



Contents lists available at ScienceDirect

## Journal of Statistical Planning and Inference

journal homepage: [www.elsevier.com/locate/jspi](http://www.elsevier.com/locate/jspi)

# On functional central limit theorems for dependent, heterogeneous arrays with applications to tail index and tail dependence estimation

Jonathan B. Hill

Department of Economics, University of North Carolina, CB 3305, Chapel Hill, NC 27599-3305, USA

## ARTICLE INFO

*Article history:*

Received 18 September 2007

Received in revised form

27 June 2008

Accepted 21 September 2008

Available online 2 October 2008

## MSC:

primary 60F17

*Keywords:*

Functional central limit theorem

Tail arrays

Tail empirical process

Extremal near epoch dependence

Hill estimator

Tail quantile process

Tail dependence

## ABSTRACT

We establish invariance principles for a large class of dependent, heterogeneous arrays. The theory equally covers conventional arrays, and inherently degenerate tail arrays popularly encountered in the extreme value theory literature including sample means and covariances of tail events and exceedances. For tail arrays we trim dependence assumptions down to a minimum leaving non-extremes and joint distributions unrestricted, covering geometrically ergodic, mixing, and mixingale processes, in particular linear and nonlinear distributed lags with long or short memory, linear and nonlinear GARCH, and stochastic volatility.

Of practical importance the limit theory can be used to characterize the functional limit distributions of a tail index estimator, the tail quantile process, and a bivariate extremal dependence estimator under substantially general conditions.

© 2008 Elsevier B.V. All rights reserved.

## 1. Introduction

In this paper we deliver invariance principles for  $L_r$ -bounded stochastic functions, including conventional arrays and inherently degenerate tail arrays, allowing significant degrees of dependence and heterogeneity. The tail array results permit the most general available limit theory for Hill's (1975) tail index estimator, a tail empirical quantile processes, and a tail dependence estimator. Together, these results have substantial practical value: extreme value estimators are frequently applied to dependent and/or heterogeneous data in finance and macroeconomics, yet available theory does not adequately cover many processes assumed in the literature (e.g. stationary ARFIMA; nonlinear GARCH processes; regime switching).

We present invariance principles of the form

$$\sum_{t=1}^{n(\xi)} X_{n,t}(u) \Rightarrow X(\xi, u) \quad \text{where } \xi \in [0, 1] \text{ and } u \in [0, 1]^{k-1}, \quad k \geq 1,$$

*E-mail address:* [jbhill@email.unc.edu](mailto:jbhill@email.unc.edu)

*URL:* <http://www.unc.edu/~jbhill>.

for mean-zero functional arrays  $\{X_{n,t}(u)\}, X_{n,t} : [0, 1] \rightarrow \mathbb{R}$ , with  $X(\zeta, u)$  a Gaussian process with *almost surely* continuous sample paths, and  $n(\zeta)$  an integer sequence. The array  $\{X_{n,t}(u)\}$  is assumed to be uniformly  $L_r$ -bounded

$$\sup_{u \in [0,1]^{k-1}, t \geq 1} (E|X_{n,t}(u)|^r)^{1/r} = O(n^{-a(r)}) \tag{1}$$

for some mapping  $a : [1, \infty) \rightarrow (0, \frac{1}{2}]$  and  $r \geq 2$ . The standard array for a stationary finite variance process  $\{X_t\}$  is

$$X_{n,t} = \frac{n^{-1/2}(X_t - E[X_t])}{[E(n^{-1/2} \sum_{t=1}^n (X_t - E[X_t]))^2]^{1/2}},$$

hence  $k = 1$  and  $a(r) = \frac{1}{2}$  at least for  $r \in [1, 2]$ .

The main invariance principle, Theorem 2.1, equally covers near epoch dependent (NED) arrays and extremal-NED tail arrays, including the sample means and covariances of tail events and exceedances (Corollary 3.3). The NED concept dates in various forms at least to Ibragimov (1962), Ibragimov and Linnik (1971), and McLeish (1975). See Gallant and White (1988) and Davidson (1994) for historical details. E-NED, due to Hill (2005), restricts NED to extremes only, it is a marginal tail property that leaves non-extremes and joint distributions unrestricted, and characterizes at least linear and nonlinear distributed lags with long or short memory, linear and nonlinear GARCH and stochastic volatility. See Section 3.

Central limit theory for dependent, heterogeneous non-tail arrays has received prolific attention, including Hall and Heyde (1980), Wooldridge and White (1988), Davidson (1992), de Jong (1997), Davidson and de Jong (2000), Wang et al. (2002), Wu and Woodroffe (2004) and Wu and Min (2005) to name a few.

Some of the best invariance principles are offered in Davidson and de Jong (2000) for NED arrays, and Wu and Min (2005) for  $L_p$ -weakly dependent processes, cf. Wu and Woodroffe (2004). We show in Section 4 that none of their results apply to *tail arrays* due to degeneracy properties. Tail arrays are critical for risk management and cost, damage and catastrophe modeling in financial, actuarial and meteorological studies. Consult Beirlant et al. (1994), Embrechts et al. (1997), Rootzén et al. (1998) and Finkenstadt and Rootzén (2003) for background theory and applications.

Consider some process  $\{X_t\}$  that takes values on  $[\infty, 0)$ , where  $\{X_t\}_{t=1}^n$  is the sample path, and  $\{k_n\}$  and  $\{b_n\}$  are real-valued sequences satisfying  $k_n \in \mathbb{N}, k_n \rightarrow \infty, k_n/n \rightarrow 0, b_n \rightarrow \infty$  and  $(n/k_n)P(X_t \geq b_n) \rightarrow 1$  as  $n \rightarrow \infty$ . Thus,  $b_n$  is the asymptotic  $k_n/n$ th  $\rightarrow 0$  quantile and any  $X_t \geq b_n$  is an extreme value.

Tail arrays include the  $b_n$ -event process

$$U_{n,t}^*(u) := k_n^{-1/2} \{I(X_t > b_n/u) - P(X_t > b_n/u)\}, \quad u = [0, 1], \tag{2}$$

the  $b_n$ -exceedance process

$$U_{n,t} := k_n^{-1/2} \{(\ln X_t/b_n)_+ - E[(\ln X_t/b_n)_+]\} \quad \text{where } (z)_+ := \max\{z, 0\}, \tag{3}$$

and an intermediate tail empirical quantile processes,

$$q_{k_n} := k_n^{-1/2} \times \ln(X_{(k_n)}/b_n) \quad \text{where } X_{(1)} \geq X_{(2)} \geq \dots \text{ are ranked } X_t \text{ 's.} \tag{4}$$

The tail array  $\{U_{n,t}^*(u)\}$  is a uniformly bounded function of a possibly infinite variance process. Central limit theorems for bounded functions of  $\alpha$ -stable moving averages are delivered in Hsing (1999), Pipiras and Taqqu (2003), and Pipiras et al. (2007). Property (1) does not require boundedness of  $X_{n,t}(u)$ , and the process  $X_t$  on which  $X_{n,t}(u)$  is based extends well beyond moving averages, the  $\alpha$ -stable laws and their domains of attraction.

Rootzén et al. (1998) analyze generic tail array sums under a mixing condition, and Hsing (1991) analyzes  $\{U_{n,t}^*(u), U_{n,t}\}$  for strong mixing processes. See also Leadbetter and Rootzén (1993). Rootzén (1995), Drees (2002) and Einmahl and Lin (2006) consider functional limit theory for iid and mixing tail empirical processes. Typically the imposed restrictions are rather abstract (Rootzén, 1995; Einmahl and Lin, 2006), or have only been exemplified for linear processes (Einmahl and Lin, 2006: asymptotically independent extremes, AR(1)). See also Csörgö et al. (1986), Mason (1988), and Einmahl (1992).

Our motivation is the use of tail arrays for tail shape and dependence characterization and estimation. We deliver an invariance principle for the  $k_n$ th-exceedance mean

$$\hat{\alpha}_{k_n}^{-1}(\zeta) := \frac{1}{k_n} \sum_{t=1}^{n(\zeta)} (\ln X_t/X_{([k_n \zeta] + 1)})_+, \tag{5}$$

a functional estimator of the index of regular variation proposed by Hill (1975) and a natural measure of tail thickness (Resnick, 1987) and therefore financial market risk (Bradley and Taqqu, 2003). The limit theory is apparently the most general available.

Similarly, we prove an invariance principle for the extremal event covariance

$$\hat{\rho}_{n(\zeta)}(h, u) = \sum_{t=1}^{n(\zeta)} U_{n,1,t-h}^*(u_1) \times U_{n,2,t}^*(u_2) \quad \text{where } u = [u_1, u_2]' \in [0, 1]^2 \tag{6}$$

for a joint process  $\{X_{1,t}, X_{2,t}\}$  where  $U_{n,i,t}^*(u_i)$  is defined by (2) for each  $X_{i,t}$ . We do not require a specified bivariate tail shape and we allow for any form of non-extremal dependence, all in contrary to extant conventions. Cf. [Ledford and Tawn \(1997, 2003\)](#), [Coles et al. \(1999\)](#), [Embrechts et al. \(2003\)](#), [Heffernan and Tawn \(2004\)](#), [Schmidt and Stadtmüller \(2006\)](#) and [Klüppelberg et al. \(2007\)](#).

Section 2 contains the main result for  $\{X_{n,t}(u)\}$ , Section 3 delivers invariance principles for tail arrays  $\{U_{n,1,t}, U_{n,1,t}^*(u)\}$ , Section 4 presents examples, and Section 5 characterizes limit laws for  $\hat{\alpha}_{k_n}^{-1}(\xi)$ ,  $q_{k_m}$  and  $\hat{\rho}_{n(\xi)}(u)$ . All proofs are relegated to Appendix A.1, and Appendix A.2 contains supporting lemmata.

Throughout  $\xrightarrow{p}$  and  $\xrightarrow{d}$ , respectively, denote convergence in probability and finite dimensional distributions, and  $\Rightarrow$  denotes weak convergence on a metric space. Gaussian elements of function spaces have zero means.  $[z]$  is the integer part of  $z$ . Write  $(z)_+ := \max\{z, 0\}$ .  $|\cdot|$  denotes the  $l_1$ -matrix norm:  $|x| := \sum |x_{ij}|$ , and  $\|\cdot\|_p$  is the usual  $L_p$ -matrix norm:  $\|x\|_p := (\sum E|x_{ij}|^p)^{1/p}$ .  $K > 0$  is always a finite constant whose value may change from line to line. Write  $\{x_n\} = \{x_n\}_{n \geq 1}$ .

## 2. Assumptions and main result

Let  $\{X_t\} = \{X_t : -\infty < t < \infty\}$  be a stochastic process on the probability measure space  $(Y, \mathfrak{F}, \mu)$ , where  $\mathfrak{F} = \sigma(\cup_{t \in \mathbb{Z}} \mathfrak{F}_t)$  and  $\mathfrak{F}_{t-1} \subset \mathfrak{F}_t \equiv \sigma(X_\tau : \tau \leq t)$ .

### 2.1. Cadlag processes

We will work with processes  $x(\xi, u)$  on the following *cadlag* space (consult [Billingsley, 1999](#))

$$D_k := D([\underline{\xi}, 1] \times [0, 1]^{k-1}), \quad k \geq 1,$$

where  $\underline{\xi} \in (0, 1]$ ,  $u \in [0, 1]^{k-1}$ , and  $x(1, u) = x(1-, u)$  for every element  $x \in D_k$ . If  $k = 1$  then by convention  $x(\xi, u) = x(\xi)$ . The martingale difference approximation argument used to prove the main result, Theorem 2.1, is greatly expedited by bounding  $\xi$  away from zero.

Let  $X_{n,t}(u)$  be  $D[0, 1]^{k-1}$ -valued and define

$$X_n(\xi, u) := \sum_{t=1}^{n(\xi)} X_{n,t}(u).$$

The functional  $n : [0, 1] \rightarrow \mathbb{N}$  is right continuous with left limits, non-decreasing,  $n(\xi) \rightarrow \infty$  as  $n \rightarrow \infty$ ,  $n(\xi_1) - n(\xi_2) \rightarrow \infty \forall \xi_1 > \xi_2$ ,  $n(0) = 0$ , and  $n(1-) = n(1) \leq n$ .

We assume  $\{X_{n,t}(u)\}$  is  $L_r$ -bounded in the sense of (1). Examples in the extreme value theory literature are detailed in Section 3.

**Definition (Functional array).**  $\{X_{n,t}(u)\}$  is an  $L_r$ -functional array of  $\{X_t\}$  if (i)  $X_{n,t}(u) \in D[0, 1]^{k-1}$ ; (ii)  $E[X_{n,t}(u)] = 0$ ; and (iii) there exists a function  $a : [1, \infty) \rightarrow (0, \frac{1}{2}]$ , such that  $\sup_{u \in [0, 1]^{k-1}, t \geq 1} \|X_{n,t}(u)\|_r = O(n^{-a(r)})$  for some  $r \geq 2$ .

**Remark 1.** We call  $a(r)$  the  $r$ th-moment index.

**Remark 2.** The standard case  $X_{n,t} = n^{-1/2}(X_t - E[X_t]) / \|n^{-1/2} \sum_{t=1}^n (X_t - E[X_t])\|_2$ , where  $\sup_{t \geq 1} \|X_t\|_2 < \infty$ , implies  $a(r) = 2$  for at least  $r \in [1, 2]$ .

**Remark 3.** Although  $L_r$ -boundedness  $\|X_{n,t}(u)\|_r \leq Kn^{-a(r)}$  uniformly in  $t$  is inherently satisfied for tail arrays  $\{U_{n,t}^*(u), U_{n,t}\}$  under minimal assumptions, it does limit the scope of the main results for non-stationary non-degenerate arrays (cf. [de Jong, 1997](#); [Davidson and de Jong, 2000](#); [Wu and Min, 2005](#)).

### 2.2. F-mixing and F-NED: tail memory and heterogeneity

The following dependence concepts are developed in [Hill \(2005\)](#) for a less general context. Let  $\{\varepsilon_t\}$  be an arbitrary, possibly vector-valued stochastic process with  $\sigma$ -algebra

$$G_t := \sigma(\varepsilon_\tau : \tau \leq t) \quad \text{where} \quad G_b^a := \sigma(\varepsilon_t : a \leq t \leq b).$$

In some minimal sense we want to be able to predict  $X_{n,t}(u)$  using  $G_t$ -measurable information  $E_{n,t}$  induced from  $\varepsilon_t$ . Examples include  $E_{n,t} = \varepsilon_t$  itself, or  $E_{n,t} = h(\varepsilon_t, \varepsilon_{t-1}, \dots)$  a Borel measurable function; but also the extreme event  $E_{n,t} = I(|\varepsilon_t| > \pi_{n,t})$ , exceedance  $(|\varepsilon_t| - \pi_{n,t})_+$  and value  $|\varepsilon_t| \times I(|\varepsilon_t| > \pi_{n,t})$  for some non-stochastic array  $\{\pi_{n,t}\}$ ,  $\pi_{n,t} \rightarrow \infty$  as  $n \rightarrow \infty$ , or Borel measurable functions of extreme values.

Typically  $E_{n,t} = \varepsilon_t$  is iid or a martingale difference. The point here is we only require  $\varepsilon_t$  to satisfy a weak mixing condition in the sense that the  $G_t$ -measurable information  $E_{n,t}$  is assumed to mix.

