Recovery of piecewise constant images using total (gradient) variation regularization

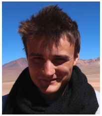
Vincent Duval

INRIA Paris (MOKAPLAN) & U. Paris-Dauphine (CEREMADE)

Journée SIGMA-MODE
Paris

30 janvier 2024

Joint work with

- Romain Petit 
- Yohann De Castro 

. . . and

- Claire Boyer
- Frédéric de Gournay
- Antonin Chambolle
- Pierre Weiss

The signal we want to recover is a superposition of point sources,

$$m_0 = \sum_{i=1}^{N} a_{0,i} \delta_{x_{\mathbf{0},i}}.$$

Many different methods

- Prony [Prony 1795, Kunis et al. '16, Sauer '17...]
- MUSIC [Schmidt 1986, Liao & Fannjiang '14...],
- ESPRIT [Kailath 1990, Li & Liao '20...]
- Matrix pencil [Hua 1988, Sarkar & Peirera 1995, Liu & Ammari '22...] ,
- ...

The Beurling LASSO (BLASSO) [Bredies & Pikkarainen'13, Azaïs & De Castro '14, Candès & Fernandez-Granda'14]

$$\min_{m \in \mathcal{M}(\Omega)} |m|(\Omega) \quad \text{s.t. } \Phi m = y. \qquad (\mathcal{P}_0(y))$$

see the review [Laville et al.'21].

The signal we want to recover is a superposition of point sources,

$$m_0 = \sum_{i=1}^{N} a_{0,i}\delta_{x_{0,i}}.$$

Many different methods

- Prony [Prony 1795, Kunis et al. '16, Sauer '17...]
- MUSIC [Schmidt 1986, Liao & Fannjiang '14...],
- ESPRIT [Kailath 1990, Li & Liao '20...]
- Matrix pencil [Hua 1988, Sarkar & Peirera 1995, Liu & Ammari '22...] ,
- ...

The Beurling LASSO (BLASSO) [Bredies & Pikkarainen'13, Azaïs & De Castro '14, Candès & Fernandez-Granda'14]

$$\min_{m\in\mathcal{M}(\Omega)} |m|(\Omega) + \frac{1}{2\lambda}\|\Phi m - y\|^2 \qquad (\mathcal{P}_\lambda(y))$$

see the review [Laville et al.'21].

Robustness of the support

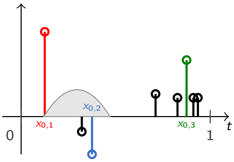
Assume that

- the unknown is $m_0 = \sum_{i=1}^N a_{0,i} \delta_{x_{\mathbf{0},i}}$,
- the observation is $y = \Phi m_0 + w$, with w some noise.

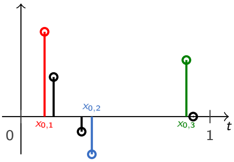
Theorem (D.-Peyré 2015)

*If some **"non-degeneracy"** assumption holds, there exists, $\alpha > 0$, $\lambda_0 > 0$ such that for $0 \leqslant \lambda \leqslant \lambda_0$ and $\|w\|_{\mathcal{H}} \leqslant \alpha\lambda$,*

- *the solution $m_{(\lambda,w)}$ to $\mathcal{P}_\lambda(y+w)$ is unique and has exactly N spikes, $m_{(\lambda,w)} = \sum_{i=1}^N a_i(\lambda,w)\delta_{x_i(\lambda,w)}$,*
- *the mapping $(\lambda, w) \mapsto (a, x)$ is $\mathscr{C}^1$.*



No Support recovery



"Support recovery"

Assume that

- the unknown is $m_0 = \sum_{i=1}^{N} a_{0,i}\delta_{x_{\mathbf{0},i}}$,
- the observation is $y = \Phi m_0 + w$, with w some noise.

Theorem (D.-Peyré 2015)

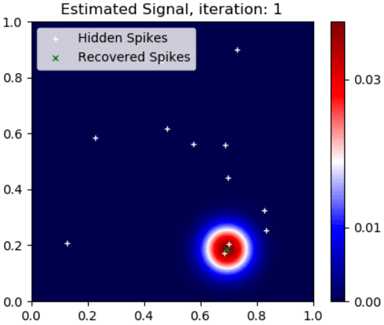
*If some **"non-degeneracy"** assumption holds, there exists, $\alpha > 0$, $\lambda_0 > 0$ such that for $0 \leqslant \lambda \leqslant \lambda_0$ and $\|w\|_{\mathcal{H}} \leqslant \alpha\lambda$,*

- *the solution $m_{(\lambda,w)}$ to $\mathcal{P}_\lambda(y+w)$ is unique and has exactly N spikes, $m_{(\lambda,w)} = \sum_{i=1}^{N} a_i(\lambda, w)\delta_{x_i(\lambda,w)}$,*
- *the mapping $(\lambda, w) \mapsto (a, x)$ is $\mathscr{C}^1$.*

The non-degeneracy holds [Poon et al. 2021] if :

- The kernel $K(x, y) \stackrel{\text{def.}}{=} \langle \Phi\delta_x, \Phi\delta_y \rangle$ is smooth and decays sufficiently as $d(x, y) \nearrow$,
- The locations $x_{0,1}, \ldots, x_{0,N}$ of the spikes are sufficiently separated.

See also the non-negative case [Denoyelle et al.'16, Poon & Peyré'19, D.'20...] and other partly smooth or mirror-stratifiable regularizations [Vaiter et al.'15, Fadili et al.'18...].



(Exp. by Q. Denoyelle)

Frank-Wolfe with local refinement steps[Bredies & Pikkarainen'13, Boyd et al.'15, Denoyelle et al.'18]

- Greedy algorithms
- Non-convex refinement steps (exploits the continuous nature of the problem)
- Yet, global convergence guarantees (convex optimization)

The signal we want to recover is piecewise constant,

$$u_0 = \sum_{i=1}^{N} a_{0,i} \mathbb{1}_{E_{0,i}}.$$

where $\partial E_{0,i}$ is "not too oscillating".

We want to reconstruct cartoon images [Meyer'01, Aujol et al.'05....]

Summary

1. Setting
2. What is the structure of the solutions?
3. Is that structure stable?
4. Algorithm and numerical results

An Inverse Problem

Unknown image $u_0 \in \mathrm{L}^2(\mathbb{R}^2)$

Unknown image $u_0 \in \mathrm{L}^2(\mathbb{R}^2)$

Observations $y_0 = \Phi u_0 \in \mathbb{R}^M$

We assume that

$$\Phi u \stackrel{\text{def.}}{=} \left(\int_{\mathbb{R}^2} u(x)\varphi_i(x)\mathrm{d}x \right)_{1 \leqslant i \leqslant M},$$

with $\{\varphi_i\}_{i=1}^M \subset L^2(\mathbb{R}^2)$.

An Inverse Problem

Unknown image $u_0 \in \mathrm{L}^2(\mathbb{R}^2)$

Observations $y_0 = \Phi u_0 \in \mathbb{R}^M$

We assume that

$$\Phi u \overset{\text{def.}}{=} \left(\int_{\mathbb{R}^2} u(x) \varphi_i(x) \mathrm{d}x \right)_{1 \leqslant i \leqslant M},$$

with $\{\varphi_i\}_{i=1}^M \subset L^2(\mathbb{R}^2)$.

Noisy observations $y = y_0 + w$

An Inverse Problem

Unknown image $u_0 \in \mathrm{L}^2(\mathbb{R}^2)$

Observations $y_0 = \Phi u_0 \in \mathbb{R}^M$

We assume that

$$\Phi u \stackrel{\text{def.}}{=} \left(\int_{\mathbb{R}^2} u(x)\varphi_i(x)\mathrm{d}x \right)_{1 \leqslant i \leqslant M},$$

with $\{\varphi_i\}_{i=1}^M \subset L^2(\mathbb{R}^2)$.

Noisy observations $y = y_0 + w$

Goal

Recover u_0 from y.

