

# BEAM HARDENING CORRECTION VIA MASS ATTENUATION DISCRETIZATION

Renliang Gu and Aleksandar Dogandžić

ECpE Department, Iowa State University  
3119 Coover Hall, Ames, IA 50011  
email: {renliang,ald}@iastate.edu

## ABSTRACT

We develop a beam-hardening correction method for polychromatic x-ray computed tomography (CT) reconstruction based on *mass attenuation coefficient discretization*. We assume that the inspected object consists of an unknown single material and that the incident x-ray spectrum is unknown. In this case, the standard photon-energy discretization of the Beer's law measurement equation leads to an excessive number of unknown parameters and scaling ambiguity. To obtain a parsimonious measurement model parametrization, we first rewrite the measurement equation in terms of integral expressions of the mass attenuation rather than photon energy. The resulting integrals can be discretized easily thanks to the fact that the range of mass attenuations is bounded and, in practice, fairly narrow. We then develop a constrained least-squares optimization approach for reconstructing the underlying object from log-scale measurements, where we impose the nonnegativity constraint to both the signal and the x-ray spectrum density estimates. We demonstrate the performance of the proposed method via a numerical example where we compare it with the standard filtered backprojection (FBP), which ignores the polychromatic nature of the measurements.

**Index Terms**— Beam hardening, computed tomography, signal reconstruction, x-ray tomography.

## 1. INTRODUCTION

Due to the bremsstrahlung phenomenon [1], x-rays generated by vacuum tubes are not monochromatic [2, 3], which causes the beam hardening effect. To describe a polychromatic x-ray source, assume that its incident energy  $\mathcal{I}^{\text{in}}$  spreads along photon energy  $\varepsilon$  following the density  $\iota(\varepsilon)$ , i.e.,

$$\int \iota(\varepsilon) d\varepsilon = \mathcal{I}^{\text{in}}. \quad (1a)$$

According to the Beer's law, the noiseless measurement collected by an energy integral detector upon traversing a straight line  $\ell = \ell(x, y)$  through a single-material object is [4–6]

$$\mathcal{I}^{\text{out}} = \int \iota(\varepsilon) \exp\left[-\mu(\varepsilon) \int \alpha(x, y) d\ell\right] d\varepsilon \quad (1b)$$

where  $\alpha(x, y)$  is the inspected object's density,  $(x, y)$  are the Cartesian coordinates, and  $\mu(\varepsilon)$  is the mass attenuation coefficient of the material, which depends on the photon energy  $\varepsilon$ .

A standard approach to simulate the polychromatic x-ray computed tomography (CT) measurements is to discretize (1a) and (1b) by approximating the corresponding integrals over photon energy with summa-

tions [3, Sec. 8.4]:

$$\mathcal{I}^{\text{in}} = \sum_{j=1}^J \iota(\varepsilon_j) \Delta\varepsilon_j \quad (2a)$$

$$\mathcal{I}^{\text{out}} = \sum_{j=1}^J \iota(\varepsilon_j) e^{-\phi^T \alpha \mu(\varepsilon_j)} \Delta\varepsilon_j \quad (2b)$$

where  $\varepsilon_0 < \varepsilon_1 < \varepsilon_2 < \dots < \varepsilon_J$  are the known discretization points along the  $\varepsilon$  axis and  $\Delta\varepsilon_j = \varepsilon_j - \varepsilon_{j-1}$  is the length of the interval between  $\varepsilon_j$  and  $\varepsilon_{j-1}$ ,  $\alpha$  is a  $p \times 1$  vector representing the two-dimensional image that we wish to reconstruct, and  $\phi$  is a  $p \times 1$  vector of weights quantifying how much each element of  $\alpha$  contributes to the x-ray attenuation on the straight-line path  $\ell$ .

The discretization (2) has been employed in beam-hardening correction schemes [5–10]. Van Gompel et al. [10] consider a “blind” scenario with *unknown* incident spectrum and materials, but assume that the number of materials is known and that each pixel is occupied by a single material; they employ the  $K$ -means clustering method to initially associate pixels to the materials and then alternate between material segmentation and updating the relative density map, incident x-ray spectrum, and mass attenuation coefficients for each material.

In the blind scenario when both  $\mu(\varepsilon)$  and  $\alpha(x, y)$  are unknown, their product suffers from scaling ambiguity, the number of unknown parameters that need to be estimated is excessive, and the sequence  $\{(\mu(\varepsilon_j), \iota(\varepsilon_j))\}$  of discretized  $\mu(\varepsilon)$  and  $\iota(\varepsilon)$  can be permuted arbitrarily, where  $\{\varepsilon_j\}_{j=1}^J$  are the discretization points over photon energy. In this paper, we discretize the Beer's law over the mass attenuation, which leads to fewer estimation parameters, and employ it to design a beam hardening correction scheme for the blind scenario.

