Sparse recovery in Inverse Problems

Silvio Fanzon

Department of Mathematics
University of Hull, UK

18 December 2025

Sapienza, Roma

SAPIENZA
UNIVERSITÀ DI ROMA

UNIVERSITY of HULL

Sparse recovery in Inverse Problems

based on joint works with

Kristian Bredies, Marcello Carioni, Francisco Romero, Daniel Walter

Outline

1. Introduction: Inverse Problems & Sparsity
2. Algorithm for sparse solutions recovery
3. Dynamic Inverse Problems
4. Application to Dynamic MRI

Supported by Austrian Science Fund (FWF) and
Christian Doppler Research Society (CDG) Project PIR27
"Mathematical methods for motion-aware medical imaging"

Outline

What is an Inverse Problem?

Mathematical models and data play a big role in modern science

- Inverse Problems: Link between model parameters and data

Direct Problem: Model Parameters $\rightsquigarrow$ Data (easy)

Inverse Problem: Data $\rightsquigarrow$ Model Parameters (hard)

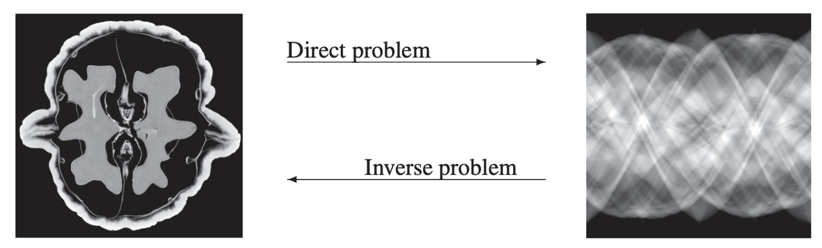
Bredies, Lorenz. *Mathematical Image Processing*. Springer (2018)

Mueller, Siltanen. *Linear and Nonlinear Inverse Problems with Practical Appl.* SIAM, 2012

van Leeuwen, Brune. *10 Lectures on Inverse Problems and Imaging.* Online link

Example: X-ray Imaging

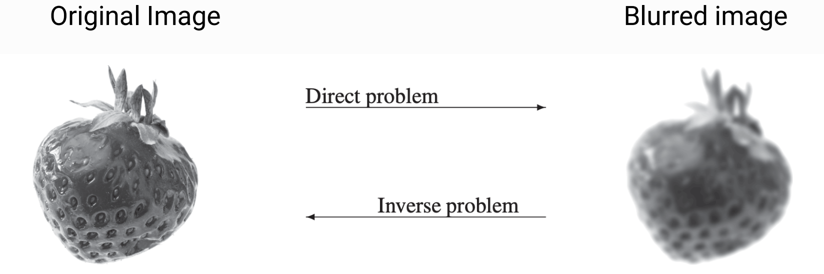
Walnut | X-ray data (sinogram form)



Direct Problem: X-rays pass through walnut, detectors measure attenuation

Inverse problem: Given many X-ray measurements from different angles, reconstruct the walnut

Example: Image Deblurring



Direct problem: Sharp image becomes blurred due to camera motion or focus issues

Inverse problem: Given the blurred image, recover the original sharp image

Famous example: Hubble Space Telescope

- Hubble Space Telescope launched in 1990
- However images were blurred due to flawed lenses (Left)
- This issue was corrected through image processing (Right)

Link to article on NASA's website

Mathematical formulation

Inverse Problem: Given data $f \in Y$, find parameters $u \in X$ such that

$$Ku = f$$

- f is data — Y data space – Banach space or $\mathbb{R}^n$
- u are parameters — X the parameters space – Same as above
- $K : X \to Y$ is Forward Operator
- K models the process to obtain the data from the parameters

Ill-posed Inverse Problems

Consider the inverse problem

$$Ku = f \tag{P}$$

Problem (P) is **well-posed** if all three conditions hold:

1. **Existence**: There exists at least one solution
2. **Uniqueness**: There exists at most one solution
3. **Stability:** The solution depends continuously on the data, i.e., there exists a constant $C > 0$ such that

$$\|u - u'\|_X \leq C \, \|f - f'\|_Y \quad \text{where} \quad Ku = f \, , \; Ku' = f'$$

Problem (P) is **ill-posed** if it is not well-posed

Measurements are noisy

Consider the inverse problem

$$Ku = f \tag{P}$$

- **Ideal world:** Measurement comes from operator $\rightsquigarrow \; f = Ku$
- **Reality:** We can only observe noisy measurements

$$f^{\varepsilon} = Ku + \varepsilon \,, \qquad \varepsilon \text{ random (unknown) noise}$$

Goal: To recover u from noisy measurement f^{ε}

Main difficulty: K^{-1} does not exist or is not continuous $\rightsquigarrow$ **ill-posedness**

Variational Regularization

$$Ku = f \tag{P}$$

- (P) might not have solution. Find **approximate** solution by **least-squares**

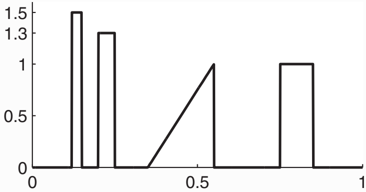
$$\min_{u \in X} \|Ku - f\|_Y^2 \tag{P'}$$

- **Problem:** Might still have **non-existence**, **non-uniqueness** and / or **instability**
- **Solution:** Replace (P) with the **regularized** least-squares problem

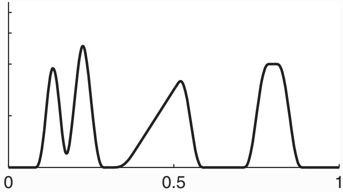
$$\min_{u \in X} \|Ku - f\|_X^2 + \alpha R(u)\,, \qquad R\colon X \to [0, +\infty]\,,\ \alpha > 0$$

1. R makes the problem well-posed and stable – if chosen properly
2. R favors certain solutions – the ones for which $R(u)$ is small

Example: 1D deconvolution



Original signal $\tilde{u} \colon [0,1] \to \mathbb{R}$



Blurred signal $f = \psi \star \tilde{u}$

- Goal: Recover $\tilde{u}$ from noisy data $f^\varepsilon = f + \varepsilon$
- This means solving the 1D-deconvolution problem:

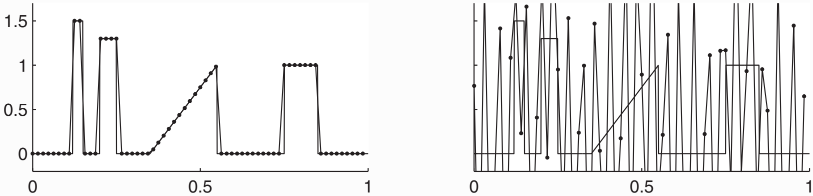
$$\text{Find } u \text{ such that } \psi \star u = f^\varepsilon$$

Naive deconvolution

Solve the discrete 1D-deconvolution problem by least-squares:

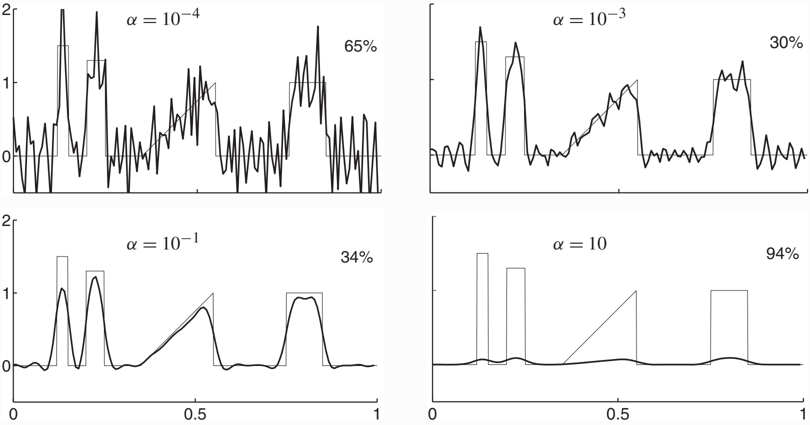
$$u^{\varepsilon} \in \underset{u \in L^2(0,1)}{\arg\min} \; \|\psi \star u - f^{\varepsilon}\|^2_{L^2(0,1)}$$

- Solution behaves well when noise $\varepsilon = 0$ but is terrible when $\varepsilon \neq 0$
  - Instability amplifies noise in the reconstruction
- Below the solid line represents the ground truth $\tilde{u}$
- We need regularizer which penalizes oscillations



u^0 $\qquad$ u^{ε}

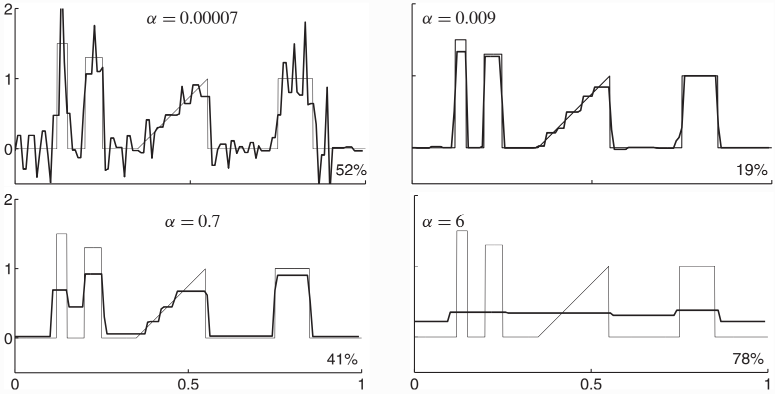
Regularize with L^2 norm: $$\min_{u \in L^2(0,1)} \|\psi \star u - f^\varepsilon\|^2_{L^2(0,1)} + \alpha \|u\|^2_{L^2(0,1)}$$



- Best is $\alpha = 10^{-1}$. Notice the smoothing effect of L^2 regularization
- Smoothness not always desirable (e.g. if u is image with sharp edges – like here)

Reg. with Total Variation (BV semi-norm): $\mathrm{TV}(u) = \sup_{\text{partitions of } [0,1]} \sum_{i=1}^{n} |u(x_i) - u(x_{i-1})|$

$$\min_{u \in L^1(0,1)} \|\psi \star u - f^\varepsilon\|^2_{L^2(0,1)} + \alpha\, \mathrm{TV}(u)$$



- Best is $\alpha = 0.009$. Notice the **sparsifying effect** of TV (the jumps)
- Sparsity assumes different forms, depending on the setting
- **Extremal points of regularizer describe features of sparse solutions**

Summary

Setting: X, Y Banach spaces, $K \colon X \to Y$ linear continuous operator

Inverse Problem: Given $f \in Y$, find $u \in X$ such that

$$Ku = f$$

Main difficulty: K^{-1} does not exist or is not continuous

Variational regularization: Given $f \in Y$, find $u \in X$ which solves

$$\min_{u \in X} \|Ku - f\|_Y^2 + \alpha R(u) \tag{P}$$

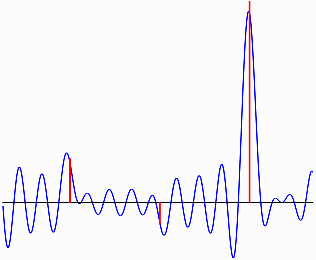
Goals of the Talk:

- Algorithm to recover sparse solutions to (P)
- Framework for regularizing dynamic inverse problems

Outline

1. Introduction to Inverse Problems & Sparsity
2. Algorithm for sparse solution recovery
3. Dynamic Inverse Problems
4. Application to Dynamic MRI

Motivation: Sparse peak recovery



Well-studied problem: Superresolution

- Solve $\qquad \mathfrak{F}u = f \qquad$ on $\quad \Omega$
- $\mathfrak{F}$ Fourier transform, $\Omega \subset \mathbb{R}^d$ finite set
- **Sparsity assumption:** $\overline{u} = \sum_{i=1}^{N} \lambda_i \delta_{x_i}$

Radon-norm regularization:

- Solve variational problem in space of Radon measures (noisy data)

$$\min_{u \in \mathcal{M}(\Omega)} \left\| \mathfrak{F}u - f \right\|_{L^2(\Omega)}^2 + \alpha \left\| u \right\|_{\mathcal{M}(\Omega)}$$

- Want to recover sparse solution $\overline{u} = \sum_{i=1}^{N} \lambda_i \delta_{x_i}$
- [Candès, Fernandez-Granda. **CPAM** (2013)] and many more

$\rightsquigarrow$ **(Fast) algorithms for a general setting**

Minimization Problem in General Setting

$$\min_{u \in X} \; F(Ku) + R(u)$$

- **Parameters:** X separable Banach space with predual X_*
- **Data:** Y Hilbert space
- **Forward operator:** $K : X \to Y$ linear and weak*-to-strong continuous
- **F** $\rightsquigarrow$ **Loss function:** **Smooth** $+$ **Strictly Convex**

$$F : Y \to [0, \infty) \qquad \left(F(y) = \|y - f\|_Y^2 \right)$$

- **R** $\rightsquigarrow$ **Regulariser:** **Convex** $+$ **1-homogeneous** $+$ **Coercive**

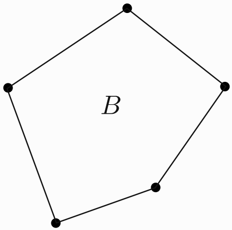
$$R : X \to [0, \infty] \qquad \text{(Promotes Sparsity)}$$

Theorem [1]: Direct method $\Longrightarrow$ Minimizer exists

[1] Bredies, Carioni, **Fanzon**, Walter. **Mathematical Programming** (2024)

Unit Ball of regularizer R

$$B := \{u \in X :\ R(u) \leq 1\}$$



[1] Bredies, Carioni, **Fanzon**, Walter. **Mathematical Programming** (2024)

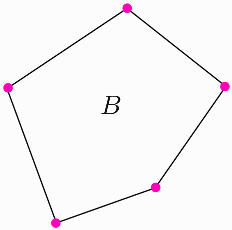
Sparsity

Unit Ball of regularizer R

$$B := \{u \in X : \; R(u) \leq 1\}$$

Extremal Points: $u \in B$ s.t.

$$\begin{cases} u = \alpha u_1 + (1-\alpha) u_2 \\ \alpha \in (0,1)\,, \; u_1, u_2 \in B \end{cases} \implies u = u_1 = u_2$$



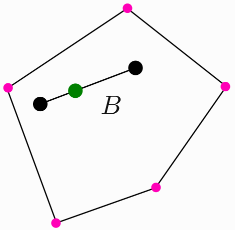
[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

Unit Ball of regularizer R

$$B := \{u \in X : \; R(u) \leq 1\}$$

Extremal Points: $u \in B$ s.t.

$$\begin{cases} u = \alpha u_1 + (1-\alpha) u_2 \\ \alpha \in (0,1)\,, \; u_1, u_2 \in B \end{cases} \implies u = u_1 = u_2$$



[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

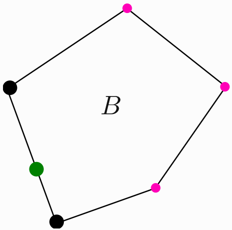
Sparsity

Unit Ball of regularizer R

$$B := \{u \in X : \ R(u) \leq 1\}$$

Extremal Points: $u \in B$ s.t.

$$\begin{cases} u = \alpha u_1 + (1-\alpha) u_2 \\ \alpha \in (0,1)\,, \ u_1, u_2 \in B \end{cases} \implies u = u_1 = u_2$$



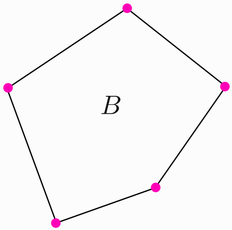
[1] Bredies, Carioni, **Fanzon**, Walter. **Mathematical Programming** (2024)

Unit Ball of regularizer R

$$B := \{u \in X : \; R(u) \leq 1\}$$

Extremal Points: $u \in B$ s.t.

$$\begin{cases} u = \alpha u_1 + (1-\alpha) u_2 \\ \alpha \in (0,1)\,, \;\; u_1, u_2 \in B \end{cases} \implies \quad u = u_1 = u_2$$



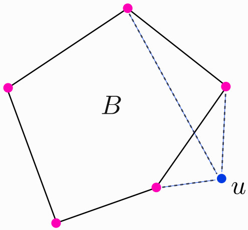
[1] Bredies, Carioni, **Fanzon**, Walter. **Mathematical Programming** (2024)

Unit Ball of regularizer R

$$B := \{u \in X : \; R(u) \leq 1\}$$

Extremal Points: $u \in B$ s.t.

$$\begin{cases} u = \alpha u_1 + (1-\alpha) u_2 \\ \alpha \in (0,1)\,, \; u_1, u_2 \in B \end{cases} \implies u = u_1 = u_2$$



Definition: $u \in X$ **sparse**

Conic combination

$$u = \sum_{i=1}^{N} \lambda_i \, u_i \,, \quad \lambda_i \geq 0 \,, \quad u_i \in \mathrm{Ext}(B)$$

