SIAM J. IMAGING SCIENCES
Vol. 17, No. 3, pp. 1539–1586

Dynamic Image Reconstruction with Motion Priors in Application to Three Dimensional Magnetic Particle Imaging*

Christina Brandt†, Tobias Kluth‡, Tobias Knopp§, and Lena Westen¶

Abstract. Various imaging modalities allow for time-dependent image reconstructions from measurements where its acquisition also has a time-dependent nature. Magnetic particle imaging (MPI) falls into this class of imaging modalities and it thus also provides a dynamic inverse problem. Without proper consideration of the dynamic behavior, motion artifacts in the reconstruction become an issue. More sophisticated methods need to be developed and applied to the reconstruction of the time-dependent sequences of images. In this context, we investigate the incorporation of motion priors in terms of certain flow-parameter-dependent PDEs in the reconstruction process of time-dependent three dimensional (3D) images in magnetic particle imaging. The present work comprises the method development for a general 3D+time setting for time-dependent linear forward operators, analytical investigation of necessary properties in the MPI forward operator, modeling aspects in dynamic MPI, and extensive numerical experiments on 3D+time imaging including simulated data as well as measurements from a rotation phantom and in vivo data from a mouse.

Key words. dynamic 3D image reconstruction, motion models, joint parameter identification, magnetic particle imaging

MSC codes. 94A08, 65K10, 92C50

DOI. 10.1137/23M1580401

Notation.

α	parameter weighting R
β	parameter weighting S
χ_X	characteristic function of a set X
γ	parameter weighting V
δ_X	indicator function of a set X
$\mathbb{R}^+$	$\mathbb{R} \cup \{+\infty\}$
$\mathcal{X}$	spatial subset
Ω	bounded space domain

*Received by the editors June 20, 2023; accepted for publication (in revised form) March 18, 2024; published electronically July 15, 2024.
https://doi.org/10.1137/23M1580401
Funding: The work of the second author was supported by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) project 426078691.

†Department of Mathematics, Universität Hamburg, 20146, Hamburg, Germany. Current address: Institute of Mathematics and Computer Science, University of Greifswald, 17489, Greifswald, Germany (christina.brandt@uni-greifswald.de).
‡Center for Industrial Mathematics, University of Bremen, 28359, Bremen, Germany (tkluth@math.uni-bremen.de).
§Institute for Biomedical Imaging, Hamburg University of Technology and University Medical Center Hamburg-Eppendorf, 22529, Hamburg, Germany (tobias.knopp@tuhh.de).
¶Department of Mathematics, Universität Hamburg, 20146, Hamburg, Germany (lena.westen@uni-hamburg.de).

ε	random noise
A	linear and bounded operator
A_t	linear and bounded operator
c	measureable image function
D	proper, convex, and lower semicontinuous data discrepancy term
h_t	small constant related to time
K	constant
m	motion model
n	number of spatial dimensions
R	spatial regularization term working on image function
S	spatial regularization term working on velocity field
T	endpoint of considered time interval
u	measured data
V	penalty term to incorporate the motion model
v	velocity field
Y	reflexive Banach space
B	linear and bounded operator
f	primal functional
g	functional to be Fenchel-conjugated
$\bar{m}$	mean magnetic moment
Δ_t	longest duration for which the concentration stays approximately static
$\mathcal{L}_\beta$	Langevin function
f_S	sampling frequency
G	full-rank gradient matrix describing the selection field
H	effective magnetic field of the MPI scanner
H_D	drive field of the MPI scanner
H_S	selection field of the MPI scanner
I_{t_s}	interval of nonzero measure including t_s
N_L	number of receive coils of the MPI scanner
s	MPI system function
T_R	repetition time of the MPI scanner
x_s	trajectory of the FFP

1. Introduction. Dynamic inverse problems emerge naturally in several imaging modalities (such as CT, MRI, SPECT, magnetic particle imaging (MPI), etc.) as often both, cause and observation, show a time-dependent behavior. Due to this fact, the field of dynamic inverse problems has been driven strongly by tomographic applications in the past. The dynamic behavior of the object in the time-dependent data acquisition process often leads to motion artifacts when using standard reconstruction techniques designed for static problems. Image quality then suffers such that dedicated methods taking into account the dynamic nature of the problem need to be used and further developed in this setting.

A general regularization framework for dynamic inverse problems is still not established, but several promising attempts are made in this direction. Approaches exploiting concepts like, for example, temporal smoothness [64, 65], dynamic programming techniques [40], explicit deformation models [28], or operator inexactness [5] have been proposed. One rather general approach for time-dependent parameter identification in dynamic systems has been made in [35, 59] by distinguishing reduced and all-at-once approaches.

For specific imaging modalities several methods have been proposed in the literature. Depending on the characteristics of the considered modality, approaches are either based on a variational formulation (e.g., [13, 42, 51, 60]), exact analytic methods (e.g., [14, 16, 27]), iterative methods (e.g., [6, 34]), approximate inversion formulas (e.g., [3, 29, 36, 37]), and recently on machine learning techniques (e.g., [30, 33, 49]). One explicit way to compensate for the dynamics is the incorporation of explicit motion models such as deformation, optical flow, etc., in the image reconstruction process, which then require the identification of motion parameters prior to or within the reconstruction step (see, e.g., [10, 11, 29, 37, 52, 53, 55, 62]). For a more detailed review on variational approaches taking into account motion priors and also learning-based approaches, we refer to the excellent recent survey [31].

The particular imaging modality being considered in the present work is MPI [23]. It is a tracer-based imaging modality exploiting the magnetization behavior of magnetic nanoparticles in a highly dynamic applied magnetic field; see also the surveys [41, 70] for an overview. From a mathematical point of view, the forward operator for static images is given by a linear Fredholm integral operator of the first kind, mapping a spatially distributed concentration (image) to a time-dependent voltage measurement. The static image reconstruction problem is already a severely ill-posed problem [18, 45], which becomes even more ill-posed in the dynamic scenario. So far, MPI has been used primarily for preclinical medical applications but it holds a significant potential for clinical applications illustrated by an increasing number of potential applications, which also rely on the solution of dynamic image reconstruction problems. One application, already suggested at the very beginning, is vascular imaging [23]. In addition, in in vivo experiments, the potential for imaging blood flow was demonstrated using healthy mice [73]. Further promising applications include long-term circulating tracers [39], tracking medical instruments [26], e.g., in angioplasty [63], cancer detection [74], cancer treatment by hyperthermia [58], and stroke monitoring [25, 54].

The time-dependent measurement process of MPI together with dynamic imaging applications requires the consideration of problems across temporal scales. In the field of MPI, a few works considered the dynamic image reconstruction problem therein so far. The general dynamic setting has been investigated from a modeling point of view in [7]. A reconstruction approach designed for periodic motion [22, 21] has been proposed which is restricted to a limited class of motion only. More recently, temporal splines have been included in the reconstruction process to increase the temporal resolution [8]. Motion models have been used for extracting optical flow information from previously obtained image reconstructions [20].

The present work considers the dynamic three dimensional (3D) image reconstruction problem in the imaging modality MPI. In particular, we consider a variational approach taking into account motion priors which allow for a simultaneous identification of parameters therein. This paper is motivated by previous work for time-dependent two dimensional (2D) images in MPI that was published in [42] as a short abstract. The numerical treatment can be interpreted as a special case of the general all-at-once approach discussed in [35, 59]. In contrast to the prior work, we particularly consider the 3D case and focus further on the analysis of the MPI model in the framework of the general joint reconstruction approach. Specific choices of motion priors are an optical flow constraint being suitable, for example, for instrument tracking and mass preservation, e.g., in the case of blood flow visualization. The theoretical part is complemented by the algorithmic solution via alternating minimization

using (stochastic) primal-dual methods combined with multiscale schemes and an intense numerical study taking into account simulated as well as measured 3D phantoms and in vivo experiments evolving in time.

The manuscript is structured as follows: in section 2 the general mathematical setting of the dynamic reconstruction approach is considered and the corresponding theoretical investigation of the MPI model is provided. Section 3 provides the numerical solution to the underlying minimization problem and in section 4 the performance of the proposed method is illustrated in a series of numerical experiments. The manuscript concludes with a conclusion and an outlook in section 5. A glossary is provided in the notation section just after the doi.

2. Mathematical problem setting and solution approach.

We start with a precise definition of the problem and the theoretical investigation of the underlying model. In subsection 2.1 we fix the general setting and define the problem to be solved. Moreover, we state the theorem for existence of minimizers of our problem. In the following subsection 2.2, we introduce a model for the forward operator of MPI and prove compactness as well as some required regularity assumptions. Subsection 2.3 then outlines differences between static and dynamic MPI and derives the associated problems.

2.1. General joint image reconstruction and motion estimation problem.

In the following, we consider a measurable time-dependent image function c on a bounded space-time domain $\Omega \times [0,T] \subset \mathbb{R}^n \times \mathbb{R}^+$, $n \in \{1,2,3\}$, $c \in L^p(0,T;BV(\Omega))$, $p > 1$. We choose the space of functions with bounded variation, as those can describe edges in images and are well suited for images with large homogeneous regions. Both features are particularly important for a wide range of medical images, which is one important field of application we have in mind.

We aim at reconstructing this image function from measured data $u : [0,T] \to Y$, $u \in L^2(0,T;Y)$ for a reflexive Banach space Y, which is obtained by inserting c as an argument into the first component of an operator $A : L^{\hat{p}}\left(0,T;L^l(\Omega)\right) \times [0,T] \to L^2(0,T;Y)$, $\hat{p} = \min(2,p)$, $l \leq \frac{n}{n-1}$, and corrupting the results with random noise $\varepsilon \in L^2(0,T;Y)$, i.e.,

$$A(c,t) + \varepsilon(t) = u(t), \quad t \in [0,T]. \tag{2.1}$$

In the present work, we consider $A(\cdot,t)$ to be a linear and bounded operator for any $t \in [0,T]$. Note that time dependency enters our inverse problem for two reasons: first, the image to be recovered depends on time and, second, the forward operator is time dependent as in other imaging applications such as in blurring of signals with time-dependent convolution kernels or time-dependent Radon data in dynamic CT.

Suppose we consider an ill-posed image reconstruction task, such that we add a spatial regularizer $R : BV(\Omega) \to \mathbb{R}$, which is assumed to be proper, convex, and lower semicontinuous and fulfills

$$R(x) \geq |x|_{BV}^p \qquad \text{for any } x \in \mathrm{BV}(\Omega). \tag{2.2}$$

In addition to the image reconstruction task, we aim at simultaneously determining a velocity field $v : \Omega \times [0,T] \to \mathbb{R}^n$, $v \in L^q(0,T;BV(\Omega)^n)$, $q > 1$, describing the motion in the data. We incorporate v into our model by using a motion model $m(c,v)$ linking images and motion as an additional constraint. Depending on the underlying assumptions, this will be either

an optical flow term or a mass conservation constraint. Again, motion estimation is an ill-posed task in general such that we add a proper, convex, and lower semicontinuous regularizer $S : BV(\Omega)^n \to \mathbb{R}$ working on the motion field. We will restrict ourselves to regularizers which again fulfill

$$S(y) \geq |y|_{BV}^q \qquad \text{for any } y \in \mathrm{BV}(\Omega)^n. \tag{2.3}$$

The joint image reconstruction and motion estimation problem can then be described as

$$\min_{c,v} \int_0^T D(A(c,t), u(t)) + \alpha R(c(\cdot,t)) + \beta S(v(\cdot,t))\,\mathrm{d}t \tag{2.4}$$

$$\text{s.t. } m(c,v) = 0 \quad \text{in } \mathcal{D}'(\Omega \times [0,T]), \tag{2.5}$$

where $D : Y \times Y \to \mathbb{R}$ denotes a proper, convex, and lower semicontinuous discrepancy term between measured and modeled data and $\mathcal{D}'(\Omega \times [0,T])$ denotes the space of distributions, i.e., of compactly supported C^∞ functions on $\Omega \times [0,T]$.

We consider two different motion models: one assuming voxel intensity constancy and one assuming mass preservation during motion. Consider first conservation of the gray value c under motion, i.e.,

$$c(x,t) - c(x + h_t v(x,t), t + h_t) = 0 \tag{2.6}$$

for all $x \in \Omega$, $t \in [0,T]$, $h_t > 0$ small. By Taylor expansion we deduce the optical flow constraint

$$m_1(c,v) = \frac{\partial}{\partial t} c + \nabla c \cdot v = 0. \tag{2.7}$$

To expand our setting to applications where the gray-value constancy is not fulfilled (having in mind for example a divergent blood flow), we introduce the mass conservation assumption, i.e., we assume

$$\int_\Omega c(x,t)\mathrm{d}x = K \qquad \forall t \in [0,T] \tag{2.8}$$

for a constant $K \in \mathbb{R}$. From the local form

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathcal{X}} c(x,t)\mathrm{d}x = \int_{\mathcal{X}} \frac{\partial}{\partial t} c(x,t)\mathrm{d}x \tag{2.9}$$

for an arbitrary subset $\mathcal{X} \subset \Omega$, we can derive the so-called mass conservation constraint by

$$\int_{\mathcal{X}} \frac{\partial}{\partial t} c\mathrm{d}x \overset{!}{=} \int_{\partial\mathcal{X}} cv \cdot \nu \mathrm{d}\mathcal{X} = -\int_{\mathcal{X}} v\nabla c\mathrm{d}x - \int_{\mathcal{X}} c\nabla \cdot v\mathrm{d}x = -\int_{\mathcal{X}} \nabla \cdot (cv)\,\mathrm{d}x, \tag{2.10}$$

where ν denotes the normal vector to $\mathrm{d}\mathcal{X}$. For the first equality, we used that the variation of mass over time must equal the flow through the boundary $\partial\mathcal{X}$. The second equality follows by the divergence theorem. Combining now (2.9) and (2.10) yields the mass conservation constraint

$$m_2(c,v) = \frac{\partial}{\partial t} c + \nabla \cdot (cv) = 0. \tag{2.11}$$

The joint image reconstruction and motion estimation problem are well-defined as minimizers exist for both motion models, as is shown in the following theorem. For clarity, note that in the following the characteristic function of a set X is denoted by χ_X, where

$$\chi_X(x) = \begin{cases} 1, & x \in X, \\ 0, & \text{otherwise}, \end{cases}$$

whereas the indicator function of a set X is denoted by δ_X, i.e.,

$$\delta_X(x) = \begin{cases} 0, & x \in X, \\ \infty, & \text{otherwise}. \end{cases}$$

Theorem 2.1 (existence of minimizers). *Consider the minimization problem* (2.4)–(2.5),

$$\min_{c,v} J(c,v) := \int_0^T D(A(c,t), u(t)) + \alpha R(c(\cdot,t)) + \beta S(v(\cdot,t))\,\mathrm{d}t$$
$$\text{s.t. } m(c,v) = 0 \quad \text{in } \mathcal{D}'(\Omega \times [0,T])$$

in $n \in \{2,3\}$ *dimensions, with* $c \in L^p(0,T;\mathrm{BV}(\Omega))$, $v \in L^q(0,T;\mathrm{BV}(\Omega)^n)$, $u \in L^2(0,T;Y)$ *over the bounded space-time domain* $\Omega \times [0,T]$, $\Omega \subset \mathbb{R}^n$ *for a reflexive Banach space* Y, *and* $1 < p,q < \infty$, $\hat{p} = \min(2,p)$. *Let* $A : L^{\hat{p}}\left(0,T;L^l(\Omega)\right) \times [0,T] \to L^2(0,T;Y)$ *be generated by bounded operators* $A_t : L^l(\Omega) \to Y$, $t \in [0,T]$, $l \le \frac{n}{n-1}$, *such that* $A_t(c(t)) = A(c,t)$ *and let* $A_t \chi_\Omega \neq 0$ *for all* $t \in [0,T]$. *Let* R *and* S *fulfill* (2.2) *and* (2.3), *i.e.,*

$$R(x) \ge |x|_{BV}^p \qquad \text{for any } x \in \mathrm{BV}(\Omega),$$
$$S(y) \ge |y|_{BV}^q \qquad \text{for any } y \in \mathrm{BV}(\Omega)^n.$$

Moreover, let $D(u_1,u_2) = \frac{1}{2}\|u_1 - u_2\|_Y^2$ *be the squared norm distance in* Y *and* m *be either* m_1 *or* m_2 *as defined in* (2.7) *and* (2.11), *respectively. Moreover, assume there exist constants* $k_\infty, k_\theta < \infty$ *such that* $\|v\|_{L^\infty(0,T,L^\infty(\Omega)^n)} \le k_\infty$ *and* $\|\nabla \cdot v\|_\theta \le k_\theta$, *where* $\theta = L^{p^* s}\left(0,T;L^{l^* k}(\Omega)\right)$ *with* $1 < s,k < \infty$, *and* $\frac{1}{p} + \frac{1}{p^*} = 1$.

Then there exists a minimizer of the problem in the set of admissible solutions

$$\left\{ (c,v) \in L^{\hat{p}}(0,T;BV(\Omega)) \times L^q(0,T;BV(\Omega)^n) \;\middle|\; \|v\|_{L^\infty(0,T;L^\infty(\Omega)^n)} \le k_\infty, \right.$$
$$\left. \|\nabla \cdot v\|_\theta \le k_\theta,\ m(c,v) = 0 \text{ in } \mathcal{D}'(\Omega \times [0,T]) \right\}. \tag{2.12}$$

Proof. The proof is based on [11, 17], but we have three main differences in our assumptions. First, our image sequence is in up to $3+t$ dimensions instead of $2+t$; second, our regularizers are of a more general form; and third, our linear operator is time dependent and has a more general range. The proof is stated in Appendix A for completeness. ∎

Remark 2.2. In practice, existence of those constants is no restrictive assumption. Bounding $\|v\|_{L^\infty(0,T;L^\infty(\Omega)^n)}$ is basically assuming a finite maximum speed, which is a physically necessary assumption. Moreover, assuming a bound on $\|\nabla \cdot v\|_\theta$ in an applied sense is bounding the compressibility of the flow, which is again a reasonable assumption.

Remark 2.3. The operator A_t is defined on $L^l(\Omega)$ for $l \leq \frac{n}{n-1}$. This limitation of topologies compatible with the operator is necessary in order to ensure the continuous embedding of $\mathrm{BV}(\Omega)$, which is the spatial domain in which we are searching for a minimizer, into the definition space of the operator.

For computational purposes, either model constraint is incorporated into the variational problem (2.4) by adding an additional penalty term V defined by

$$V\left(c(\cdot,t), \frac{\partial}{\partial t}c(\cdot,t), v(\cdot,t)\right) = \|m_i(c,v)(\cdot,t)\|^s_{L^r(\Omega)}, \quad r,s \geq 1, \quad i \in \{1,2\}, \quad t \in [0,T], \tag{2.13}$$

such that we arrive at the unconstrained minimization problem

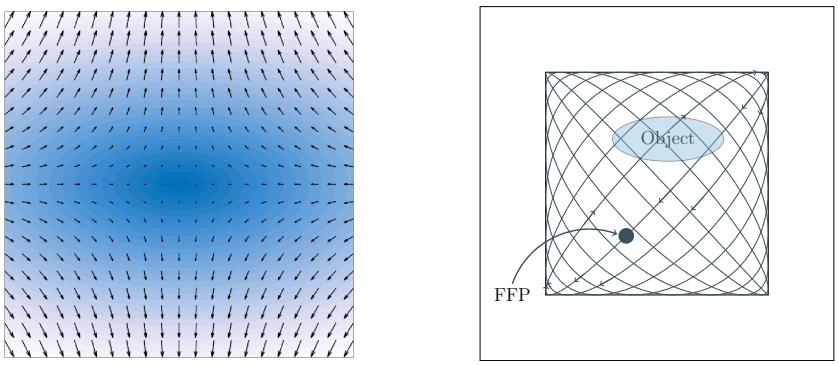
$$\min_{c,v} \int_0^T D(A(c,t), u(t)) + \alpha R(c(\cdot,t)) + \beta S(v(\cdot,t)) + \gamma V\left(c(\cdot,t), \frac{\partial}{\partial t}c(\cdot,t), v(\cdot,t)\right) \mathrm{d}t, \tag{2.14}$$

which is solved by means of alternating minimization. In [11, Lemma 3.7] it is shown that solutions of the unconstrained minimization problem (2.14) converge for $\gamma \to \infty$ to solutions of the constrained minimization problem (2.4), (2.5) in the case of the optical flow constraint in $n = 2$ dimensions. However, the proof which is based on showing Γ-convergence of the objective function of the unconstrained problem to the one of the constrained problem, can be carried out analogously for the mass conservation constraint and in $n = 3$ dimensions. We will apply this approach to the problem of dynamic MPI reconstruction in the following section and thus use the forward operator of MPI as A_t. This operator and some of its properties will be derived in the following.

Remark 2.4. Analysis of the regularization properties of the unconstrained approach (2.14) might be carried out by exploiting the established theory for Tikhonov-type regularization in Banach spaces in the nonlinear problem setting (see, e.g., [32, 66]). PDE-constraint parameter identification in the dynamic setting has also been considered as an all-at-once approach to dynamic inverse problems from a more general point of view [35, 59]. A more general investigation in this direction is beyond the scope of the present work but an extended discussion of this relation can be found in Appendix B.

2.2. Joint reconstruction for MPI. We aim for an application to the imaging modality MPI and thus need a mathematical formulation of the forward operator. The functionality is based on the nonlinear response of magnetic nanoparticles to an applied magnetic field. Modeling the nanoparticles' magnetization behavior sufficiently accurate for imaging tasks when using Lissajous-type excitations is still an open research challenge, but a simplified model, the equilibrium model (see the survey paper [41]), has been extensively used to analyze the system behavior. In line with this, we exploit the equilibrium model to verify necessary requirements of the proposed approach. In order to derive the forward operator of MPI as well as some properties, we start by giving more detailed insights into the basic principles of MPI; see also [46].

In this work, we consider an MPI scanner consisting of a time-independent magnetic field (selection field) $H_S : \mathbb{R}^3 \to \mathbb{R}^3$, which has a field-free point (FFP) in the center of the imaging device; see Figure 1(a). Other geometries like a field-free line can be considered but are



(a) (b)

Figure 1. *The time independent selection field H_S in an MPI scanner has an FFP in the middle and increasing field strength towards the boundaries (Figure 1(a)). The FFP is shifted through the region of interest by time-dependent drive fields. It typically follows a Lissajous-type curve (Figure 1(b)).*

omitted for simplicity. This field is superimposed by time-dependent magnetic fields (drive fields) $H_D : [0,T] \to \mathbb{R}^3$, which shift the FFP in space through the field of view (FOV) $\Omega \subset \mathbb{R}^3$; see Figure 1(b). By superposition, we describe the effective magnetic field $H : \mathbb{R}^3 \times [0,T] \to \mathbb{R}^3$ built by an MPI scanner by

$$H(x,t) = H_D(t) + H_S(x).$$

We assume that the selection field is a linear gradient field (in good approximation), meaning there exists a matrix $G \in \mathbb{R}^{3\times 3}$ such that $H_S(x) = Gx$. As we consider an FFP MPI device, the matrix G needs to be a full-rank matrix. Other geometries like a field-free-line scanner [72] are not considered in this work. The trajectory $x_s : [0,T] \to \mathbb{R}^3$ of the FFP can be described by

$$\begin{aligned} 0 &= H_D(t) + H_S(x) = H_D(t) + Gx_s(t) \\ \Leftrightarrow x_s(t) &= -G^{-1} H_D(t). \end{aligned}$$

Using $H(x,t) = G(x - x_s(t))$, we can compute the time derivative of the magnetic field as $\dot{H}(x,t) = -G\dot{x}_s(t)$.

