Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 0946 – 8633

Uncertainty quantification for the family-wise error rate in multivariate copula models

Jens Stange[1], Taras Bodnar[1], Thorsten Dickhaus[2]

submitted: October 30, 2013

[1] Humboldt-Universität zu Berlin
Department of Mathematics
Unter den Linden 6
10099 Berlin
E-Mail: stange@mathematik.hu-berlin.de
bodnar@mathematik.hu-berlin.de

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
E-Mail: Thorsten.Dickhaus@wias-berlin.de

No. 1862
Berlin 2013

2010 *Mathematics Subject Classification.* 62J15, 62F05, 62F03.

Key words and phrases. Delta method, Gumbel-Hougaard copula, multiple testing, simultaneous test procedure, subset pivotality.

This research was partly supported by the Deutsche Forschungsgemeinschaft via the Research Unit FOR 1735 “Structural Inference in Statistics: Adaptation and Efficiency“ (Taras Bodnar) and via grant No. DI 1723/3-1 (Jens Stange).

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

We derive confidence regions for the realized family-wise error rate (FWER) of certain multiple tests which are empirically calibrated at a given (global) level of significance. To this end, we regard the FWER as a derived parameter of a multivariate parametric copula model. It turns out that the resulting confidence regions are typically very much concentrated around the target FWER level, while generic multiple tests with fixed thresholds are in general not FWER-exhausting. Since FWER level exhaustion and optimization of power are equivalent for the classes of multiple test problems studied in this paper, the aforementioned findings militate strongly in favour of estimating the dependency structure (i. e., copula) and incorporating it in a multivariate multiple test procedure. We illustrate our theoretical results by considering two particular classes of multiple test problems of practical relevance in detail, namely, multiple tests for components of a mean vector and multiple support tests.

1 Introduction

Multiple testing is a hot topic in modern mathematical and applied statistics with a variety of applications in the life sciences like, for instance, in genetic association analyses, gene expression studies, functional magnetic resonance imaging, or brain-computer interfacing (see, e. g., Dickhaus (2013b) and references therein), as well as in economics and finance where testing the structure of an optimal portfolio plays a crucial role for the investment strategy (cf. Bodnar and Schmid (2008)). A multiple test problem is characterized by a family of $m > 1$ null hypotheses which have to be tested simultaneously based on the same data. To this end, typically (marginal) test statistics are constructed for each hypothesis.

The classical type I error criterion for multiple tests is the family-wise error rate (FWER) which is defined as the probability of at least one false rejection (type I error). Given an FWER level α, the decision rule of a multiple test is often described by local significance levels $\alpha_j, 1 \leq j \leq m$, for each marginal test. These α_j's have to be chosen such that the FWER is upper bounded by α. Two classical procedures are the Bonferroni correction (cf. Bonferroni (1935, 1936)) and the Šidák correction (Šidák (1967)), corresponding to

$$\alpha_j = \frac{\alpha}{m} \qquad \text{and} \qquad \alpha_j = 1 - (1 - \alpha)^{1/m},$$

respectively. It is well-known that both the Bonferroni and the Šidák corrections lead to FWER control for broad classes of dependent test statistics. More precisely, the Bonferroni correction is a generic procedure which can be applied under any arbitrary dependence structure, whereas the Šidák correction controls the FWER under certain forms of positive dependence among test statistics like positive lower orthant dependence or multivariate total positivity of order 2, including jointly stochastically independent test statistics; see, for instance, Block et al. (1992). On the other hand, this generic validity of the Bonferroni correction as well as the Šidák correction can lead to suboptimal power characteristics of the multiple tests if a concrete dependency structure can be assumed, meaning that their power can uniformly be improved if the dependence structure of the joint distribution of test statistics is explicitly taken into account.

The usage of copulae is highly recommendable for this purpose. The main reason is that it allows for separating the marginal distributions of test statistics from the dependence structure which is fully captured by the copula function. The application of copulae is currently becoming very popular in the theory of multiple tests (cf. Sarkar (2008); Ghosh (2011); Dickhaus and Gierl (2013); Bodnar and Dickhaus (2013)). In particular, Dickhaus and Gierl (2013) demonstrated that, under certain structural assumptions on the statistical model and the multiple test, the local significance levels $\alpha_j, 1 \leq j \leq m$, can precisely be calibrated via level sets of the copula of test statistics under the global hypothesis, i. e., when all null hypotheses are true.

However, in many practically relevant cases the copula is itself an unknown parameter which has to be estimated before multiple testing is performed. In this case, the question arises how to quantify the influence of the copula estimation on the performance of the multiple test. In particular, it is important to analyze if the empirically calibrated multiple test employing the estimated copula parameter still keeps the FWER level and outperforms a generic one which is based, for example, on the Bonferroni correction, at least with high probability. We deal with both of these problems in the present paper. First, two families of copulae, namely the family of elliptical copulae and the family of Archimedean copulae, are applied for modeling the joint distribution of test statistics. These choices are motivated by different types of limit theorems for sequences of independent and identically distributed (i.i.d.) random vectors (observables). Second, we deal with the estimation of copula parameters in detail and analyze the influence of the estimation variance (i. e., the covariance matrix of the estimator) on the performance of multiple tests. Third, we are considered with consistent bootstrap-based estimation of copula parameters when the dependence structure of test statistics can not straightforwardly be deduced from that of the original data.

The rest of the paper is structured as follows. In Section 2, we recall some theoretical background of multiple testing and copula modeling, and we formalize the connection between copulae and local significance levels. Estimation methods for copula parameters are discussed in Section 3.1, whereas their impact on the performance of multiple tests is analyzed in Section 3.2. In Section 4, we apply our theoretical findings to two important practical problems. In the first one (Section 4.1) the aim is to test hypotheses about the elements of a mean vector, while the second one deals with multiple testing of upper bounds of the supports of marginal distributions (Section 4.2). We conclude with a discussion in Section 5.

2 Notation and preliminaries

2.1 Multiple testing

Throughout the work the triple $(\mathcal{X}, \mathcal{F}, \mathcal{P})$ denotes a statistical model, where $\mathcal{P}$ is a family of probability distributions on the sample space $(\mathcal{X}, \mathcal{F})$. The family $\mathcal{P} = (\mathbb{P}_{\vartheta,\eta} : \vartheta \in \Theta, \eta \in \Xi)$ is indexed by two types of parameters, $\vartheta \in \Theta$ and $\eta \in \Xi$. We refer to the parameter $\vartheta \in \Theta$ as the parameter of interest, whereas $\eta \in \Xi$ stands for a nuisance parameter representing the dependency structure among the data. Let $\mathcal{H} = (H_j)_{j=1}^m$ denote a family of null hypotheses with $\emptyset \neq H_j \subset \Theta$ for $1 \leq j \leq m$. For a parameter $\vartheta \in \Theta$, we call the null hypothesis H_j true if $\vartheta \in H_j$ and false otherwise. A multiple test for $\mathcal{H}$ is a measurable mapping

$$\varphi = (\varphi_1, \ldots, \varphi_m)^\top : (\mathcal{X}, \mathcal{F}) \to (\{0,1\}^m, \mathfrak{P}(\{0,1\}^m)),$$

where $\varphi_j : \mathcal{X} \to \{0,1\}$ is a (non-randomized) test for H_j versus $K_j = \Theta \setminus H_j$, $1 \leq j \leq m$. We consider multiple tests of the form $\varphi = \varphi(\mathbf{T}, \mathbf{c})$, where each local test is given by $\varphi_j = \mathbf{1}(T_j > c_j)$ for a vector of real-valued test statistics $\mathbf{T} = (T_1, \ldots, T_m)^\top$, which tend to larger values under the respective alternative, and critical values $\mathbf{c} = (c_1, \ldots, c_m)^\top \in \mathbb{R}^m$.

For the calibration of φ, we aim at controlling the probability of at least one false rejection, commonly known as the FWER. For given $\vartheta \in \Theta$ and $\eta \in \Xi$, it is defined by

$$\text{FWER}_{\vartheta,\eta}(\varphi) = \mathbb{P}_{\vartheta,\eta}\left(\bigcup_{i \in I_0(\vartheta)} \{\varphi_i = 1\}\right), \tag{1}$$

where $I_0(\vartheta) = \{1 \leq j \leq m : \vartheta \in H_j\}$ denotes the index set of true hypotheses under ϑ. Notice that, although I_0 only depends on ϑ, the FWER of φ depends both on ϑ and η, because the distribution of φ (regarded as a statistic with values in $\{0,1\}^m$) typically crucially depends on the dependency structure in the data. For a predetermined value $\alpha \in (0,1)$, the multiple test φ controls the FWER at the (global) significance level α, if

$$\sup_{\vartheta \in \Theta, \eta \in \Xi} \text{FWER}_{\vartheta,\eta}(\varphi) \leq \alpha.$$

Let $H_0 := \bigcap_{j=1}^m H_j$ denote the intersection (or global) hypothesis of $\mathcal{H}$. We make the following general assumptions concerning the structure of the multiple test problem $(\mathcal{X}, \mathcal{F}, \mathcal{P}, \mathcal{H})$ and the multiple test φ.

Assumption 2.1.

(i) The nuisance parameter $\eta \in \Xi$ does not depend on the parameter $\vartheta \in \Theta$ and the marginal distribution of each T_j, $1 \leq j \leq m$, is determined by ϑ solely.

(ii) There exists a parameter $\vartheta^* \in H_0$ such that

$$\forall \vartheta \in \Theta : \forall \eta \in \Xi : \textit{FWER}_{\vartheta,\eta}(\varphi) \leq \textit{FWER}_{\vartheta^*,\eta}(\varphi). \tag{2}$$

In order to simplify notation, we put $\mathbb{P}^*_\eta = \mathbb{P}_{\vartheta^*,\eta}$ and $\textit{FWER}^*_\eta(\varphi) = \textit{FWER}_{\vartheta^*,\eta}(\varphi)$.

A parameter $\vartheta^* \in \Theta$ that satisfies (2) is called a least favorable configuration (LFC) for the FWER of φ (for fixed η). Sufficient conditions for LFCs being located in H_0 have been provided by Gabriel (1969) and Dickhaus and Stange (2013), among others. Let us also point out the strong similarity of Assumption 2.1.(ii) and the concept of subset pivotality (cf. Westfall and Young (1993), pp.42-43) which is often used in resampling-based multiple testing.

Let

$$\alpha_j = \sup_{\vartheta \in H_j} \mathbb{P}_{\vartheta,\eta}(\varphi_j = 1) = \sup_{\vartheta \in H_j} \mathbb{P}_{\vartheta,\eta}(T_j > c_j) \tag{3}$$

denote local (marginal) significance levels, when $\mathbf{c} = \mathbf{c}(\alpha)$ is chosen such that φ controls the FWER at level α. In practice, the multiple test φ is often carried out by means of the α_j and marginal p-values p_j, $1 \leq j \leq m$, which are defined by

$$p_j = \inf_{\tilde{\alpha} \in (0,1): T_j(x) > c_j(\tilde{\alpha})} \sup_{\vartheta \in H_j} \mathbb{P}_\vartheta(T_j > c_j(\tilde{\alpha})),$$

where $x \in \mathcal{X}$ denotes the actually observed data. The multiple test $\varphi(\mathbf{T}, \mathbf{c})$ is then equivalently given by $\varphi_j = \mathbf{1}(p_j \leq \alpha_j)$, $1 \leq j \leq m$. In the case of simple hypotheses $H_1, \ldots, H_m$ and continuous marginal cumulative distribution functions (cdfs) F_j of test statistics T_j under H_j, $1 \leq j \leq m$, the p-values are simply given by $p_j = 1 - F_j(T_j)$. Moreover, in the latter case p_j is exactly uniformly distributed on $[0, 1]$ under H_j. In the case of composite hypotheses, p-values are not necessarily uniformly distributed, but stochastically not larger than UNI$[0, 1]$, cf. Dickhaus (2013a). The ßtandardization"provided by transforming test statistics into p-values is useful for the interpretation of φ, especially if the test statistics have unbalanced scales. For the remainder of this work, this standardization is also used to separate the dependency structure among the test statistics (induced by η) from the marginal models, giving rise to the consideration of copula models.

2.2 Copulae

We start with a formal definition of the term copula.

Definition 2.1 (Copula)**.** *An m-dimensional copula is a multivariate distribution function on $[0, 1]^m$ with all marginal distributions equal to UNI$[0, 1]$.*

An equivalent but rather technical, geometric definition of copulae, involving quasi-monotonicity, can be found in the textbook by Nelsen (2006). The connection between marginal cdfs, joint cdfs, and copulae is given by Sklar's Theorem.

Theorem 2.1 (Sklar (1959, 1996))**.** *Let $F : \mathbb{R}^m \to [0, 1]$ be an m-dimensional distribution function, with univariate margins $F_1, \ldots, F_m : \mathbb{R} \to [0, 1]$. Then there exists an m-dimensional Copula $C : [0, 1]^m \to [0, 1]$, such that*

$$\forall (x_1, \ldots, x_m)^\top \in \mathbb{R}^m : F(x_1, \ldots, x_m) = C(F_1(x_1), \ldots, F_m(x_m)).$$

Moreover, if the marginal distribution functions are continuous, then the copula C is uniquely determined.

