

# Uncertainty Quantification for LDDMM Using a Low-rank Hessian Approximation

Xiao Yang\*, Marc Niethammer\*,†

\*Department of Computer Science and †Biomedical Research Imaging Center  
University of North Carolina at Chapel Hill, USA

**Abstract.** This paper presents an approach to estimate the uncertainty of registration parameters for the large displacement diffeomorphic metric mapping (LDDMM) registration framework. Assuming a local multivariate Gaussian distribution as an approximation for the registration energy at the optimal registration parameters, we propose a method to approximate the covariance matrix as the inverse of the Hessian of the registration energy to quantify registration uncertainty. In particular, we make use of a low-rank approximation to the Hessian to accurately and efficiently estimate the covariance matrix using few eigenvalues and eigenvectors. We evaluate the uncertainty of the LDDMM registration results for both synthetic and real imaging data.

## 1 Introduction

Image registration is critical for many medical image analysis systems to provide spatial correspondences. Consequentially, a large number of image registration methods have been developed which are able to produce high-quality spatial alignments of images. However, most image registration approaches do not provide any measures of registration uncertainty and hence do not allow a user to assess if a registration result is locally “trustworthy” or not. This is particularly problematic for highly flexible registration approaches, such as elastic or fluid registration with very large numbers of parameters to model deformations.

Different approaches to address uncertainty quantification in image registration have been proposed. For example, for rigid deformations, physical landmarks have been used to estimate the average registration error for the whole volume. [2]. Non-rigid deformations are challenging as landmarks only capture local aspects of the deformations. Instead, methods have been proposed to assess uncertainty based on probabilistic models of registration and the image itself. For B-spline models, sampling based methods have been proposed for images [5] or the optimal spline parameters [11] to create multiple registrations from which to estimate deformation uncertainty. Simpson [9] uses a variational Bayesian approach to infer the posterior distribution of B-spline parameters. Monte-Carlo sampling methods have also been explored in the context of elastic registration [8] and for LDDMM [12].

Existing methods mainly focus on parametric non-rigid registration methods such as the B-spline model or require large computational effort to sample over

high-dimensional parameter spaces [8,12] as for LDDMM [1]. Here, we develop a method to estimate uncertainty in the deformation parameters of the shooting formulation [10] of LDDMM.

We assume a local multivariate Gaussian distribution at the optimal solution, and approximate the covariance matrix through the inverse of the approximated energy Hessian to quantify uncertainty of the registration parameters. Using the Hessian of the energy to estimate the covariance matrix of parameters has been discussed for large-scale inverse problems in other application domains [3,4]. For high dimensional parameter spaces, computing the full Hessian is prohibitive due to large memory requirements. Therefore, we develop a method to compute Hessian vector products for the LDDMM energy. This allows us to efficiently compute and store an approximation of the Hessian. In particular, we directly approximate the covariance matrix by exploiting the low-rank structure of the image mismatch Hessian. Our framework therefore allows uncertainty analysis for LDDMM at a manageable computational cost.

Sec. 2 discusses the relationship of the covariance matrix of the parameter distribution and the Hessian of the energy function. Sec. 3 introduces our framework to compute the Hessian and to estimate the covariance matrix. Sec. 4 shows experimental results for both synthetic and real data, and discusses how our method extends to other registration models and its potential applications.

## 2 Covariance Matrix and Hessian of the Energy

Consider a Gaussian random vector  $\theta$  of dimension  $N_\theta$  with mean value  $\theta^*$  and covariance matrix  $\Sigma_\theta$ . Its joint probability density function can be written as

$$P(\theta) = (2\pi)^{-\frac{N_\theta}{2}} |\Sigma_\theta|^{-\frac{1}{2}} \exp \left[ -\frac{1}{2} (\theta - \theta^*)^T \Sigma_\theta^{-1} (\theta - \theta^*) \right]. \quad (1)$$

In image registration one typically minimizes the energy given by the negative log-likelihood of the posterior distribution. As the negative log-likelihood of the multivariate Gaussian is given by

$$E(\theta) = -\ln(P(\theta)) = \frac{N_\theta}{2} \ln 2\pi + \frac{1}{2} \ln |\Sigma_\theta| + \frac{1}{2} (\theta - \theta^*)^T \Sigma_\theta^{-1} (\theta - \theta^*), \quad (2)$$

computing the Hessian of  $E(\theta)$  with respect to  $\theta$  results in  $H_{E(\theta)} = \Sigma_\theta^{-1}$  and directly relates the covariance matrix of the multivariate Gaussian model to the Hessian of the energy through its inverse.

