Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2008, Vol. 44, No. 6, 1096–1127
DOI: 10.1214/07-AIHP148


ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

# Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models

Christian Genest[a] and Bruno Rémillard[b]

[a]*Département de mathématiques et de statistique, Université Laval, Québec (Québec), Canada G1K 7P4. E-mail: Christian.Genest@mat.ulaval.ca*

[b]*Service d'enseignement des méthodes quantitatives de gestion, HEC Montréal, 3000, chemin de la Côte-Sainte-Catherine, Montréal (Québec), Canada H3T 2A7. E-mail: Bruno.Remillard@hec.ca*

Received 24 May 2006; revised 3 July 2007; accepted 25 September 2007

---

**Abstract.** In testing that a given distribution $P$ belongs to a parameterized family $\mathcal{P}$, one is often led to compare a nonparametric estimate $A_n$ of some functional $A$ of $P$ with an element $A_{\theta_n}$ corresponding to an estimate $\theta_n$ of $\theta$. In many cases, the asymptotic distribution of goodness-of-fit statistics derived from the process $n^{1/2}(A_n - A_{\theta_n})$ depends on the unknown distribution $P$. It is shown here that if the sequences $A_n$ and $\theta_n$ of estimators are regular in some sense, a parametric bootstrap approach yields valid approximations for the $P$-values of the tests. In other words if $A_n^*$ and $\theta_n^*$ are analogs of $A_n$ and $\theta_n$ computed from a sample from $P_{\theta_n}$, the empirical processes $n^{1/2}(A_n - A_{\theta_n})$ and $n^{1/2}(A_n^* - A_{\theta_n^*})$ then converge jointly in distribution to independent copies of the same limit. This result is used to establish the validity of the parametric bootstrap method when testing the goodness-of-fit of families of multivariate distributions and copulas. Two types of tests are considered: certain procedures compare the empirical version of a distribution function or copula and its parametric estimation under the null hypothesis; others measure the distance between a parametric and a nonparametric estimation of the distribution associated with the classical probability integral transform. The validity of a two-level bootstrap is also proved in cases where the parametric estimate cannot be computed easily. The methodology is illustrated using a new goodness-of-fit test statistic for copulas based on a Cramér–von Mises functional of the empirical copula process.

**Résumé.** Pour tester qu'une loi $P$ donnée provient d'une famille paramétrique $\mathcal{P}$, on est souvent amené à comparer une estimation non paramétrique $A_n$ d'une fonctionnelle $A$ de $P$ à un élément $A_{\theta_n}$ correspondant à une estimation $\theta_n$ de $\theta$. Dans bien des cas, la loi asymptotique de statistiques de tests bâties à partir du processus $n^{1/2}(A_n - A_{\theta_n})$ dépend de la loi inconnue $P$. On montre ici que si les suites $A_n$ et $\theta_n$ d'estimateurs sont régulières dans un sens précis, le recours au rééchantillonnage paramétrique conduit à des approximations valides des seuils des tests. Autrement dit si $A_n^*$ et $\theta_n^*$ sont des analogues de $A_n$ et $\theta_n$ déduits d'un échantillon de loi $P_{\theta_n}$, les processus empiriques $n^{1/2}(A_n - A_{\theta_n})$ et $n^{1/2}(A_n^* - A_{\theta_n^*})$ convergent alors conjointement en loi vers des copies indépendantes de la même limite. Ce résultat est employé pour valider l'approche par rééchantillonnage paramétrique dans le cadre de tests d'adéquation pour des familles de lois et de copules multivariées. Deux types de tests sont envisagés : les uns comparent la version empirique d'une loi ou d'une copule et son estimation paramétrique sous l'hypothèse nulle ; les autres mesurent la distance entre les estimations paramétrique et non paramétrique de la loi associée à la transformation intégrale de probabilité classique. La validité du rééchantillonnage à deux degrés est aussi démontrée dans les cas où l'estimation paramétrique est difficile à calculer. La méthodologie est illustrée au moyen d'un nouveau test d'adéquation de copules fondé sur une fonctionnelle de Cramér–von Mises du processus de copule empirique.

*MSC:* 62F05; 62F40; 62H15

*Keywords:* Copula; Goodness-of-fit test; Monte Carlo simulation; Parametric bootstrap; $P$-values; Semiparametric estimation

---

## 1. Introduction

Given independent copies $X_1, \dots, X_n$ of a random vector $X$ with cumulative distribution function $F : \mathbb{R}^d \to \mathbb{R}$, suppose that it is desired to test

$$H_0\colon\ F \in \mathcal{F} = \{F_\theta\colon\ \theta \in \mathcal{O}\},$$

the hypothesis that $F$ comes from a parametric family of distributions whose members are indexed by a parameter $\theta$ belonging to an open set $\mathcal{O} \subset \mathbb{R}^p$. To achieve this goal, a natural way to proceed consists of measuring the difference between the empirical distribution function, defined for all $x \in \mathbb{R}^d$ by

$$F_n(x) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}(X_i \le x), \tag{1}$$

and a parametric estimate $F_{\theta_n}$ of $F$ derived under $H_0$ from some consistent estimate $\theta_n = T_n(X_1, \dots, X_n)$ of the true parameter value $\theta_0$. Here and in the sequel, inequalities between vectors are taken to hold componentwise.

Cramér–von Mises, Kolmogorov–Smirnov and many other standard goodness-of-fit procedures are based on statistics expressed as continuous functionals $S_n = \phi(\mathbb{G}_n^F)$ of the empirical process

$$\mathbb{G}_n^F = n^{1/2}(F_n - F_{\theta_n}).$$

Formal tests, however, require knowledge of the asymptotic null distribution of $S_n$, which often depends on the unknown value of $\theta$.

### 1.1. *The parametric bootstrap*

To solve this problem, Stute et al. [26] suggest the following "parametric bootstrap" procedure.

For some large integer $N$ and every $k \in \{1, \dots, N\}$, repeat the steps below:

(a) Given $\theta_n = T_n(X_1, \dots, X_n)$, generate $n$ independent observations $X_{1,k}^*, \dots, X_{n,k}^*$ from distribution $F_{\theta_n}$.
(b) Compute $\theta_{n,k}^* = T_n(X_{1,k}^*, \dots, X_{n,k}^*)$ and for each $x \in \mathbb{R}^d$, let

$$F_{n,k}^*(x) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}\bigl(X_{i,k}^* \le x\bigr).$$

(c) Compute $S_{n,k}^* = \phi(\mathbb{G}_{n,k}^{F^*})$, where

$$\mathbb{G}_{n,k}^{F^*} = n^{1/2}\bigl(F_{n,k}^* - F_{\theta_{n,k}^*}\bigr).$$

With the convention that large values of $S_n$ lead to the rejection of $H_0$, Stute et al. [26] show that under appropriate regularity conditions, an approximate $P$-value for the test is given by

$$\frac{1}{N}\sum_{k=1}^N \mathbf{1}\bigl(S_{n,k}^* > S_n\bigr).$$

Henze [20] obtained a similar result in the univariate discrete case. In both papers, the validity of the parametric bootstrap stems from the fact that under $H_0$ and as $n \to \infty$, $(S_n, S_{n,1}^*, \dots, S_{n,N}^*)$ converges weakly to a vector $(S, S_1^*, \dots, S_N^*)$ of mutually independent and identically distributed random variables.

### 1.2. *Motivation for the present work*

This investigation was motivated by the need to test the appropriateness of various dependence structures on the basis of a random sample

$$X_1 = (X_{11}, \dots, X_{1d}), \qquad \dots, \qquad X_n = (X_{n1}, \dots, X_{nd})$$

from a continuous random vector $X$ with cumulative distribution function $F$. Specifically, denote by $F_1, \dots, F_d$ the univariate margins of $X$ and let $C : [0, 1]^d \to [0, 1]$ be the copula for which Sklar's representation

$$F(x_1, \dots, x_d) = C\big\{F_1(x_1), \dots, F_d(x_d)\big\}$$

holds for all $x_1, \dots, x_d \in \mathbb{R}$. In fact, $C$ is simply the cumulative distribution function of $U = \xi(X)$, where $\xi : \mathbb{R}^d \to \mathbb{R}^d$ is defined for all $x_1, \dots, x_d \in \mathbb{R}$ by

$$\xi(x_1, \dots, x_d) = \big(F_1(x_1), \dots, F_d(x_d)\big). \tag{2}$$

Unless the margins are known, the vectors $U_1 = \xi(X_1), \dots, U_n = \xi(X_n)$ cannot be observed. However, a consistent estimate of $F_j$ is defined for all $t \in \mathbb{R}$ and $j \in \{1, \dots, d\}$ by

$$F_{jn}(t) = \frac{1}{n+1} \sum_{i=1}^{n} \mathbf{1}(X_{ij} \le t).$$

This uncommon choice of normalization is used because $F_{jn}$ serves later as an argument in score functions and pseudo-likelihoods that could blow up at 1. Letting

$$\xi_n(x_1, \dots, x_d) = \big(F_{1n}(x_1), \dots, F_{dn}(x_d)\big)^\top, \tag{3}$$

for all $x_1, \dots, x_d \in \mathbb{R}$, one could thus base a test of the hypothesis

$$H_0\colon\ C \in \mathcal{C} = \{C_\theta\colon\ \theta \in \mathcal{O}\} \tag{4}$$

on the pseudo-observations $\hat{U}_1 = \xi_n(X_1), \dots, \hat{U}_n = \xi_n(X_n)$. Various options are possible; two of them are briefly described below.

*Tests based on the empirical copula*

Hypothesis (4) could be tested using a Cramér–von Mises or Kolmogorov–Smirnov statistic $S_n = \phi(\mathbb{G}_n^C)$ with

$$\mathbb{G}_n^C = n^{1/2}(C_n - C_{\theta_n}),$$

where $C_{\theta_n}$ is a parametric estimate of $C_\theta$ derived from the estimation $\theta_n = T(X_1, \dots, X_n)$ of $\theta$ under $H_0$ while $C_n$ is the empirical copula, defined for all $u \in [0, 1]^d$ by

$$C_n(u) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(\hat{U}_i \le u). \tag{5}$$

This possibility is raised but quickly dismissed by Fermanian [10], due to the complexity of the weak limit of $\mathbb{G}_n^C$. See, e.g., [11,12,27] for derivations of the limit of the related empirical copula process $n^{1/2}(C_n - C_{\theta_0})$.

*Tests based on Kendall's distribution*
Another avenue explored by Wang and Wells [29] and Genest et al. [15] is to construct a test of hypothesis (4) on Kendall's distribution, i.e., the distribution function $K$ of the probability integral transform $W = F(X)$. Using the fact that one can also write $W = C(U)$, Genest and Rivest [17] and Barbe et al. [1] show that a consistent estimate of $K$ is given by the empirical distribution $K_n$ of the pseudo-observations $\hat{W}_1 = C_n(\hat{U}_1), \ldots, \hat{W}_n = C_n(\hat{U}_n)$. The latter is defined for all $w \in [0, 1]$ by

$$K_n(w) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(\hat{W}_i \le w). \tag{6}$$

Thus if $K_\theta$ denotes the distribution of $W$ when $C = C_\theta \in \mathcal{C}$, and if $K_{\theta_n}$ is a parametric estimate of $K_\theta$ derived from $\theta_n = T(X_1, \ldots, X_n)$ under the subsidiary hypothesis

$$H_0\colon\ K \in \mathcal{K} = \{K_\theta\colon\ \theta \in \mathcal{O}\}, \tag{7}$$

a goodness-of-fit test could rely on a continuous functional $S_n = \phi(\mathbb{G}_n^K)$ of

$$\mathbb{G}_n^K = n^{1/2}(K_n - K_{\theta_n}).$$

Whether hypothesis (4) is tested using $\mathbb{G}_n^C$ or the subsidiary hypothesis (7) is tested using $\mathbb{G}_n^K$, the limiting distribution of the test statistic $S_n$ does not only depend on the unknown parameter $\theta$ but also possibly on the nuisance parameters $F_1, \ldots, F_d$. Therefore, while the use of a parametric bootstrap may very well yield valid $P$-values, this conclusion cannot be reached on the basis of the results reported by Stute et al. [26], because of the presence of dependence among the sets of pseudo-observations $\hat{U}_1, \ldots, \hat{U}_n$ and $\hat{W}_1, \ldots, \hat{W}_n$.

### 1.3. *Objective and outline of the paper*

The purpose of this work is to establish the validity of the parametric bootstrap in situations where the hypothesis to be tested concerns the distribution $P$ of an unobservable $s$-variate random vector $U$, viz.

$$H_0\colon\ P \in \mathcal{P} = \{P_\theta\colon\ \theta \in \mathcal{O}\},$$

where $\mathcal{O}$ is an open subset of $\mathbb{R}^p$. Although $U$ cannot be seen, it is assumed that $U = \xi(X)$ for some function $\xi : \mathbb{R}^d \to \mathbb{R}^s$ of an observable $d$-variate random vector $X$, and that a consistent estimator $\xi_n$ of $\xi$ can be constructed from independent copies $X_1, \ldots, X_n$ of $X$.

In order to encompass procedures based on $\mathbb{G}_n^C$ and $\mathbb{G}_n^K$ as special cases, suppose that a test of $H_0$ is to be derived from a continuous functional $S_n = \phi(\mathbb{G}_n^A)$ of an abstract empirical process of the form

$$\mathbb{G}_n^A = n^{1/2}(A_n - A_{\theta_n}).$$

Here, $A_{\theta_n}$ and $A_n$ stand respectively for a parametric and a nonparametric estimate of an abstract quantity $A$ that depends on $P$. More specifically, $A$ is taken to be a function mapping a closed rectangle $\mathcal{T} \subset [-\infty, \infty]^r$ into $\mathbb{R}^s$, and $A_\theta$ denotes the form taken by $A$ when $P = P_\theta$ for some $\theta \in \mathcal{O}$. Thus for the test based on $\mathbb{G}_n^C$, one has $\mathcal{T} = [0, 1]^d$, $r = s = d$ and $A_\theta = C_\theta$; similarly, $\mathcal{T} = [0, 1]$, $r = s = 1$ and $A_\theta = K_\theta$ for the test based on $\mathbb{G}_n^K$.

The result to be shown here is that the parametric bootstrap yields a valid approximation to the null distribution of the empirical process $\mathbb{G}_n^A$ under appropriate conditions. The main requirements concern the large-sample behavior of the estimators $A_n$ of $A$ and $\theta_n$ of $\theta$ that are constructed from the pseudo-observations $\hat{U}_1 = \xi_n(X_1), \ldots, \hat{U}_n = \xi_n(X_n)$. In particular, the process $\Theta_n = n^{1/2}(\theta_n - \theta)$ needs to converge weakly, as $n \to \infty$, to a centered random variable $\Theta$. This is denoted symbolically

$$\Theta_n = n^{1/2}(\theta_n - \theta) \rightsquigarrow \Theta. \tag{8}$$

Similarly, it must be that, as $n \to \infty$,

$$\mathbb{A}_n = n^{1/2}(A_n - A) \rightsquigarrow \mathbb{A}, \tag{9}$$

i.e., $\mathbb{A}_n$ converges weakly to a centered process $\mathbb{A}$ in the space $\mathcal{D}(\mathcal{T};\mathbb{R}^s)$ of càdlàg processes from $\mathcal{T}$ to $\mathbb{R}^s$, equipped with the Skorohod topology.

Additional regularity conditions needed for the result are stated in Section 2. Although these conclusions could possibly be derived within a different framework considered by Bickel and Ren [4], the conditions given here are adapted to the current context and easier to verify than theirs. The present proofs are also different and yield interesting insights. The two-level parametric bootstrap introduced in Section 3 also appears to be novel; it is required in many applications where $A_{\theta_n}$ cannot be computed easily but can be approximated through a parametric bootstrap of its own.

The goodness-of-fit tests for copula models introduced above are revisited in Section 4. Also given there is a multivariate extension of a procedure designed by Durbin [9] for checking the fit of a univariate distribution. As a practical illustration, testing for a Gaussian copula structure is considered in Section 5 on the basis of the empirical copula (5). An explicit algorithm is also provided which can be adapted easily to test for other copula families via one- or two-level parametric bootstrapping. For a more extensive comparison of this procedure with alternative tests for copula models, see Genest et al. [16].

To avoid interrupting the flow of the presentation, most technical arguments are relegated to a series of appendices.

## 2. Validity of the one-level parametric bootstrap

Let $U_1,\dots,U_n$ be a random sample from some distribution $P$, and assume that it is desired to test the hypothesis

$$H_0\colon\ P\in\mathcal{P}=\{P_\theta\colon\ \theta\in\mathcal{O}\},$$

where $\mathcal{P}$ is a family of probability measures on $\mathbb{R}^d$ indexed by a parameter $\theta$ living in an open set $\mathcal{O}\subset\mathbb{R}^p$. The family is assumed to be identifiable, i.e., $\theta\neq\theta'\Rightarrow P_\theta\neq P_{\theta'}$.

As discussed in the Introduction, let $\mathcal{T}\subset[-\infty,\infty]^r$ be a closed rectangle and suppose that the test of $H_0$ is to be based on an abstract mapping $A:\mathcal{T}\to\mathbb{R}^s$ that depends on the true distribution $P$ of $U_1,\dots,U_n$. In particular, suppose that $A=A_\theta$ when $P=P_\theta$, and write $\mathcal{A}=\{A_\theta:\theta\in\mathcal{O}\}$. In this general context, identifiability is ensured if for every $\varepsilon>0$,

$$\inf\Big\{\sup_{t\in\mathcal{T}}\big\|A_\theta(t)-A_{\theta_0}(t)\big\|\colon\ \theta\in\mathcal{O}\text{ and }|\theta-\theta_0|>\varepsilon\Big\}>0.$$

This condition is assumed throughout, as one might otherwise have $A_\theta=A_{\theta'}$ for some $\theta\neq\theta'$ and problems could arise; see, e.g., [24]. Furthermore, the mapping $\theta\mapsto A_\theta$ is assumed to be Fréchet differentiable with derivative $\dot{A}$, i.e., for all $\theta_0\in\mathcal{O}$,

$$\lim_{h\to 0}\sup_{t\in\mathcal{T}}\frac{\|A_{\theta_0+h}(t)-A_{\theta_0}(t)-\dot{A}(t)h\|}{\|h\|}=0. \tag{10}$$

Finally, let $\theta_n=T_n(U_1,\dots,U_n)$ be a consistent estimate of $\theta$ and assume that the $\mathcal{D}(\mathcal{T};\mathbb{R}^s)$-valued process $A_n=\Upsilon_n(U_1,\dots,U_n)$ estimates $A$ consistently. Suppose specifically that the processes $\Theta_n=n^{1/2}(\theta_n-\theta)$ and $\mathbb{A}_n=n^{1/2}(A_n-A)$ have centered Gaussian limits when $n\to\infty$, as per (8) and (9).

