Optimal transport regularization of dynamic inverse problems

Silvio Fanzon

University of Graz

Based on joint works with

Kristian Bredies, Marcello Carioni and Francisco Romero

Paris, 20-21 November 2019

Plan of the Talk

1. Analytical framework for OT regularization of dynamic inverse problems (with K. Bredies)
   - *An optimal transport approach for solving dynamic inverse problems in spaces of measures.* (Preprint 2019)

2. Numerical results for sparse reconstruction in spaces of measures (with K. Bredies, M. Carioni, F. Romero)
   - *On the extremal points of the ball of the Benamou-Benier energy* (Preprint 2019)
   - *A generalized conditional gradient method for dynamic inverse problems with optimal transport regularization* (In preparation)
   - *A superposition principle for the non-homogeneous continuity equation* (In preparation)

Motivation: Motion-Aware Tomographic Reconstruction

Motion on sub-acquisition time scales $\rightsquigarrow$ **artefacts** in reconstructed images

- Imaging of lung or heart (motion cannot be suppressed)
- High-resolution imaging (sub-millimeter motion poses problems)

Workarounds: use of anaesthetics, breath-holding strategies, gating

Drawbacks: assumes periodicity (arrhythmias?). Still limited to low-resolution

Reference image

No regularizer

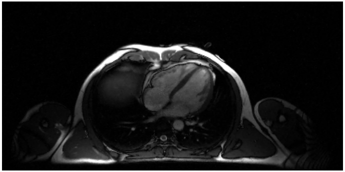
Motivation: Motion-Aware Tomographic Reconstruction

Motion on sub-acquisition time scales $\rightsquigarrow$ **artefacts** in reconstructed images

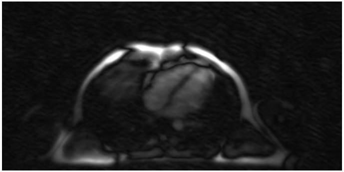
- Imaging of lung or heart (motion cannot be suppressed)
- High-resolution imaging (sub-millimeter motion poses problems)

Workarounds: use of anaesthetics, breath-holding strategies, gating

Drawbacks: assumes periodicity (arrhythmias?). Still limited to low-resolution



Reference image

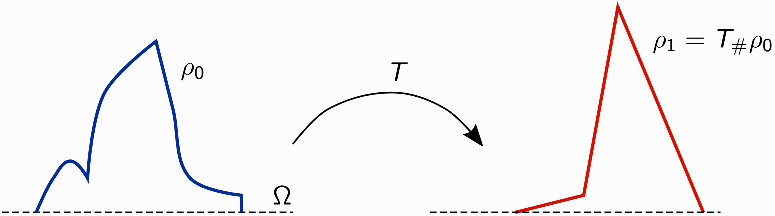


No regularizer

Proposed model: optimal transport regularization for dynamic reconstruction

Optimal Transport - Static Formulation

$\Omega \subset \mathbb{R}^d$ bounded domain, $\rho_0, \rho_1 \in \mathcal{P}(\Omega)$, $T \colon \Omega \to \Omega$ measurable displacement



Goal: move ρ_0 to ρ_1 in the cheapest way, with cost of moving mass from x to y

$$c(x,y) := |x-y|^2$$

Optimal Transport: a transport plan T solving

$$\min \left\{ \int_\Omega |T(x) - x|^2 \, d\rho_0(x) : \ T \colon \Omega \to \Omega, \ T_\# \rho_0 = \rho_1 \right\}$$

Optimal Transport - Dynamic Formulation

Idea: introduce a time variable $t \in [0,1]$ and consider evolution of ρ_t

- time dependent probability measures

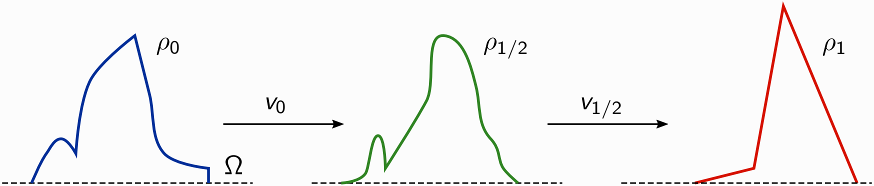
$$t \mapsto \rho_t \in \mathcal{P}(\Omega) \quad \text{for} \quad t \in [0,1]$$

- velocity field advecting ρ_t

$$v_t(x) \colon [0,1] \times \Omega \to \mathbb{R}^d$$

- (ρ_t, v_t) solves the continuity equation with initial conditions

$$\begin{cases} \partial_t \rho_t + \operatorname{div}(\rho_t v_t) = 0 \\ \text{Initial data } \rho_0, \text{ final data } \rho_1 \end{cases} \qquad \text{(CE-IC)}$$



Connection and Advantages

Theorem (Benamou-Brenier '00)

$$\min_{\substack{(\rho_t, v_t) \\ \textit{solving (CE-IC)}}} \int_0^1 \int_\Omega |v_t(x)|^2 \rho_t(x) dx\, dt = \min_{\substack{T:\, \Omega \to \Omega \\ T_\# \rho_0 = \rho_1}} \int_\Omega |T(x) - x|^2 \rho_0(x)\, dx$$

Advantages of Dynamic Formulation:

1. By introducing the momentum $m_t := \rho_t v_t$ we have

$$\int_0^1 \int_\Omega |v_t(x)|^2 \rho_t(x)\, dx\, dt = \int_0^1 \int_\Omega \frac{|m_t(x)|^2}{\rho_t(x)}\, dx\, dt$$

   which is **convex** in (ρ_t, m_t). The continuity equation becomes **linear**

$$\partial_t \rho_t + \operatorname{div} m_t = 0$$

2. we know the full trajectory ρ_t and can recover the velocity field v_t from m_t

