Topologically constrained template estimation via Morse-Smale complexes controls its statistical consistency

Nina Miolane, Susan Holmes, Xavier Pennec

Topologically constrained template estimation via Morse-Smale complexes controls its statistical consistency

*****

Abstract.
In most neuroimaging studies, one builds a brain template that serves as a reference anatomy for normalizing the measurements of each individual subject into a common space. Such a template should be representative of the population under study, in order to avoid biasing the subsequent statistical analyses. The template is often computed by iteratively registering all images to the current template and then averaging the intensities of the registered images. Geometrically, the procedure can be summarized as follows: it is the computation of the template as the "Fréchet mean" of the images projected in a quotient space. It has been argued recently that this type of algorithm could actually be asymptotically biased, therefore inconsistent. In other words, even with an infinite number of brain images in the database, the template estimate may not converge to the brain anatomy it is meant to estimate. Our paper investigates this phenomenon. We present a methodology that quantifies spatially the brain template's asymptotic bias. We identify the main variables controlling the inconsistency. This leads us to investigate the topology of the template's intensity levels sets, represented by its Morse-Smale complex. We propose a topologically constrained adaptation of the template computation, that constructs a hierarchical template with bounded bias. We apply our method to the analysis of a brain template of 136 T1 weighted MR images from the Open Access Series of Imaging Studies (OASIS) database.

Introduction. In neuroimaging, as well as in many other medical image analysis domains, a *template* is an image representing a reference anatomy. A template is computed from a database of brain images to serve as the brain image "prototype" for further analyses.

Computation of a brain template. Various methods exist to compute a brain template [15]. A first practice selects one brain image from the database as the template. If the selected subject's anatomy is far from the population mean anatomy, the template is necessarily biased towards this specific individual. Thus, the template fails at being a prototype of the population. This is why researchers consider the computation of an "unbiased template" that represents the mean anatomy better.

Such an "unbiased" template is often constructed by performing an iterative averaging of intensities and deformations [17, 24, 20]. One initializes with a template chosen among the subject images. Then, during each iteration, one registers the subjects to the current template, and computes the mean deformation. The new template is computed as the mean intensity of the subjects' images, deformed with the mean deformation. This procedure does not favor any subject's image if it does not end in a local minimum. In this sense, the procedure is called "unbiased".

The computed brain template may look blurred or sharp depending on the design chosen for the registration in the above iterative procedure. If the algorithm is designed using linear registration, the template may appear blurred. In contrast, if one uses diffeomorphic registration, the template is more likely to look sharp and the sharpness depends on the amount of regularization used [15].

Purpose and desirable properties of the brain template. Computing a template is often the first step in medical image processing because of its many applications. The template is used as a standardized 3D coordinate frame where the subject brains can be compared. The subjects

are then characterized by their *spatial diffeomorphic deformations from the template*. These deformations may serve for a statistical analysis of the subject shapes [6]. One studies the normal and pathological variations of the subjects with respect to the template. The deformations also facilitate automated segmentation, by mapping the template's already segmented regions into each subject space.

What are the desirable properties of the brain template, with respect to the applications mentioned above? First, the template should be representative of the population, removing any bias toward a specific subject during the analysis [10, 5, 7]. Second, the template should be sharply defined, so that subtle anatomical structures can be easily observed or segmented.

The brain template, an inconsistent estimator of the unique brain anatomy. However, the two desirable properties cannot be fulfilled simultaneously: there exists a trade-off, akin to the standard bias-variance trade-off in statistical learning.

First, this trade-off can be understood intuitively. Consider a database of brain images divided into two groups that have different topologies. The first group has subjects with three sulci - i.e. depressions or grooves in the cerebral cortex - in a specified brain region. The second group has subjects with only two sulci in the same region. A sharply defined template has to decide on a specific topology in this brain region, i.e. whether it shows two or three sulci. Therefore, it might not estimate correctly the brain anatomy of this population, which might be problematic for the following applications in neuroimaging. For example during a statistical analysis, registering the subjects with the three sulci topology to a template that has chosen a two sulci topology might not be reasonable [7]. A sharp template is only meaningful if the anatomical structures shown are representative of the whole population, i.e. if the population is homegeneous.

Second, the trade-off has been emphasized in recent studies investigating the bias of the template as an estimator of the population's shared anatomy. In the classical approach [2], an initial assumption states that there is a unique (brain) anatomy shared by the population. The subjects are then modeled through a generative model as random deformations of the unique brain anatomy, observed with additional noise. The unique brain anatomy is a parameter of this model. The template computation is interpreted as its estimation. One can ask about its asymptotic bias: does the template converge to the unique brain anatomy for a database with an infinite number of images?

This question has been investigated for signals, i.e. 1D images. Some authors prove the asymptotic unbiasedness of the template under the simplifying assumption of no measurement error on the observed signals [27]. Other authors have already provided examples of asymptotic bias, and therefore inconsistency, when there is measurement error [2]. Their experiments show that the template may converge to pure noise when the measurement error on simulated signals increases. A bias is shown to occur in [8] for curves estimated from a finite number of points in the presence of noise.

Recently, an asymptotic bias has been shown in the setting of Lie group actions [35, 34]. Our argument, shown in an abstract geometric context in [35, 34] but adapted here to brain images, is as follows. We look at the subspace defined by all brains with the same shape as the unique brain anatomy. We show that the curvature of this space, at the scale of the measurement noise, introduces a bias on the brain template.

Using topology to investigate the brain template's asymptotic bias. We want to link: (i) the fact that a population with two groups of different brain topologies cannot be accurately represented by a sharp template, with (ii) the mathematical results on the template's bias as an estimate of the anatomy shared by the population. The framework of [35] is based on the quotient of the space of the data space by the action of a Lie group. The data in our case are the brain images, and the Lie group action is the action of diffeomorphisms on these images. Quotienting the images by the action of diffeomorphisms amounts to filtering out any information that is invariant by diffeomorphic deformations. Thus, the quotient gives the topology of the images' level sets. Using the intuition of [35], we could quantify the brain template's asymptotic bias *using a representation of its topology.*

Quantifying the bias could enable us to decide when and where a sharply defined template makes sense. We could want a sharp template in brain regions where the intersubject anatomical variability is low and a fuzzier template when this variability is higher. Alternatively, we could consider computing several sharp brain templates using mixtures. This discussion boils down to the question: when is it reasonable to assume that a unique brain anatomy represents the whole subject population?

Furthermore, we could think about controlling the brain template's asymptotic bias by constraining its topology. Topological representations of images have been used with various objectives in the literature. For example, [12] uses a topological representation of brain images for classification of autism versus normals. Topological constraints have also been implemented for segmentation where the reconstruction of the cortical surface needs to match the brain anatomy [30, 31, 21]. However, topological representation of images or topological constraints on images have not been used to study and enforce a statistical property, like asymptotic unbiasedness.

Contributions and Outline. We use a topological representation of images - the Morse-Smale complex - to investigate and control the asymptotic bias of the brain template. We make three main contributions in this paper. First, we show how to combine geometry and topology to tackle a statistical problem in neuroimaging. We provide conjectures at the boundaries of the fields with sketches of their proofs. Second, we analyze the template as an estimator of the brain anatomy and quantify the asymptotic bias. This leads us to discuss the initial assumption of a unique anatomy. Third, we present an adaptation of the template computation algorithm that bounds the bias, through topological constraints, at the price of constructing a "smoother" template.

Section 1 presents the geometry and the topology of the template computation. We emphasize the variables that describe the bias of the brain template. Section 2 presents the chosen computational representation of these variables through Morse-Smale complices. Section 3 leverages the previous computational model to spatially identify the biased regions of the template. We thus propose an adaptation of the template computation with topological constraints bounding the bias. In Section 4 our methodology is used on the Open Access Series of Imaging Studies (OASIS) database of T1-weighted MR brain images.

1. Geometry and topology for template estimation .

This paper will not present mathematical results, rather we show how geometry and topology combine to formalize the template computation algorithm and highlight required directions for further mathematical develop-

ments.

1.1. Geometrization of the action of diffeomorphisms on images.

Brain images. We consider two- and three-dimensional images, whose domain $\Omega \subset \mathbb{R}^d$, with $d = 2, 3$, is supposed compact. We adopt the point of view of images as square-integrable functions I over the compact domain Ω, i.e. we write $I \in L_2(\Omega)$, where $L_2(\Omega)$ is a Hilbert space. The corresponding L_2 distance is invariant by volume preserving diffeomorphisms (volumorphisms). Additionally, we assume that the images are in $C^\infty(\Omega)$, which is C^∞ defined on the compact support Ω. We denote $\mathrm{Img}(\Omega)$ the set of images.

To illustrate the following concepts, we use the toy Hilbert space $\mathbb{R}^2$ where one point schematically represents one image, see Figure 1.

Diffeomorphisms. A *diffeomorphism of* Ω is a differentiable map $\phi : \Omega \to \Omega$ which is a bijection whose inverse $\phi^{(-1)}$ is also differentiable. We consider two sets of diffeomorphisms.

On the one hand, we consider $C^\infty(\Omega)$, i.e. the smooth diffeomorphisms that are the identity outside a compact support. $C^\infty(\Omega)$ can be seen as an infinite dimensional manifold [33] and forms an infinite dimensional Lie group [26]. Its Lie algebra V is the set of smooth vector fields with compact support [26]. We use this set of diffeomorphisms to present algebraic concepts.

On the other hand, we consider the set $C_{\mathrm{Id}}(\mathbb{R}^d, \mathbb{R}^d)$ defined as $C_{\mathrm{Id}}(\mathbb{R}^d, \mathbb{R}^d) = \{\phi = \mathrm{Id} + u \quad \text{for } u \in C^1_b(\mathbb{R}^d, \mathbb{R}^d)\}$, where the subscript "b" refers to functions that are bounded with bounded derivatives. These diffeomorphisms are "small", i.e. not too different from the identity. We use this set to formulate mathematical conjectures which need metric properties. If the specification between the two sets is not needed, we will refer to $\mathrm{Diff}(\Omega)$ to denote diffeomorphisms.

Action of diffeomorphisms on (brain) images. The Lie group of diffeomorphisms $\mathrm{Diff}(\Omega)$ acts on the space of images $L_2(\Omega)$ [39]: $\rho : \mathrm{Diff}(\Omega) \times \mathrm{Img}(\Omega) \to \mathrm{Img}(\Omega), \quad (\phi, I) \mapsto \phi \cdot I = I \circ \phi^{-1}$. This action is represented on Figure 1 (a), which shows an image I and its diffeomorphic deformation.

Intuition and schematic representation on $\mathbb{R}^2$*.* The statistical analysis of this paper relies on geometric considerations in the Hilbert space of images $L_2(\Omega)$ endowed with the action of an Lie group of diffeomorphisms $\mathrm{Diff}(\Omega)$. The intuition on the geometry in these abstract spaces will be given by a serie of figures, where:

$L_2(\Omega)$: Hilbert space of images is represented as $\mathbb{R}^2$: 2D surface of the paper,
$\mathrm{Diff}(\Omega)$: Lie group of diffeomorphisms is represented as $SO(2)$: Lie group of 2D rotations.

and we consider the fundamental action of $SO(2)$ on $\mathbb{R}^2$.

Figure 1(b) shows this representation. The action of a diffeomorphism ϕ on I is represented by the blue curved arrow, i.e. by the action of a 2D rotation. The action transforms the image I into another image $\phi \cdot I$, i.e. into a different point in the Hilbert space.

We note that $L_2(\Omega)$ with the action of $\mathrm{Diff}(\Omega)$ and $\mathbb{R}^2$ with the action of $SO(2)$ have different properties. For example, $\mathbb{R}^2$ and $SO(2)$ are finite dimensional and the action of $SO(2)$ is isometric with respect to the Euclidean distance on $\mathbb{R}^2$. In comparison, $L_2(\Omega)$ and $\mathrm{Diff}(\Omega)$ are infinite dimensional and the action of $\mathrm{Diff}(\Omega)$ is not isometric with respect to the

L_2 distance between images of $L_2(\Omega)$.