We introduce the notation:  $\mathbf{1}_{N \times 1}$  and  $\mathbf{0}_{N \times 1}$  the  $N \times 1$  vector of ones and zeros, “ $\circ$ ” denotes the elementwise (Hadamard) product, and  $|\cdot|$ ,  $\|\cdot\|_2$ , and “ $^T$ ” are the absolute value, Euclidean norm, and transpose, respectively. Furthermore,  $(x)_+ = \max\{x, 0\}$  is the positive-part operator,  $\text{supp}(\iota(\cdot))$  returns the support set of a function  $\iota(\cdot)$ ,  $\mathbf{y} \succeq \mathbf{0}_{N \times 1}$  denotes that all elements of a real-valued  $N \times 1$  vector  $\mathbf{y}$  are nonnegative,  $\lceil x \rceil$  is the smallest integer larger than or equal to  $x$ , and

$$\mathbb{1}_A(y) = \begin{cases} 1, & y \in A, \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

denotes the indicator function. Define the elementwise logarithm,  $k$ th power, and positive-part operators of an  $N$ -dimensional real vector  $\mathbf{x} = [x_1, \dots, x_N]^T$  as  $\ln_{\circ}(\mathbf{x}) = [\ln x_1, \dots, \ln x_N]^T$ ,  $\mathbf{x}^{\circ k} = [x_1^k, x_2^k, \dots, x_N^k]^T$ , and  $(\mathbf{x})_+ = [(x_1)_+, \dots, (x_N)_+]^T$ , respectively.

This work was supported by the NSF Industry-University Cooperative Research Program, Center for Nondestructive Evaluation, Iowa State University.

## 2. X-RAY CT MODEL VIA MASS ATTENUATION

Assume that the inspected single material is unknown and, consequently,  $\mu(\varepsilon)$  is an unknown function as well. To reduce the number of parameters that we need to estimate, we *re-parametrize* our measurement model and represent it in terms of integral expressions of  $\mu$  rather than  $\varepsilon$ .

Observe that the mass attenuation  $\mu(\varepsilon)$  and incident energy density  $\iota(\varepsilon)$  are both functions of  $\varepsilon$ , see Fig. 1. Thus, to combine the variation of these two functions and reduce the degree of freedom, we rewrite  $\iota(\varepsilon)$  as a function of  $\mu$  and set  $\mu$  as the integral variable. For invertible  $\mu(\varepsilon)$ , we define its inverse as  $\varepsilon(\mu)$ . The change of variables  $\varepsilon = \varepsilon(\mu)$  in the integral expressions (1a) and (1b) yields

$$\mathcal{I}^{\text{in}} = \int \iota(\varepsilon(\mu)) |\varepsilon'(\mu)| d\mu \quad (4a)$$

$$\mathcal{I}^{\text{out}} = \int \iota(\varepsilon(\mu)) |\varepsilon'(\mu)| e^{-\mu \int \alpha(x,y) d\ell} d\mu \quad (4b)$$

where we have also assumed that the function  $\varepsilon(\mu)$  is differentiable with derivative  $\varepsilon'(\mu) = d\varepsilon(\mu)/d\mu$ . For invertible  $\mu(\varepsilon)$ ,  $\varepsilon(\mu)$  is a decreasing function of  $\mu$ ; hence,  $|\varepsilon'(\mu)| = -\varepsilon'(\mu)$ .

All  $\mu(\varepsilon)$  encountered in practice can be divided into piecewise-continuous segments, where each segment is a differentiable monotonically decreasing function of  $\varepsilon$  [4, Tables 3 and 4] and [11, Sec. 2.3]. The points of discontinuity in  $\mu(\varepsilon)$  are referred to as  $K$ -edges and are caused by the interaction between photons and  $K$  shell electrons. A  $K$ -edge occurs only when  $\varepsilon$  reaches the binding energy of the shell electron.) An extension of (4) to this scenario is straightforward, but results in a lengthy expression.

