

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/263652382>

# Multivariate Tensor-Based Morphometry with a Right-Invariant Riemannian Distance on $GL^+(n)$

**Article** in *Journal of Mathematical Imaging and Vision* · December 2013

DOI: 10.1007/s10851-013-0479-7

---

CITATIONS

9

---

READS

146

**3 authors**, including:



**Matias Nicolas Bossa**

University of Zaragoza

**27** PUBLICATIONS **351** CITATIONS

[SEE PROFILE](#)



**S. Olmos**

University of Zaragoza

**94** PUBLICATIONS **2,409** CITATIONS

[SEE PROFILE](#)

# Multivariate Tensor-Based Morphometry with a Right-Invariant Riemannian Distance on $GL^+(n)$

Ernesto Zacur · Matias Bossa · Salvador Olmos

© Springer Science+Business Media New York 2013

**Abstract** Tensor-based morphometry (TBM) studies encode the anatomical information in spatial deformations which are locally characterized by Jacobian matrices. Current methods perform voxel-wise statistical analysis on some features, such as the Jacobian determinant or the Cauchy–Green deformation tensor, which are not complete descriptors of the local deformation. This article introduces a right-invariant Riemannian distance on the  $GL^+(n)$  group of Jacobian matrices making use of the complete geometrical information of the local deformation. A numerical recipe for the computation of the proposed distance is given. Additionally, experiments are performed on both a synthetic deformation study and a cross-sectional brain MRI study.

**Keywords** Tensor-based morphometry · Jacobian matrices · Statistics on manifolds · Right-invariant Riemannian metric

## 1 Introduction

Tensor-based morphometry (TBM) is a methodology to analyze the anatomical information encoded by the spatial transformations that map a reference template (also called an atlas) to a set of images [27]. Two main ingredients can be considered within TBM methodology: non-rigid registration and statistical analysis. The spatial mappings are estimated by means of non-rigid registration methods [43]. Non-rigid registration, also called spatial normalization, of a pair of images or spatial objects is defined as finding the spatial transformation that achieves a correspondence between the images. After the registration process it can be assumed that the anatomical information is encoded by the spatial transformation. Subsequently, voxel-wise statistical analysis is performed on features extracted from the spatial derivatives of the mappings at each location (the Jacobian matrix, hereinafter denoted by  $\mathbf{J}(\mathbf{x})$ ). Several features such as the ones described below, have been proposed in the literature.

The simplest and still most widely used feature in TBM is the Jacobian determinant. This feature has two main advantages. Firstly, it has an intuitive interpretation because it represents the local volume change. Secondly, standard univariate statistical analysis can be easily performed on the Jacobian determinant (or also on its logarithm) [18, 22, 33, 59].

The main limitation of the Jacobian determinant is that it provides a coarse description of the deformation, because it only quantifies the local volume change. In fact, there are many possible deformations without local volume change which may be relevant for a specific application. To overcome this limitation, multivariate descriptors of the Jacobian

---

E. Zacur, M. Bossa, S. Olmos, for the Alzheimer's Disease Neuroimaging Initiative.

This work was funded by research grants TEC2009-14587-C03-01 from CICYT, AMIT project CEN-20101014 from CENIT program, CIM project IPT-2011-1683-900000 from INNPACTO program, Spain.

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.ucla.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)

---

E. Zacur · M. Bossa · S. Olmos (✉)

GTC, I3A, Universidad de Zaragoza, Maria de Luna 1,  
50018 Zaragoza, Spain  
e-mail: [olmos@unizar.es](mailto:olmos@unizar.es)

E. Zacur  
e-mail: [zacur@unizar.es](mailto:zacur@unizar.es)

M. Bossa  
e-mail: [bossa@unizar.es](mailto:bossa@unizar.es)

matrix have been proposed in order to get a more complete characterization. An example of a multivariate feature is the Cauchy–Green deformation tensor,  $\mathbf{C}(\mathbf{x}) = \mathbf{J}(\mathbf{x})^T \mathbf{J}(\mathbf{x})$  (or its square root) used in [36, 37, 55].

Most of the available statistical tools are well defined for Euclidean data but not for manifold-valued data, such as Cauchy–Green deformation tensors or Jacobian matrices. In the pioneering work [51], basic statistics, such as mean and variance, were computed by means of projections on a tangent space. Principal geodesic analysis was performed on some Lie groups in [15, 24] and also on the diffeomorphism group [63]. The analysis of diffusion tensor images (DTI) has fostered the development of statistical tools on symmetric positive definite (SPD) tensors [7, 12, 19, 26, 42, 53]. More recently regression on manifold-valued data was analyzed in [25, 31, 39], a multivariate extension of the Hotelling’s  $T^2$  test was proposed for longitudinal data in [47] and an in-depth study of principal geodesic analysis using both intrinsic statistics and the tangent space approximation was done in [58]. In all these works the definition of an appropriate distance on the manifold plays a key role.

In order to identify the anatomical regions with statistically significant differences between two groups of images, voxel-wise statistical inference is performed. The inference is performed at each template coordinate using the local description of the deformations. The result is illustrated on a statistical map that reflects the spatial distribution of the voxel-wise inference. It is desired and expected to mainly obtain the same result for any template choice allowing to make general statements about the anatomical location of the findings. The template invariance requirement can be achieved by using a symmetric and transitive procedure in registration stage [20, 57, 62]. These properties ensure that an anatomical label can be assigned to each location of the images and the label assignment is the same for any template choice. To our knowledge very few registration procedures guarantee these requirements, see for example [57] where a reference image must be selected. This paper is focused on the statistical techniques for TBM studies and it is assumed that the deformation maps are obtained with a symmetric and transitive registration algorithm. Statistical tools for TBM studies use Jacobian matrices as input data which belong to the group  $\text{GL}^+(n)$  of  $n \times n$  matrices with positive determinant. It will be shown that the template invariance requirement can be achieved by using a right-invariant distance for the statistical analysis, *i.e.* the distance between any pair of Jacobian matrices is preserved under translations from the right of both matrices.

The aim of this paper is twofold. Firstly, to introduce an appropriate right-invariant distance between Jacobian matrices, in particular using a right-invariant Riemannian distance on  $\text{GL}^+(n)$ . The second aim is to illustrate some results on a synthetic study and a cross-sectional brain MRI

study using three different right-invariant distances:  $d_{DET}$  which is a distance based on Jacobian determinants;  $d_{AFF}$  which is based on Cauchy–Green deformation tensors; and  $d_{RI}$  which is a Riemannian metric on  $\text{GL}^+(n)$ . In all experiments, voxel-wise hypothesis testing is performed by means of the Cramér test [11, 60] because its statistic can be computed using only the distances between observations for both univariate and multivariate data.

The paper is organized as follows. A background on univariate and multivariate TBM is given in Sect. 2, where the template invariance property is reviewed as well as two right-invariant distances previously used in the literature. Section 3.1 presents a short summary on Riemannian metrics on manifolds and Lie groups. Using these concepts, Sect. 3.2 explains how to compute a right-invariant Riemannian distance between elements on  $\text{GL}^+(n)$ . In Sect. 4 a hypothesis test to assess inference about group comparison in TBM is presented. Results are shown on both synthetic data as well as on brain MRI from Alzheimer’s Disease Neuroimaging Initiative (ADNI) in Sect. 5. Finally, a discussion is presented in Sect. 6.

## 2 Univariate/Multivariate TBM

### 2.1 Background

Let  $\Phi = (\phi^1, \phi^2, \dots, \phi^n)^T : \Omega \rightarrow \Upsilon$  be an invertible, orientation preserving and differentiable spatial mapping (the superscript  $\cdot^j$  refers to the  $j$ -th component), where  $\Omega$  and  $\Upsilon$  are simply connected subsets of  $\mathbb{R}^n$ . Up to first order the mapping is  $\Phi(\mathbf{x} + d\mathbf{x}) = \Phi(\mathbf{x}) + \mathbf{J}(\mathbf{x})d\mathbf{x} + O(d\mathbf{x}^2)$ , where  $\mathbf{x} = (x^1, x^2, \dots, x^n)^T \in \Omega$  and  $\mathbf{J}(\mathbf{x})$ , the *Jacobian matrix of  $\Phi$  at the point  $\mathbf{x}$* , is a linear transformation. As  $\Phi$  is invertible and orientation preserving then  $\mathbf{J}(\mathbf{x})$  belongs to  $\text{GL}^+(n)$  which is the set of  $n \times n$  matrices with positive determinant. The set  $\text{GL}^+(n)$  together with the matrix multiplication operation is a Lie group [61].

The  $n^2$  elements of the matrix  $\mathbf{J}(\mathbf{x})$  are obtained from the spatial derivatives of the mapping  $\Phi$

$$\mathbf{J}(\mathbf{x}) = (\mathcal{D}\Phi)|_{\mathbf{x}} = \begin{pmatrix} \partial_1 \phi^1|_{\mathbf{x}} & \partial_2 \phi^1|_{\mathbf{x}} & \cdots & \partial_n \phi^1|_{\mathbf{x}} \\ \partial_1 \phi^2|_{\mathbf{x}} & \partial_2 \phi^2|_{\mathbf{x}} & \cdots & \partial_n \phi^2|_{\mathbf{x}} \\ \vdots & & & \vdots \\ \partial_1 \phi^n|_{\mathbf{x}} & \partial_2 \phi^n|_{\mathbf{x}} & \cdots & \partial_n \phi^n|_{\mathbf{x}} \end{pmatrix},$$

where  $(\mathcal{D}\cdot)$  is the Jacobian operator and  $\partial_i \phi^j$  is the derivative of the  $j$ -th component of  $\Phi$  along the  $i$ -th coordinate.

Let  $\mathbf{A}$  and  $\mathbf{B}$  be two instances of any kind of spatial information, such as landmark coordinates, contours, intensity images, probability densities, representing similar contents. Registration or spatial normalization of  $\mathbf{A}$  and  $\mathbf{B}$  is formulated as finding the spatial transformation  $\Phi : \text{dom}(\mathbf{A}) \rightarrow$

$\text{dom}(\mathbf{B})$  such that  $\Phi \star A \sim B$ , where  $(\Phi \star A)$  is the action of  $\Phi$  on  $A$  and  $(\cdot \sim \cdot)$  refers to an equivalence relation typically defined as a matching energy. For example: in the case of intensity images  $(\Phi \star A)(\mathbf{x}) = A(\Phi^{-1}(\mathbf{x}))$  and the integral of the squared differences is typically used as matching energy for images from the same modality [43]; for images from different modalities, mutual information is typically employed as matching energy [54]; for landmark coordinates  $[\Phi \star A]_k = \Phi([A]_k)$  where  $[A]_k$  is the  $k$ -th landmark coordinate and a commonly used matching energy is  $\sum_k \|[\Phi \star A]_k - [B]_k\|^2$ ; for the case of spatially distributed probability densities,  $(\Phi \star A)(\mathbf{x}) = \det((\mathcal{D}\Phi^{-1})|_{\mathbf{x}})A(\Phi^{-1}(\mathbf{x}))$  [66]. From now on the term “image” will be used in a broad sense and will refer to any of the previous spatial information sources.

