e_{t_2}^2e_{s_2}|dx
\end{aligned}
$$

$$
\begin{aligned}
&\leq O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_2=1}^{t_2-1}\frac{1}{t_1-t_2}\frac{1}{\sqrt{t_2-s_2}}\frac{1}{\sqrt{s_2}}\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(y)'Z(z)|dxdydz\\
&=O(1)\frac{1}{nk^2}n\ln(n)\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(y)'Z(z)|dxdydz\to 0
\end{aligned}
$$

by (2) of Lemma A.1 and $\frac{1}{k^2}\ln(n)=o(1)$.

In the second term of A_{31}, in view of (4) of Lemma A.3 we have

$$
\begin{aligned}
&|E[Z(x_{t_1})'Z(x_{t_2})Z(x_{t_2})'Z(x_{s_2})\epsilon_{j_1}e_{t_2}^2e_{s_2}|\\
&\leq O(1)\frac{1}{\sqrt{t_1-t_2}}\frac{1}{\sqrt{t_2-s_2}}\frac{1}{\sqrt{s_2}}\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(y)'Z(z)|dxdydz,
\end{aligned}
$$

although ϵ_{j_1} is contained in e_{t_2} and e_{s_2}. Taking the convergence of $\sum_{j_1=-\infty}^{t_2}|\phi_{t_1-j_1}|\leq O(1)(t_1-t_2)^{-\gamma_0+1}$ and $\gamma_0>3/2$ into account, the second term is not larger than the first term in absolute value. Next,

$$
\begin{aligned}
A_{32}=&\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=1}^{t_2-1}Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_1})e_{t_1}e_{t_2}e_{s_1}^2\\
=&\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=1}^{t_2-1}Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_1})\sum_{j_1=-\infty}^{t_1}\phi_{t_1-j_1}\epsilon_{j_1}e_{t_2}e_{s_1}^2\\
=&\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=1}^{t_2-1}Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_1})\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\epsilon_{j_1}e_{t_2}e_{s_1}^2\\
&+\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=1}^{t_2-1}Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_1})\sum_{j_1=-\infty}^{t_2}\phi_{t_1-j_1}\epsilon_{j_1}e_{t_2}e_{s_1}^2.
\end{aligned}
$$

In the first term we have $E[Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_1})\epsilon_{j_1}|\mathcal{F}_{t_2}]=\frac{1}{t_1-t_2}\xi_{t_2}$ and

$$
|\xi_{t_2}|\leq O(1)\int|Z(x)'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_1})|dx,\ a.s.
$$

and then it follows from Lemma A.3 and the structure of e_t that

$$
\begin{aligned}
&\left|\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=1}^{t_2-1}Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_1})\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\epsilon_{j_1}e_{t_2}e_{s_1}^2\right|\\
&\leq O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=1}^{t_2-1}\frac{1}{t_1-t_2}\frac{1}{\sqrt{t_2-s_2}}\frac{1}{\sqrt{s_2}}\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(y)'Z(z)|dxdydz\\
&=O(1)\frac{1}{nk^2}n\ln(n)\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(y)'Z(z)|dxdydz\to 0
\end{aligned}
$$

by (2) of Lemma A.1 and $\frac{1}{k^2}\ln(n)=o(1)$ again. Similar to A_{31}, the second term in A_{32} in absolute value is smaller than the first term, so that it is an infinitesimal as well. Moreover,

$$
A_{33}=\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_2})e_{s_1}e_{t_1}e_{s_2}e_{t_2}
$$

$$
\begin{aligned}
=&\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_2})\sum_{j_1=-\infty}^{t_1}\phi_{t_1-j_1}\epsilon_{j_1}e_{s_1}e_{s_2}e_{t_2}\\
=&\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_2})\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\epsilon_{j_1}e_{s_1}e_{s_2}e_{t_2}\\
&+\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_2})\sum_{j_1=-\infty}^{t_2}\phi_{t_1-j_1}\epsilon_{j_1}e_{s_1}e_{s_2}e_{t_2}\\
:=&A_{33}(1)+A_{33}(2).
\end{aligned}
$$

Notice that $A_{33}(1)$ is much tougher to be dealt with than $A_{33}(2)$ since in the latter we may use the convergence of $\sum_{j_1=-\infty}^{t_2}|\phi_{t_1-j_1}|\leq O(1)(t_1-t_2)^{-\gamma_0+1}$ and $\gamma_0>3/2$ as what we did before. Thus, only $A_{33}(1)=o(1)$ is shown in what follows.

By (4) of Lemma A.3, $E[Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_2})\epsilon_{j_1}|\mathcal{F}_{t_2}]=\frac{1}{t_1-t_2}\xi_{t_2}$, where

$$
\begin{aligned}
\xi_{t_2}=&\sqrt{t_1-t_2}\iint Z(w)'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_2})vh_\epsilon(v)\\
&\times\left[f_{t_1t_2/j_1}\left(\frac{w-bv-x_{t_2}^*}{\sqrt{t_1-t_2}}\right)-f_{t_1t_2/j_1}\left(\frac{w-x_{t_2}^*}{\sqrt{t_1-t_2}}\right)\right]dvdw,
\end{aligned}
\tag{B.4}
$$

in which $x_{t_2}^*=x_{t_2}+\sum_{i=t_2+1}^{t_1}\sum_{a=-\infty}^{t_2}\psi_{i-a}\epsilon_a$. Here, as $\sum_{i=t_2+1}^{t_1}\sum_{a=-\infty}^{t_2}\psi_{i-a}\epsilon_a=O_P(1)$, when t_1-t_2 is large, $\frac{1}{\sqrt{t_1-t_2}}\sum_{i=t_2+1}^{t_1}\sum_{a=-\infty}^{t_2}\psi_{i-a}\epsilon_a=o_P(1)$. Thereby, $\frac{1}{\sqrt{t_1-t_2}}x_{t_2}^*$ is asymptotically equal to $\frac{1}{\sqrt{t_1-t_2}}x_{t_2}$ in distribution according to Lemma 2 of Renyi (1958, p.223). That is, we may focus on $t_1-t_2>\sqrt{m_n}$ in the following calculation,

$$
\begin{aligned}
&\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-\sqrt{m_n}}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\frac{1}{t_1-t_2}\xi_{t_2}e_{s_1}e_{s_2}e_{t_2}\\
=&\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-\sqrt{m_n}}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\sum_{j_2=s_1+1}^{t_2}\phi_{t_2-j_2}\frac{1}{t_1-t_2}\xi_{t_2}\epsilon_{j_2}e_{s_1}e_{s_2}\\
&+\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-\sqrt{m_n}}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\sum_{j_2=-\infty}^{s_1}\phi_{t_2-j_2}\frac{1}{t_1-t_2}\xi_{t_2}\epsilon_{j_2}e_{s_1}e_{s_2},
\end{aligned}
$$

where in ξ_2, $x_{t_2}^*$ has been replaced by x_{t_2} in (B.4). For brevity we specify $\xi_{t_2}=\xi_{t_2}(x_{t_2},x_{s_1},x_{s_2})$. By virtue of (4) of Lemma A.3 again, in the first term we derive that $E[\xi_{t_2}\epsilon_{j_2}|\mathcal{F}_{s_1}]=\frac{1}{t_2-s_1}\xi_{s_1}$, where

$$
\begin{aligned}
|\xi_{s_1}|=&\sqrt{t_2-s_1}\left|\iint_{x\neq w}v_1h_\epsilon(v_1)\xi_{t_2}(x,x_{s_1},x_{s_2})\left[f_{t_2s_1}\left(\frac{x-x_{s_1}^*-b_1v_1}{\sqrt{t_2-s_1}}\right)-f_{t_2s_1}\left(\frac{x-x_{s_1}^*}{\sqrt{t_2-s_1}}\right)\right]dv_1dx\right|\\
\leq&O(1)\iint_{x\neq w}|v_1^2h_\epsilon(v_1)\xi_{t_2}(x,x_{s_1},x_{s_2})|dv_1dx\\
\leq&O(1)\int\cdots\int_{x\neq w}|Z(w)'Z(x_{s_1})Z(x)'Z(x_{s_2})|v^2h_\epsilon(v)dvdwv_1^2h_\epsilon(v_1)dv_1dx\\
=&O(1)\iint_{x\neq w}|Z(w)'Z(x_{s_1})Z(x)'Z(x_{s_2})|dwdx.
\end{aligned}
$$

Then we compute the following expectation exploiting (2) and (6) of Lemma A.3 and the structure of e_{s_1},

$$\begin{aligned}
&\iint_{x\neq w} E[|Z(w)'Z(x_{s_1})Z(x)'Z(x_{s_2})e_{s_1}e_{s_2}|]dwdx\\
\leq&\frac{1}{\sqrt{s_1-s_2}}\frac{1}{\sqrt{s_1}}\int\cdots\int_{x\neq w\neq y\neq z}|Z(w)'Z(y)Z(x)'Z(z)|dw\cdots dz\\
=&\frac{1}{\sqrt{s_1-s_2}}\frac{1}{\sqrt{s_1}}\left(\iint_{x\neq y}|Z(x_x)'Z(y)|dxdy\right)^2.
\end{aligned}$$

Finally we have

$$\begin{aligned}
&\left|\frac{d_n^2}{n^2k^2}E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-\sqrt{m_n}}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\sum_{j_2=s_1+1}^{t_2}\phi_{t_2-j_2}\frac{1}{t_1-t_2}\xi_{t_2}\epsilon_{j_2}e_{s_1}e_{s_2}\right|\\
\leq&O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-\sqrt{m_n}}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}\frac{1}{t_1-t_2}\frac{1}{t_2-s_1}\frac{1}{\sqrt{s_1-s_2}}\frac{1}{\sqrt{s_1}}\left(\iint_{x\neq y}|Z(x_x)'Z(y)|dxdy\right)^2\\
=&O(1)\frac{1}{k^2}\ln^2(n)\left(\iint_{x\neq y}|Z(x_x)'Z(y)|dxdy\right)^2\to 0,
\end{aligned}$$

due to (2) of Lemma A.1 and $\frac{1}{k^2}\ln^2(n)\to 0$ as $n\to\infty$ since $k=O(n^\kappa)$ with $\kappa>0$.

