HIERARCHICAL COMPUTATIONAL ANATOMY: UNIFYING THE MOLECULAR TO TISSUE CONTINUUM VIA MEASURE REPRESENTATIONS OF THE BRAIN

MICHAEL I. MILLER, DANIEL TWARD, AND ALAIN TROUVÉ

ABSTRACT. This paper presents a unified representation of the molecular and cellular scale descriptions of the brain with the continuum tissue scales based on a layered hierarchy of mathematical functional measures. We present a fine-to-coarse recipe for scale-space aggregation projecting functional descriptions into stable empirical probability laws that links the layers in the hierarchy. Brainspace is constructed as a metric space with metric mapping between brains provided by a hierarchy of layered Hamiltonian geodesic flows of diffeomorphisms connecting the scales. The diffeomorphisms act on the brain measures via the 3D varifold action representing "copy and paste" so that basic particle quantities that are conserved biologically are combined with greater multiplicity and not geometrically distorted. The representation unifies the classical formalism of computational anatomy for representing continuum tissue scale with non-classical generalized functions appropriate for molecular particle scales. Two applications are examined, the first histological and tissue scale data in the human brain for studying Alzheimer's disease, and the second the RNA and cell signatures of dense spatial transcriptomics mapped to the meso-scales of brain atlases.

1. INTRODUCTION

One of the striking aspects of the study of the brain in modern neurobiology is the fact that the distributions of discrete structures that make up physical tissue, from neural cells to synapses to genes and molecules, exists across nearly ten orders of magnitude in spatial scale. This paper focusses on the challenge of building multi-scale representations that simultaneously connect the quantum nano-scales of modern molecular biology for characterizing neural circuits architecture in the brain with the classical continuum representations at the anatomical gross and meso scales.

We have been highly motivated by the Cell Census Network project (BICCN [11]) which brings the nano and micron scales of single cell measures of RNA via spatial transcriptomics [44, 9, 56, 38, 41] coupled to the tissue scales of mouse atlases. The recent review on bridging scales from cells to physiology [36] motivates the mathematical framework presented herein. The recent emergence of spatial transcriptomics as method of the year [59] highlights the importance and ascendence of such approaches for understanding the dense metric structure of the brain which represent the coarse physiological atlas scales buit up from dense imaging measurements at the cellular scales.

Specifically, in our work in Alzheimer's disease the BIOCARD study [42] we are examining pathological Tau at both the micro histological and macro atlas scales of Tau particle detections, from 10-100 μm [53] and to human magnetic resonance millimeter scales for examining entire circuits in the medial temporal lobe [61]. In the mouse cell counting project we are examining single-cell spatial transcriptomics using modern RNA sequencing in dense tissue at the micron scale and its representations in the Allen atlas coordinates [54, 55].

Most noteworthy for any representation is that at the finest micro scales nothing is smooth; the distributions of cells and molecules are more well described as random quantum counting processes in space [45]. In contrast, information associated to atlasing methods at gross anatomical tissue and organ scales extend

Date: July 17, 2021.

smoothly [51, 13, 27, 37, 25, 26, 35, 10, 50, 12]. Cross-sectionally and even cross-species, gross anatomical labelling is largely repeatable, implying information transfers and changes from one coordinate system to another smoothly. This is built into the representation theory of diffeomorphisms and soft matter tissue models for which advection and transport hold [21, 52, 1, 5, 39, 32, 30], principles upon which continuum mechanics and its analogues are based, but not directly applicable for particles.

The focus of this paper is to build a coherent representation theory across scales. For this we view the micron to millimeter scales via the same representation theory called *mathematical measures*, building the finest micron scales from discrete units termed particle measures which represent molecules, synapses and cells. As they aggregate they form tissues. This measure representation allows us to understand subsets of tissues that contain discretely positioned and placed functional objects at the finest quantized scales and simultaneously pass smoothly to the classical continuum scales at which stable functional and anatomical representations exist. Since the study of the function of the brain on its geometric submanifolds -the gyri, sulci, subnuclei and laminae of cortex- are so important, we extend our general framework to exploit varifold measures [15, 3, 17] arising in the modern discipline of geometric measure theory. To be able to compare the brains we use diffeomorphisms as the comparator tool, with their action representing 3D varifold action which we formulate as "copy and paste" so that basic particle quantities that are conserved biologically are combined with greater multiplicity and not geometrically distorted as would be the case for measure transport.

The functional features are represented via singular delta-diracs at the finest micro structure scales. The functional feature is abstracted into a function space rich enough to accomodate the molecular machinery as represented by RNA or Tau particles, as well as electrophysiology associated to spiking neurons, or at the tissue scales of medical imaging dense contrasts of magnetic resonance images (MRIs). We pass to the classical function continuum via introduction of a scale-space that extends the descriptions of cortical micro-circuits to the meso and anatomical scales. This passage from the quantized features to the stochastic laws is in fact akin to the Boltzman program transferring the view from the Newtonian particles to the stable distributions describing them. For this we introduce a scale-space of kernel density transformations which allows us to retrieve the empirical averages represented by the determinism of the stochastic law consistent with our views of the macro tissue scales.

The representation provides a recipe for scale traversal in terms of a cascade of linear space scaling composed with non-linear functional feature mapping. Following the cascade implies every scale is a measure so that a universal family of measure norms can be introduced which simultaneously measure the disparety between brains in the orbit independent of the probing technology, RNA identities, Tau or amyloid histology, spike trains, or dense MR imagery.

Our brain measure model implies the existence of a sequence . This scale-space of pairs, the measure representation of the brain and the associated probing measurement technologies we call Brainspace. To formulate a consistent measurement and comparison technology on Brainspace we construct a natural metric upon it allowing us to study its geometry and connectedness. The metric between brains is constructed via a Hamiltonian which defines the geodesic connections throughout scale space, providing for the first time a hierarchical representation that unifies micro to millimeter representation in the brain and makes Brainspace into a metric space. Examples of representation and comparision are given for Alzheimer's histology integrated to magnetic resonance imaging scales, and spatial transcriptomics.

2. Results

2.1. **Measure Model of Brain Structures.** To build a coherent theory we view the micron to anatomical scales via the same representation theory building upon discrete units termed particles or atoms. As they aggregate they form tissues. This is depicted in Figure 1 in which the left panel shows mouse imaging of

CUX1 labelling of the inner layers of mouse cortex (white) and CTP2 imaging of the outer layers (green) at 2.5 micron in plane resolution. Notice the discrete nature of the cells clearly resolved which form the layers of tissue which are the global macro scale features of layer 2,3,4 which stain more prolificaly in white and the outer layers 5,6 which stain more prolifically in green.

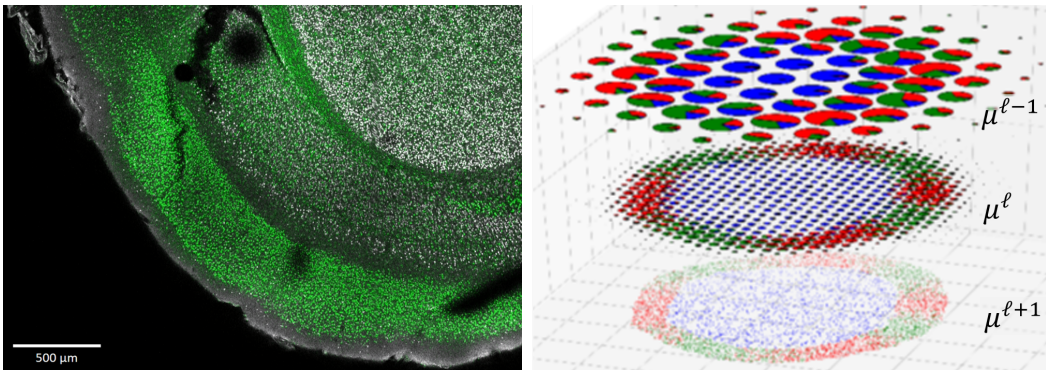

FIGURE 1. Left: Tissue from a NexCre+/-;Brn2fl/+ adult mouse mimicing a wild-type mouse with CUX1 labelling of layer ii/iii,iv and CTIP2 in layers v,vi in green. Shows sections at $2.5^2 \times 50\,\mu\text{m}^3$ 6 tile images, 1433×1973 pixels; taken from Uli Mueller. Right: Showing the abstraction of a coarse-to-fine hierarchy $\mu^{\ell-1}, \mu^{\ell}, \mu^{\ell+1}$ with fine molecular scales shown at the bottom with colors depicting $\mathcal{F}$ function ascending scales.

Our representation exists simultaneously at both the micro and tissue millimeter scales. A key aspect of anatomy is that at a micro or nano scale, information is encoded as a huge collection of pairs (x_i, f_i) where $x_i \in \mathbb{R}^d$ $(d = 2, 3)$ describes the position of a "particle" and f_i is a functional state in a given set $\mathcal{F}$ attached to it (protein or RNA signature or Tau tangle, or for single cell Neurophysiology the dynamics of neural spiking). Basically everything is deterministic, with every particle attached to its own functional state among possible functional state in $\mathcal{F}$. But zooming out, the tissue level, say mm scale, appears through the statistical distribution of its constituents with two key quantities, *the local density* of particles ρ and the *conditional probability distribution* of the functional features $\mu_x(df)$ at any location x. At position x, we no longer have a deterministic functional state but a probability distribution μ_x on functional states.

The integration of both descriptions into a common mathematical framework can be done quite naturally in the setting of mathematical measures which are mathematical constructs that are able to represent both the discrete and continuous worlds as well as natural level of approximation between both. Indeed the set $\mathcal{M}(\mathbb{R}^d \times \mathcal{F})$ of finite positive measures on $\mathbb{R}^d \times \mathcal{F}$ contains discrete measures

$$\mu = \sum_{i \in I} w_i \delta_{x_i} \otimes \delta_{f_i} \ , \tag{1}$$

where w_i is a positive weight that can encode the collection (x_i, f_i) at micro scale.

As in Boltzmann modelling we describe the features statistically at a fixed spatial scale transferring our attention to their stochastic laws modelled as conditional probabilities in $\mathcal{M}_P(\mathcal{F})$ with integral 1. For this we factor the measures into the marginal ρ on $\mathbb{R}^d$ with $\rho(dx) = \int_{\mathcal{F}} \mu(dx, df)$, and the field of probability distributions on $\mathcal{F}$ conditioned on x :

$$\mu(dx, df) = \rho(dx)\, \mu_x(df) \ , \text{ on } \mathbb{R}^d \times \mathcal{F} \ , \tag{2a}$$

with field of conditionals

$$x \mapsto \mu_x \in \mathcal{M}_P(\mathcal{F}) \ . \tag{2b}$$

Continuous tissues we abstract as brain measures μ with marginal ρ having a tissue density $\rho(dx) = \rho_c(x)dx$ with respect to the Lebesgue measure on $\mathbb{R}^d$. A fundamental link between the molecular and continuum tissue can be addressed through the law of large number since if $(x_i, f_i)_{i \geqslant 0}$ is an independent and identically distributed sample drawn from law μ/M of $\mathbb{R}^d \times \mathcal{F}$ where $M = \int_{\mathbb{R}^d \times \mathcal{F}} \mu(dx, df)$ is the total mass of such μ, then we have almost surely the weak convergence

$$\mu_N := \frac{M}{N} \sum_{i=1}^{N} \delta_{x_i} \otimes \delta_{f_i} \to \mu \ .$$

Passing from the tissue scales to the finest molecular and cellular scales behooves us to introduce a scale-space so that empirical averages which govern it are repeatable. As depicted in the right panel of Figure 1, our multi-scale model of a brain is as a sequence of measures:

$$\boldsymbol{\mu} = (\mu^\ell)_{\ell=0,1,\ldots} \ . \tag{3}$$

Our idealization of Brainspace as a sequence of measures as depicted in Figure 1 descends from the the coarse tissue scale (top) to the finest particle representation (bottom), with color representing function $f \in \mathcal{F}$, and radius space-scale.Throughout the range of scales is denoted shorthand $\ell < \ell_{max}$ to mean $0 \leqslant \ell < \ell_{max}$ with lowest scale $\ell = 0$ and upper ℓ_{max} not attained.

2.2. **Model for Crossing Scales.** The brain being a multi-scale collection of measures requires us to be able to transform from one scale to another. We do this by associating a scale-space to each particle feature by pairing to each measure a kernel function transforming it from a generalized function to a classical function $\delta_z : h \mapsto h(z)$. The kernels carry a resolution scale σ or reciprocally a bandwidth, analogous to Planck's scale.

To define this we introduce the abstract representation of our system as a collection of descriptive elements $z \in \mathcal{Z}$ made from spatial and functional features. We transform our mathematical measure $\mu(\cdot)$ on $\mathcal{Z}$ generating new measures $\mu'(\cdot)$ on $\mathcal{Z}'$ by defining correspondences via kernels $z \mapsto k(z, dz')$, with the kernel acting on the particles $K[\delta_{z_i}](dz') = k(z_i, dz')$; the measures transform as

$$K[\mu](dz') \ = \int_{\mathcal{Z}} k(z, dz')\mu(dz) \ . \tag{4}$$

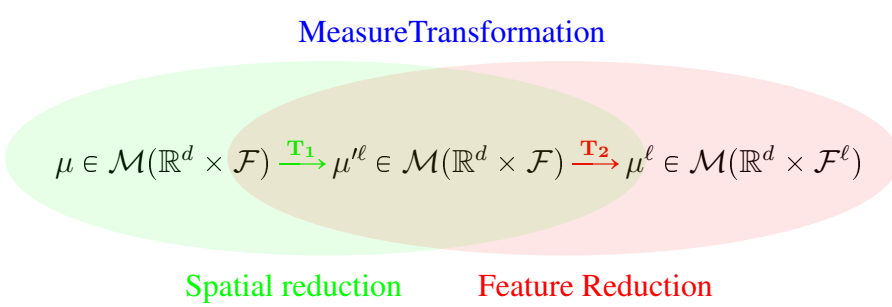

FIGURE 2. Space and function transformation shown as a composition $\mu \xrightarrow{T_1} \mu'^\ell \xrightarrow{T_2} \mu^\ell$.

Depicted in Figure 2 is the construction of our formalism for transformations of scale, linear spatial smoothing followed by non-linear transformation of the feature laws. The first operator transforms via a linear kernel $k_1((x,f),\cdot)$ leaving the feature space unchanged $\mathcal{Z} = \mathbb{R}^d \times \mathcal{F}$:

$$T_1 : \ \mu \mapsto \mu'^{\ell} = \int_{\mathbb{R}^d \times \mathcal{F}} k_1((x,f),\cdot)\, \mu(dx,df)\,. \tag{5a}$$

The second transforms nonlinearly to new features $\mathcal{F}^\ell$ via the kernel $k_2((x,\alpha),\cdot)$ transforming any feature probability $\alpha \in \mathcal{M}_P$. Decomposing $\mu'^{\ell}(dx,df) = \rho(dx)\mu'^{\ell}_x(df)$ into its marginal and conditional gives

$$T_2 : \ \mu'^{\ell} \mapsto \mu^{\ell} \ = \int_{\mathbb{R}^d} k_2((x,\mu'^{\ell}_x),\cdot)\, \rho(dx)\,. \tag{5b}$$

At every scale μ^ℓ remains a measure allowing the renormalization (5a), (5b) to be executed recursively.

