

Point Estimation of the Central Orientation of Random Rotations

Bryan Stanfill*

Ulrike Genschel

and

Heike Hofmann

Department of Statistics, Iowa State University

October 29, 2012

Abstract

Data in the form of random three-dimensional rotations have application in computer science, kinematics and materials sciences, among other areas. Estimating the central orientation from a sample of such data is an important problem, which is complicated by the fact that several different approaches exist for this, motivated by various geometrical and decision-theoretic considerations. However, little is known about how such estimators compare, especially on common distributions for location models with random rotations. We examine four location estimators, three of which are commonly found in different literatures and the fourth estimator (a projected median) is newly introduced. Our study unifies existing literature and provides a detailed numerical investigation of location estimators for three commonly used rotation distributions in statistics and materials science. While the data-generating model influences the best choice of an estimator, the proposed projected median emerges as an overall good performer, which can be suggested without particular distributional assumptions. We illustrate the estimators and our findings with data from a materials science study by approximating the central orientation of cubic crystals on the micro-surface of a metal.

Keywords: Cayley distribution, Electronic Backscatter Diffraction, Geodesic distance, Matrix Fisher distribution, Projected median, Rotation Group

*The authors gratefully acknowledge the valuable suggestions made by reviewers and the associate editor that have helped strengthening this manuscript.

1 Introduction

Data in the form of 3×3 rotation matrices find application in several scientific areas, such as biomedical engineering, computer visioning, and geological and materials sciences, where such data represent the positions of objects within some three-dimensional reference frame. For example, Rancourt et al. (2000) examine rotation matrix data in studying body positions whilst operating machinery. Fletcher et al. (2009) consider this type of orientation data in magnetic resonance imaging and in shape analysis; similar examples can be found in Schwartz and Rozumalski (2005), Pierrynowski and Ball (2009), Dai et al. (2010), or Hadani and Singer (2011). The data in our illustrative example to follow arise from a study in materials science, where 3×3 rotations represent the orientations of cubic crystals on the micro-surface of a metal specimen as measured through electron backscatter diffraction (EBSD) and “grains” within metals are composed of crystals which roughly share a common orientation; see Randle (2003) for details on EBSD data.

From a sample of orientations, an important interest is often the estimation of a main or central orientation \mathbf{S} . That is, letting the rotation group $SO(3)$ denote the collection of all 3×3 rotation matrices, observations $\mathbf{R}_1, \dots, \mathbf{R}_n \in SO(3)$ can be conceptualized as a random sample from a *location model*

$$\mathbf{R}_i = \mathbf{S}\mathbf{E}_i, \quad i = 1, \dots, n, \tag{1}$$

where $\mathbf{S} \in SO(3)$ is the *fixed* parameter of interest indicating an orientation of central tendency, and $\mathbf{E}_1, \dots, \mathbf{E}_n \in SO(3)$ denote i.i.d. *random* rotations which symmetrically perturb \mathbf{S} . The data-generating model in (1) is a rotation-matrix analog of a location model for scalar data $Y_i = \mu + e_i$, where $\mu \in \mathbb{R}$ denotes a mean and $e_i \in \mathbb{R}$ denotes an additive error symmetrically distributed around zero. This representation (1) for orientations is quite common and, in fact, a variety of parametric models exist for describing symmetrically distributed rotations \mathbf{E}_i , such as the symmetric matrix Fisher distribution (Downs, 1972), the symmetric Cayley distribution (León et al., 2006) and the circular-von Mises-based rotation distribution (Bingham et al., 2009) in the statistics literature, as well as the Bunge distribution (Bunge, 1982), the isotropic Gaussian distribution (Matthies et al., 1988; Savy-

olova and Nikolayev, 1995) and the de la Vallée Poussin distribution (Schaeben, 1997) in the materials science literature. Our goal in this paper is to summarize and compare the most frequently proposed approaches for the point estimation of \mathbf{S} based on a sample of orientation data generated by (1). Depending on the scientific literature, the approaches can be quite different.

The topic of location estimation has received considerable attention for directional data on circles or spheres, (see Fisher, 1953; Karcher, 1977; Khatri and Mardia, 1977; Fisher, 1985; Ducharme and Milasevic, 1987; Bajaj, 1988; Liu and Singh, 1992; Chan and He, 1993; Mardia and Jupp, 2000), but less is known about estimator properties with rotation data. As a compounding factor, several current approaches to estimating \mathbf{S} have arisen out of literatures having differing statistical and geometrical emphases. In the applied sciences literature, estimators of \mathbf{S} are typically based on *non-Euclidean* (i.e., Riemannian) geometry, such as the *geometric mean* (Arun et al., 1987; Horn et al., 1988; Umeyama, 1991; Moakher, 2002) or, more recently, the *geometric median* (Fletcher et al., 2008; 2009). Preferences may depend on outliers in the data, but such suggestions for estimating \mathbf{S} often do *not* consider the potential impact of the underlying data-generating mechanism. On the other hand, approaches in the statistics literature tend to motivate an estimator for \mathbf{S} through likelihood or moment-estimation principles applied to a specifically assumed distributional model (e.g., matrix Fisher or Cayley distribution) for the symmetric rotation errors \mathbf{E}_i (Downs, 1972; Jupp and Mardia, 1979; León et al., 2006; Bingham et al., 2010). Almost always, this estimator turns out to be a *projected arithmetic mean* based on *Euclidean* geometry. Hence, in addition to potential distributional assumptions, more fundamental divisions in estimation approaches may be attributable to different geometrical perspectives with rotation data.

Considering the potential effects of an underlying data generation model as well as the choice of geometry (i.e., Euclidean vs. Riemannian), the above discussion indicates a need to investigate and identify good point estimators for rotation data. In particular, because estimators in the applied sciences literatures are often selected without decision-theoretical considerations based on underlying distributions, it is of interest to understand how different location estimators behave across common distributions for rotations. In this paper,

we evaluate four estimators for \mathbf{S} in the context of the location model (1). These are either mean- or median-type estimators and based either on Euclidean- or Riemannian geometry; the Euclidean-based median estimator is introduced for the first time for $SO(3)$ data. Its inclusion is natural and its performance can be generally quite good and is (as will be demonstrated) broadly recommendable. Through simulation, we compare how these estimators perform with respect to three common probability models for symmetric rotation errors as defined in (1), namely the circular-von Mises-based distribution, the symmetric matrix Fisher distribution and the symmetric Cayley distribution. The matrix Fisher is arguably the most common distribution in the statistics literature (see Chikuse, 2003). While not noted previously, the symmetric Cayley and the de la Vallée Poussin distribution are in fact the same; the de la Vallée Poussin distribution has been advocated in the materials science literature (Schaeben, 1997). The circular-von Mises-based distribution is included because the distribution is non-regular and has been applied to EBSD data (Bingham et al., 2009). We describe how properties of error distributions for rotation data, in particular their variability and tail behavior, translate into performance differences among point estimators.

The remainder of the manuscript is organized as follows. Section 2 provides a brief background on the geometry of rotations and different distance metrics that can be used to assess overall estimation bias. Section 3 then describes the location estimators for rotation data and compares their geometric underpinnings, which serves to unify some of the existing estimation literature. Section 4 explains the design of the simulation study followed by a summary of our main findings in Section 5. Section 6 provides an illustration of the estimation methods for EBSD data in a materials science application. We provide concluding remarks and future research possibilities in Section 7.

2 Background

2.1 Geometry of Three-dimensional Orientations

Three-dimensional orientation data consist of observations belonging to the group $SO(3)$ of all 3×3 rotation matrices, where an element \mathbf{R} in $SO(3)$ is an orthogonal 3×3 matrix

(i.e., $\mathbf{R}^\top \mathbf{R} = \mathbf{I}_{3 \times 3}$) with determinant one. As $SO(3)$ is a Lie group, its elements live on a differentiable manifold. This fact is helpful in understanding the two different geometric approaches for estimating the central location $\mathbf{S} \in SO(3)$ from a sample of orientation data, referred to here as the *intrinsic* and the *embedding* estimation approaches (see also Jupp and Mardia (1989) and Mardia and Jupp (2000) for analogs with directional data). The rotation group $SO(3)$ is not closed under routine addition or scalar multiplication (i.e., operations natural to statisticians). Hence, statistical estimation approaches often *embed* the rotation group into the higher-dimensional linear space consisting of all 3×3 real matrices, denoted as $\mathcal{M}(3)$. Doing so enables the use of the familiar Euclidean geometry (and “averaging” notions) to define standard distance measures and loss criteria for obtaining location estimators (see Section 2.2 and the estimators given in Sections 3.1 and 3.2). This embedding technique has been largely applied by statisticians, typically resulting in the projected arithmetic mean of Section 3.1. See, for example, Downs (1972); Khatri and Mardia (1977) and Jupp and Mardia (1979; 1989); the Bayesian estimator used in Bingham et al. (2010) is also a concrete example of this approach as is the median-type estimator we propose in Section 3.2.

Alternatively, *intrinsic* estimation approaches use Riemannian geometry to define distances that account for the innate topology or curvature of the space $SO(3)$. In the intrinsic approach, each rotation from $SO(3)$ is associated with a skew-symmetric matrix $\Phi(\mathbf{W})$, defined as

$$\Phi(\mathbf{W}) = \begin{bmatrix} 0 & -w_3 & -w_2 \\ w_3 & 0 & -w_1 \\ w_2 & w_1 & 0 \end{bmatrix}$$

for $\mathbf{W} = (w_1, w_2, w_3)^\top \in \mathbb{R}^3$. That is, through a so-called exponential operator, we map $\Phi(\mathbf{W})$ to a rotation matrix as

$$\exp[\Phi(\mathbf{W})] = \cos(r)\mathbf{I}_{3 \times 3} + \sin(r)\Phi(\mathbf{U}) + (1 - \cos r)\mathbf{U}\mathbf{U}^\top$$

where $r = \|\mathbf{W}\|$ and $\mathbf{U} = \mathbf{W}/\|\mathbf{W}\|$. The space $\mathfrak{so}(3)$ of all skew-symmetric matrices forms the tangent space (Lie-algebra) of $SO(3)$, which is closed under familiar summation and scalar multiplication operations in the usual (i.e., element-wise) manner. The fact

that $SO(3)$ is a differentiable manifold allows a distance measure (the geodesic distance in Section 2.2) to be defined between points in $SO(3)$ according to Riemannian geometry. The resulting geodesic distance provides the basis for the “geometric” location estimators commonly found in computer science (Fletcher et al., 2008; 2009; Hartley et al., 2011) and engineering applications (Manton, 2004); see Sections 3.3 and 3.4.

Before leaving this section, it is helpful to note that each rotation matrix \mathbf{R} can be associated with a pair (r, \mathbf{U}) , where $r \in (-\pi, \pi]$ and $\mathbf{U} \in \mathbb{R}^3$, $\|\mathbf{U}\| = 1$, through

$$\mathbf{R} = \mathbf{R}(r, \mathbf{U}) = \exp[\Phi(\mathbf{U}r)] \in SO(3). \quad (2)$$

This is the so-called axis-angle representation of \mathbf{R} , where \mathbf{R} is represented by rotating the coordinate axis $\mathbf{I}_{3 \times 3}$ about the axis $\mathbf{U} \in \mathbb{R}^3$ by the angle r . In the materials science literature, \mathbf{U} and r are commonly referred to as the misorientation axis and misorientation angle of \mathbf{R} with respect to $\mathbf{I}_{3 \times 3}$; see Randle (2003).

