Statistical Reconstruction Algorithms for Polyenergetic X-ray Computed Tomography

by

Idris A. Elbakri

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering: Systems)
in The University of Michigan
2003

Doctoral Committee:

        Associate Professor Jeffrey A. Fessler, Chair
        Senior Associate Research Scientist Neil Clinthorne
        Professor Mitchell Goodsitt
        Professor Alfred Hero

To my parents, who offered me a universe of opportunity in a world of strife.

ACKNOWLEDGEMENTS

the assistance of the administrative staff.

Learning and knowledge are inherently cumulative. In this regard, I have relied on the contributions of many others. In particular, I benefited much from the dissertations of Dr. Hakan Erdogan and Dr. Mehmet Yavuz, previous members of our laboratory.

I first joined EECS as a member of the group of Dr. Emad Ebbini. I thank him for financial support and for advising me during my first year at UM.

Getting a Ph.D. was certainly a family endeavour, and my family has endured much sacrifice for the sake of my education:

To my younger brothers, Kaiss and Ashraf, for the older brother that wasn't there;

To my wife, Bayan, for her love and support and for much missed Sunday night family time;

And to my parents, to whom my debt is, simply, infinite.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

Appendix

CHAPTER I

Introduction

1.1 Motivation

Modern medical imaging was born when Wilhelm Conrad Roentgen showed that bones could be visualized by X-raying his wife's hand in 1895. Today, radiography remains one of the major medical imaging modalities. The use of X-rays for medical imaging, however, has expanded well beyond radiography. With modern X-ray modalities we are capable of seeing 3-dimensional images rich with information about the internal structure of the human body. X-ray computed tomography (CT), first introduced by Cormack [12] and Hounsfield [40], is an X-ray modality that enables the imaging of cross-sectional slices of an object, often the human body. Today, X-ray CT continues to be one of the leading clinical imaging modalities. CT is also used to image animals, industrial parts, mummies, etc. [45].

Fundamentally, CT provides a cross-sectional image of the X-ray attenuation properties of tissues within the body. CT scanners record projection measurements of the transmission of X-ray photons through an object at different angles. There are a variety of methods to reconstruct clinically useful images from CT measurements. One approach, for example, casts the CT problem as system of equations that can be solved with numerical iterations. The most dominant CT reconstruction tech-

nique, however, is filtered backprojection (FBP). Theoretically, FBP has its roots in the Radon transform and the Fourier slice theorem, which link a function and its projections to its Fourier transform. FBP is therefore analytic, and its practical implementations take advantage of the fast Fourier transform (FFT). FBP is fast and deterministic, and its properties are very well understood.

According to the Radon transform, projections of an object represent an analytic transform for that object. Inversion of the transform gives a direct solution to the reconstruction problem. The Radon transform however, assumes ideal conditions: infinite number of measurements with infinitely thin X-ray beams, noiseless data, monoenergetic X-rays, etc. In reality, there is a host of non-ideal effects and parameters, and real CT measurements never conform to ideal conditions. CT measurements suffer from noise, beam hardening, scatter, detector and X-ray tube effects, etc. These different effects, if not corrected for, generally lead to discrepancies between the true and reconstructed values of the object. These discrepancies are often called image artifacts (see [45] for a good discussion on artifact causes and corrections).

In spite of the success of FBP in CT image reconstruction, there has been significant interest recently in a different approach to image reconstruction from transmission measurements, namely statistical iterative[1] reconstruction [3,14,25,26,34,53,54, 62,71,73,76]. Statistical techniques have several attractive features [25,34,53]. They statistically model the data noise, offering the potential for better bias-variance performance. They can also model such phenomena as scatter and energy dependence leading to more accurate and artifact-free reconstruction. Statistical methods also easily incorporate the system geometry, detector response, object constraints and prior information. They are well suited for arbitrary geometries and situations with

---

[1]Strictly speaking, it is possible to have iterative but non-statistical methods, such as [79]. In this work we use the terms 'statistical' and 'iterative' interchangeably.

truncated data. Their main drawback (when compared to FBP) is longer computation times. For clinical CT images with typical sizes of 512×512 pixels or larger, statistical methods require prohibitively long computation times which hinder their use.

The interest in applying statistical methods for CT reconstruction may have been motivated by their success in emission tomography (PET and SPECT). Contributing to this success is the fact that the maximum-likelihood expectation-maximization (ML-EM) algorithm has a closed-form expression for the emission case [55]. Also, typically, PET and SPECT have low counts. For the image sizes and resolution requirements of emission modalities, the longer computational time of statistical methods is fairly easily surmounted.

Another motivation for exploring statistical reconstruction for CT is the industry's drive towards non-Radon scanning geometries where data for several slices are acquired and reconstructed simultaneously. For example, in the cone-beam geometry a larger volume of the body is irradiated with X-rays. Moreover, the introduction of helical CT in the late eighties and early nineties [45] made the acquisition of volume CT data a clinical reality. In helical CT, the patient table is translated through the scanner gantry while the X-ray beam rotates in a circular path, tracing a helix where it intersects the patient and allowing for larger volume coverage. Measurements acquired with such geometries do not conform to the Radon transform assumptions, and require rebinning and interpolation prior to FBP reconstruction. Rebinning and interpolation degrade image resolution and introduce image artifacts. Iterative methods require no prior assumptions about the geometry of the system. Conceivably, the system matrix used in iterative image reconstruction can incorporate an arbitrary geometry such as a cone-beam system, or even a cone-beam system following

a helical path.

Because statistical methods outperform FBP in low count situations in PET and SPECT, where the SNR is low, they carry the promise of lower dose for CT patients. With FBP CT reconstruction, images produced from low dose scans suffer from noise-induced streaking [41]. Since statistical methods are based on a probability distribution for measurement noise, they tend to suppress more the contributions of low signal rays in the data, leading to better image quality. Major CT manufacturers are currently exploring iterative reconstruction, and are partly motivated by the desire to deliver lower dose scans.

FBP images require data preprocessing and image post processing to produce artifact-free images. Data corrections include tube and detector effects, first-order beam hardening effects, and other nonlinear corrections. Image post processing includes correcting for scatter, partial volume and second-order beam hardening effects. Although it is possible to apply such corrections to iterative reconstruction, such an approach does not capture the full utility and elegance of statistical techniques. As will be shown later in this work, superior results are obtained if these non-ideal effects are included in an acquisition model for CT, or what we often call the CT *forward model*. A forward model, formulated for the purpose of superior quality image reconstruction, can incorporate the effects of X-ray and attenuation physics, system design and noise statistics. The purpose of this thesis is to develop iterative algorithms for CT, and in the process study some of these non-ideal effects and 'correct' their artifacts using a realistic, albeit incomplete, CT forward model.

Statistical reconstruction methods have found utility in estimating attenuation maps for attenuation correction in emission tomography from low-count transmission scans [25, 26, 30, 34, 57]. These transmission measurements are obtained from

monoenergetic radioisotope sources, and the data acquisition models are rather simple. In some instances, authors report applying the emission EM algorithm to log transmission data [18,73]. Unlike the emission EM algorithm [9,54,55], the transmission EM algorithm does not have a closed-form solution and requires mathematical approximations [9, 54, 55]. Employing the emission EM algorithm is a suboptimal choice, since the data is inherently mismatched to the assumed model [26].

Some recent work has recognized the importance of modeling at least some aspects of the CT forward problem in iterative reconstruction. Particular attention has been paid to including the beam hardening effects of the polyenergetic X-ray spectrum emitted by the source [14,22,24,75,76]. Accounting for the effects of the polyenergetic nature of the X-ray beam is important, especially in applications where accurate quantitative results are important. An example of polyenergetic modeling for a non-statistical iterative algorithm can be found in [78,79]. In addition to beam hardening, some recent work has focused on an accurate statistical model for CT measurements [23, 74, 75]. Effective statistical modeling is critical for low-dose situations where distortions in the image due to noise become dominant.

1.2   Organization of Dissertation

This chapter discuss the background and motivation behind the work presented in this thesis. Chapter II provides a general background on X-ray CT, with a brief discussion of X-ray physics and image reconstruction. Chapter III presents our work in deriving an approximate likelihood for the compound Poisson statistical distribution. Chapter IV presents a polyenergetic statistical reconstruction algorithm that is effective at reducing beam hardening artifacts. The algorithm is generalized in Chapter V with object models based on the attenuation properties of tissues. This

generalization enables the algorithm to reconstruct mixed pixels and tissues consisting mostly of mineral solutions. In Chapter VI we present a system model for the GE LightSpeed CT scanner and present the results of a preliminary evaluation of statistical reconstruction on real LightSpeed data. We conclude in Chapter VII with a summary and an outline for possible extensions of this work.

1.3 Contributions

In this thesis we report on the application of iterative reconstruction algorithms for CT image reconstruction. The common thread throughout this work is the formulation of an effective forward model for CT. Towards that end, we present the following contributions:

- A flat-detector and arc-detector fan-beam projector added to the ASPIRE image reconstruction software package [29]. The projector is geometrically accurate. It computes the areas of overlap between X-ray beams and object pixels.

- Derivation of probability distributions and approximate likelihoods based on the compound Poisson model for CT measurements. In its most general form, the log likelihood we derive accounts for the X-ray energy-dependent statistics in scintillating detectors, random diffusion of detector light photons, and additive electronic noise [23].

- A novel polyenergetic CT iterative reconstruction algorithm [21, 22]. The algorithm is based on a Poisson likelihood function that incorporates polyenergetic attenuation physics, and is implemented using surrogate functions and ordered subsets. A version of the algorithm is monotonic.

- Polyenergetic object models that generalize the applicability of the algorithm mentioned above. These models enable the algorithm to reconstruct pixels containing tissue mixtures. The *displacement model* works well for objects consisting mostly of distinct anatomical structures with mixed boundary pixels. The *solution model* is better suited for estimating mineral solution density and may be potentially useful for bone mineral density quantitative studies [24].

- Empirical verification of the above models and algorithms with simulated and real data. Real data was acquired on an experimental cone-beam system [71] and a microCT scanner (EVS Corp., now part of GE Medical Systems) courtesy of Pfizer BioImaging Center (Pfizer Inc., Ann Arbor, MI).

- Developing an algorithm based on a forward model for the GE LightSpeed CT scanner (GE Medical Systems, Milwaukee, WI). The algorithm inherently corrects for first-order beam hardening, detector afterglow, and X-ray tube off-focal radiation effects.

CHAPTER II

X-ray Computed Tomography

2.1 Computed Tomography Basics

X-ray computed tomography is an imaging modality that produces cross-sectional images of the linear attenuation coefficient of the scanned object. Fig. 2.1 illustrates the concept of a simple CT scanner, not very different in principle from the first scanner of Hounsfield. An X-ray tube emits thin X-ray beams that traverse the slice of interest. The intensity of the X-rays diminishes according to the attenuation properties of the object, and the intensity of the attenuated emerging beams is measured by a detector on the other side of the object. This process is repeated over many angles, and the CT problem becomes one of obtaining a clinically useful image from the measurements.

There are several variations on this basic system, usually having to do with the geometry of the X-ray beam. The earliest CT scanners were parallel beam scanners. For example, in the original EMI head scanner, an X-ray source was collimated to produce a pencil beam, and the source and detector were linearly translated across the scan field, and then rotated by 1 degree to the next angular position to acquire the next set of measurements [45, 56]. The scan time was 4 1/2 minutes.

The long scan time presented many image quality challenges, especially in terms

of patient motion. The second generation of CT scanners reduced the scan time by using partial fan beams as shown in Fig. 2.2. The scan time decreased with the larger translation steps and rotation angles to a fraction of a minute [56].

In third-generation CT scanners, probably the most popular configuration for clinical CT, scan time is reduced further by eliminating linear acquisition completely. An arc-shaped array of detectors and the X-ray source form a fan-beam geometry with a scanning field of view large enough to cover the slice of interest completely, in most cases. The whole arrangement rotates around the object at very high speeds, often completing a full 360° rotation in less than a second.
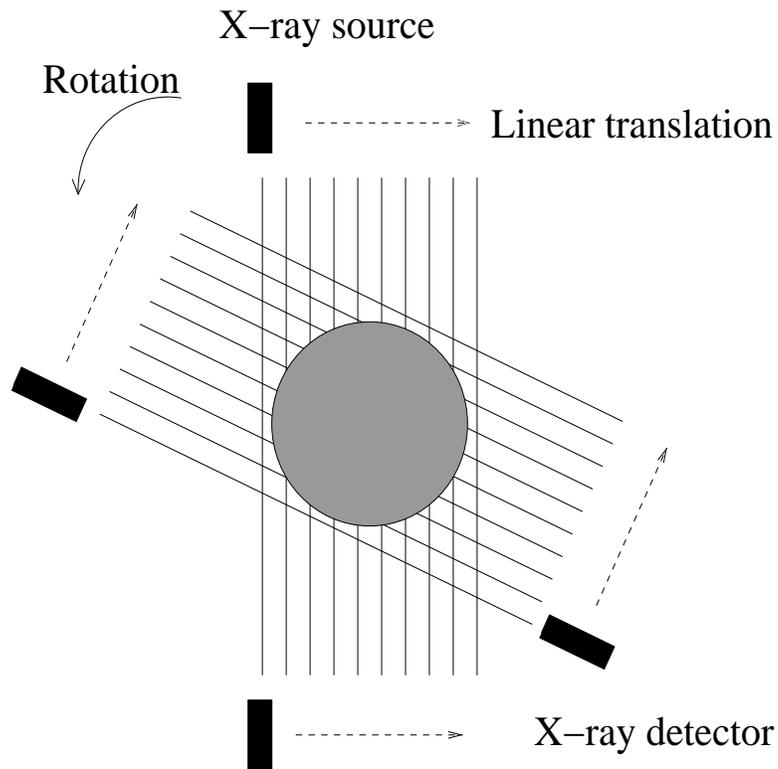
Figure 2.1: First generation CT scanner.

A schematic of a fourth-generation scanner is shown in Fig. 2.4. In this design, a rotating fan-beam impinges on a 360° stationary arc detector. This system offers
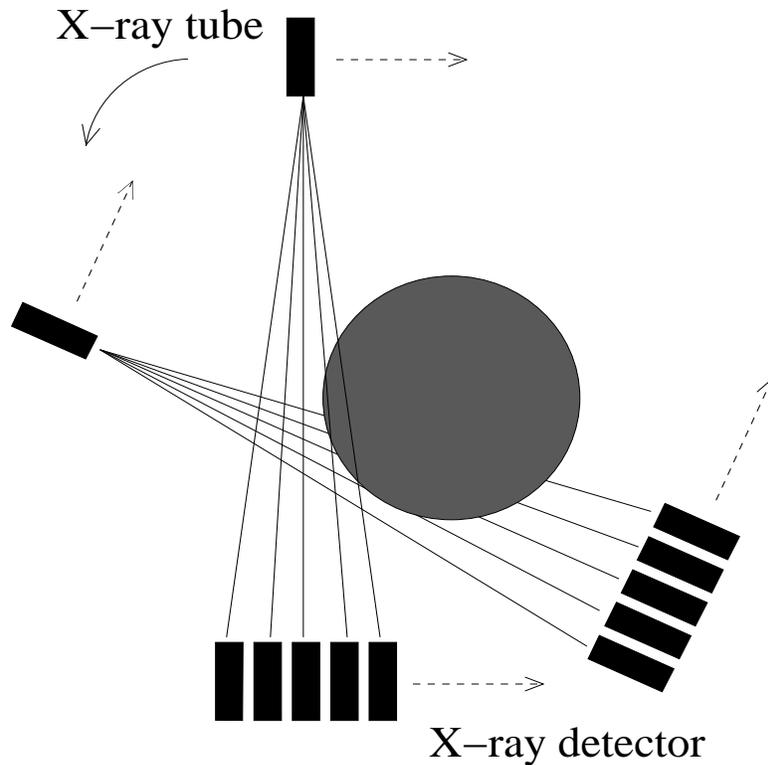
Figure 2.2: Second generation CT scanner.

better stability by allowing dynamic calibration during the scan. It also provides higher sampling density, since the sampling rate is not limited by the detector cell size, but by the rate at which measurements are taken [45]. The system is also more immune to ring artifacts [56]. The major drawbacks of this arrangement are the very high cost of the detector array and the difficulty of scatter rejection [45].

In the following sections, we will discuss some of the major components and related functions of a typical clinical CT gantry. This includes the X-ray tube and X-ray generation, X-ray detection, and CT gantry. Before that however, we will provide a brief review of X-ray physics.
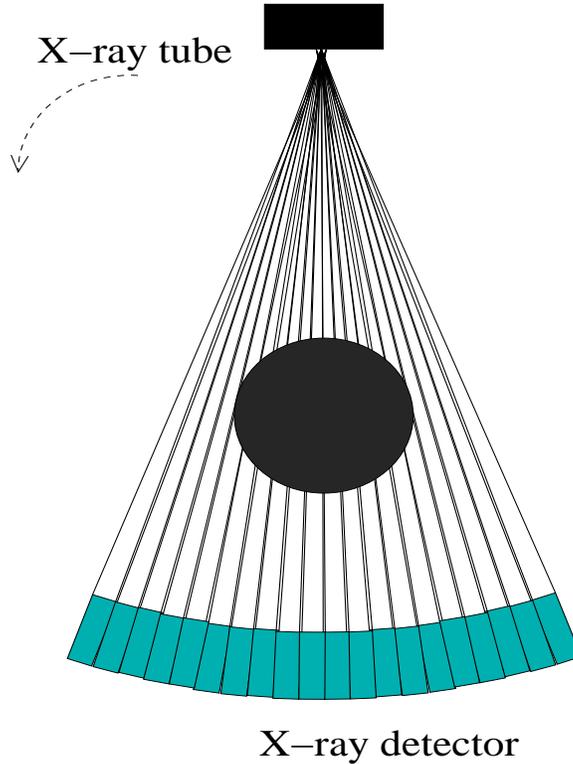
Figure 2.3: Third generation CT scanners

2.2 X-ray Physics

2.2.1 X-ray Production

X-rays belong to the family of electromagnetic radiation. The wavelength of X-rays ranges from a few picometers to a few nanometers. Diagnostic X-rays have a wavelength range of 0.1 nm to 0.01 nm. The energy of each X-ray photon is inversely proportional to its wavelength, and may be described by the following equation:

$$E = \frac{hc}{\lambda},$$
(2.1)

where h is Planck's constant (6.63×10^{-34} J s), c is the speed of light (3×10^8 m/s), and λ is the X-ray wavelength. X-ray energy is often expressed in units of electron volts (1 eV = 1.602×10^{-19} J).

X-ray photons are produced when high-speed electrons bombard a target ma-
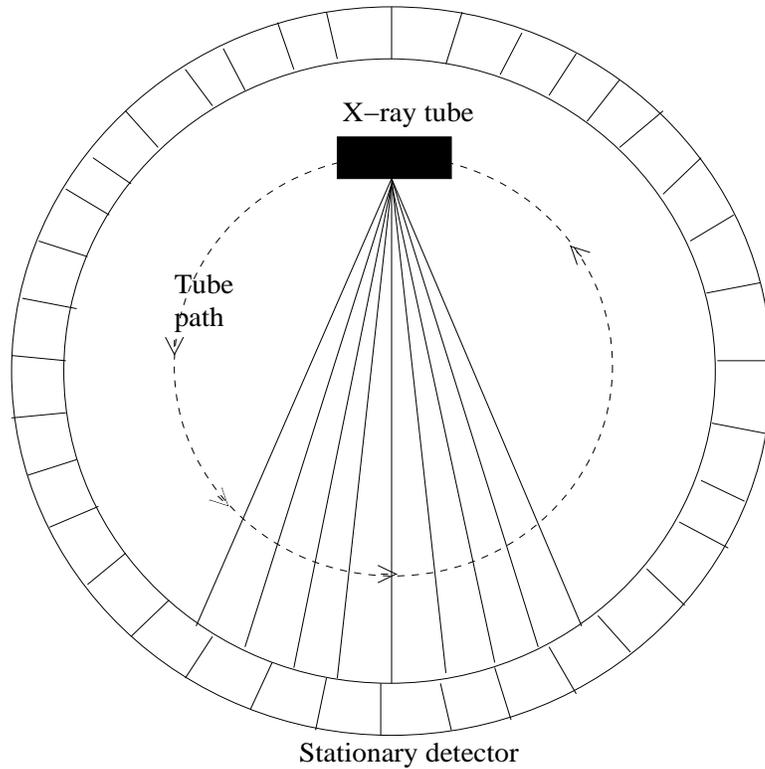
Figure 2.4: Fourth generation CT scanner.

terial. This interaction between electrons and the target takes the form of several kinds of collisions. The majority of these collisions involve energy transfer that leads to heat dissipation in the target. In fact, over 99% of the input energy dissipates as heat.

Two types of interactions lead to the production of X-rays. First, high-speed electrons traveling in the vicinity of positive nuclei experience sudden deceleration because of the attraction between the opposite charges. This produces bremsstrahlung radiation, covering a wide range of energies. It is also possible to produce bremsstrahlung radiation from the collision of an electron with a nucleus. The entire energy of the electron appears as bremsstrahlung. Energies produced by this interaction are on the upper limit of possible energies, and the overall probability

of such collisions is low.

A second type of radiation is emitted when a high-speed electron collides with and liberates an inner-shell electron of the target atom, leaving a hole in the shell. Characteristic radiation results when an electron from an outer shell fills the hole. The energy of the emitted radiation is the difference between the binding energies of the two shells.

For further reading on X-ray production, see [5, 45, 56].

2.2.2   X-ray Interaction with Matter

The linear attenuation coefficient of a material depends on the photon energy and the elemental composition of the material. In the diagnostic range of X-ray energies (below 150 keV), three mechanisms of X-ray interaction with matter dominate attenuation [5, 45, 56].

The first is coherent (or Rayleigh) scattering. The incident radiation excites electrons that in turn produce radiation at the same wavelength. Coherent scattering is material dependent and its cross section varies with Z^4 and $\mathcal{E}^{-3}$ (Z is the atomic number and $\mathcal{E}$ is the energy). Since high-Z materials are not found in abundance in the body, this effect is not very strong in the CT diagnostic range, and leads to slight broadening of the X-ray beam.

The second mechanism of X-ray absorption is the photoelectric effect, which dominates at lower energies. The X-ray photon is absorbed by interacting with and releasing a tightly bound electron. An outer-shell electron quickly fills the hole left by the ejected electron, and emits characteristic radiation in the process. At low energies, this radiation does not travel very far in tissue, and we can assume that the X-ray photon is completely absorbed.

Another important mechanism for tissue attenuation is Compton scattering where an X-ray photon collides with a free or loosely-bound electron in an outer shell. The X-ray photon is deflected at some angle (between 0 and 180 degrees) and suffers a partial loss of its energy. The scattered X-ray photon is still within the diagnostic energy range, and may exit the patient and be recorded by the detector. Scattered photons do not follow the assumed X-ray propagation geometry, and therefore are detrimental to image quality. Nearly all medical CT scanners try to minimize this effect by collimation or correction algorithms. The reader is encouraged to explore the literature for more in-depth description of these interaction mechanisms [5,45,56].

2.3 Major Components of CT Scanners

2.3.1 X-ray Tube

The basic components of the X-ray tube are the cathode and anode. Thermionic emission boils electrons off the cathode. The electrons are accelerated across the tube by a high potential difference between the cathode and anode, and collide at high speeds with the target area of the anode. This collision produces X-rays, as discussed earlier. The energy of the generated X-rays depends on the electric potential difference between the cathode and anode.

A glass or metal frame houses the cathode and anode in a vacuum. The production of X-rays is very inefficient, with over 99% of energy lost as heat. The temperature of the target can reach $2600^{o}C$. To prevent target melting, the anode rotates at very high speeds, bringing cooler areas of the target under the electron beam. Currently, tube technology offers anode rotation speeds between 8,000 and 10,000 rpm and constant X-ray production (as opposed to pulsed X-rays as in the early days of CT) for the duration of the CT scan.

Of consequence in image quality is the focal spot size and shape. In order to increase the target impact area, the focal track is at a shallow angle, typically 7°. This effect, known as the line focus principle [5, 45] increases the exposure area, but poses the problem of the focal spot size and shape becoming location dependent, especially when the viewing location is not perpendicular to the focal line.

Another tube effect that impacts image quality is off-focal radiation, caused mainly by secondary electrons [45]. Secondary electrons are emitted from the impact area of the target, and they in turn can return to the target at points outside the focal spot and produce X-rays upon impact. This results in a beam profile consisting of a high-intensity center spot surrounded by a low-intensity halo. This effect can cause degradation in low-contrast detectibility as well as shading artifacts. Although this effect can be reduced with filtration, data correction is often necessary to avoid subtle but significant image artifacts. We will discuss off-focal radiation correction in a later chapter.

2.3.2   X-ray Detector

There are several X-ray detector technologies. Third generation CT scanners use either inert gas or solid-state detectors. There are direct and indirect detection systems. A direct detector records the electrical charge which results from ionization of atoms in the detector. Gas detectors such as Xenon detectors are examples of direct detection [5]. Although low in cost, inert gas detectors suffer from low detection quantum efficiency (DQE) and are difficult to build in two dimensions for multislice CT. In this section we will focus our discussion on solid-state detectors. Description of inert gas detectors can be found in [45].

