

# Duality-based projection-domain tomography solver for splitting-based X-ray CT reconstruction

Madison G. McGaffin      Jeffrey A. Fessler

**Abstract**—Model-based image reconstruction (MBIR) for X-ray CT produces high quality images from relatively low-dose scans, but the high computational cost of MBIR algorithms prevent them from being used ubiquitously in the clinic. Variable splitting with the alternating directions methods of multipliers (ADMM) provides rapidly converging algorithms by decomposing the challenging MBIR optimization problem into an iterated sequence of simpler subproblems. Variable splitting algorithms have achieved state-of-the-art performance in 2D, but replicating those successes in 3D has proved difficult. In this paper, we consider a simple splitting algorithm that decomposes the reconstruction problem into a nonnegative denoising problem and a quadratic tomography problem. Unlike prior work, we solve the tomography problem with a novel duality-based approach that yields convergent algorithms similar to ordered subsets methods and iterated filtered backprojection. We show some promising preliminary results.

## I. INTRODUCTION

Consider the following statistical X-ray CT reconstruction problem [15]

$$\hat{\mathbf{x}} = \underset{\mathbf{x} \geq 0}{\operatorname{argmin}} \left\{ J(\mathbf{x}) = \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_{\mathbf{W}}^2 + R(\mathbf{x}) \right\}, \quad (1)$$

with noisy sinogram data  $\mathbf{y} \in \mathbb{R}^M$ , system matrix  $\mathbf{A} \in \mathbb{R}^{M \times N}$ , statistical weights  $\mathbf{W} = \operatorname{diag}_i\{w_i\}$ , and convex edge-preserving regularizer  $R$ . The optimization problem (1) is impractical or impossible to solve in closed form due to the large dimension of  $\mathbf{x}$ , the nonnegativity constraint and the nonquadratic regularizer. Consequently, solving (1) requires an iterative optimization routine, but unfortunately this is a challenging problem to solve efficiently. Gradient-based simultaneous algorithms that update all pixels of  $\mathbf{x}$  simultaneously appear to be better-equipped to take advantage of modern highly-parallel hardware. However (1) presents several challenges for gradient-based methods:

- 1) the CT projection ( $\mathbf{A}$ ) and backprojection ( $\mathbf{A}'$ ) operations, which are both required to compute an update direction for simultaneous gradient-based methods, are computationally expensive; and
- 2) the reconstruction cost function  $J$  is difficult to precondition, especially for 3D reconstruction [3], [4], [9], [18].

In this paper, we consider a variable splitting algorithm that decomposes (1) into an image-space denoising problem and a quadratic “tomography” minimization problem involving the

M. G. McGaffin and J. A. Fessler are with the Dept. of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA. (email: mcgaffin@umich.edu, fessler@umich.edu)

Supported in part by NIH grant R01-HL-098686 and equipment donations from Intel Corporation.

CT system matrix. The primary contribution of this paper is a novel duality-based approach to this tomography problem.

## II. VARIABLE SPLITTING

Variable splitting with ADMM is a technique to solve a challenging optimization problem, *e.g.*, (1), with an iterated set of simpler subproblems [9], [11]–[13]. Let  $\mathbf{v} \in \mathbb{R}^N$ . Instead of directly solving (1), we solve the constrained problem

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \min_{\mathbf{v} \geq 0} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_{\mathbf{W}}^2 + R(\mathbf{v}) \quad \text{s.t. } \mathbf{x} = \mathbf{v}. \quad (2)$$

The augmented Lagrangian for this constrained problem is

$$\mathcal{L}(\mathbf{x}, \mathbf{u}; \boldsymbol{\eta}) = \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_{\mathbf{W}}^2 + R(\mathbf{v}) + \frac{1}{2} \|\mathbf{x} - \mathbf{v} + \boldsymbol{\eta}\|_{\boldsymbol{\Gamma}}^2, \quad (3)$$

where  $\boldsymbol{\Gamma} \succ \mathbf{0}$ . The alternating directions method of multipliers (ADMM) iterations leads to two subproblems:

$$\begin{cases} \mathbf{x}^{(n+1)} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_{\mathbf{W}}^2 + \frac{1}{2} \|\mathbf{x} - \mathbf{v}^{(n)} + \boldsymbol{\eta}^{(n)}\|_{\boldsymbol{\Gamma}}^2, \\ \mathbf{v}^{(n+1)} = \underset{\mathbf{v} \geq 0}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{v} - \mathbf{x}^{(n+1)} - \boldsymbol{\eta}^{(n)}\|_{\boldsymbol{\Gamma}}^2 + R(\mathbf{v}), \\ \boldsymbol{\eta}^{(n+1)} = \boldsymbol{\eta}^{(n)} + \mathbf{x}^{(n+1)} - \mathbf{v}^{(n+1)}. \end{cases} \quad (4)$$

