Chapter 2

Weak dependence

Many authors have used one of the two following type of dependence: on the one hand mixing properties, introduced by Rosenblatt (1956) [166], on the other hand martingales approximations or mixingales, following the works of Gordin (1969, 1973) [97], [98] and Mc Leisch (1974, 1975) [127], [129]. Concerning strongly mixing sequences, very deep and elegant results have been established: for recent works, we mention the books of Rio (2000) [161] and Bradley (2002) [30]. However many classes of time series do not satisfy any mixing condition as it is quoted *e.g.* in Eberlein and Taqqu (1986) [83] or Doukhan (1994) [61]. Conversely, most of such time series enter the scope of mixingales but limit theorems and moment inequalities are more difficult to obtain in this general setting.

Between those directions, Bickel and Bühlmann (1999) [18] and simultaneously Doukhan and Louhichi (1999) [67] introduced a new idea of weak dependence. Their notion of weak dependence makes explicit the asymptotic independence between 'past' and 'future'; this means that the 'past' is progressively forgotten. In terms of the initial time series, 'past' and 'future' are elementary events given through finite dimensional marginals. Roughly speaking, for convenient functions f and g, we shall assume that

$$\mathrm{Cov}\,(f(\text{`past'}), g(\text{`future'}))$$

is small when the distance between the 'past' and the 'future' is sufficiently large. Such inequalities are significant only if the distance between indices of the initial time series in the 'past' and 'future' terms grows to infinity. The convergence is not assumed to hold uniformly on the dimension of the 'past' or 'future' involved.

The main advantage is that such a kind of dependence contains lots of pertinent examples and can be used in various situations: empirical central limit theorems are proved in Doukhan and Louhichi (1999) [67] and Borovkova, Burton and

Dehling (2001) [25], while applications to Bootstrap are given by Bickel and Bühlmann (1999) [18] and Ango Nzé *et al.*(2002) [6] and to functional estimation (Coulon-Prieur & Doukhan, 2000 [40]).
In this chapter a first section introduces the function spaces necessary to define the various dependence coefficients of the second section. They are classified in separated subsections. We shall first consider noncausal coefficients and then their causal counterparts; in both cases the subjacent spaces are Lipschitz spaces. A further case associated to bounded variation spaces is provided in the following subsection. Projective measure of dependence are included in the last subsection.

2.1 Function spaces

In this section, we give the definitions of some function spaces used in this book.

- Let m be any measure on a measurable space $(\Omega, \mathcal{A})$. For any $p \geq 1$, we denote by $\mathbb{L}^p(m)$ the space of measurable functions f from Ω to $\mathbb{R}$ such that

$$\begin{aligned} \|f\|_{p,m} &= \left(\int |f(x)|^p m(dx)\right)^{1/p} < \infty, \\ \|f\|_{\infty,m} &= \inf\left\{M > 0 \,\big/\, m(|f| > M) = 0\right\} < \infty, \quad \text{for } p = \infty. \end{aligned}$$

  For simplicity, when no confusion can arise, we shall write $\mathbb{L}^p$ and $\|\cdot\|_p$ instead of $\mathbb{L}^p(m)$ and $\|\cdot\|_{p,m}$.

Let $\mathcal{X}$ be a Polish space and δ be some metric on $\mathcal{X}$ ($\mathcal{X}$ need not be Polish with respect to δ).

- Let $\Lambda(\delta)$ be the set of Lipschitz functions from $\mathcal{X}$ to $\mathbb{R}$ with respect to the distance δ. For $f \in \Lambda(\delta)$, denote by $\mathrm{Lip}\,(f)$, f's Lipschitz constant. Let

$$\Lambda^{(1)}(\delta) = \{f \in \Lambda(\delta) \ / \ \mathrm{Lip}\,(f) \leq 1\}.$$

- Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let $\mathcal{X}$ be a Polish space and δ be a distance on $\mathcal{X}$. For any $p \in [1, \infty]$, we say that a random variable X with values in $\mathcal{X}$ is $\mathbb{L}^p$-integrable if, for some x_0 in $\mathcal{X}$, the real valued random variable $\delta(X, x_0)$ belongs to $\mathbb{L}^p(\mathbb{P})$.

Another type of function class will be used in this chapter: it is the class of functions with bounded variation on the real line. To be complete, we recall,

Definition 2.1. *A σ-finite signed measure is the difference of two positive σ-finite measures, one of them at least being finite. We say that a function h from $\mathbb{R}$ to $\mathbb{R}$ is σ-BV if there exists a σ-finite signed measure dh such that $h(x) = h(0) + dh([0, x[)$ if $x \geq 0$ and $h(x) = h(0) - dh([x, 0[)$ if $x \leq 0$ (h is left continuous). The function h is BV if the signed measure dh is finite.*

Recall also the Hahn-Jordan decomposition: for any σ-finite signed measure μ, there is a set D such that

$$\mu_+(A) = \mu(A \cap D) \geq 0, \qquad -\mu_-(A) = \mu(A \backslash D) \leq 0.$$

μ_+ and μ_- are mutually singular, one of them at least is finite and $\mu = \mu_+ - \mu_-$. The measure $|\mu| = \mu_+ + \mu_-$ is called the total variation measure for μ. The total variation of μ writes as $\|\mu\| = |\mu|(\mathbb{R})$.

