A Robust Test for Elliptical Symmetry

Ilya Soloveychik

The Hebrew University of Jerusalem

Abstract—**Most signal processing and statistical applications heavily rely on specific data distribution models. The Gaussian distributions, although being the most common choice, are inadequate in most real world scenarios as they fail to account for data coming from heavy-tailed populations or contaminated by outliers. Such problems call for the use of Robust Statistics. The robust models and estimators are usually based on elliptical populations, making the latter ubiquitous in all methods of robust statistics. To determine whether such tools are applicable in any specific case, goodness-of-fit (GoF) tests are used to verify the ellipticity hypothesis. Ellipticity GoF tests are usually hard to analyze and often their statistical power is not particularly strong. In this work, assuming the true covariance matrix is unknown we design and rigorously analyze a robust GoF test consistent against all alternatives to ellipticity on the unit sphere. The proposed test is based on Tyler's estimator and is formulated in terms of easily computable statistics of the data. For its rigorous analysis, we develop a novel framework based on the exchangeable random variables calculus introduced by de Finetti. Our findings are supported by numerical simulations comparing them to other popular GoF tests and demonstrating the significantly higher statistical power of the suggested technique.**

Index Terms—**Goodness-of-fit test, elliptical population, Tyler's estimator, robust statistics, exchangeable random variables.**

I. INTRODUCTION

The majority of methods and techniques used by statistical signal processing and data science heavily rely on various assumptions on the data such as independence of the samples, certain parametric families of possible distributions, etc. Very rarely these assumptions are confirmed on the observed samples and even if such verification attempt is made the data almost never agrees with the assumptions made. This leads to poor inference, or even to situations in which the researcher does not know the quality of the achieved results. The main reason for the lack of such tests is the technical complexity of their analysis especially when the data is far from being Gaussian. The most popular substitute for Gaussian distributions in such cases is the family of elliptical populations. Indeed, the latter is already quite flexible to reasonably approximate the heavy tails of the real-world populations or the outliers, while still allowing rigorous analytical treatment. In this paper, we focus on one of the fundamental questions accompanying any problem of signal processing or statistical inference. Namely, how reliably does the elliptical family of distributions model the data at hand. In other words, we design a novel and easy to use Goodness-of-Fit (GoF) test that efficiently and consistently validates the ellipticity assumptions. Given independently sampled data, such a test quantitatively and reliably determines whether we can assume the data to be elliptically distributed. Since an elliptical distribution is determined by its density generator function and scatter matrix[1], estimation of the scatter (or covariance) matrix of the population becomes a prerequisite for almost any ellipticity GoF test. Below we give a detailed exposition of the covariance estimation problem in such scenarios, and provide a detailed explanation of our construction.

A. Tests for Elliptical Symmetry

Numerous ellipticity GoF tests have been proposed in the statistical and signal processing literature, however, most of them lack statistical power as discussed in Section VII below in detail. Such tests are also rarely supported by provable analysis, since such analysis often becomes infeasible when the Gaussian assumptions are lifted. In addition, the computational schemes of some of these testing procedures are so complex that their usage becomes computationally infeasible even on modern machines.

In [1] Beran introduces a test based on marginal signs and ranks. That test is neither distribution-free within the family of elliptical populations nor affine-invariant. In addition, the authors do not provide practical guidelines to the choice of the basis functions involved in the test statistic making its application difficult. Baringhaus [2] proposes a test for spherical symmetry of Cramer-von Mises type based on the independence between the norm of the samples and their directions. The asymptotic distribution of this test is very hard to achieve and exploit. Dyckerhoff et al. [3] demonstrated empirically that this test can be used as a test for elliptical symmetry in dimension 2. Koltchinskii and Sakhanenko [4] design tests using the bootstrap methodology. These tests are based on a class of functions closed under orthogonal transformations and have no known asymptotic distribution, thus requiring bootstrap to get the critical values. Manzotti et al. [5] develop a test based on spherical harmonics to test whether the normalized vectors are uniformly distributed on the unit sphere. The test is computationally demanding and requires moments of order 4. Schott [6] builds a Wald-type test to compare the empirical fourth-order moments with the expected ones under elliptical symmetry. The test is relatively simple in usage since it is based on the moments of low order. However, it has very low power against several alternatives. The authors of [7] propose a Pearson χ^2-type test with multidimensional cells. Its asymptotic distribution exists only for Gaussian scenario, otherwise bootstrap techniques are required. Cassart et al. [8] construct a pseudo-Gaussian test that is most efficient against a multivariate form of Fechner-type asymmetry. The test requires finite moments of order 4.

---

[1]We assume populations have zero mean, which is a very natural assumption in most applications.

Tests based on Monte Carlo simulations can be found in [9] and [10]. The authors of [11] exploit graphical methods while [12] build conditional tests. We refer the reader to [13] and [14] for extensive surveys and performance analysis of the aforementioned tests.

B. Covariance Estimation

Elliptical GoF tests are almost never possible without explicit (e.g. in the plug-in form or through whitening the data) or implicit (e.g. spherical/isotropic GoF tests) estimation or stipulation of the covariance structure. In all of these tests, whenever the scatter matrix is unknown it must be estimated from the available data. In this section we provide a brief survey of the non-Gaussian covariance estimation literature and focus attention on Tyler's estimator which is later used in our construction.

Covariance estimation is a fundamental problem of its own in multivariate statistical analysis. It arises in diverse applications such as signal processing [15, 16], geometric functional analysis and computational geometry [17], genomics [18–20], functional MRI [21], modern social networks analysis [22, 23], empirical finance [24, 25], classical problems of clustering and discriminant analysis [26], and many other fields. Application of structured covariance matrices instead of Bayesian classifiers based on Gaussian mixture densities or kernel densities proved to be very efficient for many pattern recognition tasks, among them speech recognition, machine translation and object recognition [27].

As mentioned earlier, in most real world applications the Gaussian models become unacceptable and robust covariance estimation methods that allow the populations to be heavy-tailed or contain a small proportion of outliers are required [28, 29]. In the 70-s through the analysis of elliptical populations and their Maximum Likelihood (ML) estimates, R. A. Maronna discovered a family of covariance M-estimators [29]. These estimators turned out to be much more robust to outliers than the classical sample covariance which is the ML estimator in the Gaussian setup. The ideas of Maronna were further developed by D. E. Tyler who derived a distribution-free robust covariance matrix estimator [30]. This estimator fits any population from the Generalized Elliptical (GE) family [31] and is also a member of the M-estimators family. Tyler's estimator has become very widely used by engineers [32–35] since its discovery. Although, generally M-estimators are given as solutions to optimization programs, Tyler showed that his estimator can be obtained as a solution to a simple fixed point equation

$$\mathbf{T} = \frac{p}{n} \sum_{i=1}^{n} \frac{\mathbf{x}_i \mathbf{x}_i^\top}{\mathbf{x}_i^\top \mathbf{T}^{-1} \mathbf{x}_i}, \tag{1}$$

where $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^p$ are the collected sample vectors. To avoid the obvious scaling ambiguity (for a solution $\mathbf{T}$ to (1), $c \cdot \mathbf{T}$ is also a solution whenever $c > 0$), it is common to fix the scaling, e.g. by setting $\mathrm{Tr}(\mathbf{T}) = p$. When $\{\mathbf{x}_i\}_{i=1}^n$ are i.i.d. (independent and identically distributed) elliptical [31], their true scatter matrix $\mathbf{\Omega}$ is positive definite and $n > p$, Tyler's estimator exists with probability one and is a consistent estimator of $\mathbf{\Omega}$. In [36] Tyler also demonstrated that his

estimator can be viewed as an ML estimator of a certain distribution over a unit sphere. In elliptical populations the scatter matrix is equal to a positive multiple of the covariance matrix when the latter exists.

The elliptical and generalized elliptical classes of distributions are quite large to incorporate many known populations and they model the real non-Gaussian world behavior much better [32, 33, 37, 38] than the Gaussian distributions. In particular, the GE family includes generalized Gaussian, compound Gaussian and many other widely used distributions [31]. Elliptical populations are commonly used to model radar clutter [39], noise and interference in indoor and outdoor mobile communication channels [40] and numerous other applications.

Other robust covariance estimation approaches were also proposed, however, they have not become so much popular as M-estimators. Among them is the Stahel-Donoho estimate [41, 42], whose main idea is to detect and down-weight outliers based on their one dimensional "*outlyingness measure*". Another method proposed by P. J. Rousseeuw [43] is the Minimum Volume Ellipsoid estimate, whose name stems from the fact that among all ellipsoids containing at least half of the data points, the one defined by the Minimum Volume Ellipsoid estimate has the minimal volume. A more efficient approach, the so-called S-estimator, was later proposed by P. L. Davies [44] and deeply investigated by H. P. Lopuhaa [45] and O. Hössjer [46]. The Minimum Covariance Determinant estimate [47] is another possibility of robust covariance estimation. When prior knowledge on the estimator is available to the research, a Bayesian covariance estimator [48–50] would become natural. Shrinkage estimators in the paradigm of James-Stein estimator are a particular case of the Bayesian methodology. Shrinkage estimators of covariance matrices are computed as a conical (often convex) combination of a certain data statistic (e.g. sample covariance) and a constant matrix (e.g. the identity matrix) representing the prior [18, 51]. Robust analogs of numerous shrinkage estimators were also recently developed and thoroughly studied [52, 53].

The behavior of Tyler's estimator had been methodically investigated in various asymptotic regimes and multiple high-probability performance bounds have been developed for its analysis [34, 53–62]. However, all of these results only hold if the samples are elliptically distributed, which is easily achievable in simulation studies but can hardly be guaranteed in real applications. Therefore a much more practical question can be formulated as follows: Given the data, verify that the ellipticity assumptions can be applied to it and therefore the tools of robust statistics will yield meaningful and reliable results when applied to it. This is the question we address in our work focusing on Tyler's estimator.

