Variational curvatures and total variation regularization

José A. Iglesias

Johann Radon Institute for Computational and Applied Mathematics (RICAM)
Austrian Academy of Sciences (ÖAW)
Linz, Austria

CNA Seminar - CMU
April 8, 2020.

Funding:
(State of Upper Austria)

Introduction

We consider an ill-posed (e.g. A compact) linear operator equation

$$Au = f + w$$

where $u : \Omega \to \mathbb{R}$ for $\Omega \subset \mathbb{R}^d$, $d \geqslant 2$, and w noise (deterministic).

Assumption: the true data $u^\dagger$ to be recovered consists of distinct objects; smooth areas separated by sharp, discontinuous edges.

Introduction

We consider an ill-posed (e.g. A compact) linear operator equation

$$Au = f + w$$

where $u : \Omega \to \mathbb{R}$ for $\Omega \subset \mathbb{R}^d$, $d \geqslant 2$, and w noise (deterministic).

Assumption: the true data $u^\dagger$ to be recovered consists of distinct objects; smooth areas separated by sharp, discontinuous edges.

Total variation regularization: Assuming $A : L^q(\Omega) \to Y$ bounded with Y a Banach space and fixing $\sigma > 1$, the regularized solutions

$$u_{\alpha,w} \in \operatorname*{arg\,min}_{u \in L^q(\Omega)} \frac{1}{\sigma} \|Au - (f+w)\|_Y^\sigma + \alpha \mathrm{TV}(u)$$

balance a priori information about u encoded in having a low variation, and fidelity to the measurements available $f + w$.

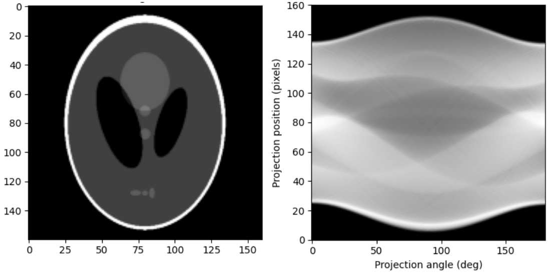
Spaces/exponents: Typically $q = d/(d-1)$, Y generic, will usually write $\sigma = 2$ (which does not always fit the assumptions) for simplicity.

TV is more than denoising

Radon transform in the plane: Integrate over lines determined by $(\theta, t) \in S^1 \times \mathbb{R}$,

$$\mathcal{R}u(\theta, t) = \int_{\{x \cdot \theta = t\}} u(x) \, \mathrm{d}\mathcal{H}^1(x).$$

$\|\mathcal{R}u\|_{H^{s+\frac{1}{2}}} \leqslant C\|u\|_{H^s}$ for compactly supported u, so inversion is ill-posed (noise is amplified). In our setting $s = 0$, $Y = L^4(\Omega)$ or $Y = L^2(\Omega)$.



Also higher dimensions: A reasonable starting model for computed tomography is the 3D analogue (X-Ray transform). Here $q \neq 2$.

Definitions

Definition: We call total variation the norm of the distributional gradient as a vector-valued Radon measure.

$$\mathrm{TV}(u) := \sup \left\{ \int_\Omega u \,\mathrm{div}\, z \, dx \;\middle|\; z \in C_0^\infty(\Omega\,;\mathbb{R}^d), \|z\|_{L^\infty(\Omega)} \leqslant 1 \right\}.$$

If finite for $u \in L^1(\Omega)$, then u has bounded variation: $u \in \mathrm{BV}(\Omega)$.

If u smooth, $\mathrm{TV}(u) = \int_\Omega |\nabla u| \, dx$.

Definitions

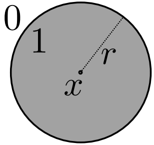
Definition: We call total variation the norm of the distributional gradient as a vector-valued Radon measure.

$$\mathrm{TV}(u) := \sup \left\{ \int_\Omega u \operatorname{div} z \, dx \;\middle|\; z \in C_0^\infty(\Omega; \mathbb{R}^d), \|z\|_{L^\infty(\Omega)} \leqslant 1 \right\}.$$

If finite for $u \in L^1(\Omega)$, then u has bounded variation: $u \in \mathrm{BV}(\Omega)$.

If u smooth, $\mathrm{TV}(u) = \int_\Omega |\nabla u| \, dx$.

For 1_E the indicatrix of a set E, define the perimeter $\mathrm{Per}(E) := \mathrm{TV}(1_E)$.



For a ball $B(x,r)$ we have $\mathrm{TV}(1_{B(x,r)}) = 2\pi r$ and the optimal z is aligned with the normal to the circle $\partial B(x,r)$.

We have in particular

$$\mathrm{Per}(A \cup B) + \mathrm{Per}(A \cap B) \leqslant \mathrm{Per}(A) + \mathrm{Per}(B)$$

generalizing 'length of the boundary' to irregular sets.

Convergence of solutions

First theoretical question: Analyze the low-noise regime in which the noise intensity and the a priori chosen parameter α tend to zero.

[1] R. Acar, C. R. Vogel. Inverse Prob. 10(6), 1994.

Convergence of solutions

First theoretical question: Analyze the low-noise regime in which the noise intensity and the a priori chosen parameter α tend to zero.

Basic convergence result[1]: If $f \in \mathrm{Ran}(A)$ and $\alpha_n, w_n \to 0$ with $\|w_n\|_Y^2/\alpha_n \leqslant C$, for

$$u_n := u_{\alpha_n, w_n} \in \operatorname*{arg\,min}_{u \in L^q(\Omega)} \frac{1}{2} \|Au - (f + w)\|_Y^2 + \alpha \mathrm{TV}(u)$$

we have (up to a subsequence) $u_n \rightharpoonup u^\dagger$ weakly in $L^{d/(d-1)}(\Omega)$ where $u^\dagger$ is a minimal TV solution of $Au = f$.

