Financial
Econometrics
Research Centre

WORKING PAPERS SERIES

WP04-03

Decoupling and Convergence to Independence with Applications to Functional Limit Theorems

Alessio Sancetta

Decoupling and Convergence to Independence with Applications to Functional Limit Theorems

Alessio Sancetta*

June 10, 2004

Abstract

This paper provides a new approach to study and apply dependence conditions. It provides exact decoupling in terms of the copula function. Based on these ideas a new dependence condition is formulated. In general we characterize convergence to independence via weak convergence ideas. In this context, this is often sufficient for our purposes. To provide a unifying approach, the most common dependence conditions used in the literature are illustrated using this approach. As an application, a CLT, stochastic equicontinuity results and a Donsker theorem are proved for dependent random variables under minimal dependence conditions.

Key Words: Copula, Coupling, Linkages, Mixing, Weak Convergence, Stochastic Equicontinuity, Regression Construction.

Classification AMS:60F05 60G17.

1 Introduction

This paper considers the problem of dependent sequences and their limiting behaviour from a new prospective. It characterizes convergence of a dependent subsequence to an independent subsequence in terms of weak convergence of measures on the unit interval.

***Aknowledgments:** I thank Paul Doukhan for useful comments and discussions. Errors and omissions are my sole resposability. Financial support from the ESRC Award RES 000-23-0400 is gratefully aknowledged. **Address for Correspondence:** Department of Applied Economics, Sidgwick Avenue, University of Cambridge, CambridgeCB3 9DE, UK. E-mail: alessio.sancetta@econ.cam.ac.uk.

The paper uses the copula function (e.g. Sklar, 1973) as a natural tool to deal with these limit theorems in the case of dependent random variables. The copula function is the joint distribution of uniform random variables in the unit hypercube. By Sklar's Theorem, any joint distribution can be rewritten in terms of the copula function, where the uniform random variables correspond to the Rosenblatt transform of the original ones or some variant of it. Suppose $X := (X_i)_{i\in\mathcal{H}}$ is a sequence of random variables indexed in $\mathcal{H} \subset \mathbb{Z}$ with law μ. Suppose X has copula C, and $\mathcal{Q}$ is a transform such that $U := \mathcal{Q}^{-1}X$ is a $[0,1]^{\mathcal{H}}$ uniform random vector with joint distribution C (this transform does not require the marginal distributions to be continuous). Then for any measurable function f, and the Lebesgue measure λ in $[0,1]^{\mathcal{H}}$ we can write

$$\mu(f) = \lambda\left((1+\gamma) f \circ \mathcal{Q}\right),$$

where γ is some measurable function that is uniquely determined by C. The paper studies the properties of γ and is concerned with conditions under which $\mu\mathcal{Q}^{-1} \to \lambda$. We define the class of functions $\mathfrak{F}$ under which weak convergence of $\gamma\lambda$ to zero holds.

This approach unifies the previous approaches via mixing conditions. In particular, I will consider several dependence conditions used in the literature and show how these can be naturally reformulated in term of the copula function. Any copula function can be written as the independent copula plus a perturbation factor (the indefinite integral of γ) that accounts for dependence in the random variables (see Rüschendorf, 1985). Therefore, by the Radon-Nikodym derivative of the original copula with respect to the independent copula, we can transform any series of dependent random variables in terms of an independent series of random variables with same marginals as the originals.

The advantage of the present formulation is that all variables are redefined on a new probability space equipped with the Lebesgue measure. As a consequence of this redefinition, the new random variables have continuous laws. This shows that the class $\mathfrak{F}$ under which weak convergence to independence holds is much larger that the usual class of continuous bounded functions under which weak convergence usually holds. This fact allows to prove new stochastic equicontinuity results by simple extension of the well known results for independent random variables.

As a final comment, it is worth noticing that the use of the copula function in statistical modelling is quite wide spread (especially in finance and economics). Therefore, a formulation of dependence in terms of it is particularly convenient for applied researchers making use of results in the probability literature.

The plan of the paper is as follows. Section 2 considers the decomposition of measures in terms of product probabilities and a sign measure $(\gamma\lambda)$ that completely captures dependence. The regression construction is discussed. In Section 3, metrics for independence are used to define suitable dependence and decoupling conditions. These conditions are defined in terms of the copula and compared to existing one. There is a direct link between the copula and some well known covariance inequalities (e.g. Viennet's (1997) inequality). Section 4 is concerned with the applications. I consider a CLT without rate assumptions. Two new stochastic equicontinuity results are proved, and as a consequence of these a Donsker theorem under minimal dependence conditions.

Throughout the paper the following notation will be used. If A is a set, ∂A stands for its boundary, and $\#A$ stands for its cardinality. Use $a \lesssim b$ to mean that b is greater than a up to a positive finite constant; $a \asymp b$ if $b \lesssim a \lesssim b$; $a \sim b$ to mean asymptotic equivalence; and $\overset{w}{\to}$, $\overset{p}{\to}$, and $\overset{a.s.}{\to}$ for weak convergence (with respect to continuous bounded functions), convergence in probability and almost surely, respectively. Use $\|...\|_{p,\mu}$ $(p \in [1,\infty])$ for the $L_p(\mu)$ norm with respect to the measure μ ($\|...\|_{\infty,\mu}$ is the essential supremum norm). For a function f on a metric space $(\mathbf{S}, d)$ define $\|f\|_L$ to be the Lipschitz seminorm, i.e.

$$\|f\|_L := \sup_{x \neq y} |f(x) - f(y)| / d(x,y)),$$

and $\|f\|_{BL} := \|f\|_{\infty,\lambda} + \|f\|_L$ for the bounded Lipschitz metric. Finally, for any function $f : \mathbb{R} \to \mathbb{R}$, use $f^{-1}(x) := \inf\{y : f(y) \geq x\}$. This notation will be used freely without further reference.

2 Coupling

2.0.1 Radon Nikodym Derivative in L_∞

Suppose μ is a measure on the infinite product sigma algebra $\mathcal{F} := \bigotimes_{i\in\mathbb{N}} \mathcal{A}_i$. Then, we write μ_K for the finite dimensional projection of μ on $\mathcal{F}_K := \bigotimes_{k=1}^{K} \mathcal{A}_i$. It is tacitly assumed that the family is consistent, e.g. $\exists$ a probability kernel Λ_k such that $\mu_{k+1} = \mu_k \otimes \Lambda_{k+1}$ for $k \geq 1$.

Let μ and ν be σ-finite measures in $(\mathbf{S}, \mathcal{A})$. If for each $A \in \mathcal{A}$, $\nu(A) = 0$ implies $\mu(A) = 0$, then μ is absolutely continuous with respect to ν, and this is written as $\mu \ll \nu$. They are called equivalent if for each $A \in \mathcal{A}$, $\nu(A) = 0$ if and only if $\mu(A) = 0$. The Radon-Nikodym theorem say that if $\mu \ll \nu$, then for all $A \in \mathcal{A}$, $\mu(A) = (\phi\nu)(A)$, where ϕ is a.s. unique and, since μ is finite, $\|\phi\|_{\infty,\nu} < \infty$

(e.g. Billingsley, 1995, p. 423). We can write $\phi = d\mu/d\nu$, and call this the Radon-Nikodym derivative. Let $f \in L_1(\nu)$. By Holder's inequality $\mu(f) \leq \|\phi\|_{\infty,\nu} \nu(|f|)$.

Consider the product probability measure $\nu_K = \bigotimes_{k=1}^K P_k$ on the product sigma algebra $\mathcal{F}_K := \bigotimes_{k=1}^K \mathcal{A}_i$ of a Polish space. Suppose μ_K is another probability measure on $\mathcal{F}_K$ with same marginals as ν_K. Then there exists a function $C_\mu^K : [0,1]^K \rightarrow [0,1]$ such that

$$\mu(A_1 \times \cdots \times A_K) = C_\mu^K (P_1(A_1), ..., P_K(A_K)) \tag{1}$$

for any hyper-rectangle of $\mathcal{F}_K$. C_μ^K is called the copula and the present extension to product probability measures of Polish spaces has been given by Scarsini (1989).

Theorem 1 *Suppose ν_K and μ_K are as above. Then $\mu_K \ll \nu_K$, and $d\mu_K/d\nu_K$ exists and it is finite ν_K-almost surely.*

Proof. The representation in (1) holds by Theorem 3.1 in Scarsini (1989). By the properties of the copula (Proposition 4(2)), $C_\mu^K = 0$ if $P_k(A_k) = 0$ for any $k = 1, ..., K$, i.e. the copula is grounded. This implies that $\mu_K \ll \nu_K$, and the remaining statements follow from the Radon-Nikodym theorem and the fact that probability measures are sigma finite measures. ■

For μ_K and ν_K as in Theorem 1, we write $d\mu_K/d\nu_K = dC_\mu^K/d\nu_K := c_\mu^K$. Suppose μ_K is a measure on the Borel sigma algebra of $\mathbb{R}^K$ and λ_K is the Lebesgue measure on the Borel sigma algebra of $[0,1]^K$. Suppose $\mathcal{Q}^K : [0,1]^K \rightarrow \mathbb{R}^K$ is a measurable map such that for any $(u_1, ..., u_K) \in [0,1]^K$

$$\mathcal{Q}^K(u_1, ..., u_K) = \left(F_1^{-1}(u_1), ..., F_K^{-1}(u_K)\right) = (x_1, ..., x_K), \tag{2}$$

where $(x_1, ..., x_K) \in \mathbb{R}^K$, and $F_k(x) := P_k((-\infty, x])$. (The simple modification in (6) takes care of discontinuous marginals, so we do not worry about this for the moment.) Suppressing the superscript K to ease notation, for some $f \in L_1(\mu)$,

$$\begin{aligned} \mu(f) &= C_\mu \circ \mathcal{Q}^{-1} f = C_\mu(f \circ \mathcal{Q}) \\ &= \int_{[0,1]^K} f\left(P_1^{-1}(\lambda_1), ..., P_K^{-1}(\lambda_K)\right) dC_\mu. \end{aligned} \tag{3}$$

To ease notation we will just omit the symbol $\circ$ for the composition of functions. The right hand side of (3) is called the hypercube transform of the integral $\mu(f)$. By trivial re-writing, $\mu(f) = \nu(f) + (\mu - \nu)(f)$ or $\mu(f) = \lambda(f\mathcal{Q}) + \lambda((c_\mu - 1) f\mathcal{Q})$. The transform with respect to $\mathcal{Q}$ allows us to compute any arbitrary integral with respect to a probability measure as an integral in the unit hypercube with respect to the copula. Then, all random variables are defined on the same probability space.

Define

$$\gamma^K := \left(c_\mu^K - 1\right). \tag{4}$$

Rüschendorf (1985) has given a complete characterization of γ^K, in terms of an integral operator and a $L_1(\lambda_K)$ function, say h. However, almost surely, h is bounded, then $h \in L_\infty\left(\lambda^K\right)$. Notice that by the Radon-Nikodym theorem, γ is a mean zero random variable.

The decomposition $\mu\mathcal{Q} = \lambda(1+\gamma)$ (dropping the superscript K) allows us to construct different decoupling equalities and inequalities (e.g. Lemma 26 in the Appendix).

