# Generalized conditional gradient methods for dynamic inverse problems

**Marcello Carioni**

(Department of Applied Mathematics and Theoretical Physics, University of Cambridge)

DGDVC 2021

Based on joint projects with Kristian Bredies, Silvio Fanzon, Francisco Romero and Daniel Walter

FUNDED BY

THE ROYAL SOCIETY

# Summary

- Generalized conditional gradient methods (GCG) for inverse problems in the space of measures
- Dynamic inverse problems with Optimal Transport regularization
- Dynamic generalized conditional gradient method (DGCG)
- Numerical simulations

# Inverse problems in the space of measures

**The total variation case (BLASSO)**

We aim at solving the following inverse problem in the space of measures:

$$\min_{\mu \in \mathcal{M}(\Omega)} \frac{1}{2}\|A\mu - y\|_Y^2 + \alpha\|\mu\|_{\mathcal{M}}$$

- $\mathcal{M}(\Omega)$ is the space of measures in $\Omega$ (bounded, open, convex set)
- $Y$ is the Hilbert space of observations
- $A : \mathcal{M}(\Omega) \to Y$ is a linear continuous observation operator (Fourier measurements)
- $\|\cdot\|_{\mathcal{M}}$ is the total variation norm.

**The total variation case (BLASSO)**
We aim at solving the following inverse problem in the space of measures:

$$\min_{\mu \in \mathcal{M}(\Omega)} \frac{1}{2}\|A\mu - y\|_Y^2 + \alpha\|\mu\|_{\mathcal{M}}$$

- $\mathcal{M}(\Omega)$ is the space of measures in $\Omega$ (bounded, open, convex set)
- $Y$ is the Hilbert space of observations
- $A : \mathcal{M}(\Omega) \to Y$ is a linear continuous observation operator (Fourier measurements)
- $\|\cdot\|_{\mathcal{M}}$ is the total variation norm.

Resolution of BLASSO is challenging due to the infinite dimensionality and non-reflexivity of the space of measure.

Two main approaches to the problem:

- **Grid-based algorithms**
  The domain is discretized. As a result the problem becomes finite dimensional (LASSO) and not difficult to solve by standard methods, e.g. proximal forward-backward splitting, IST.

  Drawbacks:

  - Resolution is limited by the size of the grid.
  - The inverse problem becomes more ill-conditioned.
  - Difficult to reconstruct spikes precisely.

Two main approaches to the problem:

- **Grid-based algorithms**

  The domain is discretized. As a result the problem becomes finite dimensional (LASSO) and not difficult to solve by standard methods, e.g. proximal forward-backward splitting, IST.

  Drawbacks:

  - Resolution is limited by the size of the grid.
  - The inverse problem becomes more ill-conditioned.
  - Difficult to reconstruct spikes precisely.

- **Grid-free algorithms**

  Optimization is performed in the space of measures directly by so called *Generalized Conditional Gradient Methods*.

  **Inverse problems in the space of measures** Bredies, K., Pikkaranien, H. K., *ESAIM: COCV*, 2013

  **The Alternating Descent Conditional Gradient Method for Sparse Inverse Problems** Boyd, N., Schiebinger, G., Recht, B. *SIAM Journal on Optimization*, 2017

  **The Sliding Frank-Wolfe Algorithm and its Application to Super-Resolution Microscopy** Denoyelle, Q., Duval, V., Peyré, G. and Soubies, E. *Inverse Problems*, 2018

# Generalized conditional gradient methods (GCG)

Generalized conditional gradient methods (GCG) are infinite dimensional generalizations of classical Frank-Wolfe type algorithms.

**Idea of the Generalized conditional gradient algorithm**

# Generalized conditional gradient methods (GCG)

Generalized conditional gradient methods (GCG) are infinite dimensional generalizations of classical Frank-Wolfe type algorithms.

## Idea of the Generalized conditional gradient algorithm

- We want to construct a sparse iterate

$$\mu^n = \sum_i c_i^n \delta_{x_i^n}$$

where $c_i^n \in \mathbb{R}$ and $x_i^n \in \Omega$ that is converging to a solution of BLASSO.

# Generalized conditional gradient methods (GCG)

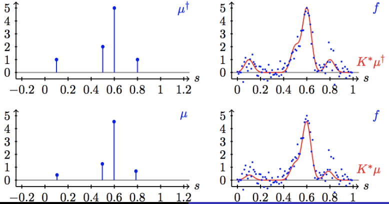
Generalized conditional gradient methods (GCG) are infinite dimensional generalizations of classical Frank-Wolfe type algorithms.

