Sparse recovery algorithms from exact and incomplete data

Massimo Fornasier

Program in Applied and Computational Mathematics
Princeton University
and
Johann Radon Institute for Computational and Applied Mathematics
Austrian Academy of Sciences

Alba di Canazei, September 6, 2007

Outline

1. Back to Compressed Sensing and beyond ℓ_1-minimization
   - Variational limits
   - Beyond ℓ_1-minimization
2. Free-discontinuity problems
   - The Mumford-Shah functional
   - Approximation by discrete functionals
   - A general discrete functional
   - Iterative thresholding algorithms for Mumford-Shah minimization
3. Adaptive frame methods for PDEs

Outline

1. Back to Compressed Sensing and beyond ℓ_1-minimization
   - Variational limits
   - Beyond ℓ_1-minimization
2. Free-discontinuity problems
   - The Mumford-Shah functional
   - Approximation by discrete functionals
   - A general discrete functional
   - Iterative thresholding algorithms for Mumford-Shah minimization
3. Adaptive frame methods for PDEs

Outline

1. Back to Compressed Sensing and beyond ℓ_1-minimization
   - Variational limits
   - Beyond ℓ_1-minimization
2. Free-discontinuity problems
   - The Mumford-Shah functional
   - Approximation by discrete functionals
   - A general discrete functional
   - Iterative thresholding algorithms for Mumford-Shah minimization
3. Adaptive frame methods for PDEs

Outline

1. Back to Compressed Sensing and beyond ℓ_1-minimization
   - Variational limits
   - Beyond ℓ_1-minimization
2. Free-discontinuity problems
   - The Mumford-Shah functional
   - Approximation by discrete functionals
   - A general discrete functional
   - Iterative thresholding algorithms for Mumford-Shah minimization
3. Adaptive frame methods for PDEs

Back to our functional: how do the parameters influence minimizers?

In our previous lecture we considered the following functional:

$$J(u,v) = \underbrace{\sum_{j=1}^{N} \| \sum_{\ell=1}^{M} T_{\ell,j} u^{\ell} - g^{j} \|_{\mathcal{K}}^{2}}_{:=\mathcal{D}(u)} + \sum_{\lambda \in \Lambda} v_\lambda \left(\sum_{\ell=1}^{M} |u_\lambda^\ell|^q \right)^{1/q}$$

$$+ \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{\ell=1}^{M} |u_\lambda^\ell|^2$$

$$+ \sum_{\lambda \in \Lambda} \theta_\lambda (\rho_\lambda - v_\lambda)^2.$$

where $u = (u_\lambda^\ell)_{\lambda \in \Lambda, \ell=1,\dots,M}$, and $v = (v_\lambda)_\lambda \geq 0$.

Variational limits and De Giorgi's Γ-convergence

Given a sequence of functionals $(F_h)_{h\in\mathbb{N}}$ on a topological space X, we wonder whether there exists a target "limit" functional F such that

$$\underbrace{\lim_h \quad F_h}_{\text{the limit is understood in some sense to be clarified}} \qquad = F$$

with the property that if $(x_h)_{h\in\mathbb{N}}$ is a sequence of minimizers of $(F_h)_{h\in\mathbb{N}}$ then

$$\lim_h x_h = x \in \operatorname{argmin}_y F(y).$$

For "nice" lower-semicontinuous functionals this convergence is given by the so-called Γ-convergence introduced by De Giorgi (1975). We emphasize that in general Γ-convergence does NOT coincide with pointwise convergence!

Just recalling notations

The second variable of the minimizers do depend on the first by the following formula:

$$\left(V_{\theta,\rho}^{(q)}(u) \right)_\lambda := \begin{cases} \rho_\lambda - \frac{1}{2\theta_\lambda}\|u_\lambda\|_q, & \|u_\lambda\|_q < 2\theta_\lambda\rho_\lambda \\ 0, & \text{otherwise} . \end{cases}$$

Approaching again ℓ_1-minimization

$$J(u,v) = J^{(q)}_{\theta,\rho,\omega}(u,v)$$
$$= \quad \|Tu-g\|^2_{\mathcal{K}} + \sum_{\lambda\in\Lambda}\omega_\lambda\|u_\lambda\|_2^2 + \sum_{\lambda\in\Lambda} v_\lambda\|u_\lambda\|_q + \sum_{\lambda\in\Lambda}\theta_\lambda(\rho_\lambda - v_\lambda)^2$$

Theorem (Fornasier and Rauhut '07)

Let $u^\star_{\theta,\rho,\omega}$ be the first component of the minimizer of $J(u,v)$, we have $u^\star_{\theta,\rho,\omega} \overset{\theta\to\infty}{\underset{\omega\to 0}{\rightharpoonup}} u^\star$ which is a minimizer of $J_{1,\rho}$. In particular,
$\Gamma - \lim_{\tau\to 0}\mathcal{J}(u, V^{(q)}_{\theta,\rho}(u)) = J_{1,\rho}(u)$,

$$J_{1,\rho}(u) = \mathcal{D}(u) + \sum_{\lambda}\rho_\lambda\|u_\lambda\|_q.$$

Approaching sparse recovery

We define

$$\underbrace{\Phi^{(q)}_{\theta,\rho,\omega}(u,v) = \sum_{\lambda\in\Lambda} v_\lambda \|u_\lambda\|_q + \sum_{\lambda\in\Lambda} \omega_\lambda \|u_\lambda\|_2^2 + \sum_{\lambda\in\Lambda} \theta_\lambda (\rho_\lambda - v_\lambda)^2}_{\text{our joint-sparsity measure}}$$

and

$$\mathcal{J}^1_\tau(u,v) := \|Tu - g\|^2_{\mathcal{K}} + \tau \Phi^{(q)}_{\theta,\rho,\omega}(u,v).$$

