

## Biased estimators on Quotient spaces

Nina Miolane, Xavier Pennec

► **To cite this version:**

Nina Miolane, Xavier Pennec. Biased estimators on Quotient spaces. Geometric Science of Information. Second International Conference, GSI 2015., Oct 2015, Palaiseau, France. Springer, 9389, pp.130-139, Lecture notes in computer science (LNCS). <10.1007/978-3-319-25040-3\_15>. <hal-01203805>

**HAL Id: hal-01203805**

**<https://hal.inria.fr/hal-01203805>**

Submitted on 23 Sep 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Biased estimators on Quotient spaces

Nina Miolane <sup>1,\*</sup>, Xavier Pennec <sup>1</sup>

<sup>1</sup> INRIA, Asclepius project-team, 2004 Route des Lucioles, BP93, F-06902 Sophia Antipolis Cedex, France, [nina.miolane@inria.fr](mailto:nina.miolane@inria.fr), +33 4 92 38 71 82

**Abstract.** Usual statistics are defined, studied and implemented on Euclidean spaces. But what about statistics on other mathematical spaces, like manifolds with additional properties: Lie groups, Quotient spaces, Stratified spaces etc? How can we describe the *interaction between statistics and geometry*? The structure of Quotient space in particular is widely used to model data, for example every time one deals with shape data. These can be shapes of constellations in Astronomy, shapes of human organs in Computational Anatomy, shapes of skulls in Palaeontology, etc. Given this broad field of applications, statistics on shapes -and more generally on observations belonging to quotient spaces- have been studied since the 1980's. However, most theories model the variability in the shapes but do not take into account the noise on the observations themselves. In this paper, we show that statistics on quotient spaces are biased and even inconsistent when one takes into account the noise. In particular, some algorithms of template estimation in Computational Anatomy are biased and inconsistent. Our development thus gives a first theoretical geometric explanation of an experimentally observed phenomenon. A biased estimator is not necessarily a problem. In statistics, it is a general rule of thumb that a bias can be neglected for example when it represents less than 0.25 of the variance of the estimator. We can also think about neglecting the bias when it is low compared to the signal we estimate. In view of the applications, we thus characterize geometrically the situations when the bias can be neglected with respect to the situations when it must be corrected.

## Introduction

In Quantum Field theory, one can roughly consider a relativistic quantum field of quarks as a function from space-time to the space of colors:  $\phi : \mathbb{R}^4 \mapsto \mathbb{C}^3$ . The function  $\phi$  associates to each point of space-time an elementary particle: the quark living in the space of colors. The space of quarks fields carries a right action of the group of symmetry of space-time and a left action of the group of quark symmetry. These group actions represent what does not change the laws of Physics.

In Computational Anatomy when we study anatomical shapes, we encounter the same mathematical structures [4]. In others words, we have often two groups acting on the same anatomical object  $\phi$  respectively on the right and on the left. These group actions represent what does not change the shape of the object.

A first example is a set of  $K$  anatomical landmarks [7, 10], which is a function from the space of labels to the real 3D space:  $\phi : [1, \dots, K] \mapsto \mathbb{R}^3$ . In this case, the manifold  $\mathcal{M}$  of landmarks sets carries a right action of the group of permutations of the labels and a left action of the group of rotations and translations of the 3D space  $SE(3)$ . Relabeling the landmarks, or rotating and translating the whole set of landmarks, does not change the anatomical shape described. A second example is a 1D-signal [9] which is a function:  $\phi : [0, 1] \mapsto \mathbb{R}$ . The manifold  $\mathcal{M}$  of 1D-signals carries a right action of the group of diffeomorphisms of  $[0, 1]$ : this action represents the reparameterizations of the signal, which do not change its shape. A third example is a parameterized surface, or more generally immersed submanifold, in  $\mathbb{R}^d$  [3] which is a function:  $\phi : \mathcal{N} \mapsto \mathbb{R}^d$ . The manifold  $\mathcal{M}$  of parameterized surfaces (resp. of immersed submanifolds) carries a right action of the group of diffeomorphisms of  $\mathcal{N}$  and a left action of the group of rotations and translations of the  $\mathbb{R}^d$  space  $SE(d)$  (poses). Again, reparameterizing, rotating or translating the surface or the submanifold does not change its shape.