C. Our Approach and Contribution

Many of the GoF tests method mentioned in Section I-A use plug-in estimates of the covariance matrix to whiten the samples before applying GoF test of uniformity over the unit sphere. Usually such plug-in estimates are based on the sample covariance matrix (e.g. [4–7] and numerous

others) which creates an unbalanced situation. Indeed, such tests assume ellipticity of the population, thus allowing it to have heavy-tails, but the tools used to estimate the covariance are not robust and therefore no performance guarantees in finite samples can be offered by these tests. In other cases, robust estimates are used, however, without rigorous studies and claims because of the significant level of complexity of the required analysis. When testing the ellipticity hypothesis, it is common to separate the question of elliptical symmetry from the radial density [31, 36]. Such a standard approach enables one to project the samples on the unit sphere as described in Sections II and III in detail and study the distribution of their normalized values. Remarkably, modulo positive scaling, such a transformation does not affect the scatter matrix of an elliptical (or GE) population and it is the use of this very technique that led to the discovery of Tyler's estimator as an ML estimate of the scatter matrix [36]. Following these ideas - for the scenario of unknown scatter matrix - we exploit Tyler's estimator to develop an asymptotically consistent GoF test against all alternatives to ellipticity on the unit p-dimensional sphere.[2] To enable analytical treatment of such a hypothesis test, we reformulate it as an asymptotic uniformity test for a certain stochastically dependent sequence of unit random vectors. The main tool used in the construction and analysis of the uniformity tests for i.i.d. scenario is the Central Limit Theorem (CLT) [63–65] which is clearly not applicable when the measurements are not independent. For our setup, we develop a novel toolbox that allows verification of the null hypothesis by resorting to the concept of *exchangeability*.

A sequence of variables is called *exchangeable* if the joint distribution of any finite subset of these variables is invariant under arbitrary permutations of their indices. Exchangeable random variables were first introduced by de Finetti [66, 67] as a direct and natural generalization of i.i.d. sequences. Interestingly, exchangeable random variables serve as one of the fundamental building blocks of the Bayesian statistics [68]. Unlike the i.i.d. case, the behavior of exchangeable sequences is much harder to analyze. We exploit certain versions of CLT and the Strong Law of Large Numbers (SLLN) for exchangeable variables to demonstrate asymptotic consistency of our test statistics built analogously to the generalized Ajne and Giné statistics [63, 64, 69] developed for the i.i.d. case.[3]

Following Tyler, our approach becomes essentially *distribution-free* within the elliptical family since we do not focus on estimating the radial density function [71]. We offer a test which is consistent against *all* alternatives to elliptical symmetry and not only certain classes of densities [5, 8, 71]. We do not use the sample covariance matrix as a plug-in estimator in particular because its convergence

to the true covariance in elliptical populations maybe very slow due to heavy tails [5], however, we emphasize that our technique *allows* the use of the sample covariance instead of Tyler's estimator. Our test is not limited to certain moments which makes it more natural and less computationally demanding [6, 8]. Unlike Monte Carlo simulations-based techniques [9, 10], our methodology is rigorous and offers a deterministic and computationally cheap algorithm. Finally, unlike previously mentioned works (and articles referenced in [71]) we believe that the methodology based on the *exchangeability* framework is the most suitable for the analysis of populations transformed by plug-in estimators. This type of analysis is usually technically more complex but in our eyes it represents the natural approach to the problem. Our theoretical studies are supported by extensive numerical simulations featuring the properties of the tests and comparing them to other available tests mentioned in earlier in Section I-A.

The rest of the article is organized as follows. In Section II, we introduce the setup and notation. The problem is formulated in Section III where we also present some of the existing tests for the known scatter case. In Section IV, we reformulate the problem and introduce necessary background on exchangeable random variables; Section IV-D provides some additional notation and auxiliary results. In Section V, we formulate the main results and discuss them. Section VI provides numerical studies of the proposed tests while Section VII demonstrates their power in comparison to other commonly used ellipticity tests. The conclusion is provided in Section VIII. Some of the proofs are postponed to the Appendix.

II. Notation and Setup

Definition 1 ([72])**.** *A vector $\mathbf{y} \in \mathbb{R}^p$ is elliptically distributed with the scatter matrix $\mathbf{\Omega} \succ 0$ and mean $\boldsymbol{\mu}$ if there exists a random vector $\mathbf{w} \in \mathcal{S}^{p-1}$ uniformly distributed over the unit $p-1$-dimensional sphere and an independent random variable $r \geqslant 0$, such that*

$$\mathbf{y} = \boldsymbol{\mu} + r \cdot \mathbf{\Omega}^{1/2}\mathbf{w}. \tag{2}$$

For example, if $r \sim \sqrt{\chi_p^2}$, then $\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Omega})$. In what follows we always assume that the data is centered, $\boldsymbol{\mu} = 0$. Let us consider the normalized vector,

$$\mathbf{x} = \frac{\mathbf{y}}{\|\mathbf{y}\|} = \frac{\mathbf{\Omega}^{1/2}\mathbf{w}}{\|\mathbf{\Omega}^{1/2}\mathbf{w}\|}, \tag{3}$$

which can be equivalently viewed as disregarding the information stored in the scalar variable r but keeping the information provided by the scatter matrix. As we see below, the distribution of $\mathbf{x}$ contains all the information about the scatter matrix $\mathbf{\Omega}$. We are going to recover the scatter matrix by sampling from the distribution of $\mathbf{x}$. Denote by $\mathbf{I} = \mathbf{I}_p$ the p-dimensional identity matrix.

---

[2]The unknown distribution is assumed to be a Lebesgue measurable probability measure over the p-dimensional Euclidean sphere as elaborated further. By *all alternatives*, it is common to understand the set of all probability measures except for the class under consideration, which in our case is the class of elliptical (or generalized elliptical) populations [4, 7, 14, 63].

[3]The Ajne statistic was originally introduced for distributions on a circle [69], the idea was extended by [70] to the 2-dimensional unit sphere and later generalized by [64] for the $p-1$-dimensional spheres. Similarly, Giné's statistic was originally defined for 1- and 2-dimensional spheres and later extended by [64] to the general dimension.

Definition 2 ([36]). *The family of real Angular Central Gaussian (ACG) distributions on $\mathcal{S}^{p-1}$ is defined by the densities of the form*

$$p(\mathbf{x}; \boldsymbol{\Omega}) = \frac{\Gamma(p/2)}{2\pi^{p/2} |\boldsymbol{\Omega}|^{1/2}} \frac{1}{(\mathbf{x}^\top \boldsymbol{\Omega}^{-1} \mathbf{x})^{p/2}}, \quad \mathbf{x} \in \mathcal{S}^{p-1}, \quad (4)$$

for $\boldsymbol{\Omega} \succ 0$ which is called the scatter matrix.

When $\mathbf{x}$ is ACG distributed with the scatter matrix $\boldsymbol{\Omega}$, we write

$$\mathbf{x} \sim \mathcal{U}(\boldsymbol{\Omega}), \quad (5)$$

in particular when $\boldsymbol{\Omega} = \mathbf{I}$ we get the uniform distribution over the unit sphere $\mathcal{U}(\mathbf{I})$. Note that ACG is not a member of the elliptical family but actually belongs to a wider class of *generalized* elliptical populations whose definition is identical to Definition 1 except for weakened assumptions on r [72]. In generalized elliptical population, r does not have to be stochastically independent of $\mathbf{w}$ and does not have to be non-negative. The following result allows us to reduce estimation of the scatter matrices of elliptical populations to the estimation of the scatter matrices of ACG vectors.

Lemma 1 ([72]). *For a random vector $\mathbf{y}$ sampled from a centered elliptical population with the scatter matrix $\boldsymbol{\Omega}$, $\mathbf{x}$ defined in (3) is ACG distributed with the same scatter matrix.*

Now assume $n > p$ i.i.d. random vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathcal{S}^{p-1}$ are sampled from $\mathcal{U}(\boldsymbol{\Omega})$, then as shown in [36] the ML estimator of the scatter matrix exists almost surely and is given by the fixed point equation (1). The solutions to this equation form a ray since the latter is invariant under multiplication of the matrix $\mathbf{T}$ by a positive constant. To resolve the ambiguity we choose $\mathbf{T}$ to satisfy $\operatorname{Tr}(\mathbf{T}) = p$, however, we note that the specific choice of the scaling does not affect any of the results presented below.

III. PROBLEM FORMULATION AND STATE OF THE ART

A. Main Goal

The problem considered in this article can be formulated as follows. Given a sequence of vectors $\{\mathbf{x}_i\}_{i=1}^n \subset \mathcal{S}^{p-1}$ sampled independently, we want to test two alternative hypotheses,

$$\mathcal{H}_0 : \mathbf{x}_1, \ldots, \mathbf{x}_n \overset{\text{i.i.d.}}{\sim} \mathcal{U}(\boldsymbol{\Omega}), \text{ for some } \boldsymbol{\Omega}, \quad (6)$$

$$\mathcal{H}_1 : \mathbf{x}_1, \ldots, \mathbf{x}_n \overset{\text{i.i.d.}}{\nsim} \mathcal{U}(\boldsymbol{\Omega}), \text{ for any } \boldsymbol{\Omega}, \quad (7)$$

and in the case of $\mathcal{H}_0$ we want to estimate the scatter matrix $\boldsymbol{\Omega}$, as well.

The test (6)-(7) is a composite hypothesis since the scatter matrix is unknown. When the scatter matrix is known, the problem can be equivalently reformulated as a uniformity test on the sphere as shown below.

B. Uniformity Tests on $\mathcal{S}^{p-1}$

Assume the scatter matrix $\boldsymbol{\Omega}$ in the hypothesis test (6)-(7) is known and introduce a derived i.i.d. sequence,

$$\mathbf{w}_i = \frac{\boldsymbol{\Omega}^{-1/2} \mathbf{x}_i}{\left\| \boldsymbol{\Omega}^{-1/2} \mathbf{x}_i \right\|}, \quad i = 1, \ldots, n. \quad (8)$$

Under $\mathcal{H}_0$, $\mathbf{w}_1, \ldots, \mathbf{w}_n \sim \mathcal{U}(\mathbf{I})$ and therefore the test (6)-(7) becomes actually a uniformity test on the unit sphere,

$$\mathcal{G}_0 : \mathbf{w}_1, \ldots, \mathbf{w}_n \overset{\text{i.i.d.}}{\sim} \mathcal{U}(\mathbf{I}), \quad (9)$$

$$\mathcal{G}_1 : \mathbf{w}_1, \ldots, \mathbf{w}_n \overset{\text{i.i.d.}}{\nsim} \mathcal{U}(\mathbf{I}). \quad (10)$$

Next, we summarize two uniformity tests on $\mathcal{S}^{p-1}$ concluding this section with Proposition 3 providing a uniformity test consistent against all alternatives on the unit sphere. Based on it, we will develop an analogous test for (6)-(7) with unknown scatter matrix in the subsequent sections. Denote by

$$V_{p-1} = \int_{\mathbf{x} \in \mathbb{S}^{p-1}} d\mathbf{x} = \frac{2\pi}{\Gamma\left(\frac{p}{2}\right)} \quad (11)$$

the area of the unit sphere. In addition, by

$$\psi_{ij} = \arccos(\mathbf{x}_i^\top \mathbf{x}_j) \quad (12)$$

we denote the angular separation (the shortest great circle distance) between $\mathbf{x}_i$ and $\mathbf{x}_j$ and by

$$N(\mathbf{y}) = |\{\mathbf{x}_i \mid \mathbf{y}^\top \mathbf{x}_i \geqslant 0\}|, \quad \mathbf{y} \in \mathbb{S}^{p-1}, \quad (13)$$

the number of points falling into the hemisphere with the pole at $\mathbf{y}$. Denote also

$$\alpha = \frac{p}{2} - 1, \quad (14)$$

$$\nu(a, b) = \binom{a + b - 2}{a - 1} + \binom{a + b - 1}{a - 1}. \quad (15)$$

The following two popular statistics and detailed investigation of their behavior can be found in [63, 69]. These results were later generalized in [64] and summarized in [65].

