CREST Centre de recherche en économie et statistique

Série des Documents de Travail

n° 2017-02
About tests of the "simplifying" assumption for conditional copulas

A. DERUMIGNY [1]
J-D.FERMANIAN [2]

[1] CREST-ENSAE. E-mail: alexis.derumigny@ensae.fr
[2] CREST-ENSAE. E-mail: jean-david.fermanian@ensae.fr

About tests of the "simplifying" assumption for conditional copulas

Alexis Derumigny*, Jean-David Fermanian†

December 18, 2016

Abstract

We discuss the so-called "simplifying assumption" of conditional copulas in a general framework. We introduce several tests of the latter assumption for non- and semiparametric copula models. Some related test procedures based on conditioning subsets instead of pointwise events are proposed. The limiting distribution of such test statistics under the null are approximated by several bootstrap schemes, most of them being new. We prove the validity of a particular semiparametric bootstrap scheme. Some simulations illustrate the relevance of our results.

Keywords: conditional copula, simplifying assumption, bootstrap.

MCS: 62G05, 62G08, 62G09.

1 Introduction

In statistical modelling and applied science more generally, it is very common to distinguish two subsets of variables: a random vector of interest (also called explained/exogenous variables) and a vector of covariates (explanatory/endogenous variables). The objective is to predict the law of the former vector given the latter vector belongs to some subset, possibly a singleton. This basic idea constitutes the first step towards forecasting some important statistical sub-products as conditional means, quantiles, volatilities, etc. Formally, consider a d-dimensional random vector $\mathbf{X}$. We are faced with two random sub-vectors $\mathbf{X}_I$ and $\mathbf{X}_J$, s.t. $\mathbf{X} = (\mathbf{X}_I, \mathbf{X}_J)$, $I \cup J = \{1, \dots, d\}$, $I \cap J = \emptyset$, and our models of interest specify the conditional law of $\mathbf{X}_I$ knowing $\mathbf{X}_J = \mathbf{x}_J$ or knowing $\mathbf{X}_J \in A_J$ for some subset $A_J \subset \mathbb{R}^{|J|}$. We use the standard notations for vectors: for any set of indices I, $\mathbf{x}_I$ means the $|I|$-dimensional vector whose arguments are the x_k, $k \in I$. For convenience and without a loss of generality, we will set $I = \{1, \dots, p\}$ and $J = \{p+1, \dots, d\}$.

Besides, the problem of dependence among the components of d-dimensional random vectors has been extensively studied in the academic literature and among practitioners in a lot of different fields. The raise of copulas for more than twenty years illustrates the need of flexible and realistic multivariate models and tools. When covariates are present and with our notations, the challenge is to study the dependence among the components of $\mathbf{X}_I$ given $\mathbf{X}_J$. Logically, the concept of conditional copulas (Patton 2006a, 2006b) has emerged. By definition, for any borel subset $A_J \subset \mathbb{R}^{d-p}$, a conditional copula of $\mathbf{X}_I$ given $(\mathbf{X}_J \in A_J)$ is denoted by $C_{I|J}(\cdot|\mathbf{X}_J \in A_J)$. This is the cdf of the random vector $(F_{1|J}(X_1|\mathbf{X}_J \in A_J), \dots, F_{p|J}(X_p|\mathbf{X}_J \in A_J))$ given $(\mathbf{X}_J \in A_J)$. Here, $F_{k|J}(\cdot|\mathbf{X}_J \in A_J)$ denotes the conditional law of X_k knowing $\mathbf{X}_J \in A_J$, $k = 1, \dots, p$. The latter conditional

*ENSAE, 3 avenue Pierre-Larousse, 92245 Malakoff cedex, France. alexis.derumigny@ensae.fr

†ENSAE, J120, 3 avenue Pierre-Larousse, 92245 Malakoff cedex, France. jean-david.fermanian@ensae.fr. This research has been supported by the Labex Ecodec.

distributions will be assumed continuous in this paper, implying the existence and uniqueness of $C_{I|J}$ (Sklar's theorem). In other words, for any $\mathbf{x}_I \in \mathbb{R}^p$,

$$\mathbb{P}\left(\mathbf{X}_I \leq \mathbf{x}_I | \mathbf{X}_J \in A_J\right) = C_{I|J}\left(F_{1|J}(x_1|\mathbf{X}_J \in A_J), \dots, F_{p|J}(x_p|\mathbf{X}_J \in A_J) \,|\, \mathbf{X}_J \in A_J\right).$$

In particular, when the conditioning events are reduced to singletons, we get that the conditional copula of $\mathbf{X}_I$ knowing $\mathbf{X}_J = \mathbf{x}_J$ is a cdf $C_{I|J}(\cdot|\mathbf{X}_J = \mathbf{x}_J)$ on $[0,1]^p$ s.t., for every $\mathbf{x}_I \in \mathbb{R}^p$,

$$\mathbb{P}\left(\mathbf{X}_I \leq \mathbf{x}_I | \mathbf{X}_J = \mathbf{x}_J\right) = C_{I|J}\left(F_{1|J}(x_1|\mathbf{X}_J = \mathbf{x}_J), \dots, F_{p|J}(x_p|\mathbf{X}_J = \mathbf{x}_J) \,|\, \mathbf{X}_J = \mathbf{x}_J\right).$$

With generalized inverse functions, an equivalent definition of a conditional copula is as follows:

$$C_{I|J}\left(\mathbf{u}_I \,|\, \mathbf{X}_J = \mathbf{x}_J\right) = F_{I|J}\left(F^{-}_{1|J}(u_1|\mathbf{X}_J = \mathbf{x}_J), \dots, F^{-}_{p|J}(u_p|\mathbf{X}_J = \mathbf{x}_J)|\mathbf{X}_J = \mathbf{x}_J\right),$$

for every $\mathbf{u}_I$ and $\mathbf{x}_J$, setting $F_{I|J}(\mathbf{x}_I|\mathbf{X}_J = \mathbf{x}_J) := \mathbb{P}\left(\mathbf{X}_I \leq \mathbf{x}_I|\mathbf{X}_J = \mathbf{x}_J\right)$.

Most often, the dependence of $C_{I|J}(\cdot|\mathbf{X}_J = \mathbf{x}_J)$ w.r.t. to $\mathbf{x}_J$ is a source of significant complexities, in terms of model specification and inference. Therefore, most authors assume the following "simplifying assumption" is fulfilled.

Assumption $(\mathcal{SA})$: the conditional copula $C_{I|J}(\cdot|\mathbf{X}_J = \mathbf{x}_J)$ does not depend on $\mathbf{x}_J$, i.e., for every $\mathbf{u}_I \in [0,1]^p$, the function $\mathbf{x}_J \in \mathbb{R}^{d-p} \mapsto C_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)$ is a constant function (that depends on $\mathbf{u}_I$).

Under $(\mathcal{SA})$, we will set $C_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J) =: C_{s,I|J}(\mathbf{u}_I)$. The latter identity means that the dependence on $\mathbf{X}_J$ across the components of $\mathbf{X}_I$ is passing only through their conditional margins. Note that $C_{s,I|J}$ is different of the usual copula of $\mathbf{X}_I$.

Remark 1. *Note that $(\mathcal{SA})$ does not imply that $C_{s,I|J}(\cdot)$ is $C_I(\cdot)$, the usual copula of $\mathbf{X}_I$. This can be checked with a simple example: let $\mathbf{X} = (X_1, X_2, X_3)$ be a trivariate random vector s.t., given X_3, $X_1 \sim \mathcal{N}(X_3, 1)$ and $X_2 \sim \mathcal{N}(X_3, 1)$. Moreover, X_1 and X_2 are independent given X_3. The latter variable may be $\mathcal{N}(0,1)$, to fix the ideas. Obviously, with our notations, $I = \{1,2\}$, $J = \{3\}$, $d = 3$ and $p = 2$. Therefore, for any couple $(u_1, u_2) \in [0,1]^2$ and any real number x_3, $C_{1,2|3}(u_1, u_2|x_3) = u_1 u_2$ and does not depend on x_3. Assumption $(\mathcal{SA})$ is then satisfied. But the copula of (X_1, X_2) is not the independence copula, simply because X_1 and X_2 are not independent.*

Basically, it is far from obvious to specify and estimate relevant conditional copula models in practice, especially when the conditioning and/or conditioned variables are numerous. Assumption $(\mathcal{SA})$ is particularly relevant with vine models (Aas et al. 2009, among others). Indeed, to build vines from a d-dimensional random vector $\mathbf{X}$, it is necessary to consider sequences of conditional bivariate copulas $C_{I|J}$, where $I = \{i_1, i_2\}$ is a couple of indices in $\{1, \dots, d\}$, $J \subset \{1, \dots, d\}$, $I \cap J = \emptyset$, and $(i_1, i_2|J)$ is a node of the vine. In other words, a bivariate conditional copula is needed at every node of any vine, and the sizes of the conditioning subsets of variables are increasing along the vine. Without additional assumptions, the modelling task becomes rapidly very cumbersome (inference and estimation by maximum likelihood). Therefore, most authors adopt our simplifying assumption $(\mathcal{SA})$ at every node of the vine. Note that the curse of dimensionality still apparently remains because conditional marginal cdfs $F_{k|J}(\cdot|\mathbf{X}_J)$ are invoked with different subsets J of increasing sizes. But this curse can be avoided by calling recursively the parametric copulas that have been estimated before (see Nagler and Czado, 2015).

Nonetheless, Assumption $(\mathcal{SA})$ has appeared to be rather restrictive, even if it may be seen as acceptable for practical reasons and in particular situations. The debate between pro and cons of the simplifying assumption is still largely open, particularly when it is called in some vine models. On one side, Hobæk-Haff et al. (2010) affirm that this $(\mathcal{SA})$ is not only required for fast, flexible, and robust inference, but that it provides "a rather good approximation, even when the simplifying assumption is far from being fulfilled by the actual model". On the other side, Acar et al. (2012) maintain that "this view is too optimistic". They propose a visual test of $(\mathcal{SA})$ when $d = 3$ and in a parametric framework. Their technique was based on local linear approximations and sequential likelihood maximizations. They illustrate the limitations of $(\mathcal{SA})$ by simulation and through real datasets. They note that "an uncritical use of the simplifying assumption may be misleading". Nonetheless, they do not provide formal test procedures. Beside, Acar et al. (2013) have proposed a formal likelihood test of $(\mathcal{SA})$ but when the conditional marginal distributions are known, a rather restrictive situation. Some authors have exhibited classes of parametric distributions for which $(\mathcal{SA})$ is satisfied: see Hobæk-Haff et al. (2010), significantly extended by Stöber et al. (2013). Nonetheless, such families are rather strongly constrained. Therefore, these two papers propose to approximate some conditional copula models by others for which the simplifying assumption is true. This idea has been developed in Spanhel and Kurz (2015) in a vine framework, because they recognize that "it is very unlikely that the unknown data generating process satisfies the simplifying assumption in a strict mathematical sense."

Therefore, there is a need for formal universal tests of $(\mathcal{SA})$. It is likely that the latter assumption is acceptable in some circumstances, whereas it is too rough in others. This means, for given subsets of indices I and J, we would like to test

$$\mathcal{H}_0 : C_{I|J}(\cdot|\mathbf{X}_J = \mathbf{x}_J) \text{ does not depend on } \mathbf{x}_J,$$

against that opposite assumption. Hereafter, we will propose several test statistics of $\mathcal{H}_0$, possibly assuming that the conditional copula belongs to some parametric family.

Note that several papers have already proposed estimators of conditional copula. Veraverbeke, Omelka and Gijbels (2011), Gijbels, Veraverbeke and Omelka (2011) and Fermanian and Wegkamp (2012) have studied some nonparametric kernel based estimators. Craiu and Sabeti (2012), Sabeti, Wei and Craiu (2014) studied bayesian additive models of conditional copulas. Recently, Schellhase and Spanhel (2016) invoke B-splines to manage vectors of conditioning variables. In a semiparametric framework, i.e. assuming an underlying parametric family of conditional copulas, numerous models and estimators have been proposed, notably Acar et al. (2011), Abegaz et al. (2012), Fermanian and Lopez (2015) (single-index type models), Vatter and Chavez-Demoulin (2015) (additive models), among others. But only a few of these papers have a focus on testing $(\mathcal{SA})$ specifically, although convergence of the proposed estimators are necessary to lead such a task in theory. Actually, some tests of $(\mathcal{SA})$ is invoked "in passing" in these papers as potential applications, but without a general approach and/or without some guidelines to evaluate p-values in practice. As an exception, in a very recent paper, Gijbels et al. (2016) have tackled the $(\mathcal{SA})$ directly through comparisons between conditional and unconditional Kendall's tau.

Example 2. *To illustrate the problem, let us consider a simple example of $(\mathcal{SA})$ in dimension 3. Assume that $p = 2$ and $d = 3$. For simplicity, let us assume that (X_1, X_2) follows a Gaussian distribution conditionally to X_3, that is :*

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \Big| X_3 = x_3 \sim \mathcal{N}\left(\begin{pmatrix} \mu_1(x_3) \\ \mu_2(x_3) \end{pmatrix}, \begin{pmatrix} \sigma_1^2(x_3) & \rho(x_3)\sigma_1(x_3)\sigma_2(x_3) \\ \rho(x_3)\sigma_1(x_3)\sigma_2(x_3) & \sigma_2^2(x_3) \end{pmatrix} . \right)$$

Obviously, $\alpha(\cdot) := (\mu_1, \mu_2, \sigma_1, \sigma_2)(\cdot)$ is a parameter that only affects the conditional margins. Moreover, the conditional copula of (X_1, X_2) given $X_3 = x_3$ is gaussian with the parameter $\rho(x_3)$. Six possible cases can then be distinguished:

a. All variables are mutually independent.

b. (X_1, X_2) is independent of X_3, but X_1 and X_2 are not independent.

c. X_1 and X_2 are both marginally independent of X_3, but the conditional copula of X_1 and X_2 depends on X_3.

d. X_1 (or X_2) and X_3 are not independent but X_1 and X_2 are independent conditionally to X_3.

e. X_1 (or X_2) and X_3 are not independent but the conditional copula of X_1 and X_2 is independent of X_3.

f. X_1 *(or* X_2*) and* X_3 *are not independent and the conditional copula of* X_1 *and* X_2 *is dependent of* X_3*.*

These six cases are summarized in the following table:

	$\rho(\cdot) = 0$	$\rho(\cdot) = \rho_0$	$\rho(\cdot)$ *is not constant*
$\alpha(\cdot) = \alpha_0$	*a*	*b*	*c*
$\alpha(\cdot)$ *is not constant*	*d*	*e*	*f*

In general, the simplifying assumption $(\mathcal{SA})$ *consists in assuming that we live in one of the cases* $\{a, b, d, e\}$*, whereas the alternative cases are* c *and* f*.*

Note that, in general, there is no reason why the conditional margins would be constant (and in most applications, they are not). Nevertheless, if we knew the marginal cdfs' were constant with respect to the conditioning variable, then the test of $(\mathcal{SA})$ (i.e. b against c) would become a classical test of independence between $\mathbf{X}_I$ and $\mathbf{X}_J$.

Testing $\mathcal{H}_0$ is closely linked to the m-sample copula problem, for which we have m different and independent samples of a p-dimensional variable $\mathbf{X}_I = (X_1, \ldots, X_p)$. In each sample k, the observations are i.i.d., with their own marginal laws and their own copula $C_{I,k}$. The m-sample copula problem consists on testing whether the m latter copulas $C_{I,k}$ are equal. Note that we could merge all samples into a single one, and create discrete variables Y_i that are equal to k when i lies in the sample k. Therefore, the m-sample copula problem is formally equivalent to testing $\mathcal{H}_0$ with the conditioning variable $\mathbf{X}_J := Y$.

Conversely, assume we have defined a partition $\{A_{1,J}, \ldots, A_{m,J}\}$ of $\mathbb{R}^{d-p}$ composed of borelian subsets such that $\mathbb{P}(\mathbf{X}_J \in A_{k,J}) > 0$ for all $k = 1, \ldots, m$, and we want to test

$$\overline{\mathcal{H}}_0 : k \in \{1, \ldots, m\} \mapsto C_{I|J}(\,\cdot\,|\mathbf{X}_J \in A_{k,J}) \text{ does not depend on } k.$$

Then, divide the sample in m different sub-samples, where any sub-sample k contains the observations for which the conditioning variable belongs to $A_{k,J}$. Then, $\overline{\mathcal{H}}_0$ is equivalent to a m-sample copula problem. Note that $\overline{\mathcal{H}}_0$ looks like a "consequence" of $\mathcal{H}_0$ when it is not the case in general (see Section 3.1), for continuous $\mathbf{X}_J$ variables.

Nonetheless, $\overline{\mathcal{H}}_0$ conveys the same intuition as $\mathcal{H}_0$. Since it can be led more easily in practice (no smoothing is required), some researchers could prefer the former assumption than the latter. That is why it will be discussed hereafter. Note that the 2-sample copula problem has already been addressed by Rémillard and Scaillet (2009), and the m-sample by Bouzebda et al. (2011). However, both paper are designed only in a nonparametric framework, and these authors have not noticed the connection with the simplifying assumption.

The goal of the paper is threefold: first, to write a "state-of-the art" of the simplifying assumption problem; second to propose some "reasonable" test statistics of $(\mathcal{SA})$ in different contexts; third, to introduce a new approach of the latter problem, through "box-related" zero assumptions and some associated test statistics. Since it is impossible to state the theoretical properties of all these test statistics, we will rely on "ad-hoc arguments" to convince the reader they are relevant, without trying to establish specific results. Globally, this paper can be considered also as a work program around $(\mathcal{SA})$ for the next years.

In Section 2, we introduce different ways of testing $(\mathcal{SA})$. We propose different test statistics under a fully nonparametric perspective, i.e. when $C_{I|J}$ is not supposed to belong into a particular parametric copula family, through some comparisons between empirical cdfs' in Subsection 2.1, or by invoking a particular independence property in Subsection 2.2. In Subsection 2.3, new tools are needed if we assume underlying parametric copulas. To evaluate the limiting distributions of such tests, we propose several bootstrap techniques (Subsection 2.4). Section 3 is related to testing $\overline{\mathcal{H}}_0$. In Subsection 3.1, we detail the relations between $\mathcal{H}_0$ and $\overline{\mathcal{H}}_0$. Then, we provide tests statistics of $\overline{\mathcal{H}}_0$ for both the nonparametric (Subsection 3.2) and the parametric framework (Subsection 3.3), as well as bootstrap methods (Subsection 3.4). In particular, we prove the validity of the so-called "parametric independent" bootstrap when testing $\overline{\mathcal{H}}_0$. The performances of the latter tests are assessed and compared by simulation in Section 4. Some of the proofs are collected in Appendix A.

2 Tests of the simplifying assumption

2.1 "Brute-force" tests of $(\mathcal{SA})$

A first natural idea is to build a test of $\mathcal{H}_0$ based on a comparison between some estimates of the conditional copula $C_{I|J}$ with and without the simplifying assumption, for different conditioning events. Such estimates will be called $\hat{C}_{I|J}$ and $\hat{C}_{s,I|J}$ respectively. Then, introducing some distance $\mathcal{D}$ between conditional distributions, a test can be based on the statistics $\mathcal{D}(\hat{C}_{I|J}, \hat{C}_{s,I|J})$. Following most authors, we immediately think of Kolmogorov-Smirnov-type statistics

$$\mathcal{T}^0_{KS,n} := \|\hat{C}_{I|J} - \hat{C}_{s,I|J}\|_\infty = \sup_{\mathbf{u}_I \in [0,1]^p} \sup_{\mathbf{x}_J \in \mathbb{R}^{d-p}} |\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}_J) - \hat{C}_{s,I|J}(\mathbf{u}_I)|, \tag{1}$$

or Cramer von-Mises-type test statistics

$$\mathcal{T}^0_{CvM,n} := \int \left(\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}_J) - \hat{C}_{s,I|J}(\mathbf{u}_I)\right)^2 \, w(d\mathbf{u}_I, d\mathbf{x}_J), \tag{2}$$

for some weight function of bounded variation w, that could be chosen as random (see below).

To evaluate $\hat{C}_{I|J}$, we propose to invoke the nonparametric estimator of conditional copulas proposed by Fermanian and Wegkamp (2012). Alternative kernel-based estimators of conditional copulas can be found in Gijbels et al. (2011), for instance.

Let us start with an iid d-dimensional sample $(\mathbf{X}_i)_{i=1,\dots,n}$. Let $\hat{F}_k$ be the marginal empirical distribution function of X_k, based on the sample $(X_{1,k},\dots,X_{n,k})$, for any $k = 1,\dots,d$. Our estimator of $C_{I|J}$ will be defined as

$$\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J) := \hat{F}_{I|J}\left(\hat{F}^-_{1|J}(u_1|\mathbf{X}_J = \mathbf{x}_J), \dots, \hat{F}^-_{p|J}(u_p|\mathbf{X}_J = \mathbf{x}_J)|\mathbf{X}_J = \mathbf{x}_J\right),$$

$$\hat{F}_{I|J}(\mathbf{x}_I|\mathbf{X}_J = \mathbf{x}_J) := \frac{1}{n}\sum_{i=1}^n K_n(\mathbf{X}_{i,J})\mathbf{1}(\mathbf{X}_{i,I} \le \mathbf{x}_I), \tag{3}$$

where

$$K_n(\mathbf{X}_{i,J}) := K_h\left(\hat{F}_{p+1}(X_{i,p+1}) - \hat{F}_{p+1}(x_{p+1}), \dots, \hat{F}_d(X_{i,d}) - \hat{F}_d(x_d)\right),$$
$$K_h(\mathbf{x}_J) := h^{-(d-p)} K\left(x_{p+1}/h, \dots, x_d/h\right),$$

and K is a $(d-p)$-dimensional kernel. Obviously, for $k \in I$, we have introduced some estimates of the marginal conditional cdfs' similarly:

$$\hat{F}_{k|J}(x|\mathbf{X}_J = \mathbf{x}_J) := \frac{1}{n}\sum_{i=1}^n K_n(\mathbf{X}_{i,J})\mathbf{1}(X_{i,k} \le x). \tag{4}$$

Obviously, $h = h(n)$ is the term of a usual bandwidth sequence, where $h(n) \to 0$ when n tends to the infinity. Since $\hat{F}_{I|J}$ is a nearest-neighbors estimator, it does not necessitate a fine-tuning of local bandwidths (except for those values $\mathbf{x}_J$ s.t. $F_J(\mathbf{x}_J)$ is close to one or zero), contrary to more usual Nadaraya-Watson techniques. In other terms, a single convenient choice of h would provide "satisfying" estimates of $\hat{C}_{I|J}(\mathbf{x}_I|\mathbf{X}_J = \mathbf{x}_J)$ for most values of $\mathbf{x}$.

Remark 3. *The estimated conditional probability $\hat{F}_{I|J}(\cdot|\mathbf{X}_J = \mathbf{x}_J)$ is not a true distribution. Nonetheless, this can be achieved through a minor modification, by replacing $\hat{F}_{I|J}(\mathbf{x}_I|\mathbf{X}_J = \mathbf{x}_J)$ by*

$$\tilde{F}_{I|J}(\mathbf{x}_I|\mathbf{X}_J = \mathbf{x}_J) := \frac{\sum_{i=1}^n K_n(\mathbf{X}_{i,J})\mathbf{1}(\mathbf{X}_{i,I} \le \mathbf{x}_I)}{\sum_{j=1}^n K_n(\mathbf{X}_{j,J})}.$$

Thus, we get similarly $\tilde{F}_{k|J}(\mathbf{x}_I|\mathbf{X}_J = \mathbf{x}_J)$ and another estimate of $C_{I|J}(\cdot|\mathbf{X}_J = \mathbf{x}_J)$.

To calculate the latter statistics (1) and (2), it is necessary to provide an estimate of the underlying conditional copula under $(\mathcal{SA})$. This could be done naively by particularizing a point $\mathbf{x}_J^* \in \mathbb{R}^{d-p}$ and by setting $\hat{C}^{(1)}_{s,I|J}(\cdot) := \hat{C}_{I|J}(\cdot|\mathbf{X}_J = \mathbf{x}_J^*)$. Since the choice of $\mathbf{x}_J^*$ is too arbitrary, an alternative could be to set

$$\hat{C}^{(2)}_{s,I|J}(\cdot) := \int \hat{C}_{I|J}(\cdot|\mathbf{X}_J = \mathbf{x}_J)\, w(d\mathbf{x}_J),$$

for some function w that is of bounded variation, and $\int w(d\mathbf{x}_J) = 1$. Unfortunately, the latter choice induce $(d-p)$-dimensional integration procedures, that becomes a numerical problem rapidly when $d-p$ is larger than three.