The total variation of an image

The **total** (gradient) **variation** of u is

$$\int_{\mathbb{R}^2} |Du| \overset{\text{def.}}{=} \sup\left\{ \int_{\mathbb{R}^2} u(x)\mathrm{div}z(x)\mathrm{d}x \;;\; z \in \mathscr{C}_c^\infty(\mathbb{R}^2), \|z\|_\infty \leqslant 1 \right\}$$
$$= \int_{\mathbb{R}^2} |\nabla u(x)|\,\mathrm{d}x \quad \text{if } u \text{ is smooth.}$$

- If $u = \mathbb{1}_E$ with $E \subseteq \mathbb{R}^2$ of class $\mathscr{C}^1$, then $\int_{\mathbb{R}^2} |D\mathbb{1}_E| = \mathcal{H}^1(\partial E) = P(E)$.
- More generally, if $E \subseteq \mathbb{R}^2$, we *define* its perimeter as $P(E) \overset{\text{def.}}{=} \int_{\mathbb{R}^2} |D\mathbb{1}_E|$.

Variational approach

Following [Rudin et al.'92, Chambolle & Lions '97,...], we consider the problems

- (noiseless setting, $y_0 = \Phi u_0$)

$$\min_{u \in L^2(\mathbb{R}^2)} \int_{\mathbb{R}^2} |Du| \quad \text{s.c.} \quad \Phi u = y_0 \qquad (\mathcal{P}_0(y_0))$$

- (noisy setting, $y = y_0 + w$)

$$\min_{u \in L^2(\mathbb{R}^2)} \int_{\mathbb{R}^2} |Du| + \frac{1}{2\lambda} \left\| \Phi u - y \right\|^2 \qquad (\mathcal{P}_\lambda(y))$$

noisy observations y — solution (well-chosen λ) — solution (high λ)

Summary

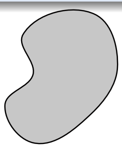
Theorem ([BCCDGW '18, Bredies & Carioni'18])

There is a solution in $\operatorname{argmin} \mathcal{P}_\lambda(y)$ *(resp.* $\operatorname{argmin} \mathcal{P}_0(y_0)$*) of the form*

$$u = \sum_{i=1}^{r} a_i \mathbb{1}_{E_i}.$$

with $r \leqslant M$ *and each* E_i *a simple set.*

The simple sets are the "simply connected" sets in the measure theoretic sense.

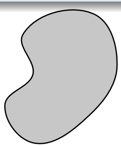


Theorem ([BCCDGW '18, Bredies & Carioni'18])

There is a solution in $\operatorname{argmin} \mathcal{P}_\lambda(y)$ *(resp.* $\operatorname{argmin} \mathcal{P}_0(y_0)$*) of the form*

$$u = \sum_{i=1}^{r} a_i \mathbb{1}_{E_i}.$$

with $r \leqslant M$ *and each* E_i *a simple set.*



The simple sets are the "simply connected" sets in the measure theoretic sense.

Idea of proof:

- Each extreme point of $\operatorname{argmin} \mathcal{P}$ belongs to a face of dimension at most $M-1$ of $\left\{u \in L^2 \ ; \ \int_{\mathbb{R}^2} |Du| \leqslant t\right\}$ where $t = TV(u)$.
- Use Carathéodory's theorem together with

Theorem ([Fleming 1957, Ambrosio et al.'01])

The extreme points of $\left\{u \in L^2 \ ; \ \int_{\mathbb{R}^2} |Du| \leqslant 1\right\}$ *are the functions of the form* $u = \pm \mathbb{1}_E / P(E)$, *where* E *is a* ***simple set****.*

The extreme points of argmin($\mathcal{P}$) are *sums of at most M indicators of simple sets*,

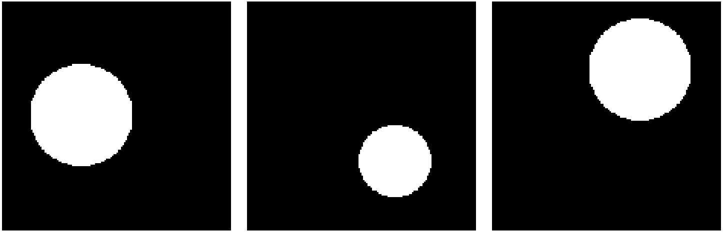
$$u = \sum_{i=1}^{M} a_i \mathbb{1}_{E_i}.$$

TV regularization promotes cartoon images

Representation of the solutions

The extreme points of argmin($\mathcal{P}$) are *sums of at most M indicators of simple sets*,

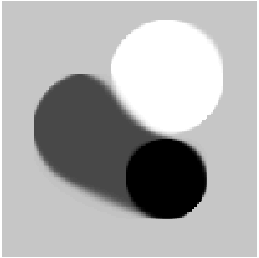
$$u = \sum_{i=1}^{M} a_i \mathbb{1}_{E_i}.$$

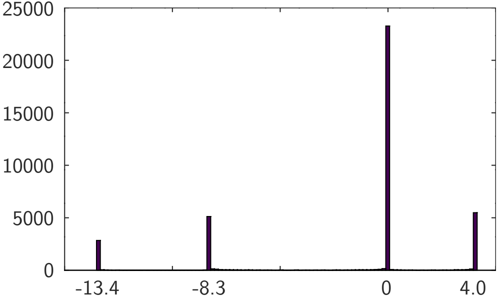


Measurement functions φ_i

The extreme points of argmin($\mathcal{P}$) are *sums of at most M indicators of simple sets*,

$$u = \sum_{i=1}^{M} a_i \mathbb{1}_{E_i}.$$

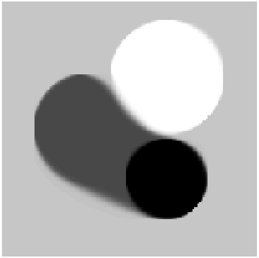


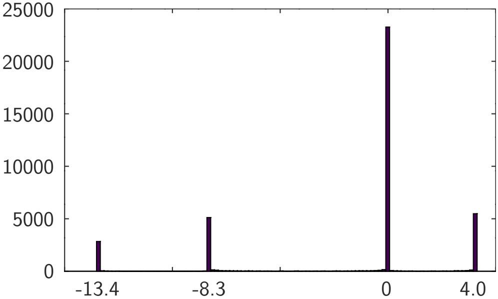


Solution to $\mathcal{P}_0(y_0)$ (left) and its histogram (right)

The extreme points of argmin($\mathcal{P}$) are *sums of at most M indicators of simple sets*,

$$u = \sum_{i=1}^{M} a_i \mathbb{1}_{E_i}.$$





Solution to $\mathcal{P}_0(y_0)$ (left) and its histogram (right)

Why M nonzero values and not $2^M - 1$??

Understanding the faces of the TV unit ball

Let F be a (linearly closed) **face** of

$$C_{\mathrm{BV}} \stackrel{\text{def.}}{=} \left\{ u \in L^2(\mathbb{R}^2) \; ; \; \int_{\mathbb{R}^2} |Du| \leqslant 1 \right\}.$$

What can we say about $u \in F$, if F has dimension k ?

☞ See [Bach 2009, Fujishige 2005] for submodular functions on a finite graph. But, in our case

- C_{BV} is not a polyhedron,
- some faces are not exposed.

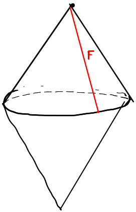
Finite-dimensional faces

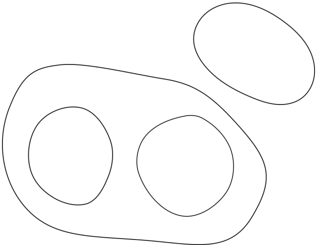
Let F be a k-**dimensional face** of $C_{\mathrm{BV}} \stackrel{\text{def.}}{=} \left\{ u \in L^2(\mathbb{R}^2) \; ; \; \int_{\mathbb{R}^2} |Du| \leqslant 1 \right\}$.

