Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Space-feature measures on meshes for mapping spatial transcriptomics

Michael I. Miller [a], Alain Trouvé [b], Laurent Younes [c,*]

[a] Center of Imaging Science and Department of Biomedical Engineering, Johns Hopkins University, United States of America
[b] Centre Giovanni Borelli (UMR 9010), Ecole Normale Supérieure Paris-Saclay, Université Paris-Saclay, France
[c] Center of imaging Science and Department of Applied Mathematics and Statistics, Johns Hopkins University, United States of America

ARTICLE INFO

Keywords:
Molecular imaging registration
Large deformation diffeomorphic metric mapping
Atlas registration

ABSTRACT

Advances in the development of largely automated microscopy methods such as MERFISH for imaging cellular structures in mouse brains are providing spatial detection of micron resolution gene expression. While there has been tremendous progress made in the field of Computational Anatomy (CA) to perform diffeomorphic mapping technologies at the tissue scales for advanced neuroinformatic studies in common coordinates, integration of molecular- and cellular-scale populations through statistical averaging via common coordinates remains yet unattained. This paper describes the first set of algorithms for calculating geodesics in the space of diffeomorphisms, what we term space-feature-measure LDDMM, extending the family of large deformation diffeomorphic metric mapping (LDDMM) algorithms to accommodate a space-feature action on marked particles which extends consistently to the tissue scales. It leads to the derivation of a cross-modality alignment algorithm of transcriptomic data to common coordinate systems attached to standard atlases.

We represent the brain data as geometric measures, termed as *space-feature measures* supported by a large number of unstructured points, each point representing a small volume in space and carrying a list of densities of *features* elements of a high-dimensional feature space. The shape of space-feature measure brain spaces is measured by transforming them by diffeomorphisms. The metric between these measures is obtained after embedding these objects in a linear space equipped with the norm, yielding a so-called "chordal metric".

1. Introduction

We are seeing a new period of method-driven renaissance in neuroanatomy, one that is distinguished by a focus on large-scale projects generating unprecedented amounts of spatially resolved brain data across multiple, complementary modalities. Recent years have seen many advances in the development of largely automated microscopy instruments for imaging cellular structures in mouse brain anatomy and function (Narasimhan et al., 2017; Osten and Margrie, 2013; Ragan et al., 2012; Zheng et al., 2013), including morphological reconstructions at the dense 3D electron microscopy (EM) (Helmstaedter, 2013) and at the mesoscale by whole-brain reconstructions (as exemplified by the BRAIN Cell Census Network (BICCN) project Ecker et al., 2017), neuronal projectomes (Oh et al., 2014; Zingg et al., 2014) and brain-wide maps of cell type distributions (Kim et al., 2017), spatial transcriptomics technologies, such as MERFISH (Chen et al., 2015; Moffitt et al., 2018; Xia et al., 2019) or STARmap (Kebschull et al., 2020; Wang* et al., 2018) generating massive amounts of gene expression data of thousands of genes at a time, as well as emergent barcoding technologies linking data on neuronal projectomes with dense transcriptional profiling at the single-cell level (Chen et al., 2019; Sun et al., 2021; Kebschull et al., 2016).

Since the publication of the Allen ISH atlas in 2006, methods for the spatial detection of gene expression have rapidly improved in their both spatial resolution and the number of genes that can be detected simultaneously. A host of different methods, including MERFISH, STARmap, seqFISH (Shah et al., 2016), and others now allows for the simultaneous measurement of a few hundreds to thousands of genes, and potentially all genes, at single molecule and hence single-cell resolution, and at the scale of whole brain sections. Spatially resolved transcriptomic data further offers an opportunity for obtaining multi-modal measurements at single-cell resolution allowing for the combination of dense spatial transcriptomics with the simultaneous measurement of single-neuron projection information using barcode sequencing in BARseq2 (Chen et al., 2019; Sun et al., 2021). Spatial transcriptomics data can also be registered to functional Ca2+ imaging data, linking gene expression and neuronal activity in behaving mice (Bugeon et al., 2021; Xu et al., 2020; Condylis et al., 2022). The importance of these technological advancements for understanding the dense metric structure of the brain by building up coarse physiological atlas scales built up from dense

* Corresponding author.
E-mail addresses: mim@jhu.edu (M.I. Miller), alain.trouve@ens-paris-saclay.fr (A. Trouvé), laurent.Younes@jhu.edu (L. Younes).

https://doi.org/10.1016/j.media.2023.103068
Received 20 January 2023; Received in revised form 18 September 2023; Accepted 19 December 2023
Available online 23 December 2023

imaging measurements at the cellular scales was recently recognized when spatial transcriptomics was selected as Nature method of the year (Xiaowei, 2021).

While disparate datasets are being collected from comparable brains and thus exist in a common underlying coordinate system, differences in data modalities and imaging technologies however map them into disparate spaces that need to be mapped to each other to allow integration and maximal impact of these datasets obtained under high expense. This is one of the principal goals of this paper to provide space-feature measure (SFM) LDMMM building on the progress made in the Computational Anatomy (CA) (Grenander and Miller, 1998; Ashburner, 2009; Pennec, 2011) community in the suite of methods called large deformation diffeomorphic metric mapping (LDDMM). Advances in CA for diffeomorphic mapping technologies to atlas coordinates at the tissue scales (Beg et al., 2005; Avants and Gee, 2004; Zheng et al., 2013; Stouffer et al., 2021) provide mapping technologies for advanced neuroinformatic studies in common coordinates. Integration of molecular scale populations through statistical averaging via common coordinates remains yet unattained at the molecular and cellular scales. Only recently have theories been put forward that extend the diffeomorphism atlasing technologies of CA (Miller et al., 2020, 2021, 2022; Grenander and Miller, 1998; Ashburner, 2009; Pennec, 2011; Beg et al., 2005; Avants and Gee, 2004; Zheng et al., 2013; Stouffer et al., 2021) to the molecular scales consistently with diffeomorphic mapping at the tissue scales, providing mapping technologies for advanced neuroinformatic studies in common coordinates, integration of molecular- and cellular-scale populations through statistical averaging via common coordinates. These theories, introduced in Miller et al. (2021, 2022), describe the geodesic equations for building correspondences between MERFISH samples using measure norms. This paper describes a new family of algorithms for calculating these geodesics, what we term space-feature-measure-, or SFM-, LDDMM, extending the family of large deformation diffeomorphic metric mapping (LDDMM) algorithms (Beg et al., 2005; Miller et al., 2006; Trouvé and Vialard, 2010; Vialard et al., 2012; Miller et al., 2015) to accommodate a space-feature action on marked particles described in Miller et al. (2021, 2022) which extends consistently to the tissue scales. We introduce and implement a mesh-based version of the method, reminiscent of existing point-based matching algorithms.

We expect that the SFM-LDDMM technologies will also be important for problems emerging now in digital pathology linking the molecular scales of histology with the tissue scales of MRI for understanding neurodegenerative diseases associated to the validation and further development of biomarkers as surrogates of molecular disease, as in Alzheimer's Disease (Stouffer et al., 2021).

This paper adapts advanced computational methods developed in the field of Computational Anatomy for neuroanatomical data analysis by a broad neuroscience research community using spatial transcriptomics datasets. We believe SFM-LDDMM can potentially be broadly used by the neuroscience community since we will describe algorithms for registering datasets collected from multiple animals as 3D stacks or as single 2D sections as well as build correspondence to the Common Coordinate Framework (CCF) template (Zeira et al., 2021) that enables standardized comparisons across different datasets and users. Further, the user will be able to overlay and segment the registered data with labels from atlases transporting the labels from the coarse tissue scale to the fine molecular scale using the diffeomorphic properties of the maps, including the Allen ARA 2008 atlas (Kim et al., 2015), CCFv3 2017 atlas (Dong, 2008) and the recently built unified Paxinos and CCFv3 atlas (Wang et al., 2020). This will allow users to perform statistical analysis on multiple sections within and across animals to extract a number of useful statistics of their data leveraging the anatomical labels and to share datasets within registered coordinate systems with other groups collaborating in different laboratories on data analysis.

The main contributions of this paper are the construction of a mesh-based version of the algorithm introduced in Miller et al. (2020, 2021, 2022), the development of the relevant algorithm and its implementation, together with the introduction and implementation of a cross-modality registration strategy. It is organized as follows. Methods are described in Section 2, starting with a general presentation of what we call *space-feature measure*, and of their discretization (Section 2.1), followed by a description of the LDDMM algorithm adapted to this kind of representation (Section 2.2). In Section 2.3, we propose an extension of our approach to cross-modality mapping, and in particular to the alignment of a spatially resolved transcriptomics dataset to an atlas with pre-labeled regions. We conclude this methods section with a remark on the analogies between the concept of space-feature measures and that of marked point processes (Section 2.4). Section 3 provides complementary details on the computation underlying the LDDMM algorithm.

Section 4 provides experimental results on synthetic and real data. It contains a description of how spatial transcriptomic data can be represented as space-feature measures (Section 4.2) and an experimental comparison with velocity-based LDDMM (Section 4.3). We conclude the paper with a discussion in Section 5.

2. Methods

2.1. Space-feature measures

2.1.1. Main concepts and definitions

Let $\mathcal{F}$ denote a "feature" space, where features typically correspond to high-dimensional measurements made by an imaging system, describing biological function. We are interested in the combined analysis of space and function, and will therefore work with the product space $\mathbb{R}^d \times \mathcal{F}$ ($d = 2$ or 3).

At macroscopic scale, an image is usually defined, using a continuum approximation, as a function $q : \mathbb{R}^d \to \mathcal{F}$. However, in biological imaging, the image values, discretized over pixels, result from the accumulated contributions (counts) of various chemical components collected in the imaged volume and are discrete in nature. Mathematically, "counting" is well represented using Dirac measures. The elementary Dirac measure $\delta_x \otimes \delta_f$ (for $x \in \mathbb{R}^d$ and $f \in \mathcal{F}$), when evaluated at a set $V \times A$ in $\mathbb{R}^d \times \mathcal{F}$, returns 1 if $x \in V$ and $f \in A$ and zero otherwise. It can be interpreted as an indication that a "basic element" (for example, a protein, or a cell) is observed at location x with feature f (which can be, for example, a protein species or a cell type). These measures can be added, taking

$$\mu = \sum_{k=1}^{n} \delta_{x_k} \otimes \delta_{f_k},$$

so that $\mu(V \times A)$ counts the number of pairs (x_k, f_k) in the $V \times A$ and can be viewed as a microscopic description of the data.

If V is a small volume, the quantity

$$\rho(V) = \frac{1}{|V|} \sum_{k=1}^{n} \delta_{x_k}(V)$$

where $|V|$ is the volume of V, measures the density of these elements. The quantity

$$\zeta_V = \frac{\sum_{k : x_k \in V} \delta_{f_k}}{\sum_{k=1}^{n} \delta_{x_k}(V)}$$

then provides a probability measure on $\mathcal{F}$ that describes the feature profile of V. The continuum approximation corresponding to macroscopic scales is obtained by fixing $x \in \mathbb{R}^d$ and letting V_x be an infinitesimal neighborhood of x with volume dx and approximating ζ_{V_x} with a limit ζ_x:

$$\mu(V_x \times A) \simeq \zeta_x(A)\rho(x)dx, \quad A \subset \mathcal{F}. \tag{1}$$

We introduce the following notation. If $\mathfrak{m}$ is a measure on $\mathbb{R}^d$ and ζ a transition probability from $\mathbb{R}^d$ to $\mathcal{F}$ (i.e., a function $x \mapsto \zeta_x$ where

ζ_x is a probability measure on $\mathcal{F}$), we define $\mathfrak{m} \bar{\otimes} \zeta$ as the measure such that

$$(\mathfrak{m} \bar{\otimes} \zeta)(U \times A) = \int_U \int_A d\zeta_x(f) d\mathfrak{m}(x) \tag{2}$$

for all measurable sets $U \subset \mathbb{R}^d$ and $A \subset \mathcal{F}$. The measure μ in Eq. (1) is equal to $(\rho\lambda) \bar{\otimes} \zeta$, where λ denotes Lebesgue's measure, with $d\lambda(x) = dx$. We will refer to measures on $\mathbb{R}^d \times \mathcal{F}$ as *space-feature measures*, and this concept provides a unified representation of microscopic and macroscopic scales. Space-feature measures are linear operators on the set of functions $F : \mathbb{R}^d \times \mathcal{F} \to \mathbb{R}$, with notation

$$(\mu \mid F) = \int_{\mathbb{R}^d \times \mathcal{F}} F(x, f) d\mu(x, f).$$

We emphasize that any function $q : \mathbb{R}^d \to \mathcal{F}$ can be considered as a space-feature measure μ_q such that $q \mapsto \mu_q$ provides a one-to-one representation of measurable functions, letting for any F,

$$(\mu_q \mid F) = \int_{\mathbb{R}^d} F(x, q(x)) dx\,. \tag{3}$$

We use diffeomorphisms to transform space-feature measures, with an action extending how they transform functions. Diffeomorphisms indeed act on functions $q : \mathbb{R}^d \to \mathcal{F}$ as $\varphi \cdot q = q \circ \varphi^{-1}$ with Eq. (3) implying

$$\begin{aligned}(\mu_{\varphi \cdot q} \mid F) &= \int_{\mathbb{R}^d} F(x, q \circ \varphi^{-1}(x)) dx \\ &= \int_{\mathbb{R}^d} |D\varphi|\, F(\varphi(x), q(x)) dx \\ &= (\mu_q \mid |D\varphi|\, F(\varphi(\cdot), \cdot))\end{aligned}$$

where $|D\varphi|$ is the absolute value of the Jacobian determinant of φ. This suggests defining the action of a diffeomorphism φ on an space-feature measure as an extension of $\mu_q \to \mu_{\varphi \cdot q}$, simply letting, for a general measure μ:

$$(\varphi \cdot \mu \mid F(\cdot, \cdot)) = (\mu \mid |D\varphi(\cdot)|\, F(\varphi(\cdot), \cdot)). \tag{4}$$

The following definition summarizes this discussion.

Definition 1. Let $\mathcal{F}$ be equipped with a σ-algebra making it a measurable space. A d-dimensional space-feature measure is a measure on the set $\mathbb{R}^d \times \mathcal{F}$.

If $\mathfrak{m}$ is a measure on $\mathbb{R}^d$ and ζ a transition probability from $\mathbb{R}^d$ to $\mathcal{F}$, the measure $\mathfrak{m} \bar{\otimes} \zeta$ in Eq. (2) is called a space-feature measure in disintegrated form.

Diffeomorphisms of $\mathbb{R}^d$ act on space-feature measures through the action defined in Eq. (4).

Remark 1. In the decomposition of Eq. (2), ζ_x only needs to be defined for x in the support of $\mathfrak{m}$.

Remark 2. If ζ is a fixed measure on $\mathcal{F}$, $\mathfrak{m} \bar{\otimes} \zeta$ is the product measure between $\mathfrak{m}$ and ζ, in which case we will prefer the standard notation $\mathfrak{m} \otimes \zeta$.

The previous discussion provides examples of space-feature measures in disintegrated form. First the "continuum space-feature measures" takes the form:

$$\mu = (\rho\,\lambda) \bar{\otimes} \zeta \tag{5a}$$

and the space-feature measure μ_q for function $q : \mathbb{R}^d \to \mathcal{F}$ has $\mathfrak{m}$ Lebesgue's measure and $\zeta_x = \delta_{q(x)}$. The discrete space-feature measure

$$\mu = \sum_{k=1}^{n} \delta_{x_k} \otimes \delta_{f_k},\ x_k \in \mathbb{R}^d, f_k \in \mathcal{F} \tag{5b}$$

has, when the x_k's are distinct, $\mathfrak{m} = \sum_{k=1}^n \delta_{x_k}$ and $\zeta_x = \delta_{q(x)}$ where $q : \mathbb{R}^d \to \mathcal{F}$ is any function such that $q(x_k) = f_k$ for $k = 1, \ldots, n$. Indeed, we can write, for any functions $F : \mathbb{R}^d \to \mathbb{R}$ and $G : \mathcal{F} \to \mathbb{R}$, and for $F \otimes G : \mathbb{R}^d \times \mathcal{F} \to \mathbb{R}$ such that $F \otimes G(x, f) = F(x)G(f)$,

$$\begin{aligned}\left(\mathfrak{m} \bar{\otimes} \zeta \mid F \otimes G\right) &= \int_{\mathbb{R}^d} \int_{\mathcal{F}} F(x) G(f) d\zeta_x(f) d\mathfrak{m}(x) \\ &= \int_{\mathbb{R}^d} F(x) G(q(x)) d\mathfrak{m}(x) \\ &= \sum_{k=1}^{n} F(x_k) G(f_k).\end{aligned}$$

Space-feature measures are the main focus of this paper, with a primary goal to develop a numerical approach allowing for their comparison. They have been introduced in Miller et al. (2021, 2022) as a tool for the analysis of spatially resolved transcriptomic images, in combination with a hierarchical modeling. (We will however only consider a single scale in the present paper.)

