GRIPS Discussion Paper 13-11

A Note on the Mixingale Limit Theorem by McLeish (1977)

Shin S. Ikeda

July 2013

National Graduate Institute for Policy Studies
7-22-1 Roppongi, Minato-ku,
Tokyo, Japan 106-8677

A Note on the Mixingale Limit Theorem by McLeish (1977)

Shin S. Ikeda*

National Graduate Institute for Policy Studies

July 1, 2013

Abstract

The logical gap in the proof of non-stationary mixingale invariance principle by McLeish (1977) is identified and fixed by a skipped sub-sampling of a partial sum process in the continuous time. The corrected proof also delivers some extensions of the previous invariance principle and several stronger versions of convergence in law.

Key Words

Mixingale, Heteroskedasticity, Mixing and Stable Convergence in Law.

*e-mail: *s-ikeda@grips.ac.jp*

McLeish (1975, Theorem 1.6; 1977, Theorem 2.4) claims a proof of some invariance principle based on a mixingale-type dependent process. Compared with his celebrated maximal inequality, however, his asymptotic normality have drawn less attention. One reason, as Wooldridge and White (1988, p.214) suggest, is the non-primitive nature of his condition (1977, (2.6)) to replace the assumption of asymptotic independent increments of Billingsley (1968, (19.14))[1]. Mixed with a logical gap remained in his original proof, the drawback of McLeish (1975, 1977) has not been fixed for more than three decades since publication. The main purpose of this note is to identify and fill the gap. The corrected proof allows several stronger versions of convergence in law.

Definition 1 *(Information Filtrations, Mixingale, $\mathcal{C}$ and $\mathcal{D}$ spaces)*

(a) *$(\Omega, \mathcal{F}, \mathcal{P})$ is a probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,1]}$ of a non-decreasing sequence of sub-σ-fields $\mathcal{F}_t$ of $\mathcal{F}$. For $p \geq 1$, $\|\cdot\|_p := (E[|\cdot|^p])^{1/p}$ defines the L_p-norm.*

(b) *$(X_{n,j}) = (X_{n,j})_{n\in\mathbb{N},j\in\mathbb{J}}$ is an array of random variables where the set of discrete indices $\mathbb{J}$ may depend on n. $(\mathcal{I}_{n,j})$ is an array of filtration of sub-σ-fields $\mathcal{I}_{n,j}$ of $\mathcal{F}$ non-decreasing with respect to j for each n. $(X_{n,j})$ is L_p-bounded if $\sup_{n,j} \|X_{n,j}\|_p < \infty$.*

(c) *An array of L_1-bounded random variables paired with a filtration, $(X_{n,j}, \mathcal{I}_{n,j})$, is an L_p-mixingale if there exist a non-negative array of heterogeneous coefficients $(c_{n,j})$ and a non-negative sequence of mixingale numbers $(\psi_r)_{r\in\mathbb{N}}$ such that*

$$\|E[X_{n,j}|\mathcal{I}_{n,j-r}]\|_p \leq c_{n,j}\psi_r, \tag{1}$$

$$\|X_{n,j} - E[X_{n,j}|\mathcal{I}_{n,j+r}]\|_p \leq c_{n,j}\psi_{r+1}, \tag{2}$$

and $\psi_r \to 0$ as $r \to \infty$. For $\phi \geq 0$, it is of size $-\phi \leq 0$ if $\psi_r = O(r^{-\phi_0})$ for some $\phi_0 > \phi$, or alternatively if $\psi_r = O(r^{-\phi-\delta})$ for some $\delta > 0$.

(d) *$(X_{n,j})$ is uniformly integrable if $\lim_{c\to\infty} \sup_{n,j} E[|X_{n,j}|1_{\{|X_{n,j}|\geq c\}}] = 0$ where $1_{\{\cdot\}}$ is the indicator function, and is uniformly square integrable if $(X_{n,j}^2)$ is uniformly integrable.*

(e) *$\mathcal{C}$ is the set of continuous functions on $[0,1]$. $\mathcal{D}$ is the set of cádlág functions on $[0,1]$ so that $x \in \mathcal{D}$ is characterized by the existence of $x(t+)$ for any $t \in [0,1)$ and $x(t-)$ for any $t \in (0,1]$, $x(t) = x(t+)$ for any $t < 1$, and $x(1) = x(1-)$. $(\mathcal{D}, d_B)$ is the complete and separable metric space where d_B is the Billingsley's metric (Davidson 2002, Section 28.3). $(\mathcal{D}, \mathbb{B})$ is the measurable space where $\mathbb{B}$ is the Borel σ-field generated by d_B.*

(f) *$\mathcal{B}_n(t) = \sum_{j=1}^{\lfloor nt\rfloor} X_{n,j}$ is the partial-sum process of a random array $(X_{n,j})$ for $t \in [0,1]$ where $\lfloor\cdot\rfloor$ is the greatest integer part. $\mathcal{B}_n(0) = 0$. Note that $\mathcal{B}_n = \mathcal{B}_n(\cdot) \in \mathcal{D}$. $\mathcal{B}_n$ induces the probability law $\mu_n(\cdot) = \mathcal{P}(\mathcal{B}_n \in \cdot)$ on $(\mathcal{D}, \mathbb{B})$. $(\mathcal{B}_n)_{n\in\mathbb{N}}$ itself, or a sequence of probability laws it induces, $(\mu_n)_{n\in\mathbb{N}}$, is called uniformly tight if for any $\epsilon > 0$ there exists a compact set $K_\epsilon \in \mathbb{B}$ such that $\sup_{n\in\mathbb{N}}\{\mu_n(\mathcal{D} \setminus K_\epsilon)\} \leq \epsilon$.*

(g) *Let $\mathcal{I}$ be a sub-σ-field of $\mathcal{F}$ and $(Z_n)_{n\in\mathbb{N}}$ be a sequence of random vectors in the d-dimensional Euclidean Borel-measurable space $(\mathbb{R}^d, \mathbb{B}(\mathbb{R}^d))$. Z_n converges $\mathcal{I}$-stably in law to the canonical*

[1]It has been a standard argument to impose a near-epoch dependence on the class of mixingale processes so that the asymptotic independent increments property hold (e.g., Pötscher and Prucha 1991).

variable Z as $n \to \infty$, denoted by $Z_n \overset{d_s(\mathcal{I})}{\to} Z$, if, for any $\mathcal{I}$-measurable bounded variable η,

$$(Z_n, \eta) \overset{d}{\to} (Z^*, \eta) \ \ \text{and} \ \ Z^* \overset{d}{=} Z. \tag{3}$$

Z_n converges $\mathcal{I}$-mixing in law to Z as $n \to \infty$, denoted by $Z_n \overset{d_m(\mathcal{I})}{\to} Z$, if (3) holds with Z^* independent of $\mathcal{I}$. An equivalent condition for $Z_n \overset{d_m(\mathcal{I})}{\to} Z$ is

$$\forall u \in \mathbb{R}^d \ \ \forall F \in \mathcal{I} \ s.t. \ \mathcal{P}(F) > 0: \ \lim_{n\to\infty} E[e^{iu'Z_n}|F] = E[e^{iu'Z^*}] \ \ \text{and} \ \ Z^* \overset{d}{=} Z. \tag{4}$$

Remarks. Definition 1-(g) follows Aldous and Eagleson (1978, Prop. 1-(B); 2-(B''); 2-(C'')). Because the conditional measure on $F \in \mathcal{F}$ with $\mathcal{P}(F) > 0$ is characterized by $\mathcal{P}(\cdot|F) = \mathcal{P}(\cdot \cap F)/\mathcal{P}(F)$, the conditional law $\mu_n(\cdot|F) = \mathcal{P}(\mathcal{B}_n \in \cdot|F)$ satisfies

$$\forall A \in \mathbb{B}(\mathbb{R}^d): \ \mu_n(A|F) = \mathcal{P}(\{\mathcal{B}_n \in A\} \cap F)/\mathcal{P}(F) \le \mathcal{P}(\mathcal{B}_n \in A)/\mathcal{P}(F) = \mu_n(A)/\mathcal{P}(F), \tag{5}$$

so that $\mu_n(\cdot|F)$ satisfies any majorant inequalities for μ_n with majorant sides divided by $\mathcal{P}(F)$.