Now we are in position to introduce

- BV_1 the space of BV functions $h : \mathbb{R} \to \mathbb{R}$ such that $\|dh\| \leq 1$.

2.2 Weak dependence

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $\mathcal{X}$ be a Polish space. Let

$$\mathcal{F} = \bigcup_{u \in \mathbb{N}^*} \mathcal{F}_u \quad \text{and} \quad \mathcal{G} = \bigcup_{u \in \mathbb{N}^*} \mathcal{G}_u \,,$$

where $\mathcal{F}_u$ and $\mathcal{G}_u$ are two classes of functions from $\mathcal{X}^u$ to $\mathbb{R}$.

Definition 2.2. *Let X and Y be two random variables with values in $\mathcal{X}^u$ and $\mathcal{X}^v$ respectively. If Ψ is some function from $\mathcal{F} \times \mathcal{G}$ to $\mathbb{R}_+$, define the $\big(\mathcal{F}, \mathcal{G}, \Psi\big)$-dependence coefficient $\epsilon(X, Y)$ by*

$$\epsilon(X, Y) = \sup_{f \in \mathcal{F}_u \, g \in \mathcal{G}_v} \frac{|\mathrm{Cov}(f(X), g(Y))|}{\Psi(f, g)} \,. \tag{2.2.1}$$

Let $(X_n)_{n \in \mathbb{Z}}$ be a sequence of $\mathcal{X}$-valued random variables. Let $\Gamma(u, v, k)$ be the set of (i, j) in $\mathbb{Z}^u \times \mathbb{Z}^v$ such that $i_1 < \cdots < i_u \leq i_u + k \leq j_1 < \cdots < j_v$. The dependence coefficient $\epsilon(k)$ is defined by

$$\epsilon(k) = \sup_{u, v} \sup_{(i,j) \in \Gamma(u,v,k)} \epsilon((X_{i_1}, \ldots, X_{i_u}), (X_{j_1}, \ldots, X_{j_v})) \,.$$

The sequence $(X_n)_{n \in \mathbb{Z}}$ is $(\mathcal{F}, \mathcal{G}, \Psi)$-dependent if the sequence $(\epsilon(k))_{k \in \mathbb{N}}$ tends to zero. If $\mathcal{F} = \mathcal{G}$ we simply denote this as $(\mathcal{F}, \Psi)$-dependence.

Remark 2.1. *Definition 2.2 above easily extends to general metric sets of indices T equipped with a distance δ (e.g. $T = \mathbb{Z}^d$ yields the case of random fields). The set $\Gamma(u, v, k)$ is then the set of (i, j) in $T^u \times T^v$ such that*

$$k = \min \{\delta(i_\ell, j_m) \,/\, 1 \leq \ell \leq u, 1 \leq m \leq v\} \,.$$

2.2.1 η, κ, λ and ζ-coefficients

In this section, we focus on the case where $\mathcal{F}_u = \mathcal{G}_u$. If f belongs to $\mathcal{F}_u$, we define $d_f = u$.
In a first time, $\mathcal{F}_u$ is the set of bounded functions from $\mathcal{X}^u$ to $\mathbb{R}$, which are Lipschitz with respect to the distance δ_1 on $\mathcal{X}^u$ defined by

$$\delta_1(x, y) = \sum_{i=1}^{u} \delta(x_i, y_i) . \tag{2.2.2}$$

In that case:

- the coefficient η corresponds to

$$\Psi(f, g) = d_f \|g\|_\infty \mathrm{Lip}\,(f) + d_g \|f\|_\infty \mathrm{Lip}\,(g) \,, \tag{2.2.3}$$

- the coefficient λ corresponds to

$$\Psi(f, g) = d_f \|g\|_\infty \mathrm{Lip}\,(f) + d_g \|f\|_\infty \mathrm{Lip}\,(g) + d_f d_g \mathrm{Lip}\,(f) \mathrm{Lip}\,(g) \,. \tag{2.2.4}$$

To define the coefficients κ and ζ, we consider for $\mathcal{F}_u$ the wider set of functions from $\mathcal{X}^u$ to $\mathbb{R}$, which are Lipschitz with respect to the distance δ_1 on $\mathcal{X}^u$, but which are not necessarily bounded. In that case we assume that the variables X_i are $\mathbb{L}^1$-integrable.

