Testing Kendall's τ for a large class of dependent sequences

Sinda Ammous

Testing Kendall's τ for a large class of dependent sequences

Sinda Ammous*

October 17, 2019

Abstract

Let $(X_i, Y_i)_{i\in\mathbb{Z}}$ be a stationary sequence of $\mathbb{R}^2$-valued random variables. To test if X_1 and Y_1 are correlated in the sense of Kendall, we propose a robust correction of the usual Kendall test, valid for a large class of dependent sequences. We also show that the condition on the dependency coefficients is optimal in a certain sense, and we illustrate our results trough different sets of simulation.

Key words: U-statistics, Kendall's tau, hypothesis testing, bivariate dependent sequences, β-mixing sequences, Markov chains.

MSC 2010: 62G10, 62G20, 62M10

Contents

1 Introduction and definitions **2**

2 Main results **4**

3 Optimality of the dependency conditions **5**

4 Simulations **7**
4.1 First example 8
4.2 Second example 9
4.3 Third example 10

5 Proofs **11**
5.1 Proof of Theorem 1 11
5.2 Proof of Proposition 3 12
5.3 Proof of Proposition 1 18

6 Appendix **22**

*MAP5 UMR CNRS 8145, Université Paris Descartes, 45 rue des Saints-Pères, 75270 Paris Cedex 06, France, sinda.ammous-kharrat@parisdescartes.fr

1 Introduction and definitions

Let (X, Y) be a couple of real-valued and continuous random variables, and let (X^*, Y^*) be an independent copy of (X, Y). The Kendall correlation coefficient τ between X and Y is then defined by

$$\tau := 2\big(\mathbb{P}\left(\{(X^* - X)(Y^* - Y) > 0\}\right) - 0.5\big)\,. \tag{1}$$

By definition, $\tau \in [-1, 1]$. If $\tau = 0$, there is no correlation in the sense of Kendall. If $Y = f(X)$ for some increasing (resp. decreasing) function f, then $\tau = 1$ (resp. $\tau = -1$). If $\tau > 0$ (resp. $\tau < 0$), there is a positive correlation (resp. negative correlation), meaning that X and Y tend to vary in the same direction (resp. the opposite direction).
Let now $(X_i, Y_i)_{i\in\mathbb{Z}}$ be a stationary sequence of $\mathbb{R}^2$-valued random variables, with the same marginal distribution as (X, Y). To test $H_0 : \tau = 0$ against $H_1 : \tau \neq 0$ from the sequence $(X_i, Y_i)_{1\leq i\leq n}$, one can use the U-statistic

$$U_n \;=\; \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{\{(X_j - X_i)(Y_j - Y_i) > 0\}}\,, \tag{2}$$

which has been studied by Esscher [8], Lindeberg [15] [16], and Kendall [14] in the case where the random variables (X_i, Y_i) are independent and identically distributed (iid). In the general iid case (that is, without assuming that X and Y are independent), the asymptotic normal distribution of $\sqrt{n}(2U_n - \tau - 1)$ is given by Hoeffding [13] (see also van der Vaart [19], Example 12.5).
In Section 2 of the present paper, we extend Hoeffding's result to the dependent case, and we propose an estimator of the limiting covariance of $\sqrt{n}(2U_n - \tau - 1)$. From these two results we derive an asymptotically valid procedure to test $H_0 : \tau = 0$ against $H_1 : \tau \neq 0$. Our results apply to a large class of dependent sequences, under a condition on the $\overline{\beta}$-dependence coefficients of the sequence $(X_i, Y_i)_{i\in\mathbb{Z}}$, that are defined below. To be complete, we show in Section 3 that the condition on the dependency coefficients is optimal in a certain sense, and we illustrate our results trough different sets of simulation (see Section 4).

Let us now introduce these dependence coefficients.

Definition 1. *Let $Z_i = (X_i, Y_i)_{i\in\mathbb{Z}}$ be a strictly stationary sequence of random variables with values in $\mathbb{R}^2$. Let P be the law of (X_0, Y_0) and $P_{(Z_i,Z_j)}$ be the law of (Z_i, Z_j). Define the σ-algebra $\mathcal{F}_0 = \sigma(Z_k, k \leq 0)$, let $P_{Z_k|\mathcal{F}_0}$ be the conditional distribution of Z_k given $\mathcal{F}_0$, and let $P_{(Z_i,Z_j)|\mathcal{F}_0}$ be the conditional distribution of (Z_i, Z_j) given $\mathcal{F}_0$.*
For any $s, t \in \mathbb{R}, z = (x, y) \in \mathbb{R}^2$, we define the function

$$f_z(s, t) := \mathbb{1}_{x\leq s}\mathbb{1}_{y\leq t} - \mathbb{P}(X_0 \leq s, Y_0 \leq t)$$

and the random variables

$$\begin{aligned} b(k) &= \sup_{z\in\mathbb{R}^2} |\, \mathbb{P}_{Z_k|\mathcal{F}_0}(f_z)\,|, \\ b(i, j) &= \sup_{(z_1,z_2)\in\mathbb{R}^2\times\mathbb{R}^2} |\, \mathbb{P}_{(Z_i,Z_j)|\mathcal{F}_0}(f_{z_1} \otimes f_{z_2}) - \mathbb{P}_{(Z_i,Z_j)}(f_{z_1} \otimes f_{z_2})\,|, \end{aligned}$$

where as usual $f_{z_1} \otimes f_{z_2}(s, t) = f_{z_1}(t)f_{z_2}(s)$. Define now the coefficients

$$\begin{aligned} \widetilde{\beta}_1(k) &= \mathbb{E}(b(k)), \\ \widetilde{\beta}_2(k) &= \max\{\widetilde{\beta}_1(k), \sup_{i>j\geq k} \mathbb{E}\,[b(i, j)]\}, \\ \delta_2(k) &= \sup_{i>j\geq k} \mathbb{E}\left(\mathbb{1}_{X_i\leq X_j}\mathbb{1}_{Y_i\leq Y_j} - \mathbb{P}(X_i \leq X_j, Y_i \leq Y_j)|\mathcal{F}_0\right), \\ \overline{\beta}_2(k) &= \max\{\widetilde{\beta}_2(k), \delta_2(k)\}. \end{aligned}$$

Let us give some comments on these definitions. Our main result (Theorem 1 below) is stated under a condition on the coefficient $\overline{\beta}_2$. The first important remark is that this coefficient is weaker than the usual β-mixing coefficient of the sequence $(X_i, Y_i)_{i\in\mathbb{Z}}$ (see Volkonskii and Rozanov [20] for the definition of the β-mixing coefficient). Hence all the results of the present paper hold true when replacing the coefficients $\widetilde{\beta}_1$, $\widetilde{\beta}_2$, $\overline{\beta}_2$ by the usual β-mixing coefficients.
But in fact the coefficient $\overline{\beta}_2$ can be computed for a large class of processes, including many non-mixing sequences in the sense of Rosenblatt [18]. This follows mostly from the paper by Dedecker and Prieur [4], which provides many examples of (possibly non-mixing) processes for which the coefficient $\widetilde{\beta}_2$ can be easily controlled (see also the monograph [2] for more examples and a comparison with other notions of dependency). The coefficient $\overline{\beta}_2$ is a bit more restrictive than $\widetilde{\beta}_2$, because of the term δ_2, which is not so easy to handle. However, in many cases (if not all) the coefficient δ_2 may be handled as $\widetilde{\beta}_2$ by following the thread of Section 6 in Dedecker-Prieur [4].
Let us give a simple example. Assume that $Z_i = (X_i, Y_i)^t$ is a $\mathbb{R}^2$-valued linear process, defined by

$$Z_i = \sum_{k=0}^{\infty} A_k \varepsilon_{i-k}\,,$$

where (ε_i) is a sequences of iid $\mathbb{R}^2$-valued random variables with mean 0 and square integrable coordinates, and A_k is a deterministic sequence of 2×2 matrices such that $\sum_{k\geq 0}|A_k|^2 < \infty$ (here $|A_k|$ is the usual norm $|A_k| = \sup_{\|x\|=1}|A_k x|$, and $\|\cdot\|$ is the euclidean norm). Let $F_{X,0}$ and $F_{Y,0}$ be the distribution functions of X_0 and Y_0, and, for $i>0$, let

$$F_{X,i}(t) = \mathbb{P}(X_i - X_0 \leq t) \quad \text{and} \quad F_{Y,i}(t) = \mathbb{P}(Y_i - Y_0 \leq t)\,.$$

If the functions $F_{X,i}, F_{Y,i}$ are uniformly Hölder of order $\gamma \in (0,1]$, meaning that there exists a positive constant C such that

$$\sup_{i\geq 0} |F_{X,i}(s) - F_{Y,i}(t)| \leq C|t-s|^{\gamma}\,,$$

then (following [4], Section 6),

$$\overline{\beta}_2(n) \leq C\left(\sum_{k\geq n}|A_k|^2\right)^{\frac{\gamma}{\gamma+2}}.$$

In particular, if $|A_k|$ is geometrically decreasing, then so is $\overline{\beta}_2(k)$ (whatever the index γ).
Note that, without extra assumptions on the distribution of ε_0, such linear processes have no reasons to be mixing in the sense of Rosenblatt. For instance, it is well known that the $\mathbb{R}$-valued linear process

$$X_i = \sum_{k\geq 0}\frac{\varepsilon_{i-k}}{2^{k+1}}, \quad \text{where } \mathbb{P}(\varepsilon_1 = -1/2) = \mathbb{P}(\varepsilon_1 = 1/2) = 1/2,$$

is not α-mixing (see for instance Bradley [1]). Hence, if $(Y_i)_{i\in\mathbb{Z}}$ is a sequence of iid random variables, independent of $(X_i)_{i\in\mathbb{Z}}$, then the sequence $(X_i, Y_i)_{i\in\mathbb{Z}}$ is not α-mixing. By contrast, for this particular example, one can check that $\overline{\beta}_2(k)$ is geometrically decreasing.

To conclude this section, let us quote that the asymptotic normality of Kendall's U-statistic for dependent sequences has been recently established by Dehling *et al.* [6] (note that these authors are able to deal with a large class of U-statistics, and that they also prove a functional central limit theorem for U-processes). Their result is valid for a large class of dependent sequences (including non-mixing sequences in the sense

of Rosenblatt [18]), with a large intersection with our class of $\overline{\beta}_2$-dependent sequences. The advantage of our approach is that we are able to prove that our condition on the coefficient $\overline{\beta}_2$ is optimal in some sense (see Proposition 2 below). Moreover, If we consider only the class of β-mixing sequences, then our results are valid under the condition $\sum_{k>0}\beta(k)<\infty$, while the condition in [6] cannot be better than $\sum_{k>0}k\beta(k)<\infty$.

2 Main results

As in the introduction, $(X_i,Y_i)_{i\in\mathbb{Z}}$ is a stationary sequence of $\mathbb{R}^2$-valued random variables, with the same marginal distribution as (X,Y), and we denote by (X^*,Y^*) an independent copy of (X,Y). Recall that Kendall's correlation coefficient τ is defined in (1). Let then

$$\pi := \frac{\tau}{2}+0.5 = \mathbb{P}\left(\{(X^*-X)(Y^*-Y)>0\}\right). \tag{3}$$

Define also

$$F(x,y)=\mathbb{P}\left(X<x,Y<y\right),\ H(x,y)=\mathbb{P}\left(X>x,Y>y\right),$$

and

$$F_X(x)=\mathbb{P}\left(X<x\right),\ H_X(x)=\mathbb{P}\left(X>x\right).$$

Our main result is the following theorem.

