



HAL
open science

Joint Image Reconstruction and Motion Estimation for Spatiotemporal Imaging

Chong Chen, Barbara Gris, Ozan Öktem

► **To cite this version:**

Chong Chen, Barbara Gris, Ozan Öktem. Joint Image Reconstruction and Motion Estimation for Spatiotemporal Imaging. *SIAM Journal on Imaging Sciences*, 2019. <hal-02386215>

HAL Id: hal-02386215

<https://hal.science/hal-02386215v1>

Submitted on 29 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

1 Joint Image Reconstruction and Motion Estimation for Spatiotemporal Imaging*

2 Chong Chen[†], Barbara Gris[‡], and Ozan Öktem[§]

3
4 **Abstract.** We propose a variational model for joint image reconstruction and motion estimation applicable to
5 spatiotemporal imaging. This model consists of two parts, one that conducts image reconstruction in
6 a static setting and another that estimates the motion by solving a sequence of coupled indirect image
7 registration problems, each formulated within the large deformation diffeomorphic metric mapping
8 framework. The proposed model is compared against alternative approaches (optical flow based
9 model and diffeomorphic motion models). Next, we derive efficient algorithms for a time-discretized
10 setting and show that the optimal solution of the time-discretized formulation is consistent with that
11 of the time-continuous one. The complexity of the algorithm is characterized and we conclude by
12 giving some numerical examples in 2D space + time tomography with very sparse and/or highly
13 noisy data.

14 **Key words.** spatiotemporal imaging, image reconstruction, motion estimation, joint variational model, shape
15 theory, large diffeomorphic deformations

16 **AMS subject classifications.** 65F22, 65R32, 65R30, 65D18, 65J22, 65J20, 65L09, 68U10, 94A12, 94A08, 92C55,
17 54C56, 57N25, 47A52

18 **1. Introduction.** Image reconstruction is challenging in a spatiotemporal setting where
19 the object being imaged undergoes a temporal evolution during the data acquisition. This
20 arises in tomographic imaging of the heart or lungs [13, 29] where the image recovery needs to
21 include estimating and compensating for the unknown motion of the organs. As an example,
22 data in positron emission tomography (PET) cardiac imaging is acquired over a relatively
23 long period of time (often in the range of minutes). Here, respiratory and cardiac motion
24 displaces organs of interest with up to 20–40 mm [48, 58]. Failing to correct for such motion
25 leads to a degradation in image quality [29].

26 Data in spatiotemporal imaging is time (or quasi-time) dependent and an obvious approach
27 is to decompose it into sub-sets (gates) such that data within each gate can be assumed to
28 originate from the object in a fixed temporal state [25, 35, 37]. As an example, in the context
29 of PET imaging of the heart or lungs, the gates correspond to the phases of breathing and/or
30 cardiac motion [23, 32, 15]. Clearly a too fine gating leads to data with worse signal-to-noise-
31 ratio, whereas too coarse gating leads to motion artefacts. Much work has been done on how
32 to optimize the gating in order to obtain the optimal balance between signal-to-noise-ratio
33 and motion artefacts [27, 56].

34 Algorithms for spatiotemporal image reconstruction that act on gated data either perform
35 image reconstruction followed by motion estimation (sequential approach) or perform these

*Submitted to the editors 2018-12-18.

Funding: The work of all authors was supported by the Swedish Foundation for Strategic Research grant No. AM13-0049. Chen was in addition supported by the Beijing Natural Science Foundation grant No. Z180002.

[†]LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (chench@lsec.cc.ac.cn).

[‡]LJLL–Laboratoire Jacques-Louis Lions, Sorbonne Université, 75005 Paris, France (bgris.maths@gmail.com).

[§]Department of Mathematics, KTH–Royal Institute of Technology, 100 44 Stockholm, Sweden (ozan@kth.se).

36 two steps jointly [7, 29]. In the sequential approach, one starts with applying static image
 37 reconstruction on data from each of the gates, resulting in a series of (low-resolution) images.
 38 Then, one selects a reconstructed image as target and registers the other reconstructed im-
 39 ages against this target. The final reconstructed image is obtained as an average of all the
 40 registered images [2, 3, 22, 24, 30, 29]. The other alternative that seeks to jointly perform
 41 image reconstruction and motion estimation is much more complex. Several approaches have
 42 been suggested for how to do this, such as [7, 47, 51, 28, 36, 9, 8, 44, 34, 52, 6, 33, 33, 10, 13].

43 The approach taken in this paper seeks to jointly perform image reconstruction and motion
 44 estimation. The motion is here modelled as deformations induced by diffeomorphisms that
 45 are provided by the large deformation diffeomorphic metric mapping (LDDMM) framework,
 46 which is a well-developed framework for diffeomorphic image registration [54, 26, 38, 5, 31,
 47 60, 55, 11, 12]. Joint image reconstruction and motion estimation can be based on the growth
 48 model in LDDMM [31] as shown in [33] for 3D space + time computed tomography (CT).
 49 It can also be used for indirect image registration as shown in [17], which also proves the
 50 resulting model is stable.

51 *Specific contributions.* The main contribution is a new variational model for joint im-
 52 age reconstruction and motion estimation in spatiotemporal imaging based on the LDDMM
 53 framework with deformable templates. As already mentioned, there are two components: one
 54 corresponding to modified static image reconstruction, and the other to estimating the mo-
 55 tion. To handle the latter, we generalize the LDDMM framework to a setting that applies to
 56 a series of coupled sequential indirect image registration problems.

57 The mathematical properties of the proposed variational model is compared against the
 58 optical flow based model in [13] and the diffeomorphic motion model in [33]. The comparison
 59 shows that the proposed model has some desirable properties in terms of the optimal solution,
 60 e.g., guaranteeing elastically large diffeomorphic deformations, averagely distributed w.r.t.
 61 time t , and non-vanishing at both the initial and the end time points, etc. Moreover, a
 62 computationally efficient gradient-based iterative scheme is presented for a time-discretized
 63 formulation. More importantly, the optimal solution of the time-discretized problem is shown
 64 to be consistent with that of the time-continuous one. Most of the computationally demanding
 65 parts relate to computations involving linearized deformations [43].

66 *Outline.* Section 2 presents a general variational model for joint image reconstruction
 67 and motion estimation. We also briefly review necessary parts of the LDDMM theory in
 68 subsection 3.1. Subsection 3.2 proposes the new variational model and subsection 3.3 makes
 69 the theoretical comparison between this and other existing models. Section 4 gives the detailed
 70 numerical algorithms associated with the proposed model. Results of numerical experiments
 71 are presented in section 5 in the context of 2D space + time tomography. Finally, we conclude
 72 with section 6 that discusses future work.

73 **2. A variational model for joint image reconstruction and motion estimation.** Spa-
 74 tiotemporal imaging often leads to the task of recovering a spatially distributed quantity
 75 (image) that exhibits temporal variations given indirect time-dependent noisy observations
 76 (measured data). Hence, both the image and its motion are unknown.

77 **2.1. General spatiotemporal inverse problem.** To formalise the notion of spatiotemporal
 78 imaging, let $f: [t_0, t_1] \times \Omega \rightarrow \mathbb{R}^k$ denote the spatiotemporal image we need to recover. Here k

79 is the number of channels/modalities (often $k = 1$) and $\Omega \subset \mathbb{R}^n$ is the spatial region containing
 80 the support of the image. Without loss of generality, the general (quasi-)time domain $[t_0, t_1]$
 81 can be re-parameterized onto $[0, 1]$.

82 The spatiotemporal inverse problem is now stated mathematically as the problem of re-
 83 constructing the spatiotemporal image $f(t, \cdot) \in \mathcal{X}$ given measured data $g(t, \cdot) \in \mathcal{Y}$ where

$$84 \quad (2.1) \quad g(t, \cdot) = \mathcal{T}(t, f(t, \cdot)) + g_{\text{noise}}(t, \cdot) \quad \text{for } t \in [0, 1].$$

85 Here, \mathcal{X} (reconstruction space) is the vector space of all possible images on a fixed domain Ω ,
 86 \mathcal{Y} (data space) is the vector space of all possible data, and $g_{\text{noise}}(t, \cdot) \in \mathcal{Y}$ is the observation
 87 noise in data. Furthermore, $\mathcal{T}(t, \cdot): \mathcal{X} \rightarrow \mathcal{Y}$ is a time-dependent forward operator, for short
 88 denoted by \mathcal{T}_t , that models how an image at time t gives rise to data in absence of noise
 89 or measurement errors (e.g., a stack of ray transforms with time varying data acquisition
 90 parameters in CT or the forward model in PET with time dependent attenuation, etc.) [41].

91 To proceed, we further specify the form of the spatiotemporal image $f(t, \cdot)$ by making
 92 use of deformable templates from shape theory.

93 **2.2. Spatiotemporal inverse problem and shape theory.** Shape theory seeks to develop
 94 quantitative tools to study shapes and their variability. The theory can be traced back to
 95 work by D’Arcy Thompson [53]. The underlying idea is that shapes of objects are represented
 96 as a deformation that acts on a template [55]. Hence, the template is the “shape invariant”
 97 part of the object whereas the set of deformations model the various shapes that can be
 98 generated from the template. Shape similarity between two objects can then be quantified as
 99 the “cost” of deforming one object into the other by means of a minimal deformation in the
 100 set of deformations. Further details are given in [60, 31, 40].

101 Bearing in mind the above, we separate the spatial and temporal components of a spa-
 102 tiotemporal image as

$$103 \quad (2.2) \quad f(t, \cdot) := \mathcal{W}(\phi_t, I) \quad \text{for some } \phi_t \in \mathcal{G} \text{ and } I \in \mathcal{X}.$$

104 Here $I: \Omega \rightarrow \mathbb{R}$ (template) is the time-independent spatial component, $\mathcal{W}: \mathcal{G} \times \mathcal{X} \rightarrow \mathcal{X}$ is a
 105 deformation operator the models how a deformation parameter $\phi_t \in \mathcal{G}$ deforms the an image,
 106 and $\phi_t: \Omega \rightarrow \Omega$ is the temporal evolution of the deformation parameter.

107 We will assume deformations are given by diffeomorphic group actions, so \mathcal{G} is a suitable
 108 subgroup of the diffeomorphic group on Ω and \mathcal{W} is given by the group action of \mathcal{G} on \mathcal{X} ,
 109 i.e., $\mathcal{W}(\phi_t, I) := \phi_t \cdot I$. The spatiotemporal inverse problem can then be written as

$$110 \quad (2.3) \quad g(t, \cdot) = \mathcal{T}_t(\phi_t \cdot I) + g_{\text{noise}}(t, \cdot) \quad \text{for } t \in [0, 1].$$

111 Notice that $f(t, \cdot) = \phi_t \cdot I$ is the spatiotemporal image at time t generated from the template
 112 I by the diffeomorphism ϕ_t . Hence, the above inverse problem calls for jointly recovering the
 113 (time-independent) template I and the curve of diffeomorphisms $t \mapsto \phi_t$.

114 **Common group actions.** In imaging, there are several diffeomorphic group actions that one
 115 could consider. A natural one merely moves intensities without changing them (geometric
 116 deformations) [17]:

$$117 \quad (2.4) \quad \phi_t \cdot I := I \circ \phi_t^{-1}.$$

118 An alternative group action corresponds to mass-preserving deformations [17, 60]:

$$119 \quad (2.5) \quad \phi_t \cdot I := |D(\phi_t^{-1})| I \circ \phi_t^{-1}.$$

120 Here, $|D(\phi)|$ denotes the determinant of the Jacobian of ϕ . This group action adjusts the
121 intensity values but preserves the total mass.

122 **2.3. Joint image reconstruction and motion estimation.** It is not difficult to observe
123 that the inverse problem in (2.3) decomposes into two sub-problems. Given a diffeomorphism
124 ϕ_t , the original problem reduces to a static image reconstruction problem where one seeks
125 to recover the template “ I ” from noisy measured data. The data represents indirect noisy
126 measurements of the template that has undergone a known diffeomorphic deformation. Con-
127 versely, given the template I , the original problem reduces to an indirect registration problem
128 at each point in time t . More precisely, the task is to recover a curve of diffeomorphisms
129 “ $t \mapsto \phi_t$ ” that registers the template at time t against a target that is known indirectly
130 through noisy measured data $g(t, \cdot)$. Such a “sequential indirect image registration problem”
131 can be seen as a temporal version of the indirect image registration problem in [17].

132 Inverse problem of the type (2.3) in imaging applications are often ill-posed. A flexible
133 framework for regularizing many of these problems is through a variational formulation [46].
134 The idea is to add regularization functional that penalizes a maximum likelihood solution and
135 thereby acts as a stabiliser by preventing overfitting. The variational formulation for (2.3)
136 reads as

$$137 \quad (2.6) \quad \min_{\substack{I \in \mathcal{X} \\ \phi_t \in \mathcal{G}}} \left\{ \int_0^1 \left[\mathcal{D}(\mathcal{T}_t(\phi_t \cdot I), g(t, \cdot)) + \mu_2 \mathcal{R}_2(\phi_t) \right] dt + \mu_1 \mathcal{R}_1(I) \right\} \quad \text{for fixed } \mu_1, \mu_2 > 0.$$

138 In the above, μ_1, μ_2 are the regularization parameters that must be chosen depending on
139 the noise level in data. Next, $\mathcal{D}: \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}_+$ above is the data discrepancy functional that
140 quantifies the mismatch in data space \mathcal{Y} . It is often taken as a suitable affine transform of
141 the negative data likelihood for the data, so minimising it amounts to computing a maximum
142 likelihood estimator. If one has data with additive Gaussian noise, then it is given by the
143 squared \mathcal{L}^2 -norm:

$$144 \quad (2.7) \quad \mathcal{D}(g_1, g_2) := \|g_1 - g_2\|_2^2 \quad \text{for } g_1, g_2 \in \mathcal{Y}.$$

145 Likewise, Poisson distributed data leads to the Kullback-Leibler (KL) divergence:

$$146 \quad (2.8) \quad \mathcal{D}(g_1, g_2) := \int_{\mathbb{M}} g_1(y) - g_2(y) \log(g_1(y)) dy \quad \text{for } g_1, g_2 \in \mathcal{Y}.$$

147 Moreover, the spatial regularization $\mathcal{R}_1: \mathcal{X} \rightarrow \mathbb{R}_+$ introduces stability by encoding priori
148 knowledge about the ground truth image. It is frequently based on a roughness prior given
149 as an \mathcal{L}^p -norm of the gradient magnitude or \mathcal{L}^1 -norm of a suitable sparse representation.
150 Typically, if $\mathcal{X} := \mathcal{H}^1(\Omega)$ then one takes the squared \mathcal{L}^2 -norm of the gradient magnitude:

$$151 \quad (2.9) \quad \mathcal{R}_1(f) := \|\nabla f\|_2^2$$

152 This choice is known to produce smooth images whereas selecting the \mathcal{L}^1 -norm of the gradient
 153 magnitude for $\mathcal{X} := \mathcal{BV}(\Omega)$ is better at preserving edges (total variation (TV) regularization)
 154 [45]:

$$155 \quad (2.10) \quad \mathcal{R}_1(f) := \|\nabla f\|_1$$

156 What remains is to describe how to generate the curve $t \mapsto \phi_t$ and to select the shape
 157 regularization $\mathcal{R}_2: \mathcal{G} \rightarrow \mathbb{R}_+$. An overall difficulty is that diffeomorphisms \mathcal{G} do not form
 158 a vector space. One option is to try re-parametrising the deformations using vector space
 159 elements. As we shall see next, this can be done within the LDDMM framework.

160 **3. A variational model for joint image reconstruction and motion estimation.** This
 161 section introduces a new variational model of the framework (2.6) based on LDDMM. We
 162 begin by recalling the basic principles of LDDMM.

163 **3.1. The LDDMM framework.** The LDDMM framework outlined here offers a powerful
 164 machinery for generating a flow of diffeomorphisms by means of a velocity field. One advantage
 165 is that the set of velocity fields form a vector space, so we have in this way re-parametrised
 166 deformations by vector space elements.

167 The idea in LDDMM is to consider a sequence of infinitesimally small vector field pertur-
 168 bations to the identity mapping. These vector fields can be seen as an instantaneous velocity
 169 field. Under certain regularity, the composition of such small deformations in the limit gener-
 170 ates a flow of diffeomorphisms given as the solution to an ordinary differential equation (ODE)
 171 [60]. More precisely, given a velocity field $\nu: [0, 1] \times \Omega \rightarrow \mathbb{R}^n$, a flow $t \mapsto \phi_t$ is generated by
 172 the following ODE:

$$173 \quad (3.1) \quad \begin{cases} \partial_t \phi_t(x) = \nu(t, \phi_t(x)) \\ \phi_0(x) = x \end{cases} \quad \text{for } x \in \Omega \text{ and } 0 \leq t \leq 1.$$

174 Note that $\phi_0 = \text{Id}$, i.e., the flow starts at the identity mapping. If the velocity field ν is
 175 sufficiently regular (see Definition 3.1), then the above ODE has a well-defined solution that
 176 is a diffeomorphism at each time point, i.e., (3.1) defines a flow of diffeomorphisms.

