Received: 23 October 2017 | Revised: 14 May 2018 | Accepted: 1 June 2018

DOI: 10.1002/bimj.201700205

RESEARCH PAPER

Biometrical Journal

Multivariate multiple test procedures based on nonparametric copula estimation

André Neumann[1] | Taras Bodnar[2] | Dietmar Pfeifer[3] | Thorsten Dickhaus[1]

[1]Institute for Statistics, University of Bremen, Bibliothekstraße 1, D-28359 Bremen, Germany

[2]Department of Mathematics, Stockholm University, Roslagsvägen 101, SE-10691 Stockholm, Sweden

[3]Institute of Mathematics, Carl von Ossietzky University of Oldenburg, D-26111 Oldenburg, Germany

Correspondence
Thorsten Dickhaus, Institute for Statistics, University of Bremen, Bibliothekstraße 1, D-28359 Bremen, Germany.
Email: dickhaus@uni-bremen.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: DI 1723/2-2 (FOR 1735)

Abstract

Multivariate multiple test procedures have received growing attention recently. This is due to the fact that data generated by modern applications typically are high-dimensional, but possess pronounced dependencies due to the technical mechanisms involved in the experiments. Hence, it is possible and often necessary to exploit these dependencies in order to achieve reasonable power. In the present paper, we express dependency structures in the most general manner, namely, by means of copula functions. One class of nonparametric copula estimators is constituted by Bernstein copulae. We extend previous statistical results regarding bivariate Bernstein copulae to the multivariate case and study their impact on multiple tests. In particular, we utilize them to derive asymptotic confidence regions for the family-wise error rate (FWER) of multiple test procedures that are empirically calibrated by making use of Bernstein copulae approximations of the dependency structure among the test statistics. This extends a similar approach by Stange et al. (2015) in the parametric case. A simulation study quantifies the gain in FWER level exhaustion and, consequently, power that can be achieved by exploiting the dependencies, in comparison with common threshold calibrations like the Bonferroni or Šidák corrections. Finally, we demonstrate an application of the proposed methodology to real-life data from insurance.

KEYWORDS

asymptotic oscillation behavior, family-wise error rate, *p*-value, risk management

1 | INTRODUCTION

Copula-based modeling of dependency structures has become a standard tool in applied multivariate statistics and quantitative risk management; see, for example, Nelsen (2006), Joe (2014), Härdle and Okhrin (2010), Embrechts, Lindskog, and McNeil (2003), and Chapter 5 of McNeil, Frey, and Embrechts (2005). The estimation of an unknown copula is key to a variety of modern multivariate statistical methods. In particular, applications of copulae to the calibration and the analysis of multiple tests have been considered by Dickhaus and Gierl (2013), Bodnar and Dickhaus (2014), Stange, Bodnar, and Dickhaus (2015), Cerqueti, Costantini, and Lupi (2012), Schmidt, Faldum, Witt, and Gerß (2014), and Schmidt, Faldum, and Gerß (2015); see also Sections 2.2.4 and 4.4 of Dickhaus (2014). Specifically, the copula-based construction of multiple testing procedures developed by Dickhaus and Gierl (2013) and Stange et al. (2015) under parametric assumptions regarding the type of dependencies among test statistics considerably extends previous approaches as in Hothorn, Bretz, and Westfall (2008) that are confined to asymptotic Gaussianity and, consequently, linear dependencies.

In the case of a parametric copula, generic estimation techniques like the (generalized) method of moments or maximum likelihood estimation are established notions; cf. Section 3.2 of Stange et al. (2015) and references therein. The empirical copula

as well as its asymptotic properties as a nonparametric estimator have been studied, among others, by Rüschendorf (1976), Deheuvels (1979), Stute (1984), and, more recently, by Bücher and Dette (2010), and Bouzebda and Zari (2013), to mention only a few references. However, similarly as multivariate histogram estimators, the empirical copula in dimension m has some undesirable properties. For example, it is discontinuous, and it typically assigns zero mass to large subsets of $[0,1]^m$, even if the sample size n is large, due to the concentration of measures phenomenon. One way to tackle these issues consists of smoothing of the empirical copula. In particular, Sancetta and Satchell (2004) proposed smoothing by Bernstein polynomials, leading to so-called Bernstein copulae. Approximation theory for Bernstein copulae has been derived by Cottin and Pfeifer (2014), and asymptotic statistical properties of Bernstein copula estimators in the bivariate case ($m = 2$) have been proven by Janssen, Swanepoel, and Veraverbeke (2012) and Belalia (2016). Functional central limit theorems for empirical copula processes have been established by Segers (2012). Applications of Bernstein copulae to modeling dependencies in non-life insurance have been considered by Diers, Eling, and Marek (2012).

In the present work, we contribute to theory and applications of Bernstein copulae in the case of a general dimension $m \geq 2$. In Section 2, we extend the asymptotic theory regarding the Bernstein copula estimator by proving its rate of convergence in infinity norm as well as its asymptotic normality in function space, for arbitrary m. Also, we provide some justifications for the proposed smoothing approach. Section 3 is then devoted to applications of Bernstein copulae for multiple testing procedures with control of the family-wise error rate (FWER), avoiding restrictive parametric dependency assumptions. The application of the central limit theorem derived in Section 2 allows for a precise quantification of the uncertainty about the realized FWER in the case that the copula of test statistics is preestimated prior to calibrating the significance thresholds of the multiple testing procedure. This extends the results of Stange et al. (2015) to the case of nonparametric copula preestimation. Section 4 demonstrates by means of a simulation study that the latter preestimation approach leads to a better exhaustion of the FWER level and thus enhances the power of the multiple testing procedure compared with traditional approaches that only take univariate marginal distributions of test statistics into account. Finally, we apply the proposed multiple testing methodology to real-life data from insurance (Section 5), and we conclude with a discussion in Section 6. Lengthy proofs and some auxiliary results are deferred to Section 7.

2 | OSCILLATION BEHAVIOR OF EMPIRICAL BERNSTEIN COPULAE

In this section, asymptotic properties of Bernstein copulas are studied. The main properties of the Bernstein estimator are consistency (Theorem 2.1) and asymptotic normality (Theorem 2.4). Some auxiliary lemmas can be found in Section 7. Nonetheless, the argumentation in this section is illustrated in some mathematical detail. More practically oriented readers might find Section 2.2 and the following sections more valuable. In Section 3, the methodology how to use this estimator in multiple testing is discussed and examples are given. The consistency of the realized FWER can be derived directly from the consistency of the Bernstein estimator. The asymptotic normality of the realized FWER follows indirectly from the asymptotic normality of the Bernstein estimator via Lemma 7.2.

Let $\mathbf{X} = (X_1, \dots, X_m)^\top$ be a random vector taking values in the probability space $(\mathcal{X}, \mathcal{F}, P)$, where $\mathcal{X} \subseteq \mathbb{R}^m$, $\mathcal{F}$ is a σ-field over $\mathcal{X}$, and P denotes the (joint) distribution of $\mathbf{X}$. The univariate marginal cumulative distribution functions (cdfs) of $\mathbf{X}$ we denote by F_j, $j = 1, \dots, m$, whereas C_X stands for the copula related to the distribution P.

Assume that $\mathbf{X}_1, \dots, \mathbf{X}_n$ are stochastically independent and identically distributed (i.i.d.) random vectors with $\mathbf{X}_1 \sim P$. Then, the marginal empirical cumulative distribution function (ecdf) $\hat{F}_{j,n}$ of $(X_{1,j}, \dots, X_{n,j})^\top$ is given by $\hat{F}_{j,n}(x_j) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{(-\infty, x_j]}(X_{i,j})$, $1 \leq j \leq m$, and the joint ecdf is defined as $\hat{H}_n(\mathbf{x}) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{(-\infty,\mathbf{x}]}(\mathbf{X}_i)$. The symbol $\mathbb{1}_{\mathcal{A}}$ denotes the indicator function of set $\mathcal{A}$ and $(-\infty, \mathbf{x}] = (-\infty, x_1] \times \ldots \times (-\infty, x_m]$. We will use an analogous bold-face notation for vectors throughout the remainder. Finally, the empirical copula $\hat{C}_{X,n}$ pertaining to $\mathbf{X}_1, \dots, \mathbf{X}_n$ is given by

$$\hat{C}_{X,n}(\mathbf{u}) = \hat{H}_n\left(\hat{\mathbf{F}}_n^{\leftarrow}(\mathbf{u})\right), \mathbf{u} \in [\mathbf{0}, \mathbf{1}],$$

with $\hat{\mathbf{F}}_n^{\leftarrow}(\mathbf{u}) = (\hat{F}_{1,n}^{\leftarrow}(u_1), ..., \hat{F}_{m,n}^{\leftarrow}(u_m))^\top$. In this, $\hat{F}_{j,n}^{\leftarrow}$ denotes the generalized inverse of the marginal ecdf in coordinate $1 \leq j \leq m$.

2.1 | Theoretical analysis

Denote the space of bounded functions on $[\mathbf{0}, \mathbf{1}]$, equipped with the supremum norm, by $(\ell^\infty([\mathbf{0}, \mathbf{1}]), \|\cdot\|_\infty)$, and the space of continuous (and bounded) functions defined on $[\mathbf{0}, \mathbf{1}]$ by $(C([\mathbf{0}, \mathbf{1}]), \|\cdot\|_\infty)$, where $\|\cdot\|_\infty$ again denotes the supremum norm.

The Bernstein copula estimation is based on the Bernstein polynomial approximation, which for a fixed copula C_X is given by the operator $B_{\mathbf{K}} : (\ell^\infty([\mathbf{0},\mathbf{1}]), \|\cdot\|_\infty) \to (C([\mathbf{0},\mathbf{1}]), \|\cdot\|_\infty)$ defined by

$$B_{\mathbf{K}}(f)(\mathbf{u}) := \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} f(\mathbf{k}/\mathbf{K}) \prod_{j=1}^{m} P_{k_j,K_j}(u_j)$$

evaluated at the function $f = C_X$, where $\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} := \sum_{k_1=0}^{K_1} \cdots \sum_{k_m=0}^{K_m}$, $\mathbf{k}/\mathbf{K} := \left(\frac{k_1}{K_1}, \ldots, \frac{k_m}{K_m}\right)^\top$,

$$P_{k,K}(u) := \binom{K}{k} u^k (1-u)^{K-k},$$

and $K_1, \ldots, K_m$ are given positive integers. The empirical Bernstein copula estimator for C_X is then given by $B_{\mathbf{K}}(\hat{C}_{X,n})$.

It is well known that continuous functions can be approximated using Bernstein polynomials. There are results on the convergence rate for continuous functions with bounded variation as well (see Chêng (1983)). For the special case of copula functions it has been proved in Corollary 3.1 of Cottin and Pfeifer (2014) that any copula function can be approximated uniformly using Bernstein polynomials.

Theorem 2.1 establishes the consistency rate of the empirical Bernstein copula estimator for any copula function C_X. This result is known for the bivariate case (see Theorem 1 in Janssen et al. (2012)).

Theorem 2.1 (**Chung–Smirnov consistency rate**). *Let m be fixed. Assume that $\sum_{j=1}^{m} K_j^{-1/2} = O(n^{-1/2}(\log\log n)^{1/2})$. Then*

$$\left\| B_{\mathbf{K}}\left(\hat{C}_{X,n}\right) - C_X \right\|_\infty = O\left(n^{-1/2}(\log\log n)^{1/2}\right) \text{ almost surely},$$

where $\|g\|_\infty := \sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} |g(\mathbf{u})|$ for $g : [\mathbf{0},\mathbf{1}] \to \mathbb{R}$.

