

AN OPTIMAL TRANSPORT APPROACH FOR SOLVING DYNAMIC INVERSE PROBLEMS IN SPACES OF MEASURES

KRISTIAN BREDIES AND SILVIO FANZON*

Abstract. In this paper we propose and study a novel optimal transport based regularization of linear dynamic inverse problems. The considered inverse problems aim at recovering a measure valued curve and are dynamic in the sense that (i) the measured data takes values in a time dependent family of Hilbert spaces, and (ii) the forward operators are time dependent and map, for each time, Radon measures into the corresponding data space. The variational regularization we propose is based on dynamic (un-)balanced optimal transport which means that the measure valued curves to recover (i) satisfy the continuity equation, *i.e.*, the Radon measure at time t is advected by a velocity field v and varies with a growth rate g , and (ii) are penalized with the kinetic energy induced by v and a growth energy induced by g . We establish a functional-analytic framework for these regularized inverse problems, prove that minimizers exist and are unique in some cases, and study regularization properties. This framework is applied to dynamic image reconstruction in undersampled magnetic resonance imaging (MRI), modelling relevant examples of time varying acquisition strategies, as well as patient motion and presence of contrast agents.

Mathematics Subject Classification. 65J20, 49J20, 35F05, 46G12, 92C55.

Received April 5, 2019. Accepted July 31, 2020.

1. INTRODUCTION

In this paper we are concerned with solving ill-posed dynamic inverse problems where the sought unknown is a curve of Radon measures. We propose to regularize such problems *via* balanced and unbalanced dynamic optimal transport and establish a functional-analytic framework that takes the specificities of dynamic inverse problems, such as the time-varying nature of the measurement process, into account. Well-posedness as well as regularization properties are proven, and the application to magnetic resonance imaging (MRI) is discussed.

Our motivation to consider dedicated strategies for dynamic inverse problems arises from the shortcomings of static reconstruction strategies for inverse problems in the presence of motion during the measurement process. In the static case, measurement data is usually continuously collected such that sufficiently many data is available to enable the unique solution of the underlying inverse problem. In this context, one has to assume that no dynamics occur during the measurement process. However, this assumption is often violated for many applications, including medical imaging techniques such as MRI and computed tomography (CT) that image,

Keywords and phrases. Dynamic inverse problems, optimal transport regularization, continuity equation, time dependent Bochner spaces, dynamic image reconstruction, dynamic MRI.

University of Graz, Institute of Mathematics and Scientific Computing, Heinrichstraße 36, 8010 Graz, Austria.

*Corresponding author: silvio.fanzon@uni-graz.at

e.g., the beating heart or the lung while breathing. Consequently, a consistent reconstruction is no longer possible and static approaches usually admit motion artifacts. A strategy to overcome this is to temporally resolve the dynamics, meaning that for each time instance during the measurement, one seeks to reconstruct a solution where only a small fraction of the necessary data is available. In addition to that, generally, in each time instance, a different part of the data set is acquired. This results in a dynamic inverse problem with time-variant forward operators and data spaces, where for each fixed time instance, the corresponding inverse problem is massively underdetermined. In order to solve such a challenging problem, both an appropriate dynamic regularization strategy as well as a suitable modelling of forward operators and data spaces is necessary.

We propose and study a regularization strategy that bases on optimal transport energies, both in a balanced and unbalanced context, see below for a detailed description. Such strategies are naturally linked with curves of Radon measures, inverse problems in the space of Radon measures and appropriate Radon-norm-based regularizers. Indeed, the fact that for each point in time, an inverse problem with underdetermined data has to be solved calls for dedicated regularization such as the intensively studied sparsity-promoting ℓ_1 -type penalties. In the discrete setting, this leads to the celebrated theory of compressed-sensing [18, 28], in which one is able to reconstruct the unknown starting from very few random measurements, yielding better stability properties. The continuous, infinite-dimensional counterpart is given by the space of Radon measures [16, 17], where the regularization can be achieved by penalizing the Radon norm and formulating the inverse problem in measure space. Recovering the unknown from very few observations is then possible since the data admits redundancies, particularly in applications to medical imaging.

In the dynamic setting, data redundancy additionally needs to be exploited by taking time correlation into account. Indeed, one can expect displacements between consecutive time samples to be small, and incorporate this information in the regularizer in order to achieve better reconstruction. In particular, the fluid-mechanics formulation of both balanced and unbalanced optimal transport, known as Wasserstein [6, 8] and Wasserstein–Fisher–Rao distance [21, 42, 44], respectively, are particularly well-suited to keep track of motion and possible mass change occurring in the ground truth. Further, this formulation is based on curves of Radon measures and is thus attractive for the regularization of dynamic inverse problems, where in each time instance, a Radon measure should be recovered. The regularization is then enforced by subjecting potential solutions of the dynamic problem, *i.e.*, curves of Radon measures, to the continuity equation, possibly with source, while at the same time penalizing displacement field and growth rate. As we will see, this approach indeed establishes a convex regularization strategy that is sparsity-promoting in each time frame, exploits data redundancy in time and intrinsically recovers the velocity field associated with the motion as well as the rate of brightness changes. Let us emphasize that in particular, the approach allows for continuous measurement in time while providing spatial Radon-norm regularization. In contrast, a straightforward generalization of [16] to the space-time cylinder would, *e.g.*, allow for measures that are singular in time and hence, not regular enough to consistently define global-in-time forward operators and data discrepancies.

Another aspect that has to be considered for dynamic inverse problems is a faithful modelling of the measurement process with respect to time, taking into account the time-varying nature of the measurements. In this paper, the latter is achieved by the construction of ad hoc Bochner-type spaces in which the data can take values in a time-dependent family of Hilbert spaces, which are correlated in time in a very weak way. This enables us to model the dynamic inverse problem with a time-dependent family of linear forward operators, mapping Radon measures to the associated data Hilbert space in each time instance, by a global-in-time forward operator that takes curves of Radon measures to the time-varying measurement data, making it thus possible to consistently define a data-mismatch term. In this respect, our model is truly dynamic and well-adapted to undersampled data as outlined above.

The overall approach then realizes a reconstruction by inverting the global forward operator subject to optimal transport penalization and continuity equation. Such an approach leads to a well-posed and convex variational problem of Tikhonov type, in which we are able to reconstruct the sought solution, along with the displacement field and mass growth rate. The main task of the paper is to establish a rigorous functional-analytic framework in which to set the problem and to obtain well-posedness and stability properties for the proposed variational

optimal transport-based regularization. We then apply our theoretical results to dynamic medical imaging, focusing on the case of undersampled MRI, showing that we are able to treat an almost arbitrary variety of sampling strategies, as well as being principally able to reconstruct the image sequence, recover the motion displacement and track the possible presence of contrast agents.

Let us shortly review the existing literature on dynamic inverse problems and optimal transport approaches for inverse problems, image processing as well as computer vision. While the theory of regularization of dynamic inverse problems is a relatively new field of research [61], regularization theory for static ill-posed inverse problems dates back several decades. In this context, the approach in this paper can be classified as Tikhonov regularization in Banach space, a well-established technique where one penalizes the data mismatch by a convex regularizer and solves the corresponding variational problem [30, 62–64]. For computer vision and image processing applications, research in the last decades focused on specific convex regularizers [57] such as, for instance, edge-preserving functionals (total variation [19, 55], total generalized variation [13, 14]), or functionals that enforce sparsity with respect to a given basis, frame or learned dictionary [1, 18, 25, 35, 65]. In this context, Radon-norm penalties can also be interpreted as sparsity-promoting regularization [16, 17].

Concerning optimal transport, the classical theory deals with the problem of transporting mass from a probability distribution into a target one, while minimizing, *e.g.*, the average squared displacement. Such a minimization problem defines a metric over the space of probability measures, called Wasserstein distance [6, 22, 56, 66]. In [8], the authors showed that the Wasserstein distance can be computed *via* solving a convex dynamic problem that corresponds to finding a geodesic path in the space of probabilities subjected to the continuity equation that minimizes the kinetic energy. This formulation is the basis for the balanced optimal-transport regularization studied in this paper. Such approach, however, intrinsically assumes mass constancy, which is not always desired in applications, *e.g.*, in mathematical imaging. In the recent years, several ways to overcome this limitation were proposed, leading to so-called unbalanced optimal transport [7, 32, 33, 36, 46, 47, 51, 52]. In this context, common strategies are to add a source term to the continuity equation and consequently, to the kinetic energy, or to allow mass to escape/enter the domain, by interpreting the boundary as an infinite reservoir. In this paper, the energy introduced independently in [21, 42, 44], known as Wasserstein–Fisher–Rao or Hellinger–Kantorovich, is used to provide a unbalanced optimal-transport regularization. The remarkable feature of such formulation is that geodesics have a clear meaning, as they can be interpreted as joint displacement and change of mass, and therefore capture the dynamics of, *e.g.*, image sequences.

Returning to inverse problems, as already mentioned, research in the dynamic framework recently gained some momentum [61], where convex regularizers that penalize the time derivative, interpret the space-time cylinder as a higher-dimensional set or enforce a spatio-temporal decomposition of low rank have been studied in the literature, most prominently in the context of medical imaging applications [24, 37, 45, 49, 58, 59, 67]. In comparison, such approaches, however, do only implicitly account for motion information in contrast to the proposed optimal-transport regularizer which explicitly yields a motion field. In this respect, the employment of optimal transport energies as regularizers for inverse problems is a very recent development. Here, existing literature mainly focuses on static inverse problems and static optimal transport leading, for instance, to Wasserstein-distance type regularization [39, 48]. In contrast, dynamic optimal transport has been utilized for specific image processing and computer vision tasks such as image interpolation [21, 38, 50]. We also mention the work [60], which appeared after the present work. In [60], the authors propose to regularize an inverse problem related to PET image reconstruction through balanced optimal transport, subsequently applying it to the problem of tracking radiolabelled cells. The regularizer they propose is similar to ours, however the forward operator they consider is static and application-specific, whereas we are able to deal with general dynamic inverse problems. Moreover, their analytical framework is greatly simplified, dealing only with discretized unknowns in space-time satisfying a discrete version of the continuity equation, rather than with actual curves of measures, which is the natural framework to obtain well-posedness, as proposed in the present paper. To the best knowledge of the authors, no other works employ dynamic optimal transport regularization for dynamic inverse problems. In particular, a framework for recovering curves of Radon measures from continuously acquired measurements does not exist to date. Let us also mention that the realization of the time-dependent Bochner spaces introduced in

this paper is new. Indeed, existing approaches usually assume that almost every data space is isomorphic, which is sufficient to model, *i.e.*, function spaces over time-varying domains [4, 23]. Such isomorphy is not required in our approach which can thus be used to model very general data acquisition strategies.

The paper is organized as follows. In the remainder of this section, we precise the mathematical setting employed for regularizing dynamic inverse problems and summarize the main theoretical results obtained (Sect. 1.1), including details on the MRI application (Sect. 1.2). In Section 2 we lay the theoretical foundations to rigorously define the optimal transport regularizer. In Section 3 we introduce and study the above mentioned class of time dependent Bochner spaces, which will be used to model the data measurements. After this preliminary part, in Section 4, we introduce the Tikhonov regularization for the dynamic inverse problem and show well-posedness as well as regularization properties. In Section 5 we apply our theoretical results to dynamic MRI, also providing examples of sampling strategies. Finally, Section 6 concludes with some perspectives for future research and some comments on the related paper [12] as well as forthcoming work, in which we perform numerical analysis for the model proposed in this paper.

1.1. Outline of the mathematical setting and main theoretical results

Let $\Omega \subset \mathbb{R}^d$ be an open and bounded domain, with $d \in \mathbb{N}$, $d \geq 1$, and consider a time variable $t \in [0, 1]$. Let H_t be a time-dependent collection of Hilbert spaces modelling the data. The time regularity required for such family will be very mild, as specified below. At each time instance t corresponds a given linear continuous forward operator K_t^* , mapping from the space of Radon measures $\mathcal{M}(\overline{\Omega})$ into H_t . We consider the following inverse problem: Given some data $f_t \in H_t$ for $t \in [0, 1]$, find a curve of Radon measures $t \in [0, 1] \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ such that

$$K_t^* \rho_t = f_t \quad \text{for a.e. } t \in [0, 1]. \quad (1.1)$$

We propose to regularize (1.1) by means of balanced/unbalanced optimal transport. This is enforced by subjecting ρ_t to the continuity equation

$$\partial_t \rho_t + \operatorname{div}(v_t \rho_t) = g_t \rho_t \quad \text{in } (0, 1) \times \overline{\Omega}, \quad (1.2)$$

where $v_t(x): (0, 1) \times \overline{\Omega} \rightarrow \mathbb{R}^d$ is a flow field transporting the mass ρ_t , while $g_t(x): (0, 1) \times \overline{\Omega} \rightarrow \mathbb{R}$ is a growth rate keeping track of mass creation and destruction, thus allowing for local mass change. We point out that no initial conditions are prescribed on ρ_t in (1.2), since in the context of the inverse problem (1.1) we only have available indirect measurements on the whole time interval $[0, 1]$. We propose to regularize (1.1) by minimizing the Tikhonov functional

$$\min_{\rho_t, v_t, g_t} \frac{1}{2} \int_0^1 \|K_t^* \rho_t - f_t\|_{H_t}^2 dt + \frac{\alpha}{2} \int_0^1 \int_{\overline{\Omega}} |v_t(x)|^2 + \delta^2 |g_t(x)|^2 d\rho_t(x) dt + \beta \int_0^1 \rho_t(\overline{\Omega}) dt, \quad (1.3)$$

subject to (1.2). Here, $\alpha, \beta > 0$ are regularization parameters, $\delta \in (0, \infty]$ is a penalty parameter and the optimization is done for the triple (ρ_t, v_t, g_t) . The second term in (1.3) is known in the literature as Wasserstein–Fisher–Rao energy for unbalanced optimal transport [21, 42, 44], and as Benamou–Brenier energy [8] for balanced optimal transport when $\delta = \infty$, enforcing $g_t = 0$ and hence mass preservation.

Our main task is to establish problem (1.3) subject to (1.2) as a regularizer for (1.1) in a rigorous functional-analytic framework. In the following we provide some details on how to make the terms appearing in (1.3) rigorous, in particular providing suitable assumptions on K_t^* and H_t . The natural space in which to cast (1.3) is given by $\mathcal{M} := \mathcal{M}(X) \times \mathcal{M}(X; \mathbb{R}^d) \times \mathcal{M}(X)$ where $X := (0, 1) \times \overline{\Omega}$. For $(\rho, m, \mu) \in \mathcal{M}$ define the transport energy as the 1-homogeneous convex functional

$$B_\delta(\rho, m, \mu) := \int_X \Psi_\delta \left(\frac{d\rho}{d\lambda}, \frac{dm}{d\lambda}, \frac{d\mu}{d\lambda} \right) d\lambda, \quad (1.4)$$

where $\lambda \in \mathcal{M}^+(X)$ is any positive measure such that $\rho, m, \mu \ll \lambda$ and for $(t, x, y) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$ we define $\Psi_\delta(t, x, y) := \frac{|x|^2 + \delta^2 y^2}{2t}$ if $t > 0$, $\Psi_\delta(0, 0, 0) := 0$ and $\Psi_\delta(t, x, y) := \infty$ in all other cases. Introduce the affine set

$\mathcal{D} := \{(\rho, m, \mu) \in \mathcal{M} : \partial_t \rho + \operatorname{div} m = \mu\}$ where the continuity equation is in the distributional sense, without initial conditions (Def. 2.1). Whenever $(\rho, m, \mu) \in \mathcal{D}$ and $B_\delta(\rho, m, \mu) < \infty$, it follows that $\rho \geq 0$, $m = v\rho$ and $\mu = g\rho$. Moreover $\rho = dt \otimes \rho_t$ with $t \mapsto \rho_t \in \mathcal{M}(\bar{\Omega})$ narrowly continuous, i.e., $t \mapsto \int_{\bar{\Omega}} \varphi(x) d\rho_t(x)$ is continuous for all $\varphi \in C(\bar{\Omega})$ (Prop. 2.4). By setting $\lambda = \rho$ in (1.4) we recover the second term in (1.3) (Prop. 2.6). Next, we outline how we define the space of measurements. Assume given a family of real Hilbert spaces $\{H_t\}_t$ for $t \in [0, 1]$, with inner products denoted by $\langle \cdot, \cdot \rangle_{H_t}$, satisfying the following.

Assumption 1.1. *There exist a Banach space D and linear continuous operators $i_t : D \rightarrow H_t$ with the properties:*

- (H1) $\|i_t\| \leq C$ for some constant $C > 0$ not depending on t ,
- (H2) $i_t(D)$ is dense in H_t ,
- (H3) the map $t \in [0, 1] \mapsto \langle i_t \varphi, i_t \psi \rangle_{H_t} \in \mathbb{R}$ is Lebesgue measurable for every fixed $\varphi, \psi \in D$.

In other words, we assume that each H_t possesses a dense subset $i_t(D)$, and such subsets are related by the time-measurability condition (H3). In particular, Assumption 1.1 allows us to define suitable notions of strong measurability and integrability for measurements $f : [0, 1] \rightarrow H$ for $H := \cup_{t \in [0, 1]} H_t$ such that $f_t \in H_t$ for a.e. t in $[0, 1]$ (see Defs. 3.2 and 3.8), leading to the definition of the measurements space

$$L^2([0, 1]; H) := \left\{ f : [0, 1] \rightarrow H : f \text{ strongly measurable, } \int_0^1 \|f_t\|_{H_t}^2 dt < \infty \right\}. \tag{1.5}$$

In Theorem 3.13 we show that (1.5) is a Hilbert space with $\langle f, g \rangle_{L^2} := \int_0^1 \langle f_t, g_t \rangle_{H_t} dt$. Notice that our construction provides a natural extension of the classic Bochner theory to the case of varying codomains, in the sense that (1.5) coincides with the classical Bochner space when $H_t = H$ for all t , with H given Hilbert space. Details about the above construction are contained in Section 3. We now address the assumptions we make on the forward operators K_t^* appearing in (1.1).

Assumption 1.2. *For a.e. $t \in [0, 1]$ the linear continuous operators $K_t^* : \mathcal{M}(\bar{\Omega}) \rightarrow H_t$ satisfy:*

- (K1) K_t^* is the adjoint of a linear continuous operator $K_t : H_t \rightarrow C(\bar{\Omega})$,
- (K2) $\|K_t\| \leq C$ for some constant $C > 0$ not depending on t ,
- (K3) the map $t \in [0, 1] \mapsto K_t^* \rho \in H_t$ is strongly measurable for every fixed $\rho \in \mathcal{M}(\bar{\Omega})$.

Under (K1)–(K3), (H1)–(H3) we have the following: if $t \mapsto \rho_t \in \mathcal{M}(\bar{\Omega})$ is narrowly continuous then the map $t \mapsto K_t^* \rho_t$ belongs to $L^2([0, 1]; H)$ (Lem. 4.2). At this point, we are ready to rigorously define the regularization functional anticipated in (1.3) as $J : \mathcal{M} \rightarrow [0, \infty]$, where

$$J(\rho, m, \mu) := \frac{1}{2} \|K_t^* \rho_t - f_t\|_{L^2}^2 + \alpha B_\delta(\rho, m, \mu) + \beta \|\rho\|_{\mathcal{M}(X)}, \tag{1.6}$$

if $(\rho, m, \mu) \in \mathcal{D}$ and $J := \infty$ otherwise. The discrepancy term in J is well defined since, if $(\rho, m, \mu) \in \mathcal{D}$ and $B_\delta(\rho, m, \mu) < \infty$, then $\rho = dt \otimes \rho_t$ with $t \mapsto \rho_t$ narrowly continuous, so that $t \mapsto K_t^* \rho_t$ belongs to $L^2([0, 1]; H)$ (Prop. 4.3). Notice that, in addition to the regularizer $B_\delta(\rho, m, \mu)$, we also included $\|\rho\|_{\mathcal{M}(X)}$ in the definition of J : This serves the purpose of enforcing weak* coercivity on J , since no initial data on ρ is prescribed. Our main theoretical results concerning existence of minimizers for J and regularization properties are summarized in the following statements, which are contained in Theorems 4.4, 4.7 and 4.10, respectively.

Theorem 1.3. *Assume (H1)–(H3), (K1)–(K3). Let $f \in L^2([0, 1]; H)$, $\alpha, \beta > 0$. Then J admits a minimizer $(\rho, m, \mu) \in \mathcal{D}$ satisfying $\rho \geq 0$, $\rho = dt \otimes \rho_t$ with $t \mapsto \rho_t$ narrowly continuous. If in addition the operators K_t^* are injective for a.e. $t \in [0, 1]$, then the minimizer is unique.*

In the next theorem, $J_{\alpha, \beta, f}$ denotes the functional J in (1.6) for $\alpha, \beta > 0$ and $f \in L^2([0, 1]; H)$.

