Error Analysis and Performance Optimization of Fast Hierarchical Backprojection Algorithms

Samit Basu, *Member, IEEE,* and Yoram Bresler, *Fellow, IEEE*

***Abstract*—We recently proposed a novel fast backprojection algorithm for reconstruction of an $N \times N$ pixel object from $O(N)$ projections in $O(N^2 \log_2 N)$ operations. In this paper, we analyze a simplified version of that algorithm, to determine the effects of various parameter choices on the algorithm's theoretical performance. We derive a bound on the variance of the per-pixel error introduced by using the hierarchical backprojection. This bound is with respect to an ensemble of input sinograms, and allows us to construct confidence intervals (for any specified level) for the per-pixel errors. The bound has a simple form, and we show how to use it to select algorithm parameters for different cost versus error tradeoffs. Simulation results show that the bound accurately predicts the performance of the algorithm over a wide range of parameter choices. These results are verified for different images, including a tomographic reconstruction from the visual human dataset (VHD). The analysis therefore provides an effective tools for the selection of parameters and operating point for the fast hierarchical backprojection algorithm.**

***Index Terms*—Error analysis, hierarchical backprojection, tomographic reconstruction.**

I. Introduction

TOMOGRAPHIC reconstruction of an image from a set of parallel beam line-integral projections, is a technique which has found a number of applications, ranging from nondestructive evaluation (NDE), and medical imaging, to synthetic aperture radar (SAR) and radioastronomy (see, e.g., [1]). When a sufficiently complete set of data is available, the reconstruction technique of choice is filtered backprojection (FBP). Unfortunately, the backprojection step in the FBP is relatively expensive, requiring $O(N^3)$ operations for reconstruction of an $N \times N$ image from $O(N)$ projections in $O(N^2 \log_2 N)$ operations.

Recently, we introduced a new fast hierarchical backprojection (FHBP) algorithm for reconstruction of an $N \times N$ object from $O(N)$ projections [2]. The idea behind the algorithm was the following. Under special conditions (specifically, a flexible radial sampling scheme), we demonstrated that it was possible to exactly decompose the backprojection of P projections onto an $N \times N$ object into the sum of four backprojections of P projections onto $N/2 \times N/2$ objects, each of which represented a quadrant of the original backprojection, centered at the origin. Using the well known essentially bandlimited properties of the Radon transform, we then replaced these backprojections by backprojections of $P/2$ projections onto $N/2 \times N/2$ objects. In doing so, we introduced radial filtering and angular decimation steps into the processing. The decomposition was applied recursively, further subdividing the $N/2 \times N/2$ objects into quadrants, until the single pixel level was reached. The results could then be aggregated into a complete reconstruction. Furthermore, by using a combination of the exact and approximate decompositions, we were able to control the cost vs. accuracy tradeoff.

The algorithm proposed in [2], as well as its simplified version analyzed in this paper, work by aggregating images of successively larger size, until the full sized backprojection image is formed. At each stage in the algorithm, each of the subimages is at the full resolution of the final image. Other fast backprojection algorithms such as the Multilevel Inversion [3], the link-based methods [4], quadtree backprojection algorithms [5], [6], and factorized backprojection [7], [8] form, at different stages of the algorithm, full-size images of successively higher resolution, until the final resolution of the backprojection image is reached.

Our simulations in [2] suggested that the new algorithm was both fast and accurate. In the formulation of the algorithm, it was necessary to specify parameters that controlled the algorithm performance. These parameters included a radial interpolation kernel, angular anti-aliasing filter, and a parameter Q that traded off accuracy for performance. For our initial studies in [2], we made the simplest possible choices for these parameters, and still obtained good results in terms of reconstruction quality.

In this paper, our goal is to analyze the FHBP algorithm to study the choice of the algorithm parameters and their effect on performance. For the sake of analytical tractability, however, we first replace the FHBP algorithm with a simpler variant. This algorithm retains most of the functionality of the original FHBP algorithm, but uses a simpler radial sampling formulation, and only one (approximate) decomposition. The cost versus accuracy tradeoff is handled differently also. The simplified FHBP algorithm decomposes the backprojections until they are a uniform L pixels in size, and these backprojections are computed exactly. Thus, by increasing L, one obtains better quality reconstructions, at the cost of higher computational effort. The resulting algorithm generally does not perform as well as the original FHBP algorithm, but the analysis of this simpler algorithm is still instructive in terms of parameter selection for the original FHBP.

We perform a stochastic error analysis of the modified FHBP algorithm, in order to study the effect of the parameter choices

Manuscript received March 15, 2000; revised February 23, 2001. This work was supported in part by the NSF under Grant CCR 99-72980, NSF Research Infrastructure Grant CDA 96-24396, and by the Mac Van Valkenburg Memorial Fellowship. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Mark R. Luettgen.

S. Basu is with the General Electric Corporate Research and Development Center, Niskayuna, NY 12039 USA (e-mail: basu@crd.ge.com).

Y. Bresler is with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: ybresler@uiuc.edu).

Publisher Item Identifier S 1057-7149(01)04527-4.

on the algorithm performance. As a result, we are able to bound the per-pixel error variance in terms of the image bandwidth and choice of interpolators. This is a point-wise variance over an ensemble of input sinograms, as opposed to a simple RMS error bound. In particular a bound on the error variance for the per-pixel error provides us with confidence intervals (for any specified level) on the per-pixel error. The paper is organized as follows. In Section II, we present the new, simplified FHBP algorithm which we propose to analyze. In Section III, we derive bounds on the pixel error variance as a function of the various parameters of interest. In Section IV, we discuss the computational complexity tradeoffs associated with the various parameters. Section V presents simulations comparing actual algorithm performance to the prediction of the bounds. Conclusions and suggestions for future research are presented in Section VI.

II. Simplified Fast Hierarchical Backprojection Algorithm

The original FHBP that we derived in [2] performed well in our phantom studies. However, the resulting algorithm is difficult to study analytically because of the nature of the processing steps involved. In this section, we derive a new, simplified FHBP algorithm [which we call the simplified fast hierarchical backprojection (SFHBP) algorithm], which is similar to the FHBP algorithm in terms of parameter choices and characteristics, but uses slightly different approximations and is simpler to analyze. The algorithm is also similar to the dual of the multilevel domain decomposition (MDD) algorithm of Boag *et al.* [9], which is a fast algorithm for reprojection that is based on a hierarchical domain-decomposition of the Radon transform.

We will use an operator and Hilbert space formulation to make the manipulations of the various steps clear. The space $L_2(\mathbb{R}^2)$ is the standard Hilbert space of square-integrable functions defined on $\mathbb{R}^2$. This space represents the set of all possible spatially continuous objects. The space $\ell_2(\mathbb{Z}^2)$ is the space of 2-D, square summable sequences. A typical element in $\ell_2(\mathbb{Z}^2)$ is the result of spatially discretizing (sampling) an element of $L_2(\mathbb{R}^2)$. The space $L_2^P(\mathbb{R})$ is the P-wise Cartesian product of $L_2(\mathbb{R})$, with an element $g_c(r, p)$ in $L_2^P(\mathbb{R})$ indexed by $p \in \{0, 1, \ldots, P-1\}$ and $r \in \mathbb{R}$. A typical element in $L_2^P(\mathbb{R})$ is the set of P radially continuous projections of an object in $L_2(\mathbb{R}^2)$. The final space of interest is $\ell_2^P(\mathbb{Z})$, which is the P-wise Cartesian product of $\ell_2(\mathbb{Z}^2)$, with an element $g(k, p)$ in $\ell_2^P(\mathbb{Z})$ indexed by $p \in \{0, 1, \ldots, P-1\}$ and $k \in \mathbb{Z}$. A typical element in $\ell_2(\mathbb{Z}^2)$ is a set of radially sampled projections of an object.

A. Discrete Backprojection Operator $\mathbf{B}_{P,N}$

Our goal is an implementation of the *discrete angle backprojection* operator: $\mathcal{B}_P$: $L_2^P(\mathbb{R}) \mapsto L_2(\mathbb{R}^2)$, defined[1] by (see, e.g., [1] and [10])

$$\mathcal{B}_P g_c(x, y) = \sum_{p=0}^{P-1} g_c(x\cos\theta_p + y\sin\theta_p, p) \tag{1}$$

[1]Our definition of backprojection does not include a $2\pi/P$ weighting for two reasons. First, this weighting is only correct for an equally spaced angle distribution. Second, we want our backprojection to be the exact adjoint of the Radon transform (see [2] for more details).

where $\theta_p \in [0, 2\pi)$. For convenience, we will assume that

$$\theta_p = \frac{2\pi p}{P} \tag{2}$$

for $p \in \{0, \ldots, P-1\}$, although in principle, the formulation of the FHBP (and the analysis) could be easily extended to other sets of view angles. Note that we are not exploiting the symmetry of the Radon transform by sampling on $[0, \pi)$, although our analysis could also be extended to that case.

Of course, implementation of (1) in practice is impossible, because it maps radially continuous projections to a continuous reconstruction. Therefore, for implementation purposes, it is necessary to start and end with discrete (and finite length) data. Given a set of sampled projections $g \in \ell_2^P(\mathbb{Z})$, an implementation of (1) involves the following steps: 1) radial interpolation from $\ell_2^P(\mathbb{Z})$ to $L_2^P(\mathbb{R})$, 2) backprojection [via (1)] from $L_2^P(\mathbb{R})$ to $L_2(\mathbb{R}^2)$, 3) spatial sampling from $L_2(\mathbb{R}^2)$ to $\ell_2(\mathbb{Z}^2)$, 4) truncation to the region of interest (ROI) from $\ell_2(\mathbb{Z}^2)$ to $\ell_2(\mathbb{Z}^2)$. The operator that combines all four of these steps is the discretized backprojection operator $\mathbf{B}$: $\ell_2^P(\mathbb{Z}) \mapsto \ell_2(\mathbb{Z}^2)$. To define it, we will first describe the operators to implement each of Steps 1)–4).

Step 1 is the radial interpolation step, which is performed by the radial interpolation operator $\mathcal{I}$: $\ell_2^P(\mathbb{Z}) \mapsto L_2^P(\mathbb{R})$

$$\mathcal{I}g(r, p) = \sum_k g(k, p)\phi(r - kT) \tag{3}$$

where ϕ is the radial interpolation kernel, and T is the radial sampling period. Note that our definition of $\mathcal{I}$ is simpler than the radial interpolation operator in [2], and this leads to some simplifications in the formulation of the fast algorithm. Step 2, of course, is computed via (1).