Proposition 1 (Generalized Ajne Test, [64, 69]). *Under the uniformity hypothesis, the Ajne statistic*

$$t_A = \frac{1}{nV_{p-1}} \int_{\mathbf{y} \in \mathbb{S}^{p-1}} \left(N(\mathbf{y}) - \frac{n}{2}\right)^2 d\mathbf{y} = \frac{n}{4} - \frac{1}{\pi n} \sum_{i<j} \psi_{ij}$$

is asymptotically distributed as $\mathcal{L}\left(\sum_{q=1}^\infty a_{2q-1}^2 K_{\nu(p-1,2q-1)}\right)$, where K_ξ are independent random variables distributed as χ_ξ^2 and

$$a_{2q-1} = \frac{(-1)^{q-1} 2^{p-2} \Gamma(\alpha+1) \Gamma(q+\alpha)(2q-2)}{\pi(q-1)!(2q+p-3)!}. \quad (16)$$

Proposition 2 (Generalized Giné Test, [63, 64]). *Under the uniformity hypothesis, the Giné statistic*

$$t_G = \frac{n}{2} - \frac{p-1}{2n} \left(\frac{\Gamma\left(\alpha+\frac{1}{2}\right)}{\Gamma(\alpha+1)}\right)^2 \sum_{i<j} \sin(\psi_{ij}) \quad (17)$$

is asymptotically distributed as $\mathcal{L}\left(\sum_{q=1}^\infty a_{2q}^2 K_{\nu(p-1,2q)}\right)$, where K_ξ are independent random variables distributed as χ_ξ^2 and

$$a_{2q}^2 = \frac{(p-1)(2q-1)}{8\pi(2q+p-1)} \left(\frac{\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma\left(q-\frac{1}{2}\right)}{\Gamma\left(q+\alpha+\frac{1}{2}\right)}\right)^2. \quad (18)$$

The following statement provides a concise and directly applicable test for uniformity under the assumption that the random vectors are sampled i.i.d. from $\mathcal{U}(\mathbf{I})$.

Proposition 3 (Uniformity test, [63, 64])**.** *Any weighted sum of t_A and t_G is consistent against all alternatives to uniformity on $\mathbb{S}^{p-1}$.*

In practice, one way to make the decision about accepting or rejecting $\mathcal{H}_0$ or $\mathcal{G}_0$ is as follows. The statistician truncates the series mentioned in the last two propositions in a data-driven manner and compares the sample values of t_A and t_G with the tables (or explicit numerical approximations) of the corresponding distributions. Another more general approach consists in replacing t_A and t_G by statistics whose expansions only have finite number of non-zero coefficients a_k (see [63] for more details). An efficient data-driven approach to the design of the uniformity tests based on a modification of the Bayesian Information Criterion was developed by [73]. Often in practice, the distributions of the statistics at hand under the null hypothesis are estimated empirically, by generating samples from the latter. This is the approach adopted in Section VII below.

In this paper we are interested in the case of unknown scatter matrix in (6)-(7). As we see below this makes the hypothesis test much more involved. In the next sections we develop analogs of generalized Ajne and Giné uniformity tests for this scenario.

IV. Problem Reformulation and Exchangeability

A. Methodology

From Theorem 3.1 from [30] we know that under $\mathcal{H}_0$ Tyler's estimator converges almost surely to the true scatter matrix when $n \to \infty$. This idea motivated our study of a new sequence of vectors, defined as follows. Under $\mathcal{H}_0$ introduced in (6), we now consider the sequence

$$\mathbf{t}_i = \frac{\mathbf{T}^{-1/2}\mathbf{x}_i}{\left\|\mathbf{T}^{-1/2}\mathbf{x}_i\right\|} \in \mathbb{S}^{p-1}, \quad i = 1, \ldots, n, \qquad (19)$$

where $\mathbf{T}$ is defined in (1). The main challenge we face in the study of $\{\mathbf{t}_i\}$ is the lack of stochastic independence unlike the case of $\{\mathbf{w}_i\}$ defined in (8). Indeed, most existing convergence results explicitly rely on independence in their derivations in such a way that any deviation from this assumption ruins the performance analysis. For example, all the results of Ajne, Giné, and Prentice utilize the CLT and thus require independence as the most crucial assumption [63–65, 69].

Next we include a brief summary of the exchangeability concept and the related toolbox. We then use it in Section IV-D to overcome the loss of independence in our analysis of the consistency of $\{\mathbf{t}_i\}$ and their statistics.

B. Exchangeable Random Variables

Definition 3. *Given a sequence $\{X_i\}$ (finite or infinite) of random variables, we say that it is exchangeable if the joint distribution of any finite subset of variables is invariant under arbitrary permutations of their indices.*

In other words, exchangeability is our indifference to the order of the measurements. This is clearly a much weaker hypothesis than independence, as any i.i.d. sequence is obviously exchangeable. In his seminal works de Finetti [66, 67] demonstrated that in certain sense every (infinite) exchangeable sequence can be represented as a composition of sequences of i.i.d. variables. This result can be viewed as the analog of Fourier decomposition in analysis, as it allows one to represent a *more complicated* exchangeable sequence as a superposition of basic building blocks - independent sequences - objects much easier accessible for analysis and reasoning.

De Finetti [66, 67] and some of his followers focused on infinite exchangeable sequences. There exist, however, finite sets of exchangeable random variables which cannot be embedded into infinite sequences, these are called *finitely exchangeable* or *non-extendable*. The analysis of extendable sequences can be reduced to the analysis of infinite sequences. On the other hand, the non-extendable sequences require quite different approaches [74]. Our sequence of samples $\{\mathbf{t}_i\}_{i=1}^n$ is an example of a non-extendable exchangeable sequence of random vectors. Indeed, their order obviously does not matter since $\mathbf{T}$ is not affected by permutations of the measurements $\{\mathbf{x}_i\}_{i=1}^n$. We can also see that this sequence is non-extendable, since addition of new random vectors $\mathbf{x}_j$ without an amendment of $\mathbf{T}$ will turn the sequence into non-exchangeable. For a detailed study of non-extendability we refer the reader to [74] and references therein.

The main result of our paper can be briefly summarized as follows. We demonstrate that the limiting behavior of the samples $\{\mathbf{t}_i\}_{i=1}^n$ is in certain sense analogous to the behavior of the vectors uniformly distributed over the unit sphere and therefore, we can apply similar tools for the hypothesis tests. Below we show how to overcome the technical challenges on this way.

C. Limit Theorems for Exchangeable Variables

To illustrate the previous section and better describe the nature of the exchangeability phenomenon and its relation to the stochastic independence, in this section we present analogs of the SLLN and CLT for triangular arrays of exchangeable variables.

Lemma 2 (Strong Law of Large Numbers for Exchangeable Arrays). *Let $\{X_{ni}\}_{n,i=1}^{\infty,n}$ be a triangular array of row-wise exchangeable random variables and $\{X_{\infty i}\}_{i=1}^{\infty}$ be a sequence of exchangeable random variables of bounded second moment such that*

1) $X_{n1} \xrightarrow{a.s.} X_{\infty 1}, \ n \to \infty,$
2) $\mathrm{var}\,(X_{n1} - X_{\infty 1}) \to 0, \ n \to \infty,$
3) $\mathbb{E}\,[X_{n1}X_{n2}] \to 0, \ n \to \infty.$

Then

$$\frac{1}{n}\sum_{i=1}^n X_{ni} \xrightarrow{a.s.} 0, \ n \to \infty. \qquad (20)$$

Proof. The proof can be found in the Appendix. $\qquad\square$

Let $k_n < n$ be two sequences of natural numbers such that

$$\frac{k_n}{n} \to \gamma \in [0, 1). \qquad (21)$$

Lemma 3 (Central Limit Theorem for Exchangeable Arrays, Theorem 2 from [75][4]). *Let $\{X_{ni}\}_{n,i=1}^{\infty,n}$ be a triangular array of row-wise exchangeable random variables such that*

1) $\mathbb{E}\left[X_{n1}X_{n2}\right] \to 0$, $n \to \infty$,

2) $\max_{1 \leqslant i \leqslant n} \frac{|X_{ni}|}{\sqrt{n}} \xrightarrow{P} 0$, $\forall n$,

3) $\frac{1}{k_n}\sum_{i=1}^{k_n} X_{ni}^2 \xrightarrow{P} 1$, $n \to \infty$.

Then

$$\sqrt{k_n}\left[\frac{1}{k_n}\sum_{i=1}^{k_n} X_{ni} - \frac{1}{n}\sum_{i=1}^{n} X_{ni}\right] \xrightarrow{L} \mathcal{N}(0, 1-\gamma), \quad n \to \infty.$$

As mentioned earlier this result provides an analog of the CLT for exchangeable sequences. However, it is important to stress its distinction from the classical CLT-type claims for the i.i.d. variables. Indeed, Lemma 3 only allows us to consider a subset of the sample of cardinality k_n smaller than the number of variables n in the row so that even their ratio must not approach one. This is a reflection of the essential difference between non-extendable exchangeable sequences and their extendable counterparts that include i.i.d. sequences as a particular case [74].