## Idea of the Generalized conditional gradient algorithm

- We want to construct a sparse iterate

$$\mu^n = \sum_i c_i^n \delta_{x_i^n}$$

where $c_i^n \in \mathbb{R}$ and $x_i^n \in \Omega$ that is converging to a solution of BLASSO.
- The next iterate $\mu^{n+1}$ is obtained by adding to $\mu^n$ a new dirac delta $\delta_{\hat{x}}$ for a suitable $\hat{x} \in \Omega$ and suitable coefficients.



- The new spike $\delta_{\hat{x}}$ is chosen to solve the following partially linearized BLASSO:

$$\min_{\mu \in \mathcal{M}(\Omega)} \langle \mu, A^*(A\mu^n - y) \rangle + \alpha \|\mu\|_{\mathcal{M}}$$

where $w^n := A^*(A\mu^n - y)$ is usually called the dual variable at step $n$ of the algorithm.

Solutions of the partially linearized problem are dirac deltas.

**Lemma (Key lemma)**

*A solution of*

$$\min_{\mu \in \mathcal{M}(\Omega)} \langle \mu, w^n \rangle + \alpha \|\mu\|_{\mathcal{M}}$$

*is given by $c\delta_{\hat{x}}$ for $c \in \mathbb{R}$ and $\hat{x} \in \Omega$ such that $|w^n(\hat{x})| = \max_{x \in \Omega} |w^n(x)|$*

- The new spike $\delta_{\hat{x}}$ is chosen to solve the following partially linearized BLASSO:

$$\min_{\mu \in \mathcal{M}(\Omega)} \langle \mu, A^*(A\mu^n - y) \rangle + \alpha \|\mu\|_{\mathcal{M}}$$

where $w^n := A^*(A\mu^n - y)$ is usually called the dual variable at step $n$ of the algorithm.

Solutions of the partially linearized problem are dirac deltas.

**Lemma (Key lemma)**

*A solution of*

$$\min_{\mu \in \mathcal{M}(\Omega)} \langle \mu, w^n \rangle + \alpha \|\mu\|_{\mathcal{M}}$$

*is given by* $c\delta_{\hat{x}}$ *for* $c \in \mathbb{R}$ *and* $\hat{x} \in \Omega$ *such that* $|w^n(\hat{x})| = \max_{x \in \Omega} |w^n(x)|$

The $n+1$ iterate is obtained as

$$\mu^{n+1} = \sum_i c_i^n \delta_{x_i^n} + c\delta_{\hat{x}} \quad (1)$$

Then the coefficients $c_i^n$ are optimized with respect to BLASSO.

- The new spike $\delta_{\hat{x}}$ is chosen to solve the following partially linearized BLASSO:

$$\min_{\mu \in \mathcal{M}(\Omega)} \langle \mu, A^*(A\mu^n - y) \rangle + \alpha \|\mu\|_{\mathcal{M}}$$

where $w^n := A^*(A\mu^n - y)$ is usually called the dual variable at step $n$ of the algorithm.

Solutions of the partially linearized problem are dirac deltas.

**Lemma (Key lemma)**

*A solution of*

$$\min_{\mu \in \mathcal{M}(\Omega)} \langle \mu, w^n \rangle + \alpha \|\mu\|_{\mathcal{M}}$$

*is given by $c\delta_{\hat{x}}$ for $c \in \mathbb{R}$ and $\hat{x} \in \Omega$ such that $|w^n(\hat{x})| = \max_{x \in \Omega} |w^n(x)|$*

The $n+1$ iterate is obtained as

$$\mu^{n+1} = \sum_i c_i^n \delta_{x_i^n} + c\delta_{\hat{x}} \tag{1}$$

Then the coefficients $c_i^n$ are optimized with respect to BLASSO.
The constructed sequence $\mu^{n+1}$ weakly* converges to a minimizer of BLASSO with sublinear rate (Bredies - Pikkaranien)

## Motivation: Motion-Aware Tomographic Reconstruction

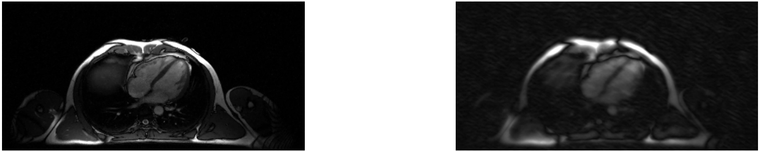
Reconstruct motion of objects (organs, blood flow) from sub-acquisition time samples.