Theorem 1. *Let $(X_i,Y_i)_{i\in\mathbb{Z}}$ be a stationary sequence of $\mathbb{R}^2$-valued random variables. Assume that*

$$\sum_{k=1}^{\infty}\widetilde{\beta}_1(k)<\infty \quad \text{and} \quad k\,\overline{\beta}_2(k)\underset{k\to+\infty}{\longrightarrow}0 \tag{4}$$

then,

$$\sqrt{n}\left(U_n-\pi\right)\underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}}\mathcal{N}\left(0,V\right),$$

where

$$V=4\ \mathrm{Var}(F(X_0,Y_0)+H(X_0,Y_0))+8\sum_{k=1}^{\infty}\mathrm{Cov}\left(F(X_0,Y_0)+H(X_0,Y_0),F(X_k,Y_k)+H(X_k,Y_k)\right). \tag{5}$$

Note that the statistic $\sqrt{n}(U_n-0.5)$ cannot be used directly to test $H_0:\tau=0$ against $H_1:\tau\neq 0$. Indeed, according to Theorem 1, the asymptotic distribution of $\sqrt{n}(U_n-0.5)$ under H_0 depends on the unknown quantity V. To resolve this problem, we propose in the next proposition a consistent estimator of V.

Proposition 1. *Let $(X_i,Y_i)_{i\in\mathbb{Z}}$ be a stationary sequence of $\mathbb{R}^2$-valued random variables. Assume that*

$$\sum_{k=1}^{\infty}\widetilde{\beta}_2(k)<\infty. \tag{6}$$

Let

$$F_n(s,t)=\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{X_i<t}\mathbb{1}_{Y_i<t},\quad H_n(s,t)=\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{X_i>s}\mathbb{1}_{Y_i>t},\quad G_n(s,t)=2\left(F_n(s,t)+H_n(s,t)\right),$$

and

$$\widehat{\gamma}(k) = \frac{1}{n}\sum_{i=1}^{n-k}(G_n(X_i, Y_i) - \bar{G}_n)(G_n(X_{i+k}, Y_{i+k}) - \bar{G}_n),$$

where $\bar{G}_n = \frac{1}{n}\sum_{i=1}^{n} G_n(X_i, Y_i)$. *Let* (a_n) *be a sequence of positive integers tending to infinity as* n *tends to infinity, such that* $a_n = o\left(\sqrt{n}/(\log n)^2\right)$. *Then,*

$$V_n = \widehat{\gamma}(0) + 2\sum_{k=1}^{a_n}\widehat{\gamma}(k)$$

converges in $\mathbb{L}^2$ *to the quantity* V *defined in* (5).

Combining Theorem 1 and Proposition 1, we obtain that, under $H_0 : \tau = 0$ and if $V > 0$, the random variables

$$T_n := \frac{\sqrt{n}(U_n - 1/2)}{\sqrt{|V_n|}} \quad \text{converges in distribution to } \mathcal{N}(0,1). \tag{7}$$

Therefore, for a significance level α, the rejection region of the corrected Kendall test is of the form $R_{n,\alpha} = \{|T_n| > q_\alpha\}$ where q_α is the quantile of order $1 - (\alpha/2)$ of the standard normal distribution.

Remark 1. *The choice of the sequence* (a_n) *is a delicate matter. If the coefficients* $\widetilde{\beta}_2(k)$ *decrease very quickly, then* a_n *should increase very slowly (it suffices to take* $a_n \equiv 0$ *in the iid setting). On the contrary, if* $\widetilde{\beta}_2(k) = O(k^{-1}(\log k)^{-a})$ *for some* $a > 1$, *then the terms in the covariance series have no reason to be small, and one should take* a_n *close to* $\sqrt{n}$ *to estimate many of these covariance terms. A data-driven criterion for choosing* a_n *is an interesting (but probably difficult) question, which is beyond the scope of the present paper.*
However, from a practical point of view, there is an easy way to proceed: one can plot the estimated covariances $\widehat{\gamma}(k)$*'s and choose* a_n *(not too large) in such a way that*

$$\widehat{\gamma}(0) + 2\sum_{k=1}^{a_n}\widehat{\gamma}(k)$$

should represent an important part of the unknown covariance series V *defined in* (5). *As we shall see in the simulations (Section 4), if the decay of the covariances terms*

$$\gamma(k) = 2\mathrm{Cov}\left(F(X_0, Y_0) + H(X_0, Y_0), F(X_k, Y_k) + H(X_k, Y_k)\right)$$

is not too slow, this provides an easy and reasonable choice for a_n.
To conclude this remark, note that an estimator of V *similar to* V_n *is also proposed in [6].*

3 Optimality of the dependency conditions

In this section, we prove that the dependency condition (4) is essentially optimal. More precisely, we give an example of a β-mixing sequences $(X_i, Y_i)_{i\in\mathbb{Z}}$ for which $\beta(n) \sim \frac{1}{n}$, and such that $\sqrt{n}\,(U_n - \pi)$ does not converge in distribution.

Proposition 2. *There exists a stationary Markov chain $(X_i)_{i\in\mathbb{Z}}$ with β-mixing coefficients $\beta(n) \sim \frac{1}{n}$ and invariant distribution $U[-\frac{1}{2},\frac{1}{2}]$, such that*

$$\frac{\sqrt{n}}{\sqrt{\log n}}(U_n(X,X^2)-0.5) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)\,, \tag{8}$$

where

$$U_n(X,X^2) = \frac{2}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{j=1\\ i<j}}^{n} \mathbb{1}_{\{(X_j-X_i)(X_j^2-X_i^2)>0\}}. \tag{9}$$

Remark 2. *Taking $Y_i = X_i^2$, we infer from Proposition 2 that $\sqrt{n}\,(U_n - 0.5)$ does not converge in distribution, and is not even stochastically bounded. This proves that the conclusion of Theorem 1 cannot be true in general if we take $\overline{\beta}_2(n) \sim \frac{1}{n}$.*

Proof. We start from the Markov chain introduced by Doukhan, Massart and Rio [7].
Let λ be the uniform distribution on $[0,1]$, and let ν be the probability with density $g(x) = 2x\mathbf{1}_{[0,1]}$. We define now a strictly stationary Markov chain by specifying its transition probabilities $K(x,A)$ as follows:

$$K(x,A) = (1-x)\delta_x(A) + x\nu(A)\,,$$

where δ_x denotes the Dirac measure. Then λ is the unique invariant probability measure of the chain with transition probabilities $K(x,\cdot)$. Let $(Z_i)_{i\in\mathbb{Z}}$ be the stationary Markov chain on $[0,1]$ with transition probabilities $K(x,\cdot)$ and invariant distribution λ. From [7], we know that the β-mixing coefficients of this chain are such that $\beta(n) \sim \frac{1}{n}$.
We now define the random variables $X_i = Z_i - 0.5$. Hence $(X_i)_{i\in\mathbb{Z}}$ is a stationary Markov chain whose β-mixing coefficients are such that $\beta(n) \sim \frac{1}{n}$. Moreover the X_i's are uniformly distributed over $[-0.5, 0.5]$. Let $Y_i = X_i^2$. As quoted in Remark 3, the statistic $\sqrt{n/\log n}(U_n(X,X^2)-0.5)$ in Proposition 2 is exactly $\sqrt{n/\log n}(U_n - 0.5)$. As in (11) (see the proof of Theorem 1), we have the Hoeffding decomposition (used with $\pi = 1/2$):

$$\frac{\sqrt{n}}{\sqrt{\log n}}(U_n - 0.5) = T_n + R_n\,, \tag{10}$$

where

$$\begin{aligned}
T_n &:= \frac{2}{\sqrt{n\log n}}\sum_{i=1}^{n}\left(F(X_i,Y_i) + H(X_i,Y_i) - 0.5\right),\\
R_n &:= \frac{2}{(n-1)\sqrt{n\log n}}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(f(Z_i,Z_j) + f(Z_j,Z_i)\right),\\
f(Z_i,Z_j) &:= \mathbb{1}_{X_i<X_j}\mathbb{1}_{Y_i<Y_j} - F(X_j,Y_j) - H(X_i,Y_i) + 1/4.
\end{aligned}$$

From the proof of Proposition 3, we get the upper bound

$$\mathbb{E}(R_n^2) \le \frac{C}{n\log n}\left(1 + \sum_{k=1}^{n} k\beta(k)\right),$$

for some positive constant C. Since $\beta(k) \sim \frac{1}{k}$, we easily infer that R_n converges to 0 in $\mathbb{L}^2$. Hence, it remains to prove (8) with T_n instead of $\sqrt{n/\log n}(U_n(X,X^2) - 0.5)$.

Let us compute $F(X, X^2)$ and $H(X, X^2)$. For $x \in \mathbb{R}$ and $y > 0$,

$$F(x,y) = P(X < x, X^2 < y) = P(X < x, |X| < \sqrt{y}) = P(X < x, -\sqrt{y} < X < \sqrt{y})$$

Consequently

$$\begin{aligned} F(x,y) &= P(\min(x, -\sqrt{y}) < X < \min(x, \sqrt{y})) \\ &= F_X(\min(x, \sqrt{y})) - F_X(\min(x, -\sqrt{y})). \end{aligned}$$

Since $\min(X, |X|) = X$ and $\min(X, -|X|) = -|X|$, we infer that

$$F(X, X^2) = F_X(X) - F_X(-|X|) = F_X(X) - H_X(|X|)\,.$$

In the same way, for $x \in \mathbb{R}$ and $y > 0$,

$$\begin{aligned} H(x,y) &= P(X > x, X^2 > y) = P(X > x, |X| > \sqrt{y}) \\ &= P(X > \max(x, \sqrt{y})) + P(x < X < -\sqrt{y}) \\ &= H_X(\max(x, \sqrt{y})) + [F_X(-\sqrt{y}) - F_X(x)]\, \mathbb{1}_{x<-\sqrt{y}} \end{aligned}$$

In our case, since $\mathbb{1}_{X<-|X|} = 0$, we infer that

$$H(X, X^2) = H_X(\max(X, |X|)) = H_X(|X|).$$

Altogether, this proves

$$F(X, X^2) + H(X, X^2) = F_X(X).$$

Note also that, since $X_i = Z_i - 0.5$, we have: $F_X(X_i) = Z_i$. Consequently

$$T_n = \frac{2}{\sqrt{n \log n}} \sum_{i=1}^{n} (Z_i - 0.5)\,.$$

Proposition 2 then follows from Lemma 1 below, whose proof will be done in Appendix.

Lemma 1. *Let $(Z_i)_{i\in\mathbb{Z}}$ be the stationary Markov chain on $[0,1]$ with transition probabilities $K(x,\cdot)$ and invariant distribution λ. Then*

$$\frac{2}{\sqrt{n \log n}} \sum_{i=1}^{n} (Z_i - 0.5) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)\,.$$

Remark 3. *This lemma can be proved by using Proposition 4 in [21]. However, checking the conditions given in that Proposition is quite long and not so easy. For the sake of clarity, we will provide a direct proof of Lemma 1, going back to the initial result by Feller [10].*

4 Simulations

In this section, we compare the usual Kendall test with our corrected test in three different cases.