Throughout we use the following array of  $\sigma$ -fields:

$$F_{n,t} := \sigma(E_{n,\tau} : \tau \leq t) \quad \text{where } F_{n,s}^t := \sigma(E_{n,\tau} : s \leq \tau \leq t) \tag{7}$$

and a sequence  $\{l_n\}$  of integer displacements where  $l_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Define  $\sigma$ -sub-fields  $A_{n,t} \in F_{n,-\infty}^t$  and  $B_{n,t+l_n} \in F_{n,t+l_n}^\infty$ , and define mixing coefficients

$$\begin{aligned} \varepsilon_{l_n} &\equiv \sup_{A_{n,t}, B_{n,t+l_n}, t \in \mathbb{Z}} |P(A_{n,t} \cap B_{n,t+l_n}) - P(A_{n,t})P(B_{n,t+l_n})|, \\ \varpi_{l_n} &\equiv \sup_{A_{n,t}, B_{n,t+l_n}, t \in \mathbb{Z}} |P(B_{n,t+l_n} | A_{n,t}) - P(B_{n,t+l_n})|. \end{aligned}$$

**Definition (F-mixing).** If  $n^{[1/2-a(r)]2\lambda} l_n^\lambda \varepsilon_{l_n} \rightarrow 0$  as  $n \rightarrow \infty$  for some  $\{l_n\}$ ,  $l_n \rightarrow \infty$ , and some  $r \geq 2$  and  $\lambda > 0$  we say  $\{\varepsilon_t\}$  is *functional-strong mixing* with size  $\lambda$ . If  $n^{[1/2-a(r)]2\lambda} l_n^\lambda \varpi_{l_n} \rightarrow 0$  as  $n \rightarrow \infty$  for some  $\{l_n\}$ ,  $l_n \rightarrow \infty$ , and some  $r \geq 2$  and  $\lambda > 0$  we say  $\{\varepsilon_t\}$  is *functional-uniform mixing* with size  $\lambda > 0$ .

**Remark.** F-mixing is simply mixing assigned to the functional  $\{E_{n,t}\}$  as  $n \rightarrow \infty$ . Hill (2005), for example, exploits a mixing tail event  $E_{n,t} = I(|\varepsilon_t| > \pi_{n,t})$ ,  $\pi_{n,t} \rightarrow \infty$  as  $n \rightarrow \infty$ , in which case  $\varepsilon_t$  is called extremal-mixing (E-mixing). Standard inequalities apply, and if  $\varepsilon_t$  is mixing then it is F-mixing. See Hill (2005) for a comparison of E-mixing and Leadbetter’s (1983) D-mixing property.

Next, we restrict tail memory and heterogeneity in  $\{X_t\}$  by assuming the functional array  $\{X_{n,t}(u)\}$  is NED. We say some stochastic array  $\{y_{n,t}\}$  is  $L_p$ -NED,  $p > 0$ , on the array of  $\sigma$ -fields  $\{F_{n,t}\}$  with size  $\lambda > 0$  if there exist arrays of non-stochastic real numbers  $\{d_{n,t}\}$  and  $\{\varphi_l\}$ , where  $d_{n,t} \geq 0$ ,  $\varphi_l \in [0, 1)$ , and  $\varphi_l = o(l^{-\lambda})$  such that

$$\|y_{n,t} - E[y_{n,t} | F_{n,t-l}^{t+l}]\|_p \leq d_{n,t} \times \varphi_l. \tag{8}$$

As  $l \rightarrow \infty$  information induced from the near epoch  $\{E_{n,\tau}\}_{\tau=l}^{t+l}$  can be used to predict  $y_{n,t}$  with vanishing prediction error in  $L_p$ -norm. The “constants”  $d_{n,t}$  permit time dependence of the  $L_p$ -norm and may satisfy  $d_{n,t} \rightarrow \infty$  as  $t \rightarrow \infty$  if there exists a trending moment. The “coefficients”  $\varphi_l$  gauge hyperbolic (i.e. “long”) memory decay. If memory is geometric (i.e. “short”)  $\varphi_l = o(\rho^{-l})$ , then size  $\lambda$  is irrelevant and therefore arbitrarily large.

Property (8) characterizes linear and nonlinear distributed lags with any degree of tail thickness and long or short memory (e.g. bounded contraction mappings, bilinear processes: see Gallant and White, 1988; Davidson, 1994), covariance stationary GARCH (Davidson, 2004), and stochastic volatility (Hill, 2008a). Further, since  $\varepsilon_t$  can in principle be anything, it can be a mixing process, and  $y_{n,t} = E_{n,t} = \varepsilon_t$  is always possible. Thus, any mixing process, including geometrically ergodic processes, can be characterized by NED (8), including nonlinear GARCH (e.g. Carrasco and Chen, 2002), nonlinear AR-GARCH, neural networks and regime switching processes (e.g. Meitz and Saikkonen, 2008).

**Definition (Functional near epoch dependence).**  $\{X_t\}$  is  $L_q$ -F-NED,  $q \geq 1$ , on  $\{F_{n,t}\}$  with size  $\lambda > 0$  if some  $L_r$ -functional-tail array  $\{X_{n,t}(u)\}$  based on  $\{X_t\}$  satisfies two conditions.

(i) For some  $\{l_n\}$ ,  $l_n \rightarrow \infty$ ,

$$\|X_{n,t}(u) - E[X_{n,t}(u) | F_{n,t-l_n}^{t+l_n}]\|_q \leq d_{n,t}(u) \times \varphi_{l_n}, \quad u \in [0, 1]^{k-1}, \tag{9}$$

where the Lebesgue measurable array  $\{d_{n,t}(u)\}$ ,  $d_{n,t} : [0, 1]^{k-1} \rightarrow \mathbb{R}_+$ , satisfies  $\sup_{u \in [0,1]^{k-1}, t \geq 1} d_{n,t}(u) = O(n^{-a(r)})$  for some  $r \geq q$ , and  $\varphi_{l_n} = o(n^{a(r)-a(q)} l_n^{-\lambda})$ . If  $k = 1$  then  $d_{n,t}(u) = d_{n,t}$ .

(ii) For  $\Delta X_{n,t}(u_1, u_2) := X_{n,t}(u_1) - X_{n,t}(u_2)$ , and  $\forall u_1, u_2 \in [0, 1]^{k-1}$ ,

$$\|\Delta X_{n,t}(u_1, u_2) - E[\Delta X_{n,t}(u_1, u_2) | F_{n,t-l_n}^{t+l_n}]\|_q \leq \left( \tilde{d}_{n,t} \times \max_{1 \leq i \leq k-1} |u_{2,i} - u_{1,i}|^{1/q} \right) \times \varphi_{l_n}, \tag{10}$$

where  $\sup_{t \geq 1} \tilde{d}_{n,t} = O(n^{-a(q)})$ .

**Remark 4.** Property (9) simply states the functional array  $\{X_{n,t}(u)\}$  is  $L_q$ -NED on  $\{F_{n,t}\}$ , and can therefore be predicted from information induced from the near epoch  $\{E_{n,\tau}\}_{\tau=l_n}^{t+l_n}$  as  $n \rightarrow \infty$ . Since  $d_{n,t}(u)\varphi_{l_n} = o(n^{-a(p)} l_n^{-\lambda})$  the scaled prediction error satisfies at least hyperbolic decay

$$n^{a(q)} \|X_{n,t}(u) - E[X_{n,t}(u) | F_{n,t-l_n}^{t+l_n}]\|_q = o(l_n^{-\lambda}).$$

**Remark 5.** In the non-functional case  $X_{n,t} = X_{n,t}(u)$  property (10) is irrelevant and (9) reduces to the  $L_q$ -NED property (8). Thus,  $L_q$ -F-NED easily captures the conventional  $L_2$ -NED setting for covariance stationary  $X_{n,t} = n^{-1/2}(X_t - E[X_t]) / \|n^{-1/2} \sum_{t=1}^n (X_t - E[X_t])\|_2$ , with  $a(2) = \frac{1}{2}$ . In this case  $n^{a(2)} \|X_{n,t} - E[X_{n,t} | F_{n,t-l_n}^{t+l_n}]\|_2 = K \|X_t - E[X_t | F_{n,t-l_n}^{t+l_n}]\|_2 = o(l_n^{-\lambda})$ .

Property (10) ensures the sequence of probability distributions associated with  $\{X_n(\xi, u)\}$  is tight. In order to ensure the limit law  $X(\xi, u)$  is also *almost surely* continuous in  $u$  we will need a partial sum of  $Z_{n,t,j}(\delta) := \sup_{|u_{1,j} - u_{2,j}| \leq \delta} |\Delta X_{n,t}(u_1, u_2)|$  to have a bounded variance. This can be ensured by assuming  $Z_{n,t,j}(\delta)$  is  $L_2$ -NED, a much stronger version of (10). Under weak conditions (e.g. [Dellacherie and Meyer, 1978](#))  $Z_{n,t,j}(\delta)$  is  $\sigma(\mathfrak{T}_t \cup F_{n,t-l_n}^{t+l_n})$ -measurable for each  $j = 1, \dots, k - 1$ , and the required extension of (10) is

$$\|Z_{n,t,j}(\delta) - E[Z_{n,t,j}(\delta) | F_{n,t-l_n}^{t+l_n}]\|_q \leq K \times \delta^{1/q} \times \tilde{d}_{n,t} \times \varphi_{l_n}, \quad j = 1, \dots, k - 1. \tag{10'}$$

### 2.3. Main result

**Assumption 1.** (a)  $\{X_{n,t}(u)\}$  is an  $L_r$ -functional array with  $r$ th-moment index  $a(r) : a(2) = \frac{1}{2}$  and  $2a(2r) > a(r)$ .

(b)  $\{X_t\}$  is  $L_2$ -F-NED in the sense of (9) with size  $\lambda = \frac{1}{2}$ . The base  $\{\varepsilon_t\}$  is F-strong mixing with size  $r/(r - 2)$  for some  $r > 2$ , or F-uniform mixing with size  $r/[2(r - 1)]$  for some  $r \geq 2$ .<sup>1</sup>

(c)  $\{X_t\}$  is  $L_2$ -F-NED in the sense of (10) with size  $\lambda = \frac{1}{2}$ , and the same mixing base  $\{\varepsilon_t\}$  as above.

(d) For some finite function  $\kappa(\xi, \delta) \geq 1$  the limit  $|n(\xi + \delta)/n(\xi) - \kappa(\xi, \delta)| \rightarrow 0$  exists  $\forall \xi \in [\underline{\xi}, 1 - \delta], \delta \in [0, 1]$ .

**Theorem 2.1.** Under Assumption 1,  $X_n(\xi, u) \Rightarrow X(\xi, u)$  on  $D_k$  where  $X(\xi, u)$  is Gaussian with independent increments,  $P(X(\xi, \cdot) \in C[\underline{\xi}, 1]) = 1$  and  $E(X_n(\xi, u)^2) = O(1)$ . Further, if (10') holds then  $P(X(\cdot, u) \in C[0, 1]^{k-1}) = 1$ .

**Remark 6.** The proof exploits a martingale difference approximation for NED  $\{X_{n,t}(u)\}$  based on Bernstein blocks, cf. [Davidson \(1992\)](#), [de Jong \(1997\)](#) and [Davidson and de Jong \(2000\)](#). The restrictions on the  $r$ th-moment index  $a(r)$  under Assumption 1(a) ensure a required Lindeberg condition is satisfied.

**Remark 7.** Assumption 1(d) expedites tightness of  $\{X_n(\xi, u)\}$ . If  $n(\xi) = [n\underline{\xi}]$  the assumption is trivial since  $\kappa(\xi, \delta) = (1 + \delta/\underline{\xi}) < \infty \forall \xi \geq \underline{\xi} > 0$ .

**Remark 8.** Suppose  $\eta(\xi, u) := \lim_{n \rightarrow \infty} \|X_n(\xi, u)\|_2$  exists for all  $\xi$  and  $u$ , where  $\eta(\xi, \cdot)$  is a non-decreasing function on  $[0, 1]$ ,  $\eta(0, \cdot) = 0, \eta(1, \cdot) = 1$ . For fixed  $u \in [0, 1]^{k-1}$  if  $\eta(\xi, \cdot) = \xi$  then  $X(\xi, \cdot)$  is Brownian motion, otherwise  $X(\xi, \cdot)$  is transformed Brownian motion.

**Remark 9.** If  $X_{n,t}(u) = X_{n,t}$  satisfies Assumption 1 and  $\lim_{n \rightarrow \infty} \|X_n(\xi)\|_2 = \eta(\xi) \in [0, 1]$ , then  $\sum_{t=1}^{n(\xi)} X_{n,t} \Rightarrow X(\xi)$  a (transformed) Brownian motion law. This is essentially a version of [Davidson and de Jong's \(2000, Theorem 3.1\)](#) invariance principle for  $L_2$ -NED arrays, and matches the generality of [Wu and Min's \(2005\)](#) results. Our  $L_r$ -functional array assumption  $\sup_{t \geq 1} \|X_{n,t}\|_r = O(n^{-a(r)})$  is simply a restricted version of Davidson and de Jong's  $L_r$ -boundedness condition. Theorem 2.1 goes much farther than either set of results, however, by including *cadlag* functionals  $X_{n,t}(u)$  and degenerate tail arrays, the subject of the next section.

A non-function central limit theorem is merely a special case of Theorem 2.1. In this case, however, Assumption 1(c)–(d) is superfluous. Let  $k = 1$  such that  $\{X_{n,t}\}$  be the  $L_r$ -functional array associated with  $X_t$ .

**Corollary 2.2.** Under Assumption 1(a)–(b)  $\sum_{t=1}^n X_{n,t} \xrightarrow{d} X$ , a Gaussian law with zero mean and finite variance.

### 3. Invariance principles for tail arrays

Assume  $X_t$  has for each  $t$  a marginal distribution  $P(X_t \leq x)$  with support on  $[0, \infty)$  and a regularly varying tail

$$\bar{F}_t(x) := P(X_t > x) = x^{-\alpha} L(x), \quad \alpha > 0 \text{ where } L(x) \text{ is slowly varying.} \tag{11}$$

See, e.g., [Bingham et al. \(1987\)](#) and [Resnick \(1987\)](#) assume  $\bar{F}_t(x)/\bar{F}_t(x-) \rightarrow 1$  as  $x \rightarrow \infty$ . Then there exist sequences  $\{k_n\}$  and  $\{b_n\}, 1 \leq k_n < n, k_n \in \mathbb{N}, k_n = o(n), k_n \rightarrow \infty$  and  $b_n \rightarrow \infty$  satisfying ([Leadbetter et al., 1983](#); Theorem 1.7.13)

$$\lim_{n \rightarrow \infty} \frac{n}{k_n} P(X_t > b_n) = 1. \tag{12}$$

<sup>1</sup> Only the F-uniform mixing case applies if  $r = 2$ .

For any  $c = [c_1, c_2]' \in \mathbb{R}^2$  construct a linear combination of the tail arrays  $\{U_{n,t}, U_{n,t}^*(u)\}$

$$Y_{n,t}(c, u) = c_1 U_{n,t} + c_2 U_{n,t}^*(u).$$

Note  $E[Y_{n,t}(c, u)] = 0$  by construction, and tail exceedances and events are inherently  $L_r$ -functional arrays.

