



**HAL**  
open science

# Wrapped statistical models on manifolds: motivations, the case $SE(n)$ , and generalization to symmetric spaces

Emmanuel Chevallier, Nicolas Guigui

## ► To cite this version:

Emmanuel Chevallier, Nicolas Guigui. Wrapped statistical models on manifolds: motivations, the case  $SE(n)$ , and generalization to symmetric spaces. Joint Structures and Common Foundations of Statistical Physics, Information Geometry and Inference for Learning, Jul 2020, Les Houches, France. hal-03154401

**HAL Id: hal-03154401**

**<https://hal.science/hal-03154401v1>**

Submitted on 28 Feb 2021

**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.

# Wrapped statistical models on manifolds: motivations, the case $SE(n)$ , and generalization to symmetric spaces

Emmanuel Chevallier<sup>1</sup>, Nicolas Guigui<sup>2</sup>

<sup>1</sup> Aix Marseille Univ, CNRS, Centrale Marseille, Institut Fresnel, Marseille 13013, France

<sup>2</sup> Université Côte d'Azur, Inria Epione, Sophia Antipolis 06902, France

SPGIL, Les Houches 2020

## Abstract

We address here the construction of wrapped probability densities on Lie groups and quotient of Lie groups using the exponential map. The paper starts by briefly reviewing the different approaches to build densities on a manifold and shows the interest of wrapped distributions. We then construct wrapped densities on  $SE(n)$  and discuss their statistical estimation. We conclude by an opening to the case of symmetric spaces.

Non-Euclidean statistics, wrapped distributions, exponential map, moment matching estimator

## 1 Introduction

Consider  $X_1, \dots, X_k$  i.i.d. random variables  $\Omega \rightarrow \mathcal{M}$  on the manifold  $\mathcal{M}$  endowed with the volume measure  $v$ . Given a realisation of these variables we would like to estimate the underlying distribution.

An important problem in density estimation on manifolds is that most algorithms lead to heavy computations, partially because the standard probability densities do not enjoy the same properties as in the vector space case. Hence the motivation for building statistical models leading to algorithms with reduced amount of computation.

Wrapped models for directional statistics are constructed from distributions on  $\mathbb{R}$  or  $\mathbb{R}^n$  wrapped around the circle or the torus by the exponential map. Define informally wrapped models on manifolds as densities pushed from tangent spaces to the manifold. Due to their interesting properties in terms of computational complexity and invariances, we chose to focus on wrapped models based on the exponential map of Lie groups or on the exponential of Riemannian manifolds arising as quotients of Lie groups. Though in most wrapped

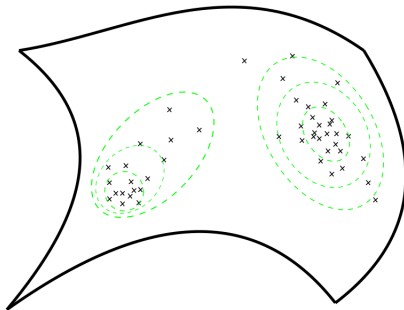


Figure 1: A set of i.i.d. draws  $x_1, \dots, x_k$  and the level lines of the estimated density.

distributions for directional statistics the densities are expressed as series, we will restrict our study to the case where distributions in tangent spaces are contained in injectivity domains of the exponential maps. Hence densities on the manifold are expressed using the Jacobian of the exponential map at only one point of the tangent space, and not by series.

In section 2, we present some classical probability densities on manifolds. In section 3, we review the relevant properties of probability densities and statistical models on manifolds, and describe wrapped statistical models. In section 4, we summarize results of [4] and describe the case  $SE(n)$ . The last section is an opening towards the general case of symmetric spaces.

