Econometric Theory, **14**, 1998, 260–284. Printed in the United States of America.

CENTRAL LIMIT AND FUNCTIONAL CENTRAL LIMIT THEOREMS FOR HILBERT-VALUED DEPENDENT HETEROGENEOUS ARRAYS WITH APPLICATIONS

XIAOHONG CHEN
University of Chicago

HALBERT WHITE
University of California, San Diego

We obtain new central limit theorems (CLT's) and functional central limit theorems (FCLT's) for Hilbert-valued arrays near epoch dependent on mixing processes, and also new FCLT's for general Hilbert-valued adapted dependent heterogeneous arrays. These theorems are useful in delivering asymptotic distributions for parametric and nonparametric estimators and their functionals in time series econometrics. We give three significant applications for near epoch dependent observations: (1) A new CLT for any plug-in estimator of a cumulative distribution function (c.d.f.) (e.g., an empirical c.d.f., or a c.d.f. estimator based on a kernel density estimator), which can in turn deliver distribution results for many Von Mises functionals; (2) a new limiting distribution result for degenerate *U*-statistics, which delivers distribution results for Bierens's integrated conditional moment tests; (3) a new functional central limit result for Hilbert-valued stochastic approximation procedures, which delivers distribution results for nonparametric recursive generalized method of moment estimators, including nonparametric adaptive learning models.

1. INTRODUCTION

Central limit theorems (CLT's) and functional central limit theorems (FCLT's) are powerful tools for obtaining asymptotic distribution results for estimators of econometric models. Recently, attention has been increasingly devoted in econometrics to the study of semiparametric and nonparametric estimators. Because such estimators take values in certain infinite-dimensional spaces, standard finite-dimensional CLT's and FCLT's are not applicable. Further, because of the time series nature of much economic data, infinite-dimensional CLT's and FCLT's for dependent processes are required.

We thank Don Andrews, Yanquin Fan, Max Stinchcombe, Peter C.B. Phillips, and two referees for helpful comments. White's participation was supported by NSF Grants SES-8921382 and SBR-9511253. Address correspondence to: Halbert White, Department of Economics 0508, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0508, USA; e-mail: hwhite@albert.ucsd.edu.

 © 1998 Cambridge University Press 0266-4666/98 $9.50

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

For a great many applications, a separable Hilbert space is a sufficiently rich infinite-dimensional space. CLT's and FCLT's for certain kinds of dependent processes in such Hilbert spaces are already available. For example, a CLT for strictly stationary strong or uniform mixing processes was given by Dehling (1983), and a CLT for martingale difference sequences was given by Jakubowski(1980). Walk (1977) gave an FCLT for Hilbert space–valued martingale difference sequences. However, neither of these dependence conditions is necessarily plausible as a general description of nonexplosive time series of the sort often encountered in economics (e.g., Andrews, 1985). More satisfactory is the mixingale condition (McLeish, 1975), which contains both mixing and martingale difference processes as special cases, and also functions of infinite histories of mixing processes (Billingsley, 1968, Sect. 21; McLeish, 1975; Gallant and White, 1988). Because these functions depend mainly on the "near epoch" of the underlying mixing process, they are called "near epoch dependent."

Hilbert space CLT's and FCLT's explicitly applicable to the mixingale case are not presently available. Our purpose here is to provide such results, complementing our earlier law of large number results for Hilbert space mixingales (Chen and White, 1996). By suitably applying the characteristic functional approach of Vakhania (1981) and Theorem 1.4.2 of Billingsley (1968) to a separable Hilbert space, the desired CLT's and FCLT's for partial sums of Hilbert space–valued mixingales follow straightforwardly. Our method of proof includes two steps: first, we establish some sufficient conditions for tightness for Hilbert-valued near epoch dependent processes; second, we demonstrate that the CLT or FCLT holds in finite-dimensional Euclidean spaces obtained by projecting the underlying Hilbert space. This demonstration is easily achieved by using results for Euclidean space–valued mixingales, such as those of McLeish (1977) or Wooldridge and White (1988, 1989).

The paper is organized as follows. Section 2 motivates our research with some specific examples. Section 3 gives CLT's for $\mathbb{H}$-valued dependent possibly heterogeneous double arrays. Section 4 gives FCLT's for such processes. Section 5 applies our theorems to obtain asymptotic distribution results for the examples of Section 2. All proofs appear in the Mathematical Appendix.

2. MOTIVATION

We begin by outlining a variety of motivating applications useful for both parametric and nonparametric estimation and inference in time series econometrics.

Example 2.1. (Plug-in estimators of a cumulative distribution function and its functionals)

Let $\{X_t; t = 1,2,\ldots\}$ be an $\mathbb{R}^d$-valued weakly stationary sequence. Suppose that X_t has a common marginal cumulative distribution function (c.d.f.) F and a corresponding density f. Let $\hat{F}_n$ be any plug-in estimator for F, i.e., $\hat{F}_n(x) = \int_{(-\infty,x]} \hat{f}_n(y)\, dy$ for any $x \in \mathbb{R}^d$, where $\hat{f}_n$ is any estimator of f. For example,

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

(i) (the kernel c.d.f. estimator $\hat{F}_{k,n}$): $\hat{f}_n(y) = [n(b_n)^d]^{-1} \sum_{t=1}^n K((y - X_t)/b_n)$ for any $y \in \mathbb{R}^d$, where K is a square-integrable nonnegative kernel with $\int K(y)dy = 1$, and $b_n \to 0$ as $n \to \infty$.

(ii) (the orthogonal series c.d.f. estimator $\hat{F}_{o,n}$): $\hat{f}_n(y) = \sum_{i=1}^{r_n} n^{-1} \sum_{i=1}^n \phi_i(X_t)\phi_i(y)$ for any $y \in \mathbb{R}^d$, where $\{\phi_i\}$ is a basis for $L_2(\mathbb{R}^d)$ and $r_n \to \infty$ as $n \to \infty$.

(iii) (the empirical distribution $\hat{F}_{e,n}$): $\hat{F}_{e,n}(x) = n^{-1} \sum_{t=1}^n \mathbf{1}(X_t \leq x)$ for any $x \in \mathbb{R}^d$.

Suppose that $\{X_t\}$ is a sequence near epoch dependent on some underlying mixing processes; then $\{\hat{F}_n - E\hat{F}_n\}$ is an $L_2(\mathbb{R}^d)$-valued near epoch dependent sequence in cases (i) and (ii). The theorems in the next sections will give a CLT in $L_2(\mathbb{R}^d)$ for $n^{1/2}[\hat{F}_n(\cdot) - F(\cdot)]$ with near epoch dependent observations $\{X_t\}$. To treat (iii), we first consider a smoothed version $\hat{F}_{se,n}(x) = n^{-1} \sum_{t=1}^n \mathcal{K}_n(x - X_t)$, where $\{\mathcal{K}_n : n \geq 1\}$ is a sequence of Lipschitz continuous distribution functions that converges as $n \to \infty$ (in a proper sense) to the step function $\mathbf{1}(u) = 1$ for $u \geq 0$; $\mathbf{1}(u) = 0$ for $u < 0$. For example, if we choose $\mathcal{K}_n(u) = \int_{(-\infty,u]} K(z/b_n)b_n^{-d}\,dz$, then we have the kernel c.d.f. estimator in (i). Clearly $\hat{F}_{se,n} - E\hat{F}_{se,n}$ is a near epoch dependent sequence in a Hilbert space, say $L_2(\mathbb{R}^d, \mu)$, where μ is some σ-finite measure, as long as $\{X_t\}$ is a near epoch dependent sequence in $\mathbb{R}^d$. Under some mild conditions, we can show that $n^{1/2}[\hat{F}_{se,n} - \hat{F}_{e,n}] = o_p(1)$ in $L_2(\mathbb{R}^d, \mu)$. We thus obtain a Hilbert-valued CLT for the empirical distribution $\hat{F}_{e,n}$ with near epoch dependent observations.

This is a powerful result. For example, it can be applied to give distribution results for $T(\hat{F}_n) - T(F)$, where $T(\cdot)$ is a smooth (Von Mises) functional of F, including such statistics as sample moments, maximum likelihood estimates, method of moments estimates, minimum distance estimates, sample quantiles, quantile regression estimates, trimmed means, and certain U- and V-statistics. (For valuable background, see Serfling, 1980, Ch. 6; Manski, 1988.) It can also be applied to derive asymptotic distribution results for the bootstrap with dependent heterogeneous data, generalizing Politis and Romano's (1994) results for the stationary strong mixing bootstrap.

Example 2.2. (Degenerate *U*-statistics)

Let $\mathcal{X}$ be a measurable space and $\{X_t\}$ be a sequence of $\mathcal{X}$-valued stationary random variables near epoch dependent on a mixing process. Suppose $K(\cdot,\cdot)$: $\mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is a real-valued symmetric square-integrable function such that $\int K(y,x)dF(x) = 0$ for almost all $y \in \mathcal{X}$, where F is the distribution of X_t. Consider the U-statistic $U_n = 2[n(n-1)]^{-1}\Sigma_{1\leq i<j\leq n}K(X_i, X_j)$. By a theorem in the next section and under conditions specified later, we can show that $nU_n \Rightarrow \sum_{j=1}^\infty \lambda_j(Z_j^2 - 1)$ as $n \to \infty$, where $\{\lambda_j\}$ is the sequence of nonzero eigenvalues for K and $\{Z_j\}$ is a sequence of correlated, real-valued normal random variables.

The existing literature gives limiting distribution results for degenerate U-statistics allowing only for stationary uniform mixing sequences (see Eagleson, 1979; Carlstein, 1988) or for stationary strong mixing sequences when the number of nonzero eigenvalues is finite (Carlstein, 1988). Our treatment significantly relaxes these conditions. Among other things, this result is well-suited to

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

deriving asymptotic distribution results for versions of the integrated conditional moment test proposed by Bierens (1982) and analyzed for martingale difference sequences by Bierens and Plöberger (1997).

The preceding examples use CLT's in Hilbert space. FCLT's in Hilbert space also have important applications in economics and econometrics. An interesting class of examples is that of the nonparametric adaptive leaning procedures studied in Chen and White (1992b, 1993). Most of the adaptive learning models in macroeconomics require that an economic agent solve a conditional Euler equation of the form $E[\Psi(Y_t, X_t, \theta)|X_t] = 0$ in real time, where X_t can contain lagged Y's, Ψ is a known function, and θ is the unknown function that the agent is learning about. For example, take $\Psi(Y_t, X_t, \theta) = Y_t - \theta$; then the agent tries to learn the conditional mean function $E[Y_t|X_t = \cdot] = \theta^*(\cdot)$ over time. The agent can learn θ^* using the following simple procedure:

$$\hat{\theta}_{n+1}(\cdot) = \hat{\theta}_n(\cdot) + n^{-1}\Psi(Y_n, X_n, \hat{\theta}_n(\cdot))K((\cdot - X_n)/b_n)/b_n^d,$$

where $K: \mathbb{R}^d \to \mathbb{R}$ is a kernel density function and b_n is the bandwidth. More generally, we consider the following algorithm.

Example 2.3. (Stochastic approximation in a Hilbert space)

Let $\mathbb{H}$ be a separable Hilbert space. Let $\overline{M}: \mathbb{H} \to \mathbb{H}$ have a unique unknown root θ^* in $\mathbb{H}$, and suppose $\overline{M}$ is observed subject to some error η_n. We can recursively estimate θ^* by taking $\hat{\theta}_0$ to be an arbitrary $\mathbb{H}$-valued random element, and

$$\hat{\theta}_{n+1} = \hat{\theta}_n + a_n M_n(Z_n, \hat{\theta}_n) \equiv \hat{\theta}_n + a_n[\overline{M}(\hat{\theta}_n) + \eta_n(Z_n, \hat{\theta}_n)], \qquad n = 1, 2, \ldots,$$

where $\{a_n\}$ is a nonincreasing sequence with $\sum_{n=1}^{\infty} a_n = \infty$, $\{Z_n\}$ is an exogenous random sequence taking values in some Hilbert space $\mathbb{G}$, and $\eta_n(Z_n, \theta) \equiv M_n(Z_n, \theta) - \overline{M}(\theta)$. This method was introduced by Robbins and Monro (1951) for $\mathbb{R}^1$ and generalized to abstract $\mathbb{H}$ by Venter (1966) with $\{\eta_n\}$ being a martingale difference process not depending on θ. Suppose that $M_n(z, \theta)$ is Lipschitz in z for each fixed θ. Then $\{M_n(Z_n, \theta^*) - E[M_n(Z_n, \theta^*)]\}$ is a Hilbert-valued near epoch dependent sequence when $\{Z_n\}$ is appropriately near epoch dependent. If the "bias" $E[M_n(Z_n, \theta^*)]$ goes to zero at a proper rate, our theorems in the next sections can establish asymptotic normality for $\hat{\theta}_n - \theta^*$, suitably normalized, via the FCLT.