177 **Definition 3.1 (Admissible space [60]).** A Hilbert space $\mathcal{V} \subset \mathcal{C}_0^1(\Omega, \mathbb{R}^n)$ is admissible if it
 178 is (canonically) embedded in $\mathcal{C}_0^1(\Omega, \mathbb{R}^n)$ with the $\|\cdot\|_{1,\infty}$ norm, i.e., there exists a constant
 179 $C > 0$ such that

$$180 \quad \|\mathbf{v}\|_{1,\infty} \leq C \|\mathbf{v}\|_{\mathcal{V}} \quad \text{for all } \mathbf{v} \in \mathcal{V}.$$

181 In the above, $\|\mathbf{v}\|_{1,\infty} := \|\mathbf{v}\|_{\infty} + \|D\mathbf{v}\|_{\infty}$ for $\mathbf{v} \in \mathcal{C}_0^1(\Omega, \mathbb{R}^n)$.

182 We now consider velocity fields that are \mathcal{L}^p in time and in an admissible space at every point
 183 in time, i.e.,

$$184 \quad (3.2) \quad \mathcal{L}^p([0, 1], \mathcal{V}) := \left\{ \nu: \nu(t, \cdot) \in \mathcal{V} \text{ and } \|\nu\|_{\mathcal{L}^p([0,1],\mathcal{V})} < \infty \text{ for } 1 \leq p \leq \infty \right\}$$

185 with the associated norm

$$186 \quad \|\nu\|_{\mathcal{L}^p([0,1],\mathcal{V})} := \left(\int_0^1 \|\nu(t, \cdot)\|_{\mathcal{V}}^p dt \right)^{1/p}.$$

187 To simplify notation, let $\mathcal{L}_{\mathcal{V}}^p(\Omega)$ denote $\mathcal{L}^p([0, 1], \mathcal{V})$ and note in particular that $\mathcal{L}_{\mathcal{V}}^2(\Omega)$ is a
 188 Hilbert space with inner product

$$189 \quad \langle \boldsymbol{\nu}, \boldsymbol{\eta} \rangle_{\mathcal{L}_{\mathcal{V}}^2(\Omega)} = \int_0^1 \langle \boldsymbol{\nu}(t, \cdot), \boldsymbol{\eta}(t, \cdot) \rangle_{\mathcal{V}} dt \quad \text{for } \boldsymbol{\nu}, \boldsymbol{\eta} \in \mathcal{L}_{\mathcal{V}}^2(\Omega).$$

190 The following theorem states that admissible velocity field generates a flow of diffeomorphisms.

191 **Theorem 3.2 ([60, 11]).** *Let \mathcal{V} be an admissible Hilbert space and $\boldsymbol{\nu} \in \mathcal{L}_{\mathcal{V}}^2(\Omega)$ be a velocity*
 192 *field. Then the ODE in (3.1) admits a unique solution $\phi^{\boldsymbol{\nu}} \in \mathcal{C}_0^1([0, 1] \times \Omega, \Omega)$, such that for*
 193 *$t \in [0, 1]$, the mapping $\phi_t^{\boldsymbol{\nu}}: \Omega \rightarrow \Omega$ is a \mathcal{C}^1 -diffeomorphism on Ω .*

194 Hence, we define

$$195 \quad (3.3) \quad \mathcal{G}_{\mathcal{V}} := \left\{ \phi: \phi = \phi_{0,1}^{\boldsymbol{\nu}} \text{ for some } \boldsymbol{\nu} \in \mathcal{L}_{\mathcal{V}}^2(\Omega) \right\},$$

196 which becomes a sub-group of the group of diffeomorphisms by **Theorem 3.2** whenever \mathcal{V} is
 197 admissible. Next, if $\phi_t^{\boldsymbol{\nu}}$ denotes the solution to the ODE in (3.1) with given $\boldsymbol{\nu} \in \mathcal{L}_{\mathcal{V}}^2(\Omega)$, then

$$198 \quad (3.4) \quad \phi_{s,t}^{\boldsymbol{\nu}} := \phi_t^{\boldsymbol{\nu}} \circ (\phi_s^{\boldsymbol{\nu}})^{-1} \quad \text{for } 0 \leq t, s \leq 1$$

199 Furthermore, $\phi_0^{\boldsymbol{\nu}} = \text{Id}$ and by (3.4) we also get

$$200 \quad (3.5) \quad \phi_t^{\boldsymbol{\nu}} = \phi_{0,t}^{\boldsymbol{\nu}}, \quad (\phi_t^{\boldsymbol{\nu}})^{-1} = \phi_{t,0}^{\boldsymbol{\nu}}.$$

201 Next, several important properties about $\mathcal{G}_{\mathcal{V}}$ are summarised in the following theorem.

202 **Theorem 3.3 ([60, 11]).** *Let \mathcal{V} be an admissible Hilbert space, $\mathcal{G}_{\mathcal{V}}$ is defined as in (3.3),*
 203 *and $d_{\mathcal{G}_{\mathcal{V}}}: \mathcal{G}_{\mathcal{V}} \times \mathcal{G}_{\mathcal{V}} \rightarrow \mathbb{R}_+$ is defined as*

$$204 \quad (3.6) \quad d_{\mathcal{G}_{\mathcal{V}}}(\phi, \psi) := \inf_{\substack{\boldsymbol{\nu} \in \mathcal{L}_{\mathcal{V}}^2(\Omega) \\ \psi = \phi \circ \phi_{0,1}^{\boldsymbol{\nu}}}} \|\boldsymbol{\nu}\|_{\mathcal{L}_{\mathcal{V}}^2(\Omega)} \quad \text{for } \phi, \psi \in \mathcal{G}_{\mathcal{V}}.$$

205 Then $\mathcal{G}_{\mathcal{V}}$ is a group for the composition of functions and $\mathcal{G}_{\mathcal{V}}$ is a complete metric space
 206 under the metric $d_{\mathcal{G}_{\mathcal{V}}}$. Furthermore, for each $\phi, \psi \in \mathcal{G}_{\mathcal{V}}$, there exists $\boldsymbol{\nu} \in \mathcal{L}_{\mathcal{V}}^2(\Omega)$ satisfying
 207 $\psi = \phi \circ \phi_{0,1}^{\boldsymbol{\nu}}$, i.e., $d_{\mathcal{G}_{\mathcal{V}}}(\phi, \psi) = \|\boldsymbol{\nu}\|_{\mathcal{L}_{\mathcal{V}}^2(\Omega)}$.

208 The metric $d_{\mathcal{G}_{\mathcal{V}}}$ can now be used to define an energy term for regularizing the registration
 209 of a template image I_0 to a target image I_1 :

$$210 \quad (3.7) \quad \inf_{\phi \in \mathcal{G}_{\mathcal{V}}} \left\{ \|\phi \cdot I_0 - I_1\|_{\mathcal{L}^2(\Omega)}^2 + \mu d_{\mathcal{G}_{\mathcal{V}}}(\text{Id}, \phi)^2 \right\} \quad \text{for fix regularization parameter } \mu > 0.$$

211 One can show that the infimum in (3.6) is reached, so it can be replaced by a minimum. Next
 212 one can also show [60, Lemma 11.3] that (3.7) is equivalent to the following variational model
 213 that optimizes over velocity fields instead of diffeomorphisms:

$$214 \quad (3.8) \quad \min_{\boldsymbol{\nu} \in \mathcal{L}_{\mathcal{V}}^2(\Omega)} \left\{ \|\phi_{0,1}^{\boldsymbol{\nu}} \cdot I_0 - I_1\|_{\mathcal{L}^2(\Omega)}^2 + \mu \int_0^1 \|\boldsymbol{\nu}(t, \cdot)\|_{\mathcal{V}}^2 dt \right\} \quad \text{with } \phi_{0,1}^{\boldsymbol{\nu}} \in \mathcal{G}_{\mathcal{V}} \text{ as in (3.4).}$$

215 In conclusion, the regularization term for image registration by LDDMM is

$$216 \quad (3.9) \quad \mathcal{R}(\phi) := d_{\mathcal{G}_{\mathcal{V}}}(\text{Id}, \phi)^2 = \int_0^1 \|\widehat{\nu}(t, \cdot)\|_{\mathcal{V}}^2 dt,$$

217 where $\widehat{\nu}$ above minimizes $\nu \mapsto d_{\mathcal{G}_{\mathcal{V}}}(\text{Id}, \phi) = \|\nu\|_{\mathcal{L}_{\mathcal{V}}^2(\Omega)}$ where $\phi = \text{Id} \circ \phi_{0,1}^{\nu}$ (such a minimiser
218 exists due to [Theorem 3.3](#)).

219 A final remark concerns the choice of \mathcal{V} . If we choose it as a reproducing kernel Hilbert
220 space (RKHS) with a symmetric and positive-definite reproducing kernel, then \mathcal{V} becomes
221 an admissible Hilbert space [11]. Using such vector fields is also advantageous from a com-
222 putational point of view as shown in [17]. We will therefore assume that \mathcal{V} is henceforth an
223 admissible RKHS.

224 **3.2. Spatiotemporal reconstruction with LDDMM.** The aim here is to formulate a spe-
225 cial case of (2.6) for solving the spatiotemporal inverse problem in (2.3). Let ϕ_t in (2.6) be
226 generated by the flow equation (3.1) as in LDDMM, i.e., $\phi_t = \phi_{0,t}^{\nu}$ for some velocity field ν .
227 If $\nu \in \mathcal{L}_{\mathcal{V}}^2(\Omega)$, then ϕ_t^{ν} is a diffeomorphism on Ω by [Theorem 3.2](#). Consequently, combining
228 [Theorem 3.3](#) with (3.9) implies that a shape regularizer \mathcal{R}_2 for the temporal deformation given
229 by $\phi_{0,t}^{\nu}$ in (2.6) can be defined as

$$230 \quad (3.10) \quad \mathcal{R}_2(\phi_{0,t}^{\nu}) := \int_0^t \|\nu(\tau, \cdot)\|_{\mathcal{V}}^2 d\tau.$$

231 The variational formulation in (2.6) for solving the spatiotemporal inverse problem in (2.3)
232 now reads as

$$233 \quad (3.11) \quad \min_{\substack{I \in \mathcal{X} \\ \nu \in \mathcal{L}_{\mathcal{V}}^2(\Omega)}} \int_0^1 \left[\mathcal{D}(\mathcal{T}_t(\phi_{0,t}^{\nu} \cdot I), g(t, \cdot)) + \mu_2 \int_0^t \|\nu(\tau, \cdot)\|_{\mathcal{V}}^2 d\tau \right] dt + \mu_1 \mathcal{R}_1(I)$$

s.t. $\phi_{0,t}^{\nu}$ solves (3.1).

234 Note that (3.11) is an ODE constrained optimization problem, which henceforth is referred
235 to as *time-continuous spatiotemporal reconstruction with LDDMM*.

236 *Remark 3.4.* Changing the order of integration for the second term in (3.11) yields the
237 following equivalent formulation:

$$238 \quad (3.12) \quad \min_{\substack{I \in \mathcal{X} \\ \nu \in \mathcal{L}_{\mathcal{V}}^2(\Omega)}} \int_0^1 \left[\mathcal{D}(\mathcal{T}_t(\phi_{0,t}^{\nu} \cdot I), g(t, \cdot)) + \mu_2(1-t) \|\nu(t, \cdot)\|_{\mathcal{V}}^2 \right] dt + \mu_1 \mathcal{R}_1(I)$$

s.t. $\phi_{0,t}^{\nu}$ solves (3.1).

239 **3.2.1. Basic properties of optimal velocity fields.** The aim here is to characterize tem-
240 poral behaviour of velocity fields that solve (3.11) when \mathcal{V} is a RKHS. For this purpose, we
241 introduce some notation that will simplify the expressions:

$$242 \quad (3.13) \quad \mathcal{D}_{g_t}(f) := \mathcal{D}(\mathcal{T}_t(f), g(t, \cdot)) \quad \text{for } f \in \mathcal{X} \text{ with given } g(t, \cdot) \in \mathcal{Y}.$$

243 By [Theorem A.4](#) (in [Appendix A](#)), the optimal velocity field in [\(3.11\)](#) satisfies

$$244 \quad (3.14) \quad \boldsymbol{\nu}(t, \cdot) = \frac{1}{2\mu_2(1-t)} \int_t^1 \mathcal{K} \left(\nabla(\phi_{0,t}^\nu \cdot I) |D(\phi_{t,\tau}^\nu)| \nabla \mathcal{D}_{g_\tau}(\phi_{0,\tau}^\nu \cdot I)(\phi_{t,\tau}^\nu) \right) d\tau \quad \text{for } 0 \leq t < 1.$$

245 Here, $\mathcal{K}(\varphi) = \int_\Omega \mathbf{K}(\cdot, y)\varphi(y) dy$, with $\mathbf{K}: \Omega \times \Omega \rightarrow \mathbb{M}_+^{n \times n}$ denoting the reproducing kernel if
246 \mathcal{V} is a RKHS.

247 By the above, the optimal velocity field can be seen as a time-averaging. Next, $\boldsymbol{\nu}(1, \cdot)$ is
248 well-defined at $t = 1$ and reads as

$$249 \quad (3.15) \quad \boldsymbol{\nu}(1, \cdot) = \frac{1}{2\mu_2} \mathcal{K}(\nabla(\phi_{0,1}^\nu \cdot I) \nabla \mathcal{D}_{g_1}(\phi_{0,1}^\nu \cdot I)).$$

250 Similarly, by [\(A.9\)](#) and [\(A.7\)](#) we get

$$251 \quad (3.16) \quad \boldsymbol{\nu}(0, \cdot) = \frac{1}{2\mu_2} \mathcal{K} \left(\nabla I \int_0^1 |D(\phi_{0,t}^\nu)| \nabla \mathcal{D}_{g_t}(\phi_{0,t}^\nu \cdot I)(\phi_{0,t}^\nu) dt \right).$$

252 Hence, the optimal velocity field in [\(3.11\)](#) is *averagely distributed w.r.t. time t* , and *non-*
253 *vanishing at both the initial and the end time points.*

254 **3.2.2. Control theoretic formulation.** The aim here is to state an equivalent optimal
255 control formulation of [\(3.11\)](#) in terms of a partial differential equation (PDE) constrained
256 optimization problem. This makes it easier to compare our proposed approach in [\(3.11\)](#)
257 against PDE based approaches for registration, as we do in [subsection 3.3](#).