Lemma 2 *Suppose $(X_i)_{i\in\mathbb{N}}$ is a sequence of random variables in a metric space $(\mathbf{S}, d)$ with tight law μ. Suppose $\mathcal{H}_{r_n}$ is a subsequence along $\{1, ..., n\}$. Suppose $\mu_{\mathcal{H}_{r(n)}}$ is the projection of μ on the sigma algebra generated by $(X_i)_{i\in\mathcal{H}_{r(n)}}$ and $\gamma^{\mathcal{H}_{r(n)}}$ is as defined above for the variables indexed in $\mathcal{H}_{r_n}$. Suppose $\lim_n \lambda\left(\gamma^{\mathcal{H}_{r(n)}} f\mathcal{Q}\right) \to 0$, $\forall f$ such that $\|f\|_{BL} \leq 1$. Then, there exists a sequence of independent random variables $(X_i^*)_{i\in\mathcal{H}_{r(n)}}$ with same marginal distributions as $(X_i)_{i\in\mathcal{H}_{r(n)}}$ such that $\lim_n (X_i)_{i\in\mathcal{H}_{r(n)}} \overset{a.s.}{\to} (X_i^*)_{i\in\mathcal{H}_{r(n)}}$ and*

$$\Pr\left(d\left((X_i)_{i\in\mathcal{H}_{r(n)}}, (X_i^*)_{i\in\mathcal{H}_{r(n)}}\right) > \epsilon\right) \leq \epsilon$$

where

$$\lambda\left(\gamma^{\mathcal{H}_{r(n)}}(f\mathcal{Q})\right) \geq \epsilon^2/4$$

for f such that $\|f\|_{BL} \leq 1$.

Proof. It is well known that if $\mu_{\mathcal{H}_{r(n)}}$ converges weakly to the product of marginals of $(X_i)_{i\in\mathcal{H}_{r(n)}}$, then, we can construct a series such that the conclusion of the theorem holds (e.g. Theorem 1.10.4, in van der Vaart and Wellner, 2000). Then, from Corollary 11.6.4 in Dudley (2002) and the discussion following this Corollary, we have the following bound for $X := (X_i)_{i\in\mathcal{H}_{r(n)}}$ and $X^* := (X_i^*)_{i\in\mathcal{H}_{r(n)}}$,

$$\inf\{\epsilon > 0 : \Pr(d(X, X^*) > \epsilon) \leq \epsilon\} \leq 2\left[\lambda\left(\gamma^{\mathcal{H}_{r(n)}}(f\mathcal{Q})\right)\right]^{1/2}.$$

■

The above results show that a successful decoupling is characterized by weak convergence of copulae to the product one. Moreover, the copula allows us to define all random variables on a fixed probability space. In this probability space, the dominating measure is tight.

2.1 The Regression Construction

A reference for almost sure construction of random variables is Rüschendorf and de Valk (1993). Suppose $\mathbf{X} := (X_1, ..., X_K)$ is a K dimensional vector of random variables with joint distribution function F. Suppose $\mathbf{V} := (V_1, ..., V_K)$ is a vector of $[0,1]$ uniform random variables independent of $\mathbf{X}$. Define the following transform $\Psi_F : \mathbb{R}^K \times [0,1]^K \to [0,1]^K$, which depends on the joint distribution $F : \mathbb{R}^K \to [0,1]$, such that

$$\Psi_F(\mathbf{x}, \mathbf{v}) = \left(\tilde{F}_1(x_1, v_1), \tilde{F}_{2|1}(x_2, v_2|x_1) ..., \tilde{F}_{K|1,...,K-1}(x_K, v_K|x_1, ..., x_{K-1})\right), \quad (5)$$

where

$$\begin{aligned}\tilde{F}_{k|1,...,k-1}(x_k, v_k|x_1, ..., x_{k-1}) \ : \ &= \Pr(X_k < x_k|X_i = x_i, i = 1, ..., k-1) \quad (6)\\ &+ v_k \Pr(X_k = x_k|X_i = x_i, i = 1, ..., k-1),\end{aligned}$$

$k = 1, ..., K$. Then define $\Psi_F^* : [0,1]^K \to \mathbb{R}^K$, such that

$$\Psi_F^*(\mathbf{u}) = \mathbf{z} = (z_1, ..., z_K),$$

where

$$z_k := F^{-1}_{k|1,...,k-1}(u_k|z_1, ..., z_{k-1}) := \inf\left\{y : F_{k|1,...,k-1}(y|z_1, ..., z_{k-1}) \geq u_k\right\},$$

and

$$F_{k|1,...,k-1}(x_k|x_1, ..., x_{k-1}) := \Pr(X_k \leq x_k|X_i = x_i, i = 1, ..., k-1).$$

Using the above notation, we have the following (Theorem 3 in Rüschendorf and de Valk, 1993).

Lemma 3 (Regression Construction) *Let* $\mathbf{X}$ *be a* K *dimensional random vector with distribution function* F*. Then*
i. $\mathbf{U} := \Psi_F(\mathbf{X}, \mathbf{V})$ *is a* K *dimensional vector of iid random variables with* $[0,1]$ *uniform distribution;*
ii. $\mathbf{Z} := \Psi_F^*(\mathbf{U})$ *is a* K *dimensional vector of random variables with distribution* F*;* $\mathbf{Z}$ *is called the regression construction of* F*;*
iii. $\mathbf{Z} = \Psi_F^*(\Psi_F(\mathbf{X}, \mathbf{V})) \overset{a.s.}{=} \mathbf{X}$.

The regression construction leads to an optimal coupling for real random variables X and Y with distribution functions F_X and F_Y. In fact a classical result of

Dall'Aglio (1956) says that $\|X-Y\|_{p,\mu}$ $(p \geq 1)$ is minimized in the class of random variables with marginals F_X and F_Y by $X = F_X^{-1}(U)$ and $Y = F_Y^{-1}(U)$ where U is a $[0,1]$ uniform random variable. Hence, without further reference, $\mathcal{Q}^{-1} : \mathbb{R} \to [0,1]$ will be understood as a continuous operator such that

$$\mathcal{Q}^{-1} X = \tilde{F}(X, V). \tag{7}$$

3 Dependence Conditions

4 The Copula Function

We recall Sklar's Theorem (e .g. Sklar, 1973). Let F be a K dimensional distribution function with one dimensional marginals $F_1, ..., F_K$, then there exists a function C from the unit K cube to the unit interval such that

$$F(x_1, ..., x_K) = C(F_1(x_1), ..., F_K(x_K)); \tag{8}$$

C is referred to as the K-Copula. If each F_k is continuous, the copula is unique. If F_k is not continuous, we use $\tilde{F}$ as in (5) to define continuous $[0,1]$ uniform random variables. Then, the copula is always understood to be the joint distribution of these uniform $[0,1]$ continuous random variables. Then, there is a unique copula corresponding to these continuous uniform marginals.

The copula satisfies the following properties (e.g. Scarsini, 1989 for property 1-5 and Rüschendorf, 1985, for property 6).

Proposition 4 (1) *C is increasing in all its arguments. Notice that throughout the paper, we use increasing to mean nondecreasing;*
(2) *C satisfies the Fréchet bounds, i.e.*

$$\max(0, u_1 + ... + u_K - (K-1)) \leq C(u_1, ..., u_K) \leq \min(u_1, ..., u_K),$$

which implies C is grounded: i.e. $C(u_1, ..., u_K) = 0$ if $u_k = 0$ for at least one k, and $C(1, ..., 1, u_k, 1, ..., 1) = u_k$, $\forall k$;
(3) $\prod_{k=1}^{K} u_k$ *is a copula for independent random variables, i.e. the product copula;*
(4) *C is Lipschitz with constant one, i.e.*

$$|C(x_1, ..., x_K) - C(y_1, ..., y_K)| \leq \sum_{k=1}^{K} |x_k - y_k|;$$

(5) *For* $k = 1, ..., K$, *let* $u_k = F_k(x_k)$ *be the distribution function of* x_k *and* $v_k = G_k(y_k)$ *be the distribution function of* $y_k = g_k(x_k)$, *where* $g_k(...)$ *is an increasing function , then* $u_1, ..., u_K$ *and* $v_1, ..., v_K$ *have same copula function;*
(6) *To any copula* $C : [0,1]^K \to [0,1]$ *there corresponds a unique function* $\mathcal{G} : [0,1]^K \to [0,1]$ *such that* $C = u_1 \cdots u_K + \mathcal{G}(u_1, ..., u_K)$.

In particular, we notice that $\mathcal{G}$ is a completely tucked function, i.e. $\mathcal{G}(u_1, ..., u_K) = 0$ for $(u_1, ..., u_K) \in \partial [0,1]^K$. Notice that $\mathcal{G}(u_1, ..., u_K)$ is the distance of the copula from the product copula. This is bounded above and below by the Fréchet bounds. For a 2-copula, the Fréchet bounds define a skewed quadrilateral where the product copula is the paraboloid inside it. All the dependence is measured by "the perturbation term" $\mathcal{G}$. One of the most important facts about copulae is that it is uniformly continuous (Proposition 4 (4)).

Suppose H_j, $j = 1, ..., r$, are discrete index sets with cardinality $\#H_j = n_j$, and $\mathcal{H}_r := \bigcup_{j=1}^{r} H_j$. Suppose $(X_i)_{i \in H_j}$ are random variables with copula C_j. Then, we may write C^{H_j} or C^{n_j} instead of C_j, whichever is more convenient. If the random variables do not have continuous marginals, C_j is understood as the unique copula arising from an application of (7).

It is shown in Genest et al. (1995) that the only 2 dimensional copula consistent with multivariate marginals is the independent one. However, using the regression construction, we can consistently define the copula of nonoverlapping copulae. This is called the Linkage function, but we will just refer to it as the copula (see Li et al., 1996, for further details, though they consider the case of absolutely continuous marginals). Therefore, we can define $C^{\mathcal{H}_r}$ to be the copula of the copulae C_j $(j = 1, ..., r)$. From the definition of copula and Proposition 4 (6), we have the following representation

$$\begin{aligned} C^{\mathcal{H}_r}(\mathbf{u}_1, ..., \mathbf{u}_r) &= \prod_{j=1}^{r} \left(\prod_{i \in H_j} u_i \right) + \mathcal{G}^{\mathcal{H}_r}(\mathbf{u}_1, ..., \mathbf{u}_r) \\ &= C^{H_1}(\mathbf{u}_1) \cdots C^{H_r}(\mathbf{u}_r) + \mathcal{R}^{\mathcal{H}_r}(\mathbf{u}_1, ..., \mathbf{u}_r), \end{aligned} \tag{9}$$

where

$$\mathcal{R}^{\mathcal{H}_r}(\mathbf{u}_1, ..., \mathbf{u}_S) := \mathcal{G}^{\mathcal{H}_r}(\mathbf{u}_1, ..., \mathbf{u}_r) - C^{H_1}(\mathbf{u}_1) \cdots C^{H_r}(\mathbf{u}_r) + \prod_{j=1}^{r} \left(\prod_{i \in H_j} u_i \right), \tag{10}$$

by obvious use of the superscripts for $\mathcal{G}$ and $\mathcal{R}$. We denote the Radon-Nikodym derivative of $\mathcal{R}^{\mathcal{H}_r}$ with respect to the Lebesgue measure in $[0,1]^{K_1} \times \cdots \times [0,1]^{K_K}$, by $\rho^{\mathcal{H}_r}$. By the same argument as in Theorem 1, $\left\|\rho^{\mathcal{H}_r}\right\|_{\infty,\lambda} < \infty$. Moreover $\mathcal{R}^{\mathcal{H}_r}$ is easily seen to satisfy the following.