Proof. The proof can be done analogously to the proof of the bivariate case considered in Janssen et al. (2012). By the triangle inequality we split the convergence of the empirical Bernstein copula estimator into an inner and outer convergence. It holds that

$$\left\| B_{\mathbf{K}}\left(\hat{C}_{X,n}\right) - C_X \right\|_\infty \le \left\| B_{\mathbf{K}}\left(\hat{C}_{X,n}\right) - B_{\mathbf{K}}\left(C_X\right) \right\|_\infty + \left\| B_{\mathbf{K}}\left(C_X\right) - C_X \right\|_\infty. \tag{2.1}$$

For the outer convergence, we get from Lemma 7.1 and our assumption that

$$\left\| B_{\mathbf{K}}\left(C_X\right) - C_X \right\|_\infty = O\left(n^{-1/2}(\log\log n)^{1/2}\right).$$

The argumentation for the inner convergence is more complicated. For the first summand in (2.1), we get

$$\left\| B_{\mathbf{K}}\left(\hat{C}_{X,n}\right) - B_{\mathbf{K}}\left(C_X\right) \right\|_\infty \le \sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} \left|\hat{C}_{X,n}(\mathbf{k}/\mathbf{K}) - C_X(\mathbf{k}/\mathbf{K})\right| \prod_{j=1}^{m} P_{k_j,K_j}(u_j) \le$$

$$\le \max_{\mathbf{k}\in\{\mathbf{0},\ldots,\mathbf{K}\}} \left|\hat{C}_{X,n}(\mathbf{k}/\mathbf{K}) - C_X(\mathbf{k}/\mathbf{K})\right|,$$

where $\{\mathbf{0},\ldots,\mathbf{K}\} := \{0,\ldots,K_1\} \times \ldots \times \{0,\ldots,K_m\}$. Let $\tilde{F}_{j,n}$ denote the marginal ecdf of $U_{i,j} := F_j(X_{i,j})$ for $j = 1,\ldots,m$ and $i = 1,\ldots,n$ and let $\tilde{H}_n$ stand for the ecdf of $\mathbf{U}_1,\ldots,\mathbf{U}_n$. Application of the identity (see, e.g. Section 3 of Swanepoel, 1986) $\tilde{F}_{j,n}^{\leftarrow}(u_j) = F_j(\hat{F}_{j,n}^{\leftarrow}(u_j))$ leads to $\hat{C}_{X,n}(\mathbf{k}/\mathbf{K}) = \tilde{H}_n(\tilde{\mathbf{F}}_n^{\leftarrow}(\mathbf{k}/\mathbf{K}))$ and

$$\left\| B_{\mathbf{K}}\left(\hat{C}_{X,n}\right) - B_{\mathbf{K}}\left(C_X\right) \right\|_\infty \le \max_{\mathbf{k}\in\{\mathbf{0},\ldots,\mathbf{K}\}} \left|\tilde{H}_n\left(\tilde{\mathbf{F}}_n^{\leftarrow}(\mathbf{k}/\mathbf{K})\right) - C_X(\mathbf{k}/\mathbf{K})\right| \le$$

$$\le \max_{\mathbf{k}\in\{\mathbf{0},\ldots,\mathbf{K}\}} \left|\tilde{H}_n\left(\tilde{\mathbf{F}}_n^{\leftarrow}(\mathbf{k}/\mathbf{K})\right) - C_X\left(\tilde{\mathbf{F}}_n^{\leftarrow}(\mathbf{k}/\mathbf{K})\right)\right| + \tag{2.2}$$

$$+ \sum_{j=1}^{m} \max_{k_j\in\{0,\ldots,K_j\}} \left|\tilde{F}_{j,n}^{\leftarrow}\left(\frac{k_j}{K_j}\right) - \frac{k_j}{K_j}\right|. \tag{2.3}$$

From Theorem 2 of Kiefer (1961) we get that the summand in (2.2) is of order $O(n^{-1/2}(\log\log n)^{1/2})$ as well as that each summand in (2.3) is of order $O(n^{-1/2}(\log\log n)^{1/2})$. This completes the proof. □

Remark 2.2. If m is not fixed, then the convergence rate in the last step of previous proof changes to $O(mn^{-1/2}(\log\log n)^{1/2})$. Hence, we get almost surely

$$\left\| B_{\mathbf{K}}\left(\hat{C}_{X,n}\right) - C_X \right\|_\infty = O\left(mn^{-1/2}\left(\log\log n\right)^{1/2}\right).$$

The next theorem is taken from Whitt (2002) and will be useful in order to show asymptotic normality of the Bernstein copula estimator.

Theorem 2.3 (**Generalized Continuous Mapping Theorem**). *Let g and g_n, $n \geq 1$, be measurable functions mapping (S, d) into (S', d'). Let the range (S', d') be separable. Let E be the set of x in S such that $g_n(x_n) \to g(x)$ fails for some sequence $\{x_n : n \geq 1\}$ with $x_n \to x$ in S. If $X_n \xrightarrow{d} X$, $n \to \infty$, in (S, d) ($\xrightarrow{d}$ denotes the convergence in distribution) and $\mathbb{P}[X \in E] = 0$, then $g_n(X_n) \xrightarrow{d} g(X)$, $n \to \infty$, in (S', d').*

Further, we need a result for the convergence of the empirical copula process $\mathbb{C}_n := n^{1/2}(\hat{C}_{X,n} - C_X)$. Let $\mathbf{u} \mapsto \gamma(\mathbf{u})$ be a C_X-Brownian bridge, that is, a zero mean Gaussian process with (almost surely) continuous paths and covariance function given by

$$\mathrm{Cov}\left(\gamma\left(\mathbf{u}\right), \gamma\left(\mathbf{v}\right)\right) = C_X\left(\mathbf{u} \wedge \mathbf{v}\right) - C_X\left(\mathbf{u}\right) C_X\left(\mathbf{v}\right)$$

for all $\mathbf{u}, \mathbf{v} \in [\mathbf{0}, \mathbf{1}]$. Denote $\gamma_j(u_j) := \gamma(1, \ldots, 1, u_j, 1, \ldots, 1)$. Then under some assumptions the process $\mathbb{C}(\mathbf{u}) := \gamma(\mathbf{u}) - \sum_{j=1}^m \partial_j C_X(\mathbf{u})\gamma_j(u_j)$ is the weak limit of the empirical copula process $\mathbb{C}_n$ in $(\ell^\infty([\mathbf{0}, \mathbf{1}]), \|\cdot\|_\infty)$ as shown in Proposition 3.1 of Segers (2012). With these two arguments we can prove a functional central limit theorem for the empirical Bernstein copula estimator.

Theorem 2.4 (**Asymptotic normality**). *Let m be fixed. Assume that the first order partial derivatives of C_X exist and are continuous. If $\mathbf{K} = K(n)$ is such that $n^{1/2} \sum_{j=1}^m K_j^{-1/2} \to 0$, $n \to \infty$, then it holds that*

$$n^{1/2} \cdot \left(B_{\mathbf{K}}\left(\hat{C}_{X,n}\right) - C_X\right) \xrightarrow{d} \mathbb{C} \text{ as } n \to \infty$$

in $(C([\mathbf{0}, \mathbf{1}]), \|\cdot\|_\infty)$.

Remark 2.5. The assumption of the existence and continuity of the first order partial derivatives on the boundaries can be weakened (cf. Condition 2.1 of Segers, 2012).

Proof. We split the empirical Bernstein copula process $n^{1/2} \cdot (B_{\mathbf{K}}(\hat{C}_{X,n}) - C_X)$ into two parts. We get

$$\begin{aligned} n^{1/2} \cdot \left(B_{\mathbf{K}}\left(\hat{C}_{X,n}\right) - C_X\right) &= B_{\mathbf{K}}\left(n^{1/2}\left(\hat{C}_{X,n} - C_X\right)\right) + n^{1/2}\left(B_{\mathbf{K}}\left(C_X\right) - C_X\right) = \\ &= B_{\mathbf{K}}\left(\mathbb{C}_n\right) + n^{1/2}\left(B_{\mathbf{K}}\left(C_X\right) - C_X\right). \end{aligned}$$

The second summand converges uniformly to zero because of Lemma 7.1 and our assumptions. The first summand is the empirical copula process $\mathbb{C}_n$ transformed by a family of operators $B_{\mathbf{K}}$, where $\mathbf{K} = \mathbf{K}(n)$.

We will use the Generalized Continuous Mapping Theorem 2.3. Let $(S, d) := (\ell^\infty([\mathbf{0}, \mathbf{1}]), \|\cdot\|_\infty)$ and $(S', d') := (C([\mathbf{0}, \mathbf{1}]), \|\cdot\|_\infty)$. Then (S', d') is a separable space, since the set of polynomials on $[\mathbf{0}, \mathbf{1}]$ with rational coefficients is a countable dense subset of S'. Further, let $g_n : S \to S'$ be defined by $g_n := B_{\mathbf{K}(n)}$ and $g : S \to S'$ be the identity function on S' and arbitrary on $S \backslash S'$. Notice that it does not matter, how g is defined on $S \backslash S'$, since we are interested in $g(\mathbb{C})$ and without loss of generality (w.l.o.g.) $\mathbb{C}$ takes values in S' (cf. Section 3 of Segers, 2012). Let E be the set of f in S such that $g_n(f_n) \to g(f)$ fails for some sequence $\{f_n : n \geq 1\}$ with $f_n \to f$ in S. Then $E \subseteq S \backslash S'$, since we can choose $f_n := f$ for $f \in S'$ and get uniform convergence by Bernstein's theorem (or by using Corollary 3.1 of Cottin and Pfeifer, 2014). Hence, $\mathbb{P}[\mathbb{C} \in E] \leq \mathbb{P}[\mathbb{C} \in S \backslash S'] = 0$.

The last thing we need to check is the weak convergence of the empirical copula process $\mathbb{C}_n$ to $\mathbb{C}$ in $(\ell^\infty([\mathbf{0}, \mathbf{1}]), \|\cdot\|_\infty)$. As already mentioned, Segers (2012) has shown this convergence under assumptions only regarding the first-order partial derivatives

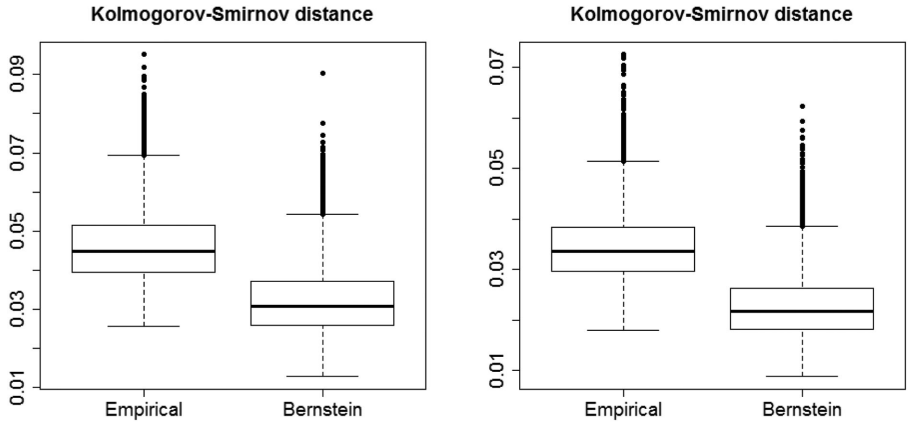

FIGURE 1 Comparison of the Bernstein copula and the empirical copula in the setting of Model 1 (left) and Model 2 (right) of Omelka et al. (2009) w.r.t. the supremum norm (Kolmogorov–Smirnov distance)

of C_X. Therefore, the proof is complete by using Proposition 3.1 of Segers (2012) and the generalized continuous mapping theorem. □

This result extends the pointwise central limit theorems of Janssen et al. (2012) and Belalia (2016) and works under weaker assumptions as well.

2.2 | The effect of smoothing

This section is meant to be an addition to the simulation study of Omelka, Gijbels, and Veraverbeke (2009). Conducting such an extensive study ourselves would go beyond the scope of this paper. Nevertheless, it is an important question how precise the Bernstein estimator is compared to other copula estimators, and this should be discussed at least to some extent.

There exists a wide variety of methods to estimate copula functions nonparametrically. Usually, the empirical copula or some sort of smoothing method is used. The Bernstein estimator studied in Section 2.1 is only one specific smoothing method among many others. Further examples comprise kernel (density) estimators (see Gijbels & Mielniczuk, 1990), and beta density estimators (see Chen, 1999). It is beyond the scope of the present work to compare all these competing approaches in detail. Generally speaking, the empirical copula is robust and universal, but it is not a copula in the strict sense, because it lacks continuity and does not have uniform margins. The Bernstein copula is a differentiable estimator, but converges rather slowly and cannot capture extreme tail dependencies (cf. Sancetta and Satchell, 2004). Recently, families of nonparametric copula estimators capable of modeling (positive) tail dependence have been studied by Pfeifer, Mändle, and Ragulina (2017). Kernel methods suffer from a boundary bias, although several modifications like the mirror approach by Schuster (1985) exist to address this problem. Beta density estimators avoid the boundary bias, but the choice of their smoothing parameter is not trivial.

Let us briefly provide some numerical justifications for smoothing of the empirical copula. In Section 3 of Omelka et al. (2009) some kernel methods have been compared in simulations under two prototypical models (Model 1 and Model 2). In Model 1, the data follow a Frank copula with parameter corresponding to Kendall's $\tau = 0.25$. In Model 2, a Clayton copula corresponding to Kendall's $\tau = 0.75$ is used.

We have applied our proposed Bernstein estimator to these models as well. Figure 1 displays the results of a simulation study under these two models. The box plots demonstrate that the estimation accuracy (measured in terms of the Kolmogorov–Smirnov distance) can be improved by smoothing. Here, we only considered smoothing by means of Bernstein polynomials, but the simulation results for various kernel methods presented by Omelka et al. (2009) are very similar. Hence, in practice it may not be most important which smoothing method to choose, while it is recommendable to smooth at all. For a more detailed overview on copula estimation methods, see Charpentier, Fermanian, and Scaillet (2007).