Dynamic inverse problem

$\Omega \subset \mathbb{R}^d$ bounded open domain, $d \geq 1$

For $t \in [0,1]$ assume given

- H_t Hilbert spaces (measurement spaces - non isomorphic)
- $K_t^* \colon \mathcal{M}(\overline{\Omega}) \to H_t$ linear continuous operators (forward-operators)

(Time dependence allows for spatial undersampling - e.g. line or point sampling)

Problem

Given some data $\{f_t\}_{t\in[0,1]}$ with $f_t \in H_t$, find a curve of measures

$$t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$$

such that they solve the dynamic inverse problem

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

Unbalanced optimal transport regularization

Consider a triple $(\rho_t,\, v_t,\, g_t)$ with

- $t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ mass density (not probability measures)
- $v_t \colon (0,1) \times \overline{\Omega} \to \mathbb{R}^d$ velocity field, $g_t \colon (0,1) \times \overline{\Omega} \to \mathbb{R}$ growth rate

We propose to regularize $K_t^* \rho_t = f_t$ via minimization in $(\rho_t,\, v_t,\, g_t)$ of

$$\underbrace{\frac{1}{2}\int_0^1 \|K_t^*\rho_t - f_t\|_{H_t}^2 \, dt}_{\text{Fidelity Term}} + \frac{\alpha}{2}\underbrace{\int_0^1 \int_{\overline{\Omega}} |v_t(x)|^2 + |g_t(x)|^2 \, d\rho_t(x) dt}_{\text{Optimal Transport Regularizer}} + \beta \underbrace{\int_0^1 \|\rho_t\| \, dt}_{\text{TV Regularizer}}$$

$$\text{s.t.} \quad \partial_t \rho + \operatorname{div}(\rho v_t) = \rho g_t \qquad \text{(Continuity Equation - No IC)}$$

- v_t keeps track of motion, g_t keeps track of contrast agent
- continuity equation enforces time “regularity”

Chizat, Peyré, Schmitzer, Vialard (Found. of Comp. Math. ’18, JFA ’18)

Liero, Mielke, Savaré (Inv. Math. ’18)

Formal definition of the Unbalanced OT Energy

Set $X := (0,1) \times \overline{\Omega}$ and consider triples $(\rho, m, \mu) \in \mathcal{M}(X) \times \mathcal{M}(X)^d \times \mathcal{M}(X)$

Define the convex, 1-homogeneous functional

$$B(\rho, m, \mu) := \int_X \Psi \left(\frac{d\rho}{d\lambda}, \frac{dm}{d\lambda}, \frac{d\mu}{d\lambda} \right) d\lambda$$

where $\lambda \in \mathcal{M}^+(X)$ is such that $\rho, m, \mu \ll \lambda$ and

$$\Psi(t, x, y) := \frac{|x|^2 + y^2}{2t} \quad \text{if} \quad t > 0, \ \ \Psi = +\infty \text{ else}$$

Proposition

B is weak lower-semicontinuous. If $B(\rho, m, \mu) < +\infty$ and $\partial_t \rho + \operatorname{div} m = \mu$ then*

- *$\rho = dt \otimes \rho_t$ for a weak*-continuous curve $t \mapsto \rho_t \in \mathcal{M}^+(\overline{\Omega})$*
- *$m = \rho v_t$ for some velocity field $v_t \colon (0,1) \times \overline{\Omega} \to \mathbb{R}^d$*
- *$\mu = \rho g_t$ for some growth rate $g_t \colon (0,1) \times \overline{\Omega} \to \mathbb{R}$*

$$B(\rho, m, \mu) = \int_0^1 \int_{\overline{\Omega}} |v_t(x)|^2 + |g_t(x)|^2 \, d\rho_t(x) \, dt$$

Sampling spaces 1/3

(H): the spaces H_t vary in a "measurable" way as $t \in [0,1]$

- $\exists$ Banach space D and $i_t \colon D \to H_t$ linear continuous
- $i_t(D) \subset H_t$ dense, $\sup_t \|i_t\| \leq C$
- for each $\varphi, \psi \in D$ the map $t \mapsto \langle i_t\varphi, i_t\psi\rangle_{H_t}$ is Lebesgue measurable

Step Functions: a map $\varphi \colon [0,1] \to D$ is a step function if

$$\varphi_t = \sum_{j=1}^{N} \chi_{E_j}(t)\,\varphi_j$$

for $\varphi_j \in D$, $E_j \subset [0,1]$ measurable, $N \in \mathbb{N}$.

Strong Measurability: a map $f \colon [0,1] \to \cup_t H_t$ with $f_t \in H_t$ is str. meas. if $\exists\ \varphi^n \colon [0,1] \to D$ step functions s.t.

$$\lim_n \|i_t\varphi_t^n - f_t\|_{H_t} = 0 \quad \text{for a.e.} \quad t \in (0,1)$$

Sampling spaces 2/3

Integrability: a str. meas. map $f\colon [0,1] \to \cup_t H_t$ with $f_t \in H_t$ is integrable if $\exists\ \varphi^n\colon [0,1] \to D$ step functions s.t.

$$\lim_n \int_0^1 \|i_t \varphi_t^n - f_t\|_{H_t}\, dt = 0$$

Theorem (SF, Bredies '19)

Let $f\colon [0,1] \to \cup_t H_t$ be strongly measurable. Then f is **integrable** *iff*

$$\int_0^1 \|f_t\|_{H_t}\, dt < \infty$$

Note: it is possible to show the Theorem after introducing suitable notions of weakly measurable and of separably valued maps $f\colon [0,1] \to \cup_t H_t$, in a way that a version of Pettis Theorem holds.

