# Multilevel Image Reconstruction with Natural Pixels

Van Emden Henson*  Mark A. Limber†

Stephen F. McCormick‡  Bruce T. Robinson§

January 14, 1995

**Abstract**

The sampled Radon transform of a 2D function can be represented as a continuous linear map $A : L_2(\Omega) \to \mathbf{R}^N$, where $(Au)_j = \langle u, \psi_j \rangle$ and $\psi_j$ is the characteristic function of a strip through $\Omega$ approximating the set of line integrals in the sample. The image reconstruction problem is: given a vector $\mathbf{b} \in \mathbf{R}^N$, find an image (or density function) $u(x, y)$ such that $Au = \mathbf{b}$. In general there are infinitely many solutions; we seek the solution with minimal 2-norm, which leads to a matrix equation $B\mathbf{w} = \mathbf{b}$, where $B$ is a square dense matrix with several convenient properties. We analyze the use of Gauss-Seidel iteration applied to the problem, observing that while the iteration formally converges, there exists a *near null space* into which the error vectors migrate, after which the iteration stalls. The null space and near null space of $B$ are characterized in order to develop a multilevel scheme. Based on the principles of the Multilevel Projection Method (PML), this scheme leads to somewhat improved performance. Its primary utility, however, is that it facilitates the development of a PML-based method for *spotlight tomography*, that is, local grid refinement over a portion of the image in which features of interest can be resolved at finer scale than is possible globally.

## 1 Introduction

In this paper, we consider a model of transmission and emission tomography and an associated image reconstruction technique. The reconstruction technique approximates a minimum norm solution to an underdetermined linear inversion problem, based on an infinite-dimensional formulation of the tomographic inversion problem. This formulation of the problem avoids the traditional square pixel discretization of the image space and leads to a smaller, but dense, matrix problem (compared to traditional algebraic reconstruction techniques). This approach leads to what have been termed "natural pixels" in [1], and the "optimal grid" in [10].

Following the development of the natural pixel discretization, we consider solution techniques for the resulting linear system. In particular, we employ Gauss-Seidel iteration, analyze its performance, and then introduce a multilevel projection method (PML) for accelerating convergence.

*Department of Mathematics, Naval Postgraduate School, email: vhenson@math.nps.navy.mil

†Auto-trol Technology Corporation, Denver, CO, email: marlim@Auto-trol.COM. Formerly at CECM, Simon Fraser University

‡Program in Applied Mathematics, University of Colorado, email: stevem@newton.colorado.edu

§Accurate Information Systems, Eatontown, NJ, email: brobinso@Accurate.COM. Formerly at Mathematics Department, Naval Postgraduate School

## 2 Image Reconstruction and the Radon Transform

We formulate the image reconstruction from projection problems in a general setting, but concentrate on a parallel beam geometry, for which we have implemented our ideas. The basic idea in tomography is that an object is subjected to a dose of radiation, either by passing X-rays through the object, or (if the object is a living patient) by administering a radiopharmeceutical. The amount of radiation leaving the object can be measured, compared with the original amount, and the difference is a measurement of the attenuation (transmission tomography) or activity (emission tomography) within the object. In parallel beam geometry, the data is collected in collimated bins, so that any activity detected in a particular bin can be attributed to the strip eminating perpendicularly out of the detector, with width equal to that of the bin.

To model this apparatus, let $u(x, y)$ be a function of the spatial variables $x$ and $y$ describing the activity in the object. Typically this is some physical quantity, such as the material density of the subject. The vector $\mathbf{f}$ represents the projection data. The data aquisition is modeled by

$$Au = \begin{pmatrix} \int_{\mathbf{R}^2} u(x,y)\psi_1(x,y)dxdy \\ \vdots \\ \int_{\mathbf{R}^2} u(x,y)\psi_N(x,y)dxdy \end{pmatrix} = \begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix} = \mathbf{f}, \tag{1}$$

where the function $\psi_k$ is the characteristic function of the $k$th strip through the image, within which passes (or emanates) the energy collected by the $k$th detector. For this to be well defined, we restrict our function space to be $L_2(\Omega)$, where $\Omega$ is a compact subset of $\mathbf{R}^2$, called the *image space*. Thus,

$$(Au)_j = \langle \psi_j, u \rangle$$

defines a continuous linear map $A : L_2(\Omega) \to \mathbf{R}^N$. The basic problem of computer assisted tomography is to reconstruct the image $u(x, y)$ from a collection of measured strip integrals, collected at various angles. When this problem can be solved, it is done through some approximate inversion of the Radon transform, which is defined as follows.

Let $u(x, y)$ be a function defined on the region $\Omega \in \mathbf{R}^2$. Letting $L$ denote any line in $\mathbf{R}^2$, the set of line integrals of $u(x, y)$, along all possible lines $L$, is a function of two variables, and is known as the *Radon transform* of $u(x, y)$, provided the integral exists. Formally,

$$[Ru](\rho, \phi) \;=\; \int_L u(x,y)ds = \int u(x,y)\delta(x\cos\phi + y\sin\phi - \rho)dxdy$$

where $\delta$ is the Dirac delta function. The line $L$ is parametrized by

$$\rho \;=\; x\cos\phi + y\sin\phi, \tag{2}$$

where $\rho$ is the signed distance from the origin and $\phi$ is an angle measured counter-clockwise from the positive x-axis. Thus, (2) determines the equation of a line in the $xy$-plane normal to the unit vector $\vec{\xi} \;=\; (\cos\phi, \sin\phi)^T$. Figure 1 shows the geometry of the Radon transform of a function $u(x, y)$ in terms of this parameterization.

Viewing the Radon transform as an operator, the image reconstruction problem can be cast as $Ru \;=\; f$, where $f$ represents the collection of measured line integrals. Given $\mathbf{f}$, a finite sampling of $f$, we model the problem $Ru = f$ with $Au = \mathbf{f}$, since each of the strip integrals $\langle \psi_j, u \rangle$ approximates a collection of line integrals, for those lines falling within the strip. Hence, the data $\mathbf{f}$ forms a sampling of the continuous Radon transform. We will refer to the set of strip integrals (1) as the *strip averaged* Radon transform.

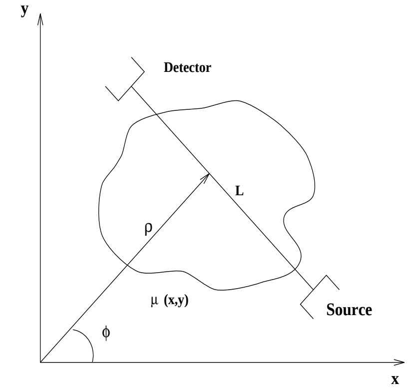


Figure 1: *The geometry of the Radon transform.*

# 3 Optimal Grid Discretization

Suppose there are $M$ angles $\phi_j$ for $j = 1 : M$, such that $\phi_1 = 0 < \phi_2 < \phi_3 \ldots < \phi_M < \pi$, and that at each angle $\phi_j$ there are $n(j)$ strips, or detector bins. Then $N = \sum_{j=1}^{M} n(j)$ gives the total number of data points. Suppose the image space $\Omega$ is some convex, compact region in $\mathbf{R}^2$, and assume that for the $j^{th}$ angle the $n(j)$ strips are parallel, non-overlapping, and entirely cover $\Omega$. Let $\psi_\ell(x, y)$ be the characteristic function of the $\ell^{th}$ strip. Then the discrete strip averaged Radon transform is the map $A : L_2(\Omega) \to \mathbf{R}^N$ defined in equation (1).

Assuming the system $Au = \mathbf{f}$ is consistent, it is underdetermined; that is, since $A$ is a linear mapping from an infinite-dimensional space, $L_2(\Omega)$, to the finite dimensional space $\mathbf{R}^N$, the null space of $A$, $NS(A)$, is infinite dimensional. If there are any solutions to $Au = \mathbf{f}$, there are infinitely many. We must select some representative solution image $u$ from the infinite number of feasible images. The minimum norm solution to the equation $Au = \mathbf{f}$ is given by $u(x, y) = A^*\mathbf{w}$, where $\mathbf{w}$ solves the $N \times N$ system

$$AA^*\mathbf{w} = \mathbf{f}.$$

We write this system as $B\mathbf{w} = \mathbf{f}$ and concentrate on efficient methods to solve it. Note that forming $A^*\mathbf{w}$ corresponds to *backprojecting* the vector $\mathbf{w}$ over the image space.

A simple formula can be used to construct the matrix $B$. Note that $A^* : \mathbf{R}^N \to L_2(\Omega)$ is defined by

$$[A^*\mathbf{w}](x, y) = \sum_{i=1}^{N} w_i \psi_i(x, y). \tag{3}$$

Since the $\psi_i$ are characteristic functions, we observe from equation (3) that the optimal image $u$ is piecewise constant on the set of polygons defined by the intersections of the strips, at all angles. This set of polygons we term the *optimal grid*, as shown in Figure 2.

The $(i, j)^{\text{th}}$ entry of $B$ can be determined by computing the $j^{\text{th}}$ entry of $B\mathbf{e}_i$ where $\mathbf{e}_i$ is the $i^{\text{th}}$ standard basis vector in $\mathbf{R}^N$. Specifically,

$$(B\mathbf{e}_i)_j = (AA^*\mathbf{e}_i)_j = (A\psi_i(x, y))_j = \int_\Omega \psi_i(x, y)\psi_j(x, y)dxdy = \langle \psi_i, \psi_j \rangle.$$

Thus, $B$ is an $N \times N$ matrix with entries

$$b_{ij} = \langle \psi_i, \psi_j \rangle. \tag{4}$$

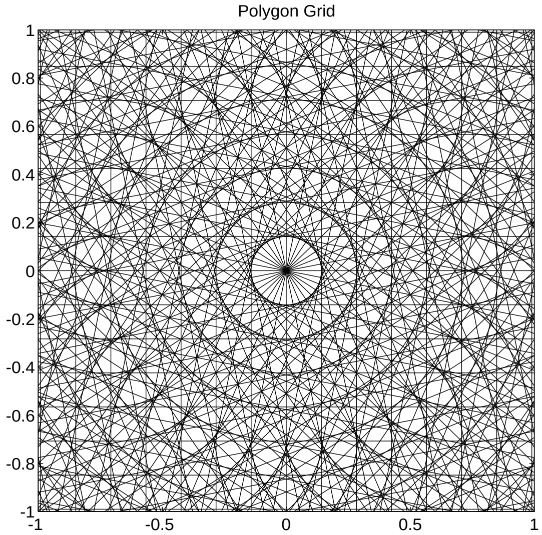


Figure 2: *A representative grid of polygons. The optimal solution is constant on each polygon. This particular optimal grid corresponds to having 16 detector bins of uniform width for each of 20 angles, taken at regular angular intervals between* 0 *and* $\pi$.