4.1 First example

In this first example, we consider two stationary sequences (X_i) and (Y_i), with (X_i) independent of (Y_i). More precisely, we shall simulate X_i and Y_i according to the auto-regressive mechanisms:

- $\begin{cases} X_i = \frac{1}{2}(X_{i-1} + \varepsilon_i) & \text{with } (\varepsilon_i)_{i\geq 1} \text{ iid, and } \varepsilon_i \backsim \mathcal{B}\left(\frac{1}{2}\right) \\ X_0 \backsim \mathcal{U}[0,1] & \text{with } X_0 \text{ independent of } (\varepsilon_i)_{i\geq 1} \end{cases}$
- $\begin{cases} Y_i = \frac{1}{2}(Y_{i-1} + \varepsilon'_i) & \text{where } (\varepsilon'_i)_{i\geq 0} \text{ iid, and } \varepsilon'_i \backsim \mathcal{B}\left(\frac{1}{2}\right) \\ Y_0 \backsim \mathcal{U}[0,1] & \text{with } Y_0 \text{ independent of } (\varepsilon'_i)_{i\geq 1} \end{cases}$

We assume moreover that $(X_0, (\varepsilon_i))$ is independent of $(Y_0, (\varepsilon'_i))$, so that the sequence (Y_i) is an independent copy of (X_i). Moreover, it is well known that the uniform distribution $\mathcal{U}[0,1]$ is the unique invariant distribution of the auto-regressive process (X_i); consequently, the two sequences (X_i) and (Y_i) are strictly stationary. Note also that the stationary process (X_i, Y_i) is not mixing in the sense of Rosenblatt (see for instance [18]). However, it follows from [4] that the coefficients $\overline{\beta}_2(k)$ converge to zero at an exponential rate.

Since, for any positive integer i, the random variable X_i is independent of Y_i, it follows that $\pi = 1/2$. Hence, the statistic T_n defined in (7) converges in distribution to the $\mathcal{N}(0,1)$ distribution as $n \to \infty$.

We now simulate the random variables (X_i, Y_i) for $i = 1, \ldots, n$, and we study the behavior of T_n for different choices of a_n (recall that a_n appears in the definition of the estimator V_n). As explained in Remark 2, the choice of a_n may be done by analyzing the graph of the auto-covariances $\widehat{\gamma}(k)$ defined in Proposition 1.

We compute T_n for different choices of n from 150 to 600. We estimate the quantities $\mathrm{Var}(T_n)$ and $\mathbb{P}(|T_n| > 1.96)$ (the estimated level) via a classical Monte-Carlo procedure, by averaging over $N = 2000$ independent trials. This procedure will also be applied in the two following Subsections 4.2 and 4.3.

If a_n is well chosen, the estimate of $\mathrm{Var}(T_n)$ should be close to 1 and the estimated level should be close to 0.05. The graph of the auto-covariances suggests a choice of $a_n = 1$ or $a_n = 2$ (see Figure 1).

The results for $a_n = 1$ and $a_n = 2$ are presented below. We also give the rejection frequency of the usual (non corrected) Kendall test.

- $a_n = 1$

n	150	200	250	300	350	400	500	600
Estimated variance	1.146	1.173	1.125	1.1667	1.138	1.072	1.147	1.115
Estimated level	0.071	0.074	0.066	0.076	0.063	0.055	0.064	0.064
Kendall test	0.119	0.131	0.129	0.133	0.126	0.116	0.136	0.128

- $a_n = 2$

n	150	200	250	300	350	400	500	600
Estimated variance	1.177	1.077	1.045	1.079	1.066	1.138	1.043	1.066
Estimated level	0.066	0.067	0.061	0.057	0.0595	0.067	0.051	0.052
Kendall test	0.132	0.127	0.1145	0.125	0.129	0.144	0.125	0.131

As suggested by Figure 1, the choice $a_n = 1$ or $a_n = 2$ gives a reasonable estimated variance.

However, for $a_n = 2$, the estimated variance is closer to 1 and the estimated level of the corrected test is closer to 0.05. For $a_n = 2$ the estimated level lies always between 5% and 7% even for moderately large samples ($n = 150$); for $n \geq 500$ it is around 0.052, which is quite satisfactory.

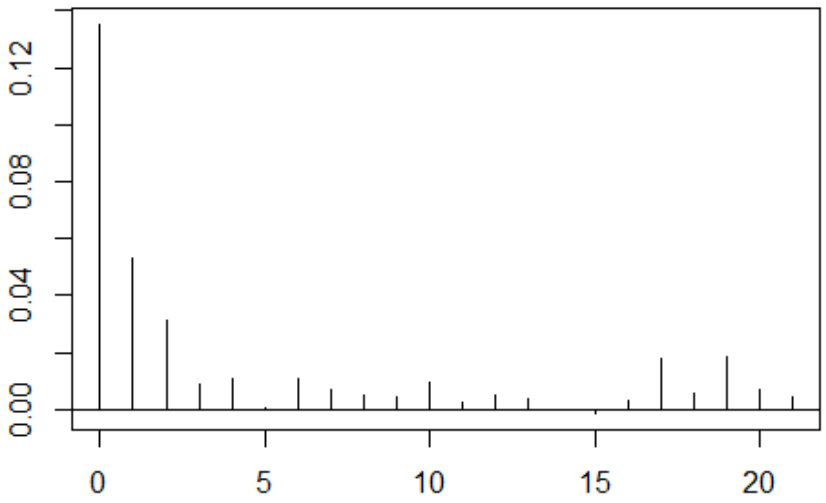

Figure 1: Graph of the auto-covariances $\widehat{\gamma}(k)$ for example 1, with $n = 150$

The estimated level of the uncorrected Kendall test is around 0.13 instead of 0.05; this is due to the fact that the usual Kendall test does not take into account the dependency of the variables.

4.2 Second example

This is an example where $(X_i, Y_i)_{i\in\mathbb{Z}}$ is a sequence of iid random vectors, but the random variables X_i and Y_i are dependent.
More precisely, let $(X_i)_{i\in\mathbb{Z}}$ and $(\varepsilon_i)_{i\in\mathbb{Z}}$ be two independent sequences of iid random variables, with $X_i \backsim \mathcal{U}[0,1]$ and $\mathbb{P}(\varepsilon_i = 1) = \mathbb{P}(\varepsilon_i = -1) = \frac{1}{2}$. Define then $Y_i = X_i\varepsilon_i$.
First, we compute the value of π.

$$\begin{aligned}
\pi &= \mathbb{P}\left((X_2 - X_1)(Y_2 - Y_1) > 0\right) \\
&= \mathbb{P}\left(\{(X_2 - X_1)^2 > 0\} \cap \{\varepsilon_1 = \varepsilon_2 = 1\}\right) + \mathbb{P}\left(\{-(X_2 - X_1)^2 > 0\} \cap \{\varepsilon_1 = \varepsilon_2 = -1\}\right) \\
&\qquad + \mathbb{P}\left(\{(X_2 - X_1)(X_2 + X_1) > 0\} \cap \{\varepsilon_1 = -1, \varepsilon_2 = 1\}\right) \\
&\qquad + \mathbb{P}\left(\{-(X_2 - X_1)(X_2 + X_1) > 0\} \cap \{\varepsilon_1 = 1, \varepsilon_2 = -1\}\right) \\
&= \frac{1}{4} + 0 + \frac{1}{8} + \frac{1}{8} = \frac{1}{2}.
\end{aligned}$$

For this example, it is clear that X_i and Y_i are not independent; however $\pi = 1/2$ (and hence $\tau = 0$) which means that there is no correlation in the sense of Kendall.

For the simulations, since we are in the usual situation where $(X_i, Y_i)_{i\in\mathbb{Z}}$ is a sequence of independent and identically distributed random vectors, we take $a_n = 0$, and we use the statistic T_n defined in (7) with $V_n = \widehat{\gamma}(0)$. The results are given below

n	150	200	250	300	350	400	500	600
Estimated variance	1.105	1.077	1.07	1.085	1.045	1.035	0.981	1.029
Estimated level	0.059	0.059	0.058	0.055	0.056	0.056	0.051	0.051
Kendall test	0.248	0.255	0.253	0.254	0.26	0.262	0.242	0.265

As expected, the estimated variance is around 1 and the estimated level of the corrected test is around 0.05, even for moderately large samples.
It is important to notice that the usual Kendall test is not well calibrated in that case, with an estimated significance level around 0.25 instead of 0.05. The reason is in fact simple: the usual Kendall test is well calibrated if X_i and Y_i are independent (because in that case the term $\mathrm{Var}(F(X_1, Y_1) + H(X_1, Y_1))$ can be explicitly computed), but it is not under the general hypothesis $H_0 : \pi = 1/2$. The conclusion is that: even in the usual case where $(X_i, Y_i)_{i\in\mathbb{Z}}$ is a sequence of iid random vectors, a correction should be made on the usual Kendall test. More precisely, to get an asymptotically well calibrated test procedure, the statistic T_n should be used with $V_n = \widehat{\gamma}(0)$.

4.3 Third example

In this last example, the X_i's are dependent random variables, and so are the Y_i's. Moreover, the variables X_i and Y_i are also dependent.
Let first (Z_i) be generated according to the auto-regressive mechanism:

$$\begin{cases} Z_i = \frac{1}{2}\left(Z_{i-1} + \varepsilon_i\right) & \text{with } (\varepsilon_i)_{i\geq 1} \text{ iid, and } \varepsilon_i \backsim \mathcal{B}\left(\frac{1}{2}\right) \\ Z_0 \backsim \mathcal{U}\left[0, 1\right] & \text{with } Z_0 \text{ independent of } (\varepsilon_i)_{i\geq 1}. \end{cases}$$

Define then $X_i = Z_i - 0.5$ and $Y_i = X_i^2$. Once again, one can easily check that $\pi = 0.5$, meaning that there is no correlation in the sense of Kendall.
For the simulations, the graph of the auto-covariances suggests to take $a_n = 4$ or $a_n = 5$ (see Figure 2). The results for $a_n = 4$ and $a_n = 5$ are given below.
• $a_n = 4$

n	150	200	250	300	350	400	500	600
Estimated variance	1.395	1.288	1.181	1.242	1.138	1.124	1.087	1.093
Estimated level	0.096	0.083	0.072	0.078	0.066	0.064	0.065	0.064
Kendall test	0.525	0.515	0.518	0.522	0.516	0.516	0.514	0.513

• $a_n = 5$

n	150	200	250	300	350	400	500	600
Estimated variance	1.383	1.341	1.178	1.284	1.155	1.113	1.051	1.044
Estimated level	0.092	0.088	0.074	0.082	0.068	0.063	0.06	0.059
Kendall test	0.512	0.53	0.50	0.522	0.522	0.495	0.502	0.507

One can see that the choices $a_n = 4$ and $a_n = 5$ lead to similar results, except for large sample ($n \geq 500$) where the estimated level is slightly better for $a_n = 5$ (around 6%), in accordance with Proposition 1.
For this example, the usual Kendall test leads to a disastrous result, with an estimated level around 51%.