Moreover, also $u_n \to u^\dagger$ strong in $L^{\bar{q}}_{\mathrm{loc}}(\Omega)$ for $\bar{q} < d/(d-1)$, and in $L^{\bar{q}}(\Omega)$ if Ω is bounded.

[1] R. Acar, C. R. Vogel. Inverse Prob. 10(6), 1994.

Convergence of solutions

First theoretical question: Analyze the low-noise regime in which the noise intensity and the a priori chosen parameter α tend to zero.

Basic convergence result[1]**:** If $f \in \mathrm{Ran}(A)$ and $\alpha_n, w_n \to 0$ with $\|w_n\|_Y^2 / \alpha_n \leqslant C$, for

$$u_n := u_{\alpha_n, w_n} \in \operatorname*{arg\,min}_{u \in L^q(\Omega)} \frac{1}{2} \|Au - (f + w)\|_Y^2 + \alpha \mathrm{TV}(u)$$

we have (up to a subsequence) $u_n \rightharpoonup u^\dagger$ weakly in $L^{d/(d-1)}(\Omega)$ where $u^\dagger$ is a minimal TV solution of $Au = f$.

Moreover, also $u_n \to u^\dagger$ strong in $L^{\bar{q}}_{\mathrm{loc}}(\Omega)$ for $\bar{q} < d/(d-1)$, and in $L^{\bar{q}}(\Omega)$ if Ω is bounded.

Convergence rates: Linear rates in Bregman distance[2], under the source condition $A^* p_0 \in \partial \mathrm{TV}(u^\dagger)$ and linear parameter choice $\|w_n\|/\alpha_n \leqslant C$.

[1] R. Acar, C. R. Vogel. Inverse Prob. 10(6), 1994.
[2] M. Burger, S. Osher. Inverse Prob. 20(5), 2004.

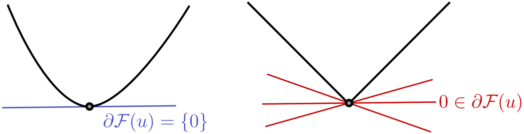
Subgradients and Bregman distances

If $\mathcal{F} : X \to \mathbb{R} \cup \{+\infty\}$, define the (set-valued) subgradient in X^* by

$$v \in \partial\mathcal{F}(u) \subset X^* \iff \mathcal{F}(\overline{u}) \geqslant \mathcal{F}(u) + \langle v, \overline{u} - u\rangle_{(X^*,X)}.$$

Fermat theorem for convex unconstrained minimization:

$$u \in \operatorname*{arg\,min}_{\overline{u} \in X} \mathcal{F}(\overline{u}) \iff 0 \in \partial\mathcal{F}(u).$$



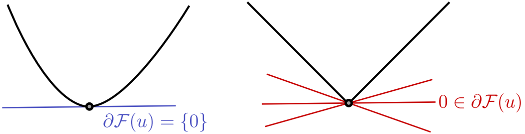
Subgradients and Bregman distances

If $\mathcal{F} : X \to \mathbb{R} \cup \{+\infty\}$, define the (set-valued) subgradient in X^* by

$$v \in \partial\mathcal{F}(u) \subset X^* \iff \mathcal{F}(\overline{u}) \geqslant \mathcal{F}(u) + \langle v, \overline{u} - u\rangle_{(X^*,X)}.$$

Fermat theorem for convex unconstrained minimization:

$$u \in \arg\min_{\overline{u}\in X} \mathcal{F}(\overline{u}) \iff 0 \in \partial\mathcal{F}(u).$$



Bregman distance: Denoting $v_0 := A^* p_0 \in \partial\mathrm{TV}(u^\dagger)$,

$$D_{v_0}(u_n, u^\dagger) := \mathrm{TV}(u_n) - \mathrm{TV}(u^\dagger) - \langle v_0, u_n - u^\dagger\rangle_{(L^2,L^2)}$$

involves $\mathrm{TV}(u_n)$, no translation to $\|u_n - u^\dagger\|_{L^2}$ without strong convexity.

However: Convergence rates in norm are possible[3] under an assumption of the type "$A^*A \geq \epsilon\,\mathrm{Id}$ on $\Sigma \supset \operatorname{supp} Du^\dagger$ with Lipschitz boundary".

[3] T. Valkonen. Regularisation, optimisation, subregularity. arXiv:2011.07575

Optimality condition I

Formal Euler-Lagrange equation: (with $Y = L^2$)

$$-\frac{A^*(Au_{\alpha,w} - f - w)}{\alpha} = \operatorname{div}\left(\frac{\nabla u_{\alpha,w}}{|\nabla u_{\alpha,w}|}\right).$$

With $u_{\alpha,w} \in C^2$ and at x with $\nabla u_{\alpha,w}(x) \neq 0$, the right hand side $\operatorname{div}(\nabla u_{\alpha,w}(x)/|\nabla u_{\alpha,w}(x)|)$ is the mean curvature of $\partial\{u_{\alpha,w} > u_{\alpha,w}(x)\}$.

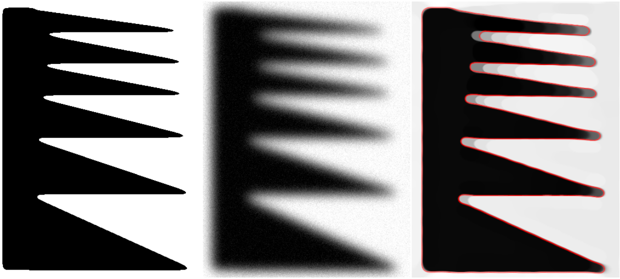
Optimality condition I

Formal Euler-Lagrange equation: (with $Y = L^2$)

$$-\frac{A^*(Au_{\alpha,w} - f - w)}{\alpha} = \operatorname{div}\left(\frac{\nabla u_{\alpha,w}}{|\nabla u_{\alpha,w}|}\right).$$

With $u_{\alpha,w} \in C^2$ and at x with $\nabla u_{\alpha,w}(x) \neq 0$, the right hand side $\operatorname{div}(\nabla u_{\alpha,w}(x)/|\nabla u_{\alpha,w}(x)|)$ is the mean curvature of $\partial\{u_{\alpha,w} > u_{\alpha,w}(x)\}$.

