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 $u \in \mathbb{R}^N$:

$$u \rightarrow Au \xrightarrow{n} f,$$

- $A$ 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

\[
\begin{align*}
\min_{u \in \mathbb{R}^N} & \quad E(u) \\
& = \sum_{1 \leq i \leq N} \varphi((\nabla_x u)_i) + \frac{\alpha}{2} \|Au - f\|_{\mathbb{R}^K}^2 \\
& = \sum_{1 \leq i \leq N} \rho(\|\nabla_x u\|_i) + \frac{\alpha}{2} \|Au - f\|_{\mathbb{R}^K}^2
\end{align*}
\]

(1)

- \(\rho(\cdot)\) is a potential function.
- \(\nabla_x\) is the forward difference operator with a specific boundary condition, e.g., the periodic or Neumann boundary condition.
2D image $f \in \mathbb{R}^{K \times K}$ is a degradation of $u \in \mathbb{R}^{N \times N}$:

$$u \rightarrow Au \xrightarrow{n} f,$$

For image restoration problem, $A$ is a linear operator such as a blur convolution.

For image reconstruction problem, $A$ is a radon transform.

$n = \{n_{i,j}, 1 \leq i, j \leq K\}$ is random noise.

$\{n_{i,j} \sim \mathcal{N}(0, \sigma^2), 1 \leq i, j \leq K\}$ are i.i.d. random variables, such as additive Gaussian noise and multiplicative noise.
Variational signal and image restoration
- minimization models

\[
\begin{align*}
\min_{u \in \mathbb{R}^{N \times N}} \quad & E_{\text{ani}}(u) = \sum_{1 \leq i,j \leq N} \left( \varphi((\nabla_x u)_{i,j}) + \varphi((\nabla_y u)_{i,j}) \right) + \frac{\alpha}{2} \|Au - f\|_{\mathbb{R}^{K \times K}}^2 \\
& = \sum_{1 \leq i,j \leq N} \left( \rho((|\nabla_x u|)_{i,j}) + \rho((|\nabla_y u|)_{i,j}) \right) + \frac{\alpha}{2} \|Au - f\|_{\mathbb{R}^{K \times K}}^2 \quad (2)
\end{align*}
\]

\[
\begin{align*}
\min_{u \in \mathbb{R}^{N \times N}} \quad & E_{\text{iso}}(u) = \sum_{1 \leq i,j \leq N} \psi((\nabla_x u)_{i,j}, (\nabla_y u)_{i,j}) + \frac{\alpha}{2} \|Au - f\|_{\mathbb{R}^{K \times K}}^2 \\
& = \sum_{1 \leq i,j \leq N} \rho(\sqrt{(\nabla_x u)_{i,j}^2 + (\nabla_y u)_{i,j}^2}) + \frac{\alpha}{2} \|Au - f\|_{\mathbb{R}^{K \times K}}^2 \quad (3)
\end{align*}
\]

- \(\nabla_x\) and \(\nabla_y\) 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.


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 $c \in \mathbb{R}$. 
A new regularizer function

\[ \overline{T}(\cdot) = \rho_\tau(\cdot) = \rho(\min(\cdot, \tau)), \]  

(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

**∅ ≠ Ω ⊆ J = \{1, \cdots , N\}.

1_Ω be its indicator function and ζ > 0 be a real number.

J_0 = \{i : (\nabla_x 1_Ω)_i = 0\}.

J_1 = \{i : (\nabla_x 1_Ω)_i \neq 0\} = J \setminus J_0.

Consider now the minimization problem (1) using truncated regularization where f = A(ζ1_Ω) ∈ \mathbb{R^K}. Denote

\[
E^\zeta(u) = \sum_{1 \leq i \leq N} \overline{T}(\| \nabla_x u_i \|) + \frac{\alpha}{2} \| A(u - \zeta 1_Ω) \|^2_{\mathbb{R^K}}.
\]
The following theorem shows the perfect recovery (i.e., contrast preservation) of the signal $\zeta 1_\Omega$ by (1) with a truncated regularization.