### 2.1. Discretization over Mass Attenuation

We discretize (4a) and (4b) in the spatial and mass attenuation domains using  $p$  pixels and  $J$  mass attenuation bins:

$$\mathcal{I}^{\text{in}} = \sum_{j=1}^J \mathcal{I}_j \quad (5a)$$

$$\mathcal{I}^{\text{out}}(\theta) = \sum_{j=1}^J \mathcal{I}_j e^{-\mu_j \Phi^T \alpha} \quad (5b)$$

where

$$\theta = (\alpha, \mathcal{I}) \quad (6a)$$

$$\mathcal{I} = [\mathcal{I}_1, \mathcal{I}_2, \dots, \mathcal{I}_J]^T \quad (6b)$$

$\mu_0 < \mu_1 < \dots < \mu_J$  are known discretization points along the  $\mu$  axis,

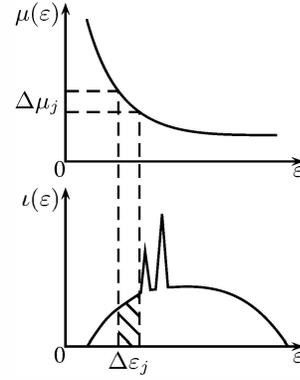
$$\mathcal{I}_j = \iota(\varepsilon(\mu_j)) |\varepsilon'(\mu_j)| \Delta\mu_j \quad (7a)$$

and  $\Delta\mu_j = \mu_j - \mu_{j-1}$  is the length of the interval between consecutive discretization points  $\mu_j$  and  $\mu_{j-1}$ . By substituting  $\varepsilon(\mu_j) = \varepsilon_j$  and  $|\varepsilon'(\mu_j)| \approx \Delta\varepsilon_j / \Delta\mu_j$  into (7a), we obtain

$$\mathcal{I}_j \approx \iota(\varepsilon_j) \Delta\varepsilon_j \quad (7b)$$

(depicted in Fig. 1) and verify the equivalence between (2a)–(2b) and (5a)–(5b). Note that (5) holds for piecewise-monotonic  $\mu(\varepsilon)$  as well, with a more complex expression for  $\mathcal{I}_j$  that generalizes (7a).

For the same number of discretization bins  $J$ , the standard photon-energy and proposed mass attenuation coefficient (MAC) discretizations yield  $p + 2J$  and  $p + J$  parameters, respectively. Intuitively, the number of functions to infer is reduced from two,  $\mu(\varepsilon)$  and  $\iota(\varepsilon)$  (photon-energy), to one,  $\iota(\mu)$  (MAC). The MAC discretization is further facilitated by the following facts:



**Fig. 1:** The mass attenuation coefficient and incident spectrum as functions of the photon energy  $\varepsilon$ .

- the mass attenuation coefficients  $\mu$  of almost all materials at any energy level are within the range  $10^{-2} \text{ cm}^2/\text{g}$  to  $10^4 \text{ cm}^2/\text{g}$ , see [4, Table 3];
- to reduce the beam hardening effect, the energy level of an x-ray scan is usually selected so that the function  $\mu(\varepsilon)$  is as flat as possible, yielding a narrow range of feasible values of  $\mu$  that is easy to discretize.

The chosen discretization points  $\{\mu_j\}_{j=1}^J$  need to have a sufficiently wide range to cover  $\mu(\text{supp}(\iota(\varepsilon)))$ ; we select them using the geometric sequence with common ratio  $q$ :

$$\mu_j = \mu_0 q^j, \quad j = 1, 2, \dots, J. \quad (8)$$

We now discuss the identifiability issues exhibited by the proposed MAC discretization. Clearly, the value of  $\mu_0$  in (8) can be arbitrary, since  $\mu_0$  can be absorbed by  $\alpha$  without affecting the value of  $\mathcal{I}^{\text{out}}$ , see also (5). If  $\mathcal{I}_1 = 0$ ,

$$\mathcal{I}^{\text{out}}((\alpha, [\mathcal{I}_1, \mathcal{I}_2, \dots, \mathcal{I}_J]^T)) \equiv \mathcal{I}^{\text{out}}((\alpha q, [\mathcal{I}_2, \mathcal{I}_3, \dots, \mathcal{I}_J, 0]^T)) \quad (9a)$$

Similarly, if  $\mathcal{I}_J = 0$ , then

$$\mathcal{I}^{\text{out}}((\alpha, [\mathcal{I}_1, \mathcal{I}_2, \dots, \mathcal{I}_J]^T)) \equiv \mathcal{I}^{\text{out}}((\alpha/q, [0, \mathcal{I}_1, \mathcal{I}_2, \dots, \mathcal{I}_{J-1}]^T)) \quad (9b)$$

Hence, if the range of  $\{\mu_j\}_{j=1}^J$  is sufficiently large to allow for zero edge elements of  $\mathcal{I}$ , then the recovery of  $\alpha$  will be correct up to a scale of common ratio  $q$ . See also Section 3 for further discussion on the selection of  $\{\mu_j\}_{j=1}^J$  and parameter identifiability.

