



Preserving Derivative Information while Transforming Neuronal Curves

Thomas L. Athey^{1,2} · Daniel J. Tward^{3,4} · Ulrich Mueller⁵ · Laurent Younes^{2,6,7,8} · Joshua T. Vogelstein^{1,2,7,8} · Michael I. Miller^{1,2,7,8}

Accepted: 31 October 2023 / Published online: 30 November 2023
© The Author(s) 2023

Abstract

The international neuroscience community is building the first comprehensive atlases of brain cell types to understand how the brain functions from a higher resolution, and more integrated perspective than ever before. In order to build these atlases, subsets of neurons (e.g. serotonergic neurons, prefrontal cortical neurons etc.) are traced in individual brain samples by placing points along dendrites and axons. Then, the traces are mapped to common coordinate systems by transforming the positions of their points, which neglects how the transformation bends the line segments in between. In this work, we apply the theory of jets to describe how to preserve derivatives of neuron traces up to any order. We provide a framework to compute possible error introduced by standard mapping methods, which involves the Jacobian of the mapping transformation. We show how our first order method improves mapping accuracy in both simulated and real neuron traces under random diffeomorphisms. Our method is freely available in our open-source Python package *brainlit*.

Keywords Neuron reconstruction · Morphology · Registration · Splines · Diffeomorphisms

Main

The brain functions as a network of chemical and electrical activity, so identifying how neurons connect across brain regions is central to understanding how the brain works, and how to treat brain diseases. Modern neuroscience techniques can image single neuron morphology at scale (Economo et al.,

2016), and subsequent neuron tracing can help discover new morphological subtypes (Winnubst et al., 2019). Due to anatomical variation, and deformations that may have occurred during tissue preparation, neuron traces need to be mapped between coordinate spaces to compare morphologies from different brain samples. Brain registration software often includes neuron mapping implementations, but these implementations have not been thoroughly characterized from a numerical analysis perspective.

This question is relevant to the ongoing work of the international neuroscience community, including the Brain Initiative Cell Census Network (BICCN), to establish comprehensive neuronal atlases of the mammalian brain (BRAIN Initiative Cell Census Network, 2021). This effort has produced many images of stained or fluorescently labeled brains, which are being used to identify whole-brain connectivity patterns. Sometimes this data is analyzed with density-based methods to measure connectivity between brain regions (Watakabe et al., 2023; Athey et al., 2023). In other cases, when individual neurons can be resolved, it is possible to generate digital neuron traces for morphological analysis (Skibbe et al., 2018; Athey et al., 2022). This paper focuses on neuron traces, which are commonly stored as a set of connected 3D coordinates, or knots, such as in the SWC format (Stockley et al., 1993; Cannon et al., 1998). The connections between the knots are classically represented as cylinders (Cannon et al., 1998), or conical

✉ Thomas L. Athey
tatheyl@jhu.edu

¹ Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA

² Institute of Computational Medicine, Johns Hopkins University, Baltimore, MD, USA

³ Department of Computational Medicine, University of California at Los Angeles, Los Angeles, CA, USA

⁴ Department of Neurology, University of California at Los Angeles, Los Angeles, CA, USA

⁵ Department of Neuroscience, Johns Hopkins University, Baltimore, MD, USA

⁶ Department of Applied Mathematics & Statistics, Johns Hopkins University, Baltimore, MD, USA

⁷ Center for Imaging Science, Johns Hopkins University, Baltimore, MD, USA

⁸ Kavli Neuroscience Discovery Institute, Johns Hopkins University, Baltimore, MD, USA

frustums (O'Halloran, 2020), but here we ignore radius information, since it is not generated by all neuron tracing methods. Consequently, the whole neuron trace is considered to be a tree of piecewise linear curves.

In order to assemble these traces into a complete picture of the various neuron morphologies in the brain, scientists need a way to map neuron traces into common coordinate systems. Several popular software applications exist for this task and are used to assemble atlases of neuron morphology. For example, Peng et al. (2021) used mBrainAligner (Qu et al., 2022), Gao et al. (2022) used the Computational Morphometry Toolkit, and the MouseLight project (Winnubst et al., 2019) used displacement fields from Fedorov et al. (2012). Existing methods use what we call *zeroth order* curve mapping in that they only map the positions of the knots (also known as trace points). However, depending on the nonlinearity of the mapping, and the continuous representation of the neuron trace, zeroth order mapping is sensitive to different samplings of the original neuronal curve (Fig. 1a,b). In other words, sampling the same curve different ways while tracing in the original image may lead to different mapped morphologies. It is critical that neuron mapping methods preserve the

geometry of digital neuron traces in order to build reliable atlases of neuron morphology, and to accurately identify deviations in diseased brains.

In this work, we introduce a method to preserve derivative information when mapping neuronal curves, and investigate the conditions under which this technique is advantageous to existing methods (Figure 1). We applied our method to both simulated data and real neuron traces from a whole mouse brain image, and the code used developed in this work is freely available in our Python package *brainlit*.

Results

Action of Diffeomorphisms on Discrete Samplings

In the following sections, we use C^k to represent the space of continuous functions with k continuous derivatives, where the domain and range can be inferred by the context. We model a neuronal branch (dendrite or axon) as a regular 3D curve $c : [0, L] \rightarrow \mathbb{R}^3$, $c \in C^k$, and $|\dot{c}| > 0$, where L , without loss of generality, is the arc length of the curve (Younes, 2010). When a neuronal curve is traced, it is typically stored