**Lemma 3.1.** *The array  $\{Y_{n,t}(c, u)\}$  satisfies  $\sup_{u \in [0,1], t \geq 1} \|Y_{n,t}(c, u)\|_r = O((k_n/n)^{1/r} k_n^{-1/2}) = O(n^{-a(r)}) \forall r > 0$ , where  $a(1) > \frac{1}{2}$ ,  $a(2) = \frac{1}{2}$  and  $2a(2r) > a(r)$ . Further  $a(r) > 1/r$  for all  $r > 2$ .*

**Remark 10.** Assumption 1(a) is therefore satisfied for  $\{U_{n,t}, U_{n,t}^*(u)\}$ .

The  $L_p$ -F-NED properties are particularly insightful for characterizing extremal dependence in  $\{X_t\}$  if  $X_{n,t}(u) = U_{n,t}^*(u)$ . Properties (9)–(10) reduce to the following *functional-extremal-NED* property.

**Definition ( $L_p$ -FE-NED).**  $\{X_t\}$  is  $L_q$  FE-NED,  $q > 0$ , on  $\{F_{n,t}\}$  defined in (7), with size  $\lambda > 0$  if for any  $u \in [0, 1]$  and  $0 \leq u_1 \leq u_2 \leq 1$ , and some  $\{l_n\}$ ,  $l_n \rightarrow \infty$ ,

$$k_n^{-1/2} \|I(X_t > b_n/u) - P(X_t > b_n/u | F_{n,t-l_n}^{t+l_n})\|_q \leq d_{n,t}(u) \times \varphi_{l_n} \tag{13}$$

and

$$k_n^{-1/2} \|I_{n,t}(u_1, u_2) - E[I_{n,t}(u_1, u_2) | F_{n,t-l_n}^{t+l_n}]\|_q \leq \left( \tilde{d}_{n,t} \times \max_{1 \leq i \leq k-1} |u_{2,i} - u_{1,i}|^{1/q} \right) \times \varphi_{l_n}, \tag{14}$$

where  $I_{n,t}(u_1, u_2) := I(b_n/u_2 < X_t < b_n/u_1)$ , and  $d_{n,t}(u)$  is Lebesgue measurable. The arrays  $\{d_{n,t}(u), \tilde{d}_{n,t}, \varphi_{l_n}\}$  satisfy  $\sup_{u \in [0,1], t \geq 1} d_{n,t}(u) = O(k_n^{-1/2} (k_n/n)^{1/r})$  and  $\sup_{t \geq 1} \tilde{d}_{n,t} = O(k_n^{-1/2} (k_n/n)^{1/r})$  for some  $r \geq q$ , and  $\varphi_{l_n} \in [0, 1]$  and  $\varphi_{l_n} = o((k_n/n)^{1/q-1/r} l_n^{-\lambda})$ .

**Remark 11.** The FE-NED property (13) is NED applied to the tail event  $I(X_t > b_n/u)$ , so it only characterizes memory and heterogeneity in extremes and says nothing about non-extremes  $X_t \leq b_n/u \rightarrow \infty$ .

**Remark 12.** Any  $L_p$ -NED process  $\{X_t\}$  with tail (11) is  $L_q$ -FE-NED for any  $q \geq 2$  (see Lemmas A.8 and A.9 in Appendix A.1), thus properties (13)–(14) characterize mixing and geometrically ergodic processes, and at least linear and nonlinear distributed lags, covariance stationary linear and nonlinear GARCH, and stochastic volatility. But FE-NED also covers GARCH data with unit or explosive roots (Hill, 2008b).

**Remark 13.** Since  $d_{n,t}(u) \times \varphi_{l_n} = o(k_n^{-1/2} (k_n/n)^{1/q} l_n^{-\lambda})$ ,  $L_2$ -FE-NED implies

$$\frac{n}{k_n} E(I(X_t > b_n/u) - P(X_t > b_n/u | F_{n,t-l_n}^{t+l_n}))^2 = o(l_n^{-2\lambda}).$$

Now expand the quadratic, and exploit properties of regularly varying functions and  $(n/k_n)P(X_t > b_n) \sim 1$  to deduce  $\{X_t\}$  is  $L_2$ -FE-NED on  $\{F_{n,t}\}$  if and only if

$$\frac{n}{k_n} E[P(X_t > b_n/u | F_{n,t-l_n}^{t+l_n})^2] = u^\alpha + o(l_n^{-2\lambda}).$$

Thus, as  $l_n \rightarrow \infty$  the extreme event predictor  $P(X_t > b_n/u | F_{n,t-l_n}^{t+l_n})$  converges to the event  $I(X_t > b_n/u)$  in  $L_2$ -norm, which satisfies  $(n/k_n)E[I(X_t > b_n/u)] \rightarrow u^\alpha$ .

**Remark 14.** The analogue to (10') is

$$k_n^{-1/2} \left\| \sup_{|u_{1,j}-u_{2,j}| \leq \delta} I_{n,t}(u_1, u_2) - E \left[ \sup_{|u_{1,j}-u_{2,j}| \leq \delta} I_{n,t}(u_1, u_2) | F_{n,t-l_n}^{t+l_n} \right] \right\|_q \leq K \times \tilde{d}_{n,t} \times \delta^{1/q} \times \varphi_{l_n}, \quad j = 1 \dots k - 1. \tag{14'}$$

**Assumption 2.** (a)  $X_t$  is  $L_q$ -FE-NED on  $\{F_{n,t}\}$  defined in (7) with size  $\lambda = \frac{1}{2}$ , and Lebesgue integrable constants  $d_{n,t}(u)$ ,  $\int_0^1 u^{-1} d_{n,t}(u) du = O(k_n^{-1/2} (k_n/n)^{1/r})$ . The base  $\{\varepsilon_t\}$  is either F-uniform mixing with size  $r/2(r - 1)$ ,  $r \geq 2$ , or F-strong mixing base with size  $r/(r - 2)$ ,  $r > 2$ .

(b) For some bounded function  $1 \leq \kappa(\xi, \delta) \leq K$ ,  $\lim_{n \rightarrow \infty} |n(\xi + \delta)/n(\xi) - \kappa(\xi, \delta)| \rightarrow 0$  exists  $\forall \xi \in [\underline{\xi}, 1 - \delta]$ ,  $\delta \in [0, 1]$ .

If (13) holds for  $q = 2$  such that the event process  $U_{n,t}^*(u)$  is  $L_2$ -NED (8), then the exceedance process  $\{U_{n,t}\}$  is also  $L_2$ -NED. Thus (13) characterizes a primitive tail memory property.

**Lemma 3.2.** Under Assumption 2(a) with  $q = 2$ :

- (i)  $\{Y_{n,t}(c, u)\}$  is  $L_2$ -NED on  $\{F_{n,t}\}$  coefficients  $\varphi_{l_n}$  and  $O(k_n^{-1/2}(k_n/n)^{1/r})$ -constants. In particular  $\{U_{n,t}\}$  is  $L_2$ -NED on  $\{F_{n,t}\}$  with coefficients  $\varphi_{l_n}$  and constants  $\alpha^{-1} K \int_0^1 u^{-1} d_{n,t}(u) du$ .
- (ii)  $\{Y_{n,t}(c, u), F_{n,t}\}$  forms a zero mean  $L_2$ -mixingale array with size  $\frac{1}{2}$  and constants  $c_{n,t} = O(n^{-1/2})$ : there exists a sequence  $\{\psi_{l_n}\}$  of positive numbers,  $\psi_{l_n} = o(l_n^{-1/2})$ , such that for some  $\{l_n\}$ ,  $l_n \rightarrow \infty$  (cf. McLeish, 1975)

$$\begin{aligned} \|E[Y_{n,t}(c, u)|F_{n,t-l_n}]\|_2 &\leq c_{n,t}\psi_{l_n}, \\ \|Y_{n,t}(c, u) - E[Y_{n,t}(c, u)|F_{n,t+l_n}]\|_2 &\leq c_{n,t}\psi_{l_{n+1}}. \end{aligned}$$

The next result is an immediate consequence of Theorem 2.1 and Lemmas 3.1 and 3.2.

**Corollary 3.3.** Under Assumption 2 with  $q = 2$ ,  $\sum_{t=1}^{n(\xi)} Y_{n,t}(c, u) \Rightarrow Y(\xi, c, u)$  on  $D_2$ , where  $Y(\xi, u, c)$  is Gaussian with independent increments,  $P(Y(\xi, \cdot) \in C[\underline{\xi}, 1]) = 1$ , and  $E(\sum_{t=1}^n Y_{n,t}(c, u))^2 = O(1)$ . Further, if (14') holds then  $P(Y(\cdot, u) \in C[0, 1]^{k-1}) = 1$ .

**Remark.** A non-functional version of Corollary 3.3 follows under only the  $L_2$ -FE-NED property (13), similar to Corollary 2.2. See Hill (2005).

Corollary 3.3 and a Cramér–Wold device for  $D_2$ -valued processes (e.g. Phillips and Durlauf, 1986) deliver the joint weak limit

$$\sqrt{k_n} \begin{bmatrix} 1/k_n \sum_{t=1}^{n(\xi)} \{I(X_t > b_n/u) - P(X_t > b_n/u)\} \\ 1/k_n \sum_{t=1}^{n(\xi)} \{(\ln X_t/b_n)_+ - E[(\ln X_t/b_n)_+]\} \end{bmatrix} \Rightarrow \begin{bmatrix} U^*(\xi, u) \\ U(\xi) \end{bmatrix}.$$

The limit processes  $\{U(\xi)\}$  and  $\{U^*(\xi, u)\}$  are Gaussian with almost surely continuous sample paths along  $\xi$ , independent increments, and finite variance. Under (14') the sample paths of  $U^*(\xi, u)$  are almost surely continuous along  $u$ .

**4. FE-NED arrays and extant FCLT's**

We now discuss extant central limit theorems that cannot include the tail arrays  $\{U_{n,t}, U_{n,t}^*(u)\}$ . Davidson and de Jong (2000, Theorem 3.1) show  $\sum_{t=1}^{K_n(\xi)} X_{n,t} \Rightarrow X(\xi)$  for some sequence of increasing, right-continuous functions  $\{K_n(\xi)\}$ , where  $\{X_{n,t}\}$  is  $L_2$ -NED (8) on some  $\{F_{n,t}\}$  with size  $\lambda = \frac{1}{2}$  and constants  $d_{n,t}$ . They require some positive non-stochastic array  $\{c_{n,t}\}$  a sequence  $\{g_n\}$ ,  $g_n \rightarrow \infty$ ,  $g_n = o(n)$ , and  $M_{n,i} := \max_{(i-1)g_n+1 \leq t \leq ig_n} c_{n,t}$  such that  $\{X_{n,t}^2/c_{n,t}^2\}$  is  $L_r$ -bounded,  $r > 2$ , and

$$\max_{1 \leq i \leq r_{n+1}} M_{n,i} = o(g_n^{-1/2}), \quad \sum_{i=1}^{r_n(\xi)} M_{n,i}^2 = O(g_n^{-1}), \quad g_n = o(K_n(\xi)), \quad \text{where } r_n(\xi) = \left\lfloor \frac{K_n(\xi)}{g_n} \right\rfloor.$$

Their framework is quite ingenious for non-degenerate arrays and allows substantial memory and heterogeneity, but cannot be satisfied for tail arrays.

**Lemma 4.1 (Hill, 2007a).** Let  $K_n(\xi)/n \rightarrow a(\xi) > 0$  a finite constant function. There does not exist an array  $\{c_{n,t}\}$  that satisfies  $M_{n,i} = o(g_n^{-1/2})$  and  $\sum_{i=1}^{r_n(\xi)} M_{n,i}^2 = O(g_n^{-1})$  such that  $\{U_{n,t}/c_{n,t}\}$  or  $\{U_{n,t}^*(u)/c_{n,t}\}$  are  $L_r$ -bounded for all  $t$  and any  $r > 2$ .

**Remark.** The imposed properties on  $c_{n,t}$  are too restrictive simply because  $U_{n,t}$  and  $U_{n,t}^*(u)$  are asymptotically degenerate.

Wu and Min (2005, Theorems 1 and 3) establish invariance principles  $\sum_{t=1}^{[n\xi]} X_t / \|\sum_{t=1}^n X_t\|_2 \Rightarrow X(\xi)$ , for stationary distributed lags  $X_t = \sum_{i=0}^\infty \psi_i a_{t-i}$  where  $a_t = h(\varepsilon_t, \varepsilon_{t-1}, \dots)$  a Borel measurable function of zero mean iid shocks  $\varepsilon_t$ ,  $E(\varepsilon_t^2) < \infty$ . Under their Theorem 1, for example, coefficient summability conditions are enforced and  $a_t$  is  $L_p$ -weakly dependent of order 1,  $p > 2$  (Wu and Min, Eq. (4)). They conclude  $\|\sum_{t=1}^n X_t\|_2 / \sqrt{n}$  is slowly varying, a property that expedites their invariance principle (see below their Eq. (38)). Their Theorem 2 tackles  $\psi_i = l(i)/n^\beta$ ,  $\frac{1}{2} < \beta < 1$ , for slowly varying  $l(i)$ .

The simplest case is their Theorem 1 for iid  $X_t$  ( $\psi_i = 0 \forall i \geq 1$ ,  $a_t = \varepsilon_t$ ) in which case the summability conditions automatically hold. Define  $S_k := \sum_{t=1}^k [I(X_t > b_n e^u) - P(X_t > b_n e^u)]$  and  $\sigma_n^2 := E[S_n^2]$ .

**Lemma 4.2** (Hill, 2007a). If  $X_t$  is iid with distribution tail (11) then  $\sigma_n^2 \rightarrow \infty$ ,  $\|E(S_n|X_0)\|_2 = 0$ , and  $\sigma_n^2/n$  is regularly varying with index 1. Further,  $\sigma_n^2/k_n$  is slowly varying.

**Remark.** Thus  $l(n) := \sigma_n^2/n$  is not slowly varying  $l(\omega n)/l(n) \rightarrow 1 \forall \omega > 0$ , but regularly varying  $l(\omega n)/l(n) \rightarrow \lambda^{-1}$  simply because  $P(X_t > b_n e^u) \sim e^{-\alpha u} (k_n/n)$  under (11) is degenerate. Wu and Min's use of slowly varying  $\sigma_n^2/n$  is supported by arguments in Wu and Woodroffe (2004, Lemma 1 and Theorem 1). But the key Lemma 1 of Wu and Woodroffe (2004) is wrong as stated, at least for tail arrays:  $\sigma_n^2 \rightarrow \infty$  and  $\|E(S_n|X_0)\|_2 = o(\sigma_n^2)$  are trivial, but clearly  $\sigma_n^2/n$  is not slowly varying. The correct scale, however, is  $1/k_n$  as shown in Corollary 3.3.

**5. Applications**

Assume  $X_t \geq 0$  a.s. has tail (11). In this section we develop a limit theory for the intermediate tail quantile function  $q_{[k_n \xi]}$ , the Hill-estimator  $\hat{\alpha}_{k_n}^{-1}(\xi)$  in (5), and the tail dependence estimator  $\hat{\rho}_{n(\xi)}(h, u)$  in (6). Let  $n(\xi) = [n\xi]$  for brevity.

5.1. Tail index estimation and the tail quantile process

We must restrict the regularly varying class (11) in order to expedite asymptotic normality in the manner of Goldie and Smith (1987, SR1). See, also, Hsing (1991) and Hill (2005).

