Optimization-based algorithm for solving the discrete x-ray transform with nonlinear partial volume effect

Buxin Chen,^a Xin Liu,^a Zheng Zhang,^a Dan Xia,^a Emil Y. Sidky,^a
and Xiaochuan Pan^{a,b,*}

^aUniversity of Chicago, Department of Radiology, Chicago, Illinois, United States

^bUniversity of Chicago, Department of Radiation and Cellular Oncology, Chicago, Illinois,
United States

Abstract

Purpose: Inverting the discrete x-ray transform (DXT) with the nonlinear partial volume (NLPV) effect, which we refer to as the NLPV DXT, remains of theoretical and practical interest. We propose an optimization-based algorithm for accurately and directly inverting the NLPV DXT.

Methods: Formulating the inversion of the NLPV DXT as a nonconvex optimization program, we propose an iterative algorithm, referred to as the nonconvex primal-dual (NCPD) algorithm, to solve the problem. We obtain the NCPD algorithm by modifying a first-order primal-dual algorithm to address the nonconvex optimization. Subsequently, we perform quantitative studies to verify and characterize the NCPD algorithm.

Results: In addition to proposing the NCPD algorithm, we perform numerical studies to verify that the NCPD algorithm can reach the devised numerically necessary convergence conditions and, under the study conditions considered, invert the NLPV DXT by yielding numerically accurate image reconstruction.

Conclusion: We have developed and verified with numerical studies the NCPD algorithm for accurate inversion of the NLPV DXT. The study and results may yield insights into the effective compensation for the NLPV artifacts in CT imaging and into the algorithm development for nonconvex optimization programs in CT and other tomographic imaging technologies.

© 2020 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.7.5.053502](https://doi.org/10.1117/1.JMI.7.5.053502)]

Keywords: nonlinear partial volume; nonconvex optimization; CT reconstruction.

Paper 20096R received Apr. 22, 2020; accepted for publication Sep. 2, 2020; published online Oct. 6, 2020.

1 Introduction

In CT imaging, a detector bin and/or x-ray-source focal spot of finite sizes sample a volume element within an imaged subject. The nonuniformity of the x-ray linear attenuation coefficient within the sampled volume element can result in the so-called nonlinear partial volume (NLPV) effect.^{1–5} Unlike the standard x-ray transform that is a linear function of the attenuation coefficient map, i.e., the image of interest, the x-ray transform with the NLPV effect becomes a nonlinear function of the image. When an algorithm developed for the former is applied to the latter, it may engender NLPV artifacts obscuring low-contrast anatomies of interest in reconstructed images.^{6–8} Effort has been devoted to developing algorithms for empirically correcting the NLPV artifacts.^{1–5,9–11}

In this work, we investigate and develop an optimization-based algorithm to accurately invert the discrete x-ray transform (DXT) with the NLPV effect numerically, which we refer to as the NLPV DXT, as such an investigation remains of theoretical and practical relevance. We first

*Address all correspondence to Xiaochuan Pan, E-mail: xpan@uchicago.edu

formulate the inverse problem as an optimization program in which the image of interest is designed as a solution of the program. The optimization program can be nonconvex in the presence of the NLPV effect, and there is a lack of algorithms for accurately solving the nonconvex optimization program. The focus of this work is thus to develop an algorithm that, under study conditions of practical relevance, can accurately invert the NLPV DXT numerically through solving the nonconvex optimization program.

We propose an iterative algorithm, referred to as the nonconvex primal-dual (NCPD) algorithm, for the nonconvex optimization program by modifying the first-order primal-dual (PD) algorithm,^{12,13} developed previously for convex optimization programs. We subsequently perform quantitative verification and characterization studies on the NCPD algorithm developed. In the verification study, we establish that the NCPD algorithm can, under the study conditions considered, invert the NLPV DXT by reconstructing numerically accurate images, whereas in the initial characterization study, we visualize the reconstruction property of the NCPD algorithm in the presence of data noise.

This work is an expansion from our previously published conference proceedings,^{14,15} with significant details added on the development of the NCPD algorithm, its convergence conditions, its adaptation to the NLPV problem, and its careful numerical verification. Convergence results of the NCPD algorithm, especially those with different computer precisions in the verification study, are being reported for the first time in this work.

2 Materials and Methods

Considering a three-dimensional (3-D) image defined on a 3-D array of identical cubic voxels, we use vector $\mathbf{f}$ of size I to denote the discrete image in a concatenated form in the order of the x , y , and z axes, with element f_i denoting the image value within voxel i , where $i \in 1, \dots, I$.

2.1 Discrete X-Ray Transform with the NLPV Effect

We assume that points L_d and L_s uniformly cover the physical areas of a detector bin and x-ray-source focal spot, as shown in Fig. 1. Total number of L rays between pair j of detector bin and focal spot thus is given by $L = L_d \times L_s$. Let the two-dimensional (2-D) detector at projection view ϕ be specified by (u, v) , a 2-D-Cartesian coordinate system. Using vector $\mathbf{g}$ of size J to indicate discrete model data in a concatenated form in the order of u , v , and ϕ , with element $g_j^{(\text{NL})}$ in detector bin j , we write the NLPV DXT as

$$g_j^{(\text{NL})}(\mathbf{f}) = -\ln \frac{1}{L} \sum_{l_j=1}^L \exp\left(-\sum_{i=1}^I a_{l_j i} f_i\right), \quad (1)$$

where $j \in 1, \dots, J$, $l_j \in 1, \dots, L$, and $a_{l_j i}$ indicates the intersection length of ray l_j within detector bin j with voxel i .