Measure representations of data have been used in different contexts in computational anatomy, mostly focusing on spatial components. Unlabeled point sets were represented by sums of Dirac measures in Glaunès et al. (2004a), and later adapted to geometric objects such as curves and surfaces with the notion of *currents* (Vaillant and Glaunès, 2005; Glaunès et al., 2008). This framework was furthermore extended by the introduction of oriented and non-oriented varifolds (Charon and Trouvé, 2013; Kaltenmark et al., 2017) and functional varifolds (Charlier et al., 2017). Note that varifolds (Almgren, 1966) were introduced as a mathematical representation of surfaces (or more generally of Riemannian manifolds), as measures in on the product space $\mathbb{R}^3 \times S^2$, where S^2 (replaced by a Grassmannian for general manifolds) is the unit sphere in $\mathbb{R}^3$, in order to facilitate the analysis of variational problems over surfaces. In that original model, the equivalent of $\mathfrak{m}$ in Definition 1 is the singular measure supported by a surface $M \subset \mathbb{R}^3$ and ζ_x is, for $x \in M$, the Dirac measure at the normal to M at x.

Remark 3. We point out that an alternate action of diffeomorphisms on measures can be defined, in which the Jacobian determinant is dropped from the right-hand side of Eq. (4). The resulting action (denoted $\varphi_\sharp \mu$) is the push-forward of the measure μ by φ. The resulting action on images (here interpreted as densities) is $\varphi_\sharp q = |D\varphi^{-1}|\, q \circ \varphi^{-1}$. This latter action is the one used in shape analysis to compare curves or surfaces (Charon and Trouvé, 2013). In our setting, where we need to compare tissues with similar compositions but different sizes, this push-forward action is not appropriate, since, say, expansion results in $|D\varphi^{-1}| < 1$ and a reduction of the original density (i.e., a sparsification of cells in tissue), which is undesirable. The action we choose throughout for space-feature measures leaves the magnitude of q unchanged, essentially creating more volume without changing the composition of the tissue.

2.1.2. A semi-discrete family of measures

Eq. (5b) describes a space-feature measure in full discrete form, which is well adapted for numerical computations. In the following, however, it will be convenient to have more flexibility on the image transition probabilities, allowing them to be non discrete (we still discretize the spatial domain using Dirac measures). This results in "semi-discrete measures", defined by

(i) A family of points $m_c \in \mathbb{R}^d$, $c \in C$, where C is a finite index set;
(ii) A list of positive numbers τ_c, $c \in C$, interpreted as infinitesimal volumes;
(iii) A second list of positive numbers, $\alpha_c \geq 0$, $c \in C$, interpreted as infinitesimal densities;
(iv) A list of probability measures on $\mathcal{F}$, ζ_c, $c \in C$;

To such a quadruple $(m_c, \tau_c, \alpha_c, \zeta_c)_{c \in C}$, we associate the space-feature measure with associated action via diffeomorphisms

$$\mu = \sum_{c \in C} \alpha_c\, \tau_c\, \delta_{m_c} \otimes \zeta_c. \tag{6a}$$

$$\varphi \cdot \mu = \sum_{c\in C} \alpha_c \, |D\varphi(m_c)| \, \tau_c \, \delta_{\varphi(m_c)} \otimes \zeta_c. \tag{6b}$$

These measures act linearly on functions on $\mathbb{R}^d \times \mathcal{F}$:

$$(\mu \mid F) = \sum_{c\in C} \alpha_c \tau_c \int_{\mathcal{F}} F(m_c, f) d\zeta_c(f).$$

Obviously, the introduction of α and τ is computationally redundant since only their product matters in the expression of μ, but their interpretation as density and volume is important in practice. Eq. (6b) can be summarized in action on quadruples, i.e.,

$$\varphi \cdot (m_c, \tau_c, \alpha_c, \zeta_c) = (\varphi(m_c), |D\varphi(m_c)| \, \tau_c, \alpha_c, \zeta_c)$$

expressing the fact that diffeomorphisms move points and change volumes without affecting tissue composition represented by density and feature.

2.1.3. Mesh-based measures

We now specialize further to the situation in which C indexes simplices in a simplicial mesh in $\mathbb{R}^d$ (i.e., tetrahedra in 3D and triangles in 2D). Letting d denote the dimension, (two or three), we define, given a nonempty index set I, a simplicial family as a collection $\boldsymbol{x} = (x_i, i \in I)$ of distinct points in $\mathbb{R}^d$ together with a family of $(d+1)$-tuples, $C = (c = (c_0, \ldots, c_d) \in I^{d+1})$ of indexes such that the simplices

$$\gamma_c(\boldsymbol{x}) = \left\{ \sum_{i=0}^{d} a_i x_{c_i}, a_0, \ldots, a_d \in [0, +\infty), a_0 + \cdots + a_d = 1 \right\} \subset \mathbb{R}^d$$

have non-empty interior with positive orientation, i.e., their volume is

$$|\gamma_c(\boldsymbol{x})| := \det(x_{c_1} - x_{c_0}, \ldots, x_{c_d} - x_{c_0})/d!, \tag{7}$$

requiring that the term in the right-hand side is positive. The simplex centers m_c are

$$m_c(\boldsymbol{x}) = \frac{1}{d+1}(x_{c_0} + \cdots + x_{c_d}).$$

This family forms a simplicial mesh of some subset D of $\mathbb{R}^d$ if the simplices only intersect at faces, edges or vertexes and their union is equal to D, but we will not need to enforce this constraint. The mesh is therefore represented by the triple $(I, C, \boldsymbol{x})$. Combined with densities $\alpha_c, c \in C$ and measures $\zeta_c, c \in C$, this mesh specifies a measure in the sense just introduced through the quadruple $(m_c(\boldsymbol{x}), |\gamma_c(\boldsymbol{x})|, \alpha_x, \zeta_c)_{c\in C}$. This is summarized in the next definition.

Definition 2. A mesh-based space-feature measure structure is given by a simplicial family $(I, C, \boldsymbol{x})$, a family of non-negative numbers $(\alpha_c, c \in C)$ and a family of probability measures on $\mathcal{F}$, $\boldsymbol{\zeta} = (\zeta_c, c \in C)$, with everything summarized as $\mathcal{T} = (I, C, \boldsymbol{x}, \boldsymbol{\alpha}, \boldsymbol{\zeta})$. The associated space-feature measure and the result of its transformation by a diffeomorphism φ are

$$\mu_{\mathcal{T}} = \sum_{c\in C} \alpha_c \, |\gamma_c(\boldsymbol{x})| \, \delta_{m_c(\boldsymbol{x})} \otimes \zeta_c \tag{8a}$$

$$\varphi \cdot \mu_{\mathcal{T}} = \sum_{\gamma\in\Gamma} \alpha_c \, |\gamma_c(\boldsymbol{x})| \, |D\varphi(m_c(\boldsymbol{x}))| \, \delta_{\varphi(m_c(\boldsymbol{x}))} \otimes \zeta_c \tag{8b}$$

Define the action of φ on $\mathcal{T}$ by $\varphi \cdot \mathcal{T} = (I, C, \varphi(\boldsymbol{x}), \boldsymbol{\alpha}, \boldsymbol{\zeta})$.

We have

$$\mu_{\varphi\cdot\mathcal{T}} = \sum_{\gamma\in\Gamma} \alpha_c \, |\gamma_c(\varphi(\boldsymbol{x}))| \, \delta_{m_c(\varphi(\boldsymbol{x}))} \otimes \zeta_c \simeq \varphi \cdot \mu_{\mathcal{T}} \ , \tag{9}$$

with the approximations $\varphi(m_c(\boldsymbol{x})) \simeq m_c(\varphi(\boldsymbol{x}))$ and

$$|D\varphi(m_c(\boldsymbol{x}))| \simeq \frac{|\gamma_c(\varphi(\boldsymbol{x}))|}{|\gamma_c(\boldsymbol{x})|}.$$

So the measure associated with the transformed mesh approximates the transformation of the measure associated with the initial mesh. The advantage of using the former is that the action only involves point transformations, which will allow us to use existing and well-established point matching procedures (and computer codes) with limited change.

2.2. LDDMM for discrete space-feature measures

Point-set, or particle, methods have been used for LDDMM in various contexts (Younes, 2019) including landmark (Joshi and Miller, 2000b; Vaillant et al., 2004; Glaunès et al., 2004a), curve (Glaunès et al., 2008; Durrleman et al., 2008a,b), surface (Glaunès et al., 2004b; Charon and Trouvé, 2013; Kaltenmark et al., 2017) or image (Allassonnière et al., 2005; Jain, 2011) registration and comparison. These methods all share the same diffeomorphic mechanism for point transformation, and differ on how discrepancies between the transformed original object and its target are penalized, through the introduction of object-specific "chordal distances" (Bauer et al., 2019) that embed shapes into a (generally much bigger) Hilbert space and measure the square norm of the difference between the representations in that space.

As described above, we use in this paper a representation of continuous or discrete biological images as measures. Such measures, which are by construction continuous linear forms on bounded continuous functions, are also continuous linear forms over reproducing kernel Hilbert spaces (RKHSs) of smooth functions, providing a two-step embedding image data → measure → H^* where H is an RKHS and H^* its topological dual. The resulting chordal distance is provided by the norm on H^*, which has a simple expression using the associated kernel. Indeed, if M is a topological space and K_M a positive kernel on M, one has, letting W be the RKHS associated with K_M

$$\|\mu\|^2_{W^*} = \int_M \int_M K_M(a, b) d\mu(a) d\mu(b) \tag{10}$$

for any measure μ on M (Glaunès et al., 2004a).

2.2.1. Space-feature-measure LDDMM

Here, $M = \mathbb{R}^d \times \mathcal{F}$ is a product space and we define K_M as the product of two positive kernels K_1 and K_2 defined on each component. This means that K_1 is defined on $\mathbb{R}^d \times \mathbb{R}^d$ with values in $\mathbb{R}$ such that, for all $n > 0$, all $x_1, \ldots, x_n \in \mathbb{R}^d$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, one has

$$\sum_{k,l=1}^{n} \lambda_k \lambda_l K_1(x_k, x_l) \geq 0.$$

The same condition is assumed for K_2, replacing $\mathbb{R}^d$ by $\mathcal{F}$. In other terms, K_1 determines the spatial part of the space-feature metric, while K_2 determines its feature part, and

$$K_M((x, f), (x', f')) = K_1(x, x') K_2(f, f').$$

A natural choice for K_1, the spatial kernel, is to use radial basis functions (such as Gaussian, or Matérn kernels (Aronszajn, 1950; Schaback and Wendland, 2006; Cheney and Light, 2009; Iske, 2018)). Feature kernels for quantitative image data can be chosen in a similar way. Kernels for categorical data are provided by positive definite matrices with size equal to the number of features, with entries equal to $K_2(f, g)$ for all pairs $f, g \in \mathcal{F}$. In the absence of any specific structure of the feature set, it is natural to use the identity matrix, i.e., $K_2(f, g) = 1$ if $f = g$ and 0 otherwise.

Following Eq. (10), the measure inner product is then defined by

$$\langle \mu, \mu' \rangle_{W^*} = \int_{\mathbb{R}^d\times\mathcal{F}} \int_{\mathbb{R}^d\times\mathcal{F}} K_1(x, x') K_2(f, f') d\mu(x, f) d\mu'(x', f'). \tag{11}$$

Let V be a space of vector fields, i.e., of functions $v : \mathbb{R}^d \to \mathbb{R}^d$. We assume that V is equipped with an inner product denoted $\langle \cdot, \cdot \rangle_V$ and associated norm $\| \cdot \|_V$ and forms furthermore a Hilbert space so that it is complete for its norm topology.[1] We also assume that elements in V

[1] Our notation $\langle \cdot, \cdot \rangle$ for inner products, applied to two elements of the same vector space, should not be confused with the previously introduced notation $(\cdot \mid \cdot)$, applied to a linear form and a vector.

have at least one continuous derivative, and more precisely that there exists a constant A such that, for any $x \in \mathbb{R}^d$ and any $v \in V$,

$$|v(x)| + |Dv(x)| \leq A\|v\|_V. \tag{12}$$

(Here, we let $|Dv(x)|$ denote any matrix norm applied to the $d \times d$ differential of v.)

Let two mesh-based measures be given as template and target respectively, associated with $\mathcal{T}^{(k)} = (I^{(k)}, C^{(k)}, \boldsymbol{x}^{(k)}, \boldsymbol{\alpha}^{(k)}, \boldsymbol{\zeta}^{(k)})$, $k = 0, 1$. Then, the mesh-based SFM-LDDMM variational problem is:

Variational Problem 1 (*Mesh-Based SFM-LDDMM*).

$$\inf_{v(\cdot)\in L^2([0,1],V)} \int_0^1 \|v(t)\|_V^2 dt + \frac{1}{\sigma^2}\|\mu_{\varphi(1)\cdot\mathcal{T}^{(0)}} - \mu_{\mathcal{T}^{(1)}}\|_{W^*}^2 \tag{13}$$

with $\qquad \partial_t \varphi(t) = v(t)\circ\varphi(t)$.

This formulation follows the common pattern of other LDDMM algorithms (Beg et al., 2005; Vaillant and Glaunès, 2005; Glaunès et al., 2004a; Charon and Trouvé, 2013; Younes, 2019). Because the action only affects vertexes, this problem can be reduced (using an RKHS argument introduced for the registration of landmarks (Joshi and Miller, 2000a; Glaunès et al., 2004a), discrete curves (Glaunès et al., 2004a) and surfaces (Vaillant and Glaunès, 2005)) to an optimization over point-set trajectories.

More precisely, Eq. (12) implies that V is a reproducing kernel Hilbert space, and because it is a space of vector fields, its kernel K_V is defined on $\mathbb{R}^d \times \mathbb{R}^d$ and takes values in the space of $d \times d$ matrices with real coefficients. This kernel is such that, for any fixed $y \in \mathbb{R}^d$, $a \in \mathbb{R}^d$: (i) the vector field $K_V(\cdot, y)a : x \mapsto K_V(x, y)a$ belongs to V and (ii)

$$\langle K_V(\cdot, y)a, v\rangle_V = a^T v(y)$$

for all $v \in V$.