D. Additional Notation and Auxiliary Results

Assume that an infinite i.i.d. sequence $\{\mathbf{x}_i\}_{i=1}^{\infty}$ is sampled under the composite $\mathcal{H}_0$ with the true scatter matrix is unknown. For every $n > p$, let the sequence of corresponding Tyler's estimators be

$$\mathbf{T}_n = \frac{p}{n}\sum_{i=1}^{n}\frac{\mathbf{x}_i\mathbf{x}_i^\top}{\mathbf{x}_i^\top\mathbf{T}_n^{-1}\mathbf{x}_i}, \quad n = p+1, \ldots, \quad (22)$$

which exist almost surely for a random sample [37, 76]. Consider a triangular array of row-wise exchangeable random vectors

$$\mathbf{t}_{ni} = \frac{\mathbf{T}_n^{-1/2}\mathbf{x}_i}{\left\|\mathbf{T}_n^{-1/2}\mathbf{x}_i\right\|} \in \mathbb{S}^{p-1}, \quad i = 1, \ldots, n, \ n = p+1, \ldots. \quad (23)$$

Note that by Definition 1, the sequence $\{\mathbf{x}_i\}_{i=1}^{\infty}$ can equivalently be defined as follows. Given a sequence $\{\mathbf{w}_i\}_{i=1}^{\infty} \sim \mathcal{U}(\mathbf{I}_p)$ of uniform i.i.d. random vectors, we look at their transforms

$$\mathbf{x}_i = \frac{\mathbf{\Omega}^{1/2}\mathbf{w}_i}{\left\|\mathbf{\Omega}^{1/2}\mathbf{w}_i\right\|}, \quad (24)$$

for some fixed but unknown $\mathbf{\Omega} \succ 0$. Define also an auxiliary sequence

$$\mathbf{t}_{\infty i} = \mathbf{w}_i. \quad (25)$$

Lemma 4. *With the notation introduced above,*

$$\mathbf{t}_{ni} \xrightarrow{a.s.} \mathbf{t}_{\infty i}, \quad n \to \infty. \quad (26)$$

Proof. The proof can be found in the Appendix. $\square$

We are now interested in the empirical distributions of the rows of the obtained triangular array, which are the finite sets

[4]To simplify the notation we assume the number of the elements in the n-th row to be n unlike the seemingly more general case of m_n variables considered in [75].

$\{\mathbf{t}_{ni}\}_{i=1}^{n}$ for every fixed $n > p$. The following CLT-type result holds in our scenario.

Proposition 4. *For the triangular array of vectors $\{\mathbf{t}_{ni}\}_{n=p+1,i=1}^{\infty,n}$ defined above,*

$$\sqrt{p}\cdot\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{t}_{ni} \xrightarrow{L} \mathcal{N}(0, \mathbf{I}), \quad n \to \infty. \quad (27)$$

Proof. The proof can be found in the Appendix. $\square$

Corollary 1. *Under $\mathcal{H}_0$, for any differentiable function $f\colon \mathbb{S}^{p-1} \to \mathbb{R}$,*

$$\sqrt{p}\cdot\frac{1}{\sqrt{n}}\sum_{i=1}^{n}f(\mathbf{t}_{ni}) \xrightarrow{L} \mathcal{N}(0, \|\nabla f(0)\|^2), \quad n \to \infty. \quad (28)$$

Proof. The proof follows the i.i.d. case verbatim using the Maclaurin expansion of f. $\square$

V. ASYMPTOTIC UNIFORMITY TESTS FOR EXCHANGEABLE VECTORS

In Section III-B we introduced statistics t_A and t_G to test the null hypothesis of uniformity for independent samples over the unit sphere $\mathbb{S}^{p-1}$. Our next statements constitute analogs of those result for the row-wise exchangeable array $\{\mathbf{t}_{ni}\}_{n=p+1,i=1}^{\infty,n}$. Denote

$$\psi_{n,ij} = \arccos(\mathbf{t}_{ni}^\top\mathbf{t}_{nj}). \quad (29)$$

Proposition 5 (Generalized Ajne Test for $\mathbf{t}_{ni}$). *Under $\mathcal{H}_0$, the Ajne statistic*

$$t_A\left(\{\mathbf{t}_{ni}\}\right) = \frac{n}{4} - \frac{1}{\pi n}\sum_{i<j}\psi_{n,ij} \quad (30)$$

is asymptotically distributed as $\mathcal{L}\left(\sum_{q=1}^{\infty} a_{2q-1}^2 K_{\nu(p-1,2q-1)}\right)$ as $n \to \infty$, where K_ξ are independent random variables distributed as χ_ξ^2 and

$$a_{2q-1} = \frac{(-1)^{q-1}2^{p-2}\Gamma(\alpha+1)\Gamma(q+\alpha)(2q-2)}{\pi(q-1)!(2q+p-3)!}. \quad (31)$$

Proof. The proof follows [63] and [64] verbatim using Corollary 1. $\square$

Definition 4. *Random variable X is said to be first-order stochastically dominated by random variable Y if*

$$\mathbb{P}[X \in A] \leqslant \mathbb{P}[Y \in A], \quad (32)$$

for any measurable set A.

Proposition 6 (Generalized Giné Test for $\bar{\mathbf{t}}_{nk}$). *Under $\mathcal{H}_0$, the Giné statistic*

$$t_G\left(\{\mathbf{t}_{ni}\}\right) = \frac{n}{2} - \frac{p-1}{2n}\left(\frac{\Gamma\left(\alpha+\frac{1}{2}\right)}{\Gamma(\alpha+1)}\right)^2\sum_{i<j}\sin(\psi_{n,ij}) \quad (33)$$

is asymptotically first-order stochastically dominated by the random variable distributed as $\sum_{q=1}^{\infty} a_{2q}^2 K_{\nu(p-1,2q)}$, where K_ξ are independent random variables distributed as χ_ξ^2 and

$$a_{2q}^2 = \frac{(p-1)(2q-1)}{8\pi(2q+p-1)}\left(\frac{\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma\left(q-\frac{1}{2}\right)}{\Gamma\left(q+\alpha+\frac{1}{2}\right)}\right)^2, \quad (34)$$

and

$$\mathbb{E}\left[t_G\left(\{\mathbf{w}_i\}\right)\right] - \mathbb{E}\left[t_G\left(\{\mathbf{t}_{ni}\}\right)\right] \sim \frac{1}{8} + \frac{1}{16p} + O\left(\frac{1}{p^2}\right), \quad n \to \infty.$$

Proof. The proof can be found in the Appendix. □

It is important to emphasize the main difference between Propositions 2 and 6. Indeed, in the former the asymptotic distribution is given by a sum of scaled χ^2-variables, while in the second one the limiting distribution is first-order stochastically dominated by the same distribution and is in fact significantly *thinner*. This discrepancy is due to the fact that $\{\mathbf{t}_{ni}\}$ are dependent in such a special way that their sample covariance matrix is exactly $\mathbf{I}$. The detailed reasoning and discussion can be found in the Appendix.

Theorem 1 (Uniformity Test for $\mathbf{t}_{ni}$). *Under $\mathcal{H}_0$, any weighted sum of $t_A\left(\{\mathbf{t}_{ni}\}\right)$ and $t_G\left(\{\mathbf{t}_{ni}\}\right)$ is consistent against all alternatives to the asymptotic uniformity of $\{\mathbf{t}_{ni}\}$ on $\mathbb{S}^{p-1}$.*

Proof. The proof follows [63] and [64] verbatim using Propositions 5 and 6. □

Remark 1. *It is also important to emphasize that an asymptotic bound to the power of the uniformity test suggested by Theorem 1 against all alternatives on the unit sphere can be easily constructed using the asymptotic normality of the scaled deviations of the Ajne and Giné statistics as shown in Section 4 of [63]. These derivations are also valid in our case, and therefore we omit them due to a lack of space.*

VI. NUMERICAL SIMULATIONS

In this section, we investigate the behavior and advantages of the criterion proposed in Theorem 1 through numerical simulations.

A. Distributions of the Statistics under the Null Hypothesis

In the first experiment, we compared the empirical distributions of $t_A\left(\{\mathbf{t}_{ni}\}\right)$ and $t_G\left(\{\mathbf{t}_{ni}\}\right)$ with their counterparts $t_A\left(\{\mathbf{w}_i\}\right)$ and $t_G\left(\{\mathbf{w}_i\}\right)$ for the independent samples playing the role of the benchmarks. In this simulation we took the true scatter matrix to be the identity $\mathbf{\Omega} = \mathbf{I}$. Figure 1 demonstrates the anticipated in Section V difference in the behavior of the Ajne and Giné statistics. More specifically, as claimed in Proposition 6 and discussed in detail in its proof, the statistic $t_G\left(\{\mathbf{t}_{ni}\}\right)$ is first-order stochastically dominated by $t_G\left(\{\mathbf{w}_i\}\right)$ due to the difference in the behavior of the quadratic term in the expansion of the statistic (75) caused by dependencies among $\{\mathbf{t}_{ni}\}$. This is in contrast to the Ajne statistic whose distributions in both cases coincide since a similar expansion into Gegenbauer polynomials involves odd degree polynomials only (for more details see the proof of Proposition 6). Note also that the theoretically predicted by Theorem 1 difference between the expected values for $p = 8$,

$$\mathbb{E}\left[t_G\left(\{\mathbf{w}_i\}\right)\right] - \mathbb{E}\left[t_G\left(\{\mathbf{t}_{ni}\}\right)\right] \sim \frac{1}{8} + \frac{1}{16p} + O\left(\frac{1}{p^2}\right) \approx 0.133 \tag{35}$$

is confirmed by the numerical simulation yielding the value of 0.131.
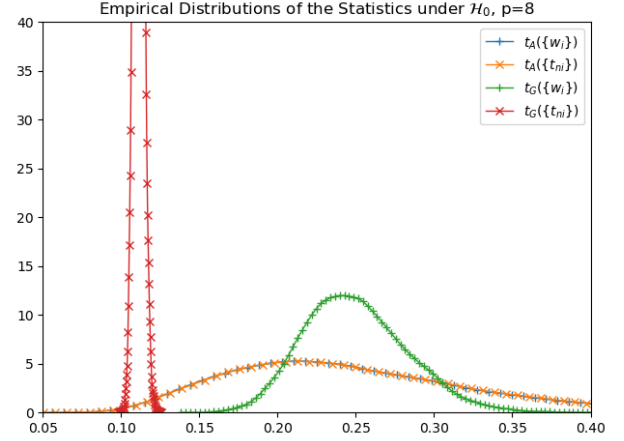

Fig. 1: Comparison of the empirical distributions of the Ajne and Giné test statistics computed for the sequences $\{\mathbf{w}_i\}$ and $\{\mathbf{t}_{ni}\}$ defined in Section IV-D with the true scatter matrix being $\mathbf{I}$.