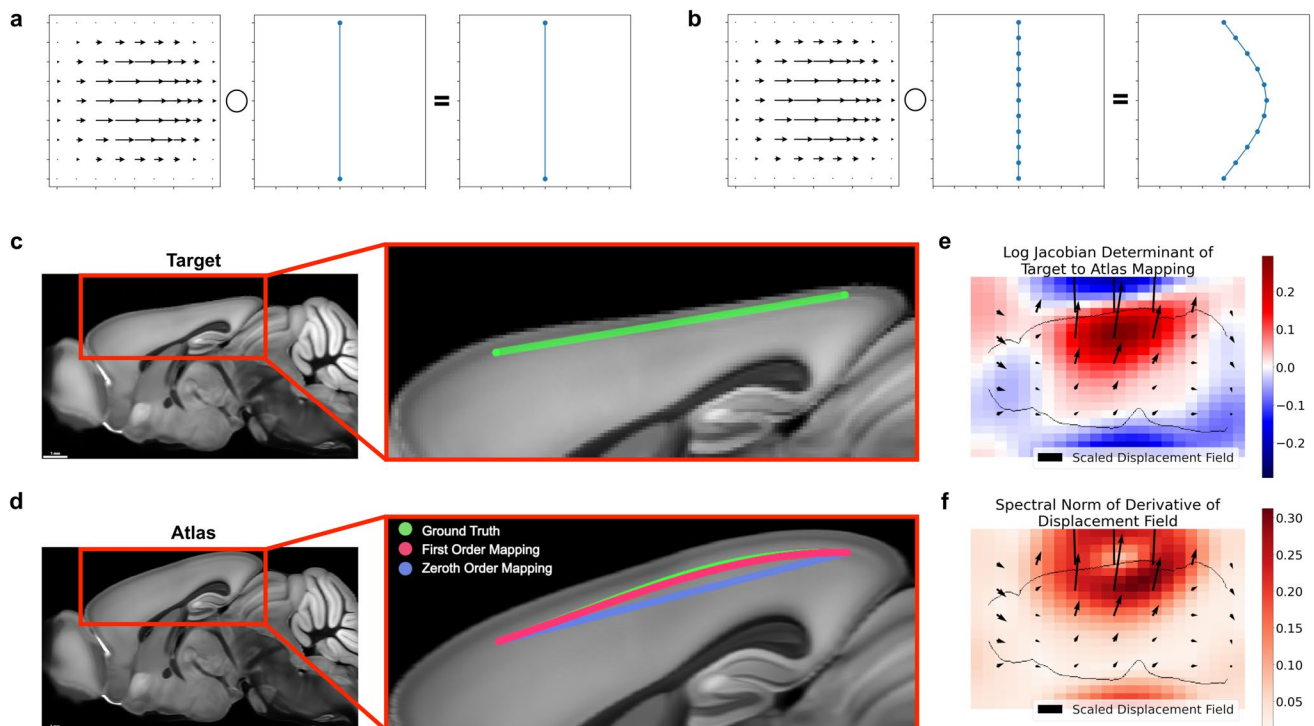


Fig. 1 Neglecting the action of a nonlinear mapping on a curve's derivatives can introduce errors. **a–b** Different samplings of a curve can lead to different results under nonlinear deformations, such as only sampling the endpoints (**a**) versus sampling several times along the curve (**b**). **c–d** Large distances between trace points can contribute to mapping inaccuracies. The green line segment following cortical layers 2/3 in a synthetic mouse brain image (**c**) is defined only by its

endpoints. Transforming only the positions of the endpoints (zeroth order mapping, **d**), is less accurate than incorporating the action on the derivatives as well (first order mapping, **d**). **e–f** Quantitative descriptions of the mapping from target to atlas via the logarithm of the Jacobian determinant, which quantifies expansion and compression (**e**), and the spectral norm of the displacement field, which plays a role in an error bound of zeroth order mapping (**f**)

as a sequence of points $\{x_i = c(t_i) : t_i < t_{i+1}\}_{i=1}^n$, where the independent variables t_i can be taken to be the indices of the points. When there is a diffeomorphism between coordinate systems $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, these traces are mapped via the group action:

$$\phi \cdot \{x_i\}_{i=1}^n = \{\phi(x_i)\}_{i=1}^n$$

We want to extend the space of traces, and the associated action, to include derivatives of the underlying curve denoted $\partial_t c$. This can be done using the jet space J^k . In our setting, $J^k = [0, L] \times X^{(k)}$, where an element of $X^{(k)}$ is a $k + 1$ -tuple $(x^0, x^1, \dots, x^k) \in (\mathbb{R}^3)^{k+1}$ representing a position and first k derivatives of a curve in $\mathbb{R}^3$. A C^k curve $c : [0, L] \rightarrow \mathbb{R}^3$ can be extended to a curve $\hat{c} : [0, L] \rightarrow X^{(k)}$ simply by adding derivatives, with $\hat{c}(t) = (c(t), \partial_t c(t), \dots, \partial_t^k c(t)) \in X^{(k)}$ (Olver, 1995).

The C^k diffeomorphisms have a natural group action on the jet space J^k , ensuring the commutation between the standard action of diffeomorphisms on curves, $(\phi, c) \mapsto \phi \circ c$ and their extensions, such that the identity $\phi \cdot \hat{c}(t) = \widehat{\phi \circ c}(t)$ holds for all curves c and times t , defining the left-hand side. For example, for $k = 2$, this provides

$$\phi \cdot (t, x^0, x^1, x^2) = (t, \phi(x^0), D\phi(x^0)x^1, D\phi(x^0)x^2 + D^2\phi(x^0)(x^1, x^1))$$

Neuron traces, as mentioned before, involve a sequence of samples with time-stamps $\{(t_i, x_i^{(k)})\}_{i=1}^n$, identified as elements of $(J^k)^n$, the n -fold Cartesian product of J^k . Our diffeomorphisms will act on such a sequence as follows:

Statement 1 For a sequence of time-stamped elements on the jet space, $T = \{(t_i, x_i^{(k)})\}_{i=1}^n$ in $(J^k)^n$, we define the action of diffeomorphisms

$$\phi \cdot T = \{(t_i, \phi \cdot x_i^{(k)})\}_{i=1}^n \quad (1)$$

The fact that this operation provides an action is an established result (Olver, 1995), and the proof is provided in the Supplement. We will define k 'th order discrete mapping to be the action in Eq. 1 of a diffeomorphism on a curve sampling that includes k derivatives. The axioms that define group actions are important to verify because they ensure that applying the identity transformation does not change the object, and that applying a composition of transformations is equivalent to applying the individual transformations successively. Further, group actions can exchange mathematical structure between the acting group and the set being acted upon, and they are at the core of several important theorems (Suksumran, 2016).

The k 'th order discrete mapping method allows us to compute the first k derivatives of the transformed curve. We will interpolate the transformed curve using splines of order $2k + 1$ that satisfy the derivative values. For example,

zeroth order mapping will produce a first order spline and first order mapping will produce a cubic Hermite spline (Spitzbart, 1960).

Error Analysis of Zeroth and First Order Mapping

Now we will examine the error introduced by zeroth order mapping, which is used by existing neuron mapping methods. First, note that under affine transformations, zeroth order mapping of piecewise linear curves introduce no error, so these results are only useful under non-affine transformations. The following results require that the curve c be parameterized by arc length. However, we note that all continuously differentiable regular curves can be reparameterized by arc length (Smale, 1958). We use $|\cdot|$ to denote the Euclidean norm for elements of $\mathbb{R}^d$, and the spectral norm for matrices.

Proposition 1 [Zeroth Order Mapping Error Bound] Say $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is a C^1 diffeomorphism and $c : [0, L] \rightarrow \mathbb{R}^3$ is a continuous, piecewise linear curve parameterized by arc length with knots $\{t_i : t_1 = 0, t_n = L, t_{i-1} < t_i\}_{i=1}^n$. For the transformed curve $f = \phi \circ c$, the zeroth order mapping defines a first order spline g which satisfies:

$$\max_{t \in [0, L]} |f(t) - g(t)| \leq \max_{i \in \{0, \dots, n\}, t \in [t_{i-1}, t_i]} \frac{1}{2} (|D\phi \circ c(t) - I| |t_i - t_{i-1}| + |\epsilon_i - \epsilon_{i-1}|) \quad (2)$$

where $\epsilon_i \triangleq c(t_i) - \phi(c(t_i))$ and $D\phi \circ c(t)$ is the Jacobian of ϕ evaluated at $c(t)$.

This shows how the error introduced by the state of the art mapping method is related to the displacement magnitude, ϵ , and the extent to which the Jacobian of the transformation, $D\phi$, differs from the identity matrix. Note that the bound in Eq. 2 goes to zero as ϕ approaches the identity map (in which case zeroth order mapping has zero error for piecewise linear curves). It depends on the arc lengths of the original curve segments and the spectral norm of $D\phi$, which is related to the finite time Lyapunov exponent ($\log |D\phi|$), a well-known quantity in field dynamics which characterizes the amount of stretching in a differentiable flow. Also, the bound applies to $\max_{t \in [0, L]} |f(t) - g(t)|$, which is not parameterization invariant, and therefore not a strictly geometric quantity. However we note that this quantity is an upper bound of the Frechet distance, which is parameterization invariant.

In this paper we demonstrate first order mapping in an effort to mitigate this mapping error. Such a method has the advantage of having superior error convergence at the knots as a consequence of Taylor's theorem. Further, we present a set of error bounds that helps clarify the advantage of first order mapping.