Theorem (Fornasier and Rauhut '07)

Let $u^\star_\tau$ be the first component of the minimizer of $\mathcal{J}^1_\tau(u,v)$, then

$$u^\star_\tau \overset{\tau\to 0}{\rightharpoonup} u^\star \in \operatorname{argmin}_{Tu=g} \Phi^{(q)}_{\theta,\rho,\omega}(u, V^{(q)}_{\theta,\rho}(u)).$$

In particular, $\Gamma - \lim_{\tau\to 0} \mathcal{J}^1_\tau(u, V^{(q)}_{\theta,\rho}(u)) = \|Tu - g\|^2$.

Outline

1. Back to Compressed Sensing and beyond ℓ_1-minimization
   - Variational limits
   - Beyond ℓ_1-minimization
2. Free-discontinuity problems
   - The Mumford-Shah functional
   - Approximation by discrete functionals
   - A general discrete functional
   - Iterative thresholding algorithms for Mumford-Shah minimization
3. Adaptive frame methods for PDEs

CS and ℓ_1-minimization

We recall that for a matrix Φ which satisfies the RIP of order $2k$ for $\delta_{2k} \leq \delta < 1/3$. Then, for any $x \in \mathbb{R}^N$ with $\#(\text{supp}(x_0)) \leq k$ and $y = \Phi x$, the decoder

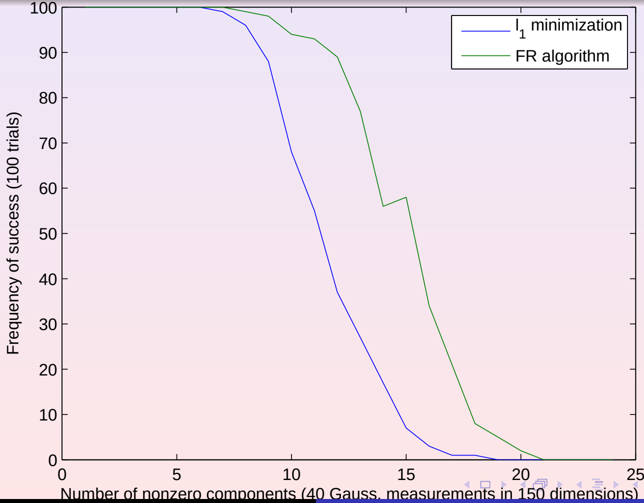
$$\Delta(y) := \text{argmin}_{\Phi z = y} \|z\|_{\ell_1^N},$$

reconstructs $x = \Delta(y)$ exactly. What happens if we define as a decoder

$$\Delta(y) := \text{argmin}_{\Phi z = y} \Phi^{(1)}_{\theta,\rho,\omega}(z, V^{(1)}_{\theta,\rho}(z)),$$

for some $M = 1$, and some parameters ρ, θ and

We surprisingly outperform the results of ℓ_1-minimization!



Outline

1. Back to Compressed Sensing and beyond ℓ_1-minimization
   - Variational limits
   - Beyond ℓ_1-minimization
2. Free-discontinuity problems
   - The Mumford-Shah functional
   - Approximation by discrete functionals
   - A general discrete functional
   - Iterative thresholding algorithms for Mumford-Shah minimization
3. Adaptive frame methods for PDEs

Free-discontinuity problems

The terminology "free-discontinuity problems" was introduced by E. De Giorgi [1] to indicate a class of variational problems that consist in the minimization of a functional, involving both volume and a surface energy, depending on a closed set K, and a function u usually smooth outside of K.

- K is not fixed a priori and it is an unknown of the problem;
- K is not a boundary in general, but a free-surface inside the domain of the problem.

[1] E. De Giorgi, *Free-discontinuity problems in calculus of variations*, Frontiers in pure and applied mathematics, a collection of papers dedicated to J.-L. Lions on the occasion of his 60^{th} birthday, R. Dautray, ed. North Holland, 1991, 55-62.

A prominent example in visual analysis

The Mumford-Shah functional is defined by

$$J(u,K) := \int_{\Omega\setminus K} \left[|\nabla u|^2 + \alpha(u-g)^2 \right] dx + \beta \mathcal{H}^{d-1}(K \cap \Omega).$$

The set Ω is a bounded open subset of $\mathbb{R}^d$, $\alpha, \beta > 0$ are fixed parameters and $g \in L^\infty(\Omega)$.
We seek for a function $u \in W^{1,2}(\Omega \setminus K)$ which approximates the datum g, smooth out of the discontinuity set K.
In visual analysis the set K is used in order to *segment* the image into connected components.

The relevant unknown

If the set K were fixed, then the minimization of $J(u, K)$ with respect to u is a relatively easy problem whose solution is given by

$$\begin{aligned} \Delta u &= \alpha(u - g), && \text{in } \Omega \setminus K, \\ \frac{\partial u}{\partial \nu} &= 0, && \text{on } \partial\Omega \cup K. \end{aligned}$$

So, the relevant unknown in free-discontinuity problems is the set K!

SBV functions

The minimization of J in both u and K is not easy because there is NO topology on the closed sets which ensures

- compactness of minimizing sequences and
- lower semicontinuity of the Hausdorff measure.

