

Sparse sensing: Fundamentals and applications in X-ray CT

Hyoung Suk Park

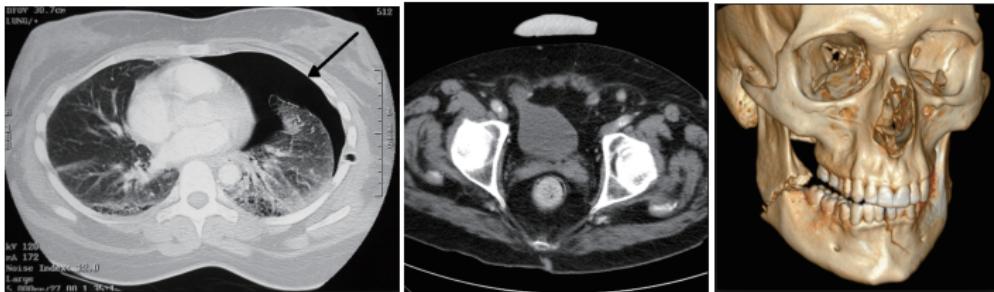
National Institute for Mathematical Sciences

January 9, 2017

2017 Winter School- CT Session

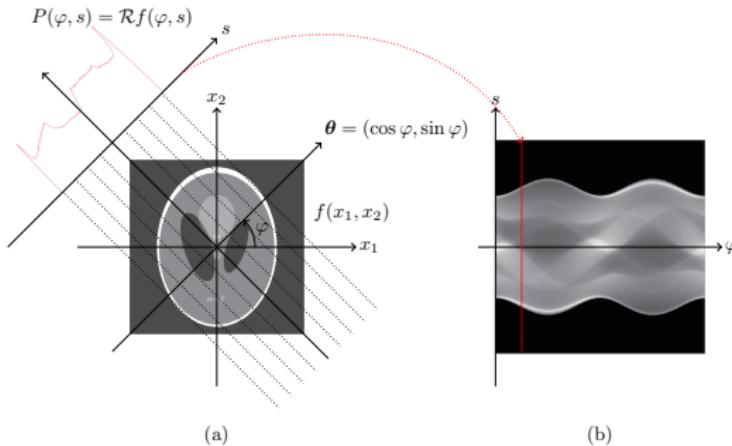
X-ray Computed Tomography

- X-ray CT: widely used imaging modality for screening and diagnosis, emergency medicine, image-guided interventions, and monitoring of therapeutic responses.



- Disadvantage: **excessive radiation dose**, risk of cancer from medical radiation exposures
 - ▶ X-ray: 0.1mSV, head CT: 2, chest CT: 8, abdomen CT: 10, nature: 2.5 ~ 10
 - ▶ Dose: 1000mSV, cancer rates are 2 to 3 times higher
 - ▶ GE, **Chest low dose CT: < 1 mSV**

Basic CT Reconstruction Algorithm



$$Rf(\varphi, s) = \int_{\mathbb{R}^2} f(\mathbf{x}) \delta(\mathbf{x} \cdot \boldsymbol{\theta} - s) d\mathbf{x} \quad : \text{Radon transform}$$

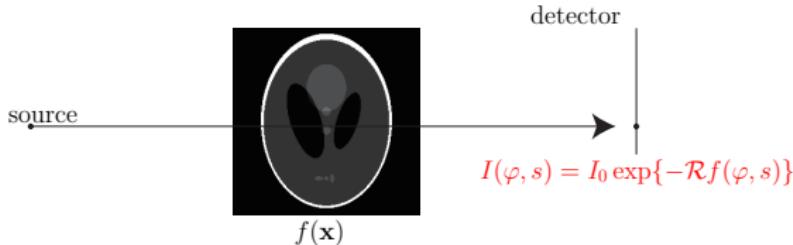
$$f(\mathbf{x}) = \frac{1}{4\pi} \mathcal{R}^* \left(\mathcal{F}^{-1}[|\omega|] * Rf \right) (\mathbf{x}) \quad : \text{Filtered backprojection}$$

$-f(\mathbf{x})$: attenuation coefficient of material at energy E and position \mathbf{x}

$-\mathcal{R}^* g(\mathbf{x}) = \int_0^{2\pi} g(\varphi, \mathbf{x} \cdot \boldsymbol{\theta}) d\varphi$: Backprojection operator

Basic CT reconstruction algorithm

- Lambert-Beer's law:



I_0 : X-ray intensity generated at X-ray source.