Sampling spaces 3/3

Definition (Data space)

$$L^2_H = \left\{ f \colon [0,1] \to \cup_t H_t : \ f_t \in H_t, \, f \text{ strongly meas }, \ \int_0^1 \|f_t\|^2_{H_t} \, dt < \infty \right\}$$

Theorem (SF, Bredies '19)

The space L^2_H *is* **Hilbert** *with the scalar product*

$$\langle f, g \rangle_{L^2_H} := \int_0^1 \langle f_t, g_t \rangle_{H_t} \, dt$$

Note: No notion of integral for $f \in L^2_H$. However $i_t^* f_t \colon [0,1] \to D^*$

- $i_t^* f_t$ is always Gelfand integrable: for $E \subset [0,1]$ measurable $\exists \ I_E(f) \in D^*$ s.t.

$$\langle I(f), \varphi \rangle_{D^*,D} = \int_0^1 \langle i_t^* f_t, \varphi \rangle_{D^*,D} \, dt \quad \text{for all} \quad \varphi \in D$$

- $i_t^* f_t$ is not Bochner integrable, as it is not strongly measurable in general (counterexamples for D non reflexive)

Forward operators and Regularized Problem

(K): the operators $K_t^* : \mathcal{M}(\overline{\Omega}) \to H_t$ satisfy

- K_t^* linear continuous and weak*-to-weak continuous
- $\sup_t \|K_t^*\| \leq C$
- for $\rho \in \mathcal{M}(\overline{\Omega})$ the map $t \mapsto K_t^* \rho$ is strongly measurable

Proposition (SF, Bredies '19)

If $t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ is weak continuous then $t \mapsto K_t^* \rho_t$ belongs to L^2_H*

Definition (Regularization)

Let $f \in L^2_H$ be some data. For $(\rho, m, \mu) \in \mathcal{M}(X) \times \mathcal{M}(X)^d \times \mathcal{M}(X)$ set

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

if $\partial_t \rho + \operatorname{div} m = \mu$, and $T_{\alpha,\beta}(\rho, m, \mu) = +\infty$ else.

Existence & Stability

Theorem (SF, Bredies '19)

Assume (H)-(K) and let $f \in L^2_H$. Then

$$\min_{(\rho,m,\mu)\in\mathcal{M}(X)^{d+2}} T_{\alpha,\beta}(\rho,m,\mu) \qquad \text{(MIN)}$$

admits a solution. If K_t^* is injective for a.e. t, then the solution is unique.

Theorem (SF, Bredies '19)

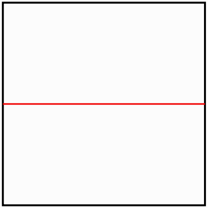
Assume (H)-(K). Let f^n be noisy data such that $f^n \to f^\dagger$ strongly in L^2_H.
Let (ρ^n, m^n, μ^n) be solution to (MIN) with par. $\alpha_n, \beta_n \to 0$ and data f^n. Then

$$(\rho^n, m^n, \mu^n) \overset{*}{\rightharpoonup} (\rho^\dagger, m^\dagger, \mu^\dagger) \ \text{ in } \ \mathcal{M}(X)^{d+2}, \quad K_t^* \rho_t^\dagger = f_t^\dagger \ \text{ for all } \ t \in [0,1]$$

$$(\rho^\dagger, m^\dagger, \mu^\dagger) \in \arg\min \alpha^* B(\rho,m,\mu) + \beta^* \|\rho\|_{\mathcal{M}(X)}, \quad \exists\, \alpha^*, \beta^* \geq 1$$

Variational reconstruction for undersampled MRI 1/2

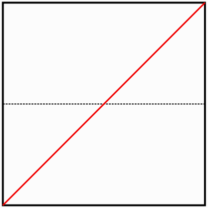
- $\Omega = [0,1]^2$ image frame, $t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ proton density
- $H_t := L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C}^N)$ with $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ sampling measures



$\sigma_t = \mathcal{H}^1 \llcorner L_t$

Variational reconstruction for undersampled MRI 1/2

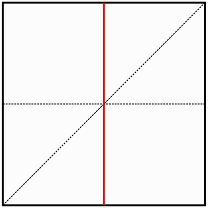
- $\Omega = [0,1]^2$ image frame, $t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ proton density
- $H_t := L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C}^N)$ with $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ sampling measures



$\sigma_t = \mathcal{H}^1 \llcorner L_t$

Variational reconstruction for undersampled MRI 1/2

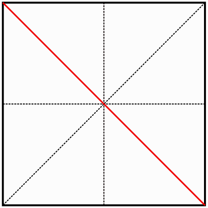
- $\Omega = [0,1]^2$ image frame, $t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ proton density
- $H_t := L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C}^N)$ with $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ sampling measures



$\sigma_t = \mathcal{H}^1 \llcorner L_t$

Variational reconstruction for undersampled MRI 1/2

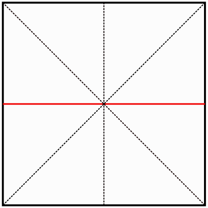
- $\Omega = [0,1]^2$ image frame, $t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ proton density
- $H_t := L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C}^N)$ with $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ sampling measures



$\sigma_t = \mathcal{H}^1 \llcorner L_t$

Variational reconstruction for undersampled MRI 1/2

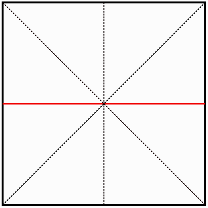
- $\Omega = [0,1]^2$ image frame, $t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ proton density
- $H_t := L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C}^N)$ with $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ sampling measures



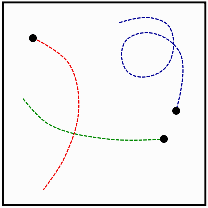
$\sigma_t = \mathcal{H}^1 \llcorner L_t$

Variational reconstruction for undersampled MRI 1/2

- $\Omega = [0,1]^2$ image frame, $t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ proton density
- $H_t := L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C}^N)$ with $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ sampling measures