**Assumption 3.** For some positive measurable function  $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$

$$L(\lambda x)/L(x) - 1 = O(g(x)) \quad \text{as } x \rightarrow \infty. \tag{15}$$

$g$  has bounded increase: there exists  $0 < D, z_0 < \infty$  and  $\tau \leq 0$  such that  $g(\lambda z)/g(z) \leq D\lambda^\tau$  some for  $\lambda \geq 1, z \geq z_0$ .  $\{k_n\}$  and  $g(\cdot)$  satisfy

$$\sqrt{k_n}g(b_n) \rightarrow 0. \tag{16}$$

**Remark.** Tails satisfying (11), (15) and (16) include  $\bar{F}(x) = cx^{-\alpha}(1 + O((\ln x)^{-\theta}))$  and  $\bar{F}(x) = cx^{-\alpha}(1 + O(x^{-\theta}))$ . The latter characterize GARCH processes (Basrak et al., 2002) and (16) holds in this case if  $k_n \sim n^\delta, \delta < 2\theta/(2\theta + \alpha)$  (Haeusler and Teugels, 1985).

Define

$$v_{1,n}^2(\xi) := E \left( k_n^{1/2} (\hat{\alpha}_{k_n}^{-1}(\xi) - \alpha^{-1}) \right)^2 \quad \text{and} \quad v_{2,n}^2(\xi) := E \left( \sum_{t=1}^{[n\xi]} U_{n,t}^* (u^{[k_n \xi]^{-1/2}}) \right)^2.$$

**Theorem 5.1.** Under Assumption 2(a) with  $q=2$  and Assumption 3 there exist Brownian motion laws  $W_1(\xi)$  and  $W_2(\xi)$  with variances  $\lim_{n \rightarrow \infty} v_{i,n}^2(\xi) < \infty, i = 1, 2$ , such that

$$k_n^{1/2} \{ \hat{\alpha}_{k_n}^{-1}(\xi) - \alpha^{-1} \} \Rightarrow W_1(\xi) \quad \text{and} \quad k_n^{1/2} \{ \ln X_{[k_n \xi]} / b_n \} \Rightarrow W_2(\xi).$$

**Remark.** Evidently the above result exists under the weakest conditions available in the literature since it covers mixing, geometrically ergodic,  $L_p$ -NED and  $L_2$ -FE-NED data. Non-functional theory exists for iid and strong mixing processes (Hall, 1982; Hall and Welsh, 1985; Hsing, 1991; Drees et al., 2004), and for data with NED extreme events (Hill, 2005). See Hill (2005) for a complete literature review and consistent kernel variance estimator of  $v_{1,n}^2(1)$ .

5.2. Tail dependence

Now consider a bivariate process  $\{X_{1,t}, X_{2,t}\}$ , where each  $X_{i,t}$  has tail (11) with index  $\alpha_i > 0$ . We want to estimate the joint/marginal tail probability discrepancy:

$$\rho_n(i, u) := \frac{n}{k_n} [P_{n,t,h}(u_1, u_2) - P_{1,n,t}(u_1)P_{2,n,t}(u_2)],$$

where  $P_{n,t,h}(u_1, u_2) := P(X_{1,t-h} > b_{1,n}/u_1, X_{2,t} > b_{2,n}/u_2)$  and  $P_{i,n,t}(u_i) := P(X_{i,t} > b_{i,n}/u_i)$ . See Hill (2007b) for comparisons of  $\rho_n(i, u)$  with tail index and copula based notions of tail dependence. The estimator  $\hat{\rho}_n(i, u)$  is defined in (6). For arbitrary  $\lambda \in \mathbb{R}^h, h \geq 1, \lambda' \lambda = 1$ , write

$$Z_{n,t}(\lambda, u, h) := \sqrt{k_n} \sum_{i=1}^h \lambda_i \times U_{1,n,t-i}^*(u_1) \times U_{2,n,t}^*(u_2)$$

so that  $\sum_{t=1}^{[n\xi]} Z_{n,t}(\lambda, u, h) = \sqrt{k_n} \sum_{i=1}^h \lambda_i \hat{\rho}_{[n\xi]}(i, u)$ .

**Lemma 5.2 (Hill, 2007a).** For all  $r \geq 1$  and finite  $h \geq 1$ ,  $\|Z_{n,t}(\lambda, u, h)\|_r = O(k_n^{-1/2}(k_n/n)^{1/r}) = O(n^{-a(r)})$ ,  $a(1) > \frac{1}{2}$ , where  $a(2) = \frac{1}{2}$ , and  $2a(2r) > a(r)$ . Further,  $a(r) > 1/r$  for all  $r > 2$ .

**Lemma 5.3 (Hill, 2007a).** Let  $\{X_{1,t}, X_{2,t}\}$  satisfy Assumption 2(a) with  $q=4$ . Then  $\{Z_{n,t}(\lambda, u, h)\}$  is  $L_2$ -NED on  $\{F_{n,t}\}$  with constants  $d_{n,t}(\lambda u)$ ,  $\sup_{u \in [0,1]^2, \lambda' \lambda = 1, t \geq 1} d_{n,t}(\lambda, u) = O(k_n^{-1/2}(k_n/n)^{1/r})$  and coefficients  $\varphi_{l_n} = o((k_n/n)^{1/2-1/r} l_n^{-1/2})$ . Moreover,  $E(Z_{n,t}(\lambda, u, h))^2 = O(1)$ .

By Lemmas 5.2 and 5.3 the conditions of Corollary 3.3 hold for  $\{Z_{n,t}(\lambda, u, h)\}$ . A Cramér–Wold device therefore suffices to prove the following claim. Write  $\hat{\rho}_{[n\xi]}^{(h)}(u) := [\hat{\rho}_{[n\xi]}^{(h)}(1, u), \dots, \hat{\rho}_{[n\xi]}^{(h)}(h, u)]'$  and  $\rho_{[n\xi]}^{(h)}(u) := [\rho_{[n\xi]}^{(h)}(1, u), \dots, \rho_{[n\xi]}^{(h)}(h, u)]'$ .

**Theorem 5.4.** If  $\{X_{1,t}, X_{2,t}\}$  satisfy Assumption 2(a) with  $q = 4$  then

$$\sqrt{k_n} \{ \hat{\rho}_{[n\xi]}^{(h)}(u) - \rho_{[n\xi]}^{(h)}(u) \} \Rightarrow W(\xi, u)$$

a Gaussian element of the  $h$ -vector space  $D_h$  with almost surely continuous sample paths along  $\xi$ , independent increments, and  $\|W(\xi, u)\|_2 < \infty$ . If each  $\{X_{i,t}\}$  satisfies (14') then  $P(W(\cdot, u) \in C[0, 1]^2) = 1$ .

**Remark 15.** Notice that although the bivariate joint tail  $P_{n,t,h}(u_1, u_2)$  is being estimated, we do not require a model nor any assumptions concerning the joint tail. Further, we do not require any restrictions on non-extremes. Compare this with the marginal iid and joint tail shape restrictions typically enforced in the literature (e.g. Ledford and Tawn, 1997; Schmidt and Stadtmüller, 2006; Klüppelberg et al., 2007).

**Remark 16.** In order for  $\hat{\rho}_{[n\xi]}^{(h)}(u)$  to be usable in practice, either arbitrary choices of the threshold sequences  $\{b_{i,n}\}$  are required, or a consistent plug-in, e.g.  $X_{i,(k_n+1)}$ . See Hill (2007b, Lemma A.2) for a proof that use of  $X_{i,(k_n+1)}$  does not affect the non-functional distribution limit of  $\hat{\rho}_{[n\xi]}^{(h)}(u)$  under Assumption 2(a). It is straightforward to show the result directly carries over to the functional limit case.

**Acknowledgment**

I would like to thank an anonymous referee for helpful comments.

**Appendix A.**

*A.1. Proofs of main results*

The proofs require the following notation. Define sequences  $g_n, l_n$  and  $r_n(\cdot)$  as follows:  $1 \leq g_n \rightarrow \infty$  and

$$1 \leq g_n = o(n(\xi)), \quad 1 \leq l_n \leq g_n - 1 \leq n(\xi) - 1, \tag{17}$$

$$l_n/g_n \rightarrow 0, \quad l_n \rightarrow \infty, \quad r_n(\delta) = \lceil n(\min\{\delta, 1\})/g_n \rceil \text{ for } \delta > 0,$$

and assume

$$g_n = o\left(n^{\min\{[2a(4)+a(2)-1]/2, 2a(2r)-a(r)\}}\right).$$

The latter is always possible as long as  $\min\{[2a(4) + a(2) - 1]/2, 2a(2r) - a(r)\} > 0$ , cf. Assumption 1(a), since  $2a(4) + a(2) > 1$  is implied by  $a(2) = \frac{1}{2}$  and  $2a(2r) > a(r)$ . For the array of  $\sigma$ -fields  $\{F_{n,t}\}$  defined by (7) construct the sub-field

$$\tilde{F}_{n,i} := \sigma\left(\bigcup_{\tau \leq ig_n} F_{n,\tau}\right) : i = 1 \dots r_n(\xi). \tag{18}$$

Now define blocks

$$Z_{n,i}(u) := \sum_{t=(i-1)g_n+l_n+1}^{ig_n} X_{n,t}(u). \tag{19}$$

For any point  $(\zeta, u)$  or sequence of points  $\{(\zeta_j, u_j)\}_{j=1}^h$ , and  $\pi \in \mathbb{R}^h, \pi' \pi = 1$ , define

$$\begin{aligned} W_{n,i}(u) &:= E[Z_{n,i}(u)|\tilde{F}_{n,i}] - E[Z_{n,i}(u)|\tilde{F}_{n,i-1}], \\ W_n(\zeta, u) &:= \sum_{i=1}^{r_n(\zeta)} W_{n,i}(u), \\ \tilde{W}_{n,i}(u, \pi) &:= \sum_{j=l}^h \pi_j W_{n,i}(u_j), \quad i = r_n(\zeta_{l-1}) + 1, \dots, r_n(\zeta_l), \quad l = 1 \dots h. \end{aligned} \tag{20}$$

Notice  $\{W_{n,i}(u), \tilde{F}_{n,i}\}$  forms a martingale difference array. The main result, Theorem 2.1, is proven by verifying the conditions of Lemma A.1 which shows  $X_n(\zeta, u) = W_n(\zeta, u) + o_p(1)$  and delivers a central limit theorem for  $W_n(\zeta, u)$ . This approach was exploited in Davidson (1992), de Jong (1997) and Davidson and de Jong (2000) for non-degenerate, and therefore non-tail  $L_2$ -NED arrays.

**Proof of Theorem 2.1.** Step 1 (weak convergence): Under Assumption 1(b), Lemma A.4(i) in Appendix A.2 shows  $\{X_{n,t}(u), F_{n,t}\}$  forms an  $L_2$ -mixingale array with constants  $c_{n,t}(u), \sup_{u \in [0,1]^{k-1}, t \geq 1} c_{n,t}(u) = O(n^{-1/2})$ , and coefficients  $\psi_{l_n} = o(l_n^{-1/2})$  of size  $\frac{1}{2}$ , hence  $E(X_n(\zeta, u)^2) = O(1)$  (McLeish, 1975).<sup>2</sup> We will show the conditions of the Lemma A.1(ii) functional central limit theorem apply. First, the  $r$ th-moment index  $a(r)$  satisfies  $2a(4) + a(2) > 1$  since  $a(2) = 1/2$  and  $2a(2r) > a(r)$  under Assumption 1(a).

Condition (a): Use McLeish’s (1975) bound to deduce

$$E \left( \sum_{i=1}^{r_n(\zeta)} \sum_{t=(i-1)g_n+1}^{(i-1)g_n+l_n} X_{n,t}(u)^2 \right) = O \left( \sum_{i=1}^{r_n(\zeta)} \sum_{t=(i-1)g_n+1}^{(i-1)g_n+l_n} c_{n,t}^2(u) \right) = O(r_n(\zeta)l_n/n) = O(l_n/g_n) = o(1),$$

hence  $\sum_{i=1}^{r_n(\zeta)} \sum_{t=(i-1)g_n+1}^{(i-1)g_n+l_n} X_{n,t}(u) \xrightarrow{P} 0$  by Chebyshev’s inequality.

Condition (b): Define index sets

$$A_i(t) := \{t : (i-1)g_n + l_n + 1 \leq t \leq ig_n\} \quad \text{and} \quad A_{n,t} = \bigcup_{i=1}^{r_n(\zeta)} A_i(t).$$

Analogous to de Jong’s (1997, A.7–A.12) argument, because  $\{X_{n,t}(u), F_{n,t}\}$  forms an  $L_2$ -mixingale array with size  $\lambda = \frac{1}{2}$ , for  $t \in A_{n,t}$  the array  $\{E[X_{n,t}(u)|\tilde{F}_{n,i-1}], F_{n,t}\}$  is an  $L_2$ -mixingale with constants  $c_{n,t}(u)\psi_{l_n}^\eta$  and coefficients  $\psi_{l_n}^{1-\eta} = o(l_n^{-1/2})$  for some sufficiently tiny  $\eta > 0$ . McLeish’s (1975) bound now gives

$$\begin{aligned} E \left( \sum_{i=1}^{r_n(\zeta)} E[Z_{n,i}(u)|\tilde{F}_{n,i-1}] \right)^2 &= O \left( \sum_{i=1}^{r_n(\zeta)} \sum_{t=(i-1)g_n+l_n+1}^{ig_n} c_{n,t}^2(u)\psi_{l_n}^{2\eta} \right) \\ &= O(r_n(\zeta)g_n n^{-1} l_n^{-\eta}) = O(l_n^{-\eta}) = o(1). \end{aligned}$$

Condition (c): The proof mimics (b).

Condition (d): Compactly write for any finite sequence of points  $\{(\zeta_j, u_j)\}_{j=1}^h$

$$Z_{n,i}(j) = Z_{n,i}(u_j) \quad \text{and} \quad W_{n,i}(j) = W_{n,i}(u_j),$$

and note from (20) we can always write

$$\sum_{i=1}^{r(\zeta_h)} \tilde{W}_{n,i}^2(u, \pi) = \sum_{l=1}^h \sum_{i=r_n(\zeta_{l-1})+1}^{r_n(\zeta_l)} \left( \sum_{j=l}^h \pi_j W_{n,i}(j) \right)^2.$$

<sup>2</sup> By Theorem 1.6 of McLeish (1975), if  $\{y_{n,t}, F_{n,t}\}$  forms an  $L_2$ -mixingale array with size  $\frac{1}{2}$  and coefficients  $\{c_{n,t}\}$  then  $E(\sum_{t=1}^n y_{n,t})^2 = O(\sum_{t=1}^n c_{n,t}^2)$ .

