PDE-constrained LDDMM via geodesic shooting and inexact Gauss-Newton-Krylov optimization using the incremental adjoint Jacobi equations

Monica Hernandez

Computer Science Department
Aragon Institute on Engineering Research (I3A)
University of Zaragoza, Spain

E-mail: mhg@unizar.es

Abstract. The class of non-rigid registration methods proposed in the framework of PDE-constrained Large Deformation Diffeomorphic Metric Mapping is a particularly interesting family of physically meaningful diffeomorphic registration methods. Inexact Gauss-Newton-Krylov optimization has shown an excellent numerical accuracy and an extraordinarily fast convergence rate in this framework. However, the Galerkin representation of the non-stationary velocity fields does not provide proper geodesic paths. In this work, we propose a method for PDE-constrained LDDMM parameterized in the space of initial velocity fields under the EPDiff equation. The derivation of the gradient and the Hessian-vector products are performed on the final velocity field and transported backward using the adjoint and the incremental adjoint Jacobi equations. This way, we avoid the complex dependence on the initial velocity field in the computations. We also avoid the computation of the adjoint equation and its incremental counterpart that has been recently identified as a subtle problem in PDE-constrained LDDMM. The proposed method provides geodesics in the framework of PDE-constrained LDDMM, and it shows performance competing with benchmark PDE-constrained LDDMM and EPDiff-LDDMM methods.

1. Introduction

In the last two decades, diffeomorphic image registration has arisen as a predominant paradigm for deformable image registration (Sotiras et al. 2013). Diffeomorphic registration methods compute transformations fundamental in Computational Anatomy applications (Kobatake & Masutani 2017). Although the differentiability and invertibility of a diffeomorphism constitute fundamental properties for its use in Computational Anatomy, the diffeomorphic constraint does not necessarily guarantee that a transformation computed with a given method is physically meaningful for the clinical domain of interest. PDE-constrained Large Deformation Diffeomorphic Metric Mapping (PDE-constrained LDDMM) methods constitute an appealing paradigm for computing transformations under plausible physical models of interest.

The first class of PDE-constrained LDDMM methods arose with the work of Younes et al. (Younes 2007), followed by the proposals in (Ashburner & Friston 2011, Vialard et al. 2011, Zhang & Fletcher 2015). In these works, the

transformations are parameterized by time-varying velocity fields that satisfy the Euler-Poincare differential (EPDiff) equation (Holm et al. 1998), an evolution partial differential equation deriving from an application of the Euler-Poincare principle. The EPDiff equation can be found in the physical model of soliton dynamics driven by the Camassa-Holm equation (Holm et al. 2004). This physical constraint allows formulating the problem in the space of initial velocity fields. This guarantees that the obtained transformations belong to geodesic paths of diffeomorphisms, which is desirable in important Computational Anatomy applications (Miller et al. 2005). For this class of methods, the dependence of the energy functional on the initial velocity field is complex, and the optimization is usually implemented using gradient-descent.

The class of PDE-constrained LDDMM methods proposed in (Hart et al. 2009, Vialard et al. 2011, Mang & Biros 2015, Mang & Biros 2016, Mang & Ruthotto 2017) is especially interesting, where the physical PDEs are imposed using hard constraints. These methods model the problem using a PDE-constrained variational formulation. Numerical optimization is approached using gradient-descent (Hart et al. 2009, Vialard et al. 2011, Mang & Biros 2015) and second-order optimization in the form of inexact reduced Newton-Krylov methods (Mang & Biros 2015, Mang & Biros 2016, Mang & Ruthotto 2017). The gradient and the Hessian-vector products are computed by the differentiation of the augmented energy functional using optimal control theory methods. The constrained optimization approach provides the versatility to impose different physical models to the computed transformations by just adding the PDEs associated to the problem as hard constraints. In particular, inexact Gauss-Newton-Krylov optimization shows an excellent numerical accuracy and an extraordinarily fast convergence rate (Mang & Biros 2015, Mang & Biros 2016, Mang & Ruthotto 2017). However, the Galerkin representation of the non-stationary velocity fields used for diffeomorphism parameterization does not provide proper geodesic paths.

The purpose of this article is to parameterize the compressible method for PDE-constrained diffeomorphic registration with inexact Gauss-Newton-Krylov optimization in (Mang & Biros 2015) in the space of initial velocity fields. Rather than computing the gradient and the Hessian-vector products from the differentiation of the augmented energy functional on the initial velocity field, we perform the computations on the final velocity field and transport the computations backward using the adjoint and the incremental adjoint Jacobi equations. This way, we avoid the complex dependence on the initial velocity field in the computations. We also avoid the computation of the adjoint equation and its incremental counterpart that has been recently identified as problematic (Mang & Ruthotto 2017). The proposed method provides geodesics in the framework of PDE-constrained LDDMM. The method has been favorably compared and evaluated with the natural formulation and derivation of the PDE-constrained LDDMM problem in (Mang & Biros 2015) in the space of initial velocity fields, the benchmark PDE-constrained LDDMM (Mang & Biros 2015), and the EPDiff-LDDMM methods in (Vialard et al. 2011, Zhang & Fletcher 2018). The evaluation has been performed with the manual segmentations of the Non-Rigid Image Registration Evaluation Project (NIREP) database (Song et al. 2010).

Methodologically, our formulation improves the existing state-of-the-art methods that may be competitive with our proposal in the following ways:

- Mang et al. (Mang & Biros 2015). The original formulation of PDE-constrained LDDMM is not a geodesic shooting method. Therefore, the solutions depart slightly from belonging to geodesic paths due to numerical errors. We provide

a parameterization for PDE-constrained LDDMM in the space of vector fields satisfying the EPDiff equation. This constraint is a guarantee that the solutions belong to geodesic paths. This property is of importance in many Computational Anatomy applications.

- Zhang et al. (Zhang & Fletcher 2018). The spatial formulation is a geodesic shooting method with gradient-descent optimization. The derivation of the gradient is performed on the final velocity field and transported backward using the adjoint Jacobi equation. The numerical implementation uses finite differences and first-order methods for PDE integration. Our method is a PDE-constrained geodesic shooting method with second-order optimization. We extend the idea of backward transport to the Hessian-vector product. The numerical implementation uses spectral differentiation with high-order methods for PDE integration. Second-order optimization improves gradient-descent in terms of the convergence rate and the computational time.

- Vialard et al. (Vialard et al. 2011). The formulation is a geodesic shooting method with gradient-descent optimization. The update equation is written in terms of the adjoint variable. The numerical implementation uses finite differences and first-order methods for PDE integration. Our improvements with respect to Zhang et al. method also apply to this method.

- Yang et al. (Yang & Niethammer 2015). The PDE-EPDiff LDDMM formulation in Appendix A would be the natural option to extend the formulation of Mang et al. to a geodesic shooting approach. The Hessian provided in Yang et al. in the context of registration uncertainty estimation is obtained using a similar derivation. However, the dependence on the initial velocity field in the computations is complex. The method involves the computation of the adjoint equation and its incremental counterpart that has been recently identified as a subtle problem in PDE-constrained LDDMM. As a result, the method shows itself numerically unstable. Our formulation results into a much simpler dependence of the initial velocity field in the computations. In addition, the idea of backward transport avoids the direct computation of the adjoint equation. As a result, our method is numerically stable.

In the following, Section 2 revisits the methods more related to our work. Section 3 presents our proposed method. Section 4 gathers the implementation details of the benchmark and our proposed method. Section 5 shows the experiments for the evaluation of our method. Finally, Section 6 gathers the most remarkable conclusions of our work.

2. Background methods

2.1. Large Deformation Diffeomorphic Metric Mapping (LDDMM)

Let I_0, and I_1 be the source and the target images defined on the image domain $\Omega \subseteq \mathbb{R}^d$. We denote the Riemannian manifold of diffeomorphisms on Ω by $Diff(\Omega)$. V is the tangent space of the Riemannian structure at the identity diffeomorphism, *id*. V is made of smooth vector fields on Ω. The Riemannian metric is defined from the scalar product in V

$$\langle v, w \rangle_V = \langle Lv, w \rangle_{L^2} = \int_\Omega \langle Lv(x), w(x) \rangle d\Omega, \qquad (1)$$

where $L = (Id - \alpha \Delta)^s, \alpha > 0, s \in \mathbb{N}$ is the invertible self-adjoint differential operator associated with the differential structure of $Diff(\Omega)$. We denote with K to the inverse of L. The degree of smoothness in V depends on the smoothness of the differential operator L.