3 | CALIBRATION OF MULTIVARIATE MULTIPLE TEST PROCEDURES

In this section, we assume that we have uncertainty about the distribution of $\mathbf{X}$. We thus consider a statistical model of the form $(\mathcal{X}, \mathcal{F}, (P_{\vartheta, C_X} : \vartheta \in \Theta, C_X \in \mathcal{C}))$. The probability measure P_{ϑ, C_X} is indexed by two parameters. The parameter C_X denotes

the copula of $\mathbf{X}$, and $\boldsymbol{\vartheta}$ is a vector of marginal parameters that refer to $F_1, \dots, F_m$. The model for the i.i.d. sample $\mathbf{X}_1, \dots, \mathbf{X}_n$ consequently reads as $(\mathcal{X}^n, \mathcal{F}^{\otimes n}, (\mathbb{P}_{\boldsymbol{\vartheta}, C_X} : \boldsymbol{\vartheta} \in \Theta, C_X \in \mathcal{C}))$, where $\mathbb{P}_{\boldsymbol{\vartheta}, C_X} = P_{\boldsymbol{\vartheta}, C_X}^{\otimes n}$.

Based on this model, we consider multiple test problems of the form $(\mathcal{X}^n, \mathcal{F}^{\otimes n}, (\mathbb{P}_{\boldsymbol{\vartheta}, C_X} : \boldsymbol{\vartheta} \in \Theta, C_X \in \mathcal{C}), \mathcal{H})$, where $\mathcal{H} = \{H_1, \dots, H_m\}$ with $\emptyset \neq H_j \subset \Theta$ for all $1 \leq j \leq m$ denotes a family of m null hypotheses regarding the parameter $\boldsymbol{\vartheta}$. The copula C_X is not the primary target of statistical inference, but a nuisance parameter in the sense that it does not depend on $\boldsymbol{\vartheta}$. This is a common setup in multiple test theory. We will mainly consider a semi-parametric situation, where Θ is of finite dimension, while $\mathcal{C}$ is a function space.

Remark 3.1. The assumption that the number of tests equals the dimension of $\mathbf{X}$ is only made for notational convenience. The case that these two quantities differ can be treated with obvious modifications.

A multiple test for a given set of hypotheses $\mathcal{H}$ is a measurable mapping $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_m) : (\mathcal{X}^n, \mathcal{F}^{\otimes n}) \to \{0,1\}^m$, where $\varphi_j(\mathbf{x}_1, \dots, \mathbf{x}_n) = 1$ for given data $\mathbf{x}_1, \dots, \mathbf{x}_n$ means rejection of the j-th null hypothesis H_j in favor of the alternative $K_j = \Theta \setminus H_j$, $1 \leq j \leq m$. We restrict our attention to multiple tests $\boldsymbol{\varphi}$ which are such that the hypotheses are rejected if the respective test statistics are large enough for given data, that is, larger than their corresponding critical values. Notationally, this mean that

$$\varphi_j = \mathbb{1}_{(c_j, \infty)}(T_j), \quad 1 \leq j \leq m, \tag{3.1}$$

where $\mathbf{T} = (T_1, \dots, T_m)^\top : \mathcal{X}^n \to \mathbb{R}^m$ denotes a vector of real-valued test statistics that tend to larger values under alternatives, and $\mathbf{c} = (c_1, \dots, c_m)^\top$ are the critical values. In many problems of practical interest, T_j will only use the marginal data $(x_{i,j})_{1 \leq i \leq n}$, for every $1 \leq j \leq m$. For example, this typically holds true if ϑ_j only corresponds to F_j, and H_j only concerns ϑ_j, for every $1 \leq j \leq m$.

For the calibration of $\mathbf{c}$, we aim at controlling the FWER in the strong sense. Strictly speaking, our procedure will only control the FWER under the global null hypothesis in the first place. However, strong control follows directly under Assumption 3.2 (a). For sufficient conditions of this assumption see Lemma 3.3.

For given $\boldsymbol{\vartheta} \in \Theta$ and $C_X \in \mathcal{C}$, the FWER is defined as the probability for at least one false rejection (type I error) of $\boldsymbol{\varphi}$ under $\mathbb{P}_{\boldsymbol{\vartheta}, C_X}$, that is,

$$\mathrm{FWER}_{\boldsymbol{\vartheta}, C_X}(\boldsymbol{\varphi}) = \mathbb{P}_{\boldsymbol{\vartheta}, C_X}\left(\bigcup_{j \in I_0(\boldsymbol{\vartheta})} \{\varphi_j = 1\} \right),$$

where $I_0(\boldsymbol{\vartheta}) = \{1 \leq j \leq m : \boldsymbol{\vartheta} \in H_j\}$ denotes the index set of true null hypotheses under $\boldsymbol{\vartheta}$. The multiple test $\boldsymbol{\varphi}$ is said to control the FWER at level $\alpha \in [0,1]$, if

$$\sup_{\boldsymbol{\vartheta} \in \Theta, C_X \in \mathcal{C}} \mathrm{FWER}_{\boldsymbol{\vartheta}, C_X}(\boldsymbol{\varphi}) \leq \alpha.$$

Notice that, although the trueness of the null hypotheses is determined by $\boldsymbol{\vartheta}$ alone, the FWER depends on $\boldsymbol{\vartheta}$ and C_X, because the dependency structure in the data typically influences the distribution of $\boldsymbol{\varphi}$ when regarded as a statistic with values in $\{0,1\}^m$.

Throughout the remainder, we assume that the following set of conditions is fulfilled.

Assumption 3.2.

(a) Letting $H_0 = \bigcap_{j=1}^m H_j$ denote the global null hypothesis of $\mathcal{H}$, there exists a least favorable configuration (LFC) $\boldsymbol{\vartheta}^* \in H_0$ such that

$$\forall C_X \in \mathcal{C} : \forall \boldsymbol{\vartheta} \in \Theta : \mathrm{FWER}_{\boldsymbol{\vartheta}, C_X}(\boldsymbol{\varphi}) \leq \mathrm{FWER}_{\boldsymbol{\vartheta}^*, C_X}(\boldsymbol{\varphi}).$$

If this assumption is fulfilled, then weak FWER control implies strong FWER control. Notice that this assumption can be weakened by considering closed test procedures, where our proposed methodology is applied to every nonempty intersection hypothesis in $\mathcal{H}$; cf. Remark 1 of Stange, Dickhaus, Navarro, and Schunk (2016) for details. However, in such a setting, the computation time for the multiple test can increase very fast with the number of hypotheses.

(b) The vector of marginal cdfs of $\mathbf{T} = (T_1, \dots, T_m)^\top$ depends on $\boldsymbol{\vartheta}$ only, and is (at least asymptotically as $n \to \infty$) known under any LFC $\boldsymbol{\vartheta}^*$. We denote the vector of marginal cdfs of $\mathbf{T} = (T_1, \dots, T_m)^\top$ under such an LFC $\boldsymbol{\vartheta}^*$ by $\mathbf{G} = (G_1, \dots, G_m)^\top$.

(c) Letting $C_T := C_{T,\vartheta^*}$ denote the copula of $\mathbf{T}$ under ϑ^* from part (b), there exists a continuously differentiable function $h : [0,1] \to [0,1]$ such that $C_T(u,\dots,u) = h(C_X(u,\dots,u))$ for all $u \in [0,1]$, where C_X is the copula of $\mathbf{X}$. The function h may be unknown. Notice that, if T_j only uses the data $(x_{i,j})_{1\le i\le n}$, for every $1 \le j \le m$, then the copula of $\mathbf{T}$ is independent of ϑ^*. The existence of h is guaranteed whenever plateaus of $u \mapsto C_X(u,\dots,u)$ occur on the same subset of [0,1] as plateaus of $u \mapsto C_T(u,\dots,u)$. In particular, h exists if $u \mapsto C_X(u,\dots,u)$ is strictly increasing. The more crucial part of the assumption is that h needs to be continuously differentiable.

The following lemma is useful in order to verify assumption (a).

Lemma 3.3. *Let $H_j : \{\vartheta \in \Theta | \vartheta_j \in \Theta_j \subseteq \mathbb{R}\}$, $1 \le j \le m$, such that the global null hypothesis H_0 is not empty and let the marginal distributions of the data in coordinate j depend on ϑ_j only. Further, assume that every test statistic T_j only uses the data $(x_{i,j})_{1\le i\le n}$. Then for all $\vartheta \in \Theta$ and any multiple test φ which is as in* (3.1)*, we can construct a parameter value $\vartheta^* \in H_0$ with*

$$FWER_{\vartheta,C_X}(\varphi) \le FWER_{\vartheta^*,C_X}(\varphi).$$

In particular, this implies that the LFC is located in H_0.

Proof. Let w.l.o.g. $I_0(\vartheta) = \{1,\dots,m_0\}$. Choose $\vartheta^* \in H_0 \neq \emptyset$ with $\vartheta_j^* = \vartheta_j$ for $j \in I_0(\vartheta)$. Then it holds that

$$\mathbb{P}_{\vartheta,C_X}\left[\bigcup_{j\in I_0(\vartheta)}\{T_j > c_j\}\right] = \mathbb{P}_{\vartheta^*,C_X}\left[\bigcup_{j\in I_0(\vartheta)}\{T_j > c_j\}\right],$$

since it is assumed that the test statistics T_j, $j \in I_0(\vartheta)$, only utilize the data from that coordinate j. Hence,

$$\begin{aligned}
\mathrm{FWER}_{\vartheta,C_X}(\varphi) &= \mathbb{P}_{\vartheta,C_X}\left[\bigcup_{j\in I_0(\vartheta)}\{T_j > c_j\}\right] = \\
&= \mathbb{P}_{\vartheta^*,C_X}\left[\bigcup_{j\in I_0(\vartheta)}\{T_j > c_j\}\right] \le \\
&\le \mathbb{P}_{\vartheta^*,C_X}\left[\bigcup_{j=1}^{m}\{T_j > c_j\}\right] = \\
&= \mathrm{FWER}_{\vartheta^*,C_X}(\varphi).
\end{aligned}$$

□

More generally, the previous lemma holds if the test statistics satisfy the so-called subset pivotality condition (see Westfall and Young, 1993; and Dickhaus and Stange, 2013). Before we start to explain the proposed method for the calibration of $\mathbf{c}$, let us illustrate prototypical example applications of our general setup.

Example 3.4.

(a) Let $\Theta = \mathbb{R}^m$ and assume that $\vartheta_j \in \mathbb{R}$ is the expected value of X_j for every $1 \le j \le m$. The j-th null hypothesis may be the one-sided null hypothesis $H_j = \{\vartheta_j \le 0\}$ with corresponding alternative $K_j = \{\vartheta_j > 0\}$. Assume that the variance of the marginal distribution of each X_j is known and w.l.o.g. equal to one. A suitable test statistic T_j is then given by $T_j(\mathbf{X}_1,\dots,\mathbf{X}_n) = \sum_{i=1}^n X_{i,j}/\sqrt{n}$. From Lemma 3.3 it follows that the LFC lies in H_0. Since the test statistics tend to get larger with increasing values of ϑ, the LFC ϑ^* equals $\mathbf{0}$. Under ϑ^*, we have that $G_j = \Phi$ (the cdf of the standard normal law on $\mathbb{R}$) is the cdf of the (asymptotic) null distribution of T_j for every $1 \le j \le m$. If the considered copula family C consists of multivariate stable copulae (meaning that the observables follow a multivariate stable distribution), then the copula C_T is of the same type as C_X, hence all parts of Assumption 3.2 are fulfilled.

(b) Let $\mathcal{X} = [0,\infty)^m$ and assume that the stochastic representations $X_j \overset{d}{=} \vartheta_j Z_j$ with $\vartheta_j > 0$ hold true for all $1 \le j \le m$, where Z_j is a random variable taking values in [0,1]. The parameter of interest in this problem is $\vartheta = (\vartheta_1,\dots,\vartheta_m)^\top \in \Theta = (0,\infty)^m$. For each coordinate j, we consider the pair of hypotheses $H_j : \{\vartheta_j \le \vartheta_j^*\}$ versus $K_j : \{\vartheta_j > \vartheta_j^*\}$, where the LFC $\vartheta^* = (\vartheta_1^*,\dots,\vartheta_m^*)^\top \in (0,\infty)^m$ (same argumentation as in (a)) is identical to the hypothesized upper bounds for the supports (or right end-points of the distributions) of the X_j's. This has applications in the context of stress testing in actuarial science

and financial mathematics (cf., e.g., Longin, 2000). Suitable test statistics are given by the component-wise maxima of the observables, that is, $T_j(\mathbf{X}_1, \dots, \mathbf{X}_n) = \max_{1 \le i \le n} X_{i,j}/\vartheta_j^*$, $1 \le j \le m$. Assuming that the tail behavior of each X_j is known such that the marginal (limiting) extreme value distribution of T_j under $\boldsymbol{\vartheta}^*$ can be derived and letting C consist of max-stable copulae, all parts of Assumption 3.2 are fulfilled here, too.

Let us remark here that these two examples have been treated under the restrictive assumption of one-parametric copula families C by Stange et al. (2015). The following lemma is taken from Dickhaus and Gierl (2013) and connects $\mathrm{FWER}_{\boldsymbol{\vartheta}, C_{\boldsymbol{X}}}(\boldsymbol{\varphi})$ with $C_{\boldsymbol{T}}$.