## 2 Some classical probability densities on manifolds

Families of densities on manifolds are often defined in order to verify particular properties like maximizing an entropy constraint, having a particular form for the maximum likelihood estimator or verifying a particular pde. For instance Gaussian distributions defined in [8] on Riemannian manifolds,

$$f_{x_0}(x) = \alpha(\gamma, x_0) e^{-\frac{d(x, x_0)^2}{2\gamma^2}} \quad (1)$$

and

$$f_{x_0}(x) = \alpha(\Gamma, x_0) e^{-\Gamma(\log_{x_0}(x), \log_{x_0}(x))} \quad (2)$$

were  $\gamma \in \mathbb{R}$  and  $\Gamma$  is a positive definite bilinear form, are the maximal entropy distributions given their first and second moments, see section 3.2 for the definition of moments. As shown in [11], on non-compact symmetric spaces distributions (1) are also such that the maximum likelihood estimator of parameter  $x_0$  is the empirical mean.

Another usual way of defining Gaussian distributions on manifolds is by using a Laplacian operator: Gaussian distributions defined as heat kernels, i.e. Green functions of the heat equation  $\frac{\partial f}{\partial t} = -\Delta f$ .

Due to its practical importance several distributions have also been proposed on spheres, such as the Fisher distributions and their anisotropic counterpart, the Kent distributions:

$$\forall \mu, x \in \mathbb{S}_2, \kappa \in \mathbb{R}, \quad f_{\mu, \kappa}(x) = \frac{\kappa}{4\pi \sinh(\kappa)} e^{\kappa \mu^T x} \quad (\text{Fisher}) \quad (3)$$

$$\forall x \in \mathbb{S}_2, \kappa, \beta \in \mathbb{R}, \quad f_{\gamma_i, \kappa, \beta}(x) = \frac{1}{c(\kappa, \beta)} e^{\kappa \gamma_1^T x + \beta (\gamma_2^T x)^2 - (\gamma_3^T x)^2}, \quad (\text{Kent}) \quad (4)$$

where  $\gamma_1 \perp \gamma_2 \perp \gamma_3 \in \mathbb{S}_2$ , see [5] and [6].

None of the previous distributions are naturally interpreted as wrapped distributions. Wrapped distributions are not defined by a theoretical property but by the procedure to construct them: given a measurable map  $\phi$  from a tangent space  $T_{x_0}\mathcal{M}$  to the manifold, wrapped densities are densities which are first defined on a tangent space and pushed forward by  $\phi$ . In the present paper, we only address the case where  $\phi$  is a diffeomorphism on a neighborhood  $U$  of  $0 \in T_{x_0}\mathcal{M}$ . Given a Lebesgue measure on  $T_{x_0}\mathcal{M}$  and a probability density  $h$  on  $T_{x_0}\mathcal{M}$  whose support is included in  $U$ , the density of the push forward at  $x \in \phi(U)$  is given by

$$f(x) = (\phi_* h)(x) = \frac{h(x)}{J(\phi^{-1}(x))}$$

where  $J$  is the Jacobian determinant of  $\phi$ . Among choices of  $\phi$  an interesting candidate is the exponential map, due to its algebraic and geometric properties. Recall that on a Lie group, the exponential map at identity is extended to arbitrary points  $g$  as

$$\exp_g = L_g \circ \exp \circ dL_{g^{-1}},$$

where  $L_g$  is the left multiplication by  $g$  and  $dL_g$  its differential. Since  $L_g \circ R_{g^{-1}} \circ \exp = \exp \circ dL_g \circ dR_{g^{-1}}$ , the definition remains the same if left multiplication are replaced by right multiplications. Extending the Lie group exponential in this way enables to push densities from arbitrary tangent spaces to the group. On specific manifolds, other projection-retraction maps might present computational advantages over the exponential map while preserving the desired invariances, especially when the Riemannian exponential cannot be computed explicitly, see [1]. Nonetheless, we chose focus on the exponential map.

### 3 Some important characteristics of statistical models on manifolds

#### 3.1 Expression of the density functions

Many of the common probability densities on manifolds do not have explicit expressions. For Gaussian distributions defined in Eq.1 and Eq.2, or for the Kent distribution on spheres, the normalizing constant has an explicit expression only in exceptional cases. Even though this

factor can be numerically estimated, it is interesting to construct densities whose expressions are fully known.

Heat kernels have a similar problem: they can be computed only on a few exceptional manifolds. Furthermore while being a natural object on Riemannian manifolds, there is usually not a canonical Laplacian on Lie groups.