Example 2.1.

a) *Gaussian copula*

The function $C_\Sigma : [0,1]^m \to [0,1]$*, with*

$$C_\Sigma : (u_1, \ldots, u_m)^\top \mapsto \Phi_m(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_m); \Sigma)$$

is an m*-dimensional Gaussian copula, where* $\Phi_m(\cdot; \Sigma)$ *denotes the distribution function of the* m*-variate normal distribution* $\mathcal{N}_m(0, \Sigma)$ *with covariance and correlation matrix equal to* Σ*, and* Φ^{-1} *is the quantile function of the univariate standard normal distribution.*

According to our general setup developed in Section 2.1 we interpret the correlation matrix as the nuisance parameter η*, i. e., we consider in the most general case the space* $\Xi = \{\Sigma \in [-1,1]^{m \times m} : \Sigma \textit{ symmetric and positive definite correlation matrix}\}$*. Note that this parameter space has dimension* $\dim(\Xi) = \mathcal{O}(m^2)$ *which usually leads to the curse of dimensionality. Many relevant applications have to deal with this problem, namely that the sample size* n *is smaller than the dimensionality of the parameter space. This requires regularizing assumptions regarding the structure of* $\Sigma \in \Xi$*, for example* $AR(p)$ *structure, Toeplitz structure, or factor structure. In Section 4, we will assume an* $AR(1)$ *structure.*

b) *General elliptical copulae*

The function $C_{\Sigma,h} : [0,1]^m \to [0,1]$*, with*

$$C_{\Sigma,h} : (u_1, \ldots, u_m)^\top \mapsto E_m(E^{-1}(u_1), \ldots, E^{-1}(u_m))$$

is an m*-dimensional elliptical copula, where* E_m *is the distribution function of an elliptical distribution with correlation matrix* Σ *and density function* $e_m(x) \propto h(x^\mathsf{T} \Sigma x),\ x \in \mathbb{R}^m$*. The function* $h : [0, \infty) \to [0, \infty)$ *is called the density generator and should satisfy* $\int_0^\infty r^{m/2-1} h(r) dr < \infty$*. The symbol* E^{-1} *denotes the quantile function of the respective marginal distribution. Members of the family of elliptical distributions besides the multivariate Gaussian distribution are, among others, the multivariate Student's* t*-distribution and the multivariate Laplace (double exponential) distribution.*

c) *Archimedean copulae*

Let $\psi : [0, \infty) \to [0,1]$ *be a nonincreasing, continuous* m*-altering function, meaning that* $(-1)^k \psi^{(k)}(x) \geq 0$ *for all* $x \in [0, \infty)$ *and* $k = 0, \ldots, m$*, with* $\psi(0) = 1$*,* $\lim_{x \to \infty} \psi(x) = 0$*, and assume that* ψ *is strictly decreasing on* $[0, \psi^{-1}(0))$*. Then* $C_\psi : [0,1]^m \to [0,1]$ *with*

$$C_\psi(\mathbf{u}) = \psi\left(\sum_{j=1}^m \psi^{-1}(u_j)\right),\ \mathbf{u} = (u_1, \ldots, u_m)^\top \in [0,1]^m,$$

is called an Archimedean copula with generator ψ*; cf. McNeil and Nešlehová (2009).*

Due to the analytic properties of ψ*, its inverse exists on* $[0, \psi^{-1}(0))$ *and it is defined by*

$$\psi^{-1}(u) := \inf\{x \in [0, \infty) : \psi(x) \leq u\}.$$

If ψ is m-altering for all $m \in \mathbb{N}$, then ψ is called completely monotone. In that case ψ may be considered as the Laplace-Stieltjes transform of a positive random variable according to Bernstein's Theorem, cf. Bernstein (1929), which is useful for a stochastic representation of a multivariate random vector following the Archimedean copula with generator ψ. For more details, in particular about the generation of (pseudo-) random samples from Archimedean copulae we refer to Hofert (2012). Finally, the class of Archimedean copulae possesses the exchangeability property, that is, if $\mathbf{U} \sim C_\psi$, then each subvector of $\mathbf{U}$ follows the same type of copula.

Within our statistical setup the most general space is

$$\Xi = \{\psi : [0, \infty) \to [0, 1] : \psi \text{ is } m\text{-monotone}\},$$

where we regard the function $\psi \in \Xi$ as an infinite-dimensional nuisance parameter. For regularization purposes, it is useful to restrict attention to parametric sub-families of generator functions and the corresponding parametric families of Archimedean copulae. A comprehensive list of such parametric Archimedean copula families is provided by Nelsen (2006), pp.116-119.

In Section 4 we will be considered with Gaussian and Archimedean copulae. While Gaussian copulae naturally arise in connection with multivariate central limit theorems, certain Archimedean copulae play an important role in connection with other types of limit laws. In this sense, our present work generalizes the results by Hothorn et al. (2008) which are only applicable under (asymptotic) normality. For an illustration, let us consider the Gumbel-Hougaard family which is defined by the generator $\psi_\eta : x \in [0, \infty) \mapsto \exp\left(-x^{1/\eta}\right),\ \eta \geq 1$. The Gumbel-Hougaard copula for a parameter $\eta \geq 1$ is consequently given by

$$G_\eta : \mathbf{u} = (u_1, \dots, u_m)^\top \in [0, 1]^m \mapsto \exp\left(-\left[\sum_{j=1}^{m}(-\log(u_j))^\eta\right]^{\frac{1}{\eta}}\right). \tag{4}$$

The Gumbel-Hougaard family is also characterized (Genest and Rivest (1989)) by the fact that it is the only family of Archimedean copulae which are max-stable, that is

$$\forall k \in \mathbb{N} : \ G_\eta(u_1, \dots, u_m)^k = G_\eta(u_1^k, \dots, u_m^k), \quad (u_1, \dots, u_m)^\top \in [0, 1]^m. \tag{5}$$

It is well known that the class of max-stable distributions coincides with the class of extreme value distributions. Thus, any Gumbel-Hougaard copula G_η can arise as the weak limit of multivariate distribution functions in the sense that, for some copula C_0,

$$\lim_{n\to\infty} \left(C_0(u_1^{1/n}, \dots, u_m^{1/n})\right)^n = G_\eta(u_1, \dots, u_m),\ (u_1, \dots, u_m)^\top \in [0, 1]^m. \tag{6}$$

Such copulae C_0 are said to be in the domain of attraction of G_η; cf. Gudendorf and Segers (2010). A sufficient condition for Archimedean copulae C_ψ which are generated by $\psi : [0, \infty) \to [0, 1]$ to be in the max domain of attraction of a Gumbel-Hougaard copula is given in the following lemma.

Lemma 2.1 (Gudendorf and Segers (2010)). *Denote by $\phi = \psi^{-1}$ the inverse function of the copula generator of the Archimedean copula C_ψ. If*

$$-\lim_{t\to 0}\frac{t\phi'(1-t)}{\phi(1-t)} = \eta \in [1,\infty), \tag{7}$$

then C_ψ is in the max domain of attraction of the Gumbel-Hougaard copula G_η.

2.3 Multiple testing in terms of copulae

Under our general Assumption 2.1 and following Dickhaus and Gierl (2013), we can upper-bound the FWER of the multiple test $\varphi = \varphi(\mathbf{T}, \mathbf{c})$ by

$$\begin{aligned}
\text{FWER}_{\vartheta,\eta}(\varphi) &\leq \text{FWER}^*_\eta(\varphi) = \mathbb{P}^*_\eta\left(\bigcup_{j=1}^m \{T_j > c_j\}\right)\\
&= 1 - \mathbb{P}^*_\eta(T_1 \leq c_1, \ldots, T_m \leq c_m),\\
&= 1 - C_\eta(F_1(c_1), \ldots, F_m(c_m)),
\end{aligned}$$

where F_j is the marginal cdf of T_j under ϑ^*, $1 \leq j \leq m$, and $C_\eta : [0,1]^m \to [0,1]$ is some m-dimensional copula indexed by the nuisance parameter $\eta \in \Xi$. Recall that Assumption 2.1.(i) implies that the dependency structure among the test statistics is entirely captured by the parameter $\eta \in \Xi$, regardless of $\vartheta \in \Theta$. The calculation above shows that the calibration of the vector $\mathbf{c}$ of critical values for FWER control of φ can be performed by means of level sets of C_η. If we denote by $u_j(\eta) = F_j(c_j(\eta)) = 1 - \alpha_j \in [0,1]$, where α_j is the local significance level from (3), then each choice from the set $C_\eta^{-1}(1-\alpha) = \{\mathbf{u} \in [0,1]^m : C_\eta(\mathbf{u}) = 1-\alpha\}$ yields valid critical values $\mathbf{c}$. Since $C_\eta^{-1}(1-\alpha)$ is an $(m-1)$-dimensional submanifold in $[0,1]^m$, one can choose a valid set of critical values by weighting hypotheses for importance, see, e. g., Roeder and Wasserman (2009). If all m hypotheses are equally important, one should simply choose $u_1 = \ldots = u_m$, yielding a unique solution for $\mathbf{c}$. In practice, as mentioned before, a null hypothesis H_j is rejected if $p_j \leq \alpha_j = 1 - u_j(\eta)$, for $1 \leq j \leq m$. Figure 1 illustrates the interrelation of global significance level, local significance level, and the copula parameter η graphically. If H_j is a composite null hypothesis, we refer to the more general definition of α_j as given in (3).

3 Empirical calibration of multiple tests

With slight abuse of notation and for ease of presentation, we let $\eta \in \Xi$ in this section denote the copula parameter of the vector $\mathbf{T}$ of test statistics rather than that of the original data, although these two quantities do not necessarily coincide.

3.1 Estimation of copula parameters

Assumption 2.1.(ii) ensures that the marginals which have to be used for FWER calibration of φ are known (because this calibration is performed under the intersection hypothesis) and,

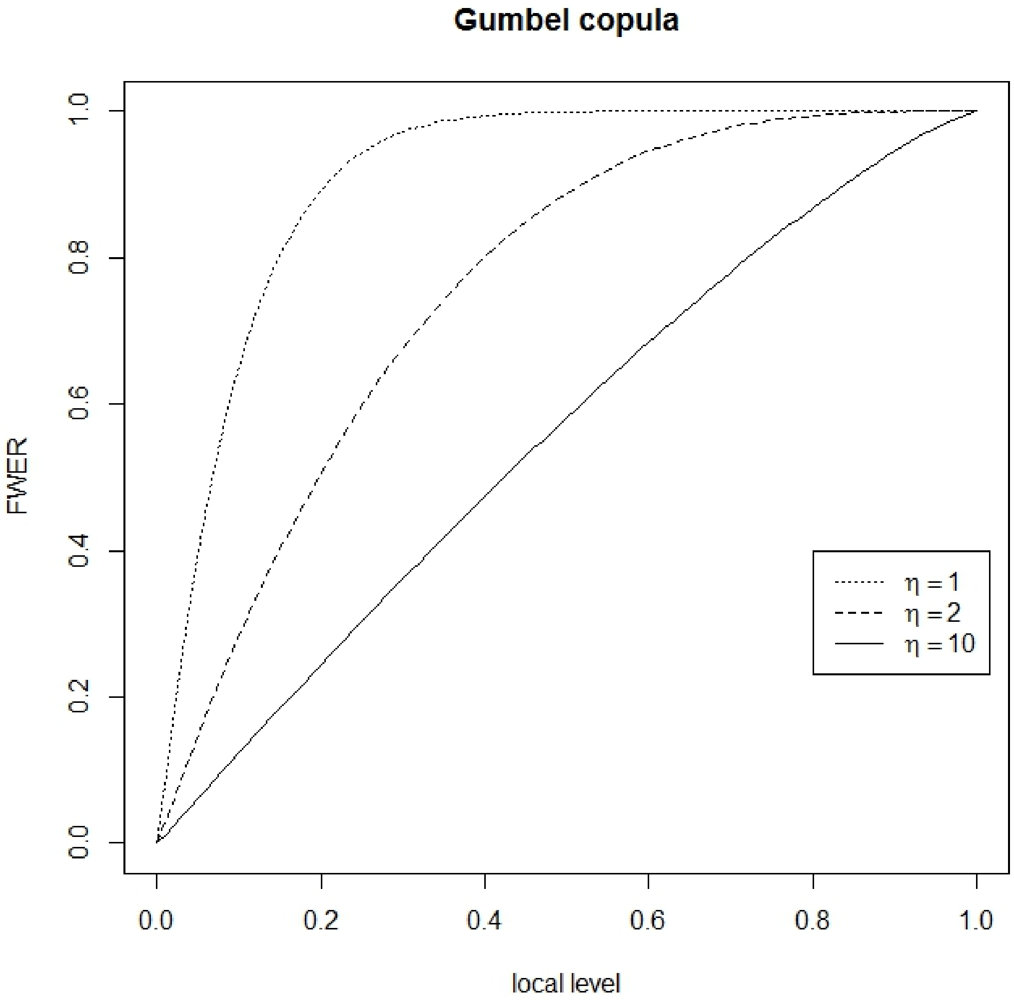

Figure 1: The upper FWER bound for $m = 12$ marginal tests as a function of its local significance levels, i. e., $\text{FWER}^*_\eta(u) = 1 - C_\eta(1-u,\dots,1-u),\ u \in [0,1]$, in the case that the dependence among test statistics is modeled by a Gumbel-Hougaard copula with varying parameter η. For a given global significance level α one can determine the corresponding equicoordinate local significance level on the abscissa. The dotted line represents independence, corresponding to a Šidák (1967) correction. For $\eta \to \infty$ the curve approaches the identity, meaning that no correction for multiplicity is necessary, since “effectively“ only one single test is performed.

consequently, only the dependence structure has to be inferred. This dependency structure is in turn fully defined by the copula parameter η. Several methods exist in the literature which can be used for estimating η. The most widely applied ones are maximum likelihood and the method of moments.

3.1.1 Maximum likelihood estimation

Maximum likelihood estimation is a well-established estimation technique in parametric models. In case of elliptical copulae the procedure is discussed in detail by Gupta et al. (2013). For Archimedean copulae the density can be calculated only if the generating function ψ is differentiable up to order $m-1$ and $\psi^{(m-1)}$ is absolutely continuous (McNeil and Nešlehová (2009)). In that case the density of C_η is given by

$$c_\eta(\mathbf{u}) = \frac{\partial^m}{\partial u_1 \dots \partial u_m} C_\eta(\mathbf{u}) = \frac{\psi^{(m)}\left(\sum_{j=1}^m \psi^{-1}(u_j)\right)}{\prod_{j=1}^m \psi'(\psi^{-1}(u_j))}, \ \mathbf{u} \in [0,1]^m.$$

Given a sample of i.i.d. random vectors $\mathbf{U}_1, \dots, \mathbf{U}_n \sim C_\eta$, the estimator of η is found by maximizing the log-likelihood function

$$\ell(\eta, \mathbf{U}_1, \dots, \mathbf{U}_n) = \sum_{i=1}^n \ell(\eta, \mathbf{U}_i) = \sum_{i=1}^n \log(c_\eta(\mathbf{U}_i)) \tag{8}$$

with respect to $\eta \in \Xi$, i. e.,

$$\hat{\eta}_{n,\mathrm{ML}} = \operatorname*{argsup}_{\eta \in \Xi} \ell(\eta, \mathbf{U}_1, \dots, \mathbf{U}_n).$$

The derivation of analytical expressions for (8) in case of five well known families of Archimedean copulae is given by Hofert et al. (2012). Moreover, the authors mention that under usual regularity assumptions (such as finiteness of the Fisher information) the maximum likelihood estimator is asymptotically efficient. Finally, it is consistent and asymptotically normally distributed. More specifically, it holds that, with $p = \dim(\eta) \in \mathbb{N}$,

$$\sqrt{n}(\hat{\eta}_{n,\mathrm{ML}} - \eta) \xrightarrow{d} \mathcal{N}_p(0, \mathcal{I}(\eta)^{-1}) \quad \text{with} \quad \mathcal{I}(\eta) = \mathbb{E}_\eta\left[\nabla\ell(\eta, \mathbf{U})\nabla\ell(\eta, \mathbf{U})^\mathsf{T}\right].$$

3.1.2 Method of moments

A further, and in most cases simpler, method to estimate the parameters of an Archimedean copula follows from the generalized method of moments; see, e. g., Hansen (1982). Since maximum likelihood estimation is typically done by numerical optimization (for instance, employing the Newton-Raphson-algorithm), the method of moment estimates often serve as initial values.