We immediately have

**Theorem** 1. *The $N \times N$ matrix $B$ is non-negative, symmetric, positive semidefinite with positive diagonal entries, $b_{ii} > 0$.*

**Proof:** This follows immediately since $\psi_k \geq 0$, $B = AA^*$, $\langle u, AA^*u\rangle = \langle A^*u, A^*u\rangle$, and $b_{ii} = \langle \psi_i, \psi_i \rangle$. ∎

Since we will employ iterative methods, it is important to identify those vectors that may cause difficulty in the iteration process. That is, we seek to characterize the eigenvectors of $\mathbf{R}^N$ that are associated with small nonzero eigenvalues of $B$. Such eigenvectors have the property that $B\mathbf{w}$ is small (in norm) compared to $\mathbf{w}$, and error vectors of this nature have residuals that are small compared to the error. We refer to them as vectors in the *near null space*, and assert that their presence in the error causes slow convergence. The near null space will be studied in section 4.2. For now, we are concerned with characterizing the null space of $B$, which is just the null space of $AA^*$:

$$NS(B) = NS(AA^*) = NS(A^*). \tag{5}$$

For this characterization we need

**Definition:** *A vector* $\mathbf{w} \in \mathbf{R}^N$ *is said to be* constant by angle *if, upon writing*

$$\mathbf{w} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_M \end{bmatrix}, \qquad \mathbf{v}_j \in \mathbf{R}^{n(j)}, \qquad \mathbf{v}_j = \begin{bmatrix} \alpha_1^j, \\ \alpha_2^j, \\ \vdots \\ \alpha_{n(j)}^j \end{bmatrix},$$

*then* $\alpha_i^j = \alpha_{i+1}^j$, $i = 1, \ldots, n(j) - 1$, *for* $j = 1 : M$.

Using this definition, we may show that

**Theorem 2.** *A vector $\mathbf{w} \in \mathbf{R}^N$ is in $NS(B)$ if and only if $\mathbf{w}$ is constant by angle and the elements of $\mathbf{w}$ sum to zero; that is, $sum_{i=1}^N w_i = 0$.*

**Proof:** Let $\{\hat{\psi}_k(x,y) | k = 1 : M\}$ be the set of strips that contain the point $(x,y)$. There is one such strip from each projection angle. Let $\{\hat{w}_k | k = 1 : M\}$ be the set of coefficients associated with those strips. Clearly, if $\mathbf{w}$ is constant by angle and sums to zero, it is in $NS(B)$, since

$$[A^*\mathbf{w}](x,y) = \sum_{i=1}^{N} w_i \psi_i(x,y) = \sum_{k=1}^{M} \hat{w}_k \hat{\psi}_k(x,y) = 0.$$

.

Conversely, suppose $\mathbf{w} \in NS(A^*)$ so that

$$\sum_{i=1}^{N} w_i \psi_i(x,y) = 0 \quad \forall (x,y) \in \Omega.$$

Writing $\mathbf{w}$ as in the definition of constant by angle, that is,

$$\mathbf{w} = [\mathbf{v}_1^T, \ldots, \mathbf{v}_M^T]^T, \quad \mathbf{v}_j \in \mathbf{R}^{n(j)}, \quad \mathbf{v}_j = [\alpha_1^j, \ldots, \alpha_{n(j)}^j]^T,$$

then the objective is to show that for the $j^{\text{th}}$ projection angle, the subvector $\mathbf{v}_j \in \mathbf{R}^{n(j)}$ is constant. Without loss of generality, we show only that $\mathbf{v}_1$ is constant.

Denote the $j^{\text{th}}$ strip at the first angle as $\Omega_j$. Consider the partitioning of the image space $\Omega$ into the set of polygons determined by the intersections of all strips at all angles *except the first.* Clearly, for each $j$, the boundary between $\Omega_j$ and $\Omega_{j+1}$ intersects the interior of at least one of these polygons. Hence, it is possible to select two points, $(x_1, y_1) \in \Omega_j$ and $(x_2, y_2) \in \Omega_{j+1}$, such that the line segment joining $(x_1, y_1)$ and $(x_2, y_2)$ lies entirely in one strip emanating from each of the other angles. That is, the line segment lies entirely in each of the strips $\Omega_{k_2}, \ldots, \Omega_{k_M}$.

By construction,

$$\begin{aligned} 0 = [A^*\mathbf{w}](x_1,y_1) - [A^*\mathbf{w}](x_2,y_2) &= \left(\alpha_j^1 - \alpha_{j+1}^1\right) + \sum_{\ell=2}^{M} w_{k_\ell} \left(\psi_{k_\ell}(x_1,y_1) - \psi_{k_\ell}(x_2,y_2)\right) \\ &= \alpha_j^1 - \alpha_{j+1}^1 \end{aligned}$$

since $\psi_j(x_1,y_1), \psi_{j+1}(x_2,y_2), \psi_{k_\ell}(x_1,y_1)$ and $\psi_{k_\ell}(x_2,y_2)$ all equal 1. It then follows that $\alpha_j^1 = \alpha_{j+1}^1$. That is, $\mathbf{w}$ is constant by angle.

Finally, since $\mathbf{w} \in NS(A^*)$, it must be that the terms sum to zero, because

$$A^*\mathbf{w}(x,y) = \sum_{k=1}^{M} \alpha_1^k = 0.$$

∎

An immediate corollary is the following.

**Corollary** 1. *The dimension of the nullspace of $B$ is $M-1$, where $M$ is the number of projection angles. A basis for $NS(B)$ is*

$$\begin{aligned}
\mathbf{w}_1 &= [\vec{1}, -\vec{1}, \vec{0}, \ldots, \vec{0}] \\
\mathbf{w}_2 &= [\vec{0}, \vec{1}, -\vec{1}, \vec{0}, \ldots, \vec{0}] \\
&\vdots \\
\mathbf{w}_{M-1} &= [\vec{0}, \ldots, \vec{0}, \vec{1}, -\vec{1}]
\end{aligned}$$

*where each constant vector $\vec{1} \in \mathbf{R}^{n(j)}$.*

# 4 Gauss-Seidel Relaxation on $B\mathbf{w} = \mathbf{f}$

Many iterative methods are available for solving equations in $B$, and, indeed, many have been applied to the general image reconstruction problem [1, 9, 18, 19, 20, 24]. Here we consider Gauss-Seidel iteration, one sweep of which may be stated as: *For $j = 1 : N$, modify the $j^{th}$ component of the vector $\mathbf{w}$ such that the $j^{th}$ component of the resulting residual vanishes.* The $j^{\text{th}}$ correction is given by

$$\mathbf{w} \leftarrow \mathbf{w} + \frac{1}{b_{jj}} \langle \mathbf{e}_j, \mathbf{f} - B\mathbf{w} \rangle \mathbf{e}_j.$$

A more common formulation arises from splitting $B$ in terms of its diagonal, upper triangular, and lower triangular parts, giving $B = D - L - U$. Then the $(n+1)^{\text{st}}$ sweep may be written as

$$\mathbf{w}^{(n+1)} = P_G \mathbf{w}^{(n)} + \mathbf{g},$$

where $\mathbf{g} = (D-L)^{-1}\mathbf{f}$ is a fixed vector and $P_G = (D-L)^{-1}U$ is known as the Gauss-Seidel *iteration matrix.* Letting $\mathbf{w}^*$ be any vector that solves $B\mathbf{w} = \mathbf{f}$, we may write an *error vector* defined as $\mathbf{z}^{(n)} = \mathbf{w}^{(n)} - \mathbf{w}^*$. It is easy to see that $\mathbf{z}^{(n+1)} = P_G \mathbf{z}^{(n)}$ and, hence, $\mathbf{z}^{(n+1)} = (P_G)^n \mathbf{z}^{(0)}$. Convergence of the iteration to $\mathbf{w}^*$ is guaranteed if the spectral radius $\rho(P_G)$ is less than one.

The matrix $B$, however, is rank deficient, so that if any solutions exist, then infinitely many solutions exist, and the iteration does not converge under all initial guesses. However, measured in the *energy semi-norm*

$$|||\mathbf{x}||| = \langle B\mathbf{x}, \mathbf{x} \rangle^{1/2},$$

Gauss-Seidel cannot diverge.

**Theorem** 3. *The energy semi-norm of the error does not increase under Gauss-Seidel iteration on $B\mathbf{w} = \mathbf{f}$. That is, $|||\mathbf{z}^{(n+1)}||| \leq |||\mathbf{z}^{(n)}|||$.*

**Proof:** A direct proof follows from the easily derived relation

$$|||\mathbf{z}^{(n+1)}||| = |||\mathbf{z}^{(n)}||| - \frac{\left\langle \mathbf{e}_j, B\mathbf{z}^{(n)} \right\rangle^2}{b_{jj}}. \tag{6}$$

∎

Gauss-Seidel applied to $B\mathbf{w} = \mathbf{f}$ cannot diverge in the energy sense, but to understand when it actually converges we first examine the related Kaczmarz iteration, applied to $Au = \mathbf{f}$.

Proof of the following may be found in [13, 18, 25]:

**Theorem** 4. *Let $L : H_1 \to H_2$ be a continuous linear operator, where $H_1$ is a Hilbert space and $H_2$ is an $N$-dimensional Hilbert space with orthonormal basis $\{v_1, v_2, \ldots, v_N\}$. Let $g \in H_2$ be given, and suppose that $Lu = g$ has a solution. Suppose $u^{(0)} \in range(L^*)$, and define the sequence $u^{(k)}$ generated by the Kaczmarz iteration by*

> *Set* $u \leftarrow u^{(k)}$.
> *For* $j = 1 : N$,
>     *Determine* $s$ *such that* $\langle v_j, L(u + sL^* v_j) - g\rangle = 0$.
>     *Set* $u \leftarrow u + sL^* v_j$.
> *Set* $u^{(k+1)} \leftarrow u$.

*Then $u^{(k)}$ converges, as $k \to \infty$, to the minimum norm solution of $Lu = g$.*

Kaczmarz iteration applied to $Au = \mathbf{f}$ uses $H_1 = L_2(\Omega)$, $H_2 = \mathbf{R}^N$ with orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_N\}$, and

$$s = \frac{\langle \mathbf{e}_j, \mathbf{f} - Au\rangle}{\langle \mathbf{e}_j, AA^* \mathbf{e}_j\rangle}. \tag{7}$$

Proof of the following may be found in [20] or [25].

**Theorem** 5. *Let $\mathbf{w}^{(0)}$ be any vector in $\mathbf{R}^N$, and let $\mathbf{w}^{(k,j)}$, for $j = 1 : N$ and $k = 1, 2, \ldots$, be the vector resulting from the $j^{th}$ step of the $k^{th}$ sweep of Gauss-Seidel iteration on $B\mathbf{w} = \mathbf{f}$, using $\mathbf{w}^{(0)}$ as the initial guess. Then the image $A^* \mathbf{w}^{(k,j)}$ is just the image $u^{(k,j)}$ resulting from the $j^{th}$ step of the $k^{th}$ sweep of the Kaczmarz iteration applied to $Au = \mathbf{f}$ with initial guess $u^{(0)} = A^* \mathbf{w}^{(0)}$.*

An immediate consequence of this theorem is the following.

**Corollary** 2. *Let $\mathbf{w}^{(0)}$ be any vector in $\mathbf{R}^N$, and let $\left\{\mathbf{w}^{(k)}\right\}$ be the sequence of vectors produced by Gauss-Seidel iteration on $B\mathbf{w} = \mathbf{f}$. If $Au = \mathbf{f}$ has a solution, then the sequence $A^* \mathbf{w}^{(k)}$ converges, as $k \to \infty$, to the minimum norm solution of $Au = \mathbf{f}$.*

### 4.1 Numerical Performance

We use the positron emission problem, as in PET and SPECT, for our model problem in developing the iterative methods presented here. Such applications are characterized by relatively small values of $N$ and $M$, so that we are dealing with fairly small computational problems. Typically, the number of bins per angle, $N$, is less than 100, as is the number of angles, $M$. Accordingly, our numerical experiments use $N = 16, 32, 64$, and $M = 10, 20, 64$. Here we report on one such test, which is very representative of the performance characteristics we have observed.

