

In *Proceedings, 2002 IEEE International Symposium on Biomedical Imaging:  
Macro to Nano*, Washington, D.C., July 7–10, 2002.

# DECOUPLING THE EQUATIONS OF REGULARIZED TOMOGRAPHY

Jonas August\*

jonas@cs.cmu.edu

Healthcare Robotics Center

The Robotics Institute

Carnegie Mellon University

5000 Forbes Avenue

Pittsburgh, PA 15213

---

\*I thank Takeo Kanade for his insightful suggestions (especially on interpolation), and Branko Jaramaz for introducing me to the metal artifact reduction problem. The NSF Advanced Computational Research program and the Heinz Endowment provided funding.

## Abstract

Deferring discretization can occasionally change our perspective on imaging problems. To illustrate, we offer a reformulation of regularized computed tomography (CT) in which the large system of coupled equations for the unknown smoothed image is *decoupled* into many smaller and simpler equations, each for a separate projection. Regularized CT thus becomes a two-stage process of (nonhomogeneous) smoothing of the projections followed by filtered backprojection. As a by-product, the repeated forward and backprojections common in iterative image reconstruction are eliminated. Despite the computational simplification, we demonstrate that this method can be used to reduce metal artifacts in X-ray CT images. The decoupling of the equations results from postponing the discretization of image derivatives that realize the smoothness constraint, allowing for this constraint to be analytically “transferred” from the image domain to the projection, or Radon, domain. Our analysis thus clarifies the role of image smoothness: it is an entirely intra-projection constraint.

## 1 Background

In the absence of noise, the basic problem of computerized tomography (CT) is to determine an unknown image  $f = f(x, y)$  from its (forward) projections, or *Radon transform*  $\mathcal{R}f$ , where  $(\mathcal{R}f)(t, \theta) := \iint f(x, y)\delta(t - x \cos \theta - y \sin \theta) dx dy$ , where  $(x, y)$  are planar coordinates,  $t$  is the location along each projection, and  $\theta \in [0, \pi)$  is the orientation of the projection. (In this paper we focus on the two dimensional problem with standard parallel-beam geometry, but the ideas readily extend to three dimensions and other scanning geometries.) Unfortunately, since

measurements are never perfect, what we actually observe are the noisy projection data  $g = g(t, \theta)$ . (To emphasize the essentials of the tomography problem, we view the unknown  $f = f(x, y)$  and the observation  $g = g(t, \theta)$  as functions, although the implementation is discrete. See §4.) We make the standard independence and locality assumptions that the likelihood  $P(g|f)$ , or conditional distribution of  $g$  given  $f$ , equals  $\prod p(g(t, \theta)|(\mathcal{R}f)(t, \theta))$ , where the product is over all  $(t, \theta)$ . (This product is nonzero because we take a finite number of factors in our implementation.) Since the likelihood depends on the particular imaging modality, we illustrate using X-ray CT for concreteness, although our technique applies more broadly (e.g., to PET). Here, therefore,  $b \exp(-g(t, \theta))$  is the observed number of X-ray photons, Poisson-distributed with mean  $b \exp(-(\mathcal{R}f)(t, \theta))$ , where  $b$  is the mean number of incident photons and  $f(x, y)$  is the attenuation coefficient at  $(x, y)$ . Our task is to infer the image  $f$  given the noisy projections  $g$ .

