supported by project webpage

Blind Polychromatic X-Ray CT Reconstruction from Poisson Measurements

顾任亮 Renliang Gu and Aleksandar Dogandžić {renliang, ald}@iastate.edu
Department of Electrical and Computer Engineering

IOWA STATE UNIVERSITY

Introduction

- Most polychromatic X-ray CT reconstruction methods assume known X-ray spectrum and materials. However,
  - the X-ray spectrum measurements based on the semiconductor detectors are usually distorted by charge trapping, escape events, and other effects [Red+09] and the corresponding correction requires highly collimated beam and special procedures [Lin+14].
  - knowing the mass-attenuation function can be challenging when the inspected material is unknown, or the inspected object is made of compound or mixture with unknown percentage of each constituent.

Our Goal: Develop a blind sparse density-map reconstruction scheme from measurements corrupted by Poisson noise.

Notation: "$\succeq$" is the elementwise version of "$\geq$"; the elementwise log $[\ln_\circ \boldsymbol{a}]_i = \ln a_i, \forall i$; soft-thresholding operator $[\mathcal{T}_\lambda(\boldsymbol{a})]_i = \operatorname{sign}(a_i)\max(|a_i| - \lambda, 0), \forall i$. $\iota^{\mathrm{L}}(s)$ is the *Laplace transform* of $\iota(\kappa)$: $\iota^{\mathrm{L}}(s) \triangleq \int \iota(\kappa) \mathrm{e}^{-s\kappa}\,\mathrm{d}\kappa$, Laplace transform with vector argument: $\boldsymbol{a}^{\mathrm{L}}_\circ(\boldsymbol{s}) = (\boldsymbol{a}^{\mathrm{L}}(s_n))_{n=1}^{N}$ obtained by stacking $\boldsymbol{a}^{\mathrm{L}}(s_n)$ columnwise, where $\boldsymbol{s} = (s_n)_{n=1}^{N}$.

Measurement Model

Denote by N the total number of measurements from all projections collected at the detector array. For the nth measurement, define its discretized line integral as $\boldsymbol{\alpha}$; stacking all N such integrals into a vector yields $\Phi\boldsymbol{\alpha}$, where

$$\Phi = [\boldsymbol{\phi}_1 \boldsymbol{\phi}_2 \cdots \boldsymbol{\phi}_N]^T \in \mathbb{R}^{N \times p} \tag{1}$$

is the known *projection matrix*. Construct *mass-attenuation spectrum* $\iota(\kappa)$ [GD13; GD16] (see Fig. 1) and expand it as

$$\iota(\kappa) = \boldsymbol{b}(\kappa)\mathcal{I} \tag{2a}$$

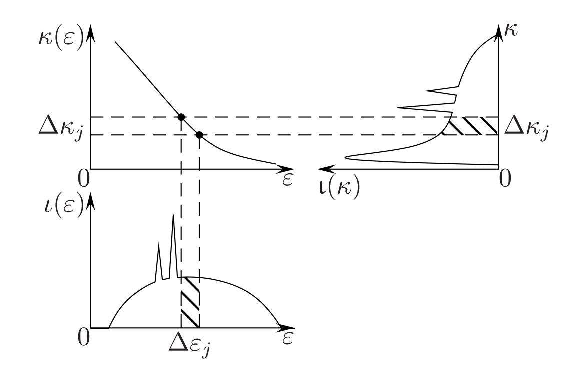


Figure 1: Relationship between mass attenuation κ, incident spectrum ι, photon energy ε, and mass-attenuation spectrum $\iota(\kappa)$.

where $\boldsymbol{b}(\kappa)$ are known $1 \times J$ B1-spline vectors with knots $\kappa_j = \kappa_0 q^j$ selected from a growing geometric series with common ratio $q > 1$, J is the number of basis functions, and

$$\mathcal{I} = (\mathcal{I}_j)_{j=1}^{J} \succeq \boldsymbol{0} \tag{2b}$$

is an *unknown* $J \times 1$ vector of corresponding basis-function coefficients; see Fig. 2.