Figure 3 displays an "exact" image, the Shepp-Logan phantom [23], from which a set of values for the right-hand side vector $\mathbf{f}$ is constructed. In generating the vector $\mathbf{f}$ we used $N = 64$ and $M = 20$. Twenty five sweeps of Gauss-Seidel on $B\mathbf{w} = \mathbf{f}$ produce the reconstructed image shown in 3.

Qualitatively, one can argue that the procedure produces a good reconstruction, in that most of the identifiable features of the original image are present. Since, in general, the exact image

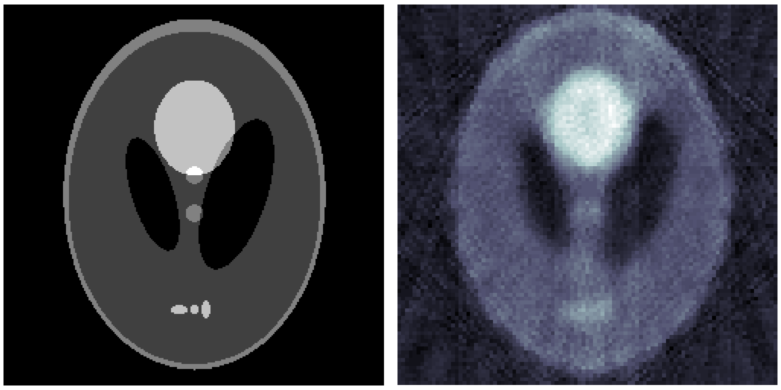

Figure 3: *An "exact" image is shown on the left, and a reconstructed image is shown on the right. The reconstruction geometry uses data collected in 64 bins of uniform width along each of 20 angles. The 20 angles are equispaced at angular intervals of* $\Delta\phi = \pi/20$ *in the interval* $[0, \pi)$*. Twenty-five sweeps of Gauss-Seidel iteration were used to reconstruct the image.*

is unknown, we use the residual $\mathbf{f} - B\mathbf{w}$ as a numerical indication of how well the method solves the problem. Figure 4 displays the logarithm of $||\mathbf{f} - B\mathbf{w}^{(n)}||_2$ as a function of $n$, the number of iteration sweeps. Noteworthy is the fact that the first few sweeps result in significant reduction in the norm of the residual, but that the improvement per sweep declines until (after approximately 10 sweeps) the residual norm remains essentially unchanged.

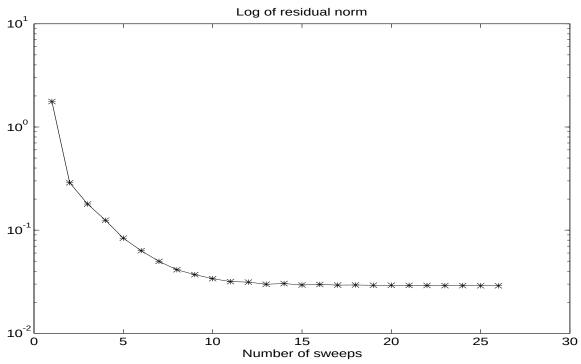


Figure 4: *Performance of the Gauss-Seidel iteration on* $B\mathbf{w} = \mathbf{f}$ *is displayed by plotting the logarithm of* $||\mathbf{f} - B\mathbf{w}^{(n)}||_2$ *as a function of* $n$*, the number of iteration sweeps. Twenty-five sweeps of Gauss-Seidel iteration were used to reconstruct the image.*

This behavior, of rapid improvement in the residual norm over the course of several sweeps followed by stagnation of the residual norm, is characteristic of many iteration methods. In the field of partial differential equations, this numerical "stalling" often occurs because relaxation eliminates the oscillatory components of the error rapidly, but is ineffectual on the remaining smooth components of the error. The stalling phenomenon is often eliminated through the use of multigrid algorithms. Shortly we will develop a multigrid method for the problem $B\mathbf{w} = \mathbf{f}$, in an attempt to address the numerical stalling. Before doing so, however, we wish to make two observations.

First, the stalling phenomenon is often unrelated to the quality of the reconstructed image when the quality is measured by the subjective standard of "looking good". While this measure is hard to quantify, and therefore not so useful to the mathematician or engineer, it is the ultimate measure applied by the end user, for example, the radiologist tasked with treating a patient. It is important to note that the reconstructed images frequently *look* good after only one or two iteration sweeps, while the numerical stalling is not apparent until much later. Figure 5 shows reconstuctions of the "exact" image of Figure 3 as they appear after 1, 2, and 4 sweeps. While subtle differences are apparent in the reconstructions, all are "good". Indeed, it is difficult to differentiate the reconstructions after 4 sweeps and 25 sweeps (Figure 3). For this reason, the residual norm may not be the appropriate indicator of reconstruction quality.

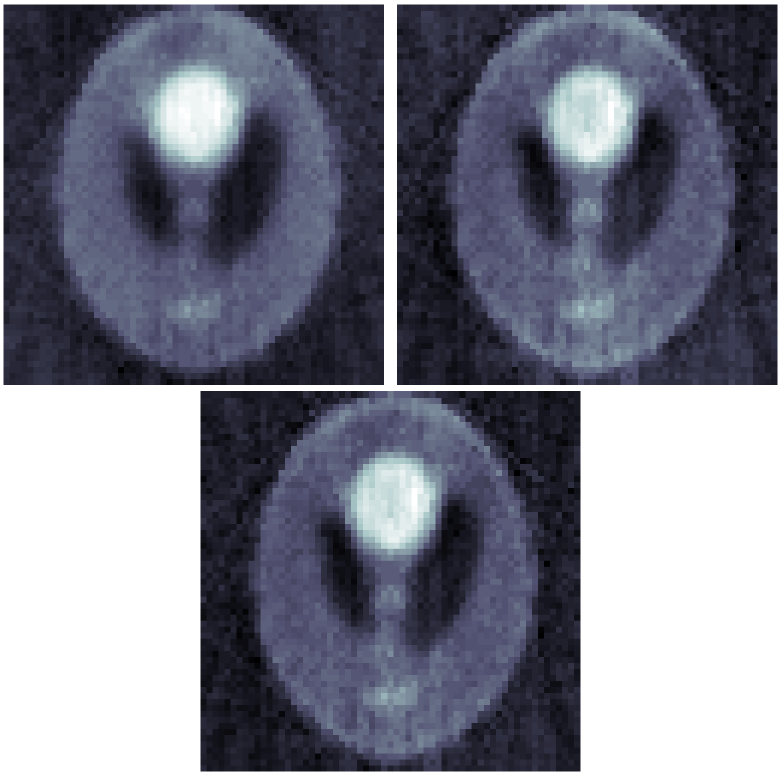

Figure 5: *Reconstructions of the image from Figure 3 using 1, 2, and 4 Gauss-Seidel sweeps are displayed clockwise from the upper left. It is difficult to distinguish these reconstructions, and even more difficult to determine which is "best".*

## 4.2 Mode analysis

The second observation we make is that it is possible to examine the performance of the Gauss-Seidel iteration on individual components of the error. For numerical partial differential equations this is often done by way of Fourier analysis [4]. However, Fourier analysis is not particularly useful in this setting, because the Fourier modes are not eigenfunctions of the continuum operator, nor are discrete Fourier modes eigenvectors of either the matrix $B$ or the iteration matrix $P_G$.

The approach we take is somewhat empirical in nature: we examine the eigenvalues and corresponding eigenvectors of the matrix $B$. Since $B$ is singular with rank $N - M + 1$, we know that zero is an eigenvalue of multiplicity $M - 1$. We are not concerned with eigenvectors corresponding to the zero eigenvalues, as they have no impact on the norm of the residual or on the reconstruction itself (their backprojections vanish).

Slow convergence of the iteration implies that the correction given by the iteration is insufficient. Since the size of the correction to the $j^{\text{th}}$ unknown is determined by the $j^{\text{th}}$ entry in the residual, then slow convergence implies that the residual is "small" compared to the error. Indeed, this has been shown to be the case for many familiar iteration schemes [3]. Since $B$ applied to the error gives the residual, troublesome components are thus errors consisting essentially of the eigenvectors associated with the small nonzero eigenvalues of $B$- the near null space components.

Figure 6 displays a plot of the eigenvalues of a representative $B$ matrix, which is typical of the spectra of all matrices we have examined. The geometry for this case has $M = 20$ angles and $N = \sum_{j=1}^{20} n(j) = 592$ total detectors. The set of eigenvalues is divided into three groups: the zero eigenvalues, the large non-zero eigenvalues, and the group of small eigenvalues $\lambda_k$ whose amplitudes decay rapidly with increasing index $k$. The last 19 eigenvalues, $\lambda_{574}$ through $\lambda_{592}$, are zero, and the associated eigenvectors form $NS(B)$. The vertical dashed line in the figure marks the division between the "good" eigenvalues ($\lambda_1$ through $\lambda_{325}$) and the eigenvalues with rapidly decaying magnitude ($\lambda_{326}$ through $\lambda_{573}$),whose associated eigenvectors form the near null space. These near null space eigenvectors are the "slow" modes that stall performance.

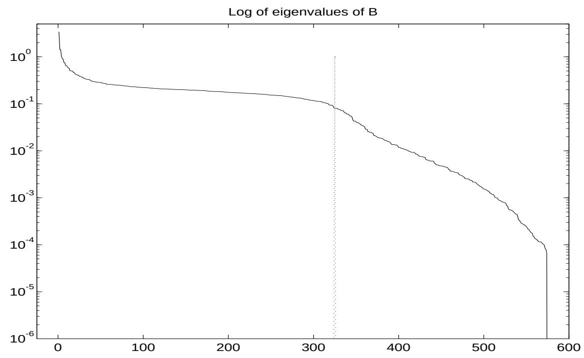


Figure 6: *The logarithms of the eigenvalues of a typical matrix $B$ are shown. $M = 20$ angles are used, so that the null space of $B$ has dimension 15. The eigenvalues between index 1 and index 325 are the "good" modes, while those between 326 and 573 are the* near null space eigenvalues.

We show this empirically in the following way. Representative modes are selected from the "good" and near null space segments of the spectrum. Each such mode is used as the right-hand side of $B\mathbf{w} = \mathbf{f}$. Gauss-Seidel relaxation is then applied, with an initial guess of 0, to solve the equation. Two important observations are obtained in this way. First, by computing the norm of the residual at the end of each sweep, we may determine the convergence factor for each mode. Second, after one sweep of Gauss-Seidel, we compute the projection of the current approximation in the directions of all the eigenvectors of $B$, and plot the resulting magnitudes against the eigenvalue index. This results in a "power spectral density" of the latest iterate. Since the initial error in such an experiment is in the direction of a single eigenvector (the right-hand side), such a spectral density plot tells us to what extent the iteration mixes modes, and if the iteration excites some of the modes, to which part of the spectrum they belong.

Figure 7 shows a typical set of results of these experiments. On the left are residual norms, as a function of iteration sweeps, for two eigenvectors. The dashed line shows the residual norms for the eigenvector corresponding to $\lambda_{15}$, a typical "good" mode, while the solid line gives the residual norms for the eigenvector corresponding to $\lambda_{540}$, a typical near null space mode. In the center is shown the power spectral density plot after one sweep with the 15-mode as the right-hand side, while on the right is the corresponding spectral density plot for the 540-mode.