## 2 Regularized Tomography

Because this inverse problem is ill-posed [1], one typically imposes extra constraints on  $f$ . In penalized maximum likelihood [2], or *regularization*, inferring  $f$  amounts to finding that  $f$  which minimizes  $-\ln P(g|f) + \rho(f)$ , where  $\rho(f)$  characterizes the extra constraint on  $f$ . Here we impose the standard smoothness constraint that uses the image gradient  $\nabla f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$  in the quadratic penalty  $\rho(f) := \beta \|\nabla f\|^2$ , where  $\beta > 0$ ,  $\|f\|^2 := \langle f, f \rangle$ , and  $\langle f_1, f_2 \rangle := \int f_1 f_2$  is an inner product. Following [3] and to simplify the presentation, we approximate  $-\ln P(g|f)$  with the quadratic form  $\|g - \mathcal{R}f\|_{\mathcal{W}}^2$ , where  $\|g\|_{\mathcal{W}}^2 = \langle g, \mathcal{W}g \rangle$  is a weighted norm with (diagonal) *weight operator*  $\mathcal{W}$  satisfying  $(\mathcal{W}g)(t, \theta) :=$

$w(t, \theta)g(t, \theta)$ . The weight  $w(t, \theta) := b \exp(-g(t, \theta))$  is small for those rays passing through dense materials such as bone or metal, and larger otherwise. This formulation of tomography requires that we solve the following “hard” optimization problem:

**Problem 1 (Regularized CT).** *Given the projection data  $g$ , find the image  $f$  that minimizes*

$$\|g - \mathcal{R}f\|_{\mathcal{W}}^2 + \beta \|\nabla f\|^2.$$

To proceed, let  $\Delta$  denote  $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ , the Laplacian in the plane. Using integration by parts and zero boundary conditions, we recall that  $\|\nabla f\|^2 = \langle f, -\Delta f \rangle$ . Then the (linear) Euler-Lagrange equation for Problem 1 is

$$\mathcal{R}^* \mathcal{W} \mathcal{R} f - \beta \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) f = \mathcal{R}^* \mathcal{W} g, \quad (1)$$

where  $f$  is unknown and  $A^*$  denotes the adjoint of linear operator  $A$  ( $\mathcal{R}^*$  is also known as the backprojection operator). By examining (1), we see that Problem 1 is hard in two related ways.

First, the problem constraints occur in two different domains. Fidelity to the data ( $\|g - \mathcal{R}f\|_{\mathcal{W}}^2$ ) is enforced in the Radon domain  $\{(t, \theta)\}$ , while smoothness ( $\|\nabla f\|^2$ ) is imposed in the image domain  $\{(x, y)\}$ . Thus we see in (1) the operators  $\mathcal{R}$  and  $\mathcal{R}^*$  for shuffling back and forth between these domains; iterative solution techniques typically compute these forward and backprojections explicitly and often at great expense or inconvenience.

Second, observe that (1) is a *coupled* equation in the two-variable function  $f = f(x, y)$ , i.e., in the large set of variables  $\{f(x, y), \text{ for all } x, y\}$  under some discretization of  $x$  and  $y$ . The coupling arises first because both  $x$ - and  $y$ - derivatives are present; in addition,  $\mathcal{R}$  and  $\mathcal{R}^*$  are integral operators, and so are not even

local. The computational difficulty in solving (1) and related tomographic problems (e.g., emission) has spawned a great deal of work in optimization [4, 5, 6].

### 3 Decoupled Regularized CT

We seek to formulate the entire regularization problem in a single domain. As we shall see, working solely in the Radon domain will decouple our large joint optimization problem into many smaller ones. But first we review the standard technique for inverting the Radon transform: filtered backprojection. Let the Fourier transform of  $g = g(t, \theta)$  with respect to  $t$  be denoted  $(\mathcal{F}_1 g)(\tau, \theta) = (2\pi)^{-\frac{1}{2}} \int g(t, \theta) e^{-i\tau t} dt$ , where  $\tau$  is the spatial frequency along the  $\theta$ -projection.