The anatomical data that we observe belong to the manifold  $\mathcal{M}$ . We are interested in the statistical analysis of their *shapes*. The corresponding shape data now belong to a quotient space  $\mathcal{Q}$ , which is the manifold  $\mathcal{M}$  quotiented by the group actions that leave the shape invariant. Thus we perform statistics on  $\mathcal{Q}$  rather than on  $\mathcal{M}$ , even if the data originally belong to  $\mathcal{M}$ . But  $\mathcal{Q}$  not a differentiable manifold in the general case. To be more precise,  $\mathcal{Q}$  is a *manifold with singularities* or a *stratified space*.

Statistics on quotient spaces or shape spaces  $\mathcal{Q}$  have been studied since the 1980's. Regarding shapes of landmarks, the theory has been first introduced by Kendall [8]. Regarding shapes of curves one can refer to [9, 6] and, in the more general setting of manifold quotiented by isometric group actions, to [5]. In all cases when one performs statistics on shape spaces, the estimation of the *mean shape* is central. A way of doing it is via parametric statistical estimation. By doing so, studies model the variability in the shapes of the objects (meaning that they assume some *variability in  $\mathcal{Q}$* ) and the variability in reparameterization or in objects pose (meaning that they assume some variability in the "orbit" of the group actions).

However, the effect of noise on the objects themselves (which is a *noise in the ambient manifold  $\mathcal{M}$* ) has not been thoroughly investigated yet. The noise on the objects in  $\mathcal{M}$  exists as it comes from imperfect measure instruments or human lack of precision, for example while positioning the landmarks on a medical image. Usual estimators were proven to be consistent in the theory without noise. But experiments have shown that they have bias and even inconsistency in the presence of noise [2]. For example, the standard estimator of the mean shape of the data, the Fréchet mean computed in the quotient space with the max-max algorithm, is biased and even inconsistent. This bias is experimentally observed and shows to be dependent on the noise level. But it is not theoretically understood so far. There is a need of a extended statistical theory on quotient spaces that takes the noise into account and quantifies the induced bias.

This paper is a first step in this direction. We suggest a geometric interpretation of the usual estimator of the mean shape. It enables to show that noise on the observations in  $\mathcal{M}$  induces bias on the estimator in  $\mathcal{Q}$ . We work in the case where the observations belong to a finite dimensional flat manifold  $\mathcal{M}$ , which is quotiented by a proper isometric action of a finite dimensional Lie group  $\mathcal{G}$ . We describe how the bias on the *statistical* estimator is controlled by the *geometry* of the quotient space  $\mathcal{Q}$ , in particular by its singularities at the scale of the noise. Even if we work in finite dimension, we provide the intuition of the behavior for infinite dimension.

In the first section, we present notions of differential geometry on quotient spaces. In the second section, we present the statistical framework and the geometric interpretation of the estimator usually computed. In the third section, we demonstrate the geometric origin of the bias on this estimator. In the fourth section, we show explicit computations of the bias for the case of  $\mathbb{R}^m$  quotiented by  $SO(m)$ , which is the common example of finite dimensional flat manifold quotiented by the isometric action of a finite dimensional Lie group.

## 1 Differential geometry of Quotient spaces $\mathcal{Q}$

We consider  $\mathcal{M}$  a finite dimensional flat Riemannian manifold and  $\mathcal{G}$  a finite dimensional Lie group. For more details on this section, we refer to [1, ?,?].

**Basis on Quotient Spaces** A *Lie group action* of  $\mathcal{G}$  on  $\mathcal{M}$  is a differentiable map  $\rho : \mathcal{G} \times \mathcal{M} \ni (g, X) \mapsto g \cdot X \in \mathcal{M}$ , such that  $e \cdot X = X$  and  $g' \cdot (g \cdot X) = (g'g) \cdot X$ . If  $\mathcal{G} \times \mathcal{M} \ni (g, X) \mapsto (g \cdot X, X) \in \mathcal{M} \times \mathcal{M}$ , is a proper mapping, i.e. if the inverse image of every compact is compact, the action is *proper*. If  $d\rho_g : T_X \mathcal{M} \rightarrow T_{g \cdot X} \mathcal{M}$  with  $\rho_g = \rho(g, \cdot)$  leaves the metric of  $\mathcal{M}$  invariant, the action is *isometric*. We only consider proper isometric actions.