Proposition 2 [Comparable Bounds for Zeroth and First Order Mapping] Say $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is a C^4 diffeomorphism and $c : [a, b] \rightarrow \mathbb{R}^3$ is a continuous, piecewise C^4 curve parameterized with knots $\{t_i : t_1 = a, t_n = b, t_{i-1} < t_i\}_{i=1}^n$. For the transformed curve $f = \phi \circ c$ defined by coordinate functions $f = (f^0, f^1, f^2)^T$, the zeroth order mapping defines a first order spline g_0 which satisfies:

$$\begin{aligned} \max_{t \in [a, b]} |f(t) - g_0(t)| &\leq \frac{\sqrt{3}}{4} \max_{t \in [a, b], j \in \{0, 1, 2\}} |\partial_t^{(4)} f^j(t)| \left(\frac{\delta}{2}\right)^4 + \\ &\quad \frac{\sqrt{3}}{2} \left(\frac{\delta}{2}\right)^2 \max_{i \in \{1 \dots n\}, j \in \{0, 1, 2\}} |\partial_t^{(3)} f^j(t_i)| \left(\frac{\delta}{2}\right) + \\ &\quad \frac{\sqrt{3}}{2} \left(\frac{\delta}{2}\right)^2 \max_{i \in \{1 \dots n\}, j \in \{0, 1, 2\}} |\partial_t^{(2)} f^j(t_i)| \end{aligned} \quad (3)$$

where $\delta \triangleq \max_{2 \leq i \leq n} |t_i - t_{i-1}|$ and $\partial_t^{(k)} f^j(t)$ is the k 'th derivative of f^j evaluated at t . Also, the first order mapping defines a third order spline g_1 , which satisfies

$$\max_{t \in [a, b]} |f(t) - g_1(t)| \leq \frac{\sqrt{3}}{4!} \max_{t \in [a, b], j \in \{0, 1, 2\}} |\partial_t^{(4)} f^j(t)| \left(\frac{\delta}{2}\right)^4 \quad (4)$$

and we note that the bound in (4) is tighter than the bound in (3). Further, there exists a transformed curve f and a set of knots $\{t_i\}_{i=1}^n$ that achieves both bounds exactly.

Thus, we have made a connection between the state of the art (zeroth order mapping) and a higher order method (first order mapping) via worst-case bounds on mapping error. The error bound for first order mapping is smaller than that for zeroth order mapping, though for any given curve, either method may produce smaller error than the other. Proofs for the propositions are in the supplement.

Software Implementation

We implemented both zeroth and first order discrete mapping in our open-source Python package *brainlit*. For first order mapping, we compute one-sided derivatives at the knots of the curve from first order splines in accordance with original SWC formulation (Stockley et al., 1993; Cannon et al., 1998). Then, once the knot positions and derivatives are transformed, we generate a continuous curve in the new space using Hermite interpolation. Further details of our implementation can be found in the Methods.

First order mapping involves more computations than zeroth order mapping. First, if the Jacobian of the transformation, $D\phi$, is not immediately available, it needs to be approximated. Our software uses a finite difference method which, to approximate $D\phi(x)$, involves calling ϕ four times, three vector addition operations, and three vector scaling

operations. Next, the derivatives at the knots need to be approximated and transformed by $D\phi$. For each line segment, our method computes one vector addition and one vector scaling to approximate one-sided derivatives, and computes two matrix-vector products to transform the derivatives (Algorithm 3). Lastly, generating and evaluating cubic splines involves more computations than first-order splines (Kincaid & Cheney, 2002). Despite these differences, first order mapping still scales linearly with the number of trace nodes, and the number of computations differs from zeroth order mapping by a constant factor of computations. In practice, both zeroth order and first order mapping take on the order of seconds for traces with thousands of nodes, which should not be a bottleneck in neuron morphology studies.

Figure 2 shows examples of our method on simulated data, compared to the zeroth order method, and the “ground truth” where we map a dense sampling of points along the first order spline of the original curve.

Application to Real Neurons

We applied our method to 20 reconstructed neurons in SWC format from a whole mouse brain image from the Janelia MouseLight project (Winnubst et al., 2019). We selected the first 20 SWC files that successfully downloaded from MouseLight’s NeuronBrowser repository and did not have repeat trace nodes. Neurons have a tree-like structure, and we split them into non-branching curves in order to apply our mapping methods. We follow a method introduced previously (Athey et al., 2021) where the root to leaf path with the longest arc length is recursively removed until the tree is reduced to non-bifurcating “branches”. The transformed branches are then reconnected to maintain the topology of the original trace, and therefore our methods generalize naturally from 3D curves, to the branching structure of neuron traces.

We generate random transformations using the Large Deformation Diffeomorphic Metric Mapping (LDDMM) framework described in Miller et al. and applied in Tward and Miller (Miller et al., 2006; Tward & Miller, 2017). We generate an initial momentum field by sampling Gaussian noise with zero mean and varying standard deviation, σ . The momentum is smoothed to construct a velocity field, and integrated in time according to the conservation laws established in Miller et al. to generate a diffeomorphic transformation (Miller et al., 2006). We generated four diffeomorphisms with σ levels of 80, 160, 320 and 640 $\mu\text{m}/\text{time}$. The position and tangent displacement profiles of these four diffeomorphisms are shown in Figure 3a. We centered the neuron traces at the origin then applied the random diffeomorphisms to compare zeroth and first order mapping to ground truth (Fig. 3b-g). Ground truth was generated by upsampling the original traces to a maximum node spacing of $2\mu\text{m}$ followed by zeroth order mapping.