2.2 Choice of Distance Metrics

The choice of geometry, i.e. Riemannian or Euclidean, results in two different metrics to measure the distance between two rotation matrices \mathbf{R}_1 and $\mathbf{R}_2 \in SO(3)$. Under the embedding approach, the natural distance metric between two random matrices is the Euclidean distance, d_E , which is induced by the Frobenius norm

$$d_E(\mathbf{R}_1, \mathbf{R}_2) = \|\mathbf{R}_1 - \mathbf{R}_2\|_F, \quad (3)$$

where $\|\mathbf{A}\|_F = \sqrt{\text{tr}(\mathbf{A}^\top \mathbf{A})}$ denotes the Frobenius norm of a matrix \mathbf{A} and $\text{tr}(\cdot)$ denotes the trace of a matrix. The Euclidean distance between two rotation matrices corresponds to the shortest cord in $\mathcal{M}(3)$ that connects both matrices. If $r \in (-\pi, \pi]$ denotes the misorientation angle in the angle-axis representation (2) of $\mathbf{R}_1^\top \mathbf{R}_2 \equiv \mathbf{R}_1^\top \mathbf{R}_2(r, \mathbf{U})$ (so that $\text{tr}(\mathbf{R}_1^\top \mathbf{R}_2) = 1 + 2 \cos r$), then $d_E(\mathbf{R}_1, \mathbf{R}_2) = 2\sqrt{(1 - \cos r)}$ holds.

By staying in the Riemannian space $SO(3)$ under the intrinsic approach, the natural distance metric becomes the Riemannian (or geodesic) distance, d_R , by which the distance

between two rotations $\mathbf{R}_1, \mathbf{R}_2 \in SO(3)$ is defined as

$$d_R(\mathbf{R}_1, \mathbf{R}_2) = \frac{1}{\sqrt{2}} \|\text{Log}(\mathbf{R}_1^\top \mathbf{R}_2)\|_F = |r|, \quad (4)$$

where $\text{Log}(\mathbf{R})$ denotes the principle logarithm of \mathbf{R} (i.e., $\text{Log}(\mathbf{R}) = \text{Log}(\mathbf{R}(\mathbf{U}, r)) = \Phi(r\mathbf{U})$ in (2)) and $r \in (-\pi, \pi]$ is the misorientation angle of $\mathbf{R}_1^\top \mathbf{R}_2$. The Riemannian distance corresponds to the length of the shortest path that connects \mathbf{R}_1 and \mathbf{R}_2 *within* the space $SO(3)$. For this reason, the Riemannian distance is often considered the more natural metric on $SO(3)$; see Moakher (2002) for this discussion along with more details on exponential/logarithmic operators related to $SO(3)$.

To exemplify the difference between the Euclidean and Riemannian distance metric, we visualize each distance in a lower dimensional example in Figure 1. In Figure 1 $\mathbf{R}_1, \mathbf{R}_2$ refer to the endpoints of two rotations on $SO(2)$. The Riemannian distance between both rotations, $d_R(\mathbf{R}_1, \mathbf{R}_2)$, corresponds to the thick curved line. The rotation of $\mathbf{R}_1\mathbf{v}$ into $\mathbf{R}_2\mathbf{v}$ follows the curved path and therefore is considered to stay in the space. The Euclidean distance between $\mathbf{R}_1, \mathbf{R}_2$, $d_E(\mathbf{R}_1, \mathbf{R}_2)$, is represented by the straight gray line. Because it is not possible for $\mathbf{R}_1\mathbf{v}$ to follow this straight path, the distance $d_E(\mathbf{R}_1, \mathbf{R}_2)$ can only be obtained by leaving $SO(2)$.

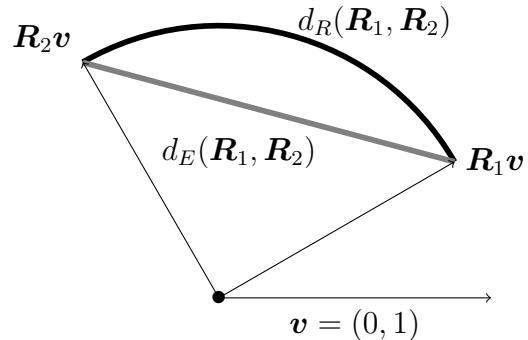


Figure 1: An illustration of the Euclidean and Riemannian distance metric on $SO(2)$.

3 Location Estimators

This section describes four estimators for the location parameter $\mathbf{S} \in SO(3)$ corresponding to orientation data generated by the model in (1). The estimators are based on two different choices. First, the choice whether to use the embedding approach, i.e. to base the estimator on the Euclidean distance metric defined in (3) or alternatively to use the intrinsic approach by employing the Riemannian distance metric as defined in (4). The second choice concerns the decision-theoretic loss functions, i.e., either using squared deviations or absolute deviations. The extent to which the choice of geometry or loss function matters in the estimation of \mathbf{S} will be an important aspect explored in Section 4. We provide an overview of all four estimators and their properties in Table 1.

3.1 The projected arithmetic mean

We begin with the definition of the arithmetic mean for orientation data, as its analog is most frequently encountered in the statistical literature for directional data (e.g. see Mardia and Jupp, 2000). For a sample of n random rotations $\mathbf{R}_i \in SO(3)$, $i = 1, 2, \dots, n$, this mean-type estimator is defined as

$$\hat{\mathbf{S}}_E = \arg \min_{\mathbf{S} \in SO(3)} \sum_{i=1}^n d_E^2(\mathbf{R}_i, \mathbf{S}) = \arg \max_{\mathbf{S} \in SO(3)} \text{tr}(\mathbf{S}^\top \bar{\mathbf{R}}) \quad (5)$$

where $\bar{\mathbf{R}} = \frac{1}{n} \sum_{i=1}^n \mathbf{R}_i$. The estimator is obtained by minimizing the sum of the squared deviations in the Euclidean sense in the ambient space $\mathcal{M}(3)$, which then is projected back into $SO(3)$. Moakher (2002), who studied the mathematical characteristics of this estimator in detail, therefore refers to it as the *projected arithmetic mean*. This estimator's appeal lies in its simplicity and statistically intuitive nature, though it has been noted that the estimator is not invariant under rigid transformations (see Moakher, 2002). However, the estimator does correspond to the maximum likelihood estimator of \mathbf{S} when the symmetrically distributed rotation errors in (1) follow a matrix Fisher distribution (Jupp and Mardia, 1979). León et al. (2006) also derived this estimator as the method of moment estimator under a Cayley distribution, and Bingham et al. (2009) showed that the projected arithmetic mean corresponds to the maximum quasi-likelihood estimator for orientation

data with rotation errors arising from the circular-von Mises-based distribution.

For a numerical implementation to calculate $\widehat{\mathbf{S}}_E$ we refer to algorithms proposed by Arun et al. (1987) and Horn et al. (1988) as well as to Umeyama (1991) for refinements of their solutions including special cases such as $\det(\bar{\mathbf{R}}) = 0$.

3.2 The projected median

A modification of the estimator from Section 3.1 is obtained by replacing the squared distances in (5) with absolute distances, leading to a median-type estimator defined as

$$\widetilde{\mathbf{S}}_E = \arg \min_{\mathbf{S} \in SO(3)} \sum_{i=1}^n d_E(\mathbf{R}_i, \mathbf{S}). \quad (6)$$

We will refer to this estimator of \mathbf{S} as the *projected median*. To the best of our knowledge, the estimator defined in (6) is the first median-type estimator for $SO(3)$ data.

Although median-type estimators exist for high dimensional and directional data, these estimators had previously not been extended to rotational data. Proposed estimators include, for example, the Euclidean median (also known as the Weber point (Bajaj, 1988)), the mediancentre (Gower, 1974) or the projection median (Durocher and Kirkpatrick, 2009) for $X_1, \dots, X_n \in \mathbb{R}^p$. Specifically, Durocher and Kirkpatrick (2009) explore the stability of the above-mentioned estimators with respect to small perturbations at a small, selected set of points. For directional and spherical data consider, for example, the circular median (Mardia, 1972), the normalized spatial median (Ducharme and Milasevic, 1987) and the Fisher median (Fisher, 1985) also known as the spherical median. Chan and He (1993), for example compare the performance of the normalized spatial median, an L_1 estimator on the sphere by He and Simpson (199) and the Fisher median for the central direction for spherical data following the von Mises-Fisher distribution. They conclude that the normalized spatial median estimator is preferable for spherical data under the von Mises-Fisher model.

We next propose an algorithm to compute the projected median. We base our method on the Weiszfeld algorithm originally given by Weiszfeld (1937). The algorithm requires an initial value that does not equal any sample point. Note that the solution is generally not

sensitive to the choice of starting points unless the data exhibit extreme spread.

1. Set $\widehat{\mathbf{S}} = \widehat{\mathbf{S}}_E$ and choose an arbitrarily small stopping rule ε .
2. For $i = 1, \dots, n$ compute $\mathbf{s}_i = \mathbf{R}_i - \widehat{\mathbf{S}}$.
3. Calculate

$$\bar{\mathbf{R}}_W = \frac{\sum_{i=1}^n \mathbf{R}_i / \|\mathbf{s}_i\|_F}{\sum_{i=1}^n 1 / \|\mathbf{s}_i\|_F}$$

which we call the weighted mean with respect to $\widehat{\mathbf{S}}$. Note that there is a theoretical probability of zero for $\|\mathbf{s}_i\|_F = 0$, in practice we do impose a lower limit on $\|\mathbf{s}_i\|_F = \delta > 0$ (for some $\delta \in \mathbf{R}$) to avoid division by zero.

4. Define $\widehat{\mathbf{S}}_{\text{new}} = \arg \max_{\mathbf{S} \in SO(3)} \text{tr}(\mathbf{S}^\top \bar{\mathbf{R}}_W)$ to be the $\mathcal{M}(3)$ projection of $\bar{\mathbf{R}}_W$.
5. If $\varepsilon > \|\widehat{\mathbf{S}} - \widehat{\mathbf{S}}_{\text{new}}\|_F$ return $\widetilde{\mathbf{S}}_E = \widehat{\mathbf{S}}_{\text{new}}$; otherwise set $\widehat{\mathbf{S}} = \widehat{\mathbf{S}}_{\text{new}}$ and return to step 2.

