Fast Splitting-Based Ordered-Subsets X-Ray CT Image Reconstruction

Hung Nien and Jeffrey A. Fessler

Abstract—Using non-smooth regularization in X-ray computed tomography (CT) image reconstruction has become more popular these days due to the recent resurgence of the classic augmented Lagrangian (AL) methods in fields such as total-variation (TV) denoising and compressed sensing (CS). For example, undersampling projection views is one way to reduce radiation dose in CT scans; however, this causes strong streak artifacts in FBP images that degrade image quality. To overcome this problem, the split Bregman (SB) method, an alias of the AL method in the context of ℓ_1-regularized image reconstruction problems, has been investigated using strong non-smooth TV and sparsity regularizations. Unfortunately, existing SB-based methods are slow due to the iterative updates for the challenging inner least-squares problem. This paper proposes to solve X-ray CT image reconstruction problems with TV or sparsity regularization using a splitting-based ordered-subsets (OS) algorithm, split OS-LALM, and evaluates the proposed algorithm using a few-view X-ray CT image reconstruction problem with TV regularization. Experimental results show that the proposed algorithm significantly accelerates the convergence of X-ray CT image reconstruction with non-smooth TV regularization over the standard (linearized) SB method and demonstrates the effectiveness of OS acceleration with splitting-based algorithms.

I. INTRODUCTION

X-ray computed tomography (CT) is a non-invasive medical procedure that images the attenuation properties, such as the density distribution, of the body. It is incredibly useful and important in the medical community, while the growing concern about radiation dose from CT scans comes from the increased use of CT procedures. In the past three decades, the average American's dose from medical exposure (not including radiotherapy) has increased from 0.54 mSv in 1982 to 3.0 mSv in 2006, where CT procedures account for about half of the collective dose from all medical procedures [1]. Compared with the natural background yearly dose of 3.6 mSv, the standard radiation dose used currently can increase the possible risk of cancers, especially for body screening with multiple scans.

Using fewer projection views in a CT scan is one way to reduce radiation dose, but such undersampling causes strong streak artifacts that degrade FBP image quality. To reduce streak artifacts, the split Bregman (SB) method [2], a fast convex optimization method using variable splitting technique, has been investigated using total-variation (TV) and sparsity regularizations. Unfortunately, existing SB-based

methods for few-view CT image reconstruction, especially for 3D CT, can be slow due to the challenging inner least-squares problem with a highly shift-variant Hessian [3–5]. For example, [3] suggested solving the inner least-squares problem of the SB method using up to 100 iterations of the conjugate gradient (CG) method, that is, hundreds of forward/back-projection pairs for a single outer-loop image update! Although the forward/back-projection operations in few-view CT are less time-consuming than in clinical CT, using hundreds of forward/back-projection pair for a single image update remains undesirable.

To solve the problem of the difficult inner least-squares problem in SB methods, Ramani *et al.* [6] introduced an additional auxiliary variable that separates the shift-variant and approximate shift-invariant parts of the statistically weighted quadratic data-fitting term so that one can find an appropriate circulant preconditioner for the better-conditioned inner problem and solve the inner problem efficiently using the preconditioned conjugate gradient (PCG) method. The acceleration is significant in 2D CT [6]; however, in 3D CT, due to the cone-beam geometry, it is much harder to find a good preconditioner for the inner least-squares problem, and the method in [6] has yet to achieve the same acceleration as in 2D CT.

Considering the same variable splitting scheme as in [6], this paper proposes to solve X-ray CT image reconstruction problem with TV or sparsity regularization using a linearized augmented Lagrangian (AL) method [7, 8] that replaces the difficult inner least-squares problem by a simple majorization-minimization procedure (a gradient descent that guarantees monotone decreasing of the cost value) and more importantly, is suitable for ordered-subsets (OS) [9] acceleration. For instance, suppose M ordered subsets are used for acceleration. The proposed splitting-based OS algorithm takes roughly $1/M$ forward/back-projection pair for a single image update! Therefore, compared with existing SB methods, we perform many more image updates in a given reconstruction time, leading to faster convergence.

