

Multigrid Tomographic Inversion With Variable Resolution Data and Image Spaces

Seungseok Oh, *Member, IEEE*, Charles A. Bouman, *Fellow, IEEE*, and Kevin J. Webb, *Fellow, IEEE*

Abstract—A multigrid inversion approach that uses variable resolutions of both the data space and the image space is proposed. Since the computational complexity of inverse problems typically increases with a larger number of unknown image pixels and a larger number of measurements, the proposed algorithm further reduces the computation relative to conventional multigrid approaches, which change only the image space resolution at coarse scales. The advantage is particularly important for data-rich applications, where data resolutions may differ for different scales. Applications of the approach to Bayesian reconstruction algorithms in transmission and emission tomography with a generalized Gaussian Markov random field image prior are presented, both with a Poisson noise model and with a quadratic data term. Simulation results indicate that the proposed multigrid approach results in significant improvement in convergence speed compared to the fixed-grid iterative coordinate descent method and a multigrid method with fixed-data resolution.

Index Terms—Computed tomography, emission tomography, image reconstruction, inverse problems, multigrid algorithms, multiresolution, transmission tomography.

I. INTRODUCTION

OVER the past decade, many important image processing applications have been formulated in the framework of inverse problems. However, a major barrier to the use of inverse problem techniques has been the computational cost of these methods, which typically require the optimization of high-dimensional, and sometimes nonquadratic, cost functions. These computational challenges are only made more difficult by concurrent trends toward larger data sets and correspondingly higher resolution images in two and higher dimensions.

Multiresolution techniques have been widely investigated as a method for reducing the computation required to solve inverse problems. The techniques have ranged from simple coarse-to-fine approaches [1]–[5], which initialize fine-scale iterations with coarse-scale solutions, to more sophisticated wavelet or multiresolution image model-based approaches, which have been applied to image segmentation [6]–[9], image restoration [10]–[15], and image reconstruction [16]–[22], [13], [23], [24].

Multigrid methods [25]–[27], which are multiresolution approaches originally developed for fast partial differential

equation (PDE) solvers, have been recently applied to inverse problems such as image reconstruction [28]–[39], optical flow estimation [40]–[44], interpolation of missing image data [45], [46], image segmentation [45], [47], [48], image analysis [40], [49]–[51], image restoration [52], and anisotropic diffusion [53]. Multigrid methods achieve fast convergence not only because coarse-scale operations are much cheaper than those at fine scale, but also because coarse grid corrections typically remove low frequency error components more effectively than fine-scale corrections. Furthermore, unlike simple coarse-to-fine approaches, they provide a systematic method to go from fine to coarse, as well as from coarse to fine, so that coarse-scale updates can be applied whenever they are expected to be effective. Since they operate directly in the space domain, multigrid algorithms can also easily enforce nonnegativity constraints, which are often necessary to obtain a physically meaningful image in tomographic reconstruction problems.

Interestingly, most of the existing work on multigrid image reconstruction has focused on applications that use a forward model described by the solution to one or more PDEs. For example, optical diffusion tomography (ODT) [32]–[34], electrical impedance tomography [29], [54], [35], bioelectric field problem [31], seismic tomography [55], and atmospheric data assimilation [36] all use a forward model that depends implicitly on the solution to a PDE. In these applications multigrid algorithms provide significant computational savings, partly because good initialization is usually not available, and partly because per iteration computation tends to be high. For example, the application of our nonlinear multigrid inversion to ODT showed the potential for very large computational savings and robust convergence with respect to various operational initializations [34]. However, relatively little work has been done on applying multigrid methods to emission and transmission tomography problems [28], [30], [39].

Conventional tomography and many other inverse problems, such as motion analysis and image deblurring, have large measurement data sets which also can be decimated at coarse scales. Some inversion approaches have used multiresolution representations of this data. For example, wavelet decomposition of projection data has been used in filtered backprojection [56]–[61] and MAP reconstruction [17], [62], [23], and a multiscale forward projection equation solver used decimated sinogram data for coarse-scale iterations [63]. Interestingly, the ordered subset expectation-maximization (OSEM) algorithm [64] does not use multiresolution data representation, but it does use only a subset of the data in each iteration. Importantly, existing multigrid methods, including our previous multigrid inversion framework [34], do not exploit the possibility of coarse representation of measurement data at coarser scales,

Manuscript received April 4, 2005; revised November 4, 2005. This work was supported by the National Science Foundation under Contracts CCR-0073357 and CCR-0431024. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Attila Kuba.

S. Oh is with the Fujifilm Software (California), Inc., San Jose, CA 95110 USA (e-mail: soh@fujifilmsoft.com).

C. A. Bouman and K. J. Webb are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907-2035 USA (e-mail: bouman@ecn.purdue.edu; webb@purdue.edu).

Digital Object Identifier 10.1109/TIP.2006.877313

and, thus, their computational gain comes only from a reduced number of unknown variables by coarse discretization of the image at coarser scales.

In this paper, we propose a multigrid method that has three important features. First, it reduces computation by changing the resolution of the data space as well as the image space. Second, it formulates the multigrid inversion problem for Bayesian reconstruction from transmission or emission data with either a Poisson or Gaussian noise model. Third, it incorporates a novel adaptive multigrid scheme which allocates computation to the scale at which the algorithm can best reduce the cost [65].

As with our previous multigrid inversion method [34], our new multigrid method formulates a consistent set of coarse-scale cost functions and moves up and down recursively in resolution to solve the original finest-scale problem. However, the important difference from our previous formulation is that the measurement data as well as the image are coarsely discretized at the coarse scale, and, thus, computation is further reduced. This is especially advantageous in applications where the data, as well as the image, have high dimensions.

An important feature of our formulation is that the choice of decimator/interpolator for the data space is independent of the choice of those for the image space. In many image processing applications, such as motion analysis and image deblurring, a measurement is available for each pixel of the image space, so the same decimation/interpolation operators may be used on both the data and images. However, in many applications, including tomography, this is not true. Thus, the flexibility in choosing the decimator/interpolator makes our proposed multigrid approach particularly suitable for tomographic image reconstruction problems.

Our simulation results show that our multigrid algorithms using variable data resolution yield better convergence speed than the iterative coordinate descent (ICD) method [66], [67] and multigrid algorithms using fixed-data resolution.

II. MULTIGRID INVERSION WITH VARIABLE RESOLUTION DATA AND IMAGE SPACES

In this section, we present a multigrid inversion approach that changes resolutions of both data and image spaces. We first present our approach for the case of measurements with additive Gaussian noise, and we then generalize the method for inversion with Poisson noise.

A. Quadratic Data Term Case

Let $Y \in \mathbb{R}^M$ be a random vector of measured data, and let $x \in \mathbb{R}^N$ be a discretized unknown image. Then, the expected value of the measurement vector is given by

$$E[Y|x] = f(x) \quad (1)$$

where $f: \mathbb{R}^N \rightarrow \mathbb{R}^M$ is known as the forward model. Our task is then to estimate the image x which produced the observations

Y . A common approach for solving this problem is to solve an associated optimization problem of the form

$$\hat{x} = \arg \min_x \{-\log p(y|x) + S(x)\} \quad (2)$$

where $p(y|x)$ is the probability density of Y given x , and $S(x)$ is a stabilizing function designed to regularize the inversion [68], [69]. If $S(x) = -\log p(x)$, where $p(x)$ is the image prior probability density, this results in the maximum *a posteriori* (MAP) estimate of x .

If the measurements Y are conditionally Gaussian given x with noise covariance matrix $(2\Lambda)^{-1}$, then the inverse is computed by minimizing the cost function

$$\|y - f(x)\|_{\Lambda}^2 + S(x) \quad (3)$$

where $\|w\|_{\Lambda}^2 = w^H \Lambda w$. By expanding the data term of (3), the cost function may be expressed within a constant as

$$c(x) = \|f(x)\|_{\Lambda}^2 + 2a^T f(x) + S(x) \quad (4)$$

where T is the transpose operator and $a = -\Lambda^T y$.

Minimizing a function such as (4) can be very computationally expensive, particularly when the image x and data y have high dimension. Our approach to reducing computation will be to formulate an approximate cost function using a coarse-scale representation of the image and data. To do this, we require methods for decimating and interpolating in both domains.

Let $x^{(q)} \in \mathbb{R}^{N^{(q)}}$ and $y^{(q)} \in \mathbb{R}^{M^{(q)}}$ denote representations of $x = x^{(0)}$ and $y = y^{(0)}$ at coarser resolution q . In order to convert between resolutions, we define the image domain decimation operator $x^{(q+1)} = I_{(q)}^{(q+1)} x^{(q)}$ and the data domain decimation operator $y^{(q+1)} = J_{(q)}^{(q+1)} y^{(q)}$. Similarly, we define the interpolation operators for image and data domains as $x^{(q)} = I_{(q+1)}^{(q)} x^{(q+1)}$ and $y^{(q)} = J_{(q+1)}^{(q)} y^{(q+1)}$, respectively. Typically, we use either pixel replication or bilinear interpolation operators and decimation operators, but the theory is applicable to a wide range of choices. Notice that, in general, $I_{(q)}^{(q+1)}$ and $J_{(q)}^{(q+1)}$ may be different.

We will assume that there is some natural way to define a coarse-scale forward model $f^{(q)}: \mathbb{R}^{N^{(q)}} \rightarrow \mathbb{R}^{M^{(q)}}$ which maps the coarse-scale image to the coarse-scale data. In practice, $f^{(q)}(\cdot)$ can result from the method used to discretize the physical problem, but, at this point, we will make few assumptions regarding its specific form. The most crucial assumption in our formulation is that

$$f^{(0)}(x^{(0)}) \cong J_{(q)}^{(0)} f^{(q)}(x^{(q)}) \quad (5)$$

Then, by replacing $f^{(0)}(x^{(0)})$ in the original finest-scale cost function (4) with an interpolated forward model $J_{(q)}^{(0)} f^{(q)}(x^{(q)})$, we have an approximate coarse-scale cost function

$$\begin{aligned} \tilde{c}^{(q)}(x^{(q)}) = & \|J_{(q)}^{(0)} f^{(q)}(x^{(q)})\|_{\Lambda}^2 \\ & + 2a^T J_{(q)}^{(0)} f^{(q)}(x^{(q)}) + S^{(q)}(x^{(q)}) \quad (6) \end{aligned}$$

where the coarse-scale stabilizing function $S^{(q)}(\cdot)$ is chosen to best approximate the original finest scale one, as described in [34], and later in Section IV-A. By defining

$$\Lambda^{(q)} = [J_{(q)}^{(0)}]^T \Lambda^{(0)} J_{(q)}^{(0)} \quad (7)$$

$$a^{(q)} = [J_{(q)}^{(0)}]^T a^{(0)} \quad (8)$$

(6) can be expressed as

$$\tilde{c}^{(q)}(x^{(q)}) = \|f^{(q)}(x^{(q)})\|_{\Lambda^{(q)}}^2 + 2a^{(q)T} f^{(q)}(x^{(q)}) + S^{(q)}(x^{(q)}). \quad (9)$$

The form of (9) is analogous to that of (4), but with quantities indexed by the scale q . As in our previous work [34], the forward model $f^{(q)}(\cdot)$ and the stabilizing function $S^{(q)}(\cdot)$ use a coarsely discretized image at each scale q , and, thus, computations are substantially reduced due to the reduced number of variables. In this work, computation is further reduced since the dimension of the forward model vector also changes with q .

We adjust the coarse-scale cost functions (9) at each scale to better match with the original fine scale one, and, thus, to produce a consistent solution. To do this, we define an adjusted cost function by appending an additional linear correction term. This yields the adjusted cost function

$$c^{(q)}(x^{(q)}) = \|f^{(q)}(x^{(q)})\|_{\Lambda^{(q)}}^2 + 2a^{(q)T} f^{(q)}(x^{(q)}) + S^{(q)}(x^{(q)}) - r^{(q)} x^{(q)} \quad (10)$$

where $r^{(q)}$ is a row vector used to adjust the function's gradient, the choice of which will be discussed later. At the finest scale, $r^{(0)} = 0$ is chosen so that $c^{(0)}(x^{(0)}) = c(x)$.

