On Nonconvex Regularized Models for Image Restoration Problems

Chunlin Wu
(Joint work with Zhifang Liu, Shuang Wen, Chao Zeng, Xue Feng)

School of Mathematical Sciences, Nankai University, China

Shanghai Jiao Tong Univ., Shanghai,
Dec., 6, 2018

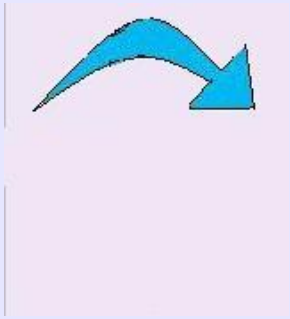
Outline

1. Variational signal and image restoration
2. A general truncated regularization framework
   - Assumptions on potential functions
   - Truncated regularization: motivation
   - Application in 2D
3. Lower bound theory
   - Lower bound theory for anisotropic model
   - Lower bound theory for isotropic model
4. Conclusions
5. References

Section 1. Variational signal and image restoration

Variational signal and image restoration - image degradation

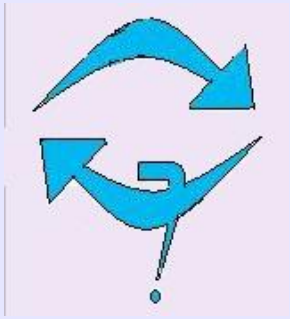
true image



blurred and noisy observation

Variational signal and image restoration - image restoration

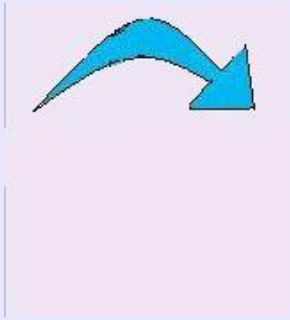
true image

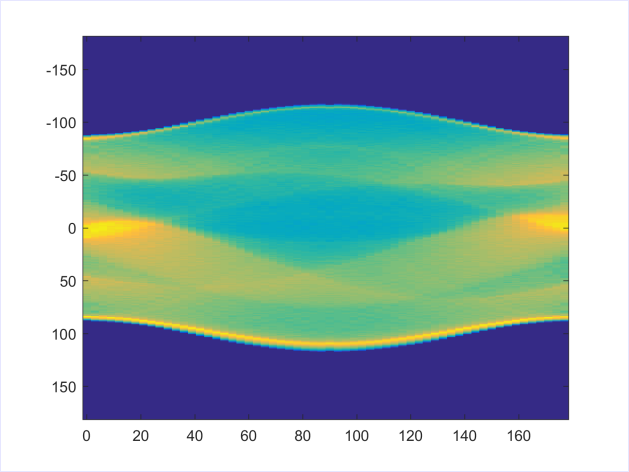


blurred and noisy observation

Variational signal and image restoration - image degradation

true image

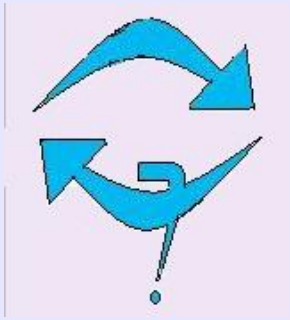


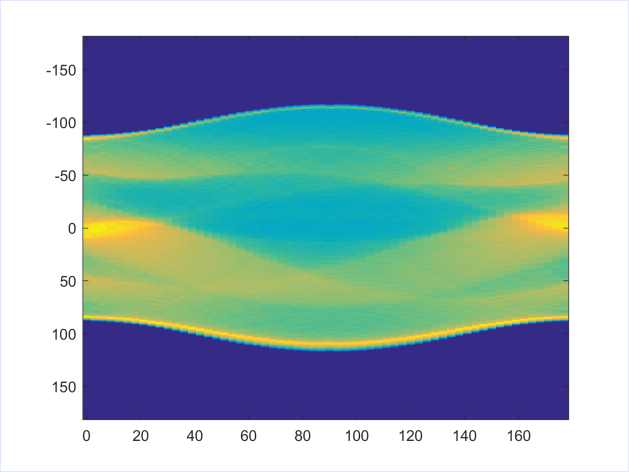


sampled data

Variational signal and image restoration - image reconstruction

true image





sampled data

Variational signal and image restoration - 1D signal

- 1D signal $f \in \mathbb{R}^K$ is a degradation of $\underline{u} \in \mathbb{R}^N$:

$$\underline{u} \to A\underline{u} \xrightarrow{n} f,$$

- A is a linear operator such as a blur convolution.
- $n = \{n_i : 1 \leq i \leq K\}$ is random noise.
- $n_i \sim \mathcal{N}(0, \sigma^2), 1 \leq i \leq K$ are i.i.d. random variables, such as additive Gaussian noise and multiplicative noise.

Variational signal and image restoration - minimization model

$$\min_{u\in\mathbb{R}^N}\left\{\begin{aligned} E(u) &= \sum_{1\le i\le N}\varphi((\nabla_x u)_i) + \frac{\alpha}{2}\|Au-f\|^2_{\mathbb{R}^K} \\ &= \sum_{1\le i\le N}\rho(|(\nabla_x u)_i|) + \frac{\alpha}{2}\|Au-f\|^2_{\mathbb{R}^K}\end{aligned}\right\}, \tag{1}$$

- $\rho(\cdot)$ is a potential function.
- ∇_x is the forward difference operator with a specific boundary condition, e.g., the periodic or Neumann boundary condition.

Variational signal and image restoration - 2D image

- 2D image $\mathbf{f} \in \mathbb{R}^{K\times K}$ is a degradation of $\underline{\mathbf{u}} \in \mathbb{R}^{N\times N}$:

$$\underline{\mathbf{u}} \to \mathbf{A}\underline{\mathbf{u}} \xrightarrow{\mathbf{n}} \mathbf{f},$$

- For image restoration problem, $\mathbf{A}$ is a linear operator such as a blur convolution.
- For image reconstruction problem, $\mathbf{A}$ is a radon transform.
- $\mathbf{n} = \{\mathbf{n}_{i,j}, 1 \le i,j \le K\}$ is random noise.
- $\{\mathbf{n}_{i,j} \sim \mathcal{N}(0,\sigma^2), 1 \le i,j \le K\}$ are i.i.d. random variables, such as additive Gaussian noise and multiplicative noise.

Variational signal and image restoration - minimization models

$$\min_{\mathbf{u}\in\mathbb{R}^{N\times N}} \left\{ \begin{aligned} E_{\text{ani}}(\mathbf{u}) &= \sum_{1\le i,j\le N} (\varphi((\nabla_x \mathbf{u})_{i,j}) + \varphi((\nabla_y \mathbf{u})_{i,j})) + \frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|^2_{\mathbb{R}^{K\times K}} \\ &= \sum_{1\le i,j\le N} (\rho(|(\nabla_x \mathbf{u})_{i,j}|) + \rho(|(\nabla_y \mathbf{u})_{i,j}|)) + \frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|^2_{\mathbb{R}^{K\times K}} \end{aligned} \right\}, \tag{2}$$

$$\min_{\mathbf{u}\in\mathbb{R}^{N\times N}} \left\{ \begin{aligned} E_{\text{iso}}(\mathbf{u}) &= \sum_{1\le i,j\le N} \psi((\nabla_x \mathbf{u})_{i,j}, (\nabla_y \mathbf{u})_{i,j}) + \frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|^2_{\mathbb{R}^{K\times K}} \\ &= \sum_{1\le i,j\le N} \rho(\sqrt{(\nabla_x \mathbf{u})^2_{i,j} + (\nabla_y \mathbf{u})^2_{i,j}}) + \frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|^2_{\mathbb{R}^{K\times K}} \end{aligned} \right\}, \tag{3}$$

- ∇_x and ∇_y are the forward difference operators with a specific boundary condition.

