Sparse X-ray CT Image Reconstruction and Blind 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 nonlinear sparse X-ray computed tomography (CT) image reconstruction method that accounts for beam hardening effects due to polychromatic X-ray sources. We adopt the blind scenario where the material of the inspected object and the incident polychromatic source spectrum are *unknown* and apply mass attenuation discretization of the underlying integral expressions that model the noiseless measurements. Our reconstruction algorithm employs constrained minimization of a penalized least-squares cost function, where nonnegativity and maximum-energy constraints are imposed on incident spectrum parameters and negative-energy and smooth ℓ_1-norm penalty terms are introduced to ensure the nonnegativity and sparsity of the density map image. This minimization scheme alternates between a nonlinear conjugate-gradient step for estimating the density map image and an active set step for estimating incident spectrum parameters. We compare the proposed method with the existing approaches, which ignore the polychromatic nature of the measurements or sparsity of the density map image.**

I. Introduction

In nondestructive evaluation (NDE) and medical imaging, there is a need to correct beam hardening effects, which occur because of the polychromatic nature of the X-rays generated by vacuum tubes [1, 2]. Photons at lower energies have larger attenuation rates, thus 'hardening' the X-ray beam as it travels. Traditional linear-model based reconstruction schemes ignore this effect and, consequently, exhibit beam hardening (e.g., cupping and streaking) artifacts [3, 4]. Beam hardening correction methods can be categorized into pre-filtering, linearization, dual-energy, and post-reconstruction approaches [5]. Reconstruction methods have recently been developed in [6–8] that aim to optimize nonlinear objective functions based on the underlying physical model; [6, 7] assume known incident polychromatic source spectrum and imaged materials, whereas [8] considers a blind scenario with *unknown* incident spectrum and imaged materials, but employs an excessive number of parameters and suffers from numerical instability [9, Sec. 3.2.1][1].

In [10], we proposed a parsimonious measurement model parametrization for the blind scenario with *unknown* incident spectrum and a single *unknown* material; this parametrization is based on discretizing the Beer's law over mass attenuation rather than photon energy, thereby reducing the number of functions to estimate from two to one: compare (5) with (2) in [10]. In this paper, we present a nonlinear X-ray computed tomography (CT) reconstruction algorithm that performs constrained minimization of a penalized least-squares (LS) cost function based on the measurement model in [10], where nonnegativity and maximum-energy constraints are imposed on the incident spectrum parameters and additive penalty terms are introduced to ensure nonnegativity and approximate sparsity of the density map image that we wish to reconstruct. To the best of our knowledge, this is the first physical-model based method for *simultaneous* sparse image reconstruction *and* beam hardening correction.

We introduce the notation: I_N, $\mathbf{1}_{N\times 1}$, and $\mathbf{0}_{N\times 1}$ are the identity matrix of size N and the $N\times 1$ vectors of ones and zeros, respectively; $|\cdot|$, $\|\cdot\|_2$, and "^{T}" are the absolute value, Euclidean norm, and transpose, respectively. Furthermore, $(x)_+ = \max\{x, 0\}$ is the positive-part operator, $\boldsymbol{a} \succeq \boldsymbol{b}$ ($\boldsymbol{a} \succ \boldsymbol{b}$) denotes that all elements of $\boldsymbol{a} - \boldsymbol{b}$ are nonnegative (positive), and $\mathbb{1}_A(y) = \begin{cases} 1, & y \in A, \\ 0, & \text{otherwise} \end{cases}$ denotes the indicator function. Define the elementwise logarithm, and positive-part operators of an N-dimensional real vector $\boldsymbol{x} = [x_1, \ldots, x_N]^T$ as $\ln_\circ(\boldsymbol{x}) = [\ln x_1, \ldots, \ln x_N]^T$, and $(\boldsymbol{x})_+ = [(x_1)_+, \ldots, (x_N)_+]^T$.