The LDDMM variational problem is given by the minimization of the energy functional

$$E(v) = \int_0^1 \langle Lv_t, v_t \rangle_{L^2} dt + \frac{1}{\sigma^2} \|I_0 \circ (\phi_1^v)^{-1} - I_1\|_{L^2}^2. \tag{2}$$

The problem is posed in the space of time-varying smooth flows of velocity fields in V, $v \in L^2([0, 1], V)$. Given the flow $v : [0, 1] \to V, v_t : \Omega \to \mathbb{R}^d \in V$, the diffeomorphism ϕ_1^v is defined as the solution at time 1 to the transport equation

$$d_t \phi_t^v = v_t \circ \phi_t^v \tag{3}$$

with initial condition $\phi_0^v = id$. The transformation $(\phi_1^v)^{-1}$ computed from the minimum of $E(v)$ is the diffeomorphism that solves the LDDMM registration problem between I_0 and I_1. The optimization of Equation 2 was originally approached in (Beg et al. 2005) using gradient-descent in $L^2([0, 1], V)$, yielding the update equation

$$v_t^{n+1} = v_t^n - \epsilon(\nabla_v E(v))_t^n. \tag{4}$$

Gauss-Newton optimization was proposed in (Hernandez 2014), showing the efficiency improvement expected with second-order optimization methods.

2.2. LDDMM in the space of initial velocity fields (EPDiff-LDDMM)

The geodesics of $Diff(\Omega)$ under the right-invariant Riemannian metric are uniquely determined by the time-varying flows of velocity fields that satisfy the Euler-Poincaré equation (EPDiff) (Holm et al. 1998)

$$\partial_t v_t = -ad_{v_t}^\dagger v_t = -Kad_{v_t}^* Lv_t = -K[(Dv_t)^T Lv_t + D(Lv_t)v_t + Lv_t \nabla \cdot v_t] \tag{5}$$

with initial condition $v_0 \in V$. The operator $ad_v^\dagger w$ is the transpose of operator

$$ad_v w = Dw \cdot v - Dv \cdot w. \tag{6}$$

LDDMM can be posed in the space of initial velocity fields

$$E(v_0) = \langle Lv_0, v_0 \rangle_{L^2} + \frac{1}{\sigma^2} \|I_0 \circ (\phi_1^v)^{-1} - I_1\|_{L^2}^2, \tag{7}$$

where $(\phi_1^v)^{-1}$ is the solution at time 1 to the transport equation of the flow v_t that satisfies the EPDiff equation for v_0.

The optimization of Equation 7 was originally approached using gradient-descent in V (Younes 2007), yielding the update equation

$$v_0^{n+1} = v_0^n + \epsilon \nabla_{v_0} E(v_0)^n. \tag{8}$$

More recently, it has been proposed to compute the gradient of the image similarity energy E_{img} at $t = 1$ and to integrate backward the reduced adjoint Jacobi equations (Bullo 1995, Hinkle et al. 2014)

$$\partial_t U_t + ad_{v_t}^\dagger U_t = 0 \text{ in } \Omega \times [0, 1) \tag{9}$$

$$\partial_t w_t + U_t - ad_{v_t} w_t + ad_{w_t}^\dagger v_t = 0 \text{ in } \Omega \times [0, 1) \tag{10}$$

with initial conditions $U(1) = \nabla_{v_1} E_{\text{img}}(v_0)$ and $w(1) = 0$, to get the gradient update at $t = 0$ (Zhang & Fletcher 2015, Zhang et al. 2017, Zhang & Fletcher 2018),

$$\nabla_{v_0} E(v_0)^n = 2v_0^n + w(0)^n. \tag{11}$$

2.3. PDE-constrained LDDMM (PDE-LDDMM)

The variational problem is given by a slight modification of Equation 2

$$E(v) = \int_0^1 \langle Lv_t, v_t \rangle_{L^2} dt + \frac{1}{\sigma^2} \|m(1) - I_1\|_{L^2}^2, \tag{12}$$

where $m(t) = I_0 \circ (\phi_t^v)^{-1}$ is computed from the solution of the state equation

$$\partial_t m(t) + \nabla m(t) \cdot v_t = 0 \text{ in } \Omega \times (0, 1], \tag{13}$$

with initial condition $m(0) = I_0$ in Ω. The combination of Equations 12 and 13 leads to a PDE-constrained optimization problem.

Optimization can be performed combining the method of Lagrange multipliers with gradient-descent (Hart et al. 2009) or second-order inexact Newton-Krylov methods (Mang & Biros 2015). The gradient is computed from

$$\partial_t m(t) + \nabla m(t) \cdot v_t = 0 \text{ in } \Omega \times (0, 1] \tag{14}$$

$$-\partial_t \lambda(t) - \nabla \cdot (\lambda(t) \cdot v_t) = 0 \text{ in } \Omega \times [0, 1) \tag{15}$$

$$(\nabla_v E(v))_t = Lv_t + \lambda(t) \cdot \nabla m(t) \text{ in } \Omega \times [0, 1] \tag{16}$$

subject to the initial and final conditions $m(0) = I_0$ and $\lambda(1) = -\frac{2}{\sigma^2}(m(1) - I_1)$. Equations 14 and 15 correspond with the state and adjoint equations, respectively.

The Hessian is computed from

$$\partial_t \delta m(t) + \nabla \delta m(t) \cdot v_t + \nabla m(t) \cdot \delta v(t) = 0 \text{ in } \Omega \times (0, 1] \tag{17}$$

$$-\partial_t \delta \lambda(t) - \nabla \cdot (\delta \lambda(t) \cdot v_t) - \nabla \cdot (\lambda(t) \cdot \delta v(t)) = 0 \text{ in } \Omega \times [0, 1) \tag{18}$$

$$(H_v E(v) \delta v)_t = L \delta v(t) + \delta \lambda(t) \cdot \nabla m(t) + \lambda(t) \cdot \nabla \delta m(t) \text{ in } \Omega \times [0, 1] \tag{19}$$

subject to the initial and final conditions $\delta m(0) = 0$ and $\delta \lambda(1) = -\frac{2}{\sigma^2} \delta m(1)$. Equations 17 and 18 correspond with the incremental state and incremental adjoint equations, respectively.

The minimization using inexact Newton-Krylov optimization yields to the update equation

$$v_t^{n+1} = v_t^n + \epsilon \delta v_t^n, \tag{20}$$

where δv^n is computed from preconditioned conjugate gradient (PCG) on the reduced system

$$H_v E(v^n) \delta v^n = -\nabla_v E(v^n), \tag{21}$$

with preconditioner K. Since the update equation is written on the reduced space V, the minimization is a reduced space optimization method.

By construction, the Hessian is positive definite in the proximity of a local minimum. However, it can be indefinite or singular far away from the solution. In this case, the search directions obtained with PCG are not guaranteed to be descent directions. In order to overcome this problem, one can use a Gauss-Newton approximation dropping expressions of $H_v E(v) \delta v$ to guarantee that the matrix is definite positive. In particular, one can drop the terms in Equations 18 and 19 involving the adjoint variable λ.

3. The proposed method (Jacobi PDE-EPDiff LDDMM)

As in PDE-LDDMM, the PDE-constrained problem is given by the minimization of the energy functional

$$E(v_0) = \langle Lv_0, v_0 \rangle_{L^2} + \frac{1}{\sigma^2} \|m(1) - I_1\|_{L^2}^2, \tag{22}$$

subject to the EPDiff and the state equations

$$\partial_t v_t + ad_{v_t}^\dagger v_t = 0 \text{ in } \Omega \times (0, 1] \tag{23}$$

$$\partial_t m(t) + \nabla m(t) \cdot v_t = 0 \text{ in } \Omega \times (0, 1], \tag{24}$$

with initial conditions $v(0) = v_0$ and $m(0) = I_0$, respectively.