For the purpose of algorithm development, the NLPV DXT is decomposed into a sum of linear and nonlinear terms,^{16–20} using Taylor expansion, as

$$g_j^{(\text{NL})}(\mathbf{f}) = g_j^{(\text{NL})}(\bar{\mathbf{f}}) + \nabla_{\mathbf{f}} g_j^{(\text{NL})}(\bar{\mathbf{f}})^{\top} (\mathbf{f} - \bar{\mathbf{f}}) + \Delta g_j^{(\text{NL})}(\mathbf{f}), \quad (2)$$

Fig. 1 For a pair of detector bin and x-ray-source focal spot of finite sizes, (a) one ray or (b) multiple rays are considered in the standard DXT or in the NLPV DXT. The detector plane is specified by coordinate system (u, v) .

where notation “ $\top$ ” denotes the transpose operation, $\bar{\mathbf{f}}$ is a known image of size I with element $\bar{f}_i$, vector $\nabla_{\mathbf{f}} g_j^{(\text{NL})}(\bar{\mathbf{f}})$ of size I is the gradient of $g_j^{(\text{NL})}(\mathbf{f})$ evaluated at $\bar{\mathbf{f}}$ with its elements given by

$$\left. \frac{\partial g_j^{(\text{NL})}}{\partial f_i} \right|_{\bar{\mathbf{f}}} = \frac{\sum_{l_j=1}^L \exp\left(-\sum_{i=1}^I a_{l_j i} \bar{f}_i\right) a_{l_j i}}{\sum_{l_j=1}^L \exp\left(-\sum_{i=1}^I a_{l_j i} \bar{f}_i\right)}, \quad (3)$$

and

$$\Delta g_j^{(\text{NL})}(\mathbf{f}) = g_j^{(\text{NL})}(\mathbf{f}) - g_j^{(\text{NL})}(\bar{\mathbf{f}}) - \nabla_{\mathbf{f}} g_j^{(\text{NL})}(\bar{\mathbf{f}})^\top (\mathbf{f} - \bar{\mathbf{f}}). \quad (4)$$

The NLPV DXT in the form of Eq. (2) reveals that, in addition to the constant term $g_j^{(\text{NL})}(\bar{\mathbf{f}})$, it comprises two components that are linearly and nonlinearly dependent upon image $\mathbf{f}$. Without loss of generality, we select the expansion point at $\bar{\mathbf{f}} = \mathbf{0}$ and thus have

$$\begin{aligned} g_j^{(\text{NL})}(\bar{\mathbf{f}}) &= 0, \\ \left. \frac{\partial g_j^{(\text{NL})}}{\partial f_i} \right|_{\bar{\mathbf{f}}} &= h_{ji}, \\ \Delta g_j^{(\text{NL})}(\mathbf{f}) &= -\ln \frac{1}{L} \sum_{l_j=1}^L \exp\left(-\sum_{i=1}^I a_{l_j i} f_i\right) - \sum_{i=1}^I h_{ji} f_i, \end{aligned} \quad (5)$$

where $h_{ji} = \sum_{l_j=1}^L \frac{a_{l_j i}}{L}$. Note that Eq. (4) and hence $\Delta g_j^{(\text{NL})}(\mathbf{f})$ in Eq. (5) are equivalent to Eq. (2). Nonetheless, showing them as the expression of $\Delta g_j^{(\text{NL})}(\mathbf{f})$ as a function of $\mathbf{f}$ is key to the algorithm development, as seen below. As a result, the data model, i.e., the NLPV DXT, in Eq. (2) is written in a matrix–vector form^{16–18} as

$$\mathbf{g}^{(\text{NL})}(\mathbf{f}) - \Delta \mathbf{g}^{(\text{NL})}(\mathbf{f}) = \mathcal{H} \mathbf{f}, \quad (6)$$

where $\Delta \mathbf{g}^{(\text{NL})}(\mathbf{f})$ is a vector of size J with element $\Delta g_j^{(\text{NL})}(\mathbf{f})$ defined in Eq. (5) and matrix $\mathcal{H}$ of size $J \times I$ has elements h_{ji} .

2.2 Methods

We develop below the NCPD algorithm to accurately reconstruct $\mathbf{f}$ from knowledge of $\mathbf{g}^{(\text{NL})}(\mathbf{f})$ by inverting the NLPV DXT in Eq. (1).

2.2.1 Optimization program

Letting $\mathbf{g}$ denote data vector, we formulate an optimization program as

$$\mathbf{f}^* = \arg \min_{\mathbf{f}} D(\mathbf{g}; \mathbf{g}^{(\text{NL})}(\mathbf{f})), \quad \text{s.t. } \|\mathbf{f}\|_{\text{TV}} \leq \gamma \quad \text{and} \quad f_i \geq 0, \quad (7)$$

where $D(\mathbf{g}; \mathbf{g}^{(\text{NL})}(\mathbf{f}))$ denotes a divergence between data $\mathbf{g}$ and model data $\mathbf{g}^{(\text{NL})}(\mathbf{f})$, $\|\mathbf{f}\|_{\text{TV}}$ is the image total variation (TV),^{21,22} γ the image-TV constraint parameter, and $f_i \geq 0$, where $i = 1, 2, \dots, I$, a non-negativity constraint on the image. Although data divergences of different forms can be designed, we consider in this work a specific data divergence as

$$D(\mathbf{g}; \mathbf{g}^{(\text{NL})}(\mathbf{f})) = \|\mathbf{g} - \mathbf{g}^{(\text{NL})}(\mathbf{f})\|_2 = \|\mathbf{g}'(\mathbf{f}) - \mathcal{H} \mathbf{f}\|_2, \quad (8)$$

which is the ℓ_2 -norm of the difference between $\mathbf{g}'$ and $\mathcal{H} \mathbf{f}$, where

$$\mathbf{g}'(\mathbf{f}) = \mathbf{g} - \Delta \mathbf{g}^{(\text{NL})}(\mathbf{f}). \quad (9)$$

Although a convex optimization program in the form of Eq. (7) is well studied, the current program can be nonconvex due to the fact that the Hessian matrix of $D(\mathbf{g}; \mathbf{g}^{(\text{NL})}(\mathbf{f}))$ can have negative eigenvalues. To the best of our knowledge, no algorithm exists for solving either theoretically exactly or numerically accurately the nonconvex optimization program in Eq. (7). This observation motivates our development of the NCPD algorithm below for solving the nonconvex optimization program.