Theorem 1.4 (Regularization). *Assume (H1)–(H3), (K1)–(K3). Let $f^\gamma, f^\dagger \in L^2([0, 1]; H)$ be noisy and exact data respectively, for noise level $\gamma > 0$.*

- (i) *Suppose that $f^n \rightarrow f^\gamma$ strongly in L^2 , $\alpha, \beta > 0$ and $(\rho^n, m^n, \mu^n) \in \arg \min J_{\alpha, \beta, f^n}$. Then, up to subsequences, (ρ^n, m^n, μ^n) converges weakly* to $(\rho, m, \mu) \in \arg \min J_{\alpha, \beta, f^\gamma}$.*
- (ii) *Assume that $\|f^{\gamma_n} - f^\dagger\|_{L^2} \leq \gamma_n$ and $\alpha_n, \beta_n \searrow 0$, such that $\gamma_n^2 / \min\{\alpha_n, \beta_n\} \rightarrow 0$. If $(\rho^n, m^n, \mu^n) \in \arg \min J_{\alpha_n, \beta_n, f^{\gamma_n}}$ then, up to subsequences, (ρ^n, m^n, μ^n) converges weakly* to $(\rho^\dagger, m^\dagger, \mu^\dagger) \in \mathcal{D}$ solving (1.1) and there exist $\alpha^*, \beta^* \in [1, \infty]$ such that*

$$(\rho^\dagger, m^\dagger, \mu^\dagger) \in \arg \min \alpha^* B_\delta(\rho, m, \mu) + \beta^* \|\rho\|_{\mathcal{M}(X)}.$$

1.2. Application to dynamic MRI

We apply the model (1.1) and its regularization (1.3) to undersampled dynamic MRI, yielding a reconstruction approach *via* convex optimization which is principally capable of capturing motion during the acquisition. A common limiting factor to medical imaging techniques and MRI in particular is acquisition speed such that, for instance, data cannot be collected sufficiently fast in order to temporally resolve the beating heart or the lung while breathing. Consequently, static reconstruction approaches lead to severe artifacts. Thus, motion has to be taken into account by considering the dynamic setting in which at each time instance, data is severely undersampled and temporal data redundancies have to be exploited. For this purpose, we show that the optimal-transport regularization framework developed in this paper can be applied, leading to a regularizer that penalizes the displacements caused by motion and intrinsically recovers the motion field as well as the growth rate.

The forward problem in undersampled dynamic MRI in two dimensions is commonly stated as follows: In each time instance t , the proton density ρ_t , a non-negative quantity, needs to be recovered from the measured data f_t . Taking coil sensitivities into account, ρ_t and f_t are linked *via* the Fourier transform. However, for each t , the Fourier data is only acquired on subsets specified by the sampling strategy, leading to each f_t generally living on a different subset of the so-called k -space and hence, being contained in a time-varying data Hilbert space H_t . Modelling the proton density ρ_t as a positive measure on the image domain $\Omega \subset \mathbb{R}^2$, denoting by K_t^* an appropriately masked Fourier transform and considering the unit time interval $[0, 1]$, the forward problem then indeed reads as $K_t^* \rho_t = f_t$ in H_t for $t \in [0, 1]$. This is made precise in the following.

Adopting the common model for parallel data acquisition (see, *e.g.*, [40, 41, 54, 58]), let $\Omega \subset \mathbb{R}^2$ be an open bounded domain representing the image domain and let the complex coil sensitivities $c_j \in C_0(\mathbb{R}^2; \mathbb{C})$ for $j = 1, \dots, N$ with $N \geq 1$ to each of the N receiver coils be given. The time-dependent sampling method is represented by a family of measures $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ for $t \in [0, 1]$. Such measures are required to satisfy some mild regularity assumptions, namely,

(M1) $\|\sigma_t\|_{\mathcal{M}(\mathbb{R}^2)} \leq C$ for a.e. $t \in [0, 1]$, where $C > 0$ does not depend on t ,

(M2) the map $t \mapsto \int_{\mathbb{R}^2} \varphi(x) d\sigma_t(x)$ is measurable for each $\varphi \in C_0(\mathbb{R}^2; \mathbb{C})$,

allowing for a variety of sampling methods, see Section 5 for details. The data space of measurements is then defined by $H_t := L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C}^N)$, interpreted as a real Hilbert space and equipped with the norm $\|h\|_{H_t}^2 := \sum_{j=1}^N \int_{\mathbb{R}^2} |h^j(x)|^2 d\sigma_t(x)$, where we denote $h = (h^1, \dots, h^N)$. The forward operators are given by $K_t^*: \mathcal{M}(\Omega) \rightarrow H_t$ defined *via*

$$K_t^* \rho_t := (\mathcal{F}(c_1 \rho_t), \dots, \mathcal{F}(c_N \rho_t)),$$

where \mathcal{F} is the Fourier transform and we interpret each $\mathcal{F}(c_j \rho_t)$ as an element of $L^2_{\sigma_t}(\mathbb{R}^2, \mathbb{C})$. In Lemma 5.4 we show that under (M1) and (M2), the spaces H_t and the forward operators K_t^* fulfill (H1)–(H3) and (K1)–(K3), respectively. In this way the hypotheses of Theorems 1.3 and 1.4 are satisfied, and we can regularize the reconstruction problem (1.1) with the functional $J: \mathcal{M} \rightarrow [0, \infty]$ defined in (1.6), which in this framework corresponds to

$$J(\rho, m, \mu) := \frac{1}{2} \sum_{j=1}^N \int_0^1 \left\| \mathcal{F}(c_j \rho_t) - f_t^j \right\|_{L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C})}^2 dt + \alpha B_\delta(\rho, m, \mu) + \beta \|\rho\|_{\mathcal{M}((0,1) \times \bar{\Omega})},$$

where the measurements $f_t = (f_t^1, \dots, f_t^N)$ belong to $L^2([0, 1]; H)$. This shows in particular that optimal-transport regularization for undersampled dynamic MRI leads to well-posed convex optimization problems. These are accessible to analysis as well as efficient and stable numerical minimization algorithms. Section 6 provides some perspectives for the latter.

2. DYNAMIC OPTIMAL TRANSPORT

The aim of this section is to provide the essential elements to define the optimal transport regularizer appearing in (1.6). We refer the reader to Appendix A.1 for measure theory definitions and results which will be needed in the following. Throughout the section, $\Omega \subset \mathbb{R}^d$ is an open bounded domain, with $d \in \mathbb{N}$, $d \geq 1$ and we set $X := (0, 1) \times \overline{\Omega}$ to be the time-space cylinder. We also define the space $\mathcal{M} := \mathcal{M}(X) \times \mathcal{M}(X; \mathbb{R}^d) \times \mathcal{M}(X)$. In Section 2.1 we introduce the concept of measure solution to the continuity equation with source

$$\partial_t \rho + \operatorname{div} m = \mu \quad \text{in } X, \tag{2.1}$$

where $(\rho, m, \mu) \in \mathcal{M}$. Here ρ represents a density, m a momentum field advecting ρ and μ a source term, accounting for local mass change. We then investigate properties of solutions $\rho \in \mathcal{M}(X)$ of (2.1). In particular in Proposition 2.2 we show that positive solutions to (2.1) disintegrate as $\rho = dt \otimes \rho_t$ with $\{\rho_t\}_{t \in [0,1]}$ Borel family of positive measures over $\overline{\Omega}$. In Proposition 2.4 we prove that, under some growth assumptions on m and μ , the curve $t \mapsto \rho_t$ is actually narrowly continuous. Finally, in Section 2.2 we introduce the optimal transport energy B_δ at (1.6), and list some of its properties in Proposition 2.6.

2.1. Continuity equation

We want to consider measure valued (distributional) solutions to the continuity equation (2.1) with suitable boundary conditions. The precise definition is as follows.

Definition 2.1. We say that $(\rho, m, \mu) \in \mathcal{M}$ is a *measure solution* to (2.1) if

$$\int_X \partial_t \varphi \, d\rho + \int_X \nabla \varphi \cdot dm + \int_X \varphi \, d\mu = 0 \quad \text{for all } \varphi \in C_c^\infty(X). \tag{2.2}$$

We remark that the above weak formulation includes zero flux boundary conditions for the momentum m on $\partial\Omega$, and no initial and final data is prescribed on ρ . Moreover one can test (2.2) against functions in $C_c^1(X)$ (see [6], Rem. 8.1.1). In the following proposition we show that positive solutions to (2.1) can be disintegrated with respect to the Lebesgue measure dt on $(0, 1)$ (see Sect. A.1.4 for details on disintegration). To this end, let $\pi: X \rightarrow (0, 1)$ be the projection on the time coordinate.

Proposition 2.2. *Assume that $(\rho, m, \mu) \in \mathcal{M}$ satisfies (2.2), with $\rho \in \mathcal{M}^+(X)$. Then ρ disintegrates, with respect to dt , as $\rho = dt \otimes \rho_t$, where $\rho_t \in \mathcal{M}^+(\overline{\Omega})$ for a.e. t . Moreover $t \mapsto \rho_t(\overline{\Omega})$ is a function of bounded variation, with distributional derivative $\pi_{\#}\mu$. In particular, if the source $\mu = 0$, then the total mass $\rho_t(\overline{\Omega})$ is constant in time.*

Proof. In order to apply Theorem A.1 we need to show that $\pi_{\#}\rho \ll dt$. Let $\tilde{\varphi} \in C_c^\infty((0, 1))$ and define $\varphi := \tilde{\varphi} \circ \pi \in C_c^\infty(X)$. By plugging φ in (2.2) we get $\int_0^1 \tilde{\varphi}' \, d(\pi_{\#}\rho) = -\int_0^1 \tilde{\varphi} \, d(\pi_{\#}\mu)$, so that $(\pi_{\#}\rho)' = \pi_{\#}\mu$ in the sense of distributions. Since $\pi_{\#}\mu \in \mathcal{M}((0, 1))$, there exists $u \in BV((0, 1))$ such that $\pi_{\#}\mu = u'$. Therefore $\pi_{\#}\rho \ll dt$ and there exists a Borel family $\rho_t \in \mathcal{M}(\overline{\Omega})$ such that $\rho = dt \otimes \rho_t$. In particular, since ρ_t is a Borel family, the map $t \mapsto \rho_t(\overline{\Omega})$ is measurable. Moreover it belongs to $L^1((0, 1))$, since $\int_0^1 |\rho_t(\overline{\Omega})| \, dt = \rho(X)$ which is finite by assumption. Finally notice that $\pi_{\#}(dt \otimes \rho_t) = \rho_t(\overline{\Omega}) \, dt$, which together with $(\pi_{\#}\rho)' = \pi_{\#}\mu$ implies that $t \mapsto \rho_t(\overline{\Omega})$ belongs to $BV((0, 1))$, with distributional derivative given by $\pi_{\#}\mu$. \square

Definition 2.3. A curve $t \in [0, 1] \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ is *narrowly continuous* if the map

$$t \mapsto \int_{\overline{\Omega}} \varphi(x) \, d\rho_t(x)$$

is continuous for every fixed $\varphi \in C(\overline{\Omega})$. We denote by $C_w([0, 1]; \mathcal{M}(\overline{\Omega}))$ the set of such curves, and by $C_w([0, 1]; \mathcal{M}^+(\overline{\Omega}))$ the set of narrowly continuous curves of positive measures.

In the next proposition we show that if (ρ, m, μ) solves the continuity equation with appropriate energy bounds, the disintegration measures ρ_t are defined for every t and are narrowly continuous. This is a well-known result for $\mu = 0$. For completeness we will carry out the proof in Appendix A.2, by adapting the argument used to prove the homogeneous version (see [6], Lem. 8.1.2).

Proposition 2.4 (Continuous representative). *Let $(\rho, m, \mu) \in \mathcal{M}$ be a solution of (2.2), with $\rho \in \mathcal{M}^+(X)$. Let $\rho_t \in \mathcal{M}^+(\overline{\Omega})$ be the disintegration of ρ with respect to dt . Assume that $m = dt \otimes v_t \rho_t$ and $\mu = dt \otimes g_t \rho_t$ with $v_t: X \rightarrow \mathbb{R}^d, g_t: X \rightarrow \mathbb{R}$ measurable functions such that*

$$\int_0^1 \int_{\overline{\Omega}} |v_t(x)| \, d\rho_t(x) \, dt < \infty \quad \text{and} \quad \int_0^1 \int_{\overline{\Omega}} |g_t(x)| \, d\rho_t(x) \, dt < \infty. \tag{2.3}$$

Then there exists a narrowly continuous curve $(t \mapsto \tilde{\rho}_t) \in C_w([0, 1]; \mathcal{M}^+(\overline{\Omega}))$ such that $\rho_t = \tilde{\rho}_t$ a.e. in $(0, 1)$. Moreover for each $\varphi \in C_c^1([0, 1] \times \overline{\Omega})$ and $0 \leq t_1 \leq t_2 \leq 1$ we have

$$\int_{t_1}^{t_2} \int_{\overline{\Omega}} (\partial_t \varphi + \nabla \varphi \cdot v_t + \varphi g_t) \, d\rho_t(x) \, dt = \int_{\overline{\Omega}} \varphi(t_2, x) \, d\tilde{\rho}_{t_2}(x) - \int_{\overline{\Omega}} \varphi(t_1, x) \, d\tilde{\rho}_{t_1}(x). \tag{2.4}$$

In the rest of the paper we will identify ρ_t with its narrowly continuous representative $\tilde{\rho}_t$ whenever the assumptions of Proposition 2.4 hold, and use the notation ρ_t .

2.2. Optimal transport energy

We want to introduce the Wasserstein–Fisher–Rao energy (or Hellinger–Kantorovich) as originally done in [20, 21, 42–44]. First, define the convex set

$$K_\delta := \left\{ (a, b, c) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} : a + \frac{1}{2} \left(|b|^2 + \frac{c^2}{\delta^2} \right) \leq 0 \right\}, \tag{2.5}$$

with $\delta \in (0, \infty]$ fixed parameter and $\frac{c^2}{\infty} = 0$ for every $c \in \mathbb{R}$. For $(t, x, y) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$ set

$$\Psi_\delta(t, x, y) := \begin{cases} \frac{|x|^2 + \delta^2 y^2}{2t} & \text{if } t > 0, \\ 0 & \text{if } t = |x| = y = 0, \\ \infty & \text{otherwise,} \end{cases}$$

where $\infty y^2 = \infty$ for $y \neq 0$ and $\infty y^2 = 0$ for $y = 0$. We have that Ψ_δ is the Legendre conjugate of the characteristic function χ_{K_δ} [21], that is,

$$\Psi_\delta(t, x, y) = \sup_{(a,b,c) \in K_\delta} (at + b \cdot x + cy) \quad \text{for each } (t, x, y) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}.$$

In particular f is convex, lower semicontinuous and 1-homogeneous.

Definition 2.5 (Transport energy). Let $(\rho, m, \mu) \in \mathcal{M}$. We define the *transport energy* as

$$B_\delta(\rho, m, \mu) := \sup \left\{ \int_X a \, d\rho + \int_X b \cdot dm + \int_X c \, d\mu : (a, b, c) \in C_0(X; K_\delta) \right\}. \tag{2.6}$$

We summarize some of the properties of the functional B_δ that will be needed throughout this paper. The proof is omitted, and it can be easily adapted from the one in Proposition 5.18 of [56].

Proposition 2.6. *The functional B_δ defined in (2.6) is convex and lower semicontinuous for the weak* convergence. Moreover it satisfies the following properties:*

- (i) $B_\delta(\rho, m, \mu) \geq 0$,
- (ii) assume that $\rho, m, \mu \ll \lambda$ for some $\lambda \in \mathcal{M}^+(X)$. Then

$$B_\delta(\rho, m, \mu) = \int_X \Psi_\delta \left(\frac{d\rho}{d\lambda}, \frac{dm}{d\lambda}, \frac{d\mu}{d\lambda} \right) d\lambda,$$

- (iii) if $B_\delta(\rho, m, \mu) < \infty$ then $\rho \geq 0$ and $m, \mu \ll \rho$,
- (iv) if $\rho \geq 0$ and $m, \mu \ll \rho$, then $m = v\rho, \mu = g\rho$ for measurable $v: X \rightarrow \mathbb{R}^d, g: X \rightarrow \mathbb{R}$ and

$$B_\delta(\rho, m, \mu) = \int_X \Psi_\delta(1, v, g) d\rho = \frac{1}{2} \int_X (|v|^2 + \delta^2 g^2) d\rho.$$

3. TIME DEPENDENT BOCHNER SPACES

In this section we construct a class of Bochner spaces of Hilbert space valued functions, where the Hilbert space can vary in time. Here the underlying measure space is the unit interval with the Lebesgue measure. This can however be easily generalized to arbitrary measure spaces. Moreover a generalization to Banach spaces valued functions seems possible, however, this is out of the scope of this paper. More precisely, we want to define a concept of integrability for functions $f: [0, 1] \rightarrow H = \cup_t H_t$, where H_t is a Hilbert space for each time t , and $f(t) \in H_t$ for a.e. t . In order to do that we will closely follow the approach to define classic Bochner spaces (see [27], Chap. II, [3], Chap. 11). In Section 3.1 we establish the functional analytic setting and assumptions under which we carry out the construction. In Section 3.2 we define suitable notions of measurability and provide the equivalent of the classic Pettis measurability theorem (see Thm. 3.5). Such result is instrumental to the following analysis, as it provides a practical characterization of strong measurability. In Section 3.3 we define integrability for maps $f: [0, 1] \rightarrow H$ and characterize it in Theorem 3.9. We then proceed to define the time dependent Bochner spaces $L^p([0, 1]; H)$. Notice that, in contrast to the classic Bochner theory, we will not define a notion of integral for integrable maps $f: [0, 1] \rightarrow H$, but only of integrability. However, a comparison with the classical theory is possible, and it will be carried out in Appendix B.2.

3.1. Functional setting

Let $\{H_t\}$ for $t \in [0, 1]$ be a family of real Hilbert spaces with norms and scalar products denoted by $\|\cdot\|_{H_t}$ and $\langle \cdot, \cdot \rangle_{H_t}$, respectively. The interval $[0, 1]$ is equipped with the Lebesgue measure. As usual, we denote by $|E|$ the measure of a set $E \subset [0, 1]$ and by χ_E its characteristic function, defined as $\chi_E(t) := 1$ if $t \in E$ and $\chi_E(t) := 0$ otherwise. Let $H := \cup_{t \in [0, 1]} H_t$. We will denote by $f: [0, 1] \rightarrow H$ maps such that $f(t) \in H_t$ for a.e. $t \in [0, 1]$. Let D be a real Banach space with norm denoted by $\|\cdot\|_D$ and duality by $\langle \cdot, \cdot \rangle_{D^*, D}$. Assume that for a.e. $t \in [0, 1]$ there exists a linear continuous operator $i_t: D \rightarrow H_t$ with the following properties:

- (H1) $\|i_t\| \leq C$ for some constant $C > 0$ not depending on t ,
- (H2) $i_t(D)$ is dense in H_t ,
- (H3) the map $t \mapsto \langle i_t \varphi, i_t \psi \rangle_{H_t}$ is Lebesgue measurable for every fixed $\varphi, \psi \in D$.

The adjoint of i_t is $i_t^*: H_t \rightarrow D^*$, defined by $\langle i_t^* h, \varphi \rangle_{D^*, D} := \langle h, i_t \varphi \rangle_{H_t}$ for all $h \in H_t, \varphi \in D$ (here we identified H_t with its dual). Notice that from (H1) it follows that $i_t^*: H_t \rightarrow D^*$ is linear continuous and such that $\|i_t^*\| \leq C$. Moreover from (H2) we have that i_t^* is injective. Throughout the section, we say that $g = f$ if the equality holds a.e. in $[0, 1]$. Moreover we say that $f_n \rightarrow f$ a.e. if $\lim_n \|f_n(t) - f(t)\|_{H_t} = 0$ for a.e. $t \in [0, 1]$.

3.2. Measurability in time dependent spaces

In this section we introduce suitable measurability notions for maps $f: [0, 1] \rightarrow H$, and prove our version of Pettis' theorem. We refer the reader to Chapter II of [27] for classic measurability definitions.

Definition 3.1 (Step function). A map $f: [0, 1] \rightarrow D$ is a *step function* if $f = \sum_{j=1}^N \chi_{E_j} \varphi_j$ with $N \in \mathbb{N}$, $\varphi_j \in D$ and E_j Lebesgue measurable pairwise disjoint subsets of $[0, 1]$.

Definition 3.2 (Measurability). Let $f: [0, 1] \rightarrow H$. We say that

(i) f is *strongly measurable* if there exists a sequence of step functions $\{f_n\}$ such that

$$\lim_n \|i_t f_n(t) - f(t)\|_{H_t} = 0 \quad \text{for a.e. } t \in [0, 1],$$

(ii) f is *weakly measurable* if $t \mapsto \langle i_t \varphi, f(t) \rangle_{H_t}$ is Lebesgue measurable for each $\varphi \in D$,
 (iii) f is *essentially separably valued* if there exist a measurable set $E \subset [0, 1]$ with $|E| = 0$ and a countable subset $S \subset D$ with the following property: for every $\varepsilon > 0$ and $t \in [0, 1] \setminus E$, there exists an element $\varphi \in S$ such that

$$\|i_t \varphi - f(t)\|_{H_t} < \varepsilon.$$

Notice that, if $H_t = H$ for each $t \in [0, 1]$, with H fixed Hilbert space, $D = H$ and $i_t \varphi := \varphi$, then Definitions 3.1 and 3.2 are equivalent to the classic ones given in Chapter II of [27].