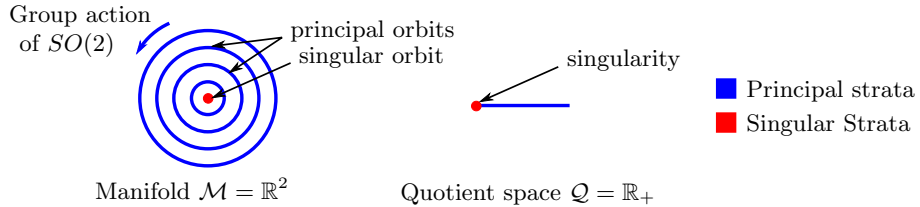
The *orbit* of  $X \in \mathcal{M}$  is defined as  $\mathcal{O}_X = \{g \cdot X \mid g \in \mathcal{G}\}$ . The orbit of  $X$  represents all the points that can be reached by moving  $X$  through the action of  $\mathcal{G}$ . The *isotropy group* of  $X$ , also called the *stabilizer* of  $X$ , is defined as  $\mathcal{G}_X = \{g \in \mathcal{G} \mid g \cdot X = X\}$ . The stabilizer of  $X$  represents the elements of  $\mathcal{G}$  that fail at moving  $X$ . The orbit  $\mathcal{O}_X$  is a submanifold of  $\mathcal{M}$  and the stabilizer  $\mathcal{G}_X$  is a Lie subgroup of  $\mathcal{G}$ . They are related by the orbit-stabilizer theorem as follows:  $\mathcal{O}_X \sim \mathcal{G}/\mathcal{G}_X$ . The orbits form a partition of  $\mathcal{M}$  and the set of orbits is called the *quotient space*  $\mathcal{Q} = \mathcal{M}/\mathcal{G}$ . Because the action is proper,  $\mathcal{Q}$  is Hausdorff and inherits a Riemannian structure from  $\mathcal{M}$ . This is precisely why we consider a proper action.

The isotropy groups  $\mathcal{G}_X$  and  $\mathcal{G}_{X'}$  of elements  $X$  and  $X'$  in the same orbit are conjugate groups in  $\mathcal{G}$ . This property enables to define the *orbit type* of an orbit as the conjugacy class ( $H$ ) of the isotropy groups of its elements. From the orbit-stabilizer theorem, we observe that a "large" orbit leads to a "small" isotropy group, so a "small" orbit type. The smallest orbit type is called the *principal orbit type*. The corresponding largest orbits are called *principal orbits*. In contrast, non principal orbits are called *singular orbits*. The corresponding

points in  $\mathcal{Q}$  are called *singularities*. The action of  $\mathcal{G}$  on  $\mathcal{M}$  is said to be *free* if there is no singular orbits. In this case,  $\mathcal{Q}$  has no singularities and is a differentiable manifold. Otherwise  $\mathcal{Q}$  is a *manifold with singularities*.

We investigate how the singular orbits are distributed in  $\mathcal{M}$ , or equivalently how the singularities are distributed in  $\mathcal{Q}$ . The partition of  $\mathcal{M}$  into orbits is a singular Riemannian foliation called the *canonical foliation of  $\mathcal{M}$* . We can gather the orbits of same orbit type ( $H$ ) and define  $\mathcal{M}_{(H)} = \{X \in \mathcal{M} \mid \mathcal{G}_X \text{ conjugate to } H\}$ . The decomposition of  $\mathcal{M}$  into the connected components of the  $\mathcal{M}_{(H)}$  is a stratification called the *orbit type stratification of  $\mathcal{M}$* . In this stratification, the principal type component is open and dense.

The decomposition of  $\mathcal{Q}$  into the corresponding components  $\mathcal{Q}_{(H)} = \mathcal{M}_{(H)}/\mathcal{G}$  also forms a stratification, whose principal type component is also dense. Thus, singularities are "sparsely distributed" in  $\mathcal{Q}$  in the sense that they are of null Lebesgue measure.



**Fig. 1.** Stratification and foliation of  $\mathcal{M}$  and stratification of  $\mathcal{Q}$  for  $\mathcal{M} = \mathbb{R}^2$  quotiented by the action of  $SO(2)$  so that  $\mathcal{Q} = \mathbb{R}^2/SO(2) = \mathbb{R}_+$ .

These foliation and the stratifications are illustrated in Figure 1 in the case of  $\mathcal{M} = \mathbb{R}^2$  quotiented by the action of  $SO(2)$  so that  $\mathcal{Q} = \mathbb{R}^2/SO(2) = \mathbb{R}_+$ . We use this 2D example throughout the paper as it is convenient for the illustrations. But of course, the theory applies to the general case.