The remainder of the paper is organized as follows. Section II introduces the problem formulation and derives the proposed splitting-based OS algorithm for solving regluarized least-squares problems. Section III considers solving few-view X-ray CT image reconstruction problem with penalized weighted least-squares (PWLS) criterion using the proposed algorithm and reports the experimental results comparing a linearized SB method with our method. Finally, we draw conclusions in Section IV.

II. PROPOSED METHOD

A. Split OS-LALM: OS-LALM with an additional split

Consider a regularized least-squares problem:

$$\hat{\mathbf{x}} \in \arg\min_{\mathbf{x}\in\Omega}\left\{\tfrac{1}{2}\left\|\mathbf{y}-\mathbf{A}\mathbf{x}\right\|_2^2 + \Phi(\boldsymbol{\Theta}\mathbf{x})\right\}, \qquad (1)$$

where $\mathbf{A}$ is the system matrix, $\mathbf{y}$ is the noisy measurement, $\boldsymbol{\Theta}$ is an analysis regularization matrix, Φ is some convex (and possibly non-smooth) potential function, and Ω denotes the convex set for a box constraint (usually the non-negativity constraint) on $\mathbf{x}$. For example, in (anisotropic) TV-regularized image restoration problems, $\boldsymbol{\Theta}$ is a finite difference matrix, and Φ is an ℓ_1 norm, probably with some weighting. The minimization problem (1) is non-trivial in general since $\boldsymbol{\Theta}$ might not be an identity matrix, and Φ can be non-smooth. One typical way to solve this problem is to use the SB method [2] that introduces an auxiliary variable for the vector $\boldsymbol{\Theta}\mathbf{x}$ and decomposes the convex optimization problem into a series of simpler penalized least-squares problems. However, when $\mathbf{A}'\mathbf{A}$ is highly shift-variant, the SB method can be slow due to the iterative inner updates.

To develop a faster algorithm, instead of solving (1) using the SB method, we consider solving an equivalent constrained minimization problem:

$$(\hat{\mathbf{x}}, \hat{\mathbf{u}}, \hat{\mathbf{v}}) \in \arg\min_{\mathbf{x},\mathbf{u},\mathbf{v}}\left\{g(\mathbf{u}) + \Phi(\mathbf{v}) + \iota_\Omega(\mathbf{x})\right\}$$
$$\text{s.t. } \mathbf{u} = \mathbf{A}\mathbf{x}, \mathbf{v} = \boldsymbol{\Theta}\mathbf{x} \quad (2)$$

using the linearized AL method [7, 8]:

$$\begin{cases}
\mathbf{x}^{(k+1)} \in \arg\min_{\mathbf{x}}\left\{\iota_\Omega(\mathbf{x}) + \breve{\theta}_k(\mathbf{x};\mathbf{x}^{(k)}) + \breve{\phi}_k(\mathbf{x};\mathbf{x}^{(k)})\right\} \\
\mathbf{u}^{(k+1)} \in \arg\min_{\mathbf{u}}\left\{g(\mathbf{u}) + \frac{\rho}{2}\left\|\mathbf{A}\mathbf{x}^{(k+1)}-\mathbf{u}-\mathbf{d}^{(k)}\right\|_2^2\right\} \\
\mathbf{v}^{(k+1)} \in \arg\min_{\mathbf{v}}\left\{\Phi(\mathbf{v}) + \frac{\eta}{2}\left\|\boldsymbol{\Theta}\mathbf{x}^{(k+1)}-\mathbf{v}-\mathbf{e}^{(k)}\right\|_2^2\right\} \\
\mathbf{d}^{(k+1)} = \mathbf{d}^{(k)} - \mathbf{A}\mathbf{x}^{(k+1)} + \mathbf{u}^{(k+1)} \\
\mathbf{e}^{(k+1)} = \mathbf{e}^{(k)} - \boldsymbol{\Theta}\mathbf{x}^{(k+1)} + \mathbf{v}^{(k+1)},
\end{cases}$$
$$(3)$$