Step 3 is the spatial sampling step, which is performed by the spatial sampling operator $\mathcal{S}$: $L_2(\mathbb{R}^2) \mapsto \ell_2(\mathbb{Z}^2)$

$$\mathcal{S}f_c(i, j) = \iint f_c(x, y)b(i - x, j - y)\,dx\,dy. \tag{4}$$

The function $b \in L_2(\mathbb{R}^2)$ is the generalized sampling kernel, and $\mathcal{S}$ corresponds to convolution of f_c with b, followed by sampling on the integer lattice. Thus, b serves as a kind of antialiasing filter prior to sampling. Step 4 is simply a truncation operator $\mathbf{K}_N$: $\ell_2(\mathbb{Z}^2) \mapsto \ell_2(\mathbb{Z}^2)$, defined by

$$\mathbf{K}_N f(i, j) = \begin{cases} f(i, j), & -\dfrac{N}{2} \le i,\ j < \dfrac{N}{2} \\ 0, & \text{else.} \end{cases} \tag{5}$$

Thus, the discrete backprojection $\mathbf{B}_{P,N}$: $\ell_2^P(\mathbb{Z}) \mapsto \ell_2(\mathbb{Z}^2)$ is defined by

$$\mathbf{B}_{P,N} = \mathbf{K}_N \mathcal{S}\mathcal{B}_P\mathcal{I}. \tag{6}$$

Alternately, we can write the discrete backprojection explicitly as

$$\mathbf{B}_{P,N}g(i, j) = \begin{cases} \displaystyle\sum_k \sum_p g(k, p)\eta_{k,j,k,p} & -\dfrac{N}{2} \le i,\ j < \dfrac{N}{2} \\ 0 & \text{else,} \end{cases} \tag{7}$$

where

$$\eta_{i,j,k,p} = \iint b(i-x,\, j-y)\phi(x\cos\theta_p + y\sin\theta_p - kT)\,dx\,dy. \tag{8}$$

Now, it is simple to prove that $\eta_{i,j,k,p} = 0$ for all i, j, k, p that satisfy the following relationship

$$|i\cos\theta_p + j\sin\theta_p - kT| \geq R_\phi + R_b \tag{9}$$

where R_ϕ is the radius of support of ϕ and R_b is the radius of support of b. If we restrict $-N/2 \leq i, j < N/2$, then $\eta_{i,j,k,p} = 0$ for all k such that

$$|k| \leq \frac{R_\phi + R_b + N/\sqrt{2}}{T}. \tag{10}$$

Equations (9) and (10) indicate that the discrete backprojection $\mathbf{B}_{P,N}$ is a local operation, in the following sense:

$$\mathbf{B}_{P,N} = \mathbf{B}_{P,N}\hat{\mathbf{K}}_N \tag{11}$$

where $\hat{\mathbf{K}}_N$: $\ell_2^P(\mathbb{Z}) \mapsto \ell_2^P(\mathbb{Z})$ is defined by

$$\hat{\mathbf{K}}_N g(k,p) = \begin{cases} g(k,p), & |k| \leq \dfrac{R_\phi + R_b + N/\sqrt{2}}{T} \\ 0, & \text{else.} \end{cases} \tag{12}$$

Although our presentation of the SFHBP algorithm is highly abbreviated, we will need one additional property of $\mathbf{B}_{P,N}$ throughout our analysis of the algorithm. Let us make the following assumptions:

- A1) radial interpolation $\mathcal{I}$ is an ideal bandlimited interpolator with bandlimit B;
- A2) P, N, and B satisfy the Nyquist criterion: $P > NB\sqrt{2}$, and T, B satisfy the Nyquist criterion: $T > \pi/B$.

Assumption A2) is directly related to the *bow tie property* of the sinogram [11], in that the two-dimensional (2-D) Fourier transform of the sinogram data is supported on a bow tie shaped region, shown in Fig. 1, from which the Nyquist criteria are determined.

Next, let us define a sinogram domain filter $\mathbf{C}_N$: $\ell_2^P(\mathbb{Z}) \mapsto \ell_2^P(\mathbb{Z})$, that is a lowpass filter with radial bandwidth BT, and an angular bandwidth of $NB/\sqrt{2}$ (see Fig. 1). Then the bow-tie spectral support of the sinogram along with Assumptions A1) and A2), imply that [2], [12] for any $V \geq 1$

$$\mathbf{B}_{L,V} = \mathbf{B}_{L,V}\mathbf{C}_V \qquad L > VB\sqrt{2}. \tag{13}$$

B. Simplified Fast Hierarchical Decomposition of the Backprojection

We now present the hierarchical decomposition of the backprojection operator that leads to the simplified fast hierarchical backprojection (SFHBP) algorithm. Proofs and details relevant to the construction of this decomposition can be found in [12]. The manipulations and arguments used in the derivation of this decomposition are also similar to the treatment in [2].

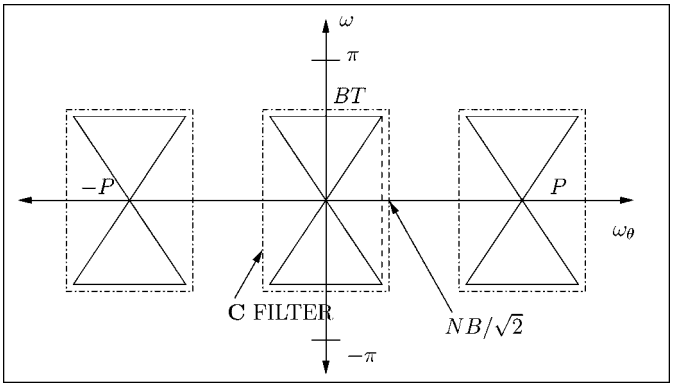

Fig. 1. Spectral support of a radially and angularly sampled sinogram with P projections from an object supported on a square of size $N \times N$ (centered at the origin), with a radial bandwidth of B, and a radial sampling interval of T. The operator $\mathbf{C}_N$ is a lowpass filter that passes only the regions shown.

To construct the decomposition, we will need to define several additional operators. The first, $\mathbf{Z}$: $\ell_2(\mathbb{Z}^2) \mapsto \ell_2(\mathbb{Z}^2)$ is a simple spatial shifter, parameterized by an offset $\boldsymbol{\delta} \in \mathbb{Z}^2$

$$\mathbf{Z}_{\boldsymbol{\delta}} f(i,j) = f(i-\delta_1,\, j-\delta_2). \tag{14}$$

The second operator $\mathbf{D}_j$: $\ell_2^{P/2^{j-1}}(\mathbb{Z}) \mapsto \ell_2^{P/2^j}(\mathbb{Z})$ is an angular decimation operator, which angularly convolves the sinogram with a lowpass filter ψ, and then downsamples by a factor of two. Thus

$$\mathbf{D}_j g(k,p) = \sum_m g(k,m)\psi(2p-m), \qquad p \in \{0,\ldots,P/2^j - 1\}. \tag{15}$$

The next operator $\mathbf{F}_{j,i}$: $\ell_2^{P/2^{j-1}}(\mathbb{Z}) \mapsto \ell_2^{P/2^{j-1}}(\mathbb{Z})$ is simply a radial shifter, which when applied to a discrete set of projections $g(k,p)$ shifts the pth projection by $\gamma_p = \{\delta_i(N/2^j)\}_1\cos\theta_p + \{\delta_i(N/2^j)\}_2\sin\theta_p$, where

$$\begin{aligned} \boldsymbol{\delta}_1(N) &= (N/4,\, N/4)', & \boldsymbol{\delta}_2(N) &= (-N/4,\, N/4)', \\ \boldsymbol{\delta}_3(N) &= (N/4,\, -N/4)', & \boldsymbol{\delta}_4(N) &= (-N/4,\, -N/4)' \end{aligned} \tag{16}$$

for $N \geq 4$, and

$$\begin{aligned} \boldsymbol{\delta}_1(2) &= (0,\,0)', & \boldsymbol{\delta}_2(2) &= (-1,\,0)', \\ \boldsymbol{\delta}_3(2) &= (0,\,-1)', & \boldsymbol{\delta}_4(2) &= (-1,\,-1)' \end{aligned} \tag{17}$$

for $N = 2$. Because the shift γ_p need not be an integer, $\mathbf{F}$ is expressed in terms of three steps:

1) interpolation to a radially continuous sinogram;
2) delay of each continuous projection by γ_p;
3) resampling of the delayed projections.

In the signal processing literature, for a fixed p, $\mathbf{F}$ is known as a fractional delay (FD) [13]. The operator $\mathbf{F}$ can be expanded as

$$\begin{aligned} &\mathbf{F}_{j,i} g(k,p) \\ &\quad = \sum_n g(n,p)\phi\left(T(k-n) - \{\delta_i(N/2^j)\}_1\cos\theta_p \right. \\ &\qquad\qquad \left. + \{\delta_i(N/2^j)\}_2\sin\theta_p\right). \end{aligned} \tag{18}$$

Finally, we also define the operator $\mathbf{A}_{j,i}$: $\ell_2^{P/2^{j-1}}(\mathbb{Z}) \mapsto \ell_2^{P/2^j}(\mathbb{Z})$ to be

$$\mathbf{A}_{j,i} = \hat{\mathbf{K}}_{N/2^j}\mathbf{D}_j\mathbf{F}_{j,i}. \tag{19}$$

Now, we introduce an additional assumption

A3) The angular filter ψ is an ideal lowpass filter with cutoff $\pi/2$ and gain 2.

Then, subject to Assumptions A1)–A3), the following decomposition holds [12]

$$\mathbf{B}_{P,N} = \sum_{i=1}^{4} \mathbf{Z}_{\boldsymbol{\delta}_i(N)}\mathbf{B}_{P/2,N/2}\mathbf{A}_{1,i}. \tag{20}$$

Equation (20) decomposes the backprojection of P projections onto an $N \times N$ image into four backprojections $\mathbf{B}_{P/2,N/2}$, each of $P/2$ projections onto an $N/2 \times N/2$ image, with some additional preprocessing denoted by $\mathbf{A}_{1,i}$ and shifts $\mathbf{Z}_{\boldsymbol{\delta}_i(N)}$. While we omit here the derivation of (20), it is important to note that it relies heavily on the bow tie support of the spectrum, or equivalently, on (13).

Equation (20), and our derivation of it in [12] differs from the decompositions presented in [2] in two ways. In [2], we constructed two decompositions of the backprojection operator, one exact and one approximate. These were used in combination to decompose the backprojection operator. In contrast, in this paper, we have constructed only one approximate decomposition—that of (20). Furthermore, (20) differs from the approximate decomposition of [2] because the processing prior to the backprojection, namely $\mathbf{A}_{1,i}$, is separable, i.e., the radial and angular processing can be done separately. This separability of the sinogram preprocessing step is the key to the analytical tractability of the SFHBP.

To use (20) in a hierarchical algorithm, we apply it recursively, decomposing each $\mathbf{B}_{P/2,N/2}$ into a sum of four backprojections $\mathbf{B}_{P/4,N/4}$. This process can continue down to the single pixel level, in which case the original $\mathbf{B}_{P,N}$ is decomposed into a sum of N^2 backprojections of the type $\mathbf{B}_{P/N,1}$, i.e., backprojection of P/N projections onto a single pixel. However, for the SFHBP algorithm, we allow for a parameter J that controls the number of times the recursion is used, after which the remaining backprojections are computed from definition (6) instead of using (20). For example, if $J = 1$, then we decompose in accordance with (20). If $J = 2$, on the other hand, then we compute

$$\mathbf{B}_{P,N} = \sum_{i=1}^{4} \mathbf{Z}_{\boldsymbol{\delta}_i(N)} \left(\sum_{l=1}^{4} \mathbf{Z}_{\boldsymbol{\delta}_l(N/2)}\mathbf{B}_{P/4,N/4}\mathbf{A}_{2,l} \right) \mathbf{A}_{1,i}. \tag{21}$$

Thus, for $J \in \{0, \ldots, \log_2 N\}$, we effectively decompose $\mathbf{B}_{P,N}$ into a sum of 4^J backprojections, each of $P/2^J$ projections onto an image of size $N/2^J$ centered at the origin.

Furthermore, in [12], we show that the cost of applying (20) recursively J times is

$$\Xi = \frac{4JP\left(R_\phi + R_b + \lceil N/\sqrt{2} \rceil\right)(R_\psi + 2R_\phi)}{T} + \frac{4PN^2(R_\phi + R_b)}{2^J T} \tag{22}$$

TABLE I
PARAMETERS THAT CONTROL THE SFHBP ALGORITHM

Parameter	Symbol
Number of projections	P
Reconstruction size	N
Number of times to apply decomposition	J
Projection interpolation kernel	ϕ
Angular anti-aliasing filter	ψ
Image discretization kernel	b
Projection sampling interval	T

floating point operations, where R_ϕ, R_ψ and R_b are the radius of support for ϕ, ψ, and b, respectively. If we choose $J = \log_2 N$, we recover an $O(NP\log_2 N)$ algorithm for reconstruction of an $N \times N$ image from $O(P)$ projections. If $P = O(N)$, then the total cost is $O(N^2 \log_2 N)$. The parameters that need to be determined for this algorithm are summarized in Table I.

If we choose ϕ and ψ to satisfy Assumptions A1) and A3), respectively, then $R_\phi = \infty$ and $R_b = P$. But for computational complexity reasons, we choose ϕ and ψ to be finite length (short) filters. These shorter filters can be used if the projections are angularly and radially oversampled, a point we will return to later. For these nonideal filters, Assumptions A1), and A3) become approximations, and in particular, (20) becomes an approximation instead of an equality.