$f(\mathbf{x})$: attenuation coefficient at position \mathbf{x} .

- Measured X-ray data:

$$\textcolor{blue}{P}(\varphi, s) = -\ln \frac{I(\varphi, s)}{I_0} + \text{Poisson noise} \quad (\neq \mathcal{R}f(\varphi, s)).$$

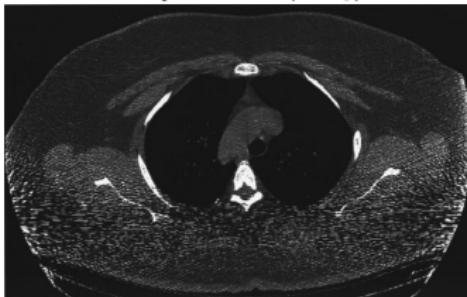
- The FBP yields

$$(\text{CT image}) \quad f_{\text{CT}}(\mathbf{x}) := \frac{1}{4\pi} \mathcal{R}^* \left(\mathcal{F}^{-1}[|\omega|] * \textcolor{blue}{P} \right) (\mathbf{x}).$$

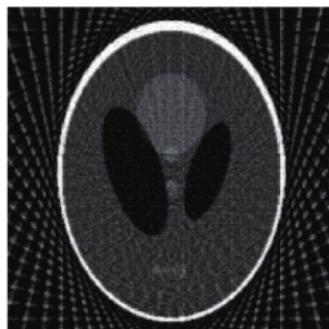
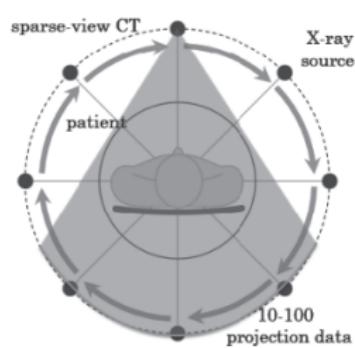
Low dose CT

- How to reduce X-ray dose

- reduce X-ray current (or I_0) → increase noise, photon starvation



- sparse view (under sampling) → causes streaking artifacts (studied as a potential strategy¹)



30 view

¹H. Qi, et al , Sparse-view computed tomography image reconstruction..., Bio-Medical Materials and Engineering, 2015.

Noise Correction Methods

Maximum-likelihood for X-ray CT

- Given measured $\mathbf{y} = \{y_j\}_{j=1}^M$ with $y_j = \text{poisson}\{I_j\}$, find $\mathbf{f} = \{f_j\}_{j=1}^N$ that maximize the probability $\Pr(\mathbf{f}|\mathbf{y})$

$$\hat{\mathbf{f}} = \arg \max_{\mathbf{f}} \Pr(\mathbf{f}|\mathbf{y})$$

- From Bayesian rule, it is equivalent to

$$\hat{\mathbf{f}} = \arg \max_{\mathbf{f}} \{\log \Pr(\mathbf{y}|\mathbf{f}) + \log \Pr(\mathbf{f}) - \log \Pr(\mathbf{y})\}$$