Nevertheless, we argue that our schematic representation makes sense for the following reasons. First, representing infinite dimensional spaces by finite dimensional ones is standard as it is difficult to draw infinite dimensions on a piece of paper. Second, we consider diffeomorphisms that transform a *brain* image into another *brain* image, i.e. an image into one that looks similar. Therefore, we restrict ourselves to "small" diffeomorphisms. In this context, we will see that we can consider their action as being isometric.

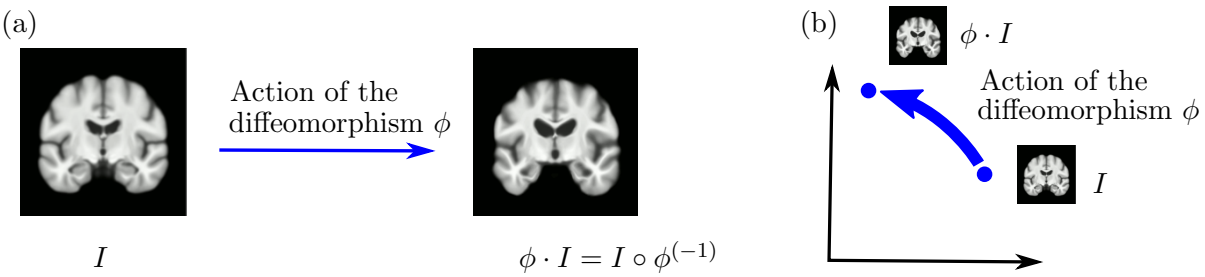

Figure 1. *Action of a diffeomorphism ϕ on a brain image I. (a) the brain image before and after the action of ϕ, (b) schematic representation of the action of ϕ on the brain image I, represented as a dot in $\mathbb{R}^2$.*

Orbit O_I of a (brain) image I. Here, we consider $\mathrm{Diff}(\Omega) = C^\infty(\Omega)$. The *orbit* O_I of a brain image I is defined as all images reachable through the action of diffeomorphisms on I: $O_I = \{I' \in \mathrm{Img}(\Omega) | \exists \phi \in \mathrm{Diff}(\Omega) \text{ s.t. } I' \circ \phi^{-1} = I\}$.

On Figure 2 (a), images I_1 and I_2 belong to the same orbit but I_3 belongs to a different orbit. We use $\mathbb{R}^2$ representing the space of images, with the action of $SO(2)$ representing the action of diffeomorphisms on Figure 2(b). The blue dotted circle on Figure 2(b) represents O the orbit of I_1 and I_2. This orbit defines a submanifold of images: in this toy illustration, the submanifold is the blue dotted circle. The red point on Figure 2(b) represents O_{I_3}, the orbit of the image I_3. This orbit contains only one point and is a submanifold of dimension 0.

We note that the orbits may be infinite dimensional in the case of diffeomorphisms acting on $L_2(\Omega)$. Furthermore the orbits are not necessarily high-dimensional spheres as the action of the diffeomorphisms is not isometric. However, by considering only "small" diffeomorphisms acting on a given image I, we move locally on the orbit of image I. We could consider writing a Taylor expansion of the orbit around I [34], where the first order gives its tangent space and the second order is a high-dimensional sphere. Therefore, "small" diffeomorphisms are consistent with a representation of the images' orbits as spheres.

Isotropy group G_I of a (brain) image I. Here, we consider $\mathrm{Diff}(\Omega) = C^\infty(\Omega)$. The *isotropy group* G_I of a brain image I is defined as the subgroup of $\mathrm{Diff}(\Omega)$ formed by the diffeomorphisms that leave I unchanged: $G_I = \{\phi \in \mathrm{Diff}(\Omega) | I \circ \phi^{-1} = I\}$. G_I describes the intrinsic symmetry of the image I: the more symmetric is I, the larger its isotropy group. All images on the same orbit have conjugate isotropy groups. Moreover, the isotropy group (also called the stabilizer) and the orbit of an image are linked by the orbit-stabilizer theorem: $O_I \sim \mathrm{Diff}(\Omega)/G_I$ in finite dimensions. The intuition is that the larger the isotropy group (and thus, the more symmetry the image has), the smaller the orbit. Figure 2(a) shows two brain

images: the isotropy group of I_1 and I_2 is larger than the isotropy group of I_3, in the sense of inclusion. As a consequence, the orbit of I_3 is "smaller" than the orbit of I_2.

We use again the representation of $\mathbb{R}^2$ to illustrate this notion. The notion of "isotropy group" dictates the dimension of a given orbit, i.e. whether we have a 1-dimensional submanifold like the blue dotted circle or a 0-dimensional submanifold like the red point on Figure 2 (b). The isotropy group of the image at the blue point on Figure 2(b) is the identity. Only the identity leaves this image at the same place. The isotropy group of the image represented by the red point is the whole Lie group of 2D rotations. Any rotation leaves this point invariant.

Going back to diffeomorphisms, we note that the isotropy group may be of infinite dimension: the isotropy group of a uniform image, i.e. I constant map over Ω, is the whole group $\mathrm{Diff}(\Omega)$. Nevertheless, the orbit-stabilizer theorem holds in infinite dimensions therefore we can rely on the intuition of "smaller" isotropy groups and "larger" orbits.

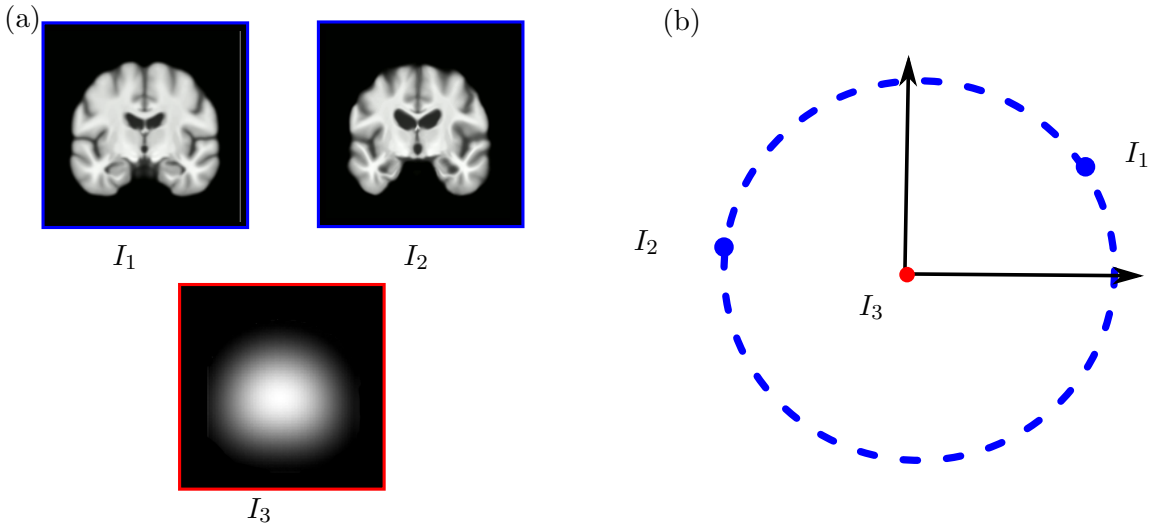

Figure 2. *Orbit and isotropy group of a brain image. (a) I_1 and I_2 belong to the same orbit: I_2 is a diffeomorphic deformations of I_1. They have conjugate isotropy groups. In contrast, I_3 belong to a different orbit and has a different isotropy group. I_3 shows more symmetry than I_1 and I_2, thus a larger isotropy group, whereas the more asymmetric details of I_1, I_2 are the sign of a smaller isotropy group. (b) The orbit of I_1, I_2 is represented as the blue dotted circle and the two images I_1, I_2 are points on this circle. The isotropy group is linked to the dimension of the image's orbit. I_1, I_2 have a smaller isotropy group, they have a circle orbit. I_3 has a larger isotropy group: its orbit is itself, i.e. the red point at $(0, 0)$.*

1.2. From geometry to topology.

Topology of (brain) images. The *topology of a brain image* I is defined as the topology of its level sets, these surfaces of Ω with constant intensity. topology refers to properties that are preserved under smooth deformations [16], i.e. conserved by the action of diffeomorphisms on I: for example, the number of holes, or the number of connected parts, see Figure 3(a).

Geometry and topology combine as follows. Two images I and I' that are diffeomorphic deformations of each other, i.e. that are on the same orbit, have the same topology. The orbit O_I itself represents the topology of image I (and I'). The set of orbits $Q = \{O_I | I \in \mathrm{Img}(\Omega)\}$, which is the *quotient space of Img*(Ω) *by the action of Diff*(Ω) is the set of the topologies.

Figure 3(b) shows how the space of images $\mathbb{R}^2$ is partitioned into orbits: blue circles and one red "singular circle", the red point at $(0,0)$. Figure 3(b) also shows $\mathbb{R}_+$, the quotient space of $\mathbb{R}^2$ by the group of 2D rotations, which schematically represents the quotient space of the space of brain images $\mathrm{Img}(\Omega)$ by the Lie group of diffeomorphisms $\mathrm{Diff}(\Omega)$. Each of the four blue circles in $\mathbb{R}^2$ becomes a blue point in the quotient space $\mathbb{R}_+$.

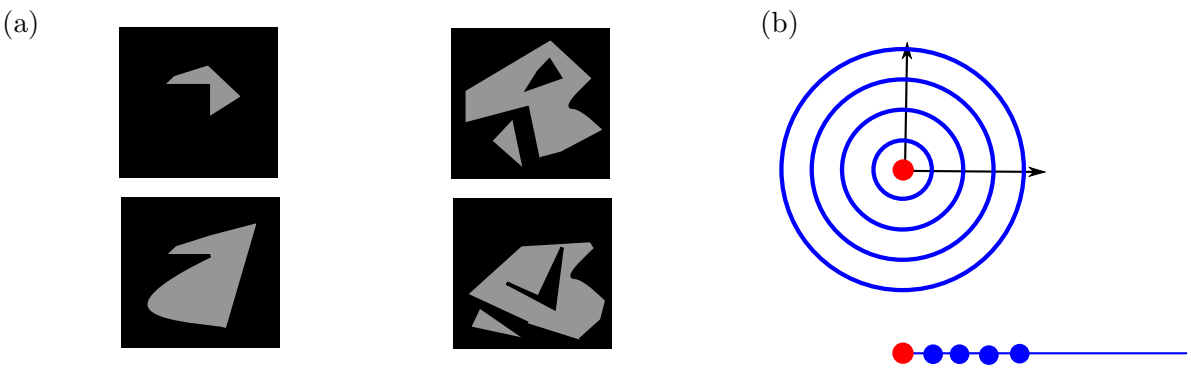

Figure 3. *(a) Images in the same colomn have same topology, images in the line have different topologies: one cannot be diffeormorphically deformed to match the other. (b) Top: schematic representation of the space of images partitioned into orbits. The two different orbit types are in blue and red respectively. Bottom: schematic representation of the quotient space of brain images Img*(Ω) *by the Lie group of diffeomorphisms Diff*(Ω).

Gathering brain images with similar topologies: orbit types. By definition, two brain images are of the same *orbit type* if their isotropy groups are conjugate subgroups in the Lie group of diffeomorphisms. In particular, brain images that belong to the same orbit have same orbit type. The type corresponding to the smallest isotropy group - in the sense of the inclusion among the subgroups of the diffeomorphism group - is sometimes called the principal type [1]. We use this appellation in this paper. For example, if an orbit has type the identity of the group of diffeomorphisms, then it is necessarily of principal type. Equivalently, the orbits of principal type are called *principal orbits*. Other orbits are called *singular orbits*.

The blue circles on Figure 3(b) have same orbit type: the images on these orbits have the identity $\{Id\}$ of the Lie group as isotropy group. They are of principal type since the identity $\{Id\}$ is the smallest subgroup - in the sense of the inclusion - of the group of rotations.

Stratification of the space of topologies. In the space of brain images, we can gather orbits of same orbit type: we gather the blue circles of Figure 3(b) into $\mathbb{R}^2 \setminus \{(0,0)\}$ on the one hand, and keep the red dot $(0,0)$ on the other hand. The orbit type itself is a submanifold of the space of brain images: $\mathbb{R}^2 \setminus \{(0,0)\}$ or $(0,0)$ in the schematic brain images space $\mathbb{R}^2$. Furthermore, these orbit type submanifolds form a *stratification*, meaning they fit together in a particularly nice way [38].