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

Chunlin Wu (Nankai U.)
Nonconvex Regularization
SJTU, Dec., 6, 2018
Application in 2D
- truncated regularization in 2D image

- Anisotropic 2D truncated regularization model:

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

(7)

- Isotropic 2D truncated regularization model:

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

(8)

which are truncated versions of (2) and (3), respectively.
Constrained optimization problem:

\[
\min_{(u, q) \in V \times Q} \left\{ \tilde{E}_{\text{iso}}(u, q) = \sum_{1 \leq i, j \leq N} \bar{T}(|q_{ij}|) + \frac{\alpha}{2} \|Au - f\|_V^2 \right\},
\]

\[
\text{s.t.} \quad q = (\nabla_x u, \nabla_y u).
\]

The augmented Lagrangian functional for the problem (9):

\[
\mathcal{L}(u, q; \lambda) = \sum_{1 \leq i, j \leq N} \bar{T}(|q_{ij}|) + \frac{\alpha}{2} \|Au - f\|_V^2 + (\lambda, q - \nabla u)_Q + \frac{\beta}{2} \|q - \nabla u\|_Q^2
\]

\[
= R(q) + \frac{\alpha}{2} \|Au - f\|_V^2 + (\lambda, q - \nabla u)_Q + \frac{\beta}{2} \|q - \nabla u\|_Q^2,
\]

- \(\nabla u = (\nabla_x u, \nabla_y u)\);
- \(\lambda \in Q\) is the Lagrangian multiplier, \(\beta > 0\) is a constant;
- \(R(q)\) is introduced to simply the notation of the regularization term.
Application in 2D
- algorithm for (8)

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

\[ q^{k+1} \in \arg\min_{q \in Q} \mathcal{L}(u^k, q; \lambda^k), \quad (10) \]
\[ u^{k+1} = \arg\min_{u \in V} \mathcal{L}(u, q^{k+1}; \lambda^k), \quad (11) \]
\[ \lambda^{k+1} = \lambda^k + \beta (q^{k+1} - \nabla u^{k+1}), \quad (12) \]
4: \textbf{end while}
The $u$—sub problem (11) is a quadratic optimization problem, whose optimality condition gives a linear system

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

- $A$ is a circulant matrix (e.g. image deblurring problem), use FFT.
- Else (e.g. CT reconstruction), use CG method.
The problem (10) reads
\[
\min_{q \in Q} \left\{ \sum_{1 \leq i, j \leq N} \overline{T}(|q_{ij}|) + (\lambda^k, q - \nabla u^k)_Q + \frac{\beta}{2} \|q - \nabla u^k\|_Q^2 \right\},
\]
which, by the monotonicity of \(\rho\) over \([0, +\infty)\), is
\[
\min_{q \in Q} \left\{ \sum_{1 \leq i, j \leq N} \min (\rho(|q_{ij}|), \rho(\tau)) + \frac{\beta}{2} |q_{i,j} - w_{i,j}|^2 \right\},
\]
where \(w = \nabla u^k - \lambda^k / \beta \in Q\). This problem is separable.
\[
\min_{z \in \mathbb{R}^2} \left\{ g(z; w) = \min (\rho(|z|), \rho(\tau)) + \frac{\beta}{2} |z - w|^2 \right\},
\]
where \(|z| = \sqrt{(z^{(1)})^2 + (z^{(2)})^2} \); and \(w \in \mathbb{R}^2, \tau > 0, \beta > 0\) are given.
Suppose $z^* = \arg\min_{z \in \mathbb{R}^2} g(z; w)$.

- 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:

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

where $t = |w|$. 
For the convenience of description, we introduce the following two functions

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

\[ \chi_2(s) = \rho(\tau) + \frac{\beta}{2} (s - t)^2. \]  
(15)
Application in 2D
- q—sub problem (10) (Cont.)