258 **Theorem 3.5.** *Let \mathcal{X} be a space of real valued functions that are sufficiently smooth, e.g.,*
259 *the space of functions with bounded variation given in [subsection 3.3.1](#). Assume next that*
260 *$I \in \mathcal{X}$ and define $f: [0, 1] \times \Omega \rightarrow \mathbb{R}$ as*

$$261 \quad (3.17) \quad f(t, \cdot) := \phi_{0,t}^\nu \cdot I \quad \text{for } 0 \leq t \leq 1 \text{ with } \phi_{0,t}^\nu \text{ given by } \a href="#">(3.4).$$

262 *Assume furthermore that $f(t, \cdot) \in \mathcal{X}$ for any $0 \leq t \leq 1$. Then, [\(3.11\)](#) is under the group*
263 *action in [\(2.4\)](#) (geometric deformation) equivalent to*

$$264 \quad (3.18) \quad \min_{\substack{f(0, \cdot) \in \mathcal{X} \\ \boldsymbol{\nu} \in \mathcal{L}_\nu^2(\Omega)}} \left\{ \int_0^1 \left[\mathcal{D}(\mathcal{T}_t(f(t, \cdot)), g(t, \cdot)) + \mu_2 \int_0^t \|\boldsymbol{\nu}(\tau, \cdot)\|_{\mathcal{V}}^2 d\tau \right] dt + \mu_1 \mathcal{R}_1(f(0, \cdot)) \right\}$$

$$\text{s.t. } \partial_t f(t, \cdot) + \langle \nabla f(t, \cdot), \boldsymbol{\nu}(t, \cdot) \rangle_{\mathbb{R}^n} = 0.$$

265 *Likewise, [\(3.11\)](#) is under the group action in [\(2.5\)](#) (mass-preserving deformation) equivalent*
266 *to*

$$267 \quad (3.19) \quad \min_{\substack{f(0, \cdot) \in \mathcal{X} \\ \boldsymbol{\nu} \in \mathcal{L}_\nu^2(\Omega)}} \left\{ \int_0^1 \left[\mathcal{D}(\mathcal{T}_t(f(t, \cdot)), g(t, \cdot)) + \mu_2 \int_0^t \|\boldsymbol{\nu}(\tau, \cdot)\|_{\mathcal{V}}^2 d\tau \right] dt + \mu_1 \mathcal{R}_1(f(0, \cdot)) \right\}$$

$$\text{s.t. } \partial_t f(t, \cdot) + \nabla \cdot (f(t, \cdot) \boldsymbol{\nu}(t, \cdot)) = 0.$$

268 *Proof.* First we consider the geometric deformation in (2.4), so (3.17) reads as

$$269 \quad (3.20) \quad f(t, \cdot) = I \circ (\phi_{0,t}^\nu)^{-1} \quad \text{for } 0 \leq t \leq 1.$$

270 Obviously, $f(0, \cdot) = I$, $f(1, \cdot) = I \circ (\phi_{0,1}^\nu)^{-1}$, and

$$271 \quad (3.21) \quad f(t, \phi_{0,t}^\nu) = I \quad \text{for } 0 \leq t \leq 1.$$

272 Differentiating (3.21) w.r.t. time t leads to

$$273 \quad \partial_t f(t, \phi_{0,t}^\nu) + \langle \nabla f(t, \phi_{0,t}^\nu), \nu(t, \phi_{0,t}^\nu) \rangle_{\mathbb{R}^n} = 0.$$

274 The above is the PDE constraint in (3.18), so a solution to (3.11) also solves (3.18).

275 We now consider the reverse implication, i.e., show that a solution to (3.18) also solves
276 (3.11). Suppose that f and ν solve (3.18). Define the diffeomorphism ψ_t that solves the ODE
277 in (3.1) with the above given ν . Since f satisfies PDE constraint in (3.18), we get

$$278 \quad \frac{d}{dt} f(t, \psi_t) = \partial_t f(t, \psi_t) + \langle \nabla f(t, \psi_t), \nu(t, \psi_t) \rangle_{\mathbb{R}^n} = 0.$$

279 Hence, $t \mapsto f(t, \psi_t)$ is constant and in particular we have

$$280 \quad f(t, \psi_t) \equiv f(0, \psi_0) = f(0, \cdot).$$

281 If $f(0, \cdot) := I$ and $\psi_t := \phi_{0,t}^\nu$, then $f(t, \cdot) = I \circ (\phi_{0,t}^\nu)^{-1}$. Hence a solution to (3.18) also solves
282 (3.11).

283 Using a mass-preserving deformation (2.5) as group action in (3.17) results in

$$284 \quad (3.22) \quad f(t, \cdot) = |D((\phi_{0,t}^\nu)^{-1})| I \circ (\phi_{0,t}^\nu)^{-1} \quad \text{for } 0 \leq t \leq 1.$$

285 We then get that $f(0, \cdot) = I$ and $f(1, \cdot) = |D((\phi_{0,1}^\nu)^{-1})| I \circ (\phi_{0,1}^\nu)^{-1}$. The symmetry of the
286 mass-preserving property furthermore yields

$$287 \quad (3.23) \quad |D(\phi_{0,t}^\nu)| f(t, \cdot) \circ \phi_{0,t}^\nu = I \quad \text{for } 0 \leq t \leq 1.$$

288 Finally, differentiating (3.23) w.r.t. t leads to the constraint in (3.19). Hence, a minimizer of
289 (3.11) with the group action given by (2.5) is also a minimizer of (3.19). Similar to the case
290 of geometric deformation, it is not difficult to prove the reverse implication. ■

291 **3.3. Comparison with existing approaches.** Here we compare (3.11) against several ex-
292 isting approaches (i.e., optical flow based model, diffeomorphic motion models).

293 **3.3.1. Optical flow based models.** Recently, [13] proposes an optical flow based varia-
294 tional model (joint TV-TV optical flow) for joint motion estimation and image reconstruction
295 in spatiotemporal imaging. The model is formulated as a PDE-constrained optimal control
296 problem where the constraint is given by a brightness constancy equation. When applied to
297 the spatiotemporal inverse problem in (2.3), it reads as

$$298 \quad (3.24) \quad \min_{\substack{f(t, \cdot) \in \mathcal{X} \\ \nu(t, \cdot) \in \mathcal{BV}(\Omega)}} \int_0^1 \left[\mathcal{D}(\mathcal{T}_t(f(t, \cdot)), g(t, \cdot)) + \mu_1 \mathcal{R}_1(f(t, \cdot)) + \mu_2 |\nu(t, \cdot)|_{\mathcal{BV}} \right] dt$$

$$\text{s.t. } \partial_t f(t, \cdot) + \langle \nabla f(t, \cdot), \nu(t, \cdot) \rangle_{\mathbb{R}^n} = 0,$$

299 where $|\cdot|_{\mathcal{BV}}$ is the semi-norm on the space of functions with bounded variation:

$$300 \quad \mathcal{BV}(\Omega) := \{u \in \mathcal{L}^1(\Omega) : |u|_{\mathcal{BV}} < \infty\} \quad \text{with} \quad |u|_{\mathcal{BV}} := \sup_{\substack{\varphi \in \mathcal{C}_0^1(\Omega, \mathbb{R}^n) \\ \|\varphi\|_\infty \leq 1}} \int_{\Omega} u(x) \operatorname{div} \varphi(x) \, dx.$$

301 In particular, $|\nu(t, \cdot)|_{\mathcal{BV}}$ denotes the sum of the semi-norm of all the elements in $\nu(t, \cdot)$ [1].

302 We will use the reformulation in (3.18) to compare the above optical flow based model
 303 from [13] with our proposed model. The use of the brightness constancy equation points to
 304 using the geometric deformation in (2.4) as a group action in (3.17), i.e., we assume (3.20)
 305 holds. It is easy to see that the constraints in (3.18) and (3.24) are equivalent. Hence, the
 306 optical flow constraint given by the brightness constancy equation in [13] is equivalent to using
 307 diffeomorphisms generated by LDDMM that act through geometric deformation.

308 By comparison, the primary distinction between (3.18) and (3.24) relates to the selection
 309 of the regularization term w.r.t. vector field $\nu(t, \cdot)$. In the model (3.24), one uses the TV
 310 functional, so the space of vector fields is assumed to be in $\mathcal{BV}(\Omega)$, which allows for a vector
 311 field that is a piecewise-constant vector-valued function distributed on Ω . In contrast, in the
 312 model (3.18), the space of vector fields reside in an admissible Hilbert space. Hence, the
 313 vector field is a sufficiently smooth vector-valued function distributed on Ω . This guarantees
 314 an elastic diffeomorphic deformation, which to some extent mimics the underlying physical
 315 mechanisms of organ motion [30, 14].

316 In addition to the above, both approaches also differ in the selection of regularization
 317 term \mathcal{R}_1 . In (3.18) one only poses restriction on the initial image $f(0, \cdot)$, whereas in (3.24)
 318 the whole time trajectory $t \mapsto f(t, \cdot)$ is regularized. The treatment in (3.18) is reasonable
 319 to some extent, since regularity properties are assumed to be preserved. More precisely, $f(t, \cdot)$
 320 is contained in the same space for all t . For example, if $f(0, \cdot) \in \mathcal{BV}(\Omega)$, then there exists
 321 a unique weak solution $f(t, \cdot) = f(0, \cdot) \circ (\phi_{0,t}^\nu)^{-1}$ to the PDE constraint of (3.18) is \mathcal{L}^∞ in
 322 time and in $\mathcal{BV}(\Omega)$ at any t [21, Theorem 4]. Hence, (3.18) has a simpler structure which is
 323 also beneficial in implementation.

324 **3.3.2. Diffeomorphic motion models.** A diffeomorphic motion model for 4D CT image
 325 reconstruction was proposed in [33]. It is based on the LDDMM growth model [31] with a
 326 time-continuous model that reads as

$$327 \quad (3.25) \quad \min_{\substack{I \in \mathcal{I} \\ \nu \in \mathcal{L}_{\mathcal{V}}^2(\Omega)}} \int_0^1 \left[\mathcal{D}(\mathcal{T}_t(\phi_{0,t}^\nu \cdot I), g(t, \cdot)) + \mu_2 \|\nu(t, \cdot)\|_{\mathcal{V}}^2 \right] dt$$

s.t. $\phi_{0,t}^\nu$ solves (3.1).

328 Compared to (3.11), the above has no regularization term \mathcal{R}_1 . Another difference relates to
 329 the selection on the shape regularization \mathcal{R}_2 . In (3.25) one uniformly weights the $\|\nu(t, \cdot)\|_{\mathcal{V}}^2$
 330 term. In contrast, (3.11) uses a non-uniformly weighting (see (3.12) for more clarity) that
 331 puts more weight on the previous time instance.

332 *Remark 3.6.* Note that in (3.11), we regularize the velocity field more at the beginning,
 333 and which is relevant because the template is selected at the initial time $t = 0$, then the

334 velocity field at time t_i has influence on the geodesic trajectory at $t \geq t_i$. Equivalently, the
 335 geodesic trajectory at time t_i depends on the velocity field at $t \leq t_i$. The earlier on we
 336 are at the velocity field, the more influence one has on the entire geodesic trajectory. Since
 337 the optimal velocity field (3.14) is averagely distributed w.r.t. time t (subsection 3.2.1), the
 338 motion of the object is close to a uniform speed. This can also result in a velocity field that
 339 is non-vanishing at the end time points (see (3.15) and (3.16)).

340 Next, by Theorem A.4 it is easy to see that the $\mathcal{L}_{\mathcal{V}}^2(\Omega)$ -norm minimizer of (3.25) satisfies

$$341 \quad (3.26) \quad \boldsymbol{\nu}(t, \cdot) = \frac{1}{2\mu_2} \int_t^1 \mathcal{K} \left(\nabla(\phi_{0,t}^{\boldsymbol{\nu}} \cdot I) |D(\phi_{t,\tau}^{\boldsymbol{\nu}})| \nabla \mathcal{D}_{g_\tau}(\phi_{0,\tau}^{\boldsymbol{\nu}} \cdot I)(\phi_{t,\tau}^{\boldsymbol{\nu}}) \right) d\tau \quad \text{for } 0 \leq t \leq 1.$$

342 It is not difficult to see from (3.26) that the optimal velocity field is *not averagely distributed*
 343 *w.r.t. time t* . In addition, the minimizer w.r.t. variations of the template satisfies

$$344 \quad (3.27) \quad \int_0^1 |D(\phi_{0,t}^{\boldsymbol{\nu}})| \nabla \mathcal{D}_{g_t}(\phi_{0,t}^{\boldsymbol{\nu}} \cdot I)(\phi_{0,t}^{\boldsymbol{\nu}}) dt = 0.$$

345 Combining (3.26) and (3.27) yields $\boldsymbol{\nu}(0, \cdot) = \boldsymbol{\nu}(1, \cdot) = 0$, i.e., the optimal velocity field accord-
 346 ing to the model (3.25) *vanishes at both the initial and end time points*.

347 To summarise, both (3.11) and the approach taken in [33] address joint reconstruction and
 348 motion estimation. Both approaches model the latter as diffeomorphic deformations generated
 349 by velocity fields within the LDDMM framework. A difference is that the regularization of
 350 the velocity field in [33] is equally weighted over the entire time trajectory. Next, the optimal
 351 velocity field vanishes at both the initial and end time points. This is not the case with the
 352 model in (3.11).

353 A variant to (3.11) based on (2.6) would regularise the template I (image at $t = 0$) instead
 354 of its entire time evolution $\phi_{0,t}^{\boldsymbol{\nu}} \cdot I$. This leads to

$$355 \quad (3.28) \quad \min_{\substack{I \in \mathcal{X} \\ \boldsymbol{\nu} \in \mathcal{L}_{\mathcal{V}}^2(\Omega)}} \int_0^1 \left[\mathcal{D}(\mathcal{T}_t(\phi_{0,t}^{\boldsymbol{\nu}} \cdot I), g(t, \cdot)) + \mu_1 \mathcal{R}_1(\phi_{0,t}^{\boldsymbol{\nu}} \cdot I) + \mu_2 \int_0^t \|\boldsymbol{\nu}(\tau, \cdot)\|_{\mathcal{V}}^2 d\tau \right] dt$$

s.t. $\phi_{0,t}^{\boldsymbol{\nu}}$ solves (3.1).

356 By Theorem A.4, the optimal velocity field for $0 \leq t < 1$ satisfies

$$357 \quad (3.29) \quad \boldsymbol{\nu}(t, \cdot) = \frac{1}{2\mu_2(1-t)} \int_t^1 \mathcal{K} \left(\nabla(\phi_{0,t}^{\boldsymbol{\nu}} \cdot I) |D(\phi_{t,\tau}^{\boldsymbol{\nu}})| \nabla \mathcal{S}_{g_\tau}(\phi_{0,\tau}^{\boldsymbol{\nu}} \cdot I)(\phi_{t,\tau}^{\boldsymbol{\nu}}) \right) d\tau$$

358 where $\mathcal{S}_{g_t}(f) := \mathcal{D}(\mathcal{T}_t(f), g(t, \cdot)) + \mu_1 \mathcal{R}_1(f)$ for fixed $g(t, \cdot) \in \mathcal{G}$ and $f \in \mathcal{X}$. The corre-
 359 sponding optimal template satisfies

$$360 \quad (3.30) \quad \int_0^1 |D(\phi_{0,t}^{\boldsymbol{\nu}})| \nabla \mathcal{S}_{g_t}(\phi_{0,t}^{\boldsymbol{\nu}} \cdot I)(\phi_{0,t}^{\boldsymbol{\nu}}) dt = 0.$$

361 Evidently, the above optimal velocity field is also a time average of the integrand. Next,
 362 $\boldsymbol{\nu}(1, \cdot)$ is well-defined at $t = 1$ and reads as

$$363 \quad \boldsymbol{\nu}(1, \cdot) = \frac{1}{2\mu_2} \mathcal{K}(\nabla(\phi_{0,1}^{\boldsymbol{\nu}} \cdot I) \nabla \mathcal{S}_{g_1}(\phi_{0,1}^{\boldsymbol{\nu}} \cdot I)).$$

364 However, by (3.29) and (3.30) at $t = 0$ we have $\boldsymbol{\nu}(0, \cdot) = 0$. Hence, the optimal velocity field in
 365 (3.28) is *averagely distributed w.r.t. time t , but vanishing at the initial time point*. In addition,
 366 as stated in subsection 3.3.1, we only need to regularize I rather than $\phi_{0,t}^\nu \cdot I$ to some extent,
 367 since I and its time evolution reside in the same space. This proposed model has a simpler
 368 structure to implement.

369 To summarise, the comparative analysis points to several advantages that come with using
 370 (3.11) over alternative approaches.

371 **4. Numerical implementation.** The first part of the numerical implementation is to derive
 372 a time-discretized formulation of (3.11).

373 **4.1. Time-discretized formulation.** Assume the time sampling of data is uniform, i.e.,
 374 there is a uniform partition of $[0, 1]$ and data is acquired at time points $\{t_i\}_{i=0}^N$ with $t_i = i/N$
 375 for $0 \leq i \leq N$. We refer to this as the *gating grid* and the time-discretized formulation of the
 376 general spatiotemporal inverse problem in (2.1) is the task to recover $t \mapsto f(t, \cdot)$ from data
 377 $g(t_i, \cdot) \in \mathcal{Y}$ where

$$378 \quad (4.1) \quad g(t_i, \cdot) = \mathcal{T}_{t_i}(f(t_i, \cdot)) + g_{\text{noise}}(t_i, \cdot) \quad \text{for } i = 1, \dots, N.$$

379 Then, a time-discretized version of (3.11) reads as

$$380 \quad (4.2) \quad \min_{\substack{I \in \mathcal{X} \\ \boldsymbol{\nu} \in \mathcal{L}_{\mathcal{Y}}^2(\Omega)}} \left\{ \frac{1}{N} \sum_{i=1}^N \left[\mathcal{D}(\mathcal{T}_{t_i}(\phi_{0,t_i}^\nu \cdot I), g(t_i, \cdot)) + \mu_2 \int_0^{t_i} \|\boldsymbol{\nu}(\tau, \cdot)\|_{\mathcal{Y}}^2 d\tau \right] + \mu_1 \mathcal{R}_1(I) \right\}$$

381 s.t. $\phi_{0,t}^\nu$ solves (3.1).

382 *Remark 4.1.* The time-discretized version (4.2) can be also written such that the image
 383 in the first gate is the template:

$$384 \quad \min_{\substack{I \in \mathcal{X} \\ \boldsymbol{\nu} \in \mathcal{L}_{\mathcal{Y}}^2(\Omega)}} \left\{ \frac{1}{N+1} \sum_{i=0}^N \left[\mathcal{D}(\mathcal{T}_{t_i}(\phi_{0,t_i}^\nu \cdot I), g(t_i, \cdot)) + \mu_2 \int_0^{t_i} \|\boldsymbol{\nu}(\tau, \cdot)\|_{\mathcal{Y}}^2 d\tau \right] + \mu_1 \mathcal{R}_1(I) \right\}$$

385 s.t. $\phi_{0,t}^\nu$ solves (3.1).