Theorem 1 *(Correcting and improving on McLeish 1975 Theorem 2.6; 1977 Theorem 2.4)*
Suppose the L_1-bounded random array paired with a filtration $(X_{n,j}, \mathcal{I}_{n,j})$ is such that

(a) it is an L_2-mixingale of size $-1/2$ with an array of heterogeneous coefficients $(c_{n,j})$,

(b) $(X_{n,j}/c_{n,j})$ is uniformly square integrable,

(c) $\sup_{s,t:0<s<t<1} \limsup_{n\to\infty} (s-t)^{-1} \sum_{j=\lfloor nt \rfloor}^{\lfloor ns \rfloor} c_{n,j}^2 < \infty$,

(d) $\mathcal{I}_{n,j}$ is the smallest σ-field containing $\mathcal{G} \cup \mathcal{H}_{n,j}$ where $(\mathcal{H}_{n,j})$ is a sub-filtration of $\mathbb{F}$ independent of a sub-σ-field $\mathcal{G}$ of $\mathcal{F}$; $(X_{n,j})_j$ is $(\mathcal{I}_{n,j})_j$-adapted with zero mean, and

(e) $h^{-1}E[(\sum_{j=\lfloor nt \rfloor}^{\lfloor n(t+h) \rfloor} X_{n,j})^2 | \mathcal{I}_{n,\lfloor ns \rfloor}] - \theta_t \overset{p}{\to} 0$ as $n^{-1} + h + (nh)^{-1} \to 0$ for any s, t such that $0 \le s < t < t + h < 1$ where $(\theta_t)_{t\in[0,1]}$ is non-negative, t-continuous, uniformly integrable and $\mathcal{G}$-measurable, and $\int_0^1 \theta_t dt$ is uniformly bounded away from zero.

Then, for each $t \in (0,1]$, there exists a limit variable $Z \sim N(0,1)$ independent of $\mathcal{G}$ such that

$$\mathcal{B}_n(t) := \left(\int_0^t \theta_v dv \right)^{-1/2} \sum_{j=1}^{\lfloor nt \rfloor} X_{n,j} \overset{d_m(\mathcal{G})}{\to} Z. \tag{6}$$

(Proof) Let us introduce another set of notations for the purpose of proof.

Definition 2 *(A skipped subsample and the $\mathcal{G}$-conditional characteristic function)*

(a) For any integer $k \ge 4$ and any partition $(t_a)_{a=1\ldots k}$ of $[0,1)$ such that $0 \le t_1 < \cdots < t_k < 1$,

$$(\bar{X}_{n,l}, \bar{\mathcal{I}}_{n,l}) := \begin{cases} (X_{n,l}, \ \mathcal{I}_{n,l}) & l \le \lfloor nt_{k-2} \rfloor \\ (X_{n,\lfloor nt_{k-1} \rfloor + l - \lfloor nt_{k-2} \rfloor - 1}, \ \mathcal{I}_{n,\lfloor nt_{k-1} \rfloor + l - \lfloor nt_{k-2} \rfloor - 1}) & l > \lfloor nt_{k-2} \rfloor \end{cases}$$

for $l \in \mathbb{N}$ and $h_j := t_j - t_{j-1}$ with $\sup_j h_j \to 0$ as $n \to \infty$.

(b) $\bar{\mathcal{B}}_n(t) := (\int_0^t \theta_v dv)^{-1/2} \sum_{l=1}^{\lfloor nt \rfloor} \bar{X}_{n,l}$, $\bar{Z}_q := \sum_{a=1}^q u_a \bar{\mathcal{B}}_n(t_a)$ for $u_a \in \mathbb{R}$, $E^*[\cdot] = E[\cdot|\mathcal{G}]$, and

$$\Phi_n^*(t,u) := E^*[e^{i\bar{Z}_{k-3}+iu\bar{\mathcal{B}}_n(t)}], \quad t > t_{k-3}, \ u \in \mathbb{R}, \ i = \sqrt{-1}.$$

(c) $\Delta_{k-1} := \bar{\mathcal{B}}_n(t_k) - \bar{\mathcal{B}}_n(t_{k-2})$. Note that $\Delta_{k-1} = (\int_0^t \theta_v dv)^{-1/2} \sum_{j=\lfloor nt_{k-1} \rfloor}^{\lfloor nt_k \rfloor} X_{n,j}$.

$(\bar{X}_{n,l}, \bar{\mathcal{I}}_{n,l})_l$ in Definition 2-(a) is a subsample of $(X_{n,j}, \mathcal{I}_{n,j})_j$ avoiding $\lfloor nt_{k-2} \rfloor < j < \lfloor nt_{k-1} \rfloor$. $\bar{Z}_q$ in Definition 2-(b) satisfies $\bar{Z}_{k-3} + u\bar{\mathcal{B}}_n(t_{k-2}) = \bar{Z}_{k-2}$ for $u = u_{k-2}$, which is assumed subsequently without loss of generality. Because the partition (t_a) is arbitrary, let us assume

$$h_k + h_{k-1} + h_{k-1}h_k^{-1} + (nh_k)^{-1} + (nh_{k-1})^{-1} \to 0 \quad as \quad n \to \infty \tag{7}$$

i.e., $h_{k-1} \to 0$ faster than $h_k \to 0$, but both are slower than $1/n \to 0$ as $n \to \infty$. Note that

$$\frac{\Phi_n^*(t_k,u) - \Phi_n^*(t_{k-2},u)}{t_k - t_{k-2}} = \frac{E^*\left[e^{i\bar{Z}_{k-2}}(e^{iu\Delta_{k-1}} - 1)\right]}{t_k - t_{k-2}}. \tag{8}$$

Davidson (2002, Lemma 11.26) and the second-order Taylor expansion with a remainder imply

$$\begin{aligned} e^{iu\Delta_{k-1}} - 1 &= iu\Delta_{k-1} - u^2\Delta_{k-1}^2/2 + c(u\Delta_{k-1}), & (9)\\ |c(u\Delta_{k-1})| &\le \min\{|u\Delta_{k-1}|^2, |u\Delta_{k-1}|^3/6\}. & (10)\end{aligned}$$

Using (7) to (10), mixingale property, t-continuity of θ_t and calculation in the appendix,

$$\forall v \in (t_{k-3}, t_k) : \frac{\Phi_n^*(t_k,u) - \Phi_n^*(v,u)}{t_k - v} = -\frac{u^2}{2} \cdot \frac{\theta_v}{\int_0^t \theta_v dv} \cdot \Phi_n^*(v,u) + \mathcal{R}_n \tag{11}$$

where $\mathcal{R}_n$ is a $\mathcal{G}$-measurable remainder asymptotically vanishing almost surely.

Conditions (a) to (c) for $(X_{n,j}, \mathcal{I}_{n,j})$ can be applied to $(\bar{X}_{n,l}, \bar{\mathcal{I}}_{n,l})$. Given (18) and the subsequent discussion in the appendix, (a) and (b) imply that $(\bar{\mathcal{B}}_n)$ is uniformly square integrable so that (11) with the second moment of Δ_{k-1} holds in the limit without the remainder. (a) to (c) combined with Lemma 4 in the appendix ensure that $(\bar{\mathcal{B}}_n)$ is also uniformly tight converging in law to a $\mathcal{P}$-a.s. continuous random process $\bar{\mathcal{B}}$. Moreover, the uniform tightness applies to conditional laws $(\mu_n(\cdot|G))_{n\in\mathbb{N}}$ for any $G \in \mathcal{G}$ with $\mathcal{P}(G) > 0$. Therefore, we can define $\Phi^*(t,u)$ for $\bar{\mathcal{B}}$ like $\Phi_n^*(t,u)$ for $\bar{\mathcal{B}}_n$ such that, $\mathcal{P}$-almost surely, $\Phi^*(t,u)$ is t-continuous with the right t-derivative coincident with the two-sided t-derivative (Billingsley 1968, p155). Divide (11) by $\Phi_n^*(v,u)$ and let $n \to \infty$ to obtain

$$\forall v \in (t_{k-3}, 1) : \frac{\partial \ln \Phi^*(v,u)}{\partial v} \overset{a.s.}{=} -\frac{u^2}{2} \cdot \frac{\theta_v}{\int_0^t \theta_v dv}. \tag{12}$$