However, we can think $u \in BV(\Omega) \cap W^{1,2}(\Omega \setminus K)$, with $K = \overline{S}_u$ (the discontinuity set of u), and minimize J in this space. Unfortunately this space is TOO large. It contains Cantor-like functions such that $\nabla u = 0$ a.e. and $\mathcal{H}^{d-1}(S_u) = 0$, dense in $L^2(\Omega)$, so that the problem is trivialized.

SBV functions

Instead it is possible to give a meaningful formulation of the functional J in $SBV(\Omega)$ (the special $BV(\Omega)$ functions, with vanishing Cantor part),

$$\mathcal{F}(u) = \int_{\Omega \setminus S_u} \left[|\nabla u|^2 + \alpha (u-g)^2 \right] dx + \beta \mathcal{H}^{d-1}(S_u).$$

Γ-convergence

Definition

Let (X, d) be a metric space and let $f, f_n : X \to [0, \infty]$ be functions. We say that $(f_n)_n$ Γ-converges to f if the following two conditions are satisfied:

i) for any sequence $(x_n)_n \subset X$ converging to x, the following holds:

$$\liminf_n f_n(x_n) \geq f(x);$$

ii) for any $x \in X$ there exists a sequence $(x_n)_n \subset X$ converging to x such that

$$\limsup_n f_n(x_n) \leq f(x);$$

Eventually the minimizers of f_n do converge to a minimizer of f.

Γ-approximation by Ambrosio and Tortorelli

Let us define

$$J_\varepsilon(u,v) := \int_\Omega \left[v|\nabla u|^2 + \alpha(u-g)^2 \right] + \frac{\beta}{2}\left(\varepsilon |\nabla v|^2 + \frac{(1-v)^2}{\varepsilon} \right) dx.$$

The function $0 \leq v \leq 1$ tends to indicate S_u for $\varepsilon \to 0$. We have

Theorem (Ambrosio-Tortorelli '92)

The functional $J_{\varepsilon_n}(u,v)$ (v depends on u in a suitable way) Γ-converges in $(L^2(\Omega))^2$ to $\mathcal{F}(u)$ for any infinitesimal sequence $(\varepsilon_n)_n$.

Outline

1. Back to Compressed Sensing and beyond ℓ_1-minimization
   - Variational limits
   - Beyond ℓ_1-minimization
2. Free-discontinuity problems
   - The Mumford-Shah functional
   - Approximation by discrete functionals
   - A general discrete functional
   - Iterative thresholding algorithms for Mumford-Shah minimization
3. Adaptive frame methods for PDEs

The Blake-Zisserman's discrete functional

Let $d = 2$, $\Omega = [0,1]^2$, and $u_{i,j} = u(hi, hj)$ a discrete function defined on $\Omega_h := \Omega \cap h\mathbb{Z}^2$, for $h > 0$. Define $W_h(t) = \min\{t^2, \beta/h\}$ the truncated quadratic potential and

$$
\begin{aligned}
J_h^\delta(u) &= h^2 \sum_{(hi,hj)\in\Omega_h} W_h\left(\frac{u_{i+1,j}-u_{i,j}}{h}\right) \\
&+ h^2 \sum_{(hi,hj)\in\Omega_h} W_h\left(\frac{u_{i,j+1}-u_{i,j}}{h}\right) \\
&+ \alpha h^2 \sum_{(hi,hj)\in\Omega_h} (u_{i,j}-g_{i,j})^2.
\end{aligned}
$$

The Blake-Zisserman's discrete functional

Theorem (Chambolle '95)

The functional $J_h^\delta(u)$ Γ-converges in $\mathcal{B}(\Omega)$ to

$$\mathcal{E}(u) = \int_{\Omega \setminus S_u} \left[|\nabla u|^2 + \alpha(u-g)^2 \right] dx + \beta \mathcal{C}(S_u),$$

for $h \to 0$, where $\mathcal{C}$ is the so-called "cab-driver" measure.

Basically $\mathcal{C}$ measures the length of a curve only through its projections along horizontal and vertical axes; for a regular C^1 curve $c = \gamma([0,1])$, with $\gamma(t) = (\gamma_1(t), \gamma_2(t)) \in \Omega$, we have

$$\mathcal{C}(c) = \int_0^1 |\gamma_1'(t)| + |\gamma_2'(t)| dt.$$

Outline

The derivation

For simplicity assume $d = 1$. We consider $\Omega = [0, 1]$, and $u_i = u(hi)$ a discrete function defined on $\Omega_h := \Omega \cap h\mathbb{Z}$, for $h > 0$. Define $W_h(t) = \min\{t^2, \frac{1}{2h}\}$ and

$$J_h^\delta(u) = h \sum_{(hi)\in\Omega_h} W_h\left(\frac{u_{i+1} - u_i}{h}\right) + h \sum_{(hi)\in\Omega_h} (u_i - g_i)^2.$$

Again, it is possible to show that this discrete functional Γ-converges to a suitable functional *a la* Mumford-Shah.