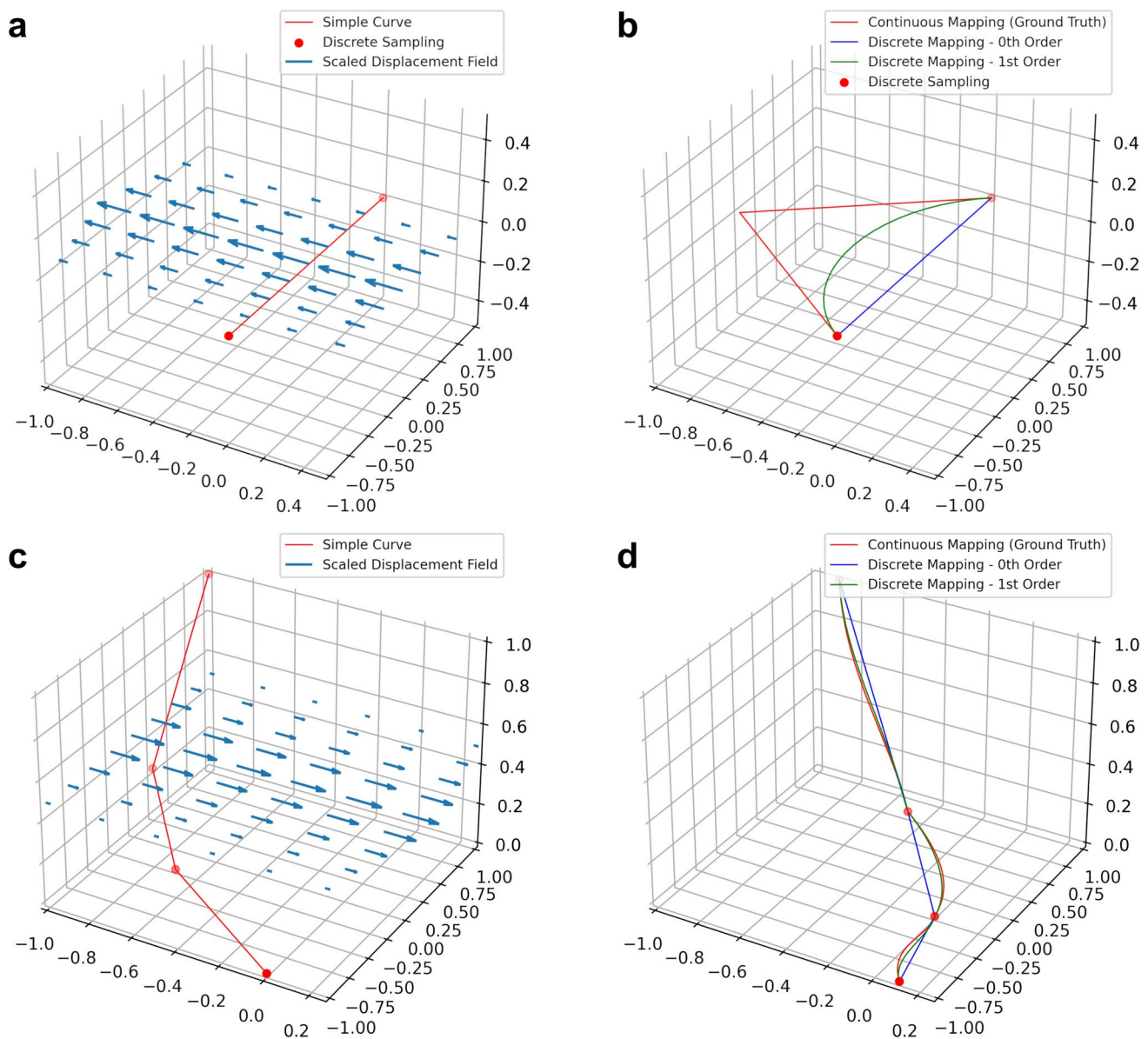


Fig. 2 Preserving derivative information can mitigate errors when transforming discretized curves. **a–b** Applying a nonlinear deformation field to a single line segment (**a**) using zeroth and first order mapping (**b**). **c–d** Applying a nonlinear deformation field to a piecewise

linear curve (**c**) using zeroth and first order mapping (**d**). Zeroth and first order discrete mapping methods are shown relative to ground truth considered to be the application of the vector field to a dense sampling of the original curves

For each neuron trace, we computed the discrete Frechet error from ground truth (Fig. 4a). We also wanted to measure which mapping method better matched the ground truth with respect to a neuron's distribution of common morphometric quantities, such as path angle, branch angle, tortuosity, and segment length. We used the Kolmogorov-Smirnov test statistic to measure how much the distribution of these quantities differed from ground truth (Fig. 4b). We performed two-sided Wilcoxon signed-rank tests for each comparison and used a Bonferroni correction across the different σ values (Fig. 4b). Lastly, we compared the discrete

Frechet errors to the average sampling period of the trace i.e. the average distance between trace nodes (Fig. 4c).

To explore the effect of downsampling neuron traces on mapped morphologies, we identified non-branching nodes in straight portions of the trace, and measured the impact of removing those nodes from the trace. Specifically, we performed both zeroth and first order mapping on the segment with the node removed, and compared it to the ground truth mapping of the original segment. We determined which fraction of nodes maintained a discrete Frechet error less than one micron, serving as an estimate for the fraction of nodes

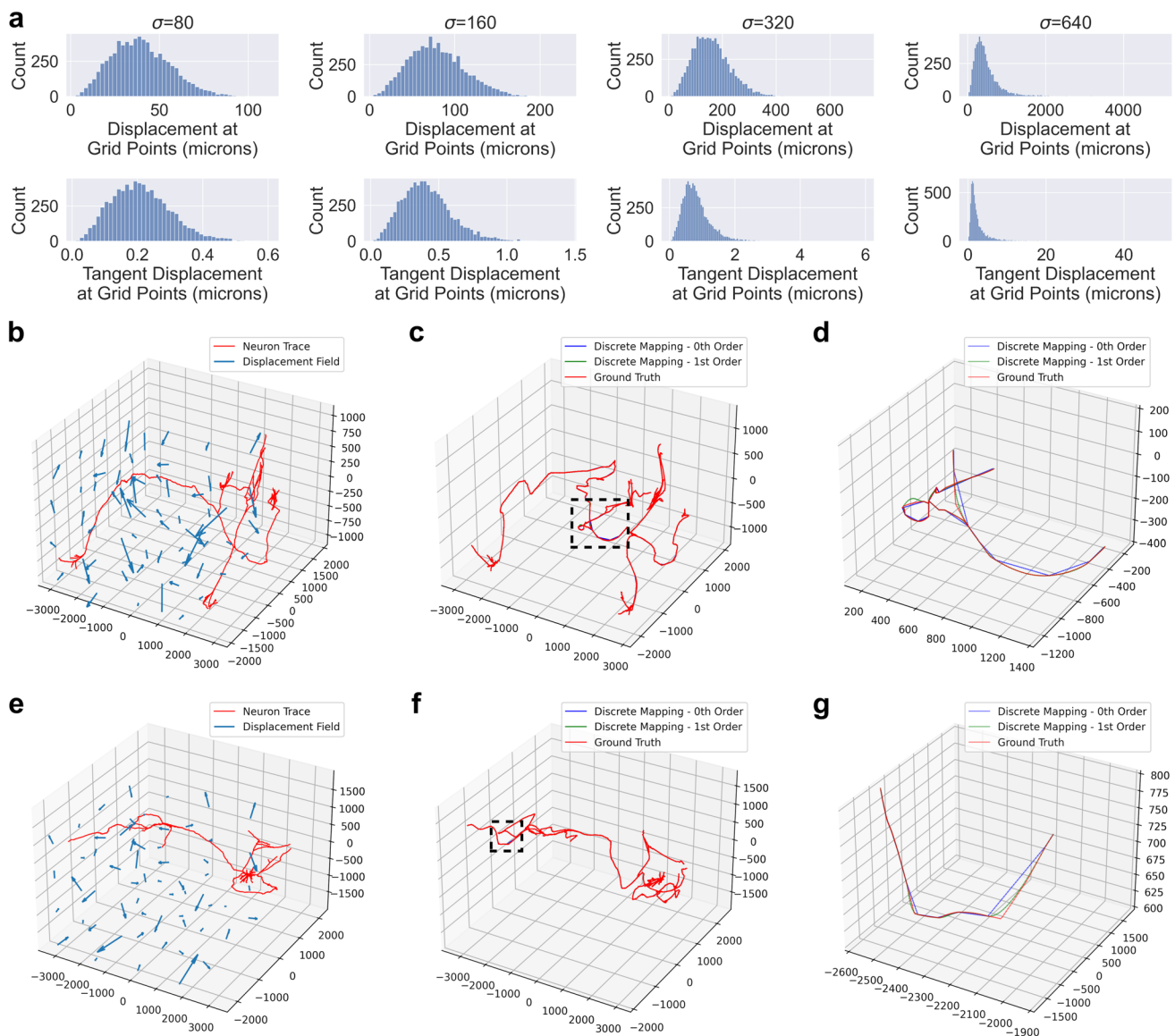


Fig. 3 Application of zeroth and first order mapping of neuron traces under diffeomorphisms derived from random Gaussian initial momenta. **a** Different values of σ produced diffeomorphisms with different position and tangent vector displacement profiles. The positions and tangent vectors sampled in the histogram were distributed as a uniform grid with a spacing of $500\mu\text{m}$. **b-g** Two examples of the diffeomorphism with

$\sigma = 640$ applied to neuron traces to produce zeroth and first order mappings, along with ground truth. Both examples show the original trace and the transformation (**b, e**), the results of the different transformation methods (**c, f**), and a zoomed in view of the region outlined by the dotted line to show discrepancies between the methods (**d, g**). Plot axes are in units of microns

which are not necessary to maintain the mapped morphology (Fig. 5c). We performed two-sided Wilcoxon signed-rank tests to compare zeroth and first order methods and used a Bonferroni correction across the different values of σ .