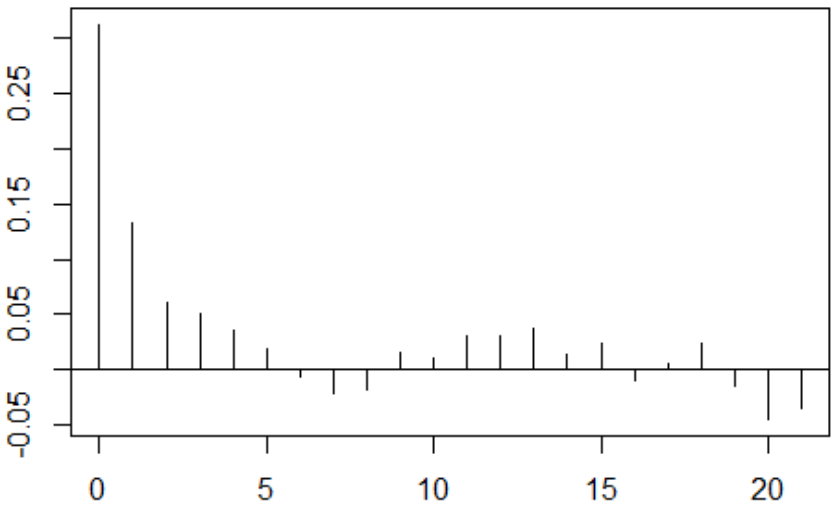

Figure 2: Graph of the auto-covariances $\widehat{\gamma}(k)$ for example 3, with n=150

5 Proofs

5.1 Proof of Theorem 1

We first recall the Hoeffding decomposition (see for instance Hoeffding [13]):

$$\sqrt{n}(U_n - \pi) = T_n + R_n \,, \tag{11}$$

$$\begin{aligned} \text{where} \quad T_n &:= \frac{2}{\sqrt{n}} \sum_{i=1}^{n} \left(F(X_i, Y_i) + H(X_i, Y_i) - \pi \right), \\ R_n &:= \frac{2}{\sqrt{n}(n-1)} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left(f(Z_i, Z_j) + f(Z_j, Z_i) \right), \\ \text{and} \quad f(Z_i, Z_j) &:= \mathbb{1}_{X_i < X_j} \mathbb{1}_{Y_i < Y_j} - F(X_j, Y_j) - H(X_i, Y_i) + \pi/2. \end{aligned}$$

The term T_n is the main term of this decomposition, and the term R_n is asymptotically negligible, as shown by the following proposition.

Proposition 3. *Let $(X_i, Y_i)_{i \in \mathbb{Z}}$ be a stationary sequence of $\mathbb{R}^2$-valued random variables. If*

$$k\, \overline{\beta}_2(k) \underset{k \to +\infty}{\longrightarrow} 0 \,,$$

then

$$R_n \underset{n \to +\infty}{\overset{\mathbb{L}^2}{\longrightarrow}} 0.$$

Let us first admit this proposition. It remains to show that, under H_0 and some condition on the $\widetilde{\beta}$-dependence coefficient,

$$T_n \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,V)\,. \tag{12}$$

According to Gordin [11] or Dedecker and Rio [5], if

$$\sum_{i=1}^{n} \|\mathbb{E}\left(F(X_i,Y_i)+H(X_i,Y_i)-\pi|\mathcal{F}_0\right)\|_1 < +\infty, \tag{13}$$

then T_n converges in distribution to a mixture of Gaussian random variables. Now, if moreover $\widetilde{\beta}_2(k)\to 0$ as $k\to\infty$ then the conditional variance in [5] is non random and (12) holds. Hence, it remains to check (13).
We first note that $\pi = 2\pi_1$ where $\pi_1 = \mathbb{P}(X^*<X, Y^*<Y)$. Indeed,

$$\begin{aligned}\pi &:= \mathbb{P}\left(\{(X^*-X)(Y^*-Y)>0\}\right)\\ &= \mathbb{P}\left(\{\{X^*<X\}\cap\{Y^*<Y\}\}\cup\{\{X^*>X\}\cap\{Y^*>Y\}\}\right)\\ &= \mathbb{P}\left(\{\{X^*<X\}\cap\{Y^*<Y\}\}\right)+\mathbb{P}\left(\{\{X^*>X\}\cap\{Y^*>Y\}\}\right) = 2\pi_1.\end{aligned}$$

In order to verify (13), we control the conditional expectation of $F(X_i,Y_i)+H(X_i,Y_i)-\pi$.
Note first that,

$$\mathbb{E}\left(F(X_1,Y_1)\right) = \mathbb{E}\left[\mathbb{1}_{(X^*<X_1, Y^*<Y_1)}\right] = \mathbb{P}\left(X^*<X_1, Y^*<Y_1\right) = \pi_1. \tag{14}$$

Hence

$$\mathbb{E}\left[F(X_i,Y_i)-\pi_1|\mathcal{F}_0\right] = \int \left[\mathbb{E}\left(\mathbb{1}_{x<X_i}\mathbb{1}_{y<Y_i}|\mathcal{F}_0\right)-\mathbb{E}\left(\mathbb{1}_{x<X_i}\mathbb{1}_{y<Y_i}\right)\right]\mathbb{P}_{(X,Y)}(dx,dy)\,,$$

and by definition of the coefficient $\widetilde{\beta}_1$,

$$\|\mathbb{E}\left[F(X_i,Y_i)-\pi_1|\mathcal{F}_0\right]\|_1 \le \int \|\mathbb{E}\left(\mathbb{1}_{x<X_i}\mathbb{1}_{y<Y_i}|\mathcal{F}_0\right)-\mathbb{E}\left(\mathbb{1}_{x<X_i}\mathbb{1}_{y<Y_i}\right)\|_1\,\mathbb{P}_{(X,Y)}(dx,dy) \le \widetilde{\beta}_1(i). \tag{15}$$

It follows that

$$\begin{aligned}&\sum_{i=1}^{n}\|\mathbb{E}\left(F(X_i,Y_i)+H(X_i,Y_i)-\pi|\mathcal{F}_0\right)\|_1\\ &\qquad\le \sum_{i=1}^{n}\|\mathbb{E}\left(F(X_i,Y_i)-\pi_1|\mathcal{F}_0\right)\|_1+\sum_{i=1}^{n}\|\mathbb{E}\left(H(X_i,Y_i)-\pi_1|\mathcal{F}_0\right)\|_1 \le 2\sum_{i=1}^{n}\widetilde{\beta}_1(i).\end{aligned}$$

Hence, (13) is satisfied as soon as $\sum_{k=1}^{\infty}\widetilde{\beta}_1(k)<\infty$, which concludes the proof of Theorem 1.

5.2 Proof of Proposition 3

We shall prove that:

$$\mathbb{E}\left(\left(\frac{2}{n\sqrt{n}}\sum_{i=1}^{n}\sum_{\substack{j=1\\ i<j}}^{n} f(Z_i,Z_j)\right)^2\right) = \frac{4}{n^3}\sum_{i=1}^{n}\sum_{\substack{j=1\\ i<j}}^{n}\sum_{k=1}^{n}\sum_{\substack{l=1\\ k<l}}^{n}\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right] \xrightarrow[n\to+\infty]{} 0\,. \tag{16}$$

There are many different cases, but it suffices to deal with the sum over the two sets

$$\Gamma_1 = \{(i,j,k,l) : i < j \leq k < l\} \quad \text{and} \quad \Gamma_2 = \{(i,j,k,l) : i < k < l < j\},$$

and the other cases can be handled in the same way (up to index permutations).
To prove the inequality (16), it remains to prove that

$$\frac{1}{n^3} \sum_{(i,j,k,l)\in\Gamma_1} |\mathbb{E}[f(Z_i,Z_j)f(Z_k,Z_l)]| \underset{n\to+\infty}{\longrightarrow} 0 \quad \text{and} \quad \frac{1}{n^3} \sum_{(i,j,k,l)\in\Gamma_2} |\mathbb{E}[f(Z_i,Z_j)f(Z_k,Z_l)]| \underset{n\to+\infty}{\longrightarrow} 0.$$

For each case we proceed in two steps:

• First case.
Step 1. Let $(i,j,k,l) \in \Gamma_{1,1} = \{(i,j,k,l) : i < j \leq k < l, l-k > k-j\}$. We start from the inequality

$$\begin{aligned}|\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right]| &= \int |\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)|Z_i=z_i,Z_j=z_j,Z_k=z_k\right]|\,\mathbb{P}_{(Z_i,Z_j,Z_k)}(\mathrm{d}z_i,\mathrm{d}z_j,\mathrm{d}z_k)\\ &\leq \int |f(z_i,z_j)|\cdot|\mathbb{E}\left[f(z_k,Z_l)|Z_i=z_i,Z_j=z_j,Z_k=z_k\right]|\,\mathbb{P}_{(Z_i,Z_j,Z_k)}(\mathrm{d}z_i,\mathrm{d}z_j,\mathrm{d}z_k).\end{aligned}$$

Since $|f(x,y)| \leq 2$,

$$\begin{aligned}&|\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right]|\\ &\leq 2\int |\mathbb{E}\left[\mathbb{1}_{x_k<X_l}\mathbb{1}_{y_k<Y_l} - F(X_l,Y_l) - H(x_k,y_k) + \pi_1|Z_i=z_i,Z_j=z_j,Z_k=z_k\right]|\,\mathbb{P}_{(Z_i,Z_j,Z_k)}(\mathrm{d}z_i,\mathrm{d}z_j,\mathrm{d}z_k)\\ &\leq 2\int |\mathbb{E}\left[\mathbb{1}_{x_k<X_l}\mathbb{1}_{y_k<Y_l} - H(x_k,y_k)|Z_i=z_i,Z_j=z_j,Z_k=z_k\right]|\,\mathbb{P}_{(Z_i,Z_j,Z_k)}(\mathrm{d}z_i,\mathrm{d}z_j,\mathrm{d}z_k)\\ &\qquad + 2\int |\mathbb{E}\left[F(X_l,Y_l) - \pi_1|Z_i=z_i,Z_j=z_j,Z_k=z_k\right]|\,\mathbb{P}_{(Z_i,Z_j,Z_k)}(\mathrm{d}z_i,\mathrm{d}z_j,\mathrm{d}z_k).\end{aligned}$$

For the first term on the right hand side, we use the fact that

$$\mathbb{E}\left[\mathbb{1}_{x_k<X_l}\mathbb{1}_{y_k<Y_l}\right] = \mathbb{P}(x_k<X_l, y_k<Y_l) = H(x_k,y_k),$$

and the definition of $\widetilde{\beta}_1$. It follows that

$$\begin{aligned}&\int |\,\mathbb{E}\left[\mathbb{1}_{x_k<X_l}\mathbb{1}_{y_k<Y_l} - H(x_k,y_k)|Z_i=z_i,Z_j=z_j,Z_k=z_k\right]|\,\mathbb{P}_{(Z_i,Z_j,Z_k)}(\mathrm{d}z_i,\mathrm{d}z_j,\mathrm{d}z_k)\\ &\leq \int \sup_{(x,y)\in\mathbb{R}^2} |\mathbb{E}\left[\mathbb{1}_{x<X_l}\mathbb{1}_{y<Y_l} - H(x,y)|Z_i=z_i,Z_j=z_j,Z_k=z_k\right]|\,\mathbb{P}_{(Z_i,Z_j,Z_k)}(\mathrm{d}z_i,\mathrm{d}z_j,\mathrm{d}z_k)\\ &\leq \left\| \sup_{(x,y)\in\mathbb{R}^2} |\mathbb{E}\left[\mathbb{1}_{x<X_l}\mathbb{1}_{y<Y_l} - H(x,y)|Z_i,Z_j,Z_k\right]| \right\|_1 \leq \widetilde{\beta}_1(l-k).\end{aligned}$$

Let us now control the second term. From (15), we have

$$\|\mathbb{E}\left[F(X_l,Y_l) - \pi_1|Z_i,Z_j,Z_k\right]\|_1 \leq \widetilde{\beta}_1(l-k).$$

Consequently,

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{1,1}}|\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right]|\leq\frac{1}{n^3}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{k=j+1}^{n}\sum_{l=k+1}^{n}4\beta_1(l-k)\mathbb{1}_{l-k>k-j}.$$