Therefore, let us randomize the "weight" functions w, to avoid multiple integrations. For instance, choose the empirical distribution of $\mathbf{X}_J$ as w, providing

$$\hat{C}^{(3)}_{s,I|J}(\cdot) := \int \hat{C}_{I|J}(\cdot|\mathbf{X}_J = \mathbf{x}_J)\, \hat{F}_J(d\mathbf{x}_J) = \frac{1}{n}\sum_{i=1}^n \hat{C}_{I|J}(\cdot|\mathbf{X}_J = \mathbf{X}_{i,J}). \tag{5}$$

An even simpler estimate of $C_{s,I|J}$, the conditional copula of $\mathbf{X}_I$ given $\mathbf{X}_J$ under the simplifying assumption, can be obtained by noting that, under $(\mathcal{SA})$, $C_{s,I|J}$ is the joint law of $\mathbf{Z}_{I|J} := (F_1(X_1|\mathbf{X}_J), \dots, F_p(X_p|\mathbf{X}_J))$ (see Property 6 below). Therefore, it is tempting to estimate $C_{s,I|J}(\mathbf{u}_I)$ by

$$\hat{C}^{(4)}_{s,I|J}(\mathbf{u}_I) := \frac{1}{n}\sum_{i=1}^n \mathbf{1}\left(\hat{F}_{1|J}(X_{i,1}|\mathbf{X}_{i,J}) \le u_1, \dots, \hat{F}_{p|J}(X_{i,p}|\mathbf{X}_{i,J}) \le u_p\right), \tag{6}$$

when $\mathbf{u}_I \in [0,1]^p$, for some consistent estimates $\hat{F}_{k|J}(x_k|\mathbf{x}_J)$ of $F_{k|J}(x_k|\mathbf{x}_J)$. A similar estimator has been promoted and studied in Gijbels et al. (2015) or in Portier and Segers (2015), but they have considered the empirical copula associated to the pseudo sample $((\hat{F}_1(X_{i1}|\mathbf{X}_{iJ}), \dots, \hat{F}_p(X_{ip}|\mathbf{X}_{iJ})))_{i=1,\dots,n}$ instead of its empirical cdf. It will be called $\hat{C}^{(5)}_{s,I|J}$. Hereafter, we will denote $\hat{C}_{s,I|J}$ one of the "averaged" estimators $\hat{C}^{(k)}_{s,I|J}$, $k > 1$ and we can forget the naive pointwise estimator $\hat{C}^{(1)}_{s,I|J}$. Therefore, under some conditions of regularity, we guess that our estimators $\hat{C}_{s,I|J}(\mathbf{u}_I)$ of the conditional copula under $(\mathcal{SA})$ will be $\sqrt{n}$-consistent and asymptotically normal. It has been proved for $C^{(5)}_{s,I|J}$ in Gijbels et al. (2015) or in Portier and Segers (2015), as a byproduct of the weak convergence of the associated process.

For practical reasons, it is important that $\hat{F}_{k|J}(x_k|\mathbf{x}_J)$ belongs to $[0,1]$. Therefore, we will use $\tilde{F}_{k|J}(x_k|\mathbf{x}_J)$ instead in the definitions of $\hat{C}_{I,J}$ and $\hat{C}_{s,I,J}$ above.

Under $\mathcal{H}_0$, we would like that the previous test statistics $\mathcal{T}^0_{KS,n}$ or $\mathcal{T}^0_{CvM,n}$ are convergent. Typically, such a property is given as a sub-product by the weak convergence of a relevant empirical process, here $(\mathbf{u}_I, \mathbf{x}_J) \in [0,1]^p \times \mathbb{R}^{d-p} \mapsto \sqrt{nh_n^{d-p}}(\hat{C}_{I|J} - C_{I|J})(\mathbf{u}_I|\mathbf{x}_J)$. Unfortunately, this will not be the case in general seing the previous process as a function indexed by $\mathbf{x}_J$, at least for wide ranges of bandwidths. Due to the difficulty of checking the tightness of the process indexed by $\mathbf{x}_J$, some alternative techniques may be required as Gaussian approximations (see Chernozhukov et al. 2014, e.g.). Nonetheless, they would lead us far beyond the scope of this paper. Therefore, we simply propose to slightly modifiy the latter test statistics, to manage only a *fixed* set of arguments $\mathbf{x}_J$. For instance, in the case of the Kolmogorov-Smirnov-type test, consider a simple grid $\chi_J := \{\mathbf{x}_{1,J}, \dots, \mathbf{x}_{m,J}\}$, and the modified test statistics

$$\mathcal{T}^{0,m}_{KS,n} := \sup_{\mathbf{u}_I \in [0,1]^p} \sup_{\mathbf{x}_J \in \chi_J} |\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}_J) - \hat{C}_{s,I|J}(\mathbf{u}_I)|.$$

In the case of the Cramer von-Mises-type test, we can approximate any integral by finite sums, possibly after a change of variable to manage a compactly supported integrand. Actually, this is how they are calculated in practice! For instance, invoking Gaussian quadratures, the modified statistics would be

$$\mathcal{T}^{0,m}_{CvM,n} := \sum_{j=1}^{m} \omega_j \left(\hat{C}_{I|J}(\mathbf{u}_{j,I}|\mathbf{x}_{j,J}) - \hat{C}_{s,I|J}(\mathbf{u}_{j,I}) \right)^2 ,$$

for some conveniently chosen constants ω_j, $j = 1, \ldots, m$. Note that the numerical evaluation of $\hat{C}_{I|J}$ is relatively costly. Since quadrature techniques require a lot less points m than "brute-force" equally spaced grids (in dimension d, here), they have to be preferred most often.

Therefore, at least for such modified test statistics, we can insure the tests are convergent. Indeed, under some conditions of regularity, it can be proved that $\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)$ is consistent and asymptotically normal, for every choice of $\mathbf{u}_I$ and $\mathbf{x}_J$ (see Fermanian and Wegkamp, 2012). And a relatively straightforward extension of their Corollary 1 would provide that, under $\mathcal{H}_0$ and for all $\mathcal{U} := (\mathbf{u}_{I,1}, \ldots, \mathbf{u}_{I,q}) \in [0,1]^{p(q+r)}$ and $\mathcal{X} := (\mathbf{x}_{J,1}, \ldots, \mathbf{x}_{J,q}) \in \mathbb{R}^{(d-p)q}$,

$$\Big\{ \sqrt{nh_n^{d-p}}(\hat{C}_{I|J} - C_{s,I|J})(\mathbf{u}_{I,1}|\mathbf{X}_J = \mathbf{x}_{J,1}), \ldots, \sqrt{nh_n^{d-p}}(\hat{C}_{I|J} - C_{s,I|J})(\mathbf{u}_{I,q}|\mathbf{X}_J = \mathbf{x}_{J,q}),$$
$$\sqrt{n}(\hat{C}_{s,I|J} - C_{s,I|J})(\mathbf{u}_{I,q+1}), \ldots, \sqrt{n}(\hat{C}_{s,I|J} - C_{s,I|J})(\mathbf{u}_{I,q+r}) \Big\} ,$$

converges in law towards a Gaussian random vector. As a consequence, $\sqrt{nh_n^{d-p}}\mathcal{T}^{0,m}_{KS,n}$ and $nh_n^{d-p}\mathcal{T}^{0,m}_{CvM,n}$ tends to a complex but not degenerate law under the $\mathcal{H}_0$.

Remark 4. *Other test statistics of $\mathcal{H}_0$ can be obtained by comparing directly the functions $\hat{C}_{I|J}(\cdot|\mathbf{X}_J = \mathbf{x}_J)$, for different values of $\mathbf{x}_J$. For instance, let us define*

$$\tilde{\mathcal{T}}^0_{KS,n} := \sup_{\mathbf{x}_J, \mathbf{x}'_J \in \mathbb{R}^{d-p}} \|\hat{C}_{I|J}(\cdot|\mathbf{x}_J) - \hat{C}_{I|J}(\cdot|\mathbf{x}'_J)\|_\infty$$
$$= \sup_{\mathbf{x}_J, \mathbf{x}'_J \in \mathbb{R}^{d-p}} \sup_{\mathbf{u}_I \in [0,1]^p} |\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}_J) - \hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}'_J)|,$$

or

$$\tilde{\mathcal{T}}^0_{CvM,n} := \int \left(\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}_J) - \hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}'_J) \right)^2 w(d\mathbf{u}_I, d\mathbf{x}_J, d\mathbf{x}'_J),$$

for some function of bounded variation w. As above, modified versions of these statistics can be obtained considering fixed $\mathbf{x}_J$-grids. Since these statistics involve higher dimensional integrals/sums than previously, they will not be studied more in depth. At first glance, their empirical performances are comparable with the previous ones.

Remark 5. *Note that the influence of $\mathbf{x}_J$ on $\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)$ is induced by the quantities $\hat{F}_k(x_k)$ only, $k \in J$, the latter ones taking only the n discrete values l/n, $l = 1, \ldots, n$. Thus, for a given $\mathbf{u}_I$, the constancy of $\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)$ with respect to $\mathbf{x}_J$ is realized if n^p values of this function are equal on a grid, or, equivalently, if*

$$\hat{C}_{I|J} \left(\mathbf{u}_I | \mathbf{X}_J = (X_{k_1,p+1}, \ldots, X_{k_{d-p},d}) \right) = \hat{C}_{I|J} \left(\mathbf{u}_I | \mathbf{X}_J = (X_{k'_1,p+1}, \ldots, X_{k'_{d-p},d}) \right) ,$$

for every choice of indices $k_1, \ldots, k_{d-p}$ and $k'_1, \ldots, k'_{d-p}$ in $\{1, \ldots, n\}$. We denote by $\mathcal{X}_{n,J}$ the set of vectors $(X_{k_1,p+1}, \ldots, X_{k_p,d-d})$, when $k_1, \ldots, k_{d-p}$ belong to $\{1, \ldots, n\}$.

The L^2-type statistics $\mathcal{T}^0_{CvM,n}$ and $\tilde{\mathcal{T}}^0_{CvM,n}$ involve at least d summations or integrals, which can become numerically expensive when the dimension of $\mathbf{X}$ is "large". Nonetheless, we are free to set convenient weight functions. To reduce the computational cost, several versions of $\mathcal{T}^0_{CvM,n}$ are particularly well-suited, by choosing conveniently the functions w. For instance, consider

$$\mathcal{T}^{(1)}_{CvM,n} := \int \left(\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}_J) - \hat{C}_{s,I|J}(\mathbf{u}_I) \right)^2 \hat{C}_I(d\mathbf{u}_I)\, \hat{F}_J(d\mathbf{x}_J),$$

where $\hat{F}_J$ and $\hat{C}_I$ denote the empirical cdf of $(\mathbf{X}_{i,J})$ and the empirical copula of $(\mathbf{X}_{i,I})$ respectively. Therefore, $\mathcal{T}^{(1)}_{CvM,n}$ simply becomes

$$\mathcal{T}^{(1)}_{CvM,n} = \frac{1}{n^2}\sum_{j=1}^n\sum_{i=1}^n \left(\hat{C}_{I|J}(\hat{U}_{i,I}|\mathbf{X}_J = \mathbf{X}_{j,J}) - \hat{C}_{s,I|J}(\hat{U}_{i,I})\right)^2, \tag{7}$$

where $\hat{U}_{i,I} = (\hat{F}_1(X_{i,1}),\ldots,\hat{F}_p(X_{i,p}))$, $i=1,\ldots,n$. Similarly, we can choose

$$\tilde{\mathcal{T}}^{(1)}_{CvM,n} := \int \left(\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}_J) - \hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}'_J)\right)^2 \hat{C}_I(d\mathbf{u}_I)\,\hat{F}_J(d\mathbf{x}_J)\,\hat{F}_J(d\mathbf{x}'_J)$$
$$= \frac{1}{n^3}\sum_{j=1}^n\sum_{j'=1}^n\sum_{i=1}^n \left(\hat{C}_{I|J}(\hat{U}_{i,I}|\mathbf{X}_J = \mathbf{X}_{j,J}) - \hat{C}_{I|J}(\hat{U}_{i,I}|\mathbf{X}_J = \mathbf{X}_{j',J})\right)^2.$$

To deal with a single summations only, it is even possible to propose to set

$$\mathcal{T}^{(2)}_{CvM,n} := \int \Big(\hat{C}_{I|J}(\hat{F}_{1|J}(x_1|\mathbf{x}_J),\ldots,\hat{F}_{p|J}(x_p|\mathbf{x}_J)|\mathbf{x}_J) - \hat{C}_{s,I|J}(\hat{F}_{1|J}(x_1|\mathbf{x}_J),\ldots,\hat{F}_{p|J}(x_p|\mathbf{x}_J))\Big)^2 \hat{F}(d\mathbf{x}_I, d\mathbf{x}_J),$$

where $\hat{F}$ denotes the empirical cdf of $\mathbf{X}$. This means

$$\mathcal{T}^{(2)}_{CvM,n} = \frac{1}{n}\sum_{i=1}^n \Big(\hat{C}_{I|J}\big(\hat{F}_{1|J}(X_{i,1}|\mathbf{X}_{i,J}),\ldots,\hat{F}_{p|J}(X_{i,p}|\mathbf{X}_{i,J})|\mathbf{X}_J = \mathbf{X}_{i,J}\big) - \hat{C}_{s,I|J}\big(\hat{F}_{1|J}(X_{i,1}|\mathbf{X}_{i,J}),\ldots,\hat{F}_{p|J}(X_{i,p}|\mathbf{X}_{i,J})\big)\Big)^2.$$

We have introduced some tests based on comparisons between empirical cdfs'. Obviously, the same idea could be applied to associated densities, as in Fermanian (2005) for instance, or even to other functions of the underlying distributions.

Since the previous test statistics are complicated functionals of some "semi-smoothed" empirical process, it is very challenging to evaluate their asymptotic laws under $\mathcal{H}_0$ analytically. In every case, these limiting laws will not be distribution free, and their calculation would be very tedious. Therefore, as usual with copulas, it is necessary to evaluate the limiting distributions of such tests statistics by a convenient bootstrap procedure (parametric or nonparametric).

2.2 Tests based on the independence property

Actually, testing $\mathcal{H}_0$ is equivalent to a test of the independence between the random vectors $\mathbf{X}_J$ and $\mathbf{Z}_{I|J} := (F_1(X_1|\mathbf{X}_J),\ldots,F_p(X_p|\mathbf{X}_J))$ strictly speaking, as proved in the following proposition.

Proposition 6. *The vectors $\mathbf{Z}_{I|J}$ and $\mathbf{X}_J$ are independent iff $C_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)$ does not depend on $\mathbf{x}_J$ for every vectors $\mathbf{u}_I$ and $\mathbf{x}_J$. In this case, the cdf of $\mathbf{Z}_{I|J}$ is $C_{s,I|J}$.*

Proof: For any vectors $\mathbf{u}_I \in [0,1]^p$ and any subset $A_J \subset \mathbb{R}^{d-p}$,

$$\mathbb{P}(\mathbf{Z}_{I|J} \le \mathbf{u}_I, \mathbf{X}_J \in A_J) = \mathbb{E}\left[\mathbf{1}(\mathbf{X}_J \in A_J)\mathbb{P}(\mathbf{Z}_{I|J} \le \mathbf{u}_I|\mathbf{X}_J)\right]$$
$$= \int \mathbf{1}(\mathbf{x}_J \in A_J)\mathbb{P}(\mathbf{Z}_{I|J} \le \mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)\,d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J)$$
$$= \int_{A_J} \mathbb{P}(F_k(X_k|\mathbf{X}_J = \mathbf{x}_J) \le u_k, \forall k \in I|\mathbf{X}_J = \mathbf{x}_J)\,d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J)$$
$$= \int_{A_J} C_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)\,d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J).$$

If $\mathbf{Z}_{I|J}$ and $\mathbf{X}_J$ are independent, then

$$\mathbb{P}(\mathbf{Z}_{I|J} \leq \mathbf{u}_I)\mathbb{P}(\mathbf{X}_J \in A_J) = \int \mathbf{1}(\mathbf{x}_J \in A_J) C_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)\, d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J),$$

for every $\mathbf{u}_I$ and A_J. This implies $\mathbb{P}(\mathbf{Z}_{I|J} \leq \mathbf{u}_I) = C_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)$ for every $\mathbf{u}_I \in [0,1]^p$ and every $\mathbf{x}_J$ in the support of $\mathbf{X}_J$. This means that $C_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)$ does not depend on $\mathbf{x}_J$, because $\mathbf{Z}_{I|J}$ does not depend on any $\mathbf{x}_J$ by definition.

Reciprocally, under $\mathcal{H}_0$, $C_{s,I|J}$ is the cdf of $\mathbf{Z}_{I|J}$. Indeed,

$$\begin{aligned}
&\mathbb{P}(\mathbf{Z}_{I|J} \leq \mathbf{u}_I) = \mathbb{P}\left(F_k(X_k|\mathbf{X}_J) \leq u_k, \forall k \in I\right) \\
&= \int \mathbb{P}\left(F_k(X_k|\mathbf{X}_J = \mathbf{x}_J) \leq u_k, \forall k \in I|\, \mathbf{X}_J = \mathbf{x}_J\right)\, d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J) \\
&= \int C_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)\, d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J) = \int C_{s,I|J}(\mathbf{u}_I)\, d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J) = C_{s,I|J}(\mathbf{u}_I).
\end{aligned}$$

Moreover, due to Sklar's Theorem, we have

$$\begin{aligned}
\mathbb{P}(\mathbf{Z}_{I|J} \leq \mathbf{u}_I, \mathbf{X}_J \in A_J) &= \int \mathbf{1}(\mathbf{x}_J \in A_J) C_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)\, d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J) \\
&= \textstyle\int \mathbf{1}(\mathbf{x}_J \in A_J) C_{s,I|J}(\mathbf{u}_I)\, d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J) = \mathbb{P}(\mathbf{Z}_{I|J} \leq \mathbf{u}_I)\mathbb{P}(\mathbf{X}_J \in A_J),
\end{aligned}$$

implying the independence between $\mathbf{Z}_{I|J}$ and $\mathbf{X}_J$. □

Then, testing $\mathcal{H}_0$ is formally equivalent to testing

$$\mathcal{H}_0^* : \mathbf{Z}_{I|J} = (F_1(X_1|\mathbf{X}_J), \ldots, F_p(X_p|\mathbf{X}_J)) \text{ and } \mathbf{X}_J \text{ are independent.}$$

Since the conditional marginal cdfs' are not observable, keep in mind that we have to work with pseudo-observations in practice, i.e. vectors of observations that are not independent. In other words, our tests of independence should be based on pseudo-samples

$$\left(\hat{F}_{1|J}(X_{i,1}|\mathbf{X}_{i,J}), \ldots, \hat{F}_{p|J}(X_{i,p}|\mathbf{X}_{i,J})\right)_{i=1,\ldots,n} := (\hat{\mathbf{Z}}_{i,I|J})_{i=1,\ldots,n},$$

for some consistent estimate $\hat{F}_{k|J}(\cdot|\mathbf{X}_J)$, $k \in I$ of the conditional cdfs', for example as defined in Equation (4). The chance of getting distribution-free asymptotic statistics will be very tiny, and we will have to rely on some bootstrap techniques again. To summarize, we should be able to apply some usual tests of independence, but replacing iid observations with (dependent) pseudo-observations.

Most of the tests of $\mathcal{H}_0^*$ rely on the joint law of $(\mathbf{Z}_{I|J}, \mathbf{X}_J)$, that may be evaluated empirically as

$$\begin{aligned}
&G_{I,J}(\mathbf{x}_I, \mathbf{x}_J) := \mathbb{P}(\mathbf{Z}_{I|J} \leq \mathbf{x}_I, \mathbf{X}_J \leq \mathbf{x}_J) \\
&\simeq \quad \hat{G}_{I,J}(\mathbf{x}) := n^{-1} \sum_{i=1}^n \mathbf{1}(\hat{\mathbf{Z}}_{i,I|J} \leq \mathbf{x}_I, \mathbf{X}_{i,J} \leq \mathbf{x}_J).
\end{aligned}$$

Now, let us propose some classical strategies to build independence tests.

- Chi-square-type tests of independence: Let $B_1, \ldots, B_N$ (resp. $A_1, \ldots, A_m$) some disjoint subsets in $\mathbb{R}^p$ (resp. $\mathbb{R}^{d-p}$).

$$\mathcal{I}_{\chi,n} = n \sum_{k=1}^N \sum_{l=1}^m \frac{\left(\hat{G}_{I,J}(B_k \times A_l) - \hat{G}_{I,J}(B_k \times \mathbb{R}^{d-p})\hat{G}_{I,J}(\mathbb{R}^p \times A_l)\right)^2}{\hat{G}_{I,J}(B_k \times \mathbb{R}^{d-p})\hat{G}_{I,J}(\mathbb{R}^p \times A_l)}.$$

- Distance between distributions:

$$\mathcal{I}_{KS,n} = \sup_{\mathbf{x}\in\mathbb{R}^d} |\hat{G}_{I,J}(\mathbf{x}) - \hat{G}_{I,J}(\mathbf{x}_I,\infty^{d-p})\hat{G}_{I,J}(\infty^p,\mathbf{x}_J)|, \text{ or}$$

$$\mathcal{I}_{2,n} = \int \left(\hat{G}_{I,J}(\mathbf{x}) - \hat{G}_{I,J}(\mathbf{x}_I,\infty^{d-p})\hat{G}_{I,J}(\infty^p,\mathbf{x}_J)\right)^2 \omega(\mathbf{x})\, d\mathbf{x},$$

for some (possibly random) weight function ω. Particularly, we can propose the single sum

$$\mathcal{I}_{CvM,n} = \int \left(\hat{G}_{I,J}(\mathbf{x}) - \hat{G}_{I,J}(\mathbf{x}_I,\infty^{d-p})\hat{G}_{I,J}(\infty^p,\mathbf{x}_J)\right)^2 \hat{G}_{I,J}(d\mathbf{x})$$

$$= \frac{1}{n}\sum_{i=1}^n \left(\hat{G}_{I,J}(\hat{\mathbf{Z}}_{i,I|J},\mathbf{X}_{i,J}) - \hat{G}_{I,J}(\hat{\mathbf{Z}}_{i,I|J},\infty^{d-p})\hat{G}_{I,J}(\infty^p,\mathbf{X}_{i,J})\right)^2.$$

- Tests of independence based on comparisons of copulas: let $\breve{C}_{I,J}$ and $\hat{C}_J$ be the empirical copulas based on the pseudo-sample $(\hat{\mathbf{Z}}_{i,I|J},\mathbf{X}_{i,J})_{i=1,\dots,n}$, and $(\mathbf{X}_{i,J})_{i=1,\dots,n}$ respectively. Set

$$\breve{\mathcal{I}}_{KS,n} = \sup_{\mathbf{u}\in[0,1]^d} |\breve{C}_{I,J}(\mathbf{u}) - \hat{C}^{(k)}_{s,I|J}(\mathbf{u}_I)\hat{C}_J(\mathbf{u}_J)|, k=1,\dots,5, \text{ or}$$

$$\breve{\mathcal{I}}_{2,n} = \int_{\mathbf{u}\in[0,1]^d} \left(\breve{C}_{I,J}(\mathbf{u}) - \hat{C}^{(k)}_{s,I|J}(\mathbf{u}_I)\hat{C}_J(\mathbf{u}_J)\right)^2 \omega(\mathbf{u})\, d\mathbf{u},$$

and in particular

$$\breve{\mathcal{I}}_{CvM,n} = \int_{\mathbf{u}\in[0,1]^d} \left(\breve{C}_{I,J}(\mathbf{u}) - \hat{C}^{(k)}_{s,I|J}(\mathbf{u}_I)\hat{C}_J(\mathbf{u}_J)\right)^2 \breve{C}_{I,J}(d\mathbf{u}).$$

The underlying ideas of the test statistics $\breve{\mathcal{I}}_{KS,n}$ and $\breve{\mathcal{I}}_{CvM,n}$ are similar to those that have been proposed by Deheuvels (1979,1981) in the case of unconditional copulas. Nonetheless, in our case, we have to calculate pseudo-samples of the pseudo-observations $(\hat{\mathbf{Z}}_{i,I|J})$ and $(\mathbf{X}_{i,J})$, instead of a usual pseudo-sample of $(\mathbf{X}_i)$.

Note that the latter techniques require the evaluation of some conditional distributions, for instance by kernel smoothing. Therefore, the level of numerical complexity of these test statistics of $\mathcal{H}_0^*$ is comparable with those we have proposed before to test $\mathcal{H}_0$ directly.

2.3 Parametric tests of $(\mathcal{SA})$

In practice, modelers often assume a priori that the underlying copulas belong to some specified parametric family $\mathcal{C} := \{C_\theta, \theta \in \Theta \subset \mathbb{R}^m\}$. Let us adapt our tests under this parametric assumption. Apparently, we would like to test

$$\breve{\mathcal{H}}_0 : C_{I|J}(\cdot|\mathbf{X}_J) = C_\theta(\cdot), \text{ for some } \theta\in\Theta \text{ and almost every } \mathbf{X}_J.$$

Actually, $\breve{\mathcal{H}}_0$ requires two different things: the fact that the conditional copula is a constant copula w.r.t. its conditioning events (test of $(\mathcal{SA})$) and, additionally, that the right copula belongs to $\mathcal{C}$ (classical composite Goodness-of-Fit test). Under this point of view, we would have to adapt "omnibus" specification tests to manage conditional copulas and pseudo observations. For instance, and among of alternatives, we could consider an amended version of Andrews (1997)'s specification test

$$CK_n := \frac{1}{\sqrt{n}} \max_{j\le n} |\sum_{i=1}^n \left[\mathbf{1}(\hat{\mathbf{Z}}_{i,I|J} \le \hat{\mathbf{Z}}_{j,I|J}) - C_{\hat{\theta}_0}(\hat{\mathbf{Z}}_{j,I|J})\right] \mathbf{1}(\mathbf{X}_{i,J}\le \mathbf{X}_{j,J})|,$$

recalling the notations in (2.2). For other ideas of the same type, see Zheng (2000) and the references therein.