By integrating both sides of (12) with respect to $v \in (s,t) \subset (t_{k-3}, 1)$ and exponentiating,

$$\Phi^*(t,u)/\Phi^*(s,u) \overset{a.s.}{=} \exp\left(-(u^2/2)\int_s^t \theta_v dv \Big/ \int_0^t \theta_v dv\right). \tag{13}$$

Set $k = 4$ and $t_{k-3} = t_1 = 0$. The continuity of $\bar{\mathcal{B}}$ and $\mathcal{P}(\bar{\mathcal{B}}(0) = 0) = 1$ imply $\Phi^*(s,u) \overset{a.s.}{\to} 1$ and therefore $\Phi_n^*(t,u) \overset{a.s.}{\to} e^{-u^2/2}$ as $s \to 0$. The calculation in the appendix reveals

$$\|\mathcal{B}_n(t) - \bar{\mathcal{B}}_n(t)\|_2 \to 0, \quad implying \ \ \mathcal{B}_n(t) - \bar{\mathcal{B}}_n(t) \overset{p}{\to} 0 \tag{14}$$

as $n\to\infty$ given (7). Because $e^{iu\{\mathcal{B}_n(t)-\bar{\mathcal{B}}_n(t)\}}-1\overset{p}{\to}0$ and $(e^{iu\{\mathcal{B}_n(t)-\bar{\mathcal{B}}_n(t)\}}-1)$ is uniformly integrable, $|E[e^{iu\mathcal{B}_n(t)}-e^{iu\bar{\mathcal{B}}_n(t)}|G]|\leq\|e^{iu\{\mathcal{B}_n(t)-\bar{\mathcal{B}}_n(t)\}}-1\|_1/\mathcal{P}(G)\to 0$ for any $G\in\mathcal{G}$ with $\mathcal{P}(G)>0$ so that $E^*[e^{iu\mathcal{B}_n(t)}-e^{iu\bar{\mathcal{B}}_n(t)}]\overset{a.s.}{\to}0$ given Lemma 1-(a) and (24) in the appendix. Consequently,

$$E^*[e^{iu\mathcal{B}_n(t)}]=E^*[e^{iu\bar{\mathcal{B}}_n(t)}]+E^*[e^{iu\mathcal{B}_n(t)}-e^{iu\bar{\mathcal{B}}_n(t)}]\overset{a.s.}{\to}e^{-u^2/2},$$

which does not depend on $\mathcal{G}$. Therefore, $\mathcal{B}_n(t)\overset{d_s(\mathcal{G})}{\to}N(0,1)$ for any $t\in(0,1)$. The continuity of $\mathcal{B}(t)$ allows the result for $t\to 1$.

□

Remarks. (e) is weaker than McLeish (1977, (2.6)) because we employ the probability limit rather than L_1 limit, and the conditional variance may not be time-homogeneous nor deterministic. The subsample $(\bar{X}_{n,l})$ is designed for ensuring $\Delta_{k-1}=\bar{\mathcal{B}}_n(t_k)-\bar{\mathcal{B}}_n(t_{k-2})\propto\sum_{j=\lfloor nt_{k-1}\rfloor}^{\lfloor nt_k\rfloor}X_{n,j}$ while $\mathcal{B}_n(t_k)-\mathcal{B}_n(t_{k-m})\propto\sum_{j=\lfloor nt_{k-m}\rfloor+1}^{\lfloor nt_k\rfloor}X_{n,j}$ for any $m\in\mathbb{N}$. $X_{n,\lfloor nt_{k-2}\rfloor}$ in $\bar{Z}_{k-2}$ and $X_{n,\lfloor nt_{k-1}\rfloor}$ in Δ_{k-1} are asymptotically separated given conditioning on $\mathcal{I}_{n,\lfloor nt_{k-2}\rfloor}$ within E^*, (7) and mixingale properties because their displacement in continuous time is $\lfloor nt_{k-1}\rfloor-\lfloor nt_{k-2}\rfloor\geq nh_{k-1}\to\infty$ as $n\to\infty$, whereas it does not disturb the asymptotic distribution as $h_{k-1}\to 0$ so that the skipped part shrinks quickly. Given $Z_q:=\sum_{a=1}^{q}u_a\mathcal{B}_n(t_a)$ based on $(X_{n,j})$, Billingsley (1968, p.160) assumes $h^{-1}E\left[e^{iZ_{k-1}}\{e^{iu(\mathcal{B}_n(t_k)-\mathcal{B}_n(t_{k-1}))}-1\}\right]\to 0$ or its variants as $h\downarrow 0$. Because $X_{n,\lfloor nt_{k-1}\rfloor+1}$ in Z_{k-1} and $X_{n,\lfloor t_{k-1}\rfloor}$ in $\mathcal{B}_n(t_k)-\mathcal{B}_n(t_{k-1})$ are always two consecutive sampling points, they cannot be separated asymptotically as $n\to\infty$. Although it is sufficient for applications to martingale-difference or interchangeable variables with weaker dependence (Billingsley 1968, Section 23, 24; Chow and Teicher 1997, Theorem 7.3.2), it is inappropriate for our case with non-trivial dependence. McLeish (1975, p.175; 1977, p.620) assumes $h^{-1}E\left[e^{iZ_{k-2}}\{e^{iu(\mathcal{B}_n(t_k)-\mathcal{B}_n(t_{k-1}))}-1\}\right]\to 0$ as $h\downarrow 0$ so that $X_{n,\lfloor nt_{k-2}\rfloor+1}$ in Z_{k-2} and $X_{n,\lfloor t_{k-1}\rfloor}$ in $\mathcal{B}_n(t_k)-\mathcal{B}_n(t_{k-1})$ are separated asymptotically. However, (8) suggests that it is not linked to Billingsley's original strategy of approximating the characteristic function by that of a Gaussian process. Theorem 1 fills this gap in a series of approximations from (8) to (11) for $(\bar{X}_{n,l})$ together with (14). It also shows that the convergence occurs in the $\mathcal{G}$-mixing sense.

Obviously, (b) and (c) are respectively implied by

(b') $\sup_{n,j}\|X_{n,j}/c_{n,j}\|_r<\infty$ for some $r>2$, and

(c') $\sup_j c_{n,j}=O(n^{-1/2})$.

The next result offers conditions alternative to (e) in Theorem 1. The size of mixing numbers below is defined similarly as that of mixingale numbers in Definition 1-(c).

Theorem 2 *For $(X_{n,j},\mathcal{I}_{n,j})$ in Theorem 1, define $\delta_{t,u}=\sum_{j=\lfloor nt\rfloor}^{\lfloor nu\rfloor}X_{n,j}$. Suppose either*

(i) $X_{n,j}=y_{n,j}U_{n,j}$ where $(y_{n,j})$ is $\mathcal{G}$-measurable and $\sup_j\|y_{n,j}\|_2=O(n^{-1/2})$, and $(U_{n,j})$ is $(\mathcal{H}_{n,j})$-adapted, L_q-bounded and α- or ϕ-mixing of size $-q/(q-2)$ for some $q>2$, or

(ii) $(X_{n,j})$ is $(\mathcal{H}_{n,j})$-adapted and α- or ϕ-mixing of any negative size. Then,

$$\forall t\in(0,1)\ \forall h\in(0,1-t):\ h^{-1}\|E[\delta^2_{t,t+h}|\mathcal{I}_{n,\lfloor nt\rfloor-r}]-E^*[\delta^2_{t,t+h}]\|_1\to 0\quad as\quad r\to\infty. \qquad (15)$$

(Proof) (i): Suppose $U_{n,j}$ is α-mixing with the α-mixing number α_m of the assumed size so that $\alpha_m=O(m^{-q/(q-2)-\epsilon})$ for some $\epsilon>0$. The proof for ϕ-mixing case is almost identical. For