The derivation

Let us denote $S_u^\delta = \{hi : \left|\frac{u_{i+1}-u_i}{h}\right| \leq \frac{1}{\sqrt{h}}\}$ and

$$v(i) = \begin{cases} 1 & i \in \Omega_h \setminus S_u^\delta \\ 0 & i \in S_u^\delta \end{cases}$$

Hence, we have

$$\begin{aligned} J_h^\delta(u) &= h \sum_{(hi)\in\Omega_h} v(i) \left|\frac{u_{i+1}-u_i}{h}\right|^2 + \frac{1}{2} \sum_{(hi)\in\Omega_h} (1-v(i))^2 \\ &+ h \sum_{(hi)\in\Omega_h} (u_i - g_i)^2. \end{aligned}$$

The derivation

Inspired by Ambrosio-Tortorelli, we can even dare to consider v as an additional variable

$$
\begin{aligned}
J_h^\delta(u,v) &= h \sum_{(hi)\in\Omega_h} v(i) \left| \frac{u_{i+1}-u_i}{h} \right|^2 + \frac{1}{2} \sum_{(hi)\in\Omega_h} (1-v(i))^2 \\
&+ h \sum_{(hi)\in\Omega_h} (u_i - g_i)^2 .
\end{aligned}
$$

Indeed the miminization w.r.t. $v \geq 0$ for u fixed gives though

$$
v(i) = \begin{cases} 1 - h \left| \frac{u_{i+1}-u_i}{h} \right|^2, & i \in \Omega_h \setminus S_u^\delta \\ 0, & i \in S_u^\delta \end{cases}
$$

The derivation

Plugging this minimal $v \geq 0$ into the functional we obtain

$$J_h^\delta(u,v) = J_h^\delta(u) \underbrace{- \frac{h^2}{2} \sum_{(hi) \in \Omega_h \setminus S_u^\delta} \left| \frac{u_{i+1} - u_i}{h} \right|^4}_{\text{Additional term}}.$$

To make disappear the additional term we need to consider (at least)

$$\mathcal{J}_h^\delta(u,v) = J_h^\delta(u,v) + \frac{h^2}{2} \sum_{(hi) \in \Omega_h \setminus S_u^\delta} \left| \frac{u_{i+1} - u_i}{h} \right|^4.$$

Equivalence of minimizers

Theorem

The u-component of the minimizers (u, v) of the functional

$$
\begin{aligned}
\mathcal{J}_h^\delta(u,v) &= h \sum_{(hi)\in\Omega_h} v(i) \left| \frac{u_{i+1}-u_i}{h} \right|^2 + \frac{1}{2} \sum_{(hi)\in\Omega_h} (1-v(i))^2 \\
&+ \frac{h^2}{2} \sum_{(hi)\in\Omega_h \setminus S_u^\delta} \left| \frac{u_{i+1}-u_i}{h} \right|^4 \\
&+ h \sum_{(hi)\in\Omega_h} (u_i - g_i)^2,
\end{aligned}
$$

are minimizers of $J_h^\delta(u)$.

Discrete derivative

For $(u(i))_{hi \in \Omega_h}$ we define the discrete derivative as the matrix D_h which maps $(u(i))_{hi \in \Omega_h}$ into $\left(\frac{u(i+1)-u(i)}{h}\right)_i$, given by

$$D_h = \frac{1}{h}\begin{pmatrix} -1 & 1 & 0 & \dots & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 \\ \dots & & & & & \\ 0 & 0 & \dots & \dots & -1 & 1 \end{pmatrix}.$$

We have

$$u = D_h^\dagger D_h u + c_u,$$

where $= D^\dagger$ is the pseudo-inverse matrix of D_h and c_u is constant vector. Observe that $\langle D_h^\dagger D_h u, c\rangle = 0$, for any constant c.

A functional of the derivative

We can reformulate the problem in terms of derivatives

$$
\begin{aligned}
\mathcal{J}_h^\delta(z,v) &= h\sum_i v(i)\,|z_i|^2 + \frac{1}{2}\sum_i (1-v(i))^2 \\
&+ \frac{h^2}{2}\sum_{(hi)\in\Omega_h\setminus S_u^\delta} |z_i|^4 \\
&+ h\sum_i (D_h^\dagger z_i - f_i)^2,
\end{aligned}
$$

where $z = D_h u$, $f = D_h^\dagger D_h g$, and $c_u = c_g$ is assumed at the minimizer.

A general discrete functional: surprise ...

On the basis of this derivation, we are interested to study the minimizers of the following general form of functionals

$$\begin{aligned} J(u,v) = \|Tu-g\|^2_{\mathcal{K}} &+ \sum_{\lambda\in\Lambda} v_\lambda |u_\lambda|^p \\ &+ \sum_{\lambda\in\Lambda} \omega_\lambda |u_\lambda|^{2p} \\ &+ \sum_{\lambda\in\Lambda} \theta_\lambda (\rho_\lambda - v_\lambda)^2, \end{aligned}$$

where Λ is a countable index set, $u = (u_\lambda)_\lambda$ is a vector in $\ell^2(\Lambda)$, $T : \ell^2(\Lambda) \to \mathcal{K}$ is a bounded operator, $g \in \mathcal{K}$, and $1 \le p \le 2$.

Outline

Convexity

Lemma.

Assume $1 < p \leq 2$ and $\rho, \omega, \theta > 0$. If $(\sigma_{\min} + \omega_\lambda)\theta_\lambda \geq \frac{p^2}{4p(2p-1)}$, $\lambda \in \Lambda$, then $J(u, v)$ is convex, and strictly convex for $(\sigma_{\min} + \omega_\lambda)\theta_\lambda > \frac{p^2}{4p(2p-1)}$.