Theorem (D.'22)

- *F is a **polytope** (finite number of extreme points)*
- *Every $u \in F$ takes at most $k+1$ nonzero values.*
- *There is a partition $\{H_i\}_{1 \leqslant i \leqslant k+2}$ of $\mathbb{R}^2$ with H_i indecomposable, such that every $u \in F$ is constant on each H_i,*

$$u = \sum_{i=1}^{k+1} t_i \mathbb{1}_{H_i}$$





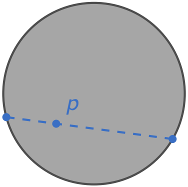
Finite-dimensional faces

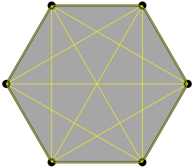
Let F be a k-**dimensional face** of $C_{\mathrm{BV}} \stackrel{\text{def.}}{=} \left\{u \in L^2(\mathbb{R}^2) \; ; \; \int_{\mathbb{R}^2} |Du| \leqslant 1\right\}$.

Theorem (D.'22)

- *F is a **polytope** (finite number of extreme points)*
- *Every $u \in F$ takes at most $k+1$ nonzero values.*
- *There is a partition $\{H_i\}_{1 \leqslant i \leqslant k+2}$ of $\mathbb{R}^2$ with H_i indecomposable, such that every $u \in F$ is constant on each H_i,*

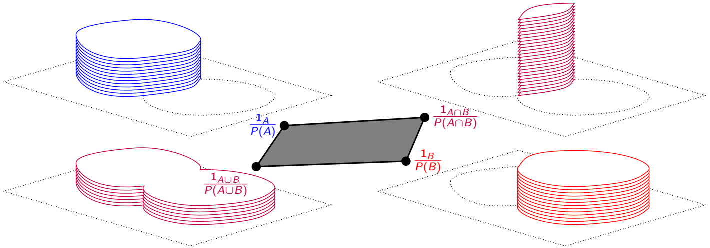
$$u = \sum_{i=1}^{k+1} t_i \mathbb{1}_{H_i}$$





In fact, almost every u in F takes **exactly** $k+1$ **nonzero values**.

Example



Generic 2-face of C_{BV}.

What about exposed faces?

- An exposed face of C_{BV} is a set of the form

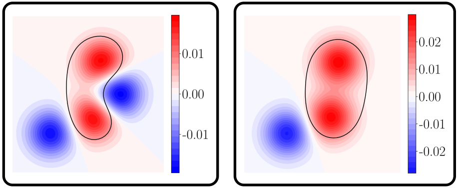
$$F = \mathrm{argmax}_{u \in C_{\mathrm{BV}}} \int_{\mathbb{R}^2} u\eta$$

given some function $\eta \in L^2(\mathbb{R}^2)$.

- The extreme points are of the form $u = \pm \mathbb{1}_E / P(E)$, where

$$E \in \left(\mathrm{argmax}_{E \in \mathbb{R}^2} \frac{\int_E \eta}{P(E)} \right) \quad \left(\text{resp.} - \frac{\int_E \eta}{P(E)} \right)$$

(generalized Cheeger problem)



What about exposed faces?

- An exposed face of C_{BV} is a set of the form

$$F = \mathrm{argmax}_{u \in C_{\mathrm{BV}}} \int_{\mathbb{R}^2} u\eta$$

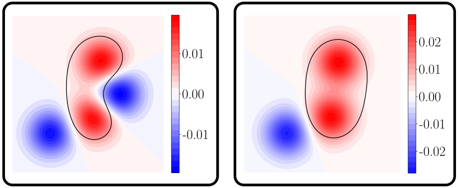
given some function $\eta \in L^2(\mathbb{R}^2)$.

- The extreme points are of the form $u = \pm \mathbb{1}_E / P(E)$, where

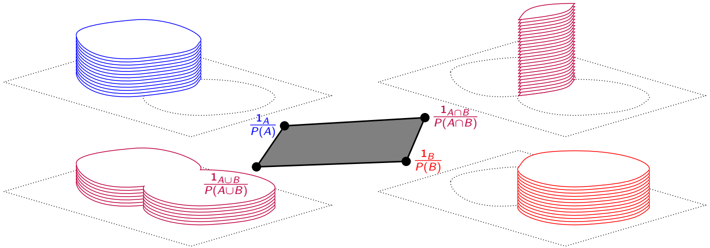
$$E \in \left(\mathrm{argmax}_{E \in \mathbb{R}^2} \frac{\int_E \eta}{P(E)} \right) \quad \left(\text{resp.} - \frac{\int_E \eta}{P(E)} \right)$$

(generalized Cheeger problem)

- If $\eta \in L^2(\mathbb{R}^2) \cap \mathscr{C}_b^1(\mathbb{R}^2)$, then E is a set of class $\mathscr{C}^3$.

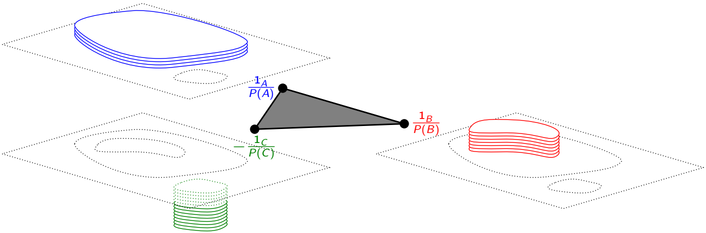


Faces exposed by $\mathscr{C}_b^1$ functions



Generic 2-face of C_{BV}.

Faces exposed by $\mathscr{C}_b^1$ functions



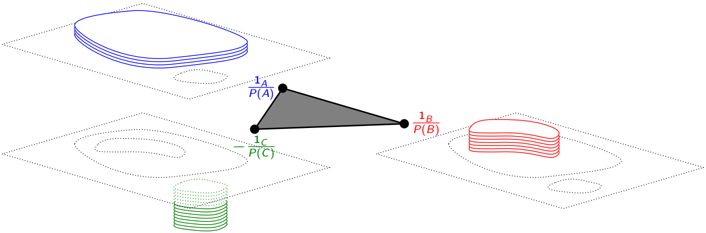
A 2-face exposed by some $\eta \in L^2(\mathbb{R}^2) \cap \mathscr{C}_b^1(\mathbb{R}^2)$

A k-face exposed by some $\eta \in L^2(\mathbb{R}^2) \cap \mathscr{C}_b^1(\mathbb{R}^2)$ is a simplex,

$$\forall u \in F, \quad u = \sum_{i=1}^{k+1} a_i \mathbb{1}_{E_i}.$$

and this decomposition is unique.

Faces exposed by $\mathscr{C}_b^1$ functions



A 2-face exposed by some $\eta \in L^2(\mathbb{R}^2) \cap \mathscr{C}_b^1(\mathbb{R}^2)$

Definition

A k-simple function is a function of the form

$$u = \sum_{i=1}^{k} a_i \mathbb{1}_{E_i}.$$

where the E_i's are $\mathscr{C}^1$ and $\partial E_i \cap \partial E_j = \emptyset$ for $i \neq j$.

Summary

1. Setting
2. What is the structure of the solutions?
3. Is that structure stable?
4. Algorithm and numerical results

$$\min_{u \in L^2(\mathbb{R}^2)} \int_{\mathbb{R}^2} |Du| + \frac{1}{2\lambda} \|\Phi u - y\|^2 \qquad (\mathcal{P}_\lambda(y))$$

$$\min_{u \in L^2(\mathbb{R}^2)} \int_{\mathbb{R}^2} |Du| \quad \text{s.c.} \quad \Phi u = y_0 \qquad (\mathcal{P}_0(y_0))$$

Proposition ([Hofmann et al., 2007])

If $\lambda^{(n)} \to 0$ and $\left\|w^{(n)}\right\|^2 / \lambda^{(n)} \to 0$, every sequence $u^{(n)}$ of solutions to $\mathcal{P}_{\lambda^{(n)}}(y^{(n)})$ has cluster points (in the weak L^2 topology), each of which is a solution to $\mathcal{P}_0(y_0)$.