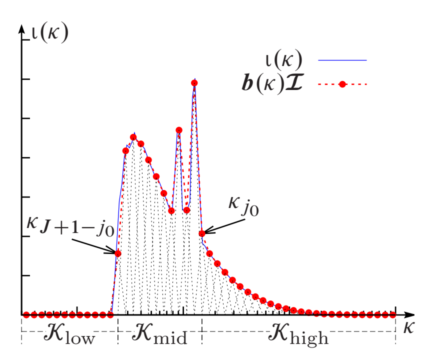


Figure 2: B1-spline approximation of $\iota(\kappa)$.

Noiseless measurements. $N \times 1$ vector of noiseless energy measurements:

$$\mathcal{I}^{\mathrm{out}}(\boldsymbol{\alpha}, \mathcal{I}) = \boldsymbol{b}^{\mathrm{L}}_\circ(\Phi\boldsymbol{\alpha})\mathcal{I} \tag{3}$$

where

$$\boldsymbol{\alpha} = (\alpha_i)_{i=1}^{p} \succeq \boldsymbol{0} \tag{4}$$

is an *unknown* $p \times 1$ vector representing the 2D image we wish to reconstruct and $\boldsymbol{b}^{\mathrm{L}}_\circ(\boldsymbol{s})$ is an *output basis-function matrix* obtained by stacking the $1 \times J$ vectors $\boldsymbol{b}^{\mathrm{L}}(s_n)$ columnwise.

Noisy measurements. For independent Poisson measurements $\mathcal{E} = (\mathcal{E}_n)_{n=1}^{N}$, the negative log-likelihood (NLL) is

$$\mathcal{L}(\boldsymbol{\alpha}, \mathcal{I}) = \boldsymbol{1}^T\left[\mathcal{I}^{\mathrm{out}}(\boldsymbol{\alpha}, \mathcal{I}) - \mathcal{E}\right] - \sum_{n, \mathcal{E}_n \neq 0} \mathcal{E}_n \ln \frac{\mathcal{I}^{\mathrm{out}}_n(\boldsymbol{\alpha}, \mathcal{I})}{\mathcal{E}_n}. \tag{5}$$

Theorem 1 (Biconvexity)

The NLL (5) is biconvex with respect to $\boldsymbol{\alpha}$ and $\mathcal{I}$ in the following set:

$$\left\{(\boldsymbol{\alpha}, \mathcal{I}) \,\middle|\, \mathcal{I}^{\mathrm{out}}(\boldsymbol{\alpha}, \mathcal{I}) \succeq \frac{(q^{j_0} - 1)^2}{q^{2j_0} + 1}\mathcal{E},\ \mathcal{I} \in \mathcal{A},\ \boldsymbol{\alpha} \in \mathbb{R}^p_+\right\} \quad \text{where} \tag{6a}$$

$$\mathcal{A} = \Big\{\mathcal{I} \in \mathbb{R}^J_+ \,\Big|\, \mathcal{I}_1 \leq \mathcal{I}_2 \leq \cdots \leq \mathcal{I}_{J+1-j_0},\ \mathcal{I}_{j_0} \geq \cdots \geq \mathcal{I}_J,$$
$$\text{and } \mathcal{I}_j \geq \mathcal{I}_{J+1-j_0},\ \forall j \in [J+1-j_0, j_0]\Big\}, \qquad j_0 \geq \lceil (J+1)/2 \rceil. \tag{6b}$$

Penalized NLL Objective Function

Our goal is to compute penalized maximum-likelihood estimates of the density-map and mass-attenuation spectrum parameters $(\boldsymbol{\alpha}, \mathcal{I})$ by solving

$$\min_{\boldsymbol{\alpha}, \mathcal{I}} f(\boldsymbol{\alpha}, \mathcal{I}) \tag{7}$$

where

$$f(\boldsymbol{\alpha}, \mathcal{I}) = \mathcal{L}(\boldsymbol{\alpha}, \mathcal{I}) + u r(\boldsymbol{\alpha}) + \mathbb{I}_{[0,+\infty)}(\mathcal{I}) \tag{8a}$$