The latter global approach is probably too demanding. Here, we prefer to isolate the initial problem that was related to $(\mathcal{SA})$ only. Therefore, let us assume that, for every $\mathbf{x}_J$, there exists a parameter $\theta(\mathbf{x}_J)$ such that $C_{I|J}(\cdot|\mathbf{x}_J) = C_{\theta(\mathbf{x}_J)}(\cdot)$. To simplify, we assume the function $\theta(\cdot)$ is continuous. Our problem is then reduced to testing the constancy of θ, i.e.

$$\mathcal{H}_0^c : \text{the function } \mathbf{x}_J \mapsto \theta(\mathbf{x}_J) \text{ is a constant, called } \theta_0.$$

For every $\mathbf{x}_J$, assume we estimate $\theta(\mathbf{x}_J)$ consistently. For instance, this can be done by modifying the standard semiparametric Canonical Maximum Likelihood methodology (Genest et al., 1995, Tsukahara, 2005): set

$$\hat{\theta}(\mathbf{x}_J) := \arg\max_{\theta\in\Theta} \sum_{i=1}^n \log c_\theta \left(\hat{F}_1(X_{i,1}|\mathbf{X}_{i,J}), \ldots, \hat{F}_p(X_{i,p}|\mathbf{X}_{i,J})\right) \cdot K_h(\mathbf{x}_J - \mathbf{X}_{i,J}),$$

through usual kernel smoothing in $\mathbb{R}^{d-p}$. Alternatively, we could consider

$$\tilde{\theta}(\mathbf{x}_J) := \arg\max_{\theta\in\Theta} \sum_{i=1}^n \log c_\theta \left(\hat{F}_{1|J}(X_{i,1}|\mathbf{X}_J = \mathbf{x}_J), \ldots, \hat{F}_{p|J}(X_{i,p}|\mathbf{X}_J = \mathbf{x}_J)\right) \cdot K_h(\mathbf{x}_J - \mathbf{X}_{i,J}),$$

instead of $\hat{\theta}(\mathbf{x}_J)$. See Abegaz et al. (2012) concerning the theoretical properties of $\tilde{\theta}(\mathbf{x}_J)$ and some choice of conditional cdfs'. Those of $\hat{\theta}(\mathbf{x}_J)$ remain to be stated precisely, to the best of our knowledge. But there is no doubt both methodologies provide consistent estimators, even jointly, under some conditions of regularity.

Under $\mathcal{H}_0^c$, the natural "unconditional" copula parameter θ_0 of the copula of the $\mathbf{Z}_{I|J}$ will be estimated by

$$\hat{\theta}_0 := \arg\max_{\theta\in\Theta} \sum_{i=1}^n \log c_\theta \left(\hat{F}_{1|J}(X_{i,1}|\mathbf{X}_{i,J}), \ldots, \hat{F}_{p|J}(X_{i,p}|\mathbf{X}_{i,J})\right). \tag{8}$$

Surprisingly, the theoretical properties of the latter estimator do not seem to have been established in the literature explicitly. Nonetheless, the latter M-estimator is a particular case of those considered in Fermanian and Lopez (2015) in the framework of single-index models when the link function is a known function (that does not depend on the index). Therefore, by adapting their assumption in the current framework, we easily obtain that $\hat{\theta}_0$ is consistent and asymptotically normal if c_θ is sufficiently regular, for convenient choices of bandwidths and kernels.

Now, there are some challengers to test $\mathcal{H}_0^c$:

- Tests based on the comparison between $\hat{\theta}(\cdot)$ and $\hat{\theta}_0$:

  $$\mathcal{T}_\infty^c := \sup_{\mathbf{x}_J\in\mathbb{R}^{d-p}} \|\hat{\theta}(\mathbf{x}_J) - \hat{\theta}_0\|, \text{ or } \mathcal{T}_2^c := \int \|\hat{\theta}(\mathbf{x}_J) - \hat{\theta}_0\|^2 \omega(\mathbf{x}_J)\, d\mathbf{x}_J,$$

  for some weight function ω.

- Tests based on the comparison between $C_{\hat{\theta}(\cdot)}$ and $C_{\hat{\theta}_0}$:

  $$\mathcal{T}_{dist}^c := \int dist\left(C_{\hat{\theta}(\mathbf{x}_J)}, C_{\hat{\theta}_0}\right) \omega(\mathbf{x}_J)\, d\mathbf{x}_J,$$

  for some distance $dist(\cdot,\cdot)$ between cdfs'.

- Tests based on the comparison between copula densities (when they exist):

  $$\mathcal{T}_{dens}^c := \int \left(c_{\hat{\theta}(\mathbf{x}_J)}(\mathbf{u}_I) - c_{\hat{\theta}_0}(\mathbf{u}_I)\right)^2 \omega(\mathbf{u}_I, \mathbf{x}_J)\, d\mathbf{u}_I\, d\mathbf{x}_J.$$

Remark 7. *It might be difficult to compute some of these integrals numerically, because of unbounded supports. One solution is to to make change of variables. For example,*

$$\mathcal{T}_2^c = \int \|\hat{\theta}(F_J^-(\mathbf{u}_J)) - \hat{\theta}_0\|^2 \omega(F_J^-(\mathbf{u}_J)) \frac{d\mathbf{u}_J}{f_J(F_J^-(\mathbf{u}_J))}.$$

Therefore, the choice $\omega = f_J$ allows us to simplify the latter statistics to $\int \|\hat{\theta}(F_J^-(\mathbf{u}_J)) - \hat{\theta}_0\|^2 d\mathbf{u}_J$, which is rather easy to evaluate. We used this trick in the numerical section below.

2.4 Bootstrap techniques for tests of $(\mathcal{SA})$

It is necessary to evaluate the limiting laws of the latter test statistics under the null. As a matter of fact, we generally cannot exhibit explicit - and distribution-free a fortiori - expressions for these limiting laws. The common technique is provided by bootstrap resampling schemes.

More precisely, let us consider a general statistics $\mathcal{T}$, built from the initial sample $\mathcal{S} := (\mathbf{X}_1, \ldots, \mathbf{X}_n)$. The main idea of the bootstrap is to construct N new samples $\mathcal{S}^* := (\mathbf{X}_1^*, \ldots, \mathbf{X}_n^*)$ following a given resampling scheme given $\mathcal{S}$. Then, for each bootstrap sample $\mathcal{S}^*$, we will evaluate a bootstrapped test statistics $\mathcal{T}^*$, and the empirical law of all these N statistics is used as an approximation of the limiting law of the initial statistic $\mathcal{T}$.

2.4.1 Some resampling schemes

The first natural idea is to invoke Efron's usual "nonparametric bootstrap", where we draw independently with replacement $\mathbf{X}_i^*$ for $i = 1, \ldots, n$ among the initial sample $\mathcal{S} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$. This provides a bootstrap sample $\mathcal{S}^* := (\mathbf{X}_1^*, \ldots, \mathbf{X}_n^*)$.

The nonparametric bootstrap is an "omnibus" procedure whose theoretical properties are well-known but that may not be particularly adapted to the problem at hand. Therefore, we will propose alternative sampling schemes that should be of interest, even if we do not state their validity on the theoretical basis. Such a task is left for further researches.

An natural idea would be to use some properties of $\mathbf{X}$ under $\mathcal{H}_0$, in particular the characterization given in Proposition 6: under $\mathcal{H}_0$, we known that $\mathbf{Z}_{i,I|J}$ and $\mathbf{X}_{i,J}$ are independent. This will be only relevant for the tests of Subsection 2.2, and for a few tests of Subsection 2.1, where such statistics are based on the pseudo-sample $(\hat{\mathbf{Z}}_{i,I|J}, \mathbf{X}_{i,J})_{i=1,\ldots,n}$. Therefore, we propose the following so-called "pseudo-independent bootstrap" scheme:

Repeat, for $i = 1$ to n,

1. draw $\mathbf{X}_{i,J}^*$ among $(\mathbf{X}_{j,J})_{j=1,\ldots,n}$;
2. draw $\hat{\mathbf{Z}}_{i,I|J}^*$ independently, among the observations $\hat{\mathbf{Z}}_{j,I|J}$, $j = 1, \ldots, n$.

This provides a bootstrap sample $\mathcal{S}^* := \big((\hat{\mathbf{Z}}_{1,I|J}^*, \mathbf{X}_{1,J}^*), \ldots, (\hat{\mathbf{Z}}_{n,I|J}^*, \mathbf{X}_{n,J}^*)\big)$.

Note that we could invoke the same idea, but with a usual nonparametric bootstrap perspective: draw with replacement a n-sample among the pseudo-observations $(\hat{\mathbf{Z}}_{i,I|J}, \mathbf{X}_{i,J})_{i=1,\ldots,n}$ for each bootstrap sample. This can be called a "pseudo-nonparametric bootstrap" scheme.

Moreover, note that we cannot draw independently $\mathbf{X}_{i,J}^*$ among $(\mathbf{X}_{j,J})_{j=1,\ldots,n}$, and beside $\mathbf{X}_{i,I}^*$ among $(\mathbf{X}_{j,I})_{j=1,\ldots,n}$ independently. Indeed, $\mathcal{H}_0$ does not imply the independence between $\mathbf{X}_I$ and $\mathbf{X}_J$. At the opposite, it makes sense to build a "conditional bootstrap" as follows:

Repeat, for $i = 1$ to n,

1. draw $\mathbf{X}_{i,J}^*$ among $(\mathbf{X}_{j,J})_{j=1,\ldots,n}$;
2. draw $\hat{\mathbf{X}}_{i,I}^*$ independently, along the estimated conditional law of $\mathbf{X}_I$ given $\mathbf{X}_J = \mathbf{X}_{i,J}^*$. This can be down by drawing a realization along the law $\hat{F}_{I|J}(\cdot|\mathbf{X}_J = \mathbf{X}_{i,J}^*)$, for instance (see (3)). This can be done easily because the latter law is purely discrete, with unequal weights that depend on $\mathbf{X}_{i,J}^*$ and $\mathcal{S}$.

This provides a bootstrap sample $\mathcal{S}^* := \big((\hat{\mathbf{X}}^*_{1,I}, \mathbf{X}^*_{1,J}), \dots, (\hat{\mathbf{X}}^*_{n,I}, \mathbf{X}^*_{n,J})\big)$.

Remark 8. *Note that the latter way of resampling is not far from the usual nonparametric bootstrap. Indeed, when the bandwidths tend to zero, once $\mathbf{x}^*_J = \mathbf{X}_{i,J}$ is drawn, the procedure above will select the other components of $\mathbf{X}_i$ (or close values), i.e. the probability that $\mathbf{x}^*_I = \mathbf{X}_{i,I}$ is "high".*

In the parametric framework, we might also want to use an appropriate resampling scheme. As a matter of fact, all the previous resampling schemes can be used, as in the nonparametric framework, but we would not take advantage of the parametric hypothesis, i.e. the fact that all conditional copulas belong to a known family. We have also to keep in mind that even if the conditional copula has a parametric form, the global model is not fully parametric, because we have not provided a parametric model neither for the conditional marginal cdfs $F_{k|J}$, $k = 1, \dots, p$, nor for the cdf of $\mathbf{X}_J$.

Therefore, we can invoke the null hypothesis $\mathcal{H}_0^c$ and approximate the real copula C_{θ_0} of $\mathbf{Z}_{I|J}$ by $C_{\hat{\theta}_0}$. This leads us to define the following "parametric independent bootstrap":

Repeat, for $i = 1$ to n,

1. draw $\mathbf{X}^*_{i,J}$ among $(\mathbf{X}_{j,J})_{j=1,\dots,n}$;
2. sample $\mathbf{Z}^*_{i,I|J,\hat{\theta}_0}$ from the copula with parameter $\hat{\theta}_0$ independently.

This provides a bootstrap sample $\mathcal{S}^* := \big((\mathbf{Z}^*_{1,I|J,\hat{\theta}_0}, \mathbf{X}^*_{1,J}), \dots, (\mathbf{Z}^*_{n,I|J,\hat{\theta}_0}, \mathbf{X}^*_{n,J})\big)$.

Remark 9. *At first sight, this might seem like a strange mixing of parametric and nonparametric bootstrap. If $|J| = 1$, we can nonetheless do a "full parametric bootstrap", by observing that all estimators of our previous test statistics do not depend on $\mathbf{X}_J$, but on realizations of $\hat{F}_J(\mathbf{X}_J)$ (see Equations (3) and (4)). Since the law of latter variable is close to a uniform distribution, it is tempting to sample $V^*_{i,J} \sim \mathcal{U}_{[0,1]}$ at the first stage, $i = 1, \dots, n$, and then to replace $\hat{F}_J(\mathbf{X}_{i,J})$ with $V^*_{i,J}$ to get an alternative bootstrap sample.*

Without using $\mathcal{H}_0^c$, we could define the "parametric conditional bootstrap" as:

Repeat, for $i = 1$ to n,

- draw $\mathbf{X}^*_{i,J}$ among $(\mathbf{X}_{j,J})_{j=1,\dots,n}$;
- sample $\mathbf{Z}^*_{i,I|J,\theta^*_i}$ from the copula with parameter $\hat{\theta}(\mathbf{X}^*_{i,J})$.

This provides a bootstrap sample $\mathcal{S}^* := \big((\mathbf{Z}^*_{1,I|J,\theta^*_i}, \mathbf{X}^*_{1,J}), \dots, (\mathbf{Z}^*_{n,I|J,\theta^*_i}, \mathbf{X}^*_{n,J})\big)$.

Note that, in several resampling schemes, we should be able to keep the same $\mathbf{X}_J$ as in the original sample, and simulate only $\mathbf{Z}^*_{i,I|J}$ in step 2, as in Andrews(1997), pages 10-11. Such an idea has been proposed by Omelka et al. (2013), in a slightly different framework and univariate conditioning variables. They proved that such a bootstrap scheme "works", after a fine-tuning of different smoothing parameters: see their Theorem 1.

2.4.2 Bootstrapped test statistics

The problem is now to evaluate the law of a given test statistic, say $\mathcal{T}$, under $\mathcal{H}_0$ by the some bootstrap techniques. We recall the main technique in the case of the classical nonparametric bootstrap. We conjecture that the idea is still theoretically sound under the other resampling schemes that have been proposed in Subsection 2.4.1.

The principle for the nonparametric bootstrap is based on the weak convergence of the underlying empirical process. Formally, if $\mathcal{S} := \{\mathbf{X}_1, \dots, \mathbf{X}_n\}$ in an iid sample in $\mathbb{R}^d$, $\mathbf{X} \sim F$ and if F_n denotes its empirical distribution, it is well-known that $\sqrt{n}\,(F_n - F)$ tends weakly in ℓ^∞ towards a d-dimensional Brownian bridge $\mathbb{B}_F$. And the nonparametric bootstrap works in the sense that $\sqrt{n}\,(F^*_n - F_n)$ converges weakly towards a process $\mathbb{B}'_F$, an independent version of $\mathbb{B}_F$, given the initial sample $\mathcal{S}$.

Due to the Delta Method, for every Hadamard-differentiable functional χ from $\ell^\infty(\mathbb{R}^d)$ to $\mathbb{R}$, there exists a random variable H_χ s.t. $\sqrt{n}\,(\chi(F_n)-\chi(F)) \Rightarrow H_\chi$. Assume a test statistics $\mathcal{T}_n$ of $\mathcal{H}_0$ can be written as a sufficiently regular functional of the underlying empirical process as

$$\mathcal{T}_n := \psi\left(\sqrt{n}\,(\chi_s(F_n)-\chi(F_n))\right),$$

where $\chi_s(F) = \chi(F)$ under the null assumption. Then, under $\mathcal{H}_0$, we can rewrite this expression as

$$\mathcal{T}_n := \psi\left(\sqrt{n}\,(\chi_s(F_n)-\chi_s(F)+\chi(F)-\chi(F_n))\right). \tag{9}$$

Given any bootstrap sample $\mathcal{S}^*$ and the associated empirical distribution F_n^*, the usual bootstrap equivalent of $\mathcal{T}_n$ is

$$\mathcal{T}_n^* := \psi\left(\sqrt{n}\,(\chi_s(F_n^*)-\chi_s(F_n)+\chi(F_n)-\chi(F_n^*))\right),$$

from Equation (9). See van der Vaart and Wellner (1996), Section 3.9, for details and mathematically sound statements.

Applying these ideas, we can guess the bootstrapped statistics corresponding to the tests statistics of $\mathcal{H}_0$, at least when the usual nonparametric bootstrap is invoked.

Let us illustrate the idea with $\mathcal{T}^0_{KM,n}$. Note that $\hat{C}_{I|J}(\cdot|\mathbf{X}_J=\cdot) = \chi_{KM}(F_n)(\cdot)$ and $\hat{C}_{s,I|J} = \chi_{s,KM}(F_n)$ for some smoothed functional χ_{KM} and $\chi_{s,KM}$. Under $\mathcal{H}_0$, $\chi_{KM} = \chi_{s,KM}$ and $\mathcal{T}^0_{KS,n} := \|\chi_{KM}(F_n)-\chi_{KM}(F)-\chi_{s,KM}(F_n)+\chi_{s,KM}(F)\|_\infty$. Therefore, its bootstrapped version is

$$\begin{aligned}\mathcal{T}^{0,*}_{KS,n} &:= \|\chi_{KM}(F_n^*)-\chi_{KM}(F_n)-\chi_{s,KM}(F_n^*)+\chi_{s,KM}(F_n)\|_\infty \\ &= \|\hat{C}^*_{I|J}-\hat{C}_{I|J}-\hat{C}^*_{s,I|J}+\hat{C}_{s,I|J}\|_\infty.\end{aligned}$$

Obviously, the functions $\hat{C}^*_{I|J}$ and $\hat{C}^*_{s,I|J}$ have been calculated as $\hat{C}_{I|J}$ and $\hat{C}_{s,I|J}$ respectively, but replacing $\mathcal{S}$ by $\mathcal{S}^*$. Similarly, the bootstrapped versions of some Cramer von-Mises-type test statistics are

$$\mathcal{T}^{0,*}_{CvM,n} := \int \left(\hat{C}^*_{I|J}(\mathbf{u}_I|\mathbf{x}_J)-\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}_J)-\hat{C}^*_{s,I|J}(\mathbf{u}_I)+\hat{C}_{s,I|J}(\mathbf{u}_I)\right)^2 w(d\mathbf{u}_I,d\mathbf{x}_J).$$

When playing with the weight functions w, it is possible to keep the same weights for the bootstrapped versions, or to replace them with some functionals of F_n^*. For instance, asymptotically, it is equivalent to consider

$$\mathcal{T}^{(1),*}_{CvM,n} := \int \left(\hat{C}^*_{I|J}(\mathbf{u}_I|\mathbf{x}_J)-\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}_J)-\hat{C}^*_{s,I|J}(\mathbf{u}_I)+\hat{C}_{s,I|J}(\mathbf{u}_I)\right)^2 \hat{C}_n(d\mathbf{u}_I)\,\hat{F}_J(d\mathbf{x}_J),\ \text{or}$$

$$\mathcal{T}^{(1),*}_{CvM,n} := \int \left(\hat{C}^*_{I|J}(\mathbf{u}_I|\mathbf{x}_J)-\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{x}_J)-\hat{C}^*_{s,I|J}(\mathbf{u}_I)+\hat{C}_{s,I|J}(\mathbf{u}_I)\right)^2 \hat{C}^*_n(d\mathbf{u}_I)\,\hat{F}^*_J(d\mathbf{x}_J).$$

Similarly, the limiting law of

$$\begin{aligned}\mathcal{T}^{(2),*}_{CvM,n} &:= \int \Big(\hat{C}^*_{I|J}(\hat{F}^*_{n,1}(x_1|\mathbf{x}_J),\dots,\hat{F}^*_{n,p}(x_p|\mathbf{x}_J)|\mathbf{x}_J) \\ &- \hat{C}_{I|J}(\hat{F}^*_{n,1}(x_1|\mathbf{x}_J),\dots,\hat{F}^*_{n,p}(x_p|\mathbf{x}_J)|\mathbf{x}_J) - \hat{C}^*_{s,I|J}(\hat{F}^*_{n,1}(x_1|\mathbf{x}_J),\dots,\hat{F}^*_{n,p}(x_p|\mathbf{x}_J)) \\ &+ \hat{C}_{s,I|J}(\hat{F}^*_{n,1}(x_1|\mathbf{x}_J),\dots,\hat{F}^*_{n,p}(x_p|\mathbf{x}_J))\Big)^2 H_n(d\mathbf{x}_I,d\mathbf{x}_J),\end{aligned}$$

given F_n is unchanged replacing H_n by H_n^*.

The same ideas apply concerning the tests of Subsection 2.2, but they require some modifications. Let H be some cdf on $\mathbb{R}^d$. Denote by H_I and H_J the associated cdf on the first p and $d-p$ components respectively. Denote by $\hat{H}$, $\hat{H}_I$ and $\hat{H}_J$ their empirical counterparts. Under $\mathcal{H}_0$, and for any measurable subsets B_I and A_J, $H(B_I\times A_J) = H(B_I)H(A_J)$. Our tests will be based on the difference

$$\begin{aligned}\hat{H}(B_I\times A_J)-\hat{H}_I(B_I)\hat{H}_J(A_J) &= (\hat{H}-H)(B_I\times A_J) \\ &- (\hat{H}_I-H_I)(B_I)\hat{H}_J(A_J)-(\hat{H}_J-H_J)(A_J)H_I(B_I).\end{aligned}$$

Therefore, a bootstrapped approximation of the latter quantity will be

$$(\hat{H}^* - \hat{H})(B_I \times A_J) - (\hat{H}_I^* - \hat{H}_I)(B_I)\hat{H}_J^*(A_J) - (\hat{H}_J^* - \hat{H}_J)(A_J)\hat{H}_I(B_I).$$

To be specific, the bootstrapped versions of our tests are specified as below.

- Chi-square-type test of independence:

$$\mathcal{I}_{\chi,n}^* := n \sum_{k=1}^{N}\sum_{l=1}^{m} \frac{1}{\hat{G}_{I,J}^*(B_k \times \mathbb{R}^{d-p})\hat{G}_{I,J}^*(\mathbb{R}^p \times A_l)} \big((\hat{G}_{I,J}^* - \hat{G}_{I,J})(B_k \times A_l) - \hat{G}_{I,J}^*(B_k \times \mathbb{R}^{d-p})\hat{G}_{I,J}^*(\mathbb{R}^p \times A_l) + \hat{G}_{I,J}(B_k \times \mathbb{R}^{d-p})\hat{G}_{I,J}(\mathbb{R}^p \times A_l)\big)^2.$$

- Distance between distributions:

$$\mathcal{I}_{KS,n}^* = \sup_{\mathbf{x}\in\mathbb{R}^d} |(\hat{G}_{I,J}^* - \hat{G}_{I,J})(\mathbf{x}) - \hat{G}_{I,J}^*(\mathbf{x}_I, \infty^{d-p})\hat{G}_{I,J}^*(\infty^p, \mathbf{x}_J) + \hat{G}_{I,J}(\mathbf{x}_I, \infty^{d-p})\hat{G}_{I,J}(\infty^p, \mathbf{x}_J)|$$

$$\mathcal{I}_{2,n}^* = \int \Big((\hat{G}_{I,J}^* - \hat{G}_{I,J})(\mathbf{x}) - \hat{G}_{I,J}^*(\mathbf{x}_I, \infty^{d-p})\hat{G}_{I,J}^*(\infty^p, \mathbf{x}_J) + \hat{G}_{I,J}(\mathbf{x}_I, \infty^{d-p})\hat{G}_{I,J}(\infty^p, \mathbf{x}_J) \Big)^2 \omega(\mathbf{x})\, d\mathbf{x},$$

  and $\mathcal{I}_{CvM,n}^*$ is obtained replacing $\omega(\mathbf{x})\, d\mathbf{x}$ by $\hat{G}_{I,J}^*(d\mathbf{x})$ (or even $\hat{G}_{I,J}(d\mathbf{x})$).