Stability analysis - the setting

$$\min_{u \in L^2(\mathbb{R}^2)} \int_{\mathbb{R}^2} |Du| + \frac{1}{2\lambda} \|\Phi u - y\|^2 \qquad (\mathcal{P}_\lambda(y))$$

$$\min_{u \in L^2(\mathbb{R}^2)} \int_{\mathbb{R}^2} |Du| \quad \text{s.c.} \quad \Phi u = y_0 \qquad (\mathcal{P}_0(y_0))$$

We know that

- For each $\lambda > 0$, each $y = y_0 + w$, there is a solution to $\mathcal{P}_\lambda(y)$ of the form $u = \sum_{i=1}^{k} a_i \mathbb{1}_{E_i}$ and $k = k(\lambda, w) \leqslant M$.
- Same for $\lambda = 0$ and y_0.

Stability analysis - the setting

$$\min_{u \in L^2(\mathbb{R}^2)} \int_{\mathbb{R}^2} |Du| + \frac{1}{2\lambda} \|\Phi u - y\|^2 \qquad (\mathcal{P}_\lambda(y))$$

$$\min_{u \in L^2(\mathbb{R}^2)} \int_{\mathbb{R}^2} |Du| \quad \text{s.c.} \quad \Phi u = y_0 \qquad (\mathcal{P}_0(y_0))$$

We know that

- For each $\lambda > 0$, each $y = y_0 + w$, there is a solution to $\mathcal{P}_\lambda(y)$ of the form $u = \sum_{i=1}^k a_i \mathbb{1}_{E_i}$ and $k = k(\lambda, w) \leqslant M$.
- Same for $\lambda = 0$ and y_0.

Assumptions

- u_0 is the unique solution to $\mathcal{P}_0(y_0)$,
- $\varphi_i \in L^2(\mathbb{R}^2) \cap \mathscr{C}_b^1(\mathbb{R}^2)$ for $1 \leqslant i \leqslant M$.

What can we say about the convergence of the E_i's?

Stability analysis - the setting

u_0 $\qquad$ $u^{(n)}$ $\qquad$ $U_n^{(t)},\ t \in \mathbb{R}$

We know that

- For each $\lambda > 0$, each $y = y_0 + w$, there is a solution to $\mathcal{P}_\lambda(y)$ of the form $u = \sum_{i=1}^k a_i \mathbb{1}_{E_i}$ and $k = k(\lambda, w) \leqslant M$.
- Same for $\lambda = 0$ and y_0.

Proposition ([Chambolle et al., 2016, Iglesias et al., 2018])

If $\lambda_n \to 0$, $\frac{\|w_n\|}{\lambda_n} \leqslant \frac{\sqrt{\pi}}{2\|\Phi^*\|}$ *+ source cond. then (up to extr.)* $u_n \to u_0$ *strictly in* $\mathrm{BV}(\mathbb{R}^2)$

$$\text{and for a.e. } t \in \mathbb{R},\ \partial U_n^{(t)} \xrightarrow{\text{Hausdorff}} \partial U_0^{(t)} \quad \text{with} \quad U^{(t)} = \begin{cases} \{u \geqslant t\} \text{ if } t \geqslant 0 \\ \{u \leqslant t\} \text{ otherwise.} \end{cases}$$

- Let u be a solution to $\mathcal{P}_\lambda(y)$ and p be the solution to *the dual problem* $\mathcal{D}_\lambda(y)$
- Then the optimality condition yields

$$\Phi^* p \in \partial \mathrm{TV}(u).$$

In other words,

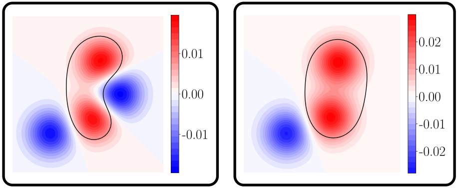
$$\frac{u}{\mathrm{TV}(u)} \in \left(\operatorname{argmax}_{v \in \mathcal{C}_{\mathrm{BV}}} \int_{\mathbb{R}^2} (\Phi^* p) u\right).$$

- Equivalently, the level sets of u must solve a *geometric variational problem* defined by p, known as the *prescribed curvature problem*.

(and similarly for $\mathcal{P}_0(y_0)$ and $\mathcal{D}_0(y_0)$).

The prescribed curvature problem

$$\min_{\substack{E \subseteq \mathbb{R}^2, \\ |E| < +\infty}} P(E) - \int_E \eta \qquad (\mathcal{Q}(\eta))$$

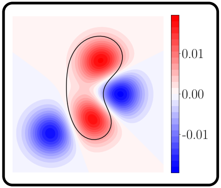


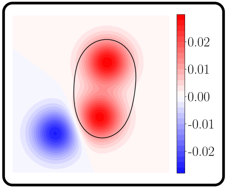
Let $\eta_{\lambda,y} \stackrel{\text{def.}}{=} \Phi^* p_{\lambda,y}$ where $p_{\lambda,y}$ is the unique solution to $\mathcal{D}_\lambda(y)$.

- $\forall t > 0,\ \{u_{\lambda,y} \geqslant t\}$ solves $\mathcal{Q}(\eta_{\lambda,y})$.
- $\forall t < 0,\ \{u_{\lambda,y} \leqslant t\}$ solves $\mathcal{Q}(-\eta_{\lambda,y})$.

The prescribed curvature problem

$$\min_{\substack{E\subseteq\mathbb{R}^2,\\ |E|<+\infty}} P(E) - \int_E \eta \qquad (\mathcal{Q}(\eta))$$





Let $\eta_{\lambda,y} \overset{\text{def.}}{=} \Phi^* p_{\lambda,y}$ where $p_{\lambda,y}$ is the unique solution to $\mathcal{D}_\lambda(y)$.

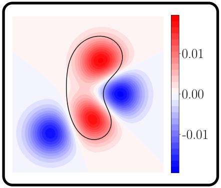
- $\forall t > 0,\ \{u_{\lambda,y} \geqslant t\}$ solves $\mathcal{Q}(\eta_{\lambda,y})$.
- $\forall t < 0,\ \{u_{\lambda,y} \leqslant t\}$ solves $\mathcal{Q}(-\eta_{\lambda,y})$.

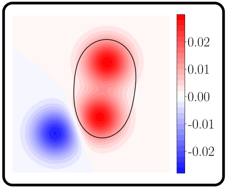
Let $\eta_{0,y_0} \overset{\text{def.}}{=} \Phi^* p_{0,y_0}$ where p_{0,y_0} is the solution to $\mathcal{D}_0(y_0)$ with minimal norm.

- $\forall t > 0,\ \{u_0 \geqslant t\}$ solves $\mathcal{Q}(\eta_{0,y_0})$.
- $\forall t < 0,\ \{u_0 \leqslant t\}$ solves $\mathcal{Q}(-\eta_{0,y_0})$.

The prescribed curvature problem

$$\min_{\substack{E\subseteq\mathbb{R}^2,\\ |E|<+\infty}} P(E) - \int_E \eta \qquad (\mathcal{Q}(\eta))$$





Let $\eta_{\lambda,y} \stackrel{\text{def.}}{=} \Phi^* p_{\lambda,y}$ where $p_{\lambda,y}$ is the unique solution to $\mathcal{D}_\lambda(y)$.

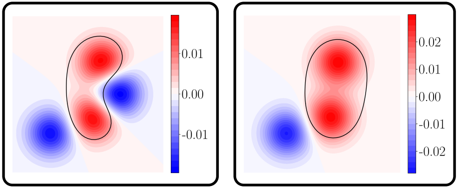
- $\forall t > 0,\ \{u_{\lambda,y} \geqslant t\}$ solves $\mathcal{Q}(\eta_{\lambda,y})$.
- $\forall t < 0,\ \{u_{\lambda,y} \leqslant t\}$ solves $\mathcal{Q}(-\eta_{\lambda,y})$.