For wrapped distributions, the expressions require knowing  $\phi^{-1}$  and  $J$ . It turns out that when  $\phi$  is an exponential map restricted to an injective domain,  $\phi^{-1}$  and  $J$  can be computed on numerous Lie groups and quotients of Lie groups.

### 3.2 Moments

Let  $\mathcal{M}$  be a Lie group or a Riemannian manifold endowed with the Haar or Riemannian measure  $\nu$ . Following [9], we say that  $\bar{x}$  is a mean of the density  $f$  when

$$E_f(\log_{\bar{x}}(x)) = \int_{\mathcal{M}} \log_{\bar{x}}(x) f(x) d\nu(x) = 0.$$

Hence  $\bar{x}$  is a mean if the vectorial mean on the distribution lifted in  $T_{\bar{x}}\mathcal{M}$  is zero. This expression assumes the existence of  $E_f(\log_{\bar{x}}(x))$ , hence the existence of a mean is related to the injectivity and surjectivity of the exponential map.

Given the expression of a density  $f$ , computing the mean is not always easy. Hence, an important property of Gaussians of Eq.1-2 is that when  $\exp_{x_0}$  is a bijection between  $T_{x_0}\mathcal{M}$  and  $\mathcal{M}$  the parameter  $x_0$  is the unique mean of the distribution. Due to the symmetry of the sphere, it can also be checked that  $\mu$  and  $\gamma_1$  are means of the Fisher and Kent distributions.

When a mean  $\bar{x}$  exists, we define the corresponding higher order moments as in the vectorial case after lifting the distribution in  $T_{\bar{x}}\mathcal{M}$ ,

$$T_n = E_f(\log_{\bar{x}}(x)^{\otimes n}) \in (T_{\bar{x}}\mathcal{M})^{\otimes n}.$$

The covariance being the second order moment, we have

$$\Sigma = E_f(\log_{\bar{x}}(x) \otimes \log_{\bar{x}}(x)). \quad (5)$$

Note that when the mean is not unique, higher-order moments are defined with respect to a mean.

Unfortunately, the parameters  $\gamma$  or  $\Gamma$  of Gaussians of the form (1)-(2) do not have an explicit link with the second moments of the distributions, see [8, 11]. Similarly, to our knowledge the covariance of the Kent distribution is not explicitly related to  $\gamma_i$ ,  $\kappa$  and  $\beta$ .

An advantage of wrapped distribution is that they can have prescribed moments, which is rarely the case for other distributions. Let  $h$  be a probability density on  $T_{\bar{x}}\mathcal{M}$  such that, its support is included in an injectivity domain of  $\exp_{\bar{x}}$ , its barycenter is 0, its covariance is  $\Sigma$  and such that  $h(x) = h(-x)$ . Let

$$f(x) = (\exp_{\bar{x}*} h)(x) = \frac{h(x)}{J(\log(x))} \quad (6)$$

be the push forward. It can be shown (it is clear) that  $\bar{x}$  is a mean of the probability density  $f$ . It follows, by definition (5), that the covariance of  $f$  with respect to  $\bar{x}$  is  $\Sigma$ . Under some assumptions on  $\Sigma$ , it can also be shown that  $\bar{x}$  is unique.

### 3.3 Invariances and estimation

Given a statistical model  $\mathcal{S}$  of densities on  $\mathcal{M}$ , the construction of estimators from a set of i.i.d. samples is a central problem. Let  $T$  be an estimator, that is to say a function which maps sets of samples  $x_1, \dots, x_k \in \mathcal{M}$  to elements in  $\mathcal{S}$ . There are mostly three criteria to evaluate the interest of  $T$  as an estimator (the quality of an estimator  $T$ ). The first criterion is how fast  $T(x_1, \dots, x_k)$  converges to  $f$  as the number of samples increases, when samples are i.i.d. from  $f \in \mathcal{S}$ . The second criterion is the ease of evaluation of  $T$ . And the third criterion is the invariance by symmetry: when a group  $G$  acts on  $\mathcal{M}$ ,  $T$  should commute with the action. Implicitly, this assumes that the statistical model  $\mathcal{S}$  is invariant under the pushforward action of  $G$  on densities, and that

$$T(g.x_1, \dots, g.x_k) = g_* . T(x_1, \dots, x_k) \in \mathcal{S}.$$

For instance, when  $\mathcal{M}$  is a Riemannian manifold,  $G$  is the group of isometries and when  $\mathcal{M}$  is a Lie group,  $G = \mathcal{M}$  is the group itself acting by left or right multiplications.

