Data-driven Model Order Reduction For Diffeomorphic Image Registration

Jian Wang[1*], Wei Xing[2*], Robert M. Kirby[2], and Miaomiao Zhang[1]

[1] Computer Science & Engineering, Washington University in St.Louis, MO, USA
[2] Scientific Computing & Imaging Institute, University of Utah, UT, USA

Abstract. This paper presents a data-driven model reduction algorithm to reduce the computational complexity of diffeomorphic image registration in the context of large deformation diffeomorphic metric mapping (LDDMM). In contrast to previous methods that repeatedly evaluate a full-scale regularization term governed by partial differential equations (PDEs) in the parameterized space of deformation fields, we introduce a reduced order model (ROM) to substantially lower the overall computational cost while maintaining accurate alignment. Specifically, we carefully construct the registration regularizer with a compact set of data-driven basis functions learned by proper orthogonal decomposition (POD), based on a key fact that the eigen spectrum decays extremely fast. This projected regularization in a low-dimensional subspace naturally leads to effective model order reduction with the underlying coherent structures well preserved. The iterative optimization involving computationally expensive PDE solvers is now carried out efficiently in a low-dimensional subspace. We demonstrate the proposed method in neuroimaging applications of pairwise image registration and template estimation for population studies.

1 Introduction

Diffeomorphic image registration has been successfully applied in the field of medical image analysis, as it maximally maintains the biological correctness of deformation fields in terms of object topology preservation. Examples of applications include alignment of functional data to a reference coordinate system [8, 27], anatomical comparison across individuals [24, 21], and atlas-based image segmentation [3, 15]. The problem of image registration is often formulated as constrained optimization over the transformation that well aligns a source image and a target image. A plethora of transformation models to today fit into various categories of parameterizations, e.g., stationary velocity fields that remain constant over time [2], and time-varying velocity fields in the framework of LDDMM [6]. We focus on the latter as it supports a distance metric in the space of diffeomorphisms that is critical for deformation-based statistical shape analysis, such as least squares mean estimator [15], geodesic regression [19, 23],

---

anatomical shape variability quantification [21], and groupwise geometrical shape comparison [24].

Despite the aforementioned advantages, one major challenge that hinders the widespread use of LDDMM is its high computational cost and large memory footprint [8, 6, 25]. The algorithm inference typically requires costly optimization, particularly on solving a full-scale regularization term defined on a dense image grid (e.g., a brain MRI with size of 128^3). Prior knowledge in the form of regularization is used to enforce the smoothness of transformation fields, also known as geodesic constraints in the space of diffeomorphisms, by solving a set of high-dimensional PDEs [17, 25]. This makes iterative optimization approaches, such as gradient descent [6], BFGS [20], or the Gauss-Newton method [4], computationally challenging. While improved computational capabilities have led to a substantial run-time reduction, such solution of a single pairwise image registration still takes tens of minutes to finish on dense 3D image grids [23].

In this paper, we aim at an approximate inference method that significantly lowers the computational complexity with little to no impact on the alignment accuracy. Our solution is motivated by the observation that smooth vector fields in the tangent space of diffeomorphisms can be characterized by a low-dimensional geometric descriptor, including a finite set of control points [10], or Fourier basis functions representing low frequencies [28, 29]. As a consequence, we hypothesize that the solution to high-dimensional PDE systems can be effectively approximated in a subspace with much lower dimensions. We develop a data-driven model reduction algorithm that constructs a low-dimensional subspace to approximate the original high-dimensional dynamic system for diffeomorphic image registration. We employ proper orthogonal decomposition (POD), a widely used technique for PDE systems, in which the approximating subspace is obtained from a discretized full-order model at selected time instances. A reduced-order model can then be constructed through Galerkin projection methods [9], where the PDEs are projected onto a compact set of POD eigen-functions.

To the best of our knowledge, this method has not yet been applied to large systems of PDEs such as the one employed in diffeomorphic registration. While we focus on the context of LDDMM, the theoretical tools developed in this paper are broadly applicable to all PDE-constrained diffeomorphic registration models. To evaluate the proposed method, we perform image registration of real 3D MR images and show that the accuracy of our estimated results is comparable to the state of the art, while with drastically lower runtime and memory consumption. We also demonstrate this method in the context of brain atlas building (mean template estimation) for efficient population studies.