The copula functions are naturally connected to measures of dependence, such as Pearson's product moment correlation or Spearman's rank correlation. Especially, for Archimedean copulae there are handy relations between the copula generating function ψ and the concordance measures which are defined as follows.

Definition 3.1. *Let $m \geq 2$ and $\mathbf{X} = (X_1, \ldots, X_m)^\top$ denote a random vector defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^m$.*

(i) Kendall's τ (Kendall (1938)):

For each pair (X_j, X_k), $1 \leq j < k \leq m$, denote by (X_j^*, X_k^*) an independent copy of (X_j, X_k) and define

$$\begin{aligned}\tau_{2,jk} &= \mathbb{P}[(X_j - X_j^*)(X_k - X_k^*) > 0] - \mathbb{P}[(X_j - X_j^*)(X_k - X_k^*) < 0] \\ &= 4\mathbb{P}(X_j \leq X_j^*, X_k \leq X_k^*) - 1.\end{aligned}$$

(ii) Coefficient of agreement (Kendall and Babington Smith (1940)):

$$\mathcal{T}_m := \binom{m}{2}^{-1} \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} \tau_{2,jk}, \ m \geq 2.$$

(iii) Multivariate Concordance (Joe (1990)):

$$\tau_m := \frac{2^m \mathbb{P}(\mathbf{X} \leq \mathbf{X}^*) - 1}{2^{m-1} - 1}, \ m \geq 2,$$

where $\mathbf{X}^*$ is an independent copy of $\mathbf{X}$.

In Definition 3.1.(i) Kendall's τ is given in the bivariate case, whereas two extensions to the multivariate case are presented in parts (ii) and (iii). The relationship of Pearson's product moment correlation as well as of Kendall's τ to the copula function is provided in the following lemma.

Lemma 3.1. *Let $X \sim F_X$ and $Y \sim F_Y$ be two random variables on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let the joint distribution of X and Y be associated with a copula C, i. e., $\mathbb{P}(X \leq x, Y \leq y) = C(F_X(x), F_Y(y))$.*

(a) The covariance of X and Y is given by

$$\begin{aligned}\sigma_{X,Y} &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \int_{\mathbb{R}^2} C(F_X(x), F_Y(y)) - F_X(x)F_Y(y)\, dx\, dy.\end{aligned}$$

(b) Let (X^*, Y^*) be an independent copy of (X, Y). Then

$$\tau_2 = 4\mathbb{P}(X \leq X^*, Y \leq Y^*) - 1 = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1\,.$$

Proof. The result of part a) goes back to the work of Höffding (1940), whereas the second statement can be found, for example, as Theorem 5.1.1 of Nelsen (2006). ■

It is remarkable to note that Kendall's τ is independent of the marginal distribution functions F_X and F_Y. This property makes this coefficient very useful in estimating the copula parameter, especially in the one-dimensional case. In the case of an Archimedean copula, Kendall's τ can be expressed in terms of the copula generating function ψ, or its parameter η, respectively. Namely, it holds that (see, e. g., Section 5.1 in Nelsen (2006))

$$\tau_2 = \tau_2(\psi) = \mathcal{T}_2 = 1 - 4\int_0^{\psi^{-1}(0)} t(\psi'(t))^2\,dt. \tag{9}$$

In some special cases, the right-hand side of (9) can further be simplified and expressed in terms of the copula parameter η. For instance, in case of the Gumbel-Hougaard family, we get $\tau_2 = \tau_2(\eta) = (\eta - 1)/\eta$. Unfortunately, there are only rare cases where closed form expressions for τ_m are available in terms of η for $m > 2$. Some formulas are provided in Genest et al. (2011). It is also worth noticing that, by the exchangeability property of Archimedean copulae, $\mathcal{T}_m$ is equal to $\tau_{2,12} = \ldots = \tau_{2,(m-1)m}$.

Next, we specify the sample counterparts of the population concordance measures given in Definition 3.1. Let $\mathbf{X}_1, \ldots, \mathbf{X}_n \sim \mathbf{X}$ be an i.i.d. sample of random vectors in $\mathbb{R}^m$. Then, with $\mathbf{X}_i = (X_{i1}, \ldots, X_{im})^\top$ for $1 \le i \le n$, the sample estimators of $\mathcal{T}_m$ or τ_m, namely $\hat{\mathcal{T}}_{m,n}$ and $\hat{\tau}_{m,n}$, are given by

$$\hat{\mathcal{T}}_{m,n} = \binom{m}{2}^{-1} \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} \left(4\binom{n}{2}^{-1} \sum_{i=1}^{n-1} \sum_{i'=i+1}^{n} \mathbf{1}(X_{ij} \le X_{i'j}, X_{ik} \le X_{i'k}) - 1 \right), \tag{10}$$

$$\hat{\tau}_{m,n} = \frac{1}{2^{m-1} - 1} \left(2^m \binom{n}{2}^{-1} \sum_{i=1}^{n-1} \sum_{i'=i+1}^{n} \mathbf{1}(\mathbf{X}_i \le \mathbf{X}_{i'}) - 1 \right). \tag{11}$$

Estimators of η are obtained by inverting $\mathcal{T}_m$ or τ_m, leading to

$$\hat{\eta}_{\tau_m,n} = \tau_m^{-1}(\hat{\tau}_{m,n}), \quad \hat{\eta}_{\mathcal{T}_m,n} = \mathcal{T}_m^{-1}(\hat{\mathcal{T}}_{m,n}).$$

Genest et al. (2011) compared the estimators (10) and (11) with each other for several Archimedean families, where it becomes obvious that $\hat{\eta}_{\tau_m,n}$ and $\hat{\eta}_{\mathcal{T}_m,n}$ perform virtually equivalently, up to a finite sample error. Moreover, both estimators are asymptotically normal, that is $\hat{\mathcal{T}}_{m,n} \overset{as.}{\sim} \mathcal{N}(\mathcal{T}_m, \sigma^2_{\mathcal{T}_m}/n)$ and $\hat{\tau}_{m,n} \overset{as.}{\sim} \mathcal{N}(\tau_m, \sigma^2_{\tau_m}/n)$. The expression for the asymptotic variance $\sigma^2_{\mathcal{T}_m} = \sigma^2_{\mathcal{T}_m,C}$ is provided in Proposition 4 of Genest et al. (2011). For Archimedean copulae this expression simplifies, due to exchangeability, to

$$\sigma^2_{\mathcal{T}_m,C_\eta} = 4\mathrm{Var}(C_\eta(U,V) + \bar{C}_\eta(U,V)). \tag{12}$$

In (12), U and V are uniformly distributed random variables with joint distribution determined by the copula C_η. The function $\bar{C}_\eta$ denotes the survival function of (U,V), that is $\bar{C}_\eta(u,v) = \mathbb{P}(U > u, V > v) = 1 - u - v + C_\eta(u,v)$. Hence, (12) can equivalently be written as

$$\sigma^2_{\mathcal{T}_m,C_\eta} = 4\left\{\mathrm{Var}(U+V) + 4\mathrm{Var}(C_\eta(U,V)) - 8\mathrm{Cov}(U, C_\eta(U,V))\right\}.$$

Further, if $\psi : [0, \infty) \to [0, 1]$ is the generator function of the Archimedean copula C_η, then the distribution function of the bivariate probability integral transform (BIPIT) $C_\eta(U, V)$ is given by

$$K_{C_\eta}(t) := \mathbb{P}(C_\eta(U, V) \leq t) = 1 - \psi^{-1}(t)\psi'(\psi^{-1}(t)), \tag{13}$$

see Genest and Rivest (1993); Nelsen (2006). The function K_{C_η} is known as Kendall's (distribution) function associated with the copula C_η and can be used for the derivation of the probability density function of $C_\eta(U, V)$ as well as in the computation of its mean and its variance. The summand $\text{Var}(U+V) = \text{Var}(U) + \text{Var}(V) + 2\text{Cov}(U, V)$ is calculated by applying Hoeffding's Lemma 3.1.(a), leading to

$$\text{Cov}(U, V) = \int_{[0,1]^2} C_\eta(u, v)\, du\, dv - 1/4, \text{ and } \text{Var}(U) = \text{Var}(V) = 1/12.$$

Finally, in order to compute $\text{Cov}(U, C_\eta(U, V))$, we suggest (if feasible) to apply the analytic representation in Corollary 4.3.5 of Nelsen (2006), or a numerical integration.

A slightly different procedure for the estimation of η was suggested by Fengler and Okhrin (2012). Let

$$g(\eta) = (\hat{\tau}_{2,12} - \tau_2(\eta), \ldots, \hat{\tau}_{2,(m-1)m} - \tau_2(\eta)) \in \mathbb{R}^{m(m-1)/2},$$

and choose an appropriate weight matrix $W \in \mathbb{R}^{\binom{m}{2} \times \binom{m}{2}}$. Then the proposed estimator of η is given by

$$\hat{\eta}_W = \underset{\eta \in \Xi}{\text{arginf}}\, g(\eta)^\mathsf{T} W g(\eta).$$

This approach leads to a weighted least squares variant of the method of moments which is based on $\hat{\mathcal{T}}_m$.

3.1.3 Resampling under the intersection hypothesis

In some cases, the sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$ can be employed in order to infer the distribution (under H_0) of the vector $\mathbf{T}$ of test statistics in a direct manner. This holds true in particular if $\mathbf{T} = \mathbf{T}(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ preserves the dependency structure of the original data. Typical examples are the empirical means in case of stable distributions and the (component-wise) maxima for the Gumbel-Hougaard copula. However, there also exist cases in which the dependency structure among the components of $\mathbf{T}(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ cannot straightforwardly be deduced from that among the components of $\mathbf{X}$. In the latter case we recommend the application of a bootstrap procedure (Efron (1979)) to estimate the distribution of $\mathbf{T}$ and thus the corresponding quantities of interest. Especially, under Assumption 2.1.(ii), resampling can be performed under the intersection hypothesis, which is conceptually very simple; cf. Westfall and Young (1993). It is important to note that there is no 'default-variant' for all kinds of statistics to achieve consistency of bootstrap estimators, as counterexamples by Bickel and Freedman (1981) show. We will return to this point in Section 4.2.

3.2 Asymptotic behaviour of empirically calibrated multiple tests

In this section we consider the asymptotic behavior of the empirically calibrated multiple tests in the general case, meaning that no strong restrictions are imposed on the copula function which determines the dependence structure among the test statistics. Let $\Phi_{\mathbf{T}} = \{\varphi(\mathbf{T}, \mathbf{c}) | \mathbf{c} \in \mathbb{R}^m\}$ be a family of multiple testing procedures based on the vector of test statistics $\mathbf{T} : \mathcal{X} \to \mathbb{R}^m$. We assume that the dependence structure of $\mathbf{T}$ is determined by the copula function C_{η_0}, where η_0 denotes the true copula parameter.

Utilization of an estimate $\hat{\eta}$ for η_0 leads to the empirically calibrated critical values $\hat{\mathbf{c}} = \mathbf{c}(\hat{\eta})$, from which we obtain the calibrated test $\hat{\varphi} = \varphi(\mathbf{T}, \hat{\mathbf{c}}) \in \Phi_{\mathbf{T}}$. As explained in Section 2.3 we define local significance levels $1 - u_j(\eta)$ by $u_j(\eta) = F_j(c_j(\eta))$ for $j = 1, \ldots, m$. Since η_0 is unknown, we approximate these local significance levels by taking an element $\mathbf{u}(\hat{\eta})$ from the set $C_{\hat{\eta}}^{-1}(1 - \alpha)$ for a given global significance level α. We assume that the following regularity conditions hold true.

Assumption 3.1. *For each $\alpha \in (0, 1)$, $g_\alpha : \eta \in \Xi \mapsto C_\eta^{-1}(1 - \alpha) \in [0, \alpha]^m$ is a well-defined and continuously differentiable function. Furthermore, the composition $C_{\eta_0} \circ g_\alpha : \eta \in \Xi \mapsto C_{\eta_0}(u_1(\eta), \ldots, u_m(\eta))$ is also continuously differentiable.*

Under Assumption 3.1, we may regard FWER$^*_{\eta_0}(\hat{\varphi})$ as a derived parameter of the dependency structure of $\mathbf{T}$. Our main theorem shows how the uncertainty about the value of η propagates itself into uncertainty about the actual (realized) FWER of the calibrated test $\hat{\varphi}$.