2.2.2 NCPD algorithm

Algorithms exist for mathematically exactly solving convex optimization programs, including those with a general form of Eq. (7). We develop below an algorithm by modifying a well-established PD algorithm for convex optimization programs to solve the nonconvex optimization program in Eq. (7).

Convexification. We first convexify the nonconvex optimization program in Eq. (7)^{16–18} by replacing $\Delta \mathbf{g}^{(\text{NL})}(\mathbf{f})$ in Eq. (9) with $\overline{\Delta \mathbf{g}^{(\text{NL})}}$, an estimate of the nonlinear term that is presumably known a priori and independent of $\mathbf{f}$. It can readily be observed now that data divergence in Eq. (8), and thus the program in Eq. (7), becomes convex and that existing algorithms for convex optimization programs can thus be applied to solving the convexified program. In this work, we tailor a specific first-order PD algorithm¹² to solve the convexified optimization program. The explicit form, detailed pseudocode, and characterization of the PD algorithm can be found in Refs. 13, 23, and 24. In particular, the updated procedure of the PD algorithm is written as

$$\mathbf{f}^{(n+1)} = \mathbf{f}^{(n)} + V(\mathbf{f}^{(n)}, \overline{\mathbf{g}}'), \quad (10)$$

where $\mathbf{f}^{(n)}$ denotes the reconstruction at iteration n , $\overline{\mathbf{g}}' = \mathbf{g} - \overline{\Delta \mathbf{g}^{(\text{NL})}}$ is independent of $\mathbf{f}^{(n)}$, and the explicit form of $V(\mathbf{f}^{(n)}, \overline{\mathbf{g}}')$ and its pseudocode can be found in Appendix A, as well as Refs. 23 and 24.

NCPD algorithm. We then propose the NCPD algorithm to solve the nonconvex optimization program in Eq. (7) by replacing $\overline{\mathbf{g}}'$ with $\mathbf{g}'(\mathbf{f}^{(n)})$ in Eq. (10) as

$$\mathbf{f}^{(n+1)} = \mathbf{f}^{(n)} + V(\mathbf{f}^{(n)}, \mathbf{g}'(\mathbf{f}^{(n)})), \quad (11)$$

where $\mathbf{g}'(\mathbf{f}^{(n)})$ is calculated using Eqs. (5) and (9) with $\mathbf{f} = \mathbf{f}^{(n)}$ at iteration n .

Although the NCPD algorithm in Eq. (11) is not shown mathematically exactly to solve the nonconvex optimization program in Eq. (7), we devise convergence conditions for the NCPD algorithm in the verification and characterization studies to be carried out below.

We first design a dimensionless data-convergence metric as

$$\tilde{D}_{\mathbf{g}}(\mathbf{f}^{(n)}) = \frac{D(\mathbf{g}; \mathbf{g}^{(\text{NL})}(\mathbf{f}^{(n)}))}{\|\mathbf{g}\|_2} \quad (12)$$

for a verification study, and as

$$\tilde{D}_{\mathbf{g}}(\mathbf{f}^{(n)}) = \frac{|D(\mathbf{g}; \mathbf{g}^{(\text{NL})}(\mathbf{f}^{(n)})) - D(\mathbf{g}; \mathbf{g}^{(\text{NL})}(\mathbf{f}^{(n-1)}))|}{\|\mathbf{g}\|_2} \quad (13)$$

for a characterization study, where $D(\mathbf{g}; \mathbf{g}^{(\text{NL})}(\mathbf{f}^{(n)}))$ is given by Eq. (8). Two more dimensionless convergence metrics for the image TV and image itself are designed as

$$\tilde{D}_{\text{TV}}(\mathbf{f}^{(n)}) = \frac{||\mathbf{f}^{(n)}||_{\text{TV}} - \gamma|}{\gamma} \quad (14)$$

and

$$\tilde{D}_{\mathbf{f}}(\mathbf{f}^{(n)}) = \frac{\|\mathbf{f}^{(n)} - \mathbf{f}_{\text{ref}}\|_2}{\|\mathbf{f}_{\text{ref}}\|_2}, \quad (15)$$

where $\mathbf{f}_{\text{ref}} = \mathbf{f}_{\text{truth}}$ for a verification study in which truth image $\mathbf{f}_{\text{truth}}$ is known and $\mathbf{f}_{\text{ref}} = \mathbf{f}^{(n-1)}$ for a characterization study in which truth image $\mathbf{f}_{\text{truth}}$ is generally unknown.

Using metrics $\tilde{D}_{\mathbf{g}}(\mathbf{f}^{(n)})$, $\tilde{D}_{\text{TV}}(\mathbf{f}^{(n)})$, and $\tilde{D}_{\mathbf{f}}(\mathbf{f}^{(n)})$ as defined, we devise mathematical convergence conditions as

$$\tilde{D}_{\mathbf{g}}(\mathbf{f}^{(n)}) \rightarrow 0, \quad \tilde{D}_{\text{TV}}(\mathbf{f}^{(n)}) \rightarrow 0, \quad \tilde{D}_{\mathbf{f}}(\mathbf{f}^{(n)}) \rightarrow 0, \quad (16)$$

as $n \rightarrow \infty$. In practice, due to finite iteration number and limited computer precision, these metrics in numerical studies tend to very small, but nonzero, values close to computer precision.