We use computational lattices to interpolate between the multi-scale continuum and the discrete representations. The core operations resamples by projecting the particles $(x_i \in \mathbb{R}^d)_{i\in I}$ to the rescaling lattice $(y_j \in Y_j \subset \mathbb{R}^d)_{j\in J}$. The kernel k_1 reweights via the fraction particle x_i shares with site Y_j:

$$\pi(x_i,\cdot) : Y_j \mapsto \pi(x_i, Y_j), \text{with } \pi(x_i,\mathbb{R}^d) = 1\,.$$

The kernel k_2 resamples features via maps ϕ from machine learning:

$$\phi : \alpha \in \mathcal{M}_P(\mathcal{F}) \mapsto \phi(\alpha) \in \mathcal{F}^\ell\,.$$

This defines $T_1 : \mu \mapsto \mu'^{\ell}$ for space normalization and $T_2 : \mu'^{\ell} \mapsto \mu^\ell$ for feature reduction onto ϕ:

$$k_1((x,f),\cdot) = \sum\nolimits_{j\in J} \pi(x,Y_j)\, \delta_{y_j} \otimes \delta_f(\cdot)\,, \tag{6a}$$

$$\mu'^{\ell} \ = \sum\nolimits_{j\in J} w_j\, \delta_{y_j} \otimes \mu'^{\ell}_{y_j} \tag{6b}$$

$$\text{with} \quad \begin{cases} w_j = \sum_{i\in I} w_i\, \pi(x_i,Y_j) \\ \mu'^{\ell}_{y_j} = \frac{1}{w_j} \sum_{i\in I} w_i\, \pi(x_i,Y_j)\, \delta_{f_i} \\ \rho = \sum_{j\in J} w_j \delta_{y_j} \end{cases} .$$

$$k_2((x,\mu'^{\ell}_x),\cdot) = \delta_x \otimes \delta_{\phi(\mu'^{\ell}_x)}(\cdot)\,, \tag{6c}$$

$$\mu^\ell = \sum\nolimits_{j\in J} w_j\, \delta_{y_j} \otimes \delta_{\phi(\mu'^{\ell}_{y_j})} \tag{6d}$$

The indexing across scales is defined disjointly $I \cap J = \varnothing$. Methods 4.1 examines these derivations.

2.3. **The Dynamical Systems Model via Varifold Action of Diffeomorphisms.** We want to measure and cluster brains by building a metric space structure. We do this by following the original program of D'Arcy Thompson building bijective correspondence. In this setting this must be done at every scale with each scale having different numbers of particles and resolutions.

We build correspondence between sample brains via dense connections of the discrete particles to the continuum at all scales using the diffeomorphism group and diffeomorphic transport. For this define the group of k-times continuously differentiable diffeomorphisms $\varphi \in G_k$ with group operation function composition $\varphi \circ \varphi'$. For any brain $\mu = \sum_{i\in I} w_i \delta_{x_i} \otimes \delta_{f_i} \in \mathcal{M}$, the diffeomorphisms act

$$(\varphi,\mu) \mapsto \varphi \cdot \mu = \sum_{i\in I} w_i |d\varphi(x_i)| \delta_{\varphi(x_i)} \otimes \delta_{f_i}\,. \tag{7a}$$

Space scales are represented as the group product, $\boldsymbol{\varphi} = (\varphi^\ell)_{\ell<\ell_{max}} \in \boldsymbol{G}_k$, acting component-wise with action

$$\boldsymbol{\varphi} \cdot \boldsymbol{\mu} := (\varphi^\ell \cdot \mu^\ell)_{\ell<\ell_{max}} \ . \tag{7b}$$

The $|d\varphi(x)|$ term in the action enables the crucial property that when a tissue is extended to a larger area, the total number of its basic constituents increase accordingly and are not conserved, in contrast to classic measure or probability transport. We call this the "copy and paste" varifold action.

Dynamics occurs through the group action generated as a dynamical system in which we introduce time-indexed vector fields $t \mapsto u_t^\ell$ which act as controls on the diffeomorphisms:

$$\dot{\varphi}_t^\ell = u_t^\ell \circ \varphi_t^\ell \ . \tag{8}$$

As depicted in Figure 3 the controls are coupled across scales by successive refinements $v^\ell, \ell < \ell_{max}$:

$$u^\ell = u^{\ell-1} + v^\ell \ , \tag{9}$$

with our convention taking $u^{-1} = 0$.

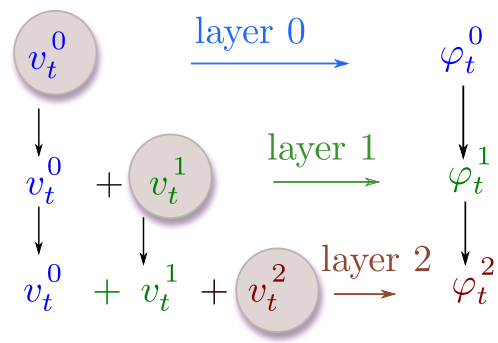

FIGURE 3. Hierarchical system, controls $u^\ell = u^{\ell-1} + v^\ell, u^0 = v^0$ and flows $\varphi^\ell, \ell = 0, 1, \ldots$.

The dynamical system model is an observer and dynamics equation:

$$\begin{cases} \boldsymbol{\varphi}_t \cdot \boldsymbol{\mu}, \ t \in [0,1] & \text{Observer} \\ \dot{\boldsymbol{\varphi}}_t = \boldsymbol{u}_t \circ \boldsymbol{\varphi}_t, \ \boldsymbol{\varphi}_0 = \mathbf{Id} & \text{Dynamics} \end{cases} . \tag{10}$$

Geodesic mapping flows a control process along paths of minimum energy respecting the boundary conditions. To optimize the kinetic energy define the hierarchy $\boldsymbol{u} := (u^\ell)_{\ell<\ell_{max}}$ with the total path $\boldsymbol{u}. := (\boldsymbol{u}_t, t \in [0,1])$. We model the controls to be smooth elements of reproducing kernel spaces (RKHS's) V_ℓ, norm $\|\cdot\|_{V_\ell}$ with $\boldsymbol{u}$ in the product $\boldsymbol{V} = \prod_{\ell<\ell_{max}} V_\ell$. The kernels are taken as diagonal $K^\ell(\cdot,\cdot) = g^\ell(\cdot,\cdot)\mathrm{id}_d$, id_d a $d \times d$ identity matrix (see [33] for non-diagonal kernels), and g^ℓ the Green's function. The green's kernel is induced by the differential operator defining the norm-square $\|L^\ell v^\ell\|_{V_\ell}^2$ with $L^\ell := ((1-(\alpha^\ell)^2\nabla^2)\mathrm{id}_d)^2$, the Laplacin weight determining the full width at half maximum (FWHM) of the Green's kernel.

Dynamics translates into a navigation in the orbit $\boldsymbol{G}_k \cdot \boldsymbol{\mu}$ and provides a metric distance between brains. Paths of minimum energy connecting any two fixed boundary conditions (BC) $\boldsymbol{\varphi}_0$ and $\boldsymbol{\varphi}_1$ where $\boldsymbol{\varphi}_1$ is accessible defines the distance extending LDDMM [6] to a hierarchy of diffeomorphisms, and is a geodesic for an associated Riemannian metric [33] on $\boldsymbol{G}_k$:

$$d_{\boldsymbol{G}_k}(\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_1)^2 := \min_{\substack{\boldsymbol{u}.\in L^2([0,1],\boldsymbol{V}): \\ \dot{\boldsymbol{\varphi}}_t = \boldsymbol{u}_t \circ \boldsymbol{\varphi}_t, \text{with BC } \boldsymbol{\varphi}_0, \boldsymbol{\varphi}_1}} \sum_{\ell<\ell_{max}} \int_0^1 \|u_t^\ell - u_t^{\ell-1}\|_{V_\ell}^2 dt \ . \tag{11}$$

The metric from $\boldsymbol{\mu}_0$ to $\boldsymbol{\mu}_1$ in the orbit $\boldsymbol{G}_k \cdot \boldsymbol{\mu}_0$ is the shortest length paths of (11) with BCs $\boldsymbol{\varphi}_1 \cdot \boldsymbol{\mu}_0 = \boldsymbol{\mu}_1$.

2.4. **The Model for Hilbert Space Construction of the Varifold Measure Norm.** The BC for matching two brains is defined using measure norms. Brains with 0 norm difference are equal; brains with small normed difference are similar. Every brain has a variable number of particles, with no apriori correspondence between particle dimensions. Measure norms accomodate these variabilities.

There are many measure norms. The varifold norm is constructed by associating an RKHS W with kernel K_W and integrating the measure against the kernel:

$$\|\mu\|^2_{W*} := \int_{\mathbb{R}^d \times \mathcal{F}} \mu(dx, df) \int_{\mathbb{R}^d \times \mathcal{F}} k_W((x,f),(y,g))\mu(dy,dg) \; . \tag{12}$$

For the varifold matching we use a Gaussian kernel that is separable over spatial position and feature values, each part having a scale parameter (or standard deviation). For spatial position and feature values the scale parameters correspond to a Euclidean distance between the position and feature points:

$$k_W((x,f),(y,g)) := \exp\left(-\frac{1}{2\sigma_x^2}|x-y|^2\right) \exp\left(-\frac{1}{2\sigma_f^2}|f-g|^2\right) \; , \tag{13}$$

where $|\cdot|$ is Euclidean norm. In varifold matching we use only Euclidean metrics, and not others such as the Fisher Rao metric discussed elsewhere. The smoothness property of the kernel implies the norm-square is smooth in perturbations. For $\mu = \sum_{i \in I} w_i \delta_{x_i} \otimes \mu_{f_i}$ this gives

$$\|\mu\|^2_{W*} = \sum_{i \in I, i' \in I} w_i w_{i'} \exp\left(-\frac{1}{2\sigma_x^2}|x_i - x_{i'}|^2\right) \exp\left(-\frac{1}{2\sigma_f^2}|f_i - f_{i'}|^2\right) \; , \tag{14}$$

The multi-scale norm becomes

$$\|\boldsymbol{\mu}\|^2_{\boldsymbol{W}*} = \sum\nolimits_{\ell < \ell_{max}} \|\mu^\ell\|^2_{W^*_\ell} \; . \tag{15}$$

2.5. **Model for Hamiltonian Control for Geodesic Brain Mapping.** Geodesic mapping solves for $\boldsymbol{u}_\cdot := (\boldsymbol{u}_t, t \in [0,1])$ minimizing the energy with the boundary condition encoded as the normed distance, $\alpha > 0$:

$$\begin{cases} \min\limits_{\boldsymbol{u}_\cdot \in L^2([0,1],\boldsymbol{V})} \dfrac{1}{2} \displaystyle\sum_{\ell < \ell_{max}} \int_0^1 \|u_t^\ell - u_t^{\ell-1}\|^2_{V_\ell} dt + \frac{\alpha}{2}\|\boldsymbol{\varphi}_1 \cdot \boldsymbol{\mu} - \boldsymbol{\mu}_{obs}\|^2_{\boldsymbol{W}*} \\ \dot{\boldsymbol{\varphi}}_t = \boldsymbol{u}_t \circ \boldsymbol{\varphi}_t, \; \boldsymbol{\varphi}_0 = \mathbf{Id} \end{cases} . \tag{16}$$

Hamiltonian control reparameterizes the flow via the multi-scale state $\boldsymbol{q}_t := (q_t^\ell)_{\ell < \ell_{max}}$:

$$\begin{aligned} & q_t := (x_{i,t} = \varphi_t(x_i), w_{i,t} = w_i |d\varphi_t|(x_i))_{i \in I} \; , \\ & \text{with } \mu(q_t) := \sum_{i \in I} w_{i,t} \delta_{x_{i,t}} \otimes \delta_{f_i} \; . \end{aligned} \tag{17}$$

The state encodes the action of (7b) $\boldsymbol{\mu}(\boldsymbol{q}_t) = \boldsymbol{\varphi}_t \cdot \boldsymbol{\mu}$ and the endpoint condition:

$$\boldsymbol{U}(\boldsymbol{q}_1) := \frac{\alpha}{2}\|\boldsymbol{\mu}(\boldsymbol{q}_1) - \boldsymbol{\mu}_{obs}\|^2_{\boldsymbol{W}*} \; . \tag{18}$$

The explicit form of the optimal control requires (i) the controls $u^\ell \in V_\ell$ are at least twice continuously differentiable, and (ii) the endpoint $\boldsymbol{U}(\boldsymbol{q}_1)$ is smoothly differentiable in $\boldsymbol{q}$ inherited from the kernel. Define $\varphi^\ell_{t,s} = \varphi^\ell_s \circ (\varphi^\ell_t)^{-1}$ and the accumulation of Green's functions $\bar{g}$ and $\nabla_1 \bar{g}(x,y)$ it's first variable gradient:

$$\bar{g}^\ell := \sum\nolimits_{j=0}^{\ell} g^j \; . \tag{19}$$

The optimal control (16) satisfies for every $\ell < \ell_{max}$:

$$\begin{cases} u_t^\ell(\cdot) = \sum_{j<\ell_{max}} \sum_{i\in I_j} \left(\bar{g}^{\ell\wedge j}(x_{i,t}^j,\cdot)p_{i,t}^{x,j} + p_{i,t}^{w,j} w_{i,t}^j \nabla_1 \bar{g}^{\ell\wedge j}(x_{i,t}^j,\cdot)\right) \\ p_{i,t}^{x,\ell} = (d\varphi_{t,1}^\ell)^T(x_{i,t}^\ell)p_{i,1}^{x,\ell} + p_{i,1}^{w,\ell} w_{i,1}^\ell \int_t^1 (d\varphi_{t,s}^\ell)^T(x_{i,t}^\ell)\,\nabla(\mathrm{div} u_s^\ell)(x_{i,s}^\ell)ds \\ p_{i,t}^{w,\ell} = p_{i,1}^{w,\ell}|d\varphi_{t,1}^\ell|(x_{i,t}^\ell) \end{cases} , \tag{20a}$$

with boundary conditions

$$p_{i,1}^{x,\ell} = -\nabla_{x_i^\ell}\boldsymbol{U}(\boldsymbol{q}_1),\ p_{i,1}^{w,\ell} = -\nabla_{w_i^\ell}\boldsymbol{U}(\boldsymbol{q}_1)\ . \tag{20b}$$

To explicitly define the gradients we introduce the field of smooth functions $h_{q^\ell}(x,f)$ on $\mathbb{R}^d \times \mathcal{F}^\ell$:

$$h_{q_1} := \alpha K_W[\mu(q_1) - \mu_{obs}]\ .$$

The kernel K_W gives the smoothness supporting gradients; for the Gaussian, $\mu = \sum_{i\in I} w_i\delta_{x_i} \otimes \mu_{f_i}$,

$$K_W[\mu](x,f) = \sum_i w_i \exp\left(-\frac{1}{2\sigma_x^2}|x-x_i|^2\right) \exp\left(-\frac{1}{2\sigma_f^2}|f-f_i|^2\right)\ .$$

Assuming h_q is continuously differentiable in x and bounded, the gradients of (20b) are given by,

$$\begin{cases} \nabla_{x_i^\ell}\boldsymbol{U}(\boldsymbol{q}_1) = w_{i,1}^\ell \nabla_{x_i} h_{q_1^\ell}(x_{i,1}^\ell, f_i^\ell) \\ \nabla_{w_i^\ell}\boldsymbol{U}(\boldsymbol{q}_1) = h_{q_1^\ell}(x_{i,1}^\ell, f_i^\ell) \end{cases} . \tag{21}$$

2.6. **Model for Merging with the Tissue Continuum.** The tissue continuum has been studied in Computational Anatomy with the continuum representation for the gradient useful for discrete computation. Tissue has density $\rho(dx) = \rho_c(x)dx$ and $\mu = \int_{\mathbb{R}^d} \rho_c(x)\delta_x \otimes \mu_x\, dx$. We take the states

$$\begin{aligned} &q_t := (\varphi_t, w_t = w_0|d\varphi_t|)\ , \\ &\text{with } \mu(q_t) := \int_{\mathbb{R}^d} w_t(x)\rho_c(x)\delta_{\varphi_t(x)} \otimes \mu_x dx\ . \end{aligned} \tag{22}$$

For initial condition $q_0 = (\mathrm{Id}, 1)$, then $\mu(q_t) = \varphi_t \cdot \mu$.