II. Beam Hardening Correction Model

An X-ray CT scan consists of multiple projections with the beam intensity measured by multiple detectors. We discretize the continuous measurement model over space and mass attenuation using p pixels and J mass attenuation bins (see [10, eq. (5)]) and model the $N\times 1$ vector of log-scale measurements at all detectors and projections as [10]

$$\boldsymbol{z} = [-\ln \mathcal{E}_1, \ldots, -\ln \mathcal{E}_N]^T = \boldsymbol{f}(\boldsymbol{\theta}) + \boldsymbol{n} = -\ln_\circ[A(\boldsymbol{\alpha})\boldsymbol{\mathcal{I}}] + \boldsymbol{n} \quad (1)$$

where $\{\mathcal{E}_i\}_{i=1}^N$ are the noisy energy measurements, the (i,j)th element of the $N\times J$ matrix $A(\boldsymbol{\alpha})$ is

$$A_{i,j}(\boldsymbol{\alpha}) = \exp(-\Phi_{(i)}^T \boldsymbol{\alpha} \mu_j), \quad (2)$$

and

- $\boldsymbol{\alpha}$ is an unknown $p\times 1$ vector representing the two-dimensional (2D) density map image that we wish to reconstruct,
- $\boldsymbol{\mathcal{I}} = [\mathcal{I}_1, \ldots, \mathcal{I}_J]^T$ is the $J\times 1$ vector of unknown incident spectrum parameters describing the discretized incident X-ray spectrum as a function of mass attenuation (see [10, eq. (7) and Fig. 1]),
- $\Phi_{(i)}$ is a known $p\times 1$ vector of weights quantifying how much each pixel of $\boldsymbol{\alpha}$ contributes to the X-ray attenuation on the straight-line path corresponding to the ith measurement,
- $\mu_0 < \mu_1 < \cdots < \mu_J$ are known mass attenuation discretization points, and
- $\boldsymbol{n}$ is zero-mean additive white noise.

Our goal is to estimate the image and incident energy density parameters

$$\boldsymbol{\theta} = \big(\boldsymbol{\alpha}, \boldsymbol{\mathcal{I}}\big). \quad (3)$$

Note that $\Phi = [\Phi_{(1)}\Phi_{(2)}\cdots\Phi_{(N)}]^T$ is the $N\times p$ Radon transform matrix for our imaging system. For Gaussian $\boldsymbol{n}$, noisy energy measurements $\{\mathcal{E}_i\}_{i=1}^N$ follow the lognormal distribution, see (1).

Constraints on density map image $\boldsymbol{\alpha}$. The nonnegativity of the density map that we wish to reconstruct yields the following constraint:

$$\boldsymbol{\alpha} \succeq \mathbf{0}_{p\times 1}. \quad (4)$$

We also assume that the image $\boldsymbol{\alpha}$ is approximately sparse in the discrete wavelet transform (DWT) domain, i.e., $\Psi^T\boldsymbol{\alpha}$ is an approximately sparse vector, where

$$\Psi = \begin{bmatrix} \boldsymbol{\psi}_1 & \boldsymbol{\psi}_2 & \ldots & \boldsymbol{\psi}_p \end{bmatrix} \quad (5)$$

is a $p\times p$ inverse DWT orthogonal matrix satisfying $\Psi^T\Psi = \Psi\Psi^T = I_p$. Note that $\boldsymbol{\alpha}$, $\boldsymbol{f}(\boldsymbol{\theta})$, and Ψ can be easily adjusted to accommodate the prior geometric contour information of the inspected object [11].

Constraints on incident spectrum $\boldsymbol{\mathcal{I}}$. We assume that the shadow of the inspected object is completely covered by the receiver array and

[1] Alarfaj was not successful implementing the blind method in [8] and discusses some of the obstacles in [9, Sec. 3.2.1]; consequently, [9] reports a non-blind version of [8] with *known* incident spectrum and object materials.

978-1-4673-3146-3/13/$31.00 ©2013IEEE

that we know the upper bound $\mathcal{I}_{\text{MAX}}^{\text{in}}$ on incident X-ray energy (obtained, e.g., from energy measurements at the detectors that have a line-of-sight view of the X-ray source); this assumption and the nonnegativity of the incident spectrum yield the following constraints on $\boldsymbol{\mathcal{I}}$:

$$\sum_{j=1}^{J} \mathcal{I}_j = \mathbf{1}^T \boldsymbol{\mathcal{I}} \leq \mathcal{I}_{\text{MAX}}^{\text{in}}, \qquad \boldsymbol{\mathcal{I}} \succeq \mathbf{0} \tag{6}$$

which can be written as a single vector linear inequality constraint as

$$B\boldsymbol{\mathcal{I}} \succeq \mathbf{b} \tag{7a}$$

where

$$B = \begin{bmatrix} I_{J\times J} & -\frac{1}{\sqrt{J}}\mathbf{1}_{J\times 1} \end{bmatrix}^T, \qquad \boldsymbol{b} = \begin{bmatrix} \mathbf{0}_{J\times 1}^T & -\frac{1}{\sqrt{J}}\mathcal{I}_{\text{MAX}}^{\text{in}} \end{bmatrix}^T. \tag{7b}$$