Proposition 5 (1) $\mathcal{R}^{\mathcal{H}_r} = \mathcal{G}^{\mathcal{H}_r} = 0$ *if* $C^{H_j}(\mathbf{u}_j)$ *is the product copula for each* $j = 1, ..., r$
(2) *if* $\mathbf{u}_j \in \{\mathbf{0}, \mathbf{1}\}$ $j = 1, ..., r$,

$$\mathcal{R}^{\mathcal{H}_r} = 0;$$

(3) *if* $\mathbf{u}_r = \mathbf{1}$,

$$\begin{aligned}
\mathcal{R}^{\mathcal{H}_r}(\mathbf{u}_1, ..., \mathbf{u}_{r-1}, \mathbf{u}_r) &= \mathcal{R}^{\mathcal{H}_{r-1}}(\mathbf{u}_1, ..., \mathbf{u}_{r-1}) \\
&= \mathcal{G}^{\mathcal{H}_{r-1}}(\mathbf{u}_1, ..., \mathbf{u}_{r-1}) - C^{H_1}(\mathbf{u}_1)\cdots C^{H_r}(\mathbf{u}_r) + \prod_{j=1}^{r-1}\left(\prod_{i\in H_j} u_i\right),
\end{aligned}$$

so that

$$\mathcal{R}^{\mathcal{H}_1}(\mathbf{u}_1) = \mathcal{G}^{\mathcal{H}_1}(\mathbf{u}_1) - C^{H_1}(\mathbf{u}_1) + \prod_{i\in H_1} u_i = 0$$

because, $\forall j$,

$$\mathcal{G}^{\mathcal{H}_j}(\mathbf{u}_j) = C^{H_j}(\mathbf{u}_j) - \prod_{i\in H_j} u_i.$$

4.1 Metrics for independence

Suppose $(X_i)_{i\in\mathcal{H}_r}$ is a sequence of random variables. By an appeal to a special case of Lemma (3), we redefine the random variables on the same probability space via the unconditional version of (6), i.e. (7). Hence, the new probability space is equipped with the measure $\lambda(1+\gamma)$. Since in this new probability space the probability measure is absolutely continuous, issues related to pointwise and uniform convergence of $\lambda(1+\gamma)$ to λ are avoided all at once. We depart from the common approach based on sigma algebras and we shall only be concern with the dependence properties in terms of the copula. Moreover, it is more convenient to restrict attention to the coarse sigma algebra of sets $\left\{(\mathbf{0}, \mathbf{u}] : \mathbf{u} \in [0,1]^K\right\}$ $(K \in \mathbb{N})$. Moreover, following Doukhan and Louhichi (1999) we express dependence conditions in terms of tuples. Consider the following distances for convergence to independence: total variation

$$\|\gamma\|_1 \to 0, \tag{11}$$

the Kolmogorov metric

$$\sup_{\mathbf{u}\in(0,1)^K} \left|\lambda\left(\gamma I_{\{(0,\mathbf{u}]\}}\right)\right|, \tag{12}$$

the Lipschitz metric

$$\sup\{|\mathbb{E}\gamma f| : \|f\|_L \leq 1\}, \tag{13}$$

the bounded Lipschitz metric

$$\sup\{|\mathbb{E}\gamma f| : \|f\|_{BL} \leq 1\}. \tag{14}$$

It is well known that (13) is equivalent to the Wasserstein distance for coupling and the transportation problem (e.g. Dudley, 2002, Rüschendorf, 1991); (14) metrizes weak convergence (e.g., van der Vaart and Wellner, 2000); by absolute continuity of the copula (12) is equivalent to (14) (Dudley, 2002, problem 4, p. 389, though weaker conditions suffice, see Dudley, problem 11, p. 390). This does not hold in general (for this reason, strong mixing may fail in some occasions; see example 13, below). If we restrict attention to uniformly integrable functions, (13) and (14) are equivelent, but not in general.

Example 6 (e.g. Dudley, 2002, p.421) *Suppose $\delta_x(s)$ is the Dirac measure at x, $\delta_x(s) = 1$ if $s = x$, 0 otherwise. Then, the measure $(\delta_0(n-1) + \delta_n)/n$ converges to δ_0 under (14), but to not under (13) (just use the definition of dirac measure and then the Lipschitz condition).*

In most cases (11) is stronger than necessary: it corresponds to beta mixing, though uniformly over a coarser sigma algebra. Unlike strong mixing, a similar comment does not apply to (12) because the copula is absolutely continuous.

4.2 Comparisons with Some Existing Dependence Conditions

We recall some well known dependence conditions. Supppose, X and Y are random variables. Then, their strong mixing coefficient is defined as

$$\alpha(\sigma(X), \sigma(Y)) = \sup\{|cov(I_A, I_B)| : A \in \sigma(X), B \in \sigma(Y)\}; \tag{15}$$

while their beta mixing coefficient,

$$\beta(\sigma(X), \sigma(Y)) = \frac{1}{2}\sup\left\{\sum_{i,j}\left|cov\left(I_{A_i}, I_{B_j}\right)\right| : A_i \in \sigma(X), B_j \in \sigma(Y)\right\},$$

where sup is taken over all possible finite partitions of $\sigma(X)$ and $\sigma(Y)$.

Strong mixing corresponds to the kolmogorov metric, while beta mixing to the total variation distance. (For details and applications, see Rio 2000.) These dependent conditions can be rewritten in terms of the copula.

Example 7 *Suppose* $(X_i)_{i\in\mathbb{Z}}$ *is a stationary sequence of random variables. Define* $\beta_i = \beta(\sigma(X_1), \sigma(X_{1+i}))$. *Viennet (1997) has proved the following inequality for the partial sum* $S_n(X)$ *of a* β*-mixing sequence of square integrable random variables* $(X_i)_{i\in\mathbb{Z}}$ *with* β*-mixing coefficients* $(\beta_i)_{i\geq 0}$,

$$var(S_n(X)) \leq 4n\mathbb{E}\left[\left(\sum_{i=1}^{n} b_i\right)^2 X^2\right],$$

where $(b_i)_{i\in\mathbb{Z}_+}$ $b_0 := 1$, $0 \leq b_i \leq 1$, *and* $\mathbb{E}b_i = \beta_i$. *Using the copula we can rediscover the inequality in Viennet (1997). Let* X *and* Y *be two random variables with marginals* F_X *and* F_Y, *copula* C. *Then, for* γ *defined as in (4),*

$$\begin{aligned} cov(X,Y) &= \int_{\mathbb{R}^2} xy\,(dC(F_X(x), F_Y(y)) - dF_X(x)\,dF_Y(y)) \\ &= \int_{\mathbb{R}^2} xy\gamma(F_X(x), F_Y(y))\,dF_X(x)\,dF_Y(y) \\ &\leq \left(\int_{\mathbb{R}^2} x^p\,|\gamma(F_X(x), F_Y(y))|\,dF_X(x)\,dF_Y(y)\right)^{\frac{1}{p}} \\ &\quad \times \left(\int_{\mathbb{R}^2} y^q\,|\gamma(F_X(x), F_Y(y))|\,dF_X(x)\,dF_Y(y)\right)^{\frac{1}{q}}, \end{aligned}$$

by Hölder's inequality with $1/p + 1/q = 1$. *Under stationarity, for* $p = q = 2$,

$$cov(X,Y) \leq \int_{\mathbb{R}^2} x^2\,|\gamma(F_X(x), F_Y(y))|\,dF_X(x)\,dF_Y(y).$$

It follows that

$$\int_{\mathbb{R}^2} |\gamma(F_X(x), F_Y(y))|\,dF_X(x)\,dF_Y(y) = 2\beta(X,Y).$$

The same result can also be written as

$$cov(X,Y) \leq \left(\int_{[0,1]^2} F_X^{-1}(u)^p\,|\gamma(u,v)|\,dudv\right)^{\frac{1}{p}} \left(\int_{[0,1]^2} F_Y^{-1}(u)^q\,|\gamma(u,v)|\,dudv\right)^{\frac{1}{q}},$$

with

$$\left(\int_{[0,1]^2} |\gamma(u,v)|\,dudv\right) = 2\beta(X,Y).$$

Therefore, the β*-mixing coefficient is just* $1/2$ *the total variation of the sign measure corresponding to* $\mathcal{G} : [0,1]^2 \to [0,1]$.

Another dependence condition is the weak dependence condition of Doukhan and Louhichi (1999). A sequence of random variables $(X_i)_{i\in\mathbb{N}}$ is called $(\theta,\mathfrak{F},\psi)$-weak dependent, if there exists a class $\mathfrak{F}$ of real valued functions, a sequence $(\theta_i)_{i\in\mathbb{N}}$ decreasing to zero at infinity, and a function ψ with arguments $(h,k,u,v)\in\mathfrak{F}^2\times\mathbb{N}^2$ such that for any u-tuple $(i_1,...,i_u)$ and any v-tuple $(j_1,...,j_v)$ with $i_1\leq\cdots\leq i_u< i_{u+r}\leq j_1\leq\cdots\leq j_v$ one has

$$\left|cov\left(h\left(X_{i_1},...,X_{i_u}\right),k\left(X_{j_1},...,X_{j_v}\right)\right)\right|\leq\psi\left(h,k,u,v\right)\theta_r$$

for all functions $h,k\in\mathfrak{F}$ that are defined respectively on $\mathbb{R}^u$ and $\mathbb{R}^v$. This dependence condition is of particular interest in our context.

Example 8 *Let h and k be bounded lipschitz functions, then, $\theta_r\to 0$ if and only if*

$$Law\left(X_{i_1},...,X_{i_u},X_{j_1},...,X_{j_v}\right)\to Law\left(X_{i_1},...,X_{i_u}\right)\times Law\left(X_{j_1},...,X_{j_v}\right),$$

which is metrized by (14) for the above tuple.

Other dependence conditions have been considered in the literature, but for economy of space they are not discussed (e.g. mixingales, McLeish, 1975, Andrews, 1988, Dedecker and Doukhan, 2003).

Because of absolute continuity of copulae, convergence under the metrics in the new probability space may be strictly weaker than the convergence requirement under the same metric in the original probability space.

Example 9 *Suppose X and Y are two random variables with copula $\lambda(1+\gamma)$. Consider the metric (12). Its equivalent in the original probability space is (15). However, if (15) is zero then (12) is zero as well, but not the other way around.*

4.3 A Decoupling Condition

Suppose $n:=(r-1)(p+q)+p$. Define the following index sets

$$H_j:=\{i:1+(j-1)(p+q)\leq i\leq p+(j-1)(p+q)\}$$

$j=1,...,r$, and $\mathcal{H}_r:=\bigcup_{j=1}^r H_j$. Define the following random variables $\mathbf{X}_j:=(X_i,i\in H_j)$ and $\boldsymbol{\xi}_j:=(\xi_i,i\in H_j)$, $i=1,...,r$. Suppose $Law(\boldsymbol{\xi}_1,....,\boldsymbol{\xi}_r)=Law(\mathbf{X}_1)\times\cdots\times Law(\mathbf{X}_r)$. For any measurable function f, we have the following decoupling result:

$$\mathbb{E}f\left(\mathbf{X}_1,...,\mathbf{X}_r\right)=\mathbb{E}f\left(\boldsymbol{\xi}_1,...,\boldsymbol{\xi}_r\right)+\mathbb{E}^{\lambda}\rho^{\mathcal{H}_r}\left(\mathbf{u}_1,...,\mathbf{u}_r\right)f\left(\mathcal{Q}_1\mathbf{u}_1,...,\mathcal{Q}_r\mathbf{u}_r\right),$$

where $\mathcal{Q}_j = \mathcal{Q}^{H_j}$ is the transform in (7), $j = 1, ..., r$, and $\mathbb{E}^\lambda$ is expectation with respect to the Lebesgue measure. With this notation it is natural to introduce the following condition.

