ON THE INDUCED DISTRIBUTION OF THE SHAPE OF THE PROJECTION OF A RANDOMLY ROTATED CONFIGURATION

H. LE,* *University of Nottingham*

D. BARDEN,** *University of Cambridge*

Abstract

Using the geometry of the Kendall shape space, in this paper we study the shape, as well as the size-and-shape, of the projection of a configuration after it has been rotated and, when the given configuration lies in a Euclidean space of an arbitrary dimension, we obtain expressions for the induced distributions of such shapes when the rotation is uniformly distributed.

Keywords: Kendall shape space; shape; size-and-shape; horizontal lift; Riemannian submersion

2010 Mathematics Subject Classification: Primary 60D05

1. Introduction

Since Kendall's seminal paper [1] on shape spaces over twenty years ago, many advances have been made in both theory and applications. A recent example is the study of the shape of the projection onto a fixed hyperplane of a given configuration after it is randomly rotated in [5] and [6] (see also [3]). As pointed out in the above papers, one motivation for this study is its importance in the field of biophysics, where, using electron microscopes, one views images of single particles in an aqueous environment in order to study the structure of biological molecules. Thus, one is viewing a planar projection of the particle rather than the three-dimensional particle itself. Moreover, the size of the particle, together with its movement, makes it impossible to control the direction of the projection. Alternatively, one might be viewing the planar projections of a group of isometric but randomly oriented particles. For the practical motivation for, and discussion of, this point of view, see [7].

We first recall Kendall's construction of the shape space as follows. A configuration of k labelled, and not totally identical, points in Euclidean m-dimensional space $\mathbb{R}^m$ is represented by an $m \times k$ matrix X^*. Without loss of generality, the configuration is assumed to have its centroid at the origin and then the matrix X^* is multiplied on the right by a certain $k \times k$ special orthogonal matrix which has the effect of producing a zero first column which is then discarded to give the pre-size-and-shape matrix X. The pre-shape matrix $X/\|X\|$ represents a point in the sphere $\mathbb{S}^{m(k-1)-1}$ and the shape $\pi(X)$ of the original configuration is the image in the shape space Σ_m^k, which is the quotient of the pre-shape sphere by the left action of SO(m) acting on the representative matrices. (Cf. [2].)

Assuming that the dimension of the Euclidean space where the configuration lies is 2, Panaretos [5] studied the diffusion, in the Kendall shape space, of the shape of the projected

configuration when the rotation evolves as a Brownian motion in SO(2). This he termed a Radon diffusion of shape. Panaretos also obtained, in the same paper, the equilibrium distribution of such a resulting shape diffusion. In [6], he studied the diffusion of the unoriented size-and-shape of the projected configuration under the assumption that the rotation evolves as a Brownian motion in SO(n) when the given configuration lies in an n-dimensional Euclidean space.

In this paper we study the induced distribution of the random shape, as well as that of the random size-and-shape, of the projected configuration when the rotation is uniformly distributed. Since the argument for the results relating to size-and-shapes are similar to those for shapes, we shall concentrate mainly on the results for shapes and give brief outlines for size-and-shapes at the end of the paper. Furthermore, since our result reduces to one of the results of [5] when the dimension of the Euclidean space in which the initial configuration lies is 2, we shall focus mainly on the case where the dimension is at least 3.

The paper is organised as follows. In Section 2 we introduce the injective map ϕ sending $v \in \mathbb{S}^{m-1}$ to the shape of the projection onto a fixed plane of the given configuration after it has been rotated by R, where R is any rotation such that Rv is the normal of the fixed plane. Our main technique is to reinterpret ϕ in terms of the map G sending v to the pre-shape of the projection of the fixed configuration onto the inversely rotated plane, that with normal v. We obtain a technical result involving the derivative of G that enables us, in Section 3, to use the geometry of shape spaces to obtain an isometric copy of the derivative ϕ_*. This makes it possible for us in Section 4 to derive the expression for the required induced distribution of the shape of the projection of the given configuration after it has been randomly rotated, in terms of the characteristic polynomial of a certain matrix. In Section 5 we further analyse this expression to obtain closed expressions when the dimension of the Euclidean space in which the original configuration lies is 3 or 4. The last section gives brief outlines on the corresponding arguments and results for the size-and-shape of the projected configuration after it has been randomly rotated.

2. The map ϕ to the shape of the projection

For a fixed unit length column vector v_0, representing a point on the unit $(m-1)$-sphere $\mathbb{S}^{m-1}$, we are interested in the shape of the projection onto the orthogonal complement $v_0^\perp$ of v_0 of a given configuration comprising k ($> m > 2$) labelled vertices in general position in $\mathbb{R}^m$ after it has been rotated by $R \in \mathrm{SO}(m)$. Note that both the orthogonal projection and rotated image of a centred configuration are also centred and that the operations of projection and rotation commute with the standard mapping from the space of centred configurations to the corresponding pre-size-and-shape space. It follows that the shape of the projection of the rotated configuration is also the shape of the projection of the rotated pre-size-and-shape of the given configuration. Thus, instead of specifying an $m \times k$ matrix representing the coordinates of the given configuration, we start from a fixed $m \times (k-1)$ matrix X of rank m representing, in standard coordinates, the pre-size-and-shape of the configuration. Without loss of generality, we also fix v_0 to be the mth standard basis element $e_m = (0, \ldots, 0, 1)^\top$. Then, the pre-size-and-shape of the above projection of the rotated image of the given configuration is $\Pi(e_m) RX$, where $\Pi(v) = I_m - vv^\top$ is the matrix representing the projection onto $v^\perp$ and I_m denotes the $m \times m$ identity matrix, and in particular $\Pi(e_m) RX$ is an $m \times (k-1)$ matrix with its final row 0. From this we derive our primary mapping

$$\psi : \mathrm{SO}(m) \to \mathscr{SS}_{m-1}^k, \qquad R \mapsto \Pi(e_m) RX, \tag{1}$$

where $\mathscr{SS}_{m-1}^k$ is the space of pre-size-and-shapes of configurations with k labelled vertices in

$\mathbb{R}^{m-1}$, embedded in $\mathcal{SS}_m^k$ in terms of the standard embedding by adding a final zero row to all its (matrix) elements.

Rotating X by R is equivalent to rotating e_m by $R^{-1} = R^\top$ and, indeed, writing $v_R = R^\top e_m$ and noting that $R^\top \Pi(Rv)R = \Pi(v)$ for any vector v and rotation R, we obtain the identity

$$\Pi(e_m)RX = R\Pi(v_R)X. \tag{2}$$

The assumption that $m > 2$ implies that the map $q \colon \mathrm{SO}(m) \to \mathbb{S}^{m-1}$, given by $R \mapsto v_R$, is not injective: $v_R = v_{\bar{R}}$ if and only if $R^\top e_m = \bar{R}^\top e_m$. This is however if and only if $\bar{R}R^\top$ is in the stabiliser $\mathrm{stab}(e_m)$ of e_m which we identify with $\mathrm{SO}(m-1)$. This gives the usual identification of $\mathbb{S}^{m-1}$ with the symmetric space $\mathrm{SO}(m)/\mathrm{SO}(m-1)$, the unit vector v corresponding to the coset of rotations R such that $v = v_R$.

Then, for any $S \in \mathrm{SO}(m-1)$, since $v_{SR} = v_R$, we have $SR\Pi(v_{SR})X = SR\Pi(v_R)X$. However, $R\Pi(v_R)X$ is the pre-size-and-shape of a configuration that lies in $e_m^\perp$ and $SR\Pi(v_R)X$ is that of that configuration rotated by S within $e_m^\perp$, so that $R\Pi(v_R)X$ and $SR\Pi(v_{SR})X$ have the same shape, which is a point in the Kendall shape space Σ_{m-1}^k of configurations with k labelled vertices in $\mathbb{R}^{m-1}$ that may be denoted by $\pi(\Pi(e_m)RX)$ in view of (2). It follows that the map in (1) induces a well-defined map

$$\phi \colon \mathbb{S}^{m-1} \to \Sigma_{m-1}^k, \qquad v \mapsto \pi(\Pi(e_m)RX), \tag{3}$$

where R is any choice such that $Rv = e_m$.