**Orbits as immersed submanifolds** An orbit  $\mathcal{O}$  is an immersed submanifold in  $\mathcal{M}$ , i.e. a differentiable manifold together with an immersion  $I : \mathcal{O} \mapsto \mathcal{M}$ . We identify  $\mathcal{O}$  with  $I(\mathcal{O})$  and we denote  $T_X\mathcal{O}$ , resp.  $N_X\mathcal{O}$ , the tangent space, resp. the normal space with respect to the metric of  $\mathcal{M}$ , of  $\mathcal{O}$  at  $X$ . The *first fundamental form* of  $\mathcal{O}$  is defined to be the induced metric on  $\mathcal{O}$ , i.e. the pull-back of the ambient (flat) metric of  $\mathcal{M}$  by  $I$ . The *second fundamental form* of  $\mathcal{O}$  is the quadratic form  $h$  defined by  $T_X\mathcal{O} \times T_X\mathcal{O} \ni (u, v) \mapsto h(v, w) = (\nabla_u w)^\perp \in N_X\mathcal{O}$ , where  $\nabla$  is the Levi-Civita connection on  $\mathcal{M}$  and  $(\nabla_v w)^\perp$  the orthogonal projection on  $N_X\mathcal{O}$ .

We can write  $\mathcal{O}$  locally around  $X$  as the graph  $G$  of a smooth function from  $T_X\mathcal{O}$  to  $N_X\mathcal{O}$ . Let  $Z$  be a point on  $\mathcal{O}$  near  $X$ . On  $T_X\mathcal{O}$  we parameterize  $Z$ , and thus  $\mathcal{O}$ , by its Riemannian Logarithm  $z = \text{Log}_X(Z)$ . The coordinates in  $T_X\mathcal{O}$  are labeled with  $i, j$ , the coordinates in  $N_X\mathcal{O}$  are labeled with  $a, b$ . Locally

around  $X$ , and equivalently around  $z = 0$ , we have the graph equation of  $\mathcal{O}$ :

$$G^a(z) = -\frac{1}{2}h_{ij}^a(0)z^i z^j + O(|z|^3) \quad (1)$$

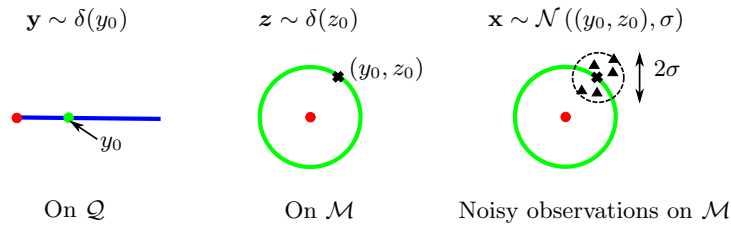
where we use Einstein summation convention: there is an implicit sum on indices that are repeated up and down. The second fundamental form  $h$  at 0 is thus the best approximation of  $\mathcal{O}$  around  $X$  as the graph of a quadratic function. It quantifies the external curvature of  $\mathcal{O}$  in  $\mathcal{M}$ .

Here we have considered a fixed orbit  $\mathcal{O}$ . But later in the paper we will consider different orbits. As the space of orbits is by definition the quotient space  $\mathcal{Q}$ , we will label the orbits by the corresponding point in  $\mathcal{Q}$ . So we will write  $\mathcal{O}_y$  for  $y \in \mathcal{Q}$ . We will thus have expressions such as  $h_{ij}^a(y, 0)$ , for the second fundamental form of the orbit represented by  $y$ .

## 2 Geometric interpretation of the template estimation

Here we interpret the statistical template estimation in Computational Anatomy with the quotient space framework. Such an estimation is usually parametric. It means that we assume a parametric model that has produced the observations, in our case: the anatomies. The (parametric) statistical estimation amounts to infer the parameters of the model from the observations, in our case: infer the template, or the mean anatomical shape, from the observed anatomies.

More precisely, one usually assumes that there is an underlying probability density of anatomical shapes, i.e. a probability density on  $\mathcal{Q}$ , whose mean is called the template. We consider here the simplest generative model with a dirac distribution in  $\mathcal{Q}$ , i.e. no variability in shapes  $\mathbf{y} \sim \delta(y_0)$ . Here  $y_0$  is the template. As we do not observe the anatomical shapes in  $\mathcal{Q}$ , but rather the anatomies in  $\mathcal{M}$ , one has to model the poses of the shape. One usually assumes that there is a probability density in the orbits. We assume here a uniform distribution on the orbits, i.e. maximal variability in poses:  $z \sim \mathcal{U}(\mathcal{O})$ . Finally, we model the noise on the observations by a Riemannian Gaussian in  $\mathcal{M}$  with isotropic noise  $\sigma$ :  $\mathbf{x} \sim \mathcal{N}_{\mathcal{M}}((\mathbf{y}, \mathbf{z}), \sigma)$  [12]. As we have an isometric action and a isotropic noise, this whole generative model is equivalent to the even simpler one:  $\mathbf{y} \sim \delta(y_0)$  and  $z \sim \delta(z_0)$  and  $\mathbf{x} \sim \mathcal{N}_{\mathcal{M}}((y_0, z_0), \sigma)$ .