Discussion

In this paper we examine the “naive” approach to mapping discretely sampled one-dimensional structures by simply transforming the positions of the knots, i.e. mapping line segments to line segments. We show that this method can

be inaccurate when the Jacobian of the transformation is non-constant. We describe how to preserve derivative information which will lead to more accurate mappings in neighborhoods of the knots. We offer an implementation of a first-order mapping technique which, empirically, is more accurate on discretely sampled differentiable curves. Our mathematical framework is described in terms of 3D curves, but we also show how our method can handle branching. We apply our method to real neuron reconstructions and show that it more accurately matches ground truth in both frechet error, and a variety of morphometric quantities.

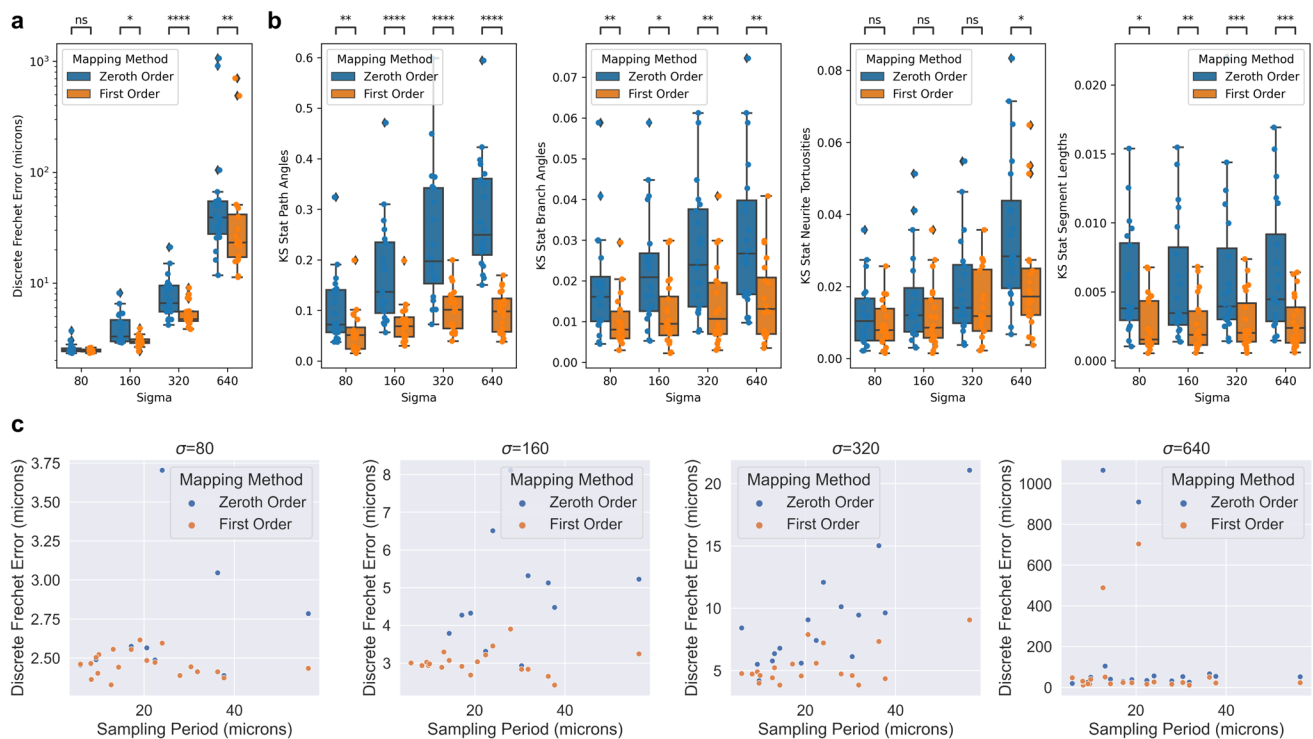


Fig. 4 Comparison of zeroth and first order mapping of neuron traces under random diffeomorphisms. **a** Discrete Frechet error was computed between the different order mappings, and ground truth. **b** Distributions of common morphometric quantities were compared to that of ground truth using the Kolmogorov-Smirnov test statistic. Differences between zeroth and first order methods were tested using Wilcoxon signed-rank test with Bonferroni correction across

different values of σ (*: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$, ****: $p \leq 0.0001$). Box plots show median, upper and lower quartiles and whiskers have a maximum length of 1.5x the interquartile range with other outlier data marked with points. **c** Relationship between discrete Frechet error and average sampling period (distance between trace points) under the random diffeomorphisms

In our experiment with real neuron reconstructions, it is important to note what we are considering ground truth. Since the original reconstructions are in SWC format, only the knot positions are known, and the neurons are typically represented as piecewise linear structures. Real neuron morphologies are not piecewise linear, and instead are continuously curving as they pass through dense brain tissue. Nonetheless, because we have no further information about the neuron trajectories, we consider the original reconstructions to be piecewise linear, and generate the ground truth mappings by transforming the straight lines between the knots.

The transformations in our experiments were generated by “shooting” a random initial momenta field (Miller et al., 2006). In neuromorphology studies, transformations are typically generated via image registration to an atlas for which several approaches exist (Toga & Thompson, 2001; Chandrashekhara et al., 2021). This work is only relevant to non-affine registration techniques since affine transformations preserve straight lines. The utility of higher order mapping depends on the extent to which the brain sample is deformed nonlinearly. In practice, investigators can look at the profiles of position and tangent vector displacements

to identify which regime (σ level) is most similar to their transformation (Fig. 3a). At low values of σ , Frechet error of both zeroth and first order methods are in the range of 1 – 10 microns (Fig. 3c), which is likely negligible for mesoscale neuromorphology. However, under more extreme transformations, the first order mapping offers a more significant improvement in both Frechet error and distributions of morphometric quantities (Fig. 3c, d).

As mentioned previously, existing mapping methods use zeroth order mapping. Investigators can use the error bound in Eq. 2 to determine whether zeroth order mapping is adequate. If Jacobian and displacement values of the transformation at hand are not easily accessible, our empirical results can offer guidance. For example, we found that under less extreme transformations ($\sigma = 80, 160$), the frechet errors remained below ten microns for both zeroth and first order methods. However, as transformations got more extreme, it became more important to either keep the sampling period small, or to use first order mapping. Specifically, if the sampling period was less than ten microns, then both zeroth and first order mapping had low error. For higher sampling periods, first order mapping offered more significant improvements.