The reduction then works as follows. Because the values of the end-cost in Eq. (13) only depends on $z_i(1) = \varphi(1, x_i^{(0)})$, $i \in I^{(0)}$, themselves uniquely specified by the values of $v(t, z_i(t))$, $t \in [0, 1]$, $i \in I^{(0)}$ via the equation $\partial_t z_i(t) = v(t, z_i(t))$, one sees that any optimal v must be such that $\|v(t)\|_V^2$ realizes, for all $t \in [0, 1]$, the minimum of $\|w\|_V^2$ among all vector fields w such that $w(z_i(t)) = v(t, z_i(t))$. Kernel interpolation theory (Aronszajn, 1950) then implies that v takes the form

$$v(t, \cdot) = \sum_{i\in I^{(0)}} K_V(\cdot, z_i(t))a_i(t) \tag{14}$$

where $\boldsymbol{a}(t) = (a_i(t), i \in I^{(0)})$ are d-dimensional vectors that provide new control variables. This gives the following reduced problem, in which we let

$$\mathcal{T}(\boldsymbol{z}) = (I^{(0)}, C^{(0)}, \boldsymbol{z}, \boldsymbol{\alpha}^{(0)}, \boldsymbol{\zeta}^{(0)}).$$

Variational Problem 2 (*Reduced Mesh-Based SFM-LDDMM*).

$$\inf_{a_i(t), t\in[0,1], i\in I^{(0)}} \sum_{i,j\in I^{(0)}} \int_0^1 a_i(t)^T K_V(z_i(t), z_j(t))a_j(t)dt + \frac{1}{\sigma^2}\|\mu_{\mathcal{T}(\boldsymbol{z}(1))} - \mu_{\mathcal{T}^{(1)}}\|_{W^*}^2 \tag{15a}$$

with

$$\partial_t z_i(t) = \sum_{j\in I^{(0)}} K_V(z_i(t), z_j(t))a_j(t), \quad z_i(0) = x_i^{(0)},\ i \in I^{(0)} . \tag{15b}$$

Remark 4. The template ($\mathcal{T}^{(0)}$) and target ($\mathcal{T}^{(1)}$) have an asymmetric role in the formulation of Problems 1 and 2. Typical data analysis studies in computational anatomy register all elements of a data set to a fixed template, justifying the asymmetry. In this context, it is important that the template be built using an approach that is permutation invariant with respect to data labels. This is the case, for example, of various methods that have been developed for shape averaging. The approach developed for surfaces in Ma et al. (2010) can, for example, be adapted to mesh-based measures with minor modifications. If desired, it is possible to symmetrize the registration problem, adapting one of several strategies introduced in the literature (Avants and Gee, 2004; Avants et al., 2008; Younes, 2007; Beg and Khan, 2007), using, for example, co-registration methods (two-shape averaging) or simultaneous estimation of a forward and inverse mapping. We are, however, not pursuing such directions in this paper.

2.2.2. Gradient of the objective function

The optimization in Problem 2 is with respect to the trajectories $\boldsymbol{a}(t) = (a_i(t), i \in I^{(0)})$. The optimal diffeomorphism $\varphi(t, \cdot)$ in Eq. (13) is then given by $\varphi(t, x) = y(t)$ where $y(\cdot)$ solves the ODE:

$$\partial_t y(t) = \sum_{i\in I^{(0)}} K(y(t), z_i(t))a_i(t)$$

with $y(0) = x$.

The gradient, with respect to $\boldsymbol{a}(\cdot)$, of the objective function in Eq. (15a) is obtained with the adjoint method and Pontryagin's maximum principle. Introduce a co-state $\boldsymbol{p}(\cdot) = (p_i(\cdot), i \in I^{(0)})$, and define the Hamiltonian, evaluated at configurations $\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{z}}, \tilde{\boldsymbol{a}}$ (that do not depend on time):

$$H(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{z}}, \tilde{\boldsymbol{a}}) = \sum_{i,j\in I^{(0)}} \tilde{p}_i^T K(\tilde{z}_i, \tilde{z}_j)\tilde{a}_j - \sum_{i,j\in I^{(0)}} \tilde{a}_i^T K(\tilde{z}_i, \tilde{z}_j)\tilde{a}_j.$$

Then the gradient is computed in two steps. One first solves the system:

$$\begin{cases} \partial_t \boldsymbol{z}(t) = \partial_{\tilde{\boldsymbol{p}}} H(\boldsymbol{p}(t), \boldsymbol{z}(t), \boldsymbol{a}(t)) \\ \partial_t \boldsymbol{p}(t) = -\partial_{\tilde{\boldsymbol{z}}} H(\boldsymbol{p}(t), \boldsymbol{z}(t), \boldsymbol{a}(t)) \end{cases} \tag{16a}$$

with boundary conditions $\boldsymbol{z}(0) = \boldsymbol{x}^{(0)}$ and:

$$\boldsymbol{p}(1) = -\nabla U(\boldsymbol{z}(1)) \tag{16b}$$

where

$$U(\boldsymbol{x}) = \frac{1}{\sigma^2}\|\mu_{\mathcal{T}(\boldsymbol{x})} - \mu_{\mathcal{T}^{(1)}}\|_{W^*}^2. \tag{16c}$$

The gradient of the objective function is then given by:

$$t \mapsto -\partial_{\tilde{\boldsymbol{a}}} H(\boldsymbol{p}(t), \boldsymbol{z}(t), \boldsymbol{a}(t))^T. \tag{16d}$$

When $K_V(x, y) = \kappa(x, y)Id_{\mathbb{R}^d}$, where κ is a scalar kernel, the second equation in Eq. (16a) gives

$$\partial_t p_i = -\sum_{j\in I^{(0)}} (p_i^T a_j + p_j^T a_i - 2a_i^T a_j)\nabla_1\kappa(z_i, z_j) \tag{17}$$

(where ∇_1 denotes the gradient with respect to the first variable) and Eq. (16d) becomes

$$\partial_{\tilde{\boldsymbol{a}}} H(\boldsymbol{p}, \boldsymbol{z}, \boldsymbol{a})^T = \sum_{j\in I^{(0)}} (p_i - 2a_i)\kappa(z_i, z_j).$$

The only computation that is specific to our situation is the evaluation of Eq. (16b) on which we now focus.

Remark 5. Rather than optimizing with respect to full trajectories $\boldsymbol{a}(\cdot)$, one can restrict the variables to those satisfying the Pontryagin maximum principle (Eq. (16a) with $\boldsymbol{p} = 2\boldsymbol{a}$) which only leaves the initial value $\boldsymbol{a}(0)$ as a free variable. While we do not provide details here, "geodesic shooting" algorithm (Vaillant et al., 2004; Allassonnière et al., 2005; Miller et al., 2006; Younes, 2007; Vialard et al., 2012) can be directly adapted in the context of SFM-LDDMM.

2.2.3. Derivative of the data attachment term in 3D and 2D

We now examine both the 3D and 2D cases using similar arguments for the computation of the derivative of U in Eq. (16c). The computation involves the weighted inward normal vectors to the faces of the

simplices. In 3D, for the tetrahedron $\gamma_c(\boldsymbol{x}) = \left\{\sum_{i=0}^3 a_i x_{c_i}, a_i \geq 0, \sum_{i=0}^3 a_i = 1\right\}$, these vectors are

$$\begin{aligned} n_{c,0} &= -(x_{c_2} - x_{c_1}) \times (x_{c_3} - x_{c_1}) \\ n_{c,1} &= (x_{c_2} - x_{c_0}) \times (x_3 - x_{c_0}) \\ n_{c,2} &= -(x_{c_1} - x_{c_0}) \times (x_{c_3} - x_{c_0}) \\ n_{c,3} &= (x_{c_1} - x_{c_0}) \times (x_{c_2} - x_{c_0}) \,. \end{aligned} \tag{18}$$

In 2D, with tetrahedra replaced by triangles, normals to triangle edges are defined as

$$\begin{aligned} n_{c,0} &= J(x_{c_2} - x_{c_1}) \\ n_{c,1} &= J(x_{c_0} - x_{c_2}) \\ n_{c,2} &= J(x_{c_1} - x_{c_0}) \end{aligned} \tag{19}$$

with $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Both Eqs. (18) and (19) obey the general definition in which, for $j \geq 1$, $n_{c,j}$ is the unique vector such that

$$n_{c,j}^T u = \det\left(x_{c_1} - x_{c_0}, \ldots, x_{c_{j-1}} - x_{c_0}, u, x_{c_{j+1}} - x_{c_0}, \ldots, x_{c_d} - x_{c_0}\right) \tag{20a}$$

for all $u \in \mathbb{R}^d$, and

$$n_{c,0}^T u = -\det\left(u, x_{c_2} - x_{c_1}, \ldots, x_{c_d} - x_{c_1}\right) \tag{20b}$$

for all $u \in \mathbb{R}^d$. They furthermore satisfy

$$\sum_{j=0}^{d} n_{c,j} = 0. \tag{20c}$$

(This property can be easily checked for $d = 2$ or 3.)

We now calculate the derivative of the data attachment term.

Proposition 1. *Let $\mathcal{T}^{(k)} = (I^{(k)}, C^{(k)}, \boldsymbol{x}^{(k)}, \boldsymbol{\alpha}^{(k)}, \boldsymbol{\zeta}^{(k)})$, $k = 0, 1$ and denote $\mathcal{T}(\boldsymbol{x}) = (I^{(0)}, C^{(0)}, \boldsymbol{x}, \boldsymbol{\alpha}^{(0)}, \boldsymbol{\zeta}^{(0)})$. Then the derivative of U in Eq. (16c) with respect to x_j is*

$$\partial_{x_j} U(\boldsymbol{x}) = \frac{2}{\sigma^2} \partial_{x_j} \langle \mu_{\mathcal{T}(\boldsymbol{x})}, \mu_{\mathcal{T}(\tilde{\boldsymbol{x}})} - \mu_{\mathcal{T}^{(1)}} \rangle_{W^*}.$$

evaluated with $\tilde{\boldsymbol{x}} = \boldsymbol{x}$, with

$$\begin{aligned} &\partial_{x_j} \langle \mu_{\mathcal{T}(\boldsymbol{x})}, \mu_{\mathcal{T}(\tilde{\boldsymbol{x}})} - \mu_{\mathcal{T}^{(1)}} \rangle_{W^*} = \\ &\sum_{c \in C : j \in c} \sum_{c' \in C'} \alpha_c \alpha'_{c'} \, |\gamma_{c'}(\boldsymbol{x}')| \langle \zeta_c, \zeta'_{c'} \rangle_{W_2^*} \left(\frac{1}{d+1} |\gamma_c(\boldsymbol{x})| \, \nabla_1 K_1(m_c(\boldsymbol{x}), m_{c'}(\boldsymbol{x}')) \right. \\ &\qquad \left. + \frac{1}{d!} K_1(m_c(\boldsymbol{x}), m_{c'}(\boldsymbol{x}')) n_c(x_j) \right) \end{aligned} \tag{21}$$

where $n_c(x_j)$ is the normal to the face opposed to x_j in $\gamma(c)$ of Eqs. (18) and (19).

See Section 3 for a proof.

2.2.4. Comparison with Beg's LDDMM image registration

We have noticed earlier that an image $q : \mathbb{R}^d \to \mathcal{F}$ could be identified with the measure $\mu_q = \lambda \bar{\otimes} \delta_{q(\cdot)}$. The inner product in Eq. (11) applied to two such measures μ_q and $\mu_{q'}$ then gives

$$\langle \mu_q, \mu_{q'} \rangle_{W^*} = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_1(x, x') K_2(q(x), q'(x')) \, dx \, dx'.$$

In particular, if images take vector values, say, $\mathcal{F} = \mathbb{R}^k$, and one chooses a Euclidean kernel for K_2, so that $K_2(f, f') = f^T f'$, one can rewrite

$$\langle \mu_q, \mu_{q'} \rangle_{W^*} = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_1(x, x') q(x)^T q'(x') \, dx \, dx'.$$

If one now assumes that K_1 has a square root (which is true for most kernels used in practice), i.e., that

$$K_1(x, y) = \int_{\mathbb{R}^d} \tilde{K}_1(x, z) \tilde{K}_1(z, y) dz,$$

where $\tilde{K}_1$ is a positive kernel, then it is easy to check that

$$\langle \mu_q, \mu_{q'} \rangle_{W^*} = \langle \tilde{K}_1 q, \tilde{K}_1 q' \rangle_{L^2} = \int_{\mathbb{R}^d} \tilde{K}_1 q(z)^T \tilde{K}_1 q'(z) dz$$

where we have let

$$\tilde{K}_1 q(z) = \int_{\mathbb{R}^d} \tilde{K}_1(z, x) q(x) \, dx.$$

In other terms, the measures $\mu_q \in W^*$ are in isometric correspondence with the functions $\tilde{K}_1 q \in L^2$ (noting that the L^2 norm is the one used in the LDDMM image matching algorithm Beg et al., 2005).

One can then consider applying LDDMM image matching to compare images considered as measures, and this does provide good results (see Fig. 8), but this approach has a notable difference with the measure model because

$$\varphi \cdot (\tilde{K}_1 q)(z) = (\tilde{K}_1 q)(\varphi^{-1}(z)) = \int_{\mathbb{R}^d} \tilde{K}_1(\varphi^{-1}(z), x) q(x) \, dx$$

differs from

$$\tilde{K}_1(q \circ \varphi^{-1})(z) = \int_{\mathbb{R}^d} \tilde{K}_1(z, \varphi(x)) |D\varphi(x)| \, q(x) \, dx.$$

In other terms, the representation $q \mapsto \tilde{K}_1 q$ is not compatible with the action of diffeomorphisms on images (it does not "commute" with it), while, by construction, the representation $q \mapsto \mu_q$ does. Note that smoothing images is a common practice with image LDDMM, as gradient descent algorithms involve computing derivatives of the compared images.

As an additional remark, we note that general space-feature measures can be also considered as images taking values in a possibly infinite dimensional vector space with an isometric representation using a square root for both space and image kernel. This representation associates with each measure $\mu = \mathfrak{m} \bar{\otimes} \zeta_x$ on $\mathbb{R}^d \times \mathcal{F}$ the function

$$q_\mu : z \mapsto \int_{\mathbb{R}^d} \int_{\mathcal{F}} \tilde{K}_1(z, x) \tilde{K}_2(\cdot, f) d\zeta_x(f) d\mathfrak{m}(x)$$

which is a mapping from $\mathbb{R}^d$ to the space of functions for $\mathcal{F}$ to $\mathbb{R}$. Now, if $\tilde{K}_2$ is a square root of K_2, in the sense that there exists a measure η_0 on $\mathcal{F}$ such that

$$K_2(f, f'') = \int_{\mathcal{F}} \tilde{K}_2(f, f') \tilde{K}_2(f', f'') d\eta_0(f'),$$

then one can see the measure inner product is isometric with the L^2 norm. When $\mathcal{F}$ is finite, this approach returns a representation of space-feature measures as vector-valued functions (with dimension $|\mathcal{F}|$). In this case, letting K_2 and $\tilde{K}_2$ be the identity kernels and using the counting measure on $\mathcal{F}$, one has simply

$$q_\mu(z) = \int_{\mathbb{R}^d} \tilde{K}_1(z, x) \zeta_x(\cdot) d\mathfrak{m}(x). \tag{22}$$

However, this representation suffers from the same limitation as the previous one regarding compatibility with the action of diffeomorphisms.

2.3. Atlasing: Crossing modality and scale

2.3.1. Variational problems

Transferring genomic, cellular and histological data to atlas coordinates is one of the mainstream examples of crossing modalities and crossing scale. Atlases (Fig. 1) are by definition often "cartoons" (Mumford and Shah, 1989) which make sense at the millimeter tissue scales but are used to interpret the finest molecular and particle scales. Similarly our work in human digital pathology brings histological micron scale of markers together with the atlas scales of Mai–Paxinos (Fig. 2). We associate to the atlas the features space of cartoon labels $\mathcal{L}$. We want to map the high-resolution gene features with associated probabilities ζ' on $\mathcal{F}$, the set of micro-scale functional features representing genomic expression or particle identity, to the tissue scale where we only have the cartoon labels.

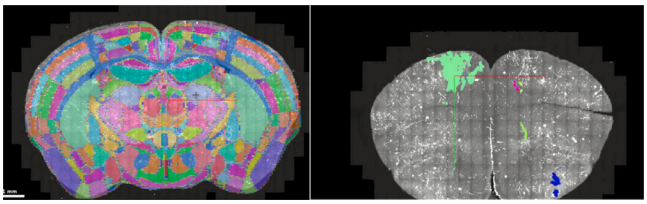

Fig. 1. Depicting Allen Atlas scale (left) and mouse micro scales right.

We assume that the measures are represented on meshes, as described in Section 2.1.3; the micro-scale fine target measure has $\mathcal{T}' = (I', C', \boldsymbol{x}', \boldsymbol{\alpha}', \boldsymbol{\zeta}')$, with

$$\mu_{\mathcal{T}'} = \sum_{c' \in C'} \alpha'_{c'} |\gamma_{c'}(\boldsymbol{x}')| \delta_{m_{c'}(\boldsymbol{x}')} \otimes \zeta'_{c'}.$$

At the coarse tissue scales, we assume that $\mathcal{L}$ is a small set of region labels forming the atlas features (typically, $|\mathcal{L}| \ll |\mathcal{F}|$). The atlas is represented as a measure defined on a mesh $\mathcal{T} = (I, C, \boldsymbol{x}, \mathbf{1}_C, \boldsymbol{\zeta})$, with $\zeta_c, c \in C$ the atlas feature probability assigning probability to region labels. The infinitesimal densities in this atlas measure are all taken equal to 1, as represented by the notation $\mathbf{1}_C$, reflecting the fact that density information is rarely provided in atlases. Through the measures ζ_c, we allow for the specification of a probabilistic atlas, and "cartoon representations" (piecewise constant images) are such that simplexes in the same region all have the same probability ζ_c which is a Dirac.