Variational signal and image restoration - convex regularization

- $\rho(s) = s^2$, Tikhonov regularization.
  Cannot preserve edges.
- $\rho(s) = s$, TV regularization.
  Cannot preserve contrast information.

Section 2. A general truncated regularization framework

Assumptions on potential functions

(AS1) $\rho(0) = 0, \rho(s) < +\infty, \forall s < +\infty$ with 0 as its strict minimizer;

(AS2) $\rho(s)$ is increasing over $[0, \infty)$;

(AS3) $\rho(s)$ is C^2 on $(0, +\infty)$;

(AS4) $\rho''(s) < 0$ strictly increases on $(0, +\infty)$ or $\rho''(s) \equiv 0$ on $(0, +\infty)$;

Assumptions on potential functions - some potential functions

when $0 < p < 1, \theta > 0, a > 2$

$\rho_1(s) = s$

$\rho_2(s) = s^p$

$\rho_3(s) = \ln(\theta s + 1)$

$\rho_4(s) = \frac{\theta s}{1+\theta s}$

$\rho_5(0) = 0, \rho_5(s) = 1$ if $s > 0$

$\rho_6(s) = \ln(\theta s^p + 1)$

$\rho_7(s) = \frac{\theta s^p}{1+\theta s^p}$

$$\rho_8(s) = \begin{cases} \theta s, & s \leq \theta \\ \frac{-s^2-\theta^2+2a\theta s}{2(a-1)}, & \theta < s < a\theta \\ \frac{(a+1)\theta^2}{2}, & s > a\theta \end{cases}$$

when $p > 1, \theta > 0$.

$\rho_9(s) = min\{\theta s^2, 1\}$

$\rho_2(s) = s^p$

$\rho_7(s) = \frac{\theta s^p}{1+\theta s^p}$

Truncated regularization: motivation - 1D signal

- Convex regularizer is impossible to perfectly recover a nonconstant signal.

Proposition (C. Wu, Z. Liu, S. Wen, Sci. China. Math. 2018)

Assume $\rho(\cdot)$ to be convex and satisfy (AS1)(AS2). If a signal $\tilde{u} \in \mathbb{R}^N$ can be recovered by the minimization problem (1) *with $f = A\tilde{u}$, then $\tilde{u} \in \mathbb{R}^N$ is a constant signal, i.e., $\tilde{u} = c(1, 1, \cdots, 1) \in \mathbb{R}^N$ for some $c \in \mathbb{R}$.*

Truncated regularization: motivation - truncated regularization framework

A new regularizer function

$$\overline{T}(\cdot) = \rho_\tau(\cdot) = \rho(\min(\cdot, \tau)), \tag{4}$$

- $\tau > 0$ is a positive real parameter.
- Flat on $(\tau, +\infty)$.
- If $\rho(\cdot)$ satisfies the basic assumptions (AS1)(AS2), $\overline{T}(\cdot) = \rho_\tau(\cdot)$ also satisfies the basic assumptions (AS1)(AS2).
- $\overline{T}(\cdot)$ is always nonconvex.

Truncated regularization: motivation - subadditivity of min function

Lemma

Given $a, b \geq 0, \tau > 0$, *then*

$$\min(a+b, \tau) \leq \min(a, \tau) + \min(b, \tau). \tag{5}$$

Proposition

Given $\tau > 0$, *if* $\rho(\cdot)$ *satisfies the subadditivity property over* $[0, +\infty)$ *and the assumptions (AS1)(AS2), then its truncated version* $\overline{T}(\cdot) = \rho(\min(\cdot, \tau))$ *also has the subadditivity property over* $[0, +\infty)$.

Truncated regularization: motivation - truncated regularization in 1D signal

- $\emptyset \neq \Omega \subsetneq J = \{1, \cdots, N\}$.
- $\mathbf{1}_\Omega$ be its indicator function and $\zeta > 0$ be a real number.
- $J_0 = \{i : (\nabla_x \mathbf{1}_\Omega)_i = 0\}$.
- $J_1 = \{i : (\nabla_x \mathbf{1}_\Omega)_i \neq 0\} = J \setminus J_0$.

Consider now the minimization problem (1) using truncated regularization where $f = A(\zeta \mathbf{1}_\Omega) \in \mathbb{R}^K$. Denote

$$E^\zeta(u) = \sum_{1 \leq i \leq N} \overline{T}(|(\nabla_x u)_i|) + \frac{\alpha}{2} \|A(u - \zeta \mathbf{1}_\Omega)\|^2_{\mathbb{R}^K}. \tag{6}$$

Truncated regularization: motivation - perfect recovery of 1D signal

The following theorem shows the perfect recovery (i.e., contrast preservation) of the signal $\zeta \mathbf{1}_\Omega$ by (1) with a truncated regularization.

Theorem (C. Wu, Z. Liu, S. Wen, Sci. China. Math. 2018)

If $\zeta > \tau + \sqrt{\frac{4\overline{T}(\tau)}{\alpha\mu_{\min}}\#J_1}$, *then the global minimizer is* $\zeta \mathbf{1}_\Omega$. *Here* $\mu_{\min} > 0$ *is the minimal eigenvalue of* $A^T A$.

Application in 2D
- truncated regularization in 2D image

- *Anisotropic 2D truncated regularization model*:

$$\min_{\mathbf{u}\in V}\left\{E_{\text{ani}}^{\overline{T}}(\mathbf{u}) = \sum_{1\le i,j\le N}\overline{T}(|(\nabla_x\mathbf{u})_{ij}|) + \overline{T}(|(\nabla_y\mathbf{u})_{ij}|) + \frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|_V^2\right\}, \tag{7}$$

- *Isotropic 2D truncated regularization model*:

$$\min_{\mathbf{u}\in V}\left\{E_{\text{iso}}^{\overline{T}}(\mathbf{u}) = \sum_{1\le i,j\le N}\overline{T}\left(\sqrt{(\nabla_x\mathbf{u})_{ij}^2 + (\nabla_y\mathbf{u})_{ij}^2}\right) + \frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|_V^2\right\}, \tag{8}$$

which are truncated versions of (2) and (3), respectively.