Proposition

The minimization problem (13) can be solved by

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

(16)

\[ s_1^* = \arg\min_{0 \leq s \leq \tau} \chi_1(s); \quad s_2^* = \arg\min_{s \geq \tau} \chi_2(s) = \max(t, \tau). \]
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 $[s_L, \tau]$. 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: \( t, \tau, \) the second order bound \( s_L \) and functions \( \chi_1(s), \chi_1'(s), \chi_2(s); \)
Ensure: \( s^*; \)

1: // Find the global minimizer of \( s_1^* = \arg \min_{0 \leq s \leq \tau} \chi_1(s). \)
2: if \( \chi_1'(s_L+) < 0 \) then
3: Find the root \( \bar{s} \) of equation \( \chi_1'(s) = 0 \) in \( [s_L, t]; \)
4: Set the feasible set \( \mathcal{X} = \{0, \min(\bar{s}, \tau)\}; \)
5: Choose \( s_1^* \in \mathcal{X} \) with \( s_1^* := \arg \min_{s \in \mathcal{X}} \chi_1(s); \)
6: else
7: Set \( s_1^* = 0; \)
8: end if
9: // Find the global minimizer of \( s_2^* = \arg \min_{t \leq \tau} \chi_2(s). \)
10: Set \( s_2^* = \max\{\tau, t\}; \)
11: // Find the global minimizer \( s^* \).
12: Choose \( s^* \) with

\[
\begin{align*}
    s^* &= \left\{ \begin{array}{ll}
    s_1^*, & \chi_1(s_1^*) < \chi_2(s_2^*), \\
    \{s_1^*, s_2^*\}, & \chi_1(s_1^*) = \chi_2(s_2^*), \\
    s_2^*, & \text{otherwise.}
    \end{array} \right.
\end{align*}
\]
Application in 2D
- convergence analysis

\((\text{AS5})\) \(A^T A (A^T A)\) is invertible.

**Theorem**

Assume that \((\text{AS1})(\text{AS2})(\text{AS3})(\text{AS4})(\text{AS5})\) hold and \(\lambda^{k+1} - \lambda^k \to 0\) as \(k \to \infty\) in the ADMM. Then any cluster point of the sequence \(\{(u^k, q^k, \lambda^k)\}\), if exists, is a KKT point of the constrained optimization problem.
Application in 2D
- experimental results

- Satellite: Size 135 × 135
  - TV: PSNR 23.30dB
  - SCAD: PSNR 24.11dB

- Blurry & Noisy:
  - PSNR 19.99dB
  - TR-TV: PSNR 23.95dB
  - TR-ℓ₂: PSNR 23.01dB

Chunlin Wu (Nankai U.)
Nonconvex Regularization
SJTU, Dec., 6, 2018
Application in 2D
- experimental results

\[ \ell_p \text{ PSNR: 23.93dB} \]
\[ \text{LN. PSNR: 23.77dB} \]
\[ \text{FRAC. PSNR: 23.78dB} \]

\[ \text{TR-} \ell_p \text{ PSNR: 24.04dB} \]
\[ \text{TR-LN. PSNR: 24.04dB} \]
\[ \text{TR-FRAC. PSNR: 23.87dB} \]
Application in 2D - experimental results

- Shepp-Logan. Size: 256 × 256
  - TV. PSNR: 25.71248dB
  - SCAD. PSNR: 34.5189dB

- Sinogram
  - TR-TV. PSNR: 29.3570dB
  - TR-ℓ2. PSNR: 29.3570dB
Application in 2D
- experimental results

- $\ell_p$. PSNR: 24.1377dB
- LN. PSNR: 25.2226dB
- FRAC. PSNR: 25.5313dB

- TR-$\ell_p$. PSNR: 25.1182dB
- TR-LN. PSNR: 28.7694dB
- TR-FRAC. PSNR: 29.1334dB
Application in 2D
- experimental results

Chunlin Wu (Nankai U.)
Nonconvex Regularization
SJTU, Dec., 6, 2018
Application in 2D
- experimental results