As mentioned in section 2, Gaussians of Eq.1 on non-compact symmetric spaces parametrized by their mean and variance verify the interesting properties: the maximum likelihood estimation of the mean is the empirical mean and the variance is a simple function of the empirical covariance. The set of Gaussians is invariant under isometries and the densities are easily computed once the normalizing constant has been numerically estimated. Hence statistical models on non-compact symmetric spaces based on Gaussians of Eq.1 have several desirable properties, their main limitation being that they only contain isotropic densities. On the other hand Gaussians of Eq.2 can model anisotropy but no results have been shown on maximum likelihood estimation and the relations between the covariance parameter  $\Gamma$  and the normalizing factor is more involved than in the isotropic case. Beyond this limitation, Gaussians of types (1) and (2) lose many of their properties when the exponential maps are not bijections, for instance on sphere and more generally on non-compact symmetric spaces. On the other hand, the estimation of parameters of the Kent distributions on spheres is not straightforward. Hence our motivation for studying wrapped statistical models.

#### 3.3.1 Wrapped statistical models

In this section we introduce the wrapped statistical models defined by a kernel  $K$ . Let  $K : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be a function such that

$$K(x \geq 1) = 0, \quad \int_{\mathbb{R}^n} K(\|x\|) = 1, \quad \int_{\mathbb{R}^n} xK(\|x\|) = 0.$$

Let

$$\int_{\mathbb{R}^n} xx^t K(\|x\|) = \beta I,$$

where  $n$  is the dimension of  $\mathcal{M}$  and  $I$  is the identity matrix. Since  $K(x \geq 1) = 0$ , we have  $0 < \beta < 1$ . Given a symmetric positive definite bilinear form  $B$  on  $T_x\mathcal{M}$ , we define the density  $h_{x,B}$  by

$$h_{x,B}(u) = K(\sqrt{\beta B(u,u)}).$$

It can be checked that  $h_{x,B}$  is a probability density on  $T_x\mathcal{M}$  with respect to the Lebesgue measure normalized by  $\beta B$ . Since  $B$  is definite it induces an isomorphism between  $T_x\mathcal{M}$  and  $T_x\mathcal{M}^*$ .  $B$  itself can thus be mapped to a bilinear map  $\Sigma$  on  $T_x\mathcal{M}^*$ . If  $M$  is the matrix of  $B$  in a basis,  $M^{-1}$  is the matrix of  $\Sigma$  on the dual basis. Again, it can be checked that  $\Sigma$  is the covariance of  $h_{x,B}$ , and we will now write  $h_{x,\Sigma}$ . The push forward of  $h_{x,\Sigma}$  by the exponential is now a density

$$f_{x,\Sigma} = \exp_{x*}(h_{x,\Sigma})$$

on  $\mathcal{M}$ . If the support of  $h_{x,\Sigma}$  is contained in an injectivity domain we have Eq.6,

$$f_{x,\Sigma} = \frac{h_{x,\Sigma}(x)}{J(\log(x))}.$$

In many examples there is an injectivity domain in the tangent space of the form  $U = \mathcal{B} + E \subset T_x\mathcal{M}$  where  $\mathcal{B}$  is an open ball in a vector space  $F$  and  $E$  a vector space such that  $T_x\mathcal{M} = E \oplus F$ . It is for instance on case on  $SE(n)$ , see section 4. Imposing that the density is supported in  $U$  can then be done in the following way. Note  $I$  the bilinear form associated with the ball  $\mathcal{B}$  and define the admissible set of covariance  $A_x$  as

$$A_x = \{\Sigma \in T_x\mathcal{M} \otimes T_x\mathcal{M} \mid \forall u \in F, \beta B_\Sigma(u,u) \geq I(u,u)\},$$

where as previously  $B_\Sigma$  is the bilinear form on  $T_x\mathcal{M} \otimes T_x\mathcal{M}$  associated with  $\Sigma$ . The statistical model on  $\mathcal{M}$  constructed from the function  $K$  is then defined by

$$\mathcal{S} = \{f_{x,\Sigma} \mid x \in \mathcal{M}, \Sigma \in A_x\}. \quad (7)$$

As mentioned in 3.2,  $x$  is not always the unique mean of  $f_{x,\Sigma}$ , nonetheless it is true for sufficiently small  $\Sigma$ .