$\kappa := \epsilon(q-2)/q > 0$, $m^{\kappa/2}\alpha_m^{(q-2)/q} = O(m^{-1-\kappa/2})$. Define $j \wedge l = \min\{j,l\}$. As $r \to \infty$,

$$\sum_{j,l=\lfloor nt\rfloor}^{\lfloor n(t+h)\rfloor} \alpha_{j\wedge l-\lfloor nt\rfloor+r}^{(q-2)/q} \leq (2nh+1)\sum_{m=r}^{\infty}\alpha_m^{(q-2)/q} \leq (2nh+1)r^{-\kappa/2}\sum_{m=r}^{\infty} m^{\kappa/2}\alpha_m^{(q-2)/q} = o(nhr^{-\kappa/2}).$$

Note that $\delta_{t,t+h}^2 = \sum_{j,l=\lfloor nt\rfloor}^{\lfloor n(t+h)\rfloor} y_{n,j}y_{n,l}U_{n,j}U_{n,l}$. For $j,l \geq \lfloor nt\rfloor$, $U_{n,j}U_{n,l}$ is measurable with respect to $\mathcal{H}_{n,\max\{j,l\}}$, which is later than $\mathcal{H}_{n,j\wedge l}$. Therefore, it is separated from $\mathcal{H}_{n,\lfloor nt\rfloor-r}$ by at least $j\wedge l-\lfloor nt\rfloor+r$ in the discrete index and the associated α-mixing number is at most $\alpha_{j\wedge l-\lfloor nt\rfloor+r}$. Using the Minkowski, Hölder, norm, α-mixing and Cauchy-Schwarz inequalities in conjunction with the above estimate and the given condition,

$$\begin{aligned}
& h^{-1}\|E[\delta_{t,t+h}^2|\mathcal{I}_{n,\lfloor nt\rfloor-r}] - E^*[\delta_{t,t+h}^2]\|_1 \\
\leq\; & 6h^{-1}\sum_{j,l=\lfloor nt\rfloor}^{\lfloor n(t+h)\rfloor}\|y_{n,j}\|_2\|y_{n,l}\|_2\|U_{n,j}\|_q\|U_{n,l}\|_q\alpha_{j\wedge l-\lfloor nt\rfloor+r}^{1-2/q} \\
\leq\; & 6h^{-1}(\sup_j\|y_{n,j}\|_2)^2(\sup_j\|U_{n,j}\|_q)^2\sum_{j,l=\lfloor nt\rfloor}^{\lfloor n(t+h)\rfloor}\alpha_{j\wedge l-\lfloor nt\rfloor+r}^{(q-2)/q} \\
=\; & o(r^{-\kappa/2}) \to 0 \quad as \quad r\to\infty.
\end{aligned}$$

(ii): Notice that $X_{n,j}$ is independent from $\mathcal{G}$. Then, we can replace $E[\cdot|\mathcal{I}_{n,j}]$ and $E[\cdot|\mathcal{G}]$ by $E[\cdot|\mathcal{H}_{n,j}]$ and $E[\cdot]$. The rest is identical to McLeish (1975, p.177).

□

Remarks. By combining McLeish (1975, Theorem 3.8; 1977, Corollary 2.11) with independence between $\mathcal{G}$ and $(\mathcal{H}_{n,j})$, we can show that the size condition for α-mixing numbers of $(U_{n,j})$ to ensure the condition (a) in Theorem 1 is equivalent to the one for the case-(i) in Theorem 2. Improving the size condition is left for a future research. (15) and Lemma 1-(c) in the appendix guarantee that it suffices to find the probability limit of $h^{-1}E[\delta_{t,t+h}^2|\mathcal{G}]$ for θ_t under Theorem 2.

Theorem 3 (FCLT) *Suppose $(X_{n,j})$ in Theorem 1 is independent of $\mathcal{G}$ and $(\theta_t)_{t\in[0,1]}$ is deterministic. For the standard Brownian motion $(W_v)_{v\in[0,1]}$ independent of $\mathcal{H}_1$,*

$$\mathcal{B}_n(\cdot) := \sum_{j=1}^{\lfloor n\cdot\rfloor} X_{n,j} \overset{d_m(\mathcal{H}_1)}{\Rightarrow} \int_0^{\cdot}\theta_v^{1/2}dW_v \quad as \quad n\to\infty.$$

(Proof) The scaling $(\int_0^t\theta_v dv)^{-1/2}$ is not necessary because (θ_t) is not random. It suffices to show that any finite-dimensional distributions conditional on each $F \in \mathcal{H}_1$ with $\mathcal{P}(F) > 0$ converge to those of an unconditionally Gaussian process because of the uniform tightness. We can identify $\mathcal{I}_{n,j}$ with $\mathcal{H}_{n,j}$ and $E^*[\cdot]$ with $E[\cdot|F]$ without loss of generality. We can apply the proof of Theorem 1 to that of $E^*[e^{iu\mathcal{B}_n(t)}] \to e^{-(u^2/2)\int_0^t\theta_v dv}$ which is independent of $\mathcal{H}_1$ for each t. (13) reads

$$\Phi^*(t,u) = \Phi^*(s,u)e^{-(u^2/2)\int_s^t\theta_v dv},$$

with $\Phi^*(s,u) = e^{-(u^2/2)\int_0^s\theta_v dv}$ by setting $k=5$, $t_1=0$, $t_2=s<t_3<t_4<t_5=t$. Repeat this argument as in Billingsley (1968, p.163) to ensure the independent incremental property and convergence of any finite-dimensional distributions.

□

Remarks. This result improves McLeish (1977, Theorem 2.4), Wooldridge and White (1988, Theorem 2.1) and De Jong (1997, Theorem 1) by replacing their time-homogeneous instantaneous variance $t \in [0,1]$ by a time-heterogeneous $(\theta_t)_{t\in[0,1]}$. The ease of proving this functional result comes at the cost of a stronger non-randomness of $(\theta_t)_{t\in[0,1]}$ than (e) in Theorem 1. We conjecture the validity of a functional $\mathcal{H}_1$-stable convergence in law if (θ_t) is $(\mathcal{H}_t)$-adapted, but its investigation is left for a future research.

Theorem 4 *If $Y_n \overset{d_s(\mathcal{I})}{\to} Y$ and $Z_n \overset{d_m(\mathcal{I})}{\to} Z$, then $(Y_n, Z_n) \overset{d_s(\mathcal{I})}{\to} (Y, Z)$. In particular, any $Y_n \overset{d_s(\mathcal{G})}{\to} Y$ and (6) under Theorem 1 occur jointly as the $\mathcal{G}$-stable convergence in law:*

$$(Y_n, \mathcal{B}_n(t)) \overset{d_s(\mathcal{G})}{\to} (Y, Z), \quad Z \sim N(0,1) \ \ \textit{independent of}\ \mathcal{G}.$$

(Proof) The proof mainly follows Barndorff-Nielsen et. al. (2008, Proposition 5). Using Definition 1-(g) and the equivalence of convergence in distribution and in characteristic function, it is sufficient to show $V := E[e^{ia'Y_n+ib'Z_n+ic'\eta}] - E[e^{ia'Y+ib'Z+ic'\eta}] \to 0$ for any real vectors a, b, c. Using the add-subtract trick, self-adjointness of conditional expectations (Kallenberg 2002, p.104) and independence of Z from $\mathcal{I}$,

$$V = E[e^{ia'Y_n}(E[e^{ib'Z_n}|\mathcal{I}] - E[e^{ib'Z}])e^{ic'\eta}] + E[(e^{ia'Y_n} - e^{ia'Y})E[e^{ib'Z}]e^{ic'\eta}]. \tag{16}$$

$E[e^{ib'Z}]$ as a constant is taken out of the second expectation. $Y_n \overset{d_s(\mathcal{I})}{\to} Y$ means $(Y_n, \eta) \overset{d}{\to} (Y, \eta)$ for any $\mathcal{I}$-measurable variable η or equivalently $E[e^{ia'Y_n+ic'\eta}] - E[e^{ia'Y+ic'\eta}] \to 0$ so that the second term in the right hand side of (16) vanishes as $n \to \infty$. The first term in the right hand side of (16) is bounded absolutely by $E[|E[e^{ib'Z_n}|\mathcal{G}] - E[e^{ib'Z}]|]$, which tends to zero if $E[e^{ib'Z_n}|\mathcal{I}] \overset{L_1}{\to} E[e^{ib'Z}]$. Because $E[e^{ib'Z_n}|F] \to E[e^{ib'Z}]$ for any $F \in \mathcal{I}$ with $\mathcal{P}(F) > 0$, (24) in the appendix ensures $E[e^{ib'Z_n}|\mathcal{I}] \overset{a.s.}{\to} E[e^{ib'Z}]$, implying $E[e^{ib'Z_n}|\mathcal{I}] \overset{L_1}{\to} E[e^{ib'Z}]$.

