

A Bayesian Approach to Statistical Shape Analysis via the Projected Normal Distribution

Luis Gutiérrez*, Eduardo Gutiérrez-Peña† and Ramsés H. Mena‡

Abstract. This work presents a Bayesian predictive approach to statistical shape analysis. A modeling strategy that starts with a Gaussian distribution on the configuration space, and then removes the effects of location, rotation and scale, is studied. This boils down to an application of the projected normal distribution to model the configurations in the shape space, which together with certain identifiability constraints, facilitates parameter interpretation. Having better control over the parameters allows us to generalize the model to a regression setting where the effect of predictors on shapes can be considered. The methodology is illustrated and tested using both simulated scenarios and a real data set concerning eight anatomical landmarks on a sagittal plane of the *corpus callosum* in patients with autism and in a group of controls.

MSC 2010 subject classifications: Primary 62F15, 62H35; secondary 62J05.

Keywords: Bookstein coordinates, identifiability, medical image, shape regression.

1 Introduction

Kendall (1977) defined a shape as the geometric feature of an object that is invariant to rigid motions and global scaling. There has since been a growing literature devoted to developing both theory and methods for statistical shape analysis. Most approaches borrow traditional ideas from geometry and statistics in order to model and analyze shapes (e.g. Small, 1996), with a more recent trend focusing on techniques for closed curves, e.g. to describe the boundary of an object (Klassen et al., 2004; Srivastava et al., 2005; Kurtek et al., 2012; Cheng et al., 2015). These extensions go back to the original question: On which space should we define the probability models required for shape analysis?

There are at least three spaces on which models can be defined (Dryden and Mardia, 2016): 1) The **configuration** space, represented by the Euclidean space \mathbb{R}^{mp} , where configurations $\mathbf{Z}_1, \dots, \mathbf{Z}_n$ of m -dimensional shapes are depicted by p landmarks. This is the space where raw objects are represented. 2) The **pre-shape** space, typically a hypersphere of unit radius in $m(p-1)$ dimensions, where configurations are represented after translations and scalings have been removed. We denote the configurations in the

*Departamento de Estadística, Pontificia Universidad Católica de Chile, lrgutier@mat.uc.cl

†Departamento de Probabilidad y Estadística, IIMAS, Universidad Nacional Autónoma de México, eduardo@sigma.iimas.unam.mx

‡Departamento de Probabilidad y Estadística, IIMAS, Universidad Nacional Autónoma de México, ramses@sigma.iimas.unam.mx

pre-shape space by $\mathbf{Z}_1^\dagger, \dots, \mathbf{Z}_n^\dagger$. 3) The **shape** space, where configurations are invariant under location, rotation and scaling. We denote configurations in this space by $\mathbf{Z}_1^*, \dots, \mathbf{Z}_n^*$. One could go from the configuration space to the pre-shape space using the transformation $\mathbf{Z}^\dagger = \frac{H\mathbf{Z}}{\|H\mathbf{Z}\|}$, where H is the Helmert sub-matrix, and from the configuration space to the shape space via Procrustes Analysis (Gower, 1975; Goodall, 1991).

Models on the pre-shape space are widely used due to the availability of the complex elliptical family of distributions, which have the property of being rotation-invariant (see, e.g. Micheas et al., 2006). Popular members of this family are the complex Bingham distribution (Kent, 1994) and the complex Watson distribution (Mardia and Dryden, 1999), with the former being a conjugate family to the latter. See also Leu and Damien (2014) for a novel Bayesian analysis of the complex Bingham distribution, and Mardia et al. (2013) for a model for the alignment of two unlabeled landmark configurations.

The rotation-invariance property can be seen as a drawback for the statistical identification of the underlying parameters. The lack of identifiability may hinder the correct interpretation of the parameters and make a Bayesian approach difficult. Indeed, the parameters of these distributions are not explicitly related to the corresponding distributional moments; thus, for example, extensions to the regression setting are not straightforward. In order to overcome these drawbacks of shape analysis on the pre-shape space, we use a statistical model on the configuration space with the similarity transformations “integrated out”, i.e. a marginal approach (Dryden and Mardia, 2016).

An alternative way to remove similarity transformations is the use of Bookstein coordinates (Bookstein, 1986), which allows us to use well-known models on the configuration space and then project them onto the shape space. Following this idea, Kume and Welling (2010) proposed a maximum-likelihood estimation procedure; however, they were not concerned with the identifiability of the model. In this paper we undertake a Bayesian analysis of the projected-normal distribution on the shape space which considers the identifiability of the model and facilitates parameter interpretation. We also derive some new properties of the projected normal distribution and extend the model to consider the effect of predictors on shapes.

The rest of the paper is organized as follows. In Section 2 we introduce the projected normal distribution on the shape space and discuss both the data structure and the transformations, including Bookstein’s coordinates. Some useful properties of the projected normal distribution are presented in Section 3. Section 4 is concerned with the identifiability analysis. To this end, constraints on the parameter space and a strategy to fix some of the parameters using the observed shape variables are proposed. Based on these constraints and a suitable reparameterization, in Section 5 we develop a Gibbs sampling algorithm for posterior inference. In Section 6, we extend the model to a regression setting. Section 7 illustrates our proposal using simulated scenarios; in particular, a Monte Carlo study allows us to visualize the asymptotic behavior of our density estimator of the projected normal distribution. In Section 8 we fit the model to a data set consisting of eight anatomical landmarks on a sagittal plane of the *corpus callosum* in patients with autism and in a group of controls. Section 9 provides some concluding remarks.

2 The projected normal distribution on the shape space

The projected normal distribution (Dryden and Mardia, 1991), is obtained by the projection of a p -dimensional normal distribution onto a manifold $\mathcal{M} \subset \mathbb{R}^p$. A common example is the projection of a bivariate normal distribution on a circle, used for the analysis of directional data (Mardia and Jupp, 2000; Jammalamadaka and SenGupta, 2001). Let us describe the data structure required for shape analysis and the transformations needed to derive the projected normal distribution on the shape space.

Let \mathbf{Z} be a $p \times 2$ random configuration matrix, and denote by $\mathbf{X} = (x_1, y_1, \dots, x_p, y_p)^T$ the $2p$ -dimensional landmarks vector with elements from \mathbf{Z} . We assume $\mathbf{X} \sim N_{2p}(\nu, \Omega)$, with N_d denoting a d -dimensional Gaussian distribution. We remove the effect of location, rotation and scale from the landmarks using the method proposed by Bookstein (1986).

2.1 Location

In order to remove the location effect, we consider the linear transformation of \mathbf{X}

$$\tilde{\mathbf{X}} = L\mathbf{X} \quad \text{with} \quad L = \begin{pmatrix} I_2 & | & \mathbf{0}_{2 \times 2(p-1)} \\ \hline A & | & I_{2(p-1)} \end{pmatrix} \quad (1)$$

and

$$A = \begin{bmatrix} -1 & 0 & -1 & 0 & \dots & -1 & 0 \\ 0 & -1 & 0 & -1 & \dots & 0 & -1 \end{bmatrix}^T.$$

This transformation shifts all landmarks, except the first one, by the vector (x_1, y_1) . The random vector $\tilde{\mathbf{X}}$ is still normally distributed on R^{2p} , with mean $\tilde{\mu} = L\nu$ and covariance matrix $\tilde{\Sigma} = L\Omega L^T$. If, in (1), we shifted the first landmark too, then it would be fixed at the location $(0,0)$ yielding a singular normal distribution. Instead, we marginalize with respect to the first landmark (x_1, y_1) from $\tilde{\mathbf{X}}$; the corresponding $(p-1)$ shifted landmarks are denoted by $\mathbf{X}^* = (x_2^*, y_2^*, \dots, x_p^*, y_p^*)^T$, so $\mathbf{X}^* \in \mathbb{R}^{2(p-1)}$, where $x_i^* = x_i - x_1$ and $y_i^* = y_i - y_1$. Then $\mathbf{X}^* \sim N_{2(p-1)}(\mu, \Sigma)$, where $\mu = (\mu_{x_2}, \mu_{y_2}, \dots, \mu_{x_p}, \mu_{y_p})^T$ and Σ is the sub-matrix of $\tilde{\Sigma}$ corresponding to the last $2(p-1)$ coordinates of $\tilde{\mathbf{X}}$.