Also, for $1\leq t_1-t_2\leq\sqrt{m_n}$, we partition s_1 into two parts: $1\leq s_1\leq t_2-\sqrt{m_n}$ and $t_2-\sqrt{m_n}+1\leq s_1\leq t_2-1$. In the first situation, since $t_2-s_1\geq\sqrt{m_n}$, $\frac{1}{\sqrt{t_2-s_1}}(x^*_{t_2}-x_{s_1})$ and $\frac{1}{\sqrt{t_2-s_1}}(x_{t_2}-x_{s_1})$ have the same distribution asymptotically, and therefore by (4) of Lemma A.3,

$$\begin{aligned}
&\frac{d_n^2}{n^2k^2}\left|E\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=2}^{t_2-\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_2})\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\epsilon_{j_1}e_{s_1}e_{s_2}e_{t_2}\right|\\
=&\frac{d_n^2}{n^2k^2}\left|\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=2}^{t_2-\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\frac{1}{t_1-t_2}E[\xi_{t_2}e_{s_1}e_{s_2}e_{t_2}]\right|\\
\leq&O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=2}^{t_2-\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}\frac{1}{t_1-t_2}\left|E\xi_{t_2}\sum_{j_2=-\infty}^{t_2}\phi_{t_2-j_2}\epsilon_{j_2}e_{s_1}e_{s_2}\right|\\
\leq&O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=2}^{t_2-\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}\frac{1}{t_1-t_2}\sum_{j_2=s_1+1}^{t_2}|\phi_{t_2-j_2}||E[\xi_{t_2}\epsilon_{j_2}e_{s_1}e_{s_2}]|\\
&+O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=2}^{t_2-\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}\frac{1}{t_1-t_2}\sum_{j_2=-\infty}^{s_1}|\phi_{t_2-j_2}||E[\xi_{t_2}\epsilon_{j_2}e_{s_1}e_{s_2}]|\\
\leq&O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=2}^{t_2-\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}\frac{1}{t_1-t_2}\frac{1}{t_2-s_1}E|\xi_{s_1}e_{s_1}e_{s_2}|\\
&+O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=2}^{t_2-\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}\frac{1}{t_1-t_2}\sum_{j_2=-\infty}^{s_1}|\phi_{t_2-j_2}|\\
&\times\frac{1}{\sqrt{t_2-s_1}}\frac{1}{\sqrt{s_1-s_2}}\frac{1}{\sqrt{s_2}}\int\cdots\int_{x\neq y\neq z\neq w}|Z(w)'Z(z)Z(x)'Z(y)|dxdydzdw
\end{aligned}$$

$$
\begin{aligned}
\leq & O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=2}^{t_2-\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}\frac{1}{t_1-t_2}\frac{1}{t_2-s_1}\frac{1}{\sqrt{s_1-s_2}}\frac{1}{\sqrt{s_2}} \\
& \times\int\cdots\int_{x\neq y\neq z\neq w}|Z(w)'Z(z)Z(x)'Z(y)|dxdydzdw \\
& + O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=2}^{t_2-\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}\frac{1}{t_1-t_2}(t_2-s_1)^{-\gamma_0+1} \\
& \times\frac{1}{\sqrt{t_2-s_1}}\frac{1}{\sqrt{s_1-s_2}}\frac{1}{\sqrt{s_2}}\int\cdots\int_{x\neq y\neq z\neq w}|Z(w)'Z(z)Z(x)'Z(y)|dxdydzdw \\
\leq & O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=2}^{t_2-\sqrt{m_n}}\sum_{s_2=1}^{s_1-1}\frac{1}{t_1-t_2}\frac{1}{t_2-s_1}\frac{1}{\sqrt{s_1-s_2}}\frac{1}{\sqrt{s_2}} \\
& \times\int\cdots\int_{x\neq y\neq z\neq w}|Z(w)'Z(z)Z(x)'Z(y)|dxdydzdw \\
= & O(1)\frac{1}{k^2}\ln^2(n)\left(\iint_{x\neq y}|Z(x)'Z(y)|dxdy\right)^2\to 0,
\end{aligned}
$$

using (2) of Lemma A.1, where $\gamma_0 > 3/2$ and $E[\xi_{t_2}\epsilon_{j_2}e_{s_1}e_{s_2}]$ in the second term is evaluated by virtue of Lemma A.3 and the structure of e_s.

In the second situation where $t_2-\sqrt{m_n}+1\leq s_1\leq t_2-1$, simply using (2) and (4) of Lemma A.3 as well as the structure of e_s, we have

$$
\begin{aligned}
& \frac{d_n^2}{n^2k^2}\left|E\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=t_2-\sqrt{m_n}+1}^{t_2-1}\sum_{s_2=1}^{s_1-1}Z(x_{t_1})'Z(x_{s_1})Z(x_{t_2})'Z(x_{s_2})\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\epsilon_{j_1}e_{s_1}e_{s_2}e_{t_2}\right| \\
= & \frac{d_n^2}{n^2k^2}\left|\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=t_2-\sqrt{m_n}+1}^{t_2-1}\sum_{s_2=1}^{s_1-1}\sum_{j_1=t_2+1}^{t_1}\phi_{t_1-j_1}\frac{1}{t_1-t_2}E[\xi_{t_2}e_{s_1}e_{s_2}e_{t_2}]\right| \\
\leq & O(1)\frac{1}{nk^2}\sum_{t_1=m_n+2}^{n}\sum_{t_2=t_1-\sqrt{m_n}+1}^{t_1-1}\sum_{s_1=t_2-\sqrt{m_n}+1}^{t_2-1}\sum_{s_2=1}^{s_1-1}\frac{1}{t_1-t_2}\frac{1}{\sqrt{t_2-s_1}}\frac{1}{\sqrt{s_1-s_2}}\frac{1}{\sqrt{s_2}} \\
& \times\left(\iint_{x\neq y}|Z(x)'Z(y)|dxdy\right)^2 \\
= & O(1)\frac{1}{k^2}\ln(m_n)\sqrt[4]{m_n}\left(\iint_{x\neq y}|Z(x)'Z(y)|dxdy\right)^2\to 0,
\end{aligned}
$$

where $\frac{1}{k^2}\ln(m_n)\sqrt[4]{m_n}=o(1)$ by choosing m_n properly. Thus, $A_{33}=o(1)$ is complete.

A_{34} has all indices different, and the only difference between A_{34} and A_{33} is the interchange of t_2 and s_1. Hence similarly $A_{34}=o(1)$. Therefore, $\frac{d_n}{nk}L'''_{1n}=o_P(1)$.

(III) Next, we shall prove $\frac{d_n}{nk}L_{3n}=o_P(1)$. For any $\delta,\epsilon>0$,

$$
\begin{aligned}
& P\left(\frac{d_n}{nk}|L_{3n}|>\delta\right)\leq P(\|\widehat{\theta}-\theta_0\|>\zeta_n\epsilon)+P\left(\frac{d_n}{nk}|L_{3n}|>\delta,\|\widehat{\theta}-\theta_0\|\leq\zeta_n\epsilon\right) \\
= & P(\|\widehat{\theta}-\theta_0\|>\zeta_n\epsilon)+P\left(\frac{d_n}{nk}|L_{3n}|I(\|\widehat{\theta}-\theta_0\|\leq\zeta_n\epsilon)>\delta\right) \\
\leq & P(\|\widehat{\theta}-\theta_0\|>\zeta_n\epsilon)+\frac{d_n}{nk\delta}E[|L_{3n}|I(\|\widehat{\theta}-\theta_0\|\leq\zeta_n\epsilon)],
\end{aligned}
$$

by virtue of Markov's inequality, where $I(\cdot)$ stands for the conventional indicator function.