385 Since (4.2) contains highly coupled arguments, it is difficult to jointly solve for the template
 386 I and the velocity field $\boldsymbol{\nu}$. An approach that is computationally more feasible is to compute I
 387 and $\boldsymbol{\nu}$ through an intertwined iterative scheme. More precisely, if the velocity field $\boldsymbol{\nu}$ is given
 388 then the spatiotemporal reconstruction problem (4.2) reduces to a static image reconstruction
 389 problem: Conversely, if the template I is fixed, then (4.2) reduces to a sequentially indirect
 390 image registration problem where we seek the velocity field $\boldsymbol{\nu}$ from time-series data that are
 391 indirect observations of the target. Formalising the above, we try to solve (4.2) through the
 392 following intertwined iterative scheme $(I^k, \boldsymbol{\nu}^k) \in \mathcal{X} \times \mathcal{L}_{\mathcal{Y}}^2(\Omega)$:

$$393 \quad (4.3) \quad \begin{cases} I^k \in \arg \min_{I \in \mathcal{X}} \mathcal{J}_{\boldsymbol{\nu}^{k-1}}(I) \\ \boldsymbol{\nu}^k \in \arg \min_{\boldsymbol{\nu} \in \mathcal{L}_{\mathcal{Y}}^2(\Omega)} \mathcal{E}_{I^k}(\boldsymbol{\nu}) \end{cases}$$

394 where the functionals $\mathcal{J}_\nu: \mathcal{X} \rightarrow \mathbb{R}$, for given $\nu \in \mathcal{L}_\gamma^2(\Omega)$, and $\mathcal{E}_I: \mathcal{L}_\gamma^2(\Omega) \rightarrow \mathbb{R}$, for given
395 $I \in \mathcal{X}$, are defined as

$$396 \quad (4.4) \quad \mathcal{J}_\nu(I) := \frac{1}{N} \sum_{i=1}^N \mathcal{D}(\mathcal{T}_{t_i}(\phi_{0,t_i}^{\nu^{k-1}} \cdot I), g(t_i, \cdot)) + \mu_1 \mathcal{R}_1(I)$$

$$397 \quad (4.5) \quad \mathcal{E}_I(\nu) := \frac{1}{N} \sum_{i=1}^N \left[\mathcal{D}(\mathcal{T}_{t_i}(\phi_{0,t_i}^\nu \cdot I_k), g(t_i, \cdot)) + \mu_2 \int_0^{t_i} \|\nu(\tau, \cdot)\|_\gamma^2 d\tau \right]$$

$$398 \quad \text{s.t. } \phi_{0,t}^\nu \text{ solves (3.1).}$$

400 To proceed, we consider the special case when the data discrepancy term is the squared
401 \mathcal{L}^2 -norm as in (2.7) and the spatial regularization is the TV functional as in (2.10). Then
402 (4.2) reads as

$$403 \quad (4.6) \quad \min_{\substack{I \in \mathcal{X} \\ \nu \in \mathcal{L}_\gamma^2(\Omega)}} \left\{ \frac{1}{N} \sum_{i=1}^N \left[\|\mathcal{T}_{t_i}(\phi_{0,t_i}^\nu \cdot I) - g(t_i, \cdot)\|_2^2 + \mu_2 \int_0^{t_i} \|\nu(\tau, \cdot)\|_\gamma^2 d\tau \right] + \mu_1 \|\nabla I\|_1 \right\}$$

404 s.t. $\phi_{0,t}^\nu$ solves (3.1).

404 Correspondingly, (4.5) become

$$405 \quad (4.7) \quad \mathcal{J}_\nu(I) = \frac{1}{N} \sum_{i=1}^N \left\| \mathcal{T}_{t_i}(\phi_{0,t_i}^\nu \cdot I) - g(t_i, \cdot) \right\|_2^2 + \mu_1 \|\nabla I\|_1$$

$$406 \quad (4.8) \quad \mathcal{E}_I(\nu) = \frac{1}{N} \sum_{i=1}^N \left[\left\| \mathcal{T}_{t_i}(\phi_{0,t_i}^\nu \cdot I) - g(t_i, \cdot) \right\|_2^2 + \mu_2 \int_0^{t_i} \|\nu(\tau, \cdot)\|_\gamma^2 d\tau \right]$$

407

408 Inserting the above into (4.3) yields the following intertwined scheme for solving (4.6):

$$409 \quad (4.9) \quad \begin{cases} I^k := \arg \min_{I \in \mathcal{X}} \mathcal{J}_{\nu^{k-1}}(I) & \text{with } \mathcal{J}_{\nu^{k-1}} \text{ given by (4.7)} \\ \nu^k := \arg \min_{\nu \in \mathcal{L}_\gamma^2(\Omega)} \mathcal{E}_{I^k}(\nu) & \text{with } \mathcal{E}_{I^k} \text{ given by (4.8).} \end{cases}$$

410 **4.2. Template reconstruction.** We henceforth consider geometric deformations where the
411 deformation operator is given as $\phi_{0,t} \cdot I := I \circ \phi_{0,t}^{-1} = I \circ \phi_{t,0}$. This is a common choice for
412 image registration [7, 9, 36, 13].

413 Next, consider the problem of computing a minimiser to $\mathcal{J}_{\nu^{k-1}}$ (static reconstruction) in
414 (4.7). This is a non-smooth TV- ℓ_2 minimization problem that is convex when the forward
415 operator $\mathcal{T}_{t_i}: \mathcal{X} \rightarrow \mathcal{Y}$ is linear. Solving it by a first order method, like a gradient descent
416 scheme as outlined in Algorithm 4.1, requires smoothing the TV-component, i.e., we seek to
417 solve

$$418 \quad (4.10) \quad \min_{I \in \mathcal{X}} \left\{ \frac{1}{N} \sum_{i=1}^N \left\| \mathcal{T}_{t_i}(I \circ \phi_{t_i,0}^\nu) - g(t_i, \cdot) \right\|_2^2 + \mu_1 \int_\Omega |\nabla I(x)|_{2,\epsilon} dx \right\},$$

419 where $|\nabla I(x)|_{2,\epsilon} := \sqrt{\sum_i (\partial_i I(x))^2 + \epsilon}$ with $\epsilon > 0$ small, e.g., $\epsilon = 10^{-12}$. This is a frequently
 420 used modification of the TV functional in imaging applications [50, 18, 16]. A gradient descent
 421 scheme for (4.10) assuming a linear \mathcal{T}_{t_i} reads as

422

$$423 \quad (4.11) \quad I^{k+1} = I^k - \alpha^k \left(\frac{2}{N} \sum_{i=1}^N |D(\phi_{0,t_i}^\nu)| \mathcal{T}_{t_i}^* \left(\mathcal{T}_{t_i}(I^k \circ \phi_{t_i,0}^\nu) - g(t_i, \cdot) \right) (\phi_{0,t_i}^\nu) \right. \\ \left. + \mu_1 \nabla^* \left(\frac{\nabla I^k}{|\nabla I^k|_{2,\epsilon}} \right) \right).$$

424

425

426 Here, α^k is the stepsize for the k -th iteration and $\mathcal{T}_{t_i}^*$ is the adjoint of \mathcal{T}_{t_i} . The iterative scheme
 427 (4.11) is used for updating the template in the intertwined scheme (4.3) for solving (4.6).

428 *Remark 4.2.* There are several optimization techniques for solving convex non-smooth
 429 problems, like minimising $I \mapsto \mathcal{J}_\nu(I)$ in (4.7). Proximal gradient methods [49, 19, 42, 20, 4]
 430 are an important class of methods that are suitable for solving non-differentiable convex
 431 optimization problems, so they can be applied for solving (4.7). However, evaluating the
 432 proximal operator of a function itself involves solving a small convex optimization problem. For
 433 this reason, these more advanced methods introduce further auxiliary variables/parameters
 434 and they come with slower convergence rates.

435 **4.2.1. Computing diffeomorphic deformations.** Updating the template requires access
 436 to diffeomorphic deformations $\phi_{t_i,0}^\nu$ and ϕ_{0,t_i}^ν for $1 \leq i \leq N$. By definition, $\phi_{s,t}^\nu$ solves the flow
 437 equation

$$438 \quad (4.12) \quad \begin{cases} \partial_t \varphi(t, x) = \nu(t, \varphi(t, x)) \\ \varphi(s, x) = x \end{cases} \quad \text{for } x \in \Omega \text{ and } 0 \leq s, t \leq 1,$$

439 where s is a fixed time point. Integrating w.r.t. time t in (4.12) yields

$$440 \quad (4.13) \quad \phi_{s,t}^\nu = \text{Id} + \int_s^t \nu(\tau, \phi_{s,\tau}^\nu) d\tau \quad \text{for } 0 \leq t \leq 1.$$

441 The time interval $[0, 1]$ is subdivided uniformly into MN parts thereby forming a dis-
 442 cretized time grid that is given as $\tau_j = j/(MN)$ for $j = 0, 1, \dots, MN$. Evidently, $\tau_{iM} = t_i$ for
 443 $i = 0, 1, \dots, N$, so each subinterval $[t_i, t_{i+1}]$ is segmented into M even parts. The M is named
 444 as a factor of discretized time degree. If $M = 1$, then $\tau_i = t_i$, implying that the discretized
 445 time grid is consistent with the gating grid. Note also that the different subintervals of the
 446 gating grid can be discretized adaptively according to the degree of motions.

447 Within a short-time interval one can approximate the diffeomorphic deformation with
 448 linearized deformations [43]. More precisely, let $s = \tau_j$, $t = \tau_{j+1}$ and τ_{j+1} in (4.13). Then, the
 449 small deformations $\phi_{\tau_i, \tau_{i-1}}^\nu$ and $\phi_{\tau_i, \tau_{i+1}}^\nu$ can be approximated by

$$450 \quad (4.14) \quad \phi_{\tau_j, \tau_{j-1}}^\nu \approx \text{Id} - \frac{1}{MN} \nu(\tau_j, \cdot), \quad \text{and} \quad \phi_{\tau_j, \tau_{j+1}}^\nu \approx \text{Id} + \frac{1}{MN} \nu(\tau_j, \cdot).$$

451 Moreover, by (3.4) we get $\phi_{\tau_j,0}^{\nu} = \phi_{\tau_{j-1},0}^{\nu} \circ \phi_{\tau_j,\tau_{j-1}}^{\nu}$, which combined with (4.14) yields

$$452 \quad (4.15) \quad \phi_{\tau_j,0}^{\nu} \approx \phi_{\tau_{j-1},0}^{\nu} \circ \left(\text{Id} - \frac{1}{MN} \nu(\tau_j, \cdot) \right) \quad \text{for } j = 1, 2, \dots, MN.$$

453 This yields the following estimate for $I \circ \phi_{\tau_j,0}^{\nu}$:

$$454 \quad (4.16) \quad I \circ \phi_{\tau_j,0}^{\nu} \approx (I \circ \phi_{\tau_{j-1},0}^{\nu}) \circ \left(\text{Id} - \frac{1}{MN} \nu(\tau_j, \cdot) \right) \quad \text{for } j = 1, 2, \dots, MN$$

455 and $I \circ \phi_{\tau_0,0}^{\nu} = I$. Similarly, (3.4) also implies $\phi_{\tau_j,t_i}^{\nu} = \phi_{\tau_{j+1},t_i}^{\nu} \circ \phi_{\tau_j,\tau_{j+1}}^{\nu}$ for $i \geq 1$, which
456 combined with (4.14) gives the following approximation:

$$457 \quad (4.17) \quad \phi_{\tau_j,t_i}^{\nu} \approx \phi_{\tau_{j+1},t_i}^{\nu} \circ \left(\text{Id} + \frac{1}{MN} \nu(\tau_j, \cdot) \right) \quad \text{for } j = iM - 1, iM - 2, \dots, 0$$

458 and $\phi_{t_i,t_i}^{\nu} = \text{Id}$. To summarize, the deformation between two images at adjacent points in the
459 discretized time grid is approximately represented by a linearized deformation.

460 **4.2.2. Computing mass-preserving deformations.** Expression for the gradient of the data
461 discrepancy term for the mass-preserving group action in (2.5) involves terms of the type

$$462 \quad |D(\phi_{0,t_i}^{\nu})| \mathcal{T}_{t_i}^* \left(\mathcal{T}_{t_i}(I \circ \phi_{t_i,0}^{\nu}) - g(t_i, \cdot) \right) \circ \phi_{0,t_i}^{\nu} \quad \text{for } i \geq 1.$$

463 Starting with the Jacobian determinant, by (4.17) we get

$$464 \quad (4.18) \quad |D(\phi_{\tau_j,t_i}^{\nu})| \approx \left(1 + \frac{1}{MN} \text{div} \nu(\tau_j, \cdot) \right) |D(\phi_{\tau_{j+1},t_i}^{\nu})| \circ \left(\text{Id} + \frac{1}{MN} \nu(\tau_j, \cdot) \right)$$

465 for $j = iM - 1, iM - 2, \dots, 0$ and where $|D(\phi_{t_i,t_i}^{\nu})| = 1$. Next, (4.17) also yields the following
466 approximation:

$$467 \quad (4.19) \quad \mathcal{T}_{t_i}^* \left(\mathcal{T}_{t_i}(I \circ \phi_{t_i,0}^{\nu}) - g(t_i, \cdot) \right) \circ \phi_{\tau_j,t_i}^{\nu} \\ 468 \quad \approx \mathcal{T}_{t_i}^* \left(\mathcal{T}_{t_i}(I \circ \phi_{t_i,0}^{\nu}) - g(t_i, \cdot) \right) \circ \phi_{\tau_{j+1},t_i}^{\nu} \circ \left(\text{Id} + \frac{1}{MN} \nu(\tau_j, \cdot) \right)$$

471 for $j = iM - 1, iM - 2, \dots, 0$. For simplicity, let

$$472 \quad (4.20) \quad \eta_{\tau,t}^{I,\nu} = |D(\phi_{\tau,t}^{\nu})| \mathcal{T}_t^* \left(\mathcal{T}_t(I \circ \phi_{t,0}^{\nu}) - g(t, \cdot) \right) \circ \phi_{\tau,t}^{\nu}.$$

473 Then multiplying (4.18) by (4.19), and using (4.20), $\eta_{0,t_i}^{I,\nu}$ for $i \geq 1$ is computed by

$$474 \quad (4.21) \quad \eta_{\tau_j,t_i}^{I,\nu} \approx \left(1 + \frac{1}{MN} \text{div} \nu(\tau_j, \cdot) \right) \eta_{\tau_{j+1},t_i}^{I,\nu} \circ \left(\text{Id} + \frac{1}{MN} \nu(\tau_j, \cdot) \right)$$

475 for $j = iM - 1, iM - 2, \dots, 0$ with $\eta_{t_i,t_i}^{I,\nu} = \mathcal{T}_{t_i}^* \left(\mathcal{T}_{t_i}(I \circ \phi_{t_i,0}^{\nu}) - g(t_i, \cdot) \right)$.

476 Based on the above derivations, the concrete implementation is given as the gradient
477 descent scheme in Algorithm 4.1.

Algorithm 4.1 Gradient descent scheme for minimizing $I \mapsto \mathcal{J}_\nu(I)$ in (4.7)

- 1: *Initialize:*
- 2: $k \leftarrow 0$.
- 3: $t_i \leftarrow \frac{i}{N}$ for $i = 0, 1, \dots, N$.
- 4: $\tau_j \leftarrow \frac{j}{MN}$ for $j = 0, 1, \dots, MN$.
- 5: Given ν .
- 6: $I^k \leftarrow I^0$. Here I^0 is a given initial template.
- 7: Spatial regularization parameter $\mu_1 > 0$.
- 8: Error tolerance $\epsilon_I > 0$, stepsize $\alpha^k = \alpha > 0$, and iteration number $K_I > 0$.
- 9: *Loop:*
- 10: Compute $I^k \circ \phi_{\tau_j, 0}^\nu$ for $1 \leq j \leq MN$ by

$$I^k \circ \phi_{\tau_j, 0}^\nu \leftarrow (I^k \circ \phi_{\tau_{j-1}, 0}^\nu) \circ \left(\text{Id} - \frac{1}{MN} \nu(\tau_j, \cdot) \right)$$

with $I^k \circ \phi_{0, 0}^\nu = I^k$.