# Dynamic inverse problems with OT regularization

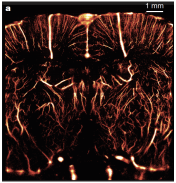
## Motivation: Motion-Aware Tomographic Reconstruction

Reconstruct motion of objects (organs, blood flow) from sub-acquisition time samples. Several difficulties:

- Motion on sub-acquisition time scales $\rightsquigarrow$ artifacts in reconstructions.
- Drawbacks: need for unrealistic assumptions (e.g. periodicity in heart beating). Still limited to low-resolution



Infimal convolution of Total Generalized Variation functionals for dynamic. M. Schloegl, M. Holler, A. Schwarzl, K. Bredies and R. Stollberger, 2017

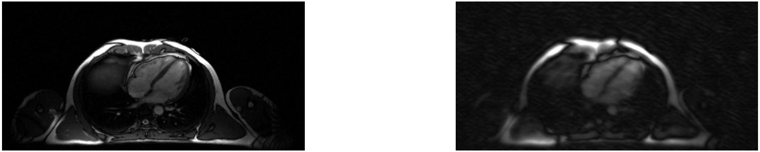



Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging. Errico et al., 2015

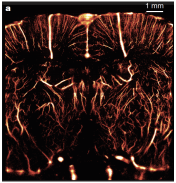
## Motivation: Motion-Aware Tomographic Reconstruction

Reconstruct motion of objects (organs, blood flow) from sub-acquisition time samples. Several difficulties:

- Motion on sub-acquisition time scales $\rightsquigarrow$ artifacts in reconstructions.
- Drawbacks: need for unrealistic assumptions (e.g. periodicity in heart beating). Still limited to low-resolution



Infimal convolution of Total Generalized Variation functionals for dynamic. M. Schloegl, M. Holler, A. Schwarzl, K. Bredies and R. Stollberger, 2017



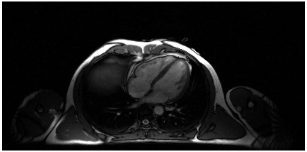

Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging. Errico et al., 2015

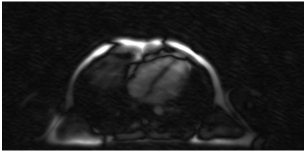
# Dynamic inverse problems with OT regularization

## Motivation: Motion-Aware Tomographic Reconstruction

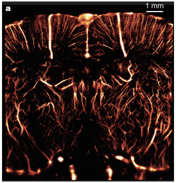
Reconstruct motion of objects (organs, blood flow) from sub-acquisition time samples. Several difficulties:

- Motion on sub-acquisition time scales $\rightsquigarrow$ artifacts in reconstructions.
- Drawbacks: need for unrealistic assumptions (e.g. periodicity in heart beating). Still limited to low-resolution





Infimal convolution of Total Generalized Variation functionals for dynamic. M. Schloegl, M. Holler, A. Schwarzl, K. Bredies and R. Stollberger, 2017



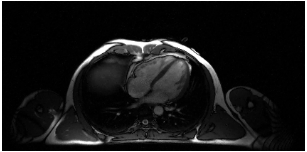

Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging. Errico et al., 2015

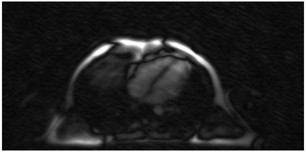
# Dynamic inverse problems with OT regularization

## Motivation: Motion-Aware Tomographic Reconstruction

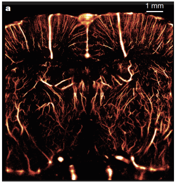
Reconstruct motion of objects (organs, blood flow) from sub-acquisition time samples. Several difficulties:

- Motion on sub-acquisition time scales $\rightsquigarrow$ artifacts in reconstructions.
- Drawbacks: need for unrealistic assumptions (e.g. periodicity in heart beating). Still limited to low-resolution





Infimal convolution of Total Generalized Variation functionals for dynamic. M. Schloegl, M. Holler, A. Schwarzl, K. Bredies and R. Stollberger, 2017




Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging. Errico et al., 2015

## Proposed model: OT regularization for dynamic reconstruction (Bredies-Fanzon, '20)

We want to reconstruct a time dependent measure $\rho_t \in \mathcal{M}(\overline{\Omega})$ for $t \in (0,1)$. At each time $t$ we have a linear observation operator

$$K_t : \mathcal{M}(\overline{\Omega}) \to H_t$$

where $H_t$ are time dependent Hilbert spaces.

## Proposed model: OT regularization for dynamic reconstruction (Bredies-Fanzon, '20)

We want to reconstruct a time dependent measure $\rho_t \in \mathcal{M}(\overline{\Omega})$ for $t \in (0,1)$. At each time $t$ we have a linear observation operator

$$K_t : \mathcal{M}(\overline{\Omega}) \to H_t$$

where $H_t$ are time dependent Hilbert spaces.
Then we solve the following variational inverse problem

$$\frac{1}{2}\int_0^1 \|K_t\rho_t - f_t\|_{H_t}^2 \, dt + \text{OT Regularizer}$$

where $f_t \in H_t$ is the data.

