

Convergent Iterative CT Reconstruction With Sparsity-Based Regularization

Sathish Ramani and Jeffrey A. Fessler

Abstract—Statistical image reconstruction for X-ray CT can provide improved image quality at reduced patient doses. An important component of statistical reconstruction methods is the regularizer. There has been increased interest in sparsity-based regularization, typically using ℓ_1 norms. The non-smooth nature of these regularizers is a challenge for iterative optimization methods and often causes slow convergence. Recently there has been renewed interest in augmented Lagrangian methods for such optimization problems, with certain variable splitting approaches [1] including the alternating direction method of multipliers (ADMM) [2]. Such algorithms have been applied successfully to image restoration problems. This paper describes an ADMM algorithm for iterative CT reconstruction using a regularized, weighted least-squares (WLS) cost function. Not only does ADMM accommodate non-smooth regularizers, but also by choosing an appropriate variable splitting it uses an inner iteration that is suitable for preconditioning using a circulant matrix (FFT) based on a kind of cone filter. Simulation results show that the proposed ADMM converges and that the cone filter preconditioner accelerates convergence.

I. INTRODUCTION

One approach to statistical image reconstruction in X-ray CT uses a regularized weighted least-squares (WLS) cost function of the following form [3]:

$$\begin{aligned} \mathbf{P0} : \hat{\mathbf{x}} &= \arg \min_{\mathbf{x}} J(\mathbf{x}), \\ J(\mathbf{x}) &= \sum_{i=1}^M \frac{w_i}{2} (y_i - [\mathbf{A}\mathbf{x}]_i)^2 + \Psi(\mathbf{x}), \end{aligned} \quad (1)$$

where $\mathbf{x} = (x_1, \dots, x_N)$ denotes the vector of N voxels of the unknown 3D image, \mathbf{y} denotes the X-ray CT projection data, w_i denotes the statistical weighting associated with the i th ray, for $i = 1, \dots, M$, M is the number of rays, \mathbf{A} is the $M \times N$ system matrix and $\Psi(\mathbf{x})$ is an edge-preserving regularizer that controls noise while attempting to preserve spatial resolution. The forward projection operation is $[\mathbf{A}\mathbf{x}]_i = \sum_{j=1}^N a_{ij}x_j$. There are two challenges that arise when developing algorithms for minimizing this cost function. The first challenge is that the Hessian of the data-fit term, $\mathbf{A}^T\mathbf{W}\mathbf{A}$, is not shift invariant due to the statistical weighting $\mathbf{W} = \text{diag}\{w_i\}$. The second challenge is that strongly edge preserving regularizers, such as those based on total variation (TV) [4] or those based on sparsity [5] are non-smooth, precluding the use of conventional gradient-based optimization methods. The algorithm described in this paper overcomes both of these challenges.

Many types of iterative algorithms have been proposed for minimizing cost functions like J , including iterative coordinate descent (ICD) methods [3], block-based coordinate descent [6], ordered-subsets (OS) algorithms based on separable quadratic surrogates (SQS) [7] and preconditioned conjugate-gradient (PCG) methods [8]. For fast computation on multi-processor computers, PCG-type methods appear to be the most amenable to efficient parallelization because they update all voxels simultaneously using all measurements. However, developing suitable preconditioners is challenging for X-ray CT because of the form of the Hessian:

$$\nabla^2 J(\mathbf{x}) = \mathbf{A}^T\mathbf{W}\mathbf{A} + \nabla^2\Psi(\mathbf{x}).$$