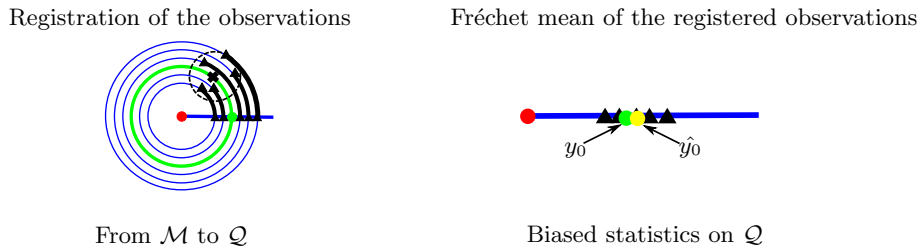


**Fig. 2.** Geometric interpretation of the three steps of the generative model.

Figure 2 illustrates the geometric picture for the generative model. The green dot in  $\mathcal{Q} = \mathbb{R}_+$  is the template, or the mean shape. The green circle in  $\mathcal{M} = \mathbb{R}^2$  is the orbit of  $y_0$ . We choose the pose  $z_0$  on this orbit to get the anatomical object represented by the black dot  $x_0 = (y_0, z_0)$  on the green circle. The observed anatomical objects are noisy observations of  $(y_0, z_0)$ . They are represented by the black triangles generated by the bivariate Gaussian of isotropic noise  $\sigma$ .

The goal of the statistical estimation is to compute an estimator  $\hat{y}_0$  of the template  $y_0$ . A standard method is the max-max algorithm with unimodal approximation, whose two steps can be geometrically interpreted as follows. First, the registration step amounts to the projection of the observations in  $\mathcal{Q}$ . Then, one computes the Fréchet mean  $\hat{y}_0$  (a generalization of the notion of mean in Riemannian manifolds) of the observations in  $\mathcal{Q}$  in order to estimate  $y_0$ .

Figure 3 illustrates the geometric picture for the estimation. First, the observations are registered: the triangles follows the black curve on their orbits (blue circles) in order to be "projected" on  $\mathcal{Q} = \mathbb{R}_+$ . Ultimately, the Fréchet mean is computed from the projected triangles on  $\mathcal{Q} = \mathbb{R}_+$ .



**Fig. 3.** Geometric interpretation of the two steps of the statistical estimation.

The algorithm produces an estimator  $\hat{y}_0$  of  $y_0$ . How good is this estimator? To answer this question, we can study its unbiasedness and its consistency. We generalize the usual definitions on linear spaces to Riemannian manifolds. Thus, the *bias of an estimator*  $\hat{y}_0$ , relative to the parameter  $y$  it is designed to estimate, is defined as:  $\text{Bias}_{y_0}[\hat{y}_0] = \text{Log}_{y_0}(\mathbb{E}_{y_0}[\hat{y}_0])$ , where  $\text{Log}$  is the Riemannian logarithm. An *unbiased estimator* has null bias. A *consistent estimator* converges in probability to the estimator is designed to estimate, as the number of observations increases. We show that  $\hat{y}_0$  is biased and inconsistent as an estimator of  $y_0$ .

### 3 Geometric foundations of bias and inconsistency

We consider observations in  $\mathcal{M}$  generated with the probability density  $F$  of our model. We show that they are equivalent to observations in  $\mathcal{Q}$  generated with a probability density  $f$  which we compute.