Typically, by the use of sinusoidal excitation functions, the movement of the FFP follows Lissajous curves. One cycle of the FFP along that curve is performed within the repetition time T_R. We note that in this paper $T_R \ll T$, i.e., we consider that the object is imaged in multiple cycles of the Lissajous curve. The change of magnetization caused by the FFP movement, induces a voltage signal u by the law of induction [46]. This signal $u = (u_i)_{i=1}^{N_L}$ is measured by the N_L receive coils of the scanner and then used for image reconstruction.

The static MPI forward operator $A : L^{\hat{p}}\left(0,T; L^l(\Omega)\right) \times [0,T] \to \left(L^2(0,T;Y)\right)^{N_L}$ for $t \in [0,T]$ for $N_L \in \mathbb{N}$ receive coil units is defined by

(2.15)

$$\begin{aligned} A_i(c_s,t) = -\int_\Omega c_s(x)\,\mu_0 m_0 R_i^T \Bigg[\left(\frac{\mathcal{L}'_\beta\left(\|H(x,t)\|_2\right)}{\|H(x,t)\|_2^2} - \frac{\mathcal{L}_\beta\left(\|H(x,t)\|_2\right)}{\|H(x,t)\|_2^3} \right) H(x,t)\, H(x,t)^T \\ + \frac{\mathcal{L}_\beta\left(\|H(x,t)\|_2\right)}{\|H(x,t)\|_2} I_3 \Bigg] \dot{H}(x,t)\, \mathrm{d}x \end{aligned}$$

for $i = 1, \dots, N_L$, where $\Omega \subset \mathbb{R}^3$ denotes the FOV, $c_s : \Omega \to \mathbb{R}^+$ denotes the static particle concentration of tracer material, μ_0 describes the permeability constant, m_0 is the absolute of a particle's magnetic moment, $R \in \mathbb{R}^{3 \times N_L}$ denotes the spatially homogeneous receive coil sensitivities (R_i being the ith column of R), and $\mathcal{L}_\beta$ describes the dilated Langevin function, which is part of the used equilibrium model assumed as the particle magnetization model. By I_3 we denote the $\mathbb{R}^{3x3}$ identity matrix. The (dilated) Langevin function $\mathcal{L}_\beta : \mathbb{R} \to \mathbb{R}$ is defined by

$$\mathcal{L}_\beta(z) = \begin{cases} \coth(\beta z) - \frac{1}{\beta z} & \text{for} \quad z \neq 0, \\ 0 & \text{else} \end{cases}$$

for a given positive parameter β.

This forward model for static MPI in three spatial dimensions can be similarly expressed by

$$u(t) = \int_\Omega s(x,t) c_s(x) \, \mathrm{d}x \tag{2.16}$$

in the time domain, where $u : [0,T] \to \mathbb{R}^{N_L}$ denotes the induced signal. The system function $s : \Omega \times [0,T] \to \mathbb{R}^{N_L}$ is described by

$$s_i(x,t) = -\mu_0 R_i^T \frac{\partial \bar{m}(x,t)}{\partial t} \tag{2.17}$$

for $i = 1, \dots, N_L$, and $\bar{m} : \Omega \times [0,T] \to \mathbb{R}^3$ denotes the mean magnetic moment of the magnetic nanoparticles, i.e.,

$$\bar{m}(x,t) = m_0 \mathcal{L}_\beta(\|H(x,t)\|_2) \frac{H(x,t)}{\|H(x,t)\|_2}.$$

We introduce this simplified notation in order to shorten the derivation of the dynamic MPI forward model in the following.

Remark 2.5. When measuring more than one drive-field cycle, the acquired signal is T_R-periodic in the static case. In case of low particle concentrations, the measurements are often averaged over some drive-field cycles to get a more robust signal with a higher signal-to-noise ratio (SNR) to reconstruct with, as the method of MPI defines a severely ill-posed problem. Note that we need measurements from at least one complete cycle of length T_R to reconstruct data on the full FOV.

We continue with certain properties of the forward operator, where we consider a specific choice of A_t, respectively, $A_{i,t}$ being discussed in more detail in subsection 2.3 on the dynamic MPI problem. By $I_{t_s} \subset [0,T]$ we denote a compact interval with nonzero measure that includes the time point t_s. We will specify the interval more precisely when describing the different time scales in MPI.

Lemma 2.6 (compactness). *Let $I_{t_s} \subset [0,T]$ be compact and with nonzero measure for arbitrary time point t_s. Further let $\Omega \subset \mathbb{R}^3$ be simply connected and bounded. Moreover, let $\|R_i\|_2 \neq 0$ for $i = 1, \dots, 3$, $x_s \in \mathcal{C}^1(I_{t_s})^3$, and $\dot{x}_s \in \mathcal{C}_b(I_{t_s})^3$, i.e., $\dot{x}_s$ continuous and bounded. Assume that $G \in \mathbb{R}^{n \times n}$ is regular and $c \in L^l(\Omega)$ for arbitrary $l > 1$.*

Then the operators $A_{i,t_s} : L^l(\Omega) \to L^2(I_{t_s})$, $i = 1, \dots, 3$, with $c \mapsto \int_\Omega c(x) s_i(x,t) \, dx$, $t \in I_{t_s}$, are compact.

Proof. Let the operator inside the square brackets in (2.15) be denoted by $F(x,t)$ to reduce the notational complexity. The MPI forward operator for the ith receive coil unit is defined by a linear Fredholm integral equation of first kind with the kernel $s_i : \Omega \times I_{t_s} \to \mathbb{R}$ given by

$$s_i(x,t) = -\mu_0 m_0 R_i^T F(x,t) G\dot{x}_s(t) \quad \text{for } t \in I_{t_s}.$$

By [45, Theorem 4.1.], we deduce that $s_i \in H^0(I_{t_s}; L^\infty(\Omega))$. The assumptions required for the theorem are fulfilled in our setting, as we shortly outline in the following. By $x_s \in \mathcal{C}^1(I_{t_s})$, we know that the drive-field $H_D(t) \in \mathcal{C}^1(I_{t_s})$ and therefore $H_D(t) \in H^1(I_{t_s})$. Moreover, the selection field $H_S(x) = Gx$ for a full rank matrix G fulfills $H_S \in L^\infty(\Omega)^n$ and the receive coil sensitivity R also fulfills $R \in L^\infty(\Omega)^n$. Thus [45, equation (4.3)] yields $s_i \in H^0(I_{t_s}; L^\infty(\Omega)) = L^2(I_{t_s}; L^\infty(\Omega))$, i.e., $s_i \in L^2\left(I_{t_s}; L^{l^*}(\Omega)\right)$ for any $1 \le l^* \le \infty$.

We now consider s_i as a Hilbert–Schmidt integral operator with

$$\|s_i\| := \left(\int_{I_{t_s}} \left(\int_\Omega |s_i(x,t)|^{l^*} \, \mathrm{d}x \right)^{\frac{2}{l^*}} \mathrm{d}t \right)^{\frac{1}{2}} = \|s_i\|_{L^2(I_{t_s}; L^{l^*}(\Omega))} < \infty \tag{2.18}$$

with $\frac{1}{l} + \frac{1}{l^*} = 1$, $1 < l < \infty$. Then by standard results from functional analysis (cf. [1, section 5.12]) it follows that the operator

$$(A_{i,t_s} c)(t) = \int_\Omega s_i(x,t)\, c(x) \, \mathrm{d}x \tag{2.19}$$

defines a compact operator from $L^l(\Omega)$ to $L^2(I_{t_s})$. ∎

Remark 2.7. By the compactness of the MPI forward operator, it is clear that the image reconstruction task is ill-posed. The degree of ill-posedness in terms of decay of the singular values was analyzed in [45]. In a standard setting, using the equilibrium magnetization model, trigonometric FFP trajectories, and a linear selection field, the singular values decay exponentially yielding a severely ill-posed problem.

Remark 2.8. Although Lemma 2.6 is formulated for $n = 3$ dimensions, it can be transferred to a lower dimensional MPI. This can be achieved by assuming the concentration to be Dirac δ-distributed with respect to the orthogonal complement of the lower dimensional affine subspace of $\mathbb{R}^3$. A detailed description and derivation can be found in [45, section 2.2].

Our goal is now to prove the regularity assumption on the forward operator needed for Theorem 2.1, i.e., $A_{t_s}\chi_\Omega \neq 0$ for arbitrary time points t_s. In order to do so, we carry out some analysis on a shifted domain. For fixed t, we use the transformation

$$\Phi_t(\xi) := G^{-1}\xi + x_s(t).$$

Applying the change of variables formula to the forward operator leads to

$$\begin{aligned}
(A_{t_s}\chi_\Omega)(t) &= -\int_\Omega \mu_0 m_0 R^T F(x,t) G\dot{x}_s(t) \, \mathrm{d}x \\
&= -\int_{\Phi_t^{-1}(\Omega)} \mu_0 m_0 R^T F(\Phi_t(\xi), t) \left|\det \nabla \Phi_t(\xi)\right| G\dot{x}_s(t) \, \mathrm{d}\xi \\
&= -\int_{\Omega_t} \mu_0 m_0 R^T \tilde{F}(\xi) \left|\det G^{-1}\right| G\dot{x}_s(t) \, \mathrm{d}\xi
\end{aligned}$$

with $\Omega_t := \Phi_t^{-1}(\Omega) = G(\Omega - x_s(t))$, $\nabla\Phi_t(\xi) = G^{-1}$, and

$$\tilde{F}(\xi) := \left(\frac{\mathcal{L}_\beta'(\|\xi\|_2)}{\|\xi\|_2^2} - \frac{\mathcal{L}_\beta(\|\xi\|_2)}{\|\xi\|_2^3}\right)\xi\xi^T + \frac{\mathcal{L}_\beta(\|\xi\|_2)}{\|\xi\|_2} I_3. \tag{2.20}$$

For this matrix we can verify the following auxiliary lemma.

Lemma 2.9 (positive definiteness). *Let $\xi \in \mathbb{R}^3 \setminus \{0\}$ be arbitrary. Then $\tilde{F}(\xi) \in \mathbb{R}^{3\times 3}$ is positive definite.*

Proof. We consider $\tilde{F}$ as defined in (2.20). Let $\xi \in \mathbb{R}^3 \setminus \{0\}$ and $x \in \mathbb{R}^3$ arbitrary. Consider first

$$x^T \left(\frac{\mathcal{L}_\beta'(\|\xi\|_2)}{\|\xi\|_2^2} - \frac{\mathcal{L}_\beta(\|\xi\|_2)}{\|\xi\|_2^3}\right)\xi\xi^T x = \left(\mathcal{L}_\beta'(\|\xi\|_2) - \frac{\mathcal{L}_\beta(\|\xi\|_2)}{\|\xi\|_2}\right)\frac{(x^T\xi)^2}{\|\xi\|_2^2}.$$

Moreover,

$$x^T \frac{\mathcal{L}_\beta(\|\xi\|_2)}{\|\xi\|_2} I_3 x = \frac{\mathcal{L}_\beta(\|\xi\|_2)}{\|\xi\|_2} x^T x$$

yields

$$\begin{aligned} x^T\tilde{F}(\xi)\,x &= \left(\mathcal{L}_\beta'(\|\xi\|_2) - \frac{\mathcal{L}_\beta(\|\xi\|_2)}{\|\xi\|_2}\right)\frac{(x^T\xi)^2}{\|\xi\|_2^2} + \frac{\mathcal{L}_\beta(\|\xi\|_2)}{\|\xi\|_2}\|x\|_2^2 \\ &= \underbrace{\mathcal{L}_\beta'(\|\xi\|_2)}_{>0}\underbrace{\frac{(x^T\xi)^2}{\|\xi\|_2^2}}_{\geq 0} + \underbrace{\frac{\mathcal{L}_\beta(\|\xi\|_2)}{\|\xi\|_2}}_{>0}\underbrace{\left(\|x\|_2^2 - \frac{(x^T\xi)^2}{\|\xi\|_2^2}\right)}_{\geq 0,\ \text{as } (x^T\xi)^2 \leq \|\xi\|_2^2\|x\|_2^2}, \end{aligned}$$

where we used the positivity of the Langevin function everywhere but in 0, the positivity of the derivative of the Langevin function, and the Cauchy–Schwarz inequality. We observe the following: if $\frac{(x^T\xi)^2}{\|\xi\|_2^2} = 0$, it follows that $(x^T\xi)^2 = 0$ and thus $(\|x\|_2^2 - \frac{(x^T\xi)^2}{\|\xi\|_2^2}) \neq 0$ if $x \neq 0$. If $(\|x\|_2^2 - \frac{(x^T\xi)^2}{\|\xi\|_2^2}) = 0$, then $\frac{(x^T\xi)^2}{\|\xi\|_2^2} \neq 0$ for $x \neq 0$. We have thus shown

$$x^T\tilde{F}(\xi)\,x > 0 \qquad \text{for } x \neq 0. \qquad \blacksquare$$

This allows us now to verify the desired regularity property of the forward operator.

Theorem 2.10 (regularity of the MPI forward operator). *Let $I_{t_s} \subset [0,T]$ be compact and with nonzero measure for arbitrary time point t_s, $\Omega \subset \mathbb{R}^3$ be simply connected and bounded. Assume that the coil sensitivities $R_i \in \mathbb{R}^3$, $i = 1,\dots,3$, fulfill*

$$\operatorname{span}\{R_i,\ i = 1,\dots,3\} = \mathbb{R}^3.$$

Moreover, let $x_s \in \mathcal{C}^1(I_{t_s})$ and assume there exists an inner point $t^* \in I_{t_s}$ such that $\|\dot{x}_s(t^*)\|_2 \neq 0$. Let the operator $A_{t_s} : L^l(\Omega) \to L^2(I_{t_s})^3$, be given by $c \mapsto \left(\int_\Omega c(x)s_i(x,t)\,dx\right)_{i=1,\dots,3}$, $t \in I_{t_s}$, according to (2.17).

Then it holds that

$$\|A_{t_s}\chi_\Omega\|_{L^2(I_{t_s})^3} > 0$$

and thus $A_{t_s}\chi_\Omega \neq 0$.

Proof. In the following, we denote $y_s := Gx_s$. As $\{R_i,\ i=1,\dots,3\}$ forms a basis of $\mathbb{R}^3$, it holds that

$$\forall t \in I_{t_s}: \quad \exists q_i \in \mathbb{R},\ i=1,\dots,3: \quad \sum_{i=1}^{3} q_i R_i = \dot{y}_s(t).$$

Therefore, it exists $\tilde{q}_i \in \mathcal{C}(I_{t_s}),\ i=1,\dots,3:\ \sum_{i=1}^{3} \tilde{q}_i R_i = \dot{y}_s$. Setting $T(R) := [R_1\ R_2\ R_3]$, continuity of those functions $\tilde{q}_i$ follows from

$$[R_1\ R_2\ R_3]\,\tilde{q} = \dot{y}_s \quad \Rightarrow \tilde{q} = T(R)^{-1}\dot{y}_s,$$

as $\left\|T(R)^{-1}\right\| < \infty$. For all t with $\dot{y}_s(t) \neq 0$, it holds that

$$\sum_{i=1}^{3} \int_{\Omega_t} \tilde{q}_i(t)\, R_i^T \tilde{F}(\xi)\, \dot{y}_s(t)\, \mathrm{d}\xi = \int_{\Omega_t} \underbrace{\dot{y}_s(t)^T \tilde{F}(\xi)\, \dot{y}_s(t)}_{>0\ \forall \xi}\, \mathrm{d}\xi > 0. \tag{2.21}$$

For $t^* \in I_{t_s}$ with $\dot{y}_s(t^*) \neq 0$ there exists, by the continuity of $\dot{x}_s$, $\varepsilon > 0$ such that

$$\begin{aligned}
0 &< \int_{(t^*-\varepsilon, t^*+\varepsilon)} \left(\sum_{i=1}^{3} \tilde{q}_i(t) \underbrace{\int_{\Omega_t} R_i^T \tilde{F}(\xi)\, \dot{y}_s(t)\, \mathrm{d}\xi}_{:=u_i(t)} \right)^2 \mathrm{d}t = \left\| \sum_{i=1}^{3} \tilde{q}_i u_i \right\|^2_{L^2(t^*-\varepsilon,\ t^*+\varepsilon)} \\
&\leq 4 \sum_{i=1}^{3} \|\tilde{q}_i u_i\|^2_{L^2(t^*-\varepsilon,\ t^*+\varepsilon)} \leq 4 \sum_{i=1}^{3} \|\tilde{q}_i u_i\|^2_{L^2(I_{t_s})} \\
&\leq 4 C_{\tilde{q}} \sum_{i=1}^{3} \|u_i\|^2_{L^2(I_{t_s})} = 4 C_{\tilde{q}} \|u\|^2_{L^2(I_{t_s})},
\end{aligned}$$

where $C_{\tilde{q}} > 0$ denotes an upper bound on $\|\tilde{q}_i\|_{L^2(I_{t_s})}$ for $i=1,..,3$, which exists as $\tilde{q}_i \in \mathcal{C}(I_{t_s})$. Denoting now

$$(A_{t_s}\chi_\Omega)(t) = -\int_{\Omega_t} \mu_0 m_0 R^T \tilde{F}(\xi) \left|\det\ G^{-1}\right| G\dot{x}_s(t)\, \mathrm{d}\xi = -\mu_0 m_0 \left|\det\ G^{-1}\right| u(t) \tag{2.22}$$

yields the result $\|A_{t_s}\chi_\Omega\|_{L^2(I_{t_s})} > 0$ and thus $A_{t_s}\chi_\Omega \neq 0$. ∎

Remark 2.11. A standard MPI scanner usually consists of orthogonal receive coils such that $\{R_i, i=1,\dots,3\}$ even forms an orthogonal basis of $\mathbb{R}^3$.

Remark 2.12. By assuming $\|\dot{x}_s(t)\|_2 \neq 0$ for all $t \in [0,T]$, it can also be shown, by the same arguments, that for the particular choice ofv $I_{t_s} = \{t_s\}$ the operator $A_{t_s} : L^l(\Omega) \to \mathbb{R}^3$ with $c \mapsto \left(\int_\Omega c(x) s_i(x,t_s)\, \mathrm{d}x\right)_{i=1,\dots,3}$ fulfills $(A_{t_s}\chi_\Omega) \neq 0$ for all $t_s \in [0,T]$. For our application, this stronger assumption can be fulfilled by choosing an MPI scanner with sinusoidal excitation frequencies for the drive field. This choice results in cosine functions for $\dot{x}_s(t)$, which do not share zeros within the time span $[0,T]$ by the choice of the frequency dividers. However, cosine excitation functions lead to sinusoidal derivatives, which share a zero at $t=0$ and thus do not fulfill the stronger assumption in general.

We have thus shown that the static MPI forward operator for a scanner with sinusoidal excitation fulfills the regularity assumption $(A_{t_s}\chi_\Omega) \neq 0$ as well as boundedness of the operators A_{t_s} for desired t_s, which is necessary to prove the existence of a minimizer in a joint motion estimation and image reconstruction problem. Note that the consideration of measurements in $L^2(I_{t_s})$ for time points t_s already indicates that in the context of MPI we need to carefully consider time on different scales as MPI is characterized by a time-dependent measurement process. If one would assume motion in c and the measurement process providing a measurement at each time on the same temporal scale, this would require reconstructing an entire spatially resolved image from one measurement in $\mathbb{R}^3$. The issue of different temporal scales will be addressed in the following part on the dynamic MPI problem.

2.3. Dynamic MPI. In the following, we will derive the dynamic MPI forward model in order to understand model inaccuracies and modeling errors, which occur in dynamic MPI.

Extending the MPI forward model to a dynamic tracer distribution $c : \Omega \times [0,T] \to \mathbb{R}^+$ and assuming a homogeneous coil sensitivity leads to

$$\begin{aligned} u_i(t) &= -\mu_0 R_i^T \frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega c(x,t)\bar{m}(x,t)\mathrm{d}x \\ &= -\mu_0 R_i^T \int_\Omega c(x,t)\frac{\partial}{\partial t}\bar{m}(x,t) + \bar{m}(x,t)\frac{\partial}{\partial t}c(x,t)\mathrm{d}x \end{aligned} \tag{2.23}$$

for $i = 1, \ldots, N_L$. This model was first mentioned in [21] but was applied to the special case of periodic motions. It was proposed in [7] to use the full dynamic model (2.23) for image reconstruction. However, although the reconstructions are of high quality, they are computationally too demanding to be implemented in practice [7]. Moreover, in the case of a fully calibrated system function it is still an unsolved problem how to obtain the full dynamic model. Limiting ourselves to dynamic distributions whose temporal change provides negligible contributions to the signal, yields

$$u_i(t) \approx -\mu_0 R_i^T \int_\Omega c(x,t)\frac{\partial}{\partial t}\bar{m}(x,t)\mathrm{d}x = \int_\Omega s_i(x,t)c(x,t)\mathrm{d}x. \tag{2.24}$$

This is a common assumption in dynamic MPI reconstruction which has also been used for periodic motion using virtual frames [21]. In the following, we will reconstruct dynamic concentrations with nonperiodic motion by suppressing artifacts by joint reconstruction of concentration and motion.

One important aspect which need to be taken into account in this context is the interplay of different temporal scales, a fine temporal scale of the measurement process itself, i.e., we need to consider time series of voltage measurements to gather the spatial information of the image, and the temporal scale of the motion. For this we denote the longest duration, for which the tracer distribution remains approximately static by Δt, i.e., we exploit the approximation

$$c(x,t) \approx c(x,t_s) \qquad \forall t \in [t_s - \Delta t, t_s] \ \forall x \in \Omega. \tag{2.25}$$

Typically, a reconstruction of the concentration $c(x,t_s)$ at time t_s requires at least data from a complete drive-field cycle, i.e., $\Delta t \geq T_R$.

This poses no problem if the motion is sufficiently slow. For $\Delta t \geq T_R$ we have a quasi-static tracer distribution that changes only slightly during one drive-field cycle and can be handled as static data in a good approximation.

However, if $\Delta t < T_R$, we face several problems which affect each other. First, additional motion artifacts occur as the tracer distribution changes within the measurement process. Second, we have a loss of spatial resolution as the FFP trajectory is not closed for specific tracer distributions and potentially large regions are not scanned for certain tracer distributions. Moreover, we currently reconstruct in the Fourier domain and not in the time domain. This is possible, as we consider data from the whole repetition time T_R as data for one time step. If we want to perform reconstruction on a finer temporal grid, we will have to window the measurements appropriately, since the signal is not periodic anymore, which poses a new research problem on its own.