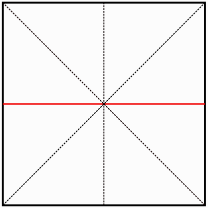
$\sigma_t = \mathcal{H}^1 \llcorner L_t$



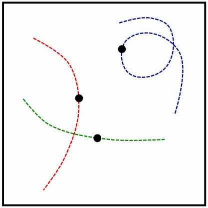
$\sigma_t = \mathcal{H}^0 \llcorner P_t$

Variational reconstruction for undersampled MRI 1/2

- $\Omega = [0,1]^2$ image frame, $t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ proton density
- $H_t := L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C}^N)$ with $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ sampling measures



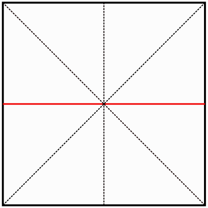
$\sigma_t = \mathcal{H}^1 \llcorner L_t$



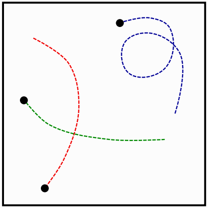
$\sigma_t = \mathcal{H}^0 \llcorner P_t$

Variational reconstruction for undersampled MRI 1/2

- $\Omega = [0,1]^2$ image frame, $t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ proton density
- $H_t := L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C}^N)$ with $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ sampling measures



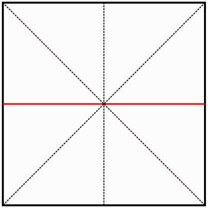
$\sigma_t = \mathcal{H}^1 \llcorner L_t$



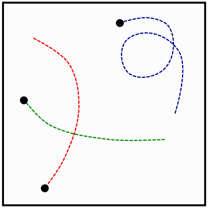
$\sigma_t = \mathcal{H}^0 \llcorner P_t$

Variational reconstruction for undersampled MRI 1/2

- $\Omega = [0,1]^2$ image frame, $t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ proton density
- $H_t := L^2_{\sigma_t}(\mathbb{R}^2; \mathbb{C}^N)$ with $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ sampling measures



$\sigma_t = \mathcal{H}^1 \llcorner L_t$



$\sigma_t = \mathcal{H}^0 \llcorner P_t$

- $K_t^* : \mathcal{M}(\overline{\Omega}) \to H_t$ masked Fourier transform

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

with $c_j \in C_0(\mathbb{R}^2; \mathbb{C})$ coil sensitivities (accounting for phase inhomogeneities)

Variational reconstruction for undersampled MRI 2/2

(M): Assume that the family $\sigma_t \in \mathcal{M}^+(\mathbb{R}^2)$ satisfies:

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

Theorem (SF, Bredies '19)

Assume (M). Let $\alpha, \beta, \delta > 0$, $f \in L^2_H$ and $c \in C_0(\mathbb{R}^2; \mathbb{C}^N)$. Then

$$\min_{\substack{(\rho,m,\mu)\in\mathcal{M}(X)^4 \\ \partial_t \rho + \operatorname{div} m = \mu}} \frac{1}{2} \sum_{j=1}^{N} \int_0^1 \|\mathfrak{F}(c_j \rho_t) - f_t\|^2_{L^2_{\sigma_t}} \, dt + \alpha B_\delta(\rho, m, \mu) + \beta \|\rho\|$$

admits a solution (ρ, m, μ) with

- $\rho = dt \otimes \rho_t$ *with $t \mapsto \rho_t$ weak* continuous*
- $m = \rho v$ *for some velocity $v \colon (0,1) \times \overline{\Omega} \to \mathbb{R}^2$*
- $\mu = \rho g$ *for some growth rate $g \colon (0,1) \times \overline{\Omega} \to \mathbb{R}^2$*

Numerical Results for Benamou-Brenier regularizer

Let $\Omega \subset \mathbb{R}^d$ be open bounded, $X := (0,1) \times \overline{\Omega}$. Consider the Benamou-Brenier energy

$$B(\rho, m) := \int_X \Psi\left(\frac{d\rho}{d\lambda}, \frac{dm}{d\lambda}\right) d\lambda$$

where $\lambda \in \mathcal{M}^+(X)$ is such that $\rho, m \ll \lambda$ and

$$\Psi(t,x) := \frac{|x|^2}{2t} \quad \text{if} \quad t > 0, \;\; \Psi = +\infty \text{ else}$$

Recall: if $\rho = dt \otimes \rho_t$ and $m = v\rho$

$$B(\rho, m) := \frac{\alpha}{2} \int_0^1 |v(t,x)|^2 \, d\rho_t(x) \, dt$$

Bibliography:

- SF, Bredies, Carioni, Romero - *On the extremal points of the ball of the Benamou-Benier energy* (Preprint 2019)
- SF, Bredies, Carioni, Romero - *A generalized conditional gradient method for dynamic inverse problems with optimal transport regularization* (In preparation)

Main Question

Definition (Regularizer)

Let $\alpha, \beta > 0$. For $(\rho, m) \in \mathcal{M}(X) \times \mathcal{M}(X)^d$ we set

$$J_{\alpha,\beta}(\rho, m) := \begin{cases} \alpha B(\rho, m) + \beta \|\rho\|_{\mathcal{M}(X)} & \text{if } \partial_t \rho + \operatorname{div} m = 0 \\ +\infty & \text{otherwise} \end{cases}$$

For $t \in [0,1]$ assume given:

- H_t family of Hilbert spaces satisfying (H)
- $K_t^* \colon \mathcal{M}(\overline{\Omega}) \to H_t$ linear continuous operators satisfying (K)

Problem

Given $f \in L^2([0,1]; H)$ compute a minimizer $(\rho, m) \in \mathcal{M}(X)^{d+1}$ for

$$T_{\alpha,\beta}(\rho, m) := \frac{1}{2} \int_0^1 \|K_t^* \rho_t - f_t\|_{H_t}^2 + J_{\alpha,\beta}(\rho, m)$$