The purpose of this section is to state additional regularity conditions on the families $\mathcal{P}$, $\mathcal{A}$ and on the sequences $A_n$ and $\theta_n$ of estimators. These requirements will ensure that a parametric bootstrap algorithm approximates correctly the limiting behavior of the empirical process

$$\mathbb{G}_n^A=n^{1/2}(A_n-A_{\theta_n}).$$

Consequently, the parametric bootstrap will also provide a suitable approximation of the asymptotic distribution of goodness-of-fit test statistics expressed as continuous functionals $S_n=\phi(\mathbb{G}_n^A)$.

The validity of the parametric bootstrap first depends on smoothness and integrability conditions on the parametric family of distributions.

***Definition 1.*** *A family $\mathcal{P}=\{P_\theta\colon\ \theta\in\mathcal{O}\}$ is said to belong to the class $\mathcal{S}(\lambda)$ for a given reference measure $\lambda$ (independent of $\theta$) if:*

1.1. *The measure* $P_\theta$ *is absolutely continuous with respect to* $\lambda$ *for all* $\theta \in \mathcal{O}$.
1.2. *The density* $p_\theta = \mathrm{d}P_\theta/\mathrm{d}\lambda$ *admits first and second order derivatives with respect to all components of* $\theta \in \mathcal{O}$. *The gradient (row) vector with respect to* $\theta$ *is denoted* $\dot{p}_\theta$, *and the Hessian matrix is represented by* $\ddot{p}_\theta$.
1.3. *For arbitrary* $u \in \mathbb{R}^d$ *and every* $\theta_0 \in \mathcal{O}$, *the mappings* $\theta \mapsto \dot{p}_\theta(u)/p_\theta(u)$ *and* $\theta \mapsto \ddot{p}_\theta(u)/p_\theta(u)$ *are continuous at* $\theta_0$, $P_{\theta_0}$ *almost surely.*
1.4. *For every* $\theta_0 \in \mathcal{O}$, *there exist a neighborhood* $\mathcal{N}$ *of* $\theta_0$ *and a* $\lambda$-*integrable function* $h : \mathbb{R}^d \to \mathbb{R}$ *such that for all* $u \in \mathbb{R}^d$, $\sup_{\theta \in \mathcal{N}} \|\dot{p}_\theta(u)\| \le h(u)$.
1.5. *For every* $\theta_0 \in \mathcal{O}$, *there exist a neighborhood* $\mathcal{N}$ *of* $\theta_0$ *and* $P_{\theta_0}$-*integrable functions* $h_1, h_2 : \mathbb{R}^d \to \mathbb{R}$ *such that for every* $u \in \mathbb{R}^d$,

$$\sup_{\theta \in \mathcal{N}} \left\| \frac{\dot{p}_\theta(u)}{p_\theta(u)} \right\|^2 \le h_1(u) \quad \textit{and} \quad \sup_{\theta \in \mathcal{N}} \left\| \frac{\ddot{p}_\theta(u)}{p_\theta(u)} \right\| \le h_2(u).$$

In the sequel, $\theta_0$ represents the true (unknown) value of $\theta$ and $P = P_{\theta_0}$. Furthermore,

$$p = p_{\theta_0}, \qquad \dot{p} = \dot{p}_{\theta_0}, \qquad \ddot{p} = \ddot{p}_{\theta_0}.$$

***Remark 1.*** *Using Condition* 1.4 *with the continuity of* $\dot{p}_\theta$ *as a function of* $\theta$ *and Lebesgue's dominated convergence theorem, one may conclude that*

$$\frac{\partial}{\partial \theta} \int p_\theta(u) g(u) \lambda(\mathrm{d}u) = \int \dot{p}_\theta(u) g(u) \lambda(\mathrm{d}u) \tag{11}$$

*for any bounded measurable function* $g : \mathbb{R}^d \to \mathbb{R}$, *not depending on* $\theta$. *In particular,* $\int \dot{p}(u)\lambda(\mathrm{d}u) = 0$. *Furthermore, if* $F_\theta$ *denotes the distribution function associated with* $P_\theta$, *the mapping* $\theta \mapsto F_\theta$ *is then Fréchet differentiable and its derivative* $\dot{F}_\theta$ *satisfies the following identity for all* $x \in \mathbb{R}^d$:

$$\dot{F}_\theta(x) = \int \dot{p}_\theta(u) \mathbf{1}(u \le x) \lambda(\mathrm{d}u). \tag{12}$$

***Remark 2.*** *When* $\mathcal{P} \in \mathcal{S}(\lambda)$, *the multivariate central limit theorem implies that if* $U_1, \ldots, U_n$ *form a random sample from* $P = P_{\theta_0}$, *then as* $n \to \infty$,

$$\mathbb{W}_{P,n} = n^{-1/2} \sum_{i=1}^{n} \frac{\dot{p}^\top(U_i)}{p(U_i)} \rightsquigarrow \mathbb{W}_P \sim \mathcal{N}(0, I_P), \tag{13}$$

*where* $\mathrm{E}(\mathbb{W}_P) = 0$ *by Remark* 1 *and* $I_P$ *is the Fisher information matrix, viz.*

$$I_P = \int \frac{\dot{p}^\top(u)\dot{p}(u)}{p(u)} \lambda(\mathrm{d}u). \tag{14}$$

The validity of the parametric bootstrap also relies on the following general notion of $\mathcal{P}$-regularity of estimators. It is cast below in terms of $A_n$ but it applies also to many other sequences in the sequel, e.g., in the case $A_n = \theta_n$.

***Definition 2.*** *Let* $U_1, \ldots, U_n$ *be a random sample from* $P = P_{\theta_0}$ *and let* $\mathbb{W}_{P,n}$ *be defined as in* (13). *A sequence* $A_n$ *is said to be* $P_{\theta_0}$-*regular for* $A = A_{\theta_0}$ *if, as* $n \to \infty$, *the process* $(\mathbb{A}_n, \mathbb{W}_{P,n})$ *with* $\mathbb{A}_n = n^{1/2}(A_n - A)$ *converges weakly in* $\mathcal{D}(\mathcal{T}; \mathbb{R}^s) \times \mathbb{R}^p$ *to a centered Gaussian pair* $(\mathbb{A}, \mathbb{W}_P)$ *and the Fréchet derivative* $\dot{A}$ *of* $A$ *defined in* (10) *satisfies* $\dot{A}(t) = \mathrm{E}\{\mathbb{A}(t)\mathbb{W}_P^\top\}$ *for every* $t \in \mathcal{T}$. *The sequence is said to be* $\mathcal{P}$-*regular for* $\mathcal{A}$ *if it is* $P_{\theta_0}$-*regular for* $A_{\theta_0}$ *at all* $\theta_0 \in \mathcal{O}$.

***Remark 3.*** *The* $\mathcal{P}$-*regularity of a sequence of estimators* $\theta_n = T_n(U_1, \ldots, U_n)$ *for* $\theta \in \mathcal{O}$ *implies that* $\Theta_n = n^{1/2}(\theta_n - \theta) \rightsquigarrow \Theta$ *as* $n \to \infty$, *where* $\Theta$ *is a centered Gaussian random vector and* $\mathrm{E}(\Theta \mathbb{W}_P^\top) = I$ *is the identity matrix.*

Now let $U_1^*, \dots, U_n^*$ be a bootstrap sample from $P_{\theta_n}$, and set

$$\theta_n^* = T_n(U_1^*, \dots, U_n^*), \qquad \Theta_n^* = n^{1/2}(\theta_n^* - \theta),$$
$$A_n^* = \Upsilon_n(U_1^*, \dots, U_n^*), \qquad \mathbb{A}_n^* = n^{1/2}(A_n^* - A).$$

The following result, whose proof is given in Appendix B, gives conditions under which the weak limits of the processes

$$\mathbb{G}_n^A = n^{1/2}(A_n - A_{\theta_n}) \quad \text{and} \quad \mathbb{G}_n^{A^*} = n^{1/2}(A_n^* - A_{\theta_n^*}).$$

are independent and identically distributed. This guarantees that a parametric bootstrap based on the process $\mathbb{A}_n$ is valid.

**Theorem 1.** *Assume that $\mathcal{P} \in \mathcal{S}(\lambda)$ and that as $n \to \infty$,*

$$(\mathbb{A}_n, \Theta_n, \mathbb{W}_{P,n}) \rightsquigarrow (\mathbb{A}, \Theta, \mathbb{W}_P) \tag{15}$$

*in $\mathcal{D}(\mathcal{T}; \mathbb{R}^s) \times \mathbb{R}^{p\otimes 2}$, where the limit is a centered Gaussian process. Let $\Gamma = \mathrm{E}(\Theta \mathbb{W}_P^\top)$ and set $a(t) = \mathrm{E}\{\mathbb{A}(t)\mathbb{W}_P^\top\}$ for every $t \in \mathcal{T}$. Then, as $n \to \infty$,*

$$(\mathbb{A}_n, \mathbb{A}_n^*, \Theta_n, \Theta_n^*) \rightsquigarrow (\mathbb{A}, \mathbb{A}^\star, \Theta, \Theta^\star)$$

*in $\mathcal{D}(\mathcal{T}; \mathbb{R}^s)^{\otimes 2} \times \mathbb{R}^{p\otimes 2}$. In the limit, $\mathbb{A}^\star = \mathbb{A}^\perp + a\Theta$ and $\Theta^\star = \Theta^\perp + \Gamma\Theta$ are defined in terms of an independent copy $(\mathbb{A}^\perp, \Theta^\perp)$ of $(\mathbb{A}, \Theta)$. If in addition $(A_n, \theta_n)$ is $\mathcal{P}$-regular for $\mathcal{A} \times \mathcal{O}$, then*

$$(\mathbb{G}_n^A, \mathbb{G}_n^{A^*}) \rightsquigarrow (\mathbb{G}^A, \mathbb{G}^{A^*}) = (\mathbb{A} - \dot{A}\Theta, \mathbb{A}^\perp - \dot{A}\Theta^\perp)$$

*in $\mathcal{D}(\mathcal{T}; \mathbb{R}^s)^{\otimes 2}$, as $n \to \infty$, and $\mathbb{G}^{A^*}$ is an independent copy of $\mathbb{G}^A$.*

## 3. A two-level parametric bootstrap

To perform a goodness-of-fit test based on a continuous functional $S_n = \phi(\mathbb{G}_n^A)$ of the process

$$\mathbb{G}_n^A = n^{1/2}(A_n - A_{\theta_n}),$$

one must compute $A_{\theta_n}$ at various points, but this is not always easily done.

For tests based on the empirical copula, for instance, one has $A_{\theta_n} = C_{\theta_n}$ and many copula families are not algebraically closed. In this case, a simple way to circumvent the problem is to generate a random sample $V_1^*, \dots, V_m^*$ from probability measure $Q_{\theta_n}$ with distribution function $C_{\theta_n}$ and for $u \in [0, 1]^d$, to approximate $C_{\theta_n}(u)$ by

$$\check{C}_n^*(u) = \frac{1}{m} \sum_{j=1}^{m} \mathbf{1}(V_j^* \le u).$$

It is typical to take $m = \lfloor \gamma n \rfloor$ for some $\gamma \in (0, \infty)$, but it will only be assumed here that $m$ is a function of $n$ such that $m/n \to \gamma \in (0, \infty)$ as $n \to \infty$.

More generally, the strategy proposed here consists of replacing $A_{\theta_n}$ by an approximation $\check{A}_n^* = \Psi_m(V_1^*, \dots, V_m^*)$ built from a random sample $V_1^*, \dots, V_m^*$ from $Q_{\theta_n} \in \mathcal{Q} = \{Q_\theta\colon \theta \in \mathcal{O}\}$. In order for this approach to make sense, it must be assumed that if $A = A_{\theta_0}$ and $\check{A}_n = \Psi_m(V_1, \dots, V_m)$ for a random sample $V_1, \dots, V_m$ from $Q = Q_{\theta_0}$, then

$$\check{\mathbb{A}}_n = n^{1/2}(\check{A}_n - A) \rightsquigarrow \check{\mathbb{A}}$$

in $\mathcal{D}(\mathcal{T}; \mathbb{R}^s)$, as $n \to \infty$ (and hence $m \to \infty$).

Given that such a process exists, here is a natural way to circumvent the lack of a closed form for $A_{\theta_n}$ in the computation of the test statistic $S_n$:

(a) Compute $\theta_n = T_n(U_1, \dots, U_n)$ and let $A_n = \Upsilon_n(U_1, \dots, U_n)$.
(b) Given $U_1, \dots, U_n$, generate a random sample $V_1^*, \dots, V_m^*$ from $Q_{\theta_n}$.
(c) Let $\breve{A}_n^* = \Psi_m(V_1^*, \dots, V_m^*)$ and compute $S_n = \phi(\mathbb{G}_n^{\breve{A}^*})$, in which $\mathbb{G}_n^{\breve{A}^*} = n^{1/2}(A_n - \breve{A}_n^*)$.

Now in order to approximate the distribution of $S_n$, a second parametric bootstrap procedure is necessary. To this end, pick $N$ large and repeat the following steps for every $k \in \{1, \dots, N\}$:

(a) Given $U_1, \dots, U_n, V_1^*, \dots, V_m^*$, generate a random sample $U_{1,k}^*, \dots, U_{n,k}^*$ from $P_{\theta_n}$.
(b) Compute $\theta_{n,k}^* = T_n(U_{1,k}^*, \dots, U_{n,k}^*)$ and let $A_{n,k}^* = \Upsilon_n(U_{1,k}^*, \dots, U_{n,k}^*)$.
(c) Given $U_1, \dots, U_n, V_1^*, \dots, V_m^*$ and $U_{1,k}^*, \dots, U_{n,k}^*$, generate a random sample $V_{1,k}^{**}, \dots, V_{m,k}^{**}$ from $Q_{\theta_{n,k}^*}$.
(d) Let $\breve{A}_{n,k}^{**} = \Psi_m(V_{1,k}^{**}, \dots, V_{m,k}^{**})$ and compute $S_{n,k}^* = \phi(\mathbb{G}_{n,k}^{\breve{A}^{**}})$, in which $\mathbb{G}_{n,k}^{\breve{A}^{**}} = n^{1/2}(A_{n,k}^* - \breve{A}_{n,k}^{**})$.

With the convention that large values of $S_n$ lead to the rejection of $H_0$, and under regularity conditions stated below, a valid approximation to the $P$-value for the test based on $S_n = \phi(\mathbb{G}_n^{\breve{A}^*})$ is given by

$$\frac{1}{N} \sum_{k=1}^{N} \mathbf{1}\left(S_{n,k}^* > S_n\right).$$

As for the standard parametric bootstrap, the validity of the above two-level extension is ensured, provided that one can show that, as $n \to \infty$, $(\mathbb{G}_n^{\breve{A}^*}, \mathbb{G}_{n,1}^{\breve{A}^{**}})$ converges weakly in $\mathcal{D}(\mathcal{T}; \mathbb{R}^s)^{\otimes 2}$ to a pair of independent and identically distributed limiting processes.

Assume that $\mathcal{Q} \in \mathcal{S}(\nu)$ for some reference measure $\nu$ (independent of $\theta$). Write $q_\theta$ for the density of $Q_\theta$, let $\dot{q}_\theta$ be the gradient (row) vector with respect to $\theta$, and denote the Hessian matrix by $\ddot{q}_\theta$. When $Q = Q_{\theta_0}$, write by extension

$$q = q_{\theta_0}, \qquad \dot{q} = \dot{q}_{\theta_0}, \qquad \ddot{q} = \ddot{q}_{\theta_0}.$$

Note that when $\mathcal{Q} \in \mathcal{S}(\nu)$, the multivariate central limit theorem implies that if $V_1, \dots, V_m$ form a random sample from $Q = Q_\theta$, then, as $n \to \infty$,

$$\mathbb{W}_{Q,n} = n^{-1/2} \sum_{i=1}^{m} \frac{\dot{q}^\top(V_i)}{q(V_i)} \rightsquigarrow \mathbb{W}_Q \sim \mathcal{N}(0, I_Q), \tag{16}$$

where in view of the fact that $m/n \to \gamma \in (0, \infty)$ as $n \to \infty$,

$$I_Q = \gamma \int \frac{\dot{q}^\top(u)\dot{q}(u)}{q(u)} \nu(\mathrm{d}u). \tag{17}$$

Now let $U_1, \dots, U_n$ and $V_1, \dots, V_m$ be two mutually independent random samples from $P = P_{\theta_0} \in \mathcal{P}$ and $Q = Q_{\theta_0} \in \mathcal{Q}$, respectively. Let $\mathbb{W}_{P,n}$ and $\mathbb{W}_{Q,n}$ be defined as in (13) and (16), respectively. Conditionally on $U_1, \dots, U_n$ and $V_1, \dots, V_m$, make the following additional assumptions:

(a) Given $\theta_n = T_n(U_1, \dots, U_n)$, the random vectors $U_1^*, \dots, U_n^*$ and $V_1^*, \dots, V_m^*$ are mutually independent random samples from $P_{\theta_n}$ and $Q_{\theta_n}$, respectively.
(b) Given $U_1^*, \dots, U_n^*$ and $V_1^*, \dots, V_m^*$ and $\theta_n^* = T_n(U_1^*, \dots, U_n^*)$, the random vectors $V_1^{**}, \dots, V_m^{**}$ are a random sample from $Q_{\theta_n^*}$.

Finally, introduce the additional notations

$$\breve{A}_n = \Psi_m(V_1, \dots, V_m), \qquad \breve{A}_n^* = \Psi_m\left(V_1^*, \dots, V_m^*\right), \qquad \breve{A}_n^{**} = \Psi_m\left(V_1^{**}, \dots, V_m^{**}\right)$$

and

$$\breve{\mathbb{A}}_n = n^{1/2}(\breve{A}_n - A), \qquad \breve{\mathbb{A}}_n^* = n^{1/2}\left(\breve{A}_n^* - A\right), \qquad \breve{\mathbb{A}}_n^{**} = n^{1/2}\left(\breve{A}_n^{**} - A\right).$$

The following result, whose proof is given in Appendix C, gives conditions under which the weak limits of the processes

$$\mathbb{G}_n^{\check{A}^*} = n^{1/2}\big(A_n - \check{A}_n^*\big) \quad \text{and} \quad \mathbb{G}_n^{\check{A}^{**}} = n^{1/2}\big(A_n^* - \check{A}_n^{**}\big)$$

are independent and identically distributed. This proves the validity of a two-level parametric bootstrap based on the process $\mathbb{A}_n$.