The setting $\Delta t < T_R$ is certainly of practical relevance, e.g., for the application of cardiovascular imaging. Mean velocities in the aorta are $12\,\mathrm{cm\,s^{-1}}$, in arteries even $45\,\mathrm{cm\,s^{-1}}$ [19] which corresponds to 2.6 mm $\mathrm{T_R}$ and 9.7 mm/$\mathrm{T_R}$, respectively, in a standard 3D MPI scanner. However, in the following we restrict ourselves to the assumption $\Delta t = T_R$ which guarantees encoding the maximum amount of spatial information in the desired MPI setting but where with potentially higher probability motion artifacts emerge. As a result we consider

$$(A_{t_s} c)(t) = \int_\Omega s(x,t)\, c(x,t_s)\, \mathrm{d}x, \quad t \in I_{t_s} := [t_s - T_R, t_s], \tag{2.26}$$

for timepoint $t_s \geq T_R$ to reconstruct $c(x,t_s)$ from u obtained on I_{t_s}. In order to reduce the effects of this approximation in our reconstruction, we jointly reconstruct images and the motion in-between time frames such that both tasks will endorse each other. If images and motion are sufficiently linked and gaps arising in the data are sufficiently interpolated, one might also use subframe reconstruction in the case of dynamic data (cf. Figure 2). In this approach, we use measurements obtained during an interval of maximum length $\Delta t < T_R$ for reconstruction of the concentration which on its own will not be sufficient but can be so, if subsequent frames are tied strong enough.

For numerical reconstruction, we need discretization in space and time. For the time discretization, the finest time scale we can use is defined by the analog-to-digital converter which converts the analog measured signal to a digital one. The different time scales as described above are illustrated in Figure 2; we are in the quasi-static setting, if not mentioned explicitly. The discretization in space is automatically obtained by the calibration of the system matrix.

We note that typical regularization terms for the image reconstruction task in MPI are covered by assumption (2.2). Depending on the specific application for which MPI is used, either the optical flow constraint (i.e., instrument tracking) or the mass conservation constraint (i.e., blood flow visualization) can be an appropriate choice for the motion model. In any case, the existence of a minimizer is guaranteed by Theorem 2.1. Before we state our numerical results for MPI reconstruction, we give some more details on the algorithms used to solve the minimization problem (2.14).

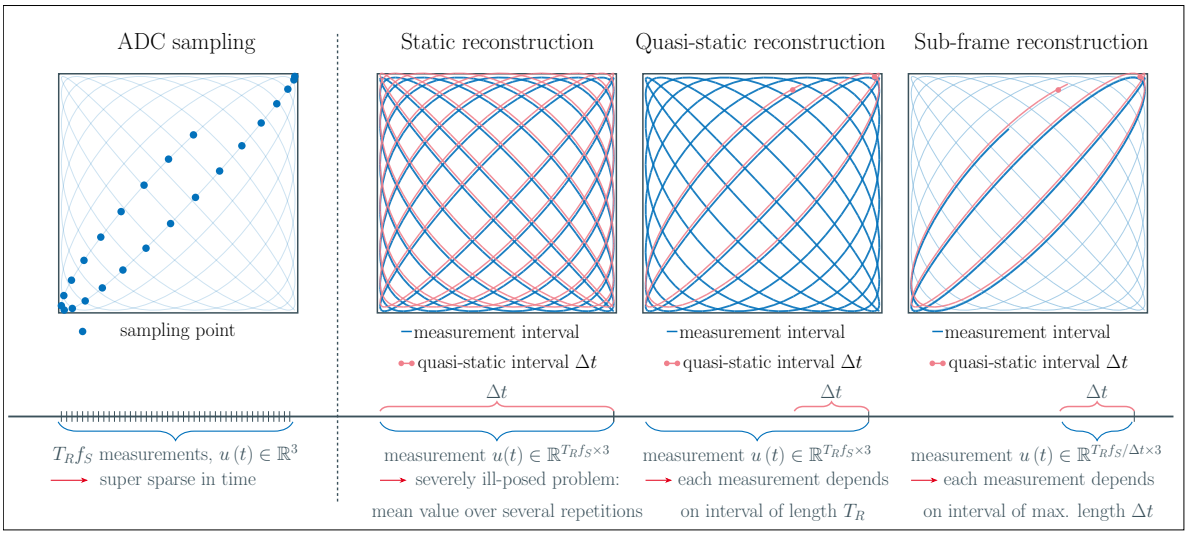


Figure 2. *The finest time scale in MPI reconstruction is defined by the sampling rate of the analog-to-digital converter, which samples measurements at a frequency f_S during the repetition time T_R. In static reconstruction, we pool all measurements obtained during one drive-field cycle and then average over the drive-field cycles in order to gain data with high SNR. The measurement interval is at most as long as the quasi-static interval. In dynamic reconstruction, the quasi-static interval Δt can be smaller than T_R. Still, we need to use data from a full drive-field cycle in order to reconstruct the full FOV (cf. (2.26)). To reduce the computational complexity and reduce motion artifacts, one could use subframe reconstruction in combination with coupling of reconstructions at the different time steps through motion. Pooled measurements are then obtained during an interval of maximum length Δt.*

3. Algorithm. To solve problem (2.14), i.e.,

$$\min_{c\geq 0,v} \int_0^T D\left(A\left(c,t\right),u\left(t\right)\right)+\alpha R\left(c(\cdot,t)\right)+\beta S\left(v(\cdot,t)\right)+\gamma V\left(c\left(\cdot,t\right),\frac{\partial}{\partial t}c\left(\cdot,t\right),v\left(\cdot,t\right)\right)\mathrm{d}t,$$

we apply alternating minimization, i.e., we solve for tracer concentration c and motion (displacement field) v in between subsequent frames alternately following the idea of [11]. We consider the algorithmic approaches to the two subproblems in the following sections. We derive the algorithmic details for $n=3$ spatial dimensions which then includes lower dimensional problems.

3.1. The motion estimation subproblem. For a given concentration sequence c, the motion estimation subproblem is defined by

$$\min_{v} \int_0^T \beta S\left(v(\cdot,t)\right)+\gamma V\left(c\left(\cdot,t\right),\frac{\partial}{\partial t}c\left(\cdot,t\right),v\left(\cdot,t\right)\right)\mathrm{d}t. \tag{3.1}$$

We derive an algorithm for each of the considered motion models.

3.1.1. The optical flow motion model. We start with the optical flow constrained problem. For a spatially but not time-dependent regularization term S our minimization problem reads

$$\min_v \int_0^T \beta S(v(\cdot,t)) + \gamma \left\| \frac{\partial}{\partial t} c(\cdot,t) + \nabla c(\cdot,t) \cdot v(\cdot,t) \right\|_{L^1(\Omega)} \mathrm{d}t, \tag{3.2}$$

where the L^1-norm could be easily replaced by an L^2-norm. However, we choose an L^1-norm so as to allow for outliers, which might exist due to noise artifacts in the reconstructed image sequences. The regularization functional S is assumed to be proper, lower semicontinuous, and convex as well as prox-tractable. For our application, we usually consider total variation (TV), i.e., L^1- or L^2-regularization of the gradient, respectively, i.e.,

$$S_1(v(\cdot,t)) = \|\nabla v(\cdot,t)\|_{L^1(\Omega)^n} \qquad \text{or} \qquad S_2(v(\cdot,t)) = \frac{1}{2}\|\nabla v(\cdot,t)\|^2_{L^2(\Omega)^n}. \tag{3.3}$$

In order to solve (3.2) with regularization functional $S = S_1$, we apply primal-dual splitting with primal functional f and g to be Fenchel-conjugated defined by

$$f(v) = \int_0^T \gamma \left\| \frac{\partial}{\partial t} c(\cdot,t) + \nabla c(\cdot,t) \cdot v(\cdot,t) \right\|_{L^1(\Omega)} \mathrm{d}t,$$
$$g(Bv) = \int_0^T \beta \|\nabla v(\cdot,t)\|_{L^1(\Omega)^n} \mathrm{d}t$$

with the linear operator

$$B = \begin{pmatrix} \nabla & 0 & 0 \\ 0 & \nabla & 0 \\ 0 & 0 & \nabla \end{pmatrix}$$

leading to the adjoint operator

$$B^* = \begin{pmatrix} -\mathrm{div} & 0 & 0 \\ 0 & -\mathrm{div} & 0 \\ 0 & 0 & -\mathrm{div} \end{pmatrix}.$$

The corresponding Fenchel-conjugated functional is given by

$$g^*(y_1,y_2,y_3) = \int_0^T \beta \left(\delta_{\{\|y_1\|_\infty \le \beta\}} + \delta_{\{\|y_2\|_\infty \le \beta\}} + \delta_{\{\|y_3\|_\infty \le \beta\}} \right) \mathrm{d}t.$$

3.1.2. The mass conservation motion model. The motion estimation problem in this case is given by

$$\min_v \int_0^T \frac{\beta}{2} \|\nabla v(\cdot,t)\|^2_{L^2(\Omega)^n} + \gamma \left\| \frac{\partial}{\partial t} c(\cdot,t) + \nabla \cdot (c(\cdot,t) \cdot v(\cdot,t)) \right\|_{L^1(\Omega)} \mathrm{d}t, \tag{3.4}$$

where we choose an L^2-norm on the gradient in order to derive the algorithmic scheme for both regularizers. In the numerical part of this manuscript, we will state clearly which regularization is used and we might combine any of the regularizers with each motion model. The problem is again solved by primal-dual splitting with

$$f(v) = 0,$$
$$g(Bv) = \int_0^T \frac{\beta}{2} \|\nabla v(\cdot,t)\|^2_{L^2(\Omega)^n} + \gamma \left\| \frac{\partial}{\partial t} c(\cdot,t) + \nabla \cdot \left(c(\cdot,t) \cdot v(\cdot,t) \right) \right\|_{L^1(\Omega)} \mathrm{d}t$$

with the linear operator

$$B = \begin{pmatrix} \nabla & 0 & 0 \\ 0 & \nabla & 0 \\ 0 & 0 & \nabla \\ \partial_x c & \partial_y c & \partial_z c \end{pmatrix}$$

using the notation $(\partial_x c)(v) = \frac{\partial}{\partial x}(cv)$ and its corresponding adjoint operator

$$B^* = \begin{pmatrix} -\mathrm{div} & 0 & 0 & -c\partial_x \\ 0 & -\mathrm{div} & 0 & -c\partial_y \\ 0 & 0 & -\mathrm{div} & -c\partial_z \end{pmatrix}.$$

This leads to the Fenchel-conjugate

$$g^*(y_1, y_2, y_3, y_4) = \int_0^T \frac{1}{2\beta}\left(\|y_1\|^2_{L^2(\Omega)^n} + \|y_2\|^2_{L^2(\Omega)^n} + \|y_3\|^2_{L^2(\Omega)^n}\right) + \gamma\left(\delta_{\{\|y_4\|_\infty \le \gamma\}} - \left\langle \frac{\partial}{\partial t} c, y_4 \right\rangle\right) \mathrm{d}t.$$

For both motion models, we can now state the saddle point problem equivalent to the original minimization problem

$$\min_v f(v) + g(Bv).$$

More particularly, the saddle point problem is given by

$$\min_v \sup_y \langle Bv, y\rangle - g^*(y) + f(v).$$

In imaging applications, such a problem is often solved by the primal-dual hybrid gradient (PDHG) method introduced by [61].

3.1.3. PDHG algorithm for motion estimation. Using the above notations, the main motion estimation algorithm is given by Algorithm 3.1.

The proximal operator $\mathrm{prox}_{\tau f}$, which occurs in the algorithm, is defined by

$$\mathrm{prox}_{\tau f}(x) = \arg\min_u f(u) + \frac{1}{2\tau}\|u - x\|^2,$$

where the norm is chosen according to the underlying spaces.

Algorithm 3.1. PDHG algorithm for motion estimation.

1: **input** initial value $\mathbf{v}^0 = (u, v, w)$, $\tilde{v}^0 = v^0$, initial value y^0, θ, step size parameters σ, τ, operators B, B^*
2: **for** k = 0,1,2,...,K-1 **do**
3: $\quad \bar{y}^{k+1} = y^k + \sigma B\tilde{\mathbf{v}}^k$
4: $\quad y^{k+1} = \mathrm{prox}_{\sigma g^*}(\bar{y}^{k+1})$ *% dual update*
5: $\quad \bar{\mathbf{v}}^{k+1} = \mathbf{v}^k - \tau B^* y^{k+1}$
6: $\quad \mathbf{v}^{k+1} = \mathrm{prox}_{\tau f}(\bar{\mathbf{v}}^{k+1})$ *% primal update*
7: $\quad \tilde{\mathbf{v}}^{k+1} = \mathbf{v}^{k+1} + \theta(\mathbf{v}^{k+1} - \mathbf{v}^k)$ *% extrapolation step*
8: **end for**

3.1.4. Operator discretization. As proposed by [11], we use forward and backward differences for discretization of the differential operator and its adjoint, respectively, for the spatial regularization terms. Within the motion constraint, we use central differences, which are self-adjoint, for the spatial derivatives and forward differences for the time derivatives to obtain a stable scheme.

3.1.5. Coarse-to-fine strategy and warping. The problem of handling large-scale motion is twofold. First, the linearized optical flow constraint is derived by a Taylor approximation, which will hold only for motion of small absolute value. Second, the absolute value of our motion estimates is bounded by the choice of our discrete differential operator.

There are two different basic approaches to the multiscale strategy. One either computes only motion increments on each scale and warps images in-between the different scales [9] or one uses the results from coarser scales for initialization but then updates the whole value in each iteration [56].

The optical flow motion model: Multiscale and warping. We use the first approach in the case of the optical flow constraint as it was shown that the warping technique implements the nonlinearized optical flow equation [9]. The process of motion estimation with a coarse-to-fine strategy and warping is illustrated in Figure 3.

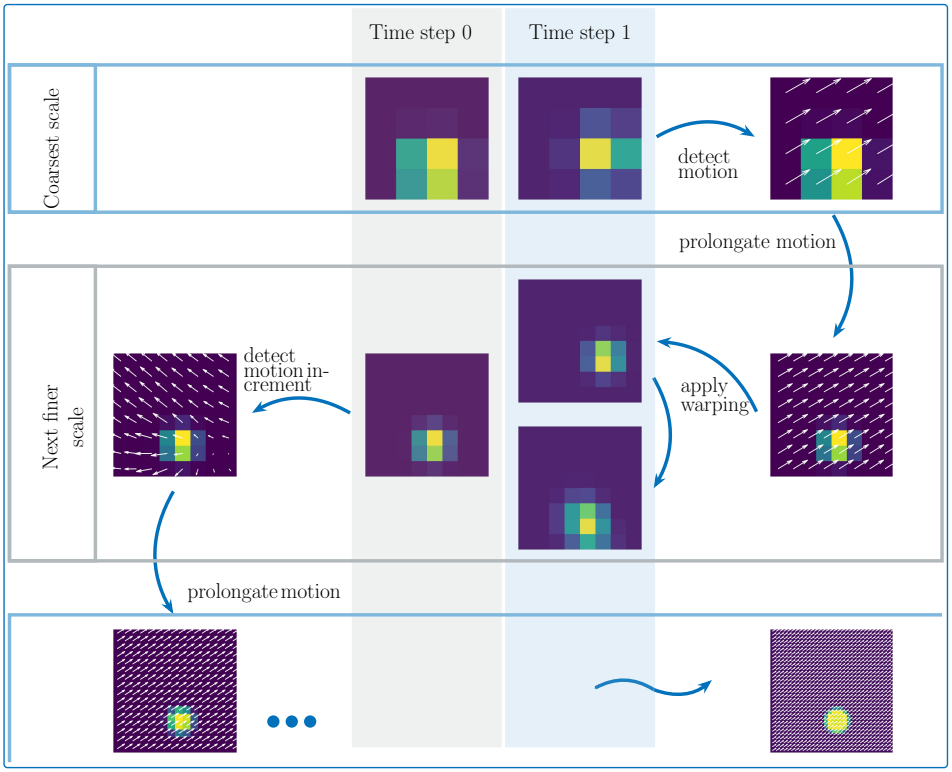


Figure 3. *The multiresolution framework in combination with warping works as follows. Starting from the coarsest scale, motion in-between the images corresponding to two succeeding time steps is detected by a motion estimation algorithm. The flow field is then prolongated to the next finer scale and applied to the image of the second time step in terms of warping. These steps are repeated on each scale.*

First, the image sequence is downsampled to the different scales yielding an image pyramid. Then, beginning on the coarsest scale, the motion in-between succeeding frames is computed and prolongated to the next finer scale. It is then applied to the image of the second time step by warping. The whole procedure is then repeated on each scale for the motion increment until the finest scale is reached. Note that image warping can be implemented by a forward mapping or a reverse mapping. When using a forward mapping, we iterate over the pixels of the source image and compute the new position after the motion is applied. The new position is then a floating point and we use linear interpolation to regrid the image. However, when using the forward mapping many source pixels can map to the same destination and some pixels in the destination image might not be covered at all. For this reason, warping is usually defined as a reverse mapping. We then iterate over the pixels of the destination image and find the origin $(x, y, z)_{old}$ in the source image for every pixel. The gray value to be transported to the destination image is then again determined by linear interpolation.

Reverse warping thus in a sense fulfills the gray value constancy assumption and can be combined with the optical flow constraint, as that constraint will lead to invertible flow fields. However, in the case of mass conservation as a motion model, there might exist sources and sinks and the flow field is then not invertible.

The mass conservation motion model: Multiscale only. In the case of the mass conservation constraint, we cannot use warping due to the aforementioned reasons. However, we need a multiscale strategy in order to deal with large motion estimates. We thus implement a coarse-to-fine strategy, where the motion estimate on each scale is initialized by the result of the last scale and updated after the increment is computed.

The algorithm used on each scale is the PDHG algorithm as explained in subsection 3.1. This allows for straightforward integration of estimates on previous scales.

3.2. The image reconstruction subproblem. The image reconstruction subproblem for a given motion estimate v is defined by

$$\min_{c\geq 0} \int_0^T D\left(A\left(c,t\right),u\left(t\right)\right) + \alpha R\left(c(\cdot,t)\right) + \gamma V\left(c\left(\cdot,t\right), \frac{\partial}{\partial t} c\left(\cdot,t\right), v\left(\cdot,t\right)\right) \mathrm{d}t\,. \tag{3.5}$$

We use fused lasso regularization as our default regularization on the concentration as that fits the expected behavior of the tracer material in MPI very well [67, 75] and also incorporate the nonnegativity constraint into the regularization. We thus have

$$\alpha R\left(c(\cdot,t)\right) = \alpha_1 \left\|c\left(\cdot,t\right)\right\|_{L^1(\Omega)} + \alpha_2 \left\|\nabla c\left(\cdot,t\right)\right\|_{L^1(\Omega)^n} + \delta_{c(\cdot,t)\geq 0}. \tag{3.6}$$

For the data fitting term, we choose either the L^1- or L^2-norm discrepancy [44], i.e.,

$$D_1\left(A\left(c,t\right),u\left(t\right)\right) = \left\|A\left(c,t\right) - u\left(t\right)\right\|_{L^1(\mathbb{R}^d)},$$
$$D_2\left(A\left(c,t\right),u\left(t\right)\right) = \frac{1}{2}\left\|A\left(c,t\right) - u\left(t\right)\right\|^2_{L^2(\mathbb{R}^d)}.$$

Either data fitting term can be combined with each motion model. For illustration, we derive one algorithm for each motion model for data fitting term D_1 in the following.

3.2.1. The optical flow motion model. As we expect the motion to exceed one pixel per frame, the linearized constraint

$$m_1(c,v) = \frac{\partial}{\partial t}c + \nabla c \cdot v = 0$$

does not hold.

We therefore implement the nonlinearized version of the gray-value constancy assumption according to (2.6), i.e., in the semidiscretized setting with respect to time step h_t we use

$$V(c(\cdot,t), c(\cdot,t+h_t), v(\cdot,t)) = \|c(\cdot + h_t v(\cdot,t), t+h_t) - c(\cdot,t)\|_{L^1(\Omega)} = \|\mathcal{W}c(t)\|_{L^1(\Omega)}$$

with

$$(\mathcal{W}c)(t) := -c(\cdot,t) + W^{t,h_t} c(\cdot, t+h_t), \tag{3.7}$$

where W^{t,h_t} maps a space-dependent image to a space-dependent image performing the reverse warping of the input image with respect to the displacement field estimate $v(\cdot,t)$. Similarly to the motion estimation subproblem, reverse warping (iterating over the destination image) is performed in order to cover all pixels of the destination image.

Applying primal-dual splitting to this setting and using the same notations as in the motion estimation subproblem leads to

$$f(c) = \int_0^T \alpha_1 \|c(\cdot,t)\|_{L^1(\Omega)} + \delta_{c(\cdot,t)\geq 0}\, \mathrm{d}t, \tag{3.8}$$

$$g(Bc) = \int_0^T \|A(c,t) - u(t)\|_{L^1(\mathbb{R}^d)} + \alpha_2 \|\nabla c(\cdot,t)\|_{L^1(\Omega)^n} + \gamma \|(\mathcal{W}c)(t)\|_{L^1(\Omega)}\, \mathrm{d}t \tag{3.9}$$

with the linear operator $B = (\, A, \nabla, \mathcal{W}\,)^T$ and its adjoint $B^* = (\, A^*, \quad -\mathrm{div}, \quad \mathcal{W}^*\,)$. The corresponding Fenchel-conjugate is given by

$$g^*(y_1,y_2,y_3) = \int_0^T \delta_{\{\|y_1\|_\infty \leq 1\}} + \langle y_1, u\rangle + \delta_{\{\|y_2\|_{3,\infty} \leq \alpha_2\}} + \delta_{\{\|y_3\|_\infty \leq \gamma\}} \mathrm{d}t. \tag{3.10}$$

3.2.2. The MC constraint. The MC constraint

$$m_2(c,v) = \frac{\partial}{\partial t}c + \nabla \cdot (cv) = 0$$

is not linearized. We therefore use the constraint directly inside our penalty term, i.e.,

$$\begin{aligned} V\left(c(\cdot,t), \frac{\partial}{\partial t}c(\cdot,t), v(\cdot,t)\right) &= \left\| \frac{\partial}{\partial t}c(\cdot,t) + \nabla \cdot \Big(c(\cdot,t)\, v(\cdot,t)\Big)\right\|_{L^1(\Omega)} \\ &= \left\| \frac{\partial}{\partial t}c(\cdot,t) + c(\cdot,t)\, \nabla \cdot \Big(v(\cdot,t)\Big) + v(\cdot,t)\cdot \nabla c(\cdot,t)\right\|_{L^1(\Omega)}. \end{aligned}$$

Applying again the notation from the motion estimation subproblem leads to

$$\begin{aligned} f(c) &= \int_0^T \alpha_1 \|c(\cdot,t)\|_{L^1(\Omega)} + \delta_{c(\cdot,t)\geq 0}\, \mathrm{d}t, \\ g(Bc) &= \int_0^T \|A(c,t) - u(t)\|_{L^1(\mathbb{R}^d)} + \alpha_2 \|\nabla c(\cdot,t)\|_{L^1(\Omega)^n} \\ &\quad + \gamma \left\| \frac{\partial}{\partial t}c(\cdot,t) + c(\cdot,t)\, \nabla \cdot \Big(v(\cdot,t)\Big) + v(\cdot,t)\cdot \nabla c(\cdot,t)\right\|_{L^1(\Omega)} \mathrm{d}t, \end{aligned}$$

where

$$B = \begin{pmatrix} A \\ \nabla \\ \frac{\partial}{\partial t} + v_1 \frac{\partial}{\partial x} + v_2 \frac{\partial}{\partial y} + v_3 \frac{\partial}{\partial z} + (\nabla \cdot v)\, I \end{pmatrix},$$

and

$$B^* = \begin{pmatrix} A^*, & -\mathrm{div}, & -\left(\frac{\partial}{\partial t} + v_1 \frac{\partial}{\partial x} + v_2 \frac{\partial}{\partial y} + v_3 \frac{\partial}{\partial z} - (\nabla \cdot v)\, I \right) \end{pmatrix}.$$

The Fenchel-conjugate corresponding to g is the same as in (3.10).

