Empirical Likelihood Test for Diagonal Symmetry

Yongli Sang[a][*] and Xin Dang[b]

[a]Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA 70504, USA

[b]Department of Mathematics, University of Mississippi, University, MS 38677, USA

January 8, 2020

Abstract

Energy distance is a statistical distance between the distributions of random variables, which characterizes the equality of the distributions. Utilizing the energy distance, we develop a nonparametric test for the diagonal symmetry, which is consistent against any fixed alternatives. The test statistic developed in this paper is based on the difference of two U-statistics. By applying the jackknife empirical likelihood approach, the standard limiting chi-square distribution with degree freedom of one is established and is used to determine critical value and p-value of the test. Simulation studies show that our method is competitive in terms of empirical sizes and empirical powers.

Keywords: Diagonal symmetry, Energy distance, Jackknife empirical likelihood, U-statistic, Wilks' theorem

MSC 2010 subject classification: 62G35, 62G20

1 Introduction

Testing the departures from symmetry is crucial in statistics, especially in nonparametric statistical science. This problem has been discussed by many works for the univariate case (see e.g. [1, 2, 3, 4, 8, 14, 16, 17, 20, 21, 25, 30]). For the multivariate case, notions of symmetry include spherical, elliptical, central, and angular symmetry. One can refer to [31] for perspectives on those symmetries. The central symmetry is considered in this paper (see e.g. [10, 11, 12, 13, 22, 34]). In this setting, we explore testing the central symmetry of a distribution around the specified center. Without loss of generality, the center at $\mathbf{0}$ is assumed and the central symmetry at $\mathbf{0}$ is called the diagonal symmetry. Suppose $\boldsymbol{X}$ is a d-variate random variable from an unknown distribution F in $\mathbb{R}^d$, then the hypothesis of interest is

$$H_0 : \boldsymbol{X} \stackrel{D}{=} -\boldsymbol{X}, \tag{1}$$

where $\stackrel{D}{=}$ denotes equality in distribution.

Often an omnibus test is based on a discrepancy measure among distributions. The energy distance ([35, 36]) is one of such measures, which is defined as follows.

Definition 1.1 (Energy distance) *Suppose that $(\boldsymbol{X}, \boldsymbol{X}')$ and $(\boldsymbol{Y}, \boldsymbol{Y}')$ are independent pairs independently from d-variate distributions F and G, respectively. Then the energy distance between*

---

[*]CONTACT Yongli Sang. Email: yongli.sang@louisiana.edu

$\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *is*

$$\mathcal{E}(\boldsymbol{X}, \boldsymbol{Y}) = 2\mathbb{E}\|\boldsymbol{X} - \boldsymbol{Y}\| - \mathbb{E}\|\boldsymbol{X} - \boldsymbol{X}'\| - \mathbb{E}\|\boldsymbol{Y} - \boldsymbol{Y}'\|, \tag{2}$$

where $\|\cdot\|$ *denotes the Euclidean distance.*

Székely and Móri ([34]) have shown that the central symmetry at $\boldsymbol{0}$ holds if and only if $\mathbb{E}\|\boldsymbol{X}+\boldsymbol{X}'\| = \mathbb{E}\|\boldsymbol{X}-\boldsymbol{X}'\|$. Utilizing this fact, we propose an empirical likelihood test for the diagonal symmetry.

Empirical likelihood (EL) tests ([9, 5, 39]) have been proposed to avoid the time-consuming permutation procedure. The EL approach ([23], [24]) enjoys effectiveness of likelihood method and flexibility of nonparametric method, and hence has been widely used. On the computational side, the EL method involves a maximization of the nonparametric likelihood supported on data subject to some constraints. If these constraints are linear, the computation of the EL method is relatively easy. However, EL loses this computational efficiency when some nonlinear constraints are involved. Jackknife empirical likelihood (JEL) ([15]) is proposed to reduce the computation burden of the EL approach when nonlinear constraints are involved. The JEL has been proved very efficient in dealing with U-statistics and has attracted statistician's interest due to the efficiency. In the literature, there are a quantity of papers utilizing the JEL to conduct tests, see [28], [29],[18], [38] and so on.

As mentioned above, suppose $\boldsymbol{X}$ and $\boldsymbol{X}'$ are i.i.d. from distribution F in $\mathbb{R}^d$, then testing the diagonal symmetry of (1) will be equivalent to testing

$$H_0' : \mathbb{E}\big[\|\boldsymbol{X} + \boldsymbol{X}'\| - \|\boldsymbol{X} - \boldsymbol{X}'\|\big] = 0. \tag{3}$$

The left-side of the equation is half of the energy distance between $\boldsymbol{X}$ and $-\boldsymbol{X}'$ and can be naturally estimated by the energy statistic that is the difference of two U-statistics. However, the energy statistic is degenerate under H_0. This does not allow us to apply the regular JEL of [15] directly. We develop an improved JEL for such a degenerate U-statistic case in this paper, which enables us to test the diagonal symmetry efficiently.

The reminder of the paper is organized as follows. In Section 2, we develop the JEL method for the diagonal symmetry test. Simulation studies are conducted in Section 3. Section 4 concludes the paper with a brief summary. All proofs are reserved to the Appendix.

2   Methodology

Let $\{\boldsymbol{X}_i\}_{i=1}^n$ be a random sample from F, and define the energy statistic,

$$U = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \Big[\|\boldsymbol{X}_i + \boldsymbol{X}_j\| - \|\boldsymbol{X}_i - \boldsymbol{X}_j\|\Big],$$

which is a one-sample U-statistic of degree 2. However, it is degenerate under H_0 which does not allow us to apply the JEL directly. We develop an improved JEL approach for degenerate U-statistics in this section.