2.2 Rotation and scale

To remove the effect of rotation and scale, we apply the following transformation:

$$\begin{aligned} h_x = x_2^*, \quad u_i &= \frac{x_i^* x_2^* + y_i^* y_2^*}{x_2^{*2} + y_2^{*2}}, \quad i = 3, \dots, p, \\ h_y = y_2^*, \quad v_i &= \frac{y_i^* x_2^* - x_i^* y_2^*}{x_2^{*2} + y_2^{*2}}, \quad i = 3, \dots, p. \end{aligned} \quad (2)$$

Note that the second landmark remains the same. If the transformations defined for u_i and v_i ($i = 3, \dots, p$) were to be applied also to the second landmark, it would move it to the location $(1, 0)$ giving rise to a singular distribution on $R^{2(p-1)}$. The vector $\mathbf{U} = (u_3, v_3, \dots, u_p, v_p)^T$ is known as the Bookstein coordinates and it is in the shape space. Letting $\mathbf{H} = (h_x, h_y)^T$ and $\det(\mathbf{J}) = (h_x^2 + h_y^2)^{p-2}$, the distribution of (\mathbf{H}, \mathbf{U}) is given by

$$f(\mathbf{H}, \mathbf{U} \mid \mu, \Sigma) = \frac{|\Sigma|^{-\frac{1}{2}}}{\sqrt{(2\pi)^{2(p-1)}}} \exp \left\{ -\frac{1}{2} (\mathbf{X}^* - \mu)^T \Sigma^{-1} (\mathbf{X}^* - \mu) \right\} |\det(\mathbf{J})|, \quad (3)$$

where $\mathbf{X}^* = \mathbf{X}^*(\mathbf{H}, \mathbf{U}) = (h_x, h_y, u_3 h_x - v_3 h_y, v_3 h_x + u_3 h_y, \dots, u_p h_x - v_p h_y, v_p h_x + u_p h_y)^T$ is the inverse transformation. Finally, note that $\mathbf{X}^* = W(\mathbf{U})\mathbf{H}$, where

$$W(\mathbf{U}) = \begin{bmatrix} 1 & 0 & u_3 & v_3 & \cdots & u_p & v_p \\ 0 & 1 & -v_3 & u_3 & \cdots & -v_p & u_p \end{bmatrix}^T.$$

2.3 Projected normal distribution

Setting $\Gamma = \Gamma(\mathbf{U}) = [W(\mathbf{U})^T \Sigma^{-1} W(\mathbf{U})]^{-1}$, $\xi = \xi(\mathbf{U}) = \Gamma W(\mathbf{U})^T \Sigma^{-1} \mu$, and denoting with $N_2(\mathbf{H} \mid \xi, \Gamma)$ the joint density of (3) can be factorized as

$$f(\mathbf{H}, \mathbf{U} \mid \mu, \Sigma) = f(\mathbf{U} \mid \mu, \Sigma) f(\mathbf{H} \mid \mathbf{U}, \mu, \Sigma) \\ = \frac{|\Gamma|^{1/2}}{(2\pi)^{p-2} |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (\mu^T \Sigma^{-1} \mu - \xi^T \Gamma^{-1} \xi) \right\} N_2(\mathbf{H} \mid \xi, \Gamma) (h_x^2 + h_y^2)^{p-2}. \quad (4)$$

Integrating (4) over \mathbf{H} , we obtain the projected normal distribution with density

$$f(\mathbf{U} \mid \mu, \Sigma) = \frac{C(\mathbf{U}) |\Gamma(\mathbf{U})|^{1/2}}{(2\pi)^{p-2} |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (\mu^T \Sigma^{-1} \mu - \xi(\mathbf{U})^T \Gamma(\mathbf{U})^{-1} \xi(\mathbf{U})) \right\}, \quad (5)$$

where $\|\mathbf{H}\| = (h_x^2 + h_y^2)^{1/2}$ and $C(\mathbf{U}) = \int \|\mathbf{H}\|^{2(p-2)} N_2(\mathbf{H}; \xi(\mathbf{U}), \Gamma(\mathbf{U})) d\mathbf{H}$. We denote this distribution by $PN_{2(p-2)}(\mu, \Sigma)$.

p	$C(\mathbf{U})$
3	$g^{(0)} = \text{tr}(\Gamma) + \xi^T \xi$
4	$g^{(1)} + [g^{(0)}]^2 = 2[\text{tr}(\Gamma)^2 + 2\xi^T \Gamma \xi] + [\text{tr}(\Gamma) + \xi^T \xi]^2$
5	$g^{(2)} + 3g^{(1)}g^{(0)} + [g^{(0)}]^3 = 8[\text{tr}(\Gamma)^3 + 3\xi^T (\Gamma)^2 \xi] + 6[\text{tr}(\Gamma)^2 + 2\xi^T \Gamma \xi][\text{tr}(\Gamma) + \xi^T \xi] + [\text{tr}(\Gamma) + \xi^T \xi]^3$

Table 1: Values of $C(\mathbf{U})$ for different values of p .

Using Theorem 3.2b.2 in Mathai and Provost (1992), the value of $C(\mathbf{U})$ can be computed via

$$C(\mathbf{U}) = \left\{ \sum_{r_1=0}^{p-3} \binom{p-3}{r_1} g^{(p-3-r_1)} \sum_{r_2=0}^{r_1-1} \binom{r_1-1}{r_2} g^{(r_1-1-r_2)} \dots \right\}, \quad (6)$$

where $g^{(k)} = 2^k k! \{ \text{tr}(\Gamma(\mathbf{U}))^{k+1} + (k+1) \xi(\mathbf{U})^T \Gamma(\mathbf{U})^k \xi(\mathbf{U}) \}$, $k = 0, 1, 2, \dots$; Table 1 shows explicit values of $C(\mathbf{U})$ for different values of p .

3 Properties of the projected normal distribution

The next result gives conditions for the symmetry of the projected normal distribution.

Theorem 1. *Consider the following partition of the parameters, $\mu = (\mu_1; \mu_2)^T$ and*

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{pmatrix},$$

where μ_1 is a vector of dimension 2×1 and Σ_{11} is a matrix of dimension 2×2 . If $\Sigma_{11} = I_2$, $\Sigma_{12} = \mathbf{0}_{2 \times 2(p-2)}$ and $\mu_1 = \mathbf{0}_2$ or $\mu_2 = \mathbf{0}_{2(p-2)}$, then the projected normal distribution $PN_{2(p-2)}(\mu, \Sigma)$ is symmetric around $\mathbf{0}_{2(p-2)}$.

The proof of Theorem 1 is given in the Supplementary Material (Gutiérrez et al., 2018). Figure 1 shows some contours of the projected normal density for the third of the $p = 3$ landmarks (recall that the first two landmarks have been ‘removed’). In panel (a), the conditions of Theorem 1 are satisfied and the distribution is symmetric around $\mathbf{0}_2$; in this case $\mu_2 = (2, 1.5)^T$ and $\Sigma = I_4$, $\mu_1 = (0, 0)^T$. In panels (b), (c) and (d), asymmetric cases are illustrated. In panel (b), $\mu_1 = (3, 6)^T$, $\mu_2 = (2, 1.5)^T$ and $\Sigma = I_4$. For panel (c) the parameters are given by $\mu_1 = (0, 0)^T$, $\mu_2 = (0, 0)^T$, $\Sigma_{11} = I_2$,

$$\Sigma_{12} = \begin{pmatrix} 2 & 3 \\ 0.8 & 1 \end{pmatrix}, \quad \text{and} \quad \Sigma_{22} = \begin{pmatrix} 17 & -1 \\ -1 & 17 \end{pmatrix}.$$

In panel (d) the parameters are $\mu_1 = (3, 6)^T$, $\mu_2 = (2, 1.5)^T$, $\Sigma_{11} = I_2$, $\Sigma_{12} = \mathbf{0}_2$ and

$$\Sigma_{22} = \begin{pmatrix} 17 & -10 \\ -10 & 17 \end{pmatrix}.$$

The parameters affect the shape of the projected normal distribution in different ways. For example, in panel (b), the parameters μ_1 and μ_2 induce asymmetry. This asymmetry affects the tails of the distribution. In panel (c), the parameter Σ_{12} induces asymmetry near the mode of the distribution and the tails remain mostly unaffected. Thus, certain choices of the parameters μ and Σ_{12} induce different kinds of asymmetry, while the parameter Σ_{22} controls dispersion and orientation of the distribution.

We now exhibit an interesting relationship between the projected normal and the Student t distributions.