Remark 3.3. Let $p \geq 1$ and $f = \sum_{j=1}^N \chi_{E_j} \varphi_j$ be a step function. Then the map $t \mapsto \|i_t f(t)\|_{H_t}^p$ is measurable and $\int_0^1 \|i_t f(t)\|_{H_t}^p dt < \infty$. Indeed, $\|i_t f(t)\|_{H_t}^2 = \sum_{j=1}^N \chi_{E_j}(t) \langle i_t \varphi_j, i_t \varphi_j \rangle_{H_t}$, so that $t \mapsto \|i_t f(t)\|_{H_t}^2$ is measurable by (H3), and hence also $t \mapsto \|i_t f(t)\|_{H_t}^p$ is. Moreover by (H1)

$$\int_0^1 \|i_t f(t)\|_{H_t}^p dt = \sum_{j=1}^N \int_{E_j} \|i_t \varphi_j\|_{H_t}^p dt \leq C^p \sum_{j=1}^N |E_j| \|\varphi_j\|_D^p < \infty.$$

Remark 3.4. It is easy to check that strong measurability is stable under sums, scalar multiplication and pointwise a.e. convergence. Moreover if $f: [0, 1] \rightarrow H$ is strongly measurable then the map $t \mapsto \|f(t)\|_{H_t}$ is Lebesgue measurable, since f can be approximated a.e. by step functions f_n and $t \mapsto \|i_t f_n(t)\|_{H_t}$ is measurable for every fixed n by Remark 3.3.

The above definitions are linked together by the analogous of the classic Pettis measurability theorem (see [27], Chap. II.1, Thm. 2).

Theorem 3.5 (Pettis). *Let $f: [0, 1] \rightarrow H$. Then f is strongly measurable if and only if f is weakly measurable and essentially separably valued.*

For a proof of this theorem, see Appendix B.1. By inspecting the proof, one can see that the following corollary holds.

Corollary 3.6. *Let $f: [0, 1] \rightarrow H$. Then f is strongly measurable if and only if it is the a.e. uniform limit of a sequence of countably valued functions $f_n: [0, 1] \rightarrow D$, that is, if*

$$\lim_n \|i_t f_n(t) - f(t)\|_{H_t} = 0$$

uniformly for a.e. $t \in [0, 1]$, for some $f_n = \sum_{j=1}^{\infty} \chi_{E_{n,j}} \varphi_{n,j}$ with $\varphi_{n,j} \in D$ and $\{E_{n,j}\}_{j \in \mathbb{N}}$ measurable and pairwise disjoint subsets of $[0, 1]$.

Proposition 3.7 (Separable case). *Assume that D is separable. Then strong measurability is equivalent to weak measurability.*

Proof. Let $S = \{\varphi_n\} \subset D$ be countable and dense. Then $\{i_t\varphi_n\}$ is dense in H_t : indeed fix $\varepsilon > 0$ and $h \in H_t$. By (H2) there exists $\varphi \in D$ such that $\|h - i_t\varphi\|_{H_t} < \varepsilon/2$. Let φ_n be such that $\|\varphi - \varphi_n\|_D < \varepsilon/2C$ where C is the constant in (H1). Then $\|h - i_t\varphi_n\|_{H_t} < \varepsilon$ by triangle inequality and (H1). Therefore H_t is separable and it is immediate to check that in this case every function $f: [0, 1] \rightarrow H$ is essentially separably valued. By Theorem 3.5 the thesis follows. \square

3.3. Integration and L^p spaces

Definition 3.8 (Integrability). Let $f: [0, 1] \rightarrow H$ be strongly measurable according to Definition 3.2. We say that f is *integrable* if there exists a sequence $\{f_n\}$ of step functions $f_n: [0, 1] \rightarrow D$ such that

$$\lim_n \int_0^1 \|i_t f_n(t) - f(t)\|_{H_t} dt = 0. \tag{3.1}$$

Notice that the definition is well posed, since the map $t \rightarrow \|i_t f_n(t) - f(t)\|_{H_t}$ is Lebesgue measurable for each fixed n (see Rem. 3.4), hence its integral is well defined. Analogously to the classic case ([27], Chap. II.2, Thm. 2), we can characterize integrability as stated in the following theorem.

Theorem 3.9 (Characterization of integrability). *Let $f: [0, 1] \rightarrow H$ be strongly measurable. Then f is integrable if and only if $\int_0^1 \|f(t)\|_{H_t} dt < \infty$.*

Proof. Assume that f is integrable. By (3.1), for sufficiently large N we have

$$\int_0^1 \|f(t)\|_{H_t} dt \leq \int_0^1 \|i_t f_N(t) - f(t)\|_{H_t} dt + \int_0^1 \|i_t f_N(t)\|_{H_t} dt < 1 + \int_0^1 \|i_t f_N(t)\|_{H_t} dt$$

and the thesis follows by Remark 3.3, since f_N is a step function. Conversely, assume that $\int_0^1 \|f(t)\|_{H_t} dt < \infty$. Since f is strongly measurable, by Corollary 3.6 there exists a sequence $\{g_n\}$ of countably valued maps $g_n: [0, 1] \rightarrow D$ such that $\|i_t g_n(t) - f(t)\|_{H_t} < 1/n$ for a.e. $t \in [0, 1]$. In particular $\|i_t g_n(t)\|_{H_t} \leq \|f(t)\|_{H_t} + 1/n$, so that $\int_0^1 \|i_t g_n(t)\|_{H_t} dt < \infty$. By construction $g_n(t) = \sum_{j=1}^{\infty} \chi_{E_{n,j}} \varphi_{n,j}$ with sets $E_{n,j} \subset [0, 1]$ measurable and pairwise disjoint. Hence there exists a sequence $\{k_n\}$ in \mathbb{N} such that $\sum_{j=k_n+1}^{\infty} \int_{E_{n,j}} \|i_t g_n(t)\|_{H_t} dt < n^{-1}$.

Therefore by setting $f_n(t) := \sum_{j=1}^{k_n} \chi_{E_{n,j}} \varphi_{n,j}$ we obtain

$$\int_0^1 \|i_t f_n(t) - f(t)\|_{H_t} dt \leq \int_0^1 \|i_t g_n(t) - f(t)\|_{H_t} dt + \int_0^1 \|i_t f_n(t) - i_t g_n(t)\|_{H_t} dt < \frac{2}{n},$$

proving that f is integrable. \square

As a corollary of the above theorem, we obtain that a suitable version of Lebesgue’s dominated convergence theorem holds in our setting. We postpone its proof to Appendix B.1.

Theorem 3.10 (Dominated convergence). *Let $f_n: [0, 1] \rightarrow H$ be a sequence of integrable functions such that $f_n \rightarrow f$ a.e. in $[0, 1]$ and that there exists $g \in L^1([0, 1])$ satisfying $\|f_n(t)\|_{H_t} \leq g(t)$ a.e. in $[0, 1]$. Then f is integrable and $f_n \rightarrow f$ strongly in $L^1([0, 1]; H)$.*

Definition 3.11 (L^p space). Fix $1 \leq p < \infty$. We define the space of the p -integrable functions

$$L^p([0, 1]; H) := \left\{ f: [0, 1] \rightarrow H : f \text{ strongly measurable, } \int_0^1 \|f(t)\|_{H_t}^p dt < \infty \right\}. \tag{3.2}$$

In (3.2) we identify functions that coincide almost everywhere. Notice that (3.2) is well posed: Indeed since f is strongly measurable, then the map $t \mapsto \|f(t)\|_{H_t}^p$ is measurable (Rem. 3.4), and hence $\int_0^1 \|f(t)\|_{H_t}^p dt$ is well defined (possibly infinite).

Remark 3.12. If $H_t \equiv H$ with H fixed Hilbert space and $D = H$, then $L^p([0, 1]; H)$ coincides with the respective classic Bochner spaces (see [27], Chap. II).

Theorem 3.13. *Let $1 \leq p < \infty$. We have that $L^p([0, 1]; H)$ is a Banach space with the norm $\|f\|_{L^p} := \int_0^1 \|f(t)\|_{H_t}^p dt$. Moreover $L^2([0, 1]; H)$ is a Hilbert space with the inner product $\langle f, g \rangle_{L^2} := \int_0^1 \langle f(t), g(t) \rangle_{H_t} dt$.*

The proof of the above theorem is postponed to Appendix B.1.

Remark 3.14 ($p = \infty$). It is possible to treat the case $p = \infty$ by defining $L^\infty([0, 1]; H)$ as the set of strongly measurable functions $f: [0, 1] \rightarrow H$ such that $\text{ess sup}_{t \in [0, 1]} \|f(t)\|_{H_t} < \infty$. By adapting the proof for classical Bochner spaces, one can show that $L^\infty([0, 1]; H)$ is a Banach space with the norm $\|f\|_{L^\infty} := \text{ess sup}_{t \in [0, 1]} \|f(t)\|_{H_t}$.

Remark 3.15 (Dual spaces). The space $L^2([0, 1]; H)$ is self-dual, being a Hilbert space. We believe that for any $p \geq 1$ one has the isometry $L^p([0, 1]; H)^* = L^q([0, 1]; H)$ with $1/p + 1/q = 1$.

Example 3.16 (Narrowly continuous curves). Let $\Omega \subset \mathbb{R}^d$ be an open bounded domain with $d \in \mathbb{N}$, $d \geq 1$. Let $\rho_t \in C_w([0, 1]; \mathcal{M}^+(\overline{\Omega}))$, $D := C(\overline{\Omega})$ with the supremum norm, $H_t := L^2_{\rho_t}(\Omega)$ with the norm $\|h\|_{H_t}^2 := \int_{\Omega} |h(x)|^2 d\rho_t(x)$. Define $i_t: D \rightarrow H_t$ by $i_t \varphi := \varphi$. It is left to the reader to check that (H1)–(H3) are satisfied. Therefore we can define the space $L^p([0, 1]; H)$ for each $p \geq 1$. Notice that in this case $L^2([0, 1]; H)$ is isometric to $L^2_{\rho}([0, 1] \times \overline{\Omega})$, where $\rho := dt \otimes \rho_t$.

Finally, we would like to mention that, although we do not define a notion of integral for maps in $L^p([0, 1]; H)$, a comparison to the classical Bochner theory is still possible. Indeed, if $f: [0, 1] \rightarrow H$, then by definition $i_t^* f: [0, 1] \rightarrow D^*$ and the codomain is a fixed Banach space. In Appendix B.2 we show that, assuming $f \in L^1([0, 1]; H)$, we always have that $i_t^* f$ is weakly* integrable (see Prop. B.2). However, Bochner integrability fails in general, as shown in Example B.3. Nevertheless, under suitable additional assumptions, one can show that Bochner integrability can be guaranteed (Prop. B.4).

4. REGULARIZATION OF DYNAMIC INVERSE PROBLEMS

In this section we define and study the properties of the optimal transport based functional (1.6), and we establish it as a regularizer for the dynamic inverse problem (1.1). Throughout the section, the functional analytic setting will be the following. Let $\Omega \subset \mathbb{R}^d$ be an open bounded domain, where $d \in \mathbb{N}$, $d \geq 1$, and define again the time space domain $X := (0, 1) \times \overline{\Omega}$. Let $\{H_t\}$ be a family of Hilbert spaces for $t \in [0, 1]$, D a Banach space and $i_t: D \rightarrow H_t$ linear operators which satisfy the assumptions (H1)–(H3) of Section 3.1. Assume given a family of linear continuous operators $K_t^*: \mathcal{M}(\overline{\Omega}) \rightarrow H_t$ such that, for a.e. $t \in [0, 1]$,

- (K1) K_t^* is the adjoint of a linear continuous operator $K_t: H_t \rightarrow C(\overline{\Omega})$,
- (K1') K_t^* is weak*-to-weak continuous,
- (K2) $\|K_t^*\| \leq C$ for some constant $C > 0$ that does not depend on t ,
- (K3) the map $t \mapsto K_t^* \rho$ is strongly measurable in the sense of Definition 3.2, for all $\rho \in \mathcal{M}(\overline{\Omega})$.

We remark that conditions (K1) and (K1') are equivalent. As before, the space of narrowly continuous curves with values in the measures and in the positive measures are denoted by $C_w([0, 1]; \mathcal{M}(\overline{\Omega}))$ and $C_w([0, 1]; \mathcal{M}^+(\overline{\Omega}))$ respectively. The dynamic inverse problem we aim to regularize is the following: Given some data $f \in L^2([0, 1]; H)$, find a narrowly continuous curve $t \mapsto \rho_t$ such that

$$K_t^* \rho_t = f_t, \quad \text{for a.e. } t \in [0, 1]. \tag{4.1}$$

We regularize (4.1) as follows. Let \mathcal{M} be defined by

$$\mathcal{M} := \mathcal{M}(X) \times \mathcal{M}(X; \mathbb{R}^d) \times \mathcal{M}(X),$$

and introduce the convex linear space of triples in \mathcal{M} satisfying the continuity equation

$$\mathcal{D} := \{(\rho, m, \mu) \in \mathcal{M} : \partial_t \rho + \operatorname{div} m = \mu \text{ in the sense of (2.2)}\}.$$

Definition 4.1 (Regularized problem). Let $f \in L^2([0, 1]; H)$ and $\alpha, \beta > 0, \delta \in (0, \infty]$. The regularizer of (4.1) is the functional $J: \mathcal{M} \rightarrow [0, \infty]$ defined by

$$J(\rho, m, \mu) := \frac{1}{2} \int_0^1 \|K_t^* \rho_t - f_t\|_{H_t}^2 dt + \alpha B_\delta(\rho, m, \mu) + \beta \|\rho\|_{\mathcal{M}(X)} \tag{4.2}$$

if $(\rho, m, \mu) \in \mathcal{D}$ and $J = \infty$ otherwise. Here $\rho = dt \otimes \rho_t$ is the disintegration of ρ with respect to time and B_δ is the transport energy defined in (2.6).

We will proceed as follows. First, in Section 4.1 we show that the inverse problem in (4.1) and the functional J in (4.2) are well defined, in the following sense: Given a triple $(\rho, m, \mu) \in \mathcal{D}$ with finite transport energy $B_\delta(\rho, m, \mu)$, then $\rho \geq 0$ and it disintegrates into $\rho = dt \otimes \rho_t$ with $t \mapsto \rho_t$ narrowly continuous. For such curves, we show in Lemma 4.2, that $t \mapsto K_t^* \rho_t$ is a measurement belonging to $L^2([0, 1]; H)$, providing well-definition for (4.1) and (4.2). In Section 4.2 we show that

$$\inf_{(\rho, m, \mu) \in \mathcal{D}} J(\rho, m, \mu) \tag{4.3}$$

admits at least one solution, and the minimizer is unique under additional assumptions on the operators K_t^* . This will be the content of Theorem 4.4. Finally, in Section 4.3, we investigate stability of the solutions to (4.3) and convergence for vanishing noise level.

4.1. Well-definition

In this section we want to show that the definition of the functional J at (4.2) is well posed. The first step is to ensure that the fidelity term is well defined, namely, that given $\rho_t \in C_w([0, 1]; \mathcal{M}(\overline{\Omega}))$ then the map $t \in [0, 1] \mapsto K_t^* \rho_t \in H_t$ belongs to $L^2([0, 1]; H)$. This fact will be established in the following lemma.

Lemma 4.2. *If $\rho_t \in C_w([0, 1]; \mathcal{M}(\overline{\Omega}))$ then $(t \mapsto K_t^* \rho_t) \in L^2([0, 1]; H)$.*

Let us postpone the proof for a moment. As a consequence of the lemma we have the following.

Proposition 4.3. *The functional J at (4.2) is well defined.*

Proof of Lemma 4.2. Part 1. Let $\rho_t \in C_w([0, 1]; \mathcal{M}(\overline{\Omega}))$. First we show that the map $t \mapsto K_t^* \rho_t$ is strongly measurable according to Definition 3.2. We do so by means of Theorem 3.5, by proving that $t \mapsto K_t^* \rho_t$ is weakly measurable and essentially separably valued.

Claim 1. The map $t \mapsto K_t^* \rho_t$ is weakly measurable as per Definition 3.2, that is, $t \mapsto \langle K_t^* \rho_t, i_t \varphi \rangle_{H_t}$ is measurable for every fixed $\varphi \in D$.

Proof of Claim 1. By definition and (K1) we have $\langle K_t^* \rho_t, i_t \varphi \rangle_{H_t} = \langle \rho_t, K_t i_t \varphi \rangle_{\mathcal{M}(\overline{\Omega}), C(\overline{\Omega})}$. Notice that the map $t \in [0, 1] \mapsto K_t i_t \varphi \in C(\overline{\Omega})$ is strongly measurable in the classic sense ([27], Chap. II). To see this, since $C(\overline{\Omega})$ is separable, by the classic Pettis theorem ([27], Chap. II.1, Thm. 2), it is enough to prove that $t \mapsto K_t i_t \varphi$ is weakly measurable, meaning that $t \mapsto \langle \rho, K_t i_t \varphi \rangle_{\mathcal{M}(\overline{\Omega}), C(\overline{\Omega})}$ is measurable for every fixed $\rho \in \mathcal{M}(\overline{\Omega})$. The latter holds because $\langle \rho, K_t i_t \varphi \rangle_{\mathcal{M}(\overline{\Omega}), C(\overline{\Omega})} = \langle K_t^* \rho, i_t \varphi \rangle_{H_t}$ and the map $t \mapsto K_t^* \rho$ is strongly measurable by assumption

(K3), and hence weakly measurable by Theorem 3.5. By definition of classic strong measurability, there exists a sequence $\{f_n\}$ of step functions $f_n : [0, 1] \rightarrow C(\bar{\Omega})$, such that $f_n(t) = \sum_{j=1}^{N_n} \chi_{E_{j,n}} f_{j,n}$ with $\{E_{j,n}\}_{j=1}^{N_n}$ measurable partition of $[0, 1]$, $f_{j,n} \in C(\bar{\Omega})$, and such that

$$\lim_{n \rightarrow \infty} \|f_n(t) - K_t i_t \varphi\|_{C(\bar{\Omega})} = 0 \tag{4.4}$$

for a.e. $t \in [0, 1]$. We have that the map $t \mapsto \langle \rho_t, f_n(t) \rangle_{\mathcal{M}(\bar{\Omega}), C(\bar{\Omega})}$ is measurable for each fixed $n \in \mathbb{N}$, since $\langle \rho_t, f_n(t) \rangle_{\mathcal{M}(\bar{\Omega}), C(\bar{\Omega})} = \sum_{j=1}^{N_n} \chi_{E_{j,n}} \int_{\bar{\Omega}} f_{j,n} d\rho_t$ and the maps $t \mapsto \int_{\bar{\Omega}} f_{j,n} d\rho_t$ are continuous by narrow continuity of $t \mapsto \rho_t$. By Proposition A.3 we have that $\sup_{t \in [0, 1]} \|\rho_t\|_{\mathcal{M}(\bar{\Omega})} < \infty$. Combining this with (4.4) yields

$$|\langle \rho_t, f_n(t) - K_t i_t \varphi \rangle_{\mathcal{M}(\bar{\Omega}), C(\bar{\Omega})}| \leq C \|f_n(t) - K_t i_t \varphi\|_{C(\bar{\Omega})} \rightarrow 0$$

as $n \rightarrow \infty$, for a.e. $t \in [0, 1]$. Hence $t \mapsto \langle \rho_t, K_t i_t \varphi \rangle_{\mathcal{M}(\bar{\Omega}), C(\bar{\Omega})}$ is measurable, being the a.e. limit of measurable maps, and the claim follows.

Claim 2. The map $t \mapsto K_t^* \rho_t$ is essentially separably valued, that is, there exists a measurable set $E \subset [0, 1]$ such that $|E| = 0$ and a countable set $S \subset D$ with the following property: for every $\varepsilon > 0$ and $t \in [0, 1] \setminus E$ there exists $\varphi \in S$ such that $\|K_t^* \rho_t - i_t \varphi\|_{H_t} < \varepsilon$.

