Magnetic Resonance Imaging 64 (2019) 122–131

Contents lists available at ScienceDirect

Magnetic Resonance Imaging

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

Original Contribution

Registration uncertainty quantification via low-dimensional characterization of geometric deformations

Jian Wang[a], William M. Wells III[b,c], Polina Golland[c], Miaomiao Zhang[a,*]

[a] Computer Science and Engineering, Washington University in St. Louis, MO, United States of America
[b] Computer Science and Artificial Intelligence Laboratory, MIT, MA, United States of America
[c] Brigham and Women's Hospital, Harvard Medical School, MA, United States of America

ARTICLE INFO

Keywords:
Bayesian image registration
Uncertainty quantification
Bandlimited space
Laplace approximation.

ABSTRACT

This paper presents an efficient approach to quantifying image registration uncertainty based on a low-dimensional representation of geometric deformations. In contrast to previous methods, we develop a Bayesian diffeomorphic registration framework in a bandlimited space, rather than a high-dimensional image space. We show that a dense posterior distribution on deformation fields can be fully characterized by much fewer parameters, which dramatically reduces the computational complexity of model inferences. To further avoid heavy computation loads introduced by random sampling algorithms, we approximate a marginal posterior by using Laplace's method at the optimal solution of log-posterior distribution. Experimental results on both 2D synthetic data and real 3D brain magnetic resonance imaging (MRI) scans demonstrate that our method is significantly faster than the state-of-the-art diffeomorphic registration uncertainty quantification algorithms, while producing comparable results.

1. Introduction

Diffeomorphic Image registration is a fundamental tool in image analysis as it provides one-to-one smooth and invertible smooth spatial correspondences (also known as diffeomorphisms) between pairwise images. Examples of applications include alignment of functional data to a reference coordinate system [1, 2], anatomical comparisons across individuals [3, 4], and atlas-based image segmentation [5, 6]. In general, image registration is an ill-posed problem since the image data are usually contaminated by unknown noise. Therefore, developing efficient measures to quantify the registration uncertainty or error is critical to fair assessments on estimated transformations, as well as subsequent improvement on the accuracy of predictive models. This also forms the basis for model-assisted decision making, for example, image-guided navigation system for neurosurgery [7, 8]. An efficient uncertainty quantification can provide important information on brain shifts (deformations) for neurosurgeons to identify residual tumor tissues during surgery; hence leading to a significant increase in the extent of tumor removal while lowering the risk of collateral tissue damage.

Existing methods have been investigated to estimate the uncertainty by having a probability distribution over the latent space of transformation parameters either in a small or large deformation setting [9, 10]. These approaches formulate Bayesian image registration as an image matching likelihood term regularized by a prior that guarantees the smoothness of deformation fields. A posterior distribution is then generated as a measure of the registration uncertainty estimation. Due to the fact that such posterior distributions do not have closed-form formulations, stochastic and sampling methods typically have been employed [7, 11, 12]. This makes the entire inference extremely challenging on a dense image grid (e.g., a 3D brain MRI with the size of 128^3). Large computational resources and efforts are required to sample over a high-dimensional space with extremely slow convergence. Also, none of the aforementioned algorithms is able to compute a full posterior covariance. These disadvantages dramatically decrease the model usability in important applications that require computational efficiency.

In this paper, we propose an efficient registration uncertainty quantification model that employs a low-dimensional representation of diffeomorphic transformations [13, 14]. Based on the key fact that the tangent space of diffeomorphisms does not develop high frequencies, we develop a Bayesian registration framework entirely in a bandlimited space with low frequencies well preserved. A posterior distribution of transformation fields is carefully designed in a truncated frequency domain, where all diffeomorphic properties remain valid. Similar to [15], we derive a Laplace approximation of the log-posterior distribution at an optimal solution to further avoid intensive computation

* Corresponding author.
E-mail address: zhang.m@wustl.edu (M. Zhang).

https://doi.org/10.1016/j.mri.2019.05.034
Received 30 January 2019; Received in revised form 21 May 2019; Accepted 22 May 2019

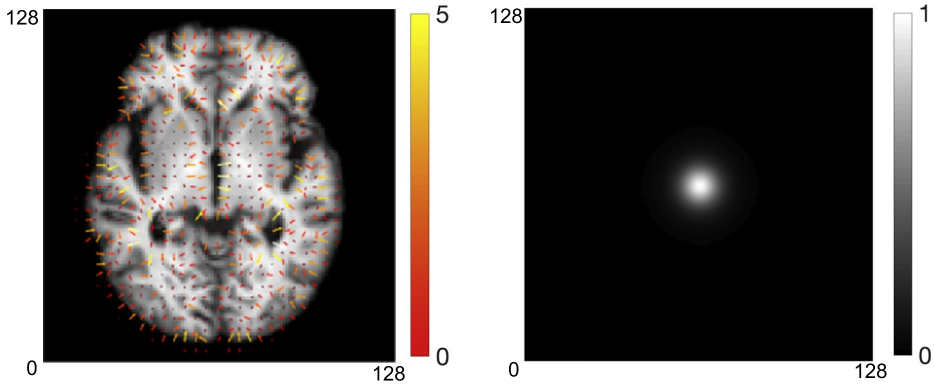

Fig. 1. Left to right: An example of 2D velocity field visualized in image space vs. Fourier space. The white dot centered in middle displays centered low frequencies.

cost of sampling methods. More specifically, we assume a complex Gaussian distribution at the mode of the posterior, with its covariance estimated by second-order information to describe uncertainties in the registration model. This is effectively done by computing an inverse Hessian of the log-posterior specifically defined in a low-dimensional bandlimited space. Our model dramatically reduces the computational complexity of approximating posterior marginals in the space of diffeomorphisms, which makes registration uncertainty analysis tractable in time. In practical, the efficiency of our method strengthens its feasibility in prospective clinical applications under a tight time constraint. This paper is an extension of a recently published conference paper [16] with (i) in-depth derivations of the Hessian computation and inference procedure provided; and (ii) comprehensive experiments to thoroughly validate the method included. We run tests on both 2D synthetic data and real 3D brain MRI scans from OASIS dataset [17]. We then compare estimated results with the state-of-the-art registration uncertainty quantification method in the full image space [15]. Experimental results show that our method is significantly faster than the baseline algorithm, while producing comparable results.

The rest of this paper is structured as follows. Section 2 briefly introduces the background of diffeomorphic image registration in the context of large deformation diffeomorphic metric mapping (LDDMM) framework [18] followed by a Fourier representation of diffeomorphisms [13, 19]. Section 3 introduces our developed Bayesian model with Section 4 presenting the inference. Section 5 demonstrates our experiments with detailed discussions.

2. Background

In this section, we first briefly review the background of diffeomorphic image registration in the context of LDDMM with geodesic shooting algorithm [18, 20]. We then introduce a recently developed Fourier representation of velocity fields in the tangent space of diffeomorphisms that dramatically speeds up the registration optimization [13, 19].

2.1. LDDMM with geodesic shooting

Let S be a source image and T be a target image defined on a torus domain $\Gamma = \mathbb{R}^d/\mathbb{Z}^d$ $(S(x),T(x) : \Gamma \to \mathbb{R})$. The problem of diffeomorphic image registration is to find the shortest path to generate time-varying diffeomorphisms $\{\psi_t\} : t \in [0,1]$ such that $S \circ \psi_1$ is similar to T, where $\circ$ is a composition operator that resamples S by the smooth mapping ψ_1. This is typically solved by minimizing an explicit energy function of LDDMM [18] as

$$E(v_t) = \text{dist}(S\circ\psi_1, T) + (\mathcal{L}v_t, v_t), \quad s.t. \quad \frac{d\psi_t}{dt} = -\mathcal{D}\psi_t \cdot v_t, \tag{1}$$

where the distance function dist$(\cdot,\cdot)$ measures the dissimilarity between images. Commonly used distance functions include sum-of-squared difference of image intensities [18], normalized cross correlation [21], and mutual information [22]. The regularization term (v_t,v_t) is a constraint that enforces spatial smoothness of transformations with being a symmetric and positive-definite differential operator. Here, $(\cdot,\cdot)$ denotes a pair between a tangent vector v_t and a momentum vector $m_t = v_t$ that is similar to an inner product, $\mathcal{D}$ is a Jacobian matrix, and $\cdot$ is an element-wise multiplication.

