Supplementary Material to: A nonparametric test for paired data

by Grzegorz Wyłupek, Institute of Mathematics, University of Wrocław

Here, we discuss the outcomes of an additional simulation study and present a description of the `R` codes. Specifically, Section S1 demonstrates influence of ε on the power of the M test. Section S2 discusses a selection of the parameter c under small, as well as large sample sizes. Small sample comparison of the tests under consideration is given in Section S3, while Section S4 contains the related results when n is large. Finally, Section S5 presents the description of the `R` codes attached in 3 separate `R` files: `Power_Functions.R`, `Generators_Alternative.R`, `Generators_Null_Model.R`.

S1. Influence of ε on the power of the M test

In this section, we illustrate the influence of ε on the powers of the WKS and M tests. For this purpose, we select from each example considered in Section 4.3 one alternative for which the powers are the largest and calculate the empirical powers of those tests for $\varepsilon \in \{0.0001, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5\}$. The results are presented in Tables S1-S9. For better comparison, the powers of the KS test are also included.

First, we recall the definitions of the aforementioned statistics. Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be the independent identically distributed random vectors coming from the population with the cumulative distribution function $H(x, y)$ with the continuous margins $F(x)$ and $G(y)$. Then, $\hat{H}(x, y) = (1/n)\sum_{i=1}^{n} \mathbb{1}(X_i \le x, Y_i \le y)$, $\hat{F}(x) = (1/n)\sum_{i=1}^{n} \mathbb{1}(X_i \le x)$, and $\hat{G}(y) = (1/n)\sum_{i=1}^{n} \mathbb{1}(Y_i \le y)$, where $\mathbb{1}$ stands for the indicator of the set $\cdot$, are the nonparametric estimates of H, F, and G, respectively. Set $I = (1/2)F + (1/2)G$, $\varepsilon = 0.0001$, $a_I = I^{-1}(\varepsilon)$, and $b_I = I^{-1}(1-\varepsilon)$, where I^{-1} is the generalized inverse.

We have

$$KS = \sqrt{n} \sup_{x \in \mathbb{R}} |\hat{F}(x) - \hat{G}(x)|,$$

$$WKS = \sup_{a_I \le x \le b_I} \left| \frac{\sqrt{n}\{\hat{F}(x) - \hat{G}(x)\}}{\sqrt{\hat{F}(x) + \hat{G}(x) - 2\hat{H}(x, x) - [\hat{F}(x) - \hat{G}(x)]^2}} \right|,$$

$$M = \max\{KS, WKS\}.$$

The larger the value of ε, the smaller the set on which the standardized difference between $\hat{F}$ and $\hat{G}$ is employed by the WKS solution, and, thereby, by M as well. Furthermore, the larger the standardized difference between $\hat{F}$ and $\hat{G}$ in the tails, the larger decrease of the power with a larger ε. Obviously, the power of M oscillates between the power of KS and the power of WKS. The results shown in the tables allow us to assess where the standardized difference between $\hat{F}$ and $\hat{G}$ is the largest. It is related to the largest decrease of power when ε changes.

Table S1 Empirical powers under *Example 1* with $\mu = 0.8$ against ε. $n = 50$, $\alpha = 0.05$, $c = 0.8$, 1000 MC runs, 1000 wild bootstrap runs. Values multiplied by 100

	ε						
Test	0.0001	0.01	0.1	0.2	0.3	0.4	0.5
KS	75.6	75.6	75.6	75.6	75.6	75.6	75.6
WKS	73.5	73.5	73.1	71.3	70.0	66.1	60.2
M	73.5	73.5	73.1	71.3	70.0	66.3	68.1

Table S2 Empirical powers under *Example 2* with $\sigma = 2.0$ against ε. $n = 50$, $\alpha = 0.05$, $c = 0.8$, 1000 MC runs, 1000 wild bootstrap runs. Values multiplied by 100

	ε						
Test	0.0001	0.01	0.1	0.2	0.3	0.4	0.5
KS	59.0	59.0	59.0	59.0	59.0	59.0	59.0
WKS	84.8	84.8	82.8	60.9	30.4	11.2	4.1
M	84.8	84.8	82.8	60.9	30.4	11.2	12.8

Table S3 Empirical powers under *Example 3* with $\sigma = 12$ against ε. $n = 50$, $\alpha = 0.05$, $c = 0.8$, 1000 MC runs, 1000 wild bootstrap runs. Values multiplied by 100