In order to study how the shape of the projected configuration changes as v does, we shall require the derivative ϕ_* of the map ϕ defined by (3). A natural way to compute this is to take a curve $\gamma(t)$ in $\mathbb{S}^{m-1}$ whose derivative $\dot{\gamma}(0)$ is a given vector u tangent to $\mathbb{S}^{m-1}$ at v and then

$$\phi_*(u) = \frac{\mathrm{d}}{\mathrm{d}t}(\phi \circ \gamma)(t)|_{t=0}$$

is tangent to Σ_{m-1}^k at the shape $\phi(v)$. Since the image of ϕ given by (3) involves $R \in \mathrm{SO}(m)$ which determines v and since the quotient map q is a Riemannian submersion, it is more convenient to represent the vector u by its horizontal lift in the tangent space to $\mathrm{SO}(m)$ and replace γ by a horizontal lift $\tilde{\gamma}$ in $\mathrm{SO}(m)$. For this, we recall that the tangent space to $\mathrm{SO}(m)$ at the identity $I = I_m$ is its Lie algebra $\mathfrak{so}(m)$ comprising the skew-symmetric $m \times m$ matrices, when $\mathrm{SO}(m)$ is regarded as a subgroup of the general linear group of all invertible matrices with its standard Riemannian metric. Then, in the tangent space to $\mathrm{SO}(m)$ at the identity I_m, the vertical subspace for the quotient map q, namely the tangent space to $\mathrm{SO}(m-1)$, comprises the matrices $A = \mathrm{diag}(A_0, 0)$, where $A_0 \in \mathfrak{so}(m-1)$, and so the horizontal subspace, the orthogonal complement to the vertical subspace, is

$$\mathcal{H}_I = \left\{ \begin{pmatrix} 0_{m-1} & h \\ -h^\top & 0 \end{pmatrix} \,\middle|\, h \in \mathbb{R}^{m-1} \right\},$$

where 0_{m-1} denotes the $(m-1) \times (m-1)$ zero matrix (cf. [4, p. 318]). Since $\mathrm{stab}(v_R) = R^\top \mathrm{stab}(e_m)R$, it follows that the horizontal subspace at $R \in \mathrm{SO}(m)$ is $\mathcal{H}_R = R^\top \mathcal{H}_I R$. Then, the corresponding tangent spaces at e_m and v_R in $\mathbb{S}^{m-1}$ are $\{He_m \mid H \in \mathcal{H}_I\}$ and $\{H^R v_R \mid H \in \mathcal{H}_I\}$, respectively, where $H^R = R^\top H R$ (*loc. cit.*).

We now consider the curve $v(t) = \mathrm{e}^{tH^R} v_R$ in $\mathbb{S}^{m-1}$, where $H \in \mathcal{H}_I$, with $v(0) = v_R$ and $\dot{v}(0) = H^R v_R = R^\top He_m \neq 0$ in $\tau_{v_R}(\mathbb{S}^{m-1})$, the tangent space at v_R to $\mathbb{S}^{m-1}$. Then,

$R(t) = R\mathrm{e}^{-tH^R}$ is the corresponding horizontal lift into $\mathrm{SO}(m)$: $v(t) = R(t)^\top e_m$ and $R(t)$ rotates the projection of X onto $v(t)^\perp$ back into $e_m^\perp$ so that, in particular, all vertices of the configuration corresponding to $R(t)\Pi(v(t))X$ lie in the same, standard, hyperplane $e_m^\perp$. The projection $\varphi\colon \mathcal{S}_{m-1}^k \to \Sigma_{m-1}^k$ from the standard pre-shape sphere to the shape space is a Riemannian submersion when restricted to the nonsingular part, the points represented by matrices of rank at least $m - 1$, where indeed the image of ϕ lies. Then

$$\phi(v(t)) = \varphi\left(\frac{(\psi \circ R)(t)}{\|(\psi \circ R)(t)\|} \right)$$

by (1) and so

$$\phi_*(\dot{v}(0)) = \varphi_*\left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{(\psi \circ R)(t)}{\|(\psi \circ R)(t)\|} \right)\bigg|_{t=0} \right). \tag{4}$$

Denoting $\Pi(v)X$ by $P(v)$ and $P(v)/\|P(v)\|$ by $G(v)$, the pre-size-and-shape and pre-shape, respectively, of the projected configuration onto $v^\perp$, the following result reinterprets the derivative

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{(\psi \circ R)(t)}{\|(\psi \circ R)(t)\|} \right)\bigg|_{t=0}$$

in terms of the tangent vector

$$\frac{\mathrm{d}G(v(t))}{\mathrm{d}t}\bigg|_{t=0},$$

which is independent of the choice of $R \in [R]$.

Lemma 1. *For any $R \in \mathrm{SO}(m)$ such that $v = v_R$ and the curves $v(t)$ and $R(t)$ defined above,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\psi(R(t))}{\|\psi(R(t))\|} \right)\bigg|_{t=0} = R\Pi(v_R)\frac{\mathrm{d}G(v(t))}{\mathrm{d}t}\bigg|_{t=0}.$$

Proof. Since $v(t) = v_{R(t)}$, we have, by (2), $\psi(R(t)) = R(t)P(v(t))$. Then

$$\frac{\mathrm{d}G(v(t))}{\mathrm{d}t}\bigg|_{t=0} = \frac{1}{\|P(v(0))\|}\frac{\mathrm{d}P(v(t))}{\mathrm{d}t}\bigg|_{t=0} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\|P(v(t))\|} \right)\bigg|_{t=0} P(v(0)),$$

which, recalling that $R(0) = R$, so that $v(0) = v_R$, and using the facts that $\|P(v(t))\| = \|\psi(R(t))\|$ and $P(v) = \Pi(v)P(v)$, gives

$$R\Pi(v_R)\frac{\mathrm{d}G(v(t))}{\mathrm{d}t}\bigg|_{t=0} = \frac{R\Pi(v_R)\mathrm{d}P(v(t))/\mathrm{d}t|_{t=0}}{\|\psi(R(0))\|} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\|\psi(R(t))\|} \right)\bigg|_{t=0} \psi(R(0)).$$

So, we only need to show that

$$\frac{\mathrm{d}\psi(R(t))}{\mathrm{d}t}\bigg|_{t=0} = R\Pi(v_R)\frac{\mathrm{d}P(v(t))}{\mathrm{d}t}\bigg|_{t=0}.$$

To see this, recall that $\dot{R}(0) = -RH^R$, that $\dot{v}(0) = H^R v_R$, and that H^R is skew-symmetric. Then,

$$\frac{\mathrm{d}}{\mathrm{d}t}\{R(t)\Pi(v(t))\}|_{t=0} = -RH^R\Pi(v_R) - R\{\dot{v}(0)v(0)^\top + v(0)\dot{v}(0)^\top\}$$

$$= -RH^R\{I - v_Rv_R^\top\} - RH^Rv_Rv_R^\top + Rv_Rv_R^\top H^R$$

$$= -R\{I - v_Rv_R^\top\}H^R.$$

However, for $H \in \mathcal{H}_I$, we have $H = e_m e_m^\top H + H e_m e_m^\top$ and so, conjugating by R, $H^R = v_R v_R^\top H^R + H^R v_R v_R^\top$. Also, since $\dot{v}(0)$ is orthogonal to $v(0)$ and $\|v(0)\| = 1$, $\Pi(v(0))\dot{\Pi}(v(0))$ reduces to $-\dot{v}(0)v(0)^\top$. The required result then follows from

$$\frac{\mathrm{d}}{\mathrm{d}t}\{R(t)\Pi(v(t))\}|_{t=0} = -RH^R v_R v_R^\top = -R\dot{v}(0)v(0)^\top = R\Pi(v_R)\frac{\mathrm{d}\Pi(v(t))}{\mathrm{d}t}\bigg|_{t=0}.$$