Let $\eta_{0,y_0} \stackrel{\text{def.}}{=} \Phi^* p_{0,y_0}$ where p_{0,y_0} is the solution to $\mathcal{D}_0(y_0)$ with minimal norm.

- $\forall t > 0,\ \{u_0 \geqslant t\}$ solves $\mathcal{Q}(\eta_{0,y_0})$.
- $\forall t < 0,\ \{u_0 \leqslant t\}$ solves $\mathcal{Q}(-\eta_{0,y_0})$.

The prescribed curvature problem

$$\min_{\substack{E \subseteq \mathbb{R}^2, \\ |E| < +\infty}} P(E) - \int_E \eta \qquad (\mathcal{Q}(\eta))$$



Let $\eta_{\lambda,y} \stackrel{\text{def.}}{=} \Phi^* p_{\lambda,y}$ where $p_{\lambda,y}$ is the unique solution to $\mathcal{D}_\lambda(y)$.

- $\forall t > 0,\ \{u_{\lambda,y} \geqslant t\}$ solves $\mathcal{Q}(\eta_{\lambda,y})$.
- $\forall t < 0,\ \{u_{\lambda,y} \leqslant t\}$ solves $\mathcal{Q}(-\eta_{\lambda,y})$.

Let $\eta_{0,y_0} \stackrel{\text{def.}}{=} \Phi^* p_{0,y_0}$ where p_{0,y_0} is the solution to $\mathcal{D}_0(y_0)$ with minimal norm.

- $\forall t > 0,\ \{u_0 \geqslant t\}$ solves $\mathcal{Q}(\eta_{0,y_0})$.
- $\forall t < 0,\ \{u_0 \leqslant t\}$ solves $\mathcal{Q}(-\eta_{0,y_0})$.

Convergence of the curvature functionals

If $\lambda^{(n)} \to 0$ and $\left\| w^{(n)} \right\| / \lambda^{(n)} \to 0$, then $\eta_{\lambda^{(n)},y^{(n)}} \to \eta_{0,y_0}$ in $L^2(\mathbb{R}^2)$ and $\mathscr{C}_b^1(\mathbb{R}^2)$.

Adapting results from [Ambrosio 2010, Maggi 2012...], we have:

Proposition (Regularity of the boundary)

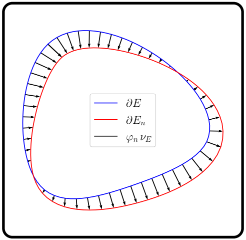
If $\eta \in \mathscr{C}^1_b(\mathbb{R}^2)$, any solution E to $\mathcal{Q}(\eta)$ is of class $\mathscr{C}^3$.

Proposition (Normal deformation)

Assume that $\eta_n \to \eta$ in $L^2(\mathbb{R}^2)$ and $\mathscr{C}^1_b(\mathbb{R}^2)$.
Then, for every $\varepsilon > 0$, there exists n_0 such that for every $n \geqslant n_0$, every solution of $\mathcal{Q}(\eta_n)$ satisfies

$$\partial E_n = (Id + \psi_n \nu_E)(\partial E)$$

for some solution E of $\mathcal{Q}(\eta)$ and $\|\psi_n\|_{C^2(\partial E)} \leqslant \varepsilon$.



Regularity results for the prescribed curvature problem

Adapting results from [Ambrosio 2010, Maggi 2012...], we have:

Proposition (Regularity of the boundary)

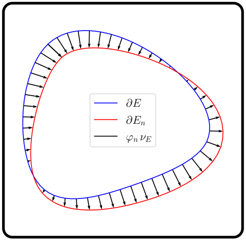
If $\eta \in \mathscr{C}^1_b(\mathbb{R}^2)$, any solution E to $\mathcal{Q}(\eta)$ is of class $\mathscr{C}^3$.

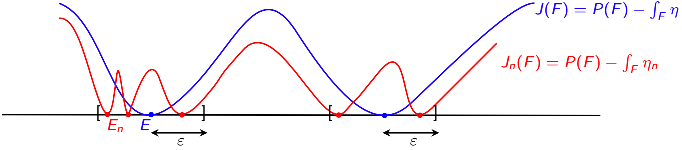
Proposition (Normal deformation)

Assume that $\eta_n \to \eta$ in $L^2(\mathbb{R}^2)$ and $\mathscr{C}^1_b(\mathbb{R}^2)$.
Then, for every $\varepsilon > 0$, there exists n_0 such that for every $n \geqslant n_0$, every solution of $\mathcal{Q}(\eta_n)$ satisfies

$$\partial E_n = (Id + \psi_n \nu_E)(\partial E)$$

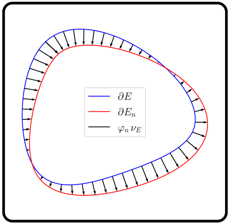
for some solution E of $\mathcal{Q}(\eta)$ and $\|\psi_n\|_{C^2(\partial E)} \leqslant \varepsilon$.





Second order shape derivatives

$$J(E) \overset{\text{def.}}{=} P(E) - \int_E \eta$$



Introduce the functional $\mathrm{j}_E : \mathrm{W}^{1,\infty}(\partial E) \to \mathbb{R}$ with

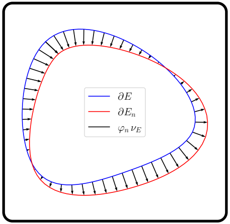
$$\psi \mapsto J(E_\psi)$$

$\partial E_\psi = (Id + \psi\, \nu_E)(\partial E)$, and look at its derivatives $j_E''(\psi)$, $j_E''(\psi)$ (see [Henrot & Pierre 2018])...

Definition

A minimizer E of J is *strictly stable* if $\forall \psi \in H^1(\partial E),\ j_E''(0)[\psi, \psi] > 0$.

Second order shape derivatives



$$J(E) \stackrel{\text{def.}}{=} P(E) - \int_E \eta$$

Introduce the functional $\mathrm{j}_E : \mathrm{W}^{1,\infty}(\partial E) \to \mathbb{R}$ with $\psi \mapsto J(E_\psi)$

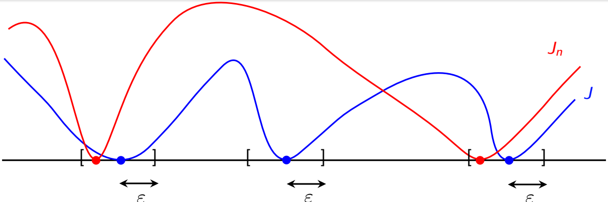
$\partial E_\psi = (Id + \psi\, \nu_E)(\partial E)$, and look at its derivatives $j''_E(\psi)$, $j''_E(\psi)$ (see [Henrot & Pierre 2018])...

Definition

A minimizer E of J is *strictly stable* if $\forall \psi \in H^1(\partial E)$, $j''_E(0)[\psi, \psi] > 0$.

Proposition

If $\eta_n \to \eta$ in $L^2(\mathbb{R}^2)$ and $\mathscr{C}^1_b(\mathbb{R}^2)$, and if E is a strictly stable minimizer of J, there exists $\varepsilon > 0$ such that for n large enough, there is at most one ψ_n s.t. $\|\psi_n\|_{C^2(\partial E)} \leqslant \varepsilon$ and E_{ψ_n} is a minimizer of J_n.