2 Background: PDE-constrained Diffeomorphic Image Registration

In this section, we briefly review the LDDMM algorithm with PDE-constrained regularization [6, 25]. Let S be the source image and T be the target image defined on a d-dimensional 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 $\{\psi_t(x)\} : t \in [0,1]$ 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 over the transformation fields ψ_t as

$$\mathrm{E}(\psi_t) = \mathrm{Dist}(S \circ \psi_1, T) + \mathrm{Reg}(\psi_t), \tag{1}$$

where the distance function $\mathrm{Dist}(\cdot, \cdot)$ measures the image dissimilarity. Commonly used distance functions include sum-of-squared difference of image intensities [6], normalized cross correlation [5], and mutual information [26]. The regularization term $\mathrm{Reg}(\cdot)$ is a constraint that enforces the spatial smoothness of transformations, arising from a distance metric on the tangent space V of diffeomorphisms, i.e., an integral over the norm of time-dependent velocity fields $\{v_t(x)\} \in V$,

$$\mathrm{Reg}(\psi_t) = \int_0^1 (Lv_t, v_t)\, dt, \quad \text{with} \quad \frac{d\psi_t}{dt} = -D\psi_t \cdot v_t, \tag{2}$$

where $L : V \to V^*$ is a symmetric, positive-definite differential operator that maps a tangent vector $v_t \in V$ into its dual space as a momentum vector $m_t \in V^*$. We typically write $m_t = Lv_t$, or $v_t = Km_t$, with K being an inverse operator of L. The notation $(\cdot, \cdot)$ denotes the pairing of a momentum vector with a tangent vector, which is similar to an inner product. Here, the operator D denotes a Jacobian matrix and $\cdot$ represents element-wise matrix multiplication.

A geodesic curve with a fixed end point is characterized by an extremum of the energy function (2) that satisfies the Euler-Poincaré differential (EPDiff) equation [1, 17]

$$\frac{\partial v_t}{\partial t} = -K \left[(Dv_t)^T \cdot m_t + Dm_t \cdot v_t + m_t \cdot \mathrm{div}\, v_t \right], \tag{3}$$

where div is the divergence. This process in Eq. (3) is known as *geodesic shooting*, stating that the geodesic path $\{\phi_t\}$ can be uniquely determined by integrating a given initial velocity v_0 forward in time by using the rule (3).

Therefore, we rewrite the optimization of Eq. (1) equivalently as

$$\mathrm{E}(v_0) = \mathrm{Dist}(S \circ \psi_1, T) + (Lv_0, v_0), \text{ s.t. Eq. (3)}. \tag{4}$$

As suggested in (4), solving the time-dependent and nonlinear registration problem requires a large number of time steps and iterations. A full-order model is not affordable in many-query or real-time context of clinical problems. For example, an image-guided navigation system that employs registration algorithms to identify residual brain tumor during the surgery [16].

3 Our Model: Data-driven Model Order Reduction of Diffeomorphic Image Registration

We develop a POD-based model order reduction algorithm, particularly for the registration regularization term governed by high-dimensional PDEs (EPDiff in
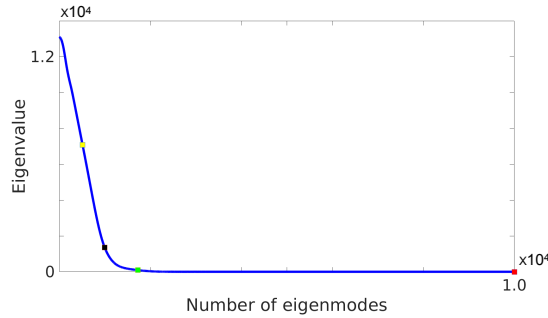
Fig. 1: An example eigenvalue plot of velocity fields generated from 2D synthetic data.

Eq. (4)), to approximate a subspace via a given set of velocity fields in an optimal least-square sense. We then derive a Galerkin projection (orthogonal projection) of EPDiff equations onto the subspace to obtain a reduced-order model.