- The problem can be reduced to

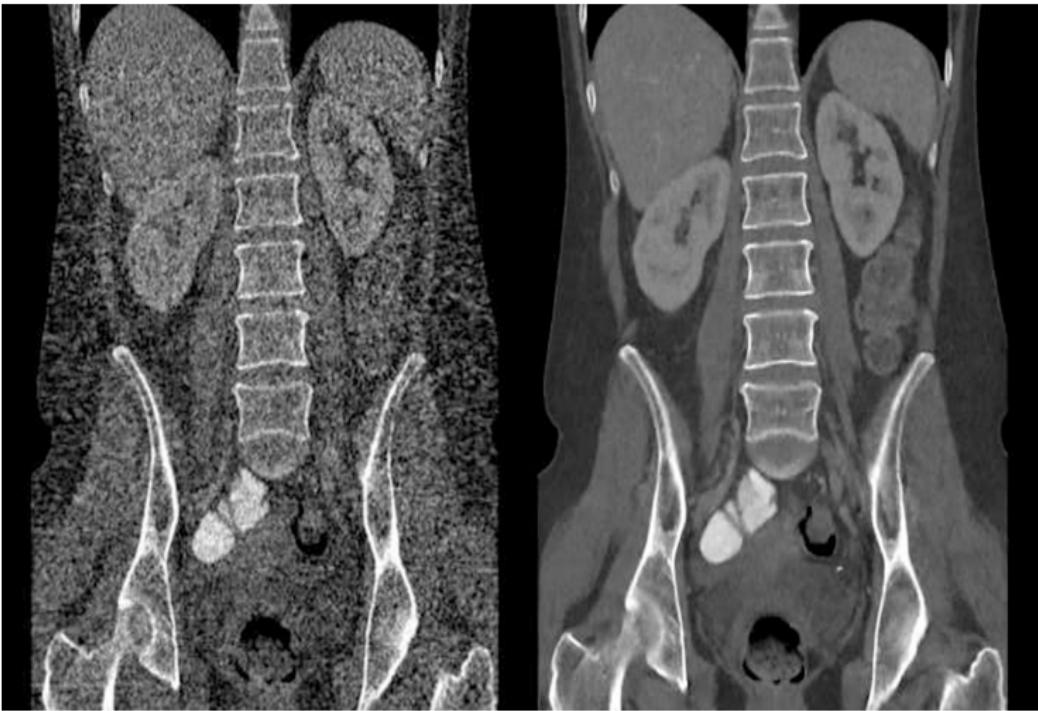
$$\hat{\mathbf{f}} = \arg \max_{\mathbf{f}} \log \Pr(\mathbf{y}|\mathbf{f}) : \text{Maximum likelihood approach}$$

- With $\Pr(y_j|\mathbf{f}) = \frac{I_j^{y_j} e^{-I_j}}{y_j!}$, second order Taylor expansion gives

$$\log \Pr(\mathbf{y}|\mathbf{f}) \approx -\frac{1}{2} (\mathbf{P} - \mathbf{Lf})^T \mathbf{W} (\mathbf{P} - \mathbf{Lf}),$$

- \mathbf{W} : diagonal matrix whose values $w_j = y_j$.
- \mathbf{L} : projection operator.

Maximum-likelihood for X-ray tomography



FBP

MBIR

One-step iterative method

- Noise reduction model:

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} (\mathbf{P} - \mathbf{L}\mathbf{f})^T \mathbf{W} (\mathbf{P} - \mathbf{L}\mathbf{f})$$

- Gradient-decent method leads to

$$\begin{aligned}\mathbf{f}^{(k)} &= \Delta t \left[\sum_{n=0}^{k-1} (\mathbf{I}_d - \Delta t \mathbf{L}^T \mathbf{W} \mathbf{L})^n \right] \mathbf{L}^T \mathbf{W} \mathbf{P} \\ &= \left[\mathbf{I}_d - (\mathbf{I}_d - \Delta t \mathbf{L}^T \mathbf{W} \mathbf{L})^k \right] \mathbf{L}^{-1} \mathbf{P}.\end{aligned}$$