The goal is to choose a regularizer that penalizes an Optimal Transport energy acting on the measure $\rho_t$ at different time instants. In this way the regularized motion will be as *natural* as possible.

An optimal transport approach for solving dynamic inverse problems in spaces of measures. K. Bredies, S. Fanzon, 2020
Dynamic Cell Imaging in PET with Optimal Transport Regularization. Schmitzer, B., Schäfers, K.P., Wirth, B., 2019

# Kinetic Formulation of Optimal Transport - Benamou Brenier Energy

# Kinetic Formulation of Optimal Transport - Benamou Brenier Energy

Given a vector field

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

we say that the pair $(\rho_t, v_t)$ solves the continuity equation with initial conditions if

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

# Kinetic Formulation of Optimal Transport - Benamou Brenier Energy

Given a vector field

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

we say that the pair $(\rho_t, v_t)$ solves the continuity equation with initial conditions if

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

**Theorem (Benamou-Brenier '00)**

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

On the left side we minimize the kinetic energy of a pair $(\rho_t, v_t)$ solving the continuity equation with initial and final data $\rho$ and $\rho_1$.

The right hand side is the classical Monge formulation of OT.

Consider the following inverse problem (Bredies-Fanzon, 20)

$$G_{\alpha,\beta}(\rho_t, v_t) = \frac{1}{2}\int_0^1 \|K_t\rho_t - f_t\|_{H_t}^2 \, dt + J_{\alpha,\beta}(\rho_t, v_t)$$

where the regularizer is

$$J_{\alpha,\beta}(\rho_t, v_t) = \underbrace{\frac{\alpha}{2}\int_0^1 \int_{\overline{\Omega}} |v_t(x)|^2 \, d\rho_t(x) dt}_{\text{Optimal Transport Regularizer}} + \beta \underbrace{\int_0^1 \|\rho_t\|_{\mathcal{M}(\overline{\Omega})} \, dt}_{\text{TV Regularizer}}$$

$$\text{s.t.} \quad \partial_t \rho_t + \mathrm{div}(\rho_t v_t) = 0 \qquad \text{(Continuity Equation)}$$

Consider the following inverse problem (Bredies-Fanzon, 20)

$$G_{\alpha,\beta}(\rho_t, v_t) = \frac{1}{2}\int_0^1 \|K_t\rho_t - f_t\|_{H_t}^2\, dt + J_{\alpha,\beta}(\rho_t, v_t)$$

where the regularizer is

$$J_{\alpha,\beta}(\rho_t, v_t) = \underbrace{\frac{\alpha}{2}\int_0^1 \int_{\overline{\Omega}} |v_t(x)|^2\, d\rho_t(x)dt}_{\text{Optimal Transport Regularizer}} + \underbrace{\beta \int_0^1 \|\rho_t\|_{\mathcal{M}(\overline{\Omega})}\, dt}_{\text{TV Regularizer}}$$

$$\text{s.t.} \quad \partial_t \rho_t + \operatorname{div}(\rho_t v_t) = 0 \qquad \text{(Continuity Equation)}$$

Now consider the substitution $\rho_t v_t = m_t \in \mathcal{M}(\overline{\Omega}, \mathbb{R}^d)$ for every $t$. Then one can rewrite the previous functional in a convex way as:

$$J_{\alpha,\beta}(\rho_t, m_t) = \underbrace{\frac{\alpha}{2}\int_0^1 \int_{\overline{\Omega}} \left(\frac{|m_t(x)|}{\rho_t}\right)^2 d\rho_t(x)dt}_{\text{Optimal Transport Regularizer}} + \underbrace{\beta \int_0^1 \|\rho_t\|_{\mathcal{M}(\overline{\Omega})}\, dt}_{\text{TV Regularizer}}$$

$$\text{s.t.} \quad \partial_t \rho_t + \operatorname{div}(m_t) = 0 \quad \text{(Continuity Equation)}$$

# A dynamic generalized conditional gradient method (DGCG)

We aim at solving the inverse problem with OT regularizer

$$\min_{(\rho_t, m_t)} \frac{1}{2} \int_0^1 \|K_t \rho_t - f_t\|_{H_t}^2 \, dt + J_{\alpha,\beta}(\rho_t, m_t)$$

using a similar approach to the generalized conditional gradient method for static measures.

# A dynamic generalized conditional gradient method (DGCG)

We aim at solving the inverse problem with OT regularizer

$$\min_{(\rho_t, m_t)} \frac{1}{2} \int_0^1 \|K_t \rho_t - f_t\|_{H_t}^2 \, dt + J_{\alpha,\beta}(\rho_t, m_t)$$

using a similar approach to the generalized conditional gradient method for static measures.

