A Central Limit Theorem for Wasserstein type distances between two different laws

Philippe Berthet, Jean-Claude Fort, Thierry Klein

A Central Limit Theorem for Wasserstein type distances between two different laws

Philippe Berthet[1], Jean Claude Fort[2], and Thierry Klein* [1,3]

[1]Institut de mathématique, UMR5219; Université de Toulouse; CNRS, UPS IMT, F-31062 Toulouse Cedex 9, France
[2]MAP5, Université Paris Descartes, France
[3]ENAC - Ecole Nationale de l'Aviation Civile, Université de Toulouse, France

Abstract

This article is dedicated to the estimation of Wasserstein distances and Wasserstein costs between two distinct continuous distributions F and G on $\mathbb{R}$. The estimator is based on the order statistics of (possibly dependent) samples of F resp. G. We prove the consistency and the asymptotic normality of our estimators.
Keywords:
Central Limit Theorems- Generelized Wasserstein distances- Empirical processes- Strong approximation- Dependent samples.
MSC Classification:
62G30, 62G20, 60F05, 60F17

Contents

*Corresponding author thierry.klein@math.univ-toulouse.fr

1 Introduction

1.1 Motivation

In this article we address the problem of estimating the distance between two different distributions with respect to a class of Wasserstein costs that we define in the sequel. The framework is very simple: two samples of *i.i.d.* real random variables taking values in $\mathbb{R}$ with continuous cumulative distribution function (*c.d.f.*) F and G are available. These samples are not necessarily independent, for instance they may be issued from simultaneous experiments. From these samples we estimate the Wasserstein distances or costs between F and G and we prove a central limit theorem (CLT).

The motivation of this work is to be found in the fast development of computer experiments. Nowadays the output of many computer codes is not only a real multidimensional variable but frequently a function computed on so many points that it can be considered as a functional output. In particular this function may be the density or *c.d.f.* of a real random variable. To analyze such outputs one needs to choose a distance to compare various *c.d.f.*. Among the large possibilities offered by the literature the Wasserstein distances are now commonly used - for more details on general Wasserstein distances we refer to [17]. Since computer codes only provide samples of the underlying distributions, the estimation of such distances are of primordial importance. Actually for univariate probability distributions the p-Wasserstein distance simply is the L^p distance of simulated random variables from a common and universal (uniform on $[0,1]$)

simulator U: $W_p^p(F,G) = \int_0^1 |F^{-1}(u) - G^{-1}(u)|^p du = \mathbb{E}|F^{-1}(U) - G^{-1}(U)|^p$, where F^{-1} is the generalized inverse of F. It is then natural to estimate $W_p^p(F,G)$ by its empirical counterpart that is $W_p^p(\mathbb{F}_n, \mathbb{G}_n)$ where $\mathbb{F}_n$ and $\mathbb{G}_n$ are the empirical $c.d.f.$ of F and G build through $i.i.d.$ samples of F and G - the two samples are possibly dependent.

Many authors were interested in the convergence of $W_p^p(\mathbb{F}_n, F)$, see *e.g.* the survey paper [4] or [9, 7, 8, 1]. Up to our knowledge there are only two recent works studying the convergence of $W_2^2(\mathbb{F}_n, \mathbb{G}_n)$ [10, 16]. In [10] very general results are obtained in the multivariate setting when the two samples are independent. However the estimator in not explicit from the data, the centering in the CLT is $\mathbb{E}W_2^2(\mathbb{F}_n, \mathbb{G}_n)$ rather than $W_2^2(F,G)$ itself, and the limiting variance is also not explicit. In [16] multivariate discrete distributions and W_2 cost are considered, again for independent samples, and the CLT is explicit. In the early work [13] a trimmed version of the mallows distance or $W_2^2(\mathbb{F}_n, \mathbb{G}_n)$ is studied, still for independent samples with an explicit CLT under implicit assumptions on the level of trimming.

To investigate more deeply the univariate setting we allow a larger class of convex costs and also dependent $i.i.d.$ samples from continuous $c.d.f.$ F and G. We look for an explicit CLT for the easily computed natural estimator, under almost minimal conditions relating F and G to the cost.

1.2 Setting

Let F and G be two $c.d.f.$ on $\mathbb{R}$. The p-Wasserstein distance between F and G is defined to be

$$W_p^p(F,G) = \min_{X\sim F, Y\sim G} \mathbb{E}|X-Y|^p, \tag{1}$$

where $X \sim F, Y \sim G$ means that X and Y are random variables with respective c.d.f. F and G. The minimum in (1) has the following explicit expression

$$W_p^p(F,G) = \int_0^1 |F^{-1}(u) - G^{-1}(u)|^p du. \tag{2}$$

The Wassertein distances can be generalized to Wasserstein costs. Given a real non negative function $c(x,y)$ of two real variables, we consider the Wasserstein cost

$$W_c(F,G) = \min_{X\sim F, Y\sim G} \mathbb{E}c(X,Y). \tag{3}$$

We restrict our study to costs for which this minimum is finite and the analogue of (2) exists.

Definition 1 *We call a good cost function any application c from $\mathbb{R}^2$ to $\mathbb{R}$ that defines a negative measure on $\mathbb{R}^2$. It satisfies the "measure property" $\mathcal{P}$,*

$$\mathcal{P} : \forall x \leqslant x' \text{ and } \forall y \leqslant y',\ c(x',y') - c(x',y) - c(x,y') + c(x,y) \leqslant 0.$$

Remark 2 *It is obvious that $c(x,y) = -xy$ satisfies the $\mathcal{P}$ property and if c satisfies $\mathcal{P}$ then any function of the form $a(x) + b(y) + c(x,y)$ also satisfies $\mathcal{P}$.*

In particular $(x-y)^2 = x^2+y^2-2xy$ *satisfies* $\mathcal{P}$. *More generally if* ρ *is a convex real function then* $c(x,y) = \rho(x-y)$ *satisfies* $\mathcal{P}$. *This is the case of* $|x-y|^p$, $p \geqslant 1$ *and for the cost associated to the* α*-quantile* $c(x,y) = (x-y)(\alpha - \mathbf{1}_{x-y<0})$.

The following theorem that can be found in [5] gives an explicit formula of W_c for cost functions satisfying property $\mathcal{P}$.

Theorem 3 (Cambanis, Simon, Stout [5]) *Let* c *satisfy the "measure property"* $\mathcal{P}$ *and* U *be a random variable uniformly distributed on* $[0,1]$, *then*

$$W_c(F,G) = \int_0^1 c(F^{-1}(u), G^{-1}(u))du = \mathbb{E}\ c(F^{-1}(U), G^{-1}(U)).$$

In view of Theorem 3, an estimator of $W_c(F,G)$ based on a sample from the joint distribution of $(F^{-1}(U), G^{-1}(U))$ seems the most natural one. Nevertheless, it is not necessary and one can sample from any coupling of the marginal c.d.f. This is very interesting in practice, since we can use experimental data without any assumption on the coupling structure. We will see that it only affects the limiting variance in the CLT but not the rate of convergence.

Let $(X_i, Y_i)_{1\leqslant i\leqslant n}$ be an *i.i.d.* sample of a random vector with distribution Π and marginal c.d.f. F and G. Write $\mathbb{F}_n$ and $\mathbb{G}_n$ the random empirical *c.d.f.* built from the two marginal samples. Let c a good cost function. Denote by $X_{(i)}$ (resp. $Y_{(i)}$) the i^{th} order statistic of the sample $(X_i)_{1\leqslant i\leqslant n}$ (resp. $(Y_i)_{1\leqslant i\leqslant n}$), i.e. $X_{(1)} \leqslant \ldots \leqslant X_{(n)}$. We have

$$W_c(\mathbb{F}_n, \mathbb{G}_n) = \frac{1}{n}\sum_{i=1}^{n} c(X_{(i)}, Y_{(i)}). \tag{4}$$

Thanks to Theorem 3, $W_c(\mathbb{F}_n, \mathbb{G}_n)$ is a natural estimator of $W_c(F,G)$. The aim of this paper is to study its asymptotic properties when $F \neq G$ and F and G are continuous. Our main result is the weak convergence of

$$\sqrt{n}\left(W_c(\mathbb{F}_n, \mathbb{G}_n) - W_c(F,G)\right). \tag{5}$$

1.3 Overview of the paper.

Organization. The paper is organized as follows. Assumptions are discussed in Section 2. In Section 3 we state our main result in the form of a CLT for $\sqrt{n}\left(W_c(\mathbb{F}_n, \mathbb{G}_n) - W_c(F,G)\right)$. A few prospects are presented in Section 4. All the results are proved in Section 5. Section 6 contains the proofs of technical results used in the previous section and complements on the assumptions.

About the assumptions. In order to control the integrals $W_c(F,G)$ and $W_c(\mathbb{F}_n, \mathbb{G}_n)$ we separate out three sets of assumptions. First, about the regularity of F and G and the separation of their tails, with the convention that G has a lighter tail. Second, on the rate of increase, the regularity, the asymptotic

expansion of c and the behaviour of $c(x,y)$ close to the diagonal $y=x$. The first two sets are hereafter labelled (FG) and (C) respectively. They allow to separately select a class of probability laws and an admissible cost. The third set is labelled (CFG) and mixes the requirements on (F,G,c) making them compatible.
Conditions (C) encompass a large class of good Wasserstein costs c, but W_1 is not included - see remark 4 below. Conditions (FG) are satisfied by all classical laws of probability. It is important to point out that conditions (FG) and (CFG) are free from the joint law Π of the two samples. Given a cost c satisfying conditions (C), conditions (FG) and (CFG) provide sufficient regularity and tail conditions on F then regularity, tail and closeness conditions on G. The nice feature is that (CFG) are almost minimal to ensure that the limiting variance σ^2 satisfies $\sigma^2(\Pi,c)<+\infty$ whatever the joint law, hence for our CLT.

Method. The (F,G,c)-dependent technique of proof we propose consists in two major steps. At the first step we combine the assumptions to show that extreme tail terms and approximations in (5) can be neglected in probability. Next, large quantiles can be centered on a larger scale and their deviation is led by the two marginal empirical quantile processes. All the assumptions (C), (FG) and (CFG) are required to control the outer integral error processes at the $\sqrt{n}$ rate. At the second step, since only the most central part of integrals eventually matters in (5) it remains to prove its weak convergence to a Gaussian law. At this stage the pertinent tool is a Brownian approximation of joint non extreme quantiles. The joint distribution naturally shows up together with the CLT rate $\sqrt{n}$.

Remark 4 *The distance W_1 does not satisfy assumption $(C3)$ since the derivative of the absolute value does not vanish at 0. This is a meaningful border case since the limiting law may now depend on the set $\{F=G\}$.*

2 Notation and assumption

2.1 Notation

Let H denote the bivariate distribution function of Π, thus

$$H(x,y)=\mathbb{P}(X\leqslant x,Y\leqslant y),\quad F(x)=H(x,+\infty),\quad G(y)=H(+\infty,y).$$

For the sake of clarity, we focus on the generic case where the c.d.f F and G have positive densities $f=F'$ and $g=G'$ supported on the whole line $\mathbb{R}$. Write F^{-1} and G^{-1} their quantile functions. The tail exponential order of decay are defined to be

$$\psi_X(x)=-\log\mathbb{P}(X>x),\quad \psi_Y(x)=-\log\mathbb{P}(Y>x),\quad x\in\mathbb{R}_+, \tag{6}$$

We introduce the density quantile functions

$$h_X=f\circ F^{-1},\quad h_Y=g\circ G^{-1},$$

and their companion functions

$$H_X(u) = \frac{1-u}{F^{-1}(u)h_X(u)}, \quad H_Y(u) = \frac{1-u}{G^{-1}(u)h_Y(u)}.$$

For $k \in \mathbb{N}_*$ denote $\mathcal{C}_k(I)$ the set of functions that are k times continuously differentiable on $I \subset \mathbb{R}$, and $\mathcal{C}_0(I)$ the set of continuous functions. Let $\mathcal{M}_2\,(m, +\infty)$ be the subset of functions $\varphi \in \mathcal{C}_2$ such that φ'' is monotone on $(m, +\infty)$. Write $RV(\gamma)$ the set of regularly varying functions at $+\infty$ with index $\gamma \geqslant 0$. We consider slowly varying functions L satisfying

$$L'(x) = \frac{\varepsilon_1(x)L(x)}{x}, \quad \varepsilon_1(x) \to 0 \text{ as } x \to +\infty. \tag{7}$$

This slight restriction is explained in the Appendix at Section 6.2.1. Then for integrability reasons we impose

$$L'(x) \geqslant \frac{l_1}{x}, \quad l_1 \geqslant 1. \tag{8}$$

When $\gamma = 0$ we define

$$RV_2^+(0, m) = \{L : L \in \mathcal{M}_2\,(m, +\infty) \text{ such that (7) and (8) hold}\}.$$

When $\gamma > 0$

$$RV_2^+(\gamma, m) = \{\varphi : \varphi \in \mathcal{M}_2\,(m, +\infty), \varphi(x) = x^\gamma L(x) \text{ such that } L' \text{ obeys (7)}\}.$$

2.2 Assumption

2.2.1 Conditions (FG).

Let $m > \max(0, F^{-1}(1/2), G^{-1}(1/2))$ be large enough to satisfies all the subsequent assumptions. Let $\overline{u} = \max(F(m), G(m)) > 1/2$. We assume that there exists $\tau_0 > 0$ such that

$(FG1)$ $\quad F, G \in \mathcal{C}_2(\mathbb{R}_+), \quad f, g > 0$ on $\mathbb{R}_+$.

$(FG2)$ $\quad (1-u)\left|(\log h(u))'\right|$ is bounded on $(\bar{u}, 1), \quad h = h_X, h_Y$.

$(FG3)$ $\quad H_X, H_Y$ are bounded on $(\bar{u}, 1)$.

$(FG4)$ $\quad \tau(u) = F^{-1}(u) - G^{-1}(u) \geqslant \tau_0, \quad u \geqslant \overline{u}$.

Remark 5 *Assumption $(FG4)$ means that the right tails of F and G are asymptotically well separated. In particular it allows translation models.*
Rewritting $(FG2)$ and $(FG3)$ with the density functions we get the following equivalent conditions

$$(FG5)\ \sup_{x>m} \frac{1-F(x)}{f(x)} \left(\frac{1}{x} + \frac{|f'(x)|}{f(x)}\right) < \infty \text{ and } \sup_{x>m} \frac{1-G(x)}{g(x)} \left(\frac{1}{x} + \frac{|g'(x)|}{g(x)}\right) < \infty.$$

At Section 6.2.2 we provide a sufficient condition for (FG1), (FG2), (FG3).

Example 6 *All classical probability laws with lighter tail than a Pareto law are (FG) admissible since they are smooth enough. An example of heavy tail is the Pareto law with parameter $p>0$ for which*

$$\psi_X(x)=p\log x,\ F^{-1}(u)=(1-u)^{-1/p},\ H_X(u)=\frac{1}{p},$$

$$h_X(u)=p(1-u)^{1+1/p},\ (1-u)\left|(\log h_X(u))'\right|=\frac{1}{p}.$$

An example of light tail is the Weibull law with parameter $q>0$ for which

$$\psi_X(x)=x^q,\ F^{-1}(u)=\left(\log(1/(1-u))\right)^{1/q},\ H_X(u)=\frac{1}{q\log(1/(1-u))},$$

$$h_X(u)=q(1-u)\left(\log(1/(1-u))\right)^{1-1/q},\ (1-u)\left|(\log h_X(u))'\right|\sim\frac{1}{q}$$

and this law is log-convex if $q<1$, log-concave if $q>1$. If ψ_X is regularly varying with index $q>0$ the previous functions are only modified by a slowly varying factor, as for the Gaussian law.

2.2.2 Conditions (C)

We consider smooth Wasserstein costs satisfying property $\mathcal{P}$. We impose $(wlog)$ that $c(x,x)=0$ and assume that, for $0<\tau_1<\tau_0$ and some $\gamma\geqslant 0$

$(C1)$ $\quad c(x,y)\geqslant 0,\quad c\in\mathcal{C}_1([-m,m]\times\mathbb{R}\cup\mathbb{R}\times[-m,m])$.

$(C2)$ $c(x,y):=\rho\left(|x-y|\right)=\exp(l(|x-y|)),(x,y)\in(m,+\infty)^2,l\in RV_2^+(\gamma,\tau_1)$.

Thus c is asymptotically smooth and symetric. Moreover we need the following contraction of $c\,(x,y)$ along the diagonal $x=y$. We assume that there exists $d(m,\tau)\to 0$ as $\tau\to 0$ such that

$(C3)$ $\quad \left|c\left(x',y'\right)-c\left(x,y\right)\right|\leqslant d(m,\tau)\left(|x'-x|+|y'-y|\right)$ for $(x,y),(x',y')\in D_m(\tau)$,

where $D_m(\tau)=\{(x,y):\max(|x|,|y|)\leqslant m,|x-y|\leqslant\tau\}$.

Remark 7 *Under $(C2)$ we have*

$$\rho(|x-y|)\leqslant\rho(\max(x,y))\leqslant\rho(x)+\rho(y),\quad (x,y)\in(m,+\infty)^2.$$

Hence

$$\sup_{x>m,y>m}\frac{c(x,y)}{\rho(x)+\rho(y)}\leqslant 1. \tag{9}$$

Example 8 *Typical costs satisfying the conditions (C) are, for $\alpha>1$,*

$$c_\alpha(x,y)=|x-y|^\alpha \tag{10}$$

and, for $\beta>0$,

$$c_\beta^-(x,y)=\exp((\log\left(1+|x-y|\right))^{1+\beta})-1,\quad c_\beta^+(x,y)=\exp(|x-y|^\beta)-1. \tag{11}$$

They satisfy $(C2)$ with $\gamma=0$, $\gamma=0$ and $\gamma=\beta$ respectively.

2.2.3 Conditions (CFG)

Recall that if $(C2)$ holds the) $l \in RV_2^+(\gamma,\tau_1)$. Now when $\gamma = 0$ in order to compare the tail functions and the cost function we need

$$\limsup_{x\to+\infty} \frac{\log(xl'(x))}{\log l(x)} = 1 - \liminf_{x\to+\infty} \frac{\log(1/\varepsilon_1(x))}{\log l(x)} = \theta_1 \in [0,1]\,, \tag{12}$$

where ε_1 is defined in (7). In the case $\gamma > 0$ we set $\theta_1 = 1$. The following crucial assumption (CFG) connects the distribution's tails with the cost function.