The enormous dynamic range of the weights $\{w_i\}$ causes the Hessian $\mathbf{A}^T\mathbf{W}\mathbf{A}$ of the data-fit term to be highly shift variant [8]. Clinthorne *et al.* [9] showed that for *unweighted* least-squares reconstruction, where $\mathbf{W} = \mathbf{I}_M$ (identity matrix of size M), one can precondition $\mathbf{A}^T\mathbf{A}$ using FFTs with a kind of cone filter. This cone filter amplifies high spatial frequencies, helping to accelerate convergence. But that cone filter is ineffective for PCG in the WLS case [8]. Delaney and Bresler [10] considered a very special type of shift invariant weighting and also demonstrated accelerated convergence, but for low-dose X-ray CT the appropriate statistical weighting does not satisfy the assumptions in [10]. Shift-variant preconditioners based on multiple FFTs were proposed in [8] for 2D transmission tomography, but never became popular due to their complexity and never were investigated for 3D problems. As described in the next section, the ADMM approach proposed here uses a kind of variable splitting that separates the weighting matrix \mathbf{W} from the system matrix \mathbf{A} , leading to an inner iteration step that requires solving a system of equations of the form $\mathbf{A}^T\mathbf{A} + \alpha\mathbf{R}^T\mathbf{R}$ for some regularization operator \mathbf{R} . This system is nearly shift invariant, so it is well suited to the kind of circulant preconditioner proposed in [9] based on FFTs and a type of regularized cone filter.

Another way to introduce a cone filter is the iterative FBP approach [11], [12]. Initially these algorithms “converge” rapidly compared to CG methods, but typically they do not have any theoretical convergence properties and “too many” iterations lead to undesirably noisy images. Furthermore it is unclear how include regularization while ensuring convergence. The ADMM proposed in this paper can use cone filters in the context of regularized cost functions like (1), while also providing a framework for establishing convergence theoretically.

The challenges described above apply regardless of the form of the regularizer. Additional challenges arise when

one chooses regularizers that are strongly edge preserving. One popular example includes total variation (TV) regularization [4] $\Psi(\mathbf{x}) = \lambda \sum_j \sqrt{\sum_{k \in \mathcal{N}_j} (x_j - x_k)^2}$, where \mathcal{N}_j denotes a neighborhood of the j th voxel. The square root function is non-differentiable at zero, precluding optimization by conventional gradient-based methods, unless the square root is modified by “corner rounding” approximations like $\sqrt{x^2} \approx \sqrt{x^2 + \epsilon}$. Even with such modifications the Hessian of the regularizer $\Psi(\mathbf{x})$ will have very high curvature which can lead to slow convergence rates for conventional gradient-based methods. Regularizers based on sparsity are also challenging, such as $\Psi(\mathbf{x}) = \|\mathbf{R}\mathbf{x}\|_1$, where \mathbf{R} denotes an analysis operator such as those based on wavelets. The ℓ_1 norm again is non-differentiable at zero, complicating optimization. Recently there has been considerable interest in developing optimization methods for image denoising and image restorations problems with such regularizers, as well as for under-sampled image reconstruction problems in MRI [13], [14]. Synthesis formulations have also been studied [5] but recent comparisons suggest that encouraging sparsity via analysis operators is preferable [15]. The proposed ADMM approach, adapted to X-ray CT from [2], accommodates a general class of regularizers including both analysis and synthesis forms and includes “conventional” edge-preserving regularizers [3].

II. ALTERNATING DIRECTION METHOD OF MULTIPLIERS

We consider a general regularizer of the form

$$\Psi(\mathbf{x}) = \lambda \sum_{n=1}^{N_r} w_n \Phi_n \left(\sum_{l=1}^L |[\mathbf{R}_l \mathbf{x}]_n|^m \right), \quad (2)$$

where λ is the regularization parameter, $\{\mathbf{R}_l\}_{l=1}^L$, $L \ll N_r$, are $N_r \times N$ matrices that represent sparsifying-operators (e.g., finite-differencing matrices, forward-transform-matrices corresponding to different sub-bands of a wavelet frame, etc), $\mathbf{R} = [\mathbf{R}_1^T \dots \mathbf{R}_L^T]^T$ and the weights $w_n > 0 \forall n$ are chosen so as to provide (nearly) uniform spatial resolution in the reconstructed output [16]. The above form of regularization includes TV, ℓ_1 -wavelets and edge-preserving regularizers.