As $\|\widehat{\theta}-\theta_0\| = O_P(\zeta_n)$ assumed in Assumption B, $P(\|\widehat{\theta}-\theta_0\| > \zeta_n\epsilon) \to 0$. Using Taylor expansion for $g(x;\theta)$ with respect to θ in a neighborhood of θ_0, we have

$$\widehat{g}(x_t) = g(x_t;\theta_0) - g(x_t;\widehat{\theta}) = l_1(x_t;\theta_0)'(\theta_0-\widehat{\theta}) - \frac{1}{2}(\theta_0-\widehat{\theta})'l_2(x_t;\bar{\theta})(\theta_0-\widehat{\theta})$$

where $\bar{\theta}$ is on the line segment connecting θ_0 and $\widehat{\theta}$. In view of this, it follows that

$$\begin{aligned}
&\frac{d_n}{nk}E(|L_{3n}|I(\|\widehat{\theta}-\theta_0\| \le \zeta_n\epsilon)) = \frac{d_n}{nk}E\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\widehat{g}(x_t)\widehat{g}(x_s)I(\|\widehat{\theta}-\theta_0\| \le \zeta_n\epsilon)\\
=&\frac{d_n}{nk}E\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)l_1(x_t;\theta_0)'(\theta_0-\widehat{\theta})l_1(x_s;\theta_0)'(\theta_0-\widehat{\theta})I(\|\widehat{\theta}-\theta_0\| \le \zeta_n\epsilon)\\
&-\frac{d_n}{nk}E\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)l_1(x_t;\theta_0)'(\theta_0-\widehat{\theta})(\theta_0-\widehat{\theta})'l_2(x_s;\bar{\theta})(\theta_0-\widehat{\theta})I(\|\widehat{\theta}-\theta_0\| \le \zeta_n\epsilon)\\
&+\frac{d_n}{4nk}E\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)(\theta_0-\widehat{\theta})'l_2(x_t;\bar{\theta})(\theta_0-\widehat{\theta})(\theta_0-\widehat{\theta})'l_2(x_s;\bar{\theta})(\theta_0-\widehat{\theta})I(\|\widehat{\theta}-\theta_0\| \le \zeta_n\epsilon)\\
=&T_1+T_2+T_3, \qquad \text{say.}
\end{aligned}$$

Using Cauchy-Schwarz inequality and $\|Z(x)\|^2 \le O(1)k$ uniformly in x, we have

$$\begin{aligned}
0 \le T_1 \le& \epsilon^2\zeta_n^2\frac{d_n}{n}E\sum_{t=1}^{n}\|l_1(x_t;\theta_0)\|^2 + 2\epsilon^2\zeta_n^2\frac{d_n}{n}E\sum_{t=2}^{n}\sum_{s=1}^{t-1}\|l_1(x_s;\theta_0)\|\|l_1(x_t;\theta_0)\|\\
\le& C\epsilon^2\zeta_n^2\frac{d_n}{n}\sum_{t=1}^{n}\frac{1}{\sqrt{t}}\int\|l_1(x;\theta_0)\|^2dx + 2C\epsilon^2\zeta_n^2\frac{d_n}{n}\sum_{t=2}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{t-s}}\frac{1}{\sqrt{s}}\iint\|l_1(x;\theta_0)\|\|l_1(y;\theta_0)\|dxdy\\
\le& O(1)\epsilon^2\zeta_n^2 + O(1)\epsilon^2\zeta_n^2\sqrt{n}\left(\int\|l_1(x;\theta_0)\|dx\right)^2 = O(1)\zeta_n^2\epsilon^2 + O(1)\epsilon^2 = o(1)
\end{aligned}$$

due to $\zeta_n^2\sqrt{n} = O(1)$ and the arbitrariness of ϵ.

Similarly, using Assumption B we have as $n \to \infty$

$$\begin{aligned}
0 \le T_3 \le& \epsilon^4\zeta_n^4\frac{d_n}{n}E\sum_{t=1}^{n}\|l_2(x_t;\bar{\theta})\|^2 + \epsilon^4\zeta_n^4\frac{d_n}{n}E\sum_{t=2}^{n}\sum_{s=1}^{t-1}\|l_2(x_t;\bar{\theta})\|\|l_2(x_s;\bar{\theta})\|\\
\le& O(1)\epsilon^4\zeta_n^4\frac{d_n}{n}\sum_{t=1}^{n}\frac{1}{\sqrt{t}}\int l^2(x)dx + O(1)\epsilon^4\zeta_n^4\frac{d_n}{n}\sum_{t=2}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{s}}\frac{1}{\sqrt{t-s}}\iint l(x)l(y)dxdy\\
\le& O(1)\epsilon^4\zeta_n^4 + O(1)\epsilon^4\zeta_n^4\sqrt{n} \to 0.
\end{aligned}$$

Notice that $T_1 = o_P(1)$ and $T_3 = o_P(1)$ imply $T_2 = o_P(1)$. Hence, it follows that $\frac{d_n}{nk}L_{3n} \to_P 0$ and then from Cauchy-Schwarz inequality that $\frac{d_n}{nk}L_{2n} \to_P 0$ as well. The proof of Theorem 3.1 is finished. □

Proof of Theorem 3.2. We only show the theorem in the case where e_t is a linear process. In view of Lemma A.4, the case where e_t takes a functional form stipulated in b(ii) of Assumption A can be shown similarly.

Under H_1 and Assumption C, we have $m(x_t) = g(x_t;\theta_1) + \triangle_n(x_t) = g(x_t;\theta_1) + \delta_n \triangle(x_t)$.

First, we shall prove the theorem holds if $\theta_0 = \theta_1$. Thus, $\widehat{e}_t = e_t + \delta_n \triangle(x_t) + g(x;\theta_0) - g(x;\widehat{\theta}) = e_t + \delta_n\triangle(x_t) + \widehat{g}(x_t)$ where $\widehat{g}(x) = g(x;\theta_1) - g(x;\widehat{\theta})$ for any real x. Write $L_n = L_{1n} + L_{2n} + L_{3n}$, where

$$\begin{aligned} L_{1n} &= \sum_{t=1}^n \sum_{s=1}^n Z(x_t)' Z(x_s)(e_t + \delta_n \triangle(x_t))(e_s + \delta_n \triangle(x_s)), \\ L_{2n} &= 2\sum_{t=1}^n \sum_{s=1}^n Z(x_t)' Z(x_s)(e_t + \delta_n \triangle(x_t))\widehat{g}(x_s), \\ L_{3n} &= \sum_{t=1}^n \sum_{s=1}^n Z(x_t)' Z(x_s)\widehat{g}(x_t)\widehat{g}(x_s). \end{aligned}$$

Observe that in the proof of Theorem 3.1, we have shown that $\frac{d_n}{nk} L_{3n} = o_P(1)$, and Cauchy-Schwarz inequality implies that $|L_{2n}| \le 2\sqrt{L_{1n}L_{3n}}$. Thus, to fulfill the proof, it suffices to show that $\frac{d_n}{nk} L_{1n} \to_P \infty$. To begin with,

$$\begin{aligned} L_{1n} =& \delta_n^2 \sum_{t=1}^n \sum_{s=1}^n Z(x_t)' Z(x_s)\triangle(x_t)\triangle(x_s) + 2\delta_n \sum_{t=1}^n \sum_{s=1}^n Z(x_t)' Z(x_s)\triangle(x_t) e_s \\ &+ \sum_{t=1}^n \sum_{s=1}^n Z(x_t)' Z(x_s) e_t e_s := L'_{1n} + 2L''_{1n} + L'''_{1n}, \quad \text{say.} \end{aligned}$$

From the proof of Theorem 3.1, we have $\frac{d_n}{nk} L'''_{1n} \to_D \sigma_e^2 L_B(1,0)$. Rewrite $L''_{1n} = L''_{1n1} + L''_{1n2} + L''_{1n3}$, where

$$\begin{aligned} L''_{1n1} &= \delta_n \sum_{t=1}^n \|Z(x_t)\|^2 \triangle(x_t) e_t, \\ L''_{1n2} &= \delta_n \sum_{t=2}^n \sum_{s=1}^{t-1} Z(x_t)' Z(x_s)\triangle(x_t) e_s, \\ L''_{1n3} &= \delta_n \sum_{t=1}^{n-1} \sum_{s=t+1}^{n} Z(x_t)' Z(x_s)\triangle(x_t) e_s. \end{aligned}$$

Note that $\frac{d_n}{nk} L''_{1n1} = o_P(1)$. In fact, by (4) of Lemma A.3 and the boundedness of $\triangle(x)$,

$$\begin{aligned} E\left(\frac{d_n}{nk} L''_{1n1}\right)^2 &= \frac{d_n^2}{n^2k^2}\delta_n^2 E\left(\sum_{t=1}^n Z(x_t)' Z(x_t)\triangle(x_t) e_t\right)^2 \\ =& \frac{d_n^2}{n^2k^2}\delta_n^2 E\sum_{t=1}^n \|Z(x_t)\|^4 \triangle^2(x_t) e_t^2 + 2\frac{d_n^2}{n^2k^2}\delta_n^2 E \sum_{t=2}^n \sum_{s=1}^{t-1} \|Z(x_t)\|^2 \triangle(x_t) e_t \|Z(x_s)\|^2 \triangle(x_s) e_s \\ \le& O(1)\frac{1}{n}\delta_n^2 \sum_{t=1}^n E[e_t^2] + 2\frac{d_n^2}{n^2k^2}\delta_n^2 E \sum_{t=2}^n \sum_{s=1}^{t-1} \|Z(x_t)\|^2 \triangle(x_t) \|Z(x_s)\|^2 \triangle(x_s) \sum_{j=-\infty}^{t} \phi_{t-j}\epsilon_j \sum_{\ell=-\infty}^{s} \phi_{s-\ell}\epsilon_\ell \\ \le& O(1)\delta_n^2 + 2\frac{d_n^2}{n^2k^2}\delta_n^2 \sum_{t=2}^n \sum_{s=1}^{t-1} \sum_{j=s+1}^{t} \phi_{t-j} \sum_{\ell=-\infty}^{s} \phi_{s-\ell} E[\|Z(x_t)\|^2 \triangle(x_t) \|Z(x_s)\|^2 \triangle(x_s)\epsilon_j\epsilon_\ell] \\ &+ 2\frac{d_n^2}{n^2k^2}\delta_n^2 \sum_{t=2}^n \sum_{s=1}^{t-1} \sum_{j=-\infty}^{s} \phi_{t-j} \sum_{\ell=-\infty}^{s} \phi_{s-\ell} E[\|Z(x_t)\|^2 \triangle(x_t) \|Z(x_s)\|^2 \triangle(x_s)\epsilon_j\epsilon_\ell] \end{aligned}$$