3. CLT'S FOR $\mathbb{H}$-VALUED DEPENDENT HETEROGENEOUS ARRAYS

3.1. Background for $\mathbb{H}$-Valued Arrays

Here we only give a brief review of needed definitions and properties; more detail can be found in Parthasarathy (1967), Araujo and Gine (1980), Ledoux and Talagrand (1991), and Chen and White (1992a, 1996). Throughout, let $\mathbb{H}$ denote a separable Hilbert space with inner product $(\cdot, \cdot)$ and norm $|\cdot|$, with the usual Borel σ-field. An $\mathbb{H}$-valued random element ($\mathbb{H}$-r.e.) W (or the distribution μ of an $\mathbb{H}$-r.e. W) has mean (or expectation) $EW \in \mathbb{H}$ if $E[(W, h)] = (EW, h)$ for all $h \in$

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

$\mathbb{H}$; it has finite rth moment if $E[|W|^r]^{1/r} < \infty$ (denoted as $\|W\|_r < \infty$ for $1 \le r < \infty$); it is *weakly second order* if $E[(W,h)^2] \equiv \int_{\mathbb{H}} (w,h)^2 \mu(dw) < \infty$ for all $h \in \mathbb{H}$. For a weakly second order $\mathbb{H}$-r.e. W with expectation $E[W]$, we define the *covariance* of W, Cov $W: \mathbb{H} \times \mathbb{H} \to \mathbb{R}$, as

$$\text{Cov } W(h,g) \equiv E[(W - E[W],h)(W - E[W],g)], \quad \text{for any } h,g \in \mathbb{H}.$$

This is a symmetric, positive, and continuous bilinear form. Alternatively, we can define a *covariance operator* $S_W: \mathbb{H} \to \mathbb{H}$ as $S_W h \equiv E((W - E[W],h)[W - E[W]])$ for any $h \in \mathbb{H}$. By the definitions we have

$$(S_W h)(g) \equiv (S_W h, g) = \text{Cov } W(h,g) \quad \text{for any } h,g \in \mathbb{H}.$$

Let $\mathcal{R}(\mathbb{H})$ be the set of all covariance operators in $\mathbb{H}$, i.e., the set of all self-adjoint positive operators in $L(\mathbb{H},\mathbb{H})$, where $L(\mathbb{H},\mathbb{H})$ denotes the space of bounded linear operators with the operator norm $\|\cdot\|$ defined as $\|S\| = \sup[\|Sh\| : \|h\| \le 1, h \in \mathbb{H}]$. For a compact operator S in $L(\mathbb{H},\mathbb{H})$, if $\sum_{j=1}^{\infty} |(Se_j,e_j)| < \infty$ for all complete orthonormal systems (CONS) $\{e_j\}$ (so that the sum $\sum_{j=1}^{\infty}(Se_j,e_j)$ is independent of the choice of $\{e_j\}$), we call S a *nuclear* operator and $\text{tr}(S) \equiv \sum_{j=1}^{\infty}(Se_j,e_j)$, the (matrix) *trace* of S. Or equivalently, a compact operator S is *nuclear* if $\sum_{j=1}^{\infty} \lambda_j([S^*S]^{1/2}) < \infty$, where $\{\lambda_j([S^*S]^{1/2}) : j \ge 1\}$ is the sequence of eigenvalues of $[S^*S]^{1/2}$, each repeated according to its multiplicity, where S^* is the adjoint of S. When a nuclear operator S is self-adjoint, we have $\text{tr}(S) = \sum_{j=1}^{\infty} \lambda_j(S)$. Let $\mathcal{S}(\mathbb{H})$ be the set of all self-adjoint positive nuclear operators. For any $S, J \in \mathcal{S}(\mathbb{H})$, we define a metric $\|S - J\|_{\text{tr}} \equiv \text{tr}(S - J)$, which is stronger than the operator norm metric $\|\cdot\|$. Then $(\mathcal{S}(\mathbb{H}), \|\cdot\|_{\text{tr}})$ is a complete separable (Polish) space. Note that $\mathcal{S}(\mathbb{H}) \subset \mathcal{R}(\mathbb{H})$, and $\mathcal{S}(\mathbb{H}) = \mathcal{R}(\mathbb{H})$ if and only if $\mathbb{H}$ is finite-dimensional. In fact, $\mathcal{S}(\mathbb{H})$ is the set of all covariance operators of Gaussian measures on $\mathbb{H}$.

An $\mathbb{H}$-r.e. $\mathcal{N}$ has a *Gaussian distribution on* $\mathbb{H}$ if for all $h \in \mathbb{H}$, the real-valued random variable $(h,\mathcal{N})$ has a Gaussian distribution on $\mathbb{R}$.

Although one can obtain various central limit results by applying conditions in Parthasarathy (1967) or in Ledoux and Talagrand (1991), it is particularly easy to obtain the following lemma by applying Vakhania's (1981) characteristic functional approach.

LEMMA 3.1. *Let* $\{W_{n,i} : 1 \le i \le a_n; n \ge 1\}$ *be a double array of* $\mathbb{H}$-r.e.*'s with zero means and finite second moments. Suppose the following conditions hold*:

(a) *for each* $0 \ne h \in \mathbb{H}$, $\Sigma_{1 \le j \le a_n}(W_{n,j},h) \Rightarrow \mathcal{N}(0,\sigma^2(h))$ *in* $\mathbb{R}$, *where* $\sigma^2(h)$ *is a positive finite number depending on* $0 \ne h \in \mathbb{H}$;
(b) $\sup_{n \ge 1}[E|\Sigma_{1 \le j \le a_n} W_{n,j}|^2] \le C < \infty$;
(c) *Let* $\{e_l\}$ *be a complete orthonormal basis of* $\mathbb{H}$ *with*

$$\lim_{k \to \infty} \sup_{n \ge 1} \Sigma_{k \le l < \infty} E[(\Sigma_{1 \le j \le a_n} W_{n,j}, e_l)^2] = 0.$$

Then $\Sigma_{1 \le j \le a_n} W_{n,j} \Rightarrow \mathcal{N}(0,S)$ *in* $\mathbb{H}$, *where* $S \in \mathcal{S}(\mathbb{H})$ *is such that* $(Sh,h) = \sigma^2(h)$ *for each* $0 \ne h \in \mathbb{H}$.

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

Condition 3.1(a) is easily satisfied by applying an appropriate real-valued central limit theorem. Conditions 3.1(b) and (c) together imply the tightness of the $\mathbb{H}$-r.e. sequence $\{\Sigma_{1\le j\le a_n} W_{n,j} : n \in \mathbb{N}\}$.

Although Vakhania's (1981) approach can also give central limit theorems for many separable reflexive Banach spaces, it is not very easy to derive FCLT's with it. On the other hand, Billingsley's (1968) "two step" approach easily delivers CLT's and FCLT's for general complete separable metric spaces. Our next $\mathbb{H}$-CLT result follows from Theorem 1.4.2 of Billingsley (1968). For this, we let $\{e_j, j \ge 1\}$ be an arbitrary complete orthonormal basis of $\mathbb{H}$ and let $\mathbb{H}_k$ be the closed linear span of $[e_j, 1 \le j \le k]$. Let $P_k : \mathbb{H} \to \mathbb{H}_k$ be the orthonormal projection operator. We also define $\mathcal{N}(0, S^k)$ as an $\mathbb{H}_k$-r.e. with a Gaussian distribution on $\mathbb{H}_k$ of mean zero and covariance S^k, where $S^k \in \mathcal{S}(\mathbb{H}_k)$ and $\mathcal{S}(\mathbb{H}_k)$ is isometric and isomorphic to the set of all symmetric, positive semidefinite $k \times k$-matrices. (We can think of S^k as an $\infty \times \infty$ matrix with zeros everywhere except in the $k \times k$ upper left corner.) For any $S \in \mathcal{S}(\mathbb{H})$, we have $P_k S P_k \in \mathcal{S}(\mathbb{H}_k)$ and $\lim_{k\to\infty} \|P_k S P_k - S\|_{\mathrm{tr}} = 0$.

LEMMA 3.2. *Let* $\{W_{n,i} : 1 \le i \le a_n; n \ge 1\}$ *be a double array of* $\mathbb{H}$-r.e.*'s with zero means and finite second moments. Suppose the following conditions hold*:

(a) *For each* $k \ge 1$, $\Sigma_{1\le j\le a_n} P_k W_{n,j} \Rightarrow \mathcal{N}(0, S^k)$ *in* $\mathbb{H}_k$ (*as* $n \to \infty$);
(b) *There exists* $S \in \mathcal{S}(\mathbb{H})$, $S \ne 0$ *such that* $\lim_{k\to\infty} \|S^k - S\|_{\mathrm{tr}} = 0$;
(c) $\limsup_n [E|\Sigma_{1\le j\le a_n}(W_{n,j} - P_k W_{n,j})|^2] \to 0$ *as* $k \to \infty$.

Then $\Sigma_{1\le j\le a_n} W_{n,j} \Rightarrow \mathcal{N}(0, S)$ *in* $\mathbb{H}$ (*as* $n \to \infty$).

Condition 3.2(a) is especially straightforward to verify in our applications, because by the isometry and isomorphism between $\mathbb{H}_k$ and $\mathbb{R}^k$ it suffices to apply a CLT in $\mathbb{R}^k$ for each k. By Corollary 2.6 in Chen and White (1992a), condition 3.2(b) is a characterization of weak convergence for a sequence of zero mean Gaussian measures in $\mathbb{H}$. Condition 3.2(c) is a simple sufficient condition for the tightness of the array $\{\Sigma_{1\le j\le a_n} W_{n,j} : n \ge 1\}$.

The following remark gives some sufficient conditions for the tightness conditions in Lemmas 3.1 and 3.2. Additional sufficient conditions can be found in Chen and White (1992a).

Remark 3.3. Let $S_n \equiv \mathrm{Cov}(\Sigma_{1\le j\le a_n} W_{n,j})$. The tightness conditions 3.1(b) and (c) in Lemma 3.1 and 3.2(c) in Lemma 3.2 can, respectively, be replaced by any one of the following:

(i) $\{S_n\}$ is a Cauchy sequence in $(\mathcal{S}(\mathbb{H}), \|\cdot\|_{\mathrm{tr}})$.
(ii) For strictly stationary Hilbert-valued arrays, $\lim_{n\to\infty} E[|\Sigma_{1\le j\le a_n} W_{n,j}|^2]$ exists and is finite.
(iii) $\sup_n \mathrm{tr}(S_n) < \infty$, and $\sup_n \Sigma_{k\le l<\infty}(S_n e_l, e_l) \to 0$ as $k \to \infty$ for a complete orthonormal basis $\{e_l\}$.
(iv) Suppose $S_n, S \in \mathcal{S}(\mathbb{H})$ such that $(S_n h, h) \to (Sh, h)$ for any $h \in \mathbb{H}$ as $n \to \infty$. Then (iii) is equivalent to $\mathrm{tr}(S_n) \to \mathrm{tr}(S)$ as $n \to \infty$.

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

3.2. CLT's for $\mathbb{H}$-Valued Near Epoch Dependent Arrays

We now make use of an extension of the mixingale concept of McLeish (1975). A large class of Hilbert-valued dependent heterogeneous arrays satisfies the following notion of $\mathbb{H}$-valued mixingales, including $\mathbb{H}$-valued arrays near epoch dependent on a mixing process. (For brevity, we refer the reader to Chen and White, 1996, for further background and details.)

