Statistics and Applications {ISSN 2454-7395 (online)}
Special Issue in Memory of Prof. C R Rao
Volume 22, No. 3, 2024 (New Series), pp 45–66
http://www.ssca.org.in/journal

Some Novel Limiting Distributions Arising in Order Restricted Inference

Sayan Ghosh[1] and Ori Davidov[2]
[1]*Department of Mathematics, Birla Institute of Technology and Sciences Pilani, Hyderabad Campus, Hyderabad 500078, India*
[2]*Department of Statistics, University of Haifa, Mount Carmel, Haifa 3498838, Israel*

Received: 21 Dember 2023; Revised: 9 February 2024; Accepted: 23 February 2024

Abstract

In this paper, we develop a methodology for testing the hypothesis that the true value of a parameter lies in the union of multiple cones against the alternative that it does not. We propose a test statistic for such problems and derive its novel asymptotic null distribution. The least favourable asymptotic null value and the corresponding least favourable asymptotic null distribution are obtained. The proposed test is uniformly more powerful than conventional tests discussed in the literature. Some illustrative examples are provided and a simulation study evaluating its performance is presented.

Key words: Hypothesis Testing; Convex Cones; Least Favourable Configuration; Asymptotic Distribution.

AMS Subject Classifications: 62F03, 62F30

1. Introduction

Testing problems are typically formulated as $H_0 : \boldsymbol{\theta} \in \boldsymbol{\Theta}_0$ versus $H_1 : \boldsymbol{\theta} \in \boldsymbol{\Theta}_1$. Usually, the null $\boldsymbol{\Theta}_0$ as well as the alternative $\boldsymbol{\Theta}_1$ are simple sets such as singletons or linear spaces. There are however various applications in which the null and/or the alternative are more complicated sets. In particular, we consider testing

$$H_0 : \boldsymbol{\theta} \in \bigcup_{i=1}^{K} \mathcal{C}_i \quad \text{versus} \quad H_1 : \boldsymbol{\theta} \notin \bigcup_{i=1}^{K} \mathcal{C}_i, \tag{1}$$

where $\mathcal{C}_1, \ldots, \mathcal{C}_K$ are arbitrary distinct convex cones in $\mathbb{R}^m$ defined by systems of linear inequalities. In this paper we address the case where $m = 2$. Some comments on the corresponding theory for general m are deferred to Section 7. It is further assumed that there exists an unconstrained estimator $\boldsymbol{S}_n$ for $\boldsymbol{\theta} \in \mathbb{R}^2$ such that as $n \to \infty$

$$\sqrt{n}(\boldsymbol{S}_n - \boldsymbol{\theta}) \Rightarrow \mathcal{N}_2(\mathbf{0}, \boldsymbol{\Sigma}) \tag{2}$$

Corresponding Author: Ori Davidov
Email: davidov@stat.haifa.ac.il

where $\Rightarrow$ denotes convergence in distribution and $\boldsymbol{\Sigma}$ is a positive definite matrix. As indicated below, there are many problems of interest that can be formulated as in (1).

Robertson and Wegman (1978) were among the first to test hypotheses of the type (1) but with $K = 1$. This setting is known in the literature as *testing against an ordering* and classified by Silvapulle and Sen (2004) as a Type B problem. In general the union of convex cones is neither convex nor a cone. Therefore Type B problems are a simple special case of (1). Other special cases of (1) in $\mathbb{R}^2$ have also been addressed in the literature. Berger and Sinclair (1984) examined the problem of testing a null hypothesis that the parameter of interest belongs to a union of linear subspaces which they applied to the testing of symmetric spacings among ordered normal means. Another paper involving linear spaces is by Berger (1997) who tested $H_0 : \min\{|\theta_1|, |\theta_2|\} = 0$ against $H_1 : \min\{|\theta_1|, |\theta_2|\} > 0$. Thus under the null the pair (θ_1, θ_2) lies on the axes whereas under the alternative it does not. If θ_i measures the effect of treatment i then the null states that at least one treatment has no effect whereas under the alternative both treatments have effects.

This paper is organized as follows. In Section 2, we discuss the preliminaries related to testing (1). We introduce relevant notations and setup for the testing problem. Then we define the proposed test statistic and show that it is identical to the likelihood ratio test statistic and to the intersection union test statistic for (1) in some cases. In Section 3 we consider the problem of testing (1) for two quadrants in $\mathbb{R}^2$. We obtain the least favourable null values and the least favourable null distribution of the proposed test statistic for finite samples. In Section 4, we consider the union of multiple distinct arbitrary convex cones in $\mathbb{R}^2$ and obtain the least favourable null values along with the least favourable null distribution of the proposed test statistic for large samples. In Section 5, some examples and testing problems in $\mathbb{R}^2$ are provided as illustration. A simulation study is performed to evaluate the proposed test in Section 6. Finally, in Section 7 we provide a brief summary of our work and some possible extensions.

2. Preliminaries

We begin with some notations. Let $\Pi_{\boldsymbol{\Sigma}}(\boldsymbol{S} \mid \mathcal{C})$ denote the projection of $\boldsymbol{S}$ onto $\mathcal{C}$ with respect to $\boldsymbol{\Sigma}$ and let $\|\boldsymbol{S}\|^2_{\boldsymbol{\Sigma}}$ be the respective norm. Note that when $\boldsymbol{\Sigma} = \boldsymbol{I}$ the latter reduces to the usual projection and the standard euclidean distance. For the hypotheses in (1), we propose the test statistic

$$T_n = \min\{n\|\boldsymbol{S}_n - \Pi_{\boldsymbol{\Sigma}}(\boldsymbol{S}_n \mid \mathcal{C}_1)\|^2_{\boldsymbol{\Sigma}}, \ldots, n\|\boldsymbol{S}_n - \Pi_{\boldsymbol{\Sigma}}(\boldsymbol{S}_n \mid \mathcal{C}_K)\|^2_{\boldsymbol{\Sigma}}\}, \tag{3}$$

where $\boldsymbol{S}_n$ was described in (2). In general the variance matrix $\boldsymbol{\Sigma}$ is unknown in which case T_n is computed with respect to a consistent estimator $\boldsymbol{\Sigma}_n$ thereof. It is clear that T_n essentially minimizes the squared distance between $\boldsymbol{S}_n$ and $\boldsymbol{\theta}$ over various values of $\boldsymbol{\theta}$ in $\boldsymbol{\Theta}_0$. The following result shows the relationship between the proposed test, the likelihood ratio test (LRT) and the intersection union test (IUT).

Theorem 1: If $\boldsymbol{S}$ follows a $\mathcal{N}_2(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ distribution with known $\boldsymbol{\Sigma}$ then the statistic (3) as a function of $\boldsymbol{S}$ is the LRT statistic for the hypotheses in (1). Moreover, (3) is the IUT statistic if and only if the cones $\mathcal{C}_1, \ldots, \mathcal{C}_K$ are all congruent.

Theorem 1 provides a meaningful motivation for using the statistic (3) when (2) holds as $n \to \infty$ as well as in situations in which $\boldsymbol{\Sigma}$ is unknown but can be consistently estimated from the data. Although under the stated conditions the LRT and the IUT statistics coincide, their critical values are in general different. Further note that if we set $\boldsymbol{G}_n = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{S}_n$ then $\sqrt{n}(\boldsymbol{G}_n - \boldsymbol{\eta}) \Rightarrow \mathcal{N}_2(\boldsymbol{0}, \boldsymbol{I})$ where $\boldsymbol{\eta} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\theta}$. In addition testing the hypotheses (1) using $\boldsymbol{S}_n$ is equivalent to testing

$$H_0 : \boldsymbol{\eta} \in \bigcup_{i=1}^{K} \mathcal{C}_i^* \quad \text{versus} \quad H_1 : \boldsymbol{\eta} \notin \bigcup_{i=1}^{K} \mathcal{C}_i^*$$

using $\boldsymbol{G}_n$ where $\mathcal{C}_i^* = \boldsymbol{\Sigma}^{-1/2}\mathcal{C}_i = \{\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\theta} : \boldsymbol{\theta} \in \mathcal{C}_i\}$ are the transformed cones. Therefore without any loss of generality we will henceforth primarily consider the case where $\boldsymbol{\Sigma} = \boldsymbol{I}$.

3. The case of two cones

We start by investigating the important special case of two quadrants. Let $\mathcal{C}_1 = \{\boldsymbol{\theta} \in \mathbb{R}^2 : \theta_1 \geq 0, \theta_2 \geq 0\}$ denote the positive quadrant and let $\mathcal{C}_2 = \{\boldsymbol{\theta} \in \mathbb{R}^2 : \theta_1 \leq 0, \theta_2 \leq 0\}$ denote the negative quadrant. Consider testing the hypotheses

$$H_0 : \boldsymbol{\theta} \in \mathcal{C}_1 \cup \mathcal{C}_2 \quad \text{against} \quad H_1 : \boldsymbol{\theta} \notin \mathcal{C}_1 \cup \mathcal{C}_2 \tag{4}$$

using a single observation $\boldsymbol{S} = (S_1, S_2)^T$ from $\mathcal{N}_2(\boldsymbol{\theta}, \boldsymbol{I})$. Note that under the null θ_1 and θ_2 are either both non-negative or both non-positive. This problem is of independent interest.

Theorem 2: Suppose that $\boldsymbol{S}$ follows $\mathcal{N}_2(\boldsymbol{\theta}, \boldsymbol{I})$. Then the LRT statistic for (4) is

$$T = \min\{S_1^2, S_2^2\}\mathbb{I}(\boldsymbol{S} \notin \mathcal{C}_1 \cup \mathcal{C}_2). \tag{5}$$

Furthermore for all $c \geq 0$ we have

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_0} \mathbb{P}_{\boldsymbol{\theta}}(T \geq c) = \frac{1}{2}\mathbb{P}(\chi_0^2 \geq c) + \frac{1}{2}\mathbb{P}(\chi_1^2 \geq c). \tag{6}$$

Equation (6), where χ_i^2 is a chi–square RV with i degrees of freedom and $\chi_0^2 \equiv 0$, provides us with a formula with which we can compute the p–values associated with the test statistic (5). The value of $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$ for which (6) holds is called the *least favourable configuration* or null value. The distribution of the statistic T when $\boldsymbol{\theta}$ is the least favourable is called the *least favourable null distribution*. The proof of Theorem 2 shows that the least favourable configurations are of the form $(0, \pm\infty)$ and $(\pm\infty, 0)$, *i.e.*, they lie on the axes at an infinite distance from the origin while the least favourable null distribution of T is given by (6). It follows that for any other value of $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$ and any c

$$\mathbb{P}_{\boldsymbol{\theta}}(T \geq c) < \frac{1}{2}\mathbb{P}(\chi_0^2 \geq c) + \frac{1}{2}\mathbb{P}(\chi_1^2 \geq c).$$

Letting $T(\theta_1, \theta_2)$ denote the LRT statistic at (θ_1, θ_2) we can restate the conclusion of Theorem 2 in the language of stochastic order relations (Shaked and Shanthikumar (2007)) as $T(\theta_1, \theta_2) \preceq_{st} T(0, \pm\infty)$ and $T(\theta_1, \theta_2) \preceq_{st} T(\pm\infty, 0)$ where $\preceq_{st}$ denotes the usual stochastic order. Both relations hold for all $(\theta_1, \theta_2) \in \boldsymbol{\Theta}_0$. It can also be shown that $T(0, 0) \preceq_{st} T(0, \theta_2)$

and $T(0,0) \preceq_{st} T(\theta_1, 0)$. In particular, $T(0,0)$ is distributed as $(1/2)\chi_0^2 + (1/2)\min\{Q_1, Q_2\}$ where Q_1 and Q_2 are independent χ_1^2 RVs.

In the proof of Theorem 2 a closed form expression for $\mathbb{P}_{\boldsymbol{\theta}}(T \geq c)$ was found facilitating the analysis and enabling one to find the least favourable configuration and null distribution. In general though, $\mathbb{P}_{\boldsymbol{\theta}}(T \geq c)$ is not amenable to a simple analysis nor is it given by a simple formula. Consequently, an asymptotic analysis yielding workable formulas of the type (6) is necessary.

4. The general case

In this section a general asymptotic theory for multiple cones is developed. First note that any convex cone in $\mathbb{R}^2$ is of the form

$$\mathcal{C} = \text{conic}(\boldsymbol{u}, \boldsymbol{v}) = \{\lambda_1 \boldsymbol{u} + \lambda_1 \boldsymbol{v} : \lambda_1 \geq 0, \lambda_2 \geq 0\},$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ are unit vectors lying on the extreme rays of $\mathcal{C}$. Further note that the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$, *i.e.*, $\measuredangle(\boldsymbol{u}, \boldsymbol{v})$ is smaller than π.

Let $\mathcal{C}_1, \ldots, \mathcal{C}_K$ be K distinct convex cones where $\mathcal{C}_i = \text{conic}(\boldsymbol{u}_i, \boldsymbol{v}_i)$ for $i = 1, \ldots, K$. For convenience it is further assumed that for $\boldsymbol{e}_1 = (1,0)^T$ we have:

$$\measuredangle(\boldsymbol{e}_1, \boldsymbol{u}_i) < \measuredangle(\boldsymbol{e}_1, \boldsymbol{v}_i)$$

for all i and

$$\measuredangle(\boldsymbol{e}_1, \boldsymbol{u}_1) < \measuredangle(\boldsymbol{e}_1, \boldsymbol{u}_2) < \cdots < \measuredangle(\boldsymbol{e}_1, \boldsymbol{u}_K).$$

Thus the cone $\mathcal{C}_1$ is the cone whose rays make the smallest angle with the positive real axis followed by the cone $\mathcal{C}_2$, and so forth. Similarly within each cone the ray associated with $\boldsymbol{u}_i$ has a smaller angle than the ray $\boldsymbol{v}_i$.