For some initial choice $v^{(0)}$, for example $v^{(0)} \equiv (\rho_\lambda)_{\lambda\in\Lambda}$, define

$$\begin{cases} u^{(n+1)} := \operatorname{argmin}_{u\in\ell_{2p}(\Lambda)} J(u, v^{(n)}) \\ v^{(n+1)} := \operatorname{argmin}_{0\le v_\lambda\le\rho_\lambda} J(u^{(n+1)}, v). \end{cases}$$

This algorithm was already studied for $p = 1$. The analysis for $1 < p \le 2$ is analogous but certain differences arise.

Theorem (Fornasier '07)

Let $1 < p \le \infty$, $\theta_\lambda(\sigma_\lambda + \omega_\lambda) > \frac{p^2}{4p(2p-1)}$, and $\omega_\lambda \ge \gamma > 0$, for all $\lambda \in \Lambda$. Then the sequence $(u^{(n)}, v^{(n)})$ generated by the algorithm converges to $(u^*, v^*) \in \ell_{2p}(\Lambda) \times \ell_{\infty,\rho^{-1}}(\Lambda)$, the unique minimizer of J. The convergence of $u^{(n)}$ is in the weak sense and the convergence of $v^{(n)}$ is componentwise.

Again, the minimization with respect to the second variable is explicit:

$$v_\lambda^{(n)} = \begin{cases} \rho_\lambda - \frac{|u_\lambda^{(n)}|^p}{2\theta_\lambda}, & |u_\lambda^{(n)}|^p \leq 2\theta_\lambda \rho_\lambda \\ 0, & \text{otherwise.} \end{cases}$$

For the minimization with respect to the first variable we have

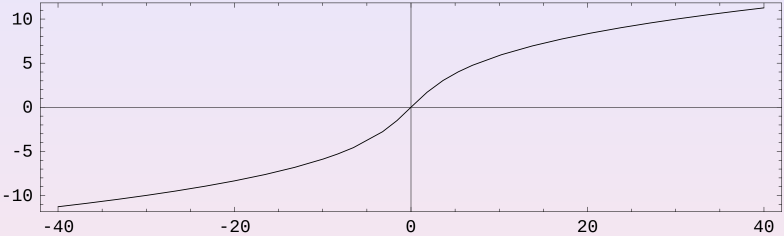
Theorem

If $\|T\| < 1$, $1 < p \leq 2$ and $\omega_\lambda \geq \gamma > 0$, for all $\lambda \in \Lambda$, then for any initial choice $u^{(0)} \in \ell_{2p}(\Lambda)$, the sequence

$$u^{(m+1)} := \mathbb{U}_{v,\omega}^{(p)}\left(u^{(m)} + T^*(g - Tu^{(m)})\right),$$

converges to the minimizer of $J(\cdot, v)$.

The thresholding operator $\mathbb{U}^{(p)}_{v,\omega}$



The operator $\mathbb{U}^{(p)}_{v,\omega}$ acts componentwise by applying curves $U^p_{v,\omega}$ as the one depicted above. It is defined as

$$\begin{aligned}\mathbb{U}^{(p)}_{v,\omega}(z) &:= \operatorname{argmin}_u \mathcal{J}^{surr}(u,v,z)\\ &:= \operatorname{argmin}_u J(u,v) + \|u-z\|^2 - \|T(u-y)\|^2.\end{aligned}$$

The single iteration algorithm

Instead of an alternating minimization we may want to consider:

$$(u^{(n+1)}, v^{(n+1)}) = \text{argmin}_{(u,v)\in\ell_{2p}(\Lambda)\times\ell_{\infty,\rho^{-1}}(\Lambda)} \mathcal{J}^{sur}(u, v, u^{(m)}).$$

Observe that in this iteration the variable v does not play any role.

Theorem (Fornasier '07)

The latter algorithm is equivalent to the following iterative thresholding algorithm:

$$u^{(n+1)} = \mathbb{H}^p_{\omega,\theta,\rho}\left(u^{(n)} + T^*(g - Tu^{(n)})\right).$$

The nonlinear thresholding operator $\mathbb{H}^p_{\omega,\theta,\rho} : \ell_2(\Lambda) \to \ell_2(\Lambda)$ *is defined componentwise* $\mathbb{H}^p_{\omega,\theta,\rho}(\xi) = (H^p_{\omega_\lambda,\theta_\lambda,\rho_\lambda}(\xi_\lambda))_{\lambda\in\Lambda}$, *where for* $z \in \mathbb{R}$,

The single iteration algorithm

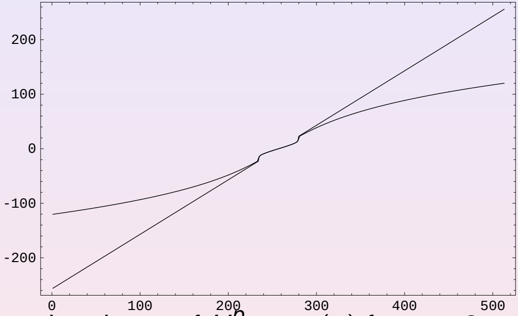
we have

$$H^p_{\omega_\lambda,\theta_\lambda,\rho_\lambda}(z) := \lim_{m\to\infty} z^{(m)},$$

and $z^{(m)}$ is defined by the algorithm:

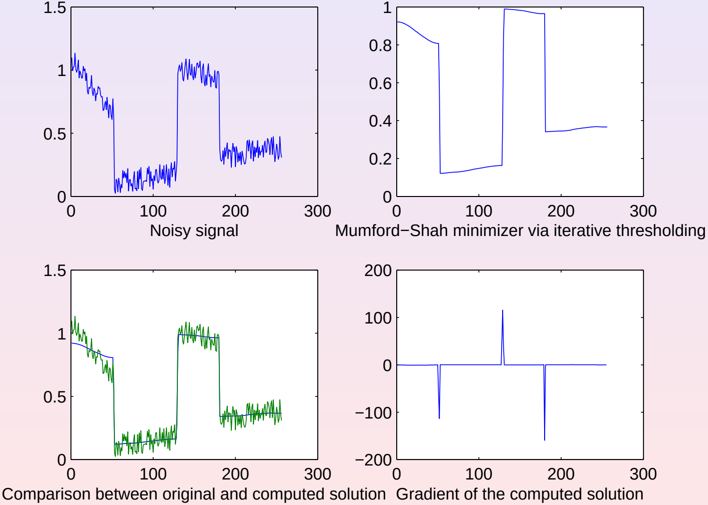
$$\begin{aligned} z^{(m+1)} &= U^p_{v^{(m)},\omega_\lambda}(z), \\ v^{(m+1)} &= \begin{cases} \rho_\lambda - \frac{|z^{(m+1)}|^p}{2\theta_\lambda}, & |z^{(m+1)}|^p \leq 2\theta_\lambda\rho_\lambda \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Typical shape of the function $H^p_{\omega_\lambda,\theta_\lambda,\rho_\lambda}(z)$



We illustrate the shape of $H^p_{\omega_\lambda,\theta_\lambda,\rho_\lambda}(z)$ for $p = 2$ and for different choice of the parameters.

Signal denoising via Mumford-Shah minimization



Signal segmentation via Mumford-Shah minimization

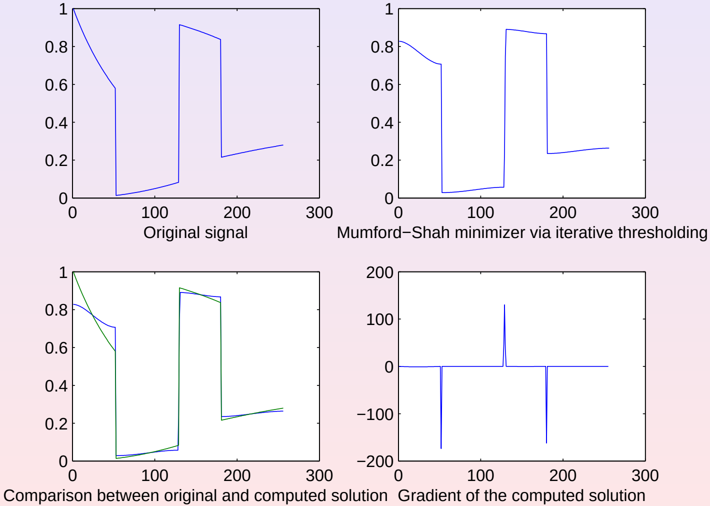


Image segmentation via Mumford-Shah minimization

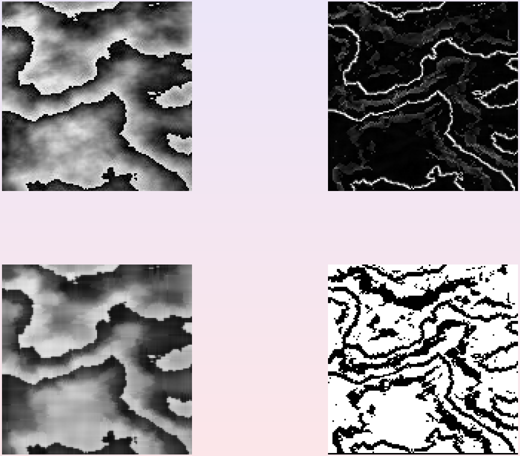


Image denoising via Mumford-Shah minimization

Original image

Noisy image

Mumford−Shah minimizer via iterative thresholding

Discrete discontinuity set

Image segmentation via Mumford-Shah minimization

Original image

The gradient of the computed minimizer

Mumford−Shah minimizer via iterative thresholding

Discrete discontinuity set

Elliptic equation: theoretical setting

- $\Omega \subset \mathbb{R}^d$ Lipschitz domain
- H Hilbert space, $H \subset L_2(\Omega) \subset H'$ Gelfand triple
- $\mathcal{L} : H \to H'$ linear, boundedly invertible

$$a(\cdot,\cdot) := \langle \mathcal{L}\cdot,\cdot\rangle : H \times H \to \mathbb{R}$$

  symmetric, continuous, H-elliptic: $a(v,v) \eqsim \|v\|_H^2$.
- task: for $f \in H'$, solve

$$\mathcal{L}u = f,$$

  i.e., find $u \in H$ s.t.

$$a(u,v) = \langle f,v\rangle, \quad v \in H.$$

Discretization via bases

Cohen, Dahmen, DeVore (1998):

- I choose wavelet basis $\Psi = \{\psi_\lambda\}_{\lambda \in \mathcal{J}} \subset L_2(\Omega)$ with

$$\|\mathbf{c}\|_{\ell_2} \eqsim \|\mathbf{c}^T D^{-1} \Psi\|_H.$$

- This implies

$$\mathcal{L}u = f \Rightarrow \mathbf{L}\mathbf{u} = \mathbf{f},$$

  where

$$\begin{aligned} \mathbf{L} &:= D^{-1}\langle \mathcal{L}\Psi, \Psi \rangle D^{-1}, \\ \mathbf{f} &:= D^{-1}\langle f, \Psi \rangle. \end{aligned}$$

- problem: construction of wavelet bases on general domains difficult/complicated!