DEFINITION 3.4. *Let $\{W_{n,i}: i = \ldots,-1,0,1,\ldots; n \geq 1\}$ be an array of sub σ-fields of $\mathcal{F}$ nondecreasing in i for each n. Then $\{W_{n,i}, \mathcal{F}^{n,i}\}$ is called an* ***$L_p(\mathbb{H})$-mixingale*** *array if there exist finite nonnegative constants $\{c_{n,i}: i \geq 1; n \geq 1\}$ and $\{\psi_m; m \geq 0\}$ with $\psi_m \to 0$ as $m \to \infty$ such that for all $i, n \geq 1$, $m \geq 0$:*

$$\|E(W_{n,i}|\mathcal{F}^{n,i-m})\|_p \leq \psi_m c_{n,i};$$
$$\|W_{n,i} - E(W_{n,i}|\mathcal{F}^{n,i+m})\|_p \leq \psi_{m+1} c_{n,i}.$$

If in addition, for each $n \geq 1$, $W_{n,i}$ is $\mathcal{F}^{n,i}$-measurable, then $\{W_{n,i}, \mathcal{F}^{n,i}\}$ is an ***adapted*** *$L_p(\mathbb{H})$-mixingale array, and the second inequality holds automatically.*

The array $\{W_{n,i}, \mathcal{F}^{n,i}\}$ is an $L_p(\mathbb{H})$-mixingale ***of size a*** *if $\sum_{m=0}^{\infty}[\psi_m]^{\delta} < \infty$ for some δ with $0 < a \leq (1/\delta)$, or equivalently, $\psi_m = O(m^{-\lambda})$ for some λ with $\lambda > a > 0$.*

DEFINITION 3.5. *Let $\{V_{n,i}: i = \ldots,-1,0,1,\ldots; n \geq 1\}$ be an array of Banach-valued random elements ($\mathbb{B}$-r.e.'s) defined on the probability space $(\Omega, \mathcal{F}, P)$, and define $\mathcal{A}_{n,a}^{n,b} \equiv \sigma(V_{n,j}, a \leq j \leq b)$ for all $a \leq b$, $-\infty \leq a, b \leq \infty$, and $n \geq 1$.*

(i) *The array $\{V_{n,i}\}$ is called a* ***strong-*** *or* ***α-mixing*** *array of $\mathbb{B}$-r.e.'s if $\lim_{m\to\infty}\alpha(m) = 0$, where $\alpha(m) \equiv \sup_{n\geq 1}\sup_i \sup[|P(AB) - P(A)P(B)|: A \in \mathcal{A}_{n,-\infty}^{n,i}, B \in \mathcal{A}_{n,i+m}^{n,\infty}]$.*

(ii) *The array $\{V_{n,i}\}$ is called a* ***uniform-*** *or* ***ϕ-mixing*** *array of $\mathbb{B}$-r.e.'s if $\lim_{m\to\infty}\phi(m) = 0$, where $\phi(m) \equiv \sup_{n\geq 1}\sup_i \sup[|\mathrm{P(B|A)} - \mathrm{P(B)}|: \mathrm{A} \in \mathcal{A}_{n,-\infty}^{n,i}, P(A) > 0, B \in \mathcal{A}_{n,i+m}^{n,\infty}]$.*

(iii) *An $\mathbb{H}$-r.e. array $\{W_{n,i}: i \geq 1; n \geq 1\}$ is called an $L_p(\mathbb{H})$-array* ***near epoch dependent*** *($L_p(\mathbb{H})$-**NED**) on $\{V_{n,i}\}$ if $\|W_{n,i}\|_p < \infty$ for all $i, n \geq 1$, and there exist constants $\{d_{n,i} \geq 0: i \geq 1; n \geq 1\}$ and $\{\mu_m \geq 0: m \geq 0\}$ with μ_m decreasing to zero as $m \to \infty$ such that*

$$\|W_{n,i} - E[W_{n,i}|\mathcal{A}_{n,i-m}^{n,i+m}]\|_p \leq \mu_m d_{n,i}.$$

Remark 3.6. Let $\{W_{n,i}\}$ be an $L_p(\mathbb{H})$-($p \geq 2$) mixingale with $\{\psi_m\}$ and $\{c_{n,i}\}$. Then for each $h \in \mathbb{H}$, $\{(W_{n,i}, h)\}$ is an $L_p(\mathbb{R})$-mixingale with $\{\psi_m\}$ and $\{c_{n,i,h}\}$. Furthermore, if there exists a complete orthonormal basis $\{e_l\}$ in $\mathbb{H}$ such that $\sum_{l=1}^{\infty}(c_{n,i,e_l})^2 < \infty$ for each fixed n, i, then one can choose $(c_{n,i})^2 = \sum_{l=1}^{\infty}(c_{n,i,e_l})^2$. The result holds true when $\{W_{n,i}\}$ is an $L_p(\mathbb{H})$-NED array, with $(d_{n,i})^2 = \sum_{l=1}^{\infty}(d_{n,i,e_l})^2 < \infty$ for each fixed n, i.

The following maximal inequality is an extension of Chen and White's (1996) Theorem 3.7 and can be proved in exactly the same way.

LEMMA 3.7. *Let $\{W_{n,i}, \mathcal{F}^{n,i}\}$ be an $L_p(\mathbb{H})$-mixingale array with $p \geq 2$ and parameters $\{\psi_m\}$ and $\{c_{n,i}\}$. If ψ_m is of size $-\frac{1}{2}$, then there exists a constant $0 <$*

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

$K < \infty$ such that for all $J \geq 0$, all $k, n \geq 1$,

$$E\left[\max_{j \leq k} |\Sigma_{J+1 \leq i \leq J+j} W_{n,i}|^2\right] \leq K \Sigma_{J+1 \leq i \leq J+k} c_{n,i}^2.$$

The next lemma describes the asymptotic behavior of covariance operators for a partial sum of a weakly stationary L_2-NED sequence of $\mathbb{H}$-r.e.'s.

LEMMA 3.8. *Let $\{W_n\}$ be a weakly stationary $\mathbb{H}$-r.e. sequence with zero means and uniformly finite L_r-norm, $r \geq 2$. Let $\{V_n\}$ be a $\mathbb{B}$-r.e. mixing sequence with $\alpha(m)$ of size $-2ar/(r-2)$, $(r > 2)$, or $\phi(m)$ of size $-ar/(r-1)$, $(r \geq 2)$. Suppose that $\{W_n\}$ is L_2-NED on $\{V_n\}$ with μ_m of size $-a$. Let S_n be the covariance operator of $n^{-1/2} \sum_{j=1}^n W_j$. Suppose there exists $\varepsilon > 0$ such that $1 < a(1+\varepsilon) < 2$.*

Then there exists $S \in \mathcal{S}(\mathbb{H})$ such that $\lim_{n \to \infty} \|S_n - S\|_{\mathrm{tr}} = 0$. Moreover,

$$\|S_n - S\|_{\mathrm{tr}} = O(n^{-a(1+\varepsilon)+1}) \|W_1\|_2,$$

where S is given by

$$(Sh, g) = E[(h, W_1)(g, W_1)] + \sum_{j=2}^{\infty} E[(h, W_1)(g, W_j)] + \sum_{j=2}^{\infty} E[(h, W_j)(h, W_1)],$$

for all $h, g \in \mathbb{H}$.

Note that $1 \leq a < 2$ will give us $1 < a(1+\varepsilon) < 2$ for some $0 < \varepsilon < 1$ in the preceding lemma.

With this machinery in place we can now state CLT's for $\mathbb{H}$-valued NED arrays as simple consequences of any CLT for finite-dimensional NED arrays (e.g., the results of Wooldridge and White, 1988, 1989) and our Lemmas 3.1, 3.2, 3.7, and 3.8. We formally present two such results. Theorem 3.9 gives a CLT for $\mathbb{H}$-valued NED double arrays that might have trending moments, thus permitting certain nonstationary processes. Corollary 3.10 gives a CLT for a weakly stationary $\mathbb{H}$-valued NED sequence.

An essential ingredient in the central limit theory of dependent random sequences is the asymptotic behavior of the covariance sequence $\{S_n\}$ of the partial sum process $\{\Sigma_{1 \leq j \leq a_n} W_{n,j}\}$. Many available CLT results directly assume that $\{S_n\}$ converges in $(\mathcal{S}(\mathbb{H}), \|\cdot\|_{\mathrm{tr}})$ to a limit $S \neq 0$ when the dependent random process is not necessarily (weakly) stationary. On the other hand, some CLT's impose more restrictive assumptions on the second or higher moments of $\{W_n\}$ and more detailed assumptions on the decay of the dependence when the dependent random process is (weakly) stationary, such as for various mixing sequences. In line with these trade-offs, Theorem 3.9 directly imposes the existence of $S \in \mathcal{S}(\mathbb{H})$, whereas Corollary 3.10 does not.

The following result is implied by Lemmas 3.1 and 3.7 and is a generalization of Wooldridge and White's (1989) Theorem 2.9 to infinite-dimensional Hilbert space.

THEOREM 3.9. *Let $\{W_{n,i} : i \geq 1; n \geq 1\}$ be a double array of $\mathbb{H}$-r.e.'s with zero means and $\|W_{n,i}\|_r < \infty$ for some $r \geq 2$. Suppose the following conditions hold*:

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

(a)
(i) $\{W_{n,i}\}$ *is* $L_2(\mathbb{H})$-NED *on* $\{V_{n,i}\}$ *with size* -1 *and constants* $\{d_{n,i}\}$;
(ii) $\{V_{n,i}\}$ *is a mixing* $\mathbb{B}$-r.e. *array with either* $\alpha(m)$ *of size* $-2r/(r-2)$, $r>2$, *or* $\phi(m)$ *of size* $-r/(r-1)$;
(iii) $\max_{1\le i\le n}(c_{n,i})^2 = O(n^{-1})$, *where* $c_{n,i} \equiv \max(d_{n,i}, \|W_{n,i}\|_r)$ *for all* n, i;
(b) *If* $r=2$, $\{|W_{n,i}/c_{n,i}|^2 : i \ge 1; n \ge 1\}$ *is uniformly integrable*.
(c) *There exists a nonsingular* $S \in \mathcal{S}(\mathbb{H})$ *such that* $\lim_{n\to\infty}\|\mathrm{Cov}(\Sigma_{1\le i\le n}W_{n,i}) - S\|_{\mathrm{tr}} = 0$.

Then $\Sigma_{1\le i\le n}W_{n,i} \Rightarrow \mathcal{N}(0,S)$ *in* $\mathbb{H}$ (*as* $n\to\infty$).

The following result is a direct consequence of Theorem 3.9 and Lemma 3.8.

COROLLARY 3.10. *Let* $W_{n,i} = n^{-1/2}W_i$ *for all* $n \ge 1$, $i = 1,\dots,n$. *Suppose* $\{W_i\}$ *satisfies all the assumptions in Lemma* 3.8 *with* $r>2$ *and of size* $a=1$. *Then* $n^{-1/2}\Sigma_{1\le i\le n}W_i \Rightarrow \mathcal{N}(0,S)$ *in* $\mathbb{H}$ (*as* $n\to\infty$), *where* S *is that given by Lemma* 3.8.

This corollary extends and implies Dehling's (1983) CLT's for $\mathbb{H}$-valued strictly stationary uniform mixing and strong mixing processes.

4. FCLT'S FOR $\mathbb{H}$-VALUED DEPENDENT HETEROGENEOUS PROCESSES

4.1. Background

Let $C_{\mathbb{H}}[0,1]$ be the set of all continuous $\mathbb{H}$-valued functions on $[0,1]$. This becomes a separable Banach space under the sup-norm $\|X\|_\infty = \sup[|X(t)| : 0 \le t \le 1]$. Similarly, for each $k \in \mathbb{N}$ we get a separable Banach space $C_{\mathbb{H}_k}[0,1]$ under the sup-norm. When $\mathbb{B} = C_{\mathbb{H}}[0,1]$ or $C_{\mathbb{H}_k}[0,1]$, we call a $\mathbb{B}$-r.e. Z a $C_{\mathbb{H}}[0,1]$-valued random element ($C_{\mathbb{H}}[0,1]$-r.e.) or a $C_{\mathbb{H}_k}[0,1]$-valued random element ($C_{\mathbb{H}_k}[0,1]$-r.e.). *A* $C_{\mathbb{H}}[0,1]$-r.e. X is a *Brownian motion* in $\mathbb{H}$ if the following conditions are satisfied:

(a) $X(0) = 0$;
(b) the increments on disjoint time intervals are independent;
(c) for all $0 \le t < t+s \le 1$, the increment $X(t+s) - X(t)$ has a Gaussian distribution on $\mathbb{H}$ with mean zero and covariance operator sS, where $S \in \mathcal{S}(\mathbb{H})$ does not depend on t, s. Or equivalently,
(c′) for all $0 \le t < t+s \le 1$ and all $h \in \mathbb{H}$, the real-valued random variable $(X(t+s) - X(t), h)$ has a Gaussian distribution on $\mathbb{R}$ with mean zero and covariance $s(Sh, h)$, where $S \in \mathcal{S}(\mathbb{H})$ does not depend on t, s.

A $C_{\mathbb{H}}[0,1]$-r.e. X is a Brownian motion in $\mathbb{H}$ with covariance $S \in \mathcal{S}(\mathbb{H})$ if and only if for every $h \in \mathbb{H}$, $\{(X(t), h) : t \in [0,1]\}$ is a Brownian motion in $\mathbb{R}$ with covariance (Sh, h). In the preceding definition, if we substitute $\mathbb{H}_k$ for $\mathbb{H}$, we get the definition of a Brownian motion in $\mathbb{H}_k$. A functional central limit theorem in $\mathbb{H}$ ($\mathbb{H}$-FCLT) means that a sequence of $C_{\mathbb{H}}[0,1]$-r.e.'s converges in distribution to a Brownian motion in $\mathbb{H}$. We obtain the following lemma by applying Theorem 1.4.2 of Billingsley (1968).