It is practical to define a function  $\mathcal{R}(\cdot, \cdot)$  that for two given images  $A$  and  $B$  returns a spatial transformation such that

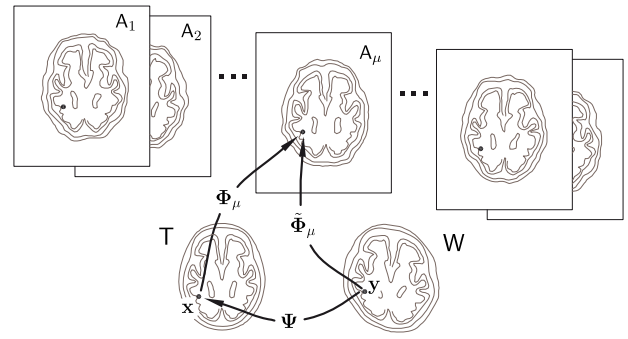
$$\mathcal{R}(A, B) \star A \sim B.$$

In order to be able to compute Jacobian matrices, the output of the function  $\mathcal{R}(\cdot, \cdot)$  must be invertible and differentiable in its whole domain. Examples of registration algorithms with these properties are in [9, 10, 13, 30, 34, 35, 56, 64]. Furthermore, in this work it is assumed that the registration process is symmetric and transitive to fulfill the invariance with respect to the template. These requirements mean  $\mathcal{R}(A, B) = \mathcal{R}(B, A)^{-1}$  and  $\mathcal{R}(B, C) \circ \mathcal{R}(A, B) = \mathcal{R}(A, C)$  for any image instance. In real practice these assumptions are rarely achieved by most registration methods. More details will be given in the discussion section.

Once the deformation  $\Phi = \mathcal{R}(A, B)$  is computed, both, the target  $B$  and the deformed instance  $\Phi \star A$  look spatially similar and a correspondence is obtained between all points of their domains  $\text{dom}(A)$  and  $\text{dom}(B)$ .

In tensor-based morphometry (TBM) the information of the local morphology is encoded by the Jacobian of the mapping registering the template  $T$  to the images. In particular, in order to assess statistical differences between two groups  $\mathcal{A} = \{A_a\}$  and  $\mathcal{B} = \{A_b\}$  at each location of the template domain a hypothesis test between the corresponding Jacobian matrices is performed. Just to fix notation, let  $\Phi_\mu = \mathcal{R}(T, A_\mu)$  be the mapping registering the template  $T$  to the image  $A_\mu$  such that  $\Phi_\mu \star T \sim A_\mu$ , and let  $\mathbf{J}_\mu(\mathbf{x})$  be the corresponding Jacobian matrix at the point  $\mathbf{x} \in \text{dom}(T)$ , i.e.  $\mathbf{J}_\mu(\mathbf{x}) = (\mathcal{D}\mathcal{R}(T, A_\mu))|_{\mathbf{x}}$ . For each location  $\mathbf{x}$  a group comparison test will be performed over the sets  $\{\mathbf{J}_a(\mathbf{x})\}$  and  $\{\mathbf{J}_b(\mathbf{x})\}$ . From here onwards, the subindexes  $a$  and  $a'$  run over the set  $\mathcal{A}$ , the subindexes  $b$  and  $b'$  run over the set  $\mathcal{B}$ , and, in general, indexes  $\mu$  and  $\nu$  run over both sets.

In many statistical tools a concept of ‘distance’ between instances is required. Therefore in order to perform statistics on Jacobian matrices a distance function  $d(\cdot, \cdot): \text{GL}^+(n) \times \text{GL}^+(n) \rightarrow \mathbb{R}^+$  must be defined.



**Fig. 1** Illustration of template invariance in a TBM study. Using the transitivity property assumption of the mappings, each image location and its corresponding location for all templates can be given with an anatomical label

## 2.2 Invariance with Respect to the Template

Let  $T$  and  $W$  be two possible templates and let  $\Psi = \mathcal{R}(W, T)$  be the differentiable and invertible mapping relating  $W$  and  $T$ , i.e.  $\Psi \star W \sim T$ . Let  $\tilde{\Phi}_\mu = \mathcal{R}(W, A_\mu)$  be the deformation to register  $W$  to  $A_\mu$ , i.e.  $\tilde{\Phi}_\mu \star W \sim A_\mu$ . The transitivity property of  $\mathcal{R}(\cdot, \cdot)$  implies that  $\tilde{\Phi}_\mu = \Phi_\mu \circ \Psi$ , which is schematically illustrated in Fig. 1. The mapping  $\Phi_\mu$  relates  $\mathbf{x} \in \text{dom}(T)$  to its corresponding point in the domain of the image  $A_\mu$ . If another template  $W$  is used, the mapping  $\tilde{\Phi}_\mu$  relates each point  $\mathbf{y} \in \text{dom}(W)$  to its corresponding point in image domain. Due to the transitivity property the mapping relating both templates will be  $\Psi = \Phi_\mu^{-1} \circ \tilde{\Phi}_\mu$ , and each point  $\mathbf{y}$  is in correspondence with the point  $\mathbf{x} = \Psi(\mathbf{y})$ . Within this framework each point  $\mathbf{x}$  of the template  $T$  can be given an anatomical label which coincides with the one in  $\mathbf{y} = \Psi^{-1}(\mathbf{x})$  of the template  $W$ . The anatomical label in template coordinates, either  $T$  or  $W$ , can be propagated to the corresponding instance location through the mappings either  $\Phi_\mu$  or  $\tilde{\Phi}_\mu$ .

Under the previous framework, the Jacobian matrix of  $\tilde{\Phi}_\mu$ , evaluated at a point  $\mathbf{y} \in \text{dom}(W)$ , is given by  $\tilde{\mathbf{J}}_\mu(\mathbf{y}) = \mathbf{J}_\mu(\Psi(\mathbf{y}))\mathbf{P}(\mathbf{y})$ , where  $\mathbf{P}(\mathbf{y}) = (\mathcal{D}\Psi)|_{\mathbf{y}}$ . Therefore, under a change of template, Jacobian matrices transform as  $\mathbf{J} \mapsto \mathbf{J}\mathbf{P}$  with  $\mathbf{P}$  being the Jacobian matrix of the deformation between the templates. In terms of the function  $\mathcal{R}(\cdot, \cdot)$  the effect of changing the template is written as

$$(\mathcal{D}\mathcal{R}(W, A_\mu))|_{\mathbf{y}} = (\mathcal{D}\mathcal{R}(T, A_\mu))|_{\mathcal{R}(W, T) \circ \mathbf{y}} (\mathcal{D}\mathcal{R}(W, T))|_{\mathbf{y}}.$$

Statistical analysis is voxel-wise performed at each template coordinate, either in  $\mathbf{x}$  coordinates for the template  $T$  or in  $\mathbf{y}$  coordinates for  $W$  (see Fig. 1). As the template is often arbitrarily chosen, the statistical analysis should be independent of that choice to avoid biased results. For a statistic based on distances, a sufficient condition to achieve template invariance is to use a distance which fulfills

$$\begin{aligned} d(\tilde{\mathbf{J}}_\mu(\mathbf{y}), \tilde{\mathbf{J}}_\nu(\mathbf{y})) &= d(\mathbf{J}_\mu(\mathbf{x})\mathbf{P}(\mathbf{y}), \mathbf{J}_\nu(\mathbf{x})\mathbf{P}(\mathbf{y})) \\ &= d(\mathbf{J}_\mu(\mathbf{x}), \mathbf{J}_\nu(\mathbf{x})), \end{aligned}$$

where  $\mathbf{x} = \Psi(\mathbf{y})$ . Therefore template invariance will hold when the statistic is based on a right-invariant distance, *i.e.* a distance fulfilling

$$d(\mathbf{J}_\mu(\mathbf{x})\mathbf{P}, \mathbf{J}_\nu(\mathbf{x})\mathbf{P}) = d(\mathbf{J}_\mu(\mathbf{x}), \mathbf{J}_\nu(\mathbf{x})) \quad (2.1)$$

for any  $\mathbf{J}_\mu(\mathbf{x}), \mathbf{J}_\nu(\mathbf{x}), \mathbf{P} \in \text{GL}^+(n)$ .

Henceforth, for brevity of the notation, the dependence of the spatial variable will be omitted, and  $\mathbf{J}$  will refer to the Jacobian matrix at the location under study.

### 2.3 Jacobian Determinant

Most TBM studies perform voxel-wise statistics on the determinant of Jacobian matrices. The determinant of a composition of deformations is the product of their determinants. Therefore, determinants belong to the group of positive numbers under multiplication. It is a Lie group and an invariant distance on it is obtained by

$$d_{DET}(\mathbf{J}_\mu, \mathbf{J}_\nu) = |\log(\det(\mathbf{J}_\mu)) - \log(\det(\mathbf{J}_\nu))|. \quad (2.2)$$

It is easy to show that  $d_{DET}(\cdot, \cdot)$  fulfills Eq. (2.1) and therefore the template invariance requirement.

The determinant of a Jacobian matrix quantifies the local volume change induced by the deformation. Note that  $d_{DET}(\mathbf{J}_\mu, \mathbf{J}_\nu) = 0$  does not imply that  $\mathbf{J}_\mu = \mathbf{J}_\nu$ . In fact,  $d_{DET}(\mathbf{J}, \mathbf{LJ}) = d_{DET}(\mathbf{J}, \mathbf{JL}) = 0$  for any matrix  $\mathbf{L}$  with  $\det(\mathbf{L}) = 1$ .

### 2.4 Deformation Tensor

In the area of continuum mechanics a commonly used feature to measure the local deformation in the Lagrangian framework is the Cauchy–Green deformation tensor [52]  $\mathbf{C} = \mathbf{J}^T \mathbf{J}$ . The tensor  $\mathbf{C}$  is a symmetric positive definite (SPD) matrix which measures local length changes [52]. Under a change of template, Jacobian matrices transform as  $\mathbf{J} \mapsto \mathbf{JP}$  and therefore Cauchy–Green deformation tensors transform as  $\mathbf{C} \mapsto \mathbf{P}^T \mathbf{CP}$ .