Application in 2D - ADMM solver

Constrained optimization problem:

$$\min_{(\mathbf{u},\mathbf{q})\in V\times Q}\left\{\tilde{E}_{\text{iso}}^{\overline{T}}(\mathbf{u},\mathbf{q})=\sum_{1\le i,j\le N}\overline{T}(|\mathbf{q}_{ij}|)+\frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|_V^2\right\},\qquad(9)$$

$$\text{s.t.}\qquad \mathbf{q}=(\nabla_x\mathbf{u},\nabla_y\mathbf{u}).$$

The augmented Lagrangian functional for the problem (9):

$$\begin{aligned}\mathcal{L}(\mathbf{u},\mathbf{q};\boldsymbol{\lambda})&=\sum_{1\le i,j\le N}\overline{T}(|\mathbf{q}_{ij}|)+\frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|_V^2+(\boldsymbol{\lambda},\mathbf{q}-\nabla\mathbf{u})_Q+\frac{\beta}{2}\|\mathbf{q}-\nabla\mathbf{u}\|_Q^2\\&=R(\mathbf{q})+\frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|_V^2+(\boldsymbol{\lambda},\mathbf{q}-\nabla\mathbf{u})_Q+\frac{\beta}{2}\|\mathbf{q}-\nabla\mathbf{u}\|_Q^2,\end{aligned}$$

- $\nabla\mathbf{u}=(\nabla_x\mathbf{u},\nabla_y\mathbf{u})$;
- $\boldsymbol{\lambda}\in Q$ is the Lagrangian multiplier, $\beta>0$ is a constant;
- $R(\mathbf{q})$ is introduced to simply the notation of the regularization term.

Application in 2D - algorithm for (8)

1: Initialization: $\mathbf{u}^0, \mathbf{q}^0, \boldsymbol{\lambda}^0$;
2: **while** stopping criteria is not satisfied **do**
3: Compute $\mathbf{q}^{k+1}$, $\mathbf{u}^{k+1}$, and update $\boldsymbol{\lambda}^{k+1}$ as follows:

$$\mathbf{q}^{k+1} \in \operatorname*{argmin}_{\mathbf{q}\in Q} \mathcal{L}(\mathbf{u}^k, \mathbf{q}; \boldsymbol{\lambda}^k), \tag{10}$$

$$\mathbf{u}^{k+1} = \operatorname*{argmin}_{\mathbf{u}\in V} \mathcal{L}(\mathbf{u}, \mathbf{q}^{k+1}; \boldsymbol{\lambda}^k), \tag{11}$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \beta(\mathbf{q}^{k+1} - \nabla \mathbf{u}^{k+1}), \tag{12}$$

4: **end while**

Application in 2D
- u−sub problem (11)

The $\mathbf{u}-$sub problem (11) is a quadratic optimization problem, whose optimality condition gives a linear system

$$\alpha \mathbf{A}^{\mathsf{T}}(\mathbf{A}\mathbf{u} - \mathbf{f}) - \nabla^{\mathsf{T}} \boldsymbol{\lambda}^k - \beta \nabla^{\mathsf{T}} \mathbf{q}^{k+1} - \beta \Delta \mathbf{u} = 0.$$

- A is a circulant matrix (e.g. image deblurring problem), use FFT.
- Else (e.g. CT reconstruction), use CG method.

Application in 2D
- q−sub problem (10)

The problem (10) reads

$$\min_{\mathbf{q}\in Q}\left\{\sum_{1\le i,j\le N}\overline{T}(|\mathbf{q}_{ij}|)+(\boldsymbol{\lambda}^k,\mathbf{q}-\nabla\mathbf{u}^k)_Q+\frac{\beta}{2}\|\mathbf{q}-\nabla\mathbf{u}^k\|_Q^2\right\},$$

which, by the monotonicity of ρ over $[0,+\infty)$, is

$$\min_{\mathbf{q}\in Q}\left\{\sum_{1\le i,j\le N}\min\big(\rho(|\mathbf{q}_{ij}|),\rho(\tau)\big)+\frac{\beta}{2}|\mathbf{q}_{i,j}-\mathbf{w}_{i,j}|^2\right\},$$

where $\mathbf{w}=\nabla\mathbf{u}^k-\boldsymbol{\lambda}^k/\beta\in Q$. This problem is separable.

$$\min_{z\in\mathbb{R}^2}\left\{g(z;w)=\min\big(\rho(|z|),\rho(\tau)\big)+\frac{\beta}{2}|z-w|^2\right\},$$

where $|z|=\sqrt{(z^{(1)})^2+(z^{(2)})^2}$; and $w\in\mathbb{R}^2,\tau>0,\beta>0$ are given.

Application in 2D
- q−sub problem (10) (Cont.)

Suppose $z^* = \underset{z\in\mathbb{R}^2}{\operatorname{argmin}}\, g(z; w)$.

- If $w = (0, 0)$, it is clear that $z^* = (0, 0)$.
- For w with $|w| \neq 0$, z^* has the same direction as w: $z^* = \frac{|z^*|}{|w|}w$. Thus to obtain z^*, it is sufficient to calculate $|z^*|$ as the minimizer of the following univariate problem:

$$\min_{s\geq 0}\left\{\chi(s;\tau,\beta,t) = \min\big(\rho(s),\rho(\tau)\big) + \frac{\beta}{2}(s-t)^2\right\}, \qquad (13)$$

where $t = |w|$.

Application in 2D
- q−sub problem (10) (Cont.)

For the convenience of description, we introduce the following two functions

$$\chi_1(s) = \rho(s) + \frac{\beta}{2}(s-t)^2, \tag{14}$$

$$\chi_2(s) = \rho(\tau) + \frac{\beta}{2}(s-t)^2. \tag{15}$$

Application in 2D
- q−sub problem (10) (Cont.)

Proposition

The minimization problem (13) *can be solved by*

$$s^* = \begin{cases} s_1^*, & \chi_1(s_1^*) < \chi_2(s_2^*), \\ \{s_1^*, s_2^*\}, & \chi_1(s_1^*) = \chi_2(s_2^*), \\ s_2^*, & \chi_1(s_1^*) > \chi_2(s_2^*), \end{cases} \tag{16}$$

where

$$s_1^* = \underset{0 \le s \le \tau}{\operatorname{argmin}} \chi_1(s); s_2^* = \underset{s \ge \tau}{\operatorname{argmin}} \chi_2(s) = \max(t, \tau).$$

Application in 2D
- q−sub problem (10) (Cont.)

Proposition

[second order lower bound] If $\rho(\cdot)$ satisfies (AS1)(AS2)(AS3)(AS4) and s^*_{loc} is a local minimizer of $\min_{s\geq 0}\chi_1(s)$, then either $s^*_{loc}=0$ or $s^*_{loc}\geq s_L$.