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

LEMMA 4.1. *Let* $\{X_n : n \geq 1\}$ *be a sequence of* $C_{\mathbb{H}}[0,1]$-r.e.*'s. Let* X^k *be a Brownian motion in* $\mathbb{H}_k$ *with* $X^k(0) = 0$, $EX^k(1) = 0$, *and* $\operatorname{Cov} X^k(1) = S^k$. *Suppose the following conditions hold*:

(a) *For each* $k \geq 1$, $P_k X_n \Rightarrow X^k$ *in* $C_{\mathbb{H}_k}[0,1]$ (*as* $n \to \infty$);
(b) $X^k \Rightarrow X$ *in* $C_{\mathbb{H}}[0,1]$ (*as* $k \to \infty$);
(c) $\limsup_n E(\sup_{t\in[0,1]}[|X_n(t) - P_k X_n(t)|^2]) \to 0$ *as* $k \to \infty$.

Then $X_n \Rightarrow X$ *in* $C_{\mathbb{H}}[0,1]$ (*as* $n \to \infty$), *where* X *is a Brownian motion in* $\mathbb{H}$ *with* $X(0) = 0$, $EX(1) = 0$, *and* $\operatorname{Cov} X(1) = S$.

4.2. FCLT's for Adapted Dependent Heterogeneous $\mathbb{H}$-Valued Processes

The usual way to establish an FCLT for real-valued adapted dependent heterogeneous arrays is via a real-valued martingale difference approximation. If the approximated martingale difference array satisfies a FCLT and if the approximation error is small, then we obtain an FCLT for the original adapted array. The same ideas work for $\mathbb{H}$-valued adapted arrays. In particular, we can well approximate $\mathbb{H}$-valued adapted mixingale arrays by $\mathbb{H}$-valued martingale difference arrays (see Chen and White, 1996). Hence we can obtain numerous useful FCLT's for $\mathbb{H}$-valued adapted mixingales based on different versions of FCLT's for $\mathbb{H}$-valued martingale difference arrays, which in turn can be obtained by applying Lemma 4.1 and various FCLT's for real-valued martingale difference arrays.

Throughout this and the next subsection, we always let $\{a_n : n \geq 1\}$ be a sequence of nondecreasing, right continuous functions $a_n : [0,1] \to \mathbb{N}$ such that $a_n(0) = 1$ and $a_n(t) - a_n(s) \to \infty$ as $n \to \infty$ for all $0 \leq s < t \leq 1$. We also denote $X_n(t) \equiv \Sigma_{1\leq i\leq a_n(t)} W_{n,i}$ if $a_n(t) > a_n(t-)$ and by linear interpolation otherwise.

THEOREM 4.2. *Let* $\{W_{n,i}, \mathcal{F}^{n,i}\}$ *be an adapted array of* $\mathbb{H}$-r.e.*'s with finite second moments. Suppose there exists a positive sequence* $\{C_{n,i}\}$ *such that*

(a) $\sup_{t\in[0,1]} |\Sigma_{1\leq i\leq a_n(t)}[W_{n,i} - Z_{n,i}]| \to 0$ *in Prob. as* $n \to \infty$, *where*

$$Z_{n,i} \equiv W_{n,i} 1(|W_{n,i}| \leq C_{n,i}) - E[W_{n,i} 1(|W_{n,i}| \leq C_{n,i}) | \mathcal{F}^{n,i-1}].$$

(b) $E[|Z_{n,i}|^2 | \mathcal{F}^{n,i-1}] < \infty$ *a.s.* $-P$;
(c) *For all* $h \in \mathbb{H}$ *and all* $\varepsilon > 0$,

$$\Sigma_{1\leq i\leq a_n(1)} E[(Z_{n,i}, h)^2 1(|(Z_{n,i}, h)| \geq \varepsilon) | \mathcal{F}^{n,i-1}] \to 0 \quad \textit{in Prob.} \quad \textit{as } n \to \infty.$$

In addition, suppose there exists $S \in \mathcal{S}(\mathbb{H})$ *such that*

(d) $\Sigma_{1\leq i\leq a_n(1)} E[|Z_{n,i}|^2 | \mathcal{F}^{n,i-1}] \to \operatorname{tr}(S)$ *in Prob. as* $n \to \infty$; *and*
(e) *For all* $h \in \mathbb{H}$, $h \neq 0$, *and* $t \in [0,1]$, $\Sigma_{1\leq i\leq a_n(t)} E[(Z_{n,i}, h)^2 | \mathcal{F}^{n,i-1}] \to t(Sh, h)$ *in Prob. as* $n \to \infty$.

Then $X_n \Rightarrow X$ *in* $C_{\mathbb{H}}[0,1]$ (*as* $n \to \infty$), *where* X *is a Brownian motion in* $\mathbb{H}$ *with* $X(0) = 0$, $EX(1) = 0$, *and* $\operatorname{Cov} X(1) = S$. *Moreover*, $\Sigma_{1\leq i\leq a_n(1)} W_{n,i} \Rightarrow \mathcal{N}(0, S)$ *in* $\mathbb{H}$ (*as* $n \to \infty$).

This theorem generalizes Theorem C of Jakubowski (1980) and Theorem 5.1 of Berger (1986).

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

Condition 4.2(c) is implied by

(c′) $\Sigma_{1\le i\le a_n(1)} E[|Z_{n,i}|^2 1(|Z_{n,i}|\ge\varepsilon)|\mathcal{F}^{n,i-1}]\to 0$ in Prob. as $n\to\infty$;

Condition 4.2(e) can be replaced by

(e′) $\Sigma_{1\le i\le a_n(t)} E[(Z_{n,i},e_l)(Z_{n,i},e_j)|\mathcal{F}^{n,i-1}]\to t(Se_l,e_j)$ in Prob. as $n\to\infty$,

for some orthonormal sequence $\{e_l\}$ in $\mathbb{H}$.

Condition 4.2(c) is the conditional Lindeberg assumption, and condition 4.2(e) is the conditional variance norming assumption. These two assumptions together imply that $(Y_n,h)\Rightarrow(X,h)$ for any fixed $h\in\mathbb{H}$, where $Y_n(t)\equiv\Sigma_{1\le i\le a_n(t)}Z_{n,i}$ if $a_n(t)>a_n(t-)$ and by linear interpolation otherwise. There are many other sufficient conditions for FCLT's for real-valued martingale difference arrays. See Gaenssler and Haeusler (1986) for a nice survey.

THEOREM 4.3. *Let* $\{W_{n,i}\}$ *and* $\{Z_{n,i}\}$ *be as in Theorem* 4.2. *Suppose the following conditions hold*:

(a) 4.2(a) *holds*;
(b) $\limsup_n \Sigma_{1\le i\le a_n(1)} E[|Z_{n,i}|^2|\mathcal{F}^{n,i-1}]\le C<\infty$ *a.s.* $-P$;
(c) *Let* $\{e_l: l\ge 1\}$ *be a complete orthonormal sequence in* $\mathbb{H}$, *with*

$$\lim_{k\to\infty}\limsup_{n\to\infty}\Sigma_{k\le l<\infty}\Sigma_{1\le i\le a_n(1)}E[(Z_{n,i},e_l)^2|\mathcal{F}^{n,i-1}]=0 \quad \text{in Prob.};$$

(d) $\{\max_{1\le i\le a_n(1)}|Z_{n,i}|: n\ge 1\}$ *is uniformly integrable*;
(e) *for all* $h\in\mathbb{H}$, $h\ne 0$, *and any* $t\in[0,1]$, $\Sigma_{1\le i\le a_n(t)}(Z_{n,i},h)^2\to t\sigma^2(h)$ *in Prob. as* $n\to\infty$, *where* $\sigma^2(h)$ *is a positive finite real number depending on* $h\in\mathbb{H}$.

Then (i) *there exists* $S\in\mathcal{S}(\mathbb{H})$ *such that* $(Sh,h)=\sigma^2(h)$ *for each* $h\in\mathbb{H}$, $h\ne 0$; *and* (ii) $X_n\Rightarrow X$ *in* $C_{\mathbb{H}}[0,1]$ (*as* $n\to\infty$), *where* X *is a Brownian motion in* $\mathbb{H}$ *with* $X(0)=0$, $EX(1)=0$, *and* $\operatorname{Cov}X(1)=S$.

Remark 4.4.

(i) Under Condition 4.2(c), condition 4.3(e) is equivalent to condition 4.4(e′).
(ii) The conclusion of Theorem 4.3 remains true if we replace conditions 4.3(b)–(e) by the following set of assumptions:
  (b′) $\lim_n[E|\Sigma_{1\le i\le a_n(1)}Z_{n,i}|^2]=C<\infty$;
  (c′) Let $\{e_l: l\ge 1\}$ be a complete orthonormal sequence in $\mathbb{H}$, with

$$\lim_{k\to\infty}\limsup_{n\to\infty}\Sigma_{k<l<\infty}E[(\Sigma_{1\le i\le a_n(1)}Z_{n,i},e_l)^2]=0;$$

  (d′) $E[\max_{1\le i\le a_n(1)}|Z_{n,i}|]\to 0$ as $n\to\infty$;
  (e′) for all $h\in\mathbb{H}$, $h\ne 0$, and $t\in[0,1]$, $\Sigma_{1\le i\le a_n(t)}E[(Z_{n,i},h)^2|\mathcal{F}^{n,i-1}]\to t\sigma^2(h)$ in Prob. as $n\to\infty$, where $\sigma^2(h)$ is a positive finite real number depending on $h\in\mathbb{H}$.

COROLLARY 4.5. *Define* $Z_{n,i}\equiv W_{n,i}1(|W_{n,i}|\le c)-E[W_{n,i}1(|W_{n,i}|\le c)|\mathcal{F}^{n,i-1}]$ *for some positive constant* c, *with* $\{W_{n,i}\}$ *the same as in Theorem* 4.2. *If conditions* 4.1(a), 4.3(c), *and* 4.3(e) *hold*, *then all the conclusions in Theorem* 4.3 *remain true*.

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

4.3. FCLT's for $\mathbb{H}$-Valued NED Processes

Although one can always obtain various FCLT's for $\mathbb{H}$-valued NED processes by the theorems in the preceding subsection 4.2, we can also establish such results by simply applying Lemma 4.1 and various FCLT's for real-valued NED processes. As an illustration, we give the following result, which is a direct generalization of Wooldridge and White's (1988) Theorem 2.11 to an infinite-dimensional Hilbert space.

THEOREM 4.6. *Let* $\{W_{n,i}: i \geq 1;\ n \geq 1\}$ *be a double array of* $\mathbb{H}$-r.e.*'s with zero means,* $\|W_{n,i}\|_r < \infty$ *for some* $r \geq 2$. *Suppose the following conditions hold*:

(a)
 (i) $\{W_{n,i}\}$ *is* $L_2(\mathbb{H})$-NED *on* $\{V_{n,i}\}$ *with* μ_m *of size* $-\gamma$, $\frac{1}{2} \leq \gamma \leq 1$, *and* $\max_{1\leq i\leq a_n(1)} d_{n,i} = O([a_n(1)]^{\gamma-1})$;
 (ii) $\{V_{n,i}\}$ *is a mixing* $\mathbb{B}$-r.e. *array with either* $\alpha(m)$ *of size* $-r/(r-2)$, $r > 2$, *or* $\phi(m)$ *of size* $-r/2(r+1)$;
 (iii) $\sup_{k,J} \limsup_{n\to\infty} k^{-1}\Sigma_{J+1\leq i\leq J+k}(c_{n,i})^2 < \infty$, *where* $c_{n,i} \equiv \max(d_{n,i}, \|W_{n,i}\|_r)$ *for all* n, i;
(b) *when* $r = 2$, $\{|W_{n,i}/c_{n,i}|^2 : i \geq 1;\ n \geq 1\}$ *is uniformly integrable for each* $0 \neq h \in \mathbb{H}$.
(c) *for all* $t \in [0,1]$, *each* $0 \neq h \in \mathbb{H}$, $\lim_{n\to\infty} E[\Sigma_{1\leq i\leq a_n(t)}(W_{n,i}, h)]^2 = t\sigma^2(h)$.
(d) $\limsup_{n\to\infty} E[|\Sigma_{1\leq i\leq a_n(1)} W_{n,i}|^2] = C < \infty$.
(e) *There exists a complete orthonormal basis* $\{e_l\}$ *such that*

$$\limsup_{n\to\infty} \Sigma_{l>k}\Sigma_{1\leq i\leq a_n(1)}(c_{n,i,e_l})^2 \to 0, \quad \text{as } k \to \infty.$$

Then (i) *there exists nonsingular* $S \in \mathcal{S}(\mathbb{H})$ *such that* $(Sh, h) = \sigma^2(h)$ *for each* $h \in \mathbb{H}$, $h \neq 0$; *and* (ii) $X_n \Rightarrow X$ *in* $C_{\mathbb{H}}[0,1]$ *(as* $n \to \infty$*), where* X *is a Brownian motion in* $\mathbb{H}$ *with* $X(0) = 0$, $EX(1) = 0$, *and* $\operatorname{Cov} X(1) = S$. *Hence,* (iii) $\Sigma_{1\leq i\leq a_n(1)} W_{n,i} \Rightarrow \mathcal{N}(0, S)$ *in* $\mathbb{H}$ *(as* $n \to \infty$*).*

Remark 4.7. Theorem 3.9 still holds if one replaces condition 3.9(c) by conditions 4.6(d) and (e), and for each $0 \neq h \in \mathbb{H}$, $\lim_{h\to\infty} E[\Sigma_{1\leq i\leq a_n(1)}(W_{n,i}, h)]^2 = \sigma^2(h)$.