For dense tissue define

$$\begin{array}{l} \bar{h}_{q_1}^\nabla(x) := \int_{\mathcal{F}} \nabla_\varphi h_{q_1}(\varphi_1(x), f)\mu_x(df) \\ \bar{h}_{q_1}(x) := \int_{\mathcal{F}} h_{q_1}(\varphi_1(x), f)\mu_x(df) \end{array} ,$$

then the gradients for the continuum are

$$\begin{cases} \nabla_{\varphi^\ell}\boldsymbol{U}(\boldsymbol{\varphi}_1, \boldsymbol{w}_1) = w_1^\ell \rho_c^\ell \bar{h}_{q_1^\ell}^\nabla \\ \nabla_{w^\ell}\boldsymbol{U}(\boldsymbol{\varphi}_1, \boldsymbol{w}_1) = \rho_c^\ell \bar{h}_{q_1^\ell} \end{cases} . \tag{23}$$

We point out this unifies with dense image matching LDDMM [6] since when $I(x) \in \mathcal{F} = \mathbb{R}^+$ is taken as deterministic, then $\mu_x = \delta_{I(x)}$, $\mu^I := \int_{\mathbb{R}^d} \delta_x \otimes \delta_{I(x)}dx$ giving

$$\varphi \cdot \mu^I = \int_{\mathbb{R}^d} \delta_x \otimes \delta_{I\circ\varphi^{-1}(x)}dx\ . \tag{24}$$

2.7. **Feature Reduction:Segmentation of MRI and Neurophysiology:** Here are two examples demonstrating feature reduction.

2.7.1. *Bayes Segmentation of MRI.* Figure 4 shows the multi-scale data from the clinical BIOCARD study [2] of Alzheimer's disease within the medial temporal lobe [29, 53, 24, 62]. The top (left) panel shows a

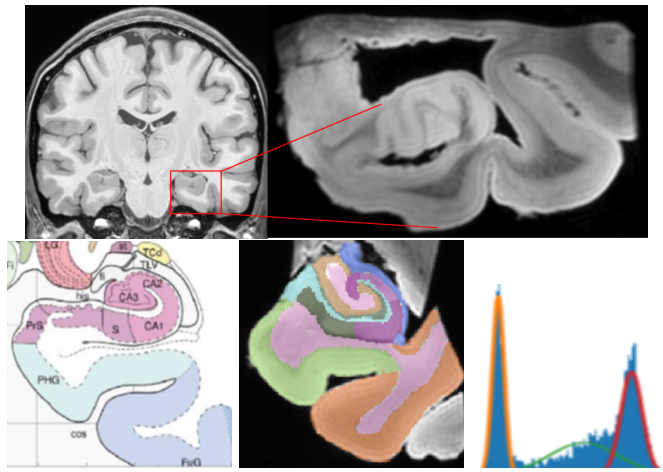

FIGURE 4. Top: Medial lobe at 1 mm and high-field 200 μm MRI. Bottom: Mai-Paxino atlas of the MRI with hippocampus and entorhinal cortex; right shows the Bayes compartments.

clinical magnetic resonance imaging (MRI) with the high-field 200 μm MRI scale (right) shown depicting the medial temporal lobe including the collateral sulcus and lateral bank of the entorhinal cortex. Bayes classifiers for parcellating the brain for studying neurodevelopmental and neurodegenerative disease performs feature reduction as a key step for segmentation at tissue scales [57, 16]. Feature reduction maps the distribution on gray levels $\mathcal{F} = [0, 255]$ to probabilities on N tissue types, defined by the integration over the decision regions $\theta_n \subset [0, 255]$:

$$\phi_n(\mu_x) = \int_{\theta_n} \mu_x(df) = p_n,\ n = 1, \ldots, N\ .$$

Figure 4 (bottom) depicts a Bayes classifier for gray, white and cerebrospinal fluid compartments generated from the temporal lobe high-field MRI section corresponding to the Mai-Paxinos section (left panel, bottom).

2.7.2. *Cellular Neurophysiology:* Single unit neurophysiology uses temporal models of spiking neurons with a "neural network" $\mu = \sum_i \delta_{x_i} \otimes \delta_{f_i}$, $x_i \in \mathbb{R}^d$, $f_i \in \mathcal{F}$ taking each neuron x_i modelled as a counting measure in time $N_i(t), t \geqslant 0$ with the spike times the feature $f_i = (t_k)_{1 \leqslant k \leqslant n_{f_i}}$:

$$\mathcal{F} = \{f = (t_k)_{1 \leqslant k \leqslant n_f} : n_f \geqslant 1, t_1 < t_2 \cdots < t_{n_f}\}\ .$$

The Poisson model with intensity $\lambda(t), t \geqslant 0$ [45] has probabilities $\Pr(N_i(t) = n) = \frac{(\int_{t_0}^t \lambda(s)ds)^n}{n!} e^{-\int_{t_0}^t \lambda(s)ds}$.

Post-stimulus time (PST) [18] and interval histograms are used to examine the instantaneous discharge rates and inter-spike interval statistics [7, 63], the first a linear dimension reduction on feature space and the second non-linear. The interval histogram abandons the requirement of maintaining the absolute phase of the signal for measuring temporal periodicity and phase locking [63, 48]. Synchrony in the PST is measured using binning $[b_i, b_{i+1})$, $i = 1, \cdots, B$ and Fourier transforms, j=$\sqrt{-1}$:

$$\phi_n^{PST}(\mu_x) = \sum_i e^{\mathrm{j}\omega_n i} \int_{\mathcal{F}} \frac{1}{n_f} \sum_{k=1}^{n_f} \mathbf{1}_{[b_i, b_{i+1})}(t_k) \mu_x(df)\ , n = 0, 1 \ldots\ .$$

The $n = 0$ frequency computes integrated rate; each phase-locked feature is complex $\phi_n \in \mathbb{C}$.

2.8. **Gaussian Scale-Space Resampling of Tau Histology:** For histology at the molecular scales the measure encodes the detected Tau and amyloid particles $\mu = \sum_i w_i \delta_{x_i} \otimes \delta_{f_i}$ for fine scale particles with function the geometric features $\mathcal{F} = \mathbb{R}^+$. Crossing from the histology accumulates the $(x_j \in \mathbb{R}^d)_{i \in I}$ Tau to the tissue lattice $(y_j \in Y_j \subset \mathbb{R}^d)_{j \in J}$ using Gaussian scale-space with normal reweighting in $\mathbb{R}^2$. Figure 5 shows the detected Tau particles as red dots at 4μ m.

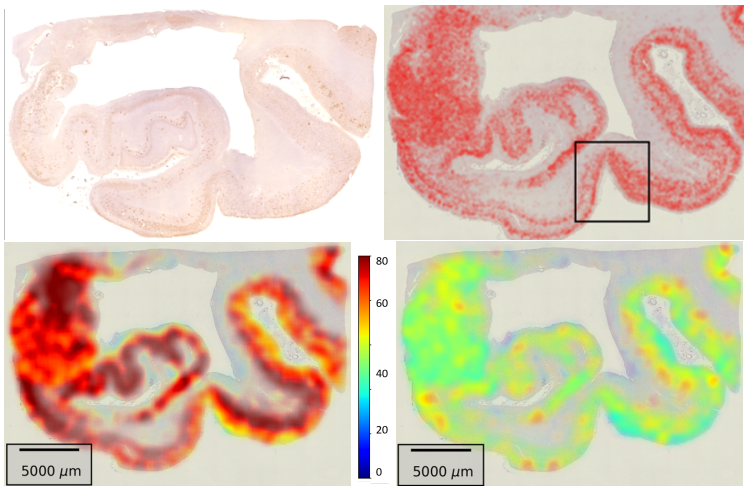

FIGURE 5. Alzheimer $4\,\mu$m Tau histology (red) from section depicted via high-field MRI Fig. 4; right shows detected Tau particles in simlar section. Box depicts trans-entorhinal region mapped in Figure 9. Bottom shows mean particle size and standard deviation at μm and tissue scales; deep red color denotes $80\,\mu\text{m}^2$ Tau area.

$$
\begin{aligned}
&\pi_\sigma(x_i, Y_j) = 1/(2\pi\sigma^2) \int_{Y_j} e^{-|y-x_i|^2/2\sigma^2} dy \,. \\
&\mu'^{\ell} = \sum_{j \in J} w_j \delta_{y_j} \otimes \mu'^{\ell}_{y_j} \\
&\mu^{\ell} = \sum\nolimits_{j \in J} w_j \delta_{y_j} \otimes \delta_{f_j^\ell} \\
&\quad \text{with} \begin{cases} w_j = \sum_{i \in I} w_i \pi_\sigma(x_i, Y_j) \\ \mu'^{\ell}_{y_j} = \frac{1}{w_j} \sum_{i \in I} w_i \pi_\sigma(x_i, Y_j)\, \delta_{f_i} \end{cases} .
\end{aligned}
\tag{25}
$$

Feature reduction to $\mathcal{F}^\ell$ is done with a moment expansion on the distribution of Tau:

$$
f_j^\ell = \phi(\mu'^{\ell}_{y_j}) = \left(\int_{\mathcal{F}} f \mu'^{\ell}_{y_j}(df), \int_{\mathcal{F}} (f)^2 \mu'^{\ell}_{y_j}(df) \right)
$$

The bottom rows shows the two moments at the tissue scale depicting the mean and variance of the particle size reconstructed from the 4μ m scale using Gaussian resampling onto the tissue lattice. The mm scale depicts the global folding property of the tissue. The color codes the mean Tau area as a function of position at the tissue scales with deep red color denoting $80\,\mu\text{m}^2$ maximum Tau area for the detected particles.

2.9. **Scale Space Resampling of RNA to Cell and Tissue Scales.** Methods in spatial-transcriptomics which have emerged for localizing and identifying cell-types via marker genes and across different cellular resolutions [43, 14, 49, 34, 56] presents the opportunity of localizing in spatial coordinates the transcriptionally

distinct cell-types. Depicted in the top row of Figure 6 are the molecular measurements at the micron scales with MERFISH [58]. The molecular measures represent RNA locations with sparse RNA features,

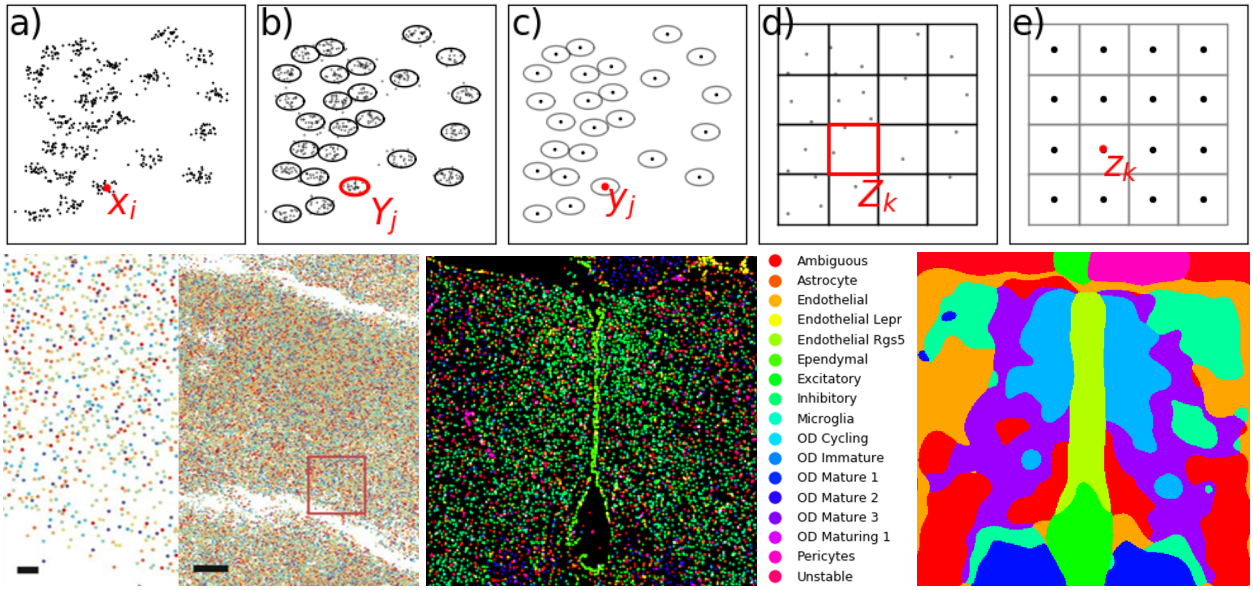

FIGURE 6. Top: Showing multi-scale renormalization of particles to cells and cell-centers to regular lattice representing the tissue. Bottom: RNA data [58] (left) shown as colored markers depicting locations of the 167 RNA gene species (bar scales $1, 10\,\mu$m); middle shows 17 cell types as colored particles clustered on cell centers with feature vectors eigen function reduction of 167 dimension RNA; right shows $K = 10$ means clustering to tissue.