Setting $l-k=t$, we get that

$$\begin{aligned}\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{1,1}}|\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right]|&\leq\frac{4}{n^3}\sum_{t=1}^{n}\beta_1(t)\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{k=j+1}^{n}\mathbb{1}_{k<t+j}\\&\leq\frac{2}{n}\sum_{t=1}^{n}t\cdot\widetilde{\beta}_1(t).\end{aligned}$$

Conclusion of Step 1: If $k\widetilde{\beta}_1(k)\underset{k\to+\infty}{\longrightarrow}0$ then

$$\frac{4}{n^3}\sum_{(i,j,k,l)\in\Gamma_{1,1}}|\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right]|\underset{n\to+\infty}{\longrightarrow}0. \tag{17}$$

Step 2. Let $(i,j,k,l)\in\Gamma_{1,2}=\{(i,j,k,l):i<j\leq k<l,k-j>l-k\}$. We start from the elementary decomposition

$$\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right]=\mathbb{E}\left[f(Z_i,Z_j)\left[f(Z_k,Z_l)-\mathbb{E}f(Z_k,Z_l)\right]\right]+\mathbb{E}f(Z_i,Z_j)\cdot\mathbb{E}f(Z_k,Z_l). \tag{18}$$

For the first term on the right hand side of (18), we work conditionally on Z_i,Z_j.

$$\begin{aligned}|\mathbb{E}\left[f(Z_i,Z_j)\left[f(Z_k,Z_l)-\mathbb{E}f(Z_k,Z_l)\right]\right]|&\leq\mathbb{E}\left[|f(Z_i,Z_j)\mathbb{E}\left[f(Z_k,Z_l)-\mathbb{E}f(Z_k,Z_l)|Z_i,Z_j\right]|\right]\\&\leq2\left\|\mathbb{E}\left[f(Z_k,Z_l)-\mathbb{E}f(Z_k,Z_l)|Z_i,Z_j\right]\right\|_1.\end{aligned}$$

Recall that $f(Z_k,Z_l)=\mathbb{1}_{X_k<X_l}\mathbb{1}_{Y_k<Y_l}-F(X_l,Y_l)-H(X_k,Y_k)+\pi_1$. Hence

$$\begin{aligned}\left\|\mathbb{E}\left[f(Z_k,Z_l)-\mathbb{E}f(Z_k,Z_l)|Z_i,Z_j\right]\right\|_1&=\left\|\mathbb{E}\left[\mathbb{1}_{X_k<X_l}\mathbb{1}_{Y_k<Y_l}-\mathbb{E}(\mathbb{1}_{X_k<X_l}\mathbb{1}_{Y_k<Y_l})|Z_i,Z_j\right]\right\|_1\\&+\left\|\mathbb{E}\left[F(X_l,Y_l)-\mathbb{E}F(X_l,Y_l)|Z_i,Z_j\right]\right\|_1\\&+\left\|\mathbb{E}\left[H(X_k,Y_k)-\mathbb{E}H(X_k,Y_k)|Z_i,Z_j\right]\right\|_1.\end{aligned}$$

We shall now give an upper bound for the last three terms by using the properties of the coefficient $\overline{\beta}_2$.

$$\left\|\mathbb{E}\left[\mathbb{1}_{X_k<X_l}\mathbb{1}_{Y_k<Y_l}-\mathbb{E}[\mathbb{1}_{X_k<X_l}\mathbb{1}_{Y_k<Y_l}]|Z_i,Z_j\right]\right\|_1\leq\delta_2(k-j)\leq\overline{\beta}_2(k-j). \tag{19}$$

From (15), we infer that

$$\left\|\mathbb{E}\left[H(X_k,Y_k)-\mathbb{E}(H(X_k,Y_k))|Z_i,Z_j\right]\right\|_1=\left\|\mathbb{E}\left[H(X_k,Y_k)-\pi_1|Z_i,Z_j\right]\right\|_1\leq\widetilde{\beta}_1(k-j)\leq\overline{\beta}_2(k-j). \tag{20}$$

For the last term, we use also the fact that $l-j>k-j$, and that the coefficient $\widetilde{\beta}_1$ is decreasing. So

$$\left\|\mathbb{E}\left[F(X_l,Y_l)-\mathbb{E}(F(X_l,Y_l))|Z_i,Z_j\right]\right\|_1\leq\widetilde{\beta}_1(l-j)\leq\widetilde{\beta}_1(k-j)\leq\overline{\beta}_2(k-j). \tag{21}$$

It follows from (19), (20) and (21) that

$$\left\|\mathbb{E}\left[f(Z_k,Z_l)-\mathbb{E}f(Z_k,Z_l)|Z_i,Z_j\right]\right\|_1\leq6\overline{\beta}_2(k-j).$$

Therefore,

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{1,2}}|\mathbb{E}\left[f(Z_i,Z_j)\left[f(Z_k,Z_l)-\mathbb{E}f(Z_k,Z_l)\right]\right]|\leq\frac{6}{n^3}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{k=j+1}^{n}\sum_{l=k+1}^{n}\overline{\beta}_2(k-j)\mathbb{1}_{k-j>l-k}.$$

Setting $k-j=s$ and $l-k=t$, we obtain

$$\begin{aligned}\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{1,2}}|\mathbb{E}\left[f(Z_i,Z_j)\left[f(Z_k,Z_l)-\mathbb{E}f(Z_k,Z_l)\right]\right]|&\leq\frac{6}{n^3}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{t=1}^{n}\sum_{s=1}^{n}\overline{\beta}_2(s)\mathbb{1}_{s>t}\\&\leq\frac{3}{n}\sum_{s=1}^{n}s\,\overline{\beta}_2(s).\end{aligned}$$

Hence, we have proved that: if $k\overline{\beta}_2(k)\underset{k\to+\infty}{\longrightarrow}0$, then

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{1,2}}|\mathbb{E}\left[f(Z_i,Z_j)\left[f(Z_k,Z_l)-\mathbb{E}f(Z_k,Z_l)\right]\right]|\underset{n\to+\infty}{\longrightarrow}0.\tag{22}$$

In the rest of the proof of Step 2, we control the second term on the right hand side of (18). We first note that

$$\begin{aligned}|\mathbb{E}f(Z_i,Z_j)|&\leq\int|\mathbb{E}\left[f(z_i,Z_j)|Z_i=z_i\right]|\,\mathbb{P}_{Z_i}(\mathrm{d}z_i)\\&\leq\int\sup_{z\in\mathbb{R}^2}|\mathbb{E}\left[f(z,Z_j)|Z_i=z_i\right]|\,\mathbb{P}_{Z_i}(\mathrm{d}z_i)\\&\leq\left\|\sup_{z\in\mathbb{R}^2}|\mathbb{E}\left[f(z,Z_j)|Z_i\right]|\right\|_1\\&\leq\left\|\sup_{z\in\mathbb{R}^2}\left|\mathbb{E}\left[\mathbb{1}_{x<X_j}\mathbb{1}_{y<Y_j}-H(x,y)|Z_i\right]\right|\right\|_1+\left\|\sup_{z\in\mathbb{R}^2}|\mathbb{E}\left[F(X_j,Y_j)-\pi_1|Z_i\right]|\right\|_1\\&\leq 2\widetilde{\beta}_1(j-i).\end{aligned}$$

In the same way

$$|\mathbb{E}f(Z_k,Z_l)|\leq 2\widetilde{\beta}_1(l-k).$$

Clearly

$$|\mathbb{E}f(Z_i,Z_j)\cdot\mathbb{E}f(Z_k,Z_l)|\leq 2\min\{\widetilde{\beta}_1(j-i),\widetilde{\beta}_1(l-k)\}.$$

Hence

$$\begin{aligned}\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{1,2}}|\mathbb{E}f(Z_i,Z_j)\cdot\mathbb{E}f(Z_k,Z_l)|&\leq\frac{2}{n^3}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{k=j+1}^{n}\sum_{l=k+1}^{n}\min\{\widetilde{\beta}_1(j-i),\widetilde{\beta}_1(l-k)\}\\&\leq\frac{2}{n}\sum_{p=1}^{n}\sum_{t=1}^{n}\min\{\widetilde{\beta}_1(p),\widetilde{\beta}_1(t)\}.\end{aligned}\tag{23}$$

Since the coefficients $\widetilde{\beta}_1(k)$ are decreasing, we get that $\min\{\widetilde{\beta}_1(p),\widetilde{\beta}_1(t)\}=\widetilde{\beta}_1(p)\mathbb{1}_{t<p}+\widetilde{\beta}_1(t)\mathbb{1}_{p\leq t}$, and consequently,

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{1,2}}|\mathbb{E}f(Z_i,Z_j)\cdot\mathbb{E}f(Z_k,Z_l)|\leq\frac{2}{n^2}\sum_{p=1}^{n}\sum_{t=1}^{n}\widetilde{\beta}_1(p)\mathbb{1}_{t<p}+\frac{2}{n^2}\sum_{p=1}^{n}\sum_{t=1}^{n}\widetilde{\beta}_1(t)\mathbb{1}_{p\leq t}\leq\frac{4}{n}\sum_{t=1}^{n}t\widetilde{\beta}_1(t).$$

Hence, we have proved that: if $k\widetilde{\beta}_1(k) \underset{k\to+\infty}{\longrightarrow} 0$, then

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{1,2}} |\mathbb{E}f(Z_i,Z_j)\cdot\mathbb{E}f(Z_k,Z_l)| \underset{n\to+\infty}{\longrightarrow} 0. \tag{24}$$

Conclusion of Step 2: From (22) and (24), we infer that, if $k\overline{\beta}_2(k) \underset{k\to+\infty}{\longrightarrow} 0$ then

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{1,2}} |\mathbb{E}[f(Z_i,Z_j)f(Z_k,Z_l)]| \underset{n\to+\infty}{\longrightarrow} 0. \tag{25}$$

Conclusion of the first case: From (17) and (25) we infer that, if $k\overline{\beta}_2(k) \underset{k\to+\infty}{\longrightarrow} 0$, then

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{1}} |\mathbb{E}[f(Z_i,Z_j)f(Z_k,Z_l)]| \underset{n\to+\infty}{\longrightarrow} 0. \tag{26}$$

• Second case.