- 11: Update $\eta_{t_i, t_i}^{I^k, \nu}$ for $1 \leq i \leq N$ by

$$\eta_{t_i, t_i}^{I^k, \nu} \leftarrow \mathcal{T}_{t_i}^* (\mathcal{T}_{t_i} (I^k \circ \phi_{t_i, 0}^\nu) - g(t_i, \cdot))$$

- 12: Compute $\eta_{0, t_i}^{I^k, \nu}$ for $1 \leq i \leq N$ by

$$\eta_{\tau_j, t_i}^{I^k, \nu} \leftarrow \left(1 + \frac{1}{MN} \text{div} \nu(\tau_j, \cdot) \right) \eta_{\tau_{j+1}, t_i}^{I^k, \nu} \circ \left(\text{Id} + \frac{1}{MN} \nu(\tau_j, \cdot) \right)$$

for $j = iM - 1, iM - 2, \dots, 0$.

- 13: Evaluate I^{k+1} by

$$I^{k+1} \leftarrow I^k - \alpha \left(\frac{2}{N} \sum_{i=1}^N \eta_{0, t_i}^{I^k, \nu} + \mu_1 \nabla^* \left(\frac{\nabla I^k}{|\nabla I^k|_{2, \epsilon}} \right) \right).$$

- 14: **If** $|I^{k+1} - I^k| > \epsilon_I$ and $k < K_I$, then $k \leftarrow k + 1$, **goto** *Loop*.
 - 15: **Output** I^{k+1} .
-

478 **4.3. Velocity field estimation.** The aim here is to provide an algorithm for minimising
 479 \mathcal{E}_I in (4.8), which amounts to sequential indirect image registration. We will for this purpose
 480 use a gradient descent scheme of the form

$$481 \quad (4.22) \quad \nu^{k+1} = \nu^k - \beta^k \nabla \mathcal{E}_I(\nu^k).$$

482 Here, β^k is the step-size in the k -th iteration, and $\nabla \mathcal{E}_I(\nu^k) \in \mathcal{L}_V^2(\Omega)$ is calculated by (B.1).

483 The challenge here lies in the computation of this gradient (subsection 4.3.1) and the final
 484 algorithm for the gradient descent scheme in (4.22) is given in Algorithm 4.2.

485 **4.3.1. Computing the gradient $\nabla\mathcal{E}_I$.** We begin by introduce notations

$$486 \quad (4.23) \quad h_{\tau,t}^{I,\nu} := \begin{cases} \eta_{\tau,t}^{I,\nu} & 0 \leq \tau \leq t \leq 1 \\ 0 & t < \tau, \end{cases} \quad \text{and} \quad \nu_{\tau,t} := \begin{cases} \nu(\tau, \cdot) & 0 \leq \tau \leq t \leq 1, \\ 0 & t < \tau. \end{cases}$$

487 **Theorem B.1** gives an expression for $\nabla\mathcal{E}_I$ where the kernel function $\mathbb{K}: \Omega \times \Omega \rightarrow \mathbb{M}_+^{n \times n}$
 488 is evaluated on points that do not move as iteration proceeds. By choosing a translation
 489 invariant kernel and points on a regular grid in Ω , we can use FFT-based convolution scheme
 490 to efficiently evaluate the velocity field at each iteration. This is computationally more feasible
 491 than letting the kernel depend on points that move in time as in the shooting method [39, 57].

492 In what follows, we write out the explicit derivations for computing $\nabla\mathcal{E}_I$. As derived in
 493 **subsection 4.2.1**, $I \circ \phi_{\tau_j,0}^\nu$ can be approximated by (4.16). The key step is now to update $h_{\tau_j,t_i}^{I,\nu}$
 494 for $\{i : t_i \geq \tau_j\}$ in (B.1). We know, by (4.23),

$$495 \quad (4.24) \quad h_{\tau_j,t_i}^{I,\nu} = \eta_{\tau_j,t_i}^{I,\nu} \quad \text{for } t_i \geq \tau_j.$$

496 By using (4.21) for $1 \leq i \leq N$, we can compute $h_{\tau_j,t_i}^{I,\nu}$ as

$$497 \quad (4.25) \quad h_{\tau_j,t_i}^{I,\nu} \approx \left(1 + \frac{1}{MN} \operatorname{div} \nu(\tau_j, \cdot)\right) h_{\tau_{j+1},t_i}^{I,\nu} \circ \left(\operatorname{Id} + \frac{1}{MN} \nu(\tau_j, \cdot)\right)$$

498 for $j = iM - 1, iM - 2, \dots, 0$ and with $h_{t_i,t_i}^{I,\nu} = \mathcal{T}_{t_i}^* \left(\mathcal{T}_{t_i} (I \circ \phi_{t_i,0}^\nu) - g(t_i, \cdot) \right)$. Hence, at $t = \tau_j$,
 499 we get from (B.1) that

$$500 \quad (4.26) \quad \nabla\mathcal{E}_I(\nu)(\tau_j, x) = -\frac{2}{N} \sum_{\{i \geq 1: t_i \geq \tau_j\}} \left[\int_{\Omega} \mathbb{K}(x, y) \nabla (I \circ \phi_{\tau_j,0}^\nu)(y) h_{\tau_j,t_i}^{I,\nu}(y) \, dy - \mu_2 \nu(\tau_j, x) \right]$$

501 for $0 \leq j \leq MN$ and $x \in \Omega$. In particular, for $j = MN$ (i.e., $\tau_j = 1$) we have

$$502 \quad \nabla\mathcal{E}_I(\nu)(1, x) = -\frac{2}{N} \left[\int_{\Omega} \mathbb{K}(x, y) \nabla (I \circ \phi_{1,0}^\nu)(y) h_{1,1}^{I,\nu}(y) \, dy - \mu_2 \nu(1, x) \right].$$

503 **Remark 4.3.** It is easy to verify that the optimal solution of the time-discretized version
 504 of the proposed model is consistent with that of the time-continuous one. Here the concept
 505 of consistency is that the time-discretized solution converges to the time-continuous solution
 506 for increasingly fine time discretization. This is however not the case for the diffeomorphic
 507 motion model (3.25) in [33]. As an example, at $\tau_j = 1$, the optimal velocity field of the
 508 time-discretized problem in [33] satisfies

$$509 \quad \nu(1, x) = \frac{1}{\mu_2} \int_{\Omega} \mathbb{K}(x, y) \nabla (I \circ \phi_{1,0}^\nu)(y) h_{1,1}^{I,\nu}(y) \, dy.$$

510 On the other hand, as derived in **subsection 3.3.2**, the optimal velocity field at $t = 1$ of its
 511 time-continuous problem satisfies $\nu(1, x) = 0$. This obviously causes inconsistencies and our
 512 consistent approach is an advantage compared to the approach in [33].

513 Finally, **Algorithm 4.2** outlines the procedure for computing the gradient descent scheme
 514 (4.22) that makes use of the above derivations.

Algorithm 4.2 Gradient descent scheme for minimizing $\boldsymbol{\nu} \mapsto \mathcal{E}_I(\boldsymbol{\nu})$ in (4.8)

- 1: *Initialize:*
- 2: $k \leftarrow 0$.
- 3: $t_i \leftarrow \frac{i}{N}$ for $i = 0, 1, \dots, N$.
- 4: $\tau_j \leftarrow \frac{j}{MN}$ for $j = 0, 1, \dots, MN$.
- 5: Fixed I .
- 6: $\boldsymbol{\nu}^k(\tau_i) \leftarrow \boldsymbol{\nu}^0(\tau_i)$, where $\boldsymbol{\nu}^0$ is a given initial velocity field.
- 7: Fixed kernel function $\mathsf{K}(\cdot, \cdot)$.
- 8: Shape regularization parameter $\mu_2 > 0$.
- 9: Error tolerance $\epsilon_\nu > 0$, stepsize $\beta^k = \beta > 0$, and maximum iterations $K_\nu > 0$.
- 10: *Loop:*
- 11: Update $I \circ \phi_{\tau_j, 0}^{\boldsymbol{\nu}^k}$ for $1 \leq j \leq MN$ by

$$I \circ \phi_{\tau_j, 0}^{\boldsymbol{\nu}^k} \leftarrow (I \circ \phi_{\tau_{j-1}, 0}^{\boldsymbol{\nu}^k}) \circ \left(\text{Id} - \frac{1}{N} \boldsymbol{\nu}^k(\tau_j, \cdot) \right)$$

with $I \circ \phi_{0, 0}^{\boldsymbol{\nu}^k} = I$.

- 12: Update $h_{t_i, t_i}^{I, \boldsymbol{\nu}^k}$ for $1 \leq i \leq N$ by

$$h_{t_i, t_i}^{I, \boldsymbol{\nu}^k} \leftarrow \mathcal{T}_{t_i}^* (\mathcal{T}_{t_i} (I \circ \phi_{t_i, 0}^{\boldsymbol{\nu}^k}) - g(t_i, \cdot)).$$

- 13: Compute $h_{\tau_j, t_i}^{I, \boldsymbol{\nu}^k}$ for $1 \leq i \leq N$ by

$$h_{\tau_j, t_i}^{I, \boldsymbol{\nu}^k} \leftarrow \left(1 + \frac{1}{MN} \text{div} \boldsymbol{\nu}(\tau_j, \cdot) \right) h_{\tau_{j+1}, t_i}^{I, \boldsymbol{\nu}^k} \circ \left(\text{Id} + \frac{1}{MN} \boldsymbol{\nu}^k(\tau_j, \cdot) \right)$$

for $j = iM - 1, iM - 2, \dots, 0$.

- 14: Compute $\nabla_\nu \mathcal{E}_I(\boldsymbol{\nu}^k)(\tau_j, \cdot)$ (using FFT for convolutions) by

$$\nabla \mathcal{E}_I(\boldsymbol{\nu}^k)(\tau_j, x)$$

$$\leftarrow -\frac{2}{N} \sum_{\{i \geq 1: t_i \geq \tau_j\}} \left[\int_{\Omega} \mathsf{K}(x, y) \nabla (I \circ \phi_{\tau_j, 0}^{\boldsymbol{\nu}^k})(y) h_{\tau_j, t_i}^{I, \boldsymbol{\nu}^k}(y) \, dy - \mu_2 \boldsymbol{\nu}^k(\tau_j, x) \right]$$

for $0 \leq j \leq MN$.

- 15: Update $\boldsymbol{\nu}^k(\tau_j, \cdot)$ for $0 \leq j \leq MN$ by:

$$\boldsymbol{\nu}^{k+1}(\tau_j, \cdot) \leftarrow \boldsymbol{\nu}^k(\tau_j, \cdot) - \beta \nabla \mathcal{E}_I(\boldsymbol{\nu}^k)(\tau_j, \cdot).$$

- 16: **If** $|\boldsymbol{\nu}^{k+1} - \boldsymbol{\nu}^k| > \epsilon_\nu$ and $k < K_\nu$, then $k \leftarrow k + 1$, **goto** *Loop*.

- 17: **Output** $\boldsymbol{\nu}^{k+1}$.
-

515 **4.4. Alternating template reconstruction and velocity field estimation.** As described
 516 in the beginning of section 4, we aim to solve (4.6) by an iterative scheme where iterates
 517 for template image I and velocity field ν are updated in an alternating manner as in (4.9).
 518 Hence, at each iterative step requires solving two sub-problems, one for updating I given ν
 519 (Algorithm 4.1) and the other for updating ν given I (Algorithm 4.2).

520 The algorithms for solving the two sub-problems are iterative, so there are inner iterations
 521 for each outer iterative step that update the template and velocity field. In order to have a
 522 computationally feasible approach, we limit the number of inner iterations to one. The final
 523 algorithm for recovering the template and velocity field is presented in Algorithm 4.3. The
 524 proposed model is non-convex so a convergence analysis of the algorithm towards a global
 optima is currently too difficult to obtain.

Algorithm 4.3 Alternately minimizing model (4.8)

- 1: *Initialize:*
 - 2: Given M, N .
 - 3: $k \leftarrow 0$.
 - 4: $t_i \leftarrow \frac{i}{N}$ for $i = 0, 1, \dots, N$. This subdivides the time $[0, 1]$ uniformly into N parts.
 - 5: $\tau_j \leftarrow \frac{j}{MN}$ for $j = 0, 1, \dots, MN$. This subdivides the time interval $[0, 1]$ uniformly into MN parts.
 - 6: Fixed kernel function $K(\cdot, \cdot)$.
 - 7: Given regularization parameters $\mu_1, \mu_2 > 0$.
 - 8: $I^k \leftarrow I^0$, where the template is initialized.
 - 9: $\nu^k(\tau_i) \leftarrow 0$, where the velocity field is initialized to a zero velocity field.
 - 10: Error tolerances $\epsilon_I, \epsilon_\nu > 0$, stepsizes $\alpha^k = \alpha > 0, \beta^k = \beta > 0$, and maximum iteration number $K > 0$.
 - 11: *Loop:*
 - 12: Let $\nu = \nu^k$. Perform lines 10–13 in Algorithm 4.1. Output I^{k+1} .
 - 13: Let $I = I^{k+1}$. Perform lines 11–15 in Algorithm 4.2. Output ν^{k+1} .
 - 14: **If** $|\nu^{k+1} - \nu^k| > \epsilon_\nu$ or $|I^{k+1} - I^k| > \epsilon_I$, and $k < K$,
 then $k \leftarrow k + 1$, **goto** *Loop*.
 - 15: **Output** I^{k+1}, ν^{k+1} .
-

525 **Complexity analysis.** Here we consider the complexity, including computational cost and
 526 memory footprint, of Algorithm 4.3. The main computationally demanding steps at each
 527 iteration is located on lines 12–13 (actually lines 10–13 in Algorithm 4.1 and lines 11–15 in
 528 Algorithm 4.2), so we restrict our complexity analysis to these parts. For ease of description,
 529 we assume $\Omega \subset \mathbb{R}^2$ and the size of the image to be reconstructed is $n \times n$ pixels.

531 On line 10 of Algorithm 4.1 and line 11 of Algorithm 4.2, we need to update $I \circ \phi_{\tau_j, 0}^\nu$
 532 for $j = 1, \dots, MN$. Moreover, each of them should be used to compute the gradient of the
 533 objective functional on line 14 of Algorithm 4.2, so they need to be stored at hand. Hence, in
 534 these two steps, the computational cost is $O(n^2MN)$ and the memory footprint is $O(n^2MN)$.

535 For line 11 of Algorithm 4.1 and line 12 of Algorithm 4.2, the $\eta_{t_i, t_i}^{I, \nu}$ (i.e., $h_{t_i, t_i}^{I, \nu}$) need to be
 536 updated and then stored for $i = 1, \dots, N$. The computational cost is $O(n^2N_dN)$, where N_d

537 is the number of data points. Actually, the N_d is at least proportional to the size of n , which
 538 is often $\sqrt{2}nN_v$ with N_v denoting the number of views. Hence, the computational cost scales
 539 as $O(n^3NN_v)$. Since the calculation for the forward and backward projections is on the fly,
 540 the required space is not too demanding.

541 Furthermore, on line 12 of [Algorithm 4.1](#) and line 13 of [Algorithm 4.2](#), for $i = 1, \dots, N$,
 542 the $\eta_{\tau_j, t_i}^{I, \nu}$ (i.e., $h_{\tau_j, t_i}^{I, \nu}$) need to be updated and stored for j from $iM - 1$ to 0, then are used to
 543 compute the gradient of the objective functional for each time point on line 14 of [Algorithm 4.2](#).
 544 Therefore, the computational cost is $O(n^2MN^2)$. For lines 11-12 of [Algorithm 4.1](#) and lines 12-
 545 13 of [Algorithm 4.2](#), the memory footprint is $O(n^2MN^2)$.

546 For line 13 of [Algorithm 4.1](#), we need to update I once, the computational cost is $O(n^2N)$
 547 and the memory footprint is $O(n^2N)$. At each time point, the FFT is used to compute the
 548 gradient of the objective functional on line 14 of [Algorithm 4.2](#). Hence the computational
 549 cost for this line is $O(MN^2n^2 \log n)$. For line 15 of [Algorithm 4.2](#), we need to update a vector
 550 field at each time point. Since a vector field would take twice more memory than a scalar
 551 field on 2D domain, we spend twice more computational cost to update that. Even so, the
 552 computational cost is $O(n^2MN)$ and the memory footprint is $O(n^2MN)$.

553 In summary, for [Algorithm 4.3](#), the computational cost scales as $O(n^3N)$ and the memory
 554 footprint scales as $O(n^2MN^2)$.

555 **5. Numerical experiments.** In this section, the proposed method for joint image recon-
 556 struction and motion estimation is applied to parallel beam tomography with very sparse or
 557 highly noisy data in spatiotemporal (e.g., 2D + time) imaging. We use the intensity-preserving
 558 group action to consider the involved deformations. The algorithms were implemented in
 559 Python using the Operator Discretization Library (<http://github.com/odlgroup/odl>), which
 560 offers GPU parallelized routines for evaluating the ray transform and its adjoint. The source
 561 code is available from <https://github.com/chongchenmath/SpatiotemporalImaging>. The nu-
 562 merical experiments ran on ThinkStation with Xeon E5-2620 v4 2.10 GHz CPU, 64Gb RAM,
 563 TITAN Xp GPU, and Ubuntu 16.04 OS. Although not a complete evaluation, the experiments
 564 illustrate the performance of the proposed method.

