Estimating axial symmetry using random projections

Alejandro Cholaquidis[a], Ricardo Fraiman[a], Manuel Hernandez-Banadik[b,*],
Stanislav Nagy[c]

[a]*Facultad de Ciencias, Universidad de la República*
[b]*Facultad de Ciencias Económicas y Administración, Universidad de la República*
[c]*Faculty of Mathematics and Physics, Charles University*

Abstract

This paper studies the problem of identifying directions of axial symmetry in multivariate distributions. Theoretical results are derived on how the measure or cardinality of the set of symmetry directions relates to spherical symmetry. The problem is framed using random projections, leading to a proof that in $\mathbb{R}^2$, agreement on two random projections is enough to identify the true axes of symmetry. A corresponding result for higher dimensions is conjectured. An estimator for the symmetry directions is proposed and proved to be consistent in the plane.

Keywords: axial symmetry, multivariate symmetry, random projections, Cramér-Wold theorem, symmetry estimation
2020 MSC: 62H12, 62G05

1. Introduction

Symmetry of probability distributions has been a well-studied property. Multivariate distributions can exhibit various types of symmetries, each characterized by invariance under a specific group of actions (see, for instance, [24]). Understanding the symmetry of a distribution is valuable in several contexts.

In medical image analysis, detecting deviations from axial symmetry is crucial, as any asymmetry could indicate pathological changes in an organ's shape. Several studies have focused on analyzing symmetry in medical images. For example, Martos and de Carvalho [16] explore symmetry analysis

---

*Corresponding author. Email address: m.hernandez.banadik@gmail.com

in three-dimensional magnetic resonance brain images, while Hogeweg et al. [14] propose a measure of axial symmetry to quantify asymmetry between lungs in frontal chest radiographs. The medical software CAD4TB version 3, designed for computer-aided detection of tuberculosis, includes a symmetry check in its routine that involves estimating an axis of symmetry [18].

Beyond medical imaging, symmetry plays a significant role in enhancing statistical methods by leveraging invariance constraints, such as in set estimation or density estimation. If the symmetry hypothesis holds, the object to be estimated becomes more constrained in nature, making it easier to estimate. Axial symmetry and related types of symmetry are particularly relevant in directional statistics; for an overview of recent research in this area, refer to Pewsey and García-Portugués [19] and the references therein.

The goal of this paper is dual. First, we theoretically characterize how the size of the set of axial symmetry directions—in terms of its measure or cardinality—implies the spherical symmetry of the entire distribution or of its marginals.

Second, we frame this question within the setting of random projections, seeking to establish a Cramér-Wold-type device for symmetry. Specifically, we investigate whether the axial symmetry of a multivariate distribution can be determined by examining its one-dimensional projections. The core idea is to determine symmetry by verifying if the one-dimensional projections of a distribution agree with those of its transformed version (e.g., reflected across an axis) over a set of directions.

As we demonstrate, this leads to a challenging technical problem. By an extension of the classical Cramér-Wold theorem [5] due to Cuesta-Albertos et al. [6], we know that while two different distributions can project along a direction to the same one-dimensional distribution, the set of such "bad" directions that cause this confusion typically has null measure. This fact is sufficient for hypothesis testing, where one can ignore these null sets (see, for instance, [12, 11] for tests of central symmetry or invariance under finitely generated group actions involving random projections).

Our scenario is different and more involved. Since our goal is to estimate the symmetry directions themselves, we must search the entire space of directions for those where the projections of the distribution and its transformed version agree. In this context, null sets of directions become critically important and cannot be disregarded. The principal question, therefore, asks how many random projections one needs to consider in order to identify the axes of symmetry of a multivariate distribution.

2

In the present contribution, we successfully solve this problem in dimension two. We prove that if a distribution and its reflection across an axis produce identical projections onto two independent, randomly chosen directions, then the identified axis is the true axis of symmetry (almost surely) and does not depend on the particular directions sampled.

The generalization of this result to higher dimensions remains an open problem; we conjecture that agreement on d independent random projections suffices in $\mathbb{R}^d$, but a proof is elusive.

Based on these results, we develop a consistent estimator for the directions of symmetry. To the best of our knowledge, this constitutes the first proposed method in the literature for this task. The estimator is well-defined for arbitrary dimension d, and is proved to be consistent in dimension $d = 2$. The proof of its consistency for $d > 2$ is analogous to the two-dimensional case, requiring only that the aforementioned higher-dimensional conjecture be true.

2. Setup and motivation

Let X be a random vector in $\mathbb{R}^d$ with the usual Euclidean inner product, and let $\mathbb{S}^{d-1} = \{u \in \mathbb{R}^d : \|u\| = 1\}$ denote the unit sphere. We denote by ν the uniform distribution on $\mathbb{S}^{d-1}$. Throughout this paper, we always assume that expected value $\mathbb{E}(X)$ of X exists, and is finite.

Definition 1. *A random vector $X \in \mathbb{R}^d$ is axially symmetric with respect to $u \in \mathbb{S}^{d-1}$ if $X - \mathbb{E}(X) \overset{d}{=} R_u(X - \mathbb{E}(X))$, where $R_u = 2uu^\top - \mathrm{I}_d$ is the orthogonal matrix representing reflection across the hyperplane orthogonal to u, $\overset{d}{=}$ denotes equality in distribution, $u^\top$ is the transpose of u, and I_d is the $d \times d$ identity matrix.*

For illustration, let us fix u to be the first canonical basis vector $u = (1, 0, \ldots, 0)^\top$. Then,

$$R_u = \begin{pmatrix} 1 & 0_{d-1}^\top \\ 0_{d-1} & -\mathrm{I}_{d-1} \end{pmatrix},$$

where 0_{d-1} is the zero vector in $\mathbb{R}^{d-1}$.

For notational simplicity, assume $\mathbb{E}(X) = 0_d$. We define the set

$$\mathcal{U} = \{u \in \mathbb{S}^{d-1} : X \overset{d}{=} R_u X\}, \tag{1}$$

3

that is the set of axes of symmetry of X.

For a univariate random variable Y, write F_Y for its distribution function. Given a direction $h \in \mathbb{S}^{d-1}$, we consider the projections $\langle X, h \rangle$ and $\langle R_u X, h \rangle$, and define the function

$$g_h(u) = \sup_{t \in \mathbb{R}} \left| F_{\langle X, h \rangle}(t) - F_{\langle R_u X, h \rangle}(t) \right|,$$

and the set

$$\mathcal{U}^h = \{ u \in \mathbb{S}^{d-1} : h^\top X \overset{d}{=} h^\top R_u X \}.$$

It is easy to see that it could also be defined by $\mathcal{U}^h = \{ u \in \mathbb{S}^{d-1} : g_h(u) = 0 \}$.

For a set of directions $\mathfrak{H} = \{ h_1, \ldots, h_k \} \subset \mathbb{S}^{d-1}$ (where k is a positive integer) we define

$$g_{\mathfrak{H}}(u) := \frac{1}{k} \sum_{j=1}^{k} g_{h_j}(u), \tag{2}$$

and

$$\mathcal{U}^{\mathfrak{H}} = \{ u \in \mathbb{S}^{d-1} : g_{\mathfrak{H}}(u) = 0 \}. \tag{3}$$

Assuming that X is symmetric with respect to u and that it fulfills a general condition on its moments, we propose an estimator of u based on a Cramér-Wold device that allows us to estimate u by the set $\mathcal{U}^{\mathfrak{H}}$, see for example [5, 6, 11] for references on the Cramér-Wold theorem and generalizations.

The set $\mathcal{U}^{\mathfrak{H}}$ from (3) may contain directions that are not axes of symmetry of X. For instance, for any $h \in \mathbb{S}^{d-1}$, it is trivial that $g_h(h) = 0$. Additionally, if the distribution of X is centrally symmetric (but not axially symmetric around any axis), then $g_h(v) = 0$ for every h orthogonal to v. This observation indicates that identifying axes of symmetry from $\mathcal{U}^{\mathfrak{H}}$ is not straightforward, and critical information may be lost when projecting onto random directions. The main theorem of this paper shows that if the set $\mathfrak{H}$ is properly defined, then $\mathcal{U}^{\mathfrak{H}}$ almost surely coincides with $\mathcal{U}$ and does not depend on $\mathfrak{H}$.