The  $\mathbf{x}$  and  $\mathbf{v}$  updates do not need to be performed exactly to ensure convergence [2]; empirically, more accurate solutions accelerate convergence.

We choose  $\boldsymbol{\Gamma}$  to be diagonal. The  $\mathbf{v}$  update is then a penalized weighted least squares denoising problem. There are many algorithms to solve this class of problem, and it can be solved quickly even for large problems and non-smooth regularizers [7], [8]. Hereafter we focus on the more challenging tomography  $\mathbf{x}$  update. Note that many splitting-based algorithms have inner steps similar to the  $\mathbf{x}$  update we study here [9], [13]. The techniques in the following section can be applied to those algorithms as well.

## III. THE TOMOGRAPHY SUBPROBLEM

Though the  $\mathbf{x}$  update in (4) is a theoretically simple unconstrained quadratic minimization problem, solving for  $\mathbf{x}^{(n+1)}$  is challenging in practice [9], [10]. The popular family of gradient-based methods share the form:

$$\begin{cases} \mathbf{r}^{(m)} = \mathbf{A}\mathbf{x}^{(m)} - \mathbf{y}, \\ \mathbf{g}^{(m)} = \mathbf{A}'\mathbf{W}\mathbf{r}^{(m)} + \boldsymbol{\Gamma}(\mathbf{x}^{(m)} - \mathbf{v}^{(n)} + \boldsymbol{\eta}^{(n)}), \\ \mathbf{d}^{(m)} = \mathbf{f}^{(m)}(\mathbf{g}^{(m)}), \\ \mathbf{x}^{(m+1)} = \mathbf{x}^{(m)} + \mathbf{d}^{(m)}. \end{cases} \quad (5)$$

The function  $\mathbf{f}^{(m)} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  is an iteration-dependent preconditioning and step-size computing step. Preconditioning modifies the search direction such that  $\mathbf{d}^{(m)}$  “points” more toward the global minimizer  $\mathbf{x}^*$ . The ideal, and unrealizable, preconditioner would rapidly implement the inverse of the cost function Hessian

$$\mathbf{f}_{\text{ideal}}(\mathbf{g}) = \mathbf{P}_{\text{ideal}}\mathbf{g} = [\mathbf{A}'\mathbf{W}\mathbf{A} + \mathbf{\Gamma}]^{-1}\mathbf{g}. \quad (6)$$

Many practical approximations to  $\mathbf{P}_{\text{ideal}}$  have been proposed involving diagonal matrices, circulant operators and the FFT [3], [4], [9], [13], [18]. Unfortunately the shift-varying nature of the Hessian, induced by the statistical weights  $\mathbf{W}$  and geometrical properties of  $\mathbf{A}$ , make designing a highly effective preconditioner challenging.

### A. Duality approach

Instead of solving the tomography problem directly with a gradient-based method, we introduce an auxiliary variable  $\mathbf{u} \in \mathbb{R}^M$  and consider the following equivalent saddle-point problem:

$$\mathbf{x}^{(n+1)} = \underset{\mathbf{x}}{\text{argmin}} \max_{\mathbf{u}} \left\{ \mathcal{S}(\mathbf{x}, \mathbf{u}) = (\mathbf{A}\mathbf{x} - \mathbf{y})'\mathbf{W}\mathbf{u} + \frac{1}{2} \left\| \mathbf{x} - \mathbf{v}^{(n)} + \boldsymbol{\eta}^{(n)} \right\|_{\mathbf{\Gamma}}^2 - \frac{1}{2} \|\mathbf{u}\|_{\mathbf{W}}^2 \right\}. \quad (7)$$