Theorem 2. *For $p = 3, 4, \dots$, let $d = 2(p - 1)$ and $q = 2(p - 2)$. If $\mathbf{U} \sim PN_q(\mathbf{0}_d, I_d)$, there exists a $a \in (0, 1]$ such that, $\tilde{\mathbf{U}} = a\mathbf{U}$ follows a q -dimensional Student t distribution. Specifically, $\tilde{\mathbf{U}} \sim t_q(\mu^a, \Sigma^a, \nu)$, with parameters $\mu^a = \mathbf{0}$, $\Sigma^a = \frac{a^2}{\nu} I_q$, and $\nu = 2$. When $p = 3$ the value of a is 1, while for $p = 4, 5, \dots$*

$$a = \left[\frac{(p-2)!}{c_j} \right]^{\frac{1}{2p-5}}, \quad \text{with } c_0 = 1 \text{ and } c_j = jc_{j-1} + (p-2)c_{j-1}, \text{ where } j = p-3.$$

The proof of Theorem 2 can be found in the Supplementary Material. For $p = 3$, Theorem 2 shows that \mathbf{U} follows a bivariate Student t distribution. In this case, $E(\mathbf{U}) =$

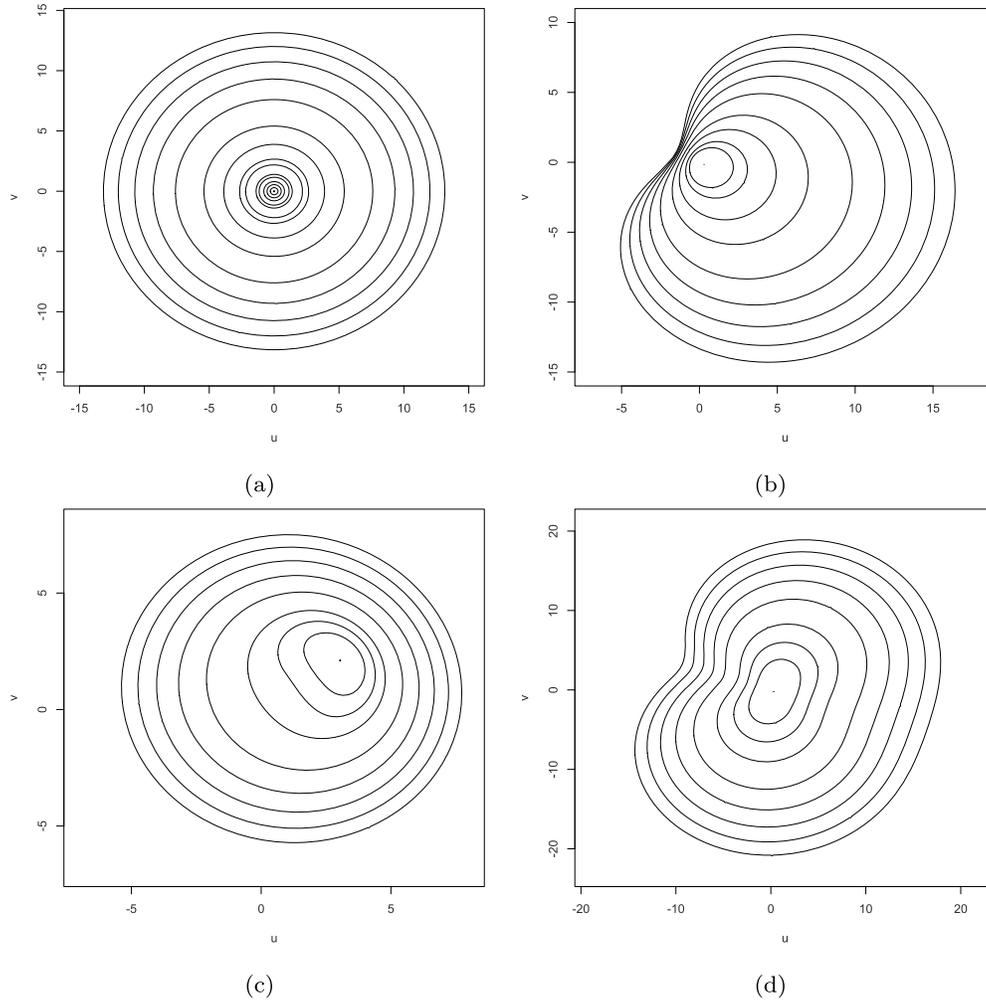


Figure 1: Contours of the projected normal distribution under different parameter values. In panel (a) the conditions of Theorem 1 are satisfied, then the distribution is symmetric. Panels (b), (c) and (d) show different kinds of asymmetry and dispersion. The asymmetry in (b) mainly affects the tails of the distribution while in (c) the region near to the mode is more affected. In (d) a different kind of dispersion combined with asymmetry is illustrated.

$\mathbf{0}_2$, while $\text{Var}(\mathbf{U})$ is undefined as there are only 2 degrees of freedom. For $p \geq 4$, the projected normal distribution with $\mu = \mathbf{0}_d$ and $\Sigma = \mathbf{I}_d$ is a rescaled version of a standard Student t distribution. Thus, when landmarks follow a normal distribution with $\mu = \mathbf{0}_d$ and $\Sigma = \mathbf{I}_d$ in the configuration space, the statistical shape analysis is reduced, after the Bookstein transformation, to that of a rescaled multivariate Student t distribution.

The invariance properties of the projected normal distribution make it appealing for statistical shape analysis. However, different sets of parameters (μ, Σ) can generate the same model, leading to lack of parameter identifiability. This is an important issue and we discuss it in some detail in the next section.

4 Identifiability analysis

The following result states in what sense the distribution (5) is not identifiable.

Proposition 1. *Let $r > 0$ be a positive constant and let R be a d -dimensional rotation matrix, where $d = 2(p - 1)$. Define $\mu_r = rR\mu$ and $\Sigma_r = r^2R\Sigma R^T$. Then*

$$f(\mathbf{U} \mid \mu_r, \Sigma_r) = f(\mathbf{U} \mid \mu, \Sigma). \tag{7}$$

The proof of Proposition 1 is given in the Supplementary Material. This result states that the projected normal distribution is invariant within the set of rescaled and/or rotated versions of μ and Σ , so that the actual parameter space is the equivalence class

$$\Theta = \{ (rR\mu, r^2R\Sigma R^T) \mid r \in \mathbb{R}^+, R \in \mathcal{SO}(2) \},$$

where $\mathcal{SO}(2)$ denotes the group of rotations in the plane. This clearly complicates the interpretability of the parameters and the comparability between different models.

The distribution on the left-hand side of (7) corresponds to the projection of the random vector $\mathbf{X}_r^* = rR\mathbf{X}^*$. If we apply transformation (2) to \mathbf{X}_r^* , the Bookstein coordinates \mathbf{U} remain the same; only \mathbf{H} is affected by the transformation. Thus the lack of identifiability is related to $\mathbf{H} = (x_2^*, y_2^*)^T$, which are unobserved variables that determine the scale and orientation of the rest of the coordinates of \mathbf{X}^* . In order to have an identifiable model, we constrain those parameters that control the location and variability of (x_2^*, y_2^*) . To this end, consider the following partitioning of the parameters

$$\mu = \begin{bmatrix} \mu_1 \\ \text{---} \\ \mu_2 \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & | & \Sigma_{12} \\ \text{---} & | & \text{---} \\ \Sigma_{21} & | & \Sigma_{22} \end{pmatrix}, \tag{8}$$

where the 2×1 vector μ_1 and the 2×2 matrix Σ_{11} are fixed and the remaining parameters are estimated. We fix the value of Σ_{11} as in Hernandez-Stumpfhauser et al. (2016), where an identifiability constraint is derived for a p -dimensional projected normal distribution with support on the unit hypersphere.

Let Σ_{11} and Σ_{22}^* be positive-definite matrices and let γ be a matrix of dimension $2(p - 2) \times 2$ with columns γ_1, γ_2 . Consider the following reparameterization

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \gamma^T \\ \gamma & \Sigma_{22}^* + \gamma\gamma^T \end{pmatrix}. \tag{9}$$

Proposition 2. *The constraint $\Sigma_{11} = I_2$ in (9) yields a positive definite matrix Σ .*

The proof of Proposition 2 is given in the Supplementary Material. We fix the value of μ_1 by means of the procedure described in the following remark.

Remark 1. The value of μ_1 can be fixed at $\mu_1^T = (\bar{x}_3^*, \bar{y}_3^*)^T$, where \bar{x}_3^* and \bar{y}_3^* are the sample means of the inverse transformations given by $\bar{x}_3^* = \bar{u}_3 x_2^* - \bar{v}_3 y_2^*$ and $\bar{y}_3^* = \bar{u}_3 y_2^* + \bar{v}_3 x_2^*$, for values of x_2^* and y_2^* obtained under the constraint $\text{Var}([x_3^*, y_3^*]) \approx I_2$. This result follows from the constraint $\Sigma_{11} = I_2$, since any procedure to fix the value of μ_1 based on the observed shape variables must consider the constraint imposed on Σ_{11} .

The values of x_2^* and y_2^* can be obtained with a simple optimization algorithm. See Algorithm 1 and the corresponding discussion in the Supplementary Material. Once μ_1 and Σ_{11} have been fixed, we are in a position to derive a Bayesian approach for inference on the parameters μ_2 , Σ_{22}^* and γ .