Each label $\ell \in \mathcal{L}$ has a specific expression pattern in the tissue, that we represent by an unknown probability distribution on $\mathcal{F}$, that we propose to estimate from data. We introduce for this purpose a parametric family of measures $(\pi_\vartheta, \vartheta \in \Theta)$ on $\mathcal{F}$ (see examples below). To each label $\ell \in \mathcal{L}$ is associated a parameter $\theta_\ell \in \Theta$ that needs to be estimated. The problem then becomes to simultaneously estimate the diffeomorphism φ of $\mathbb{R}^d$ mapping the atlas to the micro-scale measure and the parametrization vector $\theta = (\theta_\ell, \ell \in \mathcal{L})$.

Importantly, we do not assume that the π_θ's are probabilities measures, and we interpret $\pi_{\theta_\ell}(\mathcal{F})$ as a measure of the density of molecules or cells associated with label ℓ. Using this model, the imputed gene or cellular density at each site $c \in C$ becomes $\alpha_c = \sum_{\ell \in \mathcal{L}} \zeta_c(\ell) \pi_{\theta_\ell}(\mathcal{F})$ and the probability law at each site $c \in C$ is $\sum_{\ell \in \mathcal{L}} \zeta_c(\ell) \pi_{\theta_\ell} / \alpha_c$. We will assume that the densities are subject to box constraints in the form $\alpha_c^{\min} \le \alpha_c \le \alpha_c^{\max}$ for all $c \in C$, where $\alpha_c^{\min}$ and $\alpha_c^{\max}$ are known and correspond to prior expectation on the density of molecules or cells. Note that $\alpha_c^{\min} = 0$ and $\alpha_c^{\max} = \infty$ are allowed, and also $\alpha_c^{\min} = \alpha_c^{\max}$, yielding equality constraints, in case the densities are specified or estimated separately.

We therefore have a measure representation of the atlas (with imputed gene or cellular features)

$$\mu_{\mathcal{T}}^{\theta} = \sum_{c \in C} |\gamma_c(\boldsymbol{x})| \delta_{m_c(\boldsymbol{x})} \otimes \Big(\sum_{\ell \in \mathcal{L}} \zeta_c(\ell) \pi_{\theta_\ell} \Big).$$

The mapping problem is to map $\mu^{\theta}_{\varphi \cdot \mathcal{T}}$ close to $\mu_{\mathcal{T}'}$, with φ acting on meshes as defined in Section 2.1.3. This estimation can be performed using alternating minimization, looping over the estimation of φ with fixed θ and the estimation of θ with fixed φ minimizing $\|\mu^{\theta}_{\varphi(1)\cdot\mathcal{T}} - \mu_{\mathcal{T}'}\|^2_{W^*}$ with our variational problem for crossing scales.

Variational Problem 3 (*Cross-Modality SFM-LDDMM*).

$$\begin{aligned} &\inf_{\theta_\ell, \ell \in \mathcal{L}, v(\cdot) \in L^2([0,1],V)} \int_0^1 \|v(t)\|_V^2 dt + \frac{1}{\sigma^2} \|\mu^{\theta}_{\varphi(1)\cdot\mathcal{T}} - \mu_{\mathcal{T}'}\|^2_{W^*} \\ &\quad \text{with} \quad \partial_t \varphi(t) = v(t) \circ \varphi(t) \\ &\quad \text{and} \quad \alpha_c^{\min} \le \sum_{\ell \in \mathcal{L}} \zeta_c(\ell) \pi_{\theta_\ell}(\mathcal{F}) \le \alpha_c^{\max}, c \in C. \end{aligned} \tag{23}$$

Fixing $v(\cdot) \in L^2[0,1]$, with measure norm kernel a product form $K = K_1 K_2$, then we have:

$$\|\mu^{\theta}_{\varphi(1)\cdot\mathcal{T}} - \mu_{\mathcal{T}'}\|^2_{W^*} = \tag{24}$$

$$\begin{aligned} &\sum_{c_0, c_1 \in C} \sum_{\ell_0, \ell_1 \in \mathcal{L}} |\gamma_{c_0}| \, |\gamma_{c_1}| \zeta_{c_0}(\ell_0) \zeta_{c_1}(\ell_1) K_1(m_{c_0}, m_{c_1}) \int_{\mathcal{F}^2} \\ &\times K_2(f_0, f_1) d\pi_{\theta_{\ell_0}}(f_0) d\pi_{\theta_{\ell_1}}(f_1) \\ &- 2 \sum_{c \in C, c' \in C'} \sum_{\ell \in \mathcal{L}} \alpha'_{c'} |\gamma_c| \, |\gamma'_{c'}| \zeta_c(\ell) K_1(m_c, m'_{c'}) \int_{\mathcal{F}^2} \\ &\times K_2(f, f') d\pi_{\theta_\ell}(f) d\zeta'_{c'}(f') \, , \end{aligned}$$

where we have denoted for short $\gamma_c = \gamma_c(\varphi(1) \cdot \boldsymbol{x})$, $m_c = m_c(\varphi(1) \cdot \boldsymbol{x})$, $\gamma'_{c'} = \gamma_{c'}(\boldsymbol{x}')$, $m'_{c'} = m_{c'}(\boldsymbol{x}')$.

2.3.2. Special cases

Eq. (24) must be minimized in θ subject to the density constraints in Eq. (23). This generally provides a nonlinear programming problem. However, in some special cases, including those listed below, this problem boils down to quadratic programming (QP).

Example 1. Take cell types as micro-scale features, so that $(\zeta'_{c'}(f), f \in \mathcal{F})$ are the probabilities on each cell type (see Section 4.2.1). Let Θ be the space of positive measures on $\mathcal{F}$, taking $\pi_\theta(f) = \theta(f)$, $\theta(f) \ge 0$, $f \in \mathcal{F}$.

Choosing the identity kernel on $\mathcal{F}$, with $K_2(f, \tilde{f}) = 1$ if $f = \tilde{f}$ and 0 otherwise, the minimization of Eq. (24) reduces to:

$$\begin{aligned} &\inf_{\substack{\theta_\ell, \ell \in \mathcal{L} \\ \theta_\ell(f) \ge 0, f \in \mathcal{F}}} \sum_{c_0, c_1 \in C} \sum_{\ell_0, \ell_1 \in \mathcal{L}} |\gamma_{c_0}| \, |\gamma_{c_1}| \zeta_{c_0}(\ell_0) \zeta_{c_1}(\ell_1) K_1(m_{c_0}, m_{c_1}) \\ &\times \sum_{f \in \mathcal{F}} \theta_{\ell_0}(f) \theta_{\ell_1}(f) \\ &\qquad -2 \sum_{c \in C, c' \in C'} \sum_{\ell_0 \in \mathcal{L}} \alpha'_{c'} |\gamma_c| \, |\gamma'_{c'}| \zeta_c(\ell_0) K_1(m_c, m'_{c'}) \sum_{f \in \mathcal{F}} \theta_{\ell_0}(f) \zeta'_{c'}(f), \end{aligned}$$

with constraints

$$\alpha_c^{\min} \le \sum_{\ell \in \mathcal{L}} \sum_{f \in \mathcal{F}} \zeta_c(\ell) \theta_\ell(f) \le \alpha_c^{\max}, c \in C.$$

Example 2. Now consider mRNA counts on genes in $\mathcal{G}$. Take a simple model where parameters are $\theta = (\theta(g), g \in \mathcal{G}) \in [0, +\infty)^{\mathcal{G}}$ with $\pi_\theta = w_\theta \delta_{\eta_\theta}$, a weighted Dirac measure, taking $w_\theta = \sum_{g \in \mathcal{G}} \theta(g)$ and $\eta_\theta(g) = \theta(g)/w_\theta$. So, to each label ℓ is assigned an expression vector θ_ℓ with w_{θ_ℓ} representing the total expression and η_{θ_ℓ} a normalized expression. Taking the kernel as Euclidean on $\mathcal{F}$ (see remark below), $K_2(f, \tilde{f}) = \sum_{g \in \mathcal{G}} f(g) \tilde{f}(g)$, we have

$$\int_{\mathcal{F}} K_2(f, f') d\pi_{\theta_{\ell_0}}(f) d\zeta'_{c'}(f') = \sum_{g \in \mathcal{G}} w_{\theta_{\ell_0}} \eta_{\theta_{\ell_0}}(g) \bar{\theta}_{c'}(g) = \sum_{g \in \mathcal{G}} \theta_{\ell_0}(g) \bar{\theta}_{c'}(g)$$

where $\bar{\theta}_{c'}$ is the average expression $\int_{\mathcal{F}} f d\zeta'_{c'}(f)$. This gives the minimization of Eq. (24) reducing to:

$$\begin{aligned} &\inf_{\substack{\theta_\ell, \ell \in \mathcal{L} \\ \theta_\ell(g) \ge 0, g \in \mathcal{G}}} \sum_{c_0, c_1 \in C} \sum_{\ell_0, \ell_1 \in \mathcal{L}} |\gamma_{c_0}| \, |\gamma_{c_1}| \zeta_{c_0}(\ell_0) \zeta_{c_1}(\ell_1) K_1(m_{c_0}, m_{c_1}) \sum_{g \in \mathcal{G}} \theta_{\ell_0}(g) \theta_{\ell_1}(g) \\ &\qquad -2 \sum_{c \in C, c' \in C'} \sum_{\ell_0 \in \mathcal{L}} \alpha'_{c'} |\gamma_c| \, |\gamma'_{c'}| \zeta_c(\ell_0) K_1(m_c, m'_{c'}) \sum_{g \in \mathcal{G}} \theta_{\ell_0}(g) \bar{\theta}_{c'}(g) \end{aligned}$$

with constraints

$$\alpha_c^{\min} \le \sum_{\ell \in \mathcal{L}} \sum_{g \in \mathcal{G}} \zeta_c(\ell) \theta_\ell(g) \le \alpha_c^{\max}, c \in C.$$

Even though they were obtained from different models and contexts, the two examples above simplify to almost identical QP problems, respectively in $(\theta(f), f \in \mathcal{F})$ and $(\theta(g), g \in \mathcal{G})$. These problems are rephrased explicitly as QP problems below.

Algorithm 1. For Example 1., let $K_2(f, \tilde{f}) = 1$ if $f = \tilde{f}$ and 0 otherwise, defining

$$\begin{cases} A_{\ell_0, \ell_1} = \sum_{c_0, c_1 \in C} \alpha_{c_0} \alpha_{c_1} |\gamma_{c_0}| \, |\gamma_{c_1}| \zeta_{c_0}(\ell_0) \zeta_{c_1}(\ell_1) K_1(m_{c_0}, m_{c_1}) \\ b_{\ell_0}(f) = \sum_{c \in C, c' \in C'} \alpha_c \alpha'_{c'} |\gamma_c| \, |\gamma'_{c'}| \zeta_c(\ell_0) K_1(m_c, m'_{c'}) \zeta'_{c'}(f), \end{cases} \tag{25}$$

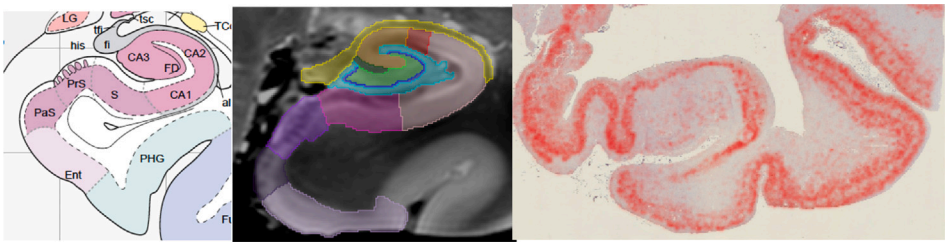

Fig. 2. Mai–Paxinos atlas section and high field MRI with Tau molecular pathology (right).

then $(\theta(f), f \in \mathcal{F})$ minimizes

$$\Phi_1(\theta) = \sum_{\ell_0,\ell_1\in\mathcal{F}} \sum_{f\in\mathcal{F}} A_{\ell_0,\ell_1}\theta_{\ell_0}(f)\theta_{\ell_1}(f) - 2\sum_{\ell_0\in\mathcal{L}}\sum_{f\in\mathcal{F}} b_{\ell_0}(f)\theta_{\ell_0}(f)$$
$$\text{subject to the constraints } \alpha_c^{\min} \le \sum_{\ell\in\mathcal{L}}\sum_{f\in\mathcal{F}} \zeta_c(\ell)\theta_\ell(f) \le \alpha_c^{\max}, c \in C. \tag{26}$$

Algorithm 2. For Example 2., let $K_2(f,\tilde{f}) = \sum_{g\in\mathcal{G}} f(g)\tilde{f}(g)$, $\bar{\theta}_{c'} = \int_{\mathcal{G}} f\, d\zeta'_{c'}$, taking A as in Eq. (25), and define

$$b_{\ell_0}(g) = \sum_{c\in C, c'\in C'} \alpha_c\alpha'_{c'}|\gamma_c|\,|\gamma'_{c'}|\zeta_c(\ell_0)K_1(m_c, m'_{c'})\bar{\theta}_{c'}, \tag{27}$$

then $(\theta(g), g \in \mathcal{G})$ minimizes

$$\Phi_2(\theta) = \sum_{\ell_0,\ell_1\in\mathcal{L}} \sum_{g\in\mathcal{G}} A_{\ell_0,\ell_1}\theta_{\ell_0}(g)\theta_{\ell_1}(g) - 2\sum_{\ell_0\in\mathcal{L}}\sum_{g\in\mathcal{G}} b_{\ell_0}(g)\theta_{\ell_0}(g)$$
$$\text{subject to the constraints } \alpha_c^{\min} \le \sum_{\ell\in\mathcal{L}}\sum_{g\in\mathcal{G}} \zeta_c(\ell)\theta_\ell(g) \le \alpha_c^{\max}, c \in C. \tag{28}$$

Remark 6. The Euclidean kernel used for K_2 in Example 2 is degenerate, in the sense that it induces a finite-dimensional feature RKHS, which can be identified to $\mathbb{R}^{\mathcal{G}}$ with the standard Euclidean norm. In this representation, probability measures on $\mathcal{F}$ are identified with their expectations, so that the metric does not differentiate between a Dirac measure δ_f and a probability measure ζ with expectation f. This explains why, in this example, the measures $\zeta'_{c'}$ were replaced with their associated average expression.

If K_1 is a spatial kernel, the product RKHS associated with K_1K_2 is identical to the RKHS of multivariate functions $x \mapsto F(x) \in \mathbb{R}^{\mathcal{G}}$, formed with the $|\mathcal{G}|$-fold tensor product of scalar RKHS's associated with K_1. Space-feature measures associated with a Euclidean image kernel are therefore identified to $|\mathcal{G}|$-dimensional vector measures.

2.4. Remark: Point processes and space-feature measures

If $\boldsymbol{N}$ is a compound point process, its realizations are discrete space-feature measures

$$N = \sum_{k=1}^{n} \delta_{x_k} \otimes \delta_{f_k}.$$

Their distribution are specified by a non-negative intensity function λ defined on $\mathbb{R}^d$ and a transition probability $(x,f) \mapsto \zeta(x,f)$ defined on $\mathbb{R}^d \times \mathcal{F}$ (so that $\zeta(x,\cdot)$ is for all x a probability on $\mathcal{F}$), such that

(i) for $\Omega \subset \mathbb{R}^d$ and $A \subset \mathcal{F}$, $\boldsymbol{N}(\Omega \times A)$ follows a Poisson distribution with parameter

$$\Lambda^{\boldsymbol{N}}(\Omega \times A) = \int_\Omega \lambda(x)\zeta(x,A)dx.$$

(ii) $\boldsymbol{N}(\Omega\times A)$ is independent of $\boldsymbol{N}(\Omega'\times A')$ if $(\Omega\times A)\cap(\Omega'\times A') = \emptyset$.