Theorem 3.1. *Let $(\mathcal{X}, \mathcal{F}, \mathcal{P})$ be a statistical model and let $\mathcal{H} = \{H_1, \ldots, H_m\}$ be a collection of hypotheses with non-empty intersection hypothesis H_0. Assume that the joint distribution of $\mathbf{T}$ is given by the copula $C_{\eta_0} \in \{C_\eta | \eta \in \Xi \subseteq \mathbb{R}^p\}$, $p \in \mathbb{N}$, and that Assumption 2.1 is fulfilled. Further, suppose that $\hat{\eta}_n : \mathcal{X} \to \Xi$ is an asymptotically normally distributed estimator of η_0, i. e., that there exists a positive definite symmetric matrix Σ_0 such that*

$$\sqrt{n}(\hat{\eta}_n - \eta_0) \xrightarrow{d} \mathcal{N}_p(0, \Sigma_0) \quad \text{as} \quad n \to \infty.$$

Let $\alpha \in (0, 1)$ be a fixed global significance level. Then, under Assumption 3.1, the following assertions hold true.

a) Consistency

$$\forall \eta_0 \in \Xi : \textit{FWER}^*_{\eta_0}(\hat{\varphi}) = 1 - C_{\eta_0}(g_\alpha(\hat{\eta}_n)) \xrightarrow{\mathbb{P}^*_{\eta_0}} 1 - C_{\eta_0}(g_\alpha(\eta_0)) = \alpha.$$

b) Asymptotic Normality

$$\forall \eta_0 \in \Xi : \sqrt{n}\left(\textit{FWER}^*_{\eta_0}(\hat{\varphi}) - \alpha\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\eta_0}),$$

where $\sigma^2_{\eta_0} = \nabla C_{\eta_0}(g_\alpha(\eta_0))^\mathsf{T} \Sigma_0 \nabla C_{\eta_0}(g_\alpha(\eta_0))$.

c) Asymptotic Confidence Region

$$\forall \delta \in (0,1) : \forall \eta_0 \in \Xi : \lim_{n\to\infty} \mathbb{P}^*_{\eta_0}\left(\sqrt{n}\frac{\textit{FWER}^*_{\eta_0}(\hat{\varphi}) - \alpha}{\hat{\sigma}_n} \leq z_{1-\delta}\right) = 1 - \delta\,,$$

where $\hat{\sigma}_n^2 : \mathcal{X} \to (0,\infty)$ is a consistent estimator of the asymptotic variance $\sigma^2_{\eta_0}$. In this, $z_\beta = \Phi^{-1}(\beta)$ denotes the β-quantile of the standard normal distribution on $\mathbb{R}$.

Proof. Part a) follows from the Continuous Mapping Theorem; see, e. g., Theorem 1.14 in DasGupta (2008). Part b) is an application of the Δ-method; see, e. g., Theorem 3.7 in DasGupta (2008). Part c) is a consequence of part b) with an additional application of Slutzky's Lemma; see, e. g., Theorem 1.5 in DasGupta (2008). ■

4 Examples

4.1 Multiple two-sided Z-tests

As in Section 3, we let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ denote a sample of i.i.d. random vectors with values in $\mathbb{R}^m$, where $\mathbf{X}_1$ is distributed as $\mathbf{X} = (X_1, ..., X_m)^\top$. Let component-wise expectations be given by $\mu_j = \mathbb{E}[X_j],\ j = 1, \ldots, m$. We assume that the marginal variances $\sigma_j^2 = \text{Var}(X_j)$ are known and, w. l. o. g., equal to 1. For a fixed vector $\mu^* = (\mu_1^*, \ldots, \mu_m^*)^\top$, we consider the family of hypotheses

$$H_j = \{\mu_j = \mu_j^*\} \text{ versus } K_j = \{\mu_j \neq \mu_j^*\}, 1 \leq j \leq m.$$

A suitable vector of test statistics $\mathbf{T}_n = (T_{1,n}, \ldots, T_{m,n})^\top$ is given by

$$T_{j,n} = T_{j,n}(\mathbf{X}_1, \ldots, \mathbf{X}_n) = |Z_{j,n}|,\ Z_{j,n} := \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_{ji} - \mu_j^*\right).$$

We consider the following two models for the distribution of $\mathbf{X}$.

Model 4.1.

(a) The random vector $\mathbf{X}$ follows an m-variate normal distribution, $\mathbf{X} \sim \mathcal{N}_m(\mu, \Sigma)$, with $(m \times m)$ covariance and correlation matrix Σ. This implies that, under the global hypothesis $H_0 = \cap_{j=1}^m H_j$, $\mathbf{Z}_n \sim \mathcal{N}_m(0, \Sigma)$, where $\mathbf{Z}_n = (Z_{1,n}, \ldots Z_{m,n})^\top$. Let $\hat{\Sigma}$ be a consistent estimator of Σ and define the empirically calibrated critical values $\hat{\mathbf{c}} = (\hat{c}_1, \ldots, \hat{c}_m)^\top = (\hat{c}_1(\hat{\Sigma}), \ldots, \hat{c}_m(\hat{\Sigma}))^\top$ as solutions of the equation

$$\mathbb{P}^*_{\hat{\Sigma}}(T_{1,n} \leq \hat{c}_1, \ldots, T_{m,n} \leq \hat{c}_m) = 1 - \alpha, \tag{14}$$

where $\mathbb{P}^*_{\hat{\Sigma}}$ refers to $\mathcal{N}_m(0, \hat{\Sigma})$. In practice the computation of the vector $\hat{\mathbf{c}}$ of two-sided normal quantiles can conveniently be performed by using the `R` *function* `qmvnorm` *from the package* `mvtnorm`*, cf. Genz and Bretz (2009).*

(b) *The random vector $\mathbf{X}$ is non-Gaussian with finite second moments. In this case, $\mathbf{Z}_n$ converges under H_0 in distribution to $\mathcal{N}_m(0, \Sigma)$ due to the multivariate central limit theorem. With a consistent estimate $\hat{\Sigma}$ of Σ, the critical values are again empirically calibrated from the Gaussian copula $C_{\hat{\Sigma}}$ as described in* (14).

It is noted that Assumption 2.1 is satisfied in both parts of Model 4.1 and that the calibration of $\hat{\mathbf{c}}$ is as in Hothorn et al. (2008).

For illustration, we choose the correlation matrix $\Sigma = (\Sigma_{ij})$ of AR(1)-structure, i. e., $\Sigma_{ij} = \rho^{|i-j|},\ i, j = 1, \ldots, m$ for $\rho \in (-1, 1)$. The parameter ρ corresponds to the nuisance parameter η in our general setup. We apply maximum likelihood estimation for ρ, where the log-likelihood function is given by

$$\ell(\rho, \mu; \mathbf{X}_1, \ldots, \mathbf{X}_n) = -\frac{mn}{2}\log(2\pi) - \frac{n}{2}\log(\text{det}(\Sigma)) - \sum_{i=1}^{n}(\mathbf{X}_i - \mu)^\top \Sigma^{-1}(\mathbf{X}_i - \mu). \quad (15)$$

The determinant in (15) equals det$(\Sigma) = (1-\rho^2)^{m-1}$, whereas the precision matrix is given by

$$\Sigma^{-1} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1 \end{pmatrix}.$$

Hence,

$$\begin{aligned} \ell(\rho, \mu; \mathbf{X}_1, \ldots, \mathbf{X}_n) \quad = \quad & -\frac{nm}{2}\log(2\pi) - \frac{n(m-1)}{2}\log(1-\rho^2) \\ & -\frac{n(m-1)}{2}\frac{S_1 - 2\rho S_p + \rho^2 S_2}{(1-\rho^2)}, \end{aligned}$$

where

$$S_p = \frac{1}{n(m-1)}\sum_{i=1}^{n}\sum_{j=1}^{m-1}(X_{j\,i} - \mu_j)(X_{j+1\,i} - \mu_{j+1}),$$

$$S_1 = \frac{1}{n(m-1)}\sum_{i=1}^{n}\sum_{j=1}^{m}(X_{j\,i} - \mu_j)^2, \text{ and } S_2 = \frac{1}{n(m-1)}\sum_{i=1}^{n}\sum_{j=2}^{m-1}(X_{j\,i} - \mu_j)^2.$$

Solving the normal equations

$$\frac{\partial}{\partial\rho}\ell(\rho, \mu; \mathbf{X}_1, \ldots, \mathbf{X}_n) = 0 \quad \text{and} \quad \frac{\partial}{\partial\mu}\ell(\rho, \mu; \mathbf{X}_1, \ldots, \mathbf{X}_n) = \mathbf{0}$$

leads to $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i$ as well as to the cubic root problem

$$\hat{\rho}^3 - S_p\hat{\rho}^2 - (1 - S_1 - S_2)\hat{\rho} - S_p = 0, \quad (16)$$

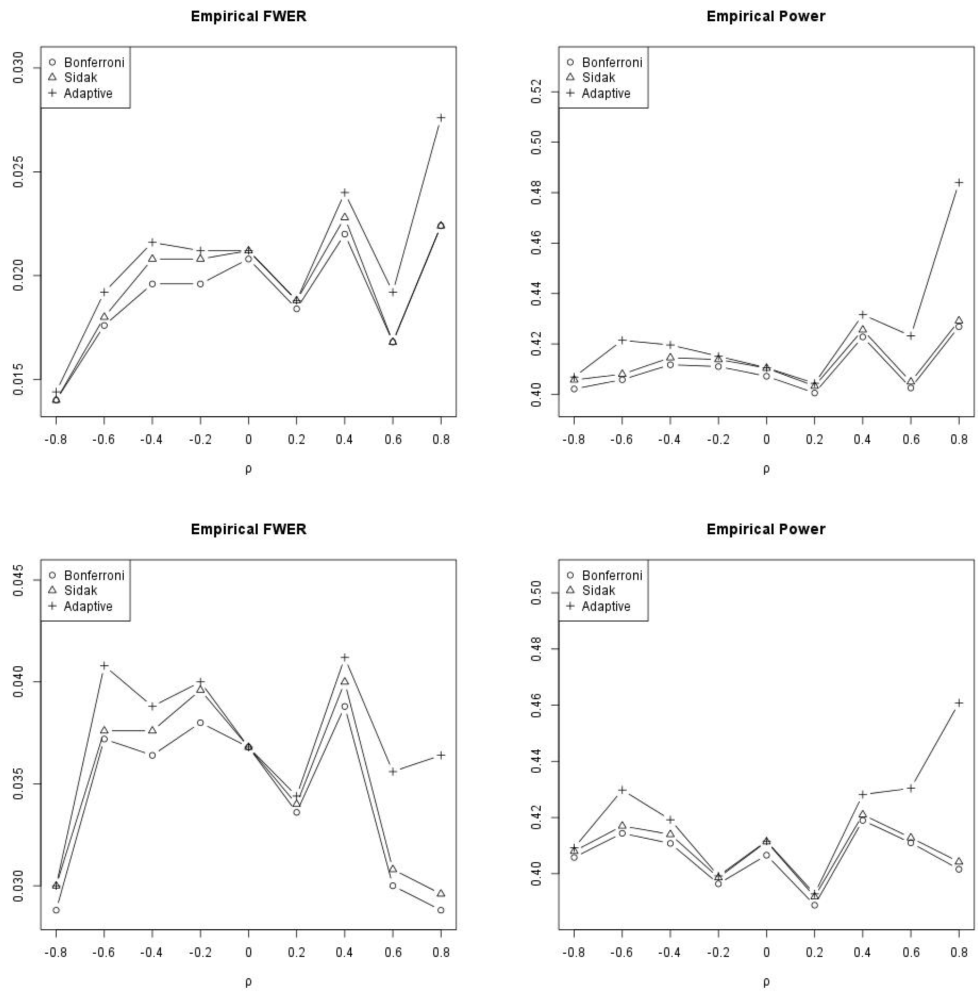

Figure 2: Empirical FWER and power as functions of $\rho \in \{-0.8, -0.6, ..., 0.6, 0.8\}$ for $m = 8$, $m_0 = 3$ (upper panel) and $m_0 = 6$ (lower panel). The datasets of size $n = 100$ were generated from a multivariate normal distribution. The results are based on $K = 2500$ independent repetitions.

where we substituted $\hat{\mu}$ for μ in S_1, S_2, and S_p. The solution of (16) can numerically be computed by using the R function `polyroot`. Asymptotic normality of the estimator $\hat{\rho}$ follows from general parametric likelihood theory.

Figures 2 to 9 display the results of a simulation study under Model 4.1. The target FWER level was set to $\alpha = 0.05$ in all simulations. Pseudo samples for Figures 2 to 5 follow the assumptions of part (a) of Model 4.1 (multivariate normal distributions), while data for Figures 6 to 9 were generated from a multivariate t-distribution with 9 degrees of freedom, constituting a special case of Model 4.1.(b). Every figure represents a different configuration of the parameters m, m_0, and ρ and is based on $K = 2500$ independent pseudo samples of size $n = 100$ (Model

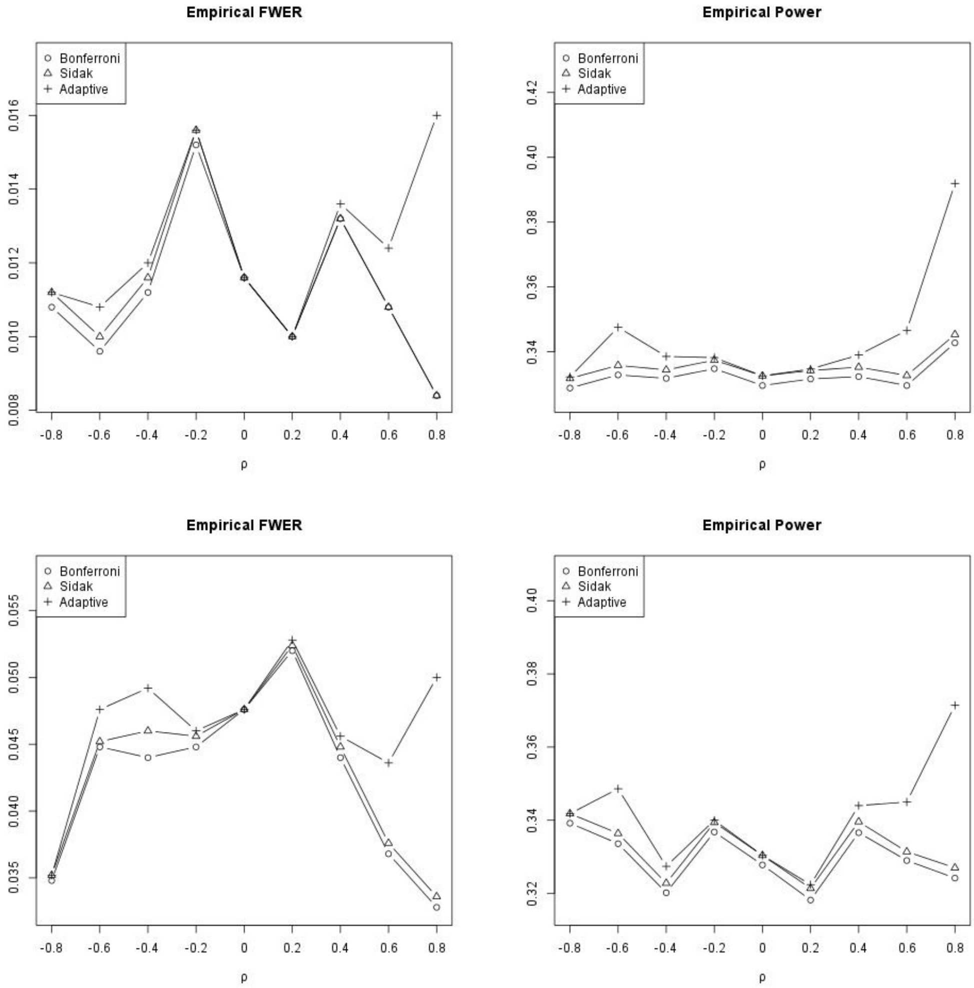

Figure 3: Empirical FWER and power as functions of $\rho \in \{-0.8, -0.6, ..., 0.6, 0.8\}$ for $m = 15$, $m_0 = 4$ (upper panel) and $m_0 = 13$ (lower panel). The datasets of size $n = 100$ were generated from a multivariate normal distribution. The results are based on $K = 2500$ independent repetitions.