5 Posterior inference

Based on the reparameterization (9), we have $\Theta = (\mu_2, \Sigma_{22}^*, \gamma)$. Posterior inference on Θ from the marginal density (5) is not straightforward, but inferences based on the joint density (3) of (\mathbf{H}, \mathbf{U}) are relatively simple provided that \mathbf{H} is regarded as a latent variable. Thus, the augmented model has density

$$f(\mathbf{H}_i, \mathbf{U}_i \mid \mu, \Sigma) = N_{2(p-1)}(\mathbf{X}_i^* \mid \mu, \Sigma) (h_{x_i}^2 + h_{y_i}^2)^{p-2}, \quad i = 1, \dots, n \quad (10)$$

where, $\mathbf{X}^* = \mathbf{X}^*(\mathbf{H}, \mathbf{U})$. The model is completed with the conjugate prior specification

$$\begin{aligned} (\mu_2 \mid \mu_{02}, \Sigma_{\mu_2}) &\sim N_{2(p-2)}(\mu_{02}, \Sigma_{\mu_2}), \\ (\gamma_1, \gamma_2 \mid \mu_\gamma, \Sigma_\gamma) &\stackrel{\text{iid}}{\sim} N_{2(p-2)}(\mu_\gamma, \Sigma_\gamma), \\ (\Sigma_{22} \mid \phi, A) &\sim IW_{2(p-2)}(\phi, A). \end{aligned} \quad (11)$$

To derive a Gibbs sampling algorithm for model (10), first consider a partition of the vector $\mathbf{X}^* = (\mathbf{X}_1^*, \mathbf{X}_2^*)^T$, where the dimensions of \mathbf{X}_1^* and \mathbf{X}_2^* are 2×1 and $2(p-2) \times 1$ respectively. Now, using the reparameterization (9) and the identifiability constraints on Σ_{11} and μ_1 , the joint density function of (3) can be factorized as

$$f(\mathbf{H}, \mathbf{U} \mid \mu, \Sigma) = N_{2(p-2)}(\mathbf{X}_2^* \mid \mu_2 + \gamma(\mathbf{X}_1^* - \mu_1), \Sigma_{22}^*) N_2(\mathbf{X}_1^* \mid \mu_1, I_2) (h_x^2 + h_y^2)^{p-2}. \quad (12)$$

Using (12), we can now derive the following Gibbs sampler algorithm.

Algorithm 2:

Updating μ_2 : The full conditional for μ_2 is $(\mu_2 \mid \dots) \sim N_{2(p-2)}(\tilde{\mu}_2, \tilde{\Sigma}_\mu)$, where

$$\tilde{\Sigma}_\mu = [n\Sigma_{22}^{-1} + \Sigma_{\mu_2}^{-1}]^{-1} \quad \text{and} \quad \tilde{\mu}_2 = \tilde{\Sigma}_\mu \left[(\Sigma_{22}^*)^{-1} \sum_{i=1}^n [\mathbf{X}_{2i}^* - \gamma(\mathbf{X}_{1i}^* - \mu_1)] + \Sigma_{\mu_2}^{-1} \mu_{02} \right].$$

Updating Σ_{22}^ :* This block of the covariance matrix follows a $2(p-2)$ -dimensional Inverse Wishart distribution, i.e. $(\Sigma_{22}^* \mid \dots) \sim IW_{2(p-2)}(\phi + n, S_{22}^* + A)$, where

$$S_{22}^* = \sum_{i=1}^n [\mathbf{X}_{2i}^* - \mu_2 - \gamma(\mathbf{X}_{1i}^* - \mu_1)][\mathbf{X}_{2i}^* - \mu_2 - \gamma(\mathbf{X}_{1i}^* - \mu_1)]^T.$$

Updating γ : $(\gamma_1 | \dots) \sim N_{2(p-2)}(\tilde{\gamma}_1, \tilde{\Sigma}_{\gamma_1})$ and $(\gamma_2 | \dots) \sim N_{2(p-2)}(\tilde{\gamma}_2, \tilde{\Sigma}_{\gamma_2})$, where

$$\begin{aligned} \tilde{\Sigma}_{\gamma_1} &= \left[\sum_{i=1}^n (y_{11i}^* - \mu_{11})^2 (\Sigma_{22}^*)^{-1} + \Sigma_{\gamma}^{-1} \right]^{-1}, \\ \tilde{\Sigma}_{\gamma_2} &= \left[\sum_{i=1}^n (y_{12i}^* - \mu_{12})^2 (\Sigma_{22}^*)^{-1} + \Sigma_{\gamma}^{-1} \right]^{-1}, \\ \tilde{\gamma}_1 &= \tilde{\Sigma}_{\gamma_1} \left[(\Sigma_{22}^*)^{-1} \sum_{i=1}^n \{ (y_{11i}^* - \mu_{11}) [\mathbf{X}_{2i}^* - \mu_2 - (y_{12i}^* - \mu_{12}) \gamma_2] \} + \Sigma_{\gamma}^{-1} \mu_{\gamma} \right] \quad \text{and} \\ \tilde{\gamma}_2 &= \tilde{\Sigma}_{\gamma_2} \left[(\Sigma_{22}^*)^{-1} \sum_{i=1}^n \{ (y_{12i}^* - \mu_{12}) [\mathbf{X}_{2i}^* - \mu_2 - (y_{11i}^* - \mu_{11}) \gamma_1] \} + \Sigma_{\gamma}^{-1} \mu_{\gamma} \right]. \end{aligned}$$

Updating the latent variables \mathbf{H}_i , $i = 1, \dots, n$: The full conditional posterior distribution of \mathbf{H} is proportional to $N_2(\mathbf{H} | \xi, \Gamma) (h_x^2 + h_y^2)^{p-2}$; see (4). This distribution cannot be sampled directly, hence we use a Metropolis-Hastings step. The proposal distribution is given by $\mathbf{H}^* \sim N_2(\mathbf{H}^{(t-1)}, aI_2)$ with $a > 0$. Thus, $\mathbf{H}^{(t)}$ is sampled as:

$$\mathbf{H}^{(t)} = \begin{cases} \mathbf{H}^* & \text{with probability } \min(\alpha, 1), \\ \mathbf{H}^{(t-1)} & \text{otherwise,} \end{cases} \quad (13)$$

where

$$\alpha = \frac{N_2(\mathbf{H}^* | \xi, \Gamma) (h_x^{*2} + h_y^{*2})^{p-2}}{N_2(\mathbf{H}^{(t-1)} | \xi, \Gamma) \left((h_x^{(t-1)})^2 + (h_y^{(t-1)})^2 \right)^{p-2}}.$$

6 Extension to shape regression

We now propose an extension of the projected normal distribution to the shape regression setting, where we include the effect of predictors through the parameter μ . Specifically, we let

$$\mathbf{X}^* \sim N_{2(p-1)}(\mu_{\mathbf{z}}, \Sigma), \quad (14)$$

where $\mu_{\mathbf{z}} = (\mu, \beta \mathbf{z})^T$, β is a $2(p-2) \times q$ matrix of regression coefficients and \mathbf{z} is a q -dimensional vector of predictors. The conjugate prior (11) can be readily extended to this regression setting. The model (14) considers the identifiability constraint on Σ and μ ; thus, only the last $p-2$ coordinates of μ are affected by the value of the predictor \mathbf{z} . In this case, instead of updating μ_2 in Algorithm 2, we have to update β . The posterior distribution for the j -th column of β is given by

$$\beta_j \sim N_{2(p-2)}(\tilde{\beta}_j, \tilde{\Sigma}_{\beta_j}), \quad (15)$$

where $\tilde{\Sigma}_{\beta_j} = [\sum_{i=1}^n z_{ji}^2 (\Sigma_{22}^*)^{-1} + \Sigma_{\beta}^{-1}]^{-1}$ and

$$\tilde{\beta}_j = \tilde{\Sigma}_{\beta_j} \left[\sum_{i=1}^n z_{ji} (\Sigma_{22}^*)^{-1} [\mathbf{X}_{2i}^* - \sum_{k \neq j} z_{ki} \beta_k - \gamma (\mathbf{X}_{1i}^* - \mu_1)] + \Sigma_{\beta}^{-1} \beta_0 \right].$$

7 Examples with simulated data

Here we provide two illustrations and a simulation study. The first example illustrates the estimation process without predictors, whereas the second considers the regression model. For the first exercise we simulated random triangles, each described using three landmarks.

We considered two cases: Scenario 1, where random triangles were simulated from a normal distribution in the configuration space with mean $\mu_1 = (1, -1, -3, 0.5, 3.5, 5)$ and covariance matrix $\Sigma_1 = 0.2I_6$; and Scenario 2, where triangles were simulated using a normal distribution with mean $\mu_2 = \mu_1$ and covariance matrix $\Sigma_2 = \{B_1, B_2, B_3\}$, where

$$B_1 = \begin{pmatrix} 0.25 & -0.2 \\ -0.2 & 0.23 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0.2 & 0.18 \\ 0.18 & 0.23 \end{pmatrix} \quad \text{and} \quad B_3 = \begin{pmatrix} 0.15 & -0.13 \\ -0.13 & 0.19 \end{pmatrix}.$$

The difference between these scenarios is the non-null correlation between the x and y coordinates of the landmarks in Scenario 2. Figure 2 shows the landmarks together with the corresponding Bookstein coordinates. The algorithm was coded in R (R Core Team, 2016).

Figure 3 shows the true and estimated densities for Scenarios 1 and 2 under vague priors and illustrates the effect of the correlation on the landmarks in Scenario 2, where the density becomes asymmetric and deformed. The density is well estimated with a sample size of 300. To study the impact of a small sample size on the estimation of the projected normal distribution, we performed a Monte Carlo study. Specifically, we simulated from the true densities of Scenarios 1 and 2, using sample sizes ranging from 30 to 300. The Hellinger distance and the Kullback–Leibler divergence between the true and the estimated densities were calculated for each sample. Figure 4 shows how these metrics decrease to zero as the sample size increases.