B. Criterion Performance for Alternatives

Following Theorem 1, to demonstrate the power of the suggested methodology in our second experiment we compared the empirical distributions of the statistics

$$s\left(\{\mathbf{t}_{ni}\}\right) = t_A\left(\{\mathbf{t}_{ni}\}\right) + t_G\left(\{\mathbf{t}_{ni}\}\right), \tag{36}$$

$$s\left(\{\mathbf{w}_i\}\right) = t_A\left(\{\mathbf{w}_i\}\right) + t_G\left(\{\mathbf{w}_i\}\right) \tag{37}$$

under the null hypotheses $\mathcal{H}_0$ and $\mathcal{G}_0$ versus their distributions under specific non-elliptical alternatives $\mathcal{H}_1$ and $\mathcal{G}_1$. The alternatives were constructed as follows. We generated the uniform sequence $\{\mathbf{w}_i\}$ as before, added a constant offset to all the obtained vectors and re-normalized them,

$$\widetilde{\mathbf{w}}_i = \frac{\mathbf{w}_i + \mathbf{a}}{\|\mathbf{w}_i + \mathbf{a}\|}. \tag{38}$$

Note that the distributions of $\widetilde{\mathbf{w}}_i$ and of $\mathbf{x}_i$ constructed from it via (24)

$$\mathbf{x}_i = \frac{\mathbf{\Omega}^{1/2}\widetilde{\mathbf{w}}_i}{\left\|\mathbf{\Omega}^{1/2}\widetilde{\mathbf{w}}_i\right\|}, \tag{39}$$

are not ACG and therefore our test should be able to discriminate between the hypotheses. In this experiment we chose $p = 5$, $\mathbf{\Omega} = \mathbf{I}$ and

$$\mathbf{a} = 0.05 \cdot \frac{1}{\sqrt{p}} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}. \tag{40}$$

Figure 2 demonstrates the 0.95-confidence bands for the distributions of the statistics $s\left(\{\mathbf{t}_{ni}\}\right)$ and $s\left(\{\mathbf{w}_i\}\right)$ as functions of the number of measurements n in the sample. The bands in Figure 2 were averaged over 50000 independent trials. We see from the graph that despite the small size of $\mathbf{a}$, already with $n = 15$ measurements the criterion allows us to easily discriminate between the hypotheses and its power is similar for both the i.i.d. (know scatter matrix) and Tyler's (unknown scatter matrix) cases.
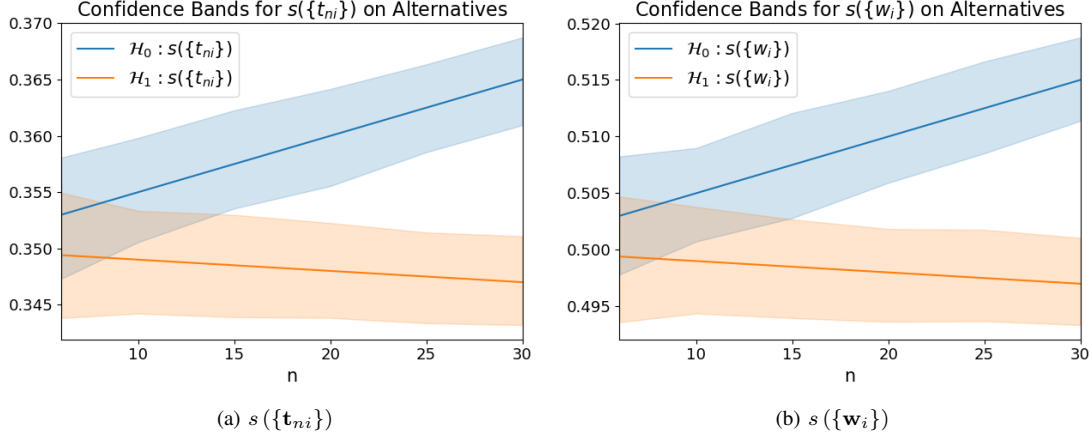
Fig. 2: Comparison of the 0.95-confidence bands for $\mathcal{H}_0$ versus $\mathcal{H}_1$ designed in section VI-B for the i.i.d. $\{\mathbf{w}_i\}$ samples and their exchangeable $\{\mathbf{t}_{ni}\}$ counterparts.

VII. COMPARISON TO OTHER TESTS

For the purpose of this comparative study, we selected the most popular tests used in the literature due to Koltchinskii and Sakhanenko [4, 14], Manzotti et al. (MPQ test) [5], Cassart (PseudoGaussian test) [8], Schott [6], and Babic (SkewOptimal test) [71]. Implementations of all these tests are available online trough the *ellipticalsymmetry*[5] package developed by [77].

In our numerical experiment, we set the dimension to $p = 5$ and the number of samples in every batch $n = 50$. Each test from the aforementioned list and the tests proposed in this paper (from Propositions 5 and 6 and Theorem 1) is invoked to classify each such batch and determine whether it is coming from an elliptical population or not. The goal is to compare the powers of the tests using their ROC (Receiver Operating Characteristic) curves. To this end, we constructed $N = 5000$ sample batches (each having n vectors of dimension p), half of which were coming from the standard normal population and labeled as *elliptical batch*, and the rest form a non-elliptical population labeled as *non-elliptical batch* and defined as follows. Let

$$\mathbf{X}_{i,k} \sim \mathcal{N}(0, \mathbf{I}_p), \quad i = 1, \ldots, N, \ k = 1, \ldots, n, \quad (41)$$

be N batches of n vectors each with all $\mathbf{X}_{i,k}$ i.i.d. standard normal. Now, we set

$$\mathbf{Y}_i = \mathbf{X}_i, \quad i = 1, \ldots, N/2 \quad (42)$$

for the first $N/2$ batches labeled as *elliptical batch* and another $N/2$ batches marked as *non-elliptical batch* we sample from a non-elliptical population designed as follows. Let

$$\mathbf{Z}_{j,i,k} \sim \mathcal{N}(\mathbf{z}_j, \mathbf{I}_p), \\ j = 1, 2, \ i = N/2 + 1, \ldots, N, \ k = 1, \ldots, n, \quad (43)$$

with the population means

$$\mathbf{z}_1 = 5 \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{z}_2 = 5 \cdot \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ \vdots \\ 1 \\ -1 \end{bmatrix} \quad (44)$$

be all i.i.d. and define

$$\mathbf{Y}_i = \mathbf{X}_i + \mathbf{Z}_1 + \mathbf{Z}_2, \quad i = N/2 + 1, \ldots, N. \quad (45)$$

Clearly, for $N/2 + 1, \ldots, N$ the vectors in the corresponding batches $\mathbf{Y}_i$ are non-elliptical and dont become GE even if normalized by their Euclidean norms. Therefore, we would expect all our tests to discriminate between *elliptical* and *non-elliptical* batches.

Figure 3 shows the ROC (Receiver Operating Characteristic) curves for all the available tests. Naturally, the highest (having maximal power β) curve at each point on the horizontal axis corresponds to the most powerful test at this False Positive Rate (α). We can see that the tests proposed in this article are more powerful than the rest. The distributions of the statistics $t_A(\{\mathbf{t}_{ni}\}), t_G(\{\mathbf{t}_{ni}\}), s(\{\mathbf{t}_{ni}\})$ under $\mathcal{H}_0$ derived in Propositions 5 and 6 were estimated empirically for this numerical study. We would like to emphasize a number of points. First, the ROC curves of all the reference tests are very close to the bisector of the first quadrant, showing their very low statistical power and making them almost equivalent to fair coin tossing. Second, every test involving $t_G(\{\mathbf{t}_{ni}\})$ has a horizontal ROC curve and statistical power of 1 for every α, meaning that it correctly classified every batch. This is due to the sharp concentration of the distribution of this statistic around a constant value value as can be seen in Figure 1. Because of that, any deviation from the expected value is a strong evidence against the ellipticity hypothesis.
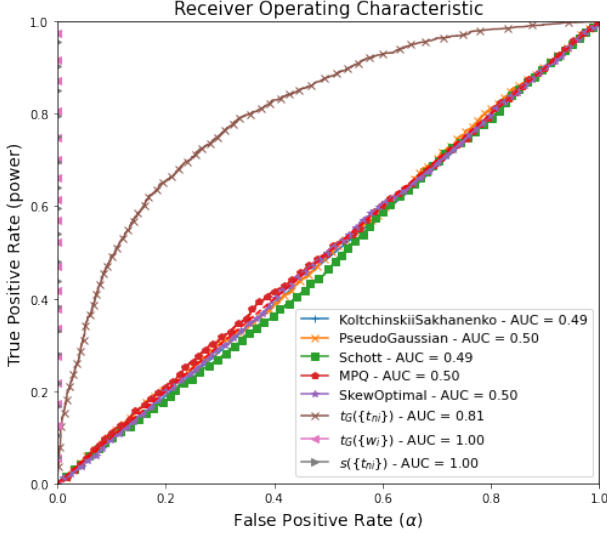
$\square$

Fig. 3: ROC curves of Koltchinskii-Sakhanenko, MPQ, PseudoGaussian, Schott, and SkewOptimal tests in comparison to the test statistic $s\left(\{\mathbf{t}_{ni}\}\right)$ defined in (36) and justified by Theorem 1.

VIII. CONCLUSION

In this paper we propose a novel elliptical symmetry test based on the ideas of robust statistics, and specifically Tyler's estimator of covariance matrix. This is an easy to apply and computationally cheap test with provable performance guarantees. In our extensive comparative studies we demonstrate that it surpasses all the other commonly exploited ellipticity tests by a large margin when it comes to their statistical power. In addition, based on the exchangeable random variables calculus introduced by de Finetti, we develop a natural mathematical framework enabling rigorous analysis of our test and numerous other tests and estimators based on dependent but exchangeable sample measurements.

APPENDIX

Proof of Lemma 2. Our proof is based on an analogous result in [78]. Both Lemmas 1 and 2 from [78] can be easily restated for our setup after replacing the Banach space E by $\mathbb{R}$ and linear functionals by scalar multiplication. In addition, note that our condition 1) immediately implies requirement (2.5) from [78]. Now, the reasoning from the proof of Theorem 1 from [78] applies verbatim. $\square$

Proof of Lemma 4. As shown in Theorem 3.1 from [30],

$$\mathbf{T}_n \xrightarrow{a.s.} \mathbf{\Omega} \succ 0, \quad n \to \infty, \tag{46}$$

therefore, starting from some n_0, $\mathbf{T}_n$ is almost surely invertible for $n \geqslant n_0$ and

$$\mathbf{T}_n^{-1/2}\mathbf{\Omega}^{1/2} \xrightarrow{a.s.} \mathbf{I}_p, \quad n \to \infty. \tag{47}$$

Now the claim follows from the definition of the sequence $\{\mathbf{t}_{ni}\}_n$,

$$\mathbf{t}_{ni} = \frac{\mathbf{T}_n^{-1/2}\mathbf{x}_i}{\left\|\mathbf{T}_n^{-1/2}\mathbf{x}_i\right\|} = \frac{\mathbf{T}_n^{-1/2}\mathbf{\Omega}^{1/2}\mathbf{w}_i}{\left\|\mathbf{T}_n^{-1/2}\mathbf{\Omega}^{1/2}\mathbf{w}_i\right\|} \xrightarrow{a.s.} \mathbf{w}_i, \quad n \to \infty.$$