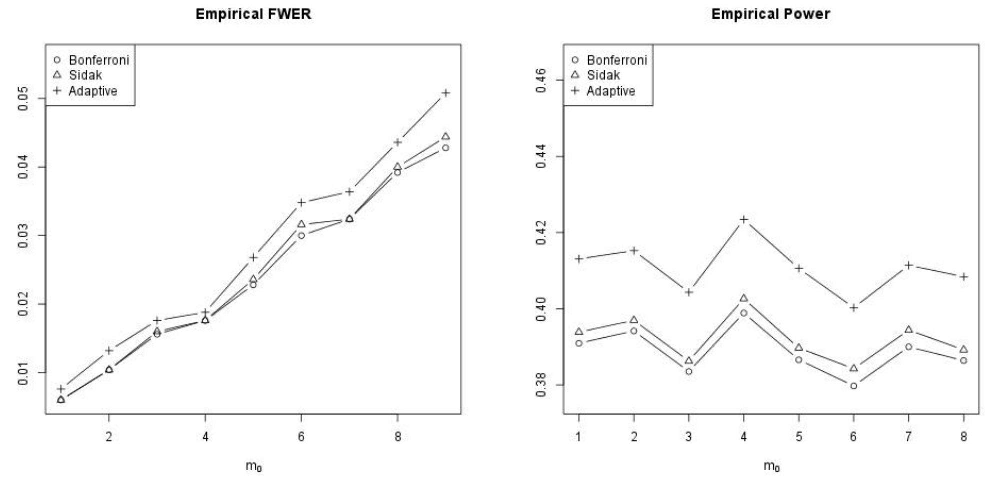

Figure 4: Empirical FWER and power as functions of $m_0 \in \{1, 2, ..., 9\}$ for $m = 9$ and $\rho = 0.6$. The datasets of size $n = 100$ were generated from a multivariate normal distribution. The results are based on $K = 2500$ independent repetitions.

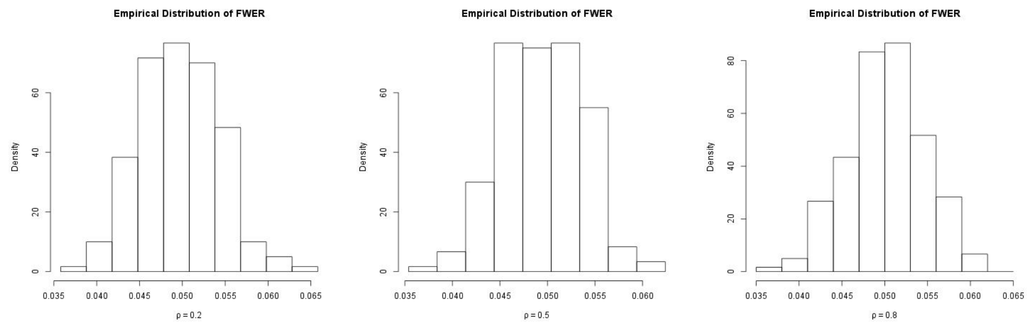

Figure 5: Histograms of the empirical FWER of empirically calibrated multiple tests for $m = m_0 = 8$, and $\rho \in \{0.2, 0.5, 0.8\}$. The datasets of size $n = 100$ were generated from a multivariate normal distribution. The results are based on 200 simulation runs with $K = 2500$ independent repetitions each.

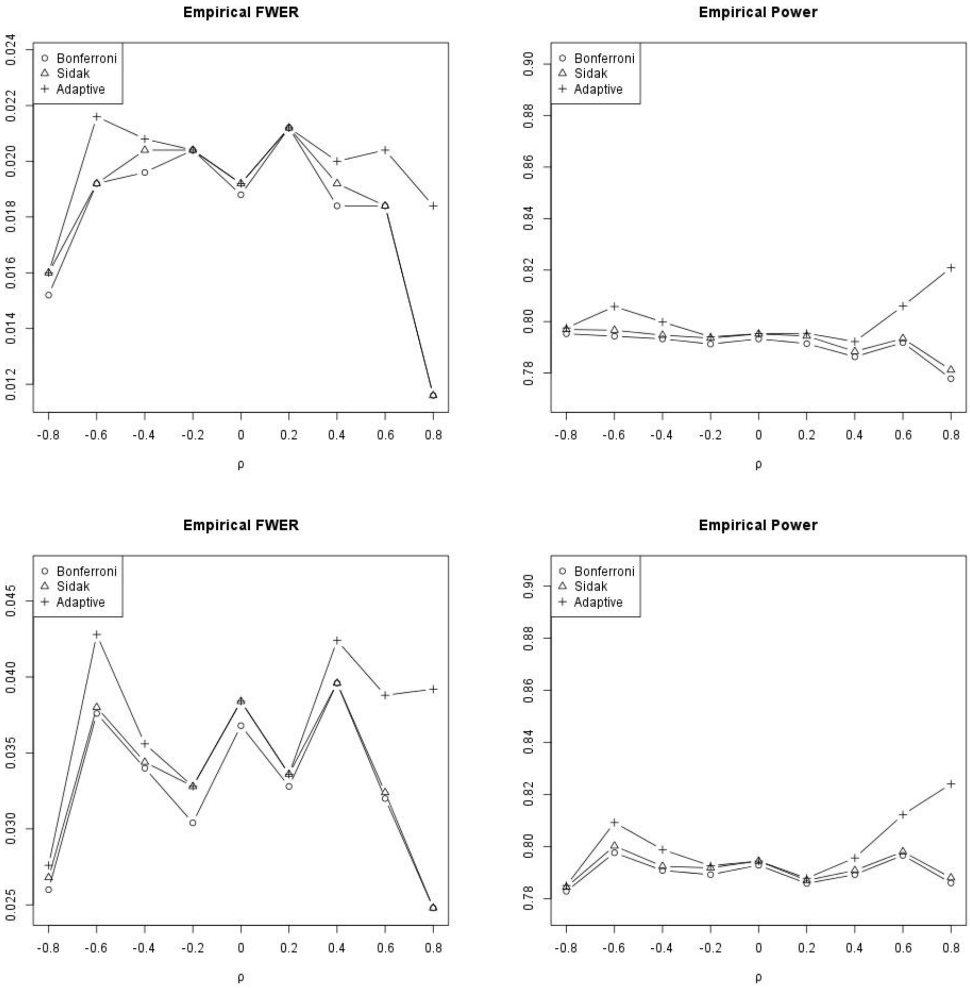

Figure 6: Empirical FWER and power as functions of $\rho \in \{-0.8, -0.6, ..., 0.6, 0.8\}$ for $m = 8$, $m_0 = 3$ (upper panel) and $m_0 = 6$ (lower panel). The datasets of size $n = 200$ were generated from a multivariate t-distribution with 9 degrees of freedom. The results are based on $K = 2500$ independent repetitions.

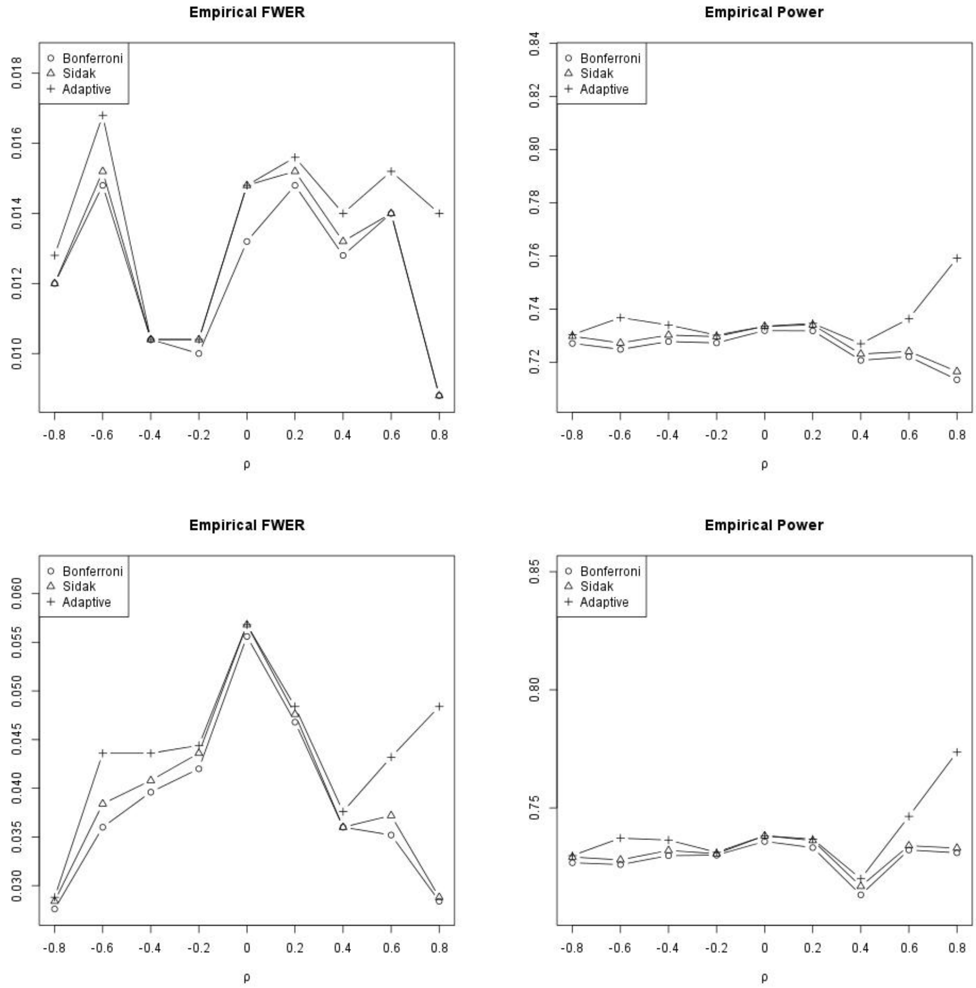

Figure 7: Empirical FWER and power as functions of $\rho \in \{-0.8, -0.6, ..., 0.6, 0.8\}$ for $m = 15$, $m_0 = 4$ (upper panel) and $m_0 = 13$ (lower panel). The datasets of size $n = 200$ were generated from a multivariate t-distribution with 9 degrees of freedom. The results are based on $K = 2500$ independent repetitions.

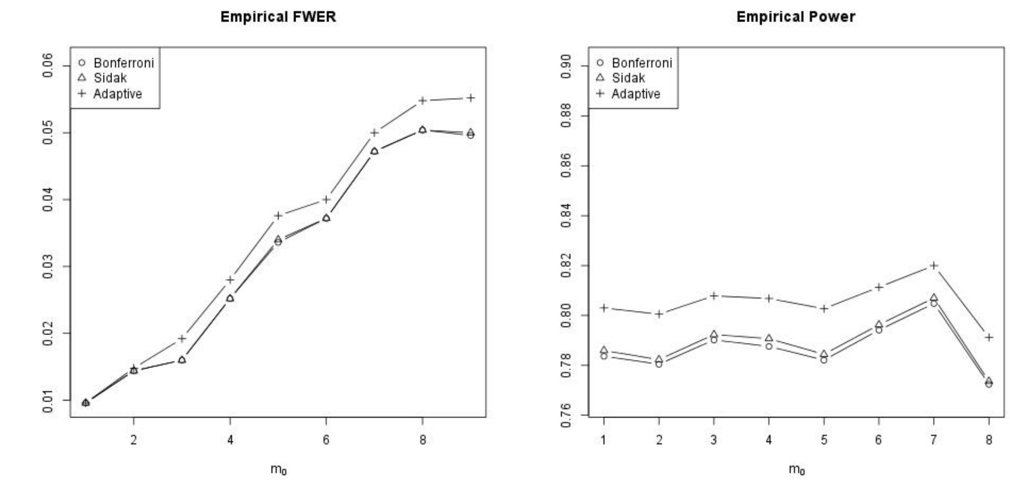

Figure 8: Empirical FWER and power as functions of $m_0 \in \{1, 2, ..., 9\}$ for $m = 9$ and $\rho = 0.6$. The datasets of size $n = 200$ were generated from a multivariate t-distribution with 9 degrees of freedom. The results are based on $K = 2500$ independent repetitions.

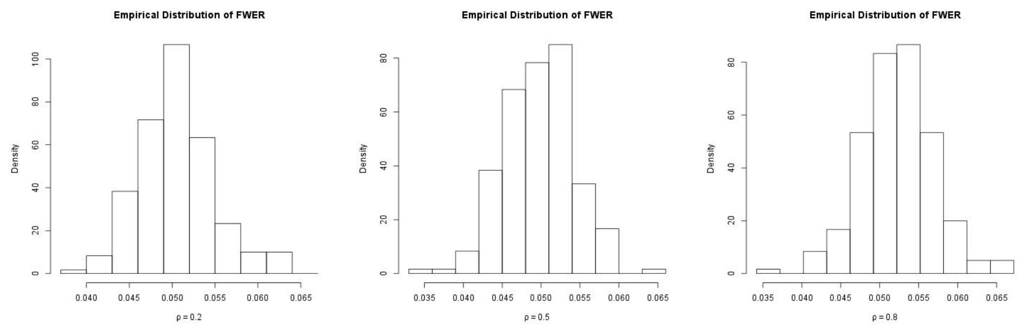

Figure 9: Histograms of the empirical FWER of empirically calibrated multiple tests for $m = m_0 = 8$, and $\rho \in \{0.2, 0.5, 0.8\}$. The datasets of size $n = 200$ were generated from a multivariate t-distribution with 9 degrees of freedom. The results are based on 200 simulation runs with $K = 2500$ independent repetitions each.