Optimization is performed combining the method of Lagrange multipliers with inexact Gauss-Newton-Krylov methods in the following way. Let $w : \Omega \times [0, 1] \to \mathbb{R}^d$ and $\lambda : \Omega \times [0, 1] \to \mathbb{R}$ be the Lagrange multipliers associated with the EPDiff and the state equations. We build the augmented Lagrangian

$$E_{aug}(v_0) = E(v_0) + \int_0^1 \langle w(t), \partial_t v(t) + ad_{v_t}^\dagger v_t \rangle_{L^2} dt$$
$$+ \int_0^1 \langle \lambda(t), \partial_t m(t) + \nabla m(t) \cdot v_t \rangle_{L^2} dt. \tag{25}$$

The image similarity energy gradient is computed at $t = 1$ from

$$\nabla_{v_1} E_{\text{img}}(v_0) = K(\lambda(1) \cdot \nabla m(1)) \tag{26}$$

integrated backward using the EPDiff and the reduced adjoint Jacobi equations

$$\partial_t v_t + ad_{v_t}^\dagger v_t = 0 \text{ in } \Omega \times (0, 1] \tag{27}$$

$$\partial_t m(t) + \nabla m(t) \cdot v_t = 0 \text{ in } \Omega \times (0, 1] \tag{28}$$

$$\partial_t U_t + ad_{v_t}^\dagger U_t = 0 \text{ in } \Omega \times [0, 1) \tag{29}$$

$$\partial_t w_t - ad_{v_t} w_t + ad_{w_t}^\dagger v_t + U_t = 0 \text{ in } \Omega \times [0, 1) \tag{30}$$

to obtain $\nabla_{v_0} E(v_0) = 2v_0 + w(0)$. With this approach, the integration of the adjoint equation is not needed.

The second-order variations of the augmented Lagrangian on w and λ yield the incremental EPDiff and incremental state equations, needed for the computation of the Hessian-vector product. Thus,

$$\partial_t \delta v_t + ad_{\delta v_t}^\dagger v_t + ad_{v_t}^\dagger \delta v_t = 0 \text{ in } \Omega \times (0, 1] \tag{31}$$

$$\partial_t \delta m(t) + \nabla \delta m(t) \cdot v_t + \nabla m(t) \cdot \delta v_t = 0 \text{ in } \Omega \times (0, 1] \tag{32}$$

with initial conditions $\delta v(0) = 0$ and $\delta m(0) = 0$.

The Hessian-vector product $H_{v_0} E(v_0) \delta v_0$ is computed from the Hessian-vector product of E_{img} at $t = 1$, which is integrated backward using the reduced incremental adjoint Jacobi equations

$$\partial_t \delta U_t + ad_{\delta v_t}^\dagger U_t + ad_{v_t}^\dagger \delta U_t = 0 \text{ in } \Omega \times [0, 1) \tag{33}$$

$$\partial_t \delta w_t - ad_{\delta v_t} w_t - ad_{v_t} \delta w_t + ad_{\delta w_t}^\dagger v_t + ad_{w_t}^\dagger \delta v_t +$$
$$\delta U_t = 0 \text{ in } \Omega \times [0, 1) \tag{34}$$

with initial conditions $\delta U(1) = K(\delta\lambda(1) \cdot \nabla m(1)) + K(\lambda(1) \cdot \nabla\delta m(1))$, and $\delta w(1) = 0$ yielding

$$H_{v_0} E(v_0)\delta v_0 = 2\delta v_0 + \delta w(0). \tag{35}$$

The proposed Gauss-Newton approximation drops $K(\lambda(1) \cdot \nabla\delta m(1))$ from the expression of $\delta U(1)$, and the ad_v and ad_w terms from the incremental adjoint Jacobi equation on δw.

   The minimization using a second-order inexact Gauss-Newton-Krylov method yields to the update equation

$$v_0^{n+1} = v_0^n + \epsilon\delta v_0^n, \tag{36}$$

where δv_0^n is computed from CG on the system

$$H_{v_0} E(\delta v_0{}^n) = -\nabla_{v_0} E(v_0^n). \tag{37}$$

We also consider PCG with the gradient and the Hessian computed on L^2 instead of V as originally proposed in (Mang & Biros 2015).

4. Implementation details

In this section we gather the most critical implementation details of the methods used in the experimental section: our Jacobi PDE-EPDiff LDDMM method, PDE-EPDiff LDDMM (i.e., the natural formulation and derivation of the PDE-constrained LDDMM problem in (Mang & Biros 2015) in the space of initial velocity fields), PDE-LDDMM (Mang & Biros 2015), and Zhang EPDiff-LDDMM (Zhang & Fletcher 2018). The implementation of Vialard EPDiff-LDDMM (Vialard et al. 2011) mimics their publicly available C++ implementation. The derivation of PDE-EPDiff LDDMM used in this work can be found in the Appendix.

4.1. Diffeomorphism parameterization

4.1.1. PDE-LDDMM. We use a parameterization of the steady and time-varying velocity fields based on the Galerkin factorization

$$v_t(x) = \sum_{r=1}^{R} b_r(t)v_r(x), x \in \Omega, \tag{38}$$

where $b_r(t)$ is the basis of Chebyshev polynomials, and R is the dimension of the basis expansion. In consequence, the numerical optimization is performed in the set of coefficients $\{v_1, \cdots, v_R\}$ instead of a dense time sampling of v. With this representation, steady and time-varying velocity field flows can be manipulated similarly, and the solutions of the state and the incremental state equations and their respective adjoints can be solved in a dense sampling of $[0, 1]$. This parameterization was proposed in (Mang & Biros 2015). In this work, we use $R = 1$ for the stationary, and $R = 5$ for the non-stationary parameterization.

4.1.2. Zhang EPDiff-LDDMM, PDE-EPDiff-LDDMM, and Jacobi PDE-EPDiff LDDMM. Diffeomorphisms are parameterized from time-varying velocity field flows satisfying the EPDiff equation. With this representation, the whole sampling of v in $[0, 1]$ needs to be stored in memory.

4.2. Regularization

In this work, we consider H^2-regularization with $L = (Id - \alpha\Delta)^s, \alpha > 0, s \in \mathbb{N}$. Regularization parameters were selected after performing the search of the optimal parameters in the registration experiments conducted in this work. As a result, we selected parameters $\alpha = 0.0025$, and $s = 2$. The parameter σ, which weights the contribution of the regularizer and the image similarity term, was selected equal to one. This selection provided the minimum image similarity error while avoiding the development of numerical problems due to weak regularization, yielding a fair comparison of all the methods.

4.3. Optimization

4.3.1. Gradient-descent. Gradient-descent optimization was used for Zhang EPDiff-LDDMM as initially proposed. Backtracking line-search was used for the selection of the step size for the first iteration. We found that the selected value provided an aceptable energy convergence rate for the remaining iterations, so we skipped the extra computations required by line-search. We show results after 50 iterations since the inner times the outer iterations in second-order methods were selected to reach this value.

4.3.2. Gauss-Newton-Krylov optimization. Gauss-Newton-Krylov optimization was used for the PDE-LDDMM methods. The step size was selected equal to 1 since it is a reasonable choice for Newton-type optimization approaches. The number of PCG and CG iterations was set equal to 5. For L^2-optimization, the PCG tolerance was selected from $\tau = \min\left(0.5, \sqrt{\frac{\|\nabla_v E(v^n)\|_2}{\|\nabla_v E(v^0)\|_2}}\right)$. For V-optimization, the CG tolerance was selected from $\tau = \min\left(0.1, \sqrt{\frac{\|\nabla_v E(v^n)\|_2}{\|\nabla_v E(v^0)\|_2}}\right)$. The optimization was stopped after 10 iterations, enough to obtain acceptable MSE values for the second-order methods.

4.4. Differentiation

The computation of differentials is approached using Fourier spectral methods as an alternative to commonly used finite difference approximations (Trefethen 2000). The use of spectral differentiation and integration was introduced in diffeomorphic registration literature in (Mang & Biros 2015). Spectral methods allow solving ODEs and PDEs to high accuracy in simple domains for problems involving smooth data. During the development phase of the algorithms, we noticed that the use of spectral differentiation allowed to resolve several numerical difficulties. Spectral differentiation provides a numerical scheme spectrally accurate in space. The scheme displays minimal numerical diffusion and does not require flux limiters. In addition, spectral differentiation allows using the same discretization models for solving the subsystems of hyperbolic and elliptic mixed-type optimality conditions.

4.5. ODE solvers and time integrals

To integrate all the differential equations involved in the PDE-LDDMM methods we used the Bogacki-Shampine Runge-Kutta method of order 3. The solutions were computed at the Chebyshev-Gauss-Lobatto discretization of the temporal domain $[0, 1]$. The number of time steps was selected sufficiently large to avoid violations

of the CFL condition, and do not observe stability problems. Once that stability is achieved, RK methods provide accurate solutions to the equations.

For stationary PDE-LDDMM, Zhang EPDiff-LDDMM, PDE-EPDiff LDDMM, and Jacobi PDE-EPDiff LDDMM, the temporal domain was discretized into 30 samples. For non-stationary PDE-LDDMM the temporal sampling needed to be increased up to 60 to yield stable results.

Time integrals were computed using spectral methods. In particular, we used Clenshaw-Curtis quadrature rule in a Chebyshev-Gauss-Lobatto discretization of the interval $[0,1]$ to avoid Runge's phenomenon. We used the same time sampling than for the ODE solvers.