The geodesic shooting algorithm states that, given an initial velocity v_0, a shortest path of diffeomorphisms $\{\psi_t\}$ can be uniquely determined by integrating the geodesic evolution equation (also known as Euler-Poincaré differential equation (EPDiff) [23, 24]) as

$$\frac{dv_t}{dt} = \text{ad}^{\dagger}_{v_t} v_t = -\mathcal{K}[(\mathcal{D}v_t)^T \cdot m_t + \mathcal{D}m_t \cdot v_t + m_t \cdot\div v_t], \tag{2}$$

where $\text{ad}^{\dagger}$ is an adjoint operator, $\mathcal{K}$ is an inverse operator of $\mathcal{L}$, and div denotes the divergence.

As a consequence, the optimization of LDDMM with geodesic shooting can be equivalently written as

$$E(v_0) = \text{dist}(S\circ\psi_1, T) + (\mathcal{L}v_0, v_0), \quad s.t. \quad \frac{d\psi_t}{dt} = -D\psi_t \cdot v_t \ \ \& \ \ \text{Eq.}. \tag{3}$$

This substantially reduces the computational complexity and improves the optimization landscape by only manipulating the initial velocity v_0 with EPDiff in Eq. (2), rather than the entire time-dependent velocity fields v_t.

2.2. Fourier representation of velocity fields

A recent work has shown that the velocity fields generated by EPDiff (2) can be efficiently captured via a low-dimensional representation in Fourier space with bandlimited signals [13, 19]. The key idea behind is that the velocity fields do not develop high frequencies and only a small amount of low frequencies contributes to the computation of generating deformation fields (as shown in Fig. 1). Therefore, we are able to capture the deformations in a bandlimited space as accurately as the original space. We next briefly review the relevant details of the method.

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a real-valued function. The Fourier transform $\mathcal{F}$ of f is given by

$$\mathcal{F}[f](\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i\langle\xi,x\rangle}\, dx, \tag{4}$$

where $x = (x_1,\ldots,x_d)$ is a d-dimensional vector of image coordinates, $\xi = (\xi_1,\ldots,\xi_d)$ is a d-dimensional vector of frequencies, and $\langle\cdot,\cdot\rangle$ denotes an inner product operator. The inverse Fourier transform $\mathcal{F}^{-1}$ of a discretized Fourier signal $\tilde{f}$

$$\mathcal{F}^{-1}[\tilde{f}](x) = \sum_{\xi} \tilde{f}(\xi) e^{2\pi i\langle\xi,x\rangle} \tag{5}$$

is an approximation of the original signal f. For vector-valued functions, such as diffeomorphisms ψ and velocity fields v, we apply (inverse) Fourier transform to each vector component separately.

Analogous to the EPDiff defined on the spatial domain of image space, we rewrite Eq. (2) in frequency domain as

$$\frac{\partial \tilde{v}_t}{\partial t} = \mathrm{ad}^{\dagger}_{\tilde{v}_t}\tilde{v}_t = -\tilde{\mathcal{K}}[(\tilde{\mathcal{D}}\tilde{v}_t)^T \bigstar \tilde{m}_t + \tilde{\mathcal{D}}\tilde{m}_t {}^{*}\tilde{v}_t + \tilde{m}_t {}^{*}(\tilde{\nabla}\cdot\tilde{v}_t)], \tag{6}$$

where $\tilde{\mathcal{K}}$ is a smoothing operator. Here $\tilde{\mathcal{L}}$ is the Fourier coefficients of a differential operator $\mathcal{L}$, which converts a vector field $\tilde{v}$ to a momentum vector by $\tilde{m} = \tilde{\mathcal{L}}\tilde{v}$, $\bigstar$ is a circular matrix-vector field auto-correlation [1], * is a circular convolution, and $\tilde{D}\tilde{v}$ is a tensor product $\tilde{D} \otimes \tilde{v}$ with $\tilde{D}(\boldsymbol{\xi}) = i\sin(2\pi\boldsymbol{\xi})$ representing the Fourier frequencies of a central difference Jacobian matrix D. The operator $\tilde{\nabla}\cdot$ is the discrete divergence operator that is computed as the sum of the Fourier coefficients of the central difference operator $\tilde{D}$ along each dimension, i.e., $\tilde{\nabla}\cdot\boldsymbol{\xi} = \sum_{j=1}^{d} i\sin(2\pi\boldsymbol{\xi}_j)$.

Since $\mathcal{K}$ on the left side of Eq. (6) is a smoothing operator (low-pass filter) that suppresses high frequencies in the Fourier domain, all computational operations in Eq. (6) are easy to implement in a truncated low-dimensional space by eliminating high frequencies. To ensure that $\tilde{f}$ represents a real-valued vector field in the spatial domain, we require $\tilde{f}(\xi_1, \ldots, \xi_d) = \tilde{f}^{*}(-\xi_1, \ldots, -\xi_d)$, where * denotes the complex conjugate.

Similarly, the diffeomorphic transformations in the frequency domain can be computed by

$$\begin{aligned} \frac{d\tilde{\psi}_t}{dt} &= -\tilde{\mathcal{D}}\tilde{\psi}_t {}^{*}\tilde{v}_t, \quad \text{or} \\ \frac{d\tilde{u}_t}{dt} &= -\tilde{v}_t - \tilde{\mathcal{D}}\tilde{u}_t {}^{*}\tilde{v}_t \ \text{ with } \ \tilde{\psi}_t = \tilde{e} + \tilde{u}_t, \end{aligned} \tag{7}$$

where $\tilde{e}$ is the Fourier coefficients of an identity element of image coordinates.

3. Our model: low-dimensional Bayesian registration uncertainty quantification

We introduce a low-dimensional Bayesian model of diffeomorphic image registration represented in the bandlimited velocity space $\tilde{V}$, with registration uncertainty explicitly encoded as latent variables of the model.

Assuming independent and identically distributed (i.i.d.) Gaussian noise on image intensities, we obtain the likelihood

$$p(T \mid S, \sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma^2)^M}\exp\left(-\frac{1}{2\sigma^2}\|S\circ\psi_1 - T\|_2^2\right), \tag{8}$$

where σ^2 is the noise variance and M is the number of image voxels. The deformation ψ_1 corresponds to $\tilde{\psi}_1$ in Fourier space via the Fourier transform $\mathcal{F}(\psi_1) = \tilde{\psi}_1$, or its inverse $\psi_1 = \mathcal{F}(\tilde{\psi}_1)$.

We define a prior on the initial velocity field $\tilde{v}_0$ to be a complex multivariate Gaussian distribution that ensures the smoothness of the geodesic path, i.e.,

$$p(\tilde{v}_0) = \frac{1}{(2\pi)^{\frac{Md}{2}}|\tilde{\mathcal{L}}^{-1}|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(\tilde{\mathcal{L}}\tilde{v}_0, \tilde{v}_0)\right), \tag{9}$$

where $|\cdot|$ is matrix determinant. In this paper, we define $\tilde{\mathcal{L}}$ as the Fourier coefficients of a commonly used Laplacian operator $\mathcal{L} = (-\alpha\Delta + I)^c$ with a positive weight parameter α and a smoothness parameter c. The Fourier transform of the Laplacian operator is given by

$$\tilde{\mathcal{L}}(\boldsymbol{\xi}) = \left(-2\alpha\sum_{j=1}^{d}(\cos(2\pi\xi_j) - 1) + 1\right)^c,$$

with $\tilde{\mathcal{K}} = 1/\tilde{\mathcal{L}}$ being an inverse operator of $\tilde{\mathcal{L}}$.