**Theorem 2.** *Assume that $\mathcal{P} \in \mathcal{S}(\lambda)$, $\mathcal{Q} \in \mathcal{S}(\nu)$ and that as $n \to \infty$,*

$$(\mathbb{A}_n, \check{\mathbb{A}}_n, \Theta_n, \mathbb{W}_{P,n}, \mathbb{W}_{Q,n}) \rightsquigarrow (\mathbb{A}, \check{\mathbb{A}}, \Theta, \mathbb{W}_P, \mathbb{W}_Q)$$

*and that the limit is a centered Gaussian process in $\mathcal{D}(\mathcal{T}; \mathbb{R}^s)^{m\otimes 2} \times \mathbb{R}^{p\otimes 3}$. Let $\Gamma = \mathrm{E}(\Theta \mathbb{W}_P^\top)$ and set $a(t) = \mathrm{E}\{\mathbb{A}(t)\mathbb{W}_P^\top\}$ and $\check{a}(t) = \mathrm{E}\{\check{\mathbb{A}}(t)\mathbb{W}_Q^\top\}$ for every $t \in \mathcal{T}$. Then, as $n \to \infty$,*

$$\big(\mathbb{A}_n, \mathbb{A}_n^*, \check{\mathbb{A}}_n, \check{\mathbb{A}}_n^*, \check{\mathbb{A}}_n^{**}, \Theta_n, \Theta_n^*\big) \rightsquigarrow \big(\mathbb{A}, \mathbb{A}^\star, \check{\mathbb{A}}, \check{\mathbb{A}}^\star, \check{\mathbb{A}}^{\star\star}, \Theta, \Theta^\star\big)$$

*in $\mathcal{D}(\mathcal{T}; \mathbb{R}^s)^{\otimes 5} \times \mathbb{R}^{p\otimes 2}$. In the limit,*

$$\mathbb{A}^\star = \mathbb{A}^\perp + a\Theta, \qquad \Theta^\star = \Theta^\perp + \Gamma\Theta, \qquad \check{\mathbb{A}}^\star = \check{\mathbb{A}}^\perp + \check{a}\Theta, \qquad \check{\mathbb{A}}^{\star\star} = \check{\mathbb{A}}^{\perp\perp} + \check{a}\Theta^\star,$$

*where $(\mathbb{A}^\perp, \Theta^\perp)$ is an independent copy of $(\mathbb{A}, \Theta)$. In addition, the processes $\check{\mathbb{A}}$, $\check{\mathbb{A}}^\perp$ and $\check{\mathbb{A}}^{\perp\perp}$ are mutually independent and identically distributed, as well as independent of $\mathbb{A}$, $\mathbb{A}^\perp$, $\Theta$ and $\Theta^\perp$. Moreover if $(A_n, \theta_n)$ is $\mathcal{P}$-regular for $\mathcal{A} \times \mathcal{O}$ and $\check{A}_n$ is $\mathcal{Q}$-regular for $\mathcal{A}$, then*

$$\big(\mathbb{G}_n^{\check{A}^*}, \mathbb{G}_n^{\check{A}^{**}}\big) \rightsquigarrow \big(\mathbb{G}^{\check{A}^*}, \mathbb{G}^{\check{A}^{**}}\big) = \big(\mathbb{A} - \check{\mathbb{A}}^\perp - \dot{A}\Theta, \mathbb{A}^\perp - \check{\mathbb{A}}^{\perp\perp} - \dot{A}\Theta^\perp\big)$$

*in $\mathcal{D}(\mathcal{T}; \mathbb{R}^s)^{\otimes 2}$, as $n \to \infty$, and $\mathbb{G}^{\check{A}^*}$ is an independent copy of $\mathbb{G}^{\check{A}^{**}}$.*

## 4. Examples of application

In this section, the validity of the one- and two-level parametric bootstrap is established in four common goodness-of-fit testing contexts. The first example considers classical tests for parametric families of random vectors; it is discussed here because the conditions under which Theorems 1 and 2 are established seem easier to verify than the requirements imposed by Stute et al. [26]. The second and the third examples are about goodness-of-fit for copula models, while the last application revisits the approach of Durbin [9] for goodness-of-fit testing of parametric families of random vectors using the probability integral transformation.

### 4.1. *Goodness-of-fit tests for parametric families*

Let $X$ be a $d$-variate random vector with continuous distribution function $F$. Suppose that it is desired to test the null hypothesis

$$H_0\colon\ F \in \mathcal{F} = \{F_\theta\colon\ \theta \in \mathcal{O}\},$$

i.e., $F = F_{\theta_0}$ for some $\theta_0 \in \mathcal{O}$. Given a random sample $X_1, \ldots, X_n$ from $F$, a natural procedure is to compare the empirical distribution function (1) to $F_{\theta_n}$, where $\theta_n = T_n(X_1, \ldots, X_n)$ is an estimation of the unknown parameter $\theta \in \mathbb{R}^p$. The test could be based, e.g., on a Cramér–von Mises or on a Kolmogorov–Smirnov functional $S_n = \phi(\mathbb{G}_n^F)$ of the empirical process

$$\mathbb{G}_n^F = n^{1/2}(F_n - F_{\theta_n}).$$

To establish the validity of the parametric bootstrap for such statistics, one can use Theorems 1 and 2 with $A_\theta = F_\theta$ and $P_\theta$ standing for the unique probability measure associated with $F_\theta$ and density $f_\theta$. Assume that $\mathcal{P} = \{P_\theta\colon \theta \in \mathcal{O}\} \in S(\lambda)$, where $\lambda$ is Lebesgue's measure. Introduce the following notation:

$$f = f_{\theta_0}, \qquad \dot{f} = \dot{f}_{\theta_0}, \qquad \ddot{f} = \ddot{f}_{\theta_0}.$$

To check the $\mathcal{P}$-regularity of $\mathcal{F}$, let $\mathbb{F}_n = n^{1/2}(F_n - F)$ and

$$\mathbb{W}_{F,n} = n^{-1/2} \sum_{i=1}^{n} \frac{\dot{f}^\top(X_i)}{f(X_i)}. \tag{18}$$

Results from [5] imply that as $n \to \infty$, $(\mathbb{F}_n, \mathbb{W}_{F,n}) \rightsquigarrow (\mathbb{F}, \mathbb{W}_F)$ in $\mathcal{D}([-\infty, \infty]^d; \mathbb{R}) \times \mathbb{R}^p$, where $\mathbb{W}_F$ is a centered Gaussian variable with variance

$$I_F = \int \frac{\dot{f}^\top(x)\dot{f}(x)}{f(x)} \lambda(\mathrm{d}x)$$

and $\mathbb{F}$ is an $F$-Brownian bridge, i.e., $\mathbb{F}$ is a continuous centered Gaussian process with covariance function

$$\operatorname{cov}\big\{\mathbb{F}(x), \mathbb{F}(y)\big\} = F(x \wedge y) - F(x)F(y),$$

where $x \wedge y = \min(x, y)$ for all $x, y \in \mathbb{R}^d$. The following result is a consequence of these observations and the fact that for all $x \in \mathbb{R}^d$,

$$E\big\{\mathbb{F}(x)\mathbb{W}_F^\top\big\} = \int \dot{f}(y)\mathbf{1}(y \le x)\lambda(\mathrm{d}y) = \dot{F}(x)$$

in view of Eq. (12).

**Proposition 1.** *Let $X_1, \ldots, X_n$ be a random sample from distribution $F = F_{\theta_0}$ for some $\theta_0 \in \mathcal{O}$. If $\mathcal{P} \in \mathcal{S}(\lambda)$, then the canonical empirical distribution function $F_n$ defined in* (1) *is $\mathcal{P}$-regular for $\mathcal{F}$.*

Next, assume that $\theta_n$ is a $\mathcal{P}$-regular sequence for $\mathcal{O}$ such that, as $n \to \infty$,

$$(\mathbb{F}_n, \Theta_n, \mathbb{W}_{F,n}) \rightsquigarrow (\mathbb{F}, \Theta, \mathbb{W}_F)$$

in $\mathcal{D}([-\infty, \infty]^d; \mathbb{R}) \times \mathbb{R}^{p\otimes 2}$. Suppose further that the limit is Gaussian, so that condition (15) is satisfied with $\mathbb{A}_n = \mathbb{F}_n$. It then follows that $(F_n, \theta_n)$ is $\mathcal{P}$-regular for $\mathcal{F} \times \mathcal{O}$ because $\mathrm{E}(\mathbb{F}\mathbb{W}_F^\top) = \dot{F} = \dot{F}_{\theta_0}$ by Proposition 1 and $\mathrm{E}(\Theta\mathbb{W}_F^\top) = I$ by the regularity hypothesis on $\theta_n$.

Finally, all the conditions of Theorems 1 and 2 are met with $\mathcal{A} = \mathcal{F}$, $A_n = F_n$ and $\check{A}_n = \check{F}_n$, where the latter is defined for all $x \in \mathbb{R}^d$ by

$$\check{F}_n(x) = \frac{1}{m} \sum_{i=1}^{m} \mathbf{1}(Y_i \le x)$$

in terms of a random sample $Y_1, \ldots, Y_m$ from $P_\theta$ that is independent of $X_1, \ldots, X_n$. Therefore, the one- and two-level parametric bootstraps yield valid approximations of the distribution of any continuous functional $S_n = \phi(\mathbb{G}_n^F)$.

In this context, the class of estimators that are $\mathcal{P}$-regular for $\mathcal{O}$ is broad, as shown below.

***Definition 3.*** *An estimator $\theta_n = T_n(X_1, \ldots, X_n)$ for $\theta \in \mathcal{O}$ is said to belong to class $\mathcal{R}$ if $n^{1/2}(\theta_n - \theta) = \Theta_n' + \mathrm{o}_P(1)$, where*

$$\Theta_n' = n^{-1/2} \sum_{i=1}^{n} J_\theta(X_i) \tag{19}$$

*is expressed in terms of a score function $J_\theta : \mathbb{R}^d \to \mathbb{R}^p$ that is square integrable with respect to $P_\theta$ and such that for all $\theta \in \mathcal{O}$, one has both*

$$\mathrm{E}_\theta\{J_\theta(X)\} = \int J_\theta(x) f_\theta(x)\lambda(\mathrm{d}x) = 0 \quad \text{and} \quad \int J_\theta(x)\dot{f}_\theta(x)\lambda(\mathrm{d}x) = I. \tag{20}$$

**Proposition 2.** *Let $\theta_n = T_n(X_1, \ldots, X_n)$ be an estimator of $\theta \in \mathcal{O}$ from the class $\mathcal{R}$. If $\mathcal{P} \in S(\lambda)$, then $(F_n, \theta_n)$ is $\mathcal{P}$-regular for $\mathcal{F} \times \mathcal{O}$.*

To establish this result, first note that each component of the vector $(\mathbb{F}_n, \Theta_n', \mathbb{W}_{F,n})$ is tight and that the finite-dimensional distributions converge by the classical multivariate central limit theorem, because each term is a sum of independent and identically distributed centered random variables. In addition, observe that $E(\Theta\mathbb{W}_F^\top) = I$ by Eq. (20).

***Example 1.*** *When it is uniquely defined and $I_F$ is non-singular, the maximum likelihood estimator belongs to $\mathcal{R}$. For, in that case, relation* (19) *holds with $J_\theta = I_F^{-1}\dot{f}_\theta^\top / f_\theta$. Furthermore, this function satisfies conditions* (20) *because of identity* (11) *and from the fact that under $P = P_{\theta_0}$,*

$$\mathrm{E}(\Theta\mathbb{W}_F^\top) = I_F^{-1} \int \frac{\dot{f}^\top(x)\dot{f}(x)}{f(x)}\lambda(\mathrm{d}x) = I_F^{-1} I_F = I.$$

***Example 2.*** *Moments estimators also belong to $\mathcal{R}$. Assume that*

$$\theta = g(\mu) \quad \text{and} \quad \mu = \int M(x) f_\theta(x)\lambda(\mathrm{d}x)$$

*for some integrable function $M : \mathbb{R}^d \to \mathbb{R}^d$ that does not depend on $\theta$. Suppose also that $g$ is continuously differentiable and that the matrix $\dot{g}$ of derivatives is non-singular. Then $g^{-1}$ exists and is continuously differentiable by the inverse function theorem. Furthermore, Slutsky's theorem implies that for all $x \in \mathbb{R}^d$,*

$$J_\theta(x) = \dot{g}\{g^{-1}(\theta)\}\{M(x) - g^{-1}(\theta)\}.$$

*This score function meets the appropriate requirements because of* (11) *and the fact that under $P$,*

$$\begin{aligned}\mathrm{E}(\Theta\mathbb{W}_F^\top) &= \dot{g}\{g^{-1}(\theta_0)\} \int h(x)\frac{\dot{f}(x)}{f(x)}\lambda(\mathrm{d}x) \\ &= \dot{g}\{g^{-1}(\theta_0)\}\left[\frac{\partial}{\partial\theta}\int M(x) f_\theta(x)\lambda(\mathrm{d}x)\right]_{\theta=\theta_0} \\ &= \dot{g}\{g^{-1}(\theta_0)\}\left[\frac{\partial}{\partial\theta}\, g^{-1}(\theta)\right]_{\theta=\theta_0} = I.\end{aligned}$$

***Example 3.*** *When it is uniquely defined, the estimator $\theta_n$ minimizing*

$$\varrho_n(\theta) = \int \{F_n(x) - F_\theta(x)\}^2 \,\mathrm{d}F_n(x)$$

*between $F_n$ and $F_\theta$ also belongs to $\mathcal{R}$, provided that*

$$\Sigma_\theta = \int \dot{F}_\theta^\top(x)\dot{F}_\theta(x) f_\theta(x)\lambda(\mathrm{d}x)$$

*is non-singular for every $\theta \in \mathcal{O}$. In this case, representation* (19) *holds with*

$$J_\theta(x) = \Sigma_\theta^{-1} \int \{\mathbf{1}(x \le y) - F_\theta(y)\}\dot{F}_\theta^\top(y) f_\theta(y)\lambda(\mathrm{d}y)$$

*for all* $x \in \mathbb{R}^d$ *and*

$$\Theta_n = \Sigma^{-1} \int \mathbb{F}_n(y) \dot{F}_\theta^\top(y) f_\theta(y) \lambda(\mathrm{d}y) + \mathrm{o}_P(1)$$

*with* $\Sigma = \Sigma_{\theta_0}$. *Thus, as* $n \to \infty$, *one has* $(\mathbb{F}_n, \Theta_n) \rightsquigarrow (\mathbb{F}, \Theta)$ *under* $P$, *and*

$$\Theta = \Sigma^{-1} \int \mathbb{F}(y) \dot{F}^\top(y) f(y) \lambda(\mathrm{d}y).$$

*Direct calculations show that* $\theta_n$ *is* $\mathcal{P}$*-regular for* $\mathcal{O}$. *For, under* $P$,

$$\mathrm{E}\big(\Theta \mathbb{W}_F^\top\big) = \Sigma^{-1} \mathrm{E}\bigg\{ \int \mathbb{F}(y) \dot{F}^\top(y) f(y) \mathbb{W}_F^\top \lambda(\mathrm{d}y) \bigg\} = \Sigma^{-1} \int \dot{F}^\top(y) \dot{F}(y) f(y) \lambda(\mathrm{d}y) = I.$$

*For conditions under which* $\theta_n$ *exists and is unique, see* [2] *or* [3].

### 4.2. *Goodness-of-fit tests for copulas*

Let $X$ be a continuous $d$-variate random vector with distribution function $F$, margins $F_1, \dots, F_d$, and unique underlying copula $C$. Suppose it is desired to test the null hypothesis

$$H_0\colon\ C \in \mathcal{C} = \{C_\theta\colon\ \theta \in \mathcal{O}\},$$

i.e., $C = C_{\theta_0}$ for some $\theta_0 \in \mathcal{O}$. Given a random sample $X_1, \dots, X_n$ from $F$, a natural way to proceed is to compare the empirical copula $C_n$ defined in (5) to a parametric estimate $C_{\theta_n}$, where $\theta_n$ is an estimation of the unknown parameter $\theta \in \mathbb{R}^p$.

In view of the fact that the dependence structure represented by the copula $C$ is invariant by strictly increasing transformations of the margins of $X$, many authors have argued that estimators $\theta_n$ of $\theta$ should be margin-free, i.e., based on the ranks of the observations, which are maximally invariant under this class of transformations. This amounts to taking $\theta_n$ as a function of the pseudo-observations $\hat{U}_i = \xi_n(X_i)$ with $\xi_n$ defined in (3).

Note that under this condition, both $C_n$ and $\theta_n$ are measurable with respect to the sigma-algebra $\mathcal{U}_n$ generated by $U_1 = \xi(X_1), \dots, U_n = \xi(X_n)$, where $\xi$ is the mapping defined in (2). Although they are unobservable, the latter variables are mutually independent copies of $U = \xi(X)$ and distributed as $C$.

For arbitrary $u \in [0,1]^d$, let

$$B_n(u) = \frac{1}{n+1} \sum_{i=1}^n \mathbf{1}(U_i \le u) \tag{21}$$

and for every $j \in \{1, \dots, d\}$ and $t \in [0,1]$, define

$$B_{jn}(t) = \frac{1}{n+1} \sum_{i=1}^n \mathbf{1}(U_{ij} \le t). \tag{22}$$

The empirical copula is then asymptotically equivalent to

$$C_n(u) = B_n\big\{B_{1n}^{-1}(u_1), \dots, B_{dn}^{-1}(u_d)\big\}$$

at every $u = (u_1, \dots, u_d) \in [0,1]^d$.