We found that RK significantly improved the convergence of Zhang EPDiff-LDDMM with respect to Euler integration, that was the integration method originally proposed in (Zhang & Fletcher 2015).

4.6. Software implementation

The experiments were run on a cluster equipped with four NVidia GeForce GTX Titan X with 12 GBS of video memory installed on a machine with an Intel Core i7-4820K working at 3.70 GHz and 64 GBS of DDR3 RAM. We developed a hybrid Matlab 2015b CPU-GPU implementation of the methods using Cuda 7.0. Stationary PDE-LDDMM and Zhang and Vialard EPDiff-LDDMM methods were computed in the GPU. However, Non-stationary PDE-LDDMM, PDE-EPDiff LDDMM, and Jacobi PDE-EPDiff LDDMM methods did not fit on the available GPU memory. We circumvented this problem implementing a hybrid CPU-GPU version of the codes. Our implementation performed the computations of the gradient and the Hessian in the GPU. These computations were gathered in the RAM memory, and then the corresponding CG or PCG iteration was computed on the CPU. We could not improve the design of our implementation due to the poor behavior of Matlab when releasing big chunks of GPU memory (for example, to clear one large variable from the GPU, one should reset the gpuDevice losing all the information uploaded to the global GPU memory, otherwise, the memory remains busy with the deallocated resource).

5. Results

The results section is divided in two. In the first part, we perform a first approximation to the evaluation of the proposed method with the 2D monkey brain images of (Beg et al. 2005). The special shape of the images to be registered allows showing interesting insights of diffeomorphic registration algorithms. In the second part, we perform a quantitative and qualitative evaluation of the proposed method with the 3D NIREP database.

5.1. PDE-EPDiff LDDMM vs Jacobi PDE-EPDiff LDDMM

The purpose of this section is to compare PDE-EPDiff LDDMM with our proposed Jacobi PDE-EPDiff LDDMM. We recall with PDE-EPDiff LDDMM to the natural formulation and derivation of the PDE-LDDMM problem (Mang & Biros 2015) in the space of initial velocity fields. The difference between both methods is on the computation of the adjoint and the incremental adjoint equations leading to $w(0)$ and $\delta w(0)$.

The experiments have been performed in the 2D monkey brain experiment of (Beg et al. 2005) (Figure 1). The images were acquired by Dr. David Van Essen at Washington University. They can be downloaded from the Center for Imaging Science site http://cis.jhu.edu/software/lddmm-volume/about.php These images are very appropriate for the initial assessment of diffeomorphic registration methods.

The comparison between PDE-EPDiff LDDMM and Jacobi PDE-EPDiff LDDMM has been performed for different sizes of temporal domain discretization ($n_t = 25, 30, 50, 100$) and the maximum number of CG or PCG iterations ($n_{PCG} = 5, 10, 15, 25$). We know that the convergence of PDE-constrained LDDMM methods is sensitive to the selection of both parameters. Optimization has been performed in V and L^2.

Figure 3 shows the image similarity error curves, $MSE_{rel} = \|m(1) - I_1\|_{L^2}^2 / \|I_0 - I_1\|_{L^2}^2$, for the different values of n_t and n_{PCG}. For optimization in V, both methods showed a similar convergence behavior. Our proposed method converged to slightly better energy values (closer to zero). For optimization in L^2, the methods showed a much worse convergence behavior. Our proposed method was able to converge to acceptable energy values for $n_{PCG} = 5$ and all the n_t values. PDE-EPDiff LDDMM was not able to converge to any acceptable energy value. The solutions blowed up for $n_{PCG} = 10$.

Figure 2 shows the warped sources for our Jacobi PDE-EPDiff LDDMM and PDE-EPDiff LDDMM at $n_t = 25$. For optimization in V, PDE-EPDiff LDDMM shows an unrealistic behavior at the brain tip shown in the target image while our method shows a plausible result. For optimization in L^2, our method showed artifacts in the solution except at $n_{PCG} = 5$. PDE-EPDiff LDDMM showed artifacts and low-quality registration results in all cases.

These experiments reveal that for PDE-EPDiff constrained LDDMM methods optimization should be performed in V rather than in L^2. PDE-EPDiff LDDMM is more unstable than our Jacobi PDE-EPDiff LDDMM. Our method is able to show plausible results in challenging locations for diffeomorphic registration. Finally, PDE-EPDiff constrained LDDMM method shows a good convergence behavior for small n_{PCG} values.
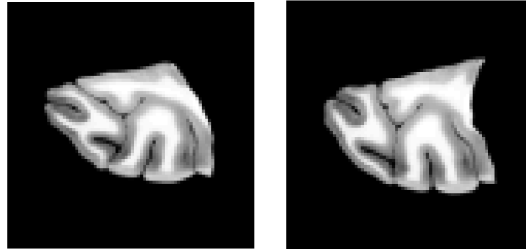
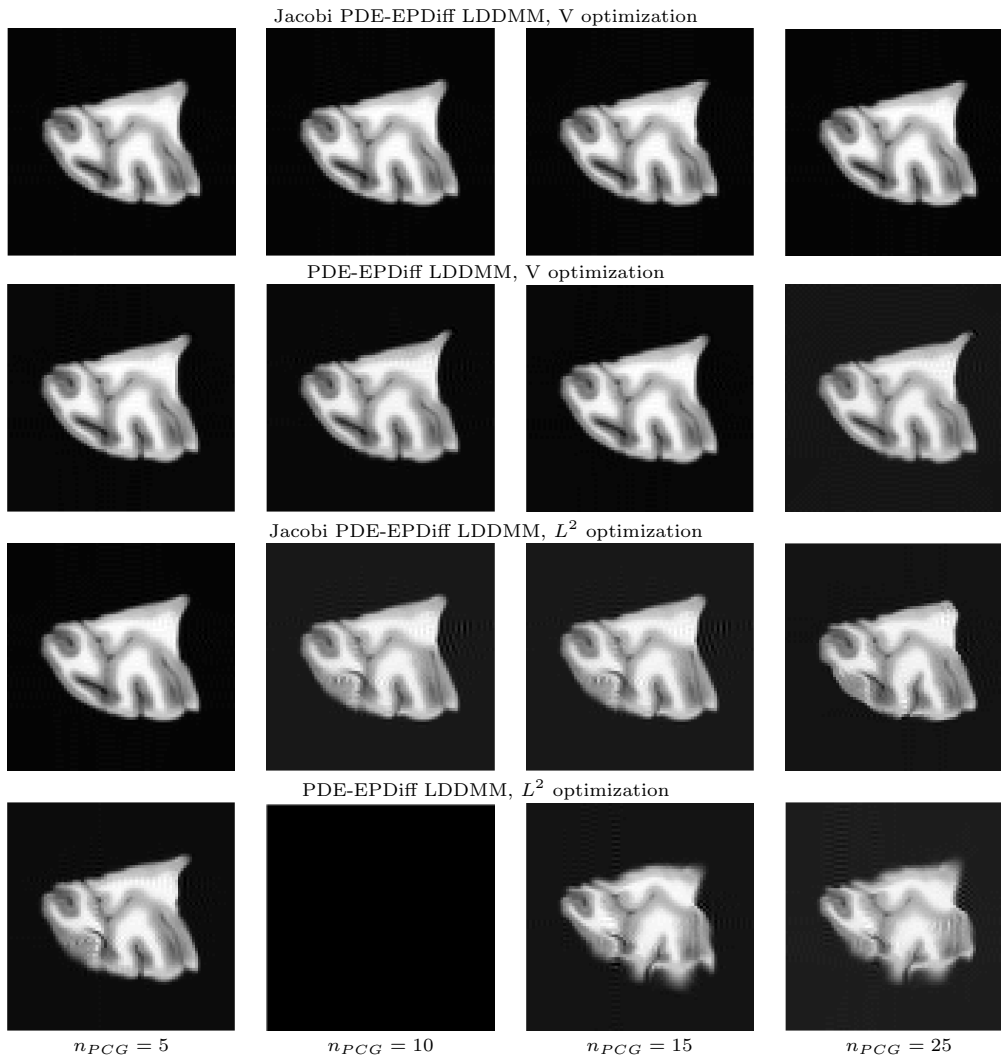Figure 1: Source and target images of the 2D monkey brain experiment.