We propose the following penalized LS cost function based on the measurement model (1) and the above image constraints:

$$L_{\nu,u}(\boldsymbol{\theta}) = \frac{1}{2}\|\boldsymbol{z} - \boldsymbol{f}(\boldsymbol{\theta})\|_2^2 + \frac{\nu}{2}\big\|(-\boldsymbol{\alpha})_+\big\|_2^2 + u\sum_{i=1}^{p}\sqrt{(\boldsymbol{\psi}_i^T\boldsymbol{\alpha})^2 + \xi} \tag{8}$$

where ν and u are scalar tuning constants for the signal nonnegativity and sparsity penalty terms and $\xi > 0$ is a smoothing factor in the signal sparsity penalty term, chosen in practice to be a small positive constant. The smooth ℓ_1-norm penalty approximation in the third summand makes gradient-descent methods applicable to minimizing (8) with respect to $\boldsymbol{\alpha}$, see also [12] where this approximation was introduced to impose signal sparsity in compressed sensing magnetic resonance imaging (MRI). Lustig *et al*. [12] employ a linear signal model; in contrast, our measurement model is nonlinear with additional parameters and constraints.

Interestingly, (8) is *biconvex* with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\mathcal{I}}$ in a biconvex set, see [13] for an introduction to biconvex functions and sets.

Adding the incident spectrum constraints (7a) leads to the following constrained minimization with linear inequality constraints:

$$\min_{\boldsymbol{\theta}, B\boldsymbol{\mathcal{I}}\succeq\mathbf{b}} L_{\nu,u}(\boldsymbol{\theta}) \tag{9}$$

that we wish to solve.

III. Minimization Algorithm

We now present our algorithm for performing the constrained penalized LS minimization in (9). Due to the difference in size and handling of the constraints on $\boldsymbol{\alpha}$ and $\boldsymbol{\mathcal{I}}$, we adopt two different schemes for their estimation and then combine these schemes in stepwise-descent manner.

In [10], the constraints on $\boldsymbol{\mathcal{I}}$ in (7a) were imposed by the logarithmic barrier function via interior-point method [14, Sec. 11.2]. Here, we impose the linear inequality constraints (7a) via an *active set method* [15, 16] (see also [17]), which removes barrier-function tuning constants, reduces the number of energy density parameters, and is faster than the interior-point method.

Define the gradient vectors and Hessian matrices of the cost function (8) with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\mathcal{I}}$, respectively:

$$\boldsymbol{g}_{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \frac{\partial L_{\nu,u}(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha}}, \qquad H_{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \frac{\partial^2 L_{\nu,u}(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha}\partial \boldsymbol{\alpha}^T}, \tag{10a}$$

$$\boldsymbol{g}_{\boldsymbol{\mathcal{I}}}(\boldsymbol{\theta}) = \frac{\partial L_{\nu,u}(\boldsymbol{\theta})}{\partial \boldsymbol{\mathcal{I}}} \qquad H_{\boldsymbol{\mathcal{I}}}(\boldsymbol{\theta}) = \frac{\partial^2 L_{\nu,u}(\boldsymbol{\theta})}{\partial \boldsymbol{\mathcal{I}}\partial \boldsymbol{\mathcal{I}}^T} \tag{10b}$$

where the gradient and Hessian in (10b) are not functions of the tuning constants ν and u.

For a row index set $Q \subseteq \{1, 2, \ldots, J+1\}$ of the matrix B in (7b), we denote by $\text{card}(Q)$ the cardinality of Q and by B_Q the $\text{card}(Q)\times J$ submatrix of B consisting of rows of B with indices in Q. Denote by B_q and b_q the qth row of B and the qth element of $\boldsymbol{b}$, respectively. Define the $J \times (J - \text{card}(Q))$ full rank matrix M_Q whose columns span the null space of B_Q, i.e., $B_Q M_Q = \mathbf{0}$.