If $m = 2$, writing $u(t) = Jv(t)$, where J is clockwise rotation by $\pi/2$, we have $P(v) = uu^\top X$ and so, since $RJR^\top = J$,

$$RP(v_R) = Ru_R u_R^\top X = RJR^\top e_2 u_R^\top X = e_1 u_R^\top X.$$

On the other hand, in this case, the map $q \colon \mathrm{SO}(2) \to \mathbb{S}^1$ is injective and the pre-shape sphere $\mathscr{S}_1^k$ *is the shape space* Σ_1^k. Hence, φ is the identity map and $\phi(v_R) = \psi(R)/\|\psi(R)\|$, so that

$$\phi_*(\dot{v}(0)) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{(\psi \circ R)(t)}{\|(\psi \circ R)(t)\|}\right)\bigg|_{t=0}.$$

However, $\Pi(v)\dot{P}(v) = u\dot{u}^\top X$. Thus, it follows from Lemma 1 and the equality $Ru = RJR^\top e_2 = Je_2 = e_1$ that

$$\phi_*(\dot{v}(0)) = R\Pi(v)\left(\frac{\mathrm{d}G(v(t))}{\mathrm{d}t}\bigg|_{t=0}\right) = Ru\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{u(t)^\top X}{\|u(t)^\top X\|}\right)\bigg|_{t=0} = e_1\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{u(t)^\top X}{\|u(t)^\top X\|}\right)\bigg|_{t=0},$$

which recovers the result of [5] in the sense that $u^\top X$ and $u^\top X/\|u^\top X\|$ are precisely the objects used there in studying the shape of the projection of a randomly rotated configuration for the case $m = 2$.

It is worth noting that the approach we take in this paper is not quite the same as that of [6]. We concentrate on the evolution of the shape of the projection of a given configuration onto a fixed hyperplane after being randomly rotated, whereas Panaretos [6] studied the shape of the projection of a given configuration onto a randomly rotated hyperplane. In terms of our current notation, $\Pi(v_R)X$ is the object considered in [6]. Although these two approaches result in the same shape, it is actually represented by points in different shape spaces. For our approach, although $\Pi(e_m)RX$ is an $m \times (k-1)$ matrix for any $R \in \mathrm{SO}(m)$, its final row is always 0 and so, by ignoring it, the shape of $\Pi(e_m)RX$ is in effect considered in the shape space Σ_{m-1}^k of configurations in $\mathbb{R}^{m-1}$ with k labelled points. While, for the approach of [6], $\Pi(v_R)X$ is treated as a pre-size-and-shape of a degenerate configuration in $\mathbb{R}^m$ and so its shape will lie in the subspace of degenerate shapes in Σ_m^k. Of course, these two spaces, Σ_{m-1}^k and the subspace of degenerate shapes in Σ_m^k, are locally isometric, the former being a Riemannian double covering of the latter. The difference between these two approaches is also reflected in the result of the lemma, as the corresponding result for the approach of [6] would require neither the projection, nor the rotation by R.

3. The derivative of ϕ

Since, as already noted, φ is a Riemannian submersion where it concerns us, φ_* maps the horizontal subspace of the tangent space to $\mathscr{S}_{m-1}^k$ at $RG(v)$ isometrically onto the tangent space to Σ_{m-1}^k at $\phi(v)$. Moreover, as R is an isometry, taking advantage of Lemma 1, we may work directly with the derivative of G to compute that of ϕ. That is, using (4), we find the horizontal

component of the projection onto $v^\perp$ of the derivative of G to obtain an isometric copy of the image of the derivative of ϕ, instead of that image itself. For the chosen $v \in \mathbb{S}^{m-1}$, let $u_1, u_2, \ldots, u_{m-1} \in \mathbb{S}^{m-1}$ be such that $\{u_1, \ldots, u_{m-1}, v\}$ forms an orthonormal basis of $\mathbb{R}^m$. Then, the projection $\Pi(v)$ onto the hyperplane $v^\perp$ with orthonormal basis $\{u_1, \ldots, u_{m-1}\}$ can be expressed as $\Pi(v) = \sum_{i=1}^{m-1} u_i u_i^\top$. Moreover, the vectors $u_1, \ldots, u_{m-1}$ form an orthonormal frame for the tangent space at v to $\mathbb{S}^{m-1}$ and the linearity of ϕ_* implies that it is sufficient to find the images of these vectors to determine ϕ_*. However, these images form a sub-frame in the tangent space to Σ_{m-1}^k at $\phi(v)$ isometric with the sub-frame in the tangent space at $G(v)$ formed by the horizontal components of the projection onto $v^\perp$ of the images of the u_i under the derivative of G. These latter projected images are the tangent vectors $W_i(v)$ at $G(v)$ to the pre-shape sphere of configurations in $v^\perp$ given by

$$W_i(v) = \Pi(v)\frac{dG(v_i(t))}{dt}\bigg|_{t=0},$$

where $v_i(t)$ is a curve in $\mathbb{S}^{m-1}$ with $v_i(0) = v$ and $\dot{v}_i(0) = u_i$. The vector $W_i(v)$ has the expression

$$W_i(v) = \Pi(v)\left\{\frac{\dot{P}(v_i(t))}{\|P(v_i(t))\|} - \frac{P(v_i(t))}{\|P(v_i(t))\|^3}\langle P(v_i(t)), \dot{P}(v_i(t))\rangle\right\}\bigg|_{t=0},$$

which, since $\Pi(v)\dot{P}(v_i(t))|_{t=0} = -u_i v^\top X$ and $\langle P(v_i(t)), \dot{P}(v_i(t))\rangle|_{t=0} = -\langle X^\top u_i, X^\top v\rangle$, gives

$$W_i(v) = \frac{1}{\|P(v)\|}\left\{\frac{G(v)}{\|P(v)\|}\langle X^\top u_i, X^\top v\rangle - u_i v^\top X\right\}. \tag{5}$$

We now calculate the required horizontal components of these tangent vectors.

Proposition 1. *For $i = 1, \ldots, m-1$, the horizontal component $W_i^h(v)$ of*

$$W_i(v) = \Pi(v)\frac{dG(v_i(t))}{dt}\bigg|_{t=0}$$

given by (5) may be expressed as

$$W_i^h(v) = \frac{1}{\|P(v)\|^3}\theta_i UU^\top X - \frac{1}{\|P(v)\|}\{u_i v^\top - UA^{(i)}U^\top\}X,$$

where $\theta_i = \langle X^\top u_i, X^\top v\rangle$, U is the $m \times (m-1)$ matrix whose ith column is u_i, $i = 1, \ldots, m-1$, and the skew-symmetric matrix $A^{(i)}$ is determined by

$$\bar{e}_i \boldsymbol{\theta}^\top - \boldsymbol{\theta}\bar{e}_i^\top = A^{(i)}\Theta - (A^{(i)}\Theta)^\top, \tag{6}$$

in which $\boldsymbol{\theta} = U^\top XX^\top v$ and $\Theta = U^\top XX^\top U$.

Proof. We need first to identify the vertical and horizontal components, with respect to φ, of the relevant tangent spaces. For this, we note that, after the projection $\Pi(v)$, we are working with the spaces of shapes, pre-shapes, and pre-size-and-shapes of configurations in the hyperplane $v^\perp$. The matrix of a configuration of k points in such a hyperplane will, in general, be an $m \times k$ matrix of rank $m - 1$, which, assuming that $Rv = e_m$, the rotation R will take to a matrix with final row 0. The operations taking these matrices to the corresponding

matrices of pre-size-and-shapes and pre-shapes are all linear or affine and so multiplication by R will take the matrix Y of a pre-shape in $v^\perp$ to that, $Z = RY$, of a pre-shape in $e_m^\perp$ which has its final row 0. Since the projection φ from the standard pre-shape sphere $\mathcal{S}_{m-1}^k$ to the shape space Σ_{m-1}^k is the quotient by $SO(m-1)$, the vertical subspace of the tangent space to $\mathcal{S}_{m-1}^k$ at the pre-shape Z is

$$\{AZ \mid A = \mathrm{diag}(A_0, 0), \; A_0 \in \mathfrak{so}(m-1)\}$$

(cf. [2]). This implies that the vertical subspace, of the tangent space at Y to the pre-shape sphere of configurations in $v^\perp$, is

$$\{A^R Y \mid A = \mathrm{diag}(A_0, 0), \; A_0 \in \mathfrak{so}(m-1)\},$$

as $SO(m-1)^R$ stabilises v and, hence, also $v^\perp$. The isometry R takes the latter space to the former and, hence, also preserves the orthogonal, horizontal subspaces. Thus, since a tangent vector W at $G(v)$ is horizontal if and only if the image vector RW at $RG(v)$ is horizontal and that is if and only if $RW(RG(v))^\top$ is symmetric (cf. [2, p. 258]), it follows that W is horizontal if and only if $WG(v)^\top$ or, equivalently, $WP(v)^\top$, is symmetric.

