Supplementary Material for: "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 document contains additional material for the main article. It contains the full derivation of the tangent linear model and the tangent linear adjoint model for computing the Hessian vector product. We also provide relative difference comparisons of different low-rank approximation methods for the 2D heart model that could not be included due to space constraints.

Appendix A Full derivation of Hessian vector product equations

We want to compute the Hessian vector product for the initial momentum version of the shooting formulation of LDDMM

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

with dynamic constraints

$$\text{momentum evolution: } m_t + \mathrm{ad}_v^* m = 0,\ m(0) = m_0, \tag{2}$$

$$\text{image evolution: } \quad I_t + \nabla I^T v = 0,\ I(0) = I_0, \tag{3}$$

$$\text{momentum-velocity transformation: } \quad m - Lv = 0. \tag{4}$$

We add the dynamic constraints into the energy using time-dependent adjoint variables $\hat{m}$, $\hat{I}$ and $\hat{v}$ through time and space $[0, 1]^d$. This gives us the augmented energy

$$\widetilde{E}(m_0, I, v, \hat{m}, \hat{I}, \hat{v}) = E(m_0) + \int_0^1 \langle \hat{m}, m_t + \mathrm{ad}_v^* m\rangle + \langle \hat{I}, I_t + \nabla I^T v\rangle + \langle \hat{v}, m - Lv\rangle dt. \tag{5}$$

Computing the second variation for this augmented energy requires computing

$$
\begin{aligned}
\sigma^2\widetilde{E} =& \frac{\partial^2}{\partial\epsilon^2}\widetilde{E}(m_0+\epsilon\delta m_0, I+\epsilon\delta I, v+\epsilon\delta v, \hat{m}+\epsilon\delta\hat{m}, \hat{I}+\epsilon\delta\hat{I}, \hat{v}+\epsilon\delta\hat{v})|_{\epsilon\to 0} \\
=& \frac{\partial^2}{\partial\epsilon^2}\Big(\langle m_0+\epsilon\delta m_0, K(m_0+\epsilon\delta m_0)\rangle + \frac{1}{\sigma^2}||I(1)+\epsilon\delta I(1)-I_1||^2+ \\
& \int_0^1 \langle \hat{m}+\epsilon\delta\hat{m}, m_t+\epsilon\delta m_t + \mathrm{ad}^*_{v+\epsilon\delta v}(m+\epsilon\delta m)\rangle + \\
& \quad \langle \hat{I}+\epsilon\delta\hat{I}, I_t+\epsilon\delta I_t + \nabla(I^T+\epsilon\delta I^T)(v+\epsilon\delta v)\rangle + \\
& \quad \langle \hat{v}+\epsilon\delta\hat{v}, m+\epsilon\delta m - L(v+\epsilon\delta v)\rangle dt\Big)|_{\epsilon\to 0}
\end{aligned} \tag{6}
$$

Computing Eq. 6 results in

$$
\begin{aligned}
\sigma^2\widetilde{E} =& \underbrace{\langle \delta m_0, 2K\delta m_0\rangle}_{1} + \underbrace{\frac{2}{\sigma^2}\langle \delta I(1), \delta I(1)\rangle}_{2} + \\
& 2\int_0^1 \Big(\underbrace{\langle \delta\hat{m}, \delta m_t + ad_v^*\delta m + ad_{\delta v}^* m\rangle}_{3} + \underbrace{\langle \hat{m}, ad_{\delta v}^*\delta m\rangle}_{4} + \\
& \quad \underbrace{\langle \delta\hat{I}, \delta I_t + \nabla I^T\delta v + \nabla\delta I^T v\rangle}_{5} + \underbrace{\langle \hat{I}, \nabla\delta I^T\delta v\rangle}_{6} + \\
& \quad \underbrace{\langle \delta\hat{v}, \delta m - L\delta v\rangle}_{7}\Big) dt
\end{aligned} \tag{7}
$$

Here we label every part of the equation to help following the computations. For the transformations in the equations below, all parts are labeled according to the number of the original component in Eq. 7. We assume periodic boundary conditions for the momentum and the veloctity. Furthermore, we assume that $\delta I(0) = 0$, i.e. the initial image is fixed. Thus we can perform the following transformations