Combining the likelihood (8) and prior (9) together, we obtain the negative log posterior distribution on the deformation parameter parameterized by $\tilde{v}_0$ as

$$-\ln p(\tilde{v}_0 \mid S, T, \sigma^2) = \frac{1}{2}(\tilde{\mathcal{L}}\tilde{v}_0, \tilde{v}_0) + \frac{\|S\circ\psi_1 - T\|_2^2}{2\sigma^2} + M\ln\sigma + \text{const.} \tag{10}$$

In most probabilistic formulations of image-based registration, the likelihood function (8), as a function of the transformation parameters, is highly non-Gaussian because of the complex spatial structure of the images. This brings difficulties in the inference of such a non-Gaussian posterior. While sampling methods have been investigated to empirically approximate the distribution [25, 26], the computation tends to be extremely time-consuming and expensive. To address this issue, we next introduce Laplace's method [27] to approximate the posterior distribution in a much more efficient manner.

3.1. Laplace approximation

The basic idea is to find the mode of the posterior as a MAP solution, and then apply a second-order Taylor series approximation for the log-posterior function that corresponds to a complex Gaussian distribution. We first minimize the negative log posterior in Eq. (10) to the optimum $\tilde{v}_0^{opt}$, which is considered to be the mean of the Gaussian distribution (details of inference will be introduced in Section 4). Estimation of covariance is not straightforward, as we will need to derive a quadratic function of log-posterior by second-order approximation.

To simplify the notation, we use $f(\tilde{v}_0) \triangleq -\ln p(\tilde{v}_0 \mid S, T, \sigma^2)$ to represent the log-posterior. The function $f(\tilde{v}_0)$ is approximated to quadratic order by using second order Taylor series expansion at the optimal solution $\tilde{v}_0$ as

$$\begin{aligned} f(\tilde{v}_0) &= f(\tilde{v}_0^{opt} + (\tilde{v}_0 - \tilde{v}_0^{opt})) \\ &\approx f(\tilde{v}_0^{opt}) + \nabla f^T(\tilde{v}_0^{opt})(\tilde{v}_0 - \tilde{v}_0^{opt}) + \frac{1}{2}(\tilde{v}_0 - \tilde{v}_0^{opt})^T\mathcal{H}f \\ &\quad (\tilde{v}_0^{opt})(\tilde{v}_0 - \tilde{v}_0^{opt}), \end{aligned}$$

where ∇ denotes the first derivative and $\mathcal{H}$ is a second-order Hessian. Since the first derivative of f vanishes at the optimal solution $\tilde{v}_0^{opt}$, we have

$$f(\tilde{v}_0) \approx f(\tilde{v}_0^{opt}) + \frac{1}{2}(\tilde{v}_0 - \tilde{v}_0^{opt})^T\mathcal{H}f(\tilde{v}_0^{opt})(\tilde{v}_0 - \tilde{v}_0^{opt}). \tag{11}$$

This indicates that the posterior is approximately a multivariate Gaussian distribution $\mathcal{N}(\tilde{v}_0^{opt}, \mathcal{H}^{-1}f(\tilde{v}_0^{opt}))$. The optimal solution $\tilde{v}^{opt}$ and the inverse Hessian correspond to the mean velocity and covariance matrix of the registration parameters respectively.

4. Inference

The algorithmic inference includes two major components: seeking for mean and covariance of the approximated posterior distribution. For mean estimation, we develop a gradient descent algorithm to minimize the negative log posterior distribution (10) w.r.t. the initial velocity $\tilde{v}_0$ and the image noise variance σ^2. For covariance estimation, we derive a second variation of Eq. (10) to compute the Hessian-vector product via a linearized forward-backward sweep. All computational steps will be introduced in the following sections.

[1] The auto-correlation operates on zero-padded signals followed by truncating back to the bandlimits in each dimension to guarantee the output remains bandlimited.

4.1. Mean estimation

Following optimal control theory [20], we add Lagrange multipliers to constrain the diffeomorphism $\tilde{\psi}_t$ to be a geodesic path in the frequency domain. This is done by introducing time-dependent adjoint variables, $\hat{\tilde{v}}_t$ and $\hat{u}_t$, and writing the augmented energy [2],

$$E(\tilde{v}_0) = -\ln p(\tilde{v}_0 \mid S, T, \sigma^2) + \int_0^1 \langle \hat{\tilde{v}}_t, \dot{\tilde{v}}_t + \mathrm{ad}^{\dagger}_{\tilde{v}_t}\tilde{v}_t\rangle + \langle \hat{u}_t, \dot{u}_t + \tilde{v}_t + \tilde{\mathcal{D}}\tilde{u}_t{}^*\tilde{v}_t\rangle \, dt, \tag{12}$$

where the last two terms correspond to Lagrange multipliers enforcing the geodesic constraint (6) and the deformation transport Eq. (7).

The optimality conditions for the adjoints $\hat{\tilde{v}}_t$, $\hat{u}_t$ are given by the following time-dependent system of ordinary differential equations, termed the adjoint equations (equivalent to error-back propagation):

$$\begin{aligned} -\dot{\hat{\tilde{v}}}_t + \mathrm{ad}_{\tilde{v}_t}\hat{\tilde{v}}_t - \mathrm{ad}^{\dagger}_{\hat{\tilde{v}}_t}\tilde{v}_t + \hat{u}_t + (\tilde{\mathcal{D}}\tilde{u}_t)^T \bigstar \hat{u}_t &= 0, \\ &= 0, \\ -\dot{\hat{u}}_t - \tilde{\mathcal{D}}\delta\hat{u}_t{}^*\tilde{v}_t - \delta\hat{u}_t{}^*(\tilde{\nabla}\cdot\tilde{v}_t) &= 0 \end{aligned} \tag{13}$$

subject to initial conditions $\hat{\tilde{v}}_1 = 0$ and $\hat{u}_1 = -\frac{1}{\sigma^2}\mathcal{F}[\langle \nabla S(1), S(1) - T\rangle]$, where $S(1) = S \circ \psi_1$. Please refer to Appendix A for more details.

After integrating the geodesic equations (state equations) Eqs. (6) (7) forward in time to $t = 1$ and then backward integrating the adjoint Eq. (13) in time to $t = 0$, the gradient of E w.r.t. $\tilde{v}_0$ is $\nabla_{\tilde{v}_0} E = \tilde{v}_0 - \hat{\tilde{v}}_0$.

Setting the gradient w.r.t. σ^2 to zero, we have a closed form update

$$\sigma^2 = \frac{1}{M}\|S(1)\circ\psi_1 - T\|_2^2.$$

4.2. Covariance estimation

To estimate the full covariance matrix as an inverse Hessian, we develop a similar forward-backward approach to compute the Hessian-vector products that involves the second variation of the augmented energy function (12). More specifically, we read off the Hessian-vector product after deriving the second variation in the direction $\delta\tilde{v}_0$ as

$$\frac{\partial^2}{\partial\epsilon^2}E(\tilde{v}_0 + \epsilon\cdot\delta\tilde{v}_0)|_{\epsilon=0} = \langle \delta\tilde{v}_0, \mathcal{H}E\ \delta\tilde{v}_0\rangle.$$

Given an initial condition $\delta\tilde{v}_0$, we compute the Hessian-vector product as

$$\mathcal{H}E\ \delta\tilde{v}_0 = \delta\tilde{v}_0 - \delta\hat{\tilde{v}}_0,$$

where $\delta\hat{\tilde{v}}_0$ is the adjoint variable of $\delta\tilde{v}_0$.