[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

Main Task

Numerical **Algorithm** to compute

$$\bar{u} \in \operatorname*{arg\,min}_{u \in X} \; F(Ku) + R(u)$$

which is **sparse**

$$\bar{u} = \sum_{i=1}^{N} \lambda_i \, u_i \,, \quad \lambda_i \geq 0 \,, \quad u_i \in \operatorname{Ext}(B)$$

Existence of sparse solutions: Proven for $K \colon X \to \mathbb{R}^n$ [1,2]

Very general setting $\rightsquigarrow$ **Important Examples:**

- Training of Machine Learning models $\rightsquigarrow$ $X = \mathbb{R}^d$
- Microstructures in Materials $\rightsquigarrow$ $X = \mathrm{BV}(\mathbb{R}^d)$ Bounded Variation
- Recovery of sparse sources $\rightsquigarrow$ $X = \mathcal{M}(\mathbb{R}^d)$ Radon Measures

---

[1] Bredies, Carioni. **Calc. Var. PDE** (2020)
[2] Boyer, Chambolle, De Castro, Duval, De Gournay, Weiss. **SIAM Optimization** (2019)

Example: Training of Machine Learning models

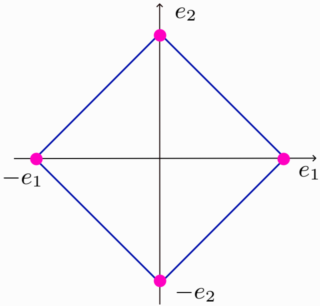
Parameters: vector $\Theta = (\theta_1, \dots, \theta_d) \in \mathbb{R}^d$

ML Model: Fit model to given data

$$\min_{\Theta \in \mathbb{R}^d} F(\theta) + \|\Theta\|_1$$

- Fidelity term F promotes data fit
- 1-norm promotes sparsity – e.g. solutions will have lots of zeros

$$\hat{\Theta} = (0, 0, \theta_{\mathbf{i}}, 0, 0, \dots, 0, 0, \theta_{\mathbf{d}})$$

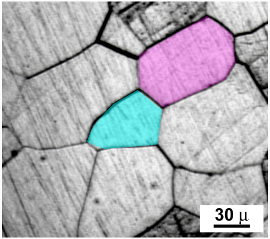


Banach space: $X = \mathbb{R}^d$

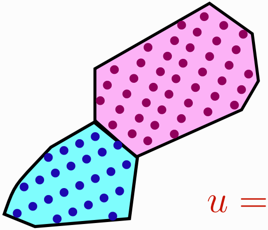
Regularizer: $\|x\|_1 := \sum_{i=1}^{d} |x_i|$

$$\mathrm{Ext}(B) = \{\pm e_i\}_{i=1}^d$$

Example: Microstructures in Materials



Polycrystalline Metal



$E_i \subset \mathbb{R}^2$

$A_i \in \mathbb{R}^{2\times 2}$

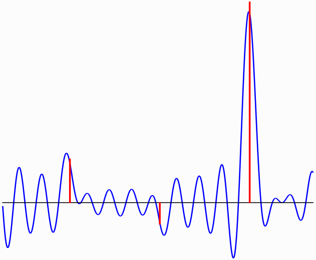
$$u = \sum_{i=1}^{N} A_i \, \chi_{E_i}$$

Banach space: $X = \mathrm{BV}(\mathbb{R}^2)$ functions of bounded variation

Regularizer: $R(u) := \|Du\|_{\mathcal{M}}$, $\mathrm{Ext}(B) = \{\chi_E : E \subset \mathbb{R}^2$ simply connected$\}$

[2] **Fanzon**, Palombaro, Ponsiglione. **SIAM Journal on Mathematical Analysis** (2019)

Example: Recovery of sparse sources



Well-studied problem: Superresolution

- Solve $\mathfrak{F}u = f$ on Ω
- $\mathfrak{F}$ Fourier transform, $\Omega \subset \mathbb{R}^d$ finite set
- **Sparsity assumption:** $\overline{u} = \sum_{i=1}^{N} \lambda_i \delta_{x_i}$

Banach space: $X = \mathcal{M}(\mathbb{R}^d)$ Radon measures

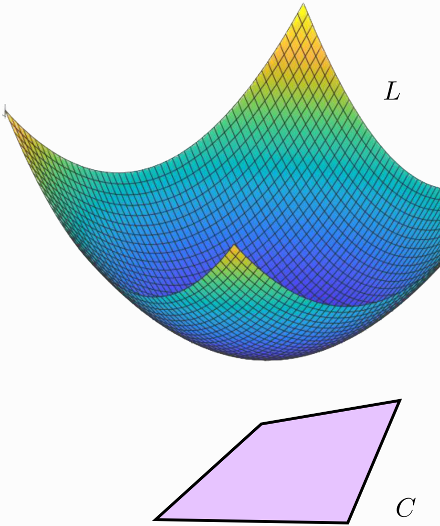
Regularizer: $R(u) := \|u\|_{\mathcal{M}}$ $\qquad \mathrm{Ext}(B) = \{\pm\delta_x : \; x \in \mathbb{R}^d\}$

Starting Point: Classic Frank-Wolfe

Problem: Constrained minimization

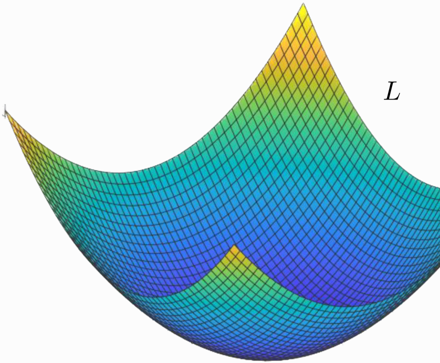
$$\min_{x \in C} \; L(x)$$

- $L \colon \mathbb{R}^N \to \mathbb{R}$ regular convex
- $C \subset \mathbb{R}^N$ convex compact set



M. Jaggi. **Proceedings of Machine Learning Research** (2013)

Starting Point: Classic Frank-Wolfe



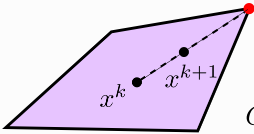
Frank-Wolfe Algorithm: Given $x^k \in C$

1. **Insertion:** Solve linearized problem

$$\min_{x \in C} \langle \nabla L(x^k), x \rangle \quad \mapsto \quad \widehat{x}$$

2. **Convex update:** Set

$$x^{k+1} := x^k + \alpha(\widehat{x} - x^k)\,, \quad \alpha := \frac{2}{k+2}$$

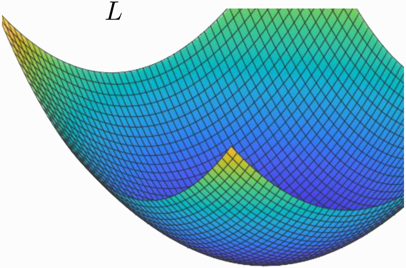


M. Jaggi. **Proceedings of Machine Learning Research** (2013)

Proposed Algorithm: Generalized Frank-Wolfe

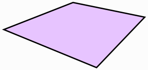
$$\min_{u \in X} L(u) + R(u)\,, \qquad L(u) = F(Ku)$$

L



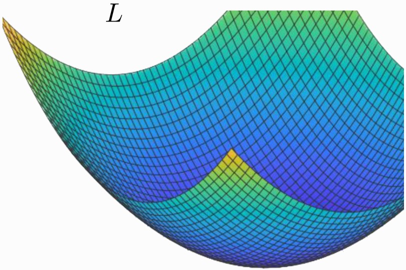
Idea: Set $B = \{R \leq 1\}$. Consider

$$\min_{u \in X} L(u) + \chi_B(u) \iff \min_{u \in B} L(u)$$



$B = \{R \leq 1\}$

[1] Bredies, Carioni, **Fanzon**, Walter. **Mathematical Programming** (2024)



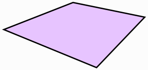
$$\min_{u\in X} L(u)+R(u)\,, \quad L(u) = F(Ku)$$

Idea: Set $B = \{R \leq 1\}$. Consider

$$\min_{u\in X} L(u) + \chi_B(u) \iff \min_{u\in B} L(u)$$

Descent Direction: Solve

$$\min_{v\in B} \langle \nabla L(u), v\rangle \quad \mapsto \quad \hat{v}$$


$B = \{R \leq 1\}$

[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

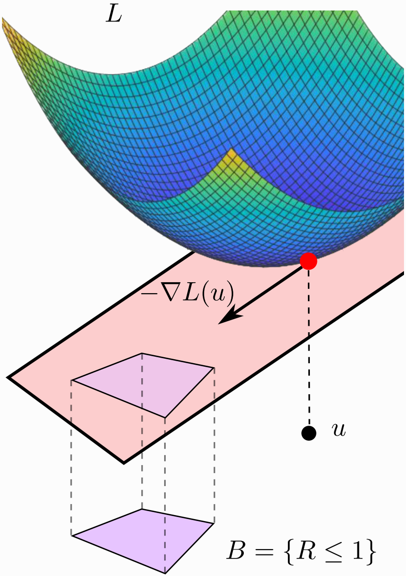
$$\min_{u \in X} L(u)+R(u)\,, \quad L(u) = F(Ku)$$

Idea: Set $B = \{R \leq 1\}$. Consider

$$\min_{u \in X} L(u) + \chi_B(u) \iff \min_{u \in B} L(u)$$

Descent Direction: Solve

$$\min_{v \in B} \langle \nabla L(u), v \rangle \quad \mapsto \quad \hat{v}$$



[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

Proposed Algorithm: Generalized Frank-Wolfe

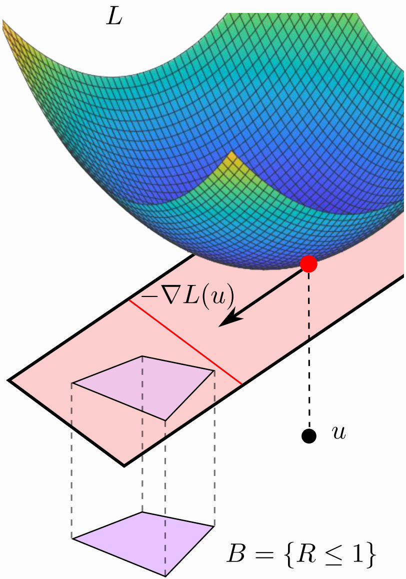
$$\min_{u \in X} L(u) + R(u)\,, \quad L(u) = F(Ku)$$

Idea: Set $B = \{R \leq 1\}$. Consider

$$\min_{u \in X} L(u) + \chi_B(u) \iff \min_{u \in B} L(u)$$

Descent Direction: Solve

$$\min_{v \in B} \langle \nabla L(u), v \rangle \quad \mapsto \quad \hat{v}$$



[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

Proposed Algorithm: Generalized Frank-Wolfe

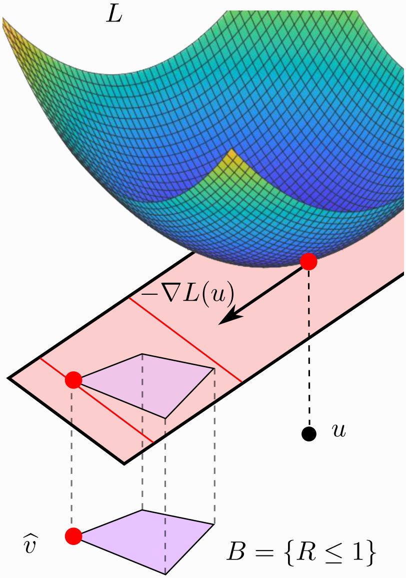
$$\min_{u \in X} L(u)+R(u)\,, \quad L(u) = F(Ku)$$

Idea: Set $B = \{R \leq 1\}$. Consider

$$\min_{u \in X} L(u) + \chi_B(u) \iff \min_{u \in B} L(u)$$

Descent Direction: Solve

$$\min_{v \in B} \langle \nabla L(u), v \rangle \quad \mapsto \quad \widehat{v}$$



[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

Proposed Algorithm: Generalized Frank-Wolfe

$$\min_{u \in X} L(u)+R(u)\,, \quad L(u) = F(Ku)$$

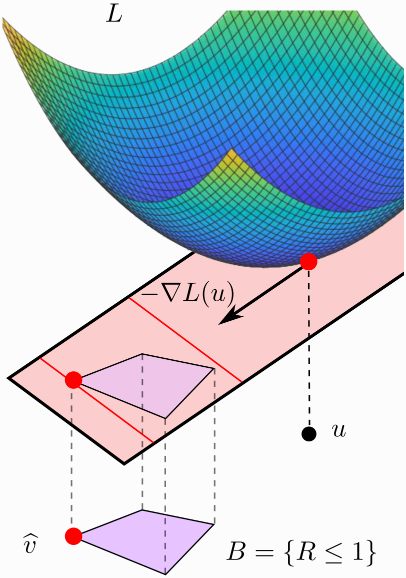
Idea: Set $B = \{R \le 1\}$. Consider

$$\min_{u \in X} L(u) + \chi_B(u) \iff \min_{u \in B} L(u)$$

Descent Direction: Solve

$$\min_{v \in B} \langle \nabla L(u), v \rangle \mapsto \hat{v}$$

Lemma [1]. $\hat{v} \in \mathrm{Ext}(B)$

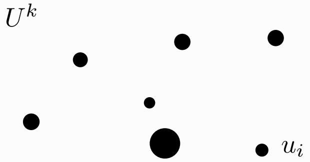


[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

Sparse k-th iterate

$$U^k = \sum_{i=1}^{n} \lambda_i \, u_i$$

$$\lambda_i \geq 0, \quad u_i \in \mathrm{Ext}(B)$$



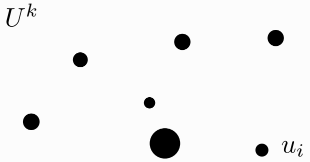
[1] Bredies, Carioni, **Fanzon**, Walter. **Mathematical Programming** (2024)

Proposed Algorithm: Generalized Frank-Wolfe

Sparse k-th iterate

$$U^k = \sum_{i=1}^{n} \lambda_i \, u_i$$

$$\lambda_i \geq 0, \quad u_i \in \mathrm{Ext}(B)$$



1. **Insertion Step:** Solve

$$\widehat{v} \in \operatorname*{arg\,max}_{v \in \mathrm{Ext}(B)} \langle \nabla L(U^k), v \rangle$$

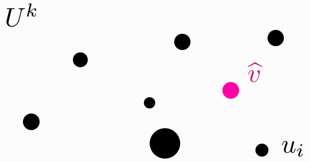
[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

Proposed Algorithm: Generalized Frank-Wolfe

Sparse k-th iterate

$$U^k = \sum_{i=1}^{n} \lambda_i \, u_i$$

$$\lambda_i \geq 0, \quad u_i \in \operatorname{Ext}(B)$$



❶ **Insertion Step:** Solve

$$\widehat{v} \in \underset{v \in \operatorname{Ext}(B)}{\arg\max} \; \langle \nabla L(U^k), v \rangle$$

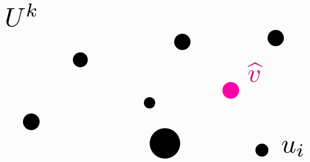
[1] Bredies, Carioni, **Fanzon**, Walter. **Mathematical Programming** (2024)

Proposed Algorithm: Generalized Frank-Wolfe

Sparse k-th iterate

$$U^k = \sum_{i=1}^{n} \lambda_i \, u_i$$

$$\lambda_i \geq 0, \quad u_i \in \mathrm{Ext}(B)$$



1. **Insertion Step:** Solve

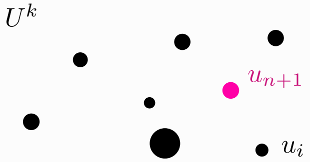
$$\widehat{v} \in \operatorname*{arg\,max}_{v \in \mathrm{Ext}(B)} \langle \nabla L(U^k), v \rangle$$

2. **Fully-Corrective Step:** Set $u_{n+1} := \widehat{v}$

[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

Proposed Algorithm: Generalized Frank-Wolfe

Sparse k-th iterate



$$U^k = \sum_{i=1}^{n} \lambda_i \, u_i$$

$$\lambda_i \geq 0, \quad u_i \in \mathrm{Ext}(B)$$

1. **Insertion Step:** Solve

$$\hat{v} \in \underset{v \in \mathrm{Ext}(B)}{\arg\max} \, \langle \nabla L(U^k), v \rangle$$

2. **Fully-Corrective Step:** Set $u_{n+1} := \hat{v}$

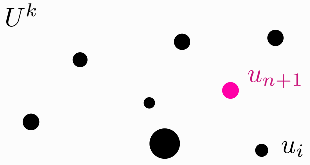
[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

Proposed Algorithm: Generalized Frank-Wolfe

Sparse k-th iterate

$$U^k = \sum_{i=1}^{n} \lambda_i \, u_i$$

$$\lambda_i \geq 0, \quad u_i \in \mathrm{Ext}(B)$$



❶ **Insertion Step:** Solve

$$\hat{v} \in \operatorname*{arg\,max}_{v \in \mathrm{Ext}(B)} \langle \nabla L(U^k), v \rangle$$