In a verification study, data $\mathbf{g}$, often referred to as ideal data, are generated using the data model in Eq. (1) from known truth image $\mathbf{f}_{\text{truth}}$, i.e., $\mathbf{g} = \mathbf{g}^{(\text{NL})}(\mathbf{f}_{\text{truth}})$, thus resulting naturally in the first necessary convergence condition in Eq. (16) with $\tilde{D}_{\mathbf{g}}(\mathbf{f}^{(n)})$ defined in Eq. (12); the second necessary convergence condition makes sure that the constraint on the image TV in Eq. (7) is satisfied; and the third condition is a sufficient condition ensuring the exact reconstruction of the truth image and thus the inversion of NLPV DXT. The goal of the verification study is to demonstrate that the NCPD algorithm can satisfy numerically the two necessary convergence conditions and especially that, for the study conditions considered, it can satisfy numerically the sufficient convergence condition.

In a characterization study, we examine the convergence property of the NCPD algorithm when it is applied to data $\mathbf{g}$ containing components inconsistent with the NLPV DXT in Eq. (1), i.e., $\mathbf{g} \neq \mathbf{g}^{(\text{NL})}(\mathbf{f}_{\text{truth}})$. In this case, $D(\mathbf{g}; \mathbf{g}^{(\text{NL})}(\mathbf{f}^{(n)}))$ approaches an unknown, nonzero value as $n \rightarrow \infty$, even when the NCPD algorithm converges. The first necessary convergence condition in Eq. (16) thus is devised with $\tilde{D}_{\mathbf{g}}(\mathbf{f}^{(n)})$ defined in Eq. (13). The second necessary convergence condition, identical to that for the verification study, again checks the TV-constraint condition in the optimization program, whereas the third necessary (instead of sufficient) convergence condition requires that the normalized change between two consecutive reconstructions diminishes as n increases. The third necessary condition is of some utility only when it is used in combination with other convergence conditions.

2.3 Numerical Study Design

Two numerical studies, a verification and a characterization study, are carried out in this work. In each of the studies, we consider the 2D-NLPV DXT for a fan-beam, circular configuration. However, the studies can readily be extended to the 3D-NLPV DXT for a cone-beam configuration. In the fan-beam configuration, the center of a 2-D-image array of $I = 512 \times 512$ identical square-shaped pixels of size 0.0488 cm coincides with the center-of-rotation of the detector-source pair; a linear detector of 60 cm in length consists of $J_m = 500$ identical bins; and the distances from the x-ray source to the center of the linear detector and to the center-of-rotation are 150 and 100 cm, respectively. Without loss of generality, a point-x-ray source, i.e., $L_s = 1$, is considered, whereas the detector bin is divided into $L_d = 5$ subelements. At each of $J_a = 1080$ views uniformly distributed over 2π , using Eq. (1), we generated from a phantom ideal data $\mathbf{g} = \mathbf{g}^{(\text{NL})}(\mathbf{f}_{\text{truth}})$ containing a total of $J = J_m \times J_a = 540,000$ measurements.

Two phantoms are used in this work. One is an ellipse phantom with metal-like, high-contrast rods and low-contrast disks, as shown in Fig. 3(a). This phantom, simply referred to hereafter as the ellipse phantom, is used to demonstrate the NLPV artifacts associated with metal or high-contrast objects in the image. The other is the FORBILD phantom²⁵ with varying levels of contrast and complexity, as shown in Fig. 4(a). The FORBILD phantom is commonly used for image reconstruction in the literature. Without loss of generality, we choose $\bar{\mathbf{f}} = \mathbf{0}$ and thus $\mathbf{g}(\bar{\mathbf{f}}) = \mathbf{0}$. In the verification study, we use ideal data $\mathbf{g} = \mathbf{g}^{(\text{NL})}(\mathbf{f}_{\text{truth}})$ calculated from the ellipse and FORBILD phantoms, whereas in the characterization study, we generate noisy data $\mathbf{g}$ by adding

Poisson noise to ideal data $\mathbf{g}^{(\text{NL})}(\mathbf{f}_{\text{truth}})$, which are inconsistent now with the NLPV DXT in Eq. (1), i.e., $\mathbf{g} \neq \mathbf{g}^{(\text{NL})}(\mathbf{f}_{\text{truth}})$. For a reference, in addition to the results of the NCPD algorithm, we also reconstruct images from ideal and noisy data using the FBP algorithm with a Hanning filter.

3 Results

3.1 Result of a Verification Study

In the verification study, with full knowledge of truth image $\mathbf{f}_{\text{truth}}$, we first calculate its TV and use it as TV parameter γ in Eq. (7). We then apply the NCPD algorithm to reconstructing the image from the ideal data through solving the nonconvex optimization program in Eq. (7). In Fig. 2, we show the convergence metrics in Eq. (16) for the verification studies with the ellipse phantom (top row) and the FORBILD phantom (bottom row). The metrics are computed as functions of the iteration number and plotted on a log-log (instead of linear-linear) scale to reveal in detail their convergence values and decreasing trends.

Based upon the verification-study results in Fig. 2 and in Appendix B, in which additional verification studies are carried out in double precision to further confirm the convergence trend of the NCPD algorithm, one can observe that the convergence metrics have decaying trends until reaching computer precision. These convergence results thus verify that the NCPD algorithm can achieve numerically the necessary convergence conditions in Eq. (16). In the absence of a theoretical proof of the convergence of the NCPD algorithm, it is of significance to rigorously demonstrate its convergence numerically. This is our motivation for carrying out the verification studies in Fig. 2 and in Appendix B. We also point out that the number of iterations taken to achieve computer precision is not a concern here as the goal of our study is to reach a solution for the nonconvex optimization program, as revealed by the decreasing trends of convergence metrics until reaching computer precision.