Extremal points of $J_{\alpha,\beta}$

Consider the convex unit ball of $J_{\alpha,\beta}$

$$C := \left\{(\rho, m) \in \mathcal{M}(X) \times \mathcal{M}(X)^d : \ J_{\alpha,\beta}(\rho, m) \leq 1\right\}$$

Definition

For $\gamma \in \mathrm{AC}^2([0,1];\overline{\Omega})$ define the measures $\rho_\gamma \in \mathcal{M}(X)$, $m_\gamma \in \mathcal{M}(X)^d$ as

$$\rho_\gamma := a_\gamma \, dt \otimes \delta_{\gamma(t)} \,, \quad m_\gamma := \dot{\gamma} \, \rho_\gamma \,, \quad a_\gamma^{-1} := \frac{\alpha}{2} \int_0^1 |\dot{\gamma}(t)|^2 \, dt + \beta$$

Theorem (Fanzon, Bredies, Carioni, Romero ’19)

The extremal points of C are characterized by

$$\mathrm{Extr}(C) = \{(0,0)\} \cup \mathcal{C}$$

where

$$\mathcal{C} := \left\{(\rho_\gamma, m_\gamma) : \ \gamma \in \mathrm{AC}^2([0,1];\overline{\Omega})\right\}$$

Idea of the proof - $\{0\} \cup \mathcal{C} \subset \mathrm{Extr}(C)$

The case $(\rho, m) = (0,0)$ is trivial. Then let $\gamma \in \mathrm{AC}^2([0,1];\overline{\Omega})$ and define

$$\rho_\gamma := a_\gamma \, dt \otimes \delta_{\gamma(t)} \,, \quad m_\gamma := \dot\gamma \, \rho_\gamma \,, \quad a_\gamma^{-1} := \frac{\alpha}{2} \int_0^1 |\dot\gamma(t)|^2 \, dt + \beta$$

We first show that $(\rho_\gamma, m_\gamma) \in C$

- (ρ_γ, m_γ) solves continuity equation: let $\varphi \in C_c^1((0,1) \times \overline{\Omega})$

$$\begin{aligned} \int_0^1 \int_\Omega \partial_t \varphi \, d\rho_\gamma + \nabla\varphi \cdot dm_\gamma &= a_\gamma \int_0^1 \partial_t \varphi(t, \gamma(t)) + \nabla\varphi(t,\gamma(t)) \cdot \dot\gamma(t) \, dt \\ &= a_\gamma \int_0^1 \frac{d}{dt} \varphi(t, \gamma(t)) \, dt = 0 \end{aligned}$$

- $J_{\alpha,\beta}(\rho_\gamma, m_\gamma) = 1$: Take $\lambda := \rho_\gamma$ and recall that $\Psi(t,x) = |x|^2/2t$,

$$\begin{aligned} J_{\alpha,\beta}(\rho_\gamma, m_\gamma) &= \alpha \int_0^1 \int_\Omega \Psi\left(\frac{d\rho_\gamma}{d\lambda}, \frac{dm_\lambda}{d\lambda} \right) \, d\lambda + \beta \, \|\rho\|_{\mathcal{M}(X)} \\ &= a_\gamma \left(\alpha \int_0^1 \int_\Omega \Psi\left(1, \dot\gamma(t)\right) \, d\delta_{\gamma(t)} \, dt + \beta \right) = a_\gamma \, a_\gamma^{-1} = 1 \end{aligned}$$

Idea of the proof - $\{0\} \cup \mathcal{C} \subset \mathrm{Extr}(C)$

Assume we can decompose

$$(\rho_\gamma, m_\gamma) = \lambda\,(\rho^1, m^1) + (1-\lambda)\,(\rho^2, m^2) \tag{D}$$

with $(\rho^j, m^j) \in C$ and $\lambda \in (0,1)$.

- Since $J_{\alpha,\beta}(\rho_\gamma, m_\gamma) = 1$, by convexity and (D) we have $J_{\alpha,\beta}(\rho^j, m^j) = 1$
- Since $J_{\alpha,\beta}(\rho^j, m^j) = 1$ then $\rho^j = a^j\, dt \otimes \rho^j_t$, $m^j = v^j\, \rho^j$ for some $a^j > 0$, $(t \mapsto \rho^j_t) \in P(\overline{\Omega})$ narrowly continuous, $v^j \colon (0,1) \times \overline{\Omega} \to \mathbb{R}^d$ measurable
- From (D) and uniqueness of disintegration

$$a_\gamma\, \delta_{\gamma(t)} = \lambda\, a^1\, \rho^1_t + (1-\lambda)\, a^2\, \rho^2_t \quad \overset{a^j > 0}{\Longrightarrow} \quad \rho^j_t = \delta_{\gamma(t)}$$

- $\partial_t \rho^j + \operatorname{div} m^j = 0$ and $\rho^j = a^j\, dt \otimes \delta_{\gamma(t)}$, $m^j = v^j \rho^j$. This forces

$$v^j(t, \gamma(t)) = \dot\gamma(t)$$

- Since $\rho^j = a^j\, dt \otimes \delta_{\gamma(t)}$ and $m^j = \dot\gamma(t)\, \rho^j \implies J_{\alpha,\beta}(\rho^j, m^j) = a^j / a_\gamma$
- Since $J_{\alpha,\beta}(\rho^j, m^j) = 1$ then $a^j = a_\gamma$. Hence $(\rho^j, m^j) = (\rho_\gamma, m_\gamma)$

Idea of the proof - $\mathrm{Extr}(C) \subset \{0\} \cup \mathcal{C}$

Let $\Omega \subset \mathbb{R}^d$ be open and bounded. Consider

$$\Gamma := \{\gamma \colon [0,1] \to \mathbb{R}^d : \gamma \ \text{ continuous } \}$$

with the supremum norm

Theorem (SF, Bredies, Carioni, Romero '19)