Step 1. Let $(i,j,k,l)\in\Gamma_{2,1}=\{(i,j,k,l): i<k<l<j, j-l>l-k\}$. Following the same strategy as to deal with $\Gamma_{1,1}$, we have that

$$|\mathbb{E}[f(Z_i,Z_j)f(Z_k,Z_l)]| \le 4\widetilde{\beta}_1(j-l).$$

Then, setting $j-l=t$, we get

$$\begin{aligned}\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{2,1}} |\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right]| &\le \frac{4}{n^3}\sum_{i=1}^{n}\sum_{k=i+1}^{n}\sum_{l=k+1}^{n}\sum_{j=l+1}^{n}\widetilde{\beta}_1(j-l)\mathbb{1}_{j-l>l-k}\\ &\le \frac{2}{n}\sum_{t=1}^{n} t\widetilde{\beta}_1(t).\end{aligned}$$

Conclusion of Step 1: If $k\,\widetilde{\beta}_1(k) \underset{k\to+\infty}{\longrightarrow} 0$ then

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{2,1}} |\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right]| \underset{n\to+\infty}{\longrightarrow} 0. \tag{27}$$

Step 2. Let $(i,j,k,l)\in\Gamma_{2,2}=\{(i,j,k,l): i<k<l<j, l-k>j-l\}$. Let also $(Z_i^*)_{i\in\mathbb{Z}}$ be an independent copy of $(Z_i)_{i\in\mathbb{Z}}$, and let us write

$$\begin{aligned}&\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right]\\ &= \mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right] - \mathbb{E}\left[f(Z_i,Z_j^*)f(Z_k,Z_l^*)\right] + \mathbb{E}\left[f(Z_i,Z_j^*)f(Z_k,Z_l^*)\right]\\ &= \mathbb{E}\left[(f(Z_i,Z_j)f(Z_k,Z_l)) - \int f(Z_i,z_j)f(Z_k,z_l)\mathbb{P}_{(Z_j,Z_l)}(\mathrm{d}z_j,\mathrm{d}z_l)\right] + \mathbb{E}\left[f(Z_i,Z_j^*)f(Z_k,Z_l^*)\right].\end{aligned}$$

Let us first control the first term on the right hand side.

$$\begin{aligned}&\left|\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l) - \int f(Z_i,z_j)f(Z_k,z_l)\mathbb{P}_{(Z_j,Z_l)}(\mathrm{d}z_j,\mathrm{d}z_l)\right]\right|\\ &= \left|\int \mathbb{E}\left[\left[f(z_i,Z_j)f(z_k,Z_l) - \int f(z_i,z_j)f(z_k,z_l)\mathbb{P}_{(Z_j,Z_l)}(\mathrm{d}z_j,\mathrm{d}z_l)\right]|Z_i=z_i,Z_k=z_k\right]\mathbb{P}_{(Z_i,Z_k)}(\mathrm{d}z_i,\mathrm{d}z_k)\right|\\ &\le \left\|\sup_{(z_1,z_2)\in\mathbb{R}^2\times\mathbb{R}^2}|\mathbb{E}\left[f(z_1,Z_j)f(z_2,Z_l)|Z_i,Z_k\right] - \mathbb{E}\left[f(z_1,Z_j)f(z_2,Z_l)\right]|\right\|_1.\end{aligned}$$

Now, by definition of f and of the coefficients $\widetilde{\beta}_2$, and since $l-k<j-k$, we easily infer that

$$\left|\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)-\int f(Z_i,z_j)f(Z_k,z_l)\mathbb{P}_{(Z_j,Z_l)}(\mathrm{d}z_j,\mathrm{d}z_l)\right]\right|\leq 9\widetilde{\beta}_2(l-k) \tag{28}$$

Consequently, setting $l-k=t$ and $j-l=p$, we get

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{2,2}}\left|\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)-\int f(Z_i,z_j)f(Z_k,z_l)\mathbb{P}_{(Z_l,Z_j)}(\mathrm{d}z_l,\mathrm{d}z_j)\right]\right|$$
$$\leq\frac{9}{n^3}\sum_{i=1}^{n}\sum_{k=i+1}^{n}\sum_{t=1}^{n}\sum_{p=1}^{n}\widetilde{\beta}_2(t)\mathbb{1}_{p<t}\leq\frac{9}{n}\sum_{t=1}^{n}t\,\widetilde{\beta}_2(t).$$

Hence, we have proved that: if $k\,\widetilde{\beta}_2(k)\underset{k\to+\infty}{\longrightarrow}0$, then

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{2,2}}\left|\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)-\int f(Z_i,z_j)f(Z_k,z_l)\mathbb{P}_{(Z_l,Z_j)}(\mathrm{d}z_l,\mathrm{d}z_j)\right]\right|\underset{n\to+\infty}{\longrightarrow}0. \tag{29}$$

Now, it remains to control the term $\mathbb{E}[f(Z_i,Z_j^*)f(Z_k,Z_l^*)]$. Since Z and Z^* are independent, one can consider that either Z or Z^* are fixed. Hence, we obtain the upper bound

$$\left|\mathbb{E}\left[f(Z_i,Z_j^*)f(Z_k,Z_l^*)\right]\right|\leq 2\cdot\min\{\widetilde{\beta}_1(j-l),\widetilde{\beta}_1(k-i)\}.$$

Therefore, setting $p=j-l$, $t=l-k$, $s=k-i$, we get

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{2,2}}\left|\mathbb{E}\left[f(Z_i,Z_j^*)f(Z_k,Z_l^*)\right]\right|\leq\frac{2}{n^3}\sum_{i=1}^{n}\sum_{k=i+1}^{n}\sum_{l=k+1}^{n}\sum_{j=l+1}^{n}\min\{\widetilde{\beta}_1(j-l),\widetilde{\beta}_1(k-i)\}$$
$$\leq\frac{2}{n}\sum_{p=1}^{n}\sum_{s=1}^{n}\min\{\widetilde{\beta}_1(p),\widetilde{\beta}_1(s)\}$$

This upper is the same as in (23) and can be handled in the same way. It follows that: if $k\widetilde{\beta}_1(k)\underset{k\to+\infty}{\longrightarrow}0$, then

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{2,2}}\left|\mathbb{E}\left[f(Z_i,Z_j^*)f(Z_k,Z_l^*)\right]\right|\underset{n\to+\infty}{\longrightarrow}0. \tag{30}$$

Conclusion of Step 2: From (29) and (30), we infer that, if $k\widetilde{\beta}_2(k)\underset{k\to+\infty}{\longrightarrow}0$ then

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{2,2}}\left|\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right]\right|\underset{n\to+\infty}{\longrightarrow}0. \tag{31}$$

Conclusion of the second case, and of the proof: From (27) and (31), we infer that

$$\frac{1}{n^3}\sum_{(i,j,k,l)\in\Gamma_{2}}\left|\mathbb{E}\left[f(Z_i,Z_j)f(Z_k,Z_l)\right]\right|\underset{n\to+\infty}{\longrightarrow}0. \tag{32}$$

Combining (26) and (32), the proof of Proposition 3 is complete.

5.3 Proof of Proposition 1

This proof consists of three steps.

First step : We first introduce the function $G(s,t) = 2(F(s,t) + H(s,t))$. Let now

$$\gamma^*(k) = \frac{1}{n}\sum_{i=1}^{n-k}(G(Z_i) - 2\pi)(G(Z_{i+k}) - 2\pi) \quad \text{and} \quad \gamma(k) = \mathrm{Cov}\left(G(Z_0), G(Z_k)\right)$$

We shall proove that $V_n^* := \gamma^*(0) + 2\sum_{k=1}^{a_n}\gamma^*(k)$ converges in $\mathbb{L}^2$ to V defined in Theorem 1.

We first note that

$$\mathbb{E}\left(\gamma^*(k)\right) = \frac{1}{n}\sum_{i=1}^{n-k}\mathbb{E}\left((G(Z_i) - 2\pi)(G(Z_{i+k}) - 2\pi)\right) = \frac{n-k}{n}\gamma(k).$$

From (15) we get that $|\gamma(k)| \leq 4\widetilde{\beta}_1(k)$. Since $\sum_{k>0}\widetilde{\beta}_1(k) < \infty$ and since $a_n \to \infty$ as $n \to \infty$, we infer from the dominated convergence theorem that $\mathbb{E}(V_n^*)$ converges to V. Hence, it remains to show that

$$\mathrm{Var}(V_n^*) \underset{n\to+\infty}{\longrightarrow} 0. \tag{33}$$

By the ergodic theorem $\gamma^*(0)$ converge in $\mathbb{L}^2$ (and almost surely) to $\mathrm{Var}(G(X_1, Y_1))$. Hence, we only deal with the term

$$\mathrm{Var}\left(\sum_{k=1}^{a_n}\gamma^*(k)\right) = \mathrm{Var}\left(\frac{1}{n}\sum_{k=1}^{a_n}\sum_{i=1}^{n-k}(G(Z_i) - 2\pi)(G(Z_{i+k}) - 2\pi)\right).$$

For the sake of clarity, let $T_i = G(Z_i) - 2\pi$, and note that $\mathbb{E}\left(T_iT_{i+k}\right) = \gamma_k$. Then

$$\begin{aligned}
\mathrm{Var}\left(\sum_{k=1}^{a_n}\gamma^*(k)\right) &= \mathrm{Var}\left(\frac{1}{n}\sum_{k=1}^{a_n}\sum_{i=1}^{n-k}T_iT_{i+k}\right) \\
&= \frac{1}{n^2}\sum_{k=1}^{a_n}\sum_{l=1}^{a_n}\sum_{i=1}^{n-k}\sum_{j=1}^{n-l}\mathrm{Cov}\left(T_iT_{i+k}, T_jT_{j+l}\right) \\
&= \frac{1}{n^2}\sum_{k=1}^{a_n}\sum_{l=1}^{a_n}\sum_{i=1}^{n-k}\sum_{j=1}^{n-l}\mathbb{E}\left[\left(T_iT_{i+k} - \gamma_k\right)\left(T_jT_{j+l} - \gamma_l\right)\right].
\end{aligned}$$

We shall now control the terms $\mathbb{E}[(T_iT_{i+k} - \gamma_k)(T_jT_{j+l} - \gamma_l)]$ with the help of the coefficients $\widetilde{\beta}$. As usual, this control depends on the gap between $i, i+k, j$ and $j+k$. Clearly, it suffices to deal with the sum over the set $\Gamma := \{i \leq j\}$. We then consider three distinct cases.

The sum over $\Gamma_1 = \{i + k \leq j\}$. In that case

$$\left|\mathbb{E}\left[\left(T_iT_{i+k} - \gamma_k\right)\left(T_jT_{j+l} - \gamma_l\right)\right]\right| \leq 2\widetilde{\beta}_2(j - i - k).$$

Then for some positive constant C, we have (changing the indexes):

$$\frac{1}{n^2}\sum_{k=1}^{a_n}\sum_{l=1}^{a_n}\sum_{i=1}^{n-k}\sum_{j=1}^{n-l}\left|\mathbb{E}\left[\left(T_iT_{i+k} - \gamma_k\right)\left(T_jT_{j+l} - \gamma_l\right)\right]\right|\mathbb{1}_{\Gamma_1} \leq \frac{2}{n^2}\sum_{k=1}^{a_n}\sum_{l=1}^{a_n}\sum_{i=1}^{n}\sum_{j\geq 0}\widetilde{\beta}_2(j) \leq \frac{Ca_n^2}{n}.$$

The sum over $\Gamma_2 = \Gamma \cap \{j \leq i+k \leq j+l\}$

$$|\mathbb{E}\left[(T_iT_{i+k} - \gamma_k)\,(T_jT_{j+l} - \gamma_l)\right]| = |\mathbb{E}\left[(T_iT_{i+k} - \gamma_k)\,T_jT_{j+l}\right]| \leq 2\widetilde{\beta}_1(j+l-i-k).$$

Then for some positive constant C, we have:

$$\frac{1}{n^2}\sum_{k=1}^{a_n}\sum_{l=1}^{a_n}\sum_{i=1}^{n-k}\sum_{j=1}^{n-l}|\mathbb{E}\left[(T_iT_{i+k} - \gamma_k)\,(T_jT_{j+l} - \gamma_l)\right]|\mathbb{1}_{\Gamma_2} \leq \frac{2}{n^2}\sum_{k=1}^{a_n}\sum_{l=1}^{a_n}\sum_{i=1}^{n}\sum_{j\geq 0}\widetilde{\beta}_1(j) \leq \frac{Ca_n^2}{n}.$$

The sum over $\Gamma_3 = \Gamma \cap \{i+k \geq j+l\}$

$$|\mathbb{E}\left[(T_iT_{i+k} - \gamma_k)\,(T_jT_{j+l} - \gamma_l)\right]| = |\mathbb{E}\left[(T_jT_{j+l} - \gamma_l)\,T_iT_{i+k}\right]| \leq 2\widetilde{\beta}_1(i+k-j-l).$$