In our method, we assume that at optimality the LDDMM energy can be locally approximated by a second order function and hence by a multivariate Gaussian distribution. In particular, we make use of the shooting based formulation of LDDMM [10] which parameterizes the spatial deformation by an initial momentum (or equivalently by an initial velocity field) and associated evolution equations describing the space deformation over time. Specifically, the registration parameter in shooting based LDDMM is the initial momentum  $m$ , which is

the dual of the initial velocity  $v$ , an element in a reproducing kernel Hilbert space  $V$ . The initial momentum belongs to  $V$ 's dual space  $V^*$ , and it is connected with  $v$  via a positive-definite, self-adjoint differential operator  $L : V \rightarrow V^*$  such that  $m = Lv$  and  $v = Km$ . Here the operator  $K$  denotes the inverse of  $L$ . The energy of LDDMM with the dynamic constraints [10] can then be written as

$$E(m_0) = \langle m_0, Km_0 \rangle + \frac{1}{\sigma^2} \|I(1) - I_1\|^2, \quad (3)$$

$$m_t + \text{ad}_v^* m = 0, \quad m(0) = m_0, \quad I_t + \nabla I^T v = 0, \quad I(0) = I_0, \quad m - Lv = 0, \quad (4)$$

where the operator  $\text{ad}_v^*$  is the dual of the negative Jacobi-Lie bracket of vector fields:  $\text{ad}_v w = -[v, w] = Dvw - Dv w$ ,  $\sigma > 0$ , and  $I_0$  and  $I_1$  are the source and the target images for registration respectively. The Hessian of this energy is

$$H_{m_0} = 2K + \frac{\partial^2 \frac{1}{\sigma^2} \|I(1) - I_1\|^2}{\partial m_0^2}. \quad (5)$$

Computing this Hessian is not straightforward, because  $I(1)$  only indirectly depends on  $m_0$  through the dynamic constraints and  $m_0$  can become very high-dimensional, making computation and storage challenging. Sec. 3 therefore discusses how to compute the Hessian and its inverse in practice.

### 3 Hessian Calculation and Covariance Estimation

#### 3.1 Calculating Hessian-vector products using the adjoint method

To avoid computation of the full Hessian we instead compute Hessian-vector products. This enables us to make use of efficient iterative methods (such as the Lanczos method) to perform eigen-decomposition of the Hessian, which we exploit to compute an approximation of the Hessian and the covariance matrix.

The equivalent of Hessian-vector products for LDDMM can be computed using the second variation of the LDDMM energy. Specifically, the second variation in the direction  $\delta m_0$  can be written as

$$\delta^2 E(m_0; \delta m_0) := \frac{\partial^2}{\partial \epsilon^2} E(m_0 + \epsilon \delta m_0)|_{\epsilon=0} = \langle \delta m_0, \nabla^2 E \delta m_0 \rangle. \quad (6)$$

Here  $\nabla^2 E$  denotes the Hessian of  $E(m_0)$ . Using this formulation, we can read off the Hessian-vector product  $\nabla^2 E \delta m_0$  from the second variation. Computing this second variation of the LDDMM shooting energy can be accomplished by linearizing both the forward equations for shooting as well as the associated adjoint equations around the optimal solution (the solution of the registration problem). Solving these equations for a given initial condition  $\delta m_0$  then allows the computation of the Hessian vector product (in a functional sense) as

$$\nabla^2 E \delta m_0 = 2K \delta m_0 - \delta \hat{m}(0). \quad (7)$$

Here,  $2K \delta m_0$  can be computed directly and  $\delta \hat{m}(0)$  is the perturbation of the adjoint of the momentum propagation constraint at  $t = 0$ , which is obtained efficiently through a forward-backward sweep through the linearized forward and adjoint equations. Please refer to the supplementary material for details.