In the verification study, metric $\tilde{D}_{\mathbf{f}}(\mathbf{f}^{(n)})$ provides a measure of the difference between the reconstructed and truth images. It can thus be used as a metric for measuring not only algorithm convergence but also, more importantly, reconstruction accuracy. As shown in Fig. 2, metric $\tilde{D}_{\mathbf{f}}(\mathbf{f}^{(n)})$ has a decaying trend until reaching computer precision. Furthermore, we display in Figs. 3 and 4 the truth images, the images reconstructed using the NCPD algorithm, and their differences. The image results, especially the difference image, together with the convergence result of $\tilde{D}_{\mathbf{f}}(\mathbf{f}^{(n)})$, suggest that under the study conditions considered, the NCPD algorithm

Fig. 2 Convergence metrics (a), (d) $\tilde{D}_{\mathbf{g}}(\mathbf{f}^{(n)})$; (b), (e) $\tilde{D}_{\text{TV}}(\mathbf{f}^{(n)})$; and (c), (f) $\tilde{D}_{\mathbf{f}}(\mathbf{f}^{(n)})$ as functions of iteration number n , plotted on a log-log scale to reveal in detail the convergence trend of the NCPD algorithm in the verification study for the ellipse phantom (top row) and FORBILD phantom (bottom row).

Fig. 3 (a) Ellipse phantom; (b), (c) images reconstructed using the NCPD and FBP algorithms; (d) difference between images (a) and (b); and (e) difference between images (a) and (c). The display windows are $[0.195, 0.205] \text{ cm}^{-1}$ (or, equivalently, $[-25, 25] \text{ HU}$) for (a)–(c), and $[-0.005, 0.005] \text{ cm}^{-1}$ for (d) and (e). For reference, the background is at 0.2 cm^{-1} .

Fig. 4 (a) FORBILD phantom; (b), (c) images reconstructed using the NCPD and FBP algorithms; (d) difference between images (a) and (b), and (e) difference between images (a) and (c). The display windows are $[0.2095, 0.2105] \text{ cm}^{-1}$ (or, equivalently, $[-2.4, 2.4] \text{ HU}$) for (a)–(c), and $[-0.0005, 0.0005] \text{ cm}$ for (d) and (e). For reference, the background is at 0.21 cm^{-1} .

accurately inverts the NLPV DXT numerically in Eq. (1). Images reconstructed from the same data using the FBP algorithm, as well as their differences from their corresponding truth images, are also shown in Figs. 3 and 4, as a reference to reveal the uncorrected NLPV effect.

3.2 Result of an Initial Characterization Study

The verification study above, in which ideal data $\mathbf{g}$ are consistent completely with the NLPV DXT in Eq. (1), i.e., $\mathbf{g} = \mathbf{g}^{(\text{NL})}(\mathbf{f}_{\text{truth}})$, is of significance as it establishes that the NCPD algorithm can, for the study conditions considered, accurately invert the NLPV DXT numerically. It is also important to characterize the convergence properties of the NCPD algorithm when applied to data containing components inconsistent with the NLPV DXT in Eq. (1) upon which the algorithm is developed. There are numerous sources of inconsistency, and a detailed and systematic characterization of the algorithm in the presence of various inconsistent components, while beyond the scope of this work, warrants future investigation. In this work, we perform, as a demonstration, an initial characterization study in which images are reconstructed from data containing Poisson noise, a component inconsistent with the NLPV DXT in Eq. (1). Specifically, the noisy data are generated by adding to the ideal data (i.e., $\mathbf{g}^{(\text{NL})}(\mathbf{f}_{\text{truth}})$) Poisson noise at a level corresponding to 10^7 photons per detector bin in an air scan at each detector bin, following the summation of exponential and prior to the logarithmic step.

In the presence of data inconsistency, even if the truth-image TV is known, it may not be desirable to use its value as the TV constraint parameter γ for yielding a reconstruction of subjectively desirable visualization and/or of other optimal quality metrics.^{23,24,26,27} Moreover, in realistic CT imaging, the TV of an image desired (i.e., to be reconstructed) is generally unknown. Therefore, we perform instead a series of image reconstructions using different values of γ to sweep empirically the parameter space of γ . In the first four columns of Fig. 5, we display images reconstructed with four selective, different values of γ , each of which was obtained using the NCPD algorithm when the necessary convergence conditions in Eq. (16) were satisfied

Fig. 5 Images reconstructed of the ellipse (top row) and FORBILD (bottom row) phantoms from their corresponding noisy data using the NCPD algorithm (a)–(d) and (f)–(i) and using the FBP algorithm with a Hanning filter (e) and (j), in the characterization studies. Specifically, the ellipse images are reconstructed with (a) $\gamma = 1003.2$, (b) $\gamma = 1075.0$, (c) $\gamma = 1085.0$, and (d) $\gamma = 1226.2$, whereas the FORBILD images are obtained with (f) $\gamma = 1184.7$, (g) $\gamma = 1406.9$, (h) $\gamma = 1451.3$, and (i) $\gamma = 1629.0$, respectively. Display windows are $[0.195, 0.205] \text{ cm}^{-1}$ (or, equivalently, $[-25, 25] \text{ HU}$) for the top row, and $[0.2095, 0.2105] \text{ cm}^{-1}$ (or, equivalently, $[-2.4, 2.4] \text{ HU}$) for the bottom row. For reference, the TV values of the truth image are 1114.7 and 1480.9 for the ellipse and FORBILD phantoms, respectively.