**Definition 1.** *Given the function  $h = h(t, \theta)$ , the **Riesz potential** is the linear operator  $\mathcal{I}^\alpha$  satisfying  $(\mathcal{F}_1 \mathcal{I}^\alpha h)(\tau, \theta) = |\tau|^{-\alpha} (\mathcal{F}_1 h)(\tau, \theta)$ .*

Given noise-free observations  $h = \mathcal{R}f$ , one can solve for the unknown  $f$  by directly implementing the following classical formula [1] for the inverse of the Radon transform:

**Fact 1 (Filtered Backprojection).**  $\mathcal{R}^{-1} = \frac{1}{4\pi} \mathcal{R}^* \mathcal{I}^{-1}$ .

By defining  $\square$  as the Laplacian  $\frac{\partial^2}{\partial t^2}$  along each projection, we note that  $\mathcal{I}^{-1}$  is the “square root” of  $-\square$ . This is because  $\mathcal{I}^{\alpha_1} \mathcal{I}^{\alpha_2} = \mathcal{I}^{\alpha_1 + \alpha_2}$  and the following:

**Fact 2.**  $\mathcal{I}^{-2} = -\square$ .

The proof is that the Fourier transform of  $-\square$  is  $\tau^2 = |\tau|^2$ .

Now recall the Fourier slice theorem, which says that the two-dimensional Fourier transform of  $f$ , evaluated at polar coordinates  $(\tau, \theta)$ , is just  $(\mathcal{F}_1 \mathcal{R}f)(\tau, \theta)$ .

Using this theorem, one can relate the two Laplacians  $\Delta$  and  $\square$  because the two-dimensional Fourier transform of  $-\Delta$  is  $u^2 + v^2$  (where  $u$  and  $v$  are spatial frequencies for  $x$  and  $y$ , respectively), or  $|\tau|^2$  in polar coordinates, which is the one-dimensional Fourier transform of  $-\square$ . For details and the extension to higher dimensions, see [1] and [7]. This crucial but simple idea is called “**intertwining**”.

**Fact 3.** *The Radon transform  $\mathcal{R}$  intertwines  $\Delta$  and  $\square$ , i.e.,*

$$\mathcal{R}\Delta = \square\mathcal{R}.$$

*Our main contribution is the realization that by applying intertwining to regularized CT, we can “decouple” this large optimization problem into an equivalent set of much smaller optimization problems.* The idea is to reformulate Problem 1 in terms of  $h = \mathcal{R}f$ . We need only transfer the smoothness constraint to the Radon domain, as the data constraint is already naturally specified there. Intertwining allows us to analytically transfer the smoothness constraint to the Radon domain, as opposed to numerically enforcing it during optimization computations.

**Proposition 1 (Smoothness Constraint to Radon Domain).**

$$\|\nabla f\|^2 = (4\pi)^{-1}\langle h, \mathcal{I}^{-3}h \rangle, \quad \text{where } h = \mathcal{R}f.$$

For proof, note that  $-\|\nabla f\|^2 = \langle f, \Delta f \rangle = \langle R^{-1}h, \Delta R^{-1}h \rangle = \langle h, \mathcal{R}^{-1*} \Delta R^{-1}h \rangle$ . Now, note that  $\mathcal{R}^{-1*} = (4\pi)^{-1}\mathcal{I}^{-1}\mathcal{R}$ , using Fact 1 and the symmetry of  $\mathcal{I}^\alpha$ . But then  $-\mathcal{R}^{-1*} \Delta R^{-1} = -(4\pi)^{-1}\mathcal{I}^{-1}\mathcal{R}\Delta R^{-1} = -(4\pi)^{-1}\mathcal{I}^{-1}\square\mathcal{R}\mathcal{R}^{-1}$ , using Fact 3. Prop. 1 follows using Fact 2 and because  $\mathcal{R}\mathcal{R}^{-1}$  is the identity operator.