### 3.2 Covariance estimation using low-rank Hessian

To estimate the covariance, a straightforward way is inverting the Hessian of the full energy. This is not feasible for standard LDDMM because the number of parameters is so large that saving or computing the inverse of the full Hessian is prohibitive<sup>1</sup>. Another possibility is to approximate the full energy Hessian. Note that the Hessian of the LDDMM energy can be separated into the Hessian of the regularization energy and the Hessian of the image mismatch energy. Thus we can separately calculate Hessian vector products for these two parts based on Eq. 7 as:

$$H_m^{\text{regularization}} \delta m_0 = 2K \delta m_0, \quad H_m^{\text{mismatch}} \delta m_0 = -\delta \hat{m}(0).$$

A simple low-rank pseudoinverse as an approximation of the Hessian would result in approximation errors for both  $H_m^{\text{regularization}}$  and  $H_m^{\text{mismatch}}$ . This can be partially avoided by using the *exact* inverse of the Hessian of the regularization combined with an *approximation* for the image mismatch Hessian. To compute the covariance matrix, we realize that for many ill-posed inverse problems, the spectrum of the absolute values of the eigenvalues of the image mismatch Hessian decays rapidly to zero. Fig. 1 shows an example of the largest 1000 absolute values of the eigenvalues for a 2D  $100 \times 140$  pixels heart registration case for initial momentum LDDMM. By computing only a few dominant eigenmodes (using an iterative eigensolver such as the Lanczos method) of the image mismatch Hessian with respect to initial momentum, we can accurately approximate the Hessian with much less memory and computational effort. Suppose we approximate the image mismatch Hessian with  $k$  dominating eigenmodes as

$$H_{m(k)}^{\text{mismatch}} \approx V_{m(k)}^T D_{m(k)} V_{m(k)}.$$

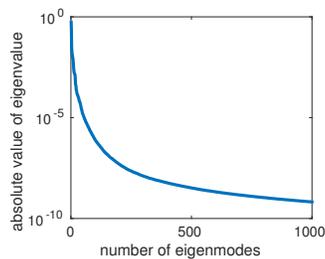
Here  $D_{m(k)}$  is a  $k \times k$  diagonal matrix, where the diagonal elements are the eigenvalues;  $V_{m(k)}$  is a  $k \times n$  matrix, where  $n$  is the number of all parameters, and each row of  $V_k$  is an eigenvector. For simplicity we write  $H_m^{\text{regularization}}$  as  $H_m^{\text{reg}}$ . We can then approximate the covariance matrix  $\Sigma_m$  as

$$\Sigma_m = (H_m^{\text{reg}} + H_m^{\text{mismatch}})^{-1} \approx (H_m^{\text{reg}} + V_{m(k)}^T D_{m(k)} V_{m(k)})^{-1}. \quad (8)$$

Since we have a closed-form solution for the inverse of  $H_m^{\text{reg}}$ , we can directly apply the Woodbury identity to the right hand side of equation 8 and obtain

$$\Sigma_m \approx (H_m^{\text{reg}})^{-1} - (H_m^{\text{reg}})^{-1} V_{m(k)}^T (D_{m(k)}^{-1} + V_{m(k)} (H_m^{\text{reg}})^{-1} V_{m(k)}^T)^{-1} V_{m(k)} (H_m^{\text{reg}})^{-1}. \quad (9)$$

<sup>1</sup> Note that this would be possible when using a landmark-based LDDMM variant where the parameterization of the deformation becomes finite-dimensional.



**Fig. 1:** First 1000 largest absolute eigenvalues of image mismatch Hessian for a  $100 \times 140$  pixels heart registration case using initial momentum LDDMM.

The advantage of this formulation is that  $D_{m(k)}^{-1} + V_{m(k)}(H_m^{\text{reg}})^{-1}V_{m(k)}^T$  is a small  $k \times k$  matrix, and its inverse can be computed easily using dense algebra.