Now, $G(v)G(v)^\top$ is already symmetric and so (5) leads to the conclusion that the vertical component of $W_i(v)$ is proportional to that of $u_i v^\top X$. This in turn implies that finding the horizontal component of $u_i v^\top X$ is sufficient for deducing $W_i^h(v)$. To achieve the former, we note that $UU^\top = \Pi(v)$ and $U^\top U = I_{m-1}$. Thus,

$$u_i v^\top X P(v)^\top = u_i v^\top X X^\top \Pi(v) = U \bar{e}_i \theta^\top U^\top, \tag{7}$$

where $\bar{e}_i$ is the ith standard basis vector in $\mathbb{R}^{m-1}$ and the $(m-1)$-dimensional column vector $\theta = U^\top X X^\top v$ has ith component $\theta_i = u_i^\top X X^\top v = \langle X^\top u_i, X^\top v \rangle$. Decomposing $u_i v^\top X = T_i^h + T_i^v$ into its horizontal and vertical components, the earlier argument then shows that $T_i^h P(v)^\top$ must be a symmetric linear combination of the basic rank one matrices $u_s u_t^\top$ which we may write as

$$T_i^h P(v)^\top = U B^{(i)} U^\top, \tag{8}$$

where the $(m-1) \times (m-1)$ matrix $B^{(i)}$ is symmetric. Similarly, since $\mathfrak{so}(m-1)$ is spanned by $\{e_s e_t^\top - e_t e_s^\top \mid 1 \leqslant s < t \leqslant m-1\}$, where $e_1, \ldots, e_{m-1}$ are the standard basis vectors of $\mathbb{R}^m$ that span the $(m-1)$-dimensional subspace $e_m^\perp$ of $\mathbb{R}^m$, the vertical component T_i^v must be of the form

$$U A^{(i)} U^\top P(v) = U A^{(i)} U^\top U U^\top X = U A^{(i)} U^\top X,$$

where $A^{(i)}$ is a skew-symmetric $(m-1) \times (m-1)$ matrix. Thus,

$$T_i^v P(v)^\top = U A^{(i)} U^\top X X^\top U U^\top = U A^{(i)} \Theta U^\top, \tag{9}$$

where the $(m-1) \times (m-1)$ matrix $\Theta = U^\top X X^\top U$ has entries $\theta_{ij} = u_i^\top X X^\top u_j = \langle X^\top u_i, X^\top u_j \rangle$. Then, (7) with (8) plus (9) gives

$$\bar{e}_i \theta^\top = B^{(i)} + A^{(i)} \Theta, \tag{10}$$

and, since $B^{(i)}$ is symmetric, (6) follows. This identity between $(m-1) \times (m-1)$ skew-symmetric matrices provides $\binom{m-1}{2} = (m-1)(m-2)/2$ independent equations for the same number of independent, not necessarily zero entries of $A^{(i)}$. Hence, we obtain the stated result.

It is possible to render (6) for the $A^{(i)}$ more explicit. For any skew-symmetric $(m-1) \times (m-1)$ matrix $A = (A_{ij})$, write $\mu(A)$ for the column vector,

$$\mu(A) = (A_{12}, A_{13}, A_{23}, A_{14}, \ldots, A_{34}, \ldots, A_{1m-1}, \ldots, A_{m-2m-1})^{\top},$$

comprising the $\binom{m-1}{2}$ elements above the diagonal in A taken in the order running down successive columns as indicated. Then, if $\boldsymbol{\alpha}_m$ is the $\binom{m-1}{2} \times (m-1)$ matrix whose ith column is $\mu(A^{(i)})$, we have

$$C_m \boldsymbol{\alpha}_m = L_m(\boldsymbol{\theta}), \tag{11}$$

where C_m and L_m are respectively the $\binom{m-1}{2} \times \binom{m-1}{2}$ and $\binom{m-1}{2} \times (m-1)$ matrices defined inductively as follows. For a $j-1$ column vector $\boldsymbol{x}$, $j \geqslant 3$, define

$$L_{j+1}\begin{pmatrix} \boldsymbol{x} \\ x_j \end{pmatrix} = \begin{pmatrix} L_j(\boldsymbol{x}) & 0_{(j-1)(j-2)/2 \times 1} \\ x_j I_{j-1} & -\boldsymbol{x} \end{pmatrix} \tag{12}$$

with $L_3((x_1, x_2)^{\top}) = (x_2, -x_1)$. Then,

$$C_{j+1} = \begin{pmatrix} C_j & L_j((\theta_{1j}, \ldots, \theta_{j-1j})^{\top}) \\ L_j((\theta_{1j}, \ldots, \theta_{j-1j})^{\top})^{\top} & E_j \end{pmatrix}, \qquad j \geqslant 3, \tag{13}$$

where $E_j = \theta_{jj} I_{j-1} + \Theta_{[j-1]}$, $\Theta_{[j-1]}$ is the submatrix of Θ formed by the intersection of its first $j-1$ rows and columns, and the induction starts from $C_3 = E_2 = \{\theta_{11} + \theta_{22}\} I_1$. Note that, since E_j is symmetric, so too, by induction, is each C_j.

That (11) is equivalent to (6) holding for each i follows from the fact that the ith column of $C_m \boldsymbol{\alpha}_m$ is $\mu(A^{(i)}\Theta - (A^{(i)}\Theta)^{\top})$, while that of $L_m(\boldsymbol{\theta})$ is $\mu(\bar{e}_i \boldsymbol{\theta}^{\top} - (\bar{e}_i \boldsymbol{\theta}^{\top})^{\top})$. This may be seen by indexing the rows of C_m and $L_m(\boldsymbol{\theta})$ and the columns of C_m by the same ordered pairs as the elements of $\mu(A)$. Then, for $1 \leqslant p < q \leqslant m-1$ and $1 \leqslant r < s \leqslant m-1$, the element in the pqth row and rsth column of C_m, the coefficient of $A_{rs}^{(i)}$ in $(A^{(i)}\Theta - (A^{(i)}\Theta)^{\top})_{pq}$, is

$$C_{(pq)(rs)} = \delta_{pr}\theta_{qs} + \delta_{qs}\theta_{pr} - \delta_{ps}\theta_{qr} - \delta_{qr}\theta_{ps}. \tag{14}$$

The symmetry and the inductive structure of C_m, as well as its independence of the matrix $A^{(i)}$, are now evident. The submatrix E_{m-1} comprises the entries $C_{(pq)(rs)}$ with $s = q = m-1$ for which expression (14) reduces to $\theta_{rp} + \delta_{pr}\theta_{m-1m-1}$, as required, since p and r range from 1 to $m-2$. The submatrix L_{m-1} comprises the entries with $q < m-1$ and $s = m-1$ for which (14) becomes

$$\delta_{pr}\theta_{qm-1} - \delta_{rq}\theta_{pm-1}, \tag{15}$$

so that the entries only involve the θ_{im-1}, $1 \leqslant i < m-1$, occurring in the last column of Θ. Writing $\boldsymbol{\vartheta}$ for this column we note that the pqth entry in the rth column is $\delta_{pr}\boldsymbol{\vartheta}_q - \delta_{rq}\boldsymbol{\vartheta}_p = (\bar{e}_r \boldsymbol{\vartheta}^{\top} - \boldsymbol{\vartheta}\bar{e}_r^{\top})_{pq}$, confirming the right-hand side of (11). Setting $q = m-2$ in (15) reduces it to $\delta_{pr}\theta_{m-2m-1} - \delta_{rm-2}\theta_{pm-1}$, as required for the last $m-3$ rows of $L_m((\theta_{1m-1}, \ldots, \theta_{m-2m-1})^{\top})$ since here p ranges from 1 to $m-3$ and r from 1 to $m-2$. Finally, for $r = m-2$, $s = m-1$, and $1 \leqslant p < q < m-2$, (15) vanishes and so gives the zero column vector in the inductive structure of L_{m-1}.