III. ERROR ANALYSIS

A. Sources of Error

Recall that to obtain a fast algorithm, we had to relax Assumptions A1) and A3), and use shorter filters ϕ and ψ. These choices will affect the quality of the approximation in each of the decomposition steps for which (20) is now only an approximation. The errors will accumulate over the multiple stages in the hierarchy resulting in errors in the final backprojection. Furthermore, the size of the error will depend not only on ϕ and ψ, but also on the other parameters in the problem (see Table I). Thus, we now turn to the problem of determining how the choice of the various parameters affects the algorithm's performance. To do so, we will employ a stochastic error analysis to derive approximate bounds on the backprojection error that depend on each of the various parameters.

Other than the relaxation of Assumptions A1) and A3), there is another potential source of error in (20). It is due to the fact that the bow-tie support result is itself an approximation (albeit a very good one, see [10], [11]). However, we expect the backprojection error to be dominated by the error due to the relaxation of Assumptions A1) and A3), and do not analyze the effects of deviations from the bowtie support. Furthermore, the bow-tie support approximation improves with the amount of radial and angular oversampling used.

B. Hierarchical Error Accumulation

We now turn to the problem of determining how the errors in the hierarchical decomposition propagate through the various stages. We will study a slightly different representation for (20)

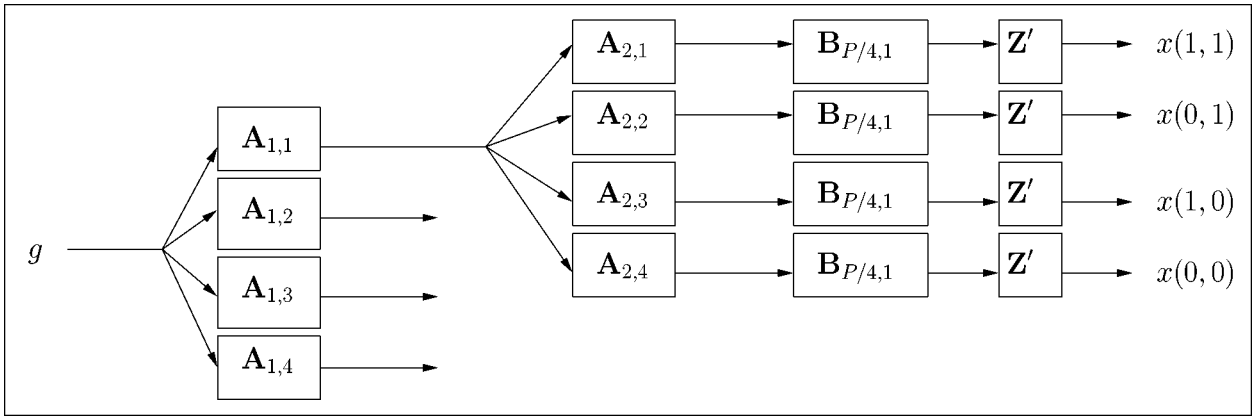

Fig. 2. SFHBP for a 4×4 image.

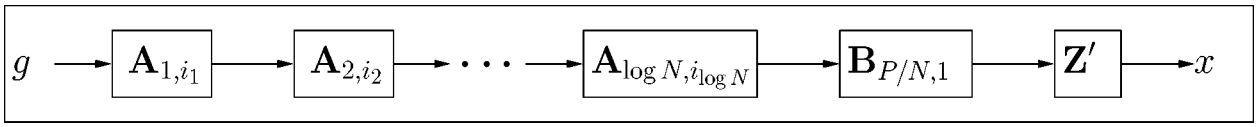

Fig. 3. Path from the input projections to any pixel in an $N \times N$ backprojection computed via the SFHBP.

which allows us to emphasize the frequency-domain characterization of the errors. First, we expand (20) via (19), to obtain

$$\mathbf{B}_{P,N} = \sum_{i=1}^{4} \mathbf{Z}_{\boldsymbol{\delta}_i(N)} \mathbf{B}_{P/2,N/2} \hat{\mathbf{K}}_{N/2} \mathbf{D}_1 \mathbf{F}_{1,i}. \tag{23}$$

Next, we apply (11) to reabsorb the radial truncation into the backprojection, yielding

$$\mathbf{B}_{P,N} = \sum_{i=1}^{4} \mathbf{Z}_{\boldsymbol{\delta}_i(N)} \mathbf{B}_{P/2,N/2} \mathbf{D}_1 \mathbf{F}_{1,i}. \tag{24}$$

Then, we apply (13) to reinsert the filtering prior to backprojection, yielding

$$\mathbf{B}_{P,N} = \sum_{i=1}^{4} \mathbf{Z}_{\boldsymbol{\delta}_i(N)} \mathbf{B}_{P/2,N/2} \mathbf{C}_{N/2} \mathbf{D}_1 \mathbf{F}_{1,i}. \tag{25}$$

Now, we define a new operator $\mathbf{E}$: $\ell_2^{P/2^{j-1}}(\mathbb{Z}) \mapsto \ell_2^{P/2^j}(\mathbb{Z})$ as

$$\mathbf{E}_{j,i} = \mathbf{C}_{N/2^j} \mathbf{D}_j \mathbf{F}_{j,i}. \tag{26}$$

Then, (25) becomes

$$\mathbf{B}_{P,N} = \sum_{i=1}^{4} \mathbf{Z}_{\boldsymbol{\delta}_i(N)} \mathbf{B}_{P/2,N/2} \mathbf{E}_{j,i}. \tag{27}$$

Equation (27) is the starting point for our study of error accumulation.

To construct the error bounds, let us first consider the simple case of backprojecting P projections onto a 4×4 image. This decomposition, corresponding to (21) for $N = 4$, is shown graphically in Fig. 2. For convenience, we combine all the concatenated shifts $\mathbf{Z}_{\boldsymbol{\delta}_i(N)}$ which shift the output into the correct location, into a single operator $\mathbf{Z}'$. We immediately note that the path from the input g to each of the pixels in the reconstruction $x(m, n)$ is the same (to within the shifts δ_i used along each path). Thus, returning to the general case of an $N \times N$ image, we will consider the system shown in Fig. 3, which maps the input g to a single pixel $x(m, n)$ via

$$\begin{aligned} x(m, n) &= \mathbf{Z}' \mathbf{B}_{P/N,1} \mathbf{E}_{\log N} \mathbf{E}_{\log N - 1} \cdots \mathbf{E}_1 g \\ &= \mathbf{Z}' \mathbf{B}_{P/N,1} \prod_{j=1}^{\log N} \mathbf{E}_j g(m, n), \end{aligned} \tag{28}$$

where we have dropped the second subscript of $\mathbf{E}$ for notational convenience (the second subscript indicated which quadrant of the backprojected image was being processed, so that these second subscripts only identify which pixel is being reconstructed). Note that we define the ordered product

$$\prod_{j=1}^{\log_N} \mathbf{E}_j = \mathbf{E}_{\log N} \mathbf{E}_{\log N - 1} \cdots \mathbf{E}_1. \tag{29}$$

Now, let us assume that $x(m, n)$ is computed using approximations $\tilde{\mathbf{E}}_j$ for the first J stages, i.e., for $j \in \{1, \ldots, J\}$. Then the resulting approximate pixel $\tilde{x}(m, n)$ is given by

$$\tilde{x}(m, n) = \mathbf{Z}' \mathbf{B}_{P/N,1} \prod_{j=J+1}^{\log_2 N} \mathbf{E}_j \prod_{j=1}^{J} \tilde{\mathbf{E}}_j g(m, n). \tag{30}$$

Thus, the per-pixel error is given by

$$|x(m, n) - \tilde{x}(m, n)| = |\mathbf{O}_{m,n} g| \tag{31}$$

where $\mathbf{O}_{m,n}$: $\ell_2^P(\mathbb{Z}) \mapsto \mathbb{R}$ defined by

$$\mathbf{O}_{m,n} = \mathbf{Z}' \mathbf{B}_{P/N,1} \prod_{j=J+1}^{\log N} \mathbf{E}_j \left(\prod_{j=1}^{J} \tilde{\mathbf{E}}_j - \prod_{j=1}^{J} \mathbf{E}_j \right). \tag{32}$$

To bound the variance of the per-pixel error

$$e(m, n) = x(m, n) - \tilde{x}(m, n) = \mathbf{O}_{m,n} g \tag{33}$$

we need a model for the distribution of the filtered projection data g.

First, assume that $g \in \mathbb{R}^{L\times P}$, where L is the number of samples in the radial direction. From (11) and (12), it follows that

$$L = \frac{2R_\phi + 2R_b + \sqrt{2}N}{T}.$$

Next, we define $\mathbf{P}$: $\mathbb{R}^{L\times P} \mapsto \mathbb{R}^{L\times P}$ as a 2-D circular convolution with a unit pulse response whose 2-D discrete Fourier transform (DFT) is 1 on the bow tie and zero elsewhere (see Fig. 4). It follows that $\mathbf{P}$ is a projection operator. Then, we make the following assumption about g

A4) The filtered projection data g satisfies $g = \mathbf{P}h$, where h is a white random process with variance σ_h^2.

Assumption A4) is equivalent to assuming that the sampled projection data, *after* ramp filtering, has a flat spectrum over the bowtie. The validity of this assumption is tested in our simulations in Section V, where we apply our bounds to images (as opposed to noise fields). From Assumption A4, we have

$$\begin{aligned}\mathrm{E}\left\{|e(m,n)|^2\right\} &= \mathrm{E}\left\{\langle \mathbf{O}_{m,n}g,\, \mathbf{O}_{m,n}g\rangle_{\ell_2^P(\mathbb{Z})}\right\}\\ &= \mathrm{E}\left\{\langle \mathbf{O}_{m,n}\mathbf{P}h,\, \mathbf{O}\mathbf{P}h\rangle_{\ell_2^P(\mathbb{Z})}\right\}. \quad (34)\end{aligned}$$

Because h is white, it follows that

$$\mathrm{E}\left\{|e(m,n)|^2\right\} = \sigma_h^2 ||\mathbf{O}\mathbf{P}||_2^2 \quad (35)$$

where the norm on the right hand side of (35) is the standard spectral operator norm

$$||\mathcal{A}||_2 = \sup_{||\mathbf{y}||_2\le 1} ||\mathcal{A}\mathbf{y}||_2. \quad (36)$$

The spectral norm is of particular interest, because it satisfies the following property

$$||\mathcal{A}\mathbf{H}||_2 \le ||\mathcal{A}||_2 ||\mathcal{H}||_2. \quad (37)$$

Furthermore, for any projection operator, such as $\mathbf{P}$, it follows that $||\mathbf{P}||_2 = 1$ [14], so that (35) becomes

$$\sigma_e(m,n) = \sqrt{\mathrm{E}\left\{|e(m,n)|^2\right\}} \le \sigma_h ||\mathbf{O}||_2. \quad (38)$$

We now substitute the definition (32) into (38), and apply (37) to obtain the following upper bound on the backprojection error variance

$$\begin{aligned}\sigma_e(m,n) \le \sigma_h ||\mathbf{Z}'||_2 \cdot ||\mathbf{B}_{P/N,1}||_2 \cdot \prod_{j=J+1}^{\log N} ||\mathbf{E}_j||_2\\ \cdot \left\| \left(\prod_{j=1}^{J} \tilde{\mathbf{E}}_j - \prod_{j=1}^{J} \mathbf{E}_j \right) \right\|_2. \quad (39)\end{aligned}$$

Let us introduce the following definitions:

$$\boldsymbol{\Delta}_j = \tilde{\mathbf{E}}_j - \mathbf{E}_j \quad (40)$$

$$||\boldsymbol{\Delta}|| = \max_{1\le j\le J} ||\boldsymbol{\Delta}_j|| \quad (41)$$

$$\left\|\tilde{\mathbf{E}}\right\| = \max_{1\le j\le J} \max\left(\left\|\tilde{\mathbf{E}}_j\right\|, ||\mathbf{E}_j|| \right). \quad (42)$$