The second variation can be accomplished by forward-sweeping the linearized geodesic constraint around the optimal solution, followed by a backward sweep of the linearized adjoint system. Introducing time-dependent adjoint variables $\delta\hat{\tilde{v}}_t$ and $\delta\hat{u}_0$, the forward linearized geodesic equations are

$$\delta\dot{\tilde{v}}_t = -\mathrm{ad}^{\dagger}_{\delta\tilde{v}_t}\tilde{v}_t - \mathrm{ad}^{\dagger}_{\tilde{v}_t}\delta\tilde{v}_t, \qquad \delta\dot{u}_t = -\tilde{\mathcal{D}}\delta\tilde{u}_t{}^*\tilde{v}_t - \tilde{\mathcal{D}}\tilde{u}_t{}^*\delta\tilde{v}_t - \delta\tilde{v}_t. \tag{14}$$

The linearized adjoint system for the backward integration is

$$\begin{aligned} \delta\dot{\hat{\tilde{v}}}_t &= \mathrm{sym}^{\dagger}_{\tilde{v}_t}\delta\hat{\tilde{v}}_t - \mathrm{sym}^{\dagger}_{\delta\tilde{v}_t}\hat{\tilde{v}}_t + \delta\hat{u}_t + (\tilde{\mathcal{D}}\tilde{u}_t)^T\bigstar\delta\hat{u}_t + (\tilde{\mathcal{D}}\delta\tilde{u}_t)^T\bigstar\hat{u}_t, \\ \delta\dot{\hat{u}}_t &= -\tilde{\mathcal{D}}\delta\hat{u}_t{}^*\tilde{v}_t - \tilde{\mathcal{D}}\hat{u}_t{}^*\delta\tilde{v}_t - \tilde{v}_t{}^*\tilde{\nabla}(\delta\hat{u}_t) - \hat{u}_t{}^*\tilde{\nabla}(\tilde{v}_t), \end{aligned} \tag{15}$$

subject to initial conditions $\delta\hat{u}_1 = -\frac{2}{\sigma^2}\mathcal{F}[\nabla S(1)\cdot\nabla S(1) + (S(1) - T)\cdot\nabla^2 S(1)]$ and $\delta\hat{\tilde{v}}_1 = 0$, with $\mathrm{sym}^{\dagger}_{\tilde{v}_t}\delta\hat{\tilde{v}}_t = \mathrm{ad}_{\delta\tilde{v}_t}\hat{\tilde{v}}_t - \mathrm{ad}^{\dagger}_{\hat{\tilde{v}}_t}\delta\tilde{v}_t$. Details of derivations are represented in Appendix A.

We summarize the inference of our algorithm in Algorithm 1.

[2] For notation simplification, we define the time derivative $\dot{\tilde{v}}_t \triangleq d\tilde{v}_t/dt$.

Algorithm 1. Model inference.

Input : source image S and target image T, step size ϵ, iterations r

Output: optimal registration $\tilde{v}_0^{opt}$, and covariance $\mathcal{H}^{-1}f(\tilde{v}_0^{opt})$

/* Mean estimation */

for $i = 1$ to r **do**

1. Forward integrating the geodesic evolution equations (6) with initial velocity $\tilde{v}_0$, then generates a collection of time-dependent velocity fields $\tilde{v}_t$;
2. Backward integrating adjoint equations (13) to get gradient $\nabla_{\tilde{v}_0 E}$;
3. $\tilde{v}_0^{opt} \leftarrow \tilde{v}_0^{opt} - \epsilon\nabla_{\tilde{v}_0 E}$

/* Covariance estimation */

for $i = 1$ to M **do**

1. Forward shooting of $\delta\tilde{v}_0$: forward-sweeping the linearized geodesic constraints (14) to generate $\delta\tilde{v}_t$;
2. Backward shooting of $\delta\tilde{v}_1$: backward integrating adjoint equation (15) to generate $\tilde{\mathcal{H}}E\delta\tilde{v}_0$;

3. Invert Hessian $\tilde{\mathcal{H}}^{-1}E\delta\tilde{v}_0 = \tilde{\mathcal{H}}^{-1}f(\tilde{v}_0^{opt})$ to generate covariance.

5. Results

To show the effectiveness of our model, we run experiments on both 2D synthetic data and 3D real brain MRI scans. We estimate the covariance by using $\alpha = 3, c = 6$ for the operator $\tilde{L}$. We set band-limited dimension of the initial velocity field as 16, which is similar to the settings used in the pairwise diffeomorphic image registration [14]. We run both our model and baseline algorithms till convergence, with the number of time steps for Euler integration in geodesic shooting being 10.

5.1. Data

We simulate a collection of 2D synthetic dataset starting from a binary circle with resolution 100×100. We then generate smooth initial velocity fields from the prior distribution $p(\tilde{v}_0)$, defined in Eq. (9). The deformed binary circles (used as our target images) are constructed by deforming the binary circle image with forward shooting the initial velocities.

The brain MRI scans of 60 subjects from the OASIS dataset [17], aged 60 to 90. All MRIs have the same resolution $128 \times 128 \times 128$ with the voxel size of $1.25 \times 1.25 \times 1.25 mm^3$. To further evaluate the accuracy of our mean estimation (registration solution), we use 10 pairs of 3D brain MRI scans with manual segmentation labels from public released ADNI dataset [28]. All images underwent the preprocessing of skull-stripping, downsampling, intensity normalization to [0,1] interval, bias field correction, and co-registration with affine transformations.

5.2. Experiments

We first evaluate the proposed approach on 2D synthetic data and compare with the state-of-the-art diffeomorphic registration uncertainty quantification algorithm defined in a full-dimensional image space [15]. This baseline algorithm is defined in the original high-dimensional image space, where the full Hessian is approximated via low-rank basis obtained by principal component analysis. The authors then compute the covariance matrix by inverting the approximated low-rank Hessian.

A visualization of uncertainty map estimated by the covariance matrix will be demonstrated. We set the initialization of initial velocity fields as zero and $\sigma = 0.01$ for both methods. We then validate our model on real 3D brain MRIs. Note that it is difficult to measure the accuracy of uncertainty methods since no registration ground truth

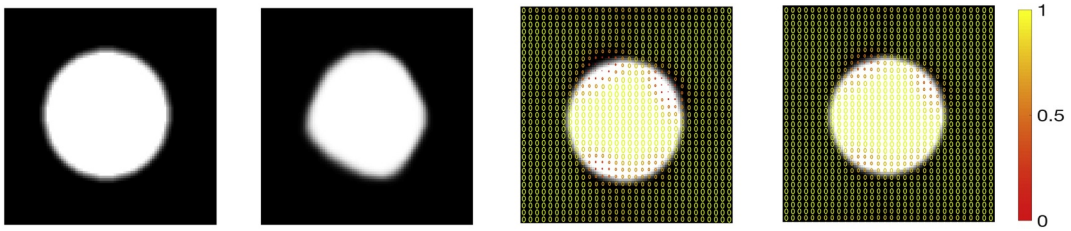

Fig. 2. Left to right: source image, target image, covariance matrix determinant estimated by the baseline algorithm and our method.

exists. In order to demonstrate the effectiveness of our method in a statistical way, we report the difference between the Hessian without low-dimensional approximation and the Hessian estimated by our method both on 2D synthetic and 3D real data. In addition, to estimate the final uncertainty maps, we show the time and memory comparison of all algorithms.

Finally, we demonstrate that another benefit of our model is the improved performance of mean estimation by standard image registration problem. Rather than finding an optimal registration result (mean) in a high-dimensional image space by using gradient descent algorithm, we instead perform an entire optimization in a low-dimensional bandlimited space. This not only makes the algorithmic inference much faster, but also reduces the risk of getting stuck in local minima. We compare our algorithm with the state-of-the-art fast diffeomorphic registration schemes: Quicksilver [29] and Voxelmorph [30]. To evaluate the estimated deformations, we perform registration-based segmentation and examine the resulting segmentation accuracy. The volume overlap, also known as Dice Similarity Coefficient (DSC), between the propagated segmentation A and the manual segmentation B for each structure is computed as $(A, B) = 2(|A| \cap |B|)/(|A| + |B|)$ where $\cap$ denotes an intersection of two regions.

5.3. Experimental results

Fig. 2 visualizes the uncertainty information estimated from both our method and the baseline algorithm performed in high-dimensional space on 2D data. We extract the local covariance matrix of each voxel and visualize it as an ellipse on the source image, with the color representing the matrix determinant. The smaller determinants are closer to the non-isotropic area (e.g., circle boundaries), which indicate more confident registration results.