- the coefficient κ corresponds to

$$\Psi(f, g) = d_f d_g \mathrm{Lip}\,(f) \mathrm{Lip}\,(g) \,, \tag{2.2.5}$$

- the coefficient ζ corresponds to

$$\Psi(f, g) = \min(d_f, d_g) \mathrm{Lip}\,(f) \mathrm{Lip}\,(g) \,. \tag{2.2.6}$$

These coefficients have some hereditary properties. For example, let $h : \mathcal{X} \to \mathbb{R}$ be a Lipschitz function with respect to δ, then if the sequence $(X_n)_{n\in\mathbb{Z}}$ is η, κ, λ or ζ weakly dependent, then the same is true for the sequence $(h(X_n))_{n\in\mathbb{Z}}$.
One can also obtain some hereditary properties for functions which are not Lipschitz on the whole space $\mathcal{X}$, as shown by Lemma 2.1 below, in the special case where $\mathcal{X} = \mathbb{R}^k$ equipped with the distance $\delta(x, y) = \max_{1\le i\le k} |x_i - y_i|$.

Proposition 2.1 (Bardet, Doukhan, León, 2006 [11]). *Let $(X_n)_{n\in\mathbb{Z}}$ be a sequence of $\mathbb{R}^k$-valued random variables. Let $p > 1$. We assume that there exists some constant $C > 0$ such that $\max_{1\le i\le k} \|X_i\|_p \le C$. Let h be a function from $\mathbb{R}^k$ to $\mathbb{R}$ such that $h(0) = 0$ and for $x, y \in \mathbb{R}^k$, there exist a in $[1, p[$ and $c > 0$ such that*

$$|h(x) - h(y)| \le c|x - y|(|x|^{a-1} + |y|^{a-1}) \,.$$

We define the sequence $(Y_n)_{n\in\mathbb{Z}}$ by $Y_n = h(X_n)$. Then,

- *if* $(X_n)_{n\in\mathbb{Z}}$ *is* η*-weak dependent, then* $(Y_n)_{n\in\mathbb{Z}}$ *also, and*

$$\eta_Y(n) = \mathcal{O}\left(\eta(n)^{\frac{p-a}{p-1}}\right);$$

- *if* $(X_n)_{n\in\mathbb{Z}}$ *is* λ*-weak dependent, then* $(Y_n)_{n\in\mathbb{Z}}$ *also, and*

$$\lambda_Y(n) = \mathcal{O}\left(\lambda(n)^{\frac{p-a}{p+a-2}}\right).$$

Remark 2.2. *The function* $h(x) = x^2$ *satisfies the previous assumptions with* $a = 2$*. This condition is satisfied by polynomials with degree* a*.*

Proof of Proposition 2.1. Let f and g be two real functions in $\mathcal{F}_u$ and $\mathcal{F}_v$ respectively. Denote $x^{(M)} = (x \wedge M) \vee (-M)$ for $x \in \mathbb{R}$. Now, for $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$, we analogously denote $x^{(M)} = (x_1^{(M)}, \ldots, x_k^{(M)})$. Assume that (i, j) belong to the set $\Gamma(u, v, r)$ defined in Definition 2.2. Define $X_{\mathbf{i}} = (X_{i_1}, \ldots, X_{i_u})$ and $X_{\mathbf{j}} = (X_{j_1}, \ldots, X_{j_v})$. We then define functions $F : \mathbb{R}^{uk} \to \mathbb{R}$ and $G : \mathbb{R}^{vk} \to \mathbb{R}$ through the relations:

- $F(X_{\mathbf{i}}) = f(h(X_{i_1}), \ldots, h(X_{i_u}))$, $F^{(M)}(X_{\mathbf{i}}) = f(h(X_{i_1}^{(M)}), \ldots, h(X_{i_u}^{(M)}))$,
- $G(X_{\mathbf{j}}) = g(h(X_{j_1}), \ldots, h(X_{j_v}))$, $G^{(M)}(X_{\mathbf{j}}) = g(h(X_{j_1}^{(M)}), \ldots, h(X_{j_v}^{(M)}))$.