To solve (1), we apply a splitting approach with auxiliary variables that separate the different terms in J . We then reformulate **P0** as the following equivalent constrained problem:

$$\mathbf{P1} : \min_{\mathbf{x}, \mathbf{z}} \left\{ f(\mathbf{z}) = \frac{1}{2} \|\mathbf{y} - \mathbf{u}\|_{\mathbf{W}}^2 + \Psi(\mathbf{v}) \right\} \text{ s.t. } \mathbf{z} = \mathbf{C}\mathbf{x}, \quad (3)$$

where $\mathbf{z} \triangleq [\mathbf{u}^T \ \mathbf{v}^T]^T \in \mathbb{R}^{N_1}$ represents the vector of auxiliary constrained variables, $N_1 = M + NL$, $\mathbf{v} = [\mathbf{v}_1^T \dots \mathbf{v}_L^T]^T$ with $\mathbf{v}_l = \mathbf{R}_l \mathbf{x}$, $l = 1, \dots, L$, and $\mathbf{C} \triangleq [\mathbf{A}^T \ \mathbf{R}^T]^T$ is a $N_1 \times N$ constraint matrix. The specific form of \mathbf{C} separates \mathbf{W} , \mathbf{A} and \mathbf{R} thereby simplifying optimization as explained next.

A. Method of Multipliers—Augmented Lagrangian Formalism

We use the classical framework of the method of multipliers [17], specifically the augmented Lagrangian (AL) formalism, for handling the constrained problem **P1**. We first construct an AL function

$$\mathcal{L}(\mathbf{x}, \mathbf{z}, \boldsymbol{\gamma}, \mu) \triangleq f(\mathbf{z}) + \boldsymbol{\gamma}^T \boldsymbol{\Lambda}(\mathbf{z} - \mathbf{C}\mathbf{x}) + \frac{\mu}{2} \|\mathbf{z} - \mathbf{C}\mathbf{x}\|_{\boldsymbol{\Lambda}^2}^2 \quad (4)$$

1. Select $\mathbf{x}^{(0)}$ and $\mu, \nu > 0$ and set $j = 0$
2. Set $\mathbf{u}^{(0)} = \mathbf{A}\mathbf{x}^{(0)}$, $\mathbf{v}^{(0)} = \mathbf{R}\mathbf{x}^{(0)}$, $\boldsymbol{\eta}_{\mathbf{u}}^{(0)} = \boldsymbol{\eta}_{\mathbf{v}}^{(0)} = \mathbf{0}$
- Repeat:**
3. Compute $\mathbf{u}^{(j+1)} = \mathbf{H}_{\mu}^{-1}(\mathbf{W}\mathbf{y} + \mu(\mathbf{A}\mathbf{x}^{(j)} + \boldsymbol{\eta}_{\mathbf{u}}^{(j)}))$
4. Compute $\mathbf{v}^{(j+1)} = \{v_{nl}^{(j+1)}\}$ using (14)
5. Obtain $\mathbf{x}^{(j+1)}$ by applying (PCG) iterations to (12)
6. $\boldsymbol{\eta}_{\mathbf{u}}^{(j+1)} = \boldsymbol{\eta}_{\mathbf{u}}^{(j)} - (\mathbf{u}^{(j+1)} - \mathbf{A}\mathbf{x}^{(j+1)})$
7. $\boldsymbol{\eta}_{\mathbf{v}}^{(j+1)} = \boldsymbol{\eta}_{\mathbf{v}}^{(j)} - (\mathbf{v}^{(j+1)} - \mathbf{R}\mathbf{x}^{(j+1)})$
8. Set $j = j + 1$
- Until stop criterion is met**

Fig. 1. ADMM for Regularized X-ray CT Reconstruction

that constitutes a Lagrange multiplier term (with multipliers $\boldsymbol{\gamma} \triangleq [\boldsymbol{\gamma}_{\mathbf{u}}^T \ \boldsymbol{\gamma}_{\mathbf{v}}^T]^T \in \mathbb{R}^{N_1}$) and an augmented quadratic penalty term with the penalty parameter $\mu > 0$ and a symmetric weighting matrix $\boldsymbol{\Lambda} \succ \mathbf{0}$. The AL scheme for solving **P1** (and thus **P0**) consists of iterating the following steps [17]:

$$(\mathbf{x}^{(j+1)}, \mathbf{z}^{(j+1)}) = \arg \min_{\mathbf{x}, \mathbf{z}} \mathcal{L}(\mathbf{x}, \mathbf{z}, \boldsymbol{\gamma}^{(j)}, \mu), \quad (5)$$

$$\boldsymbol{\gamma}^{(j+1)} = \boldsymbol{\gamma}^{(j)} + \mu \boldsymbol{\Lambda}(\mathbf{z}^{(j+1)} - \mathbf{C}\mathbf{x}^{(j+1)}). \quad (6)$$