5. APPLICATIONS

We now provide explicit conditions delivering convergence results for each of the examples of Section 2. We do not give the weakest possible conditions for the sake of simplicity and brevity.

PROPOSITION 5.1. *In Example* 2.1, *suppose* $\{X_t\}$ *has a common c.d.f.* F *in* $\mathbb{R}^d$ *and is a weakly stationary sequence* NED *of size* -1 *on some strong mixing sequence of size* $-2r/(r-2)$ *for* $r > 2$. *Then*

(i) (*the kernel c.d.f. estimator*): *In addition, suppose* $\{X_t\}$ *has a common marginal density* f *in* $\mathbb{R}^d$, *which is square-integrable and is continuously differentiable to*

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

order p inclusively and $f^{(p)}$ is square-integrable. Suppose $K \geq 0$ is symmetric and Lipschitz-continuous, with $\int K(u)\,du = 1$, $\int u_1^{p_1} \ldots u_d^{p_d} K(u)\,du = 0$ for $1 \leq \sum_{i=1}^d p_i \leq p-1$, and $\int u_1^{p_1} \ldots u_d^{p_d} K(u)\,du = C$ for $\sum_{i=1}^d p_i = p$, $0 < C < \infty$, $p > d$. Further, let $b_n \to 0$, $n(b_n)^p \to 0$, and $n(b_n)^d \to \infty$ as $n \to \infty$. Then $n^{1/2}(\hat{F}_{k,n}(\cdot) - F(\cdot)) \Rightarrow \mathcal{N}(0,S)$ in $L_2(\mathbb{R}^d)$, where for any $h, g \in L_2(\mathbb{R}^d)$,

$$(Sh,g) = \iint [F(\min(x,y)) - F(x)F(y)]h(x)g(y)\,dx\,dy$$

$$+ 2\sum_{j=2}^{\infty} \iint [P(X_1 \leq x, X_j \leq y) - F(x)F(y)]h(x)g(y)\,dx\,dy.$$

(ii) (*the empirical c.d.f. estimator*): *In addition to conditions for* (i), *we assume that f has a bounded open support $D \subset \mathbb{R}^d$ with smooth boundary, and for any $\varepsilon > 0$,*

$$n^{1/2} \sup_{x \in D} \int_{|y| > \varepsilon n^{-1/2}} |F(x-y) - F(x)| K(y/b_n) b_n^{-d}\,dy \to 0.$$

Then $n^{1/2}(\hat{F}_{e,n}(\cdot) - F(\cdot)) \Rightarrow \mathcal{N}(0,S)$ in $L_2(\mathbb{R}^d)$, where S is the same as in (i).

By the continuous mapping theorem, we immediately have the following result:

Let $T(F) = \int g(x)\,dF(x)$, where $g \in L_2(\mathbb{R}^d)$ such that $0 < (Sg,g) < \infty$. Then $n^{1/2}(T(\hat{F}_n) - T(F)) \Rightarrow \mathcal{N}(0,(Sg,g))$.

The result for the orthogonal series estimator can be proven similarly to that for the kernel c.d.f. estimator. Results for other smooth functionals of F follow easily using standard arguments. We omit these for brevity. Notice that we need additional conditions in (ii) because $\{\mathbf{1}(X_t \leq \cdot) - F(\cdot)\}$ is not necessarily a Hilbert-valued NED sequence when $\{X_t\}$ is an $\mathbb{R}^d$-valued NED sequence.

PROPOSITION 5.2. *In Example* 2.2, *suppose X_t has common distribution F with compact support on $\mathcal{X}$, a measurable space for all t. Assume $\{X_t\}$ is a strictly stationary sequence* NED *of size -1 on some strong mixing sequence of size $-2r/(r-2)$ for $r > 2$. Let $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be a symmetric function such that $\sup_{i,j} E[K^{2+\delta}(X_i, X_j)] < \infty$, $\iint K^{2+\delta}(x,y)\,dF(x)\,dF(y) < \infty$ for some $\delta > 0$, and $\int K(x,y)\,dF(x) = 0$ for almost all $y \in \mathcal{X}$. Suppose $\sum_{j=1}^{\infty} |\lambda_j| < \infty$, where $\{\lambda_j\}$ is the sequence of nonzero eigenvalues for K. Then $nU_n \Rightarrow \sum_{j=1}^{\infty} \lambda_j (Z_j^2 - 1)$ as $n \to \infty$, where $\{Z_j\}$ is a sequence of correlated real-valued normals $\mathcal{N}(0,\sigma_j^2)$, with $\sigma_j^2 = 1 + 2\sum_{i=2}^{\infty} E[g_j(X_1)g_j(X_i)]$ and*

$$\mathrm{Cov}(Z_l, Z_j) = \sum_{i=2}^{\infty} (E[g_l(X_1)g_j(X_i)] + E[g_l(X_i)g_j(X_1)]) \quad \text{for } l \neq j,$$

where $\{g_j\}$ is the sequence of nonconstant eigenfunctions for K (corresponding to $\{\lambda_j\}$).

Note that $\{Z_j\}$ becomes a sequence of independent normals $\mathcal{N}(0,1)$ if and only if $\{X_t\}$ is independent. (See the proof in the Appendix.)

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

PROPOSITION 5.3. *In Example* 2.3, *let* $a_n = 1/n$, $\overline{M}(\theta^*) = 0$, *and for some* $\beta \in (0,1)$,

$$\nu_n \equiv n^{-(1-\beta)/2}[M_n(Z_n,\theta^*) - EM_n(Z_n,\theta^*)].$$

Suppose the following conditions hold:

(a) $|\hat{\theta}_n - \theta^*| \to 0$ *a.s.* $-P$;
(b) $\overline{M}(\cdot)$ *is twice continuously Frechet differentiable at* θ^* *with first derivative* A;
(c) *There exists some* $\beta \in (0,1)$ *such that*
  (i) $\sigma^* \equiv \inf[\mathrm{Re}\,\gamma : \gamma \in \mathrm{spec}(-A)] > \beta/2$, *where* $\mathrm{spec}(-A)$ *is the spectrum of* $-A$, (*i.e.*, *for any* $u \in (0,\infty)$, $\|\exp(uA)\| < \exp(-u\sigma^*)$ *with* $\sigma^* > \beta/2$);
  (ii) $|n^{\beta/2}E[M_n(Z_n,\theta^*)] - T| \to 0$ *as* $n \to \infty$ *for some* $T \in \mathbb{H}$;
(d) *For all small* $\varepsilon > 0$ *and all* n, *there exists a measurable mapping* $D_{\varepsilon,n}$ *such that*
  (i) *for all* θ *with* $|\theta - \theta^*| \le \varepsilon$,

$$|M_n(Z_n,\theta) - M_n(Z_n,\theta^*) - D_{\epsilon,n}(Z_n,\theta^*)(\theta - \theta^*)| \le \varepsilon^{\delta},$$

  for some constant $\delta \in (1,2/\beta)$.
  (ii) $|D_{\varepsilon,n}(Z_n,\theta^*) - A| \to 0$ *a.s. as* $n \to \infty$;
(e) *The sequence* $\{\nu_n\}$ *has uniform finite rth moment* $(r > 2)$ *and is* NED *of size* -1 *on a underlying strong mixing sequence of size* $-2r/(r-2)$.

Denote $Q_n(t) \equiv n^{-1/2}\sum_{j=1}^{[nt]} \nu_j + n^{-1/2}(nt - [nt])\nu_{[nt]+1}$.

(f) *There exists a nonsingular* $S \in \mathcal{S}(\mathbb{H})$ *such that for all* $t \in [0,1]$ *and each* $h \in \mathbb{H}$, $h \ne 0$,

$$\lim_{n\to\infty} E[(Q_n(t),h)^2] = t(Sh,h).$$

Then

(i) $Q_n \Rightarrow X$ *in* $C_{\mathbb{H}}[0,1]$ *as* $n \to \infty$, *where* X *is a Brownian motion in* $\mathbb{H}$ *with* $X(0) = 0$ *and* $\mathrm{Cov}\,X(1) = S \in \mathcal{S}(\mathbb{H})$, *where*

$$(Sh,h) = \lim_n n^{-1}\sum_{j=1}^{n} E[(\nu_j,h)^2] + 2\lim_n n^{-1}\sum_{j=2}^{n}\sum_{m=1}^{j-1} E[(\nu_j,h)(\nu_{j-m},h)]$$

for any $0 \ne h \in \mathbb{H}$.

(ii) $n^{\beta/2}(\hat{\theta}_n - \theta^*) \Rightarrow \mathcal{N}(-\bar{A}^{-1}T,K)$ *as* $n \to \infty$ *in* $\mathbb{H}$, *where* K *is the unique solution of the operator equation* $\bar{A}K + K\bar{A}^* = -S$ *in* $L(\mathbb{H},\mathbb{H})$, *where* $\bar{A} \equiv (\beta/2)I + A$, *with the transpose* $\bar{A}^*$.

Chen and White (1992b) gave sufficient conditions for almost-sure convergence, rate of convergence, and asymptotic normality for $\hat{\theta}_n - \theta^*$ under a less restrictive set of sufficient conditions such as Lipschitz continuity of M_n and NED data $\{Z_n\}$. Chen and White (1993) further proved almost-sure convergence when $\{Z_n\}$ may be Granger-caused by $\{\hat{\theta}_n\}$. Here we directly assume that $\{\nu_n\}$ is NED for the sake of brevity.

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

REFERENCES

Andrews, D.W.K. (1985) A nearly independent, but non-strong mixing triangular array. *Journal of Applied Probability* 22, 729–731.

Andrews, D.W.K. (1988) Laws of large numbers for dependent non-identically distributed random variables. *Econometric Theory* 4, 458–467.

Araujo, A. & E. Gine (1980) *The Central Limit Theorem for Real and Banach-Valued Random Variables.* New York: Wiley.

Berger, E. (1986) Asymptotic behavior of a class of stochastic approximation procedures. *Probability Theory and Related Fields* 71, 517–552.

Bierens, H.J. (1982) Consistent model specification tests. *Journal of Econometrics* 20, 105–134.

Bierens, H.J. & W. Plöberger (1997) Asymptotic theory of integrated conditional moment tests. *Econometrica* 65, 1129–1152.

Billingsley, P. (1968) *Convergence of Probability Measures.* New York: Wiley.

Carlstein, E. (1988) Degenerate *U*-statistics based on non-independent observations. *Calcutta Statistical Association Bulletin* 37, 55–65.

Chen, X. & H. White (1992a) Central Limit and Functional Central Limit Theorems for Hilbert Space-Valued Dependent Processes. Department of Economics Discussion paper 92-35, University of California at San Diego.

Chen, X. & H. White (1992b) Asymptotic Properties of Some Projection-Based Robbins-Monro Procedures in a Hilbert Space. Department of Economics Discussion paper 92-46, University of California at San Diego.

Chen, X. & H. White (1993) Nonparametric Adaptive Learning with Feedback. Department of Economics Working paper, University of Chicago.

Chen, X. & H. White (1996) Laws of large numbers for Hilbert space-valued mixingales with applications. *Econometric Theory* 12, 284–304.

Dehling, H. (1983) Limit theorems for sums of weakly dependent Banach space valued random variables. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 63, 393–432.

Eagleson, G.K. (1979) Orthogonal expansions and *U*-statistics. *Australian Journal of Statistics* 21, 221–237.

Fernholz, L.T. (1991) Almost sure convergence of smoothed empirical distribution functions. *Scandinavian Journal of Statistics* 18, 255–262.

Gaenssler, P. & E. Haeusler (1986) On martingale central limit theory. In E. Eberlein & M.S. Taqqu (eds.), *Dependence in Probability and Statistics*, pp. 303–333. Boston: Birkhauser.

Gallant, A.R. & H. White (1988) *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models.* New York: Basil Blackwell.