The probability density on  $\mathcal{M}$  writes as follows:  $F(x) = \frac{1}{C} \exp(-\frac{d_{\mathcal{M}}^2(x, x_0)}{2\sigma^2})$  where  $C$  is the normalizing constant. We write  $\mathcal{O}_y$  for the orbit corresponding

to  $y \in \mathcal{Q}$ . For an isometric action we have:  $d\mathcal{M}(x) = d\mathcal{O}_y(x)d\mathcal{Q}(y)$ . Thus the probability of having an observation  $x$  projecting within the interval of quotient coordinates  $[y, y + dy]$  is:  $\mathbb{P}(\mathbf{y} \in [y, y + dy]) = \int_y^{y+dy} \left( \int_{\mathcal{O}_y} F(x)d\mathcal{O}_y(x) \right) d\mathcal{Q}(y)$ . By taking  $dy \rightarrow 0$ , we have the induced probability density of the quotient space:

$$f(y) = \int_{\mathcal{O}_y} F(x)d\mathcal{O}_y(x) = \frac{1}{c} \int_{\mathcal{O}_y} \exp\left(-\frac{d\mathcal{M}(x, x_0)^2}{2\sigma^2}\right) d\mathcal{O}_y(x) \quad (2)$$

To compute an approximation of the density  $f$ , we make the assumption that the noise is low, i.e.  $\sigma \ll 1$ . The low noise assumption enables to use Taylor expansions in the orbit coordinates and then integrate in  $\mathcal{O}_y$  locally on the ball  $\mathcal{B}_y = \mathcal{B}(x_0, \sigma) \cap \mathcal{O}_y \subset \mathcal{O}_y$ . To this aim, we introduce momenta defined for the uniform measure  $\mu_y(z)$  of the ball  $\mathcal{B}_y$  as:

$$\mathfrak{M}_{\mathcal{O}_y, p}^{i_1 \dots i_p} = \int_{\mathcal{O}_y} z^{i_1} \dots z^{i_p} \mu(dz) = \int_{\mathcal{B}_y} z^{i_1} \dots z^{i_p} dz \quad (3)$$

We get in coordinates centered at  $y_0 \in \mathcal{Q}$ :

$$f(y) = \frac{1}{2\sigma^2} \exp\left(-\frac{y_a y^a}{2\sigma^2}\right) \left( \mathfrak{M}_{\mathcal{O}_y, 0} + S(y)_{ij} \mathfrak{M}_{\mathcal{O}_y, 2}^{ij} + O(|\mathfrak{M}_{\mathcal{O}_y, 3}|) \right) \quad (4)$$

where:  $S_{ij}(y) = -\delta_{ij} + y_a h_{ij}^a(y, 0) + \frac{1}{2} h_{ki}^a(y, 0) h_{kj}^a(y, 0)$ .

The probability density  $f$  on  $Q$  differs from a normal distribution because of the  $y$ -dependent term:  $\mathfrak{M}_{\mathcal{O}_y, 0} + S(y)_{ij} \mathfrak{M}_{\mathcal{O}_y, 2}^{ij} + O(|\mathfrak{M}_{\mathcal{O}_y, 3}|)$ . This term can be interpreted as a *Taylor expansion of the bias with respect to the local geometry of the orbits*. The first order of the bias is  $\mathfrak{M}_{\mathcal{O}_y, 0}$  and corresponds to the area of  $\mathcal{B}_y$ , i.e. the area of  $\mathcal{O}_y$  seen at the scale of the noise. The second order of the bias is  $S(y)_{ij} \mathfrak{M}_{\mathcal{O}_y, 2}^{ij}$  and we recognize a contraction of the second momentum with the matrix  $S(y)$  that depends on the external curvature of  $\mathcal{O}_y$ . We expect the higher order terms to be also such contractions between higher order momenta and higher order derivatives of the external curvature of the orbits.

The expectation of  $f$  on the ball  $B(y_0, \sigma) \subset Q$  computed at the tangent space of the mean shape  $T_{y_0} \mathcal{Q}$  gives:  $\text{Log}_{y_0}(\hat{y}_0)^a = \int_{B(y_0, \sigma)} y^a f(y) d\mathcal{Q}(y)$ . We recognize the bias of the estimator  $\hat{y}_0$  in the case of an infinite number of observations. It differs from 0 because the  $f$  distribution on  $Q$  is not symmetric. Because we are in the case of an infinite number of observations, this also shows the inconsistency of the estimator  $\hat{y}_0$ . Given the expression of  $f$ , we see that bias and inconsistency depend on the external curvature of the orbits and its first derivatives at the scale of the noise. As the external curvature of orbits generally increases when we approach singularities [11], the nearer we are from a singularity of  $\mathcal{Q}$ , the larger is the bias. All in all, when one performs usual statistics on  $\mathcal{Q}$  from observations on  $\mathcal{M}$ , the singularities in  $\mathcal{Q}$  induce bias and inconsistency.