Then:

$$\begin{aligned}
|\mathrm{Cov}(F(X_{\mathbf{i}}), G(X_{\mathbf{j}}))| &\le |\mathrm{Cov}(F(X_{\mathbf{i}}), G(X_{\mathbf{j}}) - G^{(M)}(X_{\mathbf{j}}))| \\
&\quad + |\mathrm{Cov}(F(X_{\mathbf{i}}), G^{(M)}(X_{\mathbf{j}}))| \\
&\le 2\|f\|_\infty \, \mathbb{E}|G(X_{\mathbf{j}}) - G^{(M)}(X_{\mathbf{j}}))| \\
&\quad + 2\|g\|_\infty \, \mathbb{E}|F(X_{\mathbf{i}}) - F^{(M)}(X_{\mathbf{i}})| \\
&\quad + |\mathrm{Cov}(F^{(M)}(X_{\mathbf{i}}), G^{(M)}(X_{\mathbf{j}}))|
\end{aligned}$$

But we also have from the assumptions on h and Markov inequality,

$$\begin{aligned}
\mathbb{E}|G(X_{\mathbf{j}}) - G^{(M)}(X_{\mathbf{j}}))| &\le \mathrm{Lip}\, g \sum_{l=1}^{v} \mathbb{E}|h(X_{j_l}) - h(X_{j_l}^{(M)})| \\
&\le 2c \, \mathrm{Lip}\, g \sum_{l=1}^{v} \mathbb{E}\big(|X_{j_l}|^a \mathbf{1}_{|X_{j_l}|>M}\big), \\
&\le 2c\, v \, \mathrm{Lip}\, g C^p M^{a-p}.
\end{aligned}$$

The same thing holds for F. Moreover, the functions $F^{(M)} : \mathbb{R}^{uk} \to \mathbb{R}$ and $G^{(M)} : \mathbb{R}^{vk} \to \mathbb{R}$ satisfy $\mathrm{Lip}\, F^{(M)} \le 2cM^{a-1}\mathrm{Lip}\,(f)$ and $\mathrm{Lip}\, G^{(M)} \le 2cM^{a-1}$

$\mathrm{Lip}\,(g)$, and $\|F^{(M)}\|_\infty \le \|f\|_\infty$, $\|G^{(M)}\|_\infty \le \|g\|_\infty$. Thus, from the definition of weak dependence of X and the choice of $\mathbf{i}, \mathbf{j}$, we obtain respectively, if $M \ge 1$

$$\begin{aligned}\left|\mathrm{Cov}\left(F^{(M)}(X_{\mathbf{i}}), G^{(M)}(X_{\mathbf{j}})\right)\right| &\le 2c(u\mathrm{Lip}\,(f)\|g\|_\infty + v\mathrm{Lip}\,(g)\|f\|_\infty)M^{a-1}\eta(r),\\ &\le 2c(d_f\mathrm{Lip}\,(f)\|g\|_\infty + d_g\mathrm{Lip}\,(g)\|f\|_\infty)M^{a-1}\lambda(r)\\ &+ 4c^2 d_f d_g \mathrm{Lip}\,(f)\mathrm{Lip}\,(g)M^{2a-2}\lambda(r).\end{aligned}$$

Finally, we obtain respectively, if $M \ge 1$:

$$\begin{aligned}|\mathrm{Cov}(F(X_{\mathbf{i}}), G(X_{\mathbf{j}}))| &\le 2c(u\mathrm{Lip}\,f\|g\|_\infty + v\mathrm{Lip}\,g\|f\|_\infty)\\ &\qquad \times\left(M^{a-1}\eta(r) + 2C^pM^{a-p}\right),\\ &\le c(u\mathrm{Lip}\,f + v\mathrm{Lip}\,g + uv\mathrm{Lip}\,f\mathrm{Lip}\,g)\\ &\qquad \times(M^{2a-2}\lambda(r) + M^{a-p}).\end{aligned}$$

Choosing $M = \eta(r)^{1/(1-p)}$ and $M = \lambda(r)^{-1/(p+a-2)}$ respectively, we obtain the result. □

In the definition of the coefficients η, κ, λ and ζ, we assume some regularity conditions on $\mathcal{F}_u = \mathcal{G}_u$. In the case where the sequence $(X_n)_{n\in\mathbb{Z}}$ is an adapted process with respect to some increasing filtration $(\mathcal{M}_i)_{i\in\mathbb{Z}}$, it is often more suitable to work without assuming any regularity conditions on $\mathcal{F}_u$. In that case $\mathcal{G}_u$ is some space of regular functions and $\mathcal{F}_u \ne \mathcal{G}_u$. This last case is called the causal case. In the situations where both $\mathcal{F}_u$ and $\mathcal{G}_u$ are spaces of regular functions, we say that we are in the non causal case.

2.2.2 θ and τ-coefficients

Let $\mathcal{F}_u$ be the class of bounded functions from $\mathcal{X}_u$ to $\mathbb{R}$, and let $\mathcal{G}_u$ be the class of functions from $\mathcal{X}_u$ to $\mathbb{R}$ which are Lipschitz with respect to the distance δ_1 defined by (2.2.2). We assume that the variables X_i are $\mathbb{L}^1$-integrable.

- The coefficient θ corresponds to

$$\Psi(f,g) = d_g\|f\|_\infty \mathrm{Lip}\,(g)\ . \tag{2.2.7}$$

The coefficient θ has some hereditary properties. For example, Proposition 2.2 below gives hereditary properties similar to those given for the coefficients η and λ in Lemma 2.1.