3.1   Low-dimensional Subspace of Velocity Fields

Given a set of full-dimensional velocity fields $\{\mathbf{v}_t\} \in \mathcal{V}^q$, e.g., $q = 3 \times 128^3$ for a 3D discretized image grid with the size of 128^3, we are seeking an approximated subspace $\mathcal{U}^r = \mathrm{span}\{\mathbf{u}_1, \cdots, \mathbf{u}_r\} \subset \mathcal{V}^q$ $(r \ll q)$, where $\mathbf{u}_i$ is the basis, to best characterize our data. A projection from such low-dimensional subspace to the original space can be effectively performed by $\mathbf{v}_t = \mathbf{U}\boldsymbol{\alpha}_t$, where $\mathbf{U}^{q \times r} = [\mathbf{u}_1, \cdots, \mathbf{u}_r]$ and α_t is a r-dimensional vector representing factor coefficients. Here, we require the basis vectors to be orthonormal, i.e., $\mathbf{U}^T\mathbf{U} = \mathbf{I}$. The inverse projection can thus be written as $\boldsymbol{\alpha}_t = \mathbf{U}^T\mathbf{v}_t$. Our objective is to minimize the projection error defined in the tangent space of diffeomorphisms

$$\arg\min_{\mathbf{U}} \int \left(\mathbf{L}(\mathbf{v}_t - \mathbf{U}\boldsymbol{\alpha}_t), \mathbf{v}_t - \mathbf{U}\boldsymbol{\alpha}_t\right) \, dt. \tag{5}$$

where $\mathbf{L}$ is the discrete operator of L defined in Eq.(2). The minimization problem of Eq. (5), is the classic problem known as Karhunen-Loève decomposition or principal component analysis, and holds an equivalent solution to the following eigen decomposition problem of a covariance matrix $\mathbf{C}^{q \times q}$ [18, 22],

$$\mathbf{C}\mathbf{u}_i = \lambda_i\mathbf{u}_i, \text{ with } \mathbf{C} = \int \mathbf{L}(\mathbf{v}_t - \bar{\mathbf{v}})(\mathbf{v}_t - \bar{\mathbf{v}})^T dt,$$

where $\bar{\mathbf{v}} = \int \mathbf{v}_t dt$ is the mean field, and the basis $\mathbf{u}_i, i \in \{1, \cdots, r\}$ corresponds to the i-th eigenvector of $\mathbf{C}$ with associated eigenvalue λ_i. In practice, the covariance is empirically computed by a finite set of M observations (snapshots) over the full-scale dynamic system of $\mathbf{v}_t$, i.e., $\mathbf{C} \approx \frac{1}{M} \sum_{t=1}^{M} \mathbf{L}(\mathbf{v}_t - \bar{\mathbf{v}})(\mathbf{v}_t - \bar{\mathbf{v}})^T$.

   Due to a key fact that the spectrum of eigenvalues decays incredibly fast (as shown in Fig. 1), we propose to use an optimal set of eigen-functions to form

the projected subspace of velocity fields. An explicit way thus to formulate the projection error in Eq. (5) is

$$\sum_{i=r+1}^{q} \lambda_i, \text{ with } \lambda_1 > \cdots > \lambda_r \cdots > \lambda_q.$$

This closed-form solution provides an elegant way to measure the projection loss $e = 1 - (\sum_{i=1}^{r} \lambda_i / \sum_{i=1}^{q} \lambda_i)$, where r is typically chosen such as $e < 0.01$ [12, 18].

3.2 Reduced-order Regularization via Galerkin Projection

As introduced in the previous section, we developed a method to estimate a low-dimensional subspace of velocity fields that uniquely determines the geodesics of diffeomorphisms. We are now ready to construct a reduced-order model of image registration, subject to complex regularizations governed by high-dimensional PDEs (i.e., EPDiff). This procedure is known as *Galerkin projection* and has been widely used to reduce the high computational complexity of PDEs, or ODEs [13, 12, 18].

Consider the EPDiff in Eq. (3), we characterize a velocity field $\mathbf{v}_t$ by projecting it onto a finite-dimensional subspace $\mathcal{U}^r$ with much compact basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$. To simplify the notation, we drop the time index t of velocity fields in remaining sections. A discretized formulation of EPDiff equation in terms of matrix multiplication is

$$\frac{\partial \mathbf{v}}{\partial t} = -\mathbf{K} \left(\text{diag}(\mathbf{Lv})\mathbf{D}^T \mathbf{v} + \text{diag}(\mathbf{v})\mathbf{D}(\mathbf{Lv}) + \text{diag}(\mathbf{Lv}) \, \mathbf{div} \, \mathbf{v} \right),$$

$$= -\mathbf{K} \sum_{i=1}^{q} \left(\text{diag}(\mathbf{l}_i) v_i \mathbf{D}^T \mathbf{v} + v_i \mathbf{DLv} + \text{diag}(\mathbf{l}_i) v_i \, \mathbf{div} \, \mathbf{v} \right),$$