Lemma 3.5. *Let Assumption 3.2 be fulfilled. Then we have that*

$$\mathrm{FWER}_{\boldsymbol{\vartheta}, C_{\boldsymbol{X}}}(\boldsymbol{\varphi}) \le 1 - C_{\boldsymbol{T}}\left(1 - \alpha_{loc}^{(1)}, \dots, 1 - \alpha_{loc}^{(m)}\right),$$

where $\alpha_{loc}^{(j)} = 1 - G_j(c_j(\alpha))$ denotes a local significance level for the j-th marginal test problem. In practice, it is convenient to carry out the multiple testing procedure in terms of p-values $p_j = 1 - G_j(T_j)$ such that $\varphi_j = \mathbb{1}_{[0, \alpha_{loc}^{(j)})}(p_j)$.

Proof. The assertion follows from Assumption 3.2 (a) and Sklar's Theorem, since it holds that

$$\begin{aligned}\mathrm{FWER}_{\boldsymbol{\vartheta}, C_{\boldsymbol{X}}}(\boldsymbol{\varphi}) &\le \mathrm{FWER}_{\boldsymbol{\vartheta}^*, C_{\boldsymbol{X}}}(\boldsymbol{\varphi}) = \\ &= 1 - C_{\boldsymbol{T}}\left(G_1\left(c_1(\alpha)\right), \dots, G_m\left(c_m(\alpha)\right)\right) = \\ &= 1 - C_{\boldsymbol{T}}\left(1 - \alpha_{loc}^{(1)}, \dots, 1 - \alpha_{loc}^{(m)}\right).\end{aligned}$$

□

Lemma 3.5 shows that the problem of calibrating the local significance levels corresponding to $\boldsymbol{c}(\alpha)$ is equivalent to the problem of estimating the contour line of $C_{\boldsymbol{T}}$ at contour level $1 - \alpha$. Any point on that contour line defines a valid set of local significance levels. Thus, one may weight the m hypotheses for importance by choosing particular points on the contour line. If all m hypotheses are equally important it is natural to choose equal local levels $\alpha_{loc}^{(j)} \equiv \alpha_{loc}$ for all $1 \le j \le m$. This amounts to finding the point of intersection of the contour line of $C_{\boldsymbol{T}}$ at contour level $1 - \alpha$ and the "main diagonal" in the m-dimensional unit hypercube. Assumption 3.2 (c) is tailored toward this strategy and should be modified accordingly if a different weighting scheme is used.

Recall that we assume that $C_{\boldsymbol{X}}$ and, consequently, $C_{\boldsymbol{T}}$ are unknown. Based on our investigations in Section 2 and making use of Assumption 3.2 (c), we thus propose to calibrate $\boldsymbol{\varphi}$ empirically. If h is known, this can be done by solving the equation

$$h\left(B_{\mathbf{K}}\left(\hat{C}_{\boldsymbol{X},n}\right)\left(1 - \alpha_{loc}, \dots, 1 - \alpha_{loc}\right)\right) = 1 - \alpha \tag{3.2}$$

for α_{loc}. Note that this assumption is formulated for equally important hypotheses and has to be modified for different situations. If for a given α the solution of (3.2) is not unique, one should choose the smallest set of local significance levels such that (3.2) holds. We denote the solution of (3.2) by $\hat{\alpha}_{loc,n}$. This leads to the representation

$$\hat{\alpha}_{loc,n} = 1 - B_{\mathbf{K}}\left(\hat{C}_{\boldsymbol{X},n}\right)^{\leftarrow}\left(h^{\leftarrow}(1 - \alpha)\right),$$

where $B_{\mathbf{K}}(\hat{C}_{\boldsymbol{X},n})^{\leftarrow}$ is the quantile of $u \mapsto B_{\mathbf{K}}(\hat{C}_{\boldsymbol{X},n})(u, \dots, u)$. Since $B_{\mathbf{K}}(\hat{C}_{\boldsymbol{X},n})$ depends on the data, $\hat{\alpha}_{loc,n}$ is a random variable and

$$\widehat{\mathrm{FWER}}_{\boldsymbol{\vartheta}^*, C_{\boldsymbol{X}}}(\boldsymbol{\varphi}) = 1 - C_{\boldsymbol{T}}\left(1 - \hat{\alpha}_{loc,n}, \dots, 1 - \hat{\alpha}_{loc,n}\right)$$

is a random variable, too, which is distributed around the target FWER level α. The following theorem is the main result of this section and quantifies the uncertainty about the realized FWER if the empirical calibration of $\boldsymbol{\varphi}$ is performed via (3.2).

Theorem 3.6. *Let Assumption 3.2 be fulfilled. Then $\widehat{\mathrm{FWER}}_{\boldsymbol{\vartheta}^*, C_{\boldsymbol{X}}}$ has the following properties.*

a) *Consistency:*

$$\forall C_{\boldsymbol{X}} \in \mathcal{C} : \widehat{\mathrm{FWER}}_{\boldsymbol{\vartheta}^*, C_{\boldsymbol{X}}}(\boldsymbol{\varphi}) \to \alpha \text{ almost surely as } n \to \infty.$$

b) *Asymptotic Normality:*

$$\forall C_X \in \mathcal{C} : \sqrt{n}\left(\widehat{\mathrm{FWER}}_{\vartheta^*,C_X}(\boldsymbol{\varphi}) - \alpha\right) \xrightarrow{d} \mathcal{N}(0,\sigma_\alpha^2) \text{ as } n \to \infty,$$

where

$$\sigma_\alpha^2 = \frac{\sigma^2\left(C_T(1-\alpha),\dots,C_T(1-\alpha)\right)}{\left(C_X'\left(C_T(1-\alpha)\right)\right)^2} \cdot \left(C_T'\left(C_T(1-\alpha)\right)\right)^2,$$

$\sigma^2(\mathbf{u}) := \mathbb{V}[\mathbb{C}(\mathbf{u})]$*, and* C_X'*,* C_T' *denotes the first derivative of the univariate functions* $u \mapsto C_X(u,\dots,u)$*,* $u \mapsto C_T(u,\dots,u)$*, respectively.*

c) *Asymptotic Confidence Region:*

$$\forall \delta \in (0,1) : \forall C_X \in \mathcal{C} : \lim_{n\to\infty} \mathbb{P}_{\vartheta^*,C_X}\left(\sqrt{n}\frac{\widehat{\mathrm{FWER}}_{\vartheta^*,C_X}(\boldsymbol{\varphi}) - \alpha}{\hat{\sigma}_n} \le z_{1-\delta}\right) = 1-\delta\,,$$

where $\hat{\sigma}_n^2 : \mathcal{X}^n \to (0,\infty)$ *is a consistent estimator of the asymptotic variance* σ_α^2*. In this,* $z_\beta = \Phi^{-1}(\beta)$ *denotes the* β*-quantile of the standard normal distribution on* $\mathbb{R}$*.*

Proof.

a) Let $C_X \in \mathcal{C}$ be arbitrary, but fixed. Since h is continuously differentiable, h is also Lipschitz-continuous with Lipschitz constant $L > 0$. Therefore, with Theorem 2.1 we get

$$\begin{aligned}
&\left|\widehat{\mathrm{FWER}}_{\vartheta^*,C_X}(\boldsymbol{\varphi}) - \alpha\right| = \\
&= \left|1-\alpha-C_T\left(1-\hat{\alpha}_{loc,n},\dots,1-\hat{\alpha}_{loc,n}\right)\right| \\
&= \left|h\left(B_{\mathbf{K}}\left(\hat{C}_{X,n}\right)\left(1-\hat{\alpha}_{loc,n},\dots,1-\hat{\alpha}_{loc,n}\right)\right) - h\left(C_X\left(1-\hat{\alpha}_{loc,n},\dots,1-\hat{\alpha}_{loc,n}\right)\right)\right| \le \\
&\le \left\|h\left(B_{\mathbf{K}}\left(\hat{C}_{X,n}\right)\right) - h\left(C_X\right)\right\|_\infty \le \\
&\le L \cdot \left\|B_{\mathbf{K}}\left(\hat{C}_{X,n}\right) - C_X\right\|_\infty = \\
&= O\left(n^{-1/2}(\log\log n)^{1/2}\right) \text{ almost surely.}
\end{aligned}$$

b) Letting $p := h^{\leftarrow}(1-\alpha)$, Lemma 7.2 yields that

$$\begin{aligned}
\sqrt{n}\left(1-\hat{\alpha}_{loc,n} - C_X^{\leftarrow}(p)\right) &= \sqrt{n}\left(B_{\mathbf{K}}\left(\hat{C}_{X,n}\right)^{\leftarrow}(p) - C_X^{\leftarrow}(p)\right) \\
&\xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma^2\left(C_X^{\leftarrow}(p),\dots,C_X^{\leftarrow}(p)\right)}{\left(C_X'\left(C_X^{\leftarrow}(p)\right)\right)^2}\right).
\end{aligned}$$

Therefore, applying the Delta Method to $u \mapsto C_T(u,\dots,u)$, we have that

$$\begin{aligned}
&\sqrt{n}\left(\widehat{\mathrm{FWER}}_{\vartheta^*,C_X}(\boldsymbol{\varphi}) - \alpha\right) = \\
&= -\sqrt{n}\left(C_T\left(1-\hat{\alpha}_{loc,n},\dots,1-\hat{\alpha}_{loc,n}\right) - (1-\alpha)\right) = \\
&= -\sqrt{n}\left(C_T\left(1-\hat{\alpha}_{loc,n},\dots,1-\hat{\alpha}_{loc,n}\right) - C_T\left(C_X^{\leftarrow}(p),\dots,C_X^{\leftarrow}(p)\right)\right) \\
&\xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma^2\left(C_X^{\leftarrow}(p),\dots,C_X^{\leftarrow}(p)\right)}{\left(C_X'\left(C_X^{\leftarrow}(p)\right)\right)^2} \cdot \left(C_T'\left(C_X^{\leftarrow}(p)\right)\right)^2\right).
\end{aligned}$$

The result follows from the definition of p.

c) Since $\hat{\sigma}_n \to \sigma_\alpha$ almost surely and particularly, in probability for $n \to \infty$, the assertion follows directly from part b) using Slutsky's Theorem. □

If the function h is unknown, one may approximate the value of $\hat{\alpha}_{loc,n}$ with high precision by a Monte Carlo simulation for a given number M of Monte Carlo repetitions. To this end, generate $M \times n$ pseudo-random vectors that follow the estimated (joint) distribution of $\mathbf{X}$ under $\boldsymbol{\vartheta}^*$, by combining $B_{\mathbf{K}}(\hat{C}_{X,n})$ and the marginal cdfs $F_1, \dots, F_m$ of $X_1, \dots, X_m$ under the global hypothesis. From these, calculate a pseudo-sample $\mathbf{T}_1, \dots, \mathbf{T}_M$ from the distribution of $\mathbf{T}$ under $\boldsymbol{\vartheta}^*$. Then, $\mathbf{G}(\mathbf{T}_1), \dots, \mathbf{G}(\mathbf{T}_M)$ constitutes a pseudo-random sample from the estimator of C_T, and the empirical equi-coordinate $(1-\alpha)$-quantile of this pseudo-sample approximates $\hat{\alpha}_{loc,n}$. Since the number M of pseudo-random vectors to be generated is in principle unlimited, Theorem 3.6 continues to hold true if this strategy is pursued. We will make use of this approach in the more involved examples studied in Section 4 and Section 5.

4 | SIMULATION STUDY

In this section, we report the results of a simulation study regarding the FWER and the power of multiple tests that are empirically calibrated as proposed in Section 3. Assume w.l.o.g. that $I_0(\boldsymbol{\vartheta}) := \{1, ..., m_0\}$ and let $m_1 := m - m_0$. The empirical FWER is given by the relative frequency over the L simulation runs of the occurrence of at least one false rejection, i.e.,

$$\widehat{\mathrm{FWER}}_{\vartheta, C_X}(\boldsymbol{\varphi}) := L^{-1} \sum_{\ell=1}^{L} \mathbb{1}_{\bigcup_{j=1}^{m_0} \left\{ \varphi_j^{(\ell)} = 1 \right\}} \left(\mathbf{x}_1^{(\ell)}, \dots, \mathbf{x}_n^{(\ell)} \right).$$

Likewise, the empirical power is defined as the average proportion of true rejections, that is,

$$\widehat{\mathrm{power}}(\boldsymbol{\varphi}) := L^{-1} \sum_{\ell=1}^{L} \left(m_1^{-1} \sum_{j=m_0+1}^{m} \mathbb{1}_{\left\{ \varphi_j^{(\ell)} = 1 \right\}} \left(\mathbf{x}_1^{(\ell)}, \dots, \mathbf{x}_n^{(\ell)} \right) \right),$$

where $(\mathbf{x}_1^{(\ell)}, \dots, \mathbf{x}_n^{(\ell)}) \in \mathcal{X}^n$ denotes the pseudo-sample in the ℓ-th simulation run.