$\mu^{\ell+1} = \sum_{i\in I} \delta_{x_i} \otimes \delta_{f_i}$, $\mathcal{F}^{\ell+1} = \mathbb{R}^{167}$. Crossing to cells $(Y_j \subset \mathbb{R}^2)_{j\in J}$ partitions into closest particle subsets:

$$\mathscr{I}_j = \{i \in I : d(x_i, Y_j) < d(x_i, Y_{j'}), j' \neq j\}\ , j \in J\ .$$

The RNA particles $(x_i \in \mathbb{R}^d)_{i\in I}$ are resampled to the cell centers $(\delta_{y_j}, y_j \in Y_j \subset \mathbb{R}^2)_{j\in J}$ accumulating to non-sparse mixtures of RNA. The new reduced feature vector becomes the conditional probability on 17 cell-types in $\mathcal{F}^\ell \subset [0,1]^{17}$. The conditional probabilities on the RNA vectors are calculated via principle components $E_n, n = 1, 2, \ldots$, modelling $\varepsilon^\ell_{y_j}(n) = \int E_n(f)\mu'^\ell_{y_j}(df)$ features as independent, Gaussian distributed:

$$\pi(x_i, Y_j) = \begin{cases} 1 & \text{for } x_i, i \in \mathscr{I}_j \\ 0 & \text{for } x_i, i \notin \mathscr{I}_j \end{cases}\ ,$$

$$\mu^\ell = \sum\nolimits_{j\in J} w_j\, \delta_{y_j} \otimes \delta_{f^\ell_j}$$

$$\text{with} \begin{cases} w_j = |\mathscr{I}_j| \\ \mu'^\ell_{y_j} = \frac{1}{|\mathscr{I}_j|} \sum_{i\in\mathscr{I}_j} \delta_{f_i} \\ f^\ell_j(c) = \Pr\{C_{y_j} = c | \varepsilon^\ell_{y_j}(1), \varepsilon^\ell_{y_j}(2), \ldots\}\ ,\ c = 1, \ldots, 17\ . \end{cases}\ ,$$

Resampling to tissue $(\delta_{z_k}, z_k \in Z_k \subset \mathbb{R}^2)_{k\in K}$ uses normal rescaling. The new feature vector becomes the probability of the cell at any position being one of 10 tissue types $\mathcal{F}^{\ell-1} \subset [0,1]^{10}$. The probability of tissue type is calculated using 10-means clustering on the cell type probabilities. The distance for 10-means

clustering is computed using the Fisher-Rao metric [47] between the empirical feature laws $\mu'^{\ell-1}_{z_k}$. The output of 10-means are a partition of feature space $\cup_t \mathscr{F}_t = \mathcal{F}^\ell$ giving new features:

$$\pi_\sigma(y_j, Z_k) = \frac{1}{2\pi\sigma^2}\int_{Z_k} e^{-|z-y_j|^2/2\sigma^2} dz \ .$$

$$\mu^{\ell-1} = \sum\nolimits_{k\in K} w_k \delta_{z_k} \otimes \delta_{f_k^{\ell-1}}$$

$$\text{with} \begin{cases} w_k = \sum_{j\in J} |\mathscr{I}_j| \, \pi_\sigma(y_j, Z_k), \\ \mu'^{\ell-1}_{z_k} = \frac{1}{w_k}\sum_{j\in J} |\mathscr{I}_j| \, \pi_\sigma(y_j, Z_k)\, \delta_{f_j^\ell} \\ f_k^{\ell-1} = \int_{\mathscr{F}_t} \mu'^{\ell-1}_{z_k}(df) \ , \ t = 1, \ldots, 10 \ . \end{cases} .$$

Figure 6 (bottom left panels) shows the RNA forming $\mu^{\ell+1}$ depicted as colored markers corresponding to the 100's of different gene species (bar scale $1, 10$ microns). The middle panel shows the labels of the cell type feature space of 17 cell types making up μ^ℓ associated to the maximal probability in the PCA projection from a classifier on the PCA dimensions based on the mixtures of RNA at each cell location. The right panel shows the 10 tissue features associated to the 10-means procedure. In both scales the probabilities are concentrated on one class with probability 1 and the others 0.

2.10. **Geodesic Mapping for Spatial Transcriptomics, Histology and MRI.**

2.10.1. *Navigation Between Sections of Cell Identity in Spatial Transcriptomics.* Figure 7 shows sections from [56] depicting neuronal cell types via colors including excitatory cells eL2/3 (yellow), eL4 (orange), red eL5 (red), inhibitory cells ST (green), VIP (light blue), each classified via high dimensional gene expression features vectors via spatial transcriptomics.

The mathematical measure representation $\mu = \sum_i w_i \delta_{x_i} \otimes \delta_{f_i}$, $\mathcal{F} = \{\text{cell types}\}$. Crossing to atlas scales uses normal reweighting π_σ in $\mathbb{R}^2$ as above Eqn. (25) with $\mu^\ell = \sum_{j\in J} w_j \delta_{y_j} \otimes \delta_{\phi(\mu'^\ell_{y_j})}$ and feature reduction done with expectations of moments, $\mathcal{F}^\ell = \{\text{size, variance,entropy}\}$:

$$\phi(\mu'^\ell_{y_j}) = \left(\sum_{\mathcal{F}} f \mu'^\ell_{y_j}(f), \sum_{\mathcal{F}} (f)^2 \mu'^\ell_{y_j}(f), -\sum_{\mathcal{F}} \mu'^\ell_{y_j}(f) \log \mu'^\ell_{y_j}(f) \right) \ .$$

The right panels show the tissue scale feature dimensions of mean cell type size and the entropy of the conditional. The entropy is a measure of dispersion across the cell identities given by the expectation of the log probability function. An entropy of zero means that space location feature disribution μ'^ℓ_x has all its mass on 1 cell type.

Shown in Figure 7 are results transforming the spatial transcriptome of cell type [56] depicting neuronal cell types classified via high dimensional features vectors of as many as $20,000$ genes from Figure 7. Shown in the middle row is the mapping $\varphi_t \cdot \mu$ of the two sections based on the varifold measure norm Eqn. (14) a product of Gaussians between the sections showing the particles sizes; the bottom row depicts the entropy of the measure of each particle at the coarse scale.

2.10.2. *Navigation Between Sections of Histology.* Figure 8 shows navigation between the cortical folds of the $4\,\mu$m histology. Shown in panel 1 is a section showing the machine learning detection of the Tau particles. Columns 2,3, and 4 depict the template, mapped template and target showing the mathematical measure representation of the perirhinal cortex constructed from the positions and sizes at the 4μ m scale (top row) and reconstruction using Gaussian resampling onto the tissue scale (bottom row). The color codes the mean of μ_x representing Tau area as a function of position at the tissue scales with deep red of $80\,\mu\text{m}^2$ of Tau

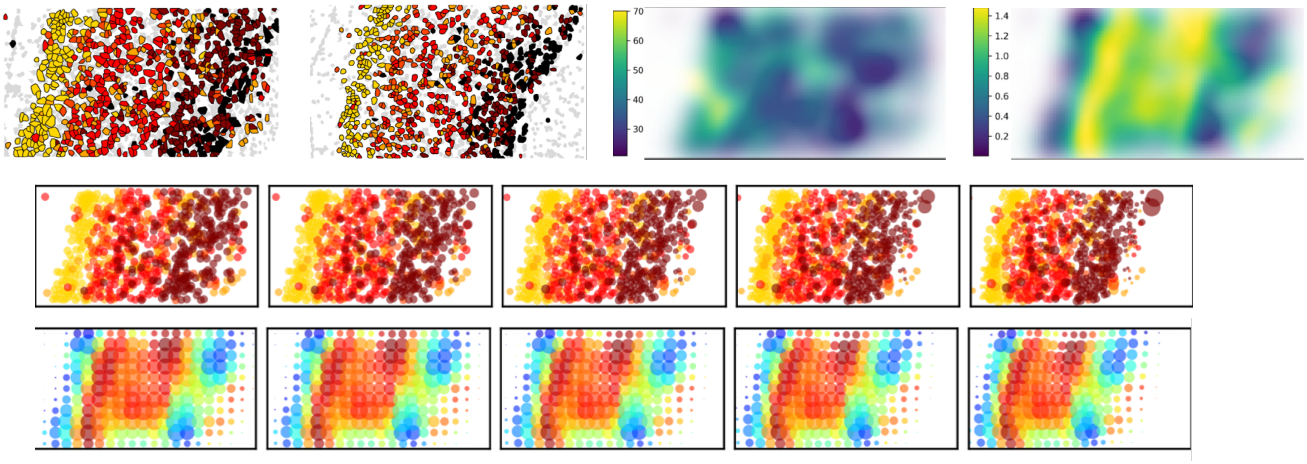

FIGURE 7. Top row: Cells depicting neuronal cell types [56]; right shows cell size and entropy of the conditional distributions for panel 1 at tissue scales. Bottom rows: Neuronal cell type from [56]; showing cell size (top row) and entropy (bottom row) of the conditional distributions. Mapping $\varphi_t \cdot \mu$ fine cell scales showing the cell types mapping with the bottom row mapping the entropy measure of each particle. Units are pixels with Gaussian resampling from fine to coarse scale was with a scale of 25 pixels on the total grid of $450 \times 200\, pix^2$; varifold norm $\sigma_x^\ell = 10\, pix, 50\, pix$ with vector field mappings $\alpha^\ell = 5\, pix, 50\, pix$ with $FWHM = 20\, pix, 145\, pix$, for $\ell = 1, 2$ respectively. Here the feature is a one hot encoded cell type for coarse scale varifold matching, and for feature is also $\sigma_f^\ell = 10, 50$. Here the feature is mean cell type (probability) and entropy in cell type.

area the maximum value. The gradients in tau tangle area between superficial and deep layers is apparent with the deep red as high as $80\, \mu$m^2 for the innermost tissue fold. The bottom left panel shows the vector field encoding of the geodesic transformation of the of the perirhinal cortex mapping between the two sections in column 1. The narrowing of the banks of the perirhinal cortex is exhibited at the tissue scale for motions order $1000\, \mu$m (brightness on scale bar).

Figure 9 shows the collateral sulcus fold at the boundary of the trans-entorhinal cortex region transforming based on the normed distances between sections with deformation motions $1000\, \mu$m in size. Shown in the two rows are the two scales depicting the transformation of the gyrus with the color representing the mean of the particle size distribution.

2.10.3. *Mapping between MRI and Histology Simultaneously.* All of the examples thus far have created the multi-scale data generated using the resampling kernels from the finest scales. As illustrated in our early figures much of the data is inherently multi-scale, with the measurement technologies generating the coarse scale representations. Shown in Figure 10 is data from Stoffer et. al. [23] illustrating our Alzheimer's study of post mortem MR images that are simultaneously collected with amyloid and Tau pathology sections. MR images have a resolution of approximately 100 μm, while pathology images have a resolution of approximately 1 μm. For computational purposes the MRI template and target images were downsampled to 759 and 693 particles, respectively with the tau tangles downsampled to 1038 and 1028 particles, respectively. We treated every pixel in the MR image as a coarse scale particle (Eqn. (24)) with image intensity as its feature value, and every detected tau tangle as a fine scale particle with a constant feature value, and performed varifold matching to align to neighboring sections. The endpoint representing the two scales is

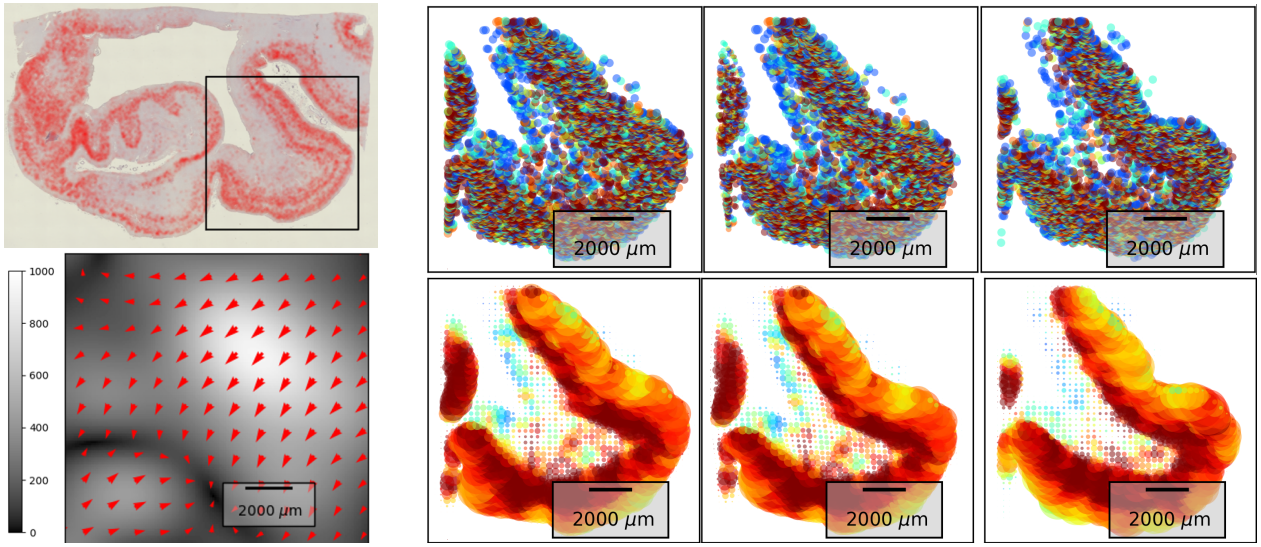

FIGURE 8. Second section 4μ m histology (similar to Figure 5) with box depicting perirhinal cortex. Columns 2, 3 and 4 show template, mapped template and target at the molecular (top) and tissue scales (bottom) showing the first moments of Tau size on the perirhinal cortex; saturated red color denotes $80\,\mu\text{m}^2$ Tau area. Mapped template shows narrowing $1000\,\mu$m of perirhinal sulcus; varifold norm $\sigma_x^\ell = 800\,\mu\text{m}, 160\,\mu\text{m}$ with vector field mappings α^ℓ with $FWHM = 580\,\mu, 3300\,\mu$, for $\ell = 1, 2$ respectively. Bottom left panel depicts vector field encoding of the geodesic mapping with associated scale bar.