$$
\begin{gathered}
3: \int_0^1 \langle \delta\hat{m}, \delta m_t\rangle \; dt = \int_0^1 \langle -\delta\hat{m}_t, \delta m\rangle \; dt + \langle \delta\hat{m}(1), \delta m(1)\rangle - \langle \delta\hat{m}(0), \delta m(0)\rangle, \\
3: \langle \delta\hat{m}, ad_v^*\delta m\rangle = \langle ad_v\delta\hat{m}, \delta m\rangle, \\
3: \langle \delta\hat{m}, ad_{\delta v}^* m\rangle = \langle ad_{\delta v}\delta\hat{m}, m\rangle = \langle -ad_{\delta\hat{m}}\delta v, m\rangle = \langle \delta v, -ad^*_{\delta\hat{m}} m\rangle, \\
4: \langle \hat{m}, ad_{\delta v}^*\delta m\rangle = \langle ad_{\delta v}\hat{m}, \delta m\rangle = \langle -ad_{\hat{m}}\delta v, \delta m\rangle = \langle \delta v, -ad^*_{\hat{m}}\delta m\rangle, \\
5: \int_0^1 \langle \delta\hat{I}, \delta I_t\rangle \; dt = \int_0^1 \langle -\delta\hat{I}_t, \delta I\rangle \; dt + \langle \delta\hat{I}(1), \delta I(1)\rangle, \\
5: \langle \delta\hat{I}, \nabla I^T\delta v\rangle = \langle \nabla I\delta\hat{I}, \delta v\rangle, \\
5: \langle \delta\hat{I}, \nabla\delta I^T v\rangle = \langle \delta I, -div(v\delta\hat{I})\rangle, \\
6: \langle \hat{I}, \nabla\delta I^T\delta v\rangle = \langle \nabla\delta I\hat{I}, \delta v\rangle = \langle \delta I, -div(\delta v\hat{I})\rangle,
\end{gathered}
$$

$$7 : \langle \delta\hat{v}, L\delta v\rangle = \langle L\delta\hat{v}, \delta v\rangle.$$

Putting these transformations into Eq. 7 we have

$$\begin{aligned}
\sigma^2\widetilde{E} = &\underbrace{\langle \delta m_0, 2K\delta m_0\rangle}_{1} + \underbrace{\frac{2}{\sigma^2}\langle \delta I(1), \delta I(1)\rangle}_{2} - \underbrace{\langle \delta\hat{m}(0), \delta m(0)\rangle}_{3} + \underbrace{\langle \delta\hat{I}(1), \delta I(1)\rangle}_{5} + \\
&\int_0^1 \Big(\underbrace{\langle \delta\hat{m}, \delta m_t + ad_v^*\delta m + ad_{\delta v}^* m\rangle + \langle \delta m, -\delta\hat{m}_t\rangle + \langle \delta m, ad_v\delta\hat{m}\rangle + \langle \delta v, -ad_{\delta\hat{m}}^* m\rangle}_{3} + \\
&\underbrace{\langle \delta m, ad_{\delta v}\hat{m}\rangle + \langle \delta v, -ad_{\hat{m}}^*\delta m\rangle}_{4} + \\
&\underbrace{\langle \delta\hat{I}, \delta I_t + \nabla I^T\delta v + \nabla\delta I^T v\rangle + \langle \delta I, -\delta\hat{I}_t\rangle + \langle \delta v, \nabla I\delta\hat{I}\rangle + \langle \delta I, -div(v\delta\hat{I})\rangle}_{5} + \\
&\underbrace{\langle \delta v, \nabla\delta I\hat{I}\rangle + \langle \delta I, -div(\delta v\hat{I})\rangle}_{6} + \underbrace{\langle \delta\hat{v}, \delta m - L\delta v\rangle + \langle \delta m, \delta\hat{v}\rangle + \langle \delta v, L\delta\hat{v}\rangle}_{7} \Big) dt.
\end{aligned}$$

Combining these terms, we get

$$\begin{aligned}
\sigma^2\widetilde{E} =&\langle \delta m_0, 2K\delta m_0 - \delta\hat{m}(0)\rangle + \langle \delta I(1), \frac{2}{\sigma^2}\delta I(1) + \delta\hat{I}(1)\rangle + \\
&\langle \delta\hat{m}(1), \delta m(1)\rangle + \\
&\int_0^1 \Big(\langle \delta\hat{m}, \delta m_t + ad_v^*\delta m + ad_{\delta v}^* m\rangle + \\
&\langle \delta\hat{I}, \delta I_t + \nabla I^T\delta v + \nabla\delta I^T v\rangle + \\
&\langle \delta\hat{v}, \delta m - L\delta v\rangle + \\
&\langle \delta v, -ad_{\delta\hat{m}}^* m - ad_{\hat{m}}^*\delta m + \nabla I\delta\hat{I} + \nabla\delta I\hat{I} - L\delta\hat{v}\rangle + \\
&\langle \delta I, -\delta\hat{I}_t - div(\delta v\hat{I} + v\delta\hat{I})\rangle + \\
&\langle \delta m, -\delta\hat{m}_t + ad_{\delta v}\hat{m} + ad_v\delta\hat{m} + \delta\hat{v} = 0\rangle \Big) dt
\end{aligned} . \tag{8}$$