$$=O(1)\delta_n^2 + 2\frac{d_n^2}{n^2k^2}\delta_n^2\sum_{t=2}^{n}\sum_{s=1}^{t-1}\frac{1}{t-s}\sum_{j=s+1}^{t}\phi_{t-j}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}E[\xi_s\|Z(x_s)\|^2\triangle(x_s)\epsilon_\ell]$$
$$+ 2\frac{d_n^2}{n^2k^2}\delta_n^2\sum_{t=2}^{n}\sum_{s=1}^{t-1}\sum_{j=-\infty}^{s}\phi_{t-j}\sum_{\ell=-\infty}^{s}\phi_{s-\ell}E[\|Z(x_t)\|^2\triangle(x_t)\|Z(x_s)\|^2\triangle(x_s)\epsilon_j\epsilon_\ell],$$

where $|\xi_s| \le O(1)\int\|Z(x)\|^2|\triangle(x)|dx \le O(1)k$ a.s. and therefore $|E[\xi_s\|Z(x_s)\|^2\triangle(x_s)\epsilon_\ell]| \le O(1)\frac{1}{\sqrt{s}}k^2$ by (2) of Lemma A.3. This ensures that the second term is $o(1)$. Meanwhile, we simply use (2) and (4) of Lemma A.3 to derive that

$$|E[\|Z(x_t)\|^2\triangle(x_t)\|Z(x_s)\|^2\triangle(x_s)\epsilon_j\epsilon_\ell]|$$
$$\le O(1)\frac{1}{\sqrt{t-s}}\int\|Z(x)\|^2|\triangle(x)|dx\frac{1}{\sqrt{s}}\int\|Z(x)\|^2|\triangle(x)|dx \le O(1)\frac{1}{\sqrt{t-s}}\frac{1}{\sqrt{s}}k^2,$$

no matter where ϵ_j and ϵ_ℓ are located. This, together with the convergence of $\sum_{j=-\infty}^{s}|\phi_{t-j}| \le O(1)(t-s)^{-\gamma_0+1}$ and $\gamma_0 > 3/2$, entails the last term to be infinitesimal.

Note also that

$$\frac{d_n}{nk}L''_{1n2} = \frac{d_n}{nk}\delta_n\sum_{t=2}^{m_n}\sum_{s=1}^{t-1}Z(x_t)'Z(x_s)\triangle(x_t)e_s + \frac{d_n}{nk}\delta_n\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}Z(x_t)'Z(x_s)\triangle(x_t)e_s,$$

where the first term is $o_P(1)$, given a certain condition on m_n. In effect,

$$\frac{d_n}{nk}\delta_n\sum_{t=2}^{m_n}\sum_{s=1}^{t-1}E|Z(x_t)'Z(x_s)\triangle(x_t)e_s| \le O(1)\frac{1}{\sqrt{n}}\delta_n\sum_{t=2}^{m_n}\sum_{s=1}^{t-1}E|e_s| = O(1)\frac{1}{\sqrt{n}}\delta_n m_n^2 = o(1)$$

as long as $\frac{1}{\sqrt{n}}m_n^2 = O(1)$. Due to this reason we only consider the case where $t > m_n$.

$$E\left(\frac{d_n}{nk}\delta_n\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}Z(x_t)'Z(x_s)\triangle(x_t)e_s\right)^2 = \frac{d_n^2}{n^2k^2}\delta_n^2E\left(\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}Z(x_t)'Z(x_s)\triangle(x_t)e_s\right)^2$$
$$=\frac{d_n^2}{n^2k^2}\delta_n^2E\sum_{t=m_n+1}^{n}\left(\sum_{s=1}^{t-1}Z(x_t)'Z(x_s)\triangle(x_t)e_s\right)^2$$
$$+ 2\frac{d_n^2}{n^2k^2}\delta_n^2E\sum_{t_1=m_n+2}^{n}\sum_{s_1=1}^{t_1-1}Z(x_{t_1})'Z(x_{s_1})\triangle(x_{t_1})e_{s_1}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_2=1}^{t_2-1}Z(x_{t_2})'Z(x_{s_2})\triangle(x_{t_2})e_{s_2}$$
$$=\frac{d_n^2}{n^2k^2}\delta_n^2E\sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\left(Z(x_t)'Z(x_s)\triangle(x_t)e_s\right)^2$$
$$+ 2\frac{d_n^2}{n^2k^2}\delta_n^2E\sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\triangle^2(x_t)e_{s_1}e_{s_2}$$
$$+ 2\frac{d_n^2}{n^2k^2}\delta_n^2E\sum_{t_1=m_n+2}^{n}\sum_{t_2=m_n+1}^{t_1-1}\sum_{s_2=1}^{t_2-1}Z(x_{t_1})'Z(x_{s_2})\triangle(x_{t_1})Z(x_{t_2})'Z(x_{s_2})\triangle(x_{t_2})e_{t_2}e_{s_2}$$
$$+ 2\frac{d_n^2}{n^2k^2}\delta_n^2E\sum_{t_1=m_n+3}^{n}\sum_{t_2=m_n+2}^{t_1-1}\sum_{s=1}^{t_2-1}Z(x_{t_1})'Z(x_s)\triangle(x_{t_1})Z(x_{t_2})'Z(x_s)\triangle(x_{t_2})e_s^2$$
$$+ 4\frac{d_n^2}{n^2k^2}\delta_n^2E\sum_{t_1=m_n+3}^{n}\sum_{t_2=m_n+2}^{t_1-1}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}Z(x_{t_1})'Z(x_{s_1})\triangle(x_{t_1})e_{s_1}Z(x_{t_2})'Z(x_{s_2})\triangle(x_{t_2})e_{s_2}$$

$$:= \sum_{i=1}^{5} T_i, \quad \text{say.}$$

Observe that, given $\mathcal{F}_s$, $\frac{1}{\sqrt{t-s}}(x_t - x_s)$ has a uniformly bounded density $f_{ts}(x)$, and hence

$$\begin{aligned}
T_1 =& \frac{d_n^2}{n^2k^2}\delta_n^2 E \sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1}\big(Z(x_t)'Z(x_s)\triangle(x_t)e_s\big)^2 \\
=& \frac{d_n^2}{n^2k^2}\delta_n^2 \sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1} E\{[E(Z(x_t)'Z(x_s)\triangle(x_t))^2|\mathcal{F}_s]e_s^2\} \\
\leq& \frac{d_n^2}{n^2k^2}\delta_n^2 \sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1} \frac{1}{\sqrt{t-s}} E\left[\int (Z(x)'Z(x_s)\triangle(x))^2 dx e_s^2\right] \\
\leq& O(1)\frac{1}{nk^2}\delta_n^2 \sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1} \frac{1}{\sqrt{t-s}} E[\|Z(x_s)\|^2 e_s^2] \\
\leq& O(1)\frac{1}{nk^2}\delta_n^2 \sum_{t=m_n+1}^{n}\sum_{s=1}^{t-1} \frac{1}{\sqrt{t-s}}\frac{1}{\sqrt{s}} \int \|Z(x)\|^2 dx = O(1)\frac{1}{k}\delta_n^2 = o(1),
\end{aligned}$$

by the orthogonality of the basis and (6) of Lemma A.3.

Ignoring the unimportant constants in each T_i in the following computation, we have

$$\begin{aligned}
T_2 =& \frac{d_n^2}{n^2k^2}\delta_n^2 E \sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\triangle^2(x_t)e_{s_1}e_{s_2} \\
=& \frac{d_n^2}{n^2k^2}\delta_n^2 E \sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\triangle^2(x_t) \sum_{j=-\infty}^{s_1}\phi_{s_1-j}\epsilon_j \sum_{\ell=-\infty}^{s_2}\phi_{s_2-\ell}\epsilon_\ell \\
=& \frac{d_n^2}{n^2k^2}\delta_n^2 E \sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\triangle^2(x_t) \sum_{j=s_2+1}^{s_1}\phi_{s_1-j}\epsilon_j \sum_{\ell=-\infty}^{s_2}\phi_{s_2-\ell}\epsilon_\ell \\
&+ \frac{d_n^2}{n^2k^2}\delta_n^2 E \sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1} Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\triangle^2(x_t) \sum_{j=-\infty}^{s_2}\phi_{s_1-j}\epsilon_j \sum_{\ell=-\infty}^{s_2}\phi_{s_2-\ell}\epsilon_\ell \\
=& \frac{d_n^2}{n^2k^2}\delta_n^2 \sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}\sum_{j=s_2+1}^{s_1}\phi_{s_1-j}\sum_{\ell=-\infty}^{s_2}\phi_{s_2-\ell} E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\triangle^2(x_t)\epsilon_j\epsilon_\ell] \\
&+ \frac{d_n^2}{n^2k^2}\delta_n^2 \sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}\sum_{j=-\infty}^{s_2}\phi_{s_1-j}\sum_{\ell=-\infty}^{s_2}\phi_{s_2-\ell} E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\triangle^2(x_t)\epsilon_j\epsilon_\ell] \\
:=& T_{21} + T_{22}.
\end{aligned}$$