Jacobi PDE-EPDiff LDDMM, V optimization

PDE-EPDiff LDDMM, V optimization

Jacobi PDE-EPDiff LDDMM, L^2 optimization

PDE-EPDiff LDDMM, L^2 optimization

$n_{PCG} = 5$ $\qquad$ $n_{PCG} = 10$ $\qquad$ $n_{PCG} = 15$ $\qquad$ $n_{PCG} = 25$

Figure 2: Warped sources for our Jacobi PDE-EPDiff LDDMM and PDE-EPDiff LDDMM at $n_t = 25$.
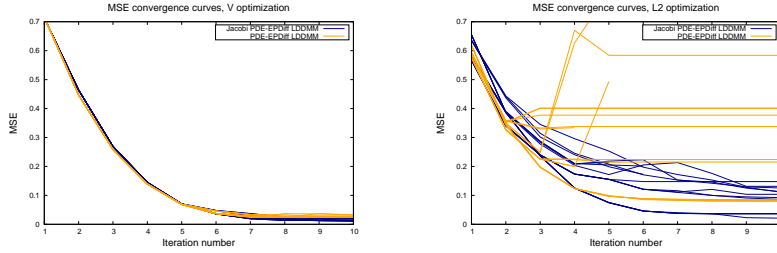
Figure 3: Image similarity error curves for PDE-EPDiff and our Jacobi PDE-EPDiff LDDMM for different temporal domain discretization sizes and maximum number of CG or PCG iterations. Left, results with V optimization. Right, results with L^2 optimization.

5.2. Evaluation of Jacobi PDE-EPDiff LDDMM

Next in this section, we evaluate the performance of our proposed method. As a baseline for the evaluation, we include the results obtained by the methods most related to our work: PDE-LDDMM with the stationary and the non-stationary parameterizations (Mang & Biros 2015), Zhang EPDiff-LDDMM in the spatial domain (Zhang & Fletcher 2018), and Vialard EPDiff-LDDMM (Vialard et al. 2011).

The experiments have been conducted on the Non-rigid Image Registration Evaluation Project database (NIREP). In this work, images were resampled into volumes of size $180 \times 210 \times 180$. Registration was carried out from the first subject to every other subject in the database, yielding 15 registrations for each method. The preprocessing details for this dataset can be found in our previous works (Hernandez 2014, Hernandez 2017).

5.2.1. Convergence analysis. Table 1 shows, averaged by the number of experiments, the relative image similarity error $MSE_{rel} = \|m(1) - I_1\|_{L^2}^2 / \|I_0 - I_1\|_{L^2}^2$, the relative gradient magnitude $\|g\|_{\infty,rel} = \frac{\|\nabla_v E(v^n)\|_\infty}{\|\nabla_v E(v^0)\|_\infty}$, the extrema of the Jacobian determinant, and the total number CG, PCG, or GD iterations. St. PDE-LDDMM showed convergence problems for the $\|g\|_{\infty,rel}$ curve in one experiment and NSt. PDE-LDDMM and PDE-EPDiff LDDMM in two experiments. These values were removed in the computation of the mean and the standard deviation. Figure 4 shows the mean and standard deviation of the MSE_{rel} and $\|g\|_{\infty,rel}$ convergence curves obtained in the ten GN iterations and the fifty GD iterations.

The lowest mean MSE_{rel} value is reached by St. PDE-LDDMM (19.18) followed by our proposed method (20.24) and NSt. PDE-LDDMM (21.11). Our proposed method outperforms PDE-EPDiff LDDMM in terms of MSE_{rel}. The worst performing method is Vialard EPDiff LDDMM. Zhang EPDiff LDDMM achieves modest MSE_{rel} values after 50 iterations. From the $\|g\|_{\infty,rel}$ values, the most remarkable result is the convergence problem shown by PDE-LDDMM and PDE-EPDiff LDDMM in several experiments while our method shows an acceptable convergence behavior in all the performed experiments even with L^2 optimization.

5.2.2. Evaluation. Figure 5 shows the accuracy of the registration results for template-based segmentation. We use the manual segmentations of the cortical

Method	Opt.	MSE_{rel}	$\|g\|_{\infty,rel}$	$\max(J(\phi_1^v)^{-1})$	$\min(J(\phi_1^v)^{-1})$	n_{its}
St. PDE-LDDMM (Mang)	GN, L^2	19.18 ± 3.34	0.08 ± 0.05	3.71 ± 0.56	0.15 ± 0.04	41.00 ± 13.02
NSt. PDE-LDDMM (Mang)	GN, L^2	21.11 ± 5.18	0.20 ± 0.11	3.63 ± 0.83	0.14 ± 0.04	25.66 ± 14.86
EPDiff LDDMM (Vialard)	GD, V	37.91 ± 2.40	0.29 ± 0.05	1.63 ± 0.15	0.27 ± 0.23	50.00 ± 0.00
EPDiff LDDMM (Zhang)	GD, V	26.32 ± 2.23	0.07 ± 0.02	2.00 ± 0.17	0.28 ± 0.04	50.00 ± 0.00
PDE-EPDiff LDDMM	GN, V	22.04 ± 5.81	0.16 ± 0.17	2.92 ± 0.69	0.13 ± 0.06	25.13 ± 6.60
Jacobi PDE-EPDiff LDDMM	GN, V	20.24 ± 2.05	0.12 ± 0.08	3.30 ± 1.12	0.11 ± 0.04	24.53 ± 4.14
Jacboi PDE-EPDiff LDDMM	GN, L^2	21.69 ± 2.87	0.28 ± 0.11	3.52 ± 1.14	0.11 ± 0.05	16.87 ± 3.50

Table 1: Mean and standard deviation of the relative image similarity error, the relative gradient magnitude, the Jacobian determinant extrema associated with the transformation $(\phi_1^v)^{-1}$, and the total number of iterations (namely, the number of PCG/CG iterations times the total number of outer iterations for Krylov methods, and the total number of iterations for gradient-descent methods). In Opt. column, GN and GD stand for Gauss-Newton and gradient descent optimization, respectively.
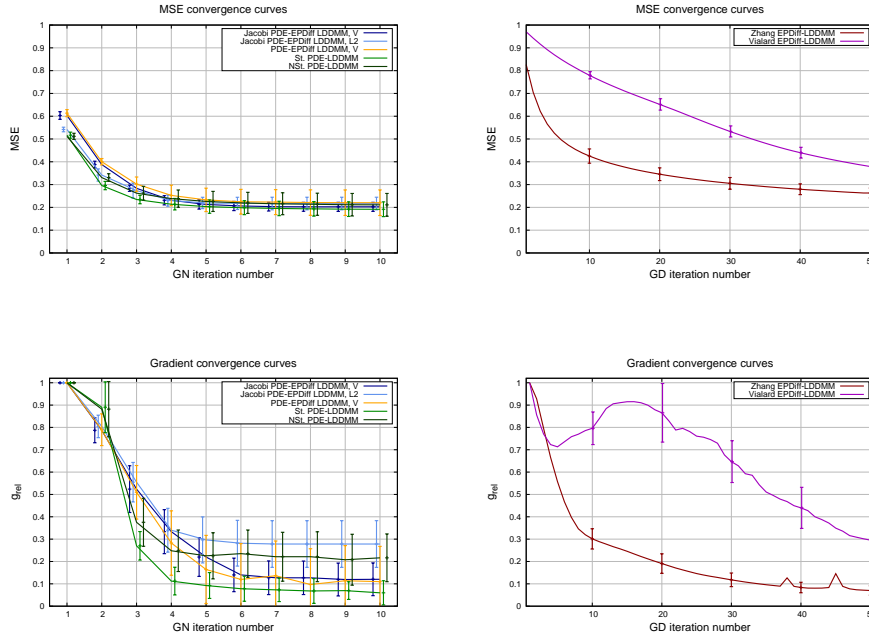

Figure 4: Top, overall MSE_{rel} convergence curves for Gauss-Newton and gradient descent methods, respectively. Bottom, overall $\|g\|_{\infty,rel}$ convergence curves. The error bars have been shifted from the GN iteration number to improve the readability of the plot. It should be noticed that the PDE-LDDMM and PDE-EPDiff LDDMM $\|g\|_{\infty,rel}$ curves show the results after removing the experiments with gradient curves not converging to zero.

anatomical structures provided with the NIREP database as a gold standard and Dice Similarity Coefficient (DSC) is selected as the performance metric similarly to (Ou et al. 2014). In addition, Figure 6 shows the DSC distribution in each of the brain structures manually segmented in the NIREP database. The evaluation of non-rigid registration with the segmentation of sufficiently locally labeled regions of interest

Method	VRAM (MBS)	1 iter time (s)	total time (s)
St. PDE-LDDMM (Mang)	1543	58.61 ± 0.55	612.12
NSt. PDE-LDDMM (Mang)	3278	370.51 ± 7.38	3866.3
EPDiff LDDMM (Vialard)	4410	18.79 ± 6.19	940.00
EPDiff LDDMM (Zhang)	1632	16.03 ± 0.34	806.81
PDE-EPDiff LDDMM	2775	128.37 ± 10.70	1303.50
Jacobi PDE-EPDiff LDDMM	2192	180.70 ± 5.71	1825.50

Table 2: Computational complexity at the resolution level of 90 x 120 x 90.

has been recommended for obtaining reliable measurements of the performance of the registration (Rohlfing 2012).