Thus we can pose Problem 1 in an equivalent, “easy” form in the Radon domain, as follows:

**Problem 2 (Decoupled Regularized CT).** Given observed projections  $g$ , find projections  $h = \mathcal{R}f$  minimizing

$$\|g - h\|_{\mathcal{W}}^2 + \beta' \langle h, \mathcal{I}^{-3}h \rangle, \quad \text{where } \beta' := (4\pi)^{-1}\beta.$$

The corresponding Euler-Lagrange equation,

$$\mathcal{W}h + \beta' \mathcal{I}^{-3}h = \mathcal{W}g, \quad (2)$$

where  $h$  is unknown, is easy exactly where (1) is hard. First, the forward and backprojections are eliminated from the optimization; backprojection need only be done once to determine  $f$  from solution  $h$ . Second, and more importantly, equation (2) is really a decoupled set of systems of equations, where each system corresponds to the unknowns  $\{h(t, \theta), \text{ for all } t\}$ , at each *fixed*  $\theta$ . This follows because operator  $\mathcal{W}$  is pointwise multiplication by a scalar and  $\mathcal{I}^{-3}$  acts only along  $t$ . Thus our “decoupled” approach to regularized CT requires solving the integral equation (2) in the unknown single variable function  $h(\cdot, \theta)$ , for each fixed  $\theta$ .

These decoupled regularization equations (2) are related to direct algebraic reconstruction tomography (DART) by emphasizing optimization entirely in the Radon domain [1] (although DART is formulated discretely). However, DART usually assumes that the weights  $w(t, \theta)$  are constant in  $t$ , which is often not the case (e.g., X-rays through bone or metal have lower weight than those through flesh). Thus decoupled regularization gains the benefits of space-varying filtering while maintaining the simplicity of DART.

One limitation of the proposed approach is that it does not impose consistency conditions [8, 9]. In addition, we assume the idealized line-integral model of projection, while discrete regularization techniques more accurately model real

scanners [12]. Nonetheless, we can still demonstrate (§4) that our method can remove noise.

## 4 Results

To solve (2), we observe that  $\mathcal{W} + \beta' \mathcal{I}^{-3}$  is a positive definite operator (if  $w(t, \theta) > 0$ ), and thus the conjugate gradient method can be applied. To discretize the equations, we sampled in  $t$  and  $\theta$  uniformly. The operator  $\mathcal{W}$  was implemented by restriction to the sample locations. The Riesz potential was implemented by taking 1-dimensional FFTs.

We applied our decoupled regularized CT method to the reduction of metal artifacts in X-ray CT (Fig. 1) [10, 11, 12]. Since the actual projection data and scanner parameters were unavailable, we simulated the projections (using Matlab) after rescaling the image pixel values (range 0 to 255) by 0.012 (Fig. 3, top left). The bright band results from the metal, and is noisier than elsewhere; the noise is obvious in the single projection in Fig. 2 (solid curve). For each fixed projection orientation  $\theta$ , the decoupled regularized CT equation (2) was solved ( $b = 10^9, \beta = 10^3$ ) to produce the nonhomogeneously smoothed projection shown in Fig. 2 (dotted curve). After smoothing each projection independently, filtered backprojection was applied to the *set* of smoothed projections (Fig. 3, top right), producing our final result (Fig. 4), which shows reduced streaking artifacts.