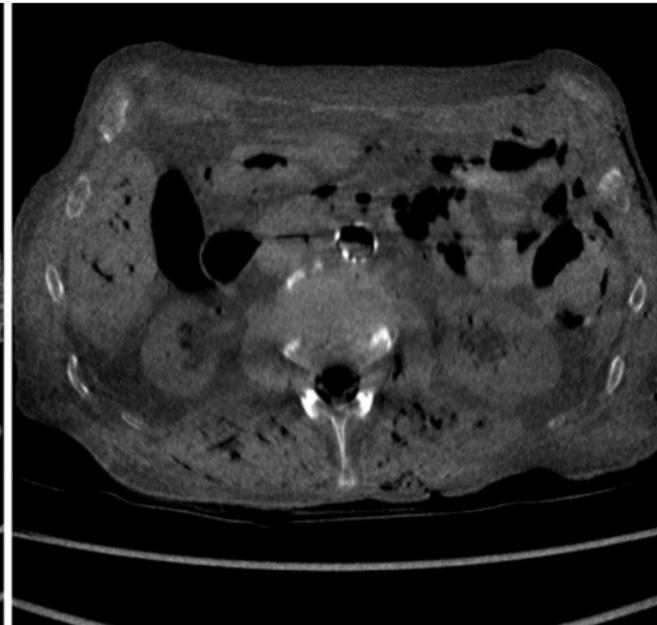
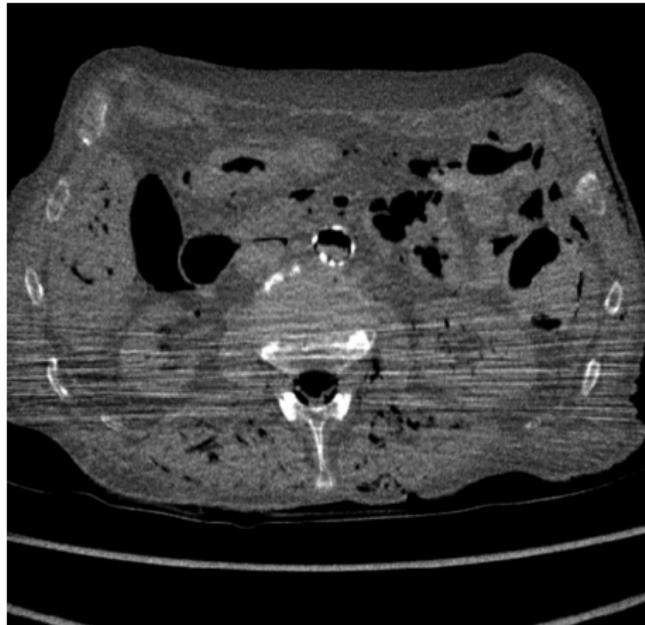
- Using some properties of Radon transform, we obtain

$$f^{(k)}(\mathbf{x}) = \frac{1}{4\pi} \mathcal{R}^* \left(\mathcal{F}^{-1}[\mathbf{H}_k] * P \right) (\mathbf{x}),$$

where H_k is the adaptive filter given by

$$H_k(\omega) = (1 - (1 - \Delta t \frac{\sigma_{ray}}{|\omega|}))^k |\omega|, \quad H_k(0) = 0.$$

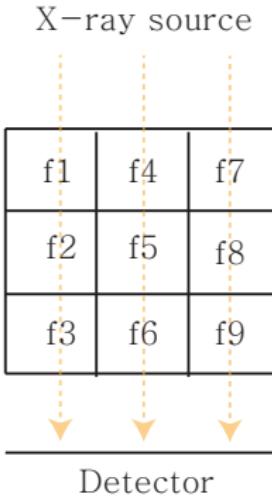
One-step iterative method



Sparse-view CT reconstruction methods

Spare view X-ray CT

- Can we reconstruct the unique solution \mathbf{f} from the under-determined system $\mathbf{L}_s \mathbf{f} = \mathbf{P}_s$?



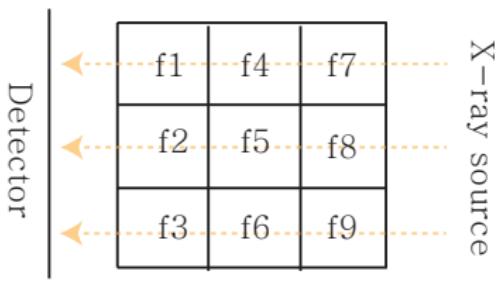
$$\mathbf{L}_s \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$$

The matrix \mathbf{L}_s is defined as:

$$\mathbf{L}_s = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Spare view X-ray CT

- Can we reconstruct the unique solution \mathbf{f} from the under-determined system
 $\mathbf{L}_s \mathbf{f} = \mathbf{P}_s$?



$$\mathbf{L}_s \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{bmatrix}$$

The equation illustrates the system of equations. The matrix \mathbf{L}_s is defined by the following 9x6 matrix:

$$\mathbf{L}_s = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Sparse driven reconstruction

- Let $\|\mathbf{x}\|_0 = \text{number of non-zeros of } \mathbf{x}$. Let $\text{spark}(\mathbf{A}) = \min\{\|\mathbf{x}\|_0 : \mathbf{x} \in N(\mathbf{A}) \setminus \{\mathbf{0}\}\}$.