Condition 10 (s) *The sequences of random variables $(X_i)_{i\in\mathbb{N}}$ satisfies Condition 10(s) if there are sequences p, q, r, such that $rp \sim n$, $p = o(q)$, $p, q, r \to \infty$ as $n \to \infty$, and*
for $s \in [1, \infty]$

$$\lim_n \left\| \rho^{\mathcal{H}_r} (\mathbf{u}_1, ..., \mathbf{u}_r) \right\|_{s,\lambda} \to 0,$$

or for $s = 0$, $\mathbf{u}_1, ..., \mathbf{u}_r$ converges weakly to the product of copulae, i.e. $\mathcal{R}^{\mathcal{H}_r} (\mathbf{u}_1, ..., \mathbf{u}_r) \to 0$ pointwise.

Remark 11 *We see that strong mixing is equivalent to Condition 10(0) if the joint distribution of the random variables is continuous (weak convergence with continuity implies uniform convergence); beta mixing is equivalent to Condition 10(1); when $\mathfrak{F}$ is the class of bounded Lipschitz functions, $(\theta, \mathfrak{F}, \psi)$-weak dependence is equivalent to Condition 10(0).*

4.4 Further Conditions

4.4.1 Condition 4(0) and the Monge-Wasserstein Distance

Let X and Y be random variables with values in a metric space $(\mathbf{S}, d)$. With the usual notation, P_X and P_Y are the laws of X and Y such that X and Y are integrable. Define $M(P_X, P_Y)$ to be the class of all joint distributions (i.e. of copulae) with marginal laws P_X and P_Y. The Wasserstein distance is defined as

$$W(P_X, P_Y) := \inf \left\{ \int_{S\times S} d(x, y)\, d\mu(x, y) : \mu \in M(P_X, P_Y) \right\}.$$

Suppose $(X_i)_{i\in\mathcal{H}_r}$ is a sequence of real random variables such that $\#\mathcal{H}_r = n = rp$. Suppose $(\xi_i)_{i\in\mathcal{H}_r}$ is a sequence such that

$$\begin{aligned} & law\left((\xi_i)_{i\in\mathcal{H}_r}\right) \\ = \; & law\left((X_i)_{i\in H_1}\right) law\left((X_i)_{i\in H_2}\right) \cdots law\left((X_i)_{i\in H_r}\right), \end{aligned}$$

constructed by regression construction (i.e. the two sequences are defined on the same probability space equipped with the Lebesgue measure). Writing P_X and P_ξ, respectively, for the laws of $X = (X_i)_{i\in\mathcal{H}_r}$ and $\xi = (\xi_i)_{i\in\mathcal{H}_r}$, and for $\rho^{\mathcal{H}_r}$ as defined above, by the Kantorovich-Rubinstein Theorem (e.g. Theorem 11.8.2 in Dudley, 2002),

$$W\left(P_X, P_\xi\right) = \sup\left\{\left|\mathbb{E}\rho^{\mathcal{H}_r} f\right| : \|f\|_L \leq 1\right\}.$$

Then, the following relationships hold

$$\left\|f\rho^{\mathcal{H}_r}\right\|_{BL_1} := \sup\left\{\left|\mathbb{E}\rho^{\mathcal{H}_r} f\right| : \|f\|_{BL} \leq 1\right\} \leq W\left(P_X, P_\xi\right), \tag{16}$$

and if the laws P_X and P_ξ are tight, by Lemma 2,

$$\Pr\left(d\left((X_i)_{i\in\mathcal{H}_r}, (\xi_i)_{i\in\mathcal{H}_r}\right) > \epsilon\right) \leq \epsilon \tag{17}$$

where

$$\left\|f\mathcal{Q}\rho^{\mathcal{H}_r}\right\|_{BL_1} \geq \epsilon^2/4. \tag{18}$$

Example 12 *Suppose* $\mathbf{S} = \mathbb{R}^n$ *so that* d *is the Euclidean norm* $\|...\|$*. By the regression construction, and Minkoski's inequality*

$$\begin{aligned} W\left(P_X, P_\xi\right) &\leq \mathbb{E}\left\|X - \xi\right\| \leq \sum_{j=1}^{r} \mathbb{E}\left[\sum_{i\in H_j}\left(X_i - \xi_i\right)^2\right]^{1/2} \\ &\leq \sum_{j=1}^{r}\sum_{i\in H_j}\left\|X_i - \xi_i\right\|_{1,\mu}. \end{aligned}$$

Define $\delta_{q,jp+i} := \left\|X_{(j-1)(p+q)+i} - \xi_{(j-1)(p+q)+i}\right\|_{1,\mu}$ *and*

$$\delta_q := \max_{\substack{1\leq i\leq p \\ 1\leq j\leq r}} \delta_{q,jp+i}.$$

Then

$$\Pr\left(d\left((X_i)_{i\in\mathcal{H}_r}, (\xi_i)_{i\in\mathcal{H}_r}\right) > 2\left(n\delta_q\right)^{1/2}\right) \leq 2\left(n\delta_q\right)^{1/2}.$$

Doukhan and Louhichi (1999) and Dedecker and Doukhan (2003) give examples of processes satisfing $\delta_q \to 0$.

4.4.2 Convergence of Measures of Concordance

Dall'Aglio (1966) has shown that if a random variable converges weakly to another and their copula converges to the maximal element (the Frechét upper bound, see Proposition 4(2)) then convergence holds in probability. Scarsini (1984) has characterized convergence to the Frechét upper bound via strong measures of concordance (see Scarsini, 1984, for definitions). Measures of concordance can then be used to describe convergence to independence (i.e. Condition 10(0), and actually Condition 10(1)).

Many multivariate measures of concordance (see Joe, 1989, for a list and description) for random variables $X_1, ..., X_K$ can be shown (up to an affine transformation) to reduce to the following

$$d_{X_1 \cdots X_K} := \mathbb{E} w \left(1 + \gamma^K\right),$$

where w is a suitably chosen monotone increasing transformation. For example

$$\phi^2_{X_1 \cdots X_K} := \mathbb{E}\left[\frac{pdf_{X_1 \cdots X_K}}{pdf_{X_1} \cdots pdf_{X_K}}\right] - 1,$$

and

$$\delta_{X_1 \cdots X_K} := \mathbb{E} \ln \left[\frac{pdf_{X_1 \cdots X_K}}{pdf_{X_1} \cdots pdf_{X_K}}\right],$$

are measures of dependence, which (by simple change of variables) can be rewritten as

$$\phi^2_{X_1 \cdots X_K} = \left\|\gamma^K\right\|_{1,\lambda_K}$$

and

$$\delta_{X_1 \cdots X_K} := E \ln \left[1 + \gamma^K\right].$$

Under suitable conditions on w, $d_{X_1 \cdots X_K} \to 0$ implies Condition 10(1), i.e. convergence in total variation.

4.4.3 Convergence of Moments

It is simple to show that any copula is completely determined by its moments (e.g., Theorem 30.1 in Billingsley, 1995). Moreover, all moments exists, hence convergence of all moments implies weak convergence (e.g. Theorem 30.2 in Billingsley, 1995). Moreover, under the following condition, only a finite number of moments is required to check Condition 10(0). Suppose $\mathbf{U}$ is the K random vector with given copula C^K, and $\Pr\left(\mathbf{U} \in \{0,1\}^{K-k} \times (0,1)^k\right) = 0$ for $k = 0, ..., K$, i.e. no mass at the upper edge of the hypercube. Since a K dimensional copula is defined on $[0,1]^K$, the high order moments of $\mathbf{U}$ converge to zero.

Example 13 (e.g. example 6.2 in Bradley, 1986) *Suppose* $(Z_t)_{t\in\mathbb{Z}}$ *is an iid Bernoulli sequence such that* $\Pr(Z = 0) = \Pr(Z = 1) = 1/2$. *Define the sequence* $(U_t)_{t\in\mathbb{Z}}$ *by*

$$U_t = \frac{1}{2} U_{t-1} + \frac{1}{2} Z_t,$$

so that the process has infinite moving average representation

$$U_t = \sum_{i=0}^{\infty} 2^{-1-i} Z_{t-i},$$

and U_t *is uniformly distributed on* $[0,1]$. *For* $T \geq 0$, *define* $\mathcal{F}_t^{t+T} := \sigma\left(U_s : t \leq s \leq t+T\right)$. *Then, the* $(U_t)_{t\in\mathbb{Z}}$ *sequence is known to have a non-trivial future tail sigma-algebra:* $\bigcap_{t=1}^{\infty} \mathcal{F}_t^{\infty}$ *coincides with* $\mathcal{F}_{-\infty}^{\infty}$, *because* U_t *is measurable with respect to* U_{t+1} *and by induction,* U_0 *is a.s. measurable with respect to* U_t. *To see this just notice that* $U_{t-1} \overset{a.s.}{=} 2U_t - \lfloor 2U_t \rfloor$ *(where* $\lfloor A \rfloor$ *is the integer part of* A*). As a consequence, the usual strong mixing coefficient* $\alpha\left(\mathcal{F}_{-\infty}^{0}, \mathcal{F}_t^{\infty}\right) \geq \alpha\left(\mathcal{F}_{-\infty}^{0}, \mathcal{F}_{-\infty}^{0}\right) = 1/4$ *because* $\mathcal{F}_{-\infty}^{0} \subset \mathcal{F}_t^{\infty}$. *On the other hand, we can check weak convergence calculating the mixed moments of* U_t *and* U_0. *Using the conditional moving average representation,*

$$U_t = 2^{-t} U_0 + \frac{1}{2} \sum_{i=0}^{t-1} 2^{-i} Z_{t-i},$$

it is easy to show that $\mathbb{E}U_tU_0 = \mathbb{E}U_t\mathbb{E}U_0 + O\left(2^{-t}\right)$, *and* $\mathbb{E}U_t^kU_0^j = \mathbb{E}U_t^k\mathbb{E}U_0^j + O\left(2^{-t}\right)$, $\forall k,j$. *Hence, for* $t \to \infty$ *the copula of* U_t *and* U_0 *converges to the product one. Hence, Condition 10*(0) *holds, i.e. pointwise convergence holds, but convergence uniformly over the sigma algebra does not.*

5 Applications

Throughout we will consider the following tuples

$$\left(X_1,...,X_p\right), \left(X_{p+q+1},...,X_{2p+q}\right), ..., \left(X_{(r-1)(p+q)+1},...,X_{(r-1)(p+q)+p}\right). \tag{19}$$

We define $(X_i)_{i\in H_j} = \left(X_{(j-1)(p+q)+1},...,X_{(j-1)(p+q)+p}\right)$, $j = 1,...,r$, and $\mathcal{H}_r = \bigcup_{j=1}^{r} H_j$. Moreover, $(\xi_i)_{i\in\mathcal{H}_r}$ will be used to denote the regression construction of a series of random variables such that

$$Law\left((\xi_i)_{i\in\mathcal{H}_r}\right) = Law\left((X_i)_{i\in H_1}\right) \times \cdots \times Law\left((X_i)_{i\in H_r}\right).$$

5.1 A Decoupling Inequality for Probabilities

We can show the following result which will be used in the sequel.