3.3 The geometric mean

As sketched in Section 2.1, the Lie group property of $SO(3)$ provides us with a convenient transform from $SO(3)$ into the tangent space $\mathfrak{so}(3)$ that is closed under addition and scalar multiplication. Obtaining the median or mean in this transformed space and projecting the result back to $SO(3)$ corresponds to the rotation that minimizes the first and second order Riemannian distances, respectively (Karcher, 1977; Moakher, 2002; Fletcher et al., 2008; 2009). Karcher (1977) made use of Riemannian manifolds to compute what is often called the Riemannian center of mass. Moakher (2002) applied Karcher's ideas to rotation matrices and defined

$$\widehat{\mathbf{S}}_R = \arg \min_{\mathbf{S} \in SO(3)} \sum_{i=1}^n d_R^2(\mathbf{R}_i, \mathbf{S}). \quad (7)$$

which was termed as the *geometric mean*. Note that the solution to (7) may not be unique. Uniqueness is tied to the property of geodesic convexity of the objective function in (7). For more information, we refer to Moakher (2002). Additionally, (7) generally does not have a closed-form solution making this estimator much more computationally intensive than its Euclidean counterpart (the projected arithmetic mean of Section 3.1). We used the algorithm proposed by Manton (2004) in our simulation study.

3.4 The geometric median

The median-type counterpart to the geometric mean was defined first in the context of spherical data by Fisher (1985) as the point on the sphere that minimizes the sum of the arc lengths to all observations in the sample. For this type of data, the resulting estimator is known as the spherical median, which is a special case of the generalized median in \mathbb{R}^d proposed by Gower (1974). For spherical data, an alternative formulation to the spherical median has been given by Liu and Singh (1992) in the framework of data depth leading, however, to the same solution.

We give an adaptation of the spherical median to rotation matrices. Recall that the shortest geodesic path between two rotations $\mathbf{R}_1, \mathbf{R}_2$ is given by the Riemannian distance $d_R(\mathbf{R}_1, \mathbf{R}_2)$. Thus the rotation matrix analog of the Fisher (1985) spherical median can be defined as

$$\tilde{\mathbf{S}}_R = \arg \min_{\mathbf{S} \in SO(3)} \sum_{i=1}^n d_R(\mathbf{R}_i, \mathbf{S}); \quad (8)$$

see also Fletcher et al. (2008; 2009). We refer to this estimator of \mathbf{S} as the *geometric median*. Hartley et al. (2011) offers an algorithm to find the geometric median in $SO(3)$.

Table 1: A summary of the presented estimators and their properties.

Name	Symbol	Distance metric	Cost function
Projected Arithmetic Mean	$\hat{\mathbf{S}}_E$	Euclidean	$\sum_{i=1}^n d_E^2$
Projected Median	$\tilde{\mathbf{S}}_E$	Euclidean	$\sum_{i=1}^n d_E$
Geometric Mean	$\hat{\mathbf{S}}_R$	Riemannian	$\sum_{i=1}^n d_R^2$
Geometric Median	$\tilde{\mathbf{S}}_R$	Riemannian	$\sum_{i=1}^n d_R$

4 Simulation Study

Section 4.1 gives an outline of the simulation design. Section 4.2 briefly describes the parametric distributional models for describing symmetric rotation errors (cf. (1)) with

differing variability that are used in the study.

4.1 Design of Simulation Study

To compare the performance of the proposed location estimators for determining the central direction \mathbf{S} given a sample of size n , we generated random rotation error samples $\mathbf{E}_1, \dots, \mathbf{E}_n$ in model (1) with sizes $n = 10, 50, 100$ and 300 . In (1), without loss of generality, we set the location parameter $\mathbf{S} = \mathbf{I}_{3 \times 3}$ (the identity matrix) and consider circular variances (denoted by ν) of $\nu = 0.25, 0.50$ and 0.75 .

For each combination of sample size, circular spread ν and choice of distribution, we generated 1,000 samples and for each sample estimated the central direction $\mathbf{S} = \mathbf{I}_{3 \times 3}$ using each of the four proposed estimators. We continue with an introduction to the distributions under consideration in the next section.

4.2 Generating Random Rotations in the Location Model

We wish to compare estimators of the (fixed) location parameter $\mathbf{S} \in SO(3)$ under three common distributional models for describing symmetric rotation errors $\mathbf{E} \in SO(3)$ in a data model $\mathbf{R} = \mathbf{SE}$ (cf. (1)): the symmetric matrix Fisher (Langevin, 1905; Downs, 1972; Khatri and Mardia, 1977; Jupp and Mardia, 1979), the symmetric Cayley (Schaeben, 1997; León et al., 2006) and the circular-von Mises-based distribution (Bingham et al., 2009). A general construction approach exists for random rotations that are symmetrically distributed around the identity matrix $\mathbf{I}_{3 \times 3}$; see Watson (1983); Bingham et al. (2009) and Hielscher et al. (2010). To this end, let $\mathbf{U} \in \mathbb{R}^3$ represent a point chosen uniformly on the unit sphere and, independently, generate a random angle r according to some circular density $C(r|\kappa)$ on $(-\pi, \pi]$, which is symmetric around 0 and where κ denotes a concentration parameter governing the spread of the circular distribution. Then, define a random rotation as $\mathbf{E} = \mathbf{E}(\mathbf{U}, r)$ using the constructive definition (2) (i.e., \mathbf{E} represents the position of $\mathbf{I}_{3 \times 3}$ upon rotating the standard coordinate frame in \mathbb{R}^3 about the random axis \mathbf{U} by the random angle r). The resulting rotation \mathbf{E} will be symmetrically distributed and its distributional type (i.e., matrix Fisher, Cayley or circular-von Mises-based) is determined by the form of the circular density $C(r|\kappa)$ for the (misorientation) angle r .

Table 2: Circular densities with respect to the Lebesgue measure and circular variance ν .

Name	Density $C(r \kappa)$	Circular variance ν
Cayley	$\frac{1}{\sqrt{\pi}} \frac{\Gamma(\kappa+2)}{\Gamma(\kappa+1/2)} 2^{-(\kappa+1)} (1 + \cos r)^\kappa (1 - \cos r)$	$\frac{3}{\kappa+2}$
matrix Fisher	$\frac{1}{2\pi[I_0(2\kappa)-I_1(2\kappa)]} e^{2\kappa \cos(r)} [1 - \cos(r)]$	$\frac{3I_0(2\kappa)-4I_1(2\kappa)+I_2(2\kappa)}{2[I_0(2\kappa)-I_1(2\kappa)]}$
circular-von Mises	$\frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(r)}$	$\frac{I_0(\kappa)-I_1(\kappa)}{I_0(\kappa)}$

The circular densities are given in Table 2, where $I_p(\cdot)$ denotes the Bessel function of order p defined as $I_p(\kappa) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(pr) e^{\kappa \cos r} dr$. To ease the interpretation of the simulation results in Section 5, we choose not to present results in terms of κ but instead use the circular variance defined as $\nu = 1 - \rho$ as a measure of spread, where $\rho = E[\cos(r)]$, which is commonly referred to as the mean resultant length. This allows us to compare the performance of the estimators for densities exhibiting the same spread. The values of κ corresponding to the chosen circular variances are given in Table 3.

Table 3: Values of κ for each rotational distribution corresponding to the circular variances.

Distribution	Circular variance		
	$\nu = 0.25$	$\nu = 0.50$	$\nu = 0.75$
Cayley	10.00	4.00	2.00
matrix Fisher	3.17	1.71	1.15
circular-von Mises	2.40	1.16	0.52

The density, with respect to the Haar measure, for each distribution of a random rotation given a circular variance of 0.75 is plotted in Figure 2. The Haar measure (or uniform distribution on $SO(3)$) acts as the dominating measure for rotations and the symmetric nature of the random rotation $\mathbf{E}_i = \mathbf{E}_i(\mathbf{U}, r)$ means that its density $f(\mathbf{E}_i|\nu) = f(r|\nu)$ can be plotted in terms of the misorientation angle r of \mathbf{E}_i in (2), which is common in materials science (Matthies et al., 1988; Savelyova and Nikolayev, 1995). Density plots for the other circular variances we consider in our simulation study are similar and therefore omitted.



Figure 2: Comparison of the distributions of interest for $\nu = 0.75$. The circular-von Mises based-distribution has the highest concentration, but surprisingly also the heaviest tail.

In Figure 2a we see that the circular-von Mises-based distribution has a much higher mass around zero than the other two distributions. Figure 2b illustrates that the Cayley distribution has the least mass around its mode, and that all three distributions cross one another twice. We will see later that this property provides a potential explanation for some of the behavior we observe in the estimators. Finally, Figure 2c shows the tail behavior of the distributions and indicates that the circular-von Mises-based distribution has the heaviest tail of the three distributions.

A visualization of a random sample of 100 rotations, one sample for each of the three angular distributions, is given in the sphere plots in Figure 3. The samples are adjusted to have an angular variance of $\nu = 0.25$. Note that Figure 3 shows only the first column of each rotation matrix, because, however rotation matrices are orthogonal with each column and row representing one element on the unit sphere. Under the assumption of independence of the angle r and the main direction, any of the other directions is expected to show a similar pattern. Figure 3 also confirms the previous findings from Figure 2; the circular-von Mises-based distribution has the highest mass in close proximity around the main direction but at the same time shows the largest numbers of ‘outlying’ rotations, i.e. rotations with a large radial distance to the center, while the Cayley and the matrix-Fisher distribution appear to be more similar with a lower concentration of mass in the center for the Cayley distribution than for the matrix-Fisher distribution.

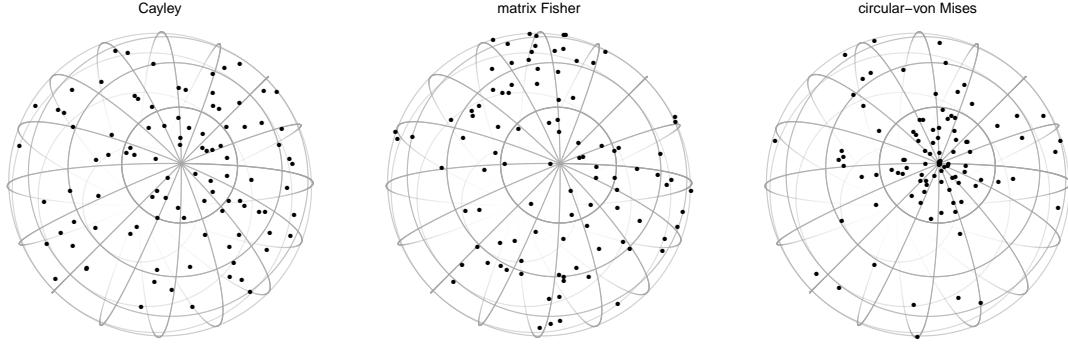


Figure 3: Sphere plots of the first column for each of the three angular distributions under an angular variance of $\nu = 0.25$

In the simulation study to follow, for generating random rotation errors based on the construction above, we used different samplers to randomly generate angles $r \in (-\pi, \pi]$ from a given circular density, recalling that the form of $C(r|\kappa)$ depends on the intended symmetric distribution for the rotation errors \mathbf{E} . We defer further details to Section A of the Appendix.