Proposition

Under the assumptions of Proposition above, we have:

(1) *If $s_L>0, \chi_1'(s_L)\geq 0$ ($s_L=0, \chi_1'(0+)\geq 0$), then $s_1^*=0$ is the unique global minimizer of $\min_{0\leq s\leq\tau}\chi_1(s)$.*

(2) *If $s_L>0, \chi_1'(s_L)<0$ ($s_L=0, \chi_1'(0+)<0$), then the equation $\chi_1'(s)=0$ has a unique root $\bar{s}$ on $[s_L,t]$. Set $\mathcal{X}=\{0,\min(\bar{s},\tau)\}$. The global minimizer of $\min_{0\leq s\leq\tau}\chi_1(s)$ is given by $s_1^*=\arg\min_{s\in\mathcal{X}}\chi_1(s)$.*

Application in 2D - algorithm for solving (13)

Require: t, τ, the second order bound s_L and functions $\chi_1(s), \chi_1'(s+), \chi_2(s)$;
Ensure: s^*;

// Find the global minimizer of $s_1^* = \arg\min_{0 \le s \le \tau} \chi_1(s)$.
if $\chi_1'(s_L+) < 0$ **then**
  Find the root $\bar{s}$ of equation $\chi_1'(s) = 0$ in $[s_L, t]$;
  Set the feasible set $\mathcal{X} = \{0, \min(\bar{s}, \tau)\}$;
  Choose $s_1^* \in \mathcal{X}$ with $s_1^* := \arg\min_{s \in \mathcal{X}} \chi_1(s)$;
else
  Set $s_1^* = 0$;
end if
// Find the global minimizer of $s_2^* = \arg\min_{t \le \tau} \chi_2(s)$.
Set $s_2^* = \max\{\tau, t\}$;
// Find the global minimizer s^*.
Choose s^* with

$$s^* = \begin{cases} s_1^*, & \chi_1(s_1^*) < \chi_2(s_2^*), \\ \{s_1^*, s_2^*\}, & \chi_1(s_1^*) = \chi_2(s_2^*), \\ s_2^*, & \text{otherwise}. \end{cases}$$

Application in 2D
- convergence analysis

(AS5) $A^\top A$ ($\mathbf{A}^\top \mathbf{A}$) is invertible.

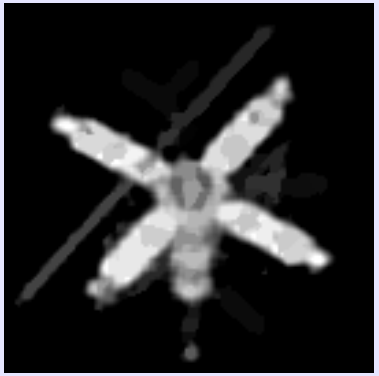
Theorem

Assume that (AS1)(AS2)(AS3)(AS4)(AS5) hold and $\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^k \to 0$ *as* $k \to \infty$ *in the ADMM. Then any cluster point of the sequence* $\{(\mathbf{u}^k, \mathbf{q}^k, \boldsymbol{\lambda}^k)\}$*, if exists, is a KKT point of the constrained optimization problem.*

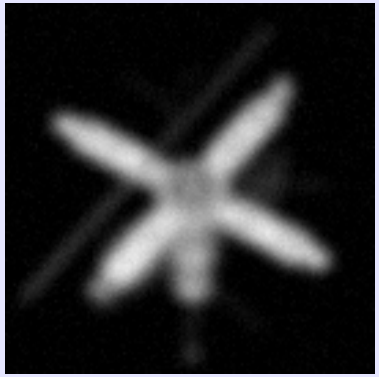
Application in 2D
- experimental results

[Satellite. Size: 135 $\times$ 135]

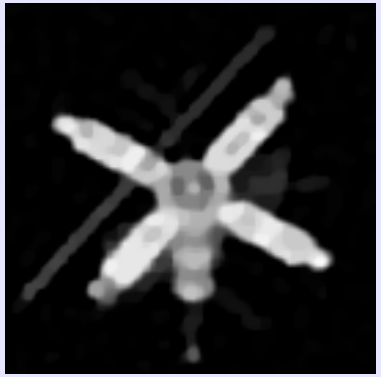
[TV. PSNR: 23.30dB]



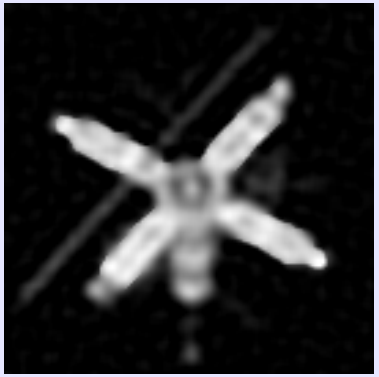
[SCAD. PSNR: 24.11dB]



[Blurry & Noisy.
PSNR: 19.99dB]

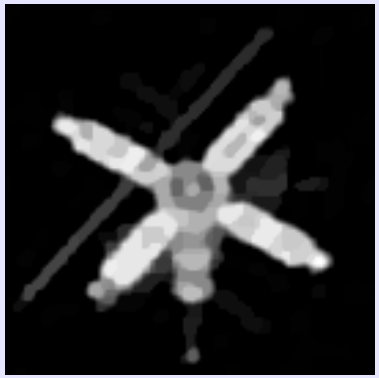


[TR-TV. PSNR: 23.95dB]

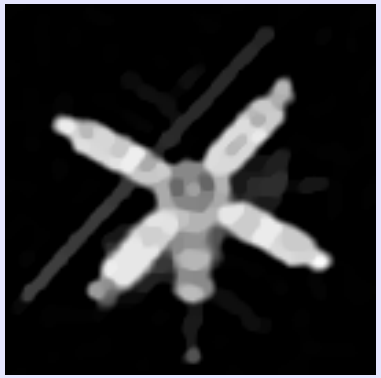


[TR-ℓ_2. PSNR: 23.01dB]

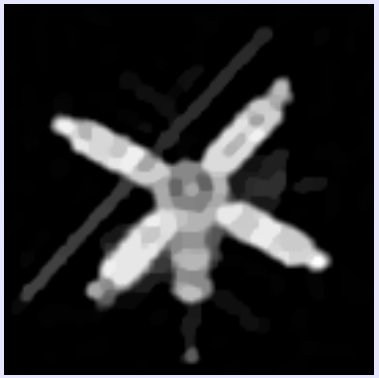
Application in 2D
- experimental results



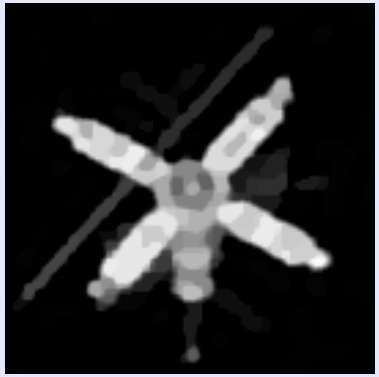
[ℓ_p. PSNR: 23.93dB]



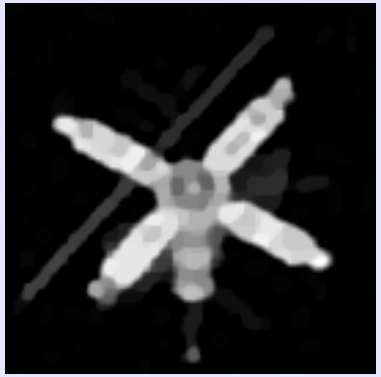
[LN. PSNR: 23.77dB]



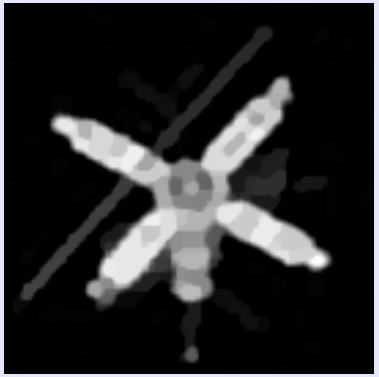
[FRAC. PSNR: 23.78dB]



[TR-ℓ_p. PSNR: 24.04dB]



[TR-LN. PSNR: 24.04dB]