We say cones $\mathcal{C}_i$ and $\mathcal{C}_j$ are adjacent if the interior of the cone $\text{conic}(\boldsymbol{v}_i, \boldsymbol{u}_j)$ is a subset of $\boldsymbol{\Theta}_1$. The angle between $\boldsymbol{v}_i$ and $\boldsymbol{u}_j$ may be smaller than $\pi/2$, between $\pi/2$ and π or larger than π. If $\measuredangle(\boldsymbol{v}_i, \boldsymbol{u}_j) \leq \pi/2$, we set

$$\mathcal{R}_{ij} = \text{conic}(\boldsymbol{v}_i, \boldsymbol{u}_j). \tag{7}$$

If $\pi/2 < \measuredangle(\boldsymbol{v}_i, \boldsymbol{u}_j) \leq \pi$, we further divide the cone $\text{conic}(\boldsymbol{v}_i, \boldsymbol{u}_j)$ into three conic regions

$$\mathcal{R}_i(\boldsymbol{v}_i) = \text{conic}(\boldsymbol{v}_i, \boldsymbol{u}_{j*}), \quad \mathcal{R}'_{ij} = \text{conic}(\boldsymbol{u}_{j*}, \boldsymbol{v}_{i*}), \quad \mathcal{R}_j(\boldsymbol{u}_j) = \text{conic}(\boldsymbol{v}_{i*}, \boldsymbol{u}_j) \tag{8}$$

where $\boldsymbol{u}_{j*}, \boldsymbol{v}_{i*} \in \text{conic}(\boldsymbol{v}_i, \boldsymbol{u}_j)$, $\boldsymbol{u}_{j*}$ is orthogonal to $\boldsymbol{u}_j$ and $\boldsymbol{v}_{i*}$ is orthogonal to $\boldsymbol{v}_i$. Finally if $\measuredangle(\boldsymbol{v}_i, \boldsymbol{u}_j) > \pi$, then we divide the region bounded by $\boldsymbol{u}_i$ and $\boldsymbol{v}_j$ into three conic regions

$$\mathcal{R}_i(\boldsymbol{u}_i) = \text{conic}(\boldsymbol{u}_i, \boldsymbol{u}_{i*}), \quad \mathcal{R}''_{ij} = \text{conic}(\boldsymbol{u}_{i*}, \boldsymbol{v}_{j*}), \quad \mathcal{R}_j(\boldsymbol{v}_j) = \text{conic}(\boldsymbol{v}_{j*}, \boldsymbol{v}_j) \tag{9}$$

where $\boldsymbol{u}_{i*}, \boldsymbol{v}_{j*} \in \text{conic}(\boldsymbol{u}_i, \boldsymbol{v}_j)$, $\boldsymbol{u}_{i*}$ is orthogonal to $\boldsymbol{u}_i$ and $\boldsymbol{v}_{j*}$ is orthogonal to $\boldsymbol{v}_j$.

Remark 1: If $K = 2$ the cones $\mathcal{C}_1$ and $\mathcal{C}_2$ are doubly adjacent. Moreover, if $\measuredangle(\boldsymbol{v}_1, \boldsymbol{u}_2) \leq \pi/2$ and $\measuredangle(\boldsymbol{v}_2, \boldsymbol{u}_1) \leq \pi/2$ then we label the regions between the cones by $\mathcal{R}_{12}$ and $\mathcal{R}_{21}$. The modification when the above mentioned angles are larger than $\pi/2$ is obvious.

The following result provides the number of possible regions between the cones or between their polar cones for various geometric arrangements of the cones.

Lemma 1: Let N, N' and N'' denote the number of regions of the type $\mathcal{R}_{ij}$, $\mathcal{R}'_{ij}$ and $\mathcal{R}''_{ij}$ respectively. Then $N + N' + N'' = K$ where $N \leq K$, $N' \leq 3$ and $N'' \leq 1$. In particular, if $N'' = 1$, then $N \leq K - 1$ and $N' \leq 1$.

It is well known that $\Pi_{\boldsymbol{I}}(\boldsymbol{S} \mid \mathcal{C}_i) = \mathbf{0}$ if and only if $\boldsymbol{S} \in \mathcal{C}_i^0$ where $\mathcal{C}_i^0$ denotes the polar cone of $\mathcal{C}_i$. Thus, it follows that $\Pi_{\boldsymbol{I}}(\boldsymbol{S} \mid \bigcup_{i=1}^{K} \mathcal{C}_i) = \mathbf{0}$ if and only if $\boldsymbol{S} \in \bigcap_{i=1}^{K} \mathcal{C}_i^0$. This event has a positive probability if the set $\bigcap_{i=1}^{K} \mathcal{C}_i^0 \supset \{\mathbf{0}\}$ and zero probability if $\bigcap_{i=1}^{K} \mathcal{C}_i^0 = \{\mathbf{0}\}$. In the first case we denote the set $\bigcap_{i=1}^{K} \mathcal{C}_i^0$ by $\mathcal{R}''_{pq}$ as defined above whereas in the latter case we set $\mathcal{R}''_{pq} = \emptyset$. In other words:

$$N'' = \begin{cases} 0 & \text{if } \bigcap_{i=1}^{K} \mathcal{C}_i^0 = \{\mathbf{0}\} \\ 1 & \text{if } \bigcap_{i=1}^{K} \mathcal{C}_i^0 \supset \{\mathbf{0}\} \end{cases},$$

We now introduce some useful additional notations.

Definition 1: Let $\mathcal{R} = \text{conic}(\boldsymbol{u}, \boldsymbol{v})$ and denote its interior angle by $\gamma = \measuredangle(\boldsymbol{u}, \boldsymbol{v})$ where $0 < \gamma \leq \pi/2$. Let $\boldsymbol{S}$ be a $\mathcal{N}_2(\mathbf{0}, \boldsymbol{I})$ RV. Then conditional on $\boldsymbol{S} \in \mathcal{R}$ we define

$$\chi^2_{1,1}(\gamma) = \mathrm{d}^2(\boldsymbol{S}, \mathrm{bd}(\mathcal{R})) = \min\{\mathrm{d}^2(\boldsymbol{S}, \mathrm{ray}(\boldsymbol{u})), \mathrm{d}^2(\boldsymbol{S}, \mathrm{ray}(\boldsymbol{v}))\} \tag{10}$$

where $\mathrm{d}(\cdot, \cdot)$ is the euclidean distance and $\mathrm{bd}(\mathcal{R})$ is the boundary of the cone $\mathcal{R}$ defined by the rays $\mathrm{ray}(\boldsymbol{u})$ and $\mathrm{ray}(\boldsymbol{v})$.

For example when $\mathcal{R}$ is any quadrant then $\gamma = \pi/2$ and $\mathbb{P}(\chi^2_{1,1}(\pi/2) \geq c) = [\mathbb{P}(\chi^2_1 \geq c)]^2$. When $0 < \gamma < \pi/2$ we have the following, numerically simple to evaluate formula.

Lemma 2: For $c \geq 0$ and $\mathcal{R}$ as in definition 1

$$\mathbb{P}(\chi^2_{1,1}(\gamma) \geq c, \boldsymbol{S} \in \mathcal{R}) = \frac{\gamma}{2\pi}(P_1 + P_2 + P_3 + P_4) \tag{11}$$

where $P_1 = \mathbb{P}(D_1 \geq \sqrt{c}, D_2 \geq \sqrt{c})$, $P_2 = \mathbb{P}(D_1 \geq \sqrt{c}, D_2 \leq -\sqrt{c})$, $P_3 = \mathbb{P}(D_1 \leq -\sqrt{c}, D_2 \geq \sqrt{c})$, $P_4 = \mathbb{P}(D_1 \leq -\sqrt{c}, D_2 \leq -\sqrt{c})$ and $(D_1, D_2)^T$ has a bivariate normal distribution with mean $\mathbf{0}$, unit variances and correlation $-\cos(\gamma)$.

Now let $\mathcal{C}_i$ and $\mathcal{C}_j$ be adjacent cones with an angle $0 < \delta \leq \pi$ between their boundaries. By definition 1, if $0 < \delta \leq \pi/2$ then $\mathcal{R} = \mathcal{R}_{ij}$ and $\gamma = \delta$. However if $\pi/2 < \delta \leq \pi$ then $\mathcal{R} = \mathcal{R}'_{ij}$ and $\gamma = \pi - \delta$. Conditional on $\boldsymbol{S} \in \mathcal{R}$, in both cases we have

$$\chi^2_{1,1}(\gamma) = \min\{\mathrm{d}^2(\boldsymbol{S}, \mathrm{ray}(\boldsymbol{u}_j)), \mathrm{d}^2(\boldsymbol{S}, \mathrm{ray}(\boldsymbol{v}_i))\} = \min\{(\boldsymbol{u}_{j*}^T \boldsymbol{S})^2, (\boldsymbol{v}_{i*}^T \boldsymbol{S})^2\}.$$

Moreover in Lemma 2, $D_1 = \boldsymbol{u}_{j*}^T \boldsymbol{S}$ and $D_2 = \boldsymbol{v}_{i*}^T \boldsymbol{S}$ so the correlation coefficient between them is $-\cos(\gamma)$ if $0 < \delta \leq \pi/2$ and $\cos(\gamma)$ if $\pi/2 < \delta \leq \pi$.

Next we consider angles associated with the cones in (1) and the regions in (7)-(9). Let $\rho_1, \ldots, \rho_K$ denote the interior angles of the cones $\mathcal{C}_1, \ldots, \mathcal{C}_K$ and set $\rho = \sum_{i=1}^K \rho_i$. By assumption $\rho > 0$. The interior angles of the cones $\mathcal{R}_{ij}$, $\mathcal{R}'_{ij}$ and $\mathcal{R}''_{ij}$, as defined in (7), (8) and (9) respectively, are all denoted by γ_{ij}. Let $\mathcal{P}$ denote the set of indices (i,j) for all pairs of adjacent cones $\mathcal{C}_i$ and $\mathcal{C}_j$ except (p,q). It is clear that for all $(i,j) \in \mathcal{P}$ we have $0 < \gamma_{ij} \leq \pi/2$. Furthermore, if $\mathcal{R}''_{pq} \neq \emptyset$, then $0 < \gamma_{pq} < \pi$ denotes the interior angle of $\mathcal{R}''_{pq}$, otherwise $\gamma_{pq} = 0$. Following (8) let $\tau_i(\boldsymbol{v}_i)$ and $\tau_j(\boldsymbol{u}_j)$ be the interior angles of the cones $\mathcal{R}_i(\boldsymbol{v}_i)$ and $\mathcal{R}_j(\boldsymbol{u}_j)$ respectively. Similarly, if $\gamma_{pq} > 0$, *i.e.*, when (9) holds let $\tau_p(\boldsymbol{u}_p)$ and $\tau_q(\boldsymbol{v}_q)$ be the interior angles of $\mathcal{R}_p(\boldsymbol{u}_p)$ and $\mathcal{R}_q(\boldsymbol{v}_q)$ respectively. Finally set $\tau = \sum_{(i,j)\in\mathcal{P}}(\tau_i(\boldsymbol{v}_i)+\tau_j(\boldsymbol{u}_j))+\tau_p(\boldsymbol{u}_p)+\tau_q(\boldsymbol{v}_q)$. Note that $\tau_i(\boldsymbol{v}_i), \tau_j(\boldsymbol{u}_j) < \pi/2$ for all $(i,j) \in \mathcal{P}$. Moreover $\tau_p(\boldsymbol{u}_p) = \tau_q(\boldsymbol{v}_q) = \pi/2$. Also $\tau = 0$ if and only only if $\gamma_{ij} \leq \pi/2$ for all $(i,j) \in \mathcal{P}$ and $\gamma_{pq} = 0$. We have

$$\rho + \sum_{(i,j)\in\mathcal{P}} \gamma_{ij} + \tau + \gamma_{pq} = 2\pi. \tag{12}$$

We are now ready to state the main results of this section.

Theorem 3: Consider $\boldsymbol{\Theta}_0$ in (1) and the statistic T_n in (3). If $n \to \infty$, then we have

$$T_n \Rightarrow \begin{cases} \chi_0^2 & \text{if } \boldsymbol{\theta} \in \text{int}(\boldsymbol{\Theta}_0) \\ \frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2 & \text{if } \boldsymbol{\theta} \in \text{ray}(\boldsymbol{\Theta}_0) \\ \frac{\rho}{2\pi}\chi_0^2 + \sum_{(i,j)\in\mathcal{P}} \frac{\gamma_{ij}}{2\pi}\chi_{1,1}^2(\gamma_{ij}) + \frac{\tau}{2\pi}\chi_1^2 + \frac{\gamma_{pq}}{2\pi}\chi_2^2 & \text{if } \boldsymbol{\theta} = \boldsymbol{0} \end{cases} \tag{13}$$

where $\text{ray}(\boldsymbol{\Theta}_0)$ is the collection of all rays generating the cones in $\boldsymbol{\Theta}_0$.

Theorem 3 provides the limiting distribution of the LRT statistic for various values of $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$. Let $T_{\boldsymbol{I}}$, $T_{\boldsymbol{R}}$ and $T_{\boldsymbol{O}}$ denote the limits of the LRT statistic when $\boldsymbol{\theta}$ is in the interior of $\boldsymbol{\Theta}_0$, on a ray of $\boldsymbol{\Theta}_0$ and the origin, respectively. For the form of the corresponding limits, see Equation (13) in the statement of Theorem 3. Clearly $T_{\boldsymbol{I}} \equiv 0$ so both $T_{\boldsymbol{O}}$ and $T_{\boldsymbol{R}}$ are stochastically larger than $T_{\boldsymbol{I}}$. It follows that the least favourable configuration and the limiting least favourable distribution are not associated with the interior points of $\boldsymbol{\Theta}_0$.