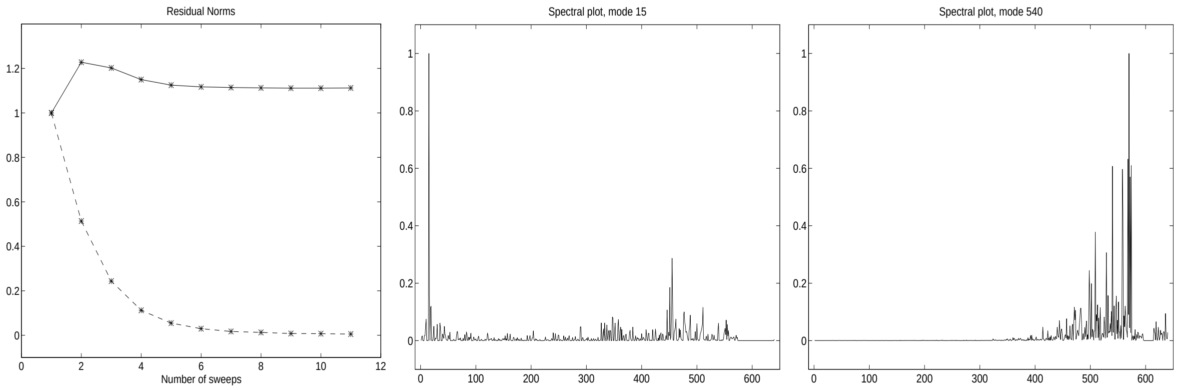


Figure 7: *The residual norms are shown for Gauss-Seidel applied to* $B\mathbf{w} = \mathbf{f}$ *using eigenvectors as the right-hand side. Observe that for mode 15 (dashed line) the iteration converges well, while for mode 540 the iteration stalls immediately. The spectral densities of the vectors resulting from one relaxation sweep is shown for mode 15 (center) and mode 540 (right).*

The results shown in Figure 7 support our assertion regarding the modes. That is, the iteration attenuates "good" error modes rapidly, while the iteration stalls on near null space error modes. In addition, we see that there is mode mixing in the spectral density plots for both cases, but that it is much more pronounced in the case of mode 540. In both cases, though, the exitation of extraneous modes occurs predominantly in the near null space band.

Finally, the images resulting from backprojecting modes 15 and 540 are shown in Figure 8, and again they are typical cases. The backprojected "good" modes generally appear as smooth, highly geometric structures in image space, often as gently undulating surfaces. The backprojected near null space modes, as the name implies, are almost invisible. They often show distinct geometric characteristics, such as narrow subparallel striping, or isolated spikes in nearly flat images.

At this stage we have a good idea what the Gauss-Seidel method achieves. Further, we have a fairly complete picture of where and why it stalls. We next develop multilevel methods for solving $B\mathbf{w} = \mathbf{f}$. We do this for three reasons. First, we believe that if the nature of the near null space modes can be accurately determined, it may be possible to design a "coarse grid correction" to treat the bad modes of the error. That is, we may be able to find a grid on which the bad modes can be anihilated efficiently. Hence, we may hope to achieve multigrid acceleration on this problem. Since the bad modes are not characterized by physical smoothness, like they are in model multigrid problems, standard coarsening is not likely to be very effective. As with other applications that do not possess standard smoothness properties, we must be careful to use what we know about these bad components to devise a coarsening process that closely matches them. Our first attempt

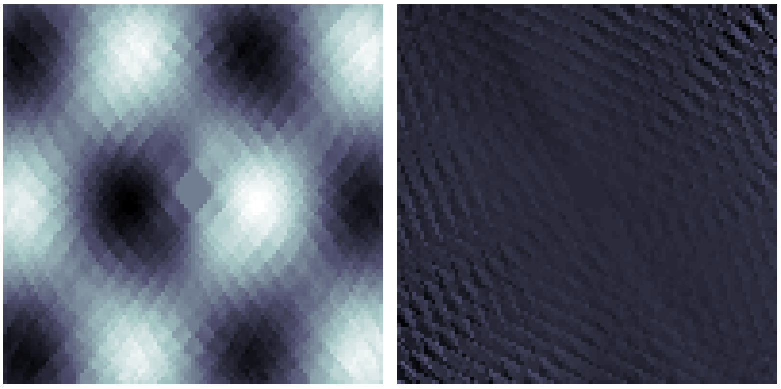

Figure 8: *The images corresponding to backprojecting mode 15 (left) and mode 540 (right) are shown. The "good" mode is a gently undulating surface, while the near null space mode is nearly black.*

is based on the presumption that, at certain scales (e.g., when the number of angles is small compared to the resolution within projections), they must be smooth within projections. We base our multigrid scheme on this idea. A second and equally important reason for using multigrid is our anticipation of the Spotlight CT problem, introduced in the final section of this paper. Finally, the correct isolation of these modes that occurs naturally and efficiently in multilevel processing may pave the way for treating them by an individualized regularization process that is better tailored to the computational objectives.

# 5 Multilevel Image Reconstruction

## 5.1 Multilevel Projection Methods (PML)

Designing a multigrid method for a new problem is a difficult task, especially when the application is far removed from the classical multigrid setting of elliptic PDEs. Multigrid has been extended to a wide variety of such problems, with varying degrees of success [6, 17, 22]; multigrid design in such instances is generally a lengthy and difficult process.

The multilevel projection methodology (PML) was developed to provide a simpler, systematic approach to multilevel algorithm design [15]. A basic tenet of PML design is that only the appropriate subspaces in which the problem is to be set need to be determined. The problem is discretized by orthogonal projections, and the projection operators in turn lead to the correct choices for intergrid transfer operators, relaxation techniques, and coarsening schemes.

To briefly describe the fundamentals of PML, let $H_1$ and $H_2$ be Hilbert spaces and $L : H_1 \to H_2$ be a linear operator. The continuum problem is to find $u \in H_1$ such that $Lu - f = 0$. *Discretization by projections* is accomplished as follows.

Let $S^h$ be a finite-dimensional subspace of $H_1$, and let $P^{S^h} : H_1 \to S^h$ be an orthogonal projection of $H_1$ onto $S^h$, where the superscript $h$ refers to a discretization parameter. We also require a finite-dimensional subspace $T^h \subset H_2$, and an orthogonal projection $P^{T^h} : H_2 \to T^h$, as well as mappings $P_{S^h} : S^h \to H_1$ and $P_{T^h} : T^h \to H_2$.

The projection operators are used to generate a discrete operator $L^h : S^h \to T^h$ by projecting the action of the continuum operator $L$ onto the subspaces, that is, the discretized problem becomes $P^{T^h} L(P^{S^h} u) = 0$, for $u \in H_1$. This allows us to define the discrete operator for the problem by

$$L^h u^h - f^h \;=\; 0, \;\text{ for }\; u^h \in S^h, \qquad \text{where} \qquad L^h \equiv P^{T^h} L P^{S^h} \;\text{ and }\; f^h = P^{T^h} f.$$

We pause here to show that the strip pixel discretization developed earlier is in fact a discretization by projection.

**Theorem 6.** *For each $j = 1 : M$, let $\Omega$ be exactly partitioned into $n(j)$ parallel non-overlapping strips and let $N = \sum_{j=1}^{M} n(j)$. Number the strips from 1:N and let $\psi_j(x, y)$ be the characteristic function of the $j^{th}$ strip. Let $S^h$ be the subspace of the Hilbert space $H_1 = L_2(\Omega)$ spanned by the set*

$$\{\psi_j\}_{j=1}^{N}.$$

*Then the matrix equation*

$$B\mathbf{w} \;=\; \mathbf{f}$$

*is a discretization by projections of the problem $Au = \mathbf{f}$, where $A$ is the strip averaged Radon transform (1) and $B$ is the $N \times N$ matrix with entries $b_{jk} = \langle \psi_j, \psi_k \rangle$.*

**Proof:** We define the various subspaces of the discretization as follows. $H_1$ and $S^h$ are defined in the statement of the theorem. Note that (3) implies that $S^h = \text{range}\,\{A^*\}$. We take $H_2 = \mathbf{R}^N$ and define the subspace $T^h$ to be $H_2$. Since $H_2 = T^h = \mathbf{R}^N$, we may take $P^{T^h} = I_N$, the $N \times N$ identity matrix. The discrete equation will then be $AP^{S^h} u \;=\; \mathbf{f}$, where $P^{S^h} u$ is the orthogonal projection of $u(x, y)$ onto $S^h$, so

$$P^{S^h} u \;=\; \sum_{k=1}^{N} w_k \psi_k(x, y) \;=\; A^{h^*} \mathbf{w}$$

for some $\mathbf{w} \in \mathbf{R}^N$. Since $P^{S^h}$ is an orthogonal projection, we must have $(u - P^{S^h} u) \perp \psi_j$ for every $\psi_j \in S^h$. Hence, for $j = 1 : N$,

$$\begin{aligned} 0 \;&=\; \left\langle u - P^{S^h} u,\, \psi_j \right\rangle \\ &=\; \left\langle u - \sum_{k=1}^{N} w_k \psi_k,\, \psi_j \right\rangle \\ &=\; \langle u,\, \psi_j \rangle \;-\; \sum_{k=1}^{N} w_k \,\langle \psi_k,\, \psi_j \rangle \,. \end{aligned}$$

Thus, if $P^{S^h} u \;=\; A^{h^*} \mathbf{w}$ is an orthogonal projection of $u(x, y)$ into $S^h$, then the vector $\mathbf{w}$ must satisfy

$$[\langle u, \psi_1 \rangle \;\; \langle u, \psi_2 \rangle, \;\; \ldots \;\; \langle u, \psi_N \rangle]^T \;=\; B\mathbf{w}. \tag{8}$$

But the left-hand side of (8) is just $Au$, so $\mathbf{w}$ must solve $Au \;= B\mathbf{w}$, and the projection-discretized form of $Au \;=\; \mathbf{f}$ is just $B\mathbf{w} = \mathbf{f}$. ∎

Henceforth, to keep track of the level we are examining, we use the notation $B^h\mathbf{w}^h = \mathbf{f}^h$, where $B^h$ is the matrix defined by (4). We also adopt superscripts for use with the characteristic functions of the strips, e.g., $\psi_k^h$.

An important observation to be made here is that it is not necessary to know the projection operators explicitly if the *condition* of orthogonal projection adequately defines the discrete operator $L^h$.

Now we can examine how the PML method makes use of discretization by projections to build a two-level solver. Let $P^{S^{2h}}$ and $P^{T^{2h}}$ be projection operators mapping the continuum spaces $H_1$ and $H_2$ into "coarse grid" subspaces $S^{2h} \subset S^h \subset H_1$ and $T^{2h} \subset T^h \subset H_2$. The coarse grid operator is given by $L^{2h} = P^{T^{2h}} L P^{S^{2h}}$.

The two main components of any multigrid problem are relaxation and coarse grid correction. The PML approach defines relaxation by decomposing the spaces $S^h$ and $T^h$ into sums (which need not be direct sums) of $m$ subspaces

$$S^h = \sum_{\ell=1}^{m} S_\ell^h \qquad \text{and} \qquad T^h = \sum_{\ell=1}^{m} T_\ell^h.$$

Any element of $S^h$ can be written as a linear combination (not necessarily unique!) of the elements of $S_\ell^h$:

$$u^h = \sum_{\ell=1}^{m} \alpha_l u_{(\ell)}^h, \quad \text{where } u_{(\ell)}^h \in S_\ell^h.$$

The $\ell^{th}$ relaxation step is defined by adding to the current approximation $u^h$ an element of the subspace $S_\ell^h$ such that the projection of the residual into $T_\ell^h$ vanishes. A relaxation sweep is made by performing the relaxation step for all $m$ subspaces. Hence, *relaxation by projections* is defined by

Relaxation: $\quad u^h \leftarrow G^h(u^h)$

For $\ell = 1,2, \ldots,$ m

1. Determine $u_{(\ell)}^h \in S_\ell^h$ such that $P^{T_\ell^h} L^h(u^h + u_{(\ell)}^h) - f^h = 0$.
2. Set $u^h \leftarrow u^h + u_{(\ell)}^h$.