Fig. 6 Convergence metrics (a), (d) $\bar{D}_g(\mathbf{f}^{(n)})$; (b), (e) $\bar{D}_{TV}(\mathbf{f}^{(n)})$; and (c), (f) $\bar{D}_f(\mathbf{f}^{(n)})$ as functions of iteration number n of the ellipse (top row) and FORBILD (bottom row) phantoms, plotted on a log–log scale to reveal in detail the convergence trend of the NCPD algorithm for reconstruction obtained with $\gamma = 1085.0$ for the ellipse phantom and with $\gamma = 1451.3$ for the FORBILD phantom.

numerically at a single-precision level of the computer used, for both ellipse phantom (top row) and FORBILD phantom (bottom row). Different values of γ yield, as expected, images of different visualizations. As a reference, images reconstructed from the same noisy data of both phantoms using the FBP algorithm are also shown in the last column of Fig. 5. In addition, for $\gamma = 1085.0$ with the ellipse phantom and $\gamma = 1451.29$ with the FORBILD phantom, we show in Fig. 6 the convergence metrics as functions of the iteration number plotted on a log–log scale.

In addition, it may be of practical interest to inspect the images at intermediate iterations. We show in Fig. (7) reconstructed images from the noisy data of the FORBILD phantom using

Fig. 7 Images reconstructed from noisy data of the FORBILD phantom using the NCPD algorithm with $\gamma = 1406.9$ at iteration (a) 60, (b) 100, (c) 300, and (d) 1000, as well as (e) the converged image. Display window is $[0.2095, 0.2105] \text{ cm}^{-1}$.

the NCPD algorithm with $\gamma = 1406.9$ at iterations 60, 100, 300, and 1000, as well as the converged image.

4 Conclusion and Discussion

In this work, we formulate the inversion of the NLPV DXT as a nonconvex optimization program and develop an algorithm, referred to as the NCPD algorithm, to accurately achieve the inversion numerically through solving the program. Quantitative studies are carried out to verify that, under the study conditions of relevance, the NCPD algorithm can accurately yield reconstructions numerically and thus in essence accurately invert the NLPV DXT numerically. We also perform an initial characterization study on the convergence properties of the NCPD algorithm when applied to data containing noise. The convergence properties of the NCPD algorithm in the presence of additional physical factors remain to be investigated and characterized. Through numerically solving the nonconvex optimization program and accurately inverting the NLPV DXT numerically, we show that the typical banding artifacts due to the NLPV effect, as observed in the FBP images as well as images from the PD algorithm without the nonlinear correction step, can be eliminated or reduced effectively by the NCPD algorithm proposed.

A Taylor-series expansion was used for linearizing the NLPV data model. The expansion point was selected at $\bar{\mathbf{f}} = \mathbf{0}$ because the linear term now becomes the linear average of the subelement rays within one detector bin, which is a physically intuitive linear approximation to the NLPV model. Different expansion points may lead to possibly nonzero zeroth-order term $\mathbf{g}(\bar{\mathbf{f}})$ and different matrices $\mathcal{H}$. Our empirical study with different expansion points other than zero, e.g., using the central ray as the linear approximation, indicate no change to the accuracy of the convergence images.

In this work, the x-ray focal spot is assumed to be a single point. However, the NLPV DXT can readily accommodate a focal spot of finite size by considering multiple rays emitting from multiple points on the focal spot, as illustrated in Fig. 1. The incorporation of a focal spot of finite size into the NLPV DXT is mathematically identical to that of the detector bin of finite size, as the NLPV DXT is formulated based upon each individual ray connecting two points on the focal spot and a detector bin. As such, this work may provide an approach to investigating the impact of the focal spot with a finite size on the NLPV effect.

Because the optimization program and the NCPD algorithm are designed for inverting the data model in a matrix form, they can be applied to a data model in a matrix form for a cone-beam geometry. However, for the circular-scanning configuration that is most widely used, the cone-beam effect on image planes away from the plane containing the circular trajectory may compound with the NLPV effect to impact the inversion accuracy. Computation efficiency is another factor to be considered in cone-beam reconstruction. Future studies are warranted for investigating these factors.

The NCPD algorithm was developed without considering in the NLPV DXT additional physical factors such as beam-hardening, scatters, detector aperture, and focal spot in realistic CT imaging. Clearly, these physical factors can impact the performance characteristics of the algorithm. Therefore, it is of significance to investigate systematically how the algorithm

responds to the physical factors, and their estimation uncertainties, in realistic CT imaging studies. The NLPV correction by the proposed method might be limited by the noise level in the data. In a high-noise level case, the noise, instead of the NLPV effect, might be the dominant source of artifacts.

Although a nonconvex optimization program only in the form of Eq. (7) is considered in this work, it is straightforward to design nonconvex optimization programs of different forms, and the NCPD algorithm can readily be tailored accordingly to address these programs. However, it remains to be established whether the NCPD algorithm can indeed accurately solve these nonconvex optimization programs numerically. The NCPD algorithm is developed by modifying the PD algorithm^{12,13,23,24} that is used for solving convex optimization programs. The same approach can readily be applied to devising additional algorithms for the nonconvex optimization program through modifying other algorithms^{28,29} for solving convex optimization programs. Again, whether these additional algorithms can solve accurately the nonconvex optimization program needs to be established.

This work reveals the possibility of a numerically accurate inversion of the NLPV DXT, and it may also yield insights into the development of practical algorithms and procedures that may effectively correct for the NLPV effect in realistic CT imaging. This work also provides a general approach to developing algorithms for solving nonconvex optimization programs.