In the following we will make the abuse of notation $N(\Omega) = N(\Omega\times\mathcal{F})$ for the total number of points in Ω (N being a realization of $\boldsymbol{N}$) and similarly write $\Lambda^{\boldsymbol{N}}(\Omega) = \Lambda^{\boldsymbol{N}}(\Omega\times\mathcal{F})$. We will write $\boldsymbol{N} \sim \text{cpp}(\lambda,\zeta)$ to indicate that $\boldsymbol{N}$ is a compound Poisson point process with intensity λ and transition probability ζ. We assume that λ is compactly supported, which allows us to represent realizations of $\boldsymbol{N}$ as finite sums.

If F is a function on $\mathbb{R}^d \times \mathcal{F}$, its pairing with N is

$$(N|F) = \sum_{k=1}^{n} F(x_k, f_k).$$

The conditional expectation of $(\boldsymbol{N}|F)$ given $n, x_1, \dots, x_n$ is

$$E((\boldsymbol{N}|F)|n, x_1, \dots, x_n) = \sum_{k=1}^{n}\int_{\mathcal{F}} F(x_k,f)d\zeta(x_k,f)$$

which is the pairing of F with the measure $\sum_{k=1}^n \delta_{x_k}\otimes\zeta(x_k,\cdot)$. One also has

$$E((\boldsymbol{N}|F)) = \int_{\mathbb{R}^d}\int_{\mathcal{F}} F(x,f)d\zeta(x,f)\lambda(x)dx$$

therefore associated with the measure $\mu_{\boldsymbol{N}} = \lambda\otimes\zeta$.

If $\boldsymbol{N} = \text{cpp}(\lambda,\zeta)$ and φ is a diffeomorphism, we define $\varphi\cdot\boldsymbol{N} := \text{cpp}(\lambda\circ\varphi^{-1}, \zeta(\varphi^{-1}(\cdot),\cdot))$, so that

$$\begin{aligned}(\mu_{\varphi\cdot\mathbf{N}}|F) &= \int_{\mathbb{R}^d}\int_{\mathcal{F}} F(x,f)d\zeta(\varphi^{-1}(x),f)\lambda(\varphi^{-1}(x))dx\\ &= \int_{\mathbb{R}^d}\int_{\mathcal{F}} |d\varphi(x)|F(\varphi(x),f)d\zeta(x,f)\lambda(x)dx\\ &= (\varphi\cdot\mu_{\mathbf{N}}|F).\end{aligned}$$

As an illustration, we now design a test statistic to assess whether the observations N_1, N_2 of two independent compound Poisson processes result from $\boldsymbol{N}_1 = \varphi_1\cdot\boldsymbol{N}$ and $\boldsymbol{N}_2 = \varphi_2\cdot\boldsymbol{N}$, where $\boldsymbol{N} = \text{cpp}(\lambda,\zeta)$ and φ_1 and φ_2 are diffeomorphisms. ($\boldsymbol{N}$ is therefore a "template" compound Poisson process.) We here assume that φ_1 and φ_2 are known, and ignore the bias resulting from the fact that they have possibly been estimated using a registration procedure also involving N_1 and N_2.

Fix subsets $\Omega \subset \mathbb{R}^d$ and $A \subset \mathcal{F}$, and assume that λ and ζ take constant values, $\bar{\lambda}$ and $\bar{\zeta}$ on Ω (so that $\bar{\lambda}$ is a non-negative number and $\bar{\zeta}$ is a measure on $\mathcal{F}$). One then has

$$\Lambda^{\boldsymbol{N}}(\Omega\times A) = \int_\Omega \lambda(x)\zeta(x,A)dx = \bar{\lambda}\,\bar{\zeta}(A)\,|\Omega|$$

and, under the assumptions above:

$$\Lambda^{\boldsymbol{N}_i}(\Omega_i\times A) = \Lambda^{\boldsymbol{N}}(\Omega\times A)\frac{|\Omega_i|}{|\Omega|} = \Lambda^{\boldsymbol{N}}(\Omega\times A)\chi_i, \quad i = 1,2,$$

with $\Omega_i = \varphi_i(\Omega)$ and $\chi_i = |\Omega_i|/|\Omega|$. So, the null hypothesis is, for given Ω, A:

$$H_0(\Omega,A):\ \frac{\Lambda^{\boldsymbol{N}_1}(\Omega_1\times A)}{\chi_1} = \frac{\Lambda^{\boldsymbol{N}_2}(\Omega_2\times A)}{\chi_2},$$

and the alternative hypothesis is

$$H_1(\Omega,A):\ \frac{\Lambda^{\boldsymbol{N}_1}(\Omega_1\times A)}{\chi_1} \neq \frac{\Lambda^{\boldsymbol{N}_2}(\Omega_2\times A)}{\chi_2}.$$

One can use the likelihood ratio test statistic to compare the two hypotheses. Letting $\hat{p} = N_1(\Omega_1\times A)/(N_1(\Omega_1\times A) + N_2(\Omega_2\times A))$ and $p = \chi_1/(\chi_1+\chi_2)$, it is given by

$$T(\Omega,A) = (N_1(\Omega_1\times A) + N_2(\Omega_2\times A))\ D(\pi_{\hat{p}}||\pi_p)\,, \tag{29}$$

$$\text{with}\quad D(\pi_{\hat{p}}||\pi_p) = \left(\hat{p}\log\frac{\hat{p}}{p} + (1-\hat{p})\log\frac{1-\hat{p}}{1-p}\right)$$

This test statistic can be computed over partitions $(\Omega_c, c \in C)$ of the support of λ (small enough to justify the constancy assumption). If one takes $A = \mathcal{F}$, the collection $T(\Omega_c, \mathcal{F})$ focus on point density only. Using in addition a partition $(A_1, \dots, A_m)$ of $\mathcal{F}$ (obtained, for example, as clusters interpreted as cell types), we obtain a complete family of statistics $T(\Omega_c, A_j)$ that provides a high-dimensional analysis of the differences between the observed realizations.

3. Theory and calculation

3.1. Proof of Proposition 1

We repeat the statement of the proposition for convenience.

Proposition. *Let $\mathcal{T}^{(k)} = (I^{(k)}, C^{(k)}, \boldsymbol{x}^{(k)}, \boldsymbol{\alpha}^{(k)}, \boldsymbol{\zeta}^{(k)})$, $k = 0, 1$ and denote $\mathcal{T}(\boldsymbol{x}) = (I^{(0)}, C^{(0)}, \boldsymbol{x}, \boldsymbol{\alpha}^{(0)}, \boldsymbol{\zeta}^{(0)})$. Then the derivative of U in Eq. (16c) with respect to x_j is*

$$\partial_{x_j} U(\boldsymbol{x}) = \frac{2}{\sigma^2}\partial_{x_j}\langle \mu_{\mathcal{T}(\boldsymbol{x})}, \mu_{\mathcal{T}(\tilde{\boldsymbol{x}})} - \mu_{\mathcal{T}^{(1)}}\rangle_{W^*}.$$

evaluated with $\tilde{\boldsymbol{x}} = \boldsymbol{x}$, with

$$\partial_{x_j}\langle \mu_{\mathcal{T}(\boldsymbol{x})}, \mu_{\mathcal{T}(\tilde{\boldsymbol{x}})} - \mu_{\mathcal{T}^{(1)}}\rangle_{W^*} =$$
$$\sum_{c\in C: j\in c}\sum_{c'\in C'} \alpha_c\alpha'_{c'}\,|\gamma_{c'}(\boldsymbol{x}')|\langle\zeta_c, \zeta'_{c'}\rangle_{W_2^*}\left(\frac{1}{d+1}|\gamma_c(\boldsymbol{x})|\,\nabla_1 K_1(m_c(\boldsymbol{x}), m_{c'}(\boldsymbol{x}')) + \frac{1}{d!}K_1(m_c(\boldsymbol{x}), m_{c'}(\boldsymbol{x}'))n_c(x_j)\right)$$

where $n_c(x_j)$ is the 3D normal to the face opposed to x_j in $\gamma(c)$ of Eq. (18).

Proof. Let $\mathcal{T}' = (I', C', \boldsymbol{x}', \boldsymbol{\alpha}', \boldsymbol{\zeta}')$. We compute the variation of $\langle \mu_{\mathcal{T}(\boldsymbol{x})}, \mu_{\mathcal{T}'}\rangle_{W^*}$ with respect to a perturbation $\boldsymbol{h} = (h_i, i \in I)$ on the vertices $\boldsymbol{x}$ at $\epsilon = 0$:

$$\partial_\epsilon\langle \mu_{\mathcal{T}(\boldsymbol{x})}, \mu_{\mathcal{T}'}\rangle_{W^*}\Big|_{\epsilon=0} = \partial_\epsilon\Big(\sum_{c\in C, c'\in C'}\alpha_c\alpha'_{c'}\,|\gamma_c(\boldsymbol{x}+\epsilon\boldsymbol{h})|\,|\gamma_{c'}(\boldsymbol{x}')|\,\langle\zeta_c, \zeta'_{c'}\rangle_{W_2^*}K_1(m_c(\boldsymbol{x}+\epsilon\boldsymbol{h}), m_{c'}(\boldsymbol{x}'))\Big)\Big|_{\epsilon=0}.$$

Differentiating gives two terms given by the derivative of the kernel and the derivative of the volume term (defined in Eq. (7)). The derivative of the kernel is

$$\partial_\epsilon K_1(m_c(\boldsymbol{x}+\epsilon\boldsymbol{h}), m_{c'}(\boldsymbol{x}'))|_{\epsilon=0} = \frac{1}{d+1}\sum_{j=0}^{d}\nabla_1 K_1(m_c(\boldsymbol{x}), m_{c'}(\boldsymbol{x}'))^T h_{c_j}.$$

The derivative of the determinant in Eq. (7) gives:

$$\partial_\epsilon|\gamma_c(\boldsymbol{x}+\epsilon\boldsymbol{h})| = \frac{1}{d!}\sum_{j=1}^{d}\det\left(x_{c_1}-x_{c_0}, \dots, x_{c_{j-1}}-x_{c_0}, h_{c_j}-h_{c_0}, x_{c_{j+1}}-x_{c_0}, x_{c_d}-x_{c_0}\right)$$
$$= \frac{1}{d!}\sum_{j=1}^{d}(h_{c_j}-h_{c_0})^T n_{c,j} = \frac{1}{d!}\sum_{j=0}^{d}h_{c_j}^T n_{c,j}$$

where the last two equations use Eqs. (20a) and (20c), respectively.

Collecting terms involving h_j gives the variation:

$$\frac{1}{d+1}\sum_{c\in C, c'\in C'}\alpha_c\alpha'_{c'}|\gamma_c(\boldsymbol{x})|\,|\gamma_{c'}(\boldsymbol{x}')|\langle\zeta_c, \zeta'_{c'}\rangle_{W_2^*}\nabla_1 K_1(m_c(\boldsymbol{x}), m_{c'}(\boldsymbol{x}'))^T\Big(\sum_{j=0}^{d}h_{c_j}\Big)$$
$$+\frac{1}{d!}\sum_{c\in C, c'\in C'}\alpha_c\alpha'_{c'}|\gamma_{c'}(\boldsymbol{x}')|\langle\zeta_c, \zeta'_{c'}\rangle_{W_2^*}K_1(m_c(\boldsymbol{x}), m_{c'}(\boldsymbol{x}'))\Big(\sum_{j=0}^{d}n_{c_j}^T h_{c_j}\Big). \tag{30}$$

Removing dependence on the perturbation direction gives the partial derivative

$$\partial_{x_j}\langle \mu_{\mathcal{T}(\boldsymbol{x})}, \mu_{\mathcal{T}'}\rangle_{W^*} =$$
$$\sum_{c\in C: j\in c}\sum_{c'\in C'}\alpha_c\alpha'_{c'}|\gamma_{c'}(\boldsymbol{x}')|\langle\zeta_c, \zeta'_{c'}\rangle_{W_2^*}\left(\frac{1}{d+1}|\gamma_c(\boldsymbol{x})|\,\nabla_1 K_1(m_c(\boldsymbol{x}), m_{c'}(\boldsymbol{x}')) + \frac{1}{d!}K_1(m_c(\boldsymbol{x}), m_{c'}(\boldsymbol{x}'))n_c(x_j)\right)$$

where $n_c(x_j)$ is the inward weighted normal to the face opposed to x_j in $\gamma(c)$. We finally get the expression of the gradient of the data attachment term as

$$\partial_{x_j}\|\mu_{\mathcal{T}(\boldsymbol{x})} - \mu_{\mathcal{T}^{(1)}}\|^2_{W^*} = 2\partial_{x_j}\langle \mu_{\mathcal{T}(\boldsymbol{x})}, \mu_{\mathcal{T}(\tilde{\boldsymbol{x}})} - \mu_{\mathcal{T}^{(1)}}\rangle_{W^*}.$$

evaluated with $\tilde{\boldsymbol{x}} = \boldsymbol{x}$. □

4. Experiments

4.1. Synthetic examples

As a first example, we consider two shapes, supported by discs in 2D with a binary image feature so that each ζ_c is a Bernoulli distribution. Both images have a peripheral and a central region. The central region in the template is larger than that in the target, even though the template is globally smaller. The registration must therefore globally expand the shape while locally contracting the region occupied by the molecule. In the template, we have ζ_c = Bernoulli(0.25) in the periphery, and ζ_c = Bernoulli(0.75) in the center. In the target, probabilities are randomized independently for each simplex and ζ_c = Bernoulli(p_c) where p_c follows a beta distribution with parameters $(0.25\sigma, 0.75\sigma)$ (resp. $(0.75\sigma, 0.25\sigma)$) in the periphery (resp. in the center) with $\sigma = 10, 5, 1$ from less noisy to more. The weights α_c are all equal to one in the template, and are independently sampled following a gamma distribution with shape/scale parameters $(10\sigma, 1/10\sigma)$ in the target. Results, together with a computation of the log-Jacobian determinant of the transformation, are displayed in Fig. 3. They show a good adequacy between the deformed template and the target, with a slight degradation with increased noise. Fig. 4 provides similar results for data simulated withing a sphere. Here, we took two noise levels, namely $\sigma = 1$ and 10.

The 2D disc has 3.6 K vertexes and 7 K triangles and the 3D ball has 25 K vertexes and 140 K tetrahedra. Computation time in 2D is about 0.9 s per gradient iterations at 64 bit precision (0.25 s at 32 bits), using limited-memory BFGS gradient descent with a line search for Wolfe's condition (Nocedal and Wright, 2006), the whole procedure requiring a few hundred iterations depending on the specified tolerance constant. Our implementation (Younes, 0000) uses GPU acceleration, as provided by the pyKeops Python package (Charlier et al., 2021), and the code was run on a 32 CPU machine (Intel Xeon W-3245) equipped with an Nvidia GeForce RTX 3090 GPU. For the 3D example, the time per iteration was approximately 45 s at 64 bit precision (35 s at 32 bits). The bulk of the computation time is in the evaluation of kernel sums in Eqs. (15b) and (17), for which pyKeops is optimized. We have not used any approximation method such as those developed in particle mesh methods (Essmann et al., 1995), fast multipole (Greengard and Rokhlin, 1987; Coifman et al., 1993), or fast Gauss (Greengard and Strain, 1991; Greengard and Sun, 1998) transforms, or biconvolved kernels (Jain and Younes, 2013). The main reason is that these methods are designed for specific classes of kernels while our implementation is flexible in this regard. Another reason is that (excluding Jain and Younes, 2013) they do not guarantee the positiveness of the approximated kernel and may result in ill-posed minimization problems in case of negative eigenvalues.