4.1.(a)) or $n = 200$ (Model 4.1.(b)) each. We put $\mu^* = \mathbf{0} \in \mathbb{R}^m$ and

$$\mu = (\mu_1, \ldots, \mu_m)^\top \quad \text{with} \quad \mu_j = \begin{cases} 0, & \text{for } j \leq m_0, \\ 0.25, & \text{otherwise,} \end{cases} \tag{17}$$

where $m = 8$ and $m_0 \in \{3, 6\}$ for Figures 2 and 6 as well as $m = 15$ and $m_0 \in \{4, 13\}$ for Figures 3 and 7. The correlation matrix Σ is assumed to follow an AR(1) structure as described above with $\rho \in \{-0.8, -0.6, ..., 0.6, 0.8\}$. In Figures 4 and 8 we analyze the behaviour of the multiple tests for different values of $m_0 \in \{1, 2, ..., 9\}$ when $m = 9$ and $\rho = 0.6$ are kept fixed. Finally, Figures 5 and 9 show histograms of the estimated realized FWER of the empirically calibrated multiple tests in case of $m = m_0 = 8$, and $\rho \in \{0.2, 0.5, 0.8\}$.

Empirical values of the FWER were calculated as the relative frequency of the occurrence of at least one type I error, i. e.,

$$\widehat{\text{FWER}} = K^{-1} \sum_{k=1}^{K} \mathbf{1}(\exists j \leq m_0 : \varphi_j^{(k)} = 1)\,,$$

where $\varphi^{(k)} = (\varphi_1^{(k)}, \ldots, \varphi_m^{(k)})^\top$ stands for the empirically calibrated multiple test in the k-th simulation run. Similarly, empirical power was computed as

$$\widehat{\text{power}} = K^{-1} \sum_{k=1}^{K} \left(m_1^{-1} \sum_{j=m_0+1}^{m} \mathbf{1}(\varphi_j^{(k)} = 1) \right), \quad m_1 = m - m_0.$$

Summarizing the results of the simulation study, we observe a very good performance of the empirically calibrated multiple tests. They exhaust the FWER level α better than the corresponding Bonferroni and Šidák corrections. The differences become more pronounced when the correlation coefficient ρ becomes larger. This result is expected since test statistics are positively correlated. Consequently, the multiple tests which are based on the Bonferroni and the Šidák corrections are unnecessarily conservative. In contrast, the empirically calibrated multiple tests allow us to capture the effect of high positive correlation among test statistics. Due to the decision structure of the considered multiple tests, better exhaustion of the FWER level directly translates into higher power, as can be verified in the corresponding figures. Finally, it is noted that these findings hold uniformly over all considered parameter settings and for both types of data distributions.

The histograms of the empirical FWER displayed in Figures 5 and 9 show that the distribution of the empirical FWER can be well approximated by a normal distribution. This observation has been confirmed by goodness-of-fit tests. Moreover, the empirical variances are very small and, as a result, the empirical FWER is well concentrated around α.

4.2 Multiple support tests

Again, we let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be a sample of i.i.d. random vectors with values in $[0, \infty)^m$. We assume that $\mathbf{X}_1$ is distributed as $\mathbf{X} = (X_1, \ldots, X_m)^\top$ with stochastic representations

$$X_j \stackrel{d}{=} \vartheta_j Z_j,\ \vartheta_j > 0,\ j = 1, \ldots, m,$$

where Z_j is a random variable taking values in $[0,1]$ with cdf $F_j : [0,1] \to [0,1]$. The parameter of interest in this problem is $\vartheta = (\vartheta_1, \ldots, \vartheta_m)^\top \in \Theta = [0,\infty)^m$. For each component $j = 1, \ldots, m$, we test the hypothesis

$$H_j : \{\vartheta_j \leq \vartheta_j^*\} \text{ versus } K_j : \{\vartheta_j > \vartheta_j^*\},$$

for a given vector $\vartheta^* = (\vartheta_1, \ldots, \vartheta_m)^\top \in [0,\infty)^m$ of hypothesized upper bounds for the supports (or right end-points of the distributions) of the X_j's. Now, suitable test statistics are given by the componentwise maxima of the observables, i.e.,

$$T_j = \max_{1 \leq i \leq n} X_{ij}/\vartheta_j^*, \; j = 1, \ldots, m.$$

It is easy to verify that Assumption 2.1.(ii) is fulfilled.

Lemma 4.1. *Let C_0 denote the copula of $\mathbf{X}$ (which is regarded as a nuisance parameter and thus, by Assumption 2.1, does not depend on $\vartheta \in \Theta$). Then, it holds that*

$$\mathbb{P}^*_{C_0}\left(\frac{X_1}{\vartheta_1^*} \leq x_1, \ldots, \frac{X_m}{\vartheta_m^*} \leq x_m\right) = C_0(F_1(x_1), \ldots, F_m(x_m)), \; x_1, \ldots, x_m \geq 0.$$

Because $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are independent we get that $T_j \sim F_j^n$ under ϑ^* for $j = 1, \ldots, m$ and therefore $F_j(T_j)^n$ is uniform on $[0,1]$ under ϑ^*. Hence, by Theorem 2.1, there exists a copula C such that

$$\mathbb{P}^*_{C_0}(T_1 \leq t_1, \ldots, T_m \leq t_m) = C(F_1(t_1)^n, \ldots, F_m(t_m)^n). \tag{18}$$

On the other hand, the vectors $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are i.i.d., which leads to

$$\begin{aligned} \mathbb{P}^*_{C_0}(T_1 \leq t_1, \ldots, T_m \leq t_m) &= \mathbb{P}^*_{C_0}\left(\bigcap_{i=1}^n \left\{\frac{X_{i1}}{\vartheta_1^*} \leq t_1, \ldots, \frac{X_{im}}{\vartheta_m^*} \leq t_m\right\}\right) \\ &= C_0(F_1(t_1), \ldots, F_m(t_m))^n. \end{aligned} \tag{19}$$

Combining (18) *and* (19) *we conclude that, for all $t_1, \ldots, t_m \geq 0$,*

$$C(F_1(t_1), \ldots, F_m(t_m)) = C_0(F_1(t_1)^{1/n}, \ldots, F_m(t_m)^{1/n})^n. \tag{20}$$

Based on Lemma 4.1, the critical values $c_j, \; j = 1, \ldots, m$, are chosen as

$$c_j = F_j^{-1}\left((1-\alpha_j)^{1/n}\right)$$

with local significance levels α_j which are obtained by an appropriate correction of the given global significance level α, depending on the copula C of $T_1, \ldots, T_m$. Unfortunately, the relationship between the copulae C and C_0 is highly non-trivial in general, meaning that the right-hand side of (20) has no analytically tractable form. However, it is tractable if C_0 belongs to the class of extreme value copulae. Hence, we consider two examples where we can exploit the fact that Gumbel-Hougaard copulae are extreme value copulae, see Section 2. It is noted that Assumption 2.1.(i) is satisfied by construction in both parts of Model 4.2.

Model 4.2.

(a) *Let $C_0 \in \{G_\eta : \eta \geq 1\}$, i. e., the copula C_0 belongs to the family of Gumbel-Hougaard copulae. Then, the copula of $\mathbf{T} = (T_1, \ldots, T_m)^\top$ coincides with the copula of $\mathbf{X}$, because of the max-stable property. In this case, η can be estimated by an appropriate method from Section 3.1. In our simulations described below, the coefficient of agreement $\mathcal{T}_m$ defined in Definition 3.1.(ii) was used.*

(b) *Assume that C_0 lies in the domain of attraction of a Gumbel-Hougaard copula G_η for some $\eta \geq 1$, where the nuisance parameter η is unknown.*

In order to estimate the copula parameter η under part (b) of Model 4.2, we propose to apply a bootstrap method with low resampling intensity. This is due to the well-known fact that the ordinary bootstrap fails for extreme value statistics. The following algorithm was used in our simulations.

Algorithm 4.1.

i) *Let $\tilde{\mathbf{X}}_1, \ldots, \tilde{\mathbf{X}}_n$ be given by $\tilde{X}_{ji} = X_{ji} / \max_{1 \leq \ell \leq n} X_{j\ell}$ for $1 \leq i \leq n$ and $1 \leq j \leq m$.*

ii) *Choose a number B of Monte Carlo repetitions and set $\nu := \lceil \sqrt{n} \rceil$ (the smallest integer larger than or equal to $\sqrt{n}$).*

iii) *For each $b = 1, \ldots, B$, draw a sample $\mathbf{X}_1^{\#b}, \ldots, \mathbf{X}_\nu^{\#b}$ of size ν with replacement out of $\tilde{\mathbf{X}}_1, \ldots, \tilde{\mathbf{X}}_n$ and calculate*

$$\mathbf{T}^{\#b} = (T_1^{\#b}, \ldots, T_m^{\#b})^\top = \max_{1 \leq \ell \leq \nu} \mathbf{X}_\ell^{\#b},$$

where the maximum is taken component-wise.

iv) *Using the vectors $\mathbf{T}^{\#1}, \ldots, \mathbf{T}^{\#B}$, determine*

$$\binom{m}{2} \tilde{\mathcal{T}}_{m,\mathrm{boot}} = \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} \left(4 \binom{B}{2}^{-1} \sum_{b=1}^{B-1} \sum_{b'=b+1}^{B} \mathbf{1}(T_j^{\#b} \leq T_j^{\#b'}, T_k^{\#b} \leq T_k^{\#b'}) - 1 \right).$$

v) *Finally, compute the estimate*

$$\hat{\eta}_{\mathrm{boot}} = 1/(1 - \tilde{\mathcal{T}}_{m,\mathrm{boot}}).$$

Theorem 4.1. *Algorithm 4.1 estimates the copula parameter η consistently as $\min\{n, B(n)\} \to \infty$.*

Proof. We apply Theorem 2 of Bickel et al. (1997). To this end, let $\mathcal{F}_0$ denote the set of all cdfs on $\mathbb{R}^m$ which are such that the copula of the component-wise maxima of i.i.d. observables is in the domain of attraction of some Gumbel-Hougaard copula. As before, we denote by $\mathbf{T} = \mathbf{T}_n$ such a vector of (properly scaled) component-wise maxima of the original data

$(\mathbf{X}_i)_{1\leq i\leq n}$, where it is assumed that $(\mathbf{X}_i)_{1\leq i\leq n}$ are i.i.d. with $\mathbf{X}_1 \sim F \in \mathcal{F}_0$. The functional $\theta_n(F)$ considered by Bickel et al. (1997) is in our case given by the coefficient of agreement of the components of $\mathbf{T}_n$. Finally, let $\mathcal{L}_n(F)$ denote the distribution of $\mathbf{T}_n$. By our general assumptions, we have weak convergence of $\mathcal{L}_n(F)$ to a non-degenerate limit law. Also, we have that $\theta_n(F)$ converges to some real constant $\theta(F)$, which is the deterministic transformation mapping the copula parameter η onto the coefficient of agreement. Now, it follows from Theorem 2 of Bickel et al. (1997) that the ideal $\nu \equiv \nu(n)$ out of n bootstrap $B_{\nu,n}(\hat{F}_n)$ (i. e., the functional in question evaluated at the uniform distribution over all possible subsets of $(\mathbf{X}_i)_{1\leq i\leq n}$ of cardinality $\nu(n)$, with replacement) estimates $\theta(F)$ consistently as $n \to \infty$, provided that $\nu(n) \to \infty$ such that $\nu(n)/n \to 0$, $n \to \infty$. The proof is completed by noticing that Algorithm 4.1 is a consistent approximation of $B_{\nu,n}(\hat{F}_n)$. ■

For an illustration, we choose C_0 from the Archimedean copula family which is generated by

$$\psi_\eta : t \mapsto \left(t^{1/\eta} + 1\right)^{-1},\ \eta \geq 1. \tag{21}$$

By (7) it holds that C_0 lies in the domain of attraction of the Gumbel copula G_η. The generation of random vectors following the chosen copula C_0 is performed by using the algorithm presented in Section 5.2 of McNeil and Nešlehová (2009).

In Figures 10 to 17 we present the results of a simulation study under Model 4.2. Similarly as in Section 4.1, $K = 2500$ independent samples of size $n = 150$ (Model 4.2.(a)) or size $n = 1600$ (Model 4.2.(b)) were generated from a Gumbel-Hougaard copula (Figures 10 to 13) and from an Archimedean copula defined by (21) (Figures 14 to 17), respectively. In all simulations, we chose F_j as the cdf of the Beta distribution with shape parameters 3 and 4, for all $1 \leq j \leq m$. We put $\vartheta^* = (2, ..., 2)^\top$ and

$$\vartheta = (\vartheta_1, \ldots, \vartheta_m)^\top \quad \text{with} \quad \vartheta_j = \begin{cases} 2, & \text{for } j \leq m_0, \\ 2.1, & \text{otherwise,} \end{cases} \tag{22}$$

for varying values of m and m_0. The copula parameter is chosen as $\eta \in \{1, 1.5, ..., 4.5, 5\}$ (Figures 10 and 11) and $\eta \in \{1, 1.25, ..., 2.75, 3\}$ (Figures 14 and 15), respectively. In Figures 12 and 16 we analyze the behaviour of the multiple tests for different values of m_0, with m and η kept fixed. Finally, Figures 13 and 17 show histograms of the estimated FWER under the global hypothesis H_0, for three different values of η each. Empirical FWER and empirical power are calculated as described in Section 4.1, and the target FWER level was set to $\alpha = 0.05$ in all simulations.