We descend (8) by alternating between 1) and 2):

1) the *nonlinear conjugate-gradient step* for $\boldsymbol{\alpha}$ [18, Sec. 14.1]:

$$\boldsymbol{\alpha}^{(i+1)} = \boldsymbol{\alpha}^{(i)} - s_{\boldsymbol{\alpha}} \frac{\boldsymbol{g}_{\boldsymbol{\alpha}}^T(\boldsymbol{\theta}^{(i)})\boldsymbol{d}^{(i)}}{\boldsymbol{d}^{(i)T} H_{\boldsymbol{\alpha}}(\boldsymbol{\theta}^{(i)})\boldsymbol{d}^{(i)}}\boldsymbol{d}^{(i)} \tag{11a}$$

where

$$\boldsymbol{\theta}^{(i)} = (\boldsymbol{\alpha}^{(i)}, \boldsymbol{\mathcal{I}}^{(i)}) \tag{11b}$$

$$\boldsymbol{e}^{(i)} = \boldsymbol{g}_{\boldsymbol{\alpha}}(\boldsymbol{\theta}^{(i)}) - \boldsymbol{g}_{\boldsymbol{\alpha}}(\boldsymbol{\theta}^{(i-1)}) \tag{11c}$$

$$\beta^{(i)} = \max\left\{0, \ \frac{\boldsymbol{g}_{\boldsymbol{\alpha}}^T(\boldsymbol{\theta}^{(i)})\boldsymbol{e}^{(i)}}{\|\boldsymbol{g}_{\boldsymbol{\alpha}}(\boldsymbol{\theta}^{(i-1)})\|_2^2}\right\} \tag{11d}$$

$$\boldsymbol{d}^{(i)} = \boldsymbol{g}_{\boldsymbol{\alpha}}(\boldsymbol{\theta}^{(i)}) + \beta^{(i)}\boldsymbol{d}^{(i-1)} \tag{11e}$$

where $0 < s_{\boldsymbol{\alpha}} \leq 1$ is the step size for updating $\boldsymbol{\alpha}$ determined via backtracking [19, Sec. 9.7] to guarantee the descent of (8) and i denotes the iteration index;

2) the *gradient projection active-set step* for $\boldsymbol{\mathcal{I}}$, where $\boldsymbol{\alpha}$ is fixed and set to $\boldsymbol{\alpha}^{(i+1)}$ obtained in 1) (see also [16] and [15, Ch. 5.2]):

a) (initialize) assign $\widetilde{\boldsymbol{\theta}}^{(i)} \leftarrow (\boldsymbol{\alpha}^{(i+1)}, \boldsymbol{\mathcal{I}}^{(i)})$ and $Q \leftarrow Q^{(i)}$, where Q denotes the *active set*, i.e., the row index set of B with the *equality* constraints, $B_Q\boldsymbol{\mathcal{I}} = \boldsymbol{b}_Q$, hold,

b) (compute the steepest descent step within the active set)

$$\boldsymbol{p} = -M_Q\Big(M_Q^T H_{\boldsymbol{\mathcal{I}}}(\widetilde{\boldsymbol{\theta}}^{(i)})M_Q\Big)^{-1}M_Q^T\boldsymbol{g}_{\boldsymbol{\mathcal{I}}}(\widetilde{\boldsymbol{\theta}}^{(i)}) \tag{12a}$$

where $\boldsymbol{p}$ is the steepest descent step of the quadratic approximation of our objective function (8) with respect to $\boldsymbol{\mathcal{I}}$, subject to the active-set constraint in the null space of B_Q,

c) (compute Lagrange multipliers)

$$\widetilde{\boldsymbol{g}}_{\boldsymbol{\mathcal{I}}} = \boldsymbol{g}_{\boldsymbol{\mathcal{I}}}(\widetilde{\boldsymbol{\theta}}^{(i)}) + H_{\boldsymbol{\mathcal{I}}}(\widetilde{\boldsymbol{\theta}}^{(i)})\boldsymbol{p} \tag{12b}$$

$$\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_{\text{card}(Q)}]^T = (B_Q B_Q^T)^{-1}B_Q\widetilde{\boldsymbol{g}}_{\boldsymbol{\mathcal{I}}} \tag{12c}$$

where $\boldsymbol{\lambda}$ is a *Lagrange multiplier* vector,

d) (prune active set) if there exists an index $k \in Q$ whose corresponding Lagrange multiplier is negative (i.e., $\lambda_{q_k} < 0$, where q_k maps the element k in Q to the corresponding row index within the matrix B_Q), then