But these assumptions are very far from the typical use cases:



Ideal $u_0 \in \{0,1\}$, blurry and noisy $f = Au_0 + w$, solution $u_{\alpha,w}$, $\partial\{u_{\alpha,w} > 0.5\}$ in red.

Optimality condition II

Rigorous optimality condition: (with j duality mapping of Y)

$$v_{\alpha,w} := -\frac{1}{\alpha} A^* j(Au_{\alpha,w} - f - w) \in \partial \mathrm{TV}(u_{\alpha,w}), \text{ or}$$

$$\mathrm{TV}(u_{\alpha,w}) - \int_\Omega v_{\alpha,w} u_{\alpha,w} \leqslant \mathrm{TV}(\overline{u}) - \int_\Omega v_{\alpha,w} \overline{u}, \text{ for all } \overline{u} \in L^q(\Omega).$$

Optimality condition II

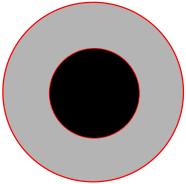
Rigorous optimality condition: (with j duality mapping of Y)

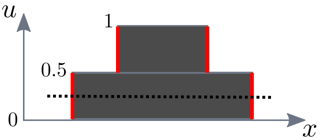
$$v_{\alpha,w} := -\frac{1}{\alpha} A^* j(Au_{\alpha,w} - f - w) \in \partial \mathrm{TV}(u_{\alpha,w}), \text{ or}$$

$$\mathrm{TV}(u_{\alpha,w}) - \int_\Omega v_{\alpha,w} u_{\alpha,w} \leqslant \mathrm{TV}(\overline{u}) - \int_\Omega v_{\alpha,w} \overline{u}, \text{ for all } \overline{u} \in L^q(\Omega).$$

Level sets? Slice with the coarea and layer-cake formulas

$$\mathrm{TV}(u) = \int_{-\infty}^{\infty} \mathrm{Per}(\{u > t\})\,\mathrm{d}t, \quad \int_\Omega \overline{u}\,\mathrm{d}x = \int_0^\infty |\{\overline{u} > t\}|\,\mathrm{d}t \text{ if } \overline{u} \geqslant 0$$





Variational curvatures of level sets

We can write optimality directly in terms of the level sets:

$$\{u_{\alpha,w} > t\} \in \arg\min_F \operatorname{Per}(F) - \int_F v_{\alpha,w} \quad (\text{for } t > 0)$$

which is the definition[4] of having **variational mean curvature** $v_{\alpha,w}$.

If everything was smooth, it would mean $v_{\alpha,w}\big|_{\partial\{u_{\alpha,w}>t\}}$ is the mean curvature of $\partial\{u_{\alpha,w} > t\}$ (as before).

[4] E. Barozzi, U. Massari, I. Tamanini, E. Gonzalez, ...

Variational curvatures of level sets

We can write optimality directly in terms of the level sets:

$$\{u_{\alpha,w} > t\} \in \arg\min_{F} \operatorname{Per}(F) - \int_F v_{\alpha,w} \quad (\text{for } t > 0)$$

which is the definition[4] of having **variational mean curvature** $v_{\alpha,w}$.

If everything was smooth, it would mean $v_{\alpha,w}\big|_{\partial\{u_{\alpha,w}>t\}}$ is the mean curvature of $\partial\{u_{\alpha,w} > t\}$ (as before).

Variational curvatures are not unique and nonlocal, since they are defined with Lebesgue/volume integrals!

[4] E. Barozzi, U. Massari, I. Tamanini, E. Gonzalez, ...

Variational curvatures of level sets

We can write optimality directly in terms of the level sets:

$$\{u_{\alpha,w} > t\} \in \arg\min_F \operatorname{Per}(F) - \int_F v_{\alpha,w} \quad (\text{for } t > 0)$$

which is the definition[4] of having **variational mean curvature** $v_{\alpha,w}$.

If everything was smooth, it would mean $v_{\alpha,w}\big|_{\partial\{u_{\alpha,w}>t\}}$ is the mean curvature of $\partial\{u_{\alpha,w} > t\}$ (as before).

Variational curvatures are not unique and nonlocal, since they are defined with Lebesgue/volume integrals!

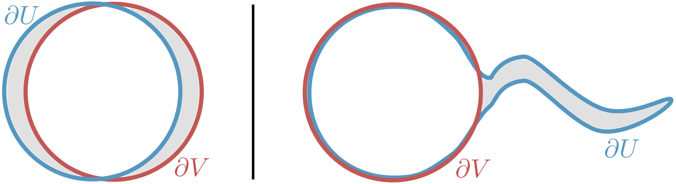
Integrability properties of $v_{\alpha,w}$ imply regularity of $\partial\{u_{\alpha,w}\}$. For the energy to not encourage small structures need to control $v_{\alpha,w}$ in L^d_{loc}, scaling:

$$\int_{B(x,r)} |v_{\alpha,w}| \leqslant r^{d-1} \|v_{\alpha,w}\|_{L^d(B(x,r))}, \text{ and } \operatorname{Per}\big(B(x,r)\big) = Cr^{d-1}.$$

[4] E. Barozzi, U. Massari, I. Tamanini, E. Gonzalez, ...

A better mode of convergence

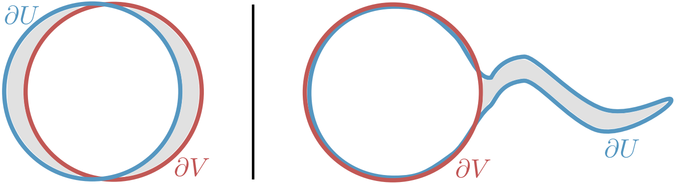
Convergence in L^p norm $\|u_n - u^\dagger\|_{L^p} \to 0$, not really adequate for measuring shapes of objects $U, V \subset \mathbb{R}^d$ with boundaries $\partial U, \partial V$:



Situations with comparable $|U \Delta V| = \|1_U - 1_V\|^p_{L^p}$.