3. Theoretical results

3.1. On symmetry under Carleman's condition

In what follows, we need a moment–determinacy guarantee for one-dimensional projections. A classical sufficient condition is Carleman's con-

4

dition (see [3]), which ensures that a distribution is uniquely determined by its moments. The Carleman condition is sufficient but not necessary: there are well–known laws with all moments finite that are nonetheless moment–indeterminate (e.g., lognormal–type families and relatives). In our setting, we assume a Carleman condition on $\|X\|$; by Lyapunov's inequality (see [2]), $\left(\mathbb{E}|h^\top X|^{2k}\right)^{1/(2k)} \leq \left(\mathbb{E}\|X\|^{2k}\right)^{1/(2k)}$, the even–moment Carleman series for each projection $h^\top X$ is dominated by that of $\|X\|$, hence each projection is moment–determinate. Combined with the Cramér–Wold device [5], this lets us pass from many identical one–dimensional projections to strong symmetry conclusions.

Definition 2. *Let $X \in \mathbb{R}^d$ be a random vector with moments of all orders*

$$m_k = \mathbb{E}\left[\,\|X\|^k\,\right], \qquad k \in \mathbb{N}.$$

We say that X satisfies the Carleman condition if

$$\sum_{k=1}^{\infty} m_k^{-\frac{1}{k}} = \infty.$$

Recall that a random vector $X \in \mathbb{R}^d$ is called spherically symmetric if for each orthogonal matrix $A \in \mathbb{R}^{d \times d}$ we have $AX \stackrel{d}{=} X$. Then we also say that the distribution P of X is spherical (or rotation-invariant).

Proposition 1. *Let $X \in \mathbb{R}^2$ be a random vector that fulfills the Carleman condition. Assume there exists a real random variable T such that for infinitely many unit directions $h \in \mathbb{S}^1$ one has $h^\top X \stackrel{d}{=} T$. Then X is spherically symmetric. In particular, $h^\top X \stackrel{d}{=} T$ for every $h \in \mathbb{S}^1$, and T is symmetric about 0_2.*

Proof. For each $k \in \mathbb{N}$ define $P_k(h) = \mathbb{E}\left[(h^\top X)^k\right]$, $\|h\| = 1$. Clearly P_k is a homogeneous polynomial of degree k in the coordinates of h; when restricted to $\mathbb{S}^1$ it is a trigonometric polynomial (hence real-analytic in the angle). By hypothesis, there is an infinite set $H \subset \mathbb{S}^1$ such that $P_k(h) = m_k := \mathbb{E}[T^k]$ for every $h \in H$. Because a non-constant trigonometric polynomial can take a fixed value only at finitely many angles, it follows that $P_k(h) \equiv m_k$ for all $h \in \mathbb{S}^1$.

Moreover, since $P_k(h)$ is constant and equal to m_k for every h and every k, the one-dimensional moments of $h^\top X$ coincide with those of T for all h.

Carleman's condition implies moment determinacy in one dimension: For any unit vector h we have

$$m_{2k}(h) := \mathbb{E}\big[\,|h^\top X|^{2k}\,\big] \ \leq\ \mathbb{E}\big[\|X\|^{2k}\big] =: M_{2k} \quad \Longrightarrow \quad m_{2k}(h)^{-\frac{1}{2k}} \geq M_{2k}^{-\frac{1}{2k}}.$$

That is, each projection $h^\top X$ satisfies Carleman's condition, so its moment sequence determines its law. We have proved that

$$h^\top X \stackrel{d}{=} T \quad \text{for every } h \in \mathbb{S}^1.$$

Now let Q be a rotation (that is, a 2×2 orthogonal matrix) in $\mathbb{R}^2$. For every $h \in \mathbb{S}^1$,

$$h^\top (QX) \ =\ (Q^\top h)^\top X \stackrel{d}{=} T.$$

Thus, every one-dimensional projection of QX has the same distribution as that of X. By the Cramér–Wold theorem we conclude $QX \stackrel{d}{=} X$ for every rotation Q; hence the law of X is rotation-invariant, i.e. X is spherically symmetric. $\qquad\square$

It is interesting to observe that without the Carleman condition, the claim of Proposition 1 is no longer true. For an example, see [20, Example 24].

Corollary 1. *Let P be a distribution on $\mathbb{R}^2$ that fulfills the Carleman condition. Assume that there are infinitely many directions of symmetry in the set $\mathcal{U}$ from (1), then the distribution of P is spherical.*

Proof. Fixing a direction $v \in \mathbb{S}^1$, by the symmetry hypothesis, there are infinitely many directions in which the projected distributions are the same, so Corollary 1 is a straightforward application of Proposition 1. $\qquad\square$

Remark 1 (Counterexample to Proposition 1 in $\mathbb{R}^3$). *Consider X to be a Gaussian vector with mean vector $(0,0,0)^\top$ and a diagonal covariance matrix with diagonal entries $(1 + \epsilon, 1, 1)$ for some $\epsilon > 0$. It is straightforward to see that every direction in the plane defined by $x_1 = 0$ (i.e., the plane orthogonal to the x_1-axis) is a direction of symmetry, despite the fact that the distribution of X is not spherical.*

Finite symmetry groups (e.g., dihedral symmetries in the plane) constrain the law along finitely many axes but do not imply spherical symmetry. In contrast, a positive surface measure of $\mathcal{U}$ provides a density point x_0 and

therefore, by Lemma 1 below, a mechanism to propagate symmetry locally along short geodesic arcs. This "percolation" converts local abundance into global invariance. The proof of the following proposition formalizes this propagation. Recall that by ν we denote the uniform distribution on $\mathbb{S}^{d-1}$.

Proposition 2. *Let $X \in \mathbb{R}^d$ be a random vector. If the set of directions of symmetry $\mathcal{U}$ of X has positive ν-measure, then the distribution of X is spherically symmetric.*

The proof of this proposition is based on the following Lemmas 1 and 2. The first one states that the set $\mathcal{U}$ of axes of symmetry is closed under conjugation.

Lemma 1. *Let $u_1, u_2 \in \mathbb{S}^{d-1}$. If $X \in \mathbb{R}^d$ is a random vector symmetric about the directions generated by u_1 and u_2, then it is also symmetric about the direction $v = R_{u_1} u_2$.*

Proof of Lemma 1. We know that $X \overset{d}{=} R_{u_1} X$ and $X \overset{d}{=} R_{u_2} X$. Our goal is to prove that $X \overset{d}{=} R_v X$, where $v = R_{u_1} u_2$. Thanks to the orthogonality of the matrix R_{u_1} we have

$$R_v = 2(R_{u_1} u_2)(R_{u_1} u_2)^\top - \mathrm{I}_d = R_{u_1}(2u_2 u_2^\top - \mathrm{I}_d)R_{u_1} = R_{u_1} R_{u_2} R_{u_1}.$$

From $X \overset{d}{=} R_{u_2} X$ and $X \overset{d}{=} R_{u_1} X$, it follows that, $R_{u_1} X \overset{d}{=} R_{u_1} R_{u_2} R_{u_1} X$. Therefore, $X \overset{d}{=} R_v X$.

$\square$

Lemma 2. *The set of directions of symmetry $\mathcal{U}$ of a random vector $X \in \mathbb{R}^d$ is closed.*

Proof. Let $\{u_\ell\}_{\ell \in \mathbb{N}}$ be a sequence in $\mathbb{S}^{d-1}$ converging to u_0 such that $X \overset{d}{=} R_{u_\ell} X$ for all ℓ. The result follows from the convergence $R_{u_\ell} \to R_{u_0}$ as $\ell \to \infty$, and the continuous mapping theorem [1, Theorem 5.5]. $\square$

Proof of Proposition 2. By Lemma 2, the set of directions of symmetry $\mathcal{U}$ is closed. So, it suffices to show that if $\mathcal{U}$ has positive measure, then it must be dense in $\mathbb{S}^{d-1}$.

By the Lebesgue density theorem [13, Corollary 261D], if $\mathcal{U}$ has positive measure, then there must exist a point $x_0 \in \mathcal{U}$ with metric density 1, i.e.,

$$\lim_{r \downarrow 0} \frac{\nu(\mathcal{U} \cap B(x_0, r))}{\nu(B(x_0, r))} = 1,$$

where $B(x_0, r) \subset \mathbb{S}^{d-1}$ denotes the open geodesic ball of radius r, centered at x_0. This implies that every neighborhood of x_0 contains points of $\mathcal{U}$.

Now, suppose for contradiction that $\mathcal{U}$ is not dense in $\mathbb{S}^{d-1}$. Then there exists an open ball $B(y, \epsilon) \subset \mathbb{S}^{d-1}$ that contains no direction of symmetry.

Consider the set of all geodesic arcs from x_0 to points in $B(y, \epsilon/4)$. Since x_0 is a point of metric density 1, the ball $B(x_0, \epsilon/4)$ must contain a point $x_1 \in \mathcal{U}$ that lies on one of these geodesic arcs. Now, reflect x_0 with respect to x_1 to get $x_2 = R_{x_1}(x_0)$. Since $x_0, x_1 \in \mathcal{U}$, by Lemma 1, also $x_2 \in \mathcal{U}$.