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = 0$$

$\text{spark}(\mathbf{A})$: smallest number of columns that are linearly dependent.

- If the under-determined system $\mathbf{Ax} = \mathbf{b}$ has the solution \mathbf{x} obeying $\|\mathbf{x}\|_0 < 1/2\text{spark}(\mathbf{A})$, then \mathbf{x} is the unique solution to the following minimization problem.

$$\min \|\mathbf{x}\|_0 \text{ subject to } \mathbf{Ax} = \mathbf{b}.$$

Relaxation of l_0 -minimization problem

- Let $\mu(\mathbf{A})$ be a mutual coherence of an m by n matrix \mathbf{A} given by

$$\mu(\mathbf{A}) = \max_{j \neq k} \frac{\mathbf{a}_j \cdot \mathbf{a}_k}{\|\mathbf{a}_j\|_2 \|\mathbf{a}_k\|_2}.$$

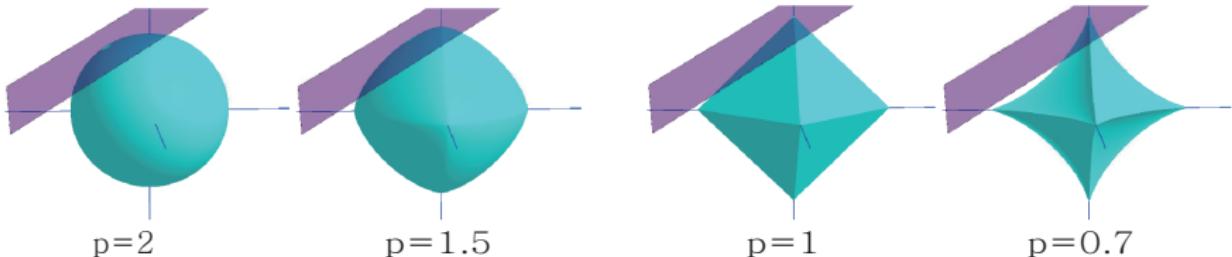
- If $\mathbf{x} \in \mathbb{R}^n$ is a s -sparse solution ($\|\mathbf{x}\|_0 = s$) to $\mathbf{Ax} = \mathbf{b}$ and

$$s < \frac{1}{2}(1 + \mu(\mathbf{A})),$$

then \mathbf{x} is the unique solution to l_1 -minimization problem

$$\min \|\mathbf{x}\|_1 \text{ subject to } \mathbf{Ax} = \mathbf{b},$$

and this is also solution to l_0 problem.



Sparse driven reconstruction

- With the aid of sparsity of $|\nabla f|$, sparse image can be reconstructed by solving the following problem.

$$\min \|\nabla f\|_1 \text{ subject to } \|\mathcal{R}f - P\|_2^2 \leq \varepsilon.$$



Original image



Absolute value of gradient

- One can use Ψf (wavelet, DCT, rank,...)

Gradient-descent method

- With L1 regularization term, data-fitting model is given by

$$\min_f \|\nabla f\|_1 + \frac{\lambda}{2} \|\mathcal{R}f - P_s\|_2^2$$

- Euler-Lagrange equation provides

$$-\nabla \cdot \left(\frac{\nabla f}{|\nabla f|} \right) + \lambda \mathcal{R}^*(\mathcal{R}f - P_s) = 0$$

- Using the gradient descent method, we have

$$\frac{\partial}{\partial t} f(\mathbf{x}, t) = \nabla \cdot \left(\frac{\nabla f}{\sqrt{|\nabla f|^2 + \beta}} \right) - \lambda \mathcal{R}^*(\mathcal{R}f - P_s), \quad \mathbf{x} = (x, y)$$