Lemma 14 *Suppose* $(X_i)_{i\in\mathcal{H}_r}$ *and* $(\xi_i)_{i\in\mathcal{H}_r}$ *are sequences of random variables as above and such that* $\#\mathcal{H}_r = n$*. Define* $S_n(X) := \sum_{j=1}^{n} X_{i_j}$*.*
i. (total variation)

$$\left|\Pr(S_n(X)\le\varepsilon)-\Pr(S_n(\xi)\le\varepsilon)\right|\le\left\|\rho^{\mathcal{H}_r}(\mathbf{u}_1,...,\mathbf{u}_r)\right\|_{1,\lambda},$$

where $\rho^{\mathcal{H}_r}(\mathbf{u}_1,...,\mathbf{u}_r)$ *is as in Condition 1(s)*
ii. (L_1 convergence)

$$\left|\Pr(S_n(X)\le\varepsilon)-\Pr(S_n(\xi)\le\varepsilon)\right|\le W(P_X,P_\xi),$$

where P_X *and* P_ξ *are, respectively, the joint probabilities of* $(X_i)_{i\in\mathcal{H}_r}$ *and* $(\xi_i)_{i\in\mathcal{H}_r}$*;*
iii. (Ky Fan convergence)

$$\left|\Pr(S_n(X)\le\varepsilon)-\Pr(S_n(\xi)\le\varepsilon)\right|\le 2\left(\left\|f\mathcal{Q}\rho^{\mathcal{H}_r}\right\|_{BL_1}\right)^{1/2}.$$

Proof. At first suppose that the copulae of $(X_i)_{i\in\mathcal{H}_r}$ and $(\xi_i)_{i\in\mathcal{H}_r}$ are unique and continuous. *i.* From Lemma 27 in the appendix, using the fact that $\|I\{x\le\varepsilon\}\|_{\infty,\lambda}\le 1$,

$$\Pr(S_n\le\varepsilon)\le\lambda\left(I_{\left\{\sum_{i=1}^{n}\mathcal{Q}_i u_i\le\varepsilon\right\}}\right)+\lambda(|\gamma^n|).$$

ii. and *iii.* From (17), (18), and (16) we have convergence in probability of $(X_i)_{i\in\mathcal{H}_r}$ to $(\xi_i)_{i\in\mathcal{H}_r}$ with the given bounds above. Notice that the usual condition: $\forall\varepsilon$ such that

$$\Pr(S_n(X)\in\partial(-\infty,\varepsilon])=\Pr(S_n(\xi)\in\partial(-\infty,\varepsilon])$$

is not required as the copula is absolutely continuous, hence no mass on the boundary. ■

Remark 15 *By absolute continuity of the copula, the same result follows (e.g. Theorem 8, p.173, in Pollard, 2002,) for* $S_n(f(X))$*, where* f *is a measurable upper or lower semicontinuous function, respectively bounded from above or below (see Definition 20 and Lemma 26 in the appendix for further details).*

5.2 A CLT

We can state and prove the following CLT result for dependent random variables.

Theorem 16 *Suppose* $(X_i)_{i\in\mathbb{Z}}$ *is a mean zero strictly stationary sequence of random variables in* $\left(\mathbb{R}^{\mathbb{Z}},\mathcal{B}\left(\mathbb{R}^{\mathbb{Z}}\right),\mu\right)$*. Set* $S_n(X) := \sum_{i=1}^{n} X_i$*. Suppose there exists a sequence* $B_n\uparrow\infty$ *such that* $\lim_n var\left(B_n^{-1}S_n(X)\right)=1$*. Then* $B_n^{-1}S_n \xrightarrow{d} Y$ *where*

$\mathbb{E}\exp\{itY\} = \exp\{it/2\}$ *in either of the following two cases:*
i. Condition 10(0) *is satisfied and* $B_n^{-2}|S_n(X)|^2$ *is asymptotically uniformly integrable;*
ii. Condition 10(1) *is satisfied and* $B_n^{-2}|S_n(X)|^2$ *is integrable.*

Remark 17 *The result of the Theorem implies that* $B_n/n = k_n > 0$ *where* k_n *is slowly varying at infinity.*

Proof. Choose sequences r_n, p_n and q_n such that $n = (p_n + q_n) r_n$. For simplicity let us drop the subscript n. Restrict p and q such that $q = \lfloor p^\alpha \rfloor$, $\alpha \in (0,1)$. Define the following index sets

$$H_j := \{i : 1 + (j-1)(p+q) \leq i \leq p + (j-1)(p+q)\}$$

$$H'_j := \{i : jp + (j-1)(p+q) \leq i \leq j(p+q)\}$$

$j = 1, ..., r$. Write $S_n = W_r + W'_r$, where $W_r := \sum_{j=1}^r Z_j$, $W'_r := \sum_{j=1}^r Z'_j$, and $Z_j := \sum_{i\in H_j} X_i$, $Z'_j := \sum_{i\in H'_j} X_i$. Define $\sigma_n^2 = var(S_n)$. We now notice that Lemma 2.6 in Berkes and Philipp (1998) is also valid under Condition $10(s)$. Then, for $\sigma_m^2 > 0$, $\sigma_n^2/\sigma_m^2 \geq \kappa(n/m)$, where κ is some function such that $\lim_j \kappa(j) = \infty$. Let $\left((\xi'_i)_{i\in H_j} : j = 1, ..., r\right)$ be sequences such that

$$Law\left((\xi'_i)_{i\in H'_j} : j = 1, ..., r\right) = Law(X_i)_{i\in H'_1} \times \cdots \times Law(X_i)_{i\in H'_r}.$$

Define $\mathcal{H}'_r := \bigcup_{j=1}^r H'_j$, and $F_{\xi'_i}$ to be the distribution function of ξ'_i. We only consider case ii; case i is actually easier. From Condition $10(1)$, Lemma 28 in the appendix and stationarity

$$\begin{aligned}
var(W'_r) &= var\left(\sum_{j=1}^r Z'_j\right) \\
&= \left\|\sum_{j=1}^r\left(\sum_{i\in H'_j}\xi'_i\right)\right\|_{2,\mu}^2 + M_n\left\|\rho^{\mathcal{H}'_r}\right\|_{1,\lambda} \\
&\quad + \left\|\rho^{\mathcal{H}'_r}\right\|_{\infty,\lambda}\left\|\sum_{j=1}^r\left(\sum_{i\in H'_j}\xi'_i\right)I_{\left\{\left|\sum_{j=1}^r\left(\sum_{i\in H'_j}\xi'_i\right)\right|^2 > M_n\right\}}\right\|_{2,\mu}^2 \\
&= r\mathbb{E}\left\|Z'_j\right\|_{2,\mu}^2 + M_n\left\|\rho^{\mathcal{H}'_r}\right\|_{1,\lambda} + r\left\|\rho^{\mathcal{H}'_r}\right\|_{\infty,\lambda}\left\|Z'_j I_{\left\{\left|Z'_j\right|^2 > M_n/r^2\right\}}\right\|_{2,\mu}^2 \\
&= \mathrm{I} + \mathrm{II} + \mathrm{III}.
\end{aligned}$$

We need to show that I$\asymp$II+III. Define

$$\left\|Z'_j\right\|_{2,\mu}^2 = \sigma_q^2,$$

and choose $M_n \asymp r^2\sigma_q^2 t_n$, where t_n is an increasing sequence. Then, for n large enough, III$\asymp r\sigma_q^2$, eventually. To control II we choose r, p and α such that $\left\|\rho^{\mathcal{H}_r'}\right\|_{1,\lambda} \asymp (rt_n)^{-1}$, for t_n increasing as slowly as possible. Hence, II+III$\asymp$I, so that

$$var\left(W_r'\right) \lesssim r\sigma_q^2.$$

and

$$\left|\frac{W_n'}{r\sigma_p}\right| \overset{p}{\lesssim} \frac{\sigma_q}{\sigma_p} \leq 1/\kappa\left(p^{1-\alpha}\right)^{1/2} \to 0.$$

(Case i holds by an application of Lemma 26 in the appendix.)

Choosing $B_n^2 = r\sigma_p^2$, and noticing that $\left|\exp\{ix\} - 1\right| \leq |x|$,

$$\left|\mathbb{E}\exp\left\{itB_n^{-1}S_n\right\} - \mathbb{E}\exp\left\{itB_n^{-1}W_r\right\}\right| \lesssim t/\kappa\left(p^{1-\alpha}\right)^{1/2}.$$

Let $\left(\left(\xi_i\right)_{i\in H_j} : j = 1, ..., r\right)$ be sequences such that

$$Law\left(\left(\xi_i\right)_{i\in H_j} : j = 1, ..., r\right) = Law\left(X_i\right)_{i\in H_1} \times \cdots \times Law\left(X_i\right)_{i\in H_r}.$$

Define $\mathcal{H}_r := \bigcup_{j=1}^r H_j$. Then,

$$\begin{aligned} & \left|\mathbb{E}\exp\left\{itB_n^{-1}W_r\right\} - \prod_{j=1}^r \mathbb{E}\exp\left\{itB_n^{-1}Z_j\right\}\right| \\ = \; & \mathbb{E}^\lambda \rho^{\mathcal{H}_r}\left(\mathbf{u}_1, ..., \mathbf{u}_r\right)\exp\left\{itB_n^{-1}\sum_{j=1}^r\left(\sum_{i\in H_j} F_{\xi_i}^{-1}\left(u_i\right)\right)\right\} \\ \leq \; & \left\|\rho^{\mathcal{H}_r}\left(\mathbf{u}_1, ..., \mathbf{u}_r\right)\right\|_{1,\lambda}, \end{aligned}$$

because $\exp\{ix\} \leq 1$, $\forall x$. (Again, apply Lemma 26 for case i). By the usual Lindeberg CLT for independent random variables with $s_n^2 := \sum_{i=1}^n var\left(X_{n,i}\right)$, for some constant C,

$$\sup_x\left|\Pr\left(S_n/s_n < x\right) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x \exp\left(-t^2/2\right)dt\right| \leq C\left(L_n(\epsilon) + \epsilon\right),$$

where

$$L_n(\epsilon) := s_n^{-2}\sum_{j=1}^n \int_{|x|\geq\epsilon s_n} x^2 dF_j(x).$$

The theorem is proved if we show that $\left\|Z_j^2/\sigma_p^2 I_{\left\{|Z_j/\sigma_p|>\sqrt{r}\epsilon\right\}}\right\|_{1,\mu} \to 0$, $\forall\epsilon > 0$, which is the case, as $\left\{|Z_j/\sigma_p| > \sqrt{r}\epsilon\right\} \downarrow \emptyset$ as $n \uparrow \infty$ for any sequence $\epsilon_n \to \epsilon$ such that $\epsilon_n\sqrt{r} \to \infty$. ■

5.3 Stochastic Equicontinuity Results

Consider the family of measurable functions $\mathfrak{F}$ from $\mathbf{S}$ to $\mathbb{R}$. Suppose d is a pseudometric on $\mathfrak{F}$. The covering number is defined as

$$N(\epsilon, \mathfrak{F}, d) := \min\left\{m : \exists f_1, ..., f_m \in \mathfrak{F} \text{ such that } \sup_{f\in\mathfrak{F}} \min_{1\leq j\leq m} d(f, f_j) \leq \epsilon\right\},$$

and $\ln N(\epsilon, \mathfrak{F}, d)$ is called the entropy. We also define the class of functions

$$\mathfrak{F}'(\epsilon, d) := \{f - g : f, g \in \mathfrak{F}, d(g, f) \leq \epsilon\},$$

and use $\|...\|_{\mathfrak{F}'(\epsilon,d)}$ as a short notation for $\sup_{f,g\in\mathfrak{F},d(g,f)\leq\epsilon} |...|$. A class of subsets $\mathcal{C}$ on a set $\mathbf{S}$ is called a V-C class if there exist a polynomial $l(...)$ such that, for every set of N points in $\mathbf{S}$, the class $\mathcal{C}$ picks at most $l(N)$ distinct subsets. A class of functions $\mathfrak{F}$ is called a V-C subgraph class if the graphs of the functions in $\mathfrak{F}$ form a V-C class of sets. That is, if we define the subgraph of a real valued function f on $\mathbf{S}$ as the following subset G_f on $\mathbf{S} \times \mathbb{R}$:

$$G_f = \{(s, t) : 0 \leq t \leq f(s) \text{ or } f(s) \leq t \leq 0\}$$

the class $G_f : f \in \mathfrak{F}$ is V-C class of set $\mathbf{S} \times \mathbb{R}$.