where $g(\mathbf{u}) \triangleq \frac{1}{2}\left\|\mathbf{y}-\mathbf{u}\right\|_2^2$, ι_Ω is the characteristic function of the convex set Ω that handles the box constraint on $\mathbf{x}$, $\mathbf{d}$ and $\mathbf{e}$ are the scaled Lagrange multipliers of the split variables $\mathbf{u}$ and $\mathbf{v}$, respectively, and $\rho > 0$ and $\eta > 0$ are the corresponding AL penalty parameters. The functions $\breve{\theta}_k(\mathbf{x};\mathbf{x}^{(k)})$ and $\breve{\phi}_k(\mathbf{x};\mathbf{x}^{(k)})$ are two separable quadratic surrogate (SQS) functions that majorize the quadratic AL penalty terms

$$\theta_k(\mathbf{x}) \triangleq \frac{\rho}{2}\left\|\mathbf{A}\mathbf{x}-\mathbf{u}^{(k)}-\mathbf{d}^{(k)}\right\|_2^2 \qquad (4)$$

and

$$\phi_k(\mathbf{x}) \triangleq \frac{\eta}{2}\left\|\boldsymbol{\Theta}\mathbf{x}-\mathbf{v}^{(k)}-\mathbf{e}^{(k)}\right\|_2^2 \qquad (5)$$

at $\mathbf{x} = \mathbf{x}^{(k)}$, respectively. Let L_1 and L_2 denote the maximum eigenvalues of $\mathbf{A}'\mathbf{A}$ and $\boldsymbol{\Theta}'\boldsymbol{\Theta}$, respectively, it follows that

$$\begin{cases}
\breve{\theta}_k(\mathbf{x};\mathbf{x}^{(k)}) \\
\quad \propto \frac{\rho}{2t_1}\left\|\mathbf{x}-\left(\mathbf{x}^{(k)}-t_1\mathbf{A}'\left(\mathbf{A}\mathbf{x}^{(k)}-\mathbf{u}^{(k)}-\mathbf{d}^{(k)}\right)\right)\right\|_2^2 \\
\breve{\phi}_k(\mathbf{x};\mathbf{x}^{(k)}) \\
\quad \propto \frac{\eta}{2t_2}\left\|\mathbf{x}-\left(\mathbf{x}^{(k)}-t_2\boldsymbol{\Theta}'\left(\boldsymbol{\Theta}\mathbf{x}^{(k)}-\mathbf{v}^{(k)}-\mathbf{e}^{(k)}\right)\right)\right\|_2^2,
\end{cases}$$
$$(6)$$

where $t_1 \triangleq 1/L_1$ and $t_2 \triangleq 1/L_2$. The majorizations remove the entanglement of $\mathbf{x}$ introduced by $\mathbf{A}$ and $\boldsymbol{\Theta}$, leading to simple inner updates in (3) using proximal mappings.

As can be seen in (3), introducing an additional split variable $\mathbf{v}$ only modestly changes the updates from the one-split linearized AL iterates [7, 8]. Letting $h_k \triangleq \iota_\Omega + \breve{\phi}_k$, the two-split linearized AL iterates (3) become the one-split linearized AL iterates with an iteration-dependent regularization term h_k, where the effect of h_k is fully determined by the $\mathbf{v}$- and $\mathbf{e}$-updates in (3)! Hence, we can easily rewrite the two-split linearized AL iterates (3) to the two-split gradient-based linearized AL iterates:

$$\begin{cases}
\mathbf{s}^{(k+1)} = \rho\nabla\ell(\mathbf{x}^{(k)}) + (1-\rho)\mathbf{g}^{(k)} \\
\mathbf{x}^{(k+1)} \in \text{prox}_{(\rho^{-1}t_1)h_k}\left(\mathbf{x}^{(k)}-\left(\rho^{-1}t_1\right)\mathbf{s}^{(k+1)}\right) \\
\mathbf{g}^{(k+1)} = \frac{\rho}{\rho+1}\nabla\ell(\mathbf{x}^{(k+1)}) + \frac{1}{\rho+1}\mathbf{g}^{(k)} \\
\mathbf{v}^{(k+1)} \in \text{prox}_{\eta^{-1}\Phi}\left(\boldsymbol{\Theta}\mathbf{x}^{(k+1)}-\mathbf{e}^{(k)}\right) \\
\mathbf{e}^{(k+1)} = \mathbf{e}^{(k)} - \boldsymbol{\Theta}\mathbf{x}^{(k+1)} + \mathbf{v}^{(k+1)},
\end{cases}$$
$$(7)$$

where ℓ denotes the quadratic data-fitting term in (1), and prox_f denotes the proximal mapping of f defined as:

$$\text{prox}_f(\mathbf{z}) \triangleq \arg\min_{\mathbf{x}}\left\{f(\mathbf{x}) + \tfrac{1}{2}\left\|\mathbf{x}-\mathbf{z}\right\|_2^2\right\}. \qquad (8)$$

Since both ι_Ω and $\breve{\phi}_k$ are separable, the $\mathbf{x}$-update of the two-split gradient-based linearized AL iterates (7) has a closed-form solution:

$$\mathbf{x}^{(k+1)} = \left[\mathbf{x}^{(k)} - \frac{1}{\rho L_1 + \eta L_2}\left(\mathbf{s}^{(k+1)} + \boldsymbol{\sigma}^{(k+1)}\right)\right]_\Omega, \quad (9)$$

where $[\cdot]_\Omega$ denotes an operator that projects a vector onto Ω, and

$$\boldsymbol{\sigma}^{(k+1)} \triangleq \eta\boldsymbol{\Theta}'\left(\boldsymbol{\Theta}\mathbf{x}^{(k)}-\mathbf{v}^{(k)}-\mathbf{e}^{(k)}\right) \qquad (10)$$

is the search direction attributed to the regularization term. Finally, the two-split gradient-based linearized AL method (7) is an extension of the one-split gradient-based linearized AL method, so we can accelerate it by using OS and the deterministic downward continuation approach proposed in [7, 8]. For the OS version, we replace the gradients in (7) with the subset gradients $M\nabla\ell_m$ for $m = 1, \ldots, M$, where $\ell_1, \ldots, \ell_M$ are M smaller quadratic functions that satisfy $\ell = \ell_1 + \cdots + \ell_M$ and the "subset balance condition" [9]:

$$\nabla\ell(\mathbf{x}) \approx M\nabla\ell_1(\mathbf{x}) \approx \cdots \approx M\nabla\ell_M(\mathbf{x}). \qquad (11)$$

When OS is used for acceleration, we call our proposed algorithm split OS-LALM, by an analogy of the SB method.

B. Applications

In this paper, we consider a regularized least-squares problem with a general composite convex regularizer $\Phi(\boldsymbol{\Theta}\mathbf{x})$. As mentioned before, when $\boldsymbol{\Theta}$ is a finite difference matrix and Φ is a weighted ℓ_1 norm, (1) becomes a TV-regularized image reconstruction problem. In this case, the $\mathbf{v}$-update in (7) can be solved efficiently uisng soft-thresholding, and the constant $L_2 \triangleq \lambda_{\max}(\boldsymbol{\Theta}'\boldsymbol{\Theta}) = 4d$, where d denotes the number of neighbors we considered in the finite difference operator. Furthermore, when $\boldsymbol{\Theta}$ is the discrete framelet transform matrix [10] and Φ is an ℓ_1 norm, (1) becomes a frame-based image

reconstruction problem [11]. In this case, the **v**-update can also be solved using soft-thresholding, and $L_2 = 1$ because the discrete framelet is a tight frame. In fact, the proposed algorithm is even more general. For example, consider introducing one more split for the box constraint on **x**. In this case, we can use non-separable (but probably tighter) quadratic surrogate functions with non-diagonal (e.g., circulant) Hessian matrices to majorize θ_k and ϕ_k in (3) because the additional split variable takes care of the box constraint.