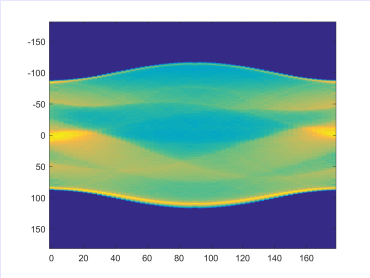
[TR-FRAC. PSNR:23.87dB]

Application in 2D
- experimental results

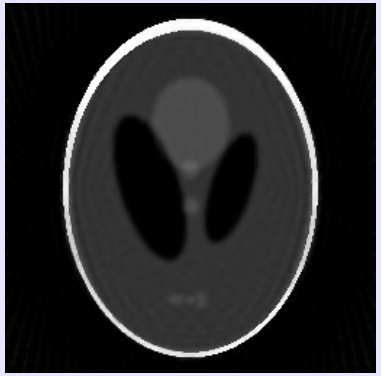
[Shepp-Logan.
Size: 256 $\times$ 256]

[TV. PSNR: 25.71248dB]

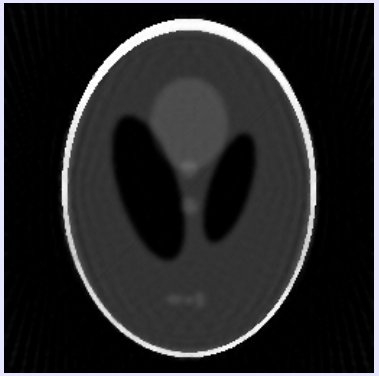
[SCAD. PSNR: 34.5189dB]



[Sinogram]

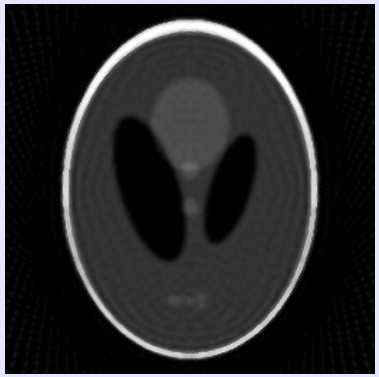


[TR-TV. PSNR: 29.3570dB]

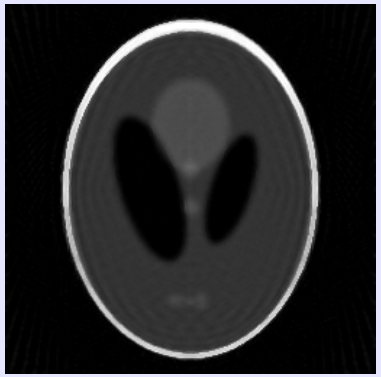


[TR-ℓ_2. PSNR: 29.3570dB]

Application in 2D
- experimental results

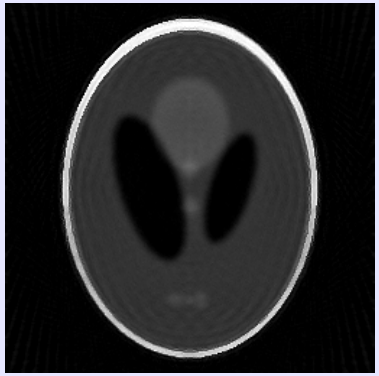


[ℓ_p. PSNR: 24.1377dB]



[LN. PSNR: 25.2226dB]

[FRAC. PSNR: 25.5313dB]

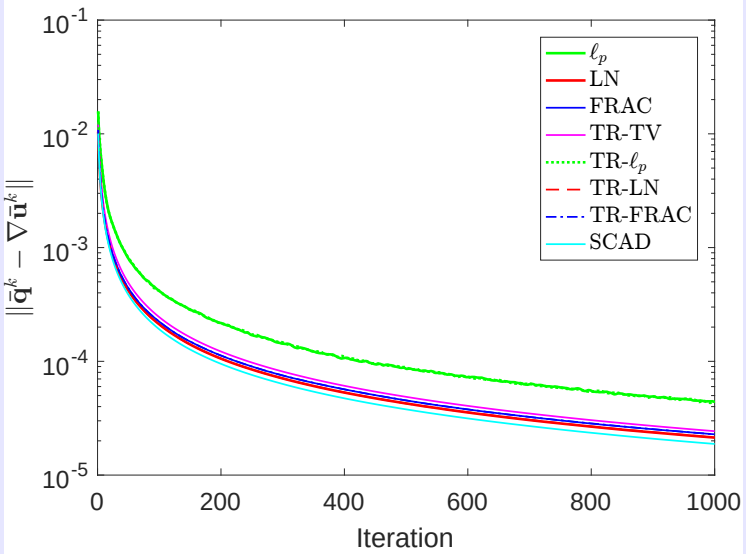


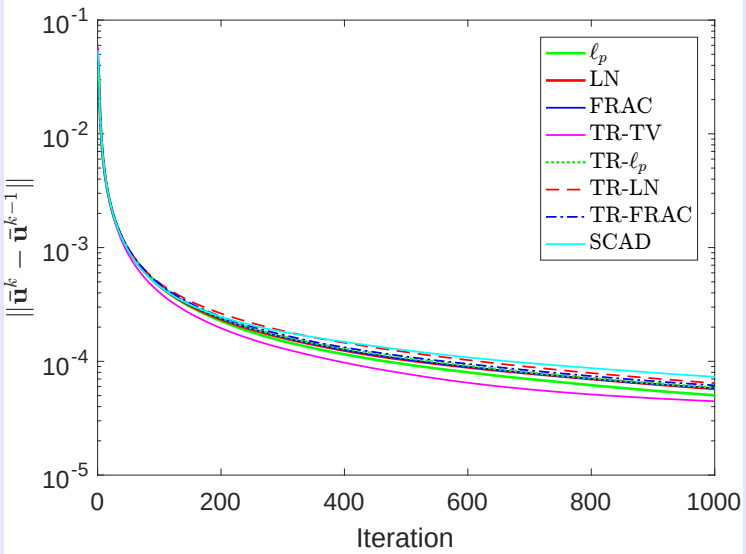
[TR-ℓ_p. PSNR: 25.1182dB]

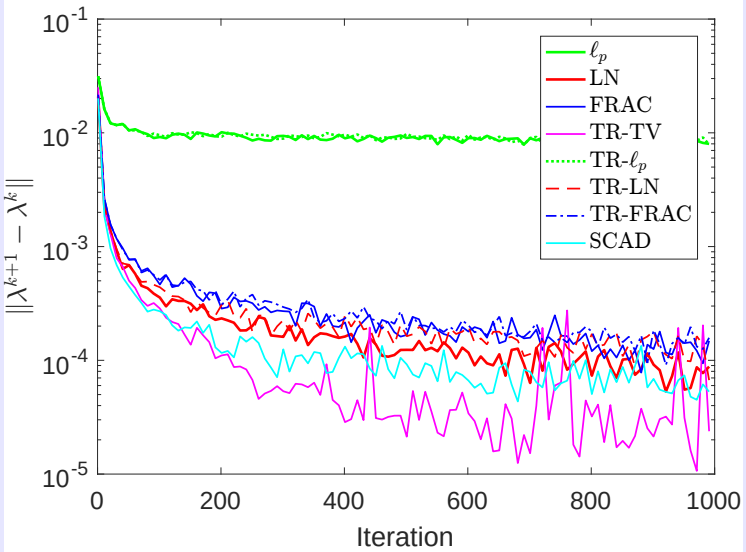
[TR-LN. PSNR: 28.7694dB]