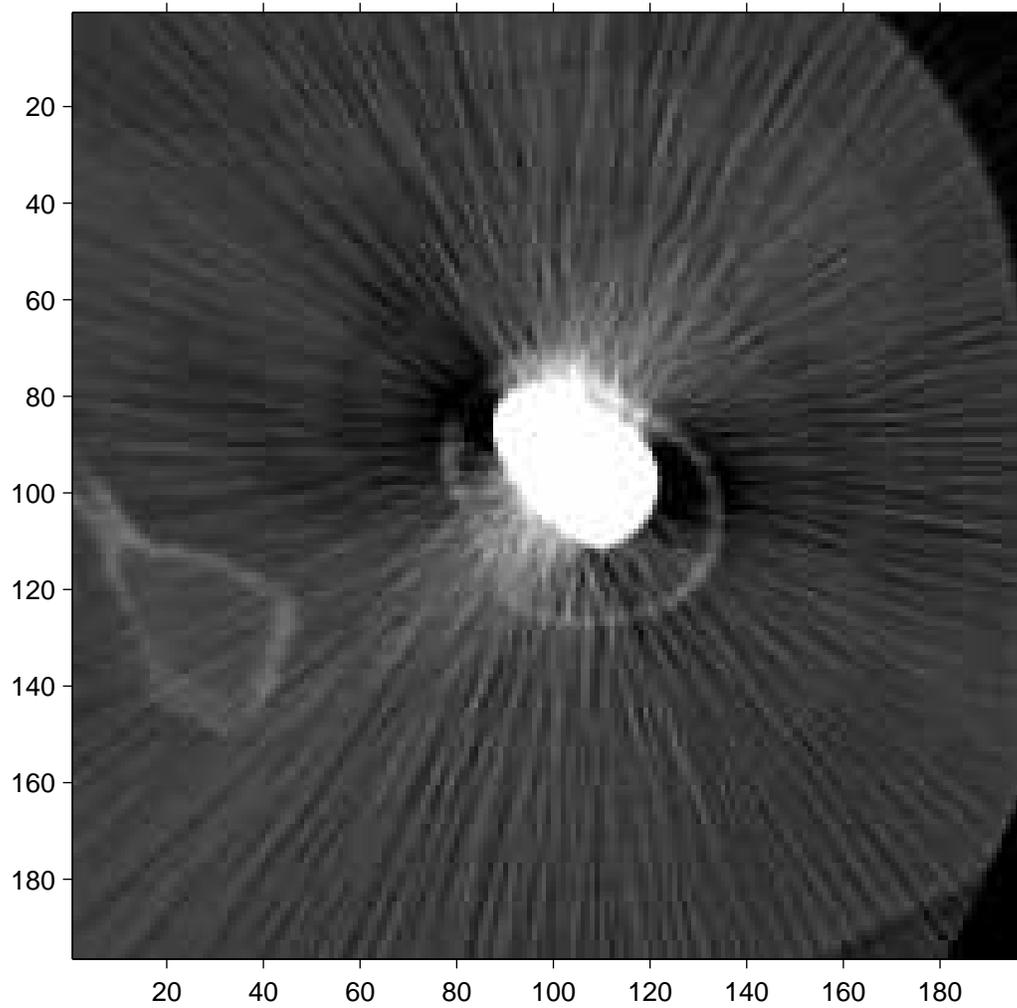


Figure 1: Streaking artifacts are due to the presence of metal (bright oval) in this portion of an X-ray CT slice of a hip.