### 2.2. Measurement Model and Estimation Algorithm

An x-ray CT scan consists of multiple projections with the beam intensity measured by multiple detectors. The vector of noiseless measurements at  $N$  detectors is  $A(\alpha) \mathcal{I}$  [see (5)], where the  $(i, j)$ th element of the  $N \times J$  matrix  $A(\alpha)$  is

$$A_{i,j}(\alpha) = \exp(-\phi_{(i)}^T \alpha \mu_j) \quad (10)$$

and  $\Phi = [\phi_{(1)}, \phi_{(2)}, \dots, \phi_{(N)}]^T$  is the Radon transform matrix of size  $N \times p$  for our imaging system.

We assume that the shadow of the inspected object is completely covered by the receiver array and wish to estimate the image  $\alpha$  from noisy measurements  $\{\mathcal{I}_i^{\text{mea}}\}_{i=1}^N$ , corresponding to  $\{\mathcal{I}_i^{\text{out}}\}_{i=1}^N$ . We impose the constraint on the incident energy  $\mathcal{I}^{\text{in}}$ :

$$\mathcal{I}^{\text{in}} = \sum_{j=1}^J \mathcal{I}_j \leq \mathcal{I}_{\text{MAX}}^{\text{in}}. \quad (11)$$

A simple possible choice for the upper bound in (11) is

$$\mathcal{I}_{\text{MAX}}^{\text{in}} = \max_{1 \leq i \leq N} \mathcal{I}_i^{\text{mea}} \quad (12)$$

where the maximum is likely achieved at a detector  $i$  that has a line-of-sight view of the x-ray source. The vector of unknown parameters is  $\theta$  in (6a), and the corresponding parameter space

$$\Theta = \{(\alpha, \mathcal{I}) \mid \alpha \succeq \mathbf{0}, \mathbf{1}^T \mathcal{I} \leq \mathcal{I}_{\text{MAX}}^{\text{in}}, \mathcal{I} \succeq \mathbf{0}\} \quad (13)$$

incorporates (11) and the facts that  $\iota(\varepsilon)$  and  $\alpha(x, y)$  are non-negative for all  $\varepsilon, x$  and  $y$ .

We adopt the following log-scale measurement model:

$$\mathbf{z} = \mathbf{f}(\theta) + \mathbf{n} = -\ln_{\circ}[A(\alpha)\mathcal{I}] + \mathbf{n} \quad (14a)$$

where

$$\mathbf{z} = [-\ln \mathcal{I}_1^{\text{mea}}, -\ln \mathcal{I}_2^{\text{mea}}, \dots, -\ln \mathcal{I}_N^{\text{mea}}]^T \quad (14b)$$

$$\mathbf{n} = [n_1, n_2, \dots, n_N]^T \quad (14c)$$

are the log-scale measurement and noise vectors, respectively. Assuming that  $\mathbf{n}$  is additive white Gaussian noise leads to the least-squares (LS) optimization problem:  $\min_{\theta \in \Theta} \|\mathbf{z} - \mathbf{f}(\theta)\|_2^2$ . A similar LS criterion has been used in [10], which, however, employs the standard discretization over photon energy and therefore estimates the mass attenuations  $\{\mu(\varepsilon_j)\}_{j=1}^J$  in addition to  $\theta$ . For known  $\alpha$ , (14a) corresponds to a generalized linear model for inference on  $\mathcal{I}$  with the Gaussian likelihood function and exponential *link function*  $e^{-z}$  [12, Sec. 2.2]. To leave some margin for the noise and discretization effects, we relax the nonnegative signal constraint  $\alpha \succeq \mathbf{0}$  and propose the following penalized LS objective function:

$$L_{\nu, t}(\theta) = \frac{1}{2} \|\mathbf{z} - \mathbf{f}(\theta)\|_2^2 + \frac{\nu}{2} \|(-\alpha)_+\|_2^2 + t \eta(\mathcal{I}) \quad (15a)$$

to be minimized with respect to  $\theta$ , where  $\nu$  and  $t$  are scalar tuning constants and the constraints on  $\mathcal{I}$  in (13) are imposed via the logarithmic barrier function [13, Sec. 11.2]

$$\eta(\mathcal{I}) = -\mathbf{1}_{J \times 1}^T \ln_{\circ}(\mathcal{I}) - \ln(\mathcal{I}_{\text{MAX}}^{\text{in}} - \mathbf{1}_{J \times 1}^T \mathcal{I}). \quad (15b)$$

**Minimization Algorithm.** Define the gradient vectors  $\mathbf{g}_{\alpha, \nu}(\theta)$ ,  $\mathbf{g}_{\mathcal{I}, t}(\theta)$  and Hessian matrices  $H_{\alpha, \nu}(\theta)$ ,  $H_{\mathcal{I}, t}(\theta)$  of the objective function (15a) with respect to  $\alpha$  and  $\mathcal{I}$ , respectively. (Here, the subscripts  $\nu$  and  $t$  emphasize the dependence on the tuning constant.)