The evaluation reveals that all the PDE-EPDiff methods performed similarly for the task of template-based segmentation with the exception of Vialard PDE-EPDiff method. PDE-LDDMM methods slightly outperformed the PDE-EPDiff methods. This may be due to the EPDiff-constraint slightly limits the ability of the PDE-EPDiff methods to further fold out the transformations in the cortical structures for the difficult task of inter-subject registration. However, the added value of our proposed method lies in the ability to provide geodesics, of importance in Computational Anatomy applications beyond the task of template-based segmentation.

5.2.3. Qualitative assessment. For a qualitative assessment of the proposed registration method, we show the registration results in a selected experiment representative of a difficult deformable registration problem. Figures 7 and 8 show the warped images and the initial velocity fields obtained in the selected experiment for all the methods considered in this experimental section.

5.2.4. Computational complexity. Finally, Table 2 shows the computational complexity of the methods at the resolution level of 90 x 120 x 90. The image size was selected to perform a fair comparison of the full GPU version of the algorithms. Our proposed method was more memory demanding than PDE-LDDMM methods and Zhang EPDiff-LDDMM but less memory demanding than PDE-EPDiff LDDMM and Vialard EPDiff-LDDMM. The computation time was among the highest.

6. Discussion and Conclusions

In this work, we have proposed a method for PDE-constrained LDDMM parameterized in the space of initial velocity fields. Our method can be regarded as an extension of the method in (Mang & Biros 2015) that allows obtaining geodesics. Rather than computing the gradient and the Hessian-vector products from the differentiation of the augmented energy functional on the initial velocity field (PDE-EPDiff LDDMM), we propose to perform the derivations of the gradient and Hessian-vector products on the final velocity field and transport the computations backward using the adjoint and the incremental adjoint Jacobi equations (Jacobi PDE-EPDiff LDDMM). This way we avoid the computation of the adjoint equation and its incremental counterpart that has been recently identified as a subtle problem in (Mang & Ruthotto 2017) and corroborated by the experiments of this work.

The comparison of our Jacobi PDE-EPDiff method with PDE-EPDiff LDDMM in the monkey brain dataset reported that our method was more stable. We found

that both methods were sensitive to the space where the optimization was performed, contrarily to PDE-LDDMM methods. We believe that CG optimization in V should be used with Newton-Krylov PDE-EPDiff methods. In addition, the experiments reported that a small number of PCG iterations should be used with PDE-EPDiff LDDMM methods to yield stable results while our Jacobi PDE-EPDiff method was stable regardless of the number of PCG iterations.

Our method has been evaluated with respect to benchmark PDE-constrained LDDMM and EPDiff-LDDMM methods in the NIREP database. Gauss-Newton-Krylov optimization provided a higher rate of convergence than gradient-descent as expected. While PDE-LDDMM and PDE-EPDiff LDDMM showed problems with the gradient energy convergence in a few cases, our method showed an acceptable convergence in all cases. The evaluation in the NIREP database showed competitive performance with respect to the EPDiff-LDDMM benchmark methods. The slight differences between PDE-LDDMM and the EPDiff methods are due to the more restrictive model imposed by the EPDiff equation on the velocity fields.

The major drawback of the proposed method is the large memory load inherent to PDE-constrained LDDMM methods and the increased computational time with respect to the benchmark methods. We will approach this problem in future work using the band-limited vector field parameterization (Zhang & Fletcher 2018, Hernandez 2018*a*, Hernandez Jul, 2018*b*). In addition, we will extend this band-limited PDE-LDDMM formulation to more complex physical models such as the incompressible and the nearly incompressible models proposed in (Mang & Biros 2015, Mang & Biros 2016). We also will extend our formulation to a symmetric formulation (Younes 2007, Avants et al. 2008, Hernandez 2014, Fleishman et al. 2017), a feature very desirable in applications involving reversible motions. Another future research direction will be to extend our formulation to different image similarity metrics such as normalized cross-correlation, mutual information, or normalized gradient fields (Modersitzki 2009) and more sophisticated metrics such as (Heinrich et al. 2012). These extensions will make our formulation suitable for other clinical applications such as cardiac, lung or abdominal registration, among others.

Acknowledgements

Appendix A. Derivation of PDE-LDDMM in the space of initial velocity fields (PDE-EPDiff LDDMM)

In this appendix, we show the equations of the gradient and the Hessian-vector product of PDE-EPDiff LDDMM method. The Hessian provided in (Yang & Niethammer 2015) in the context of registration uncertainty estimation is obtained

using a similar derivation. The variational problem is given by

$$E(v_0) = \langle Lv_0, v_0 \rangle_{L^2} + \frac{1}{\sigma^2} \|m(1) - I_1\|_{L^2}^2, \quad \text{(A.1)}$$

subject to

$$\partial_t v_t + ad_{v_t}^\dagger v_t = 0 \text{ in } \Omega \times (0, 1] \quad \text{(A.2)}$$
$$\partial_t m(t) + \nabla m(t) \cdot v_t = 0 \text{ in } \Omega \times (0, 1], \quad \text{(A.3)}$$

with initial conditions $v(0) = v_0$ and $m(0) = I_0$ in Ω.

Following (Hart et al. 2009, Mang & Biros 2015), optimization is performed combining the method of Lagrange multipliers with Newton-Krylov optimization. Let $w : \Omega \times [0, 1] \to \mathbb{R}^d$ and $\lambda : \Omega \times [0, 1] \to \mathbb{R}$ be the Lagrange multipliers associated with the EPDiff and the state equations. We build the augmented Lagrangian

$$E_{aug}(v_0) = E(v_0) + \int_0^1 \langle w(t), \partial_t v(t) + ad_{v_t}^\dagger v_t \rangle_{L^2} dt$$

$$+ \int_0^1 \langle \lambda(t), \partial_t m(t) + \nabla m(t) \cdot v_t \rangle_{L^2} dt. \quad \text{(A.4)}$$

The gradient is computed from

$$\nabla_{v_0} E(v_0) = 2v_0 - w(0), \quad \text{(A.5)}$$

where the equations leading to $w(0)$

$$\partial_t v_t + ad_{v_t}^\dagger v_t = 0 \text{ in } \Omega \times (0, 1] \quad \text{(A.6)}$$
$$\partial_t m(t) + \nabla m(t) \cdot v_t = 0 \text{ in } \Omega \times (0, 1] \quad \text{(A.7)}$$
$$-\partial_t \lambda(t) - \nabla \cdot (v\lambda(t)) = 0 \text{ in } \Omega \times [0, 1) \quad \text{(A.8)}$$
$$-\partial_t w_t + ad_{v_t} w_t - ad_{w_t}^\dagger v_t + K(\lambda(t)\nabla m(t)) = 0 \text{ in } \Omega \times [0, 1) \quad \text{(A.9)}$$

are computed from the first variation of the augmented Lagrangian.

The Hessian-vector product is computed from

$$H_{v_0} E(v_0)\delta v_0 = 2\delta v_0 - \delta w(0), \quad \text{(A.10)}$$

where the equations leading to $\delta w(0)$

$$\partial_t \delta v_t + ad_{\delta v_t}^\dagger v_t + ad_{v_t}^\dagger \delta v_t = 0 \text{ in } \Omega \times (0, 1] \quad \text{(A.11)}$$
$$\partial_t \delta m(t) + \nabla \delta m(t) \cdot v_t + \nabla m(t) \cdot \delta v_t = 0 \text{ in } \Omega \times (0, 1] \quad \text{(A.12)}$$
$$-\partial_t \delta \lambda(t) - \nabla \cdot (v\delta\lambda(t)) - \nabla \cdot (\delta v\lambda(t)) = 0 \text{ in } \Omega \times [0, 1) \quad \text{(A.13)}$$

$$-\partial_t \delta w_t + ad_{\delta v_t} w_t + ad_{v_t} \delta w_t - ad_{\delta w_t}^\dagger v_t - ad_{w_t}^\dagger \delta v_t +$$
$$K(\delta\lambda(t)\nabla m(t)) + K(\lambda(t)\nabla \delta m(t)) = 0 \text{ in } \Omega \times [0, 1) \quad \text{(A.14)}$$

are computed from the second variation of the augmented Lagrangian.