In LDDMM we could quantify the uncertainty through the covariance with respect to either the initial momentum or its corresponding initial velocity. If we use the initial momentum, the approximation of  $(D_{m(k)}^{-1} + V_{m(k)}(H_m^{\text{reg}})^{-1}V_{m(k)}^T)^{-1}$  will be accurate for small  $k$ , because the image mismatch Hessian for the initial momentum has a rapidly decreasing eigen-spectrum. However, the inverse of the regularization kernel would be  $(H_m^{\text{reg}})^{-1} = (2K)^{-1} = \frac{1}{2}L$ , which is a rough kernel. In experiments, using the rough kernel significantly increases the difference between approximations for different  $k$ 's. On the other hand, in the initial velocity formulation,  $(H_m^{\text{reg}})^{-1}$  is a smoothing kernel, and it further decreases the approximation difference for different  $k$ 's. Unfortunately, the image mismatch Hessian with respect to the initial velocity does not have a fast-decreasing spectrum due to the implicit smoothness of the initial velocity.

We want to use the rapidly decreasing eigen-spectrum of the momentum-based formulation, but at the same time we want to avoid its rough kernel when calculating the covariance matrix. Our solution is to use the low-rank approximation of the initial *momentum* image mismatch Hessian to approximate the covariance with respect to the initial *velocity*. Recall the relation between momentum and velocity:  $v = Km$ . This means the Jacobian of  $v$  with respect to  $m$  is  $J_{\frac{v}{m}} = K$ . Thus by change of variables, we can obtain our final approximation of the covariance with respect to the initial velocity as

$$\Sigma_v = J_{\frac{v}{m}} \Sigma_m J_{\frac{v}{m}}^T \approx \frac{K}{2} - \frac{1}{4}V_{m(k)}^T(D_{m(k)}^{-1} + V_{m(k)}\frac{L}{2}V_{m(k)}^T)^{-1}V_{m(k)}. \quad (10)$$

## 4 Experiments and Discussions

We evaluate our proposed model using synthetic and real data. In the following experiments,  $L$  corresponds to the invertible and self-adjoint Sobolev operator,  $L = a\Delta^2 + b\Delta + c$ , with  $a = 9 \times 10^{-4}$ ,  $b = -6 \times 10^{-2}$ , and  $c = 1$ ;  $\sigma = 0.1$ . For the eigen-decomposition we use PROPACK [6], which uses a Lanczos bidiagonalization method. [Computing 200 dominant eigenvectors for the 2D synthetic example, which gives a very accurate Hessian estimation, requires less than 3 min in Matlab; however computing the full Hessian requires more than 30 min. Hence, our method is an order of magnitude faster.](#) All images below are rescaled to a  $[0, 1]$  space range.

**Synthetic data.** The synthetic example is a simple registration of an expanding square. Figure 3 shows the source image, the target image, and the final registration result. The size of the image is  $51 \times 51$  pixels, thus the size of the Hessian is  $5202 \times 5202$ . Since this Hessian is small, we compute the full covariance matrix using finite differences as the ground truth for our covariance estimation. We compare our method with the low rank pseudoinverse using both the initial momentum Hessian and the initial velocity Hessian.

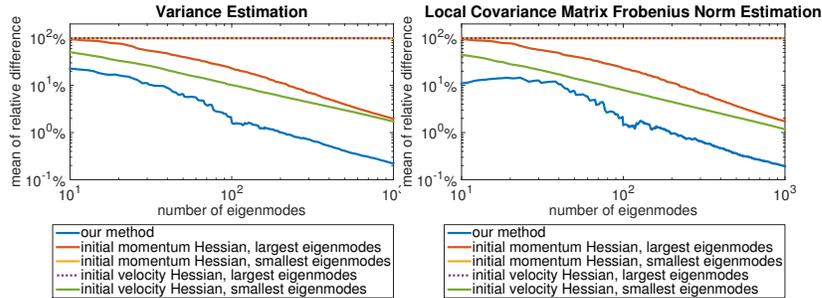


Fig. 2: Relative low-rank approximation differences for synthetic data.

We approximate two uncertainty measures: the variance of each parameter, and the spatially localized covariance matrix for each image pixel<sup>2</sup>. Fig. 2 shows the mean of relative differences with respect to the ground truth for different methods. Our method outperforms the other methods even with very few eigenmodes selected. In some test cases (see supplementary material) when only a few eigenmodes are selected (e.g., 10), the pseudoinverse of the Hessian achieves relatively better accuracy than our method, but in those cases, the relative difference can be up to 100%, making the approximation unusable. For reasonable numbers of eigenmodes (e.g., larger than 100), our method is always better.