Thus assume that $\theta_n = T_n(U_1, \dots, U_n)$ and suppose that $S_n = \phi(\mathbb{G}_n^C)$ is a continuous functional of the empirical process

$$\mathbb{G}_n^C = n^{1/2}(C_n - C_{\theta_n}).$$

To establish the validity of the parametric bootstrap for such goodness-of-fit statistics, one can use Theorems 1 and 2 with $A_\theta = C_\theta$ and $P_\theta$ standing for the unique probability measure associated with $C_\theta$ and density $c_\theta$. Assume that $\mathcal{P} = \{P_\theta\colon \theta \in \mathcal{O}\} \in S(\lambda)$, where $\lambda$ is Lebesgue's measure. Introduce the following notation:

$$c = c_{\theta_0}, \qquad \dot{c} = \dot{c}_{\theta_0}, \qquad \ddot{c} = \ddot{c}_{\theta_0}.$$

To check the $\mathcal{P}$-regularity of $\mathcal{C}$, let $\mathbb{C}_n = n^{1/2}(C_n - C)$ be the empirical copula process and write

$$\mathbb{W}_{C,n} = n^{-1/2}\sum_{i=1}^{n}\frac{\dot{c}^\top(U_i)}{c(U_i)}.$$

Using results from Chapter 5 of the book by Gänßler and Stute [12], one can then show that, as $n \to \infty$,

$$(\mathbb{B}_n, \mathbb{C}_n, \mathbb{W}_{C,n}) \rightsquigarrow (\mathbb{B}, \mathbb{C}, \mathbb{W}_C)$$

in $\mathcal{D}([0,1]^d; \mathbb{R})^{\otimes 2} \times \mathbb{R}^p$. Here, $\mathbb{W}_C$ is a centered Gaussian variable with variance

$$I_C = \int \frac{\dot{c}^\top(x)\dot{c}(x)}{c(x)}\lambda(\mathrm{d}x)$$

and $\mathbb{B}$ is a $C$-Brownian bridge. Furthermore, as shown by Gänßler and Stute [12] (but see also [10,19,27]), the limit $\mathbb{C}$ admits the representation

$$\mathbb{C}(u) = \mathbb{B}(u) - \sum_{j=1}^{d}\beta_j(u_j)\frac{\partial}{\partial u_j}C(u), \tag{23}$$

for all $u = (u_1, \dots, u_d) \in [0,1]^d$, where for each $j \in \{1, \dots, d\}$, $\beta_j$ is a classical Brownian bridge related to $\mathbb{B}$ via the equation $\beta_j(t) = \mathbb{B}(1_{t,j})$ in which $1_{t,j} = (e_1, \dots, e_d) \in \mathbb{R}^d$ with $e_i = t$ if $i = j$ and $e_i = 1$ otherwise.

Note that in view of Eq. (12),

$$E\{\mathbb{B}(u)\mathbb{W}_C^\top\} = \int \dot{c}(v)\mathbf{1}(v \le u)\lambda(\mathrm{d}v) = \dot{C}(u)$$

for all $u \in [0,1]^d$, and hence for all $t \in [0,1]$ and $j \in \{1, \dots, d\}$ one has

$$E\{\beta_j(t)\mathbb{W}_C^\top\} = \dot{C}(1_{t,j}) = 0.$$

Thus for all $u \in [0,1]^d$, representation (23) yields $E\{\mathbb{C}(u)\mathbb{W}_C^\top\} = \dot{C}(u)$.

The following result is a consequence of these observations.

**Proposition 3.** *Let $X_1, \dots, X_n$ be a random sample from distribution $F$ with unique underlying copula $C = C_{\theta_0}$ for some $\theta_0 \in \mathcal{O}$. If $\mathcal{P} \in \mathcal{S}(\lambda)$, then the empirical copula $C_n$ is $\mathcal{P}$-regular for $\mathcal{C}$.*

Next, assume that $\theta_n$ is a $\mathcal{P}$-regular sequence for $\mathcal{O}$ such that, as $n \to \infty$,

$$(\mathbb{B}_n, \Theta_n, \mathbb{W}_{C,n}) \rightsquigarrow (\mathbb{B}, \Theta, \mathbb{W}_C)$$

in $\mathcal{D}([0,1]^d; \mathbb{R}) \times \mathbb{R}^{p\otimes 2}$, where the weak limit is Gaussian. It then follows that $(C_n, \theta_n)$ is $\mathcal{P}$-regular for $\mathcal{C} \times \mathcal{O}$ because $\mathrm{E}(\mathbb{C}\mathbb{W}_C^\top) = \dot{C} = \dot{C}_{\theta_0}$ by Proposition 1 and $\mathrm{E}(\Theta\mathbb{W}_C^\top) = I$ by the regularity hypothesis on $\theta_n$.

Finally, all the conditions of Theorems 1 and 2 are met with $\mathcal{A} = \mathcal{C}$, $A_n = C_n$ and $\check{A}_n = \check{B}_n$ defined for all $u \in [0,1]^d$ by

$$\check{B}_n(u) = \frac{1}{m}\sum_{i=1}^{m}\mathbf{1}(Y_i \le u)$$

in terms of a random sample $Y_1, \dots, Y_m$ from $P_\theta$ that is independent of $X_1, \dots, X_n$. Therefore, the one- and two-level parametric bootstraps yield valid approximations of the distribution of any continuous functional $S_n = \phi(\mathbb{G}_n^C)$.

In the context of copula models, there are two main strategies for rank-based estimation of the dependence parameter $\theta$. These approaches lead to two distinct classes of estimators, which are considered separately. In the sequel,

$$H_n(u) = \big(B_{1n}(u_1), \dots, B_{dn}(u_d)\big)^\top \quad \text{and} \quad \mathbb{H}(u) = \big(\beta_1(u_1), \dots, \beta_d(u_d)\big)^\top$$

for all $u = (u_1, \dots, u_d) \in [0,1]^d$, where $B_{1n}, \dots, B_{dn}$ are defined as in (22). Thus if $H(u) = u$, then $\mathbb{H}_n = n^{1/2}(H_n - H) \rightsquigarrow \mathbb{H}$ in $\mathcal{D}([0,1]^d; \mathbb{R}^d)$, as $n \to \infty$.

***Definition 4.*** *A rank-based estimator $\theta_n = T_n(U_1, \dots, U_n)$ of $\theta$ is said to belong to class $\mathcal{R}_1$ if it can be written in the form*

$$n^{1/2}(\theta_n - \theta) = n^{-1/2} \sum_{i=1}^n J_\theta\big\{H_n(U_i)\big\} + o_P(1)$$

*in terms of a score function $J_\theta : (0,1)^d \to \mathbb{R}^p$ that satisfies the following regularity conditions for all $\theta \in \mathcal{O}$:*

(a) *$J_\theta$ is twice differentiable and $J_\theta^2$ is integrable with respect to $P_\theta$;*
(b) *$J_\theta$ is standardized, i.e.,*

$$\int J_\theta(u) c_\theta(u)\, \mathrm{d}u = 0 \quad \textit{and} \quad \int J_\theta(u) \dot{c}_\theta(u)\, \mathrm{d}u = I; \tag{24}$$

(c) *there exists a function $h_\theta$ that is integrable with respect to $P_\theta$ for which*

$$\left| \frac{\partial^2}{\partial u_i\, \partial u_j} J_\theta(u) \right| \le h_\theta(u)$$

*for all $i, j \in \{1, \dots, d\}$ and $u \in (0,1)^d$.*

A proof of the following proposition is given in Appendix D.

**Proposition 4.** *Let $\theta_n = T_n(U_1, \dots, U_n)$ be an estimator of $\theta \in \mathcal{O}$ from the class $\mathcal{R}_1$. If $\mathcal{P} \in S(\lambda)$, then $(C_n, \theta_n)$ is $\mathcal{P}$-regular for $\mathcal{C} \times \mathcal{O}$.*

***Example 4.*** *Consider the maximum pseudo-likelihood estimator investigated, e.g., by Genest et al.* [13] *and Shih and Louis* [25]. *Assume its existence and the fact that the score vector defined for all $u \in (0,1)^d$ by*

$$J(u) = I_C^{-1} \frac{\dot{c}^\top(u)}{c(u)}$$

*satisfies the regularity conditions pertaining to class $\mathcal{R}_1$. It then follows from Example* 1 *that this omnibus, rank-based estimator belongs to the class $\mathcal{R}_1$. When $\theta$ is real-valued, other examples include estimates based on the inversion of Spearman's rho or van der Waerden's coefficient. The inversion of Kendall's tau, however, falls into the class $\mathcal{R}_2$ defined below.*

***Definition 5.*** *A rank-based estimator $\theta_n = T_n(U_1, \dots, U_n)$ of $\theta$ is said to belong to class $\mathcal{R}_2$ if it can be written in the form*

$$n^{1/2}(\theta_n - \theta) = n^{-1/2} \sum_{i=1}^n J_\theta\big\{B_n(U_i)\big\} + o_P(1)$$

*in terms of a score function $J_\theta : (0,1) \to \mathbb{R}^p$ that satisfies the same regularity conditions as in Definition* 4 *and such that for all $\theta \in \mathcal{O}$,*

$$\int J_\theta\{C_\theta(u)\}c_\theta(u)\,\mathrm{d}u = 0,$$

*and*

$$\int J_\theta\{C_\theta(u)\}\dot{c}_\theta(u)\,\mathrm{d}u + \int J'_\theta\{C_\theta(u)\}\dot{C}_\theta(u)c_\theta(u)\,\mathrm{d}u = I. \tag{25}$$

A proof of the following proposition is given in Appendix E.

**Proposition 5.** *Let $\theta_n = T_n(U_1,\ldots,U_n)$ be an estimator of $\theta \in \mathcal{O}$ from the class $\mathcal{R}_2$. If $\mathcal{P} \in S(\lambda)$, then $(C_n,\theta_n)$ is $\mathcal{P}$-regular for $\mathcal{C} \times \mathcal{O}$.*

***Example 5.*** *Condition* (25) *holds for "moment-like" parameters satisfying*

$$\theta = g(\mu) \quad \textit{and} \quad \mu = \int M\{C_\theta(u)\}c_\theta(u)\,\mathrm{d}u$$

*for any integrable and continuously differentiable function $M : (0,1) \to \mathbb{R}$ that does not depend on $\theta$. Suppose that $g$ is in fact continuously differentiable, with non-singular derivative $\dot{g}$. Then, by Slutsky's theorem,*

$$J_\theta(t) = \dot{g}\{g^{-1}(\theta)\}\{M(t) - g^{-1}(\theta)\}$$

*for all $t \in (0,1)$. Condition* (25) *holds in that case, because under $P$,*

$$\begin{aligned}
\mathrm{E}(\Theta\mathbb{W}_C^\top) &= \int J\{C(u)\}\dot{c}(u)\,\mathrm{d}u + \int J'\{C(u)\}\dot{C}(u)c(u)\,\mathrm{d}u \\
&= \dot{g}(\tau_0)\left[\int M\{C(u)\}\dot{c}(u)\,\mathrm{d}u + \int M'\{C(u)\}\dot{C}(u)c(u)\,\mathrm{d}u\right] \\
&= \dot{g}(\tau_0)\left[\frac{\partial}{\partial\theta}\int M\{C_\theta(u)\}c_\theta(u)\,\mathrm{d}u\right]_{\theta=\theta_0} \\
&= \dot{g}(\tau_0)\left[\frac{\partial}{\partial\theta}g^{-1}(\theta)\right]_{\theta=\theta_0} = I
\end{aligned}$$

*with $\tau_0 = g^{-1}(\theta_0)$. Suppose, e.g., that $\theta$ is real and that Kendall's tau is defined as in* [1,21] *or* [14] *by*

$$\tau = g^{-1}(\theta) = \frac{1}{2^{d-1}-1}\left\{-1 + 2^d \int C_\theta(u)c_\theta(u)\,\mathrm{d}u\right\}.$$

*If this function and its inverse are continuously differentiable, the parameter estimate $\theta_n = g(\tau_n)$ based on the inversion of Kendall's tau belongs to $\mathcal{R}_2$.*

***Example 6.*** *When it is uniquely defined, the estimator $\theta_n$ minimizing*

$$\varrho_n(\theta) = \int \{C_n(u) - C_\theta(u)\}^2\,\mathrm{d}C_n(u)$$

*between $C_n$ and $C$ also belongs to $\mathcal{R}_2$, provided that*

$$\Sigma_\theta = \int \dot{C}_\theta^\top(u)\dot{C}_\theta(u)c_\theta(u)\,\mathrm{d}u$$

*is non-singular for all $\theta \in \mathcal{O}$. The proof is omitted, as it is similar to the argument described in Example* 3. *A natural goodness-of-fit test could thus be based on the Cramér–von Mises statistic defined as $S_n = \varrho_n(\theta_n)$. See* [27] *for conditions under which $\theta_n$ exists and is unique.*

### 4.3. *Goodness-of-fit for copulas based on Kendall's process*

Keeping the notations of Section 4.2, let $X$ be a continuous $d$-variate random vector with distribution function $F$, margins $F_1, \dots, F_d$, and unique underlying copula $C$. Suppose once again that the null hypothesis of interest is

$$H_0\colon\ C \in \mathcal{C} = \{C_\theta\colon\ \theta \in \mathcal{O}\}.$$

However, suppose that following [15] and [29], it is desired to base a goodness-of-fit test for $C$ on the probability integral transformation $W = F(X) = C(U)$.

Under the assumption that $C = C_\theta$, let the associated Kendall distribution be defined for every $w \in [0, 1]$ by

$$K_\theta(w) = \mathrm{P}\{C_\theta(U) \le w\} = \mathrm{P}\{F_\theta(X) \le w\}.$$

Given a random sample $X_1, \dots, X_n$ from $F$ and an estimator $\theta_n = T_n(U_1, \dots, U_n)$ of $\theta \in \mathbb{R}^p$, a parametric estimate of $K$ is then given by $K_{\theta_n}$. As argued by Genest and Rivest [17] and Barbe et al. [1], a reasonable test of $H_0$ can be based on a continuous functional $S_n = \phi(\mathbb{G}_n^K)$ of the process

$$\mathbb{G}_n^K = n^{1/2}(K_n - K_{\theta_n}),$$

where $K_n$ is the nonparametric estimator of $K$ defined by (6). Although these tests are not generally consistent, they are sometimes more powerful than procedures based on $C_n$; see, e.g., Genest et al. [16]. The fact that the process $\mathbb{G}_n^K$ is univariate also makes it possible to assess the fit visually, in addition to formal tests; on this point, see [17].

As in the previous section, the estimator $\theta_n$ is required to be margin-free, i.e., rank-based. Under this condition, both $K_n$ and $\theta_n$ are measurable with respect to the sigma-algebra $\mathcal{U}_n$ generated by $U_1, \dots, U_n$. For, an equivalent alternative expression for $K_n$ is given for every $w \in [0, 1]$ by

$$K_n(w) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}\{F_n(X_i) \le w\} = \frac{1}{n}\sum_{i=1}^n \mathbf{1}\{B_n(U_i) \le w\},$$

where $B_n$ is defined in (21).

To establish the validity of the parametric bootstrap for statistics of the form $S_n = \phi(\mathbb{G}_n^K)$, one can use Theorems 1 and 2 with $A_\theta = K_\theta$ and $P_\theta$ standing for the unique probability measure associated with $C_\theta$. To this end, assume that $\mathcal{P} = \{P_\theta\colon\ \theta \in \mathcal{O}\} \in S(\lambda)$, where $\lambda$ is Lebesgue's measure on $[0, 1]^d$.

If $Q_\theta$ denotes the probability measure with distribution function $K_\theta$, further assume that $\mathcal{Q} = \{Q_\theta\colon\ \theta \in \mathcal{O}\} \in S(\nu)$, where $\nu$ is Lebesgue's measure on $[0, 1]$. Finally, introduce the notations

$$K = K_{\theta_0}, \qquad k = k_{\theta_0}, \qquad \mathbb{K}_n = n^{1/2}(K_n - K).$$

To check the $\mathcal{P}$-regularity of $A_n = K_n$, consider the process defined for all $w \in [0, 1]$ by

$$\alpha_n(w) = n^{-1/2}\sum_{i=1}^n \bigl[\mathbf{1}\{C(U_i) \le w\} - K(w)\bigr]. \tag{26}$$

Theorem 1 in [1] implies that, as $n \to \infty$,

$$(\alpha_n, \mathbb{B}_n, \mathbb{K}_n, \mathbb{W}_{C,n}) \rightsquigarrow (\alpha, \mathbb{B}, \mathbb{K}, \mathbb{W}_C)$$

in $\mathcal{D}([0, 1]; \mathbb{R})^{\otimes 3} \times \mathbb{R}^p$, where $(\alpha, \mathbb{B}, \mathbb{K}, \mathbb{W}_C)$ is a continuous centered Gaussian process. In addition, $\mathbb{K}(w) = \alpha(w) - \mu(w, \mathbb{B})$, where

$$\mu(w, g) = k(w)\mathrm{E}\{g(U) | C(U) = w\},$$

is well defined for every real-valued continuous function $g$ on $[0, 1]^d$ and all $w \in [0, 1]$ by Condition II in Appendix F.

Given the calculations in Appendix G, the desired result is then the following.

**Proposition 6.** *Let $X_1, \ldots, X_n$ be a random sample from distribution $F$ with unique underlying copula $C = C_{\theta_0}$ for some $\theta_0 \in \mathcal{O}$. Suppose that $\mathcal{P} \in \mathcal{S}(\lambda)$ and $\mathcal{Q} \in \mathcal{S}(\nu)$. Assume also that the density $k_\theta$ of $Q_\theta$ satisfies Conditions* I *and* II *described in Appendix* F. *Then $K_n$ is $\mathcal{P}$-regular for $\mathcal{K} = \{K_\theta\colon \theta \in \mathcal{O}\}$.*

Now suppose that the sequence $\theta_n$ is $\mathcal{P}$-regular for $\mathcal{O}$ and that, as $n \to \infty$,

$$(\mathbb{K}_n, \Theta_n, \mathbb{W}_{C,n}) \rightsquigarrow (\mathbb{K}, \Theta, \mathbb{W}_C)$$

in $\mathcal{D}([0, 1]; \mathbb{R}) \times \mathbb{R}^{p\otimes 2}$, where the weak limit is Gaussian. It then follows that $(K_n, \theta_n)$ is $\mathcal{P}$-regular for $\mathcal{K} \times \mathcal{O}$ because $\mathrm{E}(\mathbb{K}\mathbb{W}_C^\top) = \dot{K}$ by Proposition 6 and $\mathrm{E}(\Theta\mathbb{W}_C^\top) = I$ by the regularity hypothesis on $\theta_n$.