$$r(\boldsymbol{\alpha}) = \sum_{i=1}^{p} \sqrt{\sum_{j \in \mathcal{N}_i} (\alpha_i - \alpha_j)^2} + \mathbb{I}_{[0,+\infty)}(\boldsymbol{\alpha}) \tag{8b}$$

are the penalized NLL objective function and the density-map regularization term that enforces nonnegativity and sparsity of the signal $\boldsymbol{\alpha}$ in the total-variation (TV) domain. Here, $u > 0$ is a scalar tuning constant and $\mathcal{N}_i$ is index set of neighbors of α_i, where the elements of $\boldsymbol{\alpha}$ are arranged to form a 2D image [BT09].

Corollary 1

$f(\boldsymbol{\alpha}, \mathcal{I})$ is biconvex with respect to $\boldsymbol{\alpha}$ and $\mathcal{I}$ under the conditions of Theorem 1.

Theorem 2 (Kurdyka-Łojasiewicz (KL) Property)

$f(\boldsymbol{\alpha}, \mathcal{I})$ satisfies the KL property in any compact subset $\mathbb{C} \subseteq \operatorname{dom}(f)$.

Minimization Algorithm

Iteration i for minimizing (8a) updates $\boldsymbol{\alpha}$ and $\mathcal{I}$ alternatively:

1. (NPG) Fix $\mathcal{I} = \mathcal{I}^{(i-1)}$ and descend $f(\boldsymbol{\alpha}, \mathcal{I}^{(i-1)})$ by applying a *Nesterov's proximal-gradient (NPG) step* [Nes83] for $\boldsymbol{\alpha}$:

$$\theta^{(i)} = \frac{1}{2}\left[1 + \sqrt{1 + 4\left(\theta^{(i-1)}\right)^2}\right] \tag{9a}$$

$$\overline{\boldsymbol{\alpha}}^{(i)} = \boldsymbol{\alpha}^{(i-1)} + \frac{\theta^{(i-1)} - 1}{\theta^{(i)}}\left(\boldsymbol{\alpha}^{(i-1)} - \boldsymbol{\alpha}^{(i-2)}\right) \qquad \text{Nesterov's acceleration} \tag{9b}$$

$$\boldsymbol{\alpha}^{(i)} = \arg\min_{\boldsymbol{\alpha}} \frac{1}{2\beta^{(i)}}\left\|\boldsymbol{\alpha} - \overline{\boldsymbol{\alpha}}^{(i)} + \beta^{(i)}\nabla\mathcal{L}_\iota(\overline{\boldsymbol{\alpha}}^{(i)})\right\|_2^2 + u r(\boldsymbol{\alpha}) \tag{9c}$$

where $\mathcal{L}_\iota(\boldsymbol{\alpha}) = \mathcal{L}(\boldsymbol{\alpha}, \mathcal{I}^{(i-1)})$, the minimization (9c) is computed using an inner iteration that employs the TV-based denoising method in [BT09, Sec. IV], and $\beta^{(i)} > 0$ is an adaptive step size chosen to satisfy the majorization condition:

$$\mathcal{L}_\iota(\boldsymbol{\alpha}^{(i)}) \leq \mathcal{L}_\iota(\overline{\boldsymbol{\alpha}}^{(i)}) + \left(\boldsymbol{\alpha}^{(i)} - \overline{\boldsymbol{\alpha}}^{(i)}\right)^T \nabla\mathcal{L}_\iota(\overline{\boldsymbol{\alpha}}^{(i)}) + \frac{1}{2\beta^{(i)}}\left\|\boldsymbol{\alpha}^{(i)} - \overline{\boldsymbol{\alpha}}^{(i)}\right\|_2^2 \tag{9d}$$

using a patient adaptation scheme that aims at finding the largest $\beta^{(i)}$ that satisfies (9d), see [GD15] for details. We apply *function restart* [OC13] to restore the monotonicity and improve convergence of NPG steps.