Proposition 2.2. *Let $(X_n)_{n\in\mathbb{Z}}$ be a sequence of $\mathbb{R}^k$-valued random variables. We define the sequence $(Y_n)_{n\in\mathbb{Z}}$ by $Y_n = h(X_n)$. The assumptions on $(X_n)_{n\in\mathbb{Z}}$ and on h are the same as in Lemma 2.1. Then,*

- *if $(X_n)_{n\in\mathbb{Z}}$ is θ-weak dependent, $(Y_n)_{n\in\mathbb{Z}}$ also, and*

$$\theta_Y(n) = \mathcal{O}\left(\theta(n)^{\frac{p-a}{p-1}}\right).$$

The proof of Proposition 2.2 follows the same line as the proof of Proposition 2.1 and therefore is not detailed.
We shall see that the coefficient θ defined above belongs to a more general class of dependence coefficients defined through conditional expectations with respect to the filtration $\sigma(X_j, j \le i)$.

Definition 2.3. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and $\mathcal{M}$ be a σ-algebra of $\mathcal{A}$. Let $\mathcal{X}$ be a Polish space and δ a distance on $\mathcal{X}$. For any $\mathbb{L}^p$-integrable random variable X (see § 2.1) with values in $\mathcal{X}$, we define*

$$\theta_p(\mathcal{M}, X) = \sup\{\|\mathbb{E}(g(X)|\mathcal{M}) - \mathbb{E}(g(X))\|_p \,/\, g \in \Lambda^{(1)}(\delta)\}. \tag{2.2.8}$$

Let $(X_i)_{i\in\mathbb{Z}}$ be a sequence of $\mathbb{L}^p$-integrable $\mathcal{X}$-valued random variables, and let $(\mathcal{M}_i)_{i\in\mathbb{Z}}$ be a sequence of σ-algebras of $\mathcal{A}$. On $\mathcal{X}^l$, we consider the distance δ_1 defined by (2.2.2). The sequence of coefficients $\theta_{p,r}(k)$ is then defined by

$$\theta_{p,r}(k) = \max_{\ell \le r} \frac{1}{\ell} \sup_{(i,j)\in\Gamma(1,\ell,k)} \theta_p\left(\mathcal{M}_i, (X_{j_1}, \dots, X_{j_\ell})\right). \tag{2.2.9}$$

When it is not clearly specified, we shall always take $\mathcal{M}_i = \sigma(X_k, k \le i)$.

The two preceding definitions are coherent as proved below.

Proposition 2.3. *Let $(X_i)_{i\in\mathbb{Z}}$ be a sequence of $\mathbb{L}^1$-integrable $\mathcal{X}$-valued random variables, and let $\mathcal{M}_i = \sigma(X_j, j \le i)$. According to the definition of $\theta(k)$ and to the definition 2.3, we have the equality*

$$\theta(k) = \theta_{1,\infty}(k). \tag{2.2.10}$$

Proof of Proposition 2.3. The fact that $\theta(k) \le \theta_{1,\infty}(k)$ is clear since, for any f in $\mathcal{F}_u$, g in $\mathcal{G}_v$, and any $(i,j) \in \Gamma(u,v,k)$,

$$\begin{aligned}
&\left|\mathrm{Cov}\Big(\frac{f(X_{i_1},\dots,X_{i_u})}{\|f\|_\infty}, \frac{g(X_{j_1},\dots,X_{j_v})}{v\mathrm{Lip}\,(g)}\Big)\right| \\
&\qquad \le \frac{1}{v}\left\|\mathbb{E}\Big(\frac{g(X_{j_1},\dots,X_{j_v})}{\mathrm{Lip}\,(g)}\Big|\mathcal{M}_{i_u}\Big) - \mathbb{E}\Big(\frac{g(X_{j_1},\dots,X_{j_v})}{\mathrm{Lip}\,(g)}\Big)\right\|_1 \le \theta_{1,\infty}(k).
\end{aligned}$$

To prove the converse inequality, we first notice that

$$\theta(\mathcal{M}_i, (X_{j_1}, \dots, X_{j_v}) = \lim_{k\to-\infty} \theta\left(\mathcal{M}_{k,i}, (X_{j_1}, \dots, X_{j_v})\right), \tag{2.2.11}$$

where $\mathcal{M}_{k,i} = \sigma(X_j, k \le j \le i)$. Now, letting

$$f(X_k, \ldots, X_i) = \text{sign}\big\{\mathbb{E}(g(X_{j_1}, \ldots, X_{j_v})|\mathcal{M}_{k,i}) - \mathbb{E}(g(X_{j_1}, \ldots, X_{j_v}))\big\},$$

we have that, for (i, j) in $\Gamma(1, v, k)$ and g in $\Lambda^{(1)}(\delta_1)$,

$$\begin{aligned}\|\mathbb{E}(g(X_{j_1}, \ldots, X_{j_v})|\mathcal{M}_{k,i}) &- \mathbb{E}(g(X_{j_1}, \ldots, X_{j_v}))\|_1 \\ &= \text{Cov}(f(X_k, \ldots, X_i), g(X_{j_1}, \ldots, X_{j_v})) \le v\theta(k)\ .\end{aligned}$$