Noting by $t > s_1 > s_2$ that the probabilities that any two of x_t, x_{s_1} and x_{s_2} are equal are all zero, in the calculation of the expectations in T_{21} and T_{22} we shall exclude these regions. Given $\mathcal{F}_{s_1}$, $\frac{1}{\sqrt{t-s_1}}(x_t - x_{s_1})$ has a uniformly bounded density $f_{ts_1}(x)$. Whence, in T_{21},

$$\begin{aligned}
& E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\triangle^2(x_t)|\mathcal{F}_{s_1}] \\
=& \frac{1}{\sqrt{t-s_1}} \int Z(x)'Z(x_{s_1})Z(x)'Z(x_{s_2})\triangle^2(x) f_{ts_1}\left(\frac{x-x_{s_1}}{\sqrt{t-s_1}}\right) dx.
\end{aligned}$$

Then, using (4) of Lemma A.3 yields $E[Z(x)'Z(x_{s_1})Z(x)'Z(x_{s_2})\triangle^2(x)f_{ts_1}\left(\frac{x-x_{s_1}}{\sqrt{t-s_1}}\right)\epsilon_j|\mathcal{F}_{s_2}] = \frac{1}{s_1-s_2}\xi_{s_2}(x)$ for any x, where $|\xi_{s_2}(x)| \le O(1)\int |Z(x)'Z(y)Z(x)'Z(x_{s_2})|\triangle^2(x)dy$. Thus, (2) of Lemma A.3 gives

$$\begin{aligned}
|T_{21}| =& \left|\frac{d_n^2}{n^2k^2}\delta_n^2 \sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}\sum_{j=s_2+1}^{s_1}\phi_{s_1-j}\sum_{\ell=-\infty}^{s_2}\phi_{s_2-\ell}E[Z(x_t)'Z(x_{s_1})Z(x_t)'Z(x_{s_2})\triangle^2(x_t)\epsilon_j\epsilon_\ell]\right| \\
=& \left|\frac{d_n^2}{n^2k^2}\delta_n^2 \sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}\frac{1}{\sqrt{t-s_1}}\frac{1}{s_1-s_2}\sum_{j=s_2+1}^{s_1}\phi_{s_1-j}\sum_{\ell=-\infty}^{s_2}\phi_{s_2-\ell}\int E[\xi_{s_2}(x)\epsilon_\ell]dx\right| \\
\le& O(1)\frac{1}{nk^2}\delta_n^2 \sum_{t=m_n+1}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}\frac{1}{\sqrt{t-s_1}}\frac{1}{s_1-s_2}\frac{1}{\sqrt{s_2}}\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(x)'Z(z)|\triangle^2(x)dxdydz \\
=& O(1)\frac{1}{nk^2}\delta_n^2 n\ln(n)\iiint_{x\neq y\neq z}|Z(x)'Z(y)Z(x)'Z(z)|\triangle^2(x)dxdydz = o(1),
\end{aligned}$$

by (2) of Lemma A.1.

Simply use of (2) of Lemma A.3 gives that T_{22} is smaller that T_{21} in absolute value since $\sum_{j=-\infty}^{s_2}|\phi_{s_1-j}| \le O(1)(s_1-s_2)^{-\gamma_0+1}$ and $\gamma_0 > 3/2$.

T_3 can be calculated similarly as T_2 because the only difference between them is $\triangle(x_t)^2$ in T_2 is replaced by $\triangle(x_{t_1})\triangle(x_{t_2})$ in T_3. Definitely, the replacement does not affect the calculation of the expectation.

Regarding of T_4, by virtue of conditional densities for $\frac{1}{\sqrt{t_1-t_2}}(x_{t_1}-x_{t_2})$ given $\mathcal{F}_{t_2}$, $\frac{1}{\sqrt{t_2-s}}(x_{t_2}-x_s)$ given $\mathcal{F}_s$ and making use of (6) of Lemma A.3 for the remaining expectation, we have

$$\begin{aligned}
|T_4| =& \frac{d_n^2}{n^2k^2}\delta_n^2\left|E\sum_{t_1=m_n+3}^{n}\sum_{t_2=m_n+2}^{t_1-1}\sum_{s=1}^{t_2-1}Z(x_{t_1})'Z(x_s)\triangle(x_{t_1})Z(x_{t_2})'Z(x_s)\triangle(x_{t_2})e_s^2\right| \\
\le& O(1)\frac{1}{nk^2}\delta_n^2\sum_{t_1=m_n+3}^{n}\sum_{t_2=m_n+2}^{t_1-1}\sum_{s=1}^{t_2-1}\frac{1}{\sqrt{t_1-t_2}}\frac{1}{\sqrt{t_2-s}}\frac{1}{\sqrt{s}} \\
&\times\iiint|Z(x)'Z(z)\triangle(x)Z(y)'Z(z)\triangle(y)|dxdydz \\
=& O(1)\frac{1}{k^2}\delta_n^2\sqrt{n}\int dz\left(\int|Z(x)'Z(z)\triangle(x)|dx\right)^2 \\
\le& O(1)\frac{1}{k^2}\delta_n^2\sqrt{n}\int dz\int|Z(x)'Z(z)|^2dx\int|\triangle(x)|^2dx = O(1)\frac{1}{k}\delta_n^2\sqrt{n},
\end{aligned}$$

in view of Cauchy-Schwartz inequality. Lastly,

$$\begin{aligned}
T_5 =& \frac{d_n^2}{n^2k^2}\delta_n^2E\sum_{t_1=m_n+3}^{n}\sum_{t_2=m_n+2}^{t_1-1}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}Z(x_{t_1})'Z(x_{s_1})\triangle(x_{t_1})e_{s_1}Z(x_{t_2})'Z(x_{s_2})\triangle(x_{t_2})e_{s_2} \\
=& \frac{d_n^2}{n^2k^2}\delta_n^2E\sum_{t_1=m_n+3}^{n}\sum_{t_2=m_n+2}^{t_1-1}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}Z(x_{t_1})'Z(x_{s_1})\triangle(x_{t_1})Z(x_{t_2})'Z(x_{s_2})\triangle(x_{t_2}) \\
&\times\sum_{j=-\infty}^{s_1}\phi_{s_1-j}\epsilon_j\sum_{\ell=-\infty}^{s_2}\phi_{s_2-\ell}\epsilon_\ell \\
=& \frac{d_n^2}{n^2k^2}\delta_n^2\sum_{t_1=m_n+3}^{n}\sum_{t_2=m_n+2}^{t_1-1}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}\sum_{j=s_2+1}^{s_1}\phi_{s_1-j}\sum_{\ell=-\infty}^{s_2}\phi_{s_2-\ell}
\end{aligned}$$

$$\times E[Z(x_{t_1})'Z(x_{s_1})\triangle(x_{t_1})Z(x_{t_2})'Z(x_{s_2})\triangle(x_{t_2})\epsilon_j\epsilon_\ell]$$
$$+\frac{d_n^2}{n^2k^2}\delta_n^2\sum_{t_1=m_n+3}^{n}\sum_{t_2=m_n+2}^{t_1-1}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}\sum_{j=-\infty}^{s_2}\phi_{s_1-j}\sum_{\ell=-\infty}^{s_2}\phi_{s_2-\ell}$$
$$\times E[Z(x_{t_1})'Z(x_{s_1})\triangle(x_{t_1})Z(x_{t_2})'Z(x_{s_2})\triangle(x_{t_2})\epsilon_j\epsilon_\ell]$$
$$:=T_{51}+T_{52}.$$

Once again, to use the Deric delta function in what follows the regions that any two of x_{t_1}, x_{t_2}, x_{s_1} and x_{s_2} are quale are excluded in the computation. To tackle T_{51}, again we engage the conditional densities $f_{t_1t_2}(w)$ and $f_{t_2s_1}(x)$ for $\frac{1}{\sqrt{t_1-t_2}}(x_{t_1}-x_{t_2})$ given $\mathcal{F}_{t_2}$, $\frac{1}{\sqrt{t_2-s_1}}(x_{t_2}-x_{s_1})$ given $\mathcal{F}_{s_1}$, respectively, as well as (4) of Lemma A.3 to derive

$$\begin{aligned}&E[Z(x_{t_1})'Z(x_{s_1})\triangle(x_{t_1})Z(x_{t_2})'Z(x_{s_2})\triangle(x_{t_2})\epsilon_j\epsilon_\ell]\\=&\frac{1}{\sqrt{t_1-t_2}}\frac{1}{\sqrt{t_2-s_1}}\iint_{w\neq x}E[Z(w)'Z(x_{s_1})\triangle(w)Z(x)'Z(x_{s_2})\triangle(x)\\&\times f_{t_1t_2}\left(\frac{w-x}{\sqrt{t_1-t_2}}\right)f_{t_2s_1}\left(\frac{x-x_{s_1}}{\sqrt{t_2-s_1}}\right)\epsilon_j\epsilon_\ell]dwdx\\=&\frac{1}{\sqrt{t_1-t_2}}\frac{1}{\sqrt{t_2-s_1}}\frac{1}{s_1-s_2}E[\xi_{s_2}\epsilon_\ell],\end{aligned}$$

where ξ_{s_2} satisfies

$$|\xi_{s_2}|\leq\iiint_{w\neq x\neq y}|Z(w)'Z(y)\triangle(w)Z(x)'Z(x_{s_2})\triangle(x)|dwdxdy,\quad a.s.$$

and thus by (2) of Lemma A.3, we have

$$\begin{aligned}|E[\xi_{s_2}\epsilon_\ell]|\leq&\iiint E|Z(w)'Z(y)\triangle(w)Z(x)'Z(x_{s_2})\triangle(x)\epsilon_\ell|dwdxdy\\\leq&O(1)\frac{1}{\sqrt{s_2}}\iiiint_{w\neq x\neq y\neq z}|Z(w)'Z(y)\triangle(w)Z(x)'Z(z)\triangle(x)|dwdxdydz\\=&O(1)\frac{1}{\sqrt{s_2}}\left(\iint_{w\neq y}|Z(w)'Z(y)\triangle(w)|dwdy\right)^2.\end{aligned}$$