- A test of independence based on the independence copula: Let $\breve{C}_{I,J}^*$, $\breve{C}_{I|J}^*$ and $\hat{C}_J^*$ be the empirical copulas based on a bootstrapped version of the pseudo-sample $(\hat{\mathbf{Z}}_{i,I|J}, \mathbf{X}_{i,J})_{i=1,\dots,n}$, $(\hat{\mathbf{Z}}_{i,I|J})_{i=1,\dots,n}$ and $(\mathbf{X}_{i,J})_{i=1,\dots,n}$ respectively. This version can be obtained by nonparametric bootstrap, as usual, providing new vectors $\hat{\mathbf{Z}}_{i,I|J}^*$ at every draw. The associated bootstrapped statistics are

$$\breve{\mathcal{I}}_{KS,n}^* = \sup_{\mathbf{u}\in[0,1]^d} |(\breve{C}_{I,J}^* - \breve{C}_{I,J})(\mathbf{u}) - \breve{C}_{I|J}^*(\mathbf{u}_I)\hat{C}_J^*(\mathbf{u}_J) + \breve{C}_{I|J}(\mathbf{u}_I)\hat{C}_J(\mathbf{u}_J)|,$$

$$\breve{\mathcal{I}}_{2,n}^* = \int_{\mathbf{u}\in[0,1]^d} \big((\breve{C}_{I,J}^* - \breve{C}_{I,J})(\mathbf{u}) - \breve{C}_{I|J}^*(\mathbf{u}_I)\hat{C}_J^*(\mathbf{u}_J) + \breve{C}_{I|J}(\mathbf{u}_I)\hat{C}_J(\mathbf{u}_J)\big)^2 \omega(\mathbf{u})\, d\mathbf{u},$$

$$\breve{\mathcal{I}}_{CvM,n}^* = \int_{\mathbf{u}\in[0,1]^d} \big((\breve{C}_{I,J}^* - \breve{C}_{I,J})(\mathbf{u}) - \breve{C}_{I|J}^*(\mathbf{u}_I)\hat{C}_J^*(\mathbf{u}_J) + \breve{C}_{I|J}(\mathbf{u}_I)\hat{C}_J(\mathbf{u}_J)\big)^2\, \breve{C}_{I,J}^*(d\mathbf{u}).$$

In the case of the parametric statistics, the situation is pretty much the same, as long as we invoke the nonparametric bootstrap. For instance, the bootstrapped versions of some previous test statistics are

$$(\mathcal{T}_2^c)^* := \int \|\hat{\theta}^*(\mathbf{x}_J) - \hat{\theta}(\mathbf{x}_J) - \hat{\theta}_0^* + \hat{\theta}_0\|^2 \omega(\mathbf{x}_J)\, d\mathbf{x}_J, \text{ or}$$

$$(\mathcal{T}_{dens}^c)^* := \int \Big(c_{\hat{\theta}^*(\mathbf{x}_J)}(\mathbf{u}_I) - c_{\hat{\theta}(\mathbf{x}_J)}(\mathbf{u}_I) - c_{\hat{\theta}_0^*}(\mathbf{u}_I) + c_{\hat{\theta}_0}(\mathbf{u}_I)\Big)^2 \omega(\mathbf{u}_I, \mathbf{x}_J) d\mathbf{u}_I\, d\mathbf{x}_J.$$

in the case of the nonparametric bootstrap. We conjecture that the previous techniques can be applied with the other resampling schemes that have been proposed in Subsection 2.4.1. Nonetheless, a complete theoretical study of all these alternative schemes and the statement of the validity of their associated bootstrapped statistics is beyond the scope of this paper.

Remark 10. *For the "parametric independent" bootstrap scheme, we have observed that the test powers are a lot better by considering*

$$(\mathcal{T}_2^c)^{**} := \int \|\hat{\theta}^*(\mathbf{x}_J) - \hat{\theta}_0^*\|^2 \omega(\mathbf{x}_J)\, d\mathbf{x}_J, \text{ or}$$

$$(\mathcal{T}_{dens}^c)^{**} := \int \Big(c_{\hat{\theta}^*(\mathbf{x}_J)}(\mathbf{u}_I) - c_{\hat{\theta}_0^*}(\mathbf{u}_I)\Big)^2 \omega(\mathbf{u}_I, \mathbf{x}_J) d\mathbf{u}_I\, d\mathbf{x}_J,$$

instead. The relevance of such statistics may be theoretically justified in the slightly different context of "box-type" tests in the next Section (see Theorem 14). Since our present case is close to the situation of "many small boxes", it is not surprising that we observe similar features. Note that, contrary to the nonparametric bootstrap or the "parametric conditional" bootstrap, the "parametric independent" bootstrap scheme uses $\mathcal{H}_0$.

3 Tests with “boxes”

3.1 The link with the simplifying assumption

As we have seen in Remark 1, we do not have $C_{s,I|J} = C_I$ in general, when $C_I(\mathbf{u}_I) = C_{I|J}(\mathbf{u}_I|\mathbf{X}_J \in \mathbb{R}^{d-p})$ for every $\mathbf{u}_I$. This is the hint there are some subtle relations between conditional copulas when the conditioning event is pointwise or when it is a measurable subset. Actually, to test $\mathcal{H}_0$ in Section 2, we have relied on kernel estimates and smoothing parameters, at least to evaluate conditional marginal distributions empirically. To avoid the curse of dimension (when $d-p$ is “large” i.e. larger than three in practice), it is tempting to replace the pointwise conditioning events $\mathbf{X}_J = \mathbf{x}_J$ with $\mathbf{X}_J \in A_J$ for some borelian subsets $A_J \subset \mathbb{R}^{d-p}$, $\mathbb{P}(\mathbf{X}_J \in A_J) > 0$. As a shorthand notation, we shall write $\mathcal{A}_J$ the set of all such A_J. We call them “boxes” because choosing $d-p$-dimensional rectangles (i.e. intersections of half-spaces separated by orthogonal hyperplans) is natural, but our definitions are still valid for arbitrary borelian subsets in $\mathbb{R}^{d-p}$. Technically speaking, we will assume that the functions $\mathbf{x}_J \mapsto \mathbf{1}(\mathbf{x}_J \in A_J)$ are Donsker, to apply uniform CLTs' without any hurdle. Actually, working with $\mathbf{X}_J$-“boxes” instead of pointwise will simplify a lot the picture, by avoiding smoothing and bandwidth choices.

Note that, by definition of the conditional copula of $\mathbf{X}_I$ given $(\mathbf{X}_J \in A_J)$, we have

$$\begin{aligned}&\mathbb{P}(\mathbf{X}_I \leq \mathbf{x}_I|\mathbf{X}_J \in A_J)\\ &= C_{I|J}\left(\mathbb{P}(X_1 \leq x_1|\mathbf{X}_J \in A_J), \ldots, \mathbb{P}(X_p \leq x_p|\mathbf{X}_J \in A_J)|\mathbf{X}_J \in A_J\right),\end{aligned}$$

for every point $\mathbf{x}_I \in \mathbb{R}^p$ and every subset A_J in $\mathcal{A}_J$. So, it is tempting to replace $\mathcal{H}_0$ by

$$\widetilde{\mathcal{H}}_0 : C_{I|J}(\mathbf{u}_I|\mathbf{X}_J \in A_J) \,\text{does not depend on}\, A_J \in \mathcal{A}_J, \text{for any}\, \mathbf{u}_I.$$

Unfortunately and counter-intuitively, this reasoning does not lead to a consistent test of $(\mathcal{SA})$. Indeed, under the simplifying assumption, we can see that $C_{I|J}(\mathbf{u}_I|\mathbf{X}_J \in A_J)$ **depends on** A_J in general, even if $C_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)$ **does not depend on** $\mathbf{x}_J$**!**

This is due to the nonlinear transform between conditional (univariate and multivariate) distributions and conditional copulas. In other words, for a usual d-dimensional cdf H, we have

$$H(\mathbf{x}_I|\mathbf{X}_J \in A_J) = \frac{1}{\mathbb{P}(A_J)} \int_{A_J} H(\mathbf{x}_I|\mathbf{X}_J = \mathbf{x}_J)\, d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J), \tag{10}$$

for every measurable subset $A_J \in \mathcal{A}_J$ and $\mathbf{x}_I \in \mathbb{R}^p$. At the opposite and in general, for conditional copulas,

$$C_{I|J}(\mathbf{u}_I|\mathbf{X}_J \in A_J) \neq \frac{1}{\mathbb{P}(A_J)} \int_{A_J} C_{I|J}(\mathbf{u}_I|\mathbf{X}_J = \mathbf{x}_J)\, d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J), \tag{11}$$

for $\mathbf{u}_I \in [0,1]^p$. And even if we assume $(\mathcal{SA})$, we have in general,

$$C_{I|J}(\mathbf{u}_I|\mathbf{X}_J \in A_J) \neq \frac{1}{\mathbb{P}(A_J)} \int_{A_J} C_{s,I|J}(\mathbf{u}_I)\, d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J) = C_{s,I|J}(\mathbf{u}_I). \tag{12}$$

As a particular case, taking $A_J = \mathbb{R}^{d-p}$, this means again that $C_I(\mathbf{u}_I) \neq C_{s,I|J}(\mathbf{u}_I)$.

Let us check this rather surprising feature with the example of Remark 1 for another subset A_J. Recall that $(\mathcal{SA})$ is true and that $C_{s,1,2|3}(u,v) = uv$ for every $u, v \in [0,1]$. Consider the subset $(X_3 \leq a)$, for any real number a. The probability of this event is $\Phi(a)$. Now, let us verify that

$$uv \neq H(F^{-}_{1|3}(u|X_3 \leq a), F^{-}_{2|3}(v|X_3 \leq a)|X_3 \leq a),$$

for some u, v in $(0,1)$. Clearly, for every real number x_k, we have

$$\mathbb{P}(X_k \leq x_k|X_3 \leq a) = \frac{1}{\Phi(a)} \int_{-\infty}^{a} \Phi(x_k - z)\phi(z)\, dz, k = 1, 2, \text{ and}$$

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2 | X_3 \leq a) = \frac{1}{\Phi(a)} \int_{-\infty}^{a} \Phi(x_1 - z)\Phi(x_2 - z)\phi(z)\, dz.$$

In particular, $\mathbb{P}(X_k \leq 0 | X_3 \leq a) = (1 + \Phi(-a))/2$. Therefore, set $u^* = v^* = (1 + \Phi(-a))/2$ and we get

$$H(F^{-}_{1|3}(u^* | X_3 \leq a), F^{-}_{2|3}(v^* | X_3 \leq a) | X_3 \leq a) = H(0, 0 | X_3 \leq a)$$
$$= \frac{1}{3}\left(1 + \Phi(-a) + \Phi^2(-a)\right) \neq u^* v^*.$$

In this example, $C_{s,1,2|3}(\cdot) \neq C_{1,2|3}(\cdot | X_3 \leq a)$, for every a, even if $(\mathcal{SA})$ is satisfied.

Nonetheless, getting back to the general case, we can easily provide an equivalent of Equation (10) for general conditional copulas, i.e. without assuming $(\mathcal{SA})$.

Proposition 11. *For all* $\mathbf{u}_I \in [0,1]^p$ *and all* $A_J \in \mathcal{A}_J$,

$$C_{I|J}(\mathbf{u}_I | \mathbf{X}_J \in A_J) = \frac{1}{\mathbb{P}(A_J)} \int_{A_J} \psi(\mathbf{u}_I, \mathbf{x}_J, A_J) d\mathbb{P}_{\mathbf{X}_J}(\mathbf{x}_J), \text{ with}$$

$$\psi(\mathbf{u}_I, \mathbf{x}_J, A_J)$$
$$= C_{I|J}\Bigg(F_{1|J}\Big(F^{-}_{1|J}(u_1 | \mathbf{X}_J \in A_J)\Big|\mathbf{X}_J = \mathbf{x}_J\Big), \ldots, F_{p|J}\Big(F^{-}_{p|J}(u_p | \mathbf{X}_J \in A_J)\Big|\mathbf{X}_J = \mathbf{x}_J\Big)\Bigg|\mathbf{X}_J = \mathbf{x}_J\Bigg).$$

Now, we understand why (11) (and (12) under $(\mathcal{SA})$) are not identities: the conditional copulas, given the subset A_J, still depend on the conditional margins of $\mathbf{X}_I$ given $\mathbf{X}_J$ pointwise in general.

Note that, if for every $i \in I$, X_i is independent of $\mathbf{X}_J$, then

$$F_{i|J}\Big(F^{-}_{i|J}(u_i | \mathbf{X}_J \in A_J)\Big|\mathbf{X}_J = \mathbf{x}_J\Big) = F_i\Big(F_i^{-}(u_i)\Big) = u_i.$$

When $\mathbf{X}_I$ and $\mathbf{X}_J$ are independent and when $(\mathcal{SA})$ is satisfied, there is identity between $C_{I|J}(\mathbf{u}_I | X_J \in A_J)$ and $C_I(\mathbf{u}_I)$, and then $C_{s,I|J}(\mathbf{u}_I)$. We consider such circumstances as very peculiar and do not have to be confused with $(\mathcal{SA})$. Therefore, we advise to lead a preliminary test of independence between $\mathbf{X}_I$ and $\mathbf{X}_J$ (or at least between $\mathbf{X}_i$ and $\mathbf{X}_J$ for any $i = 1, \ldots, p$) before trying to test $(\mathcal{SA})$ itself.

Now, let us revisit the characterisation of $\mathcal{H}_0$ in terms of the independence property, as in Subsection 2.2. The latter analysis is confirmed by the equivalent of Proposition 6 in the case of conditioning subsets A_J. Now, the relevant random vector would be

$$\mathbf{Z}_{I|A_J} := \left(F_{1|J}(X_1 | \mathbf{X}_J \in A_J), \ldots, F_{p|J}(X_p | \mathbf{X}_J \in A_J)\right),$$

that has straightforward empirical counterparts. Then, it is tempting to test

$$\widetilde{\mathcal{H}}_0^* : \mathbf{Z}_{I|A_J} \text{ and } (X_J \in A_J) \text{ are independent for every borelian subset } A_J \subset \mathbb{R}^{d-p}.$$

Nonetheless, it can be proved easily that this is not a test of $\mathcal{H}_0$, unfortunately.

Proposition 12. $\mathbf{Z}_{I|A_J}$ *and* $(X_J \in A_J)$ *are independent for every measurable subset* $A_J \subset \mathbb{R}^{d-p}$ *iff* $\mathbf{X}_I$ *and* $\mathbf{X}_J$ *are independent.*

Proof: For any measurable subset A_J and any $\mathbf{u}_I \in [0,1]^p$, under $\widetilde{\mathcal{H}}_0^*$, we have

$$\mathbb{P}\left(\mathbf{Z}_{I|A_J} \leq \mathbf{u}_I, \mathbf{X}_J \in A_J\right) = \mathbb{P}\left(\mathbf{Z}_{I|A_J} \leq \mathbf{u}_I\right) \mathbb{P}(\mathbf{X}_J \in A_J).$$

Consider $\mathbf{x}_I \in \mathbb{R}^p$. Due to the continuity of the conditional cdfs', there exists u_k s.t. $F_k(x_k | \mathbf{X}_J \in A_J) = u_k$, $k = 1, \ldots, p$. Then, using the invertibility of $x \mapsto F_k(x | \mathbf{X}_J \in A_J)$, we get $\mathbb{P}\left(\mathbf{Z}_{I|A_J} \leq \mathbf{u}_I, \mathbf{X}_J \in A_J\right) = \mathbb{P}\left(\mathbf{X}_I \leq \mathbf{x}_I, \mathbf{X}_J \in A_J\right)$. This implies that $\widetilde{\mathcal{H}}_0^*$ is equivalent to the following property: for every $\mathbf{x}_I \in \mathbb{R}^p$ and A_J,

$$\mathbb{P}\left(X_I \leq \mathbf{x}_I, \mathbf{X}_J \in A_J\right) = \mathbb{P}\left(\mathbf{X}_I \leq \mathbf{x}_I\right)\mathbb{P}\left(\mathbf{X}_J \in A_J\right). \;\square$$

Previously, we have exhibited a simple trivariate model where $\mathcal{H}_0$ is satisfied when $\mathbf{X}_I$ and $\mathbf{X}_J$ are not independent. Then, we see that it is not reasonable to test whether the mapping $A_J \mapsto C_{I|J}(\cdot|\mathbf{X}_J \in A_J)$ is constant over $\mathcal{A}_J$, the set of **all** A_J such that $\mathbb{P}_{\mathbf{X}_J}(A_J) > 0$, with the idea of testing $\mathcal{H}_0$.

Nonetheless, one can weaken the latter assumption, and restrict oneself to a **finite** family $\overline{\mathcal{A}}_J$ of subsets with positive probabilities. For such a family, we could test the assumption

$$\overline{\mathcal{H}}_0 : A_J \mapsto C_{I|J}(\,\cdot\,|\mathbf{X}_J \in A_J) \text{ is constant over } \overline{\mathcal{A}}_J.$$

To fix the ideas and w.l.o.g., we will consider a given family of disjoint subsets $\overline{\mathcal{A}}_J = \{A_{1,J}, \dots, A_{m,J}\}$ in $\mathbb{R}^{d-p}$ hereafter. Note the following consequence of Proposition 11.

Proposition 13. *Assume that, for all $A_J \in \overline{\mathcal{A}}_J$ and for all $i \in I$,*

$$F_{i|J}(x|\mathbf{X}_J = \mathbf{x}_J) = F_{i|J}(x|\mathbf{X}_J \in A_J), \;\; \forall \mathbf{x}_J \in A_J, x \in \mathbb{R}.$$

Then, $(\mathcal{SA})$ implies $\overline{\mathcal{H}}_0$.

Obviously, if the family $\overline{\mathcal{A}}_J$ is too big, this will be too demanding: $\overline{\mathcal{H}}_0$ will be close to a test of independence between $\mathbf{X}_I$ and $\mathbf{X}_J$, and no longer a test of $\mathcal{H}_0$. Moreover, the chosen subsets in the family $\overline{\mathcal{A}}_J$ do not need to be disjoint, even if this would be a natural choice. As a special case, if $\mathbb{R}^{d-p} \in \overline{\mathcal{A}}_J$, the previous condition is equivalent to the independence between X_i and $\mathbf{X}_J$ for every $i \in I$.

A test of $\overline{\mathcal{H}}_0$ may be relevant in a lot of situations, beside technical arguments as the absence of smoothing. First, the case of discrete (or discretized) explanatory variables $\mathbf{X}_J$ is frequent. When $\mathbf{X}_J$ is discrete and takes a value among $\{\mathbf{x}_{1,J}, \dots, \mathbf{x}_{m,J}\}$, set $A_{k,J} = \{\mathbf{x}_{k,J}\}$, $k = 1, \dots, m$. Then, there is identity between testing $\mathcal{H}_0$ and $\overline{\mathcal{H}}_0$, with $\overline{\mathcal{A}}_J = \{A_{1,J}, \dots, A_{m,J}\}$. Second, the level of precision and sharpness of a copula model is often lower than the models for (conditional) margins. To illustrate this idea, a lot of complex and subtle models to explain the dynamics of asset volatilities are available when the dynamics of cross-assets dependencies are often a lot more basic and without clear-cut empirical findings. Therefore, it makes sense to simplify conditional copula models compared to conditional marginal models. This can be done by considering only a few possible conditional copulas, associated to some events $(\mathbf{X}_J \in A_{k,J})$, $k = 1, \dots, m$. For example, Jondeau and Rockinger (2006) (the first paper that introduced conditional dependence structures, beside Patton (2006a)) proposed a Gaussian copula parameter that take a finite of values randomly, based on the realizations of some past asset returns. Third, similar situation occur with most Markov-switching copula models, where a finite set of copulas is managed. In such models, the (unobservable, in general) underlying state of the economy determines the index of the box: see Cholette et al. (2009), Wang et al. (2013), Stöber and Czado (2014), Fink et al. (2016), among others.

Therefore, testing $\overline{\mathcal{H}}_0$ is of interest per se. Even if this is not equivalent to $\mathcal{H}_0$ (i.e. $(\mathcal{SA})$) formally, the underlying intuitions are close. And, particularly when the components of the conditioning variable $\mathbf{X}_J$ are numerous, it can make sense to restrict the information set of the underlying conditional copula to a fixed number of conveniently chosen subsets A_J. And the constancy of the underlying copula when $\mathbf{X}_J$ belongs to such subsets is valuable in a lot of practical situations. Therefore, in the next subsections, we study some specific tests of $\overline{\mathcal{H}}_0$ itself.

3.2 Non-parametric tests with "boxes"

To specify such tests, we need first to estimate the conditional marginal cdfs', for instance by

$$\hat{F}_{k|J}(x|\mathbf{X}_J \in A_J) := \frac{\sum_{i=1}^n \mathbf{1}(X_{i,k} \leq x, \mathbf{X}_{i,J} \in A_J)}{\sum_{i=1}^n \mathbf{1}(\mathbf{X}_{i,J} \in A_J)},$$

for every real x and $k = 1, \ldots, p$. Similarly the joint law of $\mathbf{X}_I$ given $(\mathbf{X}_J \in A_J)$ may be estimated by

$$\hat{F}_{I|J}(\mathbf{x}_I|\mathbf{X}_J \in A_J) := \frac{\sum_{i=1}^n \mathbf{1}(X_{i,I} \leq \mathbf{x}_I, \mathbf{X}_{i,J} \in A_J)}{\sum_{i=1}^n \mathbf{1}(\mathbf{X}_{i,J} \in A_J)}.$$

The conditional copula given $(\mathbf{X}_J \in A_J)$ will be estimated by

$$C_{I|J}(\mathbf{u}_I|\mathbf{X}_J \in A_J) = \hat{F}_{I|J}(\hat{F}^{-}_{1|J}(u_1|\mathbf{X}_J \in A_J), \ldots, \hat{F}^{-}_{p|J}(u_p|\mathbf{X}_J \in A_J)|\mathbf{X}_J \in A_J).$$

Therefore, it is easy to imagine tests of $\overline{\mathcal{H}}_0$, for instance

$$\overline{\mathcal{T}}_{KM,n} := \sup_{\mathbf{u}_I \in [0,1]^d} \sup_{k,l=1,\ldots,m} |\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{X}_J \in A_{k,J}) - \hat{C}_{I|J}(\mathbf{u}_I|\mathbf{X}_J \in A_{l,J})|,$$

$$\overline{\mathcal{T}}_{CvM,n} := \sum_{k,l=1}^m \int \left(\hat{C}_{I|J}(\mathbf{u}_I|\mathbf{X}_J \in A_{k,J}) - \hat{C}_{I|J}(\mathbf{u}_I|\mathbf{X}_J \in A_{l,J})\right)^2 w(d\mathbf{u}_I),$$

for some nonnegative weight functions w, or even

$$\overline{\mathcal{T}}_{dist,n} := \sum_{k,l=1}^m dist\left(\hat{C}_{I|J}(\cdot|\mathbf{X}_J \in A_{k,J}), \hat{C}_{I|J}(\cdot|\mathbf{X}_J \in A_{l,J})\right),$$

where $dist(\cdot,\cdot)$ denotes a distance between cdfs' on $[0,1]^p$. More generally, define the matrix

$$\widehat{M}(\overline{\mathcal{A}}_J) := \left[\mathbf{1}(k \neq l)\, dist\Big(\hat{C}_{I|J}(\cdot|\mathbf{X}_J \in A_{k,J}), \hat{C}_{I|J}(\cdot|\mathbf{X}_J \in A_{l,J})\Big)\right]_{1 \leq k,l \leq m},$$

and any statistic of the form $||\widehat{M}(\overline{\mathcal{A}}_J)||$ can be used as a test statistics of $\overline{\mathcal{H}}_0$, when $||\cdot||$ is a norm on the set of $m \times m$-matrices. Obviously, it is easy to introduce similar statistics based on copula densities instead of cdfs'.

3.3 Parametric test statistics with "boxes"

When we work with subsets $A_J \in \mathbb{R}^{d-p}$ instead of pointwise conditioning events $(\mathbf{X}_J = \mathbf{x}_J)$, we can adapt all the previous parametric test statistics of Subsection 2.3. Nonetheless, the framework will be slightly modified.

Let us assume that, for every $A_J \in \overline{\mathcal{A}}_J$, $C_{I|J}(\cdot|\mathbf{X}_J \in A_J)$ belongs to the same parametric copula family $\mathcal{C} = \{C_\theta, \theta \in \Theta\}$. In other words, $C_{I|J}(\cdot|\mathbf{X}_J \in A_J) = C_{\theta(A_J)}(\cdot)$ for every $A_J \in \overline{\mathcal{A}}_J$. Therefore, we could test the constancy of the mapping $A_J \mapsto \theta(A_J)$, i.e. to test

$$\overline{\mathcal{H}}_0^c : \text{the function } k \in \{1, \ldots, m\} \mapsto \theta(A_{k,J}) \text{ is a constant called } \theta_0^b.$$

Clearly, for every $A_J \in \overline{\mathcal{A}}_J$, we can estimate $\theta(A_J)$ by

$$\hat{\theta}(A_J) := \arg\max_{\theta \in \Theta} \sum_{i=1}^n \log c_\theta \left(\hat{F}_{1|J}(X_{i,1}|\mathbf{X}_{i,J} \in A_J), \ldots, \hat{F}_{p|J}(X_{i,p}|\mathbf{X}_{i,J} \in A_J)\right) \mathbf{1}(\mathbf{X}_{i,J} \in A_J).$$

It can be proved that the estimate $\hat{\theta}(A_J)$ is consistent and asymptotically normal, by revisiting the proof of Theorem 1 in Tsukahara (2005). Here, the single difference w.r.t. the latter paper is induced by the random sample size, modifying the limiting distributions. The proof is left to the reader.

Under the null $\mathcal{H}_0^c$, we have estimated the parameter of the copula of $Z_{I|J}$ by $\hat\theta_0$ in Equation (8). And under the new zero assumption $\overline{\mathcal{H}}_0^c$, the parameter of the copula of $(F_1(X_1|\mathbf{X}_J \in A_{k,J}), \dots, F_p(X_p|\mathbf{X}_J \in A_{k,J}))$ given $(\mathbf{X}_J \in A_{k,J})$ is the same for any $k = 1, \dots, m$. We denote it $C_{\theta_0^b}$, and we can still estimate θ_0^b by the semi-parametric procedure

$$\hat\theta_0^b := \arg\max_{\theta\in\Theta} \sum_{k=1}^m \sum_{i=1}^n \log c_\theta \left(\hat F_{1|J}(X_{i,1}|\mathbf{X}_{i,J} \in A_{k,J}), \dots, \hat F_{p|J}(X_{i,p}|\mathbf{X}_{i,J} \in A_{k,J})\right) \mathbf{1}(\mathbf{X}_{i,J} \in A_{k,J}).$$

Obviously, under some conditions of regularity and under $\overline{\mathcal{H}}_0^c$, it can be proved that $\hat\theta_0^b$ is consistent and asymptotically normal, by adapting the results of Tsukahara (2005).