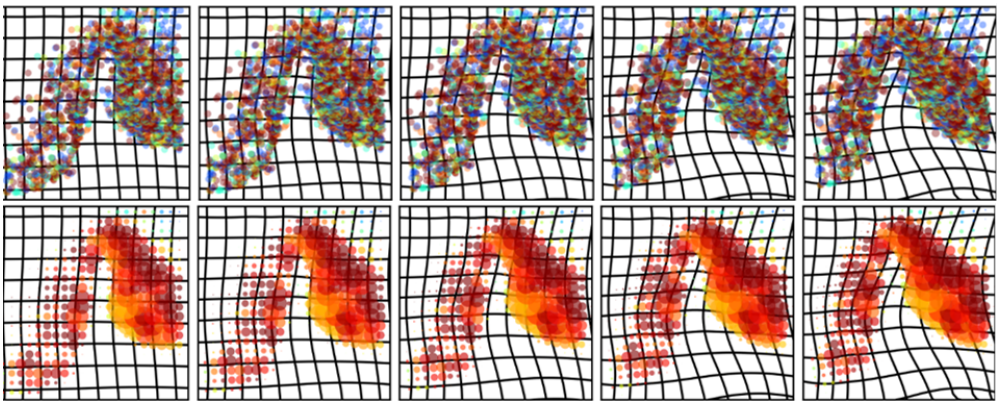

FIGURE 9. Geodesic navigation $\varphi_t \cdot \mu$ of collateral sulcus (Figure 5, box) showing $1000\,\mu$m widening of folds for molecular and tissue scales depicting the mean transforming; varifold norm $\sigma_x^\ell = 800\mu\text{m}, 160\mu\text{m}$ with vector field mappings α^ℓ with $FWHM = 580\mu\text{m}, 3300\mu\text{m}$, for $\ell = 1, 2$ respectively.

$\boldsymbol{U} = \frac{1}{2}\sum_{\ell=1,2}\left(\mu_{temp}^\ell - \mu_{obs}^\ell \,\middle|\, K_{W_\ell}[\mu_{temp}^\ell - \mu_{obs}^\ell]\right)$. For each scale norm we use a varifold kernel given by the products of Gaussian distributions with the varifold measure norm Eqn. (14) at each scale. For the MRI scale, the weights are identically $w = 1$ with the function component given by the MRI image value; for the tau particles there is no function component making $f = g$ with the kernel 1 for all values.

The top row of Figure 10 shows the imaging data for both sections. The bottom row shows the transformed template image at the fine scale. The high resolution mapping carries the kernels across all the scales as indicated by the geodesic equation (20a). Notice the global motions of the high resolution of the fine particles.

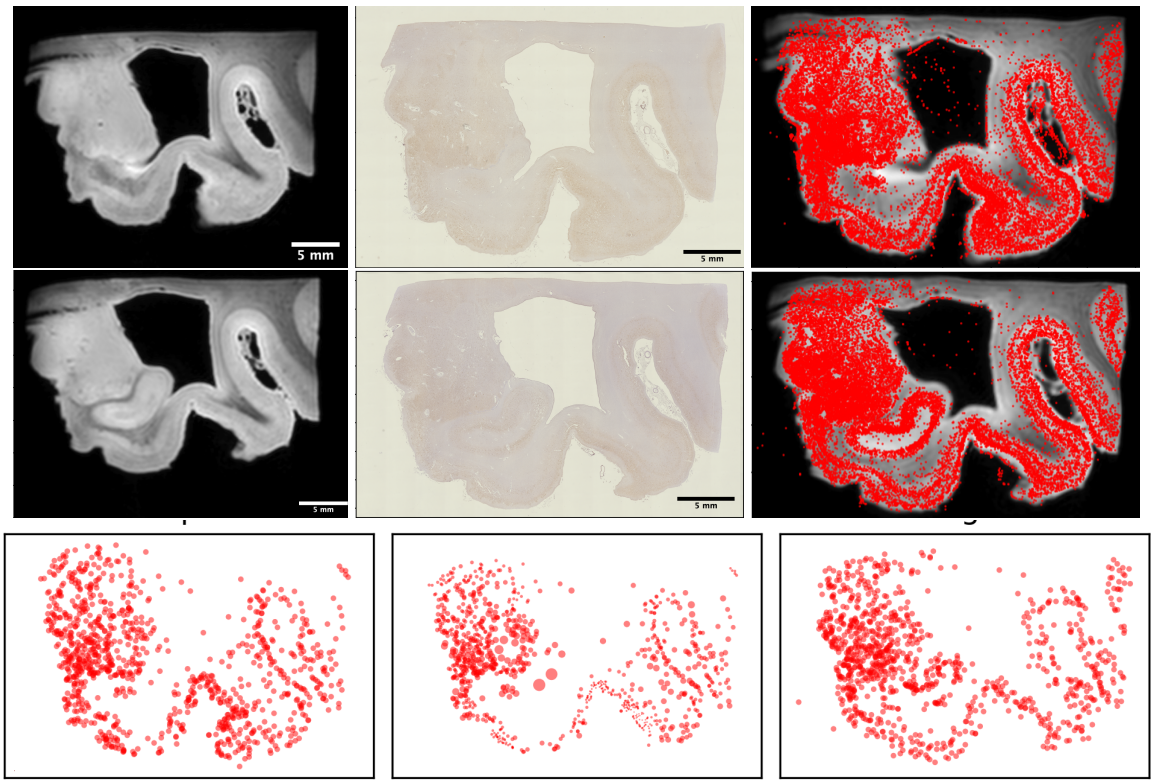

FIGURE 10. Whole brain section showing mapping MRI and histology at the multiple scales. Top row shows the MRI and Tau histology for two sections with the detected Tau particle superimposed over the MRI (right); bottom row shows the finest scales for the template, middle the template mapped, with the right shows the target; varifold norm $\sigma_x^\ell = 4000\,\mu\text{m}, 100\,\mu\text{m}$ with vector field mappings α^ℓ with $FWHM = 1500\,\mu\text{m}, 6400\,\mu\text{m}$, for $\ell = 1, 2$ respectively.

3. DISCUSSION

Computational anatomy was originally formulated as a mathematical orbit model for representing medical images at the tissue scales. The model generalizes linear algebra to the group action on images by the diffeomorphism group, a non-linear algebra, but one that inherits a metric structure from the group of diffeomorphisms [28]. The formulation relies on principles of continuity of medical images as classical functions, generalizating optical flow and advection of material to diffeomorphic flow of material, the material represented by the contrast seen in the medical imaging modality such as fiber orientation for diffusion tensor imaging, and or bold contrast for gray matter content. Unifying this representation to images built at the particle and molecular biological scale has required us to move away from classical functions, to the more modern 20th century theory of non-classical generalized functions. Mathematical measures are the proper representation as they generally reflect the property that probes from molecular biology associated to disjoints sets

are additive, the basic starting point of measure theory. Changing the model from a focus on groups acting on functions to groups acting on measures allows for a unified representation that has both a metric structure at the finest scales, as well as a unification with the tissue imaging scales.

The brain measure formulation, carries with it implicitly the notion of scale-space, i.e. the existence of a sequence of pairs across scales, the measure representation of the brain and the associated scale-space reproducing kernel Hilbert space of functions which correspond to the probing measurement technologies. As such part of the prescription of the theory is a method for crossing scales and carrying information from one scale to the other. Important to this approach is that at every scale we generate a new measure, therefore the recipe of introducing "measure norms" built from RKHS's for measuring brain disparity is universal across the hierarchy allowing us to work simultaneously with common data structures and a common formalism. Interestingly, the measure norms do not require identical particle numbers across brains in brain space at the molecular scales.

The key modelling element of brain function is that the conditional feature probability is manipulated from the quantized features to the stochastic laws. These are the analogues of the Boltzman distributions generalized to the complex feature spaces representing function. As they correspond to arbitary feature spaces not necesarily newtonian particles, we represent them simply as empirical distributions on the feature space, with the empirical measure constructed from the collapse of the fine scale to the resampled coarse scale. To model rescaling through scale-space explicitly, the two kernel transformation are used allowing us to retrieve the empirical averages represented by the determinism of the stochastic law consistent with our views of the macro tissue scales. This solves the dilemna that for the quantized atomic and micro scales cell occurence will never repeat, i.e. there is zero probability of finding a particular cell at a particular location, and conditioned on finding it once it will never be found again in the exact same location in another preparation. The properties that are stable are the probability laws with associated statistics that may transfer across organisms and species.

Importantly, our introduction of the $|d\varphi(x)|$ term in the action enables the crucial property that when a tissue is extended to a larger area, the total number of its basic constituents should increase accordingly and not be conserved. This is not traditional measure transport which is mass preserving which is not a desirable feature for biological samples. Rather we have defined a new action on measures that is reminiscent of the action on d-dimensonal varifolds [8, 22]. We call this property "copy and paste", the notion being that the brain is built on basic structuring elements that are conserved.

Successive refinement for the small deformation setting as a principled for scale hierarchy was introduced in stochastic optimization by Amit et al. [4] expanding vector fields in a basis. The notion of building multi-scale representation in the large deformation LDDMM setting was originally explored F. Riesser et al. [40] in which the kernels are represented as a sum of kernels and Sommer et al. [46] in which the kernel is represented as vector bundles. In their multi-scale setting there is a post-optimization decomposition in which the contribution of the velocity field into its different components can then each be integrated. In that multi-scale setting the basic Euler-Lagrange equation termed EPDIFF remains that of LDDMM [31]. In the setting proposed here we separate the scales before optimisation via the hierarchy of layered diffeomorphisms and use a multi-scale representation of the brain hierarchy itself which is directly associated to the differomorphism at that scale. This gives the fundamental setting of the product group of diffeomorphisms with the Euler-Lagrange equation corresponding to the sequence of layered diffeomorphisms [33].

The aggregation across scales from particle to tissue scales on lattices provides the essential link to inference on graphs. It is natural for these aggregated features with associated conditional probability laws to become the nodes in Markov random field modelling for spatial inference; see examples in spatial transcriptomics [64] and tissue segmentation [60]. Building neighborhood relations as conditional probabilities

between lattice sites from which global probabilites laws are constructed with the Hammersley-Clifford theorem [19] links us to Grenander's metric pattern theory [20] formalisms with the atoms and conditional laws $(x_i, \mu_{x_i})_{i\in I}$ at any scale playing the roles of the generators.

4. METHODS

4.1. **Method of Resampling Across Lattice Scales.** Taking $\mu = \sum_{i\in I} w_i \delta_{x_i} \otimes \delta_{f_i}$ on $\mathbb{R}^d \times \mathcal{F}$ with kernels k_1, k_2 of (6a), (6c) gives our two-step transformation. The transformation (5a) of $T_1 : \mu \mapsto \mu'^{\ell}$ gives

$$\begin{aligned} \mu'^{\ell} &= \int_{\mathbb{R}^d\times\mathcal{F}} \sum_{j\in J} \pi(x, Y_j)\, \delta_{y_j} \otimes \delta_f \sum_{i\in I} w_i\, \delta_{x_i} \otimes \delta_{f_i}\, (dx, df) \\ &= \sum_{j\in J} w_j \delta_{y_j} \otimes \frac{1}{w_j} \left(\sum_{i\in I} w_i \pi(x_i, Y_j) \delta_{f_i} \right) , \end{aligned}$$

for $w_j = \sum_{i\in I} w_i \pi(x_i, Y_j)$. Defining $\mu'^{\ell}_{y_j} = \sum_{i\in I} w_i \pi(x_i, Y_j)\delta_{f_i}/w_j$ gives (6b):

$$\mu'^{\ell} = \sum_{j\in J} w_j \delta_{y_j} \otimes \mu'^{\ell}_{y_j} .$$

The space density follows

$$\rho = \int_{\mathcal{F}} \sum_{j\in J} w_j \delta_{y_j} \otimes \mu'^{\ell}_{y_j}(df) = \sum_{j\in J} w_j \delta_{y_j} .$$

The feature transformation (5b) of $T_2 : \mu'^{\ell} \mapsto \mu^{\ell}$ gives (6d):

$$\mu^{\ell} = \int_{\mathbb{R}^d} \delta_x \otimes \delta_{\phi(\mu'^{\ell}_x)} \left(\sum_{j\in J} w_j \delta_{y_j}(dx) \right) = \sum_{j\in J} w_j \delta_{y_j} \otimes \delta_{\phi(\mu'^{\ell}_{y_j})} .$$

4.2. **Method of Comparison via Hamiltonian Control for Generating the Geodesics.** For any $\mu = \sum_{i\in I} w_i \delta_{x_i} \otimes \delta_{f_i}$, the diffeomorphisms act $(\varphi, \mu) \mapsto \varphi \cdot \mu$ according to (7a). To accomodate scale the product group for the hierarchy becomes

$$\boldsymbol{G}_k = \underbrace{G_k \times \cdots \times G_k}_{\ell_{max} \text{ times}} ,$$

with elements $\boldsymbol{\varphi} \in \boldsymbol{G}_k$ satisfying law of composition component-wise $\boldsymbol{\varphi} \circ \boldsymbol{\varphi}' = (\varphi^{\ell} \circ \varphi'^{\ell})_{\ell<\ell_{max}}$ (Figure 3). The group acts on the hierarchy componentwise according to (7a), (7b) $\boldsymbol{\varphi} \cdot \boldsymbol{\mu}$.

Dynamics occurs through the group action on the brain measures generated as a dynamical system in which the multi-scale control $t \mapsto \boldsymbol{u}_t := (u^{\ell}_t)_{\ell<\ell_{max}}$ flows the hierarchy $t \mapsto \boldsymbol{\varphi}_t$ with ODE of (8) $\dot{\boldsymbol{\varphi}}_t = \boldsymbol{u}_t \circ \boldsymbol{\varphi}_t$.

Each control u^{ℓ} is constrained to be an element of a reproducing kernel Hilbert space (RKHS) V_{ℓ} with inner-product $\langle\, ,\, \rangle_{V_{\ell}}$. The total $\boldsymbol{u}$ is in the product space $\boldsymbol{u} \in \boldsymbol{V} = \prod_{\ell<\ell_{max}} V_{\ell}$ with the norm determined by the inner-product $\langle\, ,\, \rangle_{\boldsymbol{V}} = \sum_{\ell<\ell_{max}} \langle\, ,\, \rangle_{V_{\ell}}$. The spaces are organised as a sequence of continous embeddings:

$$V^0 \hookrightarrow \cdots \hookrightarrow V^{\ell_{max}} ,$$

where $V^{\ell_{max}}$ is an additional layer containing the others. The hierarchy is connected via the continuous linear operator $\boldsymbol{A} : \boldsymbol{V} \to \boldsymbol{V}$ with $\boldsymbol{v} = \boldsymbol{A}\boldsymbol{u}$ for $u^{\ell} = u^{\ell-1} + v^{\ell}, u^0 = v^0$. The control process $\boldsymbol{u}. = (\boldsymbol{u}_t, t \in [0,1]) \in L^2([0,1], \boldsymbol{V})$ has finite square-integral with total energy

$$\mathcal{E}_{\boldsymbol{A}}(\boldsymbol{u}.) = \frac{1}{2} \int_0^1 \|\boldsymbol{A}\boldsymbol{u}_t\|^2_{\boldsymbol{V}} dt .$$

Optimal curves which minimize the integrated energy $\mathcal{E}_{\boldsymbol{A}}$ between the fixed boundary condition $\boldsymbol{\varphi}_0$ and $\boldsymbol{\varphi}_1$ which is accessible with a path of finite energy describes a geodesic for an associated Riemannian metric [33] on $\boldsymbol{G}_k$ given by $d_{\boldsymbol{G}_k}(\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_1)$ of (11) given by $\sqrt{\mathcal{E}}_{\boldsymbol{A}}(\boldsymbol{u}.)$ evaluated at its minimizer. The metric distance from $\boldsymbol{\mu}_0$ to $\boldsymbol{\mu}_1$ in the orbit $\boldsymbol{G}_k \cdot \boldsymbol{\mu}_0$ is given by the shortest length paths of (11) with boundary conditions $\boldsymbol{\varphi}_1 \cdot \boldsymbol{\mu}_0 = \boldsymbol{\mu}_1$. Existence of solutions for minimizers of (11) can be established under a broad mathematical setting. For the group G_k and $V^{\ell_{max}}$ the space of m-times continuously differentiable vector fields vanishing at infinity with $m \geqslant k \geqslant 1$ properly defined, the metric holds. Specifically, if $\mathscr{C}_0^k(\mathbb{R}^d, \mathbb{R}^d)$ is a space of $\mathscr{C}^k$ vector fields vanishing at infinity as well all its partial derivatives of order $p \leqslant k$, then for $G_k = (\mathrm{id} + \mathscr{C}_0^k(\mathbb{R}^d, \mathbb{R}^d)) \cap \mathrm{Diff}^1(\mathbb{R}^d, \mathbb{R}^d)$ and $V^{\ell_{max}} = \mathscr{C}_0^m(\mathbb{R}^d, \mathbb{R}^d)$ with $m \geqslant k$ the metric holds.