	ε						
Test	0.0001	0.01	0.1	0.2	0.3	0.4	0.5
KS	14.2	14.2	14.2	14.2	14.2	14.2	14.2
WKS	73.0	73.0	60.0	15.7	7.4	6.3	5.1
M	73.0	73.0	60.0	15.7	7.4	6.3	7.2

Table S4 Empirical powers under *Example 4* with $\beta = 0.7$ against ε. $n = 50$, $\alpha = 0.05$, $c = 0.8$, 1000 MC runs, 1000 wild bootstrap runs. Values multiplied by 100

	ε						
Test	0.0001	0.01	0.1	0.2	0.3	0.4	0.5
KS	65.4	65.4	65.4	65.4	65.4	65.4	65.4
WKS	83.5	83.5	83.0	75.2	65.3	54.3	42.1
M	83.5	83.5	83.0	75.2	65.3	54.3	47.8

Table S5 Empirical powers under *Example 5* with $\beta = 1.4$ against ε. $n = 50$, $\alpha = 0.05$, $c = 0.8$, 1000 MC runs, 1000 wild bootstrap runs. Values multiplied by 100

	ε						
Test	0.0001	0.01	0.1	0.2	0.3	0.4	0.5
KS	14.9	14.9	14.9	14.9	14.9	14.9	14.9
WKS	70.0	70.0	57.2	20.9	9.4	5.0	3.1
M	70.0	70.0	57.2	20.9	9.4	5.2	8.7

Table S6 Empirical powers under *Example 6* with $\beta = 0.6$ against ε. $n = 50$, $\alpha = 0.05$, $c = 0.8$, 1000 MC runs, 1000 wild bootstrap runs. Values multiplied by 100

	ε						
Test	0.0001	0.01	0.1	0.2	0.3	0.4	0.5
KS	70.4	70.4	70.4	70.4	70.4	70.4	70.4
WKS	96.1	96.1	96.4	87.6	68.1	53.5	39.6
M	96.1	96.1	96.4	87.6	68.1	53.5	46.7

Table S7 Empirical powers under *Example 7* with $\theta = 0.14$ against ε. $n = 50$, $\alpha = 0.05$, $c = 0.8$, 1000 MC runs, 1000 wild bootstrap runs. Values multiplied by 100

	ε						
Test	0.0001	0.01	0.1	0.2	0.3	0.4	0.5
KS	60.0	60.0	60.0	60.0	60.0	60.0	60.0
WKS	69.0	69.0	69.3	67.7	62.2	55.7	39.8
M	69.0	69.0	69.3	67.7	62.2	55.7	40.0

Table S8 Empirical powers under *Example 8* with $a = 0.4$ against ε. $n = 50$, $\alpha = 0.05$, $c = 0.8$, 1000 MC runs, 1000 wild bootstrap runs. Values multiplied by 100

	ε						
Test	0.0001	0.01	0.1	0.2	0.3	0.4	0.5
KS	29.1	29.1	29.1	29.1	29.1	29.1	29.1
WKS	77.6	77.6	77.8	57.8	14.1	4.3	3.0
M	77.6	77.6	77.8	57.8	14.1	4.3	4.6

Table S9 Empirical powers under *Example 9* with $a = 0.8$ against ε. $n = 50$, $\alpha = 0.05$, $c = 0.8$, 1000 MC runs, 1000 wild bootstrap runs. Values multiplied by 100

	ε						
Test	0.0001	0.01	0.1	0.2	0.3	0.4	0.5
KS	97.3	97.3	97.3	97.3	97.3	97.3	97.3
WKS	97.7	97.7	97.5	96.9	96.2	85.6	65.5
M	97.7	97.7	97.5	96.9	96.2	86.5	89.5

S2. A selection of the parameter c

Tables S10-S18 provide the recommendation (in boldface) for the choice of the correction c for $n = 10, 20, 30, 40$, and $n = 100, 200, 300, 400, 500$, respectively. For other sample sizes linear approximation can be applied to find a proper constant.