Then for some positive constant C, we have:

$$\frac{1}{n^2}\sum_{k=1}^{a_n}\sum_{l=1}^{a_n}\sum_{i=1}^{n-k}\sum_{j=1}^{n-l}|\mathbb{E}\left[(T_iT_{i+k} - \gamma_k)\,(T_jT_{j+l} - \gamma_l)\right]|\mathbb{1}_{\Gamma_3} \leq \frac{2}{n^2}\sum_{k=1}^{a_n}\sum_{l=1}^{a_n}\sum_{j=1}^{n}\sum_{i\geq 0}\widetilde{\beta}_1(i) \leq \frac{Ca_n^2}{n}.$$

Consequently, from the last three upper bounds, we get

$$\frac{1}{n^2}\sum_{k=1}^{a_n}\sum_{l=1}^{a_n}\sum_{i=1}^{n-k}\sum_{j=1}^{n-l}|\mathbb{E}\left[(T_iT_{i+k} - \gamma_k)\,(T_jT_{j+l} - \gamma_l)\right]|\mathbb{1}_{\Gamma} \leq \frac{3Ca_n^2}{n}.$$

Note that this result is still true on $\Gamma^c := \{i > j\}$ (interchanging i and j), in such a way that

$$\frac{1}{n^2}\sum_{k=1}^{a_n}\sum_{l=1}^{a_n}\sum_{i=1}^{n-k}\sum_{j=1}^{n-l}|\mathbb{E}\left[(T_iT_{i+k} - \gamma_k)\,(T_jT_{j+l} - \gamma_l)\right]| \leq \frac{6Ca_n^2}{n}.$$

Finally, using the fact that $a_n = o(\sqrt{n})$, we conclude that $\sum_{k=1}^{a_n}\gamma^*(k)$ converges in $\mathbb{L}^2$ to $\sum_{k\geq 1}\gamma(k)$, and

$$V_n^* \xrightarrow[n\to+\infty]{\mathbb{L}^2} V\,. \tag{34}$$

Second step : We shall prove that

$$\left\|\sup_{(s,t)\in\mathbb{R}^2}|G_n(s,t) - G(s,t)|\right\|_2^2 \leq C\frac{(\log(n))^4}{n}\sum_{k=0}^{n}\widetilde{\beta}_1(k). \tag{35}$$

In fact, it suffices to prove (35) for $F_n(s,t) - F(s,t)$, since the term involving $H_n(s,t) - H(s,t)$ can be handled similarly.
We define the empirical process by

$$\mu_n(s,t) = \sqrt{n}\,(F_n(s,t) - F(s,t))\,.$$

To simplify the rest of the proof, we reduce the interval of definition of (s,t) from $\mathbb{R}^2$ to $[0,1]^2$. We define a random variable $(U_i, V_i)_{i\in\mathbb{Z}}$ with values in $[0,1]^2$, such that $(U_i, V_i) = (F_X(X_i), F_Y(Y_i))$.

Without loss of generality, assume that the distribution functions F_X and F_Y are continuous, so that the random variables (U_i, V_i) have a uniform distribution. Then,

$$\left\| \sup_{(s,t)\in\mathbb{R}^2} |F_n(s,t) - F(s,t)| \right\|_2^2 = \left\| \sup_{(u,v)\in[0,1]^2} |\frac{1}{n}\sum_{i=1}^n \left(\mathbb{1}_{U_i\le u, V_i\le v} - \mathbb{P}(U_0\le u, V_0\le v)\right)| \right\|_2^2$$

In the rest of the proof, we apply a dyadic chaining (following [17], Chapter 7). Let K be some non-negative integer and, for any $z = (z_1, z_2)$ in the unit square $]0,1]^2$, let

$$\Pi_K(z) = (\Pi_K(z_1), \Pi_K(z_2)), \quad \text{with} \quad \Pi_K(z_1) = 2^{-K}\lceil 2^K z_1\rceil.$$

Let N be the unique integer such that $2^{N-1} < n \le 2^N$. Clearly

$$\mu_n(z) = \mu_n(z) - \mu_n(\Pi_N(z)) + \mu_n(\Pi_N(z)).$$

Consequently,

$$\sup_{z\in[0,1]^2} |\mu_n(z)| \le \underbrace{\sup_{z\in[0,1]^2} |\mu_n(z) - \mu_n(\Pi_N(z))|}_{R_N} + \underbrace{\sup_{z\in[0,1]^2} |\mu_n(\Pi_N(z))|}_{\Delta}.$$

Let us first control the main term Δ. For any $z = (z_1, z_2)$ in the unit square $]0,1]^2$, let $]0,z] =]0,z_1]\times]0,z_2]$. For any $j \in (1,2)$ and any natural integer M,

$$]0, \Pi_M(z_j)] = \bigcup_{L_j=0}^{M}]\Pi_{L_j-1}(z_j), \Pi_{L_j}(z_j)].$$

with the convention $\Pi_{-1}(z_j) = 0$. Then

$$]0, \Pi_M(z)] = \bigcup_{L\in[0,M]^2} \prod_{j=1}^{2}]\Pi_{L_j-1}(z_j), \Pi_{L_j}(z_j)].$$

Let $\mathcal{D}_L$ be the class of dyadic boxes $\prod_{i=1}^2](k_i-1)2^{-L_i}, k_i 2^{-L_i}]$ where $k = (k_1, k_2)$. Let $Z_n = \sqrt{n}(P_n - P)$ be the empirical and centered empirical measure where P denote the common marginal distribution of (U_i, V_i) and P_n the empirical measure. Define

$$\Delta_L := \sup_{S\in\mathcal{D}_L} |Z_n(S)|.$$

Then

$$\Delta \le \sum_{L\in[0,N]^2} \Delta_L. \tag{36}$$

Moreover,

$$\begin{aligned}
\|Z_n(S)\|_2^2 &= \mathbb{E}\left[\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbb{1}_{(U_i,V_i)\in S} - \mathbb{P}((U_i,V_i)\in S)\right)^2\right] \\
&\le \mathrm{Var}\left(\mathbb{1}_{(U_0,V_0)\in S}\right) + 2\sum_{k=1}^{n-1} \left|\mathbb{E}\left(\left(\mathbb{1}_{(U_0,V_0)\in S} - \mathbb{P}((U_i,V_i)\in S)\right)\left(\mathbb{1}_{(U_k,V_k)\in S} - \mathbb{P}((U_i,V_i)\in S)\right)\right)\right| \\
&\le \mathrm{Var}\left(\mathbb{1}_{(U_0,V_0)\in S}\right) + 2\sum_{k=1}^{n-1} \left|\mathbb{E}\left(\mathbb{1}_{(U_0,V_0)\in S}\left(\mathbb{1}_{(U_k,V_k)\in S} - \mathbb{P}((U_i,V_i)\in S)\right)\right)\right|.
\end{aligned}$$

Let $\mathcal{F}_0 = \sigma((U_i, V_i), i \leq 0)$, then

$$\begin{aligned}
&\leq \mathrm{Var}\left(\mathbb{1}_{(U_0,V_0)\in S}\right) + 2\sum_{k=1}^{n}\left|\mathbb{E}\left(\mathbb{1}_{(U_0,V_0)\in S}\mathbb{E}\left(\mathbb{1}_{(U_k,V_k)\in S} - \mathbb{P}((U_i,V_i)\in S)|\mathcal{F}_0\right)\right)\right| \\
&\leq \mathrm{Var}\left(\mathbb{1}_{(U_0,V_0)\in S}\right) + 2\sum_{k=1}^{n}\left|\mathbb{E}\left(\mathbb{1}_{(U_0,V_0)\in S}\, b(k)\right)\right| \\
&\leq \mathbb{E}\left(\mathbb{1}_{(U_0,V_0)\in S}\left(1 + 2\sum_{k=1}^{n} b(k)\right)\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
\|\Delta_L\|_2^2 &\leq \sum_{S\in\mathcal{D}_L}\|Z_n(S)\|_2^2 \leq \sum_{S\in\mathcal{D}_L}\mathbb{E}\left(\mathbb{1}_{(U_0,V_0)\in S}\left(1 + 2\sum_{k=1}^{n} b(k)\right)\right) \\
&\leq \mathbb{E}\left(\sum_{S\in\mathcal{D}_L}\mathbb{1}_{(U_0,V_0)\in S}\left(1 + 2\sum_{k=1}^{n} b(k)\right)\right).
\end{aligned}$$

Using the definition of $\widetilde{\beta}_1$ and the fact that $\sum_{S\in\mathcal{D}_L}\mathbb{1}_{(U_0,V_0)\in S} = 1$, we infer that

$$\|\Delta_L\|_2^2 \leq \left(1 + 2\sum_{k=1}^{n}\widetilde{\beta}_1(k)\right). \tag{37}$$

Combining (36) and (37), we obtain that

$$\|\Delta\|_2 = \sum_{L\in[0,N]^2}\|\Delta_L\|_2 \leq N^2\left(1 + 2\sum_{k=1}^{n}\widetilde{\beta}_1(k)\right)^{1/2}. \tag{38}$$

We shall now give an upper bound for the term $\mathbb{E}(R_N^2)$. Using the result of Rio [17] (Chapter 7, page 123),

$$R_N \leq 2\sqrt{n}2^{-N} + \sum_{j=1}^{2}\sup_{z_j\in[0,1]}\sqrt{n}\left(F_{n,j}(\Pi_N(z_j) + 2^{-N}) - F_{n,j}(\Pi_N(z_j))\right)$$

where $F_{n,1}$ (resp. $F_{n,2}$) is the empirical distribution function of the variables $U_1, \ldots, U_n$ (resp. $V_1, \ldots, V_n$). Let

$$\Delta_{N,j} = \sup_{z_j\in[0,1]}\sqrt{n}\left(F_{n,j}(\Pi_N(z_j) + 2^{-N}) - F_{n,j}(\Pi_N(z_j))\right)$$

In order to give an upper bound for $\Delta_{N,j}$, we use exactly the same strategy as for Δ_L (with dyadic intervals instead of dyadic boxes), which gives

$$\|\Delta_{N,j}\|_2^2 \leq 1 + 2\sum_{k=1}^{n}\widetilde{\beta}_1(k).$$

Hence

$$\|R_N\|_2 \leq 2\sqrt{n}2^{-N} + \left(1 + 2\sum_{k=1}^{n}\widetilde{\beta}_1(k)\right)^{1/2}. \tag{39}$$

Combining the inequalities (38) and (39), we obtain that there exists a positive constant K such that

$$\left\| \sup_{z\in[0,1]^2} |\mu_n(z)| \right\|_2^2 \leq Kn2^{-2N} + KN^4 \left(1 + 2\sum_{k=1}^{n} \widetilde{\beta}_1(k)\right) .$$

Taking $N = [(2\log 2)^{-1} \log n]$, we infer that there exists a positive constant C such that

$$\left\| \sup_{z\in[0,1]^2} |\mu_n(z)| \right\|_2^2 \leq C(\log(n))^4 \sum_{k=0}^{n} \beta_1(k) ,$$

which is exactly (35).

Last step :

Using (35), we see that we can replace the non-observable quantity $\gamma^*(k)$ by the estimator $\widehat{\gamma}(k)$ in the expression of V_n^*, provided that $a_n(\log(n))^2 n^{-1/2}$ tends to zero as n tends to infinity. This completes the proof of Proposition 1.