The setting is as follows. We simulate from various one-parametric copula models (namely, Frank, Clayton, Gumbel, Student's t with four degrees of freedom, and the product copula) with parameters corresponding to weak (Kendall's $\tau \approx 0.25$) and strong dependence (Kendall's $\tau \approx 0.75$), respectively. In the case of t_4-copulas we restrict our attention to the case of equi-correlation, and the parameter is the equi-correlation coefficient. For convenience (and without loss of generality), the data are marginally normally distributed with all marginal variances equal to one. In the inference procedures, however, we assume these variances to be unknown, leading to Studentized test statistics. For each $1 \le j \le m$, we let ϑ_j be the mean in coordinate j. In all simulation settings, ϑ_j is set to 0.4 under alternatives. The null hypotheses are given by $H_j : \{\vartheta_j = \vartheta_j^* = 0\}$, with two-sided alternatives. Hence, marginal two-sided t-tests are performed with multiplicity corrected local significance level. Our Bernstein procedure is compared with the widely used Bonferroni and Šidák methods.

Notice that Assumption 3.2 is fulfilled. From Lemma 3.3 we get that the LFC is indeed $\boldsymbol{\vartheta}^* = (0, \dots, 0)^\top$. Further, the marginal distribution functions of the test statistics are known (even for finite n) and the function h exists, since $u \mapsto C_X(u, \dots, u)$ is strictly increasing for the choices of C_X in this simulation study. However, the function h is unknown in contrast to the examples in Section 3.

The calculation of the Bernstein copula has been performed as in Example 4.2 of Cottin and Pfeifer (2014), which uses $K_j := n$ for all $j \in \{1, \dots, m\}$. This choice fulfills the assumption of Theorem 2.1. In order to meet the assumptions of Theorem 2.4 it would be necessary to choose K_j of slightly larger magnitude. Notice, however, that we consider small sample sizes $n \in \{20, 100\}$ in our simulations, such that asymptotic considerations do not apply here. Instead, some preliminary simulations indicated that the choice $K_j \equiv n$ is appropriate. The choice of n was motivated by the purpose to demonstrate how accurately the Bernstein estimator performs in a small sample scenario. For instance, the real data example that we will present in Section 5 has a sample size of $n = 20$. With the simulations presented here, we can thus evaluate the appropriateness of the application of the proposed methodology in this real data example.

Since the function h is assumed unknown here, we calibrate the proposed multiple test with the following algorithm that was outlined at the end of Section 3.

Algorithm 4.1.

1. *Choose a number M of Monte Carlo repetitions.*
2. *For each $b = 1, \dots, M$ draw a sample $\mathbf{U}_1^{\#b}, \dots, \mathbf{U}_n^{\#b}$ of $B_{\mathbf{K}}(\hat{C}_{\boldsymbol{X},n})$ and calculate*

$$X_{i,j}^{\#b} = \hat{\sigma}_j \cdot \Phi_j^{-1}\left(U_{i,j}^{\#b}\right) + \vartheta_j^*, 1 \le i \le n, 1 \le j \le m,$$

   where $\hat{\sigma}_j$ is the sample standard deviation of $X_{1,j}, \dots, X_{n,j}$.
3. *For all $1 \le j \le m$, compute*

$$T_j^{\#b} = T_j\left(\mathbf{X}_1^{\#b}, \dots, \mathbf{X}_n^{\#b}\right) = \left|\sqrt{n} \cdot \frac{\frac{1}{n}\sum_{i=1}^n X_{i,j}^{\#b} - \vartheta_j^*}{\hat{\sigma}_j^{\#b}}\right|$$

   and obtain the pseudo-sample

$$V_j^{\#b} = 2F_{t_{n-1}}\left(T_j^{\#b}\right) - 1$$

   from the copula of $\mathbf{T}$.
4. *Finally, calibrate $\hat{\boldsymbol{\alpha}}_{loc,n} = (\hat{\alpha}_{loc,n}^{(1)}, \dots, \hat{\alpha}_{loc,n}^{(m)})^\top$ by solving*

$$\#\left\{b \,\middle|\, V_j^{\#b} \le 1 - \hat{\alpha}_{loc,n}^{(j)} \text{ for all } 1 \le j \le m\right\} = \lceil (1-\alpha)\, M \rceil . \tag{4.1}$$

Notice that in (4.1), we implicitly weight the hypotheses. This means that the weights corresponding to the obtained $\hat{\boldsymbol{\alpha}}_{loc,n}$ depend on the simulation data, for convenience of implementation. In comparison, the classical Bonferroni and Šidák corrected local significance levels are given by

$$\alpha_{loc}^{(j)} = \frac{\alpha}{m} \text{ and } \alpha_{loc}^{(j)} = 1 - (1-\alpha)^{1/m}, 1 \le j \le m,$$

respectively.

The results are displayed in Table 1 (weak dependence with Kendall's $\tau \approx 0.25$) and Table 2 (strong dependence with Kendall's $\tau \approx 0.75$). They reveal that in this simulation study the Bernstein method performs best in the case that M is large and the proportion of true null hypotheses π_0 is not too large, that is, in these cases its empirical FWER is closer to α and its empirical power is higher than those of the generic calibrations. Under strong dependence the power of the Bernstein method increases even further. On the other hand, if all hypotheses are true then the empirical FWER for the Bernstein method can be above $\alpha = 5\%$ and M needs to be large in order to improve the empirical FWER. Surprisingly, the sample size n does not have a clear positive impact in this simulation study.

5 | APPLICATION

In this section, we analyze insurance claim data from $m = 19$ adjacent geographical regions (see Table 5). For every region $j \in \{1, \dots, 19\}$ these claims have, for confidentiality reasons, been adjusted to a neutral monetary scale. The claim amounts and types have been aggregated to full years, such that temporal dependencies are considered negligible. However, strong nonlinear spatial dependencies are likely to be present in the data. Hence, we treat each of the $n = 20$ rows in Table 5 as an independent repetition $\mathbf{X}_i = \mathbf{x}_i$ of an m-dimensional random vector $\mathbf{X} = (X_1, \dots, X_m)^\top$, where $1 \le i \le 20$ is the time index in years and $m = 19$ refers to the regions.

An important quantity for regulators and risk managers is the region-specific value-at-risk (VaR). The VaR at level p for region j is defined as the p-quantile of the (marginal) distribution of X_j, that is,

$$\mathrm{VaR}_j(p) := F_{X_j}^{\leftarrow}(p).$$

In insurance mathematics, typically considered values of p are close to one. Here, we chose $p = 0.995$. Our goal is to derive multiplicity-corrected confidence intervals for $\vartheta_j = \mathrm{VaR}_j(0.995)$, $1 \le j \le m = 19$, that are compatible with (i.e., dual to) the

TABLE 1 Comparison of empirical FWER and power regarding Bonferroni, Šidák, and Bernstein corrections under various weak dependency structures with $m = 20$, $\pi_0 = m_0/m \in \{0.5, 1\}$, $\alpha = 0.05$, $L = 1{,}000$, $M \in \{200, 1{,}000\}$, and $n \in \{20, 100\}$

				FWER			Power		
Family	π_0	M	n	**Bonferroni**	**Šidák**	**Bernstein**	**Bonferroni**	**Šidák**	**Bernstein**
Frank (2)	0.5	200	20	1.8%	1.8%	6.1%	7.5%	7.7%	14.7%
			100	1.6%	1.8%	6.9%	81.4%	81.7%	86.1%
		1000	20	2.6%	2.6%	4.2%	7.8%	8.0%	10.6%
			100	2.6%	2.6%	4.0%	82.0%	82.2%	84.6%
	1	200	20	5.2%	5.2%	14.6%			
			100	3.0%	3.1%	13.1%			
		1,000	20	5.5%	5.5%	7.8%			
			100	4.6%	4.9%	6.6%			
Gumbel (2)	0.5	200	20	2.5%	2.5%	6.4%	8.4%	8.6%	19.4%
			100	1.3%	1.5%	6.0%	80.5%	80.7%	89.9%
		1,000	20	1.3%	1.3%	3.8%	7.1%	7.2%	12.7%
			100	2.1%	2.2%	4.8%	80.6%	80.9%	88.1%
	1	200	20	1.9%	1.9%	9.8%			
			100	2.6%	2.6%	10.0%			
		1,000	20	2.7%	2.7%	5.3%			
			100	2.2%	2.2%	6.4%			
Clayton (1)	0.5	200	20	2.2%	2.2%	7.0%	7.0%	7.1%	14.3%
			100	2.1%	2.1%	6.0%	81.3%	81.5%	88.0%
		1,000	20	2.4%	2.4%	4.3%	7.0%	7.1%	9.6%
			100	1.8%	1.8%	3.9%	81.3%	81.5%	86.4%
	1	200	20	3.3%	3.4%	12.6%			
			100	4.6%	4.6%	14.7%			
		1,000	20	3.5%	3.7%	5.6%			
			100	3.6%	3.7%	7.4%			
$t_4(0.4)$	0.5	200	20	2.8%	2.8%	7.6%	6.7%	6.8%	13.4%
			100	2.0%	2.1%	8.0%	81.7%	82.0%	87.4%
		1,000	20	2.3%	2.3%	3.6%	7.3%	7.5%	10.2%
			100	2.9%	3.0%	4.0%	81.3%	81.5%	85.0%
	1	200	20	5.1%	5.1%	15.0%			
			100	4.1%	4.1%	12.8%			
		1,000	20	4.4%	4.5%	7.6%			
			100	3.3%	3.3%	6.9%			
Independence	0.5	200	20	2.5%	2.6%	8.2%	7.4%	7.6%	13.5%
			100	3.4%	3.4%	7.6%	81.8%	81.9%	86.0%
		1,000	20	2.9%	2.9%	3.9%	7.0%	7.2%	8.8%
			100	2.1%	2.2%	3.6%	81.4%	81.6%	82.8%
	1	200	20	5.3%	5.3%	14.3%			
			100	5.7%	5.8%	15.5%			
		1000	20	4.0%	4.1%	6.9%			
			100	4.2%	4.2%	7.6%			

TABLE 2 Comparison of empirical FWER and power regarding Bonferroni, Šidák, and Bernstein corrections under various strong dependency structures with $m = 20$, $\pi_0 = m_0/m \in \{0.5, 1\}$, $\alpha = 0.05$, $L = 1{,}000$, $M \in \{200, 1{,}000\}$, and $n \in \{20, 100\}$

				FWER			Power		
Family	π_0	M	n	**Bonferroni**	**Šidák**	**Bernstein**	**Bonferroni**	**Šidák**	**Bernstein**
Frank (14)	0.5	200	20	0.8%	0.8%	6.8%	8.1%	8.2%	22.5%
			100	0.6%	0.6%	7.0%	81.7%	81.9%	94.4%
		1,000	20	1.0%	1.0%	3.2%	7.9%	8.0%	18.3%
			100	0.9%	0.9%	4.1%	81.2%	81.4%	92.3%
	1	200	20	0.9%	1.0%	7.5%			
			100	1.0%	1.0%	8.8%			
		1,000	20	1.4%	1.4%	5.2%			
			100	1.1%	1.1%	5.3%			
Gumbel (4)	0.5	200	20	1.5%	1.6%	7.1%	7.7%	7.8%	23.3%
			100	0.6%	0.6%	6.2%	81.6%	81.8%	94.9%
		1,000	20	0.5%	0.5%	2.2%	7.6%	7.7%	18.1%
			100	1.1%	1.1%	4.3%	80.9%	81.1%	93.6%
	1	200	20	1.3%	1.3%	6.2%			
			100	0.9%	0.9%	7.9%			
		1,000	20	1.5%	1.5%	4.1%			
			100	1.4%	1.4%	6.3%			
Clayton (6)	0.5	200	20	0.9%	0.9%	4.8%	7.2%	7.3%	22.0%
			100	1.2%	1.3%	7.6%	81.3%	81.5%	94.9%
		1,000	20	0.8%	0.8%	3.5%	7.0%	7.1%	15.9%
			100	0.9%	0.9%	4.2%	80.8%	81.0%	93.0%
	1	200	20	1.3%	1.5%	5.9%			
			100	1.3%	1.3%	8.7%			
		1,000	20	1.4%	1.4%	4.1%			
			100	1.0%	1.0%	5.0%			
$t_4(0.9)$	0.5	200	20	1.6%	1.6%	6.9%	8.3%	8.4%	22.2%
			100	0.7%	0.8%	6.8%	80.9%	81.1%	94.3%
		1,000	20	1.0%	1.0%	2.3%	7.4%	7.6%	16.1%
			100	1.0%	1.0%	4.8%	81.4%	81.5%	93.0%
	1	200	20	1.8%	1.8%	7.8%			
			100	0.9%	1.0%	9.1%			
		1,000	20	1.5%	1.6%	4.1%			
			100	1.4%	1.4%	5.7%			

Bonferroni, Šidák, and Bernstein copula-based correction methods discussed before. To this end, let auxiliary point hypotheses be defined as $H_{\vartheta_j^*} : \{\vartheta_j = \vartheta_j^*\}$ for fixed $\vartheta_j^* > 0$. According to the Extended Correspondence Theorem (see Section 1.3 of Dickhaus, 2014), the set of all values ϑ_j^* for which $H_{\vartheta_j^*}$ is retained by a multiple test at FWER level α (leading to a local significance level $\alpha_{loc}^{(j)}$ in coordinate j) constitutes a confidence region at simultaneous confidence level $1 - \alpha$ for ϑ_j, $1 \leq j \leq m$. We set $\alpha = 5\%$.