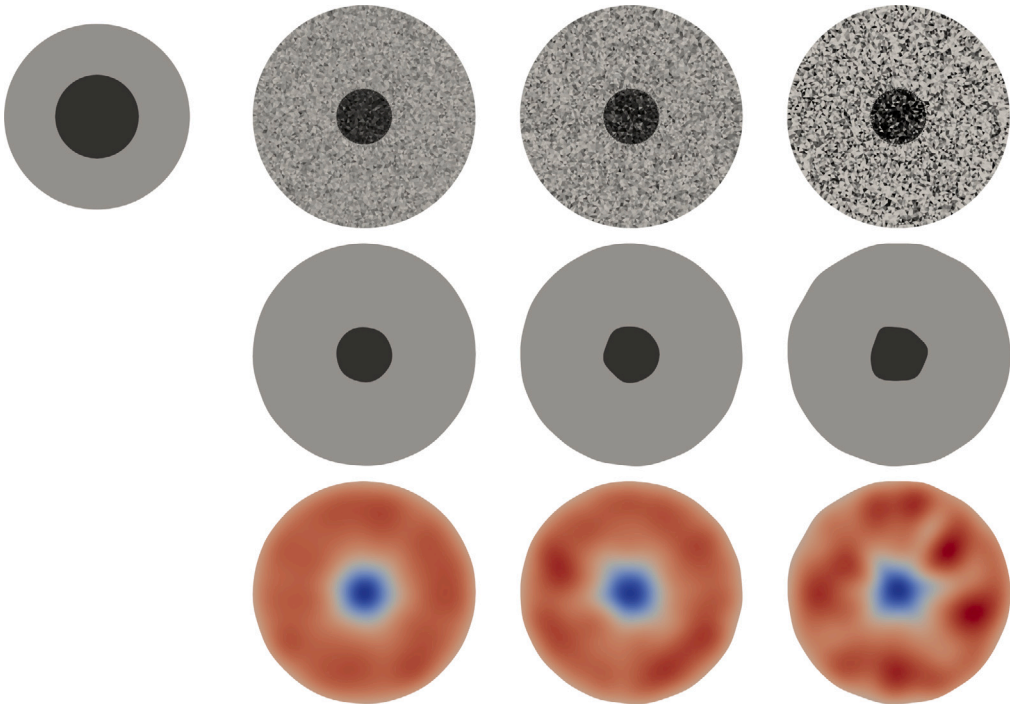

Fig. 3. First row: First image component (values between 0 and 1) of the 2D template and three targets with low, medium and high noise. Second row: Deformed templates after running LDDMM. Third row: logarithm of the Jacobian determinant of the deformation (red: expansion; blue: compression).

4.2. MERFISH image registration

4.2.1. Importing spatial transcriptomic data into space-feature measures

We now illustrate our algorithm on the registration of MERFISH images, and start with a discussion of how such data can be represented as space-feature measures.

Spatially resolved transcriptomics probe a large number of targeted mRNA molecules with high-resolution location information. After post-processing, this data can take various forms, for example represented as a 2D image with a large number of channels (associated with the measured gene set), as a list of points in space with attached gene count information (reconstructing single-cell RNAseq information combined with location Xia et al., 2019), or simply as a long list of single mRNA molecules with their detected location. We here consider a general representation that includes most situations of interest. We let $\mathcal{G}$ denote the set of targeted genes, whose size $|\mathcal{G}| = N$ can vary from several hundreds to several thousands.

We assume that the input data is a large family of spatial units, indexed by $j \in \mathcal{J}$, associating a location and a list of genes, in the form $y_j, (g_{j,k}, k = 1, \dots, n_j)$, indicating that genes $g_{j,1}, \dots, g_{j,n_j}$ were detected at location y_j (genes in the list may be repeated). A natural representation becomes the number of detections of gene g at location y_j, denoted

$$(y_j, (\nu_j(g), g \in \mathcal{G}))\ , j \in \mathcal{J}\ .$$

This representation includes raw spatially resolved transcriptomics data without additional processing with $n_j = 1$ for all j, as well as cell-centered data where y_j is the cell center and the $\nu_j(g)$'s are the gene counts associated with that cell. Preprocessing steps that cluster the raw data can be associated with this representation, the example we explore in the next section being a pre-analysis identifying single cells and cell types (Moffitt et al., 2018).

To construct our measures $\mu_{\mathcal{T}} = \sum_{c \in C} \alpha_c\, |\gamma_c(\boldsymbol{x})|\, \delta_{m_c(\boldsymbol{x})} \otimes \zeta_c$, we assume that a spatial resolution is given as a length parameter λ in μm defining a regular mesh that is first built within a bounding box containing the data and then pruned by (i) building the connected components formed by all empty simplices (containing no point y_j), and (ii) removing the component external to the data (so that empty islands within the image are retained). The pruned mesh provides the components $(I, C, \boldsymbol{x})$ supporting the space-feature measure, which must be completed by density weights and probability laws on the features, i.e., $(\alpha_c, \zeta_c), c \in C$. We now provide some possible strategies for their definition.

Gene features

The weights α_c represent densities (counts per unit volume). There are two options for their definitions, namely the density of detected mRNA molecules, or the density of points y_j, $j \in \mathcal{J}$, the latter choice making sense, e.g., if indexes $j \in \mathcal{J}$ enumerate single cells. In the first case, given that ζ_c is a probability distribution on features which we take as the genes $\mathcal{F} = \mathcal{G}$ with $f = g \in \mathcal{G}$, we let $\zeta_c(g)$ be the frequency of counts for gene $g \in \mathcal{G}$ relative to all the counts in $\gamma_c(\boldsymbol{x})$. In the second case, we let $\zeta_c(g)$ be the average of the normalized frequencies of counts over all j such that $y_j \in \gamma_c(\boldsymbol{x})$. This gives:

$$\zeta_c(g) = \frac{\sum_{j: y_j \in \gamma_c(\boldsymbol{x})} \nu_j(g)}{\sum_{\tilde{g} \in \mathcal{G}} \sum_{j: y_j \in \gamma_c(\boldsymbol{x})} \nu_j(\tilde{g})} \text{ with } \alpha_c = \frac{1}{|\gamma_c(\boldsymbol{x})|} \sum_{\tilde{g} \in \mathcal{G}} \sum_{j: y_j \in \gamma_c(\boldsymbol{x})} \nu_j(\tilde{g}) \quad (31a)$$

$$\zeta_c(g) = \frac{1}{|\{j\ :\ y_j \in \gamma_c(\boldsymbol{x})\}|} \sum_{j: y_j \in \gamma_c(\boldsymbol{x})} \frac{\nu_j(g)}{\sum_{\tilde{g} \in \mathcal{G}} \nu_j(\tilde{g})} \quad (31b)$$

$$\text{with } \alpha_c = \frac{1}{|\gamma_c(\boldsymbol{x})|} |\{j\ :\ y_j \in \gamma_c(\boldsymbol{x})\}|$$

We note that, with Eq. (31b), one disregards the information provided by the total number of counts in each cell. If one thinks of j as indexing cells in a tissue, the first choice for α_c relates to the number of counts per volume, and the second to the number of cells per volume.

Since $\mathcal{G}$ is a finite set, the image kernel is a positive definite matrix $(K_2(g, \tilde{g}), g, \tilde{g} \in \mathcal{G})$. The simplest choice for it is to use the identity, i.e., $K_2(g, \tilde{g}) = 1$ if $g = \tilde{g}$ and 0 otherwise.

Note that, in this section and in the next one, the set $\mathcal{G}$ may be replaced by a representative subset (gene panel) without any change to the discussion.

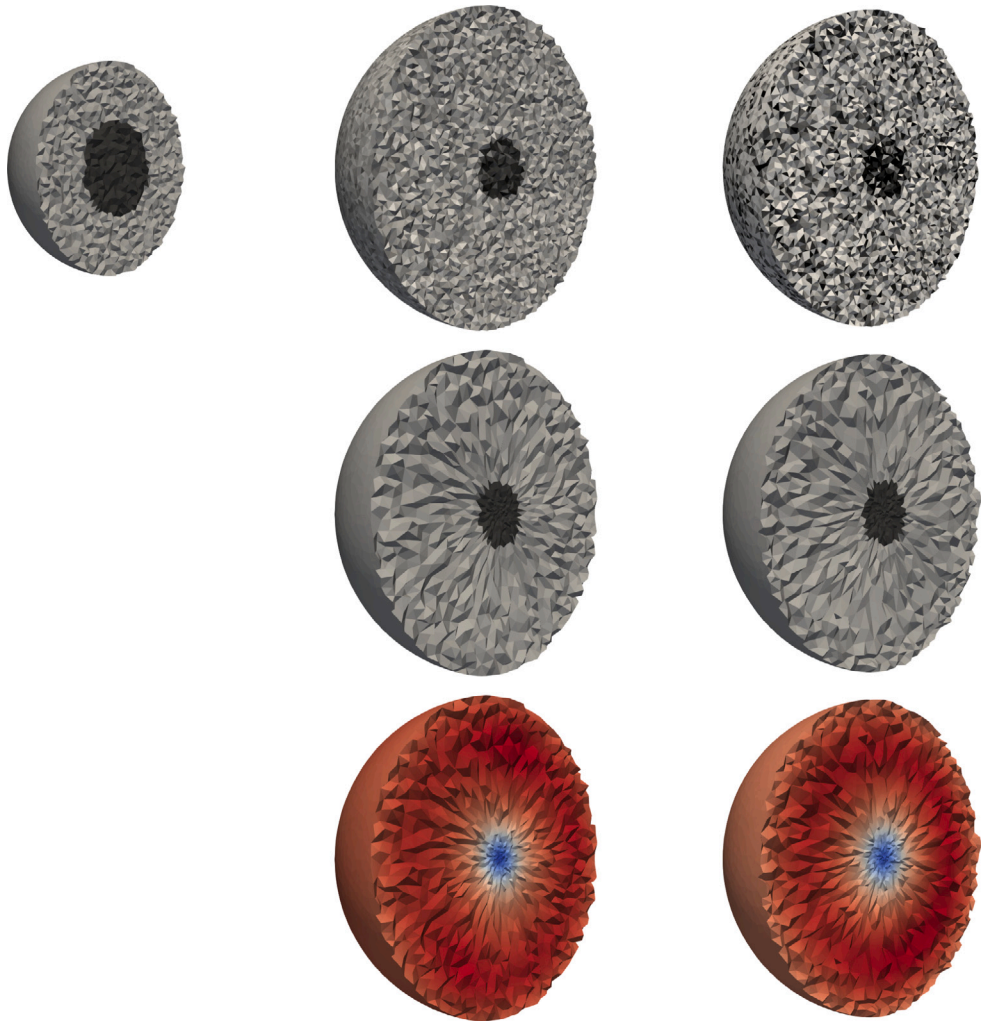

Fig. 4. First row: First image component (values between 0 and 1) of the 3D template and two targets with low and high noise. Second row: Deformed templates after running LDDMM. Third row: logarithm of the Jacobian determinant of the deformation (red: expansion; blue: compression).

RNA count features

Define the RNA count space to be features $\mathcal{F} = [0, +\infty)^{\mathcal{G}}$, that is, the set of all families $f = (f(g), g \in \mathcal{G})$ with $f(g) \geq 0$. In this context, the simplest choice is to let ζ_c be a Dirac measure at the averaged counts with mRNA count density:

$$\zeta_c = \delta_{\bar{\nu}_c}, \quad \bar{\nu}_c(g) = \frac{\sum_{j:y_j \in \gamma_c(\boldsymbol{x})} \nu_j(g)}{|\{j : y_j \in \gamma_c(\boldsymbol{x})\}|}, \tag{32}$$

$$\text{with} \quad \alpha_c = \frac{1}{|\gamma_c(\boldsymbol{x})|} |\{j : y_j \in \gamma_c(\boldsymbol{x})\}| .$$

There is a wide range of possible choices for the image kernel K_2, since we are working with quantitative data. The Gaussian kernel $K_2(\nu, \nu') = \exp(-|\nu - \nu'|^2/2\sigma^2)$ is a standard example. For our experiments in the next section, we use the product of a Euclidean and a Cauchy kernels, namely

$$K_2(\nu, \nu') = \frac{\nu^T \nu'}{\sigma^2 + |\nu - \nu'|^2}. \tag{33}$$

Note that, since elements of $\mathcal{F}$ are non-negative, the kernel K_2 can be computed in log scale, i.e., applied to $\log(1 + \nu), \log(1 + \nu')$ instead of ν, ν'.

Cell label features

Now examine the features to be cell types where we assume that the input data has been preprocessed to return cell-type labels. We let $\mathcal{F} = \{\ell_1, \ldots, \ell_p\}$, the label set, and assume that the data is a list (y_j, L_j) for locations and labels for $j \in \mathcal{J}$. The measure ζ_c can the be defined as

$$\zeta_c(\ell_k) = \frac{|\{j : y_j \in \gamma_c(\boldsymbol{x}), L_j = \ell_k\}|}{|\{j : y_j \in \gamma_c(\boldsymbol{x})\}|} \tag{34}$$

$$\text{with} \quad \alpha_c = \frac{1}{|\gamma_c(\boldsymbol{x})|} |\{j : y_j \in \gamma_c(\boldsymbol{x})\}| .$$

It is natural to use a kernel for which labels are orthogonal, i.e., $K_2(\ell, \tilde{\ell}) = 1$ if $\ell = \tilde{\ell}$ and 0 otherwise for $\ell, \tilde{\ell} \in \mathcal{F}$.

4.2.2. Experimental results

We illustrate the previous discussion with preliminary results based on MERFISH images of mouse brains (Zeng, 2022). Two-dimensional MERFISH datasets were discretized on grids with spatial resolution $\lambda = 100$ μm. Out of the 700 genes provided for the image, a subset of 10 genes with largest standard deviation was selected to build the space-feature measures. We used a Gaussian kernel for K_1 and the kernel provided in Eq. (33) for K_2 after switching to log scale. Fig. 5 provides images from two brain sections from the same mouse, the first one being used as template and the second as the target for registration. Fig. 6 shows similar results with greater deformations, using the same template, this time mapped to a brain section from a second mouse. The deformation grids are shown in Fig. 7. We see small deformation when registering the two sections from the same brain, which are quite similar, and much stronger changes for the alignment of sections from different brains.

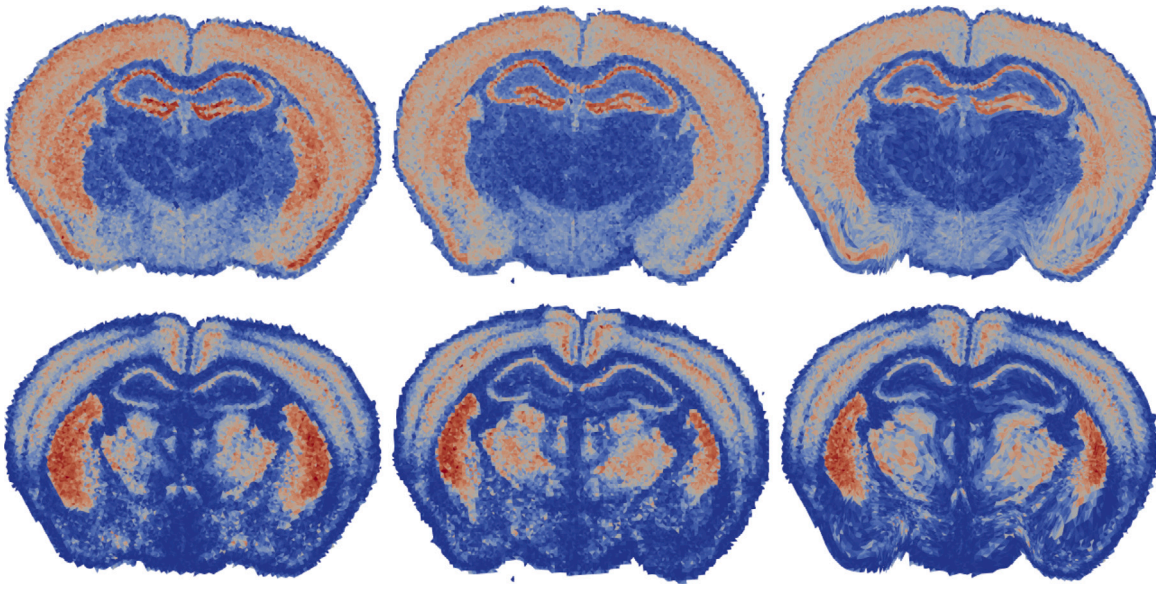

Fig. 5. MERFISH images (same mouse, nearby slices): Top row, from left to right: template with gene provided by Atp6ap1l gene counts, target of a second mouse with the same gene, template mapped based on a total of 10 genes of highest variance to the second mouse target. Second row provides the same information with Atp6ap1l replaced by Satb2. *Source:* Data from Zeng (2022).