An advantage of the AL formalism is that (5)-(6) may converge to a minimizer without having to increase $\mu \rightarrow \infty$ [17]. Absorbing the multiplier term inside the quadratic penalty and using $\boldsymbol{\eta} \triangleq [\boldsymbol{\eta}_{\mathbf{u}}^T \ \boldsymbol{\eta}_{\mathbf{v}}^T]^T = -\frac{1}{\mu} \boldsymbol{\Lambda}^{-1} \boldsymbol{\gamma}$, we write \mathcal{L} (ignoring irrelevant constants) as

$$\mathcal{L}(\mathbf{x}, \mathbf{z}, \boldsymbol{\eta}, \mu) = f(\mathbf{z}) + \frac{\mu}{2} \|\mathbf{z} - \mathbf{C}\mathbf{x} - \boldsymbol{\eta}\|_{\boldsymbol{\Lambda}^2}^2, \quad (7)$$

so (6) becomes $\boldsymbol{\eta}^{(j+1)} = \boldsymbol{\eta}^{(j)} - (\mathbf{z}^{(j+1)} - \mathbf{C}\mathbf{x}^{(j+1)})$. In standard AL formulations (e.g., [2]), $\boldsymbol{\Lambda} = \mathbf{I}_{N_1}$. But in transmission tomography, the elements of \mathbf{A} and \mathbf{R} can differ by several orders of magnitude, and it becomes crucial to balance the sub-matrices in \mathbf{C} . So we propose to use

$$\boldsymbol{\Lambda} = \begin{bmatrix} \mathbf{I}_M & \mathbf{0} \\ \mathbf{0} & \sqrt{\nu} \mathbf{I}_{NL} \end{bmatrix},$$

which results in $\mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{v}, \boldsymbol{\eta}, \mu) = f(\mathbf{z}) + \frac{\mu}{2} \|\mathbf{u} - \mathbf{A}\mathbf{x} - \boldsymbol{\eta}_{\mathbf{u}}\|^2 + \frac{\mu\nu}{2} \|\mathbf{v} - \mathbf{R}\mathbf{x} - \boldsymbol{\eta}_{\mathbf{v}}\|^2$. Including $\nu > 0$ does not affect the AL formalism.

B. Alternating Direction Minimization

The potential advantage of the splitting (3) and the AL formalism (4)-(6) is that \mathcal{L} is amenable to alternating minimization (where \mathcal{L} is minimized with respect to one variable at a time while holding the others at their most recent updates): This a numerically attractive alternative to the joint-minimization in (6) as it decouples the minimization process. So at the j th iteration, instead of (5)-(6), we perform:

$$\mathbf{u}^{(j+1)} = \arg \min_{\mathbf{u}} \mathcal{L}(\mathbf{x}^{(j)}, \mathbf{u}, \mathbf{v}^{(j)}, \boldsymbol{\eta}^{(j)}, \mu) \quad (8)$$

$$\mathbf{v}^{(j+1)} = \arg \min_{\mathbf{v}} \mathcal{L}(\mathbf{x}^{(j)}, \mathbf{u}^{(j+1)}, \mathbf{v}, \boldsymbol{\eta}^{(j)}, \mu) \quad (9)$$

$$\mathbf{x}^{(j+1)} = \arg \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \mathbf{u}^{(j+1)}, \mathbf{v}^{(j+1)}, \boldsymbol{\eta}^{(j)}, \mu) \quad (10)$$

$$\boldsymbol{\eta}^{(j+1)} = \boldsymbol{\eta}^{(j)} - (\mathbf{z}^{(j+1)} - \mathbf{C}\mathbf{x}^{(j+1)}). \quad (11)$$

Convergence of (8)-(11) to a minimizer is ensured [18, Theorem 8] provided that \mathbf{C} has full column-rank. This is readily ensured for most regularization operators \mathbf{R} for CT.