In [12, 42] a distance between SPD matrices was proposed<sup>1</sup>  $d_{SPD}(\mathbf{C}_\mu, \mathbf{C}_\nu) = \|\logm((\mathbf{C}_\mu)^{-1/2}(\mathbf{C}_\nu)(\mathbf{C}_\mu)^{-1/2})\|_F$  satisfying the following invariance:  $d_{SPD}(\mathbf{C}_\mu, \mathbf{C}_\nu) = d_{SPD}(\mathbf{P}^T \mathbf{C}_\mu \mathbf{P}, \mathbf{P}^T \mathbf{C}_\nu \mathbf{P})$  for any non-singular linear transformation  $\mathbf{P}$ . The distance  $d_{SPD}$  induces a distance between Jacobian matrices

$$d_{AFF}(\mathbf{J}_\mu, \mathbf{J}_\nu) = \left\| \logm \left( (\mathbf{J}_\mu^T \mathbf{J}_\mu)^{-1/2} (\mathbf{J}_\nu^T \mathbf{J}_\nu) (\mathbf{J}_\mu^T \mathbf{J}_\mu)^{-1/2} \right) \right\|_F \quad (2.3)$$

<sup>1</sup> $\logm(\cdot)$  denotes the matrix logarithm (the inverse of the matrix exponential) and  $\|\cdot\|_F$  the Frobenius norm.

satisfying the right-invariance property  $d_{AFF}(\mathbf{J}_\mu, \mathbf{J}_\nu) = d_{AFF}(\mathbf{J}_\mu \mathbf{P}, \mathbf{J}_\nu \mathbf{P})$  for any matrix  $\mathbf{P}$  with positive determinant, and therefore fulfills the template invariance requirement (2.1).

As in the case of  $d_{DET}$ , there may exist a pair of different elements with zero distance. Specifically  $d_{AFF}(\mathbf{J}, \mathbf{RJ}) = 0$  for any rotation matrix  $\mathbf{R}$ . Note that, for any Jacobian matrix  $\mathbf{J}$ , the set of  $\mathbf{RJ}$  is a proper subset of  $\mathbf{LJ}$ , where  $\det(\mathbf{L}) = 1$  and therefore  $d_{AFF}$  is sensitive to a larger set of deformations than  $d_{DET}$ .

Another distance between SPD matrices is commonly used in TBM studies [19, 37]. Using the Log-Euclidean framework [6, 7], a distance on SPD matrices,  $d_{LE}(\mathbf{C}_\mu, \mathbf{C}_\nu) = \|\logm(\mathbf{C}_\mu) - \logm(\mathbf{C}_\nu)\|_F$ , was proposed. However, this distance is not invariant under general linear transformations although it is invariant under similarity transformations (if  $\mathbf{G} = s\mathbf{R}$  is a rotation followed by an isotropic positive scaling, then  $d_{LE}(\mathbf{C}_\mu, \mathbf{C}_\nu) = d_{LE}(\mathbf{G}^T \mathbf{C}_\mu \mathbf{G}, \mathbf{G}^T \mathbf{C}_\nu \mathbf{G})$ ).

## 3 A Metric on $\text{GL}^+(n)$

The two previous distances  $d_{DET}$  and  $d_{AFF}$  fulfill the template invariance requirement. However they do not satisfy the property that  $d(\mathbf{J}_\mu, \mathbf{J}_\nu) = 0$  if and only if  $\mathbf{J}_\mu = \mathbf{J}_\nu$  and accordingly some differences may not be measured. To overcome this drawback a right-invariant Riemannian metric on the space of Jacobian matrices is presented in the following.

### 3.1 Invariant Riemannian Distances on Lie Groups

Let  $\mathcal{M}$  be a differentiable manifold and  $T_q \mathcal{M}$  its tangent space at the element  $q \in \mathcal{M}$ . A Riemannian metric  $(\mathcal{M}, \langle u, v \rangle_q)$  on  $\mathcal{M}$  is a smooth assignment of inner products to every tangent space where  $q \in \mathcal{M}$  and  $u, v \in T_q \mathcal{M}$  [23]. Using this assignment, the length of a curve segment  $\gamma : [t_0, t_1] \subset \mathbb{R} \rightarrow \mathcal{M}$  is defined as

$$\text{Length}(\gamma; t_0, t_1) = \int_{t_0}^{t_1} \sqrt{\langle \dot{\gamma}(s), \dot{\gamma}(s) \rangle_{\gamma(s)}} ds.$$

A geodesic segment between two elements  $q$  and  $w$  belonging to  $\mathcal{M}$  is an arc-length parameterized curve segment in  $\mathcal{M}$  which locally minimizes the length. The Riemannian distance between  $q$  and  $w$  is the length of the shortest geodesic segment connecting both elements.

Geodesics can be uniquely described by an initial point  $q \in \mathcal{M}$  and an initial velocity  $v \in T_q \mathcal{M}$ . This description is related to the Riemannian exponential function [2]  $\text{Exp}_q(v)$ . The curve  $\gamma(t) = \text{Exp}_q(tv)$  is a geodesic and the length of the segment from  $t = 0$  to  $t = 1$  is equal to  $\sqrt{\langle v, v \rangle_q}$ .

Besides, the manifold of Jacobian matrices,  $\text{GL}^+(n)$ , has a Lie group structure (defined by the matrix-matrix product) [61] and therefore their elements, curves, tangent spaces and velocities can be translated using the group action [29]. Let  $\mathcal{G} \equiv \mathcal{M}$  be a manifold which, together with a *product*

$(\bullet \bullet \bullet): \mathcal{G} \times \mathcal{G} \rightarrow \mathcal{G}$ , has a Lie group structure. Let  $e$  be the group identity and given  $q, w \in \mathcal{G}$  the following operations are defined:  $\mathbf{R}_w q = q \bullet w$  is called *right-translation*;  $\mathbf{L}_w q = w \bullet q$ , *left-translation*; and  $q^{-1}$  such that  $q^{-1} \bullet q = e$  is called *inversion*.

A right-invariant metric is a Riemannian metric which naturally arises in Lie groups. Under this metric geodesics and distances remain invariant under right-translations. The metric can be defined as an inner product at a single tangent space and propagated by translations from the right to the whole group

$$\langle u, v \rangle_q = \langle T_q \mathbf{R}_w u, T_q \mathbf{R}_w v \rangle_{\mathbf{R}_w q}, \quad (3.1)$$

where  $T_q \mathbf{R}_w$  is the tangent-lift of the  $\mathbf{R}_w$  operation [32] which translates a velocity from  $T_q \mathcal{G}$  to  $T_{\mathbf{R}_w q} \mathcal{G}$ . Usually the inner product is defined at the *group algebra*  $\mathfrak{g} \equiv T_e \mathcal{G}$ . Additionally, under a right-invariant metric

$$\mathbf{R}_w \text{Exp}_q(v) = \text{Exp}_q(v) \bullet w = \text{Exp}_{\mathbf{R}_w q}(T_q \mathbf{R}_w v)$$

and in particular, setting  $w = q^{-1}$

$$\text{Exp}_q(v) = \mathbf{R}_q \text{Exp}_e(T_q \mathbf{R}_{q^{-1}} v) = \text{Exp}_e(T_q \mathbf{R}_{q^{-1}} v) \bullet q. \quad (3.2)$$

A left-invariant metric can also be defined by replacing in Eq. (3.1) the right-translation operation  $\mathbf{R}_w$  by the left-translation  $\mathbf{L}_w$ . Otherwise, the left-invariant Riemannian exponential function can be computed from the right-invariant one by

$$\text{Exp}_q^{\text{left}}(v) = \left( \text{Exp}_{q^{-1}}^{\text{right}} \left( -(T_{q^{-1}} \mathbf{R}_q)^{-1} T_q \mathbf{L}_{q^{-1}} v \right) \right)^{-1}. \quad (3.3)$$

In the case that  $\mathcal{G}$  is a matrix group like  $\text{GL}^+(n)$ , the right- (left-) translation takes the form  $\mathbf{R}_W \mathbf{Q} = \mathbf{Q} \mathbf{W}$  (correspondingly  $\mathbf{L}_W \mathbf{Q} = \mathbf{W} \mathbf{Q}$ ) for any  $\mathbf{W}, \mathbf{Q} \in \mathcal{G}$ . Moreover,  $T_{\mathbf{Q}} \mathbf{R}_W \mathbf{V} = \mathbf{V} \mathbf{W}$  (correspondingly  $T_{\mathbf{Q}} \mathbf{L}_W \mathbf{V} = \mathbf{W} \mathbf{V}$ ) for  $\mathbf{V} \in T_{\mathbf{Q}} \mathcal{G}$ .

### 3.2 A Riemannian Right-Invariant Distance on $\text{GL}^+(n)$

Recently, a closed-form expression was given in [5] for the left-invariant Riemannian exponential on  $\text{GL}^+(n)$  for the case where  $\langle \mathbf{U}_1, \mathbf{U}_2 \rangle_{\mathbf{I}} = \text{trace}(\mathbf{U}_1^T \mathbf{U}_2)$ , with  $\mathbf{U}_1$  and  $\mathbf{U}_2$  being elements of  $T_{\mathbf{I}} \text{GL}^+(n)$ ,

$$\text{Exp}_{\mathbf{Q}}^{\text{left}}(\mathbf{V}) = \mathbf{Q} \expm(\mathbf{V}^T \mathbf{Q}^{-T}) \expm(\mathbf{Q}^{-1} \mathbf{V} - \mathbf{V}^T \mathbf{Q}^{-T}), \quad (3.4)$$

where  $\mathbf{Q} \in \text{GL}^+(n)$ ,  $\mathbf{V} \in T_{\mathbf{Q}} \text{GL}^+(n)$  and  $\expm(\mathbf{M}) = \sum_k \mathbf{M}^k / k!$  is the matrix exponential.<sup>2</sup>

<sup>2</sup>There is no global consensus about what a ‘closed-form’ expression means [16]. However, the matrix exponential function,  $\expm(\cdot)$ , can be

The right-invariant Riemannian exponential function can be obtained in closed-form applying the identity (3.3) to Eq. (3.4)

$$\begin{aligned} \text{Exp}_{\mathbf{Q}}^{\text{right}}(\mathbf{V}) &= \left( \text{Exp}_{\mathbf{Q}^{-1}}^{\text{left}} \left( -\mathbf{Q}^{-1} \mathbf{V} \mathbf{Q}^{-1} \right) \right)^{-1} \\ &= \expm(\mathbf{V} \mathbf{Q}^{-1} - \mathbf{Q}^{-T} \mathbf{V}^T) \expm(\mathbf{Q}^{-T} \mathbf{V}^T) \mathbf{Q} \end{aligned}$$

and, in particular for  $\mathbf{Q} = \mathbf{I}$ ,

$$\text{Exp}_{\mathbf{I}}^{\text{right}}(\mathbf{U}) = \expm(\mathbf{U} - \mathbf{U}^T) \expm(\mathbf{U}^T), \quad (3.5)$$

where, again, the inner product defined at the identity is  $\langle \mathbf{U}_1, \mathbf{U}_2 \rangle_{\mathbf{I}} = \text{trace}(\mathbf{U}_1^T \mathbf{U}_2)$ .

From the right-invariant Riemannian metric on  $\text{GL}^+(n)$  the following distance is induced on elements of the same connected component of the group:  $d_{RI}(\mathbf{J}_\mu, \mathbf{J}_v) = \sqrt{\langle \mathbf{V}^*, \mathbf{V}^* \rangle_{\mathbf{J}_\mu}}$ , where  $\mathbf{V}^*$  is the smallest initial velocity satisfying  $\text{Exp}_{\mathbf{J}_\mu}(\mathbf{V}) = \mathbf{J}_v$ . The distance inherits the right-invariance from the Riemannian metric, and therefore  $d_{RI}(\mathbf{J}_\mu, \mathbf{J}_v) = d_{RI}(\mathbf{J}_\mu \mathbf{P}, \mathbf{J}_v \mathbf{P})$  for any  $\mathbf{P}$  in  $\text{GL}^+(n)$ , and therefore the invariance under the template holds. In addition,  $d_{RI}$  fulfills that  $d_{RI}(\mathbf{J}_\mu, \mathbf{J}_v) = 0$  if and only if  $\mathbf{J}_\mu = \mathbf{J}_v$ .