5 Appendix A: Definition of the Operator V

The operator $V(\mathbf{f}^{(n)}, \bar{\mathbf{g}}')$ corresponds to lines 6 to 9 in the pseudocode on page 6061 in Ref. 23 and lines 6 to 8 in the pseudocode on page 7328 in Ref. 24. In detail, operator $V(\mathbf{f}^{(n)}, \bar{\mathbf{g}}')$ in Eq. (10) entails the following steps:

$$\bar{\mathbf{f}}^{(n)} = \mathbf{f}^{(n)} + \theta(\mathbf{f}^{(n)} - \mathbf{f}^{(n-1)}), \quad (17)$$

$$\mathbf{p}^{(n+1)} = \frac{\mathbf{p}^{(n)} - \sigma(\bar{\mathbf{g}}' - \mathcal{H}\bar{\mathbf{f}}^{(n)})}{1 + \sigma}, \quad (18)$$

$$\mathbf{q}'^{(n)} = \mathbf{q}^{(n)} + \sigma\alpha\nabla\bar{\mathbf{f}}^{(n)}, \quad (19)$$

$$\mathbf{q}^{(n+1)} = \mathbf{q}'^{(n)} - \sigma \frac{\mathbf{q}'^{(n)}}{|\mathbf{q}'^{(n)}|} \text{POLB}_{\alpha\gamma} \left(\frac{|\mathbf{q}'^{(n)}|}{\sigma} \right), \quad (20)$$

$$\mathbf{r}^{(n+1)} = \text{neg}(\mathbf{r}^{(n)} + \sigma\beta\bar{\mathbf{f}}^{(n)}), \quad (21)$$

$$V(\mathbf{f}^{(n)}, \bar{\mathbf{g}}') = -\tau(\mathcal{H}\mathbf{p}^{(n+1)} + \alpha\nabla\mathbf{q}^{(n+1)} + \beta\mathbf{r}^{(n+1)}), \quad (22)$$

where $\theta = 1$; σ , τ , α , and β are algorithm parameters that can be computed by calculating the largest singular values of matrices; $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{r}$ are the dual variables; POLB() denotes Projection-Onto- ℓ_1 -Ball operation; and neg() denotes nonpositivity thresholding operation. Note that $\bar{\mathbf{g}}'$ is used in Eq. (18) above. As such, once $\bar{\mathbf{g}}'$ is replaced with $\mathbf{g}'(\mathbf{f}^{(n)})$ in Eq. (18), it leads to the definition of operator $V(\mathbf{f}^{(n)}, \mathbf{g}'(\mathbf{f}^{(n)}))$ in Eq. (11) above.

6 Appendix B: Verification Studies with Double Precision

We carry out two additional verification studies on the reconstruction of the FORBILD phantom with single and double precision using ideal data $\mathbf{g} = \mathbf{g}^{(\text{NL})}(\mathbf{f}_{\text{truth}})$. Due to the increased computation load with double precision, a system matrix with smaller size is used for demonstration purposes. The truth image is down-sampled by a 1:12 ratio from that used in Fig. 4, and the ideal data is generated. As a result, image vector $\mathbf{f}_{\text{truth}}$ is of size $I = 43 \times 43 = 1849$, and data vector $\mathbf{g}$ is of size $J = 90 \times 41 = 3690$. We show in Fig. 8 the convergence metrics $\tilde{D}_{\mathbf{g}}(\mathbf{f}^{(n)})$, $\tilde{D}_{\text{TV}}(\mathbf{f}^{(n)})$, and $\tilde{D}_{\mathbf{f}}(\mathbf{f}^{(n)})$, as defined in Eqs. (12), (14), and (15), respectively, as functions of the iteration

Fig. 8 Convergence metrics (a) $\tilde{D}_g(\mathbf{f}^{(n)})$, (b) $\tilde{D}_{TV}(\mathbf{f}^{(n)})$, and (c) $\tilde{D}_f(\mathbf{f}^{(n)})$ as functions of iteration number n for the FORBILD phantom, plotted on a log–log scale to reveal in detail the convergence trends of the NCPD algorithm in verification studies with a single (solid) and double (dashed) precision. For (a) and (c), the tails of the double precision curve are zoomed in to reveal that they have plateaued even in the log–log scale.

number in the two verification studies with single and double precision. In each panel of Fig. 8, two curves are shown, representing the results obtained with a single (solid curve) and double (dashed curve) precision. Furthermore, the tails of the double precision curves in Figs. 8(a) and 8(c) are zoomed in to show details. Two key observations can be made for each convergence metric in Fig. 8. First, both single and double precision curves maintain decaying trends until plateauing at levels of the corresponding computer precision. Second, a double precision can further lower the convergence metrics beyond the level of a single precision. These convergence results thus verify that the NCPD algorithm can achieve numerically the necessary convergence conditions in Eq. (16) up to computer precision. Similar results were obtained for the ellipse phantom, which are not shown here.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

Acknowledgments

This work was supported in part by NIH R01 Grant Nos. EB026282 and EB023968 and the Grayson-Jockey Club Research. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

References

1. G. Glover and N. Pelc, “Nonlinear partial volume artifacts in x-ray computed tomography,” *Med. Phys.* **7**(3), 238–248 (1980).
2. P. M. Joseph and R. D. Spital, “The exponential edge-gradient effect in x-ray computed tomography,” *Phys. Med. Biol.* **26**(3), 473–487 (1981).
3. D. Heuscher and M. Vembar, “Reduced partial volume artifacts using spiral computed tomography and an integrating interpolator,” *Med. Phys.* **26**(2), 276–286 (1999).
4. J. Hsieh, “Nonlinear partial volume artifact correction in helical CT,” *IEEE Trans. Nucl. Sci.* **46**(3), 743–747 (1999).
5. J. Hsieh, “Partial volume artifact reduction in computed tomography,” in *Proc. 3rd IEEE Int. Conf. Image Process.*, pp. 567–570 (1996).
6. G. J. Kemerink et al., “The nonlinear partial volume effect and computed tomography densitometry of foam and lung,” *Med. Phys.* **22**(9), 1445–1450 (1995).
7. R. A. Jucius and G. X. Kambic, “Measurements of computed tomography x-ray fields utilizing the partial volume effect,” *Med. Phys.* **7**(4), 379–382 (1980).
8. J. P. Chiverton and K. Wells, “Volumetric partial volume quantification via a statistical model of 3-D voxel gradient magnitude,” in *IEEE Symp. Conf. Record Nucl. Sci.*, Vol. 7, pp. 4106–4110 (2004).