In standard multigrid, coarse grid correction is performed by restricting the residual equation to the coarse grid, solving for the error, interpolating the error to the fine grid, and adding it to the fine grid approximation. This basic process is also what PML does, though in an abstract way that is guided by the discretization. Given the coarse level subspaces $S^{2h} \subset S^h \subset H_1$ and $T^{2h} \subset T^h \subset H_2$, together with the associated projection operators, the aim is to determine an element of the coarse space $S^{2h}$ that, when added to the current approximation $u^h$, satisfies the projection of the residual equation onto $T^{2h}$. Thus, *coarse grid correction by projections* is written as

Coarse grid correction: $\quad u^h \leftarrow C^h(u^h)$

1. Determine $u^{2h} \in S^{2h}$ such that $P^{T^{2h}}\left(L(P^{S^h}u^h + P^{S^{2h}}u^{2h}) - f^h\right) = 0$.
2. Set $u^h \leftarrow u^h + u^{2h}$.

Relaxation and coarse grid correction together form a two-level PML method that is given by

Two Level PML method: $\quad u^h \leftarrow PML^h(u^h)$

- $u^h \leftarrow G^h(u^h)$.
- $u^h \leftarrow C^h(u^h)$.

The two-level PML algorithm can be converted into a multilevel scheme in just the same way that standard multigrid schemes are developed from two-grid schemes: the exact solver in the coarse grid correction is replaced by a recursive call $u^{2h} \leftarrow PML^{2h}(u^{2h})$, leading to a PML V-cycle, for example.

# 6 PML Image Reconstruction

## 6.1 Discretization and intergrid transfers

Applying PML to equations in $A$ is somewhat subtle [15], primarily because of the need to treat both projection (Radon transform) and image spaces, but with the optimal pixel discretization, applying PML to equations in $B = AA^*$ may be more direct. We have already shown that $B^h\mathbf{w}^h = \mathbf{f}^h$ is a discretization by projections of the problem $Au = \mathbf{f}$ onto $S^h$ and $T^h$. It is easy to define coarse subspaces $S^{2h}$ and $T^{2h}$ in a manner that leads to a useful multilevel algorithm.

Let $S^h$ be the span of the $N$ strip pixels $\psi_j^h$, where the $h$ is some parameter that indicates the level of the discretization (e.g., $h$ may be the width of the widest strip pixel). Suppose for simplicity that there is an even number of strip pixels for each of the $M$ views, and that we number the strips from $\psi_1^h$ to $\psi_N^h$ in a way so that two adjacent strips on any view are always numbered consecutively. Then a useful subspace $S^{2h}$ can be constructed according to

$$S^{2h} = \operatorname{span}\left\{\psi_k^{2h}\right\}_{k=1}^{N/2} \qquad \text{where} \qquad \psi_k^{2h} = \psi_{2k-1}^h + \psi_{2k}^h. \tag{9}$$

Thus, each strip pixel in the coarse subspace is the union of two adjacent fine space strip pixels. This may be viewed in physical terms as widening the aperature of the detectors, or bins.

With these coarse space strip pixels, we find that $A^{2h^*}\mathbf{w}^{2h} = \sum_{j=1}^{N/2} w_j^{2h}\psi_j^{2h}$, from which we easily obtain $A^{2h} : S^{2h} \to \mathbf{R}^{N/2}$ by $\left(A^{2h}u\right)_i = \langle\psi_i^{2h}, u\rangle$. This in turn leads us to the projection discretized coarse level problem $B^{2h}\mathbf{w}^{2h} = \mathbf{f}^{2h}$, where $B^{2h} = A^{2h}A^{2h^*}$ is an $N/2 \times N/2$ matrix with entries $b_{ij}^{2h} = \left\langle\psi_i^{2h}, \psi_j^{2h}\right\rangle$.

The multilevel scheme requires interlevel transfer operators to map grid functions between the coarse and fine levels, and a basic tenet of PML is that these operators are defined implicitly. That is, $P_{S^h}^{S^{2h}} : S^h \to S^{2h}$ and $P_{S^{2h}}^{S^h} : S^{2h} \to S^h$ are defined by

$$P_{S^{2h}} \;=\; P_{S^h} P_{S^{2h}}^{S^h}, \qquad \text{and} \qquad P^{S^{2h}} \;=\; P_{S^h}^{S^{2h}} P^{S^h}.$$

Analogous transfers can be defined for the subspaces $T^{2h} \subset T^h \subset H_2$.

To determine the intergrid transfer operators $P_{S^h}^{S^{2h}}$ and $P_{S^{2h}}^{S^h}$, we begin with a simple observation based on the definition of the coarse space strips $\psi_k^{2h}$.

**Lemma** 1. *The operators $A^h : S^h \to T^h$ and $A^{2h} : S^{2h} \to T^{2h}$ are related by*

$$A^{2h} \;=\; P_{S^h}^{S^{2h}} A^h$$

*where $P_{S^h}^{S^{2h}}$ is an $N/2 \times N$ matrix given by*

$$P_{S^h}^{S^{2h}} \;=\; \begin{pmatrix} 1 & 1 & & & & & & \\ & & 1 & 1 & & & & \\ & & & & \ddots & \ddots & & \\ & & & & & & 1 & 1 \end{pmatrix}.$$

*Furthermore, the adjoint operators $A^{h^*} : T^h \to S^h$ and $A^{2h^*} : T^{2h} \to S^{2h}$ are related by*

$$A^{2h^*} \;=\; A^{h^*} \left(P_{S^h}^{S^{2h}}\right)^T.$$

**Proof:** Let the coarse grid strip pixel $\psi_k^{2h}$ be the union of adjacent fine grid strip pixels given by $\psi_k^{2h} = \psi_{2k-1}^h + \psi_{2k}^h$, for $k = 1 : N/2$. Then

$$\left\langle \psi_k^{2h}, u \right\rangle \;=\; \left\langle \psi_{2k-1}^h + \psi_{2k}^h, u \right\rangle \;=\; \left\langle \psi_{2k-1}^h, u \right\rangle + \left\langle \psi_{2k}^h, u \right\rangle \;=\; (1 \quad 1) \begin{pmatrix} \left\langle \psi_{2k-1}^h, u \right\rangle \\ \left\langle \psi_{2k}^h, u \right\rangle \end{pmatrix}.$$

The first assertion of the lemma follows by partitioning the vector $A^h u$ into blocks consisting of pairs of adjacent entries and forming the matrix $P_{S^h}^{S^{2h}}$ by placing, for $k = 1 : N/2$, the block $(1 \quad 1)$ in the $(2k-1)^{st}$ and $(2k)^{th}$ positions of the $k^{th}$ row of an $N/2 \times N$ zero matrix. Matrix vector multiplication then yields $A^h u$.

The second assertion is established by

$$\begin{aligned} A^{2h^*} \mathbf{w}^{2h} &= \left(\psi_1^{2h} \quad \psi_2^{2h} \quad \dots \quad \psi_{N/2}^{2h}\right) \mathbf{w}^{2h} \\ &= \left(\psi_1^h + \psi_2^h \quad \psi_3^h + \psi_4^h \quad \dots \quad \psi_{N/2-1}^h + \psi_{N/2}^h\right) \mathbf{w}^{2h} \\ &= \left[A^{h^*} \left(P_{S^h}^{S^{2h}}\right)^T\right] \mathbf{w}^{2h}. \qquad \blacksquare \end{aligned}$$

The second part of the lemma verifies that the operator $P_{S^h}^{S^{2h}}$ gives a consistent definition to the adjoint of the coarse space operator $A^{2h}$, showing that $A^{2h^*} = (P_{S^h}^{S^{2h}} A^h)^* = A^{h^*} \left(P_{S^h}^{S^{2h}}\right)^T$. Thus we may define $P_{S^{2h}}^{S^h} = \left(P_{S^h}^{S^{2h}}\right)^T$. Combining this with the observation that $B^{2h} = A^{2h} A^{2h^*} = P_{S^h}^{S^{2h}} A^h A^{h^*} P_{S^{2h}}^{S^h} = P_{S^h}^{S^{2h}} B^h P_{S^{2h}}^{S^h}$, we find that the standard variational conditions of multigrid are satisfied by this discretization [5].

## 6.2 Relaxation

Following the principles of PML, we select sets of $m$ subspaces $S_\ell^h$ and $T_\ell^h$ whose unions equal $S^h$ and $T^h$, respectively. The most obvious choice is to select $m = N$ and $S_\ell^h = \text{span}\left\{\psi_\ell^h\right\}$, that is, each subspace is the span of an individual strip pixel.

> **Lemma** 2. *Let $S_\ell^h = span\left\{\psi_\ell^h\right\}$ and $T_\ell^h = span\left\{\mathbf{e}_\ell^h\right\}$ for $\ell = 1 : N$, where $\mathbf{e}_\ell^h$ is the $\ell^{th}$ standard basis vector in $T^h$. Then the PML relaxation on $A^h u^h = \mathbf{f}^h$ is implemented by performing point Gauss-Seidel iteration on the matrix equation $B^h \mathbf{w}^h = \mathbf{f}^h$.*

**Proof:** The $\ell^{th}$ step of PML relaxation consists of finding that value of $\alpha$ satisfying

$$P_\ell^{T^h}\left(AP^{S^h}(u^h + \alpha\psi_\ell^h) - \mathbf{f}^h\right) = 0,$$

where $u^h$ is the current approximation to the solution, and then modifying the approximation by $u^h \leftarrow u^h + \alpha\psi_\ell^h$. Now $\psi_\ell^h = A^{h^*}\mathbf{e}_\ell^h$, and since $P^{S^h}u^h \in S^h$, then we know that $P^{S^h}u^h = A^{h^*}\mathbf{w}^h$ for some $\mathbf{w}^h \in T^h$. Hence we seek $\alpha$ such that

$$\begin{aligned} 0 &= P_\ell^{T^h}\left(A^h(A^{h^*}\mathbf{w}^h + \alpha A^{h^*}\mathbf{e}_\ell^h) - \mathbf{f}^h\right) \\ &= P_\ell^{T^h} B^h(\mathbf{w}^h + \alpha\mathbf{e}_\ell^h) - f_\ell^h. \end{aligned}$$

Noting that $P_\ell^{T^h}$ is accomplished by forming the inner product with $\mathbf{e}_\ell^h$, then $\alpha$ must satisfy

$$(\mathbf{e}_l^h)^T(B^h\mathbf{w}^h + \alpha B^h\mathbf{e}_\ell^h) = f_\ell,$$

whose solution is given by

$$\alpha = \frac{1}{b_{\ell\ell}}(f_\ell^h - \mathbf{b}_\ell^T\mathbf{w}^h),$$

where $\mathbf{b}_\ell^T$ is the $\ell^{th}$ row of $B^h$. Hence, the $\ell^{th}$ step of the PML relaxation is

$$\mathbf{w}_\ell^h \leftarrow \mathbf{w}_\ell^h + \frac{1}{b_{\ell\ell}}(f_\ell^h - \mathbf{b}_\ell^T\mathbf{w}^h)$$

which is precisely the correction of the $\ell^{th}$ step of Gauss-Seidel applied to $B^h\mathbf{w}^h = \mathbf{f}^h$. ∎