A better mode of convergence

Convergence in L^p norm $\|u_n - u^\dagger\|_{L^p} \to 0$, not really adequate for measuring shapes of objects $U, V \subset \mathbb{R}^d$ with boundaries $\partial U, \partial V$:



Situations with comparable $|U \Delta V| = \|1_U - 1_V\|_{L^p}^p$, but very different $d_H(\partial U, \partial V)$.

Better mode: Convergence of boundaries in Hausdorff distance

$$d_H(\partial U, \partial V) := \max \left\{ \sup_{x \in \partial U} \inf_{y \in \partial V} |x - y|, \sup_{y \in \partial V} \inf_{x \in \partial U} |x - y| \right\}$$

can be seen as a geometric version of uniform convergence. Desirable to have it for the level set boundaries $\partial\{u_{\alpha,w} > t\}$ to $\partial\{u_0 > t\}$.

How to get Hausdorff convergence?

If the u_n have a common compact support, L^1_{loc} convergence $u_n \to u$ implies $|\{u_n > t\}\Delta\{u^\dagger > t\}| \to 0$ for a.e. t.

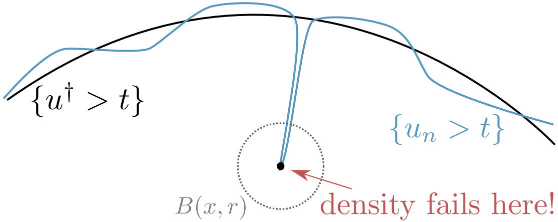
How to get Hausdorff convergence?

If the u_n have a common compact support, L^1_{loc} convergence $u_n \to u$ implies $|\{u_n > t\} \Delta \{u^\dagger > t\}| \to 0$ for a.e. t.

To improve to Hausdorff, it is sufficient to prove that there are $r_0 > 0$ and $C > 0$ such that for all $r \leqslant r_0$ and $x \in \partial\{u_n > t\}$

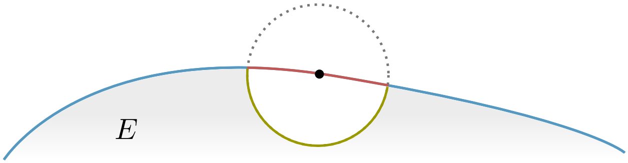
$$\frac{|\{u_n > t\} \cap B(x,r)|}{|B(x,r)|} \geqslant C \text{ and } \frac{|B(x,r) \setminus \{u_n > t\}|}{|B(x,r)|} \geqslant C,$$

density estimates, **uniform** in n, x and t.



Density estimates I

Let $E = \{u_{\alpha,w} > t\}$. To get[5] the estimate for $|E \cap B(x,r)|$, test optimality of E with $E \setminus B(x,r)$:



Using minimality and basic properties of the perimeter:

Lemma

$$\mathrm{Per}(E \cap B(x,r)) - \int_{E\cap B(x,r)} v_{\alpha,w} \leqslant 2\,\mathrm{Per}(B(x,r)\,;\, E^{(1)})$$
$$= 2\frac{d}{dr}|E \cap B(x,r)| \text{ for a.e. } r.$$

[5] E. Gonzalez, U. Massari, I. Tamanini. Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Nat. Rend. Lincei, 4(3), 1993.

Density estimates II

With Hölder inequality:

$$\begin{aligned}\left|\int_{E\cap B(x,r)} v_{\alpha,w}\right| &\leqslant |E\cap B(x,r)|^{\frac{d-1}{d}} \|v_{\alpha,w}\|_{L^d(E\cap B(x,r))}\\ &\leqslant |E\cap B(x,r)|^{\frac{d-1}{d}} \left(\|v_{\alpha,0}\|_{L^d(E\cap B(x,r))} + \|v_{\alpha,w}-v_{\alpha,0}\|_{L^d(\mathbb{R}^d)}\right).\end{aligned}$$

Density estimates II

With Hölder inequality:

$$\left|\int_{E\cap B(x,r)} v_{\alpha,w}\right| \leqslant |E\cap B(x,r)|^{\frac{d-1}{d}} \|v_{\alpha,w}\|_{L^d(E\cap B(x,r))}$$
$$\leqslant |E\cap B(x,r)|^{\frac{d-1}{d}} \left(\|v_{\alpha,0}\|_{L^d(E\cap B(x,r))} + \|v_{\alpha,w} - v_{\alpha,0}\|_{L^d(\mathbb{R}^d)}\right).$$

If we can (choosing r small enough and α large in terms of $\|w\|$) get

$$\|v_{\alpha,0}\|_{L^d(E\cap B(x,r))} < \varepsilon \text{ and } \|v_{\alpha,w} - v_{\alpha,0}\|_{L^d(\mathbb{R}^d)} < \eta,$$

the isoperimetric inequality $\mathrm{Per}(E\cap B(x,r)) \geqslant C_d|E\cap B(x,r)|^{\frac{d-1}{d}}$ and the lemma provide

$$|E\cap B(x,r)|^{\frac{d-1}{d}}(C_d - \varepsilon - \eta) \leqslant 2\frac{\mathrm{d}}{\mathrm{d}r}|E\cap B(x,r)| \text{ for a.e. } r,$$

a differential inequality giving $|E\cap B(x,r)| \geqslant Cr^d = C'|B(x,r)|$.

Strategy

This Hausdorff convergence was proved proved first[6] for $A = \mathrm{Id}$ (denoising) and $\Omega = \mathbb{R}^2$ under the conditions $\partial \mathrm{TV}(u^\dagger) \neq \emptyset$ and $\|w_n\|_{L^2(\Omega)}/\alpha_n \leqslant C$.