Indirect solid-state detectors are usually made of small blocks of scintillating

materials. An incident X-ray photon undergoes photoelectric interaction with a phosphor scintillator. Photoelectrons released travel a short distance and excite other electrons. When the excited electrons return to their ground state, they emit characteristic radiation, often in the visible or UV range. Reflective material coating the scintillator directs emitted light towards photodiodes which produce an electric signal [5, 45]. The DQE of a solid-state detector can be as high as 99% [45].

Collimators plates are often placed in front of detectors in third generation CT scanners to reject scattered photons. Since the collimator plates cover a portion of the detector area, they tend to reduce the overall efficiency of the detector.

Several detector parameters can affect image quality. These include the size and spacing of detector elements, hysteresis or radiation damage, thermal fluctuations and exponential decay of signal. Detector design often takes into consideration radiation damage and thermal fluctuations. Data pre-correction eliminates the effects of exponential decay (primary speed and afterglow) [43] and will be discussed later.

2.3.3 Scanner Gantry

Another important component of the CT scanner is the gantry. All major scanner components, weighing hundreds of pounds, are mounted on the gantry, which has to rotate at very high speed with extreme stability. The gantry must maintain angular and position accuracy, and must be free of significant vibrations that can cause image artifacts. Current slip ring technology enables collection of 1000 views in about 0.5 seconds. The gantry can also tilt to acquire slices at oblique angles.

2.4 Image Reconstruction

In this section we discuss the fundamentals of CT image reconstruction. We present the classical approach of filtered backprojection (FBP), since it is what is

used in clinical practice, and defer statistical reconstruction for in depth analysis in later chapters. We discuss the mathematical tools used to derive FBP and the underlying assumptions. For the sake of simplicity, we restrict our description to the parallel-beam geometry. Fan beam and more general geometries require some modifications that are readily available in the literature [4, 17, 27, 38, 61].

2.4.1 Simplifying Assumptions

The relationship between a two-dimensional function $\mu(x, y)$ and its complete collection of line integrals $\{p(R, \theta) : \theta \in [0, 2\pi], R \in [0, \infty)\}$ is called the Radon transform. The two dimensional image reconstruction problem is to find $\mu(x, y)$ from $p(R, \theta)$. Expressed mathematically, the line integral $p(R, \theta)$ can be written as

$$
\begin{aligned}
p(R, \theta) &= \int_{L(R,\theta)} \mu(x, y)dl \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu(x, y)\delta(x\cos(\theta) + y\sin(\theta) - R)dxdy \\
&= \int_{-\infty}^{\infty} \mu(x, y)\delta(R\cos(\theta) - l\sin(\theta), R\sin(\theta) + l\cos(\theta))dl \\
&= \int_{0}^{2\pi} \int_{0}^{\infty} \mu(r, \phi)\delta(r\cos(\theta - \phi) - R)rdrd\phi, \quad (2.2)
\end{aligned}
$$

where we switched to polar coordinates in the last equation. Fig. 2.5 illustrates the relationship between a two-dimensional function and its line integral. The set of line integrals is often referred to as the sinogram. If the line integrals of a point are stacked in an array with R varying along one axis and θ varying along the other, then the projection function will appear like a sinusoid, hence the the name sinogram.

CT measurements do not correspond directly to $p(R, \theta)$. Taking into account the polyenergetic spectrum and scatter, we can express CT measurements in the following manner:

$$
Y(R, \theta) = \int I_o(\mathcal{E})e^{-\int_{L(R,\theta)} \mu(x,y;\mathcal{E})dl}d\mathcal{E} + r(R, \theta) \quad (2.3)
$$

where $Y(R, \theta)$ is the recorded measurement at coordinates (R, θ) in the projection domain, $I_o(\mathcal{E})$ represents the source spectrum and detector response, $L(R, \theta)$ is the path of the ray of interest, $\mu(x, y)$ is the attenuation coefficient, $\mathcal{E}$ is energy and r represents the contribution of scatter.

Suppose it is possible to design an ideal scanner, where the X-rays are monoenergetic ($I_o(\mathcal{E}) = I_o \delta(\mathcal{E} - \mathcal{E}_o)$) and scatter is zero ($r(R, \theta) = 0$). Equation (2.3) simplifies to:

$$Y(R, \theta) = I_o e^{-\int_{L(R,\theta)} \mu(x,y;\mathcal{E}_o)dl} = I_o e^{-p(R,\theta)}. \tag{2.4}$$

This relationship between the measurement $Y(R, \theta)$ and the line integrals of the attenuation coefficient is the so-called Beer's law. The line integrals can be recovered under these ideal conditions by a simple log operation.

Analytical reconstruction techniques (such as FBP) have been developed for line integrals of two-dimensional functions. For this reason, the assumptions leading up to Beer's law are convenient. They allow us to 'retrieve' the line integrals by a simple log operation. As explained earlier, CT measurements never follow Beer's law. In addition to the constraints of the actual CT scanner design (finite detector size and X-ray beam width, target focal spot size, etc.), physical effects such as beam hardening and scatter contribute to the deviation of CT measurements from Beer's law.

Assuming Beer's law holds, a powerful mathematical tool, the Fourier slice theorem, makes it possible to reconstruct the two-dimensional image from its line integrals using the Fourier transform. This theorem is the subject of the next section.

2.4.2 Fourier Slice Theorem

The Fourier slice theorem, also known as the central slice theorem is the basis of (analytic reconstruction in) tomography. Simply put, the theorem states that **the one-dimensional Fourier transform of a parallel projection of an image $\mu(x, y)$ taken at angle θ equals a slice of the two-dimensional Fourier transform of the image, $\mathcal{M}(u, v)$, at an angle θ with the u-axis** [45, 52]. In other words, the Fourier transform of $p(R, \theta)$ gives the values of $\mathcal{M}(u, v)$ along a line in the (u, v) plane defined by a positive angle θ from the u-axis:

$$\mathcal{F}\{p(R, \theta)\} = \mathcal{P}(u) = \mathcal{M}(\rho \cos(\theta), \rho \sin(\theta)),$$

where $\mathcal{F}$ is the Fourier transform operator and ρ is the radial coordinate in the Fourier plane. The proof of the theorem is straightforward. Consider the case of the $\theta = 0$ projection of $\mu(x, y)$. The projection $p(R = x, 0)$ is related to the original function by the equation:

$$p(R = x, 0) = \int_{-\infty}^{\infty} \mu(x, y) dy.$$

Taking the Fourier transform with respect to x on both sides gives:

$$\mathcal{P}(u) = \int_{-\infty}^{\infty} p(x, 0) e^{-j2\pi ux} dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu(x, y) e^{-j2\pi ux} dx dy.$$

The two-dimensional Fourier transform of $\mu(x, y)$ is given by:

$$\mathcal{M}(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu(x, y) e^{-j2\pi(ux+vy)} dx dy.$$

$$\mathcal{M}(u, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu(x, y) e^{-j2\pi ux} dx dy = \mathcal{P}(u).$$

We therefore conclude that $\mathcal{P}(u) = \mathcal{M}(u, 0)$. Because the coordinate system is selected arbitrarily, the conclusion holds for all angles. Fig. 2.5 illustrates the theorem

schematically. It is possible to derive the Fourier slice theorem more formally using coordinate system rotations. The reader is referred to the literature for details [45, 52, 56].
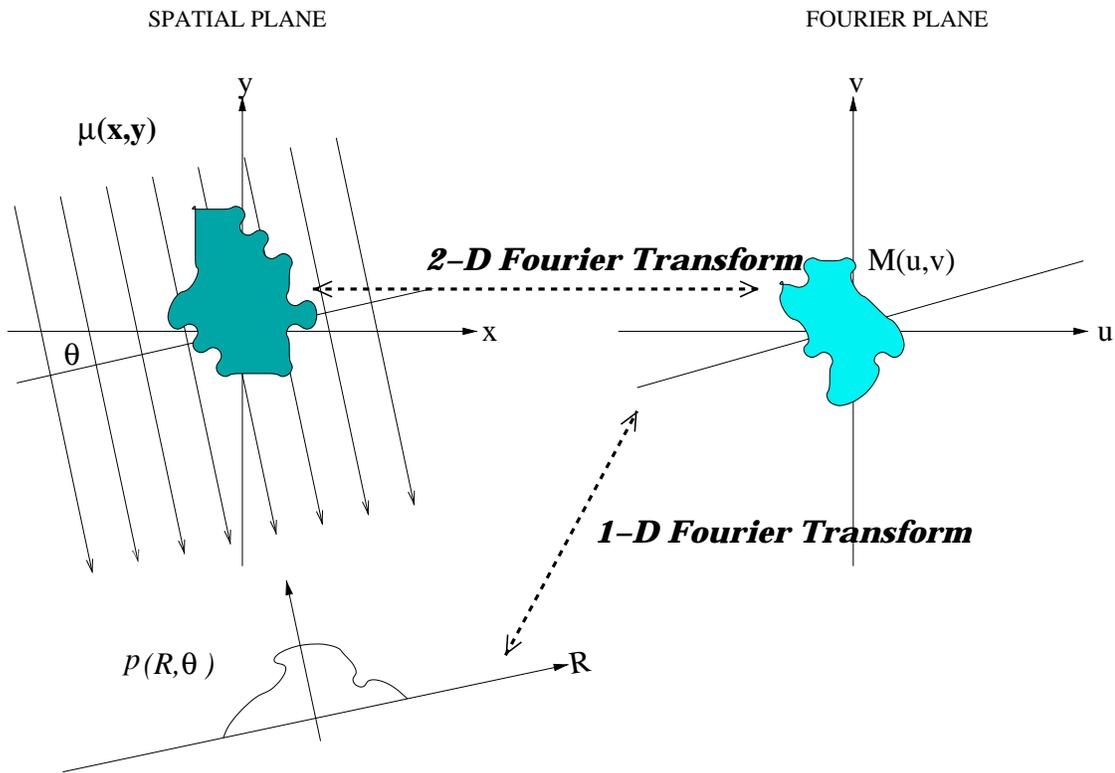
Figure 2.5: The Fourier slice theorem.

2.4.3   Filtered Backprojection

The Fourier slice theorem provides a straightforward method for tomographic reconstruction. The first step in a reconstruction algorithm based on the Fourier slice theorem is to take the one-dimensional Fourier transform of the projection at each angle. The set of one-dimensional Fourier transforms is then used to fill up the two-dimensional Fourier transform. A inverse two-dimensional Fourier transform gives the original image.

Practically speaking, direct Fourier reconstruction faces significant implementa-

tion challenges. The sampling pattern produced in Fourier space is non-Cartesian. Gridding is necessary to produce a Cartesian grid in order to take the inverse Fourier transform. Gridding or interpolation in the Fourier domain is non trivial. It is computationally demanding. Gridding errors in Fourier space are not localized in the spatial domain, and can cause artifacts across the whole image. The two-dimensional Fourier transform is also computationally non-trivial.

Filtered Backprojection (FBP) continues to be the standard method for tomographic reconstruction. FBP takes advantage of the Fourier slice theorem, but avoids gridding and two-dimensional Fourier transforms. It can be implemented efficiently with the FFT algorithm. In this section we derive FBP for the parallel beam geometry.

FBP derivation starts by looking at the expression for the inverse Fourier transform of the image [45, 52]:

$$\mu(x,y) = \int \int \mathcal{M}(u,v)e^{j2\pi(ux+vy)}dudv. \tag{2.5}$$

The first step towards manipulating (2.5) so that we can take advantage of the Fourier slice theorem is to transform the coordinate system into polar coordinates (R, θ):

$$u = \rho \cos \theta,$$

$$v = \rho \sin \theta.$$

The Jacobian of this transformation is $\rho d\rho d\theta$. Substituting this transformation in (2.5) gives

$$\mu(x,y) = \int_0^{2\pi} \int_0^\infty \mathcal{M}(\rho \cos \theta, \rho \sin \theta)e^{j2\pi\rho(x \cos \theta+y \sin \theta)}\rho d\rho d\theta.$$

The Fourier slice theorem enables us to replace $\mathcal{M}(\rho \cos \theta, \rho \sin \theta)$ with $\mathcal{P}(\rho, \theta)$. We can also take advantage of the symmetry properties of parallel-geometry projections

(namely, $\mathcal{P}(\rho, \theta + \pi) = \mathcal{P}(-\rho, \theta)$). After straight forward manipulations we can write the equation above as:

$$\mu(x, y) = \int_0^\pi \int_0^\infty \mathcal{P}(\rho, \theta)|\rho|e^{j2\pi\rho(x\cos\theta + y\sin\theta)}d\rho d\theta. \tag{2.6}$$

The inner integral (over ρ) in (2.6) is the inverse Fourier transform of $\mathcal{P}(\rho, \theta)|\rho|$. The outer integral is a backprojection. In other words, under ideal conditions, to reconstruct an image from parallel projections we can use the following steps:

1. Retrieve the line integrals by taking the log of the transmission data.

2. Take 1-dimensional Fourier transform of each line integral.

3. Multiply the Fourier transform by the ramp function $|\rho|$.

4. Take the inverse Fourier transform of the filtered line integrals.

5. Backproject the filtered line integrals.

There are several issues to consider in terms of practical computer implementations of FBP. An ideal ramp filter does not really exist. In reality the ramp filter must be set to zero beyond a certain cutoff frequency. This bandlimiting operation can lead to a small but significant DC bias that has to be corrected. In addition, the Fourier transform is typically implemented using the FFT operation. FFT is an efficient implementation of the discrete Fourier transform, which corresponds to periodic convolution in the spatial domain. To avoid aliasing artifacts, the sinogram is typically zero-padded prior to the Fourier transform and filtering operation. An important property of the ramp filter is that it tends to amplify high frequency noise. The ramp filter is often apodized to reduce noise. Designing effective filters that retain small features of the image while keeping noise to an acceptable level is

a challenging and involved process. The GE LightSpeed CT scanner (GE Medical Systems, Milwaukee, WI) offers six different types of reconstruction kernels that are useful for scanning different parts of the human body [45].

2.5   Image Display

The visualization unit used for CT image displays is called the Hounsfield unit, HU. It is also sometimes referred to as the 'CT number'. The HU is defined as:

$$HU = \frac{\mu - \mu_{\text{water}}}{\mu_{\text{water}}} \times 1000. \tag{2.7}$$

The linear attenuation coefficient is magnified by a factor of 1000. Air, with $\mu \approx 0$, has an HU value of -1000. Water is at 0 HU. Body tissues can have values ranging from a few HU to thousands.

Because the dynamic range is so large, it is often necessary to modify the gray level scale to display only a small range of HU values that correspond to the tissue of interest. The CT display window is described by its window level and window width. The window width represents that range of HU that the gray scale displays, and the window level is the mid point of that range. A typical soft tissue display window may have a window level of 20 HU and a window width of 100 HU. This means that the gray scale is displaying from -30 to 70 HU.

2.6   Conclusions

This chapter presented a brief overview of computed tomography physics, system components and conventional image reconstruction basics. Understanding X-ray physics and the operation of different scanner components will be important in following chapters as we develop a CT system model for statistical reconstruction. An

understanding of conventional image reconstruction is important for comparing the performance of statistical reconstruction with FBP reconstruction.

The discussion in this chapter was not extensive. Two important effects that will be discussed later are beam hardening and scatter. Beam hardening artifacts result from ignoring the polyenergetic nature of X-rays in the reconstruction algorithm. Scatter artifacts result from X-ray photons that are deflected in the object but still make it to the detector and are recorded. In the following chapters, we will derive algorithms for the polyenergetic reconstruction problem that can accommodate a scatter estimate if one is available. Accurate scatter estimation is an important extension of the work presented in this thesis.

CHAPTER III

X-ray CT Measurement Statistics and Likelihood

3.1 Introduction

Accurate statistical modeling forms the foundation of statistical iterative reconstruction. The statistical model leads to a cost function that is optimized by an iterative algorithm under certain constraints. In emission tomography, where detectors count individual quanta, the Poisson distribution accurately models data statistics, and the Poisson log-likelihood is used routinely for image reconstruction. The Poisson model and likelihood are also frequently used for X-ray CT. In practice, however, most CT detectors are not quanta counters and measurement statistics are dependent on the energy profile of the X-ray beam, which is usually polyenergetic.

In modern CT scanners, the X-ray source generates a polyenergetic flux of X-ray photons. The X-ray photons that are detected are converted to light photons that in turn produce photoelectrons. The current associated with these electrons is integrated and recorded digitally by an A/D channel. The number of light photons generated, and hence the recorded signal, depends on energies of the detected X-ray quanta. The energy dependence of the measurements implies that for a polyenergetic source, measurements resulting from photons at different energies will have different statistics [74]. X-ray quanta, as they individually interact with the detector, will

lead to Poisson statistics, but the overall recorded signal will not be Poisson. In fact, measurements resulting from a polyenergetic source follow compound Poisson statistics [74].

In this chapter[1], we formulate a statistical model for X-ray CT measurements based on the compound Poisson distribution. The log-likelihood for compound Poisson statistics involves infinite series and appears impractical for maximum likelihood reconstruction. We will develop approximations to the compound Poisson log-likelihood, and compare them to the ordinary Poisson likelihood and numerically-computed exact likelihood. The model we derive includes the effect of polyenergetic spectrum, Poisson light statistics and additive Gaussian noise. We derive approximations to the compound Poisson likelihood using mathematical approximations similar to the saddle-point integration method.

Section 3.2 outlines the compound Poisson process for X-ray CT detection and derives a general form for its moment generating function. Section 3.3 outlines the saddle-point approximation to integrals, which is the mathematical tool we use to approximate the compound Poisson likelihood. Sections 3.4 and 3.5 build on the saddle-point approximation in deriving log-likelihoods for the case of monoenergetic and polyenergetic X-rays, respectively. In Section 3.5.1 we incorporate the effect of additive electronic Gaussian noise and in Section 3.6 we discuss an interpretation of the saddle point in terms of the measurements moments. In Section 3.7 we present some preliminary results and in Section 3.8 conclude with a summary and discussion of future work.

---

[1]The results presented in this chapter are based on [23].

3.2 Compound Poisson Statistics

Consider the problem of photon detection in X-ray CT from its most basic principles. We focus on a single detector element and assume that the measurements for different detectors are statistically independent. Discrete photons collide with a scintillating detector, and are absorbed according to the energy-dependent detector quantum efficiency. The absorbed X-ray photons each generate an energy-dependent number of light photons. The number of incident X-ray photons is denoted by the Poisson random variable N. The number of light photons generated by each X-ray photon that is detected is also a random variable with probability mass function (p.m.f.) $P_X(x)$. We list the random variables for the sake of clarity:

- N is the Poisson random variable with unknown mean $\bar{N}$ that describes the number of X-ray photons that interact with the detector.

- X_n is a discrete random variable with p.m.f. $P_X(x)$ that denotes the number of light photons[2] generated and recorded when the nth X-ray photon interacts with the scintillator. We assume that light generation caused by an X-ray photon does not disturb subsequent scintillations, hence $\{X_n\}$ are independent and identically distributed (i.i.d).

- Y is a discrete random variable that is proportional to the total number of recorded light photons generated by the N X-ray photons that interact with the detector.

---

[2]In the case of direct detection systems, X-ray photons generate photoelectrons that form the basis of the recorded signal. The analysis in this paper applies equally to both situations, with "light photons" replaced with "photoelectrons".

Expressed mathematically,

$$Y = A \sum_{n=1}^{N} X_n \tag{3.1}$$

$$\bar{y} \triangleq E[Y] = A\bar{N}E[X], \tag{3.2}$$

where $E[\cdot]$ is the expectation operator and A denotes the overall gain factor of the recording system. For simplicity, we assume $A = 1$ hereafter, so Y denotes the total number of light photons recorded. In [74], Whiting derived a continuous probability distribution function (p.d.f.) for compound Poisson CT measurements using characteristic functions and Fourier transforms. In this work we develop most of our techniques in the discrete domain using the Z-transform. Our goal is to derive the log-likelihood from the p.m.f. of Y, by deriving its moment generating function, $g_Y(z)$.

Using iterated expectations and the properties of moment generating functions,

$$
\begin{aligned}
g_Y(z) &= E[z^Y] = E_N[E_Y[z^Y|N]] \\
&= E_N[\Pi_{n=1}^{N} E[z^{X_n}]] = E_N[E[z^X]^N] \\
&= E_N[(g_X(z))^N] \\
&= \sum_{n=0}^{\infty} (g_X(z))^n P(N = n) \\
&= \sum_{n=0}^{\infty} (g_X(z))^n \frac{e^{-\bar{N}} \bar{N}^n}{n!} \\
&= \exp(-\bar{N}(1 - g_X(z))),
\end{aligned}
\tag{3.3}
$$

where $g_X(z) = E[z^X]$ is the moment generating function of X. This result is the same as that derived by Feller [28] for the moment generating function of a compound Poisson process. Note that

$$E[Y] = g_y'(1) = \bar{N} g_x'(1) = \bar{N} E[X].$$

Ideally, we would determine the p.m.f. $P_Y(y)$ using the inverse Z-transform of $g_Y(z)$, given by the classical contour integral [63]:

$$P_Y(y) = \frac{1}{2\pi j} \oint_c z^{-y-1} g_Y(z) dz, \qquad (3.4)$$

where c is a counterclockwise closed contour that encloses the origin and that lies completely in the region of convergence (ROC) of $g_Y(z)$. This inverse is often mathematically intractable. For deriving the p.m.f. of CT measurements, we show in Appendix B that contour integration leads to likelihood expressions with infinite series. We therefore explore a mathematical approximation that gives a more practical likelihood expression. In the next section we briefly discuss saddle-point approximation and integration, which will form the basis for approximating the p.m.f in later sections.

3.3 Saddle-Point Approximation and Integration

In this section we present a generalized version of the saddle-point integration method. The saddle-point method [6, 39, 67, 80] is useful for approximating integrals that can be expressed in the exponential form

$$\frac{1}{2\pi j} \oint_c e^{\Phi(z)} dz, \qquad (3.5)$$

where z is complex and the integral is along an appropriate closed contour c in the complex plane. In the saddle-point approximation, the exponent in the integrand is expanded in a Taylor series around a real stationary (saddle) point (assuming one exists), defined to be a root of the derivative of Φ. The first derivative term in the Taylor expansion vanishes, and the second derivative term is the highest that is retained.

More generally, we will expand the exponent around a real point that is not necessarily exactly equal to the saddle point (but is close to it in practice). We retain both the first and second derivative terms of the Taylor series. Let z_o be real in the region of convergence of the integral in (3.5), then we can write

$$
\begin{aligned}
\exp[\Phi(z)] &= \exp\left[\Phi(z_o) + \Phi'(z_o)(z - z_o) + \frac{1}{2}\Phi''(z_o)(z - z_o)^2 + \sum_{l=3}^{\infty}\frac{1}{l!}\Phi^{(l)}(z_o)(z - z_o)^l\right] \\
&= \exp\left[\Phi(z_o) + \Phi'(z_o)(z - z_o) + \frac{1}{2}\Phi''(z_o)(z - z_o)^2\right] f(z, z_o),
\end{aligned}
\tag{3.6}
$$

where, based on the series expansion of the exponential function,

$$
\begin{aligned}
f(z, z_o) &= \exp\left[\sum_{l=3}^{\infty}\frac{1}{l!}\Phi^{(l)}(z_o)(z - z_o)^l\right] \\
&= 1 + \frac{\Phi^{(3)}(z_o)}{6}(z - z_o)^3 + \frac{1}{2}\left(\frac{\Phi^{(3)}(z_o)}{6}(z - z_o)^3\right)^2 + \mathrm{O}\left((z - z_o)^6\right) \\
&= 1 + F(z, z_o).
\end{aligned}
$$

The integral (3.5) becomes

$$
\frac{1}{2\pi j}\oint_c e^{\Phi(z)}dz = \frac{e^{\Phi(z_o)}}{2\pi j}\oint_c e^{\Phi'(z_o)(z - z_o) + \frac{1}{2}\Phi''(z_o)(z - z_o)^2}\left(1 + F(z, z_o)\right)dz.
\tag{3.7}
$$

Depending on the singularities of the integrand and the ROC, it may be permissible to deform the contour c into a vertical line through the real point z_o and a semicircle around the half plane that does not include z_o. This is possible, for example, if z_o were positive, and the only singularities of the integrand were at $z = 0$ and to the right of z_o. In this case the contour would be a line through z_o parallel to the imaginary axis and a semicircle in the left half plane, as shown in Fig. 3.1. If the integrand also vanishes for $\mathrm{Real}\{z\} \to -\infty$, the contribution of the semicircle section of the contour vanishes, and we can carry out the integration along the line $z = z_o + j\omega$ where ω is the imaginary variable. The integral can be rewritten as:
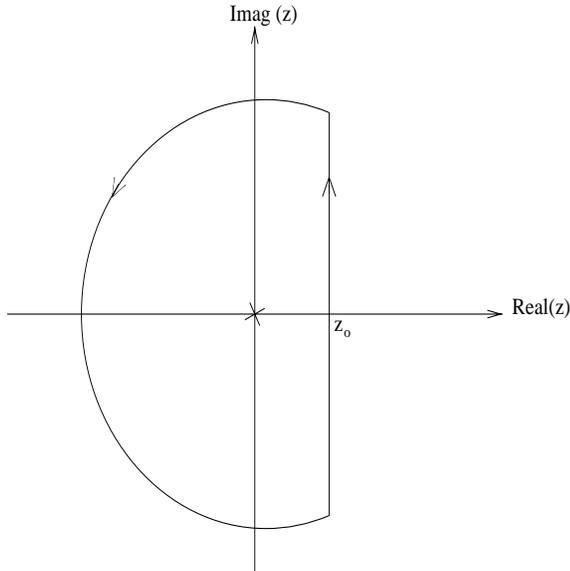
Figure 3.1: Deformation of the contour c in the complex plane into a vertical line through the saddle point and a semicircle in the left half plane at infinity. The singularities of the integrand are at $z = 0$ and $\text{Real}\{z\} = \infty$.

$$
\begin{aligned}
\frac{1}{2\pi j} \oint_c e^{\Phi(z)} dz &= e^{\Phi(z_o)} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\Phi''(z_o)\omega^2} e^{j\omega\Phi'(z_o)} \left[1 + F(z_o + j\omega, z_o)\right] \frac{d\omega}{2\pi} \\
&\approx \frac{e^{\Phi(z_o) - \frac{(\Phi'(z_o))^2}{2\Phi''(z_o)}}}{\sqrt{2\pi\Phi''(z_o)}}.
\end{aligned}
\tag{3.8}
$$

where we have used the inverse Fourier transform in the last step and assumed the integral of the remainder $F(z_o + j\omega, z_o)$ is negligible. If z_o is an actual saddle point, then $\Phi'(z_o) = 0$ and the result (3.8) reduces to the usual saddle-point approximation [39, 80]:

$$
\frac{1}{2\pi j} \oint_c e^{\Phi(z)} dz \approx \frac{e^{\Phi(z_o)}}{\sqrt{2\pi\Phi''(z_o)}}.
\tag{3.9}
$$

Applying (3.8) to the integral (3.4) yields the following approximation for the log-likelihood:

$$
\log P_Y(y) = \Phi(z_o) - \frac{(\Phi'(z_o))^2}{2\Phi''(z_o)} - \frac{1}{2}\log(2\pi\Phi''(z_o)),
\tag{3.10}
$$

where

$$
\Phi(z) = \Phi_y(z) = -(y+1)\log z + \log g_Y(z).
\tag{3.11}
$$

If $\Phi(z)$ is convex, the second derivative $\Phi''(z)$ will be positive and the approximation is guaranteed to be positive. This is a desirable result because our goal is to derive approximations for the probability mass function. To the best of our knowledge, the approximation in (3.10) and (3.8) is a new approach that generalizes the usual saddle-point method. When z_o is the root of $\Phi'(z)$, the generalized saddle-point approximation reduces to the usual one (3.9).