Proof of Claim 2. Let $T \subset [0, 1]$ be a countable dense subset. Fix $t \in T$. By (K3) the map $s \mapsto K_s^* \rho_t$ is strongly measurable and hence essentially separably valued by Theorem 3.5. Therefore there exists a measurable set $E_t \subset [0, 1]$ with $|E_t| = 0$ and a countable subset $S_t \subset D$ with the following property: for every $\varepsilon > 0$ and $s \in [0, 1] \setminus E_t$, there exists $\varphi \in S_t$ such that

$$\|K_s^* \rho_t - i_s \varphi\|_{H_s} < \varepsilon. \tag{4.5}$$

Denote by $E := \cup_{t \in T} E_t$. Since T is countable, the set E is measurable, and $|E| = 0$. Moreover let $S^0 := \cup_{t \in T} S_t$, so that $S^0 \subset D$ is countable. Define the set of averages of elements of S^0 as

$$S := \left\{ \frac{1}{k} \sum_{j=1}^k \varphi_j : k \geq 1, \varphi_j \in S^0 \right\}.$$

We have that $S \subset D$ is countable. Fix $\varepsilon > 0$, $t \in [0, 1] \setminus E$. The claim follows by showing there exists $\varphi \in S$ such that

$$\|K_t^* \rho_t - i_t \varphi\|_{H_t} < \varepsilon. \tag{4.6}$$

Indeed, by density, there exists a sequence $\{t_n\}$ in T such that $t_n \rightarrow t$ as $n \rightarrow \infty$. Since $\rho_t \in C_w([0, 1]; \mathcal{M}(\bar{\Omega}))$ it follows that $\rho_{t_n} \xrightarrow{*} \rho_t$ in $\mathcal{M}(\bar{\Omega})$. By weak*-to-weak continuity of K_t^* we have $K_t^* \rho_{t_n} \rightharpoonup K_t^* \rho_t$ weakly in H_t as $n \rightarrow \infty$. By the Banach–Saks property in H_t ([26], Chap. VIII, Thm. 1), there exists a subsequence (not relabelled) such that $\frac{1}{n} \sum_{j=1}^n K_t^* \rho_{t_j} \rightarrow K_t^* \rho_t$ strongly in H_t . Hence we can choose $N \in \mathbb{N}$ such that

$$\left\| K_t^* \rho_t - \frac{1}{N} \sum_{j=1}^N K_t^* \rho_{t_j} \right\|_{H_t} < \frac{\varepsilon}{2}. \tag{4.7}$$

Since $\{t_n\}$ is a sequence in T , by (4.5) and the definitions of S^0 and E , we have that for every $n \in \mathbb{N}$ there exists $\varphi_n \in S^0$ such that

$$\|K_t^* \rho_{t_n} - i_t \varphi_n\|_{H_t} < \frac{\varepsilon}{2}. \tag{4.8}$$

Define $\varphi := \frac{1}{N} \sum_{j=1}^N \varphi_j$, so that $\varphi \in S$. By triangle inequality, linearity of i_t , and (4.7), (4.8),

$$\begin{aligned} \|K_t^* \rho_t - i_t \varphi\|_{H_t} &\leq \left\| K_t^* \rho_t - \frac{1}{N} \sum_{j=1}^N K_t^* \rho_{t_j} \right\|_{H_t} + \left\| \frac{1}{N} \sum_{j=1}^N K_t^* \rho_{t_j} - i_t \varphi \right\|_{H_t} \\ &\leq \frac{\varepsilon}{2} + \frac{1}{N} \sum_{j=1}^N \|K_t^* \rho_{t_j} - i_t \varphi_j\|_{H_t} < \varepsilon \end{aligned}$$

which yields (4.6).

Part 2. Since $t \mapsto K_t^* \rho_t$ is strongly measurable, also $t \mapsto \|K_t^* \rho_t\|_{H_t}$ is measurable. By (K2) and Proposition A.3 we have

$$\int_0^1 \|K_t^* \rho_t\|_{H_t}^2 dt \leq C \int_0^1 \|\rho_t\|_{\mathcal{M}(\bar{\Omega})}^2 dt \leq C \sup_t \|\rho_t\|_{\mathcal{M}(\bar{\Omega})}^2 = C \left(\sup_t \|\rho_t\|_{\mathcal{M}(\bar{\Omega})} \right)^2 < \infty.$$

Hence by Theorem 3.9 we conclude that $K_t^* \rho_t$ is integrable and it belongs to $L^2([0, 1]; H)$. □

Proof of Proposition 4.3. If $J(\rho, m, \mu) < \infty$ then also $B_\delta(\rho, m, \mu) < \infty$, hence by Proposition 2.6 we have that $\rho \geq 0$, $m = v_t \rho$, $\mu = g_t \rho$ for Borel maps $v_t: X \rightarrow \mathbb{R}^d$, $g_t: X \rightarrow \mathbb{R}$ such that

$$\int_X (|v_t(x)|^2 + |g_t(x)|^2) d\rho(t, x) < \infty. \tag{4.9}$$

By assumption (ρ, m, μ) solves the continuity equation, hence (Prop. 2.2) $\rho = dt \otimes \rho_t$ for some Borel family $\rho_t \in \mathcal{M}^+(\bar{\Omega})$. In particular we have $m = dt \otimes (v_t \rho_t)$ and $\mu = dt \otimes (g_t \rho_t)$. By (4.9) and Proposition 2.4 we have that $\rho_t \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$. Therefore $K_t^* \rho_t \in L^2([0, 1]; H)$ by Lemma 4.2, and the first term in J is finite. □

4.2. Existence of minimizers

The aim of this section is to prove that the functional J defined in (4.2) admits at least one minimizer. Such minimizer is unique under suitable hypotheses on the operators K_t^* . The precise statement is the following.

Theorem 4.4. *Let $f \in L^2([0, 1]; H)$ and $\alpha, \beta > 0$, $\delta \in (0, \infty]$. Then there exists $(\rho^*, m^*, \mu^*) \in \mathcal{D}$ with $\rho^* = dt \otimes \rho_t^*$, $(t \mapsto \rho_t^*) \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$, that solves the minimization problem*

$$\inf_{(\rho, m, \mu) \in \mathcal{M}} J(\rho, m, \mu).$$

If in addition K_t^ is injective for a.e. $t \in [0, 1]$, then the minimizer is unique.*

The proof of the above theorem is based on the direct method of calculus of variations. Before proceeding to the proof, we will establish compactness and lower semicontinuity properties for the functional J . This is the object of the following two lemmas.

Lemma 4.5 (Compactness for J). *Let $f \in L^2([0, 1]; H)$ and $\alpha, \beta > 0$, $\delta \in (0, \infty]$. Assume that there exists a constant $E \geq 0$ such that the sequence $\{(\rho^n, m^n, \mu^n)\}$ in \mathcal{M} satisfies*

$$\sup_n J(\rho^n, m^n, \mu^n) \leq E. \tag{4.10}$$

Then $\rho^n = dt \otimes \rho_t^n$ for some $(t \mapsto \rho_t^n) \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$. Moreover there exists $(\rho, m, \mu) \in \mathcal{D}$ with $\rho = dt \otimes \rho_t$, $(t \mapsto \rho_t) \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$ such that, up to subsequences,

$$\begin{cases} (\rho^n, m^n, \mu^n) \overset{*}{\rightharpoonup} (\rho, m, \mu) \text{ weakly* in } \mathcal{M}, \\ \rho_t^n \overset{*}{\rightharpoonup} \rho_t \text{ weakly* in } \mathcal{M}(\bar{\Omega}), \text{ for every } t \in [0, 1]. \end{cases} \tag{4.11}$$

Proof. By the energy bound (4.10), there exists $\rho \in \mathcal{M}(X)$ such that, up to subsequences, $\rho^n \overset{*}{\rightharpoonup} \rho$ weakly* in $\mathcal{M}(X)$. From (4.10) we also have

$$\sup_n B_\delta(\rho^n, m^n, \mu^n) \leq E, \tag{4.12}$$

so that Proposition 2.6 implies $\rho^n \geq 0$, $m^n = v_t^n \rho^n$, $\mu^n = g_t^n \rho^n$ for Borel maps $v_t^n: X \rightarrow \mathbb{R}^d$, $g_t^n: X \rightarrow \mathbb{R}$ such that

$$\tilde{E} := \sup_n \int_X (|v_t^n(x)|^2 + |g_t^n(x)|^2) d\rho^n(t, x) < \infty. \tag{4.13}$$

By (4.10) we have $(\rho^n, m^n, \mu^n) \in \mathcal{D}$. Hence Proposition 2.2 implies that $\rho^n = dt \otimes \rho_t^n$ for some Borel family $\{\rho_t^n\}_{t \in [0,1]}$ in $\mathcal{M}^+(\bar{\Omega})$. In particular we have $m^n = dt \otimes (v_t^n \rho_t^n)$ and $\mu^n = dt \otimes (g_t^n \rho_t^n)$. Hence by (4.13) and Proposition 2.4 we get that $(t \mapsto \rho_t^n) \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$. Now notice that if $\rho \in \mathcal{M}^+(X)$, then by definition of B_δ (see (2.6)) we infer

$$\begin{aligned} \alpha B_\delta(\rho, m, \mu) + \beta \|\rho\|_{\mathcal{M}(X)} &= \alpha \left[\sup_{(a,b,c) \in C_0(X; K_\delta)} \int_X a d\rho + \int_X b \cdot dm + \int_X c d\mu \right] + \int_X \beta d\rho \\ &= \sup_{(a,b,c) \in C_0(X; \tilde{K}_\delta)} \int_X a d\rho + \int_X b \cdot dm + \int_X c d\mu, \end{aligned}$$

where K_δ is defined in (2.5) and

$$\tilde{K}_\delta := \left\{ (a, b, c) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} : a + \frac{1}{2\alpha} \left(|b|^2 + \frac{c^2}{\delta^2} \right) \leq \beta \right\}.$$

Therefore, by taking $0 < \varepsilon < \sqrt{2\alpha\beta} \min\{1, \delta\}$ we conclude

$$\begin{aligned} \varepsilon \|m\|_{\mathcal{M}(X; \mathbb{R}^d)} &= \sup_{\|b\|_\infty \leq \varepsilon} \int_X b \cdot dm \leq \alpha B_\delta(\rho, m, \mu) + \beta \|\rho\|_{\mathcal{M}(X)} \leq J(\rho, m, \mu), \\ \varepsilon \|\mu\|_{\mathcal{M}(X)} &= \sup_{\|c\|_\infty \leq \varepsilon} \int_X c d\mu \leq \alpha B_\delta(\rho, m, \mu) + \beta \|\rho\|_{\mathcal{M}(X)} \leq J(\rho, m, \mu). \end{aligned} \tag{4.14}$$

By combining (4.10) with the above estimates we conclude that (up to subsequences) $m^n \overset{*}{\rightharpoonup} m$ and $\mu^n \overset{*}{\rightharpoonup} \mu$ for some $m \in \mathcal{M}(X; \mathbb{R}^d)$, $\mu \in \mathcal{M}(X)$. Since \mathcal{D} is weak* closed, we get $(\rho, m, \mu) \in \mathcal{D}$. By Proposition 2.6 the functional B_δ is weak* lower semicontinuous. Therefore (4.12) implies $B_\delta(\rho, m, \mu) < \infty$ and by repeating the arguments above, we get that $\rho = dt \otimes \rho_t$, $m = dt \otimes (v_t \rho_t)$, $\mu = dt \otimes (g_t \rho_t)$ with $\rho_t \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$ and

$$\int_X (|v_t(x)|^2 + |g_t(x)|^2) d\rho(t, x) < \infty.$$

In particular $(\rho, m, \mu) \in \mathcal{D}$. We will now show the second condition in (4.11). Since (ρ^n, m^n, μ^n) solves the continuity equation, by Proposition 2.2 we have that the map $t \mapsto \rho_t^n(\bar{\Omega})$ belongs to $BV((0, 1))$ with distributional derivative given by $\pi_\# \mu^n$, where we recall that $\pi: X \rightarrow (0, 1)$ is the projection on the time coordinate. Therefore, by the embedding of $BV((0, 1))$ into $L^\infty((0, 1))$

$$\begin{aligned} \|\rho_t^n(\bar{\Omega})\|_\infty &\leq C \|\rho_t^n(\bar{\Omega})\|_{BV((0,1))} = C \int_0^1 \rho_t^n(\bar{\Omega}) dt + C \|\partial_t \rho_t^n(\bar{\Omega})\|_{\mathcal{M}((0,1))} \\ &= C \|\rho^n\|_{\mathcal{M}(X)} + C \|\pi_\# \mu^n\|_{\mathcal{M}((0,1))} \leq C \|\rho^n\|_{\mathcal{M}(X)} + C \|\mu^n\|_{\mathcal{M}(X)} \leq CE, \end{aligned} \tag{4.15}$$

where we used (4.14) and (4.10). Hence the set $\{\rho_t^n\}_{t,n}$ is uniformly bounded in $\mathcal{M}(\bar{\Omega})$, so it belongs to some set $K \subset \mathcal{M}(\bar{\Omega})$ which is weak* sequentially compact. Moreover as a consequence of Lemma A.2 we have that for every $t, s \in [0, 1]$

$$\|\rho_t^n - \rho_s^n\|_{C^1(\bar{\Omega})^*} \leq \sqrt{2C_n E_n} |t - s|^{1/2} \leq C |t - s|^{1/2},$$

where E_n, C_n are the constants defined in the same lemma. The last inequality follows from (4.13) and the fact that $\{\rho_t^n(\bar{\Omega})\}_{t,n}$ is uniformly bounded, so that the constant $C > 0$ does not depend on n . Hence by Proposition A.4 there exists a subsequence (not relabelled) and a $C^1(\bar{\Omega})^*$ -continuous curve $\tilde{\rho}_t: [0, 1] \rightarrow \mathcal{M}(\bar{\Omega})$ such that

$$\rho_t^n \xrightarrow{*} \tilde{\rho}_t \quad \text{weakly* in } \mathcal{M}(\bar{\Omega}) \quad \text{for every } t \in [0, 1]. \quad (4.16)$$

In particular $\tilde{\rho}_t$ is narrowly continuous, since it is $C^1(\bar{\Omega})^*$ -continuous (this fact can be obtained by repeating the same argument given in the proof of Prop. 2.4). Notice that (4.16) implies that $\rho^n \xrightarrow{*} \tilde{\rho}$ where $\tilde{\rho} := dt \otimes \tilde{\rho}_t$. Hence $\tilde{\rho} = \rho$. By uniqueness of the disintegration we also get that $\rho_t = \tilde{\rho}_t$ and the thesis follows. \square

Lemma 4.6 (Lower semicontinuity for J). *Let $f \in L^2([0, 1]; H)$ and $\alpha, \beta > 0$, $\delta \in (0, \infty]$. Assume that $(\rho^n, m^n, \mu^n) \in \mathcal{D}$ with $\rho^n = dt \otimes \rho_t^n$, $(t \mapsto \rho_t^n) \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$ is such that (ρ^n, m^n, μ^n) converges to (ρ, m, μ) in the sense of (4.11), where $\rho = dt \otimes \rho_t$, $(t \mapsto \rho_t) \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$. Then we have*

$$K_t^* \rho_t^n \rightharpoonup K_t^* \rho_t \quad \text{weakly in } L^2([0, 1]; H). \quad (4.17)$$

In particular J is lower semicontinuous with respect to the convergence in (4.11), that is,

$$J(\rho, m, \mu) \leq \liminf_n J(\rho^n, m^n, \mu^n). \quad (4.18)$$

Proof. Let us start by showing (4.17). To this end, fix $g \in L^2([0, 1]; H)$. By assumption we have that $\rho_t^n \xrightarrow{*} \rho_t$ weakly* in $\mathcal{M}(\bar{\Omega})$, for every $t \in [0, 1]$. In particular, by (K1'), we have $K_t^* \rho_t^n \rightharpoonup K_t^* \rho_t$ weakly in H_t for a.e. $t \in [0, 1]$, so that

$$\langle K_t^* \rho_t^n, g(t) \rangle_{H_t} \rightarrow \langle K_t^* \rho_t, g(t) \rangle_{H_t} \quad \text{for a.e. } t \in [0, 1] \quad (4.19)$$

as $n \rightarrow \infty$. By proceeding as in (4.15) we obtain

$$\|\rho_t^n(\bar{\Omega})\|_\infty \leq C \|\rho^n\|_{\mathcal{M}(X)} + C \|\mu^n\|_{\mathcal{M}(X)} \leq C,$$

for some generic constant $C \geq 0$, since ρ^n and μ^n are uniformly bounded by weak* convergence in $\mathcal{M}(X)$. Hence by Cauchy–Schwarz and by (K2) we have

$$|\langle K_t^* \rho_t^n, g(t) \rangle_{H_t}| \leq \|K_t^* \rho_t^n\|_{H_t} \|g(t)\|_{H_t} \leq C \rho_t^n(\bar{\Omega}) \|g(t)\|_{H_t} \leq C \|g(t)\|_{H_t}.$$

Since $g \in L^2([0, 1]; H)$ we have that $t \mapsto \|g(t)\|_{H_t}$ belongs to $L^1((0, 1))$. By combining the above estimate with (4.19) and invoking the classic dominated convergence theorem we conclude (4.17).

Let us now prove the remaining part of the lemma. From (4.17) we have that $K_t^* \rho_t^n - f_t$ converges weakly to $K_t^* \rho_t - f_t$ in $L^2([0, 1]; H)$, therefore by lower semicontinuity of the norm with respect to the weak convergence we have

$$\frac{1}{2} \|K_t^* \rho_t - f_t\|_{L^2}^2 \leq \liminf_n \frac{1}{2} \|K_t^* \rho_t^n - f_t\|_{L^2}^2.$$

Moreover B_δ is weak* lower semicontinuous by Proposition 2.6, thus

$$\begin{aligned} J(\rho, m, \mu) &\leq \liminf_n \frac{1}{2} \|K_t^* \rho_t^n - f_t\|_{L^2}^2 + \liminf_n \left(\alpha B_\delta(\rho^n, m^n, \mu^n) + \beta \|\rho^n\|_{\mathcal{M}(X)} \right) \\ &\leq \liminf_n J(\rho^n, m^n, \mu^n), \end{aligned}$$

and (4.18) follows. \square

Proof of Theorem 4.4. Existence. Set $\rho := dt \otimes \sigma$ with $\sigma \in \mathcal{M}^+(\bar{\Omega})$. Then $(\rho, 0, 0) \in \mathcal{D}$, and $B_\delta(\rho, 0, 0) = 0$ by (iv) in Proposition 2.6. Moreover $(t \mapsto K_t^* \sigma) \in L^2([0, 1]; H)$ by Lemma 4.2, so that $J(\rho, 0, 0) < \infty$ and the infimum is finite. Let (ρ^n, m^n, μ^n) be a minimizing sequence, that is, $J(\rho^n, m^n, \mu^n) \rightarrow \inf_{\mathcal{M}} J$ as $n \rightarrow \infty$.

Therefore $\sup_n J(\rho^n, m^n, \mu^n) \leq E$ for some constant $E \geq 0$. From Lemma 4.5 we have $\rho^n = dt \otimes \rho_t^n$ and $\rho_t^n \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$. Moreover there exists $(\rho^*, m^*, \mu^*) \in \mathcal{D}$ with $\rho^* = dt \otimes \rho_t^*$, $(t \mapsto \rho_t^*) \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$ such that, up to subsequences, (ρ^n, m^n, μ^n) converges to (ρ^*, m^*, μ^*) in the sense of (4.11). By Lemma 4.6 and the fact that (ρ^n, m^n, μ^n) is a minimizing sequence we conclude that (ρ^*, m^*, μ^*) is a minimizer for J .

Uniqueness. Assume that K_t^* is injective for a.e. $t \in [0, 1]$. The term B_δ is convex by Proposition 2.6. Also the term $\|\cdot\|_{\mathcal{M}(X)}$ is convex as it is a norm. Since minimizers are necessarily of the form $(\rho, m, \mu) \in \mathcal{D}$ with $\rho \geq 0$, in order to prove uniqueness it is sufficient to show that

$$\rho_t \mapsto \|K_t^* \rho_t - f_t\|_{L^2}^2 \tag{4.20}$$

is strictly convex for $\rho_t \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$. First, consider $\rho_t \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$ such that $\rho_t \not\equiv 0$. As a consequence the set $E := \{t \in [0, 1] : \rho_t(\bar{\Omega}) \neq 0\}$ is open (by continuity of $t \mapsto \rho_t(\bar{\Omega})$) and not empty, so $|E| > 0$. Let $F := \{t \in [0, 1] : K_t^* \text{ is injective}\}$. By assumption we have $|[0, 1] \setminus F| = 0$. Therefore

$$\|K_t^* \rho_t\|_{L^2}^2 = \int_0^1 \|K_t^* \rho_t\|_{H_t}^2 dt \geq \int_{E \cap F} \|K_t^* \rho_t\|_{H_t}^2 dt > 0 \tag{4.21}$$

since $K_t^* \rho_t \neq 0$ for $t \in E \cap F$. Now let $\rho_t^1, \rho_t^2 \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$ with $\rho_t^1 \not\equiv \rho_t^2$ and $\lambda \in (0, 1)$. The coefficient of the leading term of $\lambda \mapsto \|K_t^*(\lambda \rho_t^1 + (1 - \lambda)\rho_t^2)\|_{L^2}^2$ is $\|K_t^*(\rho_t^1 - \rho_t^2)\|_{L^2}^2$, which is non zero by (4.21), since $\rho_t^1 \not\equiv \rho_t^2$. Hence the map in (4.20) is strictly convex and we conclude. \square

4.3. Regularization properties

In this section we denote by $f^\dagger \in L^2([0, 1]; H)$ the exact data and by $f^\gamma \in L^2([0, 1]; H)$ the noisy data for the noise level $\gamma > 0$, that is, $\|f^\gamma - f^\dagger\|_{L^2} \leq \gamma$. For a datum $f \in L^2([0, 1]; H)$ and parameters $\alpha, \beta > 0$ we adopt the following notation:

$$J_{\alpha, \beta, f}(\rho, m, \mu) := \frac{1}{2} \|K_t^* \rho_t - f_t\|_{L^2}^2 + \alpha B_\delta(\rho, m, \mu) + \beta \|\rho\|_{\mathcal{M}(X)}$$

if $(\rho, m, \mu) \in \mathcal{D}$ and $J_{\alpha, \beta, f}(\rho, m, \mu) = \infty$ otherwise.

