Quadratic Regularization Design for 2-D CT

Hugo R. Shi*, *Student Member, IEEE*, and Jeffrey A. Fessler, *Fellow, IEEE*

***Abstract*—Statistical methods for tomographic image reconstruction have improved noise and spatial resolution properties that may improve image quality in X-ray computed tomography (CT). Penalized weighted least squares (PWLS) methods using conventional quadratic regularization lead to nonuniform and anisotropic spatial resolution due to interactions between the *weighting*, which is necessary for good noise properties, and the regularizer. Previously, we addressed this problem for parallel-beam emission tomography using matrix algebra methods to design data-dependent, shift-variant regularizers that improve resolution uniformity. This paper develops a fast angular integral mostly analytical (*AIMA*) regularization design method for 2-D fan-beam X-ray CT imaging, for which parallel-beam tomography is a special case. Simulation results demonstrate that the new method for regularization design requires very modest computation and leads to nearly uniform and isotropic spatial resolution in transmission tomography when using quadratic regularization.**

***Index Terms*—Fan-beam computed tomography (CT), iterative reconstruction, local impulse response, regularization, spatial resolution.**

I. Introduction

STATISTICAL image reconstruction methods for X-ray computed tomography (CT) imaging have the potential to reduce patient dose, reduce artifacts from beam hardening and metal objects, and accommodate scanning geometries that are poorly suited for conventional FBP reconstruction. Unregularized image reconstruction leads to excessively noisy images, which necessitates noise control such as a penalized weighted least squares[1] (PWLS) method, or a penalized-likelihood (PL) method [1].

Although PL and PWLS methods can control noise, interactions between a conventional quadratic roughness penalty and the *weighting* that is explicit in PWLS methods [2], and is implicit in PL methods, results in images with nonuniform and anisotropic spatial resolution, even for idealized shift-invariant imaging systems [3]. Fan-beam geometries used in X-ray CT are shift-variant and contain additional variations in spatial resolution over the field-of-view (FOV) [4]. The resulting nonuniformities and anisotropy could be avoided in part by using a conventional penalized *unweighted* least-squares (PULS) estimation method. However, without the *weighting*, PULS yields poor noise properties (akin to FBP).

Much previous work on regularization design focuses on matrix-based approaches to fit the local impulse response of the estimator to a target impulse response. A shift-variant regularizer based on the aggregate certainty of measurement rays intersecting each pixel was developed that yielded uniform but anisotropic spatial resolution [3]. Stayman parameterized the quadratic regularizer to produce uniform and isotropic spatial resolution [5] and generalized regularization design to other non-Poisson noise models [6]. [7] presents a regularization design method for uniform and isotropic spatial resolution that is not based on an explicit target point spread function (PSF), but rather focuses on circular symmetry and uniformity. Extensions to 3-D PET have also been proposed [8], [9].

We previously proposed an analytical approach to regularization design for 2-D parallel-beam emission reconstruction that uses continuous space analogs to simplify the regularization design problem [10]. This paper extends [10] by developing a mostly analytical approach to regularization design for fan-beam geometries [11]. Section II reviews the concept of a local impulse response (LIR) and discusses the design of regularizers that yield approximately uniform and isotropic spatial resolution. Section III uses the frequency domain to gain some insight into the problem and derives a minimization problem to solve for the appropriate regularization coefficients. Section IV describes efficient techniques for computing the regularization coefficients. Section V presents the results of our simulations using simulated and real X-ray CT data, and Section VI analyzes them.

II. Local Impulse Reponse

Let $\boldsymbol{y} = (y_1, \ldots, y_M)$ denote the vector of noisy transmission sinogram measurements recorded in a CT system. For simplicity, we consider the following mono-energetic formulation for the ensemble mean of the measured data:

$$\bar{y}_i(\boldsymbol{x}) = E[y_i] = b_i e^{-[\boldsymbol{A}\boldsymbol{x}]_i} + r_i \tag{1}$$

where $\boldsymbol{A}$ is the system matrix, $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$ is a discretized version of the object being imaged, b_i denotes the blank scan, r_i denotes the additive contribution of scatter, and $[\boldsymbol{A}\boldsymbol{x}]_i = \sum_{j=1}^{N} a_{ij} x_j$. The spatial resolution and therefore regularization design techniques are applicable to PL methods when appropriate weightings are chosen [3]. We estimate $\boldsymbol{x}$ by minimizing the following cost function:

$$\Psi(\boldsymbol{y}, \boldsymbol{x}) \triangleq \|\ell(\boldsymbol{y}) - \boldsymbol{A}\boldsymbol{x}\|_{\boldsymbol{W}}^2 + \zeta R(\boldsymbol{x}) \tag{2}$$

$$\hat{\boldsymbol{x}}(\boldsymbol{y}) = \arg\min_{\boldsymbol{x}} \Phi(\boldsymbol{y}, \boldsymbol{x}) \tag{3}$$

Manuscript received August 03, 2008; revised September 08, 2008. First published October 31, 2008; current version published April 29, 2009. This work was supported in part by the National Institutes of Health/National Cancer Institute (NIH/NCI) under Grant P01 CA87634 and in part by GE Medical Systems. *Asterisk indicates corresponding author.*

*H. R. Shi was with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA. He is now with the Chicago Trading Company, Chicago, IL 60605 USA (e-mail: hugoshi@umich.edu).

J. A. Fessler is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: fessler@eecs.umich.edu).

Digital Object Identifier 10.1109/TMI.2008.2007366

[1]In this paper, we write PWLS and penalized unweighted least squares (PULS) because the techniques used in this paper can be generalized for nonquadratic regularization. However, the scope of this paper is quadratic regularization and all regularizers used here are quadratic.

0278-0062/$25.00 © 2009 IEEE

where $\ell(\boldsymbol{y}) \triangleq -\ln((\boldsymbol{y}-\boldsymbol{r})/\boldsymbol{b})$ using element-wise division, $R(\boldsymbol{x})$ is a regularizer that controls noise, ζ is a scalar that determines the resolution-noise tradeoff, and $\boldsymbol{W} = \mathsf{diag}\{w_i\}$ is a statistical weighting matrix. For transmission tomography with the model in (1), usually we have the plug-in weighting $w_i = y_i$ corresponding to the PL estimator for (1) [12]. For emission tomography, the plug-in weighting is $w_i = 1/y_i$; for low count levels alternative weightings are preferable [13]. The regularization design methods in this paper apply to other statistical models than (1) by simply changing the weighting matrix $\boldsymbol{W}$.

The goal of this work is to design $R(\boldsymbol{x})$ so that the reconstructed image has approximately uniform and isotropic spatial resolution. The tools developed for this purpose could be modified to achieve other resolution design goals. Analyzing the local impulse response (LIR) is an essential part of assessing the resolution. Here we use the following definition of local impulse response for the jth pixel:

$$\begin{aligned}\boldsymbol{l}^j &\triangleq \lim_{\epsilon\to 0} \frac{\hat{\boldsymbol{x}}(\bar{\boldsymbol{y}}(\boldsymbol{x}_{\mathrm{true}} + \epsilon\boldsymbol{\delta}^j)) - \hat{\boldsymbol{x}}(\bar{\boldsymbol{y}}(\boldsymbol{x}_{\mathrm{true}}))}{\epsilon} \\ &= \nabla_{\boldsymbol{y}}\hat{\boldsymbol{x}}(\boldsymbol{y})\big|_{\boldsymbol{y}=\bar{\boldsymbol{y}}(\boldsymbol{x}_{\mathrm{true}})} \nabla_{\boldsymbol{x}}\bar{\boldsymbol{y}}(\boldsymbol{x})\big|_{\boldsymbol{x}=\boldsymbol{x}_{\mathrm{true}}} \boldsymbol{\delta}^j \end{aligned} \tag{4}$$

where $\boldsymbol{\delta}^j$ is a Kronecker impulse at the jth pixel and the gradient operations are matrices such that $[\nabla_{\boldsymbol{y}}\hat{\boldsymbol{x}}(\boldsymbol{y})]_{ji} = (\partial/\partial y_i)\hat{x}_j(\boldsymbol{y})$ and $[\nabla_{\boldsymbol{x}}\bar{\boldsymbol{y}}(\boldsymbol{x})]_{ij} = (\partial/\partial x_j)\bar{y}_i(\boldsymbol{x})$ (this definition depends on the mean of the data, rather than the mean of $\hat{\boldsymbol{x}}$ and differs slightly from that used in [3], but yields the same final expression without requiring further approximations). Evaluating (4)[3], the LIR for the PWLS estimator (3) is

$$\boldsymbol{l}^j = \boldsymbol{l}^j(\boldsymbol{x}_{\mathrm{true}}, \boldsymbol{R}) = [\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A} + \zeta\boldsymbol{R}]^{-1}\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}\boldsymbol{\delta}^j \tag{5}$$

where $\boldsymbol{R}$ is the Hessian of the regularizer $R(\boldsymbol{x})$.

As is evident from (5), the local impulse response depends on the regularizer through its Hessian, $\boldsymbol{R}$. We would like to design $\boldsymbol{R}$ such that the local impulse response $\boldsymbol{l}^j$ best matches a target response $\boldsymbol{l}^0$ at every pixel j. We could phrase this matrix optimization problem as

$$\arg\min_{\boldsymbol{R}} \sum_j \left|\left|\boldsymbol{S}^j\boldsymbol{l}^j(\boldsymbol{x}_{\mathrm{true}}, \boldsymbol{R}) - \boldsymbol{l}^0\right|\right|^2 \tag{6}$$

where $\boldsymbol{S}^j$ is a shift operator that recenters $\boldsymbol{l}^j$ to the center pixel. This matrix formulation of the design problem seems intractable, so we turn to the frequency domain to simplify the problem.

III. Frequency Domain Analysis

This section first reviews the use of discrete Fourier transforms for resolution analysis, leading to a computationally intensive approach to regularization design. We then consider continuous-space analogs that lead to simplified designs.

A. Discrete Fourier Analysis

An analytical expression for $\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}\boldsymbol{\delta}^j$ is derived in Appendix B, showing that $\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}\boldsymbol{\delta}^j$ changes slowly with j. Thus $\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}$ is approximately locally circulant. The $\boldsymbol{R}$ matrix in (5) can be designed to be locally circulant as well. This allows us to factor out a Fourier transform matrix in (5).