Fig. 3 visualizes an example of 3D brain registration uncertainty. Note that due to the difficulty of computing a full covariance matrix by inverse Hessian in a high dimensional image space, we need to use an approximated low-rank Hessian with a number of dominant eigenmodes [15]. Based on our key observation that the eigenvalues of Hessian matrix decay fast (as shown on the left panel of Fig. 4), we calculated the variances with various numbers of eigenmodes and noticed that the difference between covariances goes to zero beyond 2600 eigenmodes. Following [15], we then choose the first 2600 eigenmodes to approximate the full covariance accordingly. Both two methods show that the high uncertainty (with less confidence) appear isotropic areas (e.g., inside the ventricle), while the low uncertainty (with high

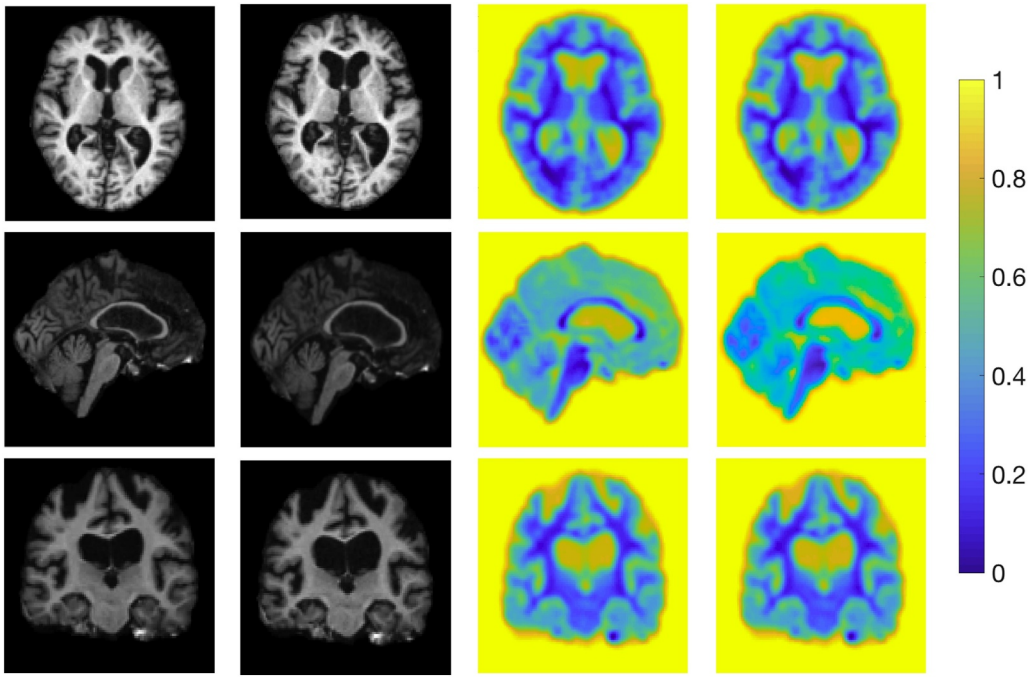

Fig. 3. Left to right: source image, target image, and uncertainty (visualized as the trace of covariance) estimated by baseline algorithm and our method.

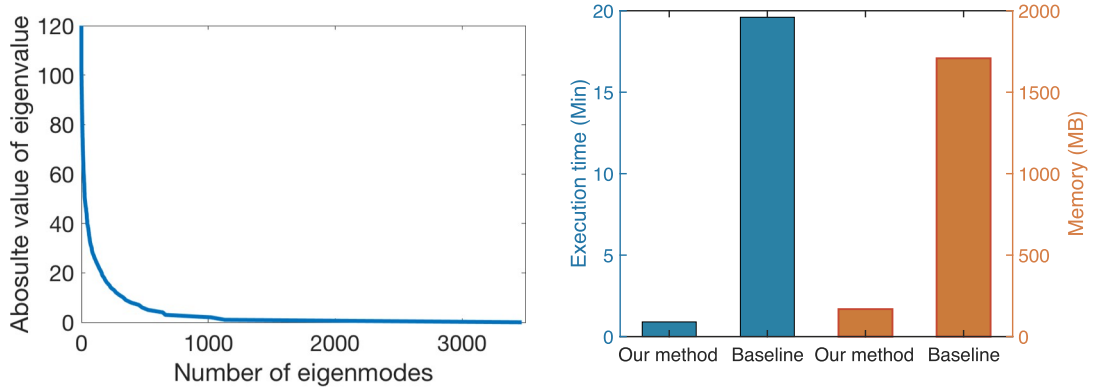

Fig. 4. Left panel: eigenvalues of the Hessian matrix estimated by the baseline algorithm; Right panel: comparison of average runtime and memory consumption for all test cases.

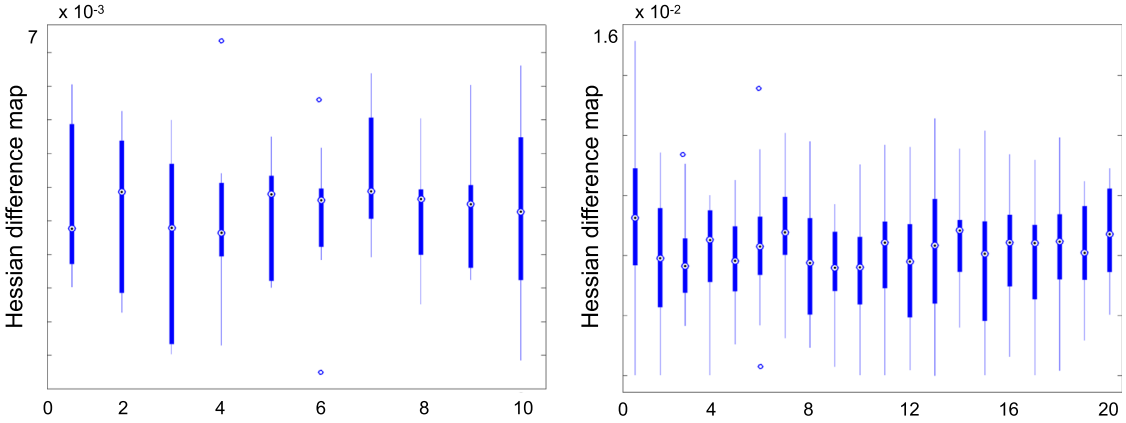

Fig. 5. Difference between the Hessian matrices estimated by our method and the full Hessian on both 2D and 3D data.

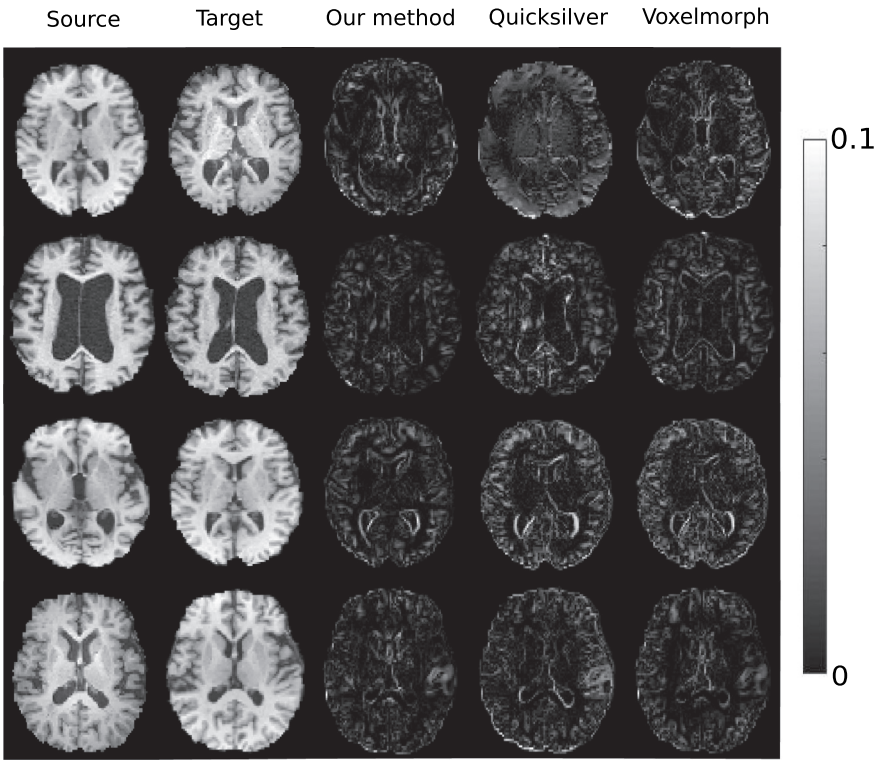

Fig. 6. Left to right: source image, target image, difference map of deformed image by estimated mean of our model, Quicksilver [29], and Voxelmorph [30].