For convenience, let us define the "box index" function $k(\mathbf{x}_J) := \sum_{k=1}^m k\mathbf{1}\{\mathbf{x}_J \in A_{k,J}\}$, for any $\mathbf{x}_J \in \mathbb{R}^{d-p}$. In other words, k is the index of the box $A_{k,J}$ that contains $\mathbf{x}_J$. It equals zero, when no box in $\overline{\mathcal{A}}_J$ contains $\mathbf{x}_J$. Let us introduce the r.v. $Y_i := k(\mathbf{X}_{i,J})$, that stores only all the needed information concerning the conditioning with respect to the variables $\mathbf{X}_{i,J}$. We can then define the empirical pseudo-observations as

$$\begin{aligned}
\mathbf{Z}_{i,I|Y} &:= \sum_{k=1}^m \left(F_{1|J}(X_{i,1}|\mathbf{X}_J \in A_{k,J}), \dots, F_{p|J}(X_{i,p}|\mathbf{X}_J \in A_{k,J})\right) \mathbf{1}\{\mathbf{X}_{i,J} \in A_{k,J}\} \\
&= \left(F_{1|J}(X_{i,1}|\mathbf{X}_J \in A_{k(\mathbf{X}_{i,J}),J}), \dots, F_{p|J}(X_{i,p}|\mathbf{X}_J \in A_{k(\mathbf{X}_{i,J}),J})\right) \\
&= \left(F_{1|Y}(X_{i,1}|Y_i), \dots, F_{p|Y}(X_{i,p}|Y_i)\right),
\end{aligned}$$

for any $i = 1, \dots, n$. Since we do not observe the conditional marginal cdfs', we define the observed pseudo-observations that we calculate in practice: for $i = 1, \dots, n$,

$$\hat{\mathbf{Z}}_{i,I|Y} := \left(\hat F_{1|J}(X_{i,1}|\mathbf{X}_J \in A_{Y_i,J}), \dots, \hat F_{p|J}(X_{i,p}|\mathbf{X}_J \in A_{Y_i,J})\right).$$

Note that we can then rewrite the previous estimators as

$$\hat\theta(A_{k,J}) = \arg\max_{\theta\in\Theta} \sum_{i=1}^n \log c_\theta \left(\hat{\mathbf{Z}}_{i,I|Y}\right) \mathbf{1}(Y_i = k), \text{ and } \hat\theta_0^b = \arg\max_{\theta\in\Theta} \sum_{i=1}^n \log c_\theta \left(\hat{\mathbf{Z}}_{i,I|Y}\right).$$

Now, let us revisit some of the previously proposed test statistics in the case of "boxes".

- Tests based on the comparison between $\hat\theta(\cdot)$ and $\hat\theta_0$:

$$\overline{\mathcal{T}}_\infty^c := \sqrt{n} \max_{k=1,\dots,m} \|\hat\theta(A_{k,J}) - \hat\theta_0\|, \ \overline{\mathcal{T}}_2^c := n \sum_{k=1}^m \|\hat\theta(A_{k,J}) - \hat\theta_0\|^2 \omega_k, \qquad (13)$$

  for some weights ω_k.

- Tests based on the comparison between $C_{\hat\theta(\cdot)}$ and $C_{\hat\theta_0}$:

$$\overline{\mathcal{T}}_{dist}^c := \sum_{k=1}^m dist(C_{\hat\theta(A_k)}, C_{\hat\theta_0})\omega_k,$$

and others.

3.4 Bootstrap techniques for tests with boxes

In the same way as in the previous section, we will need bootstrap schemes to evaluate the limiting laws of the test statistics of $\overline{\mathcal{H}}_0$ or $\overline{\mathcal{H}}_0^c$ under the null. All the nonparametric resampling schemes of Subsection 2.4.1 (in particular Efron's usual bootstrap) can be used in this framework, replacing the conditional pseudo-observations $\hat{\mathbf{Z}}_{i,I|J}$ by $\hat{\mathbf{Z}}_{i,I|Y}$, $i = 1, \dots, n$. The parametric resampling schemes of Subsection 2.4.1 can also be applied to the framework of "boxes", replacing $\hat\theta_0$ by $\hat\theta_0^b$ and $\hat\theta(\mathbf{x}_J)$ by $\hat\theta(A_J)$. In the parametric case, the bootstrapped estimates are denoted by $\hat{\theta}_0^*$ and $\hat{\theta}^*(A_J)$. They are the equivalents of $\hat{\theta}_0^b$ and $\hat{\theta}_n(A_J)$, replacing $(\hat{\mathbf{Z}}_{i,I|J}, Y_i)$ by $(\mathbf{Z}_i^*, Y_i^*)$.

The bootstrapped statistics will also be changed accordingly. Writing them explicitly is a rather straightforward exercise and we do not provide the details, contrary to Subsection 2.4. For example, the bootstrapped statistics corresponding to (13) is

$$\left(\overline{\mathcal{T}}_2^c\right)^* := n\sum_{k=1}^m \|\hat{\theta}^*(A_{k,J}) - \hat{\theta}(A_{k,J}) - \hat{\theta}_0^* + \hat{\theta}_0^b\|^2\omega_k,$$

where $\hat{\theta}_0^*$ is the result of the program $\arg\max_\theta \sum_{i=1}^n \log c_\theta\left(\hat{\mathbf{Z}}_{i,I|Y}^*\right)$, in the case of Efron's nonparametric bootstrap.

As we noticed in Remark 10, some changes are required when dealing with the "parametric independent" bootstrap. Indeed, under the alternative, we observe $\hat{\theta}^*(A_{k,J}) - \hat{\theta}_0^* \approx 0$, because we have precisely generated a bootstrap sample under $\overline{\mathcal{H}}_0^c$. As a consequence, the law of $\left(\overline{\mathcal{T}}_2^c\right)^*$ would be close to the law of $\overline{\mathcal{T}}_2^c$ but under the alternative, providing very small powers. Therefore, convenient bootstrapped test statistics of $\overline{\mathcal{H}}_0$ under the "parametric independent" scheme will be of the type

$$\left(\overline{\mathcal{T}}_2^c\right)^{**} := n\sum_{k=1}^m \|\hat{\theta}^*(A_{k,J}) - \hat{\theta}_0^*\|^2\omega_k.$$

Such a result is justified theoretically by the following theorem.

Theorem 14. *Assume that $\overline{\mathcal{H}}_0^c$ is satisfied, and that we apply the parametric independent bootstrap. Set*

$$\Theta n,0 := \sqrt{n}\big(\hat{\theta}_0 - \theta_0\big), \Theta_{n,k} := \sqrt{n}\big(\hat{\theta}(A_{k,J}) - \theta_0\big), k = 1,\dots,m,$$

$$\Theta_{n,0}^* := \sqrt{n}\big(\hat{\theta}_0^* - \theta_0\big), \text{ and } \Theta_{n,k}^* := \sqrt{n}\big(\hat{\theta}^*(A_{k,J}) - \theta_0\big), k = 1,\dots,m.$$

Then there exists two independent and identically distributed random vectors $\big(\Theta_0,\dots,\Theta_m\big)$ and $\big(\Theta_0^\perp,\dots,\Theta_m^\perp\big)$, and a real number a_0 such that

$$\Big(\Theta_{n,0},\dots,\Theta_{n,m},\Theta_{n,0}^*,\dots,\Theta_{n,m}^*\Big) \Longrightarrow \Big(\Theta_0,\dots,\Theta_m,\Theta_0^\perp + a_0\Theta_0,\dots,\Theta_m^\perp + a_0\Theta_0\Big).$$

The proof of this theorem has been postponed in Appendix A.

As a consequence of the latter result, applying the parametric independent bootstrap procedures for some test statistics based on comparisons between $\hat{\theta}_0$ and the $\hat{\theta}(A_{k,J})$ is valid. For instance, $\overline{\mathcal{T}}_2^c$ and $\left(\overline{\mathcal{T}}_2^c\right)^{**}$ will converge jointly in distribution to a pair of independent and identically distributed variables. Indeed, we have

$$\begin{aligned}
\left(\overline{\mathcal{T}}_2^c, \left(\overline{\mathcal{T}}_2^c\right)^{**}\right) &= \left(n\sum_{k=1}^m \|\hat{\theta}_{n,0}^b - \hat{\theta}_n(A_{k,J})\|^2\omega_k\,,\, n\sum_{k=1}^m \|\hat{\theta}_{n,0}^* - \hat{\theta}_n^*(A_{k,J})\|^2\omega_k\right) \\
&= \left(n\sum_{k=1}^m \|\hat{\theta}_{n,0}^b - \theta_0 + \theta_0 - \hat{\theta}_n(A_{k,J})\|^2\omega_k\,,\, n\sum_{k=1}^m \|\hat{\theta}_{n,0}^* - \theta_0 + \theta_0 - \hat{\theta}_n^*(A_{k,J})\|^2\omega_k\right) \\
&\Longrightarrow \left(\sum_{k=1}^m \|\Theta_0 - \Theta_k\|^2\omega_k\,,\, \sum_{k=1}^m \|\Theta_0^\perp + a_0\Theta_0 - \Theta_k^\perp - a_0\Theta_0\|^2\omega_k\right).
\end{aligned}$$

The same reasoning applies with $\overline{\mathcal{T}}_\infty^c$ and $\overline{\mathcal{T}}_{dist}^c$, for sufficiently regular copula families.

Remark 15. *We have to stress that the first-level bootstrap, i.e. resampling among the conditioning variables $\mathbf{Z}_{i,J}$, $i = 1,\dots,n$ is surely necessary to obtain the latter result. Indeed, it can be seen that the key proposition 16 is no longer true otherwise, because the limiting covariance functions of the two corresponding processes $\mathbb{G}_n$ and $\mathbb{G}_n^*$ will not be the same: see remark 22 below.*

4 Numerical applications

Now, we would like to evaluate the empirical performances of some of the previous tests by simulation. Such an exercise has been led by Genest et al. (2009) or Berg (2009) extensively in the case of goodness-of-fit test for unconditional copulas. Our goal is not to replicate such experiments in the case of conditional copulas and for tests of $(\mathcal{SA})$. Indeed, we have proposed dozens of test statistics and numerous bootstrap schemes. Moreover, testing $(\mathcal{SA})$ through $\mathcal{H}_0$ or some "box-type" problems through $\overline{\mathcal{H}}_0$ doubles the scale of the task. Finally, in the former case, we depend on smoothing parameters that induce additional degrees of freedom for the fine tuning of the experiments (note that Genest et al. (2009) and Berg (2009) have renounced to consider tests that require additional smoothing parameters, as the pivotal test statistics proposed in Fermanian (2005)). In our opinion, an exhaustive simulation experiment should be the topic of (at least) one additional paper. Here, we will restrict ourselves to some partial numerical elements. They should convince readers that the methods and techniques we have discussed previously provide fairly good results and can be implemented in practice safely.

Hereafter, we consider bivariate conditional copulas and a single conditioning variable, i.e. $p = 2$ and $d = 3$. The sample sizes will be $n = 500$. Concerning the bootstrap, we will resample $N = 100$ times to calculate approximated p-values. Each experiment has been repeated 500 times to calculate the percentages of rejection.

In terms of model specification, the margins of $\mathbf{X} = (X_1, X_2, X_3)$ will depend on X_3 as

$$X_1 \sim \mathcal{N}(X_3, 1),\ X_2 \sim \mathcal{N}(X_3, 1) \text{ and } X_3 \sim \mathcal{N}(0, 1).$$

We have studied the following conditional copula families: given $X_3 = x$,

- the Gaussian copula model, with a correlation parameter $\theta(x)$,
- the Student copula model, with 4 degrees of freedom and a correlation parameter $\theta(x)$,
- the Clayton copula model, with a parameter $\theta(x)$,
- the Gumbel copula model, with a parameter $\theta(x)$,
- the Frank copula model, with a parameter $\theta(x)$.

In every case, we calibrate $\theta(x)$ such that the conditional Kendall's tau $\tau(x)$ satisfies $\tau(x) = \Phi(x)\tau_{\max}$, for some constant $\tau_{\max} \in (0, 1)$. By default, $\tau_{\max}$ is equal to one. In this case, the random Kendall's tau are uniformly distributed on $[0, 1]$.

Test of $\mathcal{H}_0$: we calculate the percentage of rejections of $\mathcal{H}_0$, when the sample is drawn under the true law (level analysis) or when it is drawn under the same parametric copula family, but with varying parameters (power analysis). For example, when the true law is a Gaussian copula with a constant parameter ρ corresponding to $\tau = 1/2$, we draw samples under the alternative through a bivariate Gaussian copula whose random parameters are given by $\tau(X_3) = \Phi(X_3)$. The chosen test statistics are $\mathcal{T}^0_{CvM}$, $\tilde{\mathcal{T}}^0_{CvM}$ (nonparametrics test of $\mathcal{H}_0$), $\mathcal{I}_{\chi,n}$ and $\mathcal{I}_{2,n}$ (nonparametric tests of $\mathcal{H}_0$ based on the independence property) and $\mathcal{T}^c_2$ (a parametric test of $\mathcal{H}^c_0$). To compute these statistics, we use the estimator of the simplified copula defined in Equation (5).

Test of $\overline{\mathcal{H}}_0$: in the case of the test with boxes, the data-generating process will be

$$X_1 \sim \mathcal{N}(\gamma(X_3), 1),\ X_2 \sim \mathcal{N}(\gamma(X_3), 1) \text{ and } X_3 \sim \mathcal{N}(0, 1),$$

where $\gamma(x) = \Phi^{-1}\left(\lfloor m\Phi(X_3)\rfloor/m\right)$, so that the boxes are all of equal probability. As $m \to \infty$, we recover the continuous model for which $\gamma(x) = x$.

In the same way, we calibrate the parameter $\theta(x)$ of the conditional copulas such that the conditional Kendall's tau satisfies $\tau(X_3) = \lfloor m\Phi(X_3)\rfloor/m$. We have chosen $m = 5$ boxes of equal probability for X_3. We have only evaluated $\overline{\mathcal{T}}^c_2$ for testing $\overline{\mathcal{H}}^c_0$.

In the following tables, for the parametric tests,

- "bootNP" means the usual nonparametric bootstrap ;
- "bootPI" means the parametric independent bootstrap (where $\mathbf{Z}_{I|J}$ is drawn under $C_{\hat{\theta}_0}$ and $\mathbf{X}_J$ under the usual nonparametric bootstrap);
- "bootPC" means the parametric conditional bootstrap (nonparametric bootstrap for $\mathbf{X}_J$, and $\mathbf{X}_I$ is sampled from the estimated conditional copula $C_{\hat{\theta}(\mathbf{X}_J^*)}$).

Concerning tests of $\mathcal{H}_0$, the results are relatively satisfying. For the nonparametric tests and those based on the independence property (Tables 1 and 2) the rejection rates are large when $\tau_{\max} = 1$, and the theoretical levels (5%) are underestimated (a not problematic feature in practice). This is still the case for tests of $(\mathcal{SA})$ under a parametric copula model through $\mathcal{T}_{2,c}$: see Tables 3 and 4. The three bootstrap schemes provide similar numerical results. Remind that the bootstrapped statistics is $\mathcal{T}_{2,c}^{**}$ with bootPI. Tests of $\overline{\mathcal{H}}_0$ under a parametric framework and through $\overline{\mathcal{T}}_{2,c}$ confirm such observations.

We have tested the influence of $\tau_{\max}$: the smaller is this parameter, the smaller is the percentage of rejections under the alternative, because the simulated model tends to induce lower dependencies of copula parameters w.r.t. $\mathbf{X}_3$: see Figure 1.

We have not tried to exhibit an "asymptotically optimal" bandwidth selector for our particular testing problem. This could be the task for further research. We have preferred a basic ad-hoc procedure. In our test statistics, we smooth w.r.t. $F_3(X_3)$ (or its estimate, to be specific), whose law is uniform on $(0,1)$. A reasonable bandwidth h is given by the so-called rule-of-thumb in kernel density estimation, i.e. $h^* = \sigma(F_3(X_3))/n^{1/5} = 1/(\sqrt{12}n^{1/5}) = 0.083$. Such a choice has provided reasonable results. The typical influence of the bandwidth choice on the test results is illustrated in Figure 2. In general, the latter h^* belongs to reasonably wide intervals of "convenient" bandwidth values, so that the performances of our considered tests are not very sensitivie to the bandwidth choice.

Family	$\mathcal{T}^0_{CvM,n}$	$\tilde{\mathcal{T}}^0_{CvM,n}$	$\mathcal{I}_{\chi,n}$	$\mathcal{I}_{2,n}$
Gaussian	0	0	0	0
Student	0	0	0	1
Clayton	1	0	0	0
Gumbel	0	2	0	0
Frank	0	0	0	0

Table 1: Rejection percentages under the null (nonparametric tests, nonparametric bootstrap bootNP)

.

Family	$\mathcal{T}^0_{CvM,n}$	$\tilde{\mathcal{T}}^0_{CvM,n}$	$\mathcal{I}_{\chi,n}$	$\mathcal{I}_{2,n}$
Gaussian	97	100	100	96
Student	100	96	98	94
Clayton	100	99	98	98
Gumbel	100	97	100	97
Frank	100	100	94	32

Table 2: Rejection percentages under the alternative (nonparametric tests, nonparametric bootstrap bootNP)

.

Family	$\mathcal{T}_{2,c}$ bootPI	$\mathcal{T}_{2,c}$ bootPC	$\mathcal{T}_{2,c}$ bootNP	$\overline{\mathcal{T}}_{2,c}$ bootPI	$\overline{\mathcal{T}}_{2,c}$ bootNP
Gaussian	3	0	0	6	1
Student	5	0	3	7	1
Clayton	12	0	2	4	2
Gumbel	7	2	0	13	0
Frank	6	0	4	7	1

Table 3: Rejection percentages under the null (parametric tests)

.

Family	$\mathcal{T}_{2,c}$ bootPI	$\mathcal{T}_{2,c}$ bootPC	$\mathcal{T}_{2,c}$ bootNP	$\overline{\mathcal{T}}_{2,c}$ bootPI	$\overline{\mathcal{T}}_{2,c}$ bootNP
Gaussian	100	100	100	100	100
Student	100	100	100	100	100
Clayton	100	83	94	100	100
Gumbel	100	100	41	100	75
Frank	100	100	100	100	100

Table 4: Rejection percentages under the alternative (parametric tests)

.

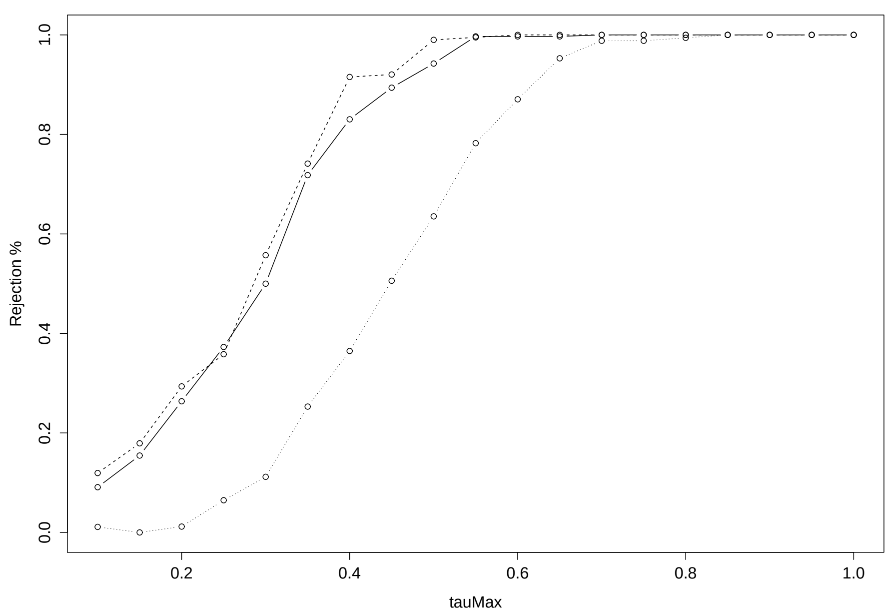

Figure 1: Rejection percentages for the statistics $\mathcal{T}_{2,c}$ as a function of τ_{max}: $\theta(x)$ is calibrated such that the conditional Kendall's tau $\tau(x)$ satisfies $\tau(x) = \Phi(x) \cdot \tau_{max}$. Solid line: bootNP. Dashed line : bootPI. Dotted line : bootPC.

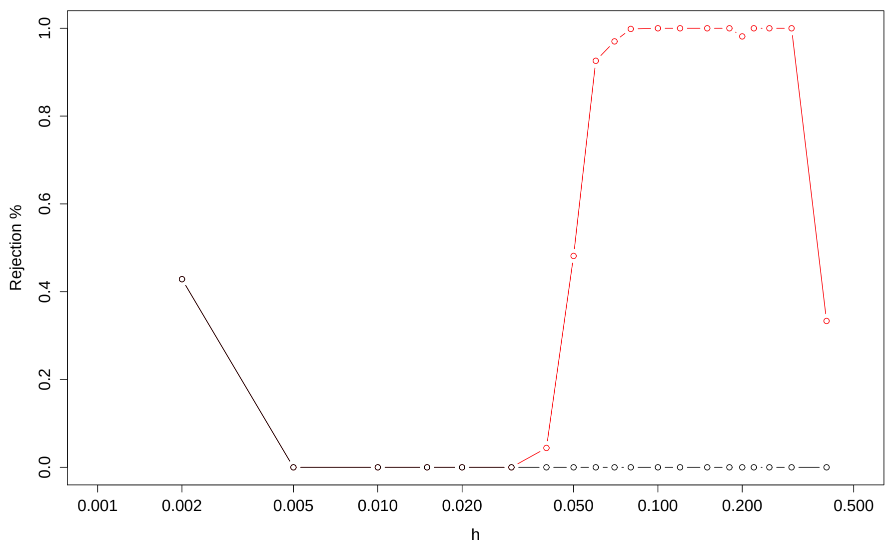

Figure 2: Rejection percentage for the statistic $\mathcal{I}_{\chi,n}$ as a function of h. The red (resp. black) line corresponds to the alternative (resp. zero) assumption.

5 Conclusion

We have provided an overview of the simplifying assumption problem, under a statistical point of view. In the context of nonparametric or parametric conditional copula models (with unknown conditional marginal distributions), numerous testing procedures have been proposed. We have developed the theory towards a slightly different but related approach, where "box-type" conditioning events replace pointwise ones. This open a new field for research that is interesting per se. Several new bootstrap procedures have been detailed, to evaluate p-values under the zero assumption in both cases. In particular, we have proved the validity of one of them (the "parametric independent" bootstrap scheme under $\overline{\mathcal{H}}_0$).

Clearly, there remains a lot of work. We have opened the Pandora box rather than provided definitive answers. Open questions are still numerous: precise theoretical convergence results of our test statistics (and others!), validity of these new bootstrap schemes, bandwidth choices, empirical performances,... All these dimensions would require further research. We have made a contribution to the landscape of $(\mathcal{SA})$-related problems, and proposed a working program for the whole copula community.

References

[1] Aas, K., Czado, C., Frigessi, A., and H. Bakken (2009). Pair-copula constructions of multiple dependence. *Insurance Math. Econom.* 44(2), 182-198.

[2] Abegaz, F., Gijbels, I. and N. Veraverbeke (2012). Semiparametric estimation of conditional copulas. *J. Multivariate Anal.* 110, $43-73$.

[3] Acar, E.F., Craiu, R.V. and F. Yao (2011). Dependence Calibration in Conditional copulas: A Nonparametric Approach. *Biometrics* 67, 445-453.

[4] Acar, E.F., Genest, C. and J. Nešlehová (2012). Beyond simplified pair-copula constructions. *J. Multivariate Anal.* 10, $74-90$.

[5] Acar, E.F., Craiu, R.V., and F. Yao (2013). Statistical testing of covariate effects in conditional copula models. *Electron. J. Stat.* 7, $2822-2850$.

[6] Andrews, D. (1997). Conditional Kolmogorov-Smirnov test. *Econometrica* 65, 1097-1128.

[7] Berg, D. (2009). Copula goodness-of-fit testing: An overview and power comparison *European J. Finance* **15**, 675-701.

[8] Bouzebda, S., Keziou, A. and T. Zari (2011). K-Sample Problem Using Strong Approximations of Empirical Copula Processes. *Math. Methods Statist.* 20, 14-29.

[9] Chernozhukov, V. Chetverikov, D. and Kato, K. (2014). Gaussian approximations of suprema of empirical processes. *Ann. Statist.* **42**, 1564-1597.

[10] Cholette, L., Heinen, A. and Valdesogo, A. (2009). Modeling international financial returns with a multivariate regime-switching copula. *J. Financial Econometrics* **7**, 437-480.

[11] Craiu, R.V. and A. Sabeti (2012). In mixed company: Bayesian inference for bivariate conditional copula models with discrete and continuous outcomes. *J. Multivariate Anal.* 110, 106-120.

[12] Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance, *Acad. Roy. Belg. Bull.Cl. Sci.* (5)65, 274–292.

[13] Deheuvels, P. (1981). A Kolmogorov–Smirnov Type Test for Independence and Multivariate Samples. *Rev. Roumaine Math. Pures Appl.* 26, 213–226.