Regarding everything inside the integration as optimality conditions, and extracting the boundary condition as $\delta m(0) = \delta m_0, \quad \delta I(0) = 0, \quad \delta\hat{m}(1) = 0, \quad \delta\hat{I}(1) = -\frac{2}{\sigma^2}\delta I(1)$, we finally get our equation for computing Hessian-vector product

$$\nabla^2 E\delta m_0 = 2K\delta m_0 - \delta\hat{m}(0) \tag{9}$$

together with a tangent linear model (TLM)

$$\begin{cases} \delta m_t + ad_{\delta v}^* m + ad_v^*\delta m = 0, \quad \delta m(0) = \delta m_0 \\ \delta I_t + \nabla\delta I^T v + \nabla I^T\delta v = 0, \quad \delta I(0) = 0 \\ \delta m - L\delta v = 0 \end{cases} \tag{10}$$

and a tangent linear adjoint model (TLAM)

$$\begin{cases} -\delta\hat{m}_t + ad_{\delta v}\hat{m} + ad_v\delta\hat{m} + \delta\hat{v} = 0, \quad \delta\hat{m}(1) = 0 \\ -\delta\hat{I}_t - div(\delta v\hat{I} + v\delta\hat{I}) = 0, \quad \delta\hat{I}(1) = -\frac{2}{\sigma^2}\delta I(1) \\ -ad^*_{\delta\hat{m}}m - ad^*_{\hat{m}}\delta m + \nabla I\delta\hat{I} + \nabla\delta I\hat{I} - L\delta\hat{v} = 0. \end{cases} \tag{11}$$

Note that the TLM and TLAM are actually the linearized versions of the forward equations:

$$\begin{cases} m_t + ad^*_v m = 0, \quad m(0) = m_0, \\ I_t + \nabla I^T v = 0, \quad I(0) = I_0, \\ m - Lv = 0, \end{cases} \tag{12}$$

and the adjoint equations:

$$\begin{cases} -\hat{m}_t + ad_v\hat{m} + \hat{v} = 0, \quad \hat{m}(1) = 0, \\ -\hat{I}_t - \mathrm{div}(v\hat{I}) = 0, \quad \hat{I}(1) = -\frac{2}{\sigma^2}(I(1) - I_1), \\ -ad^*_{\hat{m}}m + \nabla I\hat{I} - L\hat{v} = 0 \end{cases} \tag{13}$$

for LDDMM shooting. Hence, a more direct derivation could simply be linearizing the forward model and the adjoint model instead of computing the second variation. Furthermore, calculating $\delta\hat{m}(0)$ simply needs a forward-backward sweep using the TLM and TLAM.

Appendix B Supplementary result

Fig. 1 shows the relative difference of both variance estimation and local covariance matrix Frobenius norm estimation for the 2D heart case. Comparing to other methods, our method has higher relative difference when using the first 70 eigenmodes. However, the smallest mean relative difference using 70 eigenmodes is as large as 24.9% for variance estimation, and 23.5% for local covariance estimation. Thus, one cannot get a reliable estimation using such a small number of eigenmodes. When the number of eigenmodes selected is larger than 70, our method always achieves better accuracy. Furthermore, our method can achieve a reasonably high accuracy much faster than other low-rank methods. For example, while the pseudoinverse of the initial velocity Hessian needs 1000 eigenmodes to achieve a 3.87% mean relative variance difference, our method only needs 237 eigenmodes to get the same accuracy.

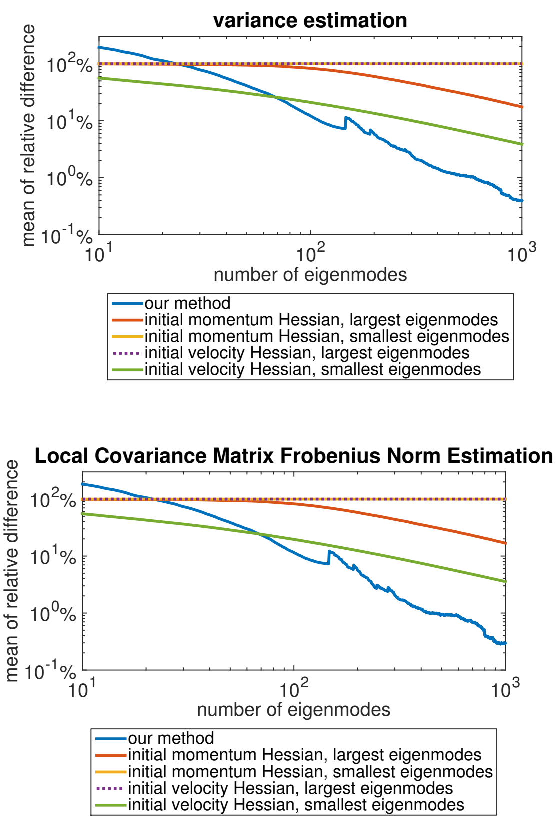

Fig. 1: Relative low-rank approximaton differences for heart data.