Table S10 Empirical Type I errors of the $\Phi^*_{M,n,\alpha,c}$ test against c under different scenarios. $n = 10$, $\alpha = 0.05$, $\varepsilon = 0.0001$, 1000 MC runs, 1000 wild bootstrap runs. Errors multiplied by 100

	$FGM(\theta)$			$Clayton(\theta)$			$Mardia(\theta)$			t_θ		
$c \setminus \theta$	-0.5	0	0.5	-0.5	0.5	1	0.5	0.7	0.9	1	3	5
1.00	1.5	2.0	1.7	1.5	2.9	1.9	2.7	3.4	8.1	1.8	2.7	2.3
0.90	0.2	0.2	0.1	0.2	0.3	0.1	0.2	0.5	7.7	0.2	0.1	0.1
0.80	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.7	0.0	0.0	0.0
0.70	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.7	0.0	0.0	0.0

Although, under the $Mardia(0.9)$ copula, the Type I error is larger than the prescribed significance level, taking into account the remaining results, we suggest $c = 1$.

Remark S1

Since the weight w_n^* takes negative values under $n = 10, 20$, we correct its form adding modulus. Such a situation also holds under $n = 30$ and the $Clayton(-0.5)$ copula, where we also incorporate such a correction.

Table S11 Empirical Type I errors of the $\Phi^*_{M,n,\alpha,c}$ test against c under different scenarios. $n = 20$, $\alpha = 0.05$, $\varepsilon = 0.0001$, 1000 MC runs, 1000 wild bootstrap runs.

Errors multiplied by 100

	$FGM(\theta)$			$Clayton(\theta)$			$Mardia(\theta)$			t_θ		
$c \setminus \theta$	-0.5	0	0.5	-0.5	0.5	1	0.5	0.7	0.9	1	3	5
1.00	15.8	17.2	19.5	13.6	21.4	19.7	18.1	14.1	8.7	15.4	17.6	16.0
0.90	11.1	12.6	13.2	10.0	13.6	11.9	11.8	9.4	4.2	9.8	11.8	10.5
0.80	7.0	7.8	9.4	7.1	8.7	7.8	7.6	5.5	1.5	5.7	7.4	6.8
0.75	5.0	5.0	5.8	4.7	6.2	4.5	5.5	3.2	0.9	3.8	5.4	5.2
0.70	3.5	3.2	3.6	3.5	3.8	3.2	3.8	1.7	0.8	3.0	4.3	3.6

Table S12 Empirical Type I errors of the $\Phi^*_{M,n,\alpha,c}$ test against c under different scenarios. $n = 30$, $\alpha = 0.05$, $\varepsilon = 0.0001$, 1000 MC runs, 1000 wild bootstrap runs.

Errors multiplied by 100

	$FGM(\theta)$			$Clayton(\theta)$			$Mardia(\theta)$			t_θ		
$c \setminus \theta$	-0.5	0	0.5	-0.5	0.5	1	0.5	0.7	0.9	1	3	5
1.00	14.6	15.7	15.9	12.8	16.4	17.7	12.1	13.6	14.3	13.9	15.9	13.0
0.90	9.6	10.0	10.2	9.4	10.9	12.5	9.2	9.0	7.0	9.8	11.1	7.8
0.80	6.8	7.6	7.5	6.4	7.2	8.0	5.8	5.0	2.4	5.6	8.0	5.2
0.75	5.7	5.8	6.0	5.3	5.1	6.2	4.6	3.9	1.7	4.5	5.5	3.8
0.70	4.6	4.5	4.9	3.5	4.3	4.3	2.9	2.3	0.9	3.7	4.7	3.1

Table S13 Empirical Type I errors of the $\Phi^*_{M,n,\alpha,c}$ test against c under different scenarios. $n = 40$, $\alpha = 0.05$, $\varepsilon = 0.0001$, 1000 MC runs, 1000 wild bootstrap runs.

Errors multiplied by 100

	$FGM(\theta)$			$Clayton(\theta)$			$Mardia(\theta)$			t_θ		
$c \setminus \theta$	-0.5	0	0.5	-0.5	0.5	1	0.5	0.7	0.9	1	3	5
1.00	11.6	12.2	13.4	10.1	14.2	16.1	13.7	14.1	16.3	11.0	12.8	12.4
0.90	9.2	9.7	9.6	7.8	10.1	10.6	9.7	9.9	10.0	8.0	9.6	8.9
0.80	4.5	5.3	5.8	5.3	5.9	6.4	6.4	5.7	5.0	4.8	5.7	5.7
0.75	3.0	3.7	4.3	3.0	4.6	5.2	4.9	4.2	3.6	3.6	4.8	4.6
0.70	2.4	2.5	3.0	2.4	3.4	2.9	3.1	2.8	2.0	2.9	3.2	3.3