With the set of coarse-scale cost functions of the form in (10), the multigrid algorithm solves the original problem by moving up and down in resolution [33], [34]. Let $x^{(q)}$ be the current solution at grid q . We would like to improve this solution by first performing iterations of fixed-grid optimization at the coarser grid $q+1$, and then using this result to correct the finer grid solution. This coarse grid update is

$$\tilde{x}^{(q+1)} \leftarrow \text{Fixed_Grid_Update}(I_{(q)}^{(q+1)} x^{(q)}, c^{(q+1)}(\cdot)) \quad (11)$$

where $\tilde{x}^{(q+1)}$ is the updated value, and the operator $\text{Fixed_Grid_Update}(x_{\text{init}}, c(\cdot))$ is any fixed-grid update algorithm designed to reduce the cost function $c(\cdot)$ starting with the initial value x_{init} . In (11), the initial condition $I_{(q)}^{(q+1)} x^{(q)}$ is formed by decimating $x^{(q)}$. We may now use this result to update the finer grid solution. We do this by interpolating the *change* in the coarser-scale solution.

$$\tilde{x}^{(q)} \leftarrow x^{(q)} + I_{(q+1)}^{(q)} (\tilde{x}^{(q+1)} - I_{(q)}^{(q+1)} x^{(q)}). \quad (12)$$

In order to ensure updates which reduce the fine-scale cost, we would like to make the fine and coarse-scale cost functions equal within an additive constant. This means we would like the equation

$$c^{(q+1)}(\tilde{x}^{(q+1)}) \cong c^{(q)} \left(x^{(q)} + I_{(q+1)}^{(q)} (\tilde{x}^{(q+1)} - I_{(q)}^{(q+1)} x^{(q)}) \right) + \text{constant} \quad (13)$$

to hold for all coarse-scale updated values of $\tilde{x}^{(q+1)}$. Our objective is then to choose a coarse-scale cost function which matches the fine cost function, as described in (13). We do this by selecting $r^{(q+1)}$ to match the gradients of the coarse and fine cost functions at the current values of $x^{(q)}$ and $x^{(q+1)} = I_{(q)}^{(q+1)} x^{(q)}$. More precisely, we enforce the condition that

$$\nabla c^{(q+1)}(x^{(q+1)}) \Big|_{x^{(q+1)} = I_{(q)}^{(q+1)} x^{(q)}} = \nabla c^{(q)}(x^{(q)}) I_{(q+1)}^{(q)} \quad (14)$$

where $\nabla c(x)$ is the row vector formed by the gradient of the function $c(\cdot)$ [33]. This condition (14) is essential to assure that the optimum solution is a fixed point of the multigrid inversion algorithm [33], and we can show how this condition can be used along with other assumptions to ensure monotone convergence of the multigrid inversion algorithm [34]. Note that in (14), the interpolation matrix $I_{(q+1)}^{(q)}$, which comes from the chain rule of differentiation, actually functions like a decimation operator because it multiplies the gradient vector on the right. Importantly, the condition (14) holds for any choice of decimation and interpolation matrices. The equality of (14) can be enforced at the current value $x^{(q)}$ by choosing

$$r^{(q+1)} \leftarrow \nabla \tilde{c}^{(q+1)}(x^{(q+1)}) \Big|_{x^{(q+1)} = I_{(q)}^{(q+1)} x^{(q)}} - \left(\nabla \tilde{c}^{(q)}(x^{(q)}) - r^{(q)} \right) I_{(q+1)}^{(q)}. \quad (15)$$

By evaluating the gradient for the cost function (4), (15) is computed by

$$r^{(q+1)} \leftarrow g^{(q+1)} - \left(g^{(q)} - r^{(q)} \right) I_{(q+1)}^{(q)} \quad (16)$$

where $g^{(q)}$ and $g^{(q+1)}$ are the gradients of the unadjusted cost function at the fine and coarse scales, respectively, given by

$$g^{(q)} \leftarrow 2 \left(f^{(q)}(x^{(q)}) \Lambda^{(q)T} + a^{(q)} \right)^T A^{(q)} + \nabla S^{(q)}(x^{(q)}) \quad (17)$$

$$g^{(q+1)} \leftarrow 2 \left(f^{(q+1)}(x^{(q+1)}) \Lambda^{(q+1)T} + a^{(q+1)} \right)^T A^{(q+1)} + \nabla S^{(q+1)}(x^{(q+1)}) \quad (18)$$

where $A^{(q)}$ denotes the gradient of the forward model or Fréchet derivative given by

$$A^{(q)} = \nabla f^{(q)}(x^{(q)}) \quad (19)$$

$$A^{(q+1)} = \nabla f^{(q+1)}(x^{(q+1)}) \Big|_{x^{(q+1)} = I_{(q)}^{(q+1)} x^{(q)}} \quad (20)$$

Assuming that

$$J_{(q+1)}^{(0)} = J_{(q)}^{(0)} J_{(q+1)}^{(q)} \quad (21)$$

the coarse-scale cost function parameters (7)–(8) can be computed iteratively by

$$\Lambda^{(q+1)} \leftarrow [J_{(q+1)}^{(q)}]^T \Lambda^{(q)} J_{(q+1)}^{(q)} \quad (22)$$

$$a^{(q+1)} \leftarrow [J_{(q+1)}^{(q)}]^T a^{(q)}. \quad (23)$$

```

main() {
  Initialize  $x^{(0)}$  with a background estimate
  For  $q = 0, 1, \dots, Q - 2$ ,  $x^{(q+1)} \leftarrow J_{(q)}^{(q+1)} x^{(q)}$ 
  For  $q = 0, 1, \dots, Q - 1$ ,  $r^{(q)} \leftarrow 0$ 
  If Gaussian noise model is used, then {
    For  $q = 0, 1, \dots, Q - 2$ ,  $\Lambda^{(q+1)} \leftarrow [J_{(q+1)}^{(q)}]^T \Lambda^{(q)} J_{(q+1)}^{(q)}$ 
    For  $q = 0, 1, \dots, Q - 2$ ,  $a^{(q+1)} \leftarrow [J_{(q+1)}^{(q)}]^T a^{(q)}$ 
  }
  If Poisson noise model is used, then {
    For  $q = 1, 2, \dots, Q - 1$ ,  $y^{(q)} \leftarrow J_{(0)}^{(q)} y^{(0)}$ 
  }
  Choose number of fixed grid iterations  $\nu_1^{(0)}, \dots, \nu_1^{(Q-1)}$  and  $\nu_2^{(0)}, \dots, \nu_2^{(Q-1)}$ 
  Repeat until converged:
     $x^{(0)} \leftarrow \text{MultigridV}(q, x^{(0)}, r^{(0)})$ 
}

 $x^{(q)} \leftarrow \text{MultigridV}(q, x^{(q)}, r^{(q)})$  {
  Repeat  $\nu_1^{(q)}$  times
     $x^{(q)} \leftarrow \text{Fixed\_Grid\_Update}(x^{(q)}, c^{(q)}(\cdot; r^{(q)}))$  //Fine grid update
  If  $q = Q - 1$ , return  $x^{(q)}$  //If coarsest scale, return result
   $x^{(q+1)} \leftarrow J_{(q)}^{(q+1)} x^{(q)}$  //Decimation
  If Gaussian noise model is used, then {
    Compute  $r^{(q+1)}$  using (15) (17) and (18)
  }
  If Poisson noise model is used, then {
    Compute  $r^{(q+1)}$  using (15) (33) and (34)
  }
   $x^{(q+1)} \leftarrow \text{MultigridV}(q + 1, x^{(q+1)}, r^{(q+1)})$  //Coarse grid update
   $x^{(q)} \leftarrow x^{(q)} + J_{(q+1)}^{(q)}(x^{(q+1)} - J_{(q)}^{(q+1)} x^{(q)})$  //Coarse grid correction
  Repeat  $\nu_2^{(q)}$  times
     $x^{(q)} \leftarrow \text{Fixed\_Grid\_Update}(x^{(q)}, c^{(q)}(\cdot; r^{(q)}))$  //Fine grid update
  Return  $x^{(q)}$  //Return result
}

```

(a)

(b)

Fig. 1. Pseudo-code specification of (a) the main routine for multigrid inversion and (b) the subroutine for the Multigrid-V inversion.

The computations of (22) and (23) are inexpensive and, in addition, can be precomputed since they are independent of the image $x^{(q)}$.

The pseudocode in Fig. 1(b) shows the Multigrid-V algorithm to solve the minimization of (4). Multigrid-V recursion is a standard multigrid method, which calls itself recursively in resolution. More specifically, it replaces the coarse-scale fixed-grid update of (11) by a recursive call of a multigrid algorithm. We solve the problem through iterative application of the Multigrid-V algorithm, as shown in Fig. 1(a). See [25]–[27], [33], and [34] for the details of Multigrid-V recursion.

B. Poisson Data Case

Some inverse problems, such as transmission and emission tomography, use Poisson measurement noise models [70], [71]. In the Poisson noise model, we assume (1) holds with the elements of Y being independent Poisson random variables. In this case, the negative log likelihood of the Poisson data is given by

$$-\log p(y|x) = \sum_{m=1}^M \{f_m(x) - y_m \log f_m(x) + \log(y_m!)\} \quad (24)$$

where M is the number of measurements and y_m is a realization of Y_m , and its corresponding regularized inverse can be solved by minimizing the cost function

$$c(x) = \sum_{m=1}^M \{f_m(x) - y_m \log f_m(x)\} + S(x). \quad (25)$$

We first compute coarse-scale measurement data using data domain decimation

$$y^{(q)} \triangleq J_{(0)}^{(q)} y^{(0)}. \quad (26)$$

In addition to (5), we also make a few assumptions, which are satisfied for most choices of data domain decimation and interpolation operators. First, we assume that the interpolated coarse-scale data approximate the fine-scale data. More formally, we say

$$y^{(0)} \cong J_{(q)}^{(0)} y^{(q)}. \quad (27)$$

Second, we assume that

$$f_m^{(0)}(x^{(0)}) \cong f_i^{(q)}(x^{(q)}) \quad \text{for } [J_{(q)}^{(0)}]_{m,i} \neq 0 \quad (28)$$

where $[B]_{m,i}$ is the (m, i) th element of matrix B . In order to understand this assumption, notice that, when $[J_{(q)}^{(0)}]_{m,i}$ is nonzero, m and i indicate the corresponding data at different resolutions. So, in this case, we would expect the two data to be approximately equal. Third, we assume that

$$\sum_{m=1}^{M^{(0)}} [J_{(q)}^{(0)}]_{m,i} = \frac{M^{(0)}}{M^{(q)}} \quad (29)$$

which insures that the average value of $y^{(0)}$ and $y^{(q)}$ are the same.