❷ **Fully-Corrective Step:** Set $u_{n+1} := \hat{v}$. Optimize coefficients

$$(\lambda_1^*, \ldots, \lambda_{n+1}^*) \in \operatorname*{arg\,min}_{\lambda_i \geq 0} \, (L + R) \left(\sum_{i=1}^{n+1} \lambda_i \, u_i \right) \quad \rightsquigarrow \quad U^{k+1} := \sum_{i=1}^{n+1} \lambda_i^* \, u_i$$

---

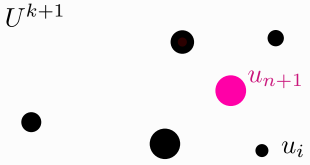
[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

Proposed Algorithm: Generalized Frank-Wolfe

Sparse k-th iterate

$$U^k = \sum_{i=1}^{n} \lambda_i \, u_i$$

$$\lambda_i \geq 0, \quad u_i \in \mathrm{Ext}(B)$$



1 Insertion Step: Solve

$$\widehat{v} \in \underset{v \in \mathrm{Ext}(B)}{\arg\max} \, \langle \nabla L(U^k), v \rangle$$

2 Fully-Corrective Step: Set $u_{n+1} := \widehat{v}$. Optimize coefficients

$$(\lambda_1^*, \dots, \lambda_{n+1}^*) \in \underset{\lambda_i \geq 0}{\arg\min} \, (L + R) \left(\sum_{i=1}^{n+1} \lambda_i \, u_i \right) \quad \rightsquigarrow \quad U^{k+1} := \sum_{i=1}^{n+1} \lambda_i^* \, u_i$$

---

[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

Proposed Algorithm: Generalized Frank-Wolfe

Sparse k-th iterate

$$U^k = \sum_{i=1}^{n} \lambda_i \, u_i$$

$$\lambda_i \geq 0, \quad u_i \in \mathrm{Ext}(B)$$

1. Non-linear problem (usually)
   - Non-linearity due to $\mathrm{Ext}(B)$
   - Expensive and / or hard to solve
2. Quadratic program – Easy to solve

1. **Insertion Step:** Solve

$$\widehat{v} \in \operatorname*{arg\,max}_{v \in \mathrm{Ext}(B)} \langle \nabla L(U^k), v \rangle$$

2. **Fully-Corrective Step:** Set $u_{n+1} := \widehat{v}$. Optimize coefficients

$$(\lambda_1^*, \ldots, \lambda_{n+1}^*) \in \operatorname*{arg\,min}_{\lambda_i \geq 0} (L+R)\left(\sum_{i=1}^{n+1} \lambda_i \, u_i\right) \quad \rightsquigarrow \quad U^{k+1} := \sum_{i=1}^{n+1} \lambda_i^* \, u_i$$

[1] Bredies, Carioni, Fanzon, Walter. **Mathematical Programming** (2024)

Convergence Analysis

Theorem [1]

U^k sparse iterate from the Generalized Frank-Wolfe Algorithm. Then

$$U^k \overset{*}{\rightharpoonup} \bar{u} \,, \qquad \bar{u} \in \arg\min \, G \,, \quad G := L + R$$

General convergence rate is **sublinear** (expected for gradient methods)

$$G(U^k) - \min G \lesssim \frac{1}{k}$$

[1] Bredies, Carioni, **Fanzon**, Walter. **Mathematical Programming** (2024)

Convergence Analysis

Theorem [1]

U^k sparse iterate from the Generalized Frank-Wolfe Algorithm. Then

$$U^k \overset{*}{\rightharpoonup} \bar{u} \,, \qquad \bar{u} \in \arg\min \, G \,, \quad G := L + R$$

General convergence rate is **sublinear** (expected for gradient methods)

$$G(U^k) - \min G \lesssim \frac{1}{k}$$

Highlight: $\bar{u}$ **sparse** $+$ **"Source Condition"** $+$ **"Quadratic Growth"**

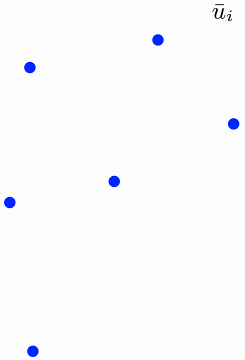
$\implies$ **linear** convergence: $G(U^k) - \min G \lesssim \dfrac{1}{2^k}$

[1] Bredies, Carioni, **Fanzon**, Walter. **Mathematical Programming** (2024)

Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \ , \qquad \bar{u}_i \in \mathrm{Ext}(B)$$

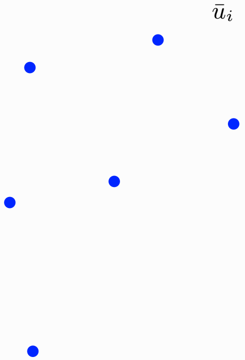


[1] Bredies, Carioni, F., Walter. **Math. Prog.** ('24) [2] Candès, Fernandez-Granda. **CPAM** ('13)

Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \ , \qquad \bar{u}_i \in \operatorname{Ext}(B)$$



❷ **(SC) Source condition:** dual variable

$$\bar{p} := \nabla L(\bar{u}) = K_* \nabla F(K\bar{u})$$

is maximized exactly at $\bar{u}_i$

$$\operatorname*{arg\,max}_{v \in \operatorname{Ext}(B)} \langle \bar{p}, v \rangle = \{\bar{u}_1, \ldots, \bar{u}_M\}$$

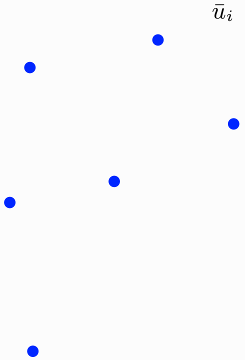
$$\max_{v \in \operatorname{Ext}(B)} \langle \bar{p}, v \rangle = 1$$

[1] Bredies, Carioni, F., Walter. **Math. Prog.** ('24) [2] Candès, Fernandez-Granda. **CPAM** ('13)

Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \,, \qquad \bar{u}_i \in \mathrm{Ext}(B)$$



❷ **(SC) Source condition:** dual variable

$$\bar{p} := \nabla L(\bar{u}) = K_* \nabla F(K\bar{u})$$

is maximized exactly at $\bar{u}_i$

$$\operatorname*{arg\,max}_{v \in \mathrm{Ext}(B)} \ \langle \bar{p}, v \rangle = \{\bar{u}_1, \dots, \bar{u}_M\}$$

$$\max_{v \in \mathrm{Ext}(B)} \ \langle \bar{p}, v \rangle = 1$$

❸ **(QG) Quadratic growth** of $\bar{p}$ around $\bar{u}_i$

$$1 - \langle \bar{p}, u \rangle \gtrsim g(u, u_i)^2 \,, \quad u \sim u_i$$

$\exists \, g \colon \mathrm{Ext}(B)^2 \to [0, \infty)$ “distance”

[1] Bredies, Carioni, F., Walter. **Math. Prog.** (’24) [2] Candès, Fernandez-Granda. **CPAM** (’13)

Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \ , \qquad \bar{u}_i \in \operatorname{Ext}(B)$$

❷ **(SC) Source condition:** dual variable

$$\bar{p} := \nabla L(\bar{u}) = K_* \nabla F(K\bar{u})$$

is maximized exactly at $\bar{u}_i$

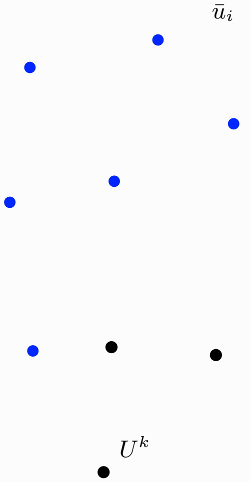
$$\operatorname*{arg\,max}_{v \in \operatorname{Ext}(B)} \langle \bar{p}, v \rangle = \{\bar{u}_1, \ldots, \bar{u}_M\}$$

$$\max_{v \in \operatorname{Ext}(B)} \langle \bar{p}, v \rangle = 1$$

❸ **(QG) Quadratic growth** of $\bar{p}$ around $\bar{u}_i$

$$1 - \langle \bar{p}, u \rangle \gtrsim g(u, u_i)^2 \, , \quad u \sim u_i$$

$\exists \, g \colon \operatorname{Ext}(B)^2 \to [0, \infty)$ "distance"



[1] Bredies, Carioni, F., Walter. **Math. Prog.** ('24) [2] Candès, Fernandez-Granda. **CPAM** ('13)

Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \,, \qquad \bar{u}_i \in \mathrm{Ext}(B)$$

❷ **(SC) Source condition:** dual variable

$$\bar{p} := \nabla L(\bar{u}) = K_* \nabla F(K\bar{u})$$

is maximized exactly at $\bar{u}_i$

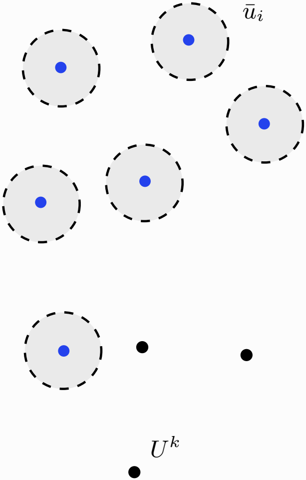
$$\operatorname*{arg\,max}_{v \in \mathrm{Ext}(B)} \langle \bar{p}, v \rangle = \{\bar{u}_1, \dots, \bar{u}_M\}$$

$$\max_{v \in \mathrm{Ext}(B)} \langle \bar{p}, v \rangle = 1$$

❸ **(QG) Quadratic growth** of $\bar{p}$ around $\bar{u}_i$

$$1 - \langle \bar{p}, u \rangle \gtrsim g(u, u_i)^2 \,, \quad u \sim u_i$$

$\exists\, g \colon \mathrm{Ext}(B)^2 \to [0, \infty)$ "distance"



[1] Bredies, Carioni, F., Walter. **Math. Prog.** ('24) [2] Candès, Fernandez-Granda. **CPAM** ('13)

Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \ , \qquad \bar{u}_i \in \mathrm{Ext}(B)$$

❷ **(SC) Source condition:** dual variable

$$\bar{p} := \nabla L(\bar{u}) = K_* \nabla F(K\bar{u})$$

is maximized exactly at $\bar{u}_i$

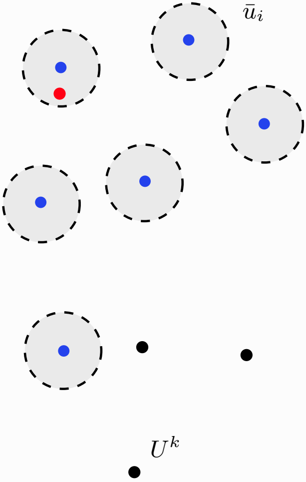
$$\underset{v \in \mathrm{Ext}(B)}{\arg\max} \ \langle \bar{p}, v \rangle = \{\bar{u}_1, \ldots, \bar{u}_M\}$$

$$\max_{v \in \mathrm{Ext}(B)} \ \langle \bar{p}, v \rangle = 1$$

❸ **(QG) Quadratic growth** of $\bar{p}$ around $\bar{u}_i$

$$1 - \langle \bar{p}, u \rangle \gtrsim g(u, u_i)^2 \, , \quad u \sim u_i$$

$\exists \, g \colon \mathrm{Ext}(B)^2 \to [0, \infty)$ "distance"



[1] Bredies, Carioni, F., Walter. **Math. Prog.** ('24) [2] Candès, Fernandez-Granda. **CPAM** ('13)

Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \,, \qquad \bar{u}_i \in \mathrm{Ext}(B)$$

❷ **(SC) Source condition:** dual variable

$$\bar{p} := \nabla L(\bar{u}) = K_* \nabla F(K\bar{u})$$

is maximized exactly at $\bar{u}_i$

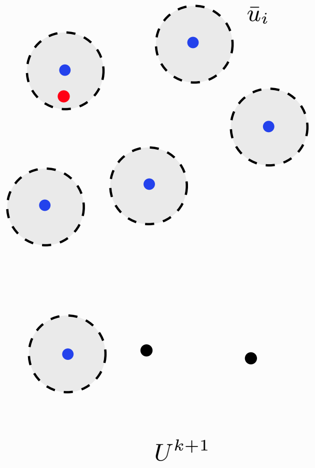
$$\arg\max_{v \in \mathrm{Ext}(B)} \langle \bar{p}, v \rangle = \{\bar{u}_1, \dots, \bar{u}_M\}$$

$$\max_{v \in \mathrm{Ext}(B)} \langle \bar{p}, v \rangle = 1$$

❸ **(QG) Quadratic growth** of $\bar{p}$ around $\bar{u}_i$

$$1 - \langle \bar{p}, u \rangle \gtrsim g(u, u_i)^2 \,, \quad u \sim u_i$$

$\exists \, g \colon \mathrm{Ext}(B)^2 \to [0, \infty)$ "distance"



[1] Bredies, Carioni, F., Walter. **Math. Prog.** ('24) [2] Candès, Fernandez-Granda. **CPAM** ('13)

Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \ , \qquad \bar{u}_i \in \mathrm{Ext}(B)$$

❷ **(SC) Source condition:** dual variable

$$\bar{p} := \nabla L(\bar{u}) = K_* \nabla F(K\bar{u})$$

is maximized exactly at $\bar{u}_i$

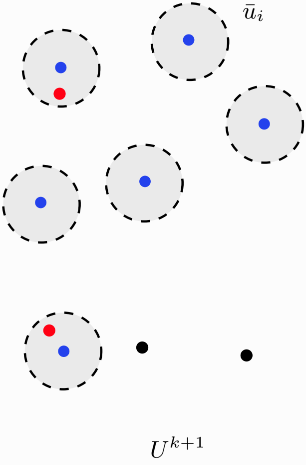
$$\operatorname*{arg\,max}_{v \in \mathrm{Ext}(B)} \ \langle \bar{p}, v \rangle = \{\bar{u}_1, \dots, \bar{u}_M\}$$

$$\max_{v \in \mathrm{Ext}(B)} \ \langle \bar{p}, v \rangle = 1$$

❸ **(QG) Quadratic growth** of $\bar{p}$ around $\bar{u}_i$

$$1 - \langle \bar{p}, u \rangle \gtrsim g(u, u_i)^2 \,, \quad u \sim u_i$$

$\exists \, g \colon \mathrm{Ext}(B)^2 \to [0, \infty)$ "distance"



[1] Bredies, Carioni, F., Walter. **Math. Prog.** ('24) [2] Candès, Fernandez-Granda. **CPAM** ('13)

Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \ , \qquad \bar{u}_i \in \mathrm{Ext}(B)$$

❷ **(SC) Source condition:** dual variable

$$\bar{p} := \nabla L(\bar{u}) = K_* \nabla F(K\bar{u})$$

is maximized exactly at $\bar{u}_i$

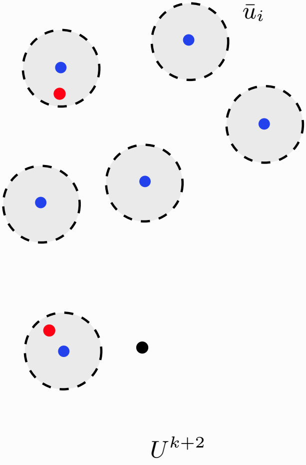
$$\operatorname*{arg\,max}_{v \in \mathrm{Ext}(B)} \ \langle \bar{p}, v \rangle = \{\bar{u}_1, \dots, \bar{u}_M\}$$

$$\max_{v \in \mathrm{Ext}(B)} \ \langle \bar{p}, v \rangle = 1$$

❸ **(QG) Quadratic growth** of $\bar{p}$ around $\bar{u}_i$

$$1 - \langle \bar{p}, u \rangle \gtrsim g(u, u_i)^2 , \quad u \sim u_i$$

$\exists \, g \colon \mathrm{Ext}(B)^2 \to [0, \infty)$ "distance"



[1] Bredies, Carioni, F., Walter. **Math. Prog.** ('24) [2] Candès, Fernandez-Granda. **CPAM** ('13)

Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \ , \qquad \bar{u}_i \in \mathrm{Ext}(B)$$

❷ **(SC) Source condition:** dual variable

$$\bar{p} := \nabla L(\bar{u}) = K_* \nabla F(K\bar{u})$$

is maximized exactly at $\bar{u}_i$

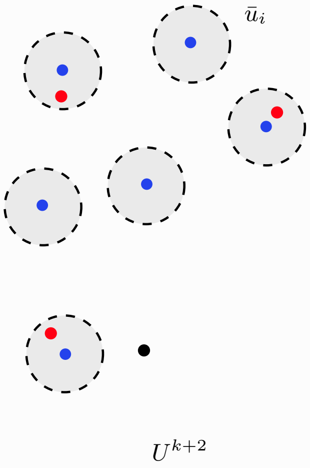
$$\operatorname*{arg\,max}_{v \in \mathrm{Ext}(B)} \ \langle \bar{p}, v \rangle = \{\bar{u}_1, \dots, \bar{u}_M\}$$

$$\max_{v \in \mathrm{Ext}(B)} \ \langle \bar{p}, v \rangle = 1$$

❸ **(QG) Quadratic growth** of $\bar{p}$ around $\bar{u}_i$

$$1 - \langle \bar{p}, u \rangle \gtrsim g(u, u_i)^2 \, , \quad u \sim u_i$$

$\exists \, g \colon \mathrm{Ext}(B)^2 \to [0, \infty)$ "distance"