Remark 2: Note that if $\boldsymbol{S}_n$ is normally distributed then the distribution of T_n at $\boldsymbol{\theta} = \boldsymbol{0}$, which we denote by $T_{\boldsymbol{O}}$, is exact.

Suppose now that a size α test is desired. Let $c_{\alpha,\boldsymbol{R}}$ and $c_{\alpha,\boldsymbol{O}}$ denote the size α critical values associated with $T_{\boldsymbol{R}}$ and $T_{\boldsymbol{O}}$ respectively. These values solve the equations

$$\mathbb{P}(T_{\boldsymbol{R}} \geq c_{\alpha,\boldsymbol{R}}) = \alpha, \quad \text{and} \quad \mathbb{P}(T_{\boldsymbol{O}} \geq c_{\alpha,\boldsymbol{O}}) = \alpha.$$

Incidentally, it is easy to see that $c_{\alpha,\boldsymbol{R}}$ is equal to $(1-2\alpha)$-quantile of the χ_1^2 distribution whereas it may be necessary to compute $c_{\alpha,\boldsymbol{O}}$ numerically. Clearly the overall limiting critical value of the test is

$$c_\alpha = \min\{c_{\alpha,\boldsymbol{R}}, c_{\alpha,\boldsymbol{O}}\}.$$

In principle, finding the appropriate limiting critical value for any α is easy. It is worth noting that there are many cases in which we have either $c_{\alpha,\boldsymbol{R}} > c_{\alpha,\boldsymbol{O}}$ or $c_{\alpha,\boldsymbol{R}} < c_{\alpha,\boldsymbol{O}}$ for all

$0 \leq \alpha \leq 1$. The first situation arises when $T_{\boldsymbol{R}} \succeq_{\text{st}} T_{\boldsymbol{O}}$ whereas the second situation arises when $T_{\boldsymbol{R}} \preceq_{\text{st}} T_{\boldsymbol{O}}$. If either order relation holds then finding the limiting critical value is immediate. However, there are situations where an ordering does not exist, *i.e.*, $c_{\alpha,\boldsymbol{R}} > c_{\alpha,\boldsymbol{O}}$ for some values of α and $c_{\alpha,\boldsymbol{R}} < c_{\alpha,\boldsymbol{O}}$ for others. To summarize, for any $c \geq 0$:

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_0} \lim_{n\to\infty} \mathbb{P}_{\boldsymbol{\theta}}(T_n \geq c) = \begin{cases} \mathbb{P}(T_{\boldsymbol{R}} \geq c) & \text{if } T_{\boldsymbol{O}} \preceq_{\text{st}} T_{\boldsymbol{R}} \\ \mathbb{P}(T_{\boldsymbol{O}} \geq c) & \text{if } T_{\boldsymbol{R}} \preceq_{\text{st}} T_{\boldsymbol{O}} \\ \max\{\mathbb{P}(T_{\boldsymbol{R}} \geq c), \mathbb{P}(T_{\boldsymbol{O}} \geq c)\} & \text{otherwise} \end{cases}. \tag{14}$$

Equation (14) helps us to compute the limiting p–values associated with the test statistic (5). The value of $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$ for which (14) holds is called the *least favourable limiting null value* of T_n. The distribution of the statistic T_n when $\boldsymbol{\theta}$ is the least favourable is called the *least favourable limiting null distribution.* In the first case of (14), any point on a ray in ray$(\boldsymbol{\Theta}_0)$ is the least favourable limiting null value of T_n and $\mathbb{P}(T_{\boldsymbol{R}} \geq c)$ is the least favourable limiting null distribution. In the second case, the origin is the least favourable limiting null value of T_n and $\mathbb{P}(T_{\boldsymbol{O}} \geq c)$ is the least favourable limiting null distribution. In the third case, their union is the least favourable limiting null value of T_n and $\max\{\mathbb{P}(T_{\boldsymbol{R}} \geq c), \mathbb{P}(T_{\boldsymbol{O}} \geq c)\}$ is the least favourable limiting null distribution. The next result shows that the least favourable limiting null distribution of the LRT statistic T_n is determined by the geometry of the cones.

Theorem 4: The least favourable limiting null distribution of (3) for testing (1) is that of $T_{\boldsymbol{O}}$ if and only if $\tau \geq \pi$ and that of $T_{\boldsymbol{R}}$ if and only if $\rho \geq \pi$.

Next we revisit the LRT and IUT for (1) in $\mathbb{R}^2$. By Theorem 1 the LRT and IUT statistics coincide if and only if all cones are congruent. As discussed earlier, the LRT rejects the null hypothesis if $T_n > c_\alpha$ where $c_\alpha = \min\{c_{\alpha,\boldsymbol{R}}, c_{\alpha,\boldsymbol{O}}\}$. The IUT rejects the null if and only if for all $i \in \{1, \ldots, K\}$ we find that $\Lambda^{(i)} > c_\alpha^{(i)}$ where $\Lambda^{(i)}$ is the LRT and $c_\alpha^{(i)}$ is the critical value for testing $H_0^{(i)} : \boldsymbol{\theta} \in \mathcal{C}_i$ against $H_1^{(i)} : \boldsymbol{\theta} \notin \mathcal{C}_i$. Thus the IUT combines K Type B problems in each of which the least favourable null value is the origin. Hence, we reject $H_0^{(i)}$ if $\Lambda^{(i)}$ is larger than the $1-\alpha$ quantile of the RV

$$\frac{\rho_i}{2\pi}\chi_0^2 + \frac{1}{2}\chi_1^2 + \frac{\pi - \rho_i}{2\pi}\chi_2^2.$$

For example, consider testing the hypotheses in (4). By Theorems 3 and 4, the null in (1) is rejected if T_n is larger than the $1-\alpha$ quantile of the RV $\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$. Since the quadrants are congruent it is easy to see that the IUT rejects the null only if T_n is larger than $(1-\alpha)$ quantile of the RV $\frac{1}{4}\chi_0^2 + \frac{1}{2}\chi_1^2 + \frac{1}{4}\chi_2^2$, which is larger than the critical value of the LRT. The following result compares the LRT and the IUT for cones in two dimensions.

Theorem 5: The LRT for (1) is asymptotically uniformly more powerful than the IUT for cones in $\mathbb{R}^2$.

Numerical examples illustrating Theorem 5 are given in Section 6.

5. Some examples and testing problems in $\mathbb{R}^2$

We begin this section by providing some synthetic examples that exemplify our notations and illustrate the applications of Theorems 3 and 4. The synthetic examples are followed by examples of problems analyzed in the literature.

5.1. Synthetic examples

In Figure 1 several examples, depicting various geometric settings, are displayed.

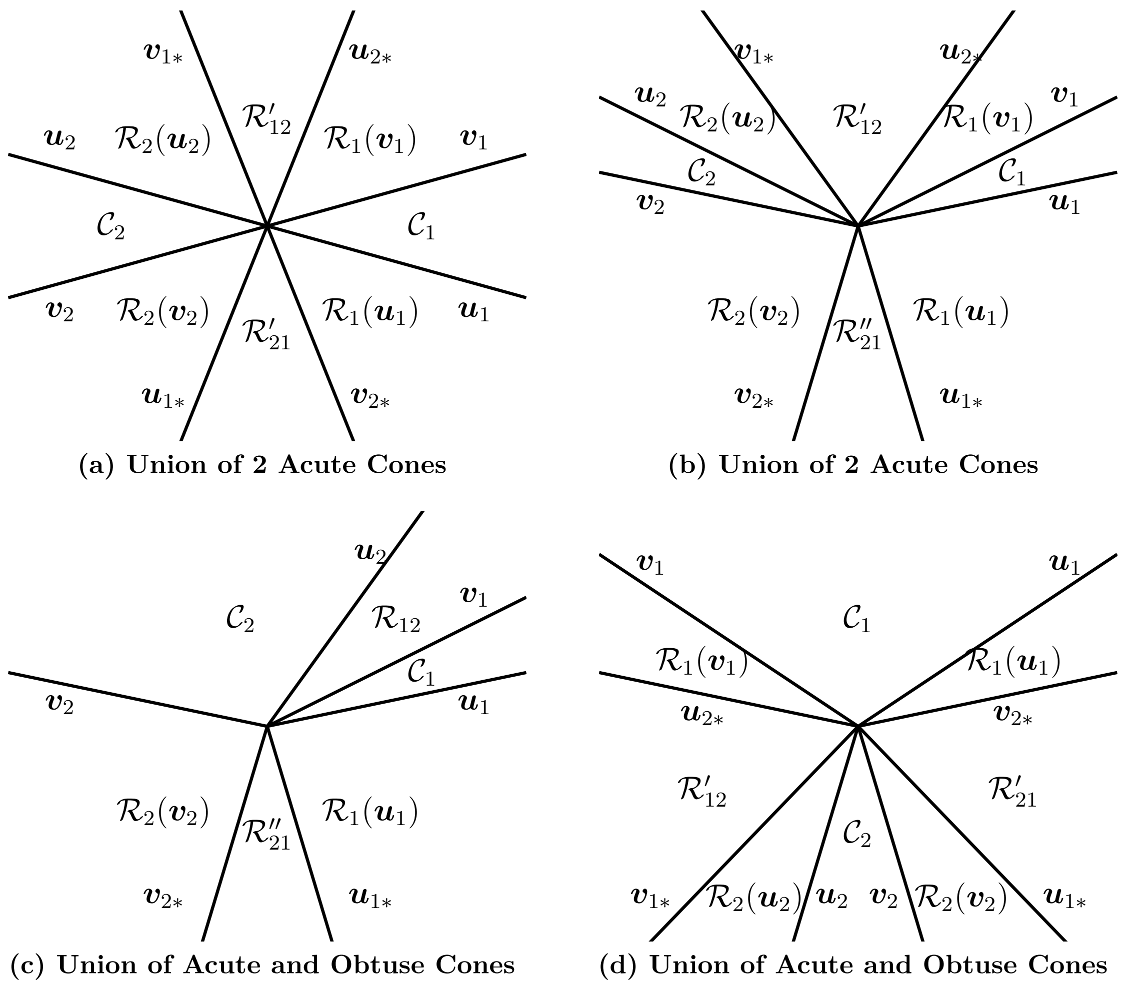


Figure 1: Partition of $\mathbb{R}^2$ by cones $\mathcal{C}_1$ and $\mathcal{C}_2$

Example 1: (**Union of Acute Cones I, Figure 1(A)**): Here $\mathcal{R}''_{21} = \emptyset$ so $\gamma_{21} = 0$. The interior angles ρ_1 and ρ_2 are both smaller than $\pi/2$ so $\rho < \pi$. If $\tau \geq \pi$, then by Theorem 4 the least favourable limiting null value of T_n is the origin and the least favourable limiting null distribution is that of $T_{\boldsymbol{O}}$. However if $\tau < \pi$ then we have an indeterminate case where there is no stochastic ordering between $T_{\boldsymbol{O}}$ and $T_{\boldsymbol{R}}$.

Example 2: (**Union of Acute Cones II, Figure 1(B)**): Here $\mathcal{R}''_{21} \neq \emptyset$ so $\gamma_{21} > 0$ and again $\rho < \pi$. Moreover $\tau_1(\boldsymbol{u}_1) = \tau_2(\boldsymbol{v}_2) = \pi/2$ so it is clear that $\tau > \pi$. Hence by Theorem 4, the least favourable limiting null value of T_n is the origin and the least favourable limiting null distribution is that of $T_{\boldsymbol{O}}$.

Example 3: (**Union of Acute and Obtuse Cones I, Figure 1(C)**): Here $\mathcal{R}''_{21} \neq \emptyset$ so $\gamma_{21} > 0$ and also $\rho < \pi$. Moreover $\tau_1(\boldsymbol{u}_1) = \tau_2(\boldsymbol{v}_2) = \pi/2$ so $\tau = \pi$. Hence by Theorem 4, the least favourable limiting null value of T_n is the origin and the least favourable limiting null distribution is that of $T_{\boldsymbol{O}}$.

Example 4: (Union of Acute and Obtuse Cones II, Figure 1(D)): Here $\mathcal{R}''_{21} = \emptyset$ so $\gamma_{21} = 0$. Let $\pi/2 < \rho_1 < \pi$ and $\rho_2 < \pi/2$ denote the interior angles. If $\rho \geq \pi$, then by Theorem 4, the least favourable limiting null value of T_n lies in ray($\boldsymbol{\Theta}_0$) and the least favourable limiting null distribution is that of $T_{\boldsymbol{R}}$. However if $\rho < \pi$, then we have an indeterminate case where there is no stochastic ordering between $T_{\boldsymbol{O}}$ and $T_{\boldsymbol{R}}$.