Jakubowski, A. (1980) On limit theorems for sums of dependent Hilbert space valued random variables. In *Lecture Notes in Statistics*, no. 2, pp. 178–187. New York: Springer-Verlag.

Ledoux, M. & M. Talagrand (1991) *Probability in Banach Spaces.* New York: Springer-Verlag.

Ljung, L., G. Pflug, & H. Walk (1992) *Stochastic Approximation and Optimization of Random Systems.* Berlin: Birkhauser Verlag.

Manski, C. (1988) *Analog Estimation Methods in Econometrics.* London: Chapman and Hall.

McLeish, D.L. (1975) A maximal inequality and dependent strong laws. *Annals of Probability* 3, 829–839.

McLeish, D.L. (1977) On the invariance principle for nonstationary mixingales. *Annals of Probability* 5, 616–621.

Parthasarathy, K.R. (1967) *Probability Measures on Metric Spaces.* New York: Academic Press.

Politis, D. & J. Romano (1994) Limit theorems for weakly dependent Hilbert space valued random variables with application to the stationary bootstrap. *Statistica Sinica* 4, 461–476.

Robbins, H. & S. Monro (1951) A stochastic approximation method. *Annals of Mathematical Statistics* 22, 400–407.

Serfling, R. (1980) *Approximation Theorems of Mathematical Statistics.* New York: Springer-Verlag.

Vakhania, N.N. (1981) *Probability Distributions on Linear Spaces.* New York: North-Holland.

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

Venter, J. (1966) On Dvoretzky stochastic approximation theorems. *Annals of Mathematical Statistics* 37, 1534–1544.
Walk, H. (1977) An invariance principle for the Robbins-Monro process in a Hilbert space. *Zeitschrift für Warscheinlichkeitstheorie und Verwandte Gebiete* 39, 135–150.
Wooldridge, J.M. & H. White (1988) Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory* 4, 210–230.
Wooldridge, J.M. & H. White (1989) Central Limit Theorems for Dependent, Heterogeneous Processes with Trending Moments. Department of Economics Discussion paper, University of California at San Diego.
Yukich, J.E. (1992) Weak convergence of smoothed empirical processes. *Scandinavian Journal of Statistics* 19, 271–279.

MATHEMATICAL APPENDIX

We provide proofs only for results that are not obvious from the text. The following lemma gives an estimate for the autocorrelations of an $\mathbb{H}$-valued NED array. It is an extension of equation (2) in Andrews (1988).

LEMMA A. *Let* $\{W_{n,i}: n, i = 0,\pm1,\pm2,\ldots\}$ *be an* $\mathbb{H}$-r.e. *double array with zero means and finite* L_r-norm $(r \geq 2)$, *and let* $\{V_{n,i}: n, i = 0,\pm1,\pm2,\ldots\}$ *be an* α- *or* ϕ-*mixing double array of* $\mathbb{B}$-r.e.'*s. Suppose that* $\{W_{n,i}\}$ *is* L_2-NED *on* $\{V_{n,i}\}$ *with constants* $\{\mu_m: m \geq 0\}$ *and* $\{d_{n,i}: n, i = 0,\pm1,\pm2,\ldots\}$. *Then for all* n *and all* $0 < a, b < m$ *with* $a + b \leq m$, *the following inequalities hold*:

(i) *when* $V_{n,i}$ *is* α-*mixing and* $r > 2$,

$$|E(W_{n,i}, W_{n,i-m})| \leq \mu_a d_{n,i} \|W_{n,i-m}\|_2 + 15^{1/2}[\alpha(b)]^{(1/2)-(1/r)} \|W_{n,i-m}\|_r \|W_{n,i}\|_2;$$

(ii) *when* $V_{n,i}$ *is* ϕ-*mixing and* $r \geq 2$,

$$|E(W_{n,i}, W_{n,i-m})| \leq \mu_a d_{n,i} \|W_{n,i-m}\|_2 + 2[\phi(b)]^{1-(1/r)} \|W_{n,i-m}\|_r \|W_{n,i}\|_2.$$

Let W, Z *be two* $\mathbb{H}$-r.e.'*s with zero means and finite* L_2-norm's. *Then we can define a cross-covariance operator* $T_{W,Z}$ *for* W *and* Z *as*

$$(T_{W,Z}h, g) \equiv E[(W, h)(Z, g)] \quad \textit{for any } h, g \in \mathbb{H}.$$

Let $\{e_l\}$ *be an arbitrary* CONS *for* $\mathbb{H}$. *Then we have*

$$E[(W, Z)] = \sum_{l=1}^{\infty} E[(W, e_l)(Z, e_l)] = \sum_{l=1}^{\infty} (T_{W,Z} e_l, e_l).$$

Another version of Lemma A is the following lemma.

LEMMA A′. *Let* $\{W_{n,i}\}$ *and* $\{V_{n,i}\}$ *satisfy all the assumptions in Lemma* A. *Then*

(i) *when* $V_{n,i}$ *is* α-*mixing and* $r > 2$,

$$\sum_{l=1}^{\infty} |E[(W_{n,i}, e_l)(W_{n,i-m}, e_l)] \leq \mu_a d_{n,i} \|W_{n,i-m}\|_2 + 15^{1/2}[\alpha(b)]^{(1/2)-(1/r)} \|W_{n,i-m}\|_r \|W_{n,i}\|_2;$$

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

(ii) *when $V_{n,i}$ is ϕ-mixing and $r \geq 2$,*

$$\sum_{l=1}^{\infty} |E[(W_{n,i}, e_l)(W_{n,i-m}, e_l)]| \leq \mu_a d_{n,i} \|W_{n,i-m}\|_2 + 2[\phi(b)]^{1-(1/r)} \|W_{n,i-m}\|_r \|W_{n,i}\|_2.$$

Proof of Lemma A. Because the elements of $\{W_{n,i}\}$ have zero means,

$$|E(W_{n,i}, W_{n,i-m})| = |E(W_{n,i}, W_{n,i-m} - EW_{n,i-m})| \leq a_1 + a_2,$$

$$a_1 \equiv |E(W_{n,i} - E[W_{n,i}|\mathcal{F}_{n,i-a}^{n,i+a}], W_{n,i-m} - EW_{n,i-m})|,$$

$$a_2 \equiv |E(E[W_{n,i}|\mathcal{F}_{n,i-a}^{n,i+a}], E[W_{n,i-m} - EW_{n,i-m}|\mathcal{F}_{n,i-a}^{n,i+a}])|.$$

By the Cauchy–Schwartz inequality,

$$a_1 \leq [E(|W_{n,i} - E(W_{n,i}|\mathcal{F}_{n,i-a}^{n,i+a})|^2)]^{1/2} [E(|W_{n,i-m} - EW_{n,i-m}|^2)]^{1/2} \leq \mu_a d_{n,i} \|W_{n,i-m}\|_2,$$

and

$$a_2 \leq [E(|E(W_{n,i}|\mathcal{F}_{n,i-a}^{n,i+a})|^2)]^{1/2} [E(|E(W_{n,i-m}|\mathcal{F}_{n,i-a}^{n,i+a}) - EW_{n,i-m}|^2)]^{1/2}.$$

By the conditional Jensen's inequality and Lemma 3.1 in Chen and White (1996),

$$a_2 \leq \|W_{n,i}\|_2 15^{1/2} [\alpha(b)]^{(1/2)-(1/r)} \|W_{n,i-m}\|_r \quad \text{for } V_{n,i} \quad \alpha\text{-mixing and} \quad r > 2;$$

or

$$a_2 \leq \|W_{n,i}\|_2 2[\phi(b)]^{1-(1/r)} \|W_{n,i-m}\|_r \quad \text{for } V_{n,i} \quad \phi\text{-mixing and} \quad r \geq 2.$$

The results follow. ∎

Proof of Lemma A′. The proof is similar to that for Lemma A. For any $h \in \mathbb{H}$,

$$|E[(W_{n,i}, h)(W_{n,i-m}, h)]| = |E[(W_{n,i}h)(W_{n,i-m} - EW_{n,i-m}, h)]| \leq a_1(h) + a_2(h),$$

$$a_1(h) \equiv |E[(W_{n,i} - E[W_{n,i}|\mathcal{F}_{n,i-a}^{n,i+a}], h)(W_{n,i-m} - EW_{n,i-m}, h)]|,$$

$$a_2(h) \equiv |E[(E[W_{n,i}|\mathcal{F}_{n,i-a}^{n,i+a}], h)E[(W_{n,i-m} - EW_{n,i-m}, h)|\mathcal{F}_{n,i-a}^{n,i+a}]]|.$$

By the Cauchy–Schwartz inequality,

$$a_1(h) \leq [E((W_{n,i} - E(W_{n,i}|\mathcal{F}_{n,i-a}^{n,i+a}), h)^2)]^{1/2} [E((W_{n,i-m} - EW_{n,i-m}, h)^2)]^{1/2},$$

and

$$a_2(h) \leq [E((E(W_{n,i}|\mathcal{F}_{n,i-a}^{n,i+a}), h)^2)]^{1/2} [E((E(W_{n,i-m}|\mathcal{F}_{n,i-a}^{n,i+a}) - EW_{n,i-m}, h)^2)]^{1/2}.$$

Let $\{e_l\}$ be any arbitrary CONS. Then,

$$\sum_{l=1}^{\infty} |E[(W_{n,i}, e_l)(W_{n,i-m}, e_l)]| \leq \sum_{l=1}^{\infty} a_1(e_l) + \sum_{l=1}^{\infty} a_2(e_l).$$

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

By the Cauchy–Schwartz inequality and Parseval's theorem,

$$\begin{aligned}\sum_{l=1}^{\infty} a_1(e_l) &\le \left[\sum_{l=1}^{\infty} E((W_{n,i} - E(W_{n,i}|\mathcal{F}_{n,i-a}^{n,i+a}), e_l)^2)\right]^{1/2} \\ &\quad \times \left[\sum_{l=1}^{\infty} E((W_{n,i-m} - EW_{n,i-m}, e_l)^2)\right]^{1/2} \\ &= [E(|W_{n,i} - E(W_{n,i}|\mathcal{F}_{n,i-a}^{n,i+a})|^2)]^{1/2}[E(|W_{n,i-m} - EW_{n,i-m}|^2)]^{1/2} \\ &\le \mu_a d_{n,i}\|W_{n,i-m}\|_2.\end{aligned}$$

Similarly, we get

$$\sum_{l=1}^{\infty} a_2(e_l) \le [E(|E(W_{n,i}|\mathcal{F}_{n,i-a}^{n,i+a})|^2)]^{1/2}[E(|E(W_{n,i-m}|\mathcal{F}_{n,i-a}^{n,i+a}) - EW_{n,i-m}|^2)]^{1/2}.$$

The rest of the proof follows that for Lemma A. ■

Proof of Lemma 3.8. First we show that $S \in \mathcal{S}(\mathbb{H})$, i.e., for each arbitrary CONS $\{e_l\}$, $\sum_{l=1}^{\infty} |(Se_l, e_l)| < \infty$. Because $\{W_n\}$ is weakly stationary, we have for any $h \in \mathbb{H}$,

$$|(Sh, h)| \le E[(W_1, h)^2] + 2\sum_{j=2}^{\infty} |E[(W_1, h)(W_j, h)]|.$$

Let $\{e_l\}$ be an arbitrary CONS; then

$$\sum_{l=1}^{\infty} |(Se_l, e_l)| \le \sum_{l=1}^{\infty} E[(W_1, e_l)^2] + 2\sum_{j=2}^{\infty}\sum_{l=1}^{\infty} |E[(W_1, e_l)(W_j, e_l)]| \equiv A_1 + 2A_2.$$

Because $A_1 = E[|W_1|^2] < \infty$, we only need to show that $A_2 < \infty$. Suppose $\{V_n\}$ is α-mixing. To apply Lemma A′(i), we set $n = j$, $m = j - 1$, and $d_n \equiv 1$. Let $a(j), b(j)$ be integers with $a(j) + b(j) \le m$. Then

$$A_2 \le \|W_1\|_2 \sum_{j=2}^{\infty} \mu_{a(j)} + \|W_1\|_r \sum_{j=2}^{\infty} 15^{1/2}[\alpha(b(j))]^{(1/2)-(1/r)}\|W_j\|_2.$$

By weak stationarity and $\sup_n \|W_n\|_r < \infty$ for $r > 2$, we have $\|W_n\|_2 = \|W_1\|_2 < \infty$ for all n. Because $\mu_m = O[m^{-a(1+\varepsilon)}]$ and $\alpha(m) = O[m^{-a(1+\varepsilon)2r/(r-2)}]$ for some $0 < \varepsilon < 1$, we have

$$A_2 \le \|W_1\|_2\left[\sum_{m=1}^{\infty} m^{-a(1+\varepsilon)} + c\sum_{m=1}^{\infty} m^{-a(1+\varepsilon)}\right], \quad \text{with } c = 15^{1/2}\|W_1\|_r.$$

Because $a(1 + \varepsilon) > 1$, we get $A_2 < \infty$ when $\{V_n\}$ is α-mixing. Similarly, we get $A_2 < \infty$ when $\{V_n\}$ is ϕ-mixing. Therefore, $S \in \mathcal{S}(\mathbb{H})$.