In the next section, we use the saddle point approximation (3.8) to examine the likelihood for the idealized case of monoenergetic X-rays. We first briefly outline the simple case of monoenergetic X-rays with no variability in light generation in the scintillator and no electronic noise. This simplified scenario will introduce our notation and illustrate some of the more salient features of our technique.

3.4 Monoenergetic X-rays

We first examine the case of an X-ray source that generates a beam at a single energy. For simplicity we assume an ideal detector with no electronic noise. We examine two simple cases, with and without light photon variability.

3.4.1 Monoenergetic X-ray Detection with Deterministic Light Generation

Suppose there is no variability in the number of light photons recorded for each absorbed X-ray photon. The p.m.f. of the number of light photons is given by:

$$P_X(x) = \begin{cases} 1, & x = x_o \\ 0, & \text{otherwise,} \end{cases} \tag{3.12}$$

where $x_o > 0$ is the number of light photons. Its value depends on the energy of the X-rays and detector design. The moment generating function for X is:

$$g_X(z) = E[z^X] = \sum_{n=0}^{\infty} z^n P_X(n) = z^{x_o}. \tag{3.13}$$

The moment generating function of the random variable Y is:

$$g_Y(z) = \exp(-\bar{N}(1 - z^{x_o})). \tag{3.14}$$

The ROC of $g_Y(z)$ is the whole plane excluding $|z| \to \infty$. As a check, when $x_o = 1$ (i.e., we have an X-ray photon counting detector), $g_Y(z)$ reduces to the ordinary Poisson generating function $\exp(-\bar{N}(1 - z))$. Alternatively, the generating function of the Poisson random variable $Y' = \frac{Y}{x_o}$ is $g_Y(z^{\frac{1}{x_o}})$ which is also the simple Poisson generating function.

To find the p.m.f. of Y, we need to take the inverse Z-transform of $g_Y(z)$. In this case $g_Y(z)$ could be inverted by inspection (see Appendix B), but for illustration purposes we use the saddle point approximation method (3.10) and (3.11). Plugging $g_Y(z)$ in (3.11) gives:

$$\Phi_y(z) = -(y + 1) \log z - \bar{N} + \bar{N} z^{x_o}. \tag{3.15}$$

For $y \geq 0$, $z \in \mathcal{R}$ and $z > 0$, the function $\Phi_y(z)$ and the integrand in (3.5) are convex. The integrand has a minimum (saddle point) at $z = z_o$, where z_o is real and $z_o \geq 0$. The saddle point z_o of $\Phi_y(z)$ is the real root of

$$\Phi_y'(z) = -\frac{(y + 1)}{z} + \bar{N} x_o z^{x_o - 1} = 0. \tag{3.16}$$

The saddle point is easily seen to be:

$$z_o(y) = \left(\frac{y + 1}{\bar{N} x_o}\right)^{\frac{1}{x_o}} = \left(\frac{y + 1}{\bar{y}}\right)^{\frac{1}{x_o}}, \tag{3.17}$$

since $\bar{y} = \bar{N} x_o$. Evaluating Φ_y and Φ_y'' at the saddle point yields

$$\begin{aligned}
\Phi_y(z_o) &= \frac{1}{x_o}\left[(y + 1)\log\frac{\bar{y}}{(y + 1)} - \bar{y} + (y + 1)\right] \\
\Phi_y''(z_o) &= (\bar{y})^{\frac{2}{x_o}} x_o (y + 1)^{1 - \frac{2}{x_o}}.
\end{aligned}$$

To ensure that the saddle-point approximation is applicable, we must look into the deformation of the contour of integration c in (3.4) and the behavior of the integrand as $|z| \to \infty$. Towards that end we must first verify the ROC of the integrand in (3.4). With the change of variable $z' = z^{x_o}$, and substituting (3.14) in (3.4) we get:

$$P_Y(y) = \frac{e^{-\bar{N}}}{x_o} \frac{1}{2\pi j} \oint_c z^{-(y'+1)} e^{\bar{N}z} dz, \tag{3.18}$$

where for simplicity we dropped the $'$ from z' and $y' = y/x_o$. It is obvious that the integrand has a pole at $z = 0$. In order to determine the ROC, we examine the real and imaginary parts of $e^{\bar{N}z}$ by writing $z = re^{j\theta}$.

$$
\begin{aligned}
e^{\bar{N}z} &= e^{\bar{N}re^{j\theta}} = e^{\bar{N}r(\cos(\theta)+j\sin(\theta))} \\[2mm]
&= e^{\bar{N}r\cos(\theta)} e^{j\bar{N}r\sin(\theta)} \\[2mm]
&= e^{\bar{N}r\cos(\theta)} \left(\cos(\bar{N}r\sin(\theta)) + j\sin(\bar{N}r\sin(\theta)) \right).
\end{aligned}
$$

$$\mathrm{Real}\{e^{\bar{N}z}\} = e^{\bar{N}r\cos(\theta)} \cos(\bar{N}r\sin(\theta)),$$

and

$$\mathrm{Imag}\{e^{\bar{N}z}\} = e^{\bar{N}r\cos(\theta)} \sin(\bar{N}r\sin(\theta)).$$

It is readily seen that the behavior of the integrand as $r = |z| \to \infty$ depends on the value of the phase angle θ. If $|\theta| \geq \pi/2$, then the real part of the exponential term will approach zero. On the other hand if $|\theta| < \pi/2$, the real part of the exponential term will diverge to $\pm\infty$. A similar result holds for the imaginary part. For large $r = |z|$, the integrand converges in the left half plane (indeed, it vanishes), and it diverges in the right half plane. It is therefore permissible to use a contour that is a line through the positive saddle point in the right half plane, parallel to the imaginary axis, and a semicircle at infinity in the left half plane.

Taking the log of the saddle point approximation (3.9) to the p.m.f. gives the log-likelihood:

$$\log P_Y(y) \equiv \frac{1}{x_o} \left[y \log \bar{y} - \bar{y} \right],$$ (3.19)

where $\equiv$ indicates that constant terms independent of $\bar{y}$ are dropped. This saddle-point approximation of the log-likelihood is equivalent to the conventional Poisson log-likelihood [55], which gives some reassurance of the utility of this method.

In image reconstruction, we want to maximize the log-likelihood over the set of solutions in object space. The dependence on the object comes from the fact that the mean number of X-ray photons, $\bar{N}$, depends on the object attenuation. In the monoenergetic problem,

$$\bar{N} = N_o \exp \left(- \int_{\mathcal{L}} \mu(x, w_o) dl \right)$$

where the integral in the exponent is a line integral over the ray $\mathcal{L}$. If we let $t = \int_{\mathcal{L}} \mu(x, w_o) dl$ denote the line integral, then the likelihood (3.19) has the usual form for transmission tomography:

$$L_y(t) \equiv \frac{y}{x_o} \log e^{-t} - N_o e^{-t}.$$ (3.20)

The result of this simple case will be useful throughout this paper. We next discuss an important generalization, where light generation is variable, as it is in practice.

3.4.2 Monoenergetic X-ray Detection with Poisson Light Statistics

In the above derivation, we have ignored the statistics of the light photons generated by the scintillation process. Light generation in scintillating phosphor is a complex process. X-ray photons that are absorbed by or scatter within the scintillator generate charge carriers that can generate light, or scatter and generate secondary

carriers that in turn can also produce light. Scintillation photons undergo a complex diffusion process, characterized by hundreds of scattering incidents per light photon [5]. K-edge effects can also have a strong impact on the amount of light generated [36]. Another complicating factor is the detector photodiodes dependence on wavelength. When X-ray photons scintillate, they produce light with a distribution of optical wavelengths, leading to wavelength-dependent statistics in the recorded signal.

As a first approximation, we assume that a detected X-ray photon will generate a Poisson number of light photons with mean number proportional to the X-ray energy. Using this assumption, we derive an approximate likelihood based on monoenergetic X-rays. For this assumption, the Poisson p.m.f. P_X is:

$$P_X(n) = \frac{x_o^n e^{-x_o}}{n!}, \tag{3.21}$$

where n represents the number of light photons generated by one of the N detected X-ray photons, and $x_o = Gw_o$ is the mean number of light photons. The scaling constant G is a characteristic of the scintillating phosphor and detector design. The moment generating function of this Poisson distribution is

$$g_X(z) = \exp\left[x_o(z-1)\right]. \tag{3.22}$$

Substituting into (3.3) yields the moment generating function of the measured signal Y

$$g_Y(z) = \exp\left[-\bar{N}(1 - e^{-x_o(1-z)})\right]. \tag{3.23}$$

The associated p.m.f. has the form (3.8) where

$$\Phi_y(z) = -(y+1)\log z - \bar{N} + \bar{N}e^{-x_o(1-z)} \tag{3.24}$$

$$\Phi'_y(z) = -\frac{(y+1)}{z} - \bar{N}x_o e^{-x_o(1-z)} \tag{3.25}$$

$$\Phi''_y(z) = \frac{(y+1)}{z^2} + \bar{N}x_o^2 e^{-x_o(1-z)}. \tag{3.26}$$

Before we can apply the saddle-point approximation, we must examine the ROC of the integrand in (3.4) and determine if the necessary contour deformation can be carried out. Plugging (3.23) in (3.4) gives:

$$P_Y(y) = \frac{e^{-\bar{N}}}{2\pi j} \oint_c z^{-(y+1)} e^{\bar{N}e^{-x_o}e^{x_o z}} dz.$$

As in the previous section, the term $z^{-(y+1)}$ introduces a pole at $z = 0$. To determine the behavior of the integrand in the remainder of the complex plane, we again examine the real and imaginary parts of the exponential term in the integrand.

$$
\begin{aligned}
e^{\bar{N}e^{-x_o}e^{x_o z}} &= e^{\bar{N}'e^{x_o r e^{j\theta}}} \\
&= \exp\left\{\bar{N}'e^{x_o r(\cos(\theta)+j\sin(\theta))}\right\} \\
&= \exp\left\{\bar{N}'e^{x_o r\cos(\theta)}e^{jx_o r\sin(\theta)}\right\} \\
&= \exp\left\{\bar{N}'e^{x_o r\cos(\theta)}\left(\cos\left(x_o r\sin(\theta)\right) + j\sin\left(x_o r\sin(\theta)\right)\right)\right\} \\
&= \exp\left\{\bar{N}'e^{x_o r\cos(\theta)}\cos(x_o r\sin(\theta))\right\}\exp\left\{j\bar{N}'e^{x_o r\cos(\theta)}\sin(x_o r\sin(\theta))\right\} \\
&= \exp\left\{\bar{N}'e^{x_o r\cos(\theta)}\cos(x_o r\sin(\theta))\right\}\left(\cos\left(\bar{N}'e^{x_o r\cos(\theta)}\sin(x_o r\sin(\theta))\right)\right. \\
&\quad \left. +j\sin\left(\bar{N}'e^{x_o r\cos(\theta)}\sin(x_o r\sin(\theta))\right)\right),
\end{aligned}
$$

where $\bar{N}' = \bar{N}e^{-x_o}$. When $|z| \to \infty$, the behavior of the integrand will depend on the phase angle θ, with the real and imaginary parts converging to ± 1 for $|\theta| \geq \pi/2$ and diverging to $\pm\infty$ otherwise. So the ROC consists of the complex plane except the origin and the right half plane at infinity. A contour consisting of a line through the

saddle point and a semicircle in the left half plane is therefore permissible. Moreover, the integrand vanishes at infinity in the left half plane.

The saddle point in this case is not available analytically from $\Phi'_y(z) = 0$. It is possible to compute the saddle point numerically, but that approach would be computationally prohibitive in the context of iterative image reconstruction, where the saddle point needs to be updated at every iteration.

Rather than use the exact saddle point, we use the saddle point (3.17) derived for the monoenergetic source (with deterministic light photon generation), based on the intuition that the saddle point (3.17) should be a close approximation of the root of (3.25). Since (3.17) is inexact in this case, we apply the generalized saddle point integration approximation discussed in Section 3.3. Evaluating Φ_y and its first two derivatives at z_o in (3.17) gives:

$$\Phi_y(z_o) = -\frac{y+1}{x_o}\log\frac{(y+1)}{\bar{N}x_o} - \bar{N} + \bar{N}e^{-x_o(1-z_o(y))} \tag{3.27}$$

$$\Phi'_y(z_o) = -(y+1)\left(\frac{\bar{N}x_o}{y+1}\right)^{\frac{1}{x_o}} + \bar{N}x_o e^{-x_o(1-z_o(y))} \tag{3.28}$$

$$\Phi''_y(z_o) = (y+1)\left(\frac{y+1}{\bar{N}x_o}\right)^{-\frac{2}{x_o}} + \bar{N}x_o^2 e^{-x_o(1-z_o(y))}. \tag{3.29}$$

It is now possible to write the approximate likelihood by plugging (3.27)-(3.29) in (3.8). The expression is long so it is not given here, but it could be used for maximum likelihood reconstruction.

To explore further, note that x_o is the mean number of light photons, which is usually in the range of hundreds to thousands [5, 36]. Assuming that x_o is large enables us to make the following approximations:

$$e^{-x_o(1-z_o)} = \exp\left[-x_o\left(1-\left(\frac{y+1}{\bar{N}x_o}\right)^{\frac{1}{x_o}}\right)\right] \approx \left(\frac{y+1}{\bar{N}x_o}\right)$$

$$\left(\frac{y+1}{\bar{N}x_o}\right)^{\frac{1}{x_o}} \approx 1,$$

which simplify Φ_y and its derivatives to

$$\begin{aligned}
\Phi_y(z_o) &\approx \frac{y}{x_o} \log \frac{\bar{N} x_o}{y} - \bar{N} \equiv \frac{1}{x_o}(y \log \bar{y} - \bar{y}) \\
\Phi_y'(z_o) &\approx 0 \\
\Phi_y''(z_o) &\approx (y+1)(1+x_o).
\end{aligned}$$

(3.30)

Under these approximations, the first derivative is zero and the second derivative is a constant. The log-likelihood is equal to $\Phi_y(z_o)$. The large optical gain approximation essentially leads to the same result as the earlier section where the likelihood was derived without optical spread. Also note that this result is equivalent to the Poisson likelihood of the variable Y/x_o. This again serves as a check, since large optical generation does in reality lead to Poisson-like statistics. Fig. 3.2 illustrates how the compound Poisson likelihood approaches the ordinary Poisson likelihood as the gain x_o increases. Another reassurance comes from (3.23). If x_o is large, $g_{\frac{Y}{x_o}}(z) = g_Y(z^{\frac{1}{x_o}}) \approx \exp[-\bar{N}(1-z)]$ which is the Poisson moment generating function. These results also justify using the Poisson likelihood for monoenergetic CT. Polyenergetic CT, however, is more complex, and is the subject of the next section.

3.5   Polyenergetic X-rays

Clinical X-ray tubes are polyenergetic, producing a continuum of photon energies. It is possible to generalize the likelihood and the saddle point approximation to the continuous X-ray spectrum case. However, when it comes to practical implementation, the continuous spectrum is approximated by a discrete sum. We will use such a discrete approximation to the continuous spectrum as our starting point for deriving an approximate likelihood.

Assume the X-ray source produces photons at L distinct energies $\{w_l\}_{l=1}^L$. The
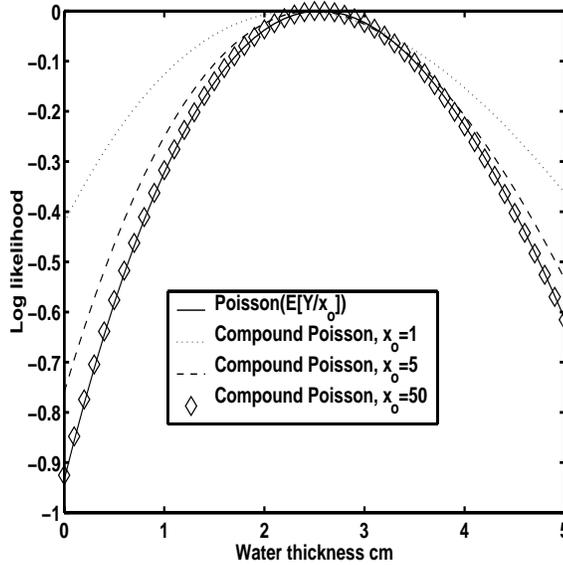
Figure 3.2: The exact compound Poisson likelihood approaches the ordinary Poisson likelihood as the number of light photons generated by X-ray quanta increases. These plots are generated using 10 X-ray photons incident on water with $\mu = 0.2$ cm^{-1}.

p.m.f. of the X-ray beam incident on the object is:

$$P[W = w_l] = p_l$$

where

$$\sum_{l=1}^{L} p_l = 1.$$

The X-ray beam traverses the object and experiences energy-dependent attenuation. The amount of attenuation is exponentially related to the path that the X-ray beam takes through the object, as expressed mathematically by the line integral. For simplicity here, we adopt the following attenuation model[3]:

$$\mu(x; w) = m(w)\rho(x), \tag{3.31}$$

---

[3]This model assumes only one material and is convenient because we can separate spectral and spatial dependence. In a more general (and realistic) model the mass attenuation coefficient would be spatially varying. See Chapter IV and Chapter V for more complex object models.

where $m(w)$ is the mass attenuation coefficient and $\rho(x)$ is the material density. We use this model because it separates the energy and spatial components of the attenuation coefficient. This model gives the following expression for the line integral:

$$\int_{\mathcal{L}} \mu(x; w) dl = m(w)s \qquad (3.32)$$

$$s \triangleq \int_{\mathcal{L}} \rho(x) dl. \qquad (3.33)$$

The energy p.m.f. of the attenuated X-ray beam is:

$$P[W = w_l; s] = \tilde{p}_l(s) \qquad (3.34)$$

where

$$\tilde{p}_l(s) \triangleq \frac{p_l e^{-m(w_l)s}}{\sum_{k=1}^{L} p_k e^{-m(w_k)s}}. \qquad (3.35)$$

The denominator is a normalizing factor that ensures that the p.m.f. sums to unity. The number of photons interacting with the detector is the Poisson random variable N with mean

$$\bar{N}(s) = N_o \sum_{l=1}^{L} p_l e^{-m(w_l)s} \qquad (3.36)$$

where N_o is the total number of photons emitted by the X-ray source.

The next step in modeling CT measurements is to add to the polyenergetic physics and compound Poisson statistics the non-trivial statistical distribution for light photons generated in the scintillator. We again assume that individual X-ray quanta generate Poisson light with mean proportional to the generating X-ray energy. In other words, the conditional light p.m.f. is

$$P[X = n | W = w] = \frac{(Gw)^n e^{-Gw}}{n!}. \qquad (3.37)$$

By total probability, the unconditional p.m.f. $P_X(x; s)$ is given by

$$P[X = n; s] = \sum_{l=1}^{L} P[X = n | W = w] P[W = w_l; s] = \sum_{l=1}^{L} \tilde{p}_l(s) \frac{(x_l)^n e^{-x_l}}{n!}, \qquad (3.38)$$

where $x_l = E[X|w = w_l] = Gw_l$ is the mean number of light photons generated by an absorbed X-ray photon with energy w_l. This p.m.f. represents a Poisson distribution at each energy/optical gain value x_l weighted by $\tilde{p}_l$ such that the overall sum adds to unity. The moment generating function exhibits the same weighting:

$$g_X(z; s) = \sum_{n=0}^{\infty} \sum_{l=1}^{L} z^n \tilde{p}_l(s) \frac{(x_l)^n e^{-x_l}}{n!} = \sum_{l=1}^{L} \tilde{p}_l(s) e^{x_l(z-1)}. \tag{3.39}$$

Substituting (3.39) into (3.3), the compound Poisson moment generating function of the total number of recorded light photons Y is

$$g_Y(z; s) = \exp\left[-\bar{N}(s)\left(1 - \sum_{l=1}^{L} \tilde{p}_l(s) e^{x_l(z-1)}\right)\right], \tag{3.40}$$

with ROC including the whole Z-plane except for $z = \infty$. As before, the p.m.f. $P_Y(y; s)$ of Y is given by an inverse Z-transform of the form (3.8), where

$$\Phi_y(z) = -(y+1)\log z - \bar{N}(s) + \bar{N}(s) \sum_{l=1}^{L} \tilde{p}_l(s) e^{-x_l(1-z)} \tag{3.41}$$

$$\Phi_y'(z) = -\frac{(y+1)}{z} + \bar{N}(s) \sum_{l=1}^{L} \tilde{p}_l(s) x_l e^{-x_l(1-z)} \tag{3.42}$$

$$\Phi_y''(z) = \frac{(y+1)}{z^2} + \bar{N}(s) \sum_{l=1}^{L} \tilde{p}_l(s) x_l^2 e^{-x_l(1-z)}. \tag{3.43}$$

These expressions are natural extensions of the monoenergetic case (3.24)-(3.26). The approximation in (3.8) is applicable because the integrand is convex for real nonnegative z, and arguments regarding the ROC and contour deformation similar to the monoenergetic case can be made. Also, the second derivative $\Phi_y''(z)$ is positive which ensures the positivity of the saddle-point approximation.

Since an analytic saddle point is not available, we pursue the analogy with the monoenergetic case and evaluate Φ_y and its derivatives at a saddle point estimate similar to (3.17),

$$\hat{z}_o(y; s) = \left(\frac{y+1}{\bar{N}(s)\bar{x}(s)}\right)^{\frac{1}{\bar{x}(s)}} = \left(\frac{y+1}{\bar{y}(s)}\right)^{\frac{1}{\bar{x}(s)}}, \tag{3.44}$$

where $\bar{x}(s) \triangleq \sum_{l=1}^{L} \tilde{p}_l(s) x_l$ is the effective mean number of light photons, and $\bar{y}(s) = N_o \sum_{l=1}^{L} p_l x_l e^{-m(w_l)s}$. Since $\bar{x}$ is likely to be in the range of tens to hundreds, we can use the large optical gain approximation, which, for large $\bar{x}$, gives,

$$e^{-x_l(1-\hat{z}_o)} = \exp\left[-x_l\left(1 - \left(\frac{y+1}{\bar{y}(s)}\right)^{\frac{1}{\bar{x}}}\right)\right] \approx \left(\frac{y+1}{\bar{y}(s)}\right)^{\frac{x_l}{\bar{x}}}.$$

This approximation simplifies Φ_y and its derivatives, resulting in the following approximate polyenergetic CT (compound Poisson) log-likelihood:

$$
\begin{aligned}
L(y;s) \approx\ & \frac{1}{\bar{x}}\left[y \log \bar{y}(s) - \bar{y}(s) + \bar{y}(s) \sum_{l=1}^{L} \tilde{p}_l(s) \left(\frac{y+1}{\bar{y}(s)}\right)^{\frac{x_l}{\bar{x}}}\right] \\
& -\frac{1}{2}\frac{\left(\bar{N}(s)\sum_{l=1}^{L}\tilde{p}_l(s)x_l\left(\frac{y+1}{\bar{y}(s)}\right)^{\frac{x_l}{\bar{x}}} - (y+1)\left(\frac{y+1}{\bar{y}(s)}\right)^{-\frac{1}{\bar{x}}}\right)^2}{\left((y+1)\left(\frac{y+1}{\bar{y}(s)}\right)^{-\frac{2}{\bar{x}}} + \bar{N}(s)\sum_{l=1}^{L}\tilde{p}_l(s)x_l^2\left(\frac{y+1}{\bar{y}(s)}\right)^{\frac{x_l}{\bar{x}}}\right)} \\
& +\frac{1}{2}\log\left[\left((y+1)\left(\frac{y+1}{\bar{y}(s)}\right)^{-\frac{2}{\bar{x}}} + \bar{N}(s)\sum_{l=1}^{L}\tilde{p}_l(s)x_l^2\left(\frac{y+1}{\bar{y}(s)}\right)^{\frac{x_l}{\bar{x}}}\right)\right]
\end{aligned}
\tag{3.45}
$$

To summarize, we have derived a likelihood approximation based on the compound Poisson statistics of X-ray CT detectors, assuming that light photons generated in the scintillator follow the Poisson distribution with means proportional to the energies of the X-ray quanta. In the next section we will generalize our model to incorporate the effect of additive Gaussian readout noise.