We partition the data set $\{\boldsymbol{X}_i\}_{i=1}^n$ into two parts: $\mathcal{X} = \{\boldsymbol{X}_1, ..., \boldsymbol{X}_{n_1}\}$ and $\mathcal{Y} = \{\boldsymbol{Y}_1, ..., \boldsymbol{Y}_{n_2}\}$, where $\boldsymbol{Y}_i = \boldsymbol{X}_{n_1+i}$, $i = 1, ..., n_2$ and $n_1 + n_2 = n$.

Define

$$U_1 = \binom{n_1}{2}^{-1} \sum_{1 \le k < j \le n_1} \|\boldsymbol{X}_k + \boldsymbol{X}_j\| := \binom{n_1}{2}^{-1} \sum_{1 \le k < j \le n_1} h_1(\boldsymbol{X}_k, \boldsymbol{X}_j)$$

$$U_2 = \binom{n_2}{2}^{-1} \sum_{1 \le k < j \le n_2} \|\boldsymbol{Y}_k - \boldsymbol{Y}_j\| := \binom{n_2}{2}^{-1} \sum_{1 \le k < j \le n_2} h_2(\boldsymbol{Y}_k, \boldsymbol{Y}_j),$$

where $h_1(\boldsymbol{s}, \boldsymbol{t}) = \|\boldsymbol{s} + \boldsymbol{t}\|$ and $h_2(\boldsymbol{s}, \boldsymbol{t}) = \|\boldsymbol{s} - \boldsymbol{t}\|$. U_1 and U_2 are unbiased estimators for $\mathbb{E}\|\boldsymbol{X} + \boldsymbol{X}'\|$ and $\mathbb{E}\|\boldsymbol{X} - \boldsymbol{X}'\|$, respectively. Thus, by (3), testing H_0 against any alternative hypothesis will be equivalent to testing $\mathbb{E}U_1 = \mathbb{E}U_2$.

In order to develop the JEL approach, the corresponding jackknife pseudo-value for U_1 and U_2 are respectively given by

$$\hat{V}_i^{(1)} = n_1 U_1 - (n_1 - 1)U_{1,n_1-1}^{(-i)}, \ i = 1, ..., n_1,$$
$$\hat{V}_i^{(2)} = n_2 U_2 - (n_2 - 1)U_{2,n_2-1}^{(-i)} \ i = 1, ..., n_2,$$

where

$$U_{1,n_1-1}^{(-i)} = \binom{n_1 - 1}{2}^{-1} \sum_{1 \le j \le k \le n_1, j,k \ne i} h_1(\boldsymbol{X}_j, \boldsymbol{X}_k),$$
$$U_{2,n_2-1}^{(-i)} = \binom{n_2 - 1}{2}^{-1} \sum_{1 \le j \le k \le n_2, j,k \ne i} h_2(\boldsymbol{Y}_j, \boldsymbol{Y}_k).$$

Those jackknife pseudo values are asymptotically independent ([15], [33]) and

$$U_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{V}_i^{(1)},$$
$$U_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} \hat{V}_i^{(2)}.$$

Under H_0, $\mathbb{E}(V_i^{(1)}) = \mathbb{E}(V_i^{(2)}) := \theta_0$, where $\theta_0 = \mathbb{E}\|\boldsymbol{X}_1 - \boldsymbol{X}_2\|$.

To apply JEL to these jackknife pseudo values, let $\boldsymbol{p} = (p_1, ..., p_{n_1})^T$ and $\boldsymbol{q} = (q_1, ..., q_{n_2})^T$ be probability vectors. Then the maximum constricted jackknife empirical likelihood is

$$L = \max_{\boldsymbol{p}, \boldsymbol{q}, \theta} \left\{ \left(\prod_{i=1}^{n_1} p_i \right) \left(\prod_{j=1}^{n_2} q_j \right) \right\},$$

subject to the following constraints

$$p_i \ge 0, \ i = 1, ..., n_1; \ \sum_{i=1}^{n_1} p_i = 1; \ \sum_{i=1}^{n_1} p_i(\hat{V}_i^{(1)} - \theta) = 0,$$
$$q_j \ge 0, \ j = 1, ..., n_2; \ \sum_{j=1}^{n_2} q_j = 1; \ \sum_{j=1}^{n_2} q_j(\hat{V}_j^{(2)} - \theta) = 0.$$

Utilizing the Lagrange multiplier approach, we have

$$p_i = \frac{1}{n_1} \frac{1}{1 + \lambda_1(\hat{V}_i^{(1)} - \theta)}, \quad i = 1, ..., n_1,$$
$$q_j = \frac{1}{n_2} \frac{1}{1 + \lambda_2(\hat{V}_j^{(2)} - \theta)}, \quad j = 1, ..., n_2,$$

where $(\lambda_1, \lambda_2, \theta)^T$ satisfies

$$\sum_{i=1}^{n_1} \frac{\hat{V}_i^{(1)} - \theta}{1 + \lambda_1 \left[\hat{V}_i^{(1)} - \theta \right]} = 0,$$

$$\sum_{j=1}^{n_2} \frac{\hat{V}_j^{(2)} - \theta}{1 + \lambda_2 \left[\hat{V}_j^{(2)} - \theta \right]} = 0,$$

$$\lambda_1 \sum_{i=1}^{n_1} \frac{-1}{1 + \lambda_1 \left[\hat{V}_i^{(1)} - \theta \right]} + \lambda_2 \sum_{j=1}^{n_2} \frac{-1}{1 + \lambda_2 \left[\hat{V}_j^{(2)} - \theta \right]} = 0. \tag{4}$$