(CFG) There exists $\theta > 1 + \theta_1$ such that $\left(\psi_X \circ l^{-1}\right)'(x) \geqslant 2 + \dfrac{2\theta}{x},\ x \geqslant l(\tau_1).$

Remark 9 *For Wasserstein distances given by $c_\alpha, \alpha > 1$, $l(x) = \alpha \log x$. We have $\gamma = 0$ and $\varepsilon_1(x) = \alpha/l(x)$ in (12) so that the restriction in (CFG) is $\theta > 1$.*

A simple sufficient condition. If we have, for some $\zeta > 2$

$$\mathbb{P}\left(X > x\right) \leqslant \frac{1}{\exp(l(x))^{\zeta}}, \quad x \in (m, +\infty)\,, \tag{13}$$

then $\psi_X(x) \geqslant \zeta l(x)$ so that (CFG) holds with arbitrarily large θ.
We use the following consequences of (CFG). Integrating (CFG) yields

$$\psi_X \circ l^{-1}(x) \geqslant 2x + 2\theta \log x + K, \quad x \geqslant l(\tau_1), \tag{14}$$

where the integrating constant K does not matter and may change from line to line. This also implies

$$\psi_X(x) \geqslant 2l(x) + 2\theta \log l(x) + K, \quad x \geqslant \tau_1, \tag{15}$$

and, more importantly for our needs, inverting (14) we obtain

$$l \circ \psi_X^{-1}(x) \leqslant \frac{x}{2} - \theta \log x + K, \quad x \geqslant \tau_1. \tag{16}$$

Now, (14) gives

$$\mathbb{P}\left(\rho(X) > x\right) = \mathbb{P}\left(l(X) > \log x\right) = \exp\left(-\psi_X \circ l^{-1}(\log x)\right) \leqslant \frac{K}{x^2(\log x)^{2\theta}}$$

and since $\theta > 1$ we have

$$\int_m^{+\infty} \sqrt{\mathbb{P}\left(\rho(X) > x\right)} dx < +\infty. \tag{17}$$

Remark 10 *This is the same kind of condition (3.4) in [4] that ensures the convergence of $W_1(\mathbb{F}_n, F)$ at rate $\sqrt{n}$. So it turns out that (17) is almost a minimal assumption in proving Theorem 14. This is clearly confirmed at Lemmas 19 and 20 establishing that the asymptotic variance of $\sqrt{n}\,(W_c(\mathbb{F}_n, \mathbb{G}_n) - W_c(F,G))$ is finite.*

Example 11 *For an over-exponential cost c_γ^+ from (11), $\gamma > 1$, (CFG) is satisfied if $\mathbb{P}(X > x) \leqslant \exp(-2x^\gamma - \delta \log x)$ with $\delta > 4(1-\gamma)$. For the Wasserstein cost c_α from (10), $\alpha > 1$, consider a Pareto law, $\psi_X(x) = \beta \log x$. Then (CFG) reads $\alpha x/\beta < x/2 - \theta \log x$, and holds if $\beta > 2\alpha$. Gaussian laws are compatible without restriction with any cost less than $\rho(x) = \exp(ax^\gamma)$, $\gamma < 2$, $a > 0$. In the case $\gamma = 2$ the variance of X has to be less than $a/4$ for (CFG) to hold, and G may be any Gaussian law different from F with smaller variance or same variance but smaller expectation.*

3 Statement of the results

3.1 Consistency

$W_c(\mathbb{F}_n, \mathbb{G}_n)$ is a consistent estimator of $W_c(F, G)$:

Theorem 12 *Assume that the good cost $c(x, y)$ is continuous, F, G are strictly increasing and $0 \leqslant c(x, y) \leqslant V(x) + V(y)$ with V a strictly increasing function such that $\mathbb{E}V(X) < +\infty$ and $\mathbb{E}V(Y) < +\infty$. Then*

$$\lim_{n\to\infty} W_c(\mathbb{F}_n, \mathbb{G}_n) = W_c(F, G) < +\infty \quad a.s.$$

3.2 A central limit theorem

Definition 13 *We say that conditions (C), (FG) and (CFG) hold if they hold for $(c(x, y), X, Y)$ as stated above and also for $(c(-x, -y), -X, -Y)$ with possibly different functions ρ, l, ψ and F again denoting the heavier tail.*

This means that the left hand tail of F and G should be reversed from $\mathbb{R}_-$ to $\mathbb{R}_+$ and obey our set of conditions and, if G has the heavier tail the couples (F, X) and (G, Y) are simply exchanged in (FG) and (CFG).

Define

$$\Pi(u, v) = \mathbb{P}\left(X \leqslant F^{-1}(u), Y \leqslant G^{-1}(v)\right)$$

then the covariance matrix

$$\Sigma(u, v) = \begin{pmatrix} \frac{\min(u,v)-uv}{h_X(u)h_X(v)} & \frac{\Pi(u,v)-uv}{h_X(u)h_Y(v)} \\ \frac{\Pi(v,u)-uv}{h_X(v)h_Y(u)} & \frac{\min(u,v)-uv}{h_Y(v)h_Y(u)} \end{pmatrix} \tag{18}$$

and the gradient

$$\nabla(u) = \left(\frac{\partial}{\partial x} c(F^{-1}(u), G^{-1}(u)), \frac{\partial}{\partial y} c(F^{-1}(u), G^{-1}(u))\right). \tag{19}$$

Our main result is the weak convergence of the empirical Wasserstein distance of (5) toward an explicit Gaussian law $\mathcal{N}$.

Theorem 14 *If (C), (FG) and (CFG) hold then*

$$\sqrt{n}\,(W_c(\mathbb{F}_n,\mathbb{G}_n)-W_c(F,G))\to_{law}\mathcal{N}(0,\sigma^2(\Pi,c))$$

with

$$\sigma^2(\Pi,c)=\int_0^1\int_0^1\nabla(u)\Sigma(u,v)\nabla(v)dudv<+\infty. \tag{20}$$

Moreover for any real sequence $\varepsilon_n\to 0$ then

$$W_{c,n}(\mathbb{F}_n,\mathbb{G}_n)=\int_{\varepsilon_n}^{1-\varepsilon_n}c(\mathbb{F}_n^{-1}(u),\mathbb{G}_n^{-1}(u))du,$$

also satisfies

$$\sqrt{n}\,(W_{c,n}(\mathbb{F}_n,\mathbb{G}_n)-W_c(F,G))\to_{law}\mathcal{N}(0,\sigma^2(\Pi,c)).$$

The result for the trimmed version $W_{c,n}$ easilly follows from the proof for W_c. Likewise slight changes in the proof of Theorem 14 yields

Theorem 15 *If (C), (FG) and (CFG) hold then*

$$\begin{aligned}\sqrt{n}\,(W_c(\mathbb{F}_n,G)-W_c(F,G))&\to_{law}\mathcal{N}(0,\sigma_x^2(F,c)),\\ \sqrt{n}\,(W_c(F,\mathbb{G}_n)-W_c(F,G))&\to_{law}\mathcal{N}(0,\sigma_y^2(G,c)),\end{aligned}$$

with

$$\begin{aligned}\sigma_x^2(F,c)&=\int_0^1\int_0^1\frac{\partial}{\partial x}c(F^{-1}(u),G^{-1}(u))\frac{\min(u,v)-uv}{h_X(u)h_X(v)}dudv<+\infty,\\ \sigma_y^2(G,c)&=\int_0^1\int_0^1\frac{\partial}{\partial y}c(F^{-1}(u),G^{-1}(u))\frac{\min(u,v)-uv}{h_Y(u)h_Y(v)}dudv<+\infty.\end{aligned}$$

In the particular case of the square Wasserstein distance and two independent samples we have

Corollary 16 *Assume that the two samples are independent, (FG) holds and $\mathbb{P}(X>x)\leqslant\dfrac{1}{x^{4+\varepsilon}}$ with $\varepsilon>0$. Then*

$$\sqrt{n}(\frac{1}{n}\sum_{i=1}^{n}(X_{(i)}-Y_{(i)})^2-W_2^2(F,G))$$

weakly converges toward a centered Gaussian random variable with variance

$$\begin{aligned}\sigma^2(F,G)=4\int_0^1\int_0^1&\left(\frac{u\wedge v-uv}{f(F^{-1}(u))f(F^{-1}(v))}+\frac{u\wedge v-uv}{g(G^{-1}(u))g(G^{-1}(v))}\right)\\ &\times(F^{-1}(u)-G^{-1}(u))(F^{-1}(v)-G^{-1}(v))dudv.\end{aligned}$$

For numerical application the following result could be useful.

Corollary 17 *Consider a family of c.d.f.* $F_{a,b}(x) = F(\frac{x-b}{a})$, $a > 0, b \in \mathbb{R}$. *Assume that* F *is symetric with variance 1, and denote* $V_4 = var(X^2)$ *where* X *has c.d.f.* F. *Then it comes*

$$\sigma^2(F_{a,b}, F_{a',b'})) = 4(a^2 + a'^2)\left((b-b')^2 + \frac{V_4}{4}(a-a')^2\right).$$

As a consequence for two distinct Gaussian laws $\mathcal{N}(\nu, \zeta^2)$ and $\mathcal{N}(\mu, \xi^2)$ we obtain $\sigma^2 = 4(\zeta^2+\xi^2)(\nu-\mu)^2 + 2(\zeta^2+\xi^2)(\zeta-\xi)^2$as in Theorem 2.2 in [14].

We now go back to our main result. It is easy to extend Theorem 14 to probability distributions supported by intervals.

Theorem 18 *Let* F *and* G *be supported by intervals. Assume that* (FG), (C) *and* (CFG) *hold. If the most lightly tailed law is compactly supported* $(FG4)$ *is discarded. Then the conclusion of Theorem 14 holds true.*

4 Conclusion

In this paper we have proved a CLT for the natural estimator of a wide class of probability distributions and Wasserstein costs. This estimator is very fastly computed. Our results concern a couple of samples having the same size but being possibly dependent, provided the marginal distributions are distinct enough. Thus it remains to handle three main problems. First, the case $F = G$ for which the speed of weak convergence could be different from the usual $\sqrt{n}$ and the limiting law could be non Gaussian. Second the case of W_1 with $F \neq G$ and $F = G$. The third problem is to extend our results to samples of different sizes without assuming independence. We will hopefully achieve these three studies in a forthcoming paper.

5 Proofs

5.1 Proof of Theorem 12

First we have

$$0 \leqslant W_c(F,G) \leqslant \int_0^1 \left(V(F^{-1}(u)) + V(G^{-1}(u))\right) du = \mathbb{E}V(X) + \mathbb{E}V(Y).$$

Since F is strictly increasing by Glivenko-Cantelli's theorem the almost sure convergence $F_n^{-1}(u) \to F^{-1}(u)$ holds for any $u \in [0,1]$. Given any $0 < \alpha < \beta < 1$, applying Dini's theorem to the increasing functions F_n^{-1} further yields

$$\lim_{n\to\infty} \sup_{u\in(\alpha,\beta)} \left|F_n^{-1}(u) - F^{-1}(u)\right| = \lim_{n\to\infty} \sup_{u\in(\alpha,\beta)} \left|G_n^{-1}(u) - G^{-1}(u)\right| = 0 \ a.s.$$

It follows, by continuity of $c(x,y)$,

$$\lim_{n\to\infty}\int_\alpha^\beta \left|c(F_n^{-1}(u),G_n^{-1}(u))-c(F^{-1}(u),G^{-1}(u))\right|du=0 \quad a.s.$$

It remains to study

$$\frac{1}{n}\sum_{i=[\beta n]}^{n} c(X_{(i)},Y_{(i)}) \leqslant \frac{1}{n}\sum_{i=[\beta n]}^{n} V(X_{(i)}) + \frac{1}{n}\sum_{i=[\beta n]}^{n} V(Y_{(i)})$$

since the lower quantiles sums can be handled similarily. Let $\beta^- < \beta$ and consider the random variable $Z_i = V(X_i)$ and $\tilde{Z}_i = 1_{Z_i \geqslant F_Z^{-1}(\beta^-)} Z_i$ where F_Z is the $c.d.f.$ of $Z = V(X)$. Since $\mathcal{E}_X(\beta^-) = \mathbb{E}\tilde{Z} \leqslant \mathbb{E}V(X) < +\infty$ it holds

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n}\tilde{Z}_i = \mathcal{E}_X(\beta^-) \quad a.s.$$

with $\mathcal{E}_X(\beta^-) \to 0$ as $\beta^- \to 1$. Since F_Z is strictly increasing and the empirical quantile of order β of $Z_1,...,Z_n$ is $V(F_n^{-1}(\beta))$ we get

$$\lim_{n\to+\infty} V(X_{([\beta n])}) = F_Z^{-1}(\beta) > F_Z^{-1}(\beta^-) \quad a.s.$$

Therefore, with probability one we ultimately have

$$\frac{1}{n}\sum_{i=[\beta n]}^{n} V(X_{(i)}) \leqslant \frac{1}{n}\sum_{i=1}^{n}\tilde{Z}_i < 2\mathcal{E}_X(\beta^-).$$

To conclude we introduce two increasing sequences $\beta_k^- < \beta_k < 1$ such that $\beta_k^- \to 1$ as $k \to +\infty$ and consider the associated $\mathcal{E}_X(\beta_k^-) = \int_{\beta_k^-}^1 V(F^{-1}(u)) \to 0$ and $\mathcal{E}_Y(\beta_k^-) = \int_{\beta_k^-}^1 V(G^{-1}(u)) \to 0$. Almost surely for all k simultaneously, using $G^{-1}(u) \leqslant F^{-1}(u)$ for u large enough, we have

$$\lim_{n\to\infty}\int_\alpha^{\beta_k} \left|c(F_n^{-1}(u),G_n^{-1}(u))-c(F^{-1}(u),G^{-1}(u))\right|du=0$$
$$\limsup_{n\to+\infty}\int_{\beta_k}^1 c(F_n^{-1}(u),G_n^{-1}(u))du \leqslant 2\mathcal{E}_X(\beta_k^-)+2\mathcal{E}_Y(\beta_k^-)$$
$$\int_{\beta_k}^1 c(F^{-1}(u),G^{-1}(u))du \leqslant \mathcal{E}_X(\beta_k^-)+\mathcal{E}_Y(\beta_k^-)$$

This proves that $W_c(\mathbb{F}_n,\mathbb{G}_n) \to W_c(F,G)$ almost surely.

5.2 Proof of Theorem 14

The proof of Theorem 14 is organised as follows.
In Section 5.2.1 we prove (20). Section 5.2.2 is dedicated to the proof of the weak

convergence of $\sqrt{n}\,(W_c(\mathbb{F}_n,\mathbb{G}_n) - W_c(F,G))$. Thanks to definition 13 we only deal with the upper part of the integral. For that purpose we split the interval $(1/2,1)$ into four parts, $(1/2,F(M))$, $(F(M),1-h_n/n)$, $(1-h_n/n,1-k_n/n)$, $(1-k_n/n,1)$, where $F(M),h_n,k_n$ will be specified further on. The first integral is the main term and the other ones will be proved to be small. We study the integral over $(1-k_n/n,1)$ in Step 1 of Section 5.2.2, the one over $(1-h_n/n,1-k_n/n)$ in Step 2 of Section 5.2.2 and the one over $(F(M),1-h_n/n)$ in Step 3 of Section 5.2.2. Finally, we deal with the main part in Step 4 of Section 5.2.2.

5.2.1 The limiting variance.

In this section we establish that (C), (FG) and (CFG) imply that $\sigma^2(\Pi,c) < +\infty$ in (20). The covariance matrix $\Sigma(u,v)$ and the gradient $\nabla(u)$ are defined at (18) and (19). It is sufficient to study the right hand tails, corresponding to the upper domain of integration $[1/2,1]^2$. As a matter of fact, this implies the same for $[0,1/2]^2$ according to Definition 13, then similar arguments hold for mixing both tails through $[1/2,1]\times[0,1/2]$ and $[0,1/2]\times[1/2,1]$ by separating the variables exactly as we show below. Hence by cutting $[1/2,1]=[1/2,\overline{u}]\cup[\overline{u},1]$ into mid quantiles and extremes we are reduced to control $\nabla(u)\Sigma(u,v)\nabla(v)$ on $[\overline{u},1]\times[\overline{u},1]$ then on $[1/2,\overline{u}]\times[1/2,1]$. The forthcoming two lemmas are then enough to conclude that (20) is true under (C), (FG) and (CFG).

Lemma 19 *Under $(C2)$, $(FG1)$, $(FG4)$ and (CFG) we have, for any $\overline{u} > F(m)$,*

$$\sigma^2(\overline{u}) = \int_{\overline{u}}^1\int_{\overline{u}}^1 \nabla(u)\Sigma(u,v)\nabla(v)dudv < +\infty.$$

Proof By $(C2)$ we have, for $x \geqslant y \geqslant m$,

$$\frac{\partial}{\partial x}c\,(x,y) = -\frac{\partial}{\partial y}c\,(x,y) = \frac{\partial}{\partial x}\rho(x-y) = l'(x-y)\rho(x-y) = \rho'(x-y).$$

By $(FG4)$ it holds $F^{-1}(u) \geqslant \tau(u) = F^{-1}(u) - G^{-1}(u) \geqslant \tau_0 > 0$ for $u > F(m)$. Thus, for $u\in[\overline{u},1]$, $\nabla(u) = (\rho'\circ\tau(u), -\rho'\circ\tau(u))$. Let us split $\sigma^2(\overline{u})$ into

$$A_1 = \int_{\overline{u}}^1\int_{\overline{u}}^1 \rho'\circ\tau(u)\frac{\min(u,v)-uv}{h_X(u)h_X(v)}\rho'\circ\tau(v)dudv$$
$$A_2 = -\int_{\overline{u}}^1\int_{\overline{u}}^1 \rho'\circ\tau(u)\frac{\Pi(v,u)-uv}{h_X(v)h_Y(u)}\rho'\circ\tau(v)dudv$$
$$A_3 = -\int_{\overline{u}}^1\int_{\overline{u}}^1 \rho'\circ\tau(u)\frac{\Pi(u,v)-uv}{h_X(u)h_Y(v)}\rho'\circ\tau(v)dudv$$
$$A_4 = \int_{\overline{u}}^1\int_{\overline{u}}^1 \rho'\circ\tau(u)\frac{\min(u,v)-uv}{h_Y(v)h_Y(u)}\rho'\circ\tau(v)dudv.$$

Observe that if $0<u<v<1$ then

$$0 \leqslant \frac{\min(u,v)-uv}{\sqrt{1-u}\sqrt{1-v}} = u\sqrt{\frac{1-v}{1-u}} \leqslant 1$$

so that we always have $0 \leqslant \min(u,v) - uv \leqslant \sqrt{1-u}\sqrt{1-v}$ and we get

$$|A_1| \leqslant \left(\int_{\overline{u}}^1 \rho' \circ \tau(u) \frac{\sqrt{1-u}}{h_X(u)} du\right)^2, \quad |A_4| \leqslant \left(\int_{\overline{u}}^1 \rho' \circ \tau(u) \frac{\sqrt{1-u}}{h_Y(u)} du\right)^2.$$