Next consider the case where $K \geq 2$. In particular, suppose each cone has interior angle η. Suppose further that angles between all adjacent cones are also equal. It follows that $\rho = K\eta$ and the angle between adjacent cones is $(2\pi - K\eta)/K$. Of course, it is assumed that $0 < \rho < 2\pi$. Clearly the angle between the adjacent cones is always smaller than π and therefore $\gamma_{pq} = 0$. Further note that τ is positive if and only if

$$\eta < \frac{2\pi}{K} - \frac{\pi}{2} \quad \text{and} \quad K \in \{2, 3\}. \tag{15}$$

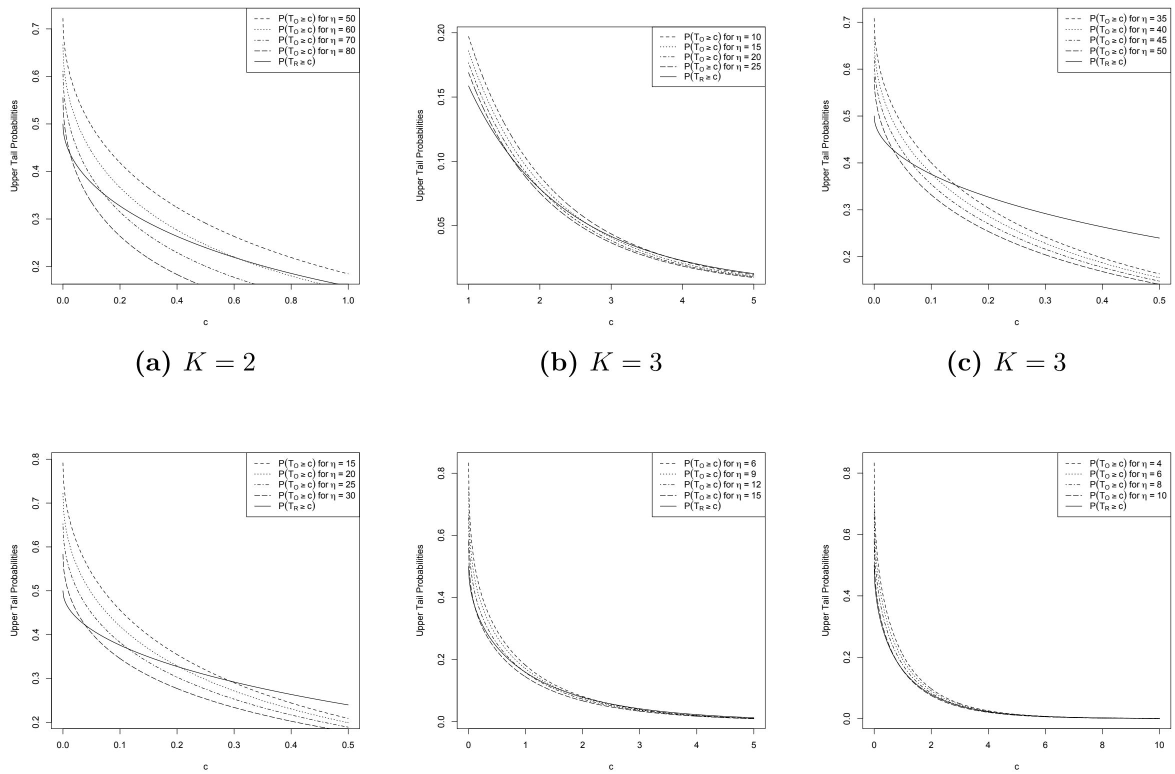
Therefore if τ is positive, then $\tau = 2(\pi - 2\eta)$ when $K = 2$ and $\tau = \pi - 6\eta$ when $K = 3$. By Theorem 3 the limiting null distribution of T_n at the origin, *i.e.*, $T_{\boldsymbol{O}}$ is

$$T_{\boldsymbol{O}} \stackrel{d}{=} \begin{cases} \frac{\eta}{\pi}\chi_0^2 + \frac{\eta}{\pi}\chi_{1,1}^2(\eta) + \frac{\pi - 2\eta}{\pi}\chi_1^2 & \text{if } K = 2 \text{ and } \eta < \frac{\pi}{2} \\ \frac{3\eta}{2\pi}\chi_0^2 + \frac{3\eta + \pi}{2\pi}\chi_{1,1}^2(\eta + \pi/3) + \frac{\pi - 6\eta}{2\pi}\chi_1^2 & \text{if } K = 3 \text{ and } \eta < \frac{\pi}{6} \\ \frac{K\eta}{2\pi}\chi_0^2 + \frac{2\pi - K\eta}{2\pi}\chi_{1,1}^2(2\pi/K - \eta) & \text{otherwise} \end{cases}. \tag{16}$$

Now we investigate the relationship between $T_{\boldsymbol{O}}$ given in (16) and $T_{\boldsymbol{R}}$ as given in (13). If $\pi/K \leq \eta < 2\pi/K$ then $\rho \geq \pi$ so by Theorem 4, the least favourable limiting null distribution of T_n is that of $T_{\boldsymbol{R}}$. Further note that $\tau \geq \pi$ if and only if $\eta \leq \pi/4$ and $K = 2$ in which case the least favourable limiting null distribution of T_n is that of $T_{\boldsymbol{O}}$. For any other values of η and K, Theorem 4 can not be used to identify the least favourable null distribution and the corresponding size α critical value so this determination must be made numerically as illustrated in Figure 2.

Figure 2(A)-(F) present plots of the tail probabilities $\mathbb{P}(T_{\boldsymbol{O}} \geq c)$ and $\mathbb{P}(T_{\boldsymbol{R}} \geq c)$ for various values of $c > 0$ and choices of K and η. In particular K and η were chosen so $\rho < \pi$. In addition $\tau < \pi$ in Figures 2(A) and 2(B) whereas $\tau = 0$ in Figures 2(C)-(F). It is clear that for the above choices the tail probabilities cross and consequently the RVs $T_{\boldsymbol{R}}$ and $T_{\boldsymbol{O}}$ are not stochastically ordered. In other words, the least favourable limiting null distribution of T_n is not the same for all size α critical values. It is also clear that whenever the tail probabilities of $T_{\boldsymbol{R}}$ and $T_{\boldsymbol{O}}$ cross then there exists a unique value c^* satisfying $\mathbb{P}(T_{\boldsymbol{O}} \geq c) \geq \mathbb{P}(T_{\boldsymbol{R}} \geq c)$ for all $c \leq c^*$ whereas $\mathbb{P}(T_{\boldsymbol{R}} \geq c) \geq \mathbb{P}(T_{\boldsymbol{O}} \geq c)$ for all $c \geq c^*$. Moreover for a fixed K, c^* is monotonically decreasing in η. Similarly if η is fixed, then c^* is monotonically decreasing in K. In the majority of cases plotted we find that $c_\alpha = c_{\alpha,\boldsymbol{R}}$.

Remark 3: Note that when $\pi/2 < \eta < \pi$, the cones are obtuse and $K \leq 3$. Thus the intersection of their polar cones contains only the origin and hence $\gamma_{pq} = 0$. Since $\rho > \pi$, by Theorem 4 the least favourable limiting null values of T_n lie in ray($\boldsymbol{\Theta}_0$) and the least favourable limiting null distribution is that of $T_{\boldsymbol{R}}$ which is the same as that for testing over union of two quadrants. Note that the null distribution of T_n at the origin is a function of dependent χ_1^2 RVs in the first case but a function of independent χ_1^2 RVs in the latter case.



(a) $K = 2$ **(b)** $K = 3$ **(c)** $K = 3$

(d) $K = 5$ **(e)** $K = 10$ **(f)** $K = 15$

Figure 2: Stochastic ordering between $T_{\boldsymbol{O}}$ and $T_{\boldsymbol{R}}$ for various no. of cones (K) and values of interior angles (η)

5.2. Examples from the literature

We conclude this section by providing some examples of relevant testing problems in $\mathbb{R}^2$ which have appeared in the literature. These examples illustrate the application of Theorems 3–5 and show the simplicity and superiority (in power) of our testing procedure as compared to those in the literature.

Example 5: Consider first testing the hypotheses $H_0 : \min\{|\theta_1|, |\theta_2|\} = 0$ against $H_1 : \min\{|\theta_1|, |\theta_2|\} > 0$. Variants of this problem have been studied by Cohen *et al.* (1983) and Berger (1997) where the IUT had been advocated. Note that by setting $\mathcal{C}_1 = \{\boldsymbol{\theta} : \theta_1 \geq 0, \theta_2 = 0\}$, $\mathcal{C}_2 = \{\boldsymbol{\theta} : \theta_1 = 0, \theta_2 \geq 0\}$, $\mathcal{C}_3 = \{\boldsymbol{\theta} : \theta_1 \leq 0, \theta_2 = 0\}$ and $\mathcal{C}_4 = \{\boldsymbol{\theta} : \theta_1 = 0, \theta_2 \leq 0\}$ we can reformulate the problem as in (1). Next, it is clear that the statistic (3) reduces to $T_n = \min\{S^2_{n,1}, S^2_{n,2}\}$. The least favourable null value and distribution are $\boldsymbol{\theta} = \boldsymbol{0}$ and $\chi^2_{1,1}(\pi/2)$ respectively and the critical value is the $\sqrt{1-\alpha}$ quantile of a χ^2_1 RV. Since the cones are congruent, T_n is the same as the IUT by Theorem 1 but more powerful than the IUT by Theorem 5.

Example 6: Laska and Meisner (1989) tested (1) with $\mathcal{C}_i = \{\boldsymbol{\theta} \in \mathbb{R}^m : \theta_i \leq 0\}$. In their formulation $\theta_i = \mu_0 - \mu_i$ where μ_0 is the mean response under treatment $\mathcal{T}_0$ and μ_i is the mean response under treatment $\mathcal{T}_i$. Thus under the null some treatments are superior to $\mathcal{T}_0$

whereas under the alternative all treatments are inferior to $\mathcal{T}_0$. This problem is known in the literature as the *sign testing problem* and has received considerable attention (e.g., Berger (1982) and Cohen *et al.* (1983)). It is easy to verify that when $m = 2$ this testing problem can be reformulated as $H_0 : \boldsymbol{\theta} \in \mathcal{Q}_2 \cup \mathcal{Q}_3 \cup \mathcal{Q}_4$ and $H_1 : \boldsymbol{\theta} \in \mathcal{Q}_1$ where $\mathcal{Q}_1, \ldots, \mathcal{Q}_4$ denote the quadrants of $\mathbb{R}^2$ in clockwise direction. Interestingly, this problem is the complement of a Type B problem since $\boldsymbol{\Theta}_1$ is a single convex cone. Berger (1982) proposed testing H_0 against H_1 using the IUT. The LRT for this problem is $T_n = \min\{S^2_{n,1}, S^2_{n,2}\}\mathbb{I}(\boldsymbol{S}_n \in \mathcal{Q}_1)$. Since $\rho = 3\pi/2$, by Theorem 4 the least favourable limiting null values of T_n are in ray($\boldsymbol{\Theta}_0$), *i.e.*, the rays defining the first quadrant, and the least favourable limiting null distribution is $\frac{1}{2}\chi^2_0 + \frac{1}{2}\chi^2_1$. By Theorem 5, the proposed test is more powerful than the IUT although they are identical by Theorem 1 due to congruence of the cones.

Example 7: Gail and Simon (1985) as well as Silvapulle (2001) tested (1) for $K = 2$ where $\mathcal{C}_1 = \{\boldsymbol{\theta} \in \mathbb{R}^m : \boldsymbol{\theta} \geq \mathbf{0}\}$ and $\mathcal{C}_2 = \{\boldsymbol{\theta} \in \mathbb{R}^m : \boldsymbol{\theta} \leq \mathbf{0}\}$. Here θ_i is the difference between the mean responses to treatments $\mathcal{T}_1$ and $\mathcal{T}_2$, say, in the i^{th} group where $i = 1, \ldots, m$. If $\mathcal{T}_1$ is more beneficial than $\mathcal{T}_2$ ($\theta_i \geq 0$) in some groups but more harmful than $\mathcal{T}_2$ ($\theta_i \leq 0$) in others, it is said that there is *crossover interaction* between treatments and groups. Thus under the null there is no crossover interaction whereas under the alternative, there is such interaction. The hypotheses of interest in $\mathbb{R}^2$ are $H_0 : \boldsymbol{\theta} \in \mathcal{C}_1 \cup \mathcal{C}_2$ and $H_1 : \boldsymbol{\theta} \notin \mathcal{C}_1 \cup \mathcal{C}_2$ where $\mathcal{C}_1$ and $\mathcal{C}_2$ are the non-negative and the non-positive quadrants. By Theorem 2, the test statistic is given by $T_n = \min\{S^2_{n,1}, S^2_{n,2}\}\mathbb{I}(\boldsymbol{S}_n \notin \mathcal{C}_1 \cup \mathcal{C}_2)$. Since $\rho = \pi$, by Theorem 4 the least favourable limiting null values of T_n lie in ray($\boldsymbol{\Theta}_0$) and the least favourable limiting null distribution is $\frac{1}{2}\chi^2_0 + \frac{1}{2}\chi^2_1$. Gail and Simon (1985) and Silvapulle (2001) did not assume that the variance is known but their statistic is of the same form as the LRT T_n.

Example 8: Berger (1989) and Liu and Berger (1995) tested (1) with $\mathcal{C}_i = \{\boldsymbol{\theta} \in \mathbb{R}^m : \boldsymbol{b}_i^T\boldsymbol{\theta} \leq 0\}$ where $\boldsymbol{\theta}$ is the mean vector of a multivariate normal distribution and $\boldsymbol{b}_i$s are non-redundant vectors. If $\boldsymbol{\theta}$ denotes a vector of means then under the null some linear combinations, e.g., contrasts, are negative whereas under the alternative, all linear combinations are positive. Consider the above problem in $\mathbb{R}^2$ where $\mathcal{C}_i = \{\boldsymbol{\theta} \in \mathbb{R}^2 : \boldsymbol{b}_i^T\boldsymbol{\theta} \leq 0\}$ (a half-space), $\boldsymbol{\theta}$ is the mean vector of a bivariate normal distribution and $\boldsymbol{b}_i$s are non-redundant vectors. Here $\boldsymbol{\Theta}_0$ is a union of multiple convex cones whereas $\boldsymbol{\Theta}_1$ is a single convex cone, which is the complement of a Type B problem. Berger (1989) and Liu and Berger (1995) applied the IUT to this problem. The LRT statistic T_n is given by (3). Since $\rho > \pi$, by Theorem 4 the least favourable asymptotic null values of T_n lie in ray($\boldsymbol{\Theta}_0$) and the least favourable asymptotic null distribution is $\frac{1}{2}\chi^2_0 + \frac{1}{2}\chi^2_1$. By Theorem 5, the proposed test based on T_n is uniformly more powerful than the IUT.