Therefore, we can rewrite the empirical likelihood function as

$$L = \prod_{i=1}^{n_1} \left(\frac{1}{n_1} \frac{1}{1 + \lambda_1(\hat{V}_i^{(1)} - \theta)} \right) \prod_{j=1}^{n_2} \left(\frac{1}{n_2} \frac{1}{1 + \lambda_2(\hat{V}_j^{(2)} - \theta)} \right).$$

Further, $\left(\prod_{i=1}^{n_1} p_i \right) \left(\prod_{j=1}^{n_2} q_j \right)$ is maximized at $p_i = \dfrac{1}{n_1}$ and $q_j = \dfrac{1}{n_2}$, $i = 1, ..., n_1, j = 1, ..., n_2$. Thus, the jackknife empirical log-likelihood ratio is

$$l = 2 \sum_{i=1}^{n_1} \log \left\{ 1 + \lambda_1 \left[\hat{V}_i^{(1)} - \theta \right] \right\} + 2 \sum_{i=1}^{n_2} \log \left\{ 1 + \lambda_2 \left[\hat{V}_i^{(2)} - \theta \right] \right\},$$

where $(\lambda_1, \lambda_2, \theta)^T$ satisfies (4).

Define $g_1(\boldsymbol{x}) = \mathbb{E}h_1(\boldsymbol{x}, \boldsymbol{X}_2) - \theta_0$, $\sigma_{g1}^2 = \text{var}\left(g_1(\boldsymbol{X}_1) \right)$, $g_2(\boldsymbol{y}) = \mathbb{E}h_2(\boldsymbol{y}, \boldsymbol{Y}_2) - \theta_0$, $\sigma_{g2}^2 = \text{var}\left(g_2(\boldsymbol{Y}_1) \right)$. We assume

- **C1.** $0 < \sigma_{g_k} < \infty$, $k = 1, 2$;

- **C2.** $n = n_1 + n_2$, $\dfrac{n_k}{n} \to \alpha_k > 0$, $k = 1, 2$ and $\alpha_1 + \alpha_2 = 1$.

Condition **C1** implies that (1) the first moment of F exists and (2) F is not a point mass distribution. Condition **C2** means that the size of one part would not dominate over the other in the partition. When $n \to \infty$, $\min(n_1, n_2) \to \infty$. With those two conditions, Wilks' theorem holds.

Theorem 2.1 *Under H_0 and the conditions **C1-C2**, we have*

$$l \xrightarrow{d} \chi_1^2, \quad as\ n \to \infty.$$

Proof. See the Appendix. ∎

With Theorem 2.1, we reject H_0 if the observed jackknife empirical log-likelihood ratio $\hat{l}$ is greater than $\chi_{1,1-\alpha}^2$, where $\chi_{1,1-\alpha}^2$ is the $100(1-\alpha)\%$ quantile of χ^2 distribution with 1 degree of freedom. The p-value of the test can be calculated by

$$p\text{-value} = P_{H_0}(\chi_1^2 > \hat{l}),$$

and the power of the test is

$$\text{power} = P_{H_a}(l > \chi_{1,1-\alpha}^2).$$

In the next theorem, we establish the consistency of the proposed test, which states that its power is tending to 1 as the sample size goes to infinity.

Theorem 2.2 *Under the conditions **C1-C2**, the proposed JEL test is consistent for any fixed alternative. That is,*

$$P_{H_a}(l > \chi^2_{1,1-\alpha}) \to 1, \quad as\ n \to \infty.$$

Proof. See the Appendix. ∎

Remark 2.1 *One can test whether the distribution is symmetric about some specified center, $\boldsymbol{\mu}$, by letting $\boldsymbol{Z} = \boldsymbol{X} - \boldsymbol{\mu}$. Then testing $\boldsymbol{X} - \boldsymbol{\mu} \overset{D}{=} \boldsymbol{\mu} - \boldsymbol{X}$ is equivalent to testing $\boldsymbol{Z} \overset{D}{=} -\boldsymbol{Z}$. However, when the center point is unknown, the profiled JEL will be applied, which is definitely a worthwhile subject for our future research.*

3 Simulation Studies

In order to assess the proposed JEL method for the symmetry testing, we conduct simulation studies in this section. Empirical sizes and powers for each method at significance level $\alpha = 0.05$ are based on 10,000 replications. The results at level $\alpha = 0.10$ are similar and hence are skipped. We compare the following methods.

JEL: our proposed JEL method. R package "dfoptim" [37] is used for solving the equation system of (4).

ET: the DISCO test of [26]. The test is implemented by the permutation procedure. Function "eqdist.etest" with the default number of replicates in R package "energy" is used [27].

M: the symmetry test about an unknown median for univariate data based on the Bonferroni measure of skewness ([21]).

CM: the symmetry test about an unknown median for univariate data using $C = \dfrac{\bar{X} - M}{s}$ as the test statistic, where $\bar{X}$, M and s are the sample mean, median and standard deviation, respectively ([4]).

MGG: a robustified version of CM ([20]). The procedure "lawstat" in R package is used.

CD: the symmetric test based on the distance of characteristic functions ([6]). Similar to the ET method, a permutation procedure is applied and the number of permutations is set to the same as the ET method.

Note that the methods of M, CM and MGG are only for univariate distributions. Generating a sample from the population and then splitting the sample into two samples will be the same as generating two independent samples from the same distribution. Therefore, in this section, we generate two independent samples with either equal ($n_1 = n_2 = 20$ or $n_1 = n_2 = 50$) or unequal sample sizes ($n_1 = 40, n_2 = 60$) from a d-dimensional distribution F.