## 6.3 Coarse grid correction

Like relaxation, coarse grid correction in the PML approach is defined by the selection of the subspaces and the implicit intergrid transfer operators. For the problem $Au = \mathbf{f}$, it is performed by finding the element $u^{2h} \in S^{2h}$ that satisfies

$$P^{T^{2h}}\left(A(P^{S^h}u^h + P^{S^{2h}}u^{2h}) - \mathbf{f}^h\right) = 0, \tag{10}$$

where $u^h$ is the current approximation in the fine space $S^h$. The correction is then given by $u^h \leftarrow u^h + u^{2h}$. Note that

$$P^{T^{2h}}AP^{S^h}u^h = P_{S^h}^{S^{2h}}P^{T^h}AP^{S^h}u^h = P_{S^h}^{S^{2h}}B^h\mathbf{w}^h$$

where $A^{h*}\mathbf{w}^h$ represents $P^{S^h}u^h$. We also know that since $u^{2h} \in S^{2h}$, there is a vector $\mathbf{y}^{2h} \in T^{2h}$ such that $u^{2h} = A^{2h*}\mathbf{y}^{2h}$. Hence, $P^{T^{2h}}AP^{S^{2h}}u^{2h} = B^{2h}\mathbf{y}^{2h}$. Noting also that $P^{T^{2h}}\mathbf{f} = P_{S^h}^{S^{2h}}\mathbf{f}$, then (10) becomes

$$P_{S^h}^{S^{2h}}B^h\mathbf{w}^h + B^{2h}\mathbf{y}^{2h} - P_{S^h}^{S^{2h}}\mathbf{f}^h = 0.$$

The correction step is thus $\mathbf{w}^h \leftarrow \mathbf{w}^h + P_{S^{2h}}^{S^h}\mathbf{y}^{2h}$. Hence, with the operators we have constructed, the PML coarsening step for this problem is formally the same as conventional multigrid:

1. Set $\mathbf{f}^{2h} = P_{S^h}^{S^{2h}}(\mathbf{f}^h - B^h\mathbf{w}^h)$.
2. Solve $B^{2h}\mathbf{y}^{2h} = \mathbf{f}^{2h}$.
3. Correct the approximation by $\mathbf{w}^h \leftarrow \mathbf{w}^h + P_{S^{2h}}^{S^h}\mathbf{y}^h$.

Of course, as with any multigrid algorithm, in practice the exact solve on the coarse grid is replaced by a recursion, so that the only time an exact solution is computed is on the coarsest subspace. To form such a recursion in the strip pixel PML setting, we need only continue defining coarser spaces $S^{jh}$, for $j = 1, 2, \ldots$. This is done by taking the strip pixels that generate the new subspace to be the pairwise union of strip pixels in the current subspace, just as was done to produce $S^{2h}$ from $S^h$. Once this is done, a PML V-cycle can be defined in the usual way.

PML V-cycle: $\quad \mathbf{w}^h \leftarrow PMLV(\mathbf{w}^h, B^h, \mathbf{f}^h)$

1. Relax $\nu_1$ times on $B^h\mathbf{x}^h = f^h$ with initial guess $\mathbf{w}^h$.
2. If $S^h$ represents the coarsest level, go to 3. Otherwise:
   (a) $\mathbf{f}^{2h} \leftarrow P_{S^h}^{S^{2h}}(f^h - B^h\mathbf{w}^h)$
   (b) $\mathbf{w}^{2h} \leftarrow 0$.
   (c) $\mathbf{w}^{2h} \leftarrow PMLV(\mathbf{w}^{2h}, B^{2h}, \mathbf{f}^{2h})$
   (d) $\mathbf{w}^h \leftarrow \mathbf{w}^h + P_{S^{2h}}^{S^h}\mathbf{w}^{2h}$.
3. Relax $\nu_2$ times on $B^h\mathbf{x}^h = f^h$ with initial guess $\mathbf{w}^h$.

### 6.4 Numerical Performance

Figures 9 and 10 display two examples of the image reconstructions obtained with the PMLV algorithm. The pair of images in Figure 9 were obtained using 20 views with 32 detectors per view, and restricting the image to lie in the unit square. The data were generated by projecting the exact image on the left, while the reconstruction of the image by PMLV is shown on the right. The reconstruction was made using 3 PMLV cycles with 2 relaxation sweeps on the downward leg of the V and one relaxation sweep on the upward leg. The Shepp-Logan phantom was used for the reconstruction in Figure 10, which was obtained from 64 views with 64 detectors per view. Again,

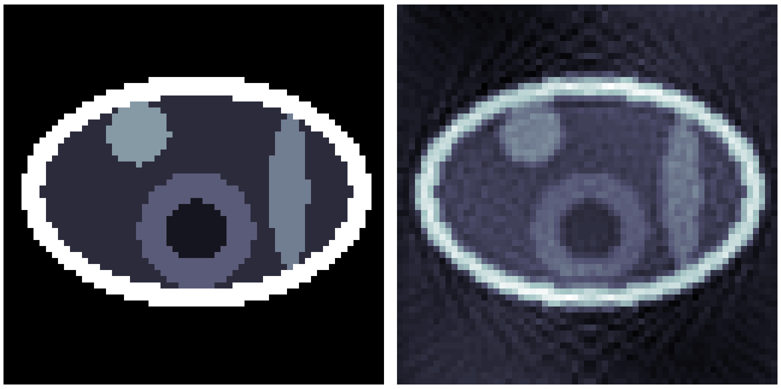

Figure 9: *The performance of the PMLV method may be observed by comparing the actual (left) and reconstructed images (right) of a "brain phantom". The reconstruction was obtained using 3 PMLV cycles with 2 relaxation sweeps on the downward leg of the V and one relaxation sweep on the upward leg. The image on the left was used to generate the data used in the reconstruction, with 20 angles and 32 detectors per angle.*

the image on the left was used to generate data for the multilevel reconstruction (right), which was made from 3 PMLV (2,1)-cycles.

To compare the performance of PMLV and Gauss-Seidel, let the work entailed by one sweep of Gauss-Seidel on the finest level be one work unit (WU), which is an $O(N^2)$ operation. Computation of the residual requires approximately one WU. The work required for one PMLV-cycle, using $\nu_1+\nu_2$ sweeps of Gauss-Seidel on each level, is then be bounded by

$$(\nu_1+\nu_2+1)(1+\frac{1}{4}+\frac{1}{16}+\cdots) \approx \frac{4}{3}(\nu_1+\nu_2+1)\,.$$

PMLV was applied to several reconstruction problems, using several different geometries. The performance of the algorithm in all tests was similar, and may be summarized by examining the results of a typical suite of experiments. In these tests the parameters $\nu_1 = 2$ and $\nu_2 = 1$ gives the number of iteration sweeps, respectively, descending and ascending through the V-cycle. Hence, one V-cycle requires approximately $\frac{16}{3}$ WU. The problem was coarsened to the coarsest possible level, giving one strip per view and a problem of size $M \times M$ at the coarsest level. Figure 11 compares the typical performance of Gauss-Seidel to the PMLV algorithm for a problem with 32 detectors over 20 angles.

It is clear from Figure 11 that, even for this relatively small problem, PMLV initially outperforms Gauss-Seidel. However, continued iteration of Gauss-Seidel eventually achieves similar results at similar costs. We believe this is due largley to the fact that the bad modes do not possess the physical smoothness characteristic of bad modes in elliptic PDE problems, so that coarsening by row-lumping within projections does not entirely succeed at eliminating the bad modes. We think that this may be caused by the problem entering a scale regime where there is close coupling between the projections. This is likely to mean that a special lumping of rows is needed, where the oscillatory but possibly regular pattern of these components across angles is taken into account. Note that the slopes at the right end of the curves indicate that further iteration may favor PMLV. However, it is important to recall that the ultimate goal is quality image reconstruction, and that

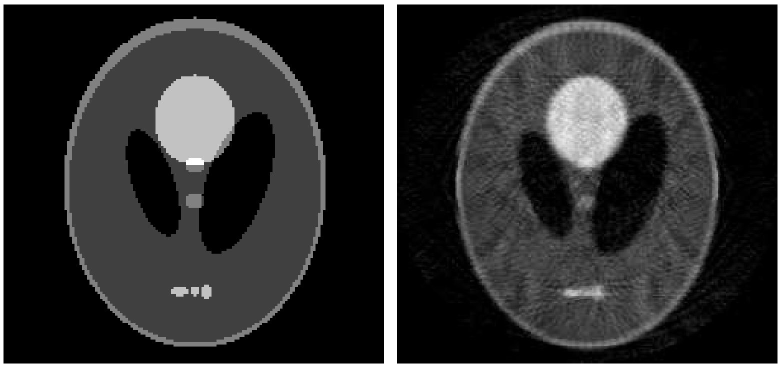

Figure 10: *The reconstruction on the right was obtained using the multilevel method from data generated from the image on the left, using 64 angles and 64 detectors per angle. The reconstruction was obtained using 3 PMLV cycles with 2 relaxation sweeps on the downward leg of the V and one relaxation sweep on the upward leg.*

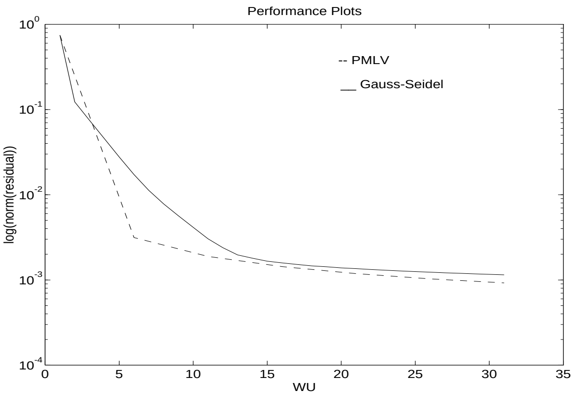


Figure 11: *Comparison of the performance of Gauss-Seidel and PMLV. The logarithm of the residual norm is plotted as a function of the number of work units required to attain it.*

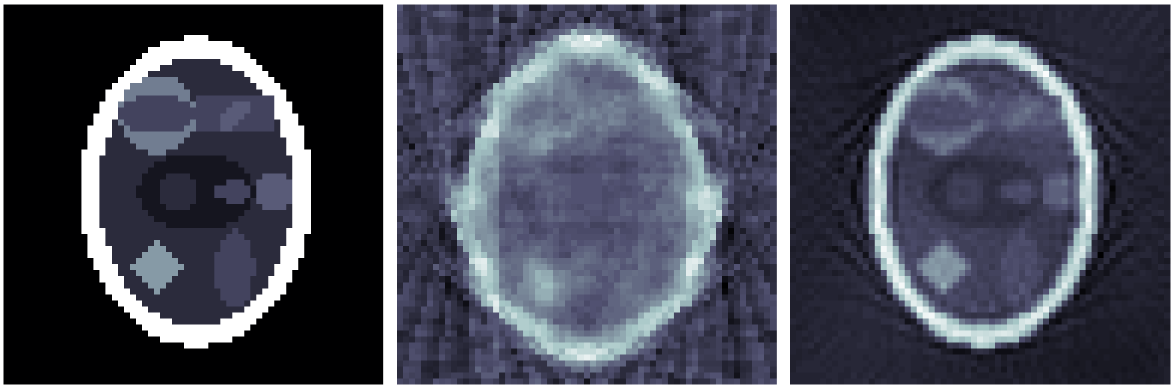

Figure 12: *Exact image (left), minimum norm solution (center), and PMLV solution (right)*

the residual norm may not be a reliable measure of success. It is also important to recall the ill-posed nature of reconstruction. Consider the problem $B\mathbf{w} = \mathbf{f}$, whose exact solution (in the least squares sense) is

$$\mathbf{w} = B^{\dagger}\mathbf{f},$$

where $B^{\dagger}$ is the psuedo-inverse of $B$ . In terms of the singular value decomposition $B = U\Sigma V^T$, the solution $\mathbf{w}$ can be expressed as

$$\mathbf{w} = \left[\sum_{i=1}^{r} \frac{1}{\sigma_i}\mathbf{v}_i\mathbf{u}_i^T\right]\mathbf{f} = \sum_{i=1}^{r} \frac{1}{\sigma_i}\mathbf{v}_i < \mathbf{u}_i, \mathbf{f} >,$$

where the $\sigma_i$'s are the singular values, $\mathbf{u}_i$ and $\mathbf{v}_i$ (the left and right singular vectors of $B$) are the columns of $U$ and $V$, and $r = \text{rank}(B)$. If there is *measurement noise* in the data, so that instead of $\mathbf{f}$ the data are $\mathbf{f}+\varepsilon$, then solution components corresponding to small singular values will magnify this noise. These are the components in the near null space that are slow to be recovered. Thus, continued iteration after the procedure stalls in an attempt to recover these slow components has the potential to corrupt the solution with magnified noise [11, 19].