3.2.3. Stochastic PDHG. To reduce the run time of the algorithm, we use the stochastic variant of PDHG introduced in 2018 by Chambolle et al. [12] to solve the problem for both motion constraints. We split the dual variable y_1 related to the data fitting term rowwise into M variables as $y_1 = (y_{1,1}, y_{1,2}, \ldots, y_{1,M})^T$. This is possible due to the separability of the MPI forward operator. Accordingly, the linear operators $B_{1,1}, \ldots, B_{1,M}$ are defined by rowwise splitting of the forward operator, i.e., $A = B_1 = (B_{1,1}, B_{1,2}, \ldots, B_{1,M})^T$. By default, we set $M = 3$, relating the data batches to the three receive channels of a standard MPI scanner. For more details on appropriate data splitting for MPI data, we refer to [75]. The resulting algorithm is stated in Algorithm 3.2. Note that $\sum_{j=1}^{3} B_j^* y_j$ in the primal update step can also be updated instead of fully computed in each iteration. In this way, the computational burden of the algorithm is reduced, as $B_i^* y_i$ has to be computed for the updated index i only, in the case $i = 1$ even for $B_{1,m}^* y_{1,m}$ only if the mth data batch is chosen.

Algorithm 3.2. Stochastic PDHG (SPDHG) algorithm for image reconstruction.

input initial value c^0, initial values y_i^0, $i \in \{1, \ldots, 3\}$, θ, step size parameters σ, τ, operators B_i, B^*
for $k = 0, 1, 2, \ldots, K-1$ **do**
  select $j \in \{1, 2, 3\}$ randomly
  if $j = 1$ **then**
    select $m \in \{1, \ldots, M\}$ randomly
    $\bar{y}_{1,m}^{k+1} = y_{1,m}^k + \sigma B_{1,m} \tilde{c}^k$
    $y_{1,m}^{k+1} = \mathrm{prox}_{\sigma g_1^*}(\bar{y}_{1,m}^{k+1})$ *% dual update*
  else
    $\bar{y_j}^{k+1} = y_j^k + \sigma B_j \tilde{c}^k$
    $y_j^{k+1} = \mathrm{prox}_{\sigma g_j^*}(\bar{y_j}^{k+1})$ *% dual update*
  end if
  $\bar{c}^{k+1} = c^k - \tau \sum_{j=1}^{4} B_j^* y_j^{k+1}$
  $c^{k+1} = \mathrm{prox}_{\tau f}(\bar{c}^{k+1})$ *% primal update*
  $\tilde{c}^{k+1} = c^{k+1} + \theta(c^{k+1} - c^k)$ *% extrapolation step*
end for

4. Numerical experiments. Our numerical experiments are divided into three sections. We first use simulated data in a 3D setting in order to be able to compare our reconstructions to a ground truth. We can thus quantify improvements by the joint approach in the image reconstruction task as well as in the motion estimation task compared to the standard MPI reconstruction scheme. The second section then uses 3D measured data of a phantom that fulfills the optical flow constraint. Measurement sequences are available for different motion speeds during measurements such that we can compare the impact of the joint approach for different velocities. In this section, we primarily focus on the image reconstruction task and we also show that obtained motion estimates are more reliable by the proposed approach compared to estimates from images reconstructed by the Kaczmarz method by comparing the trajectories derived from the obtained flow fields. In the third section, we reconstruct in vivo measurement data of a mouse heart. This application fulfills MC but not the optical flow constraint. In contrast to the previous section, we will focus on the resulting motion estimates.

4.1. Experiments on 3D simulated data. In order to have a ground truth available, we use simulated data for these first experiments. The scanner was modeled based on the preclinical MPI scanner (Bruker Biospin, Ettlingen) at the University Medical Center Hamburg-Eppendorf. The complete parameter setup can be found in Appendix C. The phantom is based on the rotation phantom used in [22] and in our experiments in subsection 4.2, but we add a movement along the z-direction yielding a ball moving along a spiral. We compare several reconstruction techniques. First, we use standard frame-by-frame static reconstructions obtained by the Kaczmarz method using standard Tikhonov L^2- regularization combined with a positivity constraint [15], which corresponds to one of the most popular MPI reconstruction techniques [70]. Moreover, we consider frame-by-frame reconstruction by the SPDHG algorithm using nonnegative fused lasso regularization (using 3D TV). In the proposed joint reconstruction setting, we test several different formulations. They all apply nonnegative fused lasso regularization on the tracer concentration and TV regularization on the motion estimates, but differ in terms of the data fitting norm and the motion model. We use both L^1- and L^2-data fitting (denoted as cases L^1-D and L^2-D) and optical flow (OF) and mass conservation (MC) constraints, resulting in a total amount of four different combinations. All joint approaches are tested on 10 frames simultaneously in the simulated data setting. The quality of the reconstructed image sequences is assessed by structural similarity index measure (SSIM).

We investigate three different noise levels on the data, namely, a no noise, a medium noise, and a high noise scenario. To obtain the noisy data for the different levels, the following procedure is used. In a first step, we add Gaussian noise with mean value 0 and standard deviations of $0\%, 50\%$, and 100%, respectively, of the maximum absolute value of the measured signal to the measurements (in the time domain) before transforming the measurements to the Fourier domain. We then use a data preprocessing similar to the one typically used for measured data. We transform matrix and data to the Fourier domain and, instead of applying an SNR threshold, we limit the maximum mixing order of the frequencies, which yields a similar outcome [69]. Moreover, the data are split into real and imaginary parts in order to perform the reconstruction on real numbers like in [67]. Due to this processing, the random noise on the data used for reconstruction then has a standard deviation of 0%, 2.7%, and 5.3%. Note that the noise-free setting is used for comparison only, but is not of practical

relevance. The high noise level reconstructions resemble what we expect of measured data, whereas the middle noise level might be achieved in an optimal experimental setup. Note that our simulated forward operator is not corrupted by noise, which defines a fundamental difference to the setting in practice, where the forward operator is typically measured.

Moreover, our simulation setting uses ideal magnetic fields, is based on the simple Langevin model for magnetization and neglects further noise sources such as background signals such that the reconstruction task is performed even in the case of noisy simulated data under ideal conditions compared to the reconstruction of measured data.

Best reconstruction parameters (obtaining highest SSIM values) are stated in Appendix D. The comparisons in Table 1 show that the Kaczmarz method yields solutions with the lowest SSIM and PSNR values as well as the highest MSE and L^1-error. For illustration, the exemplary averaged intensity projections of the reconstructed images in Figure 4 shows reconstructions for the medium noise level and further underline the quality of the results by the variational approaches, which are able to handle the noise much better than the Kaczmarz method.

Table 1

Comparison of the different algorithms in terms of SSIM, PSNR, MSE, and L^1-error of the reconstructed image sequences for three different noise levels and simulated data.

Noise level	No noise				Medium				High			
	SSIM	PSNR	MSE $\cdot 10^{-4}$	L1 $\cdot 10^{3}$	SSIM	PSNR	MSE $\cdot 10^{-4}$	L1 $\cdot 10^{3}$	SSIM	PSNR	MSE $\cdot 10^{-4}$	L1 $\cdot 10^{3}$
Kaczmarz	0.657	25.31	30	1.48	0.654	23.18	48	1.58	0.651	22.14	61	1.58
SPDHG	0.955	27.12	19	0.47	0.953	26.93	20	0.48	0.946	26.52	23	0.50
OF-L^1-D	0.962	27.79	17	0.43	0.960	27.56	18	0.44	0.955	27.62	18	0.44
OF-L^2-D	0.947	26.61	22	0.52	0.946	26.51	22	0.53	0.950	27.02	20	0.46
MC-L^1-D	0.963	27.85	17	0.42	0.961	27.60	18	0.43	0.956	27.57	18	0.43
MC-L^2-D	0.947	26.60	22	0.52	0.946	26.49	23	0.53	0.950	27.00	20	0.47

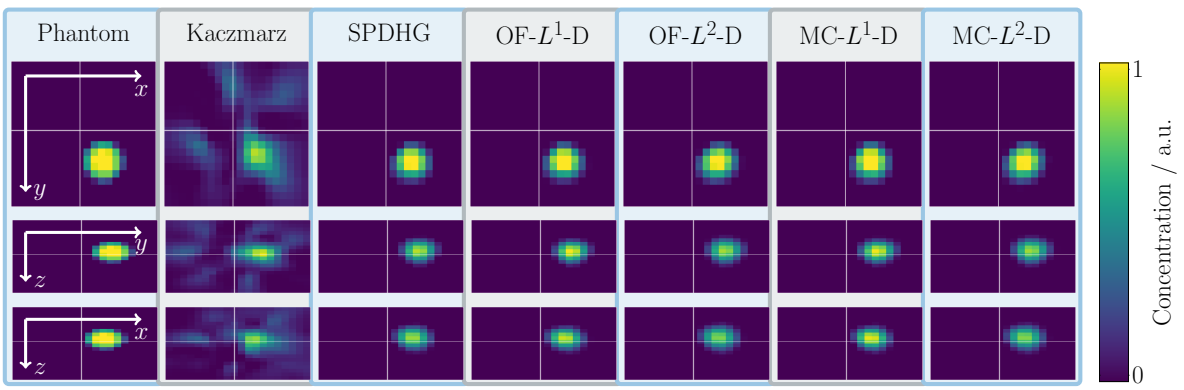


Figure 4. *Reconstructed images of the simulated data for the second time step. The upper row depicts averaged intensity projections onto the x-y-plane, the middle row projects onto the y-z-plane, and the bottom row onto the x-z-plane. Each column corresponds to one reconstruction algorithm. All images are plotted with the same colorbar viridis and windowed based on the full dynamic range of the phantom.*

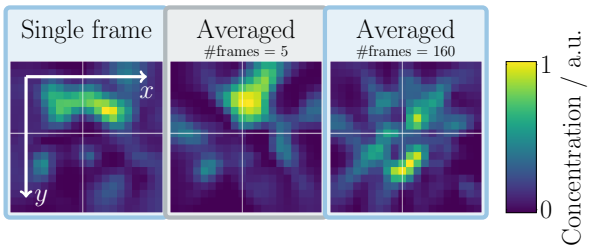


Figure 5. *Averaged intensity projections of the reconstructed images of the rotation phantom for medium noise level. Reconstructions are obtained via the Kaczmarz method using both single frame measurements as well as averaged measurements over multiple frames. The best regularization parameter* λ_{single} *weighting the Tikhonov regularization term was determined for the single frame reconstruction and then adjusted to the amount of averaging (using* $\lambda_{\text{single}}/\sqrt{\#\text{frames}}$*) in order to stay proportional to the expected noise level. The amount of frame averaging is increased from the left to the right column.*

Using a more suitable regularization method and an SPDHG algorithm considerably improves the reconstructed image sequences even when using a frame-by-frame reconstruction approach. Best results are obtained for the proposed joint reconstruction technique for an L^1-D term.

In a standard static setting, the poor image quality resulting from the Kaczmarz method under noise is improved by averaging the measurements over multiple frames. This procedure yields a higher SNR in the measurements and thus reconstructed images of higher quality with fewer noise artifacts. However, if we apply frame averaging in the dynamic case, this results in severe motion artifacts as can be seen in Figure 5. For averaging of very few frames (5 in this example) the SNR is barely improved and motion artifacts occur already, but for averaging of 160 frames the motion artifacts make it impossible to locate the object as observed in [22] as well.

4.2. Experiments on a 3D optical flow-like data set. The data used in this section was first published in [22]. The measurements were taken by a preclinical MPI scanner (Bruker Biospin, Ettlingen). The scanner uses sinusoidal excitation frequencies and thus creates 3D Lissajous-type trajectories. The drive-field frequencies are given by $f_x = 2.5\,\text{MHz}/102$, $f_y = 2.5\,\text{MHz}/96$, and $f_z = 2.5\,\text{MHz}/99$ with an amplitude of $14\,\text{mT}/\mu_0$, resulting in a repetition time of 21.54 ms. Frames are not averaged such that the temporal resolution is defined by the repetition time. The selection field gradient strength is $0.75\,\text{T/m}/\mu_0$ in the x- and y-directions and $1.5\,\text{T/m}/\mu_0$ in the z-direction. The delta sample used for system matrix acquisition was of size $2 \times 2 \times 1\ \text{mm}^3$ and covers $25 \times 25 \times 25$ positions.

The phantom consists of a rotating glass capillary and is constructed in the following way: two round disks of the same diameter as the scanning bore are connected by three rods. An additional rod is placed in the center and a glass capillary filled with tracer material ($20\,\mu$L diluted ferucarbotran with factor 1/10) is attached to it. The central rod can be attached to a screwdriver and rotated during measurements. The central rod has a diameter of 10 mm and the attached glass capillary has an inner diameter of 1.3 mm and an outer diameter of 1.7 mm. Considering the amount of tracer material, this yields a fill height of approximately 15 mm in the x-direction (corresponding to 7.5 voxel). The construction results in a circular path of the tracer material with a diameter of approximately 13 mm. Image sequences for three

different rotation speeds were obtained; the respective motion frequencies are approximately 1 Hz, 3 Hz, or 7 Hz. At a repetition time of 21.54 ms, this implies a movement by 0.9 mm, 3.1 mm, and 5.9 mm per frame along the circular path. The total displacement in the y- and z-directions is 26 mm per cycle rotation, leading to mean displacements of 0.6 mm, 2.0 mm, and 3.8 mm per frame. A more detailed description of the phantom as well as illustrations can be found in [22].

We use a standard procedure for data preprocessing, applying a frequency selection based on an SNR threshold of 5, and cutting off frequencies lower than 80 kHz (see, e.g., [43] for a formal description of the preprocessing). Moreover, we use a weighting approach by row normalization as introduced in [47].

We compare the visual quality of our reconstructed image sequences to a sequence reconstructed by a standard framewisely applied Kaczmarz algorithm using an L^2-Tikhonov regularization as well as to framewisely applied SPDHG algorithm for fused lasso regularization. Full image sequences are appended to the manuscript. Since the phantom fulfills the OF constraint and thus both possible constraints, we will use it to compare both MC and OF constraints.

Figure 6 shows that our method yields significantly fewer noisy reconstructed images than the Kaczmarz method. Moreover, the dot shaped phantom is less blurry, although still at the

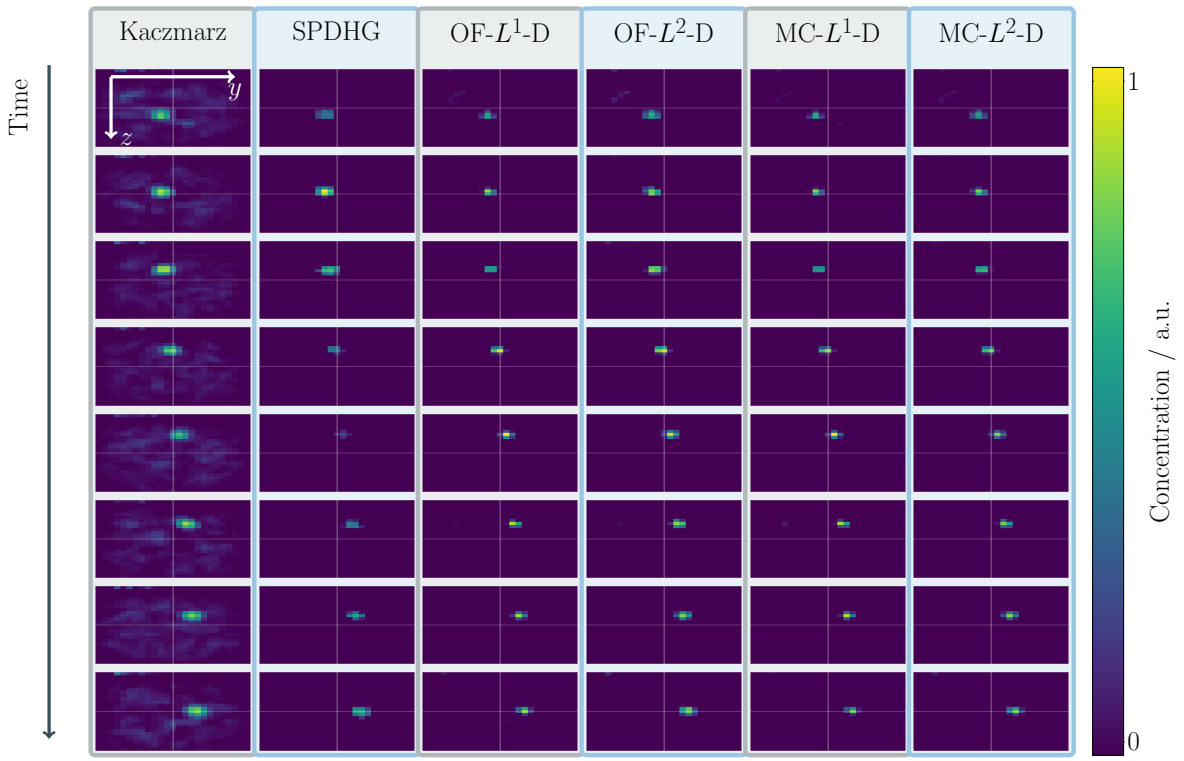


Figure 6. *Reconstructed image sequences for the* 3 Hz *rotation dataset. Each column corresponds to one algorithm, each row represents one frame in time. We display averaged intensity projections onto the y-z-plane in this figure.*

Table 2

The parameters used for the reconstruction of the 3 Hz *rotation phantom data. The early stopping index indicates at which iteration the algorithm was terminated. The parameter λ denotes the weight on the Tikhonov L^2-penalty term used by the Kaczmarz method. As defined in (3.6), α_1 denotes the weight on the L^1-penalty on the concentration, α_2 defines the weight on the TV-penalty term on the concentration. β denotes the weight on the TV-penalty on the flow and γ weights the motion model penalty term (see (3.1)).*

Algorithm	Motion model	Early stopping	λ	α_1	α_2	β	γ
Kaczmarz	None	$k = 10$	5.62				
SPDHG	None	$k = 10^3$		$5.0 \cdot 10^{-1}$	$3.0 \cdot 10^{-1}$		
Joint, L^1-D	OF	$k = 10^2$		$6.0 \cdot 10^{-1}$	$1.0 \cdot 10^{-1}$	10^{-1}	100
Joint, L^2-D	OF	$k = 10^2$		$2.5 \cdot 10^{-1}$	$1.0 \cdot 10^{+2}$	10^{-1}	100
Joint, L^1-D	MC	$k = 10^2$		$7.0 \cdot 10^{-1}$	$5.0 \cdot 10^{-1}$	10^{-1}	100
Joint, L^2-D	MC	$k = 10^2$		$2.5 \cdot 10^{-1}$	$2.5 \cdot 10^{-1}$	10^{-1}	100

correct position for each frame. The qualitative differences between the L^1-D and L^2-D terms are negligible which might be due to the chosen weighting in the data preprocessing. It can be observed that the reconstructed rod by an L^1-D term are slightly better resolved and the shape is closer to the actual phantom size, which covers less than one voxel in the y-z-plane. Framewise, the SPDHG algorithm yields fewer noise artifacts than the Kaczmarz method, but has severe problems in achieving comparable mass in each frame. All reconstruction parameters are given in Table 2. They were chosen based on visual inspection out of a wide range of tested parameters; for details on the parameter tests see Appendix D. The joint algorithm proposed here is used in combination with early stopping in order to limit the computation time, however, the reconstructed image sequences are close to convergence at that point. More precisely, the early stopping iteration is given for the image reconstruction subtask in the inner loop of the alternating algorithm. The alternations are performed 3 times. Figure 7 displays reconstructed image sequences in the other two planes. The proposed method yields reconstructed images, in which the phantom stays constant in size and has a reasonable shape (length of 7 voxels times width of 2 voxels). For the Kaczmarz method reconstructions the phantom seems to vary in size. In time frames 4 to 6, it appears significantly shorter than in other frames. Moreover, it is often difficult to distinguish between phantom and background noise.

In order to underline the improvements by the OF or mass preservation constraint, we list the coefficient of variation, i.e. standard deviation divided by the mean value, of the reconstructed mass in each frame over the first 30 frames for each reconstruction approach in Table 3. All joint reconstruction approaches outperform the Kaczmarz method based on this measure. The value of the framewise SPDHG is in completely different ranges compared to the other algorithms, which confirms the visual impression.