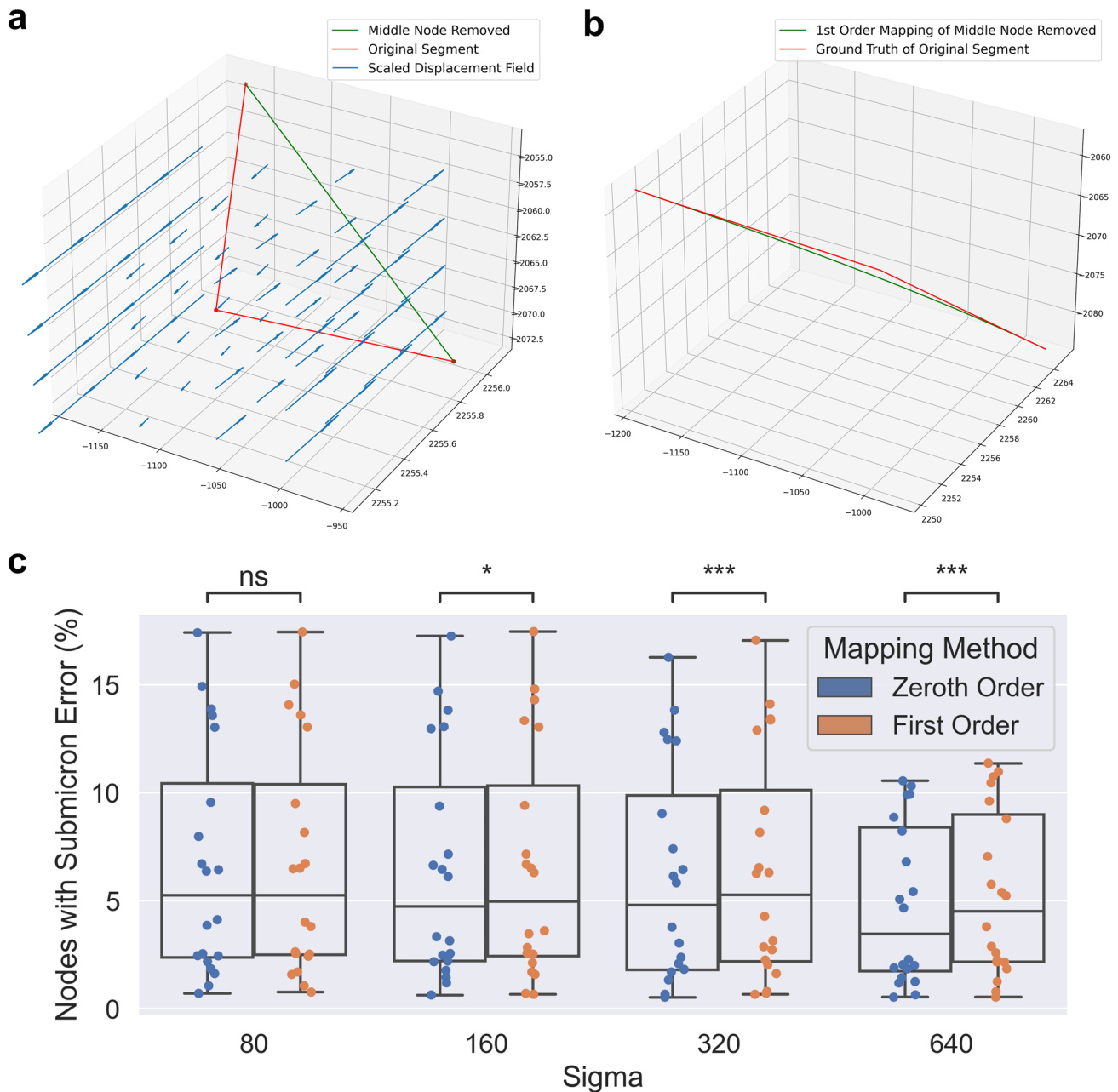


Fig. 5 Counting how many nodes in MouseLight neuron traces can be removed without affecting the mapped morphology. **a** For each non-branching node with path angle above 170 degrees, we generated a line segment with that node removed. **b** We performed first order mapping on the downsampled line segment and compared the result with the ground truth mapping of the original curve. **c** For each mapped neuron trace, we determined the fraction of nodes where

the discrete Frechet error is less than or equal to one micron under the four random diffeomorphisms. Box plots show median, upper and lower quartiles and whiskers have a maximum length of 1.5x the interquartile range with other outlier data marked with points. We performed Wilcoxon signed-rank test between the paired neuron traces at $\alpha = 0.05$ with Bonferroni correction across different values of σ (*: $p \leq 0.05$, ***: $p \leq 10^{-3}$)

Conversely our results can be used to make manual tracing more efficient. If the registration transformation, ϕ , is known a priori, and there are stretches where a neuronal branch is straight, then it is possible to compute the minimum sampling rate while still controlling the amount of error introduced during mapping to atlas coordinates.

The neuron trace files examined here are at most a couple megabytes, so this approach is not likely to produce significant data storage gains. However, it could allow manual tracers to sample more sparsely along straight stretches of axons, possibly leading to faster reconstruction. As a preliminary experiment, we computed the fraction of nodes

which could be removed, while maintaining a submicron error after first order mapping (Fig. 5). On average, around 5% of nodes achieved submicron error for both zeroth and first order mapping. This fraction decreased with larger sigma, indicating the importance of a higher sampling rate under more extreme transformations. First order mapping led to a statistically significant increase in the fraction of nodes with submicron error for $\sigma = 160, 320, 640$, though this increase was small. It is important to note that since each node was examined individually, it is not necessarily the case that removing all the nodes together would maintain submicron error. In the worst case, if all the nodes were located consecutively along the trace, only every other node could be removed to maintain submicron error. Further, it is unknown whether skipping the nodes identified in our experiment would have saved time in the MouseLight tracing protocol. A proper experiment to test this hypothesis would involve both registration and neuron reconstruction in real whole-brain images and thus is reserved as a potential avenue of future study. However, given that manual tracing remains a bottleneck and requires several person-hours per neuron (Winnubst et al., 2019), making tracing process just a couple percentage points faster would tangibly accelerate neuromorphological experiments.

It may be tempting to use our “ground-truth” mapping method, i.e. upsampling a linear interpolation then performing zeroth order mapping, as a neuron mapping method. While this may be appropriate in some settings, this approach has two primary disadvantages. First, as stated before, neurons are not piecewise linear structures so, while the knot positions can be generally regarded as lying on the neuron, the linear interpolation cannot. Therefore, it would be necessary to keep track of which knots are from the original trace, and which knots are from the upsampling in order to preserve the original trace information. This would require existing file formats to expand their metadata conventions. Secondly, for large traces, the upsampled data could become computationally cumbersome to store.

We want to highlight work in the adjacent field of neuron reconstruction where algorithms such as Li et al. (2020) can convert reconstruction knots into dense image segmentations which capture neuron trajectories at finer resolutions. Algorithms to automatically trace images of single neurons have been under development for decades (Peng et al., 2015; Athey et al., 2022). They could be adapted to generate both denser neuron samplings, and more accurate derivative estimates at the sampled points. These methods could improve both zeroth and first order mapping methods, so weighing these effects alongside the accuracy required for the given scientific goal would help determine which mapping method is appropriate.

Numerical error in mapping 3D curves, the subject of this paper, is only one source of error in neuron reconstruction studies. Error can also be introduced during the neuron tracing process. However, tracing errors generally come in the form of missing branches, or fusing two unconnected branches, rather than incorrect placement of trace points (Winnubst et al., 2019). Indeed, many tracing workflows involve semi-automated tools that “snap” the trace points to the fluorescent signal, and therefore tracing error of point placement can be assumed to be roughly equal to the resolution of the image. In the case of the MouseLight project, this was $0.3 \times 0.3 \times 1 \mu m^3$, less than the mapping errors in our experiments. A more likely location of trace error is between trace points, if the neuron does not follow a straight path between trace paths. However, in well-designed experiments, neuron tracers are trained to place enough trace points to closely follow the neuron’s trajectory. If a neuron’s morphology includes sharp turns which deviate several microns from the trace, then tracing error may dominate mapping error.