The negative logarithm of the Poisson data likelihood (24) can then be approximated as (30), shown at the bottom of the next page, where the second line comes from (5) and (27), the third from the element-by-element expansion of the data domain interpolation, the fourth from (28), the fifth from the summation order exchange, and the last from (29). Thus, an approximate coarse-scale cost function with a reduced resolution data and forward model may be expressed as

$$\tilde{c}^{(q)}(x^{(q)}) = \frac{M^{(0)}}{M^{(q)}} \sum_{m=1}^{M^{(q)}} \left[f_m^{(q)}(x^{(q)}) - y_m^{(q)} \log f_m^{(q)}(x^{(q)}) \right] + S^{(q)}(x^{(q)}). \quad (31)$$

The adjusted coarse-scale cost is then obtained by adding the gradient correction term

$$c^{(q)}(x^{(q)}) = \frac{M^{(0)}}{M^{(q)}} \sum_{m=1}^{M^{(q)}} \left\{ f_m^{(q)}(x^{(q)}) - y_m^{(q)} \log f_m^{(q)}(x^{(q)}) \right\} + S^{(q)}(x^{(q)}) - r^{(q)} x^{(q)} \quad (32)$$

where $r^{(q)}$ is computed by (16) with

$$g^{(q)} \leftarrow \frac{M^{(0)}}{M^{(q)}} \sum_{m=1}^{M^{(q)}} \left[A_{m,*}^{(q)} \left(1 - \frac{y_m^{(q)}}{f_m^{(q)}(x^{(q)})} \right) \right] + \nabla S^{(q)}(x^{(q)}) \quad (33)$$

$$g^{(q+1)} \leftarrow \frac{M^{(0)}}{M^{(q+1)}} \sum_{m=1}^{M^{(q+1)}} \left[A_{m,*}^{(q+1)} \left(1 - \frac{y_m^{(q+1)}}{f_m^{(q+1)}(x^{(q+1)})} \right) \right] + \nabla S^{(q+1)}(x^{(q+1)}) \quad (34)$$

where $A_{m,*}$ denotes the m th row of the matrix A . With this choice of coarse-scale cost functions, multigrid inversion works by the procedure specified in Fig. 1.

III. ADAPTIVE COMPUTATION ALLOCATION

The MultigridV subroutine in Fig. 1(b) specifies that $\nu_1^{(q)}$ fixed-grid iterations are performed before each coarse grid update, and $\nu_2^{(q)}$ iterations are performed after the update. The convergence speed of the algorithm can be tuned through the choices of $\nu_1^{(q)}$ and $\nu_2^{(q)}$ at each scale. In practice, the best choice of these parameters also varies with the number of MultigridV iterations. For example, coarse fixed-grid optimization is typically more important in initial iterations, while fine fixed-grid optimization is more important during later iterations when the solution is close to its final value. For this reason, we can further improve convergence speed by adaptively changing the values of $\nu_1^{(q)}$ and $\nu_2^{(q)}$ with time instead of fixing the parameters to predetermined values.

In this section, we describe how to adaptively allocate computation to the scale at which the algorithm can best reduce the cost [65]. In our adaptive scheme, we do not fix the $\nu_1^{(q)}$ and $\nu_2^{(q)}$ parameters in advance. Instead, we perform fixed-grid updates as long as they continue to effectively reduce cost. This adaptive approach can further improve convergence speed and eliminates the need to select these parameters.

First, we would like the image updates to begin at the coarsest scale since this is usually more effective when the solution is far from the optimum. To do this, we initially set $\nu_1^{(q)} = 0$, so that when proceeding from fine to coarse scale in the first multigrid-V cycle, we do not update the image and only update the r vector.

Second, when proceeding from coarse to fine scale in the first multigrid-V cycle, we perform the fixed-grid iterations until the change in the cost function falls below a threshold. More specifically, fixed-grid iterations are applied as long as the condition

$$C_1 : \Delta c^{(q)} \geq \Delta_{\max} c^{(q)} T \quad (35)$$

is satisfied, where $\Delta c^{(q)}$ is a state variable containing the reduction in cost that resulted from the most recent application of the fixed-grid optimization at grid resolution q , $\Delta_{\max} c^{(q)}$ is a state variable containing the maximum value that $\Delta c^{(q)}$ has taken on, and T is a threshold which we set to the value 0.1 in this paper. If the condition is not satisfied, the algorithm proceeds to the next scale.

Once the first multigrid cycle is complete, the adaptive multigrid algorithm compares the computational efficiency at the current scale q and at the next grid scale denoted by q_{next} , and performs the fixed-grid iteration at scale q only if it is likely to be more effective than moving to scale q_{next} . More specifically, before each fixed-grid update, a conditional test, C_2 , is evaluated. If the test is true, the fixed-grid update is performed; but if it is false, then the algorithm proceeds to the next grid scale q_{next} . This condition is given by

$$C_2 : \frac{\Delta c^{(q)}}{\text{comp}^{(q)}} \geq \frac{\Delta c^{(q_{\text{next}})}}{\text{comp}^{(q_{\text{next}})}} \quad (36)$$

where $\text{comp}^{(q)}$ is the computation required for a single fixed-grid update at scale q . Importantly, since $\Delta c^{(q)}$ and $\Delta c^{(q_{\text{next}})}$ are state variables, these values are saved from the previous pass through grid resolutions q and q_{next} .

The adaptive MultigridV algorithm is schematically summarized in Fig. 2. While some adaptive multigrid algorithms have been developed for PDE solvers [72], our adaptive scheme is unique because it uses the cost change as the criterion for adaptation. This is possible because our multigrid inversion method is based on an optimization framework [33], [34], in contrast to conventional multigrid methods which are formulated as equation solvers.

IV. APPLICATIONS TO BAYESIAN EMISSION AND TRANSMISSION TOMOGRAPHY

In this section, we apply the proposed multigrid inversion method to iterative reconstruction for emission and transmission

$$\begin{aligned}
-\log p(y|x) - \sum_{m=1}^M \log(y_m!) &= \sum_{m=1}^{M^{(0)}} \left\{ f_m^{(0)}(x^{(0)}) - y_m^{(0)} \log f_m^{(0)}(x^{(0)}) \right\} \\
&\cong \sum_{m=1}^{M^{(0)}} \left\{ \left[J_{(q)}^{(0)} f^{(q)}(x^{(q)}) \right]_m - \left[J_{(q)}^{(0)} y^{(q)} \right]_m \log f_m^{(0)}(x^{(0)}) \right\} \\
&= \sum_{m=1}^{M^{(0)}} \left\{ \sum_{i=1}^{M^{(q)}} \left[J_{(q)}^{(0)} \right]_{m,i} f_i^{(q)}(x^{(q)}) - \sum_{i=1}^{M^{(q)}} \left[J_{(q)}^{(0)} \right]_{m,i} y_i^{(q)} \log f_m^{(0)}(x^{(0)}) \right\} \\
&\cong \sum_{m=1}^{M^{(0)}} \left\{ \sum_{i=1}^{M^{(q)}} \left[J_{(q)}^{(0)} \right]_{m,i} f_i^{(q)}(x^{(q)}) - \sum_{i=1}^{M^{(q)}} \left[J_{(q)}^{(0)} \right]_{m,i} y_i^{(q)} \log f_i^{(q)}(x^{(q)}) \right\} \\
&= \sum_{i=1}^{M^{(q)}} \left\{ \left(f_i^{(q)}(x^{(q)}) - y_i^{(q)} \log f_i^{(q)}(x^{(q)}) \right) \sum_{m=1}^{M^{(0)}} \left[J_{(q)}^{(0)} \right]_{m,i} \right\} \\
&= \frac{M^{(0)}}{M^{(q)}} \sum_{i=1}^{M^{(q)}} \left[f_i^{(q)}(x^{(q)}) - y_i^{(q)} \log f_i^{(q)}(x^{(q)}) \right] \quad (30)
\end{aligned}$$

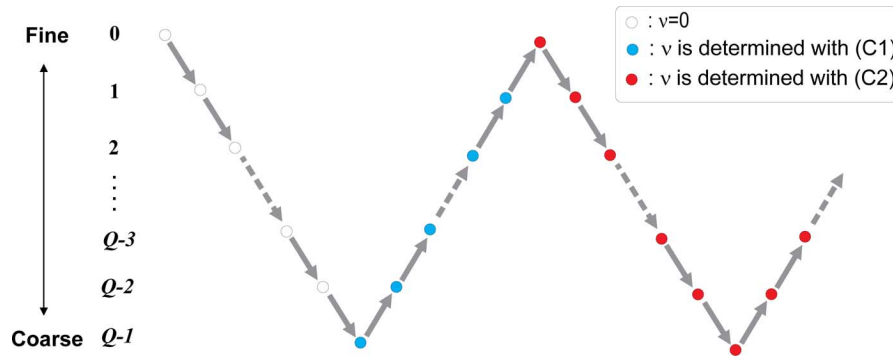


Fig. 2. Adaptive multigrid-V scheme. (Color version available online at <http://ieeexplore.ieee.org>.)

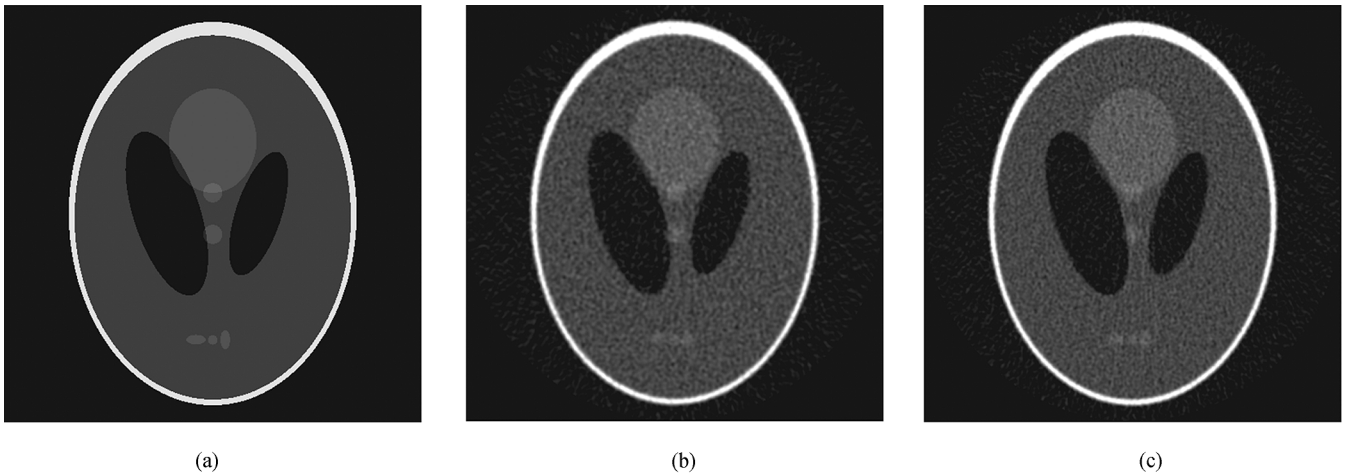


Fig. 3. (a) True phantom. (b) CBP reconstruction for emission tomography. (c) CBP reconstruction for transmission tomography.

tomography. The algorithms are formulated in a Bayesian reconstruction framework using both the quadratic data term and the Poisson noise model.

A. Multigrid Tomographic Inversion With Quadratic Data Term

Emission tomography and transmission tomography use projected photon counts y to reconstruct the image x , which consists of a cross-sectional emission rate map and a cross-sectional attenuation map, respectively. The MAP image reconstruction problem is reduced to a minimization problem with the cost function [66], [67]

$$\|\gamma - Px\|_{\Lambda}^2 + S(x) \quad (37)$$

where, for the emission case, we have

$$\gamma_m = y_m \quad (38)$$

$$\Lambda = \frac{1}{2} \text{diag} \left\{ \frac{1}{y_1}, \frac{1}{y_2}, \dots, \frac{1}{y_M} \right\} \quad (39)$$

and, for the transmission case, we have

$$\gamma_m = \log \frac{y_T}{y_m} \quad (40)$$

$$\Lambda = \frac{1}{2} \text{diag} \{y_1, y_2, \dots, y_M\} \quad (41)$$

where P is the forward projection matrix, y_T is the photon dosage per ray in the transmission case, and γ plays a role similar to y in (3).

Notice that, since (37) has the form of (3), we can use the multigrid inversion algorithm described in Section II-A to compute the MAP reconstruction. However, to do this we must specify the coarse-scale forward models, $f^{(q)}(\cdot)$, and the coarse-scale stabilizing functions, $S^{(q)}(\cdot)$.