[TR-FRAC. PSNR:29.1334dB]

Application in 2D
- experimental results







Application in 2D
- experimental results

TR-TV		TR-ℓ_p		TR-LN		TR-FRAC		SCAD	
$(\alpha, \beta) =$ (2000, 600)		$(\alpha, \beta, p) =$ (5000, 5000, 0.5)		$(\alpha, \beta, \theta) =$ (13000, 8000, 10)		$(\alpha, \beta, \theta) =$ (8000, 6000, 10)		$(\alpha, \beta) =$ (200, 100)	
τ	PSNR	τ	PSNR	τ	PSNR	τ	PSNR	$\theta(\tau = a\theta)$	PSNR
0.1	23.12	0.1	23.63	0.1	23.57	0.1	23.52	0.01	22.39
0.2	23.58	0.2	24.04	0.2	24.04	0.2	23.87	0.04	23.22
0.3	23.48	0.3	23.98	0.3	23.90	0.3	23.83	0.07	23.89
0.4	23.81	0.4	23.93	0.4	23.90	0.4	23.81	0.10	24.11
0.5	23.99	0.5	23.93	0.5	23.90	0.5	23.81	0.13	24.07
0.6	24.00	0.6	23.93	0.6	23.90	0.6	23.81	0.16	23.93
0.7	24.00	0.7	23.93	0.7	23.90	0.7	23.81	0.19	23.77
0.8	24.00	0.8	23.93	0.8	23.90	0.8	23.81	0.22	23.61
0.9	24.00	0.9	23.93	0.9	23.90	0.9	23.81	0.25	23.49
1.0	24.00	1.0	23.93	1.0	23.90	1.0	23.81	0.27	23.48

Section 3. Lower bound theory

Lower bound theory for anisotropic model

The anisotropic model,

$$\min_{\mathbf{u}\in\mathbb{R}^{n\times n}} \left\{ \begin{aligned} E_{\text{ani}}(\mathbf{u}) &= \sum_{(i,j)\in\mathbf{I}} (\varphi((\nabla_x\mathbf{u})_{i,j}) + \varphi((\nabla_y\mathbf{u})_{i,j})) + \frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|_F^2 \\ &= \sum_{(i,j)\in\mathbf{I}} (\rho(|(\nabla_x\mathbf{u})_{i,j}|) + \rho(|(\nabla_y\mathbf{u})_{i,j}|)) + \frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|_F^2 \end{aligned} \right\}, \quad (17)$$

Theorem (M. Nikolova, Multiscale Model. Simul. 2005)

For any (local) minimizer $\mathbf{u}^*$*, there exists a constant* $\theta > 0$*, such that*

$$\textit{either} \quad |(\nabla_d\mathbf{u}^*)_{i,j}| = 0 \quad \textit{or} \quad |(\nabla_d\mathbf{u}^*)_{i,j}| > \theta \quad \forall (i,j)\in\mathbf{I}, \quad (18)$$

where $d = x, y$.

Lower bound theory for anisotropic model - with box constraints

$$\min_{\mathbf{u}\in\mathbb{R}^{n\times n}} \quad \frac{\lambda}{2}\|A\mathbf{u}-\mathbf{f}\|_F^2 + \sum_{(i,j)\in\mathbf{I}} (\varphi((\nabla_x\mathbf{u})_{i,j}) + \varphi((\nabla_y\mathbf{u})_{i,j}))$$

$$s.t. \quad \underline{b}\mathbf{1} \preceq \mathbf{u} \preceq \overline{b}\mathbf{1},$$

Theorem (X. Chen, M.K. Ng, C. Zhang, IEEE TIP 2012)

For every (local) minimizer $\mathbf{u}^*$ *of the above model, we have*

$$\textit{either} \quad \|(\nabla\mathbf{u}^*)_{i,j}\| = 0 \quad \textit{or} \quad \|(\nabla\mathbf{u}^*)_{i,j}\| \geq \theta \quad \forall (i,j)\in\mathbf{I},$$

where θ *is a constant determined by* $\underline{b}, \overline{b}, \lambda, N, A$ *and* φ.

Lower bound theory for isotropic model

The isotropic model,

$$\min_{\mathbf{u}\in\mathbb{R}^{n\times n}} \left\{ \begin{aligned} E_{\text{iso}}(\mathbf{u}) &= \sum_{(i,j)\in\mathbf{I}} \psi((\nabla_x\mathbf{u})_{i,j}, (\nabla_y\mathbf{u})_{i,j}) + \frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|_F^2 \\ &= \sum_{(i,j)\in\mathbf{I}} \rho(\sqrt{(\nabla_x\mathbf{u})_{i,j}^2 + (\nabla_y\mathbf{u})_{i,j}^2}) + \frac{\alpha}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|_F^2 \end{aligned} \right\}, \tag{19}$$

Theorem (C. Zeng, C. Wu, SIAM J. Numerical Analysis 2018)

For any (local) minimizer $\mathbf{u}^*$,

$$\textit{either} \quad \|(\nabla\mathbf{u}^*)_{i,j}\|_2 = 0 \quad \textit{or} \quad \|(\nabla\mathbf{u}^*)_{i,j}\|_2 \geq \theta \quad \forall (i,j)\in\mathbf{I},$$

where θ *is a constant determined by* λ, N, A *and* ρ.

Lower bound theory for isotropic model - sketch of proof

- Rearrange into a 1D problem.

$$\mathbf{u} \to u, \mathbf{I} \to I, \mathbf{A} \to A, \mathbf{f} \to f, E_{\text{iso}}(\mathbf{u}) \to \mathcal{F}(u).$$

- Reformulation.

  Suppose u^* is a local minimizer of $\mathcal{F}(u)$.

  Let $I_0 = \{i \in I : \|\nabla_i u^*\|_2 = 0\}$, $I_1 = I \backslash I_0$, and
  $K(I_0) = \{u \in \mathbb{R}^N : \|\nabla_i u\|_2 = 0 \ \ \forall i \in I_0\}$.

  Define $\widehat{\mathcal{F}}(u) : \mathbb{R}^N \to \mathbb{R}$ by

$$\widehat{\mathcal{F}}(u) = \frac{\lambda}{2}\|Au - f\|_2^2 + \sum_{i \in I_1} \rho(\|\nabla_i u\|_2).$$

Lower bound theory for isotropic model - sketch of proof (Cont.)

- The second-order necessary conditions.
  $\widehat{\mathcal{F}}(u)$ is $\mathcal{C}^2$ on $B(u^*, \rho)$.

$$u \in K(I_0) \qquad \Rightarrow \qquad \widehat{\mathcal{F}}(u) = \mathcal{F}(u).$$

u^* is a local minimizer of $\widehat{\mathcal{F}}(u)$ over $K(I_0) \cap B(u^*, \rho)$. $\forall v \in K(I_0)$.

$$\begin{aligned} 0 \leq v^T D^2 \widehat{\mathcal{F}}(u^*) v =& \lambda \|Av\|_2^2 + \sum_{i \in I_1} \frac{\rho''(\|\nabla_i u^*\|_2)}{\|\nabla_i u^*\|_2^2} \langle \nabla_i u^*, \nabla_i v \rangle^2 \\ &+ \sum_{i \in I_1} \frac{\rho'(\|\nabla_i u^*\|_2)}{\|\nabla_i u^*\|_2^3} \left(\|\nabla_i u^*\|_2^2 \|\nabla_i v\|_2^2 - \langle \nabla_i u^*, \nabla_i v \rangle^2 \right). \end{aligned} \tag{20}$$

Lower bound theory for isotropic model - sketch of proof (Cont.)