Table S14 Empirical Type I errors of the $\Phi^*_{M,n,\alpha,c}$ test against c under different scenarios. $n = 100$, $\alpha = 0.05$, $\varepsilon = 0.0001$, 1000 MC runs, 1000 wild bootstrap runs. Errors multiplied by 100

	$FGM(\theta)$			$Clayton(\theta)$			$Mardia(\theta)$			t_θ		
$c \setminus \theta$	-0.5	0	0.5	-0.5	0.5	1	0.5	0.7	0.9	1	3	5
1.00	8.2	9.9	9.3	9.4	9.8	10.4	9.4	9.8	11.3	8.2	9.6	10.0
0.95	6.2	8.0	7.0	7.5	8.0	8.9	8.4	7.6	9.1	6.9	8.1	7.7
0.90	4.6	5.8	5.6	6.3	6.9	7.5	6.2	6.3	7.5	5.2	5.9	6.0
0.85	4.2	4.7	4.8	4.9	4.7	5.6	5.0	4.2	5.6	4.2	4.8	4.2
0.80	2.9	3.3	2.9	3.6	3.5	4.3	3.4	2.9	3.1	3.0	3.8	2.9

Table S15 Empirical Type I errors of the $\Phi^*_{M,n,\alpha,c}$ test against c under different scenarios. $n = 200$, $\alpha = 0.05$, $\varepsilon = 0.0001$, 1000 MC runs, 1000 wild bootstrap runs. Errors multiplied by 100

	$FGM(\theta)$			$Clayton(\theta)$			$Mardia(\theta)$			t_θ		
$c \setminus \theta$	-0.5	0	0.5	-0.5	0.5	1	0.5	0.7	0.9	1	3	5
1.00	6.9	7.8	8.8	6.5	9.0	9.8	6.3	6.9	11.3	7.8	8.8	7.5
0.95	5.0	5.9	6.5	5.1	6.7	7.3	4.8	5.0	8.4	6.1	6.0	6.4
0.90	3.5	4.2	5.2	4.1	5.2	6.1	3.1	4.3	6.8	5.0	5.0	4.5
0.85	2.3	3.3	3.7	3.0	4.0	4.4	2.6	3.7	4.7	4.0	3.9	3.5
0.80	1.7	2.4	2.5	2.2	2.4	3.2	1.8	2.7	3.5	3.2	3.2	2.4

Although, under the $Mardia(0.9)$ copula, the Type I error is larger than the prescribed significance level, taking into account the remaining results, we suggest $c = 0.9$.

Table S16 Empirical Type I errors of the $\Phi^*_{M,n,\alpha,c}$ test against c under different scenarios. $n = 300$, $\alpha = 0.05$, $\varepsilon = 0.0001$, 1000 MC runs, 1000 wild bootstrap runs. Errors multiplied by 100

	$FGM(\theta)$			$Clayton(\theta)$			$Mardia(\theta)$			t_θ		
$c \setminus \theta$	-0.5	0	0.5	-0.5	0.5	1	0.5	0.7	0.9	1	3	5
1.00	6.8	6.9	5.7	7.1	6.7	7.1	9.0	8.1	7.1	6.9	6.9	7.1
0.95	5.8	5.6	4.8	6.0	5.3	5.3	7.0	6.8	6.0	5.6	5.2	5.8
0.90	4.2	4.7	4.0	5.3	4.0	4.2	4.8	4.9	4.5	4.2	4.6	5.1
0.85	3.3	3.7	2.6	4.3	3.2	2.5	3.6	3.5	3.2	3.4	4.1	4.5
0.80	2.5	2.7	2.2	2.4	2.5	1.7	2.8	2.3	1.9	2.3	3.4	3.2

Table S17 Empirical Type I errors of the $\Phi^*_{M,n,\alpha,c}$ test against c under different scenarios. $n = 400$, $\alpha = 0.05$, $\varepsilon = 0.0001$, 1000 MC runs, 1000 wild bootstrap runs. Errors multiplied by 100

	$FGM(\theta)$			$Clayton(\theta)$			$Mardia(\theta)$			t_θ		
$c \setminus \theta$	-0.5	0	0.5	-0.5	0.5	1	0.5	0.7	0.9	1	3	5
1.00	7.4	6.6	7.4	7.0	7.2	7.5	5.4	6.4	8.3	6.7	8.6	7.2
0.95	5.8	5.5	5.9	5.0	5.2	5.7	4.5	5.1	6.0	5.5	7.2	5.7
0.90	4.8	4.4	4.6	4.1	3.8	4.7	3.5	3.9	4.1	4.3	5.7	4.2
0.85	3.5	3.3	3.0	3.1	3.3	2.8	2.1	3.2	3.0	3.8	5.1	3.0
0.80	2.8	2.6	2.4	2.3	2.2	2.0	1.1	2.0	2.4	2.7	3.4	2.3