The quotient space Q is also naturally partitioned into manifolds, and this partitioning is also a stratification. All in all, Q is *not* a manifold, but Q composed of manifold pieces, and those pieces are called strata. There is a partial ordering of the strata in the quotient space, using the inclusion [22].

Figure 3(b) uses the analogy of $\mathbb{R}^2$ as the space of images and $SO(2)$ as the Lie group acting on it. It shows the orbits grouped by orbit type: the color blue denotes one orbit type

and the color red another orbit type. We see for example that $Q = \mathbb{R}_+$ is stratified into one stratum being $\mathbb{R}_+^*$ - corresponding to the stratum $\mathbb{R}^2 \setminus \{(0,0)\}$ in the space of brain images - and one stratum being $\{0\}$ - corresponding to the stratum $(0,0)$ in the space of brain images.

1.3. Geometry of generative model and estimation procedure. In this subsection, we consider: $\text{Diff}(\Omega) = C_b^1(\mathbb{R}^d, \mathbb{R}^d)$.

Generative model. The n brain images $I_1, ..., I_n$ are interpreted with a generative deformable model: $I_i = \phi_i \cdot T + \epsilon_i, i = 1...n$, where each image $I_i \in \text{Img}(\Omega)$ is a diffeomorphic deformation $\phi_i \in \text{Diff}(\Omega)$ of a unique brain anatomy T, to which noise ϵ_i is added. The parameter T represents the brain anatomy shared by the population. The transformations ϕ_i's and the noises ϵ_i's are i.i.d. realizations of random variables. The transformations ϕ_i's follow a general law, which could be for example a Gaussian law in a finite-dimensional subspace of the Lie group and the ϵ_i's represent Gaussian noise on the space of images. We denote σ^2 its variance. Definitions of distributions on finite dimensional Riemannian manifolds are taken from [36] and the Gaussian distributions for infinite dimensional spaces from [28].

The model can be interpreted by a three step generative procedure illustrated schematically in Figure 4. First, there is only the shared anatomy T. Second, the template T is deformed with the diffeomorphism ϕ_i and gives a brain image $\phi_i \cdot T$. Third, we represent the measurement noise with ϵ_i, which gives the observed brain image I_i.

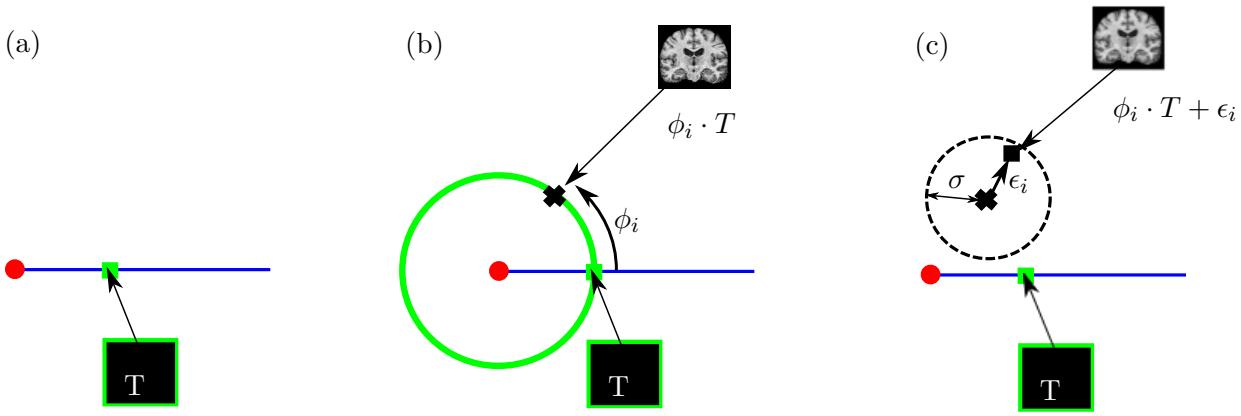

Figure 4. *Schematic illustration of the generative model of the brain images data. As before, the space of brain images is represented by the plane $\mathbb{R}^2$. (a) First step of the generative model: generate a brain anatomy. One usually assumes that there is a unique brain anatomy: T, in green. (b) Second step of the generative model: generate a deformation $\phi_i \in Diff(\Omega)$ which is used to deform the template. The brain image $\phi_i \cdot I$ belongs to the orbit of T, represented by the green circle. (c) Third step of the generative model: generate noise ϵ_i in the space of images. The brain image $\phi_i \cdot T + \epsilon_i$ does not belong to the orbit of T anymore.*

Computing the template: an estimation procedure. Computing the brain template amounts to invert the generative model: given the data, we want to estimate the parameter T. The transformations ϕ's are hidden variables of the model. The natural statistical procedure to estimate T in this context is the Expectation-Maximization (EM) algorithm [13]. The EM is an iterative procedure that maximizes the log-likelihood of the generative model with hidden variables. As such, the EM gives an asymptotically unbiased and consistent estimation of the

brain anatomy T.

In practice, one does not use the EM algorithm because it is computationally expensive, especially when dealing with tridimensional images. Most neuroimaging pipelines rely on *an approximation of* the Expectation-Maximization algorithm to estimate the brain anatomy, called "Fast Approximation with Modes" in [2]. It runs as follows. Initialize the estimate with $\hat{T} = I_1$, i.e. one of the brain images from the database. Then, iterate the following two steps until convergence [24]:

$$(1) \quad \hat{\phi}_i = \underset{\phi \in \mathrm{Diff}(\Omega)}{\operatorname{argmin}} \, d_{\mathrm{Img}(\Omega)}(\hat{T}, \phi \cdot I_i) + \lambda \mathrm{Reg}(\phi), \quad \forall i \in \{1, ..., n\},$$

$$(2) \quad \hat{T} = \underset{T \in \mathrm{Img}(\Omega)}{\operatorname{argmin}} \sum_{i=1}^{n} d_{\mathrm{Img}(\Omega)}(T, \hat{\phi}_i \cdot I_i)^2.$$

Step (1) is an estimation $\hat{\phi}_i$ of the diffeomorphisms ϕ_i, and an approximation of the E-step of the EM algorithm. In practice, each brain image I_i is registered to the current template estimate and the $\hat{\phi}_i$ is the result of this registration. The term Reg is a regularization that ensures that the optimization has a solution. Showing that the estimates $\hat{\phi}_i$'s and $\hat{T}$ exist is technical and beyond the scope of this paper.

Step (2) is the M-step of the EM algorithm: the maximization of the surrogate in the M-step amounts to the maximization of the variance of the projected data. This computes the updated template estimate, as the mean intensity of the subjects images I_i, deformed with the mean deformation of the $\hat{\phi}_i$'s.

The registration step (1) amounts to aligning the n subject images by transporting them onto their orbit (see Figure 5(b)), i.e. projecting them in the quotient space (see Figure 5 (c)). Step (2) averages the n registered images (see Figure 5 (d)).

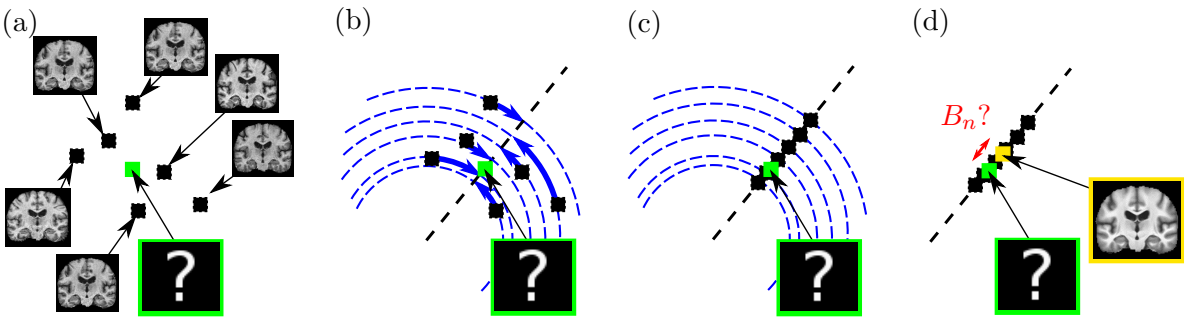

Figure 5. *Geometrization of an iteration of the template's computation: (a) n subjects images (black squares); (b) the n images are registered, they travel on their orbit (the blue circles) to get aligned; (c) registered images; (d) the empirical brain template $\hat{T}$ (in yellow) is computed as the Fréchet mean of the n registered images. How far is it from the unique anatomy T of the generative model (in green): can we quantify B_n?*

Evaluation of the procedure: definition of asymptotic bias B_∞. We evaluate the template $\hat{T}$ as an estimator of the unique brain anatomy T (see Figure 5 (d)) given n observations I_i, $i = 1...n$. We note that in other papers, T may be called the template directly and $\hat{T}$ the template's estimate.

We consider two measures of the accuracy of this estimator: its variance V_n^2 and bias B_n, which are defined as:

$$V_n^2 = \mathbb{E}((\hat{T} - \mathbb{E}(\hat{T}))^2) \qquad \text{and} \qquad B_n = \mathbb{E}(\hat{T} - T).$$

The bias for n images B_n is illustrated on Figure 5. We are interested in the asymptotic behavior $n \to +\infty$, i.e. when the number of brain images goes to infinity. One would expect that the procedure converges to the brain anatomy T it is designed to estimate. When the estimator converges, its variance is asymptotically zero: $V_\infty^2 = 0$.

1.4. Geometry of the template estimator's evaluation.

Estimation procedure interpreted as the Fréchet mean in the quotient space. We consider the estimation procedure of Equation 1.3. First, we note that the regularization term in step (1) forces the diffeomorphisms to be small, i.e. to be close to the identity and close to have an isometric action on the images. Second, we could consider the group of volumorphisms, the subgroup of diffeomorphisms that leave a volume form V invariant and is used to model incompressible fluids in continuum mechanics. In this case, using $y = \phi^{-1}(x)$ and the fact that the volume form is conserved:

$$(1)\quad d(\phi\cdot I_1, \phi\cdot I_2)^2 = \int_\Omega (I_1\circ\phi^{-1}(x) - I_2\circ\phi^{-1}(x))^2 dV(x) = \int_\Omega (I_1(y) - I_2(y))^2 dV(y) = d(I_1, I_2)^2$$

and the action is thus isometric.

In both of the above two frameworks, it is thus reasonable to trade the regularization term for the assumption that the action is isometric in our modelization.

First, we model the estimation procedure as follows.

$$(1)\quad \hat{\phi}_i = \underset{\phi\in G}{\operatorname{argmin}}\ d_M(\hat{T}, \phi\cdot I_i), \quad \forall i \in \{1, ..., n\},$$

$$(2)\quad \hat{T} = \underset{T\in M}{\operatorname{argmin}} \sum_{i=1}^{n} d_M(T, \hat{\phi}_i \cdot I_i)^2.$$

where M is a generic Riemannian manifold and G a Lie group acting on M isometrically.

This converges to a local minimum because it decreases at each step a cost bounded below by zero. The estimator computed with this procedure is:

$$(2)\qquad \hat{T} = \underset{T\in M}{\operatorname{argmin}} \sum_{i=1}^{n} \min_{\phi\in G} d_M^2(T, \phi\cdot I_i).$$

The term $\min_{\phi\in G} d_M^2(T, \phi\cdot I_i)$ is the distance in the quotient space between T and I_i. Thus Equation 2 defines the Fréchet mean on the quotient space [36].

The study of the set of solutions of algorithms in Equation 1.3 and Equation 1.4 is an interesting direction of research but beyond the scope of this paper. However, we point out that the existence and uniqueness of the solution of algorithm 1.4 is linked to the question of whether the Fréchet mean exists and is unique in the quotient space, which is studied for example in [25, 4].