We descend (15a) by alternating between (i) and (ii):

(i) the *nonlinear conjugate-gradient step* for  $\alpha$  [14, Sec. 14.1]

$$\alpha^{(i+1)} = \alpha^{(i)} - s_{\alpha} \frac{\mathbf{g}_{\alpha, \nu}^T(\theta^{(i)}) \mathbf{d}^{(i)}}{\mathbf{d}^{(i)T} H_{\alpha, \nu}(\theta^{(i)}) \mathbf{d}^{(i)}} \mathbf{d}^{(i)} \quad (16a)$$

where

$$\mathbf{e}^{(i)} = \mathbf{g}_{\alpha, \nu}(\theta^{(i)}) - \mathbf{g}_{\alpha, \nu}(\theta^{(i-1)}) \quad (16b)$$

$$\beta^{(i)} = \max \left\{ 0, \frac{\mathbf{g}_{\alpha, \nu}^T(\theta^{(i)}) \mathbf{e}^{(i)}}{\|\mathbf{g}_{\alpha, \nu}(\theta^{(i-1)})\|_2^2} \right\} \quad (16c)$$

$$\mathbf{d}^{(i)} = \mathbf{g}_{\alpha, \nu}(\theta^{(i)}) + \beta^{(i)} \mathbf{d}^{(i-1)}; \quad (16d)$$

(ii) • if

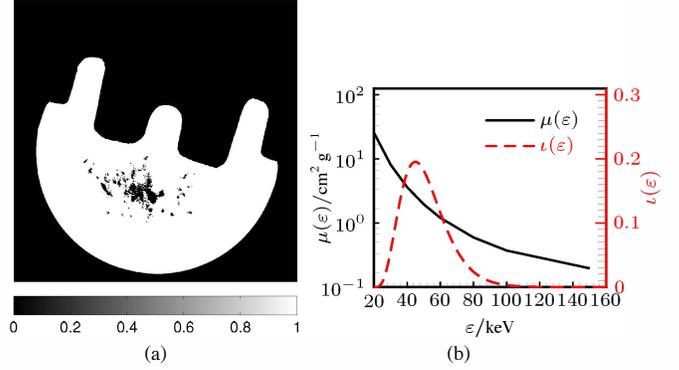
$$\mathbf{1}_{N \times 1} - \mathbf{z} + \mathbf{f}(\tilde{\theta}^{(i)}) \succeq \mathbf{0}_{N \times 1} \quad (17a)$$

holds, apply the *Newton step* for  $\mathcal{I}$

$$\mathcal{I}^{(i+1)} = \mathcal{I}^{(i)} - s_{\mathcal{I}} \left[ H_{\mathcal{I}, t}(\tilde{\theta}^{(i)}) \right]^{-1} \mathbf{g}_{\mathcal{I}, t}(\tilde{\theta}^{(i)}) \quad (17b)$$

where

$$\tilde{\theta}^{(i)} = (\alpha^{(i+1)}, \mathcal{I}^{(i)}); \quad (17c)$$



**Fig. 2:** (a) Original binary image of size  $1024^2$  and (b) mass attenuation coefficient of iron ( $Z=26$ ) and the incident x-ray spectrum density as functions of the photon energy.

- otherwise, i.e., if (17a) does not hold, simply let

$$\mathcal{I}^{(i+1)} = \mathcal{I}^{(i)}. \quad (17d)$$

Here,  $0 < s_{\alpha}, s_{\mathcal{I}} \leq 1$  are the step sizes determined via backtracking [15, Sec. 9.7] to guarantee the descent of (15a) and  $i$  denotes the iteration index.

In the nonlinear conjugate gradient step (16), we employ the Polak-Ribière formula (16c), which restarts the conjugate gradient iteration by forgetting the past search directions when  $\mathbf{g}_{\alpha, \nu}^T(\theta^{(i)}) \mathbf{e}^{(i)}$  is negative [16] to keep  $\alpha$  from going uphill. Since  $H_{\alpha, \nu}(\theta)$  has the form  $\Phi^T W(\theta) \Phi + \nu \text{diag}\{\mathbb{1}_{(-\infty, 0)}(\alpha)\}$ , where  $W(\theta)$  is an  $N \times N$  diagonal matrix, the denominator in (16a) is computed efficiently by computing and storing  $\Phi \mathbf{d}^{(i)}$ .