In order to assess the empirical sizes of the proposed JEL, we consider the following diagonal symmetric distributions. Results are reported in Table 1 for the univariate distributions and 2 for the multivariate distributions.

- Univariate case: $N(0,1)$, t_5, Laplace, Logistic distributions with location of 0 and scale of 1, and the mixture of two normal distributions with different standard deviations, 1 and 2, respectively.

- Multivariate cases $d = 2, 4, 6$:

    - $F = N(\boldsymbol{0}, \boldsymbol{I}_{d \times d})$ where $\boldsymbol{I}_{d \times d}$ is the identity matrix.

- $F = N(\mathbf{0}, \boldsymbol{\Sigma}_{d \times d})$, where $\boldsymbol{\Sigma}_{d \times d}$ is the matrix with the diagonal elements being 1 and the off-diagonal elements being 0.5.

- $F = t_5(\mathbf{0}, \boldsymbol{\Sigma}_{d \times d})$.

- F is a mixture of $N(\mathbf{0}, \boldsymbol{I}_{d \times d})$ and $N(\mathbf{0}, \boldsymbol{\Sigma}_{d \times d})$ with the mixing proportions of 0.9 and 0.1, respectively.

Table 1: Empirical sizes of each test on univariate distributions under the significance level 0.05.

Distribution	Sample Size	JEL	ET	MGG	CM	M	CD
N(0,1)	$n_1 = n_2 = 20$	.058	.044	.041	.037	.031	.055
	$n_1 = 40, n_2 = 60$	.048	.044	.049	.048	.044	.057
	$n_1 = n_2 = 50$	.048	.044	.049	.048	.044	.057
$t_5(0,1)$	$n_1 = n_2 = 20$	.075	.048	.057	.035	.032	.056
	$n_1 = 40, n_2 = 60$	.060	.049	.060	.042	.043	.058
	$n_1 = n_2 = 50$	.057	.043	.061	.041	.043	.054
Laplace(0,1)	$n_1 = n_2 = 20$	.078	.048	.062	.032	.032	.056
	$n_1 = 40, n_2 = 60$	.051	.048	.063	.033	.040	.057
	$n_1 = n_2 = 50$	.054	.048	.063	.033	.040	.057
Logistic(0,1)	$n_1 = n_2 = 20$	.069	.047	.046	.034	.032	.057
	$n_1 = 40, n_2 = 60$	.042	.048	.047	.037	.040	.058
	$n_1 = n_2 = 50$	.048	.048	.047	.037	.040	.058
$90\% N(0,1) + 10\% N(0,2)$	$n_1 = n_2 = 20$	.055	.043	.047	.038	.033	.053
	$n_1 = 40, n_2 = 60$	.041	.043	.058	.047	.047	.056
	$n_1 = n_2 = 50$	.041	.043	.054	.047	.045	.055

The proposed JEL is only compared with ET and CD approaches in Table 2 since other methods of MGG, CM and M are not valid for multivariate cases. From Tables 1 and 2 , we see that in the heavy-tailed distributions, although JEL has a slight oversize problem while CM and M are under-sized, those problems are not longer an issue when the sample size increases. Overall, the empirical sizes of all methods are fairly close to the nominal levels. Furthermore, the dimensions and the choices of n_1 and n_2 do not have too much impact on the control of the sizes. The proposed JEL test can control the sizes pretty well in different dimensions and in unequal cases.

In the second part of the simulation study, we consider the distributions not symmetric around $\mathbf{0}$ to assess the empirical powers. The following distributions in $\mathbb{R}^d$ ($d = 2, 4, 6$) are considered and results are reported in Table 3.

- $F = N(\mathbf{0.5}, \boldsymbol{I}_{d \times d})$, where $\mathbf{0.5}$ is the d-vector with all elements being 0.5.

- $F = N(\mathbf{0.5}, 16\boldsymbol{\Sigma}_{d \times d})$.

- F is a mixture of $N(\mathbf{0}, \boldsymbol{I}_{d \times d})$ and $N(\mathbf{0.5}, 16\boldsymbol{\Sigma}_{d \times d})$ with an equal mixing proportion.

- $F = t_5(\mathbf{0.5}, 16\boldsymbol{\Sigma}_{d \times d})$.

In Table 3, we observe that JEL test has a low power in $N(\mathbf{0.5}, \boldsymbol{I}_{d \times d})$, although it is consistent for any alternative. In that case, the ET and CD method achieve higher power than the JEL test. One intuitive explanation is that the JEL approach assigns more weights on the sample points close to $\mathbf{0}$ and hence the JEL loses some power to reject the symmetry. The same reason leads

6

Table 2: Empirical sizes of each test on multivariate distributions under the significance level 0.05.