565 **5.1. Spatiotemporal 2D CT.** Image reconstruction in classical static 2D CT amounts to
 566 inverting the 2D ray transform. In our spatiotemporal setting, we assume that the temporal
 567 variation comes from the object itself that undergoes motion. Hence, the forward operator
 568 $\mathcal{T}_t: \mathcal{X} \rightarrow \mathcal{Y}$, which is the 2D ray transform, is not dependent on time t :

$$569 \quad \mathcal{T}(f)(\omega, x) := \int_{\mathbb{R}} f(x + s\omega) ds \quad \text{for } \omega \in S^1 \text{ and } x \in \omega^\perp.$$

570 Here, S^1 is the unit circle and (ω, x) encodes a line on \mathbb{R}^2 with direction ω through x .

571 Moreover, consider \mathcal{V} as the space of vector fields that is a RKHS with a reproducing
 572 kernel represented by symmetric and positive definite Gaussian function $\mathsf{K}: \Omega \times \Omega \rightarrow \mathbb{M}_+^{2 \times 2}$
 573 given as

$$574 \quad (5.1) \quad \mathsf{K}(x, y) := \exp\left(-\frac{1}{2\sigma^2}\|x - y\|_2^2\right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{for } x, y \in \mathbb{R}^2 \text{ and } \sigma > 0 \text{ (kernel width).}$$

575 The images of all gates are supported on Ω . At each gate, the noise-free data per view is

576 measured by evaluating the 2D parallel beam scanning geometry. Then the additive Gaussian
 577 white noise at varying levels is added onto the noise-free data, which leads to the noise data.
 578 As in [17], the noise level in data is quantified in terms of signal-to-noise ratio (SNR) defined
 579 in logarithmic decibel (dB).

580 **5.2. Test suites and results.** The test suites seek to assess the performance against dif-
 581 ferent noise levels, and the sensitivity against various selections of regularization parameters
 582 μ_1 , μ_2 , and kernel width σ . We also compare the proposed method to TV-based static recon-
 583 struction method.

584 **5.2.1. Test suite 1: Overview performance.** Here we consider a test for evaluating the
 585 overview performance. This test uses a multi-object phantom with five gates (i.e., $N = 5$).
 586 The used phantom is shown in the last row of Figure 2, which is taken from [17].

587 The image at each gate is consisting of six separately star-like objects with grey-values
 588 over $[0, 1]$, which is digitized using 438×438 pixels. The images of all gates are supported on
 589 a fixed rectangular domain $\Omega = [-16, 16] \times [-16, 16]$. At each gate, the noise-free data per
 590 view is measured by the 2D parallel beam scanning geometry with even 620 bins, which is
 591 supported on the range of $[-24, 24]$. For gate i ($1 \leq i \leq N$), the scanning views are distributed
 592 on $[(i-1)\pi/36, \pi + (i-1)\pi/36]$ uniformly, and the view number is 12. Then three different
 593 levels of additive Gaussian white noise are added onto the noise-free data. The resulting SNR
 594 are about 4.71dB, 7.7dB, and 14.67dB, respectively. To make it more clear, we show the
 595 noise-free and noise projection data at the first view for Gate 1 with respective noise levels in
 Figure 1.

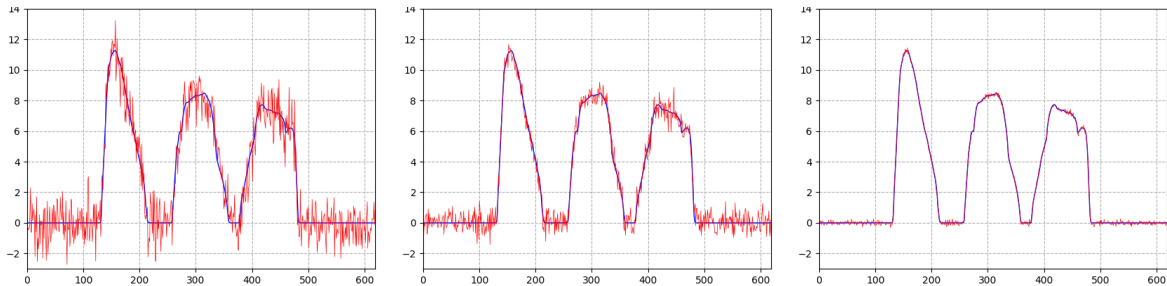


Figure 1: Data at the first view for Gate 1. The left, middle, and right figures show data at the first view for different noise levels 4.71dB, 7.7dB, and 14.67dB, respectively. The blue smooth curve is noise-free data, and the red jagged curve is noisy data.

596 The factor of discretized time degree is $M = 2$, which is defined in subsection 4.2.1. The
 597 kernel width is selected to $\sigma = 2$. The gradient stepsizes are set as $\alpha = 0.01$ and $\beta = 0.05$
 598 respectively, which should not be chosen too large or too small, otherwise could result in the
 599 algorithm not convergent or convergent too slow. First we apply Algorithm 4.1 to obtain
 600 an initial template image after 50 iterations, then use Algorithm 4.3 to solve the proposed
 601 model. Note that the above iteration number is not unchangeable, just needs enough to gain
 602 an appropriately initial template for Algorithm 4.3.

603 The regularization parameters (μ_1, μ_2) are selected as $(0.05, 10^{-7})$ for data noise level
 604

605 4.71dB, $(0.025, 10^{-7})$ for data noise level 7.7dB, and $(0.01, 10^{-7})$ for data noise level 14.67dB,
 606 respectively. The lower SNR, the larger value of μ_1 . The maximum iteration number is set to
 607 be 200, which should be large enough to guarantee a satisfying result. The runtime for each
 608 example is about 29 minutes. The reconstructed results are shown in [Figure 2](#). It is clear
 609 that the reconstructed images (rows 1–3) are close to the corresponding ground truth, even
 though the data SNR is very low.

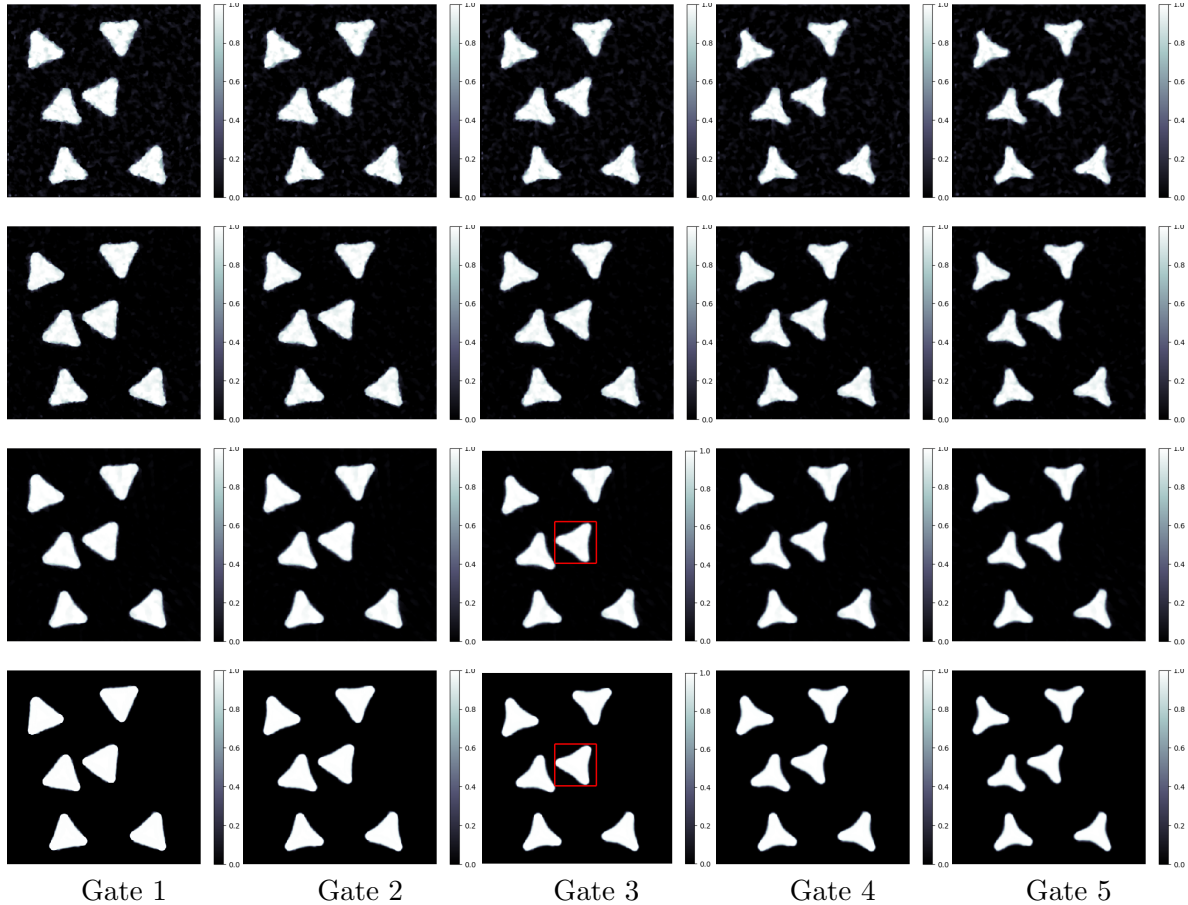


Figure 2: Multi-object phantom. Columns represent the gates and the first three rows are reconstructed spatiotemporal images for the data with noise levels 4.71dB, 7.7dB, and 14.67dB, respectively. The last row shows the ground truth for each gate. The regions of interest (the small boxes in images at the 3rd column in rows 3 and 4) is enlarged in [Figure 4](#).

610

611

612

613

614

615

Apart from the visual perception, the reconstruction is quantitatively compared using structural similarity (SSIM) and peak signal-to-noise ratio (PSNR), which is frequently used to evaluate image quality [59]. The SSIM and PSNR values are tabulated in [Table 1](#). As listed in the above table, the corresponding SSIM and PSNR values are depended on SNR of the data. The higher SNR, the larger values of SSIM and PSNR.

Figure 2	Gate 1	Gate 2	Gate 3	Gate 4	Gate 5
Row 1	0.4069 22.10	0.4208 23.02	0.4273 23.27	0.4305 23.40	0.4337 23.64
Row 2	0.5934 25.36	0.6086 27.22	0.6131 27.37	0.6149 27.66	0.6156 27.86
Row 3	0.8411 28.30	0.8523 31.49	0.8564 32.48	0.8576 32.65	0.8587 32.76

Table 1: SSIM and PSNR values of reconstructed spatiotemporal images for data with varying noise levels, see Figure 2 for the images. Comparisons are made against ground truth. Each table entry has two values, the top being the SSIM and the bottom being the PSNR. The image is shown in the corresponding row/gate in Figure 2.

616 *Comparison against static TV-regularized reconstruction.* It is well-known that tomographic
 617 reconstruction by TV-regularization outperforms other methods, such as filtered back projec-
 618 tion (FBP), when the gradient of the image is sparse. This is furthermore especially notable
 619 when data is under-sampled. In our tests we use a phantom (ground truth image) that has
 620 sparse gradient, so comparing against static TV-regularized reconstruction pitches our ap-
 621 proach against one of the best static reconstruction methods.

622 For static TV-regularized reconstruction we disregard any temporal evolution, which is
 623 equivalent to simplify the spatiotemporal problem into a static problem. The whole tomo-
 624 graphic data set will then have 60 projection views. The regularization parameter for static
 625 TV-regularized reconstruction is selected depending on the SNR of data in the same way as
 for spatiotemporal reconstruction.

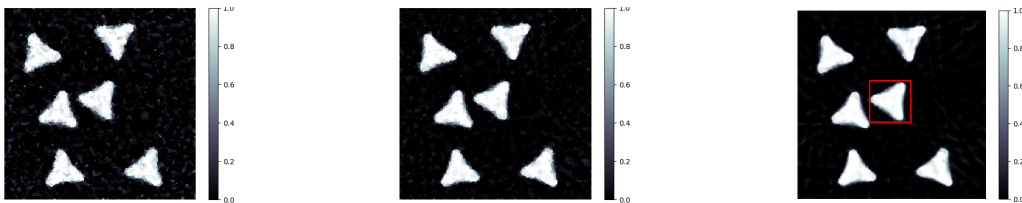


Figure 3: Static TV-regularized reconstructions for the measured data with different noise levels 4.71dB (left), 7.7dB (middle), and 14.67dB (right), respectively. The region of interest in the small box (right) is enlarged in Figure 4.

626 Reconstructions obtained by static TV-regularized reconstruction are shown in Figure 3,
 627 the edges of which become blurring against those by our method. To make it more clear,
 628 we enlarge the regions of interest in the small boxes of Figure 2 (column 3, rows 3–4) and
 629 Figure 3 for comparison, which are shown in Figure 4.

631 In addition, the corresponding SSIM and PSNR values are listed in Table 2 and compared
 632 to Table 1, the values of SSIM and PSNR for static TV-regularization are lower than those
 633 with the proposed method.



Figure 4: Regions of interest for the images in column 3 in Figure 2. Left images shows the region of interest in row 4 (ground truth), middle is for row 3 (the proposed method)) and right is from Figure 3 (static TV-regularization).

Figure 3	Gate 1	Gate 2	Gate 3	Gate 4	Gate 5
Left	0.3012	0.3163	0.3202	0.3146	0.3030
	18.57	19.94	20.42	19.98	18.80
Middle	0.4673	0.4867	0.4910	0.4840	0.4694
	20.44	22.82	23.76	22.90	20.79
Right	0.6004	0.6239	0.6291	0.6212	0.6029
	21.42	24.71	26.40	25.00	21.95

Table 2: SSIM and PSNR values of TV-regularized reconstructions compared to each ground truth. Data is from Gates 1–5 with varying noise levels, see Figure 3 for the images. Each entry has two values, where the upper is the value of SSIM and the bottom is the value of PSNR.

5.2.2. Test suite 2: Sensitivity against selections of regularization parameters.

To solve the proposed model, three regularization parameters μ_1 , μ_2 and σ need to be selected. Hence the sensitivity test should be concerned against the selections of these parameters.

As shown in the last row of Figure 5, a heart phantom with four gates (i.e., $N = 4$) is used in this test, which is originated from [31]. The image at each gate is consisting of a heart-like object with grey-values in $[0, 1]$, which is digitized using 120×120 pixels. The images of all gates are supported on a fixed rectangular domain $\Omega = [-4.5, 4.5] \times [-4.5, 4.5]$. At each gate, the noise-free data per view is measured by evaluating the 2D parallel beam scanning geometry with uniform 170 bins, which is supported on the range of $[-6.4, 6.4]$. Then the additive Gaussian white noise is added onto the noise-free data. The resulting SNR is about 14.9dB. For gate i ($1 \leq i \leq N$), the scanning views are distributed on $[(i-1)\pi/5, \pi+(i-1)\pi/5]$ evenly, which totally has 20 views. The factor of discretized time degree is $M = 2$. The gradient stepsizes are set as $\alpha = 0.01$ and $\beta = 0.05$, respectively.

We first employ Algorithm 4.1 to gain an initial template after 50 iterations, then use Algorithm 4.3 to solve the proposed model. With selecting different values for regularization parameters, after 200 iterations, the reconstructed results are obtained, as shown in Figure 5. The runtime for each example is about 3 minutes. The detailed selections of varying parameters can be referred to the caption of Figure 5.