An important strength of this approach is that it only depends on the exponential and its Jacobian, which are known on many Lie groups and quotient of Lie groups. Thus, it can be applied to numerous manifolds, it is for instance a natural way to build models on compact symmetric spaces.

### 3.3.2 Estimators in wrapped statistical models

By construction, the exponential map commutes with the action of isometries on a Riemannian manifold, and with the action of group multiplications on a Lie group. Hence by construction the model  $\mathcal{S}$  has the desired symmetry properties. Maximum likelihood estimation in  $\mathcal{S}$  is unfortunately non trivial in most cases. On the other hand,  $\mathcal{S}$  is constructed such that the moment matching estimator has a simple form.

The first step of the moment matching estimator consists in computing the empirical mean  $\hat{x}$  of the sample and is achieved with a gradient descent. The second step consists in computing the empirical covariance  $\hat{\Sigma}$  in  $T_{\hat{x}}\mathcal{M}$  and choosing a covariance  $p(\hat{\Sigma})$  in  $A_{\hat{x}}$  as close as possible to  $\hat{\Sigma}$ . This choice can be achieved in the following way. As in section 3.3.1, consider an injectivity domain of the form  $U = \mathcal{B} + E \subset T_x\mathcal{M}$ . Let  $I$  be the bilinear form associated with  $\mathcal{B}$  on the vector space  $F$  generated by  $\mathcal{B}$  and let  $f$  be the self-adjoint endomorphism on  $F$  defined by

$$\forall u, v \in F, \beta\hat{B}(u, v) = I(u, f(v)).$$

Note  $\lambda$  the smallest eigenvalue of  $f$ . If  $\lambda \geq 1$  then  $\hat{\Sigma} \in A_{\hat{x}}$ . If  $\lambda < 1$ , we have

$$\beta\hat{B}(u, u) = \lambda I(u, u) < I(u, u),$$

for some  $u \in F$ , and  $\hat{\Sigma} \notin A_x$ . It is easy to see that the smallest multiple of  $\hat{B}$  compatible with  $A_{\hat{x}}$  is  $\frac{1}{\lambda}\hat{B}$ , hence a simple choice in  $A_{\hat{x}}$  is

$$p(\hat{B}) = \frac{1}{\lambda}\hat{B}, \quad p(\hat{\Sigma}) = \lambda\hat{\Sigma}.$$

Currently, there are only few results on moments estimation on manifolds. In [2], authors proved a central limit theorem for the empirical mean on Riemannian manifolds, and the author of [10] establishes results on the moments of the empirical mean on manifolds with an affine connection. To our knowledge, there are still no results on the estimation of the higher order moments, either on Riemannian manifolds, on Lie groups, or more generally on manifolds with an affine connection.

We have seen that the moment matching estimator is easy to compute, has the desired invariances, but convergence properties remain to be shown.

## 4 Probability densities on $SE(n)$

In this section we summarize results from [4] and describe the construction of wrapped models on the special Euclidean group  $SE(n)$ .