Asymptotic bias B_∞ and curvature. We show in [35] that the asymptotic bias is non zero: $B_\infty \neq 0$. For an infinite number of brain images $n \to +\infty$, the estimate converges, but not to the brain anatomy T it was designed for. We compute in [35] a Taylor expansion of the asymptotic bias B_∞ around the noise $\sigma = 0$, in the case of a finite dimensional manifold and isometric Lie group action [34]:

$$B_\infty = \frac{\sigma^2}{2} H(T) + \mathcal{O}(\sigma^3) + \epsilon(\sigma) \tag{3}$$

where $H(T)$ denotes the mean curvature vector of the template's orbit. There is no bias when there is no measurement error $\sigma = 0$. It was observed experimentally that the bias was dependent on the measurement error [2].

The coefficient $H(T)$ depends on the template T that is being estimated. We investigate this dependency in the geometric framework of Subsection 1.2. Assume that there exists a fixed point o of the Lie group action, i.e. a point that is invariant by the whole Lie group. Consider the orbit O_T of T. As the action is isometric, the orbit belongs to a geodesic sphere S_d with center o and radius d. A geodesic sphere of radius d in a manifold - like a hypersphere of radius d in $\mathbb{R}^m$ - has mean curvature vector of norm: $||H(T)|| = \frac{(m-1)}{d}$. If we write the template's bias in the units of d, then the asymptotic bias depends on $\left(\frac{\sigma}{d}\right)^2$. In other words, the distance of the template to the singularity o, at the scale of the noise σ governs the asymptotic bias B_∞. The approach of this paper is to compute the asymptotic bias by estimating the geometric parameters. Conversely, one could compute the geometric parameters, like the external curvature of orbits for different Lie groups, from the asymptotic bias known on simulated examples.

Figure 6 shows the intuition behind this. On Figure 6(a), $\mathbb{R}^2$ schematically represents the space of brain images - the black squares represent brain images from the database, the green square is T - and the green circle is the orbit of T. The dotted circles, that have their centers on the template's orbit, represent the probability distribution of the (2D isotropic) Gaussian noise in the generative model. More precisely, they represent the level set at σ of the noise distribution. The curvature $H(T)$ controls the area in grey on Figure 6, which is the area inside the Gaussian level set that is outside T's orbit. This area is greater that the area inside T's orbit. As a consequence, the probability that the brain images are generated "outside" T's orbit is higher than the probability that they are generated inside T's orbit.

Figure 6(b) shows the registration step of the template estimation: there is a higher probability that the registered images are away from T, as if repulsed from the singularity around which the orbits warp. When one averages the registered images, one sees that the template's estimate becomes biased as it will systematically give an image that is further away than T from the quotient space's singularity, i.e. from the red dot.

Quantifying the asymptotic bias B_∞ of the brain template. In neuroimaging, the manifold is the space of brain images $\mathrm{Img}(\Omega)$ and the Lie group is the group of diffeomorphisms $\mathrm{Diff}(\Omega)$, both infinite dimensional. This paper assumes that we can apply the geometry of [34], because there are indications that B_∞ appears in the same fashion in infinite dimension, see Figure 7.

First, [34] studies the bias when the dimension of the manifold increases. They consider the finite dimensional manifold $M = \mathbb{R}^m$ with the action of $SO(m)$, i.e. a generalization of

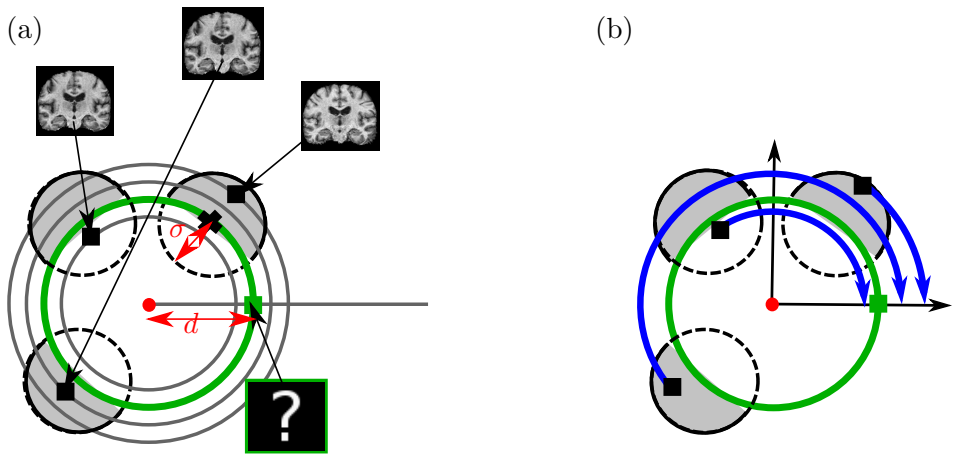

Figure 6. *Schematic illustration of the asymptotic bias in the template computation algorithm of neuroimaging. (a) The images generated with the model described above have a higher probability to be "outside" (with respect to the curvature) of the template's orbit. (b) The registration during template's estimation aligns the images. The distribution of images is unbalanced with respect to the real template.*

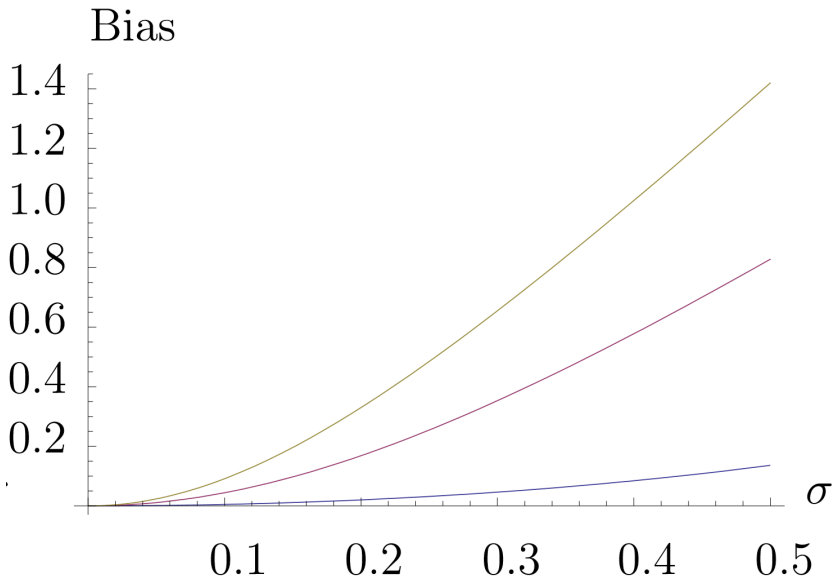

Figure 7. *Image from [34]. Take $M = \mathbb{R}^m$ with the action of $SO(m)$. The template's bias increases with σ and is more important as m increases: the blue curve shows the asymptotic bias for $m = 2$, the pink curve for $m = 10$ and the yellow curve for $m = 20$.*

the toy example $\mathbb{R}^2$ with the action of $SO(2)$ from our illustrations. We show in [34] that B_∞ increases when m increases, see Figure 7.

Then, the work of [3] shows that there exists an asymptotic bias in an infinite dimensional Hilbert space. The work of [14] provided an asymptotic behavior of the bias when the noise level σ tends to infinity. This bias is exemplified on examples of template of signals in [14],

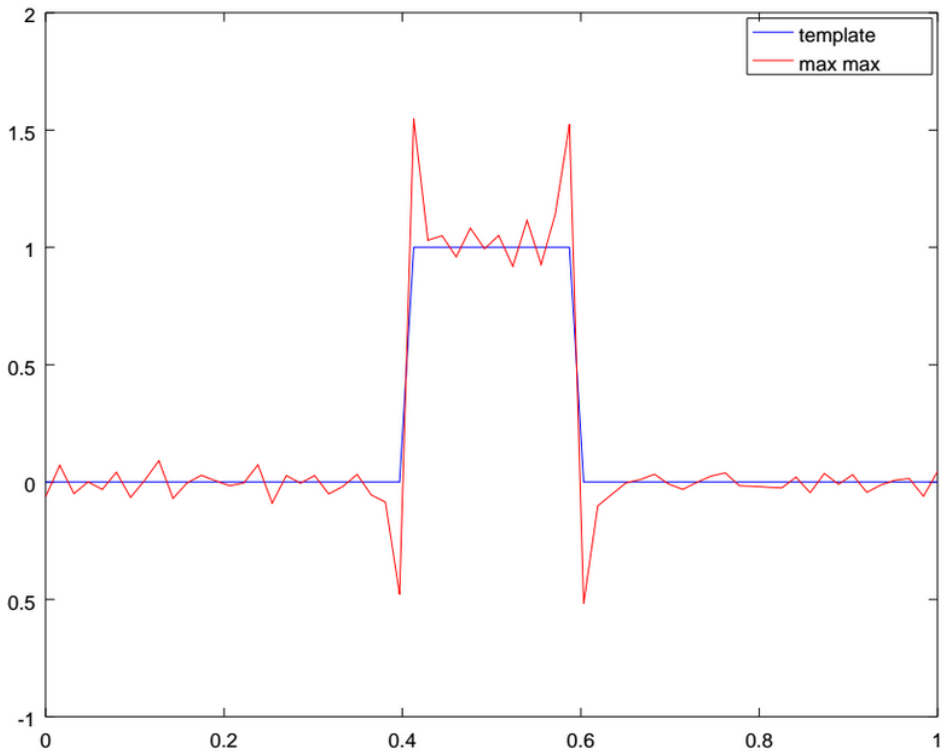

Figure 8. *Image from [14]. The real signal is shown in blue. The template, computed with $n = 10^5$ observations simulated with Gaussian noise with $\sigma = 10$, is shown in red. There is an asymptotic bias in the estimation of signal.*

see Figure 8. Although the 1D signals are discretized for the numeric implementations, they represent 1D functions that are elements of an infinite dimensional Hilbert space.

Ultimately, [9] gives a lower bound of the asymptotic bias for shapes of curves in 2D, where terms depend on derivatives of the functions representing the curve. These derivatives can be interpreted as the derivative of the action of translations, which leads to the curvature of the orbit of the given function under the translations' action. As a consequence of these examples, we assume that the intuition provided by [34] applies to neuroimaging and that the bias depends on the ratio $\left(\frac{\sigma}{d}\right)^2$.

2. Computational representation of geometry and topology . Now that we have seen that the geometrization of template estimation leads us to investigate the topology of images, we show in this section how the Morse-Smale complices can be used to encode the topology and help the estimation of the geometric parameters d and σ previously introduced.

2.1. Definition of Morse-Smale complexes for (brain) images.

Morse-Smale (intensity) functions. A real-valued smooth map $I : \Omega \to \mathbb{R}$ is a *Morse function* if all its critical points are non-degenerate (the Hessian matrix is non-singular) and no two critical points have the same function value. The intensity function I representing a bi- or tri-dimensional brain image is a Morse function, at least after a convolution with a smoothing

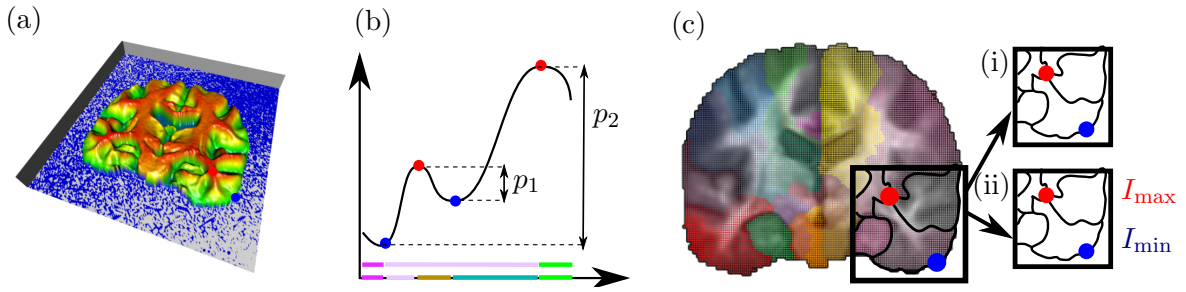

Figure 9. *(a) Intensity I on a 2D brain image visualized as height: maxima are in red, minima in blue. This is a Morse-Smale function $I : \Omega = \mathbb{R}^2 \to \mathbb{R}$. (b) Persistences $p_1 < p_2$ of two pairs min-max. A threshold $p_1 < p < p_2$ divides the domain into 3 regions (pink, violet, green), while $p < p_1$ divides it in 5 regions (pink, violet, brown, turquoise, green). (c) Computational representation of the geometry: Regions on a 2D domain, induced by a MSC with threshold $p = 0.1$: (i) The MS graph represents the isotropy group's class of the image, (ii) the labeled MS graph represents the image's orbit under the diffeomorphisms*

Gaussian [11]. In the following, I represents a brain image. *Morse theory* traditionally analyzes the topology of a manifold by studying the Morse functions on that manifold. Here, the manifold is known: it is the image domain Ω. We are not interested in the topology of Ω but rather in the topology of the functions I themselves, that is: we would like to know the distribution of their critical points. Figure 9(a) shows a 2D slice of a 3D brain image I, where the intensity is represented as the height, to better emphasize its maxima and minima: the maxima are in red and the minima in blue.