Performing the inner maximization yields the original quadratic function in (4):

$$\nabla_{\mathbf{u}}\mathcal{S} = \mathbf{W}(\mathbf{A}\mathbf{x} - \mathbf{y}) - \mathbf{W}\mathbf{u} = \mathbf{0}, \quad (8)$$

$$\mathbf{u}(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{y}, \quad (9)$$

$$\mathcal{S}(\mathbf{x}, \mathbf{u}(\mathbf{x})) = \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_{\mathbf{W}}^2 + \frac{1}{2} \|\mathbf{x} - \mathbf{v} + \boldsymbol{\eta}\|_{\mathbf{\Gamma}}^2, \quad (10)$$

so solving (7) (*i.e.*, finding the saddle point), solves the quadratic  $\mathbf{x}$  update problem.

We observe that  $\mathcal{S}(\mathbf{x}, \mathbf{u}_0)$  is convex and continuous in  $\mathbf{x}$  for all  $\mathbf{u}_0$ , and  $\mathcal{S}(\mathbf{x}_0, \mathbf{u})$  is concave and continuous in  $\mathbf{u}$  for all  $\mathbf{x}_0$ . By Sion’s minimax theorem [14], we can reverse the order of the minimization and maximization steps in (7),

$$\min_{\mathbf{x}} \max_{\mathbf{u}} \mathcal{S}(\mathbf{x}, \mathbf{u}) = \max_{\mathbf{u}} \min_{\mathbf{x}} \mathcal{S}(\mathbf{x}, \mathbf{u}), \quad (11)$$

to find the saddle point. In other words, instead of solving the image-space primal problem, we can solve the projection-domain dual problem.

We compute the minimizing value of  $\mathbf{x}$  in terms of  $\mathbf{u}$ ,

$$\mathbf{x}(\mathbf{u}) = \mathbf{v}^{(n)} - \boldsymbol{\eta}^{(n)} - \mathbf{\Gamma}^{-1}\mathbf{A}'\mathbf{W}\mathbf{u}, \quad (12)$$

and plug (12) into (7) to yield the quadratic dual problem:

$$\mathbf{u}^* = \underset{\mathbf{u}}{\text{argmax}} \left\{ D(\mathbf{u}) = -\frac{1}{2} \mathbf{u}'(\mathbf{W} + \mathbf{W}\mathbf{A}\mathbf{\Gamma}^{-1}\mathbf{A}'\mathbf{W})\mathbf{u} + \mathbf{u}'\mathbf{W}(\mathbf{A}(\mathbf{v}^{(n)} - \boldsymbol{\eta}^{(n)}) - \mathbf{y}) \right\}. \quad (13)$$

There are many options for solving (13), and as in the primal problem, a natural family of algorithms are the gradient-based

methods. The general form of a gradient-based algorithm in the dual domain is:

$$\begin{cases} \mathbf{p}^{(m)} = \mathbf{W}(\mathbf{y} + \mathbf{u}^{(m)} + \mathbf{A}(\mathbf{\Gamma}^{-1}\boldsymbol{\gamma}^{(m)} - \mathbf{v}^{(n)} + \boldsymbol{\eta}^{(n)})), \\ \mathbf{q}^{(m)} = f^{(m)}(\mathbf{p}^{(m)}), \\ \mathbf{u}^{(m+1)} = \mathbf{u}^{(m)} + \mathbf{q}^{(m)}, \\ \boldsymbol{\gamma}^{(m+1)} = \boldsymbol{\gamma}^{(m)} + \mathbf{A}'\mathbf{W}\mathbf{q}^{(m)}, \end{cases} \quad (14)$$

with  $\mathbf{u}^{(0)} = \mathbf{0}$ ,  $\boldsymbol{\gamma}^{(0)} = \mathbf{0}$ , and  $f^{(m)}$  a preconditioning and step-size computing operation in the projection domain. After a number of iterations, we update  $\mathbf{x}$  with (12) and return to the outer iterations of the ADMM algorithm (4),

$$\mathbf{x}^{(n+1)} = \mathbf{x}(\mathbf{u}^{(m)}) = \mathbf{v}^{(n)} - \boldsymbol{\eta}^{(n)} - \mathbf{\Gamma}^{-1}\boldsymbol{\gamma}^{(m)}. \quad (15)$$

As a practical matter, we can warm-start this dual problem by saving the final values of  $\mathbf{u}^{(m)}$  and  $\boldsymbol{\gamma}^{(m)}$  between  $\mathbf{x}$  updates and use those instead of  $\mathbf{u}^{(0)} = \mathbf{0}$  and  $\boldsymbol{\gamma}^{(0)} = \mathbf{0}$ .