[1] Bredies, Carioni, F., Walter. **Math. Prog.** ('24) [2] Candès, Fernandez-Granda. **CPAM** ('13)

Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \ , \qquad \bar{u}_i \in \mathrm{Ext}(B)$$

❷ **(SC) Source condition:** dual variable

$$\bar{p} := \nabla L(\bar{u}) = K_* \nabla F(K\bar{u})$$

is maximized exactly at $\bar{u}_i$

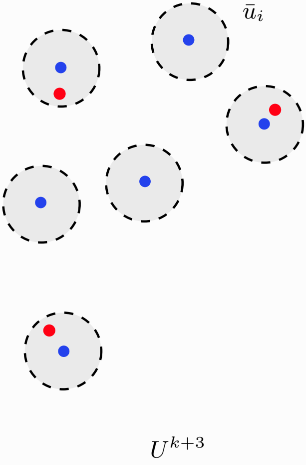
$$\arg\max_{v \in \mathrm{Ext}(B)} \ \langle \bar{p}, v \rangle = \{\bar{u}_1, \dots, \bar{u}_M\}$$

$$\max_{v \in \mathrm{Ext}(B)} \ \langle \bar{p}, v \rangle = 1$$

❸ **(QG) Quadratic growth** of $\bar{p}$ around $\bar{u}_i$

$$1 - \langle \bar{p}, u \rangle \gtrsim g(u, u_i)^2 \ , \quad u \sim u_i$$

$\exists \, g \colon \mathrm{Ext}(B)^2 \to [0, \infty)$ "distance"



[1] Bredies, Carioni, F., Walter. **Math. Prog.** ('24) [2] Candès, Fernandez-Granda. **CPAM** ('13)

Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \, , \qquad \bar{u}_i \in \mathrm{Ext}(B)$$

❷ **(SC) Source condition:** dual variable

$$\bar{p} := \nabla L(\bar{u}) = K_* \nabla F(K\bar{u})$$

is maximized exactly at $\bar{u}_i$

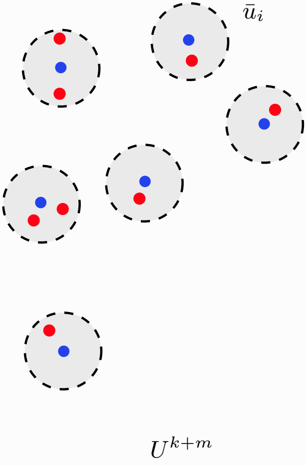
$$\operatorname*{arg\,max}_{v \in \mathrm{Ext}(B)} \langle \bar{p}, v \rangle = \{\bar{u}_1, \ldots, \bar{u}_M\}$$

$$\max_{v \in \mathrm{Ext}(B)} \langle \bar{p}, v \rangle = 1$$

❸ **(QG) Quadratic growth** of $\bar{p}$ around $\bar{u}_i$

$$1 - \langle \bar{p}, u \rangle \gtrsim g(u, u_i)^2 \, , \quad u \sim u_i$$

$\exists \, g \colon \mathrm{Ext}(B)^2 \to [0, \infty)$ "distance"



[1] Bredies, Carioni, F., Walter. **Math. Prog.** ('24) [2] Candès, Fernandez-Granda. **CPAM** ('13)

Comments on Linear Convergence Assumptions

❶ **(S) ∃ sparse minimizer** of $G := L + R$

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \, \bar{u}_i \ , \qquad \bar{u}_i \in \mathrm{Ext}(B)$$

❷ **(SC) Source condition:** dual variable

$$\bar{p} := \nabla L(\bar{u}) = K_* \nabla F(K\bar{u})$$

is maximized exactly at $\bar{u}_i$

$$\arg\max_{v \in \mathrm{Ext}(B)} \langle \bar{p}, v \rangle = \{\bar{u}_1, \ldots, \bar{u}_M\}$$

$$\max_{v \in \mathrm{Ext}(B)} \langle \bar{p}, v \rangle = 1$$

❸ **(QG) Quadratic growth** of $\bar{p}$ around $\bar{u}_i$

$$1 - \langle \bar{p}, u \rangle \gtrsim g(u, u_i)^2 \, , \quad u \sim u_i$$

$\exists \, g \colon \mathrm{Ext}(B)^2 \to [0, \infty)$ “distance”

- (S) + (SC) widely accepted
  - only proven in few cases [2]
  - can be verified numerically
- (QG) is novelty
- In applications we need to:
  - Characterize $\mathrm{Ext}(B)$
  - Define suitable distance g
  - Show (QG) under reasonable assumptions
- **Applications:** [1] Prove fast convergence of Gen. Frank-Wolfe
  - Minimum effort prob.
  - Trace-norm regularized prob.
  - Sparse source identification (heat eqn)

[1] Bredies, Carioni, F., Walter. **Math. Prog.** (’24) [2] Candès, Fernandez-Granda. **CPAM** (’13)

Application: Sparse peak recovery [1, 2]

$$\min_{u\in\mathcal{M}(\Omega)} \frac{1}{2}\|Ku - f^\varepsilon\|^2_{L^2(\Omega)} + \|u\|_{\mathcal{M}(\Omega)}$$

- **Given:** $\Omega \subset \mathbb{R}^d$ and $f^\varepsilon \in L^2(\Omega)$ noisy data
- **Forw. operator:** $K\colon \mathcal{M}(\Omega) \to L^2(\Omega)$

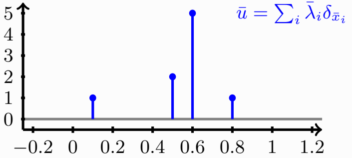
$$Ku = \psi \star u\,, \quad \psi = \text{Gauss. Kern.}$$

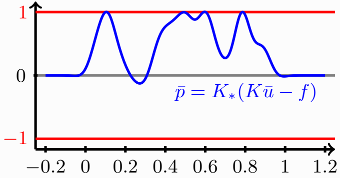
- **Extr. points:** $B = \{\|u\|_{\mathcal{M}(\Omega)} \leq 1\}$

$$\mathrm{Ext}(B) = \{\pm\delta_x : x \in \Omega\}$$

- **(S) $\exists$ sparse solution:**

$$\bar{u} = \sum_{i=1}^{M} \bar{\lambda}_i \delta_{\bar{x}_i}\,, \quad \bar{\lambda}_i > 0\,,\ \bar{x}_i \in \Omega$$





- **(SC)** $\bar{p} = K_*(K\bar{u} - f^\varepsilon) \in C(\Omega)$

$$\arg\max_{v\in\mathrm{Ext}(B)} \langle \bar{p}, v\rangle = \{\delta_{\bar{x}_1}, \dots, \delta_{\bar{x}_M}\}$$

$$\max_{v\in\mathrm{Ext}(B)} \langle \bar{p}, v\rangle = \max_{x\in\Omega} \bar{p}(x) = 1$$

[1] Bredies, Carioni, F., Walter. **Math. Prog.** (2024) [2] Pieper, Walter. **ESAIM: COCV** (2021)

Application: Sparse peak recovery [1, 2]

$$\min_{u \in \mathcal{M}(\Omega)} \frac{1}{2} \|Ku - f^\varepsilon\|^2_{L^2(\Omega)} + \|u\|_{\mathcal{M}(\Omega)}$$

- **(HP)** $\bar{p}$ strictly concave at x_i

$$\text{sign}(\bar{p}(x_i)) \, \langle \xi, \nabla^2 \bar{p}(x_i)\xi \rangle \gtrsim |\xi|^2$$

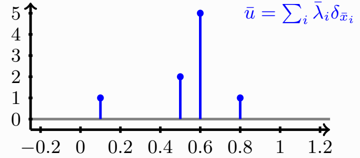
- **Metric:** $g\colon \ \mathrm{Ext}(B) \times \mathrm{Ext}(B) \to [0, \infty)$

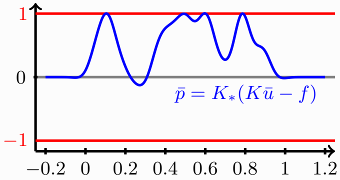
$$g(s_1\delta_{x_1}, s_2\delta_{x_2}) := |s_1 - s_2| + |x_1 - x_2|$$

- **(QG) Quadratic growth** of $\bar{p}$ around $\bar{u}_i$

$$1 - \langle \bar{p}, u \rangle \gtrsim g(u, u_i)^2 , \quad u \sim u_i$$

Theorem [1,2]: (HP) $\Longrightarrow$ (QG)
Gen. Frank-Wolfe converges **linearly**





- **(SC)** $\bar{p} = K_*(K\bar{u} - f^\varepsilon) \in C(\Omega)$

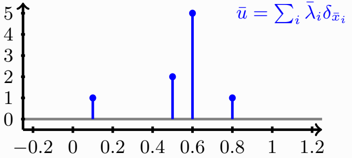
$$\arg\max_{v \in \mathrm{Ext}(B)} \langle \bar{p}, v \rangle = \{\delta_{\bar{x}_1}, \dots, \delta_{\bar{x}_M}\}$$

$$\max_{v \in \mathrm{Ext}(B)} \langle \bar{p}, v \rangle = \max_{x \in \Omega} \bar{p}(x) = 1$$

[1] Bredies, Carioni, F., Walter. **Math. Prog.** (2024) [2] Pieper, Walter. **ESAIM: COCV** (2021)

Application: Sparse peak recovery [1, 2]

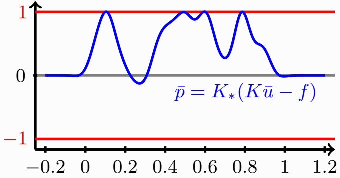
$$\min_{u \in \mathcal{M}(\Omega)} \frac{1}{2} \|Ku - f^\varepsilon\|^2_{L^2(\Omega)} + \|u\|_{\mathcal{M}(\Omega)}$$



Gen. Frank-Wolfe: $\mathrm{Ext}(B) = \{\pm\delta_x : x \in \Omega\}$

Initialize: $u^0 = 0$

Iterate: Given $u^k = \sum_{i=1}^n \lambda_i \delta_{x_i}$



1. **Insertion Step:** $p^k = K_*(Ku^k - f^\varepsilon)$

$$\max_{v \in \mathrm{Ext}(B)} \langle p^k, v\rangle = \max_{x \in \Omega} p^k(x) \quad \rightsquigarrow \quad \hat{x}$$

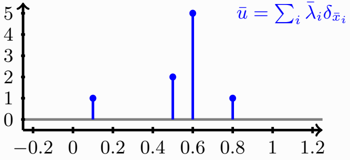
2. **Fully-corrective Step:** Solve

$$(\lambda_1^*, \dots, \lambda_{n+1}^*) \in \arg\min_{\lambda_i \geq 0} G\left(u^k + \lambda_{n+1}\delta_{\hat{x}}\right) \rightsquigarrow u^{k+1} := \left(\sum_{i=1}^n \lambda_i^* \delta_{x_i}\right) + \lambda_{n+1}^* \delta_{\hat{x}}$$

Stop: Based on Primal-Dual gap

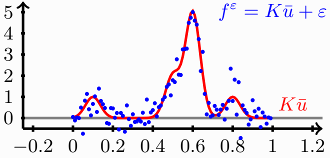
[1] Bredies, Carioni, F., Walter. **Math. Prog.** (2024) [2] Pieper, Walter. **ESAIM: COCV** (2021)

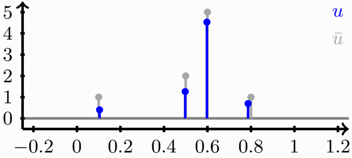
$$\min_{u\in\mathcal{M}(\Omega)} \frac{1}{2}\|Ku - f^\varepsilon\|^2_{L^2(\Omega)} + \|u\|_{\mathcal{M}(\Omega)}$$



Numerical experiment:

- Ground truth $\bar{u}$ with 4 peaks
- Noiseless data $f = K\bar{u}$
- Noisy data $f^\varepsilon = K\bar{u} + \varepsilon$
- Run Gen. Frank-Wolfe $\rightsquigarrow u$
  - u is minimizer (by Thm)
  - u correctly has 4 peaks
  - Weights of peaks are shrunk (effect of regularization)
  - Empirical **linear convergence**



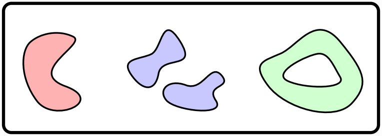


[1] Bredies, Carioni, F., Walter. **Math. Prog.** (2024) [2] Pieper, Walter. **ESAIM: COCV** (2021)

An open problem

Total variation: $X = \mathrm{BV}(\Omega)$, $\Omega \subset \mathbb{R}^d$

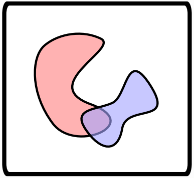
$$G(u) := F(Ku) + \|\nabla u\|_{\mathcal{M}} \,, \qquad \mathrm{Ext}(B) = \left\{ \frac{\chi_E}{\mathrm{Per}(E)} : E \subset \Omega \text{ simple} \right\}$$



Assume: sparse solution $\hat{u} = \sum_{i=1}^{M} \lambda_i \, \chi_{E_i}$

Fast convergence: Which “metric”???

$$g(E_i, E_j) := |E_i \triangle E_j|^{-1} \ \ ???$$

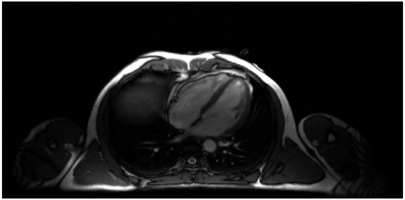


- **Connected problems:** Exact recovery, Noise Robustness

Outline

1. Introduction to Inverse Problems & Sparsity
2. Algorithm for sparse solution recovery
3. **Dynamic Inverse Problems**
4. Application to Dynamic MRI

MRI Scanner



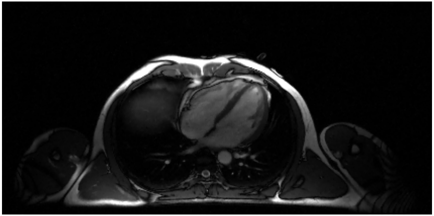
Human Heart

MRI: Medical imaging device, producing gray-scale images

$$u: \ \Omega \subset \mathbb{R}^2 \to \mathbb{R}$$

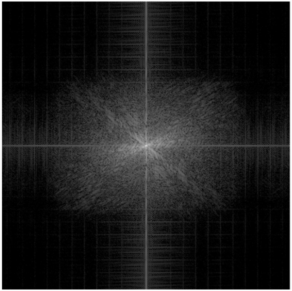
Mathematical model for MRI

Image $u \colon \Omega \to \mathbb{R}$



$\xrightarrow{\mathfrak{F}}$

$\mathfrak{F}u \colon \mathbb{R}^2 \to \mathbb{C}$



$$(\mathfrak{F}u)\,[\xi] = \frac{1}{2\pi} \int_{\mathbb{R}^2} u(x)\, e^{i\xi \cdot x}\, dx\,, \qquad \xi \in \mathbb{R}^2$$

MRI machine measures Fourier coefficients

Inverse Problem:

- Given MRI data y
- Find image $u \colon \Omega \to \mathbb{R}$ s.t.

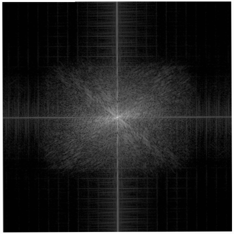
$$\mathfrak{F} u = y$$

Inverse Problem:

- Given MRI data y
- Find image $u \colon \Omega \to \mathbb{R}$ s.t.