Let $\boldsymbol{Q}$ denote an orthonormal discrete Fourier transform matrix centered at pixel j (we recenter the Fourier matrix to avoid the complex exponentials caused by noncentered impulse responses). Then $\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A} \approx \boldsymbol{Q}'\boldsymbol{\Lambda}^j\boldsymbol{Q}$, where $\boldsymbol{\Lambda}^j = \mathsf{diag}\left\{\lambda_k^j\right\}$ and $\lambda^j \triangleq F(\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}\boldsymbol{\delta}^j)$, and $\boldsymbol{R} \approx \boldsymbol{Q}'\boldsymbol{\Gamma}^j\boldsymbol{Q}$, where $\boldsymbol{\Gamma}^j = \mathsf{diag}\left\{\gamma_k^j\right\}$ and $\gamma^j \triangleq F(\boldsymbol{R}\boldsymbol{\delta}^j)$, where $F(\cdot)$ denotes the 2-D DFT [8], [14]. Then we can approximate the LIR in (5) as

$$\begin{aligned}\boldsymbol{l}^j &\approx [\boldsymbol{Q}'\boldsymbol{\Lambda}^j\boldsymbol{Q} + \zeta\boldsymbol{Q}'\boldsymbol{\Gamma}^j\boldsymbol{Q}]^{-1}\boldsymbol{Q}'\boldsymbol{\Lambda}^j\boldsymbol{Q}\boldsymbol{\delta}^j &(7)\\ &= \boldsymbol{Q}'\left[\frac{\boldsymbol{\Lambda}^j}{\boldsymbol{\Lambda}^j + \zeta\boldsymbol{\Gamma}^j}\right]\boldsymbol{Q}\boldsymbol{\delta}^j &(8)\end{aligned}$$

where the matrices in the bracketed term are diagonal and the division operation is element-wise on the diagonal. These approximations are accurate only for row and column indices that are "sufficiently close" to pixel j. The terms in brackets in (8) correspond to the local frequency response of our estimator. We would like to match these to the frequency response $\boldsymbol{L}^0$ of a target PSF (that will be discussed in Section III-B) as closely as possible, i.e., we want

$$\boldsymbol{L}^j \triangleq \frac{\boldsymbol{\Lambda}^j}{\boldsymbol{\Lambda}^j + \zeta\boldsymbol{\Gamma}^j} \approx \boldsymbol{L}^0. \tag{9}$$

Based on (9), one might consider a DFT formulation of the regularization design using the following minimization approach:

$$\hat{\boldsymbol{\Gamma}}^j = \arg\min_{\boldsymbol{\Gamma}\in\mathcal{T}} \left|\left|\boldsymbol{L}^0 - \frac{\boldsymbol{\Lambda}^j}{\boldsymbol{\Lambda}^j + \zeta\boldsymbol{\Gamma}^j}\right|\right| \tag{10}$$

where $\boldsymbol{L}^0$ is the frequency response of $\boldsymbol{l}^0$ and $\mathcal{T}$ denotes the set of possible frequency responses for the regularizer limited by its structure that will be enumerated in Section III-B. Alternatively, as a preview of the methods in Section IV-A, we can cross multiply the terms in (9) yielding the simpler design criterion

$$\hat{\boldsymbol{\Gamma}}^j = \arg\min_{\boldsymbol{\Gamma}\in\mathcal{T}} ||\boldsymbol{L}^0(\boldsymbol{\Lambda}^j + \zeta\boldsymbol{\Gamma}^j) - \boldsymbol{\Lambda}^j||. \tag{11}$$

This approach is similar to the formulation in [6]. Because $\boldsymbol{\Lambda}^j$ is the DFT of $\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}\boldsymbol{\delta}^j$, calculating $\boldsymbol{\Lambda}^j$ requires one forward projection and backward projection per pixel. Thus, regularization design based on (10) or (11) would be very slow.

Prior to [10], regularization design methods were based on discrete Fourier transforms. As shown in Appendix B, the continuous-space analog of $\boldsymbol{\Lambda}^j$ in polar coordinates (ρ, Φ) is the frequency response $w^j(\Phi)/|\rho|$, where $w^j(\Phi)$ is an expression that incorporates the Jacobian from the change of coordinates from parallel-beam to fan-beam geometry, and weights from $\boldsymbol{W}$ that correspond to rays that intersect pixel j at angle Φ. Substituting this into (9) and using continuous space analogs of $\boldsymbol{\Gamma}^j$

yields the following expression for the continuous space analog of $\boldsymbol{L}^j$:

$$\mathcal{L}^j(\rho,\Phi) \approx \frac{w^j(\Phi)\frac{1}{|\rho|}}{w^j(\Phi)\frac{1}{|\rho|} + \zeta R^j(\rho,\Phi)} = \frac{w^j(\Phi)}{w^j(\Phi) + \zeta|\rho| R^j(\rho,\Phi)} \quad (12)$$

where $R^j(\rho,\Phi)$ is the local frequency response of the regularizer near pixel j. The continuous space analogs of $\boldsymbol{\Lambda}^j$ and $\bar{\boldsymbol{\Gamma}}^j$ simplify (12) to provide a more efficient approach to regularization design than (11), as detailed below in (27).

B. Regularization Structure

Regularizers control roughness by penalizing differences between neighboring pixels. Indexing the image $\boldsymbol{x}$ as a 2-D function $x(n,m)$, we define a first-order differencing function for a regularizer that penalizes the lth direction as

$$c_l(n,m) = \frac{1}{\sqrt{n_l^2+m_l^2}}(\delta(n,m) - \delta(n-n_l, m-m_l)) \quad (13)$$

where typically $(n_l,m_l) \in \{(0,1),(1,0),(1,1),(-1,1)\}$, corresponding to horizontal, vertical, and diagonal differences. A conventional quadratic regularizer can then be expressed as

$$R(\boldsymbol{x}) = \sum_{n,m}\sum_{l=1}^{L}\frac{1}{2}|(c_l **x)(n,m)|^2 \quad (14)$$

where $**$ denotes 2-D convolution. This conventional regularizer assigns the same weight to the differences between each neighbor. In this paper, we make the regularizer spatially adaptive with the addition of weighting coefficients r_l^j as follows:

$$R(\boldsymbol{x}) = \sum_{n,m}\sum_{l=1}^{L} r_l^j \frac{1}{2}|(c_l **x)(n,m)|^2 \quad (15)$$

where j is the columnized pixel index which is a function of (n,m). The objective of this paper is to design coefficients $\{r_l^j\}$. To this end, we must analyze the local frequency response $R^j(\rho,\Phi)$ of the Hessian $\boldsymbol{R}$ of the space-variant regularizer (15).

Taking the Fourier transform of (13) yields

$$|C_l(\omega_1,\omega_2)|^2 = \frac{1}{n_l^2+m_l^2}|1-e^{-i(\omega_1 n_l+\omega_2 m_l)}|^2 = \frac{1}{n_l^2+m_l^2}(2-2\cos(\omega_1 n_l+\omega_2 m_l). \quad (16)$$

Combining (16) and (15) yields an accurate expression for the local frequency response of our regularizer that leads to the slower but more accurate Full Integral Iterative NNLS, *FIIN*, method, described in Appendix A. Next we use an approximation to simplify (16) that will lead to a fast, Angular Integral Mostly Analytical, *AIMA* method.

C. Efficient Approximation for $\boldsymbol{R}$

Using the approximation $2-2\cos(x) \approx x^2$, which we will refer to as the *AIMA* approximation, (16) simplifies to

$$|C_l(\omega_1,\omega_2)|^2 \approx \frac{1}{n_l^2+m_l^2}(\omega_1 n_l+\omega_2 m_l)^2. \quad (17)$$

One can think about isotropy intuitively in polar coordinates as eliminating angular dependence. Therefore, we convert (17) to polar frequency coordinates to simplify the analysis. Using frequency and sampling relationships from the DFT, $\omega_1 = 2\pi\Delta_x\rho\cos(\Phi)$ and $\omega_2 = 2\pi\Delta_y\rho\sin(\Phi)$, where Δ_x,Δ_y denotes pixel size and (ρ,Φ) are polar frequency coordinates. For simplicity, we assume $\Delta_x = \Delta_y = 1$ for square pixels. Then

$$\begin{aligned}|C_l(\omega_1,\omega_2)|^2 &\approx \frac{1}{n_l^2+m_l^2}(n_l 2\pi\Delta_x\rho\cos(\Phi) + m_l 2\pi\Delta_y\rho\sin(\Phi))^2 \\ &= \frac{1}{n_l^2+m_l^2}(2\pi\rho)^2(n_l\cos(\Phi)+m_l\sin(\Phi))^2 \\ &= (2\pi\rho)^2\cos^2(\Phi-\Phi_l) \end{aligned} \quad (18)$$

where $\Phi_l \triangleq \tan^{-1}(m_l/n_l)$. Taking the Fourier transform of (15) and combining with (18), our final expression for the local frequency response of the regularizer near the jth pixel is

$$R^j(\rho,\Phi) = \sum_{l=1}^{L} r_l^j (2\pi\rho)^2\cos^2(\Phi-\Phi_l). \quad (19)$$

For the usual choice of $L=4$ and for (n_l,m_l) described below (13), $\Phi_l \in \{0,\pi/2,\pi/4,3\pi/4\}$.

D. Target Local Frequency Response

Substituting (19) into (12) yields a simple expression for the local frequency response of a PWLS estimator. We want to design each regularization coefficient vector $\boldsymbol{r}^j = (r_1^j,\ldots,r_L^j)$ such that our frequency response matches that of the target as closely as possible. We know that the local frequency response associated with PULS is isotropic at the center of the field of view so we select that to be our target.

Using $(n_l,m_l) \in \{(0,1),(1,0)\}$, the squared magnitude response of a conventional regularizer in (14) using the *AIMA* approximation is

$$R_0(\rho,\Phi) = (2\pi\rho)^2 \quad (20)$$

and without the approximation is

$$R_0(\rho,\Phi) = 4 - 2\cos(2\pi\rho\cos\Phi) - 2\sin(2\pi\rho\sin\Phi). \quad (21)$$

For an unweighted cost function and a parallel beam geometry, the continuous-space frequency response that is analogous to $F(\boldsymbol{A}'\boldsymbol{A}\boldsymbol{\delta}^j)$ is $1/|\rho|$. As shown in (49) in Appendix B, for uniform weights ($w_i = 1$) we have the following local frequency response for fan-beam geometries:

$$H^j(\rho,\Phi) = \frac{2}{J(s^j(\Phi))|\rho|} \quad (22)$$

where $J(s^j(\Phi))$ is the Jacobian for the change of coordinates from parallel-beam to fan-beam geometries as defined in (44), and $s^j(\Phi)$ is an index into the sinogram based on pixel j and angle Φ. Setting $s^j(\Phi) = 0$ to correspond with the center pixel in (49), the target local frequency response is

$$\mathcal{L}^0 \triangleq \frac{\frac{2}{J(0)|\rho|}}{\frac{2}{J(0)|\rho|} + \zeta R_0(\rho, \Phi)} \tag{23}$$

$$= \frac{1}{1 + \zeta|\rho| 0.5 J(0) R_0(\rho, \Phi)} \tag{24}$$

where $\mathcal{L}^0$ is the continuous-space analog of $\boldsymbol{L}^0$ in (9). Next we will use these continuous-space analogs to solve for the regularization coefficients $\boldsymbol{r}^j$.