How can we generalize the algorithm for BLASSO to this dynamic case?

Which type of sparse iterates should we consider? In other words, which is the dynamic counterpart of the dirac deltas?

# A dynamic generalized conditional gradient method (DGCG)

We aim at solving the inverse problem with OT regularizer

$$\min_{(\rho_t, m_t)} \frac{1}{2} \int_0^1 \|K_t \rho_t - f_t\|_{H_t}^2 \, dt + J_{\alpha,\beta}(\rho_t, m_t)$$

using a similar approach to the generalized conditional gradient method for static measures.

How can we generalize the algorithm for BLASSO to this dynamic case?

Which type of sparse iterates should we consider? In other words, which is the dynamic counterpart of the dirac deltas?

Given a measure $\mu^n = (\rho_t^n, m_t^n)$ we should look at solutions of the partially linearized problem, namely

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

where the dual variable $w_t^n$ is defined as $w_t = K_t(K_t^* \rho_t^n - f_t)$ for every $t$.

Lemma (Bredies, C., Fanzon, Romero)

*There exists a solution of the partially linearized problem*

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

*given by* $c(\hat{\rho}_t, \hat{m}_t)$ *for* $(\hat{\rho}_t, \hat{m}_t) \in Ext(\{(\rho_t, m_t) : J_{\alpha,\beta}(\rho_t, m_t) \leq 1\})$

## Lemma (Bredies, C., Fanzon, Romero)

*There exists a solution of the partially linearized problem*

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

*given by* $c(\hat{\rho}_t, \hat{m}_t)$ *for* $(\hat{\rho}_t, \hat{m}_t) \in Ext(\{(\rho_t, m_t) : J_{\alpha,\beta}(\rho_t, m_t) \leq 1\})$

Therefore the sparse iterate of the algorithm will be constructed by linear combination of extremal points of the set $\{(\rho_t, m_t) : J_{\alpha,\beta}(\rho_t, m_t) \leq 1\}$

$Ext(A)$ denotes the set of its extremal points of $A$: $u \in A$ such that

$$u = \lambda u_1 + (1 - \lambda) u_2 \quad \text{for} \quad \lambda \in (0, 1) \quad \Rightarrow \quad u = u_1 = u_2$$

# Characterization of the extremal points of $J_{\alpha,\beta}$

To devise a generalized conditional gradient method (DGCG) for solving the described dynamic inverse problem. It is necessary to characterize the extremal points of the unit ball of $J_{\alpha,\beta}(\rho_t, m_t)$.

# Characterization of the extremal points of $J_{\alpha,\beta}$

To devise a generalized conditional gradient method (DGCG) for solving the described dynamic inverse problem. It is necessary to characterize the extremal points of the unit ball of $J_{\alpha,\beta}(\rho_t, m_t)$.

**Theorem (Bredies, C. , Fanzon, Romero)**

*Let* $C := \{(\rho_t, m_t) : \ J_{\alpha,\beta}(\rho_t, m_t) \leq 1\}$. *Then*

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

*where*

$$\mathcal{C} := \left\{ C(\gamma,\alpha,\beta)\,(dt \otimes \delta_{\gamma(t)}, \dot\gamma(t) dt \otimes \delta_{\gamma(t)}) : \ \gamma \in \mathrm{AC}^2([0,1];\overline{\Omega}) \right\}$$

*where* $C(\gamma,\alpha,\beta) = \left(\alpha + \beta \int_0^1 |\dot\gamma(t)|^2\, dt\right)^{-1}$.

Extremal points of the Benamou-Brenier energy are pairs of measures concentrated on absolutely continuous curves in $\overline{\Omega}$.

On the extremal points of the Benamou-Brenier energy. K. Bredies, M. Carioni, S. Fanzon, F. Romero, 2019

# Description of DGCG algorithm

# Description of DGCG algorithm

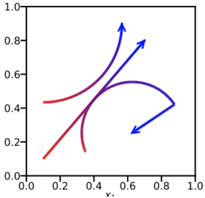
**Definition (Iterates)**

*For every* $\gamma \in \mathrm{AC}^2([0,1];\overline{\Omega})$ *we call* sparse *any measure* $\mu = (\rho, m)$ *such that*

$$\mu = \sum_{j=1}^{N} c_j \mu_{\gamma_j} = \sum_{j=1}^{N} c_j (dt \otimes \delta_{\gamma_j(t)}, \dot{\gamma}_j(t) dt \otimes \delta_{\gamma_j(t)}) \qquad (2)$$

*for some* $N \in \mathbb{N}$, $c_j > 0$ *and* $\gamma_j \in \mathrm{AC}^2([0,1];\overline{\Omega})$.