Theorem

Let $u_0 = \sum_{i=1}^k a_i \mathbb{1}_{E_i}$ be a k-simple function and assume that

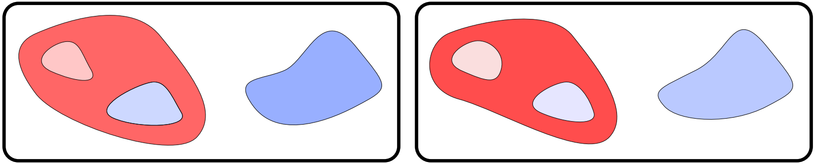
- (source condition) there exists $\eta \in \partial TV(u_0) \cap \operatorname{Im}\Phi^*$,
- each E_i is a strictly stable solution to $\mathcal{PC}(\operatorname{sign}(a_i)\eta_0)$,
- there is no other simple set solution to $\mathcal{PC}(\operatorname{sign}(a_i)\eta_0)$,
- $\{\Phi\mathbb{1}_{E_1}, \ldots, \Phi\mathbb{1}_{E_k}\}$ has full rank.

Then, if $\lambda \leqslant \lambda_0$ and $\|w\|/\lambda \leqslant \alpha$,

$$u = \sum_{i=1}^k a_i \mathbb{1}_{E_i},$$

and for $(\lambda, w) \to (0, 0)$

- $a_i \to a_{0,i}$,
- $\partial E_i = (Id + \psi\nu_E)(\partial E)$ and $\|\psi\|_{C^2(\partial E)} \to 0$.



Summary

Numerical resolution

Original

TV min with
standard scheme

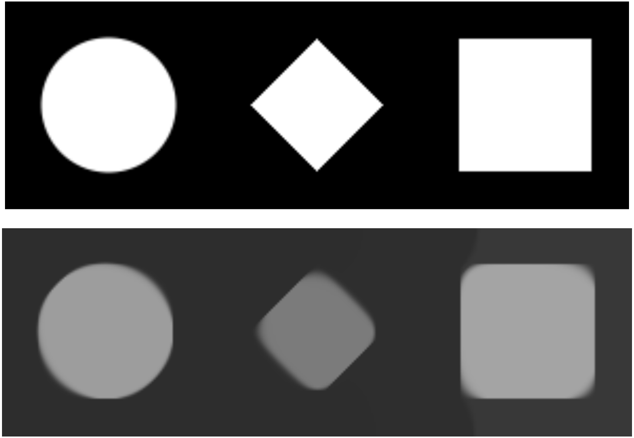


Illustration from [Tabti et al.'17]

Main issues with standard scheme:

- Blur,
- Anisotropy,
- Slow convergence with indicators of sets.

Many proposed schemes for total variation minimization on a grid have been proposed [Chambolle et al.'11, Abergel & Moisan'17, Tabti et al.'17, Condat'17, Chambolle & Pock'20..]

Goal

Exploit the structure of the solutions to design an "**off-the-grid**" algorithm, to produce

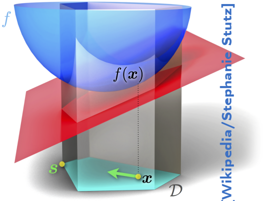
- Sharp edges
- Isotropic results

Goal: Minimize a convex differentiable function f on a compact convex set $\mathcal{D} \subset E$

Algorithm (Frank-Wolfe/Conditional gradient)

For all $k \in \mathbb{N}$, iterate

1. Linear minimization:
   $s_k \in \operatorname{argmin}_{s \in \mathcal{D}} f(x_k) + \mathrm{d}f(x_k)[s - x_k]$
2. Line search: $x_{k+1} \in \operatorname{argmin}_{x \in [x_k, s_k]} f(x)$



[Wikipedia/Stephanie Stutz]

Remarks:

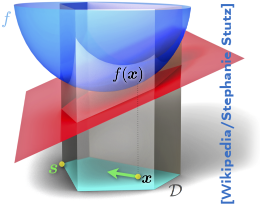
- If E is a Banach space and $\mathrm{d}f$ is Lipschitz, $f(x_k) - \min_{\mathcal{D}} f = O\left(\frac{1}{k}\right)$.
- Minimization of a linear form: OK if we can handle the **extreme points** of $\mathcal{D}$.

Goal: Minimize a convex differentiable function f on a compact convex set $\mathcal{D} \subset E$

Algorithm (Frank-Wolfe/Conditional gradient)

For all $k \in \mathbb{N}$, iterate

1. Linear minimization:
   $s_k \in \operatorname{argmin}_{s \in \mathcal{D}} f(x_k) + \mathrm{d}f(x_k)[s - x_k]$
2. Line search: $x_{k+1} \in \operatorname{argmin}_{x \in [x_k, s_k]} f(x)$



[Wikipedia/Stephanie Stutz]

Remarks:

- If E is a Banach space and $\mathrm{d}f$ is Lipschitz, $f(x_k) - \min_{\mathcal{D}} f = O\left(\frac{1}{k}\right)$.
- Minimization of a linear form: OK if we can handle the **extreme points** of $\mathcal{D}$.
- In step 2, one may choose $x_{k+1} \in \mathcal{D}$ with $f(x_{k+1}) \leqslant \min_{x \in [x_k, s_k]} f(x)$

Goal: $$\min_{u \in L^2(\mathbb{R}^2)} \Psi(u) \overset{\text{def.}}{=} \lambda \int |Du| + \frac{1}{2} \|\Phi u - y\|_{\mathcal{H}}^2$$

Differentiable? Constraint convex set $\mathcal{D}$?

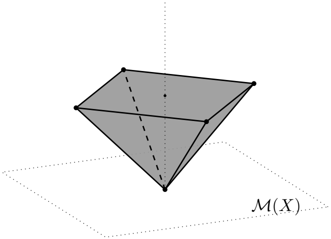
Frank-Wolfe for Total Variation recovery

Goal: $$\min_{u \in L^2(\mathbb{R}^2)} \Psi(u) \overset{\text{def.}}{=} \lambda \int |Du| + \frac{1}{2} \|\Phi u - y\|_{\mathcal{H}}^2$$

Differentiable? Constraint convex set $\mathcal{D}$?

Trick: work with the epigraph (inspired from [Harchaoui'15])

$$\min_{(t,u) \in \mathbb{R} \times L^2(\mathbb{R}^2)} \lambda t + \frac{1}{2} \|\Phi u - y\|_{\mathcal{H}}^2 \quad \text{s.t.} \quad \lambda \int |Du| \leqslant t \leqslant 1/2 \, \|y\|_{\mathcal{H}}^2$$



Algorithm (Sliding Frank-Wolfe for Total Variation Recovery)

For all $k \in \mathbb{N}$, iterate

1. Minimization:

$$\text{Set } \eta^{[k]} \stackrel{\text{def.}}{=} \frac{1}{\lambda}\Phi^*(y - \Phi u^{[k]}) \text{ and find } E_*^{[k]} \in \operatorname{argmax}_{E \subseteq \mathbb{R}^2} \pm \frac{\int_E \eta^{[k]}}{P(E)}$$

2. If $\left|\frac{\int_E \eta^{[k]}}{P(E)}\right| = 1$ then **stop**.

   Otherwise,

   - Update the support
     $\mathcal{S}^{k+1/2} = \left\{E_1^{[k]}, \ldots, E_{N_k}^{[k]}, E_*^k\right\} \stackrel{\text{def.}}{=} \left\{E_1^{[k+1/2]}, \ldots, E_{N_k+1}^{[k+1/2]}\right\}$
   - Find the amplitude (discrete LASSO):

$$a^{k+1/2} \in \operatorname*{argmin}_{a \in \mathbb{R}^{N_k+1}} \lambda \sum_i a_i P(E_i^{[k+1/2]}) + \frac{1}{2}\left\|\sum_i a_i \Phi \mathbb{1}_{E_i^{[k+1/2]}} - y\right\|_{\mathcal{H}}^2$$

   - Non-convex update of the positions and amplitudes (gradient descent)