Iterating this process, $x_{\ell+1} = R_{x_\ell}(x_{\ell-1})$, we generate a sequence $\{x_\ell\}_{\ell \in \mathbb{N}}$ of symmetry directions that must intersect $B(y, \epsilon)$ (because for the geodesic distance d_g on $\mathbb{S}^{d-1}$ we have $d_g(x_{\ell+1}, x_\ell) \leq \epsilon/4$), which contradicts the assumption that $B(y, \epsilon)$ contains no symmetry directions. Therefore, $\mathcal{U}$ must be dense in $\mathbb{S}^{d-1}$, and since it is closed, we have $\mathcal{U} = \mathbb{S}^{d-1}$, proving that the distribution is spherically symmetric. $\qquad \square$

Our main Theorem 1 (proven for $d = 2$, although we conjecture that an analogous theorem with a random sample $\{h_j\}_{j=1}^d$ holds for any $d > 2$) relies on two ingredients that isolate the true axes in the non-spherical case: For fixed $h \in \mathbb{S}^1$, the set $H(h) := \{g \in \mathbb{S}^1 : g^\top X \overset{d}{=} h^\top X\}$ is finite; otherwise, Proposition 1 would force spherical symmetry. Hence, a single direction leaves only finitely many "ambiguous" candidates for an axis. For each false axis $u \notin \mathcal{U}$, the set $M_u := \{h \in \mathbb{S}^1 : h^\top X \overset{d}{=} h^\top R_u X\}$ has ν-measure zero. Therefore, the Cartesian products $M_u \times M_u$ are ν^2-null, and Fubini's theorem shows that almost every random pair (h_1, h_2) excludes all false axes simultaneously, leaving exactly $\mathcal{U}$.

Theorem 1. *Let $X \in \mathbb{R}^2$ be a random vector that fulfills the Carleman condition. Let $\{h_j\}_{j=1}^2 \subset \mathbb{S}^1$ be a random sample of two independent directions with distribution ν. Denote by $\mathcal{U} \subset \mathbb{S}^1$ the (possibly empty) set of axes of symmetry of X. Suppose that X is not spherically symmetric, then the set of solutions to the system $h_j^\top X \overset{d}{=} h_j^\top R_u X$, $j = 1, 2$ corresponds, with probability one, to the set of true symmetry axes $\mathcal{U}$ of X.*

Proof. For $u \in \mathbb{S}^1$, denote by $M_u \subseteq \mathbb{S}^1$ the set of those $h \in \mathbb{S}^1$ that fulfill $h^\top R_u X \overset{d}{=} h^\top X$. Take $h \in \mathbb{S}^1$ fixed, and consider the distribution of $h^\top X$ in $\mathbb{R}$. Because X is not spherically symmetric, Proposition 1 gives that the set of directions $H(h) = \left\{g \in \mathbb{S}^1 : g^\top X \overset{d}{=} h^\top X\right\}$ must be finite.

We see that $h \in M_u$ only if u can be written as $(h_1 + h_2)/\|h_1 + h_2\|$ for some $h_1, h_2 \in H(h)$, $h_1 \neq -h_2$. Because there are only finitely many directions that can be written like this, we see that $h \in M_u$ only for finitely many directions $u \in \mathbb{S}^1$.

At the same time, for each $u \in \mathbb{S}^1$ such that $u \notin \mathcal{U}$, we know that M_u is of null ν-measure in $\mathbb{S}^1$ by Cuesta-Albertos et al. [6, Corollary 3.2].

Overall, we have that each $h \in \mathbb{S}^1$ is contained in finitely many sets M_u, and at the same time each M_u is of null ν-measure. Consider now the union

$$M = \bigcup_{u \notin \mathcal{U}} M_u \times M_u$$

of all the Cartesian products $M_u \times M_u \subset \mathbb{S}^1 \times \mathbb{S}^1$ over all directions $u \in \mathbb{S}^1$ that are not axes of symmetry of X. By the arguments above, for each $h \in \mathbb{S}^1$, there are only finitely many $u \notin \mathcal{U}$ such that $h \in M_u$, and for each of these finitely many u, the set M_u is of null ν-mass. Consequently, the section

$$M(h) = \{(h_1, h_2) \in M : h_1 = h\} \tag{4}$$

of M with first coordinate h is a finite union of ν-null sets in $\mathbb{S}^1$, hence it is ν-null. The mass of M with respect to the product measure ν^2 can thus be evaluated as

$$\nu^2(M) = \int_{\mathbb{S}^1} \int_{M(h)} 1 \, \mathrm{d}\,\nu(h_2) \, \mathrm{d}\,\nu(h) = \int_{\mathbb{S}^1} 0 \, \mathrm{d}\,\nu(h) = 0,$$

by Fubini's theorem [8, Theorem 4.4.5]. We have shown that for ν^2-almost every random choice of $(h_1, h_2) \in \mathbb{S}^1 \times \mathbb{S}^1$, every $u \in \mathbb{S}^1$ such that $h_i^\top R_u X \overset{d}{=} h_i^\top X$ for both $i = 1, 2$ must satisfy $u \in \mathcal{U}$, as we wanted to show. $\qquad \square$

A straightforward extension of the proof of Theorem 1 to dimension $d > 2$ is not possible, since in that case, for each h fixed, the set $M(h)$ from (4) is a union of ν-null sets that is indexed by a ν-null set (in particular, the index set may be uncountable). Such a union, however, does not need to be ν-null.

3.2. A counter-example when Carleman condition is not assumed

Let $K \subset \mathbb{R}^2$ be a symmetric convex body (a compact convex set with non-empty interior) centered at the origin. The K-norm of $x \in \mathbb{R}^2$ is defined as $\|x\|_K = \inf \{\lambda \geq 0 : x \in \lambda K\}$. The K-norm is a norm on $\mathbb{R}^2$ [22,

Section 1.7.2]. A random vector X in $\mathbb{R}^2$ is said to be K-symmetric (sometimes called pseudo-isotropic, see [9, Chapter 7], [15, Chapter 6] and [20, Chapter 6] for many references) if its characteristic function takes the form

$$\psi_X(t) = \mathbb{E}\exp(\mathrm{i}\langle t, X\rangle) = \xi(\|t\|_K) \quad \text{for all } t \in \mathbb{R}^2, \tag{5}$$

where ξ is a continuous, real-valued function.

First, fix $u \in \mathbb{S}^1$. The random vector $R_u X$ has a characteristic function

$$\psi_{R_u X}(t) = \mathbb{E}\exp(\mathrm{i}\langle t, R_u X\rangle) = \mathbb{E}\exp(\mathrm{i}\langle R_u t, X\rangle) = \psi_X(R_u t) = \xi(\|R_u t\|_K) \tag{6}$$

for $t \in \mathbb{R}^2$. Comparing (5) and (6) we see that $X \overset{d}{=} R_u X$ if and only if

$$\|t\|_K = \|R_u t\|_K \quad \text{for all } t \in \mathbb{R}^2, \tag{7}$$

which is true if and only if the body K is axially symmetric around the axis given by u.

Now, let $u \in \mathbb{S}^1$ and $h \in \mathbb{S}^1$ be given. The characteristic function of the random variable $\langle h, X\rangle$ is for $t \in \mathbb{R}$ given by

$$\psi_{\langle h, X\rangle}(t) = \mathbb{E}\exp(\mathrm{i}t\langle h, X\rangle) = \mathbb{E}\exp(\mathrm{i}\langle t\,h, X\rangle) = \xi(\|t\,h\|_K) = \xi(|t|\,\|h\|_K).$$

The characteristic function of the random variable $\langle h, R_u X\rangle$ is similarly

$$\psi_{\langle h, R_u X\rangle}(t) = \mathbb{E}\exp(\mathrm{i}t\langle h, R_u X\rangle) = \mathbb{E}\exp(\mathrm{i}\langle t\,R_u h, X\rangle) = \xi(\|t\,R_u h\|_K)$$
$$= \xi(|t|\,\|R_u h\|_K).$$