□

These results are particularly useful in the context of estimating a quadratic variation of a financial security price process given high-frequency intra-daily data with a stochastic-volatility leverage effect and a microstructure effect. For instance, Theorem 1 can generalize Barndorff-Nielsen et.al. (2008, Theorem 1 and Lemma 4) for the asymptotic normality of a cross term in their realized-kernel estimator based on a serially independent microstructure effect to those based on serially dependent one. The mixingale framework also naturally incorporates a case of diurnally heteroskedastic microstructure effect as is documented by Kalnina and Linton (2008). See Ikeda (2013a, b) for examples.

Appendix I. Results and proofs for the main theorems

Lemma 1 *(Uniformly integrable arrays)*

(a) If $X_n \xrightarrow{p} X$, $X_n \xrightarrow{L_1} X$ *if and only if* (X_n) *is uniformly integrable.*

(b) $(X_{n,j})$ *is uniformly integrable if and only if*

(i) $\sup_{n,j} E[|X_{n,j}|] < \infty$ *and (ii)* $\lim_{\mathcal{P}(A)\to 0} \sup_{n,j} E[|X_{n,j}|1_A] = 0$.

(c) If $(X_{n,j})$ *is uniformly integrable, so is* $(E[X_{n,j}|\mathcal{J}_{n,j}])$ *for any array of* σ*-fields* $(\mathcal{J}_{n,j})$.

(d) If $(X_{n,j})$ *and* $(Y_{n,j})$ *are uniformly integrable, so is* $(Y_{n,j} - X_{n,j})$.

(Proof) (a), (b): Kallenberg (2002, Proposition 4.12 and Lemma 4.10). (c): $Y_{n,j} := E[X_{n,j}|\mathcal{J}_{n,j}]$ *satisfies* $\sup_{n,j} E[|Y_{n,j}|] \leq \sup_{n,j} E[|X_{n,j}|] < \infty$ *because of (b)-(i), the triangular inequality and the law of iterated expectations. For* $\zeta > 0$ *and* $A := \{|Y_{n,j}| \geq \zeta\}$, $\mathcal{P}(A) \leq \sup_{n,j} E[|Y_{n,j}|]/\zeta \to 0$ *as* $\zeta \to \infty$ *by the Markov's inequality. The claim holds as* $E[|Y_{n,j}|1_{\{|Y_{n,j}|\geq\zeta\}}] \leq E[|X_{n,j}|1_A] \to 0$ *by the uniform integrability of* $(X_{n,j})$, $\mathcal{P}(A) \to 0$ *and Lemma 1-(b)-(ii). (d): Gut (2005, Theorem 4.6).*

Lemma 2 (Mixing inequalities; Davidson 2002, 14.2, 14.4) *For* $s \geq p > 1$, *suppose a random array* $(X_{n,j})$ *is adapted to a filtration* $(\mathcal{I}_{n,j})$, L_s*-bounded with* α*-mixing numbers and* ϕ*-mixing numbers,* (α_r) *and* (ϕ_r), *respectively. Then,*

$$\|E[X_{n,j}|\mathcal{I}_{n,j-r}] - E[X_{n,j}]\|_p \quad \leq \quad \min\{2(2^{1/p}+1)\alpha_r^{1/p-1/s}, 2\phi_r^{1-1/s}\}\|X_{n,j}\|_s. \tag{17}$$

$p = 1$ and $s = q/2$ is used in the proof of Theorem 2-(i) with multipliers of mixing numbers dominated by 6. Besides, $\|U_{n,j}U_{n,l}\|_{q/2} \leq \|U_{n,j}\|_q \|U_{n,l}\|_q$ by the Cauchy-Schwarz inequality.

Lemma 3 *A sequence of probability laws* $(\mu_n)_{n\in\mathbb{N}}$ *in* $(\mathcal{D}, \mathbb{B})$ *is uniformly tight if*

(a) $\exists N \in \mathbb{N}\ \forall n \geq N\ \forall \eta > 0\ \exists M \in (0,\infty)$*:* $\mu_n(\{x : |x(0)| > M\}) \leq \eta$.

(b) $\exists N \in \mathbb{N}\ \forall n \geq N\ \forall \epsilon > 0\ \forall \eta > 0\ \exists \delta \in (0,1)$*:* $\mu_n(\{x : w(x,\delta) \geq \epsilon\}) \leq \eta$.

Moreover, any cluster point of (μ_n), *say* μ, *satisfies* $\mu(\mathcal{C}) = 1$.

(Proof) See Appendix II.

Lemma 4 *For the* L_1*-bounded random array paired with a filtration* $(X_{n,j}, \mathcal{I}_{n,j})$ *such that*

(a) it is an L_2*-mixingale of size* $-1/2$ *with an array of heterogeneous coefficients* $(c_{n,j})$,

(b) $(X_{n,j}/c_{n,j})$ *is uniformly square integrable, and*

(c) $\sup_{s,t:0\leq t<s<1} \limsup_{n\to\infty} (s-t)^{-1} \sum_{j=\lfloor nt\rfloor}^{\lfloor ns\rfloor} c_{n,j}^2 < \infty$.

Let μ_n *be the probability law on* $(\mathcal{D}, \mathbb{B})$ *induced by* $\mathcal{X}_n(t) := \sum_{j=1}^{\lfloor nt\rfloor} X_{n,j}$. *Then,* $(\mu_n)_{n\in\mathbb{N}}$ *satisfies Lemma 3-(a) and (b).*

(Proof) Lemma 3-(a) is immediate by $\mathcal{X}_n(0) = 0$. For any $\delta \in (0,1]$ and $t \in [0, 1-\delta]$,

$$\sup_{s\in[t,t+\delta]} \frac{[\mathcal{X}_n(s) - \mathcal{X}_n(t)]^2}{\delta} = \sup_{s\in[t,t+\delta]} \frac{[\mathcal{X}_n(s) - \mathcal{X}_n(t)]^2}{(\sum_{j=\lfloor nt\rfloor}^{\lfloor ns\rfloor} c_{n,j}^2)^2} \left(\frac{\sum_{j=\lfloor nt\rfloor}^{\lfloor ns\rfloor} c_{n,j}^2}{s-t} \right)^2 \frac{(s-t)^2}{\delta}. \tag{18}$$