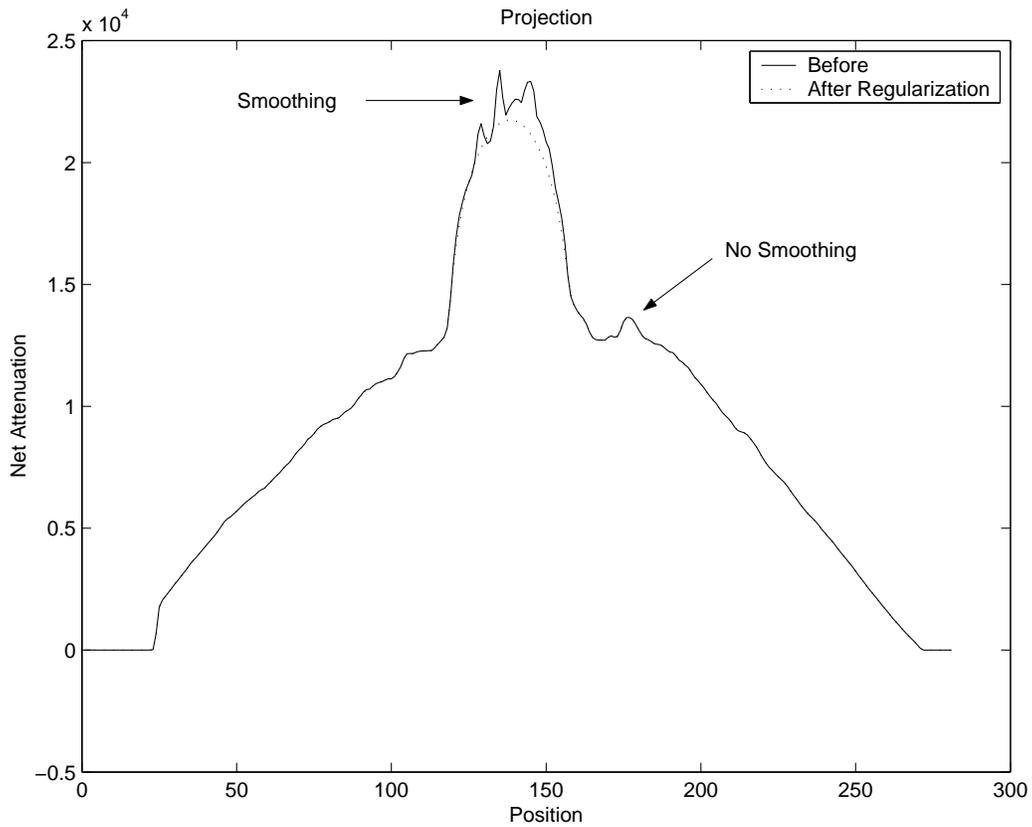


Figure 2: A simulated projection (solid curve,  $g(\cdot, \theta)$ ) oriented at  $\theta \approx 45^\circ$  is noisy particularly in the portion due to the metal (large hump in the center). The smoothing due to decoupled regularized CT (dotted curve,  $h(\cdot, \theta)$ ) is greater in the metal portion, where the weights are automatically lowest.