The iterates of the algorithm are sparse measures, $\mu^n = \sum_{j=1}^{N_n} c_j^n \mu_{\gamma_j^n}$




A generalized conditional gradient method for dynamic inverse problems with optimal transport regularization. K. Bredies, M. Carioni, S. Fanzon, F. Romero, 2020

**Insertion step:**

We add to the iterate an extremal point of the ball of the Benamou-Brenier energy to the iterate $\mu^n$. The extremal point is obtained by computing a minimizer to the partially linearized objective that, defining the dual variable $w_t^n := K_t(K_t^* \rho_t^n - f_t)$, is

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

$$= \min_{(\rho_t, m_t) \in Ext(\{J_{\alpha,\beta} \leq 1\})} \int_0^1 \langle \rho_t, w_t^n \rangle_{\mathcal{M}(\overline{\Omega}), C(\overline{\Omega})} \, dt$$

$$= \min_{\gamma \in \mathrm{AC}([0,1];\overline{\Omega})} \left( \alpha + \beta \int_0^1 |\dot{\gamma}(t)|^2 \, dt \right)^{-1} \int_0^1 w_t^n(\gamma(t)) \, dt \,, \qquad (3)$$

**Insertion step:**

We add to the iterate an extremal point of the ball of the Benamou-Brenier energy to the iterate $\mu^n$. The extremal point is obtained by computing a minimizer to the partially linearized objective that, defining the dual variable $w_t^n := K_t(K_t^* \rho_t^n - f_t)$, is

$$\begin{aligned}
&\min_{(\rho_t, m_t)} \int_0^1 \langle \rho_t, w_t^n \rangle_{\mathcal{M}(\overline{\Omega}), C(\overline{\Omega})} \, dt + J_{\alpha,\beta}(\rho_t, m_t) \\
&= \min_{(\rho_t, m_t) \in Ext(\{J_{\alpha,\beta} \leq 1\})} \int_0^1 \langle \rho_t, w_t^n \rangle_{\mathcal{M}(\overline{\Omega}), C(\overline{\Omega})} \, dt \\
&= \min_{\gamma \in \mathrm{AC}([0,1];\overline{\Omega})} \left( \alpha + \beta \int_0^1 |\dot{\gamma}(t)|^2 \, dt \right)^{-1} \int_0^1 w_t^n(\gamma(t)) \, dt \,, \qquad (3)
\end{aligned}$$

We reduce the insertion step to a variational problem in $\mathrm{AC}^2([0,1];\overline{\Omega})$ that we solve by a multistart gradient descent method.

The next iterate is constructed first by adding to $\mu^n$ the atom (extremal point) $\mu_{\gamma^*}$ associated to the minimizer of (3) denoted by $\gamma^*$.

## Coefficient optimization step

We optimize the coefficients of the linear combination

$$\sum_{j=1}^{N_n} c_j^n \mu_{\gamma_j^n} + \mu_{\gamma^*}$$

by solving

$$\min_{(c_1,c_2,\ldots,c_{N_n+1})\in\mathbb{R}_+^{N_n+1}} G_{\alpha,\beta}\left(\sum_{i=1}^{N_n} c_i \mu_{\gamma_i^n} + c_{N_n+1}\mu_{\gamma^*}\right). \tag{4}$$

## Coefficient optimization step

We optimize the coefficients of the linear combination

$$\sum_{j=1}^{N_n} c_j^n \mu_{\gamma_j^n} + \mu_{\gamma^*}$$

by solving

$$\min_{(c_1, c_2, \ldots, c_{N_n+1}) \in \mathbb{R}_+^{N_n+1}} G_{\alpha,\beta}\left(\sum_{i=1}^{N_n} c_i \mu_{\gamma_i^n} + c_{N_n+1}\mu_{\gamma^*}\right). \tag{4}$$

This is a quadratic problem, due to the linearity of the Benamou-Brenier energy on linear combination of extremal points.

## Coefficient optimization step

We optimize the coefficients of the linear combination

$$\sum_{j=1}^{N_n} c_j^n \mu_{\gamma_j^n} + \mu_{\gamma^*}$$

by solving

$$\min_{(c_1,c_2,\ldots,c_{N_n+1})\in\mathbb{R}_+^{N_n+1}} G_{\alpha,\beta}\left(\sum_{i=1}^{N_n} c_i \mu_{\gamma_i^n} + c_{N_n+1}\mu_{\gamma^*}\right). \tag{4}$$

This is a quadratic problem, due to the linearity of the Benamou-Brenier energy on linear combination of extremal points.