The second illustration is designed to assess the changes in the projected normal distribution as the predictor varies. It also allows us to explore how the model discriminates between groups. We considered a predictor z with two levels and affecting three landmarks. Specifically, z takes values in the set $\{1 : \text{sick}, 0 : \text{healthy}\}$. Here, $\mu_i = \beta_0 + \beta_1 z_i$, where $\beta_0 = (1, -1, -3, 0.5, 3.5, 5)^t$ and $\beta_1 = (0.2, 0.95, -0.8, 0.95, -2.3, 2.1)^t$. For each level of the predictor we assumed $n_1 = n_2 = 150$. We set $\Sigma = \{A_1, A_2, A_3\}$, where

$$A_1 = \begin{pmatrix} 0.2 & -0.13 \\ -0.13 & 0.2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0.2 & 0.16 \\ 0.16 & 0.2 \end{pmatrix} \quad \text{and} \quad A_3 = \begin{pmatrix} 0.21 & -0.17 \\ -0.17 & 0.21 \end{pmatrix}.$$

Figure 5 shows the true and estimated densities for sick (level 1 of the predictor) and healthy (level 2). Location, shape and orientation, are all affected by the predictor. The estimated densities are very similar to the true ones.

Note that we can classify individuals into sick and healthy, for example based on the posterior probability

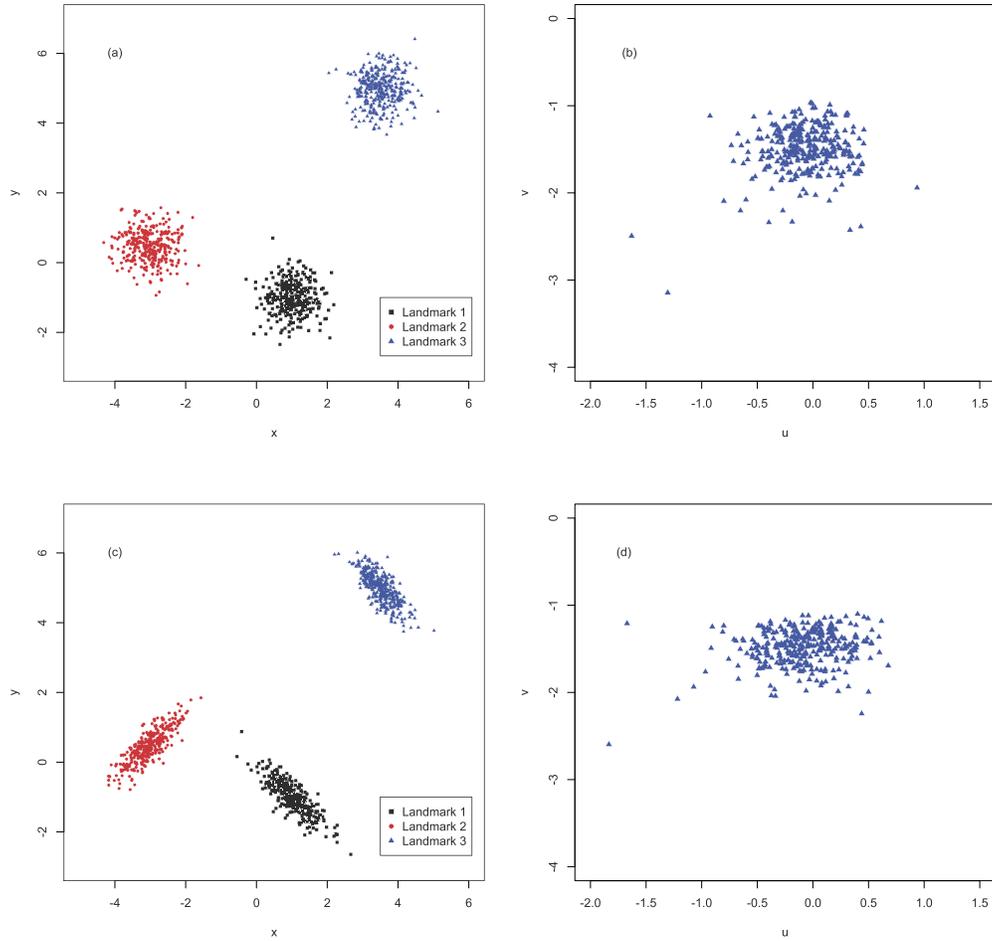


Figure 2: In Scenario 1, three uncorrelated landmarks in the configuration space were simulated; panel (a). After the Bookstein transformation, using Landmark 1 and 2, the shape information is represented in panel (b). In Scenario 2, the x and y coordinates for each landmark are correlated; panel (c). After the transformation, the shape information is displayed in panel (d).

$$\mathbb{P}(\text{sick} \mid \mathbf{U}) = \frac{\pi(\text{sick})f(\mathbf{U} \mid z = 1)}{\pi(\text{sick})f(\mathbf{U} \mid z = 1) + \pi(\text{healthy})f(\mathbf{U} \mid z = 0)}.$$

Table 2 shows the results of this classification using the prior probabilities $\pi(\text{sick}) = \pi(\text{healthy}) = 0.5$. Using the simple rule “classify the individual as sick if and only if $\mathbb{P}(\text{sick} \mid \mathbf{U}) > 0.5$ ”, the false-positive rate was 16%, while the false-negative rate was 14.7%. Thus, the shape regression model can be used as a classification tool.

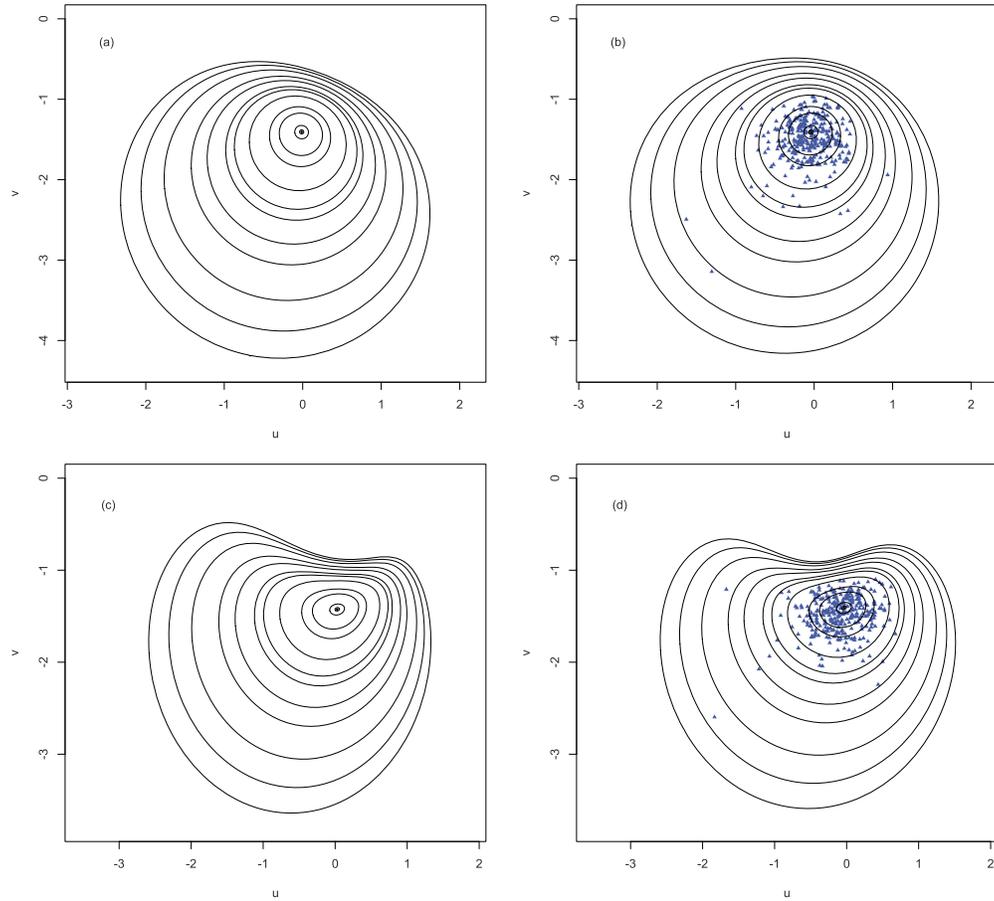


Figure 3: (a) True density for Scenario 1. (b) Estimated density for Scenario 1. (c) True density for Scenario 2. (d) Estimated density for Scenario 2.