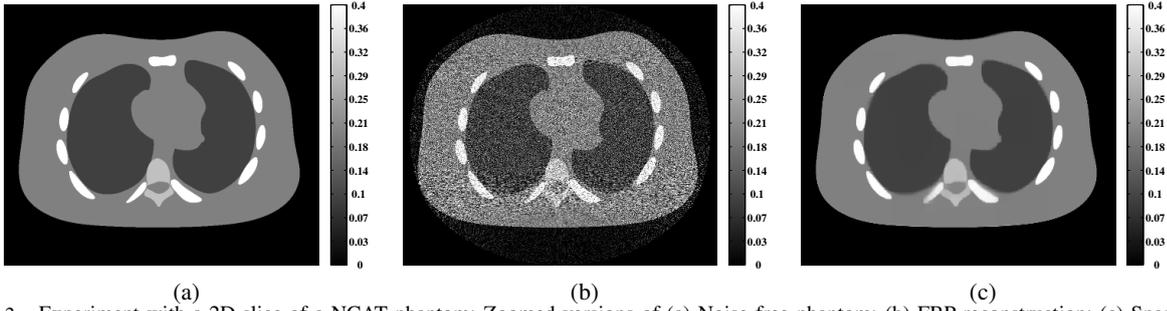


Fig. 2. Experiment with a 2D slice of a NCAT phantom: Zoomed versions of (a) Noise free phantom; (b) FBP reconstruction; (c) Sparsity-regularized Reconstruction (also the solution \mathbf{x}^* of $\mathbf{P0}$).

Sub-problems (8) and (10) involve quadratic criteria (ignoring irrelevant constants) and result in the following updates: $\mathbf{u}^{(j+1)} = \mathbf{H}_\mu^{-1}(\mathbf{W}\mathbf{y} + \mu(\mathbf{A}\mathbf{x}^{(j)} + \boldsymbol{\eta}_u^{(j)}))$ and

$$\mathbf{x}^{(j+1)} = \mathbf{H}_\nu^{-1} \begin{pmatrix} \mathbf{A}^T(\mathbf{u}^{(j+1)} - \boldsymbol{\eta}_u^{(j)}) \\ +\nu\mathbf{R}^T(\mathbf{v}^{(j+1)} - \boldsymbol{\eta}_v^{(j)}) \end{pmatrix}, \quad (12)$$

where $\mathbf{H}_\mu = (\mathbf{W} + \mu\mathbf{I})$ is a diagonal matrix that is easily inverted. Using the constraint variable \mathbf{u} has resulted in the term $\mathbf{H}_\nu = (\mathbf{A}^T\mathbf{A} + \nu\mathbf{R}^T\mathbf{R})$ that is independent of the data (i.e., \mathbf{W}) and can be “inverted” efficiently using PCG with a circulant preconditioner that is a kind of regularized cone filter. In our implementation, we applied 2 PCG iterations with warm starting (i.e., to obtain $\mathbf{x}^{(j+1)}$, PCG is initialized with $\mathbf{x}^{(j)}$). Without \mathbf{u} one would have ended up with a shift variant matrix $(\mathbf{A}^T\mathbf{W}\mathbf{A} + \nu\mathbf{R}^T\mathbf{R})$ that is difficult to precondition [8].

Using $\boldsymbol{\theta}_n = \{v_{nl}\}_{l=1}^L$, $\boldsymbol{\zeta}_n = \{\rho_{nl}^{(j)}\}_{l=1}^L$ where $\rho^{(j)} = \mathbf{R}\mathbf{x}^{(j)} + \boldsymbol{\eta}_v^{(j)}$, (9) decouples in terms of $\{\boldsymbol{\theta}_n\}_{n=1}^N$ as

$$\boldsymbol{\theta}_n^{(j+1)} = \arg \min_{\boldsymbol{\theta}_n} \left\{ \frac{\lambda w_n}{\mu\nu} \Phi_n(\|\boldsymbol{\theta}_n\|_m) + \frac{1}{2} \|\boldsymbol{\theta}_n - \boldsymbol{\zeta}_n^{(j)}\|_2^2 \right\}. \quad (13)$$