III. EXPERIMENTAL RESULTS

This section evaluates our proposed algorithm (7) using the statistically weighted few-view X-ray CT image reconstruction problem with TV regularization:

$$\hat{\mathbf{x}} \in \arg\min_{\mathbf{x} \in \Omega} \left\{ \tfrac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2_{\mathbf{W}} + \mathsf{TV}(\mathbf{x}) \right\}, \qquad (12)$$

where **A** is the system matrix of a CT scan, **y** is the noisy sinogram, **W** is the statistical weighting matrix, and $\mathsf{TV}(\cdot)$ denotes an anisotropic TV regularization term. To solve (12) using the proposed algorithm, we simply replace **A** and **y** in (1) by the weighted forward projection operator $\mathbf{W}^{1/2}\mathbf{A}$ and the weighted noisy sinogram $\mathbf{W}^{1/2}\mathbf{y}$, respectively, and let $\Theta \triangleq \mathbf{C}$ denote a finite difference matrix and Φ denote a weighted ℓ_1 norm.

Computing L_1, the maximum eigenvalue of $\mathbf{A}'\mathbf{W}\mathbf{A}$, is sometimes impractical because the power iteration might take hundreds of forward/back-projections for finding that number, while the number changes with different weighting matrix **W**, even for a fixed geometry. In practice, we simply use $\mathbf{L}_1 \triangleq \mathrm{diag}\{\mathbf{A}'\mathbf{W}\mathbf{A}\mathbf{1}\}$ to construct the SQS of the quadratic AL penalty term θ_k [7, 8]. This also provides voxel-dependent step sizes for image updates in (9). One can also generalize L_2 to a diagonal matrix $\mathbf{L}_2$ by considering a "weighted" quadratic AL penalty term of the second split [12]. However, in this paper, we just use $L_2 = 4d$ for simplicity.

We reconstructed a $512 \times 512 \times 122$ image from an undersampled chest axial CT scan. The size of the original sinogram is $888 \times 64 \times 642$ (half scan), and we uniformly undersampled the number of projection views from 642 to 81 (about 12.6% of projection views are used for reconstruction). Instead of using the standard SB-based method [3], we used a linearized SB method as the baseline reconstruction method because it has no iterative inner updates and is much easier for imposing box constraints on **x**. Let **OS-LALM-M-ρ-η** denote the proposed algorithm using M subsets with AL penalty parameters ρ and η, where "$\rho = \mathrm{c}$" denotes the deterministic downward continuation [7, 8]. When $\rho = 1$, the proposed algorithm happens to be the linearized SB method [13].

The number of subsets M is varied from 1 to 5 for investigating different amounts of OS acceleration. The AL penalty parameter η was hand-tuned for fastest convergence and remained the same throughout the experiment for fair comparison. Intuitively, η determines the step sizes for image updates in (9), especially when the deterministic downward continuation approach is used. Empirically, choosing ηL_2 that is about 2% to 10% of L_1 (or the median of the diagonal entries of $\mathbf{L}_1$) usually exhibits fast convergence of the proposed

(a)