Returning to the remark that we made at the end of the previous section on the approach of [6], the vector field in that approach, corresponding to $W_i(v)$ considered in this section, would be $dG(v_i(t))/dt|_{t=0}$ itself. However, since the Riemannian submersion would then be from $\mathcal{S}_m^k$ to Σ_m^k, rather than from $\mathcal{S}_{m-1}^k$ to Σ_{m-1}^k, and since we can apply the results from

standard shape theory directly in that situation, the vertical tangent space to $\mathcal{S}_m^k$ at the pre-shape $G(v)$ becomes $\{AP(v) \mid A \in \mathfrak{so}(m)\}$. Hence, it follows that the difference

$$\left.\frac{\mathrm{d}G(v_i(t))}{\mathrm{d}t}\right|_{t=0} - W_i(v) = -vu_i^\top X = -\{vu_i^\top - u_i^\top v\}P(v)$$

is a vertical tangent vector, which implies that the horizontal component of $\mathrm{d}G(v_i(t))/\mathrm{d}t|_{t=0}$ is identical with $W_i^h(v)$.

4. The induced distribution

We now assume that the normal v to the hyperplane onto which the given configuration is projected is uniformly distributed on $\mathbb{S}^{m-1}$ and, using the result of the previous section, investigate the resulting induced distribution of the shape of such a projection. Note that this induced distribution is also the equilibrium distribution of the diffusion on $\phi(\mathbb{S}^{m-1})$ induced by Brownian motion on the sphere $\mathbb{S}^{m-1}$. Our main result, generalising that of [5] to the case $m > 2$, is the following.

Theorem 1. *If v is uniformly distributed on the sphere $\mathbb{S}^{m-1}$ then, with respect to the volume element on $\phi(\mathbb{S}^{m-1})$, the Radon–Nikodym derivative at $\phi(v)$, defined by (3), of the shape of the projection of the configuration with pre-size-and-shape X is given by*

$$\frac{1}{\mathrm{vol}(\mathbb{S}^{m-1})\sqrt{\det(\langle W_i^h, W_j^h\rangle)}},$$

where $\|[P(v)\|^{4(m-1)}\det(\langle W_i^h, W_j^h\rangle)$ is equal to the characteristic polynomial of

$$L_m(\boldsymbol{\theta})^\top\left\{\|P(v)\|^2 C_m^{-1} - I_{\binom{m-1}{2}}\right\}L_m(\boldsymbol{\theta}) \tag{16}$$

evaluated at $\theta_v\|P(v)\|^2 - \sum_{i=1}^{m-1}\theta_i^2$. Here $W_i^h \equiv W_i^h(v)$ is given by Proposition 1, $P(v)$ is the projection of X onto the hyperplane $v^\perp$, and L_m and C_m are defined by (12) and (13) with $\theta_{ij} = \langle X^\top u_i, X^\top u_j\rangle$, $\theta_i = \langle X^\top u_i, X^\top v\rangle$, and $\theta_v = \|X^\top v\|^2$.

Proof. The uniform distribution on a compact Riemannian manifold of dimension n is, up to a normalisation constant, the volume element of the manifold that is given locally by $\sqrt{\det(g)}\,\mathrm{d}x_1\cdots\mathrm{d}x_n$, where g is the matrix of local coordinates of the Riemannian metric. Since we have chosen an orthonormal basis $\{u_1, \ldots, u_{m-1}\}$ of the tangent space to $\mathbb{S}^{m-1}$ at v, the corresponding matrix $g(v)$ will be the identity. The induced metric Φ on $\phi(\mathbb{S}^{m-1}) \subseteq \Sigma_{m-1}^k$ will then be such that

$$\phi^*(\Phi)(u_i, u_j) = \delta_{ij} = \Phi(\phi_*(u_i), \phi_*(u_j)).$$

However, writing $\Psi_{ij} = \langle\phi_*(u_i), \phi_*(u_j)\rangle$ for the inner product of the image vectors with respect to the (Kendall) Riemannian metric on $\phi(\mathbb{S}^{m-1}) \subseteq \Sigma_{m-1}^k$ and noting that $\phi_*(u_i) = \varphi_*(RW_i(v))$, we have

$$\Psi_{ij} = \langle W_i^h(v), W_j^h(v)\rangle,$$

since both φ_*, when restricted to the horizontal subspace, and the left action of R are isometries. As the matrix $\Psi = (\Psi_{ij})_{(m-1)\times(m-1)}$ is symmetric and positive definite, it has a unique positive

definite square root with inverse, say $M = (M_{ij})_{(m-1)\times(m-1)}$, and then

$$\left\langle \sum_{\ell=1}^{m-1} M_{\ell i} W_\ell^h(v), \sum_{\ell=1}^{m-1} M_{\ell j} W_\ell^h(v) \right\rangle = \delta_{ij} = \Phi(\phi_*(u_i), \phi_*(u_j)).$$

On the other hand, if we write $d\xi_i$, $i = 1, \ldots, m-1$, for the dual forms of the $\phi_*(u_i)$, the standard volume form $d\mathrm{vol}$ of the Riemannian metric on $\phi(\mathbb{S}^{m-1})$ is given by $d\mathrm{vol} = \sqrt{\det(\Psi)}\, d\xi_1 \cdots d\xi_{m-1}$. Thus, the corresponding volume form, induced from that of $\mathbb{S}^{m-1}$, will be $\{\det(\Psi)\}^{-1/2} d\mathrm{vol}$ and this, up to a normalisation constant, will also be the expression for the induced distribution on $\phi(S^{m-1})$ that we seek.

Recalling that $U^\top U = I_{m-1}$ and $UU^\top = \Pi(v)$, so that $UU^\top u_i = u_i$, and using the standard properties of the trace together with (10), the expression for $\langle W_i^h, W_j^h \rangle = \mathrm{tr}(W_i^h (W_j^h)^\top)$ arising from Proposition 1 simplifies substantially to

$$\langle W_i^h, W_j^h \rangle = \frac{1}{\|P(v)\|^2} \left\{ \delta_{ij}\theta_v - \frac{1}{\|P(v)\|^2}\theta_i\theta_j + \mathrm{tr}(A^{(i)}\Theta A^{(j)}) \right\},$$

where $\theta_v = \|X^\top v\|^2$. Then, the identity

$$-\mathrm{tr}(A^{(i)}\Theta A^{(j)}) = (\boldsymbol{\alpha}_m^\top C_m \boldsymbol{\alpha}_m)_{ij} = (\mu(A^{(i)}))^\top C_m \mu(A^{(j)})$$

may be established using (14) for the components of C_m and, noting that

$$L_m(\boldsymbol{\theta})^\top C_m^{-1} L_m(\boldsymbol{\theta}) = \boldsymbol{\alpha}_m^\top C_m \boldsymbol{\alpha}_m,$$

it follows that the matrix Ψ has the expression

$$\Psi = \frac{1}{\|P(v)\|^2} \left\{ \theta_v I_{m-1} - \frac{1}{\|P(v)\|^2}\boldsymbol{\theta}\boldsymbol{\theta}^\top - L_m(\boldsymbol{\theta})^\top C_m^{-1} L_m(\boldsymbol{\theta}) \right\}. \tag{17}$$

However, by (15) we see that

$$L_m(\boldsymbol{\theta})^\top L_m(\boldsymbol{\theta}) = -\boldsymbol{\theta}\boldsymbol{\theta}^\top + \left\{ \sum_{i=1}^{m-1} \theta_i^2 \right\} I_{m-1},$$

and so we can rewrite (17) to obtain

$$\Psi = \frac{1}{\|P(v)\|^4} \left\{ \left(\theta_v \|P(v)\|^2 - \sum_{i=1}^{m-1} \theta_i^2 \right) I_{m-1} - L_m(\boldsymbol{\theta})^\top \left(\|P(v)\|^2 C_m^{-1} - I_{\binom{m-1}{2}} \right) L_m(\boldsymbol{\theta}) \right\}. \tag{18}$$

The required result then follows from

$$\|P(v)\|^{4(m-1)} \det(\Psi) = \det\{\eta I_{m-1} - L_m(\boldsymbol{\theta})^\top \left(\|P(v)\|^2 C_m^{-1} - I_{\binom{m-1}{2}} \right) L_m(\boldsymbol{\theta})\},$$

where $\eta = \theta_v \|P(v)\|^2 - \sum_{i=1}^{m-1} \theta_i^2$.