- Using the finite difference scheme, we have

$$f(\mathbf{x}, t^{n+1}) = f(\mathbf{x}, t^n) + \Delta t \left[\frac{f_{xx}(f_y^2 + \beta) - 2f_x f_y f_{xy} + f_{yy}(f_x^2 + \beta)}{(f_x^2 + f_y^2 + \beta)^{3/2}} - \lambda \mathcal{R}^*(\mathcal{R}f(\mathbf{x}, t^n) - P_s) \right]$$

Gradient-descent method

- Two disadvantages: not exact solution, convergence speed.

$$\rightarrow \frac{\partial}{\partial t} f(\mathbf{x}, t) = \nabla \cdot \left(\frac{\nabla f}{\sqrt{|\nabla f|^2 + \beta}} \right) - \lambda \mathcal{R}^*(\mathcal{R}f - P_s)$$

$$\rightarrow \text{step size: } \Delta t \leq c \Delta x^2 \sqrt{|\nabla f|^2 + \beta}$$

Augmented Lagrangian method

$$\min_f \|\nabla f\|_1 + \frac{\lambda}{2} \|\mathcal{R}f - P\|_2^2$$

- Augmented Lagrangian method overcomes these defects. By introducing auxiliary variable $\mathbf{q} = \nabla f$, we solve

$$\begin{aligned} \min_{f, \mathbf{q}} \max_{\boldsymbol{\tau}} & \int_{\Omega} |\mathbf{q}| d\mathbf{x} + \frac{\gamma}{2} \int_0^{2\pi} \int_{\mathbb{R}} |P - \mathcal{R}f|^2 d\varphi ds \\ & + \int_{\Omega} \boldsymbol{\tau} \cdot (\mathbf{q} - \nabla f) d\mathbf{x} + \frac{r}{2} \int_{\Omega} |\mathbf{q} - \nabla f|^2 d\mathbf{x}. \end{aligned}$$

- AL algorithm

- Initialization: $\mathbf{q}^0 = \boldsymbol{\tau}^0 = \mathbf{0}$.
- For fixed multiplier $\boldsymbol{\tau}^k$, solve

$$f^{k+1} = \arg \min_f \frac{\gamma}{2} \int_0^{2\pi} \int_{\mathbb{R}} |P - \mathcal{R}f|^2 d\varphi ds - \int_{\Omega} \boldsymbol{\tau}^k \cdot \nabla f d\mathbf{x} + \frac{r}{2} \int_{\Omega} |\mathbf{q} - \nabla f|^2 d\mathbf{x} \quad (2)$$

and solve

$$\mathbf{q}^{k+1} = \arg \min_{\mathbf{q}} \int_{\Omega} |\mathbf{q}| d\mathbf{x} + \int_{\Omega} \boldsymbol{\tau}^k \cdot \mathbf{q} d\mathbf{x} + \frac{r}{2} \int_{\Omega} |\mathbf{q} - \nabla f|^2 d\mathbf{x} \quad (3)$$

- Update $\boldsymbol{\tau}^{k+1}$ by

$$\boldsymbol{\tau}^{k+1} = \boldsymbol{\tau}^k + r(\mathbf{q}^{k+1} - \nabla f^{k+1})$$

Augmented Lagrangian method

- The optimality condition of the quadratic problem (2) gives

$$(-r\nabla^2 + \gamma\mathcal{R}^*\mathcal{R})f = \gamma\mathcal{R}^*P - \nabla \cdot (r\mathbf{q} + \boldsymbol{\tau}^k),$$

- Since (3) can be rewritten as

$$\arg \min_{\mathbf{q}} \frac{1}{r} \int_{\Omega} |\mathbf{q}| d\mathbf{x} + \frac{1}{2} \int_{\Omega} |\mathbf{q} - (\nabla f - \frac{1}{r}\boldsymbol{\tau}^k)|^2 d\mathbf{x},$$

optimal value in (2) can be explicitly computed using Shrinkage operator:

$$\mathbf{q} = \frac{1}{r} shrink(r\nabla f - \boldsymbol{\tau}^k, 1),$$

where Shrinkage operator *shrink* defined by

$$shrink(\mathbf{x}, \gamma) = \frac{\mathbf{x}}{|\mathbf{x}|} \max(|\mathbf{x}| - \gamma, 0).$$



Thank
you!