We infer that

$$\frac{1}{v}\theta(\mathcal{M}_{k,i}, (X_{j_1}, \ldots, X_{j_v}) \le \theta(k)$$

and we conclude from (2.2.11) that $\theta_{1,\infty}(k) \le \theta(k)$. The proof is complete. □
Having in view the coupling arguments in § 5.3, we now define a variation of the coefficient (2.2.8) where we exchange the order of $\|.\|_p$ and the supremum. This is the same step as passing from $\alpha-$mixing to $\beta-$mixing, which is known to ensure nice coupling arguments (see Berbee, 1979 [16]).

Definition 2.4. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and $\mathcal{M}$ a σ-algebra of $\mathcal{A}$. Let $\mathcal{X}$ be a Polish space and δ a distance on $\mathcal{X}$. For any $\mathbb{L}^p-$integrable (see § 2.1)) $\mathcal{X}$-valued random variable X, we define the coefficient τ_p by:*

$$\tau_p(\mathcal{M}, X) = \left\| \sup_{g \in \Lambda^{(1)}(\delta)} \left\{ \int g(x)\mathbb{P}_{X|\mathcal{M}}(dx) - \int g(x)\mathbb{P}_X(dx) \right\} \right\|_p \tag{2.2.12}$$

where $\mathbb{P}_X$ is the distribution of X and $\mathbb{P}_{X|\mathcal{M}}$ is a conditional distribution of X given $\mathcal{M}$. We clearly have

$$\theta_p(\mathcal{M}, X) \le \tau_p(\mathcal{M}, X)\,. \tag{2.2.13}$$

Let $(X_i)_{i\in\mathbb{Z}}$ be a sequence of $\mathbb{L}^p$-integrable $\mathcal{X}$-valued random variables. The coefficients $\tau_{p,r}(k)$ are defined from τ_p as in (2.2.9).

2.2.3 $\tilde{\alpha}$, $\tilde{\beta}$ and $\tilde{\phi}$-coefficients.

In the case where $\mathcal{X} = (\mathbb{R}^d)^r$, we introduce some new coefficients based on indicator of quadrants. Recall that if x and y are two elements of $\mathbb{R}^d$, then $x \le y$ if and only if $x_i \le y_i$ for any $1 \le i \le d$.

Definition 2.5. *Let $X = (X_1, \ldots, X_r)$ be a $(\mathbb{R}^d)^r$-valued random variable and $\mathcal{M}$ a σ-algebra of $\mathcal{A}$. For t_i in $\mathbb{R}^d$ and x in $\mathbb{R}^d$, let $g_{t_i,i}(x) = \mathbf{1}_{x\le t_i} - \mathbb{P}(X_i \le t_i)$. Keeping the same notations as in Definition 2.4, define for $t = (t_1, \ldots, t_r)$ in $(\mathbb{R}^d)^r$,*

$$L_{X|\mathcal{M}}(t) = \int \prod_{i=1}^r g_{t_i,i}(x_i)\mathbb{P}_{X|\mathcal{M}}(dx) \quad \textit{and} \quad L_X(t) = \mathbb{E}\prod_{i=1}^r g_{t_i,i}(X_i).$$

Define now the coefficients

1. $\tilde{\alpha}(\mathcal{M}, X) = \sup_{t \in (\mathbb{R}^d)^r} \|L_{X|\mathcal{M}}(t) - L_X(t)\|_1$.

2. $\tilde{\beta}(\mathcal{M}, X) = \Big\| \sup_{t \in (\mathbb{R}^d)^r} |L_{X|\mathcal{M}}(t) - L_X(t)| \Big\|_1$.

3. $\tilde{\phi}(\mathcal{M}, X) = \sup_{t \in (\mathbb{R}^d)^r} \|L_{X|\mathcal{M}}(t) - L_X(t)\|_\infty$.

Remark 2.3. *Note that if $r = 1, d = 1$ and $\delta(x, y) = |x - y|$, then, with the above notation,*

$$\tau_1(\mathcal{M}, X) = \int \|L_{X|\mathcal{M}}(t)\|_1 dt\,.$$

The proof of this equality follows the same lines than the proof of the coupling property of τ_1 (see Chapter 5, proof of Lemma 5.2).