For comparison, we also present the result for static TV-regularized reconstruction in

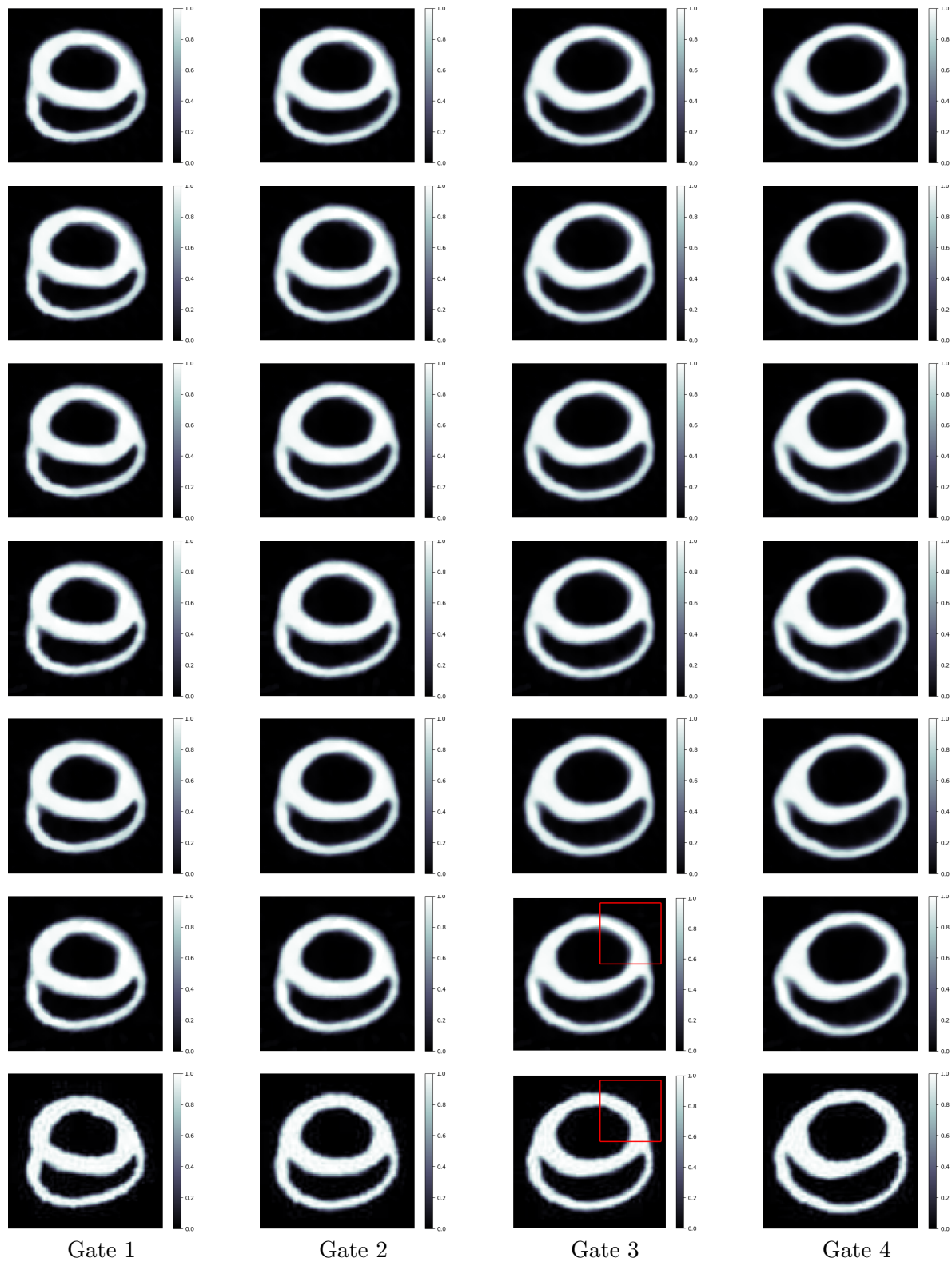


Figure 5: Heart phantom. The columns are the 4 gates and the first 6 rows are reconstructed spatiotemporal images with parameter pairs (μ_1, μ_2, σ) chosen as $(0.01, 10^{-7}, 1.0)$, $(0.01, 10^{-6}, 1.0)$, $(0.01, 10^{-7}, 0.5)$, $(0.005, 10^{-7}, 0.5)$, $(0.01, 10^{-6}, 0.5)$, and $(0.005, 10^{-6}, 0.5)$. The last row shows the ground truth for each gate. The regions of interest in the small boxes (column 3, rows 6–7) are enlarged in [Figure 6](#).

653 Figure 6 as we did in the first test. As shown in Figure 5, the related reconstructed results are
 654 almost the same and close to the counterpart ground truth. However, the reconstructed result
 655 by static TV-regularization in Figure 6 is severely degraded. To illustrate this, we enlarge
 656 the regions of interest in the small boxes of Figure 5 (column 3, rows 6-7) and Figure 6 for
 comparison, which are shown in Figure 6.

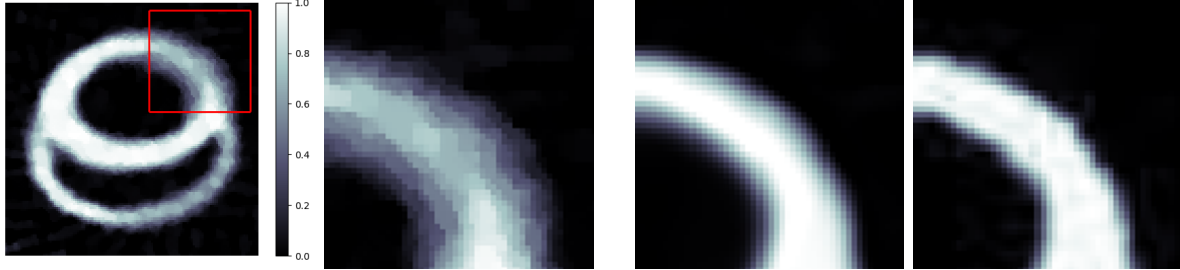


Figure 6: Leftmost image shows reconstruction using static TV-regularization from the measured data with comparable noise level with the region of interest shown separately in the image of the right. The next image to the right is the same region taken from column 3, row 6 in Figure 5 (the proposed method). Finally, the rightmost image is from column 3, row 7 in Figure 5 (ground truth).

657 Furthermore, SSIM and PSNR values are reported in Table 3. As can be seen, the SSIM
 658 and PSNR values for the proposed method are better than those obtained by static TV-
 659 regularized reconstruction. Furthermore, they are quite similar for different parameter choices.
 660

661 As shown in Table 3, these values are a little bit decreased when the value of kernel
 662 parameter σ is changed from 1.0 to 0.5 with fixed μ_1 and μ_2 , as compared the values between
 663 row 1 and row 3, also row 2 and row 4, for instance. Therefore, this test demonstrates that to
 664 some extent the proposed method is not sensitive to the precise selection of the regularization
 665 parameters under the visual perception and the quantitative comparison (SSIM and PSNR).
 666 However, those values are selected too big or too small, which would causes over- or under-
 667 regularized results.
 668

669 As indicated in subsection 3.3, the optimal velocity field of the proposed model is nonva-
 670 nishing at both the initial and the end time points. To verify this standpoint in numerical
 671 way, we display the computed optimal velocity field at times $t = 0$ and $t = 1$ in Figure 7 for
 672 the example with parameter pair $(0.005, 10^{-7}, 0.5)$ on Row 4 of Figure 5.

673 **6. Conclusions and the future work.** A general framework of variational model has been
 674 investigated for joint image reconstruction and motion estimation in spatiotemporal imag-
 675 ing, which is based on the deformable templates from shape theory. Along this framework,
 676 we proposed a new variational model for solving the above joint problem using the prin-
 677 ciple of LDDMM. The proposed model is equivalent to a PDE-constrained optimal control
 678 problem. Based on the equivalency, we made a mathematical comparison against the joint
 679 TV-TV optical flow based model [13], which showed that our method can guarantee elasti-
 680 cally diffeomorphic deformations, and is of benefit to the practical computation additionally.

	Gate 1	Gate 2	Gate 3	Gate 4
Row 1	0.8928	0.9382	0.9340	0.9235
	24.25	28.44	27.64	26.28
Row 2	0.8960	0.9415	0.9346	0.9234
	24.30	28.47	27.67	26.37
Row 3	0.9103	0.9497	0.9459	0.9343
	25.33	29.41	28.97	27.78
Row 4	0.8940	0.9368	0.9361	0.9291
	25.13	29.06	28.60	27.65
Row 5	0.9087	0.9472	0.9462	0.9336
	25.14	29.30	28.83	27.68
Row 6	0.8884	0.9339	0.9358	0.9295
	25.23	29.06	28.65	27.74
Static TV	0.5641	0.7310	0.7458	0.5969
	14.09	19.09	18.96	14.01

Table 3: Rows 1–6: SSIM and PSNR values of reconstructed spatiotemporal images for varying values of the regularization parameters μ_1 , μ_2 , and the kernel width σ . Each table entry has two values, the top being the SSIM and the bottom being the PSNR. The image is shown in the corresponding row/gate in Figure 5. The bottom row reports SSIM and PSNR values of TV-regularized reconstructions compared to each ground truth from gates 1–4 by the measured data with comparable noise level.

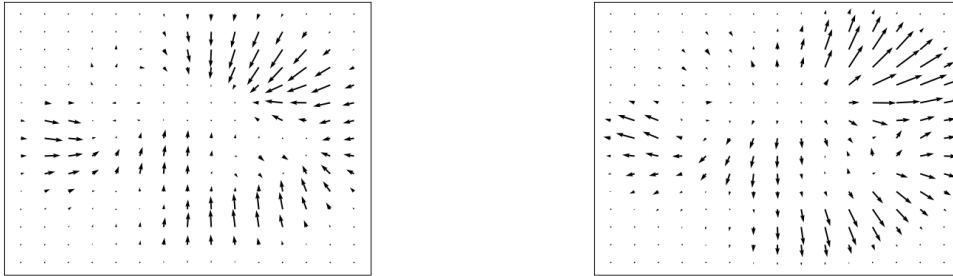


Figure 7: The computed optimal velocity fields at times $t = 0$ (left) and $t = 1$ (right) for the example with parameter pair $(0.005, 10^{-7}, 0.5)$ on row 4 in Figure 5.

681 Furthermore, the theoretical comparison was also performed between the proposed model and
682 other diffeomorphic motion models, which demonstrated that the optimal velocity field of our
683 model is distributed w.r.t. time t averagely, and non-vanishing at both the initial and the end
684 time points. We also presented an efficient computational method for the time-discretized for-
685 mulation and showed that its optimal solution is consistent with that of the time-continuous
686 one. This is not the case for the diffeomorphic motion model in [33].

687 An alternating gradient descent algorithm was designed to solve the time-discretized pro-
 688 posed model, where the main calculations were only based on the easily-implemented linearized
 689 deformations. For spatiotemporal (2D space + time) parallel beam tomographic imaging, the
 690 computational cost of the algorithm is then $O(n^3NN_v)$ and its memory requirement scales as
 691 $O(n^2MN^2)$. With [Algorithm 4.3](#), we have evaluated the performance of the proposed model
 692 in dealing with the 2D space + time tomography in the case of very sparse and/or highly
 693 noisy data. As shown in these visual and quantitative results, the new method can yield
 694 reconstructed spatiotemporal images of high quality for the above difficult problems.

695 The future work will focus on the theoretical analysis of the proposed model, such as the
 696 existence and uniqueness of the solution, the convergence analysis of the proposed algorithm,
 697 and its extensions and applications to more complicated modalities in spatiotemporal imaging.

698 **Appendix A. Optimality conditions.** To derive optimality conditions for [\(3.11\)](#), we begin
 699 with the following result, which is proved in [\[60\]](#).

700 **Lemma A.1.** Let $\nu, \eta \in \mathcal{L}_V^2(\Omega)$, $\phi_{0,t}^\nu$ denotes the solution to the ODE in [\(3.1\)](#) with given
 701 ν at time t , and $\phi_{s,t}^\nu$ is defined as in [\(3.4\)](#). Then,

$$702 \quad (\text{A.1}) \quad \left. \frac{d}{d\epsilon} \phi_{s,t}^{\nu+\epsilon\eta}(x) \right|_{\epsilon=0} = \int_s^t D(\phi_{\tau,t}^\nu)(\phi_{s,\tau}^\nu(x)) \left(\eta(\tau, \phi_{s,\tau}^\nu(x)) \right) d\tau \quad \text{for } x \in \Omega \text{ and } 0 \leq s, t \leq 1.$$

703 [Lemma A.1](#) can be used to prove the following statement:

704 **Lemma A.2.** Let the assumptions in [Lemma A.1](#) hold and assume $I \in \mathcal{L}^2(\Omega, \mathbb{R})$ is differ-
 705 entiable. Then, under the group action in [\(2.4\)](#) we have

$$706 \quad (\text{A.2}) \quad \left. \frac{d}{d\epsilon} (\phi_{0,t}^{\nu+\epsilon\eta} \cdot I)(x) \right|_{\epsilon=0} = - \int_0^t \left\langle \nabla(\phi_{0,\tau}^\nu \cdot I)(\phi_{t,\tau}^\nu(x)), \eta(\tau, \phi_{t,\tau}^\nu(x)) \right\rangle_{\mathbb{R}^n} d\tau \quad \text{for } x \in \Omega.$$

707 *Proof.* By the chain rule we get

$$708 \quad (\text{A.3}) \quad \left. \frac{d}{d\epsilon} (\phi_{0,t}^{\nu+\epsilon\eta} \cdot I)(x) \right|_{\epsilon=0} = \left\langle \nabla I(\phi_{t,0}^\nu(x)), \left. \frac{d}{d\epsilon} \phi_{t,0}^{\nu+\epsilon\eta}(x) \right|_{\epsilon=0} \right\rangle_{\mathbb{R}^n}.$$

709 Using [Lemma A.1](#), we know

$$710 \quad (\text{A.4}) \quad \left. \frac{d}{d\epsilon} \phi_{t,0}^{\nu+\epsilon\eta}(x) \right|_{\epsilon=0} = - \int_0^t D(\phi_{\tau,0}^\nu)(\phi_{t,\tau}^\nu(x)) \left(\eta(\tau, \phi_{t,\tau}^\nu(x)) \right) d\tau.$$

711 Inserting [\(A.4\)](#) into [\(A.3\)](#), we immediately prove [\(A.2\)](#). ■

712 The following result is a direct consequence of the above definition and [Lemma A.2](#).

713 **Lemma A.3.** Let the assumptions in [Lemma A.2](#) hold and define the data discrepancy
 714 functional $\mathcal{D}_{gt}: \mathcal{X} \rightarrow \mathbb{R}$ as in [\(3.13\)](#). If \mathcal{D}_{gt} is differentiable, then

$$715 \quad (\text{A.5}) \quad \left. \frac{d}{d\epsilon} \mathcal{D}_{gt}(\phi_{0,t}^{\nu+\epsilon\eta} \cdot I) \right|_{\epsilon=0} \\
 716 \quad = \int_0^t \left\langle -|D(\phi_{\tau,t}^\nu)| \nabla \mathcal{D}_{gt}(\phi_{0,t}^\nu \cdot I)(\phi_{\tau,t}^\nu) \nabla(\phi_{0,\tau}^\nu \cdot I), \eta(\tau, \cdot) \right\rangle_{\mathcal{L}^2(\Omega, \mathbb{R}^n)} d\tau. \\
 717 \\
 718$$

Appendix B. First-order variation of \mathcal{E}_I .

Theorem B.1. *Let the assumptions in Lemma A.2 hold and $\mathcal{E}_I: \mathcal{L}_\gamma^2(\Omega) \rightarrow \mathbb{R}$ is given as in (4.8). Assume furthermore that \mathcal{V} is a RKHS with a reproducing kernel $\mathbf{K}: \Omega \times \Omega \rightarrow \mathbb{M}_+^{n \times n}$. Then, the $\mathcal{L}_\gamma^2(\Omega)$ -gradient of \mathcal{E}_I is*

$$(B.1) \quad \nabla \mathcal{E}_I(\boldsymbol{\nu})(t, x) = -\frac{2}{N} \int_{\Omega} \mathbf{K}(x, y) \nabla(I \circ \phi_{t,0}^{\boldsymbol{\nu}})(y) \sum_{\{i \geq 1: t_i \geq t\}} h_{t,t_i}^{I,\boldsymbol{\nu}}(y) dy + \frac{2\mu_2}{N} \sum_{\{i \geq 1: t_i \geq t\}} \boldsymbol{\nu}_{t,t_i}(x),$$

for $0 \leq t \leq 1$ and $x \in \Omega$.