Discretization via frames

Dahlke, Fornasier, Raasch, Stevenson (2003-2006)

- choose a frame $\Psi = \{\psi_\lambda\}_{\lambda \in \mathcal{J}} \subset L_2(\Omega)$ with

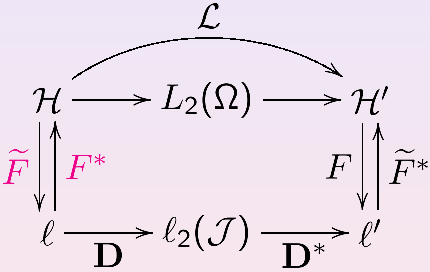
$$\|g\|_{L_2} \eqsim \|\langle g, \Psi \rangle\|_{\ell_2}.$$

- Ψ is a Gelfand frame for (H, L_2, H') if

$$F^* : \ell \to H : c \mapsto c^T \Psi, \quad \tilde{F} : H \to \ell : g \mapsto \langle g, \tilde{\Psi} \rangle$$

are bounded.

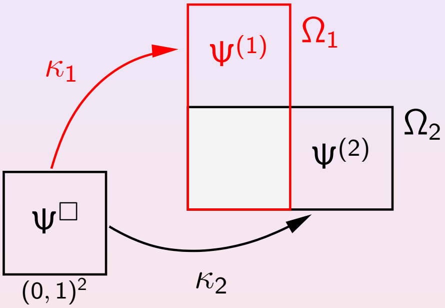
Mapping diagram



Construction of wavelet Gelfand frames (Dahlke, Fornasier, Raasch '04)

- $H = H_0^t(\Omega)$;
- reference Riesz basis $\Psi^\square \subset H_0^t(\square)$, $\square := (0,1)^d$;
- overlapping decomposition $\Omega = \sum_{i=1}^n \Omega_i$;
- $\kappa_i : \square \to \Omega_i$, C^m-diffeomorphisms, $m \geq t$;
- appropriate lifting yields Gelfand frames $\Psi = \bigcup_{i=1}^n \Psi_i$.

Gelfand frame construction



Solution of the discrete problem

The operator $\mathbf{L} : \ell_2(\mathcal{J}) \to \ell_2(\mathcal{J})$ is going to be bounded, symmetric, positive, but has nontrivial kernel. Nevertheless we have

$$\mathbf{L} : ran(L) \to ran(L)$$

is boundedly invertible;

- (ideal) damped Richardson iteration:

$$(R) \quad \mathbf{u}^{(n+1)} = \mathbf{u}^{(n)} + \mathbf{r}^{(n)}, \quad \mathbf{r}^{(n)} = \mathbf{f} - \mathbf{L}\mathbf{u}^{(n)}$$

- (ideal) steepest descent method

$$(SD) \quad \mathbf{u}^{(n+1)} = \mathbf{u}^{(n)} + \frac{\langle \mathbf{r}^{(n)}, \mathbf{r}^{(n)} \rangle}{\langle \mathbf{L}\mathbf{r}^{(n)}, \mathbf{r}^{(n)} \rangle} \mathbf{r}^{(n)}.$$

Basic discrete procedures

Assume that we have the following procedures at our disposal:

- **RHS**$[\varepsilon, \mathbf{g}] \to \mathbf{g}_\varepsilon$: determines for $\mathbf{g} \in \ell_2(\mathcal{J})$ a finitely supported $\mathbf{g}_\varepsilon \in \ell_2(\mathcal{J})$ such that
$$\|\mathbf{g} - \mathbf{g}_\varepsilon\|_{\ell_2(\mathcal{J})} \leq \varepsilon; \tag{1}$$
- **APPLY**$[\varepsilon, \mathbf{N}, \mathbf{v}] \to \mathbf{w}_\varepsilon$: determines for $\mathbf{N} \in B(\ell_2(\mathcal{J}))$ and for a finitely supported $\mathbf{v} \in \ell_2(\mathcal{J})$ a finitely supported $\mathbf{w}_\varepsilon$ such that
$$\|\mathbf{N}\mathbf{v} - \mathbf{w}_\varepsilon\|_{\ell_2(\mathcal{J})} \leq \varepsilon; \tag{2}$$
- **COARSE**$[\varepsilon, \mathbf{v}] \to \mathbf{v}_\varepsilon$: determines for a finitely supported $\mathbf{v} \in \ell_2(\mathcal{J})$ a finitely supported $\mathbf{v}_\varepsilon \in \ell_2(\mathcal{J})$ with at most N significant coefficients, such that
$$\|\mathbf{v} - \mathbf{v}_\varepsilon\|_{\ell_2(\mathcal{J})} \leq \varepsilon. \tag{3}$$