In order to investigate the impact of our approach to the motion estimation, we compare the motion fields obtained during the joint reconstruction task to motion fields obtained from the sequences reconstructed by the Kaczmarz method. In Figure 8(a) we compare motion fields for the 3 Hz dataset by comparing the estimated particle trajectories obtained from the different motion estimates. The best possible particle trajectories (left column, depicted in white) are obtained step by step from the reconstructed images and are almost identical for both methods. For guidance, we fitted a circular path with the correct diameter of approximately 13 mm as approximate ground truth to the plot (depicted in green). However,

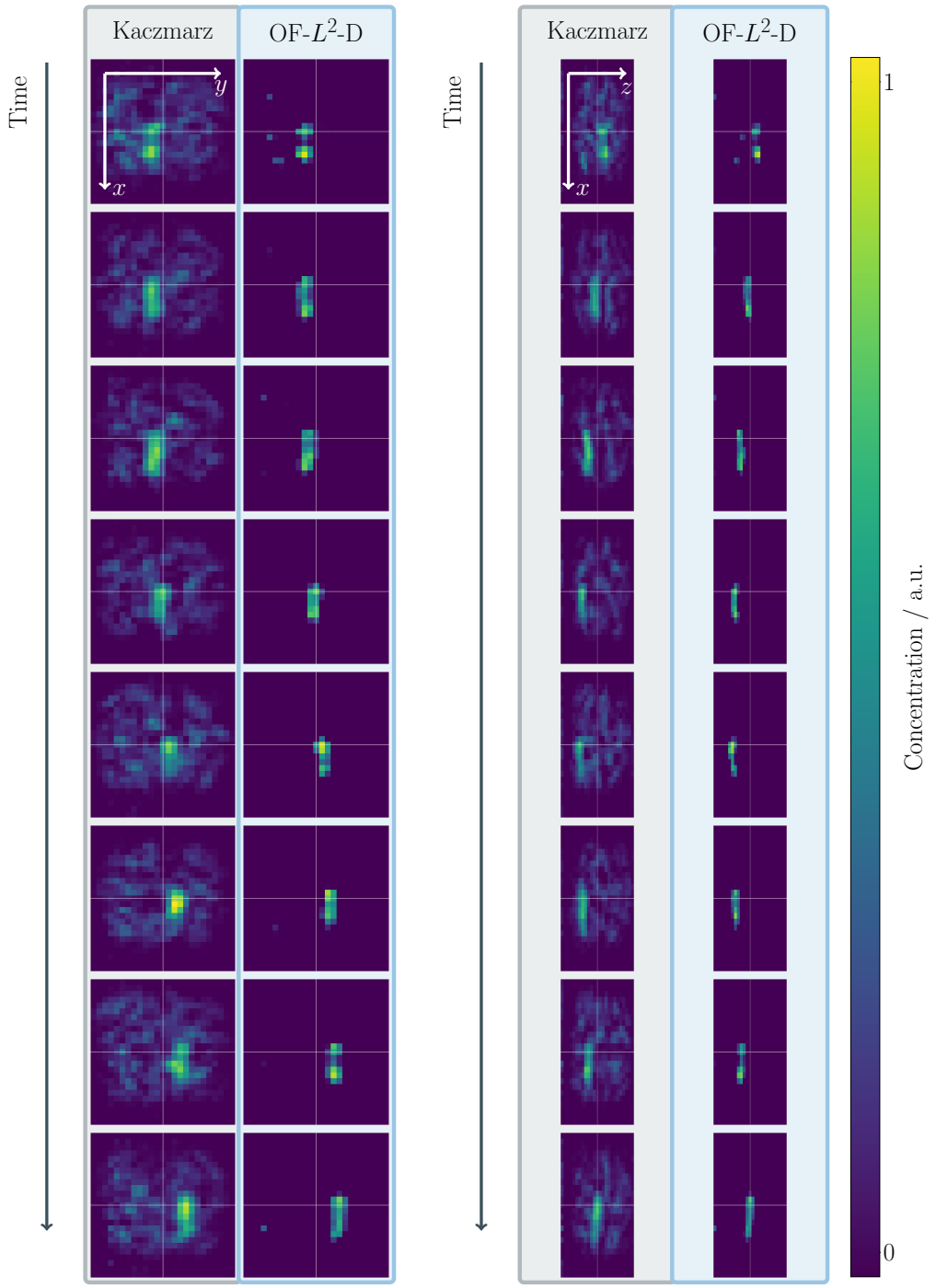


Figure 7. *Reconstructed image sequence of the* 3 Hz *rotation phantom for the Kaczmarz method (applied frame-by-frame) and the joint OF-L^2-D algorithm. This figure displays averaged intensity projections onto the y-x- and z-x-planes for the first eight timeframes (each row corresponds to one step in time). The phantom has a size of* 15 mm *in the x-direction and* 1.3 mm *in the z- and y-direction. This translates to approximately* 7–8 *voxels in the x-direction, one voxel in the y-direction, and* 1–2 *voxels in the z-direction, depending on the specific position of the phantom.*

Table 3

The coefficient of variation of the total reconstructed mass in each frame over the first 30 *frames of the sequence for the* 3 Hz *dataset.*

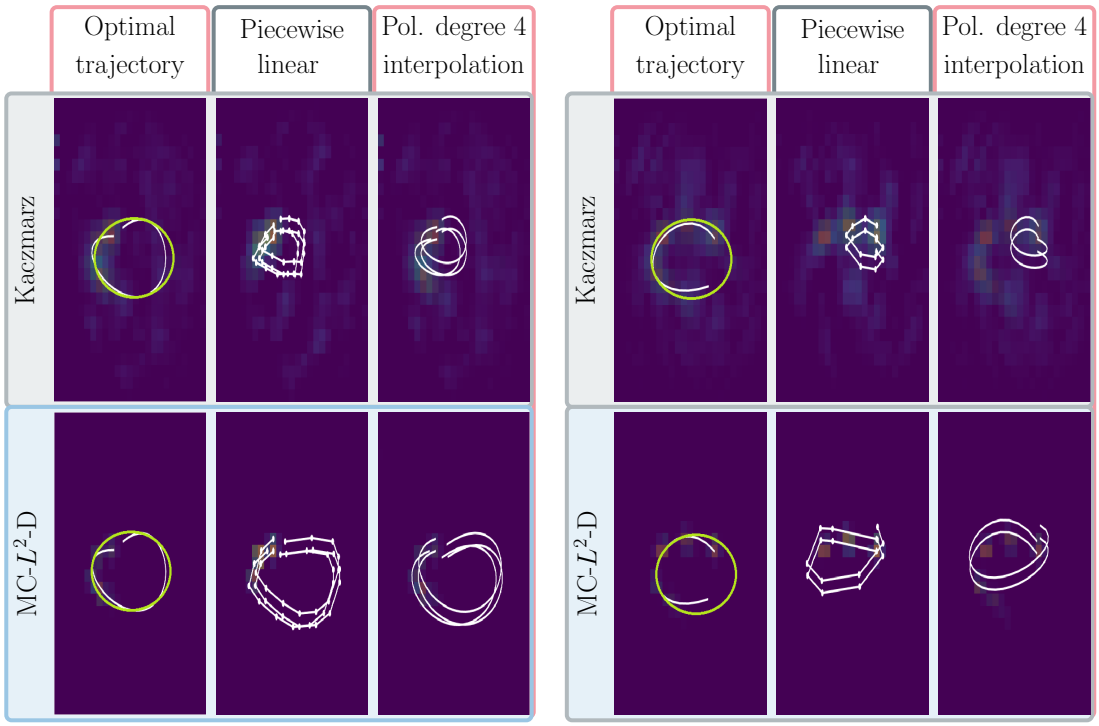
Algorithm	Kaczmarz	SPDHG	OF-L^1-D	OF-L^2-D	MC-L^1-D	MC-L^2-D
Coefficient of variation	0.149	0.556	0.116	0.141	0.110	0.129

the motion field obtained from the Kaczmarz reconstructions highly underestimates the magnitude of motion, probably due to the severe noise artifacts in the image sequence. The piecewise linear as well as the smoothed approximation of the trajectory computed with the help of the motion field is too narrow in the case of the Kaczmarz reconstructed image sequence. The motion estimates obtained by the joint MC-L^2-D algorithm, however, are close to the optimal trajectory. This visual impression is underlined by the computed radii of the best-fit circles through the data points indicated in the middle column. The motion field based on Kaczmarz-reconstructed images yields a best-fit circle with radius 3.7 mm, whereas the proposed approach obtains a circle with radius 6.7 mm. The phantoms setup shows that the ground truth radius is in-between 6.3 mm and 6.7 mm. Moreover, the mean distance traveled on the circular path per frame (averaged over the first 34 frames) is 1.9 mm and 3.2 mm for the motion based on the Kaczmarz method and the joint approach, respectively. The ground truth in this case is a mean distance of 3.1 mm per frame. Noting that the spatial resolution is 2 mm and 1 mm in the two directions, respectively, the radius as well as the mean distance traveled per frame are very close to the ground truth.

Figure 8(b) shows similar results for the 7 Hz dataset. The more challenging image reconstruction task leads to severe problems in the motion estimation task based on the images reconstructed by the Kaczmarz method. Our method again yields trajectory estimates that are close to the best possible guess based on the image sequences.

Reconstruction of the rotation phantom datasets for different speeds shows that the image reconstruction task benefits from the joint approach. The more appropriate priors (in comparison to the Tikhonov regularization used by the Kaczmarz method) minimize noise artifacts in the reconstruction. Moreover, the reconstructed mass over the time frames is more constant due to the additional MC prior. Moreover, the motion estimates are also improved by the joint approach. The enhanced image sequences lead to more accurate motion estimates. In summary, the experiments underline the increased quality of the results by the joint approach. A two-step scheme using the Kaczmarz method for the image reconstruction has significantly fewer parameters to tune, but the reconstructed images as well as motion estimates are not of comparable quality (see Figures 6 to 8). A two-step scheme based on SPDHG would yield images and motion estimates of improved quality very similar to the ones by the joint approach, but it has a similar number of parameters to tune. Thus, the slightly improved quality of the results by the joint approach emphasizes its appropriateness.

Video sequences of reconstructed images for the 3 Hz and 7 Hz datasets reconstructed frame-by-frame by the Kaczmarz method as well as a joint approach are available as additional material (M158040_01.gif [local/web 10.9MB], M158040_02.gif [local/web 15.8MB], M158040_03.gif [local/web 17.5MB], M158040_04.gif [local/web 2.31MB], M158040_05.gif [local/web 2.39MB]).



(a) Particle trajectories for frames 3 to 15 of the 3 Hz dataset.

(b) Particle trajectories for frames 3 to 8 of the 7 Hz dataset.

Figure 8. *The left column displays the particles trajectory as observed in the reconstruction and smoothed by a 4th order polynomial approximation (in white) as well as a circle with correct diameter fitted to the curve (in green). The middle column shows the piecewise linear trajectories, which are obtained by adding up the computed motion at each time step. The right column shows a 4th order polynomial approximation to the curve in the middle.*

4.3. Experiments on a 3D in vivo blood flow dataset. The in vivo mouse data were first published in [24] in the context of presenting a highly sensitive gradiometric receive coil. The system matrix has been calibrated with a small capillary of size 0.7 mm in order to capture the fine structures of the cardiovascular system. This yields a system matrix of size $46 \times 36 \times 19$ voxels capturing a volume of size $32.2 \times 25.2 \times 13.3$ mm^3. The measurement sequence is collected by the above-mentioned gradiometric receive coil for an injected bolus of volume 10 μL. No time averaging was applied resulting in a repetition time of 21.54 ms as for the previous phantom experiment. The expected flow during the inflow of the bolus is about 5 cm s^{-1} [38], which would translate into a movement of about 1 mm per frame. One thus already expects motion-related inconsistencies, for which the OF constraint is beneficial. We use the same data prepocessing as for the rotation phantom data with an additional correction for background signal, where dedicated empty measurements are subtracted from measurements and system matrix [71]. Anatomic background information on the inner tissue structure is available from MRI scans and is combined with the MPI reconstructions as grayscale background images in order to interpret the resulting images. Exemplary images of the particle inflow into the cardiovascular system of the mouse heart via the vena cava inferior are displayed in Figure 9.

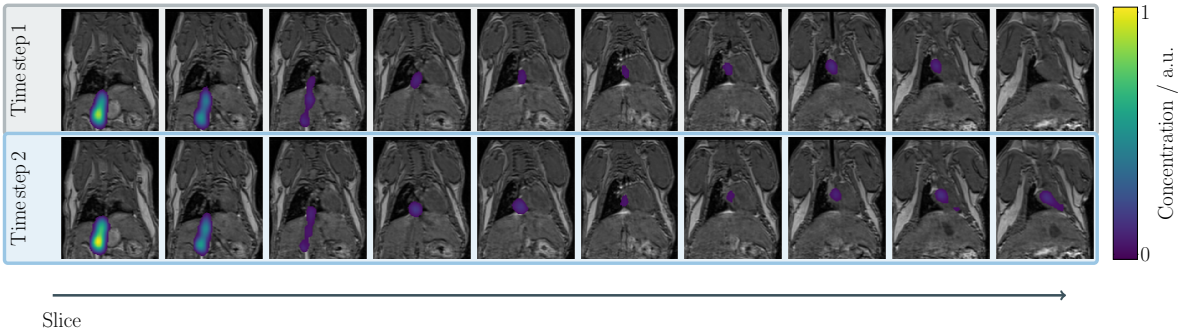


Figure 9. *Coronal slices of reconstructed concentration obtained from the in vivo mouse dataset using the MC-L^2-D method with an L^2 penalty term for v. Each column corresponds to one slice of tissue; MRI data are integrated for background information. Two different time steps are depicted in the different rows. Concentrations larger than* 0.4% *of the maximum concentration are illustrated.*

A full video sequence showing the reconstructed images for the MC-L^1-D algorithm with the parameters $\alpha_1 = 0.05$, $\alpha_2 = 100$, $\beta = 0.1$, and $\gamma = 100$ is appended to this manuscript (M158040_01.gif [local/web 10.9MB], M158040_02.gif [local/web 15.8MB], M158040_03.gif [local/web 17.5MB], M158040_04.gif [local/web 2.31MB], M158040_05.gif [local/web 2.39MB]). In the absence of a ground truth reference we now focus on the reconstructed motion fields. This is particularly useful, as MPI has potential applications in cardiovascular imaging [38]. As the blood flow through the cardiovascular system is not as homogeneous in space as for the previous phantoms, we do not use regularization on the gradient of the flow. Instead we use standard L^2 Tikhonov regularization on the flow fields. The motion estimation subproblem thus reads

$$\min_v \int_0^T \beta \left\| v\left(\cdot,t\right)\right\|^2_{L^2(\Omega)^n} + \gamma \left\| \frac{\partial}{\partial t} c\left(\cdot,t\right) + \nabla \cdot \left(c\left(\cdot,t\right) \cdot v\left(\cdot,t\right)\right)\right\|_{L^1(\Omega)} \mathrm{d}t.$$

The dual update step in Algorithm 3.1 simplifies as the Tikhonov term belongs to the primal problem. The primal update step consists of a different but easy to compute proximal operator. Parts of the reconstructed image sequence as well as the estimated motion field in 2D slices for a later time point after the particle inflow are depicted in Figure 10. The flow is illustrated in two different ways. First, we have a quiver plot where the direction and strength of the flow at a certain position is indicated by an arrow and its direction and length. The strength of the flow is clearly visible in this plot. However, direction is difficult to observe, which is why we add a color wheel plot, where each direction is linked to a certain color. Additionally, the flow in the third spatial direction, i.e., orthogonal to the slices depicted in Figure 10, is illustrated in Figure 11. In general, we observe flow estimates, which are plausible with respect to the underlying structure provided by the MRI image and illustrate the blood flow through the cardiovascular system. Such flow information is already used in other imaging modalities like 4-dimensional-flow magnetic resonance imaging [68] and thus would be very interesting for future clinical real-world applications of MPI.

5. Conclusion and outlook. In the present work we considered a joint image reconstruction and motion estimation approach in the dynamic inverse problem setting of MPI where a

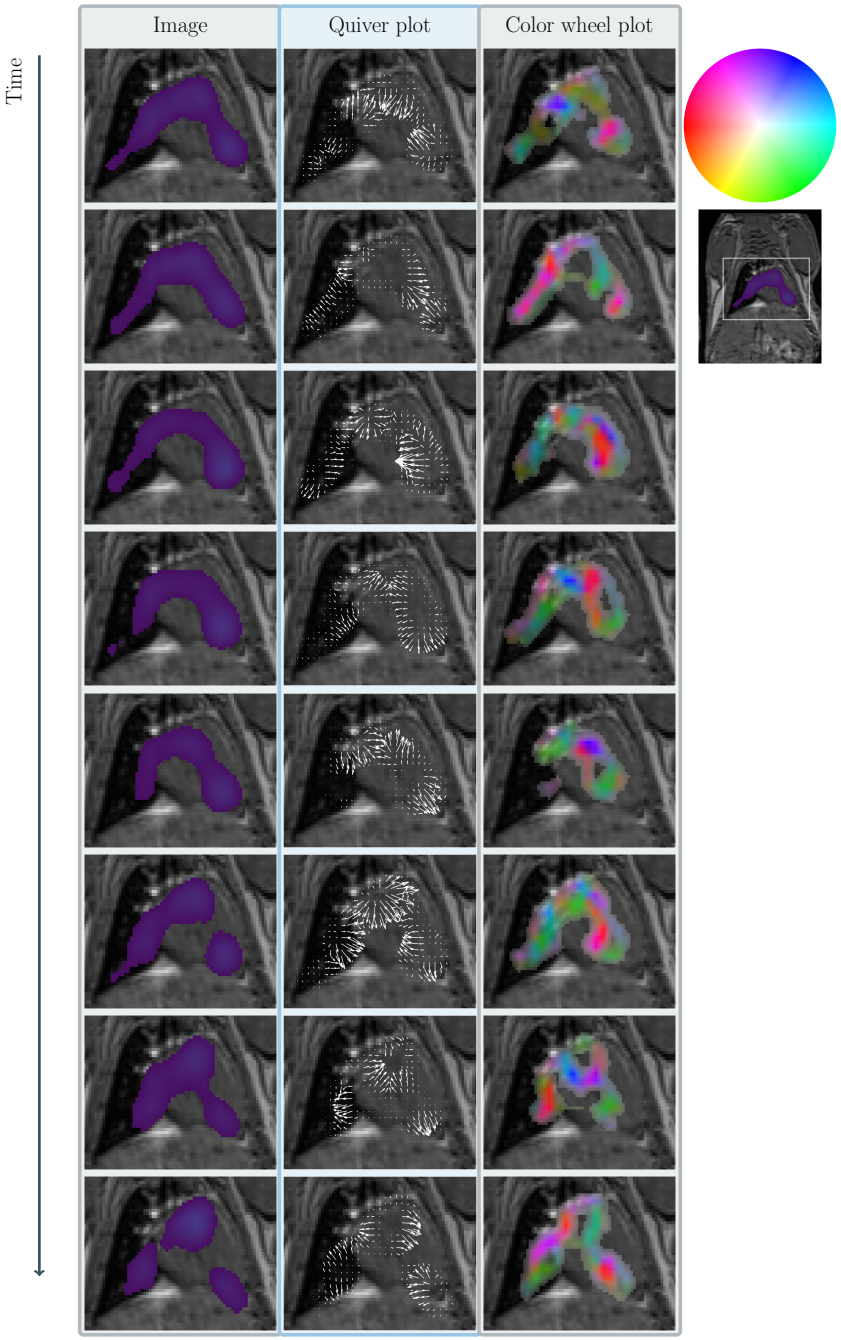


Figure 10. *Reconstructed image sequences for the in vivo flow dataset using the MC-L^2-D method with an L^2-penalty term for v. For illustrative purposes we depict a slice corresponding to height* $[9.1, 9.8]$mm *(index* 14*) of the* 3*D concentration and the* 3*D+time flow over eight subsequent frames. The left column depicts the reconstructed concentration in the snippet indicated in the most-right column (same colormap as before). Concentrations larger than* 0.4% *of the maximum concentration are illustrated. The second left plot indicates the flow fields by a quiver plot, whereas the second right column depicts the flow field by a color wheel plot using the color wheel in the right in order to illustrate the direction of the flow.*

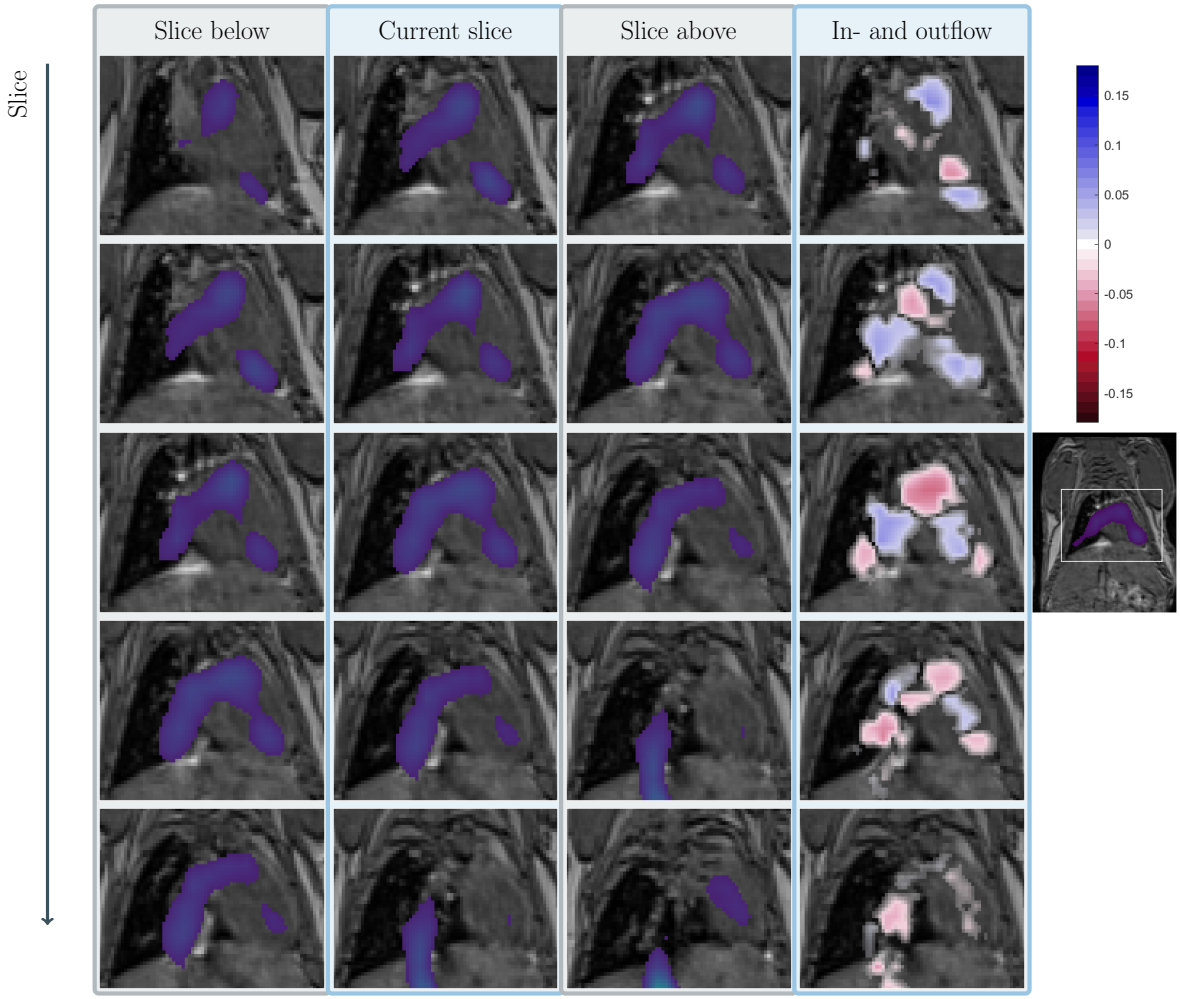


Figure 11. *Reconstructed image at one time step for the in vivo flow dataset using the MC-L^2-D method with an L^2-penalty term for v. For illustrative purposes we depict three consecutive slices of the 3D concentration in the first three columns (same colorbar as before) and the 3D+time flow leaving the current slice in the most-right column. Concentrations larger than 0.6% of the maximum concentration are illustrated. The colorbar in the right identifies the direction of the in- and out-flow into the current slice, where a positive flow means into the slice above and a negative flow means going into the direction of the slice below.*

time-dependent sequence of 3D concentration images is reconstructed. The joint image and flow parameter estimation is based on [11] taking into account an OF or a mass preservation constraint for 2D image processing tasks, which is in the present work extended to the 3D+time case in MPI investigating the existence of a solution to the minimization problem in this setting. This particularly includes a theoretical discussion of the required properties in the MPI equilibrium model. In addition an efficient solution to the problem is developed using alternating minimization with primal-dual algorithms in both subproblems. The image reconstruction subproblem is solved by stochastic PDHG exploiting the structure of the forward operator in MPI. The algorithm solving the motion estimation subproblem combines the PDHG algorithm with a multiscale scheme and warping. The theoretical part is

accompanied by an extensive numerical investigation in the context of MPI based on simulated data showing the advantages of the proposed joint approach compared to standard MPI reconstruction algorithms. In addition the proposed method is evaluated on measured 3D+time data including rotation phantoms as well as in vivo data from a mouse. Here, not only the improved reconstruction quality in the joint approach is discussed but also the benefits of having access to the jointly estimated motion parameters are illustrated.