- $Q \leftarrow Q \setminus k$,
- go back to step 2b,

otherwise continue to 2e,

e) (compute the largest possible step size for updating $\boldsymbol{\mathcal{I}}$)

$$s_{\boldsymbol{\mathcal{I}},\text{MAX}} = \arg\max_{B(\boldsymbol{\mathcal{I}}^{(i)} + s\,\boldsymbol{p})\succeq\boldsymbol{b}} s \tag{12d}$$

which is the largest step size for updating $\boldsymbol{\mathcal{I}}$ that satisfies the constraints in (7a),

f) (compute the step size for updating $\boldsymbol{\mathcal{I}}$) determine $s_{\boldsymbol{\mathcal{I}}}$ via backtracking, initialized with $\min\{1, s_{\boldsymbol{\mathcal{I}},\text{MAX}}\}$,

g) (grow active set) if $s_{\boldsymbol{\mathcal{I}}} = s_{\boldsymbol{\mathcal{I}},\text{MAX}}$ then

- find index q of the newly activated constraint, i.e., $q \in \{1, 2, \ldots, J+1\} \setminus Q$ such that $B_q\big(\boldsymbol{\mathcal{I}}^{(i)} + s_{\boldsymbol{\mathcal{I}},\text{MAX}}\boldsymbol{p}\big) = b_q$,
- update the active set accordingly: $Q \leftarrow Q \cup \{q\}$.

h) update the iterate of $\boldsymbol{\mathcal{I}}$ and its active set:

$$\boldsymbol{\mathcal{I}}^{(i+1)} \leftarrow \boldsymbol{\mathcal{I}}^{(i)} + s_{\boldsymbol{\mathcal{I}}}\,\boldsymbol{p} \tag{12e}$$

$$Q^{(i+1)} \leftarrow Q. \tag{12f}$$

Here, 1) aims at unconstrained minimization of the penalized LS cost function in (8) with respect to $\boldsymbol{\alpha}$ whereas 2) aims at constrained minimization of the LS cost function with respect to $\boldsymbol{\mathcal{I}}$. [The constraints (7a) do not involve $\boldsymbol{\alpha}$ and penalty terms in (8) are not functions of $\boldsymbol{\mathcal{I}}$.]

In step 1), we seek an update of $\boldsymbol{\alpha}$ that reduces $L_{\nu,u}((\boldsymbol{\alpha}, \boldsymbol{\mathcal{I}}))$ and apply the nonlinear conjugate gradient method to achieve this goal. In the nonlinear conjugate gradient step (11), we use the Polak-Ribière formula (11d), which restarts the conjugate gradient iteration

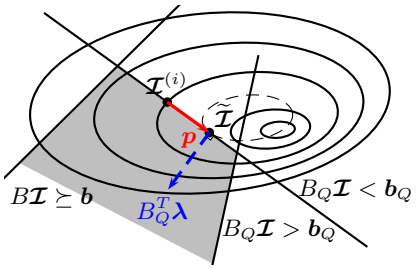
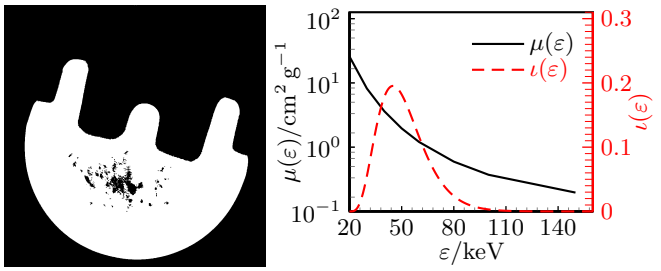


Figure 1: Illustration of the gradient projection active set step.