5 Results

In this section we summarize and present the main findings of the simulation study for estimating the central direction $\mathbf{S} = \mathbf{I}_{3 \times 3}$ with the four proposed estimators of Section 3. We quantify the estimation error between the true location $\mathbf{S} = \mathbf{I}_{3 \times 3}$ and an estimate $\widehat{\mathbf{S}}$ using the geodesic distance, i.e.

$$d_R(\mathbf{S}, \widehat{\mathbf{S}}) = \frac{1}{\sqrt{2}} \|\text{Log}(\mathbf{S}^\top \widehat{\mathbf{S}})\|_F, \quad \text{where } \widehat{\mathbf{S}} = \widehat{\mathbf{S}}_E, \widehat{\mathbf{S}}_R, \widetilde{\mathbf{S}}_E \text{ or } \widetilde{\mathbf{S}}_R. \quad (9)$$

Note that for any two rotations \mathbf{R}_1 and \mathbf{R}_2 in $SO(3)$ it holds that their Euclidean distance $d_E(\mathbf{R}_1, \mathbf{R}_2) = 2\sqrt{2} \sin[d_R(\mathbf{R}_1, \mathbf{R}_2)/2]$. We refer to Section B of the Appendix for a short proof of this result. Hence, results using d_E would prove equivalent, albeit on a smaller scale.

Figure 4 displays boxplots of the estimation errors for each of the considered rotation distributions and choice of circular spread ν for a sample of size $n = 100$. Additionally, Table 8 provides the root mean square error (RMSE) as a measure of precision as well as the

mean estimation error for $\nu = 0.25$ in Figure 4, i.e. the top three plots. Despite skewness in some of the plotted error distributions the *median estimation error* was quantitatively similar to the mean estimation error and is therefore not reported in Table 8 or the following.

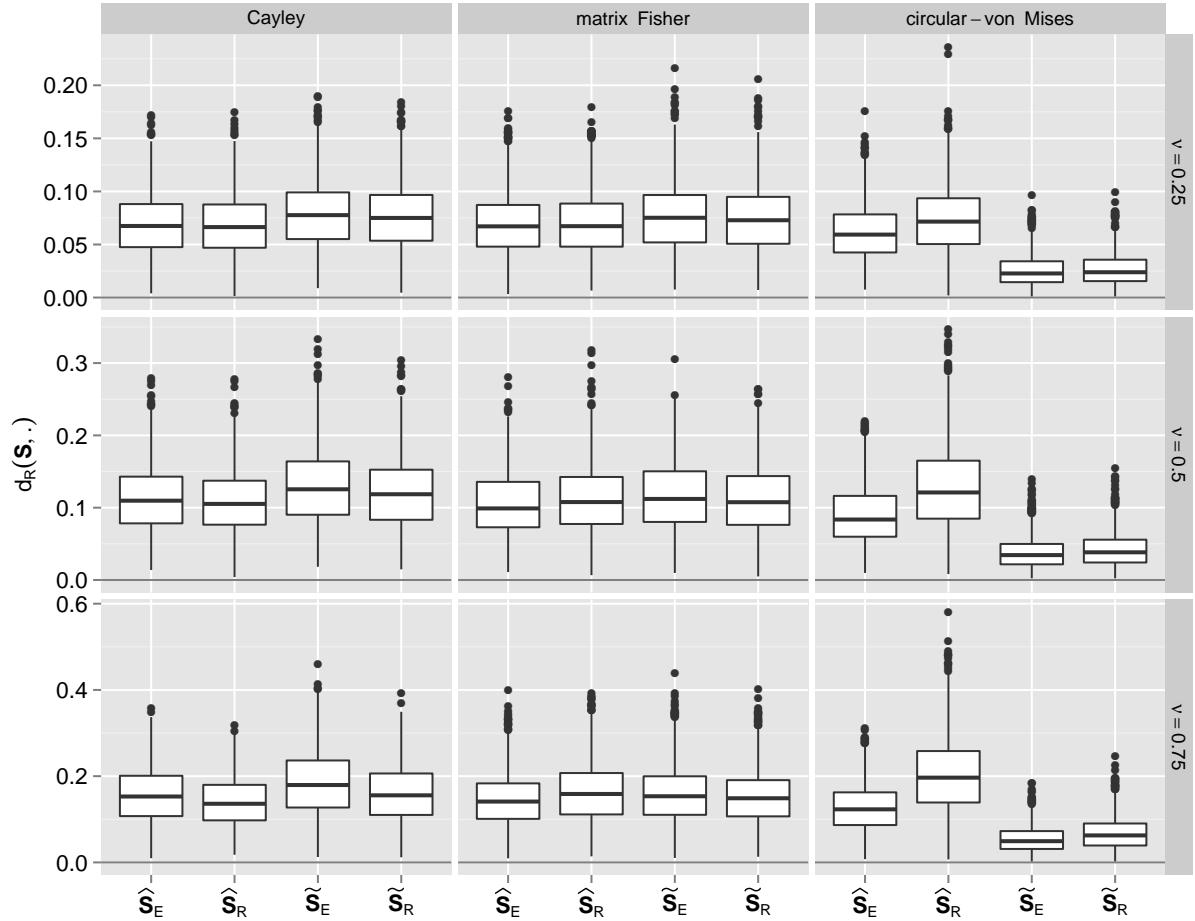


Figure 4: Boxplots of the estimation errors for each rotation distribution and level of ν , $n = 100$.

First and foremost the results suggest that different location estimators emerge as preferable depending on the type of distribution for the rotation errors in (1). For the circular-von Mises-based distribution both median-type estimators (\tilde{S}_E and \tilde{S}_R) are superior with respect to the estimation error while for the Cayley and matrix Fisher models the mean-type estimators (\hat{S}_E and \hat{S}_R) perform better though on a much less pronounced scale.

While preferences within the median- and mean-type estimators are visible, these generally disappear as the variability in the data, i.e. ν decreases. For the Cayley and the matrix Fisher distribution the overall pattern of estimation is very similar. $\widehat{\mathbf{S}}_E$ and $\widehat{\mathbf{S}}_R$ typically exhibit a little less spread and a slightly lower average estimation error than $\widetilde{\mathbf{S}}_E$ and $\widetilde{\mathbf{S}}_R$ but differences between all four estimators lessen as ν becomes smaller. Figure 4 further shows that the estimation error is a function of the circular spread ν ; as ν decreases the range of the observed estimation errors decreases within each rotation model and for each of the four estimators. The same is true for the mean estimation error, and RMSE in Table 8. For a plot similar to Figure 4 in which the estimation errors are given as a function of sample size for $\nu = 0.75$ refer to the Appendix.

Table 4: Numerical summaries of the estimation error by distribution, $n = 100$, $\nu = 0.25$.

Estimator	Cayley		matrix Fisher		circular-von Mises	
	Mean error	RMSE	Mean error	RMSE	Mean error	RMSE
$\widehat{\mathbf{S}}_R$	0.069	0.075	0.070	0.076	0.074	0.081
$\widehat{\mathbf{S}}_E$	0.070	0.076	0.070	0.076	0.062	0.067
$\widetilde{\mathbf{S}}_R$	0.077	0.083	0.075	0.081	0.027	0.031
$\widetilde{\mathbf{S}}_E$	0.079	0.086	0.077	0.083	0.026	0.030

Figure 2c may provide an explanation for why the circular-von Mises-based distribution so clearly distinguishes between the mean- and median-type estimators. Upon closer examination of the tail of all three rotation densities (expressed in terms of the misorientation angle r as in Figure 2) the circular-von Mises-based distribution exhibits the heaviest tail with respect to the Haar measure, see Figure 2c. As a consequence, *more extreme* observations are more likely to be sampled under the circular-von Mises-based distribution compared to the Cayley and matrix Fisher models suggesting that a median-type estimator is more favorable.

We use Figure 7 to examine the extent to which the tail-behavior accounts for the observed differences in the mean- and median-type estimators. Figure 7 displays the tail weight for each sample (i.e. the proportion of observations in the sample considered to come from the tail of the distribution) plotted against the difference in errors for the mean- and

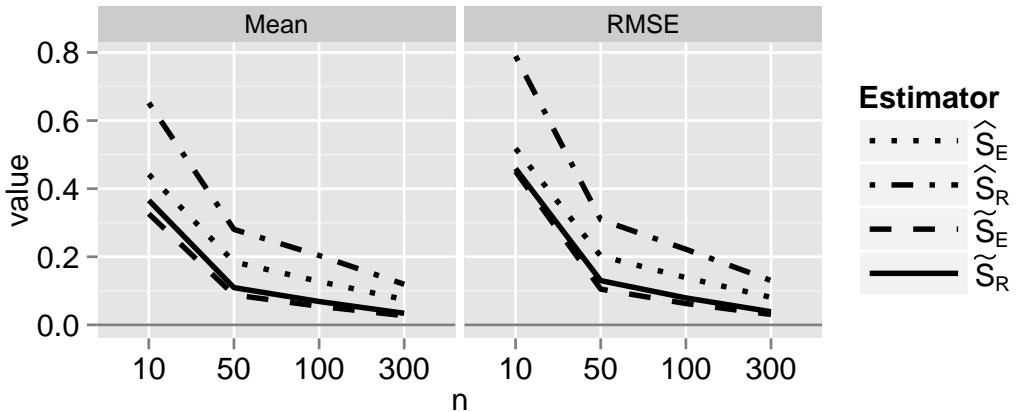


Figure 5: Plot of the estimation error for all levels of n for the circular-von Mises-based distribution, $\nu = 0.75$.

median-type estimators. The results shown in Figure 7 are with respect to the Euclidean geometry-based estimators \widehat{S}_E and \widetilde{S}_E . Similar results are obtained for the Riemannian geometry-based estimators \widehat{S}_R and \widetilde{S}_R and therefore are omitted. Note that we define the tail to be the average crossing point at which the distributions cross for the second time, see Figure 2c. From Figure 7 we can see that in general there appears to be a positive relationship between the proportion of observations in the tail and the relative difference in estimator error indicating that with increasing tail weight the error of the mean estimator indeed increases at a higher rate than does the error of the median estimator. That is, as the tail becomes heavier the projected median is preferable to the projected mean more often. Values on the y -axis not only indicate which estimator is preferable, but also by how much. For example, a y -value of 0 suggests that the two estimators result in equivalent estimates in terms of their distance from the true \mathbf{S} . The more extreme a y -value is in the negative direction, e.g. close to -0.5 and beyond, the closer the projected-mean is to the truth in comparison to the projected median. At -0.5 the error associated with the projected median is roughly three times that of the projected mean. Likewise, large positive y -values around 1 signify that the error by the projected median is roughly zero compared to the error of the projected median.