In the light of Propositions 4–6, one then gets the following result.

**Proposition 7.** *Let $\theta_n = T_n(U_1, \ldots, U_n)$ be an estimator of $\theta \in \mathcal{O}$ from the class $\mathcal{R}_1 \cup \mathcal{R}_2$. Suppose that $\mathcal{P} \in \mathcal{S}(\lambda)$, $\mathcal{Q} \in \mathcal{S}(\nu)$ and that the density $k_\theta$ of $Q_\theta$ satisfies Conditions* I *and* II *described in Appendix* F. *Then the sequence $(K_n, \theta_n)$ is $\mathcal{P}$-regular for $\mathcal{K} \times \mathcal{O}$.*

Finally, given that $A_n \neq C_n$ in this particular application, the question of what should serve as $\check{A}_n$ must be addressed. Two natural choices are:

(a) Generate a random sample $V_1, \ldots, V_m$ from $P_\theta$, define

$$\hat{W}_i = \frac{1}{m}\sum_{j=1}^{m} \mathbf{1}(V_j \le V_i)$$

for each $i \in \{1, \ldots, m\}$ and, for all $w \in [0, 1]$, let

$$\check{A}_n(w) = \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}(\hat{W}_i \le w). \tag{27}$$

(b) Generate a random sample $W_1, \ldots, W_m$ from $Q_\theta$ with associated distribution function $K_\theta$; then for each $w \in [0, 1]$, let

$$\check{A}_n(w) = \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}(W_i \le w). \tag{28}$$

The conditions for the regularity of these estimators are delineated in the following result, whose proof is immediate from Propositions 1 and 6.

**Proposition 8.** *Suppose that $\mathcal{P} \in \mathcal{S}(\lambda)$, $\mathcal{Q} \in \mathcal{S}(\nu)$ and that the density $k_\theta$ of $Q_\theta$ satisfies Conditions* I *and* II *described in Appendix* F. *Then the sequence $\check{A}_n$ is $\mathcal{P}$-regular for $\mathcal{K}$ when defined by* (27) *and it is $\mathcal{Q}$-regular for $\mathcal{K}$ when defined by* (28).

With either one of these choices for $\check{A}_n$, therefore, the conditions are assembled for the application of Theorems 1 and 2. Consequently, the one- and two-level parametric bootstraps yield valid approximations of the distribution of any continuous functional $S_n = \phi(\mathbb{G}_n^K)$.

### 4.4. *Goodness-of-fit testing using Durbin's approach*

In the same context as in Section 4.1, but calling on the notion of probability integral transform discussed in Section 4.3, one could base a test of

$$H_0\colon\ F \in \mathcal{F} = \{F_\theta\colon\ \theta \in \mathcal{O}\}$$

on the distribution $K$ of $F(X)$, i.e., Kendall's distribution. A parametric estimate under $H_0$ is given for all $w \in [0, 1]$ by

$$D_n(w) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\big\{F_{\theta_n}(X_i) \le w\big\},$$

where $\theta_n = T(X_1, \dots, X_n)$ is a consistent estimator of $\theta$.

A goodness-of-fit test could thus be based on some continuous functional of

$$\mathbb{G}_n^D = n^{1/2}(D_n - K_{\theta_n}).$$

This proposal is considered by Durbin [9] in the univariate case, where it can be seen to yield a consistent test. Although the multivariate extension investigated here is not always consistent in dimension $d > 1$, its univariate character allows for a graphical assessment of goodness-of-fit, which may be an advantage in some circumstances.

To establish the validity of the parametric bootstrap for statistics of the form $S_n = \phi(\mathbb{G}_n^D)$, one can use Theorems 1 and 2 with $A_\theta = K_\theta$ and $P_\theta$ standing for the unique probability measure associated with $F_\theta$. To this end, assume that $\mathcal{P} = \{P_\theta\colon\ \theta \in \mathcal{O}\} \in S(\lambda)$, where $\lambda$ is Lebesgue's measure on $\mathbb{R}^d$. If $Q_\theta$ denotes the probability measure with distribution function $K_\theta$, further assume that $\mathcal{Q} = \{Q_\theta\colon\ \theta \in \mathcal{O}\} \in S(\nu)$, where $\nu$ is Lebesgue's measure on [0, 1].

Let $K = K_{\theta_0}$ and $k = k_{\theta_0}$ as before, and introduce also

$$\mathbb{D}_n = n^{1/2}(D_n - K).$$

To check the $\mathcal{P}$-regularity of $A_n = D_n$, note that the process $\alpha_n$ defined in (26) can also be written as follows for all $w \in [0, 1]$:

$$\alpha_n(w) = n^{-1/2}\sum_{i=1}^{n}\big[\mathbf{1}\big\{F(X_i) \le w\big\} - K(w)\big].$$

If $\mathbb{W}_{F,n}$ is defined as in (18), one can then call on the methodology developed in [18] to show that under $P$,

$$(\alpha_n, \mathbb{D}_n, \mathbb{W}_{F,n}) \rightsquigarrow (\alpha, \mathbb{D}, \mathbb{W}_F)$$

in $\mathcal{D}([0, 1]; \mathbb{R})^{\otimes 2} \times \mathbb{R}^p$, as $n \to \infty$. Here, the weak limit $(\alpha, \mathbb{D}, \mathbb{W}_F)$ is a continuous centered Gaussian process, and for all $w \in [0, 1]$,

$$\mathbb{D}(w) = \alpha(w) - \kappa(w, \dot{F})\Theta$$

with $\kappa$ defined for every real-valued continuous function $g$ on $[0, 1]^d$ and all $w \in [0, 1]$ by

$$\kappa(w, g) = k(w)\mathrm{E}\big\{g(X)|F(X) = w\big\}.$$

Note that it follows from a remark at the end of Appendix G that $\mu(w, \dot{C}) = \kappa(w, \dot{F})$ when both $\dot{F}$ and $\dot{C}$ exist.

Given the calculations in Appendix H, the desired result is then the following.

**Proposition 9.** *Let $X_1, \dots, X_n$ be a random sample from distribution $F = F_{\theta_0}$ for some $\theta_0 \in \mathcal{O}$. Suppose that $\mathcal{P} \in \mathcal{S}(\lambda)$ and $\mathcal{Q} \in \mathcal{S}(\nu)$. Assume also that the density $k_\theta$ of $Q_\theta$ satisfies Conditions* I *and* II *described in Appendix* F. *Then $D_n$ is $\mathcal{P}$-regular for $\mathcal{K} = \{K_\theta : \theta \in \mathcal{O}\}$.*

Now suppose that the sequence $\theta_n$ is $\mathcal{P}$-regular for $\mathcal{O}$ and that, as $n \to \infty$,

$$(\mathbb{D}_n, \Theta_n, \mathbb{W}_{F,n}) \rightsquigarrow (\mathbb{D}, \Theta, \mathbb{W}_F)$$

in $\mathcal{D}([0,1];\mathbb{R}) \times \mathbb{R}^{p\otimes 2}$, where the weak limit is Gaussian. It then follows that $(K_n, \theta_n)$ is $\mathcal{P}$-regular for $\mathcal{K} \times \mathcal{O}$ because $\mathrm{E}(\mathbb{D}\mathbb{W}_F^\top) = \dot{K}$ by Proposition 9 and $\mathrm{E}(\Theta\mathbb{W}_F^\top) = I$ by the regularity hypothesis on $\theta_n$.

In the light of Proposition 2, one then gets the following result.

**Proposition 10.** *Let $\theta_n = T_n(X_1,\ldots,X_n)$ be an estimator of $\theta \in \mathcal{O}$ from the class $\mathcal{R}$. Suppose that $\mathcal{P} \in \mathcal{S}(\lambda)$, $\mathcal{Q} \in \mathcal{S}(\nu)$ and that the density $k_\theta$ of $Q_\theta$ satisfies Conditions* I *and* II *described in Appendix* F. *Then the sequence $(D_n, \theta_n)$ is $\mathcal{P}$-regular for $\mathcal{K} \times \mathcal{O}$.*

Finally, given that $A_n \neq F_n$, the issue of what should serve as $\check{A}_n$ must again be addressed. Here, the most natural choices are:

(a) Generate a random sample $Y_1,\ldots,Y_m$ from $P_\theta$, define

$$\hat{W}_i = \frac{1}{m}\sum_{j=1}^{m} \mathbf{1}(Y_j \le Y_i)$$

for $i \in \{1,\ldots,m\}$ and, for all $w \in [0,1]$, let

$$\check{A}_n(w) = \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}(\hat{W}_i \le w). \tag{29}$$

(b) Generate a random sample $W_1,\ldots,W_m$ from $Q_\theta$ with associated distribution function $K_\theta$; then for each $w \in [0,1]$, let

$$\check{A}_n(w) = \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}(W_i \le w). \tag{30}$$

The conditions for the regularity of these estimators are delineated in the following result, whose proof is immediate from Propositions 1 and 9.

**Proposition 11.** *Suppose that $\mathcal{P} \in \mathcal{S}(\lambda)$, $\mathcal{Q} \in \mathcal{S}(\nu)$ and that the density $k_\theta$ of $Q_\theta$ satisfies Conditions* I *and* II *described in Appendix* F. *Then the sequence $\check{A}_n$ is $\mathcal{P}$-regular for $\mathcal{K}$ when defined by* (29) *and it is $\mathcal{Q}$-regular for $\mathcal{K}$ when defined by* (30).

With either one of these choices for $\check{A}_n$, therefore, the conditions are assembled for the application of Theorems 1 and 2. Consequently, the one- and two-level parametric bootstraps yield valid approximations of the distribution of any continuous functional $S_n = \phi(\mathbb{G}_n^D)$.

## 5. An illustration

Stute et al. [26] and Henze [20] show the usefulness of the parametric bootstrap in testing the goodness-of-fit of multivariate continuous and discrete distributions, respectively. This methodology is also applied with success by Genest et al. [15,16] and Dobrić and Schmid [8] in copula modeling contexts. This section, therefore, is limited to a short illustration.

Consider the problem of testing that subject to appropriate transformations of its margins, a continuous $d$-variate random vector $X$ is Gaussian. In other words, suppose that one wants to check whether there exists a $d \times d$ correlation matrix $\Sigma$ for which the underlying copula $C$ of $X$ is of the form

$$C_\Sigma(u) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \int_{-\infty}^{\Phi^{-1}(u_1)} \cdots \int_{-\infty}^{\Phi^{-1}(u_d)} \exp\left(-\frac{1}{2} z^\top \Sigma^{-1} z\right) \mathrm{d}z_d \cdots \mathrm{d}z_1,$$

where $z = (z_1, \dots, z_d)^\top$, $u = (u_1, \dots, u_d) \in (0,1)^d$, and $\Phi$ denotes the cumulative distribution function of a standard $\mathcal{N}(0,1)$ random variable.

This problem, which is of current interest in finance, is considered, e.g., by Breymann et al. [6] and Malevergne and Sornette [23]. The tests they propose are based on specific properties of the multivariate Gaussian distribution. Because the asymptotic behavior of their statistics is unwieldy, however, they approximate it by the distribution that would obtain if the copula parameters and the univariate margins of the data were known. Unfortunately, this yields unreliable $P$-values as shown, e.g., by Dobrić and Schmid [8]. Thanks to the parametric bootstrap, however, it is possible to bypass these issues entirely.

For the *general* problem of testing hypothesis (4) that a copula $C$ belongs to a given parametric copula family $\mathcal{C} = \{C_\theta\colon \theta \in \mathcal{O}\}$, a simple "blanket procedure" would be to reject $H_0$ for large values of the Cramér–von Mises statistic

$$\mathfrak{S}_n = n \int \{C_n(u) - C_{\theta_n}(u)\}^2 \, \mathrm{d}C_n(u) = \sum_{i=1}^{n} \{C_n(\hat{U}_i) - C_{\theta_n}(\hat{U}_i)\}^2. \tag{31}$$

When $\theta_n$ is a rank-based estimator of $\theta$ from the class $\mathcal{R}_1 \cup \mathcal{R}_2$,

$$\mathfrak{S}_n = \int \{\mathbb{G}_n^C(u)\}^2 \, \mathrm{d}C(u) + \mathrm{o}_P(1)$$

is an approximation of a continuous functional of the empirical process $\mathbb{G}_n^C = n^{1/2}(C_n - C_{\theta_n})$. Theorem 1 then guarantees that the parametric bootstrap yields valid $P$-values for $\mathfrak{S}_n$, provided that $\mathcal{P} \in \mathcal{S}(\lambda)$. One would resort to a one- or two-level procedure, depending whether the exact value of $C_\theta(u)$ could or could not be computed easily at each $u = \hat{U}_i$, $i \in \{1, \dots, n\}$.

The requirements are met for Gaussian copulas. In that case, Klaassen and Wellner [22] show that an efficient rank-based estimation of $\theta = \Sigma$ is given by the matrix $\theta_n = (\hat{\sigma}_{jk})$ whose entries are the van der Waerden correlations, viz.

$$\hat{\sigma}_{jk} = \sum_{i=1}^{n} \Phi^{-1}\left(\frac{R_{ij}}{n+1}\right) \Phi^{-1}\left(\frac{R_{ik}}{n+1}\right) \Big/ \sum_{i=1}^{n} \Phi^{-1}\left(\frac{i}{n+1}\right)^2, \tag{32}$$

for all $j, k \in \{1, \dots, d\}$. Here, $R_{ij}$ is the rank of $X_{ij}$ among $X_{1j}, \dots, X_{nj}$ for all $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, d\}$.

The one- and two-level parametric bootstrap algorithms are detailed below in the general case. A user could call on either one, depending whether a numerical integration routine is available or not for $C_\theta(u)$ at arbitrary $u \in [0,1]^d$.

*One-level parametric bootstrap procedure for* $\mathfrak{S}_n$

1. Convert the data into rank vectors $R_i = (R_{i1}, \dots, R_{id})^\top$, $i \in \{1, \dots, n\}$.
2. Put $\hat{U}_i = R_i/(n+1)$ for $i \in \{1, \dots, n\}$ and for all $u \in [0,1]^d$, let

$$C_n(u) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(\hat{U}_i \le u).$$

3. Estimate $\theta$ by a rank-based estimator $\theta_n$; e.g., use formula (32) when testing that the copula is Gaussian.
4. Compute $\mathfrak{S}_n$ using formula (31).
5. Pick $N$ large and repeat the following steps for every $k \in \{1, \dots, N\}$:
   (a) Generate a random sample $X_{1,k}^*, \dots, X_{n,k}^*$ from copula $C_{\theta_n}$ and compute the associated rank vectors $R_{1,k}^*, \dots, R_{n,k}^*$.
   (b) Put $\hat{U}_{i,k}^* = R_{i,k}^*/(n+1)$ for $i \in \{1, \dots, n\}$ and for all $u \in [0,1]^d$, let

$$C_{n,k}^*(u) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(\hat{U}_{i,k}^* \le u).$$

(c) Construct an estimate $\theta^*_{n,k}$ of $\theta$ by the same rank-based method as in Step 3; e.g., when testing that the copula is Gaussian, substitute $\hat{U}^*_{i,k} = R^*_{i,k}/(n+1)$ for $\hat{U}_i = R_i/(n+1)$ in formula (32).

(d) Compute

$$\mathfrak{S}^*_{n,k} = \sum_{i=1}^{n} \{C^*_{n,k}(\hat{U}^*_{i,k}) - C_{\theta^*_{n,k}}(\hat{U}^*_{i,k})\}^2.$$

An approximate $P$-value for the test based on $\mathfrak{S}_n$ is then given by

$$\frac{1}{N}\sum_{k=1}^{N} \mathbf{1}(\mathfrak{S}^*_{n,k} > \mathfrak{S}_n).$$

*Two-level parametric bootstrap procedure for $\mathfrak{S}_n$*

1. Convert the data into rank vectors $R_i = (R_{i1}, \ldots, R_{id})^\top$, $i \in \{1, \ldots, n\}$.
2. Put $\hat{U}_i = R_i/(n+1)$ for $i \in \{1, \ldots, n\}$ and for all $u \in [0,1]^d$, let

$$C_n(u) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(\hat{U}_i \le u).$$

3. Estimate $\theta$ by a rank-based estimator $\theta_n$; e.g., use formula (32) when testing that the copula is Gaussian.
4. Pick $m$ much larger than $n$:
   (a) Generate a random sample $V^*_1, \ldots, V^*_m$ from copula $C_{\theta_n}$.
   (b) Approximate $C_{\theta_n}(u)$ at each $u \in [0,1]^d$ by

$$\check{C}^*_n(u) = \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}(V^*_i \le u).$$

   (c) Compute

$$\mathfrak{S}_n = n \int \{C_n(u) - \check{C}^*_n(u)\}^2 \, \mathrm{d}C_n(u) = \sum_{i=1}^{n} \{C_n(\hat{U}_i) - \check{C}^*_n(\hat{U}_i)\}^2.$$

5. Pick $N$ large and repeat the following steps for every $k \in \{1, \ldots, N\}$:
   (a) Generate a random sample $X^*_{1,k}, \ldots, X^*_{n,k}$ from copula $C_{\theta_n}$ and compute the associated rank vectors $R^*_{1,k}, \ldots, R^*_{n,k}$.
   (b) Put $\hat{U}^*_{i,k} = R^*_{i,k}/(n+1)$ for $i \in \{1, \ldots, n\}$ and for all $u \in [0,1]^d$, let

$$C^*_{n,k}(u) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(\hat{U}^*_{i,k} \le u).$$

   (c) Construct an estimate $\theta^*_{n,k}$ of $\theta$ by the same rank-based method as in Step 3; e.g., when testing that the copula is Gaussian, substitute $\hat{U}^*_{i,k} = R^*_{i,k}/(n+1)$ for $\hat{U}_i = R_i/(n+1)$ in formula (32).
   (d) For the same integer $m$ as in Step 4:
      (i) Generate a random sample $V^{**}_{1,k}, \ldots, V^{**}_{m,k}$ from copula $C_{\theta^*_{n,k}}$.
      (ii) Approximate $C_{\theta^*_{n,k}}(u)$ at each $u \in [0,1]^d$ by

$$\check{C}^{**}_{n,k}(u) = \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}(V^{**}_{i,k} \le u).$$

(iii) Compute

$$\mathfrak{S}_{n,k}^* = \sum_{i=1}^{n} \left\{ C_{n,k}^*\left(\hat{U}_{i,k}^*\right) - \check{C}_{n,k}^{**}\left(\hat{U}_{i,k}^*\right) \right\}^2.$$

An approximate $P$-value for the test based on the Cramér–von Mises statistic $\mathfrak{S}_n$ is then given by

$$\frac{1}{N} \sum_{k=1}^{N} \mathbf{1}\left(\mathfrak{S}_{n,k}^* > \mathfrak{S}_n\right).$$

To illustrate the validity of the parametric bootstrap for the Cramér–von Mises statistic $\mathfrak{S}_n$, 10,000 random samples of size $n = 250$ were generated from the bivariate Gaussian copula with correlation $\theta = 1/4$. The null hypothesis was then tested at the 5% level using $N = 1000$ bootstrap samples and a numerical integration routine for $C_\theta$. As an alternative, the copula was also estimated by a two-level bootstrap procedure using $m = 100{,}000$.