Consider the bound of $|A_1|$ first. By $(C2)$, ρ' is $\mathcal{C}_1(m,+\infty)$ and positive. Now, as $u \to 1$, $\tau(u) \geqslant \tau_0 > 0$ is either unbounded or bounded. In both cases we have

$$0 < \rho'(\tau(u)) \leqslant \max\left(\rho' \circ F^{-1}(u), \sup_{\tau_0 < x \leqslant l_2} \rho'(x)\right) \leqslant k_1 \rho' \circ F^{-1}(u)$$

for $k_1 \geqslant 1$ since by Proposition 32 the increasing function ρ is convex on $(l_2,+\infty)$ under $(C2)$. Observe that ρ is also invertible, so that $\rho(X)$ has quantile function, density function and density quantile function respectively given by

$$F^{-1}_{\rho(X)} = \rho \circ F^{-1}, \quad f_{\rho(X)} = \frac{f \circ \rho^{-1}}{\rho' \circ \rho^{-1}}, \quad h_{\rho(X)} = f_{\rho(X)} \circ F^{-1}_{\rho(X)} = \frac{h_X}{\rho' \circ F^{-1}}. \tag{21}$$

Recalling that (CFG) implies (17), the change of variable $x = \rho \circ F^{-1}(u)$ yields

$$\begin{aligned}\frac{1}{k_1} \int_{\overline{u}}^1 \rho' \circ \tau(u) \frac{\sqrt{1-u}}{h_X(u)} du &\leqslant \int_{F(m)}^1 \rho' \circ F^{-1}(u) \frac{\sqrt{1-u}}{h_X(u)} du \\ &= \int_{F(m)}^1 \frac{\sqrt{1-u}}{h_{\rho(X)}(u)} du \\ &= \int_{\rho(m)}^{+\infty} \sqrt{\mathbb{P}\left(\rho(X) > x\right)} dx < +\infty.\end{aligned}$$

Having proved that $|A_1| < +\infty$ let us next study the upper bound of $|A_4|$. Under $(C2)$ and (12) we have, for some $\varepsilon_1(x) \to \gamma$,

$$\rho'(x) = l'(x)\rho(x) = \varepsilon_1(x) \frac{l(x)}{x} \rho(x) \leqslant (1+\gamma) \frac{l(x)^{\theta_1'}}{x} \rho(x)$$

where $\theta_1' \in (\theta_1, \theta - 1)$ if $\gamma = 0$, and $\theta_1' = 1$ if $\gamma > 0$. It then follows from the change of variable $u = G(x)$ that, by setting $\phi = G^{-1} \circ F = \psi_Y^{-1} \circ \psi_X$,

$$\begin{aligned}&\int_{\overline{u}}^1 \rho' \circ \tau(u) \frac{\sqrt{1-u}}{h_Y(u)} du \\ &\leqslant (1+\gamma) \int_{G^{-1}(\overline{u})}^{+\infty} \frac{(l \circ \phi^{-1}(x))^{\theta_1'}}{\phi^{-1}(x)} \rho \circ \phi^{-1}(x) \sqrt{\mathbb{P}\left(Y > x\right)} dx.\end{aligned} \tag{22}$$

Now, by $(FG4)$ we have

$$x \leqslant \phi^{-1}(x) = F^{-1} \circ G(x) = \psi_X^{-1} \circ \psi_Y(x) = \psi_X^{-1}\left(\log\left(\frac{1}{\mathbb{P}\left(Y > x\right)}\right)\right)$$

thus by (16) we have

$$l \circ \phi^{-1}(x) \leqslant \frac{1}{2} \log\left(\frac{1}{\mathbb{P}(Y > x)}\right) - \theta \log\log\left(\frac{1}{\mathbb{P}(Y > x)}\right) + K.$$

We can bound (22) from above by

$$\begin{aligned}
&(1+\gamma)\int_{\phi(m)}^{+\infty} \frac{(l\circ\phi^{-1}(x))^{\theta_1'}}{\phi^{-1}(x)} \exp\left(l\circ\phi^{-1}(x)\right)\sqrt{\mathbb{P}(Y>x)}dx \\
&\leqslant K\int_{\phi(m)}^{+\infty} \frac{(\psi_Y(x))^{\theta_1'-\theta}}{\psi_X^{-1}\circ\psi_Y(x)}dx \\
&\leqslant K\int_{\phi(m)}^{+\infty} \frac{1}{x\,(\psi_Y(x))^{\theta-\theta_1'}}dx \\
&\leqslant K\int_{\phi(m)}^{+\infty} \frac{1}{x\,(l(x))^{\theta-\theta_1'}}dx.
\end{aligned}$$

The last inequality comes from $\psi_Y(x) > \psi_X(x)$ by $(FG4)$. If $\gamma > 0$ then $\theta - \theta_1' = \theta - 1 > 0$ and $l(x) > x^{\gamma/2}$ hence the bounding integral is finite. If $\gamma = 0$ then $l(x) \geqslant \log x$ by (8) and having enforced $\theta - \theta_1 > \theta - \theta_1' > 1$ also makes the above integral finite. We have shown that $|A_4| < +\infty$. It remains to bound $A_2 = A_3$. Since F and G are continuous it holds

$$\begin{aligned}
\Pi(u,v) &\leqslant \min\left(\mathbb{P}\left(X \leqslant F^{-1}(u)\right), \mathbb{P}(Y \leqslant G^{-1}(v))\right) = \min(u,v) \\
\Pi(u,v) &\geqslant \mathbb{P}\left(X \leqslant F^{-1}(u)\right) + \mathbb{P}(Y \leqslant G^{-1}(v)) - 1 = u + v - 1
\end{aligned}$$

and thus

$$\begin{aligned}
\Pi(u,v) - uv &\leqslant \min(u,v) - uv \leqslant \sqrt{1-u}\sqrt{1-v} \\
\Pi(u,v) - uv &\geqslant u + v - 1 - uv = -(1-u)(1-v)
\end{aligned}$$

which proves that $|\Pi(u,v) - uv| \leqslant \sqrt{1-u}\sqrt{1-v}$. Hence $A_2 = A_3$ satisfies

$$\begin{aligned}
|A_2| &\leqslant \int_{\overline{u}}^{1} \rho'\circ\tau(v)\frac{\sqrt{1-v}}{h_X(v)}dv \int_{\overline{u}}^{1} \rho'\circ\tau(u)\frac{\sqrt{1-u}}{h_Y(u)}du \\
&\leqslant k_1^2 \int_{F(m)}^{1} \rho'\circ F^{-1}(v)\frac{\sqrt{1-v}}{h_X(v)}du \int_{F(m)}^{1} \rho'\circ F^{-1}(u)\frac{\sqrt{1-u}}{h_Y(u)}du
\end{aligned}$$

where these integrals are already proved to be finite. Finally $\sigma^2(\overline{u}) = A_1 + A_2 + A_3 + A_4 < +\infty$. ■

Lemma 20 *Under $(C1)$, $(C2)$, $(FG1)$, $(FG4)$ and (CFG) we have, for any $\overline{u} > F(m)$,*

$$\begin{aligned}
\sigma_-^2(\overline{u}) &= \int_{1/2}^{\overline{u}}\int_{1/2}^{1} \nabla(u)\Sigma(u,v)\nabla(v)dudv < +\infty, \\
\sigma_+^2(\overline{u}) &= \int_{\overline{u}}^{1}\int_{1/2}^{\overline{u}} \nabla(u)\Sigma(u,v)\nabla(v)dudv < +\infty.
\end{aligned}$$

Proof Since F^{-1} and G^{-1} are bounded on $[1/2, \overline{u}]$ we have, by $(C1)$, that $\nabla(u)$ exists and is bounded on $[1/2, \overline{u}]$. Likewise $(FG1)$ ensures that h_X and h_Y are bounded on $[1/2, \overline{u}]$ hence $\Sigma(u,v)$ is bounded on $[1/2, \overline{u}]^2$. As a consequence,

$$A_0 = \int_{1/2}^{\overline{u}} \int_{1/2}^{\overline{u}} \nabla(u)\Sigma(u,v)\nabla(v)dudv, \quad |A_0| < +\infty.$$

By $(C2)$ we have $\nabla(u) = (\rho' \circ \tau(u), -\rho' \circ \tau(u))$ on $[\overline{u}, 1]$, thus

$$\begin{aligned}
A_{01} &= \int_{1/2}^{\overline{u}} \int_{\overline{u}}^{1} \frac{\partial}{\partial x} c\left(F^{-1}(u), G^{-1}(u)\right) \frac{\min(u,v) - uv}{h_X(u)h_X(v)} \rho' \circ \tau(u)dudv \\
A_{02} &= -\int_{1/2}^{\overline{u}} \int_{\overline{u}}^{1} \frac{\partial}{\partial y} c\left(F^{-1}(u), G^{-1}(u)\right) \frac{\Pi(v,u) - uv}{h_X(v)h_Y(u)} \rho' \circ \tau(u)dudv \\
A_{03} &= -\int_{1/2}^{\overline{u}} \int_{\overline{u}}^{1} \frac{\partial}{\partial x} c\left(F^{-1}(u), G^{-1}(u)\right) \frac{\Pi(u,v) - uv}{h_X(u)h_Y(v)} \rho' \circ \tau(u)dudv \\
A_{04} &= \int_{1/2}^{\overline{u}} \int_{\overline{u}}^{1} \frac{\partial}{\partial y} c\left(F^{-1}(u), G^{-1}(u)\right) \frac{\min(u,v) - uv}{h_Y(v)h_Y(u)} \rho' \circ \tau(u)dudv.
\end{aligned}$$

Along the same arguments as in Lemma 19 we have

$$|A_{01}| \leqslant I_X J_X, \quad |A_{02}| \leqslant I_Y J_X, \quad |A_{03}| \leqslant I_X J_Y, \quad |A_{04}| \leqslant I_Y J_Y,$$

where, by the previous boundedness argument on $[1/2, \overline{u}]$,

$$\begin{aligned}
I_X &= \left(\int_{1/2}^{\overline{u}} \left| \frac{\partial}{\partial x} c\left(F^{-1}(u), G^{-1}(u)\right) \right| \frac{\sqrt{1-u}}{h_X(u)} \right) < +\infty \\
I_Y &= \left(\int_{1/2}^{\overline{u}} \left| \frac{\partial}{\partial y} c\left(F^{-1}(u), G^{-1}(u)\right) \right| \frac{\sqrt{1-u}}{h_Y(u)} \right) < +\infty
\end{aligned}$$

and by (CFG), (14), (15) and (17) on $[\overline{u}, 1]$,

$$\begin{aligned}
\frac{J_X}{k_1} &= \left(\int_{\overline{u}}^{1} \rho' \circ F^{-1}(v) \frac{\sqrt{1-v}}{h_X(v)} dv \right) < +\infty \\
\frac{J_Y}{k_1} &= \left(\int_{\overline{u}}^{1} \rho' \circ F^{-1}(v) \frac{\sqrt{1-v}}{h_Y(v)} dv \right) < +\infty.
\end{aligned}$$

Therefore $\sigma_-^2(\overline{u}) = A_0 + A_{01} + A_{02} + A_{03} + A_{04} < +\infty$. In the same way the result holds for $\sigma_+^2(\overline{u})$. ■

5.2.2 Proof of the weak convergence

Step1: Extreme Values

In this first step we show that the contribution of extremes is negligible despite the rate $\sqrt{n}$. Without information on joint laws of extreme values we treat separately the upper tail of the integrals $W_c(\mathbb{F}_n, \mathbb{G}_n)$ and $W_c(F,G)$. Indeed the latter in not a centering of the former at the very end of tails so that the empirical quantile processes cannot help.

Let K_n be a positive increasing sequence such that

$$K_n \to +\infty, \quad \frac{K_n}{\log\log n} \to 0. \tag{23}$$

Define

$$k_n = \frac{\sqrt{n}}{K_n \exp\left(l \circ \psi_X^{-1}(\log n + K_n)\right)}. \tag{24}$$

Under $(C2)$ and $(FG1)$ we have $l \circ \psi_X^{-1}(x) \to +\infty$ as $x \to +\infty$ thus $k_n = o\left(\sqrt{n}/K_n\right)$. Moreover, by (16) and (23) for any $\theta' \in (1,\theta)$ and all n large enough it holds

$$k_n \geqslant \frac{c}{K_n} \exp\left(-\frac{K_n}{2} + \theta \log(\log n + K_n)\right) > (\log n)^{\theta'}. \tag{25}$$

Hence we have $k_n/\log\log n \to +\infty$ and $k_n/\sqrt{n} \to 0$. Let us define

$$D_n = \int_{1-k_n/n}^{1} c\left(F^{-1}(u), G^{-1}(u)\right) du,$$
$$S_n = \int_{1-k_n/n}^{1} c\left(\mathbb{F}_n^{-1}(u), \mathbb{G}_n^{-1}(u)\right) du = \frac{1}{n}\sum_{i=n-[k_n]}^{n} c\left(X_{(i)}, Y_{(i)}\right).$$

Lemma 21 1. *Assume that $(C2)$, $(FG1)$, $(FG4)$ and (CFG) hold. Then*

$$\sqrt{n}D_n \to 0.$$

2. *Under $(C2)$ and (CFG), we have*

$$\sqrt{n}S_n \to 0 \quad \textit{in probability.}$$

Proof

1. By C_2 and FG_4 we have

$$D_n = \int_{1-k_n/n}^{1} \rho\left(F^{-1}(u) - G^{-1}(u)\right) du \leqslant \int_{1-k_n/n}^{1} w(u) du$$

where

$$w(u) = \exp\left(l \circ F^{-1}(u)\right) = \exp\left(l \circ \psi_X^{-1}\left(\log\left(\frac{1}{1-u}\right)\right)\right).$$

Under (CFG), for $\theta > 1$ it holds, by (16),

$$l \circ F^{-1}(u) \leqslant \frac{1}{2}\log\left(\frac{1}{1-u}\right) - \theta \log\log\left(\frac{1}{1-u}\right) + K$$

thus, as $n \to +\infty$,

$$\int_{1-k_n/n}^{1} w(u)du \leqslant \left[-\frac{K\sqrt{1-u}}{\left(\log\left(1/(1-u)\right)\right)^{\theta}} \right]_{1-k_n/n}^{1} = \frac{K\sqrt{k_n/n}}{\left(\log(n/k_n)\right)^{\theta}} \to 0$$

so that $w(u)$ is integrable on $(\overline{u}, 1)$. By (CFG) $\varphi = \psi_X \circ l^{-1}$ satisfies

$$\left(\varphi^{-1}\right)'(x) = \frac{1}{\varphi' \circ \varphi^{-1}(x)} \leqslant \frac{1}{2 + 2\theta/\varphi^{-1}(x)}$$

and for x large enough,

$$\left(\varphi^{-1}\right)'(x) = \left(l \circ \psi_X^{-1}\right)'(x) \leqslant \frac{1}{2 + 2\theta/(x/2 - \theta \log x + K)} < \frac{1}{2} - \frac{\theta}{x}. \quad (26)$$

We then have

$$\left(-\left(1-u\right)w(u)\right)' = w(u)\left(1 - \left(l \circ \psi_X^{-1}\right)'\left(\log\left(\frac{1}{1-u}\right)\right)\right) > \frac{w(u)}{2}$$

which gives

$$\int_{1-k_n/n}^{1} w(u)du \leqslant 2\left[-\left(1-u\right)w(u)\right]_{1-k_n/n}^{1} \leqslant \frac{2k_n}{n} w\left(1 - \frac{k_n}{n}\right),$$

since $\lim_{u\to 1}\left(1-u\right)w(u) = 0$. Recalling (24) it follows that for n large enough,

$$\begin{aligned}
\sqrt{n}D_n &\leqslant \frac{2k_n}{\sqrt{n}} \exp\left(l \circ \psi_X^{-1}\left(\log\left(\frac{n}{k_n}\right)\right)\right) \\
&\leqslant \frac{2}{K_n} \exp\left(l \circ \psi_X^{-1}\left(\log\left(\frac{n}{k_n}\right)\right) - l \circ \psi_X^{-1}(\log n + K_n)\right).
\end{aligned}$$

By (23), (24) and (16) with $\theta > 1$ we get

$$\begin{aligned}
\log\left(\frac{n}{k_n}\right) &\sim \frac{\log n}{2} + \log K_n + l \circ \psi_X^{-1}(\log n + K_n) \\
&\leqslant \log n + \frac{K_n}{2} + \log K_n - \theta \log(\log n + K_n) + K
\end{aligned}$$

hence $\sqrt{n}D_n \leqslant 2/K_n \to 0$ as $n \to +\infty$ since $l \circ \psi_X^{-1}$ is increasing.

2. Next we control S_n the stochastic sum of extreme values. Fix $\delta > 0$ and consider the events

$$A_n = \left\{\sqrt{n}S_n \geqslant 4\delta\right\}, \quad B_{n,X} = \left\{X_{(n-[k_n])} > m\right\}, \quad B_{n,Y} = \left\{Y_{(n-[k_n])} > m\right\}.$$

We have

$$\mathbb{P}\left(A_n\right) \leqslant \mathbb{P}\left(A_n \cap B_{n,X} \cap B_{n,Y}\right) + \mathbb{P}\left(B_{n,X}^c\right) + \mathbb{P}\left(B_{n,X}^c\right).$$

Since F and G are strictly increasing we obviously have, for $\xi > 0$ and $u_0 = F(m+\xi)$, as $n \to +\infty$,

$$\begin{aligned}\mathbb{P}\left(B_{n,X}^c\right) &= \mathbb{P}\left(\mathbb{F}_n^{-1}\left(1-\frac{k_n}{n}\right) < m\right) \\ &\leqslant \mathbb{P}\left(\mathbb{F}_n^{-1}(u_0) < F^{-1}(u_0) - \xi\right) \to 0\end{aligned}$$

and likewise, $\mathbb{P}\left(B_{n,Y}^c\right) \to 0$. By (9) we can write, under $B_{n,X} \cap B_{n,Y}$,

$$\begin{aligned}\sqrt{n} S_n &\leqslant \frac{1}{\sqrt{n}} \sum_{i=n-[k_n]}^{n} \left(\rho\left(X_{(i)}\right) + \rho\left(Y_{(i)}\right)\right) \\ &\leqslant \frac{k_n+1}{\sqrt{n}} \left(\rho\left(X_{(n)}\right) + \rho\left(Y_{(n)}\right)\right)\end{aligned}$$

hence $\mathbb{P}\left(A_n \cap B_{n,X} \cap B_{n,Y}\right) \leqslant \mathbb{P}\left(C_{n,X}\right) + \mathbb{P}\left(C_{n,Y}\right)$ where

$$C_{n,X} = \left\{\rho\left(X_{(n)}\right) \geqslant \delta \frac{\sqrt{n}}{k_n}\right\}, \quad C_{n,Y} = \left\{\rho\left(Y_{(n)}\right) \geqslant \delta \frac{\sqrt{n}}{k_n}\right\}.$$

Now we have, by $(FG4)$ and since $X_1, ..., X_n$ are independent,

$$\mathbb{P}\left(C_{n,Y}\right) \leqslant \mathbb{P}\left(C_{n,X}\right) = 1 - \left(1 - \mathbb{P}\left(\rho(X) > \delta \frac{\sqrt{n}}{k_n}\right)\right)^n$$

then combining $\rho^{-1}(x) = l^{-1}(\log x)$ with (24) gives,

$$\mathbb{P}\left(\rho(X) > \delta \frac{\sqrt{n}}{k_n}\right) = \exp\left(-\psi_X \circ l^{-1}\left(\log \delta + l \circ \psi_X^{-1}\left(\log n + \log K_n\right) + \log K_n\right)\right)$$

Now by (CFG) $\psi_X \circ l^{-1}$ is increasing. As soon as $\log K_n > |\log \delta|$ we get

$$\mathbb{P}\left(\rho(X) > \delta \frac{\sqrt{n}}{k_n}\right) \leqslant \exp\left(-\psi_X \circ l^{-1}\left(l \circ \psi_X^{-1}\left(\log n + \log K_n\right)\right)\right) = \frac{1}{nK_n},$$

which yields

$$\mathbb{P}\left(C_{n,X}\right) \leqslant 1 - \exp\left(-\frac{K}{K_n}\right) \to 0.$$

We conclude that

$$\begin{aligned}\mathbb{P}\left(A_n\right) &\leqslant \mathbb{P}\left(A_n \cap B_{n,X} \cap B_{n,Y}\right) + \mathbb{P}\left(B_{n,X}^c\right) + \mathbb{P}\left(B_{n,X}^c\right) \\ &\leqslant \mathbb{P}\left(C_{n,X}\right) + \mathbb{P}\left(C_{n,X}\right) + \mathbb{P}\left(B_{n,X}^c\right) + \mathbb{P}\left(B_{n,X}^c\right)\end{aligned}$$

satisfies $\mathbb{P}\left(A_n\right) \to 0$.