We introduce the notions of integral lines, ascending and descending manifolds that are needed to define Morse-Smale (intensity) functions. An *integral line* is a maximal path in the image domain Ω whose tangent vector correspond to the intensity gradient ∇I, the gradient of I, at every point. This notion comes from autonomous ordinary differential equation, where it represents the trajectory of a system verifying:

$$\frac{dx}{dt}(x) = \nabla I(x) \tag{4}$$

Each integral line starts and ends at critical points of I, where the gradient ∇I is zero. *Ascending $A(x_i)$ and descending $D(x_j)$ manifolds* of respective extrema x_i and x_j are defined as:

$$\begin{aligned} A(x_i) &= \{x \in \Omega \quad | \quad \text{The integral line going through } x \text{ ends at } x_i\} \\ D(x_j) &= \{x \in \Omega \quad | \quad \text{The integral line going through } x \text{ starts at } x_j\} \end{aligned}$$

Take the two manifolds $A(x_i)$ and $D(x_j)$ in Ω and assume they intersect at a point $p \in \Omega$. Let T_A (resp. T_D) denotes the set of all vectors tangent to $A(x_i)$ (resp. $D(x_j)$) at p. If every vector in Ω is the sum of a vector in T_A and a vector in T_D, then $A(x_i)$ and $D(x_j)$ are said to intersect transversely at the point p. The intensity function I defining the brain image is *Morse-Smale* if the ascending and descending manifolds only intersect transversely. We assume in the following that all brain images I are Morse-Smale.

Morse-Smale complex and persistence. The *Morse-Smale complex* of a Morse-Smale function is the set of intersections $A(x_i) \cap D(x_j)$, over all combinations of extrema (x_i, x_j) [19]. The Morse-Smale complex includes regions (i.e., sub-manifolds of Ω) of dimensions 0 through D, where D is the dimension of the domain Ω, i.e. $D = 2$ or $D = 3$ for our purposes. The *Morse-Smale (MS) complex of* I is a partition of the domain Ω into regions defined by the set of integral lines that share common starting and ending points. The interior of each region is monotonic with respect to the intensity I: a region contains no critical points and has a single local minimum and maximum on its boundary, see for example Figure 9(c)(ii) where the maximum $I_{\max}$ and minimum $I_{\min}$ are shown on the boundary of the grey region. The MS complex can also be seen as a graph on the brain image domain Ω whose nodes are the critical points of the brain image intensity.

The *persistence of a critical point* x_i *of* I is the amount of change in intensity I required to remove this critical point:

$$p(x_i) = |I(x_i) - I(n(x_i))| \tag{5}$$

where $n(x_i)$ is the critical point closest to x_i *in intensity*, among the critical points connected to x_i by an integral line [16]. The persistence of x_i is a measure of its significance as a critical point, i.e. importance of the topological feature. Figure 9(b) illustrates the definition of persistence on a 1D example. The function represented has 4 critical points: two minima and two maxima. The figure shows how they pair, as well as their persistence. On the x-axis, colors show the regions of the corresponding 1D Morse-Smale complex.

Beside this usual definition of persistence of a critical point, we define here the *persistence of a region of the Morse-Smale complex* as the amount of change in intensity required to remove this region from the MS complex, and more precisely:

$$p(\text{region}) = |I_{\max} - I_{\min}| \tag{6}$$

where $I_{\max}$ and $I_{\min}$ are respectively the maximum and the minimum in intensity of this region. In contrast to the definition of the persistence of a critical point, we do not rely on the saddle points, but only on the extrema i.e. the minima and maxima.

Hierarchy of Morse-Smale complexes. The notion of persistence of a region enables the definition of a *hierarchy of MS complexes* of one brain image I [19, 16]. One uses the ordering given by persistence to successively remove topological features from the image I. One starts with the MS complex of the brain image I defined above and one recursively removes the critical points with minimal persistence. This leads to a nested series of successively simplified Morse-Smale complexes. At each level, some of the MS regions are merged into a single region. Ultimately the Morse-Smale complex consists of only one region which is the entire domain Ω.

The persistence introduces a notion of scale at which the Morse-Smale complex of I is considered. One keeps only the nodes whose persistence is above the threshold. Figure 9(b) shows that the one dimensional domain is partitioned differently if one takes a threshold p below p_1 or between p_1 and p_2. We say that a Morse-Smale complex is represented at a given persistence level. At the scale of the persistence threshold p, the intensity is considered monotonic on each region of the MS.

We note that this MS hierarchy is different from a Gaussian scale space (GSS) hierarchy of images [37]. The latter takes critical points across smoothing scales and not across persistence levels.

2.2. Computing Morse-Smale complexes of (brain) images in practice. The previous definitions are relevant to (continuous) Morse-Smale theory and apply strictly for a continuous intensity function I. Nevertheless, the MS complex, introduced in terms of ascending and descending manifolds, can be computed for discrete brain images as follows [16]. We choose the approach of line integrals to compute the Morse-Smale complex. Other approaches may also be considered like the Delaunay triangulation [12].

Computing the Morse-Smale of a brain image. We compute the Morse-Smale complex of a brain image, which will later be the brain *template* image. Our input are $\{x_i, I_i\}$, i.e. the intensity values $\{I_i\}$ on a grid $\{x_i\}$ of Ω. We compute the integral lines of the intensity gradient, which we then gather to get the regions of the Morse-Smale complex. For each element of the grid x_i, following the gradient ∇I leads to computing the integral line going through x_i and in particular its starting and ending points [16]. The domain Ω can be approximated via a k nearest-neighbor graph and one computes the integral lines by considering the connectivity of the graph. Then, elements x_i's with same starting and ending points belong to the same Morse-Smale region. This gives the partition of the domain Ω and therefore the Morse-Smale complex. We remark that x_i's necessarily belong to a 3-dimensional (for a tridimensional image) component of the Morse-Smale complex because the 0-, 1- and 2-dimensional components have measure zero.

Figure 9(c) shows the Morse-Smale complex of the 2D slice of a 3D brain image for level of persistence of $p = 0.1$. The image's 2D domain is divided in different regions, represented by the different colors. The quadrant shows part of the underlying Morse-Smale graph. The red dot represents a maximum in intensity, and the blue dot a minimum of intensity. They are nodes of the underlying graph on the domain Ω.

Morse-Smale (MS) graph and labeled Morse-Smale (MS) graph. There are two ways of representing the Morse-Smale graph corresponding to the computed Morse-Smale complex. Both will be useful for analyzing the template's asymptotic bias. One can consider the graph as the set of nodes and edges, without any intensity information at the nodes. We simply call this graph the *Morse-Smale (MS) graph*: this is the graph illustrated on Figure 9(c)(i). Alternatively, one can label the nodes with the intensity information. We call this graph the *labeled Morse-Smale (MS) graph*: this is the graph illustrated on Figure 9(c)(ii). Both of these graphs are oriented, the edges being directed in the direction of the intensity gradient from a node to the next.

2.3. Template's computation and Morse-Smale complexes. We show how the MS complex of an image can represent its geometry and in particular isotropy group.

Lie algebra of the isotropy group and intensity gradient of the brain template. The template is an image $I \in \mathrm{Img}(\Omega)$. We consider the Lie group of diffeomorphisms $\mathrm{Diff}(\Omega) = C^\infty(\Omega)$ and its Lie algebra V, see Subsection 1.1.

Lemma 1. *Take $\epsilon > 0$ and consider the Lie algebra:*

$$V_I = \{v \in V \ \ s.t.: \ I \circ Exp(tv) = I, \forall |t| < \epsilon\} \tag{7}$$

By construction, the exponential of the elements of V_I are in the isotropy group of I. For $v \in V_I$, we have:

$$\forall x \in \Omega, \quad \nabla I(x)^T . v(x) = 0 \tag{8}$$

Proof. Take $v \in V_I$ and $|t| < \epsilon$. Its group exponential is a diffeomorphisms in the isotropy group G_I, which can be written:

$$\phi = \exp(tv) = Id + tv + \mathcal{O}(t^2)$$

Then, the equation above leads to:

$$\begin{aligned} I(x) &= I(x - tv(x) + \mathcal{O}(t^2)) \\ I(x) &= I(x) - DI(x).(tv(x)) + \mathcal{O}(t^2) \\ 0 &= DI(x).(tv(x)) + \mathcal{O}(t^2) \end{aligned}$$

The identification of the coefficients in this Taylor expansion leads to:

$$\nabla I(x).v(x) = 0$$

A vector field of V_I, which is in the Lie algebra of the isotropy group of the image I, is perpendicular the image's gradient at any point x of the image's domain Ω. ∎

We note that this lemma does not give a characterization of the vector fields in V_I. It gives the inclusion: $V_I \subset \{v | \forall x \in \Omega, \nabla I(x).v(x) = 0\}$. Thus, it allows to control the complexity of V_I, and thus of some of the isotropy group's Lie algebra. It represents a first step towards the following conjecture.

Conjecture 1. *The Lie algebra of the isotropy group G_I of the brain image I is constituted of vector fields that are everywhere perpendicular to the image's gradient.*

Sketch of Proof 1. *The proof will study the relations between the following three sets of vector fields:*

- *the Lie algebra $\mathfrak{g}_I$ of G_I,*
- *the set $V_I = \{v \in V \text{ s.t.: } I \circ Exp(tv) = I, \forall |t| < \epsilon\}$,*
- *the set $V_\perp = \{v | \forall x \in \Omega, \nabla I(x).v(x) = 0\}$.*

• *Proving $\mathfrak{g}_I = V_I$ by proving the double inclusion: (a) $\mathfrak{g}_I \subset V_I$ and (b) $V_I \subset \mathfrak{g}_I$.*
The inclusion (a) may be difficult to prove directly because of the lack of characterization of the infinite dimensional Lie algebra $\mathfrak{g}_I$. The inclusion (b) could make use of the above Lemma which shows that $Exp(V_I) \subset G_I$. We need to use the group logarithm in order to transform an inclusion on Lie groups into an inclusion on Lie algebra. The group logarithm is defined as the reciprocal of the group exponential, on its domain of bijectivity. As the domain of bijectivity is not well characterized in the general case, it is difficult to make use of this.
• *Proving $V_I = V_\perp$ by proving the double inclusion: (a) $V_I \subset V_\perp$ and (b) $V_\perp \subset V_I$.*
The inclusion (a) is proven in Lemma 1. The inclusion (b) may be proven in two steps. First, we would write the full Taylor expansion of $I \circ Exp(tv)$ where v is an element of $V_\perp$ and

show that the orders above the 1st give 0. At this point, the expression of the Taylor expansion is unclear, as well as if the condition $v \perp \nabla I$ is sufficient to make the orders above the 1st make 0. In a second step, we would need to translate the inclusion written in terms of Lie groups as an inclusion written in terms of Lie algebra, which is not trivial according to the paragraph above.

To understand the intuition behind this conjecture, consider a uniform image, i.e. with a constant intensity. In this case, there is no restrictions a priori on the vector fields of the Lie algebra of the image's isotropy group. Thus, the isotropy group is as large as it can be. We get that the isotropy group of an image with constant intensity is the whole group of diffeomorphisms.