Our model assumptions are analogous to those from the examples in the previous sections. It can be shown (cf. our argumentation in Example 3.4 (a)) that Assumption 3.2 (a) and (b) are fulfilled. On the other hand, it is difficult to check Assumption 3.2 (c) in many applications. For example, in the simulation study reported in Section 4 we used the fact that the data were simulated under some suitable copula families.

In quantitative risk management, it is common practice to model the excess distribution of X_j over some given threshold u_j by a generalized Pareto distribution (GPD) (cf., e.g., Section 7.2.2 of McNeil et al., 2005).

Definition 5.1 (Definition 7.16 of McNeil et al. (2005)). For shape parameter $\xi \in \mathbb{R}$ and scale parameter $\beta > 0$, the cdf of the GPD is given by

$$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi x/\beta)^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp(-x/\beta), & \xi = 0, \end{cases}$$

where $x \geq 0$ if $\xi \geq 0$ and $0 \leq x \leq -\beta/\xi$ if $\xi < 0$.

In the remainder, we make the following assumption.

Assumption 5.2. For every $1 \leq j \leq m = 19$ there exists a threshold u_j and parameter values ξ_j and β_j such that

$$\mathbb{P}\left[X_j - u_j \leq x \,\middle|\, X_j > u_j\right] \approx G_{\xi_j,\beta_j}(x)$$

for all $x \geq 0$.

Under Assumption 5.2, an approximation of the VaR at level p for region j is given by

$$\mathrm{VaR}_{\xi_j,\beta_j}(p) \approx u_j + \frac{\beta_j}{\xi_j}\left(\left(\frac{1-p}{1 - F_{X_j}(u_j)}\right)^{-\xi_j} - 1\right) =: q_j(\xi_j, \beta_j),$$

provided that $p \geq F_{X_j}(u_j)$. For ease of notation, we let $\vartheta_j = q_j(\xi_j, \beta_j)$ in the sequel.

For computational convenience, we carried out the test for $H_{\vartheta_j^*}$ as a confidence-region test in the sense of Aitchison (1964) based on the family

$$\left(H_{\xi_j^*,\beta_j^*} : \left\{\xi_j = \xi_j^*, \beta_j = \beta_j^*\right\} \,\middle|\, \beta_j^* > 0, \xi_j^* \in \mathbb{R}\right) \tag{5.1}$$

of point hypotheses. Namely, the test procedure works as follows.

Algorithm 5.3.

1. *Test each $H_{\xi_j^*,\beta_j^*}$ by an arbitrary level $\alpha_{loc}^{(j)}$ test, where $\alpha_{loc}^{(j)}$ denotes a multiplicity-corrected significance level based on the Bonferroni, Šidák, or Bernstein copula calibration, respectively.*
2. *Let a confidence region $C_{\xi_j,\beta_j}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ at confidence level $1 - \alpha_{loc}^{(j)}$ for (ξ_j, β_j) be defined as the set of all parameter values (ξ_j^*, β_j^*) for which $H_{\xi_j^*,\beta_j^*}$ is retained.*
3. *Reject $H_{\vartheta_j^*}$ at level $\alpha_{loc}^{(j)}$, if the set $\{(\xi_j^*, \beta_j^*) : q_j(\xi_j^*, \beta_j^*) = \vartheta_j^*\}$ has an empty intersection with $C_{\xi_j,\beta_j}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$.*

Due to Algorithm 5.3, it suffices to construct point hypothesis tests for (5.1). A standard technique for testing parametric hypotheses is to perform a likelihood ratio test. In the risk management context, this method is described in Appendix A.3.5 of McNeil et al. (2005). Define the random variable $N_{u_j} := \#\{1 \leq i \leq n | X_{i,j} > u_j\}$ and let $\tilde{X}_{1,j}, \ldots, \tilde{X}_{N_{u_j},j}$ denote the corresponding sub-sample for region j. Then the excesses $Y_{1,j}, \ldots, Y_{N_{u_j},j}$ over u_j are defined by

$$Y_{i,j} := \tilde{X}_{i,j} - u_j.$$

The test statistic for testing $H_{\xi_j^*,\beta_j^*}$ is then given by

$$T_j\left(Y_{1,j,}\ldots, Y_{N_{u_j},j}; \xi_j^*, \beta_j^*\right) := -2\log\Lambda\left(Y_{1,j,}\ldots, Y_{N_{u_j},j}; \xi_j^*, \beta_j^*\right),$$

where the likelihood ratio Λ is defined by

$$\Lambda\left(Y_{1,j,}\ldots, Y_{N_{u_j},j}; \xi_j^*, \beta_j^*\right) := \frac{L\left(Y_{1,j,}\ldots, Y_{N_{u_j},j}; \xi_j^*, \beta_j^*\right)}{\sup_{(\xi,\beta)} L\left(Y_{1,j,}\ldots, Y_{N_{u_j},j}; \xi, \beta\right)}$$

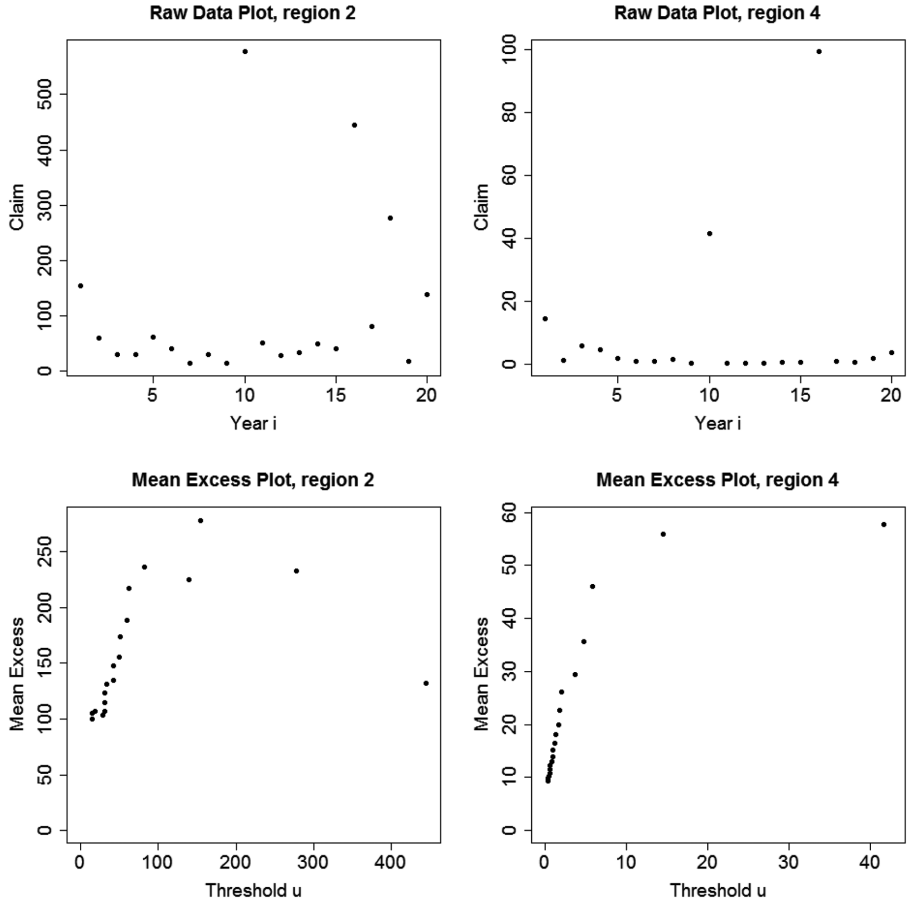

FIGURE 2 Raw data and mean excess plots for regions 2 and 4. The graphs in the upper panel display the data from Table 5 for $j \in \{2, 4\}$, respectively. The graphs in the lower panel show the corresponding mean excess plots

with log-likelihood function

$$\log L\left(Y_{1,j,\ldots,}Y_{N_{u_j},j};\xi,\beta\right) = -N_{u_j}\log\beta - \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{N_{u_j}}\log\left(1+\xi\frac{Y_{i,j}}{\beta}\right).$$

Under $H_{\xi_j^*,\beta_j^*}$, T_j is asymptotically χ^2-distributed with two degrees-of-freedom. This means that the (asymptotic) confidence interval $C_{\xi_j,\beta_j}(\mathbf{x}_1,\ldots,\mathbf{x}_n)$ in the second step of Algorithm 5.3 is given by

$$C_{\xi_j,\beta_j}\left(\mathbf{X}_1,\ldots,\mathbf{X}_n\right) = \left\{\left(\xi_j^*,\beta_j^*\right) : T_j\left(Y_{1,j,\ldots,}Y_{N_{u_j},j};\xi_j^*,\beta_j^*\right) \le F_{\chi_2^2}^{-1}\left(1-\alpha_{loc}^{(j)}\right)\right\}. \tag{5.2}$$

Utilizing (5.2), the confidence region $[\vartheta_j^{\text{lower}}, \vartheta_j^{\text{upper}}]$ for ϑ_j based on the third step of Algorithm 5.3 is constructed by finding the minimum value $\vartheta_j^{\text{lower}} = \min q_j(\xi_j^*,\beta_j^*)$ and the maximum value $\vartheta_j^{\text{upper}} = \max q_j(\xi_j^*,\beta_j^*)$, where (ξ_j^*,β_j^*) are located on the boundary of $C_{\xi_j,\beta_j}(\mathbf{x}_1,\ldots,\mathbf{x}_n)$.

A graphical method for the determination of a suitable threshold u_j is based on the mean excess plot in coordinate j; see Section 7.2.2 of McNeil et al. (2005) for details. Namely, all possible values u of u_j are plotted against the mean of the values of $Y_{1,j,\ldots,}Y_{N_u,j}$. If the GPD model is appropriate, the plot should yield an approximately linear graph for arguments exceeding u_j. Usually the few largest values of u are ignored, because they lead to very small values of N_u.

For example, Figure 2 shows the mean excess plots for the two regions 2 and 4. The mean excess plot for region 2 is approximately linear when ignoring the three smallest and the four largest values of u. This means that a suitable threshold u_2 would be

TABLE 3 Estimated parameters $\hat{\xi}_j$, $\hat{\beta}_j$, $1 \le j \le 19$, for the region-specific GPD models. Estimation has been performed via maximum likelihood

$\hat{\xi}_j$	0.41	1.17	0.75	1.43	0.87	1.51	1.10	0.30	0.49	0.79
	0.56	0.98	1.00	0.73	0.47	0.81	1.08	0.60	0.89	
$\hat{\beta}_j$	19.59	22.21	18.41	0.82	1.10	1.56	4.57	9.75	2.91	6.46
	0.64	0.99	5.12	3.42	20.34	4.52	6.98	1.96	1.64	

TABLE 4 Lower confidence bounds $\vartheta_j^{\text{lower}}$ for the 99.5% VaR, $1 \le j \le 19$, obtained by the Bonferroni, the Šidák, and the Bernstein copula method, respectively. The results for the Bernstein method rely on $M = 1{,}000$ Monte Carlo repetitions in Algorithm 5.4

Bonferroni	89.08	283.30	126.20	19.41	10.00	36.68	62.57	39.45	14.62	51.14
	3.74	10.13	53.74	25.43	101.62	37.11	84.99	12.79	14.80	
Šidák	89.22	284.03	126.46	19.48	10.03	36.81	62.75	39.51	14.64	51.25
	3.75	10.15	53.82	25.47	101.78	37.20	85.20	12.81	14.83	
Bernstein	91.59	287.32	127.61	19.81	10.13	37.37	63.54	38.82	14.73	51.74
	3.78	10.25	52.89	26.27	99.90	37.58	82.71	12.91	14.98	

between 18.815 and 28.316. Similarly, the mean excess plot for region 4 is approximately linear when ignoring the two largest values of u, hence $u_4 < 0.321$. Based on such considerations, we chose the thresholds $\mathbf{u} = (u_1, \dots, u_{19})^\top$ given by

$$\mathbf{u} := (1.0, 28.0, 9.0, 0.3, 0.2, 0.4, 2.6, 1.2, 0.4, 1.1, 0.1, 0.2, 22.5, 1.6, 3.2, 0.2, 12.5, 1.2, 0.5)^\top.$$

Finally, it remains to determine the local significance levels $(\alpha_{loc}^{(j)})_{1 \le j \le 19}$. In the case of the Bonferroni or the Šidák method, this is trivial. To calibrate the local significance levels with the Bernstein method, we employed a modified version of Algorithm 4.1 based on the empirical excess distribution. Algorithm 5.4 yields a resampling-based approximation of the copula of the vector $\mathbf{T} = (T_1, \dots, T_m)^\top$ of the region-specific likelihood ratio test statistics.

Algorithm 5.4.