■

Step2: Centered high order quantiles

This section ends the part of the proof of Theorem 14 devoted to the secondary order. We split the arguments into the three lemmas below. Remind that k_n is defined at (24). Let introduce

$$h_n = n^{\beta}, \quad \beta \in \left(\frac{1}{2}, 1\right), \quad I_n = \left[1 - \frac{h_n}{n}, 1 - \frac{k_n}{n}\right], \tag{27}$$

and define the centered random integral of non extreme tail quantiles to be

$$T_n = \int_{1-h_n/n}^{1-k_n/n} \left(c\left(\mathbb{F}_n^{-1}(u), \mathbb{G}_n^{-1}(u)\right) - c\left(F^{-1}(u), G^{-1}(u)\right)\right) du.$$

Lemma 22 *Under* $(C2)$*,* (FG) *and* (CFG) *we have*

$$\lim_{n\to+\infty} \sqrt{n} T_n = 0 \quad a.s.$$

The proof of this lemma is based on the two following lemmas whose proof are postponed in the appendix. In order to bound T_n we first evaluate the quantile empirical processes

$$\beta_n^X(u) = \sqrt{n}(\mathbb{F}_n^{-1}(u) - F^{-1}(u)), \quad \beta_n^Y(u) = \sqrt{n}(\mathbb{G}_n^{-1}(u) - G^{-1}(u)). \tag{28}$$

Lemma 23 *Define* $\Delta_n = [\overline{u}, 1 - k_n/n]$*. Under* $(FG1)$ *and* $(FG2)$ *we have*

$$\limsup_{n\to+\infty} \sup_{u\in\Delta_n} \frac{|\beta_n(u)|\, h(u)}{\sqrt{(1-u)\log\log n}} \leqslant 4 \quad a.s.$$

where $(\beta_n, h) = (\beta_n^X, h_X)$ *or* $(\beta_n, h) = (\beta_n^Y, h_Y)$*.*

In the next key lemma we have to carefully check that the conditions given at Proposition 31 are almost surely met on $I_n \subset \Delta_n$. For $u \in I_n$ and $n \geqslant 3$ define

$$\varepsilon_n(u) = \varepsilon_n^X(u) - \varepsilon_n^Y(u), \quad \varepsilon_n^X(u) = \frac{\beta_n^X(u)}{\sqrt{n}}, \quad \varepsilon_n^Y(u) = \frac{\beta_n^Y(u)}{\sqrt{n}}. \tag{29}$$

Lemma 24 *Assume that* $(C2)$*,* (FG) *and* (CFG) *hold. Then there exists* $K_2 > 0$ *such that*

$$\limsup_{n\to+\infty} \sup_{u\in I_n} \frac{\left|c\left(\mathbb{F}_n^{-1}(u), \mathbb{G}_n^{-1}(u)\right) - c\left(F^{-1}(u), G^{-1}(u)\right)\right|}{\rho' \circ F^{-1}(u) \;\; |\varepsilon_n(u)|} \leqslant K_2 \quad a.s.$$

Proof of Lemma 22
Remind notation from (24), (27) and (29). By Lemma 24 it holds, with probability one, for all n large enough

$$|T_n| \leqslant K \int_{1-h_n/n}^{1-k_n/n} \rho' \circ F^{-1}(u)\, |\varepsilon_n(u)|\, du.$$

We proceed as in the proof of Lemma 19 where similar integrable functions show up, however they have now to be integrated to sharply evaluate $\sqrt{n}\,|T_n|$. From Lemma 23 it follows, with probability one, that for all n large and all $u \in I_n \subset \Delta_n$,

$$|\varepsilon_n(u)| \leqslant \left|\frac{\beta_n^X(u)}{\sqrt{n}}\right| + \left|\frac{\beta_n^Y(u)}{\sqrt{n}}\right| \leqslant 5\sqrt{\frac{\log\log n}{n}}\left(\frac{\sqrt{1-u}}{h_X(u)} + \frac{\sqrt{1-u}}{h_Y(u)}\right). \tag{30}$$

We then compute separately the following two integrals

$$\sqrt{n}\,|T_n| \leqslant 5K\sqrt{\log\log n}\left(\int_{1-h_n/n}^{1-k_n/n} t_X(u)du + \int_{1-h_n/n}^{1-k_n/n} t_Y(u)du\right)$$

where, for $Z = X, Y$ we write $t_Z(u) = \rho' \circ F^{-1}(u)\dfrac{\sqrt{1-u}}{h_Z(u)}$.

First integral. Since ρ is convex by Proposition 32 we can use (21) as in the proof of Lemma 19 to justify the change of variable $u = F \circ \rho^{-1}(x)$ then apply (6) to $\rho^{-1}(x) = l^{-1}(\log x)$ and rewrite the first integral as

$$\begin{aligned}\int_{1-h_n/n}^{1-k_n/n} t_X(u)du &= \int_{1-h_n/n}^{1-k_n/n} \frac{\sqrt{1-u}}{h_{\rho(X)}(u)}du \\ &= \int_{b(n/h_n)}^{b(n/k_n)} \sqrt{\mathbb{P}\left(\rho(X) > x\right)}dx \\ &= \int_{b(n/h_n)}^{b(n/k_n)} \exp\left(-\frac{1}{2}\psi_X \circ l^{-1}(\log x)\right)dx\end{aligned}$$

where, by (CFG) reformulated into (16),

$$b(x) = \rho \circ F^{-1}\left(1 - \frac{1}{x}\right) = \exp\left(l \circ \psi_X^{-1}(\log x)\right) \leqslant \frac{K\sqrt{x}}{(\log x)^{\theta}}. \tag{31}$$

Equation (17) justifies that t_X is integrable since $\theta > 1$ and, by (14),

$$\exp\left(-\frac{1}{2}\psi_X \circ l^{-1}(\log x)\right) \leqslant \frac{K}{x(\log x)^{\theta}}.$$

Now observe that $\varphi = \psi_X \circ l^{-1}$ satisfies $\varphi' = (\psi_X'/l') \circ l^{-1}$ and (CFG) reads

$$\varphi'(x) \geqslant 2 + \frac{2\theta}{x}, \quad x > l(\tau_1),$$

so that we have, for all $x > b(n/h_n) > l(\tau_1)$,

$$\begin{aligned}\left(-x\exp\left(-\frac{1}{2}\varphi(\log x)\right)\right)' &= \left(\frac{1}{2}\varphi'(\log x) - 1\right)\exp\left(-\frac{1}{2}\varphi(\log x)\right) \\ &\geqslant \frac{\theta}{\log x}\exp\left(-\frac{1}{2}\varphi(\log x)\right).\end{aligned}$$

Therefore it holds, thanks to the upper bound (31) and since $b(x)$ is increasing,

$$\begin{aligned}\int_{1-h_n/n}^{1-k_n/n} t_X(u)du &\leqslant \frac{\log b(n/k_n)}{\theta}\int_{b(n/h_n)}^{b(n/k_n)} \frac{\theta}{\log x}\exp\left(-\frac{1}{2}\psi_X\circ l^{-1}(\log x)\right)dx \\ &\leqslant \frac{\log b(n)}{\theta}\left[-x\exp\left(-\frac{1}{2}\psi_X\circ l^{-1}(\log x)\right)\right]_{b(n/h_n)}^{b(n/k_n)} \\ &\leqslant \frac{K\log n}{\theta}\frac{b(n/h_n)}{\sqrt{n/h_n}} \\ &= \frac{K}{\theta(1-\beta)^\theta(\log n)^{\theta-1}}\end{aligned}$$

since $h_n = n^\beta$. This proves that

$$\lim_{n\to+\infty}\sqrt{\log\log n}\int_{1-h_n/n}^{1-k_n/n} t_X(u)du = 0.$$

Second integral. Next consider

$$J_n = \int_{1-h_n/n}^{1-k_n/n} t_Y(u)du = \int_{1-h_n/n}^{1-k_n/n} l'\circ F^{-1}(u)\frac{\sqrt{1-u}}{h_Y(u)}\rho\circ F^{-1}(u)du.$$

By (49) and (7), under $(C2)$ we have $l'(x) = \varepsilon_1(x)l(x)/x$ with $\varepsilon_1(x)\to\gamma$ as $x\to+\infty$. If $\gamma=0$ the rate of $\varepsilon_1(x)$ is given by (12) and we pick $\theta_1'\in(\theta_1,\theta-1)$. If $\gamma>0$ let $\theta_1'=1$. Recall that $\phi^{-1}=F^{-1}\circ G=\psi_X^{-1}\circ\psi_Y$. Start with

$$\begin{aligned}J_n &\leqslant (1+\gamma)\int_{1-h_n/n}^{1-k_n/n}\frac{(l\circ F^{-1}(u))^{\theta_1'}}{F^{-1}(u)}\frac{\sqrt{1-u}}{h_Y(u)}\rho\circ F^{-1}(u)du \\ &= (1+\gamma)\int_{G^{-1}(1-h_n/n)}^{G^{-1}(1-k_n/n)}\frac{(l\circ\phi^{-1}(x))^{\theta_1'}}{\phi^{-1}(x)}\exp\left(l\circ\phi^{-1}(x)\right)\sqrt{\mathbb{P}(Y>x)}dx.\end{aligned}$$

Observe that (CFG) and (16) imply

$$l\circ\phi^{-1}(x) = l\circ\psi_X^{-1}\circ\psi_Y(x)\leqslant\frac{\psi_Y(x)}{2}-\theta\log\psi_Y(x)+K.$$

Since $\psi_X^{-1}\circ\psi_Y(x)\geqslant x$ by $(FG4)$ and $\psi_Y'(x)\geqslant K/x$ by $(FG5)$ it readily follows, for $\theta-\theta_1'>1$ and $K>0$,

$$\begin{aligned}J_n &\leqslant (1+\gamma)\int_{G^{-1}(1-h_n/n)}^{G^{-1}(1-k_n/n)}\frac{(\psi_Y(x))^{\theta_1'-\theta}}{\psi_X^{-1}\circ\psi_Y(x)}dx \\ &\leqslant K\int_{G^{-1}(1-h_n/n)}^{G^{-1}(1-k_n/n)}\frac{\psi_Y'(x)}{(\psi_Y(x))^{\theta-\theta_1'}}dx \\ &= K\left[\frac{-1}{(\psi_Y(x))^{\theta-\theta_1'-1}}\right]_{\psi_Y^{-1}(\log(n/h_n))}^{\psi_Y^{-1}(\log(n/k_n))} \\ &\leqslant \frac{K}{((1-\beta)\log n)^{\theta-\theta_1'-1}}\end{aligned}$$

therefore

$$\lim_{n\to+\infty}\sqrt{\log\log n}\int_{1-h_n/n}^{1-k_n/n} t_Y(u)du = 0.$$

As a conclusion, the almost sure upper bound of $\sqrt{n}\,|T_n|$ tends to zero. ∎

Step 3: Upper middle order quantiles

At (27) we have defined $h_n = n^\beta$ with $\beta \in (1/2,1)$ to be chosen. Let us introduce

$$I_{M,n} = \left(F(M), 1-\frac{h_n}{n}\right), \quad M > m. \tag{32}$$

Since $F(M) > F(m) = \overline{u}$ and (39) in Section 6.1.2 holds we have by $(C2)$

$$\begin{aligned} U_{M,n} &= \int_{F(M)}^{1-h_n/n} \left(c\left(\mathbb{F}_n^{-1}(u), \mathbb{G}_n^{-1}(u)\right) - c\left(F^{-1}(u), G^{-1}(u)\right)\right) du \\ &= \int_{F(M)}^{1-h_n/n} \rho\left(|\tau(u)+\varepsilon_n(u)|\right) - \rho\left(\tau(u)\right) du \end{aligned}$$

where $\varepsilon_n(u)$ is as in (29). In order to control the last integral, we expand ρ and make use of a distribution free Brownian approximation of the joint quantile processes.

Lemma 25 *Assume $(C2)$, (FG) and (CFG). For any $\varepsilon > 0$ and $\lambda > 0$ we can find $M > m$ such that, for all n large enough,*

$$\mathbb{P}\left(\sqrt{n}\,|U_{M,n}| > \lambda\right) < \varepsilon.$$

Proof

1. Under $(C2)$ we have $l'(x) = \varepsilon_1(x)l(x)/x$ where $\varepsilon_1(x) \to \gamma$ as $x \to +\infty$ thus ε_1 is bounded on $(M,+\infty)$. Moreover, (CFG) ensures that

$$l \circ \psi_Y^{-1}(x) \leqslant l \circ \psi_X^{-1}(x) < x$$

whereas (15) and (8) entails that $\psi_X(x) > 2l(x) \geqslant 2\log x$ thus

$$F^{-1}(u) = \psi_X^{-1}\left(\log\left(\frac{1}{1-u}\right)\right) < \frac{1}{\sqrt{1-u}}$$

for all $u \in I_{M,n}$ and $x \in F(I_{M,n})$. Under $(FG4)$ we have $\tau(u) = F^{-1}(u) - G^{-1}(u) \geqslant \tau_0$ for $u \in I_{M,n}$. Hence by choosing $M > m$ and $K > 0$ sufficiently large, (30) and $(FG3)$ imply that it almost surely eventually holds

$$\begin{aligned}
&\sup_{u\in I_{M,n}} \varepsilon_1 \circ \tau(u) \frac{l\circ\tau(u)}{\tau(u)} |\varepsilon_n(u)| \\
&\leqslant K \sup_{u\in I_{M,n}} l\circ F^{-1}(u)(H_X(u)F^{-1}(u) + H_Y(u)G^{-1}(u))\sqrt{\frac{\log\log n}{n(1-u)}} \\
&\leqslant K\sqrt{\log\log n} \sup_{u\in I_{M,n}} \frac{l\circ\psi_X^{-1}(\log(1/(1-u))}{\sqrt{n(1-u)}} F^{-1}(u) \\
&\leqslant K\sqrt{n\log\log n} \sup_{u\in I_{M,n}} \frac{\log(1/(1-u)}{n(1-u)} \\
&\leqslant K\frac{\log n}{h_n}\sqrt{n\log\log n}
\end{aligned}$$

which vanishes since $\beta > 1/2$ in (27). We have shown that

$$\lim_{n\to+\infty} \sup_{u\in I_{M,n}} |\varepsilon_n(u)|\, l'\circ\tau(u) = 0 \quad a.s. \tag{33}$$

2. By (33), the second part of Proposition 31 can be applied for all large n. It says that

$$\rho\left(|\tau(u)+\varepsilon_n(u)|\right) - \rho\left(\tau(u)\right) = k_0(\tau(u),\varepsilon_n(u))\rho'\circ\tau(u)\varepsilon_n(u)$$

where, by (53),

$$\lim_{\delta_0\to 0} \sup_{\tau(u)>\tau_0} \sup_{|\varepsilon_n(u)|l'\circ\tau(u)\leqslant\delta_0} |k_0(\tau(u),\varepsilon_n(u)) - 1| = 0$$

which can be reformulated through (33) into $k_1(u) = k_0(\tau(u),\varepsilon_n(u))$ and

$$\lim_{n\to+\infty} \sup_{u\in I_{M,n}} |k_1(u)-1| = 0 \quad a.s. \tag{34}$$

Thus, given any $\vartheta\in(0,1)$ the random function $k_1(u)$ is such that $k_1(u)\in(1-\vartheta,1+\vartheta)$ for all $u\in I_{M,n}$ and

$$\sqrt{n}U_{M,n} = \int_{F(M)}^{1-h_n/n} k_1(u)\rho'\circ\tau(u)\left(\beta_n^X(u)+\beta_n^Y(u)\right)du.$$

From now on we work on the probability space of Theorem 28. This allows us to write

$$\sqrt{n}U_{M,n} = \int_{F(M)}^{1-h_n/n} k_1(u)\rho'\circ\tau(u)\left(\frac{B_n^X(u)+Z_n^X(u)}{h_X(u)} + \frac{B_n^Y(u)+Z_n^Y(u)}{h_Y(u)}\right)du$$

where $\left(U_{M,n}, B_n^X, Z_n^X, B_n^Y, Z_n^Y, k_1\right)$ are built together on Ω^* in such a way that for some small $\xi>0$ independent of the law Π,

$$\lim_{n\to+\infty} n^\xi \sup_{u\in I_{M,n}} \left|Z_n^X(u)\right| = \lim_{n\to+\infty} n^\xi \sup_{u\in I_{M,n}} \left|Z_n^Y(u)\right| = 0 \quad a.s. \tag{35}$$

and B_n^X, B_n^Y are Brownian briges define at (44). Therefore k_1 obeys (34). Let set $\sqrt{n}\, U_{M,n} = N_{M,n} + R_{M,n} + S_{M,n}$ with

$$N_{M,n} = \int_{F(M)}^{1-h_n/n} \rho' \circ \tau(u) \left(\frac{B_n^X(u)}{h_X(u)} + \frac{B_n^Y(u)}{h_Y(u)} \right) du$$

$$R_{M,n} = \int_{F(M)}^{1-h_n/n} k_1(u) \rho' \circ \tau(u) \left(\frac{Z_n^X(u)}{h_X(u)} + \frac{Z_n^Y(u)}{h_Y(u)} \right) du$$

$$S_{M,n} = \int_{F(M)}^{1-h_n/n} (k_1(u)-1) \rho' \circ \tau(u) \left(\frac{B_n^X(u)}{h_X(u)} + \frac{B_n^Y(u)}{h_Y(u)} \right) du$$

3. We first deal with $R_{M,n}$. Since $\rho'(x)$ is increasing by Proposition 32, $(C2)$ implies $l'(x) < Kl(x)/x$ with $K > \gamma$ and (CFG) entails $l \circ \psi_Y^{-1}(x) \leqslant l \circ \psi_X^{-1}(x) \leqslant x/2 - \theta \log x$ by (14) we readily have

$$\begin{aligned}
&\left| \int_{F(M)}^{1-h_n/n} \frac{\rho' \circ \tau(u)}{h_X(u)} Z_n^X(u) du \right| \\
&\leqslant \frac{K}{n^\xi} \int_{F(M)}^{1-h_n/n} \frac{l \circ \psi_X^{-1}(\log(1/(1-u))}{F^{-1}(u) h_X(u)} \exp\left(l \circ \psi_X^{-1}(\log(1/(1-u))) \right) du \\
&\leqslant \frac{K}{n^\xi} \int_{F(M)}^{1-h_n/n} \frac{\log(1/(1-u))}{(\log(1/(1-u)))^\theta} \frac{H_X(u)}{(1-u)^{3/2}} du
\end{aligned}$$

which is, by using $(FG3)$ and $\theta > 1$ then choosing $\beta \in (1-\xi, 1)$, less than

$$\frac{K}{n^\xi} \int_{F(M)}^{1-h_n/n} \frac{1}{(1-u)^{3/2}} du < K n^{-\xi/2}.$$

The same bound holds for h_Y since $F^{-1} > G^{-1}$ and

$$\left| \int_{F(M)}^{1-h_n/n} \frac{\rho' \circ \tau(u)}{h_Y(u)} Z_n^Y(u) du \right| \leqslant \frac{K}{n^\xi} \int_{F(M)}^{1-h_n/n} \frac{G^{-1}(u)}{F^{-1}(u)} \frac{\log(1/(1-u))}{(\log(1/(1-u)))^\theta} \frac{H_Y(u)}{(1-u)^{3/2}} du.$$

By (34), (35) and the above bounds we have almost surely for n large enough $|R_{M,n}| \leqslant 2Kn^{-\xi/2} \longrightarrow 0$.