3.5.1 Gaussian Readout Noise

The compound Poisson model would be sufficient to describe CT measurements if the detectors were ideal. In reality, measurements suffer from an additive component caused by noise from the system electronics. This noise is assumed to be independent of the quanta measurements and to follow a Gaussian distribution with zero mean

and variance σ^2. The model in (3.2) generalizes to

$$U = Y + R = \sum_{n=1}^{N} X_n + R \tag{3.46}$$

where, as before, X_n denotes the number of light photons generated by the nth X-ray photon, N is Poisson with mean $\bar{N}$, and $R \sim \mathcal{N}(0, \sigma^2)$. The random variables Y and R are assumed to be independent.

Since U is the sum of a discrete and continuous random variables, it is itself continuous. The moment generating function for U is therefore the Laplace transform of its probability density function $f_U(u)$. By total probability, the p.d.f. of U is

$$f_U(u) = \sum_{k=0}^{\infty} P_Y(k) f_R(u - k). \tag{3.47}$$

The moment generating function of U is

$$
\begin{aligned}
g_U(\zeta) &= E[e^{-u\zeta}] = \int e^{-u\zeta} \sum_{k=0}^{\infty} P_Y(k) f_R(u - k) du \\
&= \sum_{k=0}^{\infty} P_Y(k) \int e^{-u\zeta} f_R(u - k) du \\
&= \sum_{k=0}^{\infty} P_Y(k) e^{-k\zeta} \int e^{-u\zeta} f_R(u) du \\
&= g_Y(z)|_{z=e^{-\varsigma}} g_R(\zeta),
\end{aligned} \tag{3.48}
$$

where $g_R(\zeta) = e^{\frac{\zeta^2 \sigma^2}{2}}$ is the well-known moment generating function of the Gaussian density. Using the polyenergetic moment generating function (3.23), the moment generating function of the overall recorded signal U is

$$g_U(\zeta) = \exp\left[-\bar{N}(s) \left(1 - \sum_{l=1}^{L} \tilde{p}_l(s) e^{-x_l(1 - e^{-\varsigma})} \right) + \frac{\zeta^2 \sigma^2}{2} \right]. \tag{3.49}$$

The probability density function is given by the inverse Laplace transform of $g_U(\zeta)$ and, as in earlier cases, can be approximated using the saddle point technique.

$$
\begin{aligned}
f_U(u) &= \frac{1}{2\pi j} \int e^{\zeta y} g_U(\zeta) d\zeta \\
&= \frac{1}{2\pi j} \int e^{\Phi_u(\zeta)} d\zeta,
\end{aligned} \tag{3.50}
$$

where

$$\Phi_u(\zeta) = \zeta u - \bar{N}(s) + \bar{N}(s) \sum_{l=1}^{L} \tilde{p}_l(s) e^{-x_l(1-e^{-\zeta})} + \frac{\sigma^2 \zeta^2}{2} \qquad (3.51)$$

$$\Phi'_u(\zeta) = u - \bar{N}(s) \sum_{l=1}^{L} \tilde{p}_l(s) x_l e^{-x_l - \zeta + x_l e^{-\zeta}} + \sigma^2 \zeta \qquad (3.52)$$

$$\Phi''_u(\zeta) = \bar{N}(s) \sum_{l=1}^{L} \tilde{p}_l(s) x_l (1 + x_l e^{-\zeta}) e^{-\zeta} e^{-x_l(1-e^{-\zeta})} + \sigma^2. \qquad (3.53)$$

The second derivative $\Phi''_u(\zeta)$ is positive in this case as well, leading to a positive saddle-point approximation. Since there is no analytical solution to $\Phi'_u(\zeta) = 0$, we again use an approximation to the saddle point similar in form to the simple monoenergetic saddle point (3.17). In terms of ζ, the approximate saddle point is

$$\hat{\zeta}_o(y; s) = -\frac{1}{\bar{x}(s)} \log \frac{u}{\bar{u}(s)}, \qquad (3.54)$$

where $\bar{u}(s) = \bar{N}(s)\bar{x}(s)$. Evaluating Φ_u and its derivatives at $\hat{\zeta}_o$ and using the large optical gain approximation gives:

$$\Phi_u(\hat{\zeta}_o) \approx -\frac{u}{\bar{x}(s)} \log \frac{u}{\bar{u}(s)} - \bar{N}(s) + \bar{N}(s) \sum_{l=1}^{L} \tilde{p}_l(s) \left(\frac{u}{\bar{u}(s)}\right)^{\frac{x_l}{\bar{x}(s)}} + \frac{1}{2} \sigma^2 \left(\log \left(\frac{u}{\bar{u}(s)}\right)^{\frac{1}{\bar{x}(s)}}\right)^2$$

$$\Phi'_u(\hat{\zeta}_o) \approx u - \bar{N}(s) \sum_{l=1}^{L} \tilde{p}_l(s) x_l \left(\frac{u}{\bar{u}(s)}\right)^{\frac{x_l}{\bar{x}(s)}} - \sigma^2 \log \left(\frac{u}{\bar{u}(s)}\right)^{\frac{1}{\bar{x}(s)}}$$

$$\Phi''_u(\hat{\zeta}_o) \approx \bar{N}(s) \sum_{l=1}^{L} \tilde{p}_l(s) x_l (1 + x_l) \left(\frac{u}{\bar{u}(s)}\right)^{\frac{x_l}{\bar{x}(s)}} + \sigma^2.$$

The log-likelihood follows directly from (3.10) and (3.11).

Accounting for the effect of Gaussian readout noise is important for imaging systems where this component contributes significantly to the measurements. This may not the case in clinical CT scanners, where readout noise is kept very low, but may be useful for systems with CCD camera detectors, such as some of the commercially available microCT scanners.

3.5.2 First and Second Moments

It is also useful at this point to present formulas for the first and second moments of U, which correspond to the first and second derivative of its moment generating function evaluated at zero, respectively. Taking the derivatives of g_U in (3.49) is straight forward, and we simply state the results:

$$E_s[U] = \bar{N}(s) \sum_{l=1}^{L} x_l \tilde{p}_l \tag{3.55}$$

$$E_s[U^2] = \left(\bar{N}(s) \sum_{l=1}^{L} x_l \tilde{p}_l \right)^2 + \bar{N}(s) \sum_{l=1}^{L} x_l^2 \tilde{p}_l + \sigma^2 \tag{3.56}$$

$$\mathrm{Var}_s[U] = \bar{N}(s) \sum_{l=1}^{L} x_l^2 \tilde{p}_l + \sigma^2. \tag{3.57}$$

The subscript s of the moments operators makes explicit the dependence of the moments on the line integral s. The first and second moments are useful in computing a likelihood function based on a Gaussian distribution:

$$L(u; s) = -\frac{1}{2\mathrm{Var}_s[U]} (u - E_s[U])^2 - \frac{1}{2} \log \mathrm{Var}_s[U]. \tag{3.58}$$

3.6 Moments Interpretation of the Saddle Point

In this section we interpret the saddle point in terms of the moments of the random variable. We show that the saddle point is, in some sense, a measure of the difference between the random variable and its ensemble mean. To illustrate this point, we write the function Φ_u and its first derivatives in terms of the moment

generating function:

$$
\begin{aligned}
\Phi_u(\zeta) &= \zeta u + \log g_U(\zeta) \\
\Phi'_u(\zeta) &= u + \frac{g'_U(\zeta)}{g_U(\zeta)} \\
\Phi''_u(\zeta) &= \frac{g''_U(\zeta)g_U(\zeta) - (g'_U(\zeta))^2}{g_U^2(\zeta)}.
\end{aligned}
\tag{3.59}
$$

Recall that $g_U(0) = 1$, $g'_U(0) = -E[U]$ and $g''_U(0) = E[U^2]$. Expanding the first derivative Φ'_u in a first order Taylor expansion around $\zeta = 0$ and evaluating the series at its root ζ_o gives

$$
\Phi'_u(\zeta_o) \approx \Phi'_u(0) + \Phi''_u(0)\zeta_o \approx 0.
\tag{3.60}
$$

Solving for ζ_o in (3.60) gives

$$
\zeta_o(u) \approx -\frac{\Phi'_u(0)}{\Phi''_u(0)} = -\frac{u + \frac{g'_U(0)}{g_U(0)}}{\frac{g''_U(0)g_U(0) - (g'_U(0))^2}{g_U^2(0)}} = \frac{E[U] - u}{\text{Var}\{U\}}.
\tag{3.61}
$$

The case of the Gaussian random variable illustrates this result nicely. Let U be Gaussian with mean m and variance σ^2. The moment generating function of U is

$$
g_U(\zeta) = e^{-m\zeta + \frac{\zeta^2\sigma^2}{2}},
$$

and

$$
\begin{aligned}
\Phi_u(u) &= -m\zeta + u\zeta + \frac{\zeta^2\sigma^2}{2} \\
\Phi'_u(u) &= -m + u + \zeta\sigma^2 = 0.
\end{aligned}
$$

The saddle point follows immediately:

$$
\zeta_o^{\text{Gaussian}} = \frac{m - u}{\sigma^2} = \frac{E[U] - u}{\text{Var}\{U\}}.
\tag{3.62}
$$

If the saddle point of (3.54) is expressed with a 1st order Taylor series about $u = E[u]$, the resulting approximation will essentially be (3.62). In addition, the saddle point of (3.54) involves the log of the ratio of the recorded signal u and its mean, which is a measure of their dissimilarity. We have empirically observed that the saddle point approximation proposed in this section is close to (3.54) (when $\sigma^2 = 0$) for a both low and high SNR.

The value of the saddle point approximation (3.61) is mostly interpretive, since (3.54) is simple to compute. The fact that the moments interpretation and form of the saddle point agrees with our earlier saddle point approximations increases our confidence in the use of those approximations.

3.7   Preliminary Results

We compare the approximate compound Poisson likelihood to a numerically computed exact likelihood and to the ordinary Poisson likelihood. We use 50 kVp X-rays incident on 2.5 cm of water and plot the likelihoods over a range of water thicknesses. The effective energy of the spectrum is 31.5 keV and we set the detector gain $G = 1$. Fig. 3.3 illustrates likelihood plots for high and low signal situations, both with an additive noise component. The bottom of Fig. 3.3 represents a situation with severe X-ray photon deprivation. We can see that the saddle-point approximate likelihood is closer to the exact likelihood than the ordinary Poisson likelihood function. The top of Fig. 3.3 is not a case of very high signal level, yet the ordinary Poisson likelihood is very close to the exact one. X-ray flux in CT can reach up to millions of quanta. This may shed some light on why the Poisson likelihood works well with conventional CT applications. There are however, even clinically, situations when the measurements are extremely photon-deprived, and the proposed approximate
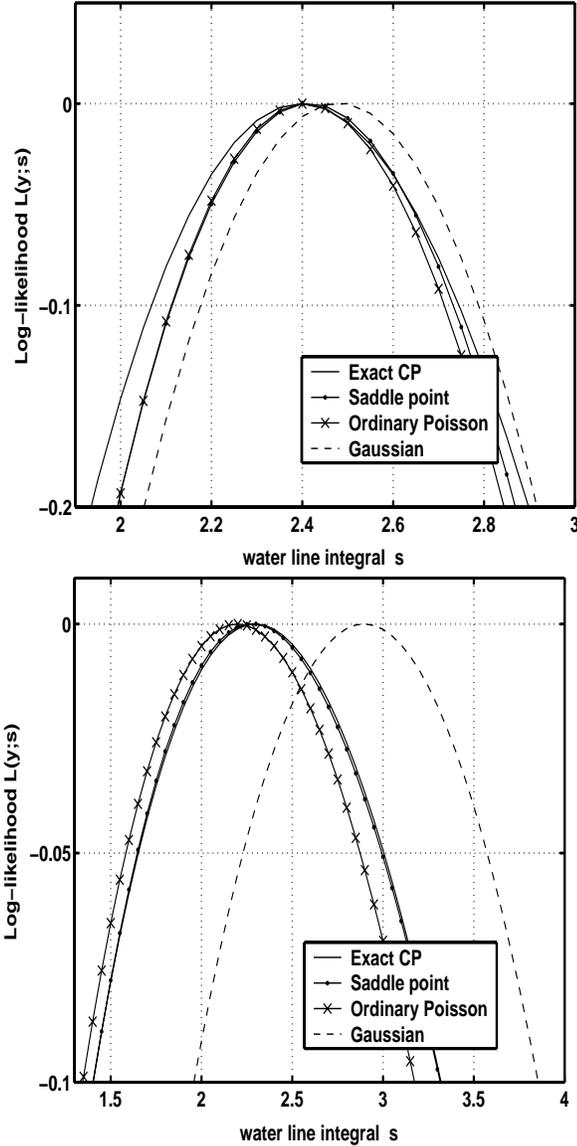
Figure 3.3: Polyenergetic exact and approximate compound Poisson likelihoods and the ordinary Poisson likelihood. 50 kVp X-rays incident on 2.5 cm of water. Effective incident energy is 31.5 keV. Detector gain set to unity. Average number of light photons ≈ 33 photons/keV. Top: 50 incident photons generate a detector signal of 671 optical photons. Noise level approximately 5%. Bottom: 5 incident photons generate a detector signal of 72 and an additive noise level of about 20%.

likelihood may be useful for image reconstruction.

To further assess the different approaches in a situation more similar to those encountered in image reconstruction, we computed the polyenergetic likelihoods that

are actually the sums of several other likelihoods. Because the measurements are statistically independent, the log-likelihoods add. We used a 50 kVp spectrum incident on six different thicknesses of water (3.5 cm - 5.2 cm). The incident flux was 25 photons and additive noise of 4% - 7% of the recorded signal was included. Detector gain was set to unity. Fig. 3.4 shows plots of the individual likelihoods and their cumulative sum (exact), and the approximate compound Poisson and ordinary Poisson likelihoods.

3.8   Conclusion

In this chapter we developed a statistical model for X-ray CT measurements that is more sophisticated and more realistic than the Poisson model. The statistical model accounts for energy-dependent statistics in the measurements, and can be described as a compound Poisson distribution. We also compared the exact compound Poisson likelihood with a saddle-point approximation and the ordinary Poisson likelihood.

Preliminary results show that the compound Poisson saddle-point approximate likelihood is more accurate than the ordinary Poisson likelihood, especially in situations of low counts. The ordinary Poisson model works well with moderate to high signal levels. More analysis is needed to understand the usefulness of the compound Poisson likelihood in statistical image reconstruction. The compound Poisson model may be useful for certain source/detector configurations where photon deprivation is very acute, as in some microCT scanners with CCD detectors. For the applications and X-ray flux levels considered in the remainder of this thesis, ordinary Poisson statistics will suffice.

There are additional possible refinements of the statistical model. The model can
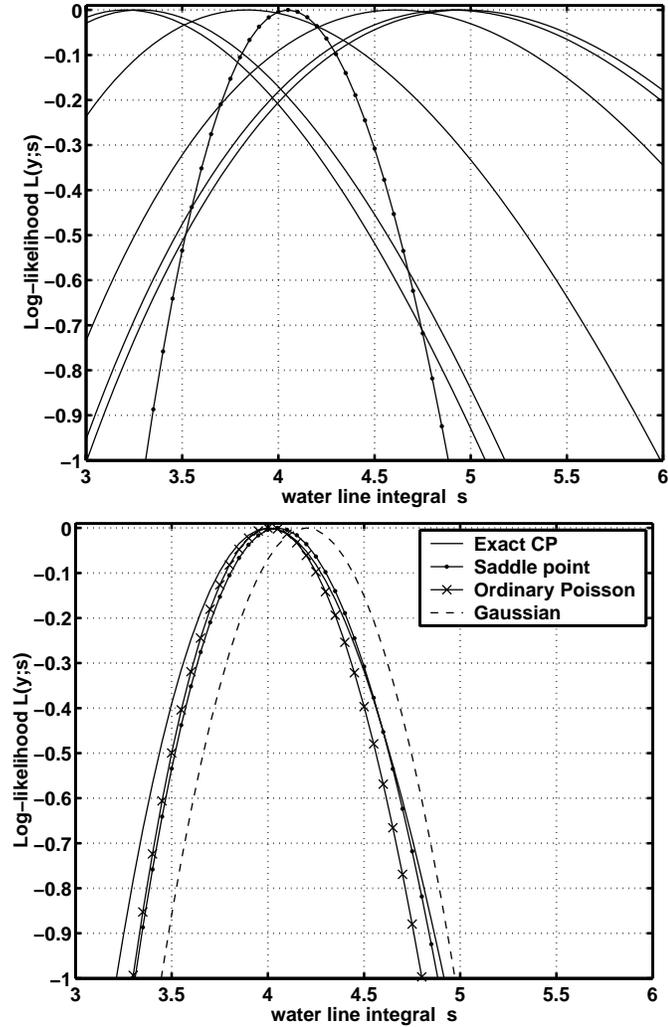
Figure 3.4: Polyenergetic exact and approximate compound Poisson likelihoods and the ordinary Poisson likelihood for 50 kVp X-rays incident on different thickness of water. Effective incident energy is 31.5 keV. Detector gain set to unity. Average number of light photons ≈ 33 photons/keV. Top: plots of individual likelihoods (solid line) and their cumulative sum (dotted line). Bottom: Plots of the exact, compound Poisson and ordinary Poisson cumulative log-likelihoods.

include, for example, wavelength dependence of the detector photodiodes, leakage current dependence on the accumulated charge and nonlinearities in the electronics. Statistical iterative reconstruction with the compound Poisson likelihood may lead to the practical advantage of lowering X-ray dose with little sacrifice in image quality.

CHAPTER IV

Statistical Image Reconstruction Algorithm for Polyenergetic CT

4.1 Introduction

Unlike the conventional filtered back projection (FBP) method, iterative methods are based on models for the measurement statistics and physics, and for this reason have the potential for improved bias-variance and image artifact performance. It is important that the assumed statistical and physics models are true to the data acquisition process, since mismatches can cause significant image degradation.

This chapter focuses on beam hardening effects [8, 20, 49, 52, 81]. Beam hardening is a process whereby the average energy of the X-ray beam increases as the beam propagates through a material because lower-energy X-rays are preferentially attenuated. X-rays traversing different paths through an object will emerge with different spectra, leading to inconsistencies in the data (in the Radon sense) and reconstruction artifacts if one ignores the energy dependence of the measurements.

If uncorrected, beam hardening generally leads to a reduction in the reconstructed attenuation coefficient [8]. Thick bones also generate dark streaks [52]. In soft tissue, non-uniform depression of image values, or 'cupping', results [19]. In addition, bone areas can 'spill over' into soft tissue, leading to a perceived increase in the attenuation

coefficient [49].

Beam hardening artifacts and correction schemes have been investigated since the early days of clinical X-ray CT. They arise in clinical practice because the X-ray attenuation properties of human tissue are energy-dependent, and X-ray sources in typical clinical scanners emit broad spectra. Using monoenergetic photon sources would eliminate beam hardening artifacts but is impractical for diagnostic CT because of SNR considerations. Beam hardening correction methods are therefore necessary for reconstructing artifact-free attenuation coefficient images from polyenergetic measurements. There are a variety of schemes for eliminating beam hardening artifacts in FBP images, including dual-energy imaging and post-processing techniques [2, 8, 44, 49, 59, 60]

The dual-energy approach [2, 33, 71] doubles the number of measurements in the CT reconstruction problem, and enables reconstruction of separate bone and soft tissue images. It is useful for tissue characterization and quantitative CT, but its major drawback is the requirement for two separate energy measurements.

Preprocessing approaches are based on the assumption that the energy dependence of soft tissue is similar to that of water. Knowledge of the energy dependence of the attenuation coefficient of water provides a one-to-one mapping between monoenergetic and polyenergetic measurements [49, 58, 60]. In preprocessing, one simply maps (or pre-corrects) the sinogram data to monoenergetic values and then one reconstructs the image from the corrected sinogram. Preprocessing works well with soft-tissue objects, but is poor when high Z materials, such as bone, are present. Preprocessing is often the first step in bone correction algorithms [49, 60].

The method described by Joseph and Spital (henceforth referred to as JS) [48, 49, 65] is a post-processing technique that corrects for soft tissue and dense bone

distortions. The method involves an initial FBP reconstruction and segmentation of the image into soft tissue and bone images. The segmented images are forward projected. The bone projection provides an estimate of the amount of nonlinear beam hardening distortion, which is then corrected for. Applying FBP to the corrected data provides an image with reduced artifacts. It is also possible to iterate the procedure several times to improve the resulting image. Yan *et al.* developed an iterative, but non-statistical, beam hardening correction method [79]. It also assumes two categories of materials and iteratively computes their volume fraction at each pixel.

Statistical reconstruction algorithms for monoenergetic transmission measurements have been very effective in providing attenuation maps for attenuation correction in emission tomography [25,26,30]. Such algorithms are based on a transmission model similar to Beer's law. Researchers have also applied monoenergetic algorithms to CT [13,53,57] with monoenergetic X-rays either explicitly or implicitly assumed. Since clinical CT measurements are polyenergetic, we show later in this chapter that images reconstructed from polyenergetic CT measurements using a monoenergetic reconstruction algorithm [22,26] suffer from beam hardening artifacts. This further points out the need for a polyenergetic statistical algorithm.

Recently, De Man *et al.* proposed a polyenergetic statistical approach [14] that models the object attenuation as a linear combination of the attenuation properties of two base substances. The algorithm uses the polyenergetic source spectrum and does not need a pre-segmented image. The algorithm can reconstruct mixed pixels but is not guaranteed to monotonically increase the likelihood each iteration.

In this chapter[1], we present and demonstrate the effectiveness of a polyenergetic statistical reconstruction algorithm for X-ray CT. The algorithm is based on a

---

[1]The results presented in this chapter are based on [22].

physical model that accounts for the polyenergetic X-ray source spectrum and the measurement nonlinearities caused by energy-dependent attenuation. The algorithm iteratively minimizes surrogate functions to the Poisson likelihood. Although one version of the algorithm is monotonic, we use ordered subsets [47] to accelerate the algorithm, which compromises monotonicity. The algorithm can also easily accommodate scatter estimates. Similar to the conventional approach (JS), the proposed algorithm requires a segmentation map of the different tissues in the object. We relax this requirement in the next chapter.

This chapter is organized as follows. The next section discusses polyenergetic X-ray CT and derives the iterative algorithm. In Section 4.3 we present simulation results that compare JS-corrected FBP with the proposed statistical reconstruction algorithm. Section 4.4 summarizes the results and outlines potential extensions of this work.

4.2 Polyenergetic X-ray CT

4.2.1 Statistical Model

We model the measurements as independently distributed Poisson random variables [25] that are contaminated by extra background counts, caused primarily by scatter. The approach generalizes to more sophisticated models [33]. Additive detector read-out noise can be accounted for in a several ways [67]. We assume the following statistical model for the measurements:

$$Y_i \sim \text{Poisson}\{\int I_i(\mathcal{E}) e^{-\int_{L_i} \mu(x,y,\mathcal{E})dl} d\mathcal{E} + r_i\}, \quad i = 1, ..., N, \tag{4.1}$$

where $\mu(x, y, \mathcal{E})$ is the unknown spatially- and energy-dependent attenuation map of the object. The integral in the exponent is taken over the line L_i and $I_i(\mathcal{E})$ incorporates the energy dependence of both the incident ray source spectrum and

the detector sensitivity. The assumed known scatter mean is denoted by r_i and N is the number of rays.

4.2.2 Object Model for Polyenergetic X-ray CT

We parameterize object space using square pixels. We assume that the object is comprised of K known materials (tissues). For any particular tissue, we model the attenuation coefficient as the product of the mass attenuation coefficient and the tissue density [2, 49, 71]. For pixel j, the attenuation coefficient is the weighted sum of the attenuation coefficients of all the tissues present in the pixel. Expressed mathematically,

$$\mu(x, y; \mathcal{E}) = \sum_{j=1}^{p} \mu_j(\mathcal{E}) b_j(x, y), \tag{4.2}$$

where $b_j(x, y)$ is the square-pixel basis function and $\mu_j(\mathcal{E})$ is the unknown energy-dependent linear attenuation coefficient of the materials in pixel j. Expressing $\mu_j(\mathcal{E})$ using the mass attenuation coefficient/density model gives

$$\mu_j(\mathcal{E}) = m(\mathcal{E})\rho_j = \rho_j \sum_{k=1}^{K} m_k(\mathcal{E}) f_k^j \tag{4.3}$$

where ρ_j is the density of pixel j, $\{m_k(\mathcal{E})\}_{k=1}^{K}$ are the mass attenuation coefficient of the K tissues that constitute the object and f_k^j is the fraction of tissue k in pixel j. We assume that $\{m_k(\mathcal{E})\}_{k=1}^{K}$ are known functions[2]. With the introduction of the tissue fractions, the number of unknowns appears to be $K \times p$. Oftentimes in clinical practice, $K = 2$ (bone and soft tissue) so the number of unknowns in the polyenergetic problem is twice that of a monoenergetic problem[3] This seems inevitable in dealing with the polyenergetic problem, since at each pixel the reconstruction algorithm must estimate both density and energy information.