Calling $(c_1^*, c_2^*, \ldots, c_{N_n+1}^*)$ a solution of (4) we obtain the next iterate as

$$\mu^{n+1} = \sum_{j=1}^{N_n} c_j^* \mu_{\gamma_j^n} + c_{N_n+1}^* \mu_{\gamma^*}$$

**Theorem (Bredies, C., Fanzon, Romero)**

*Let $\mu^n = (\rho_t^n, m_t^n)$ an iterate of the DGCG algorithm. Then $\mu^n$ converges weakly* to a minimizer of $G_{\alpha,\beta}$ with sublinear rate.*

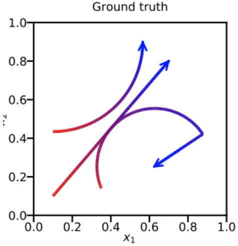
# Numerical simulations

We now want to test our DGCG algorithm to reconstruct observations for the inverse problem

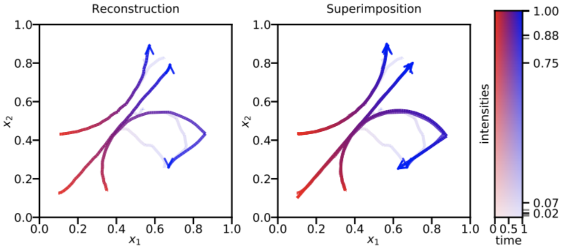
$$\min_{(\rho_t, m_t)} \frac{1}{2} \int_0^1 \|K_t \rho_t - f_t\|_{H_t}^2 \, dt + J_{\alpha,\beta}(\rho_t, m_t)$$

where $J_{\alpha,\beta}$ is the Benamou-Brenier energy.

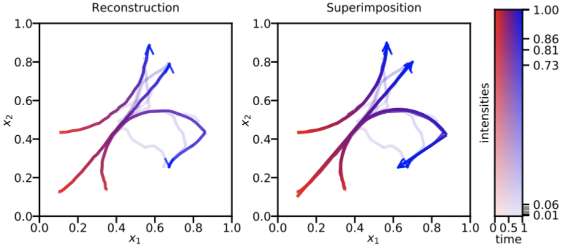
- The domain is $\overline{\Omega} = [0,1]^2$
- The observation operators $K_t : \mathcal{M}(\overline{\Omega}) \to H_t$ are time dependent Fourier measurements that at each time detect different sets of Fourier frequencies of $\rho_t$ (severely ill-posed).
- The data $f_t$ is the image by $K_t$ of sparse dynamic measures (with 20% and 60% of noise)




(A) Considered ground-truth.

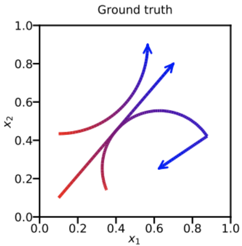



(B) Reconstruction with noiseless data.

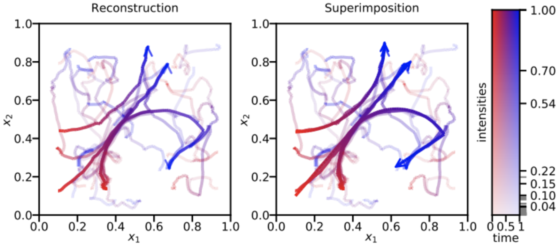



(C) Reconstruction with 20% of relative noise.

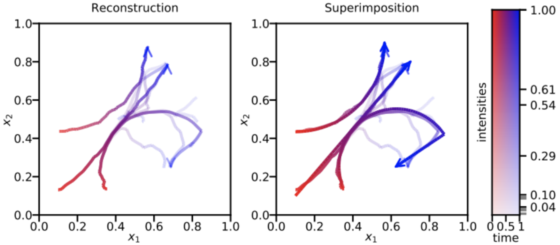
# Reconstructions with 60% of noise:




(A) Considered ground-truth.




(B) Reconstruction with parameter choice $\alpha = \beta = 0.1$.

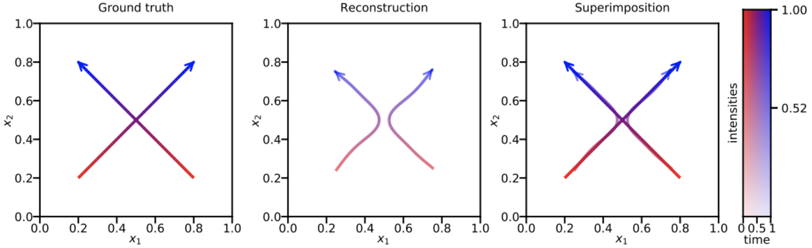



(C) Reconstruction with parameter choice $\alpha = \beta = 0.3$.