Experiments have shown, in fact, that it is possible to drive the residual norm to zero, finding one of the solutions to the linear system, and have reconstructed images that are of poor subjective quality, perhaps worse than that of early iterates. Figure 12, for example, displays an exact image and two reconstructions of the data for that image. One reconstruction is made by computing $\mathbf{w} = B^{\dagger}\mathbf{f}$, while the other is made by running 3 PMLV cycles. The residual norms are $1.8 \times 10^{-13}$ (pseudoinverse) and $4 \times 10^{-4}$ (PMLV), but it can be seen that PMLV has produced a somewhat better reconstruction.

Such problems require some form of regularization to prevent the ill-posedness from completely corrupting the approximation. One way to do this [18] is to stop iterating when the algorithm begins to stall. An ad hoc approach to this is to measure the difference between successive residual norms, and stop iterating when a tolerance is achieved. A potentially more effective stopping criteria exists [20], based on a newly developed convergence theory for multilevel algorithms [7], but this is beyond the scope of interest here.

# 7 Spotlight CT Image Reconstruction

Often, one desires high resolution in a certain region of the image, for example, where a tumor is suspected. Discretizing the entire image space at a fine resolution may be impractical, as this leads to extremely large systems of equations. An attractive alternative is to discretize the region of special interest at a fine resolution and the remaining image space at a coarser resolution, leading to a *composite grid* problem. This is called *spotlighting* the region of interest. Numerous multigrid methods have been developed for handling composite grid problems [2, 16, 21]. One such method that we develop in the next section is a consequence of PML methodology.

## 7.1 Fast Adaptive Composite Grid (FAC)

The spotlight CT problem is essentially a composite grid problem, in which an operator equation $Lu = f$ must be solved on a composite grid $\Omega^{\underline{h}}$ comprised of a global coarse grid $\Omega^{2h}$ and one or more local refinement grid $\Omega^h$ (the refinement grid may itself be a composite grid, which permits recursive refinement). Fast adaptive composite grid methods (FAC) were developed [14, 16] in order to utilize multigrid technology to treat such problems efficiently. It comes from the PML methodolgy by simply restricting the fine grid subspaces to local collections of detectors.

FAC succeeds because it handles the composite grid as a nested sequence of regular grids that can be treated independently using virtually any regular-grid method. The key ingredients lie in having appropriate representations of the operator and intergrid transfer operators. Thus, grid functions $u$ must be representable on the composite, global, and refinement grids $(u^{\underline{h}}, u^{2h}, u^h)$, and operators must exist to transfer grid functions between these grids ( $I^h_{\underline{h}} : \Omega^{\underline{h}} \to \Omega^h$, $I^{2h}_{\underline{h}} : \Omega^{\underline{h}} \to \Omega^{2h}$, $I^{\underline{h}}_h : \Omega^h \to \Omega^{\underline{h}}$, and $I^{\underline{h}}_{2h} : \Omega^{2h} \to \Omega^{\underline{h}}$). Finally, it is critical that the operator $L$ be representable on these grids $(L^{\underline{h}}, L^{2h}, L^h)$. In general, the details of these operators and representations can be developed in a straightforward fashion once the grids are defined; the details, however, are very technical and can be found in [12, 14, 16]. Once the operators and grids are defined, FAC proceeds in the following two-step sequence:

**Step 1:** *Set* $\quad f^{2h} \leftarrow I^{2h}_{\underline{h}}(f^{\underline{h}} - L^{\underline{h}}u^{\underline{h}}), \quad u^{2h} \leftarrow (L^{2h})^{-1} f^{2h}, \quad u^{\underline{h}} \leftarrow u^{\underline{h}} + I^{\underline{h}}_{2h} u^{2h}$

**Step 2:** *Set* $\quad f^{h} \leftarrow I^{h}_{\underline{h}}(f^{\underline{h}} - L^{\underline{h}}u^{\underline{h}}), \quad u^{h} \leftarrow (L^{h})^{-1} f^{h}, \quad u^{\underline{h}} \leftarrow u^{\underline{h}} + I^{\underline{h}}_{h} u^{h}.$

Despite this formal definition, FAC need not utilize exact solvers on the global coarse grid or the refinement patch. Historically, FAC has been used predominantly with iterative methods [14].

## 7.2 The spotlight grid

We utilize FAC methodology to devise a discretization for the spotlight CT problem. We begin with a global grid $\Omega^{2h}$ generated by the natural pixels $\psi^{2h}_j$. We next add a refinement grid $\Omega^h$ by forming strips $\psi^h_j$. For each view we choose pairs of strips whose union conforms exactly to one of the global strips. As a very simple example, consider the discretization given by three views, each consisting of four strip pixels (Figure 13, left). The resulting global (coarse grid) operator is the $12 \times 12$ matrix $B_{2h,2h}$, the matrix whose $(j,k)^{th}$ entry is $b^{2h}_{jk} = \left\langle \psi^{2h}_j, \psi^{2h}_k \right\rangle$. The refinement grid $\Omega^h$ is formed by partitioning one strip from each view into two strips. Hence $\psi^{2h}_2$ generates $\psi^h_1$ and $\psi^h_2$. Similarly, strips 3 and 4 on the refinement grid are a partition of $\psi^{2h}_7$ and $\psi^{2h}_{11}$ generates refinement

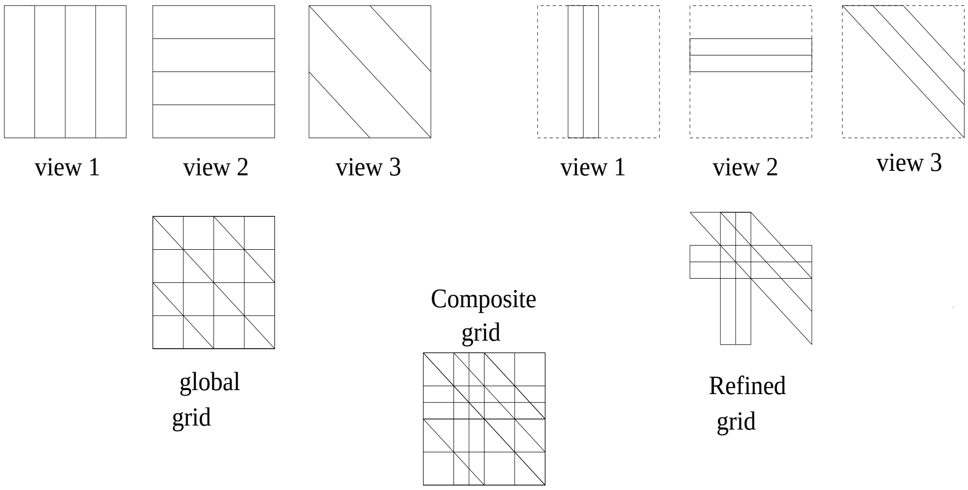


Figure 13: *The global grid $\Omega^h$ (bottom left) is generated by strips from three views (top left), while the refinement patch (bottom right) is generated by finer strips from the three views (top right). The union of all strips generates the composite grid (bottom center).*

strips 5 and 6 (Figure 13, right). The refinement strips $\psi_k^h$ generate a refinement patch operator $B_{h,h}$, formed in the usual manner by $b_{jk}^h = \langle \psi_j^h, \psi_k^h \rangle$, giving a $6 \times 6$ matrix in the simple example. The composite grid $\Omega^{\underline{h}}$ formed in this manner is shown at the bottom of Figure 13.

The basic assumption of the discretization, that the unknown image is a linear combination of the strip functions, is unchanged. The total number of strips is $N = N_{2h} + N_h$, the sum of the number of strips on the global grid and the refinement patch. Hence, the image $u = A^{\underline{h}*}\mathbf{w}$ is given by (3) where $\psi_j = \psi_j^{2h}$ if $1 \leq j \leq N_{2h}$ and $\psi_j = \psi_{j-N_{2h}}^h$ if $N_{2h} < j \leq N$. We can define the composite grid version of $A^{\underline{h}}$ by $(Au)_j = \langle \psi_j, u \rangle$ using the same ordering of the $\psi_j$'s. The discretized composite grid operator $B_{\underline{h},\underline{h}}$ may then be computed in the standard way, by $B_{\underline{h},\underline{h}} = A^{\underline{h}} A^{\underline{h}*}$, and it has, as its $(j,k)^{th}$ entry, the element $b_{jk}^{\underline{h}} = \langle \psi_j, \psi_k \rangle$. This method accounts for the interaction of the refinement grid with the global coarse grid, including the inner products between the fine strips and coarse strips. In fact, this leads to a natural partitioning of the composite grid operator as

$$B_{\underline{h},\underline{h}} = \begin{pmatrix} B_{2h,2h} & B_{2h,h} \\ B_{h,2h} & B_{h,h} \end{pmatrix}.$$

We observe that by the ordering of the $\psi_j$'s and the definition of $A^{\underline{h}*}$ we are led to a natural definition of the composite grid unknown, namely $u^{\underline{h}} = (\mathbf{w}^{2h} \ \ \mathbf{w}^h)^T$. Naturally, we must have a compatible composite grid data vector $\mathbf{f}^{\underline{h}} = (\mathbf{f}^{2h} \ \ \mathbf{f}^h)^T$. Conceivably, the coarse grid data and the refinement data could be acquired in separate recordings, but it is more likely that a single data set be generated, from which the coarse and refinement data are derived. The composite grid problem $B_{\underline{h},\underline{h}}\mathbf{w}^{\underline{h}} = \mathbf{f}^{\underline{h}}$ then becomes

$$\begin{pmatrix} B_{2h,2h} & B_{2h,h} \\ B_{h,2h} & B_{h,h} \end{pmatrix} \begin{pmatrix} \mathbf{w}^{2h} \\ \mathbf{w}^h \end{pmatrix} = \begin{pmatrix} \mathbf{f}^{2h} \\ \mathbf{f}^h \end{pmatrix}. \tag{11}$$