IV. *AIMA* Approach

This section describes the angular integral mostly analytical (*AIMA*) approach. This approach is appropriately named because the continuous space approximations removes all dependence on ρ leaving a single integral over the angular coordinate Φ. This also results in a very simple problem that can be solved mostly analytically.

A. Solving for Regularization Coefficients

We try to design regularization coefficients $\boldsymbol{r}^j$ to match our designed local frequency response (12) to the target local frequency response (24) as follows:

$$\frac{w^j(\Phi)}{w^j(\Phi) + \zeta|\rho| R^j(\rho, \Phi)} \approx \frac{1}{1 + \zeta|\rho| 0.5 J(0) (2\pi\rho)^2}. \tag{25}$$

Cross multiplying and simplifying yields

$$\begin{aligned} w^j(\Phi)\zeta|\rho| 0.5 J(0)(2\pi\rho)^2 &\approx \zeta|\rho| R^j(\rho, \Phi) \\ w^j(\Phi) 0.5 J(0)(2\pi\rho)^2 &\approx \sum_{l=1}^{L} r_l^j (2\pi\rho)^2 \cos^2(\Phi - \Phi_l) \\ \tilde{w}^j(\Phi) \triangleq w^j(\Phi) 0.5 J(0) &\approx \sum_{l=1}^{L} r_l^j \cos^2(\Phi - \Phi_l). \end{aligned} \tag{26}$$

We design $\boldsymbol{r}^j$ by minimizing the difference between both sides of (26)

$$\boldsymbol{r}^j = \underset{\boldsymbol{r} \geq 0}{\arg\min} \frac{1}{2\pi} \int_0^{2\pi} \left(\tilde{w}^j(\Phi) - \sum_{l=1}^{L} r_l \cos^2(\Phi - \Phi_l) \right)^2 d\Phi. \tag{27}$$

We solve for L coefficients using the above minimization for each pixel j. We constrain the L coefficients to be nonnegative which ensures that the penalty function is convex. This expression also applies to parallel-beam geometries, where $J(0) = 1$.

We can think of the minimization in (27) as a projection of $\tilde{w}^j(\Phi)$ onto the space spanned by $\{\cos^2(\cdot - \Phi_l)\}$, which allows us to greatly simplify the problem and derive a computationally efficient and mostly analytical solution to the regularization design problem. Expanding these cosines in terms of a three term basis that is orthonormal with respect to the inner product $\langle f, g \rangle = (1/2\pi) \int_0^{2\pi} f(\Phi) g(\Phi) d\Phi$ yields

$$\cos^2(\Phi - \Phi_l) = \frac{1}{2} \cdot 1 + \frac{\cos(2\Phi_l)}{2\sqrt{2}} \left[\sqrt{2} \cos(2\Phi) \right] + \frac{\sin(2\Phi_l)}{2\sqrt{2}} \left[\sqrt{2} \sin(2\Phi) \right]. \tag{28}$$

The three orthonormal basis functions are

$$\begin{aligned} p_1(\Phi) &= 1 \\ p_2(\Phi) &= \sqrt{2} \cos(2\Phi) \\ p_3(\Phi) &= \sqrt{2} \sin(2\Phi). \end{aligned}$$

Using (27), we write $\sum_{l=1}^{L} r_l^j \cos^2(\Phi - \Phi_l) = \boldsymbol{P}\boldsymbol{T}\boldsymbol{r}^j$, where $\boldsymbol{P}$ is an operator whose columns are p_1, p_2, and p_3, and $\boldsymbol{T}$ is a $3 \times L$ matrix of linear combination coefficients whose lth column is $[1/2 \quad \cos(2\Phi_l)/2\sqrt{2} \quad \sin(2\Phi_l)/2\sqrt{2}]^T$. $\boldsymbol{T}$ is computed by taking the inner products of $\cos^2(\Phi - \Phi_l)$ and p_1, p_2, and p_3, i.e., $\boldsymbol{T}_{mn} = \int \cos^2(\Phi - \Phi_m) p_n(\Phi) d\Phi$.

Using (28), the minimization problem (27) simplifies to the following expression:

$$\boldsymbol{r}^j = \underset{\boldsymbol{r} \geq 0}{\arg\min} \|\boldsymbol{T}\boldsymbol{r} - \boldsymbol{b}^j\|^2 \tag{29}$$

where $\boldsymbol{b}^j \triangleq \boldsymbol{P}^* \tilde{\boldsymbol{w}}^j(\cdot)$ and $\boldsymbol{P}^*$ denotes the adjoint of $\boldsymbol{P}$, i.e., $b_k^j = \int p_k(\Phi) \tilde{w}^j(\Phi) d\Phi$, $k = 1, 2, 3$.

B. Zeros in the Hessian

If there are too many zeros in $\boldsymbol{r}^j$, there will be zeros in the Hessian, possibly degrading $\hat{\boldsymbol{x}}$. This can cause elongated impulse responses that may contribute to streak artifacts in the reconstructed image. This phenomenon occurred in [10] however we did not notice the artifacts due to the coarser spatial resolution in PET. It is present when using *AIMA* with the ring phantom in Section V. The problem improves when using *FIIN* method of Appendix A, thus we believe this phenomenon is caused by the approximation error in (17) as well as the nonnegativity constraint. For *AIMA*, and to be safe using *FIIN*, we modify (29) to ensure that an adequate number of r_l^j values are greater than some $\epsilon_l^j > 0$. Requiring the penalty coefficients for the vertical and horizontal directions be nonzero is sufficient to eliminate zeros in the Hessian. (A similar constraint could be created using the 2 diagonals instead of the vertical and horizontal neighbors; however the approximation in (17) is worse for diagonal neighbors). We turn to previous work to select ϵ_l^j. In [3], we derived a certainty based weighting using a spatially variant $(\kappa^j)^2 = (\sum_i a_{ij}^2 \tilde{w}_i)/(\sum_i a_{ij}^2)$ that seeks to provide uniform spatial resolution. In terms of the continuous space analogs used in this paper, $(\kappa^j)^2 = (1/\pi) \int_0^{\pi} \tilde{w}^j(\Phi) d\Phi$. This regularization design method can be implemented using the regularization structure presented in this paper by setting $r_l^j = (\kappa^j)^2$ for $(m_l, n_l) = (0, 1), (1, 0)$. This approach provides a convenient nominal value for the regularization strength at each pixel. We define the lower constraint vector $\boldsymbol{\epsilon}^j$ such that $\epsilon_l^j = \alpha(\kappa^j)^2$ for vertical and horizontal neighbors, $(m_l, n_l) = (0, 1), (1, 0)$, and $\epsilon_l^j = 0$ for all other neighbors. Using a nonzero coefficient α improves noise at the expense of isotropy. For the results presented in this paper, we used $\alpha = 0.1$.

Now we formulate our problem so that nonnegative least squares (NNLS) algorithms will accommodate this new constraint. Let $\bar{\boldsymbol{r}}^j \triangleq \boldsymbol{r}^j + \boldsymbol{\epsilon}^j$. Solving with the constraint of $\boldsymbol{r}^j \geq 0$ ensures that $\bar{\boldsymbol{r}}^j \geq \boldsymbol{\epsilon}^j$. Substituting $\bar{\boldsymbol{r}}^j$ into (29) yields $\boldsymbol{T}\bar{\boldsymbol{r}}^j - \boldsymbol{b}^j = \boldsymbol{T}(\boldsymbol{r}^j + \boldsymbol{\epsilon}^j) - \boldsymbol{b}^j = \boldsymbol{T}\boldsymbol{r}^j - (\boldsymbol{b}^j - \boldsymbol{T}\boldsymbol{\epsilon}^j)$. So our final cost function for regularization design is

$$\begin{aligned}\hat{\boldsymbol{r}}^j &= \arg\min_{\boldsymbol{r}\geq 0} \|\boldsymbol{T}\boldsymbol{r} - (\boldsymbol{b}^j - \boldsymbol{T}\boldsymbol{\epsilon}^j)\| \\ &= \arg\min_{\boldsymbol{r}\geq 0} \|\boldsymbol{T}\boldsymbol{r} - \bar{\boldsymbol{b}}^j\| \end{aligned} \tag{30}$$

where $\bar{\boldsymbol{b}}^j \triangleq \boldsymbol{b}^j - \boldsymbol{T}\boldsymbol{\epsilon}^j$. We use coefficients $\bar{\boldsymbol{r}}^j = \hat{\boldsymbol{r}}^j + \boldsymbol{\epsilon}^j$ in the regularizer (15). We next solve (30) analytically. It is this analytical solution that makes the *AIMA* more efficient than the *FIIN* method of Appendix A.

C. Mostly Analytical Solution

Using a second-order neighborhood ($L = 4$) we select (n_l, m_l) to be $(1,0),(0,1),(1,1),(1,-1)$ leading to the following Φ_l: $\Phi_1 = 0$, $\Phi_2 = \pi/2$, $\Phi_3 = \pi/4$, $\Phi_4 = -\pi/4$. So the terms in (30) are

$$\boldsymbol{T} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}$$

$$\bar{\boldsymbol{b}}^j = \begin{bmatrix} d_1^j \\ \sqrt{2}d_2^j \\ \sqrt{2}d_3^j \end{bmatrix}, \; \boldsymbol{d}^j = \begin{bmatrix} \frac{1}{2\pi}(1-\alpha)\int_0^{2\pi} \tilde{w}^j(\Phi)d\Phi \\ \frac{1}{2\pi}\int_0^{2\pi} \tilde{w}^j(\Phi)\cos(2\Phi)d\Phi \\ \frac{1}{2\pi}\int_0^{2\pi} \tilde{w}^j(\Phi)\sin(2\Phi)d\Phi \end{bmatrix}. \tag{31}$$

Observe that $\boldsymbol{T}\boldsymbol{\epsilon}^j = [\alpha(\kappa^j)^2 \; 0 \; 0]^T$, so $\boldsymbol{\epsilon}^j$ affects only d_1. Ignoring α, d_1 is the continuous space analog of the certainty-based regularization weighting in [3], d_2 is related to the horizontal and vertical directions, and d_3 is related to the diagonal directions.