Suppose $(X_i)_{i\in\mathbb{Z}}$ is a sequence of mean zero random variables such that $(X_i)_{1\leq i\leq n}$ is adapted to the Borel sigma algebra $\mathcal{B}(\mathbb{R}^n)$. For $f \in \mathfrak{F}$ with envelope F such that $\|F\|_{s,\mu} < \infty$, for some $2 < s < \infty$. consider the following empirical process $v_n(f) := n^{-\frac{1}{2}} \sum_{i=1}^n (f(X_i) - \mathbb{E}f(X_i))$. Then define the following metric $d(f, g) = \|f - g\|_{2,\mu}$.

Condition 18 *Let $(Z_i)_{i\in\mathbb{N}}$ be a sequence of mean zero random variables with uniformly integrable second moment. Then, the following hold*

$$\left\|n^{-1/2} S_n(Z)\right\|_{2,\mu} \leq b_s M_s,$$

where $b_s > 0$ is a constant depending on s only, $M_s := \max_{i\in\mathbb{N}} \|Z_i\|_{s,\mu}$, and

$$s := \sup\left\{t > 2 : \max_{i\in\mathbb{N}} \|Z_i\|_{t,\mu} \vee b_t < \infty\right\}.$$

This condition may hold in several cases:

Example 19 *Suppose $(Z_i)_{i\in\mathbb{N}}$ is as above, but stationary and strong mixing with mixing coefficients bounded above by the non increasing function $\alpha(\lfloor t \rfloor)$, $t \in \mathbb{R}_+$. Then, for $1/p + 1/q = 1$,*

$$\left\|n^{-1/2} S_n(Z)\right\|_{2,\mu}^2 \leq M_p^2 \left(1 + 8 \sum_{i=1}^n \alpha(i)^{1/q}\right),$$

and $b_p = \lim_n \left(1 + 8\sum_{i=1}^{n} \alpha(i)^{1/q}\right)^{1/2}$. *Condition 18 is satisfied if there exists a* $p > 2$ *such that the right hand side is finite.*

Definition 20 *A function* $f : \mathbf{S} \to \mathbb{R}$ *is called lower semicontinuous if* $\{x \in \mathbf{S} : f(x) > t\}$ *is an open set for each fixed* t. *A function* $f : \mathbf{S} \to \mathbb{R}$ *is called upper semicontinuous if* $\{x \in \mathbf{S} : f(x) < t\}$ *is an open set for each fixed* t.

Remark 21 *If* f *is lower semicontinuous, then,* $-f$ *is upper semicontinuous. So we will just consider lower semicontinuous functions.*

An example of lower semicontinuous function is the indicator of an open set. We have the following.

Theorem 22 *Suppose* $\mathfrak{F}$ *is a class of measurable lower semicontinuous functions with envelope function* F *satisfying* $\mathbb{E}F^s < \infty$, $2 < s < \infty$, *and, for any probability measure* Q *and finite constants* v *and* b,

$$N\left(x, \mathfrak{F}, \|...\|_{2,Q}\right) \leq \left(\frac{b\,\|F\|_{2,Q}}{x}\right)^v,$$

such that

$$\int_0^{\delta} \sqrt{\ln N\left(x, \mathfrak{F}, \|...\|_{2,Q}\right)}dx < \infty.$$

Suppose $(X_i)_{i\in\mathbb{Z}}$ *is a mean zero sequence of strictly stationary random variables in* $\left(\mathbb{R}^{\mathbb{Z}}, \mathcal{B}\left(\mathbb{R}^{\mathbb{Z}}\right), \mu\right)$ *such that Condition 10(0) is satisfied with* $p = o(n/\ln n)$. *Suppose that* $(f(X_i))_{i\in\mathbb{Z}}$ *satisfies Condition 18. Then, to any* $\eta > 0$ *and* $\varepsilon > 0$, *there exists an* $\epsilon > 0$ *such that*

$$\lim_{n\to\infty} \sup \Pr\left(\|v_n(f)\|_{\mathfrak{F}'(\epsilon,d)} > \eta\right) < \varepsilon.$$

Proof. Define the following sets

$$H_j := \{i : 1 + (2p_n)(j-1) \leq i \leq p_n + 2p_n(j-1)\}$$

$$H'_j := \{i : 1 + p_n j \leq i \leq 2p_n j\}$$

(in this proof use p_n instead of p to avoid ambiguities in the argument of the proof). We show that we can use sequences $(\xi_i)_{i\in H_j}$ $(j = 1, ..., r)$, such that

$$law\left((\xi_i)_{i\in H_1}, ..., (\xi_i)_{i\in H_r}\right) = law\left((X_i)_{i\in H_1}\right) \times \cdots \times law\left((X_i)_{i\in H_r}\right).$$

(In this case, we choose $p_n \asymp q_n$, which is weaker than what is used in Condition 10.) For $x, y \in \mathbb{R}$, suppose $A := \{x, y : x + y > c, c > 0\}$ and $B := \{x : x > c/2, c > 0\} \cup \{y : y > c/2, c > 0\}$. Then, $A \subset B$. By this remark

$$\begin{aligned} &\lim_{n\to\infty} \Pr\left(\|v_n(f)\|_{\mathfrak{F}'(\epsilon,d)} > \eta\right) \\ \leq\ & 2\lim_{n\to\infty} \Pr\left(\left\|n^{-\frac{1}{2}} \sum_{j=1}^{r} \sum_{i\in H_j} f(X_i)\right\|_{\mathfrak{F}'(\epsilon,d)} > \eta/4\right). \end{aligned}$$

We use the regression construction to construct a sequence $(\xi_i)_{i\in\mathcal{H}_r}$, $\mathcal{H}_r := \bigcup_{j=1}^{r} H_j$, with the properties mentioned above. Then, by Condition 10(0), Lemma 14 and the continuous mapping theorem, for any a.s. continuous function f, there exists a sequence $\tau_{n,p} \to 0$, as $p \to \infty$, such that

$$\begin{aligned} &\lim_{n\to\infty} \Pr\left(\left\|n^{-\frac{1}{2}} \sum_{j=1}^{r} \sum_{i\in H_j} f(X_i)\right\|_{\mathfrak{F}'(\epsilon,d)} > \eta/4\right) \\ \leq\ & \Pr\left(\left\|n^{-\frac{1}{2}} \sum_{j=1}^{r} \sum_{i\in H_j} f(\xi_i)\right\|_{\mathfrak{F}'(\epsilon,d)} > \eta/4\right) + \tau_{n,p}. \end{aligned} \tag{20}$$

Now notice that Lemma 2.1 in Arcones and Yu (1994) establishes equicontinuity in probability for empirical processes based on underlying beta mixing sequences. Close inspection of the proof of their Lemma 2.1 shows that their result is still valid if we can decouple into a sequence of block independent random variables, with a suitable block size $p_n \to \infty$ as $n \to \infty$. In fact, using their notation, set their $a_n = p_n$, and their $p = s$. Then, they choose a truncation region $\delta_n n^{1/(2s-2)}$ $(s > 2)$, for some sequence δ_n going to zero slowly enough, so that $\delta_n n^{1/(2s-2)} \to \infty$. More generally, their truncation region can be written as $\delta_n b_n$ and only needs to satisfy $p_n b_n^p \asymp n/\ln n$. By Condition 18, $s > 2$, and we have $b_n \asymp [n/(p_n \ln n)]^{1/s} \to \infty$, by the conditions in Theorem 22. Finally, their Lemma 2.1 only requires our Condition 18. Hence, the result follows setting their $c_{p,\beta} = b_s$ and closely applying all their arguments. ■

We can also establish equicontinuity in probability under bracketing conditions. Given two functions l and u, the bracket $[l, u]$ is the set of all functions f such that $l \leq f \leq u$. Then, an ϵ-bracket is a bracket $[l, u]$ with $d(l, u)$, and the bracketing number $N_{[]}(\epsilon, \mathfrak{F}, d)$ is the minimum number of ϵ-brackets needed to cover $\mathfrak{F}$. Its logarithm is called the entropy with bracketing. It is possible to avoid the restrictive

Condition 10 by introducing a new metric. Suppose $(X_i)_{i\in\mathbb{N}}$ is a sequence of random variables. Then,

$$\|f\|_{2,\gamma} := \mathbb{E}^{\lambda}\left(f\mathcal{Q}(U)\, f\mathcal{Q}(V)\left(\sum_{i\geq 1}\tilde{\gamma}^{\{1,i\}}(U,V)\right)\right),$$

where $\tilde{\gamma}^{\{1,1\}}(u,v) := \delta(u-v)$, δ is the Dirac measure at zero, and

$$\tilde{\gamma}^{\{1,i\}}(U,V) := sign\left(\mathbb{E}^{\lambda} f\mathcal{Q}(U)\, f\mathcal{Q}(V)\,\gamma^{\{1,i\}}(U,V)\right)\gamma^{\{1,i\}}(U,V),$$

where $\gamma^{\{1,i\}}(u,v)$ $i \geq 2$ is as in (4) for the random variables $(X_j)_{j\in\{1,i\}}$. This metric bounds from above the variance of the sum $S_n(f(X))$. In particular, for $(X_i)_{i\in\mathbb{N}}$ with associated sign measure having density $\gamma^{\mathbb{N}}$, it computes the following

$$\|f\|_{2,\gamma} = \left(\mathbb{E}f(X) + \sum_{i>1}\left|\mathbb{E}f(X_1)\, f(X_i) - \mathbb{E}f(X_1)\,\mathbb{E}f(X_i)\right|\right)^{1/2}.$$

Remark 23 *Suppose $(X_i)_{i\in\mathbb{N}}$ is a stationary sequence of random variables. If $\mathbb{E}f(X_1)^2 < \infty$, then, $f(X_1)\, f(X_i)$, is uniformly integrable for $i > 1$. This follows from domination: $f(X_1)\, f(X_i) < 2f(X_1)^2$.*

We have the following.

Theorem 24 *Suppose $(X_i)_{i\in\mathbb{N}}$ is a sequence of stationary real random variables with associated sign measure having density $\gamma^{\mathbb{N}}$ and satisfying Condition 5(0). Suppose $\mathfrak{F}$ is a class of measurable lower semicontinuous functions with envelope F such that such $\|F\|_{2,\gamma} < \infty$ and*

$$\int N_{[]}\left(\epsilon, \mathfrak{F}, \|...\|_{2,\gamma}\right) d\epsilon < \infty.$$

Then, $v_n(f)$ is equicontinuous in probability.