(b)
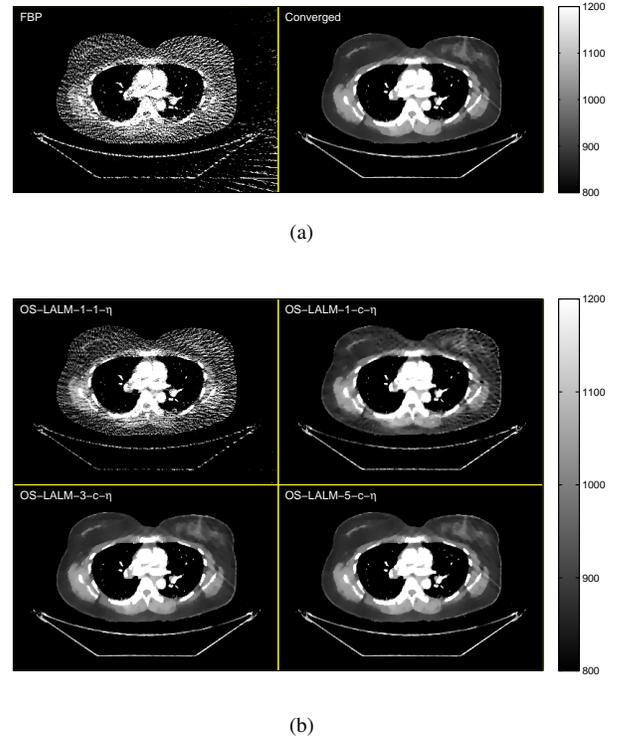
Fig. 1: Chest CT: cropped images [displayed from 800 to 1200 HU] from the central transaxial plane of (a) the initial FBP image $\mathbf{x}^{(0)}$ and the reference reconstruction $\mathbf{x}^\star$, and (b) the reconstructed images $\mathbf{x}^{(100)}$ at the 100th iteration using the proposed algorithm with different AL penalty parameters. When $\rho = 1$, the proposed algorithm reverts to the linearized SB method.

algorithm. Finally, the total number of iterations is set to be 100. In this case, 100 undersampled forward/back-projection pairs, about 13 full forward/back-projection pairs, are used for the reconstruction.

Figure 1 shows the initial FBP image and the almost converged reference reconstruction together with the reconstructed images of the proposed algorithm with different parameters. As can be seen in Figure 1(a), the initial FBP image exhibits strong streak artifacts due to the undersampled projection views, and these streak artifacts are reduced significantly in the reference reconstruction by applying TV regularization. Figure 1(b) demonstrates the effectiveness of our proposed algorithm. With the deterministic downward continuation (i.e., $\rho = \mathrm{c}$), the proposed algorithm shows less streak artifacts in the reconstructed images, and the reduction is more effective for larger M. Figure 2 shows the convergence rate curves (RMS differences between the reconstructed image $\mathbf{x}^{(k)}$ and the reference reconstruction $\mathbf{x}^\star$ as a function of iteration) of the proposed algorithm with different AL penalty parameters. As can be seen in Figure 2, the proposed algorithm shows substantial acceleration with continuation and ordered subsets. For example, the RMS difference of OS-LALM-5-c-η reaches 10 HU within 50 iterations, while the linearized SB method (OS-LALM-1-1-η) is still far away from the optimum (about
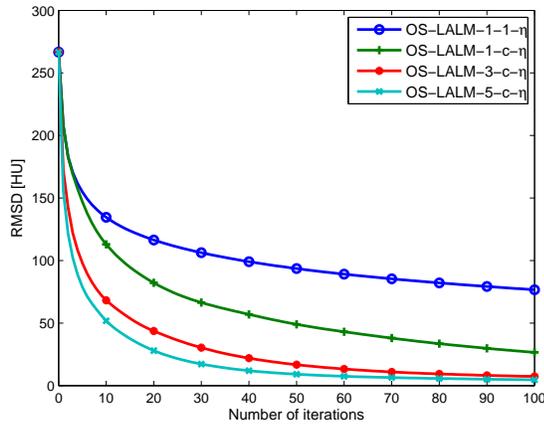
Fig. 2: Chest CT: RMS differences between the reconstructed image $\mathbf{x}^{(k)}$ and the reference reconstruction $\mathbf{x}^\star$ as a function of iteration using the proposed algorithm with different AL penalty parameters. When $\rho = 1$, the proposed algorithm reverts to the linearized SB method.