- Proof by induction.
  Let $\Xi := \{\mu : \mu = \|\nabla_i u^*\|_2, i \in I_1\}$. Then $r = \#\Xi \leqslant \#I_1 < N$:

  $$\Xi = \{\mu_1, \mu_2, \ldots, \mu_r\},$$

  where $\mu_1 > \mu_2 > \cdots > \mu_r$.
  Define $\bar{I}^j = \{i \in I : \|\nabla_i u^*\|_2 = \mu_j\} \subseteq I_1$.
  Let $\#\bar{I}^j = \Upsilon_j$ and $\Upsilon = (\Upsilon_1, \ldots, \Upsilon_r) \in \mathbb{Z}_+^r$, where $\mathbb{Z}_+ := \{t \in \mathbb{Z} : t > 0\}$.

Lower bound theory for isotropic model - sketch of proof (Cont.)

- Establish a lower bound for μ_1.

  We choose $\bar{v}^1 \in \mathbb{R}^N$ satisfies

$$\bar{\mathbf{v}}^1_{1,1} = 0, \nabla_i \bar{v}^1 = \frac{1}{\mu_1} \nabla_i u^* \qquad \forall i \in I. \tag{21}$$

  Then,

$$\|\bar{\mathbf{v}}^1\|_F^2 < \frac{N(N+1)(2N+1)}{6}. \tag{22}$$

  Equation (20) becomes

$$0 \leq \bar{v}^{1^T} D^2 \widehat{\mathcal{F}}(u^*) \bar{v}^1 = \lambda \|A\bar{v}^1\|_2^2 + \sum_{i \in I_1} \rho''(\|\nabla_i u^*\|_2) \frac{\|\nabla_i u^*\|_2^2}{\mu_1^2}.$$

Lower bound theory for isotropic model - sketch of proof (Cont.)

$$0 \leqslant \lambda \|A^T A\|_2 \|\bar{v}^1\|_2^2 + \sum_{i \in \bar{I}^1} \rho''(\|\nabla_i u^*\|_2) \frac{\|\nabla_i u^*\|_2^2}{\mu_1^2}$$

$$[\text{using } \#\bar{I}^1 = \Upsilon_1] \; \lambda \|A^T A\|_2 \Omega + \Upsilon_1 \rho''(\mu_1)$$

$$\Longrightarrow \quad \rho''(\mu_1) > -\frac{\lambda \Omega \|A^T A\|_2}{\Upsilon_1}.$$

Since $\rho''(0+) = -\infty$, the constant given by

$$\theta_1 = \inf \left\{ t > 0 : \rho''(t) = -\frac{\lambda \Omega \|A^T A\|_2}{\Upsilon_1} \right\}$$

is well defined and finite. Then $\mu_1 > \theta_1$.

Lower bound theory for isotropic model - sketch of proof (Cont.)

- Establish a lower bound for μ_s.

  Assume $\mu_j > \theta_j > 0$, where $1 \leqslant j \leqslant s-1$ with $1 < s < r$.

  Let $\bar{I} = \bigcup_{j=1}^{s-1} \bar{I}^j$. We choose $\bar{\mathbf{v}}^s \in \mathbb{R}^{n\times n}$ satisfying

$$(\nabla \bar{\mathbf{v}}^s)_{i_x,i_y} = \left(\frac{\nabla_i^x u^*}{\mu_s}, \frac{\nabla_i^y u^*}{\mu_s}\right) \qquad \forall i \in I \setminus \bar{I}. \tag{23}$$

$$\bar{\mathbf{v}}^s_{\boldsymbol{\kappa}_i} = 0 \text{ for a given index } \boldsymbol{\kappa}_i \in B_1(\underline{\mathbf{I}}^i), \forall i \in \{1,\ldots,h\};$$

$$\bar{\mathbf{v}}^s_{\boldsymbol{\kappa}} = 0, \forall \boldsymbol{\kappa} \in \mathbf{I} \setminus B_1(\mathbf{I} \setminus \bar{\mathbf{I}}). \tag{24}$$

Lower bound theory for isotropic model - sketch of proof (Cont.)

$$\frac{|\nabla_i^x u^*|}{\mu_s} \leqslant 1, \frac{|\nabla_i^y u^*|}{\mu_s} \leqslant 1, \forall i \in I \setminus \bar{I}.$$

$$|\bar{\mathbf{v}}_{\boldsymbol{\kappa}}^s| < N.$$

$$|(\nabla_x \bar{\mathbf{v}}^s)_{\boldsymbol{\kappa}}| < 2N, |(\nabla_y \bar{\mathbf{v}}^s)_{\boldsymbol{\kappa}}| < 2N.$$

$$\|(\nabla \bar{\mathbf{v}}^s)_{\boldsymbol{\kappa}}\|_2^2 = |(\nabla_x \bar{\mathbf{v}}^s)_{\boldsymbol{\kappa}}|^2 + |(\nabla_y \bar{\mathbf{v}}^s)_{\boldsymbol{\kappa}}|^2 < 8N^2. \tag{25}$$

Lower bound theory for isotropic model - sketch of proof (Cont.)

Then,

$$
\begin{aligned}
\bar{v}^{s^T} D^2 \widehat{\mathcal{F}}(u^*) \bar{v}^s =& \lambda \|A\bar{v}^s\|_2^2 + \sum_{i \in I_1 \setminus \bar{I}} \rho''(\|\nabla_i u^*\|_2) \frac{\|\nabla_i u^*\|_2^2}{\mu_s^2} \\
&+ \sum_{i \in \bar{I}} \frac{\rho''(\|\nabla_i u^*\|_2)}{\|\nabla_i u^*\|_2^2} \langle \nabla_i u^*, \nabla_i \bar{v}^s \rangle^2 \\
&+ \sum_{i \in \bar{I}} \frac{\rho'(\|\nabla_i u^*\|_2)}{\|\nabla_i u^*\|_2^3} (\|\nabla_i u^*\|_2^2 \|\nabla_i \bar{v}^s\|_2^2 - \langle \nabla_i u^*, \nabla_i \bar{v}^s \rangle^2) \\
\leqslant& \lambda \|A^T A\|_2 \|\bar{v}^s\|_2^2 + \sum_{i \in I_1 \setminus \bar{I}} \rho''(\|\nabla_i u^*\|_2) \frac{\|\nabla_i u^*\|_2^2}{\mu_s^2} \\
&+ \sum_{i \in \bar{I}} \frac{\rho'(\|\nabla_i u^*\|_2)}{\|\nabla_i u^*\|_2^3} \|\nabla_i u^*\|_2^2 \|\nabla_i \bar{v}^s\|_2^2
\end{aligned}
$$

Lower bound theory for isotropic model - sketch of proof (Cont.)

$$
\begin{aligned}
&< \lambda \|A^T A\|_2 \Omega + \sum_{i \in \bar{I}^s} \rho''(\|\nabla_i u^*\|_2) \frac{\|\nabla_i u^*\|_2^2}{\mu_s^2} \\
&\quad + 8N^2 \sum_{j=1}^{s-1} \sum_{i \in \bar{I}^j} \frac{\rho'(\|\nabla_i u^*\|_2)}{\|\nabla_i u^*\|_2} \\
&= \lambda \|A^T A\|_2 \Omega + \sum_{i \in \bar{I}^s} \rho''(\mu_s) + 8N^2 \sum_{j=1}^{s-1} \Upsilon_j \frac{\rho'(\mu_j)}{\mu_j} \\
&\leqslant \lambda \|A^T A\|_2 \Omega + \Upsilon_s \rho''(\mu_s) + 8N^2 \sum_{j=1}^{s-1} \Upsilon_j \frac{\rho'(\theta_j)}{\theta_j}.
\end{aligned}
$$

Lower bound theory for isotropic model - sketch of proof (Cont.)