The fine-scale forward model is given by the linear transformation

$$f(x) = Px. \quad (42)$$

The coarse-scale forward model also has the linear form

$$f^{(q)}(x^{(q)}) = P^{(q)}x^{(q)} \quad (43)$$

where $P^{(q)}$ is an $M^{(q)} \times N^{(q)}$ coarse-scale projection matrix given by

$$P^{(q+1)} \triangleq J_{(q)}^{(q+1)} P^{(q)} I_{(q+1)}^{(q)}. \quad (44)$$

Note that $P^{(q+1)}$ in (44) can be precomputed and stored since it is independent of the images.

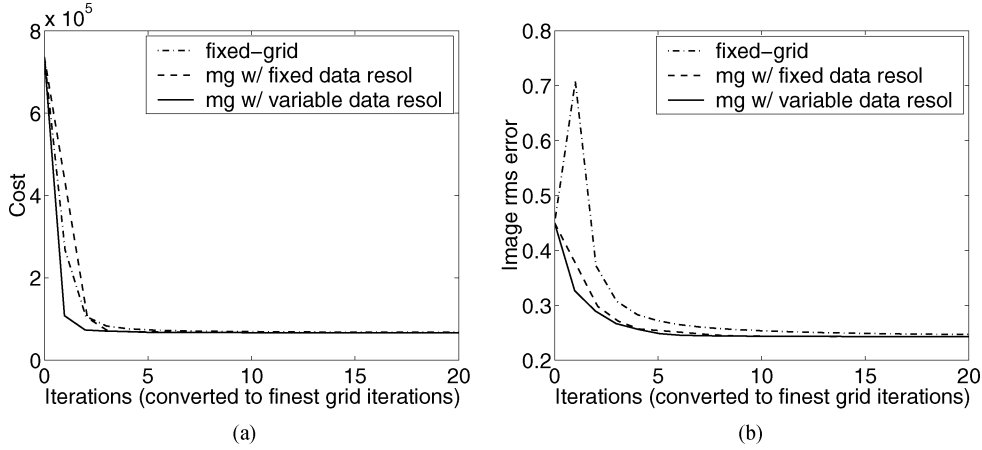


Fig. 4. Convergence in emission tomography with quadratic data term in terms of (a) cost function and (b) image rms error.

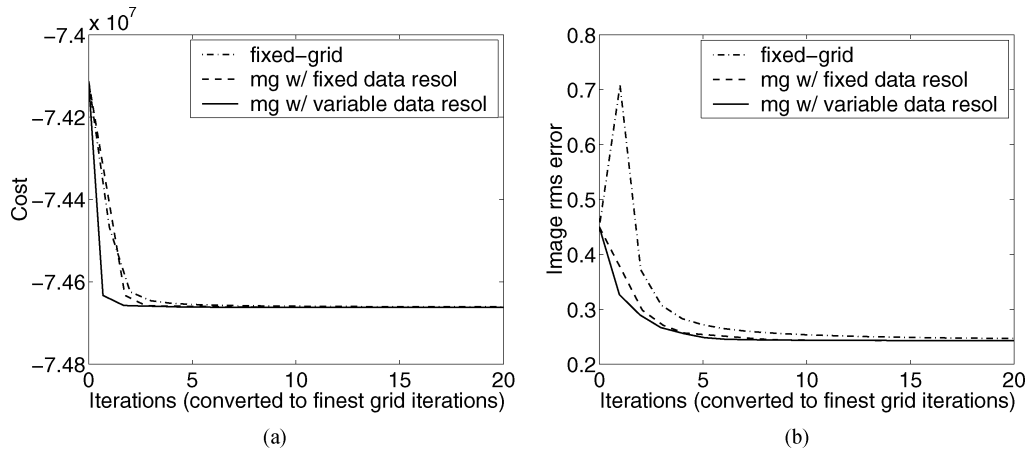


Fig. 5. Convergence in emission tomography with the Poisson noise model in terms of (a) cost function and (b) image rms error.

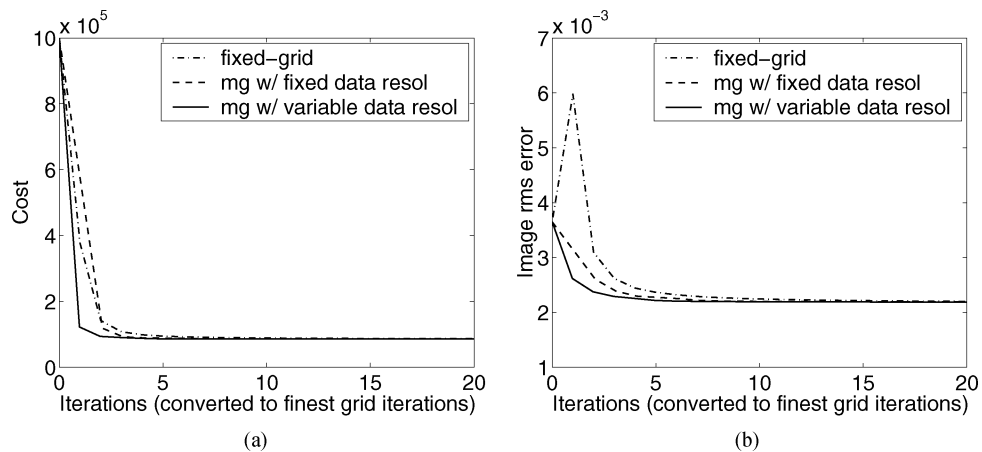


Fig. 6. Convergence in transmission tomography with quadratic data term in terms of (a) cost function and (b) image rms error.

Although, in principle, our multigrid inversion framework can work with any choice of data domain interpolator $J_{(q+1)}^{(q)}$ and decimator $J_{(q)}^{(q+1)}$, we need to choose them carefully to retain computational efficiency. We choose $J_{(q+1)}^{(q)}$ so that each

row has only one nonzero element, and, thus, the resulting coarse-scale weight matrix $\Lambda^{(q)}$ given by (22) is diagonal. For this reason, we interpolate using pixel replication along both the displacement and angle dimensions of the sinogram data. In other words, $J_{(q+1)}^{(q)}$ interpolates the sinogram data with the

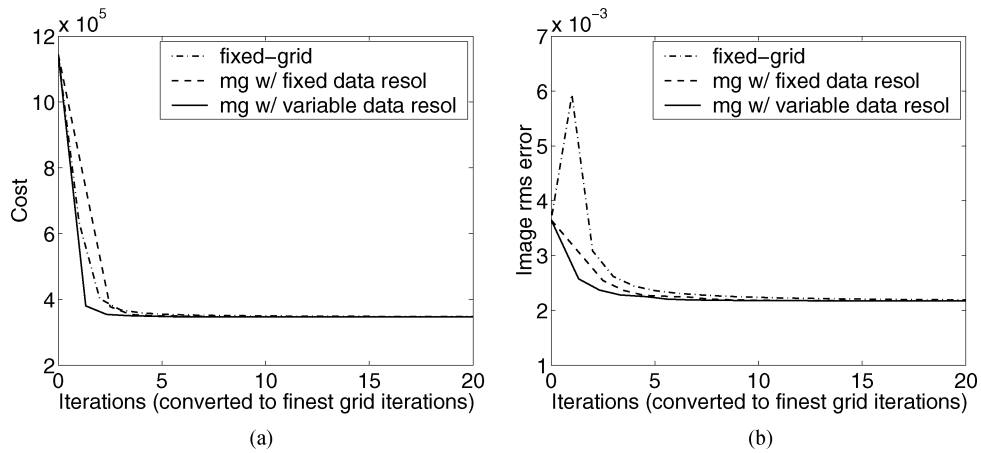


Fig. 7. Convergence in transmission tomography with the Poisson noise model in terms of (a) cost function and (b) image rms error.

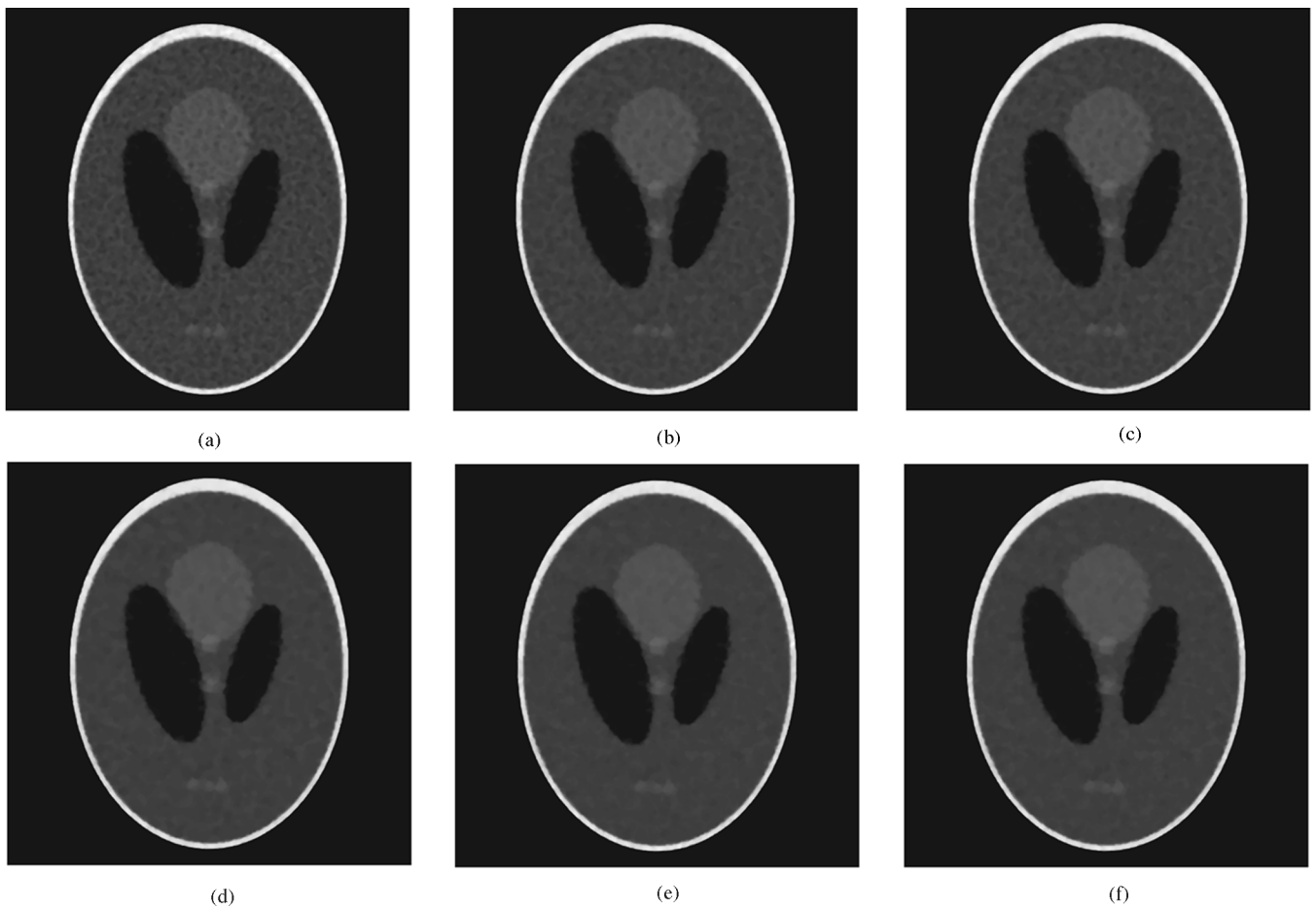


Fig. 8. Reconstructions for emission tomography with quadratic data term: Fixed-grid algorithm with (a) seven iterations, (b) 14 iterations, (c) 28 iterations, and (d) 50 iterations. (e) Multigrid algorithm with fixed-data resolution (7.79 iterations) and (f) multigrid algorithm with variable data resolution (5.94 iterations).

one-dimensional interpolation matrix

$$(45) \quad \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

along both the angle and displacement axes. We choose the decimator to have the adjoint form of the interpolator, giving

$$(46) \quad J_{(g)}^{(q+1)} = \frac{1}{2} [J_{(g+1)}^{(g)}]^T$$

Note that some other interpolation matrices, including the popular bilinear interpolator, do not preserve the sparsity of weight matrix $\Lambda^{(g)}$ at coarse scales.