It can be checked that the matrix in (16) is never of full rank, so that $\theta_v \|P(v)\|^2 - \sum_{i=1}^{m-1} \theta_i^2$ is always a factor of $\det(\langle W_i^h, W_j^h \rangle)$.

Since $\|P(v)\|^2 = \text{tr}(\Theta)$, Theorem 1 expresses the induced distribution in terms of the inner products and squared norms of the vectors $X^\top u_i$ and $X^\top v$, which we have arranged in the matrix Θ, the vector $\boldsymbol{\theta}$, and the scalar θ_v. Alternatively, we may express this result in terms of the eigenvalues λ_i^2, $1 \leqslant i \leqslant m$, of $XX^\top$ and those $\lambda_i(v)^2$, $1 \leqslant i \leqslant m-1$, of $P(v)P(v)^\top$ that are nonzero: since we have assumed that $\text{rank}(X) = m$, $\text{rank}(P(v))$ will be $m-1$ and there will indeed be $m-1$ nonzero eigenvalues of $P(v)P(v)^\top$. Since the determinant of Ψ that we require is invariant under an orthogonal transformation of the basis $\{u_1, \ldots, u_{m-1}\}$, we assume without loss of generality that $\{u_1, \ldots, u_{m-1}\}$ may be chosen so that Θ is a diagonal matrix. Under this assumption, by (14), C_m becomes the diagonal matrix with $C_{(pq)(pq)} = \theta_{pp} + \theta_{qq}$. The θ_{ii} are then the nonzero eigenvalues of $P(v)P(v)^\top$ and, if V is the orthogonal matrix comprising U with the column v appended,

$$V^\top XX^\top V = \begin{pmatrix} \Theta & \boldsymbol{\theta} \\ \boldsymbol{\theta}^\top & \theta_v \end{pmatrix} \tag{19}$$

has the same characteristic polynomial as $XX^\top$. Expanding these two polynomials and equating coefficients, we obtain identities that allow us to express θ_j, $1 \leqslant j \leqslant m-1$, and θ_v as functions of symmetric polynomials in the λ_i^2 and those in various subsets of the $\lambda_i(v)^2$. For example, equating the coefficients in degree $m-2$, we obtain

$$\sum_{1 \leqslant i < j \leqslant m} \lambda_i^2 \lambda_j^2 = \theta_v \sum_{i=1}^{m-1} \theta_{ii} + \sum_{1 \leqslant i < j \leqslant m-1} \theta_{ii}\theta_{jj} - \sum_{i=1}^{m-1} \theta_i^2,$$

which allows us to rewrite the parameter η in Theorem 1 as

$$\theta_v \|P(v)\|^2 - \sum_{i=1}^{m-1} \theta_i^2 = \eta = \sum_{1 \leqslant i < j \leqslant m} \lambda_i^2 \lambda_j^2 - \sum_{1 \leqslant i < j \leqslant m-1} \lambda_i(v)^2 \lambda_j(v)^2.$$

Thus, the coefficients in the characteristic polynomial (16) and η, and, hence, also the Radon–Nikodym derivative of Theorem 1, are all functions solely of the eigenvalues of $XX^\top$ and those of $P(v)P(v)^\top$.

It follows from the expressions for C_m and $L_m(\boldsymbol{\theta})$ given in the previous section that replacing X by aX multiplies these, and also $\|P(v)\|^2$, by a^2 and, hence, that the Radon–Nikodym derivative of Theorem 1 is independent of the size of the initial given configuration. Furthermore, that expression is universal in the following sense. If we have two configurations whose pre-size-and-shape are related by $X_1 = XS$, where $S \in \text{SO}(k-1)$, then $XX^\top$ and $X_1 X_1^\top$ have the same eigenvalues, and the corresponding $P(v)P(v)^\top$ and $P_1(v)P_1(v)^\top$ also have the same eigenvalues. Hence, the discussion of the previous paragraph implies that the expressions for the Radon–Nikodym derivatives corresponding to X and X_1 given by Theorem 1 are identical and so its dependence on the given configuration is only via the eigenvalues of $XX^\top$ and of $P(v)P(v)^\top$. However, the domain $\phi(\mathbb{S}^{m-1})$ in Σ_{m-1}^k on which the expression of Theorem 1 lies does depend on the choice of S. Nevertheless, since the right action of S induces an isometry on the shape space, the subspace $\phi(\mathbb{S}^{m-1})$ corresponding to X is isometric with that corresponding to X_1.

Recall that the pseudo-singular values decomposition of X is $X = R(\Lambda, 0)S$, where $R \in \text{SO}(m)$, $S \in \text{SO}(k-1)$, and Λ is a diagonal matrix whose diagonal entries are the positive square roots of the eigenvalues of $XX^\top$ except that, when $k = m+1$, the sign of the smallest one is the same as that of $\det(X)$. Recall also that, in terms of its pseudo-singular values decomposition, the shape of X is determined by the pair $(\Lambda/\|\Lambda\|, S)$. The above discussion

implies that the expression for the Radon–Nikodym derivative gives us some information on Λ. We shall see in the next section that, up to the sign of the smallest diagonal element when $k = m + 1$, $\Lambda/\|\Lambda\|$ can be determined by the extreme values of the Radon–Nikodym derivative when $m = 2$ or 3.

5. Low-dimensional cases

In this section we compute explicit closed expressions for the Radon–Nikodym derivative in the cases $m = 2$, 3, and 4 in order to gain more insight into the induced distribution. First, we note that, when $m = 2$, as pointed out in the discussion following the proof of Lemma 1, $P(v)/\|P(v)\|$ is the shape of $P(v)$ and so, in terms of our notation, $\phi_*(u) = RW(v_R)$, where we have denoted u_1 by u and $W_1 = W_1^h$ by W. Thus, in this case, since both $L_2(\theta)$ and C_2 are vacuous and since $\|P(v)\|^2 = \|u^\top X\|^2 = \theta_{11}$, the Radon–Nikodym derivative of the distribution, induced by uniformly distributed v, of the shape of the projected configuration is $1/(2\pi\|W(v)\|)$, where

$$\theta_{11}^2\|W\|^2 = \eta = \theta_{11}\theta_v - \theta_1^2 = \det\begin{pmatrix} \theta_{11} & \theta_1 \\ \theta_1 & \theta_v \end{pmatrix} = \det\{XX^\top\}$$

by (18) and (19). Since $\det\{XX^\top\} = \lambda_1^2\lambda_2^2$, we recover the result of [5] that, for the case $m = 2$, the Radon–Nikodym derivative is

$$\frac{1}{2\pi}\frac{\|P(v)\|^2}{\lambda_1\lambda_2},$$

where the λ_i are the positive square roots of λ_i^2. Note that, in this case, for a given configuration, the Radon–Nikodym derivative depends only on the size of the projected configuration relative to that of the given configuration.

If w_1 and w_2 are the unit eigenvectors of $XX^\top$ corresponding to the eigenvalues λ_1^2 and λ_2^2, respectively, then $v = \cos s\, w_1 + \sin s\, w_2$ for some s, and the nonzero eigenvalue $\lambda(v)^2 = \theta_{11} = \|P(v)\|^2$ of $P(v)P(v)^\top$ is $\sin^2 s\, \lambda_1^2 + \cos^2 s\, \lambda_2^2$ and so the above Radon–Nikodym derivative can be expressed as

$$\frac{1}{2\pi}\left\{\frac{\lambda_1}{\lambda_2}\sin^2 s + \frac{\lambda_2}{\lambda_1}\cos^2 s\right\},$$

which, assuming that $\lambda_1 \geqslant \lambda_2$, clearly has the maximum value λ_1/λ_2 and the minimum value λ_2/λ_1. With the further assumption that $\lambda_1^2 + \lambda_2^2 = 1$, these two extreme values of the Radon–Nikodym derivative determine the two eigenvalues of $XX^\top$. Hence, it is possible to recover certain information on the shape of the initial given configuration by the nature of this Radon–Nikodym derivative.