Theorem 4.7 (Stability). *Assume that $f^n \rightarrow f^\gamma$ strongly in $L^2([0, 1]; H)$ and that*

$$(\rho^n, m^n, \mu^n) \in \arg \min_{(\rho, m, \mu) \in \mathcal{M}} J_{\alpha, \beta, f^n}(\rho, m, \mu). \tag{4.22}$$

Then $(\rho^n, m^n, \mu^n) \in \mathcal{D}$ with $\rho^n = dt \otimes \rho_t^n$, $\rho_t^n \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$. Moreover (ρ^n, m^n, μ^n) admits a subsequence converging in the sense of (4.11). The limit of each converging subsequence of (ρ^n, m^n, μ^n) is a minimizer of $J_{\alpha, \beta, f^\gamma}$.

Proof. A sequence $(\rho^n, m^n, \mu^n) \in \mathcal{D}$ satisfying (4.22) exists by Theorem 4.4. By the same theorem it also follows that $\rho^n = dt \otimes \rho_t^n$ with $\rho_t^n \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$. We have

$$J_{\alpha, \beta, f^\gamma}(\rho^n, m^n, \mu^n) \leq 2J_{\alpha, \beta, f^n}(\rho^n, m^n, \mu^n) + \|f_t^\gamma - f_t^n\|_{L^2}^2.$$

Since (ρ^n, m^n, μ^n) is a minimizer for J_{α, β, f^n} we can test (4.22) against $(0, 0, 0) \in \mathcal{D}$ to obtain

$$\sup_n J_{\alpha, \beta, f^\gamma}(\rho^n, m^n, \mu^n) \leq \|f_t^n\|_{L^2}^2 + \|f_t^n - f_t^\gamma\|_{L^2}^2 < \infty,$$

where last inequality follows from the convergence $f^n \rightarrow f^\gamma$ in $L^2([0, 1]; H)$. By applying Lemma 4.5 to $J_{\alpha, \beta, f^\gamma}$, there exists $(\tilde{\rho}, \tilde{m}, \tilde{\mu}) \in \mathcal{D}$, with $\tilde{\rho} = dt \otimes \tilde{\rho}_t$, $\tilde{\rho}_t \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$ such that (ρ^n, m^n, μ^n) converges to $(\tilde{\rho}, \tilde{m}, \tilde{\mu})$

in the sense of (4.11). We are left to show that $(\tilde{\rho}, \tilde{m}, \tilde{\mu})$ is a minimizer for $J_{\alpha, \beta, f^\gamma}$. Since $\rho_t^n \xrightarrow{*} \tilde{\rho}_t$ for every $t \in [0, 1]$, by Lemma 4.6 and the convergence $f^n \rightarrow f^\gamma$ we have $(K_t^* \rho_t^n - f_t^n) \rightharpoonup (K_t^* \tilde{\rho}_t - f_t^\gamma)$ weakly in $L^2([0, 1]; H)$. Also by Lemma 4.6,

$$J_{\alpha, \beta, f^\gamma}(\tilde{\rho}, \tilde{m}, \tilde{\mu}) \leq \liminf_n J_{\alpha, \beta, f^n}(\rho^n, m^n, \mu^n) \leq \liminf_n J_{\alpha, \beta, f^n}(\rho, m, \mu) = J_{\alpha, \beta, f^\gamma}(\rho, m, \mu)$$

for every $(\rho, m, \mu) \in \mathcal{M}$, since (4.22) holds and $f^n \rightarrow f^\gamma$. Hence $(\tilde{\rho}, \tilde{m}, \tilde{\mu})$ is a minimizer. \square

We are now interested in studying properties of the minimizers of $J_{\alpha, \beta, f^\gamma}$ for vanishing noise level, that is, for data such that $\|f^\gamma - f^\dagger\|_{L^2} \leq \gamma$ for every $\gamma \geq 0$. To this end, we need to understand how the regularization term

$$\alpha B_\delta(\rho, m, \mu) + \beta \|\rho\|_{\mathcal{M}(X)}$$

behaves for fixed argument $(\rho, m, \mu) \in \mathcal{D}$. Since multiple parameters are involved, we will also allow α and β to take the value ∞ . We define

$$\beta \|\cdot\|_{\mathcal{M}(X)} := I_{\{0\}} \quad \text{if } \beta = \infty, \tag{4.23}$$

where $I_{\{0\}}$ denotes the convex indicator function of the set $\{0\}$. In order to give a similar definition for the case $\alpha = \infty$ we first need to characterize the subset of \mathcal{D} where $B_\delta(\rho, m, \mu) = 0$.

Proposition 4.8. *Assume that $(\rho, m, \mu) \in \mathcal{D}$. We have that $B_\delta(\rho, m, \mu) = 0$ if and only if $m = 0$, $\mu = 0$ and $\rho = dt \otimes \sigma$ for some $\sigma \in \mathcal{M}^+(\bar{\Omega})$.*

Proof. By Proposition 2.6 point (iv) we have that $B_\delta(dt \otimes \sigma, 0, 0) = 0$ for any $\sigma \in \mathcal{M}^+(\bar{\Omega})$. Conversely, assume that $(\rho, m, \mu) \in \mathcal{D}$ is such that $B_\delta(\rho, m, \mu) = 0$. In particular the energy is finite, so points (iii)–(iv) of Proposition 2.6 imply that $\rho \geq 0$, $m = v\rho$, $\mu = g\rho$ for some Borel maps $v: X \rightarrow \mathbb{R}^d$, $g: X \rightarrow \mathbb{R}$, and we have $B_\delta(\rho, m, \mu) = \frac{1}{2} \int_X (|v|^2 + \delta^2 g^2) d\rho$. Since $B_\delta(\rho, m, \mu) = 0$ and $\rho \geq 0$, we conclude that $m = 0$ and $\mu = 0$. By assumption $(\rho, 0, 0)$ solves the continuity equation in the sense of (2.2), therefore Proposition 2.2 guarantees that $\rho = dt \otimes \rho_t$ for some Borel family $\rho_t \in \mathcal{M}^+(\bar{\Omega})$. Since $v = 0$ and $g = 0$ a.e. in X , we can apply Proposition 2.4 and conclude that $(t \mapsto \rho_t) \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$. In particular, for every $0 \leq t_1 \leq t_2 \leq 1$ and $\varphi(t, x) := a(x)$ with $a \in C^1(\bar{\Omega})$, formula (2.4) reads $\int_{\bar{\Omega}} a(x) d\rho_{t_1}(x) = \int_{\bar{\Omega}} a(x) d\rho_{t_2}(x)$. By a density argument one can show that the previous holds for all $a \in C(\bar{\Omega})$, and hence $\rho_t = \rho_0$ for every $t \in [0, 1]$. The thesis follows by setting $\sigma := \rho_0$. \square

Proposition 4.8 motivates the following definition:

$$\alpha B_\delta := I_Z \quad \text{if } \alpha = \infty, \tag{4.24}$$

where $Z := \{(\rho, 0, 0) \in \mathcal{D} : \rho = dt \otimes \sigma, \sigma \in \mathcal{M}^+(\bar{\Omega})\}$. We are now in the position to define minimal energy solutions of the inverse problem

$$K_t^* \rho_t = f_t^\dagger \quad \text{for a.e. } t \in [0, 1]. \tag{4.25}$$

Definition 4.9 (Minimal energy solution). Let $f^\dagger \in L^2([0, 1]; H)$ and $\alpha^*, \beta^* \in [1, \infty]$, $\delta \in (0, \infty]$. We say that $(\rho^\dagger, m^\dagger, \mu^\dagger) \in \mathcal{M}$ is a *minimal energy solution* of (4.25) with parameters α^*, β^* if

$$(\rho^\dagger, m^\dagger, \mu^\dagger) \in \arg \min_{(\rho, m, \mu) \in \mathcal{D}} \left\{ \alpha^* B_\delta(\rho, m, \mu) + \beta^* \|\rho\|_{\mathcal{M}(X)} : K_t^* \rho_t = f_t^\dagger \quad \text{a.e. in } [0, 1] \right\}.$$

In the following theorem we show that the minimizers for vanishing noise level converge in the sense of (4.11) to an energy minimizing solution of the inverse problem (4.25).

Theorem 4.10 (Convergence for vanishing noise level). *Let $f^\dagger \in L^2([0, 1]; H)$ be the exact data and $\{f^n\}$ be a sequence of noisy data such that $\|f^n - f^\dagger\|_{L^2} \leq \gamma_n$, $\gamma_n \rightarrow 0$ as $n \rightarrow \infty$. Let $\alpha_n, \beta_n > 0$ be such that*

$$\alpha_n, \beta_n \searrow 0 \quad \text{and} \quad \frac{\gamma_n^2}{\alpha_n}, \frac{\gamma_n^2}{\beta_n} \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty. \tag{4.26}$$

Let $c_n := \min\{\alpha_n, \beta_n\}$, $\tilde{\alpha}_n := \alpha_n/c_n$, $\tilde{\beta}_n := \beta_n/c_n$ so that, up to subsequences, $\tilde{\alpha}_n \rightarrow \alpha^$ and $\tilde{\beta}_n \rightarrow \beta^*$ as $n \rightarrow \infty$, with $\alpha^*, \beta^* \in [1, \infty]$. Assume there exists $(\rho^\dagger, m^\dagger, \mu^\dagger) \in \mathcal{D}$ satisfying $\rho^\dagger = dt \otimes \rho_t^\dagger$, $\rho_t^\dagger \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$, (4.25) and*

$$\alpha^* B_\delta(\rho^\dagger, m^\dagger, \mu^\dagger) + \beta^* \|\rho^\dagger\|_{\mathcal{M}(X)} < \infty. \tag{4.27}$$

Let $(\rho^n, m^n, \mu^n) \in \mathcal{M}$ be such that

$$(\rho^n, m^n, \mu^n) \in \underset{(\rho, m, \mu) \in \mathcal{M}}{\operatorname{arg\,min}} J_{\alpha_n, \beta_n, f^n}(\rho, m, \mu). \tag{4.28}$$

Then $\rho^n = dt \otimes \rho_t^n$ with $\rho_t^n \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$, and (ρ^n, m^n, μ^n) converges to (ρ^, m^*, μ^*) in the sense of (4.11), up to subsequences. Moreover, every such weak* limit of (ρ^n, m^n, μ^n) is a minimal energy solution of (4.25) with parameters α^* and β^* .*

Proof. First notice that $c_n \rightarrow 0$ and $\gamma_n^2/c_n \rightarrow 0$ as $n \rightarrow \infty$ by (4.26). If $\tilde{\alpha}_n \rightarrow \infty$ or $\tilde{\beta}_n \rightarrow \infty$, we set $\alpha^* := \infty$ and $\beta^* := \infty$ respectively. If either of the sequences do not diverge to ∞ , it is possible to find accumulation points $\alpha^*, \beta^* \in [1, \infty)$. In particular, up to extracting a subsequence, we can assume that $\tilde{\alpha}_n \rightarrow \alpha^*$ and $\tilde{\beta}_n \rightarrow \beta^*$ as $n \rightarrow \infty$. A sequence (ρ^n, m^n, μ^n) satisfying (4.28) exists by Theorem 4.4. By the same theorem it also follows that $\rho^n = dt \otimes \rho_t^n$ with $\rho_t^n \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$. By testing (4.28) against $(\rho^\dagger, m^\dagger, \mu^\dagger)$ and using (4.25) and (4.26) we get

$$J_{\alpha_n, \beta_n, f^n}(\rho^n, m^n, \mu^n) \leq \frac{\gamma_n^2}{2} + \alpha_n B_\delta(\rho^\dagger, m^\dagger, \mu^\dagger) + \beta_n \|\rho^\dagger\|_{\mathcal{M}(X)} \rightarrow 0. \tag{4.29}$$

In particular $(K_t^* \rho_t^n - f_t^\dagger) \rightarrow 0$ in $L^2([0, 1]; H)$. Since by assumption $f^n \rightarrow f^\dagger$, we obtain $K_t^* \rho_t^n \rightarrow f_t^\dagger$ in $L^2([0, 1]; H)$. Dividing the inequality at (4.29) by c_n , taking the limes superior and keeping (4.26) in mind yields

$$\limsup_n \left(\tilde{\alpha}_n B_\delta(\rho^n, m^n, \mu^n) + \tilde{\beta}_n \|\rho^n\|_{\mathcal{M}(X)} \right) \leq \alpha^* B_\delta(\rho^\dagger, m^\dagger, \mu^\dagger) + \beta^* \|\rho^\dagger\|_{\mathcal{M}(X)}. \tag{4.30}$$

Notice that the right hand side in (4.30) is always bounded, thanks to definitions (4.23), (4.24) and assumption (4.27). By definition we have $\tilde{\alpha}_n, \tilde{\beta}_n \geq 1$, thus

$$\limsup_n J_{1,1,f^\dagger}(\rho^n, m^n, \mu^n) \leq \limsup_n \left(\tilde{\alpha}_n B_\delta(\rho^n, m^n, \mu^n) + \tilde{\beta}_n \|\rho^n\|_{\mathcal{M}(X)} \right) < \infty$$

where we used (4.30) and the convergence $K_t^* \rho_t^n \rightarrow f_t^\dagger$. Therefore an application of Lemma 4.5 guarantees the existence of $(\rho^*, m^*, \mu^*) \in \mathcal{D}$ with $\rho^* = dt \otimes \rho_t^*$, $\rho_t^* \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$, such that, up to subsequences, (ρ^n, m^n, μ^n) converges to (ρ^*, m^*, μ^*) in the sense of (4.11). In particular by Lemma 4.6 we have $K_t^* \rho_t^n \rightharpoonup K_t^* \rho_t^*$ weakly in $L^2([0, 1]; H)$. Since we already proved that $K_t^* \rho_t^n \rightarrow f_t^\dagger$, by uniqueness of the weak limit we have

$$K_t^* \rho_t^* = f_t^\dagger \quad \text{a.e. in} \quad [0, 1]. \tag{4.31}$$

We are left to show that (ρ^*, m^*, μ^*) is an energy minimizing solution of (4.25). By Lemma 4.6

$$\begin{aligned} \alpha^* B_\delta(\rho^*, m^*, \mu^*) + \beta^* \|\rho^*\|_{\mathcal{M}(X)} &\leq \liminf_n \left(\tilde{\alpha}_n B_\delta(\rho^n, m^n, \mu^n) + \tilde{\beta}_n \|\rho^n\|_{\mathcal{M}(X)} \right) \\ &\leq \limsup_n \left(\tilde{\alpha}_n B_\delta(\rho^n, m^n, \mu^n) + \tilde{\beta}_n \|\rho^n\|_{\mathcal{M}(X)} \right) \\ &\leq \alpha^* B_\delta(\rho^\dagger, m^\dagger, \mu^\dagger) + \beta^* \|\rho^\dagger\|_{\mathcal{M}(X)} \end{aligned} \tag{4.32}$$

where we used (4.30) and that $\tilde{\alpha}_n \rightarrow \alpha^*$, $\tilde{\beta}_n \rightarrow \beta^*$. Replacing $(\rho^\dagger, m^\dagger, \mu^\dagger)$ by an arbitrary solution of (4.25) with finite energy B_δ , the argument can be repeated, and from (4.31), (4.32) we conclude that (ρ^*, m^*, μ^*) is an energy minimizing solution of (4.25). \square

5. APPLICATION TO DYNAMIC UNDERSAMPLED MRI

We will now detail on the application of the above results to dynamic magnetic resonance imaging as outlined in the introduction. Let $\Omega \subset \mathbb{R}^2$ be an open bounded domain representing the image frame, and $c_j \in C_0(\mathbb{R}^2; \mathbb{C})$ for $j = 1, \dots, N$ with $N \geq 1$ be the coil sensitivities. Let $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ for $t \in [0, 1]$ be a family of measures such that

- (M1) $\|\sigma_t\|_{\mathcal{M}(\mathbb{R}^2)} \leq C$ for a.e. $t \in [0, 1]$, where $C > 0$ does not depend on t ,
- (M2) the map $t \mapsto \int_{\mathbb{R}^2} \varphi(x) d\sigma_t(x)$ is measurable for each $\varphi \in C_0(\mathbb{R}^2; \mathbb{C})$.

Let $D := C_0(\mathbb{R}^2; \mathbb{C}^N)$ be the Banach space normed by $\|\varphi\|_\infty := \max_{\{j=1, \dots, N\}} \max_{x \in \mathbb{R}^2} |\varphi^j(x)|$, where $\varphi = (\varphi^1, \dots, \varphi^N)$. Define Hilbert spaces $H_t := L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C}^N)$, equipped with the norm $\|h\|_{H_t}^2 := \sum_{j=1}^N \int_{\mathbb{R}^2} |h^j(x)|^2 d\sigma_t(x)$, where $h = (h^1, \dots, h^N)$. Define $i_t: D \rightarrow H_t$ as the identity map, acting component-wise. Note that here we are interpreting D and H_t as real vector spaces. For a measure $\rho \in \mathcal{M}(\bar{\Omega}; \mathbb{C})$ we denote its Fourier transform as

$$\mathcal{F}\rho(x) := \frac{1}{2\pi} \int_{\mathbb{R}^2} e^{-i\omega \cdot x} d\rho(\omega),$$

so that $\mathcal{F}\rho \in C(\mathbb{R}^2; \mathbb{C})$. Notice that in the above definition we extend ρ to be zero outside of $\bar{\Omega}$. For each $t \in [0, 1]$, define the linear operator $K_t^*: \mathcal{M}(\bar{\Omega}) \rightarrow H_t$ as

$$K_t^* \rho := (\mathcal{F}(c_1 \rho), \dots, \mathcal{F}(c_N \rho)). \tag{5.1}$$

In the MRI context, the family of measures $\rho_t \in \mathcal{M}^+(\bar{\Omega})$ for $t \in [0, 1]$ represents the proton density at each time step. Given some data $f \in L^2([0, 1]; H)$, we want to reconstruct a solution to the dynamic inverse problem

$$K_t^* \rho_t = f_t \quad \text{for a.e. } t \in [0, 1]. \tag{5.2}$$

As proposed in the previous sections, we relax the problem to measures $\rho \in \mathcal{M}(X)$, with $X := (0, 1) \times \bar{\Omega}$, and minimize the functional J introduced in (4.2). Under the assumptions (M1) and (M2) the functional J admits at least one minimizer, and minimizers are unique under suitable additional assumptions. This claim is the object of the following theorem.

Theorem 5.1. *Let $\alpha, \beta > 0$, $\delta \in (0, \infty]$, $f \in L^2([0, 1]; H)$. Let $\{\sigma_t\}$ for $t \in [0, 1]$ be a family of Radon measures in $\mathcal{M}^+(\mathbb{R}^2)$ satisfying (M1) and (M2). Let $c_1, \dots, c_N \in C_0(\mathbb{R}^2; \mathbb{C})$ be coil sensitivities. Then the regularization of (5.2) according to*

$$\inf_{(\rho, m, \mu) \in \mathcal{D}} \frac{1}{2} \sum_{j=1}^N \int_0^1 \left\| \mathcal{F}(c_j \rho_t) - f_t^j \right\|_{L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C})}^2 dt + \alpha B_\delta(\rho, m, \mu) + \beta \|\rho\|_{\mathcal{M}(X)}$$

admits a solution $(\rho^, m^*, \mu^*) \in \mathcal{D}$ with $\rho^* = dt \otimes \rho_t^*$ and the curve $t \mapsto \rho_t^*$ belonging to $C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$. If in addition the supports of the measures σ_t have non empty interior for a.e. $t \in [0, 1]$, and the vector of coil sensitivities satisfies $c(x) \neq 0$ for every $x \in \bar{\Omega}$, then the minimizer is unique.*

Before proceeding with the proof, we want to show how this analytical framework allows us to treat a wide variety of sampling patterns. We will give two examples.