Next, define

$$\gamma(J) = \left\| \prod_{j=1}^{J} \tilde{\mathbf{E}}_j - \prod_{j=1}^{J} \mathbf{E}_j \right\| \quad (43)$$

for $1 \le J \le \log_2 N$. Then

$$\begin{aligned}\gamma(J) &= \left\| \prod_{j=1}^{J} \tilde{\mathbf{E}}_j - \tilde{\mathbf{E}}_J \prod_{j=1}^{J-1} \mathbf{E}_j + \tilde{\mathbf{E}}_J \prod_{j=1}^{J-1} \mathbf{E}_j - \prod_{j=1}^{J} \mathbf{E}_j \right\|\\ &\le \left\|\tilde{\mathbf{E}}_J\right\| \gamma(J-1) + \left\|\tilde{\mathbf{E}}_J - \mathbf{E}_J\right\| \prod_{j=1}^{J-1} ||\mathbf{E}_j||\\ &= \left\|\tilde{\mathbf{E}}_J\right\| \gamma(J-1) + ||\boldsymbol{\Delta}_J|| \prod_{j=1}^{J-1} ||\mathbf{E}_j||. \quad (44)\end{aligned}$$

Equation (44) can be bounded by

$$\gamma(J) \le \left\|\tilde{\mathbf{E}}\right\| \gamma(J-1) + ||\boldsymbol{\Delta}|| \left\|\tilde{\mathbf{E}}\right\|^{J-1}. \quad (45)$$

Recognizing that $\gamma(1) \le ||\boldsymbol{\Delta}||$, it follows from (45) that

$$\gamma(J) \le J||\boldsymbol{\Delta}|| \left\|\tilde{\mathbf{E}}\right\|^{J-1}. \quad (46)$$

Thus

$$\left\| \prod_{j=1}^{J} \tilde{\mathbf{E}}_j - \prod_{j=1}^{J} \mathbf{E}_j \right\| \le J||\boldsymbol{\Delta}|| \left\|\tilde{\mathbf{E}}\right\|^{J-1}. \quad (47)$$

Combining (47) and (42) with (39) yields

$$\sigma_e(m,n) \le J||\mathbf{Z}'||_2 ||\mathbf{B}_{P/N,1}||_2 ||\boldsymbol{\Delta}||_2 \left\|\tilde{\mathbf{E}}\right\|_2^{\log_2 N - 1} \sigma_h. \quad (48)$$

Equation (48) encapsulates the error accumulation in the hierarchical algorithm. To form a computable bound, we need expressions for the worst-case per stage error $||\boldsymbol{\Delta}||$, as well as the worst-case gain $||\tilde{\mathbf{E}}||$.

C. Per-Stage Error

We now turn to the problem of computing the per-stage error, which is a prerequisite for using (48). To calculate $||\boldsymbol{\Delta}_j||_2$ (and thus $||\boldsymbol{\Delta}||$), we expand $\tilde{\mathbf{E}}_j - \mathbf{E}_j$ as

$$||\boldsymbol{\Delta}_j||_2 = \left\| \mathbf{C}_{N/2^j} \left(\tilde{\mathbf{D}}_j \tilde{\mathbf{F}}_{j,i} - \mathbf{D}_j \mathbf{F}_{j,i} \right) \right\|_2 \quad (49)$$

where $\tilde{\mathbf{D}}$ and $\tilde{\mathbf{F}}$ are the computationally efficient approximations to $\mathbf{D}$ and $\mathbf{F}$, respectively. Suppressing the display of the subscripts for notational convenience, we have

$$\begin{aligned}||\boldsymbol{\Delta}_j||_2 &= \left\| \mathbf{C} \left(\tilde{\mathbf{D}}\tilde{\mathbf{F}} - \tilde{\mathbf{D}}\mathbf{F} + \tilde{\mathbf{D}}\mathbf{F} - \mathbf{D}\mathbf{F} \right) \right\|_2\\ &= \left\| \mathbf{C}\tilde{\mathbf{D}} \left(\tilde{\mathbf{F}} - \mathbf{F} \right) + \mathbf{C} \left(\tilde{\mathbf{D}} - \mathbf{D} \right) \mathbf{F} \right\|_2.\end{aligned}$$

Next, we note that because $\tilde{\mathbf{D}}$ involves only angular processing, it does not change the radial bandwidth. Thus, $\mathbf{C}\tilde{\mathbf{D}} = \mathbf{C}\tilde{\mathbf{D}}\tilde{\mathbf{C}}$,

where $\hat{\mathbf{C}}$ is a radial bandlimiter, with radial bandwidth BT. Hence

$$\begin{aligned}||\mathbf{\Delta}_j||_2 &= \left\|\mathbf{C}\tilde{\mathbf{D}}\left(\tilde{\mathbf{F}}-\mathbf{F}\right)+\mathbf{C}\left(\tilde{\mathbf{D}}-\mathbf{D}\right)\mathbf{F}\right\|_2 \\ &= \left\|\mathbf{C}\tilde{\mathbf{D}}\hat{\mathbf{C}}\left(\tilde{\mathbf{F}}-\mathbf{F}\right)+\mathbf{C}\left(\tilde{\mathbf{D}}-\mathbf{D}\right)\mathbf{F}\right\|_2 \\ &\le \left\|\mathbf{C}\tilde{\mathbf{D}}\right\|_2\left\|\hat{\mathbf{C}}\left(\tilde{\mathbf{F}}-\mathbf{F}\right)\right\|_2+\left\|\mathbf{C}\left(\tilde{\mathbf{D}}-\mathbf{D}\right)\right\|_2\end{aligned} \tag{50}$$

where in the last step, we have used $||\mathbf{F}||_2 = 1$. We will write (50) as

$$||\mathbf{\Delta}_j||_2 \le \nu_j^{\tilde{D}}\varepsilon_j^F + \varepsilon_j^D \tag{51}$$

where

$$\nu_j^{\tilde{D}} = \left\|\mathbf{C}_{N/2^j}\tilde{\mathbf{D}}_j\right\|_2 \tag{52}$$

$$\varepsilon_j^F = \max_i \left\|\hat{\mathbf{C}}\left(\tilde{\mathbf{F}}_{j,i}-\mathbf{F}_{j,i}\right)\right\|_2 \tag{53}$$

$$\varepsilon_j^D = \left\|\mathbf{C}_{N/2^j}\left(\tilde{\mathbf{D}}_j-\mathbf{D}_j\right)\right\|_2. \tag{54}$$

The constants $\nu_j^{\tilde{D}}$, ε_j^F and ε_j^D fully characterize the error introduced in each stage of the SFHBP. The constant $\nu_j^{\tilde{D}}$ can be interpreted as the gain of the angular decimator, while ε_j^D is the relative error of the angular decimator. The constant ε_j^F is the relative error of the radial shifter. We now turn to bounding each of these constants so as to better understand the effect of parameter choice on the algorithm performance.

1) Fractional Delay Relative Error: To bound ε_j^F, we first recall that the operator $\mathbf{F}_{j,i}$ is a fractional delay operator, where each projection undergoes a shift δ_p which depends upon j and i. We can determine the worst-case performance by looking at a one dimensional FD problem, and taking the worst case shift. Thus

$$\varepsilon_j^F \le \varepsilon^F \tag{55}$$

where ε^F is simply the worst-case one dimensional error for a fractional delay. To compute ε^F, we first note that from (18) it follows that our approximate FD applied to a signal $g(n)$ has the form of a standard FIR filtering operation

$$\tilde{h}(n) = \sum_k g(k)\tilde{q}_\delta(n-k) \tag{56}$$

where $\tilde{q}_\delta(k) \in \ell_2$ is an approximation to the ideal FD filter, and depends on a shift $\delta \in [-0.5, 0.5]$. From (53), we would like to upper bound $||\hat{\mathbf{C}}(\tilde{\mathbf{F}}_{j,i}-\mathbf{F}_{j,i})||$. But $\hat{\mathbf{C}}$, $\tilde{\mathbf{F}}_{j,i}$ and $\mathbf{F}_{j,i}$ are all shift invariant (convolutional) operators, that act on each projection by filtering. Thus, they commute, and

$$\varepsilon_J^F = \max_i \left\|\left(\tilde{\mathbf{F}}_{j,i}-\mathbf{F}_{j,i}\right)\hat{\mathbf{C}}\right\|. \tag{57}$$

Equivalently, we need to find a bound on the approximation error $||\tilde{\mathbf{F}}_{j,i}-\mathbf{F}_{j,i}||$ for signals $g \in \ell_2^P(\mathbb{Z})$ that are radially bandlimited, with radial bandwidth BT. The ideal $h(n)$ is thus given by

$$h(n) = \sum_k g(k)q_\delta(n-k) \tag{58}$$

where $q_\delta(k)$ is the unit pulse response of a filter with transfer function $Q_\delta(\omega) = e^{-j\omega\delta}$ for $\omega \in [-BT, BT]$, and BT is the radial bandwidth of g.

The error in using $\tilde{q}$ instead of q is then simply

$$h(n) - \tilde{h}(n) = \sum_k g(k)\left(q_\delta(n-k) - \tilde{q}_\delta(n-k)\right). \tag{59}$$

Taking the discrete time Fourier transform (DTFT) of both sides, yields

$$\begin{aligned}H(\omega) - \tilde{H}(\omega) &= G(\omega)\left(Q_\delta(\omega) - \tilde{Q}_\delta(\omega)\right) \\ &= G(\omega)\left(e^{j\omega\delta} - \tilde{Q}_\delta(\omega)\right).\end{aligned} \tag{60}$$

Squaring and integrating over $[-\pi, \pi]$ yields

$$\begin{aligned}&\frac{1}{2\pi}\int_{|\omega|\le BT}\left|H(\omega)-\tilde{H}(\omega)\right|^2 d\omega \\ &\quad\le \max_{|\omega|\le BT}\left|e^{j\omega\delta}-\tilde{Q}_\delta(\omega)\right|^2 \frac{1}{2\pi}\int_{-\pi}^{\pi}|G(\omega)|^2\,d\omega.\end{aligned} \tag{61}$$

Hence, by Parseval's Theorem

$$\left\|h-\tilde{h}\right\|_2 \le \sup_\delta \max_{|\omega|\le BT}\left|e^{j\omega\delta}-\tilde{Q}_\delta(\omega)\right| ||g||_2. \tag{62}$$

Thus, we obtain the following upper bound for ε_j^F

$$\varepsilon_j^F \le \varepsilon^F \le \sup_\delta \max_{|\omega|\le BT}\left|e^{j\omega\delta}-\tilde{Q}_\delta(\omega)\right| \tag{63}$$

where $\tilde{Q}_\delta$ is the DTFT of the FIR approximation to a FD with delay δ. Furthermore, from (63), it follows that ε^F is simply the worst-case peak ripple in the passband error of $\tilde{Q}_\delta$ with respect to δ.

Remarks: 1) Although in principle the FD filters Q_δ are all derived from ϕ via (18), we will permit the additional freedom of designing the filters $\tilde{q}$ independent of the choice of ϕ. We do this because ε^F depends on the minimax error, which can be optimized for the discrete filters. 2) Different filter orders can be used for different δ, to minimize the right hand side of (63).

2) Angular Decimation Gain and Relative Error: Next, we turn to the angular processing in the form of $\mathbf{D}$ and $\tilde{\mathbf{D}}$. Substituting an approximate filter $\tilde{\psi}$ into (15), we have

$$\tilde{\mathbf{D}}_j g(k,p) = \sum_{m=0}^{P-1} g(k,m)\tilde{\psi}(2p-m). \tag{64}$$

Recall that g is periodic with respect to its second argument, with period P. We will not assume that $\tilde{\psi}$ is periodic, although it is a standard DSP result that convolution of a periodic sequence with an aperiodic filter can be mapped to the convolution of two periodic sequences with the same period [15].