The system (31) is under-determined, which is somewhat intuitive since one can obtain approximately isotropic smoothing using only the horizontal and vertical neighbors, or only the diagonal neighbors. For the purposes of regularization design, an under-determined situation is desirable since it allows us to use the "extra" degrees-of-freedom to ensure nonnegativity even when anisotropic regularization is needed.

We could solve the minimization (30) using an iterative *NNLS algorithm* [15, p.158]. However, using the properties of $\boldsymbol{T}$ and $\boldsymbol{d}$, we can avoid iterations almost entirely by solving (30) *analytically* using the KKT conditions. When $\alpha = 0$, $\sqrt{d_2^2 + d_3^2} \leq d_1$. as outlined in Appendix C. This inequality is usually true for the values of α used (typically around 0.1), however for pixels where it is not, the problem would have to be solved using a NNLS algorithm rather than this analytical solution. To use NNLS, one must add Tikhonov regularization to help (30) converge to a minimum norm solution which ensures that $\boldsymbol{r}^j$ is a continuous function of $\boldsymbol{d}^j$ which is a continuous function of $\tilde{w}^j(\Phi)$. For more details on using NNLS to solve the regularization problem, see [9, Sec. IV-A]. For the results presented in this paper, all pixels of both simulations satisfied $\sqrt{d_2^2 + d_3^2} \leq d_1$ for all pixels and NNLS was not used. The structure of $\boldsymbol{T}$ leads to eight-fold symmetry that simplifies analysis. If $d_2 < 0$ we can solve for $\boldsymbol{r}$ using $|d_2|$ and then swap r_1 with r_2. If $d_3 < 0$ we can solve for $\boldsymbol{r}$ using $|d_3|$ and then swap r_3 with r_4. If $d_3 > d_2$ we can solve for $\boldsymbol{r}$ with d_2 and d_3 interchanged, and then swap r_1 with r_3 and r_2 with r_4. Therefore, hereafter we focus on cases where $0 \leq d_3 \leq d_2 \leq d_1$. Fig. 1 shows these first octant cases, numbered according to the number of nonzero elements of $\boldsymbol{r}$.

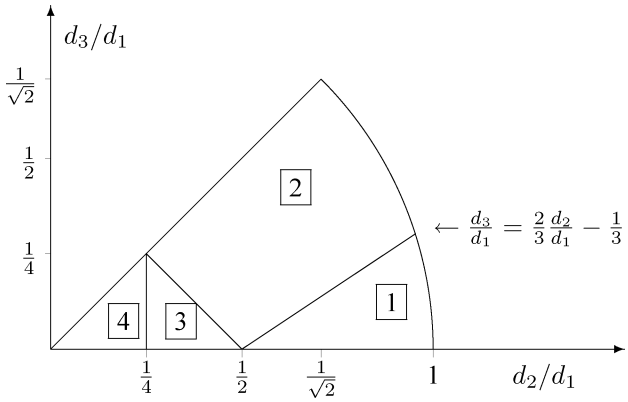

Fig. 1. First octant of quadratic penalty design space showing the four regions where different constraints are active.

1) If $d_2 \geq (1/2)d_1$ and $d_3 \leq (2/3)d_2 - (1/3)d_1$, then $r_1 = (4/3)(d_1 + d_2)$, $r_2 = r_3 = r_4 = 0$.
2) If $d_3 \geq (2/3)d_2 - (1/3)d_1$ and $d_3 + d_2 \geq (1/2)d_1$, then $r_1 = (8/5)[(1/2)d_1 + (3/2)d_2 - d_3]$, $r_2 = r_4 = 0$, $r_3 = (12/5)[d_3 - ((2/3)d_2 - (1/3)d_1)]$.
3) If $d_3 + d_2 \leq (1/2)d_1$ and $d_2 \geq (1/4)d_1$, then there are multiple nonnegative $\boldsymbol{r}$ that exactly solve $\nabla\Psi(\boldsymbol{r}) = \boldsymbol{0}$. The minimum-norm solution is $r_1 = 4d_2$, $r_2 = 0$, $r_3 = d_1 - 2d_2 + 2d_3$, $r_4 = 2[(1/2)d_1 - (d_2 + d_3)]$.
4) If $d_2 \leq (1/4)d_1$, then there are multiple nonnegative $\boldsymbol{r}$ that are exact solutions. The natural choice is the minimum-norm $\boldsymbol{r}$ given by the pseudo-inverse solution $\boldsymbol{r} = \boldsymbol{T}^\dagger\boldsymbol{d}$, where $r_1 = 2((1/4)d_1 + d_2)$, $r_2 = 2((1/4)d_1 - d_2)$, $r_3 = 2((1/4)d_1 + d_3)$, $r_4 = 2((1/4)d_1 - d_3)$.

The analytical solution presented above is for the usual first-order differences (13). For higher-order differences or neighborhoods, it would appear to become increasingly cumbersome to solve (27) analytically, so an iterative NNLS approach may be more appealing. This can still be practical since $\boldsymbol{T}$ is quite small. The analytical solution above is a continuous function of $\boldsymbol{d}$, which in turn is a continuous function of $\tilde{w}^j(\Phi)$. This continuity property would seem to be desirable for avoiding artifacts in the reconstructed images.

For practical implementation, we simply discretize the integrals in (31) [16]. This presents interpolation issues in extracting a discretized version of $\tilde{w}^j(\Phi)$, from $\boldsymbol{W}$. For the fan beam case, we use the analytical formula for $\tilde{w}^j(\Phi)$ presented in Appendix B (48). Those equations are presented with continuous space coordinates s that spans the length of the detector, and β, the angle of rotation of the system. In discrete implementation, we round these off to the nearest neighbor to get indices into $\boldsymbol{W}$. The analytical solution presented in this section is very efficient over all values of j, and the bulk of the computation time is spent computing $\tilde{w}^j(\Phi)$ which has the compute time of approximately one back-projection. For further speed optimizations, down-sampled back-projections could be explored. In practice, $\tilde{w}^j(\Phi)$ is fairly smooth and the basis functions that we use to approximate $\tilde{w}^j(\Phi)$ are even smoother. One can achieve good results by calculating $\tilde{w}^j(\Phi)$ for fewer angles

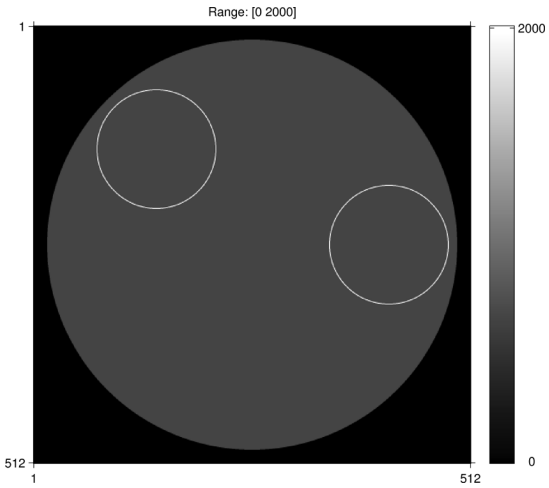

Fig. 2. Ring phantom used for simulation study.

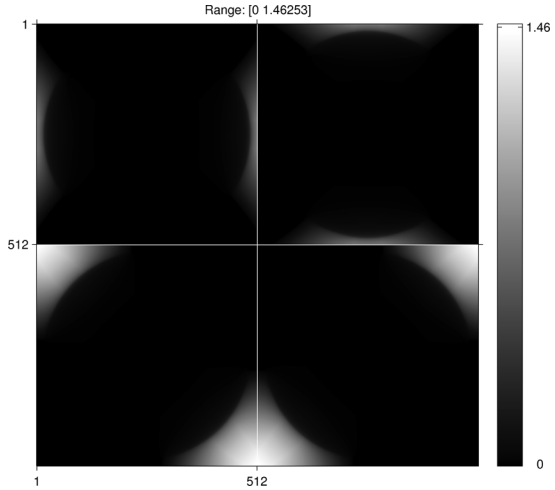

Fig. 3. Regularization penalty coefficients used in reconstruction of ring phantom. The four images are r_l^j for $l = 1, \ldots, 4$.

and interpolating the rest. The performance and efficacy of such a scheme still needs to be explored.

V. Results

We first investigated imaging a phantom consisting of two uniform rings that highlight the effects of nonuniformities and anisotropy [7]. Afterwards, we studied real CT data.

We simulated a 2-D third-generation fan-beam CT system using distance-driven forward and backprojections [17]. The rotation center is 40.8 cm from the detector, and the source is 94.9 cm from the detector. The axis of rotation is at the center of the object. The simulated imaging system has 888 rays per view spaced 1 mm apart, and 984 evenly spaced view angles over a 2π rotation. The reconstructed images consisted of a 512×512 grid of 1 mm pixels. We chose a ζ in (24) such that the target PSF has a full-width at half-maximum (FWHM) of 3.18 mm.

We simulated a noiseless fan-beam sinogram without scatter using a phantom consisting of a background disk and two rings each of thickness 1 mm shown in Fig. 2. We generated the noiseless sinogram by taking line-integrals through the analytical phantom with the same system geometry. We used the plug-in weighting $w_i = y_i$ for this experiment. Fig. 3 shows penalty coefficients $\{\boldsymbol{r}^j\}$ designed using the method of Section IV-C that yielded the following results. The top images show coefficients in the horizontal and vertical neighbors, and the bottom images show coefficients for diagonal neighbors. There are substantial spatial variations in these coefficients.

We reconstructed images $\hat{\boldsymbol{x}}$ using several methods. We first created an image uniformly blurred by the target PSF. We then reconstructed images using the following

1) conventional regularization, $\boldsymbol{r}^j = 1$, $(m_l, n_l) = (0,1), (1,0)$;
2) certainty based regularization [3];
3) *AIMA* method with $\alpha = 0.1$;
4) *AIMA* method with $\alpha = 0$;
5) *FIIN* method with $\alpha = 0$.

Fig. 4 is a closeup of the right-most ring reconstructed with the various methods listed above and Fig. 5 is a closeup of the left-most ring.