90 HU in RMS difference). Finally, note that algorithms with OS typically are not convergent. The algorithms enter a "limit cycle" in which updates stop approaching the optimum. However, as can be seen in Figure 2, the cyan curve (OS-LALM-5-c-η) is below 5 HU at the 100th iteration and keeps decreasing after that. This demonstrates the gradient error tolerance of our proposed splitting-based OS algorithm.

IV. CONCLUSION

In this paper, we proposed a splitting-based ordered-subsets (OS) algorithm, split OS-LALM, for solving weighted least-squares X-ray computed tomography (CT) image reconstruction problems with a general composite convex regularizer. To demonstrate our proposed algorithm, we investigated solving a few-view X-ray CT image reconstruction problem with total-variation (TV) regularization. Experimental results showed that the proposed algorithm exhibits fast convergence rate and excellent gradient error tolerance when OS is used for acceleration. The same technique can also be applied to 3D clinical CT with complicated (e.g., non-smooth) regularization terms.

REFERENCES

[1] "Ionizing radiation exposure of the population of the United States," Tech. Rep. 160, National Council on Radiation Protextion and Measurements (NCRP), 2009.

[2] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM J. Imaging Sci.*, vol. 2, no. 2, pp. 323–43, 2009.

[3] B. Vandeghinste, B. Goossens, J. D. Beenhouwer, A. Pizurica, W. Philips, S. Vandenberghe, and S. Staelens, "Split-Bregman-based sparse-view CT reconstruction," in *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med*, pp. 431–4, 2011.

[4] B. Vandeghinste, B. Goossens, R. Van Holen, C. Vanhove, A. Pizurica, S. Vandenberghe, and S. Staelens, "Combined shearlet and TV regularization in sparse-view CT reconstruction," in *Proc. 2nd Intl. Mtg. on image formation in X-ray CT*, pp. 37–40, 2012.

[5] Y. Li, P. T. Lauzier, J. Tang, and G.-H. Chen, "Bregman regularized statistical image reconstruction method and application to prior image constrained compressed sensing (PICCS)," in *Proc. SPIE 8668 Medical Imaging 2013: Phys. Med. Im.*, p. 86683A, 2013.

[6] S. Ramani and J. A. Fessler, "A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 31, pp. 677–88, Mar. 2012.

[7] H. Nien and J. A. Fessler, "Accelerating ordered-subsets X-ray CT image reconstruction using the linearized augmented Lagrangian framework," in *Proc. SPIE 9033 Medical Imaging 2014: Phys. Med. Im.*, p. 903332, 2014.

[8] H. Nien and J. A. Fessler, "Fast X-ray CT image reconstruction using the linearized augmented Lagrangian method with ordered subsets," *arXiv: 1402.4381*, 2014. Submitted to *IEEE Trans. Med. Imag.*.

[9] H. Erdoğan and J. A. Fessler, "Ordered subsets algorithms for transmission tomography," *Phys. Med. Biol.*, vol. 44, pp. 2835–51, Nov. 1999.

[10] I. Daubechies, B. Han, A. Ron, and Z. Shen, "Framelets: MRA-based constructions of wavelet frames," *Appl. Comput. Harmon. Anal.*, vol. 14, pp. 1–46, Jan. 2003.

[11] J. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," *SIAM J. Multiscale Model. Simul.*, vol. 8, no. 2, pp. 337–69, 2009.

[12] H. Nien and J. A. Fessler, "Combining augmented Lagrangian method with ordered subsets for X-ray CT reconstruction," in *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med*, pp. 280–3, 2013.

[13] H. Nien and J. A. Fessler, "A convergence proof of the split Bregman method for regularized least-squares problems," *arXiv: 1402.4371*, 2014.