E.g., in Fig 2, while 100 eigenmodes result in a relative error of around 1% for our method, around 1000 are required for the other approaches for similar accuracy.

To visualize the uncertainty information, we extract the local covariance matrices from the approximated covariance matrix and visualize these matrices as ellipses on the source image. Fig. 3 shows the uncertainty visualization for the synthetic data. The ellipses are estimated using 200 dominant eigenmodes from the image mismatch Hessian. The color indicates the determinant of local covariance matrices. The closer to the center of the square, the smaller the determinant is, meaning a more confident registration result closer to the center. Furthermore, the uncertainty along the edge is larger than the uncertainty perpendicular to the edge, indicating the aperture problem of image registration.

**2D heart image.** We use cardiac data from the Sunnybrook cardiac MR database [7]. The image corresponds to a beating heart of a 63 years normal individual at two different time points.

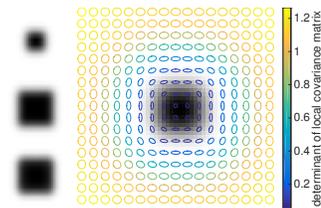
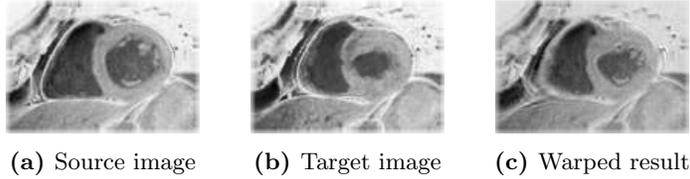


Fig. 3: Square registration case. Left: top to bottom: source image, target image, warped result. Right: Uncertainty visualization by ellipses, mapped on source image.

<sup>2</sup> In 2D this amounts to computing  $\begin{pmatrix} \sigma_{xx}^2(i, j) & \sigma_{xy}^2(i, j) \\ \sigma_{xy}^2(i, j) & \sigma_{yy}^2(i, j) \end{pmatrix}$  for each pixel location  $(i, j)$ , i.e., not considering non-local cross-variances.



**Fig. 4:** Heart registration test case

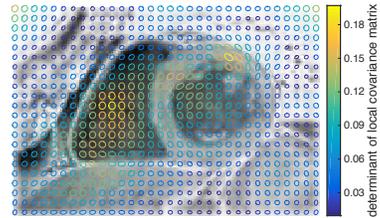
We cropped the axial images to a common 2D rectangular region around the heart. Fig. 4 shows the heart image and registration result. The size of the heart image is  $100 \times 140$  pixels, resulting in a  $28000 \times 28000$  Hessian. We select the top 500 eigenmodes for approximation and achieve a mean relative difference for variance of 1.13%, and for the Frobenius norm of the local covariance matrix of 0.94%. Using the pseudoinverse of the full Hessian gives a mean relative difference of 6.81% and 6.27% for the initial velocity Hessian, and 31.17% and 30.04% for the initial momentum Hessian.

Fig. 5 shows the uncertainty visualization for the initial velocity on the source image. From the image we see that the area inside the ventricle has high uncertainty, indicating low deformation confidence in the isotropic area. Also, there exists high uncertainty at the upper right edge of the atrium. This indicates high uncertainty for shifting along the edge.

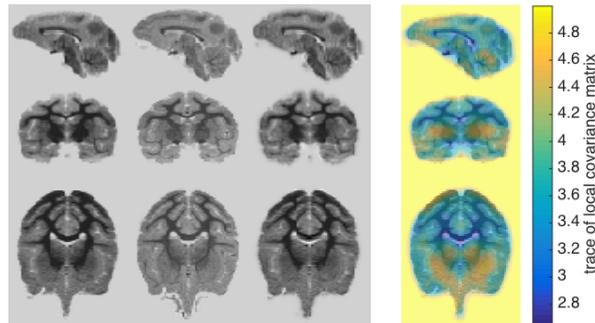
**3D brain image.** Here our data is two MR images ( $75 \times 90 \times 60$  voxels) of a macaque monkey at 6 months and 12 months of age. The size of this Hessian is  $1,215,000 \times 1,215,000$ . We calculate the largest 1000 eigenmodes for covariance approximation. For visualization of uncertainty, we use the trace of the local covariance matrix. Fig. 6 shows the 3D test case as well as the uncertainty visualization. We can see that although the deformation is very small despite of overall intensity change, our method can still capture isotropic areas that have very small changes. These areas have a higher uncertainty compared to others.