(a) (b)

Figure 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.

by forgetting the past search directions when $\boldsymbol{g}_{\boldsymbol{\alpha}}^T(\boldsymbol{\theta}^{(i)})\boldsymbol{e}^{(i)}$ is negative [20] to keep $\boldsymbol{\alpha}$ from going uphill. Since $H_{\boldsymbol{\alpha}}(\boldsymbol{\theta})$ has the form

$$H_{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \Phi^T W(\boldsymbol{\theta})\Phi + \nu\,\mathrm{diag}\left(\mathbb{1}_{(-\infty,0)}(\boldsymbol{\alpha})\right) + u\Psi U(\boldsymbol{\alpha})\Psi^T \quad (13)$$

where both $W(\boldsymbol{\theta})$ and $U(\boldsymbol{\alpha})$ are diagonal matrices, the denominator in (11a) is computed efficiently by computing and storing $\Phi\boldsymbol{d}^{(i)}$ and $\Psi^T\boldsymbol{d}^{(i)}$.

In step 2), we seek an update of $\boldsymbol{\mathcal{I}}$ that reduces $L_{\nu,u}((\boldsymbol{\alpha},\boldsymbol{\mathcal{I}}))$ subject to the inequality constraints (7a). To achieve this goal, we apply the gradient-projection active set method, which searches for *active restrictions* (defined as restrictions that are fulfilled with equality) and meanwhile approximates the quadratic steepest descent step for the simplified problem under these equality constraints.

Fig. 1 illustrates the gradient projection active-set step in 2). The plane is the space represented by the elements of vector $\boldsymbol{\mathcal{I}}$, where the gray shadow is defined by the constraints, $B\boldsymbol{\mathcal{I}} \succeq \boldsymbol{b}$. The solid- and dashed-line ellipses depict contour graphs of the objective function $L_{\nu,u}((\boldsymbol{\alpha},\boldsymbol{\mathcal{I}}))$ and its quadratic approximation around $L_{\nu,u}((\boldsymbol{\alpha},\boldsymbol{\mathcal{I}}^{(i)}))$, respectively. For the current active set is Q, $\widetilde{\boldsymbol{\mathcal{I}}} = \boldsymbol{\mathcal{I}}^{(i)} + \boldsymbol{p}$ minimizes the quadratic approximation under the active constraints and, consequently, the expected gradient at $\widetilde{\boldsymbol{\mathcal{I}}}$, $B_Q^T\boldsymbol{\lambda}$, is perpendicular to $\boldsymbol{p}$. The active set Q chosen by step 2d) guarantees that $B_Q^T\boldsymbol{\lambda}$ ultimately points in the correct direction, the same side with $B_Q\boldsymbol{\mathcal{I}} \succ \boldsymbol{b}_Q$. In this particular example, Q remains unchanged and the largest possible step size $s_{\boldsymbol{\mathcal{I}},\mathrm{MAX}}$ is greater than 1 because $\widetilde{\boldsymbol{\mathcal{I}}}$ does not go beyond the right boundary of the shadow area. Consequently, the backtracking search for $s_{\boldsymbol{\mathcal{I}}}$ in step 2f) is initialized by 1. In other cases, when $\widetilde{\boldsymbol{\mathcal{I}}}$ goes beyond the right boundary, one more constraint will be added to the active set; when $B_Q^T\boldsymbol{\lambda}$ points to the opposite direction, the current active constraint will be deactivated.

The increment vector $\boldsymbol{p}$ is determined by projecting the gradient $\boldsymbol{g}_{\boldsymbol{\mathcal{I}}}(\widetilde{\boldsymbol{\theta}}^{(i)})$ onto the constraints' null space (12a) in the metric of $H_{\boldsymbol{\mathcal{I}}}(\widetilde{\boldsymbol{\theta}}^{(i)})$, which is equivalent to the update step in Jamshidian's gradient projection (GP) active set algorithm. This step is also similar to that in the *constrained scoring algorithm* [17], where Fisher information matrix for $\boldsymbol{\mathcal{I}}$ is used instead of the Hessian matrix, see (12a).

Note that the steepest descent vector $\boldsymbol{p}$ works very well as an increment candidate (or is a good increment candidate) when the quadratic approximation fits the objective (8). In practice, the step size $s_{\boldsymbol{\mathcal{I}}}$ is usually restricted by the linear constraints (7a) and backtracking line search procedure.