[14] Fermanian, J-D (2005). Goodness-of-fit tests for copulas. *J. Multivariate Anal.* **95**, 119-152.

[15] Fermanian J-D and M. Wegkamp (2012). Time-dependent copulas. *J. Multivariate Anal.* 110, 19-29.

[16] Fermanian J.-D. and O. Lopez (2015). Single-index copulas. Working paper Crest 2015-12.

[17] Fink, H., Klimova, Y., Czado, C. and Stöber, J. (2016). Regime switching vine copula models for global equity and volatility indices. arXiv:1604.05598.

[18] Genest, C., Ghoudi, K. and L.-P. Rivest (1995). A semiparametric estimation procedure of dependenceparameters in multivariate families of distributions. *Biometrika* 82, 543-552.

[19] Genest, C. and B. Rémillard (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Ann. Inst. Henri Poincaré Probab. Stat.* 44 (6), 1096-1127.

[20] Genest, C., Rémillard, B. and A. Beaudoin (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance Math. Econom.* 44, 199-213

[21] Gijbels, I., Veraverbeke, N. and M. Omelka (2011). Conditional copulas, association measures and their applications. *Comput. Statist. Data Anal.* 55 1919-1932.

[22] Gijbels, I., Veraverbeke, N.and M. Omelka (2015). Estimation of a Copula when a Covariate Affects only Marginal Distributions. *Scand. J. Stat.* 42 1109-1126.

[23] I. Gijbels, M. Omelka, N. Veraverbeke (2016). Nonparametric testing for no covariate effects in conditional copulas. *Statistics*, online.

[24] Hobæk Haff, I., Aas, K. and A. Frigessi (2010). On the simplified pair-copula construction–simply useful or too simplistic? *J. Multivariate Anal.* 101 1296–1310.

[25] Jondeau, E. and Rockinger, M. (2006). The copula-garch model of conditional dependencies: An international stock market application. *J. of Internat. Money and Finance* **25**, 827-853.

[26] Nagler, T. and C. Czado (2016). Evading the curse of dimensionality in nonparametric density estimation with simplified vine copulas. arXiv:1503.03305

[27] Omelka, M., Veraverbeke, N. and I. Gijbels (2013). Bootstrapping the conditional copula. *J. Statist. Plann. Inference* 143, 1-23.

[28] Patton, A. (2006a) Modelling Asymmetric Exchange Rate Dependence, *Internat. Econom. Rev.* 47, 527-556.

[29] Patton, A. (2006b) Estimation of multivariate models for time series of possibly different lengths, *J. Appl. Econometrics* 21, 147-173.

[30] Portier, F. and J. Segers (2015). On the weak convergence of the empirical conditional copula under a simplifying assumption. arXiv:1511.06544.

[31] Rémillard, B. and O. Scaillet (2009). Testing for Equality between Two copulas. *J. Multivariate Anal.* 100, 377-386.

[32] Sabeti, A., Wei, M. and R.V. Craiu (2014). Additive models for conditional copulas *Stat* 3, 300-312.

[33] Spanhel, F. and M.S. Kurz (2015). Simplified vine copula models: Approximations based on the simplifying assumption. arXiv:1510-06971.

[34] Stöber, J., Joe, H. and C. Czado (2013). Simplified pair copula constructions—limitations and extensions. *J. Multivariate Anal.* 119 101–118.

[35] Stöber, J., Czado, C. (2014). Regime switches in the dependence structure of multivariate financial data. *Computational Statistics and Data Analysis* **76**, 672-686.

[36] Tsukahara, H. (2005). Semiparametric estimation in copula models. *Canad. J. Statist.* 33, 357-375.

[37] van der Vaart, A. and J. Wellner (1996). *Weak convergence and empirical processes.* Springer.

[38] Vatter, T. and V. Chavez-Demoulin (2015). Generalized additive models for conditional dependence structures. *J. Multivariate Anal.* 141, 147-167.

[39] Veraverbeke, N., Omelka, M. and I. Gijbels (2011). Estimation of a Conditional Copula and Association Measures. *Scand. J. Stat.* 38, 766-780.

[40] Wang, Y.-C., Wu, J.-L. and Lai, Y.-H. (2013). A Revisit to the Dependence Structure between the Stock and Foreign Exchange Markets: A Dependence-Switching Copula Approach. *J. Banking Finance* **37**, 1706-1719.

[41] Zheng, J.X. (2000). A consistent test of conditional parametric distributions. *Econometric Theory* 16 667-691.

A Proof of Theorem 14

A.1 Preliminaires

Let $(\mathbf{Z}_i)_{i=1,\dots,n}$ be a sequence of i.i.d random vectors in $[0,1]^p$, $\mathbf{Z}_i$ being drawn from the true cdf C_{θ_0}. They have the same law as the previously called vectors $\mathbf{Z}_{i,I|A_J}$ or $\mathbf{Z}_{i,I|Y}$ under the zero assumption $\overline{\mathcal{H}}_0^c$. Let $(\mathbf{X}_{i,J})_{i=1,\dots,n}$ be a sequence of i.i.d random vectors in $\mathbb{R}^{d-p}$, $\mathbf{X}_{i,J} \sim F_J$. Let $(\mathbf{Z}_i^*)_{i=1,\dots,n}$ be an independent sequence of i.i.d random vectors in $[0,1]^p$, where $\mathbf{Z}_i^* \sim C_{\theta_0}$ exactly as $\mathbf{Z}_i$. The three samples $(\mathbf{Z}_i)$, $(\mathbf{X}_{i,J})$ and $(\mathbf{Z}_i^*)$ are mutually independent. Let $(\mathbf{X}_{i,J}^*)_{i=1,\dots,n}$ be a sequence of i.i.d random vectors in $\mathbb{R}^{d-p}$, which are drawn from $F_{n,J}$, the empirical cdf of $\mathbf{X}_{1,J},\dots \mathbf{X}_{n,J}$, and independently of both $(\mathbf{Z}_i)$ and $(\mathbf{Z}_i^*)$.

In the following, we shall use the notation $f \otimes g := (x,y) \mapsto f(x)g(y)$ when f, g are two real functions, possibly from different spaces. Set $l(\theta,\cdot) := \log c_\theta(\cdot)$. We will need some conditions of regularity.

Assumption (R): $(\theta, \mathbf{u}_I) \mapsto l(\theta,\mathbf{u}_I)$ is three times differentiable with respect to θ, for every $\mathbf{u}_I \in (0,1)^p$. Moreover, for every $\epsilon > 0$,

$$\mathbb{E}\left[\sup_{\theta | \|\theta-\theta_0\|\leq \epsilon} \sup_{\{\mathbf{z} | \|\mathbf{z}-\mathbf{Z}_i\| \leq \|\hat{\mathbf{Z}}_{i,I|Y}-\mathbf{Z}_i\|\}} || \frac{\partial^3 l}{\partial \theta^3}(\theta,\mathbf{z}) || \right] < +\infty.$$

The latter technical assumption can be weakened through some trimming techniques, as in Fermanian and Lopez (2015). Since this would require to change the definitions of the parametric estimators, we do not try to improve towards this direction. We will set $\dot{c}_\theta := \partial c_\theta/\partial\theta$ and $\ddot{c}_\theta := \partial^2 c_\theta/\partial\theta^2$.

We associate to every $\mathbf{X}_{i,J}$ (resp. $\mathbf{X}_{i,J}^*$) its corresponding index Y_i (resp. Y_i^*) s.t. $\mathbf{X}_{i,J} \in A_{Y_i}$ (resp. $\mathbf{X}_{i,J}^* \in A_{Y_i^*}$). For convenience, we assume that $(A_k)_{k=1,\dots,m}$ is a partition of $\mathbb{R}^{d-p}$. Otherwise, we have to restrict our sample to the observations for which $X_{i,J}$ belongs to some "box" A_k, $k=1,\dots,m$. Therefore, denote by C_n, C_n^*, $P_{n,Y}$ and $P_{n,Y}^*$ the empirical laws of $(\mathbf{Z}_i)$, $(\mathbf{Z}_i^*)$, (Y_i) and (Y_i^*) respectively. The joint law of $(\mathbf{Z}_1, Y_1)$ (resp. $(\mathbf{Z}_1, \mathbf{X}_{1,J})$) will be denoted by $\overline{G} := C_{\theta_0} \otimes P_Y$ (resp. $\overline{G} := C_{\theta_0} \otimes F_J$), with $P_Y(k) = \mathbb{P}(Y = k)$, $k=1,\dots,m$. Denote by G_n (resp. $\overline{G}_n$) the empirical law of $(\mathbf{Z}_i, Y_i)_{i=1,\dots,n}$ (resp. $(\mathbf{Z}_i, \mathbf{X}_{i,J})_{i=1,\dots,n}$) Moreover, G_n^* and $\overline{G}_n^*$ will be the empirical distributions of $(\mathbf{Z}_i^*, Y_i^*)_{i=1,\dots,n}$ and $(\mathbf{Z}_i^*, \mathbf{X}_{i,J}^*)_{i=1,\dots,n}$ respectively. Let $\mathcal{P}_n$ be the joint probability distribution of

$$\left(\mathbf{Z}_i, Y_i, \mathbf{Z}_i^*, Y_i^*\right)_{i=1,\dots,n} \in \left([0,1]^p \times \{1,\dots,m\}\right)^{\otimes 2n}.$$

The following proposition is key. It will be proved in Subsection A.3.

Proposition 16. *Consider the empirical process defined on* $[0,1]^p \times \mathbb{R}^{d-p}$ *by*

$$\overline{\mathbb{G}}_n(\mathbf{z},\mathbf{x}_J) := \sqrt{n}(\overline{G}_n - \overline{G})(\mathbf{z},\mathbf{x}_J) := \frac{1}{\sqrt{n}}\sum_{i=1}^n \{\mathbf{1}\big((\mathbf{Z}_i,\mathbf{X}_{i,J}) \leq (\mathbf{z},\mathbf{x}_J)\big) - C_{\theta_0}(\mathbf{z})F_J(\mathbf{x}_J)\},$$

and the corresponding bootstrapped empirical process

$$\overline{\mathbb{G}}_n^*(\mathbf{z},\mathbf{x}_J) := \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{1}\Big((\mathbf{Z}_i^*,\mathbf{X}_{i,J}^*) \leq (\mathbf{z},\mathbf{x}_J)\Big) - C_{\theta_0}(\mathbf{z})\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{1}(\mathbf{X}_{i,J}\leq \mathbf{x}_J),$$

or, equivalently, $\overline{\mathbb{G}}_n^* = \sqrt{n}(\overline{G}_n^* - C_{\theta_0}\otimes F_{n,J})$. *Then there exist two independent and identically distributed Gaussian processes* $\overline{\mathbb{A}}_G$ *and* $\overline{\mathbb{A}}_G^\perp$ *such that* $\big(\overline{\mathbb{G}}_n\,,\,\overline{\mathbb{G}}_n^*\big)$ *converges to* $(\overline{\mathbb{A}}_G\,,\,\overline{\mathbb{A}}_G^\perp)$ *weakly in* $\Big(\ell^\infty\big([0,1]^p\times\mathbb{R}^{d-p}\big)\Big)^2$.

As a Corollary, we deduce the same results when the discrete variables Y_i replace the variables $\mathbf{X}_{i,J}$.

Proposition 17. *Under the assumptions of Proposition 16, let the empirical process defined on* $[0,1]^p\times\{1,\dots,m\}$ *by*

$$\begin{aligned}\mathbb{G}_n(\mathbf{z},k) &:= \sqrt{n}(G_n - G)(\mathbf{z},k)\\ &:= \frac{1}{\sqrt{n}}\sum_{i=1}^n \{\mathbf{1}\big(\mathbf{Z}_i\leq \mathbf{z}, Y_i = k\big) - C_{\theta_0}(\mathbf{z})P_Y(k)\},\end{aligned}$$

and its bootstrapped empirical process

$$\mathbb{G}_n^*(\mathbf{z},k) := \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{1}\big(\mathbf{Z}_i^*\leq \mathbf{z}, Y_i^* = k\big) - C_{\theta_0}(\mathbf{z})\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{1}(Y_i = k),$$

or equivalently $\mathbb{G}_n^* = \sqrt{n}(G_n^* - C_{\theta_0}\otimes P_{n,Y})$, $P_{n,Y}(k)$ *being the empirical proportion of* $\mathcal{S}_n$*-observations into* A_k. *Then, there exist two independent and identically distributed processes* $\mathbb{A}_G$ *and* $\mathbb{A}_G^\perp$ *such that* $\big(\mathbb{G}_n\,,\,\mathbb{G}_n^*\big)$ *converges to* $(\mathbb{A}_G\,,\,\mathbb{A}_G^\perp)$ *weakly in* $\Big(\ell^\infty\big([0,1]^p\times\{1,\dots,m\}\big)\Big)^2$.

Remark 18. *The covariance function of* $\mathbb{A}_G$ *(or* $\mathbb{A}^\perp$*) is given by*

$$\begin{aligned}&\mathbb{E}[\mathbb{A}_G(\mathbf{z},y)\mathbb{A}_G(\mathbf{z}',y')] = \lim_n \mathbb{E}[\mathbb{G}_n(\mathbf{z},y)\mathbb{G}_n(\mathbf{z}',y')]\\ &= \mathbf{1}(y=y')\mathbb{P}(Y=y)C_{\theta_0}(\mathbf{z}\wedge\mathbf{z}') - \mathbb{P}(Y=y)\mathbb{P}(Y=y')C_{\theta_0}(\mathbf{z})C_{\theta_0}(\mathbf{z}').\end{aligned}$$

As a "toolbox", we will need the following lemma.

Lemma 19. *Let* $\hat\theta_0^b$ *and* $\hat\theta(A_k)$ *be the estimators based on the pseudo-sample* $(\hat{\mathbf{Z}}_{i,I|Y},Y_i)_{i=1,\dots,n}$ *(and then on the sample* $(\mathbf{Z}_i,Y_i)_{i=1,\dots,n}$*) as*

$$\hat\theta_0^b := \arg\max_{\theta\in\Theta}\sum_{i=1}^n \log c_\theta(\hat{\mathbf{Z}}_{i,I|Y}),\ \text{and}$$

$$\hat\theta(A_k) := \arg\max_{\theta\in\Theta}\sum_{i=1}^n \log c_\theta(\hat{\mathbf{Z}}_{i,I|Y}).\mathbf{1}(Y_i=k),\ k=1,\dots,m.$$

We will assume they lie in the interior of Θ. *Set* $\Theta_{n,0} := \sqrt{n}\big(\hat\theta_0^b - \theta_0\big)$, *and, for* $k=1,\dots,m$, $\Theta_{n,k} := \sqrt{n}\big(\hat\theta(A_k)-\theta_0\big)$. *Moreover, for any distribution* H *on* $[0,1]^p\times\{1,\dots,m\}$, *set*

$$\psi_{k,1}(H) := \int \frac{\partial l}{\partial\theta}\left(\theta_0, \left(\frac{\int \mathbf{1}\{z_q^1\leq z_q^2, y^1=y^2\}dH(\mathbf{z}^1,y^1)}{\int \mathbf{1}\{y^1=k\}dH(\mathbf{z}^1,y^1)}\right)_{q=1,\dots,p}\right)\mathbf{1}\{y^2=k\}\,dH(\mathbf{z}^2,y^2),$$

$$\psi_{k,2}(H) := \int \frac{\partial^2 l}{\partial\theta^2}\left(\theta_0, \left(\frac{\int \mathbf{1}\{z_q^1\leq z_q^2, y^1=y^2\}dH(\mathbf{z}^1,y^1)}{\int \mathbf{1}\{y^1=k\}\,dH(\mathbf{z}^1,y^1)}\right)_{q=1,\dots,p}\right)\mathbf{1}\{y^2=k\}\,dH(\mathbf{z}^2,y^2).$$

(i) For $k=1,\dots,m$,

$$\Theta_{n,k} = -\frac{\sqrt{n}\psi_{k,1}(G_n)}{\psi_{k,2}(G_n)} + o_P(1).$$

(ii) For every discrete law P_Y with values in $\{1,\dots,m\}$, the corresponding distribution $\tilde G := C_{\theta_0} \otimes P_Y$ satisfies $\psi_{k,1}(\tilde G) = 0$.

(iii) $\psi_1 := (\psi_{1,1},\dots,\psi_{m,1})$ is Hadamard-differentiable at every cdf H, and its differential is given by

$$\dot\psi_{k,1}(H)(h) = \int \frac{\partial l}{\partial \theta}\left(\theta_0, \left(\frac{\int \mathbf{1}\{z_q^1 \le z_q^2, y^1 = y^2\} dH(\mathbf{z}^1, y^1)}{\int \mathbf{1}\{y^1 = k\} dH(\mathbf{z}^1,y^1)}\right)_{q=1,\dots,p}\right) \mathbf{1}\{y^2 = k\}\, dh(\mathbf{z}^2, y^2)$$
$$+ \sum_{j=1}^p \int \frac{\partial^2 l}{\partial\theta\,\partial z_j}\left(\theta_0, \left(\frac{\int \mathbf{1}\{z_q^1 \le z_q^2, y^1 = y^2\} dH(\mathbf{z}^1, y^1)}{\int \mathbf{1}\{y^1 = k\} dH(\mathbf{z}^1,y^1)}\right)_{q=1,\dots,p}\right) \mathbf{1}\{y^2 = k\}$$
$$\cdot \left(\frac{\int \mathbf{1}\{z_j^1 \le z_j^2, y^1 = y^2\} dh(\mathbf{z}^1, y^1)}{\int \mathbf{1}\{y^1 = k\} dH(\mathbf{z}^1,y^1)} - \frac{\int \mathbf{1}\{z_j^1 \le z_j^2, y^1 = y^2\} dH(\mathbf{z}^1, y^1) \int \mathbf{1}\{y^1 = k\} dh(\mathbf{z}^1,y^1)}{\left(\int \mathbf{1}\{y^1 = y^2\} dH(\mathbf{z}^1,y^1)\right)^2}\right) dH(\mathbf{z}^2, y^2)$$

(iv)

$$\Theta_{n,0} = -\frac{\sum_{k=1}^m \sqrt{n}\big(\psi_{k,1}(G_n)\big)}{\sum_{k=1}^m \psi_{k,2}(G_n)} + o_P(1) = \frac{\sum_{k=1}^m \psi_{k,2}(G_n)\Theta_{n,k}}{\sum_{k=1}^m \psi_{k,2}(G_n)} + o_P(1)$$

Proof : Note that $\hat{\mathbf{Z}}_{i,I|J}$ is an explicit measurable function of the sample $(\mathbf{Z}_{i,I|J})_{i=1,\dots,n}$. Indeed, for any $i=1,\dots,n$ and $q=1,\dots,p$,

$$\begin{aligned}
\hat Z_{i,q|Y} :=& \hat F_{n,q}(X_{i,q} | \mathbf{X}_J \in A_{Y_i,J}) \\
:=& \frac{\sum_{j=1}^n \mathbf{1}\{X_{j,q} \le X_{i,q}, \mathbf{X}_{j,J} \in A_{Y_i,J}\}}{\sum_{j=1}^n \mathbf{1}\{\mathbf{X}_{j,J} \in A_{k(\mathbf{X}_{i,J}),J}\}} \\
=& \frac{\sum_{j=1}^n \mathbf{1}\{F_q(X_{j,q}|\mathbf{X}_J \in A_{Y_j,J}) \le F_q(X_{i,q}|\mathbf{X}_J \in A_{Y_j,J}), Y_j = Y_i\}}{\sum_{j=1}^n \mathbf{1}\{Y_j = Y_i\}} \\
=& \frac{\sum_{j=1}^n \mathbf{1}\{F_q(X_{j,q}|\mathbf{X}_J \in A_{Y_j,J}) \le F_q(X_{i,q}|\mathbf{X}_J \in A_{Y_i,J}), Y_j = Y_i\}}{\sum_{j=1}^n \mathbf{1}\{Y_j = Y_i\}} \\
=& \frac{\sum_{j=1}^n \mathbf{1}\{Z_{j,q|Y} \le Z_{i,q|Y}, Y_j = Y_i\}}{\sum_{j=1}^n \mathbf{1}\{Y_j = Y_i\}}.
\end{aligned} \tag{14}$$

We deduce that $\hat\theta_0^b$ and $\hat\theta(A_k)$ are measurable functions of the unobservable random variables $\mathbf{Z}_{i,I|Y}$ and Y_i, for $i=1,\dots,n$.

(i). Let $k \in \{1,\dots,m\}$. Applying successively the first order condition for the estimator $\hat\theta(A_k)$ and some Taylor series expansions, we have

$$\begin{aligned}
0 &= \frac{1}{n}\sum_{i=1}^n \frac{\partial l}{\partial\theta}(\hat\theta(A_k), \hat{\mathbf{Z}}_{i,I|J}) \mathbf{1}\{Y_i = k\} \\
&= B_n^{1,k} - B_n^{2,k}\big(\hat\theta(A_{k,J}) - \theta_0\big) + o_P\big(\hat\theta(A_{k,J}) - \theta_0\big), \text{ with}
\end{aligned}$$

$$B_n^{1,k} := \frac{1}{n}\sum_{i=1}^n \frac{\partial l}{\partial\theta}\big(\theta_0, \hat{\mathbf{Z}}_{i,I|J}\big)\mathbf{1}\{Y_i = k\} \text{ and } B_n^{2,k} := -\frac{1}{n}\sum_{i=1}^n \frac{\partial^2 l}{\partial\theta^2}\big(\theta_0, \hat{\mathbf{Z}}_{i,I|J}\big)\mathbf{1}\{Y_i = k\},$$

implying

$$\Theta_{n,k} := \sqrt{n}\big(\hat\theta(A_{k,J}) - \theta_0\big) = \frac{\sqrt{n}B_n^{1,k}}{B_n^{2,k}} + o_P\big(\Theta_{n,k}\big).$$

Now, invoking (14), let us compute the numerator of this expression:

$$\begin{aligned}
B_n^{1,k} &= \frac{1}{n}\sum_{i=1}^n \frac{\partial l}{\partial \theta}\left(\theta_0, \left(\frac{\sum_{j=1}^n \mathbf{1}\{Z_{j,q}\le Z_{i,q}, Y_q = k\}}{\sum_{j=1}^n \mathbf{1}\{Y_j = k\}}\right)_{q=1,\dots,p}\right)\mathbf{1}\{Y_i = k\} \\
&= \int \frac{\partial l}{\partial \theta}\left(\theta_0, \left(\frac{\int \mathbf{1}\{z_q^1 \le z_q^2, y^1 = k\} dG_n(\mathbf{z}^1, y^1)}{\int \mathbf{1}\{y^1 = k\} dG_n(\mathbf{z}^1, y^1)}\right)_{q=1,\dots,p}\right)\mathbf{1}\{y^2 = k\} dG_n(\mathbf{z}^2, y^2) \\
&= \psi_{k,1}(G_n).
\end{aligned}$$

In the same way, the denominator can be rewritten as

$$\begin{aligned}
B_n^{2,k} &= -\int \frac{\partial^2 l}{\partial \theta^2}\left(\theta_0, \left(\frac{\int \mathbf{1}\{z_q^1 \le z_q^2, y^1 = k\} dG_n(\mathbf{z}^1, y^1)}{\int \mathbf{1}\{y^1 = k\} dG_n(\mathbf{z}^1, y^1)}\right)_{q=1,\dots,p}\right)\mathbf{1}\{y^2 = k\} dG_n(\mathbf{z}^2, y^2) \\
&= -\psi_{k,2}(G_n).
\end{aligned}$$

(ii). We now prove the second part of the lemma. Since $\tilde{G} = C_{\theta_0} \otimes F_Y$, we get

$$\begin{aligned}
\psi_{k,1}(\tilde{G}) &:= \int \frac{\partial l}{\partial \theta}\left(\theta_0, \left(\frac{\int \mathbf{1}\{z_q^1 \le z_q^2, y^1 = k\} d\tilde{G}(\mathbf{z}^1, y^1)}{\int \mathbf{1}\{y^1 = k\} d\tilde{G}(\mathbf{z}^1, y^1)}\right)_{q=1,\dots,p}\right)\mathbf{1}\{y^2 = k\} d\tilde{G}(\mathbf{z}^2, y^2) \\
&= \int \frac{\partial l}{\partial \theta}\left(\theta_0, \left(\frac{\mathbb{P}\{Z_q^1 \le z_q^2, Y^1 = k\}}{\mathbb{P}\{Y^1 = k\}}\right)_{q=1,\dots,p}\right)\mathbf{1}\{y^2 = k\} d\tilde{G}(\mathbf{z}^2, y^2) \\
&= \int \frac{\partial l}{\partial \theta}\left(\theta_0, \left(\mathbb{P}\{Z_q^1 \le z_q^2\}\right)_{q=1,\dots,p}\right) dC_{\theta_0}(\mathbf{z}^2) \int \mathbf{1}\{y^2 = k\} dF_Y(y^2) \\
&= \mathbb{P}\{Y = k\} \int \frac{\partial l}{\partial \theta}\left(\theta_0, \mathbf{z}^2\right) dC_{\theta_0}(\mathbf{z}^2) = 0.
\end{aligned}$$

(iii). We remark that the law G appears three times in $\psi_{k,1}$: two times in the log-density l and one time at the end of the main integral. By separating the effect of a change from H to $H + h$ in the main integral only (first term of the differential) and the effect of a change in l, and using the standard rule of differential calculus (l is differentiable), we obtain the second part of the given result.