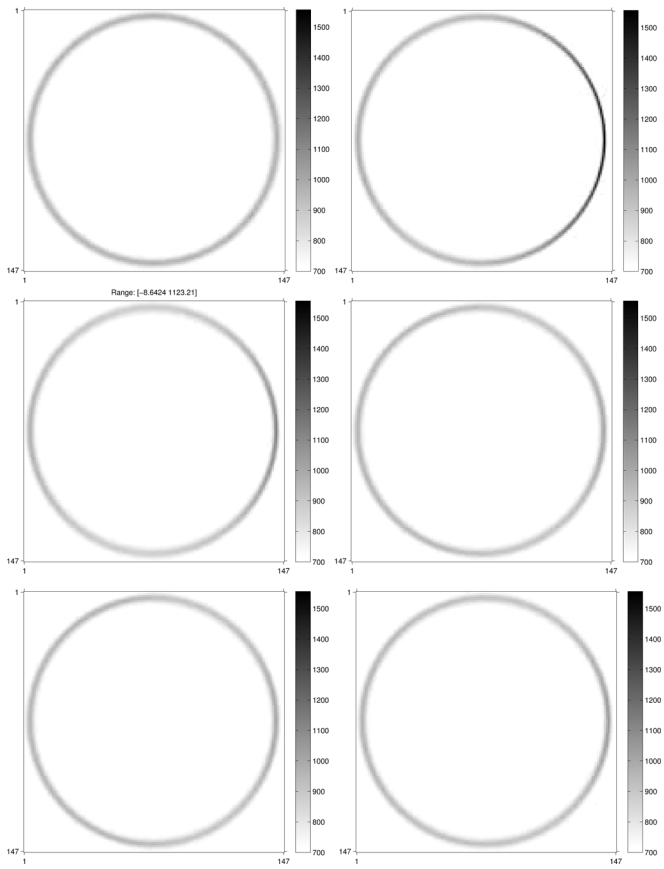

Fig. 4. Images of right-most ring, Upper-Left: uniformly blurred by target PSF. Upper-Right: reconstructed using conventional regularization. Mid-Left: reconstructed using certainty-based regularization. Mid-Right: reconstructed using *AIMA* regularization, with $\alpha = 0.1$. Lower-Left: reconstructed using *AIMA* regularization, with $\alpha = 0$. Lower-Right: reconstructed using *FIIN* regularization, with $\alpha = 0$.

Figs. 6 and 7 show profiles around the two rings of the reconstructed images using the various regularization methods relative to the mean intensity of the rings from our target, PULS reconstruction with conventional regularization. This verifies that *AIMA* and *FIIN* improve resolution uniformity. Here, 0 radians

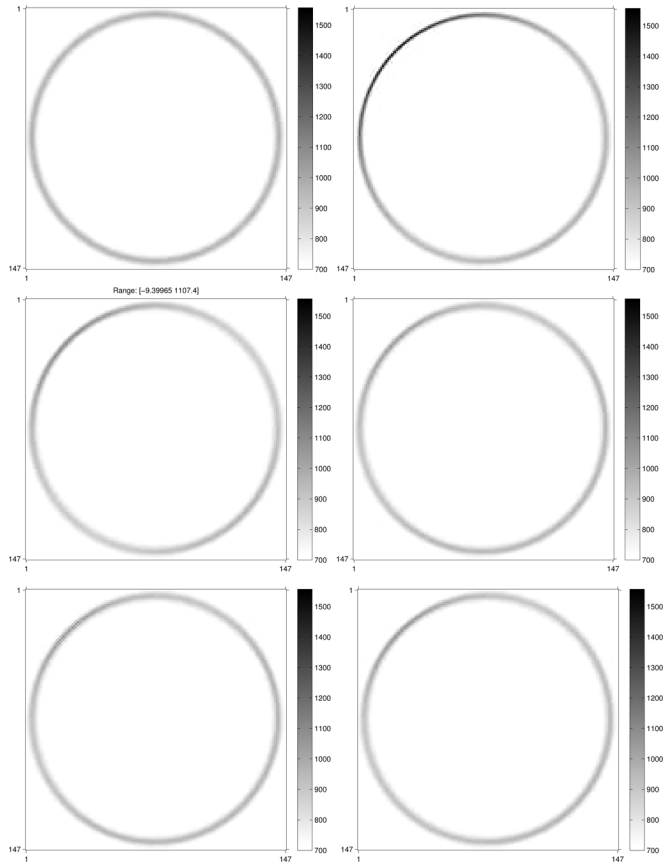

Fig. 5. Images of left-most ring, Upper-Left: uniformly blurred by target PSF. Upper-Right: reconstructed using conventional regularization. Mid-Left: reconstructed using certainty-based regularization. Mid-Right: reconstructed using *AIMA* regularization, with $\alpha = 0.1$. Lower-Left: reconstructed using *AIMA* regularization, with $\alpha = 0$. Lower-Right: reconstructed using *FIIN* regularization, with $\alpha = 0$.

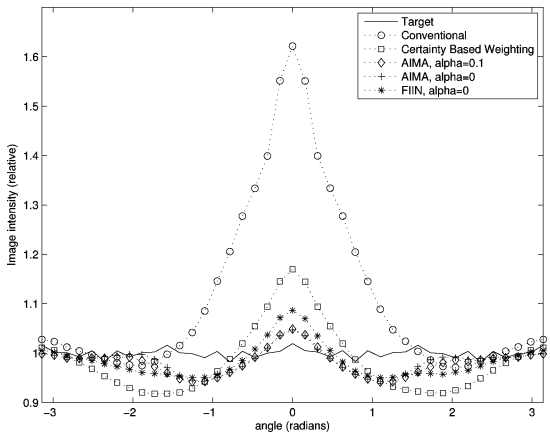

Fig. 6. Profiles through the right-most ring from each reconstructed image.

corresponds to the rightmost point of that ring and the angles are measured clockwise.

The second study used a similar imaging geometry using one slice of real CT data from a GE scanner described in [18] with weightings $\boldsymbol{W}$ computed in the same way as [19]. ζ was selected such that the target PSF had a FWHM of 1.51 mm.

Fig. 8 displays an image reconstructed by PWLS using conventional regularization. The impulse response locations

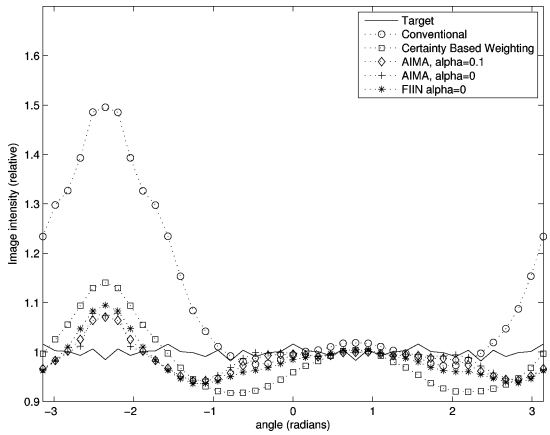

Fig. 7. Profiles through the left-most ring from each reconstructed image.

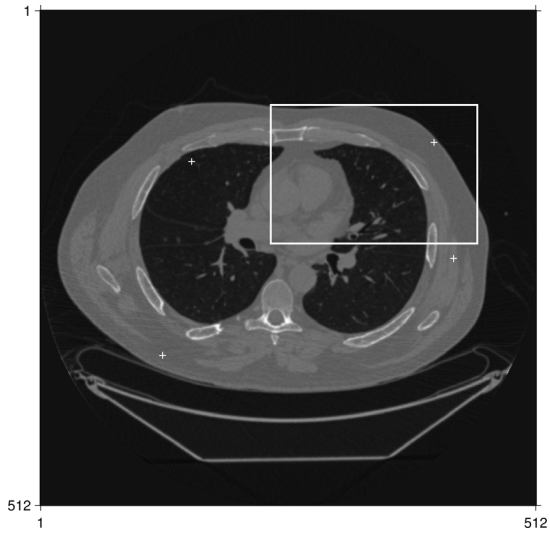

Fig. 8. Reconstruction with conventional regularization without windowing with impulse responses marked.

are denoted with crosses and a region is marked which will be displayed in Fig. 9. Fig. 9 show windowed reconstructions using conventional regularization, certainty based regularization, *AIMA* with $\alpha = 0.1$ and $\alpha = 0$, and *FIIN* with $\alpha = 0$.

Figs. 10–13 show local impulse responses for the five regularization methods at several locations calculated analytically using (5). These figures show from left to right, the target impulse response, and local impulses responses for conventional regularization, certainty based regularization [3], *AIMA* with $\alpha = 0.1$, *AIMA* with $\alpha = 0$, and *FIIN* with $\alpha = 0$. Contour plots of the LIR are displayed below at 0.9, 0.75, 0.5, 0.25, and 0.1 of the maximum value of the target PSF. The LIR becomes more anisotropic near the edge of the FOV. Our Fourier-based regularization scheme compensates for this anisotropy better than the certainty-based approach of [3].

To quantify the performance of these regularizers on real CT data, we computed the PSF of the regularizers at every 10 pixels within the body. Then we calculated the FWHM of the PSFs at 181 evenly spaced angles, and computed the root mean square (rms) error between the actual FWHM and the FWHM of the target. Histograms of these errors are displayed in Fig. 14. The