Because ξ is continuous and $\xi(0_d) = \mathbb{E}\exp(\mathrm{i}\langle 0_d, X\rangle) = 1$, we obtain that $\langle h, X\rangle \overset{d}{=} \langle h, R_u X\rangle$ if and only if

$$\|h\|_K = \|R_u h\|_K. \tag{8}$$

Consider now the special convex body K displayed in Figure 1 (the beige body with black boundary line). The body K is symmetric around the origin (that is, $K = -K$), but is clearly not axially symmetric around any $u \in \mathbb{S}^1$. By [10, Theorem 1], see also [20, Theorem 12], there exists a K-symmetric random vector X with this particular choice of K. The distribution of X follows a multivariate extension of the Cauchy distribution. Because K is not axially symmetric around any $u \in \mathbb{S}^1$, formula (7) gives that X is not
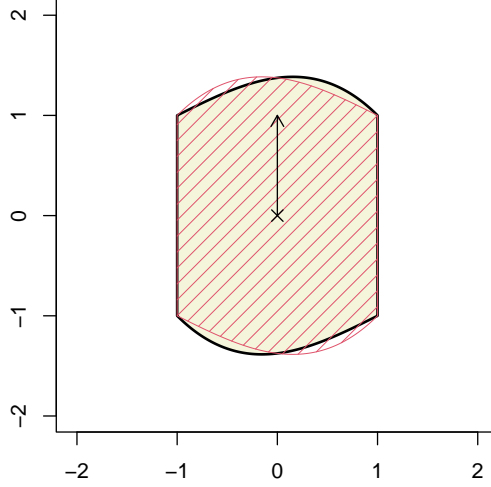
Figure 1: The convex body K (beige body with black boundary line) and its reflection $R_u K$ around the vertical axis given by $u = (0,1)^\top \in \mathbb{S}^1$ (the shaded body in red). The bodies K and $R_u K$ are not identical, meaning that K is not axially symmetric around u (or any other vector in $\mathbb{S}^1$).

axially symmetric around any $u \in \mathbb{S}^1$.

On the other hand, in Figure 1 we see that for all $x \in \mathbb{R}^2$ in the line segments on the boundary of K (that is, for $x \in K$ whose first coordinate is either 1 or -1) we have for $u = (0,1)^\top$

$$\|R_u\, x\|_K = \|x\|_K = 1.$$

This means that (8) is satisfied for all such x, and the same is true for all $h = x/\|x\| \in \mathbb{S}^1$. We see that even though there is no axis of symmetry of X, if we sample $\mathfrak{H} = \{h_1, h_2\}$ independently from the uniform distribution on $\mathbb{S}^1$, then with positive probability we get a non-empty set $\mathcal{U}^{\mathfrak{H}}$ containing $u = (0,1)^\top$.

4. A plug-in based estimator

Suppose that we observe a random sample $\aleph_n = \{X_1, \ldots, X_n\}$ from a distribution P in $\mathbb{R}^d$. Our aim is to estimate the axes of symmetry of P based on the random sample $\aleph_n$. To define the first estimator that we propose, we divide the sample $\aleph_n$, randomly into two balanced groups. After reindexing, we obtain $\aleph_1 = \{X_1, \ldots, X_{\lfloor n/2 \rfloor}\}$ and $\aleph_2 = \{X_{\lfloor n/2 \rfloor + 1}, \ldots, X_n\}$. For $k \in$

11

$\mathbb{N}, k \geq 1$, consider a random set $\mathfrak{H} = \{h_1, \ldots, h_k\}$, sampled from the uniform distribution ν^k over $(\mathbb{S}^{d-1})^k$, we consider the function $g_{\mathfrak{H}}$ defined by (2) and the empirical versions of the involved quantities:

$$F^n_{\langle X,h \rangle}(t) = \frac{1}{|\aleph_1|} \sum_{X_i \in \aleph_1} \mathbf{1}\{\langle X_i, h \rangle \leq t\},$$

$$F^n_{\langle R_u X,h \rangle}(t) = \frac{1}{|\aleph_2|} \sum_{X_i \in \aleph_2} \mathbf{1}\{\langle R_u X_i, h \rangle \leq t\},$$

$$\widehat{g}_{n,h}(u) = \sup_{t \in \mathbb{R}} \left| F^n_{\langle X,h \rangle}(t) - F^n_{\langle R_u X,h \rangle}(t) \right|,$$

$$\widehat{g}_{n,\mathfrak{H}}(u) = \frac{1}{k} \sum_{j=1}^{k} \widehat{g}_{n,h_j}(u),$$

where $|A|$ stands for the cardinality of a finite set A, and $\mathbf{1}\{A\}$ is the indicator of A giving 1 if A is true and 0 otherwise. To simplify the notation, we will omit the $\mathfrak{H}$ in $\widehat{g}_{n,\mathfrak{H}}$ and $g_{\mathfrak{H}}$.

As a naive approach, one might be tempted to estimate $\mathcal{U}$ by finding the minimum of $\widehat{g}_n$ over $\mathbb{S}^{d-1}$, but this is not a good solution. To see this, note that it is clear that if g is convex and if $\widehat{g}_n \to g$ uniformly as $n \to \infty$, then $\lim_{n \to \infty} \arg \min \widehat{g}_n = \arg \min g$, a.s. However, there is no guarantee for g not to have several minima, and in this case, the most powerful statement in this setup can be derived from Rockafellar and Wets [21, Theorem 7.33]: if $\widehat{g}_n$ converges uniformly to g, then $\limsup_{n \to \infty} \arg \min \widehat{g}_n \subset \arg \min g$, a.s. In particular, there is no guarantee of being able to recover all the axes of symmetry of P using $\widehat{g}_n$.

For this reason, we use a different approach. We consider an ϵ_n-level set of $\widehat{g}_n$, defined as the set

$$\mathcal{U}_n(\epsilon_n) := \left\{ u \in \mathbb{S}^{d-1} : \widehat{g}_n(u) < \epsilon_n \right\}, \tag{9}$$

where $\epsilon_n = 1/\beta_n \to 0$ as $n \to \infty$, with

$$\beta_n = \frac{\sqrt{n}}{\log n}. \tag{10}$$

We will demonstrate that $\mathcal{U}_n(\epsilon_n)$ converges to $\mathcal{U}$ in the Hausdorff distance (see Theorem 2 below). A common strategy for showing that empirical level

sets converge to their population counterparts is to establish the uniform convergence of the sequence of functions to its population version (see [17, Theorem 2.1], [4] and references therein). This result is formalized in the following proposition.

Proposition 3. *Given $\widehat{g}_n$ and g as introduced above,*

$$\sup_{v \in \mathbb{S}^{d-1}} \frac{\sqrt{n}}{a_n} |\widehat{g}_n(v) - g(v)| \to 0 \ as \ n \to \infty, \ almost \ surely,$$

where $a_n = \sqrt{\delta_n \log n}$, with $\delta_n \to \infty$. In this result, the almost sure convergence is understood both with respect to the random sample $\aleph_n$, and the random directions $\mathfrak{H}$.

Proof. Let us introduce some additional notation. For $v \in \mathbb{S}^{d-1}$ and fixed $h \in \mathbb{S}^{d-1}$, define $\mathfrak{B}_{v,t,h} = \{x \in \mathbb{R}^d : \langle R_v x, h \rangle \leq t\}$. Then, for fixed h_0,

$$\mathbb{P}\{\langle R_v X, h \rangle \leq t \,\big|\, h = h_0\} = P(\mathfrak{B}_{v,t,h_0}).$$

Let $P_n(A) = (1/n) \sum_{i=1}^n \mathbf{1}\{X_i \in A\}$ be the empirical measure based on the sample $\aleph_n$. Let us prove first that there exists $0 < C \leq 4e^8$ such that for all $n > 2(d+1)$,

$$\mathbb{P}\left\{\sup_{v \in \mathbb{S}^{d-1}} |\widehat{g}_n(v) - g(v)| > \epsilon\right\} \leq 2C(n^{2d+2} + 1)e^{-n\epsilon^2/2}. \tag{11}$$

We will first condition on fixed $\mathfrak{H}$, so for all $h \in \mathbb{S}^{d-1}$ and $v \in \mathbb{S}^{d-1}$, we have

$$
\begin{aligned}
|\widehat{g}_{n,h}(v) - g_h(v)| &= \left| \sup_{t \in \mathbb{R}} \left| F^n_{\langle X,h \rangle}(t) - F^n_{\langle R_v X, h \rangle}(t) \right| - \sup_{t \in \mathbb{R}} \left| F_{\langle X,h \rangle}(t) - F_{\langle R_v X, h \rangle}(t) \right| \right| \\
&\leq \sup_{t \in \mathbb{R}} \left| \left(F^n_{\langle X,h \rangle}(t) - F_{\langle X,h \rangle}(t) \right) - \left(F^n_{\langle R_v X, h \rangle}(t) - F_{\langle R_v X, h \rangle}(t) \right) \right| \\
&\leq \sup_{t \in \mathbb{R}} \left| F^n_{\langle X,h \rangle}(t) - F_{\langle X,h \rangle}(t) \right| + \sup_{t \in \mathbb{R}} \left| F^n_{\langle R_v X, h \rangle}(t) - F_{\langle R_v X, h \rangle}(t) \right| \\
&= \sup_{t \in \mathbb{R}} \left| F^n_{\langle X,h \rangle}(t) - F_{\langle X,h \rangle}(t) \right| + \sup_{t \in \mathbb{R}} \left| P_n(\mathfrak{B}_{v,t,h}) - P(\mathfrak{B}_{v,t,h}) \right|.
\end{aligned}
\tag{12}
$$