<table>
<thead>
<tr>
<th>TR-TV</th>
<th>TR-ℓₚ</th>
<th>TR-LN</th>
<th>TR-FRAC</th>
<th>SCAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>(α, β) = (2000, 600)</td>
<td>(α, β, p) = (5000, 5000, 0.5)</td>
<td>(α, β, θ) = (13000, 8000, 10)</td>
<td>(α, β, θ) = (8000, 6000, 10)</td>
<td>(α, β) = (200, 100)</td>
</tr>
<tr>
<td>τ</td>
<td>PSNR</td>
<td>τ</td>
<td>PSNR</td>
<td>τ</td>
</tr>
<tr>
<td>---</td>
<td>------</td>
<td>---</td>
<td>------</td>
<td>---</td>
</tr>
<tr>
<td>0.1</td>
<td>23.12</td>
<td>0.1</td>
<td>23.63</td>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
<td>23.58</td>
<td>0.2</td>
<td>24.04</td>
<td>0.2</td>
</tr>
<tr>
<td>0.3</td>
<td>23.48</td>
<td>0.3</td>
<td>23.98</td>
<td>0.3</td>
</tr>
<tr>
<td>0.4</td>
<td>23.81</td>
<td>0.4</td>
<td>23.93</td>
<td>0.4</td>
</tr>
<tr>
<td>0.5</td>
<td>23.99</td>
<td>0.5</td>
<td>23.93</td>
<td>0.5</td>
</tr>
<tr>
<td>0.6</td>
<td>24.00</td>
<td>0.6</td>
<td>23.93</td>
<td>0.6</td>
</tr>
<tr>
<td>0.7</td>
<td>24.00</td>
<td>0.7</td>
<td>23.93</td>
<td>0.7</td>
</tr>
<tr>
<td>0.8</td>
<td>24.00</td>
<td>0.8</td>
<td>23.93</td>
<td>0.8</td>
</tr>
<tr>
<td>0.9</td>
<td>24.00</td>
<td>0.9</td>
<td>23.93</td>
<td>0.9</td>
</tr>
<tr>
<td>1.0</td>
<td>24.00</td>
<td>1.0</td>
<td>23.93</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Section 3. Lower bound theory
The anisotropic model,

\[
E_{\text{ani}}(u) = \sum_{(i,j) \in I} (\varphi((\nabla_x u)_{i,j}) + \varphi((\nabla_y u)_{i,j})) + \frac{\alpha}{2} \|A u - f\|_F^2
\]

\[
= \sum_{(i,j) \in I} (\rho(\|\nabla_x u_{i,j}\|) + \rho(\|\nabla_y u_{i,j}\|)) + \frac{\alpha}{2} \|A u - f\|_F^2
\]

(17)

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

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

\[
\text{either } |(\nabla_d u^*)_{i,j}| = 0 \quad \text{or} \quad |(\nabla_d u^*)_{i,j}| > \theta \quad \forall (i,j) \in I,
\]

(18)

*where \(d = x, y\).*
Lower bound theory for anisotropic model - with box constraints

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

s.t. $\underline{b}1 \preceq u \preceq \bar{b}1$,

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

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

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

where $\theta$ is a constant determined by $b, \bar{b}, \lambda, N, A$ and $\varphi$. 
Lower bound theory for isotropic model

The isotropic model,

\[
\begin{align*}
\min_{u \in \mathbb{R}^{n \times n}} & \quad \left\{ E_{\text{iso}}(u) = \sum_{(i,j) \in I} \psi((\nabla_x u)_{i,j}, (\nabla_y u)_{i,j}) + \frac{\alpha}{2} \|Au - f\|_F^2 \right. \\
& \quad \left. \quad = \sum_{(i,j) \in I} \rho(\sqrt{(\nabla_x u)_{i,j}^2 + (\nabla_y u)_{i,j}^2}) + \frac{\alpha}{2} \|Au - f\|_F^2 \right\},
\end{align*}
\]

(19)

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

For any (local) minimizer \( u^* \),

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

where \( \theta \) is a constant determined by \( \lambda, N, A \) and \( \rho \).
Lower bound theory for isotropic model  
- sketch of proof

- Rearrange into a 1D problem.

\[ u \rightarrow u, \ I \rightarrow I, \ A \rightarrow A, \ f \rightarrow f, \ E_{\text{iso}}(u) \rightarrow \mathcal{F}(u). \]