$$= -\mathbf{K} \sum_{i=1}^{q} \left(\text{diag}(\mathbf{l}_i)\mathbf{D}^T + \mathbf{DL} + \text{diag}(\mathbf{l}_i) \, \mathbf{div} \right) v_i \mathbf{v}, \tag{6}$$

where $\mathbf{v}$ is a q-dimensional vector, and $\text{diag}(\cdot)$ converts a vector to a diagonal matrix. The matrices $\mathbf{L}^{q \times q}$, $\mathbf{K}^{q \times q}$, $\mathbf{D}^{q \times q}$, and $\mathbf{div}^{q \times q}$ denote discrete analogs of the differential operator L with its inverse K, Jacobian matrix D, and divergence div obtained by finite difference schemes respectively. Here, $\mathbf{l}_i$ is the i-th column of the matrix $\mathbf{L}$ and v_i is the i-th element of vector $\mathbf{v}$.

By defining a composite operator $\mathbf{A}_i^{q \times q} \triangleq \mathbf{K}(\text{diag}(\mathbf{l}_i)\mathbf{D}^T + \mathbf{DL} + \text{diag}(\mathbf{l}_i) \, \mathbf{div})$, we write Eq. (6) as

$$\frac{\partial \mathbf{v}}{\partial t} = \sum_{i=1}^{q} \mathbf{A}_i v_i \mathbf{v}. \tag{7}$$

Next, we derive a reduced-order model of EPDiff via Galerkin projection by plugging $\mathbf{v} = \mathbf{U}^{q \times r}\boldsymbol{\alpha}$ into Eq. (7). We then have

$$\frac{\partial \mathbf{U}\boldsymbol{\alpha}}{\partial t} = \sum_{i=1}^{q} \mathbf{A}_i(\mathbf{U}\boldsymbol{\alpha})_i\mathbf{U}\boldsymbol{\alpha},$$

$$\frac{\partial \boldsymbol{\alpha}}{\partial t} = \mathbf{U}^T \sum_{i=1}^{q} \mathbf{A}_i(\sum_{j=1}^{r} \mathrm{U}_{ij}\alpha_j)\mathbf{U}\boldsymbol{\alpha} = \sum_{j=1}^{r}\sum_{i=1}^{q} \mathbf{U}^T\mathbf{A}_i\mathbf{U}\mathrm{U}_{ij}\alpha_j\boldsymbol{\alpha} \triangleq \sum_{j=1}^{r} \tilde{\mathbf{A}}_j\alpha_j\boldsymbol{\alpha}, \quad (8)$$

where U_{ij} the element of $\mathbf{U}$ in the i-th row and j-th column. Here, we define $\tilde{\mathbf{A}}_j^{r \times r} = \sum_{i=1}^{q} \mathbf{U}^T\mathbf{A}_i\mathbf{U}\mathrm{U}_{ij}$ as a reduced-order model operator of A_j. It is worthy to mention that the computation of $\tilde{\mathbf{A}}_i$ is a one-time cost accomplished offline. No further update is needed once a proper subspace is sought. Solution to this reduced-order model can be found by employing commonly used temporal differential schemes, e.g., Euler or Runge-Kutta Method, with an initial condition $\boldsymbol{\alpha}_0 = \mathbf{U}^T\mathbf{v}_0$.

4   ROM for Diffeomorphic Image Registration

In this section, we present a reduced-order model of LDDMM algorithm with geodesic shooting for diffeomorphic image registration. We run gradient descent on a projected initial velocity, represented by the loading coefficient $\boldsymbol{\alpha}_0$, entirely in a low-dimensional subspace. A geodesic path consequently generates a flow of diffeomorphisms by Eq. (2) after constructing the time-dependent velocity fields back in its original space using $\mathbf{v}_t = \mathbf{U}\boldsymbol{\alpha}_t$.

The redefined energy function of LDDMM in Eq. (4) with sum-of-squared dissimilarity between images is

$$\mathrm{E}(\boldsymbol{\alpha}_0) = \frac{1}{2\sigma^2}\|S \circ \psi_1 - T\|_2^2 + (\mathbf{L}\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_0), \;\; s.t. \text{ Eq. (8)}. \tag{9}$$

Here, we adopt a commonly used Laplacian operator $\mathbf{L} = (-\beta\Delta + \mathbf{I})^c$, where β is a positive weight parameter, c controls the level of smoothness, and $\mathbf{I}$ is an identity matrix.