Let $\mathfrak{H} = \{h_1, \ldots, h_k\} \subset \mathbb{S}^{d-1}$ be fixed, then

$$
\begin{aligned}
\sup_{v \in \mathbb{S}^{d-1}} |\widehat{g}_n(v) - g(v)| &= \sup_{v \in \mathbb{S}^{d-1}} \left| \frac{1}{k} \sum_{j=1}^{k} \widehat{g}_{n,h_j}(v) - g_{h_j}(v) \right| \\
&\leq \sup_{v \in \mathbb{S}^{d-1}} \frac{1}{k} \sum_{j=1}^{k} |\widehat{g}_{n,h_j}(v) - g_{h_j}(v)| \\
&\leq \sup_{v,h \in \mathbb{S}^{d-1}} |\widehat{g}_{n,h}(v) - g_h(v)|.
\end{aligned}
\tag{13}
$$

Therefore,

$$
\mathbb{P} \left\{ \sup_{v \in \mathbb{S}^{d-1}} |\widehat{g}_n(v) - g(v)| > \epsilon \right\} = \mathbb{E}_{\mathfrak{H}} \left[\mathbb{P} \left(\sup_{v \in \mathbb{S}^{d-1}} |\widehat{g}_n(v) - g(v)| > \epsilon | \mathfrak{H} \right) \right],
$$

and

$$
\mathbb{P} \left\{ \sup_{v \in \mathbb{S}^{d-1}} |\widehat{g}_n(v) - g(v)| > \epsilon \Big| \mathfrak{H} \right\} \leq \mathbb{P} \left\{ \sup_{h \in \mathbb{S}^{d-1}, t \in \mathbb{R}} \left| F_{\langle X,h \rangle}^n(t) - F_{\langle X,h \rangle}(t) \right| > \frac{\epsilon}{2} \right\}
$$
$$
+ \mathbb{P} \left\{ \sup_{v,h \in \mathbb{S}^{d-1}, t \in \mathbb{R}} |P_n(\mathfrak{B}_{v,t,h}) - P(\mathfrak{B}_{v,t,h})| > \frac{\epsilon}{2} \right\}.
$$

Let $\mathcal{B} = \{\mathfrak{B}_{v,t,h} : v, h \in \mathbb{S}^{d-1}, t \in \mathbb{R}\}$. Each set $\mathfrak{B}_{v,t,h}$ can be expressed as $\{x \in \mathbb{R}^d : f_{v,t,h}(x) \geq 0\}$, where

$$
f_{v,t,h}(x) = t - \langle R_v x, h \rangle = t - (R_v^\top h)^\top x = t - w^\top x,
$$

with $w = R_v^\top h \in \mathbb{S}^{d-1}$. Since R_v is an orthogonal matrix and h varies over $\mathbb{S}^{d-1}$, w also varies over $\mathbb{S}^{d-1}$.

We see that the class of functions $\mathcal{F} = \{f_{v,t,h} : \mathbb{R}^d \to \mathbb{R} : v, h \in \mathbb{S}^{d-1}, t \in \mathbb{R}\}$ consists of linear functions of the form $f_{w,t}(x) = t - w^\top x$, where $w \in \mathbb{S}^{d-1}$ and $t \in \mathbb{R}$. It follows that $\mathcal{B}$ corresponds to the set of half-spaces in $\mathbb{R}^d$, whose Vapnik Chervonenkis (VC) dimension is at most $d + 1$. By Theorems 12.8 and 13.3 in Devroye et al. [7], for all $n > 2(d + 1)$, we have

$$
\mathbb{P} \left\{ \sup_{v,h \in \mathbb{S}^{d-1}, t \in \mathbb{R}} |P_n(\mathfrak{B}_{v,t,h}) - P(\mathfrak{B}_{v,t,h})| > \epsilon \right\} \leq C(n^{2(d+1)} + 1)e^{-2n\epsilon^2},
$$

14

where $C \leq 4e^8$. Similarly,

$$\mathbb{P}\left\{\sup_{h \in \mathbb{S}^{d-1}, t \in \mathbb{R}} \left|F^n_{\langle X,h \rangle}(t) - F_{\langle X,h \rangle}(t)\right| > \epsilon\right\} \leq C(n^{2(d+1)} + 1)e^{-2n\epsilon^2}.$$

Then, we have that

$$\begin{aligned}
\mathbb{P}\left\{\sup_{v \in \mathbb{S}^{d-1}} \frac{\sqrt{n}}{a_n} |\widehat{g}_n(v) - g(v)| > \epsilon\right\} &\leq 2C(n^{2d+2} + 1)e^{-n\epsilon^2 a_n^2/(2n)} \\
&\leq 2C(n^{2d+2} + 1)n^{-\epsilon^2\delta_n/2} \\
&= 2C(n^{2d+2-\epsilon^2\delta_n/2} + n^{-\epsilon^2\delta_n/2}).
\end{aligned}$$

The assumption $\delta_n \to \infty$ implies that, for every $\epsilon > 0$, there exists $N(= N(\epsilon))$ such that if $n \geq N$, then

$$2C(n^{2d+2-\epsilon^2\delta_n/2} + n^{-\epsilon^2\delta_n/2}) < n^{-2}.$$

Combining these inequalities, we obtain the desired bound (11). Lastly, the result is a direct application of Borel-Cantelli's lemma [8, Theorem 8.3.4]. $\square$

Lemma 3. *Assume that P is an absolutely continuous distribution in $\mathbb{R}^d$. Then for all $\mathfrak{H}$, the function g given by (2) is continuous.*

Proof. Consider $\mathfrak{H}$ fixed. Similarly as in (12) and (13), for any $u_\ell \to u$ in $\mathbb{S}^{d-1}$ as $\ell \to \infty$ we have that

$$|g(u_\ell) - g(u)| \leq \max_{j=1,\dots,k} \sup_{t \in \mathbb{R}} \left|F_{\langle R_{u_\ell}X, h_j \rangle}(t) - F_{\langle R_u X, h_j \rangle}(t)\right|. \tag{14}$$

Because P is absolutely continuous, for each $\mathfrak{H}$ the distribution function $F_{\langle R_u X, h_j \rangle}$ is continuous as a function of u. At the same time, since $u_\ell \to u$ in $\mathbb{S}^{d-1}$, we have that $\langle R_{u_\ell}X, h_j \rangle$ converges in distribution to $\langle R_u X, h_j \rangle$ for each $j = 1, \dots, k$, meaning that as $\ell \to \infty$, the right-hand side of (14) vanishes. $\square$

The following theorem states that the level set $\mathcal{U}_n(\epsilon_n)$ from (9), for a properly defined sequence ϵ_n, converges almost surely to $\mathcal{U}$ in the Hausdorff

15

distance. That is, with probability one,

$$d_H(\mathcal{U}, \mathcal{U}_n(\epsilon_n)) := \max\left\{\sup_{a\in\mathcal{U}} d_E(a, \mathcal{U}_n(\epsilon_n)), \sup_{c\in\mathcal{U}_n(\epsilon_n)} d_E(c, \mathcal{U})\right\} \to 0, \text{ as } n \to \infty,$$

where $d_E(a, B) = \inf_{b\in B}\|a - b\|$ is the Euclidean distance of a and a set B.

Theorem 2. *Assume that P is an absolutely continuous distribution in $\mathbb{R}^d$. Let β_n be as in* (10), *and $\mathcal{U}_n(\epsilon_n)$ as in* (9). *Define $\epsilon_n = 1/\beta_n$; then*

$$d_H(\mathcal{U}, \mathcal{U}_n(\epsilon_n)) \to 0 \quad \text{almost surely, as } n \to \infty.$$

The almost sure convergence in Theorem 2 is again considered with respect to both the random sample $\aleph_n$ and the set of directions $\mathfrak{H}$.

Proof. Let us assume that all the random variables involved are defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$. In this proof, given $\omega \in \Omega$, the superindex ω denotes the dependence of the related object on the selected ω.

From Proposition 3 applied with $\delta_n = \log(n)$, there exists $\Omega_0 \in \mathcal{A}$, such that, $\mathbb{P}(\Omega_0) = 1$ and, for every $\omega \in \Omega_0$, it happens that

$$\beta_n\|\widehat{g}_n^\omega - g\|_\infty \to 0, \quad \text{as } n \to \infty,$$

(notice that g and β_n have no superindex because they are not random).