The results for the one-level parametric bootstrap are given in the first line of Table 1, along with those of a similar experiment carried out with sample size $n = 500$. Figures for the two-level bootstrap (not reported) are very similar. The quality of the approximation is seen to be excellent.

To check the power of the goodness-of-fit test based on $\mathfrak{S}_n$, samples of size $n = 250$ and 500 were also generated from Student copulas with various degrees of freedom (df) but the same correlation as under the Gaussian model; for additional information about this class of meta-elliptical copulas, refer to [7]. Those results may also be found in Table 1. As expected, the test quickly gains in power as the degrees of freedom get smaller. For examples involving other copula models and extensive comparisons with alternative goodness-of-fit tests, see [16].

## 6. Conclusion

This paper shows the validity of the parametric bootstrap for testing the goodness-of-fit of a class $\mathcal{P} = \{P_\theta \colon \theta \in \mathcal{O}\}$ of probability measures whenever the sequences $A_n$ and $\theta_n$ of estimators of $A_\theta$ and $\theta$ are $\mathcal{P}$-regular. For situations where the distribution function associated with $P_\theta$ is not available in closed form, a two-level extension of the parametric bootstrap is also developed and shown to be valid.

The results proved herein are obtained under conditions that are generally easier to verify than those of [26] or [3]. While these authors limited their investigation to parametric contexts, the approach described here also applies in semiparametric settings and is illustrated in four situations commonly encountered in practice.

In particular, a one- or two-level parametric bootstrap approach is valid in goodness-of-fit testing for copula models using either the empirical copula process or its associated Kendall process, as discussed by Genest et al. [15,16]. In the latter paper as in [8], the authors also consider tests of goodness-of-fit based on Rosenblatt's transformation. It is easy to check that the parametric bootstrap methodology also applies to this case.

Table 1
Percentage of rejection of the null hypothesis that a copula is Gaussian, based on the statistic $\mathfrak{S}_n$ whose 95% critical value is estimated from $N = 1000$ parametric bootstrap samples. The power, observed over 10,000 replicates, is presented as a function of the sample size $n$ and of the meta-elliptical copula used to generate the data. In all models, the correlation $\theta$ between the two variables is fixed at $1/4$.

| Copula model | $n = 250$ | $n = 500$ |
|---|---|---|
| Gaussian | 5.09 | 5.08 |
| Student (20 df) | 6.45 | 6.13 |
| Student (10 df) | 8.28 | 9.38 |
| Student (5 df) | 14.88 | 21.20 |
| Student (2.5 df) | 43.51 | 75.46 |
| Student (2 df) | 63.17 | 94.22 |
| Student (1.5 df) | 87.56 | 99.93 |

## Appendix A. Auxiliary results

Let $\mathcal{P} = \{P_\theta\colon \theta \in \mathcal{O}\}$ be a parametric family of distributions and assume that $\mathcal{P} \in \mathcal{S}(\lambda)$ for some reference measure $\lambda$ which is independent of $\theta$. Let $U_1, \ldots, U_n, U_1^*, \ldots, U_n^*$ be mutually independent observations from $P = P_{\theta_0}$ for some $\theta_0 \in \mathcal{O}$. Write $p_\theta = \mathrm{d}P_\theta/\mathrm{d}\lambda$ and $p = p_{\theta_0}$. Finally, let $\theta_n = T_n(U_1, \ldots, U_n)$ be an estimator of $\theta$ and introduce

$$\Theta_n = n^{1/2}(\theta_n - \theta_0), \qquad \mathbb{W}_{P,n} = n^{-1/2}\sum_{i=1}^n \frac{\dot{p}^\top(U_i)}{p(U_i)}, \qquad \ell_n = \sum_{i=1}^n \log\left\{\frac{p_{\theta_n}(U_i^*)}{p(U_i^*)}\right\}.$$

The following result, which concerns the weak limit of the pair $(\ell_n, \Theta_n)$, is instrumental in establishing Theorem 1.

**Lemma 1.** *Suppose that the sequence $(\mathbb{W}_{P,n}, \Theta_n)$ converges weakly, as $n \to \infty$, and that the joint distribution of the limit $(\mathbb{W}_P, \Theta)$ is $\mathcal{N}(0, \Sigma)$ with*

$$\Sigma = \begin{pmatrix} I_P & \Gamma^\top \\ \Gamma & \Lambda \end{pmatrix}, \quad \Gamma = \mathrm{E}\bigl(\Theta\mathbb{W}_P^\top\bigr), \quad \Lambda = \mathrm{E}\bigl(\Theta\Theta^\top\bigr)$$

*and $I_P$ defined as in* (14). *There exists an independent copy $\mathbb{W}_P^\perp$ of $\mathbb{W}_P$, also independent of $\Theta$, such that, as $n \to \infty$,*

$$(\ell_n, \Theta_n) \rightsquigarrow \bigl(\Theta^\top W_P^\perp - \Theta^\top I_P\Theta/2, \Theta\bigr).$$

**Proof.** When $\mathcal{P} \in \mathcal{S}(\lambda)$, the sequence

$$\mathbb{W}_{P,n}^* = n^{-1/2}\sum_{i=1}^n \frac{\dot{p}^\top(U_i^*)}{p(U_i^*)}$$

is known to have a weak limit, say $\mathbb{W}_P^\perp$, which has the same distribution as $\mathbb{W}_P$ but is independent from it and from $\Theta$. When $\|\Theta_n\| \le M$, one can write

$$\begin{aligned}\ell_n &= \sum_{i=1}^n \bigl[\log\bigl\{p_{\theta_n}\bigl(U_i^*\bigr)\bigr\} - \log\bigl\{p\bigl(U_i^*\bigr)\bigr\}\bigr] \\ &= \Theta_n^\top\mathbb{W}_{P,n}^* + \frac{1}{2}\Theta_n^\top\left[\frac{1}{n}\sum_{i=1}^n\left\{\frac{\ddot{p}(U_i^*)}{p(U_i^*)} - \frac{\dot{p}^\top(U_i^*)\dot{p}(U_i^*)}{p^2(U_i^*)}\right\}\right]\Theta_n + R_n,\end{aligned}$$

where

$$|R_n| \le \frac{M^2}{2n}\sum_{i=1}^n \sup_{\|\theta-\theta_0\|\le Mn^{-1/2}}\left\{\left\|\frac{\ddot{p}_\theta(U_i^*)}{p_\theta(U_i^*)} - \frac{\ddot{p}(U_i^*)}{p(U_i^*)}\right\| + \left\|\frac{\dot{p}_\theta^\top(U_i^*)\dot{p}_\theta(U_i^*)}{p_\theta^2(U_i^*)} - \frac{\dot{p}^\top(U_i^*)\dot{p}(U_i^*)}{p^2(U_i^*)}\right\|\right\}$$

can be made arbitrarily small with probability close to one because of part 1.4 of Definition 1. Using the tightness of the sequence $\Theta_n$ and the fact that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n\left\{\frac{\ddot{p}(U_i^*)}{p(U_i^*)} - \frac{\dot{p}^\top(U_i^*)\dot{p}(U_i^*)}{p^2(U_i^*)}\right\} = -I_P \quad P \text{ almost surely,}$$

one can see that, as $n \to \infty$, $\ell_n \rightsquigarrow \Theta^\top\mathbb{W}_P^\perp - \Theta^\top I_P\Theta/2$, whence the result. $\square$

Next, assume that $\mathcal{Q} = \{Q_\theta\colon \theta \in \mathcal{O}\} \in \mathcal{S}(\nu)$ for some reference measure $\nu$ which is independent of $\theta$. Let $V_1, \ldots, V_m$, $V_1^*, \ldots, V_m^*$, $V_1^{**}, \ldots, V_m^{**}$ be observations from $Q = Q_{\theta_0}$. Write $q_\theta = \mathrm{d}Q_\theta/\mathrm{d}\nu$ and $q = q_{\theta_0}$. Assume that

$$U_1, \ldots, U_n, \qquad U_1^*, \ldots, U_n^*, \qquad V_1, \ldots, V_m, \qquad V_1^*, \ldots, V_m^*, \qquad V_1^{**}, \ldots, V_m^{**},$$

are mutually independent and let $\mathcal{P}_n$ denote their joint probability measure.

Let $\theta_n^* = T_n(U_1^*, \dots, U_n^*)$ be an estimator of $\theta$ and set

$$\Theta_n^* = n^{1/2}\big(\theta_n^* - \theta\big).$$

The following result, which is required for the proof of Theorem 2, concerns the joint limiting behavior of the sequences $\Theta_n$, $\Theta_n^*$ and the logarithm of

$$\frac{\mathrm{d}\mathcal{P}_n^*}{\mathrm{d}\mathcal{P}_n} = \left\{\prod_{i=1}^n \frac{p_{\theta_n}(U_i^*)}{p(U_i^*)}\right\} \times \left\{\prod_{i=1}^m \frac{q_{\theta_n}(V_i^*)}{q(V_i^*)}\right\} \times \left\{\prod_{i=1}^m \frac{q_{\theta_n^*}(V_i^{**})}{q(V_i^{**})}\right\}.$$

**Lemma 2.** *Suppose that the sequence* $(\mathbb{W}_{P,n}, \mathbb{W}_{Q,n}, \Theta_n)$ *converges weakly, as* $n \to \infty$, *and that the distribution of the limit* $(\mathbb{W}_P, \mathbb{W}_Q, \Theta)$ *is* $\mathcal{N}(0, \Delta)$ *with*

$$\Delta = \begin{pmatrix} I_P & 0 & \Gamma^\top \\ 0 & I_Q & 0 \\ \Gamma & 0 & \Lambda \end{pmatrix}, \quad \Gamma = \mathrm{E}\big(\Theta \mathbb{W}_P^\top\big), \quad \Lambda = \mathrm{E}\big(\Theta\Theta^\top\big),$$

*where* $I_P$ *and* $I_Q$ *are defined as in* (14) *and* (17), *respectively. There exist an independent copy* $(\mathbb{W}_P^\perp, \Theta^\perp)$ *of* $(\mathbb{W}_P, \Theta)$ *and mutually independent copies* $\mathbb{W}_Q^\perp$ *and* $\mathbb{W}_Q^{\perp\perp}$ *of* $\mathbb{W}_Q$ *that are independent of* $\mathbb{W}_P$ *and* $\mathbb{W}_P^\perp$, $\Theta$ *and* $\Theta^\perp$, *such that, as* $n \to \infty$,

$$\left(\log\left(\frac{\mathrm{d}\mathcal{P}_n^*}{\mathrm{d}\mathcal{P}_n}\right), \Theta_n, \Theta_n^*\right)$$
$$\rightsquigarrow \left(\Theta^\top \mathbb{W}_P^\perp + \Theta^\top \mathbb{W}_Q^\perp - \frac{1}{2}\Theta^\top I_P \Theta - \frac{1}{2}\Theta^\top I_Q \Theta + \big(\Theta^\perp\big)^\top \mathbb{W}_Q^{\perp\perp} - \frac{1}{2}\big(\Theta^\perp\big)^\top I_Q \Theta^\perp, \Theta, \Theta^\perp\right).$$

**Proof.** Let $\mathbb{W}_P^\perp$ be the weak limit of the sequence $\mathbb{W}_{P,n}^*$, as in the proof of Lemma 1. When $\mathcal{Q} \in \mathcal{S}(\nu)$, the sequences

$$\mathbb{W}_{Q,n} = n^{-1/2}\sum_{i=1}^m \frac{\dot{q}^\top(V_i)}{q(V_i)}, \qquad \mathbb{W}_{Q,n}^* = n^{-1/2}\sum_{i=1}^m \frac{\dot{q}^\top(V_i^*)}{q(V_i^*)}, \qquad \mathbb{W}_{Q,n}^{**} = n^{-1/2}\sum_{i=1}^m \frac{\dot{q}^\top(V_i^{**})}{q(V_i^{**})}$$

are known to have weak limits, denoted $\mathbb{W}_Q$, $\mathbb{W}_Q^\perp$ and $\mathbb{W}_Q^{\perp\perp}$, respectively. Moreover, the latter are mutually independent and identically distributed, in addition to being independent of $\mathbb{W}_P$, $\mathbb{W}_P^\perp$, $\Theta$ and $\Theta^\perp$.

Proceeding as in the proof of Lemma 1, one can deduce that when $\|\Theta_n\| \le M$,

$$\log\left(\frac{\mathrm{d}\mathcal{P}_n^*}{\mathrm{d}\mathcal{P}_n}\right) = \Theta_n^\top \mathbb{W}_{P,n}^* + \Theta_n^\top \mathbb{W}_{Q,n}^* - \frac{1}{2}\Theta_n^\top I_P \Theta_n - \frac{1}{2}\Theta_n^\top I_Q \Theta_n + \big(\Theta_n^*\big)^\top \mathbb{W}_{Q,n}^{**} - \frac{1}{2}\big(\Theta_n^*\big)^\top I_Q \Theta_n^* + R_n,$$

where $|R_n|$ can be made arbitrarily small with probability close to one. Given that the sequence $(\Theta_n)$ is tight, the conclusion follows by construction and the fact that $m/n \to \gamma$, as $n \to \infty$. □

## Appendix B. Proof of Theorem 1

The proof is based on Le Cam's Third Lemma as stated, e.g., by van der Vaart and Wellner [28]. Thus, assume at first that $U_1, \dots, U_n, U_1^*, \dots, U_n^*$ are mutually independent random vectors with probability measure $P = P_{\theta_0}$ for some $\theta_0 \in \mathcal{O}$. Denote by $\mathcal{P}_n$ their joint probability measure.

Let $\theta_n = T_n(U_1, \dots, U_n)$ and $\theta_n^* = T_n(U_1^*, \dots, U_n^*)$ be estimators of $\theta$ and write

$$\Theta_n = n^{1/2}(\theta_n - \theta), \qquad \Theta_n^* = n^{1/2}\big(\theta_n^* - \theta\big).$$

Similarly, let $A_n = \Upsilon_n(U_1, \dots, U_n)$ and $A_n^* = \Upsilon_n(U_1^*, \dots, U_n^*)$ be estimators of $A = A_{\theta_0}$ and introduce

$$\mathbb{A}_n = n^{1/2}(A_n - A), \qquad \mathbb{A}_n^* = n^{1/2}\bigl(A_n^* - A\bigr).$$

Under the conditions of the theorem, the joint limiting distribution of $\Theta_n$ and $\mathbb{A}_n$ is Gaussian so without loss of generality, the former may be treated as a component of the latter. This is done below both for $\Theta_n$ and $\Theta_n^*$.

Define $\mathcal{P}_n^*$ by

$$\frac{\mathrm{d}\mathcal{P}_n^*}{\mathrm{d}\mathcal{P}_n} = \exp(\ell_n) = \prod_{i=1}^n \frac{p_{\theta_n}(U_i^*)}{p(U_i^*)}.$$

Note that under $\mathcal{P}_n^*$, $U_1, \dots, U_n$ are mutually independent with probability measure $P$, while conditionally on the sigma-algebra $\mathcal{U}_n$ generated by $U_1, \dots, U_n$, the random vectors $U_1^*, \dots, U_n^*$ are mutually independent with probability measure $P_{\theta_n}$.

By hypothesis, $\mathbb{A}_n^*$ is independent of $U_1, \dots, U_n$ and has the same distribution as $\mathbb{A}_n$ under $\mathcal{P}_n$. Therefore, it follows from Lemma 1 that, as $n \to \infty$,

$$\left(\frac{\mathrm{d}\mathcal{P}_n^*}{\mathrm{d}\mathcal{P}_n}, \mathbb{A}_n, \mathbb{A}_n^*\right) \rightsquigarrow \bigl(\zeta, \mathbb{A}, \mathbb{A}^\perp\bigr)$$

under $\mathcal{P}_n$, where $\mathbb{A}^\perp$ is an independent copy of $\mathbb{A}$ and

$$\zeta = \exp\bigl(\Theta^\top \mathbb{W}_P^\perp - \Theta^\top I_P \Theta/2\bigr).$$

Moreover, $\mathrm{E}(\zeta) = \mathrm{E}(\zeta\,|\,\Theta) = 1$ because $\mathbb{W}_P^\perp$ is distributed as $\mathcal{N}(0, I_P)$ and independent of $\Theta$.

Invoking Le Cam's Third Lemma, one can now see that $\mathcal{P}_n^*$ is contiguous with respect to $\mathcal{P}_n$. Furthermore if $Y_n$ is an arbitrary sequence of random vectors such that $Y_n \rightsquigarrow Y$ under $\mathcal{P}_n$, as $n \to \infty$, then $Y_n \rightsquigarrow Y^\star$ under $\mathcal{P}_n^*$ also. Moreover, the limit $Y^\star$ is such that for any bounded continuous function $L$,

$$\mathrm{E}\bigl\{L\bigl(Y^\star\bigr)\bigr\} = \mathrm{E}\bigl\{\zeta L(Y)\bigr\}.$$

In particular, as $n \to \infty$, $(\mathbb{A}_n, \Theta_n) \rightsquigarrow (\mathbb{A}, \Theta)$ under $\mathcal{P}_n^*$, because $\mathbb{W}_P^\perp$ is independent of $(\mathbb{A}, \Theta)$. Accordingly,

$$\mathrm{E}\bigl\{\zeta L(\mathbb{A}, \Theta)\bigr\} = \mathrm{E}\bigl\{\mathrm{E}(\zeta\,|\,\Theta)\, L(\mathbb{A}, \Theta)\bigr\} = \mathrm{E}\bigl\{L(\mathbb{A}, \Theta)\bigr\}$$

for any bounded continuous function $L : \mathcal{D}(\mathcal{T}; \mathbb{R}^s) \times \mathbb{R}^p \to \mathbb{R}$.