Our goal has been to extend this result to linear inverse problems, higher dimensions, bounded domains, and possibly without source condition.

[6] A. Chambolle, V. Duval, G. Peyré, and C. Poon. Inverse Prob. 33(1), 2017.

Strategy

This Hausdorff convergence was proved proved first[6] for $A = \mathrm{Id}$ (denoising) and $\Omega = \mathbb{R}^2$ under the conditions $\partial \mathrm{TV}(u^\dagger) \neq \emptyset$ and $\|w_n\|_{L^2(\Omega)}/\alpha_n \leqslant C$.

Our goal has been to extend this result to linear inverse problems, higher dimensions, bounded domains, and possibly without source condition.

General proof scheme:

- Fenchel dual problem lives in in $(L^q)^* = L^{q/(q-1)}$ and involves $v_{\alpha,w}$ that should belong to L^d.
- Strong L^d convergence of noiseless dual variables $v_{\alpha,0}$.
- Stability estimates for $v_{\alpha,w}$ with respect to w.
- Both of these together imply uniform density estimates of $\{u_{\alpha,w} > t\}$ to improve the mode of convergence.

[6] A. Chambolle, V. Duval, G. Peyré, and C. Poon. Inverse Prob. 33(1), 2017.

The dual problem

Using the characterization for one-homogeneous functionals

$$v \in \partial \mathrm{TV}(u) \iff v \in \partial \mathrm{TV}(0) \text{ and } \int_\Omega vu = \mathrm{TV}(u),$$

the Fenchel dual of our problem is to maximize among $p \in Y^*$ such that $A^* p \in \partial \mathrm{TV}(0)$, the quantity

$$\mathcal{D}_{\alpha,w}(p) := \langle p, f + w \rangle_{(Y^*, Y)} - \frac{\alpha}{2} \|p\|^2_{Y^*}.$$

The dual problem

Using the characterization for one-homogeneous functionals

$$v \in \partial \mathrm{TV}(u) \iff v \in \partial \mathrm{TV}(0) \text{ and } \int_\Omega vu = \mathrm{TV}(u),$$

the Fenchel dual of our problem is to maximize among $p \in Y^*$ such that $A^* p \in \partial \mathrm{TV}(0)$, the quantity

$$\mathcal{D}_{\alpha,w}(p) := \langle p, f + w \rangle_{(Y^*,Y)} - \frac{\alpha}{2} \|p\|_{Y^*}^2 .$$

A is continuous and the fidelity term as well, so strong duality holds:

$$\inf_{u \in L^q(\Omega)} \mathcal{F}_{\alpha,w}(u) = \sup_{A^* p \in \partial \mathrm{TV}(0)} \mathcal{D}_{\alpha,w}(p),$$

the maximizer $p_{\alpha,w}$ is unique and we have the (already familiar) optimality condition

$$v_{\alpha,w} := A^* p_{\alpha_w} \in \partial \mathrm{TV}(u_{\alpha,w}).$$

Stability of the dual wrt noise

$$p_{\alpha,w} = \operatorname*{arg\,min}_{\substack{p \in Y^* \\ A^* p \in \partial \mathrm{TV}(0)}} \left\|\alpha^{-1}(f+w)\right\|_Y^2 - \frac{2}{\alpha}\left(\langle p, f+w\rangle_{(Y^*,Y)} - \frac{\alpha}{2}\left\|p\right\|_{Y^*}^2\right).$$

Stability of the dual wrt noise

$$p_{\alpha,w} = \underset{\substack{p \in Y^* \\ A^* p \in \partial \mathrm{TV}(0)}}{\arg\min} \; \|\alpha^{-1}(f+w)\|_Y^2 - \frac{2}{\alpha}\left(\langle p, f+w\rangle_{(Y^*,Y)} - \frac{\alpha}{2}\|p\|_{Y^*}^2\right).$$

This parallels the formula for projections onto a convex set in a Hilbert space (which are nonexpansive, so they give a linear stability estimate)

$$\inf_{\substack{p \in H \\ A^* p \in \partial \mathrm{TV}(0)}} \left\|\frac{f+w}{\alpha}\right\|_H^2 - 2\left\langle p, \, \frac{f+w}{\alpha}\right\rangle_H + \|p\|_H^2 = \inf_{\substack{p \in H \\ A^* p \in \partial \mathrm{TV}(0)}} \left\|p - \frac{f+w}{\alpha}\right\|_H^2.$$

Stability of the dual wrt noise

$$p_{\alpha,w} = \operatorname*{arg\,min}_{\substack{p\in Y^* \\ A^*p\in\partial\mathrm{TV}(0)}} \|\alpha^{-1}(f+w)\|_Y^2 - \frac{2}{\alpha}\left(\langle p, f+w\rangle_{(Y^*,Y)} - \frac{\alpha}{2}\|p\|_{Y^*}^2\right).$$

This parallels the formula for projections onto a convex set in a Hilbert space (which are nonexpansive, so they give a linear stability estimate)

$$\inf_{\substack{p\in H \\ A^*p\in\partial\mathrm{TV}(0)}} \left\|\frac{f+w}{\alpha}\right\|_H^2 - 2\left\langle p, \frac{f+w}{\alpha}\right\rangle_H + \|p\|_H^2 = \inf_{\substack{p\in H \\ A^*p\in\partial\mathrm{TV}(0)}} \left\|p - \frac{f+w}{\alpha}\right\|_H^2.$$

Our dual can be interpreted as a generalized (not metric!) projection[7] which is also stable assuming uniform convexity of Y^*:

$$\|p_{\alpha,w} - p_{\alpha,0}\|_{Y^*} \leqslant \rho_{Y,2}\left(\frac{\|w\|_Y}{2\alpha}\right),$$

applying A^* we can pick α with $\|v_{\alpha,w} - v_{\alpha,0}\|_{L^{q'}(\Omega)} \leqslant \eta$ for any $\eta > 0$.