## 4.1 Wrapped models on $SE(n)$

As a manifold  $SE(n)$  is a product between  $SO(n)$  and  $\mathbb{R}^n$  but the group structure is a semi-direct product:

$$\begin{aligned} SE(n) &= SO(n) \ltimes \mathbb{R}^n \\ (R, t)(R', t') &= (RR', Rt' + t) \end{aligned}$$

and elements of its Lie algebra are parametrized by couples  $(A, T)$  where  $A$  is a skew-symmetric matrix and  $T \in \mathbb{R}^n$ . Recall that a skew-symmetric matrix can be block-diagonalized with 2 by 2 rotations on the diagonal, followed by a 0 when the dimension is odd. For each  $n$  by  $n$  skew-symmetric matrix  $A$ , we note  $\theta_1, \dots, \theta_{\lfloor \frac{n}{2} \rfloor}$  the set of angles of the 2 by 2 rotations. Let  $U$  be the subset in  $T_e SE(n)$  defined by

$$U = \{u = (A, T) \mid \theta_i \in [-\pi, \pi], i = 1, \dots, \frac{n(n-1)}{2}\}$$

It can be checked that the exponential map on  $U$  is a bijection. Hence we can define the logarithm on  $SE(n)$  as the inverse of the exponential on  $U$ . Isometries  $(R, t)$  of  $SE(n)$  can be embedded in  $GL(n+1)$  as

$$\begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \in GL_{n+1}(\mathbb{R}).$$

which enable to compute the exponential map of  $SE(n)$  using the matrix exponential of  $GL(n+1)$  and the logarithm using the matrix principal logarithm.

In order to be consistent with the previous section 3.3.1,  $U$  should in fact be restricted to a domain  $\tilde{U}$  of the form  $\mathcal{B} \times E$ ,

$$\tilde{U} = \{u = (A, T) \mid \sqrt{\text{trace}(AA^T)} < \pi\}.$$

The differential of the exponential map on Lie groups at  $u$  in the Lie algebra is given by the following formula:

$$d \exp_u = dL_{\exp u} \circ \left( \sum_{k \geq 0} \frac{(-1)^k}{(k+1)!} ad_u^k \right).$$

It can be checked that on  $SE(n)$ , volume forms induced by left invariant fields of basis are also right invariant: the group has a bi-invariant Haar measures. Fix an arbitrary reference basis  $e_1, \dots, e_p$  of the Lie algebra  $T_e SE(n)$  and consider the corresponding left invariant field of basis. The computation of the Jacobian determinant  $J$  in a left (or right) invariant basis field gives,

$$\begin{aligned} J = \det(d \exp_u) &= \left( \prod_i 2 \frac{1 - \cos(\theta_i)}{\theta_i^2} \right)^\alpha \times \dots \\ &\dots \prod_{i < j} \left( 4 \frac{1 + \cos(\theta_i + \theta_j)}{(\theta_i + \theta_j)^2} \cdot \frac{1 + \cos(\theta_i - \theta_j)}{(\theta_i - \theta_j)^2} \right) \end{aligned}$$

where  $\alpha = 1$  when  $n$  is even and  $\alpha = 2$  when  $n$  is odd and  $\theta_i$  are the angles of the planar rotations of the block diagonalization of  $A$ . This simplifies to  $J(\theta, T) = \left| 2 \frac{1 - \cos(\theta)}{\theta^2} \right|$  on  $SE(2)$ .

Given a kernel  $K$  on  $\mathbb{R}_+$  we can define the wrapped probability distribution

$$f_{g,\Sigma}(\exp_g(u)) = \frac{1}{J(u)\sqrt{\det(\Sigma)}} K\left(\sqrt{\beta u^t \Sigma^{-1} u}\right),$$

where  $u$  and  $\Sigma$  are respectively a coordinate vector and a positive definite matrix expressed in the reference left invariant basis field.

## 4.2 Density estimation on $SE(n)$

As noted in section 3.3.2, there are no theoretical results on the convergence the moment matching estimator. Hence we will only mention the empirical results obtained in [4] for the case  $SE(2)$ . Samples are drawn from a uniform distribution  $f_{e,\Sigma}$  on a ball in the injectivity domain in the Lie algebra. The empirical moments are then computed using the Python package `geomstats`, see [7], and the error between the original density and the estimated densities are computed using the  $L^1$  distance between densities  $\|f_{e,\Sigma} - f_{\hat{x},\hat{\Sigma}}\|_1$ . The  $L^1$  distance between densities is the total variation between measures and present the interest of being independent of the reference measure. The  $L^1$  error is compared to the same estimation error in  $\mathbb{R}^3$  where the density is estimated directly in  $T_e SE(2) \sim \mathbb{R}^3$  using a vectorial moment matching estimator. Results displayed in Fig.2 indicate that the errors of the  $SE(2)$  and  $\mathbb{R}^3$  problems are asymptotically related by a factor close to 1.