---

[2]Data available for anatomical tissues and other materials from [46]

[3]Although we develop the algorithm for an arbitrary K, in practice the number of materials is limited by the spectral properties of human tissues in the diagnostic range. Typically, $K = 2$ or $K = 3$.

Multi-energy scanning addresses the increase in the number of unknowns by essentially providing a proportional increase in the information available [2,69]. Dual-energy imaging has already shown promising results with iterative reconstruction when the object is assumed to consist of bone and soft tissue only [33,71].

Another possible strategy is to assume that each pixel is either completely bone or completely soft tissue, and that a segmentation map of the object is available *a priori*. This approach is used in the JS method, and we will adopt it here as we derive the algorithm [22]. This approach works well when a good FBP image of an object with distinct anatomical features is available. The FBP reconstruction provides an initial image and a segmentation map for the iterative algorithm. This approach does not allow mixed pixels, which are likely to arise at tissue boundaries. It also will not work for objects containing mineral solutions. We will generalize this restrictive model in the next chapter [24]. For now, we set $f_j^k = 1$ if the jth pixel belongs to the kth material, and $f_j^k = 0$ otherwise.

Denoting the system matrix by $\mathbf{A} = \{a_{ij}\}$ where $a_{ij} = \int_{L_i} b_j(x, y) \, dl$, we can write the line integral of the attenuation coefficient as

$$
\begin{aligned}
\int_{L_i} \mu(x, y; \mathcal{E}) \, dl &= \sum_{j=1}^{p} \sum_{k=1}^{K} m_k(\mathcal{E}) \rho_j f_k^j(\rho_j) \int_{L_i} b_j(x, y) \, dl \\
&= \sum_{j=1}^{p} \sum_{k=1}^{K} m_k(\mathcal{E}) \rho_j f_k^j(\rho_j) a_{ij}.
\end{aligned}
\tag{4.4}
$$

We make the following definitions:

$$
a_{ij}^k \triangleq a_{ij} f_j^k
\tag{4.5}
$$

$$
s_i^k(\rho) \triangleq \sum_{j=1}^{p} a_{ij} f_j^k(\rho_j) \rho_j
\tag{4.6}
$$

$$
= \sum_{j=1}^{p} a_{ij}^k \rho_j
\tag{4.7}
$$

$$
\underline{s}_i(\rho) = (s_i^1, s_i^2, ..., s_i^K),
\tag{4.8}
$$

where s_i^k denotes the "effective thickness" (actually, the area density) of the kth tissue type along the ith ray. From (4.1), (4.3) and the definitions above, the mean of the measured data along path L_i is

$$
\begin{aligned}
E[Y_i|\rho] &= \int I_i(\mathcal{E}) \exp\left(-\sum_{k=1}^{K} m_k(\mathcal{E}) s_i^k(\rho)\right) d\mathcal{E} + r_i \\
&= \int I_i(\mathcal{E}) e^{-\underline{m}'(\mathcal{E})\underline{s}_i(\rho)} d\mathcal{E} + r_i \triangleq \bar{Y}_i(\underline{s}_i(\rho)) + r_i
\end{aligned}
\tag{4.9}
$$

where $\underline{m}'(\mathcal{E}) = [m_1(\mathcal{E}), ..., m_K(\mathcal{E})]$. We have expressed the measurements in terms of the vector function $\underline{s}_i$ which has as its elements the line integrals of the K different material densities. Given the X-ray spectrum, we tabulate the values of $\bar{Y}_i(\cdot)$ and its gradient $\nabla \bar{Y}_i(\cdot) = [\frac{\partial \bar{Y}_i}{\partial s^1}, \ldots, \frac{\partial \bar{Y}_i}{\partial s^K}]$ over the range of arguments $[s^1, \ldots, s^K]$ that correspond to representative objects.

The goal of the algorithm is to estimate the density coefficient vector $\rho = [\rho_1, \ldots, \rho_p]'$. Rather than estimating K vector quantities of length p, each representing the density of one material, the assumption of non-overlapping materials enables us to keep the number of unknowns equal to p, as is the case in the monoenergetic model. This is possible only if prior segmentation of the object is available. This segmentation is also necessary for the JS technique [49].

4.2.3   Polyenergetic Model Cost Function

We now express the Poisson negative log-likelihood in terms of the vector density ρ and the vector function $\underline{s}_i$. To derive the algorithm, we use the optimization transfer principle three times: first using the multiplicative convexity property [15]; second using parabola surrogates [25] and lastly De Pierro's additive convexity trick [16]. The successive applications of the optimization transfer principle yield a separable and simple surrogate function that is easier to minimize than the negative log-likelihood.

Recall that the function $\bar{Y}_i(\underline{s}_i(\rho))$ in (4.9) represents the ideal expected value of the measurement Y_i at the ith detector. Using $E[Y_i|\rho]$ in (4.9) gives the following negative log-likelihood in the polyenergetic case:

$$-L(\rho) = \sum_{i=1}^{N} -Y_i \log(E[Y_i|\rho]) + E[Y_i|\rho] \tag{4.10}$$

$$-L(\rho) = \sum_{i=1}^{N} h_i(\bar{Y}_i(\underline{s}_i(\rho)) + r_i) \tag{4.11}$$

$$h_i(t) \triangleq -Y_i \log t + t. \tag{4.12}$$

When the system matrix $\mathbf{A}$ has full column rank and the data is noise-free, minimizing the negative likelihood would give a perfect result. In reality, the data is noisy and maximum likelihood (ML) will give a very noisy reconstruction due to the ill-posedness of the problem, hence the need for regularization.

We regularize by adding a penalty term to the likelihood function. We consider pairwise regularization of the following form:

$$R(\mu) = \sum_{j=1}^{p} \sum_{k \in \mathcal{N}_j} \psi(\mu_j - \mu_k) \tag{4.13}$$

where ψ is a potential function and $\mathcal{N}_j$ is some neighborhood of pixel j. For our simulation and experimental results, we used the convex edge-preserving Huber penalty

$$\psi(x; \delta) = \begin{cases} \frac{x^2}{2}, & x < \delta \\ \delta|x| - \frac{\delta^2}{2}, & x \geq \delta. \end{cases} \tag{4.14}$$

Combining the likelihood with a penalty gives a penalized-likelihood (PL) cost function:

$$\Phi(\rho) = -L(\rho) + \beta R(\rho) \tag{4.15}$$

where β is a scalar that controls the tradeoff between the data-fit and the penalty terms. The goal of the reconstruction technique becomes to minimize (4.15) subject

to certain object constraints such as non-negativity:

$$\hat{\rho} = \begin{array}{c} \text{argmin} \\ \rho \geq 0 \end{array} \Phi(\rho). \tag{4.16}$$

An iterative algorithm is needed to perform the minimization (4.16). For simplicity, we focus on the likelihood term in the algorithm derivation. Similar results follow for the penalty term.

The difficulty arises with the argument of $h_i(\cdot)$, which is nonlinear in (4.11). Our first goal is to move the integral in (4.9) outside the (convex) function h_i. Towards that end, define:

$$t_i(\mathcal{E}, \underline{s}) \triangleq e^{-\underline{m}'(\mathcal{E})\underline{s}} + \bar{r}_i \tag{4.17}$$

$$\bar{r}_i \triangleq \frac{r_i}{\displaystyle\int I_i(\mathcal{E})d\mathcal{E}} \tag{4.18}$$

$$b_i^n(\mathcal{E}) \triangleq \frac{\bar{Y}_i(\underline{s}_i^n)}{t_i(\mathcal{E}, \underline{s}_i^n)}, \tag{4.19}$$

where $\underline{s}_i^n = \underline{s}_i(\rho^n)$. With the above definitions and (4.9):

$$\bar{Y}_i(\underline{s}) + r_i = \int I_i(\mathcal{E})t_i(\mathcal{E}, \underline{s}) \, d\mathcal{E}$$

$$= \int \frac{I_i(\mathcal{E})}{b_i^n(\mathcal{E})} t_i(E, \underline{s}) b_i^n(\mathcal{E}) \, d\mathcal{E}. \tag{4.20}$$

Since,

$$\int \frac{I_i(\mathcal{E})}{b_i^n(\mathcal{E})} d\mathcal{E} = 1 \tag{4.21}$$

$$\frac{I_i(\mathcal{E})}{b_i^n(\mathcal{E})} \geq 0, \tag{4.22}$$

we can use the convexity of the function $h_i(t)$ in (4.12) as follows [15, 82]:

$$h_i(\bar{Y}_i(\underline{s}) + r_i) = h_i\left(\int \frac{I_i(\mathcal{E})}{b_i^n(\mathcal{E})} t_i(\mathcal{E}, \underline{s}) b_i^n(\mathcal{E}) \, d\mathcal{E}\right)$$

$$\leq \int \frac{I_i(\mathcal{E})}{b_i^n(\mathcal{E})} h_i\left(t_i(\mathcal{E}, \underline{s}) b_i^n(\mathcal{E})\right) \, d\mathcal{E}.$$

Combining with (4.11) gives the following surrogate Q_1 for the negative of the polyenergetic log-likelihood:

$$
\begin{aligned}
-L(\rho) &= \sum_{i=1}^{N} h_i \left(\bar{Y}_i(\underline{s}_i(\rho)) + r_i \right) & (4.23) \\
&\leq \sum_{i=1}^{N} \int \frac{I_i(\mathcal{E})}{b_i^n(\mathcal{E})} h_i \left(t_i(\mathcal{E}, \underline{s}_i(\rho)) b_i^n(\mathcal{E}) \right) d\mathcal{E} \\
&\stackrel{\triangle}{=} Q_1(\rho; \rho^n). & (4.24)
\end{aligned}
$$

It is straightforward to verify that Q_1 satisfies the conditions of the optimization transfer principle [31]. The surrogate Q_1 is simpler than the actual likelihood because the energy integral is outside of the log operation. It is not, however, quadratic. We next apply optimization transfer to Q_1 to derive a paraboloidal surrogate. Such a surrogate is desirable because it is easily minimized. The first step is to express h_i using a quadratic surrogate:

$$
\begin{aligned}
h_i(t_i(\mathcal{E}, \underline{s}) b_i^n(\mathcal{E})) &= h_i \left(b_i^n(\mathcal{E}) e^{-\underline{m}'(\mathcal{E}) \underline{s}} + b_i^n(\mathcal{E}) \bar{r}_i \right) \\
&\stackrel{\triangle}{=} g_i^n(\underline{m}'(\mathcal{E}) \underline{s}, \mathcal{E}) \\
&\leq q_i^n(\underline{m}'(\mathcal{E}) \underline{s}, \mathcal{E}), & (4.25)
\end{aligned}
$$

where

$$
q_i^n(l, \mathcal{E}) = g_i^n(l_i^n, \mathcal{E}) + \dot{g}_i^n(l_i^n, \mathcal{E})(l - l_i^n) + \frac{1}{2} C_i^n(\mathcal{E})(l - l_i^n)^2. \quad (4.26)
$$

We must choose the curvature $C_i^n(\mathcal{E})$ to ensure that $q_i^n(l, \mathcal{E})$ satisfies the conditions for a surrogate (if we seek a monotone algorithm). Combining (4.25) and (4.24), the overall paraboloidal surrogate is:

$$
Q_2(\rho; \rho^n) = \sum_{i=1}^{N} \int \frac{I_i(\mathcal{E})}{b_i^n(\mathcal{E})} q_i^n(\underline{m}'(\mathcal{E}) \underline{s}_i(\rho), \mathcal{E}) d\mathcal{E}. \quad (4.27)
$$

Next we derive a separable surrogate that lends itself easily to parallelization

using De Pierro's additive convexity trick [15, 16]. First, define

$$\lambda_{ij}(\mathcal{E}) \triangleq \sum_{k=1}^{K} m_k(\mathcal{E}) a_{ij}^k$$

$$\mathbf{\Lambda}(\mathcal{E}) \triangleq \{\lambda_{ij}(\mathcal{E})\}.$$

We rewrite the density line integrals as follows

$$\underline{m}'(\mathcal{E})\,\underline{s}_i(\rho) = \sum_{j=1}^{p} \sum_{k=1}^{K} m_k(\mathcal{E}) a_{ij}^k \rho_j = \sum_{j=1}^{p} \lambda_{ij}(\mathcal{E}) \rho_j$$

$$= \sum_{j=1}^{p} \alpha_{ij} \left(\frac{\lambda_{ij}(\mathcal{E})}{\alpha_{ij}} (\rho_j - \rho_j^n) + [\mathbf{\Lambda}(\mathcal{E})\rho^n]_i \right),$$

where the α_{ij}'s satisfy

$$\sum_{j=1}^{p} \alpha_{ij} = 1, \, \forall\, i, \, \alpha_{ij} \geq 0.$$

Using the convexity of q_i^n (in its first argument) yields

$$q_i^n(\underline{m}'(\mathcal{E})\underline{s}_i(\rho), \mathcal{E}) = q_i^n \left(\sum_{j=1}^{p} \alpha_{ij} \left(\frac{\lambda_{ij}(\mathcal{E})}{\alpha_{ij}} (\rho_j - \rho_j^n) + [\mathbf{\Lambda}(\mathcal{E})\rho^n]_i \right), \mathcal{E} \right)$$

$$\leq \sum_{j=1}^{p} \alpha_{ij} q_i^n \left(\frac{\lambda_{ij}(\mathcal{E})}{\alpha_{ij}} (\rho_j - \rho_j^n) + [\mathbf{\Lambda}(\mathcal{E})\rho^n]_i, \mathcal{E} \right). \qquad (4.28)$$

The final separable paraboloidal surrogate to minimize is:

$$Q(\rho; \rho^n) = \sum_{j=1}^{p} \sum_{i=1}^{N} \int \frac{I_i(\mathcal{E})}{b_i^n(\mathcal{E})} \alpha_{ij} q_i^n \left(\frac{\lambda_{ij}(\mathcal{E})}{\alpha_{ij}} (\rho_j - \rho_j^n) + [\mathbf{\Lambda}(\mathcal{E})\rho^n]_i, \mathcal{E} \right) d\mathcal{E}. \qquad (4.29)$$

4.2.4 Iterative Algorithm for Polyenergetic CT

To derive the proposed algorithm, take the first derivative of the surrogate and set it equal to zero. This gives the following update expression (ignoring regularization):

$$\rho_j^{n+1} = \left[\rho_j^n - \frac{\left. \frac{\partial Q(\rho; \rho^n)}{\partial \rho_j} \right|_{\rho=\rho^n}}{\left. \frac{\partial^2 Q(\rho; \rho^n)}{\partial \rho_j^2} \right|_{\rho=\rho^n}} \right]_+ , j = 1, \ldots, p. \qquad (4.30)$$

The derivatives of the surrogate, evaluated at the current iterate ρ^n are:

$$
\begin{aligned}
\left.\frac{\partial Q}{\partial \rho_j}\right|_{\rho=\rho^n} &= \sum_{i=1}^{N} \int \frac{I_i(\mathcal{E})}{b_i^n(\mathcal{E})} \lambda_{ij}(\mathcal{E}) \dot{g}_i^n(\underline{m}'(\mathcal{E})\underline{s}_i^n, \mathcal{E}) d\mathcal{E} \\
&= \sum_{i=1}^{N} \left(\frac{Y_i}{\bar{Y}_i(\underline{s}_i^n)} - 1\right) \times \int \lambda_{ij}(\mathcal{E}) I_i(\mathcal{E}) e^{-\underline{m}'(\mathcal{E})\underline{s}_i^n} d\mathcal{E} \\
&= \sum_{i=1}^{N} \sum_{k=1}^{K} a_{ij} f_j^k \left(\frac{Y_i}{\bar{Y}_i(\underline{s}_i^n)} - 1\right) \nabla_k \bar{Y}_i(\underline{s}_i^n) \\
&= -\left.\frac{\partial L}{\partial \rho_j}\right|_{\rho=\rho^n} \\
\frac{\partial^2 Q}{\partial \rho_j^2} &= \sum_{i=1}^{N} \int \frac{I_i(\mathcal{E})}{b_i^n(\mathcal{E})} \frac{C_i^n(\mathcal{E})}{\alpha_{ij}} \lambda_{ij}^2(\mathcal{E}) \, d\mathcal{E}.
\end{aligned}
\tag{4.31}
$$

The second derivative in (4.31) has two terms that are iteration dependent, $b_i^n(\mathcal{E})$ and the curvature $C_i^n(\mathcal{E})$. The curvature, in particular, influences the rate of convergence of the algorithm [25]. We next explore some possibilities for it.

4.2.5 Curvature

If one desires a monotonic algorithm, then it is necessary to choose curvatures such that (4.26) satisfies the condition of the optimization transfer principle. A simple choice for the curvature is the maximum second derivative in the feasible region for the projections. The closed form expression for the maximum curvature is [25]:

$$
C_i^n(\mathcal{E}) = \left[b_i^n(\mathcal{E}) - \frac{Y_i \bar{r}_i}{(1 + \bar{r}_i)^2} \right]_+ \leq [b_i^n(\mathcal{E})]_+ .
\tag{4.32}
$$

This inequality always holds since $r_i \geq 0$ and $Y_i \geq 0$. We can use the simpler right hand side of (A.5) and still have a monotonic algorithm. This is equivalent to using the maximum curvature when the background term $\bar{r}_i$ is small.

The curvature affects the step size that the algorithm takes towards the mini-mizer. The maximum curvature results in small steps, and hence a slowly converging

algorithm. Plugging the right hand side of (A.5) in (4.31) gives the following:

$$\frac{\partial^2 Q}{\partial \rho_j^2} = \sum_{i=1}^{N} \int \frac{I_i(\mathcal{E})}{\alpha_{ij}} \lambda_{ij}^2(\mathcal{E}) \, d\mathcal{E}. \tag{4.33}$$

The above equation has no iteration-dependent terms and can be easily precomputed.

Another possible curvature, given in [25], is optimal in the sense that it satisfies the conditions of optimization transfer while keeping the step size as large as possible. The optimal curvature must be computed at every iteration. It therefore accelerates convergence, but requires more computation per iteration.

4.2.6   Precomputed Curvature

By relaxing the monotonicity requirement, we can develop faster algorithms. Since we use ordered subsets to implement the algorithms, monotonicity is compromised anyway.

We can choose a curvature in (4.26) such that $q_i^n(l, \mathcal{E}) \approx g_i^n(l, \mathcal{E})$, rather than requiring inequality. In this case the paraboloids are quadratic approximations to the likelihood that are updated at every iteration. A reasonable curvature to use is the second derivative of g_i^n evaluated at the point that minimizes the function, $l_i^{\min} = \log(b_i^n(\mathcal{E})/(Y_i - b_i^n(\mathcal{E})\bar{r}_i))$. The curvature becomes:

$$\begin{aligned} C_i^n(\mathcal{E}) &= \ddot{g}_i^n \left(\log \frac{b_i^n(\mathcal{E})}{Y_i - b_i^n(\mathcal{E})\bar{r}_i}, \mathcal{E} \right) \\ &\approx Y_i. \end{aligned} \tag{4.34}$$

To simplify matters further, define

$$\mathcal{E}_{\text{eff}} \triangleq \frac{\displaystyle\int \mathcal{E} I_i(\mathcal{E}) d\mathcal{E}}{\displaystyle\int I_i(\mathcal{E}) d\mathcal{E}} \tag{4.35}$$

as the effective energy of the X-ray beam. We now make the approximation of evaluating $\lambda_{ij}(\mathcal{E})$ in (4.31) at $\mathcal{E}_{\text{eff}}$ and pull it out of the integral. The remaining

energy terms integrate to unity. With the assumption of non-overlapping materials and

$$\alpha_{ij} = \frac{\sum_k a_{ij}^k}{\sum_k \sum_j a_{ij}^k} = \frac{a_{ij}}{\sum_j a_{ij}} = \frac{a_{ij}}{\gamma_i},$$

the second derivative reduces to:

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial \rho_j^2} &\approx \sum_{i=1}^{N} \frac{\lambda_{ij}^2(\mathcal{E}_{\mathrm{eff}})}{\alpha_{ij}} Y_i \\
&\approx \sum_{k=1}^{K} m_k^2(\mathcal{E}_{\mathrm{eff}}) \sum_{i=1}^{N} a_{ij} \gamma_i Y_i \qquad (4.36) \\
&\triangleq d_j.
\end{aligned}
$$

This expression is completely independent of iteration. It can be pre-computed and stored, further accelerating the algorithm. Here is the overall ordered subsets version of the algorithm:

- Precompute d_j using (4.36).

- Tabulate $\bar{Y}_i(\cdot)$ and $\nabla \bar{Y}_i(\cdot)$ over an appropriate range of break points.

- initialize with $\hat{\rho}$.

- **for** each iteration $n = 1, \ldots, \mathrm{niter}$

  - **for** each subset $S = 1, \ldots, M$

    * compute $\hat{s}_i^k = \sum_{j=1}^{p} a_{ij}^k \hat{\rho}_j$ for $k = 1, \ldots, K$. Set $\hat{\underline{s}}_i = [\hat{s}_i^1, \ldots, \hat{s}_i^K]$.

    * compute $\bar{Y}_i(\hat{\underline{s}}_i)$ and $\nabla \bar{Y}_i(\hat{\underline{s}}_i)$ by interpolation from the precomputed tables.

    * evaluate

    $$\hat{N}_j = \sum_{i \in S} \sum_{k=1}^{K} a_{ij} f_j^k \left(\frac{Y_i}{\bar{Y}_i(\hat{\underline{s}}_i)} - 1 \right) \nabla_k \bar{Y}_i(\hat{\underline{s}}_i)$$

$*$ compute for $j = 1, \ldots, p$

$$\hat{\rho}_j = \left[\hat{\rho}_j - \frac{M\hat{N}_j + \beta \left. \dfrac{\partial S}{\partial \rho_j} \right|_{\rho = \hat{\rho}}}{d_j + \beta \left. \dfrac{\partial^2 S}{\partial \rho_j^2} \right|_{\rho = \hat{\rho}}} \right]_+ , \qquad (4.37)$$

– **end**

• **end**

where S in (4.37) denotes an appropriate surrogate for the regularization penalty. If the optimal curvature [25] or maximum curvature (A.5) are used, this algorithm will monotonically decrease the cost function each iteration when one subset is used. Using ordered subsets and the pre-computed curvature destroys monotonicity, but significantly accelerates progress in the early iterations.

4.3   Simulation Results

We assess the effectiveness of our algorithm with simulated polyenergetic Poisson X-ray projections. We assume a parallel beam geometry, but the algorithm applies equally well to other geometries. The image FOV is 40 cm and the rotation range is 180°. The simulated measurements are free of scatter and detector readout noise.

We simulate polyenergetic transmission data with the polyenergetic 140 kVp spectrum (mean 67.12 keV and standard deviation 17.76 keV) shown in Fig. 4.1 and a blank scan of 4.87×10^6 counts/detector. The blank scan value is realistic and mimics a 120 kVp, 170mAs scan protocol [75]. The spectrum was obtained from Monte Carlo simulations of the setup in [70] and includes the effects of detector response.

We reconstruct the simulated data with FBP, monoenergetic statistical algorithm (PWLS-OS) [22, 26] and polyenergetic statistical algorithm. We use the soft-tissue

and JS methods to remove beam hardening artifacts in the FBP image. We also pick the FBP reconstruction parameters to give comparable resolution and noise properties to the statistical algorithms (the statistical algorithms can perform better in this regard, but our purpose is to illustrate beam hardening correction). All runs of the statistical algorithms use a Huber penalty with $\beta = 1000$, $\delta = 0.01$ cm^{-1} for monoenergetic reconstruction, and $\delta = 0.1$ g/cc for polyenergetic reconstruction. The parameter β is determined by trial and error and the parameter δ is chosen to preserve the contrast between soft tissue and bone.

The first phantom, shown in Fig. 4.2 is a 256×256 density phantom consisting of four high-density 'bone' disks ($\rho = 2$ g/cc) immersed in a water disk ($\rho = 1$ g/cc), which is surrounded by air ($\rho = 0$ g/cc). The pixel size is 1.6 mm. The data were simulated over 500 angular steps and 600 radial bins, 1.3 mm each. The colorbar adjacent to Fig. 4.2 illustrates the gray scale window used to view all the images in the figure. The iterative algorithms ran for 20 iterations and 20 subsets.

Fig. 4.2b-c show the reconstructed images (scaled to display density values) when algorithms that do not correct for beam hardening are used. Both FBP and PWLS-OS exhibit beam hardening artifacts: reduction in overall pixel values and dark streaks between high-density regions. Fig. 4.2c shows the importance of developing iterative algorithms based on polyenergetic physics.

Fig. 4.3a-b illustrate the image corrected by soft-tissue preprocessing and by the JS technique. The soft-tissue method, available on commercial scanners, leaves substantial artifacts. The JS technique post-corrects for most of the artifacts, but some persist. Simple thresholding (with threshold 1.5 g/cc) of the soft-tissue corrected FBP image provides the segmentation required by the JS technique.