1. *For every* $1 \le j \le m$, *estimate the parameters* ξ_j *and* β_j *of the excess distribution of* X_j *via maximum likelihood and calculate* N_{u_j}.
2. *Choose a number* M *of Monte Carlo repetitions.*
3. *For each* $1 \le b \le M$ *draw a pseudo sample* $\mathbf{U}_1^{\#b}, \dots, \mathbf{U}_n^{\#b}$ *from the (empirical) Bernstein copula* $B_{\mathbf{K}}(\hat{C}_{\boldsymbol{X},n})$ *and calculate the corresponding GPD excesses*

$$Y_{i,j}^{\#b} = G_{\hat{\xi}_j,\hat{\beta}_j}^{\leftarrow}\left(U_{(i),j}^{\#b}\right), 1 \le i \le N_{u_j}, 1 \le j \le m,$$

   where $U_{(i),j}^{\#b}$ *denotes the* i*-th reverse order statistic of* $(U_{i,j}^{\#b})_{1 \le i \le n}$.
4. *For each* $1 \le j \le m$, *compute* $T_j^{\#b} = T_j(Y_{1,j}^{\#b}, \dots, Y_{N_{u_j},j}^{\#b}; \hat{\xi}_j, \hat{\beta}_j)$, *and obtain the pseudo-sample*

$$V_j^{\#b} = \hat{G}_{j,M}\left(T_j^{\#b}\right), 1 \le j \le m$$

   from the copula of $\mathbf{T}$.
5. *Finally, calibrate* $\hat{\boldsymbol{\alpha}}_{loc,n} = (\hat{\alpha}_{loc,n}^{(1)}, \dots, \hat{\alpha}_{loc,n}^{(m)})^\top$ *by solving*

$$\#\left\{b \,\middle|\, V_j^{\#b} \le 1 - \hat{\alpha}_{loc,n}^{(j)} \text{ for all } 1 \le j \le m\right\} = \lceil (1-\alpha) M \rceil.$$

Table 3 displays the parameter estimates for the region-specific GPD models, and Table 4 displays the lower bounds $(\vartheta_j^{\text{lower}})_{1 \le j \le m}$ of the region-specific confidence intervals for the 99.5% VaR obtained by the Bonferroni, Šidák, and Bernstein copula calibration, respectively.

Similarly as in Algorithm 4.1, an implicit weighting has been employed for the determination of the local significance levels $(\alpha_{loc}^{(j)})_{1 \le j \le m}$ in Algorithm 5.4. Therefore, the confidence bounds obtained with the Bernstein copula method are not guaranteed to be more informative (i.e., larger) than the ones obtained by the Bonferroni or the Šidák methods for all regions. However, we

TABLE 5 Insurance claim data from 19 adjacent geographical regions over 20 years

		Region j							
Raw data $x_{i,j}$		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
year i	1	23.664	154.664	40.569	14.504	10.468	7.464	22.202	17.682
	2	1.080	59.545	3.297	1.344	1.859	0.477	6.107	7.196
	3	21.731	31.049	55.973	5.816	14.869	20.771	3.580	14.509
	4	28.990	31.052	30.328	4.709	0.717	3.530	6.032	6.512
	5	53.616	62.027	57.639	1.804	2.073	4.361	46.018	22.612
	6	29.950	41.722	12.964	1.127	1.063	4.873	6.571	11.966
	7	3.474	14.429	10.869	0.945	2.198	1.484	4.547	2.556
	8	10.020	31.283	21.116	1.663	2.153	0.932	25.163	3.222
	9	5.816	14.804	128.072	0.523	0.324	0.477	3.049	7.791
	10	170.725	576.767	108.361	41.599	20.253	35.412	126.698	71.079
	11	21.423	50.595	4.360	0.327	1.566	64.621	5.650	1.258
	12	6.380	28.316	3.740	0.442	0.736	0.470	3.406	7.859
	13	124.665	33.359	14.712	0.321	0.975	2.005	3.981	4.769
	14	20.165	49.948	17.658	0.595	0.548	29.350	6.782	4.873
	15	78.106	41.681	13.753	0.585	0.259	0.765	7.013	9.426
	16	11.067	444.712	365.351	99.366	8.856	28.654	10.589	13.621
	17	6.704	81.895	14.266	0.972	0.519	0.644	8.057	18.071
	18	15.550	277.643	26.564	0.788	0.225	1.230	26.800	64.538
	19	10.099	18.815	9.352	2.051	1.089	6.102	2.678	4.064
	20	8.492	138.708	46.708	3.680	1.132	1.698	165.600	7.926

9	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
12.395	18.551	1.842	4.100	46.135	14.698	44.441	7.981	35.833	10.689	7.299
1.436	3.720	0.429	1.026	7.469	7.058	4.512	0.762	14.474	9.337	0.740
17.175	87.307	0.209	2.344	22.651	4.117	26.586	3.920	13.804	2.683	3.026
0.682	3.115	0.521	0.696	31.126	1.878	29.423	6.394	18.064	1.201	0.894
1.581	11.179	2.715	1.327	40.156	4.655	104.691	28.579	17.832	1.618	3.402
15.676	24.263	4.832	0.701	16.712	11.852	29.234	7.098	17.866	5.206	5.664
0.456	1.137	0.268	0.580	11.851	2.057	11.605	0.282	16.925	2.082	1.008
1.581	5.477	0.741	0.369	3.814	1.869	8.126	1.032	14.985	1.390	1.703
4.079	7.002	0.524	6.554	5.459	3.007	8.528	1.920	5.638	2.149	2.908
21.762	64.582	9.882	6.401	106.197	44.912	191.809	90.559	154.492	36.626	36.276
0.626	3.556	1.052	8.277	22.564	8.961	19.817	16.437	25.990	2.364	6.434
0.894	3.591	0.136	0.364	28.000	7.574	3.213	1.749	12.735	1.744	0.558
2.006	1.973	1.990	15.176	57.235	23.686	110.035	17.373	7.276	2.494	0.525
2.921	6.394	0.630	0.762	25.897	3.439	8.161	3.327	24.733	2.807	1.618
2.180	3.769	0.770	15.024	36.068	1.613	6.127	8.103	12.596	4.894	0.822
9.589	19.485	0.287	0.464	24.211	38.616	51.889	1.316	173.080	3.557	11.627
5.515	13.163	0.590	2.745	16.124	2.398	20.997	2.515	5.161	2.840	3.002
2.637	80.711	0.245	0.217	12.416	4.972	59.417	3.762	24.603	7.404	19.107
2.373	2.057	0.415	0.351	10.707	2.468	10.673	1.743	27.266	1.368	0.644
2.972	5.237	0.566	0.708	22.646	6.652	14.437	63.692	113.231	7.218	2.548

observe improvements in almost all regions j. It is remarkable that this expected behavior of the Bernstein copula calibration can already be verified for the rather moderate sample size of $n = 20$, because the likelihood ratio tests and the Bernstein copula calibration are both based on asymptotic considerations.

We omitted the values of $(\vartheta_j^{\text{upper}})_{1\le j\le m}$, because they are uninformative (extremely large). This is in line with the fact that all scale parameter estimates $\hat{\xi}_j$ in Table 3 are positive. For $\xi \ge 0$, the GPD has infinite support, thus the modeled 99.5% VaR tends to be very large.

6 | DISCUSSION

We have derived a nonparametric approach to the calibration of multiple testing procedures that take the joint distribution of test statistics into account. In contrast to previous approaches that were restricted to cases with low-dimensional copula parameters, the Bernstein copula-based approximation of the local significance levels proposed in the present work can be applied under almost no assumptions regarding the dependency structures among test statistics or p-values, respectively. This makes the proposed methodology an attractive choice for data the dependency structure of which has not been explicitly modeled prior to the statistical analysis. Furthermore, our empirical results on simulated as well as on real-life data indicate the gain in power which is possible by the consideration of the dependency structure among test statistics in the calibration of the multiple test. This is particularly important for modern applications with high dimensionality of, but also pronounced dependencies in the data.

On the other hand, Theorem 3.6 provides a precise asymptotic performance guarantee for the empirically calibrated multiple test, meaning that a sharp upper bound for its realized FWER can be obtained, at least asymptotically for large sample sizes. This is in contrast to most of the existing resampling-based multiple test procedures like the "max T" and "min P" tests proposed by Westfall and Young (1993), which are obvious competitors of our approach.

Future work shall explore the case that some qualitative assumptions regarding the dependency structure are at hand. For example, it will be interesting to quantify the uncertainty of the FWER of a multiple testing procedure that is calibrated by assuming an Archimedean p-value copula as in Bodnar and Dickhaus (2014). In this case, nonparametric estimation of the copula generator function as for instance proposed by Lambert (2007) will lead to an empirical calibration of the multiple test.

7 | AUXILIARY RESULTS

In this section, two auxiliary lemmas are formulated and proved. The first lemma is used in the proofs of Theorem 2.1 and Theorem 2.4. The second lemma follows from Theorem 2.4 and is used in Theorem 3.6.

Lemma 7.1. *It holds that*

$$\left\| B_{\mathbf{K}}\left(C_X\right) - C_X \right\|_\infty \le \frac{1}{2}\sum_{j=1}^{m} K_j^{-1/2},$$

where $\|g\|_\infty := \sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} |g(\mathbf{u})|$ *for* $g : [\mathbf{0},\mathbf{1}] \to \mathbb{R}$.

Proof. We get

$$\begin{aligned}
\left\| B_{\mathbf{K}}\left(C_X\right) - C_X \right\|_\infty &\le \sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} \left| C_X\left(\mathbf{k}/\mathbf{K}\right) - C_X\left(\mathbf{u}\right)\right| \prod_{j=1}^{m} P_{k_j,K_j}\left(u_j\right) \le \\
&\le \sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} \sum_{j_1=1}^{m} \left| \frac{k_{j_1}}{K_{j_1}} - u_{j_1} \right| \cdot \prod_{j_2=1}^{m} P_{k_{j_2},K_{j_2}}\left(u_{j_2}\right) \le \\
&\le \frac{1}{2}\sum_{j=1}^{m} K_j^{-1/2},
\end{aligned}$$

where the second inequality follows from the Lipschitz property of multivariate copula (cf. Section 2 of Sancetta and Satchell, 2004). For the last inequality we use the fact that $P_{k_j,K_j}(u_j)$ is the probability mass function of the binomial distribution for each

$u_j \in [0,1]$ and $j = 1, \dots, m$. Therefore, by the Jensen inequality it follows that

$$\sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} \sum_{j_1=1}^{m} \left| \frac{k_{j_1}}{K_{j_1}} - u_{j_1} \right| \cdot \prod_{j_2=1}^{m} P_{k_{j_2},K_{j_2}}\left(u_{j_2}\right) =$$

$$= \sum_{j=1}^{m} \sup_{u_j\in[0,1]} \sum_{k_j=0}^{K_j} \left| \frac{k_j}{K_j} - u_j \right| P_{k_j,K_j}\left(u_j\right) \leq$$

$$\leq \sum_{j=1}^{m} \sup_{u_j\in[0,1]} \left(\sum_{k_j=0}^{K_j} \left(\frac{k_j}{K_j} - u_j \right)^2 P_{k_j,K_j}\left(u_j\right) \right)^{1/2} =$$

$$= \sum_{j=1}^{m} \sup_{u_j\in[0,1]} \left(\frac{u_j\left(1-u_j\right)}{K_j} \right)^{1/2} = \frac{1}{2} \sum_{j=1}^{m} K_j^{-1/2}.$$

□

Lemma 7.2. *Let $p \in (0,1)$. Suppose that $C'_X(C_X^{\leftarrow}(p)) > 0$ exists, then*

$$n^{1/2}\left(B_{\mathbf{K}}\left(\hat{C}_{X,n}\right)^{\leftarrow}(p) - C_X^{\leftarrow}(p) \right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma^2\left(C_X^{\leftarrow}(p), \dots, C_X^{\leftarrow}(p)\right)}{\left(C'_X\left(C_X^{\leftarrow}(p)\right)\right)^2}\right),$$

where $\sigma^2(\mathbf{u}) = \mathbb{V}[\mathbb{C}(\mathbf{u})]$, C'_X is the first derivative of $u \mapsto C_X(u, \dots, u)$, and $C_X^{\leftarrow}$, $B_{\mathbf{K}}(\hat{C}_{X,n})^{\leftarrow}$ is the quantile of $u \mapsto C_X(u, \dots, u)$, $u \mapsto B_{\mathbf{K}}(\hat{C}_{X,n})(u, \dots, u)$, respectively.

Remark 7.3. In order to prove this lemma, we need a slightly extended version of Theorem 2.4. Let $\mathbf{u} \in [\mathbf{0},\mathbf{1}]$ and $\mathbf{u}_n := \mathbf{u} \pm \boldsymbol{\epsilon}_n$, where $\boldsymbol{\epsilon}_n \to \mathbf{0}$ for $n \to \infty$, such that $\mathbf{u}_n \in [\mathbf{0},\mathbf{1}]$ for all $n \in \mathbb{N}$. Then under the assumptions of Theorem 2.4 it holds that

$$n^{1/2} \cdot \left(B_{\mathbf{K}}\left(\hat{C}_{X,n}\right)\left(\cdot \pm \boldsymbol{\epsilon}_n\right) - C_X\left(\cdot \pm \boldsymbol{\epsilon}_n\right) \right) \xrightarrow{d} \mathbb{C}$$

in $(C_X([\mathbf{0},\mathbf{1}]), \|\cdot\|_\infty)$.