Thus, $\tilde{\mathbf{D}}$ consists of the following steps: 1) convolution of $g(k,\cdot)$ with $\tilde{\psi}$ for each k, downsampling of the result by a factor of two in the second argument. The approximation error of $\tilde{\mathbf{D}}$ to $\mathbf{D}$ clearly does not depend on the first index k. So we consider instead a one-dimensional (1-D) problem in which a periodic signal g with period P is convolved with a nonideal low-pass filter prior to downsampling. Furthermore, from (52), (54), and Fig. 1, it follows that the output is postprocessed with $\mathbf{C}$, which bandlimits the result.

Let $\tilde{\Psi}(\omega)$ denote the DTFT of $\tilde{\psi}$. Let h denote the decimated convolution of g with $\tilde{\psi}$, i.e.,

$$h(n)=\sum_{k=0}^{P-1} g(k)\tilde{\psi}(2n-k). \tag{65}$$

Clearly, $h(n)$ is periodic with period $P/2$. Let $H(l)$ denote the $P/2$-point DFT of one period of $h(n)$, and let $G(l)$ denote the P-point DFT of one period of G. Then (65) can be rewritten in the Fourier domain (after accounting for $\mathbf{C}$) as

$$H(l)=\begin{cases}\dfrac{1}{2}\Big[G(l)\tilde{\Psi}\Big(\dfrac{2\pi l}{P}\Big)+G(l+P/2) \\ \qquad\cdot\tilde{\Psi}\Big(\dfrac{2\pi(l+P/2)}{P}\Big)\Big] & |l|\le NB\sqrt{2}/4 \\ 0, & \text{else.}\end{cases} \tag{66}$$

Now, taking the squared magnitude of both sides of (66) and applying the Cauchy–Schwartz inequality to the right hand side yields

$$|H(l)|^2\le\tfrac{1}{4}\left(|G(l)|^2+|G(l+P/2)|^2\right)\cdot\left[\left|\tilde{\Psi}\left(\frac{2\pi l}{P}\right)\right|^2+\left|\tilde{\Psi}\left(\frac{2\pi(l+P/2)}{P}\right)\right|^2\right] \tag{67}$$

for $|l|\le NB\sqrt{2}/4$.

Replacing the term in square brackets by its maximum, summing over l, and applying Parseval's Theorem yields

$$\|h\|\le\frac{1}{\sqrt{2}}\max_{|l|\le NB\sqrt{2}/4}\cdot\sqrt{\left|\tilde{\Psi}\left(\frac{2\pi l}{P}\right)\right|^2+\left|\tilde{\Psi}\left(\frac{2\pi(l+P/2)}{P}\right)\right|^2}\cdot\|g\|. \tag{68}$$

From this expression, it follows that

$$\nu_j^{\tilde{D}}\le\frac{1}{\sqrt{2}}\max_{|l|\le NB\sqrt{2}/2^{j+1}}\cdot\sqrt{\left|\tilde{\Psi}\left(\frac{2\pi l}{P/2^{j-1}}\right)\right|^2+\left|\tilde{\Psi}\left(\frac{2\pi(l+P/2^j)}{P/2^{j-1}}\right)\right|^2}. \tag{69}$$

Using similar manipulations, we also arrive at the following upper bound for ε_j^D:

$$\varepsilon_j^D\le\frac{1}{\sqrt{2}}\max_{|l|\le NB\sqrt{2}/2^{j+1}}\cdot\sqrt{\left|\tilde{\Psi}\left(\frac{2\pi l}{P/2^{j-1}}\right)-2\right|^2+\left|\tilde{\Psi}\left(\frac{2\pi(l+P/2^j)}{P/2^{j-1}}\right)\right|^2}. \tag{70}$$

We can further simplify (69) and (70) by replacing the j-dependent expressions with the following bounds:

$$\nu_j^{\tilde{D}}\le\nu^{\tilde{D}}\le\frac{1}{\sqrt{2}}\max_{|\omega|\le\pi NB\sqrt{2}/(2P)}\sqrt{\left|\tilde{\Psi}(\omega)\right|^2+\left|\tilde{\Psi}(\omega+\pi)\right|^2} \tag{71}$$

$$\varepsilon_j^D\le\varepsilon^D\le\frac{1}{\sqrt{2}}\max_{|\omega|\le\pi NB\sqrt{2}/(2P)}\sqrt{\left|\tilde{\Psi}(\omega)-2\right|^2+\left|\tilde{\Psi}(\omega+\pi)\right|^2}. \tag{72}$$

An intuitive interpretation for $\nu^{\tilde{D}}$ and ε^D can be found in (71) and (72), respectively. Specifically, $\nu^{\tilde{D}}$ is roughly the peak in-band magnitude plus the peak out of band ripple of the $\tilde{\psi}$ filter. Similarly, ε^D is roughly the sum of the peak in-band ripple plus the peak out of band ripple for the $\tilde{\psi}-\psi$ filter. Thus, if $\tilde{\psi}=\psi$, then $\nu^{\tilde{D}}=\sqrt{2}$, and $\varepsilon^D=0$. Note that (71) and (72) depend on the amount of angular oversampling. In particular, because of Assumption A2), it follows that

$$\frac{NB\sqrt{2}}{P}<1$$

which in turn affects the set of frequencies over which (71) and (72) are computed.

3) Level-Independent and Simplified Bounds: At this point, we have obtained computable expressions for the constants ε_j^F, ε_j^D and $\nu_j^{\tilde{D}}$ in terms of the Fourier transforms of the various filters involved. We have also derived bounds ε^F, ε^D and $\nu^{\tilde{D}}$ on these constants that do not depend on the level j. Thus, (51) becomes

$$\|\mathbf{\Delta}_j\|\le\nu^{\tilde{D}}\varepsilon^F+\varepsilon^D \tag{73}$$

which together with (41) implies that we can take $\|\mathbf{\Delta}\|=\nu^{\tilde{D}}\varepsilon^F+\varepsilon^D$. Next, we make the following simplification. Recall that $\nu^{\tilde{D}}$ is equivalent to the peak in-band magnitude plus a small ripple term. To first order, $\nu^{\tilde{D}}\approx\sqrt{2}=\|\mathbf{D}_j\|_2$. Thus

$$\|\mathbf{\Delta}\|=\sqrt{2}\varepsilon^F+\varepsilon^D. \tag{74}$$

Furthermore

$$\left\|\tilde{\mathbf{E}}_j\right\|_2\le\left\|\mathbf{C}_{N/2^j}\tilde{\mathbf{D}}_j\right\|_2\|\mathbf{F}_j\|_2=\nu_j^{\tilde{D}}\|\mathbf{F}_j\|_2. \tag{75}$$

It is simple to show $||\mathbf{F}_j||_2$ is the peak magnitude in the passband of the fractional delay filters. Thus, a good approximation is $||\mathbf{F}_j|| \approx 1$, so that $||\tilde{\mathbf{E}}_j||_2 \leq \nu_j^{\tilde{D}} = \sqrt{2}$, and

$$\left\|\tilde{\mathbf{E}}\right\|_2 = \sqrt{2}. \tag{76}$$

Combining (74), (76) and (48), and noting that $||\mathbf{Z}'||_2 = 1$, yields the following bound:

$$\sigma_e(m, n) \leq J||\mathbf{B}_{P/N,1}||_2 \left(\sqrt{2}\varepsilon^F + \varepsilon^D\right)\sqrt{\frac{N}{2}}\,\sigma_h. \tag{77}$$

Equation (77) can be further simplified by writing $||\mathbf{B}_{P/N,1}||$ explicitly. From (7), it follows that

$$||\mathbf{B}_{P/N,1}||_2 = \sqrt{\sum_k \sum_p \eta^2_{0,0,k,p}}. \tag{78}$$

Consider a circularly symmetric image sampling kernel b [see (4) or (8)] with a 2-D Fourier transform $\check{b}(\omega)$, where ω is the radial frequency. Then it can be shown by means of Parseval's Theorem and the Fourier Slice Theorem (see [10]), that subject to Assumption A1)

$$||\mathbf{B}_{P/N,1}||_2 = \sqrt{\frac{P}{N\pi T}}\left(\int_0^B \left|\check{b}(\omega)\right|^2 d\omega\right)^{1/2}. \tag{79}$$

For the special case that b is the indicator function for a disk of radius 1/2 (which is the choice we use in our implementation in Section V), the right hand side of (79) can be expressed in terms of hypergeometric functions [16]. For simplicity, we prefer to bound (79) by

$$\begin{aligned}||\mathbf{B}_{P/N,1}||_2 &\leq \sqrt{\frac{P}{N\pi T}}\left(\int_0^\infty \left|\check{b}(\omega)\right|^2 d\omega\right)^{1/2} \\ &= \sqrt{\frac{2P}{3NT}}.\end{aligned} \tag{80}$$

Substituting (80) into (77) yields

$$\sigma_e(m, n) \leq J\sqrt{\frac{P}{3T}}\left(\sqrt{2}\varepsilon^F + \varepsilon^D\right)\sigma_h. \tag{81}$$

Equation (81) is simple in form, and bounds the error variance for the (m, n)th pixel. But the right hand side does not depend on (m, n). Thus, we can replace (81) by the following uniform bound

$$\max_{-N/2 \leq m,\, n < N/2} \sigma_e(m, n) \leq J\sqrt{\frac{P}{3T}}\left(\sqrt{2}\varepsilon^F + \varepsilon^D\right)\sigma_h. \tag{82}$$

Equation (82) serves as the basis for our analysis of the parameter choices on the backprojection performance. In addition to serving as a useful design tool, (82) also provides us with confidence intervals for errors in the reconstruction. Indeed, from (32), it follows that each error term $e(m, n)$ is a weighted sum of random variables. Furthermore, this weighted sum involves at least P terms with roughly equal weights. Hence, we can appeal to the Central Limit Theorem (CLT) to argue that the errors $e(m, n)$ should have an asymptotic Gaussian distribution [17]. For large P, we can then construct confidence intervals based on (82). For example, we can expect that over 99% of the errors in the (m, n)th pixel (over an ensemble of sinograms) will lie in the interval $[-3c, 3c]$, where $c = J\sqrt{P/3T}(\sqrt{2}\varepsilon^F + \varepsilon^D)\sigma_h$. We will also return to (82) as a possible means of selecting an operating point for the algorithm.

Note that bound (82) increases with P. This would appear to suggest that the error increases as the number of views is increased. But our definition of the backprojection operation [see, e.g., (1)] is unweighted by $1/P$, so that the signal level in the backprojection also increases with P. When the error is normalized by standard deviation of the signal (to form a signal to noise ratio), the dependence on $\sqrt{P}$ cancels out.

IV. Discussion of Bounds

Prior to an experimental study of the bounds we have derived, it is interesting to note that we can already draw some conclusions from the forms of the bounds. In this section, we will state those conclusions, which will then be tested via simulations in the next section.

A. Choice of J

The first point of interest is the choice of the parameter J. Note that (82) depends *linearly* on J, so that doubling J (or equivalently halving the size of the smallest backprojection that is decomposed) doubles the approximation error. On the other hand, the cost expression (22) indicates that the cost decreases roughly exponentially fast with increased J. Hence, the bounds suggest that J should be chosen as large as possible (i.e., $J = \log_2 N$), with the errors controlled via the radial and angular oversampling and filter lengths.

B. Radial and Angular Oversampling and Filter Lengths

Another conclusion that can be drawn from (81) is that the error bound depends on ε^F and ε^D in an additive fashion. This means that the approximations $\tilde{\phi}$ and $\tilde{\psi}$ to the ideal radial and angular filters should be chosen so that they contribute roughly equal errors to the output. Put another way, the bound depends on $\sqrt{2}\varepsilon^F + \varepsilon^D$. Thus, if ε^F is made very small (by taking long FD filters) relative to ε^D, then the overall accuracy of the algorithm becomes dominated by ε^D.