In this work the Riemannian distance is computed by solving the following problem:

$$\underset{\mathbf{V} \in T_{\mathbf{J}_\mu} \text{GL}^+(n)}{\text{minimize}} \left\| \text{Exp}_{\mathbf{J}_\mu}(\mathbf{V}) - \mathbf{J}_v \right\|_F^2, \quad (L1)$$

where the Frobenius norm of the Euclidean difference is chosen as the objective function for simplicity.

In general, the Riemannian exponential function under a right-invariant metric is a surjective mapping to the connected component of the group [28]. As the group  $\text{GL}^+(n)$  consists of only one connected component, then the existence of an initial velocity to generate a geodesic connecting  $\mathbf{J}_\mu$  and  $\mathbf{J}_v$  is guaranteed for any pair.

Using the right-invariance property in Eq. (3.2), it is possible to reformulate the problem (L1) as

$$\underset{\mathbf{U} \in \mathfrak{gl}(n)}{\text{minimize}} \left\| \text{Exp}_{\mathbf{I}}(\mathbf{U}) - \mathbf{J}_v \mathbf{J}_\mu^{-1} \right\|_F^2, \quad (L2)$$

where  $\mathfrak{gl}(n) \equiv \mathbb{R}^{n \times n}$  is the Lie algebra of  $\text{GL}^+(n)$ . Due to the right-invariance of the metric,  $\langle \mathbf{U}^*, \mathbf{U}^* \rangle_{\mathbf{I}} = \langle \mathbf{V}^*, \mathbf{V}^* \rangle_{\mathbf{J}_\mu}$  and therefore,  $d_{RI}(\mathbf{J}_\mu, \mathbf{J}_v) = \sqrt{\langle \mathbf{U}^*, \mathbf{U}^* \rangle_{\mathbf{I}}}$ , where  $\mathbf{U}^*$  is the smallest solution of (L2).

accurately and efficiently computed by most scientific software packages like Mathematica and Matlab [45].

In order to compute  $\mathbf{U}^*$ , a descent procedure can be used to solve (L2). The derivative with respect to  $\mathbf{U}$  of the objective function  $E(\mathbf{U}; \mathbf{J}_v \mathbf{J}_\mu^{-1}) = \|\text{Exp}_{\mathbf{I}}(\mathbf{U}) - \mathbf{J}_v \mathbf{J}_\mu^{-1}\|_F^2$  is

$$\mathcal{D}_{\mathbf{U}} E(\mathbf{U}; \mathbf{J}_v \mathbf{J}_\mu^{-1}) = 2 \left( \overline{\text{Exp}_{\mathbf{I}}(\mathbf{U})} - \overline{\mathbf{J}_v \mathbf{J}_\mu^{-1}} \right)^T \mathcal{D}_{\mathbf{U}} \text{Exp}_{\mathbf{I}}(\mathbf{U}),$$

where  $\overline{\mathbf{M}}$  denote the  $n^2$  dimensional vector resulting by the stacking of the columns of  $\mathbf{M}$ .  $\mathcal{D}$  is the Fréchet derivative operator: for a matrix function of matrix argument  $\mathbf{F}: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{p \times q}$ ,  $\mathcal{D}_{\mathbf{M}} \mathbf{F}(\mathbf{M})$  is a  $(pq) \times (mn)$  linear operator fulfilling  $[\mathcal{D}_{\mathbf{M}} \mathbf{F}(\mathbf{M})]^{i,j} = \partial_{\overline{\mathbf{M}}} \mathbf{F}(\mathbf{M})^i$ . Some rules to compute derivatives of matrix functions are in [1, Chap. 13], [40].

Taking derivatives of Eq. (3.5) it is obtained<sup>3</sup>

$$\begin{aligned} \mathcal{D}_{\mathbf{U}} \text{Exp}_{\mathbf{I}}(\mathbf{U}) &= (\mathbf{I}_n \otimes \text{expm}(\mathbf{U} - \mathbf{U}^T)) \text{dexpm}(\mathbf{U}^T) \mathbf{K}_{nn} \\ &\quad + (\text{expm}(\mathbf{U}) \otimes \mathbf{I}_n) \text{dexpm}(\mathbf{U} - \mathbf{U}^T) (\mathbf{I}_{n^2} - \mathbf{K}_{nn}), \end{aligned} \quad (3.6)$$

where  $\mathbf{I}_m$  is the  $m \times m$  identity matrix,  $\mathbf{K}_{nn}$  is the commutation matrix [1] defined by  $\mathbf{K}_{nn} \overline{\mathbf{M}} = \overline{\mathbf{M}^T}$  for an  $n \times n$  matrix  $\mathbf{M}$ , and  $\text{dexpm}(\cdot)$  is the Fréchet derivative of the matrix exponential function (see Appendix A).

The optimization problem to compute  $d_{RI}$  can be solved by means of a descent procedure along the negated gradient or a modified direction [50] with a line-search strategy.

Once the optimal initial velocity  $\mathbf{U}^*$  solving (L2) is obtained, the length of the geodesic segment is  $\sqrt{\langle \mathbf{U}^*, \mathbf{U}^* \rangle_{\mathbf{I}}} = \sqrt{\text{trace}(\mathbf{U}^{*T} \mathbf{U}^*)} = \|\mathbf{U}^*\|_F$ . If its Riemannian exponentiation generates the shortest curve segment between  $\mathbf{I}$  and  $\mathbf{J}_v \mathbf{J}_\mu^{-1}$ , then  $d_{RI}(\mathbf{J}_\mu, \mathbf{J}_v) = d_{RI}(\mathbf{I}, \mathbf{J}_v \mathbf{J}_\mu^{-1}) = \|\mathbf{U}^*\|_F$ .

While the existence of a zero minimizer is guaranteed, a serious drawback of this formulation is the non-uniqueness of the velocities satisfying  $\text{Exp}_{\mathbf{I}}(\mathbf{U}) = \mathbf{J}_v \mathbf{J}_\mu^{-1}$ . There may exist different initial velocities which generate geodesic segments between  $\mathbf{J}_\mu$  and  $\mathbf{J}_v$  (like in the rotations case where adding a rotation of  $2\pi$  the same rotation is obtained). Finding the shortest one in the general case is an open problem.

An initial estimate for  $\mathbf{U}$  is needed to start the optimization. It can be obtained by solving the initial velocity problem in a simpler subgroup of  $\text{GL}^+(n)$  where the solution is given by the matrix logarithm.

Let  $(\text{SO}(n) \times \text{S}^+(1))$  be the direct product of the rotation group in  $\mathbb{R}^n$  and the group of isotropic positive scalings. Its algebra,  $(\mathfrak{so}(n) \oplus \mathfrak{s}(1))$ , is the set of matrices of the form  $\mathbf{S} + \alpha \mathbf{I}_n$  where  $\mathbf{S}$  is an  $n \times n$  skew-symmetric matrix

and  $\alpha \in \mathbb{R}$ . Under the metric  $\langle \mathbf{U}, \mathbf{V} \rangle_{\mathbf{I}} = \text{trace}(\mathbf{U}^T \mathbf{V})$ , the Riemannian exponential over  $(\text{SO}(n) \times \text{S}^+(1))$  generates curves which are also geodesics in  $\text{GL}^+(n)$ , i.e.  $(\text{SO}(n) \times \text{S}^+(1))$  is a totally geodesic subgroup of  $\text{GL}^+(n)$  under the given metric [44]. For a matrix  $\mathbf{H} = \mathbf{S} + \alpha \mathbf{I}_n$  the Riemannian exponential function (3.5) results in  $\text{expm}(\mathbf{H} - \mathbf{H}^T) \text{expm}(\mathbf{H}^T) = \text{expm}(\mathbf{H})$  because in this case,  $(\mathbf{H} - \mathbf{H}^T)$  and  $\mathbf{H}^T$  commute. Therefore, the matrix logarithm of a matrix  $\mathbf{G} \in (\text{SO}(n) \times \text{S}^+(1))$  gives the solution of the problem (L2) when it is restricted to the group  $(\text{SO}(n) \times \text{S}^+(1))$ . The proposed initial estimate is the logarithm of the closest matrix, in the Frobenius norm sense, belonging to  $(\text{SO}(n) \times \text{S}^+(1))$  to the target matrix  $\mathbf{J}_v \mathbf{J}_\mu^{-1}$

$$\mathbf{U}^{\text{initial}} = \text{logm} \left( \arg \min_{\mathbf{G} \in (\text{SO}(n) \times \text{S}^+(1))} \left\| \mathbf{G} - \mathbf{J}_v \mathbf{J}_\mu^{-1} \right\|_F^2 \right). \quad (3.7)$$

To compute this, let  $\mathbf{J}_v \mathbf{J}_\mu^{-1} = \mathbf{Z} \mathbf{D} \mathbf{X}^T$  be the singular value decomposition of the target matrix, then the matrix  $\mathbf{G}^* = \text{trace}(\mathbf{D})/n \mathbf{Z} \text{diag}(1, 1, \dots, 1, \det(\mathbf{Z} \mathbf{X}^T)) \mathbf{X}^T$  is the solution of the minimization problem in Eq. (3.7).

Theoretical details about convergence and illustrative examples of performance when using gradient descent and Gauss–Newton algorithms are given in Appendix B.

## 4 Cramér Test

Three different distance functions over Jacobian matrices have been considered in previous sections:  $d_{DET}$ ,  $d_{AFF}$  and  $d_{RI}$ . Most statistical analysis techniques are defined in terms of a distance function between observations, and the Cramér two-sample test is used in this work [11, 60, 65]. This test was selected because its statistic can be computed using only the distances between the instances. In addition, it can be directly used on both univariate and multivariate data. The statistic for the Cramér two-sample test is computed at each template location  $\mathbf{x}$  as

$$\begin{aligned} \sigma(\mathbf{x}) &= \frac{n_a n_b}{n_a + n_b} \left( \frac{1}{n_a n_b} \sum_{a=1}^{n_a} \sum_{b=1}^{n_b} D_{ab} - \frac{1}{2n_a^2} \sum_{a=1}^{n_a} \sum_{a'=1}^{n_a} D_{aa'} \right. \\ &\quad \left. - \frac{1}{2n_b^2} \sum_{b=1}^{n_b} \sum_{b'=1}^{n_b} D_{bb'} \right), \end{aligned}$$

where  $D_{\mu\nu} = d(\mathbf{J}_\mu(\mathbf{x}), \mathbf{J}_\nu(\mathbf{x}))$  is the inter-element distance matrix and  $n_a$  and  $n_b$  are the cardinalities of the sets  $\mathcal{A}$  and  $\mathcal{B}$ , respectively.