Therefore,

$$\begin{aligned}|T_{51}|\leq&O(1)\frac{1}{nk^2}\delta_n^2\sum_{t_1=m_n+3}^{n}\sum_{t_2=m_n+2}^{t_1-1}\sum_{s_1=2}^{t_2-1}\sum_{s_2=1}^{s_1-1}\frac{1}{\sqrt{t_1-t_2}}\frac{1}{\sqrt{t_2-s_1}}\frac{1}{s_1-s_2}\\&\times\frac{1}{\sqrt{s_2}}\left(\iint_{w\neq y}|Z(w)'Z(y)\triangle(w)|dwdy\right)^2\\=&O(1)\frac{1}{k^2}\sqrt{n}\ln(n)\delta_n^2\left(\iint_{w\neq y}|Z(w)'Z(y)\triangle(w)|dwdy\right)^2=o(1)\frac{1}{k}\sqrt{n}\delta_n^2\end{aligned}$$

by (2) of Lemma A.1.

Because of $\sum_{j=-\infty}^{s_2}|\phi_{s_1-j}|=O(1)(s_1-s_2)^{-\gamma_0+1}$, $\gamma_0>3/2$ and a routine derivation, it is easy to obtain $|T_{52}|=o(1)\frac{1}{k}\sqrt{n}\delta_n^2$ in view of the proof for $|T_{51}|$. This gives $|T_5|=o(1)\frac{1}{k}\sqrt{n}\delta_n^2$.

From all the above and $T_{41} = O(1)\frac{1}{k}\sqrt{n}\delta_n^2$, it follows that $\frac{d_n}{nk}L''_{1n2} = O_P(n^{1/4}k^{-1/2}\delta_n)$. Similar calculation yields the same result for $\frac{d_n}{nk}L''_{1n3}$. Therefore, $\frac{d_n}{nk}L''_{1n} = O_P(n^{1/4}k^{-1/2}\delta_n)$.

Finally, we have as $n \to \infty$

$$\begin{aligned}\frac{d_n}{nk}L'_{1n} =& \delta_n^2 \frac{d_n}{nk}\sum_{t=1}^n\sum_{s=1}^n Z(x_t)'Z(x_s)\triangle(x_t)\triangle(x_s)\\ =&\frac{\delta_n^2 d_n}{nk}\sum_{t=1}^n\sum_{s=1}^n\sum_{i=0}^{k-1}\mathscr{H}_i(x_t)\mathscr{H}_i(x_s)\triangle(x_t)\triangle(x_s) = \frac{\delta_n^2 d_n}{nk}\sum_{i=0}^{k-1}\left(\sum_{t=1}^n \mathscr{H}_i(x_t)\triangle(x_t)\right)^2\\ =&\frac{\delta_n^2 n}{d_n k}\sum_{i=0}^{k-1}\left(\frac{d_n}{n}\sum_{t=1}^n \mathscr{H}_i(d_n x_{tn})\triangle(d_n x_{tn})\right)^2 \geq \frac{\delta_n^2\sqrt{n}}{|\psi|k}\left(\frac{d_n}{n}\sum_{t=1}^n \mathscr{H}_{i_0}(d_n x_{tn})\triangle(d_n x_{tn})\right)^2 \to_P \infty\end{aligned}$$

because $\left(\frac{d_n}{n}\sum_{t=1}^n \mathscr{H}_{i_0}(d_n x_{tn})\triangle(d_n x_{tn})\right)^2 \to_D \left(\int \mathscr{H}_{i_0}(x)\triangle(x)dx L_B(1,0)\right)^2$ by the continuous mapping theorem and $\delta_n^2\sqrt{n}/k \to \infty$ by Assumption C, where $i_0 \geq 0$ is the first integer such that $\int \mathscr{H}_{i_0}(x)\triangle(x)dx \neq 0$. Obviously, such i_0 does exist, otherwise for any i, $\int \mathscr{H}_i(x)\triangle(x)dx = 0$ so that $\triangle(x)$ is orthogonal with every $\mathscr{H}_i(x)$ in Hilbert space $L^2(\mathbb{R})$. This amounts to saying that $\triangle(x)$ is a zero function. Thus, we finish the proof in the case of $\theta_0 = \theta_1$.

We now turn to the case that $\theta_0 \neq \theta_1$. In this case, $\widehat{e}_t = y_t - g(x_t, \widehat{\theta}) = e_t + \delta_n\triangle(x_t) + g(x_t,\theta_1) - g(x_t,\widehat{\theta}) = e_t + \delta_n\triangle(x_t) + \widehat{g}(x_t) + g(x_t,\theta_1) - g(x_t,\theta_0)$. This means that under the alternative we have two departure functions from the null, one is the local departure $\delta_n\triangle(x_t)$ and another one is a global departure $g(x_t,\theta_1) - g(x_t,\theta_0)$.

If $g(x,\theta_1) - g(x,\theta_0) = 0$ almost everywhere on $\mathbb{R}$, the global departure would not play any role since $g(x_t,\theta_1) - g(x_t,\theta_0) = 0$ almost surely. In what follows we suppose there exists a set $S \subset \mathbb{R}$ with positive Lebesgue measure such that $g(x,\theta_1) - g(x,\theta_0) \neq 0$ for any $x \in S$. This gives rise to an extra term in $\frac{d_n}{nk}L_n$,

$$\begin{aligned}&\frac{d_n}{nk}\sum_{t=1}^n\sum_{s=1}^n Z(x_t)'Z(x_s)[g(x_t,\theta_1) - g(x_t,\theta_0)][g(x_t,\theta_1) - g(x_t,\theta_0)]\\ \geq&\frac{\sqrt{n}}{|\psi|k}\left(\frac{d_n}{n}\sum_{t=1}^n \mathscr{H}_{i_0}(d_n x_{tn})[g(d_n x_{tn},\theta_1) - g(d_n x_{tn},\theta_0)]\right)^2 \to_P \infty\end{aligned}$$

because similarly the squared part converges to $\left(\int \mathscr{H}_{i_0}(x)[g(x,\theta_1) - g(x,\theta_0)]dx L_B(1,0)\right)^2$ in distribution by the continuous mapping theorem and $\sqrt{n}/k \to \infty$, where $i_0 \geq 0$ is the first integer such that $\int \mathscr{H}_{i_0}(x)[g(x,\theta_1) - g(x,\theta_0)]dx \neq 0$. Such i_0 does exist as discussed for $\triangle(x)$. Without δ_n^2, this quantity is much bigger than the $\frac{d_n}{nk}L'_{1n}$ in the case of $\theta_0 = \theta_1$. In other words, when $\theta_1 \neq \theta_0$ the global departure would dominate the local departure. Accordingly, the theorem still holds. $\square$

Proof of Lemma 3.1. It follows from Hualde and Robinson (2011) that $\widehat{\rho}$ is consistent, $\widehat{\rho} \to_P \rho_0$. Meanwhile, the series $\sum_{j=0}^{\infty}\psi_j(\widehat{\rho})$ is convergent uniformly in $\widehat{\rho}$ because of (a) in Assumption A and the compactness of $\mathbb{S}$. The uniform convergence gives the continuity of $\widehat{\psi}$ on $\widehat{\rho}$. Hence $\widehat{\psi} \to_P \psi$ and $\widehat{d}_n/d_n \to_P 1$ by the continuous mapping theorem.

Next, we shall show $\widehat{\sigma}_e^2 \to_P \sigma_e^2$. Under the null,

$$\begin{aligned}\widehat{\sigma}^2 =& \frac{1}{n}\sum_{t=1}^n \widehat{e}_t^2 = \frac{1}{n}\sum_{t=1}^n (y_t - g(x_t,\widehat{\theta}))^2 = \frac{1}{n}\sum_{t=1}^n (e_t + g(x_t,\theta) - g(x_t,\widehat{\theta}))^2 \\ =& \frac{1}{n}\sum_{t=1}^n (e_t + \widehat{g}(x_t))^2 = \frac{1}{n}\sum_{t=1}^n e_t^2 + \frac{1}{n}\sum_{t=1}^n \widehat{g}^2(x_t) + 2\frac{1}{n}\sum_{t=1}^n e_t\widehat{g}(x_t),\end{aligned}$$

where $\widehat{g}(x) := g(x,\theta) - g(x,\widehat{\theta})$ for any real x.