The Hessian of the LS cost  $\|\mathbf{z} - \mathbf{f}(\theta)\|_2^2$ ,

$$A^T(\alpha) \text{diag}\{[\mathbf{1}_{N \times 1} - \mathbf{z} + \mathbf{f}(\theta)] \circ [A(\alpha)\mathcal{I}]^{\circ-2}\} A(\alpha)$$

is positive definite and  $\|\mathbf{z} - \mathbf{f}(\theta)\|_2^2$ , as well as (15a), is a convex function of  $\mathcal{I}$ , if  $\mathbf{1}_{N \times 1} - \mathbf{z} + \mathbf{f}(\theta) \succeq \mathbf{0}_{N \times 1}$ , justifying the condition (17a) in (ii) above.

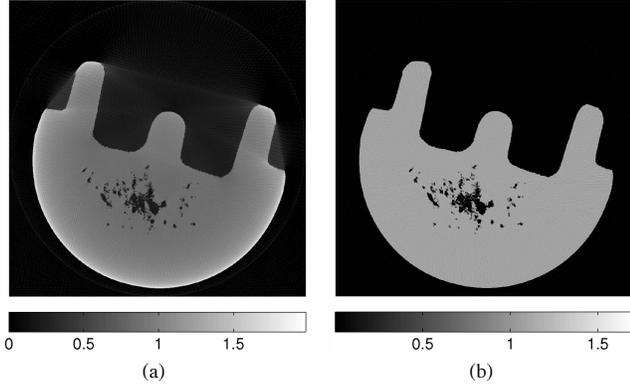
We decrease the value of  $t$  at each instance where  $\|\mathcal{I}^{(i+1)} - \mathcal{I}^{(i)}\|_2^2$  is sufficiently small until smaller than  $\epsilon/(J+1)$ , where  $\epsilon$  determines the convergence accuracy (see [13, Sec. 11.2]). This strategy is called the *barrier* (or *path-following*) method [13, Sec. 11.3].

### 3. NUMERICAL EXAMPLE

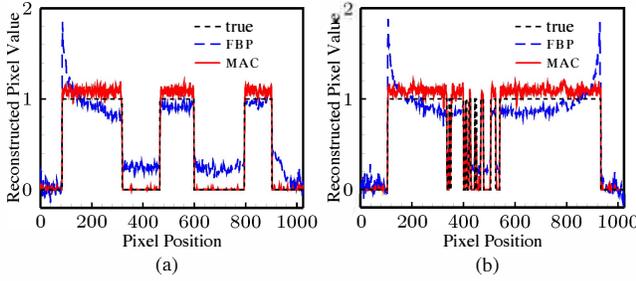
#### 3.1. Simulated Beam Hardening Correction

We construct a simulation example based on a binary  $1024 \times 1024$  image in Fig. 2(a) that corresponds to a real x-ray CT reconstruction of a metal casting, obtained by thresholding the pixel values of a reconstruction in [17, Fig. 5(b)]. The inspected object, assumed to be made of iron, contains irregularly shaped inclusions. We assume that the x-ray signal does not attenuate as it passes through the inclusions. The functional dependence of the mass attenuation coefficients on the photon energy for iron has been obtained by spline interpolation of the corresponding measurements from the National Institute of Standards and Technology (NIST) database [4]. The spectrum  $\iota(\varepsilon)$  of the incident x-ray is modeled as a scaled and shifted Gamma(5, 1) probability density function (pdf) in range from 20 keV to 150 keV:

$$\iota(\varepsilon) = \text{Gamma}\left(\frac{4(\varepsilon - 20)}{25} \mid 5, 1\right) \mathbb{1}_{[20, 150]}(\varepsilon) \quad (18)$$



**Fig. 3:** (a) The FBP and (b) MAC reconstructions.

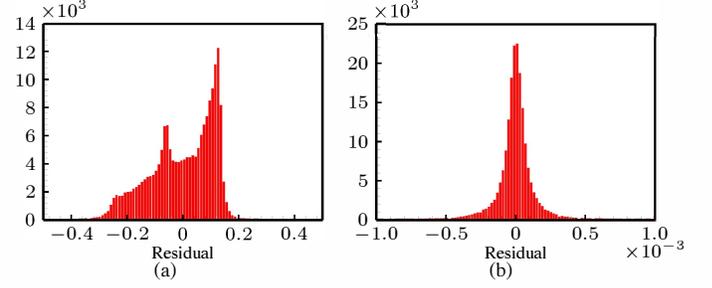


**Fig. 4:** (a) The 500-th and (b) 700-th row profiles of the true image and FBP and MAC reconstructions.

see Fig. 2(b) and [18, Sec. 3.3] for the definition of  $\text{Gamma}(x|\alpha, \beta)$ . We simulated the polychromatic sinogram measurements using the conventional photon-energy discretization (2) with 130 equi-spaced discretization points over the range 20 keV to 150 keV that approximates well the support of  $\iota(\varepsilon)$ , see Fig. 2(b). The Radon transform matrix  $\Phi$  and its transpose  $\Phi^T$  are constructed using nonuniform Fast Fourier Transform (NUFFT) [19] with the full circular mask [20], see also [2, Sec. 3.3] which describes the construction of the Radon transform and adjoint operators.