The results presented in Figures 10 to 17 are even stronger than the ones observed in Figures 2 to 9. The performance of the calibrated multiple tests is much better than that of the corresponding Bonferroni and Šidák tests. This result holds uniformly over all considered values of η. For instance, if $\eta = 4$ under Model 4.2.(a), the power is about two times larger for the empirically calibrated multiple test. Both the Bonferroni and the Šidák corrections lead to markedly undersized multiple tests as soon as η deviates from 1 ($\eta = 1$ corresponds to the case of independent test statistics). The obtained findings are almost identical for all considered values of m and m_0. Similarly to the results of Section 4.1, the histograms displayed in Figures 13 and 17 show that the distribution of the empirical FWER can be well approximated by a normal

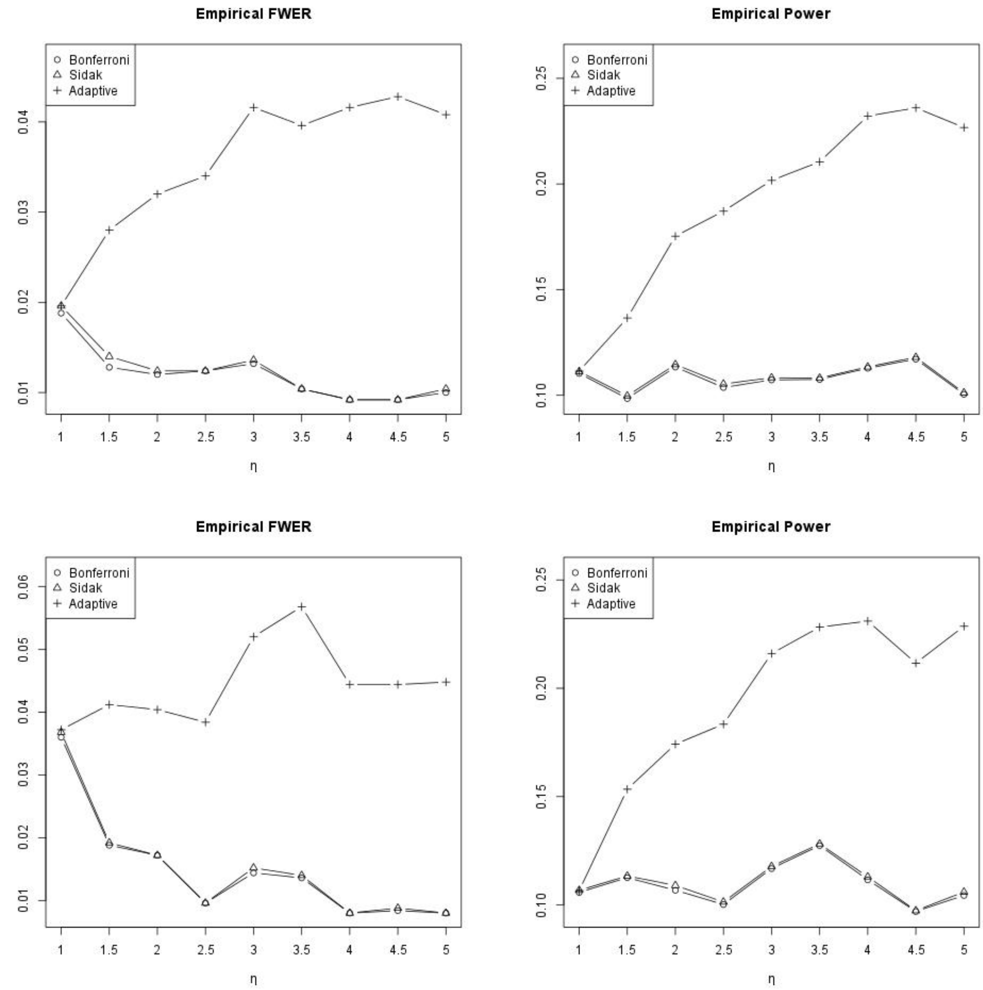

Figure 10: Empirical FWER and power as functions of $\eta \in \{1, 1.5, ..., 4.5, 5\}$ for $m = 8$, $m_0 = 3$ (upper panel) and $m_0 = 6$ (lower panel). The datasets of size $n = 150$ were generated from a Gumbel-Hougaard copula. The results are based on $K = 2500$ independent repetitions.

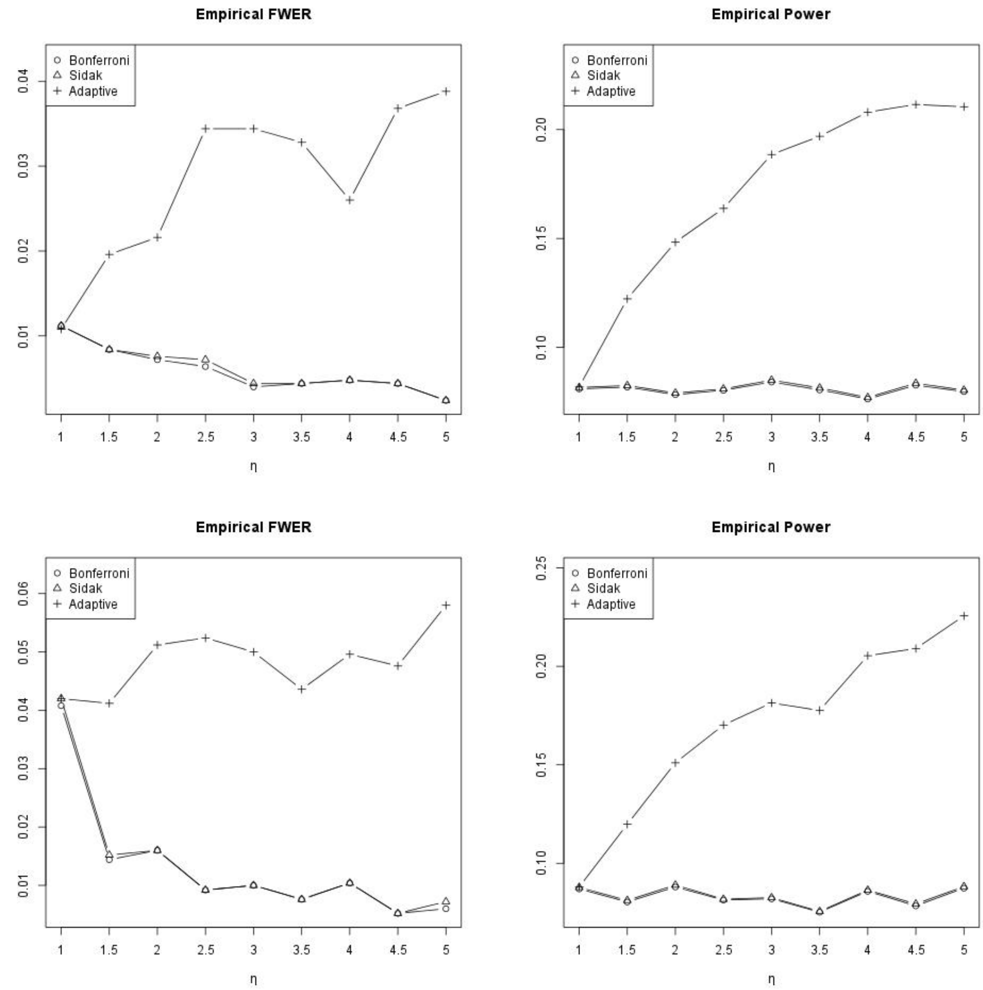

Figure 11: Empirical FWER and power as functions of $\eta \in \{1, 1.5, ..., 4.5, 5\}$ for $m = 15$, $m_0 = 4$ (upper panel) and $m_0 = 13$ (lower panel). The datasets of size $n = 150$ were generated from a Gumbel-Hougaard copula. The results are based on $K = 2500$ independent repetitions.

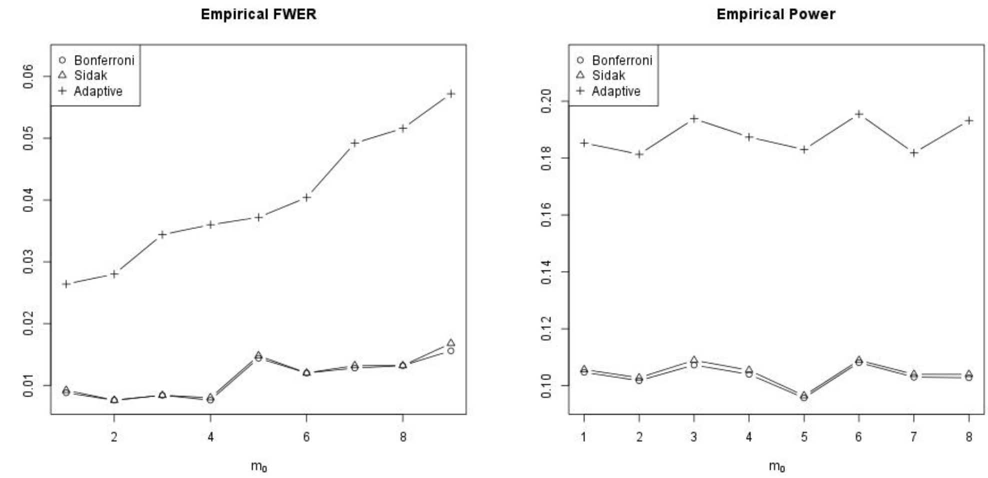

Figure 12: Empirical FWER and power as functions of $m_0 \in \{1, 2, ..., 9\}$ for $m = 9$ and $\eta = 2.5$. The datasets of size $n = 150$ were generated from a Gumbel-Hougaard copula. The results are based on $K = 2500$ independent repetitions.

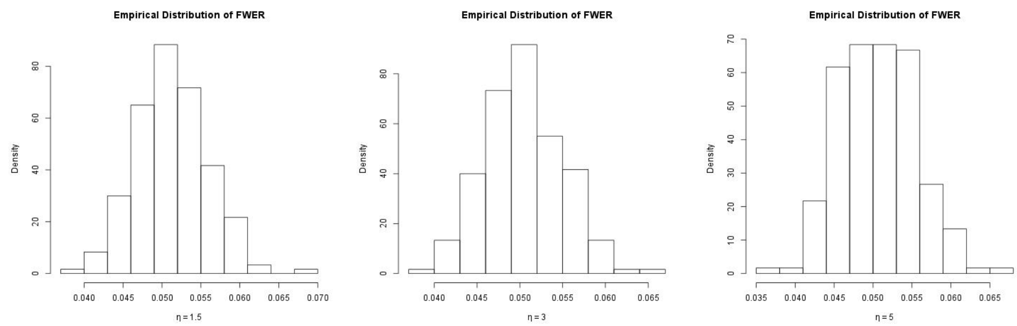

Figure 13: Histograms of the empirical FWER of empirically calibrated multiple tests for $m = m_0 = 8$ and $\eta \in \{1.5, 3.0, 5.0\}$. The datasets of size $n = 150$ were generated from a Gumbel-Hougaard copula. The results are based on 200 simulation runs with $K = 2500$ independent repetitions each.

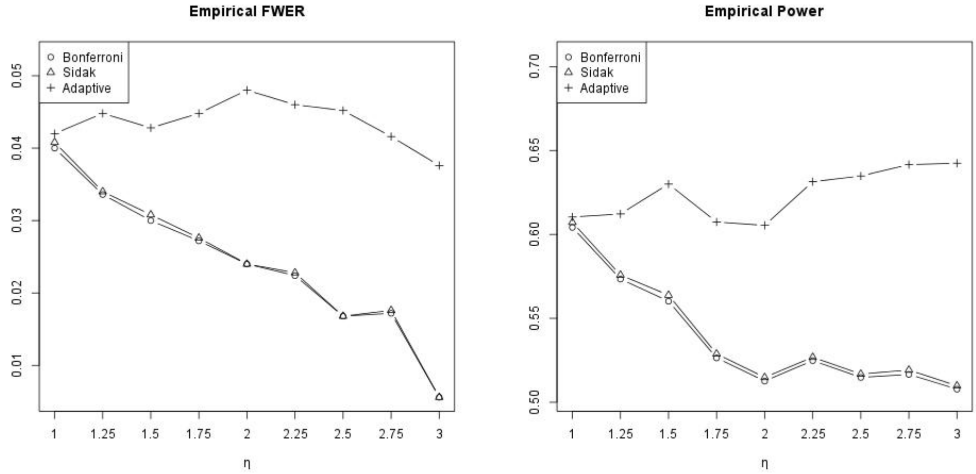

Figure 14: Empirical FWER and power as functions of $\eta \in \{1, 1.25, ..., 2.75, 3\}$ for $m = 4$ and $m_0 = 2$. The datsets of size $n = 1600$ were generated from an Archimedean copula defined by (21). The results are based on $K = 2500$ independent repetitions with $B = 400$ bootstrap replications in Algorithm 4.1 each.

distribution (again confirmed by goodness-of-fit tests). The empirical variances are again very small, implying that the empirical FWER is well concentrated. Figure 17, however, reflects a slightly liberal behavior of the empirically calibrated multiple tests under Model 4.2.(b), at least in case of $m = m_0$, if η is small (see in particular the left graph in Figure 17). Further simulations (not shown here) confirm that this liberal behavior attenuates with growing values of η, as reflected by the middle and the right graph in Figure 17. Since small values of η correspond to a low degree of dependency among test statistics, one may in practice apply a Šidák correction whenever the estimated value of η is below a certain threshold η_{lower} (say). Our simulations suggest to choose η_{lower} in the range of $[2, 3]$. Another possibility consists in adjusting the nominal value of α based on computer simulations under H_0.

5 Discussion

First, let us mention that approximate confidence regions for the main parameter ϑ can straightforwardly be deduced from the empirically calibrated vector $\hat{\mathbf{c}}$ of critical values by virtue of the extended correspondence theorem, see Section 4.1 of Finner (1994). If, in contrast, the main focus is on power of the multiple test, then it is recommendable to consider step-down variants of the considered multiple tests as nicely described by Romano and Wolf (2005). Their construction principle is particularly easy to apply if Assumption 2.1 holds true.