### B. Projection-domain group coordinate ascent algorithm

Ordered subsets (OS) algorithms [1], [5] partition the system matrix, weights and data into disjoint subsets by view,  $\{S_k\}_{k=1}^K$ , such that

$$\mathbf{A}'\mathbf{W}(\mathbf{A}\mathbf{x}^{(m)} - \mathbf{y}) \approx \frac{N_{\beta}}{|S_k|} \sum_{\beta \in S_k} \mathbf{A}_{\beta}'\mathbf{W}_{\beta}(\mathbf{A}_{\beta}\mathbf{x}^{(m)} - \mathbf{y}_{\beta}). \quad (16)$$

The right hand side of (16) is used in OS methods as an approximate gradient that requires the forward- and back-projection of only  $|S_k|$  views instead of  $N_{\beta}$ . This gradient approximation is accurate when  $\mathbf{x}^{(m)}$  is far from the solution and when the subsets contain enough views, but for  $\mathbf{x}^{(m)}$  near the solution and smaller subsets, OS algorithms without relaxation approach limit cycles around the solution.

In the projection (dual) domain, an analogy to ordered subsets is group coordinate ascent (GCA). A GCA algorithm divides  $\mathbf{u}$  into  $K$  groups,  $\mathbf{u} = [\mathbf{u}_0, \dots, \mathbf{u}_{K-1}]$  and iteratively maximizes over each group while holding the others constant. Even if the maximization is performed approximately with a minorize-maximize step, the algorithm is convergent [16].

The following algorithm for solving (13) uses  $f^{(m)}$  to apply the linear preconditioner  $\mathbf{P}_k$  and computes the optimal step length for each group. For all groups  $k = 0, \dots, K-1$  perform

the following:

$$\begin{cases} \mathbf{p}_k^{(m)} &= \mathbf{W}_k \left( \mathbf{y} + \mathbf{u}_k^{(m)} \right. \\ &\quad \left. + \mathbf{A}_k \left( \Gamma^{-1} \gamma^{(m+\frac{k}{K})} - \mathbf{v}^{(n)} + \boldsymbol{\eta}^{(n)} \right) \right), \\ \mathbf{q}_k^{(m)} &= \mathbf{P}_k \mathbf{p}_k^{(m)}, \\ \mathbf{d}_k^{(m)} &= \mathbf{A}_k' \mathbf{W}_k \mathbf{q}_k^{(m)}, \\ \alpha_k^{(m)} &= - \frac{\left( \mathbf{q}_k^{(m)} \right)' \mathbf{p}_k^{(m)}}{\left( \mathbf{q}_k^{(m)} \right)' \mathbf{W}_k \mathbf{q}_k^{(m)} + \left( \mathbf{d}_k^{(m)} \right)' \Gamma^{-1} \mathbf{d}_k^{(m)}}, \\ \mathbf{u}_k^{(m+1)} &= \mathbf{u}_k^{(m)} + \alpha_k^{(m)} \mathbf{q}_k^{(m)}, \\ \gamma^{(m+\frac{k+1}{K})} &= \gamma^{(m+\frac{k}{K})} + \alpha_k^{(m)} \mathbf{d}_k^{(m)}. \end{cases} \quad (17)$$

These updates are iterated for  $m = 1, \dots, N_{\text{iter-dual}}$ . In the context of an ADMM algorithm with warm-starting,  $N_{\text{iter-dual}} = 1$  or 2 appears to be sufficient.