## 5 Towards a generalization to symmetric spaces

As seen in previous sections, the main challenges in building wrapped models are to compute the Jacobian determinants and to find injectivity domains for the exponential maps. We show here that symmetric spaces are an interesting class of manifolds regarding the computation of the Jacobian determinant. Consider a manifold  $\mathcal{M}$  with a connection  $\nabla$ . The manifold  $\mathcal{M}$  is said to be locally symmetric when

$$\nabla R = 0$$

where  $R$  is the curvature tensor of  $\nabla$ . Recall that Jacobi fields  $Y(t)$  along a geodesic  $\gamma$  are solution of the Jacobi equation,

$$\nabla_\gamma(\nabla_\gamma Y)(t) + R(\gamma'(t), Y(t))\gamma'(t) = 0$$

or

$$\nabla_\gamma(\nabla_\gamma Y)(t) + \tilde{R}(t)(Y(t)) = 0$$

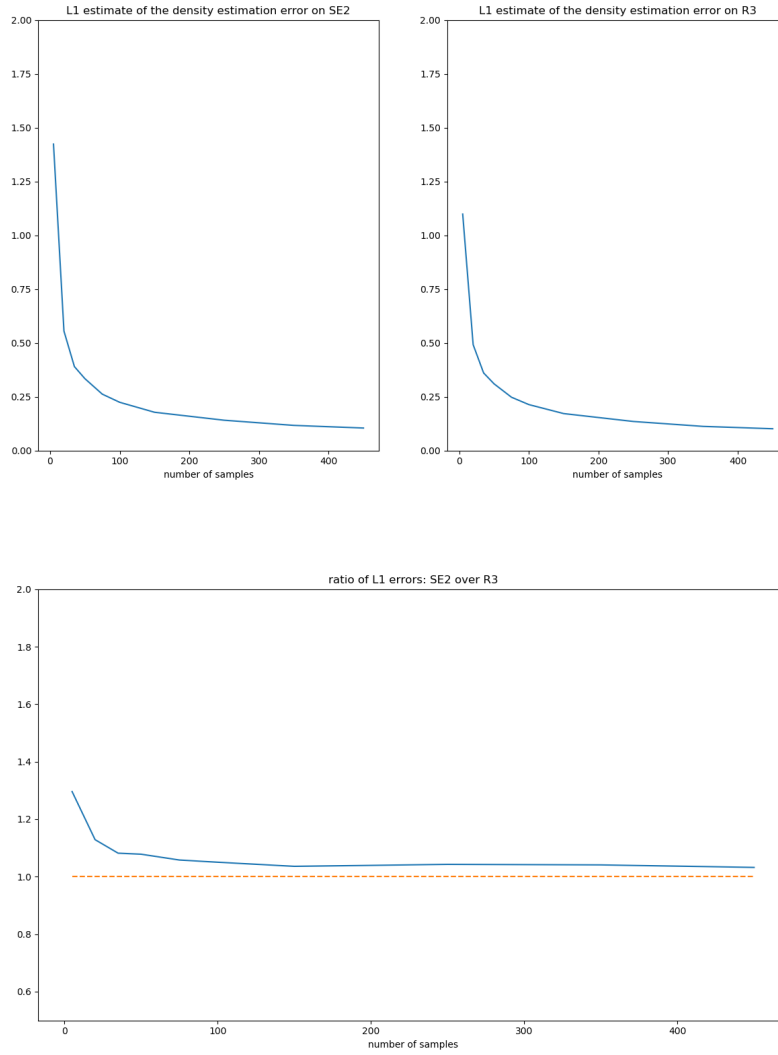


Figure 2:  $L^1$  errors and their ratios on  $SE(2)$  and  $T_eSE(2) \sim \mathbb{R}^3$ , see ([4]).