In the definition of the coefficients θ and τ, we have used the class of functions $\Lambda^{(1)}(\delta)$. In the case where $d = 1$, we can define the coefficients $\tilde{\alpha}(\mathcal{M}, X)$, $\tilde{\beta}(\mathcal{M}, X)$ and $\tilde{\phi}(\mathcal{M}, X)$ with the help of bounded variation functions. This is the purpose of the following lemma:

Lemma 2.1. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $X = (X_1, \ldots, X_r)$ a $\mathbb{R}^r$-valued random variable and $\mathcal{M}$ a σ-algebra of $\mathcal{A}$. If f is a function in BV_1, let $f^{(i)}(x) = f(x) - \mathbb{E}(f(X_i))$. The following relations hold:*

1. $\tilde{\alpha}(\mathcal{M}, X) = \sup_{f_1, \ldots, f_r \in BV_1} \left\| \mathbb{E}\left(\prod_{i=1}^r f_i^{(i)}(X_i) \Big| \mathcal{M} \right) - \mathbb{E}\left(\prod_{i=1}^r f_i^{(i)}(X_i) \right) \right\|_1$.

2. $\tilde{\beta}(\mathcal{M}, X) = \left\| \sup_{f_1, \ldots, f_r \in BV_1} \left| \int \prod_{i=1}^r f_i^{(i)}(x_i) \left(\mathbb{P}_{X|\mathcal{M}} - \mathbb{P}_X \right)(dx) \right| \right\|_1$.

3. $\tilde{\phi}(\mathcal{M}, X) = \sup_{f_1, \ldots, f_r \in BV_1} \left\| \mathbb{E}\left(\prod_{i=1}^r f_i^{(i)}(X_i) | \mathcal{M} \right) - \mathbb{E}\left(\prod_{i=1}^r f_i^{(i)}(X_i) \right) \right\|_\infty$.

Remark 2.4. *For $r = 1$ and $d = 1$, the coefficient $\tilde{\alpha}(\mathcal{M}, X)$ was introduced by Rio (2000, equation 1.10c [161]) and used by Peligrad (2002) [140], while $\tau_1(\mathcal{M}, X)$ was introduced by Dedecker and Prieur (2004a) [45]. Let $\alpha(\mathcal{M}, \sigma(X))$, $\beta(\mathcal{M}, \sigma(X))$ and $\phi(\mathcal{M}, \sigma(X))$ be the usual mixing coefficients defined respectively by Rosenblatt (1956) [166], Rozanov and Volkonskii (1959) [187] and Ibragimov (1962) [110]. Starting from Definition 2.5 one can easily prove that*

$$\tilde{\alpha}(\mathcal{M}, X) \le 2\alpha(\mathcal{M}, \sigma(X)),\ \tilde{\beta}(\mathcal{M}, X) \le \beta(\mathcal{M}, \sigma(X)),\ \tilde{\phi}(\mathcal{M}, X) \le \phi(\mathcal{M}, \sigma(X)).$$

Proof of Lemma 2.1. Let f_i be a function in BV_1. Assume without loss of generality that $f_i(-\infty) = 0$. Then

$$f_i^{(i)}(x) = -\int (\mathbf{1}_{x \le t} - \mathbb{P}(X_i \le t))\, df_i(t)\,.$$

Hence,

$$\int \prod_{i=1}^{k} f_i^{(i)}(x_i) \mathbb{P}_{X|\mathcal{M}}(dx) = (-1)^k \int \Big(\int \prod_{i=1}^{k} g_{t_i,i}(x_i) \mathbb{P}_{X|\mathcal{M}}(dx) \Big) \prod_{i=1}^{k} df_i(t_i)\,,$$

and the same is true for $\mathbb{P}_X$ instead of $\mathbb{P}_{X|\mathcal{M}}$. From these inequalities and the fact that $|df_i|(\mathbb{R}) \le 1$, we infer that

$$\sup_{f_1,\dots,f_k \in BV_1} \Big| \int \prod_{i=1}^{k} f_i^{(i)}(x_i) \mathbb{P}_{X|\mathcal{M}}(dx) - \int \prod_{i=1}^{k} f_i^{(i)}(x_i) \mathbb{P}_X(dx) \Big| \le \sup_{t \in \mathbb{R}^r} \left| L_{X|\mathcal{M}}(t) - L_X(t) \right|\,.$$

The converse inequality follows by noting that $x \mapsto \mathbf{1}_{x \le t}$ belongs to BV_1 . □
The following proposition gives the hereditary properties of these coefficients.