Analogous to arguments in de Jong (1997, A.13–A.17), it follows

$$\begin{aligned} & \sup_{\pi' \pi=1} \left\| \sum_{l=1}^h \sum_{i=r_n(\xi_{l-1})+1}^{r_n(\xi_l)} \left( \sum_{j=l}^h \pi_j Z_{n,i}(j) \right)^2 - \sum_{i=1}^{r_n(\xi_h)} \tilde{W}_{n,i}^2(u, \pi) \right\|_1 \\ &= \sup_{\pi' \pi=1} \left\| \sum_{l=1}^h \sum_{i=r_n(\xi_{l-1})+1}^{r_n(\xi_l)} \left[ \left( \sum_{j=l}^h \pi_j Z_{n,i}(j) \right)^2 - \left( \sum_{j=l}^h \pi_j W_{n,i}(j) \right)^2 \right] \right\|_1 \\ &\leq K \sum_{l=1}^h \sum_{j_1, j_2=l}^h \sum_{i=r_n(\xi_{l-1})+1}^{r_n(\xi_l)} \|Z_{n,i}(j_1) - W_{n,i}(j_1)\|_2 \times \|Z_{n,i}(j_2)\|_2 \\ &= O \left( \sum_{l=1}^h \sum_{i=r_n(\xi_{l-1})+1}^{r_n(\xi_l)} \left( \sum_{t \in A_i(t)} c_{n,t}^2 \psi_{l_n}^{2\eta} \right)^{1/2} \left( \sum_{t \in A_i(t)} c_{n,t}^2 \right)^{1/2} \right) \\ &= O \left( \sum_{l=1}^h r_n(\xi_l) \times [g_n n^{-1} l_n^{-\eta}]^{1/2} \times [g_n n^{-1}]^{1/2} \right) = O(l_n^{-\eta/2}) = o(1). \end{aligned}$$

Now use Assumption 1(a, b) and Lemma A.4(iii) to deduce  $\sup_{\pi' \pi=1} |\sum_{i=1}^{r_n(\xi_h)} \tilde{W}_{n,i}^2(u, \pi) - 1| \xrightarrow{P} 0$ .

Condition (e): For any  $\xi, \delta \in [\underline{\xi}, 1]$  and  $u \in [0, 1]^{k-1}$ , and some integer sequence  $\{r_n^*(\xi)\}$  satisfying

$$0 \leq [\delta r_n^*(\xi)] \leq r_n(\xi + \delta) - r_n(\xi),$$

use Minkowski's inequality to deduce

$$\begin{aligned} \left\| \sum_{i=1}^{[\delta r_n^*(\xi)]} W_{n,i}(u) \right\|_2 &\leq \left\| \sum_{i=1}^{[\delta r_n^*(\xi)]} \sum_{t=(i-1)g_n+l_n+1}^{ig_n} X_{n,t}(u) \right\|_2 \\ &+ \left\| \sum_{i=1}^{[\delta r_n^*(\xi)]} \sum_{t=(i-1)g_n+l_n+1}^{ig_n} (X_{n,t}(u) - E[X_{n,t}(u)|\tilde{F}_{n,i}]) \right\|_2 \\ &+ \left\| \sum_{i=1}^{[\delta r_n^*(\xi)]} \sum_{t=(i-1)g_n+l_n+1}^{ig_n} E[X_{n,t}(u)|\tilde{F}_{n,i-1}] \right\|_2. \end{aligned} \tag{21}$$

Under the maintained assumptions  $\{X_{n,t}(u), F_{n,t}\}$  forms an  $L_2$ -mixingale array with size  $\lambda = \frac{1}{2}$  and constants  $c_{n,t}(u) = O(n^{-1/2})$ . Similarly, for each  $t \in A_{n,t}$ ,  $\{E[X_{n,t}(u)|\tilde{F}_{n,i-1}], F_{n,t}\}$  and  $\{X_{n,t}(u) - E[X_{n,t}(u)|\tilde{F}_{n,i}], F_{n,t}\}$  form  $L_2$ -mixingale sequences with size  $\lambda = \frac{1}{2}$  and constants  $c_{n,t}(u)\psi_{n,l_n}^\eta$  for infinitesimal  $\eta > 0$  (de Jong, 1997). Apply McLeish's (1975) bound to each right-hand side term in (21) and note  $\psi_{n,l_n}^\eta = O(l_n^{-\eta/2}) = o(1)$  to obtain

$$\begin{aligned} E \left( \sum_{i=1}^{[\delta r_n^*(\xi)]} W_{n,i}(u) \right)^2 &= O \left( \sum_{i=1}^{[\delta r_n^*(\xi)]} \sum_{t=(i-1)g_n+l_n+1}^{ig_n} c_{n,t}^2(u) \right) \\ &= \delta \times O(r_n^*(\xi)g_n/n). \end{aligned}$$

Now, by Assumption 1(d) and the construction of  $r_n^*(\xi)$  there exists a finite mapping  $\kappa(\xi, \delta) \geq 1$  satisfying

$$\frac{\delta r_n^*(\xi)g_n}{n} \leq \frac{\delta r_n^*(\xi)g_n}{n(\xi)} \leq \left( \frac{r_n(\xi + \delta)}{r_n(\xi)} - 1 \right) \times (1 + o(1)) \rightarrow \kappa(\xi, \delta) - 1 < \infty.$$

But if  $\delta r_n^*(\xi)g_n/n$  is bounded then so is  $r_n^*(\xi)g_n/n$  since  $\delta \in [0, 1]$ . The proof is complete by exploiting the martingale difference property of  $\{W_{n,i}(u), \tilde{F}_{n,i}\}$ :

$$\sum_{i=1}^{[\delta r_n^*(\xi)]} E(W_{n,i}(u))^2 = E \left( \sum_{i=1}^{[\delta r_n^*(\xi)]} W_{n,i}(u) \right)^2 \leq \delta \times \kappa \quad \text{for large } n \text{ and some } \kappa \geq 1.$$

Condition (f): For any  $u \in [0, 1]^{k-1}$  and  $u_{j,1}, u_{j,2} \in [0, 1], j = 1 \dots k - 1$ , write

$$\Delta X_{n,t}(u_{j,1}, u_{j,2}) := X_{n,t}(\{u_i\}_{i \neq j}, u_{j,1}) - X_{n,t}(\{u_i\}_{i \neq j}, u_{j,2}).$$

The F-NED property (10) and Lemma A.4(ii) imply  $\{\Delta X_{n,t}(u_{j,1}, u_{j,2}), F_{n,t}\}$  forms an  $L_2$ -mixingale array with size  $\frac{1}{2}$  and constants  $\tilde{c}_{n,t} \times \max_{1 \leq i \leq k-1} |u_{2,i} - u_{1,i}|^{1/2}, \sup_{t \geq 1} \tilde{c}_{n,t} = O(n^{-1/2})$ . Now mimic (21) and the subsequent argument to obtain

$$\begin{aligned} \|W_n(\cdot, u_{j,1}) - W_n(\cdot, u_{j,2})\|_2 &\leq \left\| \sum_{i=1}^{r_n(\xi)} \sum_{t=(i-1)g_n+1}^{ig_n} \Delta X_{n,t}(u_{j,1}, u_{j,2}) \right\|_2 \\ &\quad + \left\| \sum_{i=1}^{r_n(\xi)} \sum_{t=(i-1)g_n+1}^{ig_n} (\Delta X_{n,t}(u_{j,1}, u_{j,2}) - E[\Delta X_{n,t}(u_{j,1}, u_{j,2}) | \tilde{F}_{n,i}]) \right\|_2 \\ &\quad + \left\| \sum_{i=1}^{r_n(\xi)} \sum_{t=(i-1)g_n+1}^{ig_n} E[\Delta X_{n,t}(u_{j,1}, u_{j,2}) | \tilde{F}_{n,i-1}] \right\|_2 \\ &= O \left( \left( \sum_{i=1}^{r_n(\xi)} \sum_{t=(i-1)g_n+1}^{ig_n} \tilde{c}_{n,t}^2 \right)^{1/2} \times |u_{j,1} - u_{j,2}|^{1/2} \right) \\ &= |u_{j,1} - u_{j,2}|^{1/2} \times O((r_n(\xi)g_n/n)^{1/2}) \leq K \times |u_{j,1} - u_{j,2}|^{1/2}. \end{aligned}$$

Since the right-hand side is not a function of  $\xi$  the inequality is uniform in  $\xi \in [\underline{\xi}, 1]$ .

Condition (g): Use  $\Delta X_{n,t}(u_{j,1}, u_{j,2})$  from above and define

$$\begin{aligned} Z_{n,t,j}(\delta) &:= \sup_{u_{1,j} \leq u_{2,j} \leq u_{1,j} + \delta} |\Delta X_{n,t}(u_{1,j}, u_{2,j})|, \\ W_{n,t,j}(\delta) &:= E[Z_{n,t,j}(\delta) | \tilde{F}_{n,i}] - E[Z_{n,t,j}(\delta) | \tilde{F}_{n,i-1}]. \end{aligned}$$

Property (10') implies  $\{Z_{n,t,j}(\delta)\}$  is  $L_2$ -NED on  $\{F_{n,t}\}$  with constants  $K\delta^{1/2}\tilde{d}_{n,t}, \sup_{t \geq 1} \tilde{d}_{n,t} = O(n^{-a(2)})$ , and coefficients  $\varphi_{ln} = o(n^{a(r)-a(2)}l_n^{-\lambda})$ , where  $a(2) = \frac{1}{2}$  under Assumption 1(a). An argument identical to the proof of Lemma A.4(ii) reveals  $\{Z_{n,t,j}(\delta), F_{n,t}\}$  forms an  $L_2$ -mixingale array with size  $\frac{1}{2}$  and constants  $\tilde{c}_{n,t} = \delta^{1/2} \times O(n^{-1/2})$ . It is similarly easy to show  $\{W_{n,t,j}(\delta), F_{n,t}\}$  forms an  $L_2$ -mixingale array with size  $\frac{1}{2}$  and constants  $\tilde{e}_{n,t} = \delta^{1/2} \times O(n^{-1/2})$  (cf. de Jong, 1997). McLeish's (1975) bound again delivers

$$\begin{aligned} E \left( \sup_{u_{1,j} \leq u_{2,j} \leq u_{1,j} + \delta} |W_n(\cdot, u_{1,j}) - W_n(\cdot, u_{2,j})| \right)^2 &\leq E \left( \sum_{i=1}^{r_n(\xi)} \sum_{t=(i-1)g_n+1}^{ig_n} W_{n,t,j}(\delta)^2 \right) \\ &= O \left( \sum_{t=1}^n \tilde{e}_{n,t}^2 \right) \leq K\delta. \end{aligned}$$

Step 2 (increments): For any  $0 < \xi_k < \xi_l < 1$  and  $u_k, u_l \in [0, 1]^{k-1}$  we need to show

$$E[\{X_n(\xi_k, u_k) - X_n(\xi_{k-1}, u_{k-1})\} \times \{X_n(\xi_l, u_l) - X_n(\xi_{l-1}, u_{l-1})\}] = o_p(1).$$

First, decompose the increment  $X_n(\xi_k, u_k) - X_n(\xi_{k-1}, u_{k-1})$ :

$$\begin{aligned} X_n(\xi_k, u_k) - X_n(\xi_{k-1}, u_{k-1}) &= [X_n(\xi_k, u_k) - X_n(\xi_{k-1}, u_k)] + [X_n(\xi_{k-1}, u_k) - X_n(\xi_{k-1}, u_{k-1})] \\ &= A_{n,k} + B_{n,k}. \end{aligned}$$

Analogous to arguments in Davidson and de Jong (2000, pp. 635–636 and Lemma A.3), use McLeish's (1975) bound, the properties  $\sup_{u \in [0,1]^{k-1}, t \geq 1} c_{n,t}(u) = O(n^{-1/2}), n(\xi_l) - n(\xi_k) \rightarrow \infty, l > k$ , and  $n(1) \leq n$  to deduce for arbitrary  $\delta \in [0, 1]$

$$\begin{aligned} |E[A_{n,k}A_{n,l}]| &\leq \left\| \sum_{t=n(\xi_{k-1})+1}^{n(\xi_k)} X_{n,t}(u_k) \right\|_2 \times \left\| \sum_{t=n(\xi_{l-1})+1}^{n(\xi_{l-1}+\delta)} X_{n,t}(u_l) \right\|_2 + \sum_{t,s=1}^{n(1)} E[|X_{n,t}(u_k)X_{n,s}(u_l)| \times I(|s-t| \geq n(\xi_{l-1}+\delta) - n(\xi_k))] \\ &= O \left( \sup_{0 < \xi_i < 1} [(n(\xi_i) - n(\xi_{i-1}))/n] \right) + o_p(1) = o_p(1). \end{aligned}$$

Similarly, since under Lemma A.4(ii)  $\{X_{n,t}(u_l)X_{n,t}(u_{l-1}), F_t\}$  forms an  $L_2$ -mixingale array with size  $\frac{1}{2}$  and constants proportional to  $\tilde{c}_{n,t}$ ,  $\sup_{t \geq 1} \tilde{c}_{n,t} = O(n^{-1/2})$ , for arbitrary  $\delta > 0$

$$\begin{aligned} |E[A_{n,k}B_{n,l}]| &\leq \left\| \sum_{t=n(\xi_{k-1})+1}^{n(\xi_{k-1}+2\delta)} X_{n,t}(u_k) \right\|_2 \times \left\| \sum_{t=1}^{n(\xi_{l-1})} (X_{n,t}(u_l) - X_{n,t}(u_{l-1})) \right\|_2 \\ &+ \left\| \sum_{t=n(\xi_{k-1}+2\delta)+1}^{n(\xi_k)} X_{n,t}(u_k) \right\|_2 \times \left\| \sum_{t=n(\xi_{k-1}+\delta)+1}^{n(\xi_{l-1})} (X_{n,t}(u_l) - X_{n,t}(u_{l-1})) \right\|_2 \\ &+ \sum_{s,t=1}^{n(1)} E[|X_{n,t}(u_k)(X_{n,s}(u_l) - X_{n,s}(u_{l-1}))| \times I(|s-t| \geq n(\xi_{k-1} + 2\delta) - n(\xi_{k-1} + \delta))] \\ &= O\left(\left(\sum_{t=n(\xi_{k-1})+1}^{n(\xi_{k-1}+2\delta)} c_{n,t}^2(u_k)\right)^{1/2} \left(\sum_{t=1}^{n(\xi_{l-1})} \tilde{c}_{n,t}^2\right)^{1/2}\right) \\ &+ O\left(\left(\sum_{t=n(\xi_{k-1}+2\delta)+1}^{n(\xi_k)} c_{n,t}^2(u_k)\right)^{1/2} \left(\sum_{t=n(\xi_{k-1}+\delta)+1}^{n(\xi_{l-1})} \tilde{c}_{n,t}^2\right)^{1/2}\right) + o_p(1) \\ &= o_p(1), \end{aligned}$$

because  $c_{n,t}(u)$  and  $\tilde{c}_{n,t}$  are uniformly  $O(n^{-1/2})$ ,  $0 < \xi_k < \xi_l < 1$  and  $n(1) \leq n$ .