The Hamiltonian method reduces the parameterization of the vector field to the essential dynamics of the particles that encode the flow with states $q_t = (x_{i,t} = \varphi_t(x_i), w_{i,t} = w_i |d\varphi_t|(x_i))_{i \in I}$ and state dynamics of (17). We write the dynamics explicitly as a linear function of the control,

$$\xi_{q_t}(u_t) := (u_t(x_{i,t}), w_{i,t}\, \mathrm{div} u_t(x_{i,t}))_{i \in I} \ ,$$

with $\boldsymbol{\xi}_{\boldsymbol{q}_t}(\boldsymbol{u}_t) := (\xi_{q_t^\ell}(u_t^\ell))_{\ell < \ell_{max}}$.

The optimal control problem satisfying (16) reparameterized in the states becomes, for $\alpha > 0$:

$$\begin{cases} \min\limits_{\boldsymbol{u}. \in L^2([0,1], \boldsymbol{V})} \dfrac{1}{2} \displaystyle\int_0^1 \|\boldsymbol{A}\boldsymbol{u}_t\|_{\boldsymbol{V}}^2 dt + \dfrac{\alpha}{2} \|\boldsymbol{\mu}(\boldsymbol{q}_1) - \boldsymbol{\mu}_{obs}\|_{\boldsymbol{W}^*}^2 \\ \dot{\boldsymbol{q}}_t = \boldsymbol{\xi}_{\boldsymbol{q}_t}(\boldsymbol{u}_t), \ \boldsymbol{\mu}(\boldsymbol{q}_0) = \boldsymbol{\mu} \end{cases} . \tag{26}$$

The Hamiltonian controls the states via co-states $\boldsymbol{p} = (\boldsymbol{p}^x, \boldsymbol{p}^w)$ introduced as

$$\begin{aligned} \mathcal{H}(\boldsymbol{q}, \boldsymbol{p}, \boldsymbol{u}) &= (\ \boldsymbol{p} \mid \boldsymbol{\xi}_{\boldsymbol{q}}(\boldsymbol{u})) - \frac{1}{2} \|\boldsymbol{A}\boldsymbol{u}\|_{\boldsymbol{V}}^2 \ , \\ &\text{with } (\ \boldsymbol{p} \mid \boldsymbol{\xi}_{\boldsymbol{q}}(\boldsymbol{u})) = \textstyle\sum_{\ell < \ell_{max}} \left(\sum_{i \in I_\ell} \langle p_i^{x,\ell}, u^\ell(x_i^\ell) \rangle_{\mathbb{R}^d} + p_i^{w,\ell} w_i^\ell \mathrm{div} u^\ell(x_i^\ell) \right) \ . \end{aligned} \tag{27}$$

The Pontryagin maximum [32] gives the optimum control $\partial_{u^\ell} \mathcal{H}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{u}) = 0$, with dynamics for every ℓ,

$$\begin{cases} \dot{q}^\ell = \ \ \partial_{p^\ell} \mathcal{H}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{u}) \\ \dot{p}^\ell = -\partial_{q^\ell} \mathcal{H}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{u}) \end{cases} . \tag{28}$$

Statement 1. *We assume that $V^{\ell_{max}} = \mathscr{C}_0^m(\mathbb{R}^d, \mathbb{R}^d)$ with $m \geqslant 2$. If $\boldsymbol{u}.$ is a solution of the optimal control problem* (26) *then there exists time-dependent co-state such that for every $i \in I_\ell$ and $\ell < \ell_{max}$:*

$$\begin{cases} \dot{q}_{i,t}^\ell = (u_t(x_{i,t}^\ell), w_{i,t}^\ell \operatorname{div} u_t(x_{i,t}^\ell)) \\ \dot{p}_{i,t}^{x,\ell} = -(du_t)^T(x_{i,t}^\ell) p_{i,t}^{x,\ell} - p_{i,t}^{w,\ell} w_{i,t}^\ell \nabla(\operatorname{div} u_t)(x_{i,t}^\ell) \\ \dot{p}_{i,t}^{w,\ell} = -p_{i,t}^{w,\ell} \operatorname{div} u_t(x_{i,t}^\ell) \\ u_t^\ell(\cdot) = \sum_{j < \ell_{max}} \sum_{i \in I_\ell} \left(\bar{g}^{\ell \wedge j}(x_{i,t}^j, \cdot) p_{i,t}^{x,j} + p_{i,t}^{w,j} w_{i,t}^j \nabla_1 \bar{g}^{\ell \wedge j}(x_{i,t}^j, \cdot) \right) \end{cases} . \tag{29}$$

The Hamiltonian differential equations are derived in Appendix A using calculus of variations; the averaged Green's function $\bar{g}$ is Eqn. (19). The endpoint condition $\boldsymbol{U}(\boldsymbol{q}_1)$ is added to the Hamiltonian with smoothness of $\boldsymbol{U}(\boldsymbol{q}_1)$ in the state giving absolutely integrable solutions to the Hamitonian dynamics. We omit the superscripts ℓ below since co-states and states and flows are at the same scale.

Statement 2. *Assume $\boldsymbol{q} \to \boldsymbol{U}(\boldsymbol{q})$ is $\mathscr{C}^1$ in $\boldsymbol{q}$, then the co-state integral equations* (20a) *flowing from $t = 1$ solves the Hamiltonian differential equations* (29)*; the integral equations flowing from $t = 0$ satisfy:*

$$\begin{cases} p^x_{i,t} = (d\varphi_{t,0})^T(x_{i,t})p^x_{i,0} - p^w_{i,0}w_{i,0}\displaystyle\int_0^t (d\varphi_{t,s})^T(x_{i,t})\,\nabla(\mathrm{div} u_s)(x_{i,s})ds \\ p^w_{i,t} = p^w_{i,0}|d\varphi_{t,0}|(x_{i,t}) \end{cases} . \tag{30}$$

Proof. First take $p^w_{i,t} = p^w_{i,0}|d\varphi_{t,0}|(x_{i,t})$ and show it satisfies (29). For $w_{i,t} = w_i|d\varphi_t|(x_i)$, $x_{i,t} = \varphi_t(x_i)$ then

$$\begin{aligned} \frac{d}{dt}\, w_{i,t}p^w_{i,t} &= \frac{d}{dt}\, w_{i,0}|d\varphi_t|(x_i)p^w_{i,0}|d\varphi_t^{-1}|(x_{i,t}) = 0 \\ &\Rightarrow\ \dot{p}^w_{i,t} = -\frac{p^w_{i,t}\dot{w}_{i,t}}{w_{i,t}} = -p^w_{i,t}\mathrm{div}\, u_t(x_{i,t}) \end{aligned} \tag{31}$$

The $p^x_{i,t}$ satisfies (29) since it is a linear inhomogeneous differential equation of the form $\dot{a}_t = -(du)^T(x_{i,t})a_t - b_t$. We rewrite the integral solution using $\varphi_{t,s} := \varphi_s \circ \varphi_t^{-1}$ and the identities:

$$\begin{aligned} &(d\varphi_{t,s})^T = (d\varphi_t^{-1})^T d\varphi_s^T \circ \varphi_t^{-1}\ , \\ &(d\varphi_{t,0})^T(x_{i,t}) = (d\varphi_t^{-1})^T \circ \varphi_t(x_i)\ , \\ &p^x_{i,t} = (d\varphi_t^{-1})^T \circ \varphi_t(x_i)p^x_{i,0} - p^w_{i,0}w_{i,0}\int_0^t (d\varphi_t^{-1})^T \circ \varphi_t(x_i)d\varphi_s^T(x_i)\nabla(\mathrm{div}\, u_s)(x_{i,s})dx\ . \end{aligned}$$

Differentiating requires the identity $\frac{d}{dt}\,(d\varphi_t^{-1})^T \circ \varphi_t = -(du_t)^T \circ \varphi_t\,(d\varphi_t^{-1})^T \circ \varphi_t$ (see below):

$$\begin{aligned} \dot{p}^x_{i,t} &= -(du_t)^T(x_{i,t})\left((d\varphi_t^{-1})^T \circ \varphi_t(x_i)p^x_{i,0} - p^w_{i,0}w_{i,0}\int_0^t (d\varphi_t^{-1})^T \circ \varphi_t(x_i)d\varphi_s^T(x_i)\nabla(\mathrm{div}\, u_s)(x_{i,s})dx\right) \\ &\quad -p^w_{i,0}w_{i,0}(d\varphi_t^{-1})^T \circ \varphi_t(x_i)d\varphi_t^T(x_i)\nabla(\mathrm{div}\, u_t)(x_{i,t}) \\ &= -(du_t)^T(x_{i,t})p^x_{i,t} - p^w_{i,t}w_{i,t}\nabla(\mathrm{div}\, u_t)(x_{i,t})\ , \end{aligned}$$

where the last equality uses $w_{i,0}p^w_{i,0} = w_{i,t}p^w_{i,t}$ from (31). □

The remaining identity follows from $\frac{d}{dt}((d\varphi_t)^{-1}\, d\varphi_t) = 0$:

$$\begin{aligned} \frac{d}{dt}(d\varphi_t)^{-1} &= -(d\varphi_t)^{-1}(d\dot{\varphi}_t)(d\varphi_t)^{-1} \\ &= -(d\varphi_t)^{-1}\, du_t \circ \varphi_t = (d\varphi_t^{-1}) \circ \varphi_t\, du_t \circ \varphi_t\ . \end{aligned}$$

We point out that the optimal control (20a) is written in the endpoint with the boundary condition (20b). Using the form of $\frac{d}{dt}((d\varphi_t)^{-1}\, d\varphi_t)$ gives $p^x_{i,t}$ in the endpoint satisfies the differential equation. For $p^w_{i,t}$ then constancy $p^w_{i,t}w_{i,t} = p^w_{i,1}w_{i,1}$ with $|d\varphi_{t,1}|(x_{i,t}) = |d\varphi_1|(x_i)|d\varphi_t^{-1}|(x_{i,t})$ gives the endpoint:

$$p^w_{i,t} = \frac{p^w_{i,1}w_{i,1}}{w_{i,t}} = \frac{p^w_{i,1}|d\varphi_1(x_i)|}{|d\varphi_t(x_i)|} = p^w_{i,1}|d\varphi_{t,1}|(x_{i,t})\ .$$

4.3. **Method of Gradient Matching Condition.** To make explicit the gradients $\nabla\boldsymbol{U}(\boldsymbol{q}_1)$ we introduce the duality bracket notation defined for all h smooth functions on $\mathbb{R}^d \times \mathcal{F}$,

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

The measure norm with associated RKHSs W with kernel K_W, gives norm-square

$$\|\mu\|^2_{W^*} = (\mu \mid K_W[\mu])\ .$$

For $\mu(q_1^\ell) = \sum_{i\in I_\ell} w_{i,1}^\ell \delta_{x_{i,1}^\ell} \otimes \delta_{f_i^\ell}$ for each scale $\ell < \ell_{max}$, with

$$h_{q_1^\ell}(x,f) := \alpha K_W[\mu(q_1) - \mu_{obs}^\ell](x,f),\ \text{on}\, \mathbb{R}^d \times \mathcal{F}^\ell\ ,$$

defines the endpoint

$$\boldsymbol{U}(\boldsymbol{q}_1) = \frac{1}{2}\sum_{\ell<\ell_{max}} \left(\mu(q_1^\ell) - \mu_{obs}^\ell \,\middle|\, h_{q_1^\ell}\right)\ .$$

Statement 3. *For h_q continuously differentiable in x and bounded, $\mathscr{C}_b^{1,0}$, the gradients of* (20b) *are given by,*

$$\begin{aligned} \nabla_{x_i^\ell}\boldsymbol{U}(\boldsymbol{q}_1) &= w_{i,1}^\ell\, \nabla_{x_i} h_{q_1^\ell}(x_{i,1}^\ell, f_i^\ell) \\ \nabla_{w_i^\ell}\boldsymbol{U}(\boldsymbol{q}_1) &= h_{q_1^\ell}(x_{i,1}^\ell, f_i^\ell) \end{aligned}\ . \tag{32}$$

Proof. The perturbation is over all scales $\boldsymbol{q}(\varepsilon) = (q^\ell(\varepsilon))_{\ell<\ell_{max}}$, $q_1^\ell(\varepsilon) = (x_{i,1}^\ell + \varepsilon\xi_i^{x,\ell}, w_{i,1}^\ell + \varepsilon\xi_i^{w,\ell})_{i\in I_\ell}$, which decomposes; the dependence on scale is implied, $q_1 = (x_{i,1}, w_{i,1})_{i\in I}$, $\mu(q_1) = \sum_{i\in I} w_{i,1}\delta_{x_{i,1}} \otimes \delta_{f_i}$:

$$\begin{aligned} \frac{d}{d\varepsilon}\frac{\alpha}{2}\|\mu(q_1(\varepsilon)) - \mu_{obs}\|^2_{W^*}|_{\varepsilon=0} &= \frac{d}{d\varepsilon}\frac{\alpha}{2}(\mu(q_1(\varepsilon)) - \mu_{obs} \mid K_W[\mu(q_1(\varepsilon)) - \mu_{obs}])|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon}(\mu(q_1(\varepsilon)) \mid \alpha\, K_W[\mu(q_1) - \mu_{obs}])|_{\varepsilon=0} \\ &= \sum_{i\in I}\frac{d}{d\varepsilon}\left((w_{i,1} + \varepsilon\xi_i^w)\, h_{q_1}(x_{i,1} + \varepsilon\xi_i^x, f_i)\right)|_{\varepsilon=0}\ . \\ &= \sum_{i\in I}\left(\langle w_{i,1}\nabla_{x_i} h_{q_1}(x_{i,1}, f_i), \xi_i^x\rangle + h_{q_1}(x_{i,1}, f_i)\xi_i^w)\right) = 0\ . \end{aligned}$$

□

The technical conditions to ensure that the brain measures are elements of W^* the dual of W is the kernels K_W are densely and continuously embedded in bounded continuous functions $\mathscr{C}_b(\mathbb{R}^d \times \mathcal{F}, \mathbb{R})$ so that the signed measure spaces $\mathcal{M}_s(\mathbb{R}^d \times \mathcal{F})$ of brains are continuously embedded in the dual spaces $\mathscr{C}_b(\mathbb{R}^d \times \mathcal{F}, \mathbb{R})^*$.