Table S18 Empirical Type I errors of the $\Phi^*_{M,n,\alpha,c}$ test against c under different scenarios. $n = 500$, $\alpha = 0.05$, $\varepsilon = 0.0001$, 1000 MC runs, 1000 wild bootstrap runs. Errors multiplied by 100

	$FGM(\theta)$			$Clayton(\theta)$			$Mardia(\theta)$			t_θ		
$c \setminus \theta$	-0.5	0	0.5	-0.5	0.5	1	0.5	0.7	0.9	1	3	5
1.00	7.1	6.7	6.1	5.7	5.5	6.6	7.7	7.5	8.5	7.4	7.3	6.7
0.95	5.1	5.6	4.5	4.6	4.3	5.2	6.2	6.6	7.1	5.0	5.7	6.0
0.90	3.5	4.7	3.3	3.5	3.5	4.3	4.8	5.1	6.0	3.6	4.3	4.8
0.85	2.4	2.7	2.8	2.5	2.9	2.8	3.5	3.3	4.6	2.8	3.1	3.7
0.80	1.7	1.9	2.2	2.4	2.3	1.5	2.4	2.3	3.9	2.1	2.2	2.6

Under $n = 10$ and a really strong dependence structure, i.e., $Mardia(0.9)$, 9 of 10 coordinates of the data are the same, on average. For instance, $X_1 = Y_1, \ldots, X_9 = Y_9$, and only $X_{10} \neq Y_{10}$. As a result, the distribution of the test statistic is pretty discrete. Actually, M only takes 6 values. This is the reason of selection $c = 1$. Under larger sample sizes the problem vanishes and for $n = 20, 30, 40, 50, 100, 200, 300, 400, 500$, we observe that c very slowly tends to 1 as the asymptotic result asserts.

S3. A small sample comparison

Here, we investigate the Type I errors and powers of the tests under consideration for $n = 20$. It can be seen that such a small number of the observations disturbs to keep the level well to BMP, M-C, $GPRY$, QE, VGH, and KP. The problems with the Type I error control of the VGH solution under $Mardia(\theta)$ copula has the same origin as in the case where $n = 50$.

Table S19 Empirical Type I errors under different scenarios. $n = 20$, $\alpha = 0.05$, $\varepsilon = 0.0001$, $c = 0.75$, 1000 MC runs, 1000 wild bootstrap runs. Errors multiplied by 100

	$FGM(\theta)$			$Clayton(\theta)$			$Mardia(\theta)$			t_θ			
Test\\θ	-0.5	0	0.5	-0.5	0.5	1	0.5	0.7	0.9	1	3	5	Method
W	4.8	4.8	4.3	5.7	4.6	4.5	4.6	3.8	1.0	4.9	4.9	5.8	Asymptotic
BMP	6.4	6.3	6.4	6.0	7.6	6.8	6.5	5.3	3.5	7.4	6.2	7.0	Asymptotic
M-C	3.0	2.4	1.4	4.0	1.4	1.0	1.2	0.5	0.1	5.2	5.4	7.0	Smoothed bootstrap
$GPRY$	8.7	8.3	8.2	9.3	7.8	8.5	8.9	6.1	-	7.7	8.3	7.5	Asymptotic approximation
QE	7.8	7.6	5.7	8.9	5.7	7.1	6.1	5.5	3.4	2.5	5.8	6.6	Bayesian bootstrap
VGH	4.6	5.0	4.1	6.3	4.4	5.1	11.0	56.2	99.8	5.1	5.4	5.8	Finite sample
KP	5.1	5.4	5.7	4.6	6.2	6.2	5.2	5.2	10.0	3.6	4.9	5.8	Wild bootstrap
KS	5.0	5.8	3.3	5.0	4.0	2.8	4.8	2.9	1.6	4.1	4.3	4.4	Wild bootstrap
M	5.0	5.0	5.8	4.7	6.2	4.5	5.5	3.2	0.9	3.8	5.4	5.2	Wild bootstrap

Remark S2
In the case of $Mardia(0.9)$, there is a problem with calculation of the Type I error of the $GPRY$ test. Specifically, finding the inverse covariance matrix of the score vector is impossible since the determinant is equal to zero.