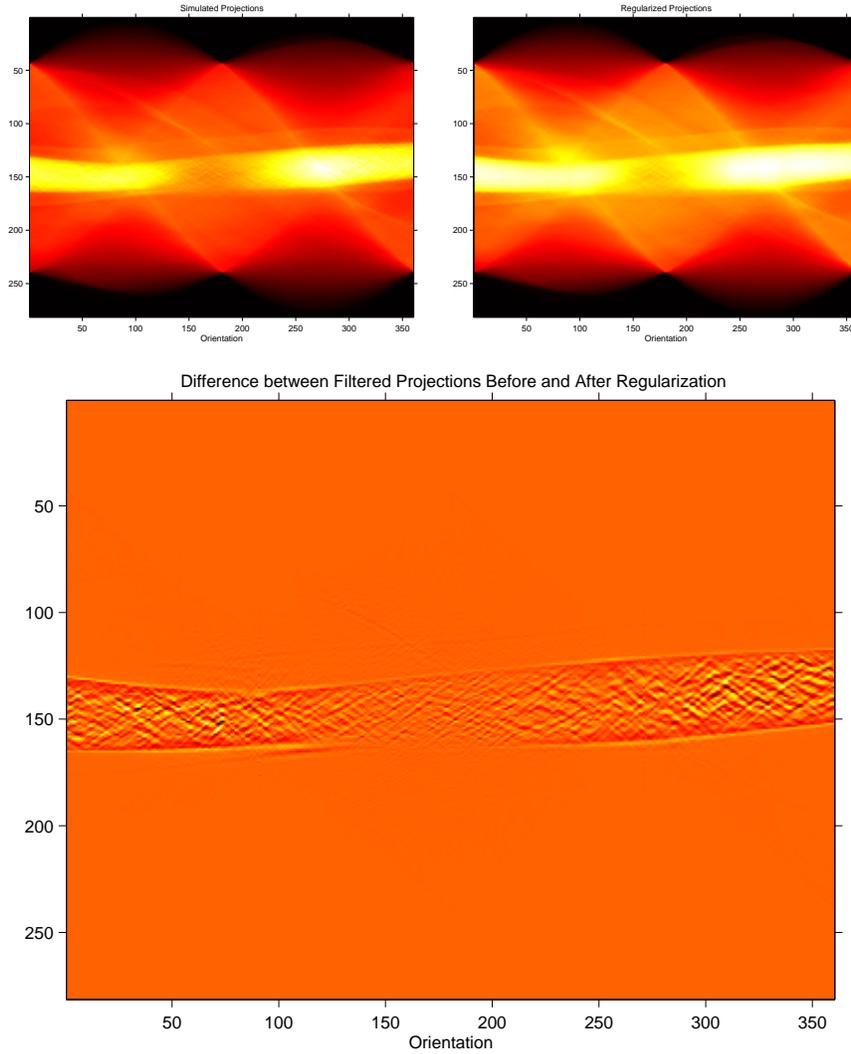


Figure 3: The simulated projection data  $g(t, \theta)$  (top left), or Radon transform, of Fig. 1 shows a bright band due to the metal. The key smoothing action of decoupled regularized CT ( $h(t, \theta)$ , top right) is localized on the metal band, as emphasized (bottom) in the result of applying  $\mathcal{I}^{-1}$  (a kind of differentiation and the first step in filtered backprojection) to the difference  $g - h$ .