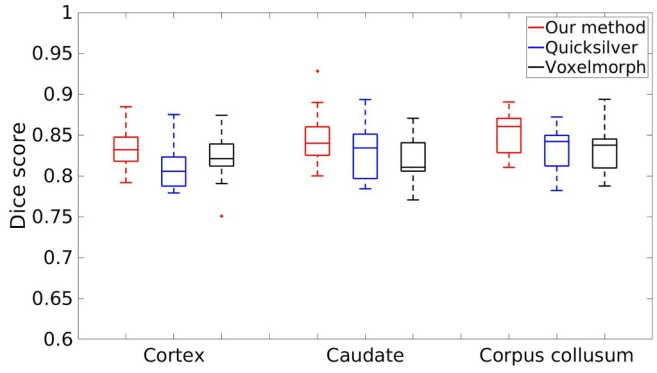

Fig. 7. Dice score evaluated on the brain structure, cortex, caudate and corpus collusum of our method, Quicksilver, and Voxelmorph.

confidence) appears around non-isotropic areas (e.g. ventricle boundaries). The right panel of Fig. 4 reports the comparison of time and memory consumption. Our algorithm offers significant improvements in computational efficiency. It is worth mentioning that while we adopted FLASH [14] for fast registration in our paper, the proposed that Bayesian framework can be generalized to other registration models with the same level of efficiency, e.g., stationary velocity fields that remain constant over time [31], or projected subspace of velocity fields that characterizes deformations in a much lower dimensional space [32].

Fig. 5 reports the absolute value of difference between the Hessian estimated by our method and the full Hessian on ten pairs of 2D synthetic data and twenty pairs of 3D MRI scans. The low errors on both 2D and 3D experiments indicate that our estimation is fairly close to the optimal uncertainty result.

Fig. 6 displays example error maps between deformed source images by the estimated mean (an optimal of registration) from all methods and target images. It shows that our method achieves better registration results (a.k.a., smaller errors) than the two baseline algorithm [29, 30].

Fig. 7 reports the dice evaluations on brain cortex over 10 pairs of images for all methods. Our algorithm produces better dice score than the other two baseline algorithms.

6. Conclusion

We presented a novel Bayesian model for registration uncertainty quantification in the space of diffeomorphic transformations. In contrast to previous approaches, our method significantly reduced the computational cost of the registration posterior inference effectively via (i) a low-dimensional representation of velocity fields that are associated with deformation fields; and (ii) an efficient Laplace's approximation to the posterior distribution, where the covariance was fully estimated by second-order methods. This work is the first step toward efficient probabilistic models of registration uncertainty quantification based on high-dimensional geometric transformations. Our future work will be investigating sampling methods to thoroughly assess our developed model uncertainty, and further extending its application to real clinical settings, e.g., real-time image-guided navigation system for neurosurgery. While in this paper we focused on the representation of time-dependent velocity fields in the context of LDDMM, our method is general to other transformation parameterizations such as stationary velocity fields [31].

Acknowledgments

This work is sponsored by NIH grants P41EB015898 and P41EB015902.

Appendix A. Appendix

This section includes the supplementary materials of computational derivatives for gradient term and second-order Hessian approximation.

A.1. Formulation of first derivative: gradient computation

As we derived in Section 2, the augmented energy with the adjoint variables $\hat{\tilde{v}}_t$, $\hat{u}_t$ can be formed as follows after adding geodesic constraints as Lagrangian multipliers,

$$\begin{aligned} E(\tilde{v}_0) = & -\ln p(\tilde{v}_0 \mid S, T, \sigma^2) + \int_0^1 \langle \hat{\tilde{v}}_t, \dot{\tilde{v}}_t + \underbrace{\mathrm{ad}^{\dagger}_{\tilde{v}_t} \tilde{v}_t}_{1} \rangle \\ & + \langle \hat{u}_t, \dot{u}_t + \underbrace{\tilde{v}_t + \tilde{\mathcal{D}} u_t * \tilde{v}_t}_{2} \rangle \, dt. \end{aligned} \tag{16}$$

Next, we derive the variation of $\tilde{E}$ w.r.t. the variables $\tilde{v}_0$ and $\tilde{u}_0$ for 1,2 respectively. To simplify the notation, we drop the subscript of time index t in the following sections.

Each part of the first derivative of $\partial_{\tilde{v}} E$ is,

$$1: \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=0} \int_0^1 \langle \hat{\tilde{v}}, \partial_t(\tilde{v} + \epsilon\delta\tilde{v}) + \mathrm{ad}^{\dagger}_{\tilde{v}+\epsilon\delta\tilde{v}}(\tilde{v} + \epsilon\delta\tilde{v}) \rangle$$

$$2: \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=0} \left(\int_0^1 \langle \hat{u}, \tilde{v} + \epsilon\delta\tilde{v} + \tilde{\mathcal{D}} u * (\tilde{v} + \epsilon\delta\tilde{v}) \rangle \right)$$

$$
\begin{aligned}
1,\,2:\ & \int_0^1 \langle \hat{v},\, \delta\dot{\tilde{v}} + \mathrm{ad}^{\dagger}_{\tilde{v}}\delta\tilde{v} + \mathrm{ad}^{\dagger}_{\delta\tilde{v}}\tilde{v}\rangle + \int_0^1 \langle \hat{u},\, \delta\tilde{v} + \widetilde{\mathcal{D}}u^{*}\delta\tilde{v}\rangle \\
&= \langle \hat{v},\, \delta\tilde{v}\rangle|_{t=0}^{t=1} \\
&\quad - \int_0^1 \langle \dot{\hat{v}},\, \delta\tilde{v}\rangle + \int_0^1 \langle \mathrm{ad}_{\tilde{v}}\hat{v} - \mathrm{ad}^{\dagger}_{\hat{v}}\tilde{v},\, \delta\tilde{v}\rangle \\
&\quad + \int_0^1 \langle \hat{u} + (\widetilde{\mathcal{D}}\tilde{u})^{T}\bigstar\hat{u},\, \delta\tilde{v}\rangle \\
&= \langle \hat{v},\, \delta\tilde{v}\rangle|_{t=0}^{t=1} \\
&\quad + \int_0^1 \langle -\dot{\hat{v}} + \mathrm{ad}_{\tilde{v}}\hat{v} - \mathrm{ad}^{\dagger}_{\hat{v}}\tilde{v} + \hat{u} + (\widetilde{\mathcal{D}}\tilde{u})^{T}\bigstar\hat{u},\, \delta\tilde{v}\rangle.
\end{aligned}
\tag{17}
$$

Each part of the first derivative $\partial_{\tilde{u}}E$ is

$$
\begin{aligned}
1,\,2:\ \partial_{\tilde{u}}E &= \frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0}\left(\int_0^1 \langle \hat{u},\, \partial_t(\tilde{u} + \epsilon\delta\tilde{u}) + \widetilde{\mathcal{D}}(\tilde{u} + \epsilon\delta\tilde{u})^{*}\tilde{v}\rangle\right) \\
&= \int_0^1 \langle \hat{u},\, \delta\dot{u} + \widetilde{\mathcal{D}}\delta\tilde{u}^{*}\tilde{u}\rangle \\
&= \langle \hat{u},\, \delta\tilde{u}\rangle|_{t=0}^{t=1} - \int_0^1 \langle \dot{\hat{u}},\, \delta\tilde{u}\rangle - \int_0^1 \langle \delta\tilde{u},\, \widetilde{\mathcal{D}}\tilde{u}^{*}\tilde{v} + \tilde{u}^{*}\widetilde{\nabla}\cdot\tilde{v}\rangle \\
&= \langle \hat{u},\, \delta\tilde{u}\rangle|_{t=0}^{t=1} - \int_0^1 \langle \delta\tilde{u},\, \dot{\hat{u}} + \widetilde{\mathcal{D}}\tilde{u}^{*}\tilde{v} + \tilde{u}^{*}\widetilde{\nabla}\cdot\tilde{v})\rangle.
\end{aligned}
\tag{18}
$$