(iv). As in the proof of (i), we apply successively the first order condition for $\hat{\theta}_{n,0}^b$ and some Taylor series expansion to get

$$0 = \frac{1}{n}\sum_{i=1}^n \frac{\partial l}{\partial \theta}(\hat{\theta}_0^b, \hat{\mathbf{Z}}_{i,I|Y}) = B_n^1 - \left(\hat{\theta}_0^b - \theta_0\right) B_n^2 + o_P\left(\hat{\theta}_0^b - \theta_0\right), \text{ with}$$

$$B_n^1 := \frac{1}{n}\sum_{i=1}^n \frac{\partial l}{\partial \theta}\left(\theta_0, \hat{\mathbf{Z}}_{i,I|Y}\right) = \sum_{k=1}^m B_n^{1,k} \text{ and } B_n^2 := -\frac{1}{n}\sum_{i=1}^n \frac{\partial^2 l}{\partial \theta^2}\left(\theta_0, \hat{\mathbf{Z}}_{i,I|Y}\right) = \sum_{k=1}^m B_n^{2,k}.$$

We deduce

$$\begin{aligned}
\Theta_{n,0} := \sqrt{n}\left(\hat{\theta}_0^b - \theta_0\right) &= \frac{\sqrt{n} B_n^1}{B_n^2} + o_P\left(\Theta_{n,0}\right) = \frac{\sqrt{n}\sum_{k=1}^m B_n^{1,k}}{\sum_{k=1}^m B_n^{2,k}} + o_P\left(\Theta_{n,0}\right) \\
&= \frac{\sqrt{n}\sum_{k=1}^m \psi_{k,1}(G_n)}{\sum_{k=1}^m -\psi_{k,2}(G_n)} + o_P\left(\Theta_{n,0}\right) \\
&= \frac{\sum_{k=1}^m \psi_{k,2}(G_n)\Theta_{n,k}}{\sum_{k=1}^m \psi_{k,2}(G_n)} + o_P\left(\Theta_{n,0}\right). \quad \square
\end{aligned}$$

Lemma 20. *Let* ℓ_n *be defined by*

$$\ell_n := \sum_{i=1}^n \log\left(\frac{c_{\hat\theta_0^b}(\mathbf{Z}_i^*)}{c_{\theta_0}(\mathbf{Z}_i^*)}\right).$$

If there exists a random vector Θ_0 *such that* $\Theta_{n,0} \Longrightarrow \Theta_0$ *under* $\mathcal{P}_n$*, then we have*

$$\ell_n = \Theta_0^T \mathbb{W}^\perp - \frac{1}{2}\Theta_0^T I_0 \Theta_0 + o_P(1),$$

where $\mathbb{W}^\perp \sim \mathcal{N}(0, I_0)$ *is independent of the sample* $\left(\mathbf{Z}_{i,I|Y}, Y_i\right)_{i=1,\dots,n}$ *and* I_0 *is the Fisher information matrix*

$$I_0 := \mathbb{E}_{C_{\theta_0}}\left[\frac{\dot{c}_{\theta_0}^T(\mathbf{Z})\,\dot{c}_{\theta_0}(\mathbf{Z})}{c_{\theta_0}^2(\mathbf{Z})}\right].$$

Proof : By a Taylor expansion, we obtain

$$\begin{aligned}\ell_n &= \sum_{i=1}^n \{l\left(\hat\theta_0^b, \mathbf{Z}_i^*\right) - l\left(\theta_0, \mathbf{Z}_i^*\right)\} \\ &= \left(\hat\theta_0^b - \theta_0\right)^T \sum_{i=1}^n \frac{\partial l}{\partial\theta}\left(\theta_0, \mathbf{Z}_i^*\right) + \frac{1}{2}\left(\hat\theta_0^b - \theta_0\right)^T \sum_{i=1}^n \frac{\partial^2 l}{\partial\theta^2}\left(\theta_0, \mathbf{Z}_i^*\right)\left(\hat\theta_0^b - \theta_0\right) + R_n \\ &= \Theta_{n,0}^T\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\partial l}{\partial\theta}\left(\theta_0, Z_i^*\right)\right] + \frac{1}{2}\Theta_{n,0}^T\left[\frac{1}{n}\sum_{i=1}^n \frac{\partial^2 l}{\partial\theta^2}\left(\theta_0, Z_i^*\right)\right]\Theta_{n,0} + R_n.\end{aligned}$$

First, we have

$$\begin{aligned}R_n &\leq Cst||\hat\theta_0^b - \theta_0||^3 \sup_{\theta \mid \|\theta-\theta_0\| \leq \|\hat\theta_0^b - \theta_0\|} ||\sum_{i=1}^n \frac{\partial^3 l}{\partial\theta^3}\left(\theta, Z_i^*\right)|| \\ &\leq Cst||\Theta_{n,0}||^3 \cdot \sup_{\theta \mid \|\theta-\theta_0\| \leq \|\hat\theta_0^b - \theta_0\|} ||\frac{1}{n}\sum_{i=1}^n \frac{\partial^3 l}{\partial\theta^3}\left(\theta, Z_i^*\right)|| \cdot \frac{1}{\sqrt{n}} = O_P\left(\frac{1}{\sqrt{n}}\right),\end{aligned}$$

by Assumption (R). By the usual CLT, we know that $\frac{1}{\sqrt{n}}\sum_{i=1}^n \partial l/\partial\theta\left(\theta_0, Z_i^*\right) \longrightarrow \mathbb{W}^\perp$. $\mathbb{W}^\perp$ is independent of $\left(\mathbf{Z}_{i,I|Y}, Y_i\right)_{i=1,\dots,n}$ as a limit of a sequence of variables that have the same property. Using the law of large numbers, we have also

$$\frac{1}{n}\sum_{i=1}^n \frac{\partial^2 l}{\partial\theta^2}\left(\theta_0, Z_i^*\right) = \frac{1}{n}\sum_{i=1}^n \frac{\ddot{c}_\theta}{c_\theta}(Z_i^*) - \frac{\dot{c}_\theta^T\dot{c}_\theta}{c_\theta^2}(Z_i^*) \Longrightarrow 0 - I_0 \quad \square$$

A.2 Proof of Theorem 14

We first reason under $\mathcal{P}_n$ as in Theorem 1 in Genest and Rémillard (2008). By Proposition 17, under $\mathcal{P}_n$, there exist two independent and identically distributed processes $\mathbb{A}_G$ and $\mathbb{A}_G^\perp$ such that

$$\sqrt{n}\Big(G_n - C_{\theta_0}\otimes P_Y\,,\, G_n^* - C_{\theta_0}\otimes P_{n,Y}\Big) \Longrightarrow (\mathbb{A}_G\,,\,\mathbb{A}_G^\perp),$$

weakly in $\Big(\ell^\infty\big([0,1]^p \times \{1,\dots,m\}\big)\Big)^2$. By (iii) of Lemma 19, ψ_1 is Hadamard-differentiable and so, using the functional Delta-method, we deduce

$$\sqrt{n}\Big(\psi_1(G_n) - \psi_1(C_{\theta_0}\otimes P_Y)\,,\,\psi_1(G_n^*) - \psi_1(C_{\theta_0}\otimes P_{n,Y})\Big) \Longrightarrow \Big(\dot\psi_1(G)(\mathbb{A}_G)\,,\,\dot\psi_1(G)(\mathbb{A}_G^\perp)\Big).$$

By (ii) of Lemma 19, $\psi_1(C_{\theta_0} \otimes P_Y) = \psi_1(C_{\theta_0} \otimes P_{n,Y}) = 0$, implying

$$\sqrt{n}\Big(\psi_{1,1}(G_n), \dots, \psi_{m,1}(G_n)\,,\, \psi_{1,1}(G_n^*), \dots, \psi_{m,1}(G_n^*)\Big)$$
$$\Longrightarrow \Big(\dot{\psi}_{1,1}(G)(\mathbb{A}_G), \dots, \dot{\psi}_{m,1}(G)(\mathbb{A}_G)\,,\, \dot{\psi}_{1,1}(G)(\mathbb{A}_G^{\perp}), \dots, \dot{\psi}_{m,1}(G)(\mathbb{A}_G^{\perp})\Big).$$

By Slutsky's theorem, we have

$$\sqrt{n}\Big(\frac{\psi_{1,1}(G_n)}{\psi_{1,2}(G_n)}, \dots, \frac{\psi_{m,1}(G_n)}{\psi_{m,2}(G_n)}\,,\, \frac{\psi_{1,1}(G_n^*)}{\psi_{1,2}(G_n^*)}, \dots, \frac{\psi_{m,1}(G_n^*)}{\psi_{m,2}(G_n^*)}\Big)$$
$$\Longrightarrow \Big(\frac{\dot{\psi}_{1,1}(G)(\mathbb{A}_G)}{\psi_{1,2}(G)}, \dots, \frac{\dot{\psi}_{m,1}(G)(\mathbb{A}_G)}{\psi_{m,2}(G)}\,,\, \frac{\dot{\psi}_{1,1}(G)(\mathbb{A}_G^{\perp})}{\psi_{1,2}(G)}, \dots, \frac{\dot{\psi}_{m,1}(G)(\mathbb{A}_G^{\perp})}{\psi_{m,2}(G)}\Big).$$

By (i) of Lemma 19, the latter convergence result implies

$$\Big(\Theta_{n,1}, \dots, \Theta_{n,m}\,,\, \Theta_{n,1}^*, \dots, \Theta_{n,m}^*\Big)$$
$$\Longrightarrow \Big(\frac{\dot{\psi}_{1,1}(G)(\mathbb{A}_G)}{-\psi_{1,2}(G)}, \dots, \frac{\dot{\psi}_{m,1}(G)(\mathbb{A}_G)}{-\psi_{m,2}(G)}\,,\, \frac{\dot{\psi}_{1,1}(G)(\mathbb{A}_G^{\perp})}{-\psi_{1,2}(G)}, \dots, \frac{\dot{\psi}_{m,1}(G)(\mathbb{A}_G^{\perp})}{-\psi_{m,2}(G)}\Big)$$
$$=: \Big(\Theta_1, \dots, \Theta_m\,,\, \Theta_1^{\perp}, \dots, \Theta_m^{\perp}\Big).$$

Moreover, $\Big(\Theta_1, \dots, \Theta_m\Big)$ and $\Big(\Theta_1^{\perp}, \dots, \Theta_m^{\perp}\Big)$ are independent and identically distributed under $\mathcal{P}_n$, by construction. Because of (iv) of Lemma 19, $\Theta_{n,0}$ can asymptotically be seen as a mean of the $\Theta_{n,k}$ and this provides

$$\Theta_{n,0} = \frac{\sum_{k=1}^m \psi_{k,2}(G_n)\Theta_{n,k}}{\sum_{k=1}^m \psi_{k,2}(G_n)} \Longrightarrow \frac{\sum_{k=1}^m \psi_{k,2}(G)\Theta_k}{\sum_{k=1}^m \psi_{k,2}(G)} =: \Theta_0.$$

Therefore, by the continuous mapping theorem, we deduce

$$\Big(\Theta_{n,0}, \dots, \Theta_{n,m}\,,\, \Theta_{n,0}^*, \dots, \Theta_{n,m}^*\Big) \Longrightarrow \Big(\Theta_0, \dots, \Theta_m\,,\, \Theta_0^{\perp}, \dots, \Theta_m^{\perp}\Big),$$

and we still have that $\Big(\Theta_0, \dots, \Theta_m\Big)$ and $\Big(\Theta_0^{\perp}, \dots, \Theta_m^{\perp}\Big)$ are independent and identically distributed under $\mathcal{P}_n$.

Now, we will work under $\mathcal{P}_n^*$ the probability measure over $\big([0,1]^p \times \{1, \dots, m\}\big)^{\otimes 2n}$ whose density with respect to $\mathcal{P}_n$ is

$$\frac{d\mathcal{P}_n^*}{d\mathcal{P}_n}(\mathbf{z}_1, y_1, \dots, \mathbf{z}_n, y_n, \mathbf{z}_1^*, y_1^*, \dots, \mathbf{z}_n^*, y_n^*) = \prod_{i=1}^n \frac{c_{\hat{\theta}_0^b}(\mathbf{z}_i^*)}{c_{\theta_0}(\mathbf{z}_i^*)},$$

where $\hat{\theta}_0^b$ is the estimator of θ_0 when applied to the "sample" $(\mathbf{z}_1, y_1, \dots, \mathbf{z}_n, y_n)$. We remark that

$$\frac{d\mathcal{P}_n^*}{d\mathcal{P}_n}(\mathbf{Z}_1, Y_1, \dots, \mathbf{Z}_n, Y_n, \mathbf{Z}_1^*, Y_1^*, \dots, \mathbf{Z}_n^*, Y_n^*) = \exp(\ell_n).$$

Since we have shown that $\Theta_{n,0} \Longrightarrow \Theta_0$ under $\mathcal{P}_n$, use Lemma 20 and obtain

$$\ell_n = \Theta_0^T \mathbb{W}^{\perp} - \frac{1}{2}\Theta_0^T I_0 \Theta_0 + o_P(1).$$

Therefore, under $\mathcal{P}_n$, we have

$$\left(\frac{d\mathcal{P}_n^*}{d\mathcal{P}_n}, \Theta_{n,0}, \dots, \Theta_{n,m}, \Theta_{n,0}^*, \dots, \Theta_{n,m}^*\right) \Longrightarrow \big(\zeta, \Theta_0, \dots, \Theta_m, \Theta_0^{\perp}, \dots, \Theta_m^{\perp}\big),$$

where $\zeta := \exp\big(\Theta_0^T \mathbb{W}^{\perp} - \Theta_0^T I_0 \Theta_0/2\big)$. Note that $\mathbb{E}[\zeta] = \mathbb{E}[\mathbb{E}[\zeta|\Theta_0]] = 1$ because Θ_0 and $\mathbb{W}^{\perp}$ are independent, and $\mathbb{W}^{\perp} \sim \mathcal{N}(0, I_0)$. This corresponds to condition (iii) of Theorem

3.10.5 of Van der Waart and Wellner (1996), and we deduce $\mathcal{P}_n^*$ is contiguous with respect to $\mathcal{P}_n$. We can then apply Le Cam's Third Lemma (Theorem 3.10.7 of Van der Waart and Wellner (1996)), and we get that, under $\mathcal{P}_n^*$,

$$\left(\Theta_{n,0},\dots,\Theta_{n,m},\Theta_{n,0}^*,\dots,\Theta_{n,m}^*\right) \Longrightarrow \left(\tilde\Theta_0,\dots,\tilde\Theta_m,\Theta_0^*,\dots,\Theta_m^*\right),$$

where $\mathbb{E}[\chi(\tilde\Theta_{0:m},\Theta_{0:m}^*)] = \mathbb{E}[\zeta\chi(\Theta_{0:m},\Theta_{0:m}^\perp)]$ for any simple function χ. Choose w_1 and $w_2 \in \mathbb{R}^{m+1}$ and set $\Sigma := Var\left[\Theta_{0:m}\right]$. Then, we have

$$\begin{aligned}
&\mathbb{E}[\exp(iw_1^T\tilde\Theta_{0:m} + iw_2^T\Theta_{0:m}^*)] = \mathbb{E}[\zeta\exp(iw_1^T\Theta_{0:m} + iw_2^T\Theta_{0:m}^\perp)] \\
&= \mathbb{E}[\exp(\Theta_0^T\mathbb{W}^\perp - \Theta_0^TI_0\Theta_0/2 + iw_1^T\Theta_{0:m} + iw_2^T\Theta_{0:m}^\perp)] \\
&= \mathbb{E}\Big[\exp(iw_1^T\Theta_{0:m} - \Theta_0^TI_0\Theta_0/2)\,\mathbb{E}[\exp(\Theta_0^T\mathbb{W}^\perp + iw_2^T\Theta_{0:m}^\perp)\,|\,\Theta_{0:m}]\Big] \\
&= \mathbb{E}\left[\exp(iw_1^T\Theta_{0:m} - \Theta_0^TI_0\Theta_0/2)\exp\left(\frac{1}{2}\left(-w_2^T\Sigma w_2 + \Theta_0^TI_0\Theta_0 + 2iw_2\mathbb{E}[{\Theta_{0:m}^\perp}^T\mathbb{W}^\perp]\Theta_0\right)\right)\right] \\
&= \mathbb{E}\left[\exp\left(iw_1^T\Theta_{0:m} - w_2^T\Sigma w_2/2 + iw_2\mathbb{E}[{\Theta_{0:m}^\perp}^T\mathbb{W}^\perp]\Theta_0\right)\right] \\
&= \mathbb{E}\left[\exp\left(iw_1^T\Theta_{0:m} + iw_2\Theta_{0:m}^\perp + iw_2\mathbb{E}[{\Theta_{0:m}^\perp}^T\mathbb{W}^\perp]\Theta_0\right)\right].
\end{aligned}$$

Therefore, we have proven the following equality:

$$\left(\tilde\Theta_0,\dots,\tilde\Theta_m,\Theta_0^*,\dots,\Theta_m^*\right) \overset{\text{law}}{=} \left(\Theta_0,\dots,\Theta_m,\Theta_0^\perp + a_0\Theta_0,\dots,\Theta_m^\perp + a_m\Theta_0\right),$$

where $a_k = \mathbb{E}[{\Theta_k^\perp}^T\mathbb{W}^\perp]$. To finish the proof, it remains to show that $a_k = a_0$ for all $k \in \{1,\dots,m\}$, i.e.

$$\mathbb{E}[{\Theta_0^\perp}^T\mathbb{W}^\perp] = \mathbb{E}[{\Theta_k^\perp}^T\mathbb{W}^\perp].$$

First, we know from the proof of Lemma 19 that $\Theta_{k,n} = -\dot\psi_{k,1}(G)(\mathbb{A}_G)/\psi_{k,2}(G) + o_P(1)$, $k = 1,\dots,m$ and $\Theta_{0,n} = -\dot\psi_{0,1}(G)(\mathbb{A}_G)/\psi_{0,2}(G) + o_P(1)$, where

$$\psi_{0,1}(G) := \int \frac{\partial l}{\partial\theta}\left(\theta_0, \left(\frac{\int \mathbf{1}\{z_q^1 \le z_q^2, y^1 = y^2\}dG(\mathbf{z}^1,y^1)}{\int \mathbf{1}\{y^1 = y^2\}dG(\mathbf{z}^1,y^1)}\right)_{q=1,\dots,p}\right) dG(\mathbf{z}^2,y^2),$$

and

$$\psi_{0,2}(G) := \int \frac{\partial^2 l}{\partial\theta^2}\left(\theta_0, \left(\frac{\int \mathbf{1}\{z_q^1 \le z_q^2, y^1 = y^2\}dG(\mathbf{z}^1,y^1)}{\int \mathbf{1}\{y^1 = y^2\}dG(\mathbf{z}^1,y^1)}\right)_{q=1,\dots,p}\right) dG(\mathbf{z}^2,y^2).$$

This implies $\Theta_k = -\dot\psi_{k,1}(G)(\mathbb{A}_G)/\psi_{k,2}(G)$, $k = 0,\dots,m$.

Actually, the reasoning is exactly the same when dealing with $\Theta_{k,n}^*$ and $\Theta_k^\perp$, $k = 0,\dots,m$, replacing $\mathbb{A}_G$ by $\mathbb{A}_G^\perp$. We get

$$\Theta_k^\perp = -\frac{\dot\psi_{k,1}(G)(\mathbb{A}_G^\perp)}{\psi_{k,2}(G)}, \text{ and } \Theta_0^\perp = -\frac{\dot\psi_{0,1}(G)(\mathbb{A}_G^\perp)}{\psi_{0,2}(G)}.$$

Second, note that, when $k = 1,\dots,m$,

$$\begin{aligned}
\psi_{k,2}(G) &:= \int \frac{\partial^2 l}{\partial\theta^2}\left(\theta_0, \left(\frac{\int \mathbf{1}\{z_q^1 \le z_q^2, y^1 = k\}dG(\mathbf{z}^1,y^1)}{\int \mathbf{1}\{y^1 = k\}\,dG(\mathbf{z}^1,y^1)}\right)_{q=1,\dots,p}\right)\mathbf{1}\{y^2 = k\}\,dG(\mathbf{z}^2,y^2) \\
&= \mathbb{P}(Y = k)\int \frac{\partial^2 l}{\partial\theta^2}\left(\theta_0, \left(\int \mathbf{1}\{z_q^1 \le z_q^2\}dC_{\theta_0}(\mathbf{z}^1)\right)_{q=1,\dots,p}\right) dC_{\theta_0}(\mathbf{z}^2) \\
&= \mathbb{P}(Y = k)\int \frac{\partial^2 l}{\partial\theta^2}(\theta_0,\mathbf{z})\; dC_{\theta_0}(\mathbf{z}^2) \\
&= \mathbb{P}(Y = k)\psi_{0,2}(G).
\end{aligned}$$

Third, let us calculate $\dot{\psi}_{k,1}(G)(h)$, $k=0,1,\ldots,m$. From Lemma 19, we have

$$\dot{\psi}_{k,1}(G)(h) = \int \frac{\partial l}{\partial\theta}\left(\theta_0,\mathbf{z}^2\right)\mathbf{1}\{y^2=k\}\,dh(\mathbf{z}^2,y^2) + \sum_{j=1}^p \int \frac{\partial^2 l}{\partial\theta\,\partial z_j}\left(\theta_0,\mathbf{z}^2\right)\mathbf{1}\{y^2=k\}$$
$$\cdot\left(\frac{\int \mathbf{1}\{z_j^1\leq z_j^2, y^1=y^2\}dh(\mathbf{z}^1,y^1)}{\int \mathbf{1}\{y^1=y^2\}dG(\mathbf{z}^1,y^1)} - \frac{\int \mathbf{1}\{z_j^1\leq z_j^2, y^1=y^2\}dG(\mathbf{z}^1,y^1)\int \mathbf{1}\{y^1=y^2\}dh(\mathbf{z}^1,y^1)}{\left(\int \mathbf{1}\{y^1=y^2\}dG(\mathbf{z}^1,y^1)\right)^2}\right) dG(\mathbf{z}^2,y^2).$$

for $k=1,\ldots,m$. Since $G=C_{\theta_0}\otimes F_Y$, we can simplify the latter equalities:

$$\dot{\psi}_{k,1}(G)(h) = \int \frac{\partial l}{\partial\theta}\left(\theta_0,\mathbf{z}^2\right)\mathbf{1}\{y^2=k\}\,dh(\mathbf{z}^2,y^2) + \mathbb{P}(Y=k)\sum_{j=1}^p \int \frac{\partial^2 l}{\partial\theta\,\partial z_j}\left(\theta_0,\mathbf{z}^2\right)$$
$$\cdot\left(\frac{\int [\mathbf{1}\{z_j^1\leq z_j^2, y^1=y^2\} - z_j^2\mathbf{1}\{y^1=y^2\}]dh(\mathbf{z}^1,y^1)}{\int \mathbf{1}\{y^1=y^2\}dG(\mathbf{z}^1,y^1)}\right) dC_{\theta_0}(\mathbf{z}^2).$$

Since $\dot{\psi}_{0,1}(G)(h)=\sum_{k=1}^m \dot{\psi}_{k,1}(G)(h)$, we have

$$\dot{\psi}_{0,1}(G)(h) = \int \frac{\partial l}{\partial\theta}\left(\theta_0,\mathbf{z}^2\right)\,dh(\mathbf{z}^2,y^2) + \sum_{j=1}^p \int \frac{\partial^2 l}{\partial\theta\,\partial z_j}\left(\theta_0,\mathbf{z}^2\right)$$
$$\cdot\left(\frac{\int [\mathbf{1}\{z_j^1\leq z_j^2, y^1=y^2\} - z_j^2\mathbf{1}\{y^1=y^2\}]dh(\mathbf{z}^1,y^1)}{\int \mathbf{1}\{y^1=y^2\}dG(\mathbf{z}^1,y^1)}\right) dC_{\theta_0}(\mathbf{z}^2)dF_Y(\mathbf{y}^2).$$

Then, we can rewrite $\dot{\psi}_{k,1}(G)(h) = M_1(h,k)+\mathbb{P}(Y=k)M_3(h,k)$ and $\dot{\psi}_{0,1}(G)(h) = M_2(h)+\sum_{k'=1}^m \mathbb{P}(Y=k')M_3(h,k')$, where

$$M_1(h,k) := \int \frac{\partial l}{\partial\theta}\left(\theta_0,\mathbf{z}^2\right)\mathbf{1}\{y^2=k\}\,dh(\mathbf{z}^2,y^2),\quad M_2(h) := \int \frac{\partial l}{\partial\theta}\left(\theta_0,\mathbf{z}^2\right)\,dh(\mathbf{z}^2,y^2),$$
$$M_3(h,k) := \sum_{j=1}^p \int \frac{\partial^2 l}{\partial\theta\,\partial z_j}\left(\theta_0,\mathbf{z}^2\right)\left(\frac{\int [\mathbf{1}\{z_j^1\leq z_j^2, y^1=k\} - z_j^2\mathbf{1}\{y^1=k\}]dh(\mathbf{z}^1,y^1)}{\int \mathbf{1}\{y^1=k\}dG(\mathbf{z}^1,y^1)}\right) dC_{\theta_0}(\mathbf{z}^2).$$