To briefly summarize the results presented in Figure S1, it can be stated that the order of the competitive solutions is similar to the one when $n = 50$. It appears that 20 observations is to small to notice nice properties of the new solution. It should be also emphasized that too large Type I error of the most competitive solutions under $\mathcal{H}$ implies their larger powers under $\mathcal{A}$.

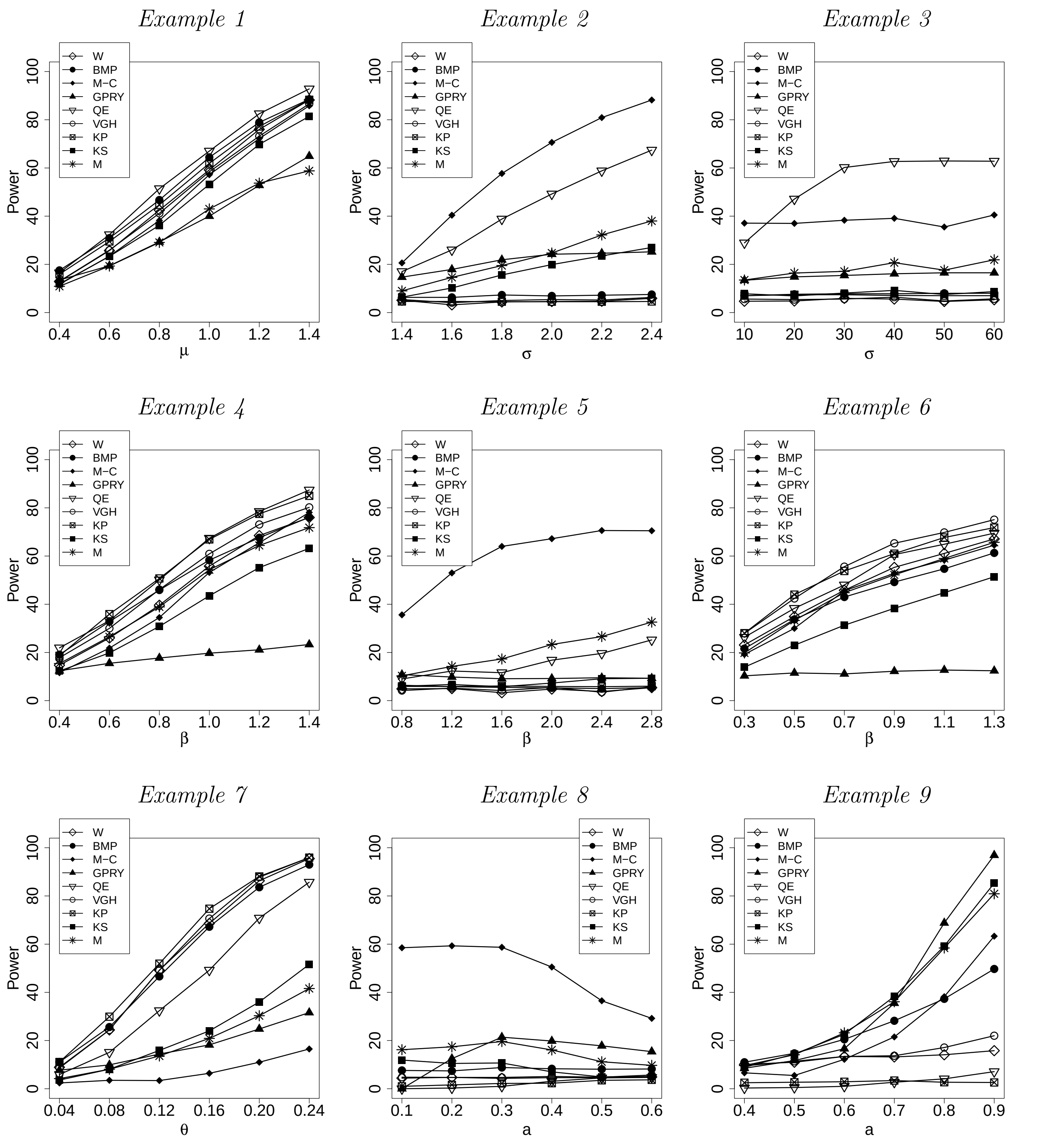

Fig. S1. Empirical powers against a parameter. $\alpha = 0.05$, $n = 20$, $\varepsilon = 0.0001$, $c = 0.75$. Based on 1000 MC runs and 1000 wild bootstrap runs. Values multiplied by 100.