Intensity gradient and MS graph. We now present a lemma showing that the MS graph can be used to computationally represent the isotropy group of an image. Take two images I_1 and I_2. Assume that their MS graphs are the same, regardless of the nodes' positions.

Lemma 2. *Assume we can map the level sets of I_1 to the level sets of I_2. This implies that we can map the partition of the domain I_1 induced by its Morse-Smale to the partition of the domain of I_2.*

Take a part $w \subset \Omega$ of this partition. There exists a diffeomorphism ψ and a function κ such that: $\nabla I_2(x) = \kappa(x).d^*\psi(x).\nabla I_1 \circ \psi(x), \forall x \in w$.

Proof. We consider a part $w \subset \Omega$ of the partition of the domain of I_1 induced by its Morse-Smale complex. On this part, there exists a function f such that:

$$I_1 = f \circ I_2 \circ \psi \tag{9}$$

The function f gives the mapping of intensity levels on each level set. Differentiating the previous equation with the chain rule gives:

$$dI_1 = df(I_2 \circ \psi).dI_2 \circ \psi.d\psi \tag{10}$$

where $df(I_2 \circ \psi)$ is a scalar which we note κ. Taking the adjoint gives:

$$\nabla I_1 = \kappa.\nabla I_2 \circ \psi.d^*\psi \tag{11}$$

■

This lemma is the first step towards the following conjecture.

Conjecture 2. *Two images with same MS graphs have same isotropy group.*

Sketch of Proof 2. *The images I_1 and I_2 have the same MS graph. The graph of I_1, taken with the nodes and edges positions on Ω, can be diffeomorphically deformed on the graph of I_2. We take ψ_1 a diffeomorphism that realizes the graphs' matching.*

Now, $I_1 \circ \psi_1^{(-1)}$ and I_2 share the same MS graph, taken with the nodes and edges' positions on Ω. We consider one cell of this graph.

We consider the integral lines of the respective gradients $\nabla\psi_1.\nabla I_1 \circ \psi_1$ and ∇I_2 on the cell. Both define a "parallel" partition of the cell. As a consequence, the set of integral lines of the gradient of $I_1 \circ \psi_1^{(-1)}$ can be mapped diffeomorphically to the set of integral lines of the gradient of I_2. We take ψ_2 a diffeormorphism that realizes the matching of the integral lines.

At this step, the notion of parallel partition would need to be written as a definition, together with the proof of the existence of ψ_2.

Here, we would need to combine the ψ_2's obtained on the different cells of the MS complex into one transformation on Ω that we also write ψ_2. A study of the behavior of ψ_2 on the edges of the MS complex should show that ψ_2 is itself a diffeomorphism on Ω. Then, we write $\psi = \psi_2 \circ \psi_1$. At the end of this step, I_1 is transformed to $I_1 \circ \psi^{(-1)}$.

The gradients of $I_1 \circ \psi^{(-1)}$ and I_2 share the same integral lines. Therefore ψ also maps the level sets of I_1 to the level sets of I_2. Using Lemma 2, we conclude.

In others words: if two images I_1, I_2 have the same MS graph, then I_1 can be diffeomorphically deformed so that its intensity gradient is parallel at every point to the intensity gradient of I_2.

From Lemma 1, the sets $\{v_1 | \forall x \in \Omega, \nabla I_1(x).v_1(x) = 0\}$ and $\{v_2 | \forall x \in \Omega, \nabla I_2(x).v_2(x) = 0\}$ control the isotropy groups of I_1 and I_2. If I_1 and I_2 have same MS graph, any vector field in the first set can be diffeomorphically deformed to get a vector field in the second set, and conversely. As a consequence, the MS graph represents the image's isotropy group. From Section 1, we know that the isotropy group controls, in turn, the orbit's type of the image i.e. to which stratum the image belongs.

Furthermore, we note that we could have considered the labeled MS graph of the image, i.e. the MS graph with intensities at the nodes, see Figure 9(c)(ii). The labeled MS graph controls the orbit of the image: images in the same orbit have the same topology but also same intensities.

3. Topology quantifies and controls the template's asymptotic bias.

This Section gathers the elements of Sections 1 and 2 to quantify the asymptotic bias in the brain template computation. We use Morse-Smale complexes to quantify, and then control, the bias.

3.1. Quantify the template inconsistency.

Understand and estimate the geometric parameter d. The distance d is the distance of the current image to a brain image with larger isotropy group, measured in sum of squared differences of intensities, see Figure 6. How can we measure this distance d locally on the template's image? From Section 1, we know that the isotropy group becomes larger when the image is "more symmetric". From Section 2, we know that the isotropy group becomes larger when the image topology becomes simpler. Thus, the distance d is a distance in intensity from the template image to a similar image with simpler topology.

We want to express this distance locally on the template image. Modifying the intensity locally on the template image modifies the image itself and may simplify its topology. For example, modifying the intensity locally in a region of the image can suppress a min-max pair and the image becomes "more symmetric". Thus we describe the distance d locally on the image by the amount of intensity needed to be changed in this region, so that the topology is simplified.

We quantify the local intensity needed to simplify the template image's topology using the Morse-Smale complex representation of Section 2. Let be given the Morse-Smale complex of the template image. The intensity needed to simplify the image's topology is, by definition, the intensity needed to simplify the Morse-Smale graph. We consider the partition of the

image's domain Ω induced by the Morse-Smale complex. For each region of the partition, the intensity needed to simplify the topology can be represented by the amount of intensity needed to remove the min-max pair of the region:

$$\hat{d}(\text{region}) = \hat{T}_{\max} - \hat{T}_{\min} = p_{\hat{T}}(\text{region}) \tag{12}$$

This quantifies the importance of the region as a representative of the brain anatomy: if the intensity difference between the region's min and max is low, then one can assume that this min-max pair has been created by chance because of the noise on the images. We see that the notion of persistence defined in Section 2 estimates the first geometric parameter d.

Understand and estimate the geometric parameter σ. Now we turn to the second geometric parameter that causes the asymptotic bias: the standard deviation σ of the noise, see again Equation 3 and Figure 6. The standard deviation σ of the noise is a parameter of the generative model that we assume has produced the observed images of the subjects brain anatomies. The parameter σ is unknown but it can be estimated from the observed images. Since we want to compute the asymptotic bias locally, we are interested in estimating the parameter σ locally, and for example on a region of the Morse-Smale complex of the template image. We estimate it as the average of the variability in intensity of the registered images in the region:

$$\hat{\sigma}(\text{region}) = \frac{1}{\#x} \sum_{x \in \text{region}} \hat{\sigma}(x), \tag{13}$$

where $\hat{\sigma}(x)$ is the variability in intensity of the registered images at the voxel x, and serves as an estimate of the noise at this voxel. This quantifies the amount of noise in this region. The larger the standard deviation σ of the noise, the more chances for the template to show min-max pairs that appeared by chance. One could use other estimator of the standard deviation, for example the sample standard deviation. Future work may investigate these estimators and their impact on the estimator of the template's bias.

Compute the asymptotic bias using the persistence of the whitened brain template. The local estimates of the geometric parameters d and σ enable us to estimate the asymptotic bias locally on a brain region:

$$\hat{B}_\infty(\text{region}) = \left(\frac{\hat{d}(\text{region})}{\hat{\sigma}(\text{region})} \right)^{-2}, \tag{14}$$

We emphasize here that $\hat{B}_\infty$ is an *estimate* of the asymptotic bias B_∞ (of the brain template estimation), and not an exact computation.

We link the estimate $\hat{B}_\infty$ to the definition of persistence in the Morse-Smale complex framework. First, we define the *whitened brain template estimate* $\hat{t}$ of $\hat{T}$ as:

$$\forall x \in \Omega, \quad \hat{t}(x) = \frac{\hat{T}(x)}{\hat{\sigma}(x)}. \tag{15}$$

In other words, we divide the brain template intensity of each voxel x by the estimation of the standard deviation of the noise at this voxel $\hat{\sigma}(x)$. This whitens the noise all over the brain template.

We assume that the critical points of $\hat{t}$ are close to the critical points of $\hat{T}$ and consider the Morse-Smale complex of the whitened template. We further assume that: $\hat{\sigma}_{\text{region}} \simeq \hat{\sigma}(\max) \simeq \hat{\sigma}(\min)$, where $\hat{\sigma}(\max), \hat{\sigma}(\min)$ are the variabilities at the respective min and max of the Morse-Smale complex. We can write:

$$\hat{B}_\infty(\text{region}) \simeq \left(\frac{\hat{T}_{\max}}{\hat{\sigma}(\max)} - \frac{\hat{T}_{\min}}{\hat{\sigma}(\min)} \right)^{-2} = \left(\hat{t}(\max) - \hat{t}(\min)\right)^{-2} = p_{\hat{t}}(\text{region})^{-2}, \tag{16}$$

where we recognize the persistence $p_{\hat{t}}(\text{region})$ of the corresponding region of the whitened template $\hat{t}$. This links the estimation of the asymptotic bias to the persistence of the whitened template's Morse-Smale complex. This shows how a topological property of the image in fact represents a statistical property of this image as the estimate of the brain template.

Hierarchy of the whitened template. The persistence of the whitened template quantifies locally the asymptotic bias, i.e. how far the brain template is from the unique brain anatomy of the generative model. Is there a statistical interpretation of the hierarchies of Morse-Smale complexes, introduced in Section 2? Let us consider another Morse-Smale of the whitened template's hierarchy, i.e. a Morse-Smale computed at a given persistence threshold $p_{\text{threshold}}$. There is an asymptotic bias threshold that corresponds, which we can write: $p_{\text{threshold}}^{-1/2}$. The regions kept in the new Morse-Smale are those having a persistence higher than the persistence threshold $p_{\text{threshold}}$, i.e. those having an asymptotic bias lower than the asymptotic bias threshold $p_{\text{threshold}}^{-1/2}$.

Therefore, if we can impose the topology of the brain template to match the new Morse-Smale of threshold $p_{\text{threshold}}^{-1/2}$, we control its asymptotic bias. This means that we preserve only the min-max pairs shown on the Morse-Smale graph chosen. It eliminates the min-max pairs that have been created by chance, because the noise on the images was at a similar level than the intensity signal on these regions. The next subsection explains how to impose the topology of a given Morse-Smale on the template's image.

3.2. Controlling the template's asymptotic bias by constraining its topology.

We are given the template's image and we want to force its asymptotic bias to be below a threshold, so that it is closer to estimating the anatomy of the database, i.e. the anatomy shared by the subject brains. The development above suggests to compute the Morse-Smale complex with a persistence threshold corresponding to the desired bias threshold. Then, enforcing template's topology to match the Morse-Smale complex will control its asymptotic bias. This enforcement procedure is called "Topological denoising".

3.2.1. Topological denoising.

Topological denoising is a procedure for smoothing an image, like our template image, while preserving topological features [23, 18]. The input of the procedure is the intensity function defining the template $\hat{T} : \Omega \to \mathbb{R}$ and a MS complex with intensity values at its nodes. Enforcing the template's topology to match the MS complex means that we compute $\hat{T}' : \Omega \to \mathbb{R}$ which is a smoothed version of original template estimate $\hat{T}$ containing only the intensity min-max pairs specified by the MS complex chosen. $\hat{T}'$ should be otherwise as close as possible to the original template estimate $\hat{T}$ in terms of intensity. The values and positions of the MS extrema are preserved, while all other extrema are removed

from the brain template estimate. Such procedure provides control over the topology of the brain image $\hat{T}$.