The following result shows that, in contrast to the case $m = 2$, when $m = 3$, the Radon–Nikodym derivative depends, for a given configuration, on the eigenvalues of $P(v)P(v)^\top$, in addition to the dependence on the relative size of the projected configuration.

Theorem 2. *If v is uniformly distributed on the sphere $\mathbb{S}^2$ then, with respect to the volume element on $\phi(\mathbb{S}^2)$, the Radon–Nikodym derivative at $\phi(v)$ of the shape of the projection of the configuration with pre-size-and-shape X is given by*

$$\frac{1}{4\pi}\frac{\|P(v)\|^4}{\lambda_1^2\lambda_2^2 + \lambda_1^2\lambda_3^2 + \lambda_2^2\lambda_3^2 - \lambda_1(v)^2\lambda_2(v)^2},$$

where λ_1^2, λ_2^2, and λ_3^2 are the three eigenvalues of $XX^\top$, and $\lambda_1(v)^2$ and $\lambda_2(v)^2$ are the two nonzero eigenvalues of $P(v)P(v)^\top$.

Proof. When $m = 3$, C_3 is just $E_2 = \text{tr}(\Theta) = \|P(v)\|^2$ and so (18) becomes

$$\Psi = ((\langle W_i^h, W_j^h \rangle))_{2 \times 2} = \frac{\theta_v(\theta_{11} + \theta_{22}) - \theta_1^2 - \theta_2^2}{\|P(v)\|^4} I_2.$$

However, on the one hand, $\text{tr}\{(V^\top X X^\top V)^{-1}\} = \text{tr}\{(X X^\top)^{-1}\} = \lambda_1^{-2} + \lambda_2^{-2} + \lambda_3^{-2}$, where the λ_i^2 are the three eigenvalues of $X X^\top$, and, on the other hand,

$$\text{tr}\{(V^\top X X^\top V)^{-1}\} = \det\{(V^\top X X^\top V)^{-1}\}\text{tr}\{\text{adj}(V^\top X X^\top V)\}$$
$$= \det\{(X X^\top)^{-1}\}\{\theta_v\theta_{11} + \theta_v\theta_{22} + \theta_{11}\theta_{22} - \theta_1^2 - \theta_2^2\}$$

by (19), given our choice of u_1 and u_2 such that $\theta_{12} = 0$. Since $\det(X X^\top) = \lambda_1^2\lambda_2^2\lambda_3^2$, we have

$$\theta_v\theta_{11} + \theta_v\theta_{22} + \theta_{11}\theta_{22} - \theta_1^2 - \theta_2^2 = \lambda_1^2\lambda_2^2 + \lambda_1^2\lambda_3^2 + \lambda_2^2\lambda_3^2,$$

and then since, with $\theta_{12} = 0$, θ_{11} and θ_{22} are the nonzero eigenvalues $\lambda_1(v)^2$ and $\lambda_2(v)^2$ of $P(v)P(v)^\top$,

$$\|W_i^h(v)\|^2 = \frac{1}{\|P(v)\|^4}\{\lambda_1^2\lambda_2^2 + \lambda_1^2\lambda_3^2 + \lambda_2^2\lambda_3^2 - \lambda_1(v)^2\lambda_2(v)^2\},$$

giving the stated closed form for the Radon–Nikodym derivative of Theorem 1 when $m = 3$.

By writing the eigenvalues of $P(v)P(v)^\top$ as functions of v and the eigenvalues of $X X^\top$, it can be checked that, if the three eigenvalues of $X X^\top$ are in decreasing order, $\lambda_1^2 \geqslant \lambda_2^2 \geqslant \lambda_3^2$, then the maximum and minimum values of the Radon–Nikodym derivative are respectively $(\lambda_1^2 + \lambda_2^2)/\lambda_3^2$ and $(\lambda_2^2 + \lambda_3^2)/\lambda_1^2$. These two extreme values are sufficient to determine the three eigenvalues of $X X^\top / \|X\|^2$.

As m increases, the expression for the Radon–Nikodym derivative becomes rapidly more complex. For $m = 4$, we have the following result.

Theorem 3. *If v is uniformly distributed on the sphere $\mathbb{S}^3$ then, with respect to the volume element on $\phi(\mathbb{S}^3)$, the Radon–Nikodym derivative at $\phi(v)$ of the shape of the projection of the configuration in $\mathbb{R}^4$ with pre-size-and-shape X is given by*

$$\frac{1}{2\pi^2}\|P(v)\|^4\left\{\frac{\prod_{1 \leqslant i < j \leqslant 3}(\lambda_i(v)^2 + \lambda_j(v)^2)}{\sum_{1 \leqslant i < j \leqslant 4}\lambda_i^2\lambda_j^2 - \sum_{1 \leqslant i < j \leqslant 3}\lambda_i(v)^2\lambda_j(v)^2}\right\}^{1/2}$$

$$\times \left\{\left(\sum_{1 \leqslant i < j \leqslant 4}\lambda_i^2\lambda_j^2 - \sum_{1 \leqslant i < j \leqslant 3}\lambda_i(v)^2\lambda_j(v)^2\right)\left(\sum_{1 \leqslant i < j < \ell \leqslant 4}\lambda_i^2\lambda_j^2\lambda_\ell^2 - \prod_{i=1}^3\lambda_i(v)^2\right)\right.$$

$$\left. - (\|X\|^2 - \|P(v)\|^2)\prod_{i=1}^4\lambda_i^2\right\}^{-1/2},$$

where λ_i^2, $1 \leqslant i \leqslant 4$, are the four eigenvalues of $X X^\top$ and $\lambda_i(v)^2$, $1 \leqslant i \leqslant 3$, are the three nonzero eigenvalues of $P(v)P(v)^\top$.

Proof. We denote the matrix in (16) with $m = 4$ by Γ and the three eigenvalues of Γ by κ_1, κ_2, and κ_3. Since

$$\det\{L_4(\theta_1, \theta_2, \theta_3)\} = \det\begin{pmatrix} \theta_2 & -\theta_1 & 0 \\ \theta_3 & 0 & -\theta_1 \\ 0 & \theta_3 & -\theta_2 \end{pmatrix} = 0,$$

we have $\det(\Gamma) = 0$ and at least one eigenvalue, κ_3 say, is 0. Then

$$\Gamma = \begin{pmatrix} \dfrac{\theta_{33}}{\theta_{11}+\theta_{22}}\theta_2^2 + \dfrac{\theta_{22}}{\theta_{11}+\theta_{33}}\theta_3^2 & -\dfrac{\theta_{33}}{\theta_{11}+\theta_{22}}\theta_1\theta_2 & -\dfrac{\theta_{22}}{\theta_{11}+\theta_{33}}\theta_1\theta_3 \\[2ex] -\dfrac{\theta_{33}}{\theta_{11}+\theta_{22}}\theta_1\theta_2 & \dfrac{\theta_{33}}{\theta_{11}+\theta_{22}}\theta_1^2 + \dfrac{\theta_{11}}{\theta_{22}+\theta_{33}}\theta_3^2 & -\dfrac{\theta_{11}}{\theta_{22}+\theta_{33}}\theta_2\theta_3 \\[2ex] -\dfrac{\theta_{22}}{\theta_{11}+\theta_{33}}\theta_1\theta_3 & -\dfrac{\theta_{11}}{\theta_{22}+\theta_{33}}\theta_2\theta_3 & \dfrac{\theta_{22}}{\theta_{11}+\theta_{33}}\theta_1^2 + \dfrac{\theta_{11}}{\theta_{22}+\theta_{33}}\theta_2^2 \end{pmatrix}$$

and

$$\det(\eta I_3 - \Gamma) = \eta^3 - \eta^2 \mathrm{tr}(\Gamma) + \eta \kappa_1 \kappa_2. \tag{20}$$