Let us fix an element $\omega \in \Omega_0$. Then, there exists $n_0^\omega \in \mathbb{N}$ such that for all $n > n_0^\omega$, we have

$$0 \le \widehat{g}_n^\omega(u) < g(u) + \frac{1}{\beta_n} \quad \text{for all } u \in \mathbb{S}^{d-1}.$$

Therefore, if $u \in \mathcal{U}$, for all $n > n_0^\omega$, we have that $\widehat{g}_n^\omega(u) < 1/\beta_n$, and consequently $\mathcal{U} \subset \mathcal{U}_n^\omega(\epsilon_n)$ for all $n > n_0^\omega$.

Additionally, if $u \in \mathcal{U}_n^\omega(\epsilon_n)$, and $n \ge n_0^\omega$, then $u \in g^{-1}([0, 2/\beta_n))$ so $\mathcal{U}_n^\omega(\epsilon_n) \subset g^{-1}([0, 2/\beta_n))$. Therefore, if we prove that $d_H(\mathcal{U}, g^{-1}([0, 2/\beta_n))) \to 0$, the claim will follow; notice that there is no superindex here because none of the involved terms is random.

Proceeding by contradiction, let us assume that there exists $\kappa > 0$ and a sequence $\{u_n\}_{n=1}^\infty \subset \mathbb{S}^{d-1}$ with $u_n \in g^{-1}([0, 2/\beta_n))$ and $d_E(u_n, \mathcal{U}) > \kappa$ for every $n \in \mathbb{N}$. Then, there exists a subsequence $u_{n_\ell} \to u_0 \in \mathbb{S}^{d-1}$ as $\ell \to \infty$. Because $g(u_n) \le 2/\beta_n \to 0$ as $n \to \infty$, we obtain by Lemma 3 that

16

$g(u_0) = \lim_{\ell \to \infty} g(u_{n_\ell}) = 0$, meaning that $u_0 \in \mathcal{U}$. That contradicts the fact that $d_E(u_0, \mathcal{U}) \geq \kappa$. $\qquad\square$

5. Simulation study

To demonstrate the performance of our estimator, we conducted a simulation study in which we generated datasets with known axes of symmetry and evaluated whether our method could accurately recover them. We considered two situations in the plane, named **Scenario 1** and **Scenario 2**, and for each of them we studied the performance with varying n (the sample size) and k (the number of random directions considered).

Scenario 1. A bivariate normal distribution with unit variances and a correlation coefficient of $\rho = 0.7$. A sample drawn from this distribution of size $n = 1000$ is depicted in Figure 2a.

Scenario 2. A uniform distribution on the square $[-1, 1]^2$. A sample consisting of $n = 10000$ data points is depicted in Figure 2b.

In these figures, the true axes of symmetry are represented by solid gray lines. Additionally, a black line illustrates the empirical estimate of the function g, presented in a polar coordinate system. This choice of representation enhances the understanding of the radial relationships within the data.

For both **Scenarios 1** and **2**, we independently simulated 500 datasets and ran our algorithm on each one.

To quantify the error, we need to find the angles at which $\widehat{g}_n$ reaches its local minima. To tackle this optimization problem, we create an equally spaced grid $\{u_j\}_{j=1}^{200} \subset \mathbb{S}^1$ and approximate

$$\arg\min_{u \in \mathbb{S}^1} \widehat{g}_n(u)$$

by

$$\arg\min_{u \in \{u_1, \ldots, u_{200}\}} \widehat{g}_n(u).$$

The random directions $h_1, \ldots, h_k \in \mathbb{S}^1$ are drawn as independent random vectors with uniform distribution on $\mathbb{S}^1$.

One complication here is that $\widehat{g}_n$ is a noisy estimate of the function g, so it will have multiple spurious local minima. Identifying the number of

true (non-spurious) local minima, and their locations, leads to an interesting and challenging task. To address this issue we used the AMPD algorithm presented in [23].[1]

In Tables 1a to 1e we show the frequency of the number of local minima detected by the AMPD algorithm for **Scenario 2** for each sample size. We considered $n = 200, 500, 1000, 2000$ and 5000, and the number of random projections was set at $k = 200$ at first; later, we will take into account the choice of this parameter. For **Scenario 1**, our algorithm correctly detected the two local minima in 100% of the replications for all sample sizes greater than 200. For $n = 200$, it detected an incorrect number of local minima in only 2 out of 500 replications.

(a) Sample size $n = 200$					
1	2	3	4	5	5+
0.15	0.48	0.15	0.18	0.02	0.02

(b) Sample size $n = 500$					
1	2	3	4	5	5+
0.11	0.23	0.13	0.50	0.01	0.01

(c) Sample size $n = 1000$					
1	2	3	4	5	5+
0.03	0.14	0.03	0.79	0.00	0.00

(d) Sample size $n = 2000$					
1	2	3	4	5	5+
0.00	0.02	0.00	0.97	0.00	0.00

(e) Sample size $n = 5000$					
1	2	3	4	5	5+
0.00	0.00	0.00	1.00	0.00	0.00

Table 1: Frequency of number of local minimas detected for **Scenario 2**.

Once we have the local minima of $\widehat{g}_n$, and restricting to the cases where the estimated number of minima is correct, we matched each estimated axis to its closest true axis. We then computed the average angular distance

---

[1]In particular, our procedure is designed specifically for the two-dimensional setting. If one is interested in working in higher dimensions, a specific optimization algorithm is required, as it involves optimizing a noisy, non-convex, and discretized function defined on the sphere. This makes the estimation in dimension $d > 2$ substantially more challenging and is outside the scope of our present simulation study.
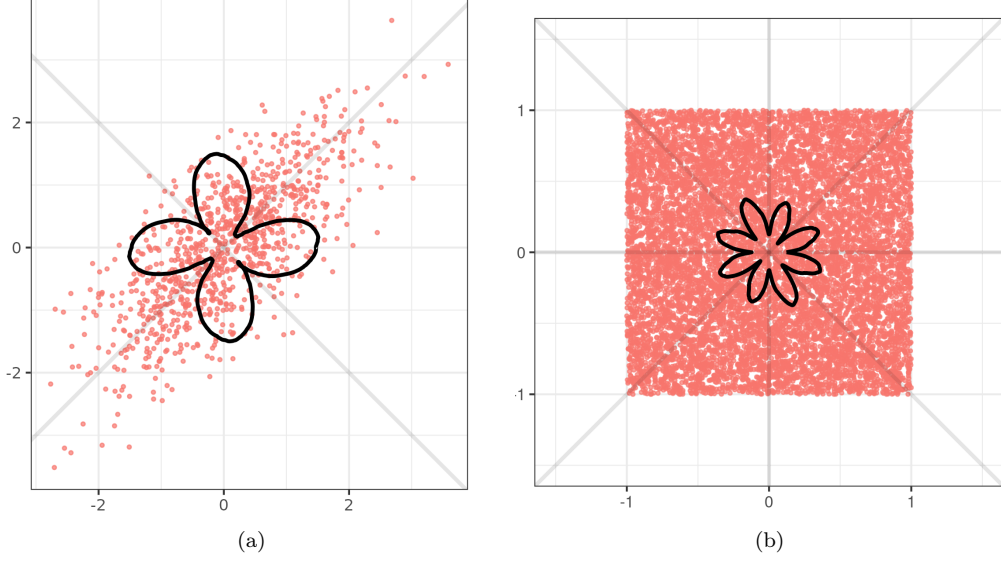
Figure 2: In each subfigure, simulated data are depicted as red points; the black line illustrates the estimate of the function g in a polar plot format, and the gray lines indicate the true axes of symmetry. Subfigure (a) represents a bivariate normal distribution (**Scenario 1**), while subfigure (b) represents a uniform distribution within a square (**Scenario 2**).

between these matched pairs. Figure 3 show these errors for **Scenarios 1** and **2**, and $k = 200$.

In a second numerical experiment, we vary the number of random projections, exploring $k = 2, 10, 50, 100, 150,$ and 200. The frequency with which the algorithm correctly identifies the number of minima is presented in Tables 2 and 3 for both scenarios. Restricting again to replications where the true number of local minima is detected, the corresponding estimation error is shown in Figure 4 and Figure 5.