Given that, as $n \to \infty$, $(\mathbb{A}_n, \mathbb{A}_n^*) \rightsquigarrow (\mathbb{A}, \mathbb{A}^\perp)$ under $\mathcal{P}_n$, a similar argument implies that $(\mathbb{A}_n, \mathbb{A}_n^*) \rightsquigarrow (\mathbb{A}, \mathbb{A}^\star)$ under $\mathcal{P}_n^*$ with

$$\mathrm{E}\bigl\{L\bigl(\mathbb{A}, \mathbb{A}^\star\bigr)\bigr\} = \mathrm{E}\bigl\{\zeta L\bigl(\mathbb{A}, \mathbb{A}^\perp\bigr)\bigr\} \tag{B.1}$$

for any bounded continuous function $L : \mathcal{D}(\mathcal{T}; \mathbb{R}^s)^{\otimes 2} \to \mathbb{R}$. Furthermore, $\mathbb{A}^\star$ is càdlàg whenever $\mathbb{A}$ is càdlàg, and it is continuous if $\mathbb{A}$ is continuous.

Next, fix $\omega_j \in \mathbb{R}^{\ell s}$ and let $s_{j1}, \dots, s_{j\ell} \in \mathcal{T}$ for $j \in \{1, 2\}$. Let also

$$Z_1 = \bigl(\mathbb{A}(s_{11})^\top, \dots, \mathbb{A}(s_{1\ell})^\top\bigr)^\top,$$

$$Z_2 = \bigl(\mathbb{A}^\perp(s_{21})^\top, \dots, \mathbb{A}^\perp(s_{2\ell})^\top\bigr)^\top,$$

$$Z^\star = \bigl(\mathbb{A}^\star(s_{21})^\top, \dots, \mathbb{A}^\star(s_{2\ell})^\top\bigr)^\top.$$

Further define $\Sigma_j = \mathrm{E}(Z_j Z_j^\top)$ for $j \in \{1, 2\}$ and put $a_2^\top = \mathrm{E}(\mathbb{W}_P^\perp Z_2^\top)$.

Exploiting identity (B.1), multivariate normality and the independence between $(\Theta, \mathbb{A})$ and $(\mathbb{W}_P^\perp, \mathbb{A}^\perp)$, one finds

$$\begin{aligned}
\mathrm{E}\bigl\{\exp\bigl(i\omega_1^\top Z_1 + i\omega_2^\top Z^\star\bigr)\bigr\} &= \mathrm{E}\bigl\{\zeta \exp\bigl(i\omega_1^\top Z_1 + i\omega_2^\top Z_2\bigr)\bigr\} \\
&= \mathrm{E}\bigl\{\exp\bigl(i\omega_1^\top Z_1 + i\omega_2^\top Z_2 + \Theta^\top \mathbb{W}_P^\perp - \Theta^\top I_P \Theta/2\bigr)\bigr\} \\
&= \mathrm{E}\bigl\{\exp\bigl(i\omega_1^\top Z_1 - \omega_2^\top \Sigma_2 \omega_2/2 + i\omega_2^\top a_2 \Theta\bigr)\bigr\},
\end{aligned}$$

where the last equality follows upon conditioning on $Z_1$ and $\Theta$. Similarly,

$$\mathrm{E}\bigl\{\exp\bigl(i\omega_1^\top Z_1 - \omega_2^\top \Sigma_2 \omega_2/2 + i\omega_2^\top a_2 \Theta\bigr)\bigr\} = \mathrm{E}\bigl[\exp\bigl\{i\omega_1^\top Z_1 + i\omega_2^\top (Z_2 + a_2\Theta)\bigr\}\bigr].$$

Consequently, $\mathbb{A}^\star$ is a centered Gaussian process. Furthermore, the finite-dimensional distributions of $(\mathbb{A}, \mathbb{A}^\star)$ agree with those of $(\mathbb{A}, \mathbb{A}^\perp + a\Theta)$, where $a(t) = \mathrm{E}\{\mathbb{A}(t)\mathbb{W}_P^\top\}$ for every $t \in \mathcal{T}$. As a result, the processes $(\mathbb{A}, \mathbb{A}^\star)$ and $(\mathbb{A}, \mathbb{A}^\perp + a\Theta)$ are identically distributed, as claimed.

To establish the second assertion, it suffices to remark that together with the above result, condition (10) implies

$$\begin{aligned}
\mathbb{G}_n^A &= n^{1/2}(A_n - A_{\theta_n}) = \mathbb{A}_n - \dot{A}\Theta_n + \mathrm{o}_P(1), \\
\mathbb{G}_n^{A^*} &= n^{1/2}\bigl(A_n^* - A_{\theta_n^*}\bigr) = \mathbb{A}_n^* - \dot{A}\Theta_n^* + \mathrm{o}_P(1).
\end{aligned}$$

If the sequence $(A_n, \theta_n)$ is also regular for $\mathcal{A} \times \mathcal{O}$, it follows that, as $n \to \infty$, $(\mathbb{G}_n^A, \mathbb{G}_n^{A^*}) \rightsquigarrow (\mathbb{A} - \dot{A}\Theta, \mathbb{A}^\star - \dot{A}\Theta^\star)$. Now in this case, $a = \dot{A}$ for the process $\mathbb{A}_n$ while $a = \mathrm{E}(\Theta \mathbb{W}_P^\top) = I$ for the process $\Theta_n$, which was assimilated into $\mathbb{A}_n$. Therefore,

$$\mathbb{A}^\star - \dot{A}\Theta^\star = \mathbb{A}^\perp + \dot{A}\Theta - \dot{A}\bigl(\Theta^\perp + \Theta\bigr) = \mathbb{A}^\perp - \dot{A}\Theta^\perp$$

is an independent copy of $\mathbb{A} - \dot{A}\Theta$, which completes the proof.

## Appendix C. Proof of Theorem 2

The proof is similar to that of Theorem 1 but based on Lemma 2. Thus, fix $\theta_0 \in \mathcal{O}$ and consider at first two sets of mutually independent random vectors $U_1, \dots, U_n$, $U_1^*, \dots, U_n^*$ and $V_1, \dots, V_m$, $V_1^*, \dots, V_m^*$, $V_1^{**}, \dots, V_m^{**}$ from probability measures $P = P_{\theta_0}$ and $Q = Q_{\theta_0}$, respectively.

Denote by $\mathcal{P}_n$ the joint probability measure of these $2n + 3m$ random vectors. Given estimators $\theta_n = T_n(U_1, \dots, U_n)$ and $\theta_n^* = T_n(U_1^*, \dots, U_n^*)$ of $\theta$, define another probability measure $\mathcal{P}_n^*$ by

$$\frac{\mathrm{d}\mathcal{P}_n^*}{\mathrm{d}\mathcal{P}_n} = \left\{\prod_{i=1}^n \frac{p_{\theta_n}(U_i^*)}{p(U_i^*)}\right\} \times \left\{\prod_{i=1}^m \frac{q_{\theta_n}(V_i^*)}{q(V_i^*)} \times \frac{q_{\theta_n^*}(V_i^{**})}{q(V_i^{**})}\right\}.$$

Note that under the conditions of the theorem, the $2n + 3m$ random vectors have the same distribution as in Lemma 2 under $\mathcal{P}_n^*$. Assuming without loss of generality that $\theta_n$ is a component of $A_n$, it thus follows that, as $n \to \infty$, the vector of processes

$$\left(\frac{\mathrm{d}\mathcal{P}_n^*}{\mathrm{d}\mathcal{P}_n}, \mathbb{A}_n, \mathbb{A}_n^*, \check{\mathbb{A}}_n, \check{\mathbb{A}}_n^*, \check{\mathbb{A}}_n^{**}, \mathbb{W}_{P,n}, \mathbb{W}_{P,n}^*, \mathbb{W}_{Q,n}, \mathbb{W}_{Q,n}^*, \mathbb{W}_{Q,n}^{**}\right)$$

converges weakly in $\mathbb{R} \times \mathcal{D}(\mathcal{T}; \mathbb{R}^s)^{\otimes 5} \times \mathbb{R}^{p \otimes 5}$, under $\mathcal{P}_n$, to a limit of the form

$$\bigl(\check{\zeta}, \mathbb{A}, \mathbb{A}^\perp, \check{\mathbb{A}}, \check{\mathbb{A}}^\perp, \check{\mathbb{A}}^{\perp\perp}, \mathbb{W}_P, \mathbb{W}_P^\perp, \mathbb{W}_Q, \mathbb{W}_Q^\perp, \mathbb{W}_Q^{\perp\perp}\bigr).$$

In this limit, $(\mathbb{A}^\perp, \mathbb{W}_P^\perp)$ is an independent copy of $(\mathbb{A}, \mathbb{W}_P)$, while $(\check{\mathbb{A}}^\perp, \mathbb{W}_Q^\perp)$ and $(\check{\mathbb{A}}^{\perp\perp}, \mathbb{W}_Q^{\perp\perp})$ are independent copies of $(\check{\mathbb{A}}, \mathbb{W}_Q)$. Furthermore, $\mathbb{A}$, $\mathbb{A}^\perp$, $\mathbb{W}_P$ and $\mathbb{W}_P^\perp$ are mutually independent of $\check{\mathbb{A}}$, $\check{\mathbb{A}}^\perp$, $\check{\mathbb{A}}^{\perp\perp}$, $\mathbb{W}_Q$, $\mathbb{W}_Q^\perp$, $\mathbb{W}_Q^{\perp\perp}$. Finally,

$$\check{\zeta} = \exp\bigl\{\Theta^\top \mathbb{W}_P^\perp + \Theta^\top \mathbb{W}_Q^\perp + \bigl(\Theta^\perp\bigr)^\top \mathbb{W}_Q^{\perp\perp} - \Theta^\top I_P \Theta/2 - \Theta^\top I_Q \Theta/2 - \bigl(\Theta^\perp\bigr)^\top I_Q \Theta^\perp/2\bigr\},$$

and it can be checked easily that $\mathrm{E}(\check{\zeta}) = 1$.

It now follows from Le Cam's Third Lemma that $\mathcal{P}_n^*$ is contiguous with respect to $\mathcal{P}_n$, and if $Y_n$ is an arbitrary sequence of random variables such that $Y_n \rightsquigarrow Y$ under $\mathcal{P}_n$, as $n \to \infty$, then $Y_n \rightsquigarrow Y^\star$ under $\mathcal{P}_n^*$ also. Moreover, the limit $Y^\star$ is such that for any bounded continuous function $L$,

$$\mathrm{E}\{L(Y^\star)\} = \mathrm{E}\{\check{\zeta} L(Y)\}.$$

Next, proceeding as in the proof of Theorem 1, one can see (separating $\theta_n$ from $A_n$ for added clarity) that

$$\mathrm{E}\{L(\mathbb{A}, \mathbb{A}^\star, \check{\mathbb{A}}, \check{\mathbb{A}}^\star, \check{\mathbb{A}}^{\star\star}, \Theta, \Theta^\star)\} = \mathrm{E}\{\check{\zeta} L(\mathbb{A}, \mathbb{A}^\perp, \check{\mathbb{A}}, \check{\mathbb{A}}^\perp, \check{\mathbb{A}}^{\perp\perp}, \Theta, \Theta^\perp)\},$$

where for $a$, $\check{a}$ and $\Gamma$ given in the statement of the theorem,

$$\Theta^\star = \Theta^\perp + \Gamma\Theta, \qquad \mathbb{A}^\star = \mathbb{A}^\perp + a\Theta, \qquad \check{\mathbb{A}}^\star = \check{\mathbb{A}}^\perp + \check{a}\Theta, \qquad \check{\mathbb{A}}^{\star\star} = \check{\mathbb{A}}^{\perp\perp} + \check{a}\Theta^\star.$$

To verify this assertion, fix $\omega_1, \omega_2, \omega_3, \omega_4, \omega_5 \in \mathbb{R}^{\ell m}$ and for $j \in \{1, 2, 3\}$, let $s_{j1}, \ldots, s_{j\ell}, t_j = t_{j1}, \ldots, t_{j\ell} \in \mathcal{T}$. Next, set

$$\begin{aligned}
Z_1 &= \left(\mathbb{A}(s_{11})^\top, \ldots, \mathbb{A}(s_{1\ell})^\top\right)^\top, & Z_2 &= \left(\check{\mathbb{A}}(t_{11})^\top, \ldots, \check{\mathbb{A}}(t_{1\ell})^\top\right)^\top, \\
Z_3 &= \left(\mathbb{A}^\perp(s_{21})^\top, \ldots, \mathbb{A}^\perp(s_{2\ell})^\top\right)^\top & Z_4 &= \left(\check{\mathbb{A}}^\perp(t_{21})^\top, \ldots, \check{\mathbb{A}}^\perp(t_{2\ell})^\top\right)^\top, \\
Z_5 &= \left(\check{\mathbb{A}}^{\perp\perp}(t_{31})^\top, \ldots, \check{\mathbb{A}}^{\perp\perp}(t_{3\ell})^\top\right)^\top, & Z^\star &= \left(\mathbb{A}^\star(s_{21})^\top, \ldots, \mathbb{A}^\star(s_{2\ell})^\top\right)^\top, \\
\check{Z}^\star &= \left(\check{\mathbb{A}}^\star(t_{21})^\top, \ldots, \check{\mathbb{A}}^\star(t_{2\ell})^\top\right)^\top, & \check{Z}^{\star\star} &= \left(\check{\mathbb{A}}^{\star\star}(t_{31})^\top, \ldots, \check{\mathbb{A}}^{\star\star}(t_{3\ell})^\top\right)^\top,
\end{aligned}$$

and let $\Sigma_j = \mathrm{E}(Z_j Z_j^\top)$ for $j \in \{1, \ldots, 5\}$. Further set

$$a_3^\top = \mathrm{E}(\mathbb{W}_P^\perp Z_3^\top), \qquad a_4^\top = \mathrm{E}(\mathbb{W}_Q^\perp Z_4^\top), \qquad a_5^\top = \mathrm{E}(\mathbb{W}_Q^{\perp\perp} Z_5^\top)$$

and

$$\begin{aligned}
\check{\zeta}_1 &= \exp\left(\Theta^\top \mathbb{W}_P^\perp - \Theta^\top I_P \Theta/2\right), \\
\check{\zeta}_2 &= \exp\left(\Theta^\top \mathbb{W}_P^\perp - \Theta^\top I_P \Theta/2 + \Theta^\top \mathbb{W}_Q^\perp - \Theta^\top I_Q \Theta/2\right).
\end{aligned}$$

It is then possible to develop

$$\Omega = \mathrm{E}\{\exp(i\omega_1^\top Z_1 + i\omega_2^\top Z^\star + i\omega_3^\top Z_3 + i\omega_4^\top \check{Z}^\star + i\omega_5^\top \check{Z}^{\star\star})\} = \mathrm{E}\left\{\check{\zeta} \exp\left(i \sum_{j=1}^{5} \omega_j^\top Z_j\right)\right\}$$

as follows, exploiting multivariate normality and independence as appropriate:

$$\begin{aligned}
\Omega &= \mathrm{E}\left\{\check{\zeta}_2 \exp\left(i \sum_{j=1}^{4} \omega_j^\top Z_j\right) \exp\left(-\omega_5^\top \Sigma_5 \omega_5/2 + i\omega_5^\top a_5 \Theta^\perp\right)\right\} \\
&= \mathrm{E}\left\{\check{\zeta}_1 \exp\left(i \sum_{j=1}^{3} \omega_j^\top Z_j\right) \exp\left(-\omega_5^\top \Sigma_5 \omega_5/2 + i\omega_5^\top a_5 \Theta^\perp\right) \exp\left(-\omega_4^\top \Sigma_4 \omega_4/2 + i\omega_4^\top a_4 \Theta\right)\right\} \\
&= \mathrm{E}\left\{\left(i \sum_{j=1}^{2} \omega_j^\top Z_j\right) \exp\left(-\omega_5^\top \Sigma_5 \omega_5/2 - \omega_5^\top a_5 \Lambda a_5^\top \omega_5/2 + i\omega_5^\top \Gamma a_5 \Theta\right)\right. \\
&\qquad \left. \times \exp\left(-\omega_4^\top \Sigma_4 \omega_4/2 + i\omega_4^\top a_4 \Theta - \omega_3^\top \Sigma_3 \omega_3/2 + i\omega_3^\top a_3 \Theta + \omega_5^\top a_5 \Xi \omega_3\right)\right\}.
\end{aligned}$$

In the last expression, $\Gamma = \mathrm{E}(\Theta \mathbb{W}_P^\top)$ and $\Lambda = \mathrm{E}(\Theta\Theta^\top)$ are as defined in Lemma 2, while $\Xi = \mathrm{E}(\Theta^\perp Z_3^\top)$. The first part of Theorem 2 is thus proved, because

$$\Omega = \mathrm{E}\bigl[\exp\bigl[i\omega_1^\top Z_1 + i\omega_2^\top Z_2 + i\omega_3^\top(Z_3 + a_3\Theta) + i\omega_4^\top(Z_4 + a_4\Theta) + i\omega_5^\top\{Z_5 + a_5\bigl(\Theta^\perp + \Gamma\Theta\bigr)\}\bigr]\bigr].$$

To establish the second claim, one can proceed along the same lines as in the proof of the second part of Theorem 1. Given that $(A_n, \theta_n)$ is $\mathcal{P}$-regular for $\mathcal{A} \times \mathcal{O}$ and $\check{A}_n$ is $\mathcal{Q}$-regular for $\mathcal{A}$, one has $a = \check{a} = \dot{A}$ and $\Gamma = I$. It follows that as $n \to \infty$, the pair $(\mathbb{G}_n^{\check{A}^*}, \mathbb{G}_n^{\check{A}^{**}})$ converges weakly in $\mathcal{D}(\mathcal{T}; \mathbb{R}^s)^{\otimes 2}$ to $(\mathbb{G}^{\check{A}^\star}, \mathbb{G}^{\check{A}^{\star\star}})$, where $\mathbb{G}^{\check{A}^\star} = \mathbb{A} - \mathbb{B}^\star = \mathbb{A} - \check{\mathbb{A}}^\perp - \dot{A}\Theta$ and

$$\mathbb{G}^{\check{A}^{\star\star}} = \mathbb{A}^\star - \mathbb{B}^{\star\star} = \mathbb{A}^\perp + \dot{A}\Theta - \check{\mathbb{A}}^{\perp\perp} - \dot{A}\bigl(\Theta^\perp + \Theta\bigr) = \mathbb{A}^\perp - \check{\mathbb{A}}^{\perp\perp} - \dot{A}\Theta^\perp.$$

As the latter is clearly an independent copy of $\mathbb{G}^{\check{A}^\star}$, the proof is complete.