We next explore the effect the choice of geometry, i.e. Riemannian vs. Euclidean, has on

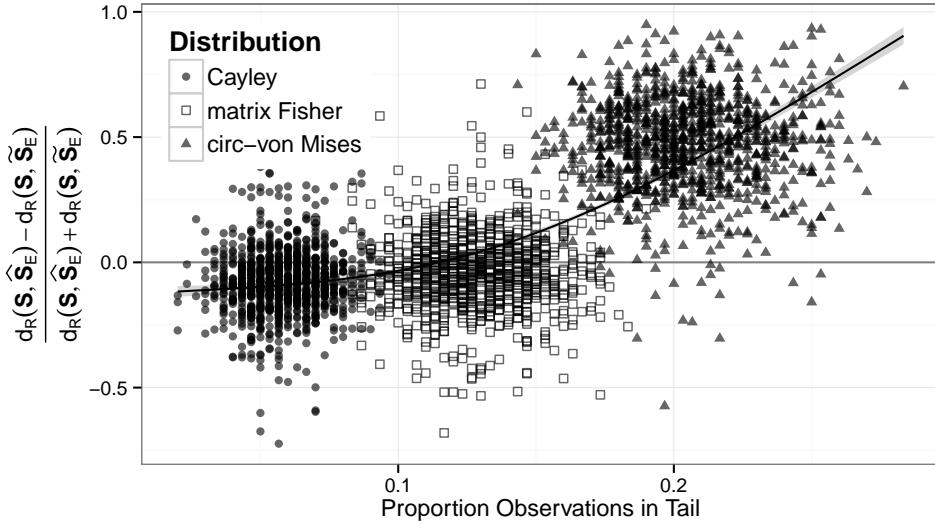


Figure 6: The proportion of observations in the tail against the difference in projected mean and median errors for simulated data with $n = 300$. Different symbols indicate different error distributions.

the estimation error for both types of loss functions. To provide more insight into possible differences we plot the estimation error resulting from d_E (x -axis) versus the estimation error resulting from d_R (y -axis) for $n = 100$ and $\nu = .25, \nu = .75$, respectively; see Figures 8 (a) and (b). Mean-type estimators are represented by black dots whereas median-type estimators correspond to light gray dots. A point falls below the 45° line if the error on the y -axis corresponding to the Riemannian geometry-based estimator is smaller in comparison to the error plotted on the x -axis for the Euclidean geometry-based estimator. The same holds true vice versa if a point falls above the 45° line. For example, in Figure 8, \tilde{S}_R tends to yield less estimation error than \tilde{S}_E for the Cayley distribution as most of the points fall below the identity line while the Riemannian distance-based \tilde{S}_R results in greater errors for \mathbf{S} for the circular-von Mises-based distribution. This result supports results for \tilde{S}_E and \tilde{S}_R seen in Figure 4.

In Tables 9 and 10 we support Figure 8 with an exact count (expressed as a percentage) of how often d_R resulted in a smaller estimation error than d_E . Additionally, we show the average amount by which the d_R - and d_E -based estimates deviate from one another. We denote the latter quantity by $\bar{\delta}$ in Table 9 where $\delta = d_R(\tilde{S}_E, \mathbf{S}) - d_R(\tilde{S}_R, \mathbf{S})$. Our previous results suggested the use of median-type estimators for the circular-von Mises-

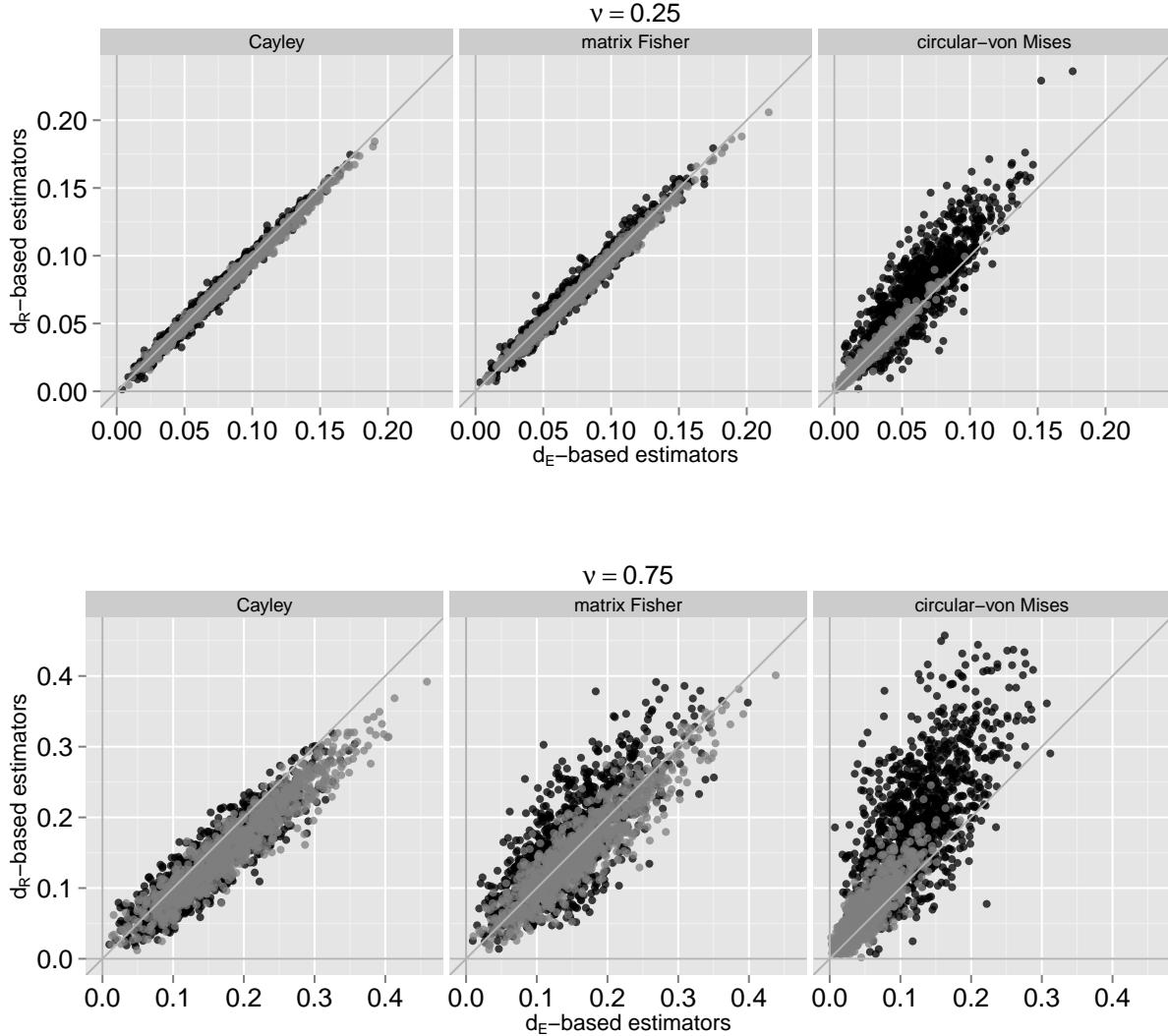


Figure 7: Comparison of the estimation errors resulting from the embedding (x -axis) and intrinsic (y -axis) approaches, $n = 100$. The mean-type estimators are in black dots and the median-type estimators are in light gray.

based distribution which favors $\tilde{\mathbf{S}}_E$ over $\tilde{\mathbf{S}}_R$ as $d_R(\tilde{\mathbf{S}}_E, \mathbf{S}) < d_R(\tilde{\mathbf{S}}_R, \mathbf{S})$ most of the time. For the Cayley and the matrix Fisher model the preference is reversed, typically $\tilde{\mathbf{S}}_E$ exhibits a larger spread (cf. Figure 5).

Similarly to the median-type estimators, d_R is the preferred metric for the Cayley distribution especially when ν is large. For the matrix Fisher distribution the preference is less clear, especially for less variable data, but as ν increases the Euclidean-based mean yields generally a smaller estimation error more often. In summary,

Table 5: Average reduction in estimation error by using $\tilde{\mathbf{S}}_R$ instead of $\tilde{\mathbf{S}}_E$, $\delta = d_R(\tilde{\mathbf{S}}_E, \mathbf{S}) - d_R(\tilde{\mathbf{S}}_R, \mathbf{S})$ and percentage of samples for which $d_R(\tilde{\mathbf{S}}_R, \mathbf{S}) < d_R(\tilde{\mathbf{S}}_E, \mathbf{S})$.

		Cayley		matrix Fisher		circular-von Mises	
	n	$\bar{\delta}$	%	$\bar{\delta}$	%	$\bar{\delta}$	%
$\nu = 0.25$	10	0.008	0.743	0.006	0.725	-0.005	0.328
	50	0.003	0.783	0.002	0.697	-0.002	0.327
	100	0.002	0.789	0.002	0.712	-0.001	0.308
	300	0.001	0.781	0.001	0.711	-0.001	0.284
$\nu = 0.50$	10	0.031	0.772	0.017	0.662	-0.019	0.335
	50	0.013	0.811	0.005	0.620	-0.008	0.282
	100	0.009	0.809	0.005	0.660	-0.005	0.302
	300	0.005	0.804	0.002	0.658	-0.003	0.255
$\nu = 0.75$	10	0.089	0.821	0.034	0.633	-0.040	0.322
	50	0.037	0.866	0.009	0.597	-0.021	0.238
	100	0.025	0.858	0.007	0.603	-0.014	0.240
	300	0.014	0.850	0.003	0.589	-0.007	0.218

Table 6: Average reduction in estimation error by using $\hat{\mathbf{S}}_R$ instead of $\hat{\mathbf{S}}_E$, $\delta = d_R(\hat{\mathbf{S}}_E, \mathbf{S}) - d_R(\hat{\mathbf{S}}_R, \mathbf{S})$ and percentage of samples for which $d_R(\hat{\mathbf{S}}_R, \mathbf{S}) < d_R(\hat{\mathbf{S}}_E, \mathbf{S})$.

		Cayley		matrix Fisher		circular-von Mises	
	n	$\bar{\delta}$	%	$\bar{\delta}$	%	$\bar{\delta}$	%
$\nu = 0.25$	10	0.001	0.521	-0.002	0.450	-0.034	0.128
	50	0.001	0.531	-0.001	0.435	-0.016	0.209
	100	0.001	0.565	-0.000	0.469	-0.013	0.201
	300	0.000	0.588	-0.000	0.486	-0.007	0.239
$\nu = 0.50$	10	0.010	0.592	-0.018	0.434	-0.101	0.157
	50	0.007	0.645	-0.011	0.392	-0.055	0.145
	100	0.004	0.642	-0.007	0.393	-0.038	0.162
	300	0.003	0.642	-0.004	0.393	-0.023	0.157
$\nu = 0.75$	10	0.016	0.668	-0.096	0.357	-0.210	0.171
	50	0.026	0.741	-0.036	0.338	-0.096	0.150
	100	0.017	0.735	-0.024	0.346	-0.076	0.111
	300	0.009	0.728	-0.012	0.344	-0.045	0.127

- the choice of location estimator can depend on the rotation error distribution in the location model (1). For the matrix Fisher and the Cayley distribution the projected arithmetic mean $\widehat{\mathbf{S}}_E$ and the geometric mean $\widehat{\mathbf{S}}_R$ are, respectively, preferable though $\widetilde{\mathbf{S}}_E$ and $\widetilde{\mathbf{S}}_R$ are not far behind especially when the circular spread is smaller. For the circular-von Mises-based distribution the projected median $\widetilde{\mathbf{S}}_E$ should be used.
- However, a significant finding of these simulation results is that the (Euclidean-based) projected median $\widetilde{\mathbf{S}}_E$ is a generally good location estimator across rotation error models. For the circular-von Mises-based estimation, this generally has the best performance, while for the Cayley or matrix Fisher distributions, this estimator is often quite comparable to the best estimator. In other words, an estimator $\widetilde{\mathbf{S}}_E$ not previously considered for rotation matrices in the literature appears to be generally suggestible, particularly in small samples and without knowledge of the underlying rotation error distribution.