The Gauss-Newton approximation drops $K(\lambda(t) \cdot \nabla \delta m(t))$ and the ad_v and ad_w terms from the incremental equation of δw_t. It should be noticed that this method

solves the adjoint and incremental adjoint equations for λ and $\delta\lambda$ and computes the body force $K(\lambda(t)\nabla m(t))$ and the incremental counterpart. In contrast, Zhang EPDiff-LDDMM and Jacobi EPDiff-LDDMM approach the computation of the body force equivalent by solving the adjoint Jacobi and the adjoint and incremental adjoint Jacobi equations, respectively.
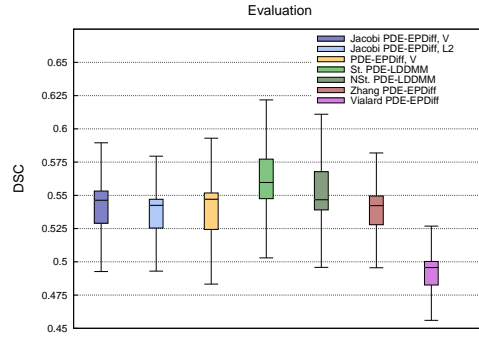
Figure 5: Evaluation results. Volume overlap obtained by the registration methods measured in terms of the DSC between the deformed and the corresponding manual segmentations. Box and whisker plots represent the distribution of the DSC values averaged over the 32 regions of interest.
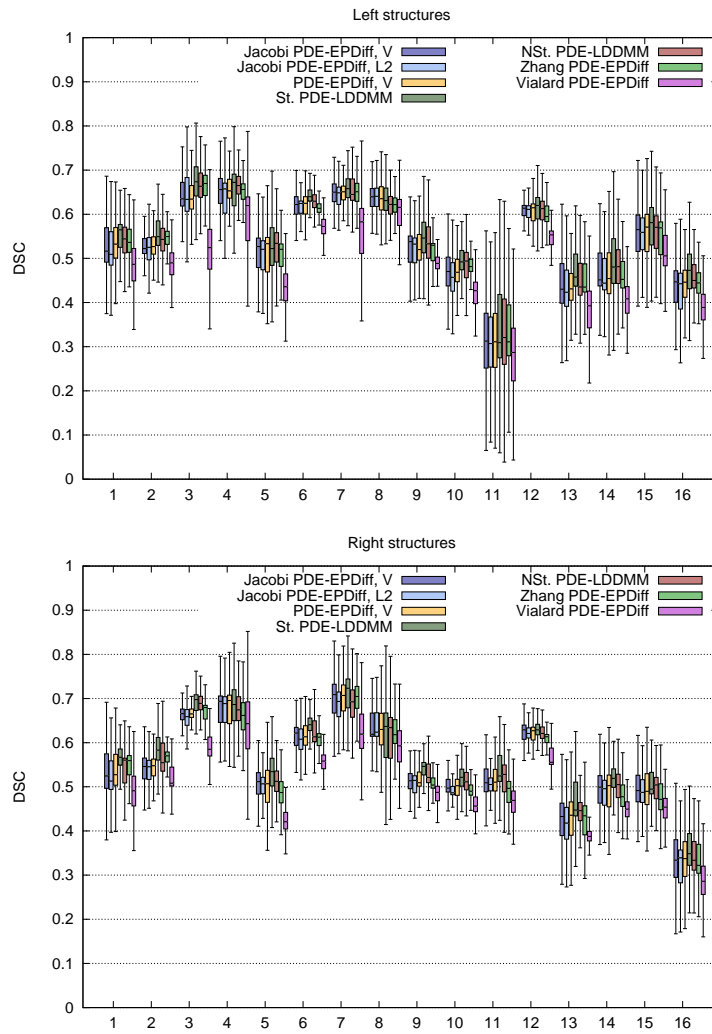

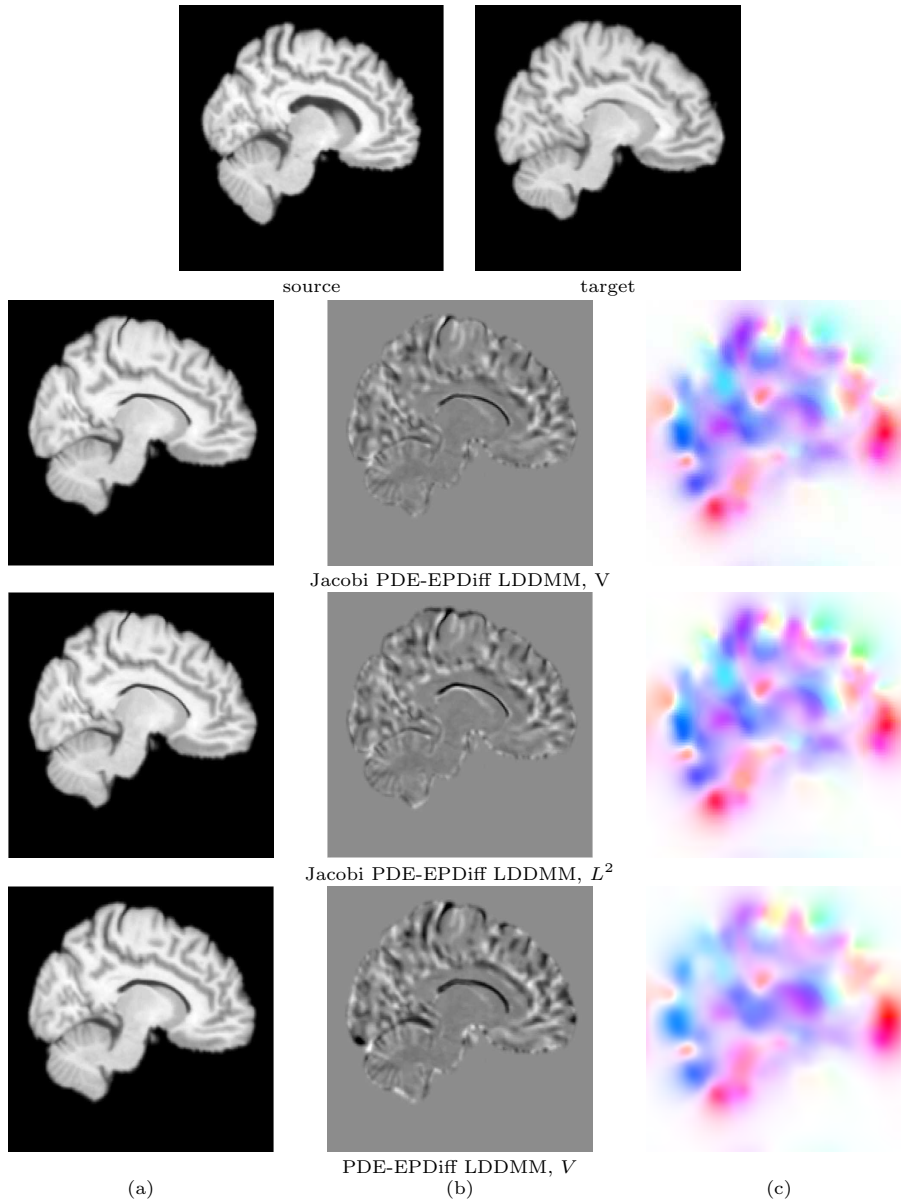Figure 6: Evaluation results. Volume overlap in the 32 regions of interest.