IV. Numerical examples

We construct a simulation example based on a binary 1024×1024 image in Fig. 2a that corresponds to a real X-ray CT reconstruction of a metal casting, obtained by thresholding the pixel values of a reconstruction in [21, Fig. 5(b)]. The inspected object, assumed to be made of iron, contains irregularly shaped inclusions. We simulated the polychromatic sinogram using the mass attenuation of iron and incident spectrum in Fig. 2b and constructed the Radon transform and adjoint matrices Φ and Φ^T using nonuniform Fast Fourier Transform (NUFFT) [22] with the circular mask [11]. We compare

- the standard filtered backprojection (FBP) method,
- the proposed reconstruction obtained upon convergence of the iteration (11)–(12) for solving the constrained penalized LS (CPLS) problem (9) (labeled CPLS), and
- our method in [10] (labeled MAC), which uses the mass attenuation coefficient (MAC) discretization to correct for beam hardening but ignores signal sparsity, see also the discussion in Section III.

To initialize the CPLS and MAC iterations, we chose the same $\boldsymbol{\alpha}$ and $\boldsymbol{\mathcal{I}}$ that we employed to initialize the MAC method in [10, Sec. 3]. The value of μ_0 can be an arbitrary positive number and we selected the remaining $J = 17$ mass attenuation discretization points $\{\mu_j\}_{j=1}^J$ using a geometric sequence $\mu_j = \mu_0 q^j$ that spans the range $\mu_J/\mu_1 = 10^3$ with the common ratio $q = (\mu_J/\mu_1)^{1/(J-1)}$; this selection is sufficiently wide to cover the range of significant μs, see Fig. 2b.

Further, we choose $\mathcal{I}_{\mathrm{MAX}}^{\mathrm{in}} = 1$ based on the (normalized) energy measurements at the detectors that have a line-of-sight view of the X-ray source (see [10, eq. (12)]) and set the nonnegativity and sparsity tuning constants to

$$\nu = 1, \qquad u = 1.5 \times 10^{-4} \quad (14)$$

which yield good reconstruction performance. The nonnegativity and logarithmic barrier tuning parameters for the MAC method have also been chosen for good performance and set as in [10]. We run 2000 iterations for both CPLS and MAC methods.

We use the relative square error (RSE)

$$\mathrm{RSE} = 1 - \left(\frac{\widehat{\boldsymbol{\alpha}}^T\boldsymbol{\alpha}_{\mathrm{true}}}{\|\widehat{\boldsymbol{\alpha}}\|_2\|\boldsymbol{\alpha}_{\mathrm{true}}\|_2}\right)^2 \quad (15)$$

as the reconstruction performance metric. Note that (15) is invariant to scaling of $\widehat{\boldsymbol{\alpha}}$ and that the blind CPLS and MAC reconstructions can be determined only up to a scaling factor because the (arbitrary) value of the discretization point μ_0 can be absorbed by $\boldsymbol{\alpha}$, see (2).

Figs. 3a–3c show the reconstructions and corresponding RSEs achieved by the three methods. Figs. 3d and 3e show the 500th and 700th rows of the reconstructed images and the true image in Fig. 2a. Note that the 500th row cuts through the 'bay area' whereas the 700th row cuts through the region with inclusions. The MAC and CPLS reconstructions can be determined only up to a scaling factor, which explains the mismatch between the reconstructed and true low-signal levels in Figs. 3d and 3e.

Since FBP does not account for the beam hardening effect, it achieves a fairly high RSE due to the cupping and streaking artifacts [3, 4]: 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. Both CPLS and MAC correct the beam