Substituting h by $\mathbb{A}_G^\perp$, we get

$$\Theta_k^\perp = -\frac{\dot{\psi}_{k,1}(G)(\mathbb{A}_G^\perp)}{\psi_{k,2}(G)} = -\frac{M_1(\mathbb{A}_G^\perp,k)}{\mathbb{P}(Y=k)\psi_{0,2}(G)} - \frac{M_3(\mathbb{A}_G^\perp,k)}{\psi_{0,2}(G)},$$
$$\Theta_0^\perp = -\frac{\dot{\psi}_{0,1}(G)(\mathbb{A}_G^\perp)}{\psi_{0,2}(G)} = -\frac{M_2(\mathbb{A}_G^\perp)}{\psi_{0,2}(G)} - \frac{\sum_{k'=1}^m \mathbb{P}(Y=k')M_3(\mathbb{A}_G^\perp,k')}{\psi_{0,2}(G)}.$$

Fourth, since $\mathbb{W}^\perp$ is the weak limit of $\sum_{i=1}^n \frac{\partial l}{\partial\theta}(\theta_0,\mathbf{Z}_i^*)/\sqrt{n}$ under $\mathcal{P}_n$, this implies $\mathbb{W}^\perp = \dot{\psi}_3(G)(\mathbb{A}_G^\perp)$, with

$$\psi_3(G) = \int \frac{\partial l}{\partial\theta}(\theta_0,\mathbf{z})\,dG(\mathbf{z},y),\ \text{and}\ \dot{\psi}_3(G)(h) = \int \frac{\partial l}{\partial\theta}(\theta_0,\mathbf{z})\,dh(\mathbf{z},y).$$

Finally, by (i) of the following Lemma 21, we have $\mathbb{E}[M_1(\mathbb{A}_G^\perp,k)^T\mathbb{W}^\perp] = \mathbb{P}(Y=k)\mathbb{E}[M_2(\mathbb{A}_G^\perp)^T\mathbb{W}^\perp]$, By (ii) of the latter lemma, we have $\mathbb{E}[M_3(\mathbb{A}_G^\perp,k)^T\mathbb{W}^\perp] = \mathbb{E}[M_3(\mathbb{A}_G^\perp,k')^T\mathbb{W}^\perp]$ for all k and k'. Finally, we obtain $\mathbb{E}[{\Theta_0^\perp}^T\mathbb{W}^\perp] = \mathbb{E}[{\Theta_k^\perp}^T\mathbb{W}^\perp]$, which finishes the proof. □

Lemma 21. *Assume that $\overline{\mathcal{H}}_0^c$ is satisfied. Then,*

(i) For $k=1,\ldots,m$,

$$\mathbb{E}\left[\int \frac{\partial l}{\partial \theta^T}(\theta_0,\mathbf{z})\,\mathbf{1}\{y=k\}\,d\mathbb{A}_G^{\perp}(\mathbf{z},y)\int \frac{\partial l}{\partial \theta}\left(\theta_0,\mathbf{z}'\right)\,d\mathbb{A}_G^{\perp}(\mathbf{z}',y')\right]$$
$$= \mathbb{P}(Y=k)\,\mathbb{E}\left[\int \frac{\partial l}{\partial \theta^T}(\theta_0,\mathbf{z})\,d\mathbb{A}_G^{\perp}(\mathbf{z},y)\int \frac{\partial l}{\partial \theta}\left(\theta_0,\mathbf{z}'\right)\,d\mathbb{A}_G^{\perp}(\mathbf{z}',y')\right].$$

(ii) The expectations

$$\mathbb{E}\left[\int \frac{\partial^2 l}{\partial \theta^T\,\partial z_j}\left(\theta_0,\mathbf{z}^2\right)\left(\frac{\int[\mathbf{1}\{z_j^1\le z_j^2,y^1=k\}-z_j^2\mathbf{1}\{y^1=k\}]d\mathbb{A}_G^{\perp}(\mathbf{z}^1,y^1)}{\int \mathbf{1}\{y^1=k\}dG(\mathbf{z}^1,y^1)}\right)dC_{\theta_0}(\mathbf{z}^2)\right.$$
$$\left.\cdot\int \frac{\partial l}{\partial \theta}\left(\theta_0,\mathbf{z}^3\right)\,d\mathbb{A}_G^{\perp}(\mathbf{z}^3,y^3)\right]$$

do not depend on $k=1,\dots,m$.

Proof : (i) By simple calculations, we obtain

$$\mathbb{E}\left[\int \frac{\partial l}{\partial \theta^T}(\theta_0,\mathbf{z})\,\mathbf{1}\{y=k\}\,d\mathbb{A}_G^{\perp}(\mathbf{z},y)\int \frac{\partial l}{\partial \theta}\left(\theta_0,\mathbf{z}'\right)\,d\mathbb{A}_G^{\perp}(\mathbf{z}',y')\right]$$
$$=\int \frac{\partial l}{\partial \theta^T}(\theta_0,\mathbf{z})\,\mathbf{1}\{y=k\}\frac{\partial l}{\partial \theta}\left(\theta_0,\mathbf{z}'\right)d_{\mathbf{z},y,\;z',y'}\left(\mathbb{E}\left[\mathbb{A}_G^{\perp}(\mathbf{z},y)\mathbb{A}_G^{\perp}(\mathbf{z}',y')\right]\right)$$
$$=\int \frac{\partial l}{\partial \theta^T}(\theta_0,\mathbf{z})\,\mathbf{1}\{y=k\}\frac{\partial l}{\partial \theta}\left(\theta_0,\mathbf{z}'\right)\left\{\delta_{y'=y}d\mathbb{P}(y)[dC_{\theta_0}(\mathbf{z})\delta_{\mathbf{z}'=\mathbf{z}}+dC_{\theta_0}(\mathbf{z}')\delta_{\mathbf{z}=\mathbf{z}'}]\right.$$
$$\left.-dC_{\theta_0}(\mathbf{z})dC_{\theta_0}(\mathbf{z}')d\mathbb{P}(y)d\mathbb{P}(y')\right\}$$
$$=2\mathbb{P}(Y=k)\int \frac{\partial l}{\partial \theta^T}(\theta_0,\mathbf{z})\,\frac{\partial l}{\partial \theta}(\theta_0,\mathbf{z})\,dC_{\theta_0}(\mathbf{z})-\mathbb{P}(Y=k)\int \frac{\partial l}{\partial \theta^T}(\theta_0,\mathbf{z})\,dC_{\theta_0}(\mathbf{z})\cdot\int \frac{\partial l}{\partial \theta}(\theta_0,\mathbf{z})\,dC_{\theta_0}(\mathbf{z}).$$

By summing up the latter identities w.r.t. $k=1,\dots,m$, we prove (i).

(ii) For convenience, let us write $\phi_2(\mathbf{z}) := \partial^2 l/(\partial\theta^T\,\partial z_j)\,(\theta_0,\mathbf{z})$ and $\phi_3(\mathbf{z}) := \partial l(\theta_0,\mathbf{z})/\partial\theta^T$. We get the result if we prove that

$$A_{1,k} := \mathbb{E}\left[\int \phi_2(\mathbf{z}_2)\left(\frac{\int \mathbf{1}\{z_j^1\le z_j^2\}\mathbf{1}\{y^1=k\}d\mathbb{A}_G^{\perp}(\mathbf{z}^1,y^1)}{\int \mathbf{1}\{y^1=k\}dG(\mathbf{z}^1,y^1)}\right)dC_{\theta_0}(\mathbf{z}^2)\int \phi_3(\mathbf{z}_3)\,d\mathbb{A}_G^{\perp}(\mathbf{z}^3,y^3)\right] \text{ and}$$
$$A_{2,k} := \mathbb{E}\left[\int \phi_2(\mathbf{z}_2)\left(\frac{\int z_j^2\mathbf{1}\{y^1=k\}d\mathbb{A}_G^{\perp}(\mathbf{z}^1,y^1)}{\int \mathbf{1}\{y^1=k\}dG(\mathbf{z}^1,y^1)}\right)dC_{\theta_0}(\mathbf{z}^2)\int \phi_3(\mathbf{z}_3)\,d\mathbb{A}_G^{\perp}(\mathbf{z}^3,y^3)\right]$$

do not depend on k. We will do the task for $A_{1,k}$, $k=1,\dots,m$, and the calculations will be similar for $A_{2,k}$. Note that

$$A_{1,k}=\frac{1}{\mathbb{P}(Y=k)}\int \phi_2(\mathbf{z}_2)\mathbf{1}\{z_j^1\le z_j^2\}\mathbf{1}\{y^1=k\}\phi_3(\mathbf{z}_3)\,dC_{\theta_0}(\mathbf{z}^2)d_{\mathbf{z}^1,y^1,\mathbf{z}^3,y^3}\mathbb{E}\left[\mathbb{A}_G^{\perp}(\mathbf{z}^1,y^1)\mathbb{A}_G^{\perp}(\mathbf{z}^3,y^3)\right]$$
$$=\frac{1}{\mathbb{P}(Y=k)}\int \phi_2(\mathbf{z}_2)\mathbf{1}\{z_j^1\le z_j^2\}\mathbf{1}\{y^1=k\}\phi_3(\mathbf{z}_3)\,dC_{\theta_0}(\mathbf{z}^2)$$
$$\left\{\delta_{y^3=y^1}d\mathbb{P}(y^1)[dC_{\theta_0}(\mathbf{z}^1)\delta_{\mathbf{z}^3=\mathbf{z}^1}+dC_{\theta_0}(\mathbf{z}^3)\delta_{\mathbf{z}^1=\mathbf{z}^3}]-C_{\theta_0}(\mathbf{z}^1)dC_{\theta_0}(\mathbf{z}^3)d\mathbb{P}(y^1)d\mathbb{P}(y^3)\right\}.$$

We deduce that

$$A_{1,k}=2\int \phi_2(\mathbf{z}_2)\mathbf{1}\{z_j^1\le z_j^2\}\phi_3(\mathbf{z}^1)\,dC_{\theta_0}(\mathbf{z}^1)dC_{\theta_0}(\mathbf{z}^2)$$
$$-\int \phi_2(\mathbf{z}_2)\mathbf{1}\{z_j^1\le z_j^2\}\phi_3(\mathbf{z}^3)\,dC_{\theta_0}(\mathbf{z}^1)dC_{\theta_0}(\mathbf{z}^2)dC_{\theta_0}(\mathbf{z}^3),$$

that does not depend on k. □

A.3 Proof of Proposition 16

As usual with the nonparametric bootstrap, we rewrite the bootstrapped empirical process by counting the number of times every observation of the initial sample is drawn:

$$d\overline{G}_n^* = \frac{1}{n}\sum_{i=1}^n M_{n,i}\delta_{(\mathbf{Z}_i^*,\mathbf{X}_{i,J})},$$

where $M_{n,i}$ denotes the number of times $(\mathbf{X}_{i,J}^*)$ has been redrawn in a n-size bootstrap resampling with replacement. It is well-known that $M_n := (M_{n,1},\ldots,M_{n,n})$ follows a multinomial distribution $\mathcal{M}(n,n^{-1},\ldots,n^{-1})$: its mean is n and the associated probabilities are $1/n,\ldots,1/n$. In other words, $\overline{\mathbb{G}}_n^*(\mathbf{z},\mathbf{x}_J) = \frac{1}{\sqrt{n}}\sum_{i=1}^n M_{n,i}\{\mathbf{1}((\mathbf{Z}_i^*,\mathbf{X}_{i,J}) \leq (\mathbf{z},\mathbf{x}_J)) - C_{\theta_0}(\mathbf{z})F_{n,J}(\mathbf{x}_J)\}$.

We can remove the dependence between the random components $M_{n,i}$, $i=1,\ldots,n$ by a "Poissonization" procedure. We mimic van der Vaart and Wellner (1996), p.346: instead of drawing n times the initial observations, this is done N_n times, where N_n follows a Poisson distribution with mean n and N_n is independent of the initial sample. Then, the n variables $M_{N_n,1},\ldots,M_{N_n,n}$ are i.i.d. Poisson random variables with mean one. And we can build the new process as

$$\tilde{\mathbb{G}}_n^*(\mathbf{z},\mathbf{x}_J) := \frac{1}{\sqrt{n}}\sum_{i=1}^n M_{N_n,i}\{\mathbf{1}_{(\mathbf{Z}_i^*,\mathbf{X}_{i,J})\leq(\mathbf{z},\mathbf{x}_J)} - C_{\theta_0}(\mathbf{z})F_{n,J}(\mathbf{x}_J)\}.$$

Actually, the distance between $\overline{\mathbb{G}}_n^*$ and $\tilde{\mathbb{G}}_n^*$ is negligible. Indeed, for every $(\mathbf{z},\mathbf{x}_J)$,

$$\Delta_n(\mathbf{z},\mathbf{x}_J) := (\tilde{\mathbb{G}}_n^* - \overline{\mathbb{G}}_n^*)(\mathbf{z},\mathbf{x}_J) = \frac{1}{\sqrt{n}}\sum_{i=1}^n (M_{N_n,i}-M_{n,i})\{\mathbf{1}_{(\mathbf{Z}_i^*,\mathbf{X}_{i,J})\leq(\mathbf{z},\mathbf{x}_J)} - C_{\theta_0}(\mathbf{z})F_{n,J}(\mathbf{x}_J)\}$$

is centered. Moreover, by independence between the observations and by the resampling scheme, we have

$$\begin{aligned}
\mathbb{E}[\|\Delta_n\|_\infty^2] = \mathbb{E}[\sup_{\mathbf{z},\mathbf{x}_J}\Delta_n^2(\mathbf{z},\mathbf{x}_J)] &\leq \frac{1}{n^2}\sum_{i,j=1}^n \mathbb{E}[|(M_{N_n,i}-M_{n,i})(M_{N_n,j}-M_{n,j})|] \\
&\leq \frac{1}{n^2}\sum_{i,j=1}^n \mathbb{E}[(M_{N_n,i}-M_{n,i})^2]^{1/2}\mathbb{E}[(M_{N_n,j}-M_{n,j})^2]^{1/2} \\
&\leq \mathbb{E}[(M_{N_n,1}-M_{n,1})^2],
\end{aligned}$$

because the sequence $(M_{N_n,i}-M_{n,i})_{i=1,\ldots,n}$ is exchangeable. Given $N_n = k$, the i-th variable $|M_{N_n,i}-M_{n,i}|$ is binomial with the parameters $(|k-n|,1/n)$, i.e.

$$P(|M_{k,i}-M_{n,i}| = l) = C_{|k-n|}^l \frac{1}{n^l}\left(1-\frac{1}{n}\right)^{|k-n|-l},\ l = 0,\ldots,|k-n|.$$

Therefore, we obtain

$$\mathbb{E}[(M_{N_n,i}-M_{n,i})^2] = \sum_{k=0}^\infty \exp(-n)\frac{n^k}{k!}\left\{\frac{|k-n|}{n}(1-\frac{1}{n}) + \left(\frac{|k-n|}{n}\right)^2\right\}.$$

Simple calculations provide

$$\sum_{k=0}^\infty \exp(-n)\frac{n^k}{k!}\frac{|k-n|}{n} = \frac{2n^n}{n!}\exp(-n) \sim \left(\frac{2}{\pi n}\right)^{1/2},$$

by Stirling's formula, and

$$\sum_{k=0}^{\infty} \exp(-n)\frac{n^k}{k!}\left(\frac{k-n}{n}\right)^2 = \frac{\exp(-n)}{n^2}\sum_{k=0}^{\infty}\frac{n^k}{k!}(k(k-1)+k(1-2n)+n^2) = \frac{1}{n}\cdot$$

We deduce $\mathbb{E}[(M_{N_n,i}-M_{n,i})^2] = O(n^{-1/2})$ and $\mathbb{P}\left(\|\Delta_n\|_\infty > \epsilon\right) \longrightarrow 0$, when n tends to the infinity, given almost all sequences $\mathcal{S}_n := (\mathbf{Z}_i, \mathbf{X}_{i,J})_{i=1,\dots,n}$. This means that we can safely replace $\overline{\mathbb{G}}_n^*$ by $\tilde{\mathbb{G}}_n^*$, and the theorem follows if we prove the weak convergence of $(\overline{\mathbb{G}}_n, \tilde{\mathbb{G}}_n^*)$.

Note that we can rewrite

$$\begin{aligned}
\tilde{\mathbb{G}}_n^*(\mathbf{z},\mathbf{x}_J) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n (M_{N_n,i}-1)\left\{\mathbf{1}(\mathbf{Z}_i^*, X_{i,J}) \le (\mathbf{z},\mathbf{x}_J) - C_{\theta_0}(\mathbf{z})F_J(\mathbf{x}_J)\right\} \\
&+ \frac{1}{\sqrt{n}}\sum_{i=1}^n \left\{\mathbf{1}(\mathbf{Z}_i^*, X_{i,J}) \le (\mathbf{z},\mathbf{x}_J) - C_{\theta_0}(\mathbf{z})F_J(\mathbf{x}_J)\right\} - C_{\theta_0}(\mathbf{z})\sqrt{n}\left(F_{n,J}-F_J\right)(\mathbf{x}_J) \\
&+ \left(1-\frac{1}{n}\sum_{i=1}^n M_{N_n,i}\right) C_{\theta_0}(\mathbf{z})\sqrt{n}\left(F_{n,J}-F_J\right)(\mathbf{x}_J) \\
&:= \tilde{\mathbb{G}}_{n,1}^*(\mathbf{z},\mathbf{x}_J) + \tilde{\mathbb{G}}_{n,2}^*(\mathbf{z},\mathbf{x}_J) - \mathbb{G}_{n,3}(\mathbf{z},\mathbf{x}_J) + R_n(\mathbf{z},\mathbf{x}_J).
\end{aligned}$$

Obviously, the last remaining term is $o_P(1)$ uniformly w.r.t. $(\mathbf{z},\mathbf{x}_J)$, and it can be forgotten. Moreover, since the variables $(M_{N_n,i}-1)_{i=1,\dots,n}$ are i.i.d., centered with variance one and independent of the data, we can invoke some multiplier bootstrap results. Consider we live in the space $\mathcal{W} := [0,1]^p \times [0,1]^p \times \mathbb{R}^{d-p}$ that is related to our observations $W_i := (\mathbf{Z}_i, \mathbf{Z}_i^*, \mathbf{X}_{i,J})$, $i=1,\dots,n$. The true distribution of W_i under the null is P_W, whose cdf is $C_{\theta_0}\otimes C_{\theta_0}\otimes F_J$. Applying Corollary 2.9.3. in van der Vaart and Wellner (1996), the sequence of processes

$$(\mathbb{W}_n, \mathbb{W}_n^*) := \left(n^{-1/2}\sum_{i=1}^n (\delta_{W_i}-P_W), n^{-1/2}\sum_{i=1}^n (M_{N_n,i}-1)(\delta_{W_i}-P_W)\right)$$

converges weakly in $\ell^\infty(\mathcal{F})\times\ell^\infty(\mathcal{F})$ to a vector of independent Gaussian processes, where $\mathcal{F}$ denotes any Donsker class of measurable functions from $\mathcal{W}$ to $\mathbb{R}$.

Now, let us consider the class $\mathcal{F}$ of functions

$$f_{\mathbf{z}_0,\mathbf{z}'_0,\mathbf{x}_{J,0}} : (\mathbf{z},\mathbf{z}',\mathbf{x}_J) \mapsto \mathbf{1}(\mathbf{z}\le\mathbf{z}_0, \mathbf{z}'\le\mathbf{z}'_0, \mathbf{x}_J \le \mathbf{x}_{J,0}),$$

for any triplet $(\mathbf{z}_0,\mathbf{z}'_0,\mathbf{x}_{J,0})$ in $[0,1]^p\times[0,1]^p\times\overline{\mathbb{R}}^{d-p}$. Note that $\mathcal{F}$ is Donsker, that $\tilde{\mathbb{G}}_{n,1}^*(\mathbf{z},\mathbf{x}_J) = \mathbb{W}_n^* f_{1,\mathbf{z},\mathbf{x}_J}$, $\tilde{\mathbb{G}}_{n,2}^*(\mathbf{z},\mathbf{x}_J) = \mathbb{W}_n f_{1,\mathbf{z},\mathbf{x}_J}$ and that $\tilde{\mathbb{G}}_{n,3}^*(\mathbf{z},\mathbf{x}_J) = C_{\theta_0}(\mathbf{z})\mathbb{W}_n f_{1,1,\mathbf{x}_J}$. By the permanence of the Donsker property (see Section 2.10 in van der Vaart and Wellner, 1996), and the continuity of C_{θ_0}, the process $\tilde{\mathbb{G}}_n^*$ converges in $\ell^\infty([0,1]^p\times\mathbb{R}^{d-p})$ to a gaussian process $\overline{\mathbb{A}}^\perp$. Obviously, $\overline{\mathbb{G}}_n$ tends in distribution in $\ell^\infty([0,1]^p\times\mathbb{R}^{d-p})$ to a Gaussian process $\overline{\mathbb{A}}$, whose covariance function is given by

$$\mathbb{E}\left[\overline{\mathbb{G}}_n(\mathbf{z},\mathbf{x}_J)\overline{\mathbb{G}}_n(\mathbf{z}',\mathbf{x}'_J)\right] = C_{\theta_0}(\mathbf{z}\wedge\mathbf{z}')F_J(\mathbf{x}_J\wedge\mathbf{x}'_J) - C_{\theta_0}(\mathbf{z})F_J(\mathbf{x}_J)C_{\theta_0}(\mathbf{z}')F_J(\mathbf{x}'_J),$$

for every $\mathbf{z},\mathbf{z}',\mathbf{x}_J,\mathbf{x}'_J)$. By some standard calculations, we check that $\mathbb{E}[\tilde{\mathbb{G}}_n^*(\mathbf{z},\mathbf{x}_J)\tilde{\mathbb{G}}_n^*(\mathbf{z}',\mathbf{x}'_J)] = \mathbb{E}[\overline{\mathbb{G}}_n(\mathbf{z},\mathbf{x}_J)\overline{\mathbb{G}}_n(\mathbf{z}',\mathbf{x}'_J)]$ for every couples $(\mathbf{z},\mathbf{x}_J)$ and $(\mathbf{z}',\mathbf{x}'_J)$, implying that $\overline{\mathbb{A}}$ and $\overline{\mathbb{A}}^\perp$ have the same covariance functions. Moreover, the two limiting processes $\overline{\mathbb{A}}$ and $\overline{\mathbb{A}}^\perp$ are uncorrelated because

$$\mathbb{E}[\overline{\mathbb{G}}_n(\mathbf{z},\mathbf{x}_J)\tilde{\mathbb{G}}_n^*(\mathbf{z}',\mathbf{x}'_J)] = \mathbb{E}[\overline{\mathbb{G}}_n(\mathbf{z},\mathbf{x}_J)\mathbb{E}[\tilde{\mathbb{G}}_n^*(\mathbf{z}',\mathbf{x}'_J)|\mathcal{S}_n]] = 0,$$

for every couples $(\mathbf{z},\mathbf{x}_J)$ and $(\mathbf{z}',\mathbf{x}'_J)$. Therefore, the $\overline{\mathbb{A}}$ and $\overline{\mathbb{A}}^\perp$ are two independent versions of the same Gaussian process.

Remark 22. *If there were no resampling of the observations* $\mathbf{X}_{i,J}$ *at the first level, this would no longer be true. Indeed, the corresponding bootstrapped process would be given by*

$$\mathbb{G}_n^{**}(\mathbf{z},\mathbf{x}_J) := \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{1}(\mathbf{X}_{i,J} \leq \mathbf{x}_J)\left\{\mathbf{1}(\mathbf{Z}_i^* \leq \mathbf{z}) - C_{\theta_0}(\mathbf{z})\right\},$$

implying

$$\mathbb{E}\left[\overline{\mathbb{G}}_n^{**}(\mathbf{z},\mathbf{x}_J)\overline{\mathbb{G}}_n^{**}(\mathbf{z}',\mathbf{x}'_J)\right] = F_J(\mathbf{x}_J \wedge \mathbf{x}'_J)[C_{\theta_0}(\mathbf{z}\wedge\mathbf{z}') - C_{\theta_0}(\mathbf{z})C_{\theta_0}(\mathbf{z}')],$$

that is different of $\mathbb{E}\left[\overline{\mathbb{G}}_n(\mathbf{z},\mathbf{x}_J)\overline{\mathbb{G}}_n(\mathbf{z}',\mathbf{x}'_J)\right]$.

To conclude, we apply Corollary 1.4.5. in van der Vaart and Wellner (1996): for every bounded nonnegative Lipschitz function h and $\tilde{h}$,

$$\begin{aligned}\mathbb{E}[h(\overline{\mathbb{G}}_n)\tilde{h}(\tilde{\mathbb{G}}_n^*)] - \mathbb{E}[h(\overline{\mathbb{A}})\tilde{h}(\overline{\mathbb{A}}^{\perp})] &= \mathbb{E}[h(\overline{\mathbb{G}}_n)\left(\mathbb{E}[\tilde{h}(\tilde{\mathbb{G}}_n^*)|\mathcal{S}_n] - \mathbb{E}[\tilde{h}(\overline{\mathbb{A}}^{\perp})]\right)] \\ + \quad & \mathbb{E}[\big(h(\overline{\mathbb{G}}_n) - \mathbb{E}[h(\overline{\mathbb{A}})]\big)]\mathbb{E}[\tilde{h}(\overline{\mathbb{A}}^{\perp})].\end{aligned}$$

The first (resp. second) term tends to zero by the weak convergence of $\tilde{\mathbb{G}}_n^*$ (resp. $\overline{\mathbb{G}}_n$). This concludes the proof. □