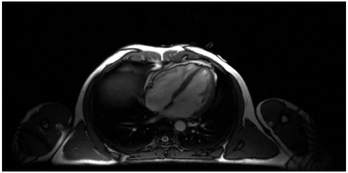
$$\mathfrak{F}u = y$$

Ideal World: Fourier transform is invertible. Unique solution is $u = \mathfrak{F}^{-1}y$



Data y

$$\xrightarrow{\mathfrak{F}^{-1}}$$



Reconstruction u

MRI Inverse Problem

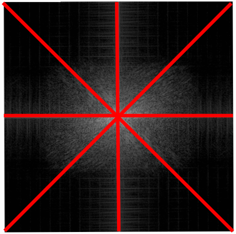
Reality: Things are not straightforward

- Machine is slow in acquiring data $\Longrightarrow$ can only sample **limited data**
- Measurement process is inherently **noisy**

Reality: Things are not straightforward

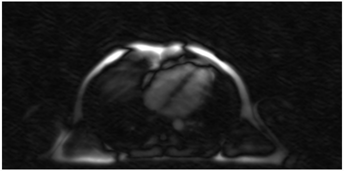
- Machine is slow in acquiring data $\implies$ can only sample **limited data**
- Measurement process is inherently **noisy**

Issue: Plain inversion $\rightsquigarrow$ poor reconstructions



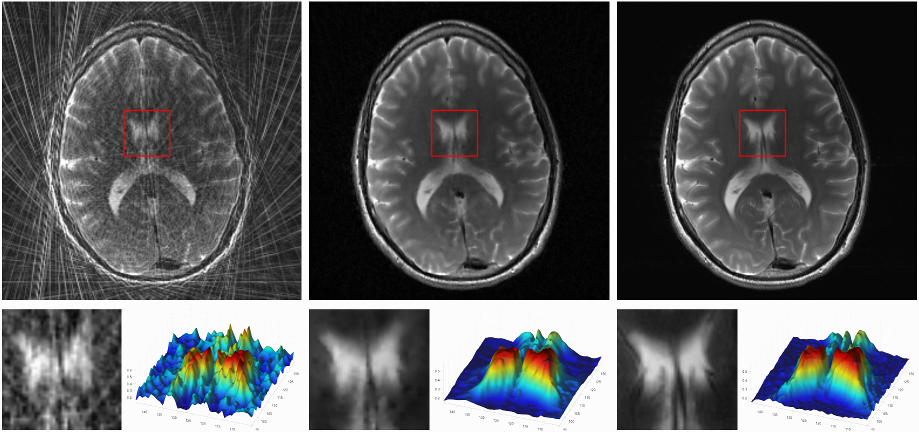
Undersampled noisy data y

$$\xrightarrow{\mathfrak{F}^{-1}}$$



Reconstruction u

Benchmark Regularizer: TGV



Unders. Noisy Data
Least Squares

Unders. Noisy Data
Regularized (TGV)

Full Data
Least Squares

Bredies, Kunisch, Pock. Total Generalized Variation. **SIAM Imaging** (2010)

Motivation: Undersampled Dynamic MRI

Goal: Dynamic MRI $\rightsquigarrow$ **Motion** is big challenge to accurate reconstructions

- High resolution imaging
- Imaging moving organs

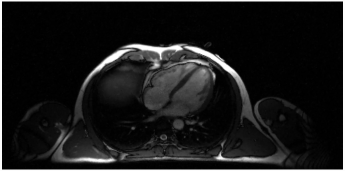
Motivation: Undersampled Dynamic MRI

Goal: Dynamic MRI $\rightsquigarrow$ **Motion** is big challenge to accurate reconstructions

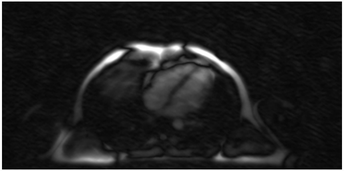
- High resolution imaging
- Imaging moving organs

Dynamic IP: Reconstruct movie u_t from undersampled data series y_t

$$\mathfrak{F}(u_t) = y_t \quad \text{for all} \quad t \in [0,1]$$



Fully sampled data



Undersampled data

Solution: We need regularization for **dynamic inverse problems**

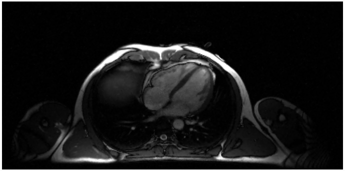
Motivation: Undersampled Dynamic MRI

Goal: Dynamic MRI $\rightsquigarrow$ **Motion** is big challenge to accurate reconstructions

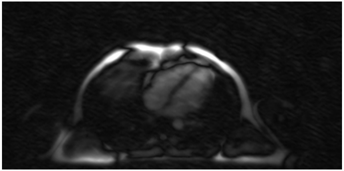
- High resolution imaging
- Imaging moving organs

Dynamic IP: Reconstruct movie u_t from undersampled data series y_t

$$\mathfrak{F}(u_t) = y_t \quad \text{for all} \quad t \in [0,1]$$



Fully sampled data



Undersampled data

Solution: We need regularization for **dynamic inverse problems**

Dynamic Inverse Problem

- **Images:** Radon Measures $\mu \in \mathcal{M}(\Omega)$ ($\Omega \subset \mathbb{R}^N$ bounded closed domain)
- **Data spaces:** Hilbert spaces H_t for $t \in [0,1]$
- **Measurement Operators:** linear continuous maps

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

- **Data points:** Curve $t \mapsto y_t$ with $y_t \in H_t$

[2] Bredies, **Fanzon**. **ESAIM: Mathematical Modelling and Numerical Analysis** (2020)

Dynamic Inverse Problem

- **Images:** Radon Measures $\mu \in \mathcal{M}(\Omega)$ ($\Omega \subset \mathbb{R}^N$ bounded closed domain)
- **Data spaces:** Hilbert spaces H_t for $t \in [0,1]$
- **Measurement Operators:** linear continuous maps

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

- **Data points:** Curve $t \mapsto y_t$ with $y_t \in H_t$

Dynamic Inverse Problem: Find **curve** of measures $t \mapsto \mu_t \in \mathcal{M}(\Omega)$ s.t.

$$K_t \mu_t = y_t \quad \text{for all} \quad t \in [0,1] \tag{P}$$

Assumptions: weak time-regularity for $\{H_t\}_t$ and K_t^* (wk*-measurability)

[2] Bredies, Fanzon. **ESAIM: Mathematical Modelling and Numerical Analysis** (2020)

Dynamic Inverse Problem

- **Images:** Radon Measures $\mu \in \mathcal{M}(\Omega)$ ($\Omega \subset \mathbb{R}^N$ bounded closed domain)
- **Data spaces:** Hilbert spaces H_t for $t \in [0,1]$
- **Measurement Operators:** linear continuous maps

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

- **Data points:** Curve $t \mapsto y_t$ with $y_t \in H_t$

Dynamic Inverse Problem: Find **curve** of measures $t \mapsto \mu_t \in \mathcal{M}(\Omega)$ s.t.

$$K_t \mu_t = y_t \quad \text{for all} \quad t \in [0,1] \qquad \text{(P)}$$

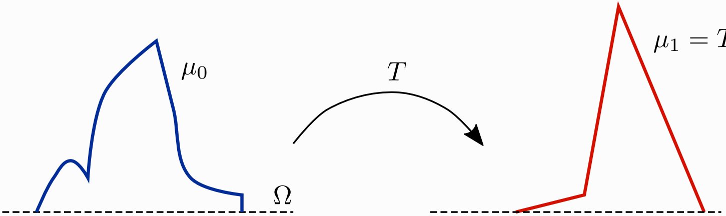
Assumptions: weak time-regularity for $\{H_t\}_t$ and K_t^* (wk*-measurability)

Proposal: Regularize (P) with an **Optimal Transport Energy**

[2] Bredies, Fanzon. **ESAIM: Mathematical Modelling and Numerical Analysis** (2020)

Optimal Transport - Static Formulation

$\Omega \subset \mathbb{R}^d$ bounded domain, $\mu_0, \mu_1 \in \mathcal{P}(\Omega)$, $T\colon \Omega \to \Omega$ measurable displacement



Goal: move μ_0 to μ_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 $\hat{T}$ solving

$$\hat{T} \in \arg\min \left\{ \int_\Omega |T(x)-x|^2 \, d\mu_0(x) \, : \, T\colon \Omega \to \Omega, \, T_\# \mu_0 = \mu_1 \right\}$$

Optimal Transport - Dynamic Formulation

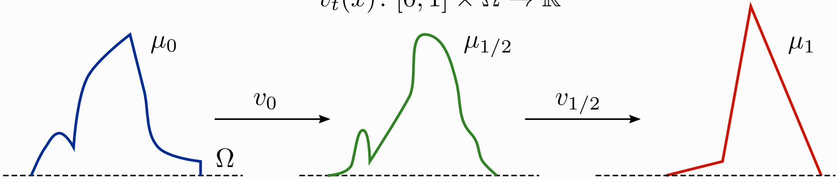
Idea: introduce a time variable $t \in [0,1]$ and consider the density evolution

- time dependent probability measures

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

- μ_t is advected by the velocity field

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



Dynamic model: (μ_t, v_t) solves the continuity equation with initial conditions

$$\begin{cases} \partial_t \mu_t + \operatorname{div}(\mu_t v_t) = 0 \\ \text{Initial data } \mu_0, \text{ final data } \mu_1 \end{cases} \tag{CE-IC}$$

Benamou-Brenier Formula

Theorem: Benamou-Brenier [1]

$$\min_{\substack{(\mu_t, v_t) \\ \text{solving (CE-IC)}}} \int_0^1 \int_\Omega |v_t(x)|^2 \, \mu_t(x) dx \, dt = \min_{\substack{T \colon \Omega \to \Omega \\ T_\# \mu_0 = \mu_1}} \int_\Omega |T(x) - x|^2 \, \mu_0(x) \, dx$$

Advantages of Dynamic Formulation:

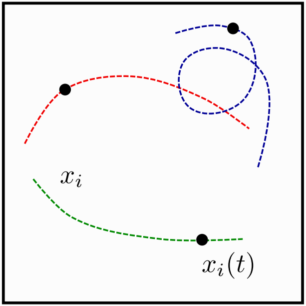
1. By introducing the momentum $m_t := \rho_t v_t$ we have
$$\int_0^1 \int_\Omega |v_t(x)|^2 \, \mu_t(x) \, dx \, dt = \int_0^1 \int_\Omega \frac{|m_t(x)|^2}{\mu_t(x)} \, dx \, dt$$
which is **convex** in (μ_t, m_t)
2. The continuity equation becomes **linear**
$$\partial_t \mu_t + \operatorname{div} m_t = 0$$
3. We know the full trajectory μ_t and can recover the velocity field v_t from m_t

[1] Benamou, Brenier. **Numerische Mathematik** (2000)

Optimal Transport Regularization

Trajectories: Curve of measures

$$t \mapsto \mu_t \in \mathcal{M}(\Omega)\,, \qquad t \in [0,1]$$



$$\mu_t = \sum_i \delta_{x_i(t)}$$

[2] Bredies, **Fanzon**. **ESAIM: Mathematical Modelling and Numerical Analysis** (2020)

Optimal Transport Regularization

Trajectories: Curve of measures

$$t \mapsto \mu_t \in \mathcal{M}(\Omega)\,, \qquad t \in [0,1]$$

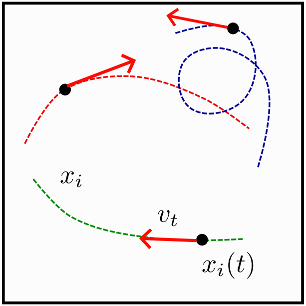
Assumptions:

- μ_t satisfies **Continuity Equation**

$$\partial_t \mu_t + \mathrm{div}(v_t \mu_t) = 0$$

for some velocity field (to find)

$$v_t : \mathbb{R}^2 \to \mathbb{R}^2$$



$$\mu_t = \sum_i \delta_{x_i(t)}$$

[2] Bredies, Fanzon. **ESAIM: Mathematical Modelling and Numerical Analysis** (2020)

Optimal Transport Regularization

Trajectories: Curve of measures

$$t \mapsto \mu_t \in \mathcal{M}(\Omega)\,, \qquad t \in [0,1]$$

Assumptions:

- μ_t satisfies **Continuity Equation**

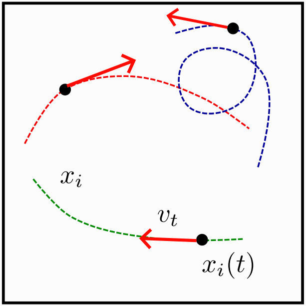
  $$\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0$$

  for some velocity field (to find)

  $$v_t : \mathbb{R}^2 \to \mathbb{R}^2$$

- Finite **Kinetic Energy**

  $$\int_0^1 \int_{\mathbb{R}^2} |v_t(x)|^2 \, d\mu_t(x) \, dt < \infty$$



$$\mu_t = \sum_i \delta_{x_i(t)}$$

[2] Bredies, Fanzon. **ESAIM: Mathematical Modelling and Numerical Analysis** (2020)

Optimal Transport Regularization

Minimization Problem: Given data $t \mapsto y_t \in H_t$

$$K_t \mu_t = y_t \quad \rightsquigarrow \quad \min_{\mu, v} \; L(\mu) + R(\mu, v)$$

[2] Bredies, **Fanzon**. **ESAIM: Mathematical Modelling and Numerical Analysis** (2020)

Optimal Transport Regularization

Minimization Problem: Given data $t \mapsto y_t \in H_t$

$$K_t \mu_t = y_t \quad \rightsquigarrow \quad \min_{\mu, v} \; L(\mu) + R(\mu, v)$$

- $L \rightsquigarrow$ **Loss Function:** Fits $t \mapsto \mu_t$ to data $t \mapsto y_t$ (Generalized Bochner spaces [2])

$$L(\mu) := \int_0^1 \|K_t \mu_t - y_t\|_{H_t}^2 \, dt$$

[2] Bredies, **Fanzon**. **ESAIM: Mathematical Modelling and Numerical Analysis** (2020)

Optimal Transport Regularization

Minimization Problem: Given data $t \mapsto y_t \in H_t$

$$K_t \mu_t = y_t \quad \rightsquigarrow \quad \min_{\mu, v} \; L(\mu) + R(\mu, v)$$

- $L \rightsquigarrow$ **Loss Function:** Fits $t \mapsto \mu_t$ to data $t \mapsto y_t$ (Generalized Bochner spaces [2])

$$L(\mu) := \int_0^1 \|K_t \mu_t - y_t\|_{H_t}^2 \; dt$$

- $R \rightsquigarrow$ **Regularizer:** Connected to **Optimal Transport** (Benamou-Brenier formula)

$$R(\mu, v) := \underbrace{\int_0^1 \int_\Omega |v_t(x)|^2 \; d\mu_t(x) \, dt}_{\text{Kinetic Energy}} + \underbrace{\int_0^1 \|\mu_t\|_{\mathcal{M}(\Omega)} \; dt}_{\text{Radon Norm}}$$

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

- **Theorem [2]:** R provides stable regularization for the dynamic inverse problem

[2] Bredies, **Fanzon**. **ESAIM: Mathematical Modelling and Numerical Analysis** (2020)

Extremal Points

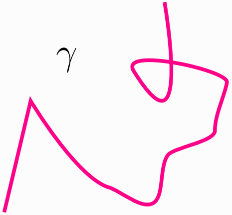
$$R(\mu, v) := \int_0^1 \int_\Omega |v_t(x)|^2 \, d\mu_t(x) \, dt + \int_0^1 \|\mu_t\|_{\mathcal{M}(\Omega)} \, dt$$

$$\text{s.t.} \quad \partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0 \qquad \text{(CE)}$$

Theorem [1]

Let $B = \{R \leq 1\}$. Then $\operatorname{Ext}(B)$ are measures $t \mapsto \mu_t$ supported on **Sobolev Curves**

$$t \mapsto \mu_t = \delta_{\gamma(t)} \, , \qquad \gamma \in H^1([0,1]; \Omega)$$



Superposition Principle [2]: μ_t solves (CE) $\iff \exists \, \sigma \in P(H^1)$ concentrated on sol. to

$$\dot{\gamma}(t) = v(t, \gamma(t)) \ \text{ and s.t. } \int_\Omega \varphi \, d\mu_t = \int_{H^1} \varphi(\gamma(t)) \, d\sigma(\gamma) \, , \ \forall \, \varphi \in C(\Omega)$$

Proof Idea: μ_t extr. for $R \overset{\text{SP}}{\iff} \sigma$ extr. for $\|\cdot\|_{\mathcal{M}(H^1)} \overset{\text{Known}}{\iff} \sigma = \delta_\gamma \overset{\text{SP}}{\iff} \mu_t = \delta_{\gamma(t)}$

[1] Bredies, Carioni, Fanzon, Romero. **Bulletin London Mathematical Society** (2021)
[2] Ambrosio. **Inv. Math.** (2004)

Non-homogeneous case

- Homogeneous continuity equation implies mass preservation

$$\mu_t(\Omega) \quad \text{is constant for all } t$$

- Restrictive in applications – e.g. contrast agent in MRI
- Modify the regularizer to allow change of mass
- Based on **Unbalanced OT** – a.k.a. Hellinger-Kantorovich distance [2,3]

$$R(\mu, v) := \int_0^1 \int_\Omega |v_t(x)|^2 + |g_t(x)|^2 \, d\mu_t(x) \, dt + \int_0^1 \|\mu_t\|_{\mathcal{M}(\Omega)} \, dt$$

$$\text{s.t.} \quad \partial_t \mu_t + \operatorname{div}(v_t \mu_t) = g_t \mu_t \qquad \text{(CE)}$$