(a) and (b) imply that $\{\sup_{s\in[t,t+\delta]}[\mathcal{X}_n(s) - \mathcal{X}_n(t)]^2/(\sum_{j=\lfloor nt\rfloor}^{\lfloor ns\rfloor} c_{n,j}^2)^2\}_{n\in\mathbb{N}}$ is uniformly integrable (Davidson 2002, Theorem 16.13; McLeish 1977, Proof of Theorem 2.4), and so is the supremum of the left hand side of (18) with respect to $\delta \in (0,1]$, $t \in [0, 1-\delta]$ and $n \in \mathbb{N}$ by (c) and $(s-t)^2 \leq \delta$. The uniform integrability of $\{\sup_{s\in[t,t+\delta]} \delta^{-1}[\mathcal{X}_n(s) - \mathcal{X}_n(t)]^2\}_{\delta\in(0,1],t\in[0,1-\delta],n\in\mathbb{N}}$ guarantees that one can pick some δ sufficiently small and some $N \in \mathbb{N}$ such that for any $n \geq N$, $t \in [0, 1-\delta]$, and $\eta', \epsilon > 0$,

$$\begin{aligned}
& \delta^{-1}\mu_n(\{x : \sup_{s\in[t,t+\delta]} |x(s) - x(t)| \geq \epsilon/2\}) \\
= \; & \delta^{-1}\mathcal{P}(\sup_{s\in[t,t+\delta]} |\mathcal{X}_n(s) - \mathcal{X}_n(t)| \geq \epsilon/2) \\
= \; & \frac{1}{(\epsilon/2)^2} E\left[\delta^{-1}(\epsilon/2)^2 1_{\{\sup_{s\in[t,t+\delta]} \delta^{-1}|\mathcal{X}_n(s)-\mathcal{X}_n(t)|^2 \geq \delta^{-1}(\epsilon/2)^2\}}\right] \\
\leq \; & \frac{1}{(\epsilon/2)^2} E\left[\sup_{s\in[t,t+\delta]} \delta^{-1}|\mathcal{X}_n(s) - \mathcal{X}_n(t)|^2 1_{\{\sup_{s\in[t,t+\delta]} \delta^{-1}|\mathcal{X}_n(s)-\mathcal{X}_n(t)|^2 \geq \delta^{-1}(\epsilon/2)^2\}}\right] \\
\leq \; & \frac{\eta'}{(\epsilon/2)^2}
\end{aligned}$$

because $(\epsilon/2)^2/\delta$ can be sufficiently large. Since $\eta'/(\epsilon/2)^2$ can be made arbitrarily smaller than $\eta/2$ for any $\eta > 0$. Therefore, $\exists\delta \in (0,1]\ \exists N \in \mathbb{N}\ \forall n \geq N\ \forall t \in [0, 1-\delta]\ \forall\epsilon > 0\ \forall\eta > 0$:

$$\delta^{-1}\mu_n(\{x : \sup_{s\in[t,t+\delta]} |x(s) - x(t)| \geq \epsilon/2\}) \leq \eta/2. \tag{19}$$

By multiplying both sides by δ and taking the supremum over $t \in [0, 1-\delta]$, we have

$$\begin{gathered}
\exists\delta \in (0,1], \exists N \in \mathbb{N}, \forall n \geq N, \forall\epsilon > 0, \forall\eta > 0 : \\
\sup_{t\in[0,1-\delta]} \mu_n \left(\left\{ x : \sup_{s\in[t,t+\delta]} |x(s) - x(t)| \geq \epsilon/2 \right\} \right) \leq \eta\delta/2.
\end{gathered} \tag{20}$$

Davidson (2002, Lemma 27.13) implies Lemma 3-(b).

□

Proof of (11)

Since $t_k - t_{k-2} = h_k + h_{k-1}$, (11) follows immediately once the next relation is established:

$$\left| \frac{\Phi_n^*(t_k, u) - \Phi_n^*(t_{k-2}, u)}{h_k + h_{k-1}} + \frac{u^2}{2} \cdot \frac{\theta_{t_{k-2}}}{\int_0^t \theta_v dv} \cdot \Phi_n^*(t_{k-2}, u) \right| \leq \mathcal{R}_n \overset{a.s.}{\to} 0. \tag{21}$$

For any $\zeta > 0$,

$$\begin{aligned} h_k^{-1}E^*[|c(u\Delta_{k-1})|] &\leq h_k^{-1}E^*[\min\{|u\Delta_{k-1}|^2, |u\Delta_{k-1}|^3/6\}] \\ &= h_k^{-1}E^*[\min\{|u\Delta_{k-1}|^2, |u\Delta_{k-1}|^3/6\}(1_{\{|\Delta_{k-1}|^2\geq\zeta h_k\}} + 1_{\{|\Delta_{k-1}|^2<\zeta h_k\}})] \\ &\leq |u|^2E^*[h_k^{-1}\Delta_{k-1}^2 1_{\{|\Delta_{k-1}|^2\geq\zeta h_k\}}] + (|u|^3/6)h_k^{-1}E^*[|\Delta_{k-1}|^3 1_{\{|\Delta_{k-1}|^3<(\zeta h_k)^{3/2}\}}] \\ &\leq |u|^2E^*[h_k^{-1}\Delta_{k-1}^2 1_{\{h_k^{-1}|\Delta_{k-1}|^2\geq\zeta\}}] + |u|^3\zeta^{3/2}h_k^{1/2}. \end{aligned} \quad (22)$$

By substituting (22) and (9) in the right hand side of (8) and using the add-subtract trick, law of iterated expectations, triangular inequality, $|e^{i\bar{Z}_{k-2}}| = 1$ and $(h_k + h_{k-1})^{-1} \leq h_k^{-1}$, the left hand side of (8) is dominated by

$$\begin{aligned} &|u|E^*[|E[h_k^{-1}\Delta_{k-1}|\mathcal{I}_{n,\lfloor nt_{k-2}\rfloor}]|] + (u^2/2)E^*[|E[(h_{k-1}+h_k)^{-1}\Delta_{k-1}^2|\mathcal{I}_{n,\lfloor nt_{k-2}\rfloor}] - \theta_{t_{k-2}}|] \\ &\qquad + u^2E^*[h_k^{-1}\Delta_{k-1}^2 1_{\{h_k^{-1}\Delta_{k-1}^2\geq\zeta\}}] + |u|^3\zeta^{3/2}h_k^{1/2}. \end{aligned} \quad (23)$$

Let us employ the notation $E^*[A_n] = E[A_n|\mathcal{G}](\omega)$ for $\omega \in \Omega$ to emphasize that E^* is a $\mathcal{G}$-measurable random variable. Because $\Omega = \bigcup_{F\in\mathcal{G}_1} F$,

$$\begin{aligned} &\mathcal{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} E^*[A_n] \neq c\right\}\right) = \mathcal{P}\left(\left\{\omega \in \bigcup_{F\in\mathcal{G}_1} F : \lim_{n\to\infty} E[A_n|\mathcal{G}](\omega) \neq c\right\}\right) \\ &\leq \sum_{F\in\mathcal{G}_1:\mathcal{P}(F)>0} \mathcal{P}\left(\left\{\omega \in F : \lim_{n\to\infty} E[A_n|\mathcal{G}](\omega) \neq c\right\}\right) = \sum_{F\in\mathcal{G}:\mathcal{P}(F)>0} \mathcal{P}\left(\lim_{n\to\infty} E[X_n|F] \neq c\right) \end{aligned} \quad (24)$$

for some constant c. To establish $E^*[A_n] \overset{a.s.}{\to} c$, therefore, it is sufficient to confirm $E[A_n|G] \to c$ for each $G \in \mathcal{G}$ with $\mathcal{P}(G) > 0$. Motivated by this fact, let us replace E^* in (23) by $E[\cdot|G]$ for $G \in \mathcal{G}$ with $\mathcal{P}(G) > 0$. Using a similar technique as in (5), the G-conditional version of (23) is dominated by the following divided by $\mathcal{P}(G)$:

$$\begin{aligned} &|u|\|E[h_k^{-1}\Delta_{k-1}|\mathcal{I}_{n,\lfloor nt_{k-2}\rfloor}]\|_1 + (u^2/2)\|E[(h_{k-1}+h_k)^{-1}\Delta_{k-1}^2|\mathcal{I}_{n,\lfloor nt_{k-2}\rfloor}] - \theta_{t_{k-2}}\|_1 \\ &\qquad + u^2E[h_k^{-1}\Delta_{k-1}^2 1_{\{h_k^{-1}\Delta_{k-1}^2\geq\zeta\}}] + |u|^3\zeta^{3/2}h_k^{1/2}. \end{aligned} \quad (25)$$