The Gaussians kernels (12) satisfy our technical conditions and determine the smoothness required for the measure norm $\|\mu\|^2_{W^*} = (\mu \mid K_W[\mu])$. For measures μ_0, μ_1 with particles indexed by i and j, respectively, the norm-square reduces to

$$\begin{aligned} \|\mu_0 - \mu_1\|^2_{W^*} = &\sum_{i,i'} w_i w_{i'} \exp\left(\frac{1}{2\sigma_x^2}|x_i - x_{i'}|^2\right)\exp\left(\frac{1}{2\sigma_f^2}|f_i - f_{i'}|^2\right) \\ &- 2\sum_{i,j} w_i w_j \exp\left(\frac{1}{2\sigma_x^2}|x_i - x_j|^2\right)\exp\left(\frac{1}{2\sigma_f^2}|f_i - f_j|^2\right) \\ &+ \sum_{j,j'} w_j w_{j'} \exp\left(\frac{1}{2\sigma_x^2}|x_j - x_{j'}|^2\right)\exp\left(\frac{1}{2\sigma_f^2}|f_j - f_{j'}|^2\right)\ . \end{aligned}$$

For data carrying position information but no feature values (such as tau tangle locations), each f_i, f_j is constant and the resulting exponential terms are all 1.

4.3.1. *The Tissue Continuum.* The tissue continuum has been studied extensively in MRI and Computational Anatomy with the continuum representation for the endpoint gradient $\nabla_{\boldsymbol{q}}\boldsymbol{U}$ useful in image matching in its discretized form on the computational lattices. The continuum measure has $\rho(dx) = \rho_c(x)dx$ and is given according to $\mu = \int_{\mathbb{R}^d} \rho_c(x)\delta_x \otimes \mu_x\, dx$. The state process $t \mapsto q_t := (\varphi_t, w_t)$ with

$$\mu(q_t) := \int_{\mathbb{R}^d} w_t(x)\rho_c(x)\delta_{\varphi_t(x)} \otimes \mu_x dx \ .$$

The initial condition $q_0 = (\mathrm{Id}, 1)$ implies $\mu(q_t) := \varphi_t \cdot \mu$. The average of h_q over the feature space determines the boundary term variation. Denote $\nabla_\varphi h_{q_1}(a,b)$ the gradient of h_{q_1} with respect to a,

$$\begin{array}{l} \bar{h}^{\nabla}_{q_1}(x) := \int_{\mathcal{F}} \nabla_\varphi h_{q_1}(\varphi_1(x), f)\mu_x(df) \\ \bar{h}_{q_1}(x) := \int_{\mathcal{F}} h_{q_1}(\varphi_1(x), f)\mu_x(df) \end{array} .$$

Statement 4. *The gradients are given by*

$$\begin{array}{l} \nabla_{\varphi^\ell}\boldsymbol{U}(\boldsymbol{\varphi}_1, \boldsymbol{w}_1) = w_1^\ell \rho_c^\ell \bar{h}^{\nabla}_{q_1^\ell} \\ \nabla_{w^\ell}\boldsymbol{U}(\boldsymbol{\varphi}_1, \boldsymbol{w}_1) = \rho_c^\ell \bar{h}_{q_1^\ell} \end{array} . \tag{33}$$

Proof. With $q_1 = (\varphi_1, w_1)$ take the variation $\varphi_1 \to \varphi_1(\varepsilon) = \varphi_1 + \varepsilon\delta\varphi$:

$$\begin{aligned} \frac{d}{d\varepsilon}\frac{\alpha}{2}\|\mu(\varphi_1(\varepsilon), w_1) - \mu_{obs}\|^2_{W^*}|_{\varepsilon=0} &= \frac{d}{d\varepsilon}\frac{\alpha}{2}(\mu(\varphi_1(\varepsilon), w_1) - \mu_{obs} \mid K_W[\mu(\varphi_1(\varepsilon), w_1) - \mu_{obs}])|_{\varepsilon=0} \ . \\ &= \frac{d}{d\varepsilon}(\mu(\varphi_1(\varepsilon), w_1) \mid h_{q_1})|_{\varepsilon=0} \\ &= \int_{\mathbb{R}^d\times\mathcal{F}} w_1(x)\rho_c(x)\frac{d}{d\varepsilon}h_{q_1}(\varphi_1(\varepsilon)(x), f)|_{\varepsilon=0}\mu_x(df)dx \\ &= \int_{\mathbb{R}^d} \langle w_1(x)\,\rho_c(x)\bar{h}^{\nabla}_{q_1}(x), \delta\varphi_1(x)\rangle dx = 0 \ . \end{aligned}$$

Take the variation $w_1 \to w_1(\varepsilon) = w_1 + \varepsilon\delta w$:

$$\begin{aligned} \frac{d}{d\varepsilon}\frac{\alpha}{2}\|\mu(\varphi_1, w_1(\varepsilon)) - \mu_{obs}\|^2_{W^*}|_{\varepsilon=0} &= \frac{d}{d\varepsilon}\frac{\alpha}{2}(\mu(\varphi_1, w_1(\varepsilon)) - \mu_{obs} \mid K_W[\mu(\varphi_1, w_1(\varepsilon)) - \mu_{obs}])|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon}(\mu(\varphi_1, w_1(\varepsilon)) \mid h_{q_1})|_{\varepsilon=0} \\ &= \int_{\mathbb{R}^d\times\mathcal{F}} \frac{d}{d\varepsilon}w_1(\varepsilon)(x)|_{\varepsilon=0}\rho_c(x)h_{q_1}(\varphi_1(x), f)\mu_x(df)dx \\ &= \int_{\mathbb{R}^d} \rho_c(x)\bar{h}_{q_1}(\varphi_1(x), f)\delta w_1(x)dx \ . \end{aligned}$$

We note that computing the variation $\frac{d}{d\varepsilon}U(\varphi_1(\varepsilon), w_1(\varepsilon))|_{\varepsilon=0}$ requires U as a function of φ is $\mathscr{C}^1$ for $\varphi \in G_k \subset \mathrm{Diff}^k_{\mathrm{id}}(\mathbb{R}^d, \mathbb{R}^d)$ when $k \geqslant 2$. □

5. Aknowledgement

This work was supported by the National Institutes of Health (NIH) (www.nih.gov) grants R01EB020062 (MM), R01NS102670 (MM), U19AG033655 (MM), and R01MH105660 (MM), the National Science Foundation (NSF) (www.nsf.gov) 16-569 NeuroNex contract 1707298 (MM), the Computational Anatomy Science Gateway (DT and MM) as part of the Extreme Science and Engineering Discovery Environment (XSEDE Towns et al., 2014), which is supported by the NSF grant ACI1548562, Johns Hopkins University Alzheimer's

Disease Research Center with NIH grant P50AG05146, the Dana Foundation's (www.dana.org) clinical neuroscience research program, and the Kavli Neuroscience Discovery Institute (kavlijhu.org) supported by the Kavli Foundation (www.kavlifoundation.org) (DT, MM, and JT). MM 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 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. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

APPENDIX A. PROOF OF STATEMENT 1

Since the states (x_t, w_t), co-states (p^x_t, p^w_t) controls u_t, v_t and flows φ_t are all functions of time in the statement we suppress it in the notation.

Appendix Statement 1. *We assume that $V^{\ell_{max}} = \mathscr{C}^m_0(\mathbb{R}^d, \mathbb{R}^d)$ with $m \geqslant 2$. If $\boldsymbol{u}_\cdot$ is a solution of the optimal control problem* (26) *then there exists time-dependent co-state $(t \mapsto \boldsymbol{p}_t)$ such that for any $i \in I_\ell$ for all ℓ:*

$$\left\{\begin{array}{rcl} \dot{q}^\ell_{i,t} & = & (u^\ell_t(x^\ell_{i,t}), w^\ell_{i,t} \operatorname{div} u^\ell_t(x^\ell_{i,t})) \\ \dot{p}^{x,\ell}_{i,t} & = & -(du_t)^T(x^\ell_{i,t})p^{x,\ell}_{i,t} - p^{w,\ell}_{i,t} w^\ell_{i,t} \nabla(\operatorname{div} u_t)(x^\ell_{i,t}) \\ \dot{p}^{w,\ell}_{i,t} & = & -p^{w,\ell}_{i,t} \operatorname{div} u_t(x^\ell_{i,t}) \end{array}\right. \tag{34a}$$

The optimal control satisfies $\partial_{u^\ell}\mathcal{H}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{u}) = 0$ and $\boldsymbol{v} = \boldsymbol{A}\boldsymbol{u}$ satisfies for any $\ell_0 < \ell_{max}$ (with convention $v^{\ell_{max}} = 0$):

$$\left\{\begin{array}{l} v^{\ell_0}(x) = u^{\ell_0}(x) - u^{\ell_0 - 1}(x) = \displaystyle\sum_{\ell_0 \leqslant \ell < \ell_{max}} \sum_{i \in I_\ell} \Big(g^{\ell_0}(x^\ell_i, x) p^{x,\ell}_i + p^{w,\ell}_i w^\ell_i \, \nabla_1 g^{\ell_0}(x^\ell_i, x) \Big) \\ u^{\ell_0}(x) = \displaystyle\sum_{\ell < \ell_{max}} \sum_{i \in I_\ell} \Big(\bar{g}^{\ell_0 \wedge \ell}(x^\ell_i, x) p^{x,\ell}_i + p^{w,\ell}_i w^\ell_i \, \nabla_1 \bar{g}^{\ell_0 \wedge \ell}(x^\ell_i, x) \Big) \ . \end{array}\right. \tag{34b}$$

where for any $\ell < \ell_{max}$, $L^\ell : V_\ell \to V_\ell{}^*$ is the isometry such that $\langle \tilde{u}^\ell, \tilde{v}^\ell \rangle_{V_\ell} = (L^\ell \tilde{u}^\ell \mid \tilde{v}^\ell)$ for any $\tilde{u}^\ell,\ \tilde{v}^\ell \in V_\ell$ with $(\cdot \mid \cdot)$ the dual bracket and $\nabla_1 g(a, b)$ denotes the gradient of g^ℓ with respect to the first variable a.

Proof. Under the assumption $V^{\ell_{max}} \hookrightarrow \mathscr{C}^2_0(\mathbb{R}^d, \mathbb{R}^d)$ then we have $(\boldsymbol{u}, \boldsymbol{q}) \mapsto \boldsymbol{\xi_q}(\boldsymbol{u})$ is $\mathscr{C}^2$ and standard results of optimal control theory apply establishing that $\dot{\boldsymbol{q}} = \partial_{\boldsymbol{p}}\mathcal{H}, \dot{\boldsymbol{p}} = -\partial_{\boldsymbol{q}}\mathcal{H}, \partial_{\boldsymbol{u}}\mathcal{H} = 0$. Taking the variation for $\partial_{q^\ell}\mathcal{H}$ of the Hamiltonian for scale ℓ varies $x^\ell_i(\varepsilon) = x^\ell_i + \varepsilon \xi^{x,\ell}_i$, $w^\ell_i(\varepsilon) = w^\ell_i + \varepsilon \xi^{w,\ell}_i$:

$$\frac{d}{d\varepsilon}\big(p^\ell \ \big| \xi(q^\ell(\varepsilon), u^\ell)\big)_{|\varepsilon=0} = \sum_{i \in I_\ell} \Big(\langle p^{x,\ell}_i, du^\ell(x^\ell_i)\xi^{x,\ell}_i \rangle + p^{w,\ell}_i w^\ell_i \langle \nabla(\operatorname{div} u^\ell)(x^\ell_i), \xi^{x,\ell}_i \rangle + p^{w,\ell}_i (\operatorname{div} u^\ell)(x^\ell_i)\xi^{w,\ell}_i \Big) \tag{35}$$

so that we get the last two equations of (29). To calculate $\partial_{u^\ell}\mathcal{H} = 0$, define $u^\ell(\varepsilon) = u^\ell + \varepsilon\xi$ implying $v^{\ell+1}(\varepsilon) = v^{\ell+1} - \varepsilon\xi$, $v^\ell(\varepsilon) = v^\ell + \varepsilon\xi$ for $\xi \in V_\ell$. We have

$$\begin{aligned} &\frac{d}{d\varepsilon}\left(\big(p^\ell \ \big| \xi(q^\ell, u^\ell(\varepsilon))\big) - \frac{1}{2}\sum_{j=\ell}^{\ell+1} \big(L^j v^j(\varepsilon) \ \big| v^j(\varepsilon)\big) \right)_{|\varepsilon=0} \\ &= \sum_{i \in I_\ell} \Big(\langle p^{x,\ell}_i, \xi(x^\ell_i) \rangle + p^{w,\ell}_i w^\ell_i (\operatorname{div} \xi)(x^\ell_i) \Big) - \big(L^\ell v^\ell - L^{\ell+1} v^{\ell+1} \ \big| \xi^u \big) = 0 \ . \end{aligned} \tag{36}$$

After summation of (36) for $\ell \geqslant \ell_0$, we get for any $\xi \in V^{\ell_0}$ that

$$(L^{\ell_0} v^{\ell_0} \mid \xi) = \sum_{\ell \geqslant \ell_0} \sum_{i \in I_\ell} \Big(\langle p^{x,\ell}_i, \xi(x^\ell_i) \rangle + p^{w,\ell}_i w^\ell_i (\operatorname{div} \xi)(x^\ell_i) \Big) \tag{37}$$

Now, for any $x, \alpha \in \mathbb{R}^d$, consider $\delta_x^\alpha \in V^{\ell_0 *}$ such that $(\delta_x^\alpha \mid \tilde{v}^\ell) = \langle \tilde{v}^\ell(x), \alpha \rangle_{\mathbb{R}^d}$ for any $\tilde{v}^{\ell_0} \in V^{\ell_0}$. The reproducing property on V^{ℓ_0} gives $\langle K^{\ell_0}\delta_x^\alpha,\ K^{\ell_0}\delta_y^\beta \rangle_{V^{\ell_0}} = g^{\ell_0}(x,y)\langle \alpha, \beta \rangle$. We get from (37) for $\xi = K^{\ell_0}\delta_x^\alpha = g^{\ell_0}(.,x)\alpha$ that

$$\langle v^{\ell_0}(x), \alpha \rangle = (L^{\ell_0} v^{\ell_0} \mid K^{\ell_0}\delta_x^\alpha) = \sum_{\ell \geqslant \ell_0} \sum_{i \in I_\ell} \Big(g^{\ell_0}(x_i^\ell, x)\langle p_i^{x,\ell}, \alpha \rangle + p_i^{w,\ell} w_i^\ell \langle \nabla_1 g^{\ell_0}(x_i^\ell, x), \alpha \rangle \Big)$$

so that we get the first equality above of (34b) for v^{ℓ_0}.