Finally,  $0 < \xi_k < \xi_l < 1$ ,  $n(\xi_k) < n(\xi_l) < n$  and the  $L_2$ -mixingale property of  $X_{n,t}(u_l) - X_{n,t}(u_{l-1})$  imply

$$\begin{aligned} |E[B_{n,k}B_{n,l}]| &\leq \left\| \sum_{t=1}^{n(\xi_{k-1})} (X_{n,t}(u_k) - X_{n,t}(u_{k-1})) \right\|_2 \times \left\| \sum_{t=1}^{n(\xi_{l-1})} (X_{n,t}(u_l) - X_{n,t}(u_{l-1})) \right\|_2 \\ &= O\left(\left(\sum_{t=1}^{n(\xi_{k-1})} \tilde{c}_{n,t}^2\right)^{1/2} \left(\sum_{t=1}^{n(\xi_{l-1})} \tilde{c}_{n,t}^2\right)^{1/2}\right) \\ &= O((n(\xi_{k-1})/n)^{1/2}(n(\xi_{l-1})/n)^{1/2}) = o(1). \quad \square \end{aligned}$$

**Proof of Corollary 2.2.** The claim follows from conditions (a)–(d) of Lemma A.1, which are verified in the proof of Theorem 2.1 under Assumption 1(a,b).  $\square$

**Proof of Lemma 3.1.** By Lemma A.2 in Hill (2005) for any  $r \geq 1$

$$\lim_{n \rightarrow \infty} (n/k_n)^{1/r} k_n^{1/2} \|U_{n,t}^*(u)\|_r \leq K \quad \text{and} \quad \lim_{n \rightarrow \infty} (n/k_n)^{1/r} k_n^{1/2} \|U_{n,t}\|_r \leq K.$$

Since  $k_n \rightarrow \infty$  as  $n \rightarrow \infty$  it follows  $k_n^{-(1/2-1)} n^{-1/2} > n^{-1/2}$ ,  $k_n^{-(1/2-1/2)} n^{-1/2} = n^{-1/2}$ , and  $k_n^{-(1/2-1/r)} n^{-1/r} < n^{-1/r} \forall r > 2$ . Hence for some  $a(r)$  each  $\{\|U_{n,t}\|_r, \|U_{n,t}^*(u)\|_r\} = O(n^{-a(r)})$ , where  $a(1) > 1/2$ ,  $a(2) = \frac{1}{2}$ , and  $a(r) > 1/r \forall r > 2$ . Similarly,

$$\begin{aligned} O(n^{-2a(2r)}) &= O(k_n^{-(1/2-1/2r)} n^{-1/2r})^2 = O(k_n^{-(1-1/r)} n^{-1/r}) \\ &= O(k_n^{-(1/2-1/r)} n^{-1/r}) k_n^{-1/2} = O(n^{-a(r)}) \times k_n^{-1/2} \leq O(n^{-a(r)}), \end{aligned}$$

hence  $2a(2r) > a(r)$ .  $\square$

**Proof of Lemma 3.2.** The  $L_2$ -NED and  $L_2$ -mixingale claims follow from Lemma 3 of Hill (2005) and a change of variable  $e^{-v} = u$ ,  $v \geq 0$ . The bound  $E(\sum_{t=1}^n Y_{n,t}(c, u))^2 = O(1)$  follows from McLeish (1975).  $\square$

**Proof of Theorem 5.1.** The claim for  $\hat{\alpha}_{[n\xi]}^{-1}$  follows from Lemma A.7 and Corollary 3.3:

$$k_n^{1/2} \{ \hat{\alpha}_{k_n}^{-1}(\zeta) - \alpha^{-1} \} = \sum_{t=1}^{[n\xi]} \{ U_{n,t} - \alpha^{-1} U_{n,t}^*(u^{[k_n\xi]^{-1/2}}) \} + o_p(1) \Rightarrow W_1(\zeta).$$

Now consider  $X_{([k_n\xi])}$ . We infer from Theorem 2.4 of Hsing (1991) that under Assumption 3 for some process  $Z(\zeta)$  if  $\sum_{t=1}^{[n\xi]} U_{n,t}^*(u^{k_n^{-1/2}}) \Rightarrow Z(\zeta)$  then  $k_n^{1/2} \ln X_{([k_n\xi])}/b_n \Rightarrow Z(\zeta)$ . Corollary 3.3 completes the proof:  $\sum_{t=1}^{[n\xi]} U_{n,t}^*(u^{[k_n\xi]^{-1/2}}) \Rightarrow W_2(\zeta)$  for some Gaussian element  $W_2(\zeta)$  of  $D[\underline{\zeta}, 1]$  with variance  $\lim_{n \rightarrow \infty} v_{2,n}^2(\zeta) = \lim_{n \rightarrow \infty} E(\sum_{t=1}^{[n\xi]} U_{n,t}^*(u^{[k_n\xi]^{-1/2}}))^2 < \infty$ .  $\square$

A.2. Supporting lemmata

The following results exploit the constructions in (17)–(20). First, Lemma A.1 delivers an omnibus central limit theorem and functional central limit theorem for functional arrays  $X_{n,t}(u)$ .

**Lemma A.1.** Let  $\{X_{n,t}(u)\}$  be an  $L_r$ -functional tail array with  $r$ th-moment index  $a(r)$ ,  $2a(4) + a(2) > 1$ , and let  $\|X_n(\zeta, u)\|_2 = O(1)$ . Assume  $\{g_n\}$  in (17) satisfies  $g_n = o(n^{[2a(4)+a(2)-1]/2})$ . Consider the following conditions:

- (a)  $\sum_{i=1}^{r_n(\zeta)} \sum_{t=(i-1)g_n+1}^{(i-1)g_n+I_n} X_{n,t}(u) \xrightarrow{P} 0$ ,
- (b)  $\sum_{i=1}^{r_n(\zeta)} E[Z_{n,i}(u) | \tilde{F}_{n,i-1}] \xrightarrow{P} 0$ ,
- (c)  $\sum_{i=1}^{r_n(\zeta)} (Z_{n,i}(u) - E[Z_{n,i}(u) | \tilde{F}_{n,i}]) \xrightarrow{P} 0$ ,
- (d)  $\sup_{\pi' \pi=1} \left| \sum_{i=1}^{r_n(\zeta_n)} \tilde{W}_{n,i}^2(u, \pi) - 1 \right| \xrightarrow{P} 0$ ,
- (e)  $\sum_{i=1}^{[\delta r_n^*(\zeta)]} E(W_{n,i}(u))^2 \leq \delta \times \kappa, \quad \forall \delta \in [0, 1]$

for some  $\kappa \geq 1$  and some sequence  $\{r_n^*(\zeta)\}$  satisfying  $0 \leq [\delta \times r_n^*(\zeta)] \leq r_n(\zeta + \delta) - r_n(\zeta)$ ; and for each  $j=1 \dots k-1$  and any  $u_1, u_2 \in [0, 1]^{k-1}$

(f)  $\sup_{\zeta \in [\underline{\zeta}, 1]} \|W_n(\zeta, u_{1,j}) - W_n(\zeta, u_{2,j})\|_2 \leq K|u_{1,j} - u_{2,j}|^{1/2}$ ,

and for each  $\zeta \in [\underline{\zeta}, 1]$

(g)  $E \left( \sup_{u_{1,j} \leq u_{2,j} \leq u_{1,j} + \delta} |W_n(\zeta, u_{1,j}) - W_n(\zeta, u_{2,j})| \right)^2 \leq K\delta$ .

- (i) For any points  $\zeta \in (0, 1]$  and  $u \in [0, 1]^{k-1}$  under (a)–(d)  $X_n(\zeta, u) \xrightarrow{d} X(\zeta, u)$ , a zero mean Gaussian law with finite variance.
- (ii) Under (a)–(f)  $X_n(\zeta, u) \Rightarrow X(\zeta, u)$  a Gaussian element of  $D_k$  with covariance function  $E[X(\zeta_i, u_i)X(\zeta_j, u_j)]$  and  $P(X(\zeta, \cdot) \in C[\underline{\zeta}, 1]) = 1$ . If additionally (g) holds then  $P(X(\cdot, u) \in C[0, 1]^{k-1}) = 1$ .

**Remark.** Conditions (a)–(c) imply  $X_n(\zeta, u)$  is approximable by martingale differences  $\{W_{n,i}(u)\}$ . Condition (d) ensures convergence of the finite dimensional distributions of  $\{W_n(\zeta, u)\}$ , and conditions (e) and (f) ensure the sequence  $\{W_n(\zeta, u)\}$  is uniformly tight with respect to  $\zeta$  and  $u$ , respectively. Thus, (a)–(d) deliver a CLT for tail arrays  $X_n(\cdot, \cdot)$ , (a)–(e) an FCLT for  $X_n(\zeta, \cdot)$ , and (a)–(f) an FCLT for  $X_n(\zeta, u)$ .

**Lemma A.2.** Under conditions (a)–(d) of Lemma A.1,  $W_n(\zeta, u) \rightarrow W(\zeta, u)$  with respect to finite dimensional distributions, where  $W(\zeta, u)$  is Gaussian with covariance function  $E[W(\zeta_i, u_i)W(\zeta_j, u_j)]$ .

**Lemma A.3.** Under conditions (a)–(f) of Lemma A.1 the sequence  $\{W_n(\zeta, u)\}$  is uniformly tight on  $[\underline{\zeta}, 1] \times [0, 1]^{k-1}$ , and  $P(W(\zeta, \cdot) \in C[\underline{\zeta}, 1]) = 1$ . Under condition (g)  $P(W(\cdot, u) \in C[0, 1]^{k-1}) = 1$ .

**Lemma A.4.** (i) Under Assumption 1(a,b)  $\{X_{n,t}(u), F_{n,t}\}$  forms an  $L_2$ -mixingale array with coefficients  $\psi_{l_n} = o(l_n^{-1/2})$  and constants  $c_{n,t}(u)$ ,  $\sup_{u \in [0,1]^{k-1}, t \geq 1} c_{n,t}(u) = O(n^{-1/2})$ :

$$\begin{aligned} \|E[X_{n,t}(u)|F_{n,t-l_n}]\|_2 &\leq c_{n,t}(u)\psi_{l_n}, \\ \|X_{n,t}(u) - E[X_{n,t}(u)|F_{n,t+l}]\|_2 &\leq c_{n,t}(u)\psi_{l_{n+1}}. \end{aligned}$$

(ii) Write  $\Delta X_{n,t}(u_1, u_2) := X_{n,t}(u_1) - X_{n,t}(u_2)$ . Under Assumption 1(a,c)  $\{\Delta X_{n,t}(u_1, u_2), F_{n,t}\}$  forms an  $L_2$ -mixingale array with coefficients  $\psi_{l_n} = o(l_n^{-1/2})$  and constants  $\tilde{c}_{n,t} \times \max_{1 \leq j \leq k-1} |u_{2j} - u_{1j}|^{1/2}$ ,  $\sup_{t \geq 1} \tilde{c}_{n,t} = O(n^{-1/2})$ :

$$\begin{aligned} \|E[\Delta X_{n,t}(u_1, u_2)|F_{n,t-l_n}]\|_2 &\leq \left( \tilde{c}_{n,t} \max_{1 \leq j \leq k-1} |u_{2j} - u_{1j}|^{1/2} \right) \psi_{l_n}, \\ \|\Delta X_{n,t}(u_1, u_2) - E[\Delta X_{n,t}(u_1, u_2)|F_{n,t+l}]\|_2 &\leq \left( \tilde{c}_{n,t} \max_{1 \leq i \leq k-1} |u_{2,i} - u_{1,i}|^{1/2} \right) \psi_{l_{n+1}}. \end{aligned}$$

(iii) Let  $\{(\zeta_i, u_i)\}_{i=1}^h$  be arbitrary points on  $[\underline{\zeta}, 1] \times [0, 1]^{k-1}$ ,  $h \geq 1$ . Under Assumption 1(a,b)

$$\sup_{\pi' \neq \pi} \left| \sum_{l=1}^h \sum_{i=r_n(\zeta_{l-1})+1}^{r_n(\zeta_l)} \left( \sum_{j=l}^h \pi_j Z_{n,i}(u_j) \right)^2 - 1 \right| \rightarrow 0, \quad \forall \{(\zeta_j, u_j)\}_{j=1}^h.$$

**Lemma A.5** (de Jong, 1997, Lemma 4). For any  $L_2$ -mixingale array  $\{Y_{n,t}, F_{n,t}\}$  with size  $\lambda = \frac{1}{2}$  and constants  $c_{n,t}$ ,  $\sup_{t \geq 1} c_{n,t} = O(n^{-1/2})$

$$\lim_{n \rightarrow \infty} \left| \sum_{i=1}^{r_{n(1)}} \sum_{k=i+1}^{r_{n(1)}} \sum_{t=(i-1)g_n+l_n+1}^{ig_n} \sum_{s=(k-1)g_n+l_n+1}^{kg_n} E[Y_{n,s}Y_{n,t}] \right| = 0.$$

**Lemma A.6.** Under Assumption 1(a,b)  $\sum_{i=1}^{r_n(\zeta)} (Z_{n,i}^2(u) - E[Z_{n,i}^2(u)]) \rightarrow 0 \forall \zeta, u \in [\underline{\zeta}, 1] \times [0, 1]^{k-1}$ .

**Lemma A.7** (Hill, 2007a). Under the conditions of Theorem 5.1 for any  $u \in [0, 1]$

$$k_n^{1/2}(\hat{\alpha}_{k_n}^{-1}(\zeta) - \alpha^{-1}) = \sum_{t=1}^{\lfloor n\zeta \rfloor} (U_{n,t} - \alpha^{-1} U_{n,t}^*(u^{\lfloor k_n \zeta \rfloor^{-1/2}})) + o_p(1).$$

**Lemma A.8** (Hill, 2007a). Let  $\{X_t\}$  be  $L_p$ -NED,  $p \in [0, 2]$ , on  $\{F_{n,t}\}$  with constants  $d_t$ ,  $\sup_{t \geq 1} d_t < \infty$ , and coefficients  $\vartheta_{l_n}$  of size  $\lambda > 0$ . Then  $\{X_t\}$  satisfies the  $L_q$ -FE-NED property (13) for any  $q \geq 2$  and some displacement sequence  $\{l_n\}$ ,  $l_n \rightarrow \infty$ , with constants  $d_{n,t}(u)$  and coefficients  $\varphi_{l_n}$  of size  $\lambda \times \min\{p, 1\}/(2q)$ . In particular,  $d_{n,t}(u)$  is Lebesgue measurable,  $\sup_{u \in [0,1], t \geq 1} d_{n,t}(u) = O((k_n/n)^{1/r} k_n^{-1/2})$  and  $\int_0^1 u^{-1} d_{n,t}(u) du = O(k_n^{-1/2} (k_n/n)^{1/r})$  for some  $r \geq q$ . Further  $\varphi_{l_n} \in [0, 1]$  uniformly in and  $\varphi_{l_n} = o((k_n/n)^{1/p-1/r} l_n^{-\lambda})$ .

**Lemma A.9** (Hill, 2007a). Let  $\{X_t\}$  be  $L_p$ -NED on  $\{F_{n,t}\}$  with constants  $d_t$ ,  $\sup_{t \geq 1} d_t < \infty$ , and coefficients  $\vartheta_{l_n}$  of size  $\lambda > 0$ . Then  $\{X_t\}$  satisfies the  $L_q$ -FE-NED property (14) for any  $q \geq 2$  with constants  $\tilde{d}_{n,t}$ ,  $\sup_{t \geq 1} \tilde{d}_{n,t} = O((k_n/n)^{1/r} k_n^{-1/2})$  for some  $r \geq q$  and coefficients  $\varphi_{l_n}$ .

**Proof of Lemma A.1.** It suffices to prove  $X_n(\zeta, u)$  convergences in finite dimensional distributions, and  $\{X_n(\zeta, u)\}$  is tight on  $[\underline{\zeta}, 1] \times [0, 1]^{k-1}$ .