Formally, the original Topological denoising problem is written as the minimization [23]:

$$\begin{aligned} &\text{argmin}_{T'} \sum_{x_i \in \Omega} ||\hat{T}(x_i) - T'(x_i)||^2 + \int_\Omega ||\Delta T'||^2 \\ &\text{s.t. } T'(x_i) = \hat{T}(x_i) \text{ for } x_i \text{ a node of the MS complex} \\ &\quad T'(x_j) > T'(x_i) \text{ for } x_j \text{ neighbor of } x_i \text{ and } x_i \text{ minimum} \\ &\quad T'(x_j) < T'(x_i) \text{ for } x_j \text{ neighbor of } x_i \text{ and } x_i \text{ maximum} \\ &\quad T'(x_i) > \min_{\text{neighbor } x_j} T'(x_j) \text{ for } x_i \text{ not an extremum} \\ &\quad T'(x_i) < \max_{\text{neighbor } x_j} T'(x_j) \text{ for } x_i \text{ not an extremum.} \end{aligned}$$

The non linear inequality constraints make this optimization problem hard to solve. The solution suggested by [23] is to compute a *representative function* u that verifies the last four inequality constraints. Given this function u, the topological denoising problem becomes:

$$\begin{aligned} &\text{argmin}_{T'} \sum_{x_i \in \Omega} ||\hat{T}(x_i) - T'(x_i)||^2 + \int_\Omega ||\Delta T'||^2 \\ &\text{s.t. } T'(x_i) = \hat{T}(x_i) \text{ for } x_i \text{ a node of the MS complex} \\ &\quad (T'(x_i) - T'(x_j))(u(x_i) - u(x_j)) > 0, \text{ for } (x_i, x_j) \text{ a pair of neighbors,} \end{aligned}$$

where the last constraint means that the direction of T' shall be aligned with the direction of u. This alternative optimization problem is easily solved [23].

The representative function u can be computed by solving the Dirichlet problem:

$$\begin{aligned} &\text{argmin}_{u} \int_\Omega ||\nabla u||^2 \\ &\text{s.t. } u(x_i) = 0 \text{ for } x_i \text{ a minimum} \\ &\quad u(x_i) = 1 \text{ for } x_i \text{ a maximum.} \end{aligned}$$

Minimizers of the Dirichlet energy are harmonic functions, and their properties guarantee that x_i and x_j are minima and maxima and that u contains no other extrema inside the MS regions. We refer to [23] for further details.

Figure 10 shows examples of topological denoising. The topology to be enforced is represented by the red and blue dots, which are nodes of the MS complex: red for intensity maxima and blue for intensity minima. On the left example, the circle motifs that were inducing undesirable minima and maxima are removed. On the right example, two of the initial four maxima in the center of the image are removed too. Only the topology dictated by the input Morse-Smale complex is preserved.

3.2.2. Integrating the topological denoising in the template computation pipeline. The original template's computation is performed with the algorithm of [24] and use the LCC

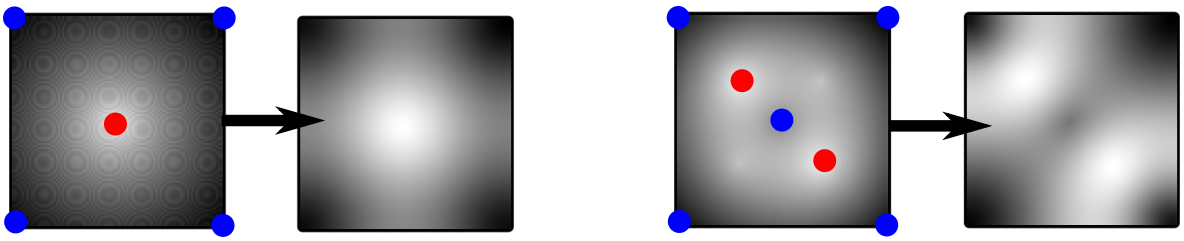

Figure 10. *Topological denoising on two toy examples. We impose topological constraints on the initial images, on the left in both cases: minima in blue and maxima in red. The arrows denote the action of the topological denoising and point to the output image.*

Log-demons for the registrations [29]. We adapt it by adding a Topological denoising step, in order to control the template's asymptotic bias.

Algorithm 1 shows the adapted procedure. One initiates with the template being one of the subject images: $\hat{T}_1 = I_1$. At each iteration k of the template's computation, one registers the subject images to the current template $\hat{T}_k$ and performs the average of the registered images' intensities to get a first version of the updated template $\hat{T}_{k+1}$. So far, this matches the usual template estimation procedure. Our adaptation is what follows. The MS complex of the updated template $\hat{T}_{k+1}$ is computed, using the R package msr [16]. Then, the updated template $\hat{T}_{k+1}$ is smoothed using Topological denoising, see Figure 12. These steps are iterated until convergence.

Algorithm 1 Controlled brain template estimation

Input: Images $\{I_i\}_{i=1}^n$, noise variance σ^2, persistence threshold $p_{\text{threshold}}$
Initialization:
$\hat{T}_1 = I_1$ (one of the subjects images)
$k = 1$
Repeat:
Non-linearly register the images to $\hat{T}_k$, i.e. compute ϕ_k^i: $J_k^i \simeq I^i \circ \phi_k^i$
Compute the mean deformation: $\bar{\phi}_k$
Register subject image: $L_k^i = I^i \circ \phi_k^i \circ \bar{\phi}_k^{-1}$
Compute the mean intensity image for template iteration: $\hat{T}_{k+1} = \frac{1}{n}\sum_{i=1}^n L_k^i$
Compute the MS complex of $\hat{T}_{k+1}$ at persistence level p
Topological denoising of T_{k+1} using the MS complex
$k \leftarrow k + 1$
until convergence: $||\hat{T}_k - \hat{T}_{k+1}|| < \epsilon$
Output: $\hat{T}_k$

The main parameter controlling this adapted procedure is the asymptotic bias threshold, i.e. the persistence threshold $p_{\text{threshold}}$ for the MS complex computation. The next section discusses the choice of this parameter $p_{\text{threshold}}$. Varying the threshold $p_{\text{threshold}}$ leads to the construction of a hierarchy of templates. The other parameter is σ, which is the noise on the

subject images. Either one knows it from the experimental design, or one estimates it with the variability of the registered subject images, as we did in Section 2.

4. Experimental results. This section presents experimental results on the quantification of the template's asymptotic bias and the adapted algorithm that bounds this bias. We use the Open Access Series of Imaging Studies (OASIS) database consisting of 136 T1 weighted MR images of brains [32].

4.1. Quantification of the template inconsistency. We quantify the asymptotic bias locally on the brain template computed from the OASIS database with the usual procedure. This shows how faithfully the computed template represents human brain anatomy for the neuroimaging studies.

First, we produce maps showing the local asymptotic bias directly with a color code superimposed on the original tridimensional template image, see Figure 11. We call these maps the *asymptotic bias maps*. A green color indicates a low asymptotic bias for the region and a red color indicates a high asymptotic bias on the region.

The scale for the color code corresponds to a logarithmic scale, and more precisely to $\mathrm{SNR}_{\mathrm{dB}}$, where:

$$\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10} \left[\left(\frac{d}{\sigma} \right)^2 \right]. \tag{17}$$

The scale is thus in dB, as the decibel is the logarithmic unit that expresses the ratio of two values of a physical quantity, which is the squared intensity in our case. This unit emphasizes that the quantification of the asymptotic bias depends on a signal-noise ratio (SNR). Indeed, one can consider that the signal is d, which is the template's intensities representing the brain anatomies and the "noise" is σ, the intersubject variability after registration. The larger is the SNR, the lower is the asymptotic bias on the brain template.

We compute several maps, see (c)-(d)-(e) on Figure 11 for the same brain template. The difference between the maps is the Morse-Smale complexes's persistence threshold used to compute the asymptotic bias. The threshold is increased from left to right on Figure 11 (c)-(d)-(e). Increasing the threshold makes more and more regions appear and these are more and more biased: they become colored in orange-red.

The asymptotic bias maps have the following interpretation with respect to neuroimaging. The maps show regions, in orange-red, where the template's brain structures are small with respect to the subjects' variability in the database. In these orange-red regions, it is not reasonable to have a sharply defined template, because the structures may have appeared by chance, by registration of noise between the different subjects. In other words, the maps reveal brain regions where the assumption of a unique anatomy in the subject population may break down.

4.2. Topological denoising for a consistent template.

Choice of the persistence threshold. Each map of Figure 11 (c)-(d)-(e) represents the asymptotic bias of the brain template we would obtain if we were constraining the image to the topology of the corresponding Morse-Smale. The persistence threshold gives a way to investigate the trade-off between asymptotic unbiasedness and sharpness of the template. On the

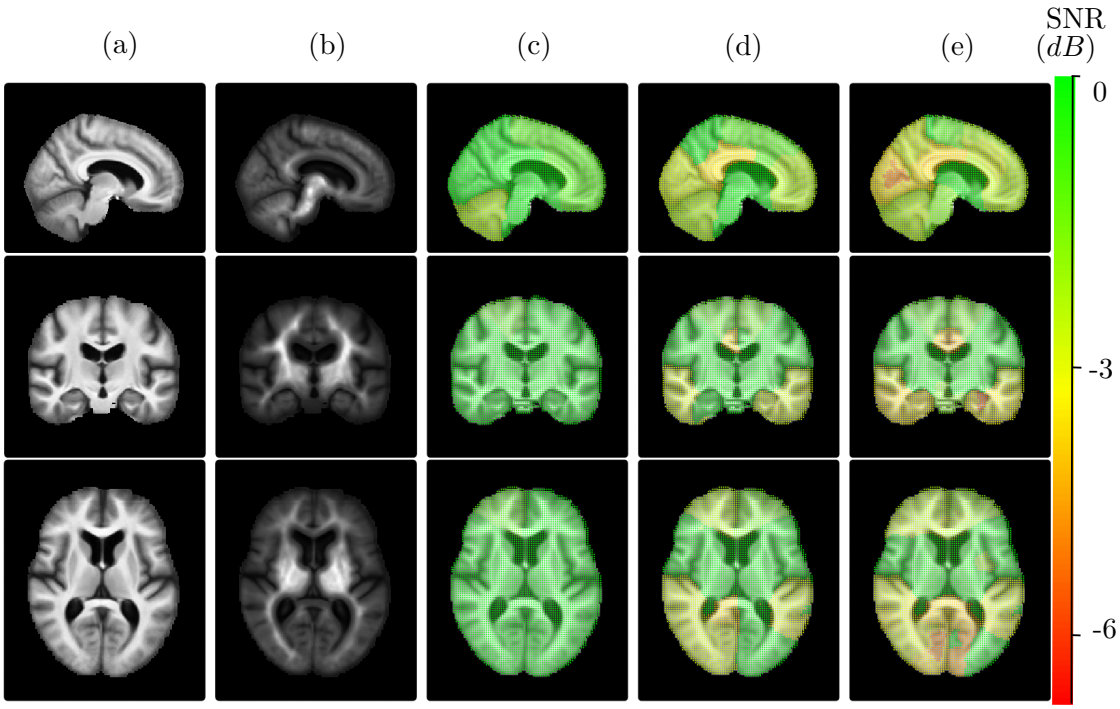

Figure 11. *Investigation of the template's consistency as an estimator of a unique anatomy. (a) Template. (b) Template whitened by the intersubject variability. (c) Region-wise inconsistency for a threshold= 1.3, (d) for threshold = 2, (e) for threshold = 4 (dimensionless).*

one hand, a complex topology - i.e. a low persistence threshold - implies an important asymptotic bias on the template, which may not represent faithfully the brain anatomy shared by the subjects in the OASIS database. On the other hand, a topology that is too simple - i.e. a high persistence threshold - has no chance of representing a brain anatomy at all. If we want to look at small brain structures, we have to allow for some precision in the topology.

Therefore, which topology shall we choose in this trade-off of asymptotic unbiasedness versus sharpness? If the local intensity of the computed template is below the noise, there is no hope to compute a consistent template. As in the 1D example of [2], if the noise is of the same order of magnitude as the signal, the template may estimate the noise instead of the signal. Thus it makes sense to choose an inconsistency threshold between -1 and 0 dB, that expresses the limit situation where signal (intensity on the brain image) and noise are of the same order of magnitude.

Applying topological denoising to control the brain template's bias. We apply the methodology of Section 3.2 to enforce the asymptotic bias to be below a threshold, using Topological denoising. Enforcing the unbiasedness in the procedure enables us to build the template of Figure 12. As a proof of concept, we have run it on the subject coronal slices of the OASIS database. Following the development above, we bound the asymptotic bias by setting the SNR threshold to -0.8 dB. We observe that the brain regions that were the more biased - i.e.

in orange-red in Figure 11 are now blurred. Thus, Topological denoising decides where the sharply defined brain template makes sense as a representative of the shared brain anatomy, and blurs it where it does not.