S4. A large sample comparison

Here, we investigate the Type I errors and powers of the tests under consideration for $n = 200$. It can be seen that such a number of observations allows the tests to keep the level well. An exception is the VGH solution under the $Mardia(\theta)$ copula, which corresponds to a well recognizable problem under the strong dependence structure of the data at hand. The larger the n, the larger the Type I error of that solution.

Table S20 Empirical Type I errors under different scenarios. $n = 200$, $\alpha = 0.05$, $\varepsilon = 0.0001$, $c = 0.9$, 1000 MC runs, 1000 wild bootstrap runs. Errors multiplied by 100

	$FGM(\theta)$			$Clayton(\theta)$			$Mardia(\theta)$			t_θ			
Test\\θ	-0.5	0	0.5	-0.5	0.5	1	0.5	0.7	0.9	1	3	5	Method
W	4.7	4.4	4.6	4.6	4.6	4.6	4.6	3.5	0.8	5.8	6.0	5.2	Asymptotic
BMP	5.5	5.8	5.5	5.5	6.1	5.4	5.7	5.8	5.1	5.3	4.1	5.4	Asymptotic
M-C	2.8	1.8	1.6	4.9	1.1	0.8	2.2	0.7	0.0	3.5	4.7	6.2	Smoothed bootstrap
$GPRY$	5.3	5.8	5.4	5.1	5.5	5.1	5.1	4.9	4.7	3.8	4.8	5.2	Asymptotic approximation
QE	5.6	5.6	5.6	4.5	5.2	5.2	4.7	4.1	5.3	1.3	5.3	5.1	Bayesian bootstrap
VGH	5.1	5.3	4.7	5.4	5.3	5.7	54.9	100.0	100.0	5.5	6.7	4.8	Finite sample
KP	5.2	5.3	5.4	5.2	5.4	5.7	5.2	5.7	5.2	4.0	5.0	5.4	Wild bootstrap
KS	4.9	5.1	6.1	4.3	5.4	5.5	4.2	4.4	5.4	4.5	4.2	4.9	Wild bootstrap
M	3.5	4.2	5.2	4.1	5.2	6.1	3.1	4.3	6.8	5.0	5.0	4.5	Wild bootstrap

To briefly summarize the results presented in Figure S2, it can be stated that the order of the competitive solutions is also similar to the one for $n = 50$. Under $n = 200$ observations, the advantages of the new M solution are clearly seen.