### C. Comparison to image-space algorithms

While (17) is written in a convenient form for implementation, the behavior of the algorithm is unclear. To compare (17) and traditional ordered subsets algorithms, we rewrite some of the steps in terms of  $\mathbf{x}$  using (15):

$$\begin{cases} \mathbf{r}_k^{(m)} &= \mathbf{A}_k \mathbf{x}^{(m+\frac{k}{K})} - \mathbf{y}_k, \\ \mathbf{u}_k^{(m+1)} &= \mathbf{u}_k + \alpha_k^{(m)} \mathbf{P}_k \mathbf{W}_k \left( \mathbf{u}_k^{(m)} - \mathbf{r}_k^{(m)} \right), \\ \mathbf{x}^{(m+\frac{k+1}{K})} &= \mathbf{x}^{(m+\frac{k}{K})} \\ &\quad - \alpha_k^{(m)} \Gamma^{-1} \mathbf{A}_k' \mathbf{W}_k \mathbf{P}_k \mathbf{W}_k \left( \mathbf{r}_k - \mathbf{u}_k \right), \end{cases} \quad (18)$$

with  $\mathbf{x}^{(0)} = \mathbf{v}^{(n)} - \boldsymbol{\eta}^{(n)}$ .

Recall that the group assignments are arbitrary and each loop through all the groups requires only one total forward and back-projection. The algorithm given in (18) will converge to the solution to the  $\mathbf{x}$ -update equation,  $\mathbf{x}^{(n+1)}$ , though possibly not monotonically in the primal cost function. This flexibility in subset selection is very different from traditional OS algorithms that require judicious design of the subsets to satisfy the subset approximation (16).

The step size computation guarantees monotone convergence in the dual function, so we have some flexibility in designing the projection-domain preconditioners  $\mathbf{P}_k$ . One interesting choice is

$$\mathbf{P}_k^{\text{IFBP}} = \mathbf{W}_k^{-1} \mathbf{F}_k \mathbf{W}_k^{-1}, \quad (19)$$

where the  $\mathbf{F}_k$  are positive-definite ramp filters. If  $\Gamma^{-1}$  is chosen to correct for the nonuniform spatial sampling of the CT system, (18) resembles an iterated filtered backprojection algorithm for solving the  $\mathbf{x}$  update subproblem (7):

$$\mathbf{x}^{(m+\frac{k+1}{K})} = \mathbf{x}^{(m+\frac{k}{K})} - \alpha_k^{(m)} \Gamma^{-1} \mathbf{A}_k' \mathbf{F}_k (\mathbf{A}_k \mathbf{x} - \mathbf{y}_k - \mathbf{u}_k). \quad (20)$$

### D. Simplified step size computation

As the number of groups increases, the forward- and back-projections are no longer the only considerable computational costs of the dual group coordinate ascent algorithm (17). In particular the step size computation, which includes an inner product on two image-sized vectors, becomes relatively computationally expensive. To mitigate this we replace the step size computation with a minorize-maximize step. Let  $\mathbf{M}_k \succeq \mathbf{A}_k \Gamma^{-1} \mathbf{A}_k'$  be diagonal. All the entries of  $\mathbf{A}_k$  and  $\Gamma$  are nonnegative, so we use the easily-computed majorizer [1]:

$$\mathbf{M}_k = \text{diag}_i \{ [\mathbf{A}_k \Gamma^{-1} \mathbf{A}_k']_i \}. \quad (21)$$

The following modified dual ascent algorithm performs no inner products and only a few diagonal matrix multiplications:

$$\begin{cases} \mathbf{p}_k^{(m)} &= \mathbf{W}_k \left( \mathbf{y} + \mathbf{u}_k^{(m)} \right. \\ &\quad \left. + \mathbf{A}_k \left( \Gamma^{-1} \gamma^{(m+\frac{k}{K})} - \mathbf{v}^{(n)} + \boldsymbol{\eta}^{(n)} \right) \right), \\ \mathbf{q}_k^{(m)} &= -(\mathbf{W}_k + \mathbf{W}_k \mathbf{M}_k \mathbf{W}_k)^{-1} \mathbf{p}_k^{(m)}, \\ \mathbf{u}_k^{(m+1)} &= \mathbf{u}_k^{(m)} + \mathbf{q}_k^{(m)}, \\ \gamma^{(m+\frac{k+1}{K})} &= \gamma^{(m+\frac{k}{K})} + \mathbf{A}_k' \mathbf{W}_k \mathbf{q}_k^{(m)}. \end{cases} \quad (22)$$

This algorithm is more practical than (17) when the number of groups is very large (*e.g.*, one group for each view). We do not expect that using a majorizer will considerably slow convergence, because  $\mathbf{A}_k \Gamma^{-1} \mathbf{A}_k'$  shrinks as the number of views in the  $k$ th group decreases.