We have

$$\rho''(\mu_s) > -\frac{\lambda\Omega\|A^TA\|_2}{\Upsilon_s} - 8N^2\sum_{j=1}^{s-1}\frac{\Upsilon_j}{\Upsilon_s}\frac{\rho'(\theta_j)}{\theta_j}. \tag{26}$$

Define

$$\theta_s = \inf\left\{t > 0 : \rho''(t) = -\frac{\lambda\Omega\|A^TA\|_2}{\Upsilon_s} - 8N^2\sum_{j=1}^{s-1}\frac{\Upsilon_j}{\Upsilon_s}\frac{\rho'(\theta_j)}{\theta_j}\right\}.$$

Then $\mu_s > \theta_s$.

Lower bound theory for isotropic model - with box constraints

$$\min_{\mathbf{u}\in\mathbb{R}^{n\times n}} \quad \frac{\lambda}{2}\|\mathcal{A}\mathbf{u}-\mathbf{f}\|_F^2 + \sum_{(i,j)\in\mathbf{I}} \rho(\|(\nabla\mathbf{u})_{i,j}\|)$$
$$s.t. \quad \underline{b}\mathbf{1} \preceq \mathbf{u} \preceq \overline{b}\mathbf{1},$$

Theorem (C. Zeng, C. Wu, Adv. Comput. Math. 2018)

For every (local) minimizer $\mathbf{u}^*$ *of the above model, we have*

$$\textit{either} \quad \|(\nabla\mathbf{u}^*)_{i,j}\| = 0 \quad \textit{or} \quad \|(\nabla\mathbf{u}^*)_{i,j}\| \geq \theta \quad \forall (i,j)\in\mathbf{I},$$

where θ *is a constant determined by* $\underline{b}, \overline{b}, \lambda, N, A$ *and* ρ.

On ℓ_0 gradient regularized model with box constraints - lower bound theory

$$\min_{\mathbf{u}\in\mathbb{R}^{n\times n}} \quad \frac{\lambda}{2}\|\mathcal{A}\mathbf{u}-\mathbf{f}\|_F^2 + \sum_{(i,j)\in\mathbf{I}} \rho_5(\|(\nabla\mathbf{u})_{i,j}\|)$$
$$s.t. \quad \underline{b}\mathbf{1} \preceq \mathbf{u} \preceq \overline{b}\mathbf{1},$$

Theorem (X. Feng, C. Wu, C. Zeng, Inverse Problems 2018)

Suppose that $\mathbf{u}^*$ *is a global minimizer of* ℓ_0 *gradient regularized model with box constraints. Then,*

$$\textit{either} \quad \|(\nabla\mathbf{u}^*)_{i,j}\| = 0 \quad \textit{or} \quad \|(\nabla\mathbf{u}^*)_{i,j}\| \geq \theta \quad \forall (i,j)\in\mathbf{I},$$

where $\theta = \min\{\frac{\sqrt{5}-1}{2\sqrt{\lambda N}\|A\|_2}, \frac{\sqrt{2}|\underline{b}-\overline{b}|}{2}\}$.

On ℓ_0 gradient regularized model with box constraints - uniqueness of global minimizer

Theorem (X. Feng, C. Wu, C. Zeng, Inverse Problems 2018)

Suppose that A (the matrix form of $\mathcal{A}$) has full column rank. Then, there exists a subset $Z \subset \mathbb{R}^{n\times n}$, whose Lebesgue measure is zero, such that if $\mathbf{f} \in \mathbb{R}^{n\times n}\backslash Z$, any two local minimizers have different energy values. Consequently, the variational problem has a unique global solution.

On ℓ_0 gradient regularized model with box constraints - piecewise constant dependency on λ

Theorem (X. Feng, C. Wu, C. Zeng, Inverse Problems 2018)

The set of the global minimizers $\mathcal{U}_g(\lambda)$ has piecewise constant dependency on the parameter λ. Meanwhile, the corresponding optimal objective value has piecewise linear dependency on λ.

Section 4. Conclusions

Conclusions

- Any convex regularization, is impossible to recover the ground truth.
- Presented a general truncation regularization framework.
- Optimization in 2D with implementation and convergence.
- Experiments numerically showed advantages of our method.
- Give lower bound theory for isotropic models.

Section 5. References

Rudin L I, Osher S, and Fatemi E. Nonlinear total variation based noise removal algorithms. Phys D, 1992, 60: 259–268.

Nikolova M. Analysis of the recovery of edges in images and signals by minimizing nonconvex regularized least-squares. Multiscale Model Simul, 2005, 4: 960–991.

Wang Y, Yang J, Yin W, et al. A new alternating minimization algorithm for total variation image reconstruction. SIAM J Imaging Sci, 2008, 1: 248–272.

Wu C and Tai X-C. Augmented lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and high order models. SIAM J Imaging Sci, 2010, 3: 300–339.

Xu Z, Chang X, Xu F, et al. $L_{1/2}$ regularization: A thresholding representation theory and a fast solver. IEEE Trans Neural Netw Learn Syst, 2012, 23: 1013–1027.

Chen X., Ng M. K., and Zhang C., Non-Lipschitz-regularization and box constrained model for image restoration, IEEE Trans. Image Process., 21 (2012), pp. 4709C4721.

Hintermüller M and Wu T. Nonconvex TV^q-models in image restoration: Analysis and a trust-region regularization-based superlinearly convergent solver. SIAM J Imaging Sci, 2013, 6: 1385–1415.

Li G and Pong T K. Global convergence of splitting methods for nonconvex composite optimization. SIAM J Optim, 2015, 25: 2434–2460.

Wang Y, Yin W, and Zeng J. Global convergence of ADMM in nonconvex nonsmooth optimization. arXiv preprint arXiv:1511.06324v4, 2016.

Wu C., Liu Z., and Wen S.. A general truncated regularization framework for contrast-preserving variational signal and image restoration: Motivation and implementation. Sci. China. Math., 2018.

Zeng C., and Wu C.. On the edge recovery property of noncovex nonsmooth regularization in image restoration. SIAM Journal on Numerical Analysis, 2018, 56, 1168-1182.

Feng X., Wu C., and Zeng C.. On the local and global minimizers of ℓ_0 gradient regularized model with box constraints for image restoration. Inverse Problems, 2018.

Thank you for your attention.