Using the equalities

$$\{\theta_{ii}+\theta_{jj}\}\theta_{ii}\theta_{jj} = \|P(v)\|^2 \theta_{ii}\theta_{jj} - \prod_{\ell=1}^{3}\theta_{\ell\ell}$$

and

$$\{\theta_{ii}+\theta_{jj}\}^2 = \|P(v)\|^2\{\theta_{ii}+\theta_{jj}\} + \theta_{ii}\theta_{jj} - \sum_{1\leqslant \ell_1 < \ell_2 \leqslant 3} \theta_{\ell_1\ell_1}\theta_{\ell_2\ell_2}$$

for $i \neq j$, we can check that

$$\left\{\prod_{1\leqslant i<j\leqslant 3}(\theta_{ii}+\theta_{jj})\right\}\mathrm{tr}(\Gamma) = \left\{2\prod_{i=1}^{3}\theta_{ii} - \|P(v)\|^2\sum_{1\leqslant i<j\leqslant 3}\theta_{ii}\theta_{jj}\right\}\sum_{i=1}^{3}\theta_i^2$$
$$+ \|P(v)\|^4 \sum_{i\neq j\neq \ell}\theta_i^2\{\theta_{jj}+\theta_{\ell\ell}\} - \|P(v)\|^2\sum_{i\neq j\neq \ell}\theta_i^2\theta_{jj}\theta_{\ell\ell}$$

and that

$$\left\{\prod_{1\leqslant i<j\leqslant 3}(\theta_{ii}+\theta_{jj})\right\}\kappa_1\kappa_2 = \sum_{i=1}^{2}\theta_i^3\left\{\|P(v)\|^2\sum_{i\neq j\neq \ell}\theta_i^2\theta_{jj}\theta_{\ell\ell} - \prod_{i=1}^{3}\theta_{ii}\sum_{i=1}^{3}\theta_i^2\right\}.$$

Thus, using the relationships of θ_v and θ_i with the eigenvalues of $XX^\top$ and $P(v)P(v)^\top$ given in the previous section, detailed computation using (20) gives the stated result, since $2\pi^2$ is the volume of $\mathbb{S}^3$.

6. The induced distribution of the size-and-shape

Rather than considering the distribution of the shape of the projection of a given configuration induced by the uniform distribution on $\mathbb{S}^{m-1}$, one may be interested in that of the size-and-shape of the projection of a given configuration after it has been randomly rotated. Our previous arguments for the shape of such a projection can easily be adapted to the case of its size-and-shape. In the following, we shall merely outline the main steps of the corresponding results, rather than repeating them in full, since they are similar to those given in the previous sections.

Firstly, the map corresponding to (3) becomes

$$\tilde{\phi}: \mathbb{S}^{m-1} \to S\Sigma_{m-1}^{k}, \qquad v \mapsto \tilde{\pi}(\Pi(e_m)RX),$$

where $S\Sigma_{m-1}^{k}$ denotes the (Kendall) size-and-shape space of configurations in $\mathbb{R}^{m-1}$ with k labelled points and $\tilde{\pi}(\Pi(e_m)RX)$ denotes the size-and-shape of $\Pi(e_m)RX$. If

$$\tilde{\varphi}: \mathbb{R}_{m-1}^{k-1} = (\mathbb{R}^{m-1})^{k-1} \to S\Sigma_{m-1}^{k}$$

is the projection from the standard pre-size-and-shape space to the size-and-shape space, which is similarly a Riemannian submersion when it is restricted to the nonsingular part, then a similar argument to that for (4) gives analogously that $\tilde{\phi}(v(t)) = \tilde{\varphi}((\psi \circ R)(t))$ and so

$$\tilde{\phi}_*(\dot{v}(0)) = \tilde{\varphi}_*\left(\frac{\mathrm{d}(\psi \circ R)(t)}{\mathrm{d}t}\bigg|_{t=0}\right).$$

The result corresponding to that of Lemma 1 is then

$$\frac{\mathrm{d}\psi(R(t))}{\mathrm{d}t}\bigg|_{t=0} = R\Pi(v_R)\frac{\mathrm{d}P(v(t))}{\mathrm{d}t}\bigg|_{t=0}$$

and $\tilde{W}_i(v) = \Pi(v)\dot{P}(u_i) = -u_i v^\top X$. This in turn implies that the horizontal component of $\tilde{W}_i(v)$ is

$$\tilde{W}_i^h(v) = -\{u_i v^\top - U A^{(i)} U^\top\}X,$$

where the matrix $A^{(i)}$ is that determined by (6) as for the case of the corresponding shape of the projection of the given configuration and, consequently, the matrix $\tilde{\Psi} = (\langle \tilde{W}_i^h, \tilde{W}_j^h \rangle)$ simplifies slightly to become

$$\theta_v I_{m-1} - L_m(\boldsymbol{\theta})^\top C_m^{-1} L_m(\boldsymbol{\theta}),$$

where C_m and $L_m(\boldsymbol{\theta})$ are still those defined in Section 3. Thus, the Radon–Nikodym derivative of the size-and-shape of the projection of the given configuration at $\tilde{\phi}(v)$, induced by uniformly distributed $v \in \mathbb{S}^{m-1}$ is given by $1/\{\mathrm{vol}(\mathbb{S}^{m-1})\sqrt{\det(\langle \tilde{W}_i^h, \tilde{W}_j^h \rangle)}\}$, where $\det(\langle \tilde{W}_i^h, \tilde{W}_j^h \rangle)$ is equal to the characteristic polynomial of

$$L_m(\boldsymbol{\theta})^\top C_m^{-1} L_m(\boldsymbol{\theta})$$

evaluated at θ_v.

In particular, when $m = 3$,

$$\tilde{\Psi} = (\langle \tilde{W}_i^h(v), \tilde{W}_j^h(v) \rangle)_{2 \times 2} = \theta_v I_2 - \frac{1}{\theta_{11} + \theta_{22}}\begin{pmatrix} \theta_2^2 & -\theta_1\theta_2 \\ -\theta_1\theta_2 & \theta_1^2 \end{pmatrix}.$$

Hence, its determinant has the closed form given by

$$\frac{\|X\|^2 - \|P(v)\|^2}{\|P(v)\|^2}\{\lambda_1^2\lambda_2^2 + \lambda_1^2\lambda_3^2 + \lambda_2^2\lambda_3^2 - \lambda_1(v)^2\lambda_2(v)^2\}\left(= \theta_v\|P(v)\|^6\sqrt{\det(\langle W_i^h, W_j^h \rangle)}\right)$$

and so we have the following result.

Theorem 4. *The Radon–Nikodym derivative, with respect to the volume element on $\tilde{\phi}(\mathbb{S}^2)$, of the distribution of the size-and-shape of the projection of the given configuration, induced from the uniform distribution on $\mathbb{S}^2$, is*

$$\frac{\|P(v)\|}{4\pi\sqrt{\|X\|^2 - \|P(v)\|^2}}\frac{1}{\sqrt{\lambda_1^2\lambda_2^2 + \lambda_1^2\lambda_3^2 + \lambda_2^2\lambda_3^2 - \lambda_1(v)^2\lambda_2(v)^2}}.$$

Note that, when $m = 2$, the expression for the corresponding Radon–Nikodym derivative can be found similarly to be $1/(2\pi\|W(v))\|) = 1/(2\pi\sqrt{\theta_v}) = 1/(2\pi\sqrt{\|X\|^2 - \|P(v)\|^2})$.

References

[1] KENDALL, D. G. (1984). Shape manifolds, Procrustean metrics, and complex projective spaces. *Bull. London Math. Soc.* **16,** 81–121.

[2] KENDALL, D. G., BARDEN, D., CARNE, T. K. AND LE, H. (1999). *Shape and Shape Theory*. John Wiley, Chichester.

[3] KENDALL, W. S. AND LE, H. (2009). Statistical shape theory. In *New Perspectives in Stochastic Geometry*, eds W. S. Kendall and I. Molchanov, Oxford University Press, pp. 384–373.

[4] O'NEILL, B. (1983). *Semi-Riemannian Geometry*. Academic Press, New York.

[5] PANARETOS, V. M. (2006). The diffusion of Radon shape. *Adv. Appl. Prob.* **38,** 320–335.

[6] PANARETOS, V. M. (2008). Representation of Radon shape diffusions via hyperspherical Brownian motion. *Math. Proc. Camb. Phil. Soc.* **145,** 457–470.

[7] PANARETOS, V. M. (2009). On random tomography with unobservable projection angles. *Ann. Statist.* **37,** 3272–3306.