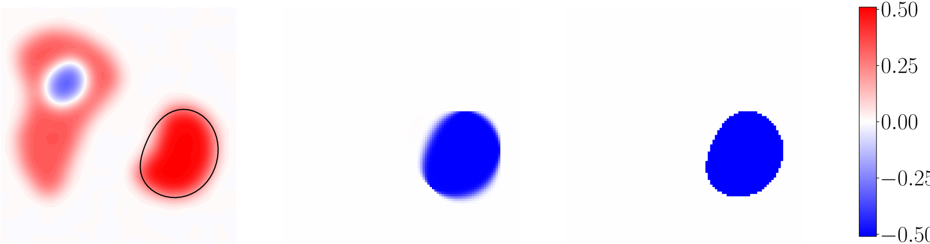
$$(\{a_i^{[k+1]}\}, \{E_i^{[k+1]}\}) \in \operatorname*{descent}_{(a,E)} \left(\sum_i a_i P(E_i^{[k+1/2]}) + \frac{1}{2}\left\|\sum_i a_i \Phi \mathbb{1}_{E_i^{[k+1/2]}} - y\right\|_{\mathcal{H}}^2\right)$$

Solving the Cheeger problem

Linear minimization step:

$$\text{argmax}_{E\subseteq\mathbb{R}^2} \pm \frac{\int_E \eta^{[k]}}{P(E)}$$

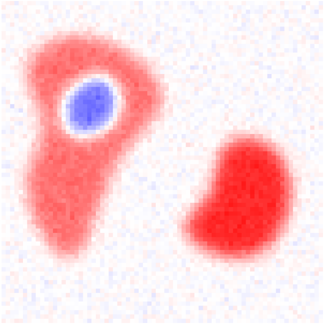
- There is a solution which is simply connected
- Resolution using polygonal curve evolution
- Initialization with the output some proximal algorithm on a rough grid ([Carlier et al.'09])



Weight η_k — Solution on a grid — Level set

Example

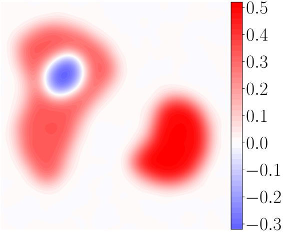
Unkown function u^*



Observation $y = \Phi u^* + n$

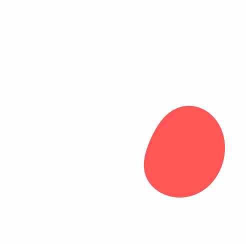
Example

Weight $\eta^{[1]} = \frac{1}{\lambda}\Phi^* y$



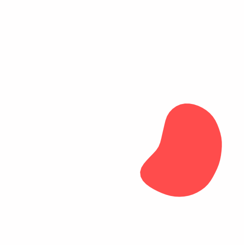
Example

Cheeger set $E_*^{[1]} = \text{argmin}_E \frac{\int_E \eta^{[1]}}{P(E)}$



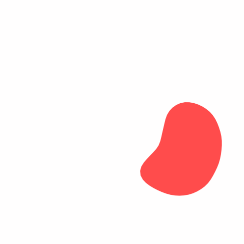
Example

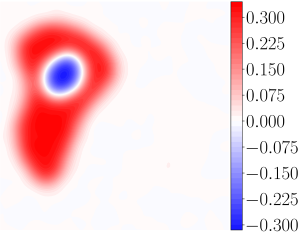
$u^{[1]} = \text{descent}\,(a, E)(\Psi(a\mathbb{1}_E))$



Example

$u^{[1]} = \text{descent}\,(a, E)(\Psi(a\mathbb{1}_E))$

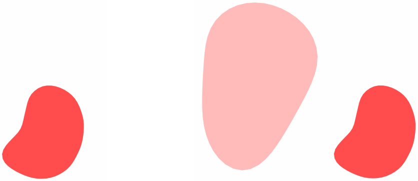




Weight $\eta^{[2]} = \frac{1}{\lambda}\Phi^*(y - \Phi u^{[1]})$

Example

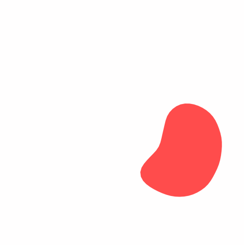
$u^{[1]} = \text{descent}\,(a, E)(\Psi(a\mathbb{1}_E))$

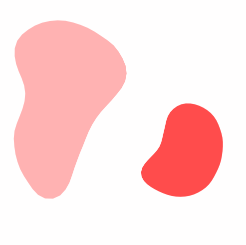


Cheeger set $E_*^{[2]} = \text{argmax}_E \pm \frac{\int_E \eta^{[2]}}{P(E)}$

Example

$u^{[1]} = \text{descent}\,(a, E)(\Psi(a\mathbb{1}_E))$



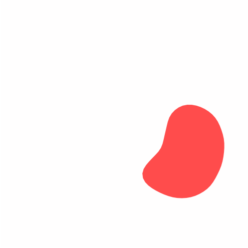


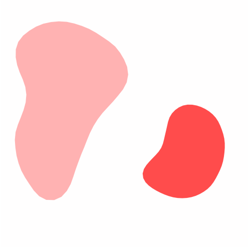
$u^{[2]} = \text{descent}\,(a, E)(\Psi(a_1\mathbb{1}_{E_1} + a_2\mathbb{1}_{E_2}))$

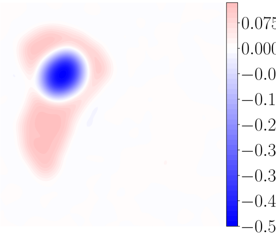
Example

$u^{[1]} = \text{descent}\,(a, E)(\Psi(a\mathbb{1}_E))$

Weight $\eta^{[3]} = \frac{1}{\lambda}\Phi^*(y - \Phi u^{[2]})$





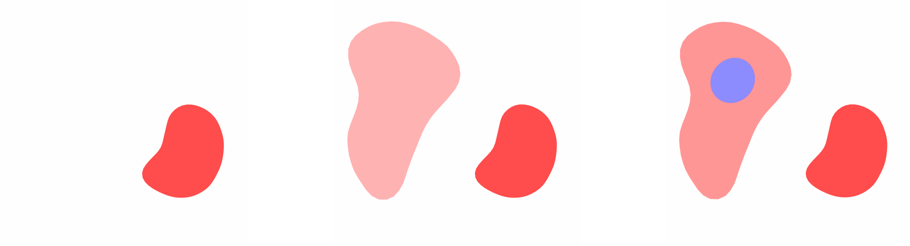


$u^{[2]} = \text{descent}\,(a, E)(\Psi(a_1\mathbb{1}_{E_1} + a_2\mathbb{1}_{E_2}))$

Example

$u^{[1]} = \text{descent}\,(a, E)(\Psi(a\mathbb{1}_E))$

Cheeger set $E_*^{[3]} = \text{argmax}_E \pm \frac{\int_E \eta^{[3]}}{P(E)}$

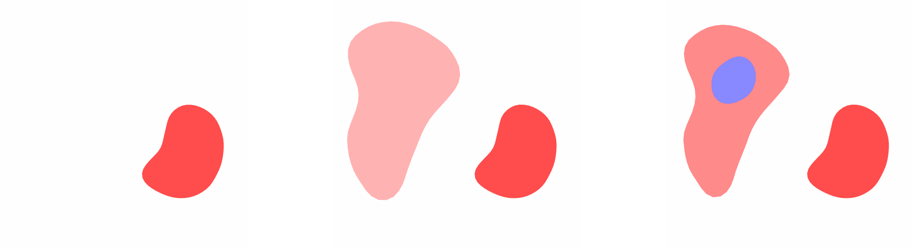


$u^{[2]} = \text{descent}\,(a, E)(\Psi(a_1\mathbb{1}_{E_1} + a_2\mathbb{1}_{E_2}))$

Example

$u^{[1]} = \text{descent}\,(a, E)(\Psi(a\mathbb{1}_E))$

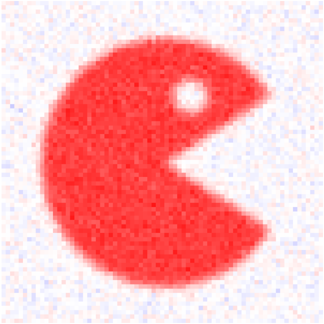
$u^{[3]} = \text{descent}\,(a, E)\left(\Psi(\sum_{i=1}^{3} a_i \mathbb{1}_{E_i 1})\right)$