(i) $h_k^{-1}\|E[\Delta_{k-1}|\mathcal{I}_{n,\lfloor nt_{k-2}\rfloor}]\|_1$: using Lemma 4-(a), the Minkowski/Cauchy-Schwarz/norm inequalities, $\lfloor nt_{k-1}\rfloor - \lfloor nt_{k-2}\rfloor \geq \lfloor nh_{k-1}\rfloor$ and mixingale properties,

$$\begin{aligned} h_k^{-1}\|E[\Delta_{k-1}|\mathcal{I}_{n,\lfloor nt_{k-2}\rfloor}]\|_1 &\leq h_k^{-1}\|E[\Delta_{k-1}|\mathcal{I}_{n,\lfloor nt_{k-2}\rfloor}]\|_2 \\ &\leq \left(h_k^{-1}\sum_{j=\lfloor nt_{k-1}\rfloor}^{\lfloor nt_k\rfloor} c_{n,j}^2\right)^{1/2}\left(h_k^{-1}\sum_{r=\lfloor nh_{k-1}\rfloor}^{\infty} \psi_r^2\right)^{1/2}. \end{aligned}$$

The condition-(c) bounds the first square root. (ψ_r) of size $-1/2$ implies $\psi_r^2 = O(r^{-1-2\epsilon})$ for some $\epsilon > 0$. Then, $r^\epsilon\psi_r^2 = O(r^{-1-\epsilon})$ and therefore

$$h_k^{-1}\sum_{r=\lfloor nh_{k-1}\rfloor}^{\infty} \psi_r^2 \leq h_k^{-1}\lfloor nh_k\rfloor^{-\epsilon}\sum_{r=\lfloor nh_{k-1}\rfloor}^{\infty} r^\epsilon\psi_r^2 = o(h_k^{-1}\lfloor nh_k\rfloor^{-\epsilon}),$$

which tends to zero if $h_k \propto n^{-a}$ and $h_{k-1} \propto n^{-b}$ with $b = \epsilon/(1+\epsilon) \in (0,1)$ and $0 < a < b$.

(ii) $\|E[(h_{k-1}+h_k)^{-1}\Delta_{k-1}^2|\mathcal{I}_{n,\lfloor nt_{k-2}\rfloor}]-\theta_{t_{k-2}}\|_1$: using $(h_k+h_{k-1})^{-1} = h_k^{-1}\{1-(h_k+h_{k-1})^{-1}h_{k-1}\}$ and $(h_k + h_{k-1})^{-1}h_{k-1} \le h_k^{-1}h_{k-1}$ from (7), it has the majorant

$$\|E[h_k^{-1}\Delta_{k-1}^2|\mathcal{I}_{n,\lfloor nt_{k-2}\rfloor}] - \theta_{t_{k-1}}\|_1 + (h_k^{-1}h_{k-1})\|E[h_k^{-1}\Delta_{k-1}^2|\mathcal{I}_{n,\lfloor nt_{k-2}\rfloor}]\|_1 + \|\theta_{t_{k-1}} - \theta_{t_{k-2}}\|_1.$$

Define $Y_{n,k} := E[h_k^{-1}\Delta_{k-1}^2|\mathcal{I}_{n,\lfloor nt_{k-2}\rfloor}]$. The first term tends to zero if $Y_{n,k} - \theta_{t_{k-1}}$ does in L_1 as $n \to \infty$. Condition-(e) and Lemma 1-(a) guarantee that it is equivalent to the uniform integrability of $(Y_{n,k} - \theta_{t_{k-1}})_{n,k\in\mathbb{N}}$. Now the claim follows by the uniform integrability of $(\theta_t)_{t\in[0,1]}$ as is assumed in (e), that of $(h_k^{-1}\Delta_{k-1}^2)$ confirmed in Lemma 4, and Lemma 1-(b), (c), (d). By Lemma 4 and Lemma 1-(b)-(i), $(h_k^{-1}\Delta_{k-1}^2)$ is L_1-bounded. Therefore, the second term is dominated by $h_k^{-1}h_{k-1}\|h_k^{-1}\Delta_{k-1}^2\|_1 = O(h_k^{-1}h_{k-1}) = o(1)$ given $h_k^{-1}h_{k-1} \to 0$ from (7). The last term tends to zero by the $\mathcal{P}$-a.s. continuity of θ_t and $h_{k-1} \to 0$.

(iii) $E[h_k^{-1}\Delta_{k-1}^2 1_{\{h_k^{-1}|\Delta_{k-1}|^2 \ge \zeta\}}] \to 0$ as $\zeta \to \infty$ because $(h_k^{-1}\Delta_{k-1}^2)$ is uniformly integrable.

(iv) $\zeta^{3/2}h_k^{1/2} \to 0$ by designing $\zeta = o(h_k^{-1/3})$ for $h_k \to 0$.

(21) holds because (25) vanishes $\mathcal{P}$-almost surely. The case of $v = t_{k-1}$ follows by combining (11), (7) and continuity of θ_t. The specified range of v in (12) is justified because t_{k-2} is arbitrary as long as it is greater than t_{k-3}, even if t_k and t_{k-1} approach t_{k-2} under (7).

□

Proof of (14)

Apply the law of iterated expectations, add-subtract and multiply-divide tricks and the Minkowski's inequality to obtain the estimate

$$\|\mathcal{B}_n(t) - \bar{\mathcal{B}}_n(t)\|_2 \le \|\mathcal{B}_n(t_{k-1}) - \mathcal{B}_n(t_{k-2})\|_2 \le h_{k-1}^{1/2}\left\|\left(\int_0^t \theta_v dv\right)^{-1}\theta_{t_{k-2}}\right\|_1^{1/2}$$
$$+h_{k-1}^{1/2}\left\|\left(\int_0^t \theta_v dv\right)^{-1} E\left[h_{k-1}^{-1}\left(\sum_{j=\lfloor nt_{k-2}\rfloor+1}^{\lfloor nt_{k-1}\rfloor} X_{n,j}\right)^2 - \theta_{t_{k-2}}\middle|\mathcal{I}_{n,\lfloor ns\rfloor}\right]\right\|_1^{1/2}$$

Because $\int_0^t \theta_v dv$ is uniformly bounded away from zero, the same argument as for (ii) in the proof of Theorem 1 reveals that both terms tend to zero given (7).

□

Appendix II. Technical lemmas for uniform tightness

Lemma 5 (Davidson 2002, Theorem 28.12; Billingsley 1968, Theorem 14.3) *$K \subset \mathcal{D}$ is relatively compact, i.e., the closure of K, denoted by $cl(K)$, is compact in $(\mathcal{D}, d_B)$, if*

(a) $\sup_{x\in K}\sup_{t\in[0,1]}|x(t)| < \infty$, *and*

(b) $\lim_{\delta\to\infty}\sup_{x\in K} w'(x,\delta) = 0$ *where* $w'(x,\delta) := \inf_{\Pi_\delta}\max_{i=1\ldots r}\sup_{s,t\in[t_{i-1},t_i)}|x(s) - x(t)|$, *and* Π_δ *is a partition* $(t_i)_{i=1,\ldots,r}$ *of* $[0,1]$ *such that* $r \le \lfloor 1/\delta\rfloor$ *and* $\min_{i=2,\ldots,r}(t_i - t_{i-1}) > \delta$.

Lemma 5 is about the topological structure of $\mathcal{D}$ and is independent of any induced laws on it.

Lemma 6 (If part of Davidson 2002, Theorem 28.13) *Let $w'(x,\delta)$ be as in Lemma 5. A sequence of probability laws (μ_n) on $(\mathcal{D}, d_B, \mathbb{B})$ is uniformly tight if*

(a) $\exists N \in \mathbb{N}\ \forall n \geq N\ \forall \eta > 0\ \exists M < \infty$: $\mu_n(\{x : \sup_{t\in[0,1]} |x(t)| > M\}) \leq \eta$, and

(b) $\exists N \in \mathbb{N}\ \forall n \geq N\ \forall \eta > 0\ \forall \epsilon > 0\ \exists \delta \in (0,1)$: $\mu_n(\{x : w'(x,\delta) \geq \epsilon\}) \leq \eta$.