2. (BFGS) Set the design matrix $A = \boldsymbol{b}^{\mathrm{L}}_\circ(\Phi\boldsymbol{\alpha}^{(i)})$, treat it as known, and minimize the regularized NLL function $f(\boldsymbol{\alpha}^{(i)}, \mathcal{I})$ with respect to $\mathcal{I}$, i.e., solve

$$\mathcal{I}^{(i)} = \arg\min_{\mathcal{I} \succeq \boldsymbol{0}} \boldsymbol{1}^T(A\mathcal{I} - \mathcal{E}) - \sum_{n, \mathcal{E}_n \neq 0} \mathcal{E}_n \ln \frac{[A\mathcal{I}]_n}{\mathcal{E}_n} \tag{10}$$

using the inner limited-memory Broyden-Fletcher-Goldfarb-Shanno with box constraints (L-BFGS-B) iteration [Byr+95], initialized by $\mathcal{I}^{(i-1)}$.

Iterate between Steps 1 and 2 until the relative distance of consecutive iterates of the density map $\boldsymbol{\alpha}$ does not change significantly:

$$\left\|\boldsymbol{\alpha}^{(i)} - \boldsymbol{\alpha}^{(i-1)}\right\|_2 < \epsilon\left\|\boldsymbol{\alpha}^{(i)}\right\|_2 \tag{11}$$

where $\epsilon > 0$ is the convergence threshold. The convergence criteria for the inner TV-denoising and L-BFGS-B iterations are chosen to trade off the accuracy and speed of the inner iterations and provide sufficiently accurate solutions to (9c) and (10).

Remark 1 (Monotonicity)

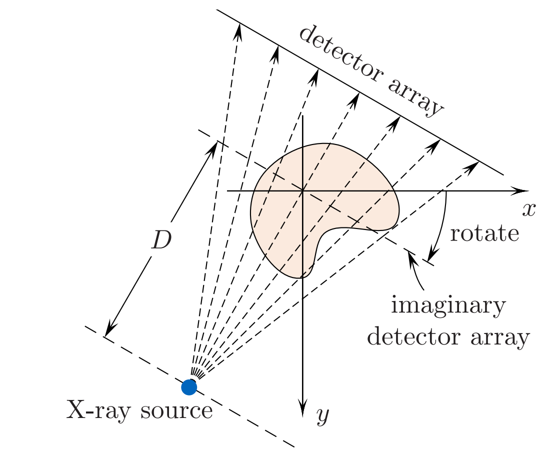
Under the condition (6a) of Theorem 1, the NPG-BFGS iteration with function restart is monotonically non-increasing:

$$f\left(\boldsymbol{\alpha}^{(i)}, \mathcal{I}^{(i)}\right) \leq f\left(\boldsymbol{\alpha}^{(i-1)}, \mathcal{I}^{(i-1)}\right) \qquad \forall i. \tag{12}$$

Numerical Example

Performance metric is the relative square error (RSE) of an estimate $\widehat{\boldsymbol{\alpha}}$ of the signal coefficient vector:

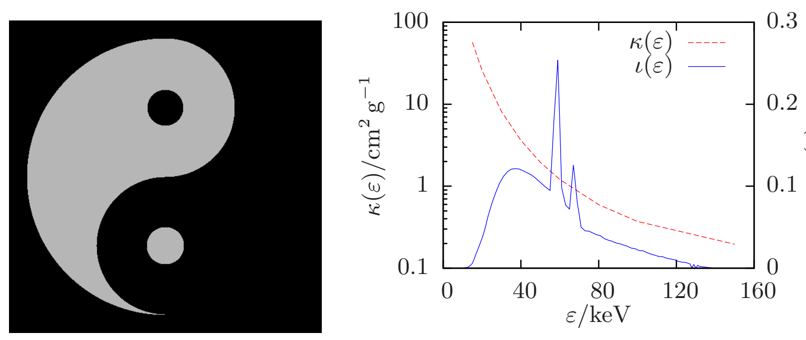
$$\mathrm{RSE}\{\widehat{\boldsymbol{\alpha}}\} = 1 - \left(\frac{\widehat{\boldsymbol{\alpha}}^T\boldsymbol{\alpha}_{\mathrm{true}}}{\|\widehat{\boldsymbol{\alpha}}\|_2\|\boldsymbol{\alpha}_{\mathrm{true}}\|_2}\right)^2.$$