6. Simulation study

We performed a small simulation study comparing the power of the LRT to that of the IUT. We considered $K = 2$ cones and both congruent (C) as well as non-congruent (NC) pairs of cones. See Table 1 for the settings of the study. We fixed $n = 100$ and $\alpha = 0.05$. The critical values were computed by simulation at the least favourable null values. For computing power, the point in the alternative for union of quadrants is of the form (θ_1, θ_2) where θ_1 and θ_2 have different signs; otherwise it is of the form $(0, \theta_2)$ where $\theta_2 > 0$. These points were chosen so that the LRT has a power of around 0.8. From Table 2 it is observed that in each of the settings in Table 1 the LRT is more powerful than the IUT for congruent

Table 1: Congruent and Non-congruent pairs of cones under various settings

Settings	Geometry	Type	Cones
1	$\rho = \pi$	Congruent	$\{\theta_1 \geq 0, \theta_2 \geq 0\}$,$\{\theta_1 \leq 0, \theta_2 \leq 0\}$
2	$\rho = \pi$	Non-congruent	$\{\theta_2 \geq -\sqrt{3}\theta_1, \theta_2 \leq \sqrt{3}\theta_1\}$,$\{\sqrt{3}\theta_2 \geq \theta_1, \sqrt{3}\theta_2 \leq -\theta_1\}$
3	$\rho > \pi$	Congruent	$\{\theta_2 \geq -\theta_1, \theta_2 \leq 2\theta_1\}$,$\{\theta_2 \geq \theta_1, \theta_2 \leq -2\theta_1\}$
4	$\rho > \pi$	Non-congruent	$\{\theta_2 \geq -\theta_1, \theta_2 \leq 3\theta_1\}$,$\{\theta_2 \geq \theta_1, \theta_2 \leq -2\theta_1\}$
5	$\tau = \pi$	Congruent	$\{\theta_2 \geq \theta_1, \theta_2 \leq 2\theta_1\}$,$\{\theta_2 \geq -\theta_1, \theta_2 \leq -2\theta_1\}$
6	$\tau = \pi$	Non-congruent	$\{\theta_2 \geq \theta_1, \theta_2 \leq 2.1\theta_1\}$, $\{\theta_2 \geq -\theta_1, \theta_2 \leq -2\theta_1\}$
7	$\tau > \pi$	Congruent	$\{4\theta_2 \geq \theta_1, 3\theta_2 \leq \theta_1\}$,$\{4\theta_2 \geq -\theta_1, 3\theta_2 \leq -\theta_1\}$
8	$\tau > \pi$	Non-congruent	$\{4\theta_2 \geq \theta_1, 3\theta_2 \leq \theta_1\}$,$\{4\theta_2 \geq -\theta_1, 2\theta_2 \leq -\theta_1\}$

as well as non-congruent pairs of cones. Although not reported in Table 2, it is observed that the powers of the LRT and the IUT decrease or increase as $\boldsymbol{\theta}$ is closer to or further from $\mathbf{0}$. Moreover the ratio of the power of the LRT to that of the IUT increases or decreases as $\boldsymbol{\theta}$ is closer to or further from $\mathbf{0}$, and equals 1 for large $\boldsymbol{\theta}$.

Table 2: Powers of LRT ($\mathbf{P_{LRT}}$) and IUT ($\mathbf{P_{IUT}}$) under settings in Table 1 for $\alpha = 0.05$ **and selected** $\boldsymbol{\theta} \in \boldsymbol{\Theta}_1$

Settings	$\boldsymbol{\theta}$	$\mathrm{P_{LRT}}$	$\mathrm{P_{IUT}}$
1	(-0.29,0.29)	0.8009	0.6814
2	(0,0.5)	0.8014	0.6988
3	(0,0.66)	0.8084	0.6623
4	(0,0.82)	0.8069	0.6863
5	(0,0.75)	0.8042	0.7367
6	(0,0.76)	0.7989	0.7299
7	(0,0.32)	0.8080	0.7023
8	(0,0.33)	0.8021	0.6894

7. Discussion

As noted in the introduction, the existing literature has focused on testing (1) in situations where the null parameter space is either a linear subspace or single convex cone. In this paper, we develop a general framework to address multiple cone problems in two dimensions. We consider situations where the null parameter space can be expressed as the union of multiple closed convex cones in $\mathbb{R}^2$, which encompasses a large class of problems. We propose a test statistic which is equivalent to the LRT under normality and coincides with the the IUT in some special cases. Since the finite sampling distributions of these test statistics usually do not have closed-form expressions, we derive their asymptotic null distributions. We also obtain their least favourable asymptotic null values and the corresponding least favourable asymptotic null distributions based on the geometry of the cones. These distributions are used to determine the size α critical values, which depend on the stochastic ordering of the test statistics. Finally, we show that our tests are uniformly more powerful than the conventional IUTs discussed in the literature. In future, we hope to address the more challenging problem of testing (1) for arbitrary convex cones in any finite dimension.

In fact the scope of (1) is much broader than the references in Sections 1 and 3. Some important classes of problems which can be formulated as (1) in higher dimensions include problems of model selection arising in order restricted inference (Mack and Wolfe (1981), Pan and Wolfe (1996), Pan (1997), Rueda *et al.* (2016), Wei *et al.* (2019), Panda (2019), Larriba *et al.* (2016), Larriba *et al.* (2020), Peddada *et al.* (2003), Peddada *et al.* (2005)); problems in the theory of ranking and selection; and problems in mathematical psychology which involve the verification of transitivity axioms underlying social choice theory (Oliveira *et al.* (2018), Iverson and Falmagne (1985), Tversky (1969), Regenwetter *et al.* (2011), Davis-Stober (2009), Myung *et al.* (2005), Heck and Davis-Stober (2019)). For example, in the theory of ranking and selection, Nettleton (2009) considered the problem of testing for the supremacy of a multinomial cell probability. The supremacy of the K^{th} cell probability is established by rejecting the null hypothesis $H_0 : \theta_K \leq \max\{\theta_1, \ldots, \theta_{K-1}\}$ where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)^T$ denotes the vector of multinomial cell probabilities. Clearly the null can be rewritten as $H_0 : \boldsymbol{\theta} \in \bigcup_{j=1}^{K-1} \mathcal{C}_j$, where $\mathcal{C}_j = \{\boldsymbol{\theta} \in \mathrm{P} : \theta_K \leq \theta_j\}$ and P is the set of K-dimensional probability vectors whose components sum to 1.

Acknowledgments

The work of Sayan Ghosh was conducted while being a post doctoral fellow at the University of Haifa. The work of Ori Davidov was partially supported by the Israel Science Foundation Grants No. 456/17 and 2200/22 and gratefully acknowledged.

References

Berger, R. L. (1982). Multiparameter hypothesis testing and acceptance sampling. *Technometrics*, **24**, 295–300.

Berger, R. L. (1989). Uniformly more powerful tests for hypotheses concerning linear inequalities and normal means. *Journal of the American Statistical Association*, **84**, 192–199.

Berger, R. L. (1997). *Advances in Statistical Decision Theory and Applications*. Birkhäuser, Boston.

Berger, R. L. and Sinclair, D. F. (1984). Testing hypotheses concerning unions of linear subspaces. *Journal of the American Statistical Association*, **79**, 158–163.

Cohen, A., Gatsonis, C., and Marden, J. I. (1983). *Hypothesis Tests and Optimality Properties in Discrete Multivariate Analysis*. Academic Press, New York.

Davis-Stober, C. P. (2009). Analysis of multinomial models under inequality constraints: applications to measurement theory. *Journal of Mathematical Psychology*, **53**, 1–13.

Gail, M. and Simon, R. (1985). Testing for qualitative interactions between treatment effects and patient subsets. *Biometrics*, **41**, 361–372.

Heck, D. W. and Davis-Stober, C. P. (2019). Multinomial models with linear inequality constraints: Overview and improvements of computational methods for Bayesian inference. *Journal of Mathematical Psychology*, **91**, 70–87.

Iverson, G. and Falmagne, J. C. (1985). Statistical issues in measurement. *Mathematical Social Sciences*, **10**, 131–153.

Larriba, Y., Rueda, C., Fernández, M., and Peddada, S. (2016). Order Restricted Inference for Oscillatory Systems for Detecting Rhythmic Signals. *Nucleic Acids Research*, **44**, 1–8.

Larriba, Y., Rueda, C., Fernández, M., and Peddada, S. (2020). Order restricted inference in chronobiology. *Statistics in Medicine*, **39**, 265–278.

Laska, E. M. and Meisner, M. J. (1989). Testing whether identified treatment is best. *Biometrics*, **45**, 1139–1151.

Liu, H. and Berger, R. L. (1995). Uniformly more powerful, one-sided tests for hypotheses about linear inequalities. *Annals of Statistics*, **23**, 55–72.

Mack, G. A. and Wolfe, D. A. (1981). K-Sample Rank Tests for Umbrella Alternative. *Journal of the American Statistical Association*, **76**, 175–181.

Myung, J. I., Karabatsos, G., and Iverson, G. J. (2005). A Bayesian approach to testing decision making axioms. *Journal of Mathematical Psychology*, **49**, 205–225.

Nettleton, D. (2009). Testing for the supremacy of a Multinomial cell probability. *Journal of the American Statistical Association*, **104**, 1052–1059.

Oliveira, I. F. D., Ailon, N., and Davidov, O. (2018). A new and flexible approach to the analysis of paired comparison data . *Journal of Machine Learning Research*, **19**, 1–29.

Pan, G. (1997). Confidence subset containing the unknown peaks of an umbrella ordering. *Journal of the American Statistical Association*, **92**, 307–314.

Pan, G. and Wolfe, D. A. (1996). Comparing groups with umbrella orderings. *Journal of the American Statistical Association*, **91**, 311–317.

Panda, S. (2019). The arrival of circadian medicine. *Nature Reviews Endocrinology*, **15**, 67–69.

Peddada, S., Harris, S., Zajd, J., and Harvey, E. (2005). ORIOGEN: order restricted inference for ordered gene expression data. *Bioinformatics*, **21**, 3933–3934.

Peddada, S., Lobenhofer, E. K., Li, L., Afshari, C. A., Weinberg, C. R., and Umbach, D. M. (2003). Gene selection and clustering for time-courseand dose–response microarray experimentsusing order-restricted inference. *Bioinformatics*, **19**, 834–841.

Regenwetter, M., Dana, J., and Davis-Stober, C. P. (2011). Transitivity of preferences. *Psychological Review*, **118**, 42–56.

Robertson, T. and Wegman, E. J. (1978). Likelihood ratio tests for order restrictions in exponential families. *Annals of Statistics*, **6**, 485–505.

Rueda, C., Ugarte, M. D., and Militino, A. F. (2016). Checking unimodality using isotonic regression: an application to breast cancer mortality rates. *Stochastic Environmental Research and Risk Assessment*, **30**, 1277–1288.

Saumard, A. and Wellner, J. A. (2014). Log-concavity and strong log-concavity: A review. *Statistics Surveys*, **8**, 45–114.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*. Springer, New York.

Silvapulle, M. J. (2001). Tests against qualitative interaction: Exact critical values and robust tests. *Biometrics*, **57**, 1157–1165.

Silvapulle, M. J. and Sen, P. K. (2004). *Constrained Statistical Inference.* Wiley, New York.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, **76**, 31–48.

Wei, Y., Wainwright, M. J., and Guntuboyina, A. (2019). The geometry of hypothesis testing over convex cones: generalized likelihood ratio tests and minimax radii. *Annals of Statistics*, **47**, 994–1024.

APPENDIX

PROOFS

Proof of Theorem 1:

Proof: Since $\boldsymbol{S} \sim \mathcal{N}_2(\boldsymbol{\theta}, \boldsymbol{\Sigma})$, the kernel of the log-likelihood is given by

$$l(\boldsymbol{\theta}) = -\frac{1}{2}(\boldsymbol{S}-\boldsymbol{\theta})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{S}-\boldsymbol{\theta}) = -\frac{1}{2}\|\boldsymbol{S}-\boldsymbol{\theta}\|^2_{\boldsymbol{\Sigma}}. \tag{17}$$

It follows that the global, unrestricted MLE of $\boldsymbol{\theta}$ is $\widehat{\boldsymbol{\theta}} = \boldsymbol{S}$ and

$$l(\widehat{\boldsymbol{\theta}}) = 0. \tag{18}$$

The restricted MLE solves

$$\begin{aligned}\widetilde{\boldsymbol{\theta}} &= \arg\max\{l(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \bigcup_{i=1}^{K} \mathcal{C}_i\} = \arg\min\{\|\boldsymbol{S}-\boldsymbol{\theta}\|^2_{\boldsymbol{\Sigma}} : \boldsymbol{\theta} \in \bigcup_{i=1}^{K} \mathcal{C}_i\} \\ &= \arg\min\{\|\boldsymbol{S}-\boldsymbol{\theta}\|^2_{\boldsymbol{\Sigma}} : \boldsymbol{\theta} \in \{\widetilde{\boldsymbol{\theta}}_1, \ldots, \widetilde{\boldsymbol{\theta}}_K\}\}\end{aligned}$$

where for $i = 1, \ldots, K$ we define $\widetilde{\boldsymbol{\theta}}_i = \arg\min\{\|\boldsymbol{S}-\boldsymbol{\theta}\|^2_{\boldsymbol{\Sigma}} : \boldsymbol{\theta} \in \mathcal{C}_i\}$ which is nothing but the projection of $\boldsymbol{S}$ on $\mathcal{C}_i$ with respect to $\boldsymbol{\Sigma}$ denoted by $\Pi_{\boldsymbol{\Sigma}}(\boldsymbol{S}_n \mid \mathcal{C}_i)$. In other words

$$\widetilde{\boldsymbol{\theta}} = \arg\min\{\|\boldsymbol{S} - \Pi_{\boldsymbol{\Sigma}}(\boldsymbol{S} \mid \mathcal{C}_i)\|^2_{\boldsymbol{\Sigma}} : i \in \{1, \ldots, K\}\}, \tag{19}$$

so

$$l(\widetilde{\boldsymbol{\theta}}) = -\frac{1}{2}\min\{\|\boldsymbol{S} - \Pi_{\boldsymbol{\Sigma}}(\boldsymbol{S} \mid \mathcal{C}_i)\|^2_{\boldsymbol{\Sigma}} : i \in \{1, \ldots, K\}\}. \tag{20}$$

Now the LRT statistic is given by

$$\Lambda = 2\{l(\widehat{\boldsymbol{\theta}}) - l(\widetilde{\boldsymbol{\theta}})\}$$

which, using (18) and (20), reduces to

$$\Lambda = \min\{\|\boldsymbol{S} - \Pi_{\boldsymbol{\Sigma}}(\boldsymbol{S} \mid \mathcal{C}_i)\|^2_{\boldsymbol{\Sigma}} : i \in \{1, \ldots, K\}\} \tag{21}$$

as claimed in (21) with $\boldsymbol{S}_n$ replaced by $\boldsymbol{S}$.