Neuron traces are not the only 3D curves being mapped in neuroscience. Tractographic data in diffusion MRI involves similar data structures and mapping algorithms, so we believe the relevance of our work extends beyond neuron tracing studies in mice (Tournier et al., 2019). We believe this work could also be extended to other geometric objects

Input: x_0, x_1, ϕ

Output: $Y_{GT} = [y_0, y_1, \dots, y_N]$

$L \leftarrow |x_1 - x_0|$

$N \leftarrow \text{ceil}(L/2) + 1$

for $k \leftarrow 0$ **to** $N - 1$ **do**

$X[k] \leftarrow x_0 + 2k \frac{x_1 - x_0}{|x_1 - x_0|}$

$Y_{GT}[k] \leftarrow \phi(X[k])$

end for

$X[N] \leftarrow x_1$

$Y_{GT}[N] \leftarrow \phi(X[N])$

▷ Start (x_0) and end (x_1) coordinates in microns, diffeomorphism ϕ

▷ Zeroth order mapping of upsampled segment

▷ Length of segment

▷ Sample every 2 microns along segment

▷ Transform coordinate according to ϕ

Algorithm 1 Ground Truth Mapping

Input: x_0, x_1, ϕ ▷ Start (x_0) and end (x_1) coordinates in microns, diffeomorphism ϕ
Output: $Y_0 = [y_0, y_1, \dots, y_N]$ ▷ Discrete sampling of zeroth order mapping
 $L \leftarrow |x_1 - x_0|$ ▷ Length of segment
 $y_0 \leftarrow \phi(x_0)$ ▷ Transform coordinate according to ϕ
 $y_1 \leftarrow \phi(x_1)$
 $S_0 \leftarrow \text{fit_first_order_spline}([0, L], [y_0, y_1])$ ▷ Make spline using *splprep* method in SciPy [29]
 $N \leftarrow \text{ceil}(L/2) + 1$
for $k \leftarrow 0$ **to** $N - 1$ **do**
 $Y_0[k] \leftarrow S_0(2k)$
end for
 $Y_0[N] \leftarrow S_0(L)$

Algorithm 2 Zeroth Order Mapping

other than branching curves, such as closed networks of curves, or surfaces.

As brain mapping efforts expand to include other species, such as non-human primates, we believe it will remain important to design algorithms that mitigate numerical error when handling digital neuron traces. Non-human primate brains are orders of magnitudes larger than those in mice, and also have shown more inter-individual variability, presenting new challenges for brain mapping (Ose et al., 2022).

Methods

Software Implementation

Our mapping framework was described in terms of non-branching curves in 3D, but can be naturally extended to neuron traces that have tree-like topology, i.e. have branching, but no loops. We decompose a branching trace into “branches” by recursively removing the root to leaf path with the longest arc length (Athey et al., 2021). The trace points

where branching occurs are copied into each of their associated branches e.g. the trace point of at a bifurcation will be part of two neuron branches. Thus, the tree-like graph of trace points is decomposed into non-branching subgraphs.

Each branch is considered to be a first order spline, i.e. the curve is piecewise linear. In this representation, mapping a neuron branch can be decomposed into mapping a collection of line segments that connect consecutive trace points, $x_i = c(t_i), x_{i+1} = c(t_{i+1})$. Algorithm 1 illustrates our process for generating ground truth, where we linearly interpolate the line segment every 2 microns, then apply ϕ to each coordinate. Algorithm 2 illustrates the zeroth order mapping method, where we apply ϕ to the endpoints of the line segment, then perform linear interpolation. Algorithm 3 illustrates our first order mapping method, where $\frac{x_{i+1}-x_i}{|x_{i+1}-x_i|}$ is the (one-sided) derivative at each endpoint, then we use ϕ to map the endpoint positions, and $D\phi$ to map the endpoint derivatives. The mapped positions and derivatives are used to define a cubic Hermite spline (Kincaid & Cheney, 2002). Specifically, we use the SciPy implementation of cubic Hermite splines (Virtanen et al., 2020).

Input: $x_0, x_1, \phi, D\phi$ ▷ Start (x_0) and end (x_1) coordinates in microns, diffeomorphism ϕ and its Jacobian $D\phi$
Output: $Y_1 = [y_0, y_1, \dots, y_N]$ ▷ Discrete sampling of first order mapping
 $L \leftarrow |x_1 - x_0|$ ▷ Length of segment
 $dx_0 \leftarrow \frac{x_1 - x_0}{L}$
 $dx_1 \leftarrow \frac{x_1 - x_0}{L}$
 $y_0 \leftarrow \phi(x_0)$ ▷ Transform coordinate according to ϕ
 $y_1 \leftarrow \phi(x_1)$
 $dy_0 \leftarrow D\phi(x_0) \cdot dx_0$ ▷ Transform derivative according to $D\phi$
 $dy_1 \leftarrow D\phi(x_1) \cdot dx_1$
 $S_1 \leftarrow \text{fit_cubic_hermite_spline}([0, L], [y_0, y_1], [dy_0, dy_1])$ ▷ Make spline using *PPoly* class in SciPy [29]
 $N \leftarrow \text{ceil}(L/2) + 1$
for $k \leftarrow 0$ **to** $N - 1$ **do**
 $Y_1[k] \leftarrow S_1(2k)$
end for
 $Y_1[N] \leftarrow S_1(L)$

Algorithm 3 First Order Mapping

After mapping is applied to each branch, the branches are re-connected by matching and “fusing” the trace points that were copied into multiple branches. This way, the topology of the original branching trace is preserved. Figure 3b–g depicts two examples of branching neuron traces which were transformed with both zeroth and first order mapping. The mapped morphologies are written to SWC format.

Quantitatively Comparing Curves

We define the Frechet distance between neuron traces to be the maximal Frechet distance between matched branches. Since we are comparing different mappings of the same neuron, matching branches to each other is trivial. We used the package from Jekel et al. to compute discrete frechet distance (Jekel et al., 2019). Discrete Frechet distance is an approximation of, and upper bound to Frechet distance (Eiter & Mannila, 1994). We used nGauge to load the transformed SWC files and compute morphometric quantities. We used SciPy to perform Kolmogorov-Smirnov statistics (Walker et al., 2022; Virtanen et al., 2020).

Further details about our implementation can be found in our open-source Python package brainlit: <http://brainlit.neurodata.io/>.

Information Sharing Statement

The neuron traces used in this work were from the MouseLight project’s NeuronBrowser website: <https://ml-neuronbrowser.janelia.org/>. Specifically, we used traces: AA-1087, 1089-1093, 1149, 1192, 1196, 1223, 1413, 1417, 1477, 1493, 1537-1540, 1543 and 1548. The code used in this work is available in our open-source Python package brainlit: <http://brainlit.neurodata.io/>.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12021-023-09648-0>.

Acknowledgements This work is supported by the National Institutes of Health grants RF1MH121539, P41EB015909, R01NS086888, U19AG033655, R01AG066184-01 the National Science Foundation grants 2031985, 2014862, and the CAREER award and the National Institute of General Medical Sciences Grant T32GM119998. We thank the MouseLight team at HHMI Janelia for providing us with access to this data, and answering our questions about it.

Author Contributions T.L.A. developed the mathematical details and the software, executed the experiments and contributed to the writing of this paper. D.J.T. contributed to the experimental design, and writing of the paper. U.M. contributed to the conceptual development, and data availability. L.Y. formulated the mathematical framework, contributed to the experimental design and the writing of the paper. J.T.V. contributed to the scope and design of the experiments. M.I.M. contributed to the scope of the experiments and the writing of the paper.

Declarations

Ethical Approval Under a license agreement between AnatomyWorks and the Johns Hopkins University, Dr. Miller and the University are entitled to royalty distributions related to technology described in the study discussed in this paper. Dr. Miller is a founder of and holds equity in AnatomyWorks. This arrangement has been reviewed and approved by the Johns Hopkins University in accordance with its conflict of interest policies.

Inclusion and Ethics Statement The purpose of this work is to investigate a computational technique that is being used in the brain mapping community. The roles and responsibilities of the authors were discussed prior to the research. Efforts were made to make the work accessible by, for example, providing data used in the experiments, and maintaining an open-source repository of the code with extensive documentation.

Research Animals Statement The mouse projection neuron traces came from the MouseLight project’s NeuronBrowser website and the experimental protocols that generated this data can be found in Winnubst et al. (2019).