Now we show $\|S_n - S\|_{\text{tr}} \to 0$ as $n \to \infty$, i.e., for each arbitrary CONS $\{e_l\}$, $\sum_{l=1}^{\infty} |((S_n - S)e_l, e_l)| < \infty$. By weak stationarity we have

$$\begin{aligned}(S_n h, g) &= E[(W_1, h)(W_1, g)] + n^{-1}\sum_{j=2}^{n} (n - j + 1)E[(W_1, h)(W_j, g)] \\ &\quad + n^{-1}\sum_{j=2}^{n} (n - j + 1)E[(W_j, h)(W_1, g)].\end{aligned}$$

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

Because

$$(Sh,g) = E[(W_1,h)(W_1,g)] + \sum_{j=2}^{\infty} E[(W_1,h)(W_j,g)] + \sum_{j=2}^{\infty} E[(W_j,h)(W_1,g)],$$

it follows that $|((S - S_n)e_l,e_l)| \le 2[A_3^l + A_4^l]$, where

$$A_3^l \equiv |\Sigma_{n+1\le j<\infty} E[(W_1,e_l)(W_j,e_l)]|; \qquad A_4^l \equiv |\Sigma_{2\le j\le n}[(j-1)/n]E[(W_1,e_l)(W_j,e_l)]|.$$

It suffices to show that $A_3 \equiv \sum_{l=1}^{\infty} A_3^l < \infty$ and $A_4 \equiv \sum^{l=1^{\infty}} A_4^l < \infty$ for any CONS $\{e_l\}$. We get the following upper bounds for A_3 and A_4 just as for A_2:

$$A_3 \le \Sigma_{n+1\le j<\infty} \sum_{l=1}^{\infty} |E[(W_1,e_l)(W_j,e_l)]| \le \|W_1\|_2(1+c)\Sigma_{n\le m<\infty} m^{-a(1+\varepsilon)},$$

$$A_4 \le \Sigma_{2\le j<n}[(j-1)/n] \sum_{l=1}^{\infty} |E[(W_1,e_l)(W_j,e_l)]|$$

$$\le \|W_1\|_2(1+c)\Sigma_{1\le m<n-1}(m/n)m^{-a(1+\varepsilon)}.$$

Hence, for each arbitrary CONS $\{e_l\}$,

$$\sum_{l=1}^{\infty} |((S_n - S)e_l,e_l)| \le \|W_1\|_2(1+c)[\Sigma_{n\le m<\infty} m^{-a(1+\varepsilon)} + n^{-1}\Sigma_{1\le m<n-1} m^{-a(1+\varepsilon)+1}].$$

Because for any $i > j > 0$, $0 < q \ne 1$,

$$\Sigma_{j\le m<i} m^{-q} = O([i^{-q+1} - j^{-q+1}]/[-q+1]),$$

for any $1 < a(1+\varepsilon) < 2$, we have

$$\Sigma_{n\le m<\infty} m^{-a(1+\varepsilon)} = O(-n^{-a(1+\varepsilon)+1}/[-a(1+\varepsilon)+1]) = O(n^{-a(1+\varepsilon)+1}),$$

and

$$n^{-1}\Sigma_{1\le m<n-1} m^{-a(1+\varepsilon)+1} = O(n^{-1}[(n-1)^{-a(1+\varepsilon)+2} - 1]/[-a(1+\varepsilon)+2])$$

$$\le n^{-a(1+\varepsilon)+1} - n^{-1} = O(n^{a(1+\varepsilon)+1}).$$

Hence, for each arbitrary CONS $\{e_l\}$,

$$\sum_{l=1}^{\infty} |((S_n - S)e_l,e_l)| = O(n^{-a(1+\varepsilon)+1})\|W_1\|_2.$$

This completes the proof. ■

Proof of Theorem 3.9. We prove this by verifying Lemma 3.1. For any fixed $h \in \mathbb{H}$, $h \ne 0$, $\{(W_{n,t},h)\}$ is a real-valued array of near epoch dependent process. Conditions 3.9(a)–(c) imply that $\{(W_{n,t},h)\}$ satisfies Assumptions A.1, A.2′, and A.3 in Wooldridge and White's (1989) Theorem 2.9, which implies $\sum_{t=1}^{n}(W_{n,t},h) \Rightarrow \mathcal{N}(0,(Sh,h))$ as $n \to \infty$. Thus condition 3.1(a) is satisfied. By Remark 3.3, Lemma 3.7, and condition 3.9(c), conditions 3.1(b) and (c) are satisfied. The result now follows from Lemma 3.1. ■

Because the proofs of Theorem 4.2, Theorem 4.3, and Corollary 4.5 are very similar, we just give a proof of Theorem 4.3 here.

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

Proof of Theorem 4.3. By definition, $\{Z_{n,i}, \mathcal{F}^{n,i}\}$ is an $\mathbb{H}$-valued martingale difference array. For each fixed $h \in \mathbb{H}$, $\{(Z_{n,i}, h), \mathcal{F}^{n,i}\}$ is a real-valued martingale difference array, and $\{(Y_n(t), h) : t \in [0,1]\}$ is a $C_{\mathbb{R}}[0,1]$-valued random sequence. By Gaenssler and Haeusler's (1986) Theorem 2.3 and conditions 4.3(d) and (e),

$$(Y_n, h) \Rightarrow X_h \quad \text{in } C_{\mathbb{R}}[0,1], \quad \text{and} \quad \Sigma_{1 \le i \le b_n(1)} (Z_{n,i}, h) \Rightarrow \mathcal{N}(0, \sigma^2(h)) \quad \text{in } \mathbb{R},$$

where X_h is a Brownian motion in $C_{\mathbb{R}}[0,1]$ with $X_h(1) = \mathcal{N}(0, \sigma^2(h))$. Let T_n be the conditional covariance operator of $\Sigma_{1 \le i \le b_n(1)} Z_{n,i}$. Then $\lim_{n \to \infty} (T_n h, h) = \sigma^2(h)$ for any $h \in \mathbb{H}$. Given any CONS $\{e_l\}$ of $\mathbb{H}$, condition 4.3(b) implies that, for any $k \in \mathbb{N}$,

$$\sum_{l=1}^{k} \lim_n (T_n e_l, e_l) = \lim_n \sum_{l=1}^{k} (T_n e_l, e_l) \le \limsup_n \sum_{l=1}^{\infty} (T_n e_l, e_l) \le C < \infty.$$

Letting $k \to \infty$, we get for any CONS $\{e_l\}$ of $\mathbb{H}$,

$$\sum_{l=1}^{\infty} \sigma^2(e_l) = \sum_{l=1}^{\infty} \lim_n (T_n e_l, e_l) \le C < \infty.$$

Hence there exists $S \in \mathcal{S}(\mathbb{H})$ satisfying (i).

For each fixed $k \in \mathbb{N}$, $\{P_k Z_{n,i}, \mathcal{F}^{n,i}\}$ is an $\mathbb{H}_k$-valued martingale difference array. Define $P_k Y_n(t) \equiv \Sigma_{1 \le i \le b_n(t)} P_k Z_{n,i}$ if $b_n(t) > b_n(t-)$ and by linear interpolation otherwise. By the preceding one-dimensional CLT and FCLT result and the Cramer–Wold device, $\{P_k Y_n\}$ satisfies 4.1(a) with X^k a Brownian motion in $\mathbb{H}_k$, with $X^k(0) = 0$, $EX^k(1) = 0$, and Cov $X^k(1) = P_k S P_k$, where S is given by the result (i). Now 4.1(b) holds with $S^k \equiv P_k S P_k$. Because for each fixed $k \ge 1$, $\{Y_n(t) - P_k Y_n(t) : t \in [0,1]\}$ is again a martingale in $C_{\mathbb{H}}[0,1]$, Doob's inequality and condition 4.3(c) imply 4.1(c):

$$\limsup_n E\Big(\sup_{t \in [0,1]} [|Y_n(t) - P_k Y_n(t)|^2] \Big) \le \limsup_n 4E[|Y_n(1) - P_k Y_n(1)|^2] \to 0,$$

where the limit is taken as $k \to \infty$. Now Lemma 4.1 implies that $Y_n \Rightarrow X$ in $C_{\mathbb{H}}[0,1]$ and $\Sigma_{1 \le i \le b_n(1)} Z_{n,i} \Rightarrow \mathcal{N}(0, S)$ in $\mathbb{H}$. This together with condition 4.3(a) implies result (ii). ■

Proof of Theorem 4.6. For any fixed $h \in \mathbb{H}$, $h \ne 0$, $\{(W_{n,i}, h)\}$ is a real-valued array with zero mean, finite rth moment, NED on $\{V_{n,i}\}$. Given conditions 4.6(a)–(c), $\{(W_{n,i}, h)\}$ satisfies Assumptions A.1, B.2, and A.3–A.5 in Wooldridge and White (1988). By Wooldridge and White's (1988) Theorem 2.11, $(X_n, h) \Rightarrow X_h$ in $C_{\mathbb{R}}[0,1]$, where X_h is a Brownian motion in $C_{\mathbb{R}}[0,1]$ with $X_h(1) = \mathcal{N}(0, \sigma^2(h))$. This together with condition 4.6(d) implies the existence of $S \in \mathcal{S}(\mathbb{H})$ satisfying (i). By the same proof as for Theorem 4.3, we have that conditions 4.1(a) and (b) are satisfied. Now Remark 3.6, Lemma 3.7, and condition 4.6(e) imply condition 4.1(c). The result (ii) follows from Lemma 4.1; result (iii) follows from the continuous mapping theorem. ■

Proof of Proposition 5.1.

(i) In the following proof we denote $K_b(u) \equiv K(u/b_n) b_n^{-d}$ for any $u \in \mathbb{R}^d$, $\hat{f}_n(y) = n^{-1} \sum_{t=1}^{n} K_b(y - X_t)$ for any $y \in \mathbb{R}^d$, and

$$\hat{F}_n(x) = \int_{\mathbb{R}^d} \hat{f}_n(y) 1(y \le x)\, dy = n^{-1} \sum_{t=1}^{n} \int_{\mathbb{R}^d} K_b(y - X_t) 1(y \le x)\, dy.$$

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

We need to show (1) $n^{1/2}(\hat{F}_n(\cdot) - E\hat{F}_n(\cdot)) \Rightarrow \mathcal{N}(0,S)$ in $L_2(\mathbb{R}^d)$; and (2) $n^{1/2}(E\hat{F}_n(\cdot) - F(\cdot)) = o(1)$ in $L_2(\mathbb{R}^d)$.

(2) is implied by $n\int[\int(\int K(u)[f(y + b_n u) - f(y)]du)1(y \le x)dy]^2 dx \to 0$ as $n \to \infty$, which is automatically satisfied by the assumptions on K and f and $nb_n^p \to 0$ for $p > d$.