Analogous to solving the optimal control problems in [28], we compute the gradient term by using a forward-backward sweep scheme. Below are the general steps for gradient computation (please refer to Alg. 1 for more details):

(i) Compute the gradient $\nabla_{\boldsymbol{\alpha}_1} \mathrm{E}$ of the energy (9) at $t = 1$ by integrating both the diffeomorphism ψ_t and the projected velocity field α_t forward in time, i.e.,

$$\nabla_{\boldsymbol{\alpha}_1} \mathrm{E} = \mathbf{K}\left(\frac{1}{\sigma^2}(S \circ \psi_1 - T) \cdot \nabla(S \circ \psi_1)\right). \tag{10}$$

(ii) Bring the gradient $\nabla_{\boldsymbol{\alpha}_1} \mathrm{E}$ back to $t = 0$. We obtain $\nabla_{\boldsymbol{\alpha}_0} \mathrm{E}$ by integrating reduced adjoint Jacobi field equations [7] backward in time as

$$\frac{d\hat{\alpha}}{dt} = -\mathrm{ad}_\alpha^\dagger \hat{h}, \quad \frac{d\hat{h}}{dt} = -\hat{\alpha} - \mathrm{ad}_\alpha\hat{h} + \mathrm{ad}_{\hat{h}}^\dagger\alpha, \tag{11}$$

where $\mathrm{ad}^\dagger$ is an adjoint operator and $\hat{h}, \hat{\alpha} \in V$ are introduced adjoint variables with an initial condition $\hat{h} = 0, \hat{\alpha} = \nabla_{\boldsymbol{\alpha}_1} \mathrm{E}$ at $t = 1$.

Algorithm 1: Optimization of Reduced order model for diffeomorphic image registration

---

Input: source image S, target image T
/* Online optimization */
1 Initialize $\boldsymbol{\alpha}_0 = 0$;
2 **repeat**
   | /* Forward shooting of $\boldsymbol{\alpha}_0$ */
3 | forward integrate the reduced-order model of EPDiff equation (8) to
   | generate $\{\boldsymbol{\alpha}_t\}$ at discrete time points;
   | /* Compute the diffeomorphic transformations ψ_t */
4 | integrate the transformation fields ψ by using (2) after constructing velocity
   | fields back to the original space via $\mathbf{v}_t = \mathbf{U}^T \boldsymbol{\alpha}_t$;
   | /* Compute gradient at time point $t = 1$ */
5 | compute the gradient $\nabla_{\boldsymbol{\alpha}_1} E$ by (10);
   | /* Propagate gradient back to time point $t = 0$ */
6 | integrate the reduced adjoint Jacobi field equations (11) backward in time
   | to obtain $\nabla_{\boldsymbol{\alpha}_0} E$.
7 | Update $\boldsymbol{\alpha}_0 \leftarrow \boldsymbol{\alpha}_0 - \varepsilon \nabla_{\boldsymbol{\alpha}_0} E$, where ε is the step size;
8 **until** *convergence*;

---

5 Experimental Evaluation

To demonstrate the effectiveness of our proposed model, we compare its performance with the state-of-the-art vector momentum LDDMM [23] in applications of pairwise image registration and atlas building. For fair comparison, we use $\beta = 3, c = 6$ for the L operator, and $\sigma = 0.01$ with 10 time-steps for Euler integration across all baseline algorithms.

Data. We applied the algorithm to 3D brain MRI scans from a public released resource Open Access Series of Imaging Studies (OASIS) for Alzheimer's disease [11]. The data includes fifty healthy subjects as well as disease, aged 60 to 96. To better evaluate the estimated transformations, we employed another fifty 3D brain MRI scans with manual segmentations from Alzheimer's Disease Neuroimaging Initiative(ADNI) [14]. All MRIs are of dimension $128 \times 128 \times 128$ with the voxel size of $1.25mm^3$. The images underwent downsampling, skull-stripping, intensity normalization, bias field correction and co-registration with affine transformation.

Experiments. We first tested our algorithm for pairwise image registration at different levels of projected dimension $r = 4^3, 8^3, 12^3, 20^3$ and compared the total energy formulated in Eq. (9). In order to find an optimal basis, we ran parallel programs of the full-scale EPDiff equation in (3) and generated a collection of snapshots to perform POD effectively. Since the learning process of basis functions were conducted offline with one-time cost for all experiments, we only focused on the exact runtime, memory consumption, and convergence rate of our model after the fact.