Conflicts of Interest The remaining authors have no conflicts of interest to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Athey, T. L., Teneggi, J., Vogelstein, J. T., Tward D. J., Mueller, U., & Miller, M. I. (2021). Fitting splines to axonal arbors quantifies relationship between branch order and geometry. *Frontiers in Neuroinformatics*.
- Athey, T. L., Tward, D. J., Mueller, U., Vogelstein, J. T., & Miller, M. I. (2022). Hidden markov modeling for maximum probability neuron reconstruction. *Communications Biology*, 5(1), 1–11.
- Athey, T. L., Wright, M. A., Pavlovic, M., Chandrashekhkar, V., Deisseroth, K., Miller, M. I., & Vogelstein, J. T. (2023). Brainline: An open pipeline for connectivity analysis of heterogeneous whole-brain fluorescence volumes. *Neuroinformatics*, 1–3.
- BRAIN Initiative Cell Census Network (BICCN). (2021). A multi-modal cell census and atlas of the mammalian primary motor cortex. *Nature*, 598(7879), 86–102.
- Cannon, R. C., Turner, D. A., Pyapali, G. K., & Wheal, H. V. (1998). An on-line archive of reconstructed hippocampal neurons. *Journal of Neuroscience Methods*, 84(1–2), 49–54.
- Chandrashekhkar, V., Tward, D. J., Crowley, D., Crow, A. K., Wright, M. A., Hsueh, B. Y., Gore, F., Machado, T. A., Branch, A., Rosenblum, J. S., et al. (2021). Cloudreg: automatic terabyte-scale cross-modal brain volume registration. *Nature Methods*, 18(8), 845–846.

- Economo, M. N., Clack, N. G., Lavis, L. D., Gerfen, C. R., Svoboda, K., Myers, E. W., & Chandrashekar, J. (2016). A platform for brain-wide imaging and reconstruction of individual neurons. *Elife*, 5, e10566.
- Eiter, T., & Mannila, H. (1994). Computing discrete fréchet distance.
- Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., et al. (2012). 3d slicer as an image computing platform for the quantitative imaging network. *Magnetic Resonance Imaging*, 30(9), 1323–1341.
- Gao, L., Liu, S., Gou, L., Hu, Y., Liu, Y., Deng, L., Ma, D., Wang, H., Yang, Q., Chen, Z., et al. (2022). Single-neuron projectome of mouse prefrontal cortex. *Nature Neuroscience*, 25(4), 515–529.
- Jekel, C. F., Venter, G., Venter, M. P., Stander, N., & Haftka, R. T. (2019). Similarity measures for identifying material parameters from hysteresis loops using inverse analysis. *International Journal of Material Forming*, ISSN 1960–6214. <https://doi.org/10.1007/s12289-018-1421-8>
- Kincaid, D., & Cheney, W. (2002). *Numerical analysis: mathematics of scientific computing*. Pacific Grove, CA: Brooks/Cole Thomson Learning.
- Li, S., Quan, T., Zhou, H., Huang, Q., Guan, T., Chen, Y., Xu, C., Kang, H., Li, A., Fu, L., et al. (2020). Brain-wide shape reconstruction of a traced neuron using the convex image segmentation method. *Neuroinformatics*, 18(2), 199–218.
- Miller, M. I., Trounev, A., & Younes, L. (2006). Geodesic shooting for computational anatomy. *Journal of Mathematical Imaging and Vision*, 24(2), 209–228. ISSN 0924-9907. <https://doi.org/10.1007/s10851-005-3624-0>
- Olver, P. J. (1995). *Equivalence. Invariants and Symmetry*: Cambridge University Press, Cambridge.
- Ose, T., Autio, J. A., Ohno, M., Frey, S., Uematsu, A., Kawasaki, A., Takeda, C., Hori, Y., Nishigori, K., Nakako, T., et al. (2022). Anatomical variability, multi-modal coordinate systems, and precision targeting in the marmoset brain. *NeuroImage*, 250, 118965.
- O'Halloran, D. M. (2020). Module for swc neuron morphology file validation and correction enabled for high throughput batch processing. *PloS one*, 15(1), e0228091.
- Peng, H., Meijering, E., & Ascoli, G. A. (2015). From diadem to bigneuron.
- Peng, H., Xie, P., Liu, L., Kuang, X., Wang, Y., Qu, L., Gong, H., Jiang, S., Li, A., Ruan, Z., et al. (2021). Morphological diversity of single neurons in molecularly defined cell types. *Nature*, 598(7879), 174–181.
- Qu, L., Li, Y., Xie, P., Liu, L., Wang, Y., Wu, J., Liu, Y., Wang, T., Li, L., Guo, K., et al. (2022). Cross-modal coherent registration of whole mouse brains. *Nature Methods*, 19(1), 111–118.
- Skibbe, H., Reiser, M., Nakae, K., Watakabe, A., Hata, J., Mizukami, H., Okano, H., Yamamori, T., & Ishii, S. (2018). Pat-probabilistic axon tracking for densely labeled neurons in large 3-d micrographs. *IEEE Transactions on Medical Imaging*, 38(1), 69–78.
- Smale, S. (1958). Regular curves on riemannian manifolds. *Transactions of the American Mathematical Society*, 87(2), 492–512.
- Spitzbart, A. (1960). A generalization of hermite's interpolation formula. *The American Mathematical Monthly*, 67(1), 42–46.
- Stockley, E. W., Cole, H. M., Brown, A. D., & Wheal, H. V. (1993). A system for quantitative morphological measurement and electrotonic modelling of neurons: three-dimensional reconstruction. *Journal of Neuroscience Methods*, 47(1–2), 39–51.
- Suksumran, T. (2016). Gyrogroup actions: A generalization of group actions. *Journal of Algebra*, 454, 70–91.
- Toga, A. W., & Thompson, P. M. (2001). The role of image registration in brain mapping. *Image and Vision Computing*, 19(1–2), 3–24. ISSN 0262-8856. [https://doi.org/10.1016/S0262-8856\(00\)00055-X](https://doi.org/10.1016/S0262-8856(00)00055-X)
- Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H., & Connelly, A. (2019). Mrtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage*, 202, 116137.
- Tward, D. J., & Miller, M. I. (2017). On the complexity of human neuroanatomy at the millimeter morphome scale: Developing codes and characterizing entropy indexed to spatial scale. *Frontiers in Neuroscience*, 11. ISSN 1662-453X. <https://www.frontiersin.org/articles/10.3389/fnins.2017.00577>.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, 17, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Walker, L. A., Williams, J. S., Li, Y., Roossien, D. H., Lee, W. J., Michki, N. S., & Cai, D. (2022). nGauge: Integrated and extensible neuron morphology analysis in python. *Neuroinformatics*, 20(3), 755–764. ISSN 1559-0089. <https://doi.org/10.1007/s12021-022-09573-8>
- Watakabe, A., Skibbe, H., Nakae, K., Abe, H., Ichinohe, N., Rachmadi, M. F., Wang, J., Takaji, M., Mizukami, H., Woodward, A., et al. (2023). Local and long-distance organization of prefrontal cortex circuits in the marmoset brain. *Neuron*.
- Winnubst, J., Bas, E., Ferreira, T. A., Wu, Z., Economo, M. N., Edson, P., Arthur, B. J., Bruns, C., Rokicki, K., Schauder, D., et al. (2019). Reconstruction of 1,000 projection neurons reveals new cell types and organization of long-range connectivity in the mouse brain. *Cell*, 179(1), 268–281.
- Younes, L. (2010). *Shapes and diffeomorphisms* (Vol. 171). Berlin: Springer.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.