We apply Theorem 3.9 to show (1). Let $n^{1/2}(\hat{F}_n(\cdot) - E\hat{F}_n(\cdot)) = \sum_{t=1}^n W_{n,t}(\cdot)$, with

$$W_{n,t}(\cdot) = n^{-1/2}\left[\int_{\mathbb{R}^d} K_b(y - X_t)1(y \le \cdot)\,dy - \int_{\mathbb{R}^d} E[K_b(y - X_t)]1(y \le \cdot)\,dy\right].$$

Clearly $\{W_{n,t}\}$ is a zero-mean $L_2(\mathbb{R}^d)$-valued sequence NED of size -1 on some underlying mixing process. Now

$$\begin{aligned} &E[(|W_{n,t}(\cdot)|_{L_2(\mathbb{R}^d)})^4] \\ &= n^{-2}\left[E\left(\int A(z)\,dz\right)^2 + \left(\int C(z)\,dz\right)^2 + E\left(\int B(z)\,dz\right)^2\right. \\ &\quad + 2\int C(z)\,dz \quad E\int A(z)\,dz \\ &\quad \left. - 2E\left(\int A(z)\,dz \int B(z)\,dz\right) - 2\int C(z)\,dz \quad E\int B(z)\,dz\right], \end{aligned}$$

where

$$A(z) \equiv \left[\int K_b(y - X_t)1(y \le z)\,dy\right]^2,$$

$$C(z) \equiv \left[\int E(K_b(y - X_t))1(y \le z)\,dy\right]^2,$$

$$B(z) \equiv 2\int K_b(y - X_t)1(y \le z)\,dy \int E(K_b(y - X_t))1(y \le z)\,dy.$$

It is easy to check that $\sup_{1\le t\le n}\|W_{n,t}(\cdot)\|_4 \le \text{const.}\, n^{-1/2}$. Hence conditions 3.9(a) and (b) are satisfied with $r = 4$ and $c_{n,t} = \text{const.}\, n^{-1/2}$. To verify condition 3.9(c), it suffices to show that (3) $\sup_n E|\sum_{t=1}^n W_{n,t}(\cdot)|^2] \le$ const. and (4) $\lim_{n\to\infty} E[(\sum_{t=1}^n W_{n,t}, h)(\sum_{s=1}^n W_{n,t}, g)] = (Sh, g)$ for any $h, g \in L_2(\mathbb{R}^d)$. Lemma 3.7 implies (3). To verify (4), we denote $E[(\sum_{t=1}^n W_{n,t}, h)(\sum_{s=1}^n W_{n,s}, g)] \equiv [a_{1,n} - a_{3,n}] + [a_{2,n} - a_{4,n}]$, where

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

$$a_{1,n} \equiv n^{-1} \sum_{t=1}^{n} E\Bigg[\iint K_b(y - X_t) 1(y \le x)\, dy\, h(x)\, dx \times \iint K_b(y' - X_t) 1(y' \le x')\, dy' g(x')\, dx' \Bigg],$$

$$a_{2,n} \equiv n^{-1} \Sigma_{s \ne t} E\Bigg[\iint K_b(y - X_t) 1(y \le x)\, dy\, h(x)\, dx \times \iint K_b(y' - X_s) 1(y' \le x')\, dy' g(x')\, dx' \Bigg],$$

$$a_{3,n} \equiv n^{-1} \sum_{t=1}^{n} \iint EK_b(y - X_t) 1(y \le x)\, dy\, h(x)\, dx \times \iint EK_b(y' - X_t) 1(y' \le x')\, dy' g(x')\, dx',$$

$$a_{4,n} \equiv n^{-1} \Sigma_{s \ne t} \iint EK_b(y - X_t) 1(y \le x)\, dy\, h(x)\, dx \times \iint EK_b(y' - X_s) 1(y' \le x')\, dy' g(x')\, dx'.$$

Because X_t has common marginal density f, we have

$$a_{1,n} = \iiiiint K_b(y - z) K_b(y' - z) f(z) 1(y \le x)\, dy\, h(x) 1(y' \le x') \times g(x')\, dz\, dx\, dx'\, dy\, dy'.$$

By change of variable, we get

$$\begin{aligned} a_{1,n} &= \iiint \Bigg[\int K(u) 1(u \le (x - z)/b)\, du \int K(u') 1(u' \le (x' - z)/b)\, du' \Bigg] \\ &\quad \times f(z) h(x) g(x')\, dz\, dx\, dx' \\ &= \iiint [\bar{K}((x - z)/b) \bar{K}((x' - z)/b)] f(z) h(x) g(x')\, dz\, dx\, dx', \end{aligned}$$

where $\bar{K}(x) \equiv \int K(u) 1(u \le x)\, du$ for any $x \in \mathbb{R}^d$. Because K is a nonnegative density function, $\bar{K}$ is a c.d.f. function. Because $\sup_y |(b_n)^{-d} EK((y - X_t)/b) - f(y)| \to 0$ as $n \to \infty$, we have

$$\begin{aligned} \lim_{n \to \infty} [a_{1,n} - a_{3,n}] &= \iint \Bigg[\int f(z) 1(z \le \min(x, x'))\, dz - \int f(y) 1(y \le x)\, dy \int f(y') 1(y' \le x')\, dy' \Bigg] \\ &\quad \times h(x) g(x')\, dx\, dx' \\ &= \iint [F(\min(x, x')) - F(x) F(x')] h(x) g(x')\, dx\, dx'. \end{aligned}$$

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

Similarly,

$$a_{2,n} = n^{-1}\Sigma_{t\neq s}\iiiint\!\!\iiint K_b(y-z)K_b(y'-z')1(y\le x)1(y'\le x')$$
$$\times h(x)g(x')f_{X_t,X_s}(z,z')\,dx\,dx'\,dy\,dy'\,dz\,dz'.$$

Again by change of variable,

$$a_{2,n} = n^{-1}\Sigma_{t\neq s}\iiiint$$
$$\times\left[\int K(u)1(u\le (x-z)/b)\,du\int K(u')1(u'\le (x'-z')/b)\,du'\right]$$
$$\times f_{X_t,X_s}(z,z')h(x)g(x')\,dx\,dx'\,dz\,dz'$$
$$= n^{-1}\Sigma_{t\neq s}\iiiint\,[\bar{K}((x-z)/b)\bar{K}((x'-z')/b)]$$
$$\times f_{X_t,X_s}(z,z')h(x)g(x')\,dx\,dx'\,dz\,dz'.$$

Because f_{X_t,X_s} only depends on $|t-s|$, we have

$$\lim_{n\to\infty}[a_{2n}-a_{4,n}] = 2\sum_{j=2}^{\infty}\iint$$
$$\times\left[\iint f_{1,j}(z,z')1(z\le x,z'\le x')\,dz\,dz'\right.$$
$$\left.-\int f(y)1(y\le x)\,dy\int f(y')1(y'\le x')\,dy'\right]$$
$$\times h(x)g(x')\,dx\,dx'$$
$$= 2\sum_{j=2}^{\infty}\iint[P(X_1\le x, X_j\le x')$$
$$-F(x)F(x')]h(x)g(x')\,dx\,dx'.$$

Now we obtain (1) by applying Theorem 3.9.

(ii) By applying a proof similar to that in Fernholz (1991) or Yukich (1992), we can show that $n^{1/2}[\hat{F}_{se,n}-\hat{F}_{e,n}] = o_p(1)$ in $L_2(\mathbb{R}^d)$. The result now follows from (i). ■

Proof of Proposition 5.2. Let $L_2(\mathcal{X},F)$ be the space of all square-integrable functions on $\mathcal{X}$ with respect to F. Define $T_K: L_2(\mathcal{X},F)\to L_2(\mathcal{X},F)$ as $T_K f(x)\equiv E[K(X_1,x)f(X_1)]$ for all $f\in L_2(\mathcal{X},F)$. Then T_K is a compact self-adjoint linear operator with eigenvalues $\{\lambda_j\}$ and eigenfunctions $\{g_j\}$ satisfying $\int K(x,y)g_j(y)\,dF(y) = \lambda_j g_j(x)$; $\int g_j(x)\,dF(x) = 0$ for all $j\ge 1$; $g_0\equiv 1$, $(\lambda_0 = 0)$; $\int g_j(x)g_l(x)\,dF(x) = \delta_{jl}$, where $\delta_{jl} = 1$ for $j = l$, $\delta_{jl} = 0$ for $j\neq l$. Now $\iint K^2(x,y)\,dF(x)\,dF(y) = \sum_{l=1}^{\infty}(\lambda_l)^2 < \infty$, and we have the following eigenfunction expansion: $K(x,y) = \sum_{l=1}^{\infty}\lambda_l g_l(x)g_l(y)$ for all $x,y\in\mathcal{X}$, where the convergence is in

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

mean square with respect to the product measure $dF \times dF$. In addition, the convergence is also almost everywhere in x, y (with respect to $dF \times dF$) as a result of $\sum_{j=1}^{\infty} |\lambda_j| < \infty$. Hence

$$U_n = 2[n(n-1)]^{-1}\Sigma\Sigma_{1\le i<j\le n}K(X_i,X_j) = 2[n(n-1)]^{-1}\Sigma\Sigma_{1\le i<j\le n}\sum_{l=1}^{\infty}\lambda_l g_l(X_i)g_l(X_j).$$

Because $\sum_{j=1}^{\infty}|\lambda_j| < \infty$, we can define another Hilbert space:

$$\mathbb{H} = \left\{f = (f_1, f_2, \ldots) \in L_2(\mathcal{X}, F) : \sum_{l=1}^{\infty}|\lambda_l|(f_l)^2 < \infty\right\},$$

where $f_l \equiv \int f(x)g_l(x)\,dF(x)$, with inner product $(f,h) \equiv \sum_{l=1}^{\infty}\lambda_l f_l h_l$ and norm $|f| \equiv |\sum_{l=1}^{\infty}\lambda_l(f_l)^2|^{1/2}$. Denote $e_l(\cdot) \equiv g_l(\cdot)/\lambda_l$ for $l \ge 1$; clearly $\{e_l\}$ is a complete orthonormal basis for $\mathbb{H}$. In fact, $e_l = (0,\ldots,0,1/\lambda_l,0,\ldots)$. Let $W_t = (g_1(X_t), g_2(X_t), \ldots)$; then $\{W_t\}$ is a sequence of $\mathbb{H}$-valued stationary random variables NED of size -1 on some strong mixing process of size $-2r/(r-2)$ for $r > 2$. Now $EW_t = 0$ because K is degenerate. By Corollary 3.10, $n^{-1/2}\sum_{t=1}^{n}W_t \Rightarrow \mathcal{N}(0,S)$ in $\mathbb{H}$, with $S \in \mathcal{S}(\mathbb{H})$ such that, for any l, j,

$$\begin{aligned}(Se_l, e_j) &= \lim_{n\to\infty} E\left[\left(n^{-1/2}\sum_{t=1}^{n}W_t, e_l\right)\left(n^{-1/2}\sum_{s=1}^{n}W_s, e_j\right)\right] \\ &= \lim_{n\to\infty} n^{-1}E\left[\left(\sum_{t=1}^{n}g_l(X_t)\right)\left(\sum_{s=1}^{n}g_j(X_s)\right)\right], \text{ thus}\end{aligned}$$

$$(Se_l, e_j) = E[g_l(X_1)g_j(X_1)] + \sum_{i=2}^{\infty}E[g_l(X_1)g_j(X_i)] + \sum_{i=2}^{\infty}E[g_j(X_1)g_l(X_i)].$$

By the continuous mapping theorem,

$$\left|n^{-1/2}\sum_{t=1}^{n}W_t\right|^2 \Rightarrow |\mathcal{N}(0,S)|^2 = \sum_{l=1}^{\infty}\lambda_l(Z_l)^2,$$

where $\{Z_l\}$ is a sequence of correlated real-valued normals $\mathcal{N}(0,\sigma_l^2)$, with $\mathrm{Cov}(Z_l,Z_j) \equiv (Se_l,e_j)$, and $\sigma_l^2 \equiv (Se_l,e_l)$. Now

$$nU_n = 2(n-1)^{-1}\Sigma_{1\le i<j\le n}(W_i,W_j) = (n-1)^{-1}\left|\sum_{t=1}^{n}W_t\right|^2 - (n-1)^{-1}\sum_{t=1}^{n}|W_t|^2.$$

Because $E\{n^{-1}\sum_{t=1}^{n}|W_t|^2\} \to \sum_{l=1}^{\infty}\lambda_l < \infty$ as $n \to \infty$, we obtain $nU_n \Rightarrow \sum_{l=1}^{\infty}\lambda_l((Z_l)^2 - 1)$. ■

Proof of Proposition 5.3. Result (i) is a direct consequence of conditions 5.3(e) and (f) on $\{\nu_n\}$, Lemmas 3.7 and 3.8, and Theorem 4.6.

We can prove result (ii) by applying Theorem 1.5.8 in Ljung, Pflug, and Walk (1992, p. 32). We have

$$\begin{aligned}\hat{\theta}_{n+1} - \theta^* = \hat{\theta}_n - \theta^* &+ n^{-1}[M_n(Z_n,\hat{\theta}_n) - M_n(Z_n,\theta^*)] \\ &+ n^{-1}[M_n(Z_n,\theta^*) - EM_n(Z_n,\theta^*)] + n^{-1}EM_n(Z_n,\theta^*).\end{aligned}$$

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press

By condition 5.3(a), for any small $\varepsilon > 0$, there exists N_ε sufficiently large that for all $n \geq N_\varepsilon$, we have $|\hat{\theta}_n - \theta^*| < \varepsilon$. This, together with conditions 5.3(c.ii) and (d), implies that for all $n \geq N_\varepsilon$, the following holds almost surely:

$$\hat{\theta}_{n+1} - \theta^* = \hat{\theta}_n - \theta^* + n^{-1}A(\hat{\theta}_n - \theta^*) + n^{-(1+\beta)/2}\nu_n + n^{-(1+(\beta/2))}T.$$

This, together with result (i) and conditions 5.3(b)–(e), implies that all the conditions for Theorem 1.5.8 in Ljung, Pflug, and Walk (1992) are satisfied, so that result (ii) holds. ■

https://doi.org/10.1017/S0266466698142056 Published online by Cambridge University Press