Distribution	Dimension	Sample Size	JEL	ET	CD
$N(\mathbf{0}, \boldsymbol{I}_{d\times d})$	$d=2$	$n_1 = n_2 = 20$	.051	.046	.053
		$n_1 = 40, n_2 = 60$	.044	.043	.053
		$n_1 = n_2 = 50$	.045	.043	.055
	$d=4$	$n_1 = n_2 = 20$	.048	.045	.056
		$n_1 = 40, n_2 = 60$	.043	.043	.056
		$n_1 = n_2 = 50$	.046	.046	.054
	$d=6$	$n_1 = n_2 = 20$	.041	.045	.055
		$n_1 = 40, n_2 = 60$	.044	.044	.053
		$n_1 = n_2 = 50$	.047	.046	.055
$N(\mathbf{0}, \boldsymbol{\Sigma}_{d\times d})$	$d=2$	$n_1 = n_2 = 20$	.053	.046	.054
		$n_1 = 40, n_2 = 60$	.047	.046	.054
		$n_1 = n_2 = 50$	.045	.044	.055
	$d=4$	$n_1 = n_2 = 20$	.052	.042	.056
		$n_1 = 40, n_2 = 60$	.043	.046	.055
		$n_1 = n_2 = 50$	.045	.044	.054
	$d=6$	$n_1 = n_2 = 20$	.051	.046	.054
		$n_1 = 40, n_2 = 60$	.046	.043	.050
		$n_1 = n_2 = 50$	.044	.046	.055
$90\% N(\mathbf{0}, \boldsymbol{I}_{d\times d}) + 10\% N(\mathbf{0}, \boldsymbol{\Sigma}_{d\times d})$	$d=2$	$n_1 = n_2 = 20$	.050	.043	.052
		$n_1 = 40, n_2 = 60$	.043	.046	.055
		$n_1 = n_2 = 50$	.046	.041	.055
	$d=4$	$n_1 = n_2 = 20$	.046	.043	.054
		$n_1 = 40, n_2 = 60$	.045	.043	.054
		$n_1 = n_2 = 50$	.045	.040	.051
	$d=6$	$n_1 = n_2 = 20$	.045	.045	.057
		$n_1 = 40, n_2 = 60$	.047	.044	.051
		$n_1 = n_2 = 50$	.047	.045	.056
$t_5(\mathbf{0}, \boldsymbol{\Sigma}_{d\times d})$	$d=2$	$n_1 = n_2 = 20$	.066	.043	.051
		$n_1 = 40, n_2 = 60$	.055	.048	.058
		$n_1 = n_2 = 50$	.051	.044	.054
	$d=4$	$n_1 = n_2 = 20$	.102	.045	.058
		$n_1 = 40, n_2 = 60$	.051	.047	.051
		$n_1 = n_2 = 50$	.047	.046	.053
	$d=6$	$n_1 = n_2 = 20$	.180	.046	.054
		$n_1 = 40, n_2 = 60$	.068	.046	.053
		$n_1 = n_2 = 50$	.056	.046	.057

Table 3: Empirical powers of each test under the significance level 0.05.

Distribution	Dimension	Sample Size	JEL	ET	CD
$N(\mathbf{0.5}, \boldsymbol{I}_{d\times d})$	$d=2$	$n_1 = n_2 = 20$	.158	.973	.944
		$n_1 = 40, n_2 = 60$	.390	1.00	1.00
		$n_1 = n_2 = 50$	.416	1.00	1.00
	$d=4$	$n_1 = n_2 = 20$	.274	.999	.997
		$n_1 = 40, n_2 = 60$	.642	1.00	1.00
		$n_1 = n_2 = 50$	.686	1.00	1.00
	$d=6$	$n_1 = n_2 = 20$	.402	1.00	1.00
		$n_1 = 40, n_2 = 60$	.803	1.00	1.00
		$n_1 = n_2 = 50$	.846	1.00	1.00
$N(\mathbf{0.5}, 16\boldsymbol{\Sigma}_{d\times d})$	$d=2$	$n_1 = n_2 = 20$	.867	.119	.080
		$n_1 = 40, n_2 = 60$	.573	.248	.116
		$n_1 = n_2 = 50$	.535	.262	.117
	$d=4$	$n_1 = n_2 = 20$	.993	.137	.069
		$n_1 = 40, n_2 = 60$	.933	.287	.094
		$n_1 = n_2 = 50$	.904	.299	.093
	$d=6$	$n_1 = n_2 = 20$	.996	.144	.068
		$n_1 = 40, n_2 = 60$	.990	.310	.083
		$n_1 = n_2 = 50$	.979	.322	.080
$50\% N(\mathbf{0}, \boldsymbol{I}_{d\times d}) + 50\% N(\mathbf{0.5}, 16\boldsymbol{\Sigma}_{d\times d})$	$d=2$	$n_1 = n_2 = 20$	.521	.061	.059
		$n_1 = 40, n_2 = 60$	.220	.099	.062
		$n_1 = n_2 = 50$	.207	.109	.063
	$d=4$	$n_1 = n_2 = 20$	.934	.063	.057
		$n_1 = 40, n_2 = 60$	.822	.120	.053
		$n_1 = n_2 = 50$	.804	.120	.060
	$d=6$	$n_1 = n_2 = 20$	.987	.066	.058
		$n_1 = 40, n_2 = 60$	.942	.128	.055
		$n_1 = n_2 = 50$	.937	.132	.054
$t_5(\mathbf{0.5}, 16\boldsymbol{\Sigma}_{d\times d})$	$d=2$	$n_1 = n_2 = 20$	.972	.102	.075
		$n_1 = 40, n_2 = 60$	.936	.204	.116
		$n_1 = n_2 = 50$	.932	.204	.114
	$d=4$	$n_1 = n_2 = 20$	.990	.118	.079
		$n_1 = 40, n_2 = 60$	.994	.233	.101
		$n_1 = n_2 = 50$	.992	.239	.102
	$d=6$	$n_1 = n_2 = 20$	.988	.127	.070
		$n_1 = 40, n_2 = 60$	.995	.260	.097
		$n_1 = n_2 = 50$	.997	.262	.096

to low powers for the JEL test in detecting location differences of two distributions [29]. Such a phenomenon is also common for the tests based on the density function approach, as mentioned in [19]. However, when the covariance is departure from the identity, especially for high dimensions, there are less sample points around **0**. JEL is much more powerful than the ET and CD approaches. One can clearly find that the ET and CD approaches are dull to detect the asymmetry when the scatter matrix of the distribution is not identity, and the JEL method is efficient to detect this type of asymmetry, especially for high dimensions. It can also be seen clearly from the results that the empirical powers increase as the dimension increases from $d = 2$ to $d = 6$, which is consistent with the results in [6] and [35].