The iterative reconstruction based on the polyenergetic model is shown in
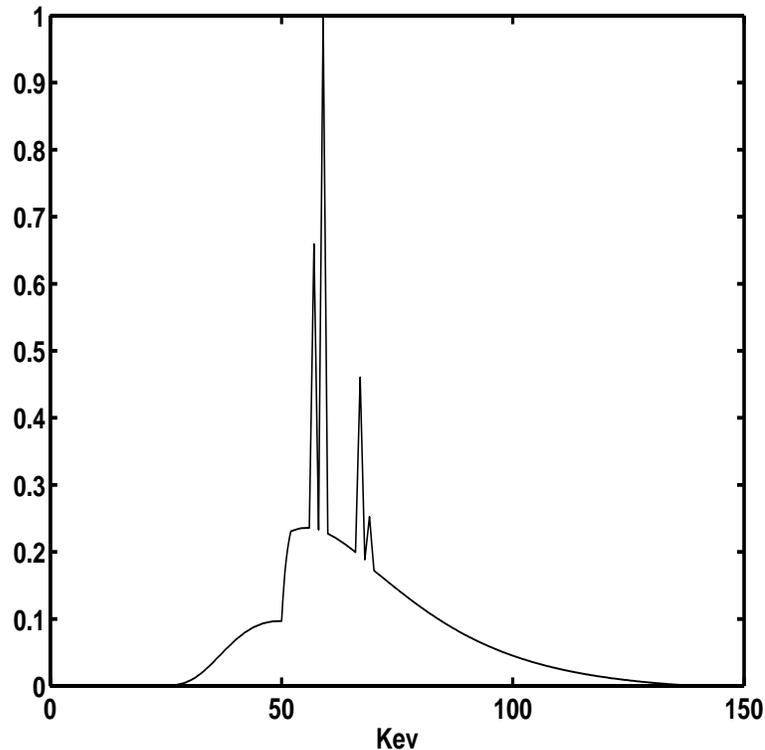
Figure 4.1: 140 kVp energy spectrum

Fig. 4.3c. The algorithm, was initialized with the JS corrected image. The object was classified into bone or soft tissue by segmenting the JS corrected FBP reconstruction using a density threshold of 1.5 g/cc. We choose the segmentation threshold such that the number of mismatched pixels with the true object classification is minimum. In a more realistic setting the true object is not available, and we address the segmentation issue in later work. The iterative algorithm significantly reduces artifacts, relative to JS, since it inherently accounts for the effects of the broad energy spectrum. The profile plots of the JS and the polyenergetic statistical algorithm images in Fig. 4.4 further delineate the difference in performance between the two methods. Table 4.1 lists the root mean squared (RMS) error of all the methods, relative to the true object, and shows that the polyenergetic statistical reconstruction has the

69

a. Density Phantom   **b. Uncorrected FBP**   **c. Monoenergetic Statistical Reconstruction**

1.2

1

0.8
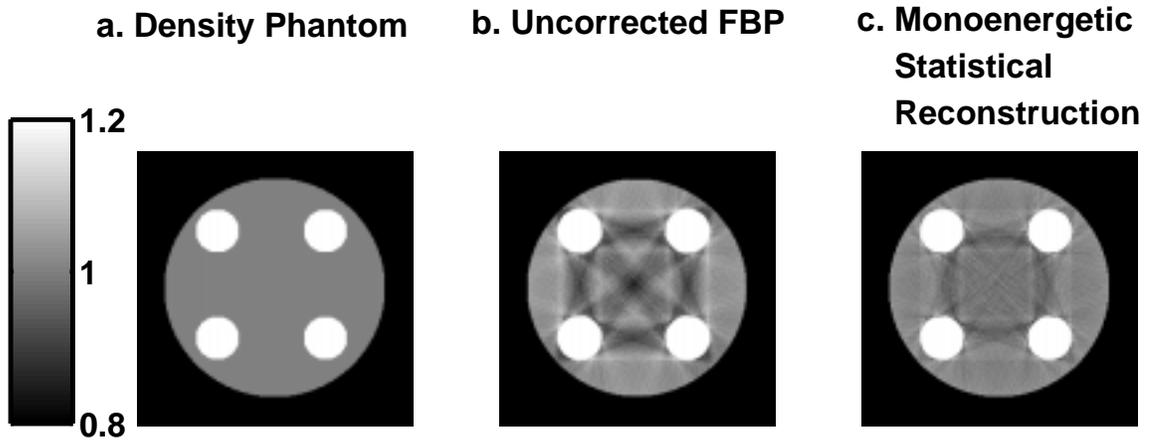
Figure 4.2: Bone/water density phantom results

a. Soft–tissue corrected FBP   **b. JS corrected FBP**   **c. Polyenergetic Statistical Reconstruction**

1.2

1

0.8
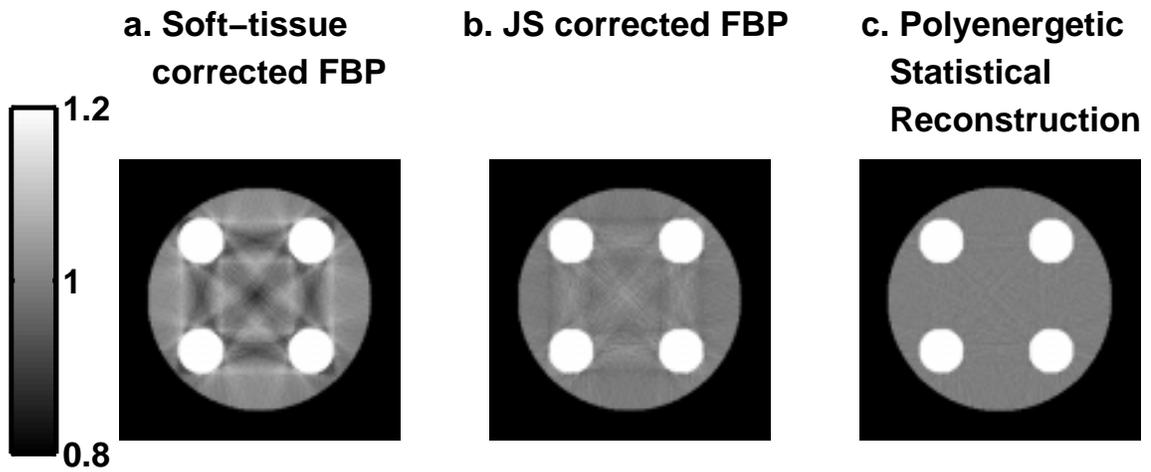
Figure 4.3: Bone/water density phantom results

Reconstruction Method	RMS Error
FBP	11.9 %
Soft Tissue	16.6 %
Joseph and Spital	4.9 %
Monoenergetic Iterative	9.6 %
Polyenergetic Iterative	2.2 %

Table 4.1: Root mean squared error (RMS) of the different reconstructions of the phantom in Fig. 4.2.

lowest error among all methods used. To compute the RMS error (and to display the images) for FBP and monoenergetic statistical reconstruction, the images were scaled by the appropriate mass attenuation coefficients to give density values.

To gain more confidence in our approach, we performed an additional experiment with this phantom where we simulated and reconstructed noise-free data (not shown). The results showed that polyenergetic iterative reconstruction was significantly more effective in reducing artifacts than JS, even when there was no noise. This is due to the fact that the iterative algorithm models the beam spectrum completely, whereas the JS method is approximate, even for noise-free data.

We also applied the different algorithms to the 512×250 object shown in Fig. 4.5a. We created this "true" object by manually segmenting a previously acquired real CT image, then assigning density to each anatomical structure. The density of the bones is in the range $1.6 - 2$ g/cc and the soft tissue densities vary from 0.9 to 1.1 g/cc. The pixel size is 0.8 mm and the sinogram has 700 angular bins and 500 radial bins, 1.0 mm each. The statistical algorithms ran for 10 iterations with 50 subsets. We use the JS image to initialize the polyenergetic iterative technique and the soft-tissue corrected image to determine the distribution of bone and soft-tissue regions with a threshold of 1.6 cm/gm^3. This threshold was chosen to minimize pixel mismatch with the true object.
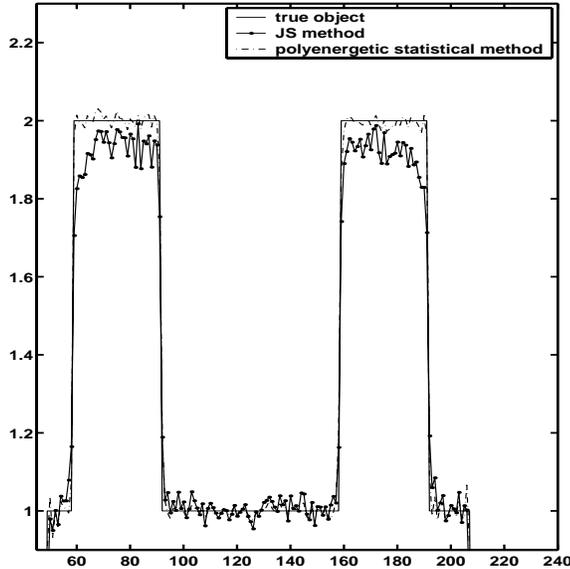
Figure 4.4: Profile plots of reconstructed images

The results are shown in Fig. 4.5 and Fig. 4.6. Uncorrected FBP and monoenergetic iterative algorithm images (scaled to display density values) suffer from beam-hardening artifacts, with streaks visible in the vicinity of bones. The soft-tissue, JS and polyenergetic iterative results are also shown, with the latter clearly yielding less artifacts. Table 4.2 also lists the RMS error for all methods, with polyenergetic statistical reconstruction having the smallest error.

For comparison, we also performed the polyenergetic reconstruction (not shown) with bone and water classification obtained from segmenting the true object. This yielded very similar results, so using the soft-tissue corrected image to determine the f_j^k's appears to be a practical approach.

Like the JS method, the algorithm we propose requires knowledge of the spectrum (or tables of log data) and a pre-segmented image, but gives considerably improved density images. With the knowledge of the mass attenuation coefficient, one can scale the resulting density images to obtain attenuation coefficient information at

a. True object

b. Unocrrected FBP

c. Monoenergetic statistical reconstruction

0.8   1   1.2
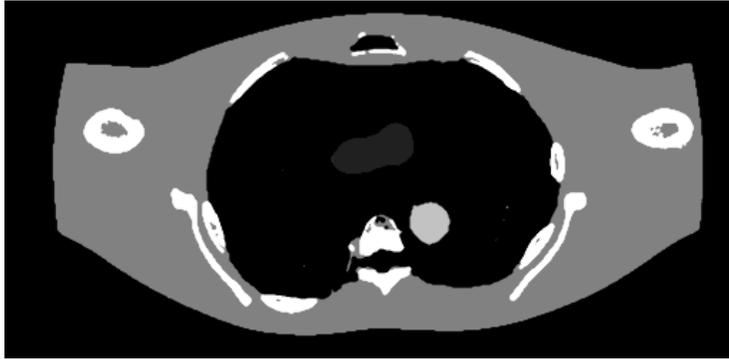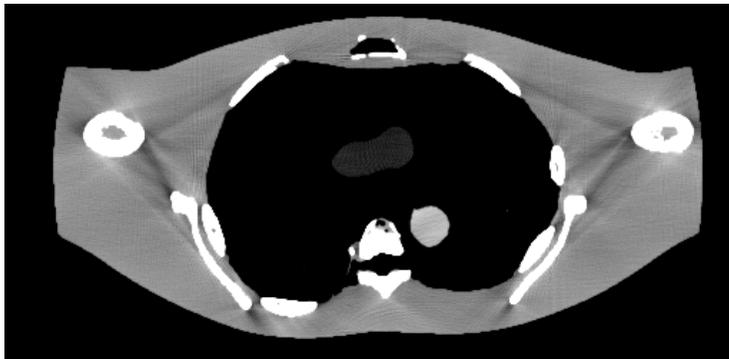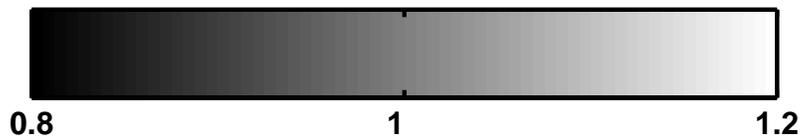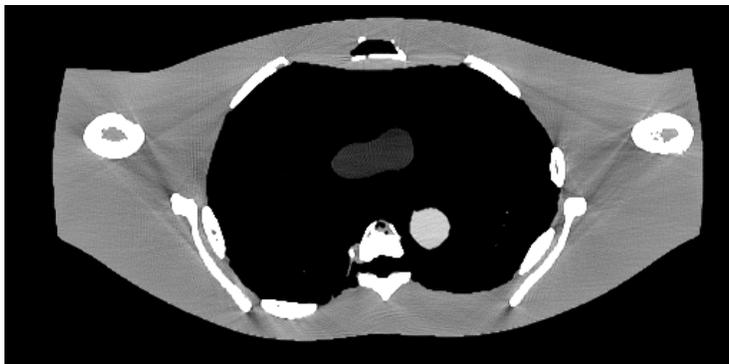
Figure 4.5: True object, uncorrected FBP and monoenergetic statistical reconstructions of simulated polyenergetic data.

Reconstruction Method	RMS Error
FBP	8.2 %
Soft Tissue	17.1 %
Joseph and Spital	6.1 %
Monoenergetic Iterative	6.8 %
Polyenergetic Iterative	2.5 %

Table 4.2: Root mean squared error (RMS) of the different reconstructions of the phantom in Fig. 4.5.

any energy using (4.3).

4.4   Conclusion

We have introduced a statistical iterative reconstruction algorithm for energy dependent X-ray attenuation that produces images with significantly reduced beam hardening artifacts. The algorithm is applicable for an arbitrary number of non-overlapping materials, and we demonstrate its effectiveness for bone and soft tissue objects. Unlike most other transmission CT iterative algorithms, our algorithm is based on a realistic polyenergetic model. Fig. 4.2c and Fig. 4.5c illustrate the severe artifacts that result when an iterative algorithm based on a monoenergetic model reconstructs an image from polyenergetic data.

By successive applications of the optimization transfer principle, the statistical algorithm minimizes a separable paraboloidal surrogate, hence it is parallelizable and fairly simple to implement. We also use ordered subsets and pre-computed surrogate curvatures to accelerate convergence and reduce computation. When one subset is used with appropriate curvatures, the algorithm monotonically decreases the cost function. This is about the most that can be said about convergence since the cost function is inherently not convex.

When compared with the post-processing technique of Joseph and Spital [49],

the statistical algorithm yielded fewer artifacts. The JS method estimates the line-integral dependent nonlinearity for each sinogram bin and then recalculates the line integrals. The statistical method needs no such post-processing since it inherently accounts for the nonlinearities. This is likely the reason for its superior performance.

Our results suggest that one can estimate the distribution of materials by thresholding a good FBP image, corrected with the soft-tissue technique. In the results of the earlier section, we chose the segmentation threshold to minimize the mismatch between the segmentation map and the true object. This is not possible in practice, and thresholding will be a subjective error-prone process. One could possibly improve on this approach by regenerating the segmentation after running one or more iterations of the proposed algorithm. Another limitation of pre-segmentation is that it does not permit pixels to contain mixtures of different materials, limiting the accuracy of the algorithm. To improve the accuracy of the algorithm, especially at material boundaries, we need to augment the set of variables with material fractions to be estimated at every pixel. In the next chapter, we will present two generalized object models that permit pixels to contain tissue mixtures and that do away with pre-segmentation. One of the models, the *displacement model* is appropriate for objects containing distinct anatomical structures, with mixed pixels occurring mostly at tissue boundaries. The other model, the *solution model*, is better suited for estimating the density of mineral solutions.

a. Soft–tissue corrected FBP

b. JS corrected FBP

c. Polyenergetic Statistical Reconstruction

0.8   1   1.2
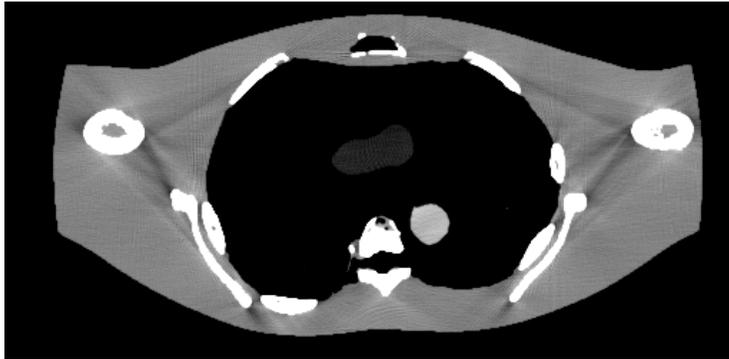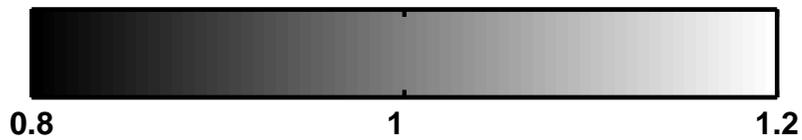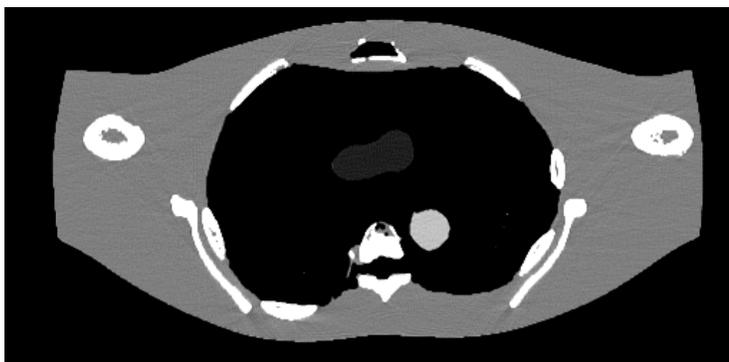
Figure 4.6: Soft-tissue, JS-corrected FBP and polyenergetic statistical reconstructions of simulated polyenergetic data.

CHAPTER V

Object Models for Polyenergetic CT for Quantitative Applications

5.1 Introduction

In the previous chapter we presented an iterative algorithm for reconstruction of polyenergetic X-ray CT data. The algorithm was based on a cost function that inherently accounted for polyenergetic X-ray attenuation. Since the algorithm incorporated energy dependence, constraints were necessary to limit the number of unknowns to equal the number of pixels. To keep the algorithm derivation simple, we assumed that an initial image could be accurately segmented into soft tissue and bone.

Requiring a segmented initial image was admittedly a restrictive condition, but it was similar to the requirements of the classical beam hardening method (JS). The polyenergetic CT algorithm was effective in giving reconstructions free of beam hardening artifacts, but it could not give accurate density values at boundary pixels or in objects where the water-bone decomposition did not hold. In this chapter[1], we augment the algorithm with object models that remove the pre-segmented image requirement. Constraints on the object are still necessary to keep the number of

---

[1]The results presented in this chapter are based on [24].

unknowns equal to the number of pixels.

In addition to the object models, we examine the robustness of the algorithm to mismatches in the X-ray source spectrum with simulated data. We also assess the algorithm performance in terms of the accuracy of the reconstructed values of mineral solution densities from real measurements [24].

5.2 Polyenergetic Statistical Model

As in the previous chapter, we assume the following statistical model for the measurements:

$$Y_i \sim \text{Poisson}\{\int I_i(\mathcal{E})e^{-\int_{L_i} \mu(x,y,\mathcal{E})dl}d\mathcal{E} + r_i\}, \quad i = 1, ..., N, \tag{5.1}$$

where $\mu(x, y, \mathcal{E})$ is the unknown spatially- and energy-dependent attenuation map of the object. The integral in the exponent is taken over the line L_i and $I_i(\mathcal{E})$ incorporates the energy dependence of both the incident ray source spectrum and the detector sensitivity. The assumed known scatter mean is denoted by r_i and N is the number of rays.

We parameterize object space using square pixels. We assume that the object is comprised of K known materials (tissues) and each pixel may contain a mixture of those materials. We model the attenuation coefficient as the product of the mass attenuation coefficient and the tissue density [2, 49, 71]. For pixel j, the attenuation coefficient is the weighted sum of the attenuation coefficients of all the tissues present in the pixel. Expressed mathematically,

$$\mu(x, y; \mathcal{E}) = \sum_{j=1}^{p} \mu_j(\mathcal{E})b_j(x, y), \tag{5.2}$$

where $b_j(x, y)$ is the square-pixel basis function and $\mu_j(\mathcal{E})$ is the unknown energy-dependent attenuation coefficient of the materials in pixel j. Expressing $\mu_j(\mathcal{E})$ using

the mass attenuation coefficient/density model gives

$$\mu_j(\mathcal{E}) = m(\mathcal{E})\rho_j = \sum_{k=1}^{K} m_k(\mathcal{E}) f_k^j \rho_j \tag{5.3}$$

where ρ_j is the density of pixel j, $\{m_k(\mathcal{E})\}_{k=1}^{K}$ are the mass attenuation coefficient of the K tissues that constitute the object and f_k^j is a unitless fraction that describes the contribution of material k to attenuation in pixel j. Unlike the work presented in the previous chapter, we do not assume prior knowledge of a water-bone segmentation map that predetermines the tissue fractions. We propose an alternative strategy that avoids segmentation yet can be applied to a single-kVp scan. Using the principle that high-density voxels tend to be composed of bone, we model the fraction of material k in pixel j as a predetermined function of the density of the pixel:

$$\mu_j(\mathcal{E}) = \sum_{k=1}^{K} m_k(\mathcal{E})\rho_j f_k^j(\rho_j). \tag{5.4}$$

By predefining the $\{f_k^j\}$ functions, there is only one unknown (ρ_j) per voxel, and reconstruction from a single scan is possible. Denoting the system matrix by $\mathbf{A} = \{a_{ij}\}$ where $a_{ij} = \int_{L_i} b_j(x, y)\, dl$, we can write the line integral of the attenuation coefficient as

$$\int_{L_i} \mu(x, y; \mathcal{E})\, dl = \sum_{j=1}^{p} \sum_{k=1}^{K} m_k(\mathcal{E})\rho_j f_k^j(\rho_j) \int_{L_i} b_j(x, y)\, dl$$

$$= \sum_{j=1}^{p} \sum_{k=1}^{K} m_k(\mathcal{E})\rho_j f_k^j(\rho_j) a_{ij}. \tag{5.5}$$

From (5.1), and following the notation of the Chapter IV, the mean of the measured data along path L_i is

$$E[Y_i|\rho] = \int I_i(\mathcal{E}) \exp\left(-\sum_{k=1}^{K} m_k(\mathcal{E}) s_i^k(\rho)\right) d\mathcal{E} + r_i$$

$$= \int I_i(\mathcal{E}) e^{-\underline{m}'(\mathcal{E})\underline{s}_i(\rho)} d\mathcal{E} + r_i \triangleq \bar{Y}_i(\underline{s}_i(\rho)) + r_i \tag{5.6}$$

where $\underline{m}'(\mathcal{E}) = [m_1(\mathcal{E}), ..., m_K(\mathcal{E})]$. We have expressed the mean measurements as a function of the vector $\underline{s}_i$ which has as its elements the line integrals of the K different material densities. Because $\bar{Y}_i(\underline{s}_i(\rho))$ is strongly nonlinear, it is easier to work in the log domain. Define

$$F_i(\underline{s}_i(\rho)) \triangleq - \log \frac{\bar{Y}_i(\underline{s}_i(\rho))}{I_i}, \tag{5.7}$$

where $I_i \triangleq \int I_i(\mathcal{E}) d\mathcal{E}$. For monoenergetic X-rays, F_i would be linear in $\underline{s}_i$. For polyenergetic X-rays, F_i's departure from linearity depends on the density and thickness of the object scanned and the range of X-ray energies. Rewriting (5.6) using F_i gives

$$E[Y_i|\rho] = I_i e^{-F_i(\underline{s}_i(\rho))} + r_i. \tag{5.8}$$

From knowledge of the X-ray spectrum or equivalent measurements, we can determine $F_i(\cdot)$ and its gradient, often in the form of a polynomial approximation [10]. We next discuss the tissue fraction functions $\{f_k^j(\rho_j)\}$.

5.3   Tissue Fraction Functions

The function $f_k^j(\rho_j)$ specifies a model for the fraction of the kth material in pixel j. The actual tissue fraction may depend on more than the pixel density. Ideally, we would want to estimate tissue fractions independently of density or any other parameter. However, that is apparently not possible without additional measurements, as pointed out in the previous section. It is therefore necessary to make model assumptions and impose constraints. The constraint imposed by the JS method and in [22] is that each $f_k^j(\rho_j)$ is either 1 or 0 as specified by the FBP segmentation. Our goal is to improve upon that model.