6 Data Application

The data in our example represent the orientations of cubic crystals on a Nickel surface measured by electron backscatter diffraction. The data were obtained using a 14-fold technical replicate in each location. A central interest in EBSD data is the identification of so-called *grain maps* – grains are defined as regions on the surface with nearly identical main direction \mathbf{S} . This makes the estimation of \mathbf{S} crucial to the field.

The left hand side of Figure 9 shows a map of the surface colored by angle between projected median $\widetilde{\mathbf{S}}_E$ and the identity in each location. Distinct spatial structures become apparent. After applying all four of our estimators to each location on the surface, we noticed large differences between the estimates in some of the locations. On the right hand side of Figure 9 the differences in angles between projected median and mean estimate for each location are shown. Largest differences are along boundaries between structures visible on the left of this figure.

Figure 10 shows a parallel coordinate plot of the data of one of these locations overlaid by the corresponding estimates: each coefficient is plotted on a vertical axis corresponding to

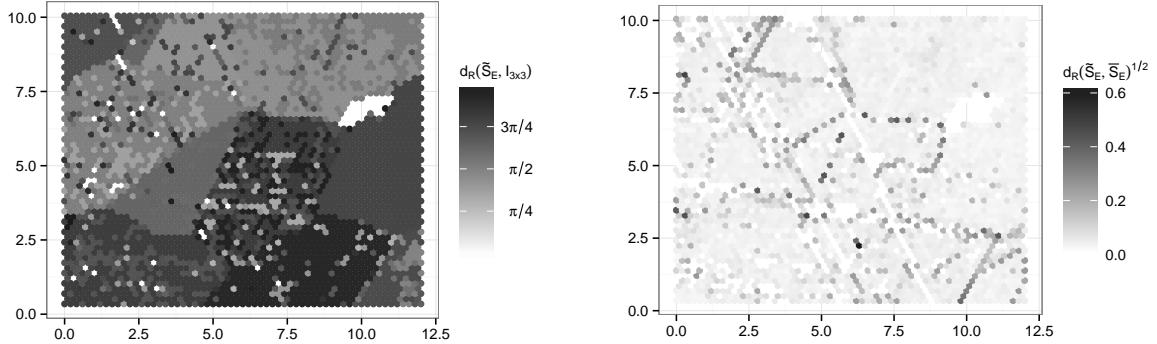


Figure 8: Grain map of the investigated Nickel surface (left). Each dot corresponds to one location, shading shows the angle of the projected median estimate $\tilde{\mathbf{S}}_E$ and the identity. On the right, differences between the projected median and mean estimates for each location are shown. Largest differences fall on the boundaries between grains.

its value in $[-1,1]$. Coefficients of the same matrix are connected by a line. Three clusters of rotation matrices become apparent. Note that the values are jittered – we used a small perturbation in form of a rotation matrix with a small mis-orientation angle with respect to the identity to draw the rotation matrices in each cluster apart slightly in order to avoid over-plotting and demonstrate cluster sizes. The lines for the rotations are overlaid by the four estimates of main rotations (in color).

Note that the figure shows a sample of a single location, yet the clustering in the data suggests clearly the presence of *several* main directions. The same holds true for locations in the vicinity which happen to coincide with the boundaries between grains.

The median estimates, $\tilde{\mathbf{S}}_E$ and $\tilde{\mathbf{S}}_R$, are not nearly as much affected by the clustering in the sample as are the means, and reliably estimate the main direction of the largest group of rotations within the sample.

7 Recommendations and Conclusions

To estimate the central orientation \mathbf{S} given a random sample of three-dimensional orientations from (1) the scientific literature suggests a variety of approaches. These approaches differ largely with respect to the geometry (Riemannian vs. Euclidean) in which the es-

Table 7: List of all rotations in a single location exhibiting a large difference between mean and median estimators. We observe two tight clusters of rotations and one outlying rotation.

x_{11}	x_{12}	x_{13}	x_{21}	x_{22}	x_{23}	x_{31}	x_{32}	x_{33}
-0.999	-0.000	-0.043	0.043	0.077	-0.996	0.004	-0.997	0.077
-0.999	-0.006	-0.047	0.047	0.079	-0.996	0.009	-0.997	0.079
-0.999	-0.005	-0.044	0.044	0.080	-0.996	0.009	-0.997	0.080
-0.999	-0.005	-0.047	0.047	0.079	-0.996	0.009	-0.997	0.078
-0.999	-0.004	-0.044	0.044	0.079	-0.996	0.007	-0.997	0.078
-0.999	-0.008	-0.048	0.047	0.079	-0.996	0.011	-0.997	0.079
-0.998	-0.008	-0.056	0.055	0.076	-0.996	0.012	-0.997	0.075
-0.998	-0.010	-0.054	0.054	0.073	-0.996	0.014	-0.997	0.072
-0.871	-0.061	-0.487	0.446	0.510	-0.735	0.203	-0.858	0.472
-0.874	-0.059	-0.483	0.441	0.513	-0.737	0.204	-0.857	0.474
-0.872	-0.053	-0.487	0.442	0.516	-0.734	0.212	-0.855	0.473
-0.870	-0.060	-0.489	0.447	0.512	-0.733	0.207	-0.857	0.472
-0.874	-0.066	-0.482	0.445	0.510	-0.736	0.198	-0.858	0.475
-0.708	-0.315	-0.632	0.609	0.181	-0.773	0.358	-0.932	0.063

timation is done, assumptions about the underlying data-generating mechanism and the choice of loss function when defining suitable estimators. The main goal of this paper was to explore the extent to which the above mentioned differences affect the estimation of \mathbf{S} . Our simulation study showed that the choice of loss function depends on the underlying data-generating mechanism. For the circular-von Mises-based model median-type estimators are recommended while for the Cayley and matrix Fisher model the mean-type estimators show less estimation error and variability. As noted in Section 1 the applied sciences generally pursue estimation of \mathbf{S} without considering the distributional underpinnings. This can be a pitfall. Restricting ourselves to the three rotation distributions under consideration, if indeed nothing is known about the underlying data-generating mechanism we suggest to use either median-type estimator, i.e. $\tilde{\mathbf{S}}_E$ or $\tilde{\mathbf{S}}_R$. The overall estimation error, even under mis-specification will be much less than the potential estimation error resulting from either mean-type estimator. The correct choice of geometry also depends on the underlying distributional model, however, the effects are not as large. The Riemannian distance metric d_R is generally not to recommend with a mean-type estimator due to the

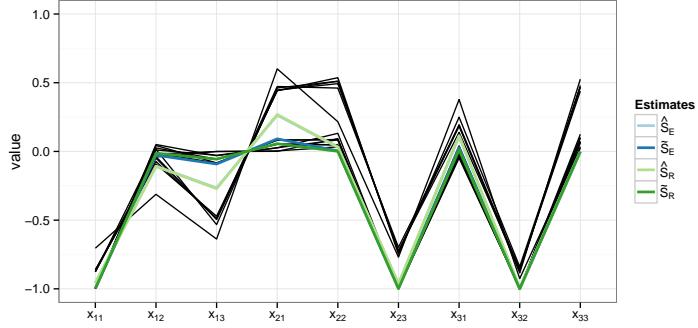


Figure 9: Parallel coordinate plot of all nine coefficients of the rotation matrices in a single location. Three clusters of rotation matrices are apparent.

resulting increase in variability in the estimation errors. This fact holds across all three distributions and is much expressed than for any of the other three estimators. Lastly, the extent to which all four estimators disagree relative to each other depend on the circular variance ν ; the estimators differ more when ν is large and tend to agree more and more as the data become more concentrated.

The scope of this study can be extended to other frequently encountered distributional model, especially because we restricted the simulation to symmetric perturbations around \mathbf{S} in (1). At least as important, if not even more is the extension of the studied point estimators to interval estimators. The latter requires a significant improvement in computing time before it can be practically implemented. Both of these tasks will be pursued by the authors in the near future.

Acknowledgements: We authors wish to thank ... for collecting and providing the EBSD data.

References

- K. Arun, T. Huang, and S. Blostein. Least-squares fitting of two 3-d point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):698–700, 1987.
- C. Bajaj. The algebraic degree of geometric optimization problems. *Discrete & Computational Geometry*, 3(1):177–191, 1988.
- D. Best and N. Fisher. Efficient simulation of the von Mises distribution. *Applied Statistics*, 28(2):152–157, 1979. ISSN 0035-9254.
- M. Bingham, D. Nordman, and S. Vardeman. Modeling and inference for measured crystal orientations and a tractable class of symmetric distributions for rotations in three dimensions. *Journal of the American Statistical Association*, 104(488):1385–1397, 2009.
- M. Bingham, D. Nordman, and S. Vardeman. Finite-sample investigation of likelihood and Bayes inference for the symmetric von Mises-Fisher distribution. *Computational Statistics & Data Analysis*, 54(5):1317–1327, 2010.
- H. Bunge. *Texture Analysis in Material Science*. Butterworth, London, 1982.
- Y. Chan and X. He. On median-type estimators of direction for the von Mises-Fisher distribution. *Biometrika*, 80(4):869–875, 1993.
- Y. Chikuse. *Statistics on Special Manifolds*. Springer Verlag, 2003. ISBN 0387001603.
- Y. Dai, J. Trumpf, H. Li, N. Barnes, and R. Hartley. Rotation averaging with application to camera-rig calibration. *Computer Vision–ACCV 2009*, pages 335–346, 2010.
- T. Downs. Orientation statistics. *Biometrika*, 59(3):665–676, 1972.
- G. Ducharme and P. Milasevic. Spatial median and directional data. *Biometrika*, 74(1):212–215, 1987.
- S. Durocher and D. Kirkpatrick. The projection median of a set of points. *Computational Geometry*, 42(5):364–375, 2009.

- N. Fisher. Spherical Medians. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 342–348, 1985.
- R. Fisher. Dispersion on a Sphere. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, pages 295–305, 1953. ISSN 0080-4630.
- P. Fletcher, S. Venkatasubramanian, and S. Joshi. Robust Statistics on Riemannian Manifolds via the Geometric Median. In *2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8. IEEE, 2008.
- P. Fletcher, S. Venkatasubramanian, and S. Joshi. The geometric median on riemannian manifolds with application to robust atlas estimation. *NeuroImage*, 45(1):S143–S152, 2009.
- J. Gower. The mediancentre. *Applied Statistics*, 23(3):466–470, 1974.
- R. Hadani and A. Singer. Representation theoretic patterns in three-dimensional cryo-electron microscopy ii – the class averaging problem. *Foundations of Computational Mathematics*, pages 1–28, 2011.
- R. Hartley, K. Aftab, and J. Trumpf. L1 rotation averaging using the Weiszfeld algorithm. In *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3041–3048. IEEE, 2011.
- X. He and D. Simpson. Robust Direction Estimation. *The Annals of Statistics*, 20(1):351–369, 199. ISSN 0090-5364.
- R. Hielscher, H. Schaeben, and S. H. Orientation distribution within a single hematite crystal. *Mathematical Geosciences*, 42:359–375, 2010.
- B. Horn, H. Hilden, and S. Negahdaripour. Closed-form solution of absolute orientation using orthonormal matrices. *Journal of the Optical Society of America*, 5(7):1127–1135, 1988.
- P. Jupp and K. Mardia. Maximum likelihood estimators for the matrix von Mises-Fisher and Bingham distributions. *The Annals of Statistics*, 7(3):599–606, 1979. ISSN 0090-5364.