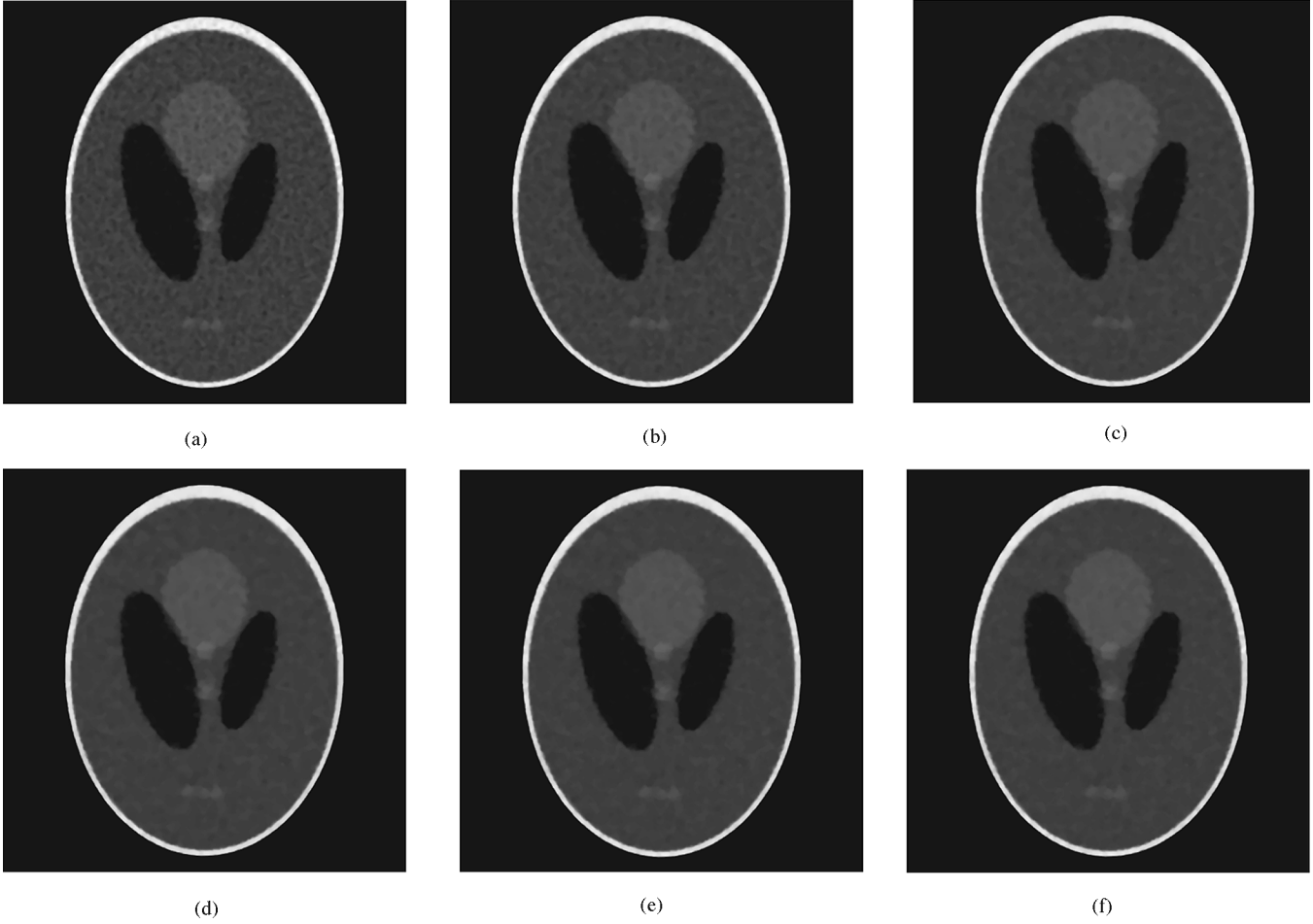


Fig. 9. Reconstructions for emission tomography with the Poisson noise model: Fixed-grid algorithm with (a) seven iterations, (b) 14 iterations, (c) 28 iterations, and (d) 50 iterations. (e) Multigrid algorithm with fixed-data resolution (8.06 iterations) and (f) multigrid algorithm with variable data resolution (5.31 iterations).

For the image prior model, we use the generalized Gaussian Markov random field (GGMRF) model [73], which is known to effectively enforce smoothness while preserving edges in tomographic reconstruction. In this case, the stabilizing function is given by

$$S(x) = \frac{1}{p\sigma^p} \sum_{\{i,j\} \in \mathcal{N}} b_{i-j} |x_i - x_j|^p \quad (47)$$

where σ is a normalization parameter, $1 \leq p \leq 2$ controls the degree of edge smoothness, the set \mathcal{N} consists of all pairs of adjacent pixels, and b_{i-j} is a weight given to the pair of pixels i and j . We use the corresponding coarse-scale stabilizing functions [34]

$$S^{(q)}(x^{(q)}) = \frac{1}{p(\sigma^{(q)})^p} \sum_{\{i,j\} \in \mathcal{N}} b_{i-j} |x_i^{(q)} - x_j^{(q)}|^p \quad (48)$$

where $\sigma^{(q)}$ is given by $\sigma^{(q)} = 2^{q(1-d/p)} \cdot \sigma^{(0)}$, and d is the dimensionality of the problem. The gradient terms of the stabi-

lizing function used in (17), (18), (33), and (34) are computed by

$$\frac{\partial}{\partial x_n} S(x) = \frac{1}{(\sigma^{(q)})^p} \sum_{j \in \mathcal{N}_n} b_{n-j} |x_n^{(q)} - x_j^{(q)}|^{p-1} \text{sgn}(x_n^{(q)} - x_j^{(q)}). \quad (49)$$

B. Multigrid Tomographic Inversion for Poisson Data Model

In the emission case, the photon count Y_m for the m th detector or detector pair is known to be described by the Poisson distribution (24) with mean and variance

$$f_m(x) = P_{m,*}x \quad (50)$$

where $P_{m,*}$ is the m th row of the matrix P . For this case, the MAP image reconstruction problem is reduced to minimizing the cost function (25) with the Poisson mean (50). We also use the coarse-scale projection matrix of (44).

A similar method can be used for the transmission case, but with the Poisson mean given by

$$f_m(x) = y_T \exp(-P_{m,*}x). \quad (51)$$

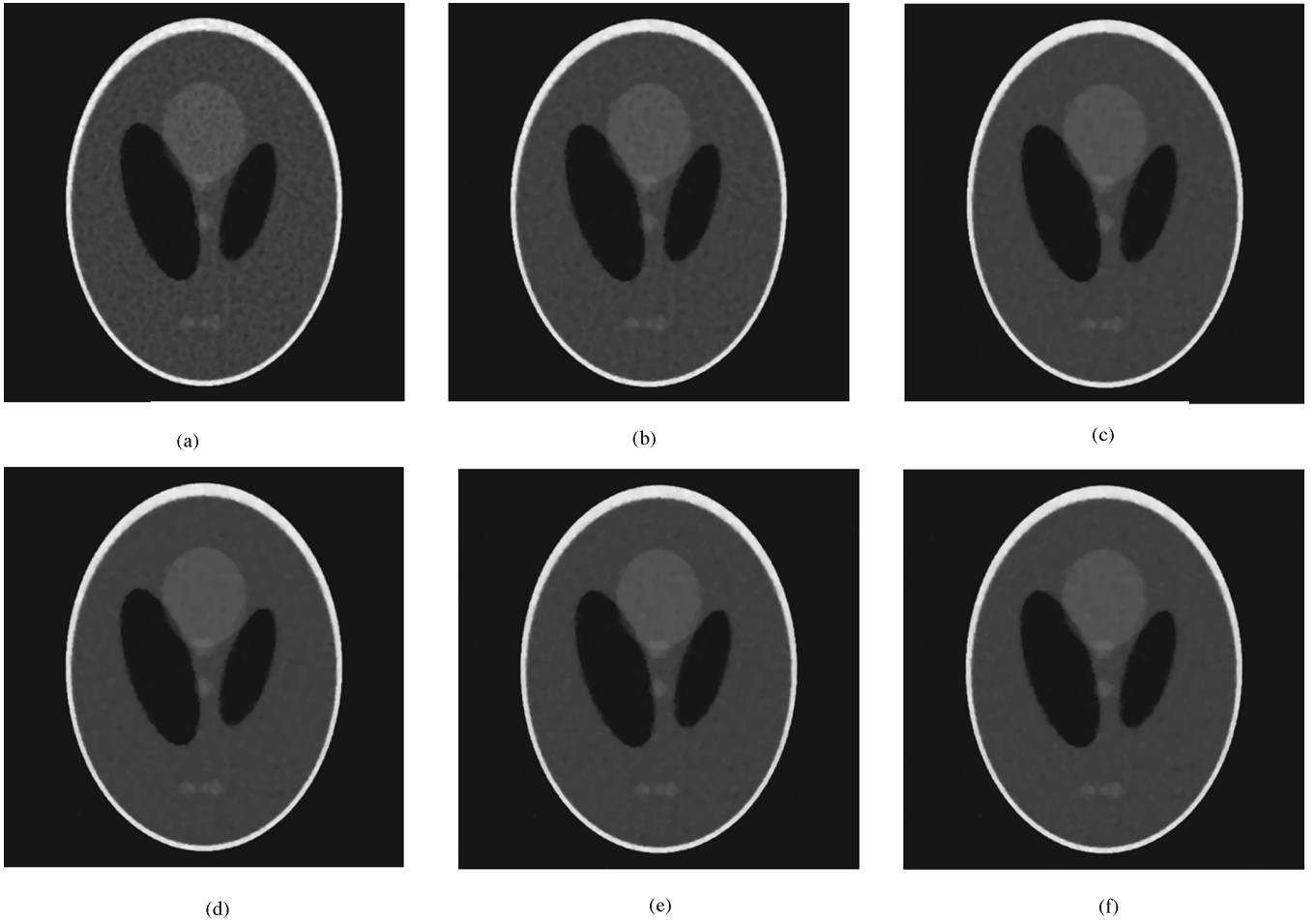


Fig. 10. Reconstructions for transmission tomography with quadratic data term: Fixed-grid algorithm with (a) seven iterations, (b) 14 iterations, (c) 28 iterations, and (d) 50 iterations. (e) Multigrid algorithm with fixed-data resolution (7.48 iterations) and (f) multigrid algorithm with variable data resolution (5.81 iterations).

We use the coarse-scale Poisson mean vector computed by

$$f_m^{(q)}(x^{(q)}) = y_T \exp(-P_{m,*}^{(q)} x^{(q)}) \quad (52)$$

where $P^{(q)}$ is once again given by (44).

Both emission and transmission cases use the same interpolation/decimation matrices and coarse-scale stabilizing functions as described in Section IV-A.