Although not necessary, we require $f_k^j(\rho_j) \geq 0$ and $\sum_k f_k^j(\rho_j) = 1$ for simplicity. The crux of our beam hardening model lies in our choice of the tissue fraction func-
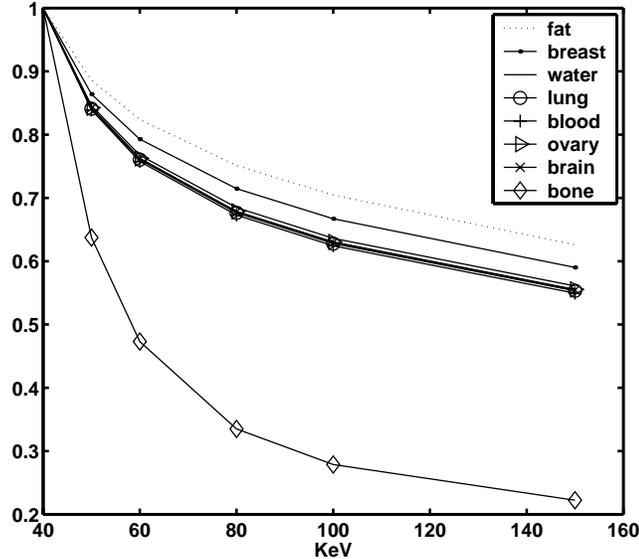
Figure 5.1: Mass attenuation coefficient of human tissues normalized at 40 keV.

tions. To design reasonable models that are as consistent as possible with attenuation physics, we examine the X-ray spectral characteristics of human tissues.

Fig. 5.1 illustrates the normalized mass attenuation coefficients of several human tissues in the diagnostic X-ray energy range [46]. Several tissues (ovary, brain, muscle, lung, blood) have spectral characteristics that are almost identical to those of water. Fat (adipose tissue) and breast tissue differ somewhat from water, and bone is significantly different. All soft tissues, including fat and breast, have densities in the range $0.95 - 1.06$ g/cc. Because soft tissues are spectrally similar to water, and have densities close to that of water, it is reasonable to use water as a base substance. Cortical bone is another natural choice for a base substance. The model could be augmented to include more base substances such as Iodine. We use $K = 2$ in (5.3) (water and bone) and model the attenuation coefficient of tissue as follows:

$$\mu_j(\mathcal{E}) \approx \left(m_w(\mathcal{E}) f_w^j(\rho_j) + m_b(\mathcal{E}) f_b^j(\rho_j) \right) \rho_j \qquad (5.9)$$

Tissue	Density g/cc	f_w
Fat (adipose tissue)	0.95	1.0
Breast	1.02	1.0
Water	1.0	1.0
Lung	1.05	0.99
Blood	1.06	0.99
Ovary	1.05	0.99
Brain	1.04	0.99
Muscle	1.05	1.0
Eye lens	1.07	1.0
Testis	1.04	1.0
Cortical Bone	1.92	0.0

Table 5.1: Densities and water fractions of human tissues. Most soft tissues have densities close to that of water. The water fraction f_w is computed from a weighted least squares fit to (5.10).

where $m_w(\mathcal{E})$ and $m_b(\mathcal{E})$ are the mass attenuation coefficients of the first (water) and second (bone) base substances. The pixel tissue fraction functions f_w^j and f_b^j determine to what extent the tissue in pixel j is spectrally water-like or bone-like, depending on the tissue density.

To gain more confidence in our model and insight into possible forms of f_w^j and f_b^j we used weighted (by a typical spectrum) least squares to compute the coefficients f_w and f_b in

$$m(\mathcal{E}) = m_w(\mathcal{E})f_w + m_b(\mathcal{E})f_b \qquad (5.10)$$

for the biological substances listed in Fig. 5.1. We also imposed the constraints that the solutions be in the interval $[0, 1]$, in addition to $f_w + f_b = 1$. The results are listed in table 5.1. From table 5.1, we observe that the water coefficient f_w is almost 1.0 for soft tissues and 0.0 for bone, reaffirming that soft tissues are spectrally similar to water and that water and bone can be used as distinct base substances. The pluses and circles in Fig. 5.2 illustrate the water and bone coefficients of the different tissues, respectively.
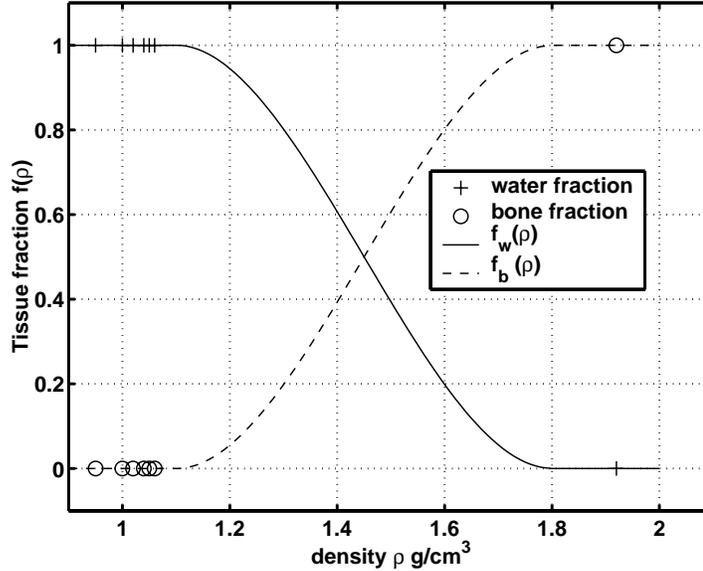
Figure 5.2: Water and bone fractions computed using constrained least squares fit to (5.10), and the displacement model third order polynomial tissue fraction functions $f_w(\rho)$, $f_b(\rho)$.

One possible scheme that emerges from examining Fig. 5.2 is to set $f_w(\rho) = 1$ when the density is less than some threshold value and 0 otherwise. This would be closely related to our earlier approach, where pixels are forced to be either bone or water and cannot contain mixtures of tissues [22].

Another approach would be to linearly interpolate the $f_j^k(\rho_j)$ functions between the water (density 1.0 g/cc) and bone (density 1.92 g/cc) coefficients. This implies that the energy dependence of substances between water and bone are linear combinations of the energy dependence of bone and water. This is similar to the approach of De Man *et al.,* [14]. The disadvantage of the piece-wise linear approach lies in the fact that many tissue types do not fit this spectral linear combination model. In fact, table 5.1 shows that soft tissues are better spectrally modeled as equivalent to water. Also, the derivatives of $f_j^k(\rho_j)$ would have complicating discontinuities in the piecewise linear approach.

5.3.1 Displacement Model

To model mixed pixels, we propose using the functions plotted in Fig. 5.2 using the solid and dotted lines. These are third-order polynomial functions of the tissue density. They have continuous first- and second-order derivatives, and satisfy $f_w(\rho) + f_b(\rho) = 1$. In essence, this choice models tissues that have densities close to that of water and that are spectrally similar to water as 'dense water'. The model allows tissue density to vary from that of water, while assuming it is spectrally equivalent to water. This approach avoids the technical complications posed by the derivatives of piecewise linear functions, and has the potential for better accuracy, particularly for tissues that are spectrally similar to water but have densities larger than water.

The exact shape of the smooth curves is subjective and may be situation dependent. If it were known that an object contained material with density in the mid-range of say, 1.3-1.6 g/cc, the curves can be altered so that the algorithm can correctly account for that material. The polynomials in Fig. 5.2 work well for anatomical structures like the phantom shown in Fig. 5.6.

The polynomials shown in Fig. 5.2 do not match the partial volume fractions for tissue mixtures, i.e, they are *not* mass fractions. Fig. 5.3 illustrates the mass fractions of several soft tissues mixed with bone. For example, a 50% by mass mixture of bone and water would have a density of 1.315 g/cc. Mixtures of other soft tissues with bone would have different 50% density points and their mass fractions would be different functions of density. To estimate mixtures accurately, an algorithm would have to determine what mixture(s) a particular pixel contains and then parameterize the mass fractions accordingly. This is a prohibitively complex problem. Our major concern in designing the polynomials of Fig. 5.2 was to model the spectral properties of *homogeneous* soft-tissue pixels as equivalent to water. Accurately determining the
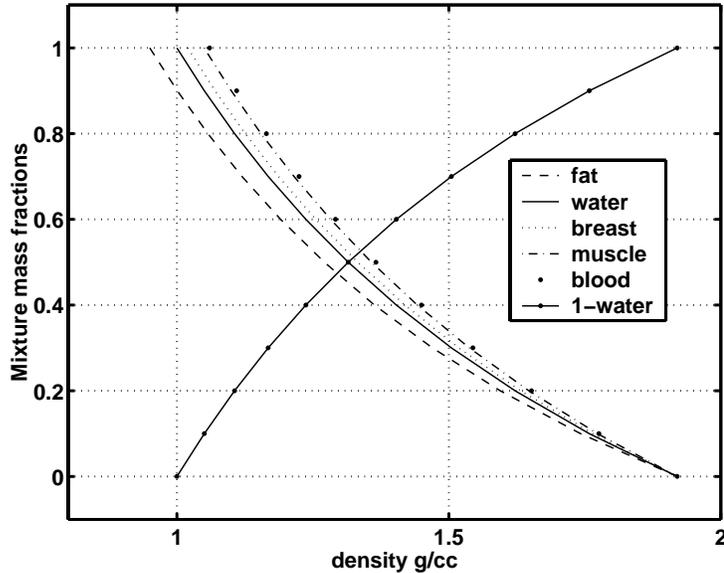
Figure 5.3: Partial volume soft tissue (mass) fractions of mixtures of the different soft tissues with cortical bone. The bone fraction is equal to the soft tissue fraction minus 1.

densities of pixels on tissue boundaries, inherently limited by system resolution, was a secondary consideration. In situations where higher mixture accuracy is desired, the tissue fraction functions can be redesigned to match the mass fractions over a limited density range.

We call this model the *displacement model*, because it is designed for situations where materials occupy distinct spatial regions, with mixed pixels arising mostly at the boundaries. This model is appropriate for regions of the body with dense and thick skeletal bones. Cortical bone is only about 12% water (by mass) and it contains significant concentrations high Z materials (Z$\geq$8) (22.5% Ca, 10.3 % P, by mass) [77] . We therefore treat it as a distinct tissue type. This is possible because its density value is also far enough from that of water. Other kinds of bones, on the other hand, are better described as mineral solutions in water, with mineral composition of a few percents [77].
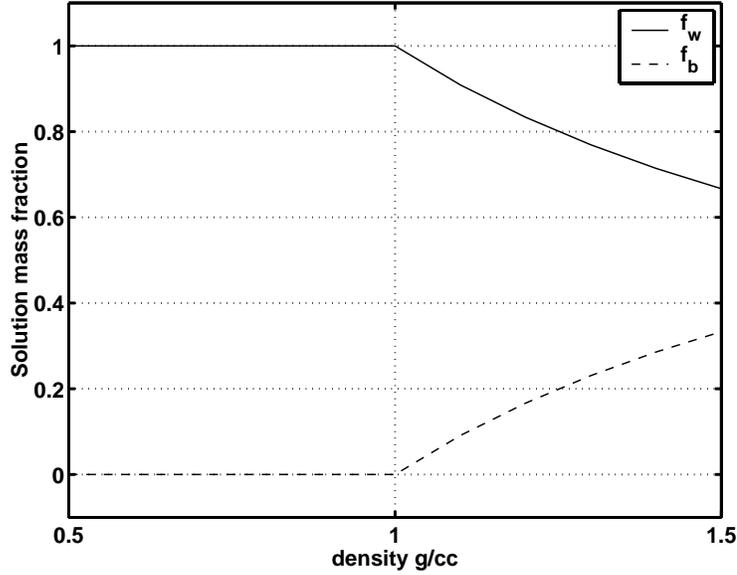
Figure 5.4: Solution model tissue fraction functions $f_w(\rho)$, $f_b(\rho)$. The fractions in this model represent the mass fractions of water and mineral as explained in (5.13).

5.3.2 Solution Model

We also consider another approach that we call the *solution model*, where we assume the object contains a mineral solution in water (or some other solvent). In this model, the water-mineral solution density is greater than or equal to the density of water, and the attenuation coefficient can be modeled as follows:

$$\mu(\mathcal{E}) \approx m_w(\mathcal{E})\rho_w + m_b(\mathcal{E})(\rho - \rho_w) = \rho\left(m_w(\mathcal{E})\frac{\rho_w}{\rho} + m_b(\mathcal{E})(1 - \frac{\rho_w}{\rho})\right), \quad (5.11)$$

for $\rho > \rho_w$. The fraction functions for this model are shown in Fig. 5.4 and are given by:

$$f_w(\rho) = \begin{cases} 1 & \text{if } \rho \le \rho_w \\ \frac{\rho_w}{\rho} & \text{if } \rho > \rho_w \end{cases} \quad (5.12)$$

$$f_b(\rho) = \begin{cases} 0 & \text{if } \rho \le \rho_w \\ 1 - \frac{\rho_w}{\rho} & \text{if } \rho > \rho_w \end{cases} = 1 - f_w(\rho) \quad (5.13)$$
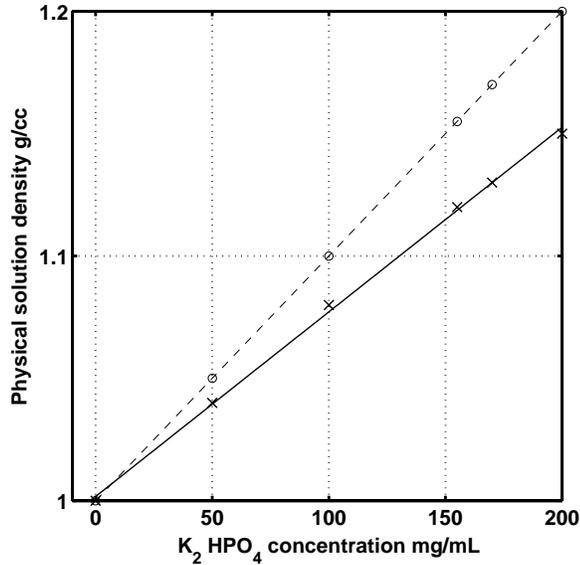
Figure 5.5: Physical density of potassium phosphate solution as a function of the solution concentration. The displacement effect causes the density to depend nonlinearly on concentration. One cubic centimeter of solution will have less than 1 gram of water. The solid line is a second-order polynomial fit to data points (x) reported in [11, 64]. Without post correction, the solution model assumes the density depends on concentration according to the linear dash-circle line (- - -o- -).

where ρ_w is the density of water (or, more generally, the solvent). This model effectively assumes that the object consists of water everywhere, whose density (in some regions) has increased because of mineral content. The fraction functions in this model are the mass fractions of water and mineral in the solution. This model is designed for cases when bone is not very dense, and is more like a water-calcium solution than solid bone mineral. One advantage of this approach is that it makes no assumptions regarding the density of the mineral.

The solution model ignores what is known to chemists as the displacement effect[2] which arises because the amount of water per cc varies nonlinearly with mineral concentration [11,64]. In other words, the physical density of a 100 mg/mL solution will

---

[2]Not to be confused with the *displacement model* discussed in the previous section.

not be exactly equal to 1.1 g/cc. This effect is small at low concentrations. Although it is possible to reparameterize (5.11) to account for this effect, we do not incorporate it at this point. We can correct for the error introduced by ignoring the effect of displacement in solution by applying a simple polynomial correction, based on empirical measurements of density versus concentration, to an image reconstructed with the proposed solution model.

We developed the solution model in hindsight as we tried to apply our algorithm to reconstruct real data generated from a phantom consisting of potassium phosphate solutions of different concentrations. This approach gave better quantitative results than the displacement model (see Section 5.5 below).

5.3.3   Problem Statement

After choosing one of the models for the fractions $f_k^j(\rho)$, the reconstruction problem becomes one of estimating the density vector ρ from the measurements, using the model (5.8). We use a likelihood-based estimation approach. The negative log-likelihood for independent Poisson measurements is given by:

$$-L(\rho) = \sum_{i=1}^{N} E[Y_i|\rho] - Y_i \log(E[Y_i|\rho]) \tag{5.14}$$

where $E[Y_i|\rho]$ is given by (5.6). When the system matrix $\mathbf{A}$ has full column rank and the data is noise-free, minimizing the negative likelihood would give a perfect result. In reality, the data is noisy and maximum likelihood (ML) will give a very noisy reconstruction due to the ill-posedness of the problem, hence the need for regularization. We regularize using the convex edge-preserving Huber penalty given in (4.14).

Combining the likelihood and penalty gives a penalized-likelihood (PL) cost func-

tion:

$$\Phi(\rho) = -L(\rho) + \beta R(\rho) \tag{5.15}$$

where β is a scalar that controls the tradeoff between the data-fit and the penalty terms. The goal of the reconstruction technique becomes to minimize (5.15) subject to certain object constraints such as non-negativity:

$$\hat{\rho} = \underset{\rho \geq 0}{\operatorname{argmin}} \, \Phi(\rho). \tag{5.16}$$

The next section describes an iterative technique for solving (5.16) approximately.

5.3.4 Polyenergetic X-ray CT Iterative Algorithm

The goal of the iterative algorithm is to find the minimizer in (5.16). The optimization transfer principle [15, 16, 22, 25] has proven to be a very effective tool in simplifying such problems. Stated simply, optimization transfer enables us to replace a complex likelihood like (5.15) with surrogate cost functions that are simpler and easier to minimize. Optimization transfer methods can ensure that the cost function decreases monotonically each iteration. Because the Poisson likelihood includes a scatter term (r_i) and because of the nonlinearity in $F_i(s)$, the likelihood is not convex. Monotonicity is all that can be claimed about the convergence of the algorithm.

Along the lines of our earlier work [22], we apply successive surrogates to the penalized likelihood cost function, ultimately resulting in a cost function that is quadratic and separable in ρ. Quadratic functions are more easily minimized, and separability allows the algorithm to update all pixels simultaneously.

We omit the details and refer the reader to [22] and to Appendix A. The resulting algorithm is a diagonally-preconditioned gradient descent method of the following

form:

$$\rho^{n+1} = \left[\rho^n - \mathbf{D}^{-1}\nabla\Phi(\rho^n)\right]_+, \tag{5.17}$$

where $\nabla\Phi(\rho^n)$ is the gradient of the cost function, $\mathbf{D}$ is a diagonal matrix that influences the rate of convergence and and monotonicity of the algorithm, and $[\cdot]_+$ enforces the nonnegativity constraint. For a monotonic algorithm, the elements of $\mathbf{D}$ must be such that a surrogate satisfies the conditions of optimization transfer. It is possible to derive a version of $\mathbf{D}$ that gives a monotone algorithm [25, 32, 82] but the elements of the matrix will have to be updated at every iteration, resulting in a more computationally expensive approach. Since monotonicity will be compromised anyway by using ordered subsets to accelerate the algorithm, a precomputed approximate $\mathbf{D}$ will suffice. By making certain assumptions about the object [22], we derive an approximate precomputed "curvature" matrix using the second derivative of the quadratic surrogate evaluated at the initial image:

$$d_j = m_w^2(\mathcal{E}_{\text{eff}}) \sum_{i=1}^{N} a_{ij}^2 Y_i,$$

where d_j is the jth entry of the diagonal matrix $\mathbf{D}$ and $\mathcal{E}_{\text{eff}}$ is the effective energy of the X-ray spectrum defined as:

$$\mathcal{E}_{\text{eff}} \triangleq \frac{\displaystyle\int \mathcal{E} I_i(\mathcal{E}) d\mathcal{E}}{\displaystyle\int I_i(\mathcal{E}) d\mathcal{E}}. \tag{5.18}$$

See Appendix A for details. With this approximation, the method is not guaranteed to be monotonic. One could check the cost function at each iteration and apply the monotonic update to those rare cases where the cost function does not decrease using the approximate method. In our experience, with a good starting image, such as a FBP reconstruction, the cost function always decreases. As far as the computational load of the algorithm, each iteration involves one forward projection and one

backprojection. An iteration is therefore roughly equivalent to two runs of FBP.

5.4 Simulation Results

We assessed the effectiveness of our algorithm with simulated transmission polyenergetic Poisson X-ray projections. We simulated a parallel beam geometry, but the algorithm applies equally well to other geometries. We used a 100 kVp spectrum (incident mean 45 keV and standard deviation 14.8 keV) using the code available from [7] with 10 mm Aluminum filtration at the source to give a spectrum shape similar to clinical practice.

Fig. 5.6 shows the phantom used. The phantom is based on on the NCAT phantom [72], where in addition to (water-like) soft tissue and bone, we have manually introduced regions of breast, fat (adipose tissue), blood and soft-tissue lesions with contrast differences of -4% to 6%. The bones and soft tissue have densities $1.8 - 2.0$ g/cc and $0.2 - 1.0$ g/cc respectively. The densities for fat, bone and blood are identical to those cited in table 5.1. The measurements are simulated using the actual attenuation properties of the different tissues, not just those of water and bone. The sinogram has 1200 radial bins with 0.055 mm ray spacing and 1200 angular steps over 180°. We did not simulate scatter. The blank scan value was 4.87×10^6 counts/detector. The ray with the lowest SNR had 77 counts. We generated pseudo-random Poisson variates for each sinogram element. The original phantom was 1024×1024 (0.055 mm / pixel) but we reconstructed the data on a 512×512 grid (0.11 mm/pixel) to generate the effect of mixed pixels.

The data was reconstructed by uncorrected FBP, JS-corrected FBP, the polyenergetic statistical algorithm discussed in Chapter IV with a pre-segmented JS image and an idealized oracle segmentation, and the proposed polyenergetic algorithm. We
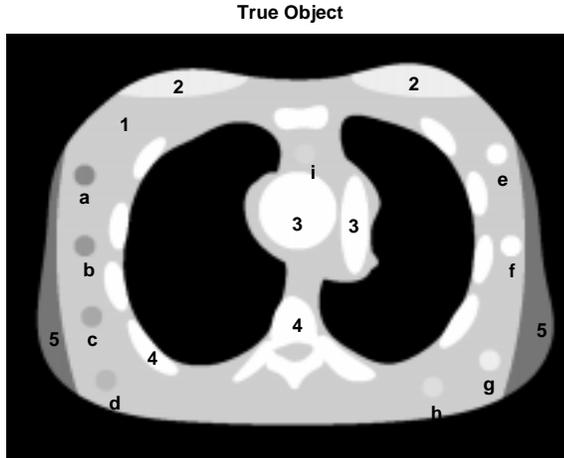
Figure 5.6: Simulation experiment. True 1024 x 1024 object containing soft tissue (1), breast (2), blood (3), bone (4) and fat (5). The water lesions are (a) -4%, (b) -3%, (c) -2%, (d) -1%, (e) 6%, (f) 3%, (g) 2%, (h) 1%, (i) 0.5% different from 1 g/cc. Window level = -45 HU and window width = 150 HU.

used a Hanning window in all FBP reconstructions, and post-processed the FBP images with a 5 x 5 median filter. We generated the pre-segmented JS image for the statistical algorithm of the previous chapter using a pixel value threshold of 1.5. The JS correction also used the idealized oracle segmentation. The oracle segmentation combined every 4 pixels of the true object to either one bone pixel or one water pixel. The iterative algorithm in [22] accommodates only water and bone, and requires prior knowledge of their distribution in the image. With oracle segmentation, we assume this algorithm has 'perfect' prior knowledge. This is not a practical case, but is included for the sake of comparison to provide an upper bound on performance.

For the iterative methods, we included edge-preserving shift-variant regularization [35]. The iterative algorithms ran until the absolute normalized maximum difference between successive iterations fell below a prespecified small value. To achieve a maximum difference of less than 2×10^{-3}, we used 21 iterations with 50 subsets. Fig. 5.8 illustrates how the maximum difference decreases with iteration for the pro-

Uncorrected FBP

Oracle JS FBP
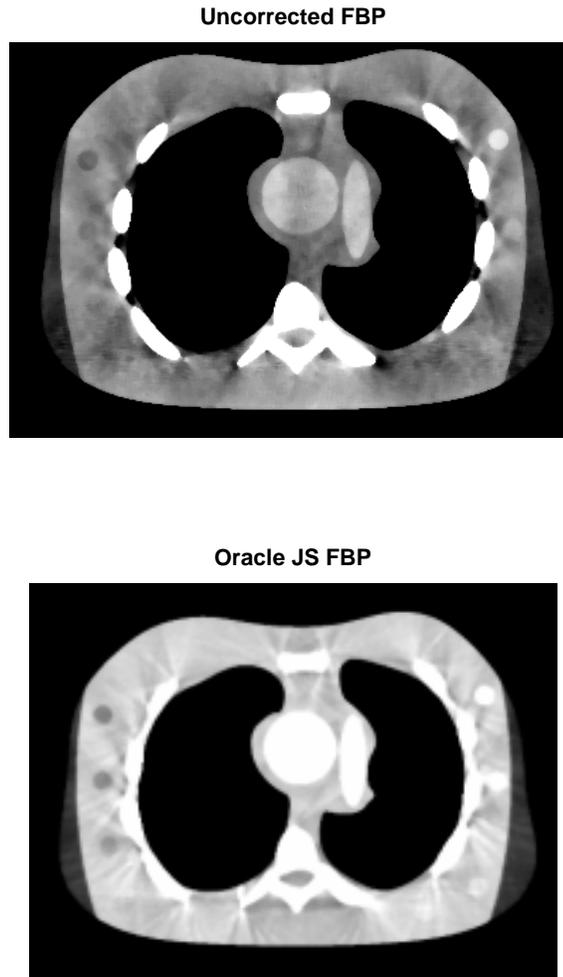
Figure 5.7: Simulation experiment. Top: uncorrected FBP (scaled by ratio of water density to attenuation coefficient). Bottom: JS correction post processed with a 5 x 5 median filter. Window level = -45 HU and window width = 150 HU.
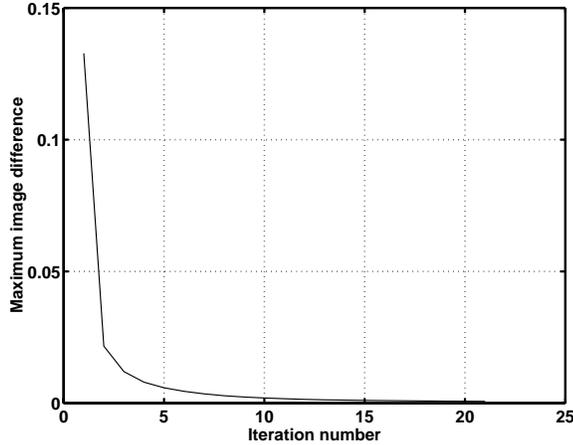
Figure 5.8: Absolute maximum difference between successive iterations (normalized by the maximum pixel of the current iteration) of the proposed algorithm with simulated data of the phantom in Fig. 5.6.

posed algorithm.