This is a L -dimensional denoising problem that can be solved either iteratively (using a general purpose gradient-descent method) for a general Φ_n or exactly for some specific instances of Φ_n and m [19]. For the special case of ℓ_1 -regularization ($\Phi_n(x) = x$, $m = 1$), (13) further decouples in to L 1D problems whose solutions are given by soft-thresholding:

$$v_{nl}^{(j+1)} = \text{soft}\{\rho_{nl}^{(j)}, \lambda w_n / \mu\nu\}, \quad l = 1, \dots, L, \quad (14)$$

where $\text{soft}(d, \lambda) = \text{sign}(d) \max(|d| - \lambda, 0)$.

Based on (3)-(14), Fig. 1 presents our ADMM algorithm for solving $\mathbf{P0}$. With the exception of Step 5, all the steps can be implemented exactly. Steps 3 and 4 are independent and may therefore be executed in parallel. The parameters μ and ν govern the convergence speed of ADMM and do not influence the solution of $\mathbf{P0}$. We used $\nu = \frac{1}{100} \frac{\sigma_{\max}\{\mathbf{A}^T\mathbf{A}\}}{\sigma_{\max}\{\mathbf{R}^T\mathbf{R}\}}$, where $\sigma_{\max}\{\mathbf{A}\}$ is the maximum eigenvalue of \mathbf{A} and $\mu = \text{median}\{w_i\}$ as they provided good convergence speeds for ADMM in all our experiments.

III. RESULTS

We performed preliminary evaluations using 2D CT simulations and a 2D CT phantom scan. The proposed method is readily applicable to 3D problems but our initial implementation is in Matlab so we focused on small 2D cases. For Ψ , we used a ℓ_1 -regularization with finite differences

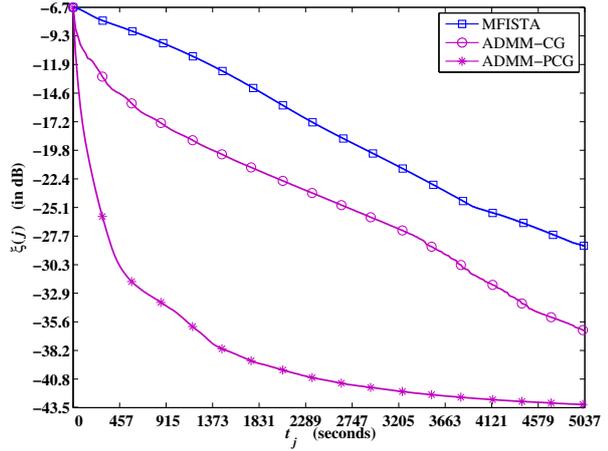


Fig. 3. Experiment with NCAT phantom: Plot of $\xi(j)$ as a function time t_j for MFISTA, and ADMM with unpreconditioned (ADMM-CG) and preconditioned CG inner-iterations (ADMM-PCG).

and chose $w_n = [\mathbf{A}^T\mathbf{W}\mathbf{1}]_n / [\mathbf{A}^T\mathbf{1}]_n$ [16]. To quantify the convergence speed, we computed the normalized ℓ_2 -distance $\xi(j) = 20 \log_{10}(\|\mathbf{x}^{(j)} - \mathbf{x}^*\|_2 / \|\mathbf{x}^*\|_2)$ between $\mathbf{x}^{(j)}$ and \mathbf{x}^* (a solution of $\mathbf{P0}$) as a function of algorithm run-time t_j . We obtained \mathbf{x}^* using the Monotone Fast Iterative Shrinkage Thresholding Algorithm (MFISTA) [20] (with Chambolle-type iterations [15] for the inner-step [20, Equation 5.3]) which is a state-of-the-art method that does not require any “corner rounding” to handle (2) and is directly applicable to $\mathbf{P0}$. We used a 12-core PC with 2.67 GHz Intel Xeon processors. We used the FBP reconstruction as our initial guess $\mathbf{x}^{(0)}$.