Next, we turn to the question of radial and angular oversampling. Let us assume that the input data is sampled radially and angularly at the relevant Nyquist rates. From (63), it follows that the radial component of the error is the ripple in the FIR approximation to an allpass filter with a phase response of $e^{j\omega/2}$, and the angular component of the error is determined by the passband and stopband ripple in a short approximation to an ideal angular low pass filter with cutoff $\pi/2$ and no transition band. It is well known that both of these errors will be large, because the desired frequency response is discontinuous [15]. Thus, it follows that some radial and angular oversampling are required for the errors to be made small.

We can draw some further conclusions for the angular filtering case by employing the well known empirical relationship between filter order and approximation error for minimax-optimal lowpass filters (see [15, p. 480]). It has been demonstrated

that for filters with a ripple of μ in the passband and stopband, the filter order necessary to achieve that ripple is

$$M_a \propto \frac{\log \mu}{\Delta\omega} \tag{83}$$

where $\Delta\omega$ is the width of the transition band. If the original projection data were angularly sampled at the Nyquist rate $P > NB\sqrt{2}$, then upsampling angularly by a factor of $K_a = P/NB\sqrt{2}$ means that the width of the transition band is given by

$$\Delta\omega = \frac{\pi(K_a - 1)}{K_a}.$$

Thus, noting that ε^D depends directly on μ, the relationship between the approximation error ε^D, the filter order M, and the oversampling factor of K_a, is given by

$$\varepsilon^D \approx e^{-(\beta_a M_a (K_a - 1)/K_a)}. \tag{84}$$

From (84), it follows that the approximation error decays fairly slowly with increasing oversampling but fast with respect to filter order M_a. Similar conclusions can be drawn starting from Lagrange and the so-called central sampling formulae (see [18] and [19], respectively).

For the radial processing, results on FDs can be found in [20], [21], which suggest a similar relationship between filter order, and approximation error and bandwidth as in the angular case. In particular, for a passband ripple of μ, and a "transition band" of $\Delta\omega$, the filter order M_r satisfies (83). For a signal upsampled radially by a factor of $K_r = \pi/BT$, we have that

$$\Delta\omega = \frac{\pi(K_r - 1)}{K_r}.$$

Thus, we conclude that

$$\varepsilon^F \approx e^{-(\beta_r M_r (K_r - 1)/K_r)}. \tag{85}$$

Again, the error bounds decrease exponentially fast in the filter order but rather slowly in the amount of oversampling. Equations (84) and (85) can be substituted into (82) to yield the following expression:

$$\begin{aligned}\max_{m,n} \sigma_e(m, n) &\le J\sqrt{K_a K_r}\sqrt{\frac{P}{3T}}\left(\sqrt{2}e^{-(\beta_r M_r (K_r - 1)/K_r)}\right. \\ &\qquad \left. + e^{-(\beta_a M_a (K_a - 1)/K_a)}\right)\sigma_h \end{aligned} \tag{86}$$

where we have also substituted $P \to K_a P$, $T \to T/K_r$. Furthermore, Ξ can also be modified to account for the additional cost of applying the SFHBP algorithm to upsampled projections (as well as using potentially longer filters in the radial and angular filtering steps). The resulting expression is

$$\begin{aligned}\Xi = \frac{4K_a K_r P}{T}&\left(J\left(M_r + R_b + \left\lceil N/\sqrt{2}\right\rceil\right)\right. \\ &\left. \cdot (M_a + 2M_r) + \frac{N^2(R_\phi + R_b)}{2^J}\right). \end{aligned} \tag{87}$$

TABLE II
PARAMETER CHOICES COMMON TO THE VARIOUS SIMULATIONS

Parameter & Symbol	Value
Projection interpolation kernel (ϕ)	$\operatorname{sinc}(t/T)\cos\left(\frac{\pi t}{10T}\right)$ for $\lvert t\rvert < 5T$
Radius of support of ϕ (R_ϕ)	5
Radial sampling interval (T)	1.0
Image discretization kernel (b)	$b(x, y) = 1$ for $x^2 + y^2 \le \frac{1}{4}$
Radius of support of b (R_b)	1/2
Radial Bandwidth (B)	π

C. Design Considerations

The bounds suggest the following scheme for parameter selection:

1) choose J as large as possible (i.e., decompose to the single pixel level);
2) choose $\tilde{\phi}$ and $\tilde{\psi}$ so that $\sqrt{2}\varepsilon^F$ and ε^D contribute equally to the total error;
3) keep oversampling (though necessary both angularly and radially) to a minimum.

In principle, with (86) and (87), the problem of parameter selection could be reduced to an optimization problem, in which the cost Ξ is fixed, and the error is minimized, or the error is fixed and the cost Ξ is minimized. The resulting problem is a constrained integer programming problem that can be solved by exhaustive or other techniques [22]. This automated approach to parameter selection, while promising, is beyond the scope of this paper. We will instead focus on simple "rule of thumb" results that can be found through numerical simulations.

V. SIMULATIONS

In this section, we present some numerical simulations which demonstrate the usefulness of the bounds in characterizing performance of the SFHBP algorithm. Our goal is to demonstrate that the bounds correctly predict the various trends of algorithm performance as a function of parameter choice so that they can be used for design purposes. In particular, we will verify and refine the conclusions as to parameter selection that we discussed in the previous section. These experiments also serve to verify that the various approximations made in the derivation of the bounds are valid. In the course of the various simulations, some of our parameter choices were common throughout. These parameters, and the related constants are summarized in Table II.

A. Error and Bound Evaluation

To evaluate the error, in principle, we need to compute backprojections of an ensemble of images using the direct and SFHBP algorithms, and then study the error variance at each pixel and compare it to the upper bound. In practice, however, we instead approximate the error field $(\mathbf{B} - \tilde{\mathbf{B}})g$ as an ergodic random process and compare the rms error

$$\sigma_e \approx \frac{\left\|\mathbf{B}g - \tilde{\mathbf{B}}g\right\|_2}{N} \tag{88}$$

to the upper bound (86). Thus we compare the spatially averaged error variance to the bound instead of computing an average over an ensemble of images. Furthermore, we scale both the estimate

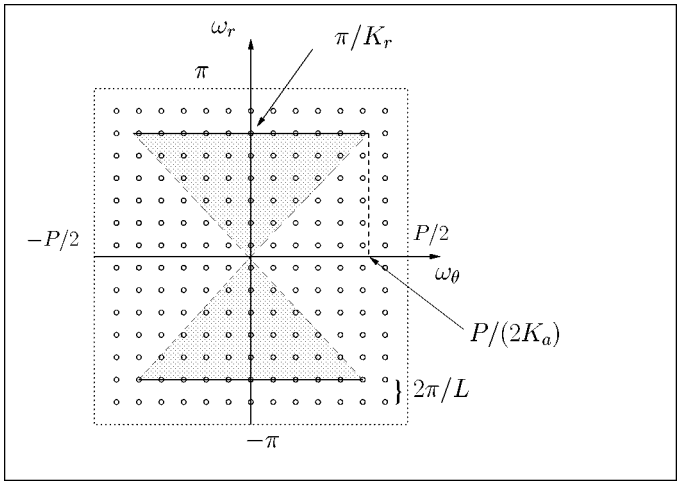

Fig. 4. For the circulant–block–circulant bow tie projection operator, the rank of the matrix $\mathbf{P}_0$ is determined by the number of samples that fall in the bow tie region (shaded).

and the bound by the rms backprojected value, so that errors can be expressed in percentages.

To compute (86), we also need to estimate σ_h from g. Let σ_g^2 denote the average signal power in the sinogram

$$\sigma_g^2 = \frac{\|g\|_2^2}{S} \tag{89}$$

where S is the number of points in the support of $\mathbf{g}$. From the law of large numbers and Fig. 4, it can be shown [12] that

$$\sigma_g^2 \approx \frac{1}{2K_aK_r}\,\sigma_h^2 \tag{90}$$

or equivalently

$$\sigma_h \approx \|g\|_2\sqrt{\frac{2K_aK_r}{S}}. \tag{91}$$

We can thus relate σ_h to the measured ℓ_2 norm of the projection data. The last step is to incorporate this estimate into the bound. Combining (88) with (91) and (82) yields the following bound:

$$\frac{\left\|\mathbf{B}g - \tilde{\mathbf{B}}g\right\|_2}{\|\mathbf{B}g\|_2} \le J\,\frac{\|g\|_2}{\|\mathbf{B}g\|_2}\sqrt{\frac{2PN^2K_aK_r}{3ST}}\left(\sqrt{2}\varepsilon^F + \varepsilon^D\right). \tag{92}$$

B. Filter Order and Oversampling

Our first goal was to determine an empirical rule relating the radial and angular filter orders. To that end, we performed a series of experiments in which we fixed the bandwidth B, angular and radial oversampling (K_a and K_r respectively), and designed the relevant minimax-optimal angular and radial filters. The angular filter was designed using the standard Parks–McClellan algorithm [15]. The radial filter was a FD filter with delay 1/2 designed using the extended Parks–McClellan algorithm developed in [23]. More details on the problem of minimax-optimal design of FD filters can be found in [20], [21]. We then calculated $\sqrt{2}\varepsilon^F$ and ε^D for filters of different length and tabulated them. The results of a typical such experiment are shown in Fig. 5, which indicates the results for $K_a = K_r = 10/9$.

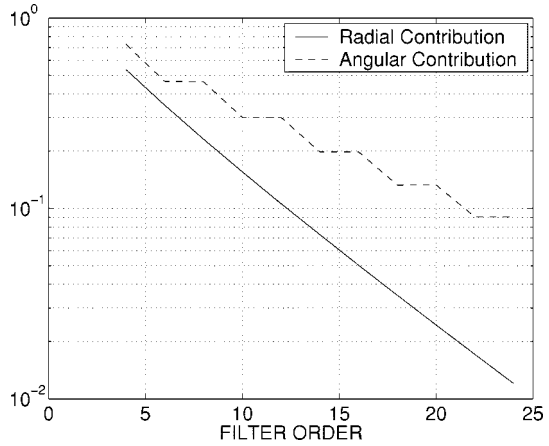

Fig. 5. Contributions of the radial and angular processing to the error bound as a function of filter order.

From Fig. 5, two features are immediately noticeable. The first feature is that ε^D decreases in "steps," so that angular filters of order $4m-2$ and $4m$ yield the same error. This behavior is a result of the fact that $\tilde{\psi}(n)$ is a "half band" filter [24], and that (except for $n=0$) has zero taps for all even values of n. Furthermore, convolution with these half band filters requires half the computational effort that we assumed in the derivation of Ξ.

The second feature of Fig. 5 is that the decay rates for the radial and angular errors satisfy $\beta_r \approx 2\beta_a \approx 2.4$, so that for a given error, the angular filter needs to be twice as long as the radial filter to contribute an equal amount of error to the bound. The most likely explanation for this phenomenon is that the transition bands for the radial filter are of the form $\pm[(1-\alpha)\pi, \pi]$, while the transition bands for the angular filter are of the form $\pm[\alpha\pi/2, \pi - \alpha\pi/2]$. The transition bands for the radial filter, when considered on $[0, 2\pi]$, form a larger transition band of twice the width, while the same is not true for the angular filter (the two transition bands are separated by a stop band). We conjecture that this is the reason for the difference in decay rates. Furthermore, this difference motivates us to choose

$$M_a = 2M_r. \tag{93}$$

C. Effective Signal-to-Noise Ratios and Optimal Tradeoffs

At this point, we can make some predictions about the tradeoff between oversampling and filter orders. An understanding of those tradeoffs is facilitated by rephrasing the bounds in terms of an effective signal-to-noise ratio (SNR). Consider the approximate algorithm with the exact sinogram data g, and the direct algorithm $\mathbf{B}_{P,N}$ with noisy data $g'(k,p) = g(k,p) + w(k,p)$, where $w(k,p)$ is a white noise process with variance σ_w^2. The noise $w(k,p)$ will contribute an error variance of $\sigma_f^2(m,n)$ to $\mathbf{B}_{P,N}g'(m,n)$. Recall that $\sigma_\varepsilon^2(m,n)$ is the variance of $(\mathbf{B}_{P,N} - \tilde{\mathbf{B}}_{P,N})g(m,n)$, i.e., the per-pixel error variance due to the approximations.