By setting Eq. (18) to zero, we have $\dot{\hat{u}}_t = -\widetilde{\mathcal{D}}\delta\hat{u}_t{}^{*}\tilde{v}_t - \delta\hat{u}_t{}^{*}(\widetilde{\nabla}\cdot\tilde{v}_t)$. Putting Eq. (17) and Eq. (18) together, we obtain the adjoint equations of Eq. (13), subject to initial conditions

$$
\begin{aligned}
\hat{v}_1 &= 0, \\
\hat{u}_1 &= -\frac{1}{\sigma^2}\langle \nabla S_1,\, S_1 - T\rangle.
\end{aligned}
$$

A.2. Formulation of second-order derivation: Hessian computation

We develop a similar forward-backward approach to compute the second derivatives of $\tilde{v}$ and $\tilde{u}$. We first derive the second-order forward shooting equation of by adding auxiliary variables $\delta\hat{v}$ and $\delta\hat{u}$ on the geodesic shooting EpDiff Eq. (6) and deformation transport equation Eq. (7) respectively. For $\delta\tilde{v}$,

$$
\begin{aligned}
\partial^2_{\delta\tilde{v}}E &= \frac{\partial^2}{\partial\epsilon}\Big|_{\epsilon=0}\int_0^1 \langle \delta\hat{v},\, \partial_t(\delta\tilde{v} + \epsilon\delta\hat{v}) + \mathrm{ad}^{\dagger}_{\tilde{v}+\epsilon\delta\hat{v}}(\tilde{v} + \epsilon\delta\hat{v})\rangle \\
&= \frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0}\int_0^1 \langle \delta\hat{v},\, \delta\dot{\tilde{v}} + \mathrm{ad}^{\dagger}_{\tilde{v}+\epsilon\delta\hat{v}}\delta\tilde{v} + \mathrm{ad}^{\dagger}_{\delta\tilde{v}}(\tilde{v} + \epsilon\delta\hat{v})\rangle \\
&\quad + \int_0^1 \langle \delta\hat{v},\, \partial_t(\delta\hat{v} + \epsilon\delta\hat{v}) + \mathrm{ad}^{\dagger}_{\tilde{v}+\epsilon\delta\hat{v}}(\tilde{v} + \epsilon\delta\hat{v})\rangle \\
&= \int_0^1 \langle \delta\hat{v},\, \delta\dot{\tilde{v}} + \mathrm{ad}^{\dagger}_{\tilde{v}}\delta\tilde{v} + \mathrm{ad}^{\dagger}_{\delta\tilde{v}}\tilde{v}\rangle + \langle \delta\tilde{v},\, \dot{\tilde{v}} + \mathrm{ad}^{\dagger}_{\tilde{v}}\tilde{v}\rangle.
\end{aligned}
\tag{19}
$$

We then derive the forward shooting equation for $\delta\tilde{u}$ as follows,

$$
\begin{aligned}
\partial_{\delta\tilde{u}}E &= \frac{\partial^2}{\partial\epsilon}\Big|_{\epsilon=0}\int_0^1 \langle \delta\hat{u},\, \partial_t(\tilde{u} + \epsilon\delta\hat{u}) + \tilde{v} + \epsilon\delta\hat{v} + \widetilde{\mathcal{D}}(\tilde{u} + \epsilon\delta\hat{u})^{*}(\tilde{v} + \epsilon\delta\hat{v})\rangle \\
&= \frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0}\int_0^1 \langle \delta\hat{u},\, \dot{\delta\tilde{u}} + \delta\tilde{v} + \widetilde{\mathcal{D}}\delta\tilde{u}^{*}\tilde{v} + \widetilde{\mathcal{D}}(\tilde{u} + \epsilon\delta\hat{u})^{*}\delta\tilde{v}\rangle \\
&\quad + \langle \delta\hat{u},\, \partial_t(\tilde{u} + \epsilon\delta\hat{u}) + \tilde{v} + \epsilon\delta\hat{v} + \widetilde{\mathcal{D}}(\tilde{u} + \epsilon\delta\hat{u})^{*}(\tilde{v} + \epsilon\delta\hat{v})\rangle \\
&= \int_0^1 \langle \delta\hat{u},\, \delta\dot{\tilde{u}} + \delta\tilde{v} + \widetilde{\mathcal{D}}\delta\tilde{u}^{*}\tilde{v} + \widetilde{\mathcal{D}}\tilde{u}^{*}\delta\tilde{v}\rangle + \int_0^1 \langle \delta\hat{u},\, \dot{u} + \tilde{v} + \widetilde{\mathcal{D}}\tilde{u}^{*}\tilde{v}\rangle.
\end{aligned}
\tag{20}
$$

By setting Eq. (19) and Eq. (20) to zero, we obtain the second-order forward shooting for $\delta\tilde{v}$ and $\delta\tilde{u}$ as stated in Eq. (14).

Next we construct the augmented energy by adding second-order forward shooting equations as Lagrangian multipliers.

$$
\begin{aligned}
\mathrm{E}(\delta\hat{v}_0) \;&=\; -\ln p(\hat{v}_0 \mid S, T, \sigma^2) \\
&\quad + \int_0^1 \langle \delta\hat{v},\, \delta\dot{\tilde{v}} + \mathrm{ad}^{\dagger}_{\delta\tilde{v}}\tilde{v} + \mathrm{ad}^{\dagger}_{\tilde{v}}\delta\tilde{v}\rangle \\
&\quad + \langle \delta\hat{u},\, \delta\dot{u} + \tilde{\mathcal{D}}\delta\tilde{u}^{*}\tilde{v} - \tilde{\mathcal{D}}\tilde{u}^{*}\delta\tilde{v} - \delta\tilde{v}\rangle.
\end{aligned}
$$

Similarly, we derive the second-order backward shooting equation by adding auxiliary variables $\delta\hat{v}$ and $\delta\hat{u}$ on part 1 and 2. For $\delta\tilde{v}$,

$$
\begin{aligned}
\partial^2_{\delta\tilde{v}}E \;&=\; \frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0} \int_0^1 \langle \delta\hat{v},\, \partial_t(\delta\tilde{v} + \epsilon\delta\hat{v}) + \mathrm{ad}_{\delta\tilde{v}+\epsilon\delta\hat{v}}\tilde{v} + \mathrm{ad}_{\tilde{v}}(\delta\tilde{v} + \epsilon\delta\hat{v})\rangle \\
&\quad + \langle \delta\hat{u},\, \partial(\delta\tilde{v} + \epsilon\delta\hat{v}) + \tilde{\mathcal{D}}\delta\tilde{u}^{*}\tilde{v} + \tilde{\mathcal{D}}\tilde{u}^{*}(\delta\tilde{v} + \epsilon\hat{v}) + (\delta\tilde{v} + \epsilon\delta\hat{v})\rangle \\
&= \langle \delta\hat{v},\, \delta\tilde{v}\rangle|_{t=0}^{t=1} + \int_0^1 \langle \delta\hat{v},\, -\delta\dot{\tilde{v}} + \mathrm{ad}_{\tilde{v}}\delta\tilde{v} - \mathrm{ad}^{\dagger}_{\delta\tilde{v}}\tilde{v} + \delta\tilde{u} + (\tilde{\mathcal{D}}\tilde{u})^{T}\bigstar\delta\tilde{u}\rangle \\
&\quad + \int_0^1 \langle \delta\hat{v},\, -\delta\dot{\tilde{v}} + \mathrm{ad}_{\delta\tilde{v}}\tilde{v} - \mathrm{ad}^{\dagger}_{\tilde{v}}\delta\tilde{v} + (\tilde{\mathcal{D}}\delta\tilde{u})^{T}\bigstar\tilde{u}\rangle.
\end{aligned}
\tag{21}
$$