In summary, the present manuscript provides a comprehensive investigation of the dynamic reconstruction problem in MPI particularly taking into account motion priors and it thus paves the way for various directions of future work. On the theoretical as well as numerical side, future works include, for example, an extension allowing for the consideration of an in- and outflow into the field-of-view being included in the mass preservation or OF constraint or inclusion of prior knowledge about, e.g., edges obtained from MRI images. On the implementation side, we note that we did not focus on efficiency yet. The runtime is, with 41 s per frame for one alternating iteration compared to 5 s per frame for the Kaczmarz method, still to too large to be usable in real-time applications. Thus, an important future work is to reduce the reconstruction time using, e.g., parallelization and/or matrix compression techniques [48].

In the field of application of MPI, future works include further qualitative as well as quantitative evaluations with respect to image reconstruction quality as well as accuracy in flow estimation. This requires a careful experimental design to obtain access to ground truth information not only of the phantom but also of the flow field. In the present work, the flow is solely included in the motion model with respect to the concentration but recent studies [57] indicate that there also might be an explicit influence on the obtained measurement. In this case the obtained measurement have an immediate dependence on the flow field and it thus needs to be taken into account explicitly as an argument of the forward operator A, i.e., consider $A(c, v, t)$ in the general problem setting. Suitable models for the operator A in the context of MPI and the general theoretical investigation of this approach remain future work.

Appendix A. Proof of Theorem 2.1.

Proof. We follow the proof of [11, 17], but we have three main differences in our assumptions. First, our image sequence is in 3+t dimensions; second, our regularizers are of a more general form; and third, our linear operator is time dependent. We follow the technique of the direct method in the calculus of variations. The line of argumentation is as follows:

1. The functional $J : L^{\hat{p}}(0,T;BV(\Omega)) \times L^q(0,T;BV(\Omega)^n) \to \mathbb{R}$ is bounded from below such that $\inf J(c,v) > -\infty$. This implies the existence of a minimizing sequence of admissible$(c_m, v_m) \in L^{\hat{p}}(0,T;BV(\Omega)) \times L^q(0,T;BV(\Omega)^n), m \in \mathbb{N}$, (here, admissible means $c_m \in L^p(0,T;BV(\Omega))$, $\|v_m\|_{L^\infty(0,T;L^\infty(\Omega)^n)} \leq k_\infty$, $\|\nabla \cdot v_m\|_\theta \leq k_\theta$, $m_i(c_m, v_m) = 0$ in $\mathcal{D}'(\Omega \times [0,T])$), such that $\lim_{m\to\infty} J(c_m, v_m) = \inf J(c,v)$.
2. The sublevel sets of J are weak-* compact. As J is continuous, the sublevel sets are closed. It remains to show for this step in the argumentation that the sublevel sets are bounded in the norm. This then implies the existence of a subsequence $(c_{m_k}, v_{m_k}) \overset{*}{\rightharpoonup} (c^*, v^*)$, which then identifies the candidate (c^*, v^*) for the minimizer.
3. The functional J is lower semicontinuous with respect to the weak-* topology (as a sum of such functionals), which yields

$$\inf J(c,v) \leq J(c^*, v^*) \leq \liminf_{m\to\infty} J(c_{m_k}, v_{m_k}) = \inf J(c,v).$$

4. In order to obtain the weak-* closedness of the admissible set (in the weak-* topology), we prove convergence of the constraint in a distributional sense, i.e., we verify that (c^*, v^*) is admissible. We consider both the OF and MC constraints and complete the proofs in 3D.

Weak- compactness of sublevel sets:* Denote the sublevel set to $\nu \in \mathbb{R}$ by

$$S_\nu := \left\{ (c,v) \in L^{\hat{p}}\left(0,T;BV\left(\Omega\right)\right) \times L^q\left(0,T;BV\left(\Omega\right)^n\right) : J\left(c,v\right) \leq \nu, \|v\|_{L^\infty(0,T;L^\infty(\Omega)^n)} \leq k_\infty \right\}.$$

We will use that $S_\nu \subset L^{\hat{p}}\left(0,T;BV\left(\Omega\right)\right) \times L^q\left(0,T;BV\left(\Omega\right)^n\right) \cong X^*$ for X^* being the dual space of a Banach space X. Therefore S_ν is weak-* compact if and only if S_ν is closed in weak-* topology (this holds as J is continuous) and the sublevel set is bounded in the norm.

It remains to show that $\|c\|_{L^{\hat{p}}(0,T;BV(\Omega))} = (\int_0^T (\|c\|_{L^1(\Omega)} + |c|_{\mathrm{BV}(\Omega)})^{\hat{p}} \mathrm{d}t)^{1/\hat{p}} \leq k$ and $\|v\|_{L^q(0,T;BV(\Omega)^n)} \leq k$ for a $k < \infty$.

Let $(c,v) \in S_\nu$ (i.e, it holds additionally $\|v\|_{L^\infty(0,T;L^\infty(\Omega)^n)} \leq k_\infty$).

We start with the bound on the norm of c: from $(c,v) \in S_\nu$ it follows that

$$\int_0^T \frac{1}{2} \left\| \left(A_t c\right)(t) - u\left(t\right) \right\|_Y^2 \mathrm{d}t = \frac{1}{2} \left\| A_t c - u \right\|_Y^2 \leq \nu$$
$$\Rightarrow \left(A_t c - u\right) \in Y \quad \text{a.e. in } \left[0,T\right].$$

Define $k_A(t) := \|(A_t c)(t) - u(t)\|_Y$, then

$$\int_0^T k_A^{\hat{p}} \mathrm{d}t = \int_0^T \|A_t c - u\|_Y^{\hat{p}} \, \mathrm{d}t$$

is bounded for $1 \leq \hat{p} \leq 2$ due to the continuous embedding $L^2(0,T;Y) \hookrightarrow L^{\hat{p}}(0,T;Y)$.

To bound $\|c\|_{L^{\hat{p}}(0,T;L^1(\Omega))}$, we start by an upper bound for arbitrary but fixed $t \in [0,T]$. Defining $\bar{c} = \frac{1}{|\Omega|} \int_\Omega c(x,t) \mathrm{d}x$, $c_0 = c(\cdot,t) - \bar{c}$ we note that

$$\int_\Omega c_0 \, \mathrm{d}x = 0.$$

Moreover,

$$|c_0|_{\mathrm{BV}(\Omega)} = |c\left(\cdot,t\right)|_{\mathrm{BV}(\Omega)} \leq k \int_0^T |c\left(\cdot,\tau\right)|_{\mathrm{BV}(\Omega)}^p \, \mathrm{d}\tau \leq \frac{k\nu}{\alpha}$$

for a constant $k \in \mathbb{R}$. With Poincaré–-Wirtinger, it follows that there exists $k_1 > 0$ such that

$$\|c_0\|_{L^l(\Omega)} \leq k_1 \left|c\right|_{\mathrm{BV}(\Omega)} \leq k_1 \frac{k\nu}{\alpha} =: \tilde{k}_1 \tag{A.1}$$

for $l \leq \frac{n}{n-1}$. We derive the following estimate:

$$\begin{aligned}
&\|A_t \bar{c}\|_Y^2 - 2 \|A_t \bar{c}\|_Y \left(\|A_t\| \, \|c_0\|_{L^l(\Omega)} + \|u\|_Y \right) \\
&\quad \leq \|A_t \bar{c}\|_Y^2 - 2 \|A_t c_0 - u\|_Y \, \|A_t \bar{c}\|_Y \\
&\quad \leq \|A_t c_0 - u\|_Y^2 + \|A_t \bar{c}\|_Y^2 - 2 \|A_t c_0 - u\|_{L^2(Y)} \, \|A_t \bar{c}\|_Y \\
&\quad = \left(\|A_t c_0 - u\|_Y - \|A_t \bar{c}\|_Y \right)^2 \\
&\quad \leq \|A_t c_0 - u + A_t \bar{c}\|_Y^2 \\
&\quad = \|A_t \left(c_0 + \bar{c}\right) - u\|_Y^2 = k_A^2(t).
\end{aligned}$$

Using the estimate of (A.1) and $u \in L^2(0,T;Y)$ yields

$$0 \leq \|A_t\| \, \|c_0\|_{L^1(\Omega)} + \|u\|_Y \leq \|A_t\| \, \tilde{k}_1 + \|u\|_Y := k_2(t).$$

Note that, as $\|A_t\|$ is bounded for all t, we also have a bound on the time integral over k_2. We can now use the above estimate to obtain the following bound on $\|A_t\bar{c}\|_Y$:

$$\begin{aligned}
&\|A_t\bar{c}\|_Y^2 - 2\|A_t\bar{c}\|_Y \left(\|A_t\| \, \|c_0\|_{L^1(\Omega)} + \|u\|_Y\right) \\
&\quad + \left(\|A_t\| \, \|c_0\|_{L^1(\Omega)} + \|u\|_Y\right)^2 \leq k_A^2(t) + k_2^2(t) \\
&\Leftrightarrow \left(\|A_t\bar{c}\|_Y - \|A_t\| \, \|c_0\|_{L^1(\Omega)} + \|u\|_Y\right)^2 \leq k_A^2(t) + k_2^2(t) \\
&\Rightarrow \|A_t\bar{c}\|_Y \leq \left(k_A^2(t) + k_2^2(t)\right)^{1/2} + \|A_t\| \, \|c_0\|_{L^1(\Omega)} + \|u\|_Y \\
&\leq \left(k_A^2(t) + k_2^2(t)\right)^{1/2} + k_2(t) := k_3(t) \, .
\end{aligned}$$

Finally, we deduce a bound on $\|c(\cdot,t)\|_{L^1(\Omega)}$ by

$$\begin{aligned}
&0 \leq \|c(\cdot,t)\|_{L^1(\Omega)} \leq k_4 \|c(\cdot,t)\|_{L^1(\Omega)} = k_4 \|c_0 + \bar{c}\|_{L^1(\Omega)} \\
&\leq k_4 \left(\|c_0\|_{L^1(\Omega)} + |\bar{c}| \, \|\chi_\Omega\|_{L^1(\Omega)}\right) \leq k_4 \left(\tilde{k}_1 + \frac{k_3(t) \, \|\chi_\Omega\|_{L^1(\Omega)}}{\|A_t\chi_\Omega\|_Y}\right),
\end{aligned}$$

where we used that by assuming $\|A_t\chi_\Omega\|_Y \neq 0$ it holds that

$$|\bar{c}| \, \|A_t\chi_\Omega\|_Y = \|A_t\bar{c}\|_Y \leq k_3(t)$$

and thus

$$|\bar{c}| \leq \frac{k_3(t)}{\|A_t\chi_\Omega\|_Y}.$$

We now have

$$\int_0^T \|c(\cdot,t)\|_{L^1(\Omega)}^{\hat{p}} \, \mathrm{d}t \leq \int_0^T \left(k_4 \left(\tilde{k}_1 + \frac{k_3(t) \, \|\chi_\Omega\|_{L^1(\Omega)}}{\|A_t\chi_\Omega\|_Y}\right)\right)^{\hat{p}} \mathrm{d}t \leq k_M < \infty \quad \text{for} \quad 1 \leq \hat{p} \leq 2,$$

where k_M exists as the expression is bounded for all t and thus the supremum is also bounded. Putting it all together, we arrive with $j := \frac{p}{\hat{p}}$ at

$$
\begin{aligned}
\|c\|^{\hat{p}}_{L^{\hat{p}}(0,T;\mathrm{BV}(\Omega))} &= \int_0^T \|c\|^{\hat{p}}_{\mathrm{BV}(\Omega)}\,\mathrm{d}t \\
&\le 2^{\hat{p}-1}\left(\int_0^T |c|^{\hat{p}}_{\mathrm{BV}(\Omega)}\,\mathrm{d}t + \int_0^T \|c\|^{\hat{p}}_{L^1(\Omega)}\,\mathrm{d}t\right) \\
&\le 2^{\hat{p}-1}\int_0^T |c|^{\hat{p}}_{\mathrm{BV}(\Omega)}\,\mathrm{d}t + 2^{\hat{p}-1}k_M \\
&\le 2^{\hat{p}-1}\left(\int_0^T \left|\chi_{[0,T]}\right|^{j^*}\mathrm{d}t\right)^{\frac{1}{j^*}}\left(\int_0^T |c|^{\hat{p}j}_{\mathrm{BV}(\Omega)}\,\mathrm{d}t\right)^{\frac{1}{j}} + 2^{\hat{p}-1}k_M \quad \text{with } \frac{1}{j}+\frac{1}{j^*}=1, \\
&= 2^{\hat{p}-1}k_M + 2^{\hat{p}-1}\left\|\chi_{[0,T]}\right\|_{L^{j^*}(0,T)}\left(\int_0^T |c|^{p}_{\mathrm{BV}(\Omega)}\,\mathrm{d}t\right)^{\frac{1}{j}} \\
&\le 2^{\hat{p}-1}\left(k_M + \left\|\chi_{[0,T]}\right\|_{L^{j^*}(0,T)}\left(\frac{\nu}{\alpha}\right)^{\frac{1}{j}}\right) < \infty.
\end{aligned}
$$

Consider now the norm of v: By assumption, we know that there exists k_∞ such that $\|v\|_{L^\infty(\Omega)^n} \le k_\infty$ almost everywhere on $[0,T]$. We proceed by the following estimate:

$$
\begin{aligned}
\|v\|^q_{L^q(0,T;\mathrm{BV}(\Omega)^n)} &= \int_0^T \left(\|v\|_{L^1(\Omega)^n} + |v|_{\mathrm{BV}(\Omega)^n}\right)^q \mathrm{d}t \\
&\le 2^{q-1}\left(\int_0^T \|v\|^q_{L^1(\Omega)^n}\,\mathrm{d}t + \int_0^T |v|^q_{\mathrm{BV}(\Omega)^n}\,\mathrm{d}t\right) \\
&\le 2^{q-1}\left(\int_0^T \left(k_\infty\,|\Omega|\right)^q \mathrm{d}t + \frac{\nu}{\beta}\right) \\
&= 2^{q-1}\left(T\left(k_\infty\,|\Omega|\right)^q + \frac{\nu}{\beta}\right) < \infty.
\end{aligned}
$$

For the second inequality, we used that $v \in S_\nu$ and thus $\int_0^T \beta\,|v|^q_{\mathrm{BV}(\Omega)^n}\,\mathrm{d}t \le \nu$ by assumption. We conclude that the sublevel sets S_ν are bounded in the norm for fixed $\nu \in \mathbb{R}$, as $\|v\|_{L^q(0,T;\mathrm{BV}(\Omega)^n)}$ and $\|c\|_{L^{\hat{p}}(0,T;\mathrm{BV}(\Omega))}$ are bounded and the bounds depend on ν only. Moreover, we know that $\mathrm{BV}(\Omega)$ is isometrically isomorphic to the dual space of some Banach space Z and thus

$$
\begin{aligned}
L^{\hat{p}}\left(0,T;\mathrm{BV}(\Omega)\right) &\cong \left(L^{p^*}\left(0,T;Z\right)\right)^* \quad \text{with} \quad \frac{1}{\hat{p}}+\frac{1}{p^*}=1, \\
L^{q}\left(0,T;\mathrm{BV}(\Omega)^n\right) &\cong \left(L^{q^*}\left(0,T;Z^n\right)\right)^* \quad \text{with} \quad \frac{1}{q}+\frac{1}{q^*}=1.
\end{aligned}
$$

Consequently, both spaces are dual spaces and thus $S_\nu \subset L^{\hat{p}}(0,T;BV(\Omega)) \times L^q(0,T;BV(\Omega)^n) \cong X^*$ for X^* being the dual of a Banach space X (here $X = L^{p^*}(0,T;Z) \times L^{q^*}(0,T;Z^n)$). Moreover, S_ν is closed in weak-* topology as J is continuous and the sublevel sets are bounded in the norm. It follows that S_ν is weak-* compact by Banach–Alaoglu.

Lower semicontinuity of the functional J with respect to the weak- topology:* As a sum of weak-* lower semicontinuous functionals, J itself is weak-* lower semicontinuous; cf. [17].

Convergence of the constraint: OF: Let $(c_m, v_m) \in L^{\hat{p}}(0,T;BV(\Omega)) \times L^q(0,T;BV(\Omega)^n)$, $m \in \mathbb{N}$, be an admissible sequence (i.e., $c_m \in L^p(0,T,BV(\Omega))$, $\|v_m\|_{L^\infty(0,T;L^\infty(\Omega)^n)} \le k_\infty$,

$\|\nabla \cdot v_m\|_\theta \leq k_\theta$, $m_1(c,v) = 0$ in $\mathcal{D}'(\Omega \times [0,T]))$, which also fulfills $(c_m, v_m) \in S_\nu$ for some $\nu \in \mathbb{R}$. Then c_m and v_m are bounded and there exist c and v such that by passing over to a subsequence (again denoted by c_m, v_m) we have

$$c_m \overset{*}{\rightharpoonup} c, \quad v_m \overset{*}{\rightharpoonup} v.$$

We want to show that

$$\frac{\partial}{\partial t} c_m + \nabla c_m \cdot v_m \longrightarrow \frac{\partial}{\partial t} c + \nabla c \cdot v \quad \text{in } \mathcal{D}'(\Omega \times [0,T]),$$

at least in a distributional sense. Therefore we need strong convergence of at least one factor of $\nabla c_m \cdot v_m$ in a certain sense. In order to use the lemma of Aubin–Lions, we need boundedness of the time derivative of the sequence c_m as well as some specific embeddings. We start with the bound for $\frac{\partial}{\partial t} c$:

$$\begin{aligned}
\left| \int_0^T \int_\Omega \frac{\partial}{\partial t} c\varphi \, \mathrm{d}x\mathrm{d}t \right| &= \left| \int_0^T \int_\Omega c \nabla \cdot (v\varphi) \, \mathrm{d}x\mathrm{d}t \right| \quad \forall \varphi \in \mathcal{C}_C^\infty(\Omega \times (0,T)) \\
&\leq \int_0^T \|c\nabla \cdot (v\varphi)\|_{L^1(\Omega)} \, \mathrm{d}t \\
&\leq \int_0^T \|c\|_{L^l(\Omega)} \|\nabla \cdot (v\varphi)\|_{L^{l^*}(\Omega)} \, \mathrm{d}t \quad \left(\text{Hoelder}: \frac{1}{l} + \frac{1}{l^*} = 1, \ l \leq \frac{n}{n-1} \right) \\
&\leq \int_0^T \|c\|_{L^l(\Omega)} \left(\|\varphi \nabla \cdot v\|_{L^{l^*}(\Omega)} + \|v \cdot \nabla \varphi\|_{L^{l^*}(\Omega)} \right) \mathrm{d}t \quad \text{(Minkowksi)} \\
&= \underbrace{\int_0^T \|c\|_{L^l(\Omega)} \|\varphi \nabla \cdot v\|_{L^{l^*}(\Omega)} \, \mathrm{d}t}_{(I)} + \underbrace{\int_0^T \|c\|_{L^l(\Omega)} \|v \cdot \nabla \varphi\|_{L^{l^*}(\Omega)} \, \mathrm{d}t}_{(II)}.
\end{aligned}$$

The first term (I) can be estimated by

$$\begin{aligned}
&\int_0^T \|c\|_{L^l(\Omega)} \|\varphi \nabla \cdot v\|_{L^{l^*}(\Omega)} \, \mathrm{d}t = \left\| \|c\|_{L^l(\Omega)} \|\varphi \nabla \cdot v\|_{L^{l^*}(\Omega)} \right\|_{L^1(0,T)} \\
&\leq \|c\|_{L^p(0,T;L^l(\Omega))} \|\varphi \nabla \cdot v\|_{L^{p^*}(0,T;L^{l^*}(\Omega))} \quad \left(\text{Hoelder}: \frac{1}{p} + \frac{1}{p^*} = 1, \ 1 < p \leq 2 \right), \\
&= \|c\|_{L^p(0,T;L^l(\Omega))} \left\| \|\varphi \nabla \cdot v\|_{L^{l^*}(\Omega)} \right\|_{L^{p^*}(0,T)} \\
&\leq k_c \left\| \|\varphi\|_{L^{l^* k^*}(\Omega)} \|\nabla \cdot v\|_{L^{l^* k}(\Omega)} \right\|_{L^{p^*}(0,T)} \quad \left(\text{Hoelder}: \frac{1}{l^* k} + \frac{1}{l^* k^*} = \frac{1}{l^*}, \ 1 \leq k \leq \infty \right), \\
&\leq k_c \|\varphi\|_{L^{p^* s^*}(0,T;L^{l^* k^*}(\Omega))} \|\nabla \cdot v\|_{L^{p^* s}(0,T;L^{l^* k}(\Omega))} \quad \left(\text{Hoelder}: \frac{1}{p^* s} + \frac{1}{p^* s^*} = \frac{1}{p^*}, \ 1 \leq s < \infty \right) \\
&\leq k_c \|\varphi\|_{L^{p^* s^*}(0,T;L^{l^* k^*}(\Omega))} k_\theta,
\end{aligned}$$

where we used that c is bounded in $L^p(0,T;\mathrm{BV}(\Omega))$ by assumption and as $\mathrm{BV}(\Omega) \hookrightarrow L^l(\Omega)$ there exists a constant k_c such that $\|c\|_{L^p(0,T;L^l(\Omega))} \leq k_c$. For the second term (II) we find

$$\begin{aligned}
&\int_0^T \|c\|_{L^l(\Omega)} \|v\cdot\nabla\varphi\|_{L^{l^*}(\Omega)}\,\mathrm{d}t \qquad \text{(A.2)}\\
&\quad\le \int_0^T \|c\|_{L^l(\Omega)} \||v|\cdot|\nabla\varphi|\|_{L^{l^*}(\Omega)}\,\mathrm{d}t \quad (\text{Cauchy}-\text{Schwarz inequality}),\\
&\quad\le \int_0^T \|c\|_{L^l(\Omega)} \|v\|_{L^\infty(\Omega)^n} \|\nabla\varphi\|_{L^{l^*}(\Omega)^n}\,\mathrm{d}t\\
&\quad\le \int_0^T \|c\|_{L^l(\Omega)} k_\infty \|\varphi\|_{W^{1,l^*}(\Omega)}\,\mathrm{d}t\\
&\quad\le k_\infty \|c\|_{L^p(0,T;L^l(\Omega))} \|\varphi\|_{L^{p^*}(0,T;W^{1,l^*}(\Omega))} \quad \left(\text{Hoelder}: \frac{1}{p}+\frac{1}{p^*}=1,\ 1<p\le 2\right)\\
&\quad\le k_\infty k_c \|\varphi\|_{L^{p^*}(0,T;W^{1,l^*}(\Omega))}. \qquad \text{(A.3)}
\end{aligned}$$

Combining those results for (I) and (II) yields

$$\begin{aligned}
\left|\int_0^T\int_\Omega \frac{\partial}{\partial t} c\varphi \mathrm{d}x\mathrm{d}t\right| &\le k_c k_\theta \|\varphi\|_{L^{p^*s^*}(0,T;L^{l^*k^*}(\Omega))} + k_\infty k_c \|\varphi\|_{L^{p^*}(0,T;W^{1,l^*}(\Omega))}\\
&\le k_c\left(k_\theta \|\varphi\|_{L^{p^*s^*}(0,T;L^{l^*k^*}(\Omega))} + k_\infty k_6 \|\varphi\|_{L^{p^*s^*}(0,T;W^{1,l^*}(\Omega))}\right).
\end{aligned}$$

The last estimate was obtained by using $s^*>1$ and thus $L^{p^*s^*}(\Omega)\hookrightarrow L^{p^*}(\Omega)$, which defines the constant k_6 by $\|x\|_{L^{p^*}(\Omega)}\le k_6\|x\|_{L^{p^*s^*}(\Omega)}$. We now use the dimension dependent embedding

$$W^{1,l^*}(\Omega)\hookrightarrow L^{l^*k^*}(\Omega),$$

which exists for all $1<k^*<\infty$, as $l^*\ge n$ and then $l^*k^*\ge l^*$.