Figure 7: Image registration results. Sagittal view of (a) the warped sources, (b) the intensity differences after registration, and the (c) initial velocity fields for our proposed method and PDE-EPDiff LDDMM.
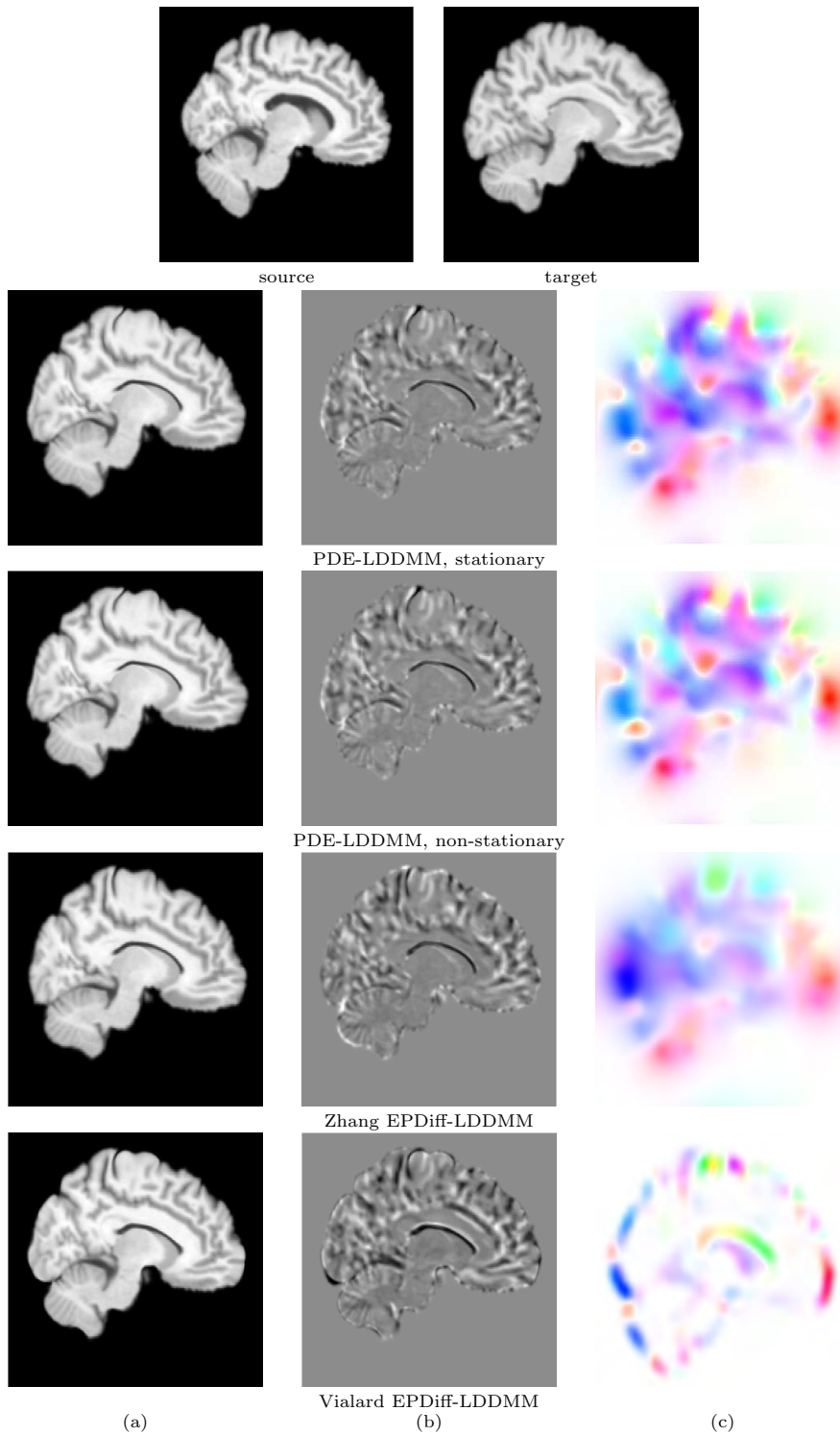
Figure 8: Image registration results. Sagittal view of (a) the warped sources, (b) the intensity differences after registration, and (c) the initial velocity fields for the benchmark methods considered in the comparison.

References

Ashburner J & Friston K J 2011 Diffeomorphic registration using geodesic shooting and Gauss–Newton optimisation *Neuroimage* **55(3)**, 954 – 967.

Avants B B, Epstein C L, Grossman M & Gee J C 2008 Symmetric dieomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain *Med. Image Anal.* **12**, 26 – 41.

Beg M F, Miller M I, Trouve A & Younes L 2005 Computing large deformation metric mappings via geodesic flows of diffeomorphisms *Int. J. Comput. Vision* **61 (2)**, 139–157.

Bullo F 1995 Invariant affine connections and controllability on Lie groups *Techical report for Geometric Mechanics, California Institute of Technology* .

Fleishman G M, Fletcher P T & Thompson P M 2017 Symmetric interleaved geodesic shooting in diffeomorphisms *Proc. of International Conference on Information Processing and Medical Imaging (IPMI'17), Lecture Notes in Computer Science* .

Hart G L, Zach C & Niethammer M 2009 An optimal control approach for deformable registration *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'09)* .

Heinrich M P, Jenkinson M, Bhushan M, Matin T, Gleeson F V, Brady M & Schnabel J A 2012 Mind: Modality independent neighborhood descriptor for multi-modal deformable registration *Med. Image Anal.* **16(7)**, 1423 – 1435.

Hernandez M 2014 Gauss-Newton inspired preconditioned optimization in large deformation diffeomorphic metric mapping *Phys. in Med. and Biol.* **59(20)**.

Hernandez M 2017 Primal-Dual convex optimization in Large Deformation Diffeomorphic Metric Mapping: LDDMM meets robust regularizers *Phys. in Med. and Biol.* .

Hernandez M 2018*a* Band-limited stokes large deformation diffeomorphic metric mapping *IEEE J. of Biom. and Health Inf.* .

Hernandez M Jul, 2018*b* Efficient Gauss-Newton-Krylov momentum conservation constrained PDE-LDDMM using the band-limited vector field parameterization *https://arxiv.org/abs/1807.11560* .

Hinkle J, fletcher P T & Joshi S 2014 Intrinsic polynomials for regression on riemannian manifolds *J. Math. Imaging Vis.* **50(1-2)**, 32 – 52.

Holm D D, Marsden J E & Ratiu T S 1998 The Euler-Poincare equations and semidirect products with applications to continuum theories *Adv. in Math.* **137**, 1 – 81.

Holm D D, Ratnanather J T, Trouve A & Younes L 2004 Soliton dynamics in computational anatomy *Neuroimage* **23**, 170 – 178.

Kobatake H & Masutani Y 2017 *Computational anatomy based on whole body imaging* Springer Germany.

Mang A & Biros G 2015 An inexact Newton-Krylov algorithm for constrained diffeomorphic image registration *SIAM J. Imaging Sciences* **8(2)**, 1030–1069.

Mang A & Biros G 2016 Constrained H1 regularization schemes for diffeomorphic image registration *SIAM J. Imaging Sciences* .

Mang A & Ruthotto L 2017 A lagrangian GaussNewtonKrylov solver for mass- and intensity-preserving diffeomorphic image registration *SIAM J. Sci. Comput.* **39(5)**, B860 – B885.

Miller M I, Trouve A & Younes L 2005 Geodesic shooting for Computational Anatomy *J. Math. Imaging Vis.* .

Modersitzki J 2009 FAIR: Flexible Algorithms for Image Registration *SIAM* .

Ou Y, Akbari H, Bilello M, Da X & Davatzikos C 2014 Comparative evaluation of registration algorithms in different brain databases with varying difficulty: results and insights *IEEE Trans. Med. Imaging* **33(10)**, 2039 – 2065.

Rohlfing T 2012 Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable *IEEE Trans. Med. Imaging* **31(2)**, 153 – 163.

Song J H, Christensen G E, Hawley J A, Wei Y & Kuhl J G 2010 Evaluating image rgistration using NIREP *WBIR 2010, LNCS* **6204**, 140 – 150.

Sotiras A, Davatzikos C & Paragios N 2013 Deformable medical image registration: A survey *IEEE Trans. Med. Imaging* **32(7)**, 1153 – 1190.

Trefethen L N 2000 Spectral methods in matlab *SIAM, Philadelphia* .

Vialard F X, Risser L, Rueckert D & Cotter C J 2011 Diffeomorphic 3D image registration via geodesic shooting using an efficient adjoint calculation *Int. J. Comput. Vision* **97(2)**, 229 – 241.

Yang X & Niethammer M 2015 Uncertainty quantification for LDDMM using a low-rank hessian approximation *Proc. of the 18th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'15), Lecture Notes in Computer Science*

9350, 289 – 296.

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

Zhang M & Fletcher P T 2015 Finite-dimensional Lie algebras for fast diffeomorphic image registration *Proc. of International Conference on Information Processing and Medical Imaging (IPMI'15), Lecture Notes in Computer Science* .

Zhang M & Fletcher T 2018 Fast diffeomorphic image registration via Fourier-Approximated Lie algebras *Int. J. Comput. Vision* .

Zhang M, Liao R, Dalca A V, Truk E A, Luo J, Grant P E & Golland P 2017 Frequency diffeomorphisms for efficient image registration *Proc. of International Conference on Information Processing and Medical Imaging (IPMI'17), Lecture Notes in Computer Science* pp. 559–570.