Example 5.2 (Continuous sampling). Let $\Omega := (-\frac{1}{2}, \frac{1}{2})^2$ and for $t \in [0, 1]$ define the line $L_t := (-\frac{1}{2}, \frac{1}{2}) \times \{t - \frac{1}{2}\}$. Set $\sigma_t := \mathcal{H}^1 \llcorner L_t$, that is, the restriction of the 1-dimensional Hausdorff measure to the lines L_t . It is immediate to check that $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ satisfies (M1) and (M2): indeed $\|\sigma_t\|_{\mathcal{M}^+(\mathbb{R}^2)} = \mathcal{H}^1(L_t) = 1$, while the map $t \mapsto \int_{\mathbb{R}^2} \varphi(x) d\sigma_t(x)$ is continuous for φ in $C_0(\mathbb{R}^2; \mathbb{C})$. In the same way we can treat radial sampling, by setting L_t to be a collection of diameters through the origin, evolving in time (see Example B.3).

Example 5.3 (Compressed-sensing sampling). In this example we propose to sample along a finite collection of moving points in an open bounded domain $\Omega \subset \mathbb{R}^2$. To be more specific, fix $M \in \mathbb{N}$, $M \geq 1$ and for every $j = 1, \dots, M$ let $t \in [0, 1] \mapsto x_t^j \in \Omega$ be a measurable curve. For a.e. $t \in [0, 1]$ define $P_t := \{x_t^1, \dots, x_t^M\}$ and $\sigma_t := \mathcal{H}^0 \llcorner P_t = \sum_{j=1}^M \delta_{x_t^j}$. Notice that (M1) is satisfied since $\|\sigma_t\|_{\mathcal{M}(\mathbb{R}^2)} = M$. Given a map $\varphi \in C_0(\mathbb{R}^2; \mathbb{C})$ we have that $t \mapsto \int_{\mathbb{R}^2} \varphi(x) d\sigma_t(x) = \sum_{j=1}^M \varphi(x_t^j)$ is measurable by construction. Therefore also (M2) is satisfied.

We now want to prove Theorem 5.1. Before that, we need a preliminary lemma, stating that under (M1) and (M2) the above definitions of D, H_t, i_t, K_t^* satisfy assumptions (H1)–(H3) and (K1)–(K3).

Lemma 5.4. *Assume that (M1) and (M2) hold. The spaces D, H_t and the operators i_t satisfy (H1)–(H3). Moreover the operators K_t^* in (5.1) satisfy (K1)–(K3).*

Proof. Notice that i_t is linear and continuous, with $\|i_t\|^2 \leq N \|\sigma_t\|_{\mathcal{M}(\mathbb{R}^2)}$. In particular (H1) follows from (M1). Moreover (H2) is trivially satisfied. Finally for $\varphi, \psi \in D$ we have

$$t \mapsto \langle i_t \varphi, i_t \psi \rangle_{H_t} = \Re \left(\sum_{j=1}^N \int_{\mathbb{R}^2} \varphi^j(x) \overline{\psi^j(x)} d\sigma_t(x) \right),$$

which is measurable by (M2), as it is the real part of a sum of measurable maps. Hence (H3) is also satisfied. Let us now show that (K1)–(K3) hold. For $\rho \in \mathcal{M}(\overline{\Omega})$ we have

$$|\mathcal{F}(c_j \rho)(x)| = \frac{1}{2\pi} \left| \int_{\mathbb{R}^2} e^{-i\omega \cdot x} c_j(\omega) d\rho(\omega) \right| \leq \frac{\|c_j\|_\infty}{2\pi} \|\rho\|_{\mathcal{M}(\overline{\Omega})}. \tag{5.3}$$

Hence, each $\mathcal{F}(c_j \rho)$ is square integrable with respect to σ_t , so that K_t^* maps $\mathcal{M}(\overline{\Omega})$ into H_t . Moreover, by the above estimate we also have

$$\|K_t^* \rho\|_{H_t}^2 = \sum_{j=1}^N \int_{\mathbb{R}^2} |\mathcal{F}(c_j \rho)(x)|^2 d\sigma_t(x) \leq \frac{N}{4\pi^2} \|c\|_\infty^2 \|\rho\|_{\mathcal{M}(\overline{\Omega})}^2 \|\sigma_t\|_{\mathcal{M}(\mathbb{R}^2)},$$

where $c := (c_1, \dots, c_N)$ is the vector of coil sensitivities. Therefore K_t^* is continuous, with $\|K_t^*\|^2 \leq \frac{N}{4\pi^2} \|c\|_\infty^2 \|\sigma_t\|_{\mathcal{M}(\mathbb{R}^2)}$ and (K2) is satisfied because of assumption (M1). Let us show that K_t^* is weak*-to-weak continuous. To this end, let $\{\rho_n\}$ in $\mathcal{M}(\overline{\Omega})$ be such that $\rho_n \xrightarrow{*} \rho$. Since ρ_n, ρ are supported in the compact set $\overline{\Omega}$, it follows that

$$\mathcal{F}(c_j \rho_n) \rightarrow \mathcal{F}(c_j \rho) \quad \text{pointwise in } \mathbb{R}^2, \tag{5.4}$$

as $n \rightarrow \infty$. Moreover, by weak* convergence we have $\sup_n \|\rho_n\|_{\mathcal{M}(\overline{\Omega})} < \infty$. As a consequence of (5.3), there exists a constant $C \geq 0$ such that

$$\sup_n \|\mathcal{F}(c_j \rho_n)\|_\infty \leq C \quad \text{for every } j = 1, \dots, N. \tag{5.5}$$

By invoking the dominated convergence theorem in conjunction with (5.4) and (5.5) we conclude that $K_t^* \rho_n \rightharpoonup K_t^* \rho$ weakly in $L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C}^N)$. Hence (K1') is satisfied and, as a consequence, K_t^* is the adjoint of some linear continuous operator $K_t: H_t \rightarrow C(\overline{\Omega})$. Finally we need to show (K3): that the map $t \mapsto K_t^* \rho$ is strongly

measurable according to Definition 3.2, for every fixed $\rho \in \mathcal{M}(\overline{\Omega})$. Notice that the space $D = C_0(\mathbb{R}^2; \mathbb{C}^N)$ is separable, hence by Proposition 3.7 it is sufficient to show that $t \mapsto K_t^* \rho$ is weakly measurable according to Definition 3.2. However, since $\mathcal{F}(c_j \rho) \in C(\mathbb{R}^2; \mathbb{C})$ is bounded (see (5.3)) and maps in D are continuous, this is an immediate consequence of (M2). \square

Proof of Theorem 5.1. The existence of a minimizer follows from Proposition 5.4 and of Theorem 4.4. For the uniqueness, from Theorem 4.4, it is sufficient to check that the operators $K_t^* : \mathcal{M}(\overline{\Omega}) \rightarrow H_t$ are injective for a.e. $t \in [0, 1]$. To this end, choose $t \in [0, 1]$ such that $\text{supp } \sigma_t$ has non empty interior, and let $\rho \in \mathcal{M}(\overline{\Omega})$ be such that $K_t^* \rho = 0$. In particular $\mathcal{F}(c_j \rho) = 0$ in $\text{supp } \sigma_t$, for every $j = 1, \dots, N$. Since $\mathcal{F}(c_j \rho)$ is analytic and $\text{supp } \sigma_t$ contains an open ball, we conclude that $\mathcal{F}(c_j \rho) = 0$ in \mathbb{R}^2 . By injectivity of the Fourier transform we have that $c_j \rho = 0$, and since we are assuming that $c(x) \neq 0$ for every $x \in \overline{\Omega}$, we conclude that $\rho = 0$. \square

6. CONCLUSIONS AND PERSPECTIVES

In the paper, we have shown that it is possible to successfully use energy functionals that are associated with a dynamic formulation of optimal transport as regularization functionals for dynamic inverse problems that aim at the recovery of measure-valued curves. Let us point out some future directions of research. On the one hand, the focus of the paper is on regularizers that penalize mass transport by the squared distance as well as mass growth in terms of quadratic costs for growth rate. Thus, a generalization to other convex optimal transport energies (such as, *e.g.*, the p -th power of the euclidean distance) in an appropriate dynamic context (*i.e.*, where the dynamic formulation involves a continuity type equation), would be interesting and seems to be possible. On the other hand, the regularized problems involve, in addition to the transport energy, a Radon-norm term which corresponds to a penalization of the total mass. Also here, a generalization to other regularization functionals should be possible, provided that one can still ensure boundedness of the total mass. This way, it might be possible to impose, *e.g.*, spatial smoothness of the solution curve.

Finally, we would like to mention that a numerical optimization algorithm for the solution of the regularized problem is currently under preparation, in the general setting and also with focus on the application to dynamic MRI. The numerical solution to the dynamic optimal transport problem outside the inverse problems context is addressed for example in [8, 21, 50], where the authors employ proximal splitting algorithms, after a careful discretization of the continuity equation constraint by means of staggered grids, and, subsequently, of the energy. The recent application to dynamic PET imaging [60] also builds on these algorithms. The method we propose, based on Frank-Wolfe-type algorithms [15, 29, 34], is inherently discretization-free and therefore genuinely cast in the space of Radon measures, in the spirit of [16, 53]. Generally, such algorithms are attractive for obtaining sparse solutions for inverse problems in terms of extremal points of the regularizer [9, 10]. In our setting, the idea is to linearize the fidelity term in (1.3) around some initial guess, and then proceed to minimize (a suitably modified version) of the functional obtained. The key part of the analysis is that such linearized problem admits a solution which is an extremal point of the unit ball of the regularizer, making the minimization problem numerically accessible, although non-convex in general. Indeed, as shown in [12], the extremal points of the Benamou–Brenier regularizer are given by measures concentrated on curves (with a certain regularity), yielding an optimization problem in some Sobolev space. It is then possible to show convergence of the algorithm in the space of measures. Such an approach is then particularly well-suited to recover sparse solutions, given by travelling Dirac deltas with the addition of noise, making it attractive for medical applications in which the tracking of point-sources is relevant. With this approach one can, in principle, deal also with the unbalanced optimal transport case. The key ingredient is of course the characterization of the extremal points of the Wasserstein–Fisher–Rao energy [11]. Such characterization bases on a measure-theoretic superposition principle for the non-homogeneous continuity equation, which is in itself a novel result.

APPENDIX A. MEASURE THEORY

A.1. Measure theory preliminaries

In this paper we follow the definitions and notations of [5]. In particular, scalar or vectorial measures will always be defined on the Borel σ -algebra $\mathcal{B}(X)$ of some locally compact, separable metric space X . Given a measure μ , we denote with $|\mu|$ its *total variation*. We always assume that $|\mu|$ is at least locally finite. The set of \mathbb{R}^m -valued measures for which $|\mu|(X) < \infty$ is denoted by $\mathcal{M}(X; \mathbb{R}^m)$ and $\mathcal{M}(X) := \mathcal{M}(X; \mathbb{R})$, while the set of positive finite measures is denoted by $\mathcal{M}^+(X)$.

A.1.1. Absolute continuity, support and restriction

Given $\mu \in \mathcal{M}(X)$ and $\nu \in \mathcal{M}(X; \mathbb{R}^m)$, we say that ν is *absolutely continuous* with respect to μ , in symbols $\nu \ll \mu$, if $|\nu|(E) = 0$ whenever $\mu(E) = 0$, $E \in \mathcal{B}(X)$. If μ and σ are real or vector valued measures, we say that they are mutually singular, $\mu \perp \nu$, if there exists a set $E \in \mathcal{B}(X)$ such that $|\mu|(E) = 0$ and $|\nu|(X \setminus E) = 0$. For a measure μ its *support*, denoted by $\text{supp } \mu$, is the closure of the set of points $x \in X$ such that $|\mu|(U) > 0$ for every neighbourhood U of x . If $A \in \mathcal{B}(X)$ and $\mu \in \mathcal{M}(X; \mathbb{R}^m)$, the *restriction* of μ to A is the measure $\mu \llcorner A$ defined as $\mu \llcorner A(E) := \mu(A \cap E)$ for every $E \in \mathcal{B}(X)$.

A.1.2. Push-forward

Let Y be a locally compact, separable metric space. A map $f: X \rightarrow Y$ is *Borel* if $f^{-1}(E) \in \mathcal{B}(X)$ for each $E \in \mathcal{B}(Y)$. If μ is a real or vector valued measure on X we define the push-forward of μ through f as the measure $f_{\#}\mu$ on Y , defined by $f_{\#}\mu(E) := \mu(f^{-1}(E))$ for each $E \in \mathcal{B}(Y)$. If φ is a real or vector valued map on Y integrable with respect to $f_{\#}\mu$ then $\int_Y \varphi d(f_{\#}\mu) = \int_X \varphi \circ f d\mu$. We recall that if f is continuous and *proper* ($f^{-1}(K)$ is compact if $K \subset Y$ is compact) and μ is finite, then also $f_{\#}\mu$ is finite.

A.1.3. Convergences

Let $\{\mu_n\}$ be a sequence of measures on X . We say that μ_n *narrowly converges* to μ , in symbols $\mu_n \rightarrow \mu$, if $\int_X \varphi d\mu_n \rightarrow \int_X \varphi d\mu$ for all $\varphi \in C_b(X)$, i.e., φ continuous and bounded. We say that μ_n *weak* converges* to μ , in symbols $\mu_n \xrightarrow{*} \mu$, if $\int_X \varphi d\mu_n \rightarrow \int_X \varphi d\mu$ for all $\varphi \in C_0(X)$. We recall that $\varphi \in C_0(X)$ if $\varphi \in C(X)$ and for each $\varepsilon > 0$ there exists $K \subset X$ compact such that $|\varphi| < \varepsilon$ in $X \setminus K$. Note that if X is compact, then narrow convergence and weak* convergence coincide.

A.1.4. Disintegration

Let X, Y be locally compact, separable metric spaces. Consider $\{\mu_x : x \in X\}$, a family of measures on Y . We say that the family $\{\mu_x\}$ is *Borel* if the map $x \in X \mapsto \mu_x(B)$ is Borel measurable for every $B \subset Y$ measurable. Such condition implies that for every bounded Borel function $\varphi: X \times Y \rightarrow \mathbb{R}$, the map $x \in X \mapsto \int_Y \varphi(x, y) d\mu_x(y)$ is Borel measurable (see [5], Prop. 2.26). We now state the disintegration theorem. For this version and the following properties, see Sections 2.3 and 2.4 of [2] and Theorem 2.28 of [5].

Theorem A.1 (Disintegration). *Let X, Y be locally compact, separable metric spaces. Let μ be a real (resp. vector valued) measure on $X \times Y$, $\pi: X \times Y \rightarrow X$ the projection on the first factor, and ν a positive measure on X , with the property that $\pi_{\#}|\mu| \ll \nu$. Then there exists a Borel family $\{\mu_x : x \in X\}$ of real (resp. vector valued) measures on Y such that $\mu = \nu \otimes \mu_x$, that is,*

$$\int_{X \times Y} f(x, y) d\mu(x, y) = \int_X \left(\int_Y f(x, y) d\mu_x(y) \right) d\nu(x),$$

for every $f \in L^1(X \times Y; |\mu|)$. A family $\{\mu_x\}$ such that $\mu = \nu \otimes \mu_x$ is called a *disintegration* of μ with respect to ν .

In the setting of Theorem A.1 the following properties hold:

- (i) If $\{\tilde{\mu}_x\}$ is another disintegration of μ with respect to ν , then $\mu_x = \tilde{\mu}_x$ for ν -a.e. x .
- (ii) If μ is finite, then also μ_x is finite for ν -a.e. x .
- (iii) Let $E \in \mathcal{B}(X)$, $F \in \mathcal{B}(Y)$. Then $\mu(E \times F) = 0$ if and only if $\mu_x(F) = 0$ for ν -a.e. $x \in E$.

A.2. Narrow continuity results

As in Section 2.1 we denote by $C_w([0, 1]; \mathcal{M}(\bar{\Omega}))$ the set of narrowly continuous curves $t \mapsto \rho_t$ and by $C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$ the set of positive narrowly continuous curves. The remaining notations are the same as in Section 2.

Lemma A.2. *Let $\rho_t \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$, $\rho := dt \otimes \rho_t$, $m = dt \otimes (v_t \rho_t)$, $\mu = dt \otimes (g_t \rho_t)$, with $v_t : X \rightarrow \mathbb{R}^d$, $g_t : X \rightarrow \mathbb{R}$ measurable. Assume that $\partial_t \rho + \operatorname{div} m = \mu$ in the sense of (2.2) and*

$$E := \int_X (|v_t(x)|^2 + |g_t(x)|^2) d\rho_t(x) dt < \infty.$$

Set $m := \min_{t \in [0, 1]} \rho_t(\bar{\Omega})$, $M := \max_{t \in [0, 1]} \rho_t(\bar{\Omega})$ and $C := 4(m + E)$. Then $M \leq C$ and

$$\left| \int_{\bar{\Omega}} \varphi (d\rho_t - d\rho_s) \right| \leq \|\varphi\|_{C^1} \sqrt{2CE} |t - s|^{1/2}, \quad \text{for all } \varphi \in C^1(\bar{\Omega}), \quad 0 \leq s \leq t \leq 1.$$

We remark that the above lemma is an easy generalization of Lemma 2.2 in [42], where the authors prove the same result under the restriction that $v_t = \nabla_x g_t$.

Proof. Since $\rho_t \in C_w([0, 1]; \mathcal{M}^+(\bar{\Omega}))$, we have $m, M < \infty$. Arguing as in the proof of Proposition 2.4 below, we obtain that for all $\varphi \in C^1(\bar{\Omega})$, the weak derivative of $\rho_t(\varphi) := \int_{\bar{\Omega}} \varphi d\rho_t$ satisfies

$$\rho'_t(\varphi) = \int_{\bar{\Omega}} (\nabla \varphi(x) \cdot v_t(x) + \varphi(x) g_t(x)) d\rho_t(x).$$

In particular by applying twice the Cauchy–Schwarz inequality we get

$$\begin{aligned} |\rho_t(\varphi) - \rho_s(\varphi)| &= \left| \int_s^t \int_{\bar{\Omega}} (\nabla \varphi(x) \cdot v_\lambda(x) + \varphi(x) g_\lambda(x)) d\rho_\lambda(x) d\lambda \right| \\ &\leq \sqrt{M} \left(\int_s^t \int_{\bar{\Omega}} |\nabla \varphi(x) \cdot v_\lambda(x) + \varphi(x) g_\lambda(x)|^2 d\rho_\lambda(x) d\lambda \right)^{1/2} |t - s|^{1/2} \\ &\leq \|\varphi\|_{C^1} \sqrt{2ME} |t - s|^{1/2}. \end{aligned}$$

The proof is concluded if we show that $M \leq C$. By applying the above estimate to $\varphi \equiv 1$ we get $|\rho_t(\bar{\Omega}) - \rho_s(\bar{\Omega})| \leq \sqrt{2ME}$. If we pick s and t such that $\rho_s(\bar{\Omega}) = m$ and $\rho_t(\bar{\Omega}) = M$, then we get $M \leq m + \sqrt{2ME} \leq m + \sqrt{2}(M + E)/2$, from which follows $M \leq 4(m + E) = C$. \square

Proposition A.3. *If $\rho_t \in C_w([0, 1]; \mathcal{M}(\bar{\Omega}))$ then $\sup_{t \in [0, 1]} \|\rho_t\|_{\mathcal{M}(\bar{\Omega})} < \infty$.*

Proof. The curve ρ_t defines a family $\{\rho_t\}_{t \in [0, 1]}$ of functionals in $\mathcal{M}(\bar{\Omega})$, via $\varphi \mapsto \int_{\bar{\Omega}} \varphi d\rho_t$. By narrow continuity, the map $t \mapsto \int_{\bar{\Omega}} \varphi d\rho_t$ is continuous for each $\varphi \in C(\bar{\Omega})$, yielding that $\sup_{t \in [0, 1]} \int_{\bar{\Omega}} \varphi d\rho_t < \infty$. The principle of uniform boundedness then implies the thesis. \square

Proposition A.4 (A refined version of Ascoli–Arzelà’s theorem). *Let $K \subset \mathcal{M}(\bar{\Omega})$ be sequentially weak* compact and $\rho_t^n : [0, 1] \rightarrow \mathcal{M}(\bar{\Omega})$ be such that $\rho_t^n \in K$ for all $t \in [0, 1]$, $n \in \mathbb{N}$ and*

$$\limsup_n \|\rho_t^n - \rho_s^n\|_{C^1(\bar{\Omega})^*} \leq \omega(t, s) \quad \text{for every } t \in [0, 1],$$

where $\omega : [0, 1] \times [0, 1] \rightarrow [0, \infty)$ is continuous, symmetric and such that $\omega(t, t) = 0$ for every $t \in [0, 1]$. Then there exists a $C^1(\bar{\Omega})^*$ -continuous curve $\rho_t : [0, 1] \rightarrow \mathcal{M}(\bar{\Omega})$ and a subsequence $\{\rho_t^{n_k}\}$ such that $\rho_t^{n_k} \xrightarrow{*} \rho_t$ for every $t \in [0, 1]$.