The effective SNR is defined as the SNR in the measurements g' such that $\sigma_f^2(m, n) = \sigma_e^2(m, n)$ for some $\alpha < 1$. At this effective SNR, errors introduced by the fast algorithm have variance equal to the error variance due to measurement noise. For example, if the true measurement SNR is 3 dB *below* the effective SNR, then the errors due to the FHBP contribute no more than an additional 50% in error variance to the backprojected variance.

To compute the effective SNR, we first calculate σ_f^2. It can be shown that

$$\sigma_f^2 = ||\mathbf{B}_{P,1}||_2^2 \sigma_w^2. \tag{94}$$

From (80), it follows that for the circular pixel choice of b, (94) is[2]

$$\sigma_f^2 = \frac{2P}{3T}\sigma_w^2. \tag{95}$$

Now, from (82), we know that the fast backprojection contributes a variance of

$$\sigma_e^2 \leq J^2 \frac{P}{3T}\left(\sqrt{2}\varepsilon^F + \varepsilon^D\right)^2 \sigma_h^2. \tag{96}$$

If we use the relationship between σ_g and σ_h from (90), we can rewrite (96) as

$$\sigma_e^2 < J^2 \frac{2PK_aK_r}{3T}\left(\sqrt{2}\varepsilon^F + \varepsilon^D\right)^2 \sigma_g^2. \tag{97}$$

Suppose we look at the case $\sigma_e^2 = \sigma_f^2$. Then from (97) and (95), we find that

$$\sigma_w^2 = J^2 K_a K_r \left(\sqrt{2}\varepsilon^F + \varepsilon^D\right)^2 \sigma_g^2. \tag{98}$$

We can also divide σ_g^2 by σ_w^2 to obtain an expression for the effective SNR

$$\frac{\sigma_g^2}{\sigma_w^2} = \frac{1}{K_aK_rJ^2\left(\sqrt{2}\varepsilon^F + \varepsilon^D\right)^2}. \tag{99}$$

When we choose the algorithm parameters, we will trade this effective SNR for algorithm speed. To measure the speed, we calculated the speedup using (87) and the parameters in Table II. The resulting speedup for the case $J = \log_2 N$ is

$$\frac{15N}{K_aK_r\left(M\log_2 N\sqrt{2}+3\right)}. \tag{100}$$

We can now plot the effective SNR against the speedup, and obtain an intuitively meaningful assessment of the bounds. Fig. 6 shows plots of the effective SNR (in decibels) versus the speedup for $K = K_a = K_r \in \{1.25, 1.5, 2.0\}$, and $M \in \{4, 6, \ldots, 20\}$. Smaller M lead to larger speedups for a fixed amount of oversampling. Calculations were performed assuming $N = 2^{10}$.

From Fig. 6, we can immediately determine optimal operating points for the fast algorithm, given an estimate of the measurement noise. For example, if the projections are corrupted by noise to an SNR of 40 dB, then we can operate the algorithm so that the effective SNR of the fast algorithm is 43 dB. This choice guarantees that the errors in the backprojection contribute half as much power to the errors in the final image as the errors in the measurements themselves. Furthermore, from Fig. 6, we can determine that the optimal (in terms of maximizing speedup) choice is $K = 1.5, M = 12$, with a resulting speedup of almost $50\times$.

[2]Equation (80) remains valid when P/N is replaced by P

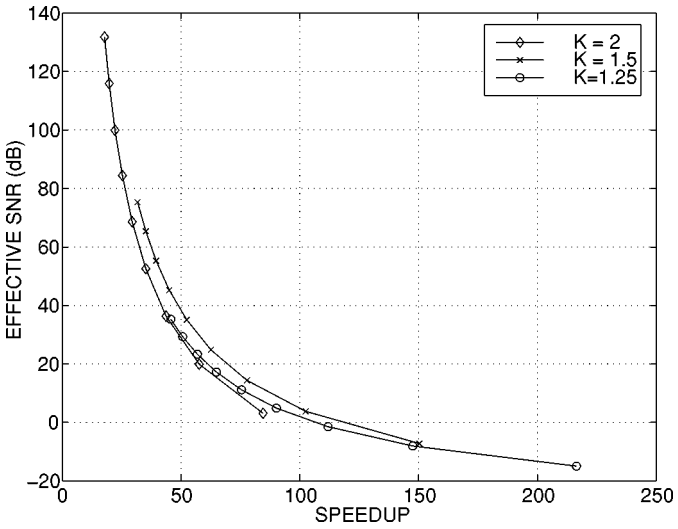

Fig. 6. Normalized speedup versus error dependance for the bounds as a function of filter order M (increasing speedup for smaller M) and oversampling factor K.

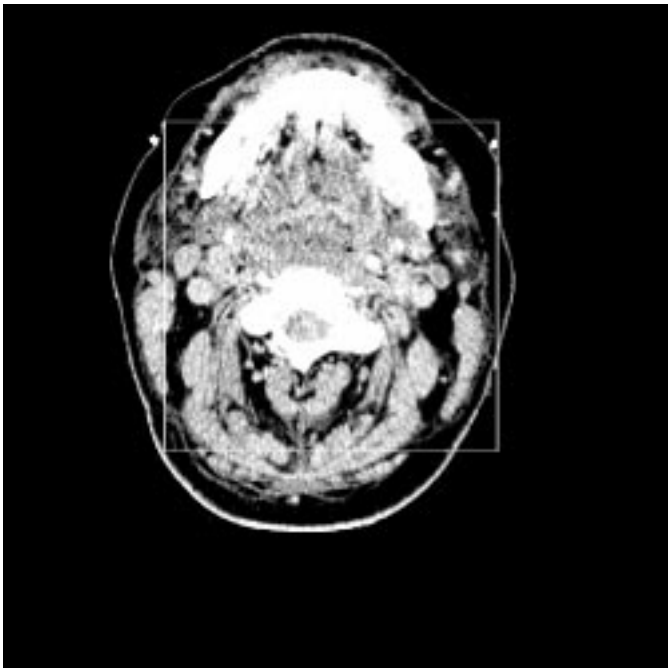

Fig. 7. Phantom used for the $N = 512$ pixel experiments. For the $N = 256$ pixel experiments, the subimage indicated by the white rectangle was used.

D. Bound Verification

Our second goal in this section is to verify the usefulness of the bounds in predicting the behavior of the SFHBP algorithm with respect to parameter choices. We were unable to obtain data from a CT scanner for testing the bounds, so we elected to generate synthetic projections from a CT reconstruction in the visual human dataset (VHD) [25]. Specifically, we used slice 1194 from the Visual Female dataset, which is shown in Fig. 7 windowed to [0, 100] Hounsfield units for display purposes only. Note that the slice is fairly challenging to reconstruct accurately

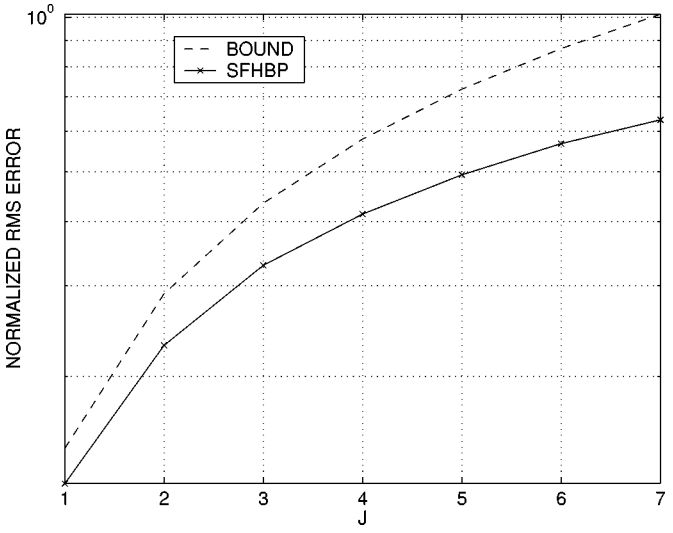

Fig. 8. SFHBP performance versus the bound for various choices of J.

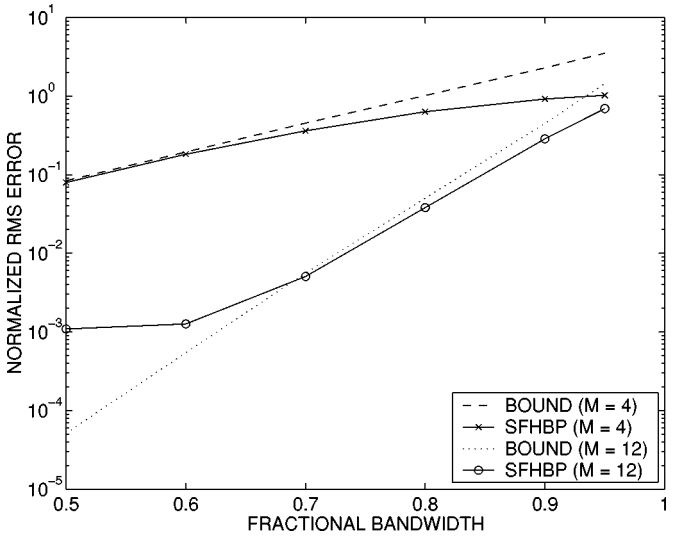

Fig. 9. SFHBP performance versus the bound for various choices of the fractional radial bandwidth.

due to the high contrast from the high density bone regions and the structure in the surrounding soft tissue. The original scan was 512×512 pixels in size. For our experiments, we used the 256×256 subimage of the phantom indicated by the rectangular region in Fig. 7.

The first experiment was to verify the behavior of the cost and error bounds with respect J. To that end, we took the 256×256 phantom subimage, and computed $P = 1024$ projections using a discrete approximation to the Radon transform with $T = 1$, found by taking the adjoint of (1). The projections were then ramp filtered with the Ram–Lak filter [26] with a cutoff at $BT = 0.8\pi$. This cutoff is equivalent to $K_r = 1.25$ and $K_a = 1.13$. We then computed an exact backprojection via (1), and a series of fast backprojections using the SFHBP algorithm with different choices for the parameter J. The radial filters were fourth-order minimax-optimal fractional delays, that were computed using the techniques described in [21]. The angular filter was a sixth-order lowpass filter designed using the **remez** command in MATLAB [27]. The error bound and the actual algorithm errors are compared in Fig. 8. The agreement between the bound and the algorithm performance is good, suggesting that the derived expressions for the error bounds and cost accurately model the effects of choosing J.

The second experiment was to verify the behavior of the error bounds as a function of the oversampling factor $K_a = K_r = K$. The same phantom was used as in the previous experiment, with the same number of projections. In order to simulate the oversampling process without implementing noninteger oversampling, we simply bandlimited the projections to $((K-1)/K)\pi$ in the Ram–Lak filter. This serves to simulate the effect of radial and angular oversampling on the error behavior of the SFHBP algorithm.[3] The filters were designed optimally for each choice of $(K-1)/K \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$. For each K, both an exact backprojection, and a fast backprojection (with $J = 7$) were computed. The error and bound were computed via (92). The results are shown in Fig. 9. Two different choices for radial filter length are shown. For the shorter filter $M = 4$, the bounds are accurate for the range of K. For the longer $M = 12$ filter, the bounds fail for errors smaller than 0.1%. This suggests that the failure of the bounds is due to errors in the "exact" backprojection. In particular, our exact backprojection uses a radial filter ϕ (see Table II) that is not ideal. Thus, errors are present in the exact backprojection which dominate the errors of the fast algorithm for long filter lengths or large amounts of oversampling.

[3]Although this simulated oversampling does not allow us verify the effects of oversampling on Ξ, these are straightforward.