Let $t \mapsto \rho_t \in P(\overline{\Omega})$ be narrowly continuous and $v \colon (0,1) \times \overline{\Omega} \to \mathbb{R}^d$. Assume

$$\partial_t \rho_t + \mathrm{div}(\rho_t v) = 0\,, \quad \int_0^1 \int_\Omega |v(t,x)|^2 \, d\rho_t(x)\, dt < +\infty$$

There exists $\sigma \in P(\Gamma)$ such that

$$\int_{\overline{\Omega}} \varphi(x)\, d\rho_t(x) = \int_\Gamma \varphi(\gamma(t))\, d\sigma(\gamma) \quad \text{for all} \quad \varphi \in C(\overline{\Omega}),\ t \in [0,1]$$

Moreover σ is concentrated on curves $\gamma \in AC^2([0,1];\overline{\Omega})$ such that

$$\dot\gamma(t) = v(t,\gamma(t))$$

L. Ambrosio (Inv. Math. '04) for $\Omega = \mathbb{R}^d$

Idea of the proof - $\mathrm{Extr}(C) \subset \{0\} \cup \mathcal{C}$

- Smooth ρ_t and v so that $\partial_t \rho_t + \mathrm{div}(\rho_t v) = 0$ holds and the ODE

$$\frac{d}{dt} X_t(x) = v(t, X_t(x)) \,, \quad X_0(x) = x$$

admits a global solution $X_t(x) \colon [0,1] \times \mathbb{R}^d \to \mathbb{R}^d$. Freezing time, we have

$$\rho_t = (X_t)_\# \rho_0 \tag{A}$$

- Interpret X as a map $X \colon \mathbb{R}^d \to \Gamma$ via $x \mapsto (t \mapsto X_t(x))$ and define

$$\sigma := (X)_\# \rho_0 \in P(\Gamma) \tag{B}$$

The representation formula holds for fixed $t \in [0,1]$:

$$\int_\Gamma \varphi(\gamma(t)) \, d\sigma(\gamma) \overset{(B)}{=} \int_{\mathbb{R}^d} \varphi(X_t(x)) \, d\rho_0(x) \overset{(A)}{=} \int_{\mathbb{R}^d} \varphi(x) \, d\rho_t(x) \tag{F}$$

Then one can pass to the limit to get thesis (hard part is the concentration)

- Assume (ρ, m) is an extremal point: Since $J_{\alpha,\beta}(\rho, m) \leq 1$ we can apply the Theorem and represent ρ via (F). Extremality forces

$$\sigma = \delta_{\gamma^*} \quad \text{for some} \quad \gamma^* \in AC^2([0,1]; \overline{\Omega}) \quad \overset{(F)}{\Longrightarrow} \quad \text{Thesis}$$

Discrete time sampling and finite-dimensional data

Fix $N \geq 1$ times $0 < t_1 < t_2 < \cdots < t_N < 1$ and let

- H_i finite dimensional Hilbert space, $\mathcal{H} := \bigtimes_{i=1}^{N} H_i$
- $K_i^* \colon \mathcal{M}(\overline{\Omega}) \to H_i$ linear and weak*-continuous

Inverse problem: for $(f_1, \ldots, f_N) \in \mathcal{H}$ find a curve $t \mapsto \rho_t \in \mathcal{M}(\overline{\Omega})$ such that

$$K_i^* \rho_{t_i} = f_i \quad \text{for} \quad i = 1, \ldots, N$$

Regularization: we regularize via

$$T_{\alpha,\beta}(\rho, m) := \frac{1}{2} \sum_{i=1}^{N} \left\| K_i \rho_{t_i} - f_i \right\|_{H_i}^2 + \alpha B(\rho, m) + \beta \left\| \rho \right\|_{\mathcal{M}(X)}$$

Sparse minimizers

Theorem (Fanzon, Bredies, Carioni, Romero '19)

The minimization problem

$$\min_{(\rho,m)\in\mathcal{M}(X)^{d+1}} \frac{1}{2}\sum_{i=1}^{N} \|K_i\rho_{t_i} - f_i\|_{H_i}^2 + \alpha B(\rho,m) + \beta\,\|\rho\|_{\mathcal{M}(X)}$$

admits a sparse minimizer of the form

$$(\rho^*, m^*) = \sum_{i=1}^{p} c_i\,(\rho_{\gamma_i}, m_\gamma)$$

where $c_i > 0$, $\gamma_i \in \mathrm{AC}^2([0,1];\overline{\Omega})$ *and* $p \leq \dim\mathcal{H}$.

K. Bredies, M. Carioni (Calc. Var. PDEs '19)

C. Boyer, A. Chambolle, Y. De Castro, V. Duval, F. De Gournay, P. Weiss (SIAM Opt. '19)

Sparse reconstruction (time continuous case)

Definition (Regularizer)

Let $\alpha, \beta > 0$. For $(\rho, m) \in \mathcal{M}(X) \times \mathcal{M}(X)^d$ we set

$$J_{\alpha,\beta}(\rho, m) := \begin{cases} \alpha B(\rho, m) + \beta \left\|\rho\right\|_{\mathcal{M}(X)} & \text{if } \partial_t \rho + \operatorname{div} m = 0 \\ +\infty & \text{otherwise} \end{cases}$$

For $t \in [0, 1]$ assume given:

- H_t family of Hilbert spaces satisfying (H)
- $K_t^* : \mathcal{M}(\overline{\Omega}) \to H_t$ linear continuous operators satisfying (K)

Problem

Given $f \in L^2([0,1]; H)$ compute a minimizer $(\rho, m) \in \mathcal{M}(X)^{d+1} \times \mathcal{M}(X)^d$ for

$$T_{\alpha,\beta}(\rho, m) := \frac{1}{2} \int_0^1 \left\|K_t^* \rho_t - f_t\right\|_{H_t}^2 + J_{\alpha,\beta}(\rho, m)$$