The proof is essentially the same as that of Theorem 2.4. Notice that Lemma 7.1 and Bernstein's theorem hold uniformly. This means that we can use Lemma 7.1 directly again and Bernstein's theorem with an additional argument. We used Bernstein's theorem to show the uniform convergence of $g_n(f) \to g(f)$ for $n \to \infty$ and all $f \in S'$. Recall that f is any continuous function on the compact set $[\mathbf{0},\mathbf{1}]$. We need to show that $g_n(f) \to g(f)$ for $n \to \infty$ still holds uniformly when we transform the argument $\mathbf{u}$ of $g_n(f)$ to $\mathbf{u}_n$. We get that

$$\sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} \left| g_n(f)\left(\mathbf{u}_n\right) - g(f)(\mathbf{u}) \right| = \sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} \left| B_{\mathbf{K}(n)}(f)\left(\mathbf{u}_n\right) - f(\mathbf{u}) \right| \leq$$

$$\leq \left\| B_{\mathbf{K}(n)}(f) - f \right\|_\infty + \sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} \left| f\left(\mathbf{u}_n\right) - f(\mathbf{u}) \right|.$$

The first summand again converges to 0 because of Bernstein's theorem. The second summand converges to 0 because of the uniform continuity of f. The function g'_n defined by $g'_n(f)(\mathbf{u}) := g_n(f)(\mathbf{u}_n) = B_{\mathbf{K}(n)}(f)(\mathbf{u}_n)$ is then used in the generalized continuous mapping theorem instead of g_n.

Proof. We argue similarly to the proof of Theorem A in Section 2.3.3 of Serfling (1980). Fix $p \in (0,1)$ and let

$$G_n(t) := \mathbb{P}\left[\frac{n^{1/2}\left(B_{\mathbf{K}}\left(\hat{C}_{X,n}\right)^{\leftarrow}(p) - C_X^{\leftarrow}(p) \right)}{\tilde{\sigma}} \leq t \right],$$

where $\tilde{\sigma} := \frac{\sigma(C_X^{\leftarrow}(p),\ldots,C_X^{\leftarrow}(p))}{C_X'(C_X^{\leftarrow}(p))}$. Let $u_n := t\tilde{\sigma}n^{-1/2} + C_X^{\leftarrow}(p)$. We have

$$G_n(t) = \mathbb{P}\left[B_{\mathbf{K}}\left(\hat{C}_{X,n}\right)^{\leftarrow}(p) \le u_n\right] =$$
$$= \mathbb{P}\left[p \le B_{\mathbf{K}}\left(\hat{C}_{X,n}\right)(u_n,\ldots,u_n)\right]$$

Put $c_{nt} := \frac{n^{1/2}(C_X(u_n,\ldots,u_n)-p)}{\sigma(u_n,\ldots,u_n)}$. Then it holds that

$$G_n(t) = \mathbb{P}\left[-c_{nt} \le Z_n\right],$$

where $Z_n := \frac{n^{1/2}(B_{\mathbf{K}}(\hat{C}_{X,n})(u_n,\ldots,u_n)-C_X(u_n,\ldots,u_n))}{\sigma(u_n,\ldots,u_n)}$. Furthermore, we get

$$\Phi(t) - G_n(t) = \mathbb{P}\left[Z_n < -c_{nt}\right] - (1-\Phi(t)) =$$
$$= \mathbb{P}\left[Z_n < -c_{nt}\right] - \Phi\left(-c_{nt}\right) + \Phi(t) - \Phi\left(c_{nt}\right). \tag{7.1}$$

Since C_X and $\partial_j C_X$, $1 \le j \le m$, are continuous, we have

$$\lim_{n\to\infty} c_{nt} = \lim_{n\to\infty}\left(t \cdot \frac{\tilde{\sigma}}{\sigma\left(u_n,\ldots,u_n\right)} \cdot \frac{C_X\left(u_n,\ldots,u_n\right) - C_X\left(C_X^{\leftarrow}(p),\ldots,C_X^{\leftarrow}(p)\right)}{t\tilde{\sigma}n^{-1/2}}\right) =$$
$$= t \cdot \frac{\tilde{\sigma}}{\sigma\left(C_X^{\leftarrow}(p),\ldots,C_X^{\leftarrow}(p)\right)} \cdot C_X'\left(C_X^{\leftarrow}(p)\right) =$$
$$= t.$$

Next, we utilize Remark 7.3 (restricted to the point $\mathbf{u} := (C_X^{\leftarrow}(p),\ldots,C_X^{\leftarrow}(p))$ with $\mathbf{u}_n := (u_n,\ldots,u_n)$) and Polya's Theorem (see Section 1.5.3 of Serfling, 1980) to show uniform convergence of the distribution function of Z_n to the standard normal distribution function. Since Φ is continuous, we have

$$\lim_{n\to\infty}\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left[Z_n \le x\right] - \Phi(x)\right| = 0.$$

Using these two properties, (7.1) results in

$$\lim_{n\to\infty}\left|\Phi(t) - G_n(t)\right| \le \lim_{n\to\infty}\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left[Z_n < x\right] - \Phi(x)\right| + \lim_{n\to\infty}\left|\Phi(t) - \Phi\left(c_{nt}\right)\right| =$$
$$= 0.$$

□

ACKNOWLEDGMENTS

Funding of the first author has been provided by the Deutsche Forschungsgemeinschaft via the Research Unit FOR 1735 "Structural Inference in Statistics: Adaptation and Efficiency". We are grateful to two anonymous reviewers and to the Associate Editor for constructive suggestions which have lead to an improved presentation. Special thanks are due to the organizers of MCP 2017 for the successful conference.

CONFLICT OF INTEREST

The authors have declared no conflict of interest.

ORCID

Thorsten Dickhaus http://orcid.org/0000-0003-3084-3036

REFERENCES

Aitchison, J. (1964). Confidence-region tests. *Journal of Royal Statistics Society Series B*, *26*, 462–476.

Belalia, M. (2016). On the asymptotic properties of the Bernstein estimator of the multivariate distribution function. *Statistics and Probability Letters*, *110*, 249–256.

Bodnar, T., & Dickhaus, T. (2014). False discovery rate control under Archimedean copula. *Electronic Journal of Statistics*, *8*(2), 2207–2241.

Bouzebda, S., & Zari, T. (2013). Strong approximation of empirical copula processes by Gaussian processes. *Statistics*, *47*(5), 1047–1063.

Bücher, A., & Dette, H. (2010). A note on bootstrap approximations for the empirical copula process. *Statistics and Probability Letters*, *80*(23–24), 1925–1932.

Cerqueti, R., Costantini, M., & Lupi, C. (2012). A copula-based analysis of false discovery rate control under dependence assumptions. Economics & Statistics Discussion Paper 065/12, Università degli Studi del Molise, Dipartimento di Scienze Economiche, Gestionali e Sociali (SEGeS).

Charpentier, A., Fermanian, J.-D., & Scaillet, O. (2007). The estimation of copulas: Theory and practice. In J. Rank (Ed.), *Copulas: From theory to application in finance* (pp. 35–62). London, UK: Risk Books.

Chen, S. X. (1999). Beta kernel estimators for density functions. *Computational Statistics and Data Analysis*, *31*(2), 131–145.

Chêng, F. H. (1983). On the rate of convergence of Bernstein polynomials of functions of bounded variation. *Journal of Approximation Theory*, *39*(3), 259–274.

Cottin, C., & Pfeifer, D. (2014). From Bernstein polynomials to Bernstein copulas. *Journal of Applied Functional Analysis*, *9*(3–4), 277–288.

Deheuvels, P. (1979). La fonction de dependance empirique et ses propriétés. Un test non paramètrique d'independance. *Bull. Cl. Sci., V. Sér., Acad. R. Belg.*, *65*, 274–292.

Dickhaus, T. (2014). *Simultaneous statistical inference with applications in the life sciences*. Berlin Heidelberg, DE: Springer-Verlag.

Dickhaus, T., & Gierl, J. (2013). Simultaneous test procedures in terms of p-value copulae. *Proceedings on the 2nd Annual International Conference on Computational Mathematics, Computational Geometry & Statistics (CMCGS 2013)*, pp. 75–80. Global Science and Technology Forum (GSTF).

Dickhaus, T., & Stange, J. (2013). Multiple point hypothesis test problems and effective numbers of tests for control of the family-wise error rate. *Calcutta Statistical Association Bulletin*, *65*(257–260), 123–144.

Diers, D., Eling, M., & Marek, S. D. (2012). Dependence modeling in non-life insurance using the Bernstein copula. *Insurance Mathematics and Economics*, *50*(3), 430–436.

Embrechts, P., Lindskog, F., & McNeil, A. (2003). Modelling dependence with copulas and applications to risk management. In S. Rachev (Ed.), *Handbook of heavy tailed distributions in finance* (pp. 329–384). Amsterdam, NL: Elsevier Science B.V.

Gijbels, I., & Mielniczuk, J. (1990). Estimating the density of a copula function. *Communications in Statistics Theory and Methods*, *19*(2), 445–464.

Härdle, W. K., & Okhrin, O. (2010). De copulis non est disputandum–Copulae: an overview. *AStA Advances in Statistical Analysis*, *94*(1), 1–31.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363.

Janssen, P., Swanepoel, J., & Veraverbeke, N. (2012). Large sample behavior of the Bernstein copula estimator. *Journal of Statistical Planning and Inference*, *142*(5), 1189–1197.

Joe, H. (2014). *Dependence modeling with copulas*. Boca Raton, FL: CRC Press.

Kiefer, J. (1961). On large deviations of the empiric D. F. of vector chance variables and a law of the iterated logarithm. *Pacific Journal of Mathematics*, *11*, 649–660.

Lambert, P. (2007). Archimedean copula estimation using Bayesian splines smoothing techniques. *Computational Statistics and Data Analysis*, *51*(12), 6307–6320.

Longin, F. M. (2000). From value at risk to stress testing: The extreme value approach. *Journal of Banking and Finance*, *24*, 1097–1130.

McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative risk management. Concepts, techniques, and tools*. Princeton, NJ: Princeton University Press.

Nelsen, R. B. (2006). *An introduction to copulas*. 2nd ed. Springer series in statistics. New York, NY: Springer.

Omelka, M., Gijbels, I., & Veraverbeke, N. (2009). Improved kernel estimation of copulas: Weak convergence and goodness-of-fit testing. *The Annals of Statistics*, *37*(5B), 3023–3058.

Pfeifer, D., Mändle, A., & Ragulina, O. (2017). New copulas based on general partitions-of-unity and their applications to risk management (part II). *Dependence Modeling*, *5*(1), 246–255.

Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *The Annals of Statistics*, *4*, 912–923.

Sancetta, A., & Satchell, S. (2004). The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory*, *20*(03), 535–562.

Schmidt, R., Faldum, A., & Gerß, J. (2015). Adaptive designs with arbitrary dependence structure based on Fisher's combination test. *Statistical Methods and Applications*, *24*(3), 427–447.

Schmidt, R., Faldum, A., Witt, O., & Gerß, J. (2014). Adaptive designs with arbitrary dependence structure. *Biometrical Journal*, *56*(1), 86–106.

Schuster, E. F. (1985). Incorporating support constraints into nonparametric estimators of densities. *Communication in Statistics A—Theory Methods*, *14*(5), 1123–1136.

Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli*, *18*(3), 764–782.

Serfling, R. J. (1980). *Approximation theorems of mathematical statistics*. New York, NY: Wiley.

Stange, J., Bodnar, T., & Dickhaus, T. (2015). Uncertainty quantification for the family-wise error rate in multivariate copula models. *AStA Advances in Statistical Analysis*, *99*(3), 281–310.

Stange, J., Dickhaus, T., Navarro, A., & Schunk, D. (2016). Multiplicity- and dependency-adjusted *p*-values for control of the family-wise error rate. *Statistics and Probability Letters*, *111*, 32–40.

Stute, W. (1984). The oscillation behavior of empirical processes: The multivariate case. *Annals of Probability*, *12*(2), 361–379.

Swanepoel, J. W. H. (1986). A note on proving that the (modified) bootstrap works. *Communications in Statistics A—Theory Methods*, *15*(11), 3193–3203.

Westfall, P. H., & Young, S. S. (1993). *Resampling-based multiple testing: examples and methods for p-value adjustment*. Wiley series in probability and mathematical statistics. Applied probability and statistics. New York, NY: Wiley.

Whitt, W. (2002). *Stochastic-process limits: An introduction to stochastic-process limits and their application to queues*. Springer series in operations research. New York, NY: Springer-Verlag.

SUPPORTING INFORMATION

Additional Supporting Information including source code to reproduce the results may be found online in the supporting information tab for this article.

How to cite this article: Neumann A, Bodnar T, Pfeifer D, Dickhaus T. Multivariate multiple test procedures based on nonparametric copula estimation. *Biometrical Journal*. 2019;61:40–61. https://doi.org/10.1002/bimj.201700205