One could be interested in a template, that would be sharp *and* unbiased. In this case, one could consider dropping the assumption of a unique anatomy and consider multiple templates, i.e. use a mixture model. Further work is needed to investigate the construction of a stratified template, which would add a new stratification every time a region's asymptotic bias crosses the threshold $B_\infty \sim 1dB$.

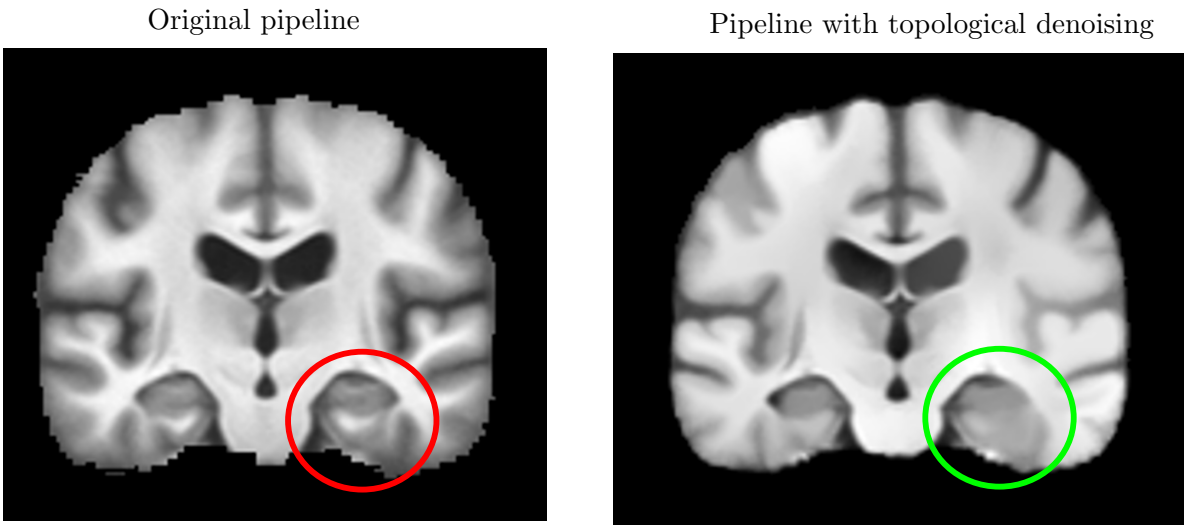

Figure 12. *Results of Topological Denoising integration in the pipeline for the template's estimation. Left: Template from pipeline without topological denoising. Right: Template with the topologically constrained pipeline. Inconsistent regions from Figure 11 are now blurred.*

Conclusion and perspectives. Computations of templates have been used in the medical imaging literature for at least 15 years. This paper investigates such computations as the estimations of a unique anatomy shared by the population. We have presented a topological method to quantify the asymptotic bias of the template. This is, to our knowledge, the first attempt to assess the bias of such procedures.

Our methodology builds a bridge between the diffeomorphic registration framework of Medical Imaging and Morse-Smale theory. This link is an interesting application of topology in itself. There are some limitations from the technical point of view as we provide sketches of proves for our conjectures. As such, this paper opens the door to mathematical developments at the boundary of Differential geometry and topology.

Our Morse-Smale framework identifies biased regions in the brain template in Section 3. In these regions, a sharp template might not be desirable. We control the template's asymptotic bias by adding a Topological Denoising step in its iterative computation, creating a trade-off between sharpness and unbiasedness. Our methodology is illustrated on a real database of 136 brain images in Section 4. It shows how the Topological Denoising blurrs the regions that

were the most biased. We control the template's bias at the price of dropping its sharpness.

It would be very interesting to be able to keep both the unbiasedness and the sharpness of the brain template. In fact, the template being biased can be seen as an indication that the assumption of a unique anatomy within the population should be relaxed. One could think about estimating a mixture of several templates or stratified templates. Each of the templates would represent only a subset of the brain population. This subset would have a lower variability. Therefore, the parameter σ will be decreasing and the bias too. This will allow for templates that are sharper and still unbiased.

REFERENCES

[1] D. Alekseevsky, A. Kriegl, M. Losik, and P. W. Michor, *The Riemannian geometry of orbit spaces. the metric, geodesics, and integrable systems*, Publ. Math. Debrecen, 62 (2003).

[2] S. Allassonnière, Y. Amit, and A. Trouvé, *Towards a coherent statistical framework for dense deformable template estimation*, Journal of the Royal Statistical Society, 69 (2007), pp. 3–29.

[3] S. Allassonnière, L. Devilliers, and X. Pennec, *Estimating the template in the total space with the Fréchet mean on quotient spaces may have a bias.*, in Proceedings of the fifth international workshop on Mathematical Foundations of Computational Anatomy (MFCA'15), 2015, pp. 131–142.

[4] M. Arnaudon, F. Barbaresco, and L. Yang, *Medians and means in Riemannian geometry: existence, uniqueness and computation.* working paper or preprint, Nov. 2011, https://hal.archives-ouvertes.fr/hal-00640626.

[5] J. Ashburner and K. Friston, *Why voxel-based morphometry should be used.*, Neuroimage, 14 (2001), pp. 1238–1243.

[6] J. Ashburner, C. Hutton, R. Frackowiak, I. Johnsrude, C. Price, and K. Friston, *Identifying global anatomical differences: Deformation-based morphometry*, Human Brain Mapping, 6 (1998), pp. 348–357.

[7] S. Baloch and C. Davatzikos, *Morphological appearance manifolds in computational anatomy: Group-wise registration and morphological analysis*, NeuroImage, 45 (2009), pp. S73 – S85.

[8] J. Bigot and B. Charlier, *On the consistency of Fréchet means in deformable models for curve and image analysis*, Electronic Journal of Statistics, (2011), pp. 1054–1089.

[9] J. Bigot and S. Gadat, *A deconvolution approach to estimation of a common shape in a shifted curves model*, Ann. Statist., 38 (2010), pp. 2422–2464.

[10] F. Bookstein, *Voxel-based morphometry should not be used with imperfectly registered images*, Neuroimage, 14 (6) (2001), pp. 1454–1462.

[11] U. V. Boscain, J. Duplaix, J.-P. Gauthier, and F. Rossi, *Anthropomorphic image reconstruction via hypoelliptic diffusion.*, SIAM J. Control and Optimization, 50 (2012), pp. 1309–1336.

[12] M. K. Chung, V. Singh, P. T. Kim, K. M. Dalton, and R. J. Davidson, *Topological Characterization of Signal in Brain Images Using Min-Max Diagrams*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 158–166.

[13] A. P. Dempster, N. M. Laird, and D. B. Rubin, *Maximum likelihood from incomplete data via the em algorithm*, Journal of the Royal Society, Series B, 39 (1977), pp. 1–38.

[14] L. Devilliers, X. Pennec, and S. Allassonniere, *Inconsistency of template estimation with the frchet mean in quotient space*, Proceedings of IPMI conference., (2017).

[15] A. Evans, A. Janke, D. Collins, and S. Baillet, *Brain templates and atlases*, Neuroimage, 62(2) (2012), pp. 911–922.

[16] S. Gerber and K. Potter, *Data analysis with the morse-smale complex: The msr package for r*, Journal of Statistical Software, 050 (2012).

[17] A. Guimond, J. Meunier, and J.-P. Thirion, *Automatic computation of average brain models*, in Proc. of First Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'98), vol. 1496 of LNCS, Cambridge, USA, October 1998, Springer, pp. 631–640.

[18] D. Günther, A. Jacobson, J. Reininghaus, H.-P. Seidel, O. Sorkine-Hornung, and

T. Weinkauf, *Fast and memory-efficient topological denoising of 2D and 3D scalar fields*, IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE VIS), 20 (2014), pp. 2585–2594.

[19] A. Gyulassy, P.-T. Bremer, B. Hamann, and V. Pascucci, *A practical approach to morse-smale complex computation: Scalability and generality*, IEEE Transactions on Visualization and Computer Graphics, 14 (2008), pp. 1619–1626.

[20] M. Hadj-Hamou, M. Lorenzi, N. Ayache, and X. Pennec, *Longitudinal Analysis of Image Time Series with Diffeomorphic Deformations: A Computational Framework Based on Stationary Velocity Fields*, Frontiers in Neuroscience, (2016).

[21] X. Han, *Topology preserving level set method for geometric deformable models*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 25 (2003), pp. 755–768.

[22] B. Hughes, *Geometric topology of stratified spaces.*, Electron. Res. Announc. Am. Math. Soc., 02 (1996), pp. 73–81.

[23] A. Jacobson, T. Weinkauf, and O. Sorkine, *Smooth shape-aware functions with controlled extrema*, Computer Graphics Forum (Proc. SGP), 31 (2012), pp. 1577–1586.

[24] S. Joshi, B. Davis, B. M. Jomier, and G. G. B, *Unbiased diffeomorphic atlas construction for computational anatomy*, Neuroimage, 23 (2004), pp. 151–160.

[25] W. S. Kendall, *Probability, convexity, and harmonic maps with small image I: Uniqueness and fine existence*, Proceedings of the London Mathematical Society, s3-61 (1990), pp. 371–406.

[26] A. Kriegl and P. Michor, *The Convenient Setting of Global Analysis*, Mathematical Surveys, American Mathematical Society, 1997, https://books.google.fr/books?id=l-XxBwAAQBAJ.

[27] S. A. Kurtek, A. Srivastava, and W. Wu, *Signal estimation under random time-warpings and nonlinear signal alignment*, in Advances in Neural Information Processing Systems 24, J. Shawe-taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Weinberger, eds., 2011, pp. 675–683.

[28] M. A. Lifshits, *Infinite-Dimensional Gaussian Distributions*, Springer Netherlands, Dordrecht, 1995, pp. 68–83.

[29] M. Lorenzi, N. Ayache, G. B. Frisoni, and X. Pennec, *LCC-Demons: a robust and accurate symmetric diffeomorphic registration algorithm*, NeuroImage, 81 (2013), pp. 470–483, doi:10.1016/j.neuroimage.2013.04.114, https://hal.inria.fr/hal-00819895.

[30] D. MacDonald, N. Kabani, D. Avis, and A. C. Evans, *Automated 3-d extraction of inner and outer surfaces of cerebral cortex from mri*, NeuroImage, 12 (2000), pp. 340–356.

[31] J.-F. Mangin, V. Frouin, I. Bloch, J. Régis, and J. López-Krahe, *From 3d magnetic resonance images to structural representations of the cortex topography using topology preserving deformations*, Journal of Mathematical Imaging and Vision, 5 (1995), pp. 297–318.

[32] D. Marcus, T. Wang, J. Parker, J. Csernansky, J. Morris, and R. Buckner, *Open access series of imaging studies (oasis): Cross-sectional mri data in young, middle aged, nondemented, and demented older adults.*, Journal of Cognitive Neuroscience., 19 (2007), pp. 1498–1507.

[33] J. Milnor, *Curvatures of left invariant metrics on lie groups*, Advances in Mathematics, 21 (1976), pp. 293 – 329.

[34] N. Miolane, S. Holmes, and X. Pennec, *Template shape estimation: correcting an asymptotic bias*, SIAM Journal of Imaging Science., (2017).

[35] N. Miolane and X. Pennec, *Biased estimators on quotient spaces*, Proceedings of the 2nd international of Geometric Science of Information (GSI'2015), (2015).

[36] X. Pennec, *Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements*, Journal of Mathematical Imaging and Vision, 25 (2006), pp. 127–154.

[37] J. Reininghaus, N. Kotava, D. Günther, J. Kasten, H. Hagen, and I. Hotz, *A scale space based persistence measure for critical points in 2d scalar fields*, IEEE Trans. Vis. Comput. Graph., 17 (2011), pp. 2045–2052.

[38] S. Weinberger, *The Topological Classification of Stratified Spaces*, Chicago Lectures in Mathematics, University of Chicago Press, 1994, https://books.google.fr/books?id=Q8RF_--FYVIC.

[39] L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences, Springer London, Limited, 2012.