## IV. PRELIMINARY EXPERIMENTS

As a preliminary experiment, we instantiated a  $1024 \times 1024 \times 192$ -pixel XCAT phantom and generated simulated data for an axial scan with a GE Lightspeed scanner [17] with 888 channels, 64 rows, and 984 views. We simulated Poisson noise and reconstructed onto a  $512 \times 512 \times 96$ -pixel grid without the nonnegativity constraint and using a 26-neighbor edge-preserving regularizer with the smooth Fair potential. All calculations were performed on an NVIDIA Tesla C2050.

The proposed variable splitting algorithm used  $\Gamma = \text{diag}_j \{ [\mathbf{A}' \mathbf{W} \mathbf{A}]_{jj} \}$ . We solved the denoising subproblem with group coordinate descent [8] and performed the  $\mathbf{x}$  update with the tomography solver using one view per group and the simplified step size calculation (22). We looped through all the views in the tomography algorithm once per outer iteration and visited the views in random order.

To provide a preliminary comparison, we plotted cost function against iteration for the proposed algorithm, ordered subsets with separable quadratic surrogates Nesterov's 1983 first-order acceleration [6] with 8, 10, 12 and 16 subsets. All algorithms perform one forward- and back-projection per outer iteration; each algorithm took approximately one minute per (outer) iteration. See Figure 1. The proposed algorithm converges faster than all the accelerated OS algorithms. Figure 3 illustrates the cost function of each algorithm as a function of iteration, and Figure 2 shows the center slice of  $\mathbf{x}^{(5)}$  after five iterations of the proposed algorithm.

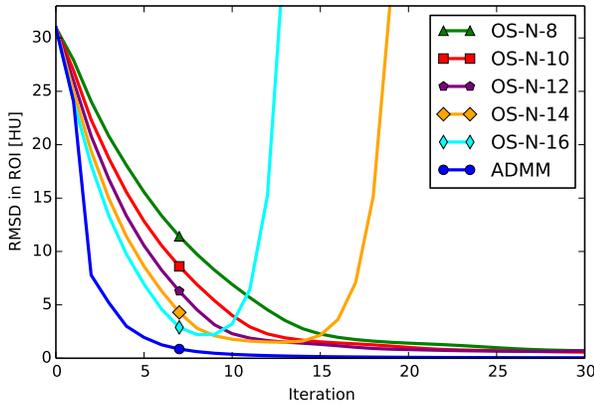


Fig. 1: Root mean squared difference of the proposed algorithm and OS-SQS-Nesterov by iteration to a converged reference image. The proposed algorithm converges considerably more quickly than OS-SQS-Nesterov with any tested number of subsets.

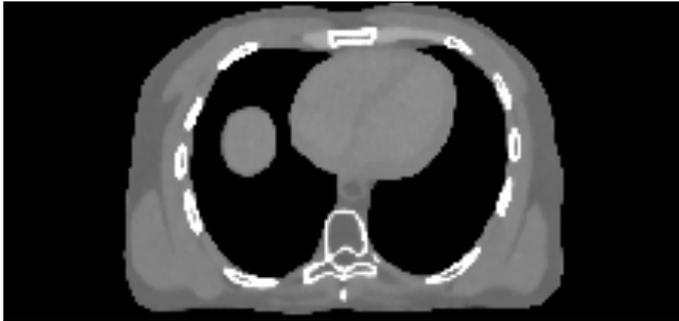


Fig. 2: Image after 5 iterations of proposed algorithm; center slice shown in a [800,1200] HU window.

## V. CONCLUSIONS AND FUTURE WORK

We proposed a novel duality-based method to solve an inner quadratic tomography subproblem for a variable splitting algorithm. The reconstruction algorithm appears to converge rapidly in preliminary experiments, and we intend to perform a more thorough investigation of its performance.

The duality-based approach to the tomography problem presented here is applicable beyond the simple splitting-based algorithm in this paper, and it offers an efficient solution to the challenging 3D “tomography problem” that appears in many splitting-based reconstruction algorithms. Future work will explore more efficient methods to solve the projection-domain dual problem.