Proof. From Condition 10(0) and semicontinuity of the class $\mathfrak{F}$, we refer to the previous proof and 20, using the same notation to obtain

$$\Pr\left(\|v_n(f)\|_{\mathfrak{F}'(\epsilon,d)} > \eta\right) \leq \Pr\left(\left\|n^{-\frac{1}{2}}\sum_{j=1}^{r}\sum_{i\in H_j} f(\xi_i)\right\|_{\mathfrak{F}'(\epsilon,d)} > \eta/4\right) + \tau_{n,p},$$

for some sequence $\tau_{n,p} \to 0$, as $p \to \infty$. By Bernstein inequality

$$\Pr\left(\left|\sum_{j=1}^{r}\sum_{i\in H_j} f(\xi_i)\right| > x\right) \leq 2\exp\left\{-\frac{1}{2}\frac{x^2}{\left\|\sum_{j=1}^{r}\sum_{i\in H_j} f(\xi_i)\right\|_{2,\mu}^2 + xp\,\|f(\xi_i)\|_{\infty,\mu}/3}\right\}. \tag{22}$$

Then, for any finite set of functions $\mathfrak{G}$, (22) together with Lemma 2.2.10 in van der Vaart and Wellner (2000) imply that

$$\left\|\frac{1}{n^{1/2}}\sum_{j=1}^{r}\sum_{i\in H_j} f(\xi_i)\right\|_{\mathfrak{G}} \lesssim \max_{f\in\mathfrak{G}}\|f(\xi_i)\|_{\infty,\mu}\frac{p\ln(\#\mathfrak{G})}{n^{1/2}}+\max_{f\in\mathfrak{G}}\left\|\sum_{j=1}^{r}\sum_{i\in H_j} f(\xi_i)\right\|_{2,\mu}\left(\frac{\ln(\#\mathfrak{G})}{n}\right)^{1/2}.$$

We now prepare for an application of Theorem 8 in Pollard (2002). Using the new norm,

$$\left\|\sum_{j=1}^{r}\sum_{i\in H_j} f(\xi_i)\right\|_{2,\mu} \leq n^{1/2}\|f(\xi_i)\|_{2,\gamma},$$

so that

$$\begin{aligned}\left\|\frac{1}{n^{1/2}}\sum_{j=1}^{r}\sum_{i\in H_j} f(\xi_i)\right\|_{\mathfrak{G}} &\lesssim \|f(\xi_i)\|_{\infty,\mu}\frac{p\ln(\#\mathfrak{G})}{n^{1/2}}+\|f(\xi_i)\|_{2,\gamma}\ln(\#\mathfrak{G})^{1/2}\\ &= \epsilon l_{\mathfrak{G}}\left(1+b\frac{pl_{\mathfrak{G}}}{\epsilon n^{1/2}}\right), \qquad (23)\end{aligned}$$

setting $b := \|f(\xi_i)\|_{\infty,\mu}$ and $l_{\mathfrak{G}} := \ln(\#\mathfrak{G})^{1/2}$. Define $\bar{\gamma}_i := \left|\mathbb{E}^{\lambda} uv\gamma_i(u,v)\right|$. Then choose the following truncation region $\left\{|f| > \|f\|_{2,\gamma}/\sqrt{R(x)}\right\}$, where

$$R(x) := \sum_{i\geq 1}\int_0^{\bar{\gamma}_x\wedge\bar{\gamma}_i} du \geq x\bar{\gamma}_x.$$

By Chebishev's inequality and the fact that $\|f\|_{2,\mu} \leq \|f\|_{2,\gamma}$,

$$\mathbb{E}|f|\left\{|f| > \|f\|_{2,\gamma}/\sqrt{R(x)}\right\} \leq \|f\|_{2,\gamma}\sqrt{R(x)}. \qquad (24)$$

Finally, choose $\epsilon/b = 1/\sqrt{R(p)} \leq 1/\sqrt{p\bar{\gamma}_p}$, and $l_{\mathfrak{G}}/\sqrt{n} = \sqrt{\bar{\gamma}_p/p}$. With this choice

$$\left\|\frac{1}{n^{1/2}}\sum_{j=1}^{r}\sum_{i\in H_j} f(\xi_i)\right\|_{\mathfrak{G}} \lesssim \epsilon l_{\mathfrak{G}}.$$

Then, we apply a chaining argument with adaptive truncation. To balance the terms, we choose p such that $\tau_{p,n} \asymp \epsilon l_{\mathfrak{G}}$. Then, assumptions i-vi in Pollard (2002) hold. Assumptions i-iii are automatically satisfied. Assumption iv is satisfied using the maximal inequality (23). Assumption v is satisfied using the L_1 norm. Assumption vi is satisfied using (24). For the final details we refer the reader to Pollard (2002). ∎

5.4 Invariance Principle for Empirical Processes

Theorems 22 and 24 imply an invariance principle for empirical processes. We only state the result in terms of Theorem 24.

Theorem 25 *Suppose $(X_i)_{i\in\mathbb{N}}$ is a sequence of stationary real random variables with associated sign measure having density $\gamma^{\mathbb{N}}$ and satisfying Condition 10(0), and such that var $(S_n(X))/n$ is uniformly integrable. Suppose $\mathfrak{F}$ is a class of measurable lower semicontinuous functions with envelope F such that such $\|F\|_{2,\gamma} < \infty$ and*

$$\int N_{[]}\left(\epsilon, \mathfrak{F}, \|...\|_{2,\gamma}\right) d\epsilon < \infty.$$

Then $v_n(f)$ is Donsker.

Proof. It directly follows from Theorems 16 and 24, and the fact that the integral condition on the bracketing numbers implies that the class $\mathfrak{F}$ is totally bounded. This implies a tight Gaussian limit (e.g. example 1.5.10 in van der Vaart and Wellner, 2000). ■

A Additional Lemmata

Lemma 26 *Let μ and ν be probability measures as in Theorem 1, and f a uniformly integrable measurable map which is lower semi continuous with values in $\mathbb{R}$. If $\mu \to \nu$ with respect to the class of continuous bounded functions, then*

$$\mu f = \nu f + \lambda(\gamma f \mathcal{Q}) \to \nu f.$$

Proof. Use the decomposition $\mu f = \nu f + \lambda\gamma f\mathcal{Q}$. Suppose f is continuous, then, weak convergence implies $\lambda\gamma f\mathcal{Q} \to 0$ by uniform integrability (e.g. Theorem 1.11.3 in van der Vaart and Wellner, 2000). If the function is only lower semi continuous, then by uniform integrability, we can just assume f to be bounded from below by a finite constant M. Use an approximation with respect to Lipschitz bounded functions (e.g. simply extend Lemma 7 in Pollard, 2002a, to the multivariate case). Then the result follows by monotone approximation and continuity of the copula. ■

Lemma 27 *Consider the following measurable map $f : [0,1]^K \to \mathbb{R}$ and let μ be a probability measure in $[0,1]^K$ absolutely continuous with respect to the Lebesgue measure λ. Let γ be defined as in (4). Then,*

$$\mu f = \lambda f + (\mu - \lambda) f \leq \lambda f + \|\gamma\|_p \|f\|_q,$$

for $1/p+1/q=1$.

Proof. Use Hölder's inequality. ■

Lemma 28 *Let* $X_1,...,X_n$ *be a sequence of random variables, and* $S_n(X) := \sum_{i=1}^n X_i$*. Then, for any* $M_n>0$,

$$\mathbb{E}\left(S_n(X)^2\right) = \mathbb{E}\left|S_n(\xi)\right|^2 + M_n\|\gamma^n\|_{1,\lambda} + \|\gamma^n\|_{\infty,\lambda}\,\mathbb{E}\left|S_n(\xi)\right|^2 I_{\left\{|S_n(\xi)|^2>M_n\right\}},$$

where $\xi_1,...,\xi_n$ *is an independent sequence with same marginal laws as* $X_1,...,X_n$*, and* $\gamma=\gamma^n$ *is as in (4).*

Proof. Consider the representation

$$\mathbb{E}\left(S_n(X)^2\right) = \lambda\left(S_n(\mathcal{Q}u)^2\right) + \lambda\left(\gamma^n S_n(\mathcal{Q}u)^2\right),$$

and rewrite

$$S_n(\mathcal{Q}u)^2 = \left|S_n(\mathcal{Q}u)\right|^2 I_{\left\{|S_n(\mathcal{Q}u)|^2\le M_n\right\}} + \left|S_n(\mathcal{Q}u)\right|^2 I_{\left\{|S_n(\mathcal{Q}u)|^2> M_n\right\}},$$

substitute in the previous display and use Hölder's inequality on the first term on the right hand side of the last display. ■

References

[1] Andrews, D.W.K. (1988) Laws of Large Numbers for Dependent Non-Identically Distributed Random Variables. Econometric Theory 4, 458-467.

[2] Berkes, I. and W. Philipp (1998) Limit Theorems for Mixing Sequences without Rate Assumptions. Annals of Probability 26, 805-831.

[3] Billingsley, P. (1995) Probability and Measure. New York: Wiley.

[4] Dall'Aglio G. (1956) Sugli Estremi dei Momenti delle Funzioni di Ripartizione Doppia. Annali Scuola Normale Superiore di Pisa 10, 35-74.

[5] Dall'Aglio G. (1966) Osservazioni sulla Convergenza in Distribuzione e in Probabilità. Giornale dell'Istituto Italiano degli Attuari 29, 255-269.

[6] Dedecker, J. and P. Doukhan (2003) A New Covariance Inequality and Applications. Stochastic Processes and their Applications 106, 63-80.

[7] Devore, R.A. and G.G. Lorentz (1993) *Constructive Approximations*. Berlin: Springer.

[8] Doukhan, P., P. Massart and E. Rio (1994). The Functional Central Limit Theorem for Strongly Mixing Processes. Annales de l'Institut Henri Pointcaré, Probabilités et Statistiques 30, 63-82.

[9] Doukhan, P. and S. Louhichi (1999) A New Weak Dependence Condition and Applications to Moment Inequalities. Stochastic Processes and Applications 84, 313-342.

[10] Dudley, R.M. (2002) Real Analysis and Probability. Cambridge: Cambridge University Press.

[11] Genest, C., J.J. Quesada Molina and J.A. Rodriguez Lallena (1995). De l'Impossibilité de Construire des Loi à Marges Multidimesionelles Données à Partir de Copules. Comptes Rendus des Séances de l'Académie des Sciences. Série I. Mathématique 320, 723-726.

[12] Joe, H. (1989) Relative Entropy Measures of Multivariate Dependence. Journal of the American Statistical Association 84, 157-164.

[13] Li, H., M. Scarsini and M. Shaked (1996) Linkages: a Tool for the Construction of Multivariate Distributions with Given Nonoverlapping Multivariate Marginals. Journal of Multivariate Analysis 56, 20-41.

[14] McLeish, D.L. (1975) A Maximal Inequality and Dependent Strong Laws. Annals of Probability 3, 829-839.

[15] Pollard, D. (2002) Maximal Inequalities via Bracketing with Adaptive Truncation. Annales de l'Institut Henri Pointcaré, Probabilités et Statistiques 38, 1039-1052.

[16] Pollard, D. (2002a) A User's Guide to Measure Theoretic Probability. Cambridge: Cambridge University Press.

[17] Rio, E. (1993) Covariance Inequalities for Strongly Mixing Processes. Annales de l'Institut Henri Pointcaré, Probabilités et Statistiques 29, 587-597.

[18] Rio, E. (2000) Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants. Paris: Springer.

[19] Rüschendorf, L. (1985) Construction of Multivariate Distributions with Given Marginals. Annals of the Institute of Statistical Mathematics 28, 225-233.