Proof of Proposition 4. As above, we can equivalently rewrite $\mathbf{t}_{ni}$ as

$$\mathbf{t}_{ni} = \frac{\mathbf{T}_n^{-1/2}\mathbf{\Omega}^{1/2}\mathbf{w}_i}{\left\|\mathbf{T}_n^{-1/2}\mathbf{\Omega}^{1/2}\mathbf{w}_i\right\|}, \quad i = 1, \ldots, n, \ \ n = p+1, \ldots,$$

which is just a useful representation as clearly $\mathbf{\Omega}$ is not revealed to the researcher. Fix a vector $\mathbf{a} \in \mathbb{R}^p$ of unit norm $\|\mathbf{a}\| = 1$ and consider the following triangular array of row-wise exchangeable random variables

$$X_{ni} = \sqrt{p} \cdot \mathbf{a}^\top \mathbf{t}_{ni}, \quad i = 1, \ldots, n, \ \ n = p+1, \ldots. \tag{48}$$

Let us study the properties of $\{X_{ni}\}_{n=p+1,i=1}^{\infty,n}$. First, consider

$$\mathbb{E}\left[X_{n1}X_{n2}\right] = p\,\mathbf{a}^\top \mathbb{E}\left[\mathbf{t}_{n1}\mathbf{t}_{n2}^\top\right]\mathbf{a}, \tag{49}$$

Lemma 4 implies that

$$\mathbb{E}\left[\mathbf{t}_{n1}\mathbf{t}_{n2}^\top\right] \to \mathbb{E}\left[\mathbf{w}_1\mathbf{w}_2^\top\right] = 0, \quad n \to \infty, \tag{50}$$

therefore,

$$\mathbb{E}\left[X_{n1}X_{n2}\right] \to 0, \quad n \to \infty. \tag{51}$$

Next, note that

$$\frac{|X_{ni}|}{\sqrt{n}} = \sqrt{p}\frac{|\mathbf{a}^\top \mathbf{t}_{ni}|}{\sqrt{n}} \leqslant \sqrt{p}\frac{\|\mathbf{a}\|\,\|\mathbf{t}_{ni}\|}{\sqrt{n}} = \sqrt{\frac{p}{n}} \to 0. \tag{52}$$

Finally, let us show that

$$\frac{1}{k_n}\sum_{i=1}^{k_n} X_{ni}^2 \xrightarrow{P} 1, \quad n \to \infty. \tag{53}$$

Indeed,

$$\frac{1}{k_n}\sum_{i=1}^{k_n} X_{ni}^2 = p\frac{1}{k_n}\sum_{i=1}^{k_n}\mathbf{a}^\top \mathbf{t}_{ni}\mathbf{t}_{ni}^\top\mathbf{a} = p\,\mathbf{a}^\top\left[\frac{1}{k_n}\sum_{i=1}^{k_n}\mathbf{t}_{ni}\mathbf{t}_{ni}^\top\right]\mathbf{a}.$$

For the sample covariance we obtain,

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\mathbf{t}_{ni}\mathbf{t}_{ni}^\top = \frac{1}{k_n}\sum_{i=1}^{k_n}\frac{\mathbf{T}_n^{-1/2}\mathbf{\Omega}^{1/2}\mathbf{w}_i\mathbf{w}_i^\top\mathbf{\Omega}^{1/2}\mathbf{T}_n^{-1/2}}{\left\|\mathbf{T}_n^{-1/2}\mathbf{\Omega}^{1/2}\mathbf{w}_i\right\|^2}$$

$$= \mathbf{T}_n^{-1/2}\mathbf{\Omega}^{1/2}\left[\frac{1}{k_n}\sum_{i=1}^{k_n}\frac{\mathbf{w}_i\mathbf{w}_i^\top}{\left\|\mathbf{T}_n^{-1/2}\mathbf{\Omega}^{1/2}\mathbf{w}_i\right\|^2}\right]\mathbf{\Omega}^{1/2}\mathbf{T}_n^{-1/2}$$

$$\xrightarrow{a.s.} \frac{1}{p}\mathbf{I}, \quad n \to \infty, \tag{54}$$

and therefore,

$$\frac{1}{k_n}\sum_{i=1}^{k_n} X_{ni}^2 \xrightarrow{P} \mathbf{a}^\top\mathbf{I}\,\mathbf{a} = \|\mathbf{a}\|^2 = 1, \quad n \to \infty. \tag{55}$$

The rest of the proof is based on the argument proposed in [79]. Assume without loss of generality that n is an even number and set

$$k_n = \frac{n}{2}. \tag{56}$$

Consider now the following sequence,

$$Y_{ni} = \begin{cases} X_{ni}, & i \leqslant k_n, \\ -X_{ni}, & i > k_n. \end{cases} \tag{57}$$

Clearly, the new sequence is exchangeable with the same joint distribution as the original sequence. Indeed, under $\mathcal{H}_0$ the joint distribution of $\{\mathbf{t}_{ni}\}$ is invariant under multiplication of any of the random vectors by -1 and $\mathbf{T}_n$ is an even function of $\mathbf{t}_{ni}$. Now all the conditions of Lemma 3 are satisfied for the new sequence $\{Y_{ni}\}$ and we obtain,

$$\sqrt{\frac{n}{2}} \left[\frac{2}{n} \sum_{i=1}^{n/2} Y_{ni} - \frac{1}{n} \sum_{i=1}^{n} Y_{ni} \right] \xrightarrow{L} \mathcal{N}\left(0, \frac{1}{2}\right), \quad n \to \infty. \tag{58}$$

Note that

$$\frac{2}{n} \sum_{i=1}^{n/2} Y_{ni} - \frac{1}{n} \sum_{i=1}^{n} Y_{ni} \tag{59}$$

$$= \frac{2}{n} \sum_{i=1}^{n/2} X_{ni} - \frac{1}{n} \left[\sum_{i=1}^{n/2} X_{ni} - \sum_{i=n/2+1}^{n} X_{ni} \right]$$

$$= \frac{1}{n} \sum_{i=1}^{n/2} X_{ni} + \frac{1}{n} \sum_{i=n/2+1}^{n} X_{ni} = \frac{1}{n} \sum_{i=1}^{n} X_{ni},$$

to obtain

$$\sqrt{\frac{n}{2}} \cdot \frac{1}{n} \sum_{i=1}^{n} X_{ni} \xrightarrow{L} \mathcal{N}\left(0, \frac{1}{2}\right), \quad n \to \infty, \tag{60}$$

or

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_{ni} \xrightarrow{L} \mathcal{N}(0, 1), \quad n \to \infty. \tag{61}$$

By the definition of X_{ni} we get

$$\sqrt{p} \cdot \mathbf{a}^\top \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{t}_{ni} \xrightarrow{L} \mathcal{N}(0, 1), \quad n \to \infty. \tag{62}$$

Finally, recall that the vector $\mathbf{a}$ was chosen arbitrarily to conclude the proof. □

For a sequence of vectors $\{\mathbf{y}_i\}_{i=1}^n$, denote their sample covariance by

$$\mathbf{S}_{\mathbf{y},n} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{y}_i \mathbf{y}_i^\top. \tag{63}$$

Lemma 5 (Theorem 6.2 from [80]). *Let $\{\mathbf{y}_i\}_{i=1}^n \subset \mathbb{S}^{p-1}$ be a set of $n \geqslant p$ vectors, then*

$$\operatorname{Tr}\left(\mathbf{S}_{\mathbf{y},n}^2\right) \geqslant \frac{1}{p}. \tag{64}$$

Proof of Proposition 6. The difference in the behavior of the Ajne and Giné statistics stems from the fact that the former is a sum of Gegenbauer polynomials of odd orders involving only monomials of odd powers, while the latter reads as a sum of Gegenbauer polynomials of even orders involving only monomials of even powers [63, 64]. Next we explain this in more detail.

Gegenbauer (ultraspherical) polynomial [81] of index α and order $q \geqslant 2$ is defined as

$$C_q^\alpha(z) = \sum_{k=0}^{\lfloor q/2 \rfloor} (-1)^k \frac{\Gamma(q-k+\alpha)}{\Gamma(\alpha)k!(q-2k)!} (2z)^{q-2k}. \tag{65}$$

Note that the Gegenbauer polynomials of odd/even order involves monomials of only odd/even order, respectively.

In order to analyze the Giné statistic (33), we use the following expansion of $\sin\theta$ into Gegenbauer polynomials in $\cos\theta$ from [64],

$$\frac{1}{2} - \frac{p-1}{2} \left(\frac{\Gamma\left(\alpha + \frac{1}{2}\right)}{\Gamma(\alpha+1)}\right)^2 \sin\theta$$

$$= \sum_{q=1}^{\infty} \frac{(p-1)(2q-1)(4q+p-2)}{(p-2)(2q+p-1)8\pi}$$

$$\times \left(\frac{\Gamma\left(\alpha + \frac{1}{2}\right)\Gamma\left(q - \frac{1}{2}\right)}{\Gamma(q+\alpha+\frac{1}{2})}\right)^2 C_{2q}^\alpha(\cos\theta),$$

where we remind the reader that

$$\alpha = \frac{p}{2} - 1. \tag{66}$$

Using (65) we can write,

$$\frac{1}{2} - \frac{p-1}{2} \left(\frac{\Gamma\left(\alpha + \frac{1}{2}\right)}{\Gamma(\alpha+1)}\right)^2 \sin\theta = \sum_{r=0}^{\infty} \gamma_{2r}(\alpha, p) \cos^{2r}\theta$$

$$= \gamma_0(\alpha, p) + \gamma_2(\alpha, p) \cos^2\theta + \sum_{r=2}^{\infty} \gamma_{2r}(\alpha, p) \cos^{2r}\theta.$$

Below we use the explicit form of $\gamma_2(\alpha, p)$,

$$\gamma_2(\alpha, p) = \sum_{q=1}^{\infty} \frac{(p-1)(2q-1)(4q+p-2)}{(p-2)(2q+p-1)8\pi}$$

$$\times \left(\frac{\Gamma\left(\alpha + \frac{1}{2}\right)\Gamma\left(q - \frac{1}{2}\right)}{\Gamma\left(q+\alpha+\frac{1}{2}\right)}\right)^2 \zeta_{2q,2}^\alpha, \tag{67}$$

where

$$\zeta_{2q,2}^\alpha = 2(-1)^{q-1} \frac{\Gamma(q+1+\alpha)}{\Gamma(\alpha)(q-1)!}, \tag{68}$$

is the weight of z^2 in $C_{2q}^\alpha(z)$ defined as in (65). Thus, we obtain

$$\gamma_2(\alpha, p) = \sum_{q=1}^{\infty} (-1)^{q-1} \frac{(p-1)(2q-1)(4q+p-2)}{(p-2)(2q+p-1)4\pi} \tag{69}$$

$$\times \left(\frac{\Gamma\left(\frac{p}{2} - \frac{1}{2}\right)\Gamma\left(q - \frac{1}{2}\right)}{\Gamma\left(q + \frac{p}{2} - \frac{1}{2}\right)}\right)^2 \frac{\Gamma\left(q + \frac{p}{2}\right)}{\Gamma\left(\frac{p}{2} - 1\right)(q-1)!}.$$