V. NUMERICAL RESULTS

In this section, we compare three algorithms: the proposed multigrid algorithms with variable data resolution; the multigrid algorithms with fixed-data resolution; and the fixed-grid ICD algorithm [66], [67]. The multigrid algorithms with fixed-data resolution here means a special case of the algorithms proposed in Sections II–IV, where all resolutions used the same data by simply using identity matrices for $J_{(q)}^{(q+1)}$ and $J_{(q+1)}^{(q)}$ for all q . We tested the algorithms for Bayesian reconstruction in emission and transmission tomography using the modified Shepp–Logan phantom [74] shown in Fig. 3(a). The width and the height of the bounding rectangle was 20 cm, and the two-dimensional region was discretized with 513×513 pixels. In the

emission case, the brighter regions correspond to higher emission; and in the transmission case, the brighter regions correspond to higher absorption, with a peak absorption coefficient of 0.05 cm^{-1} . Projection data was simulated using 180 uniformly spaced angles, each with 512 uniformly spaced projections. The projection beam was assumed to have a triangular beam profile with a width of two times the projection spacing. In the emission case the total photon count per projection data was approximately 1.68×10^6 photons. In the transmission case, the dosage y_T per ray was 800 photons. Measurements were simulated as independent Poisson random variables. The same data set was used for both the quadratic data term-based reconstruction and the Poisson model-based reconstruction.

Reconstructions were performed on 513×513 pixels. All three algorithms were initialized with the convolution backprojection (CBP) reconstructions shown in Fig. 3(b) and (c). The CBP algorithm was implemented for a generalized Hamming reconstruction filter with frequency response $H(\omega) = H_{id}(\omega)(0.5 + 0.5 \cos(\pi\omega/\omega_c))$ for $|\omega| < \omega_c$, where $H_{id}(\omega)$ is the ideal ramp filter. The cutoff frequency ω_c was chosen to yield minimum image root-mean-square error (RMSE), which was $\omega_c = 0.6\pi$ for transmission tomography and $\omega_c = 0.5\pi$ for emission tomography.

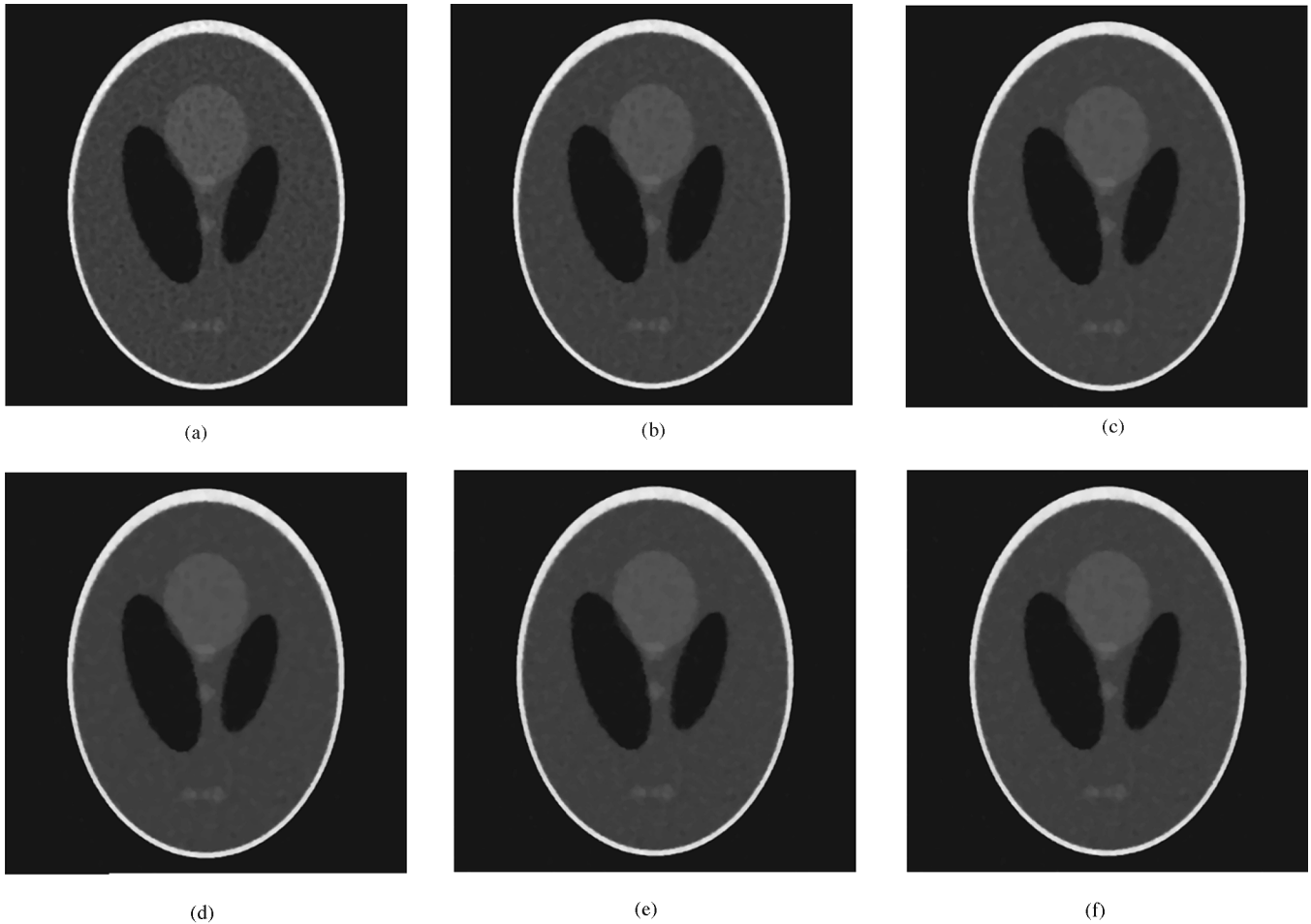


Fig. 11. Reconstructions for transmission tomography with the Poisson noise model: Fixed-grid algorithm with (a) eight iterations, (b) 16 iterations, (c) 32 iterations, and (d) 50 iterations. (e) Multigrid algorithm with fixed-data resolution (9.06 iterations) and (f) multigrid algorithm with variable data resolution (6.46 iterations).

Both multigrid algorithms used a three level multigrid-V recursion, and used the fixed-grid ICD algorithm [66], [67] with random-order pixel updates. We chose the ν parameters in Fig. 1(b), which control the number of fixed-grid update iterations at each scale, adaptively, as described in Section III. For fair comparison, we scaled the iteration number by the theoretical computational complexity. A detailed description for the conversion can be found in the Appendix. The CBP computation is not included in the computational complexity since the CBP initialization is of negligible cost compared with the ensuing computation.

The image prior model parameters used were an eight point neighborhood GGMRF with $p = 1.2$, and $b_{j-k} = (2 - \sqrt{2})/4$ for nearest neighbors and $b_{j-k} = (\sqrt{2} - 1)/4$ for diagonal neighbors. We chose the image prior variance parameter to be $\sigma = 0.0025 \text{ cm}^{-1}$ in the transmission case and $\sigma = 0.4 \text{ cm}^{-1}$ in the emission case. These values were lower than the optimal parameters yielding minimum image RMSE, but they resulted in qualitatively better reconstructions in spite of a slightly larger RMSE.

Figs. 4(a), 5(a), 6(a), and 7(a) compare the convergence speed of the algorithms in terms of the cost function. For both imaging modalities and both data likelihood functions, the multigrid algorithm with variable data resolution converged at least twice

as fast as the multigrid algorithm with fixed-data resolution. Importantly, although the convergence of the fixed-grid ICD algorithms in the initial few iterations is comparable with that of the multigrid algorithms with fixed-data resolution, they eventually require many more iterations (30 \sim 50 iterations) to reduce the cost to the value to which the multigrid algorithms with variable data resolution converged in 5 \sim 8 iterations.

Figs. 4(b), 5(b), 6(b), and 7(b) compare the convergence speed of the algorithms in terms of RMSE of reconstructed images. For all the cases, the multigrid algorithm with variable data resolution converged fastest. The fixed-grid algorithm behaved poorly at the first iteration, and it produced some salt and pepper noise by overshooting in some image pixel updates. Again, the fixed-grid algorithm required about 30 \sim 50 iterations to reduce image RMSE to the value that the multigrid algorithms converged to in 5 \sim 8 iterations. Since the convexity of the cost function excludes the possibility of being trapped into a local minimum, the difference in convergence speed is probably due to the fact that there are some error components which the fixed-grid optimization cannot effectively remove.

The convergence plots show that all the algorithms eventually converged to the same cost and RMSE, which should be a natural consequence of the convex optimization function. However, although the cost decrease rate of the multigrid algorithms and

the fixed-grid algorithm were similar for the initial iterations, the RMSE convergence results indicate that they converged following different optimization trajectories. The trajectory of the multigrid algorithms are perhaps more favorable because they yielded significantly smaller RMSE image error before full convergence.

Figs. 8 and 9 show the reconstructed images for emission tomography with the Poisson noise model and the quadratic approximation of data likelihood, respectively, and Figs. 10 and 11 show the reconstructed images for transmission tomography. For all cases, the final reconstruction quality was quantitatively and qualitatively almost the same for the three algorithms. However, the fixed-grid algorithm yielded poorer image quality even with twice or four times the computation that the multigrid methods required to converge. For example, the fixed-grid reconstructions in Fig. 9(b) and (c) with 14 and 28 iterations, respectively, were visually worse than the multigrid reconstructions with only 8.06 or 5.31 iterations, which are shown in Fig. 9(e) and (f). Fig. 12 shows corresponding region of interest (ROI) images from Fig. 9. We adjusted the contrast of Fig. 12 linearly from Fig. 9 so that the ROI images show the visual differences more clearly. Comparing Figs. 9–12 with Fig. 3, the reconstructions by all the statistical methods improve the image quality compared to the CBP reconstruction.

In summary, the proposed multigrid algorithm significantly saved computation as compared with the fixed-grid ICD algorithm initialized with the CBP reconstruction. It was also found that in intermediate iterations of the multigrid algorithms yielded images that appeared to have better subjective visual quality than the fixed-grid ICD algorithm. Since the optimization functions were all convex in the problems we considered, the fully converged reconstructions were, as expected, approximately the same for all the algorithms. However, in nonconvex optimization problems, such as ODT [34], or seismic tomography [55], multigrid algorithms may have the potential to provide better image quality than fixed-grid algorithms [34].

VI. CONCLUSION

The multigrid inversion methods with variable resolution data and image spaces were proposed. In formulating a set of optimization functions at different scales, the algorithm changes grid resolution of both measurement data space and image space, and, thus, improves computational efficiency further than the previous multigrid inversion methods which changes resolutions in the image space only. Application to conventional transmission and emission tomography problems demonstrated substantially reduced computation relative to the fixed-grid ICD algorithm and our previous multigrid inversion with fixed-data resolution.

APPENDIX COMPUTATIONAL COMPLEXITY

In this Appendix, we analyze the computational cost of the multigrid inversion algorithms. We use the number of multiplications/divisions (and the number of additional exponentiations in the Poisson transmission case) as a measure of computational complexity.

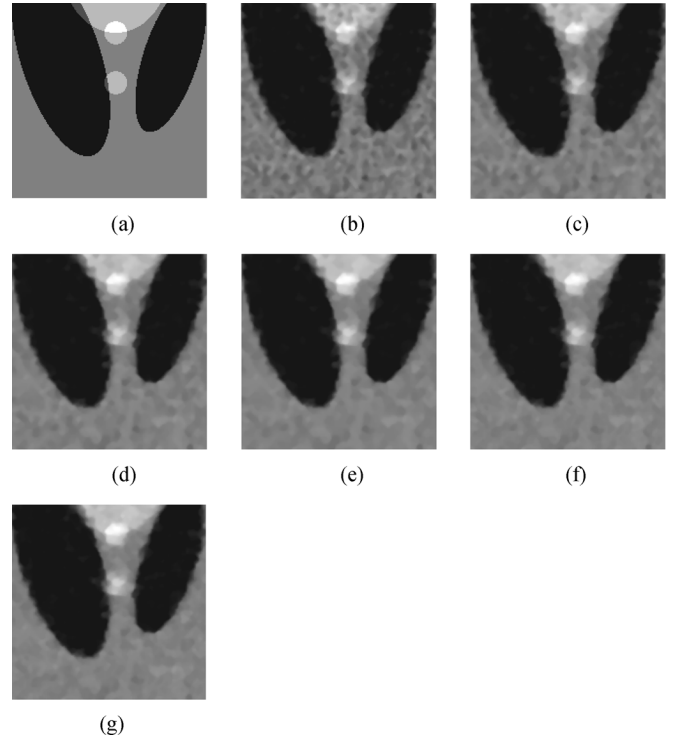


Fig. 12. ROI comparison for Fig. 9: (a) Original image; fixed-grid algorithm reconstructions with (b) seven iterations, (c) 14 iterations, (d) 28 iterations, and (e) 50 iterations. (f) Multigrid algorithm reconstruction with fixed-data resolution (8.06 iterations) and (g) multigrid algorithm reconstruction with variable data resolution (5.31 iterations).