From an empirical point of view, we see that our estimation method is consistent. This confirms our theoretical analysis from Section 3. Further, there appears to be little gain in choosing $k > 50$. The estimation algorithm performs well for a sufficiently large sample size, which is expected given its completely nonparametric framework that typically requires larger amounts of data. The results also show that the method works better for the Gaussian data (**Scenario 1**) than for the uniform data (**Scenario 2**).
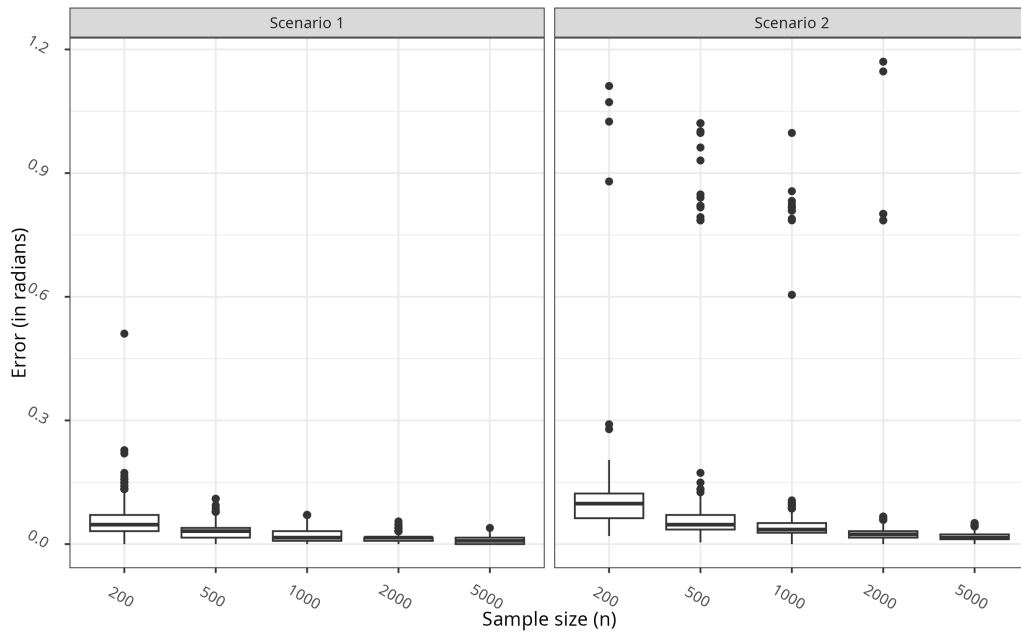
19

Figure 3: Average angular error for different sample sizes for **Scenario 1** and **2**, for $k = 200$ projections.

	$k = 2$	$k = 10$	$k = 50$	$k = 100$	$k = 150$	$k = 200$
$n = 200$	0.864	0.960	0.996	0.998	0.998	0.996
$n = 500$	0.900	0.992	1.000	1.000	1.000	1.000
$n = 1000$	0.892	0.986	1.000	1.000	1.000	1.000
$n = 2000$	0.888	0.992	1.000	1.000	1.000	1.000
$n = 5000$	0.896	0.992	1.000	1.000	1.000	1.000

Table 2: Frequency of correct number of minima detected, for **Scenario 1**.

	$k = 2$	$k = 10$	$k = 50$	$k = 100$	$k = 150$	$k = 200$
$n = 200$	0.140	0.146	0.176	0.160	0.182	0.176
$n = 500$	0.318	0.406	0.472	0.482	0.486	0.498
$n = 1000$	0.456	0.646	0.764	0.786	0.788	0.790
$n = 2000$	0.656	0.878	0.966	0.968	0.964	0.970
$n = 5000$	0.872	0.988	0.998	0.998	1.000	1.000

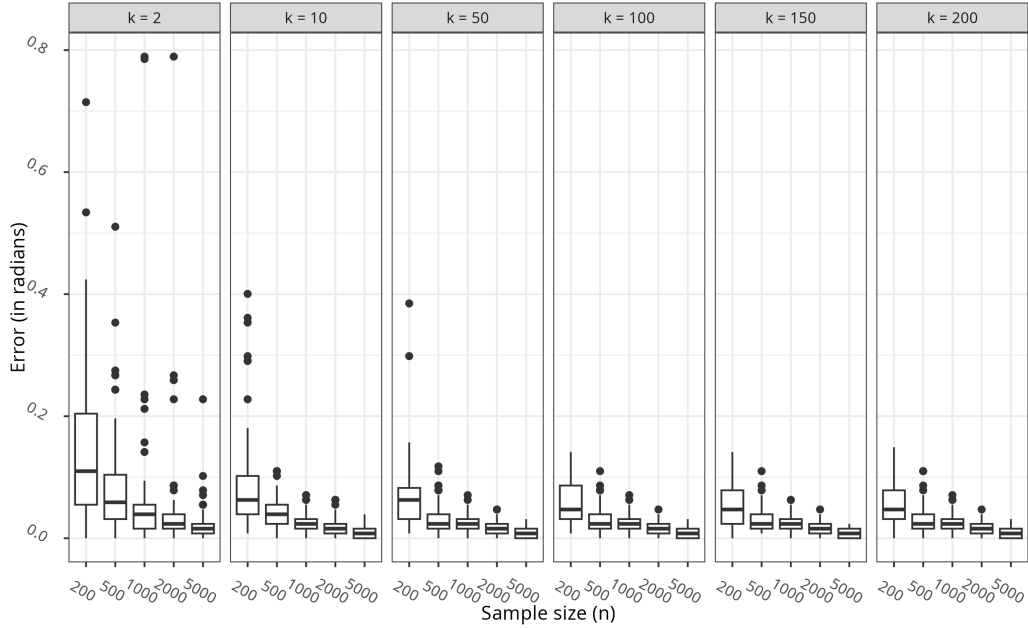Table 3: Frequency of correct number of minima detected, for **Scenario 2**.

Figure 4: Average angular error when considering different sample sizes, and different values of k, for **Scenario 1**.
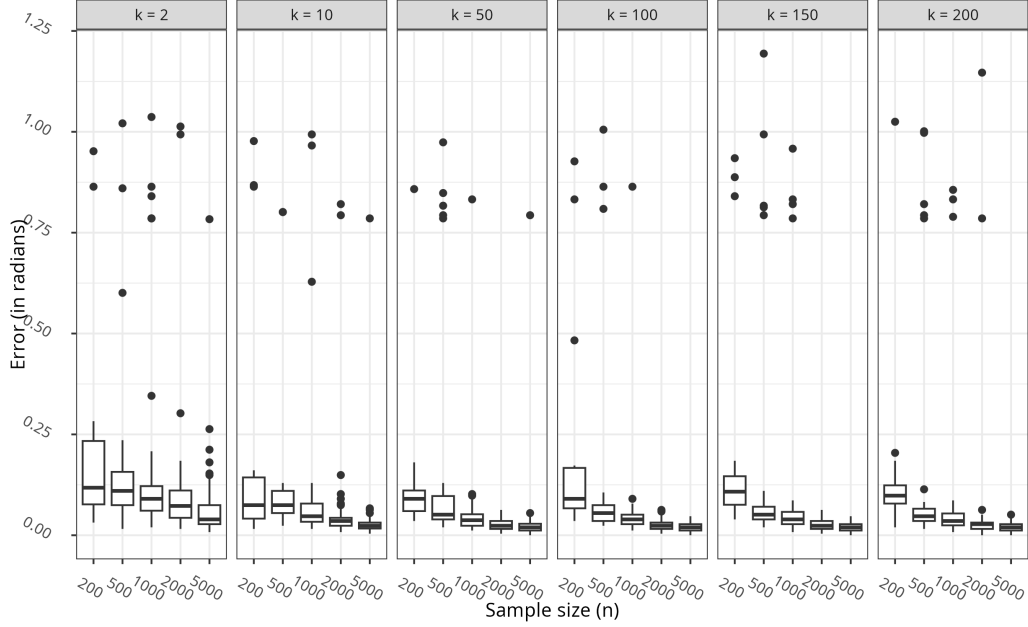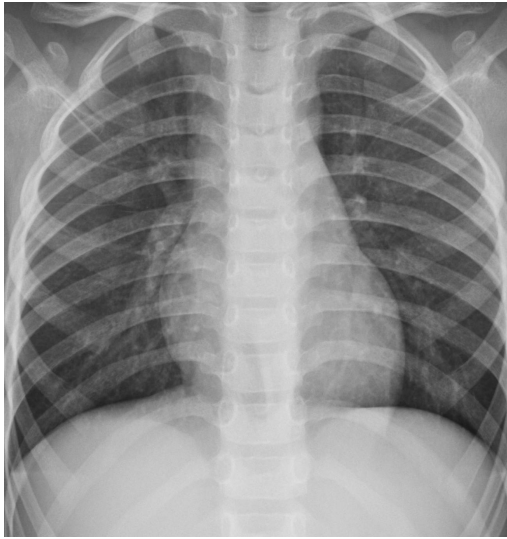
21

Figure 5: Average angular error when considering different sample sizes, and different values of k, for **Scenario 2**.
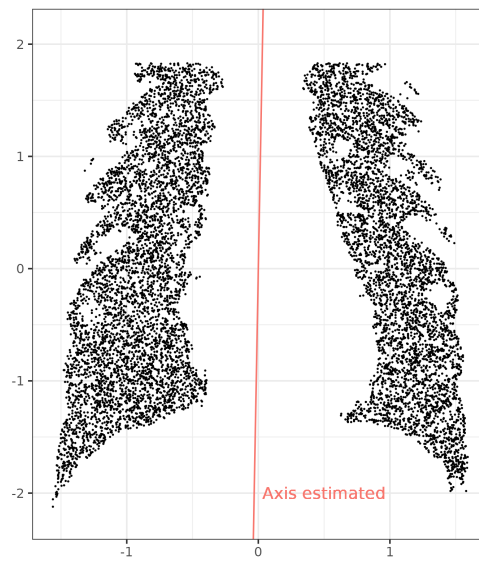
6. Application to real data: analysis of medical images

We conclude by presenting a methodological application of our estimation technique to the field of medical image analysis.