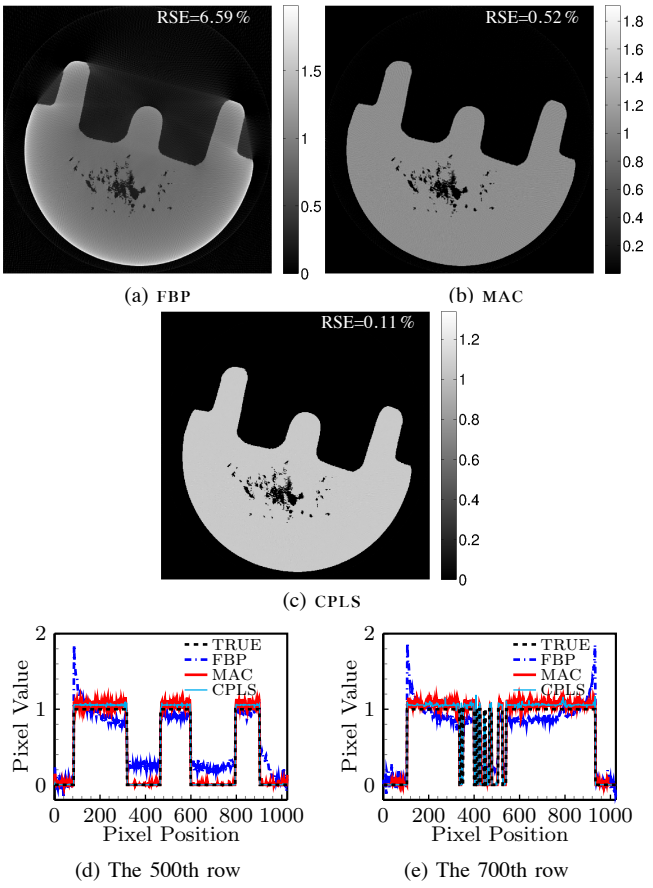


Figure 3: The (a) FBP, (b) MAC and (c) CPLS reconstructions of full 180-projection data with their (d) 500-th and (e) 700-th row profiles compared to that of true image.

hardening effect successfully and achieve RSEs that are an order of magnitude smaller than FBP. Imposing approximate signal sparsity has a denoising effect on the resulting reconstruction and significantly reduces the corresponding RSE, compare Figs. 3b and 3c and the row profiles of CPLS and MAC in Figs. 3d and 3e.

V. Conclusion

Further research will include analyzing the convergence of the reconstruction algorithm in Section III, e.g., establishing conditions under which it converges to a local maximum of (9); we expect that the biconvexity of (9) will be useful for such analysis. We will also generalize the proposed mass attenuation discretization and sparse signal reconstruction methods to handle multiple materials.

Acknowledgment

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

References

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

[2] J. Hsieh, *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*, 2nd ed. Bellingham, WA: SPIE, 2009.

[3] J. Barrett and N. Keat, "Artifacts in CT recognition and avoidance," *Radiographics*, vol. 24, no. 6, pp. 1679–1691, 2004.

[4] J. Nuyts *et al*., "Modelling the physics in the iterative reconstruction for transmission computed tomography," *Phys. Med. Biol*., vol. 58, no. 12, R63–R96, 2013.

[5] M. Krumm, S. Kasperl, and M. Franz, "Reducing non-linear artifacts of multi-material objects in industrial 3D computed tomography," *NDT & E Int*., vol. 41, no. 4, pp. 242–251, 2008.

[6] 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.

[7] ——, "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.

[8] G. Van Gompel *et al*., "Iterative correction of beam hardening artifacts in CT," *Med. Phys*., vol. 38, S36–S49, 2011.

[9] M. Alarfaj, "A comparative study of two methods for the correction of beam hardening artifacts in X-ray computed tomography," Master's thesis, Department of Mathematics, Oregon State University, Corvallis, OR, Dec. 2012.

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

[11] A. Dogandžić, R. Gu, and K. Qiu, "Mask iterative hard thresholding algorithms for sparse image reconstruction of objects with known contour," *Proc. Asilomar Conf. Signals, Syst. Comput*., Pacific Grove, CA, Nov. 2011, pp. 2111–2116.

[12] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[13] J. Gorski, F. Pfeuffer, and K. Klamroth, "Biconvex sets and optimization with biconvex functions: a survey and extensions," *Math. Methods Oper. Res*., vol. 66, no. 3, pp. 373–407, 2007.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge University Press, 2004.

[15] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*. San Diego, CA: Academic Press, 1981.

[16] M. Jamshidian, "On algorithms for restricted maximum likelihood estimation," *Comput. Stat. Data. Anal*., vol. 45, no. 2, pp. 137–157, 2004.

[17] T. Moore, B. Sadler, and R. Kozick, "Maximum-likelihood estimation, the Cramér–Rao bound, and the method of scoring with parameter constraints," *IEEE Trans. Signal Process*., vol. 56, no. 3, pp. 895–908, 2008.

[18] 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.

[19] W. H. Press *et al*., *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. New York: Cambridge University Press, 2007.

[20] 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.

[21] K. Qiu and A. Dogandžić, "Sparse signal reconstruction via ECME hard thresholding," *IEEE Trans. Signal Process*., vol. 60, pp. 4551–4569, Sep. 2012.

[22] J. A. Fessler and B. P. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Trans. Signal Process*., vol. 51, no. 2, pp. 560–574, Feb. 2003.