## REFERENCES

- [1] H. Erdoğan and J. A. Fessler. Ordered subsets algorithms for transmission tomography. *Phys. Med. Biol.*, 44(11):2835–51, November 1999.
- [2] J. Eckstein and D. P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(1-3):293–318, April 1992.
- [3] J. A. Fessler and S. D. Booth. Conjugate-gradient preconditioning methods for shift-variant PET image reconstruction. *IEEE Trans. Im. Proc.*, 8(5):688–99, May 1999.

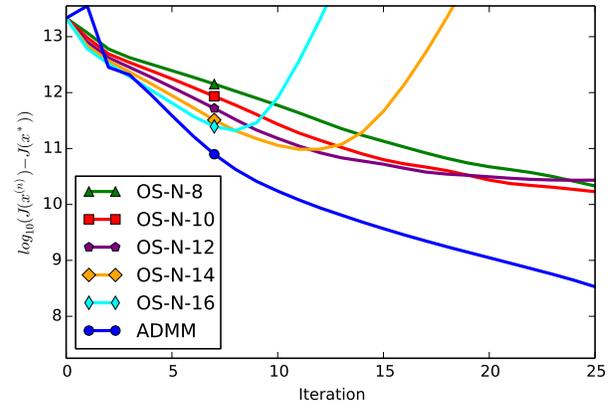


Fig. 3: Cost function values of the proposed algorithm and OS-SQS-Nesterov by iteration relative to the converged reference.

- [4] L. Fu, Z. Yu, J-B. Thibault, B. D. Man, M. G. McGaffin, and J. A. Fessler. Space-variant channelized preconditioner design for 3D iterative CT reconstruction. In *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med.*, pages 205–8, 2013.
- [5] H. M. Hudson and R. S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans. Med. Imag.*, 13(4):601–9, December 1994.
- [6] D. Kim, S. Ramani, and J. A. Fessler. Accelerating X-ray CT ordered subsets image reconstruction with Nesterov’s first-order methods. In *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med.*, pages 22–5, 2013.
- [7] M. McGaffin and J. A. Fessler. Sparse shift-varying FIR preconditioners for fast volume denoising. In *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med.*, pages 284–7, 2013.
- [8] M. G. McGaffin and J. A. Fessler. Fast edge-preserving image denoising via group coordinate descent on the GPU. In *Proc. SPIE 9020 Computational Imaging XII*, page 90200P, 2014.
- [9] M. G. McGaffin, S. Ramani, and J. A. Fessler. Reduced memory augmented Lagrangian algorithm for 3D iterative X-ray CT image reconstruction. In *Proc. SPIE 8313 Medical Imaging 2012: Phys. Med. Im.*, page 831327, 2012.
- [10] H. Nien and J. A. Fessler. Combining augmented Lagrangian method with ordered subsets for X-ray CT image reconstruction. In *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med.*, pages 280–3, 2013.
- [11] H. Nien and J. A. Fessler. Splitting-based statistical X-ray CT image reconstruction with blind gain correction. In *Proc. SPIE 8668 Medical Imaging 2013: Phys. Med. Im.*, page 86681J, 2013.
- [12] L. Pfister and Y. Bresler. Model-based iterative tomographic reconstruction with adaptive sparsifying transforms. In *Proc. SPIE 9020 Computational Imaging XII*, 2014.
- [13] S. Ramani and J. A. Fessler. A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction. *IEEE Trans. Med. Imag.*, 31(3):677–88, March 2012.
- [14] M. Sion. On general minimax theorems. *Pacific J. Math.*, 8(1):171–6, 1958.
- [15] J-B. Thibault, K. Sauer, C. Bouman, and J. Hsieh. A three-dimensional statistical approach to improved image quality for multi-slice helical CT. *Med. Phys.*, 34(11):4526–44, November 2007.
- [16] P. Tseng and S. Yun. A coordinate gradient descent method for nonsmooth separable minimization. *Mathematical Programming*, 117(1-2):387–423, 2009.
- [17] J. Wang, T. Li, H. Lu, and Z. Liang. Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose X-ray computed tomography. *IEEE Trans. Med. Imag.*, 25(10):1272–83, October 2006.
- [18] Z. Yu, L. Fu, D. Pal, J-B. Thibault, C. A. Bouman, and K. D. Sauer. Nested loop algorithm for parallel model based iterative reconstruction. In *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med.*, pages 197–200, 2013.