In [14], a method is presented for measuring symmetry in medical images, which offers a feature that improves disease detection. The first step involves determining an axis of symmetry. Subsequently, each point inside a shape on one side of the image (e.g., a point on the right lung) is compared to its symmetric counterpart in the shape on the other side of the image (i.e., the left lung). This comparison process results in a local symmetry measure. By averaging these local symmetry measures over the entire image, a global symmetry measure is obtained. This quantification helps in analyzing and detecting potential asymmetries or deviations from perfect symmetry in the medical images, which can be indicative of pathological changes or abnormalities. The combination of local and global symmetry measures provides valuable information for disease detection and medical image analysis.

22

Figure 6: (a) A frontal chest radiograph. The region of interest (the lungs) is highlighted. (b) A set of independent random points sampled uniformly from within the lung shape. The red line indicates the axis of symmetry estimated from this point cloud.

One issue here is that, despite existing protocols on medical imaging, in practice, the axis of symmetry is not known with certainty and must be estimated. The authors suggest estimating it as the direction that optimizes the proposed symmetry measure. In this matter, we will use our estimator to identify the axis of symmetry in a different way, which allows us to correctly align the images. By utilizing our estimator, we provide an alternative approach for estimating the axis of symmetry, contributing to improved medical image analysis and disease detection.
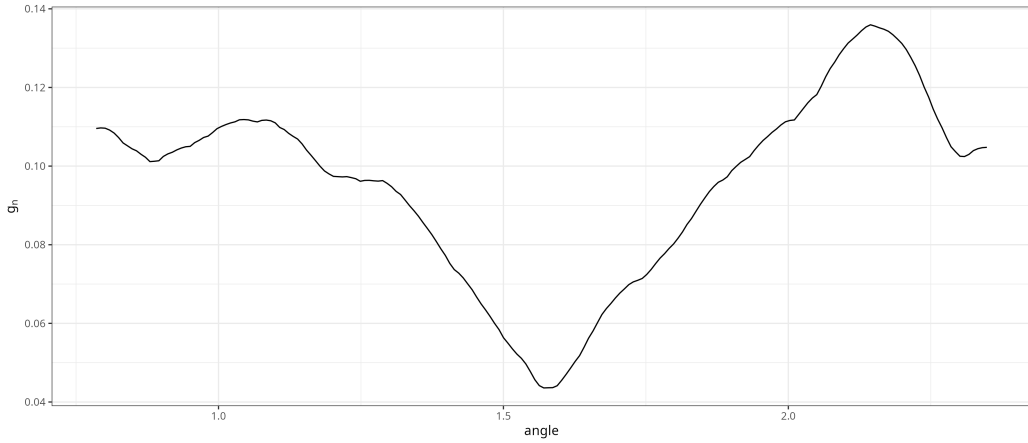

Figure 7: Plot of the estimating function $\widehat{g}_n$ for the lungs data.

A way to compute our estimator is to draw a sample of random points inside the shapes that we are interested in analyzing. Suppose we have a grayscale image like the one shown in Figure 6a, and we are interested in analyzing anomalies in the lungs. We can select pixels with a normalized intensity below a chosen threshold (here, 0.45, where 0 represents pure black and 1 pure white). We then draw a subsample of $n = 10000$ points from these pixel locations with some added jitter, as shown in Figure 6b.

By using this sampling procedure, we obtain a sample $\aleph_n = \{X_1, \ldots, X_n\}$ from an unknown distribution F_X of $X \in \mathbb{R}^2$ in the plane. If we assume that there exists a vector $u \in \mathbb{S}^1$ that makes $X \stackrel{d}{=} R_u X$, we would be able to estimate it consistently. It is important to note that this sampling procedure is intended solely for estimating the axis of symmetry, not for measuring the symmetry features that would require more information.

In Figure 7, the estimator $\widehat{g}_n$ is plotted. The figure clearly shows a unique global minimum in the direction of the angle close to $\pi/2$, where the column

is expected to be. The estimated axis of symmetry is displayed as the red line in Figure 6b.

References

[1] Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley & Sons, New York.

[2] Billingsley, P. (1995). *Probability and Measure*. John Wiley & Sons, New York, third edition.

[3] Carleman, T. (1922). Sur le problème des moments. *Comptes Rendus de l'Académie des Sciences de Paris*, 174:1680–1682.

[4] Cholaquidis, A., Fraiman, R., and Moreno, L. (2022). Level set and density estimation on manifolds. *Journal of Multivariate Analysis*, 189:104925.

[5] Cramér, H. and Wold, H. (1936). Some theorems on distribution functions. *The Journal of the London Mathematical Society*, 11(4):290–295.

[6] Cuesta-Albertos, J. A., Fraiman, R., and Ransford, T. (2007). A sharp form of the Cramér–Wold theorem. *Journal of Theoretical Probability*, 20(2):201–209.

[7] Devroye, L., Györfi, L., and Lugosi, G. (1996). *A Probabilistic Theory of Pattern Recognition*. Springer, New York.

[8] Dudley, R. M. (2002). *Real Analysis and Probability*. Cambridge University Press, Cambridge. Revised reprint of the 1989 original.

[9] Fang, K. T., Kotz, S., and Ng, K. W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman & Hall, London.

[10] Ferguson, T. S. (1962). A representation of the symmetric bivariate Cauchy distribution. *The Annals of Mathematical Statistics*, 33(4):1256–1266.

[11] Fraiman, R., Moreno, L., and Ransford, T. (2024). Application of the Cramér–Wold theorem to testing for invariance under group actions. *TEST*, 33(2):379–399.

[12] Fraiman, R., Moreno, L., and Vallejo, S. (2017). Some hypothesis tests based on random projection. *Computational Statistics*, 32(3):1165–1189.

[13] Fremlin, D. H. (2003). *Measure Theory, Volume 2.* Torres Fremlin, Colchester.

[14] Hogeweg, L., Sánchez, C. I., Maduskar, P., Philipsen, R. H. H. M., and van Ginneken, B. (2017). Fast and effective quantification of symmetry in medical images for pathology detection: Application to chest radiography. *Medical Physics*, 44(6):2242–2256.

[15] Koldobsky, A. (2005). *Fourier Analysis in Convex Geometry.* American Mathematical Society, Providence, RI.

[16] Martos, G. and de Carvalho, M. (2018). Discrimination surfaces with application to region-specific brain asymmetry analysis. *Statistics in Medicine*, 37(11):1859–1873.

[17] Molchanov, I. S. (1998). A limit theorem for solutions of inequalities. *Scandinavian Journal of Statistics*, 25(1):235–242.

[18] Murphy, K., Habib, S. S., Zaidi, S. M. A., Khowaja, S., Khan, A., Melendez, J., Scholten, E. T., Amad, F., Schalekamp, S., Verhagen, M., Philipsen, R. H. H. M., Meijers, A., and van Ginneken, B. (2020). Computer aided detection of tuberculosis on chest radiographs: An evaluation of the CAD4TB v6 system. *Scientific Reports*, 10(1):5492.

[19] Pewsey, A. and García-Portugués, E. (2021). Recent advances in directional statistics. *TEST*, 30(1):1–58.

[20] Ranošová, H. (2023). α-symmetric measures. Master's thesis, Charles University.

[21] Rockafellar, R. T. and Wets, R. J.-B. (1998). *Variational Analysis.* Springer, Berlin.

[22] Schneider, R. (2014). *Convex Bodies: The Brunn–Minkowski Theory.* Cambridge University Press, Cambridge, expanded edition.

[23] Scholkmann, F., Boss, J., and Wolf, M. (2012). An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals. *Algorithms*, 5(4):588–603.

[24] Serfling, R. J. (2006). Multivariate symmetry and asymmetry. In Kotz, S., Read, C. B., Balakrishnan, N., and Vidakovic, B., editors, *Encyclopedia of Statistical Sciences*. John Wiley & Sons.