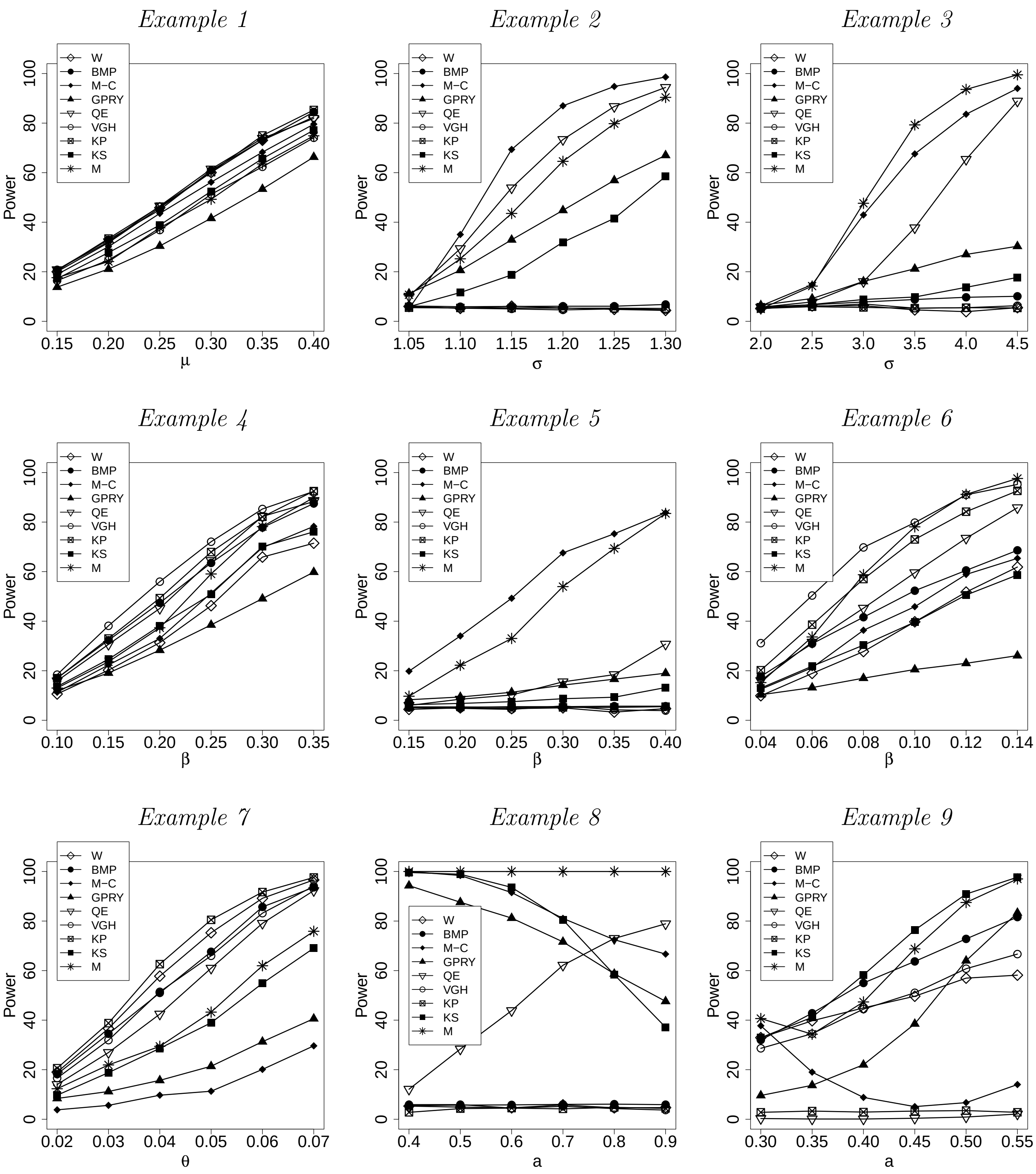

Fig. S2. Empirical powers against a parameter. $\alpha = 0.05$, $n = 200$, $\varepsilon = 0.0001$, $c = 0.9$. Based on 1000 MC runs and 1000 wild bootstrap runs. Values multiplied by 100.

S5. A description of the `R` codes

In the `R` file `Power_Functions.R` a software for computing values of the power functions of the investigated tests, i.e., W, BMP, M-C, $GPRY$, QE, VGH, KP, KS, and M, under the alternative: *Example 1* with the parameter $\mu = 0.8$, under $\alpha = 0.05$, $n = 50$, 1000 MC runs, 1000 wild bootstrap runs, is provided. The values are multiplied by 100. Additional comments on the functions, parameters, and the generator can be found in the file.

In the file `Generators_Alternative.R` the generators for the alternatives defined in Examples 1-9 are yielded. Copying a selected generator and pasting it to the file `Power_Functions.R` in a proper place allows one to calculate the powers of the tests under the given alternative. A selection of the parameter defining the alternative, sample size `n`, parameter `c`, and critical value of the VGH test (`cv.VGH` in the file `Power_Functions.R`) enables one to calculate the powers under different configurations of the parameters.

In the `R` file `Generators_Null_Model.R` the generators for the models obeying the null hypohtesis are given. Copying a selected generator and pasting it to the file `Power_Functions.R` in a proper place allows one to calculate the Type I errors of the tests under the given model. A selection of the parameter defining the hypothesis, the sample size `n`, the parameter `c`, and the critical value of the VGH test (`cv.VGH` in the file `Power_Functions.R`) enables one to calculate the Type I errors under different configurations of the parameters.

Acknowledgments

The author is grateful to the Editor, Associate Editor, and Reviewers for the comments which led to the improvement in the presentation. He would like to express appreciation to P. Martínez-Camblor for sharing the `R` code, L. Gutiérrez for sending the `R` package `BNPpairedtest`, and J.-F. Quessy for comments on his multiplier. Calculations have been carried out at Wrocław Centre for Networking and Supercomputing (`http://www.wcss.wroc.pl`) under grant No. 199. M. Jedrzejewski's aid in ensuring optimal use of the Centre's resources is acknowledged with much gratitude, likewise.

Funding

The research has been supported by the grant 2017/25/B/ST1/00500 from the National Science Centre, Poland.

ORCID

Grzegorz Wyłupek https://orcid.org/0000-0001-7731-8963