4. As $N_{M,n}$ is the sum of two linear functionals of Brownian bridges it is a mean zero Gaussian random variable with variance

$$\sigma^2(M,n) = \int_{F(M)}^{1-h_n/n} \int_{F(M)}^{1-h_n/n} \rho' \circ \tau(u) \rho' \circ \tau(v) \Xi(u,v) du dv$$

where

$$\begin{aligned}
\Xi(u,v) &= cov\left(\frac{B_n^X(u)}{h_X(u)} + \frac{B_n^Y(u)}{h_Y(u)}, \frac{B_n^X(v)}{h_X(v)} + \frac{B_n^Y(v)}{h_Y(v)} \right) \\
&= \frac{\min(u,v) - uv}{h_X(u) h_X(v)} + \frac{\Pi(v,u) - uv}{h_X(v) h_Y(u)} + \frac{\Pi(u,v) - uv}{h_X(u) h_Y(v)} + \frac{\min(u,v) - uv}{h_Y(v) h_Y(u)}.
\end{aligned}$$

Therefore by Lemma 19 taken in $\overline{u} = F(M)$ we see that $\sigma^2(M,n) \to \sigma^2(M)$ as $n \to \infty$ and $\sigma^2(M) \to 0$ as $M \to +\infty$. On an other hand the random variable $\int_{F(M)}^{1-h_n/n} \rho' \circ \tau(u) \left(\frac{B_n^X(u)}{h_X(u)} + \frac{B_n^Y(u)}{h_Y(u)} \right) du$ is $a.s.$ finite. Thus $\int_{F(M)}^{1-h_n/n} \rho' \circ \tau(u) \left| \frac{B_n^X(u)}{h_X(u)} + \frac{B_n^Y(u)}{h_Y(u)} \right| du$ is $a.s.$ finite. Hence

$$|S_{M,n}| \leqslant \sup_{u \in I_{M,n}} |k_1(u) - 1| \int_{F(M)}^{1-h_n/n} \rho' \circ \tau(u) \left| \frac{B_n^X(u)}{h_X(u)} + \frac{B_n^Y(u)}{h_Y(u)} \right| du,$$

which $a.s.$ tend to 0 when $n \to \infty$.
As a conclusion, for any $\varepsilon > 0$ and $\lambda > 0$ we can find $M = M(\varepsilon, \lambda) > m$ such that

$$\begin{aligned} \mathbb{P}\left(\sqrt{n}\,|U_{M,n}| > \lambda\right) &\leqslant \mathbb{P}\left(|N_{M,n}| > \frac{\lambda}{3}\right) + \mathbb{P}\left(|R_{M,n}| > \frac{\lambda}{3}\right) + \mathbb{P}\left(|S_{M,n}| > \frac{\lambda}{3}\right) \\ &\leqslant \frac{\sigma^2(M,n)}{(\lambda/3)^2} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} < \varepsilon, \end{aligned}$$

for all $n > n(\varepsilon, \lambda, M)$.

■

Step 4: Centered middle order quantiles

Define

$$I_M = (F(-M), F(M)), \quad M > m,$$

and consider the centered random integral

$$\mathbb{M}_{M,n} = \int_{I_M} \left(c\left(\mathbb{F}_n^{-1}(u), \mathbb{G}_n^{-1}(u)\right) - c\left(F^{-1}(u), G^{-1}(u)\right) \right) du.$$

In order to conclude the proof of Theorem 14 it remains to exploit the Brownian approximation of the joint quantile processes β_n^X and β_n^Y defined at (28) to accurately approximate $\sqrt{n}\mathbb{M}_{M,n}$. Recalling (19) let us write

$$\nabla_x(u) = \frac{\partial}{\partial x} c(F^{-1}(u), G^{-1}(u)), \quad \nabla_y(u) = \frac{\partial}{\partial y} c(F^{-1}(u), G^{-1}(u))$$

and

$$\sqrt{n}\mathbb{N}_{M,n} = \int_{I_M} \left(\nabla_x(u) \beta_n^X(u) + \nabla_y(u) \beta_n^Y(u) \right) du.$$

Lemma 26 *Assume (C), (FG) and (CFG). Then for any $\delta > 0$, any $\varepsilon > 0$ and any $M > m' > m$ there exists $n(\varepsilon, \delta, M)$ such that for all $n > n(\varepsilon, \delta, M)$,*

$$\mathbb{P}\left(\left| \sqrt{n}\mathbb{M}_{M,n} - \sqrt{n}\mathbb{N}_{M,n} \right| > \varepsilon \right) \leqslant \delta.$$

Proof

1. Under $(FG1)$, h_X and h_Y are away from 0 on I_M and we write

$$\eta_M = \min\left(\inf_{u\in I_M} h_X(u), \inf_{u\in I_M} h_Y(u)\right) > 0.$$

   We keep working on the probability space of Theorem 28. In particular, since $I_M \subset \mathcal{I}_n$ we can aplly again Theorem 28 and get the analogue of (35)

$$\mathbb{P}\left(\sup_{u\in I_M}\left|\frac{Z_n^X(u)}{h_X(u)}\right| > \frac{1}{n^\xi}\right) = o(1), \mathbb{P}\left(\sup_{u\in I_M}\left|\frac{Z_n^Y(u)}{h_Y(u)}\right| > \frac{1}{n^\xi}\right) = o(1). \quad (36)$$

   Introduce the event

$$A_n(M,C) = \left\{\sup_{u\in I_M}\left|\mathbb{F}_n^{-1}(u) - F^{-1}(u)\right| + \left|\mathbb{G}_n^{-1}(u) - G^{-1}(u)\right| \leqslant \frac{4C}{\sqrt{n}}\right\}.$$

   By (36), for any $\delta > 0$ one can find $C_\delta > 0$ so large that, for all n large enough,

$$\begin{aligned}
&\mathbb{P}\left(A_n(M,C_\delta)^c\right) \\
&= \mathbb{P}\left(\sup_{u\in I_M}\sqrt{n}\left|\mathbb{F}_n^{-1}(u) - F^{-1}(u)\right| + \sqrt{n}\left|\mathbb{G}_n^{-1}(u) - G^{-1}(u)\right| > 4C_\delta\right) \\
&\leqslant \mathbb{P}\left(\sup_{u\in I_M}\left|\frac{B_n^X(u)}{h_X(u)}\right| > C_\delta\right) + \mathbb{P}\left(\sup_{u\in I_M}\left|\frac{B_n^Y(u)}{h_Y(u)}\right| > C_\delta\right) + o(1) \\
&\leqslant 2\mathbb{P}\left(\sup_{u\in I_M}|B(u)| > \eta_M C_\delta\right) + \frac{\delta}{2} \\
&\leqslant \delta
\end{aligned}$$

   where B denotes a standard Brownian bridge.

2. Since $F \neq G$ and F, G are continuous, for any $\tau_1 \in (0,\tau_0)$ there exists an open interval $I(\tau_1) \subset I_M$ such that $|\tau(u)| > \tau_1$ for $u \in I(\tau_1)$, provided that $m > 0$ is chosen large enough. By taking $M > m$, by $(FG4)$ we further have $\tau_M = \sup_{u\in I_M}|\tau(u)| \geqslant \tau_0 > \tau_1$. Thus

$$\begin{aligned}
D_M^+(\tau_1) &= \{u : \tau_1 < |\tau(u)| \leqslant \tau_M\} \cap I_M \\
D_M^-(\tau_1) &= \{u : |\tau(u)| \leqslant \tau_1\} \cap I_M
\end{aligned}$$

   are such that $I(\tau_1) \subset D_M^+(\tau_1) \neq \emptyset$ and $D_M^-(\tau_1) \subset I_M$ is possibly empty, and $D_M^+(\tau_1) \cup D_M^-(\tau_1) = I_M$. By $(C3)$, for any (x,y), $(x',y') \in D_m(\tau)$,

$$|c(x',y') - c(x,y)| \leqslant d(m,\tau)\left(|x'-x| + |y'-y|\right)$$

   with $d(m,\tau) \to 0$ as $\tau \to 0$ and m is fixed. Observe that $u \in D_M^-(\tau_1) = D_m^-(\tau_1)$ if, and only if, $\left(F^{-1}(u), G^{-1}(u)\right) \in D_m(\tau_1)$. Let $\tau_1' \in (\tau_1,\tau_0)$ and $m' \in (m,M)$.

Now, given M and C_δ, if $A_n(M, C_\delta)$ is true for a large enough n then $\left(\mathbb{F}_n^{-1}(u), \mathbb{G}_n^{-1}(u)\right) \in D_{m'}(\tau_1')$ whenever $\left(F^{-1}(u), G^{-1}(u)\right) \in D_m(\tau_1) \subset D_{m'}(\tau_1')$ and $u \in I_M$. Thus, under the event $A_n(M, C_\delta)$ it holds

$$\begin{aligned}
\sqrt{n}\mathbb{M}_{M,n}^{-}(\tau_1) &:= \sqrt{n}\int_{u \in D_M^-(\tau_1)} \left| c\left(\mathbb{F}_n^{-1}(u), \mathbb{G}_n^{-1}(u)\right) - c\left(F^{-1}(u), G^{-1}(u)\right)\right| du \\
&\leqslant \sqrt{n}\int_{u \in D_M^-(\tau_1)} d(m', \tau_1') \left(\left|\mathbb{F}_n^{-1}(u) - F^{-1}(u)\right| + \left|\mathbb{G}_n^{-1}(u) - G^{-1}(u)\right|\right) du \\
&\leqslant 4C_\delta d(m', \tau_1').
\end{aligned}$$

3. The main term is

$$\sqrt{n}\mathbb{M}_{M,n}^{+}(\tau_1) := \sqrt{n}\int_{u \in D_M^+(\tau_1)} \left(c\left(\mathbb{F}_n^{-1}(u), \mathbb{G}_n^{-1}(u)\right) - c\left(F^{-1}(u), G^{-1}(u)\right)\right) du.$$

Under the event $A_n(M, C_\delta)$ the Taylor expansion of $c(F^{-1}(u), G^{-1}(u))$ is justified on $D_M^+(\tau_1)$, that is away from the diagonal. As a matter of fact, under $(C1)$ we have, for x, y in $(-M, M)$ such that $|x - y| \geqslant \tau$,

$$\left| c\left(x + \varepsilon_x, y + \varepsilon_y\right) - c\left(x, y\right) - \nabla_x(x, y)\varepsilon_x - \nabla_y(x, y)\varepsilon_y\right| \leqslant \lambda(M, \tau)\Theta\left(|\varepsilon_x| + |\varepsilon_y|\right),$$

where $\Theta\left(s\right)/s \to 0$ as $s \to 0$ for M and τ_1 fixed. Then the expansion of $c(F^{-1}(u), G^{-1}(u))$ on $u \in D_M^+(\tau_1)$ can be writen as

$$c\left(\mathbb{F}_n^{-1}(u), \mathbb{G}_n^{-1}(u)\right) - c(F^{-1}(u), G^{-1}(u) = \left(\nabla_x(u)\beta_n^X(u) + \nabla_y(u)\beta_n^Y(u)\right) + \mathcal{R}_n(u).$$

We have

$$\begin{aligned}
&\left|\sqrt{n}\mathbb{M}_{M,n}^{+}(\tau_1) - \int_{u \in D_M^+(\tau_1)} \left(\nabla_x(u)\beta_n^X(u) + \nabla_y(u)\beta_n^Y(u)\right) du\right| \\
&\leqslant \sqrt{n}\left|\int_{u \in D_M^+(\tau_1)} \mathcal{R}_n(u) du\right| \\
&\leqslant \lambda(M, \tau_1)\sqrt{n}\Theta\left(\frac{1}{\sqrt{n}} \sup_{u \in I_M} \left|\beta_n^X(u)\right| + \left|\beta_n^Y(u)\right|\right)
\end{aligned}$$

As $|\mathbb{M}_{M,n} - \mathbb{M}_{M,n}^{+}(\tau_1)| \leqslant \mathbb{M}_{M,n}^{-}(\tau_1)$, whenever $A_n(M, C_\delta)$ is true we have

$$\begin{aligned}
&\left|\sqrt{n}\mathbb{M}_{M,n} - \int_{u \in D_M^+(\tau_1)} \left(\nabla_x(u)\beta_n^X(u) + \nabla_y(u)\beta_n^Y(u)\right) du\right| \\
&\leqslant \sqrt{n}\mathbb{M}_{M,n}^{-}(\tau_1) + \lambda(M, \tau_1)\sqrt{n}\Theta\left(\frac{4C_\delta}{\sqrt{n}}\right)
\end{aligned}$$

where $\sqrt{n}\Theta\left(4C_\delta\sqrt{n}\right) \to 0$ as $n \to +\infty$. We also have $D_M^-(\tau_1) = I_M \setminus D_M^+(\tau_1) \subset I_M$ and ∇_x, ∇_y are bounded on I_M thus

$$\left|\int_{u\in D_M^-(\tau_1)} \left(\nabla_x(u)\beta_n^X(u) + \nabla_y(u)\beta_n^Y(u)\right) du\right|$$
$$\leqslant 2m\frac{4C_\delta}{\sqrt{n}} \sup_{u\in I_M} |\nabla_x(u)| + |\nabla_y(u)| .$$

Hence under $A_n(M, C_\delta)$ $|\sqrt{n}\mathbb{M}_{M,n} - \sqrt{n}\mathbb{N}_{M,n}|$ is bounded by

$$4C_\delta d(m', \tau_1') + \lambda(M, \tau_1)\sqrt{n}\Theta\left(\frac{4C_\delta}{\sqrt{n}}\right) + 2m\frac{4C_\delta}{\sqrt{n}} \sup_{u\in I_M} |\nabla_x(u)| + |\nabla_y(u)|$$

Therefore, for any $\delta > 0$, any $\varepsilon > 0$ and any triplet $M > m' > m$ we can choose τ_1 and $\tau_1' > \tau_1$ so small that $4C_\delta d(m', \tau_1') \leqslant \varepsilon/2$. Then there exists $n(\varepsilon, \delta, M)$ such that for all $n > n(\varepsilon, \delta, M)$,

$$\mathbb{P}\left(\left|\sqrt{n}\mathbb{M}_{M,n} - \sqrt{n}\mathbb{N}_{M,n}\right| > \varepsilon\right) \leqslant \mathbb{P}\left(A_n(M, C_\delta)^c\right) \leqslant \delta.$$

■

Step 5: Conclusion

Now recall that $\sqrt{n}\left(W_c(\mathbb{F}_n, \mathbb{G}_n) - W_c(F, G)\right) = \sqrt{n}D_n + \sqrt{n}S_n + \sqrt{n}T_n + \sqrt{n}U_{M,n} + \sqrt{n}\mathbb{M}_{M,n}$. By Steps 1, 2 $\sqrt{n}D_n + \sqrt{n}S_n + \sqrt{n}T_n$ converges to zero in probability. Hence, we only need to prove the weak convergence of $\sqrt{n}U_{M,n} + \sqrt{n}\mathbb{M}_{M,n}$. Let $\mathbb{X}_\infty$ be a centered Gaussian random variable with variance $\sigma^2(\Pi, c)$. For any B-bounded r-Lipschitz function Φ, we have

$$\mathbb{E}\left[\left|\Phi\left(\sqrt{n}(U_{M,n} + \mathbb{M}_{M,n})\right) - \Phi\left(\mathbb{X}_\infty\right)\right|\right]$$
$$\leqslant \mathbb{E}\left[\left|\Phi\left(\sqrt{n}(U_{M,n} + \mathbb{M}_{M,n})\right) - \Phi\left(\sqrt{n}\mathbb{M}_{M,n}\right)\right|\right] + \mathbb{E}\left[\left|\Phi\left(\sqrt{n}\mathbb{M}_{M,n}\right) - \Phi\left(\mathbb{X}_\infty\right)\right|\right]$$

Dealing with the first right hand term we have

$$\begin{aligned}
&\mathbb{E}\left[\left|\Phi\left(\sqrt{n}(U_{M,n} + \mathbb{M}_n)\right) - \Phi\left(\sqrt{n}\mathbb{M}_{M,n}\right)\right|\right] \\
&\quad = \mathbb{E}\left[\left|\Phi\left(\sqrt{n}(U_{M,n} + \mathbb{M}_{M,n})\right) - \Phi\left(\sqrt{n}\mathbb{M}_{M,n}\right)\right| \mathbb{1}_{|\sqrt{n}U_{M,n}|>\lambda}\right] \\
&\quad + \mathbb{E}\left[\left|\Phi\left(\sqrt{n}(U_{M,n} + \mathbb{M}_{M,n})\right) - \Phi\left(\sqrt{n}\mathbb{M}_{M,n}\right)\right| \mathbb{1}_{|\sqrt{n}U_{M,n}|\leqslant\lambda}\right] \\
&\quad \leqslant 2B\mathbb{P}\left(|\sqrt{n}U_{M,n}| > \lambda\right) r\lambda
\end{aligned}$$

By lemma 25 we can make $2B\mathbb{P}\left(|\sqrt{n}U_{M,n}| > \lambda\right) r\lambda$ as small as we want by choosing λ small enough and M large enough.
We now consider the second right hand term

$$\mathbb{E}\left[\left|\Phi\left(\sqrt{n}\mathbb{M}_{M,n}\right) - \Phi\left(\mathbb{X}_\infty\right)\right|\right] \leqslant$$
$$\mathbb{E}\left[\left|\Phi\left(\sqrt{n}\mathbb{M}_{M,n}\right) - \Phi\left(\sqrt{n}\mathbb{N}_{M,n}\right)\right|\right] + \mathbb{E}\left[\left|\Phi\left(\sqrt{n}\mathbb{N}_{M,n}n\right) - \Phi\left(\mathbb{X}_\infty\right)\right|\right]$$

By lemma 26 the term $\mathbb{E}\left[\left|\Phi\left(\sqrt{n}\mathbb{M}_{M,n}\right)-\Phi\left(\sqrt{n}\mathbb{N}_{M,n}\right)\right|\right]$ can be made as small as desired. As $\sqrt{n}\mathbb{N}_{M,n}$ is a Gaussian random variable with variance

$$\sigma^2(M,\Pi,c)=\int_{F(-M)}^{F(M)}\int_{F(-M)}^{F(M)}\nabla(u)\Sigma(u,v)\nabla(v)dudv \tag{37}$$

that converges to $\sigma^2(\Pi,c)$, the term $\mathbb{E}\left[\left|\Phi\left(\sqrt{n}\mathbb{N}_{M,n}\right)-\Phi\left(\mathbb{X}_\infty\right)\right|\right]$ is small enough for large enough M. This achieves the proof of Theorem 14.