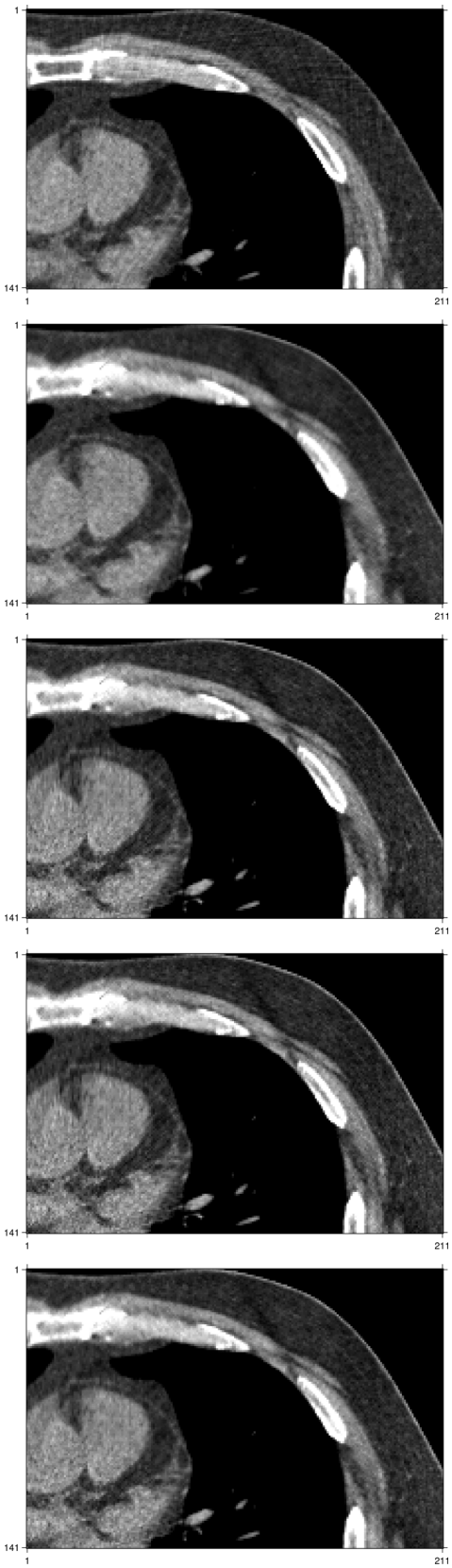

Fig. 9. Reconstructions windowed from 800–1200 HUs using, from top to bottom, conventional regularization, certainty based regularization, *AIMA* with $\alpha = 0.1$, *AIMA* with $\alpha = 0$, *FIIN* with $\alpha = 0$.

mean of the rms errors for conventional regularization, certainty based regularization, the *AIMA* method with $\alpha = 0.1$, the *AIMA* method with $\alpha = 0$, and the *FIIN* method with $\alpha = 0$ are 2.7, 2.7, 2.3, 2.5, and 2.0, respectively.

Fig. 10. Impulse Responses at $(-100, -100)$. From left to right, target, conventional regularization, certainty based regularization, *AIMA* regularization with $\alpha = 0.1$, *AIMA* regularization with $\alpha = 0$, *FIIN* regularization with $\alpha = 0$.

Fig. 11. Impulse Responses at $(150, -120)$. From left to right, target, conventional regularization, certainty based regularization, *AIMA* regularization with $\alpha = 0.1$, *AIMA* regularization with $\alpha = 0$, *FIIN* regularization with $\alpha = 0$.

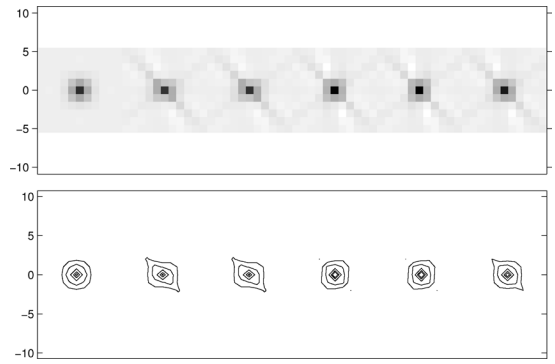

Fig. 12. Impulse Responses at (170,0). From left to right, target, conventional regularization, certainty based regularization, *AIMA* regularization with $\alpha = 0.1$, *AIMA* regularization with $\alpha = 0$, *FIIN* regularization with $\alpha = 0$.

VI. Discussion

A. Spatial Resolution Properties

Figs. 4 and 5 provide a qualitative understanding of the spatial resolution properties of various regularization methods for the ring phantom. The phantom used consists of rings of uniform intensity and uniform width, thus images with uniform spatial

Fig. 13. Impulse Responses at $(-130, 100)$. From left to right, target, conventional regularization, certainty based regularization, *AIMA* regularization with $\alpha = 0.1$, *AIMA* regularization with $\alpha = 0$, *FIIN* regularization with $\alpha = 0$.

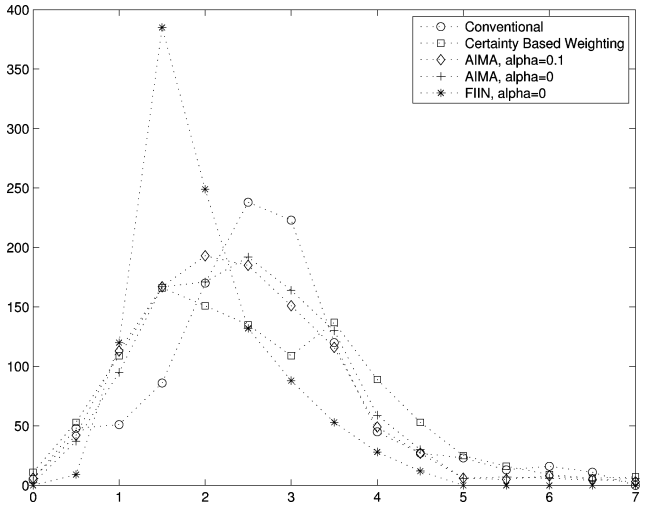

Fig. 14. Plots of the error histogram for different impulse responses.

resolution should have rings with uniform width and uniform intensity. Conventional regularization creates rings with sharper spatial resolution near the edge of the field of view. Certainty based regularization improves ring uniformity but anisotropy remains. The last three reconstructions using *AIMA* with $\alpha = 0.1$, *AIMA* with $\alpha = 0$ and *FIIN* provide rings that look almost identical, and which have a more uniform ring width than conventional, and certainty based regularization.

Figs. 6 and 7 show the amplitude of the rings traced clockwise. This confirms our initial assessment, that certainty based regularization achieves more uniform spatial resolution than conventional regularization, and that the *AIMA* and *FIIN* methods are all very similar and outperform the previous approaches.

Fig. 9 display a quadrant of windowed reconstructions using real CT data to illustrate the images produced using these regularization design methodologies. However, since we do not know the "truth" for this data, these images provide only a qualitative illustration of the effect of regularization design on spatial resolution. The impulse responses in Figs. 10–13 illustrate the effect of regularization design on spatial resolution at various locations. These figures confirm that *AIMA* and *FIIN* methods improved isotropy over conventional and certainty-based regularization. The histogram plot of Fig. 14 and the mean of the PSF errors mentioned previously confirm this. *AIMA* has lower FWHM rms error both certainty based regularization and conventional regularization. *FIIN* has the lowest FWHM rms error, however it is much slower than *AIMA*.

The resulting impulse responses from the *AIMA* and *FIIN* methods are not completely isotropic. This may seem to contradict the dramatic improvement these regularization design methods achieved with the ring phantom. However, recall that we are trying to approximate $\tilde{w}^j(\Phi)$ using three basis functions (see Section IV-A) for the *AIMA*, and four basis functions (see Appendix A) for *FIIN*. With real data, $\tilde{w}^j(\Phi)$ is a complicated function that cannot be parameterized using three or four basis functions. This aspect, along with the nonnegativity constraint limits the performance of any regularization design technique with a finite number of parameters. $\tilde{w}^j(\Phi)$ is much simpler for simple phantoms like the ring phantom, so *AIMA* yields better results there. Extensions of this regularization design to higher order penalties have the potential for more basis functions, and better performance.

The analysis of this paper focuses on the resolution nonuniformities cased by statistical weightings, not the resolution variation due to detector response and magnification. A more general regularization design with similar parameterization is discussed in [6]. Using the techniques in this paper to account for these effects is an open problem.

B. Computation Time

AIMA is quite efficient. Computing certainty based regularization takes the time of about 1 backprojection. In AIMA, we must first compute $\tilde{w}^j(\Phi)$ which takes the time of about one backprojection, and then solve the analytical solution which is very fast. FIIN also requires one backprojection to compute $\tilde{w}^j(\Phi)$, however it then has to run a nonnegative least squares problem for every pixel. Though this is a fairly small NNLS problem, T is 4×4, it adds much compute time since it must be calculated for each pixel. In general, due to the faster compute times, we recommend AIMA with $\alpha = 0.1$. If accuracy is more important than compute time, FIIN can be used instead.

VII. Conclusion and Future Work

The results presented indicate that we have a functional and efficient regularization design structure that yields approximately uniform and isotropic spatial resolution for 2-D fan-beam systems. The methods also apply to the simpler case of 2-D (parallel-beam) PET as described in [10]. In the future, we will extend these methods to 3-D PET [9] and cone-beam CT. We will also study the noise properties. Another open question is how to design nonquadratic regularization [20].

Appendix A
Slower Regularization Design

This appendix describes the slower Double Integral Iterative NNLS, *FIIN*, method that does not use the *AIMA* approximation $2 - 2\cos(x) \approx x^2$ in (17) in the expression of the local frequency response of $\boldsymbol{R}$. This design method is appropriately named because it uses the full integral over all frequency domain polar coordinates instead of just the angular coordinates. The matrices involved in this method are slightly more complex

than those in *AIMA* so an iterative technique is needed. Converting (16) in Section III-C into polar coordinates and combining with (15) yields

$$R^j(\rho,\Phi)=\sum_{l=1}^{L}\frac{r_l^j}{n_l^2+m_l^2}(2-2\cos(u_l(\rho,\Phi)))$$

where $u_l(\rho,\Phi)=n_l 2\pi\Delta_x\rho\cos(\Phi)+m_l 2\pi\Delta_y\rho\sin(\Phi)$. The target we use in regularization is the LIR associated with PULS at the center of the image, expressed in (23).

Following (25), we try to match the LIR at each pixel j to the target

$$\begin{aligned}\frac{w^j(\Phi)}{w^j(\Phi)+\beta|\rho|R^j(\rho,\Phi)}&\approx\frac{1}{1+\beta|\rho|0.5J(0)R_0(\rho,\Phi)}\\ w^j(\Phi)+\beta|\rho|R^j(\rho,\Phi)&\approx w^j(\Phi)(1+\beta|\rho|0.5J(0)R_0(\rho,\Phi))\\ \beta|\rho|R^j(\rho,\Phi)&\approx\beta|\rho|0.5J(0)w^j(\Phi)R_0(\rho,\Phi))\\ R^j(\rho,\Phi)&\approx\tilde{w}^j(\Phi)R_0(\rho,\Phi))\end{aligned}$$

where $\tilde{w}^j(\Phi)=0.5J(0)w^j(\Phi)$. To design $\boldsymbol{R}$, we use a method similar to that of Section IV-A and project $\tilde{w}^j(\Phi)R_0(\rho,\Phi))$ onto the space spanned by $\{1-\cos(u_l(\rho,\Phi))\}$ that can be orthonormalized into L basis functions, $\{p_l(\rho,\Phi)\}$ using Gramm–Schmidt. Then our regularization design problem simplifies to

$$\boldsymbol{r}^j=\arg\min_{\boldsymbol{r}\geq 0}\|\boldsymbol{T}\boldsymbol{r}-\boldsymbol{b}^j\|^2 \tag{32}$$

where $\boldsymbol{b}^j$ is a vector of inner products between $\tilde{w}^j(\Phi)R_0(\rho,\Phi))$ and the L orthonormal basis functions $\{p_l(\rho,\Phi)\}$, $b_k^j=\int\int p_k(\rho,\Phi)\tilde{w}^j(\Phi)R_0(\rho,\Phi)d\rho d\Phi$, and $\boldsymbol{T}$ is a $L\times L$ matrix whose elements are the inner products between $2-2\cos(u_l(\rho,\Phi))$ and $\{p_l(\rho,\Phi)\}$, $\boldsymbol{T}_{mn}=\int\int 2-2cos(u_m(\rho,\Phi))p_n(\rho,\Phi)d\rho d\Phi$. We define $\boldsymbol{P}$ to be an operator whose columns are $\{p_1,p_2,\ldots p_L\}$ so $R(\rho,\Phi)=\boldsymbol{P}\boldsymbol{T}\boldsymbol{r}^j$. This design problem is then solved with an NNLS algorithm. This more accurate version is slower than that presented in Section IV-A because there is no apparent analytical solution similar to the one presented in Section IV-C.