8 Application

The *corpus callosum* is a wide flat bundle of neural fibers beneath the cortex of the brain at the longitudinal fissure. Its function is to connect the left and right cerebral hemispheres and facilitates inter-hemispheric communication. Many studies have reported a reduction in the size or volume of the *corpus callosum* in patients diagnosed with Autism Spectrum Disorder (ASD), see, e.g., Piven et al. (1997), Hardan et al. (2000), Brambilla et al. (2003), Vidal et al. (2006), Hardan et al. (2009) and Prigge et al. (2013). However, only a few of these studies have considered possible differences in the shape of the *corpus callosum* between patients with autism and healthy individuals. Indeed, potential differences in shape might reveal focal abnormalities helping to pinpoint which cortical regions are involved in the pathophysiology of autism (c.f. Casanova et al., 2011).

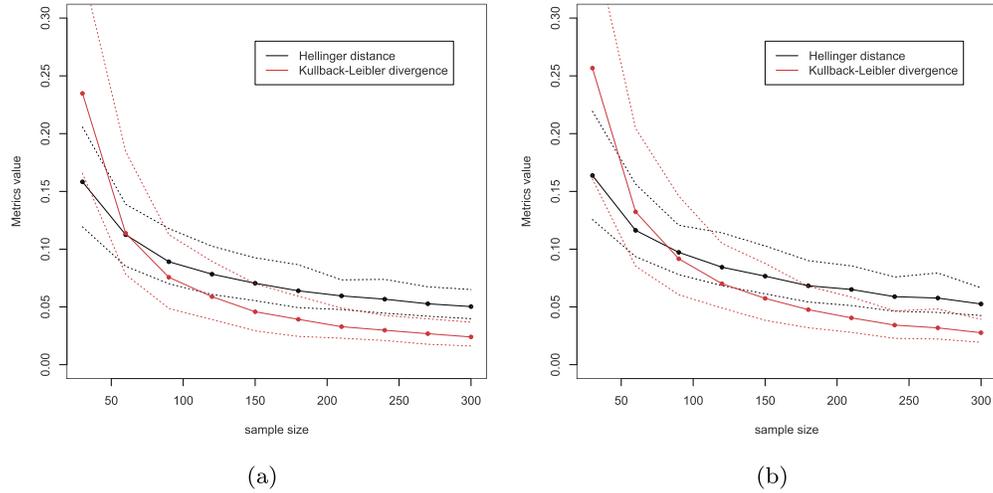


Figure 4: Monte Carlo estimation of the Hellinger distance and Kullback–Leibler divergence based on 100 replicates. Mean values (continuous lines), 2.5% and 97.5% percentiles (dotted lines): (a) Scenario 1; (b) Scenario 2.

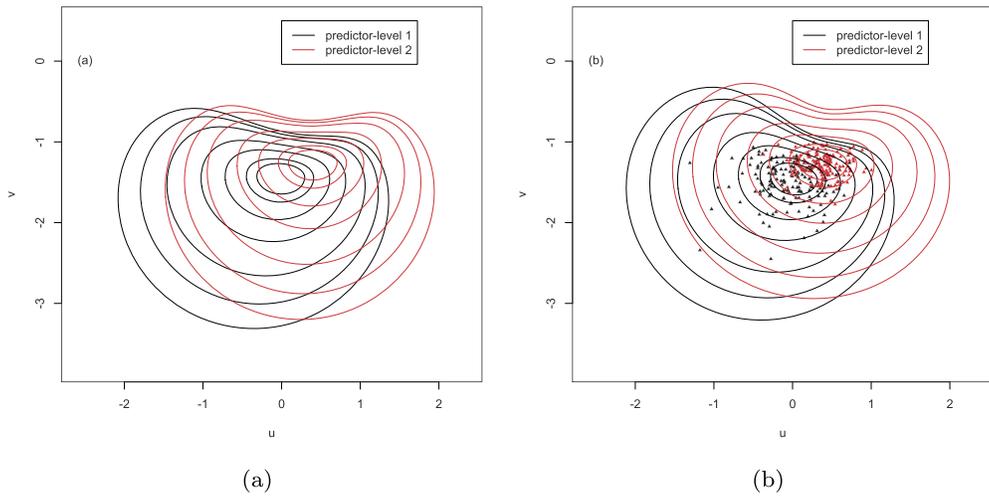


Figure 5: (a) True densities for the two levels of the predictor. (b) Estimated densities.

With this in mind, we produced a dataset consisting of eight anatomical landmarks that describe the shape of the *corpus callosum*, both in children diagnosed with autism and in a control group. For our study, the landmarks were located in the sagittal plane view of the *corpus callosum* using magnetic resonance images (MRI) of the brain. These MRI images were obtained from the public repository ABIDE (Autism Brain Imaging Data Exchange) (Di Martino et al., 2014) located at <http://www.loni.usc.edu>.

Health condition	Healthy	Sick	% Correct classification
Healthy	126	24	84.0%
Sick	22	128	85.3%
Total			84.7%

Table 2: Classification rates.

The MRI images were downloaded and stored in NifTI format, and then loaded into R using the `oro.nifti` library (Whitcher et al., 2011). Once in R, the images were explored using the `EBImage` library (Pau et al., 2010). During this exploration, a trained technician – working under the supervision of a radiologist – selected a sagittal view of the *corpus callosum* for each patient. Each sagittal view was saved as an image and loaded into the `StereoMorph` library (Oslen and Westneat, 2015), with the purpose of locating and digitalizing the landmarks. Figure 6 shows the sagittal plane view of the *corpus callosum* of a typical patient, together with the location of the eight anatomical landmarks. Locating the landmarks can be a difficult task, since there is considerable variability between the patients (see Figure S.1 in the Supplementary Material).

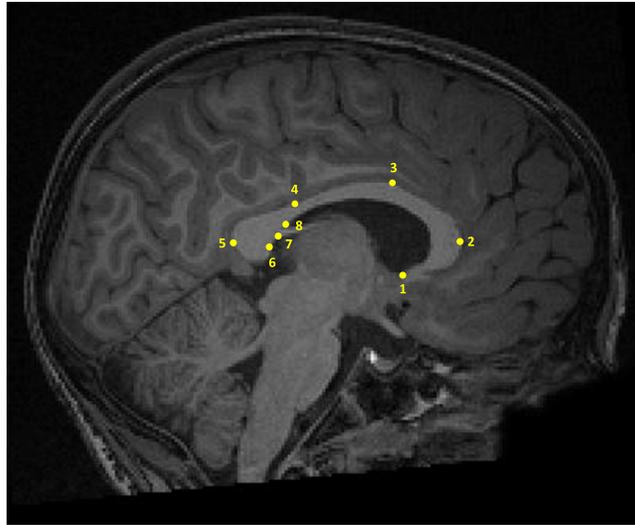
Figure 6: Location of eight anatomical landmarks on the sagittal view of the *corpus callosum*.

Table 3 shows the sample sizes for each level of disease and sex. The ages of the patients range from 6 to 13 years.

We used landmarks 1 and 2 to perform the Bookstein transformation, and then fitted models that included disease, sex and age as predictors; we also included some interactions. In particular, we considered the following three models: *model 1*, including the main effects only; *model 2*, which considers the main effects and the interaction

	Male	Female	Total
Autism	153	24	177
Control	145	37	182
Total	298	61	359

Table 3: Sample size for each condition.

between disease and age; and *model 3*, which includes the main effects and the interaction between sex and disease. In all three models we considered the reference-cell parametrization, i.e. the parameter β_1 represents the baseline given by female patients of the control group; β_2 quantifies the effect of the disease (1 : yes, 0 : no); β_3 corresponds to gender, coded as (1 : male, 0 : female). Additionally, in *model 2*, β_4 quantifies the effect of the interaction disease \times gender; similarly, in *model 3*, β_4 represents the effect of the interaction sex \times disease. In each of these three models we used vague priors on all of the parameters.

Landmark	Disease β_2			Sex β_3			Age β_4		
	mean	2.5%	97.5%	mean	2.5%	97.5%	mean	2.5%	97.5%
u_3	-0.119	-0.357	0.126	-0.242	-0.534	0.055	0.039	-0.031	0.115
v_3	-0.013	-0.346	0.339	-0.043	-0.524	0.448	-0.201	-0.312	-0.100
u_4	-0.035	-0.374	0.315	-0.195	-0.587	0.235	0.108	-0.001	0.236
v_4	-0.137	-0.505	0.267	0.353	-0.164	0.932	-0.201	-0.316	-0.066
u_5	-0.016	-0.417	0.410	-0.091	-0.569	0.416	0.247	0.132	0.432
v_5	-0.036	-0.532	0.446	0.235	-0.346	0.872	-0.172	-0.310	-0.014
u_6	-0.028	-0.363	0.307	-0.219	-0.628	0.210	0.168	0.074	0.298
v_6	-0.085	-0.488	0.317	0.356	-0.143	0.899	-0.171	-0.283	-0.039
u_7	-0.075	-0.358	0.225	-0.200	-0.545	0.168	0.129	0.045	0.247
v_7	0.05	-0.349	0.442	0.261	-0.244	0.824	-0.194	-0.309	-0.068
u_8	-0.075	-0.327	0.176	-0.141	-0.437	0.183	0.085	0.016	0.176
v_8	-0.142	-0.556	0.279	0.278	-0.236	0.836	-0.139	-0.261	-0.027

Table 4: Posterior means and credible sets for β_2 , β_3 and β_4 in *model 1*.