For $\delta\tilde{u}$,

$$
\begin{aligned}
\partial^2_{\delta\tilde{u}}E \;&=\; \frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0} \int_0^1 \langle \delta\hat{u},\, \partial_t(\delta\tilde{u} + \epsilon\delta\hat{u}) + \mathrm{ad}_{\delta\tilde{v}}\tilde{v} + \mathrm{ad}_{\tilde{v}}(\delta\tilde{v})\rangle \\
&\quad + \langle \delta\hat{u},\, \partial_t(\delta\tilde{u} + \epsilon\delta\hat{u}) + \tilde{\mathcal{D}}(\delta\tilde{u} + \epsilon\delta\hat{u})^{*}\tilde{v} + \tilde{\mathcal{D}}\tilde{u}^{*}\delta\tilde{v} - \delta\tilde{v}\rangle \\
&= \frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0} \int_0^1 \langle \delta\hat{u},\, \partial_t(\delta\tilde{u} + \epsilon\delta\hat{u}) + \tilde{\mathcal{D}}(\tilde{u} + \epsilon\delta\hat{u})^{*}\tilde{v}\rangle \\
&\quad + \langle \delta\hat{u},\, \partial_t(\delta\tilde{u} + \epsilon\delta\hat{u}) + \tilde{\mathcal{D}}(\tilde{u} + \epsilon\delta\hat{u})^{*}\delta\tilde{v}\rangle \\
&= \langle \delta\hat{u},\, \delta\tilde{u}\rangle|_{t=0}^{t=1} + \int_0^1 \langle -\delta\hat{u},\, \delta\dot{\tilde{u}} + \tilde{\mathcal{D}}\delta\tilde{u}^{*}\tilde{v} + \delta\tilde{v}^{*}\tilde{\nabla}\cdot\tilde{v}\rangle \\
&\quad + \int_0^1 \langle -\delta\hat{u},\, \delta\dot{\tilde{u}} + \tilde{\mathcal{D}}\tilde{u}^{*}\delta\tilde{v} + \tilde{u}^{*}\tilde{\nabla}\cdot\delta\tilde{v}\rangle.
\end{aligned}
\tag{22}
$$

By setting Eq. (21) and Eq. (22) to 0, we arrive the second-order backward equations of $\delta\tilde{v}$ and $\delta\tilde{u}$ as Eq. (15), subject to initial conditions $\delta\hat{u}_1 = -\frac{2}{\sigma^2}\mathcal{F}[\nabla S(1)\cdot\nabla S(1) + (S(1) - T)\cdot\nabla^2 S(1)]$ and $\delta\hat{v}_1 = 0$.

References

[1] Christensen GaryE, Rabbitt RichardD, Miller MichaelI. Deformable templates using large deformation kinematics. IEEE transactions on image processing 1996;5(10):1435–47.

[2] Christensen GaryE, Rabbitt RichardD, Miller MichaelI. A deformable neuroanatomy textbook based on viscous fluid mechanics. 27th Ann. Conf. on Inf. Sciences and Systems. 1993. p. 211–6.

[3] Miaomiao Zhang, PThomas Fletcher. Bayesian principal geodesic analysis in diffeomorphic image registration. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2014. p. 121–8.

[4] Miaomiao Zhang, PThomas Fletcher. Bayesian principal geodesic analysis for estimating intrinsic diffeomorphic image variability. Medical image analysis 2015;25(1):37–44.

[5] John Ashburner, Friston KarlJ. Unified segmentation. Neuroimage 2005;26(3):839–51.

[6] Sarang Joshi, Brad Davis, Matthieu Jomier, Guido Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage 2004;23:S151–60.

[7] Petter Risholm, Steve Pieper, Eigil Samset, Wells WilliamM. Summarizing and visualizing uncertainty in non-rigid registration. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2010. p. 554–61.

[8] Jie Luo, Matt Toews, Ines Machado, Sarah Frisken, Miaomiao Zhang, Frank Preiswerk. A feature-driven active framework for ultrasound-based brain shift compensation. arXiv preprint arXiv:180307682 2018.

[9] Petter Risholm, Eigil Samset, William Wells. Bayesian estimation of deformation and elastic parameters in non-rigid registration. International Workshop on Biomedical Image Registration. Springer; 2010. p. 104–15.

[10] Demian Wassermann, Matthew Toews, Marc Niethammer, William Wells. Probabilistic diffeomorphic registration: representing uncertainty. International Workshop on Biomedical Image Registration. Springer; 2014. p. 72–82.

[11] Jan Kybic. Bootstrap resampling for image registration uncertainty estimation without ground truth. IEEE Transactions on Image Processing 2010;19(1):64–73.

[12] Le Folgoc Loic, Herve Delingette, Antonio Criminisi, Nicholas Ayache. Quantifying registration uncertainty with sparse Bayesian modelling. IEEE transactions on medical imaging 2017;36(2):607–17.

[13] Miaomiao Zhang, PThomas Fletcher. Finite-dimensional lie algebras for fast diffeomorphic image registration. International Conference on Information Processing in Medical Imaging. Springer; 2015. p. 249–60.

[14] Miaomiao Zhang, Ruizhi Liao, Dalca AdrianV, Turk EsraA, Jie Luo, PEllen Grant. Frequency diffeomorphisms for efficient image registration. International Conference on Information Processing in Medical Imaging. Springer; 2017. p. 559–70.

[15] Xiao Yang, Marc Niethammer. Uncertainty quantification for LDDMM using a low-rank Hessian approximation. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2015. p. 289–96.

[16] Jian Wang, Wells WilliamM, Polina Golland, Miaomiao Zhang. Efficient laplace approximation for Bayesian registration uncertainty quantification. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2018. p. 880–8.

[17] Marcus DanielS, Wang TracyH, Jamie Parker, Csernansky JohnG, Morris JohnC, Buckner RandyL. Cross-sectional MRI data in young, middle aged, nondemented and demented older adults. Cognitive Neuroscie 2007:1489–507.
[18] MFaisal Beg, Miller MichaelI, Alain Trouvé, Laurent Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. International journal of computer vision 2005;61(2):139–57.
[19] Miaomiao Zhang, PThomas Fletcher. Fast diffeomorphic image registration via Fourier-approximated lie algebras. International Journal of Computer Vision 2018:1–13.
[20] FrançoisXavier Vialard, Laurent Risser, Daniel Rueckert, Cotter ColinJ. Diffeomorphic 3d image registration via geodesic shooting using an efficient adjoint calculation. International Journal of Computer Vision 2012;97(2):229–41.
[21] Avants BrianB, Epstein CharlesL, Murray Grossman, Gee JamesC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Medical image analysis 2008;12(1):26–41.
[22] Leventon M, Wells III WilliamM, Grimson WEricsL. Multiple view 2d-3d mutual information registration. Image Understanding Workshop. 20. Citeseer; 1997. p. 21.
[23] Arnol'd VI. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. Ann Inst Fourier 1966;16:319–61.
[24] Miller MI, Trouvé A, Younes L. Geodesic shooting for computational anatomy. Journal of Mathematical Imaging and Vision 2006;24(2):209–28. https://doi.org/10.1007/s10851-005-3624-0.
[25] Miaomiao Zhang, Nikhil Singh, PThomas Fletcher. Bayesian estimation of regularization and atlas building in diffeomorphic image registration. International conference on information processing in medical imaging. Springer; 2013. p. 37–48.
[26] Simpson IvorJA, Woolrich MarkW, ManuelJorge Cardoso, Cash DavidM, Marc Modat, Schnabel JuliaA. A bayesian a p. oach for spatially adaptive regularisation in non-rigid registration. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2013. p. 10–8.
[27] WingHung Wong, Li Bing. Laplace expansion for posterior densities of nonlinear functions of parameters. Biometrika 1992;79(2):393–8.
[28] Jr CliffordR Jack, Bernstein MattA, Fox NickC, Paul Thompson, Gene Alexander, Danielle Harvey. The Alzheimer's disease neuroimaging initiative (ADNI): Mri methods. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine 2008;27(4):685–91.
[29] Xiao Yang, Roland Kwitt, Martin Styner, Marc Niethammer. Quicksilver: fast predictive image registration-a deep learning approach. NeuroImage 2017;158:378–96.
[30] Guha Balakrishnan, Amy Zhao, Sabuncu MertR, John Guttag, Dalca AdrianV. An unsupervised learning model for deformable medical image registration. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018. p. 9252–60.
[31] Vincent Arsigny, Olivier Commowick, Xavier Pennec, Nicholas Ayache. A log-Euclidean framework for statistics on diffeomorphisms. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2006. p. 924–31.
[32] Jian Wang, Wei Xing, RobertMichael Kirby, Miaomiao Zhang. Data-driven model order reduction for diffeomorphic image registration. International conference on information processing in medical imaging. 2019.