This yields

$$\begin{aligned}
\left|\int_0^T\int_\Omega \frac{\partial}{\partial t} c\varphi \mathrm{d}x\mathrm{d}t\right| &\le k_c\left(k_\theta \|\varphi\|_{L^{p^*s^*}(0,T;W^{1,l^*}(\Omega))} + k_\infty k_6 \|\varphi\|_{L^{p^*s^*}(0,T;W^{1,l^*}(\Omega))}\right)\\
&= \|\varphi\|_{L^{p^*s^*}(0,T;W^{1,l^*}(\Omega))}\left(k_c k_\theta + k_c k_\infty k_6\right).
\end{aligned}$$

We conclude that $\frac{\partial}{\partial t}c$ acts as a bounded linear functional on $L^{p^*s^*}(0,T;W_0^{1,l^*}(\Omega))$ and thus $\frac{\partial}{\partial t}c\in(L^{p^*s^*}(0,T;W_0^{1,l^*}(\Omega)))^* = L^{\frac{ps}{p+s-1}}(0,T;W^{-1,l}(\Omega))$.

We now use the lemma of Aubin–Lions [4, 50].

Remark A.1. The lemma of Aubin–Lions reads let X,Y,Z be Banach spaces with $X\Subset Y$, $Y\hookrightarrow Z$, let c_i be a sequence of bounded functions in $L^p(0,T;X)$, and $\frac{\partial}{\partial t}cc_i$ be bounded in $L^q(0,T;Z)$ with either $q=1$ and $1\le p<\infty$ or $q>1$ and $1\le p\le\infty$. Then c_i is relatively compact in $L^p(0,T;Y)$. We want to identify the Banach spaces as $X=\mathrm{BV}(\Omega)$, $Y=L^r(\Omega)$, and $Z=W^{-1,l}(\Omega)$.

The embedding $\mathrm{BV}(\Omega)\Subset L^r(\Omega)$ holds for $r<\frac{n}{n-1}$ [2], i.e.,

$$\begin{cases} r<2, & n=2,\\ r<1.5, & n=3. \end{cases}$$

It remains to identify for which r the continuous embedding $L^r(\Omega) \hookrightarrow W^{-1,l}(\Omega)$ exists. The embedding holds if $W^{1,l^*}(\Omega) \hookrightarrow L^{r^*}(\Omega)$ which by the Sobolev embedding theorem for the different dimensions and $l^* \geq n$ translates to

$$W^{1,l^*}(\Omega) \hookrightarrow L^{r^*}(\Omega) \text{ for } l^* \leq r^* < \infty. \tag{A.4}$$

In terms of r, this results in $1 < r \leq l$.

Applying Aubin–Lions thus yields $\{c \in L^p(0,T;\mathrm{BV}(\Omega)) : \exists k > 0 \text{ s.t. } \|c\|_{L^p(0,T;\mathrm{BV}(\Omega))} \leq k, \frac{\partial}{\partial t}c + v \cdot \nabla c = 0 \text{ in } \mathcal{D}'(\Omega \times [0,T])\}$ is relatively compact in $L^p(0,T;L^r(\Omega))$ for r satisfying

$$\begin{cases} 1 \leq r < 2, & n = 2, \\ 1 \leq r < \frac{3}{2}, & n = 3. \end{cases} \tag{A.5}$$

The sequence c_m thus converges strongly even to c in $L^p(0,T;L^r(\Omega))$.

We are now settled to prove convergence of the constraint. In the following, let $\varphi \in \mathcal{C}_0^\infty(\Omega \times [0,T])$. We start with the time derivative,

$$\int_0^T \int_\Omega \left(\left(\frac{\partial}{\partial t} c \right)_m - \frac{\partial}{\partial t} c \right) \varphi \mathrm{d}x\mathrm{d}t = - \int_0^T \int_\Omega (c_m - c) \frac{\partial}{\partial t} \varphi \mathrm{d}x\mathrm{d}t,$$

by integration by parts. We know that $c_m \overset{*}{\rightharpoonup} c$ in $L^p(0,T;\mathrm{BV}(\Omega)) \cong \left(L^{p^*}(0,T;Z)\right)^*$ for $\mathrm{BV}(\Omega)$ being isometrically isomorphic to the dual space of Z. As $\varphi \in \mathcal{C}_0^\infty(\Omega \times [0,T])$, it follows that $\frac{\partial}{\partial t}\varphi \in L^{p^*}(0,T;Z)$ and this yields

$$- \int_0^T \int_\Omega (c_m - c) \frac{\partial}{\partial t} \varphi \mathrm{d}x\mathrm{d}t \longrightarrow 0 \quad \text{for } m \longrightarrow \infty.$$

For the product term, it holds that

$$\begin{aligned} - \int_0^T \int_\Omega (\nabla c_m \cdot v_m - \nabla c \cdot v) \varphi \, \mathrm{d}x \, \mathrm{d}t &= \int_0^T \int_\Omega c_m \nabla \cdot (v_m \varphi) - c \nabla \cdot (v \varphi) \, \mathrm{d}x \, \mathrm{d}t \\ &= \underbrace{\int_0^T \int_\Omega (c_m - c) \nabla \cdot (v_m \varphi) \, \mathrm{d}x \, \mathrm{d}t}_{(I)} + \underbrace{\int_0^T \int_\Omega c \nabla \cdot (\varphi v_m - \varphi v) \, \mathrm{d}x \, \mathrm{d}t}_{(II)}. \end{aligned}$$

For (I) we obtain

(A.6)

$$\begin{aligned} &\int_0^T \int_\Omega (c_m - c) \nabla \cdot (v_m \varphi) \, \mathrm{d}x\mathrm{d}t \\ &\quad \leq \int_0^T \|c_m - c\|_{L^r(\Omega)} \|\nabla \cdot (v_m \varphi)\|_{L^{r^*}(\Omega)} \, \mathrm{d}t \quad \left(\text{Hoelder}: \frac{1}{r} + \frac{1}{r^*} = 1, \ 1 < r < \infty \right) \\ &\quad \leq \|c_m - c\|_{L^p(0,T;L^r(\Omega))} \|\nabla \cdot (v_m \varphi)\|_{L^{p^*}(0,T;L^{r^*}(\Omega))} \quad \left(\text{Hoelder}: \frac{1}{p} + \frac{1}{p^*} = 1, \ 1 < p \leq 2 \right) \\ &\quad \leq \|c_m - c\|_{L^p(0,T;L^r(\Omega))} \left(\|\nabla \varphi \cdot v_m\|_{L^{p^*}(0,T;L^{r^*}(\Omega))} + \|\varphi \nabla \cdot (v_m)\|_{L^{p^*}(0,T;L^{r^*}(\Omega))} \right) \\ &\quad \leq \|c_m - c\|_{L^p(0,T;L^r(\Omega))} \left(\underbrace{\|\nabla \varphi \cdot v_m\|_{L^{p^*}(0,T;L^{r^*}(\Omega))}}_{(a)} \right. \end{aligned}$$

(A.7)

$$+ \underbrace{\|\varphi\|_{L^{p^*s^*}(0,T;L^{r^*}(\Omega))}}_{(b)} \underbrace{\|\nabla\cdot(v_m)\|_{L^{p^*s}(0,T;L^{r^*}(\Omega))}}_{(c)} \Bigg) \quad \left(\frac{1}{p^*s} + \frac{1}{p^*s^*} = \frac{1}{p^*},\ 1<s<\infty \right).$$

We note that (b) is bounded as φ is a test function. Consider now (c): $\|\nabla\cdot(v_m)\|_{L^{p^*s}(0,T;L^{l^*k}(\Omega))}$ is bounded by k_θ by assumption. We thus need an embedding

$$L^{p^*s}(0,T;L^{l^*k}(\Omega)) \hookrightarrow L^{p^*s}(0,T;L^{r^*}(\Omega))$$

and therefore $r^* \le l^*k$ which is equivalent to $r \ge \frac{l^*k}{l^*k-1}$. We need to adjust the bounds on k as follows.

- *Case* $n=2$: We have $1 \le r < 2$ and $l^* \ge 2$. $1 \le \frac{l^*k}{l^*k-1}$ is always fulfilled, but in order to guarantee $\frac{l^*k}{l^*k-1} < 2$, we have to restrict k such that $[\frac{l^*k}{l^*k-1}, 2) \neq \emptyset$. This is ensured by $l^*k > 2$, which is equivalent to $k > 1$, as $l^* \ge 2$.
- *Case* $n=3$: We have $1 \le r < \frac{3}{2}$ and $l^* \ge 3$. $1 \le \frac{l^*k}{l^*k-1}$ is always fulfilled, but in order to guarantee $\frac{l^*k}{l^*k-1} < \frac{3}{2}$, we have to restrict k such that $[\frac{l^*k}{l^*k-1}, \frac{3}{2}) \neq \emptyset$. This is ensured by $l^*k > 3$, which is equivalent to $k > 1$, as $l^* \ge 3$.

It remains to show that (a) is bounded. This holds as φ is a test function and thus $\nabla\varphi$ is bounded and $\|v\|_{L^\infty(\Omega)^n} \le k_\infty$. We showed that $(I) \longrightarrow 0$ as $m \longrightarrow \infty$. Now consider (II), i.e.,

(A.8)

$$\int_0^T \int_\Omega c\nabla\cdot(\varphi v_m - \varphi v)\,\mathrm{d}x\,\mathrm{d}t = \underbrace{\int_0^T \int_\Omega c\varphi\nabla\cdot(v_m - v)\,\mathrm{d}x\,\mathrm{d}t}_{(a)} + \underbrace{\int_0^T \int_\Omega c\,(v_m - v)\cdot\nabla\varphi\,\mathrm{d}x\mathrm{d}t}_{(b)}.$$

Consider (a) first. By assumption, $\nabla\cdot(v_m - v) \in L^{p^*s}\left(0,T;L^{l^*k}(\Omega)\right)$. We need to show $c\varphi \in \left(L^{p^*s}\left(0,T;L^{l^*k}(\Omega)\right)\right)^*$. Starting from $c \in L^p(0,T;\mathrm{BV}(\Omega))$, the first embedding needed is $\mathrm{BV}(\Omega) \hookrightarrow L^{\frac{l^*k}{l^*k-1}}(\Omega) = L^{(l^*k)^*}(\Omega)$.

- *Case* $n=2$: $\mathrm{BV}(\Omega) \hookrightarrow L^{\frac{l^*k}{l^*k-1}}(\Omega)$ exists continuously for $1 \le \frac{l^*k}{l^*k-1} \le 2$. This is fulfilled for $l^*k \ge 2$, which is ensured by $l^* \ge 2$ and $k > 1$.
- *Case* $n=3$: $\mathrm{BV}(\Omega) \hookrightarrow L^{\frac{l^*k}{l^*k-1}}(\Omega)$ exists for $1 \le \frac{l^*k}{l^*k-1} \le 1.5$. Again, this holds for $l^*k \ge 3$, which is ensured by $l^* \ge 3$ and $k > 1$.

Note that $(p^*s)^* = \frac{ps}{ps-p+1}$ and $\frac{ps}{ps-p+1} < p$ if $1 < p$, which again holds by assumption.

In summary, we have

$$L^p(0,T;\mathrm{BV}(\Omega)) \hookrightarrow L^p\left(0,T;L^{(l^*k)^*}(\Omega)\right) \hookrightarrow L^{(p^*s)^*}\left(0,T;L^{(l^*k)^*}(\Omega)\right).$$

As φ and its derivatives are bounded, $c\varphi \in L^{(p^*s)^*}(0,T;L^{(l^*k)^*}(\Omega))$. Moreover, from $v_m \overset{*}{\rightharpoonup} v$ it follows that $\nabla\cdot v_m \overset{*}{\rightharpoonup} \nabla\cdot v$ and thus altogether

$$\int_0^T \int_\Omega c\varphi\nabla\cdot(v_m - v)\,\mathrm{d}x\mathrm{d}t \longrightarrow 0 \text{ for } m \longrightarrow \infty.$$

Considering (b), we first denote that

$$\mathrm{BV}(\Omega) \hookrightarrow L^1(\Omega) \;\Rightarrow\; L^p(0,T;\mathrm{BV}(\Omega)) \hookrightarrow L^p\left(0,T;L^1(\Omega)\right) \hookrightarrow L^1\left(0,T;L^1(\Omega)\right).$$

Therefore, we have $c \in L^1(0,T;L^1(\Omega))$ and by the boundedness of φ and its derivatives, we obtain $c\nabla\varphi \in L^1(0,T;L^1(\Omega)^n)$. Moreover, by $\|v\|_{L^\infty(\Omega)^n} \leq k_\infty$ we see that $(v_m - v) \in L^\infty(0,T;L^\infty(\Omega)^n)$. Now by

$$L^\infty(0,T;L^\infty(\Omega)^n) \cong \left(L^1\left(0,T;L^1(\Omega)^n\right)\right)^*$$

and the weak-* convergence of v_m to v, we reach

$$\int_0^T \int_\Omega c\nabla\varphi \cdot (v_m - v)\,\mathrm{d}x\mathrm{d}t \longrightarrow 0 \text{ for } m \longrightarrow \infty.$$

Combining all the results yields the convergence of the constraint

$$\int_0^T \int_\Omega \left(\left(\left(\frac{\partial}{\partial t}c\right)_m - \nabla c_m \cdot v_m\right) - \left(\frac{\partial}{\partial t}c - \nabla c \cdot v\right)\right)\varphi \mathrm{d}x\mathrm{d}t \longrightarrow 0 \text{ for } m \longrightarrow \infty.$$

Convergence of the constraint: MC: Let $(c_m, v_m) \in L^{\hat{p}}(0,T;BV(\Omega)) \times L^q(0,T;BV(\Omega)^n)$, $m \in \mathbb{N}$, be an admissible sequence (i.e., $c_m \in L^p(0,T,BV(\Omega))$, $\|v_m\|_{L^\infty(0,T;L^\infty(\Omega)^n)} \leq k_\infty$, $\|\nabla \cdot v_m\|_\theta \leq k_\theta$, $m_2(c,v) = 0$ in $\mathcal{D}'(\Omega \times [0,T])$), which also fulfills $(c_m, v_m) \in S_\nu$ for some $\nu \in \mathbb{R}$. Then c_m and v_m are bounded and there exist c and v such that by passing over to a subsequence (again denoted by c_m, v_m) we have

$$c_m \overset{*}{\rightharpoonup} c, \quad v_m \overset{*}{\rightharpoonup} v.$$

We want to show that

$$\left(\frac{\partial}{\partial t}c\right)_m + \nabla \cdot (c_m v_m) \longrightarrow \frac{\partial}{\partial t}c + \nabla \cdot (cv) \quad \text{in } \mathcal{D}'(\Omega \times [0,T]),$$

at least in a distributional sense. We start again with a bound on $\frac{\partial}{\partial t}c$:

$$\begin{aligned}
\left|\int_0^T \int_\Omega \frac{\partial}{\partial t}c\varphi \mathrm{d}x\mathrm{d}t\right| &= \left|\int_0^T \int_\Omega -\nabla \cdot (cv)\,\varphi \mathrm{d}x\mathrm{d}t\right| = \left|\int_0^T \int_\Omega (cv) \cdot \nabla\varphi \mathrm{d}x\mathrm{d}t\right| \\
&\leq \int_0^T \int_\Omega |(cv) \cdot \nabla\varphi|\,\mathrm{d}x\mathrm{d}t \\
&\leq \int_0^T \|c\|_{L^l(\Omega)} \|v \cdot \nabla\varphi\|_{L^{l^*}(\Omega)}\,\mathrm{d}t \quad \left(\text{Hoelder}: \frac{1}{l} + \frac{1}{l^*} = 1,\ l \leq \frac{n}{n-1}\right), \\
&\leq k_\infty k_c \|\varphi\|_{L^{p^*}(0,T;W^{1,l^*}(\Omega))} \quad (\text{c.f. (B.2)} - \text{(B.3)}).
\end{aligned} \tag{A.9}$$

We thus know that $\frac{\partial}{\partial t}c$ acts as a bounded linear functional on $L^{p^*}\left(0,T;W^{1,l^*}(\Omega)\right)$ and thus $\frac{\partial}{\partial t}c \in L^p\left(0,T;W^{-1,l}(\Omega)\right)$. The lemma of Aubin–Lions can be applied similarly to the OF case

and yields strong convergence of $c_m \longrightarrow c$ in $L^p(0,T;L^r(\Omega))$ with the same constraints on r. Also, the arguments for the convergence of the time derivative

$$-\int_0^T \int_\Omega (c_m - c)\, \frac{\partial}{\partial t}\varphi \mathrm{d}x\mathrm{d}t \longrightarrow 0 \quad \text{for } m \longrightarrow \infty$$

are the same as in the OF case. We now come to the product term. It remains to show that

$$\nabla \cdot (c_m v_m) \rightharpoonup \nabla \cdot (cv)$$

$$\begin{aligned}
&-\int_0^T \int_\Omega (\nabla \cdot (c_m v_m) - \nabla \cdot (cv))\, \varphi \mathrm{d}x\mathrm{d}t = \int_0^T \int_\Omega (c_m v_m - cv \pm c v_m) \cdot \nabla \varphi \mathrm{d}x\mathrm{d}t \\
&= \int_0^T \int_\Omega ((c_m - c)\, v_m + c\,(v_m - v)) \cdot \nabla \varphi \,\mathrm{d}x\,\mathrm{d}t \\
&\le \int_0^T \|c_m - c\|_{L^r(\Omega)}\, \|v_m \cdot \nabla \varphi\|_{L^{r^*}(\Omega)} + \int_\Omega c\,(v_m - v) \cdot \nabla \varphi \mathrm{d}x\mathrm{d}t \quad \left(\frac{1}{r} + \frac{1}{r^*} = 1\right) \\
&\le \underbrace{\|c_m - c\|_{L^p(0,T;L^r(\Omega))}\, \|v_m \cdot \nabla \varphi\|_{L^{p^*}(0,T;L^{r^*}(\Omega))}}_{\longrightarrow 0 \quad \text{for } m \longrightarrow \infty \text{ (cf. (B.7)}(a)\text{)}} + \underbrace{\int_0^T \int_\Omega c\,(v_m - v) \cdot \nabla \varphi \mathrm{d}x\mathrm{d}t}_{\longrightarrow 0 \quad \text{for } m \longrightarrow \infty \text{ (cf. (B.8)}(b)\text{)}}. \qquad \text{(A.10)}
\end{aligned}$$

Appendix B. Extension Remark 2.4. Analysis of the regularization properties of the unconstrained approach (2.14) might be carried out by exploiting the established theory for Tikhonov-type regularization in Banach spaces in the nonlinear problem setting (see, e.g., [32, 66]). Even if the image reconstruction problem with respect to c is linear, the reconstruction of the tuple (c, v) is nonlinear due to the PDE constraint defined by the motion prior m_i. Regularization properties of the proposed method can be guaranteed by suitable choices of penalty terms S and R, e.g., in line with the previous setting are choices such that $\|R(c(\cdot))\|_{L^1(0,T)}$ and $\|S(v(\cdot))\|_{L^1(0,T)}$ are proper, convex, and weak-* lower semicontinuous functionals with respect to $(c, v) \in L^{\hat{p}}(0,T;BV(\Omega)) \times L^q(0,T;BV(\Omega)^n)$. The constraint $\|v\|_{L^\infty(0,T;L^\infty(\Omega)^n)} \le k_\infty$ is then again sufficient to guarantee weak-* closedness of the sublevel sets (see the proof of Theorem 2.1). The crucial and more sophisticated remaining part is to show that the nonlinear operator $F : \mathrm{dom}(F) \subset U \to L^2(0,T,Y) \times L^s(0,T,L^r(\Omega))$ for certain $s, r \ge 1$ with

$$(c, v) \mapsto (A(c(\cdot), \cdot), m_i(c, v)) \qquad \text{(B.1)}$$

is weak-*-to-weak continuous with respect to a suitable space

$$U \subset (L^{\hat{p}}(0,T;BV(\Omega)) \times L^q(0,T;BV(\Omega)^n))$$

and a suitable weak-* closed domain $\mathrm{dom}(F)$ of F. In the first component one has to deal with a bounded linear operator mapping to the reflexive Banach space $L^2(0,T;Y)$. It would thus be sufficient to verify that the linear operator is weak-*-to-weak-* continuous with respect to c. For the second component we need weak-*-to-weak continuity with respect to the tuple. The interplay between weak-*-to-weak continuity and a suitable choice of spaces replaces the

motion constraint being fulfilled in a distributional sense in the setting of Theorem 2.1. Choosing $L^s(0,T,L^r(\Omega))$ includes the motion constraint in a stronger sense which requires further regularity of c, e.g., we need existence of $\frac{\partial}{\partial t}c$ by making suitable assumptions on U implying weak convergence of the derivatives of c. A simple but more restrictive example would be the reflexive space $U = H^1(0,T;H^1(\Omega)) \times L^2(0,T;L^2(\Omega))$ (where weak-* and weak convergence are equivalent) and $p = q = r = s = 2$, where $\mathrm{dom}(F) = \{(c,v) \in U | \|v\|_{L^\infty(0,T;L^\infty(\Omega)^n)} \leq k_\infty \wedge \|\nabla c\|_{L^\infty(0,T;L^\infty(\Omega)^n)} \leq k'_\infty\}$. A more general investigation in this direction is beyond the scope of the present work but under these assumptions the regularization properties would then be given by the general theory for noisy measurements $(u^\varepsilon 0)$; see, for example, [32]. This perspective on PDE-constraint parameter identification in the dynamic setting has also been considered as an all-at-once approach to dynamic inverse problems from a more general point of view [35, 59].