Generalized conditional gradient method 1/2

First one can replace

$$\min_{(\rho,m)} T_{\alpha,\beta}(\rho,m) := \frac{1}{2}\int_0^1 \|K_t^* \rho_t - f_t\|_{H_t}^2 + J_{\alpha,\beta}(\rho,m)$$

by the equivalent problem

$$\min_{(\rho,m)} \tilde{T}_{\alpha,\beta}(\rho,m) := \frac{1}{2}\int_0^1 \|K_t^* \rho_t - f_t\|_{H_t}^2 + \varphi(J_{\alpha,\beta}(\rho,m)) \qquad \text{(P)}$$

where $\varphi \colon \mathbb{R} \to [0,\infty]$ is for example

$$\varphi(t) := t + \chi_{\{s \leq M_0\}}(t)\,, \quad M_0 := \frac{1}{2}\int_0^1 \|f_t\|_{H_t}^2 \, dt$$

Then one approximates (P) by linearizing the quadratic term around $(\tilde{\rho}, \tilde{m})$

$$\min_{(\rho,m)} -\int_0^1 \langle \rho_t, w_t \rangle_{\mathcal{M}(\overline{\Omega}),C(\overline{\Omega})} \, dt + \varphi(J_{\alpha,\beta}(\rho,m))\,, \quad w_t := -K_t(K_t^* \tilde{\rho}_t - f_t)$$

Generalized conditional gradient method 2/2

Consider the convex unit ball of $J_{\alpha,\beta}$

$$C := \left\{(\rho, m) \in \mathcal{M}(X)^{d+1} : \, J_{\alpha,\beta}(\rho, m) \leq 1\right\}$$

and denote by $\mathrm{Extr}(C)$ its extremal points

Theorem (SF, Bredies, Carioni, Romero '19)

Assume (H)-(K). Let $f \in L^2_H$ and fix $t \mapsto \tilde{\rho}_t \in \mathcal{M}(\overline{\Omega})$ narrowly continuous. Set

$$w_t := -K_t(K_t^* \tilde{\rho}_t - f_t)$$

Then there exists a solution $(\rho^*, m^*) \in \mathrm{Extr}(C)$ to the problem

$$\min_{(\rho,m)\in C} -\int_0^1 \langle \rho_t, w_t \rangle_{\mathcal{M}(\overline{\Omega}), C(\overline{\Omega})} \, dt \qquad \text{(L)}$$

Moreover there exists $M \geq 0$ such that $(M\rho^*, Mm^*)$ is a solution to

$$\min_{(\rho,m)} -\int_0^1 \langle \rho_t, w_t \rangle_{\mathcal{M}(\overline{\Omega}), C(\overline{\Omega})} \, dt + \varphi(J_{\alpha,\beta}(\rho, m))$$

Algorithm

Let $f \in L^2_H$ be given. Initialize $(\rho^0, m^0) := (0,0)$ in $\mathcal{M}(X) \times \mathcal{M}(X)^d$

1. (Insertion) Assume given $\gamma_j \in AC^2([0,1];\overline{\Omega})$ pairwise distinct, $c_j > 0$ and set

$$(\rho^n, m^n) := \sum_j c_j\,(\rho_{\gamma_j}, m_{\gamma_j})$$

   Compute the dual variable $w_t := -K_t(K_t^* \rho_t^n - f_t)$ and solve

$$\gamma^* \in \operatorname*{arg\,min}_{\gamma \in AC^2([0,1];\overline{\Omega})} -\left(\frac{\alpha}{2}\int_0^1 |\dot{\gamma}(t)|^2\,dt + \beta\right)^{-1} \int_0^1 w_t(\gamma(t))\,dt$$

   Set $(\rho^{n+1/2}, m^{n+1/2}) := (\rho^n, m^n) + (\rho_{\gamma^*}, m_{\gamma^*}) = \sum_j c_j\,(\rho_{\gamma_j}, m_{\gamma_j})$

2. (Optimization) Solve the quadratic problem

$$\bar{C} = (\bar{c}_j)_j \in \operatorname*{arg\,min}_{c_j \geq 0}\; T_{\alpha,\beta}\left(\rho^{n+1/2}, m^{n+1/2}\right)$$

   Define

$$(\rho^{n+1}, m^{n+1}) := \sum_j \bar{c}_j\,(\rho_{\gamma_j}, m_{\gamma_j})$$

Convergence

Define the functional distance

$$r(\rho, m) := T_{\alpha,\beta}(\rho, m) - \min T_{\alpha,\beta}$$

Theorem (SF, Bredies, Carioni, Romero '19)

Let $f \in L^2_H$, $\alpha, \beta > 0$ and $(\rho^n, m^n) \in \mathcal{M}(X)^{d+1}$ be the sequence in the Algorithm

- (ρ^n, m^n) *is minimizing with*

$$r(\rho^n, m^n) \leq \frac{C}{n} \qquad \text{(B)}$$

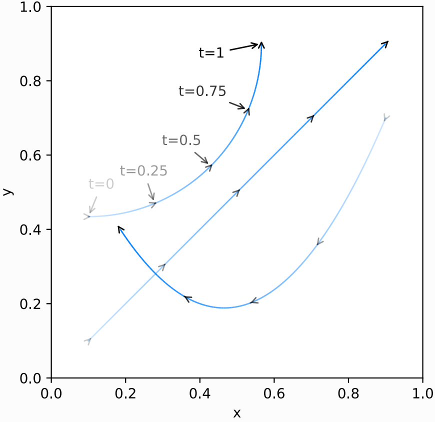
  where $C > 0$ depends only on f, α, β

- *Each weak* accumulation point of (ρ^n, m^n) is a minimum for $T_{\alpha,\beta}$*