# References


- **A generalized conditional gradient method for dynamic inverse problems with optimal transport regularization.** K. Bredies, M. Carioni, S. Fanzon, F. Romero, 2020
- **On the extremal points of the ball of the Benamou-Brenier energy** Bredies, K., Carioni, M., Fanzon, F., Romero, F., 2019
- **Inverse problems in the space of measures** Bredies, K., Pikkaranien, H. K., *ESAIM: COCV*, 2013
- **An optimal transport approach for solving dynamic inverse problems in spaces of measures** Bredies, K., Fanzon, S., 2019
- **Dynamic Cell Imaging in PET with Optimal Transport Regularization** Schmitzer, B., Schäfers, K.P., Wirth, B., 2019

## Effect of the non-uniqueness in the measure representation:




(A) Considered ground-truth. (B) Computed reconstruction with parameters $(\alpha, \beta) = (0.5, 0.5)$.

# Summary of the algorithm and convergence rate

**Algorithm 1:** Core algorithm

**Input:** Data $f_t \in H_t$, parameters $\alpha, \beta > 0$, forward operators $K_t : \mathcal{M}(\overline{\Omega}) \to H_t$

$\boldsymbol{c} \leftarrow (), \quad \boldsymbol{\gamma} \leftarrow ()$

**for** $n = 0, 1, \dots$ **do**

$N_n \leftarrow |\gamma|, \quad \mu^n \leftarrow \sum_{j=1}^{N_n} c_j \mu_{\gamma_j}, \quad w_t^n \leftarrow K_t(K_t^* \rho_t^n - f_t)$

/* Insertion step */

$\gamma^* \leftarrow \arg\min_{\gamma \in \mathrm{AC}^2} \int_0^1 \langle \rho_t^n, w_t^n \rangle_{\mathcal{M}(\overline{\Omega}), C(\overline{\Omega})} \, dt$

$\boldsymbol{\gamma} \leftarrow (\boldsymbol{\gamma}, \gamma^*)$

/* Coefficients optimization step */

$\boldsymbol{c}^* \leftarrow \arg\min_{(c_1, \dots, c_{N_n+1}) \in \mathbb{R}_+^{N_n+1}} T_{\alpha,\beta} \left( \sum_{j=1}^{N_n} c_j \mu_{\gamma_j} + c_{N_n+1} \mu_{\gamma^*} \right)$

$N_{n+1} \leftarrow \#\{j : c_j^* > 0\}$

$(\boldsymbol{\gamma}, \boldsymbol{c}^*) \leftarrow \texttt{delete\_zero\_weighted}(\boldsymbol{\gamma}, \boldsymbol{c}^*)$

$\mu^{n+1} \leftarrow \sum_{j=1}^{N_{n+1}} c_j^* \mu_{\gamma_j}$

# Summary of the algorithm and convergence rate

**Algorithm 1:** Core algorithm

**Input:** Data $f_t \in H_t$, parameters $\alpha, \beta > 0$, forward operators $K_t : \mathcal{M}(\overline{\Omega}) \to H_t$

$\boldsymbol{c} \leftarrow (), \quad \boldsymbol{\gamma} \leftarrow ()$

**for** $n = 0, 1, \ldots$ **do**

$N_n \leftarrow |\gamma|, \quad \mu^n \leftarrow \sum_{j=1}^{N_n} c_j \mu_{\gamma_j}, \quad w_t^n \leftarrow K_t(K_t^* \rho_t^n - f_t)$

`/* Insertion step */`

$\gamma^* \leftarrow \arg\min_{\gamma \in \mathrm{AC}^2} \int_0^1 \langle \rho_t^n, w_t^n \rangle_{\mathcal{M}(\overline{\Omega}), C(\overline{\Omega})} \, dt$

$\boldsymbol{\gamma} \leftarrow (\boldsymbol{\gamma}, \gamma^*)$

`/* Coefficients optimization step */`

$\boldsymbol{c}^* \leftarrow \arg\min_{(c_1, \ldots, c_{N_n+1}) \in \mathbb{R}_+^{N_n+1}} T_{\alpha,\beta} \left( \sum_{j=1}^{N_n} c_j \mu_{\boldsymbol{\gamma}_j} + c_{N_n+1} \mu_{\gamma^*} \right)$

$N_{n+1} \leftarrow \#\{j : c_j^* > 0\}$

$(\boldsymbol{\gamma}, \boldsymbol{c}^*) \leftarrow \texttt{delete\_zero\_weighted}(\boldsymbol{\gamma}, \boldsymbol{c}^*)$

$\mu^{n+1} \leftarrow \sum_{j=1}^{N_{n+1}} c_j^* \mu_{\gamma_j}$