Appendix C. Simulation parameters. The scanner is modeled as follows: The drive-field frequencies are given by $f_x = 2.5\,\mathrm{MHz}/102$, $f_y = 2.5\,\mathrm{MHz}/96$, and $f_z = 2.5\,\mathrm{MHz}/99$ with an amplitude of $14\,\mathrm{mT}/\mu_0$, resulting in a repetition time of 21.54 ms. The selection field gradient strength is $0.5\,\mathrm{T/m}/\mu_0$ in the x- and y-directions and $1.0\,\mathrm{T/m}/\mu_0$ in the z-direction. The magnetic nanoparticles are modeled as stated in Table 4.

We model ideal magnetic fields and use the Langevin magnetization model. In the forward operator voltage measurements for three receive channels in respective unit vector directions are used. To avoid an inverse crime, simulations are performed on a finer grid compared to the reconstruction, i.e., we use 40^3 voxels of size 1 mm $\times$ 1 mm $\times$ 0.5 for simulation and 20^3 voxels of size 2 mm $\times$ 2 mm $\times$ 1 for reconstruction. For reconstruction, the time-dependent signals are transformed to the Fourier domain, a frequency selection is performed, and the real and imaginary parts are split, i.e., concatenated.

Appendix D. Information on parameter search regions for the numerical experiments. In each subsection in section 4, we stated the parameters used for the illustrative reconstructed image sequences. These were obtained by detailed parameter searches in each case and for each algorithm.

D.1. The simulated data experiments. For the Kaczmarz algorithm, we tested for the early stopping index k and the Tikhonov regularization parameter λ. The parameter λ was tested in the range between 10^{-5} and 10^3, k in-between 1 and 10.

Frame-by-frame the SPDHG algorithm has two parameters, α_1 corresponding to the L^1-penalty term and α_2 corresponding to the TV penalty term. The testing range for α_1 was in-between 10^{-3} and 10^{-1}, for α_2 in-between 10^{-9} and 10^{-5}.

Table 4
Physical constants and parameters for modeling the tracer particles.

Parameter	Value
Permeability constant μ_0	$4\pi \cdot 10^{-7}\,\mathrm{N/A^2}$
Particle core diameter	$2 \cdot 10^{-8}$ m
Core saturation magnetization	$0.6\,\mathrm{T}/\mu_0$

Parameter tests for the joint approaches were performed on 10 time steps simultaneously within the ranges stated in Table 5.

The resulting image sequences for each parameter combination was tested for the SSIM value. The best parameter combination is stated in Table 6.

D.2. The rotation phantom data experiments. For the Kaczmarz algorithm, we tested for the early stopping index k and the Tikhonov regularization parameter λ. The parameter λ was tested in the range between 10^{-4} and 30, k in-between 1 and 100.

Frame-by-frame the SPDHG algorithm has two parameters, α_1 corresponding to the L^1-penalty term and α_2 corresponding to the TV penalty term. The testing range for α_1 was in-between 10^{-4} and 1, for α_2 in-between 10^{-4} and 10.

Parameter tests for the joint approaches were performed on 15 time steps simultaneously within the ranges stated in Table 7.

The resulting image sequences were visually inspected and the most convincing ones were chosen. The search area was iteratively reduced and refined until no visual differences were observed.

The resulting parameters for the 3 Hz dataset are stated in Table 2 and below for completeness. Parameter tests for the 1 Hz and 7 Hz datasets are performed analogously and yield similar results.

Table 5

Parameter search area for the joint approaches for the simulated data.

Parameter	Min. value	Max. value
α_1	10^{-3}	10^{-1}
α_2	10^{-8}	10^{-5}
β	10^{-2}	1
γ	10^{-5}	1

Table 6

The best parameters obtained for the reconstruction of the simulated data in terms of SSIM value.

Algorithm	Motion model	Early stopping	λ	α_1	α_2	β	γ
Kaczmarz	None	$k=3$	10^3				
SPDHG	None	$k=2\cdot10^3$		10^{-2}	10^{-7}		
Joint, L^1-D	OF	$k=10^2$		10^{-2}	10^{-7}	10^{-1}	10^{-4}
Joint, L^2-D	OF	$k=10^2$		10^{-2}	10^{-7}	10^{-1}	10^{-4}
Joint, L^1-D	MC	$k=10^2$		10^{-2}	10^{-8}	10^{-1}	10^{-4}
Joint, L^2-D	MC	$k=10^2$		10^{-2}	10^{-5}	10^{-1}	10^{-4}

Table 7

Parameter search area for the joint approaches for the rotation phantom data.

Parameter	Min. value	Max. value
α_1	10^{-2}	10
α_2	10^{-2}	100
β	10^{-2}	1
γ	1	200

Table 8
The parameters used for the reconstruction of the 3 Hz *rotation phantom data.*

Algorithm	Motion model	Early stopping	λ	α_1	α_2	β	γ
Kaczmarz	None	$k = 10$	5.62				
SPDHG	None	$k = 10^3$		$5.0 \cdot 10^{-1}$	$3.0 \cdot 10^{-1}$		
Joint, L^1-D	OF	$k = 10^2$		$6.0 \cdot 10^{-1}$	$1.0 \cdot 10^{-1}$	10^{-1}	100
Joint, L^2-D	OF	$k = 10^2$		$2.5 \cdot 10^{-1}$	$1.0 \cdot 10^{+2}$	10^{-1}	100
Joint, L^1-D	MC	$k = 10^2$		$7.0 \cdot 10^{-1}$	$5.0 \cdot 10^{-1}$	10^{-1}	100
Joint, L^2-D	MC	$k = 10^2$		$2.5 \cdot 10^{-1}$	$2.5 \cdot 10^{-1}$	10^{-1}	100

Table 9
Parameter search area for the MC L^1-D algorithm for the in vivo mouse data reconstruction.

Parameter	Min. value	Max. value
α_1	10^{-2}	10^{-1}
α_2	10^{-3}	200
β	10^{-3}	10^{-1}
γ	1	1000

D.3. The in vivo mouse data data experiments. Motivated by the specific application, we use the MC constrained L^1-algorithm. The parameter search area is indicated in Table 9, tests were performed on 30 time steps simultaneously.

The resulting reconstruction parameters are given by $\alpha_1 = 0.05$, $\alpha_2 = 100$, $\beta = 0.1$, and $\gamma = 100$.

REFERENCES

[1] H. W. Alt, *Linear Functional Analysis: An Application-Oriented Introduction*, Springer, London, 2016.
[2] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Clarendon Press, Oxford, 2000.
[3] S. Arridge, P. Fernsel, and A. Hauptmann, *Joint reconstruction and low-rank decomposition for dynamic inverse problems*, Inverse Probl. Imaging, 16 (2022), pp. 483–523, https://doi.org/10.3934/ipi.2021059.
[4] J. Aubin, *Un théorème de compacité*, C. R. Acad. Sc. Paris, 256 (1963), pp. 5042–5044.
[5] S. E. Blanke, B. N. Hahn, and A. Wald, *Inverse problems with inexact forward operator: Iterative regularization and application in dynamic imaging*, Inverse Problems, 36 (2020), 124001.
[6] C. Blondel, R. Vaillant, G. Malandain, and N. Ayache, *3D tomographic reconstruction of coronary arteries using a precomputed 4D motion field*, Phys. Med. Biol., 49 (2004), pp. 2197–2208.
[7] C. Brandt and C. Schmidt, *Modeling magnetic particle imaging for dynamic tracer distributions*, Sensing and Imaging, 22 (2021), 45.
[8] C. Brandt and C. Schmidt, *Motion compensation for non-periodic dynamic tracer distributions in multi-patch magnetic particle imaging*, Phys. Med. Biol., 67 (2022), 085005.
[9] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, *High accuracy optical flow estimation based on a theory for warping*, in European Conference on Computer Vision, T. Pajdla and J. Matas, eds., Springer, Berlin, 2004, pp. 25–36.
[10] M. Burger, H. Dirks, L. Frerking, A. Hauptmann, T. Helin, and S. Siltanen, *A variational reconstruction method for undersampled dynamic x-ray tomography based on physical motion models*, Inverse Problems, 33 (2017), 124008.
[11] M. Burger, H. Dirks, and C.-B. Schönlieb, *A variational model for joint motion estimation and image reconstruction*, SIAM J. Imaging Sci., 11 (2018), pp. 94–128.

[12] A. Chambolle, M. J. Ehrhardt, P. Richtárik, and C.-B. Schönlieb, *Stochastic primal-dual hybrid gradient algorithm with arbitrary sampling and imaging applications*, SIAM J. Optim., 28 (2018), pp. 2783–2808.
[13] C. Chen, B. Gris, and O. Öktem, *A new variational model for joint image reconstruction and motion estimation in spatiotemporal imaging*, SIAM J. Imaging Sci., 12 (2019), pp. 1686–1719.
[14] C. Crawford, K. King, C. Ritchie, and J. Godwin, *Respiratory compensation in projection imaging using a magnification and displacement model*, IEEE Trans. Med. Imaging, 15 (1996), pp. 327–332.
[15] A. Dax, *On row relaxation methods for large constrained least squares problems*, SIAM J. Sci. Comput., 14 (1993), pp. 570–584.
[16] L. Desbat, S. Roux, and P. Grangeat, *Compensation of some time dependent deformations in tomography*, IEEE Trans. Med. Imaging, 26 (2007), pp. 261–269.
[17] H. Dirks, *Variational Methods for Joint Motion Estimation and Image Reconstruction*, Ph.D. thesis, Universitäts-und Landesbibliothek Münster, Münster, Germany, 2015.
[18] W. Erb, et al., *Mathematical analysis of the* 1*D model and reconstruction schemes for magnetic particle imaging*, Inverse Problems, 34 (2018), 055012.
[19] A. Fasano and A. Sequeira, *Hemomath: The Mathematics of Blood*, MS&A. Model. Simul. Appl. 18, Springer, Cham, Switzerland, 2017.
[20] J. Franke, R. Lacroix, H. Lehr, M. Heidenreich, U. Heinen, and V. Schulz, *MPI flow analysis toolbox exploiting pulsed tracer information - An aneurysm phantom proof*, Int. J. Magn. Part. Imag., 3 (2017).
[21] N. Gdaniec, M. Boberg, M. Möddel, P. Szwargulski, and T. Knopp, *Suppression of motion artifacts caused by temporally recurring tracer distributions in multi-patch magnetic particle imaging*, IEEE Trans. Med. Imaging, 39 (2020), pp. 3548–3558.
[22] N. Gdaniec, M. Schlüter, M. Möddel, M. G. Kaul, K. Krishnan, A. Schlaefer, and T. Knopp, *Detection and compensation of periodic motion in magnetic particle imaging*, IEEE Trans. Med. Imaging, 36 (2017), pp. 1511–1521.
[23] B. Gleich and J. Weizenecker, *Tomographic imaging using the nonlinear response of magnetic particles*, Nature, 435 (2005), pp. 1214–1217.
[24] M. Gräser, T. Knopp, P. Szwargulski, T. Friedrich, A. von Gladiss, M. Kaul, K. M. Krishnan, H. Ittrich, G. Adam, and T. M. Buzug, *Towards picogram detection of superparamagnetic iron-oxide particles using a gradiometric receive coil*, Sci. Rep., 7 (2017), pp. 1–13.
[25] M. Gräser, et al., *Human-sized magnetic particle imaging for brain applications*, Nature Commun., 10 (2019), 1936.
[26] J. Haegele, J. Rahmer, B. Gleich, J. Borgert, H. Wojtczyk, N. Panagiotopoulos, T. Buzug, J. Barkhausen, and F. Vogt, *Magnetic particle imaging: Visualization of instruments for cardiovascular intervention*, Radiology, 265 (2012), pp. 933–938.
[27] B. Hahn, *Reconstruction of dynamic objects with affine deformations in dynamic computerized tomography*, J. Inverse Ill-Posed Probl., 22 (2014), pp. 323–339.
[28] B. N. Hahn, *Efficient algorithms for linear dynamic inverse problems with known motion*, Inverse Problems, 30 (2014), 035008.
[29] B. N. Hahn, *Motion estimation and compensation strategies in dynamic computerized tomography*, Sensing and Imaging, 18 (2017), pp. 1–20.
[30] A. Hauptmann, S. Arridge, F. Lucka, V. Muthurangu, and J. A. Steeden, *Real-time cardiovascular MR with spatio-temporal artifact suppression using deep learning–proof of concept in congenital heart disease*, Magn. Reson. Med., 81 (2019), pp. 1143–1156.
[31] A. Hauptmann, O. Öktem, and C. Schönlieb, *Image reconstruction in dynamic inverse problems with temporal models*, in Handbook of Mathematical Models and Algorithms in Computer Vision and Imaging: Mathematical Imaging and Vision, Springer, Cham, Switzerland, 2021, pp. 1–31.
[32] B. Hofmann, B. Kaltenbacher, C. Poeschl, and O. Scherzer, *A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators*, Inverse Problems, 23 (2007), 987.
[33] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox, *Flownet* 2.0*: Evolution of optical flow estimation with deep networks*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Piscataway, NJ, 2017, pp. 2462–2470.

[34] A. Isola, A. Ziegler, T. Koehler, W. Niessen, and M. Grass, *Motion-compensated iterative cone-beam CT image reconstruction with adapted blobs as basis functions*, Phys. Med. Biol., 53 (2008), pp. 6777–6797.
[35] B. Kaltenbacher, *All-at-once versus reduced iterative methods for time dependent inverse problems*, Inverse Problems, 33 (2017), 064002.
[36] A. Katsevich, *An accurate approximate algorithm for motion compensation in two-dimensional tomography*, Inverse Problems, 26 (2010), 065007.
[37] A. Katsevich, M. Silver, and A. Zamayatin, *Local tomography and the motion estimation problem*, SIAM J. Imaging Sci., 4 (2011), pp. 200–219.
[38] M. G. Kaul, J. Salamon, T. Knopp, H. Ittrich, G. Adam, H. Weller, and C. Jung, *Magnetic particle imaging for in vivo blood flow velocity measurements in mice*, Phys. Med. Biol., 63 (2018), 064001.
[39] A. Khandhar, P. Keselman, S. Kemp, R. Ferguson, P. Goodwill, S. Conolly, and K. Krishnan, *Evaluation of PEG-coated iron oxide nanoparticles as blood pool tracers for preclinical magnetic particle imaging*, Nanoscale, 9 (2017), pp. 1299–1306.
[40] S. Kindermann and A. Leitão, *On regularization methods for inverse problems of dynamic type*, Numer. Funct. Anal. Optim., 27 (2006), pp. 139–160.
[41] T. Kluth, *Mathematical models for magnetic particle imaging*, Inverse Problems, 34 (2018), 083001, http://iopscience.iop.org/article/10.1088/1361-6420/aac535.
[42] T. Kluth, B. Hahn, and C. Brandt, *Spatio-temporal concentration reconstruction using motion priors in magnetic particle imaging*, in 9th International Workshop on Magnetic Particle Imaging 2019, T. Knopp and T. M. Buzug, eds., Infinite Science Publishing, Hingham, MA, 2019, pp. 23–24.
[43] T. Kluth and B. Jin, *Enhanced reconstruction in magnetic particle imaging by whitening and randomized SVD approximation*, Phys. Med. Biol., 64 (2019), 125026, https://doi.org/10.1088/1361-6560/ab1a4f.
[44] T. Kluth and B. Jin, *L1 data fitting for robust numerical reconstruction in magnetic particle imaging: Quantitative evaluation on* `Open MPI dataset`, Int. J. Magn. Part. Imag., 6 (2020), 2012001, https://journal.iwmpi.org/index.php/iwmpi/article/view/146.
[45] T. Kluth, B. Jin, and G. Li, *On the degree of ill-posedness of multi-dimensional magnetic particle imaging*, Inverse Problems, 34 (2018), 095006, https://doi.org/10.1088/1361-6420/aad015.
[46] T. Knopp and T. M. Buzug, *Magnetic Particle Imaging: An Introduction to Imaging Principles and Scanner Instrumentation*, Springer, Berlin, 2012.
[47] T. Knopp, J. Rahmer, T. F. Sattel, S. Biederer, J. Weizenecker, B. Gleich, J. Borgert, and T. M. Buzug, *Weighted iterative reconstruction for magnetic particle imaging*, Phys. Med. Biol., 55 (2010), 1577.
[48] T. Knopp and A. Weber, *Local system matrix compression for efficient reconstruction in magnetic particle imaging*, Adv. Math. Phys., 2015 (2015), 472818.
[49] T. Li, M. Zhang, W. Qi, E. Asma, and J. Qi, *Deep learning based joint pet image reconstruction and motion estimation*, IEEE Trans. Med. Imaging, 41 (2022), pp. 1230–1241, https://doi.org/10.1109/TMI.2021.3136553.
[50] J.-L. Lions, *Quelques méthodes de résolution de problemes aux limites non linéaires*, Dunod, Paris, 1969.
[51] J. Liu, X. Zhang, X. Zhang, H. Zhao, Y. Gao, D. Thomas, D. Low, and H. Gao, *5D respiratory motion model based image reconstruction algorithm for 4D cone-beam computed tomography*, Inverse Problems, 31 (2015), 115007.
[52] W. Lu and T. R. Mackie, *Tomographic motion detection and correction directly in sinogram space*, 47 (2002), pp. 1267–1284.
[53] F. Lucka, N. Huynh, M. Betcke, E. Zhang, P. Beard, B. Cox, and S. Arridge, *Enhancing compressed sensing 4D photoacoustic tomography by simultaneous motion estimation*, SIAM J. Imaging Sci., 11 (2018), pp. 2224–2253, https://doi.org/10.1137/18M1170066.
[54] P. Ludewig, M. Gräser, N. D. Forkert, F. Thieben, J. Rández-Garbayo, J. Rieckhoff, K. Lessmann, F. Förger, P. Szwargulski, T. Magnus, and T. Knopp, *Magnetic particle imaging for assessment of cerebral perfusion and ischemia*, Wiley Interdiscip. Rev.: Nanomed. Nanobiotechnol., 14 (2022), e1757.

[55] D. Manke, K. Nehrke, and P. Börnert, *Novel prospective respiratory motion correction approach for free-breathing coronary MR angiography using a patient-adapted affine motion model*, Magn. Reson. Med., 50 (2003), pp. 122–131.
[56] E. Meinhardt-Llopis, J. S. Pérez, and D. Kondermann, *Horn-Schunck optical flow with a multi-scale strategy*, Image Process. Online, 2013 (2013), pp. 151–172.
[57] M. Möddel, A. Schlömerkemper, T. Knopp, and T. Kluth, *Limitations of current MPI models in the context of fluid dynamics*, Int. J. Magn. Part. Imag., 9 (2023), 2303078.
[58] K. Murase, M. Aoki, N. Banura, K. Nishimoto, A. Mimura, T. Kuboyabu, and I. Yabata, *Usefulness of magnetic particle imaging for predicting the therapeutic effect of magnetic hyperthermia*, Open J. Med. Imag., 5 (2015), 85.
[59] T. T. N. Nguyen, *Landweber-Kaczmarz for parameter identification in time-dependent inverse problems: All-at-once versus reduced version*, Inverse Problems, 35 (2019), 035009.
[60] R. Otazo, E. Candès, and D. Sodickson, *Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components*, Magn. Reson. Med., 73 (2015), pp. 1125–1136.
[61] T. Pock, D. Cremers, H. Bischof, and A. Chambolle, *An algorithm for minimizing the Mumford-Shah functional*, in 2009 IEEE 12th International Conference on Computer Vision, IEEE, Piscataway, NJ, 2009, pp. 1133–1140.
[62] M. Reyes, G. Malandain, P. Koulibaly, M. González-Ballester, and J. Darcourt, *Model-based respiratory motion compensation for emission tomography image reconstruction*, Phys. Med. Biol., 52 (2007), pp. 3579–3600.
[63] J. Salamon, M. Hofmann, C. Jung, M. G. Kaul, F. Werner, K. Them, R. Reimer, P. Nielsen, A. vom Scheidt, G. Adam, T. Knopp, and H. Ittrich, *Magnetic particle/magnetic resonance imaging: In-vitro MPI-guided real time catheter tracking and 4D angioplasty using a road map and blood pool tracer approach*, PLOS one, 11 (2016), e0156899.
[64] U. Schmitt and A. Louis, *Efficient algorithms for the regularization of dynamic inverse problems:* I. *Theory*, Inverse Problems, 18 (2002), pp. 645–658.
[65] U. Schmitt, A. Louis, C. Wolters, and M. Vauhkonen, *Efficient algorithms for the regularization of dynamic inverse problems:* II. *Applications*, Inverse Problems, 18 (2002), pp. 659–676.
[66] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski, *Regularization Methods in Banach Spaces*, Radon Ser. Comput. Appl. Math. 10, de Gruyter, Berlin, 2012.
[67] M. Storath, C. Brandt, M. Hofmann, T. Knopp, J. Salamon, A. Weber, and A. Weinmann, *Edge preserving and noise reducing reconstruction for magnetic particle imaging*, IEEE Trans. Med. Imaging, 36 (2016), pp. 74–85.
[68] A. Sträter, A. Huber, J. Rudolph, M. Berndt, M. Rasper, E. J. Rummeny, and J. Nadjiri, *4d-flow MRI: Technique and applications*, in RöFo-Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren, Vol. 190, Georg Thieme, Stuttgart, 2018, pp. 1025–1035.
[69] P. Szwargulski and T. Knopp, *Influence of the receive channel number on the spatial resolution in magnetic particle imaging*, Int. J. Magn. Part. Imag., 3 (2017).
[70] N. G. T. Knopp and M. Möddel, *Magnetic particle imaging: From proof of principle to preclinical applications*, Phys. Med. Biol., 62 (2017), R124, https://doi.org/10.1088/1361-6560/aa6c99.
[71] K. Them, M. G. Kaul, C. Jung, M. Hofmann, T. Mummert, F. Werner, and T. Knopp, *Sensitivity enhancement in magnetic particle imaging by background subtraction*, IEEE Trans. Med. Imaging, 35 (2015), pp. 893–900.
[72] J. Weizenecker, B. Gleich, and J. Borgert, *Magnetic particle imaging using a field free line*, J. Phys. D, 41 (2008), 105009.
[73] J. Weizenecker, B. Gleich, J. Rahmer, H. Dahnke, and J. Borgert, *Three-dimensional real-time in vivo magnetic particle imaging*, Phys. Med. Biol., 54 (2009), L1.
[74] E. Y. Yu, M. Bishop, B. Zheng, R. M. Ferguson, A. P. Khandhar, S. J. Kemp, K. M. Krishnan, P. W. Goodwill, and S. M. Conolly, *Magnetic particle imaging: A novel in vivo imaging platform for cancer detection*, Nano Lett., 17 (2017), pp. 1648–1654, https://doi.org/10.1021/acs.nanolett.6b04865.
[75] L. Zdun and C. Brandt, *Fast MPI reconstruction with non-smooth priors by stochastic optimization and data-driven splitting*, Phys. Med. Biol., 66 (2021), 175004.