[20] Rüschendorf, L. and V. de Valk (1993) On Regression Representation of Stochastic Processes. Stochastic Processes and their Applications 46, 183-198.

[21] Scarsini, M. (1984) Strong Measures of Concordance. Stochastica 8, 201-218.

[22] Scarsini, M. (1989) Copulae of Probability Measures on Product Spaces. Journal of Multivariate Analysis 31, 201-219.

[23] Sklar, A. (1973) Random Variables, Joint Distribution Functions, and Copulas. Kybernetika 9, 449-460.

[24] Van der Vaart, A. and J.A. Wellner (2000) Weak Convergence of Empirical Processes. Springer Series in Statistics. New York: Springer.

[25] Viennet, G. (1997) Inequalities for Absolutely Regular Sequences: Application to Density Estimation. Probability Theory and Related Fields 107, 467-492.

List of other working papers:

2004

1. Xiaohong Chen, Yanqin Fan and Andrew Patton, Simple Tests for Models of Dependence Between Multiple Financial Time Series, with Applications to U.S. Equity Returns and Exchange Rates, WP04-19
2. Valentina Corradi and Walter Distaso, Testing for One-Factor Models versus Stochastic Volatility Models, WP04-18
3. Valentina Corradi and Walter Distaso, Estimating and Testing Sochastic Volatility Models using Realized Measures, WP04-17
4. Valentina Corradi and Norman Swanson, Predictive Density Accuracy Tests, WP04-16
5. Roel Oomen, Properties of Bias Corrected Realized Variance Under Alternative Sampling Schemes, WP04-15
6. Roel Oomen, Properties of Realized Variance for a Pure Jump Process: Calendar Time Sampling versus Business Time Sampling, WP04-14
7. Richard Clarida, Lucio Sarno, Mark Taylor and Giorgio Valente, The Role of Asymmetries and Regime Shifts in the Term Structure of Interest Rates, WP04-13
8. Lucio Sarno, Daniel Thornton and Giorgio Valente, Federal Funds Rate Prediction, WP04-12
9. Lucio Sarno and Giorgio Valente, Modeling and Forecasting Stock Returns: Exploiting the Futures Market, Regime Shifts and International Spillovers, WP04-11
10. Lucio Sarno and Giorgio Valente, Empirical Exchange Rate Models and Currency Risk: Some Evidence from Density Forecasts, WP04-10
11. Ilias Tsiakas, Periodic Stochastic Volatility and Fat Tails, WP04-09
12. Ilias Tsiakas, Is Seasonal Heteroscedasticity Real? An International Perspective, WP04-08
13. Damin Challet, Andrea De Martino, Matteo Marsili and Isaac Castillo, Minority games with finite score memory, WP04-07
14. Basel Awartani, Valentina Corradi and Walter Distaso, Testing and Modelling Market Microstructure Effects with an Application to the Dow Jones Industrial Average, WP04-06
15. Andrew Patton and Allan Timmermann, Properties of Optimal Forecasts under Asymmetric Loss and Nonlinearity, WP04-05
16. Andrew Patton, Modelling Asymmetric Exchange Rate Dependence, WP04-04
17. Alessio Sancetta, Decoupling and Convergence to Independence with Applications to Functional Limit Theorems, WP04-03
18. Alessio Sancetta, Copula Based Monte Carlo Integration in Financial Problems, WP04-02
19. Abhay Abhayankar, Lucio Sarno and Giorgio Valente, Exchange Rates and Fundamentals: Evidence on the Economic Value of Predictability, WP04-01

2002

1. Paolo Zaffaroni, Gaussian inference on Certain Long-Range Dependent Volatility Models, WP02-12
2. Paolo Zaffaroni, Aggregation and Memory of Models of Changing Volatility, WP02-11
3. Jerry Coakley, Ana-Maria Fuertes and Andrew Wood, Reinterpreting the Real Exchange Rate - Yield Diffential Nexus, WP02-10
4. Gordon Gemmill and Dylan Thomas , Noise Training, Costly Arbitrage and Asset Prices: evidence from closed-end funds, WP02-09
5. Gordon Gemmill, Testing Merton's Model for Credit Spreads on Zero-Coupon Bonds, WP02-08
6. George Christodoulakis and Steve Satchell, On th Evolution of Global Style Factors in the MSCI Universe of Assets, WP02-07
7. George Christodoulakis, Sharp Style Analysis in the MSCI Sector Portfolios: A Monte Caro Integration Approach, WP02-06

8. George Christodoulakis, Generating Composite Volatility Forecasts with Random Factor Betas, WP02-05
9. Claudia Riveiro and Nick Webber, Valuing Path Dependent Options in the Variance-Gamma Model by Monte Carlo with a Gamma Bridge, WP02-04
10. Christian Pedersen and Soosung Hwang, On Empirical Risk Measurement with Asymmetric Returns Data, WP02-03
11. Roy Batchelor and Ismail Orgakcioglu, Event-related GARCH: the impact of stock dividends in Turkey, WP02-02
12. George Albanis and Roy Batchelor, Combining Heterogeneous Classifiers for Stock Selection, WP02-01

2001

1. Soosung Hwang and Steve Satchell , GARCH Model with Cross-sectional Volatility; GARCHX Models, WP01-16
2. Soosung Hwang and Steve Satchell, Tracking Error: Ex-Ante versus Ex-Post Measures, WP01-15
3. Soosung Hwang and Steve Satchell, The Asset Allocation Decision in a Loss Aversion World, WP01-14
4. Soosung Hwang and Mark Salmon, An Analysis of Performance Measures Using Copulae, WP01-13
5. Soosung Hwang and Mark Salmon, A New Measure of Herding and Empirical Evidence, WP01-12
6. Richard Lewin and Steve Satchell, The Derivation of New Model of Equity Duration, WP01-11
7. Massimiliano Marcellino and Mark Salmon, Robust Decision Theory and the Lucas Critique, WP01-10
8. Jerry Coakley, Ana-Maria Fuertes and Maria-Teresa Perez, Numerical Issues in Threshold Autoregressive Modelling of Time Series, WP01-09
9. Jerry Coakley, Ana-Maria Fuertes and Ron Smith, Small Sample Properties of Panel Time-series Estimators with I(1) Errors, WP01-08
10. Jerry Coakley and Ana-Maria Fuertes, The Felsdtein-Horioka Puzzle is Not as Bad as You Think, WP01-07
11. Jerry Coakley and Ana-Maria Fuertes, Rethinking the Forward Premium Puzzle in a Non-linear Framework, WP01-06
12. George Christodoulakis, Co-Volatility and Correlation Clustering: A Multivariate Correlated ARCH Framework, WP01-05
13. Frank Critchley, Paul Marriott and Mark Salmon, On Preferred Point Geometry in Statistics, WP01-04
14. Eric Bouyé and Nicolas Gaussel and Mark Salmon, Investigating Dynamic Dependence Using Copulae, WP01-03
15. Eric Bouyé, Multivariate Extremes at Work for Portfolio Risk Measurement, WP01-02
16. Erick Bouyé, Vado Durrleman, Ashkan Nikeghbali, Gael Riboulet and Thierry Roncalli, Copulas: an Open Field for Risk Management, WP01-01

2000

1. Soosung Hwang and Steve Satchell , Valuing Information Using Utility Functions, WP00-06
2. Soosung Hwang, Properties of Cross-sectional Volatility, WP00-05
3. Soosung Hwang and Steve Satchell, Calculating the Miss-specification in Beta from Using a Proxy for the Market Portfolio, WP00-04
4. Laun Middleton and Stephen Satchell, Deriving the APT when the Number of Factors is Unknown, WP00-03
5. George A. Christodoulakis and Steve Satchell, Evolving Systems of Financial Returns: Auto-Regressive Conditional Beta, WP00-02
6. Christian S. Pedersen and Stephen Satchell, Evaluating the Performance of Nearest Neighbour Algorithms when Forecasting US Industry Returns, WP00-01

1999

1. Yin-Wong Cheung, Menzie Chinn and Ian Marsh, How do UK-Based Foreign Exchange Dealers Think Their Market Operates?, WP99-21
2. Soosung Hwang, John Knight and Stephen Satchell, Forecasting Volatility using LINEX Loss Functions, WP99-20
3. Soosung Hwang and Steve Satchell, Improved Testing for the Efficiency of Asset Pricing Theories in Linear Factor Models, WP99-19
4. Soosung Hwang and Stephen Satchell, The Disappearance of Style in the US Equity Market, WP99-18
5. Soosung Hwang and Stephen Satchell, Modelling Emerging Market Risk Premia Using Higher Moments, WP99-17
6. Soosung Hwang and Stephen Satchell, Market Risk and the Concept of Fundamental Volatility: Measuring Volatility Across Asset and Derivative Markets and Testing for the Impact of Derivatives Markets on Financial Markets, WP99-16
7. Soosung Hwang, The Effects of Systematic Sampling and Temporal Aggregation on Discrete Time Long Memory Processes and their Finite Sample Properties, WP99-15
8. Ronald MacDonald and Ian Marsh, Currency Spillovers and Tri-Polarity: a Simultaneous Model of the US Dollar, German Mark and Japanese Yen, WP99-14
9. Robert Hillman, Forecasting Inflation with a Non-linear Output Gap Model, WP99-13
10. Robert Hillman and Mark Salmon , From Market Micro-structure to Macro Fundamentals: is there Predictability in the Dollar-Deutsche Mark Exchange Rate?, WP99-12
11. Renzo Avesani, Giampiero Gallo and Mark Salmon, On the Evolution of Credibility and Flexible Exchange Rate Target Zones, WP99-11
12. Paul Marriott and Mark Salmon, An Introduction to Differential Geometry in Econometrics, WP99-10
13. Mark Dixon, Anthony Ledford and Paul Marriott, Finite Sample Inference for Extreme Value Distributions, WP99-09
14. Ian Marsh and David Power, A Panel-Based Investigation into the Relationship Between Stock Prices and Dividends, WP99-08
15. Ian Marsh, An Analysis of the Performance of European Foreign Exchange Forecasters, WP99-07
16. Frank Critchley, Paul Marriott and Mark Salmon, An Elementary Account of Amari's Expected Geometry, WP99-06
17. Demos Tambakis and Anne-Sophie Van Royen, Bootstrap Predictability of Daily Exchange Rates in ARMA Models, WP99-05
18. Christopher Neely and Paul Weller, Technical Analysis and Central Bank Intervention, WP99-04
19. Christopher Neely and Paul Weller, Predictability in International Asset Returns: A Re-examination, WP99-03
20. Christopher Neely and Paul Weller, Intraday Technical Trading in the Foreign Exchange Market, WP99-02
21. Anthony Hall, Soosung Hwang and Stephen Satchell, Using Bayesian Variable Selection Methods to Choose Style Factors in Global Stock Return Models, WP99-01

1998

1. Soosung Hwang and Stephen Satchell, Implied Volatility Forecasting: A Compaison of Different Procedures Including Fractionally Integrated Models with Applications to UK Equity Options, WP98-05
2. Roy Batchelor and David Peel, Rationality Testing under Asymmetric Loss, WP98-04
3. Roy Batchelor, Forecasting T-Bill Yields: Accuracy versus Profitability, WP98-03
4. Adam Kurpiel and Thierry Roncalli , Option Hedging with Stochastic Volatility, WP98-02
5. Adam Kurpiel and Thierry Roncalli, Hopscotch Methods for Two State Financial Models, WP98-01