Now note that testing (1) is equivalent to testing $\bigcup_{i=1}^{K} H_0^{(i)}$ versus $\bigcap_{i=1}^{K} H_1^{(i)}$ where $H_0^{(i)}$: $\boldsymbol{\theta} \in \mathcal{C}_i$ and $H_1^{(i)} : \boldsymbol{\theta} \notin \mathcal{C}_i$. It is clear that the LRT statistic for individual tests $H_0^{(i)}$ versus $H_1^{(i)}$, each of which is a Type B problem (Silvapulle and Sen (2004)), is $\Lambda^{(i)} = \|\boldsymbol{S} - \tilde{\boldsymbol{\theta}}_i\|^2 = \|\boldsymbol{S} - \Pi_{\boldsymbol{\Sigma}}(\boldsymbol{S} \mid \mathcal{C}_i)\|^2$. Since the least favourable null value for any Type B problem is the origin $H_0^{(i)}$ is rejected if $\Lambda^{(i)}$ is larger than $c_\alpha^{(i)}$, the $1-\alpha$ quantile of the RV

$$\sum_{k=1}^{K} w_k(\mathcal{C}_i, \boldsymbol{\Sigma})\chi_k^2. \tag{22}$$

The IUT rejects the null in (1) if and only if $\Lambda^{(i)} > c_\alpha^{(i)}$ for every i. Note that $c_\alpha^{(i)} = c_\alpha^{(j)}$ if and only if the weights in (22) satisfy $w_k(\mathcal{C}_i, \boldsymbol{\Sigma}) = w_k(\mathcal{C}_j, \boldsymbol{\Sigma})$ for all $k = 1, \ldots, K$ or in other words that all cones are congruent. Since the critical values $c_\alpha^{(i)}$ are equal for each of the K tests, it follows that the IUT statistic is

$$\min\{\Lambda^{(1)}, \ldots, \Lambda^{(K)}\},$$

which is the same as the LRT statistic in (21) as a function of $\boldsymbol{S}$. □

Proof of Theorem 2:

Proof: By Theorem 1 the LRT statistic for (4) is (3) which reduces to

$$T = \min\{\|\boldsymbol{S} - \Pi(\boldsymbol{S} \mid \mathcal{C}_1)\|_2^2, \|\boldsymbol{S} - \Pi(\boldsymbol{S} \mid \mathcal{C}_2)\|_2^2\}. \tag{23}$$

Note that when $\boldsymbol{S} \in \mathcal{C}_1$ then $\Pi(\boldsymbol{S} \mid \mathcal{C}_1) = \boldsymbol{S}$ so $\|\boldsymbol{S} - \Pi(\boldsymbol{S} \mid \mathcal{C}_1)\|_2^2 = 0$ and similarly when $\boldsymbol{S} \in \mathcal{C}_2$. Thus if $\boldsymbol{S} \in \mathcal{C}_1 \cup \mathcal{C}_2$ then $T = 0$. Next, if $\boldsymbol{S} \notin \mathcal{C}_2 \cup \mathcal{C}_2$ then for $i \in \{1, 2\}$, $\Pi(\boldsymbol{S} \mid \mathcal{C}_i) = (S_1, 0)^T$ or $(0, S_2)^T$ so $\|\boldsymbol{S} - \Pi(\boldsymbol{S} \mid \mathcal{C}_i)\|_2^2 = S_1^2$ or S_2^2. Consequently,

$$T = \min\{S_1^2, S_2^2\}\mathbb{I}(\boldsymbol{S} \notin \mathcal{C}_1 \cup \mathcal{C}_2)$$

as claimed. Let $c > 0$. Then for any $\boldsymbol{\theta} \in \mathbb{R}^2$ we have

$$\begin{aligned}
\mathbb{P}_{\boldsymbol{\theta}}(T \geq c) &= \mathbb{P}_{\boldsymbol{\theta}}(\min\{S_1^2, S_2^2\} \geq c, \boldsymbol{S} \notin \mathcal{C}_1 \cup \mathcal{C}_2) \\
&= \mathbb{P}_{\boldsymbol{\theta}}(S_1^2 \geq c, S_2^2 \geq c, S_1 \geq 0, S_2 \leq 0) + \mathbb{P}_{\boldsymbol{\theta}}(S_1^2 \geq c, S_2^2 \geq c, S_1 \leq 0, S_2 \geq 0) \\
&= \mathbb{P}_{\boldsymbol{\theta}}(S_1 \geq \sqrt{c})\mathbb{P}_{\boldsymbol{\theta}}(S_2 \leq -\sqrt{c}) + \mathbb{P}_{\boldsymbol{\theta}}(S_1 \leq -\sqrt{c})\mathbb{P}_{\boldsymbol{\theta}}(S_2 \geq \sqrt{c}) \\
&= [1 - \Phi(\sqrt{c} - \theta_1)]\Phi(-\sqrt{c} - \theta_2) + \Phi(-\sqrt{c} - \theta_1)[1 - \Phi(\sqrt{c} - \theta_2)]. \qquad (24)
\end{aligned}$$

Here, as usual, $\phi(\cdot)$ and $\Phi(\cdot)$ denote the density and distribution function of a standard normal RV. Denote $\mathbb{P}_{\boldsymbol{\theta}}(T \geq c)$ by $H(\boldsymbol{\theta}) = H(\theta_1, \theta_2)$. Our goal is to maximize $H(\theta_1, \theta_2)$ over $(\theta_1, \theta_2) \in \boldsymbol{\Theta}_0$. We will first show that if $\boldsymbol{\theta} \in \mathcal{C}_1$ and $\theta_1 > \theta_2 > 0$ then

$$H(\theta_1, \theta_2) < H(\theta_1, 0). \tag{25}$$

Since $\theta_1 > \theta_2 > 0$, it follows that

$$\begin{aligned}
H(\theta_1, 0) - H(\theta_1, \theta_2) &= [1 - \Phi(\sqrt{c} - \theta_1)][\Phi(-\sqrt{c}) - \Phi(-\sqrt{c} - \theta_2)] \\
&\quad - \Phi(-\sqrt{c} - \theta_1)[\Phi(\sqrt{c}) - \Phi(\sqrt{c} - \theta_2)] \\
&> [1 - \Phi(\sqrt{c} - \theta_2)][\Phi(-\sqrt{c}) - \Phi(-\sqrt{c} - \theta_2)] \\
&\quad - \Phi(-\sqrt{c} - \theta_2)[\Phi(\sqrt{c}) - \Phi(\sqrt{c} - \theta_2)]. \qquad (26)
\end{aligned}$$

Let $p = 1 - \Phi(\sqrt{c} - \theta_2)$, $q = \Phi(-\sqrt{c}) - \Phi(-\sqrt{c} - \theta_2)$, $r = \Phi(-\sqrt{c} - \theta_2)$ and $s = \Phi(\sqrt{c}) - \Phi(\sqrt{c} - \theta_2)$. Thus establishing (25) is equivalent to showing that $pq > rs$ or $p/s > r/q$. Observe that p, q, r and s are all strictly positive. Furthermore $p > q$, $p > r$, $p > s$, $s > q$ and $p > q + r$, from which we deduce that $p/s > (q + r)/s$. It follows that showing $(q + r)/s > r/q$ will complete the proof of (25). Suppose the latter does not hold, *i.e.*, $(q + r)/s \leq r/q$ which in turn implies that $q^2 \leq r(s - q) < 0$. Since $q > 0$, we have a contradiction. Thus $(q + r)/s > r/q$ and consequently $pq > rs$ so (25) holds. A similar argument can be used to show that if $\boldsymbol{\theta} \in \mathcal{C}_1$ which satisfy $\theta_2 > \theta_1 > 0$ then

$$H(\theta_1, \theta_2) < H(0, \theta_2). \tag{27}$$

Next we consider the case where $\theta_1 = \theta_2 = \theta > 0$. Now,

$$H(\theta, \theta) = 2\Phi(-\sqrt{c} - \theta))\Phi(-\sqrt{c} + \theta) < 2[\Phi(-\sqrt{c})]^2 = H(0, 0),$$

where the inequality above is a consequence of the log–concavity of $\Phi(\cdot)$ (see Saumard and Wellner (2014)). Thus,

$$H(\theta, \theta) < H(0, 0). \tag{28}$$

It follows from (25), (27) and (28) that for any $\boldsymbol{\theta}$ in the interior of $\mathcal{C}_1$ there exists a $\boldsymbol{\theta}^*$ on the boundary of $\mathcal{C}_1$ for which

$$H(\boldsymbol{\theta}^*) > H(\boldsymbol{\theta}). \tag{29}$$

Repeating the above arguments we can show that (29) holds also for $\boldsymbol{\theta} \in \mathcal{C}_2$. Thus $\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} H(\boldsymbol{\theta})$ is attained on the set $\{(0, x) \cup (x, 0) : x \in \mathbb{R}\}$, *i.e.*, the boundary of $\boldsymbol{\Theta}_0$. Next consider the function $H(\theta_1, 0)$ with $\theta_1 \geq 0$. Clearly,

$$H(\theta_1, 0) = \Phi(-\sqrt{c})[1 - \Phi(\sqrt{c} - \theta_1) + \Phi(-\sqrt{c} - \theta_1)] \tag{30}$$

and therefore

$$\frac{\partial}{\partial \theta_1} H(\theta_1, 0) = \Phi(-\sqrt{c})[\phi(\sqrt{c} - \theta_1) - \phi(-\sqrt{c} - \theta_1)] \geq 0$$

since $\phi(\sqrt{c} - \theta_1) \geq \phi(-\sqrt{c} - \theta_1)$ whenever $\theta_1 \geq 0$. This implies that

$$\sup_{\boldsymbol{\theta}_1 \geq 0} H(\theta_1, 0) = \lim_{\theta_1 \to \infty} H(\theta_1, 0) = \Phi(-\sqrt{c}) = 1 - \Phi(\sqrt{c}) = \mathbb{P}(\mathcal{N}(0, 1) \geq \sqrt{c}) = \frac{1}{2}\mathbb{P}(\chi_1^2 \geq c).$$

Since the function $H(\theta_1, \theta_2)$ is permutation invariant and odd we have

$$\sup_{\boldsymbol{\theta}_1 \geq 0} H(\theta_1, 0) = \sup_{\boldsymbol{\theta}_1 \leq 0} H(\theta_1, 0) = \sup_{\boldsymbol{\theta}_2 \geq 0} H(0, \theta_2) = \sup_{\boldsymbol{\theta}_2 \leq 0} H(0, \theta_2) = \frac{1}{2}\mathbb{P}(\chi_1^2 \geq c). \tag{31}$$

Furthermore it is easy to see that:

$$\begin{aligned}\lim_{\theta_1 \to \infty} \mathbb{P}_{(\theta_1, 0)}(T = 0) &= \lim_{\theta_1 \to -\infty} \mathbb{P}_{(\theta_1, 0)}(T = 0) = \lim_{\theta_2 \to \infty} \mathbb{P}_{(0, \theta_2)}(T = 0) = \lim_{\theta_2 \to -\infty} \mathbb{P}_{(0, \theta_2)}(T = 0) \\ &= \frac{1}{2}\chi_0^2.\end{aligned} \tag{32}$$

It now follows from (31) and (32) that for $c \geq 0$

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} \mathbb{P}_{\boldsymbol{\theta}}(T \geq c) = \frac{1}{2}\mathbb{P}(\chi_0^2 \geq c) + \frac{1}{2}\mathbb{P}(\chi_1^2 \geq c). \tag{33}$$

as claimed. □

Proof of Lemma 1:

Proof: Note that $\mathcal{R}_{ij}$, $\mathcal{R}'_{ij}$ and $\mathcal{R}''_{ij}$ denote regions between the boundaries of adjacent cones or their polar cones. Since there are K such regions, we have $N + N' + N'' = K$. The upper bound for each summand is attained if all the regions between various pairs of adjacent cones are of the same type. If the angle between every pair of adjacent cones is less than $\pi/2$, then there are K regions of the type $\mathcal{R}_{ij}$, *i.e.*, $N \leq K$. If the angle between every pair of adjacent cones is between $\pi/2$ and π, then we have $K \leq 3$ and hence at most 3 regions of the type $\mathcal{R}'_{ij}$, *i.e.*, $N' \leq 3$. If the angle between some pair of adjacent cones is greater than π, then the angles between all other pairs of such cones are each less than π. Hence there is at most one region of the type $\mathcal{R}''_{ij}$, *i.e.*, $N'' \leq 1$. Suppose now that $N'' = 1$. If the angles between all other pairs of adjacent cones are each less than $\pi/2$, then there are $K-1$ regions of the type $\mathcal{R}_{ij}$, *i.e.*, $N \leq K-1$. Moreover there can be at most one other pair of adjacent cones with an angle between $\pi/2$ and π between them, and hence at most one region of the type $\mathcal{R}'_{ij}$, *i.e.*, $N' \leq 1$. □