In the simulation, we used a 1024×1024 2D slice of the NCAT phantom [21] and numerically generated a 888×984 noisy sinogram (with GE LightSpeed fan-beam geometry [22]) corresponding to a mono-energetic source with 2.5×10^4 incident photons per ray. We reconstructed 512×512 images over a 65cm FOV. Fig. 2 compares standard FBP and sparsity-regularized (\mathbf{x}^*) reconstructions. The regularized output has less noise than the FBP output. Fig. 3 plots ξ for the two versions ADMM-CG and ADMM-PCG corresponding to unpreconditioned and preconditioned CG, respectively, for Step 5. We also included MFISTA for completeness. In this experiment, both versions of ADMM are faster than MFISTA and ADMM-PCG converges to \mathbf{x}^* faster than ADMM-CG.

We scanned a large CIRS phantom on a GE HD scanner using a 80kVp source potential and a 150 mA tube current

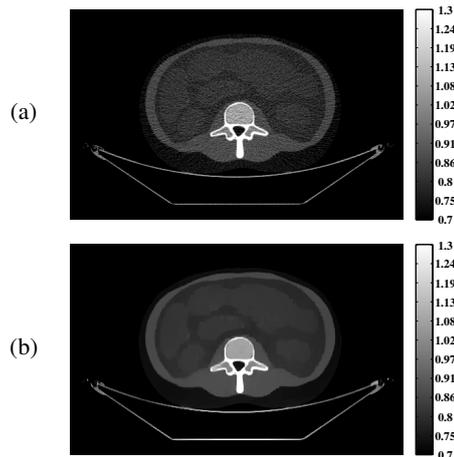


Fig. 4. Experiment with a 2D slice of a real phantom data: Zoomed versions of (a) FBP reconstruction and (b) Sparsity-regularized Reconstruction (also the solution \mathbf{x}^* of P0).

with a 1-second 360 degree rotation. We reconstructed a single 1.25mm thick slice from the 888 by 984 view sinogram. The reconstructed images are 512 x 512 over a 50cm FOV and are displayed in Fig. 4. The sparsity-regularized output has less noise but exhibits blocky image-patches that are typical for ℓ_1 regularization. It may be desirable to use a smoothed edge-preserving regularization (e.g., Huber) to reduce such block-artifacts. Fig. 5 shows ξ for this experiment. ADMM-CG is slower than MFISTA, but ADMM-PCG is fastest among all algorithms.

IV. SUMMARY & CONCLUSIONS

We have described a new iterative reconstruction algorithm, Alternating Direction Method of Multipliers (ADMM), for X-ray CT. ADMM can accommodate non-smooth regularizers (like TV and sparsity encouraging approaches) as well as conventional edge-preserving regularization. The method has an inner step that involves solving a system of equations based on $\mathbf{A}^T \mathbf{A}$, and this step is amenable to preconditioning using FFTs and a type of cone filter. Preliminary 2D CT results show that the proposed algorithm (ADMM-PCG) converges fairly rapidly and that the cone filter greatly accelerates the convergence rate as predicted. The next step is to evaluate the method with 3D helical and axial CT scans of patients and compare it with other algorithms in the literature.

REFERENCES

- [1] M. A. T. Figueiredo and José M Bioucas-Dias. Deconvolution of Poissonian images using variable splitting and augmented Lagrangian optimization, 2009.
- [2] M. V. Afonso, José M Bioucas-Dias, and Mário A T Figueiredo. An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems, 2009. Submitted to IEEE Trans. Image Processing.
- [3] 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.
- [4] E. Y. Sidky and X. Pan. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Phys. Med. Biol.*, 53(17):4777–808, September 2008.
- [5] X. Jia, B. Dong, Y. Lou, and S. B. Jiang. GPU-based iterative cone beam CT reconstruction using tight frame regularization, 2010. Preprint. arxiv:1008.2042v1.