- P. Jupp and K. Mardia. A unified view of the theory of directional statistics. *International Statistical Review*, 57(3):261–294, 1989.
- H. Karcher. Riemannian center of mass and mollifier smoothing. *Communications on pure and applied mathematics*, 30(5):509–541, 1977.
- C. Khatri and K. Mardia. The von Mises-Fisher matrix distribution in orientation statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):95–106, 1977. ISSN 0035-9246.
- P. Langevin. Magnetism and the theory of the electron. *Annales de chimie et de physique*, 5:70, 1905.
- C. León, J. Massé, and L. Rivest. A statistical model for random rotations. *Journal of Multivariate Analysis*, 97(2):412–430, 2006.
- R. Liu and K. Singh. Ordering directional data: concepts of data depth on circles and spheres. *The Annals of Statistics*, pages 1468–1484, 1992. ISSN 0090-5364.
- J. Manton. A globally convergent numerical algorithm for computing the centre of mass on compact lie groups. In *8th Conference on Control, Automation, Robotics and Vision, (ICARCV)*, volume 3, pages 2211–2216. IEEE, 2004.
- K. Mardia. *Statistics of Directional Data*. London: Academic Press, 1972.
- K. Mardia and P. Jupp. *Directional Statistics*. Wiley Chichester, 2000.
- S. Matthies, J. Muller, and G. Vinel. On the Normal Distribution in the Orientation Space. *Textures and Microstructures*, 10:77–96, 1988.
- M. Moakher. Means and averaging in the group of rotations. *SIAM Journal on Matrix Analysis and Applications*, 24(1):1–16, 2002.
- M. Pierrynowski and K. Ball. Oppugning the assumptions of spatial averaging of segment and joint orientations. *Journal of biomechanics*, 42(3):375–378, 2009.

- D. Rancourt, L. Rivest, and J. Asselin. Using orientation statistics to investigate variations in human kinematics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 49(1):81–94, 2000. ISSN 1467-9876.
- V. Randle. *Microtexture Determination and its Applications*. London: Maney for The Institute of Materials, Minerals and Mining, 2003.
- T. Savyolova and D. Nikolayev. Normal Distribution on the Rotation Group SO(3). *Textures and Microstructures*, 29:201–233, 1995.
- H. Schaeben. A simple standard orientation density function: The hyperspherical de la vallée poussin kernel. *Phys. Stat. Sol. (B)*, 200:367–376, 1997.
- M. Schwartz and A. Rozumalski. A new method for estimating joint parameters from motion data. *Journal of Biomechanics*, 38(1):107–116, 2005.
- S. Umeyama. Least squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:376–380, 1991.
- G. Watson. Statistics on spheres, university of arkansas lecture notes in the mathematical sciences, vol. 6, 1983.
- E. Weiszfeld. Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tohoku Mathematics Journal*, 43:355–386, 1937.

A Sampling Process

A.1 Circular-von Mises-based distribution

To simulate a set of random rotations from the circular-von Mises-based distribution we follow the algorithm proposed by Best and Fisher (1979). The algorithm is available in the IMSL Library (1991) and is implemented as follows. Let $\mu = 0$ denote the mean of the target angular distribution and κ its concentration parameter. We define constants a, b and d as $a \equiv 1 + \sqrt{1 + 4\kappa^2}$, $b \equiv (a - \sqrt{2a})$, $d \equiv (1 + b^2)/2b$. In steps one, two and four we generate three new observations u_1, u_2 and u_3 , each from a uniform distribution defined over the interval $(0, 1)$.

1. Set $z = \cos(\pi u_1)$, $f = (1 + dz)/(z + d)$ and $c = \kappa(d - f)$.
2. If $c(2 - c) - u_2 > 0$ go to step 4.
3. If $\log(c/u_2) + 1 - c < 0$ return to step 1.
4. Set $r = \text{sign}(u_3 - 0.5) \cos^{-1}(f)$
5. Then r follows a von Mises (κ) distribution.

A.2 Cayley distribution

To simulate rotation matrices from a Cayley distribution we make use of a result given in León et al. (2006). If the angle r follows a Cayley distribution it holds that $\frac{1+\cos r}{2} \sim \text{Beta}(\kappa + 1/2, 3/2)$. Hence, angles according to a Cayley distribution can be simulated through composition: we simulate a Bernoulli trial Y with outcomes -1 and 1 having probability 0.5 and an observation X from a $\text{Beta}(\kappa + 1/2, 3/2)$ distribution and then set $r = \frac{Y}{2} \cos^{-1}(2X - 1)$.

A.3 matrix Fisher distribution

Simulation from the matrix Fisher distribution is achieved through a rejection algorithm. Let $C_F(r|\kappa)$ denote the matrix von Mises-Fisher density as given in Table 2 and $Y \sim \text{Uniform}(-\pi, \pi]$.

1. Define $M = \frac{1}{2\kappa} e^{2\kappa-1} \frac{1}{I_0(2\kappa) - I_1(2\kappa)}$
2. Generate $U \sim \text{Uniform}(0, 1)$ and $Y \sim \text{Uniform}(-\pi, \pi]$, where U and Y are independent.
3. If $U < \frac{1}{M} C_F(Y|\kappa)$, accept Y ; otherwise return to step (2)

Given a set of randomly generated angles r_1, \dots, r_n we can now generate the corresponding set of rotation matrices as follows:

1. Generate a point uniformly on the unit sphere

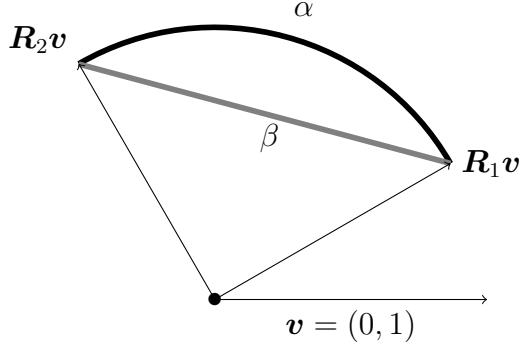
$$\mathbf{U} = (u_1, u_2, u_3)^\top = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)^\top$$

where $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$.

2. Given an angle of rotation, r_i , generated as described above from an angular distribution symmetric about 0 and with concentration κ rotate \mathbf{I} about \mathbf{U} by r_i radians.

A.4 Comparison of Euclidean and Riemannian Metrics When Measuring Distance to the Central Direction

The results section of this paper only compared the distances of each estimator from the central direction in terms of Riemannian distance. This section is intended to show that a comparison based on the Euclidean distance will not lead to different results. Recall that the Riemannian distance between two rotations is the shortest geodesic curve that connects those two rotations and the Euclidean distance is the shortest cord that connects those two rotations. Refer to the picture below for a two-dimensional simplification of the idea.



B No Name yet

Let $\mathbf{R}_1\mathbf{v}$ and $\mathbf{R}_2\mathbf{v}$ be two observations in $SO(2)$. The Riemannian distance between these two observations is given by α and is indicated with the thick black arc. The Euclidean distance is given by β and is indicated by the gray line running through the circle. Because this is the unit circle, the Riemannian distance is also the angle in the center of the circle thus, using basic geometry its clear that half of the Euclidean distance is the sine of half of the Riemannian distance, i.e.

$$\beta = 2 \sin\left(\frac{\alpha}{2}\right).$$

Extending this to $SO(3)$ gives the following proposition: for all \mathbf{R}_1 and \mathbf{R}_2 in $SO(3)$

$$d_E(\mathbf{R}_1, \mathbf{R}_2) = 2^{3/2} \sin\left(\frac{d_R(\mathbf{R}_1, \mathbf{R}_2)}{2}\right). \quad (10)$$

Proof:

Let $\mathbf{o}_1, \mathbf{R}_2 \in SO(3)$ be given and define $\text{tr}(\mathbf{R}_1^\top \mathbf{R}_2) = 1 + 2 \cos(\theta)$ then $|\theta| = d_R(\mathbf{R}_1, \mathbf{R}_2)$. Notice that (10) holds trivially for $d_E(\mathbf{R}_1, \mathbf{R}_2) = d_R(\mathbf{R}_1, \mathbf{R}_2) = 0$ so consider the case

$|\theta| > 0$. By definition of $d_E(\mathbf{R}_1, \mathbf{R}_2)$ we have the following:

$$\begin{aligned}
d_E(\mathbf{R}_1, \mathbf{R}_2)^2 &= \|\mathbf{R}_1 - \mathbf{R}_2\|_F^2 \\
&= \text{tr} [(\mathbf{R}_1 - \mathbf{R}_2)^\top (\mathbf{R}_1 - \mathbf{R}_2)] = \text{tr} [(\mathbf{R}_1^\top - \mathbf{R}_2^\top)(\mathbf{R}_1 - \mathbf{R}_2)] \\
&= \text{tr} [\mathbf{R}_1^\top \mathbf{R}_1 + \mathbf{R}_2^\top \mathbf{R}_2 - \mathbf{R}_2^\top \mathbf{R}_1 - \mathbf{R}_1^\top \mathbf{R}_2] \\
&= \text{tr} [2\mathbf{I} - \mathbf{R}_2^\top \mathbf{R}_1 - \mathbf{R}_1^\top \mathbf{R}_2] \\
&= 2\text{tr}(\mathbf{I}) - \text{tr}(\mathbf{R}_2^\top \mathbf{R}_1) - \text{tr}(\mathbf{R}_1^\top \mathbf{R}_2) \\
&= 6 - 2\text{tr}(\mathbf{R}_1^\top \mathbf{R}_2) = 6 - 2(1 + 2\cos(\theta)) \\
&= 6 - 2 - 4\cos(\theta) = 8 \left(\frac{1 - \cos(\theta)}{2} \right) \\
&= 8 \sin^2 \left(\frac{\theta}{2} \right) \\
&= \left[2^{3/2} \sin \left(\frac{|\theta|}{2} \right) \right]^2 \\
&= \left[2^{3/2} \sin \left(\frac{d_R(\mathbf{R}_1, \mathbf{R}_2)}{2} \right) \right]^2
\end{aligned}$$

Taking square roots on both sides and noticing that $d_E(\mathbf{R}_1, \mathbf{R}_2) \geq 0$ and $d_R(\mathbf{R}_1, \mathbf{R}_2) \geq 0$ gives (10).