$u^{[2]} = \text{descent}\,(a, E)(\Psi(a_1\mathbb{1}_{E_1} + a_2\mathbb{1}_{E_2}))$

Another example

Unkown function u^*



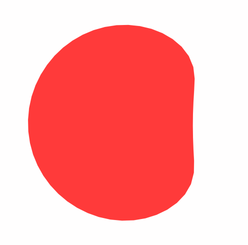
Observation $y = \Phi u^* + n$

Another example

Weight $\eta^{[1]} = \frac{1}{\lambda}\Phi^* y$

Another example

Cheeger set $E_*^{[1]} = \mathrm{argmax}_E \frac{\int_E \eta^{[1]}}{P(E)}$

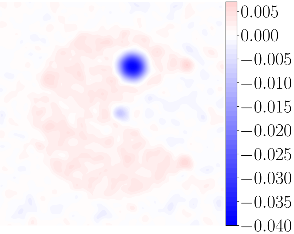


Another example

$$u^{[1]} = \mathrm{descent}\,(a, E)\,(\Psi(a\mathbb{1}_E))$$

Another example

$u^{[1]} = \text{descent}\,(a, E)\,(\Psi(a\mathbb{1}_E)))$



Weight $\eta^{[2]} = \frac{1}{\lambda}\Phi^*(y - \Phi u^{[1]})$

Another example

$u^{[1]} = \text{descent}\,(a, E)\,(\Psi(a\mathbb{1}_E)))$

Cheeger set $E_*^{[2]} = \text{argmax}_E - \frac{\int_E \eta^{[2]}}{P(E)}$

Another example

$u^{[1]} = \text{descent}\,(a, E)\,(\Psi(a\mathbb{1}_E)))$

$u^{[2]} = \text{descent}\,(a, E)\,(\Psi(a_1\mathbb{1}_{E_1} + a_2\mathbb{1}_{E_2})))$

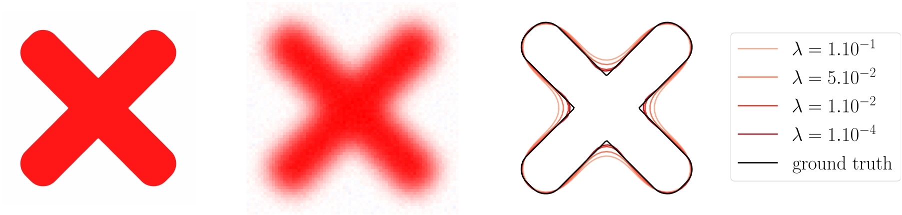
Jump set   Solution   Unkown u_*

Typical behavior of total variation regularization

- Loss of contrast,
- Rounding of the corners.

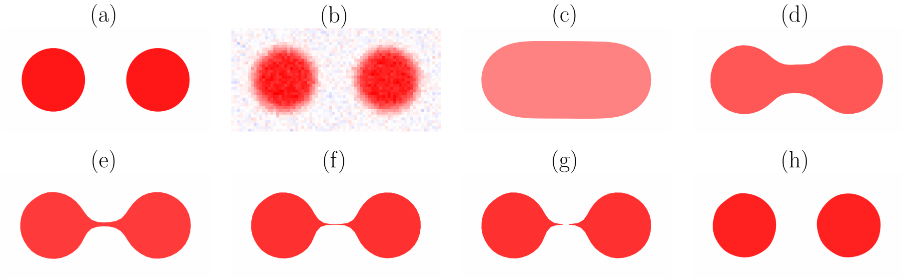
Curvature

$$\min_u \lambda \int |Du| + \frac{1}{2} \|\Phi u - y\|^2$$



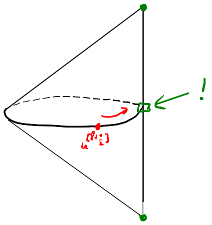
Topology changes

- For the Cheeger problem, there is a simply connected set
- In the non-convex refinement step, topology changes might occur.
- Handling the topology changes is not mandatory for global convergence, but it might yield better convergence / cleaner iterates.



Topology changes

- For the Cheeger problem, there is a simply connected set
- In the non-convex refinement step, topology changes might occur.
- Handling the topology changes is not mandatory for global convergence, but it might yield better convergence / cleaner iterates.



A more complex image

Unkown function u^*

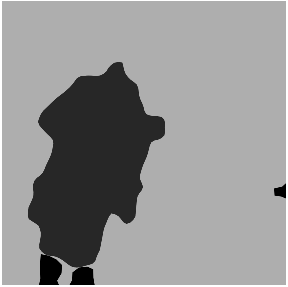
Observation $y = \Phi u^* + n$

Reconstruction $u^{[k]}$

A more complex image

Unkown function u^*

Observation $y = \Phi u^* + n$



Reconstruction $u^{[k]}$

A more complex image

Unkown function u^*

Observation $y = \Phi u^* + n$

Reconstruction $u^{[k]}$

A more complex image

Unkown function u^*

Observation $y = \Phi u^* + n$

Reconstruction $u^{[k]}$

A more complex image

Unkown function u^*

Observation $y = \Phi u^* + n$

Reconstruction $u^{[k]}$

A more complex image

Unkown function u^*

Observation $y = \Phi u^* + n$

Reconstruction $u^{[k]}$

Local convergence

- The global convergence is guaranteed by the convex framework
- In practice it can be better thanks to the shape optimization.

Can we bound the error? Prove the early convergence of the algorithm?

- Difficult to study
- Only partial results in the radial case, $N = 3$ or 4
- Ongoing work.

Conclusion

- A representation of the solutions is given by elementary arguments of convex analysis
- For some images, this representation is stable
- A gridless algorithm to take advantage of this structure

Thank you for your attention!

Faces and extreme points of convex sets for the resolution of inverse problems, V. Duval *Habilitation thesis (2022)*

Towards Off-the-grid Algorithms for Total Variation Regularized Inverse Problems, Y. De Castro, V. Duval, R. Petit *Journal of Mathematical Imaging and Vision (2022)*

Exact recovery of the support of piecewise constant images via total variation regularization, Y. De Castro, V. Duval, R. Petit *arXiv preprint: arxiv:2307.03709 (2023)*

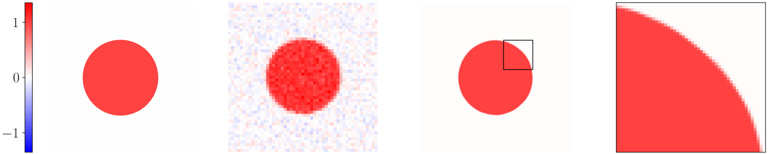
References I

Chambolle, A., Duval, V., Peyré, G., and Poon, C. (2016).
Geometric properties of solutions to the total variation denoising problem.
Inverse Problems, 33(1):015002.

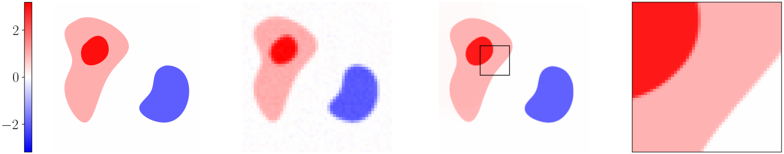
Hofmann, B., Kaltenbacher, B., Pöschl, C., and Scherzer, O. (2007).
A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators.
Inverse Problems, 23(3):987–1010.

Iglesias, J. A., Mercier, G., and Scherzer, O. (2018).
A note on convergence of solutions of total variation regularized linear inverse problems.
Inverse Problems, 34(5):055011.

Numerical examples



u_0 $y_0 + w$ approx. of $u_{\lambda,w}$



u_0 $y_0 + w$ approx. of $u_{\lambda,w}$