We validated registration results by examining the accuracy of propagated delineations for cortex (Cor), caudate (Caud), and corpus collusum (CC). After aligning all volumes to a reference image, we transformed the manual segmentation from the reference to other volumes by using the estimated deformations. We evaluated dice similarity coefficient (volume overlap) between the propagated segmentation and the manual segmentation for each structure.

We also ran both our method and the baseline algorithm to build an atlas from a set of 3D brain MRIs. We initialized the template image as an average of image intensities, and set the projected dimension as $r = 20^3$ as that was shown to be optimal in our eigen plots. In this experiment, we used a message passing interface (MPI) parallel programming implementation for all methods, and distributed data on four processors in total.

Results. Fig. 2 reports the total energy in formulation (9) averaged over 10 random selected pairs of test images for different values of projected dimensions. Our method arrives at the same solution at $r = 12^3$ and higher, which indicates that the estimated subspace has fairly recovered the result of full-scale registration algorithms. Fig. 2 also provides runtime and memory consumption across all three methods, including the baseline algorithm vector momemtum LDDMM. Our algorithm has substantially lower computational cost than vector momentum LDDMM performed in a full-dimensional space.
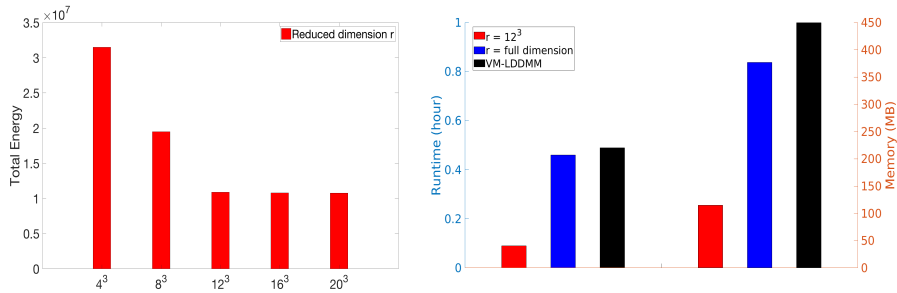

Fig. 2: Left: average total energy for different values of projected dimensions $r = 4^3, 8^3, 12^3, 16^3, 20^3$. Right: exact runtime and memory consumption for all methods.

Fig. 3 reports segmentation volume overlap on different brain structures, estimated from both our method and the baseline algorithm. It show that our algorithm is able to achieve comparable results, while offering significant improvements in computational efficiency. The right panel of Fig. 3 illustrates results for an example case from the study. We observe that the delineations achieved by transferring manual segmentations from the reference frame to the coordinate system of the target frame align fairly well with the manual segmentations. The left panel of Fig. 4 shows the axial and coronal slices from 12 of
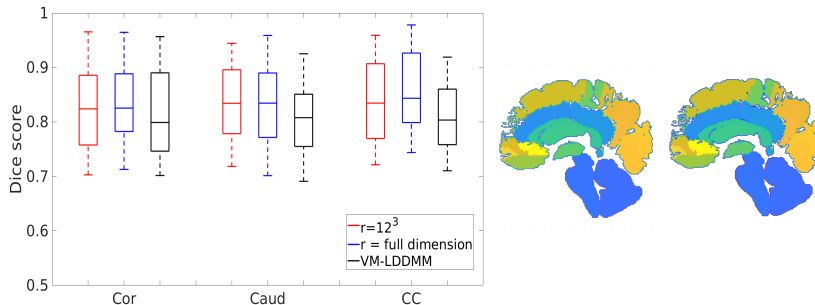

Fig. 3: Left: volume overlap between manual segmentations and propagated segmentations of three important regions cortex (Cor), caudate (Caud), and corpus collusum (CC); Middle: example ground truth segmentation; Right: propagated segmentation with three structures obtained by our method. 2D slices are shown for visualization only, all computations are carried out fully in 3D.

the selected 3D MRI dataset. The right panel demonstrates the atlas image estimated by our algorithm, followed by the atlas estimated by vector momenta LDDMM. The difference image between the two atlas results shows that our algorithm generated a very similar atlas to vector momenta LDDMM, but at a fraction of the time and memory cost (as illustrated on the right bottom panel of Fig. 4).

Fig. 5 shows the eigenvalue spectrum and convergence plot for both image registration (top) and atlas building (bottom). It is clear to see that our method conducted in a low-dimensional space is able to arrive at the same solution as the full dimensional scenario. We outperform the baseline algorithm vector momentum LDDMM, i.e., lower energy at the optimal solution.