Theorem [1]: R is stable regularizer for the dynamic inverse problem

[1] Bredies, Fanzon. **ESAIM: Mathematical Modelling and Numerical Analysis** (2020)
[2] Chizat, Peyré, Schmitzer, Vialard. **Found. of Comp. Math** (2018)
[3] Liero, Mielke, Savaré. **Inv. Math.** (2018)

Extremal Points – Non-homogeneous case

$$R(\mu, v) := \int_0^1 \int_\Omega |v_t(x)|^2 + |g_t(x)|^2 \, d\mu_t(x) \, dt + \int_0^1 \|\mu_t\|_{\mathcal{M}(\Omega)} \, dt$$

$$\text{s.t.} \quad \partial_t \mu_t + \operatorname{div}(v_t \mu_t) = g_t \mu_t \qquad \text{(CE)}$$

Theorem [1]

Let $B = \{R \leq 1\}$. Then $\operatorname{Ext}(B)$ are curves of measures of the form

$$t \mapsto \mu_t = h(t)\delta_{\gamma(t)}$$

- $h, \sqrt{h} \in \mathrm{AC}^2(0,1)$, $\gamma \in C(\{h > 0\}; \Omega)$, $\sqrt{h}\gamma \in \mathrm{AC}^2([0,1]; \mathbb{R}^d)$
- $\{h > 0\}$ is connected

Proof Idea: Novel Probabilistic Superposition Principle to (CE)

[1] Bredies, Carioni, Fanzon. **Communications in PDEs** (2022)

Numerical optimization

Dynamic IP: Given $t \mapsto y_t \in H_t$ find $t \mapsto \mu_t \in \mathcal{M}(\Omega)$ s.t.

$$K_t \mu_t = y_t \quad \text{for all} \quad t \in [0,1]$$

Optimal Transport Regularization: $\min_{\mu,v} \; L(\mu) + R(\mu, v)$

$$L = \int_0^1 \|K_t\mu_t - y_t\|_{H_t}^2 \,, \qquad R = \int_0^1 \int_\Omega |v_t(x)|^2 \, dt + \int_0^1 \|\mu_t\|_{\mathcal{M}(\Omega)} \, dt$$

$$\text{s.t. } \partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0$$

- **Given:** $\Omega \subset \mathbb{R}^d$ and $t \mapsto y_t \in H_t$ data
- **Forw. operator:** $K \colon \mathcal{M}(\Omega) \to H_t$
- **Extr. points:** $B = \{R \leq 1\}$

$$\operatorname{Ext}(B) = \left\{ t \mapsto \delta_{\gamma(t)} : \gamma \in H^1([0,1]; \Omega) \right\}$$

Generalized Frank-Wolfe [1]

$$\min_{\mu,v} \int_0^1 \|K_t\mu_t - y_t\|_{H_t}^2 + \int_0^1 \int_\Omega |v_t(x)|^2 \, dt + \int_0^1 \|\mu_t\|_{\mathcal{M}(\Omega)} \, dt$$

s.t. $\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0$

Algorithm: **Initialize:** $\mu^0 = 0$ **Iterate:** Given $\mu^k = \sum_{i=1}^n \lambda_i \delta_{\gamma_i}$

❶ **Insertion Step:** $p_t^k = K_*(K\mu_t^k - y_t)$ $\quad p_t^k \in L^\infty([0,1]; C(\Omega))$

$$\max_{w \in \mathrm{Ext}(B)} \langle p^k, w\rangle = \max_{\gamma \in H^1([0,1];\Omega)} \left(\int_0^1 |\dot\gamma(t)|^2 \, dt + 1 \right)^{-1} \int_0^1 p_t^k(\gamma(t)) \, dt \quad \rightsquigarrow \quad \hat\gamma$$

❷ **Fully-corrective Step:** Solve

$$\lambda_i^* \in \operatorname*{arg\,min}_{\lambda_i \geq 0} \; G\left(\mu^k + \lambda_{n+1}\delta_{\hat\gamma}\right) \quad \rightsquigarrow \quad \mu^{k+1} := \left(\sum_{i=1}^n \lambda_i^* \delta_{\gamma_i}\right) + \lambda_{n+1}^* \delta_{\hat\gamma}$$

[1] Bredies, Carioni, Fanzon, Romero. **Found. of Computational Mathematics** (2023)

Convergence Analysis

Theorem [1]

μ^k sparse iterate from the Generalized Frank-Wolfe Algorithm. Then

$$\mu^k \overset{*}{\rightharpoonup} \bar{\mu}\,, \qquad \bar{\mu} \in \arg\min\, G\,, \quad G := L + R$$

General convergence rate is **sublinear** (expected for gradient methods)

$$G(\mu^k) - \min G \lesssim \frac{1}{k}$$

Work in Progress: $\bar{\mu}$ **sparse** $+$ **"Source Condition"** $+$ **"Quadratic Growth"**

$\Longrightarrow$ **linear** convergence: $$G(\mu^k) - \min G \lesssim \frac{1}{2^k}$$

[1] Bredies, Carioni, **Fanzon**, Romero. **Found. of Computational Mathematics** (2023)

Details and additional tweaks

- Solve the curve insertion problem

$$\hat{\gamma} \in \operatorname*{arg\,max}_{\gamma \in H^1([0,1];\Omega)} \left(\int_0^1 |\dot{\gamma}(t)|^2 \, dt + 1 \right)^{-1} \int_0^1 p_t^k(\gamma(t)) \, dt$$

  via gradient descent with suitable stepsize rule

Theorem [1]

Under suitable regularity assumptions, the gradient descent procedure converges subsequentially to stationary points and strongly in $\mathrm{AC}^2([0,1];\Omega)$.

- Multiple starts with suitable initial guess (crossovers, random curves, etc.) to increase chance to obtain global minimizer
- Multiple insertion $\rightsquigarrow$ insert all obtained stationary points

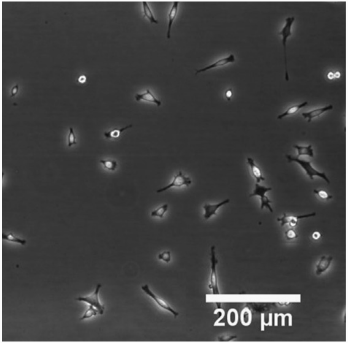
[1] Bredies, Carioni, Fanzon, Romero. **Found. of Computational Mathematics** (2023)

Outline

1. Introduction to Inverse Problems & Sparsity
2. Algorithm for sparse solution recovery
3. Dynamic Inverse Problems
4. **Application to Dynamic MRI**

Imaging Method $\Longrightarrow$

- Stars from ground-based telescope
- Microbubbles in blood vessels
- Cell migration



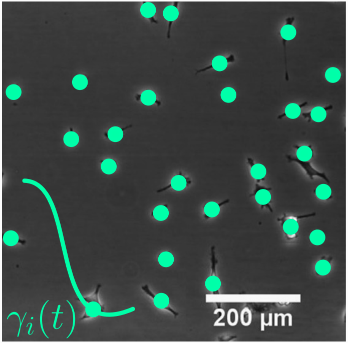
Microscopy image of cells

Image from: Yang, Venkataraman, Styles, et al. **Journal of Biomechanics** (2016)

Motivation: Particle Tracking

Imaging Method $\Longrightarrow$

- Stars from ground-based telescope
- Microbubbles in blood vessels
- Cell migration



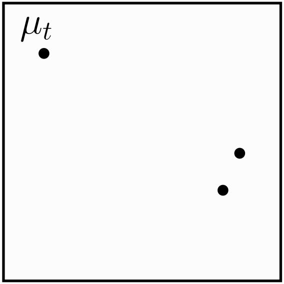
Microscopy image of cells

$$\mu_t = \sum_{i=1}^{M} \delta_{\gamma_i(t)}$$

Image from: Yang, Venkataraman, Styles, et al. **Journal of Biomechanics** (2016)

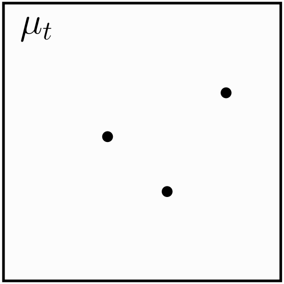
Application: Peak tracking for Dynamic MRI

y_t



$$\mu_t = \sum_{i=1}^{3} \delta_{\gamma_i(t)}$$

Frame-by-Frame: MRI Data $y_t \rightsquigarrow$ Image μ_t

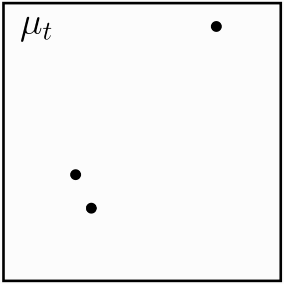


$$\mu_t = \sum_{i=1}^{3} \delta_{\gamma_i(t)}$$

Frame-by-Frame: MRI Data $y_t \rightsquigarrow$ Image μ_t

Application: Peak tracking for Dynamic MRI

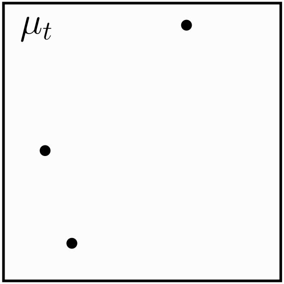
y_t



$$\mu_t = \sum_{i=1}^{3} \delta_{\gamma_i(t)}$$

Frame-by-Frame: MRI Data $y_t \rightsquigarrow$ Image μ_t

y_t

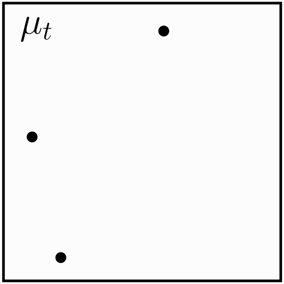


$$\mu_t = \sum_{i=1}^{3} \delta_{\gamma_i(t)}$$

Frame-by-Frame: MRI Data y_t $\rightsquigarrow$ Image μ_t

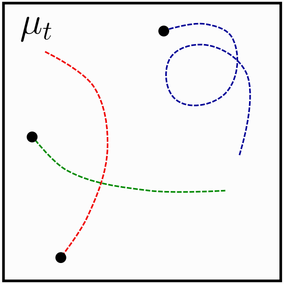
Application: Peak tracking for Dynamic MRI

y_t



$$\mu_t = \sum_{i=1}^{3} \delta_{\gamma_i(t)}$$

Frame-by-Frame: MRI Data y_t $\rightsquigarrow$ Image μ_t

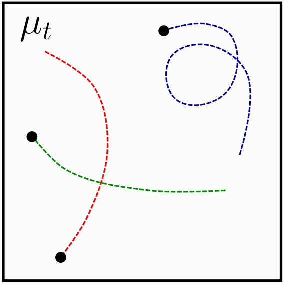


$$\mu_t = \sum_{i=1}^{3} \delta_{\gamma_i(t)}$$

Frame-by-Frame: MRI Data $y_t \rightsquigarrow$ Image $\mu_t \implies$ Interpolate Trajectories

Application: Peak tracking for Dynamic MRI

y_t



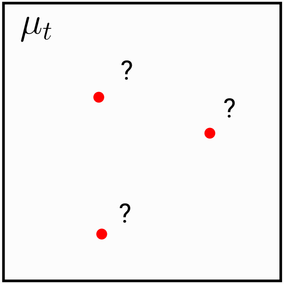
$$\mu_t = \sum_{i=1}^{3} \delta_{\gamma_i(t)}$$

Frame-by-Frame: MRI Data $y_t \rightsquigarrow$ Image $\mu_t \implies$ Interpolate Trajectories

Issue: Motion $\implies$ Low Scan Time $\implies$ **Low Data** $\mathbf{y}_t$

Application: Peak tracking for Dynamic MRI

y_t



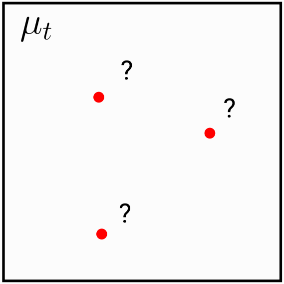
$$\mu_t = \sum_{i=1}^{3} \delta_{\gamma_i(t)}$$

Frame-by-Frame: MRI Data y_t $\rightsquigarrow$ Image μ_t $\Longrightarrow$ Interpolate Trajectories

Issue: Motion $\Longrightarrow$ Low Scan Time $\Longrightarrow$ **Low Data** y_t $\rightsquigarrow$ **Particles?**

Application: Peak tracking for Dynamic MRI

y_t



$$\mu_t = \sum_{i=1}^{3} \delta_{\gamma_i(t)}$$

Frame-by-Frame: MRI Data $y_t \rightsquigarrow$ Image $\mu_t \implies$ Interpolate Trajectories

Issue: Motion $\implies$ Low Scan Time $\implies$ **Low Data** $y_t \rightsquigarrow$ **Particles?**

Global-in-Time: Full Time-Series $t \mapsto y_t \rightsquigarrow$ Trajectories $t \mapsto \mu_t$

The Dynamic Undersampled MRI problem

Dynamic IP MRI: Given $t \mapsto y_t \in \mathbb{C}^{M_t}$ find $t \mapsto \mu_t \in \mathcal{M}(\Omega)$ s.t.

$$K_t \mu_t = y_t \quad \text{for all} \quad t \in [0,1]$$

Fourier Transform: For $\mu \in \mathcal{M}(\Omega)$

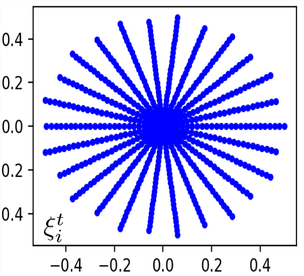
$$\hat{\mu} \colon \mathbb{C} \to \mathbb{C}\,, \quad \hat{\mu}\,[\xi] := \frac{1}{2\pi} \int_{\mathbb{R}^2} e^{i\xi \cdot x}\, d\mu(x)$$

Sampling Frequencies: M_t points

$$\xi_1^t\,, \ldots\,, \xi_{M_t}^t \in \mathbb{C}$$

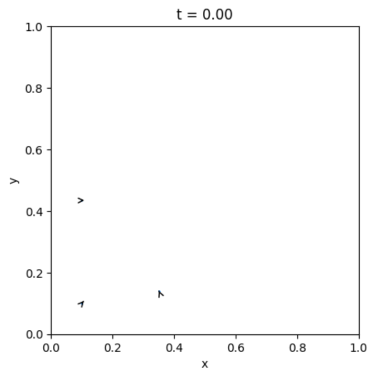
Forward operators: linear cont. $K_t \colon \mathcal{M}(\Omega) \to \mathbb{C}^{M_t}$

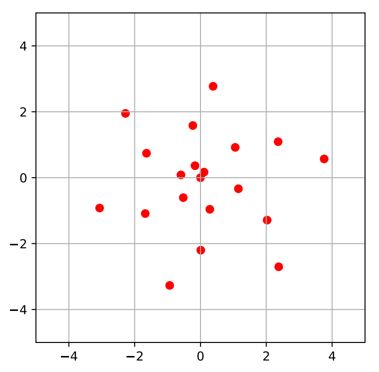
$$K_t \mu := \left(\hat{\mu}[\xi_1^t]\,, \ldots\,, \hat{\mu}[\xi_{M_t}^t]\right)$$



Example: Radial Sampling

Experiment 1: Spiral sampling (static)





Ground truth: Curve of measures

$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)} + \delta_{\gamma_3(t)}$$

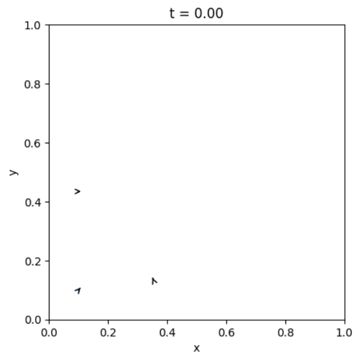
Sampling Freq: $\xi_1, \ldots, \xi_{20}$

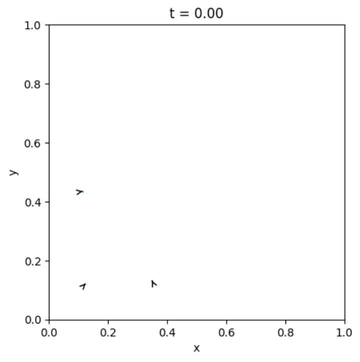
Data: Defined by

$$y_t := K\bar{\mu}_t + 20\% \text{ Noise}$$

Experiment 1: Spiral sampling (static)

Algorithm: Generalized Frank-Wolfe $\rightsquigarrow \quad t \mapsto \mu_t^k = \sum_{i=1}^{M} \lambda_i \, \delta_{\gamma_i(t)}$





Ground truth: Curve of measures

$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)} + \delta_{\gamma_3(t)}$$