9. B. De Man, "Iterative reconstruction for reduction of metal artifacts in CT," PhD Thesis, PhD Dissertation, University of Leuven, Belgium (2001).
10. Y. Zou, E. Y. Sidky, and X. Pan, "Partial volume and aliasing artefacts in helical cone-beam CT," *Phys. Med. Biol.* **49**(11), 2365–2375 (2004).
11. J. Stayman et al., "Overcoming nonlinear partial volume effects in known-component reconstruction of cochlear implants," *Proc. SPIE* **8668**, 86681L (2013).
12. A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imaging Vision* **40**(1), 120–145 (2011).
13. E. Y. Sidky, J. H. Jørgensen, and X. Pan, "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle–Pock algorithm," *Phys. Med. Biol.* **57**(10), 3065–3091 (2012).
14. X. Liu et al., "Optimization-based reconstruction for correcting non-linear partial volume artifacts in CT," *Proc. SPIE* **10948**, 109482Q (2019).
15. X. Liu et al., "A preliminary study on explicit compensation for the non-linear-partial-volume effect in CT," *Proc. SPIE* **11072**, 110720F (2019).
16. B. Chen et al., "Image reconstruction and scan configurations enabled by optimization-based algorithms in multispectral CT," *Phys. Med. Biol.* **62**(22), 8763–8793 (2017).
17. B. Chen et al., "Algorithm-enabled partial-angular-scan configurations for dual-energy CT," *Med. Phys.* **45**(5), 1857–1870 (2018).
18. X. Pan et al., "Non-convex optimization-based reconstruction in multispectral CT," in *Proc. 5th Int. Conf. Image Format. X-Ray CT*, pp. 373–376 (2018).
19. B. Chen et al., "Non-convex Chambolle–Pock algorithm for multispectral CT," in *Proc. 5th Int. Conf. Image Format. X-Ray CT*, pp. 377–381 (2018).
20. Y. Zhao, X. Zhao, and P. Zhang, "An extended algebraic reconstruction technique (E-ART) for dual spectral CT," *IEEE Trans. Med. Imaging* **34**(3), 761–768 (2015).
21. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D* **60**(1–4), 259–268 (1992).
22. E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *J. X-Ray Sci. Technol.* **14**(2), 119–139 (2006).
23. Z. Zhang et al., "Investigation of optimization-based reconstruction with an image-total-variation constraint in PET," *Phys. Med. Biol.* **61**(16), 6055–6084 (2016).
24. D. Xia et al., "Optimization-based image reconstruction with artifact reduction in C-arm CBCT," *Phys. Med. Biol.* **61**(20), 7300–7333 (2016).
25. Z. Yu et al., "Simulation tools for two-dimensional experiments in x-ray computed tomography using the FORBILD head phantom," *Phys. Med. Biol.* **57**(13), N237–N252 (2012).
26. J. Bian et al., "Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT," *Phys. Med. Biol.* **55**(22), 6575–6599 (2010).
27. X. Han et al., "Algorithm-enabled low-dose micro-CT imaging," *IEEE Trans. Med. Imaging* **30**(3), 606–620 (2011).
28. S. Ramani and J. A. Fessler, "A splitting-based iterative algorithm for accelerated statistical x-ray CT reconstruction," *IEEE Trans. Med. Imaging* **31**(3), 677–688 (2012).
29. S. Boyd et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends® Mach. Learn.* **3**(1), 1–122 (2011).

Buxin Chen received his BS degree in physics from Nanjing University, Nanjing, China, in 2008, his MS degree in medical physics from Columbia University, New York, in 2011, and his PhD in medical physics from The University of Chicago in 2017. Currently, he is a staff scientist in the Department of Radiology at The University of Chicago. His research interests include CT system development and image reconstruction.

Xin Liu received his BS degree in physics from Xiaogan College, China, in 2001, his MS degree in applied mathematics from Hubei University, Wuhan, China, in 2004, and his PhD in optical engineering from Huazhong University of Science and Technology, Wuhan, in 2007. He was a visiting scholar in the Department of Radiology at The University of Chicago at the time of this work. Currently, he is a lecturer at the College of Physics and Optoelectronic Engineering,

Shenzhen University. His research interests include x-ray phase-contrast imaging and CT image reconstruction.

Zheng Zhang received his BS degree in physics from University of Science and Technology of China, Hefei, China, in 2007, his MA degree in physics from Johns Hopkins University, Baltimore, Maryland, in 2009, and his PhD in medical physics from The University of Chicago in 2016. Currently, he is a staff scientist at the Department of Radiology, The University of Chicago. His research interests are image reconstruction in cone-beam computed tomography (CBCT) and positron emission tomography (PET).

Dan Xia received his BS and MS degrees from Tsinghua University, Beijing, in 1999 and 2002, and his PhD in medical physics from The University of Chicago in 2009. Currently, he is a staff scientist with the Department of Radiology, The University of Chicago. His research interests are in CT image reconstruction and CT system calibration.

Emil Y. Sidky received his BS degree in physics and mathematics from the University of Wisconsin, Madison, in 1987 and his PhD in physics from The University of Chicago in 1993. He held academic positions in physics at the University of Copenhagen and Kansas State University. He joined The University of Chicago in 2001, where he is currently a research professor. His current interests are in CT image reconstruction and large-scale optimization.

Xiaochuan Pan received his BS degree from Beijing University, Beijing, China, in 1982, his MS degree from the Institute of Physics, Chinese Academy of Sciences, Beijing, in 1986, and his master's and PhD degrees from The University of Chicago in 1988 and 1991, respectively, all in physics. Currently, he is a professor with the Department of Radiology, The University of Chicago. His research interests include physics, algorithms, and applications of tomographic imaging.