- Reformulation.
  Suppose \( u^* \) is a local minimizer of \( \mathcal{F}(u) \).
  Let \( I_0 = \{ i \in I : \| \nabla_i u^* \|_2 = 0 \} \), \( I_1 = I \setminus I_0 \), and
  \( K(I_0) = \{ u \in \mathbb{R}^N : \| \nabla_i u \|_2 = 0 \ \forall i \in I_0 \} \).
  Define \( \hat{\mathcal{F}}(u) : \mathbb{R}^N \rightarrow \mathbb{R} \) by

  \[
  \hat{\mathcal{F}}(u) = \frac{\lambda}{2} \| Au - f \|_2^2 + \sum_{i \in I_1} \rho(\| \nabla_i u \|_2).
  \]
The second-order necessary conditions.
\( \hat{F}(u) \) is \( C^2 \) on \( B(u^*, \rho) \).

\[ u \in K(I_0) \quad \Rightarrow \quad \hat{F}(u) = F(u). \]

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

\[
0 \leq v^T D^2 \hat{F}(u^*) v = \lambda \|Av\|_2^2 + \sum_{i \in I_1} \rho''(\|\nabla_i u^*\|_2) \frac{\|\nabla_i u^*\|_2^2}{\|\nabla_i u^*\|_2^2} \langle \nabla_i u^*, \nabla_i v \rangle^2 \\
+ \sum_{i \in I_1} \rho'(\|\nabla_i u^*\|_2) \frac{\|\nabla_i u^*\|_2^3}{\|\nabla_i u^*\|_2^3} (\|\nabla_i u^*\|_2^2 \|\nabla_i v\|_2^2 - \langle \nabla_i u^*, \nabla_i v \rangle^2). \tag{20}
\]
Proof by induction.
Let $\Xi := \{ \mu : \mu = \| \nabla_i u^* \|_2, i \in I_1 \}$. Then $r = \# \Xi \leq \# I_1 < N$:

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

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 $\# \bar{I}_j = \gamma_j$ and $\gamma = (\gamma_1, \ldots, \gamma_r) \in \mathbb{Z}^r_+$, where $\mathbb{Z}_+ := \{ t \in \mathbb{Z} : t > 0 \}$. 
Establish a lower bound for $\mu_1$.

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

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

Then,

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

Equation (20) becomes

$$
0 \leq \bar{v}^1^T D^2 \hat{F}(u^*) \bar{v}^1 = \lambda \|A\bar{v}^1\|_2^2 + \sum_{i \in I_1} \rho'' \left(\|\nabla_i u^*\|_2\right) \frac{\|\nabla_i u^*\|_2^2}{\mu_1^2}.
$$
Lower bound theory for isotropic model
- sketch of proof (Cont.)

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

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

\[ \Rightarrow \rho''(\mu_1) > -\frac{\lambda \Omega \|A^T A\|_2}{\Upsilon_1}. \]

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 $\mu_s$.

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

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

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

(23)

$\overline{v}^s_{\kappa_i} = 0$ for a given index $\kappa_i \in B_1(I^i), \forall i \in \{1, \ldots, h\}$;

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

(24)
Lower bound theory for isotropic model
- sketch of proof (Cont.)

\[
\frac{\nabla^x_i u^*}{\mu_s} \leq 1, \quad \frac{\nabla^y_i u^*}{\mu_s} \leq 1, \quad \forall i \in I \setminus \bar{I}.
\]

\[|\nabla_s^s| < N.\]

\[|\nabla^x_s^s| < 2N, \quad |\nabla^y_s^s| < 2N.\]

\[\|\nabla^s_s\|_2^2 = |\nabla^x_s^s|^2 + |\nabla^y_s^s|^2 < 8N^2. \quad (25)\]
Then,

\[ \tilde{v}^s \mathbf{D}^2 \mathbf{\hat{F}}(u^*) \tilde{v}^s = \lambda \| A \tilde{v}^s \|_2^2 + \sum_{i \in I_1 \setminus \tilde{I}} \rho''(\| \nabla_i u^* \|_2) \frac{\| \nabla_i u^* \|_2^2}{\mu_s^2} \]