## Appendix D. Proof of Proposition 4

Set $J = J_{\theta_0}$ and let $J'(u)$ be the $p \times d$ matrix of partial derivatives of $J(u)$ with respect to the components of $u = (u_1, \ldots, u_d) \in (0, 1)^d$. It is then easy to check that

$$\begin{aligned}\Theta_n &= n^{-1/2}\sum_{i=1}^n J(U_i) + n^{-1/2}\sum_{i=1}^n\bigl[J\bigl\{H_n(U_i)\bigr\} - J(U_i)\bigr] + \mathrm{o}_P(1)\\ &= n^{-1/2}\sum_{i=1}^n J(U_i) + \frac{1}{n}\sum_{i=1}^n J'(U_i)\mathbb{H}_n(U_i) + \mathrm{o}_P(1).\end{aligned}$$

It follows from results in Section 3.2 of [19] that if

$$\Theta_n^\dagger = n^{-1/2}\sum_{i=1}^n J(U_i),$$

then, as $n \to \infty$,

$$\bigl(\mathbb{H}_n, \Theta_n^\dagger, \Theta_n\bigr) \rightsquigarrow \bigl(\mathbb{H}, \Theta^\dagger, \Theta\bigr)$$

in $\mathcal{D}([0, 1]^d; \mathbb{R}^d) \times \mathbb{R}^{p\otimes 2}$, where the weak limit is a continuous centered Gaussian process in which

$$\Theta = \Theta^\dagger + \int J'(u)\mathbb{H}(u)c(u)\,\mathrm{d}u.$$

Under these conditions, it follows that, as $n \to \infty$,

$$\bigl(\mathbb{C}_n, \Theta_n, \mathbb{W}_{C,n}\bigr) \rightsquigarrow (\mathbb{C}, \Theta, \mathbb{W}_C)$$

in $\mathcal{D}([0, 1]^d; \mathbb{R}^d) \times \mathbb{R}^{p\otimes 2}$, where the limit is a continuous, centered Gaussian process. Moreover, the sequence $\theta_n$ is $\mathcal{P}$-regular for $\mathcal{O}$. For, under $P$,

$$\mathrm{E}\bigl(\Theta\mathbb{W}_C^\top\bigr) = \mathrm{E}\bigl(\Theta^\dagger\mathbb{W}_C^\top\bigr) + \int J'(u)\mathrm{E}\bigl\{\mathbb{H}(u)\mathbb{W}_C^\top\bigr\}c(u)\,\mathrm{d}u = \int J(u)\dot{c}(u)\,\mathrm{d}u = I$$

in view of (24) and the fact that $\mathrm{E}\{\mathbb{H}(u)\mathbb{W}_C^\top\} = 0$ for all $u \in [0, 1]^d$.

## Appendix E. Proof of Proposition 5

Set $J = J_{\theta_0}$ and $J'$ as in Appendix D. One can then see that

$$\Theta_n = n^{-1/2} \sum_{i=1}^{n} J\{C(U_i)\} + \frac{1}{n} \sum_{i=1}^{n} J'\{C(U_i)\}\mathbb{B}_n(U_i) + \mathrm{o}_P(1).$$

Hence if

$$\Theta_n^{\ddagger} = n^{-1/2} \sum_{i=1}^{n} J\{C(U_i)\},$$

results in Section 3.2 of [18] imply that, as $n \to \infty$,

$$(\mathbb{B}_n, \Theta_n^{\ddagger}, \Theta_n) \rightsquigarrow (\mathbb{B}, \Theta^{\ddagger}, \Theta)$$

in $\mathcal{D}([0, 1]^d; \mathbb{R}) \times \mathbb{R}^{p\otimes 2}$, where the weak limit is a continuous centered Gaussian process with

$$\Theta = \Theta^{\ddagger} + \int J'\{C(u)\}\mathbb{B}(u)c(u)\,\mathrm{d}u.$$

Moreover, the sequence $\theta_n$ is $\mathcal{P}$-regular for $\mathcal{O}$. For, under $P$,

$$\begin{aligned}
\mathrm{E}(\Theta \mathbb{W}_C^{\top}) &= \mathrm{E}(\Theta^{\ddagger}\mathbb{W}_C^{\top}) + \int J'\{C(u)\}\mathrm{E}\{\mathbb{B}(u)\mathbb{W}_C^{\top}\}c(u)\,\mathrm{d}u \\
&= \int J\{C(u)\}\dot{c}(u)\,\mathrm{d}u + \int J'\{C(u)\}\dot{C}(u)c(u)\,\mathrm{d}u = I,
\end{aligned}$$

in view of (25) and the fact that $B_n$ is $\mathcal{P}$-regular for $\mathcal{C}$.

## Appendix F. Smoothness conditions for the existence of Kendall's process

***Condition I.*** *For all $\theta \in \mathcal{O}$, the distribution function $K_\theta$ of $C(U)$ admits a density $k_\theta$ which is continuous on $\mathcal{O} \times (0, 1]$ and such that $k_\theta(w) = \mathrm{o}\{w^{-1/2}\log^{-1/2-\varepsilon}(1/w)\}$ for some $\varepsilon > 0$, as $w \to 0$.*

***Condition II.*** *For all $\theta \in \mathcal{O}$, there exists a version of the conditional distribution of the vector $U$ given $C(U) = w$ such that, for any continuous real-valued function $g$ on $[0, 1]^d$, the mapping $w \mapsto \mu(w, g) = k_\theta(w)\mathrm{E}\{g(U)|C(U) = w\}$ is continuous on $(0, 1]$ with $\mu(1, g) = k(\theta, 1)\, g(1, \dots, 1)$.*

## Appendix G. Proof of Proposition 6

To show that the sequence $K_n$ is $\mathcal{P}$-regular for $\mathcal{K}$, it remains to see that $E(\mathbb{K}\mathbb{W}_C^{\top}) = \dot{K}$. To this end, first observe that together with Conditions I and II in Appendix F, the smoothness assumptions on $k_\theta$ imply

$$\dot{K}(w) = \int_0^w \dot{k}(t)\,\mathrm{d}t$$

for all $w \in [0, 1]$ and

$$\dot{K}(1) = \left[\frac{\partial}{\partial\theta}\int_0^1 k_\theta(w)\,\mathrm{d}w\right]_{\theta=\theta_0} = 0.$$

Similarly, the conditions on $c_\theta$ are such that $\int \dot{c}(u)\,\mathrm{d}u = 0$.

Write $a(w) = \mathrm{E}\{\mathbb{K}(w)\mathbb{W}_C^\top\}$ for all $w \in [0,1]$. To show that $a = \dot{K}$, let $\ell : [0,1] \to \mathbb{R}$ be an arbitrary continuous function and write

$$L(w) = \int_0^w \ell(t)\,\mathrm{d}t$$

for arbitrary $w \in [0,1]$. Interchanging the order of integration, one finds

$$\int_0^1 \dot{K}(w)\ell(w)\,\mathrm{d}w = \int_0^1 \int_0^w \dot{k}(t)\ell(w)\,\mathrm{d}t\,\mathrm{d}w = -\int_0^1 \dot{k}(t)L(t)\,\mathrm{d}t.$$

Using the fact that $\mathrm{E}\{\mathbb{B}(u)\mathbb{W}_C^\top\} = \dot{C}(u)$ for all $u \in [0,1]^d$, one then gets

$$\begin{aligned}\int_0^1 \dot{K}(w)\ell(w)\,\mathrm{d}w &= -\left[\frac{\partial}{\partial\theta}\int_0^1 k_\theta(w)L(w)\,\mathrm{d}w\right]_{\theta=\theta_0}\\ &= -\left[\frac{\partial}{\partial\theta}\int c_\theta(u)L\big\{C_\theta(u)\big\}\,\mathrm{d}u\right]_{\theta=\theta_0}\\ &= -\int \dot{c}(u)L\big\{C(u)\big\}\,\mathrm{d}u - \int c(u)\ell\big\{C(u)\big\}\dot{C}(u)\,\mathrm{d}u. \qquad \text{(G.1)}\end{aligned}$$

Similarly,

$$\begin{aligned}\int_0^1 \mathrm{E}\big\{\mu(w,\mathbb{B})\mathbb{W}_C^\top\big\}\ell(w)\,\mathrm{d}w &= \int_0^1 \mu(w,\dot{C})\ell(w)\,\mathrm{d}w\\ &= \int_0^1 k(w)\mathrm{E}\big\{\dot{C}(U)|C(U)=w\big\}\ell(w)\,\mathrm{d}w\\ &= \int c(u)\ell\big\{C(u)\big\}\dot{C}(u)\,\mathrm{d}u. \qquad \text{(G.2)}\end{aligned}$$

Finally,

$$\begin{aligned}\int_0^1 \mathrm{E}\big\{\alpha(w)\mathbb{W}_C^\top\big\}\ell(w)\,\mathrm{d}w &= \int_0^1 \int \ell(w)\dot{c}(u)\mathbf{1}\big\{C(u) \le w\big\}\,\mathrm{d}u\,\mathrm{d}w\\ &= \int \dot{c}(u)\big[L(1) - L\big\{C(u)\big\}\big]\,\mathrm{d}u\\ &= -\int \dot{c}(u)L\big\{C(u)\big\}\,\mathrm{d}u\\ &= \int_0^1 \dot{K}(w)\ell(w)\,\mathrm{d}w + \int_0^1 \mathrm{E}\big\{\mu(w,\mathbb{B})\mathbb{W}_C^\top\big\}\ell(w)\,\mathrm{d}w. \qquad \text{(G.3)}\end{aligned}$$

Upon substitution of (G.1) and (G.2) into (G.3), one finds

$$\int_0^1 \big\{a(w) - \dot{K}(w)\big\}\ell(w)\,\mathrm{d}w = 0.$$

As the choice of $\ell$ is arbitrary, one may conclude.

Note that as a by-product of the proof, one finds $\dot{K}(w) = \mathrm{E}\{\alpha(w)\mathbb{W}_C^\top\} - \mu(w,\dot{C})$ for all $w \in [0,1]$.

## Appendix H. Proof of Proposition 9

To show that the sequence $D_n$ is $\mathcal{P}$-regular for $\mathcal{K}$, it remains to check that $E(\mathbb{D}\mathbb{W}_F^\top) = \dot{K}$. To this end, write $a(w) = \mathrm{E}\{\mathbb{D}(w)\mathbb{W}_F^\top\}$ for all $w \in [0,1]$ and let $\ell : [0,1] \to \mathbb{R}$ be an arbitrary continuous function. Let also $L(w)$ denote its integral on the interval $[0, w]$, as in Appendix G.

Write $a(w) = \mathrm{E}\{\mathbb{D}(w)\mathbb{W}_F^\top\}$ for every $w \in [0,1]$. Proceeding as in the proof of Proposition 6, one finds

$$\begin{aligned}
\int_0^1 \dot{K}(w)\ell(w)\,\mathrm{d}w &= -\left[\frac{\partial}{\partial\theta}\int_0^1 k_\theta(w)L(w)\,\mathrm{d}w\right]_{\theta=\theta_0} \\
&= -\left[\frac{\partial}{\partial\theta}\int f_\theta(x)L\big\{F_\theta(x)\big\}\,\mathrm{d}x\right]_{\theta=\theta_0} \\
&= -\int \dot{f}(x)L\big\{F(x)\big\}\,\mathrm{d}x - \int f(x)\ell\big\{F(x)\big\}\dot{F}(x)\,\mathrm{d}x. \qquad \text{(H.1)}
\end{aligned}$$

Similarly,

$$\begin{aligned}
\int_0^1 \mathrm{E}\big\{\kappa(w,\dot{F})\Theta\mathbb{W}_F^\top\big\}\ell(w)\,\mathrm{d}w &= \int_0^1 \kappa(w,\dot{F})\ell(w)\,\mathrm{d}w \\
&= \int_0^1 k(w)\mathrm{E}\big\{\dot{F}(X)|F(X)=w\big\}\ell(w)\,\mathrm{d}w \\
&= \int f(x)\ell\big\{F(x)\big\}\dot{F}(x)\,\mathrm{d}x \qquad \text{(H.2)}
\end{aligned}$$

and

$$\begin{aligned}
\int_0^1 \mathrm{E}\big\{\alpha(w)\mathbb{W}_F^\top\big\}\ell(w)\,\mathrm{d}w &= \int_0^1\int \ell(w)\dot{f}(x)\mathbf{1}\big\{F(x)\le w\big\}\,\mathrm{d}x\,\mathrm{d}w \\
&= \int \dot{f}(x)\big[L(1) - L\big\{F(x)\big\}\big]\,\mathrm{d}x \\
&= -\int \dot{f}(x)L\big\{F(x)\big\}\,\mathrm{d}x. \qquad \text{(H.3)}
\end{aligned}$$

Upon substitution of (H.1) and (H.2) into (H.3), one finds

$$\int_0^1 \big\{a(w) - \dot{K}(w)\big\}\ell(w)\,\mathrm{d}w = 0.$$

As the choice of $\ell$ is arbitrary, one may conclude.

## Acknowledgments

Funding in support of this work was provided by the Natural Sciences and Engineering Research Council of Canada, the Fonds québécois de la recherche sur la nature et les technologies, and the Institut de finance mathématique de Montréal.

## References

[1] P. Barbe, C. Genest, K. Ghoudi and B. Rémillard. On Kendall's process. *J. Multivariate Anal.* **58** (1996) 197–229. MR1405589

[2] R. Beran. Minimum distance procedures. In *Nonparametric Methods* 741–754. *Handbook of Statistics* **4**. North-Holland, Amsterdam, 1984. MR0831734

[3] R. Beran and P. W. Millar. A stochastic minimum distance test for multivariate parametric models. *Ann. Statist.* **17** (1989) 125–140. MR0981440
[4] P. J. Bickel and J.-J. Ren. The bootstrap in hypothesis testing. In *State of the Art in Probability and Statistics (Leiden, 1999)* 91–112. *IMS Lecture Notes Monogr. Ser.* **36**. Inst. Math. Statist., Beachwood, OH, 2001. MR1836556
[5] P. J. Bickel and M. J. Wichura. Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* **42** (1971) 1656–1670. MR0383482
[6] W. Breymann, A. Dias and P. Embrechts. Dependence structures for multivariate high-frequency data in finance. In *Selected Proceedings from Quantitative Methods in Finance, 2002 (Cairns/Sydney)* **3** 1–14, 2003. MR1972372
[7] S. Demarta and A. J. McNeil. The *t* copula and related copulas. *Internat. Statist. Rev.* **73** (2005) 111–129.
[8] J. Dobrić and F. Schmid. A goodness of fit test for copulas based on Rosenblatt's transformation. *Comput. Statist. Data Anal.* **51** (2007) 4633–4642.
[9] J. Durbin. Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.* **1** (1973) 279–290. MR0359131
[10] J.-D. Fermanian. Goodness-of-fit tests for copulas. *J. Multivariate Anal.* **95** (2005) 119–152. MR2164126
[11] J.-D. Fermanian, D. Radulović and M. H. Wegkamp. Weak convergence of empirical copula processes. *Bernoulli* **10** (2004) 847–860. MR2093613
[12] P. Gänßler and W. Stute. *Seminar on Empirical Processes*. Birkhäuser Verlag, Basel, 1987. MR0902803
[13] C. Genest, K. Ghoudi and L.-P. Rivest. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82** (1995) 543–552. MR1366280
[14] C. Genest, J.-F. Quessy and B. Rémillard. Tests of serial independence based on Kendall's process. *Canad. J. Statist.* **30** (2002) 441–461. MR1944373
[15] C. Genest, J.-F. Quessy and B. Rémillard. Goodness-of-fit procedures for copula models based on the probability integral transformation. *Scand. J. Statist.* **33** (2006) 337–366. MR2279646
[16] C. Genest, B. Rémillard and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance Math. Econom.* **43** (2008). In press.
[17] C. Genest and L.-P. Rivest. Statistical inference procedures for bivariate Archimedean copulas. *J. Amer. Statist. Assoc.* **88** (1993) 1034–1043. MR1242947
[18] K. Ghoudi and B. Rémillard. Empirical processes based on pseudo-observations. In *Asymptotic Methods in Probability and Statistics (Ottawa, ON, 1997)* 171–197. North-Holland, Amsterdam, 1998. MR1661480
[19] K. Ghoudi and B. Rémillard. Empirical processes based on pseudo-observations. II. The multivariate case. In *Asymptotic Methods in Stochastics* 381–406. *Fields Inst. Commun.* **44**. Amer. Math. Soc., Providence, RI, 2004. MR2106867
[20] N. Henze. Empirical-distribution-function goodness-of-fit tests for discrete models. *Canad. J. Statist.* **24** (1996) 81–93. MR1394742
[21] M. N. Jouini and R. T. Clemen. Copula models for aggregating expert opinions. *Oper. Res.* **44** (1996) 444–457.
[22] C. A. J. Klaassen and J. A. Wellner. Efficient estimation in the bivariate normal copula model: Normal margins are least favourable. *Bernoulli* **3** (1997) 55–77. MR1466545
[23] Y. Malevergne and D. Sornette. Testing the Gaussian copula hypothesis for financial assets dependences. *Quant. Finance* **3** (2003) 231–250. MR1999654
[24] D. Pollard. The minimum distance method of testing. *Metrika* **27** (1980) 43–70. MR0563412
[25] J. H. Shih and T. A. Louis. Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* **51** (1995) 1384–1399. MR1381050
[26] W. Stute, W. González-Manteiga and M. Presedo-Quindimil. Bootstrap based goodness-of-fit tests. *Metrika* **40** (1993) 243–256. MR1235086
[27] H. Tsukahara. Semiparametric estimation in copula models. *Canad. J. Statist.* **33** (2005) 357–375. MR2193980
[28] A. W. van der Vaart and J. A. Wellner. *Weak Convergence and Empirical Processes*. Springer, New York, 1996. MR1385671
[29] W. Wang and M. T. Wells. Model selection and semiparametric inference for bivariate failure-time data (with discussion). *J. Amer. Statist. Assoc.* **95** (2000) 62–76. MR1803141