Proposition 2.4. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, X an $\mathbb{R}^r$-valued, random variable and $\mathcal{M}$ a σ-algebra of $\mathcal{A}$. Let $g_1, \dots, g_r$ be any nondecreasing functions, and let $g(X) = (g_1(X_1), \dots, g_r(X_r))$. We have the inequalities $\tilde\alpha(\mathcal{M}, g(X)) \le \tilde\alpha(\mathcal{M}, X)$, $\tilde\beta(\mathcal{M}, g(X)) \le \tilde\beta(\mathcal{M}, X)$ and $\tilde\phi(\mathcal{M}, g(X)) \le \tilde\phi(\mathcal{M}, X)$. In particular, if F_i is the distribution function of X_i, we have $\tilde\alpha(\mathcal{M}, F(X)) = \tilde\alpha(\mathcal{M}, X)$, $\tilde\beta(\mathcal{M}, F(X)) = \tilde\beta(\mathcal{M}, X)$ and $\tilde\phi(\mathcal{M}, F(X)) = \tilde\phi(\mathcal{M}, X)$.*

Notations 2.1. For any distribution function F, we define the generalized inverse as

$$F^{-1}(x) = \inf \big\{ t \in \mathbb{R} \big/ F(t) \ge x \big\}. \tag{2.2.14}$$

For any non-increasing càdlàg function $f : \mathbb{R} \to \mathbb{R}$ we analogously define the generalized inverse

$$f^{-1}(u) = \inf\{t / f(t) \le u\}.$$

Proof of Proposition 2.4. The fact that $\tilde\alpha(\mathcal{M}, g(X)) \le \tilde\alpha(\mathcal{M}, X)$ is immediate, from its definition. We infer that $\tilde\alpha(\mathcal{M}, F(X)) \le \tilde\alpha(\mathcal{M}, X)$. Applying the first result once more, we obtain that $\tilde\alpha(\mathcal{M}, F^{-1}(F(X))) \le \tilde\alpha(\mathcal{M}, F(X))$. To conclude, it suffices to note that $F^{-1} \circ F(X) = X$ almost surely, so that $\tilde\alpha(\mathcal{M}, X) \le \tilde\alpha(\mathcal{M}, F(X))$. Of course, the same arguments apply to $\tilde\beta(\mathcal{M}, X)$ and $\tilde\phi(\mathcal{M}, X)$. □
We now define the coefficients $\tilde\alpha_r(k)$, $\tilde\beta_r(k)$ and $\tilde\phi_r(k)$ for a sequence of σ-algebras and a sequence of $\mathbb{R}^d$-valued random variables.

Definition 2.6. *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space. Let* $(X_i)_{i\in\mathbb{Z}}$ *be a sequence of* $\mathbb{R}^d$*-valued random variables, and let* $(\mathcal{M}_i)_{i\in\mathbb{Z}}$ *be a sequence of* σ*-algebras of* $\mathcal{A}$*. For* $r \in \mathbb{N}^*$ *and* $k \geq 0$*, define*

$$\tilde{\alpha}_r(k) = \max_{1\leq l\leq r} \sup_{(i,j)\in\Gamma(1,l,k)} \tilde{\alpha}(\mathcal{M}_i, (X_{j_1}, \ldots, X_{j_l}))\,. \tag{2.2.15}$$

The coefficients $\tilde{\beta}_r(k)$ *and* $\tilde{\phi}_r(k)$ *are defined in the same way. When it is not clearly specified, we shall always take* $\mathcal{M}_i = \sigma(X_k, k \leq i)$*.*

2.2.4 Projective measure of dependence

Sometimes, it is not necessary to introduce a supremum over a class of functions. We can work with the simple following projective measure of dependence

Definition 2.7. *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space, and* $\mathcal{M}$ *a* σ*-algebra of* $\mathcal{A}$*. Let* $p \in [1, \infty]$*. For any* $\mathbb{L}^p-$*integrable real valued random variable define*

$$\gamma_p(\mathcal{M}, X) = \|\mathbb{E}(X|\mathcal{M}) - \mathbb{E}(X)\|_p. \tag{2.2.16}$$

Let $(X_i)_{i\in\mathbb{Z}}$ *be a sequence of* $\mathbb{L}^p-$*integrable real valued random variables, and let* $(\mathcal{M}_i)_{i\in\mathbb{Z}}$ *be a sequence of* σ*-algebras of* $\mathcal{A}$*. The sequence of coefficients* $\gamma_p(k)$ *is then defined by*

$$\gamma_p(k) = \sup_{i\in\mathbb{Z}} \gamma_p(\mathcal{M}_i, X_{i+k})\,. \tag{2.2.17}$$

When it is not clearly specified, we shall always take $\mathcal{M}_i = \sigma(X_k, k \leq i)$*.*

Remark 2.5. *Those coefficients are defined in Gordin (1969) [97], if* $p \geq 2$ *and in Gordin (1973) [98] if* $p = 1$*. Mc Leish (1975a) [128] and (1975b) [129] uses these coefficients in order to derive various limit theorems. Let us notice that*

$$\gamma_p(\mathcal{M}, X) \leq \theta_p(\mathcal{M}, X)\,. \tag{2.2.18}$$