\[ + \sum_{i \in \tilde{I}} \rho''(\| \nabla_i u^* \|_2) \frac{\| \nabla_i u^* \|_2^2}{\| \nabla_i u^* \|_2} \langle \nabla_i u^*, \nabla_i \tilde{v}^s \rangle^2 \]

\[ + \sum_{i \in \tilde{I}} \rho'(\| \nabla_i u^* \|_2) \frac{\| \nabla_i u^* \|_2^2 \| \nabla_i \tilde{v}^s \|_2^2 - \langle \nabla_i u^*, \nabla_i \tilde{v}^s \rangle^2}{\| \nabla_i u^* \|_2^2} \]

\[ \leq \lambda \| A^T A \|_2 \| \tilde{v}^s \|_2^2 + \sum_{i \in I_1 \setminus \tilde{I}} \rho''(\| \nabla_i u^* \|_2) \frac{\| \nabla_i u^* \|_2^2}{\mu_s^2} \]

\[ + \sum_{i \in \tilde{I}} \rho'(\| \nabla_i u^* \|_2) \frac{\| \nabla_i u^* \|_2^2}{\| \nabla_i u^* \|_2^2} \| \nabla_i u^* \|_2^2 \| \nabla_i \tilde{v}^s \|_2^2 \]
Lower bound theory for isotropic model - sketch of proof (Cont.)

$$<\lambda \|A^T A\|_2\Omega + \sum_{i \in \mathcal{I}^s} \rho''(\|\nabla_i u^*\|_2) \frac{\|\nabla_i u^*\|^2}{\mu_s^2}$$

$$+ 8N^2 \sum_{j=1}^{s-1} \sum_{i \in \mathcal{I}^j} \rho'(\|\nabla_i u^*\|_2) \frac{\|\nabla_i u^*\|_2}{\|\nabla_i u^*\|_2}$$

$$= \lambda \|A^T A\|_2\Omega + \sum_{i \in \mathcal{I}^s} \rho''(\mu_s) + 8N^2 \sum_{j=1}^{s-1} \Upsilon_j \frac{\rho'(\mu_j)}{\mu_j}$$

$$\leq \lambda \|A^T A\|_2\Omega + \Upsilon_s \rho''(\mu_s) + 8N^2 \sum_{j=1}^{s-1} \Upsilon_j \frac{\rho'(\theta_j)}{\theta_j}.$$
Lower bound theory for isotropic model
- sketch of proof (Cont.)

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

\[
\min_{u \in \mathbb{R}^{n \times n}} \frac{\lambda}{2} \|Au - f\|_F^2 + \sum_{(i,j) \in I} \rho(\|\nabla u_{i,j}\|)
\]

s.t. \( b1 \preceq u \preceq b1 \),


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

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

where \( \theta \) is a constant determined by \( b, \bar{b}, \lambda, N, A \) and \( \rho \).
On $\ell_0$ gradient regularized model with box constraints - lower bound theory

\[
\min_{u \in \mathbb{R}^{n \times n}} \frac{\lambda}{2} \|Au - f\|_F^2 + \sum_{(i,j) \in I} \rho_5(\|\nabla u_{i,j}\|)
\]

s.t. \hspace{1em} b1 \leq u \leq \bar{b}1,

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

Suppose that $u^*$ is a global minimizer of $\ell_0$ gradient regularized model with box constraints. Then,

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

where $\theta = \min\{\frac{\sqrt{5} - 1}{2\sqrt{\lambda N}\|A\|_2}, \frac{\sqrt{2}|b - \bar{b}|}{2}\}.$
On $\ell_0$ 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 $A$) has full column rank. Then, there exists a subset $Z \subset \mathbb{R}^{n \times n}$, whose Lebesgue measure is zero, such that if $f \in \mathbb{R}^{n \times n} \setminus Z$, any two local minimizers have different energy values. Consequently, the variational problem has a unique global solution.
On $\ell_0$ gradient regularized model with box constraints - piecewise constant dependency on $\lambda$

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

The set of the global minimizers $U_g(\lambda)$ has piecewise constant dependency on the parameter $\lambda$. Meanwhile, the corresponding optimal objective value has piecewise linear dependency on $\lambda$. 
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


Thank you for your attention.