4 Conclusion

In this paper, we have developed a consistent JEL test for diagonal symmetry based on energy statistics. The standard limiting chi-square distribution with one degree of freedom is established and is used to conduct hypothesis testings without a permutation procedure. Numerical studies confirm the advantages of the proposed method under a variety of situations. One of important contributions of this paper is to develop a powerful nonparametric method for diagonal symmetry test. This can also be extended to test the central symmetry around some specified centers. However, if the center is not provided, we need to estimate the center, which is definitely a worthwhile subject for future research.

5 Appendix

Define $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$,

$$W_{1n}(\theta, \boldsymbol{\lambda}) = \frac{1}{n} \sum_{i=1}^{n_1} \frac{\hat{V}_i^{(1)} - \theta}{1 + \lambda_1 \left[\hat{V}_i^{(1)} - \theta \right]},$$

$$W_{2n}(\theta, \boldsymbol{\lambda}) = \frac{1}{n} \sum_{i=1}^{n_2} \frac{\hat{V}_i^{(2)} - \theta}{1 + \lambda_2 \left[\hat{V}_i^{(2)} - \theta \right]},$$

$$W_{3n}(\theta, \boldsymbol{\lambda}) = \frac{1}{n} \lambda_1 \sum_{i=1}^{n_1} \frac{-1}{1 + \lambda_1 \left[\hat{V}_i^{(1)} - \theta \right]} + \frac{1}{n} \lambda_2 \sum_{i=1}^{n_2} \frac{-1}{1 + \lambda_2 \left[\hat{V}_i^{(2)} - \theta \right]}.$$

Lemma 5.1 (Hoeffding, 1948) *If* $\mathbb{E}\left[h_1^2(X_1, X_2)\right] < \infty$, $\mathbb{E}\left[h_2^2(Y_1, Y_2)\right] < \infty$, $\sigma_{g_1}^2 > 0$, $\sigma_{g_2}^2 > 0$ *then*
$$\frac{\sqrt{n_1}(U_1 - \theta_0)}{2\sigma_{g_1}} \xrightarrow{d} N(0,1) \quad as \ n_1 \to \infty,$$

$$\frac{\sqrt{n_2}(U_2 - \theta_0)}{2\sigma_{g_2}} \xrightarrow{d} N(0,1) \quad as \ n_2 \to \infty.$$

Lemma 5.2 *Let* $S_j = \frac{1}{n_j} \sum_{i=1}^{n} \left[\hat{V}_i^{(j)} - \theta_0 \right]^2$, $j = 1, 2$. *Under the conditions of Lemma 5.1,*

$$S_j = 4\sigma_{g_j}^2 + o_P(1), \ n_j \to \infty, \ j = 1, 2.$$

Proof. See Lemma A.3 in [15]. ∎

9

Lemma 5.3 (Liu, Liu and Zhou, 2018) *Under the conditions of Lemma 5.1, and under H_0, with probability tending to one as $n \to \infty$, there exists a root $\tilde{\theta}$ of*

$$W_{jn}(\theta, \boldsymbol{\lambda}) = 0, \ \ j = 1, 2, 3,$$

such that $|\tilde{\theta} - \theta_0| < \delta$, where $\delta = n^{-1/3}$.

Let $\tilde{\eta} = (\tilde{\theta}, \tilde{\boldsymbol{\lambda}})^T$ be the solution to the above equations, and $\eta_0 = (\theta_0, 0, 0)^T$. By expanding $W_{jn}(\tilde{\eta})$ at η_0, we have, for $j = 1, 2, 3$,

$$0 = W_{jn}(\eta_0) + \frac{\partial W_{jn}}{\partial \theta}(\eta_0)(\tilde{\theta} - \theta_0) + \frac{\partial W_{jn}}{\partial \lambda_1}(\eta_0)\tilde{\lambda}_1 + \frac{\partial W_{jn}}{\partial \lambda_2}(\eta_0)\tilde{\lambda}_2 + R_{jn}, \tag{5}$$

where $R_{jn} = \dfrac{1}{2}(\tilde{\eta} - \eta_0)^T \dfrac{\partial^2 W_{jn}(\eta^*)}{\partial \eta \partial \eta^T}(\tilde{\eta} - \eta_0)$ where η^* lies between η_0 and $\tilde{\eta}$. One can easily prove that $R_{jn} = o_P(n^{-1/2})$.

Lemma 5.4 $Cov(U_1, U_2) = 0$.

Proof of Theorem 2.1

By Lemma 5.2 and equation (5),

$$\begin{pmatrix} W_{1n}(\eta_0)) \\ W_{2n}(\eta_0) \\ 0 \end{pmatrix} = W \begin{pmatrix} \tilde{\lambda}_1 \\ \tilde{\lambda}_2 \\ \tilde{\theta} - \theta_0 \end{pmatrix} + o_P(n^{-1/2}),$$

where

$$W = \begin{pmatrix} \alpha_1 \sigma_1^2 & 0 & \alpha_1 \\ 0 & \alpha_2 \sigma_2^2 & \alpha_2 \\ \alpha_1 & \alpha_2 & 0 \end{pmatrix},$$

$\sigma_j^2 = 4\sigma_{g_j}^2, j = 1, 2$.