Statistical significance of  $\sigma(\mathbf{x})$  can be assessed by means of a random permutation test of the group labels [49]. The null hypothesis is that instances from both groups follow the same distribution. The distribution of the statistic under the null hypothesis is empirically estimated and it is rejected for

<sup>3</sup>Let  $\mathbf{A}$  and  $\mathbf{B}$  be two  $n \times n$  matrices, then  $\overline{\mathbf{AB}} = (\mathbf{B}^T \otimes \mathbf{I}_n) \overline{\mathbf{A}}$  and therefore  $\mathcal{D}_{\mathbf{A}}(\mathbf{AB}) = \mathbf{B}^T \otimes \mathbf{I}_n$ . Similarly,  $\overline{\mathbf{AB}} = (\mathbf{I}_n \otimes \mathbf{A}) \overline{\mathbf{B}}$  and therefore  $\mathcal{D}_{\mathbf{B}}(\mathbf{AB}) = \mathbf{I}_n \otimes \mathbf{A}$ .

large enough values of  $\sigma(\mathbf{x})$ . The  $p$ -value is the proportion of the permutations having a  $\sigma(\mathbf{x})$  value larger or equal than the value without relabeling. Note that the statistic for a random permutation can be obtained without recomputing the distances between the Jacobian matrices which is the most intensive computational task.

In a TBM study the hypothesis test is performed at each template location. It is well-known that a multiple comparison problem appears when performing a large number of tests. Two criteria can be used for correcting the increase of type I error: family wise error rate (FWE) and false discovery rate (FDR) [14]. In this work the FDR criterion was used for the correction of the proportion of false-positives among the rejected null hypotheses.

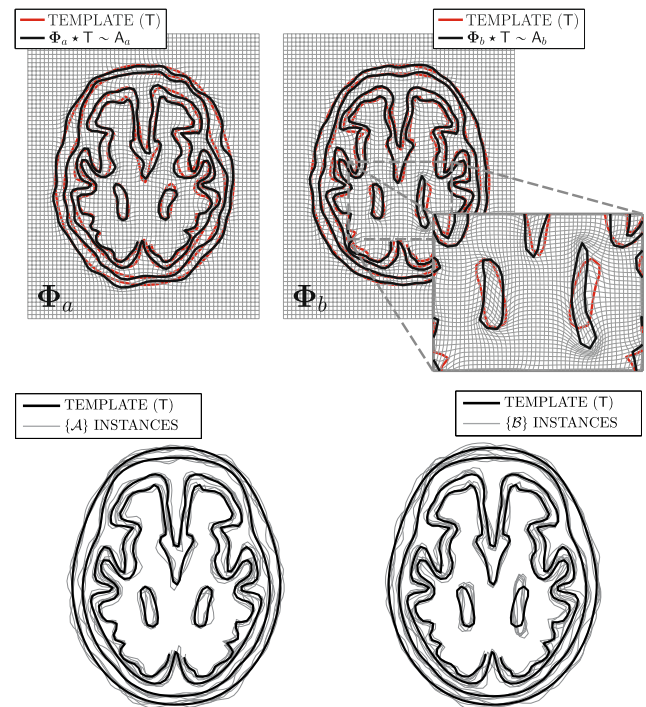
## 5 Results

### 5.1 Synthetic Study

A synthetic study was designed to illustrate the results of the three right-invariant distances in a TBM application environment.

In order to facilitate the visualization of the results the experiment was performed generating deformations on a 2D image. Two sets of 50 random deformations were generated starting from a synthetic template. The first set was designed to represent the anatomical variability within a ‘control’ group, while the second set is aimed at representing a ‘pathological’ group with group-driven anatomical differences and intra-group variability. Representative examples of the deformation maps are illustrated in Fig. 2. The contours are drawn as a visualization aid, but the actual input data for the analysis are the set of spatial deformations  $\Phi_\mu$ . The bottom row in Fig. 2 shows 5 instances from ‘control’ and ‘pathological’ groups depicting the intra-group variabilities.

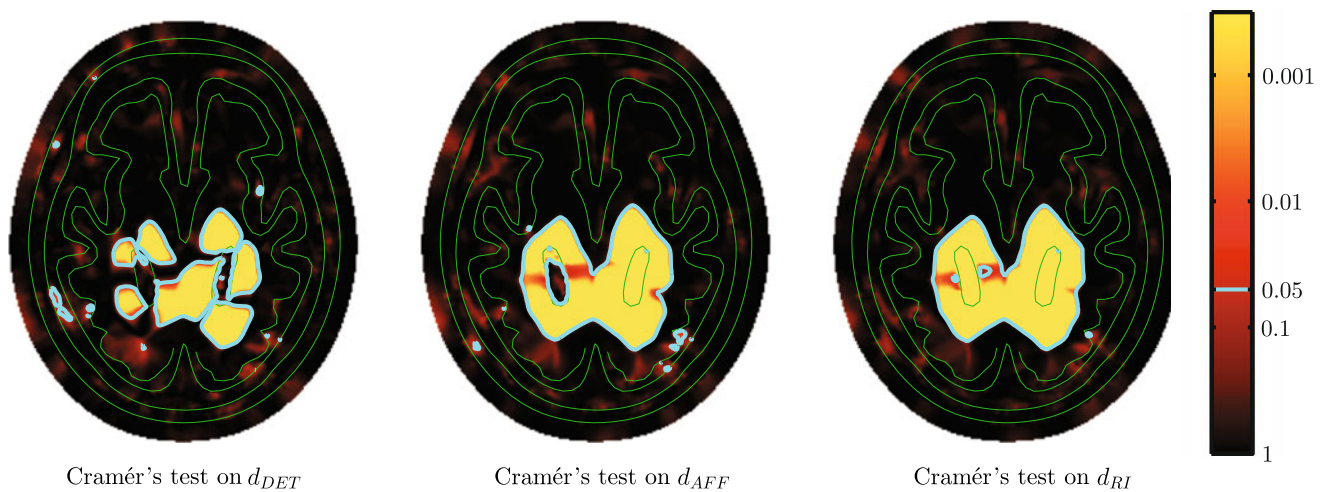
Deformation instances from ‘control’ group  $\mathcal{A}$  were modeled as smooth and invertible random deformations. Instances from ‘pathological’ group  $\mathcal{B}$  were modeled as the composition of a common group-driven deformation and random deformations similar to the ones used for ‘control’ group. The group-driven deformation was designed to produce the following changes on the ‘subcortical structures’: a clockwise rotation of the structure in the ‘left hemisphere’ of 15 degrees; a counter-clockwise rotation of the structure in the ‘right hemisphere’ of 15 degrees and a subsequent anisotropic scaling with factors 0.7 and 1/0.7 along the horizontal and vertical directions respectively. Note that the group-driven deformation preserves the volume of the ‘subcortical structures’ but the surrounding regions suffer a more complex deformation due to the continuity of the mapping. This effect can be seen in the zoomed panel of Fig. 2.



**Fig. 2** Synthetic data for ‘control’ group (left column) and ‘pathological’ group (right column). Top row: example of the deformation maps  $\Phi_a$  and  $\Phi_b$  mapping the template to given instances. Bottom row: five contours illustrating the intra-group anatomical variability

All deformations were obtained by integration of stationary velocity fields. Velocity fields were parameterized with cubic splines on a  $220 \times 280$  regular grid and circular boundary conditions were imposed. The deformed location  $\Phi(\mathbf{x})$  was obtained by the integration along the velocity field up to time 1 of a particle with initial position  $\mathbf{x}$  [17]. Accordingly, the corresponding Jacobian matrix  $\mathbf{J}(\mathbf{x})$  is computed by integrating up to time 1 the local deformation guided by the Jacobian of the velocity field [8]. With this parametrization it is guaranteed that deformation maps result in invertible and differentiable mappings and therefore their Jacobian matrices belong to  $GL^+(2)$ . In this way, the typical numerical issues by using a finite difference scheme on grid locations are avoided. Whenever compositions of deformations are required, the corresponding velocity fields are successively integrated and interpolations of the deformation mappings are avoided.

Statistical analyses were performed using the previous defined distances:  $d_{DET}$ ,  $d_{AFF}$  and  $d_{RI}$ . At each grid location of the template Cramér tests were computed and a set of 100,000 random permutations were performed in order to assess the  $p$ -value. In order to correct for multiple comparisons, the FDR criterion was used [14]. In addition, a Student’s  $t$ -test was performed on the log of the Jacobian matrix determinant because this is one of the most widely used approaches in TBM studies.



**Fig. 3** Statistical maps of FDR corrected  $p$ -values for Cramér test based on the three distances. Template contours are illustrated in green colour for localization purposes. Cyan contours show the boundary of the regions with significant differences with the criterion of  $p_{FDR} \leq 0.05$

Figure 3 shows the FDR corrected  $p$ -value maps corresponding to the Cramér tests. The statistical map of the Student's  $t$ -test on the logarithm of the Jacobian determinant (not shown in the figure) was very similar to the one obtained by the Cramér test on  $d_{DET}$ . The distance  $d_{DET}$  was not able to detect statistically significant differences in local volume change in the interior of the 'subcortical structures'. However, significant differences were found at outer regions surrounding these structures. This behavior shows that the deformations driven by a rotation of a structure surrounded by a static region mainly generates local volume changes outside the structure.

As was expected, the statistical map using  $d_{AFF}$  shows significant differences in the interior of the 'right subcortical structure', where there is an anisotropic scaling of the structure. However, no significant differences were found in the interior of the 'left subcortical structure' because the deformations were mainly a rotation. Significant differences were also found in the surrounding regions.

Regarding the statistical map using  $d_{RI}$ , it can be seen that significant differences were found in the interior of both 'subcortical structures'. Looking at  $d_{AFF}$  and  $d_{RI}$  statistical maps it is clearly seen that the main differences are located in the interior of the left 'subcortical structure' which is the only region with a pure rotation. An increasing sensitivity of the statistical test is obtained with a more complete geometrical descriptions of the local deformation.