The above statement is a particular case of Proposition 3.3.1 from [6], since $(\mathcal{M}(\overline{\Omega}), \|\cdot\|_{C^1(\overline{\Omega})^*})$ is a metric space and the $C^1(\overline{\Omega})^*$ -norm is weak* sequentially lower semicontinuous.

Proof of Proposition 2.4. Let $a \in C_c^\infty((0, 1))$, $b \in C^\infty(\overline{\Omega})$. Set $\varphi(t, x) := a(t)b(x)$, so that φ is a test function for (2.2). Define $\rho_t(b) := \int_{\overline{\Omega}} b(x) d\rho_t(x)$. Notice that $t \mapsto \rho_t(b)$ is measurable since ρ_t is a Borel family. Moreover $t \mapsto \rho_t(b)$ belongs to $L^1((0, 1))$ since $\int_0^1 |\rho_t(b)| dt \leq \|b\|_{C^1(\overline{\Omega})} \|\rho\|_{\mathcal{M}(X)}$ where $\|b\|_{C^1(\overline{\Omega})} := \max\{\|b\|_\infty, \|\nabla b\|_\infty\}$. Testing (2.2) against φ yields that $\rho'_t(b) = \int_{\overline{\Omega}} [\nabla b(x) \cdot v_t(x) + b(x)g_t(x)] d\rho_t(x)$ weakly. Notice that for a.e. $t \in [0, 1]$ we have

$$|\rho'_t(b)| \leq \|b\|_{C^1(\overline{\Omega})} V(t), \quad V(t) := \int_{\overline{\Omega}} (|v_t(x)| + |g_t(x)|) d\rho_t(x). \tag{A.1}$$

In particular $V \in L^1((0, 1))$ by assumption (2.3), so that $\rho_t(b) \in W^{1,1}((0, 1))$ with

$$\|\rho_t(b)\|_{W^{1,1}((0,1))} \leq \|b\|_{C^1(\overline{\Omega})} \left(\|\rho\|_{\mathcal{M}(X)} + \int_0^1 V(t) dt \right). \tag{A.2}$$

By the embedding $W^{1,1}((0, 1)) \hookrightarrow C([0, 1])$, there exists a unique $\tilde{\rho}_t(b) \in C([0, 1])$ such that $\tilde{\rho}_t(b) = \rho_t(b)$ for a.e. $t \in [0, 1]$, and $\|\tilde{\rho}_t(b)\|_\infty \leq C \|b\|_{C^1(\overline{\Omega})}$ where $C > 0$ does not depend on b , thanks to (A.2). Moreover

$$\tilde{\rho}_t(b) - \tilde{\rho}_s(b) = \int_s^t \rho'_\lambda(b) d\lambda \quad \text{for all } s, t \in [0, 1]. \tag{A.3}$$

By density of $C^\infty(\overline{\Omega})$ in $C^1(\overline{\Omega})$, for each $t \in [0, 1]$ the map $b \mapsto \tilde{\rho}_t(b)$ can be uniquely extended to an element of $C^1(\overline{\Omega})^*$, since the maps $b \mapsto \rho_t(b)$ are linear and the extensions $\tilde{\rho}_t(b)$ are unique. This defines a bounded curve $t \mapsto \tilde{\rho}_t$ in $C^1(\overline{\Omega})^*$. Such curve is uniformly continuous in $[0, 1]$, since (A.1) and (A.3) imply $|\tilde{\rho}_t(b) - \tilde{\rho}_s(b)| \leq \|b\|_{C^1(\overline{\Omega})} \int_s^t V(\lambda) d\lambda$ for every $b \in C^1(\overline{\Omega})$. We are left to prove that $\tilde{\rho}_t \in \mathcal{M}^+(\overline{\Omega})$ for each $t \in [0, 1]$. This follows from the fact that there is a Borel set $E \subset [0, 1]$ with $|[0, 1] \setminus E| = 0$ such that $\rho_t \in \mathcal{M}^+(\overline{\Omega})$ for every $t \in E$ and $\{\rho_t\}_{t \in E}$ is weak* sequentially precompact in $\mathcal{M}^+(\overline{\Omega})$, since by Proposition 2.2 we have that $\{\rho_t(\overline{\Omega})\}$ is a.e. bounded as $t \mapsto \rho_t(\overline{\Omega})$ belongs to $BV((0, 1))$. The weak* continuity of the curve $t \mapsto \tilde{\rho}_t$ in $\mathcal{M}^+(\overline{\Omega})$ automatically follows from the one in $C^1(\overline{\Omega})^*$: indeed let $\varphi \in C(\overline{\Omega})$ and let $\{\varphi_n\}$ be a sequence in $C^1(\overline{\Omega})$ such that $\|\varphi_n - \varphi\|_\infty \rightarrow 0$ as $n \rightarrow \infty$. Then it is immediate to check that $\sup_{t \in [0, 1]} |\tilde{\rho}_t(\varphi_n) - \tilde{\rho}_t(\varphi)| \rightarrow 0$ so that $t \mapsto \tilde{\rho}_t(\varphi)$ is continuous, since it is uniform limit of continuous maps $t \mapsto \tilde{\rho}_t(\varphi_n)$. Finally let $\varphi \in C_c^1([0, 1] \times \overline{\Omega})$, $0 \leq t_1 \leq t_2 \leq 1$ and define $\varphi_\varepsilon(t, x) := a_\varepsilon(t)\varphi(t, x)$, where $a_\varepsilon \in C_c^\infty((t_1, t_2))$ is such that $0 \leq a_\varepsilon(t) \leq 1$, $\lim_{\varepsilon \rightarrow 0} a_\varepsilon(t) = \chi_{(t_1, t_2)}(t)$ for almost every $t \in [0, 1]$ and $\lim_{\varepsilon \rightarrow 0} a'_\varepsilon = \delta_{t_1} - \delta_{t_2}$ weakly* in $\mathcal{M}([0, 1])$. Testing (2.2) against φ_ε and passing to the limit as $\varepsilon \rightarrow 0$ (by continuity of $t \mapsto \int_{\overline{\Omega}} \varphi(t, x) d\tilde{\rho}_t(x)$ and (2.3)) yields (2.4). \square

APPENDIX B. TIME-DEPENDENT BOCHNER SPACES

In this appendix we assume (H1)–(H3) as in Section 3.1. Definitions of step functions, strong measurability, weak measurability, separably valued and integrability are as in Sections 3.2 and 3.3.

B.1. Auxiliary results and proofs of Section 3

Here we state and prove a suitable version of Egoroff’s theorem, as well as present the proof of Theorems 3.5, 3.10, and 3.13.

Proposition B.1 (Egoroff). *Let $f_n, f: [0, 1] \rightarrow H$ be strongly measurable and such that, for a.e. $t \in [0, 1]$, $\lim_n \|f_n(t) - f(t)\|_{H_t} = 0$. Then for each fixed $\varepsilon > 0$ there exists a Lebesgue measurable set $E \subset [0, 1]$ with $|E| < \varepsilon$ and such that $f_n \rightarrow f$ uniformly in $[0, 1] \setminus E$, that is,*

$$\lim_n \sup_{t \in [0, 1] \setminus E} \|f_n(t) - f(t)\|_{H_t} = 0. \tag{B.1}$$

Proof. The proof follows by replacing absolute values with the H_t norms in the proof of the classic Egoroff theorem. Indeed, since $f_n, f: [0, 1] \rightarrow H$ are assumed to be strongly measurable, the map $t \mapsto \|f_n(t) - f(t)\|_{H_t}$ is Lebesgue measurable (see Rem. 3.4). Then the sets $E_k(n) := \cup_{m \geq k} \{t \in [0, 1] : \|f_m(t) - f(t)\|_{H_t} \geq 1/n\}$ are measurable for each fixed $n, k \in \mathbb{N}$. Moreover, for n fixed, we have that $E_{k+1}(n) \subset E_k(n)$ and $|E_k(n)| \searrow 0$ as $k \rightarrow \infty$, since we are assuming that $f_n \rightarrow f$ a.e. in $[0, 1]$. Let $\{k_n\}$ be an increasing sequence of indices such that $|E_{k_n}(n)| < \varepsilon/2^n$. It is immediate to see that the measurable set $E := \cup_{n=1}^\infty E_{k_n}(n)$ satisfies (B.1). \square

Proof of Theorem 3.5. Part 1. Assume that f is strongly measurable. Hence there exists a sequence $\{f_n\}$ of step functions $f_n: [0, 1] \rightarrow D$ with $f_n = \sum_{j=1}^{N_n} \chi_{E_{n,j}} \varphi_{n,j}$ such that $\|i_t f_n(t) - f(t)\|_{H_t} \rightarrow 0$ a.e. in $[0, 1]$. We claim that f is weakly measurable: Indeed fix $\varphi \in D$ and define $\theta(t) := \langle i_t \varphi, f(t) \rangle_{H_t}$, $\theta_n(t) := \langle i_t \varphi, i_t f_n(t) \rangle_{H_t}$ for $t \in [0, 1]$. Clearly θ_n is measurable for fixed n , by (H3). Moreover using Cauchy–Schwarz and (H1) yields $|\theta_n(t) - \theta(t)| \leq C \|\varphi\|_D \|i_t f_n(t) - f(t)\|_{H_t}$, so that $\theta_n \rightarrow \theta$ for a.e. $t \in [0, 1]$, implying that θ is measurable, and hence f is weakly measurable. We will now show that f is essentially separably valued. By definition $i_t f_n$ is strongly measurable and $i_t f_n \rightarrow f$ a.e. hence Proposition B.1 implies that for every $n \in \mathbb{N}$ there exists a measurable set $E_n \subset [0, 1]$ with $|E_n| \leq 1/n$ and such that $\|i_t f_n(t) - f(t)\|_{H_t} \rightarrow 0$ uniformly on $[0, 1] \setminus E_n$. Define the countable set $S := \cup_{n=1}^\infty f_n([0, 1]) \subset D$. Let $E := \cap_{n=1}^\infty E_n$ so that $|E| = 0$. Fix $\varepsilon > 0$ and $t \in [0, 1] \setminus E$. Hence there exists an index $n \in \mathbb{N}$ such that $t \in [0, 1] \setminus E_n$. By uniform convergence we conclude that $\|i_t f_n(t) - f(t)\|_{H_t} < \varepsilon$, for sufficiently large n . Therefore Definition 3.2 (iii) is satisfied by setting $\varphi := f_n(t)$.

Part 2. Let f be weakly measurable and essentially separably valued. Let $S = \{\varphi_n\} \subset D$ be countable and $E \subset [0, 1]$ with $|E| = 0$ satisfying Definition 3.2. For $t \in [0, 1]$ define $\psi_n(t) := 1/\|i_t \varphi_n\|_{H_t}$ if $i_t \varphi_n \neq 0$ and $\psi_n(t) := 0$ otherwise. Notice that ψ_n is Lebesgue measurable for each fixed n , since $t \mapsto \|i_t \varphi_n\|_{H_t}$ is measurable by (H3). We will now show that

$$\|f(t)\|_{H_t} = \sup_n \psi_n(t) |\langle i_t \varphi_n, f(t) \rangle_{H_t}| \quad \text{for all } t \in [0, 1] \setminus E. \tag{B.2}$$

Indeed, the supremum never exceeds $\|f(t)\|_{H_t}$ by the Cauchy–Schwarz inequality. Conversely, if $f(t) = 0$ the equality is trivial, hence assume $f(t) \neq 0$. Fix $0 < \varepsilon < \|f(t)\|_{H_t}/2$. Since f is essentially separably valued, there exists $\varphi_n \in S$ such that $\|i_t \varphi_n - f(t)\|_{H_t} < \varepsilon$. In particular $i_t \varphi_n \neq 0$. Then

$$\begin{aligned} \|f(t)\|_{H_t} &< \varepsilon + \|i_t \varphi_n\|_{H_t} = \varepsilon + \psi_n(t) |\langle i_t \varphi_n, i_t \varphi_n \rangle_{H_t}| \\ &\leq \varepsilon + \psi_n(t) |\langle i_t \varphi_n, i_t \varphi_n - f(t) \rangle_{H_t}| + \psi_n(t) |\langle i_t \varphi_n, f(t) \rangle_{H_t}| \\ &\leq 2\varepsilon + \psi_n(t) |\langle i_t \varphi_n, f(t) \rangle_{H_t}| \leq 2\varepsilon + \sup_n \psi_n(t) |\langle i_t \varphi_n, f(t) \rangle_{H_t}|, \end{aligned}$$

and since ε is arbitrarily small we conclude. Notice that the map $t \mapsto |\langle i_t \varphi_n, f(t) \rangle_{H_t}|$ is measurable by weak measurability of f . Thus $t \mapsto \psi_n(t) |\langle i_t \varphi_n, f(t) \rangle_{H_t}|$ is measurable, being product of measurable maps. Since the countable pointwise supremum of measurable functions is measurable, by (B.2) we conclude that $t \mapsto \|f(t)\|_{H_t}$ is measurable. Also the map $\theta_n(t) := \|f(t) - i_t \varphi_n\|_{H_t}$ is measurable at n fixed, as

$$\|f(t) - i_t \varphi_n\|_{H_t}^2 = \|f(t)\|_{H_t}^2 - 2\langle i_t \varphi_n, f(t) \rangle_{H_t} + \|i_t \varphi_n\|_{H_t}^2$$

is a sum of measurable functions, where the second element is measurable by weak measurability of f and the third by (H3). Fix $\varepsilon > 0$ and define the measurable sets $E_n := \{t \in [0, 1] : \theta_n(t) < \varepsilon\}$ and the map $g: [0, 1] \rightarrow D$ by setting $g(t) := \varphi_n$ if $t \in E_n \setminus \cup_{j=1}^{n-1} E_j$ for some n , and $g(t) = 0$ otherwise. Note that for $t \in [0, 1] \setminus E$ there exists some index n such that $\|f(t) - i_t \varphi_n\|_{H_t} < \varepsilon$. Therefore, by picking the smallest n such that this condition is verified, we have $g(t) = \varphi_n$. Since this is true for each $t \in [0, 1] \setminus E$, this means that $\|f(t) - i_t g(t)\|_{H_t} < \varepsilon$ a.e. in $[0, 1]$. Hence we can approximate f essentially uniformly by $i_t g(t)$ with g countably valued. By choosing $\varepsilon = 1/n$ we obtain countably valued functions $g_n: [0, 1] \rightarrow D$ such that

$$\|f(t) - i_t g_n(t)\|_{H_t} < 1/n \quad \text{for every } t \in E, \tag{B.3}$$

where $|[0, 1] \setminus E| = 0$. Note that by definition $g_n = \sum_{j=1}^\infty \chi_{E_{n,j}} \varphi_{n,j}$ with $\{E_{n,j}\}_{j \in \mathbb{N}}$ measurable partition of $[0, 1]$. Therefore for every $n \in \mathbb{N}$, there exists k_n such that the set $\cup_{j=1}^{k_n} E_{n,j}$ satisfies

$$\left| [0, 1] \setminus \cup_{j=1}^{k_n} E_{n,j} \right| < \frac{1}{n^2}. \tag{B.4}$$

Set $F_n := \cap_{s=n}^\infty \cup_{j=1}^{k_s} E_{s,j}$, $F := \cup_{n=1}^\infty F_n$. In this way $|[0, 1] \setminus F| = 0$ by (B.4), since $|[0, 1] \setminus F_n| \leq \sum_{s=n}^\infty \frac{1}{s^2} \rightarrow 0$ as $n \rightarrow \infty$. Now define step functions $f_n: [0, 1] \rightarrow D$ obtained by truncating g_n , that is, $f_n := \sum_{j=1}^{k_n} \chi_{E_{n,j}} \varphi_{n,j}$. If we prove that

$$\lim_n \|f(t) - i_t f_n(t)\|_{H_t} = 0 \quad \text{for every } t \in F \cap E, \tag{B.5}$$

then we conclude that f is strongly measurable, since $|[0, 1] \setminus (F \cap E)| = 0$. In order to show (B.5), fix $\varepsilon > 0$ and $t \in F \cap E$. By using (B.3) and (H1) we have that for all $n \in \mathbb{N}$

$$\begin{aligned} \|f(t) - i_t f_n(t)\|_{H_t} &\leq \|f(t) - i_t g_n(t)\|_{H_t} + \|i_t g_n(t) - i_t f_n(t)\|_{H_t} \\ &< \frac{1}{n} + C \left\| \sum_{j=k_n+1}^\infty \chi_{E_{n,j}}(t) \varphi_{n,j} \right\|_D. \end{aligned} \tag{B.6}$$

Since $t \in F$, by definition there exists an index N_t such that $t \in F_{N_t}$. Hence $t \in \cup_{j=1}^{k_n} E_{n,j}$ for every $n \geq N_t$, so that $\sum_{j=k_n+1}^\infty \chi_{E_{n,j}}(t) \varphi_{n,j} = 0$ for each $n \geq N_t$. Set $n_{\varepsilon,t} := \max\{N_t, 1/\varepsilon\}$. From (B.6) we have $\|f(t) - i_t f_n(t)\|_{H_t} < \varepsilon$ for every $n \geq n_{\varepsilon,t}$, implying (B.5). \square

Proof of Theorem 3.10. Since $f_n \rightarrow f$ a.e. the map f is strongly measurable and $\theta_n(t) := \|f_n(t) - f(t)\|_{H_t}$ is Lebesgue measurable. By assumption we have that $\theta_n \rightarrow 0$ and $\theta_n \leq 2g$ a.e. in $[0, 1]$. Therefore, by the classic dominated convergence theorem, each θ_n is integrable and $f_n \rightarrow f$ strongly in $L^1([0, 1]; H)$. To conclude integrability for f it is sufficient to employ triangle inequality, integrability of θ_n and Theorem 3.9. \square

Proof of Theorem 3.13. The fact that $\|\cdot\|_{L^p}$ is a norm follows immediately from the classic case as well as the fact that the map $t \mapsto \|f(t)\|_{H_t}^p$ is measurable for each $p \geq 1$, when f is assumed to be strongly measurable (see Rem. 3.4). Moreover $\langle \cdot, \cdot \rangle_{L^2}$ is an inner product, since the spaces H_t are Hilbert. In order to show completeness, it is sufficient to follow the lines of the proof of the classic Riesz–Fischer theorem. Let $1 \leq p < \infty$ and let $\{f_n\}$ be a Cauchy sequence in $L^p([0, 1]; H)$. In any normed linear space, a Cauchy sequence having a convergent subsequence converges to the same limit. Therefore, up to extracting a subsequence, we can assume that

$$\|f_m - f_n\|_{L^p} < \frac{1}{2^n} \quad \text{for every } n \text{ and } m > n. \tag{B.7}$$

For every $n \in \mathbb{N}$ define measurable sets $E_n := \{t \in [0, 1] : \|f_{n+1}(t) - f_n(t)\|_{H_t} \geq 1/n^2\}$, so that $\chi_{E_n}(t)/n^2 \leq \|f_{n+1}(t) - f_n(t)\|_{H_t}$ a.e. in $[0, 1]$ and $|E_n| < n^{2p}/2^{np}$ by (B.7). In particular one has $\sum_n |E_n| < \sum_n n^{2p}/2^{np} < \infty$. Define $F_n := \cup_{m \geq n} E_m$, so that $\{F_n\}$ is a nested sequence of measurable sets, with $|F_n| \leq \sum_{m \geq n} m^{2p}/2^{mp} \rightarrow 0$ as $n \rightarrow \infty$. Finally set $F := \cap_n F_n$, which satisfies $|F| = 0$. By definition, if $t \in [0, 1] \setminus F$, then $\|f_{n+1}(t) - f_n(t)\|_{H_t} < n^{-2}$ for n sufficiently large. Hence for $t \in [0, 1] \setminus F$, $m > n$ and n sufficiently large one has $\|f_m(t) - f_n(t)\|_{H_t} \leq \sum_{j \geq n} \|f_{j+1}(t) - f_j(t)\|_{H_t} < \sum_{j \geq n} 1/j^2$, and since $\sum_{j \geq n} 1/j^2 \rightarrow 0$ as $n \rightarrow \infty$, we conclude that $\{f_n(t)\}$ is a Cauchy sequence in H_t . For $t \in [0, 1] \setminus F$ denote by $f(t) \in H_t$ the strong limit of $\{f_n(t)\}$, which exists since H_t is complete. For $t \in F$ set $f(t) = 0$. This defines a map $f: [0, 1] \rightarrow H$, which is strongly measurable since it is the a.e. pointwise limit of a sequence of strongly measurable maps (see Rem. 3.4). Moreover, by the a.e. pointwise convergence, we also have that $\|f_n(t)\|_{H_t} \rightarrow \|f(t)\|_{H_t}$ as $n \rightarrow \infty$ for a.e. t . Since the maps $t \mapsto \|f_n(t)\|_{H_t}^p, t \mapsto \|f(t)\|_{H_t}^p$ are measurable, we can apply Fatou’s lemma and conclude that $\int_0^1 \|f(t)\|_{H_t}^p dt \leq \liminf_n \int_0^1 \|f_n(t)\|_{H_t}^p dt$, which is bounded since $\{f_n\}$ is a Cauchy sequence in $L^p([0, 1]; H)$. Hence $f \in L^p([0, 1]; H)$. Finally, one more application of Fatou’s lemma combined with (B.7) yields $\|f_n - f\|_{L^p} \rightarrow 0$ as $n \rightarrow \infty$. \square

B.2. Comparison with classic Bochner theory

In this section we will investigate integrability properties for $i_t^* f: [0, 1] \rightarrow D^*$ when $f \in L^1([0, 1]; H)$. Since the codomain of $i_t^* f$ is the fixed space D^* , it makes sense to check whether $i_t^* f$ is integrable in a classic sense. Specifically, in Proposition B.2, we will see that $i_t^* f$ is always Gelfand integrable. On the other hand, $i_t^* f$ is not always Bochner integrable, as we show in Example B.3. The main impediment is that $i_t^* f$ is not strongly measurable in general. Finally in Proposition B.4 we will give sufficient conditions under which $i_t^* f$ is Bochner integrable.