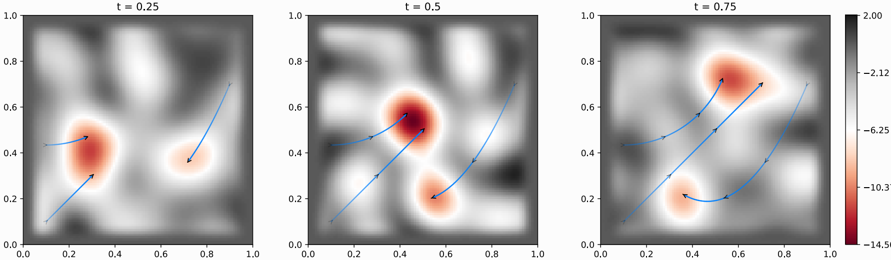
Improving convergence: In the simulations we obtain linear convergence. Therefore one could expect that (B) can be improved (current proof does not take advantage of coefficients optimization step)

Simulations 1/4

- $\Omega = [0,1]^2$, $\sigma = \mathcal{H}^1 \llcorner s$ where $s =$ spiral in Ω
- $H_t := L^2_\sigma(\mathbb{R}^2; \mathbb{C})$ (time independent)
- $K_t^* \colon \mathcal{M}(\overline{\Omega}) \to H_t$ masked Fourier transform (time independent)
- Data is $f_t = K_t^* \rho_t$ where ρ_t are the curves in picture (no noise)



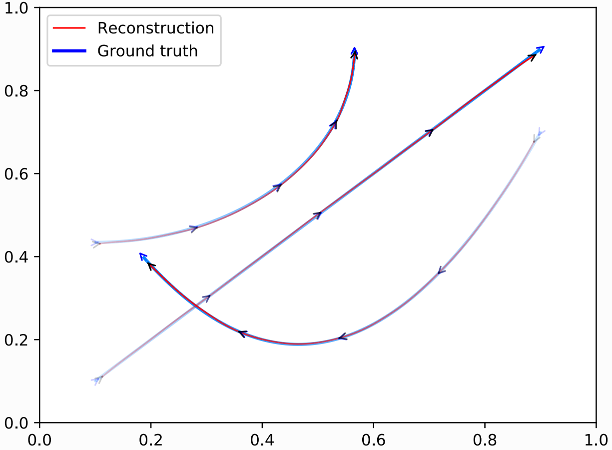
Simulations 2/4



Dual variable first iteration $w_t := -K_t f_t$

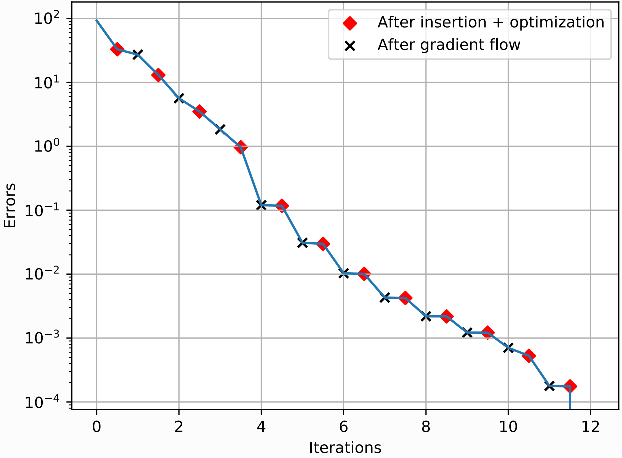
(Recall that $w_t \in C(\overline{\Omega})$ for each $t \in [0,1]$)

Simulations 3/4



Reconstructed trajectories

Simulations 4/4



Convergence plot: exhibits linear rate

$$\text{Error} = T_{\alpha,\beta}(\rho^n, m^n) - T_{\alpha,\beta}(\rho^{n+1}, m^{n+1})$$

Further directions: unbalanced OT case

Consider the regularizer

$$J_{\alpha,\beta}(\rho, m, \mu) := \alpha \int_X \Psi\left(\frac{d\rho}{d\lambda}, \frac{dm}{d\lambda}, \frac{d\mu}{d\lambda}\right) d\lambda + \beta \|\rho\|_{\mathcal{M}(X)} \quad \text{if} \quad \partial_t \rho + \operatorname{div} m = \mu$$

where $\lambda \in \mathcal{M}^+(X)$ is such that $\rho, m, \mu \ll \lambda$ and $\Psi(t, x, y) := \frac{|x|^2 + y^2}{2t}$ if $t > 0$

Further direction: carry out same analysis for the regularizer $J_{\alpha,\beta}$ above

Key ingredients:

- Characterization of the extremal points of $J_{\alpha,\beta}$, which are of the form

$$\rho = h(t)\, dt \otimes \delta_{\gamma(t)}\,, \quad m = \dot{\gamma}\rho\,, \quad \mu = \frac{\dot{h}}{h}\,\rho$$

  where $h\colon [0,1] \to [0,\infty)$, $\gamma\colon [0,1] \to \overline{\Omega}$ satisfy certain regularity properties

- Characterization is based on a superposition principle for $\partial_t \rho + \operatorname{div} m = \mu$

SF, Bredies, Carioni, Romero - A superposition principle for the non-homogeneous continuity equation (In preparation)

Conclusions and Perspectives

Conclusions:

- Introduced rigorous framework for optimal transport regularization of time dependent inverse problems
- Application to variational reconstruction for undersampled MRI
- Characterization of the extremal points of the Benamou-Brenier regularizer
- Numerical algorithm for dynamic spike reconstruction

Perspectives:

- Linear convergence for the conditional gradient method (in progress...)
- Extremal points for the unbalanced transport regularizer (almost done!)
- Numerical analysis for the unbalanced transport regularizer (in progress...)

Thank You!