## 5.2 ADNI Dataset

A subset of T1-MRI brain images was selected from the ADNI database ([adni.loni.ucla.edu](http://adni.loni.ucla.edu)). The ADNI was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering,

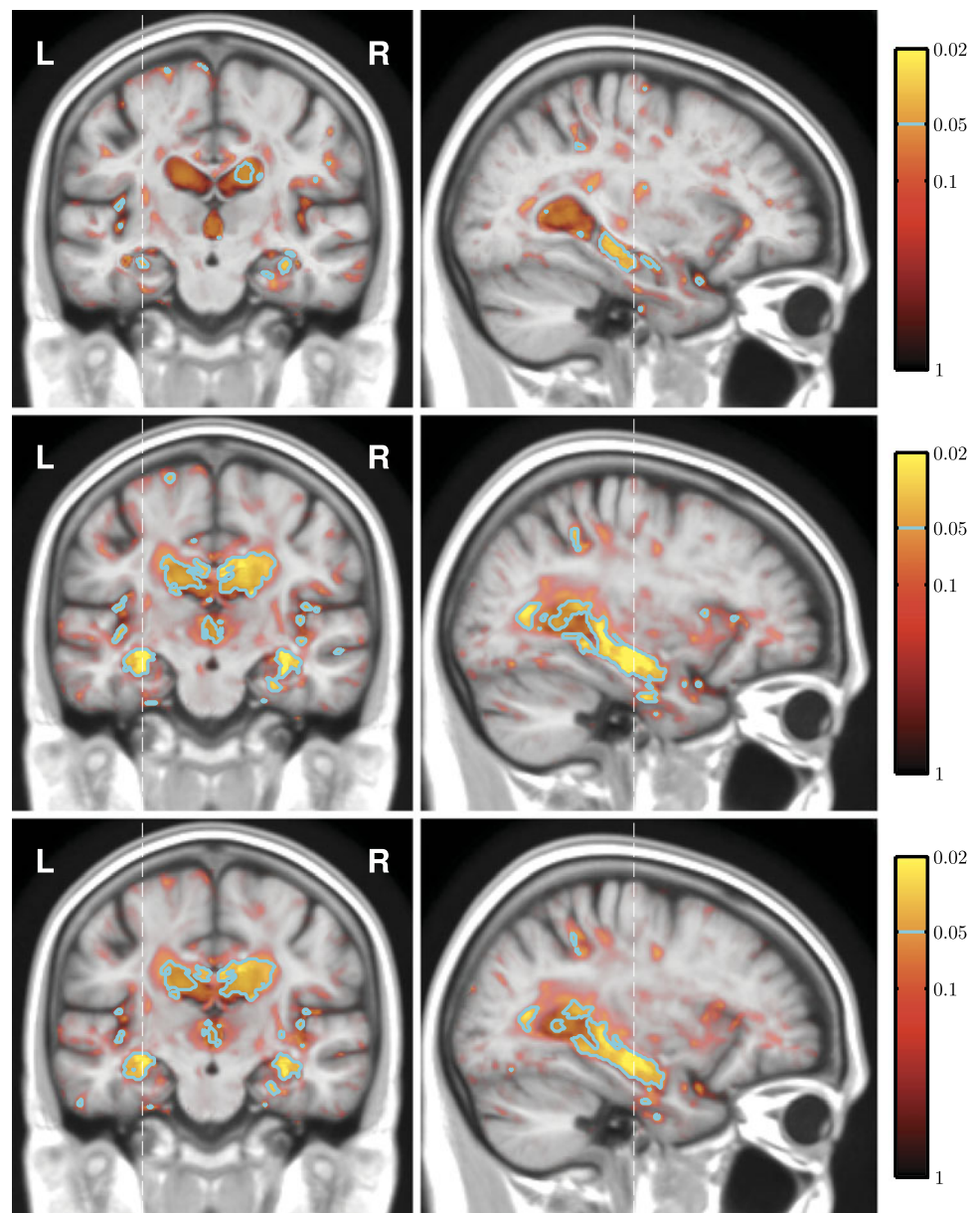
**Table 1** Descriptive statistics of demographic and clinical information (mean $\pm$ std). \* denotes statistically significant differences under Student's  $t$ -test ( $p$ -value  $\leq 0.05$ )

	AGE	GENDER	MMSE*	memory CDR*
NOR group	72.2 $\pm$ 0.3	10/10	29.3 $\pm$ 1.0	0.0 $\pm$ 0.0
AD group	72.2 $\pm$ 0.5	10/10	23.3 $\pm$ 1.8	0.9 $\pm$ 0.3

the Food and Drug Administration, private pharmaceutical companies and non-profit organizations, as a 60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment and early Alzheimer's disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials.

Twenty elderly control subjects (denoted here as NOR group) and twenty AD patients (denoted as AD group) were selected from the database. As the brain atrophy is affected by factors such as age and gender, subjects were selected such that they were gender-matched, using a very narrow age interval. The demographic and clinical variables of the selected subjects are summarized in Table 1 where it can be seen that there were not significant differences between subject groups on age and gender variables. As expected, clinical scores, such as MiniMental State Examination (MMSE) [21] or the memory score in the Clinical Dementia Ratio (CDR) [46], were significantly different between both groups under the Student's  $t$ -test.

**Fig. 4** Illustrative coronal (*left*) and sagittal (*right*) views of FDR-corrected  $p$ -value maps of the Cramér test performed using three right-invariant distances:  $d_{DET}$  (*top row*),  $d_{AFF}$  (*middle row*) and  $d_{RI}$  (*bottom row*). L and R denote left and right hemisphere respectively. The dashed vertical lines show the location of the sagittal and coronal slices



The MRI brain images were calibrated with phantom-based geometric corrections to ensure consistency among scans acquired at different sites. Additional image corrections included geometric distortion correction, bias field correction and geometrical scaling. The pre-processed images are available to the scientific community and were downloaded from the ADNI website.<sup>4</sup>

The template  $T$  was built from 40 MRI of elderly control subjects as it is described in [18], being a  $1\text{ mm} \times 1\text{ mm} \times 1\text{ mm}$  voxel-size image where the brain parenchyma occupies about 2 million voxels. The deformation fields  $\Phi_\mu$  were estimated using a SVF diffeomorphic registration [18, 30]

between the template  $T$  and each MRI instance. The output of the SVF registration method is a velocity field that characterizes via time-integration the deformation mapping  $\Phi_\mu$ . Jacobian matrices were computed by integration (as was explained in the Synthetic study Sect. 5.1) ensuring that every  $J(\mathbf{x})$  belong to  $GL^+(3)$ .

Voxel-wise Cramér tests were performed using the three right-invariant distances  $d_{DET}$ ,  $d_{AFF}$  and  $d_{RI}$ . Critical values were estimated by means of a permutation test using 100,000 permutations. After, the  $p$ -values maps were corrected with the FDR criterion. Figure 4 shows a coronal and sagittal illustrative slice of the corrected  $p$ -value map for each distance.

It must be noted that the hypothesis tests were performed on very small samples (20 subjects per group) and the sta-

<sup>4</sup><http://www.adni-info.org>

tistical power is limited by this condition. Larger samples would require to covariate with fixed effects such as age and this statistical problem is harder than the group comparison considered here.

The three tests provide statistical maps which are meaningful according to the pathophysiological knowledge of the disease. As was expected, the number of voxels with significant group differences in the  $d_{DET}$  map was much smaller than in  $d_{AFF}$  and  $d_{RI}$  maps. The differences between  $d_{DET}$  and  $d_{AFF}-d_{RI}$  maps might be due to the higher statistical power achieved when considering the whole multivariate Jacobian matrix information rather than the simple scalar determinant. Regarding the visual comparison between  $d_{AFF}$  and  $d_{RI}$  maps, they look very similar. This suggests that pure rotations were not relevant on this experiment which is limited by the small sample sizes.

Regarding computational issues,  $d_{RI}$  is more expensive than  $d_{DET}$  and  $d_{AFF}$ . In our implementation, that uses the Gauss–Newton strategy and a relative error of  $10^{-10}$  as stopping criterion, the average computation time for  $d_{RI}$  was 50 msec. On the other hand, average computation times for  $d_{DET}$  and  $d_{AFF}$  were 0.02 msec and 1 msec, respectively.

## 6 Discussion

Many of the methodological contributions on TBM studies usually belong to the non-rigid registration stage while a few works are devoted to investigate the statistical analysis of the deformations. Most statistical analyses make use of Jacobian determinant. A more complete descriptor of the local deformation is considered in this work formulated as a Riemannian distance between Jacobian matrices. Results were illustrated on both synthetic and brain MRI data.

As was argued in the introduction, TBM is built on the anatomical correspondence between images and a template which is achieved when the template invariance requirement holds. A transitive registration method must be used to obtain the anatomical correspondence from any template choice. This work focused on the statistical analysis of deformations assuming that a transitive registration procedure is used. We show that under this assumption the template invariance of the statistical map can be achieved by using a statistic based on a right-invariant distance.

Three different right-invariant distances on  $GL^+(n)$  were considered in this work. The first two distances have been previously used in the literature: a distance between Jacobian determinants  $d_{DET}$  and a distance between deformation tensors  $d_{AFF}$ . The third one is a Riemannian distance between Jacobian matrices  $d_{RI}$ .

The three distances quantify different aspects of the local deformations. The determinant of the Jacobian matrix measures the local atrophy or expansion. The Cauchy–Green

deformation tensor measures the amount of symmetric deformation up to rotations. Also, the proposed distance  $d_{RI}$ , can be geometrically interpreted as a measure of the local deformation that quantifies the Riemannian energy between two Jacobian matrices. If a pure rotational transformation acts on a region, there is no local strain and therefore  $d_{AFF}$  is not sensitive to such deformations. In contrast,  $d_{RI}$  also quantifies local rotations. This qualitative difference was nicely observed in the synthetic example, in particular at the left ‘subcortical structure’.

As the geometrical interpretation of the local volume change is intuitive it may be interesting to decompose the multivariate descriptors in terms of volume change and the remaining deformation. The group  $GL^+(n)$  can be decomposed by the direct product  $(SL(n) \times S^+(1))$  of the *Special linear group* of  $n \times n$  matrices with determinant equal to 1, and the *isotropic positive scale group*. With this decomposition, it is possible to describe volume changes. In the present work, the selected Riemannian metric was  $\langle \mathbf{U}_1, \mathbf{U}_2 \rangle_I = \text{trace}(\mathbf{U}_1^T \mathbf{U}_2)$ . This metric corresponds to an identity metric tensor in the canonical basis of the algebra  $\mathfrak{gl}(n)$  and it was selected because a closed-form solution of the Riemannian exponential is known. However, a large family of possible metric tensors could have been selected. For a particular application, it is possible to adjust the metric tensor to quantify different aspects of the geometric deformation. Unluckily the known closed-form solution will be not valid anymore and numerical procedures with high computational load would be required.

As was mentioned in the introduction, the transitivity property is not fully accomplished by most of the registration algorithms currently available. In a general case the mappings  $\Phi_\mu$  and  $\tilde{\Phi}_\mu$  from the templates  $T$  and  $W$  respectively to each instance  $\mu$  shown in Fig. 1 can be written using an instance-dependent right-translation

$$\tilde{\Phi}_\mu = \Phi_\mu \circ \Psi_\mu. \quad (6.1)$$

The transitivity of the registration is obtained when the equivalence relation  $(\cdot \sim \cdot)$  between any pair of images is satisfied by a unique deformation map. Therefore, the non-transitivity implies that there exists more than one mapping registering any two images which can be understood as an uncertainty in the mapping  $\mathcal{R}(A, B)$ . From Eq. (6.1) it is straightforward that

$$\Phi_\mu \circ \Psi_\mu \circ \tilde{\Phi}_\mu^{-1} = \text{identity map}.$$

In the case of a transitive registration algorithm is used, all right-translations  $\Psi_\mu$  in Eq. (6.1) will be the same,  $\Psi_\mu = \mathcal{R}(W, T) = \Psi$  for all instances  $\mu$ . The difference between the mappings  $\Phi_\mu \circ \Psi \circ \tilde{\Phi}_\mu^{-1}$  and the identity map gives a measurement of the impact of the non-transitivity in a TBM study. When these differences are small enough relative to

the intergroup difference of the deformations, then the non-transitivity is not a relevant issue.