Proposition B.2. *Assume that $f \in L^1([0, 1]; H)$. Then $i_t^* f$ is Gelfand integrable in D^* , that is, for each measurable set $E \subset [0, 1]$ there exists an element $I_E(i_t^* f) \in D^*$ such that*

$$\langle I_E(i_t^* f), \varphi \rangle_{D^*, D} = \int_E \langle i_t^* f, \varphi \rangle_{D^*, D} dt, \quad \text{for all } \varphi \in D. \tag{B.8}$$

Proof. Let $\varphi \in D$. By duality one has $\langle i_t^* f(t), \varphi \rangle_{D^*, D} = \langle f(t), i_t \varphi \rangle_{H_t}$. Therefore $i_t^* f$ is weak* measurable since $f: [0, 1] \rightarrow H$ is weak measurable. By (H1),

$$\int_0^1 |\langle i_t^* f(t), \varphi \rangle_{D^*, D}| dt = \int_0^1 |\langle i_t \varphi, f(t) \rangle_{H_t}| dt \leq C \|\varphi\|_D \int_0^1 \|f(t)\|_{H_t} dt < \infty$$

since f is integrable. This shows that $t \mapsto \langle i_t^* f(t), \varphi \rangle_{D^*, D}$ belongs to $L^1([0, 1])$ for each $\varphi \in D$. Hence $i_t^* f$ is Gelfand integrable by Theorem 11.52 in [3], and (B.8) holds. \square

Example B.3 (Radial sampling). Let $\Omega := B_1(0) = \{x \in \mathbb{R}^2 : |x| < 1\}$ and for $t \in [0, 1]$ define the lines through the origin $S_t := \{(\cos(\pi t)s, \sin(\pi t)s) : |s| < 1\}$, so that $S_t \subset \Omega$. Define $D = C_0(\Omega)$ equipped with the supremum norm. Hence $D^* = \mathcal{M}(\Omega)$. Define $H_t := L^2_{\sigma_t}(S_t)$ with inner product $\langle h_1, h_2 \rangle_{H_t} := \int_{S_t} h_1 h_2 d\sigma_t$, $\sigma_t := \mathcal{H}^1 \llcorner S_t$ where \mathcal{H}^1 is the 1-dimensional Hausdorff measure. Finally define $i_t: D \rightarrow H_t$ by $i_t \varphi = \varphi|_{S_t}$. It is straightforward to check that (H1)–(H3) are satisfied, so that we can consider the space $L^2([0, 1]; H)$ defined as in (3.2).

We will now construct a map f belonging to $L^2([0, 1]; H)$, compute the Gelfand integral of $i_t^* f: [0, 1] \rightarrow D^*$ and show that $i_t^* f$ is not Bochner integrable. To this end, notice that for a map $h: \Omega \rightarrow \mathbb{R}$ such that $h/|\cdot| \in L^1(\Omega)$ we have that

$$\int_0^1 \int_{S_t} h d\mathcal{H}^1 dt = \int_{\Omega} \frac{h(x)}{\pi|x|} dx. \tag{B.9}$$

Note that (B.9) is an easy consequence of the classical coarea formula Theorem 3.11 of [31], and its proof is left to the reader. Let now $\tilde{f}: \Omega \rightarrow \mathbb{R}$ be such that $\tilde{f} \not\equiv 0$ and $\tilde{f}/|\cdot| \in L^2(\Omega)$. By applying (B.9) to $|\tilde{f}|^2$ and by the assumptions on \tilde{f} we have that $\tilde{f}|_{S_t}$ belongs to H_t for a.e. $t \in [0, 1]$. Define $f: [0, 1] \rightarrow H$ by $f(t) := \tilde{f}|_{S_t}$. Notice that f is strongly measurable, since \tilde{f} can be approximated in Ω by C_0 functions. Moreover by (B.9) we infer

$$\int_0^1 \|f(t)\|_{H_t}^2 dt = \int_0^1 \int_{S_t} |\tilde{f}|^2 d\mathcal{H}^1 dt = \int_{\Omega} \frac{|\tilde{f}(x)|^2}{\pi|x|} dx,$$

which is finite by assumption on \tilde{f} . Hence f is integrable by Theorem 3.9, and it belongs to $L^2([0, 1]; H)$. The Gelfand integral of $i_t^* f$, which exists by Proposition B.2, is given by

$$I(i_t^* f) = \frac{\tilde{f}}{\pi|\cdot|} d\mathcal{H}^2 \llcorner \Omega.$$

The above follows immediately by applying (B.9) to $\varphi \tilde{f}$ with $\varphi \in D$. However $i_t^* f$ is not Bochner integrable, since it is not strongly measurable in the classic sense ([27], Chap. II): For every $E \subset [0, 1]$ with $|E| = 0$ we

have that the set $i_t^* f([0, 1] \setminus E)$ is not norm separable in $\mathcal{M}(\Omega)$. Indeed, it is easy to show that for a.e. $w \neq t$

$$\|f(t)\mathcal{H}^1 \llcorner S_t - f(w)\mathcal{H}^1 \llcorner S_w\|_{\mathcal{M}(\Omega)} = \int_{S_t} |\tilde{f}| d\mathcal{H}^1 + \int_{S_w} |\tilde{f}| d\mathcal{H}^1.$$

Since $\tilde{f} \not\equiv 0$, we infer that $i_t^* f(F)$ is a discrete set for any $F \subset [0, 1]$ with $|F| > 0$. Therefore $i_t^* f(F)$ is norm separable if and only if F is countable, which is never the case. Hence $i_t^* f$ is not essentially separably valued, and the classic Pettis theorem ([27], Chap. II.1, Thm. 2) implies that $i_t^* f$ is not strongly measurable and hence not Bochner integrable.

Proposition B.4. *Assume that $i_t^* i_t(S) \subset D^*$ is essentially norm separable for each countable set $S \subset D$ and that D is reflexive. If $f \in L^1([0, 1]; H)$ then $i_t^* f: [0, 1] \rightarrow D^*$ is Bochner integrable.*

Proof. Since $f: [0, 1] \rightarrow H$ is strongly measurable, it is also weakly measurable and essentially separably valued by Theorem 3.5. We start by showing that weak measurability for f implies weak measurability for $i_t^* f$ in the classic sense. Indeed by reflexivity the canonical injection $j: D \rightarrow D^{**}$ is also surjective. Therefore for each $\varphi^{**} \in D^{**}$ there exists a unique $\varphi \in D$ with $j(\varphi) = \varphi^{**}$ and we have $\langle \varphi^{**}, i_t^* f(t) \rangle_{D^{**}, D^*} = \langle f(t), i_t \varphi \rangle_{H_t}$, which is measurable since f is weakly measurable. Now let $S = \{\varphi_n\} \subset D$ and $E \subset [0, 1]$ measurable with $|E| = 0$ and such that Definition 3.2 is satisfied. Therefore for every $\varepsilon > 0$ and $t \in [0, 1] \setminus E$ there exists $\varphi_n \in S$ such that $\|i_t \varphi_n - f(t)\|_{H_t} < \varepsilon/C$. Therefore by (H1) we can estimate $\|i_t^* i_t \varphi_n - i_t^* f(t)\|_{D^*} \leq C \|i_t \varphi_n - f(t)\|_{H_t} < \varepsilon$. Hence the points of $i_t^* f([0, 1] \setminus E)$ are arbitrarily close to $\{i_t^* i_t \varphi_n, n \in \mathbb{N}, t \in [0, 1] \setminus E\}$. Since $i_t^* i_t(S)$ is assumed to be essentially separable in D^* , there exists a measurable set $F \subset [0, 1]$ with $|F| = 0$ such that $\{i_t^* i_t \varphi_n, n \in \mathbb{N}, t \in [0, 1] \setminus F\}$ is separable in D^* . By defining $\tilde{E} := E \cup F$ we obtain that also $i_t^* f([0, 1] \setminus \tilde{E})$ is norm separable, hence $i_t^* f$ is essentially separably valued in the classic sense. By the classic Pettis theorem ([27], Chap. II.1, Thm. 2) we conclude that $i_t^* f$ is strongly measurable. Finally (H2) implies that $\int_0^1 \|i_t^* f(t)\|_{D^*} dt \leq C \int_0^1 \|f(t)\|_{H_t} dt < \infty$. By Chapter II.2 and Theorem 2 of [27], it follows that $i_t^* f$ is Bochner integrable. \square

Acknowledgements. The authors gratefully acknowledge support by the Christian Doppler Research Association (CDG) and Austrian Science Fund (FWF) through the Partnership in Research project PIR-27 “Mathematical methods for motion-aware medical imaging”. The Institute of Mathematics and Scientific Computing, to which the authors are affiliated, is a member of NAWI Graz (<http://www.nawigraz.at/en/>). The authors are further members of/associated with BioTechMed Graz (<https://biotechmedgraz.at/en/>).

REFERENCES

- [1] M. Aharon, M. Elad and A. Bruckstein, *K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation*. *IEEE Trans. Signal Process.* **54** (2006) 4311–4322.
- [2] G. Alberti, S. Bianchini and G. Crippa, *A uniqueness result for the continuity equation in two dimensions*. *J. Eur. Math. Soc.* **16** (2014) 201–234.
- [3] C.D. Aliprantis and K. Border, *Infinite Dimensional Analysis*. Springer-Verlag, Berlin, Heidelberg (2006).
- [4] A. Alphonse, C. Elliott and B. Stinner, *An abstract framework for parabolic PDEs on evolving spaces*. *Port. Math.* **72** (2015) 1–46.
- [5] L. Ambrosio, N. Fusco and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Science Publications. Clarendon Press (2000).
- [6] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*. *Lectures in Mathematics*. ETH Zürich. Birkhäuser, Basel (2006).
- [7] J.-D. Benamou, *Numerical resolution of an “unbalanced” mass transport problem*. *ESAIM: M2AN* **37** (2003) 851–868.
- [8] J.-D. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem*. *Numer. Math.* **84** (2000) 375–393.
- [9] C. Boyer, A. Chambolle, Y. De Castro, V. Duval, F. De Gournay and P. Weiss, *On representer theorems and convex regularization*. *SIAM J. Optim.* **29** (2019) 1260–1281.
- [10] K. Bredies and M. Carioni, *Sparsity of solutions for variational inverse problems with finite-dimensional data*. *Calc. Var. Part. Differ. Equ.* **59** (2019) 14.
- [11] K. Bredies, M. Carioni and S. Fanzon, *A superposition principle for the inhomogeneous continuity equation with Hellinger–Kantorovich-regular coefficients*. Preprint [arXiv:2007.06964](https://arxiv.org/abs/2007.06964) (2020).

- [12] K. Bredies, M. Carioni, S. Fanzon and F. Romero, On the extremal points of the ball of the Benamou–Brenier energy. Preprint [arXiv:1907.11589](https://arxiv.org/abs/1907.11589) (2019).
- [13] K. Bredies and M. Holler, Regularization of linear inverse problems with total generalized variation. *J. Inverse Ill-Posed Prob.* **22** (2014) 871–913.
- [14] K. Bredies, K. Kunisch and T. Pock, Total generalized variation. *SIAM J. Imaging Sci.* **3** (2010) 492–526.
- [15] K. Bredies, D.A. Lorenz and P. Maass, A generalized conditional gradient method and its connection to an iterative shrinkage method. *Comput. Optim. Appl.* **42** (2009) 173–193.
- [16] K. Bredies and H.K. Pikkariainen, Inverse problems in spaces of measures. *ESAIM: COCV* **19** (2013) 190–218.
- [17] E.J. Candès and C. Fernandez-Granda, Towards a mathematical theory of super-resolution. *Commun. Pure Appl. Math.* **67** (2014) 906–956.
- [18] E.J. Candès, J.K. Romberg and T. Tao, Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* **59** (2006) 1207–1223.
- [19] A. Chambolle and P.-L. Lions, Image recovery via total variation minimization and related problems. *Numer. Math.* **76** (1997) 167–188.
- [20] L. Chizat, G. Peyré, B. Schmitzer and F.-X. Vialard, Unbalanced optimal transport: dynamic and Kantorovich formulations. *J. Funct. Anal.* **274** (2018) 3090–3123.
- [21] L. Chizat, G. Peyré, B. Schmitzer and F.-X. Vialard, An interpolating distance between optimal transport and Fisher–Rao metrics. *Found. Comput. Math.* **18** (2018) 1–44.
- [22] M. Cuturi and G. Peyré, Computational optimal transport. *Found. Trends Mach. Learn.* **11** (2019) 355–607.
- [23] G. Dal Maso and R. Toader, On the Cauchy problem for the wave equation on time-dependent domains. *J. Differ. Equ.* **266** (2019) 3209–3246.
- [24] M. Dannhauer, E. Lämmel, C.H. Wolters and T.R. Knösche, Spatio-temporal regularization in linear distributed source reconstruction from EEG/MEG: a critical evaluation. *Brain Topogr.* **26** (2013) 229–246.
- [25] I. Daubechies, M. Defrise and C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.* **57** (2004) 1413–1457.
- [26] J. Diestel, Sequences and Series in Banach Spaces. Springer-Verlag, New York (1984).
- [27] J. Diestel and J. Uhl, Vector Measures. American Mathematical Society (1977).
- [28] D.L. Donoho, Compressed sensing. *IEEE Trans. Inform. Theory* **52** (2006) 1289–1306.
- [29] V. Duval and G. Peyré, Exact support recovery for sparse spikes deconvolution. *Found. Comput. Math.* **15** (2015) 1315–1355.
- [30] H.W. Engl, M. Hanke and A. Neubauer, Regularization of Inverse Problems. Springer (2000).
- [31] L.C. Evans and R.F. Gariépy, Measure Theory and Fine Properties of Functions. CRC Press (2015).
- [32] A. Figalli, The optimal partial transport problem. *Arch. Ratio. Mech. Anal.* **195** (2010) 533–560.
- [33] A. Figalli and N. Gigli, A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions. *J. Math. Pures Appl.* **94** (2010) 107–130.
- [34] M. Frank and P. Wolfe, An algorithm for quadratic programming. *Nav. Res. Logist. Q.* **3** (1956) 95–110.
- [35] V. Frikel, Sparse regularization in limited angle tomography. *Appl. Comput. Harmon. Anal.* **34** (2013) 117–141.
- [36] W. Gangbo, W. Li, S. Osher and M. Puthawala, Unnormalized optimal transport. *J. Comput. Phys.* **399** (2019) 108940.
- [37] M. Holler and K. Kunisch, On infimal convolution of TV-type functionals and applications to video and image reconstruction. *SIAM J. Imaging Sci.* **7** (2014) 2258–2300.
- [38] R. Hug, E. Maitre and N. Papadakis, Multi-physics optimal transportation and image interpolation. *ESAIM: M2AN* **49** (2015) 1671–1692.
- [39] J. Karlsson and A. Ringh, Generalized Sinkhorn iterations for regularizing inverse problems using optimal mass transport. *SIAM J. Imaging Sci.* **10** (2017) 1935–1962.
- [40] F. Knoll, K. Bredies, T. Pock and R. Stollberger, Second order total generalized variation (TGV) for MRI. *Magn. Reson. Med.* **65** (2011) 480–491.
- [41] F. Knoll, C. Clason, K. Bredies, M. Uecker and R. Stollberger, Parallel imaging with nonlinear reconstruction using variational penalties. *Magn. Reson. Med.* **67** (2012) 34–41.
- [42] S. Kondratyev, L. Monsaingeon and D. Vorotnikov, A new optimal transport distance on the space of finite Radon measures. *Adv. Differ. Equ.* **21** (2016) 1117–1164.
- [43] M. Liero, A. Mielke and G. Savaré, Optimal transport in competition with reaction: the Hellinger–Kantorovich distance and geodesic curves. *SIAM J. Math. Anal.* **48** (2016) 2869–2911.
- [44] M. Liero, A. Mielke and G. Savaré, Optimal Entropy-Transport problems and a new Hellinger–Kantorovich distance between positive measures. *Invent. Math.* **211** (2018) 969–1117.
- [45] S.G. Lingala, Y. Hu, E. DiBella and M. Jacob, Accelerated dynamic MRI exploiting sparsity and low-rank structure: k - t SLR. *IEEE T. Med. Imaging* **30** (2011) 1042–1054.
- [46] D. Lombardi and E. Maitre, Eulerian models and algorithms for unbalanced optimal transport. *ESAIM: M2AN* **49** (2015) 1717–1744.
- [47] J. Maas, M. Rumpf, C. Schönlieb and S. Simon, A generalized model for optimal transport of images including dissipation and density modulation. *ESAIM: M2AN* **49** (2015) 1745–1769.
- [48] L. Métivier, R. Brossier, Q. Mérigot, E. Oudet and J. Virieux, An optimal transport approach for seismic tomography: application to 3D full waveform inversion. *Inverse Prob.* **32** (2016) 115008.

- [49] R. Otazo, E. Candès and D.K. Sodickson, Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. *Magn. Reson. Med.* **73** (2015) 1125–1136.
- [50] N. Papadakis, G. Peyré and E. Oudet, Optimal transport with proximal splitting. *SIAM J. Imaging Sci.* **7** (2014) 212–238.
- [51] B. Piccoli and F. Rossi, Generalized Wasserstein distance and its application to transport equations with source. *Arch. Ratio. Mech. Anal.* **211** (2014) 335–358.
- [52] B. Piccoli and F. Rossi, On properties of the generalized Wasserstein distance. *Arch. Ratio. Mech. Anal.* **222** (2016) 1339–1365.
- [53] K. Pieper and D. Walter, Linear convergence of accelerated conditional gradient algorithms in spaces of measures. Preprint [arXiv:1904.09218](https://arxiv.org/abs/1904.09218) (2019).
- [54] K.P. Pruessmann, Encoding and reconstruction in parallel MRI. *NMR Biomed.* **19** (2006) 288–299.
- [55] L.I. Rudin, S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithms. *Phys. D* **60** (1992) 259–268.
- [56] F. Santambrogio, Optimal Transport for Applied Mathematicians. Birkhäuser, Basel (2015).
- [57] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier and F. Lenzen, Variational Methods in Imaging. Springer (2009).
- [58] M. Schloegl, M. Holler, A. Schwarzl, K. Bredies and R. Stollberger, Infimal convolution of total generalized variation functionals for dynamic MRI. *Magn. Reson. Med.* **78** (2017) 142–155.
- [59] U. Schmitt and A.K. Louis, Efficient algorithms for the regularization of dynamic inverse problems: I. Theory. *Inverse Prob.* **18** (2002) 645.
- [60] B. Schmitzer, K.P. Schäfers and B. Wirth, Dynamic cell imaging in PET with optimal transport regularization. *IEEE T. Med. Imaging* **39** (2019) 1626–1635.
- [61] T. Schuster, B. Hahn and M. Burger, Dynamic inverse problems: modelling–regularization–numerics. *Inverse Prob.* **34** (2018) 040301.
- [62] T. Schuster, B. Kaltenbacher, B. Hofmann and K.S. Kazimierski, Regularization Methods in Banach Spaces. Walter de Gruyter (2012).
- [63] A.N. Tikhonov and V. Arsenin, Solutions of Ill-Posed Problems. Wiley (1977).
- [64] A.N. Tikhonov, A.S. Leonov and A.G. Yagola, Nonlinear Ill-Posed Problems. Chapman & Hall (1998).
- [65] B. Vandeghinste, B. Goossens, R.V. Holen, C. Vanhove, A. Pižurica, S. Vandenberghe and S. Staelens, Iterative CT reconstruction using shearlet-based regularization. *IEEE Trans. Nucl. Sci.* **60** (2013) 3305–3317.
- [66] C. Villani, Optimal Transport: Old and New. Springer, Berlin, Heidelberg (2008).
- [67] J. Weickert and C. Schnörr, Variational optic flow computation with a spatio-temporal smoothness constraint. *J. Math. Imaging Vis.* **14** (2001) 245–255.