Since the last series is telescopic we conclude that in particular

$$\gamma_2(\alpha, p) > 0. \tag{70}$$

Let us compute the first few terms of this series,

$$\gamma_2(\alpha, p) \tag{71}$$

$$= \frac{p(p+2)}{4(p+1)(p-1)} - \frac{3p(p+2)(p+6)}{8(p-1)(p+1)^2(p+3)} + O\left(\frac{1}{p^2}\right)$$

$$= \frac{1}{4} + \frac{1}{8p} + O\left(\frac{1}{p^2}\right).$$

Recall that due to (29),

$$\cos(\psi_{n,ij}) = \mathbf{t}_{ni}^\top \mathbf{t}_{nj}, \tag{72}$$

and therefore,

$$\sin(\psi_{n,ij}) = \sin(\psi_{n,ji}), \tag{73}$$

together with

$$\sin(\psi_{n,ii}) = 0. \tag{74}$$

Now we can see that the Giné statistic reads as

$$
\begin{aligned}
t_G\left(\{\mathbf{t}_{ni}\}\right) &= \frac{n}{2} - \frac{p-1}{2n}\left(\frac{\Gamma\left(\alpha+\frac{1}{2}\right)}{\Gamma(\alpha+1)}\right)^2 \sum_{i<j} \sin(\psi_{n,ij}) \\
&= \frac{n}{2} - \frac{p-1}{4n}\left(\frac{\Gamma\left(\alpha+\frac{1}{2}\right)}{\Gamma(\alpha+1)}\right)^2 \sum_{i,j=1}^{n} \sin(\psi_{n,ij}) \\
&= \frac{n}{2} - \frac{1}{2n}\left[\frac{p-1}{2}\left(\frac{\Gamma\left(\alpha+\frac{1}{2}\right)}{\Gamma(\alpha+1)}\right)^2 \sum_{i,j=1}^{n} \sin(\psi_{n,ij})\right] \\
&= \frac{n}{2} + \frac{1}{2n}\sum_{i,j=1}^{n}\left[\gamma_0(\alpha,p) + \gamma_2(\alpha,p)\cos^2\theta\right. \\
&\qquad\qquad\left. + \sum_{r=2}^{\infty}\gamma_{2r}(\alpha,p)\cos^{2r}\theta - \frac{1}{2}\right] \\
&= \frac{n}{4} + \frac{n\gamma_0(\alpha,p)}{2} + \frac{1}{2n}\sum_{i,j=1}^{n}\left[\gamma_2(\alpha,p)\left(\mathbf{t}_{ni}^\top\mathbf{t}_{nj}\right)^2\right. \\
&\qquad\qquad\left. + \sum_{r=2}^{\infty}\gamma_{2r}(\alpha,p)\left(\mathbf{t}_{ni}^\top\mathbf{t}_{nj}\right)^{2r}\right]. \tag{75}
\end{aligned}
$$

Note that

$$\sum_{i,j=1}^{n}\left(\mathbf{t}_{ni}^\top\mathbf{t}_{nj}\right)^2 = \sum_{k,l=1}^{p}\left[\sum_{i=1}^{n}\mathbf{t}_{ni}^{(k)}\mathbf{t}_{ni}^{(l)}\right]^2 \tag{76}$$

$$= \mathrm{Tr}\left(\left[n\mathbf{S}_{\mathbf{t},n}\right]^2\right) = n^2\mathrm{Tr}\left(\mathbf{S}_{\mathbf{t},n}^2\right),$$

where we denote

$$\mathbf{t} = \begin{bmatrix}\mathbf{t}^{(1)} \\ \vdots \\ \mathbf{t}^{(p)}\end{bmatrix}. \tag{77}$$

In our setup, the sample covariance matrix satisfies the following relation,

$$
\begin{aligned}
\mathbf{S}_{\mathbf{t},n} &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{t}_{ni}\mathbf{t}_{ni}^\top = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{T}_n^{-1/2}\mathbf{x}_i\mathbf{x}_i^\top\mathbf{T}_n^{-1/2}}{\left\|\mathbf{T}_n^{-1/2}\mathbf{x}_i\right\|^2} \tag{78}\\
&= \frac{1}{n}\mathbf{T}_n^{-1/2}\sum_{i=1}^{n}\frac{\mathbf{x}_i\mathbf{x}_i^\top}{\mathbf{x}_i^\top\mathbf{T}_n^{-1}\mathbf{x}_i}\mathbf{T}_n^{-1/2} = \frac{1}{p}\mathbf{I},
\end{aligned}
$$

and therefore,

$$\mathrm{Tr}\left(\mathbf{S}_{\mathbf{t},n}^2\right) = \frac{1}{p^2}\mathrm{Tr}\left(\mathbf{I}\right) = \frac{1}{p}. \tag{79}$$

By Lemma 5, for $\{\mathbf{w}_i\}_{i=1}^{n}$ i.i.d. uniformly distributed over $\mathbb{S}^{p-1}$ with $n \geqslant p$,

$$\mathrm{Tr}\left(\mathbf{S}_{\mathbf{w},n}^2\right) \geqslant \frac{1}{p} = \mathrm{Tr}\left(\mathbf{S}_{\mathbf{t},n}^2\right). \tag{80}$$

Since $\gamma_2(\alpha,p) > 0$, from (75) we infer that $t_G\left(\{\mathbf{w}_i\}\right)$ first-order stochastically dominates $t_G\left(\{\mathbf{t}_{ni}\}\right)$.

Recall that the limiting distribution of the spectrum of $p\,\mathbf{S}_{\mathbf{w},n}$ is given by the Marchenko-Pastur law [82] whose second moment gives us the following asymptotic equivalence,

$$\mathbb{E}\left[\mathrm{Tr}\left(\left[p\,\mathbf{S}_{\mathbf{w},n}\right]^2\right)\right] \sim p + \frac{p^2}{n}, \quad n \to \infty. \tag{81}$$

As a consequence,

$$
\begin{aligned}
&\frac{1}{2n}\left(n^2\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{S}_{\mathbf{w},n}^2\right)\right] - n^2\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{S}_{\mathbf{t},n}^2\right)\right]\right) \\
&\qquad \sim \frac{n^2}{2n}\left(\frac{1}{p} + \frac{1}{n} - \frac{1}{p}\right) \sim \frac{1}{2}, \quad n \to \infty,
\end{aligned}
$$

and from (75) we conclude,

$$
\begin{aligned}
\mathbb{E}\left[t_G\left(\{\mathbf{w}_i\}\right)\right] - \mathbb{E}\left[t_G\left(\{\mathbf{t}_{ni}\}\right)\right] &\sim \frac{\gamma_2(\alpha,p)}{2} \\
&\sim \frac{1}{8} + \frac{1}{16p} + O\left(\frac{1}{p^2}\right), \quad n \to \infty. \\
&\qquad\qquad\qquad\qquad\qquad\Box
\end{aligned}
$$

REFERENCES

[1] R. Beran, "Testing for ellipsoidal symmetry of a multivariate density," *The Annals of Statistics*, pp. 150–162, 1979.

[2] L. Baringhaus, "Testing for spherical symmetry of a multivariate distribution," *The Annals of Statistics*, pp. 899–917, 1991.

[3] R. Dyckerhoff, C. Ley, and D. Paindaveine, "Depth-based runs tests for bivariate central symmetry," *Annals of the Institute of Statistical Mathematics*, vol. 67, no. 5, pp. 917–941, 2015.

[4] V. Koltchinskii and L. Sakhanenko, "Testing for ellipsoidal symmetry of a multivariate distribution," *High Dimensional Probability II*, pp. 493–510, 2000.

[5] A. Manzotti, F. J. Pérez, and A. J. Quiroz, "A statistic for testing the null hypothesis of elliptical symmetry," *Journal of Multivariate Analysis*, vol. 81, no. 2, pp. 274–285, 2002.

[6] J. R. Schott, "Testing for elliptical symmetry in covariance-matrix-based analyses," *Statistics & Probability Letters*, vol. 60, no. 4, pp. 395–404, 2002.

[7] F. W. Huffer and C. Park, "A test for elliptical symmetry," *Journal of Multivariate Analysis*, vol. 98, no. 2, pp. 256–281, 2007.

[8] D. Cassart, M. Hallin, and D. Paindaveine, "Optimal detection of Fechner-asymmetry," *Journal of Statistical Planning and Inference*, vol. 138, no. 8, pp. 2499–2525, 2008.

[9] C. Diks and H. Tong, "A test for symmetries of multivariate probability distributions," *Biometrika*, vol. 86, no. 3, pp. 605–614, 1999.

[10] L. X. Zhu and G. Neuhaus, "Nonparametric Monte Carlo tests for multivariate distributions," *Biometrika*, vol. 87, no. 4, pp. 919–928, 2000.

[11] R. Z. Li, K. T. Fang, and L. X. Zhu, "Some QQ probability plots to test spherical and elliptical symmetry," *Journal of Computational and Graphical Statistics*, vol. 6, no. 4, pp. 435–450, 1997.

[12] L. X. Zhu and G. Neuhaus, "Conditional tests for elliptical symmetry," *Journal of Multivariate Analysis*, vol. 84, no. 2, pp. 284–298, 2003.

[13] R. J. Serfling, "Multivariate symmetry and asymmetry," *Encyclopedia of statistical sciences*, vol. 8, pp. 5338–5345, 2006.

[14] L. Sakhanenko, "Testing for ellipsoidal symmetry: A comparison study," *Computational statistics and data analysis*, vol. 53, no. 2, pp. 565–581, 2008.

[15] E. J. Kelly, "An adaptive detection algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, no. 2, pp. 115–127, 1986.

[16] I. Soloveychik and A. Wiesel, "Group symmetric robust covariance estimation," *IEEE Transactions on Signal Processing*, vol. 64, no. 1, pp. 244–257, 2016.

[17] R. Adamczak, A. Litvak, A. Pajor, and N. Tomczak-Jaegermann, "Quantitative estimates of the convergence of the empirical covariance matrix in log-concave ensembles," *Journal of American Mathematic Society*, vol. 23, no. 2, pp. 535–561, 2010.

[18] J. Schäfer and K. Strimmer, "A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics," *Statistical applications in genetics and molecular biology*, vol. 4, no. 1, 2005.