6 Appendix

6.1 Proof of auxiliary results

6.1.1 Proof of Lemma 23

Remind that $\Delta_n=[\overline{u},1-k_n/n]$ where $k_n/\log\log n\to+\infty$ and $k_n/n\to 0$ comes from (24) and (25). Let us study $(\beta_n,h)=(\beta_n^X,h_X)$ in Lemma 23. Under $(FG1)$ we have $f>0$ on $\mathbb{R}$ thus the random variables $U_i=F(X_i)$ are independent, uniformely distributed on $[0,1]$ and such that $X_{(i)}=F^{-1}(U_{(i)})$. Let $\mathbb{F}_{U,n}$ and $\mathbb{F}_{U,n}^{-1}$ denote the empirical cdf and quantile functions associated to $U_1,...,U_n$ so that $\mathbb{F}_n=\mathbb{F}_{U,n}\circ F$ and $\mathbb{F}_n^{-1}=F^{-1}\circ\mathbb{F}_{U,n}^{-1}$. Write $q_n(u)=\mathbb{F}_{U,n}^{-1}(u)-u$. By [6] we have

$$\limsup_{n\to\infty}\sup_{u\in\Delta_n}\frac{\sqrt{n}q_n(u)}{\sqrt{(1-u)\log\log n}}\leqslant 4 \quad a.s. \tag{38}$$

Since $(FG1)$ ensures that h_X is $\mathcal{C}_1$ on Δ_n the following expansion almost surely asymptotically holds,

$$\begin{aligned}&\sup_{u\in\Delta_n}\left|\left(F^{-1}(u+q_n(u))-F^{-1}(u)\right)h_X(u)-q_n(u)\right|\\ &=\sup_{u\in\Delta_n}\left|\left(\frac{q_n(u)}{h_X(u)}+\frac{q_n^2(u)}{2}\left(\frac{1}{h_X(u)}\right)'_{u=u^*}\right)h_X(u)-q_n(u)\right|\\ &\leqslant A_nB_n\end{aligned}$$

where $|u-u^*|\leqslant|q_n(u)|$ and, by (38),

$$A_n=\sup_{u\in\Delta_n}\frac{q_n^2(u)}{2\,(1-u)}\leqslant K\frac{\log\log n}{n}$$

whereas, by $(FG2)$,

$$\begin{aligned}
B_n &= \sup_{u\in\Delta_n} (1-u)\, h_X(u) \left| \left(\frac{1}{h_X(u)}\right)'_{u=u^*} \right| \\
&\leqslant \sup_{u\in\Delta_n} (1-u^*) h_X(u^*) \left| \left(\frac{1}{h_X(u)}\right)'_{u=u^*} \right| \sup_{u\in\Delta_n} \frac{1-u}{1-u^*} \frac{h_X(u)}{h_X(u^*)} \\
&\leqslant K \sup_{u\in\Delta_n} \frac{1-u}{1-u^*} \sup_{u\in\Delta_n} \frac{h_X(u)}{h_X(u^*)}.
\end{aligned}$$

Now, (38) shows that the random sequence

$$\sup_{u\in\Delta_n} \left| \frac{1-u^*}{1-u} - 1 \right| \leqslant \sup_{u\in\Delta_n} \frac{1}{\sqrt{1-u}} \sup_{u\in\Delta_n} \left| \frac{q_n(u)}{\sqrt{1-u}} \right| \leqslant 5\sqrt{\frac{n}{k_n}} \sqrt{\frac{\log\log n}{n}}$$

almost surely tends to 0. Moreover $(FG2)$ implies that

$$\left| (\log h_X(u))' \right| \leqslant K \left(\log \frac{1}{1-u} \right)'$$

so that $|\log h_X(u_2) - \log h_X(u_1)| \leqslant K(\log(1-u_1) - \log(1-u_2))$ for any $u_1 < u_2$ in Δ_n. Therefore, the random sequence

$$\sup_{u\in\Delta_n} \frac{h_X(u)}{h_X(u^*)} \leqslant \sup_{u\in\Delta_n} \max \left((\frac{1-u^*}{1-u}), (\frac{1-u}{1-u^*}) \right)^K$$

almost surely tends to 1. We have shown that it almost surely ultimately holds

$$\begin{aligned}
\sup_{u\in\Delta_n} \left| \frac{\beta_n^X(u) h_X(u) - \sqrt{n} q_n(u)}{\sqrt{(1-u)\log\log n}} \right| &\leqslant A_n B_n \sqrt{\frac{n}{\log\log n}} \sup_{u\in\Delta_n} \frac{1}{\sqrt{1-u}} \\
&\leqslant 10K \sqrt{\frac{\log\log n}{k_n}}
\end{aligned}$$

which proves Lemma 23, by (38) again.

6.1.2 Proof of Lemma 24.

In view of (25) and (27) we eventually have $I_n \subset \Delta_n$. Hence Lemma 23 and $(FG3)$ imply that, almost surely, for all n large

$$\begin{aligned}
\sup_{u\in I_n} \frac{\left| \mathbb{F}_n^{-1}(u) - F^{-1}(u) \right|}{F^{-1}(u)} &\leqslant 2K_0 \sup_{u\in I_n} \frac{\sqrt{1-u}}{F^{-1}(u) h_X(u)} \sqrt{\frac{\log\log n}{n}} \\
&= 2K_0 \sup_{u\in I_n} H_X(u) \sqrt{\frac{\log\log n}{n(1-u)}} \leqslant K \sqrt{\frac{\log\log n}{k_n}}.
\end{aligned}$$

The same bound holds for $\left|\mathbb{G}_n^{-1}(u) - G^{-1}(u)\right|/G^{-1}(u)$. By (25) we then get

$$\lim_{n\to+\infty} \sup_{u\in I_n} \frac{\left|\varepsilon_n^Y(u)\right|}{F^{-1}(u)} \leqslant \lim_{n\to+\infty} \sup_{u\in I_n} \frac{\left|\varepsilon_n^Y(u)\right|}{G^{-1}(u)} = \lim_{n\to+\infty} \sup_{u\in I_n} \frac{\left|\varepsilon_n^X(u)\right|}{F^{-1}(u)} = 0 \quad a.s.$$

so that $\sup_{u\in I_n} |\varepsilon_n(u)|/F^{-1}(u)$ almost surely vanishes. Under $(FG1)$ the law of large numbers for $\mathbb{F}_n$ and $\mathbb{G}_n$ readily implies

$$\lim_{n\to+\infty} \mathbb{F}_n^{-1}\left(1-\frac{h_n}{n}\right) = \lim_{n\to+\infty} \mathbb{G}_n^{-1}\left(1-\frac{h_n}{n}\right) = +\infty \quad a.s.$$

Therefore for any $q_0 > 0$, all n large enough and all $u \in I_n$, it holds

$$\min\left(\mathbb{F}_n^{-1}(u), F^{-1}(u), \mathbb{G}_n^{-1}(u), G^{-1}(u)\right) > m, \quad |\varepsilon_n(u)| < q_0 F^{-1}(u) \tag{39}$$

which implies, by $(C2)$ and for $\tau(u) = F^{-1}(u) - G^{-1}(u)$,

$$c\left(\mathbb{F}_n^{-1}(u), \mathbb{G}_n^{-1}(u)\right) - c\left(F^{-1}(u), G^{-1}(u)\right) = \rho\left(|\tau(u)+\varepsilon_n(u)|\right) - \rho\left(\tau(u)\right).$$

Case 1. Assume that $\gamma = 0$ in $(C2)$. By Proposition 32 ρ' is increasing and

$$\left|\rho\left(|\tau(u)+\varepsilon_n(u)|\right) - \rho\left(\tau(u)\right)\right| \leqslant \rho'\left(\tau(u) + |\varepsilon_n(u)|\right)|\varepsilon_n(u)|.$$

Observe that if

$$\liminf_{u\to 1} \frac{G^{-1}(u)}{F^{-1}(u)} = q_1 > 0$$

then the result follows with $K_2 = 1$ since by taking $0 < q_0 < q_1 \leqslant 1$ in (39) we ultimately have, with probability one,

$$\rho'\left(\tau(u) + |\varepsilon_n(u)|\right) = \rho'\left(F^{-1}(u)\left(1 - \frac{G^{-1}(u)}{F^{-1}(u)} + \frac{|\varepsilon_n(u)|}{F^{-1}(u)}\right)\right) \leqslant \rho'\left(F^{-1}(u)\right).$$

If $q_1 = 0$, let us control $\rho'\left(\tau(u) + |\varepsilon_n(u)|\right) \leqslant \rho'\left(F^{-1}(u)\left(1 + |\varepsilon_n(u)|/F^{-1}(u)\right)\right)$. Remind (48) and the fact that l is increasing whereas l' is decreasing, by (7) and (8). For $y > x$, $x \to +\infty$, $y \sim x$ we have $l(x) \leqslant l(y) \leqslant l(2x) \sim l(x)$ and

$$\frac{\rho'(y)}{\rho'(x)} = \frac{l'(y)}{l'(x)}\frac{\rho(y)}{\rho(x)} \leqslant \frac{\rho(y)}{\rho(x)} = \exp\left(l(y) - l(x)\right) \leqslant \exp\left(l'(x)(y-x)\right).$$

Therefore, by (7), (12) and $(FG3)$, taking $\theta_1' \in (\theta_1, \theta - 1)$ yields

$$\begin{aligned} 1 &\leqslant \frac{1}{\rho' \circ F^{-1}(u)}\rho'\left(F^{-1}(u)\left(1 + \frac{|\varepsilon_n(u)|}{F^{-1}(u)}\right)\right) \\ &\leqslant \exp\left(l' \circ F^{-1}(u)\,|\varepsilon_n(u)|\right) \\ &= \exp\left(\varepsilon_1 \circ F^{-1}(u) l \circ F^{-1}(u) \frac{|\varepsilon_n(u)|}{F^{-1}(u)}\right) \\ &\leqslant \exp\left(\left(l \circ F^{-1}\left(1 - \frac{k_n}{n}\right)\right)^{\theta_1'} K\sqrt{\frac{\log\log n}{k_n}}\right) \end{aligned}$$

provided n is large enough and $u \in I_n$. Moreover (14) implies

$$l \circ F^{-1}\left(1 - \frac{k_n}{n}\right) = l \circ \psi_X^{-1}\left(\log\left(\frac{n}{k_n}\right)\right) \leqslant l \circ \psi_X^{-1}(\log n) \leqslant \log n. \quad (40)$$

By choosing θ' in (25) such that $\theta > \theta' > 1 + \theta_1' \geqslant \max(1, 2\theta_1')$ we get

$$\lim_{n\to+\infty} \sup_{u\in I_n} \frac{\rho'\left(\tau(u) + |\varepsilon_n(u)|\right)}{\rho' \circ F^{-1}(u)} \leqslant 1 \quad a.s.$$

which yields the result with $K_2 = 1$ again.

Case 2. Assume that $\gamma > 1$ in $(C2)$. Since l' is now increasing the above argument fails to guaranty that $\rho'(x) \sim \rho'(y)$ as $y \sim x$ are sufficiently close. Instead we check the sufficient condition in Proposition 31. The function $l(x)/x$ is increasing as it is regularly varying with index $\gamma - 1 > 0$. Recall also that (CFG) yields (40) and that $H = H_X + H_Y$ is bounded under $(FG3)$. As a consequence of $I_n \subset \Delta_n$ and Lemma 23 we almost surely have, for all n large,

$$\begin{aligned}
\sup_{u\in I_n} \frac{l\circ\tau(u)}{\tau(u)} |\varepsilon_n(u)| &\leqslant 2K_0 \sup_{u\in I_n} l \circ F^{-1}(u) H(u) \sqrt{\frac{\log\log n}{n(1-u)}} \\
&\leqslant 2K_0 l \circ F^{-1}\left(1 - \frac{k_n}{n}\right) \sqrt{\frac{\log\log n}{k_n}} \sup_{u\in I_n} H(u) \\
&\leqslant K \frac{\log n}{\sqrt{k_n}} \sqrt{\log\log n} \sup_{u\in I_n} H(u). \quad (41)
\end{aligned}$$

Since $\theta > 2$ in (CFG) choosing $\theta' \in (2, \theta)$ in (25) makes the upper bound in (41) vanish. Therefore, under (CFG) the requirements of Proposition 31 are almost surely ultimately fulfiled with

$$x_0 = \tau_0, \quad x = \tau(u), \quad |\varepsilon| = |\varepsilon_n(u)| \leqslant \frac{\delta_0}{l'(x)} = \frac{\delta_0}{l' \circ \tau(u)} = \frac{\delta_0 \tau(u)}{\gamma l \circ \tau(u)}, \quad u \in I_n,$$

which entails that, for all n large enough and $K_2 = k_0$,

$$\left|\rho\left(|\tau(u) + \varepsilon_n(u)|\right) - \rho\left(\tau(u)\right)\right| \leqslant k_0 \rho' \circ \tau(u) \left|\varepsilon_n(u)\right| \leqslant K_2 \rho' \circ F^{-1}(u) \left|\varepsilon_n(u)\right|. \quad (42)$$

Case 3. Assume that $0 < \gamma \leqslant 1$ in $(C2)$. Since $l(x)/x$ is either decreasing or, if $\gamma = 1$, not even monotone, $l \circ \tau(u)/\tau(u)$ cannot be compared to the worse case $\tau(u) \sim F^{-1}(u)$ directly. However, by Proposition 31, if $u \in I_n$ is such that $|\varepsilon_n(u)| \leqslant \delta_0/l'(\tau(u))$ then (42) holds. Consider

$$I_n^- = \left\{u \in I_n : |\varepsilon_n(u)| > \frac{\delta_0}{l' \circ \tau(u)}\right\}.$$

Since $l'(x) \sim \gamma l(x)/x$ and $\rho(x) \sim x\rho'(x)/\gamma l(x)$ as $x \to +\infty$, for any $0 < x_0 < \tau_0$ we can find $\xi_0 > 1/\gamma$ such that

$$\rho(x) \leqslant \xi_0 \rho'(x) \frac{x}{l(x)}, \quad x \geqslant x_0. \quad (43)$$

Let $\xi_1 > \gamma/\delta_0$ and assume n so large that $l \circ \tau(u) > 1/\xi_1$ and $\tau(u) \geqslant \tau_0$ for $u \in I_n$. Any $u \in I_n^-$ then satisfies

$$\begin{aligned}\tau_0 &\leqslant \max\left(\tau(u), |\varepsilon_n(u)|\right) \\ &\leqslant \max\left(\delta_0 \xi_1 \frac{l \circ \tau(u)}{l' \circ \tau(u)}, |\varepsilon_n(u)|\right) \\ &\leqslant \frac{x_n(u)}{2} := \xi_1 l \circ \tau(u) \left|\varepsilon_n(u)\right|.\end{aligned}$$

By (43) and the fact that $l(x)$ is increasing it follows that

$$\begin{aligned}\left|\rho\left(|\tau(u)+\varepsilon_n(u)|\right) - \rho\left(\tau(u)\right)\right| &\leqslant \rho\left(\tau(u) + |\varepsilon_n(u)|\right) \\ &\leqslant \xi_0 \rho'(x_n(u)) \frac{x_n(u)}{l \circ \tau(u)} \\ &= 2\xi_0\xi_1 \rho'(x_n(u)) \left|\varepsilon_n(u)\right|.\end{aligned}$$

Using (40) as for (41) we almost surely eventually have

$$\frac{1}{2\xi_1} \sup_{u \in I_n^-} \frac{x_n(u)}{F^{-1}(u)} = \sup_{u \in I_n^-} l \circ \tau(u) \frac{|\varepsilon_n(u)|}{F^{-1}(u)} \leqslant K \frac{\log n}{\sqrt{k_n}} \sqrt{\log\log n} \sup_{u \in I_n^-} H(u)$$

and the upper bound tends to 0 provided that $2 < \theta' < \theta$ from (25). As a conclusion, $x_n(u) \leqslant F^{-1}(u)$ on I_n^- even if $|\varepsilon_n(u)|$ is large and it asymptotically holds, for $K_2 = \max(k_0, 2\xi_0\xi_1)$,

$$\left|\rho\left(|\tau(u)+\varepsilon_n(u)|\right) - \rho\left(\tau(u)\right)\right| \leqslant K_2 \rho'(F^{-1}(u)) \left|\varepsilon_n(u)\right|, \quad u \in I_n.$$

6.1.3 Strong approximation of the joint quantile processes

In this section $(FG1)$ and $(FG2)$ are crucially required to justify the key approximation used at steps 4 and of the main proof. Let k_n be defined as in (24), thus $k_n/n \to 0$, $k_n/\log\log n \to +\infty$. Consider $\mathcal{I}_n = (k_n/n, 1-k_n/n)$ which contains both $I_{M,n}$ from (32) and Δ_n from (27). As in (28) write $\beta_n^X = \sqrt{n}(\mathbb{F}_n^{-1} - F^{-1})$ and $\beta_n^Y = \sqrt{n}(\mathbb{G}_n^{-1} - G^{-1})$ the quantile processes associated to each sample. Our goal is to derive a coupling of

$$\left\{(\beta_n^X(u), \beta_n^Y(u)) : u \in \mathcal{I}_n\right\} \quad \text{and} \quad \left\{\left(\frac{B_n^X(u)}{h_X(u)}, \frac{B_n^Y(u)}{h_Y(u)}\right) : u \in \mathcal{I}_n\right\}$$

where (B_n^X, B_n^Y) are two marginal standard Brownian Bridges

$$B_n^X(u) = \mathbb{B}_n\left(\mathcal{H}_{F^{-1}(u)}\right), \quad \mathcal{H}_{x_0} = \left\{(x,y) : x \leqslant x_0\right\}, \tag{44}$$

$$B_n^Y(u) = \mathbb{B}_n\left(\mathcal{H}^{G^{-1}(u)}\right), \quad \mathcal{H}^{y_0} = \left\{(x,y) : y \leqslant y_0\right\}, \tag{45}$$

indexed by $u \in [0,1]$ and driven by a sequence $\mathbb{B}_n$ of Π-Brownian Bridge indexed by the collection $\mathcal{C}$ of half planes $\mathcal{H}_{x_0}$ or $\mathcal{H}^{y_0}$. In other words, $\mathbb{B}_n$ is a zero mean Gaussian process indexed by $\mathcal{C}$ having covariance

$$cov(\mathbb{B}_n(A), \mathbb{B}_n(B)) = \Pi(A \cap B) - \Pi(A)\Pi(B)$$

for $A, B \in \mathcal{C}$, and B_n^X are centered Gaussian processes with covariance

$$
\begin{aligned}
cov(B_n^X(u), B_n^X(v)) &= \Pi(\mathcal{H}_{F^{-1}(u)} \cap \mathcal{H}_{F^{-1}(v)}) - uv = \min(u,v) - uv \\
cov(B_n^Y(u), B_n^Y(v)) &= \Pi(\mathcal{H}^{G^{-1}(u)} \cap \mathcal{H}^{G^{-1}(v)}) - uv = \min(u,v) - uv \\
cov(B_n^X(u), B_n^Y(v)) &= \Pi(\mathcal{H}_{F^{-1}(u)} \cap \mathcal{H}^{G^{-1}(v)}) - uv = L(u,v) - uv
\end{aligned}
$$

for $u, v \in [0,1]$, where the usual copula function $L(u,v) = H(F^{-1}(u), G^{-1}(v))$ measures the distortion between Π and $P \otimes Q$ on all quadrants, half spaces and then rectangles.