We compare the standard filtered backprojection (FBP) method using the ramp filter [21, Sec. 3.4.7] (applied to the log-scale measurements, as is done in practice [2, Sec. 4.1]) and the proposed reconstruction obtained upon convergence of the iteration (16)–(17) (labeled MAC), respectively.

We initialize the MAC iteration with the FBP reconstruction  $\hat{\alpha}_{\text{FBP}}$ :  $\alpha^{(0)} = \hat{\alpha}_{\text{FBP}}$  and denote by  $\alpha^{(+\infty)}$  the MAC reconstruction obtained after 10000 iterations. Simultaneously multiplying the reconstruction  $\alpha$  by  $q$  or  $1/q$  and shifting the entries of the corresponding energy parameter vectors  $\mathcal{I}$  by one element to the left or right, respectively, will lead to the same response function  $f(\theta)$ , see (9). To ensure that the main lobe of the final energy parameter vector estimate  $\mathcal{I}^{(+\infty)}$  in its center, we set all but one element of the initial  $\mathcal{I}^{(0)}$  to zero and select the nonzero element in the middle:  $\mathcal{I}_{\lceil J/2 \rceil}^{(0)} = 1$ ; consequently,  $f(\theta^{(0)}) = \Phi \alpha^{(0)} \mu_{\lceil J/2 \rceil}$ . As discussed in Section 2.1, the value of  $\mu_0$  can be arbitrary and we select it so that  $\mu_{\lceil J/2 \rceil} = 1$  and thus our initial response function  $f(\theta^{(0)}) = \Phi \alpha^{(0)} = \Phi \hat{\alpha}_{\text{FBP}}$  is identical to that of the standard FBP reconstruction. We selected  $J = 17$  discretization points  $\{\mu_j\}_{j=1}^J$  spanning the range  $\mu_J/\mu_1 = 10^3$  using (8) with the



**Fig. 5:** The residual histograms from (a) the FBP and (b) MAC reconstructions.

common ratio  $q = (\mu_J/\mu_1)^{1/(J-1)}$ . Since the range of ‘true’  $\mu$  used to generate the sinogram is from  $0.19 \text{ cm}^2/\text{g}$  to  $26 \text{ cm}^2/\text{g}$  [see Fig. 2(b)], our selection  $\mu_J/\mu_1 = 10^3$  is sufficiently wide to cover the range of significant  $\mu$ s. Further, we choose  $\mathcal{I}_{\text{MAX}}^n$  according to (12), let  $\nu = 1$  and decrease value of  $t$  from 1 as well.

Figs. 3(a) and 3(b) show the standard FBP and MAC reconstructions, respectively. Since FBP does not account for the beam hardening effect, its reconstruction exhibits the cupping and streaking artifacts commonly associated with the beam-hardening phenomenon [22]: the FBP reconstruction in Fig. 3(a) shows decreasing material density towards the center of the inspected object and existence of nonzero object density in the ‘bay area’ of the object where the true density is zero.

Figs. 4(a) and 4(b) show the 500th and 700th rows of the true image in Fig. 2(a) and the FBP and MAC reconstructions. Note that the 500th and 700th rows cut through the ‘bay area’ and the region with inclusions, respectively. Recall that the MAC reconstructions can be determined only up to a scaling factor, which explains the mismatch between the MAC reconstructed and true high-signal levels in Fig. 4.

Figs. 5(a) and 5(b) show the histograms of the residuals  $z - \Phi \hat{\alpha}_{\text{FBP}}$  and  $z - f(\theta^{(+\infty)})$  for the FBP and MAC reconstructions. In Fig. 5(a), the peaks in the histogram around zero with a valley in between, which indicates the disagreement between the linear monochromatic measurement model employed by FBP and the measurements. In contrast, the histogram in Fig. 5(b) is symmetric around zero and close to the Gaussian distribution.