To begin with, we shall show that $\frac{1}{n}\sum_{t=1}^n e_t^2 \to_P \sigma_e^2$. First, suppose $e_t = \sum_{j=0}^\infty \phi_j \epsilon_{t-j}$ stipulated in b(i) of Assumption A.

$$\begin{aligned}&E\left(\frac{1}{n}\sum_{t=1}^n e_t^2 - \sigma_e^2\right)^2 = \frac{1}{n^2}\sum_{t=1}^n E(e_t^2-\sigma_e^2)^2 + 2\frac{1}{n^2}E\sum_{t=2}^n\sum_{s=1}^{t-1}(e_t^2-\sigma_e^2)(e_s^2-\sigma_e^2) \\ =& \frac{1}{n}Var(e_1^2) + 2\frac{1}{n^2}E\sum_{t=2}^n\sum_{s=1}^{t-1}(e_t^2-\sigma_e^2)(e_s^2-\sigma_e^2) \\ =& o(1) + \frac{2}{n^2}E\sum_{t=2}^n\sum_{s=1}^{t-1}\left(\sum_{j=-\infty}^t \phi_{t-j}^2(\epsilon_j^2-1) + \sum_{j=-\infty}^t\sum_{j_1=-\infty,\neq j}^t \phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}\right)(e_s^2-\sigma_e^2) \\ =& \frac{2}{n^2}E\sum_{t=2}^n\sum_{s=1}^{t-1}\left(\sum_{j=-\infty}^s \phi_{t-j}^2(\epsilon_j^2-1) + \sum_{j=-\infty}^s\sum_{j_1=-\infty,\neq j}^s \phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}\right)(e_s^2-\sigma_e^2) \\ =& 2\frac{1}{n^2}E\sum_{t=2}^n\sum_{s=1}^{t-1}\sum_{j=-\infty}^s \phi_{t-j}^2(\epsilon_j^2-1)\sum_{\ell=-\infty}^s \phi_{s-\ell}^2(\epsilon_\ell^2-1) \\ &+ 2\frac{1}{n^2}E\sum_{t=2}^n\sum_{s=1}^{t-1}\sum_{j=-\infty}^s\sum_{j_1=-\infty,\neq j}^s \phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1}\sum_{\ell=-\infty}^s\sum_{\ell_1=-\infty,\neq \ell}^s \phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1} \\ &+ 2\frac{1}{n^2}E\sum_{t=2}^n\sum_{s=1}^{t-1}\sum_{j=-\infty}^s \phi_{t-j}^2(\epsilon_j^2-1)\sum_{\ell=-\infty}^s\sum_{\ell_1=-\infty,\neq \ell}^s \phi_{s-\ell}\phi_{s-\ell_1}\epsilon_\ell\epsilon_{\ell_1} \\ &+ 2\frac{1}{n^2}E\sum_{t=2}^n\sum_{s=1}^{t-1}\sum_{\ell=-\infty}^s \phi_{s-\ell}^2(\epsilon_\ell^2-1)\sum_{j=-\infty}^s\sum_{j_1=-\infty,\neq j}^s \phi_{t-j}\phi_{t-j_1}\epsilon_j\epsilon_{j_1} \\ =& 2\frac{1}{n^2}\sum_{t=2}^n\sum_{s=1}^{t-1}\sum_{j=-\infty}^s \phi_{t-j}^2\phi_{s-j}^2E(\epsilon_j^2-1)^2 + 4\frac{1}{n^2}\sum_{t=2}^n\sum_{s=1}^{t-1}\sum_{j=-\infty}^{s-1}\sum_{j_1=j+1}^s \phi_{t-j}\phi_{t-j_1}\phi_{s-j}\phi_{s-j_1}E\epsilon_j^2\epsilon_{j_1}^2 \\ \leq& O(1)\frac{1}{n^2}\sum_{t=2}^n\sum_{s=1}^{t-1}\sum_{j=-\infty}^s \phi_{t-j}^2 + 4\frac{1}{n^2}\sum_{t=2}^n\sum_{s=1}^{t-1}\sum_{j=-\infty}^{s-1}\sum_{j_1=j+1}^s |\phi_{t-j}\phi_{t-j_1}| \\ =& O(1)\frac{1}{n^2}\sum_{t=2}^n\sum_{s=1}^{t-1}(t-s)^{-2\gamma_0+1} + 4\frac{1}{n^2}\sum_{t=2}^n\sum_{s=1}^{t-1}(t-s)^{-2\gamma_0+2} = o(1)\end{aligned}$$

because $\sum_{t=2}^n\sum_{s=1}^{t-1}(t-s)^{-2\gamma_0+1} = O(1)\sum_{t=2}^n[1-t^{-2\gamma_0+2}] = O(n)$ and $\sum_{t=2}^n\sum_{s=1}^{t-1}(t-s)^{-2\gamma_0+2} = \sum_{t=2}^n[1-t^{-2\gamma_0+3}] = O(n)$ by virtue of $\gamma_0 > 3/2$.

Second, suppose $e_t = \wp(\epsilon_t,\cdots,\epsilon_{t-m_0+1};\eta_t,\cdots,\eta_{t-m_1+1})$ stipulated in b(ii) of Assumption A.

Let $\bar{m} = \max(m_0, m_1)$. Notice that whenever $|t-s| > \bar{m}$, e_t and e_s are independent. Whence,

$$E\left(\frac{1}{n}\sum_{t=1}^{n} e_t^2 - \sigma_e^2\right)^2 = \frac{1}{n^2}\sum_{t=1}^{n} E(e_t^2-\sigma_e^2)^2 + 2\frac{1}{n^2}E\sum_{t=2}^{n}\sum_{s=1}^{t-1}(e_t^2-\sigma_e^2)(e_s^2-\sigma_e^2)$$
$$=\frac{1}{n}Var(e_1^2) + 2\frac{1}{n^2}E\sum_{t=\bar{m}+1}^{n}\sum_{s=t-\bar{m}}^{t-1}(e_t^2-\sigma_e^2)(e_s^2-\sigma_e^2) \leq o(1) + 2\frac{1}{n^2}n\bar{m}Var(e_1^2) = o(1).$$

Next, we shall show $\frac{1}{n}\sum_{t=1}^{n}\widehat{g}^2(x_t) = o_P(1)$. For the sake of convenience, denote $\widehat{G}_n = \frac{1}{n}\sum_{t=1}^{n}\widehat{g}^2(x_t)$. Hence, for any $\delta, \varepsilon > 0$,

$$P(\widehat{G}_n > \delta) \leq P(\|\widehat{\theta}-\theta\| > \zeta_n\varepsilon) + P(\widehat{G}_n > \delta, \|\widehat{\theta}-\theta\| \leq \zeta_n\varepsilon)$$
$$\leq P(\|\widehat{\theta}-\theta\| > \zeta_n\varepsilon) + \frac{1}{\delta}E[\widehat{G}_n I(\|\widehat{\theta}-\theta\| \leq \zeta_n\varepsilon)],$$

by virtue of Markov's inequality, where $I(\cdot)$ stands for the conventional indicator function.

As $\|\widehat{\theta}-\theta_0\| = O_P(\zeta_n)$ assumed in Assumption B, $P(\|\widehat{\theta}-\theta_0\| > \zeta_n\epsilon) \to 0$. Using Taylor expansion for $g(x;\theta)$ with respect to θ in a neighborhood of θ_0, we have

$$\widehat{g}(x_t) = g(x_t;\theta_0) - g(x_t;\widehat{\theta}) = l_1(x_t;\theta_0)'(\theta_0-\widehat{\theta}) - \frac{1}{2}(\theta_0-\widehat{\theta})'l_2(x_t;\bar{\theta})(\theta_0-\widehat{\theta})$$

where $\bar{\theta}$ is on the line segment connecting θ_0 and $\widehat{\theta}$. It follows that

$$E[\widehat{G}_n I(\|\widehat{\theta}-\theta\| \leq \zeta_n\varepsilon)] = \frac{1}{n}\sum_{t=1}^{n}E[\widehat{g}^2(x_t)I(\|\widehat{\theta}-\theta\| \leq \zeta_n\varepsilon)]$$
$$=\frac{1}{n}\sum_{t=1}^{n}E[l_1(x_t;\theta_0)'(\theta_0-\widehat{\theta}) - \frac{1}{2}(\theta_0-\widehat{\theta})'l_2(x_t;\bar{\theta})(\theta_0-\widehat{\theta})]^2 I(\|\widehat{\theta}-\theta\| \leq \zeta_n\varepsilon)$$
$$\leq 2\frac{1}{n}\sum_{t=1}^{n}E[l_1(x_t;\theta_0)'(\theta_0-\widehat{\theta})]^2 I(\|\widehat{\theta}-\theta\| \leq \zeta_n\varepsilon)$$
$$+\frac{1}{2n}\sum_{t=1}^{n}E[(\theta_0-\widehat{\theta})'l_2(x_t;\bar{\theta})(\theta_0-\widehat{\theta})]^2 I(\|\widehat{\theta}-\theta\| \leq \zeta_n\varepsilon)$$
$$\leq 2\zeta_n^2\varepsilon^2\frac{1}{n}\sum_{t=1}^{n}E\|l_1(x_t;\theta_0)\|^2 + \zeta_n^4\varepsilon^4\frac{1}{2n}\sum_{t=1}^{n}E\|l_2(x_t;\bar{\theta})\|^2$$
$$\leq O(1)\zeta_n^2\varepsilon^2\frac{1}{n}\sum_{t=1}^{n}\frac{1}{\sqrt{t}}\int\|l_1(x;\theta_0)\|^2dx + O(1)\zeta_n^4\varepsilon^4\frac{1}{n}\sum_{t=1}^{n}\frac{1}{\sqrt{t}}\int\|l(x)\|^2dx$$
$$=O(1)\zeta_n^2\varepsilon^2\frac{1}{\sqrt{n}} + O(1)\zeta_n^4\varepsilon^4\frac{1}{\sqrt{n}} = o(1),$$

which, together with the convergence of $\frac{1}{n}\sum_{t=1}^{n}e_t^2$, implies $\frac{1}{n}\sum_{t=1}^{n}e_t\widehat{g}(x_t) = o_P(1)$. The assertion is proved. □