**Discussions.** Although we focus on LDDMM, our method has much wider applicability. Using the Hessian-vector product for efficient low-rank Hessian optimization is relevant for many other non-parametric registration approaches formulated with similar regularization and image similarity measures, such as optic flow and registration by stationary velocity fields, since they are usually over-parameterized. Even for low-dimensional parameterization methods such as landmark LDDMM, our method could be used to compute the full Hessian. Our method could also be used as in [9] for uncertainty-based smoothing, and for surgical- or radiation-treatment planning to inform margins. Finally, comparing our method with sampling-based methods for LDDMM [12] will be useful.

**Support.** This research is supported by NSF EECS-1148870 and EECS-0925875.



**Fig. 5:** Uncertainty visualization of heart test case mapped on source image.



**Fig. 6:** 3D monkey brain test case. Left to right: source image, target image, warped result, visualization of trace of local covariance matrix on the source image.

## References

1. Beg, M., Miller, M., Trouné, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *IJCV* 61(2), 139–157 (2005) [2](#)
2. Fitzpatrick, J., West, J.: The distribution of target registration error in rigid-body point-based registration. *IEEE Trans. Med. Imaging* 20(9), 917–927 (2001) [1](#)
3. Flath, H.P., Wilcox, L.C., Akçelik, V., Hill, J., van Bloemen Waanders, B., Ghattas, O.: Fast algorithms for Bayesian uncertainty quantification in large-scale linear inverse problems based on low-rank partial Hessian approximations. *SIAM Journal on Scientific Computing* 33(1), 407–432 (2011) [2](#)
4. Kalmikov, A.G., Heimbach, P.: A hessian-based method for uncertainty quantification in global ocean state estimation. *SIAM Journal on Scientific Computing* 36(5), S267–S295 (2014) [2](#)
5. Kybic, J.: Bootstrap resampling for image registration uncertainty estimation without ground truth. *IEEE Trans. Med. Imaging* 19(1), 64–73 (2010) [1](#)
6. Larsen, R.M.: Lanczos bidiagonalization with partial reorthogonalization (1998) [5](#)
7. Radau, P., Lu, Y., Connelly, K., Paul, G., Dick, A., Wright, G.: Evaluation framework for algorithms segmenting short axis cardiac MRI. *The MIDAS Journal* 49 (2009) [6](#)
8. Risholm, P., Janoos, F., Norton, I., Golby, A.J., Wells, W.M.: Bayesian characterization of uncertainty in intra-subject non-rigid registration. *Medical Image Analysis* 17(5), 538–555 (2013) [1](#), [2](#)
9. Simpson, I., Woolrich, M., Groves, A., Schnabel, J.: Longitudinal brain mri analysis with uncertain registration. In: Fichtinger, G., Martel, A., Peters, T. (eds.) *MICCAI 2011, Part II, LNCS*, vol. 6892, pp. 647–654. Springer, Heidelberg (2011) [1](#), [7](#)
10. Vialard, F.X., Risser, L., Rueckert, D., Cotter, C.J.: Diffeomorphic 3d image registration via geodesic shooting using an efficient adjoint calculation. *IJCV* 97(2), 229–241 (2012) [2](#), [3](#)
11. Watanabe, T., Scott, C.: Spatial confidence regions for quantifying and visualizing registration uncertainty. In: *Biomedical Image Registration*, pp. 120–130. Springer (2012) [1](#)
12. Zhang, M., Singh, N., Fletcher, P.: Bayesian estimation of regularization and atlas building in diffeomorphic image registration. In: Gee, J., Joshi, S., Pohl, K., Wells, W., Zillei, L. (eds.) *IPMI 2013, LNCS*, vol. 7917, pp. 37–48. Springer Berlin Heidelberg (2013) [1](#), [2](#), [7](#)