Now since $u^{\ell_0} = \sum_{k \leqslant \ell_0} v^k$ we deduce the equality for u^{ℓ_0} given in (34b). □

REFERENCES

[1] Special issue on computational anatomy and functional imagery. *NeuroImage*, 23, 2004.

[2] M. Albert, A. Soldan, R. Gottesman, G. McKhann, N. Sacktor, L. Farrington, M. Grega, R. Turner, B. Lu, S. Li, M.-C. Wang, O. Selnes, and B. R. Team. Cognitive changes preceding clinical symptom onset of mild cognitive impairment and relationship to apoe genotype. *Current Alzheimer Research*, 11(8):773âĂŞ784, 2014.

[3] W. Allard. On the first variation of a varifold. *Ann. of Math.*, 95:417–491, 1972.

[4] Y. Amit, U. Grenander, and M. Piccioni. Structural image restoration through deformable templates. *Journal of the American Statistical Association*, 86(414):376–381, 1991.

[5] J. Ashburner. Computational anatomy with the spm software. *Magnetic Resonance Imaging*, 27:1163–1174, October 2009.

[6] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes. Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, 2005.

[7] E. D. Boer and P. Kuyper. Triggered correlation. *IEEE Transactions on Bio-Medical Engineering*, 15(3), July 1968.

[8] N. Charon and A. Trouvé. The varifold representation of nonoriented shapes for diffeomorphic registration. *SIAM Journal on Imaging Sciences*, 6(4):2547–2580, 2013.

[9] K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, and X. Zhuang. Spatially resolved, highly multiplexed rna profiling in single cells. *Science*, 348(6233), 2015.

[10] N. Chuang, S. Mori, A. Yamamoto, H. Jiang, X. Ye, X. Xu, L. J. Richards, J. Nathans, M. I. Miller, A. W. Toga, et al. An mri-based atlas and database of the developing mouse brain. *Neuroimage*, 54(1):80–89, 2011.

[11] J. R. Ecker, D. H. Geschwind, A. R. Kriegstein, J. Ngai, P. Osten, D. Polioudakis, A. Regev, N. Sestan, I. R. Wickersham, and H. Zeng. The brain initiative cell census consortium: Lessons learned toward generating a comprehensive brain cell atlas. *Neuron*, 96(3):542 – 557, 2017.

[12] D. C. V. Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, and K. Ugurbil. The wu-minn human connectome project: An overview. *NeuroImage*, 80:62 – 79, 2013. Mapping the Connectome.

[13] A. Evans, D. Collins, S. Mills, E. Brown, R. Kelly, and T. Peters. 3d statistical neuroanatomical models from 305 mri volumes. page 1813âĂŞ1817, 1993.

[14] J. Fan, N. Salathia, R. Liu, G. E. Kaeser, Y. C. Yung, J. L. Herma, F. Kaper, J.-B. Fan, K. Zhang, J. Chun, and P. V. Kharchenko. Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. *Nature Methods*, 13:241âĂŞ244, 2016.

[15] H. Federer. *Geometric measure theory*. Number 153. Springer, 1969.

[16] B. Fischl, D. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. Dale. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–55, January 2002.

[17] J. F.J. Almgren and J. Taylor. The geometry of soap bubbles and soap films. *Scientific American*, pages 82–93, July 1976.

[18] G. Gerstein and N. Kiang. An approach to the quantitative analysis of electrophysiological data from single neurons. *Biophysical journal*, 1(1):15âĂŞ28, 1960.

[19] G. GR. A theorem about random fields. *Bulletin of the London Mathematical Society*, 5(1):81–84, 1973.

[20] U. Grenander. *General Pattern Theory*. Number ISBN 978-0198536710. Oxford Science Publications, 1994.

[21] U. Grenander and M. I. Miller. Computational anatomy: An emerging discipline. *Quarterly of Applied Mathematics*, 56(4):617–694, 1998.

[22] H.-W. Hsieh and N. Charon. Metrics, quantization and registration in varifold spaces. *Foundations of Computational Mathematics*, pages 1–45, 2021.

[23] S. KM, W. Z, X. E, L. KS, L. P, M. S, T. JC, S. A, A. MS, M. MI, and T. DJ. Quantifying patterns of pathology: A characterization of the 3d spatial distribution of tau tangles and amyloid plaques in alzheimerâĂŹs disease using multi-modal image registration with a scattering transform. 2021.

[24] S. Kulason, D. J. Tward, T. Brown, C. S. Sicat, C.-F. Liu, J. T. Ratnanather, L. Younes, A. Bakker, M. Gallagher, M. Albert, M. I. Miller, and the Alzheimer's Disease Neuroimaging Initiative. Cortical thickness atrophy in the transentorhinal cortex in mild cognitive impairment. *Neuroimage Clin.*, 21:101617, 2019.

[25] J. W. Lewis and D. C. Van Essen. Mapping of architectonic subdivisions in the macaque monkey, with emphasis on parieto-occipital cortex. *J Comp Neurol*, 428(1):79–111, 2000.

[26] J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, D. L. Collins, P. Thompson, M. MacDonald, D. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. Le Goualher, D. Boomsma, T. Cannon, , B. Kawashima, and R. Mazoyer. A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (ICBM). *Phil. Trans. R. Soc. Lond. B Biol. Sci.*, 356:1293–1322, 2001.

[27] J. Mazziotta, A. W. Toga, A. Evans, P. Fox, and J. Lancaster. Probabilistic atlas of the human brain: Theory and rationale for its development. *NeuroImage*, 2:89–101, 1995.

[28] M. Miller and L. Younes. Group Actions, Homeomorphisms, and Matching: A General Framework. *International Journal of Computer Vision*, 41(4):61–84, 2001.

[29] M. Miller, L. Younes, J. Ratnanather, T. Brown, H. Trinh, E. Postell, D. Lee, M.-C. Wang, S. Mori, R. O'Brien, and M. Albert. The diffeomorphometry of temporal lobe structures in preclinical alzheimer's disease. *NeuroImage: Clinical*, 3:352–360, 12 2013.

[30] M. I. Miller, S. Arguillere, D. J. Tward, and L. Younes. Computational anatomy and diffeomorphometry: A dynamical systems model of neuroanatomy in the soft condensed matter continuum. pages 1–24, 2018.

[31] M. I. Miller, A. Trouvé, and L. Younes. Geodesic Shooting for Computational Anatomy. *Journal of Mathematical Imaging and Vision*, 24(2):209–228, 2006.

[32] M. I. Miller, A. Trouvé, and L. Younes. Hamiltonian systems and optimal control in computational anatomy: 100 years since d'arcy thompson. *Annual Review of Biomed Engineering*, (17):447–509, November 4 2015.

[33] M. I. Miller, D. J. Tward, and A. TrouvÃl'. Coarse-to-fine hamiltonian dynamics of hierarchical flows in computational anatomy. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 3760–3765, 2020.

[34] J. R. Moffitt, D. Bambah-Mukku, S. W. Eichhorn, E. Vaughn, K. Shekhar, J. D. Perez, N. D. Rubinstein, J. Hao, A. Regev, C. Dulac, and X. Zhuang. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science*, 362(6416), 2018.

[35] S. Mori, S. Wakana, L. Nagae-Poetscher, and P. C. V. Zijl. Elsevier, 2005.

[36] M. P. Nagle, G. S. Tam, E. Maltz, Z. Hemminger, and R. Wollman. Bridging scales: From cell biology to physiology using in situ single-cell technologies. *Cell systems*, 12(5):388âĂŤ400, May 2021.

[37] M. Ono, S. Kubik, and C. D. Abernathey. *Atlas of the Cerebral Sulci*. Georg Thieme Verlag, 1990.

[38] I. Peikon, J. Kebschull, V. Vagin, D. Ravens, Y. Sun, E. Brouzes, I. J. CorrÃła, D. Bressan, and A. Zador. Using high-throughput barcode sequencing to efficiently map connectomes. *Nucleic Acids Res.*, 12(45), July 7 2017.

[39] X. Pennec. From Riemannian Geometry to Computational Anatomy. *Elements*, 2011.

[40] L. Risser, F.-X. Vialard, R. Wolz, M. Murgasova, D. D. Holm, and D. Rueckert. Simultaneous Multi-scale Registration Using Large Deformation Diffeomorphic Metric Mapping. *IEEE Transactions on Medical Imaging*, 30(10):1746–59, 2011.

[41] S. G. Rodriques, R. R. Stickels, A. Goeva, C. A. Martin, E. Murray, Vanderburg, C. R., J. Welch, L. M. Chen, F. Chen, and E. Z. Macosko. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science*, 363(6434):1463–1467, 2019.

[42] N. Sacktor, A. Soldan, M. Grega, F. L, Q. Cai, M. Wang, R. Gottesman, R. Turner, A. M, and B. R. Team. The biocard index: A summary measure to predict onset of mild cognitive impairment. alzheimer dis assoc disord. *Alzheimer Disease Association Disord*, 31(2):114–119, Apr-Jun 2017.

[43] R. Satij, J. A. Farrell, D. Gennert, F. Schier, Alexander, and A. Regev. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology*, 33, 2015.

[44] S. Shah, E. Lubeck, W. Zhou, and L. Cai. In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron*, 92:342–357, 10 2016.

[45] D. Snyder and M. Miller. *Random Point Processes in Time and Space*. Number ISBN-13: 978-1461278214. Springer-Verlag, 1991.

[46] S. Sommer, F. Lauze, M. Nielsen, and X. Pennec. Sparse multi-scale diffeomorphic registration: The kernel bundle framework. *Journal of Mathematical Imaging and Vision*, 46(292), 2013.
[47] A. Srivastava, W. Wu, S. Kurtek, E. Klassen, and J. S. Marron. Registration of functional data using fisher-rao metric, 2011.
[48] G. J. Srulovicz P. A central spectrum model: a synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum. *J Acoust Soc Am.*, 4(73):1266–76, April 1983.
[49] P. L. Ståhl, F. Salmén, S. Vickovic, A. Lundmark, J. F. Navarro, J. Magnusson, S. Giacomello, M. Asp, J. O. Westholm, M. Huss, A. Mollbrink, S. Linnarsson, S. Codeluppi, Å. Borg, F. Pontén, P. I. Costea, P. Sahlén, J. Mulder, O. Bergmann, J. Lundeberg, and J. Frisén. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, 353(6294):78–82, 2016.
[50] S. M. Sunkin, L. Ng, C. Lau, T. D. andTerri L. Gilbert, C. L. Thompson, M. Hawrylycz, and C. Dang. Allen brain atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Res.*, D996âĂŞD1008(41(Database issue)), Jan 2013.
[51] J. Talairach and P. Tournoux. *Co-Planar Stereotaxis Atlas of the Human Brain*. Georg Thiem Verlag, Stuttgart, 1988.
[52] P. M. Thompson and A. W. Toga. A framework for computational anatomy. *Computing and Visualization in Science*, 5(1):13–34, 2002.
[53] D. Tward, T. Brown, Y. Kageyama, J. Patel, Z. Hou, S. Mori, M. Albert, J. Troncoso, and M. Miller. Diffeomorphic registration with intensity transformation and missing data: Application to 3d digital pathology of alzheimer's disease. *Frontiers*, 2019.
[54] D. Tward, X. Li, B. Huo, B. Lee, P. Mitra, and M. Miller. 3d mapping of serial histology sections with anomalies using a novel robust deformable registration algorithm. volume LNCS, volume 11846, October 10 2019.
[55] Q. Wang, S.-L. Ding, Y. Li, J. Royall, D. Feng, P. Lesnar, N. Graddis, M. Naeemi, B. Facer, A. Ho, et al. The allen mouse brain common coordinate framework: a 3d reference atlas. *Cell*, 181(4):936–953, 2020.
[56] X. Wang*, W. Allen*, M. Wright, E. Sylwestrak, N. Samusik, S. Vesuna, K. Evans, C. Liu, C. Ramakrishnan, J. Liu, G. P. Nolan, F.-A. Bava, and K. Deisseroth. Three-dimensional intact-tissue-sequencing of single-cell transcriptional states. *Science*, 2018.
[57] W. WM, G. WL, K. R, and J. FA. Adaptive segmentation of mri data. *IEEE Trans Med Imaging*, 4(15):429–42, 1996.
[58] C. Xia, J. Fan, G. Emanuel, J. Hao, and X. Zhuang. Spatial transcriptome profiling by merfish reveals subcellular rna compartmentalization and cell cycle-dependent gene expression. *Proceedings of the National Academy of Sciences*, 116(39):19490–19499, 2019.
[59] A. Xiaowei. Method of the year 2020: spatially resolved transcriptomics. *Nature Methods*, 18(1), 2021.
[60] Z. Y, B. M, and S. S. Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging*, 20(1):45–57, January 2001.
[61] L. Younes, M. Albert, and M. I. Miller. Inferring changepoint times of medial temporal lobe morphometric change in preclinical alzheimer's disease. *NeuroImage: Clinical*, 5:178 – 187, 2014.
[62] L. Younes, M. Albert, A. Moghekar, A. Soldan, C. Pettigrew, and M. I. Miller. Identifying changepoints in biomarkers during the preclinical phase of alzheimerâĂŹs disease. *Frontiers in Aging Neuroscience*, 11:74, 2019.
[63] S. M. Young ED. Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. *The Journal of the Acoustical Society of America*, (66):1381–1403.
[64] E. Zhao, M. R. Stone, X. Ren, T. Pulliam, P. Nghiem, J. H. Bielas, and R. Gottardo. Bayesspace enables the robust characterization of spatial gene expression architecture in tissue sections at increased resolution. *bioRxiv*, 2020.

DEPARTMENT OF BIOMEDICAL ENGINEERING & KAVLI NEUROSCIENCE DISCOVERY INTITUTE & CENTER OF IMAGING SCIENCE, THE JOHN HOPKINS UNIVERSITY, 301 CLARK HALL, BALTIMORE, MD 21218
E-mail address, M. I. Miller: mim@cis.jhu.edu

UNIVERSITY OF CALIFORNIA AT LOS ANGELES
E-mail address, D. Tward: dtward@mednet.ucla.edu

CENTRE GIOVANNI BORELLI (UMR 9010), ECOLE NORMALE SUPÉRIEURE PARIS-SACLAY, UNIVERSITÉ PARIS-SACLAY, 4 RUE DES SCIENCES, F-91190 GIF SUR YVETTE, FRANCE
E-mail address, A. Trouvé: alain.trouve@ens-paris-saclay.fr