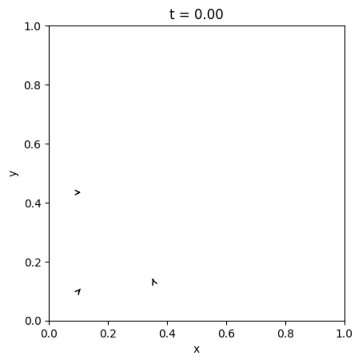
Reconstruction: from data

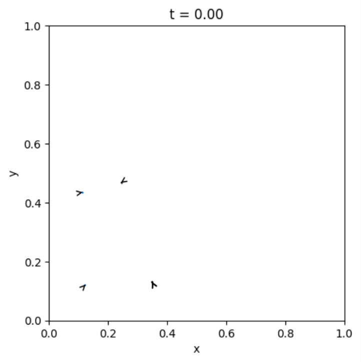
$$y_t = K\bar{\mu}_t + 20\% \text{ Noise}$$

(Thresholded at 0.05)

Experiment 1: Spiral sampling (static)

Algorithm: Generalized Frank-Wolfe $\rightsquigarrow \quad t \mapsto \mu_t^k = \sum_{i=1}^{M} \lambda_i \, \delta_{\gamma_i(t)}$





Ground truth: Curve of measures

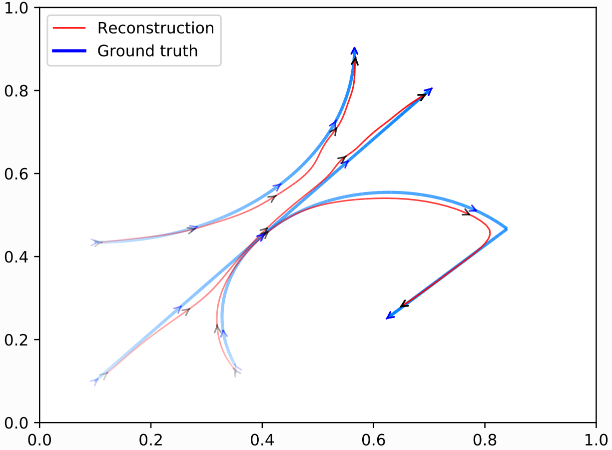
$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)} + \delta_{\gamma_3(t)}$$

Reconstruction: from data

$$y_t = K\bar{\mu}_t + 20\% \text{ Noise}$$

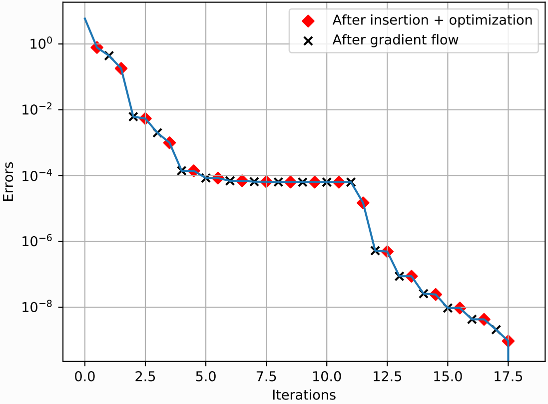
(No Thresholding)

Experiment 1: Spiral sampling (static)



Reconstructed trajectories

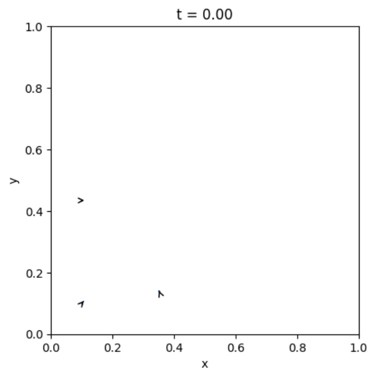
Experiment 1: Spiral sampling (static)

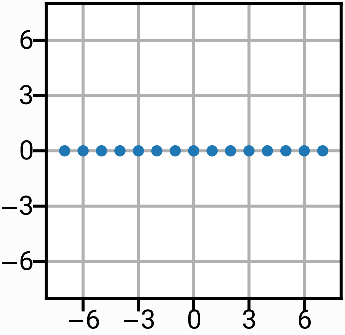


Convergence plot: exhibits linear rate

$$\text{Error} = G(\mu^k) - G(\mu^{k+1})$$

Experiment 2: Dynamic sampling on lines





$t = 0$

Ground truth: Curve of measures

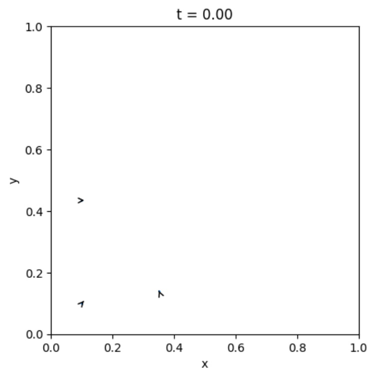
$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)} + \delta_{\gamma_3(t)}$$

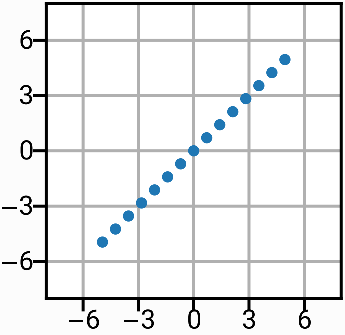
Sampling Freq: $\xi_1^t, \dots, \xi_{15}^t$

Data: Defined by

$$y_t := K_t \bar{\mu}_t + 20\% \text{ Noise}$$

Experiment 2: Dynamic sampling on lines





$t = 1/50$

Ground truth: Curve of measures

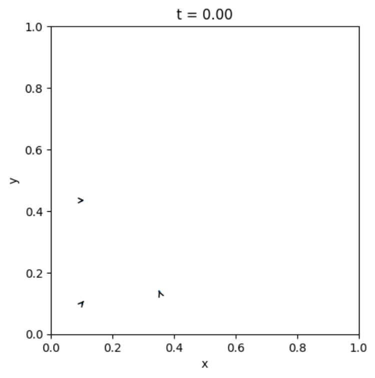
$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)} + \delta_{\gamma_3(t)}$$

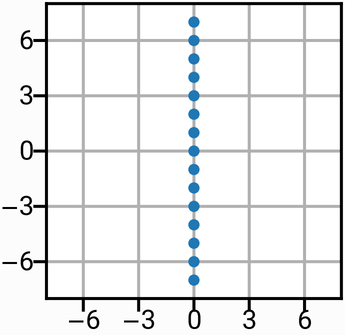
Sampling Freq: $\xi_1^t, \dots, \xi_{15}^t$

Data: Defined by

$$y_t := K_t \bar{\mu}_t + 20\% \text{ Noise}$$

Experiment 2: Dynamic sampling on lines





$t = 2/50$

Ground truth: Curve of measures

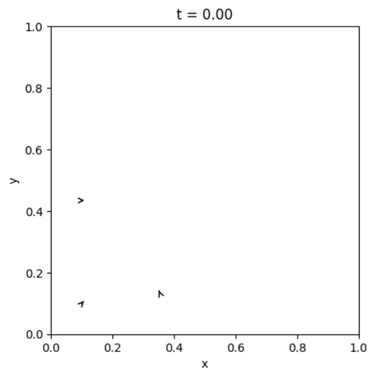
$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)} + \delta_{\gamma_3(t)}$$

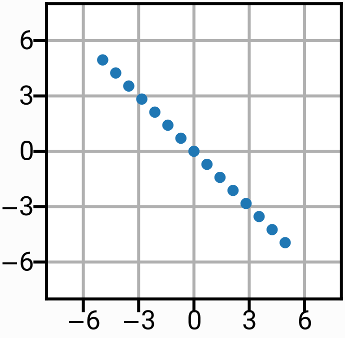
Sampling Freq: $\xi_1^t, \dots, \xi_{15}^t$

Data: Defined by

$$y_t := K_t \bar{\mu}_t + 20\% \text{ Noise}$$

Experiment 2: Dynamic sampling on lines





$t = 3/50$

Ground truth: Curve of measures

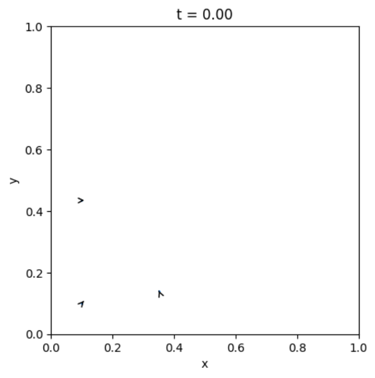
$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)} + \delta_{\gamma_3(t)}$$

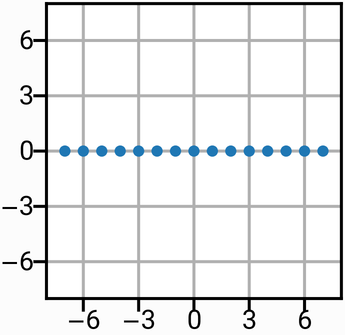
Sampling Freq: $\xi_1^t, \ldots, \xi_{15}^t$

Data: Defined by

$$y_t := K_t \bar{\mu}_t + 20\% \text{ Noise}$$

Experiment 2: Dynamic sampling on lines





$t = 4/50$

Ground truth: Curve of measures

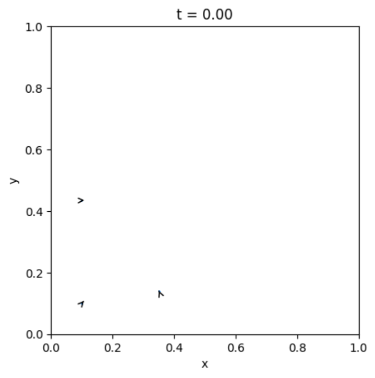
$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)} + \delta_{\gamma_3(t)}$$

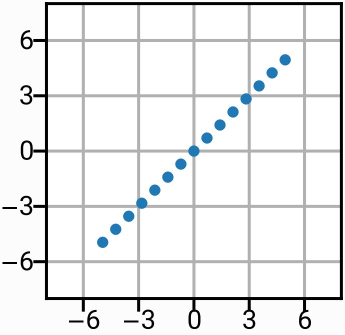
Sampling Freq: $\xi_1^t, \dots, \xi_{15}^t$

Data: Defined by

$$y_t := K_t \bar{\mu}_t + 20\%$$ Noise

Experiment 2: Dynamic sampling on lines





$t = 5/50$

Ground truth: Curve of measures

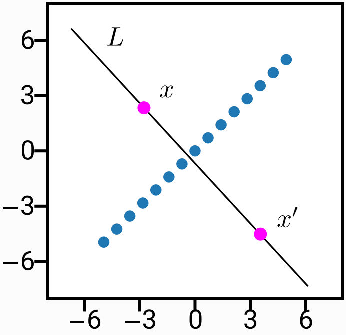
$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)} + \delta_{\gamma_3(t)}$$

Sampling Freq: $\xi_1^t, \dots, \xi_{15}^t$

Data: Defined by

$$y_t := K_t \bar{\mu}_t + 20\% \text{ Noise}$$

Experiment 2: Dynamic sampling on lines



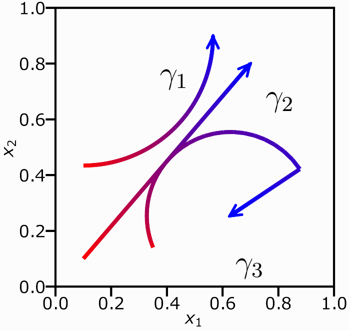
- L line orthogonal to the line of sampling frequences at time t

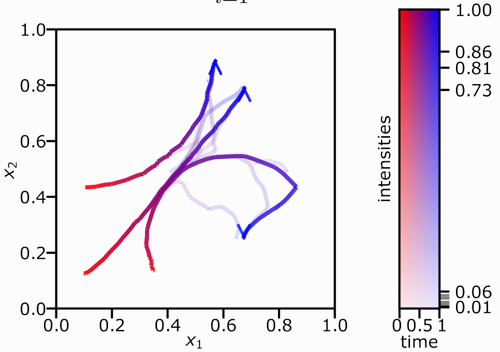
$$x, x' \in L \quad \Longrightarrow \quad K_t \delta_x = K_t \delta_{x'}$$

- Impossible to locate source along L at time t
- Only way to locate source is to enforce time regularity
- Example showcases need for time regularization

Experiment 2: Dynamic sampling on lines

Algorithm: Generalized Frank-Wolfe $\rightsquigarrow \quad t \mapsto \mu_t^k = \sum_{i=1}^{M} \lambda_i \, \delta_{\gamma_i(t)}$





Ground truth: Curve of measures

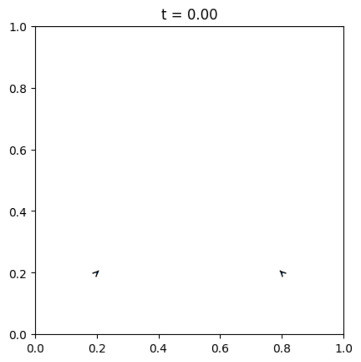
$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)} + \delta_{\gamma_3(t)}$$

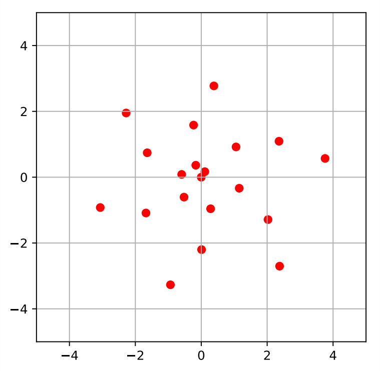
Reconstruction: from data

$$y_t = K_t \bar{\mu}_t + 20\% \text{ Noise}$$

Remarkable reconstructions – considering unfavorable sampling pattern

Experiment 3: Crossing





Ground truth: Curve of measures

$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)}$$

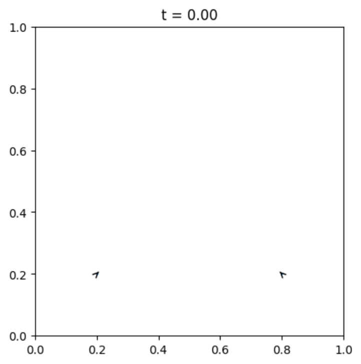
Sampling Freq: $\xi_1, \dots, \xi_{20}$

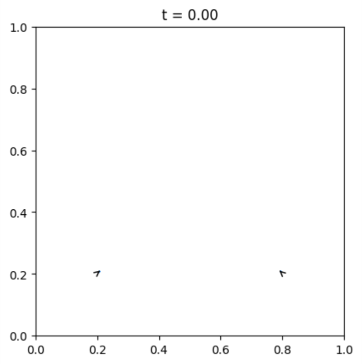
Data: Defined by

$$y_t := K\bar{\mu}_t + 20\% \text{ Noise}$$

Experiment 3: Crossing

Algorithm: Generalized Frank-Wolfe $\rightsquigarrow \quad t \mapsto \mu_t^k = \sum_{i=1}^{M} \lambda_i \, \delta_{\gamma_i(t)}$





Ground truth: Curve of measures

$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)}$$

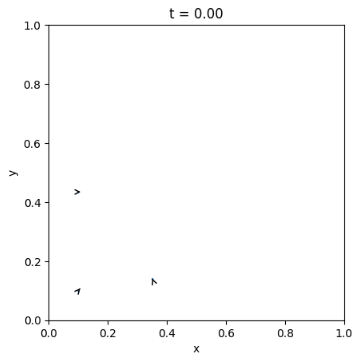
Reconstruction: from data

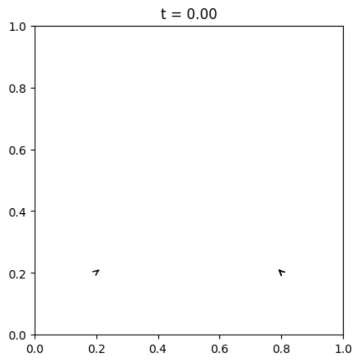
$$y_t = K\bar{\mu}_t + 20\% \text{ Noise}$$

(Thresholded at 0.01)

Experiment 3: Crossing

Algorithm: Generalized Frank-Wolfe $\rightsquigarrow \quad t \mapsto \mu_t^k = \sum_{i=1}^{M} \lambda_i \, \delta_{\gamma_i(t)}$





Ground truth: Curve of measures

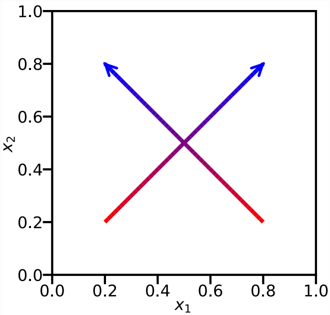
$$\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)}$$

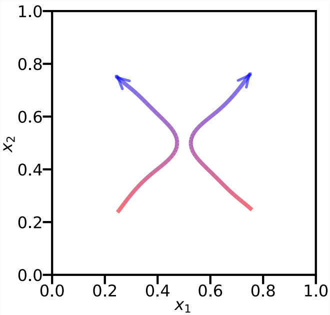
Reconstruction: from data

$$y_t = K\bar{\mu}_t + 20\% \text{ Noise}$$

(No Thresholding)

Experiment 3: Crossing





Ground truth: $\bar{\mu}_t := \delta_{\gamma_1(t)} + \delta_{\gamma_2(t)}$ **Reconstruction:** $\tilde{\mu}_t := \delta_{\tilde{\gamma}_1(t)} + \delta_{\tilde{\gamma}_2(t)}$

Question: Why do reconstructed trajectories differ from ground truth ones?