The Cramér test was selected in this work because its statistic can be computed in terms of the inter-element distances solely. It may be of interest to extend the use of the presented Riemannian distance to another statistical tools such as Hotelling's  $T^2$  test or regression. Unluckily in both cases, an optimization will be required to estimate the intrinsic or Karcher mean [53] on the  $GL^+(n)$  manifold and the computational cost will be even larger. Moreover, up to our knowledge, there is no parametric method to assess significance of hypothesis tests on Jacobian matrices and the use of random permutations would require to compute the intrinsic mean for each permutation.

**Acknowledgements** This work was funded by research grants TEC2009-14587-C03-01 from CICYT, AMIT project CEN-20101014 from CENIT program, CIM project IPT-2011-1638-900000 from IN-PACTO program, Spain.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129 and K01 AG030514.

## Appendix A: Fréchet Derivative of the Matrix Exponential

To compute the Fréchet derivative  $\mathcal{D}_U \text{Exp}_I(\mathbf{U})$  given in Eq. (3.6) an expression for  $\text{dexpm}(\mathbf{M}) \equiv \mathcal{D}_M \text{expm}(\mathbf{M})$  is required. This is the linear operator containing the derivatives of each element of  $\text{expm}(\mathbf{M})$  with respect to perturbation  $\mathbf{a}$  on each element of  $\mathbf{M}$  and results in an  $n^2 \times n^2$  matrix.

There are different attempts to compute  $\mathcal{D}_M \text{expm}(\mathbf{M})$  [3, 4, 38, 48]. In our computations we have used the approach given in [41], where for an analytic matrix function  $\mathbf{F}(\mathbf{M})$ :

$$\mathbf{F}\left(\begin{pmatrix} \mathbf{M} & \mathbf{P} \\ \mathbf{0} & \mathbf{M} \end{pmatrix}\right) = \begin{pmatrix} \mathbf{F}(\mathbf{M}) & d_r \mathbf{F}(\mathbf{M} + r\mathbf{P}) \\ \mathbf{0} & \mathbf{F}(\mathbf{M}) \end{pmatrix}$$

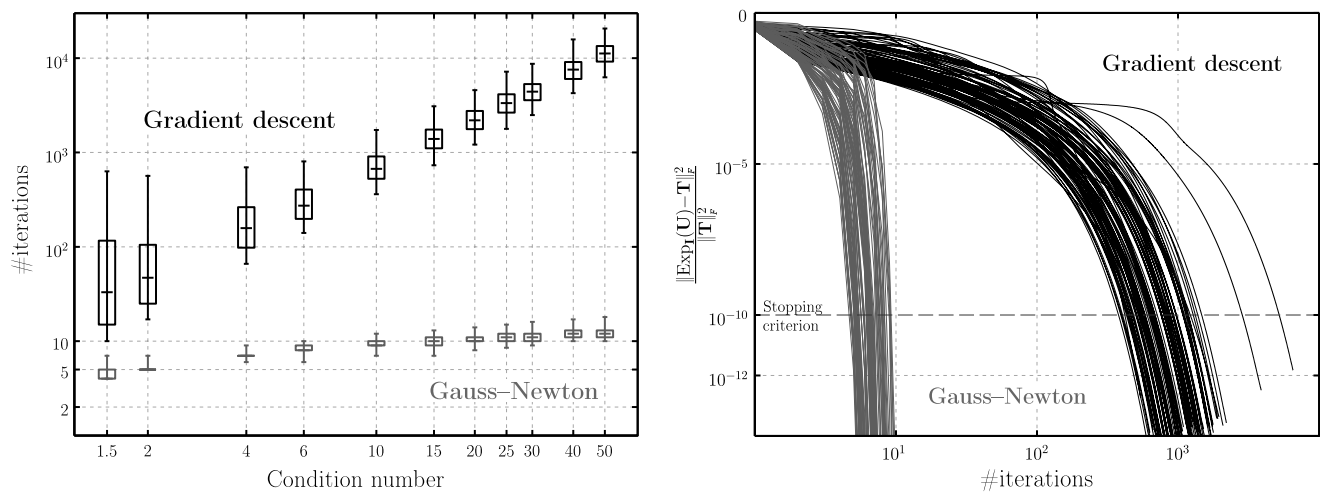
and if  $\mathbf{P}^j$  is the  $j$ -th canonical perturbation, then the vectorization of the upper-right submatrix  $d_r \mathbf{F}(\mathbf{M} + r\mathbf{P}^j)$  is the  $j$ -th column of the matrix  $\mathcal{D}_M \mathbf{F}(\mathbf{M})$ .

## Appendix B: Convergence and Performance of Descent Procedures

Below, arguments are given for using a descent method to compute the initial velocity of a geodesic starting at the identity  $\mathbf{I}$  to a target matrix  $\mathbf{T} \in GL^+(n)$ . The existence of such geodesic is guaranteed because the set  $GL^+(n)$  is connected. Moreover,  $\text{Exp}_I(\cdot)$  in Eq. (3.5) is surjective on  $GL^+(n)$  and therefore there always exists at least one zero-minimizer of the objective function  $E(\mathbf{U}; \mathbf{T})$ . In fact, it may happen that there exists many (even an infinite number of) different geodesics from the identity to  $\mathbf{T}$  and all their corresponding initial velocities are zero-minimizers of  $E(\mathbf{U}; \mathbf{Q})$ . Therefore, the objective function  $E(\mathbf{U}; \mathbf{Q})$  is in general not convex in  $gl(n)$  and its minimization procedure has many basins of attraction. Nevertheless, it can be shown that all local minima have zero energy and therefore they are also global minima. This can be proved by noticing that the function  $\text{Exp}_I(\cdot)$  maps open subsets of  $\mathbb{R}^{n \times n} \equiv gl(n)$  around  $\mathbf{U}$  to open subsets of  $\mathbb{R}^{n \times n} \supset GL^+(n)$  around  $\text{Exp}_I(\mathbf{U})$ . The same applies for the function  $\text{Exp}_I^{-1}(\cdot)$ . Moreover, the Frobenius norm is a convex function in  $\mathbb{R}^{n \times n}$ . Then, for any  $\mathbf{U} \in gl(n)$  either,  $E(\mathbf{U}; \mathbf{Q})$  is zero or there is an  $\mathbf{U}'$  in a neighborhood of  $\mathbf{U}$  with a lower value of the objective function.

It is remarkable that for some target matrices there may exist descending *valleys* in the objective function which extends up to infinity. For example, in the extreme case of  $\mathbf{T} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ , there exists a descent valley in  $gl(n)$  along the direction  $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$  whose Riemannian exponential ends up at the null matrix. Although we could not find a simple proof, we conjecture that those *attraction basins towards infinity* have a zero measure in  $gl(n)$ . For the previous matrix  $\mathbf{T}$ , descent procedures starting from any symmetric matrix tend to the null matrix, but by slightly perturbing with a skew-symmetric matrix the procedure converges to a global minimum. For a continuous descending path  $\mathbf{U}(t)$  such that  $\|\mathbf{U}(t)\|_F \rightarrow +\infty$  for  $t \rightarrow +\infty$ , it can be proved that  $\lim_{t \rightarrow +\infty} \text{Exp}_I(\mathbf{U}(t))$ , either it is a singular matrix or it does not exist. Then, to circumvent those *valleys towards infinity* it is convenient to provide to the descent algorithm with a control in the determinant of  $\text{Exp}_I(\mathbf{U}_k)$  (with  $\mathbf{U}_k$  the current iteration of the algorithm) and perform a perturbation if the algorithm is converging to a singular matrix.

To show the performance of descent strategies the following experiment was performed: one thousand  $3 \times 3$  random matrices were generated with positive determinant and with predefined condition numbers; for each target matrix a solution  $\mathbf{U}^*$  of the problem (L2) was found by using two



**Fig. 5** Performance of gradient descent and Gauss–Newton strategies using an exact line-search. *Left panel*: box-plot showing 5, 25, 50, 75 and 95 percentiles of the number of iterations needed to reach the

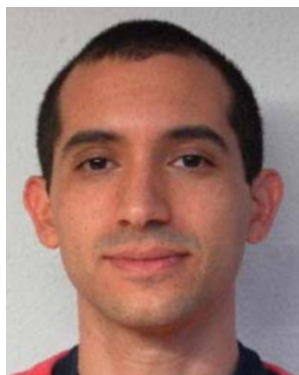
stopping criterion for target matrices with different condition numbers. *Right panel*: evolution of the relative error for 100 random target matrices with a condition number of 10

descent strategies, along the negated gradient direction and along the Gauss–Newton direction, both performed with an exact line-search; the stopping criterion was set to a relative error in the Frobenius norm smaller than  $10^{-10}$ . Figure 5 shows in the left panel the number of iterations needed to reach the stopping criterion for different condition numbers of the target matrix. Right panel in Fig. 5 shows the relative error for 100 random target matrices with a condition number of 10 for both descent strategies. It is very clear that the Gauss–Newton strategy requires a much smaller number of iterations than gradient descent.