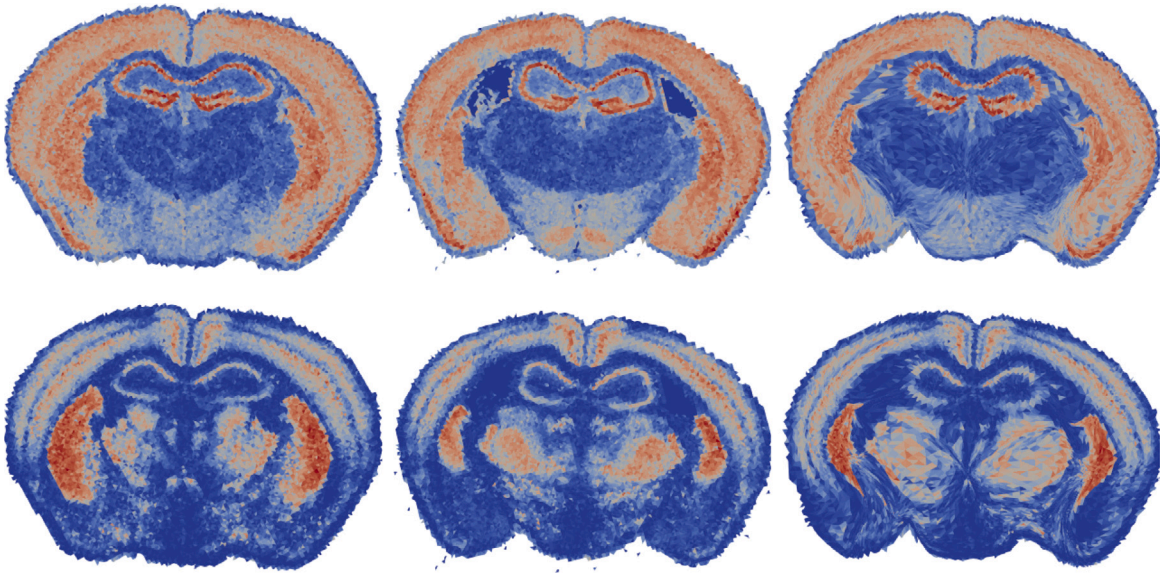

Fig. 6. MERFISH images (different mice, visually matched slices): Top row, from left to right: template with gene provided by Atp6ap1l gene counts, target of a second mouse with the same gene, template mapped based on a total of 10 genes of highest variance to the second mouse target. Second row provides the same information with Atp6ap1l replaced by Satb2. The first column (template) is the same as in Fig. 5.
Source: Data taken from Zeng (2022).

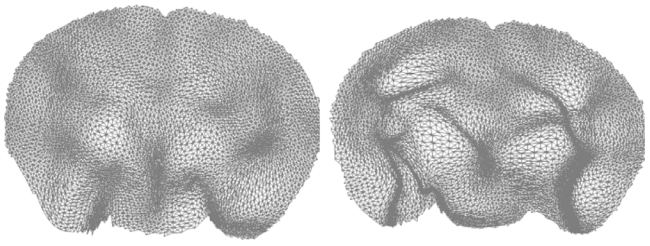

Fig. 7. Deformation grids for the two mappings shown above, Figures Fig. 5, Fig. 6. *Source:* Data from Zeng (2022).

4.3. Using velocity-based image matching

We here illustrate the discussion made Section 2.2.4 on the definition of a method akin (but not equivalent) to SFM-LDDMM, based on the L^2 comparison of smoothed measures. In this context, for finite feature sets, Eq. (22) is used to generate images from measures and the algorithm developed in Beg et al. (2005) can then be used to minimize, given two measures μ_0 and μ_1 (and using the notation in Eq. (22))

$$\int_0^1 \|v(t)\|_V^2 \, dt + \frac{1}{\sigma^2} \int_{\mathbb{R}^d} |q_{\mu_0}(\varphi^{-1}(1,y)) - q_{\mu_1}(y)|^2 \, dy. \tag{35}$$

with $\partial_t \varphi = v \circ \varphi$. We emphasize here the fact that, in order to use the correct action of diffeomorphisms on functions, one should use $q_{\varphi \cdot \mu_0}$ rather than $q_{\mu_0} \circ \varphi^{-1}$ in the end cost, and that the interpretation of

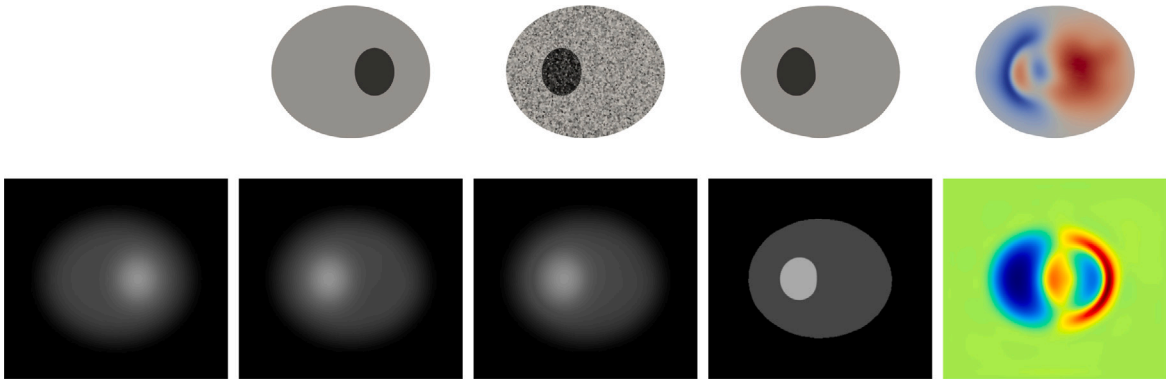

Fig. 8. Comparison between measure matching and image matching. First row is measure matching with, from left to right: Template image, noisy target image, deformed template, log Jacobian determinant of the transformation. Second row is LDDMM velocity-based matching with, from left to right: blurred template and target images used as inputs of registration, deformed blurred template, deformation applied to the original unblurred template, log Jacobian determinant.

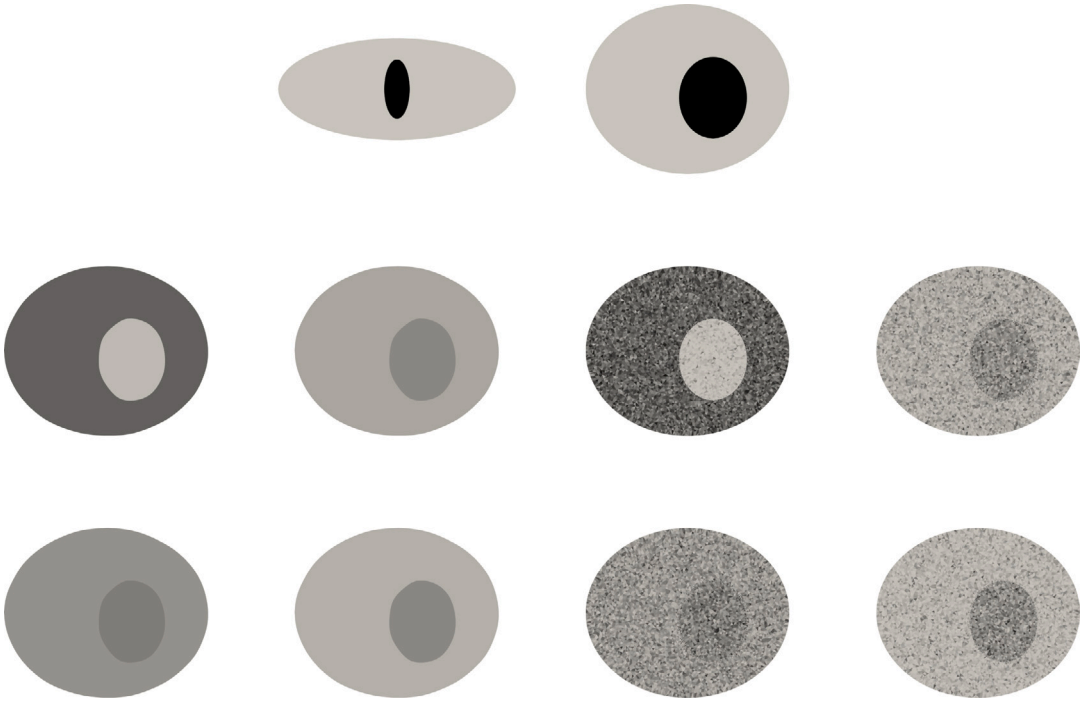

Fig. 9. Atlas mapping: First row: Template (observed) and target (unobserved) labels. Second row: first and second feature probabilities for the deformed template (left, estimated) and the observed target (right, observed). Third row: same information for the third and fourth feature probabilities.

$q_{\mu_0}\circ\varphi^{-1}$ in terms of tissue transformation is unclear. However, Eq. (35) can be directly implemented using one of several existing implementations of the LDDMM algorithm for image matching. The results are presented in Fig. 8. In this experiment, both methods provide good quality registration, but using different transformations as illustrated by the computed Jacobian determinant. Both methods used the same kernel for $\|\cdot\|_V$, up to discretization.

4.4. Atlas mapping

We present a synthetic illustration of the approach defined in Section 2.3 with a binary template with two labels (say, 1 and 2), provided as the left image in the first row of Fig. 9 and a target with four classes, generated using a ground truth transition matrix

$$\Theta_0 = \begin{pmatrix}\theta_{01}\\ \theta_{02}\end{pmatrix} = \begin{pmatrix}0.05 & 0.3 & 0.35 & 0.3\\ 0.5 & 0.15 & 0.25 & 0.1\end{pmatrix}.$$

The target is generated starting from the right image in the first row of Fig. 9, whose cells are labeled 1 or 2. Using the notation of Algorithm 1, the observable probability $\zeta'_{c'}$ in cell c' of the target is then sampled using a Dirichlet distribution with parameters $\sigma\theta_{0k}$, where $k = 1$ or 2 is the true label of cell c'. We then use Algorithm 1 to estimate Θ_0 and register the target to the atlas (without knowing the ground-truth labels of the target). Results are provided in Fig. 9, which shows the four estimated feature probabilities in the deformed template compared to target observed probabilities. In the application of Algorithm 1, we used $\alpha_c^{min} = \alpha_c^{max}$ independent of c and equal to the average value of $\alpha'_{c'}$ in the target image.

The algorithm also returns an estimate of Θ_0, given by

$$\hat{\Theta}_0 = \begin{pmatrix}0.051 & 0.297 & 0.351 & 0.300\\ 0.500 & 0.151 & 0.249 & 0.100\end{pmatrix},$$

showing excellent accuracy compared to the ground truth.

5. Discussion

The family of algorithms presented here provides the basis for future mapping technologies that allow for the representation of massive lists

of molecular and cellular descriptions of the human body with the tissue scales of radiological and pathological imaging. Unifying the molecular and image scales represents an important step forward in brain mapping. The central representation is the brain as a measure defined on the direct product of space and function.

The algorithms described allow for the molecular computational anatomy mapping program to continue in the vein of D'Arcy Thompson, computing normed distances between brains. A basic principle calculates similarity by acting diffeomorphisms which transforms one brain onto the other measuring the size of the transformation. Central to the theory proposed here is the action preserving the density of the quantized objects as space is transformed. This emphasizes the representation as containing two objects, the density ρ on $\mathbb{R}^d$ and the field of conditional distributions $(\zeta_x, x \in \mathbb{R}^d)$ representing function over space. The measures norms introduced for placing the measures of brain space into a normed-space score both the density measure as well as function measure.

The space-feature-measure brainspace represents both space and function. Because the transformations defined act on space, the variation of the norm with respect to the group action becomes the variation of space through the measure space kernel weighted by the direct inner product measuring alignment of the features.

Interestingly the measure action we derive makes the molecular scale representation mathematically consistent with the tissue scale representation associated to MRI imaging and atlasing at 100 micron–1 millimeter scale. We explicitly define several features including RNA and cell-centered features. In all the cases the features are represented as empirical probability laws over the RNA or cell identity feature spaces.

As part of the atlasing method we examine several algorithms for transferring the high-resolution gene features to the atlas tissue scales by inferring transition from labels to features. We demonstrate that this carries us into a family of quadratic programming problems in which the imputed feature laws are constrained to be probability measures.

We also examine the family of optimal test statistics for the spatial transcriptomic setting and show that the Kullback–Leibler divergence plays a central role in characterizing discriminability. The KL-distance is calculated between the empirical feature laws $\zeta_x, x \in \mathbb{R}^d$ under different hypotheses for the brain measures.

Declaration of competing interest

Michael Miller owns a founder share of Anatomy Works with the arrangement being managed by Johns Hopkins University in accordance with its conflict of interest policies. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data availability

Software is publicly available on bitbucket. The authors do not own the data that was made available to them.

Acknowledgments

Authors would like to acknowledge the Allen Institute for their support via the data contribution.

This work was supported by the National Institutes of Health (NIH), United States grants R01EB020062, R01NS102670, U19AG033655, P41-EB031771, and R01MH105660; the National Science Foundation (NSF), United States grant 2309683, and 16–569 NeuroNex contract, United States 1707298; and the Computational Anatomy Science Gateway as part of the Extreme Science and Engineering Discovery Environment, United States (XSEDE Towns et al. 2014), which is supported by the National Science Foundation, United States grant ACI1548562, and the Kavli Neuroscience Discovery Institute supported by the Kavli Foundation , United States.

Code availability

Programs used to generate experiments in this paper are available in the *py-lddmm* repository (Younes, 0000).

References

Allassonnière, S., Trouvé, A., Younes, Laurent, 2005. Geodesic shooting and diffeomorphic matching via textured meshes. In: EMMCVPR 2005. Springer, pp. 365–381.

Almgren, Frederick J., 1966. Plateau's Problem: An Invitation to Varifold Geometry, Vol. 13. American Mathematical Soc..

Aronszajn, Nachman, 1950. Theory of reproducing kernels. Trans. Amer. Math. Soc. 68 (3), 337–404.

Ashburner, J., 2009. Computational anatomy with the SPM software. Magn. Reson. Imaging 27, 1163–1174.

Avants, Brian B, Epstein, Charles L, Grossman, Murray, Gee, James C, 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. 12 (1), 26–41.

Avants, Brian, Gee, James C., 2004. Geodesic estimation for large deformation anatomical shape averaging and interpolation. Neuroimage 23, S139–S150.

Bauer, Martin, Charon, Nicolas, Younes, Laurent, 2019. Metric registration of curves and surfaces using optimal control. In: Handbook of Numerical Analysis, Vol. 20. Elsevier, pp. 613–646.

Beg, Mirza Faisal, Khan, Ali, 2007. Symmetric data attachment terms for large deformation image registration. IEEE Trans. Med. Imaging 26 (9), 1179–1189.

Beg, M. Faisal, Miller, Michael I., Trouvé, Alain, Younes, Laurent, 2005. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. Int. J. Comput. Vis. (ISSN: 09205691) 61 (2), 139–157. http://dx.doi.org/10.1023/B:VISI.0000043755.93987.aa.

Bugeon, Stephane, Duffield, Joshua, Dipoppa, Mario, Ritoux, Anne, Prankerd, Isabelle, Nicolout-sopoulos, Dimitris, Orme, David, Shinn, Maxwell, Peng, Han, Forrest, Hamish, Viduolyte, Aiste, Reddy, Charu Bai, Isogai, Yoh, Carandini, Matteo, Harris, Kenneth D., 2021. A transcriptomic axis predicts state modulation of cortical interneurons. bioRxiv http://dx.doi.org/10.1101/2021.10.24.465600, 2021.10.24.465600.

Charlier, Benjamin, Charon, Nicolas, Trouvé, Alain, 2017. The fshape framework for the variability analysis of functional shapes. Found. Comput. Math. 17, 287–357.

Charlier, Benjamin, Feydy, Jean, Glaunès, Joan Alexis, Collin, François-David, Durif, Ghislain, 2021. Kernel operations on the gpu, with autodiff, without memory overflows. J. Mach. Learn. Res. 22 (1), 3457–3462.

Charon, N., Trouvé, A., 2013. The varifold representation of nonoriented shapes for diffeomorphic registration. SIAM J. Imaging Sci. 6 (4), 2547–2580. http://dx.doi.org/10.1137/130918885.

Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S., Zhuang, X., 2015. RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. Science 348 (6233), aaa6090. http://dx.doi.org/10.1126/science.aaa6090, ISSN 1095-9203 (Electronic) 0036-8075 (Linking).

Chen, X., Sun, Y.C., Zhan, H., Kebschull, J.M., Fischer, S., Matho, K., Huang, Z.J., Gillis, J., Zador, A.M., 2019. High-throughput mapping of long-range neuronal projection using in situ sequencing. Cell 179 (3), 772–786 e19. http://dx.doi.org/10.1016/j.cell.2019.09.023, ISSN 1097-4172 (Electronic) 0092-8674 (Linking).

Cheney, Elliott Ward, Light, William Allan, 2009. A Course in Approximation Theory, Vol. 101. American Mathematical Soc..

Coifman, Ronald, Rokhlin, Vladimir, Wandzura, Stephen, 1993. The fast multipole method for the wave equation: A pedestrian prescription. IEEE Antennas Propag. Mag. 35 (3), 7–12.

Condylis, C., Ghanbari, A., Manjrekar, N., Bistrong, K., Yao, S., Yao, Z., Nguyen, T.N., Zeng, H., Tasic, B., Chen, J.L., 2022. Dense functional and molecular readout of a circuit hub in sensory cortex. Science 375 (6576), eabl5981. http://dx.doi.org/10.1126/science.abl5981, ISSN 1095-9203 (Electronic) 0036-8075 (Linking).

Dong, Hong Wei, 2008. The Allen Reference Atlas: A Digital Color Brain Atlas of the C57Bl/6J Male Mouse. John Wiley and Sons Inc, ISBN: 0470054085.

Durrleman, Stanley, Pennec, Xavier, Trouvé, Alain, Ayache, Nicholas, 2008a. Sparse approximation of currents for statistics on curves and surfaces. Med. Image Comput. Comput.-Assist. Interv. 11 (Pt 2), 390–398.

Durrleman, Stanley, Pennec, Xavier, Trouvé, Alain, Thompson, Paul, Ayache, Nicholas, 2008b. Inferring brain variability from diffeomorphic deformations of currents: an integrative approach.. Med. Image Anal. 12 (5), 626–637.

Ecker, Joseph R., Geschwind, Daniel H., Kriegstein, Arnold R., Ngai, John, Osten, Pavel, Polioudakis, Damon, Regev, Aviv, Sestan, Nenad, Wickersham, Ian R., Zeng, Hongkui, 2017. The BRAIN initiative cell census consortium: Lessons learned toward generating a comprehensive brain cell atlas. Neuron (ISSN: 0896-6273) 96 (3), 542–557. http://dx.doi.org/10.1016/j.neuron.2017.10.007.

Essmann, Ulrich, Perera, Lalith, Berkowitz, Max L, Darden, Tom, Lee, Hsing, Pedersen, Lee G, 1995. A smooth particle mesh Ewald method. J. Chem. Phys. 103 (19), 8577–8593, Publisher: American Institute of Physics.

Glaunès, Joan, Qiu, Anqi, Miller, Michael I., Younes, Laurent, 2008. Large deformation diffeomorphic metric curve matching. Int. J. Comput. Vis. 80 (3), 317–336.

Glaunès, Joan, Trouvé, Alain, Younes, Laurent, 2004a. Diffeomorphic matching of distributions: A new approach for unlabelled point-sets and sub-manifolds matching. In: Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, Vol. 2. Ieee, pp. II–712.

Glaunès, J., Vaillant, M., Miller, M.I., 2004b. Landmark matching via large deformation diffeomorphisms on the sphere. J. Math. Imaging Vision (20), 179–200.

Greengard, Leslie, Rokhlin, Vladimir, 1987. A fast algorithm for particle simulations. J. Comput. Phys. 73 (2), 325–348.

Greengard, Leslie, Strain, John, 1991. The fast Gauss transform. SIAM J. Sci. Stat. Comput. 12 (1), 79–94.

Greengard, Leslie, Sun, Xiaobai, 1998. A new version of the fast Gauss transform. Doc. Math. 3, 575–584.

Grenander, Ulf, Miller, Michael I., 1998. Computational anatomy: An emerging discipline. Quart. Appl. Math. (ISSN: 0033569X) 56 (4), 617–694.

Helmstaedter, M., 2013. Cellular-resolution connectomics: challenges of dense neural circuit reconstruction. Nature Methods 10 (6), 501–507. http://dx.doi.org/10.1038/nmeth.2476, ISSN 1548-7105 (Electronic) 1548-7091 (Linking).

Iske, Armin, 2018. Approximation Theory and Algorithms for Data Analysis. Springer.

Jain, Aastha, 2011. Particle Methods for Diffeomorphic Registration (Ph.D. thesis). Johns Hopkins University.

Jain, Aastha, Younes, Laurent, 2013. A kernel class allowing for fast computations in shape spaces induced by diffeomorphisms. J. Comput. Appl. Math. 245, 162–181, Publisher: Elsevier.

Joshi, S., Miller, M.I., 2000a. Landmark matching via large deformation diffeomorphisms. IEEE Trans. Image Process. 9 (8), 1357–1370.

Joshi, S.C., Miller, M.I., 2000b. Landmark matching via large deformation diffeomorphisms. IEEE Trans. Image Process. 9 (8), 1357–1370. http://dx.doi.org/10.1109/83.855431, ISSN.

Kaltenmark, Irene, Charlier, Benjamin, Charon, Nicolas, 2017. A general framework for curve and surface comparison and registration with oriented varifolds. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3346–3355.

Kebschull, J.M., Richman, E.B., Ringach, N., Friedmann, D., Albarran, E., Kolluru, S.S., Jones, R.C., Allen, W.E., Wang, Y., Cho, S.W., Zhou, H., Ding, J.B., Chang, H.Y., Deisseroth, K., Quake, S.R., Luo, L., 2020. Cerebellar nuclei evolved by repeatedly duplicating a conserved cell-type set. Science 370 (6523), http://dx.doi.org/10.1126/science.abd5059, ISSN 1095-9203 (Electronic) 0036-8075 (Linking).

Kebschull, J.M., Garcia da Silva, P., Reid, A.P., Peikon, I.D., Albeanu, D.F., Zador, A.M., 2016. High-throughput mapping of single-neuron projections by sequencing of barcoded RNA. Neuron 91 (5), 975–987. http://dx.doi.org/10.1016/j.neuron.2016.07.036, ISSN 1097-4199 (Electronic) 0896-6273 (Linking).

Kim, Y., Venkataraju, K.U., Pradhan, K., Mende, C., Taranda, J., Turaga, S.C., Arganda-Carreras, I., Ng, L., Hawrylycz, M.J., Rockland, K.S., Seung, H.S., Osten, P., 2015. Mapping social behavior-induced brain activation at cellular resolution in the mouse. Cell Rep. 10 (2), 292–305. http://dx.doi.org/10.1016/j.celrep.2014.12.014, ISSN 2211-1247 (Electronic).

Kim, Y., Yang, G.R., Pradhan, K., Venkataraju, K.U., Bota, M., Garcia Del Molino, L.C., Fitzgerald, G., Ram, K., He, M., Levine, J.M., Mitra, P., Huang, Z.J., Wang, X.J., Osten, P., 2017. Brain-wide maps reveal stereotyped cell-type-based cortical architecture and subcortical sexual dimorphism. Cell 171 (2), 456–469 e22. http://dx.doi.org/10.1016/j.cell.2017.09.020, ISSN 1097-4172 (Electronic) 0092-8674 (Linking).

Ma, J., Miller, M.I., Younes, L., 2010. A Bayesian generative model for surface template estimation. J. Biomed. Imaging 2010 (16).

Miller, M.I., D.J., Tward, A., Trouve, 2021. Hierarchical computational anatomy: Unifying the molecular to tissue continuum via measure rperesentations of the brain. bioRxiv.

Miller, M.I., Trouvé, A., Younes, L., 2006. Geodesic shooting for computational anatomy. J. Math. Imaging Vis. 24, 209–228.

Miller, Michael I., Trouvé, Alain, Younes, Laurent, 2015. Hamiltonian systems and optimal control in computational anatomy: 100 years since D'Arcy Thompson. Annu. Rev. Biomed Eng. 17, 447–509.

Miller, M.I., Tward, D.J., Trouvé, A., 2020. Coarse-to-fine Hamiltonian dynamics of hierarchical flows in computational anatomy. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 3760–3765. http://dx.doi.org/10.1109/CVPRW50498.2020.00438.

Miller, Michael, Tward, Daniel, Trouve, Alain, 2022. Molecular computational anatomy: Unifying the molecular to tissue continuum via measure represention of the brain. Sci. Partn. J..

Moffitt, Jeffrey R., Bambah-Mukku, Dhananjay, Eichhorn, Stephen W., Vaughn, Eric, Shekhar, Karthik, Perez, Julio D., Rubinstein, Nimrod D., Hao, Junjie, Regev, Aviv, Dulac, Catherine, Zhuang, Xiaowei, 2018. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. Science (ISSN: 0036-8075) 362 (6416), http://dx.doi.org/10.1126/science.aau5324.

Mumford, David Bryant, Shah, Jayant, 1989. Optimal approximations by piecewise smooth functions and associated variational problems. Commun. Pure Appl. Math..

Narasimhan, Arun, Umadevi Venkataraju, Kannan, Mizrachi, Judith, Albeanu, Dinu F., Osten, Pavel, 2017. Oblique light-sheet tomography: fast and high resolution volumetric imaging of mouse brains. bioRxiv http://dx.doi.org/10.1101/132423.

Nocedal, Jorge, Wright, Stephen, 2006. Numerical Optimization. Springer Science & Business Media.

Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., Mortrud, M.T., Ouellette, B., Nguyen, T.N., Sorensen, S.A., Slaughterbeck, C.R., Wakeman, W., Li, Y., Feng, D., Ho, A., Nicholas, E., Hirokawa, K.E., Bohn, P., Joines, K.M., Peng, H., Hawrylycz, M.J., Phillips, J.W., Hohmann, J.G., Wohnoutka, P., Gerfen, C.R., Koch, C., Bernard, A., Dang, C., Jones, A.R., Zeng, H., 2014. A mesoscale connectome of the mouse brain. Nature 508 (7495), 207–214. http://dx.doi.org/10.1038/nature13186, ISSN 1476-4687 (Electronic) 0028-0836 (Linking).

Osten, P., Margrie, T.W., 2013. Mapping brain circuitry with a light microscope. Nature Methods 10 (6), 515–523. http://dx.doi.org/10.1038/nmeth.2477, ISSN 1548-7105 (Electronic) 1548-7091 (Linking).

Pennec, X., 2011. From Riemannian geometry to computational anatomy. Elements.

Ragan, T., Kadiri, L.R., Venkataraju, K.U., Bahlmann, K., Sutin, J., Taranda, J., Arganda-Carreras, I., Kim, Y., Seung, H.S., Osten, P., 2012. Serial two-photon tomography for automated ex vivo mouse brain imaging. Nature Methods 9 (3), 255–258, ISSN 1548-7105 (Electronic) 1548-7091 (Linking).

Schaback, Robert, Wendland, Holger, 2006. Kernel techniques: from machine learning to meshless methods. Acta Numer. 15, 543–639.

Shah, Sheel, Lubeck, Eric, Zhou, Wen, Cai, Long, 2016. In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. Neuron 92, 342–357. http://dx.doi.org/10.1016/j.neuron.2016.10.001.

Stouffer, Kaitlin M, Wang, Zhenzhen, Xu, Eileen, Lee, Karl, Lee, Paige, Miller, Michael I., Tward, Daniel J, 2021. From picoscale pathology to decascale disease: Image registration with a scattering transform and varifolds for manipulating multiscale data. In: International Workshop on Multimodal Learning for Clinical Decision Support. Springer, pp. 1–11.

Sun, Y.C., Chen, X., Fischer, S., Lu, S., Zhan, H., Gillis, J., Zador, A.M., 2021. Integrating barcoded neuroanatomy with spatial transcriptional profiling enables identification of gene correlates of projections. Nat. Neurosci. 24 (6), 873–885. http://dx.doi.org/10.1038/s41593-021-00842-4, ISSN 1546-1726 (Electronic) 1097-6256 (Linking).

Trouvé, Alain, Vialard, François-Xavier, 2010. Shape splines and stochastic shape evolutions: A second order point of view. Quart. Appl. Math. 26.

Vaillant, Marc, Glaunès, Joan, 2005. Surface matching via currents. In: Proceedings of Information Processing in Medical Imaging (IPMI 2005). In: Lecture Notes in Computer Science, vol. 3565, pp. 381–392.

Vaillant, M., Miller, M.I., Younes, L., Trouvé, A., 2004. Statistics on diffeomorphisms via tangent space representation. NeuroImage 23, 161–169.

Vialard, F.-X., Risser, L., Rueckert, D., Cotter, C.J., 2012. 3D image registration via geodesic shooting using and efficient adjoint calculation. J. Int. J. Comput. Vis. 97 (2), 229–241.

Wang*, Xiao, Allen*, WE, Wright, M, Sylwestrak, E, Samusik, N, Vesuna, S, Evans, K, Liu, C, Ramakrishnan, C, Liu, J, Nolan, G P, Bava, F-A, Deisseroth, K, 2018. Three-dimensional intact-tissue-sequencing of single-cell transcriptional states. Science http://dx.doi.org/10.1126/science.aat5691.

Wang, Quanxin, Ding, Song-Lin, Li, Yang, Royall, Josh, Feng, David, Lesnar, Phil, Graddis, Nile, Naeemi, Maitham, Facer, Benjamin, Ho, Anh, et al., 2020. The Allen mouse brain common coordinate framework: a 3D reference atlas. Cell 181 (4), 936–953.

Xia, Chenglong, Fan, Jean, Emanuel, George, Hao, Junjie, Zhuang, Xiaowei, 2019. Spatial transcriptome profiling by MERFISH reveals subcellular rna compartmentalization and cell cycle-dependent gene expression. Proc. Natl. Acad. Sci. (ISSN: 0027-8424) 116 (39), 19490–19499. http://dx.doi.org/10.1073/pnas.1912459116.

Xiaowei, A., 2021. Method of the year 2020: spatially resolved transcriptomics. Nature Methods 18 (1).

Xu, S., Yang, H., Menon, V., Lemire, A.L., Wang, L., Henry, F.E., Turaga, S.C., Sternson, S.M., 2020. Behavioral state coding by molecularly defined paraventricular hypothalamic cell type ensembles. Science 370 (6514), http://dx.doi.org/10.1126/science.abb2494, ISSN 1095-9203 (Electronic) 0036-8075 (Linking).

Younes, Laurent, Py-LDDMM (Bitbucket repository), lyounes@bitbucket.org/laurent_younes/py-lddmm.git.

Younes, L., 2007. Jacobi fields in groups of diffeomorphisms and applications. Quart. Appl. Math. 65, 113–134.

Younes, Laurent, 2019. Shapes and Diffeomorphisms, second ed. In: Applied Mathematical Sciences, Springer-Verlag, Berlin Heidelberg, ISBN: 978-3-662-58495-8, http://dx.doi.org/10.1007/978-3-662-58496-5.

Zeira, Ron, Land, Max, Raphael, Benjamin J., 2021. Alignment and integration of spatial transcriptomics data. bioRxiv http://dx.doi.org/10.1101/2021.03.16.435604, 2021.03.16.435604.

Zeng, Hongkui, 2022. Unpublished data.

Zheng, Ting, Yang, Zhongqing, Li, Anan, Lv, Xiaohua, Zhou, Zhenqiao, Wang, Xiaojun, Qi, Xiaoli, Li, Shiwei, Luo, Qingming, Gong, Hui, 2013. Visualization of brain circuits using two-photon fluorescence micro-optical sectioning tomography. Opt. Express (ISSN: 1094-4087) 21 (8), 9839–9850.

Zingg, B., Hintiryan, H., Gou, L., Song, M.Y., Bay, M., Bienkowski, M.S., Foster, N.N., Yamashita, S., Bowman, I., Toga, A.W., Dong, H.W., 2014. Neural networks of the mouse neocortex. Cell 156 (5), 1096–1111. http://dx.doi.org/10.1016/j.cell.2014.02.023, ISSN 1097-4172 (Electronic) 0092-8674 (Linking).