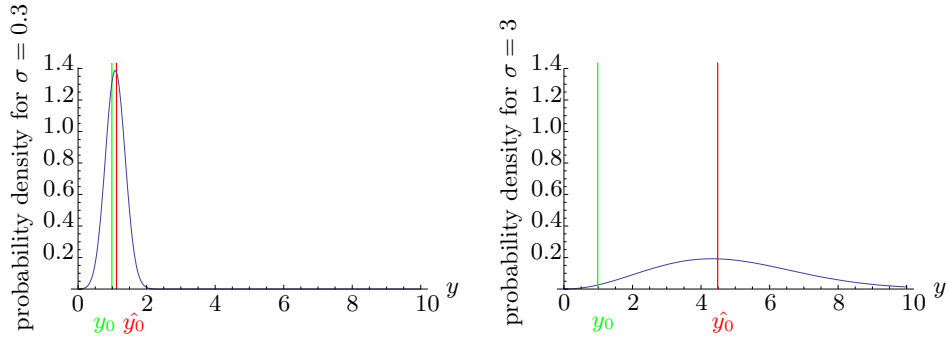
#### 4 An illustration on the quotient $\mathcal{Q} = \mathbb{R}^m / SO(m)$

We consider the case  $\mathbb{R}^m$  quotiented by  $SO(m)$ , which is a common example of a finite dimensional flat manifold quotiented by an isometric Lie group action. We perform the computations globally without the low noise assumption.

Generating observations in  $\mathcal{M} = \mathbb{R}^m$  with a multivariate normal law and then projecting to  $\mathcal{Q} = \mathbb{R}_+$  is equivalent to generating observations directly in  $\mathcal{Q}$  with the following probability density:

$$f_m(y) = \frac{2^{1-m/2}}{\sigma^m y^{m-1}} \exp\left(-\frac{y^2 + y_0^2}{2\sigma^2}\right) {}_0\tilde{F}_1\left(\frac{m}{2}, \frac{y^2 y_0^2}{4\sigma^2}\right) \quad (5)$$

where  ${}_0\tilde{F}_1$  is a regularized hypergeometric function. Figure 4 shows that an increase in the noise  $\sigma$  induces that the expectation of  $f$  is shifted away from  $y_0$ . The expectation is precisely the estimator  $\hat{y}_0$  for the case of an infinite number of observations. Thus we see the bias increasing with the noise level. We also note that the probability density and therefore the bias depend on the dimension  $m$ .



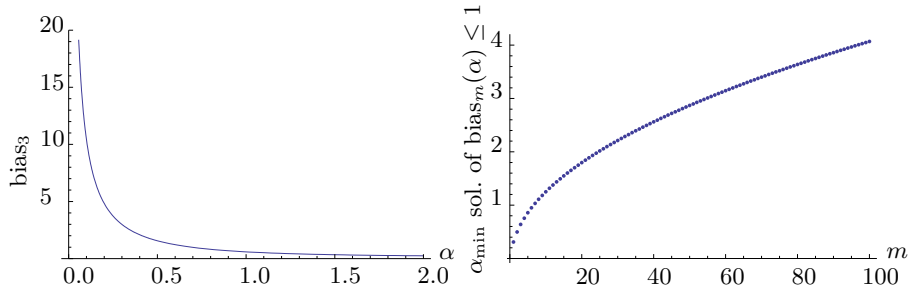
**Fig. 4.** Induced distributions on  $\mathcal{Q} = \mathbb{R}_+$  for  $m = 3$  and noise level  $\sigma = 0.3$  (left) and  $\sigma = 3$  (right). In green, the original mean. In red, the estimator of the mean.

The expectation of the probability density  $f_m$  on the quotient  $\mathbb{R}_+$  is  $\mathbb{E}_m(\alpha) = y_0 \cdot e_m(\alpha)$  where  $\alpha^2 = \frac{y_0^2}{2\sigma^2}$  and  $e_m(\alpha) = \frac{1}{\alpha} \Gamma\left(\frac{m+1}{2}\right) {}_1\tilde{F}_1\left(-\frac{1}{2}, \frac{m}{2}, -\alpha^2\right)$ . Because  $\mathbb{R}_+$  is linear, the bias writes:

$$\text{Bias}_{y_0}(\hat{y}_0) = \mathbb{E}[\hat{y}_0] - y_0 = y_0 \cdot (e_m(\alpha) - 1) = y_0 \cdot \text{bias}_m(\alpha) \quad (6)$$