[7] Alber, Notik. Comm. Appl. Nonlinear Anal., 1995.

Strong convergence of noiseless duals

The formal limits of the primal and dual problems when $\alpha, w \to 0$ write

$$\inf\{\mathrm{TV}(u) \ \mid \ u \in L^q(\Omega),\, Au = f\} \text{ attained by } u^\dagger, \text{ and}$$

$$\sup_{A^*p \in \partial \mathrm{TV}(0)} \langle p\,,\, f\rangle_{(Y^*,Y)} = \sup_{A^*p \in \partial \mathrm{TV}(0)} \int_\Omega (A^*p) u^\dagger.$$

Strong convergence of noiseless duals

The formal limits of the primal and dual problems when $\alpha, w \to 0$ write

$$\inf\{\mathrm{TV}(u) \;\mid\; u \in L^q(\Omega),\, Au = f\} \text{ attained by } u^\dagger, \text{ and}$$

$$\sup_{A^*p \in \partial \mathrm{TV}(0)} \langle p\,,\, f\rangle_{(Y^*,Y)} = \sup_{A^*p \in \partial \mathrm{TV}(0)} \int_\Omega (A^*p)u^\dagger.$$

The fidelity term became a constraint (discontinuous), so no strong duality. Attainment in the dual for $u^\dagger, p_0$ with **source condition**

$$A^*p_0 \in \partial \mathrm{TV}(u^\dagger).$$

This also gives that the **noiseless** $p_{\alpha_n,0}$ converge strongly in Y^*, to the choice p^* with minimal Y^* norm
(homogeneity of TV and Radon-Riesz property of Y^*).

Strong convergence of noiseless duals

The formal limits of the primal and dual problems when $\alpha, w \to 0$ write

$$\inf\{\mathrm{TV}(u) \mid u \in L^q(\Omega),\, Au = f\} \text{ attained by } u^\dagger, \text{ and}$$

$$\sup_{A^* p \in \partial \mathrm{TV}(0)} \langle p\,, f\rangle_{(Y^*, Y)} = \sup_{A^* p \in \partial \mathrm{TV}(0)} \int_\Omega (A^* p) u^\dagger.$$

The fidelity term became a constraint (discontinuous), so no strong duality. Attainment in the dual for $u^\dagger, p_0$ with **source condition**

$$A^* p_0 \in \partial \mathrm{TV}(u^\dagger).$$

This also gives that the **noiseless** $p_{\alpha_n,0}$ converge strongly in Y^*, to the choice p^* with minimal Y^* norm
(homogeneity of TV and Radon-Riesz property of Y^*).

In consequence the sequence $(v_{\alpha_n,0}) = (A^* p_{\alpha_n,0})$ is L^d-equiintegrable: for each $\varepsilon > 0$ there is r_0 for which

$$\int_{B(x,r_0)} |v_{\alpha_n,0}|^d \leqslant \varepsilon^d \text{ for all } x \text{ and } n.$$

Geometric convergence under source condition

We assumed that the $\{u_{\alpha,w}\}$ have a common compact support, but for $q = d/(d-1)$ this can be proved again with stability and equiintegrability at infinity

$$\int_{\Omega \setminus B(0,R)} |v_{\alpha,0}|^d \leqslant \varepsilon^d \text{ for } R > R_0.$$

Geometric convergence under source condition

We assumed that the $\{u_{\alpha,w}\}$ have a common compact support, but for $q = d/(d-1)$ this can be proved again with stability and equiintegrability at infinity

$$\int_{\Omega \setminus B(0,R)} |v_{\alpha,0}|^d \leqslant \varepsilon^d \text{ for } R > R_0.$$

Theorem (I., Mercier, Scherzer '18 (d=2), I., Mercier '20 (d>2))

Let $q = d/(d-1)$, $\sigma \leqslant d/(d-1)$, Ω either $\mathbb{R}^d$ or a bounded domain with Neumann or Dirichlet (geometric restrictions!) boundary, $\alpha_n, w_n \xrightarrow[n\to\infty]{} 0$ related by

$$\alpha_n^{\frac{1}{\sigma-1}} \geqslant C(d, \Omega, Y, \sigma)\|w_n\|_Y,$$

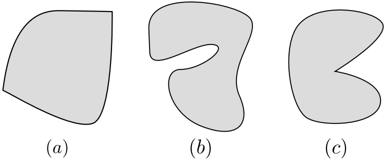
and assume the source condition $\mathrm{Ran}(\mathrm{A}^*) \cap \partial\mathrm{TV}(u^\dagger) \neq \emptyset$.

Then for u_n a subsequence of minimizers such that $u_n \to u^\dagger$ in L^1_{loc} and almost every $t > 0$, the level set boundaries $\partial\{u_n > t\}$ converge to $\partial\{u^\dagger > t\}$ in the sense of Hausdorff convergence.

Density estimates with boundary conditions

Dirichlet: Prescribe zero values outside Ω which has a variational curvature κ_Ω with $\kappa_\Omega\big|_{\mathbb{R}^d\setminus\Omega} \in L^d(\mathbb{R}^d \setminus \Omega)$,

$$\kappa_{\alpha,w} = v_{\alpha,w}1_\Omega + \kappa_\Omega 1_{\mathbb{R}^d\setminus\Omega} \text{ also a curvature for } \{u_{\alpha,w} > t\}.$$

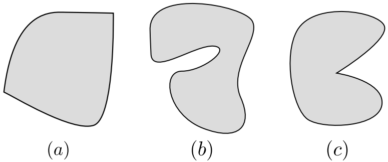


(a) and (b) are allowed, (c) is not.