Answer: They don't! When regarded as measures they are basically the same

$$dt \otimes \bar{\mu}_t \approx dt \otimes \tilde{\mu}_t$$

Regularizer is Dynamic OT $\implies$ Particles chose shortest path

What to do? Maybe could include curvature penalization

Conclusion

1. Algorithm for computing **sparse** solutions to

$$\min_{u \in X} \; F(Ku) + R(u)$$

in **Banach** space

Conclusion

1. Algorithm for computing **sparse** solutions to

$$\min_{u \in X} F(Ku) + R(u)$$

   in **Banach** space

2. **Linear** convergence if solution is **Sparse** + **"Source Condition"** + **"Quadratic Growth"**

Conclusion

1. Algorithm for computing **sparse** solutions to

$$\min_{u \in X} F(Ku) + R(u)$$

   in **Banach** space

2. **Linear** convergence if solution is **Sparse** + **"Source Condition"** + **"Quadratic Growth"**

3. General framework for **dynamic** inverse problems

Conclusion

❶ Algorithm for computing **sparse** solutions to

$$\min_{u \in X} \; F(Ku) + R(u)$$

in **Banach** space

❷ **Linear** convergence if solution is **Sparse** + **“Source Condition”** + **“Quadratic Growth”**

❸ General framework for **dynamic** inverse problems

❹ Application to **Dynamic MRI**

Thank You!

References

Generalized Frank-Wolfe Algorithm

[1] Bredies, Carioni, **Fanzon**, Walter. **Mathematical Programming** (2024)

Particles Tracking + Dynamic Inverse Problems

[2] **Fanzon**, Bredies. **ESAIM: Mathematical Modelling and Numerical Analysis** (2020)

[3] Bredies, Carioni, **Fanzon**, Romero. **Bulletin London Mathematical Society** (2021)

[4] Bredies, Carioni, **Fanzon**. **Communications in PDEs** (2022)

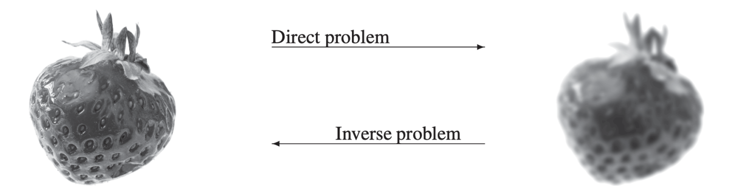
[5] Bredies, Carioni, **Fanzon**, Romero. **Found. of Computational Mathematics** (2023)

FWF

Christian Doppler Forschungsgesellschaft

Supported by Austrian Science Fund (FWF) and
Christian Doppler Research Society (CDG) Project PIR27
“Mathematical methods for motion-aware medical imaging”

Infinite dimensional Example: Deblurring



Original Image $u\colon \Omega \to \mathbb{R}$ Blurred image $f\colon \Omega \to \mathbb{R}$

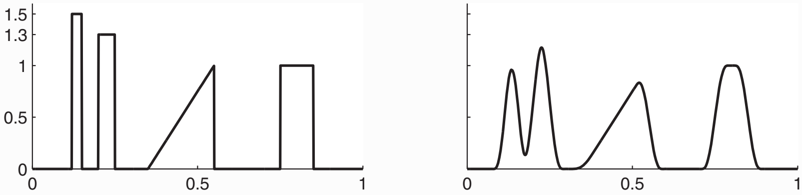
- Deblurring can be achieved by **deconvolution**

  Solve $\quad Ku = f\,, \qquad (Ku)(x) = \displaystyle\int_{\mathbb{R}^2} \psi(y)u(x-y)\,dy = (\psi \star u)(x)$

  with ψ suitable kernel (e.g. point-spread function)

- $K\colon L^2(\Omega) \to L^2(\Omega)$ compact operator $\implies$ K^{-1} unbounded (ill-posed)

Simpler case: 1D deconvolution

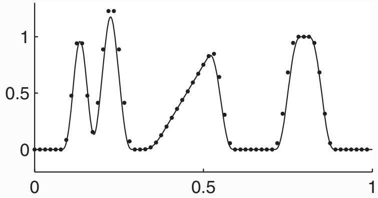


Original signal $\tilde{u} \colon [0,1] \to \mathbb{R}$

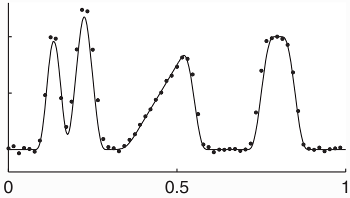
Blurred signal $f = \psi \star \tilde{u}$

- Goal: Recover $\tilde{u}$ from noisy data $f^{\varepsilon} = f + \varepsilon$
- This means solving the 1D-deconvolution problem: Find u such that

$$\psi \star u = f^{\varepsilon}$$



Discrete f with $n = 64$ grid points



Add 1% noise to obtain $f^\varepsilon \in \mathbb{R}^{64}$

- The convolution $\psi \star u$ can be discretized using Riemann sums
- The discrete inverse problem is therefore: Find $u \in \mathbb{R}^{64}$ such that

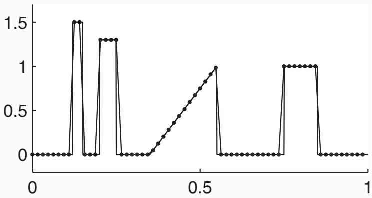
$$Ku = f^\varepsilon \,, \qquad K \in \mathbb{R}^{64 \times 64}$$

Naive deconvolution

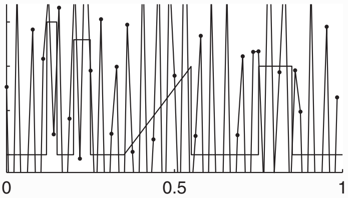
Solve the discrete 1D-deconvolution problem: Find $u \in \mathbb{R}^{64}$ such that

$$Ku = f^\varepsilon \,, \qquad K \in \mathbb{R}^{64\times 64}$$

- The naive solution is $u = K^{-1} f^\varepsilon$
- This behaves well when $\varepsilon = 0$ but is terrible when $\varepsilon \neq 0$
- Below the solid line represents the ground truth $\tilde{u}$
- We need regularizer which penalizes oscillations



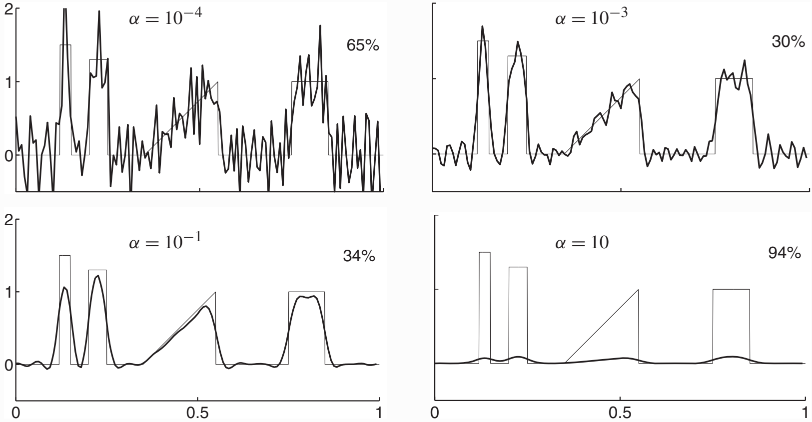
$$u = K^{-1} f$$



$$u = K^{-1} f^\varepsilon$$

Regularize the discrete inverse problem with the ℓ_2 norm:

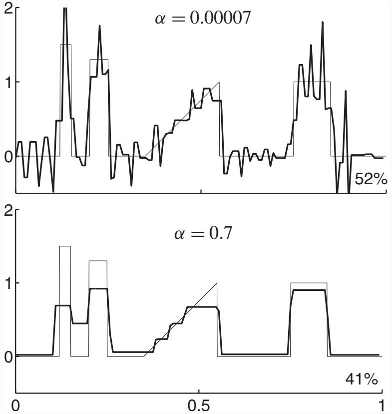
$$\min_{u \in L^2(0,1)} \|Ku - f^\varepsilon\|^2_{L^2(0,1)} + \alpha \|u'\|^2_{L^2(0,1)}$$



Best is $\alpha = 10^{-1}$. Notice the smoothing effect of ℓ_2 Smoothness is not always desirable (e.g. if u is image with sharp edges)

Regularize inverse problem with the Total Variation (BV) semi-norm:

$$\min_{u \in L^1(0,1)} \|Ku - f^\varepsilon\|^2_{L^2(0,1)} + \alpha \, \mathrm{TV}(u)$$



Best is $\alpha = 0.009$. Notice the sparsifying effect of TV (the jumps)

Elementary example: Matrix inversion

Given $f \in \mathbb{R}^m$ and a matrix $K \in \mathbb{R}^{m \times n}$ we want to find $u \in \mathbb{R}^n$ such that

$$Ku = f + \varepsilon \tag{P}$$

What could go wrong:

1. $m > n \implies$ Range$(K) \neq \mathbb{R}^m \implies$ No solution when $f + \varepsilon \notin$ Range(K)
2. $m < n \implies$ ker$(K) \neq \{0\} \implies$ There are several solutions
3. $m = n$ and K^{-1} exists: However condition number $\kappa = \lambda_1 / \lambda_n$ could be large. Then K is almost singular and

$$\left\| K^{-1} \varepsilon \right\| \approx \frac{\|\varepsilon\|}{\lambda_n} \quad \implies \quad \text{Naive reconstruction is dominated by noise}$$

$$\tilde{u} = u + K^{-1} \varepsilon \quad \implies \quad \text{instability}$$

Therefore (P) is in general **ill-posed**

Least-squares solution

Given $f \in \mathbb{R}^m$ and a matrix $K \in \mathbb{R}^{m \times n}$ we want to find $u \in \mathbb{R}^n$ such that

$$Ku = f \tag{P}$$

- (P) might not have solution. Find approximate solution by **least-squares**

$$\min_{u \in \mathbb{R}^n} \; \|Ku - f\|_2^2 \tag{P'}$$

  with $\|\cdot\|_2$ the Euclidean norm

- (P') always has the explicit solution (seen by differentiation)

$$\tilde{u} = (K^T K)^{-1} K^T f$$

- **Problem 1:** Solution to (P') **not unique** (if K is not injective)
- **Problem 2:** Solution might be **instable** (depends on eigenvalues of $K^T K$)

Variational Regularization

Given $f \in \mathbb{R}^m$ and a matrix $K \in \mathbb{R}^{m \times n}$ we want to find $u \in \mathbb{R}^n$ such that

$$Ku = f \tag{P}$$

- **Question:** Non uniqueness and / or instability. What to do?
- **Answer:** Replace (P) with the **regularized** least-squares problem

$$\min_{u \in \mathbb{R}^n} \|Ku - f\|_2^2 + \alpha R(u)$$

with $R \colon \mathbb{R}^n \to [0, +\infty]$ **regularizer** and $\alpha > 0$ to be chosen

1. R promotes stability (if chosen properly)
2. R selects only some solutions (the ones for which $R(u)$ is small)

First Example: ℓ_2 regularization

Regularize using the ℓ^2 norm:

$$\min_{u \in \mathbb{R}^n} \|Ku - f\|_2^2 + \alpha \|u\|_2^2 \tag{P}$$

- (P) is known as Ridge-regression in Statistics
- (P) always has the explicit solution (seen by differentiation)

$$\tilde{u} = (K^T K + \alpha I)^{-1} K^T f$$

- (P) more stable because eigenvalues of $K^T K + \alpha I$ are away from zero
- ℓ_2 norm **shrinks components** $\implies$ mitigates effects of noise

Second Example: ℓ_1 regularization

Regularize using the ℓ^1 norm

$$\min_{u \in \mathbb{R}^n} \|Ku - f\|_2^2 + \alpha \|u\|_1 \tag{P}$$

- (P) is known as LASSO-regression is statistics
- (P) always admits a solution (no explicit formula available)
- ℓ_1 norm automatically **sets some components to zero** $\rightsquigarrow$ **sparsity**

$$\tilde{u} = (0, 0, 0, *, 0, 0, \dots, 0, *, 0, 0, 0, 0)$$

- Desirable when n is large (many parameters), as it simplifies the model

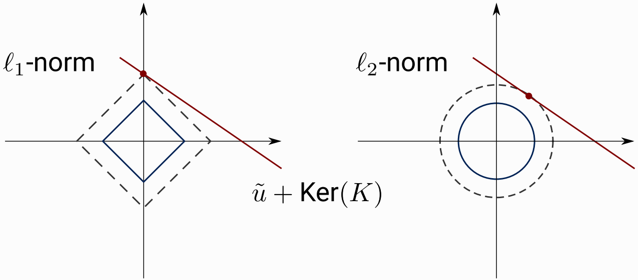
Models like GPT-5 have 10s of trillions of parameters

Why does ℓ_1 set components to zero?

$$\min_{u \in \mathbb{R}^n} \|Ku - f\|_2^2 + \alpha \|u\|_1 \qquad \min_{u \in \mathbb{R}^n} \|Ku - f\|_2^2 + \alpha \|u\|_2^2$$

$$\Updownarrow \qquad \Updownarrow$$

$$\min_{\|u\|_1 \leq s} \|Ku - f\|_2^2 \qquad \min_{\|u\|_2 \leq s} \|Ku - f\|_2^2$$



Extremal points are different $\Longrightarrow$ ℓ_1 and ℓ_2 select different solutions

Extremal points of regularizer describe features of sparse solutions

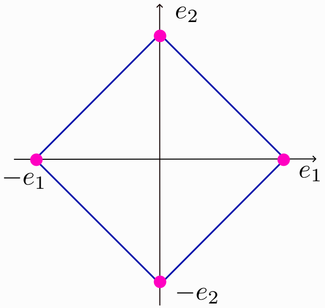
Example: Portfolio Optimization

Portfolio: Vector

$$P = (w_1, \dots, w_d) \qquad w_i = \text{capital to invest in asset } i$$

Sparsity: Invest in few assets

$$P = (0, 0, \mathbf{w_i}, 0, 0, \dots, 0, 0, \mathbf{w_d}) \implies \text{lower managing fees}$$



Banach space: $X = \mathbb{R}^d$

Regularizer: $\|x\|_1 := \sum_{i=1}^{d} |x_i|$

$$\mathrm{Ext}(B) = \{\pm e_i\}_{i=1}^d$$

Extremal Points

$$R(\mu, v) := \int_0^1 \int_\Omega |v_t(x)|^2 \, d\mu_t(x) \, dt + \int_0^1 \|\mu_t\|_{\mathcal{M}(\Omega)} \, dt$$

$$\text{s.t.} \quad \partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0 \qquad \text{(CE)}$$

[3] Bredies, Carioni, **Fanzon**, Romero. **Bulletin London Mathematical Society** (2021)

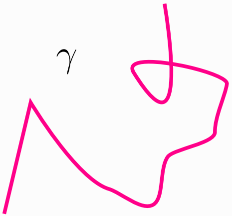
$$R(\mu, v) := \int_0^1 \int_\Omega |v_t(x)|^2 \, d\mu_t(x) \, dt + \int_0^1 \|\mu_t\|_{\mathcal{M}(\Omega)} \, dt$$

s.t. $\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0$ (CE)

Theorem [3]

Let $B = \{R \leq 1\}$. Then $\operatorname{Ext}(B)$ are measures $t \mapsto \mu_t$ supported on **Sobolev Curves**

$$t \mapsto \mu_t = \delta_{\gamma(t)} \,, \qquad \gamma \in H^1([0,1];\mathbb{R}^2)$$



[3] Bredies, Carioni, **Fanzon**, Romero. **Bulletin London Mathematical Society** (2021)

Extremal Points

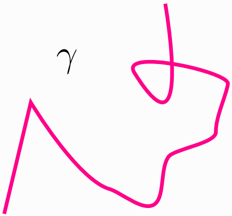
$$R(\mu, v) := \int_0^1 \int_\Omega |v_t(x)|^2 \, d\mu_t(x) \, dt + \int_0^1 \|\mu_t\|_{\mathcal{M}(\Omega)} \, dt$$

$$\text{s.t.} \quad \partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0 \qquad \text{(CE)}$$

Theorem [3]

Let $B = \{R \leq 1\}$. Then $\operatorname{Ext}(B)$ are measures $t \mapsto \mu_t$ supported on **Sobolev Curves**

$$t \mapsto \mu_t = \delta_{\gamma(t)} \, , \qquad \gamma \in H^1([0,1]; \mathbb{R}^2)$$



Proof Idea: Probabilistic Superposition Principle for measure solutions to

$$\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0 \qquad (= g_t \mu_t \;)$$

[3] Bredies, Carioni, **Fanzon**, Romero. **Bulletin London Mathematical Society** (2021)