[19] J. Xie and P. M. Bentler, "Covariance structure models for gene expression microarray data," *Structural Equation Modeling*, vol. 10, no. 4, pp. 566–582, 2003.

[20] A. Hero and B. Rajaratnam, "Hub discovery in partial correlation graphs," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 6064–6078, 2012.

[21] G. Derado, F. D. Bowman, and C. D. Kilts, "Modeling the spatial and temporal dependence in fMRI data," *Biometrics*, vol. 66, no. 3, pp. 949–957, 2010.

[22] S. L. Lauritzen, "Graphical models," *Oxford University Press*, 1996.

[23] O. Banerjee, L. El Ghaoui, and A. d'Aspremont, "Model selection through sparse maximum likelihood estimation for multivariate Gaussian or binary data," *The Journal of Machine Learning Research*, vol. 9, pp. 485–516, 2008.

[24] J. Bai and S. Shi, "Estimating high dimensional covariance matrices and its applications," *Annals of Economics and Finance*, vol. 12, no. 2, pp. 199–215, 2011.

[25] O. Ledoit and M. Wolf, "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection," *Journal of Empirical Finance*, vol. 10, no. 5, pp. 603–621, 2003.

[26] J. H. Friedman, "Regularized discriminant analysis," *Journal of the American statistical association*, vol. 84, no. 405, pp. 165–175, 1989.

[27] J. Dahmen, D. Keysers, M. Pitz, and H. Ney, "Structured covariance matrices for statistical image object recognition," *Mustererkennung*, pp. 99–106, 2000.

[28] P. J. Huber, "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964.

[29] R. A. Maronna, "Robust M-estimators of multivariate location and scatter," *The annals of statistics*, pp. 51–67, 1976.

[30] D. E. Tyler, "A distribution-free M-estimator of multivariate scatter," *The Annals of Statistics*, vol. 15, no. 1, pp. 234–251, 1987.

[31] G. Frahm, "Generalized elliptical distributions: theory and applications," *Universität zu Köln*, 2004.

[32] Y. I. Abramovich, N. K. Spencer, and M. D. Turley, "Time-varying autoregressive (TVAR) models for multiple radar observations," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1298–1311, 2007.

[33] F. Pascal, Y. Chitour, J. P. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound Gaussian noise: Existence and algorithm analysis," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 34–48, 2008.

[34] F. Bandiera, O. Besson, and G. Ricci, "Knowledge-aided covariance matrix estimation and adaptive detection in compound-Gaussian noise," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5391–5396, 2010.

[35] A. Wiesel and T. Zhang, "Structured robust covariance estimation," *Foundations and Trends in Signal Processing*, vol. 8, no. 3, pp. 127–216, 2015.

[36] D. E. Tyler, "Statistical analysis for the angular central Gaussian distribution on the sphere," *Biometrika*, vol. 74, no. 3, pp. 579–589, 1987.

[37] A. Wiesel, "Geodesic convexity and covariance estimation," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6182–6189, 2012.

[38] T. Zhang, A. Wiesel, and M. S. Greco, "Multivariate generalized gaussian distribution: Convexity and graphical models," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 4141–4148, 2013.

[39] E. Conte, M. Longo, and M. Lops, "Modelling and simulation of non-Rayleigh radar clutter," *IEE Proceedings F on Radar and Signal Processing*, vol. 138, no. 2, pp. 121–130, 1991.

[40] D. Middleton, "Man-made noise in urban environments and transportation systems: Models and measurements," *IEEE Transactions on Communications*, vol. 21, no. 11, pp. 1232–1241, 1973.

[41] W. A. Stahel, "Breakdown of covariance estimators," *Fachgruppe für Statistik, Eidgenössische Techn. Hochsch.*, 1981.

[42] D. L. Donoho, "Breakdown properties of multivariate location estimates," *PhD qualifying paper, Dept. Statistics, Harvard Univ.*, 1982.

[43] P. J. Rousseeuw, "Multivariate estimation with high breakdown point," *Mathematical Statistics and Applications*, vol. 8, pp. 283–297, 1985.

[44] P. L. Davies, "Asymptotic behaviour of S-estimates of multivariate location parameters and dispersion matrices," *Annals of Statistics*, pp. 1269–1292, 1987.

[45] H. P. Lopuhaa, "On the relation between S-estimators and M-estimators of multivariate location and covariance," *Annals of Statistics*, pp. 1662–1683, 1989.

[46] O. Hössjer, "On the optimality of S-estimators," *Statistics and Probability Letters*, vol. 14, no. 5, pp. 413–419, 1992.

[47] P. J. Rousseeuw and K. V. Driessen, "A fast algorithm for

the minimum covariance determinant estimator," *Technometrics*, vol. 41, no. 3, pp. 212–223, 1999.

[48] T. Leonard and J. S. J. Hsu, "Bayesian inference for a covariance matrix," *The Annals of Statistics*, vol. 20, no. 4, pp. 1669–1696, 1992.

[49] I. Alvarez, J. Niemi, and M. Simpson, "Bayesian inference for a covariance matrix," *arXiv preprint arXiv:1408.4050*, 2014.

[50] R. Yang and J. O. Berger, "Estimation of a covariance matrix using the reference prior," *The Annals of Statistics*, pp. 1195–1211, 1994.

[51] O. Ledoit and M. Wolf, "Honey, i shrunk the sample covariance matrix," *The Journal of Portfolio Management*, vol. 30, no. 4, pp. 110–119, 2004.

[52] Y. Chen, A. Wiesel, and A. O. Hero, "Robust shrinkage estimation of high-dimensional covariance matrices," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4097–4107, 2011.

[53] R. Couillet and M. R. McKay, "Large dimensional analysis and optimization of robust shrinkage covariance matrix estimators," *arXiv preprint arXiv:1401.4083*, 2014.

[54] I. Soloveychik and A. Wiesel, "Performance analysis of Tyler's covariance estimator," *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 418–426, 2015.

[55] O. Besson and Y. I. Abramovich, "On the fisher information matrix for multivariate elliptically contoured distributions," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1130–1133, 2013.

[56] M. Greco and F. Gini, "Cramér-Rao lower bounds on covariance matrix estimation for complex elliptically symmetric distributions," *IEEE Transactions on Signal Processing*, vol. 61, no. 24, pp. 6401–6409, 2013.

[57] L. Duembgen, "The asymptotic behavior of Tyler's M-estimator of scatter in high dimension," *Ann. Inst. Statist. Math*, vol. 50, pp. 471–491, 1997.

[58] G. Frahm and K. Glombek, "Semicircle law of Tyler's M-estimator for scatter," *Statistics & Probability Letters*, vol. 82, no. 5, pp. 959–964, 2012.

[59] T. Zhang, X. Cheng, and A. Singer, "Marchenko-Pastur Law for Tyler's and Maronna's M-estimators," *arXiv preprint arXiv:1401.3424*, 2014.

[60] R. Couillet, F. Pascal, and J. W. Silverstein, "Robust M-estimation for array processing: a random matrix approach," *IEEE Transactions on Signal Process*, 2012.

[61] ——, "Robust estimates of covariance matrices in the large dimensional regime," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7269–7278, 2014.

[62] F. Pascal, Y. Chitour, J.-P. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound Gaussian noise: Existence and algorithm analysis," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 34–48, 2007.

[63] E. Giné, "Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms," *The Annals of statistics*, pp. 1243–1266, 1975.

[64] M. J. Prentice, "On invariant tests of uniformity for directions and orientations," *The Annals of Statistics*, vol. 6, no. 1, pp. 169–176, 1978.

[65] E. Garcia-Portugués and T. Verdebout, "An overview of uniformity tests on the hypersphere," *arXiv preprint arXiv:1804.00286*, 2018.

[66] B. De Finetti, "Funzione caratteristica di un fenomeno aleatorio," *Atti del Congresso Internazionale dei Matematici: Bologna del 3 al 10 de settembre di 1928*, pp. 179–190, 1929.

[67] ——, "La prévision: ses lois logiques, ses sources subjectives," *Annales de l'institut Henri Poincaré*, vol. 7, no. 1, pp. 1–68, 1937.

[68] D. M. Cifarelli and E. Regazzini, "De finetti's contribution to probability and statistics," *Statistical Science*, pp. 253–282, 1996.

[69] B. Ajne, "A simple test for uniformity of a circular distribution," *Biometrika*, vol. 55, no. 2, pp. 343–354, 1968.

[70] R. J. Beran, "Testing for uniformity on a compact homogeneous space," *Journal of Applied Probability*, vol. 5, no. 1, pp. 177–195, 1968.

[71] S. Babic, L. Gelbgras, M. Hallin, and C. Ley, "Optimal tests for elliptical symmetry: specified and unspecified location," *ULB–Universite Libre de Bruxelles*, 2019.

[72] G. Frahm and U. Jaekel, "Tyler's M-estimator, random matrix theory, and generalized elliptical distributions with applications to finance," *Technical report, Helmut Schmidt University*, 2007.

[73] P. E. Jupp, "Data-driven Sobolev tests of uniformity on compact Riemannian manifolds," *The Annals of Statistics*, vol. 36, no. 3, pp. 1246–1260, 2008.

[74] T. Konstantopoulos and L. Yuan, "On the extendibility of finitely exchangeable probability measures," *Transactions of the American Mathematical Society*, vol. 371, no. 10, pp. 7067–7092, 2019.

[75] N. C. Weber, "A martingale approach to central limit theorems for exchangeable random variables," *Journal of Applied Probability*, vol. 17, no. 3, pp. 662–673, 1980.

[76] A. Wiesel, "Unified framework to regularized covariance estimation in scaled gaussian models," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 29–38, 2012.

[77] S. Babić, C. Ley, and M. Palangetić, "Elliptical symmetry tests in R," *arXiv preprint arXiv:2011.12560*, 2020.

[78] R. L. Taylor and R. F. Patterson, "Strong laws of large numbers for arrays of row-wise exchangeable random elements," *International Journal of Mathematics and Mathematical Sciences*, vol. 8, no. 1, pp. 135–144, 1985.

[79] I. Soloveychik, "Central limit theorem for symmetric exchangeable random variables," *arXiv preprint arXiv:2006.10819*, 2020.

[80] J. J. Benedetto and M. Fickus, "Finite normalized tight frames," *Advances in Computational Mathematics*, vol. 18, no. 2-4, pp. 357–385, 2003.

[81] M. Reimer, "Multivariate polynomial approximation," *Birkhäuser*, vol. 144, 2012.

[82] V. A. Marcenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Sbornik: Mathematics*, vol. 1, no. 4, pp. 457–483, 1967.