Density estimates with boundary conditions

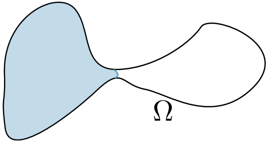
Dirichlet: Prescribe zero values outside Ω which has a variational curvature κ_Ω with $\kappa_\Omega\big|_{\mathbb{R}^d \setminus \Omega} \in L^d(\mathbb{R}^d \setminus \Omega)$,

$$\kappa_{\alpha,w} = v_{\alpha,w} 1_\Omega + \kappa_\Omega 1_{\mathbb{R}^d \setminus \Omega} \text{ also a curvature for } \{u_{\alpha,w} > t\}.$$

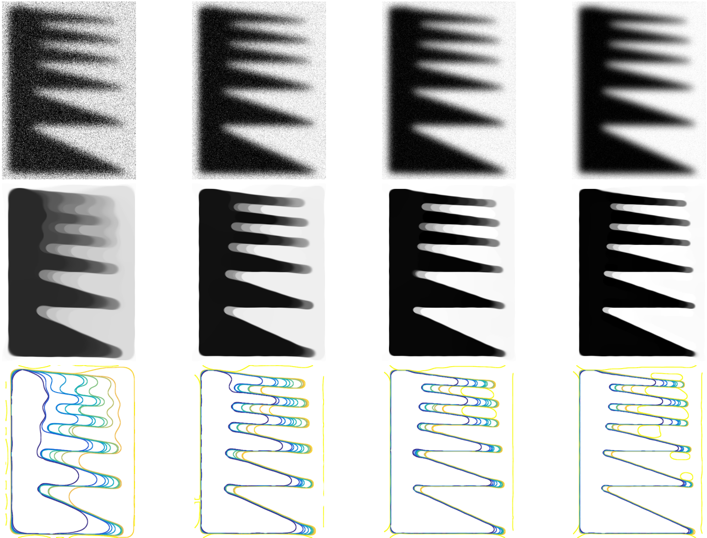


(a) and (b) are allowed, (c) is not.

Neumann: Variation in Ω, isoperimetric inequality fails but we have the BV Poincaré-Sobolev inequality (η depends on embedding constant).



A numerical example



Deblurring with Dirichlet boundary, $\alpha = 1, 1/4, 1/16, 1/64$ and $\|w\|_{L^2} = \alpha/10$.

Boundedness under source condition

Theorem (Bredies, I., Mercier, soon)

Under the same conditions (minus the geometric restriction on Ω), and for the full sequence,

$$\|u_n\|_{L^\infty(\Omega)} \leqslant C\left(u^\dagger, \sup_n \alpha_n^{-\frac{1}{\sigma-1}} \|w_n\|_Y\right) < +\infty,$$

and in consequence, if Ω is bounded and $u_n \to u^\dagger$ in L^1, also

$$u_n \to u^\dagger \text{ strongly in } L^p \text{ for all } p \in (1, \infty).$$

Boundedness under source condition

Theorem (Bredies, I., Mercier, soon)

Under the same conditions (minus the geometric restriction on Ω), and for the full sequence,

$$\|u_n\|_{L^\infty(\Omega)} \leqslant C\Big(u^\dagger, \sup_n \alpha_n^{-\frac{1}{\sigma-1}} \|w_n\|_Y\Big) < +\infty,$$

and in consequence, if Ω is bounded and $u_n \to u^\dagger$ in L^1, also

$$u_n \to u^\dagger \text{ strongly in } L^p \text{ for all } p \in (1,\infty).$$

Remarks:

The measurements $f + w$ could be unbounded. And even if they were bounded, there is A to deal with.

Analogously if $\partial \mathrm{TV}(u^\dagger) \neq \emptyset$, then $u^\dagger$ itself is bounded.

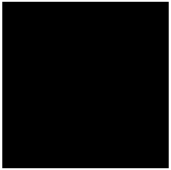
The bound is not explicit, since we use equiintegrability of $v_{\alpha,0}$.

Source condition and L^d curvatures

The source condition is not harmless: for the indicatrix of $S = (0,1)^2 \subset \mathbb{R}^2$, we have no curvature in L^2 and $\partial_{L^2}\mathrm{TV}(1_S) = \emptyset$.

This is because 'cutting' a corner K_δ of side length δ reduces perimeter by a fixed factor, which would contradict $v_S \in L^2$:

$$C\delta \leqslant \mathrm{Per}(S) - \mathrm{Per}(S \setminus K_\delta) \leqslant \int_{K_\delta} v_S \leqslant \delta \|v_S\|^2_{L^2(K_\delta)},$$



Original ($w = 0$)

'Denoised' ($A = \mathrm{Id}$), large α

But the square satisfies density estimates. And without noise, one can find a solution explicitly and the Hausdorff convergence still holds.

Spaces for curvatures and regularity

So far: To prove density estimates one uses (local) bounds in L^d, which in turn prevent mild singularities which we want to still handle.

Strong regularity: For $p > d$ and $d < 8$, a set E having a variational curvature in L^p implies[8] that $\partial E \in C^{1,\alpha}$.

Regularity with L^d curvature? bi-Hölder local parametrizations of the boundary[9], bi-Lipschitz[10] for $d = 2$. There are examples[11] of planar sets with singular points but with variational mean curvatures in L^2.

What is the optimal space? A more general (critical) setting for variational curvatures is the Morrey space $L^{1,d-1}$, but there are counterexamples[12] with no density estimates.

---

[8] U. Massari. Rend. Sem. Mat. Univ. Padova, 53, 1975.

[9] E. Paolini. Manuscripta Math., 97(1), 1998.

[10] L. Ambrosio, E. Paolini. Ricerche Mat. 48(suppl.), 1999.

[11] E. Gonzalez, U. Massari, I. Tamanini. Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Nat. Rend. Lincei, 4(3), 1993.

[12] Q. Li, M. Torres. Adv. Calc. Var. 12(2), 2019.

Denoising by comparison I

For denoising we can work on $v_{\alpha,w} = \alpha^{-1}(f + w - u_{\alpha,w})$ pointwise.