Proof of Lemma 2:

Proof: First note that $\mathrm{d}^2(\boldsymbol{S}, \mathrm{ray}(\boldsymbol{u})) = (\boldsymbol{u}_*^T \boldsymbol{S})^2$ and similarly $\mathrm{d}^2(\boldsymbol{S}, \mathrm{ray}(\boldsymbol{v})) = (\boldsymbol{v}_*^T \boldsymbol{S})^2$. Since $\boldsymbol{S} \sim \mathcal{N}_2(\boldsymbol{0}, \boldsymbol{I}$, both $(\boldsymbol{u}_*^T \boldsymbol{S})^2$ and $(\boldsymbol{v}_*^T \boldsymbol{S})^2$ are distributed as χ_1^2 RVs. Moreover the correlation coefficient between $\boldsymbol{u}_*^T \boldsymbol{S}$ and $\boldsymbol{v}_*^T \boldsymbol{S}$ is $\boldsymbol{u}_*^T \boldsymbol{v}_*$. Since $\measuredangle(\boldsymbol{u}_*, \boldsymbol{v}_*) = \pi - \measuredangle(\boldsymbol{u}, \boldsymbol{v}) = \pi - \gamma$, we have $\boldsymbol{u}_*^T \boldsymbol{v}_* = \cos(\measuredangle(\boldsymbol{u}_*, \boldsymbol{v}_*)) = \cos(\pi - \gamma) = -\cos(\gamma)$. Let $D_1 = \boldsymbol{u}_*^T \boldsymbol{S}$ and $D_2 = \boldsymbol{v}_*^T \boldsymbol{S}$ so $\mathrm{Var}(D_1) = \mathrm{Var}(D_2) = 1$ and the correlation coefficient between D_1 and D_2 is $\boldsymbol{u}_*^T \boldsymbol{v}_* = -\cos(\gamma)$. Further we have $\chi_{1,1}^2(\gamma) = \min\{D_1^2, D_2^2\}$ which is a function of the length of $\boldsymbol{S}$. Since $\boldsymbol{S} \sim \mathcal{N}_2(\boldsymbol{0}, \boldsymbol{I})$, the length and direction of $\boldsymbol{S}$ are independently distributed. Thus

$$
\begin{aligned}
\mathbb{P}(\chi_{1,1}^2(\gamma) \geq c, \boldsymbol{S} \in \mathcal{R}) &= \mathbb{P}(\boldsymbol{S} \in \mathcal{R})\mathbb{P}(\chi_{1,1}^2(\gamma) \geq c) \\
&= \frac{\gamma}{2\pi}\mathbb{P}(\min\{D_1^2, D_2^2\} \geq c) \\
&= \frac{\gamma}{2\pi}\mathbb{P}(D_1^2 \geq c, D_2^2 \geq c) \\
&= \frac{\gamma}{2\pi}[1 - \mathbb{P}(-\sqrt{c} \leq D_1 \leq \sqrt{c}) - \mathbb{P}(D_1 \leq -\sqrt{c}, -\sqrt{c} \leq D_2 \leq \sqrt{c}) \\
&- \mathbb{P}(D_1 \geq \sqrt{c}, -\sqrt{c} \leq D_2 \leq \sqrt{c})] \\
&= \frac{\gamma}{2\pi}[\mathbb{P}(D_1 \geq \sqrt{c}, D_2 \geq \sqrt{c}) + \mathbb{P}(D_1 \geq \sqrt{c}, D_2 \leq -\sqrt{c}) \\
&+ \mathbb{P}(D_1 \leq -\sqrt{c}, D_2 \geq \sqrt{c}) + \mathbb{P}(D_1 \leq -\sqrt{c}, D_2 \leq -\sqrt{c})]
\end{aligned}
$$

where $(D_1, D_2)^T$ has a bivariate normal distribution with mean $\boldsymbol{0}$, unit variances and correlation $-\cos(\gamma)$. □

Proof of Theorem 3:

Proof: Recall that the LRT statistic T_n for (1) is

$$
T_n = \min\{n\|\boldsymbol{S}_n - \Pi(\boldsymbol{S}_n \mid \mathcal{C}_1)\|^2, \ldots, n\|\boldsymbol{S}_n - \Pi(\boldsymbol{S}_n \mid \mathcal{C}_K)\|^2\}, \tag{34}
$$

which minimizes the squared distance between $\boldsymbol{S}_n$ and each of the K cones. If $\boldsymbol{S}_n \in \mathcal{C}_i$ for any i, then $\Pi(\boldsymbol{S}_n \mid \mathcal{C}_i) = \boldsymbol{S}_n$, thus $T_n = 0$. If $\boldsymbol{S}_n$ lies in $\mathcal{R}_i(\boldsymbol{u}_i)$ for some i, then $\Pi(\boldsymbol{S}_n \mid \mathcal{C}_i) = (\boldsymbol{u}_i^T \boldsymbol{S}_n)\boldsymbol{u}_i$ and $\|\boldsymbol{S}_n - \Pi(\boldsymbol{S}_n \mid \mathcal{C}_r)\|^2 > \|\boldsymbol{S}_n - \Pi(\boldsymbol{S}_n \mid \mathcal{C}_i)\|^2$ where $r \neq i$, thus $T_n = n(\boldsymbol{u}_{i*}^T \boldsymbol{S}_n)^2$. Similarly T_n can be obtained for $\boldsymbol{S}_n$ lying in $\mathcal{R}_i(\boldsymbol{v}_i)$ or $\mathcal{R}_j(\boldsymbol{u}_j)$ or $\mathcal{R}_j(\boldsymbol{v}_j)$. If $\boldsymbol{S}_n$ lies in $\mathcal{R}_{ij}$ or $\mathcal{R}'_{ij}$ for some $(i,j) \in \mathcal{P}$, then $\Pi(\boldsymbol{S}_n \mid \mathcal{C}_i) = (\boldsymbol{v}_i^T \boldsymbol{S}_n)\boldsymbol{v}_i$ and $\Pi(\boldsymbol{S}_n \mid \mathcal{C}_j) = (\boldsymbol{u}_j^T \boldsymbol{S}_n)\boldsymbol{u}_j$. Also $\|\boldsymbol{S}_n - \Pi(\boldsymbol{S}_n \mid \mathcal{C}_r)\|^2 > \max\{\|\boldsymbol{S}_n - \Pi(\boldsymbol{S}_n \mid \mathcal{C}_i)\|^2, \|\boldsymbol{S}_n - \Pi(\boldsymbol{S}_n \mid \mathcal{C}_j)\|^2\}$ where $r \notin \{i,j\}$, thus $T_n = \min\{n(\boldsymbol{u}_{j*}^T \boldsymbol{S}_n)^2, n(\boldsymbol{v}_{i*}^T \boldsymbol{S}_n)^2\}$ Finally if $\boldsymbol{S}_n \in \mathcal{R}''_{pq}$, then $\boldsymbol{S}_n \in \mathcal{C}_i^0$ and $\Pi(\boldsymbol{S}_n \mid \mathcal{C}_i) = \boldsymbol{0}$ for every i, thus $T_n = n\|\boldsymbol{S}_n\|^2$. To summarize,

$$T_n = \begin{cases} 0 & \text{if } \boldsymbol{S}_n \in \mathcal{C}_i \\ \min\{n(\boldsymbol{u}_{j*}^T \boldsymbol{S}_n)^2, n(\boldsymbol{v}_{i*}^T \boldsymbol{S}_n)^2\} & \text{if } \boldsymbol{S}_n \in \mathcal{R}_{ij} \text{ or } \mathcal{R}'_{ij} \\ n(\boldsymbol{u}_{i*}^T \boldsymbol{S}_n)^2 \text{ or } n(\boldsymbol{v}_{i*}^T \boldsymbol{S}_n)^2 & \text{if } \boldsymbol{S}_n \in \mathcal{R}_i(\boldsymbol{u}_i) \text{ or } \mathcal{R}_i(\boldsymbol{v}_i) \\ n\|\boldsymbol{S}_n\|^2 & \text{if } \boldsymbol{S}_n \in \mathcal{R}''_{pq} \end{cases} \tag{35}$$

for all $(i,j) \in \mathcal{P}$. Next, we evaluate the limiting distribution of T_n for various values of $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$. Suppose first that $\boldsymbol{\theta} \in \text{int}(\boldsymbol{\Theta}_0)$, *i.e.*, $\boldsymbol{\theta}$ lies in the interior of $\boldsymbol{\Theta}_0$. If so, the ball $\mathcal{B}(\boldsymbol{\theta}, \delta)$ is a subset of $\text{int}(\boldsymbol{\Theta}_0)$ for some $\delta > 0$. Since $\boldsymbol{S}_n$ is consistent for $\boldsymbol{\theta}$ it follows that $\mathbb{P}(\boldsymbol{S}_n \in \mathcal{B}(\boldsymbol{\theta}, \epsilon)) \to 1$ as $n \to \infty$ for all $\epsilon < \delta$. Therefore $T_n \xrightarrow{p} 0$ and consequently $\boldsymbol{\theta} \in \text{int}(\boldsymbol{\Theta}_0)$ implies that

$$T_n \Rightarrow \chi_0^2, \tag{36}$$

as $n \to \infty$. Next, consider the situation when $\boldsymbol{\theta} \in \text{ray}(\boldsymbol{\Theta}_0)$. Without loss of generality let $\boldsymbol{\theta} = \lambda \boldsymbol{u}_1$ for some fixed $\lambda > 0$, *i.e.*, $\boldsymbol{\theta}$ lies on one of the rays generating the cone $\mathcal{C}_1$. The ray through $\boldsymbol{\theta}$ partitions $\mathcal{B}(\boldsymbol{\theta}, \epsilon)$ into two half circles $\mathcal{B}_1 \subset \mathcal{C}_1$ and $\mathcal{B}_2 \subset \mathcal{R}_1(\boldsymbol{u}_1)$. Observe that $\mathcal{B}_1 - \boldsymbol{\theta} = \boldsymbol{\theta} - \mathcal{B}_2$ where $\mathcal{B}_1 - \boldsymbol{\theta} = \{\boldsymbol{S} - \boldsymbol{\theta} \mid \boldsymbol{S} \in \mathcal{B}_1\}$ and $\boldsymbol{\theta} - \mathcal{B}_2 = \{\boldsymbol{\theta} - \boldsymbol{S} \mid \boldsymbol{S} \in \mathcal{B}_2\}$. The distribution of $\boldsymbol{S}_n$ is spherically symmetric around $\boldsymbol{\theta}$ so $\boldsymbol{S}_n - \boldsymbol{\theta} \stackrel{d}{=} \boldsymbol{\theta} - \boldsymbol{S}_n$. Consequently,

$$\begin{aligned} \mathbb{P}(\boldsymbol{S}_n \in \mathcal{B}_1) &= \mathbb{P}(\boldsymbol{S}_n - \boldsymbol{\theta} \in \mathcal{B}_1 - \boldsymbol{\theta}) = \mathbb{P}(\boldsymbol{S}_n - \boldsymbol{\theta} \in \boldsymbol{\theta} - \mathcal{B}_2) \\ &= \mathbb{P}(\boldsymbol{\theta} - \boldsymbol{S}_n \in \boldsymbol{\theta} - \mathcal{B}_2) = \mathbb{P}(\boldsymbol{S}_n - \boldsymbol{\theta} \in \mathcal{B}_2 - \boldsymbol{\theta}) = \mathbb{P}(\boldsymbol{S}_n \in \mathcal{B}_2). \end{aligned}$$

Moreover, since $\boldsymbol{S}_n$ is consistent for $\boldsymbol{\theta}$ we have $\mathbb{P}(\boldsymbol{S}_n \in \mathcal{B}(\boldsymbol{\theta}, \epsilon)) \to 1$ so $\mathbb{P}(\boldsymbol{S}_n \in \mathcal{B}_1) = \mathbb{P}(\boldsymbol{S}_n \in \mathcal{B}_2) \to 1/2$ as $n \to \infty$. Thus, $\mathbb{P}(\boldsymbol{S}_n \in \mathcal{C}_1) = \mathbb{P}(\boldsymbol{S}_n \in \mathcal{R}_1(\boldsymbol{u}_1)) = 1/2 + o_P(1)$ and the LRT statistic equals

$$T_n = 0 \times \mathbb{I}(\boldsymbol{S}_n \in \mathcal{C}_1) + n(\boldsymbol{u}_{1*}^T \boldsymbol{S}_n)^2 \times \mathbb{I}(\boldsymbol{S}_n \in \mathcal{R}_1(\boldsymbol{u}_1)) + o_P(1).$$

Observe that $\boldsymbol{S}_n = \boldsymbol{\theta} + \bar{\boldsymbol{Z}}_n$ where $\bar{\boldsymbol{Z}}_n$ is the average of n IID $\mathcal{N}_2(\boldsymbol{0}, \boldsymbol{I})$ RVs as $n \to \infty$. It follows that

$$n(\boldsymbol{u}_{1*}^{T/} \boldsymbol{S}_n)^2 = n(\boldsymbol{u}_{1*}^T (\boldsymbol{\theta} + \bar{\boldsymbol{Z}}_n))^2 = n(\boldsymbol{u}_{1*}^T (\lambda \boldsymbol{u} + \bar{\boldsymbol{Z}}_n))^2 = (\boldsymbol{u}_{1*}^T (\sqrt{n} \bar{\boldsymbol{Z}}_n))^2 \Rightarrow \chi_1^2$$

and therefore, if $\boldsymbol{\theta} = \lambda \boldsymbol{u}_1$

$$T_n \Rightarrow \frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2. \tag{37}$$