The function  $\text{bias}_m(\alpha)$  is more precisely the bias of any estimator  $\hat{y}_0$  in units of the parameter  $y_0$  it is designed to estimate. It entirely depends in the variable  $\alpha$ . First,  $\alpha$  can be seen as a signal over noise ratio (SNR): we interpret  $y_0$  as the signal we seek to recover. But we can also interpret  $\alpha$  geometrically: it is the ratio of the distance  $y_0$  to the singularity 0 of  $\mathcal{Q} = \mathbb{R}_+$  over the noise  $\sigma$ . The most favorable conditions are when there is no noise with respect to the signal, or equivalently when we are far away from the singularity at the scale

of the noise:  $\alpha \rightarrow \infty$ . In this case, we have:  $\text{bias}_m(\alpha) \xrightarrow{\alpha \rightarrow \infty} 0$ . In contrast, the less favorable conditions are when the noise is preponderant with respect to the signal, or equivalently when we are close to the singularity at the scale of the noise:  $\alpha \rightarrow 0$ . In this case, we have:  $\text{bias}_m(\alpha) \xrightarrow{\alpha \rightarrow 0} \infty$  and more precisely  $\text{bias}_m(\alpha) = \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \frac{1}{\alpha} + O(\alpha)$ . The plot of  $\text{bias}_3$  is shown on the left side of Figure 5. Moreover, bias increases when we increase the dimension  $m$ . The right side of Figure 5 shows the minimal ratio  $\alpha$  one needs if  $\text{bias}_m(\alpha) \leq 1$  is required. This leads to think that there is bias in infinite dimension!



**Fig. 5.** Bias function of the SNR  $\alpha$  (left). Influence of the dimension on the bias (right).

## Conclusion

There is bias in statistics on quotient spaces  $\mathcal{Q} = \mathcal{M}/\mathcal{G}$  for noisy observations in  $\mathcal{M}$ . For a finite dimensional flat Riemannian manifold  $\mathcal{M}$ , local computations at the scale of noise show that the bias depends on geometric properties of  $\mathcal{Q}$ , more precisely on its singularities. Global computations on  $\mathbb{R}^m$  quotiented by  $SO(m)$  further emphasize that bias cannot be neglected as soon as signal and noise are of the same order, or equivalently as soon as we are close to a singularity at the scale of the noise. Additionally, the increase of the bias with the dimension leads to think that the same phenomenon exists in infinite dimension. Further developments will involve computations for non flat manifolds together with an algorithm to correct the bias. Ultimately, one should generalize the study to the infinite dimensional case.

## References

1. Alekseevsky, D., Kriegl, A., Losik, M., Michor, P.W.: The riemannian geometry of orbit spaces. the metric, geodesics, and integrable systems (2001)
2. Allasonnière, S., Amit, Y., Trouvé, A.: Towards a coherent statistical framework for dense deformable template estimation. Journal of the Royal Statistical Society. 69(1), pp. 3–29 (2007)

3. Bauer, M., Bruveris, M., Michor, P.: Overview of the geometries of shape spaces and diffeomorphism groups. *Journal of Mathematical Imaging and Vision* 50(1-2), 60–97 (2014)
4. Grenander, U., Miller, M.: Computational anatomy: An emerging discipline. *Quarterly of Applied Mathematics*; LVI(4):617-694. (1998)
5. Huckemann, S., Hotz, T., Munk, A.: Intrinsic shape analysis: Geodesic principal component analysis for riemannian manifolds modulo lie group actions. *Statistica Sinica* 20, 1–100 (2010)
6. Joshi, S., Kaziska, D., Srivastava, A., Mio, W.: Riemannian structures on shape spaces: A framework for statistical inferences. In: *Statistics and Analysis of Shapes*, pp. 313–333 (2006)
7. Kendall, D.G.: Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society* 16(2), 81–121 (1984)
8. Kendall, D.G.: A survey of the statistical theory of shape. *Statistical Science* 4(2), pp. 87–99 (1989)
9. Kurtek, S.A., Srivastava, A., Wu, W.: Signal estimation under random time-warpings and nonlinear signal alignment. In: *Advances in Neural Information Processing Systems* 24, pp. 675–683 (2011)
10. Le, H., Kendall, D.G.: The riemannian structure of euclidean shape spaces: A novel environment for statistics. *The Annals of Statistics* 21(3), 1225–1271 (09 1993)
11. Lytchak, A., Thorbergsson, G.: Curvature explosion in quotients and applications. *J. Differential Geom.* 85(1), 117–140 (05 2010)
12. Pennec, X.: Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision* 25(1), 127–154 (2006), <http://dx.doi.org/10.1007/s10851-006-6228-4>