Moreover, $N \lesssim N_{\min}$ holds, $N_{\min}$ being the minimal number

The algorithm

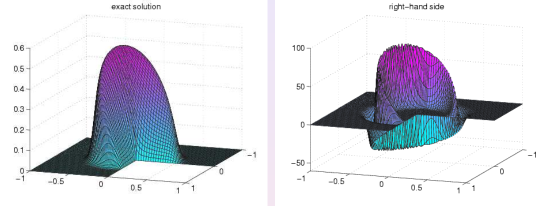
SOLVE$[\varepsilon, \mathbf{L}, \mathbf{f}] \to \mathbf{u}_\varepsilon$:
Let $\theta < 1/3$ and $K \in \mathbb{N}$ be fixed such that $3\rho^K < \theta$.
$i := 0,\ \mathbf{v}^{(0)} := 0,\ \varepsilon_0 := \|\mathbf{L}^{-1}_{|\operatorname{ran}(\mathbf{L})}\| \|\mathbf{f}\|_{\ell_2(\mathcal{J})}$
While $\varepsilon_i > \varepsilon$ do
$i := i + 1$
$\varepsilon_i := 3\rho^K \varepsilon_{i-1}/\theta$
$\mathbf{f}^{(i)} := \mathbf{RHS}[\frac{\theta\varepsilon_i}{6\alpha K}, \mathbf{f}]$
$\mathbf{v}^{(i,0)} := \mathbf{v}^{(i-1)}$
For $j = 1, ..., K$ do
$\mathbf{v}^{(i,j)} := \mathbf{v}^{(i,j-1)} - \alpha(\mathbf{APPLY}[\frac{\theta\varepsilon_i}{6\alpha K}, \mathbf{L}, \mathbf{v}^{(i,j-1)}] - \mathbf{f}^{(i)})$
od
$\mathbf{v}^{(i)} := \mathbf{COARSE}[(1-\theta)\varepsilon_i, \mathbf{v}^{(i,K)}]$
od
$\mathbf{u}_\varepsilon := \mathbf{v}^{(i)}$.

The proof of convergence and optimality

Convergence and optimal complexity:

- (R) Riesz basis case: Cohen, Dahmen, DeVore (2000)
- (R) Frame case: Stevenson (2003)
- (SD) Riesz basis case: Dahmen, Urban, Vorloeper (2002) Canuto, Urban (2003)
- (SD) Dahlke, Fornasier, Raasch, Stevenson, Werner (2005)

Numerical examples



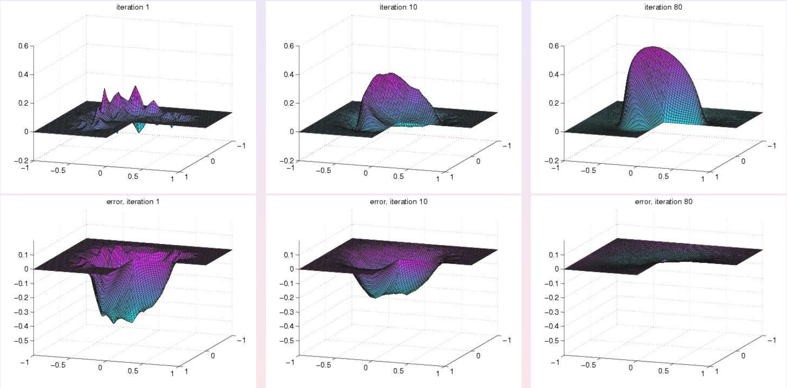
Let us consider the Poisson's equation on the *L*-shaped domain:

$$-\Delta u = f, \quad u|_{\partial\Omega} = 0,$$

and choose f with singularity at the re-entrant corner.

Numerical examples

some (SD) iterations:



Acknowledgment

- Program and Applied and Computational Mathematics (PACM), Princeton University & Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences
- European Union's Outgoing International Marie Curie Fellowship MOIF-CT-2006-039438

Acknowledgment

Co-authors

Maria Charina Inst. for Appl. Math., University of Dortmund Germany
Costanza Conti Energ. Dept. “Sergio Stecco”, University of Florence Italy
Stephan Dahlke Numer./Wavelet-Anal. Group, Univ. of Marburg Germany
Ingrid Daubechies Dept. Math., Princeton University USA
Ronal A. DeVore Dept, of Math., University of South Carolina, USA
Hans G. Feichtinger Numer. Harmon. Anal. Group, Univ. of Vienna Austria
Karlheinz Groechenig Numer. Harmon. Anal. Group, Univ. of Vienna Austria
Laura Gori, MeMoMat, Univ. of Rome “La Sapienza Italy”
Sinan Güntürk Courant Institute, New York University, USA
Ignace Loris, Dienst Theor. Natuurk., Vrije Universiteit Brussel Belgium
Riccardo March IAC “Mauro Picone” Rome, Nat. Counc. of Research Italy

Acknowledgement

Co-authors

Peter Markowich RICAM-OEAW/Dept. Maths, University of Vienna Austria
Maria Morandi Cecchi Depart. of Pure and Appl. Math., Univ. of Padua Italy
Francesca Pitolli MeMoMat, Univ. of Rome “La Sapienza Italy”
Thorsten Raasch Numer./Wavelet-Anal .Group, Univ. of Marburg Germany
Holger Rauhut Numer. Harmon. Anal. Group, Univ. of Vienna Austria
Carola B. Schönlieb, Faculty of Math., Univ. of Vienna Austria
Gabriele Steidl Fac. of Math. and Comput. Sci., Univ. of Mannheim Germany
Rob Stevenson Dept. of Math., Univ. of Utrecht Holland
Gerd Teschke, Konrad-Zuse-Zentrum fuer Technomat. Berlin Germany
Joel Tropp, Applied and Comput. Math., Eng. Appl. Sciences, Caltech, USA
Domenico Toniolo Department of Physics “G. Galilei”, Univ. of Padua Italy
Manuel Werner AG Numer./Wavelet-Anal. Group,Univ. of Marburg Gesrmany