6   Conclusion

We presented a data-driven model reduction algorithm for diffeomorphic image registration in the context of LDDMM with geodesic shooting. Our method is the first to simulate the high-dimensional dynamic system of diffeomorphisms in
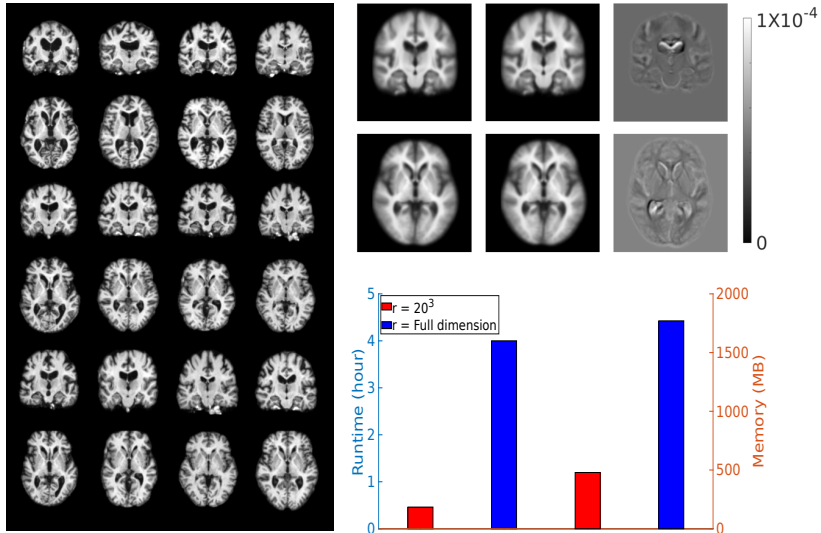
Fig. 4: Left: axial view and coronal view of twelve example brain MRIs selected from dataset. Right top: atlas images estimated by our method and vector momentum LD-DMM with difference map shown by side. Right bottom: a comparison of exact runtime and memory consumption.
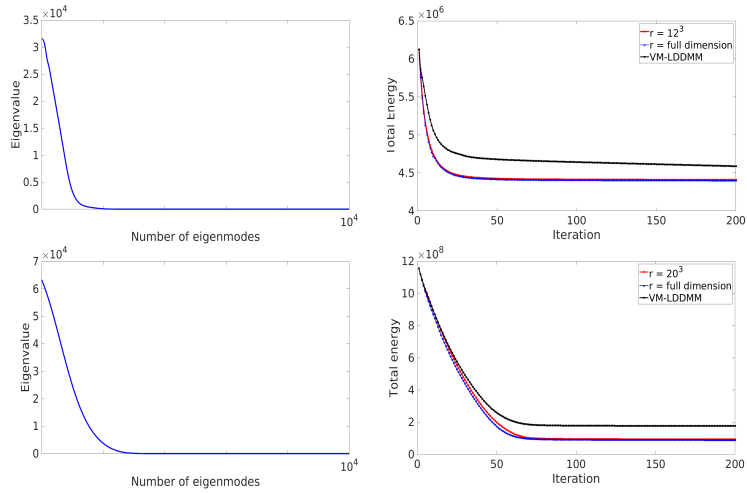

Fig. 5: Top to bottom: results of pairwise image registration vs. atlas building. Left to right: eigenvalue spectrum of velocity fields vs. total energy with an optimal projected dimension, a full dimension of our method, and vector momemtum LDDMM.

an approximated subspace via proper orthogonal decomposition and Galerkin projection. This approach substantially reduces the computational cost of diffeomorphic registration algorithms governed by high-dimensional PDEs, while preserving comparative accuracy. The theoretical tools developed in this paper are broadly applicable to all PDE-constrained diffeomorphic registration models with gradient-based optimization.