□

B.1 Additional Results

Figure 5 illustrates the behavior of the estimators as a function of the sample size. Results are displayed for each level of n for $\nu = 0.75$. As to be expected, the estimation error decreases as the sample size increases for all three distributions. For small samples, $n = 10$, the estimator exhibiting the largest amount of variability is the geodesic mean $\hat{\mathbf{S}}_R$. This behavior is consistent for all three distributions. While the estimator's variability lessens considerably for the Cayley and matrix Fisher distribution as n increases, the estimator remains the most variable estimator for the circular-von Mises-based distribution. A possible explanation for this behavior is that the algorithm to estimate $\hat{\mathbf{S}}_R$ uses a random sample point in its initiating step. For small samples or samples with several extreme observations it is likely to start the algorithm far from the true center, which in turn may cause the

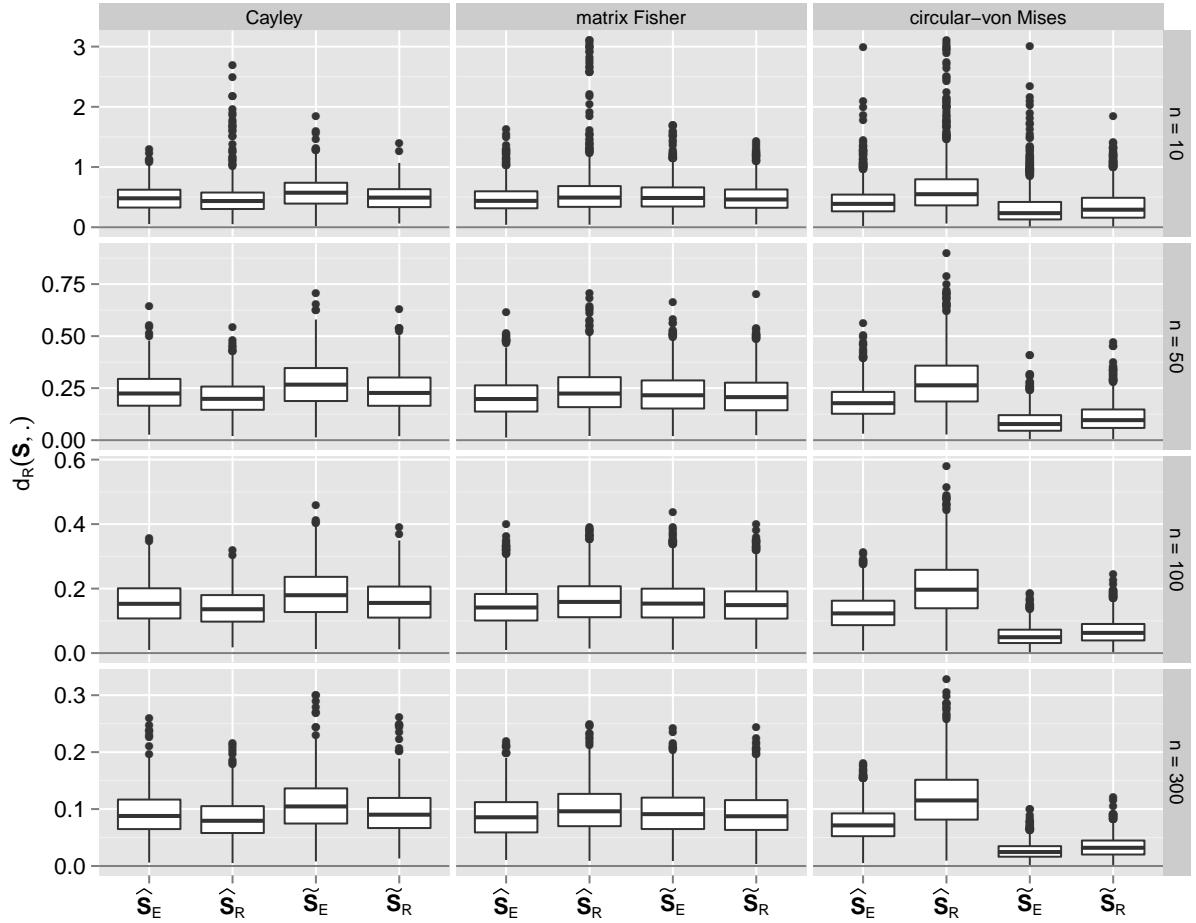


Figure 10: Box-plots of the estimation error for each rotation distribution and level of n , $\nu = 0.75$.

algorithm to get stuck in a local minimum and to fail to converge globally. In practice we suggest the algorithm be started at some other location estimate of \mathbf{S} , e.g. at $\hat{\mathbf{S}}_E$. In simulations where $\hat{\mathbf{S}}_E$ was used as a starting point for the algorithm we observed similar results with less variability in the estimates of $\hat{\mathbf{S}}_R$ but for the purpose of a fair comparison we started the algorithm at a random sample point. We refer to Figure 6 for a comparison of all estimators for the circular-von Mises-based distribution case when $\nu = 0.75$.

Table 8: Numerical summaries of the estimation error by distribution, $n = 100$, $\nu = 0.25$. The mean error (standard deviation) and RMSE are reported here.

Estimator	Cayley		matrix Fisher		circular-von Mises	
	Mean error	RMSE	Mean error	RMSE	Mean error	RMSE
$\widehat{\mathbf{S}}_R$	0.069 (0.0009)	0.075	0.070 (0.0009)	0.076	0.074 (0.0010)	0.081
$\widehat{\mathbf{S}}_E$	0.070 (0.0009)	0.076	0.070 (0.0009)	0.076	0.062 (0.0008)	0.067
$\widetilde{\mathbf{S}}_R$	0.077 (0.0010)	0.083	0.075 (0.0010)	0.081	0.027 (0.0005)	0.031
$\widetilde{\mathbf{S}}_E$	0.079 (0.0010)	0.086	0.077 (0.0010)	0.083	0.026 (0.0005)	0.030

Table 9: Average reduction in estimation error by using $\widetilde{\mathbf{S}}_R$ instead of $\widetilde{\mathbf{S}}_E$, $\delta = d_R(\widetilde{\mathbf{S}}_E, \mathbf{S}) - d_R(\widetilde{\mathbf{S}}_R, \mathbf{S})$ with standard error and percentage of samples for which $d_R(\widetilde{\mathbf{S}}_R, \mathbf{S}) < d_R(\widetilde{\mathbf{S}}_E, \mathbf{S})$.

	n	Cayley				matrix Fisher				circular-von Mises		
		$\bar{\delta}$	SE($\bar{\delta}$)	%		$\bar{\delta}$	SE($\bar{\delta}$)	%		$\bar{\delta}$	SE($\bar{\delta}$)	%
$\nu = 0.25$	10	0.0078	0.0004	0.7430	0.0060	0.0004	0.7250	-0.0053	0.0006	0.3280		
	50	0.0031	0.0001	0.7830	0.0024	0.0001	0.6970	-0.0018	0.0001	0.3270		
	100	0.0022	0.0001	0.7890	0.0018	0.0001	0.7120	-0.0013	0.0001	0.3080		
	300	0.0012	0.0001	0.7810	0.0010	0.0001	0.7110	-0.0008	0.0000	0.2840		
$\nu = 0.50$	10	0.0315	0.0509	0.7720	0.0171	0.0531	0.6620	-0.0192	0.0617	0.3350		
	50	0.0126	0.0149	0.8110	0.0055	0.0177	0.6200	-0.0081	0.0144	0.2820		
	100	0.0085	0.0100	0.8090	0.0047	0.0120	0.6600	-0.0047	0.0087	0.3020		
	300	0.0049	0.0057	0.8040	0.0024	0.0066	0.6580	-0.0027	0.0043	0.2550		
$\nu = 0.75$	10	0.0315	0.0016	0.7720	0.0171	0.0017	0.6620	-0.0192	0.0020	0.3350		
	50	0.0126	0.0005	0.8110	0.0055	0.0006	0.6200	-0.0081	0.0005	0.2820		
	100	0.0085	0.0003	0.8090	0.0047	0.0004	0.6600	-0.0047	0.0003	0.3020		
	300	0.0049	0.0002	0.8040	0.0024	0.0002	0.6580	-0.0027	0.0001	0.2550		
	10	0.0895	0.0042	0.8210	0.0340	0.0040	0.6330	-0.0396	0.0062	0.3220		
	50	0.0366	0.0011	0.8660	0.0093	0.0013	0.5970	-0.0213	0.0011	0.2380		
	100	0.0250	0.0008	0.8580	0.0069	0.0009	0.6030	-0.0140	0.0007	0.2400		
	300	0.0140	0.0004	0.8500	0.0033	0.0005	0.5890	-0.0072	0.0003	0.2180		

Table 10: Average reduction in estimation error by using $\widehat{\mathbf{S}}_R$ instead of $\widehat{\mathbf{S}}_E$, $\delta = d_R(\widehat{\mathbf{S}}_E, \mathbf{S}) - d_R(\widehat{\mathbf{S}}_R, \mathbf{S})$ with standard error and percentage of samples for which $d_R(\widehat{\mathbf{S}}_R, \mathbf{S}) < d_R(\widehat{\mathbf{S}}_E, \mathbf{S})$.

	n	Cayley			matrix Fisher			circular-von Mises		
		$\bar{\delta}$	SE($\bar{\delta}$)	%	$\bar{\delta}$	SE($\bar{\delta}$)	%	$\bar{\delta}$	SE($\bar{\delta}$)	%
$\nu = 0.25$	10	0.0011	0.0004	0.5210	-0.0016	0.0005	0.4500	-0.0344	0.0030	0.1280
	50	0.0007	0.0002	0.5310	-0.0011	0.0003	0.4350	-0.0156	0.0007	0.2090
	100	0.0007	0.0001	0.5650	-0.0004	0.0002	0.4690	-0.0126	0.0005	0.2010
	300	0.0005	0.0001	0.5880	-0.0003	0.0001	0.4860	-0.0070	0.0003	0.2390
$\nu = 0.50$	10	0.0099	0.0013	0.5920	-0.0178	0.0027	0.4340	-0.1011	0.0051	0.1570
	50	0.0069	0.0006	0.6450	-0.0107	0.0011	0.3920	-0.0545	0.0018	0.1450
	100	0.0043	0.0004	0.6420	-0.0067	0.0008	0.3930	-0.0385	0.0013	0.1620
	300	0.0025	0.0002	0.6420	-0.0040	0.0004	0.3930	-0.0234	0.0007	0.1570
$\nu = 0.75$	10	0.0163	0.0062	0.6680	-0.0958	0.0105	0.3570	-0.2101	0.0113	0.1710
	50	0.0257	0.0012	0.7410	-0.0356	0.0045	0.3380	-0.0955	0.0032	0.1500
	100	0.0169	0.0009	0.7350	-0.0240	0.0043	0.3460	-0.0763	0.0021	0.1110
	300	0.0091	0.0005	0.7280	-0.0124	0.0009	0.3440	-0.0446	0.0012	0.1270

Table 11: Results from an analysis of variance on the log transformed error data. Each unique sample size (n), circular variance (ν) and distribution combination is treated as a fixed block. From this we can conclude that there is a statistically significant difference between the estimators.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Blocks	35	84515.71	2414.73	6805.73	$< 1 \times 10^{-5}$
Estimator	3	2714.88	904.96	2550.56	$< 1 \times 10^{-5}$
Residuals	143961	51078.63	0.35		
Total	143999	138309.2			

Table 12: Results from an analysis of variance on the log transformed error data for $n = 100$ and $\nu = 0.25$ and treating distributions as blocks. From this we can conclude that there is a statistically significant difference between the estimators.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Distribution	2	927.61	463.80	1371.56	$< 1 \times 10^{-5}$
Estimator	3	228.43	76.14	225.17	$< 1 \times 10^{-5}$
Residuals	11994	4055.85	0.34		
Total	11999	5211.88			