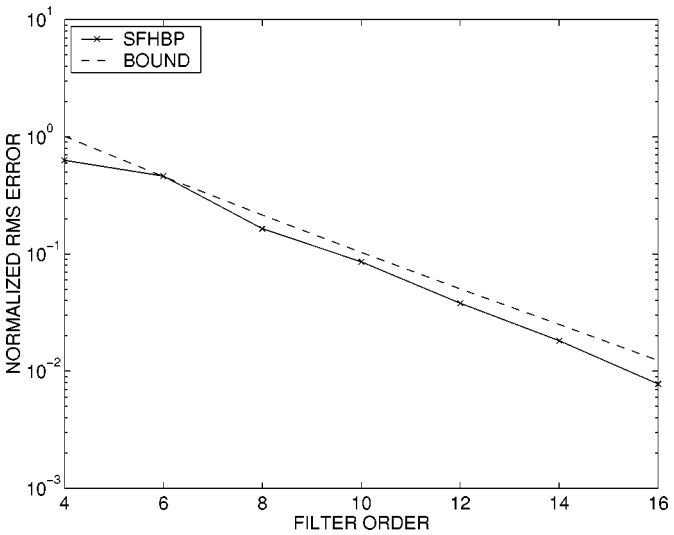

Fig. 10. SFHBP performance versus the bound for various choices of the filter order M.

The third experiment was designed to verify the predicted cost and error behaviors as a function of the filter order $M_a = 2M_r = M$. The phantom and P were the same as in the previous experiment. The cutoff of the ramp filter was set to 0.8π, simulating a radial oversampling factor of $K_r = 1.25$ and an angular oversampling factor of $K_a = 1.13$ in the SFHBP. The filter order M was then varied over the range $M \in \{4, 6, 8, \ldots, 16\}$, and the normalized error calculated. The result is shown, with a comparison to (92), in Fig. 10. The agreement between the bound and the measured performance

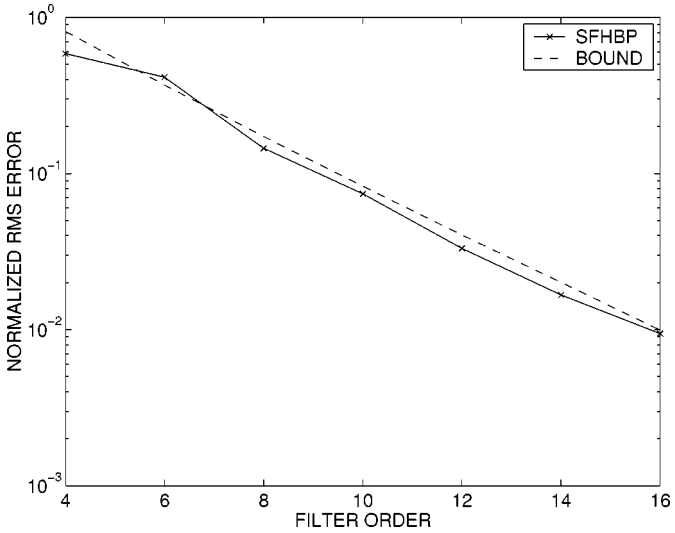

Fig. 11. SFHBP performance versus the bound for various choices of the filter order M, using the Shepp–Logan head phantom image.

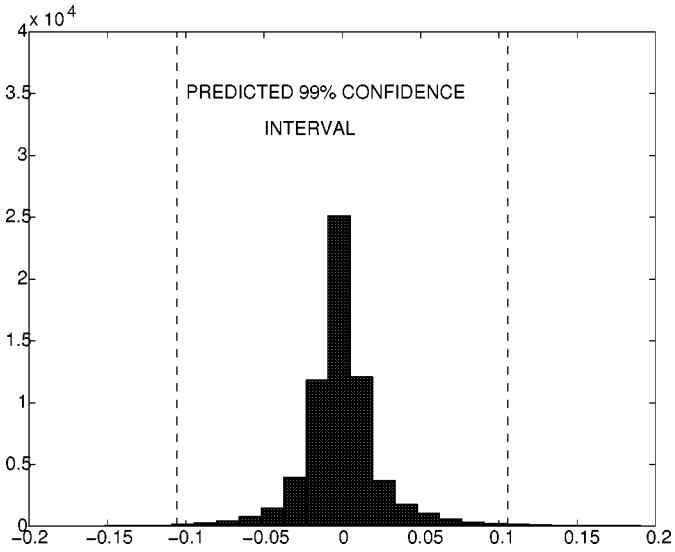

Fig. 12. Histogram of the pixel errors for $M = 12$ with $P = 1024$ projections from the VHD phantom. Indicated is the 99% confidence interval as predicted by the bounds and the ergodicity assumption. Over 97% of the errors fall within the predicted interval.

is good, suggesting that the bounds are an accurate predictor of the backprojection error as a function of filter length as well. To verify that the bounds were accurate for other images, we also reran this experiment using a 256×256 Shepp–Logan head phantom image instead of the VHD image [28]. The results of this experiment, shown in Fig. 11, demonstrate that the bounds are accurate for this image as well.

As a final experiment, we examined the actual distribution of the pixel errors. In Fig. 12, a histogram is shown of the errors for the backprojection from the previous experiment with the VHD data and $M = 12$. Also indicated is the confidence interval corresponding to three standard deviations computed according to the bounds. Thus, from the CLT argument given previously and the ergodic assumption, we expect 99% of the pixel errors to fall within this interval. For this sample image, over 97% of the errors do fall within the interval, suggesting that the Gaussian approximation and ergodic assumptions are justified for modeling the per-pixel error distribution.

VI. Conclusions

We have presented an error analysis for a simplified version of our FHBP algorithm. The analysis lead to useful information on the behavior of the reconstruction error as a function of interpolator length, oversampling, and the parameter J, which determined the number of times the backprojection was decomposed. Simulations on a slice from the Visual Human Dataset support the bounds, and verify the proposed schemes for parameter selection. Future work might build upon these error bounds by allowing for different interpolator lengths at different stages, or focus on error bounds that do not use bandlimited models.

References

[1] S. R. Deans, *The Radon Transform and Some of its Applications*. New York: Wiley, 1983.

[2] S. Basu and Y. Bresler, "$O(n^2 \log_2 n)$ filtered backprojection reconstruction algorithm for tomography," *IEEE Trans. Image Processing*, vol. 9, pp. 1760–1773, Oct. 2000.

[3] A. Brandt, J. Mann, M. Brodski, and M. Galun, "A fast and accurate multilevel inversion of the radon transform," *SIAM J. Appl. Math.*, vol. 60, no. 2, pp. 437–462, 1999.

[4] S. Nilsson, "Fast backprojection," Tech. Rep. LiTH-ISY-R-1865, ISSN 1400-3902, Dept. Elect. Eng., Linköping Univ., S-58 183 Linköping, Sweden, July 1996.

[5] J. McCorkle and M. Rofheart, "An order $n^2 \log(n)$ backprojector algorithm for focusing wide-angle wide-bandwidth arbitrary arbitrary-motion synthetic aperature radar," *Proc. SPIE*, vol. 2747, pp. 25–36, 1996.

[6] S.-M. Oh, J. H. McClellan, and M. C. Cobb, "Multi-resolution mixed-radix quadtree SAR image focusing algorithms," in *Proc. 1999 Advanced Sensors Consortium Federated Laboratory Symp.*, Feb. 1999, pp. 139–143.

[7] O. Seger, M. Herberthson, and H. Hellsten, "Real time SAR processing of low frequency ultra wide band radar data," in *Proc. EUSAR*, 1998, pp. 489–492.

[8] L. Ulander, H. Hellsten, and G. Stenstrom, "Synthetic aperature radar processing using fast factorised backprojection," in *Proc. EUSAR 2000*, 2000, pp. 753–756.

[9] A. Boag, Y. Bresler, and E. Michielssen, "A multilevel domain decomposition algorithm for fast $O(N^2 \log N)$ reprojection of tomographic images," *IEEE Trans. Image Processing*, vol. 9, pp. 1573–1582, Sept. 2000.

[10] F. Natterer, *The Mathematics of Computerized Tomography*. New York: Wiley, 1986.

[11] P. A. Rattey and A. G. Lindgren, "Sampling the 2-D radon transform," *IEEE Trans. Acoust. Speech, Signal, Processing*, vol. ASSP-29, no. 5, pp. 994–1002, 1981.

[12] S. K. Basu, "Fast algorithms for tomography," Ph.D. dissertation, Univ. Illinois, Urbana-Champaign, 2000.

[13] T. Laakso, V. Välimäki, M. Karjalainen, and U. Laine, "Splitting the unit delay: tools for fractional delay filter design," *IEEE Signal Processing Mag.*, pp. 30–60, Jan. 1996.

[14] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.

[15] A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[16] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series and Products*, 5th ed. New York: Academic, 1994.

[17] H. Stark and J. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[18] R. Radzyner and P. Bason, "An error bound for Lagrange interpolation of low-pass functions," *IEEE Trans. Inform. Theory*, pp. 669–671, 1972.

[19] F. Natterer, "Efficient evaluation of oversampled functions," *J. Comput. Appl. Math.*, vol. 14, pp. 303–309, 1986.

[20] W. Putnam and J. Smith, "Design of fractional delay filters using convex optimization," in *IEEE Workshop Applications Signal Processing Audio Acoustics*, 1997.

[21] S. Basu and Y. Bresler, "An empirical study of minimax-optimal fractional delays for lowpass signals," *IEEE Trans. Circuits Syst.*, submitted for publication.

[22] M. Syslo, N. Deo, and J. Kowalik, *Discrete Optimization Algorithms: With Pascal Programs*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
[23] L. J. Karam and J. H. McClellan, "Complex Chebyshev approximation for FIR filter design," *IEEE Trans. Circuits Syst.*, pp. 207–216, Mar. 1995.
[24] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[25] "Visual human dataset for the female," Nat. Lib. Med., Nat. Inst. Health, Bethesda, MD, 1998.
[26] A. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. New York: IEEE Press, 1988.
[27] *Matlab Reference Guide*. Natick, MA: MathWorks, Inc., 1997.
[28] L. A. Shepp and B. F. Logan, "The Fourier reconstruction of head section," *IEEE Trans. Nucl. Sci.*, vol. NS-21, no. 1, pp. 21–43, 1974.

Samit Basu (M'00) received the B.S. degree with honors in electrical engineering from the University of Delaware, Newark, in 1995, and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois, Urbana-Champaign, in 1998 and 2000, respectively.

He is currently an Electrical Engineer with the Computed Tomography Program of General Electric's Corporate Research and Development Laboratory, Niskayuna, NY. His research interests include tomography, inverse problems, and signal and image processing.

Yoram Bresler (F'99) received the B.Sc. (cum laude) and M.Sc. degrees from the Technion—Israel Institute of Technology, Haifa, in 1974 and 1981 respectively, and the Ph.D. degree from Stanford University, Stanford, CA, in 1985, all in electrical engineering.

From 1974 to 1979, he was an Electronics Engineer with the Israeli Defense Force. From 1979 to 1981, he was a Consultant for the Flight Control Lab at the Technion, developing algorithms for autonomous TV aircraft guidance. From 1985 to 1987, he was a Research Associate with the Information Systems Laboratory, Stanford University, working on sensor array processing and medical imaging. In 1987, he joined the University of Illinois, Urbana-Champaign, where he is currently a Professor with the Department of Electrical and Computer Engineering and the Bioengineering Program, and Research Professor at the Coordinated Science Laboratory. In 1995–1996, he was on sabbatical at the Technion. His current research interests include multidimensional and statistical signal processing and their applications to inverse problems in imaging and sensor arrays. He is currently is on the editorial board of *Machine Vision and Applications*

Dr. Bresler was an Associate Editor for the IEEE TRANSACTIONS FOR IMAGE PROCESSING in 1992–1993, and a member of the IEEE Image and Multidimensional Signal Processing Technical Committee from 1994 to 1998. In 1988 and 1989, he received the Senior Paper Awards from the IEEE Acoustics, Speech, and Signal Processing Society. He is the recipient of a 1991 NSF Presidential Young Investigator Award, the Technion Fellowship in 1995, and the Xerox Senior Award for Faculty Research in 1998. He was named a "University of Illinois Scholar" in 1999, and was appointed an Associate at the University of Illinois Center for Advanced Study in 2001.