Proof. From Lemma A.2 it is not difficult to derive

$$\begin{aligned} \frac{d}{d\epsilon} \mathcal{E}_I(\boldsymbol{\nu} + \epsilon \boldsymbol{\eta}) \Big|_{\epsilon=0} &= \frac{1}{N} \sum_{i=1}^N \int_0^{t_i} \left\langle -2\eta_{\tau,t_i}^{I,\boldsymbol{\nu}} \nabla(I \circ \phi_{\tau,0}^{\boldsymbol{\nu}}), \boldsymbol{\eta}(\tau, \cdot) \right\rangle_{\mathcal{L}^2(\Omega, \mathbb{R}^n)} d\tau \\ &\quad + \frac{\mu_2}{N} \sum_{i=1}^N \int_0^{t_i} \left\langle 2\boldsymbol{\nu}(\tau, \cdot), \boldsymbol{\eta}(\tau, \cdot) \right\rangle_{\mathcal{V}} d\tau \\ &= \int_0^1 \left\langle -\frac{2}{N} \sum_{i=1}^N h_{\tau,t_i}^{I,\boldsymbol{\nu}} \nabla(I \circ \phi_{\tau,0}^{\boldsymbol{\nu}}), \boldsymbol{\eta}(\tau, \cdot) \right\rangle_{\mathcal{L}^2(\Omega, \mathbb{R}^n)} d\tau \\ &\quad + \int_0^1 \left\langle \frac{2\mu_2}{N} \sum_{i=1}^N \boldsymbol{\nu}_{\tau,t_i}(\cdot), \boldsymbol{\eta}(\tau, \cdot) \right\rangle_{\mathcal{V}} d\tau \\ &= \int_0^1 \left\langle -\frac{2}{N} \sum_{\{i \geq 1: t_i \geq t\}} h_{t,t_i}^{I,\boldsymbol{\nu}} \nabla(I \circ \phi_{t,0}^{\boldsymbol{\nu}}), \boldsymbol{\eta}(t, \cdot) \right\rangle_{\mathcal{L}^2(\Omega, \mathbb{R}^n)} dt \\ &\quad + \int_0^1 \left\langle \frac{2\mu_2}{N} \sum_{\{i \geq 1: t_i \geq t\}} \boldsymbol{\nu}_{t,t_i}(\cdot), \boldsymbol{\eta}(t, \cdot) \right\rangle_{\mathcal{V}} dt. \end{aligned}$$

The last two equations are obtained by inserting (4.23). Combining the above with (A.11) proves (B.1). \blacksquare

Acknowledgments. The authors would like to thank Alain Trouvé for his helpful discussions. The authors also would like to express their gratitude to the anonymous reviewers for their comments, which helped to improve the presentation of this paper.

REFERENCES

- [1] G. AUBERT AND P. KORNPÖBST, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, vol. 147 of Applied Mathematical Sciences, Springer-Verlag, 2002.
- [2] W. BAI AND M. BRADY, *Regularized B-spline deformable registration for respiratory motion correction in PET images*, Phys. Med. Biol., 54 (2009), pp. 2719–2736.
- [3] W. BAI AND M. BRADY, *Motion correction and attenuation correction for respiratory gated PET images*, IEEE Trans. Med. Imaging, 30 (2011), pp. 351–365.
- [4] H. H. BAUSCHKE AND P. L. COMBETTES, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, 2ed ed., 2017.

- 779 [5] F. M. BEG, M. I. MILLER, A. TROUVÉ, AND L. YOUNES, *Computing large deformation metric mappings*
780 *via geodesic flow of diffeomorphisms*, International Journal of Computer Vision, 61 (2005), pp. 139–
781 157.
- 782 [6] R. BHAGALIA, J. PACK, S. MILLER, AND M. IATROU, *Nonrigid registration-based coronary artery motion*
783 *correction for cardiac computed tomography*, Med. Phys., 39 (2012), pp. 4245–4254.
- 784 [7] M. BLUME, N. NAVAB, AND M. RAFECAS, *Joint image and motion reconstruction for PET using a*
785 *B-spline motion model*, Phys. Med. Biol., 57 (2012), pp. 8249–8270.
- 786 [8] M. BREHM, P. PAYSAN, M. OELHAFEN, AND M. KACHELRIESS, *Artifact-resistant motion estimation*
787 *with a patient-specific artifact model for motion-compensated cone-beam CT*, Med. Phys., 40 (2013),
788 p. 101913.
- 789 [9] M. BREHM, P. PAYSAN, M. OELHAFEN, P. KUNZ, AND M. KACHELRIESS, *Self-adapting cyclic registration*
790 *for motion-compensated cone-beam CT in image-guided radiation therapy*, Med. Phys., 39 (2012),
791 pp. 7603–7618.
- 792 [10] M. BREHM, S. SAWALL, J. MAIER, S. SAUPPE, AND M. KACHELRIESS, *Cardiorespiratory motion-*
793 *compensated micro-CT image reconstruction using an artifact model-based motion estimation*, Med.
794 Phys., 42 (2015), pp. 1948–1958.
- 795 [11] M. BRUVERIS AND D. D. HOLM, *Geometry of image registration: The diffeomorphism group and momen-*
796 *tum maps*, in Geometry, Mechanics, and Dynamics: The Legacy of Jerry Marsden, C. D. E., D. D.
797 Holm, G. Patrick, and T. Ratiu, eds., vol. 73 of Fields Institute Communications, Springer-Verlag,
798 2015, pp. 19–56.
- 799 [12] M. BRUVERIS AND F.-X. VIALARD, *On completeness of groups of diffeomorphisms*, J. Eur. Math. Soc.,
800 19 (2017), pp. 1507–1544.
- 801 [13] M. BURGER, H. DIRKS, AND C. SCHÖNLIEB, *A Variational Model for Joint Motion Estimation and Image*
802 *Reconstruction*, SIAM Journal on Imaging Sciences, 11 (2018), pp. 94–128.
- 803 [14] M. BURGER, J. MODERSITZKI, AND L. RUTHOTTO, *A Hyperelastic Regularization Energy for Image*
804 *Registration*, SIAM J. Sci. Comput., 35 (2013), pp. B132–B148.
- 805 [15] F. BÜTHER, M. DAWOOD, L. STEGGER, F. WÜBBELING, M. SCHÄFERS, O. SCHOBER, AND K. P.
806 SCHÄFERS, *List mode-driven cardiac and respiratory gating in PET*, J. Nucl. Med., 50 (2009), pp. 674–
807 681.
- 808 [16] B. CHEN, Z. ZHANG, E. Y. SIDKY, D. XIA, AND X. PAN, *Image reconstruction and scan configurations*
809 *enabled by optimization-based algorithms in multispectral CT*, Phys. Med. Biol., 62 (2017), pp. 8763–
810 8793.
- 811 [17] C. CHEN AND O. ÖKTEM, *Indirect Image Registration with Large Diffeomorphic Deformations*, SIAM
812 Journal on Imaging Sciences, 11 (2018), pp. 575–617.
- 813 [18] C. CHEN AND G. XU, *Gradient-flow-based semi-implicit finite-element method and its convergence analysis*
814 *for image reconstruction*, Inverse Problems, 28 (2012), p. 035006.
- 815 [19] C. CHEN AND G. XU, *The Linearized Split Bregman Iterative Algorithm and Its Convergence Analysis*
816 *for Robust Tomographic Image Reconstruction*, Tech. Report 13-66, UCLA CAM Reports, 2013.
- 817 [20] C. CHEN AND G. XU, *A new linearized split Bregman iterative algorithm for image reconstruction in*
818 *sparse-view X-ray computed tomography*, Comput. Math. Appl., 71 (2016), pp. 1537–1559.
- 819 [21] K. CHEN AND D. A. LORENZ, *Image sequence interpolation using optimal control*, Journal of Mathemat-
820 ical Imaging and Vision, 41 (2011), pp. 222–238.
- 821 [22] M. DAWOOD, F. BÜTHER, X. JIANG, AND K. P. SCHÄFERS, *Respiratory motion correction in 3-D PET*
822 *data with advanced optical flow algorithms*, IEEE Trans. Med. Imaging, 27 (2008), pp. 1164–1175.
- 823 [23] M. DAWOOD, F. BÜTHER, N. LANG, O. SCHOBER, AND K. P. SCHÄFERS, *Respiratory gating in positron*
824 *emission tomography: a quantitative comparison of different gating schemes*, Med. Phys., 34 (2007),
825 pp. 3067–3076.
- 826 [24] M. DAWOOD, N. LANG, X. JIANG, AND K. P. SCHÄFERS, *Lung motion correction on respiratory gated*
827 *3-d PET/CT images*, IEEE Trans. Med. Imaging, 25 (2006), pp. 1164–1175.
- 828 [25] M. DAWOOD, L. STEGGER, X. JIANG, O. SCHOBER, M. SCHÄFERS, AND K. P. SCHÄFERS, *Optimal*
829 *number of respiratory gates in positron emission tomography: a cardiac patient study*, Med. Phys., 36
830 (2009), pp. 1775–1784.
- 831 [26] P. DUPUIS, U. GRENANDER, AND M. MILLER, *Variational Problems on Flows of Diffeomorphisms for*
832 *Imaging Matching*, Quarterly of Applied Mathematics, 56 (1998), pp. 587–600.

- 833 [27] W. v. ELMPT, J. HAMILL, J. JONES, D. DE RUYSSCHER, P. LAMBIN, AND M. ÖLLERS, *Optimal gating*
834 *compared to 3D and 4D PET reconstruction for characterization of lung tumors*, Eur. J. Nucl. Med.
835 Mol. Imaging, 38 (2011), pp. 843–855.
- 836 [28] H. GAO, J. CAI, Z. SHEN, AND H. ZHAO, *Robust principal component analysis-based four-dimensional*
837 *computed tomography*, Phys. Med. Biol., 56 (2011), pp. 3181–3198.
- 838 [29] F. GIGENGACK, X. JIANG, M. DAWOOD, AND K. SCHÄFERS, *Motion Correction in Thoracic Position*
839 *Emission Tomography*, Springer Briefs in Electrical and Computer Engineering, Springer, 2015.
- 840 [30] F. GIGENGACK, L. RUTHOTTO, M. BURGER, C. H. WOLTERS, X. JIANG, AND K. P. SCHÄFERS, *Motion*
841 *correction in dual gated cardiac PET using mass-preserving image registration*, IEEE Trans. Med.
842 Imaging, 31 (2012), pp. 698–712.
- 843 [31] U. GRENANDER AND M. MILLER, *Pattern Theory. From Representation to Inference*, Oxford University
844 Press, 2007.
- 845 [32] W. GROOTJANS, L.-F. DE GEUS-OEI, A. P. W. MEEUWIS, S. CHARLOTTE, M. GOTTHARDT, W. OYEN,
846 AND E. VISSER, *Amplitude-based optimal respiratory gating in positron emission tomography in pa-*
847 *tients with primary lung cancer*, Eur. Radio., 24 (2014), pp. 3242–3250.
- 848 [33] J. HINKLE, M. SZEGEDI, B. WANG, B. SALTER, AND S. JOSHI, *4D CT image reconstruction with diffe-*
849 *morphic motion model*, Medical Image Analysis, 16 (2012), pp. 1307–1316.
- 850 [34] X. JIA, Y. LOU, B. DONG, Z. TIAN, AND S. JIANG, *4D computed tomography reconstruction from few-*
851 *projection data via temporal non-local regularization*, in Medical Image Computing and Computer-
852 Assisted Intervention – MICCAI 2010, T. Jiang, N. Navab, J. Pluim, and M. Viergever, eds., vol. 6361
853 of Lecture Notes in Computer Science, Springer-Verlag, 2010, pp. 143–150.
- 854 [35] N. LANG, M. DAWOOD, F. BÜTHER, O. SCHOBER, M. SCHÄFERS, AND K. P. SCHÄFERS, *Organ move-*
855 *ment reduction in PET/CT using dual-gated list-mode acquisition*, Med. Phys., 16 (2006), pp. 93–100.
- 856 [36] J. LIU, X. ZHANG, H. ZHAO, Y. GAO, D. THOMAS, D. LOW, AND H. GAO, *5D respiratory motion model*
857 *based image reconstruction algorithm for 4D cone-beam computed tomography*, Inverse Problems, 31
858 (2015), p. 115007.
- 859 [37] G. LUCIGNANI, *Respiratory and cardiac motion correction with 4D PET imaging: shooting at moving*
860 *targets*, Eur. J. Nucl. Med. Mol. Imaging, 36 (2009), pp. 315–319.
- 861 [38] M. I. MILLER, A. TROUVÉ, AND L. YOUNES, *On the metrics and Euler-Lagrange equations of computa-*
862 *tional anatomy*, Annual Reviews of Biomedical Engineering, 4 (2002), pp. 375–405.
- 863 [39] M. I. MILLER, A. TROUVÉ, AND L. YOUNES, *Geodesic shooting for computational anatomy*, Journal of
864 Mathematical Imaging and Vision, 24 (2006), pp. 209–228.
- 865 [40] M. I. MILLER, A. TROUVÉ, AND L. YOUNES, *Hamiltonian systems and optimal control in computational*
866 *anatomy: 100 years since D’arcy Thompson*, Annual Review of Biomedical Engineering, 17 (2015),
867 pp. 447–509.
- 868 [41] F. NATTERER, *The Mathematics of Computerized Tomography*, vol. 32 of Classics in Applied Mathematics,
869 SIAM, 2001.
- 870 [42] H. NIEN AND J. A. FESSLER, *Fast X-ray CT image reconstruction using a linearized augmented La-*
871 *grangian method with ordered subsets*, IEEE Trans. Med. Imag., 34 (2015), pp. 388–399.
- 872 [43] O. ÖKTEM, C. CHEN, N. O. DOMANIÇ, P. RAVIKUMAR, AND C. BAJAJ, *Shape based image reconstruction*
873 *using linearised deformations*, Inverse Problems, 33 (2017), p. 035004.
- 874 [44] L. RITSCHL, S. SAWALL, M. KNAUP, A. HESS, AND M. KACHELRIESS, *Iterative 4D cardiac micro-CT*
875 *image reconstruction using an adaptive spatio-temporal sparsity prior*, Phys. Med. Biol., 57 (2012),
876 pp. 1517–1525.
- 877 [45] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Phys.
878 D, 60 (1992), pp. 259–268.
- 879 [46] O. SCHERZER, M. GRASMAIR, H. GROSSAUER, M. HALTMEIER, AND F. LENZEN, *Variational Methods in*
880 *Imaging*, vol. 167 of Applied Mathematical Sciences, Springer-Verlag, New York, 2009.
- 881 [47] H. SCHUMACHER, J. MODERSITZKI, AND B. FISCHER, *Combined reconstruction and motion correction in*
882 *SPECT imaging*, IEEE Trans. Nucl. Sci., 56 (2009), pp. 73–80.
- 883 [48] A. J. SCHWARZ AND M. O. LEACH, *Implications of respiratory motion for the quantification of 2D MR*
884 *spectroscopic imaging data in the abdomen*, Phys. Med. Biol., 45 (2000), pp. 2105–2116.
- 885 [49] E. SIDKY, J. JØRGENSEN, AND X. PAN, *Convex optimization problem prototyping for image reconstruction*
886 *in computed tomography with the Chambolle–Pock algorithm*, Physics in medicine and biology, 57

- 887 (2012), p. 3065.
- 888 [50] E. SIDKY, C. KAO, AND X. PAN, *Accurate image reconstruction from few-views and limited-angle data*
889 *in divergent-beam CT*, Journal of X-ray Science and Technology, 14 (2006), pp. 119–139.
- 890 [51] K. TAGUCHI AND H. KUDO, *Motion Compensated Fan-Beam Reconstruction for Nonrigid Transformation*,
891 IEEE Trans. Med. Imaging, 27 (2008), pp. 907–917.
- 892 [52] Q. TANG, J. CAMMIN, S. SRIVASTAVA, AND K. TAGUCHI, *A fully four-dimensional, iterative motion*
893 *estimation and compensation method for cardiac CT*, Med. Phys., 39 (2012), pp. 4291–4305.
- 894 [53] D. THOMPSON, *On Growth and Form*, Cambridge University Press, New York, 1945.
- 895 [54] A. TROUVÉ, *Diffeomorphisms Groups and Pattern Matching in Imaging Analysis*, International Journal
896 of Computer Vision, 28 (1998), pp. 213–221.
- 897 [55] A. TROUVÉ AND L. YOUNES, *Shape spaces*, in Handbook of Mathematical Methods in Imaging, S. Otmar,
898 ed., Springer-Verlag, 2015, pp. 1759–1817.
- 899 [56] A. VAN DER GUCHT, B. SERRANO, F. HUGONNET, B. PAULMIER, N. GARNIER, AND M. FARAGGI,
900 *Impact of a new respiratory amplitude-based gating technique in evaluation of upper abdominal PET*
901 *lesions*, Eur. J. Radiol., 83 (2014), pp. 509–515.
- 902 [57] F.-X. VIALARD, L. RISSER, D. RUECKERT, AND C. J. COTTER, *Diffeomorphic 3D image registration via*
903 *geodesic shooting using an efficient adjoint calculation*, International Journal of Computer Vision, 97
904 (2012), pp. 229–241.
- 905 [58] Y. WANG, E. VIDAN, AND G. W. BERGMAN, *Cardiac motion of coronary arteries: variability in the rest*
906 *period and implications for coronary MR angiography*, Radiology, 213 (1999), pp. 751–758.
- 907 [59] Z. WANG, A. C. BOVIK, H. R. SHEIKH, AND E. P. SIMONCELLI, *Image quality assessment: From error*
908 *visibility to structural similarity*, IEEE Trans. Image Process., 13 (2004), pp. 600–612.
- 909 [60] L. YOUNES, *Shapes and Diffeomorphisms*, vol. 171 of Applied Mathematical Sciences, Springer-Verlag,
910 2010.