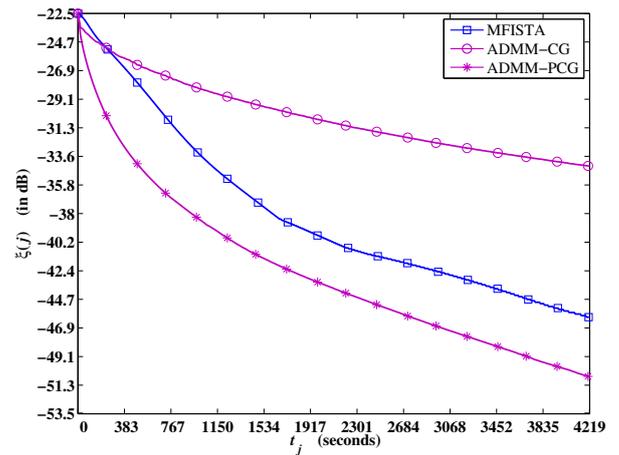


Fig. 5. Experiment with real phantom data: Plot of $\xi(j)$ as a function time t_j for MFISTA, and ADMM with unpreconditioned (ADMM-CG) and preconditioned CG inner-iterations (ADMM-PCG).

- [6] T. M. Benson, B. K. B. D. Man, L. Fu, and J-B. Thibault. Block-based iterative coordinate descent. In *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, 2010.
- [7] C. Kamphuis and F. J. Beekman. Accelerated iterative transmission CT reconstruction using an ordered subsets convex algorithm. *IEEE Trans. Med. Imag.*, 17(6):1001–5, December 1998.
- [8] 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.
- [9] N. H. Clinthorne, T. S. Pan, P. C. Chiao, W. L. Rogers, and J. A. Stamos. Preconditioning methods for improved convergence rates in iterative reconstructions. *IEEE Trans. Med. Imag.*, 12(1):78–83, March 1993.
- [10] A. H. Delaney and Y. Bresler. A fast and accurate Fourier algorithm for iterative parallel-beam tomography. *IEEE Trans. Im. Proc.*, 5(5):740–53, May 1996.
- [11] J. W. Wallis and T. R. Miller. Rapidly converging iterative reconstruction algorithms in single-photon emission computed tomography. *J. Nuc. Med.*, 34(10):1793–800, October 1993.
- [12] J. Nuyts, B. De Man, P. Dupont, M. Defrise, P. Suetens, and L. Mortelmans. Iterative reconstruction for helical CT: A simulation study. *Phys. Med. Biol.*, 43(4):729–37, April 1998.
- [13] T. Goldstein and S. Osher. The split Bregman method for L1-regularized problems. *SIAM J. Imaging Sci.*, 2(2):323–43, 2009.
- [14] S. Ramani and J. A. Fessler. Parallel MR image reconstruction using augmented Lagrangian methods. *IEEE Trans. Med. Imag.*, 2010. To appear as TMI-2010-0275.
- [15] I. W. Selesnick and Mário A T Figueiredo. Signal restoration with overcomplete wavelet transforms: comparison of analysis and synthesis priors. In *spie-7446*, page 74460D, 2009. Wavelets XIII.
- [16] J. A. Fessler and W. L. Rogers. Spatial resolution properties of penalized-likelihood image reconstruction methods: Space-invariant tomographs. *IEEE Trans. Im. Proc.*, 5(9):1346–58, September 1996.
- [17] D. P. Bertsekas. Multiplier methods: A survey. *Automatica*, 12(2):133–45, March 1976.
- [18] J. Eckstein and D. P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *mathprogr*, 55(1-3):293–318, April 1992.
- [19] C. Chau, P. L. Combettes, J-C. Pesquet, and Valérie R Wajs. A variational formulation for frame-based inverse problems. *Inverse Prob.*, 23(4):1495–518, August 2007.
- [20] A. Beck and M. Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Trans. Im. Proc.*, 18(11):2419–34, November 2009.
- [21] W. P. Segars and B. M. W. Tsui. Study of the efficacy of respiratory gating in myocardial SPECT using the new 4-D NCAT phantom. *IEEE Trans. Nuc. Sci.*, 49(3):675–9, June 2002.
- [22] 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.