Appendix B
Analysis of Grammian Operator

This appendix considers fan-beam geometries and uses continuous-space analysis to analyze the Fourier transform of the Grammian operator $\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}$ to simplify the regularization design problem in (12). One can use polar coordinates Φ,ρ, and continuous-space analysis to separate the angular and radial components of $\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}$. Using this framework, isotropy can be thought of as eliminating dependence on the angular component Φ.

Fan-Beam Geometry: Fig. 15 illustrates the fan-beam geometry that we consider. Let P be the rotation isocenter. $D_{0\mathrm{d}}$ denotes the distance from the point P to the detector, $D_{\mathrm{s}0}$ denotes the distance from the X-ray source to P, and D_{fs} denotes the distance from the X-ray source to the focal point of the detector arc. Define $D_{\mathrm{sd}}\triangleq D_{0\mathrm{d}}+D_{\mathrm{s}0}$ to be the total distance from the X-ray source to the center of the detector, and $D_{\mathrm{fd}}\triangleq D_{\mathrm{sd}}+D_{\mathrm{fs}}$ to be the total distance from the focal point to the center of the detector. This formulation encompasses a variety of system configurations by allowing the detector focal point to differ from the X-ray source location. For flat detectors, $D_{\mathrm{fs}}=\infty$. For third-generation X-ray CT systems, $D_{\mathrm{fs}}=0$. For fourth generation X-ray CT systems, $D_{\mathrm{fs}}=-D_{\mathrm{s}0}$.

Let $s\in[-s_{\max},s_{\max}]$ denote the (signed) *arc length* along the detector, where $s=0$ corresponds to the detector center. Assuming detector elements are equally spaced along the detector, arc length is the natural parameterization. The various angles have the following relationships:

$$\begin{aligned}\alpha(s)&=\frac{s}{D_{\mathrm{fd}}}\\ \gamma(s)&=\tan^{-1}\left(\frac{(D_{\mathrm{fd}})\sin\alpha(s)}{(D_{\mathrm{fd}})\cos\alpha(s)-D_{\mathrm{fs}}}\right)\end{aligned}$$

where the two most important cases are

$$\gamma(s)=\begin{cases}\alpha(s), & D_{\mathrm{fs}}=0\\ \frac{\tan^{-1}s}{D_{\mathrm{sd}}}, & D_{\mathrm{fs}}=\infty.\end{cases} \tag{33}$$

The relationship between s and γ is

$$s=\begin{cases}(D_{\mathrm{fd}})\left[\gamma-\arcsin\left(\frac{D_{\mathrm{fs}}}{D_{\mathrm{fd}}}\sin\gamma\right)\right], & 0\leq D_{\mathrm{fs}}<\infty\\ D_{\mathrm{sd}}\tan\gamma, & D_{\mathrm{fs}}=\infty.\end{cases} \tag{34}$$

The ray corresponding to detector element s and angle β is

$$\mathcal{L}(s,\beta)=\{(x,y):x\cos\varphi(s,\beta)+y\sin\varphi(s,\beta)=r(s)\}$$

where

$$\begin{aligned}\varphi(s,\beta)&\triangleq\beta+\gamma(s)\\ r(s)&\triangleq D_{\mathrm{s}0}\sin\gamma(s).\end{aligned} \tag{35}$$

The range of r depends on the position of the X-ray source and the size of the detector

$$|r(s)|\leq r_{\max}\triangleq D_{\mathrm{s}0}\sin\gamma_{\max} \tag{36}$$

where $\gamma_{\max}\triangleq\gamma(s_{\max})$ and $s_{\max}$ is half the total detector arc length. The radius $r_{\max}$ defines the circular *FOV* of the imaging system. The *fan angle* is $2\gamma_{\max}$.

The line-integral projection $p(s,\beta)$ of f along $\mathcal{L}(s,\beta)$ is[2]

$$\begin{aligned}&p(s,\beta)\\ &=\int_{\mathcal{L}(\mathrm{s},\beta)}f(x,y)\mathrm{d}\ell\\ &=\int\int f(x,y)\delta(x\cos\varphi(s,\beta)+y\sin(\varphi(s,\beta))-r(s))\mathrm{d}x\mathrm{d}y\end{aligned} \tag{37}$$

for $|s|\leq s_{\max}$ and $0\leq\beta<\beta_{\max}$. We require $\beta_{\max}\geq\pi+2\gamma_{\max}$ to ensure complete sampling of the FOV (otherwise the impulse response would be highly anisotropic) .

The usual inner product for fan-beam projection space is

$$\langle p_1,p_2\rangle=\int_{-s_{\max}}^{s_{\max}}\int_0^{\beta_{\max}}p_1(s,\beta)p_2(s,\beta)\mathrm{d}s\mathrm{d}\beta.$$

[2]Practically speaking, the integral should be restricted to the FOV: $\sqrt{x^2+y^2}\leq r_{\max}$, but this restriction would complicate analysis by introducing a shift variance into the problem, so we ignore it.

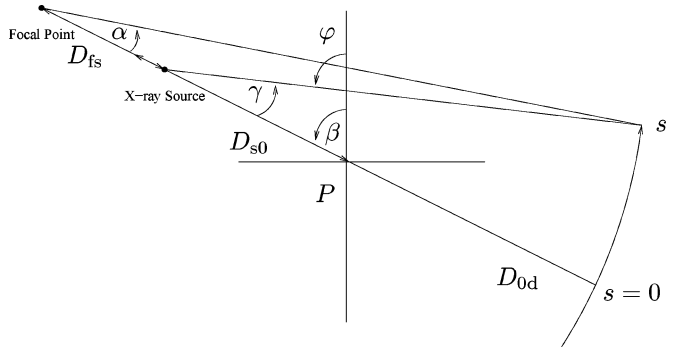

Fig. 15. Illustration of fan beam geometry.

Using this inner product in projection space, and the usual $\mathcal{L}^2$ inner product in image space, the adjoint of $\boldsymbol{\mathcal{P}}$ is given by

$$(\boldsymbol{\mathcal{P}}^* p)(x,y) = \int_{-\mathrm{s}_{\max}}^{s_{\max}} \int_0^{\beta_{\max}} p(s,\beta) \delta(x\cos\varphi(s,\beta) + y\sin\varphi(s,\beta) - r(s)) p(s,\beta) \mathrm{d}s\mathrm{d}\beta \quad (38)$$

where $r(s)$ and $\varphi(s,\beta)$ were defined in (35). We will next extend a common derivation of backproject then filter (BPF) tomographic reconstruction to accommodate a user defined weighting, and then change the coordinates from parallel-beam to fan-beam space.

Parallel-Beam Grammian Analysis: When we analyze the local impulse response, we typically consider recentered local impulse functions. In this derivation we will start with an uncentered local impulse response, and center it at the end by removing a phase term in the frequency domain. The uncentered local impulse response of the Grammian operator is shown in

$$\boldsymbol{\mathcal{P}}^*\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{P}}\boldsymbol{\delta}^j(x,y) = \int_0^{\pi}\int_{-\infty}^{\infty} \delta(x^j\cos(\phi) + y^j\sin(\phi) - r) w(r,\phi)\delta(x\cos(\phi) + y\sin(\phi) - r) dr d\phi.$$

Using the sampling property with the first δ, define

$$w^j(\phi) \triangleq w(r,\phi) = w(x^j\cos(\phi) + y^j\sin(\phi),\phi). \quad (39)$$

We denote the Fourier transform of $\delta^j(x,y)$ as $F^j(\rho,\Phi)$. Then, using the Fourier slice theorem

$$\begin{aligned}\boldsymbol{\mathcal{P}}^*\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{P}}\boldsymbol{\delta}^j(x,y) &= \int_0^{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F^j(\rho,\phi) e^{i2\pi\rho r} d\rho \\ &\quad w^j(\phi)\delta(x\cos\phi + y\sin\phi - r) dr d\phi \\ &= \int_0^{\pi}\int_{-\infty}^{\infty} F^j(\rho,\Phi) \\ &\quad w^j(\phi) e^{i2\pi\rho(x\cos\phi + y\sin\phi)} d\rho d\phi. \end{aligned} \quad (40)$$

This is nearly the inverse Fourier transform in signed polar coordinates except for a ρ scale factor. Dividing by ρ

$$\boldsymbol{\mathcal{P}}^*\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{P}}\boldsymbol{\delta}^j(x,y) = \int_0^{\pi}\int_{-\infty}^{\infty} \frac{w^j(\phi)F(\delta^j(x,y))}{\rho} e^{i2\pi\rho(x\cos\phi + y\sin\phi)} d\rho d\phi.$$