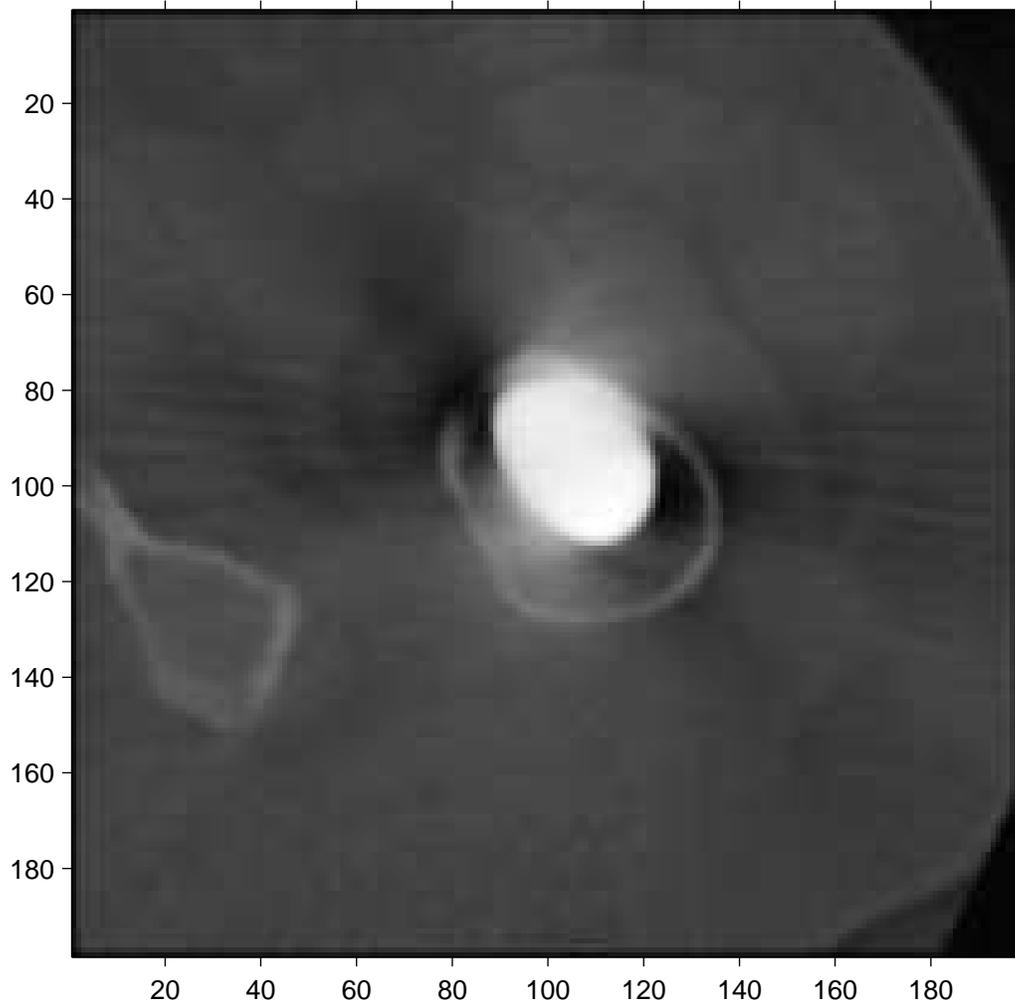


Figure 4: The result of decoupled regularized CT for reducing streaking artifacts. Observe that the smoothing effect is primarily near the metal; the rest of the image is still sharp.

## 5 Conclusion

By studying regularized tomography in the continuous domain, we were able to decouple a linear equation in a two-variable function into a one-parameter family of linear equations in single-variable functions. Although we emphasized a quadratic approximation of the X-ray CT likelihood, our result only hinges upon Prop. 1, which is independent of the imaging modality. For example, one can set up decoupled nonlinear equations to fully capture the Poisson likelihood. This technique can also be applied in three dimensions by extending Facts 1-3. Finally, the gradient smoothness term can be replaced with related higher-order derivative penalties while preserving the decoupling of the regularization.

## References

- [1] F. Natterer, *The Mathematics of Computerized Tomography*, Wiley, Chichester, 1986.
- [2] Jeffrey A. Fessler and Scott D. Booth, “Conjugate-gradient preconditioning methods for shift-variant pet image reconstruction,” *IEEE Trans. on Image Proc.*, vol. 8, no. 5, pp. 688–699, 1999.
- [3] Ken Sauer and Charles Bouman, “A local update strategy for iterative reconstruction from projections,” *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 534–548, 1993.
- [4] Jeffrey A. Fessler, Edward P. Ficaro, Neal H. Clinthorne, and Kenneth Lange, “Grouped-coordinate ascent algorithms for penalized-likelihood

- transmission reconstruction,” *IEEE Trans. on Medical Imaging*, vol. 16, no. 2, pp. 166–175, 1997.
- [5] Erkan U. Mumcoglu, Richard M. Leahy, and Simon R. Cherry, “Bayesian reconstruction of pet images: methodology and performance analysis,” *Phys. Med. Biol.*, vol. 41, pp. 1777–1807, 1996.
- [6] Jolyon Browne and Alvaro R. De Pierro, “A row-action alternative to the em algorithm for maximizing likelihoods in emission tomography,” *IEEE Trans. on Medical Imaging*, vol. 15, no. 5, pp. 687–699, 1996.
- [7] Sigurdur Helgason, *The Radon Transform*, Birkhauser, Boston, 2nd edition, 1999.
- [8] R. M. Lewitt and R. H. T. Bates, “Image reconstruction from projections: Iii: Projection completion methods (theory),” *Optik*, vol. 50, no. 3, pp. 189–204, 1978.
- [9] Hiroyuki Kudo and Tsuneo Saito, “Sinogram recovery with the method of convex projections for limited-data reconstruction in computed tomography,” *J. Opt. Soc. Am. A*, vol. 8, no. 7, pp. 1148–1160, 1991.
- [10] Ge Wang, Troy Frei, and Michael W. Vannier, “Fast iterative algorithm for metal artifact reduction in x-ray ct,” *Acad. Radiol.*, vol. 7, pp. 607–614, 2000.
- [11] Jiang Hsieh, “Adaptive streak artifact reduction in computed tomography resulting from excessive x-ray photon noise,” *Med. Phys.*, vol. 25, no. 11, pp. 2139–2147, 1998.

- [12] B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens, "Reduction of metal streak artifacts in x-ray computed tomography using a transmission maximum a posteriori algorithm," *IEEE Trans. on Nuclear Science*, vol. 47, no. 3, pp. 977–981, 2000.