## 4. CONCLUSION

Further research will include

- developing a sparse signal reconstruction method based on the proposed beam-hardening correction scheme [e.g., by adding a sparsifying penalty term to the objective function (15a) or by adding a hard-thresholding step to the iteration step (i) in (16), along the lines of [23]] and demonstrating that imposing signal sparsity and other signal constraints (e.g., the geometric shape of the inspected object) will have a denoising effect on the reconstructed signals, allowing us to handle limited-angle projections or significantly undersampled measurements;
- iteratively refining the selection of the mass attenuation discretization points based on the obtained estimates of the incident energy density parameters  $\mathcal{I}$ ;
- generalizing the proposed MAC discretization to handle multiple materials.

## References

- [1] *Bremsstrahlung*, accessed 27-November-2012. [Online]. Available: <http://en.wikipedia.org/wiki/Bremsstrahlung>.
- [2] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. New York: IEEE Press, 1988.
- [3] J. Hsieh, *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*, 2nd ed. Bellingham, WA: SPIE, 2009.
- [4] J. H. Hubbell and S. M. Seltzer, "Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients 1 keV to 20 MeV for elements  $Z = 1$  to 92 and 48 additional substances of dosimetric interest," National Inst. Standards and Technology, Ionizing Radiation Div., Gaithersburg, MD, Tech. Rep. NISTIR 5632, 1995.
- [5] I. A. Elbakri and J. A. Fessler, "Statistical image reconstruction for polyenergetic X-ray computed tomography," *IEEE Trans. Med. Imag.*, vol. 21, no. 2, pp. 89–99, 2002.
- [6] I. A. Elbakri and J. A. Fessler, "Segmentation-free statistical image reconstruction for polyenergetic X-ray computed tomography with experimental validation," *Phys. Med. Biol.*, vol. 48, no. 15, pp. 2453–2477, 2003.
- [7] B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens, "An iterative maximum-likelihood polychromatic algorithm for CT," *IEEE Trans. Med. Imag.*, vol. 20, no. 10, pp. 999–1008, 2001.
- [8] J. Williamson, B. Whiting, J. Benac, R. Murphy, G. Blaine, J. O'Sullivan, D. Politte, and D. Snyder, "Prospects for quantitative computed tomography imaging in the presence of foreign metal bodies using statistical image reconstruction," *Med. Phys.*, vol. 29, p. 2404, 2002.
- [9] J. A. O'Sullivan and J. Benac, "Alternating minimization algorithms for transmission tomography," *IEEE Trans. Med. Imag.*, vol. 26, no. 3, pp. 283–297, Mar. 2007.
- [10] G. Van Gompel, K. Van Slambrouck, M. Defrise, K. Batenburg, J. de Mey, J. Sijbers, and J. Nuyts, "Iterative correction of beam hardening artifacts in CT," *Med. Phys.*, vol. 38, S36–S49, 2011.
- [11] W. Huda, *Review of Radiologic Physics*, 3rd ed. Baltimore, MD: Lippincott Williams & Wilkins, 2010.
- [12] P. McCullagh and J. Nelder, *Generalized Linear Models*, 2nd ed. New York: Chapman and Hall, 1989.
- [13] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge University Press, 2004.
- [14] J. R. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain," Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep. CMU-CS-94-125, 1994. [Online]. Available: <http://www.cs.cmu.edu/~jrs/jrspapers.html>.
- [15] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. New York: Cambridge University Press, 2007.
- [16] J. C. Gilbert and J. Nocedal, "Global convergence properties of conjugate gradient methods for optimization," *SIAM J. Optim.*, vol. 2, no. 1, pp. 21–42, 1992.
- [17] K. Qiu and A. Dogandžić, "Sparse signal reconstruction via ECME hard thresholding," *IEEE Trans. Signal Process.*, vol. 60, pp. 4551–4569, Sep. 2012.
- [18] G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed. Pacific Grove, CA: Duxbury Press, 2002.
- [19] J. Fessler and B. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 560–574, Feb. 2003.
- [20] A. Dogandžić, R. Gu, and K. Qiu, "Mask iterative hard thresholding algorithms for sparse image reconstruction of objects with known contour," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 2011, pp. 2111–2116.
- [21] J. A. Fessler, "Analytical tomographic image reconstruction methods," Nov. 2009, [Online]. Available: <http://web.eecs.umich.edu/~fessler/book/c-tomo-prop.pdf>.
- [22] J. Barrett and N. Keat, "Artifacts in CT recognition and avoidance," *Radiographics*, vol. 24, no. 6, pp. 1679–1691, 2004.
- [23] T. Blumensath, "Compressed sensing with nonlinear observations and related nonlinear optimisation problems," May 2012. arXiv: 1205.1650 [cs.LG].