Obviously (37) remains unchanged if $\boldsymbol{\theta}$ lies on any other ray of $\boldsymbol{\Theta}_0 \setminus \{\mathbf{0}\}$. Finally, suppose $\boldsymbol{\theta} = \mathbf{0}$. Here $\sqrt{n}\boldsymbol{S}_n \sim \mathcal{N}_2(\mathbf{0}, \boldsymbol{I})$ as $n \to \infty$ which is spherically symmetric so the direction and length of $\boldsymbol{S}_n$ are statistically independent. We have already seen that $n(\boldsymbol{u}_{j*}^T \boldsymbol{S}_n)^2$ and $n(\boldsymbol{v}_{i*}^T \boldsymbol{S}_n)^2$ are each distributed as a χ_1^2 RV whereas $n\|\boldsymbol{S}_n\|^2$ is distributed as a χ_2^2 RV as $n \to \infty$. It follows from (35) that: (i) if

$$\boldsymbol{S}_n \in \bigcup_{i=1}^{K} \mathcal{C}_i,$$

an event that has probability $\rho/2\pi$, then $T_n \Rightarrow \chi_0^2$; (ii) if $(i,j) \in \mathcal{P}$ and $\boldsymbol{S}_n \in \mathcal{R}_{ij}$ or $\mathcal{R}'_{ij}$, an event that has probability $\gamma_{ij}/2\pi$, then $T_n \Rightarrow \chi_{1,1}^2(\gamma_{ij})$; (iii) if $(i,j) \in \mathcal{P}$ and

$$\boldsymbol{S}_n \in \bigcup_{(i,j)\in\mathcal{P}} (\mathcal{R}_i(\boldsymbol{u}_i) \cup \mathcal{R}_i(\boldsymbol{v}_i) \cup \mathcal{R}_j(\boldsymbol{u}_j) \cup \mathcal{R}_j(\boldsymbol{v}_j)),$$

an event that has probability $\tau/2\pi$, then $T_n \Rightarrow \chi_1^2$; and (iv) if $\boldsymbol{S}_n \in \mathcal{R}''_{pq}$, an event that has probability $\gamma_{pq}/2\pi$, then $T_n \Rightarrow \chi_2^2$. Putting it all together we find that when $\boldsymbol{\theta} = \mathbf{0}$, we have

$$T_n \Rightarrow \frac{\rho}{2\pi}\chi_0^2 + \sum_{(i,j)\in\mathcal{P}} \frac{\gamma_{ij}}{2\pi}\chi_{1,1}^2(\gamma_{ij}) + \frac{\tau}{2\pi}\chi_1^2 + \frac{\gamma_{pq}}{2\pi}\chi_2^2. \tag{38}$$

Equations (36), (37) and (38) establish the result. □

Proof of Theorem 4:

Proof: First we prove that the given conditions are sufficient. Recall that for any $c \geq 0$

$$\mathbb{P}(T_{\boldsymbol{R}} \geq c) = \frac{1}{2}\mathbb{P}(\chi_0^2 \geq c) + \frac{1}{2}\mathbb{P}(\chi_1^2 \geq c) \tag{39}$$

and

$$\mathbb{P}(T_{\boldsymbol{O}} \geq c) = \frac{\rho}{2\pi}\mathbb{P}(\chi_0^2 \geq c) + \sum_{(i,j)\in\mathcal{P}} \frac{\gamma_{ij}}{2\pi}\mathbb{P}(\chi_{1,1}^2(\gamma_{ij}) \geq c) + \frac{\tau}{2\pi}\mathbb{P}(\chi_1^2 \geq c) + \frac{\gamma_{pq}}{2\pi}\mathbb{P}(\chi_2^2 \geq c), \tag{40}$$

where τ or γ_{pq} may be equal to 0. If $\tau \geq \pi$ then both

$$(A_{\boldsymbol{O}}) \quad \frac{\tau}{2\pi} \geq \frac{1}{2} \quad \text{and} \quad (B_{\boldsymbol{O}}) \quad \frac{\rho}{2\pi} < \frac{1}{2}$$

hold. By $(B_{\boldsymbol{O}})$ the first two terms on the right hand side of (40) are larger than the first term on the right hand side of (39). By $(A_{\boldsymbol{O}})$ the same is true when comparing the last two terms in (40) to the second term of (39). Therefore $\mathbb{P}(T_{\boldsymbol{O}} \geq c) > \mathbb{P}(T_{\boldsymbol{R}} \geq c)$ so

$$T_{\boldsymbol{O}} \succeq_{\text{st}} T_{\boldsymbol{R}}.$$

Hence, we conclude that the least favourable limiting null distribution for T_n is that of $T_{\boldsymbol{O}}$ when $\tau \geq \pi$.

Now suppose that $\rho \geq \pi$ then $\gamma_{pq} = 0$. This is because if $\gamma_{pq} > 0$, then $\tau \geq \pi$ so $\rho < \pi$. We have

$$(A_{\boldsymbol{R}}) \ \ \frac{\rho}{2\pi} \geq \frac{1}{2} \quad \text{and} \quad (B_{\boldsymbol{R}}) \ \ \sum_{(i,j)\in\mathcal{P}} \gamma_{ij} + \tau \leq \pi.$$

Hence,

$$\begin{aligned}
\mathbb{P}(T_{\boldsymbol{O}} \geq c) &= \frac{\rho}{2\pi}\mathbb{P}(\chi_0^2 \geq c) + \sum_{(i,j)\in\mathcal{P}} \frac{\gamma_{ij}}{2\pi}\mathbb{P}(\chi_{1,1}^2(\gamma_{ij}) \geq c) + \frac{\tau}{2\pi}\mathbb{P}(\chi_1^2 \geq c) \\
&< \frac{1}{2}\mathbb{P}(\chi_0^2 \geq c) + \frac{\sum_{(i,j)\in\mathcal{P}} \gamma_{ij} + \tau}{2\pi}\mathbb{P}(\chi_1^2 \geq c) \\
&\leq \frac{1}{2}\mathbb{P}(\chi_0^2 \geq c) + \frac{1}{2}\mathbb{P}(\chi_1^2 \geq c) = \mathbb{P}(T_{\boldsymbol{R}} \geq c)
\end{aligned}$$

where the first inequality is a consequence of $(A_{\boldsymbol{R}})$ and the second of $(B_{\boldsymbol{R}})$. Thus

$$T_{\boldsymbol{R}} \succeq_{\text{st}} T_{\boldsymbol{O}},$$

i.e., the least favourable limiting null distribution for T_n is that of $T_{\boldsymbol{R}}$ when $\rho \geq \pi$.

Next we prove that the above conditions are necessary. Suppose first that $\tau < \pi$. Then $\gamma_{pq} = 0$ since $\gamma_{pq} > 0$ implies $\tau \geq \pi$. For $c > 0$, we have

$$\begin{aligned}
\mathbb{P}(T_{\boldsymbol{O}} \geq c) - \mathbb{P}(T_{\boldsymbol{R}} \geq c) &= \sum_{(i,j)\in\mathcal{P}} \frac{\gamma_{ij}}{2\pi}\mathbb{P}(\chi_{1,1}^2(\gamma_{ij} \geq c) + (\frac{\tau}{2\pi} - \frac{1}{2})\mathbb{P}(\chi_1^2 \geq c) \\
&< (\sum_{(i,j)\in\mathcal{P}} \frac{\gamma_{ij}}{2\pi} + \frac{\tau}{2\pi} - \frac{1}{2})\mathbb{P}(\chi_1^2 \geq c) \\
&= \frac{\sum_{(i,j)\in\mathcal{P}} \gamma_{ij} + \tau - \pi}{2\pi}\mathbb{P}(\chi_1^2 \geq c) \\
&= \frac{\pi - \rho}{2\pi}\mathbb{P}(\chi_1^2 \geq c). \qquad (41)
\end{aligned}$$

If $\rho \geq \pi$ then the RHS of (41) is negative so $\mathbb{P}(T_{\boldsymbol{O}} \geq c) < \mathbb{P}(T_{\boldsymbol{R}} \geq c)$ for all $c > 0$. However, if $\rho < \pi$ then the RHS of (41) is positive. So there is at least one $c > 0$ satisfying $\mathbb{P}(T_{\boldsymbol{O}} \geq c) < \mathbb{P}(T_{\boldsymbol{R}} \geq c)$. Thus the condition $\tau \geq \pi$ is necessary for the limiting null distribution of T_n to be that of $T_{\boldsymbol{O}}$.

Finally suppose that $\rho < \pi$. If $\gamma_{pq} > 0$, then $\tau \geq \pi$ so $\mathbb{P}(T_{\boldsymbol{R}} \geq c) < \mathbb{P}(T_{\boldsymbol{O}} \geq c)$ for all $c > 0$ as shown earlier. If $\gamma_{pq} = 0$ then using (41), we have for $c > 0$

$$\begin{aligned}
\mathbb{P}(T_{\boldsymbol{O}} \geq c) - \mathbb{P}(T_{\boldsymbol{R}} \geq c) &= \sum_{(i,j)\in\mathcal{P}} \frac{\gamma_{ij}}{2\pi}\mathbb{P}(\chi_{1,1}^2(\gamma_{ij} \geq c) + (\frac{\tau}{2\pi} - \frac{1}{2})\mathbb{P}(\chi_1^2 \geq c) \\
&> (\frac{\tau}{2\pi} - \frac{1}{2})\mathbb{P}(\chi_1^2 \geq c). \qquad (42)
\end{aligned}$$

If $\tau \geq \pi$ then the RHS of (42) is positive so $\mathbb{P}(T_{\boldsymbol{R}} \geq c) < \mathbb{P}(T_{\boldsymbol{O}} \geq c)$ for all $c > 0$. However if $\tau < \pi$ then the RHS of (42) is negative. So there is at least one $c > 0$ satisfying $\mathbb{P}(T_{\boldsymbol{R}} \geq c) < \mathbb{P}(T_{\boldsymbol{O}} \geq c)$. Thus the condition $\rho \geq \pi$ is necessary for the limiting null distribution of T_n to be that of $T_{\boldsymbol{R}}$. □

Proof of Theorem 5:

Proof: Let $\Lambda^{(i)}$ denote the LRT statistic for testing $H_0^{(i)} : \boldsymbol{\theta} \in \mathcal{C}_i$ against $H_1^{(i)} : \boldsymbol{\theta} \notin \mathcal{C}_i$ and let $c_\alpha^{(i)}$ denote its critical value. Clearly, (cf. Silvapulle and Sen (2004)) $c_\alpha^{(i)}$ is the $1-\alpha$ quantile of the RV

$$\frac{\rho_i}{2\pi}\chi_0^2 + \frac{1}{2}\chi_1^2 + \frac{\pi - \rho_i}{2\pi}\chi_2^2. \tag{43}$$

The size α IUT rejects the null $\bigcup_{i=1}^{K} H_0^{(i)}$ if and only if $\Lambda^{(i)} > c_\alpha^{(i)}$ for every i. Recall that the size α LRT for (1) rejects the null if $T_n > c_\alpha$ where c_α is its critical value as discussed in Section 3. The cones $\mathcal{C}_1, \ldots, \mathcal{C}_K$ satisfy one of three possibilities: (I) $\rho \geq \pi$; (II) $\tau \geq \pi$; or (III) $\rho < \pi, \tau < \pi$. If (I) holds then by Theorems 3 and 4 the asymptotic critical value for T_n is the $1-\alpha$ quantile of the RV

$$\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2. \tag{44}$$

It follows that $c_\alpha < \min\{c_\alpha^{(1)}, \ldots, c_\alpha^{(K)}\}$ so the LRT has higher power than the IUT. If (II) holds then the asymptotic critical value for T_n is the $1-\alpha$ quantile of the RV

$$\frac{\rho}{2\pi}\chi_0^2 + \sum_{(i,j)\in\mathcal{P}} \frac{\gamma_{ij}}{2\pi}\chi_{1,1}^2(\gamma_{ij}) + \frac{\tau}{2\pi}\chi_1^2 + \frac{\gamma_{pq}}{2\pi}\chi_2^2. \tag{45}$$

Observe that

$$\sum_{(i,j)\in\mathcal{P}} \frac{\gamma_{ij}}{2\pi}\chi_{1,1}^2(\gamma_{ij}) + \frac{\tau}{2\pi}\chi_1^2 + \frac{\gamma_{pq}}{2\pi}\chi_2^2 \preceq_{st} \frac{\sum_{(i,j)\in\mathcal{P}} \gamma_{ij} + \tau}{2\pi}\chi_1^2 + \frac{\gamma_{pq}}{2\pi}\chi_2^2 \tag{46}$$

and since (II) holds we have

$$\frac{\sum_{(i,j)\in\mathcal{P}} \gamma_{ij} + \tau}{2\pi}\chi_1^2 + \frac{\gamma_{pq}}{2\pi}\chi_2^2 \preceq_{st} \frac{1}{2}\chi_1^2 + \frac{\pi - \rho}{2\pi}\chi_2^2 \tag{47}$$

where the upper bound on the left hand side of (47) is attained when τ along with all $\gamma_{ij} \in \mathcal{P}$ are minimized and γ_{pq} is maximized. Combining (46) and (47), we conclude that the RV in (45) is stochastically smaller than the RV

$$\frac{\rho}{2\pi}\chi_0^2 + \frac{1}{2}\chi_1^2 + \frac{\pi - \rho}{2\pi}\chi_2^2 \tag{48}$$

which is itself stochastically smaller than the RV in (43). Thus $c_\alpha < \min\{c_\alpha^{(1)}, \ldots, c_\alpha^{(K)}\}$ so the LRT is more powerful than the IUT. Finally if (III) holds then c_α is the $1-\alpha$ quantile of the RV in (44) for some values of α and of the RV in (45) for others so the LRT is more powerful. Note that in each of the three cases the rejection probability of the null for the LRT is greater than that for the IUT for all values of $\boldsymbol{\theta} \in \boldsymbol{\Theta}_1$ so the LRT is asymptotically uniformly more powerful than the IUT for (1). □