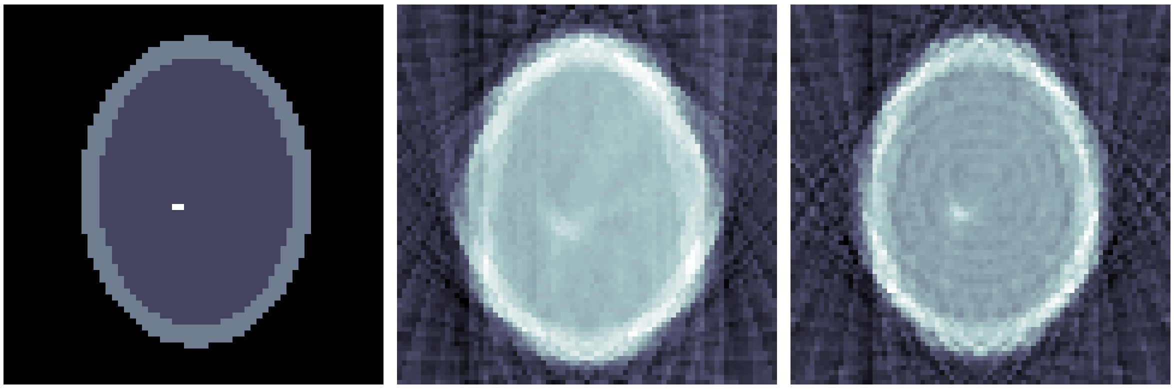

Figure 14: *The "exact" image used to generate data for the spotlight tomography problem is shown on the left. In the center is shown the PML reconstruction on the global coarse grid $\Omega^{2h}$, using data generated for 20 angles with 32 strips per angle. On the right is the spotlight reconstruction, generated using data for strips half the width of the global coarse grid, over the central region of the image.*

### 7.3 FAC Implementation

It can be shown [14, 20] that FAC in this setting is equivalent to applying two steps of *block* Gauss-Seidel iteration to the system (11). That is, FAC takes the two step form:

$$\textbf{Step 1:}\ \ \textit{Set}\quad \mathbf{w}^{2h} \leftarrow B^{-1}_{2h,2h}(\mathbf{f}^{2h} - B_{2h,h}\mathbf{w}^{h}),$$

$$\textbf{Step 2:}\ \ \textit{Set}\quad \mathbf{w}^{h} \leftarrow B^{-1}_{h,h}(\mathbf{f}^{h} - B_{h,2h}\mathbf{w}^{2h})$$

These steps are formal, of course, since we know that $B_{2h,2h}$ is singular. In fact, we take $\mathbf{w} \leftarrow B^{-1}\mathbf{f}$ to mean "solve $B\mathbf{w} = \mathbf{f}$", which need not be done with exact solvers. In principle we may apply any method to these subproblems: ART, filtered backprojection, Fourier methods. A natural choice is an iterative method, such as Gauss-Seidel or multigrid. Noting that each of the steps are solving a "residual" equation on one of the grids, this process may be viewed as one of multilevel correction.

The composite grid operator $B_{\underline{h},\underline{h}}$ possesses a host of useful and interesting properties [20], related to a family of useful properties generated by the discretization method and inherent in the global and local operators $B_{2h,2h}$ and $B_{h,h}$. Space limitations do not permit elaboration here, nor is there room for a performance analysis. A rigorous treatment of the method, including a performance assesment, is forthcoming [8, 20].

We demonstrate the promise of the spotlight method with a simple example. A "brain" phantom is generated, consisting of a uniform grey region within the skull (high-density elliptical ring). Embedded in the grey region is a small square high-density region. Data is generated by integrating the product of this image with the characteristic functions of the strips representing a 20 angle, 32 bins-per-angle discretization. The square of high density has width equal to one half the width of the individual strips making up the global coarse grid. The "exact" image is shown on the left of Figure 14, while the global coarse grid reconstruction is shown in the center.

A single-level refinement region is generated by refining one half of the strips in the center of the set for each angle, using strips of half the width of those on the global coarse grid. The

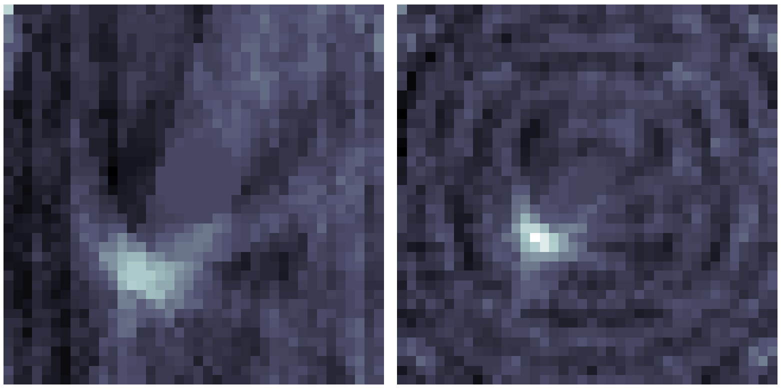

Figure 15: *The central portions of global coarse grid (left) and spotlight (right) reconstructions of Figure 14 are shown in this figure. The image reconstructed from the composite grid data clearly shows the object of interest, while the global coarse grid image fails to resolve it.*

reconstructed image using the spotlight method with 1 cycle of FAC is shown on the right of Figure 14. Both the global coares grid and refinement grid portions of the composite grid were solved using 3 V-cycles, each with 2 relaxation sweeps on the downward leg and 1 relaxation sweep on the upward leg. The high-density region, which does not show in the global coarse-grid reconstruction, appears in the spotlight reconstruction. This is demonstrated a bit more clearly in Figure 15, in which only the central region of each of the reconstructions is shown.

We chose this example because of its clarity for the reader. However, it does not really illustrate the practical benefits of FAC because it costs essentially the same as would solving the globally refined problem. In practice, we envision that FAC will be used for spotlighting smaller features of the image, and to much finer detail. That is, rather than refining one half of the strips on each angle by splitting them once, we foresee refining a much smaller region, such as one tenth of the strips along each angle, to a resolution four or eight times that of the global coarse grid. In such settings the benefits of spotlighting would be very substantial.

## 8 Concluding Remarks

The results presented here are encouraging, in that they demonstrate that multilevel methodology can be applied to the image reconstruction problem with some hope of success. The benefits of multilevel reconstruction, the way we have developed it, remain somewhat limited, although we do see images of quality equal to those produced by Gauss-Seidel, achieved at somewhat lower cost.

We believe that this limitation may stem from restricting the coarsening to be within projections. That is, we have reduced only the number of detectors per angle, not the number of angles themselves. Evidence gathered by examining the near null space components suggests that angle coarsening is essential for efficiency at coarse grain resolution, which multigrid methods always face. This is currently being explored.

The results presented here show great promise in the area of spotlight tomography, for cases where a finer resolution image is needed over portions of the image space. It is not feasible to compute entire images at the fine resolution, since such problems lead to extremely large, dense systems. As the simple example shows, however, PML can be used to formulate the spotlight problem to use FAC technology in a way that may lead to practical algorithms.

## References


[1] M. H. Bounocore, W. R. Brody, and A. Macovski, *A natural pixel decomposition for two dimensional image reconstruction*, IEEE Transactions on biomedical engineering, BME-28 (1981), pp. 69–78.

[2] A. Brandt, *Multi-level adaptive techniques (MLAT) for partial differential equations: ideas and software*, in Proceedings of the symposium on mathematical software, 1977, New York, NY, 1977, Academic Press, pp. 277–318.

[3] ——, *Multi-level adaptive solutions to boundary value problems*, Mathematics of Computation, 31 (1987), pp. 333–390.

[4] ——, *Rigorous local mode analysis of multigrid*, in Preliminary proceedings of the fourth Copper Mountain conference on multigrid methods, J. Mandel and S. F. McCormick, eds., 1989.

[5] W. L. Briggs, *A Multigrid Tutorial*, Society for Industrial and Applied Mathematics, Philadelphia, 1987.

[6] K. J. Cavanaugh and V. E. Henson, *A multilevel cost-space approach to solving the balanced long transportation problem*, in Proceedings of the 6th Copper Mountain Conference on Multigrid Methods, vol. CP-3224, NASA Conference Publications, 1993.

[7] C. C. Douglas and J. Douglas, *A unified convergence theory for abstract multigrid or multilevel algorithms, serial and parallel*, SIAM Journal of Numerical Analysis, 30 (1993), pp. 136–158.

[8] V. E. Henson, M. A. Limber, S. F. McCormick, and B. T. Robinson, *Multilevel projection methods for spotlight computed tomography*, in preparation.

[9] G. T. Herman, *Image Reconstruction from Projections*, Academic Press, Orlando, Florida, 1980.

[10] M. A. Limber, T. A. Manteuffel, S. F. McCormick, and D. S. Sholl, *Optimal resolution in maximum entropy image reconstruction from projections with multigrid acceleration*, in Proceedings of the Sixth Annual Copper Mountain Conference on Multigrid Methods, 1993.

[11] A. K. Louis, *Medical imaging: the state of the art and future development*, Inverse Problems, 8 (1992), pp. 709–738.

[12] J. Mandel and S. F. McCormick, *Iterative solution of elliptic equations with refinement: the model multilevel case*, in Domain decomposition methods, T. F. Chan, ed., Philadelphia, PA, 1987, Society for Industrial and Applied Mathematics, pp. 81–92.

[13] S. F. McCormick, *The methods of Kaczmarz and row orthogonalization for solving linear equations and least squares problems in Hilbert space*, Indiana University Mathematics Journal, 26 (1977), pp. 1137–1150.

[14] ——, *Multilevel adaptive methods for partial differential equations*, vol. 6 of Frontiers in applied mathematics, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1989.

[15] ——, *Multilevel projection methods for partial differential equations*, vol. 62 of CBMS-NSF regional conference series in applied mathematics, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.

[16] S. F. McCormick and J. Thomas, *The fast adaptive composite grid method (FAC) for elliptic boundary value problems*, Math. Comp., 46 (1986), pp. 439–456.

[17] S. F. McCormick and J. G. Wade, *Multilevel parameter estimation*, in Proceedings of the 5th Copper Mountain Conference on Multigrid Methods, 1991.

[18] F. Natterer, *The Mathematics of Computerized Tomography*, John Wiley and Sons, New York, N.Y., 1986.

[19] T.-S. Pan, A. E. Yagle, N. H. Clinthorne, and W. L. Rogers, *Acceleration and filtering in the generalized Landweber iteration using a variable shaping matrix*, IEEE Transactions on Medical Imaging, 12 (1993), pp. 278–286.

[20] B. T. Robinson, *A Multilevel Approach to the Algebraic Image Reconstruction Problem*, PhD thesis, Naval Postgraduate School, 1994.

[21] U. Rüde, *Mathematical and computational techniques for multilevel adaptive methods*, vol. 13 of Frontiers in applied mathematics, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1993.

[22] J. Ruge and K. Stüben, *Algebraic multigrid (AMG)*, in Multigrid methods, S. F. McCormick, ed., vol. 3 of Frontiers in applied mathematics, Philadelphia, PA, 1987, Society for Industrial and Applied Mathematics, pp. 131–177.

[23] L. A. Shepp and B. F. Logan, *The Fourier reconstruction of a head section*, IEEE Trans. Nucl. Sci., NS-21 (1974), pp. 21–43.

[24] H. Stark, J. W. Woods, I. Paul, and R. Hingorani, *Direct Fourier reconstruction in computer tomography*, IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-29 (1981), pp. 237–244.

[25] K. Tanabe, *Projection method for solving a singular system of linear equations and its apllications*, Numerical Mathematics, 17 (1971), pp. 203–214.