Lemma

Assume that the finite perimeter sets E_1 and E_2 admit variational mean curvatures κ_1 and κ_2 respectively, and such that $\kappa_1 < \kappa_2$ in $E_1 \setminus E_2$. Then $|E_1 \setminus E_2| = 0$, that is $E_1 \subseteq E_2$ up to Lebesgue measure zero.

Write

$$\mathrm{Per}(E_1) - \int_{E_1} \kappa_1 \leqslant \mathrm{Per}(E_1 \cap E_2) - \int_{E_1 \cap E_2} \kappa_1,$$

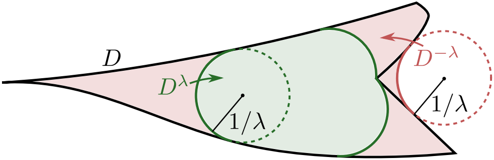
$$\mathrm{Per}(E_2) - \int_{E_2} \kappa_2 \leqslant \mathrm{Per}(E_1 \cup E_2) - \int_{E_1 \cup E_2} \kappa_2.$$

Summing and using $\mathrm{Per}(E_1) + \mathrm{Per}(E_2) \leqslant \mathrm{Per}(E_1 \cap E_2) + \mathrm{Per}(E_1 \cup E_2)$,

$$\int_{E_1 \setminus E_2} \kappa_2 \leqslant \int_{E_1 \setminus E_2} \kappa_1.$$

Denoising by comparison II

Let D be a bounded set of finite perimeter. Then D has[13] at least one variational mean curvature in $\kappa_D \in L^1(\mathbb{R}^d)$, minimal in $L^p(D)$ norm for all $p > 1$ among such curvatures.



With $g \in L^1(\mathbb{R}^d)$ with $g > 0$ consider

$$D^\lambda \in \operatorname*{arg\,min}_{E \subset D} \operatorname{Per}(E) - \lambda \int_E g, \text{ and } \mathbb{R}^d \setminus D^{-\lambda} \in \operatorname*{arg\,min}_{F \subset \mathbb{R}^d \setminus D} \operatorname{Per}(F) - \lambda \int_F g.$$

For $\lambda \to \infty$ the D^λ exhaust D, so for $x \in D$ we can define

$$\kappa_D(x) := \inf \left\{ \lambda g(x) \;\middle|\; \lambda > 0 \text{ and } x \in E^\lambda \right\}.$$

[13] E. Barozzi, E. Gonzalez, and I. Tamanini. Proc. Amer. Math. Soc., 99(2), 1987.

Denoising by comparison III

Theorem (I., Mercier '21)

Let $q = \sigma = d/(d-1)$, $\Omega = \mathbb{R}^d$, $A = \mathrm{Id}$, α_n, w_n *related by* $\alpha_n^{\frac{1}{\sigma-1}} \geqslant C(d)\|w_n\|$, *and* $f = 1_D$ *for* D *bounded of finite perimeter.*

Then for u_n *a subsequence of minimizers and almost every* $t > 0$, *the level set boundaries* $\partial\{u_n > t\}$ *converge to* ∂D *in the sense of Hausdorff convergence.*

Compare with $\kappa_D \in L^1(\mathbb{R}^d)$, which acts as replacement of $p_0 \in \partial \mathrm{TV}(f)$. Denoising reduces curvature pointwise:

$$|v_{\alpha,0}| \leqslant |\kappa_D| \text{ and } \mathrm{sign}(v_{\alpha,0}) = \mathrm{sign}(\kappa_D).$$

Since D is arbitrary, the density estimates are not uniform anymore. But using the additional bound obtained by comparing with balls

$$|\kappa_D| \leqslant \mathrm{dist}(\cdot, \partial D)$$

the Hausdorff convergence still holds.

Summary

- For TV regularization the dual variable is the variational curvature of the level sets of the solution.
- In dimension $d > 2$, one needs to go outside Hilbert spaces: the natural space for the curvatures is L^d.

Summary

- For TV regularization the dual variable is the variational curvature of the level sets of the solution.
- In dimension $d > 2$, one needs to go outside Hilbert spaces: the natural space for the curvatures is L^d.
- Uniform convexity of $Y, Y^* \implies$ Dual stability wrt noise (S).
  Source condition $\implies$ Noiseless dual convergence (C).
  (S)+(C)+param. choice $\implies$ Hausdorff convergence, L^∞ bounds.

Summary

- For TV regularization the dual variable is the variational curvature of the level sets of the solution.
- In dimension $d > 2$, one needs to go outside Hilbert spaces: the natural space for the curvatures is L^d.
- Uniform convexity of $Y, Y^* \implies$ Dual stability wrt noise (S).
  Source condition $\implies$ Noiseless dual convergence (C).
  (S)+(C)+param. choice $\implies$ Hausdorff convergence, L^∞ bounds.
- For denoising and $f = 1_D$, convergence by comparison with $\kappa_D \in L^1$.

Summary

- For TV regularization the dual variable is the variational curvature of the level sets of the solution.
- In dimension $d > 2$, one needs to go outside Hilbert spaces: the natural space for the curvatures is L^d.
- Uniform convexity of $Y, Y^* \implies$ Dual stability wrt noise (S).
  Source condition $\implies$ Noiseless dual convergence (C).
  (S)+(C)+param. choice $\implies$ Hausdorff convergence, L^∞ bounds.
- For denoising and $f = 1_D$, convergence by comparison with $\kappa_D \in L^1$.

- J. A. Iglesias, G. Mercier, and O. Scherzer. A note on convergence of solutions of total variation regularized linear inverse problems. Inverse Problems, 34(5):055011, 2018. arXiv:1711.06495
- J. A. Iglesias, G. Mercier. Influence of dimension on the convergence of level-sets in total variation regularization. ESAIM. Control, Optimisation and Calculus of Variations, 26:52, 2020. arXiv:1811.12243
- J. A. Iglesias, G. Mercier. Convergence of level sets in total variation denoising through variational curvatures in unbounded domains. SIAM Journal on Mathematical Analysis, 53(2):1509-1545, 2021. arXiv:2005.13910