W is nonsingular. Therefore,

$$\begin{pmatrix} \tilde{\lambda}_1 \\ \tilde{\lambda}_2 \\ \tilde{\theta} - \theta_0 \end{pmatrix} = W^{-1} \begin{pmatrix} W_{1n}(\eta_0) \\ W_{2n}(\eta_0) \\ 0 \end{pmatrix} + o_P(n^{-1/2}).$$

Under H_0, $\sigma_1^2 = \sigma_2^2 := \sigma^2$.

$$\tilde{\theta} - \theta_0 = W_{1n}(\eta_0) + W_{2n}(\eta_0) + o_p(n^{-1/2}).$$

By [15], we have

$$\tilde{\lambda}_j = \frac{U_j - \tilde{\theta}}{\tilde{S}_j} + o_P(n^{-1/2}), \ \ j = 1, 2,$$

where $\tilde{S}_j = \dfrac{1}{n_j} \sum_{i=1}^{n_j} (\hat{V}_i^{(j)} - \tilde{\theta})^2$. It can be easily to check that $\tilde{S}_j = \sigma_j^2 + o_p(1)$, $j = 1, 2$. By the proof of Theorem 1 in [15],

$$l(\theta_0) = \left[\sum_{j=1}^{2} n_j \frac{(U_j - \tilde{\theta})^2}{\sigma_j^2} \right] (1 + o_P(1)).$$

10

By simple algebra,

$$\sum_{j=1}^{2} n_j \frac{(U_j - \tilde{\theta})^2}{\sigma_j^2} = (\sqrt{n}W_{1n}(\eta_0), \sqrt{n}W_{2n}(\eta_0),) \times \boldsymbol{A}^T \boldsymbol{D} \boldsymbol{A} \times (\sqrt{n}W_{1n}(\eta_0), \sqrt{n}W_{2n}(\eta_0))^T + o_p(1),$$

(6)

where

$$\boldsymbol{D} = \begin{pmatrix} \alpha_1/\sigma_1^2 & 0 \\ 0 & \alpha_2/\sigma_2^2 \end{pmatrix}$$

and

$$\boldsymbol{A} = \begin{pmatrix} 1/\alpha_1 - 1 & -1 \\ -1 & 1/\alpha_2 - 1 \end{pmatrix}.$$

By Hoeffding decomposition and Lemma 5.4,

$$\sqrt{n} \begin{pmatrix} W_{1n}(\eta_0) \\ W_{2n}(\eta_0) \end{pmatrix} \to N(\boldsymbol{0}, \boldsymbol{\Sigma}),$$

where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \alpha_1 \sigma_1^2 & 0 \\ 0 & \alpha_2 \sigma_2^2 \end{pmatrix}.$$

Hence, under H_0, the empirical log likelihood ratio converges in distribution to $\sum_{i=1}^{2} \omega_i \chi_i^2$, where $\chi_i^2, i = 1, 2$ are two independent chi-square random variables with one degree of freedom, and $\omega_i, i = 1, 2$ are eigenvalues of $\boldsymbol{\Sigma}_0^{1/2} \boldsymbol{A}^T \boldsymbol{D}_0 \boldsymbol{A} \boldsymbol{\Sigma}^{1/2}$, where

$$\boldsymbol{\Sigma}_0 = \sigma^2 \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}$$

and

$$\boldsymbol{D}_0 = \frac{1}{\sigma^2} \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}.$$

After a simple algebra calculation, $\boldsymbol{\Sigma}_0^{1/2} \boldsymbol{A}^T \boldsymbol{D}_0 \boldsymbol{A} \boldsymbol{\Sigma}_0^{1/2}$ has two eigenvalues 0, 1. □

Proof of Theorem 2.2 Under H_a, $\mathbb{E}U_k$, $k = 1, 2$ will be different. Let $\mathbb{E}U_k = \theta_k$, $k = 1, 2$. From (6),

$$
\begin{aligned}
l &= \sum_{k=1}^{2} \frac{n_k(U_k - \tilde{\theta})^2}{\tilde{S}_i^2} + o(1) \\
&= \sum_{k=1}^{2} \left[\frac{\sqrt{n_k}(U_k - \theta_k)}{\tilde{S}_k} + \frac{\sqrt{n_k}(\theta_k - \tilde{\theta})}{\tilde{S}_k} \right]^2 + o(1),
\end{aligned}
$$

which will be divergent since at least one of $\dfrac{\sqrt{n_k}(\theta_k - \tilde{\theta})^2}{\tilde{S}_k}, k = 1, 2$ will diverge to ∞. □

References

[1] Antille, A., Kersting, G., and Zucchini, W. (1982). Tests symmetry. *J. Amer. Statist. Assoc.*, **77**, 639-646.

[2] Bhattacharya, P., Gastwirth, J. Kersting, G., and Wright, A. (1982). Two modified Wilcoxon test for symmetry about an unknown location parameter. *Biometrika*, **69**, 377-382.

[3] Boos, D. (1982). A test for symmetry associated with the Hodges-Lehmann estimator. *J. Amer. Statist. Assoc.*, **77**, 647-651.

[4] Cabilio, P. and Masaro, J. (1996). A simple test of symmetry about an unknown median. *Canad. J. Statis.*, **24**(3), 349-361.

[5] Cao, R. and Van Keilegom, I. (2006). Empirical likelihood tests for two-sample problems via nonparametric density estimation. *Can. J. Stat.*, **34**(1), 61-77.

[6] Chen, F., Meintanis, S.G. and Zhu, L. (2019). On some characterizations and multidimensional criteria for testing homogeneity, symmetry and independence. *J. Multivariate Anal.*, **173**, 125-144.

[7] Chen, J., Variyath, A. M. and Abraham, B. (2008). Adjusted empirical likelihood and its properties. *J. Comput. Graph. Statist.* **17** (2), 426-443.

[8] Doksum, K., Fenstad, G., and Aaberge, R. (1977). Plots and tests for symmetry. *Biometrika*, **64**(3), 473-487.