Proof of Theorem 3.3. This is obvious due to Lemma 3.1, Theorem 3.1 and Theorem 3.2. □

Proof of Theorem 3.4. The theorem can be demonstrated almost the same as Theorem 3.1 except that

$$\frac{d_n}{nk}\sum_{t=1}^{n}\|Z(x_t)\|^2\varphi^2(x_t) = \frac{d_n}{n}\sum_{t=1}^{n}\mathscr{T}_k(x_t)\varphi^2(x_t)$$

$$=\frac{d_n}{n}\sum_{t=1}^{n}\mathscr{T}(x_t)\varphi^2(x_t)+\frac{d_n}{n}\sum_{t=1}^{n}[\mathscr{T}_k(x_t)-\mathscr{T}(x_t)]\varphi^2(x_t)\to_D\int\mathscr{T}(x)\varphi^2(x)dxL_B(1,0)$$

by Lemma A.1 and Wang and Phillips (2009a). Thus, the limit is different from that of Theorem 3.1. □

Proof of Theorem 3.5. The theorem can be demonstrated identically as Theorem 3.2. □

Proof of Theorem 4.1. (1) Denote $\widehat{g}^*(x_t)=g(x_t;\widehat{\theta})-g(x_t;\widehat{\theta}^*)$ and hence $y_t^*-g(x_t;\widehat{\theta}^*)=\widehat{g}^*(x_t)+e_t^*$.

$$\begin{aligned}L_n^* =&\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)(y_t^*-g(x_t;\widehat{\theta}^*))(y_s^*-g(x_s;\widehat{\theta}^*))\\ =&\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)(\widehat{g}^*(x_t)+e_t^*)(\widehat{g}^*(x_s)+e_s^*)\\ =&\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)e_t^*e_s^*+2\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)\widehat{g}^*(x_t)e_s^*\\ &+\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)\widehat{g}^*(x_t)\widehat{g}^*(x_s)\\ :=&L_{1n}^*+2L_{2n}^*+L_{3n}^*,\qquad \text{say.}\end{aligned}$$

Notice that

$$\begin{aligned}L_{1n}^* =&\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)e_t^*e_s^*=\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)\widehat{e}_t\widehat{e}_s\tau_t\tau_s\\ =&\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)(e_t+\widehat{g}(x_t))(e_s+\widehat{g}(x_s))\tau_t\tau_s\\ =&\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)e_te_s\tau_t\tau_s+\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)\widehat{g}(x_t)\widehat{g}(x_s)\tau_t\tau_s\\ &+2\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)\widehat{g}(x_t)e_s\tau_t\tau_s\\ =&\sum_{t=1}^{n}Z(x_t)'Z(x_t)e_t^2+\sum_{t=1}^{n}Z(x_t)'Z(x_t)e_t^2(\tau_t^2-1)+\sum_{t=1}^{n}\sum_{s=1,\neq t}^{n}Z(x_t)'Z(x_s)e_te_s\tau_t\tau_s\\ &+\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)\widehat{g}(x_t)\widehat{g}(x_s)\tau_t\tau_s+2\sum_{t=1}^{n}\sum_{s=1}^{n}Z(x_t)'Z(x_s)\widehat{g}(x_t)e_s\tau_t\tau_s\end{aligned}$$

where, as defined before, $\widehat{g}(x_t)=g(x_t;\theta_0)-g(x_t;\widehat{\theta})$.

Let $F(x)$ be the distribution function of $\sigma_e^2L_B(1,0)$ and l_α be the $1-\alpha$ quantile of $F(x)$, that is, $F(l_\alpha)=1-\alpha$. It follows from Theorem 3.1 and the properties of the sequence τ_t that as $n\to\infty$

$$P^*\left(\frac{d_n}{nk}L_n^*<x\right)\to F(x) \tag{B.5}$$

holds for any $x\in\mathbb{R}$ in probability with respect to the distribution of the sample $\mathcal{W}_n$. Hence,

$$P^*\left(\frac{d_n}{nk}L_n^*>l_\alpha\right)\to 1-F(l_\alpha)=\alpha,\qquad \text{in probability} \tag{B.6}$$

which, together with $P^*\left(\frac{d_n}{nk}L_n^* > l_\alpha^*\right) = \alpha$ by definition, implies that $l_\alpha^* \to_P l_\alpha$.

Note that the result of Theorem 3.1 and (B.5) are tantamount to that as $n \to \infty$

$$P^*\left(\frac{d_n}{nk}L_n^* < x\right) - P\left(\frac{d_n}{nk}L_n < x\right) \to_P 0, \quad \forall\, x \in \mathbb{R}. \tag{B.7}$$

Recalling the definition of l_α^* again, (B.7) indicates $\lim_{n\to\infty} P\left(\frac{d_n}{nk}L_n > l_\alpha^*\right) = \alpha$, as required.

(2) This part is an implication of Theorem 3.2.

(3) The proof is similar to that of (1), so we omit for brevity.

(4) It is a consequence of Theorem 3.4. □

References

Ai, C. and Chen, X. (2003). Efficient estimation of models with conditional moment restrictions containing unknown functions. *Econometrica*, 71:1795–1843.

Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley & Sons, Inc., New York.

Edelman, A. and Rao, N. R. (2005). Random matrix theory. *Acta Numerica*, Cambridge University Press:1–65.

Gao, J. (2007). *Nonlinear Time Series: semi- and non-parametric methods.* Chapman & Hall, New York.

Gao, J., King, M. L., Lu, Z., and Tjøstheim, D. (2009a). Nonparametric specification testing for nonlinear time series with nonstationarity. *Econometric Theory*, 25:1869–1892.

Gao, J., King, M. L., Lu, Z., and Tjøstheim, D. (2009b). Specification testing in nonlinear and nonstationary time series autoregression. *Annals of Statistics*, 37(68):3893–3928.

Gao, J. and Phillips, P. C. B. (2013a). Functional coefficient nonstationary regression with non- and semi-parametric cointegration. *http://ideas.repec.org/p/msh/ebswps/2013-16.html*, 13:1–56.

Gao, J. and Phillips, P. C. B. (2013b). Semiparametric estimation in triangular system equations with nonstationarity. *Journal of Econometrics*, 176:59–79.

Gao, J., Tong, H., and Wolff, R. (2002). Model specification tests in nonparametric stochastic regression models. *Journal of Multivariate Analysis*, 83:324–359.

Gao, J., Wang, Q., and Yin, J. (2011). Specification testing in nonlinear time series with long-range dependence. *Econometric Theory*, 27:260–284.

Gelf́and, I. M. and Shilov, G. E. (1964). *Generalized Functions.* Academic Press, New York.

Hall, P., Horowitz, J., and Jing, B. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika*, 82:561–574.

Hong, S. H. and Phillips, P. C. B. (2010). Testing linearity of cointegrating relations with an application to purchasing power parity. *Journal of Business and Economic Statistics*, 28(1):96–114.

Hualde, J. and Robinson, P. M. (2011). Gaussian pseudo-maximum likelihood estimation of fractional time series models. *Annals of Statistics*, 39(6):3152–3181.

Karlsen, H. A., Mykelbust, T., and Tjøstheim, D. (2007). Nonparametric estimation in a nonlinear cointegration type model. *Annals of Statistics*, 35:252–299.

Karlsen, H. A. and Tjøstheim, D. (2001). Nonparametric estimation in null recurrent time series. *Annals of Statistics*, 29:372–416.

Newey, W. and Powell, J. (2003). Instrumental variable estimation of nonparametric models. *Econometrica*, 71:1565–1578.

Newey, W., Powell, J., and Vella, F. (1999). Nonparametric estimation of triangular simultaneous equations models. *Econometrica*, 67:565–603.

Park, J. Y. and Phillips, P. C. B. (1999). Asymptotics for nonlinear transformations of integrated time series. *Econometric Theory*, 15:269–298.

Park, J. Y. and Phillips, P. C. B. (2001). Nonlinear regression with integreted time series. *Econometrica*, 69(1):117–161.

Renyi, A. (1958). On mixing sequences of sets. *Acta Mathematica Hungarica*, 9:215–228.

Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion.* A series of comprehensive studies in mathematics 293. Springer-Verlag.

Szego, G. (1975). *Orthogonal Polynomials.* Colloquium publications XXIII. American Mathematical Association, Providence, Rhode Island.

Tao, T. (2012). *Topics in Random Matrix Theory.* Graduate studies in Mathematics, v-132. America Mathematic Soceity.

Wang, Q., Lin, Y. X., and Gulati, C. M. (2003). Asymptotics for general fractionally integrated processes with applications to unit root tests. *Econometric Theory*, 19:143–164.

Wang, Q. and Phillips, P. C. B. (2009a). Asymptotic theory for local time density estimation and nonparametric cointegreting regression. *Econometric Theory*, 25:710–738.

Wang, Q. and Phillips, P. C. B. (2009b). Structure nonparametric cointegrating regression. *Econometrica*, 77:1901–1948.

Wang, Q. and Phillips, P. C. B. (2012a). A specification test for nonlinear nonstationary models. *Annals of Statistics*, 40(2):727–758.

Wang, Q. and Phillips, P. C. B. (2012b). Supplement to "A specification test for nonlinear nonstationary models". *DOI:10.1214/12-AOS975SUPP.*

Wigner, E. P. (1958). On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67(2):325–327.