For model comparison, we computed the Conditional Predictive Ordinates (CPO_i) (Chen et al., 2000), summarized in the log-pseudo marginal likelihood statistic $LPML = \sum_{i=1}^n \log(CPO_i)$ (Geisser and Eddy, 1979). The values of the $LPML$ for each model were -1555.39 , -1559.167 and -1550.228 for models 1, 2 and 3, respectively. Even though *model 3* attained the highest $LPML$, the 95% credible intervals for the sex-disease interaction coefficients suggest these terms can be removed from the model. Since there does not seem to be an effect of the interaction terms for either of the models 2 and 3, we opted for *model 1* which has the second highest $LPML$. The posterior mean for each of the parameters, together with the corresponding credible intervals for *model 1*, are shown in Table 4. These results suggest that age does have an effect on the shape of the *corpus callosum* for landmarks 3 to 8. On the other hand, all the credible intervals for the coefficients of disease and sex include the value zero, suggesting a lack of effect of those variables on the shape of the *corpus callosum*.

With the aim of visualizing the distributions of some of the landmarks in the shape space, we computed the marginal posterior predictive distributions with *model 1*. Figure 7 shows these predictive distributions for comparison purposes.

The distributions of cases and controls for females and males all look quite similar, while the predictive distributions for the female-control and male-autism groups show small differences. As expected, larger differences in the shape of the *corpus callosum* are observed with changes in age. Furthermore, our results agree with those found in He et al. (2010), who used nine anatomical landmarks to compare the shape of the *corpus callosum* in autistic patients and controls. That said, the results for our data set are not conclusive so as to highlight significant differences in the shape due to the disease. Indeed, He et al. (2010) did not find significant differences in their studies based on landmarks. However, by comparing the boundaries of the *corpus callosum* after automatic or semi-automatic segmentation in the MRI images, they found significant differences in the distance between the interior genu and the posterior-most section of the *corpus callosum*. Finally, it is important to mention that our regression model allows us to perform comparisons controlling the effect of covariates such as sex and age. Here, age is an important factor that helps explain the observed differences in the shape of the *corpus callosum*.

9 Discussion

Our approach offers a simple way to perform statistical shape analysis. The definition of a probability model in the configuration space and its projection onto the shape space allows us to work with a well-known model in the configuration space. This approach, together with the constraints to attain identifiability, facilitates the Bayesian analysis of the model and the inclusion of predictors. In the particular case of the projection of a standard normal distribution, we have found connections with the Student t distribution that gives us some insight about the tail behavior of the projected normal distribution.

Typically, the parameters of a projected distribution have a different interpretation than in the original space. In our case, μ (representing the mean in the configuration space) becomes an asymmetry parameter in the shape space and also affects the tails of the distribution. Similarly, Σ (the variance-covariance matrix in the configuration space) represents the dispersion in the shape space and also affects the symmetry of the distribution. In our extension to the regression case, we are indexing the parameter μ with the predictors. Thus, the tails of the projected normal distribution are more affected by the changes in the predictors than other aspects of the distribution.

Even though the use of the Bookstein coordinates induces correlation in the shape variables, in our approach we are estimating the full covariance structure; thus, it is not really a problem. The proposed strategy can also be applied to other distributions on the configuration space and their projections onto the shape space. We are currently studying the use of infinite mixture models on the configuration space, which would allow us to model data sets of landmarks with heavy tails or asymmetries in the configuration space.

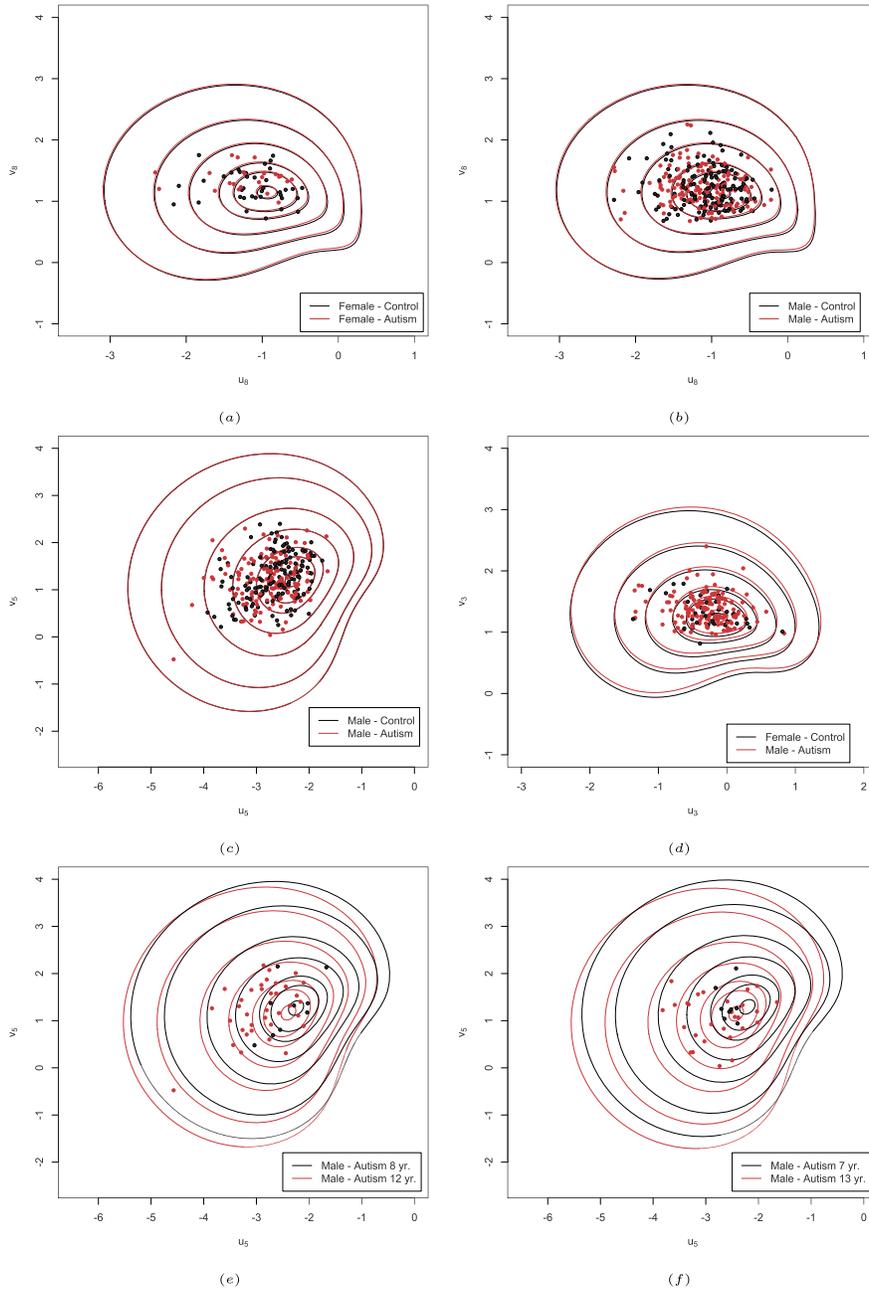


Figure 7: Marginal posterior predictive distributions: (a) female control vs female autism for landmark 8; (b) male control vs male autism for landmark 8; (c) male control vs male autism for landmark 5; (d) female control vs male autism for landmark 3; (e) male autism 8 yr. vs male autism 12 yr.; (f) male autism 7 yr. vs male autism 13 yr.

Supplementary Material

Supplementary Material for ‘A Bayesian approach to statistical shape analysis via the projected normal distribution’ (DOI: [10.1214/18-BA1113SUPP](https://doi.org/10.1214/18-BA1113SUPP); .pdf). The online Supplementary Material contains the proofs of Theorems 1 and 2, as well as those of Propositions 1 and 2. It also contains Algorithm 1 from Section 4 and some images relating to the application described in Section 8.