For simplicity, we make three assumptions. First, all the data-independent vectors and matrices, such as P , Λ , and a , are pre-computed and stored. Second, the ratio M_0/M is approximately constant across resolutions, where M_0 is the average number of nonzero projections associated with each image pixel. Finally, we neglect the computational cost required for decimation and interpolation. In other words, we assume that the main computational cost at resolution q consists of the fixed-grid update on $x^{(q)}$ and the computation of $r^{(q)}$.

The ICD iteration typically has complexity of $O(M_0N)$, where N is the number of pixels. Thus, one ICD iteration at scale q requires only 16^{-q} times the computations at the finest scale for the variable data resolution case, and 4^{-q} times the computation at the finest scale for fixed-data resolution case. This is also true for the $r^{(q)}$ computation, which is computed only once when the inversion proceeds from scale q to $q+1$.

Then, in a similar manner to [34], the complexity of one MultigridV iteration is given by

$$\sum_{q=0}^{Q-2} \left[16^{-q} \times \text{Comp}_x \times \nu^{(q)} + \text{Comp}_r \right] + 16^{-(Q-1)} \times \text{Comp}_x \times \nu^{(Q-1)} \quad (53)$$

for the variable data resolution case and

$$\sum_{q=0}^{Q-2} \left[4^{-q} \times \text{Comp}_x \times \nu^{(q)} + \text{Comp}_r \right] + 4^{-(Q-1)} \times \text{Comp}_x \times \nu^{(Q-1)} \quad (54)$$

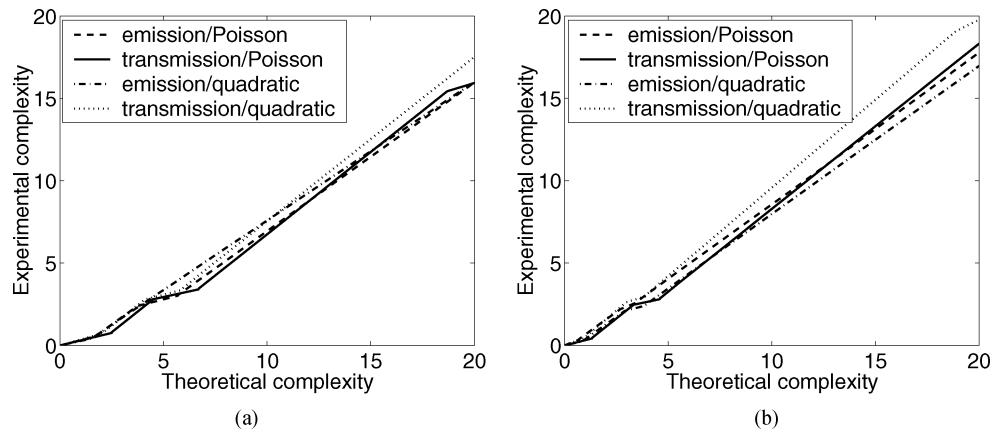


Fig. 13. Comparison between the theoretical complexity and the measure CPU time for the multigrid algorithms with (a) fixed-data resolution and (b) variable data resolution.

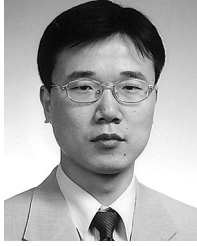
for fixed-data resolution case, where Comp_x is the complexity for one ICD iteration at the finest scale, Comp_r is the complexity for updating r vector at the finest scale, and $\nu^{(q)} = \nu_1^{(q)} + \nu_2^{(q)}$ is the number of iterations of fixed-grid update at scale q . The ratio of Comp_r to Comp_x is $2/3$ for the quadratic cases, $2/5$ for the Poisson emission case, and 1 for the Poisson transmission case, where we conservatively assume that the exponentiations dominate the complexity. The formulas (53) and (54) were used to scale the iteration number in Section V.

Fig. 13 compares the theoretical complexity, computed with (53) and (54), with a measured experimental complexity in terms of the CPU time. The experimental complexity is the elapsed CPU time divided by the average CPU time for one iteration of the fixed-grid ICD algorithm. The elapsed CPU time was measured for every iterations on all grid resolutions. It was measured on a linux machine with an AMD 2.0-GHz Athlon CPU and 2-GByte memory. The experimental complexity for multigrid algorithms was consistently a little lower than the theoretical complexity. Interestingly, we found that the coarse-scale ICD iterations took substantially shorter time than the theoretical complexity anticipates, which might be an effect of the better cache locality when solving the small-scale problem.^f

REFERENCES

- [1] M. V. Ranganath, A. P. Dhawan, and N. Mullani, "A multigrid expectation maximization reconstruction algorithm for positron emission tomography," *IEEE Trans. Med. Imag.*, vol. 7, no. 4, pp. 273–278, Dec. 1988.
- [2] T. Pan and A. E. Yagle, "Numerical study of multigrid implementations of some iterative image reconstruction algorithms," *IEEE Trans. Med. Imag.*, vol. 10, no. 4, pp. 572–588, Dec. 1991.
- [3] A. B. Milstein, S. Oh, J. S. Reynolds, K. J. Webb, C. A. Bouman, and R. P. Millane, "Three-dimensional Bayesian optical diffusion tomography using experimental data," *Opt. Lett.*, vol. 27, pp. 95–97, Jan. 2002.
- [4] S. Oh, A. B. Milstein, R. P. Millane, C. A. Bouman, and K. J. Webb, "Source-detector calibration in three-dimensional Bayesian optical diffusion tomography," *J. Opt. Soc. Amer. A*, vol. 19, no. 10, pp. 1983–1993, Oct. 2002.
- [5] A. B. Milstein, S. Oh, K. J. Webb, C. A. Bouman, Q. Zhang, D. A. Boas, and R. P. Millane, "Fluorescence optical diffusion tomography," *Appl. Opt.*, vol. 42, pp. 3081–3094, Jun. 2003.
- [6] Z. Kato, M. Berthod, and J. Zerubia, "Parallel image classification using multiscale Markov random fields," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Minneapolis, MN, Apr. 27–30, 1993, vol. 5, pp. 137–140.
- [7] C. A. Bouman and M. Shapiro, "A multiscale random field model for Bayesian image segmentation," *IEEE Trans. Image Process.*, vol. 3, no. 2, pp. 162–177, Mar. 1994.
- [8] M. L. Comer and E. J. Delp, "Segmentation of textured images using a multiresolution Gaussian autoregressive model," *IEEE Trans. Image Process.*, vol. 8, no. 3, pp. 408–420, Mar. 1999.
- [9] J.-M. Laferte, P. Perez, and F. Heitz, "Discrete Markov image modeling and inference on the quadtree," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 390–404, Mar. 2000.
- [10] R. D. Nowak, "Shift invariant wavelet-based statistical models and 1/f processes," presented at the IEEE DSP Workshop 1998.
- [11] K. Chou, A. Willsky, A. Benveniste, and M. Basseville, "Recursive and iterative estimation algorithms for multi-resolution stochastic processes," in *Proc. 28th Conf. Decision and Control*, Tampa, FL, Dec. 13–15, 1989, vol. 2, pp. 1184–1189.
- [12] H.-C. Yang and R. Wilson, "Adaptive image restoration using a multiresolution Hopfield neural network," in *Proc. 5th Int. Conf. Image Processing and its Applications*, Edinburgh, U.K., Jul. 4–6, 1995, pp. 198–202.
- [13] R. Nowak and E. D. Kolaczyk, "A multiscale MAP estimation method for Poisson inverse problems," in *Proc. 32nd Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 1–4, 1998, vol. 2, pp. 1682–1686.
- [14] R. D. Nowak and E. D. Kolaczyk, "A statistical multiscale framework for Poisson inverse problems," *IEEE Trans. Inf. Theory*, vol. 46, no. 5, pp. 1811–1825, Aug. 2000.
- [15] G. Wang, J. Zhang, and G. Pan, "Solution of inverse problems in image processing by wavelet expansion," *IEEE Trans. Image Process.*, vol. 4, no. 5, pp. 579–591, May 1995.
- [16] B. Sahiner and A. Yagle, "Image reconstruction from projections under wavelet constraints," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3579–3584, Dec. 1993.
- [17] M. Bhatia, W. C. Karl, and A. S. Willsky, "Wavelet-based method for multiscale tomographic reconstruction," *IEEE Trans. Med. Imag.*, vol. 15, no. 1, pp. 92–101, Jan. 1996.
- [18] N. Lee and B. J. Lucier, "Wavelet methods for inverting the radon transform with noisy data," *IEEE Trans. Image Process.*, vol. 10, no. 1, pp. 79–94, Jan. 2001.
- [19] A. Delaney and Y. Bresler, "Multiresolution tomographic reconstruction using wavelets," *IEEE Trans. Image Process.*, vol. 4, no. 6, pp. 799–813, Jun. 1995.
- [20] W. Zhu, Y. Wang, Y. Deng, Y. Yao, and R. Barbour, "A wavelet-based multiresolution regularization least squares reconstruction approach for optical tomography," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 210–217, Apr. 1997.
- [21] Z. Wu, G. T. Herman, and J. A. Browne, "Edge preserving reconstruction using adaptive smoothing in wavelet domain," in *Proc. IEEE Nuclear Science Symp. Medical. Imaging Conf.*, San Francisco, CA, Nov. 6, 1993, vol. 3, pp. 1917–1921.
- [22] S. S. Saquib, C. A. Bouman, and K. Sauer, "A non-homogeneous MRF model for multiresolution Bayesian estimation," in *Proc. IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, Sep. 16–19, 1996, vol. 2, pp. 445–448.
- [23] T. Frese, C. A. Bouman, and K. Sauer, "Adaptive wavelet graph model for Bayesian tomographic reconstruction," *IEEE Trans. Image Process.*, vol. 11, no. 7, pp. 756–770, Jul. 2002.