Decompose  $X_n(\zeta, u)$ :

$$\begin{aligned}
 X_n(\zeta, u) &= \sum_{i=1}^{r_n(\zeta)} W_{n,i}(u) + \sum_{t=r_n(\zeta)g_n+1}^{n(\zeta)} X_{n,t}(u) + \sum_{i=1}^{r_n(\zeta)} E[Z_{n,i}(u)|\tilde{F}_{n,i-1}] \\
 &\quad + \sum_{i=1}^{r_n(\zeta)} (Z_{n,i}(u) - E[Z_{n,i}(u)|\tilde{F}_{n,i-1}]) + \sum_{i=1}^{r_n(\zeta)} \sum_{t=(i-1)g_n+1}^{(i-1)g_n+l_n} X_{n,t}(u) \\
 &= \sum_{i=1}^{r_n(\zeta)} W_{n,i}(u) + \sum_{t=r_n(\zeta)g_n+1}^{n(\zeta)} X_{n,t}(u) + R_n(u),
 \end{aligned}
 \tag{22}$$

say. By the definition of an  $L_r$ -functional array and  $r_n(\zeta) = \lfloor n(\zeta)/g_n \rfloor$

$$\left\| \sum_{t=r_n(\zeta)g_n+1}^{n(\zeta)} X_{n,t}(u) \right\|_1 = O((n(\zeta) - r_n(\zeta)g_n)n^{-a(1)}) = o(1),$$

and conditions (a)–(c) imply  $R_n(u) = o_p(1)$ , hence

$$X_n(\zeta, u) = \sum_{i=1}^{r_n(\zeta)} W_{n,i}(u) + o_p(1) = W_n(\zeta, u) + o_p(1).$$

Lemma A.2 proves the CLT claim (i), and the FCLT claim (ii) follows from Lemmas A.2 and A.3.  $\square$

**Proof of Lemma A.2.** Pick any  $\pi \in \mathbb{R}^h$ ,  $\pi' \pi = 1$ , and notice for any finite collection  $\{(\zeta_j, u_j)\}_{j=1}^h$ ,  $\zeta_1 \leq \dots \leq \zeta_h$ , by construction, cf. (20),

$$\sum_{j=1}^h \pi_j W_n(\zeta_j, u_j) = \sum_{i=1}^{r_n(\zeta_h)} \tilde{W}_{n,i}(u, \pi).$$

By construction  $\{\tilde{W}_{n,i}(u, \pi), \tilde{F}_{n,i}\}$  forms a martingale difference array. We will show under conditions (a)–(d) of Lemma A.1

$$\sum_{i=1}^{r_n(\zeta_h)} \tilde{W}_{n,i}(u, \pi) \xrightarrow{d} N \left( \mathbf{0}, \lim_{n \rightarrow \infty} E \left( \sum_{i=1}^{r_n(\zeta_h)} \tilde{W}_{n,i}(u, \pi) \right)^2 \right)$$

is a consequence of Theorem 2.3 of [McLeish \(1974\)](#), where  $\lim_{n \rightarrow \infty} \|\sum_{i=1}^{r_n(\zeta_h)} \tilde{W}_{n,i}(u, \pi)\|_2 \leq K$  follows from (22) and  $\|X_n(\zeta, u)\|_2 = O(1)$ . A Cramér–Wold device completes the proof.

We need only verify McLeish’s conditions (a)–(c). McLeish’s condition (c) is condition (d) of Lemma A.1. Moreover, McLeish’s conditions (a) and (b) apply under the Lindeberg condition  $\sum_{i=1}^{r_n(\zeta_h)} E[\tilde{W}_{n,i}^2(u, \pi) \times I(|\tilde{W}_{n,i}(u, \pi)| > \varepsilon)] \rightarrow 0$  for any  $\varepsilon > 0$ . Since  $r_n(\zeta_h) \leq r_n(1)$ , and  $r_n(1)g_n \sim n(1) \leq n$ , for any  $\varepsilon > 0$  it follows

$$\begin{aligned}
 &r_n(\zeta_h) \times \max_{i \geq 1} E[\tilde{W}_{n,i}^2(u, \pi) I(|\tilde{W}_{n,i}(u, \pi)| > \varepsilon)] \\
 &\leq r_n(1) \times \max_{i \geq 1} \|\tilde{W}_{n,i}(u, \pi)\|_4^2 \|\tilde{W}_{n,i}(u, \pi)\|_2 \varepsilon^{-1} \\
 &\leq r_n(1) \times |\pi| \varepsilon^{-1} \times \max_{i \geq 1} \left\{ \sup_{u \in [0,1]^{k-1}} \|W_{n,i}(u)\|_4^2 \sup_{u \in [0,1]^{k-1}} \|W_{n,i}(u)\|_2 \right\} \\
 &\leq K \times r_n(1) \times g_n^2 \sup_{u \in [0,1]^{k-1}, t \geq 1} \|X_{n,t}(u)\|_4^2 \times g_n \sup_{u \in [0,1]^{k-1}, t \geq 1} \|X_{n,t}(u)\|_2 \\
 &= O(r_n g_n^2 n^{-2a(4)} g_n n^{-a(2)}) = O(g_n^2 n^{1-2a(4)-a(2)}) = o(1).
 \end{aligned}$$

The first inequality is Hölder’s and Markov’s; the second and third are Jensen’s and Minkowski’s; and the last line follows from the definition of a functional array and  $g_n = o(n^{[2a(4)+a(2)-1]/2})$ . The Lindeberg condition is therefore satisfied which completes the proof.  $\square$

**Proof of Lemma A.3.** Define for each element  $\theta \in \{\zeta, u_1, \dots, u_{k-1}\}$

$$w''_{\theta}(X_n, \delta) := \sup_{\theta_1 \leq \theta \leq \theta_2; \theta_2 - \theta_1 \leq \delta} \{|X_n(\cdot, \theta) - X_n(\cdot, \theta_1)| \wedge |X_n(\cdot, \theta) - X_n(\cdot, \theta_2)|\},$$

$$w_{\theta}(X_n, \delta) := \sup_{\theta \leq \theta' \leq \theta + \delta} |X_n(\cdot, \theta) - X_n(\cdot, \theta')|.$$

According to Proposition 1 of Hill (2007a), cf. Bickel and Wichura (1971) and Neuhaus (1971), it suffices to show for every  $\varepsilon, \eta > 0$  there exists a  $\delta \in [0, 1]$  and  $n_0 \in \mathbb{N}$  such that

$$P(w''_{\zeta}(W_n(\zeta, u), \delta) > \varepsilon/k) \leq \eta/k, \quad n \geq n_0, \tag{23}$$

$$P(w''_{u_i}(W_n(\zeta, u), \delta) > \varepsilon/k) \leq \eta/k, \quad n \geq n_0, \quad i = 1 \dots k - 1, \tag{24}$$

and for any  $\varepsilon > 0$

$$\begin{aligned} \lim_{\delta \rightarrow 0} P(|W_n(\zeta = 1, \cdot) - W_n(\zeta = 1 - \delta, \cdot)| > \varepsilon), \\ \lim_{\delta \rightarrow 0} P(|W_n(\cdot, u_i = 1) - W_n(\cdot, u_i = 1 - \delta)| > \varepsilon), \quad i = 1 \dots k - 1. \end{aligned} \tag{25}$$

Lemma A.3.1 and the property (cf. Billingsley, 1999)

$$w''_{\zeta}(W_n(\zeta, u), \delta) \leq w_{\zeta}(W_n(\zeta, u), k\delta), \quad \delta \in [0, 1/k]$$

imply (23), and Lemmas A.3.2 and A.3.3, respectively, establish (24) and (25).

Finally, Lemma A.3.1 and Theorem 2 of Wichura (1969) ensure  $P(W(\zeta, \cdot) \in C[\underline{\zeta}, 1]) = 1$ . If condition (g) of Lemma A.1 holds then Lemma A.3.4 guarantees  $P(W(\cdot, u) \in C[0, 1]^{k-1}) = 1$ .  $\square$

**Lemma A.3.1.** For each  $\varepsilon, \eta > 0$  there exists some  $\delta \in (0, 1/k]$  and  $N_0 \in \mathbb{N}$  such that  $P(w_{\zeta}(W_n(\zeta, u), \delta) > \varepsilon/k) \leq \eta \delta \forall n \geq N_0$ .

**Proof.** Drop the argument  $u$  and write  $W_n(\zeta) = W_n(\zeta, \cdot)$  and  $W_{n,i} = W_{n,i}(u)$ . Let  $Z \sim N(0, 1)$  and  $\kappa \geq 1$  be chosen below. Choose any

$$\lambda > \max\{\varepsilon/\sqrt{2}, 2k^2 \times E|Z|^3 \times \kappa/\eta\varepsilon^2\}$$

and fix

$$\delta = \frac{\varepsilon^2}{\kappa 2k\lambda^2} \leq \frac{1}{k}.$$

We can always find a sequence of positive integers  $\{r_n^*(\zeta)\}$  satisfying  $0 \leq [\delta r_n^*(\zeta)] \leq r_n(\zeta + \delta) - r_n(\zeta)$  and  $0 \leq r_n^*(\zeta) \leq n(\zeta)$  such that

$$\begin{aligned} P\left(\sup_{\zeta \leq \zeta' \leq \zeta + \delta} |W_n(\zeta') - W_n(\zeta)| > \lambda v_n\right) &\leq P\left(\sup_{1 \leq j \leq [\delta r_n^*(\zeta)]} \left| \sum_{i=1}^{r_n(\zeta)+j} W_{n,i} - \sum_{i=1}^{r_n(\zeta)} W_{n,i} \right| > \lambda v_n\right) \\ &\leq E \left| \sum_{i=r_n(\zeta)+1}^{r_n(\zeta)+[\delta r_n^*(\zeta)]} W_{n,i}/v_n \right|^3 \lambda^{-3}, \end{aligned} \tag{26}$$

where  $v_n := \|\sum_{i=r_n(\zeta)+1}^{r_n(\zeta)+[\delta r_n^*(\zeta)]} W_{n,i}\|_2$ , and the second inequality is Kolmogorov's. By construction

$$\frac{E|Z|^3}{\lambda^3} < \frac{\eta\varepsilon^2}{\lambda^2 2k^2 \kappa} = \frac{\eta\delta}{k}$$

and from Lemma A.2 for some standard normal  $Z$

$$\sum_{i=r_n(\zeta)+1}^{r_n(\zeta)+[\delta r_n^*(\zeta)]} W_{n,i}/v_n \xrightarrow{d} Z.$$

Thus, there exists a sufficiently large  $N_0$  such that  $\forall n \geq N_0$

$$E \left| \sum_{i=r_n(\xi)+1}^{r_n(\xi)+[\delta r_n^*(\xi)]} W_{n,i}/v_n \right|^3 \lambda^{-3} \leq \frac{\eta \varepsilon^2}{\lambda^2 2k^2 \kappa} = \frac{\eta \delta}{k} \leq \eta \delta. \tag{27}$$

Now, by condition (e) of Lemma A.1 and the martingale difference property of  $\{W_{n,i}, \tilde{F}_{n,i}\}$ , for some finite  $\kappa \geq 1$

$$v_n^2 = E \left( \sum_{i=r_n(\xi)+1}^{r_n(\xi)+[\delta r_n^*(\xi)]} W_{n,i} \right)^2 = \sum_{i=r_n(\xi)+1}^{r_n(\xi)+[\delta r_n^*(\xi)]} E(W_{n,i})^2 \leq \delta \times \kappa.$$

Hence, for sufficiently large  $N_0$  and all  $n \geq N_0$ , and  $\delta = \varepsilon^2 / [\kappa 2k\lambda^2]$ , we have  $\lambda v_n \leq \lambda \delta^{1/2} \kappa^{1/2} = \varepsilon/k$ . The claim now follows from (26)–(27):  $P(w_\xi(W_n(\xi, u), \delta) > \varepsilon/k) \leq P(\sup_{\xi \leq \xi' \leq \xi+\delta} |W_n(\xi') - W_n(\xi)| > \lambda v_n) \leq \eta \delta \forall n \geq N_0$ .  $\square$

**Lemma A.3.2.** *There exists for each  $j = 1 \dots k - 1$  and every  $\varepsilon, \eta > 0$  some  $\delta \in [0, 1/k]$  and  $N_0 \in \mathbb{N}$  such that  $P(w'_{u_j}(W_n(\xi, u), \delta) > \varepsilon/k) \leq \eta/k \forall n \geq N_0$ .*

**Proof.** Write  $W_n(u_j) = W_n(\xi, \{u_i\}_{i \neq j}, u_j)$  and let  $0 \leq u_{1j} \leq u_{2j} \leq u_{3j} \leq 1$ . By condition (f) of Lemma A.1 and  $r_n(\xi) = [n(\xi)/g_n]$

$$\begin{aligned} E[|W_n(u_{1j}) - W_n(u_{2j})| |W_n(u_{2j}) - W_n(u_{3j})|] &\leq \|W_n(u_{1j}) - W_n(u_{2j})\|_2 \|W_n(u_{2j}) - W_n(u_{3j})\|_2 \\ &= \left\| \sum_{i=1}^{r_n(\xi)} (W_{n,i}(u_{1j}) - W_{n,i}(u_{2j})) \right\|_2 \times \left\| \sum_{i=1}^{r_n(\xi)} (W_{n,i}(u_{2j}) - W_{n,i}(u_{3j})) \right\|_2 \\ &= |u_{2j} - u_{1j}|^{1/2} \times |u_{3j} - u_{2j}|^{1/2} \times O(r_n(\xi)g_n/n) \leq K \times |u_{3j} - u_{1j}|, \end{aligned}$$

for some finite  $K > 0$ . Now apply (13.13)–(13.14) of Billingsley (1999).  $\square$

**Lemma A.3.3.** *For every  $\varepsilon > 0$ ,  $\lim_{\delta \rightarrow 0} P(|W_n(\xi = 1, \cdot) - W_n(\xi = 1 - \delta, \cdot)| > \varepsilon) = 0$  and  $\lim_{\delta \rightarrow 0} P(|W_n(\cdot, u_j = 1) - W_n(\cdot, u_j = 1 - \delta)| > \varepsilon) = 0$ ,  $j = 1 \dots k - 1$ .*

**Proof.** Use conditions (e) and (f) of Lemma A.1 to deduce

$$\begin{aligned} P(|W_n(\xi = 1, \cdot) - W_n(\xi = 1 - \delta, \cdot)| > \varepsilon) &\leq \varepsilon^{-2} \times (r_n(1) - r_n(1 - \delta))^2 \sup_{i \geq 1} \|W_{n,i}(u)\|_2^2 \\ &\leq \varepsilon^{-2} \times (r_n(1) - r_n(1 - \delta))^2 \sup_{i \geq 1} \|W_{n,i}(u)\|_2^2 \\ &\leq K(r_n(1) - r_n(1 - \delta))^2 \rightarrow 0 \end{aligned}$$

as  $\delta \rightarrow 0$  since  $r(\cdot) = [n(\cdot)/g_n]$  and  $n(1 - \delta) = n(1)$ . Finally, (f) and Chebychev’s inequality imply for any  $j = 1 \dots k - 1$

$$P(|W_n(\cdot, u_j = 1) - W_n(\cdot, u_j = 1 - \delta)| > \varepsilon) \leq \varepsilon^{-2} K \times \delta. \quad \square$$