References

- Bookstein, F. (1986). “Size and shape spaces for landmark data in two dimensions.” *Statistical Science*, 1(2): 181–242. [428](#), [429](#)
- Brambilla, P., Hardan, A., di Nemi, S. U., Perez, J., Soares, J. C., and Barale, F. (2003). “Brain anatomy and development in autism: review of MRI studies.” *Brain Research Bulletin*, 61: 557–569. [438](#)
- Casanova, M. F., El-Baz, A., Elnakib, A., Switala, A. E., Williams, E. L., Williams, D. L., Minshew, N. J., and Conturo, T. E. (2011). “Quantitative analysis of the shape of the corpus callosum in patients with autism and comparison individuals.” *Autism*, 15(2): 223–238. [438](#)
- Chen, M. H., Shao, Q. M., and Ibrahim, J. C. (2000). *Monte Carlo Methods in Bayesian Computation*. Springer, New York. [MR1742311](#). doi: <https://doi.org/10.1007/978-1-4612-1276-8>. [441](#)
- Cheng, W., Dryden, L., and Huang, X. (2015). “Bayesian registration of functions and curves.” *Bayesian Analysis*. <http://projecteuclid.org/euclid.ba/1433162661>. [MR3471998](#). doi: <https://doi.org/10.1214/15-BA957>. [427](#)
- Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., Anderson, J. S., Assaf, M., Bookheimer, S. Y., Dapretto, M., Deen, B., Delmonte, S., Dinstein, I., Ertl-Wagner, B., Fair, D. A., Gallagher, L., Kennedy, D. P., Keown, C. L., Keysers, C., Lainhart, J. E., Lord, C., Luna, B., Menon, V., Minshew, N. J., Monk, C. S., Mueller, S., Muller, R.-A., Nebel, M. B., Nigg, J. T., O’Hearn, K., Pelphrey, K. A., Peltier, S. J., Rudie, J. D., Sunaert, S., Thioux, M., Tyszka, J. M., Uddin, L. Q., Verhoeven, J. S., Wenderoth, N., Wiggins, J. L., Mostofsky, S. H., and Milham, M. P. (2014). “The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism.” *Molecular Psychiatry*, 19(6): 659–667. [439](#)
- Dryden, I. and Mardia, K. (1991). “General shape distributions in a plane.” *Advances in Applied Probability*, 23(2): 259–276. [MR1104079](#). doi: <https://doi.org/10.2307/1427747>. [429](#)
- Dryden, I. and Mardia, K. (2016). *Statistical Shape Analysis, with Applications in R (2nd ed.)*. John Wiley and Sons, Chichester. [MR1646114](#). [427](#), [428](#)
- Geisser, S. and Eddy, W. F. (1979). “A predictive approach to model selection.” *Journal of the American Statistical Association*, 74: 153–160. [MR0529531](#). [441](#)

- Goodall, C. (1991). “Procrustes methods in the statistical analysis of shape.” *Journal of the Royal Statistical Society, Series B*, 53: 285–339. [MR1108330](#). 428
- Gower, J. (1975). “Generalized Procrustes analysis.” *Psychometrika*, 40: 33–50. [MR0405725](#). doi: <https://doi.org/10.1007/BF02291478>. 428
- Gutiérrez, L., Gutiérrez-Peña, E., and Mena, R. H. (2018). “Supplementary Material for ‘A Bayesian approach to statistical shape analysis via the projected normal distribution’.” *Bayesian Analysis*. doi: <https://doi.org/10.1214/18-BA1113SUPP>. 431
- Hardan, A. Y., Minshe, N. J., and Keshavan, M. S. (2000). “Corpus callosum size in autism.” *Neurology*, 55: 1033–1036. 438
- Hardan, A. Y., Pabalan, M., Gupta, N., Bansal, R., Melhem, N. M., Fedorov, S., Keshavan, M. S., and Minshe, N. J. (2009). “Corpus callosum volume in children with autism.” *Psychiatry Research: Neuroimaging*, 174: 57–61. 438
- He, Q., Duan, Y., Karsch, K., and Miles, J. (2010). “Detecting corpus callosum abnormalities in autism based on anatomical landmarks.” *Psychiatry Research: Neuroimaging*, 183: 126–132. 442
- Hernandez-Stumpfhauer, D., Breit, F., and van der Woerd, M. (2016). “The general projected normal distribution of arbitrary dimension: Modeling and Bayesian inference.” *Bayesian Analysis*. <http://projecteuclid.org/euclid.ba/1453211962>. [MR3597569](#). doi: <https://doi.org/10.1214/15-BA989>. 433
- Jammalamadaka, S. R. and SenGupta, A. (2001). *Topics in Circular Statistics*. World Scientific. [MR1836122](#). doi: <https://doi.org/10.1142/9789812779267>. 429
- Kendall, D. (1977). “The diffusion of shape.” *Advances in Applied Probability*, 9: 428–430. 427
- Kent, J. (1994). “The complex Bingham distribution and shape analysis.” *Journal of The Royal Statistical Society B*, 56: 285–299. [MR1281934](#). 428
- Klassen, E., Srivastava, A., Mio, W., and Joshi, S. H. (2004). “Analysis of planar shapes using geodesic paths on shape spaces.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26: 372–383. 427
- Kume, A. and Welling, M. (2010). “Maximum likelihood estimation for the Offset-Normal shape distributions using EM.” *Journal of Computational and Graphical Statistics*, 19(3): 702–723. [MR2732499](#). doi: <https://doi.org/10.1198/jcgs.2010.09190>. 428
- Kurtek, S., Srivastava, A., Klassen, E., and Ding, Z. (2012). “Statistical modeling of curves using shapes and related features.” *Journal of the American Statistical Association*, 107: 1152–1165. [MR3010902](#). doi: <https://doi.org/10.1080/01621459.2012.699770>. 427
- Leu, R. and Damien, P. (2014). “Bayesian shape analysis of the complex Bingham distribution.” *Journal of Statistical Planning and Inference*, 149: 183–200. [MR3199904](#). doi: <https://doi.org/10.1016/j.jspi.2013.11.016>. 428

- Mardia, K. and Dryden, I. L. (1999). “The complex Watson distribution and shape analysis.” *Journal of the Royal Statistical Society, Series B*, 61: 913–926. MR1722247. doi: <https://doi.org/10.1111/1467-9868.00210>. 428
- Mardia, K., Fallaize, C., Barber, S., Jackson, R., and Theobald, D. (2013). “Bayesian alignment of similarity shapes.” *Annals of Applied Statistics*, 7: 989–1009. MR3113498. doi: <https://doi.org/10.1214/12-AOAS615>. 428
- Mardia, K. V. and Jupp, P. E. (2000). *Directional Statistics*. Wiley. MR1828667. 429
- Mathai, A. M. and Provost, S. B. (1992). *Quadratic Forms in Random Variables: Theory and Applications (Statistics: Textbooks and Monographs)*. CRC Press. MR1192786. 430
- Micheas, A., Dey, D., and Mardia, K. (2006). “Complex elliptical distributions with application to shape analysis.” *Journal of Statistical Planning and Inference*, 136: 2961–2982. MR2256214. doi: <https://doi.org/10.1016/j.jspi.2005.01.002>. 428
- Oslen, A. and Westneat, M. (2015). “StereoMorph: an R package for the collection of 3D landmarks and curves using a stereo camera set-up.” *Methods in Ecology and Evolution*, 6: 351–356. 440
- Pau, G., Fuchs, F., Sklyar, O., Boutros, M., and Huber, W. (2010). “EBImage – an R package for image processing with applications to cellular phenotypes.” *Bioinformatics*, 26(7): 979–981. 440
- Piven, J., Bailey, J., Ranson, B. J., and Arndt, S. (1997). “An MRI study of the corpus callosum in autism.” *American Journal of Psychiatry*, 154: 1051–1056. 438
- Prigge, M., Lange, N., Bigler, E. D., Merkley, T. L., Neeley, E. S., Abildskov, T. J., Froehlich, A. L., Nielsen, J. A., Cooperrider, J. R., Cariello, A. N., Ravichandran, C., Alexander, A. L., and Lainhart, J. E. (2013). “Corpus callosum area in children and adults with autism.” *Research in Autism Spectrum Disorders*, 7: 221–234. 438
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>. 436
- Small, C. G. (1996). *The Statistical Theory of Shape*. Berlin: Springer. MR1418639. doi: <https://doi.org/10.1007/978-1-4612-4032-7>. 427
- Srivastava, A., Joshi, S., Kaziska, D., and Wilson, D. (2005). “Applications of planar shape analysis to image-based inferences.” In *Acoustics, Speech, and Signal Processing*, 1037–1040. Proceedings. (ICASSP ’05). IEEE International Conference. 427
- Vidal, C. N., Nicolson, R., DeVito, T. J., Hayashi, K. M., Geaga, J. A., Drost, D. J., Williamson, P. C., Rajakumar, N., Sui, Y., Dutton, R. A., Toga, A. W., and Thompson, P. M. (2006). “Mapping corpus callosum deficits in autism: an index of aberrant cortical connectivity.” *Biological Psychiatry*, 60(3): 218–225. 438
- Whitcher, B., Schmid, V., and Thornton, A. (2011). “Working with the DICOM and NIfTI Data Standards in R.” *Journal of Statistical Software*, 44(6): 1–28. URL <http://www.jstatsoft.org/v44/i06/> 440

Acknowledgments

The first, second and third authors were partially funded by grants Fondecyt-11140013, PAPIIT-UNAM IN106114-3 and CONACyT-241195, respectively. The work was initiated during a visiting research position at IIMAS-UNAM (Cátedra IIMAS) by the first author and concluded during a Fulbright Visiting Scholar position of the third author at the University of Texas at Austin. The authors wish to thank the Editor, Associate Editor, and Reviewer for helpful suggestions that improved the presentation of the paper, as well as Dr. Mauricio Canals and M.Sc. Stephanie Mesías for their dedication in the location of the landmarks in hundreds of images. Funding sources for the ABIDE data sets are listed at http://fcon_1000.projects.nitrc.org/indi/abide/index.html.