The coupling is achieved at Theorem 28 simply by combining the strong approximation of the empirical process (see [3])

$$
\Lambda_n(A) = \sqrt{n}(\Pi_n(A) - \Pi(A)),\ A \in \mathcal{C},\ \Pi_n = \frac{1}{n}\sum_{i \leqslant n} \delta_{(X_i, Y_i)}
$$

with the usual quantile transform and classical results for real quantiles. This result has an interest by itself as it is valid whatever the joint law Π satisfying the marginal conditions $(FG1)$ and $(FG2)$.

Remark 27 *Theorem 28 remains valid for the d marginal quantile processes of a law Π in $\mathbb{R}^d$ provided each marginal laws obeys $(FG1)$ and $(FG2)$, with obviously no change in the proof for $d = 2$.*

Theorem 28 *Assume that F, G satisfy $(FG1)$ and $(FG2)$. One can built on the same probability space the sequence $\{(X_n, Y_n)\}$ and a sequence of versions of $\left\{(B_n^X(u), B_n^Y(u)) : u \in \mathcal{I}_n\right\}$ such that*

$$
\beta_n^X(u) = \frac{B_n^X(u) + Z_n^X(u)}{h_X(u)}, \quad \beta_n^Y(u) = \frac{B_n^Y(u) + Z_n^Y(u)}{h_Y(u)}
$$

satisfies, for some $\xi > 0$,

$$
\lim_{n \to +\infty} n^{\xi} \sup_{u \in \mathcal{I}_n} \left|Z_n^X(u)\right| = \lim_{n \to +\infty} n^{\xi} \sup_{u \in \mathcal{I}_n} \left|Z_n^Y(u)\right| = 0 \quad a.s.
$$

Moreover we can take

$$
(B_n^X(u), B_n^Y(u)) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (G_k^X(u), G_k^Y(u))
$$

where $\left\{(G_k^X(u), G_k^Y(u)) : u \in (0,1)\right\}$ is a sequence of independent versions of Brownian Bridges (G^X, G^Y) such that $cov(G^X(u), G^Y(v)) = L(u,v) - uv$.

Proof Define the two marginal empirical processes to be, for $x \in \mathbb{R}$,

$$
\begin{aligned}
\alpha_n^X(x) &= \sqrt{n}(\mathbb{F}_n(x) - F(x)) = \Lambda_n(\mathcal{H}_x), \\
\alpha_n^Y(x) &= \sqrt{n}(\mathbb{G}_n(x) - G(x)) = \Lambda_n(\mathcal{H}^x).
\end{aligned}
$$

Under $(FG1)$ the random variables $U_i = F(X_i)$ and $V_i = G(Y_i)$ are uniform on $(0,1)$. Write $\alpha_n^{X,U}$ and $\alpha_n^{Y,V}$ the uniform empirical process associated to $U_1, ..., U_n$ and $V_1, ..., V_n$ respectively. Also write $\mathbb{F}_{X,U,n}$ and $\mathbb{F}_{X,U,n}^{-1}$ the empirical c.d.f. and quantile functions then $\beta_n^{X,U}(u) = \sqrt{n}(\mathbb{F}_{X,U,n}^{-1}(u) - u)$. Likewise write $\mathbb{F}_{Y,V,n}$, $\mathbb{F}_{Y,V,n}^{-1}$ and $\beta_n^{Y,V}$. Clearly $\alpha_n^{X,U}$ and $\alpha_n^{Y,V}$ are not independent, neither are $\beta_n^{X,U}$ and $\beta_n^{Y,V}$. What is next obtained for X is also valid for Y.

Under $(FG1)$ and $(FG2)$ the arguments given in Section 5.1.1 yield that

$$\lim_{n\to+\infty} \frac{\sqrt{n}}{\log\log n} \sup_{u\in\mathcal{I}_n} \left| h_X(u)\beta_n^X(u) - \beta_n^{X,U}(u) \right| = 0 \quad a.s. \tag{46}$$

since $\beta_n^{X,U} = \sqrt{n}q_n$ and the supremum is showed to be less than $\sqrt{n}A_nB_n$ with the almost sure bounds such that $A_n < K(\log\log n)/n$ and $B_n \to 0$ as $n \to +\infty$. By [2] and [11] we also have

$$\limsup_{n\to+\infty} \frac{n^{1/4}}{\sqrt{\log n}(\log\log n)^{1/4}} \sup_{u\in\mathcal{I}_n} \left| \beta_n^{X,U}(u) + \alpha_n^{X,U}(u) \right| \leqslant \frac{1}{2^{1/4}} \quad a.s. \tag{47}$$

thus for any $\xi < 1/4$ it holds

$$\lim_{n\to+\infty} n^{\xi} \sup_{u\in\mathcal{I}_n} \left| h_X(u)\beta_n^X(u) + \alpha_n^{X,U}(u) \right| = 0 \quad a.s.$$

It is important here that (46) and (47) holds true for $\beta_n^{X,U}$ and $\beta_n^{Y,U}$ simultaneously with probability one whatever the underlying probability space. Hence, recalling that $\alpha_n^{X,U} = \alpha_n^X \circ F^{-1}$, $\Pi_n(\mathcal{H}_{F^{-1}(u)}) = \mathbb{F}_n(F^{-1}(u))$ and $\Pi(\mathcal{H}_{F^{-1}(u)}) = u$ it follows that

$$\lim_{n\to+\infty} n^{\xi} \sup_{u\in\mathcal{I}_n} \left| h_X(u)\beta_n^X(u) + \Lambda_n(\mathcal{H}_{F^{-1}(u)}) \right| = 0 \quad a.s.$$

$$\lim_{n\to+\infty} n^{\xi} \sup_{u\in\mathcal{I}_n} \left| h_Y(u)\beta_n^Y(u) + \Lambda_n(\mathcal{H}^{G^{-1}(u)}) \right| = 0 \quad a.s.$$

on any probability space. It remains to approximate Λ_n uniformly on $\mathcal{C}$. The collection of sets $\mathcal{C}$ is a VC-class of order 3 thus satisfies the uniform entropy condition (VC) used in [3] with $v_0 = 2(3-1) = 4$. By their Proposition 1 taken with $\theta = 2$ there exists a probability space on which the sequence $\{(X_n, Y_n)\}$ can be built together with a sequence $\mathbb{B}_n$ of Π-Brownian Bridges indexed by $\mathcal{C}$ such that

$$\mathbb{P}\left(\sup_{A\in\mathcal{C}} |\Lambda_n(A) + \mathbb{B}_n(A)| \geqslant \frac{K}{n^{\beta_2}} \right) \leqslant \frac{1}{n^2}$$

where we take $\beta_2 > 1/22$ to avoid the $\log n$ factor. Note that since $\mathbb{B}_n$ and $-\mathbb{B}_n$ have the same law, we choose to approximate with $-\mathbb{B}_n$. Consider in particular $\mathcal{H}_n^X = \left\{ \mathcal{H}_{F^{-1}(u)} : u \in \mathcal{I}_n \right\} \subset \mathcal{C}$ and define $B_n^X(u) = \mathbb{B}_n\left(\mathcal{H}_{F^{-1}(u)}\right)$. On the previous probability space it holds

$$\begin{aligned} &\limsup_{n\to+\infty} n^{\beta_2} \sup_{u\in\mathcal{I}_n} \left| \alpha_n^X \circ F^{-1}(u) + B_n^X(u) \right| \\ &= \limsup_{n\to+\infty} n^{\beta_2} \sup_{A\in\mathcal{H}_n^X} |\Lambda_n(A) + \mathbb{B}_n(A)| \leqslant K \quad a.s. \end{aligned}$$

the above comparision between $h_X(u)\beta_n^X(u)$ and $\alpha_n^X \circ F^{-1}(u)$ gives in turn, for $\xi < \max(1/4, \beta_2) = \beta_2$ and $Z_n^X(u) = h_X(u)\beta_n^X(u) - B_n^X(u)$,

$$\limsup_{n\to+\infty} n^{\xi} \sup_{u\in\mathcal{I}_n} \left|Z_n^X(u)\right| = 0 \quad a.s.$$

In the same way we simultaneously obtain, for $Z_n^Y(u) = h_Y(u)\beta_n^Y(u) - B_n^Y(u)$,

$$\limsup_{n\to+\infty} n^{\xi} \sup_{u\in\mathcal{I}_n} \left|Z_n^Y(u)\right| = 0 \quad a.s.$$

The processes B_n^X and B_n^Y are joint through the leading process $\mathbb{B}_n$, whence the covariance $cov(B_n^X(u), B_n^Y(v)) = L(u,v) - uv$. The second statement to be proved follows by applying Theorem 1 of [3] in place of Proposition 1. If $\beta_2 > 0$ is chosen small enough the approximating process can be built in the form $\mathbb{B}_n = \sum_{k=1}^n \mathbb{B}_k^*/\sqrt{n}$ where $\{\mathbb{B}_k^* : k \geqslant 1\}$ is a sequence of independent Π-Brownian Bridges. Since $\mathbb{B}_n$ is again a Π-Brownian Bridge, $G_k^X(u) = \mathbb{B}_k^*(\mathcal{H}_{F^{-1}(u)})$ and $G_k^Y(u) = \mathbb{B}_k^*(\mathcal{H}^{G^{-1}(u)})$ are standard Brownian Bridges with the desired correlation structure. ■

6.2 Complements on assumptions

6.2.1 Regular an smooth slow variation

In this section, we present the regular and slow variation properties needed for assumption (C2). For more details we refer to [12, 15]. For $k \in \mathbb{N}_*$ write $\mathcal{C}_k$ the set of functions that are k times continuously differentiable on $\mathbb{R}$, and $\mathcal{C}_0$ the set of continuous functions. Let $\mathcal{M}_k(m, +\infty)$ be the subset of functions $\varphi \in \mathcal{C}_k$ such that $\varphi^{(k)}$ is monotone on $(m, +\infty)$, and hence φ, φ', φ'',...,$\varphi^{(k)}$ are also monotone on $(m, +\infty)$ by changing m. Let $\mathcal{M}_0(m, +\infty)$ denote the set of continuous functions monotone on $(m, +\infty)$. Write $RV(\gamma)$ the set of regularly varying functions at $+\infty$ with index $\gamma \in \mathbb{R}$. They are of the form $x^{\gamma}L(x)$ with $L \in RV(0)$, which means that given any $\lambda > 0$,

$$\lim_{x\to+\infty} \frac{L(\lambda x)}{L(x)} = 1. \tag{48}$$

If $L \in RV(0)$ is monotone on $(m, +\infty)$ then L is equivalent at $+\infty$ to a function in $\mathcal{C}_\infty(m, +\infty) \cap RV(0)$. Therefore, at the first order, it is not a restriction to assume that functions of $RV(\gamma)$ are in $\mathcal{M}_k(m, +\infty)$ as well. Problems however arise with respect to differentiation. In particular, two apparently close slowly varying functions may have very different local variations. First consider the smooth regular variation. Let introduce

$$RV_k(\gamma, m) = RV(\gamma) \cap \mathcal{M}_k(m, +\infty), \quad \gamma \neq 0.$$

The following statements are taken as $x \to +\infty$. Assuming that $k \geqslant 1$ and $\gamma \neq 0$, if $\varphi \in RV_k(\gamma, m)$ then φ' is monotone, so that it holds, by the monotone density theorem,

$$\varphi'(x) \sim \frac{\gamma\varphi(x)}{x}. \tag{49}$$

This implies that $\varphi' \in RV_{k-1}(\gamma-1,m)$ and, whenever $k \geqslant 2$ and $\gamma \neq 1$, φ'' in turns satisfies $\varphi'' \in RV_{k-2}(\gamma-2,m)$ and

$$\varphi''(x) \sim \frac{(\gamma-1)\varphi'(x)}{x} \sim \frac{\gamma(\gamma-1)\varphi(x)}{x^2}. \tag{50}$$

For $L \in RV(0)$ it holds, by Karamata's theorem,

$$\frac{\int_m^x L'(t)\left(\frac{L(t)}{tL'(t)}\right)dt}{\int_m^x L'(t)dt} = \frac{1}{L(x)}\int_m^x \frac{L(t)}{t}dt \to +\infty.$$

Hence the function $L(t)/tL'(t)$ is unbounded and, if $L \in \mathcal{C}_1(m,+\infty)$, continuous on $(m,+\infty)$. It is not very restrictive to exclude functions $L(t)/tL'(t)$ that are asymptotically oscillating and not going to infinity. We thus assume (7).
For instance, if $L(x) = \varphi(\log x)$ where $\varphi \in RV_2(\gamma,m)$ and $\gamma > 0$ then $\varepsilon_1(x) \sim \gamma/\log x$. Likewise, if $L(x) = \varphi(L_1(x))$ where $\varphi \in RV_2(\gamma,m)$ and $\gamma > 0$ then we get $\varepsilon_1(x) \sim \gamma x L_1'(x)/L_1(x)$. Also remind the well known representation, for $x \in (m,+\infty)$,

$$L(x) = d_0(x)\exp\left(\int_m^x \frac{\varepsilon_0(t)}{t}dt\right), \quad d_0(x) \to d_0 > 0, \quad \varepsilon_0(x) \to 0.$$

If $d_0(x)$ is constant then $d_0 = L(m)$ and $\varepsilon_0(x) = \varepsilon_1(x)$ from (7). More generally, (7) is equivalent to $xd_0'(x) \to 0$ and we have $\varepsilon_1(x) = \varepsilon_0(x) + xd_0'(x)$.

6.2.2 A sufficient condition for (FG)

In this section we provide a sufficient condition to

$$(FG2)\ \sup_{x>m}(1-F(x))\frac{|f'(x)|}{f^2(x)} < +\infty, \quad (FG3)\ \sup_{u>F(m)} H_X(u) < +\infty,$$

based on standard regular variation or smooth slow variation. Starting from

$$F(x) = 1-\exp(-\psi_X(x)), \quad F^{-1}(u) = \psi_X^{-1}(\log(1/(1-u))),$$
$$f(x) = \psi_X'(x)\exp(-\psi_X(x)), \quad h_X(u) = (1-u)\psi_X' \circ \psi_X^{-1}(\log(1/(1-u))),$$

we have

$$H_X(u) = \frac{1-u}{F^{-1}(u)h_X(u)} = \left(\log\psi_X^{-1}\right)'(\log(1/(1-u)))$$

thus $(FG3)$ holds whenever $\left(\log\psi_X^{-1}\right)'(x)$ is bounded, or $1/x\psi_X'(x)$ is bounded. Conversely, $(FG3)$ implies that $F^{-1}(u) = O(1/(1-u)^K)$ for K bounding H_X since $\left(\log F^{-1}(u)\right)' = H_X(u)/(1-u)$. In the same vein, $(FG2)$ is equivalent to

$$\sup_{m<x<+\infty}\left|\left(\frac{1}{\psi_X'(x)}\right)'\right| < +\infty \tag{51}$$

since $f'(x) = \left(-\psi''_X(x) - \psi'^2_X(x)\right)\exp(-\psi_X(x))$ and

$$(1-F(x))\,\frac{|f'(x)|}{f^2(x)} = \left|\frac{\psi''_X(x)+\psi'^2_X(x)}{\psi'^2_X(x)}\right| = \left|\frac{\psi''_X(x)}{\psi'^2_X(x)}+1\right| = \left|\left(\frac{1}{\psi'_X(x)}\right)' - 1\right|.$$

Proposition 29 *If $\psi_X \in RV_2^+(0,m)$ then F satisfies (FG). If $\psi_X \in RV_2(\gamma_1,m)$ for some $\gamma_1 > \gamma_0 > 0$ and, if $\gamma_1 = 1$ assuming also that $\psi_X(x) = xL(x)$ with $L' \in RV_1(-1,m)$ and (7), then F satisfies (FG) and $(FG3)$ can be replaced by*

$$H_X(u) \leqslant \frac{1}{\gamma_0 \log(1/(1-u))}, \quad u > F(m). \tag{52}$$

Proof Clearly ψ_X is $\mathcal{C}_2$, increases to infinity, $F(x) = 1 - e^{-\psi_X(x)}$ has unbounded right tail and $f(x) = \psi'_X(x)\exp(-\psi_X(x))$ is $\mathcal{C}_1$ which yields $(FG1)$. Next we check $(FG2)$ and $(FG3)$ in the two cases.
Case 1. Assume that $\gamma_1 > 0$. If $\gamma_1 \neq 1$ then (49) and (50) give, as $x \to +\infty$,

$$\left(\frac{1}{\psi'_X(x)}\right)' = \frac{\psi''_X(x)}{\psi'^2_X(x)} \sim \frac{\gamma_1 - 1}{\gamma_1 \psi_X(x)} \to 0.$$

If $\gamma_1 = 1$ then $\psi_X(x) = xL(x)$ with $L \in RV_2(0,m)$, $L' \in RV_1(-1,m)$ and (7) thus $L'(x) \sim -xL''(x)$ and $xL'(x)/L(x) \to 0$ which entails

$$\left|\left(\frac{1}{\psi'_X(x)}\right)'\right| = \left|\frac{\psi''_X(x)}{\psi'^2_X(x)}\right| = \frac{|2L'(x)+xL''(x)|}{(L(x)+xL'(x))^2} \leqslant K\frac{|L'(x)|}{L^2(x)} \to 0.$$