[9] Einmahl, J. and McKeague, I. (2003). Empirical likelihood based hypothesis testing. *Bernoulli*, **9**(2), 267-290.

[10] Einmahl, J. and Gan, Z. (2016). Testing for central symmetry. *J. Statist. Plann. Inference*, **169**, 27-33.

[11] Ghosh, S. and Ruymgaart, F.H. (1992). Applications of empirical characteristic in some multivariate problems. *Canad. J. Statis.*, **20**, 429-440.

[12] Heathcote, C.R., Rachev, S.T. and Cheng, B. (1995). Testing multivariate symmetry. *J. Multivariate Anal.*, **54**, 91-112.

[13] Henze, N., Klar, B. and Meintanis, S.G. (2003). Invariant tests for symmetry about an unspecified point based on empirical characteristic function. *J. Multivariate Anal.*, **87**, 275-297.

[14] Hollander, M. (1988). Testing for symmetry, in: S. Kotz, N.L. Johnson (Eds). *Encyclopedia Statist. Sci.*, Wiley, New York, 211-216.

[15] Jing, B., Yuan, J. and Zhou, W. (2009). Jackknife empirical likelihood. *J. Amer. Statist. Assoc.* **104**, 1224-1232.

[16] Koutrouvelis, I. (1985). Distribution-free procedures for location and symmetry inference problems based on the empirical characteristic function. *Scand. J. Statist.* **12**, 257-269.

[17] Koziol, J. (1985). A note on testing symmetry with estimated parameters. *Statist. Prob. Lett.* **3**, 227-230.

[18] Liu, Y., Liu, Z. and Zhou, W. (2018). A test for equality of two sample distributions via integrating characteristic functions. *Statist. Sinica*, DOI: 10.5705/ss.202017.0236.

[19] Martìnez-Camblor, P. and de Uña-Álvarez, J. (2009). Non-parametric *k*-sample tests: Density functions vs distribution functions. *Comput. Statist. Data Anal.* **53**, 3344-3357.

[20] Miao, W., Gel, Y.R., and Gastwirth, J.L. (2006). A new test of symmetry about an unknown median. In Hsiung A, Zhang C, Ying Z (eds.), *Random Walk, Sequential Analysis and Related Topics-A Festschrift in Honor of Yuan-Shih Chow* , 199-214. World Scientific Publisher, Singapore.

[21] Mira, A. (1999). Distribution-free test for symmetry based on Bonferroni's measure. *J. App. Stat.*, **26**(8), 959-972.

[22] Neuhaus, G. and Zhu, L.X. (1998). Permutation test for reflected symmetry. *J. Multivariate Anal.*, **67**, 129-153.

[23] Owen, A. (1988). Empirical likelihood ratio confidence intervals for single functional. *Biometrika* **75**, 237-249.

[24] Owen, A. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18**, 90-120.

[25] Randles, R.H., Flinger, M.A., Policello, G.E., Wolfe, G.A. (1980). An asymptotically distribution-free test for symmetry versus asymmetry. *J. Amer. Statist. Assoc. Ann.*, **75**, 168-172.

[26] Rizzo, M.L. and Székely, G.J. (2010). Disco Analysis: A nonparametric extension of analysis of variance. *Ann. Appl. Stat.* **4**, 1034-1055.

[27] Rizzo, M. and Székely, G.J. (2017). E-statistics: multivariate inference via the energy of data. R package version 1.7-2. https://CRAN.R-project.org/package=energy.

[28] Sang, Y., Dang, X. and Zhao, Y. (2019). Jackknife empirical likelihood methods for Gini correlation and their equality testing. *J. Statist. Plann. Inference*, **199**, 45-59.

[29] Sang, Y., Dang, X. and Zhao, Y. (2019). Jackknife empirical likelihood approach for K-sample tests. Submitted. https://arxiv.org/abs/1908.00477.

[30] Schuster, E. and Barker, R. (1987). Using the bootstrap in testing symmetry versus asymmetry. *Comm. Statist. Simul. Comp.* **16**, 69-84.

[31] Serfling, R.J. (2006). Multivariate symmetry and asymmetry. *Encyclopedia Statist. Sci.*, **8**, 5338-5345.

[32] Sen, P.K. (1977). Tied down Wiener process approximations for aligned rank order statistics and some applications. *Ann. Statist.*, **5**, 1107-1123.

[33] Shi, X. (1984). The approximate independence of jackknife pseudo-values and the bootstrap methods. *Journal of Wuhan Institute Hydra-Electric Engineering*, **2**, 83-90.

[34] Székely, G.J. and Móri, T.F. (2001). A characteristic measure of asymmetry and its application for testing diagonal symmetry. *Commun. Stat. Theory Methods*, **30**:8-9, 1633-1639.

[35] Székely, G.J. and Rizzo, M.L. (2013). Energy statistics: A class of statistics based on distances. *J. Stat. Plan. Infer.*, **143**, 1249-1272.

[36] Székely, G.J. and Rizzo, M.L. (2017). The energy of data, *Ann. Rev. Stat. Appl.*, **4** (1), 447-479.

[37] Varadhan R., Borchers, H.W. and Varadhan, M. R.(2018). R.: Package "dfoptim", available at: https://cran.r-project.org/web/packages/ dfoptim/index.html.

[38] Wan,Y., Liu, Z. and Deng, M. (2018). Empirical likelihood test for equality of two distributions using distance of characteristic functions. *Statistics*, **52** (6), 1379-1394

[39] Zhang, J. and Wu, Y. (2007). K-sample tests based on the likelihood ratio. *Comput. Statist. Data Anal.* **51**, 4682-4691.