## References

1. Abadir, K.M., Magnus, J.R.: Matrix Algebra. Cambridge Univ. Press, Cambridge (2005)
2. Absil, P.A., Mahony, R., Sepulchre, R.: Optimization Algorithms on Matrix Manifolds. Princeton Univ. Press, Princeton (2009)
3. Al-Mohy, A.H.: Algorithms for the matrix exponential and its Fréchet derivative. Ph.D. thesis, University of Manchester (2010)
4. Al-Mohy, A.H., Higham, N.J.: Computing the Fréchet derivative of the matrix exponential, with an application to condition number estimation. *SIAM J. Matrix Anal. Appl.* **30**(4), 1639–1657 (2009)
5. Andruchow, E., Larotonda, G., Recht, L., Varela, A.: The left invariant metric in the general linear group. *arXiv preprint* (2011)
6. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Fast and simple calculus on tensors in the log-Euclidean framework. In: *Medical Image Computing and Computer Assisted Intervention*, pp. 115–122. Springer, Berlin (2005)
7. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.* **29**(1), 328–347 (2007)
8. Ashburner, J.: A fast diffeomorphic image registration algorithm. *NeuroImage* **38**(1), 95–113 (2007)
9. Ashburner, J., Friston, K.J.: Diffeomorphic registration using geodesic shooting and Gauss–Newton optimisation. *NeuroImage* **55**(3), 954–967 (2011)
10. Avants, B.B., Schoenemann, P.T., Gee, J.C.: Lagrangian frame diffeomorphic image registration: morphometric comparison of human and chimpanzee cortex. *Med. Image Anal.* **10**(3), 397–412 (2006)
11. Baringhaus, L., Franz, C.: On a new multivariate two-sample test. *J. Multivar. Anal.* **88**(1), 190–206 (2004)
12. Batchelor, P.G., Moakher, M., Atkinson, D., Calamante, F., Connelly, A.: A rigorous framework for diffusion tensor calculus. *Magn. Reson. Med.* **53**(1), 221–225 (2005)
13. Beg, M.F., Miller, M.I., Trounev, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int. J. Comput. Vis.* **61**(2), 139–157 (2005)
14. Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995)
15. Boisvert, J., Cheriet, F., Pennec, X., Labelle, H., Ayache, N.: Geometric variability of the scoliotic spine using statistics on articulated shape models. *IEEE Trans. Med. Imaging* **27**(4), 557–568 (2008)
16. Borwein, J.M., Crandall, R.E.: Closed forms: what they are and why we care. *Not. Am. Math. Soc.* **60**, 50–65 (2013)
17. Bossa, M., Zacur, E., Olmos, S.: Algorithms for computing the group exponential of diffeomorphisms: performance evaluation. In: *CVPR Workshops*, pp. 23–28. IEEE (2008)
18. Bossa, M., Zacur, E., Olmos, S.: ADNI: Tensor-based morphometry with stationary velocity field diffeomorphic registration: application to ADNI. *NeuroImage* **51**(3), 956–969 (2010)
19. Boucher, M., Evans, A., Siddiqi, K.: Anisotropic diffusion of tensor fields for fold shape analysis on surfaces. In: *Information Processing in Medical Imaging*, pp. 271–282. Springer, Berlin (2011)
20. Christensen, G.E., Johnson, H.J.: Invertibility and transitivity analysis for nonrigid image registration. *J. Electron. Imaging* **12**(1), 106–117 (2003)
21. Cockrell, J.R., Folstein, M.F.: Mini-mental state examination (MMSE). *Psychopharmacol. Bull.* **24**(4), 689–692 (1988)
22. Davatzikos, C., Genc, A., Xu, D., Resnick, S.M.: Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. *NeuroImage* **14**(6), 1361–1369 (2001)
23. Do Carmo, M.P.: Riemannian Geometry. Birkhauser, Basel (1992)

24. Fletcher, P.T.: Statistical variability in nonlinear spaces: application to shape analysis and DT-MRI. Ph.D. thesis, Department of Computer Science, University of North Carolina (2004)
25. Fletcher, P.T.: Geodesic regression and the theory of least squares on Riemannian manifolds. *Int. J. Comput. Vis.* **105**(2), 171–185 (2013)
26. Fletcher, P.T., Joshi, S.: Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Process.* **87**(2), 250–262 (2007)
27. Frackowiak, R.S.J.: Human brain function. In: *Morphometry*, pp. 707–724. Academic Press, New York (2004)
28. Gallier, J.: Notes on differential geometry and Lie groups. Unpublished, accessed from <http://www.cis.upenn.edu/~jean/gbooks/manif.html>
29. Hall, B.C.: Lie groups, Lie algebras, and Representations: An Elementary Introduction. Springer, Berlin (2003)
30. Hernandez, M., Bossa, M., Olmos, S.: Registration of anatomical images using paths of diffeomorphisms parameterized with stationary vector field flows. *Int. J. Comput. Vis.* **85**(3), 291–306 (2009)
31. Hinkle, J., Muralidharan, P., Fletcher, P.T., Joshi, S.: Polynomial regression on Riemannian manifolds. arXiv preprint (2012)
32. Holm, D.D., Schmah, T., Stoica, C., Ellis, D.C.P.: Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions. Oxford Univ. Press, Oxford (2009)
33. Hua, X., Leow, A.D., Parikshak, N., Lee, S., Chiang, M., Toga, A.W., Jack, C.R., Weiner, M.W., Thompson, P.M.: ADNI: Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: an MRI study of 676 AD, MCI, and normal subjects. *NeuroImage* **43**(3), 458–469 (2008)
34. Joshi, S.C., Miller, M.I.: Landmark matching via large deformation diffeomorphisms. *IEEE Trans. Image Process.* **9**(8), 1357–1370 (2000)
35. Karacali, B., Davatzikos, C.: Estimating topology preserving and smooth displacement fields. *IEEE Trans. Med. Imaging* **23**(7), 868–880 (2004)
36. Lepore, N., Brun, C.A., Chiang, M., Chou, Y., Dutton, R.A., Hayashi, K.M., Lopez, O.L., Aizenstein, H.J., Toga, A.W., Becker, J.T., Thompson, P.M.: Multivariate statistics of the Jacobian matrices in tensor based morphometry and their application to HIV/AIDS. In: *Medical Image Computing and Computer Assisted Intervention*, pp. 191–198. Springer, Berlin (2006)
37. Lepore, N., Brun, C., Chou, Y.Y., Chiang, M.C., Dutton, R.A., Hayashi, K.M., Luders, E., Lopez, O.L., Aizenstein, H.J., Toga, A.W., Becker, J.T., Thompson, P.M.: Generalized tensor-based morphometry of HIV/AIDS using multivariate statistics on deformation tensors. *IEEE Trans. Med. Imaging* **27**(1), 129–141 (2008)
38. Li, C., Sheng, Y., Wang, M.: An effective method to compute Fréchet derivative of matrix exponential and its error analysis. *J. Inf. Comput. Sci.* **7**, 1854–1859 (2010)
39. Machado, L., Silva Leite, F., Krakowski, K.: Higher-order smoothing splines versus least squares problems on Riemannian manifolds. *J. Dyn. Control Syst.* **16**(1), 121–148 (2010)
40. Magnus, J.R., Neudecker, H.: Matrix Differential Calculus with Applications in Statistics and Econometrics. Wiley, New York (1988)
41. Mathias, R.: A chain rule for matrix functions and applications. *SIAM J. Matrix Anal. Appl.* **17**(3), 610–620 (1996)
42. Moakher, M.: A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.* **26**(3), 735–747 (2005)
43. Modersitzki, J.: Numerical methods for image registration. Oxford Univ. Press, Oxford (2004)
44. Modin, K., Perlmutter, M., Marsland, S., McLachlan, R.: On Euler–Arnold equations and totally geodesic subgroups. *J. Geom. Phys.* **61**(8), 1446–1461 (2011)
45. Moler, C., Van Loan, C.: Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Rev.* **45**, 3–49 (2003)
46. Morris, J.C.: The clinical dementia rating (CDR): current version and scoring rules. *Neurology* **43**(11), 2412–2414 (1993)
47. Muralidharan, P., Fletcher, P.T.: Sasaki metrics for analysis of longitudinal data on manifolds. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1027–1034 (2012)
48. Najfeld, I., Havel, T.F.: Derivatives of the matrix exponential and their computation. *Adv. Appl. Math.* **16**(3), 321–375 (1995)
49. Nichols, T.E., Holmes, A.P.: Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* **15**(1), 1–25 (2002)
50. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, Berlin (2006)
51. Pennec, X.: Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *J. Math. Imaging Vis.* **25**(1), 127–154 (2006)
52. Pennec, X.: Left-invariant Riemannian elasticity: a distance on shape diffeomorphisms? In: *MICCAI Workshop on Mathematical Foundations of Computational Anatomy*, pp. 1–13 (2006)
53. Pennec, X., Fillard, P., Ayache, N.: A Riemannian framework for tensor computing. *Int. J. Comput. Vis.* **66**(1), 41–66 (2006)
54. Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A.: Mutual-information-based registration of medical images: a survey. *IEEE Trans. Med. Imaging* **22**(8), 986–1004 (2003)
55. Ridgway, G.: Statistical analysis for longitudinal MR imaging of dementia. Ph.D. thesis, University College London (2009)
56. Rueckert, D., Aljabar, P., Heckemann, R.A., Hajnal, J.V., Hammers, A.: Diffeomorphic registration using B-splines. In: *Medical Image Computing and Computer Assisted Intervention*, pp. 702–709. Springer, Berlin (2006)
57. Škrinjar, O., Bistoquet, A., Tagare, H.: Symmetric and transitive registration of image sequences. *J. Biomed. Imaging* **208**, 14 (2008)
58. Sommer, S., Lauze, F., Hauberg, S., Nielsen, M.: Manifold valued statistics, exact principal geodesic analysis and the effect of linear approximations. In: *European Conference on Computer Vision*, pp. 43–56 (2010)
59. Studholme, C., Drapaca, C., Iordanova, B., Cardenas, V.: Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change. *IEEE Trans. Med. Imaging* **25**(5), 626–639 (2006)
60. Székely, G.J., Rizzo, M.L.: Testing for equal distributions in high dimension. *InterStat* (Nov) (2004)
61. Tapp, K.: Matrix Groups for Undergraduates. AMS, Providence (2005)
62. Thompson, P.M., Toga, A.W.: Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations. *Med. Image Anal.* **1**(4), 271–294 (1997)
63. Vaillant, M., Miller, M.I., Younes, L., Trounev, A.: Statistics on diffeomorphisms via tangent space representations. *NeuroImage* **23**(Suppl. 1), S161–S169 (2004)
64. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Symmetric log-domain diffeomorphic registration: a demons-based approach. In: *Medical Image Computing and Computer Assisted Intervention*, pp. 754–761. Springer, Berlin (2008)
65. Whitcher, B., Wisco, J.J., Hadjikhani, N., Tuch, D.S.: Statistical group comparison of diffusion tensors via multivariate hypothesis testing. *Magn. Reson. Med.* **57**(6), 1065–1074 (2007)
66. Younes, L., Arrate, F., Miller, M.I.: Evolutions equations in computational anatomy. *NeuroImage* **45**(1), S40–S50 (2009)



**Ernesto Zacur** received the M.Sc. degree in physics from the Instituto Balseiro, Bariloche, Argentina, in 2002. In 2011, he obtained the Ph.D. degree in biomedical engineering at the University of Zaragoza. He is currently at the Communications Technologies Group (GTC) of the Aragon Institute of Engineering Research (I3A) at the University of Zaragoza. His current research interests include morphometry and medical image analysis.



**Matias Bossa** received the M.Sc. degree in physics from the Instituto Balseiro, Argentina, in 2003. He is currently working toward the Ph.D. degree in biomedical engineering at the Aragon Institute for Engineering Research (I3A), University of Zaragoza, Zaragoza, Spain. His current research interests include manifold valued statistical methods, mathematical tools for image analysis, specially those related to large deformation, image registration, and their applications to the analysis of medical images.



**Salvador Olmos** obtained Industrial Engineering degree and the corresponding Ph.D. at Universitat Politècnica de Catalunya (Barcelona, Spain) in 1994 and 1998 respectively. Currently he is Associate Professor at Universidad de Zaragoza (Spain) in the Electronics Engineering and Communications Department. As a researcher he is member of the Aragon Institute of Engineering Research. From Aug 2000 to Aug 2001 he visited Lund University (Sweden) as a post-doc researcher funded by Spanish Ministry of Education (Spain). His research experience initially focused in the field signal processing of biomedical signals. Since 2005 he is involved in medical image analysis, including non-rigid registration, computational anatomy, brain morphometry.