We compare
- NPG-BFGS method,
- NPG for known mass attenuation spectrum $\iota(\kappa)$;
- linearized basis pursuit denoising (linearized BPDN), which applies the NPG approach to solve the BPDN problem [BT09]:

  $\min_{\boldsymbol{\alpha}} 0.5\|\boldsymbol{y} - \Phi\boldsymbol{\alpha}\|_2^2 + u' r(\boldsymbol{\alpha})$, where $\boldsymbol{y} = (\iota^{\mathrm{L}})_\circ^{-1}(\mathcal{E})$ are the linearized measurements,
- the traditional filtered backprojection (FBP) method without [KS88, Ch. 3] and with linearization [Her79], i.e., based on the 'data' $\boldsymbol{y} = -\ln_\circ \mathcal{E}$ and $\boldsymbol{y} = (\iota^{\mathrm{L}})_\circ^{-1}(\mathcal{E})$.

In Fig. 4, average RSEs of the methods that do not assume knowledge of the mass-attenuation spectrum $\iota(\kappa)$ are shown using solid lines whereas dashed lines represent methods that assume known $\iota(\kappa)$. Red and blue colors present methods that *do* and *do not* employ signal-sparsity regularization, respectively.



(a) (b)

Figure 3: (a) Density-map image and (b) mass attenuation and incident X-ray spectrum as functions of the photon energy ε.

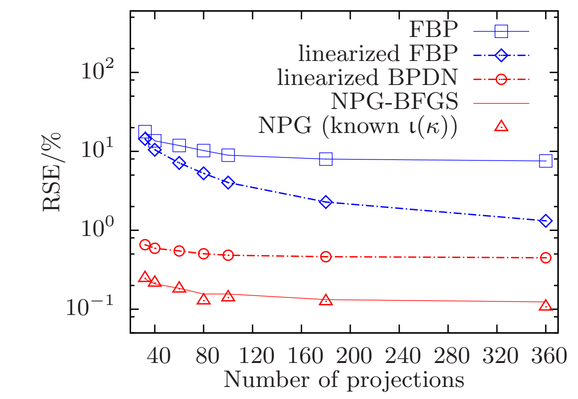


Figure 4: Average RSEs as functions of the number of projections.

(a) FBP (b) linearized FBP (c) linearized BPDN (d) NPG known $\iota(\kappa)$ (e) NPG-BFGS

Figure 5: Reconstructions from 60 projections.

References

R. Redus *et al.*, "Characterization of CdTe detectors for quantitative X-ray spectroscopy", *IEEE Trans. Nucl. Sci.*, vol. 56, no. 4, pp. 2524–2532, 2009.

Y. Lin *et al.*, "An angle-dependent estimation of CT x-ray spectrum from rotational transmission measurements", *Med. Phys.*, vol. 41, no. 6, p. 062 104, 2014.

R. Gu and A. Dogandžić, "Beam hardening correction via mass attenuation discretization", in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Vancouver, Canada, May 2013, pp. 1085–1089.

——, "Blind X-ray CT image reconstruction from polychromatic Poisson measurements", *IEEE Trans. Comput. Imag.*, 2016, to appear.

A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems", *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, 2009.

Y. Nesterov, "A method of solving a convex programming problem with convergence rate $O(1/k^2)$", in *Sov. Math. Dokl.*, vol. 27, 1983, pp. 372–376.

R. Gu and A. Dogandžić, "Projected Nesterov's proximal-gradient signal recovery from compressive Poisson measurements", in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 2015, to appear.

B. O'Donoghue and E. Candès, "Adaptive restart for accelerated gradient schemes", *Found. Comput. Math.*, pp. 1–18, Jul. 2013.

R. H. Byrd *et al.*, "A limited memory algorithm for bound constrained optimization", *SIAM J. Sci. Comput.*, vol. 16, no. 5, pp. 1190–1208, 1995.

A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. New York: IEEE Press, 1988.

G. T. Herman, "Correction for beam hardening in computed tomography", *Phys. Med. Biol.*, vol. 24, no. 1, pp. 81–106, 1979.