6 Appendix

In this section, we prove Lemma 1. We first recall some elementary facts given in [3]. Let $(T_k)_{k\geq 0}$ be the sequence of stopping times defined by

$$T_0 = \inf\{i > 1 : Z_i \neq Z_{i-1}\} \quad \text{and} \quad T_k = \inf\{i > T_{k-1} : Z_i \neq Z_{i-1}\} \quad \text{for } k > 0.$$

Let $\tau_k = T_{k+1} - T_k$. The random variables $(Z_{T_k}, \tau_k)_{k\geq 0}$ are iid, Z_{T_k} has law ν, and the conditional distribution of τ_k given $Z_{T_k} = x$ is the geometric distribution $\mathcal{G}(x)$. Note that τ_0 has a weak moment of order 2: there exists $c > 0$ such that $\mathbb{P}(\tau_0 > x) \leq cx^{-2}$ for any $x > 0$.
Let now $N(n) = \inf\{T_k : T_k \leq n\}$. We can write

$$\sum_{i=1}^{n} (Z_i - 0.5) = (T_0 - 1)(Z_1 - 0.5) + \sum_{k=0}^{N(n)-1} \tau_k (Z_{T_k} - 0.5) + (n - T_{N(n)})(Z_{T_{N(n)}} - 0.5) . \tag{40}$$

Clearly $(T_0 - 1)(Z_0 - 0.5)/\sqrt{n\log n}$ converges to 0 in probability, so this term is negligible for the convergence in distribution. Let us now consider the last term in (40), which is a bit more complicated. We first note that

$$|(n - T_{N(n)})(Z_{T_{N(n)}} - 0.5)| \leq n - T_{N(n)} . \tag{41}$$

Let t be any integer in $[0, n-2]$, then

$$\begin{aligned}
\mathbb{P}(n - T_{N(n)} > t) &= \sum_{k=1}^{n-t-1} \mathbb{P}(T_{N(n)} = k) = \sum_{k=1}^{n-t-1} \sum_{l=1}^{k} \mathbb{P}(T_l = k, N(n) = l) \\
&= \sum_{k=1}^{n-t-1} \sum_{l=1}^{k} \mathbb{P}(T_l = k, T_{l+1} > n) = \sum_{k=1}^{n-t-1} \sum_{l=1}^{k} \mathbb{P}(T_l = k, \tau_l > n - k) .
\end{aligned}$$

Since T_l is independent of τ_l, and since the τ_i's are iid, we get

$$\mathbb{P}(n - T_{N(n)} > t) = \sum_{k=1}^{n-t-1} \sum_{l=1}^{k} \mathbb{P}(T_l = k)\mathbb{P}(\tau_0 > n - k) \leq \sum_{k=1}^{n-t-1} \mathbb{P}(\tau_0 > n - k) .$$

Recall that τ_0 has a weak moment of order 2. Hence

$$\mathbb{P}(n - T_{N(n)} > t) \leq c \sum_{k=1}^{n-t-1} \frac{1}{(n-k)^2} \leq \frac{c'}{t} \tag{42}$$

for some $c' > 0$. From (41) and (42) we easily infer that $(n - T_{N(n)})(Z_{T_{N(n)}} - 0.5)/\sqrt{n \log n}$ converges to zero in probability.
From these considerations, we see that to prove Lemma 1, it is equivalent to prove that

$$\frac{2}{\sqrt{n \log n}} \sum_{k=0}^{N(n)-1} \tau_k(Z_{T_k} - 0.5) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1) \,. \tag{43}$$

To do this, the main point is to prove that the random variable $\tau_0(Z_{T_0} - 0.5)$ has a weak moment of order 2, and to apply a result of Feller [10] on the domain of attraction of the normal distribution. So, let us compute the tails of $\tau_0(Z_{T_0} - 0.5)$. If $t > 0$,

$$\mathbb{P}(\tau_0(Z_{T_0} - 0.5) > t) = 2 \int_{1/2}^{1} x(1-x)^{[t/(x-0.5)]} dx \,,$$

and one can easily see that

$$\lim_{t\to\infty} t^2 \mathbb{P}(\tau_0(Z_{T_0} - 0.5) > t) = 0 \,.$$

Now, for $t > 0$,

$$\mathbb{P}(\tau_0(Z_{T_0} - 0.5) < -t) = \mathbb{P}(\tau_0(0.5 - Z_{T_0}) > t) = 2 \int_{0}^{1/2} x(1-x)^{[t/(0.5-x)]} dx \,,$$

Hence

$$\mathbb{P}(\tau_0(Z_{T_0} - 0.5) < -t) = \frac{2}{t^2} \int_{0}^{t/2} y(1-(y/t))^{[t/(0.5-(y/t))]} dy \,,$$

and by the dominated convergence theorem, we obtain that

$$\lim_{t\to\infty} t^2 \mathbb{P}(\tau_0(Z_{T_0} - 0.5) < -t) = 2 \int_{0}^{\infty} y \exp(-2y) dy = 0.5 \,.$$

It follows that $\tau_0(Z_{T_0} - 0.5)$ is in the domain of attraction of the normal distribution (see Feller [10] and also Gouëzel [12], Section 1.2.2, for a short exposition), and that

$$\frac{\sqrt{2}}{\sqrt{n \log n}} \sum_{k=0}^{n-1} \tau_k(Z_{T_k} - 0.5) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1) \,. \tag{44}$$

To conclude, it remains to replace n by $N(n)$ in (44), in order to get (43). More precisely, we shall prove that,

$$\frac{1}{\sqrt{n \log n}} \left| \sum_{k=0}^{[n/\mathbb{E}(\tau_0)]-1} \tau_k(Z_{T_k} - 0.5) - \sum_{k=0}^{N(n)-1} \tau_k(Z_{T_k} - 0.5) \right| \xrightarrow[n\to+\infty]{\mathbb{P}} 0 \,. \tag{45}$$

Note that $\mathbb{E}(\tau_0) = 2$, and recall that $N(n)/n$ converges almost surely to $1/\mathbb{E}(\tau_0) = 1/2$.

Let us prove (45). Let $\epsilon > 0, \delta > 0$, and let $W_k = \tau_k(Z_{T_k} - 0.5)$. We have

$$\mathbb{P}\left(\left|\sum_{k=0}^{[n/2]-1} W_k - \sum_{k=0}^{N(n)-1} W_k\right| > \epsilon\sqrt{n\log n}\right) \leq \mathbb{P}(|N(n) - [n/2]| > \delta n) + \mathbb{P}\left(\left|\sum_{k=0}^{[n/2]-1} W_k - \sum_{k=0}^{N(n)-1} W_k\right| > \epsilon\sqrt{n\log n}\,, |N(n) - [n/2]| \leq \delta n\right) . \quad (46)$$

Now, the first term on right hand in (46) tends to zero as $n \to \infty$. For the second term, we easily see that it is smaller than

$$2\mathbb{P}\left(\max_{1\leq k\leq \delta n}\left|\sum_{i=1}^{k} W_i\right| > \epsilon\sqrt{n\log n}\right) .$$

Using Etemadi's inequality [9], we get that

$$\mathbb{P}\left(\max_{1\leq k\leq \delta n}\left|\sum_{i=1}^{k} W_i\right| > \epsilon\sqrt{n\log n}\right) \leq 3\max_{1\leq k\leq \delta n}\mathbb{P}\left(3\left|\sum_{i=1}^{k} W_i\right| > \epsilon\sqrt{n\log n}\right) . \quad (47)$$

Since $(n\log n)^{-1/2}\sum_{i=1}^{n} W_k$ converges in distribution (see (44)), we easily see that

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\max_{1\leq k\leq \delta n}\mathbb{P}\left(3\left|\sum_{i=1}^{k} W_i\right| > \epsilon\sqrt{n\log n}\right) = 0\,.$$

Going back to (46), we obtain that, for any $\epsilon > 0$,

$$\lim_{n\to\infty}\mathbb{P}\left(\left|\sum_{k=0}^{[n/2]-1} W_k - \sum_{k=0}^{N(n)-1} W_k\right| > \epsilon\sqrt{n\log n}\right) = 0\,,$$

and (45) is proved.
Combining (44) and (45), and bearing in mind that $\mathbb{E}(\tau_0) = 2$, we infer that

$$\frac{2}{\sqrt{n\log n}}\sum_{k=0}^{N(n)-1}\tau_k(Z_{T_k} - 0.5) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)\ , \quad (48)$$

which is exactly (43). This completes the proof of Lemma 1.

Acknowledgements. I would like to thank my PhD advisors, Jérôme Dedecker and Céline Duval, for their advice during the preparation of this article. I also thank Florence Merlevède for very useful discussions about the optimality of the results, and for pointing the reference [21].

References

[1] Richard C. Bradley. Basic properties of strong mixing conditions. In *Dependence in probability and statistics (Oberwolfach, 1985)*, volume 11 of *Progr. Probab. Statist.*, pages 165–192. Birkhäuser Boston, Boston, MA, 1986.

[2] Jérôme Dedecker, Paul Doukhan, Gabriel Lang, José Rafael León R., Sana Louhichi, and Clémentine Prieur. *Weak dependence: with examples and applications*, volume 190 of *Lecture Notes in Statistics*. Springer, New York, 2007.

[3] Jérôme Dedecker, Sébastien Gouëzel, and Florence Merlevède. Large and moderate deviations for bounded functions of slowly mixing Markov chains. *Stoch. Dyn.*, 18(2):38 pages, 2018.

[4] Jérôme Dedecker and Clémentine Prieur. An empirical central limit theorem for dependent sequences. *Stochastic Process. Appl.*, 117(1):121–142, 2007.

[5] Jérôme Dedecker and Emmanuel Rio. On the functional central limit theorem for stationary processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 36(1):1–34, 2000.

[6] Herold Dehling, Daniel Vogel, Martin Wendler, and Dominik Wied. Testing for changes in Kendall's tau. *Econometric Theory*, 33(6):1352–1386, 2017.

[7] Paul Doukhan, Pascal Massart, and Emmanuel Rio. The functional central limit theorem for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 30(1):63–82, 1994.

[8] Fredrick Esscher. On a method of determining correlation from the ranks of the variates. *Scandinavian Actuarial Journal*, 1924(1):201–219, 1924.

[9] Nasrollah Etemadi. On some classical results in probability theory. *Sankhyā Ser. A*, 47(2):215–221, 1985.

[10] William Feller. *An introduction to probability theory and its applications. Vol. II.* John Wiley & Sons, Inc., New York-London-Sydney, 1966.

[11] Mikhail Gordin. Abstract of communication T.1: A-K. *International Conference on Probability Theory, Vilnius*, 1973.

[12] Sébastien Gouëzel. Central limit theorem and stable laws for intermittent maps. *Probab. Theory Related Fields*, 128(1):82–122, 2004.

[13] Wassily Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics*, 19:293–325, 1948.

[14] Maurice G. Kendall. A new measure of rank correlation. *Biometrika*, 30:277–283, 1938.

[15] Jarl Waldemar Lindeberg. Über die correlation. *VI Skand. Mathematikerkongre i Kobenhawn*, pages 437–446, 1925.

[16] Jarl Waldemar Lindeberg. Some remarks of the mean error of the percentage of correlation. *Nordic Statistical Journal*, 1:137–141, 1929.

[17] Emmanuel Rio. *Asymptotic theory of weakly dependent random processes*, volume 80. Springer, 2017.

[18] Murray Rosenblatt. A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U.S.A.*, 42:43–47, 1956.

[19] Aad W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.

[20] V. A. Volkonskiĭ and Yu. A. Rozanov. Some limit theorems for random functions. I. *Theor. Probability Appl.*, 4:178–197, 1959.

[21] Ou Zhao, Michael Woodroofe, and Dalibor Volný. A central limit theorem for reversible processes with nonlinear growth of variance. *J. Appl. Probab.*, 47(4):1195–1202, 2010.