Proof of Lemma 3

The proof traces Davidson (2002, Theorem 28.14) with minor modifications. $w'(x,\delta/2) \geq \epsilon$ implies $w(x,\delta) \geq \epsilon$ because $w'(x,\delta/2) \leq w(x,\delta)$ (Davidson 2002, p.458). Combined with (ii), $\mu_n(\{x : w'(x,\delta/2) \geq \epsilon\}) \leq \mu_n(\{x : w(x,\delta) \geq \epsilon\}) \leq \eta$, which is the condition (b) in Lemma 6. Select any $\epsilon, \eta' > 0$. For such arbitrary $\eta' > 0$, (i) implies the existence of $M' \in (0,\infty)$ such that $\mu_n(\{x : |x(0)| > M'\}) \leq \eta'$. Define $k := 1 + \lfloor 1/\delta \rfloor$ so that $k\delta > 1$ or $1/k < \delta$. For $i = 1 \dots k$ and $t \in [0,1]$, it is possible to find $\delta > 0$ and k such that

$$\mu_n(\{x : \sup_{t\in[0,1]} |x(ti/k) - x(t(i-1)/k)| \geq \epsilon\}) \leq \mu_n(\{x : w(x,\delta) \geq \epsilon\}) \leq \eta' \tag{26}$$

because $ti/k - t(i-1)/k = t/k \leq 1/k < \delta$ so that $\sup_{t\in[0,1]} |x(ti/k) - x(t(i-1)/k)| \leq w(x,\delta)$, allowing the application of (ii). By the add-subtract and telescopic-sum argument combined with the triangular inequality, $|x(t)| \leq |x(0)| + \sum_{i=1}^{k} |x(ti/k) - x(t(i-1)/k)|$ so that $\sup_{t\in[0,1]} |x(t)| \geq \alpha$ implies $|x(0)| + \sum_{i=1}^{k} \sup_{t\in[0,1]} |x(ti/k) - x(t(i-1)/k)| \geq \alpha$. For any non-negative A, B,

$$\mu_n(A + B \geq \alpha + \beta) \leq \mu_n(\{A \geq \alpha\} \cup \{B \geq \beta\}) \leq \mu_n(A \geq \alpha) + \mu_n(B \geq \beta).$$

By a repeated application of this inequality,

$$\begin{aligned} \mu_n(\{x : \sup_{t\in[0,1]} |x(t)| > M' + k\epsilon\}) &\leq \mu_n(\{x : |x(0)| + \sum_{i=1}^{k} \sup_{t\in[0,1]} |x(ti/k) - x(t(i-1)/k)| > M' + k\epsilon\}) \\ &\leq \mu_n(\{x : |x(0)| > M'\}) + \sum_{i=1}^{k} \mu_n(\{x : w(x,\delta/2) \geq \epsilon\}) \leq (1+k)\eta'. \end{aligned}$$

Because η' and ϵ can be arbitrarily small whereas δ and k are selected accordingly, we can make $\eta := (1+k)\eta'$ arbitrarily small and select $M := M' + k\epsilon \in (0,\infty)$ accordingly to guarantee the condition (a) in Lemma 6. Consequently, the sequence (μ_n) is uniformly tight.

Suppose $\mu = \lim_{k\to\infty} \mu_{n_k}$, i.e. the limit of a converging subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ of $(\mu_n)_{n\in\mathbb{N}}$. For $\epsilon = \eta = 1/j$, define $A := \{x : w(x,\delta) \geq \epsilon\}$ and $B_j := \{x : w(x,\delta_j) \geq 1/j\}$ for some $\delta_j > 0$. We can select (δ_j) as a decreasing sequence of positive reals such that $B_j \subset int(A)$, which is the interior of A. Because the weak convergence of $(\mu_{n_k})_{k\in\mathbb{N}}$ is characterized by $\liminf_{k\to\infty} \mu_{n_k}(\mathcal{O}) \geq \mu(\mathcal{O})$ for any open set $\mathcal{O} \in \mathbb{B}$ (Davidson 2002, Theorem 26.10-(c)), $int(A) \subset A$ and the condition (ii) ensure

$$\mu(B_j) \leq \mu(int(A)) \leq \liminf_{k\to\infty} \mu_{n_k}(int(A)) \leq \liminf_{k\to\infty} \mu_{n_k}(A) \leq \eta = 1/j. \tag{27}$$

This means that $\mu(\bigcap_{j\geq m} B_j) = 0$ for any finite $m \in \mathbb{N}$. If we define $B := \bigcup_{m=1}^{\infty} \bigcap_{j=m}^{\infty} B_j$, $\mu(B) = \mu(\bigcup_{m=1}^{\infty} \bigcap_{j=m}^{\infty} B_j) \leq \sum_{m=1}^{\infty} \mu(\bigcap_{j=m}^{\infty} B_j) = 0$ or $\mu(B^c) = 1$. It suffices to show $B^c \subset \mathcal{C}$. Because $B^c = \bigcap_{m=1}^{\infty} \bigcup_{j=m}^{\infty} B_j^c = \{x : \forall m \geq 1; \exists j \geq m : w(x,\delta_j) < 1/j\}$, any $x \in B^c$ satisfies $\lim_{\delta\to 0} w(x,\delta) = 0$. This is the case when x has a continuous path.

Finally, the above result applies to the conditional laws $\mu_n(\cdot|G)$ for any $G \in \mathcal{G}$ with $\mathcal{P}(G) > 0$ for any majorant sides divided by $\mathcal{P}(G)$:

- $\mu_n(\{x : w'(x, \delta/2) \geq \epsilon\}|G) \leq \eta/\mathcal{P}(G)$ but the right hand side can be arbitrarily small;
- $\mu_n(\{x : \sup_{t\in[0,1]} |x(t)| > M' + k\epsilon\}|G) \leq (1 + k)\eta'/\mathcal{P}(G)$ similarly; and
- for $\mu(\cdot|G) = \lim_{k\to\infty} \mu_{n_k}(\cdot|G)$, $\mu(B|G) = 0$ or $\mu(B^c|G) = 1$, implying $\mu^*(\mathcal{C}) \overset{a.s.}{=} 1$ because $B^c \subset \mathcal{C}$ regardless of the conditioning and a similar argument as for (24) implies that $\mathcal{P}(\{\omega \in \Omega : \mu^*(B)(\omega) \neq 0\}) \leq \sum_{G\in\mathcal{G}:\mathcal{P}(G)>0} \mathcal{P}(\mu(B|G) \neq 0) = 0$.

□

References

Barndorff-Nielsen, O. E., P. Hansen, A. Lunde and N. Shephard (2008). Designing Realized Kernels to Measure the Ex Post Variation of Equity Prices in the Presence of Noise. *Econometrica, Vol. 76, No. 6*, 1481-1536.

Billingsley, P. (1968). *Convergence of Probability Measures*, 1st ed., John Wiley and Sons.

Chow, Y.S. and H. Teicher (1997). *Probability Theory*, 3rd ed., Springer.

Davidson, J. (2002) *Stochastic Limit Theory*, 2nd ed., Oxford University Press.

De Jong, R. M. (1997). Central Limit Theorems for Dependent Heterogeneous Random Variables. *Econometric Theory, Vol. 13*, 353-367.

Gut, A. (2005) *Probability: A Graduate Course.* Springer.

Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed., Springer.

Kalnina, I. and O. Linton (2008). Estimating Quadratic Variation Consistently in the Presence of Endogenous and Diurnal Measurement Error. *Journal of Econometrics, Vol. 147*, 47-59.

McLeish, D.L. (1975). Invariance Principles for Dependent Variables. *Z. Wahrscheinlichkeitstheorie verw. Gebiete 32*, 165-178.

McLeish, D.L. (1977). On the Invariance Principle for Nonstationary Mixingales. *Annals of Probability*, Vol. 5 No. 4, 616-621.

Pötscher, P. M. and I. R. Prucha (1991). Basic Structure of the Asymptotic Theory in Dynamic Nonlinear Econometric Models, Part I: Consistency and Approximation Concepts. *Econometric Reviews, Vol. 10*, 125-216.

Wooldridge, J. and H. White (1988). Some Invariance Principles and Central Limit Theorems for Dependent Heterogeneous Processes. *Econometric Theory, Vol.4*, 210-230.