The local impulse response is recentered, $h^j(x,y) = \boldsymbol{\mathcal{P}}^*\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{P}}\boldsymbol{\delta}^j(x + x^j, y + y^j)$ which eliminates the complex phase term $F^j(\rho,\phi)$. Then, the local frequency response is

$$H^j(\rho,\Phi) = \frac{w^j(\Phi)}{\rho}. \quad (41)$$

Fan Beam Grammian Analysis: The natural indexing for fan-beam data is arc-length s and angle β. The analogs to r,ϕ for parallel-beam systems are $r(s),\varphi(s,\beta)$ as defined in (35). The weighting expression $w(r,\phi)$ is indexed as $w(s,\beta)$ in fan-beam coordinates. We start by looking at the fan-beam projection in terms of the analogs to parallel-beam coordinates

$$\begin{aligned}(\boldsymbol{\mathcal{P}}^*\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{P}}\boldsymbol{\delta}^j)(x,y) = \int_{-\beta_{\max}}^{\beta_{\max}}\int_{-\mathrm{s}_{\max}}^{s_{\max}} & w(s,\beta) \\ & \delta(x^j\cos(\varphi(s,\beta)) + y^j\sin(\varphi(s,\beta)) - r(s)) \\ & \delta(x\cos(\varphi(s,\beta)) + y\sin(\varphi(s,\beta)) - r(s)) ds d\beta.\end{aligned}$$

We use the change of variables

$$r' = r(s) = D_{\mathrm{s0}}\sin\gamma(s) \quad (42)$$

$$\varphi' = \varphi(s,\beta) = \beta + \gamma(s) \quad (43)$$

as defined in (35) which has the corresponding Jacobian determinant

$$J(s) = |D_{\mathrm{s0}}\cos\gamma(s)||\dot{\gamma}(s)|. \quad (44)$$

Then

$$\begin{aligned}(\boldsymbol{\mathcal{P}}^*\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{P}}\boldsymbol{\delta}^j)(x,y) = \int_0^{2\pi}\int_{-r_{\max}}^{r_{\max}} & w(r',\varphi') \\ & \delta(x^j\cos\varphi' + y^j\sin\varphi' - r') \\ & \delta(x\cos\varphi' + y\sin\varphi' - r')\frac{1}{J(s(r'))} dr' d\varphi'.\end{aligned}$$

In this expression

$$\begin{aligned} w(r,\varphi) &\triangleq w(s(r),\beta(r,\varphi)) \\ s(r) &= \gamma^{-1}\left(\arcsin\left(\frac{r}{D_{\mathrm{s0}}}\right)\right) \\ \beta(r,\varphi) &= \varphi - \arcsin\left(\frac{r}{D_{\mathrm{s0}}}\right). \end{aligned}$$

Using the sampling property of the first δ as in the parallel beam case

$$w^j(\varphi') \triangleq w(x^j\cos\varphi' + y^j\sin\varphi',\varphi') \quad (45)$$

$$s^j(\varphi') \triangleq s(r')|_{r'=x^j\cos\varphi' + y^j\sin\varphi'}. \quad (46)$$

Again, let $F^j(\rho,\Phi)$ denote the Fourier transform of $\delta^j(x,y)$. Then, using the Fourier slice theorem

$$\begin{aligned}&(\boldsymbol{\mathcal{P}}^*\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{P}}\boldsymbol{\delta}^j)(x,y) \\ &= \int_{-r_{\max}}^{r_{\max}}\int_0^{2\pi}\int_{-\infty}^{\infty} \frac{F^j(\rho,\Phi)w^j(\varphi')}{J(s^j(\varphi'))} e^{i2\pi\rho r'} \\ &\quad \delta(x\cos\varphi' + y\sin\varphi' - r') d\rho d\varphi' dr'\end{aligned}$$

$$= \int_0^{2\pi} \int_{-\infty}^{\infty} \frac{F^j(\rho,\Phi) w^j(\varphi)}{J(s^j(\varphi'))} e^{i2\pi\rho(x\cos\varphi' + y\sin\varphi')} d\rho d\varphi'$$
$$= \int_0^{\pi} \int_{-\infty}^{\infty} \frac{F^j(\rho,\Phi) w^j(\varphi')}{J(s^j(\varphi'))} e^{i2\pi\rho(x\cos\varphi' + y\sin\varphi')} d\rho d\varphi'$$
$$+ \int_0^{\pi} \int_{-\infty}^{\infty} \frac{F^j(\rho,\Phi) w^j(\varphi'')}{J(s^j(\varphi''))} e^{i2\pi\rho(x\cos\varphi'' + y\sin\varphi'')} d\rho d\varphi''$$

where $\varphi'' = \varphi' + \pi$. This is similar to the parallel-beam derivation except that we have two integrals. We can convert each integral into the inverse Fourier transform as we did in the parallel-beam case and strip out the phase term $F^j(\rho,\Phi)$ by recentering the local impulse response. The local frequency responses of the Grammian operator is

$$H^j(\rho,\Phi) = \frac{\tilde{w}^j(\Phi)}{|\rho|} \tag{47}$$

where

$$\tilde{w}^j(\Phi) = \frac{1}{J(s^j(\Phi))} \left[w^j(\varphi)\big|_{\varphi=\Phi} + w^j(\varphi)\big|_{\varphi=\Phi+\pi} \right]. \tag{48}$$

Because of the absolute value function in (44), $J(s)$ is invariant to the π phase shift. For the case where we have uniform weighting, $w(s,\beta) = 1$ and therefore $\boldsymbol{W} = \boldsymbol{I}$, and (47) simplifies to

$$H^j(\rho,\Phi) = \frac{2}{J(s^j(\Phi))|\rho|}. \tag{49}$$

We use this equation in the calculation of a target local impulse response (24).

Appendix C
Derivations of the Analytical Solution

In this appendix, we show that $\sqrt{d_2^2 + d_3^2} \le d_1$, as claimed below (31) in Section IV-C. Squaring the integrals in (31), we have

$$d_1^2 = \frac{1}{\pi} \int_0^{\pi} \int_0^{\pi} w^j(X) w^j(Y) dX dY$$
$$d_2^2 = \frac{1}{\pi} \int_0^{\pi} \int_0^{\pi} \cos(2X) w^j(X) \cos(2Y) w^j(Y) dX dY$$
$$d_3^2 = \frac{1}{\pi} \int_0^{\pi} \int_0^{\pi} \sin(2X) w^j(X) \sin(2Y) w^j(Y) dX dY.$$

In particular

$$d_2^2 + d_3^2 = \frac{1}{\pi} \int_0^{\pi} \int_0^{\pi} w^j(X) w^j(Y) [\cos(2X)\cos(2Y) + \sin(2X)\sin(2Y)] dX dY$$
$$= \frac{1}{\pi} \int_0^{\pi} \int_0^{\pi} w^j(X) w^j(Y) \cos(2X - 2Y) dX dY.$$

Thus, $d_2^2 + d_3^2 \le d_1^2$ since $w^j(\Phi) \ge 0$ for all Φ and $\cos(2X - 2Y) \le 1$ for all X and Y.

Acknowledgment

The authors would like to thank GE and J. B. Thibault for giving them the real CT data used in this paper. The authors would also like to thank the reviewers who suggested that we examine approximation in (17) which led to Appendix A.

References

[1] J. A. Fessler, "Statistical image reconstruction methods for transmission tomography," in *Handbook of Medical Imaging*, M. Sonka and J. M. Fitzpatrick, Eds. Bellingham, WA: SPIE, 2000, vol. 2, Medical Image Processing and Analysis, pp. 1–70.

[2] J. A. Fessler, "Penalized weighted least-squares image reconstruction for positron emission tomography," *IEEE Trans. Med. Imag.*, vol. 13, pp. 290–300, Jun. 1994.

[3] J. A. Fessler and W. L. Rogers, "Spatial resolution properties of penalized-likelihood image reconstruction methods: Space-invariant tomographs," *IEEE Trans. Image Process.*, vol. 5, no. 9, pp. 1346–58, Sep. 1996.

[4] P. M. Joseph, "The influence of modulation transfer function shape on CT image quality," *Radiology*, vol. 145, no. 1, pp. 179–85, Oct. 1982.

[5] J. W. Stayman and J. A. Fessler, "Regularization for uniform spatial resolution properties in penalized-likelihood image reconstruction," *IEEE Trans. Med. Imag.*, vol. 19, no. 6, pp. 601–15, Jun. 2000.

[6] J. W. Stayman and J. A. Fessler, "Compensation for nonuniform resolution using penalized-likelihood reconstruction in space-variant imaging systems," *IEEE Trans. Med. Imag.*, vol. 23, no. 3, pp. 269–84, Mar. 2004.

[7] J. Nuyts and J. A. Fessler, "A penalized-likelihood image reconstruction method for emission tomography, compared to post-smoothed maximum-likelihood with matched spatial resolution," *IEEE Trans. Med. Imag.*, vol. 22, no. 9, pp. 1042–52, Sep. 2003.

[8] J. Qi and R. M. Leahy, "Resolution and noise properties of MAP reconstruction for fully 3-D PET," *IEEE Trans. Med. Imag.*, vol. 19, no. 5, pp. 493–506, May 2000.

[9] H. Shi and J. A. Fessler, "Quadratic regularization design for 3-D cylindrical PET," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, 2005, vol. 4, pp. 2301–5.

[10] J. A. Fessler, "Analytical approach to regularization design for isotropic spatial resolution," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, 2003, vol. 3, pp. 2022–6.

[11] H. Shi and J. A. Fessler, "Quadratic regularization design for fan beam transmission tomography," in *Proc. SPIE*, 2005, vol. 5747, pp. 2023–33.

[12] J. A. Fessler, Resolution Properties of Regularized Image Reconstruction Methods Dept. Elct. Eng. Comput. Sci., Commun. Signal Process. Lab., Univ. Michigan, Ann Arbor, Tech. Rep. 297, 1995.

[13] Q. Li, E. Asma, J. Qi, J. R. Bading, and R. M. Leahy, "Accurate estimation of the Fisher information matrix for the PET image reconstruction problem," *IEEE Trans. Med. Imag.*, vol. 23, no. 9, pp. 1057–64, Sep. 2004.

[14] 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.*, vol. 12, no. 1, pp. 78–83, Mar. 1993.

[15] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. Upper Saddle River, NJ: Prentice-Hall, 1974.

[16] J. A. Fessler, Matlab Tomography Toolbox 2004 [Online]. Available: http://www.eecs.umich.edu/~xfessler

[17] B. De Man and S. Basu, "Distance-driven projection and backprojection," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, 2002, vol. 3, pp. 1477–80.

[18] 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.*, vol. 34, no. 11, pp. 4526–44, Nov. 2007.

[19] J.-B. Thibault, C. A. Bouman, K. D. Sauer, and J. Hsieh, "A recursive filter for noise reduction in statistical iterative tomographic imaging," in *Proc. SPIE*, 2006, vol. 6065, p. 60650X.

[20] S. Ahn and R. M. Leahy, "Spatial resolution properties of nonquadratically regularized image reconstruction for PET," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Apr. 2006, pp. 287–90.