where  $\tilde{R}(t)$  is the linear map  $Y(t) \mapsto R(\gamma'(t), Y(t))\gamma'(t)$ . Let  $u$  be a tangent vector in  $T_x\mathcal{M}$  and  $Y_u$  be the Jacobi Field along  $\gamma$  with initial condition  $Y(0) = 0 \in T_x\mathcal{M}$  and  $(\nabla_{\gamma'(0)}Y)(0) = u$ . For  $t > 0$ , the differential of the exponential at  $t\gamma'(0)$  in the direction  $u$  is given by,

$$\frac{\partial \exp_x}{\partial u}(t\gamma'(0)) = \frac{1}{t}Y(t).$$

Hence the Jacobian determinant along  $t\gamma'(0)$  is given by

$$J(t) = \det(A(t))$$

where  $A$  follows the second order differential equation

$$\nabla_\gamma(\nabla_\gamma A)(t) + \tilde{R}(t)A(t) = 0.$$

When the space is locally symmetric, the coefficient  $\tilde{R}(t)$  is constant. Hence the symmetric property is a natural framework to study the Jacobian of the exponential map.

## 6 Conclusion

Across the paper, we have seen how wrapped models can be constructed on manifolds with an exponential map. We have analysed the case of the Lie group  $SE(n)$  and mentioned Riemannian manifolds, though such wrapped models can be considered on any manifold with an affine connection. The main limitation of these models is that the underlying manifold should have an exponential map, a log map and a Jacobian which can be computed at a reasonable cost. Note that this is often the case when the manifold is a homogeneous space. Future efforts should focus on three problems. The first one is to obtain convergence results for the density estimation using wrapped models. The second one is to understand and characterize the shape of the injectivity domains in the different possible settings, such as Lie groups and Riemannian symmetric spaces, see [12]. And the third problem is to identify on which manifolds with an affine connection the Jacobian of the exponential map can be computed explicitly. As shown in section 5, symmetric spaces seem to be an interesting setting to study this property. Note however that there are known examples outside of this class where the Jacobian is explicit, see for instance the Wasserstein metric on Gaussian distributions [3].

## References

- [1] Absil, P.-A., Malik, J., Projection-like retractions on matrix manifolds. *SIAM Journal on Optimization*, 2012, vol. 22, no 1, p. 135-158.
- [2] Bhattacharya, R., Patrangenaru, V., Large sample theory of intrinsic and extrinsic sample means on manifolds: II. *Annals of statistics*, 2005, p.1225-1259.

- [3] Chevallier, E., Kalunga, E., Angulo, J., Kernel density estimation on spaces of Gaussian distributions and symmetric positive definite matrices. *SIAM J. Imaging Sci.*, 2017, 10, 191–215.
- [4] Chevallier, E., Guigui, N., “A Bi-Invariant Statistical Model Parametrized by Mean and Covariance on Rigid Motions.” *Entropy*, 2020, 22(4), 432, <https://doi.org/10.3390/e22040432>.
- [5] Fisher, R. A., Dispersion on a sphere. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 1953, 217(1130), 295-305.
- [6] Kent, J. T., The Fisher-Bingham distribution on the sphere. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1982, 44(1), 71-80.
- [7] Miolane, N., Guigui, N., Le Brigant, A., Mathe, J., Hou, B., et al.. *Geomstats: A Python Package for Riemannian Geometry in Machine Learning*. 2020. hal-02536154
- [8] Pennec, X. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *J. Math. Imaging Vis.*, 2006, 25, 127.
- [9] Pennec, X., Arsigny, V., Exponential barycenters of the canonical Cartan connection and invariant means on Lie groups. In *Matrix Information Geometry*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 123–166.
- [10] Pennec, X. Curvature effects on the empirical mean in Riemannian and affine Manifolds: a non-asymptotic high concentration expansion in the small-sample regime. 2019. arXiv preprint arXiv:1906.07418.
- [11] Said, S., Hajri, H., Bombrun, L., and Vemuri, B. C., Gaussian distributions on Riemannian symmetric spaces: statistical learning with structured covariance matrices. *IEEE Transactions on Information Theory*, 2017, 64(2), 752-772.
- [12] Yang, L. (2007). Injectivity radius for non-simply connected symmetric spaces via Cartan polyhedron. arXiv preprint math/0703521.