**Lemma A.3.4.** *Let  $v_{j,n} := \|\sup_{u_{1j} \leq u_{2j} \leq u_{1j}+\delta} |W_n(\cdot, u_{1j}) - W_n(\cdot, u_{2j})|\|_2 \geq v_j > 0$  uniformly in  $n$ . There exists for every  $\varepsilon, \eta > 0$  some  $\delta \in [0, 1/k]$  and  $N_0 \in \mathbb{N}$  such that  $P(w_{u_j}(W_n(\xi, u), \delta) > \varepsilon/k) \leq \eta \delta \forall n \geq N_0$ .*

**Proof.** For large  $K > 1$  choose any  $\lambda > \max\{\varepsilon/[k^{1/2}K^{1/2}], K^{1/2}/[\eta^{1/2}v_j^{1/2}]\}$  and fix  $\delta = \varepsilon^2/[k^2\lambda^2K] \leq 1/k$ . Condition (g) of Lemma A.1 and Chebyshev’s inequality imply  $\lambda v_{j,n} \leq \lambda K^{1/2} \delta^{1/2} = \varepsilon/k$  and

$$\begin{aligned} P \left( \sup_{u_{1j} \leq u_{2j} \leq u_{1j}+\delta} |W_n(\cdot, u_{1j}) - W_n(\cdot, u_{2j})| > \lambda v_{j,n} \right) &\leq E \left( \sup_{u_{1j} \leq u_{2j} \leq u_{1j}+\delta} |W_n(\cdot, u_{1j}) - W_n(\cdot, u_{2j})| \right)^2 \times \lambda^{-2} v_{j,n}^{-2} \\ &\leq \frac{K\delta}{\lambda^2 v_{j,n}^2} \leq \frac{\eta \delta v_j}{v_{j,n}^2} \leq \eta \delta. \end{aligned}$$

**Proof of Lemma A.4.** Claim (i): Recall  $F_{n,t} = \sigma(E_{n,\tau} : \tau \leq t)$  where  $E_{n,t}$  is  $\sigma(\varepsilon_\tau : \tau \leq t)$ -measurable. Under Assumption 1(b) if  $\varepsilon_t$  is F-strong mixing with coefficients  $\varepsilon_{l_n}$ , then  $E_{n,t}$  is strong mixing, hence Theorem 17.5 of Davidson (1994) implies

$$\|X_{n,t}(u) - E[X_{n,t}(u)|F_{n,t-l_n}]\|_2 \leq \max\{\|X_{n,t}(u)\|_r, d_{n,t}(u)\} \times \max(6\varepsilon_{l_n}^{1/2-1/r}, \varphi_{l_n}).$$

By assumption  $n^{[1/2-a(r)]2r/[r-2]}l_n^{r/[r-2]}\varepsilon_{l_n} = o(1)$ ,  $\sup_{u \in [0,1]^{k-1}, t \geq 1} \|X_{n,t}(u)\|_r = O(n^{-a(r)})$ ,  $\sup_{u \in [0,1]^{k-1}, t \geq 1} d_{n,t}(u) = O(n^{-a(r)})$ , and  $\varphi_{l_n} = o(n^{a(r)-a(2)} \times l_n^{-1/2})$  where  $a(2) = \frac{1}{2}$  under the functional array definition. We may therefore write for sufficiently large  $K > 0$

$$\begin{aligned} \|X_{n,t}(u) - E[X_{n,t}(u)|F_{n,t-l_n}]\|_2 &\leq (K \times n^{-1/2}) \times \max\{(n^{[a(r)-1/2]2r/[r-2]}\varepsilon_{l_n})^{1/2-1/r}, n^{1/2-a(r)}\varphi_{l_n}\} \\ &= c_{n,t} \times \psi_{l_n}, \end{aligned}$$

say, where  $\sup_{t \geq 1} c_{n,t} = O(n^{-1/2})$  is trivial and  $\psi_{l_n} = o(l_n^{-1/2})$  follows from  $n^{[1/2-a(r)]2\lambda}\varepsilon_{l_n} = o(l_n^{-\lambda})$  where  $\lambda = r/[r-2]$  and  $n^{1/2-a(r)}\varphi_{l_n} = o(l_n^{-1/2})$ . A similar argument holds for the remaining mixingale bound  $\|X_{n,t}(u) - E[X_{n,t}(u)|F_{n,t+l_n}]\|_2 \leq c_{n,t}(u)\psi_{l_{n+1}}$  and in the F-uniform mixing case (cf. Davidson, 1994, p. 265).

Claim (ii): An identical argument applies to  $\Delta X_{n,t}(u_1, u_2)$ . Since the  $L_2$ -NED constants of  $\Delta X_{n,t}(u_1, u_2)$  are  $\tilde{d}_{n,t} \times \max_{1 \leq i \leq k-1} |u_{2,i} - u_{1,i}|^{1/2}$ , the mixingale constants are  $\tilde{c}_{n,t} \times \max_{1 \leq i \leq k-1} |u_{2,i} - u_{1,i}|^{1/2}$  for some  $\sup_{t \geq 1} \tilde{c}_{n,t} = O(n^{-1/2})$ .

Claim (iii): The limit  $\sup_{\pi} \pi \pi = 1 \left| \sum_{l=1}^h \sum_{i=r_n(\xi_{l-1})+1}^{r_n(\xi_l)} \left\{ \sum_{j=1}^h \pi_j Z_{n,i}(u_j) \right\}^2 - 1 \right| \xrightarrow{p} 0$  for any sequence of points  $\{(\xi_i, u_i)\}_{i=1}^h$  follows from Lemmas A.5 and A.6 and an argument identical to de Jong's (1997, A.39–A.41).

**Proof of Lemma A.6.** Because  $u \in [0, 1]^{k-1}$  and  $\xi \in [\underline{\xi}, 1]$  are arbitrary, the claim follows from Lemma A.4 of Hill (2005).  $\square$

## References

- Basrak, B., Davis, R.A., Mikosch, T., 2002. A characterization of multivariate regular variation. *Ann. Appl. Probab.* 12, 908–920.
- Beirlant, J., Teugels, J., Vynckier, P., 1994. Extremes in non-life insurance. In: Galambos, J., Lechner, J., Simui, E. (Eds.), *Extreme Value Theory and Applications*. Kluwer, Dordrecht.
- Bickel, P.J., Wichura, M.J., 1971. Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* 42, 1656–1670.
- Billingsley, P., 1999. *Convergence of Probability Measures*. Wiley, New York.
- Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. *Regular Variation*. Cambridge University Press, Great Britain.
- Bradley, B., Taqqu, M., 2003. Financial risk and heavy tails. In: Rachev, S. (Ed.), *Handbook of Heavy Tailed Distributions in Finance*. Elsevier, Amsterdam.
- Carrasco, Chen, 2002. Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* 18, 17–39.
- Coles, S., Heffernan, J., Tawn, J., 1999. Dependence measures for extreme value analysis. *Extremes* 2, 339–365.
- Csörgő, M., Csörgő, S., Horváth, L., Mason, D.M., 1986. Weighted empirical and quantile processes. *Ann. Probab.* 14, 31–85.
- Davidson, J., 1992. The central limit theorem for globally non-stationary near-epoch-dependent functions of mixing processes. *Econometric Theory* 8, 313–334.
- Davidson, J., 1994. *Stochastic Limit Theory*. Oxford University Press, Oxford.
- Davidson, J., 2004. Moment and Memory Properties of Linear Conditional Heteroscedasticity Models, and a New Model. *J. Bus. Econ. Statist.* 22, 16–29.
- Davidson, J., de Jong, R.M., 2000. The functional central limit theorem and weak convergence to stochastic integrals. *Econometric Theory* 16, 621–642.
- de Jong, R.M., 1997. Central limit theorems for dependent heterogeneous random variables. *Econometric Theory* 13, 353–367.
- Dellacherie, C., Meyer, P.-A., 1978. *Probabilities and Potential*. North-Holland, Amsterdam.
- Drees, H., 2002. Tail empirical processes under mixing conditions. In: Dehling, H., Mikosch, T., Sørensen's, M. (Eds.), *Empirical Process Techniques for Dependent Data*. Birkhauser, Berlin.
- Drees, H., Ferreira, A., de Haan, L., 2004. On maximum likelihood estimation of the extreme value index. *Ann. Appl. Probab.* 14, 1179–1201.
- Einmahl, J.H.J., 1992. Limit theorems for tail processes with application to intermediate quantile estimation. *J. Statist. Plann. Inference* 32, 127–145.
- Einmahl, J.H.J., Lin, T., 2006. Asymptotic normality of extreme value estimators on  $C[0,1]$ . *Ann. Statist.* 34, 469–492.
- Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. *Modelling Extremal Events for Insurance and Finance*. Springer, New York.
- Embrechts, P., Lindskog, F., McNeil, A., 2003. Modelling dependence with copulas and applications to risk management. In: Rachev, S. (Ed.), *Handbook of Heavy Tailed Distributions in Finance*. Elsevier, Amsterdam.
- Finkenstadt, B., Rootzén, H., 2003. *Extreme Values in Finance, Telecommunications and the Environment*. Chapman & Hall, London.
- Gallant, A.R., White, H., 1988. *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. Basil Blackwell: Oxford.
- Goldie, C.M., Smith, R.L., 1987. Slow variation with remainder: theory and applications. *Quart. J. Math.* 38, 45–71.
- Haeusler, E., Teugels, J.L., 1985. On asymptotic normality of Hill's estimator for the exponent of regular variation. *Ann. Statist.* 13, 743–756.
- Hall, P., 1982. On Some Estimates of an Exponent of Regular Variation. *J. Roy. Statist. Soc. Ser. B* 44, 37–42.
- Hall, P., Heyde, C.C., 1980. *Martingale Limit Theory and its Applications*. Academic Press, New York.
- Hall, P., Welsh, A.H., 1985. Adaptive estimates of parameters of regular variation. *Ann. Statist.* 13, 331–341.
- Heffernan, J., Tawn, J.A., 2004. A conditional approach to multivariate extreme values. *J. Roy. Statist. Soc. Ser. B* 66, 497–546.
- Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. *Ann. Math. Statist.* 3, 1163–1174.
- Hill, J.B., 2005. On tail index estimation for dependent, heterogeneous data. Working paper, Department of Economics, University of North Carolina, Chapel Hill. Available at: ([www.unc.edu/~jbhill/hill\\_het\\_dep.pdf](http://www.unc.edu/~jbhill/hill_het_dep.pdf)).
- Hill, J.B., 2007a. Technical appendix for "On functional central limit theorems for dependent, heterogeneous arrays with applications to tail index and tail dependence estimation". Available at: ([www.unc.edu/~jbhill/tech\\_append\\_tail\\_fclt.pdf](http://www.unc.edu/~jbhill/tech_append_tail_fclt.pdf)).
- Hill, J.B., 2007b. Robust non-parametric tests of extremal dependence. Working paper, Department of Economics, University of North Carolina, Chapel Hill. Available at: ([www.unc.edu/~jbhill/extremal\\_spillover.pdf](http://www.unc.edu/~jbhill/extremal_spillover.pdf)).
- Hill, J.B., 2008a. Extremal memory of stochastic volatility with applications to tail shape and tail dependence inference. Working paper, Department of Economics, University of North Carolina, Chapel Hill. Available at: ([www.unc.edu/~jbhill/stoch\\_vol\\_tails.pdf](http://www.unc.edu/~jbhill/stoch_vol_tails.pdf)).
- Hill, J.B., 2008b. Tail and non-tail memory with applications to distributed lags and central limit theory. Working paper, Department of Economics, University of North Carolina, Chapel Hill. Available at: ([www.unc.edu/~jbhill/stoch\\_vol\\_tails.pdf](http://www.unc.edu/~jbhill/stoch_vol_tails.pdf)).
- Hsing, T., 1991. On tail index estimation using dependent data. *Ann. Statist.* 19, 1547–1569.

- Hsing, T., 1999. On the asymptotic distributions of partial sums of functionals of infinite-variance moving averages. *Ann. Probab.* 27, 1579–1599.
- Ibragimov, I.A., 1962. Some limit theorems for stationary processes. *Theory Probab. Appl.* 7, 349–382.
- Ibragimov, I.A., Linnik, Y.V., 1971. *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhof: Berlin.
- Klüppelberg, C., Kuhn, G., Peng, L., 2007. Estimating the tail dependence function of an elliptical distribution. *Bernoulli* 13, 229–251.
- Leadbetter, M.R., 1983. Extremes and local dependence in stationary sequences. *Z. Wahrsch. Verw. Geb* 65, 291–306.
- Leadbetter, M.R., Rootzén, H., 1993. On central limit theory for families of strongly mixing additive set functions. In: Cambanis, S., Sen, P.K. (Eds.), *Festschrift in Honour of Gopinath Kallianpur*. Springer, New York, pp. 211–223.
- Leadbetter, M.R., Lindgren, G., Rootzén, H., 1983. *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.
- Ledford, A.W., Tawn, J.A., 1997. Modeling dependence within joint tail regions. *J. Roy. Statist. Soc. Ser. B* 59, 475–499.
- Ledford, A.W., Tawn, J.A., 2003. Diagnostics for dependence within time series extremes. *J. Roy. Statist. Soc. Ser. B* 65, 521–543.
- Mason, D., 1988. A strong invariance theorem for the tail empirical process. *Ann. Inst. H. Poincaré* 24, 491–506.
- McLeish, D.L., 1974. Dependent central limit theorems. *Ann. Probab.* 2, 620–628.
- McLeish, D.L., 1975. A maximal inequality and dependent strong laws. *Ann. Probab.* 3, 829–839.
- Meitz, M., Saikkonen, P., 2008. Stability of nonlinear AR-GARCH models. *J. Time Ser. Anal.* 29, 453–475.
- Neuhas, G., 1971. On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* 42, 1285–1295.
- Phillips, P.C.B., Durlauf, S.N., 1986. Multiple time series regression with integrated processes. *Rev. Econom. Stud.* 53, 473–495.
- Pipiras, V., Taqqu, M.S., 2003. Central limit theorems for partial sums of bounded functionals of infinite-variance moving averages. *Bernoulli* 9, 833–855.
- Pipiras, V., Taqqu, M.S., Abry, P., 2007. Bounds for the covariance of functions of infinite variance stable random variables with applications to central limit theorems and wavelet-based estimation. *Bernoulli* 13, 1091–1123.
- Resnick, S., 1987. *Extreme Values, Regular Variation and Point Processes*. Springer, New York.
- Rootzén, H., 1995. The tail empirical process for stationary sequences. Report 1995:9, Department of Mathematical Statistics, Chalmers University of Technology.
- Rootzén, H., Leadbetter, M., de Haan, L., 1998. On the distribution of tail array sums for strongly mixing stationary sequences. *Ann. Probab.* 8, 868–885.
- Schmidt, R., Stadtmüller, U., 2006. Non-parametric estimation of tail dependence. *Scand. J. Statist.* 33, 307–335.
- Wang, Q.Y., Lin, Y.X., Gulati, C.M., 2002. The invariance principle for linear processes with applications. *Econometric Theory* 18, 119–139.
- Wichura, M.J., 1969. Inequalities with applications to the weak convergence of random processes with multidimensional time parameters. *Ann. Math. Statist.* 40, 681–687.
- Wooldridge, J.M., White, H., 1988. Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory* 4, 210–230.
- Wu, W.B., Min, M., 2005. On linear processes with dependent innovations. *Stochastic Process. Appl.* 115, 939–958.
- Wu, W.B., Woodroffe, M., 2004. Martingale approximations for sums of stationary processes. *Ann. Probab.* 32, 1674–1690.