Whence (51) and $(FG2)$.
Whatever $\gamma_1 > 0$, ψ_X is continuous and strictly increasing with inverse $\psi_X^{-1} \in RV_2(1/\gamma_1, \psi_X^{-1}(m))$. By using again (49) we obtain

$$\left(\log \psi_X^{-1}(x)\right)' = \frac{1}{\psi_X^{-1}(x)\psi'_X \circ \psi_X^{-1}(x)} \sim \frac{1}{\gamma_1 \psi_X \circ \psi_X^{-1}(x)} = \frac{1}{\gamma_1 x}$$

as $x \to +\infty$, which is bounded. This implies $(FG3)$ and more accurately (52) since for $\gamma_1 > \gamma_0 > 0$ and m sufficiently large,

$$H_X(u) = \left(\log \psi_X^{-1}\right)'(\log(1/(1-u))) \leqslant \frac{1}{\gamma_0 \log(1/(1-u))}, \quad u > F(m).$$

Case 2. If $\gamma_1 = 0$ then $x\psi'_X(x) = \varepsilon_1(x)\psi_X(x) \geqslant l_1 \geqslant 1$ with $\varepsilon_1(x) \to 0$ as $x \to +\infty$, by (7) and (8). Now $l_1/x \leqslant \psi'_X(x) \leqslant \psi_X(x)/x$ implies that $\psi'_X \in RV_1(-1,m)$ and (50) yields $\psi''_X(x) \sim -\psi'_X(x)/x = -\varepsilon_1(x)\psi_X(x)/x^2$. It ensues $(1/\psi'_X(x))' = \psi''_X(x)/\psi'^2_X(x) \sim -1/\varepsilon_1(x)\psi_X(x)$. Therefore the upper bound in $(FG2)$ is, for $x \in (m,+\infty)$,

$$\left|\left(\frac{1}{\psi'_X(x)}\right)' - 1\right| = 1 + \frac{1}{\varepsilon_1(x)\psi_X(x)} \leqslant \frac{l_1+1}{l_1}$$

and the upper bound in $(FG3)$ is

$$\left(\log \psi_X^{-1}(x)\right)' = \frac{1}{\psi_X^{-1}(x)\psi_X' \circ \psi_X^{-1}(x)} = \left(\frac{1}{\varepsilon_1 \psi_X}\right) \circ \psi_X^{-1}(x) \leqslant \frac{1}{l_1}.$$

Note that if $\gamma_1 > 0$ we have $\varepsilon_1(x) \to \gamma_1$ so the second equality in the left-hand side yields back the sharper bound $1/\gamma_1 x$ used for (52). ∎

Corollary 30 *Let (C) hold with $\gamma > 0$. Assume that F and G satisfy $(FG4)$ together with the condition of Proposition 29 with $\gamma_1 > 0$. Assume that (CFG) holds with $\theta > 1$. Then the conclusion of Theorem 14 remains true.*

Proof The result was proved for $\gamma > 0$ and $\theta > 2$. The only changes needed for $\theta > 1$ are at cases 2 and 3 in the proof of Lemma 24 at Section 5.1.2. We have $\psi_Y \geqslant \psi_X$, $\psi_X \in RV_2(\gamma_1, m)$ and $\gamma_1 \geqslant \gamma > 0$ by (15). Applying (52) from Proposition 29 yields

$$\sup_{u \in I_n} H(u) \leqslant \sup_{u \in I_n} \frac{2}{\gamma_0 \log(1/(1-u))} \leqslant \frac{2}{(1-\beta)\gamma_0 \log n}$$

and this extra $1/(\log n)$ makes the bounding sequence in (41) tends to 0 provided that $1 < \theta' < \theta$ in (25). ∎

6.2.3 Consequences of (C2)

Proposition 31 *Assume $(C2)$. Then it holds*

$$\rho(|x+\varepsilon|) - \rho(x) = k_0(x,\varepsilon)\rho'(x)\varepsilon$$

where, for any $x_0 > \tau_1$,

$$\lim_{\delta_0 \to 0} \sup_{x > x_0} \sup_{|\varepsilon| l'(x) \leqslant \delta_0} |k_0(x,\varepsilon) - 1| = 0. \tag{53}$$

In particular, there exists $\delta_0 > 0$ and $k_0 > 0$ such that, for all $x > x_0$ and $|\varepsilon| \leqslant \delta_0 / l'(x)$ we have $|\rho(|x+\varepsilon|) - \rho(x)| \leqslant k_0 \rho'(x) |\varepsilon|$.

Proof Fix $x_0 > \tau_1 > 0$ and let $M > x_0$ be as large as needed below. If $\varepsilon = 0$ then (53) requires that $k_0(x,0) = 1$ for $x > x_0$. For $\varepsilon \neq 0$ we distinguish between $x \in (x_0, M)$ and $x \geqslant M$. In the first case, since $\rho \in \mathcal{C}_2$ under $(C2)$ the Taylor expansion of ρ holds uniformely on (x_0, M). Namely, for any δ_0 small enough, $x \in (x_0, M)$ and $|\varepsilon| \leqslant \varepsilon_0 = \delta_0 / \inf\{l'(x) : x \in (x_0, M)\} < x_0 - \tau_1$ we have

$$\rho(|x+\varepsilon|) - \rho(x) = k_0(x,\varepsilon)\rho'(x)\varepsilon, \quad k_0(x,\varepsilon) = 1 + \frac{\rho''(x^*)}{2\rho'(x)}\varepsilon,$$

with $x^* \in (x_0 - \varepsilon_0, M + \varepsilon_0)$ and $|k_0(x,\varepsilon) - 1| \leqslant K\delta_0$ where $K < +\infty$ depends on x_0, M, ρ. We deduce that, for any $M > x_0$,

$$\lim_{\delta_0 \to 0} \sup_{x_0 < x < M} \sup_{|\varepsilon| l'(x) \leqslant \delta_0} |k_0(x,\varepsilon) - 1| = 0. \tag{54}$$

If $x \geqslant M$ then $l'(x) > 0$ and we intend to expand

$$\rho(|x+\varepsilon|) - \rho(x) = \rho(x)\left(\exp(l(|x+\varepsilon|) - l(x)) - 1\right). \tag{55}$$

Case 1. Assume $(C2)$ with $\gamma > 0$. Write $l(x) = x^\gamma L(x)$ where $L \in \mathcal{RV}_2(0, \tau_1)$ satisfies (7). For any $\delta_0 \in (0, \gamma l(M)/4)$ define

$$\Delta_0 = \{(x,\varepsilon) : x \geqslant M,\ |\varepsilon|\, l'(x) \leqslant \delta_0\}. \tag{56}$$

By (49), for M large enough and $(x,\varepsilon) \in \Delta_0$ it holds $l'(x) > \gamma l(x)/2x$, which implies $|\varepsilon|/x \leqslant 2\delta_0/\gamma l(x) < 1/2$ and $|x+\varepsilon| = x + \varepsilon > M/2$. Therefore $\sup_{(x,\varepsilon)\in\Delta_0} |\varepsilon|/x \to 0$ as $\delta_0 \to 0$ and we have, for $(x,\varepsilon) \in \Delta_0$,

$$\begin{aligned}\frac{l(x+\varepsilon) - l(x)}{x^\gamma} &= \left(1 + \frac{\varepsilon}{x}\right)^\gamma L(x+\varepsilon) - L(x)\\ &= \frac{\gamma\varepsilon}{x}(1 + \delta_1(x,\varepsilon))L(x+\varepsilon) + L(x+\varepsilon) - L(x)\end{aligned} \tag{57}$$

where $\sup_{(x,\varepsilon)\in\Delta_0} |\delta_1(\varepsilon, x)| \to 0$ as $\delta_0 \to 0$. By (7) we also have, for $(x,\varepsilon) \in \Delta_0$,

$$|L(x+\varepsilon) - L(x)| \leqslant \sup_{|y-x|\leqslant|\varepsilon|} |L'(y)|\,|\varepsilon| = \sup_{|y-x|\leqslant|\varepsilon|} |\varepsilon_1(y)|\,\frac{L(y)}{y}\,|\varepsilon|$$

where $\varepsilon_1(y) \to 0$ as $y > x - |\varepsilon| > M/2 \to +\infty$. Moreover, for $\delta = 2\delta_0/\gamma l(M)$,

$$\frac{1}{L(x)} \sup_{|y-x|\leqslant|\varepsilon|} L(y) = \sup_{1-|\varepsilon|/x<\lambda<1+|\varepsilon|/x} \frac{L(\lambda x)}{L(x)} \leqslant \sup_{1-\delta<\lambda<1+\delta} \frac{L(\lambda x)}{L(x)}$$

and the second term has limit 1 as $x \to +\infty$ since $L \in RV(0)$. Hence for any $\varepsilon_1 > 0$, assuming M so large that $\sup_{y>M/2} |\varepsilon_1(y)| < \varepsilon_1/4$ and δ_0 small ensures that, for $(x,\varepsilon) \in \Delta_0$,

$$|L(x+\varepsilon) - L(x)| \leqslant \frac{\varepsilon_1}{3}\frac{L(x)}{x - |\varepsilon|}\,|\varepsilon| \leqslant \frac{\varepsilon_1}{2}\frac{|\varepsilon|}{x}L(x)$$

and (57) reads

$$l(x+\varepsilon) - l(x) = (1 + \delta_2(x,\varepsilon))\frac{\gamma x^\gamma L(x)}{x}\varepsilon = (1 + \delta_3(x,\varepsilon))l'(x)\varepsilon$$

then (55) gives

$$\frac{\rho(x+\varepsilon) - \rho(x)}{l'(x)\rho(x)\varepsilon} = \frac{\exp(l(x+\varepsilon) - l(x)) - 1}{l'(x)\varepsilon} = 1 + \delta_4(x,\varepsilon) = k_0(x,\varepsilon)$$

with $\sup_{(x,\varepsilon)\in\Delta_0} |\delta_k(\varepsilon, x)| < \varepsilon_1$ for $k = 2,3,4$. We have proved that for any $\varepsilon_1 > 0$ there exists M such that

$$\lim_{\delta_0\to 0} \sup_{x\geqslant M} \sup_{|\varepsilon| l'(x)\leqslant\delta_0} |k_0(x,\varepsilon) - 1| \leqslant \varepsilon_1$$

which yields (53) when combined to (54).
Case 2. Assume $(C2)$ with $\gamma = 0$. Since $l \in \mathcal{RV}_2^+(0, \tau_1)$, (7) and (8) give

$$\varepsilon_1(x) = \frac{xl'(x)}{l(x)} \geqslant \frac{l_1}{l(x)} > 0$$

where $\varepsilon_1(x) \to 0$. Thus $l'(x) \to 0$ as $x \to +\infty$ and $l' \in \mathcal{RV}_2(-1, \tau_1)$. Thus l' is decreasing on $(M, +\infty)$ since $l' \in \mathcal{M}_2(\tau_1, +\infty)$. Condider $\delta_0 \in (0, 1/2l_1)$ and define Δ_0 as in (56). For $(x, \varepsilon) \in \Delta_0$ it holds

$$\frac{|\varepsilon|}{x} \leqslant l_1 \frac{|\varepsilon|}{x} \leqslant l(x)\varepsilon_1(x)\frac{|\varepsilon|}{x} = |\varepsilon|\, l'(x) \leqslant \delta_0$$

hence $|x + \varepsilon| = x + \varepsilon > M/2$ again, and

$$l'(x + |\varepsilon|)\, |\varepsilon| \leqslant |l(x + \varepsilon) - l(x)| \leqslant l'(x - |\varepsilon|)\, |\varepsilon|$$

where, since $l''(x) \sim -l'(x)/x$ by (50),

$$0 \leqslant \frac{l'(x - |\varepsilon|) - l'(x)}{l'(x - |\varepsilon|)} \leqslant \sup_{x - |\varepsilon| \leqslant y \leqslant x} \frac{|l''(y)\varepsilon|}{l'(y)} \leqslant \frac{|\varepsilon|}{x - |\varepsilon|} \leqslant 2\delta_0,$$
$$0 \leqslant \frac{l'(x) - l'(x + |\varepsilon|)}{l'(x)} \leqslant \sup_{x \leqslant y \leqslant x + |\varepsilon|} \frac{|l''(y)\varepsilon|}{l'(y)} \leqslant \frac{|\varepsilon|}{x} \leqslant \delta_0.$$

We deduce that for $k = 1, 2$ and $\sup_{(x,\varepsilon) \in \Delta_0} |\delta_k(\varepsilon, x)| \to 0$ as $\delta_0 \to 0$ it holds

$$l(x + \varepsilon) - l(x) = (1 + \delta_1(x, \varepsilon))l'(x)\varepsilon$$

for all $(x, \varepsilon) \in \Delta_0$ and, by (55),

$$\frac{\rho(x + \varepsilon) - \rho(x)}{l'(x)\rho(x)\varepsilon} = \frac{\exp(l(x + \varepsilon) - l(x)) - 1}{l'(x)\varepsilon} = 1 + \delta_2(x, \varepsilon) = k_0(x, \varepsilon)$$

thus (53) follows. ■
Several arguments exploit the asymptotic convexity of ρ which follows from$(C2)$:

Proposition 32 *Under $(C2)$ the function $\rho(x)$ is convex on $(l_2, +\infty)$ for some $l_2 > 0$. If moreover $l_1 > 1$ it is strictly convex.*

Proof We have to show that $\rho''(x) = (l''(x) + l'(x)^2)\rho(x) \geqslant 0$ if $(C2)$ holds. In the case $1 \neq \gamma > 0$ we have, by (49) and (50), as $x \to +\infty$,

$$l'(x) \sim \frac{\gamma l(x)}{x}, \quad l''(x) \sim \frac{\gamma(\gamma - 1)l(x)}{x^2} \ll l'(x), \quad \frac{l''(x)}{l'(x)} \sim \frac{\gamma - 1}{x},$$

thus there exists $l_2 > l^{-1}(1/\gamma)$ such that all $x > l_2$ satisfy $l'(x) > 0$ and

$$l''(x) + l'(x)^2 \sim l'(x)\left(\frac{\gamma - 1}{x} + \frac{\gamma l(x)}{x}\right) \geqslant \frac{l'(x)}{x}(\gamma l(x) - 1) > 0.$$

If $\gamma = 1$ then $l(x) = xL(x)$ and $l''(x) = 2L'(x) + xL''(x) \sim L'(x)$ whereas $l'(x)^2 \sim (L(x) + xL'(x))^2 \sim L^2(x)$. Since $L'(x)/L^2(x) = \varepsilon_1(x)/xL(x) \to 0$ we have $l''(x) + l'(x)^2 > 0$ for $x > l_2$. If $\gamma = 0$ in $(C2)$ then by (7) and (8) we have

$$xl'(x) = \varepsilon_1(x)l(x) \geqslant l_1 \geqslant 1, \quad \varepsilon_1(x) \to 0,$$

hence

$$\frac{l_1}{x} \leqslant l'(x) \leqslant \frac{l(x)}{x}$$

and $l' \in RV_1^+(-1, 0)$. Now by (50) we get, as $x \to +\infty$,

$$l'(x) \sim \frac{\varepsilon_1(x)l(x)}{x}, \quad l''(x) \sim -\frac{l'(x)}{x} = -\frac{\varepsilon_1(x)l(x)}{x^2}, \quad \frac{l''(x)}{l'(x)} \sim -\frac{1}{x},$$

so that

$$l''(x) + l'(x)^2 \sim \frac{l'(x)}{x}\left(\varepsilon_1(x)l(x) - 1\right) \geqslant \frac{l'(x)}{x}\left(l_1 - 1\right).$$

Therefore if $l_1 > 1$ $\rho(x)$ is strictly convex on $(l_2, +\infty)$ for l_2 large enough. It remains convex for $l_1 = 1$. ■

References

[1] P. C. Álvarez-Esteban, E. del Barrio, J. A. Cuesta-Albertos, and C. Matrán. Uniqueness and approximate computation of optimal incomplete transportation plans. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(2):358–375, 2011.

[2] R. R. Bahadur. A note on quantiles in large samples. *Ann. Math. Statist.*, 37:577–580, 1966.

[3] P. Berthet and D. M. Mason. Revisiting two strong approximation results of Dudley and Philipp. In *High dimensional probability*, volume 51 of *IMS Lecture Notes Monogr. Ser.*, pages 155–172. Inst. Math. Statist., Beachwood, OH, 2006.

[4] S. G. Bobkov and M. Ledoux. One-dimensional empirical measures, order statistics and kantorovich transport distances. *To appear in: Memoirs of the AMS*, Preprint 2016.

[5] S. Cambanis, G. Simons, and W. Stout. Inequalities for $Ek(X, Y)$ when the marginals are fixed. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 36(4):285–294, 1976.

[6] M. Csörgö and P. Révész. Strong approximations of the quantile process. *Ann. Statist.*, 6(4):882–894, 1978.

[7] E. del Barrio, E. Giné, and C. Matrán. Central limit theorems for the Wasserstein distance between the empirical and the true distributions. *Ann. Probab.*, 27(2):1009–1071, 1999.

[8] E. del Barrio, E. Giné, and C. Matrán. Correction: "Central limit theorems for the Wasserstein distance between the empirical and the true distributions" [Ann. Probab. **27** (1999), no. 2, 1009–1071; MR1698999 (2000g:60034)]. *Ann. Probab.*, 31(2):1142–1143, 2003.

[9] E. del Barrio, E. Giné, and F. Utzet. Asymptotics for L_2 functionals of the empirical quantile process, with applications to tests of fit based on weighted Wasserstein distances. *Bernoulli*, 11(1):131–189, 2005.

[10] E. del Barrio and J.-M. Loubès. Central limit theorems for empirical transportation cost in general dimension. *arXiv:1705.01299v2*, may 2017.

[11] J. Kiefer. Deviations between the sample quantile process and the sample df. In *Nonparametric Techniques in Statistical Inference (Proc. Sympos., Indiana Univ., Bloomington, Ind., 1969)*, pages 299–319. Cambridge Univ. Press, London, 1970.

[12] T. Mikosch, Operations Research EURANDOM European Institute for Statistics, Probability, and their Applications. *Regular Variation, Subexponentiality and Their Applications in Probability Theory.* EURANDOM report. Eindhoven University of Technology, 1999.

[13] A. Munk and C. Czado. Nonparametric validation of similar distributions and assessment of goodness of fit. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 60(1):223–241, 1998.

[14] T. Rippl, A. Munk, and A. Sturm. Limit laws of the empirical Wasserstein distance: Gaussian distributions. *J. Multivariate Anal.*, 151:90–109, 2016.

[15] E. Seneta. *Regularly varying functions.* Lecture Notes in Mathematics, Vol. 508. Springer-Verlag, Berlin-New York, 1976.

[16] M. Sommerfeld and A. Munk. Inference for Empirical Wasserstein Distances on Finite Spaces. *ArXiv e-prints*, October 2016.

[17] C. Villani. *Topics in Optimal Transportation.* American Mathematical Society, 2003.