References

1. Arnol'd, V.I.: Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. Ann. Inst. Fourier **16**, 319–361 (1966)
2. Arsigny, V., Commowick, O., Pennec, X., Ayache, N.: A log-euclidean framework for statistics on diffeomorphisms. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 924–931. Springer (2006)
3. Ashburner, J., Friston, K.J.: Unified segmentation. Neuroimage **26**(3), 839–851 (2005)
4. Ashburner, J., Friston, K.J.: Diffeomorphic registration using geodesic shooting and gauss–newton optimisation. NeuroImage **55**(3), 954–967 (2011)
5. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Medical image analysis **12**(1), 26–41 (2008)
6. Beg, M.F., Miller, M.I., Trouvé, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. International journal of computer vision **61**(2), 139–157 (2005)
7. Bullo, F.: Invariant affine connections and controllability on lie groups. Tech. rep., technical Report for Geometric Mechanics, California Institute of Technology (1995)
8. Christensen, G.E., Rabbitt, R.D., Miller, M.I.: Deformable templates using large deformation kinematics. IEEE Transactions on Image Processing **5**(10), 1435–1447 (1996)
9. Cockburn, B., Shu, C.W.: Tvb runge-kutta local projection discontinuous galerkin finite element method for conservation laws. ii. general framework. Mathematics of computation **52**(186), 411–435 (1989)
10. Durrleman, S., Prastawa, M., Gerig, G., Joshi, S.: Optimal data-driven sparse parameterization of diffeomorphisms for population analysis. In: Biennial International Conference on Information Processing in Medical Imaging. pp. 123–134. Springer (2011)
11. Fotenos, A.F., Snyder, A., Girton, L., Morris, J., Buckner, R.: Normative estimates of cross-sectional and longitudinal brain volume decline in aging and ad. Neurology **64**(6), 1032–1039 (2005)
12. Hajek, B., Wong, E.: Stochastic processes in information and dynamical systems (1989)
13. Holmes, P., Lumley, J.L., Berkooz, G., Rowley, C.W.: Turbulence, coherent structures, dynamical systems and symmetry. Cambridge university press (2012)

14. Jack Jr, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., L. Whitwell, J., Ward, C., et al.: 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 **27**(4), 685–691 (2008)
15. Joshi, S., Davis, B., Jomier, M., Gerig, G.: Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage **23**, S151–S160 (2004)
16. Luo, J., Toews, M., Machado, I., Frisken, S., Zhang, M., Preiswerk, F., Sedghi, A., Ding, H., Pieper, S., Golland, P., et al.: A feature-driven active framework for ultrasound-based brain shift compensation. arXiv preprint arXiv:1803.07682 (2018)
17. Miller, M.I., Trouvé, A., Younes, L.: Geodesic shooting for computational anatomy. Journal of Mathematical Imaging and Vision **24**(2), 209–228 (2006). https://doi.org/10.1007/s10851-005-3624-0
18. Newman, A.J.: Model reduction via the karhunen-loeve expansion part i: An exposition. Tech. rep. (1996)
19. Niethammer, M., Huang, Y., Vialard, F.X.: Geodesic regression for image time-series. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 655–662. Springer (2011)
20. Polzin, T., Niethammer, M., Heinrich, M.P., Handels, H., Modersitzki, J.: Memory efficient lddmm for lung ct. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 28–36. Springer (2016)
21. Qiu, A., Younes, L., Miller, M.I.: Principal component based diffeomorphic surface mapping. Medical Imaging, IEEE Transactions on **31**(2), 302–311 (2012)
22. Shah, A., Xing, W., Triantafyllidis, V.: Reduced-order modelling of parameter-dependent, linear and nonlinear dynamic partial differential equation models. Proceedings. Mathematical, physical, and engineering sciences **473**(2200) (2017)
23. Singh, N., Hinkle, J., Joshi, S., Fletcher, P.T.: A vector momenta formulation of diffeomorphisms for improved geodesic regression and atlas construction. In: International Symposium on Biomedial Imaging (ISBI) (April 2013)
24. Vaillant, M., Miller, M.I., Younes, L., Trouvé, A.: Statistics on diffeomorphisms via tangent space representations. NeuroImage **23**, S161–S169 (2004)
25. Vialard, F.X., Risser, L., Rueckert, D., Cotter, C.J.: Diffeomorphic 3d image registration via geodesic shooting using an efficient adjoint calculation. International Journal of Computer Vision **97**(2), 229–241 (2012)
26. Wells, W., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R.: Multi-modal volume registration by maximization of mutual information. Medical image analysis (1996)
27. Zhang, M., Singh, N., Fletcher, P.T.: Bayesian estimation of regularization and atlas building in diffeomorphic image registration. In: Information Processing in Medical Imaging. pp. 37–48. Springer (2013)
28. Zhang, M., Fletcher, P.T.: Finite-dimensional lie algebras for fast diffeomorphic image registration. In: International Conference on Information Processing in Medical Imaging. pp. 249–260. Springer (2015)
29. Zhang, M., Liao, R., Dalca, A.V., Turk, E.A., Luo, J., Grant, P.E., Golland, P.: Frequency diffeomorphisms for efficient image registration. In: International conference on information processing in medical imaging. pp. 559–570. Springer (2017)