- [24] E. L. Miller, L. Nicolaides, and A. Mandelis, "Nonlinear inverse scattering methods for thermal wave slice tomography: A wavelet domain approach," *J. Opt. Soc. Amer. A*, vol. 15, no. 6, pp. 1545–1556, Jun. 1998.
- [25] A. Brandt, "Multi-level adaptive solutions to boundary value problems," *Math. Comput.*, vol. 31, no. 138, pp. 333–390, Apr. 1977.
- [26] U. Trottenberg, C. W. Oosterlee, and A. Schueller, *Multigrid*. London, U.K.: Academic, 2000.
- [27] W. L. Briggs, V. E. Henson, and S. F. McCormick, *A Multigrid Tutorial*, 2nd ed. Philadelphia, PA: SIAM, 2000.
- [28] C. A. Bouman and K. Sauer, "Nonlinear multigrid methods of optimization in Bayesian tomographic image reconstruction," in *Proc. SPIE Conf. Neural and Stochastic Methods in Image and Signal Processing*, San Diego, CA, Jul. 19–24, 1992, vol. 1766, pp. 296–306.
- [29] S. F. McCormick and J. G. Wade, "Multigrid solution of a linearized, regularized least-squares problem in electrical impedance tomography," *Inv. Probl.*, vol. 9, pp. 697–713, 1993.
- [30] V. E. Henson, M. A. Limber, S. F. McCormick, and B. T. Robinson, "Multilevel image reconstruction with natural pixels," *SIAM J. Sci. Comput.*, vol. 17, pp. 193–216, 1996.
- [31] C. R. Johnson, M. Mohr, U. Ruede, A. Samsonov, and K. Zyp, T. J. Barth, T. F. Chan, and R. Haimes, Eds., "Multilevel methods for inverse bioelectric field problems," in *Lecture Notes in Computational Science and Engineering – Multiscale and Multiresolution Methods: Theory and Applications*. Heidelberg, Germany: Springer-Verlag, 2001, vol. 20.
- [32] J. C. Ye, C. A. Bouman, R. P. Millane, and K. J. Webb, "Nonlinear multigrid optimization for Bayesian diffusion tomography," presented at the IEEE Int. Conf. Image Processing Kobe, Japan, Oct. 25–28, 1999.
- [33] J. C. Ye, C. A. Bouman, K. J. Webb, and R. P. Millane, "Nonlinear multigrid algorithms for Bayesian optical diffusion tomography," *IEEE Trans. Image Process.*, vol. 10, no. 6, pp. 909–922, Jun. 2001.
- [34] S. Oh, A. B. Milstein, C. A. Bouman, and K. J. Webb, "A general framework for nonlinear multigrid inversion," *IEEE Trans. Image Process.*, vol. 14, no. 1, pp. 125–140, Jan. 2005.
- [35] R. Gandlin and A. Brandt, "Two multigrid algorithms for inverse problem in electrical impedance tomography," presented at the Copper Mountain Conf. Multigrid Methods Copper Mountain, CO, Mar. 30–Apr. 4, 2003.
- [36] A. Brandt and R. Gandlin, "Multigrid for atmospheric data assimilation: Analysis," in *Proc. Hyperbolic Problems: Theory, Numerics and Applications*, Pasadena, CA, Mar. 2002, pp. 369–376.
- [37] A. Brandt, "Multiscale and multiresolution methods: Theory and applications," in *Multiscale Scientific Computation: Review 2001*, T. J. Barth, T. F. Chan, and R. Haimes, Eds. Heidelberg, Germany: Springer-Verlag, 2001, pp. 3–96.
- [38] A. Brandt and D. Ron, "Multigrid solvers and multilevel optimization strategies," in *Multilevel Optimization and VLSICAD*, J. Cong and J. R. Shinnerl, Eds. Boston, MA: Kluwer, 2002, pp. 1–69.
- [39] J. A. O'Sullivan and J. Benac, "Alternating minimization multigrid algorithms for transmission tomography," in *Proc. SPIE Conf. Computational Imaging II*, San Jose, California, Jan. 2004, pp. 216–21.
- [40] D. Terzopoulos, "Image analysis using multigrid relaxation methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 2, pp. 129–139, Mar. 1986.
- [41] E. Enkelmann, "Investigations of multigrid algorithms for the estimation of optical flow fields in image sequences," *Comput. Vis. Graph. Image Process.*, vol. 43, pp. 150–177, 1988.
- [42] E. Memin and P. Perez, "Dense estimation and object-based segmentation of the optical flow with robust techniques," *IEEE Trans. Image Process.*, vol. 7, no. 5, pp. 703–719, May 1998.
- [43] P. Hellier, C. Barillot, E. Mémin, and P. Pérez, "Hierarchical estimation of a dense deformation field for 3-d robust registration," *IEEE Trans. Med. Imag.*, vol. 20, no. 5, pp. 388–402, May 2001.
- [44] S. Ghosal and P. Vanek, "Fast algebraic multigrid for discontinuous optical flow estimation Center Comput. Math., Univ. Colorado, Denver, Tech. Rep. UCD-CCM-025, 1994.
- [45] M. Unser, "Multigrid adaptive image processing," in *Proc. IEEE Int. Conf. Image Processing*, Washington, DC, Oct. 1995, vol. I, pp. 49–52.
- [46] M. Arigovindan, M. Sühling, P. Hunziker, and M. Unser, "Multigrid image reconstruction from arbitrarily spaced samples," in *Proc. IEEE Int. Conf. Image Processing*, Rochester, NY, Sep. 22–25, 2002, vol. III, pp. 381–384.
- [47] D. L. Pham and J. L. Prince, "Adaptive fuzzy segmentation of magnetic resonance images," *IEEE Trans. Med. Imag.*, vol. 18, no. 9, pp. 737–752, Sep. 1999.
- [48] E. Sharon, A. Brandt, and R. Basri, "Fast multiscale image segmentation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Jun. 2000, pp. 70–77.
- [49] D. Terzopoulos, "The computation of visible-surface representations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 4, pp. 417–438, Jul. 1988.
- [50] R. Kimmel and I. Yavneh, "An algebraic multigrid approach for image analysis," *SIAM J. Sci. Comput.*, vol. 24, no. 4, pp. 1218–1231, 2003.
- [51] S. Henn and K. Witsch, "A multigrid approach for minimizing a nonlinear functional for digital image matching," *Computing*, vol. 64, pp. 339–348, 2000.
- [52] K. Zhou and C. K. Rushforth, "Image restoration using multigrid methods," *Appl. Opt.*, vol. 30, no. 20, pp. 2906–2912, Jul. 1991.
- [53] S. T. Acton, "Multigrid anisotropic diffusion," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 280–291, Mar. 1998.
- [54] L. Borcea, "Nonlinear multigrid for imaging electrical conductivity and permittivity at low frequency," *Inv. Probl.*, vol. 17, pp. 329–359, Apr. 2001.
- [55] J. Pan, L. Boschi, and A. M. Dziewonski, "Multiresolution in Seismic Tomography: Application of Multigrid Methods [Online]. Available: <http://www.seismology.harvard.edu/pan/multigrid/paper.pdf>, to be published
- [56] T. Olson and J. DeStefano, "Wavelet localization of the Radon transform," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2055–2067, Aug. 1994.
- [57] T. Olson, "Optimal time-frequency projections for localized tomography," *Ann. Biomed. Eng.*, vol. 23, pp. 622–636, 1995.
- [58] B. Sahiner and A. Yagle, "Region-of-interest tomography using exponential radial sampling," *IEEE Trans. Image Process.*, vol. 4, no. 8, pp. 1120–1127, Aug. 1995.
- [59] R. Rashid-Farrokhi, K. J. R. Liu, C. A. Berenstein, and D. Walnut, "Wavelet-based multiresolution local tomography," *IEEE Trans. Image Process.*, vol. 6, no. 10, pp. 1412–1430, Oct. 1997.
- [60] S. Y. Zhao, G. Welland, and G. Wang, "Wavelet sampling and localization schemes for the Radon transform in two dimensions," *SIAM J. Appl. Math.*, vol. 57, pp. 1749–1762, 1997.
- [61] S. Y. Zhao, "Wavelet filtering for filtered backprojection in computed tomography," *Appl. Comput. Harmon. Anal.*, vol. 6, pp. 346–373, 1999.
- [62] M. Bhatia, W. C. Karl, and A. S. Willsky, "Tomographic reconstruction and estimation based on multiscale natural -pixel bases," *IEEE Trans. Image Process.*, vol. 6, no. 3, pp. 463–478, Mar. 1997.
- [63] 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.
- [64] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 601–609, Dec. 1994.
- [65] S. Oh, A. B. Milstein, C. A. Bouman, and K. J. Webb, "Adaptive nonlinear multigrid inversion with applications to Bayesian optical diffusion tomography," presented at the IEEE Workshop Statistical Signal Processing St. Louis, MO, Sep. 2003.
- [66] K. Sauer and C. A. Bouman, "A local update strategy for iterative reconstruction from projections," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 534–548, Feb. 1993.
- [67] C. A. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Trans. Image Process.*, vol. 5, no. 3, pp. 480–492, Mar. 1996.
- [68] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 11, pp. 721–741, Nov. 1984.
- [69] E. Levitan and G. Herman, "A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography," *IEEE Trans. Med. Imag.*, vol. MI-6, no. 5, pp. 185–192, Sep. 1987.
- [70] A. J. Rockmore and A. Macovski, "A maximum likelihood approach to emission image reconstruction from projections," *IEEE Trans. Nucl. Sci.*, vol. NS-23, pp. 1428–1432, 1976.
- [71] —, "A maximum likelihood approach to transmission image reconstruction from projections," *IEEE Trans. Nucl. Sci.*, vol. NS-24, pp. 1929–1935, 1977.
- [72] S. F. McCormick, *Multilevel Adaptive Methods for Partial Differential Equations*. Philadelphia, PA: SIAM, 1989.
- [73] C. A. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. Image Process.*, vol. 2, no. 3, pp. 296–310, Jul. 1993.
- [74] S. A. Shepp and B. F. Logan, "The Fourier reconstruction of a head section," *IEEE Trans. Nucl. Sci.*, vol. NS-21, pp. 21–43, 1974.



Seungseok Oh (S'00–M'06) received the B.S. and M.S. degrees in electrical engineering from Seoul National University, Seoul, Korea, in 1997 and 1999, respectively, and the Ph.D. degree from Purdue University, West Lafayette, IN, in 2005.

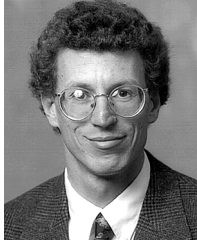
From 1999 to 2000, he was with the Hanaro Te-lococom, Seoul, Korea. In 2005, he joined the Medical Research Division of Fujifilm Software (California), Inc., San Jose, CA, as a Research Scientist. His area of research interest lies mainly within the fields of image processing. His work in the fields includes to-

mography, inverse problems, medical image analysis, and video coding.



Kevin J. Webb (S'81–M'84–SM'98–F'04) received the B.Eng. and M.Eng. degrees from the Royal Melbourne Institute of Technology, Australia, in 1978 and 1983, respectively, the M.S.E.E. degree from the University of California, Santa Barbara, in 1981, and the Ph.D. degree from the University of Illinois, Urbana, in 1984.

He is a Professor with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN.



Charles A. Bouman (S'86–M'89–SM'97–F'01) received the B.S.E.E. degree from the University of Pennsylvania, Philadelphia, in 1981, the M.S. degree from the University of California, Berkeley, in 1982, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 1989.

From 1982 to 1985, he was a full staff member at the Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge. In 1989, he joined the faculty of Purdue University, West Lafayette, IN, where he holds the rank of Professor with a primary appoint-

ment in the School of Electrical and Computer Engineering and a secondary appointment in the School of Biomedical Engineering. His research focuses on the use of statistical image models, multiscale techniques, and fast algorithms in applications, including medical and electronic imaging.

Dr. Bouman is a Fellow of the American Institute for Medical and Biological Engineering (AIMBE), a Fellow of the society for Imaging Science and Technology (IS&T), a member of the SPIE professional society, a recipient of IS&T's Raymond C. Bowman Award for outstanding contributions to digital imaging education and research, and a University Faculty Scholar of Purdue University. He is currently the Editor-in-Chief Elect of the IEEE TRANSACTIONS ON IMAGE PROCESSING, Co-Chair of the SPIE/IS&T Symposium on Electronic Imaging, Secretary of the IEEE Biomedical Image and Signal Processing Technical Committee, and a member of the Steering Committee for the IEEE TRANSACTIONS ON MEDICAL IMAGING. He has been an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING and the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. He has also been the Awards Chair for the ICIP 1998 organizing committee, general Co-Chair of the 2006 SPIE/IS&T Symposium on Electronic Imaging, Co-Chair of the SPIE/IS&T conferences on Visual Communications and Image Processing 2000 (VCIP), and a member of the IEEE Image and Multidimensional Signal Processing Technical Committee, a Vice President of Publications and a member of the Board of Directors for the IS&T Society, and he is the Founder and Co-Chair of the SPIE/IS&T conference on Computational Imaging.