Second, one may ask why the empirical calibration of $\hat{\mathbf{c}}$ is in the present paper performed via the pre-estimation of η and not directly via resampling of the original data and application of the 'max T' or 'min P' algorithms suggested by Westfall and Young (1993). In the case of marginal k-sample problems with $k \geq 2$, the permutation methods of Westfall and Young (1993)

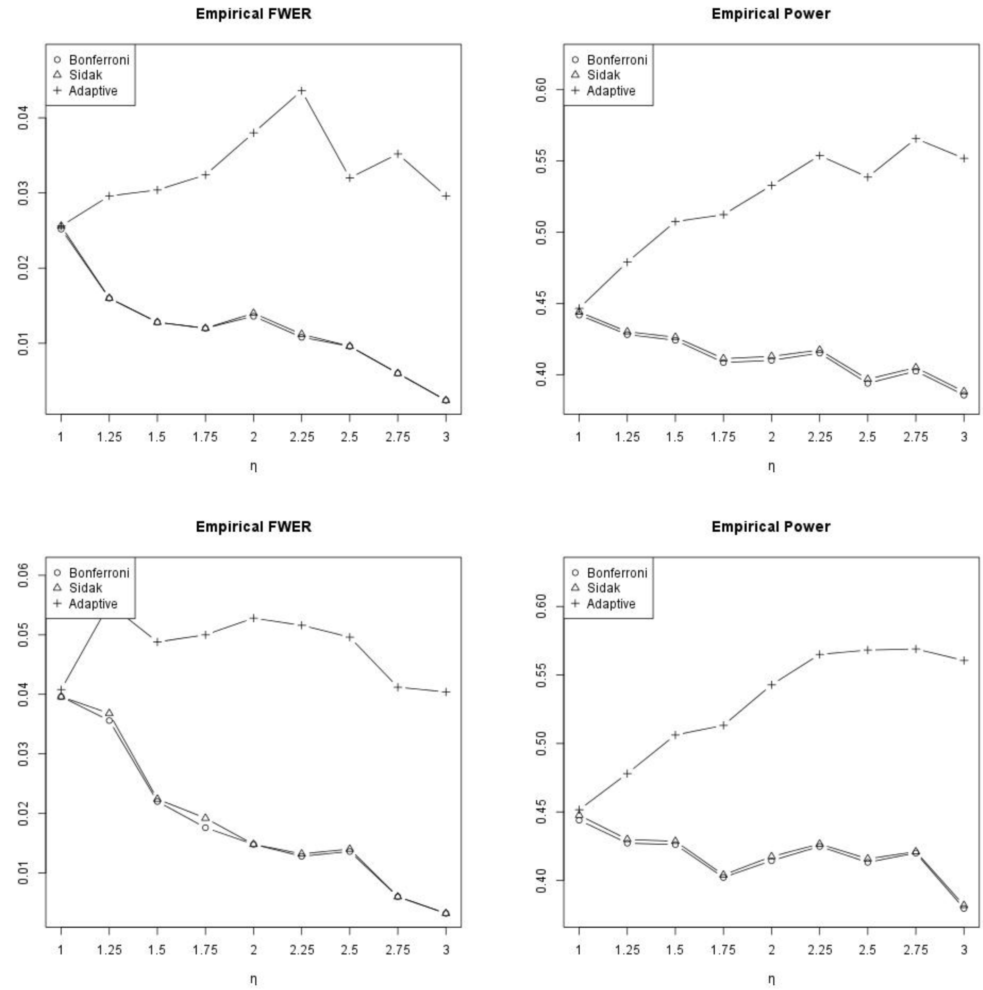

Figure 15: Empirical FWER and power as functions of $\eta \in \{1, 1.25, ..., 2.75, 3\}$ for $m = 8$, $m_0 = 3$ (upper panel) and $m_0 = 6$ (lower panel). The datsets of size $n = 1600$ were generated from an Archimedean copula defined by (21). The results are based on $K = 2500$ independent repetitions with $B = 400$ bootstrap replications in Algorithm 4.1 each.

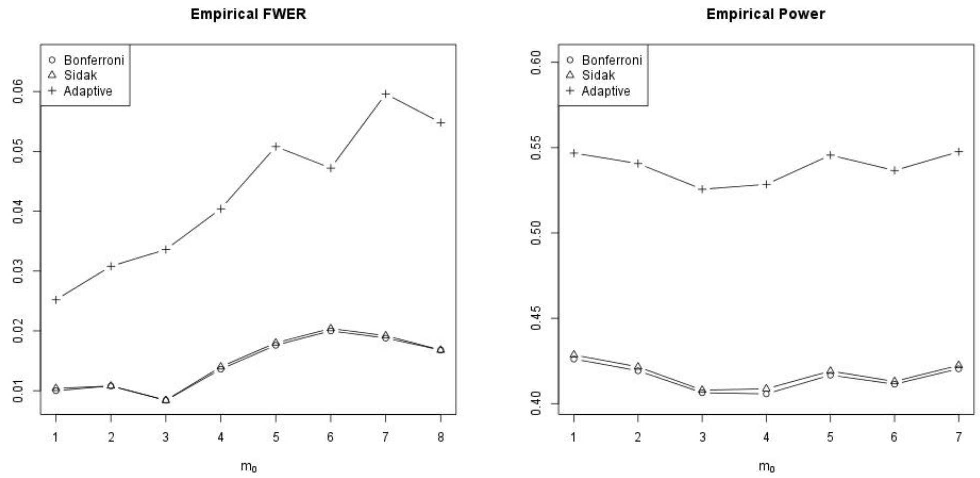

Figure 16: Empirical FWER and power as functions of $m_0 \in \{1, 2, ..., 8\}$ for $m = 8$ and $\eta = 2$. The datsets of size $n = 1600$ were generated from an Archimedean copula defined by (21). The results are based on $K = 2500$ independent repetitions with $B = 400$ bootstrap replications in Algorithm 4.1 each.

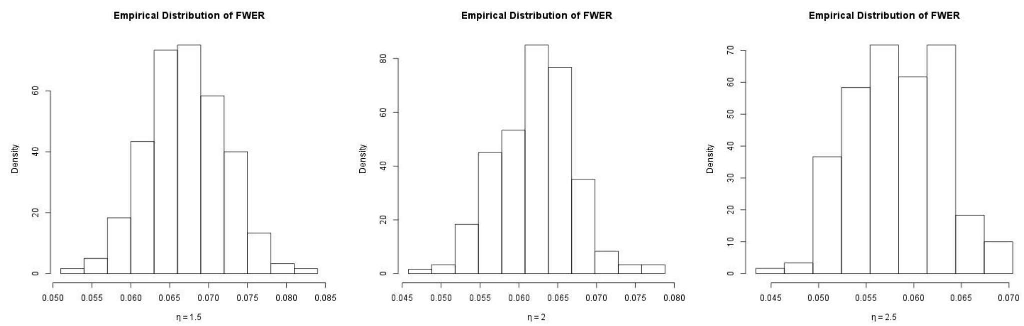

Figure 17: Histograms of the empirical FWER of empirically calibrated multiple tests for $m = m_0 = 4$ and $\eta \in \{1.5, 2.0, 2.5\}$. The datsets of size $n = 1600$ were generated from an Archimedean copula defined by (21). The results are based on 200 simulation runs with $K = 2500$ independent repetitions each, where $B = 400$ bootstrap replications in Algorithm 4.1 were performed.

are indeed an attractive alternative and even asymptotically optimal as shown by Meinshausen et al. (2011). However, in the cases of the one-sample problems studied in Section 4, marginal test statistics are invariant under data permutations such that this method is not applicable. Bootstrapping quantiles of a high-dimensional random vector appears to be much less reliable than bootstrapping a lower-dimensional copula parameter. Hence, if a parametric copula model can be assumed or even deduced by limit theorems, the approach of the present paper seems to be the better choice. Furthermore, the uncertainty of the estimation can precisely be quantified by applying Theorem 3.1, implying confidence statements about the realized FWER.

Finally, from the practical point of view, it is interesting to explore which type of copula is appropriate for which type of real-world application, especially if no theoretical results are at hand. This topic, however, is beyond the scope of the present work and deferred to future research. A promising nonparametric approach consists of modeling dependency structures by Bernstein copulae, see Diers et al. (2012) and Cottin and Pfeifer (2013).

References

Bernstein, S., 1929. Sur les fonctions absolument monotones. Acta Mathematica 52 (1), 1–66. URL `http://dx.doi.org/10.1007/BF02592679`

Bickel, P., Götze, F., van Zwet, W., 1997. Resampling fewer than n observations: Gains, losses, and remedies for losses. Stat. Sin. 7 (1), 1–31.

Bickel, P. J., Freedman, D., 1981. Some asymptotic theory for the bootstrap. Ann. Stat. 9, 1196–1217.

Block, H. W., Costigan, T., Sampson, A. R., 1992. Product-type probability bounds of higher order. Probab. Eng. Inf. Sci. 6 (3), 349–370.

Bodnar, T., Dickhaus, T., 2013. False Discovery Rate Control under Archimedean Copula. Preprint, arXiv:1305.3897.

Bodnar, T., Schmid, W., 2008. A test for the weights of the global minimum variance portfolio in an elliptical model. Metrika 67, 127–143.

Bonferroni, C. E., 1935. Il calcolo delle assicurazioni su gruppi di teste. Studi in onore Salvatore Ortu Carboni 13-60.

Bonferroni, C. E., 1936. Teoria statistica delle classi e calcolo delle probabilita. Pubbl. d. R. Ist. Super. di Sci. Econom. e Commerciali di Firenze 8. Firenze: Libr. Internaz. Seeber.

Cottin, C., Pfeifer, D., 2013. From Bernstein polynomials to Bernstein copulas. Journal of Applied Functional Analysis, to appear.

DasGupta, A., 2008. Asymptotic Theory of Statistics and Probability. Springer Texts in Statistics. Springer.

Dickhaus, T., 2013a. Randomized p-values for multiple testing of composite null hypotheses. J. Stat. Plann. Inference 143 (11), 1968–1979.

Dickhaus, T., 2013b. Simultaneous Statistical Inference with Applications in the Life Sciences. Springer, forthcoming.

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). Global Science and Technology Forum (GSTF), pp. 75–80.

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, to appear.

Diers, D., Eling, M., Marek, S. D., 2012. Dependence modeling in non-life insurance using the Bernstein copula. Insur. Math. Econ. 50 (3), 430–436.

Efron, B., 1979. Bootstrap methods: Another look at the jackknife. Ann. Stat. 7, 1–26.

Fengler, M. R., Okhrin, O., 2012. Realized Copula. SFB 649 Discussion Paper 2012-034, Sonderforschungsbereich 649, Humboldt-Universität zu Berlin, Germany, available at http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2012-034.pdf.

Finner, H., 1994. Testing Multiple Hypotheses: General Theory, Specific Problems, and Relationships to Other Multiple Decision Procedures. Habilitationsschrift. Fachbereich IV, Universität Trier.

Gabriel, K. R., 1969. Simultaneous test procedures - some theory of multiple comparisons. Ann. Math. Stat. 40, 224–250.

Genest, C., Nešlehová, J., Ben Ghorbal, N., 2011. Estimators based on Kendall's tau in multivariate copula models. Aust. N. Z. J. Stat. 53 (2), 157–177.

Genest, C., Rivest, L.-P., 1989. A characterization of Gumbel's family of extreme value distributions. Statistics & Probability Letters 8 (3), 207–211.

Genest, C., Rivest, L.-P., 1993. Statistical inference procedures for bivariate Archimedean copulas. J. Am. Stat. Assoc. 88 (423), 1034–1043.

Genz, A., Bretz, F., 2009. Computation of multivariate normal and t probabilities. Lecture Notes in Statistics 195. Berlin: Springer.

Ghosh, D., 2011. Generalized Benjamini-Hochberg procedures using spacings. Technical report, Penn State University.

Gudendorf, G., Segers, J., 2010. Extreme-value copulas. In: Jaworski, P., Durante, F., Härdle, W. K., Rychlik, T. (Eds.), Copula Theory and Its Applications. Lecture Notes in Statistics. Springer Berlin Heidelberg, pp. 127–145.
URL http://dx.doi.org/10.1007/978-3-642-12465-5_6

Gupta, A., Varga, T., Bodnar, T., 2013. Elliptically Contoured Models in Statistics and Portfolio Theory. Springer.

Hansen, L. P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Hofert, M., 2012. A stochastic representation and sampling algorithm for nested Archimedean copulas. Journal of Statistical Computation and Simulation 82 (9), 1239–1255.
URL http://www.tandfonline.com/doi/abs/10.1080/00949655.2011.574632

Hofert, M., Mächler, M., McNeil, A. J., 2012. Likelihood inference for Archimedean copulas in high dimensions under known margins. J. Multivariate Anal. 110, 133–150.

Höffding, W., 1940. Maßstabinvariante Korrelationstheorie. Schr. math. Inst. Inst. angew. Math. Univ. Berlin 5, 181–233.

Hothorn, T., Bretz, F., Westfall, P., Jun 2008. Simultaneous inference in general parametric models. Biom. J. 50 (3), 346–363.

Joe, H., 1990. Multivariate concordance. J. Multivariate Anal. 35 (1), 12–30.

Kendall, M. G., 1938. A new measure of rank correlation. Biometrika, Cambridge, 30, 81-93.

Kendall, M. G., Babington Smith, B., 1940. On the method of paired comparisons. Biometrika, Cambridge, 31, 324-345.

McNeil, A. J., Nešlehová, J., 2009. Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. Ann. Stat. 37, 3059–3097.

Meinshausen, N., Maathuis, M. H., Bühlmann, P., 2011. Asymptotic optimality of the Westfall-Young permutation procedure for multiple testing under dependence. Ann. Stat. 39 (6), 3369–3391.

Nelsen, R. B., 2006. An introduction to copulas. 2nd ed. Springer Series in Statistics. New York, NY: Springer.

Roeder, K., Wasserman, L., 2009. Genome-wide significance levels and weighted hypothesis testing. Statistical Science 24 (4), 398–413.

Romano, J. P., Wolf, M., 2005. Exact and approximate stepdown methods for multiple hypothesis testing. J. Am. Stat. Assoc. 100 (469), 94–108.

Sarkar, S., 2008. Rejoinder: On methods controlling the false discovery rate. Sankhyā: The Indian Journal of Statistics, Series A 70 (2), 183–185.

Šidák, Z., 1967. Rectangular confidence regions for the means of multivariate normal distributions. J. Am. Stat. Assoc. 62, 626–633.

Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris 8, 229–231.

Sklar, A., 1996. Random variables, distribution functions, and copulas - a personal look backward and forward. In: Distributions with Fixed Marginals and Related Topics. Institute of Mathematical Statistics, Hayward, CA, pp. 1–14.

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. Wiley, New York.