5.4.1 Image Reconstruction Results

Since algorithms with beam hardening correction usually reconstruct density values, we display the reconstructed images using a *density* Hounsfield unit (HU) defined as:

$$HU = \frac{\rho - \rho_{\text{water}}}{\rho_{\text{water}}} \times 1000.$$

If one wishes to display attenuation Hounsfield units, the density images can be segmented and each material scaled by its mass attenuation coefficient at some energy.

Fig. 5.7 shows the FBP image reconstruction results and Fig. 5.9 shows the iterative reconstruction results. The uncorrected FBP image suffered from severe beam hardening artifacts. The statistical iterative reconstructions exhibit significantly better artifact reduction than the Joseph and Spital image because they inherently model the nonlinearity and reconstruct the data accordingly. The segmentation-free algorithm reconstruction is comparable to the oracle pre-segmented iterative reconstruc-

Pre−segmented Iterative Image
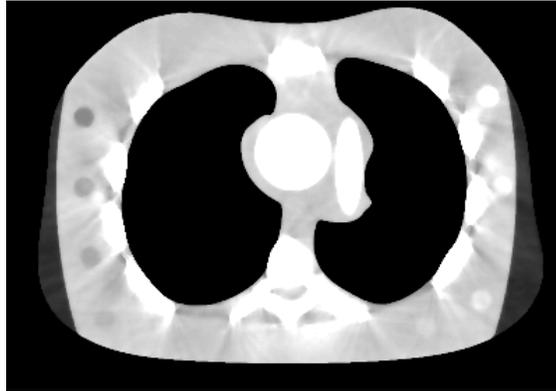
Oracle−segmented Iterative Image
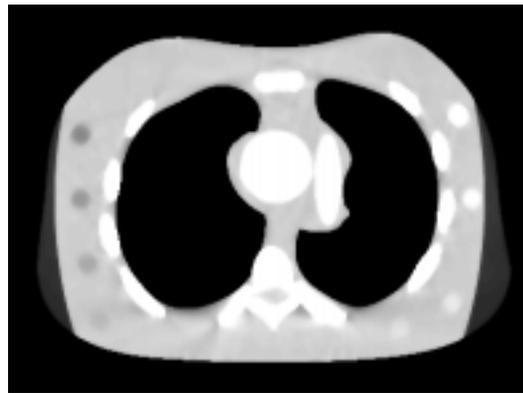
Segmentation−free Iterative Image
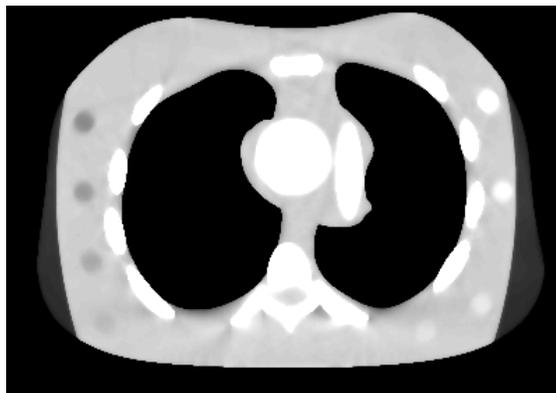
Figure 5.9: Simulation experiment reconstruction results. Top: pre-segmented iterative algorithm. Middle: Oracle pre-segmented iterative reconstruction. Bottom: proposed statistical reconstruction algorithm. Window level = -45 HU and window width = 150 HU.

tion. The JS-presegmented iterative image suffers from significant artifacts when compared to the oracle or segmentation-free results. Although it may be possible to reduce the artifacts by adjusting the segmentation threshold, this is a subjective and error-prone process that is unlikely to eliminate the artifacts completely. This points out the limitation of the pre-segmentation based statistical algorithm. None of the reconstruction algorithms produced an image where breast tissue could be visually discerned from the surrounding water.

In addition to qualitative assessment of artifact reduction, we quantify the performance of the different algorithms using the percent normalized root mean squared error (NRMSE) in uniform regions within each tissue type. The uniform regions were selected far from edges, and contained between 100 and 300 pixels. It is difficult to quantify the performance of the algorithms at tissue boundaries because of the intentional mismatch between image-domain grids used for data generation and reconstruction.

Table 5.2 shows the NRMSE results. The proposed iterative method quantitatively outperforms JS-corrected FBP with oracle segmentation for all tissues except fat and breast. It also has better overall NRMSE performance than the pre-segmented iterative reconstruction, and is comparable in performance to the idealized oracle pre-segmented iterative algorithm image.

Table 5.2 suggests that the proposed algorithm gives more accurate results than JS FBP and the pre-segmented iterative algorithm for all tissues, except fat and breast tissue. We observed that fat (adipose tissue) and breast are generally underestimated by the algorithms. Fat is less dense and less attenuating than water. Since the proposed algorithm uses the spectral attenuation properties of water, it is reasonable for the estimated fat density to be less than the true value, to compensate for

Reconstruction Method	Water RMS, σ	Fat RMS, σ	Blood RMS, σ	Breast RMS, σ	Bone RMS, σ
Oracle JS FBP	0.6, 0.5	3.8, 0.5	1.1, 0.7	1.6, 0.2	7.4, 0.3
Pre-segmented Iterative	0.5, 0.5	4.1, 0.3	0.3, 0.3	2.3, 0.1	0.4, 0.2
Segmentation-free Iterative	0.16, 0.15	4.1, 0.13	0.4, 0.1	2.1, 0.1	0.2, 0.2
Oracle-segmented Iterative	0.15, 0.1	4.2, 0.17	0.4, 0.1	2.1, 0.14	0.2, 0.2

Table 5.2: Percent normalized root mean squared error and percent standard deviation of the different reconstructions of the altered NCAT phantom in Fig. 5.6.
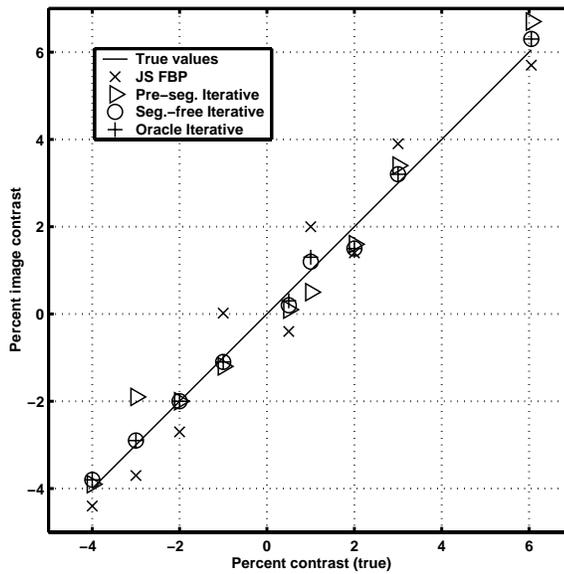
Figure 5.10: Percent contrast of the soft tissue lesions in the phantom of Fig. 5.6 for the different reconstruction algorithms.

the higher attenuation introduced by the water model. A similar argument follows for breast tissue. In the case of breast tissue, however, tissue density is very close to that of water, and breast tissue is only slightly less attenuating than water. This explains the difficulty in producing images where breast is visually distinguishable from water. At the effective energy of the 100 kVp spectrum used, the difference in attenuation coefficient between water and breast is less than 1%, corresponding to an attenuation Hounsfield unit difference of 6 HU.

Fig. 5.10 illustrates the performance of the different algorithms in terms of low contrast detectibility. The phantom has 9 soft tissue lesions with density HU values of -40, -30, -20, -10, 5, 10, 20, 30, 60. The 5 HU lesion, corresponding to a 0.5% difference from the actual density of water, could not be discerned from any of the reconstructions. The other lesions were visible in both the JS FBP and statistical reconstructions using a narrow display window. Fig. 5.10 shows that the iterative reconstruction images are generally closer to the true contrast levels than the JS FBP image.

5.4.2   Robustness to Spectrum Model Mismatch

Measuring the X-ray source spectrum is a challenging problem that requires careful measurements and calibrations. Measuring the spectrum directly is difficult, and indirect methods are often used. One possibility is to attempt to estimate the function F_i in (5.7) by measuring phantoms of different thicknesses and using polynomial fits on the data. Another approach is to solve for the spectrum from a discretized version of the transmission equation using measurements taken on phantoms with varying thicknesses [66, 79]. Regardless, attempts to measure the scanner spectrum are likely to give approximate results. Moreover, the kVp setting on X-ray scanners may be imperfect, and the spectrum may change with tube drift over time.

It is therefore important to investigate the robustness of our reconstruction method to mismatches between the spectrum that generates data and the assumed spectrum used to tabulate F_i and its gradient.

We reconstructed 100 kVp data generated from a phantom similar to the one described in the previous section, but without the soft-tissue lesions. In addition to the matched 100 kVp algorithm, we used algorithms based on 80, 90, 110, and 120

kVp spectra. The (incident) mean energies of these spectra deviate by 11.2%, 5.2%, 4.7% and 9.0% respectively, from that of the 100 kVp spectrum. This experiment reflects a practical situation where the kVp setting is inaccurate, because of tube drift or other effects. Fig. 5.11 shows the spectra used.
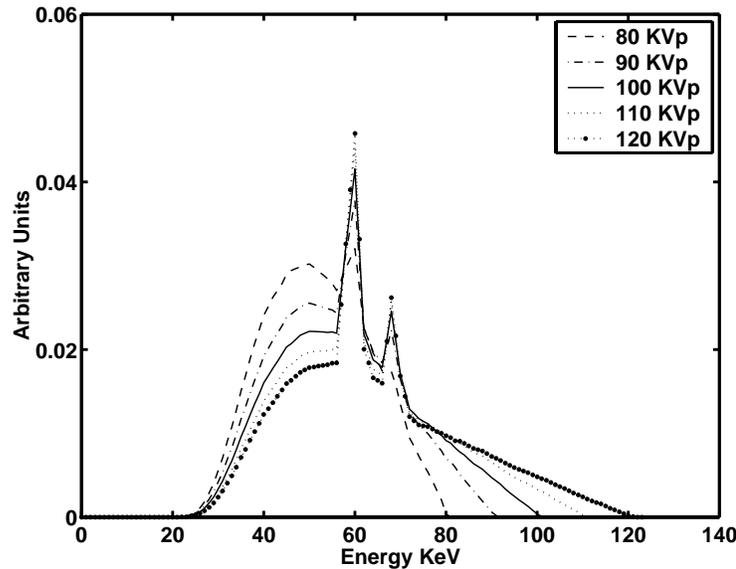
Figure 5.11: Energy spectra used to examine effect of spectral mismatch in section 5.4.2.

The images shown in Fig. 5.12 were reconstructed using the algorithm parameters listed in the previous section. The only difference between the different reconstructions was the spectrum used to tabulate F_i. The display window is the same for all the images. Spectral mismatch resulted in images exhibiting different 'brightness' levels, and inaccurate pixel density values, as further illustrated by Fig. 5.13. Algorithms with lower kVp settings underestimated density values and vice versa. The error introduced in estimating tissue densities by spectral deviations is illustrated in Fig. 5.14.

The error introduced by spectral mismatch is generally higher for higher density bones. This is reasonable since dense bones are highly attenuating and the devi-

(a) 80 kVp

(b) 90 kVp
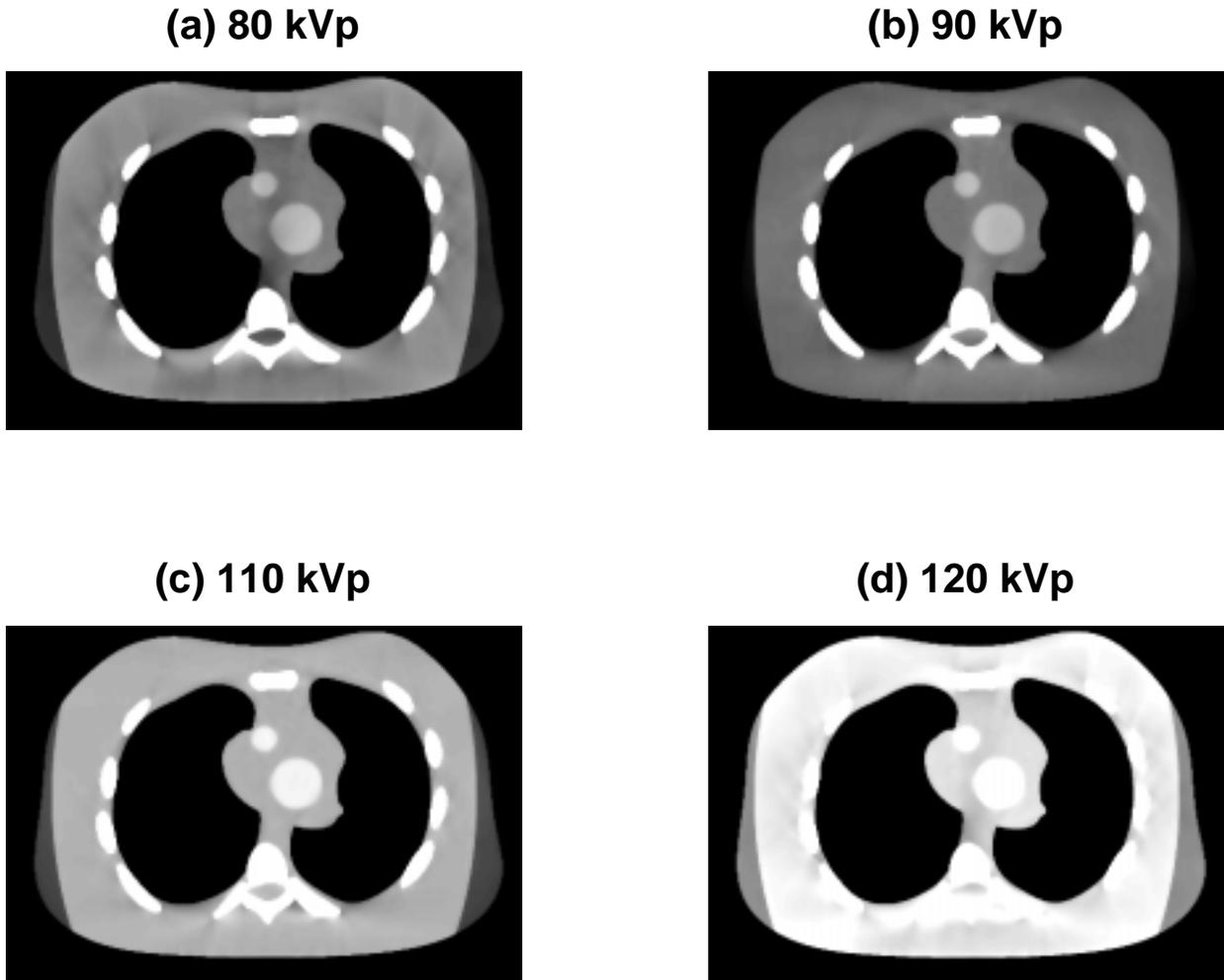
(c) 110 kVp

(d) 120 kVp

Figure 5.12: Spectrum mismatch results. 100 kVp data reconstructed with (a) 80 kVp algorithm; (b) 90 kVp algorithm; (c) 110 kVp algorithm; (d) 120 kVp algorithm. Window level=-25 HU. Window width=250 HU.
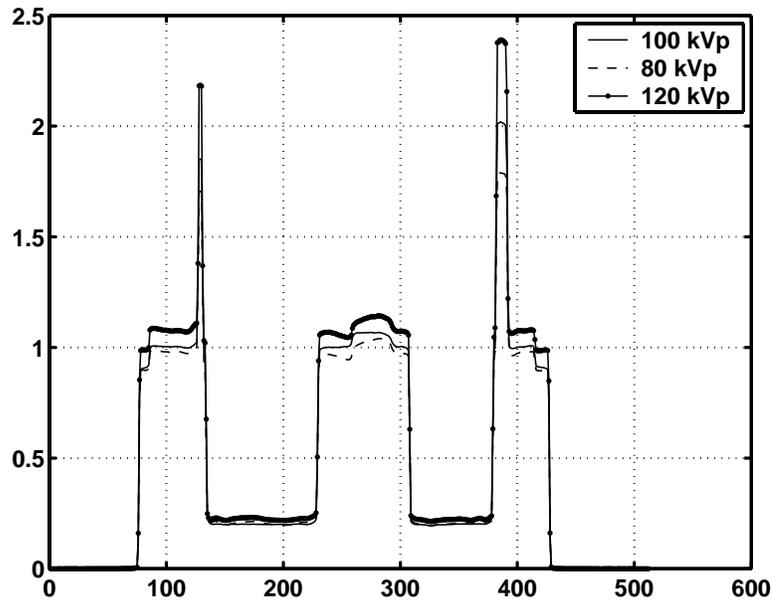
Figure 5.13: Spectrum mismatch results. Profile plots of one row of the reconstructed images. The plots show that a lower kVp mismatch causes the algorithm to underestimate tissue density. Higher kVp mismatch has the opposite effect.
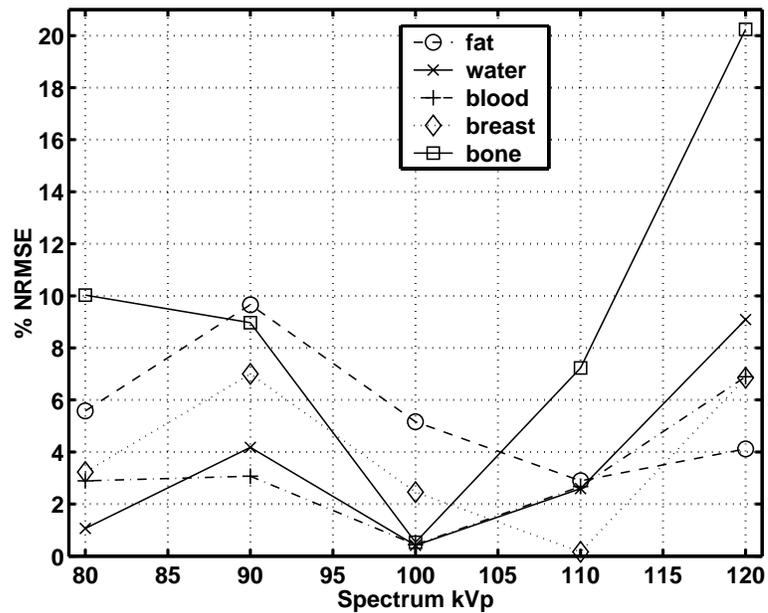
Figure 5.14: Normalized root mean squared error for different tissues of the phantom caused by different kVp settings.

ation will be increasingly nonlinear. This can also be seen from the scatter plots of Fig. 5.15. The solid lines in Fig. 5.15 are the identity lines for density values of the image reconstructed using the correct spectrum. The dotted scatter plots illustrate the deviation of the images reconstructed with mismatched spectra from the matched spectrum case. As the figure illustrates, the deviation is non-linear and increases with object density, and results in overestimation and underestimation of density values for higher and lower kVp settings, respectively.

Fig. 5.14 also shows an interesting pattern for fat and breast tissues. Their error patterns are very similar and they have minimum error at 110 kVp. Fat and breast tissues have similar attenuation properties (see Fig. 5.1), and are less attenuating than water. Since the algorithm with exact spectrum underestimates fat, it is reasonable that a higher kVp setting that tends to overestimate tissue densities will give less error. Note also that blood and water, which have very similar attenuation properties, have similar error patterns.

The acceptable degree of deviation between the actual and assumed spectra will depend on the application. We note that the density errrors are closely related to the percent errors of the tube potentials. Clearly, in quantitative applications, as close a match as possible is desirable. If the application requires only visual and qualitative assessments, then a mismatch of 20 kV in the kVp scanner setting may be tolerable.

5.5 Real Data Results

5.5.1 Mineral Density Measurement

To validate our method with real data, we obtained real transmission data acquired on the table-top system described in [70]. The phantom consisted of a water cylinder (15 cm in diameter) with five embedded smaller cylinders (2.5 cm diameter)
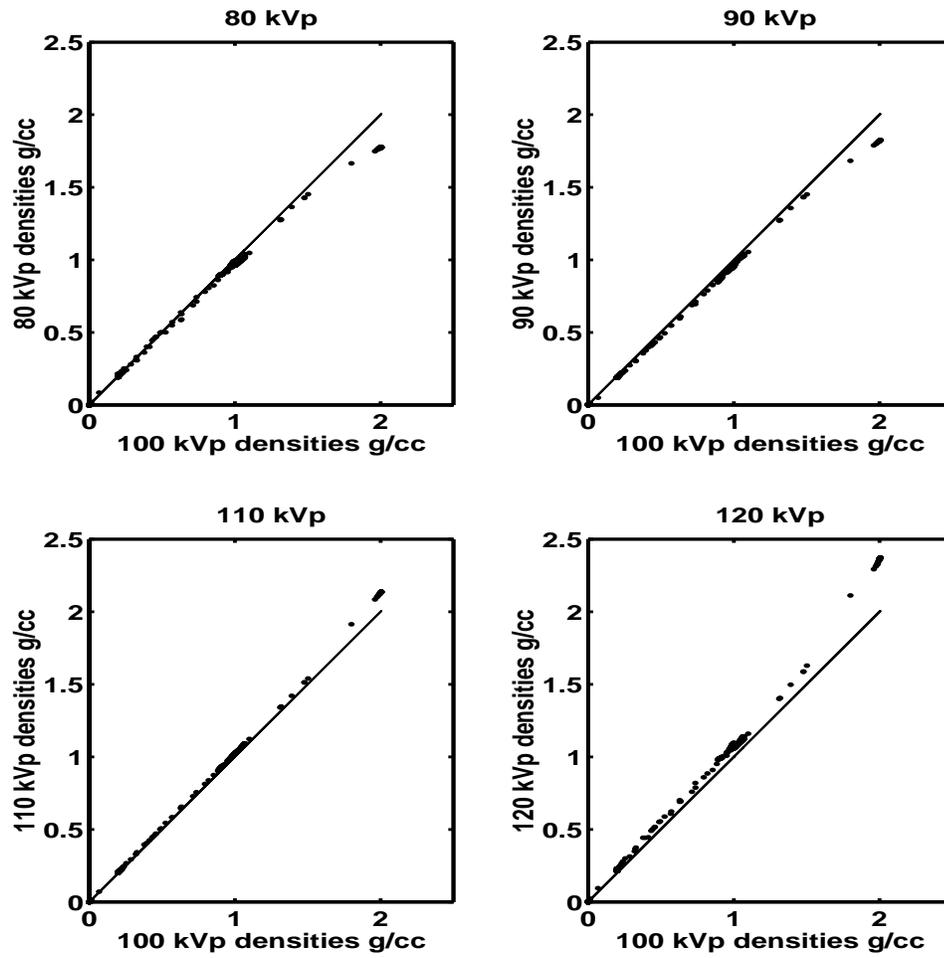
Figure 5.15: Spectrum mismatch scatter plot. The solid line is the identity line of the density values of the image reconstructed with a perfectly matched spectrum. The dotted lines are scatter plots that illustrate the deviations caused by spectrum mismatch.

filled with different concentrations of (dry) potassium phosphate dibasic (K_2HPO_4) solution. The powder was mixed with water to form the following solutions: 50, 100, 141, 150 an 200 mg/mL, corresponding to densities[3] of 1.04, 1.08, 1.108, 1.115 and 1.153 g/cc respectively. This powder is quite soluble in water, so we used the solution model (5.12) and (5.13) for image reconstruction. This phantom was intended to mimic a range of trabecular bone densities.

Fan-beam data (834 angular view, 576 detector bins) were collected at 60kVp using a 360° scan with axial collimation and no scatter grid. The beam was filtered with 1 mm Aluminum and 0.5 mm Copper and the spectrum estimated from Monte Carlo simulations. The data was reconstructed on 256×256 grids using the proposed polyenergetic statistical algorithms (10 iterations, 36 subsets) with both the displacement and solution models, as well as water-corrected FBP. We applied a polynomial correction to the solution model image to account for the displacement effect in solution. The polynomial coefficients were computed to fit density and concentration data reported for K_2HPO_4 in [11, 64].

5.5.2 Scatter Estimation

Initial reconstruction suffered from streaking that was typical of the kind caused by scatter, but no scatter estimate was available with the data. To estimate scatter, we computed a difference sinogram between the real sinogram and a synthetic sinogram corresponding to a numerical water phantom based on a segmentation of an initial reconstruction. We excluded sinogram rays passing through the potassium phosphate regions from the difference sinogram and replaced them with interpolated values. We applied smoothing to the difference sinogram to estimate scatter. Fig. 5.17 shows the estimated scatter sinogram. The measured transmission sino-

[3]These densities were not measured directly, but were calculated based on data in [11, 64].
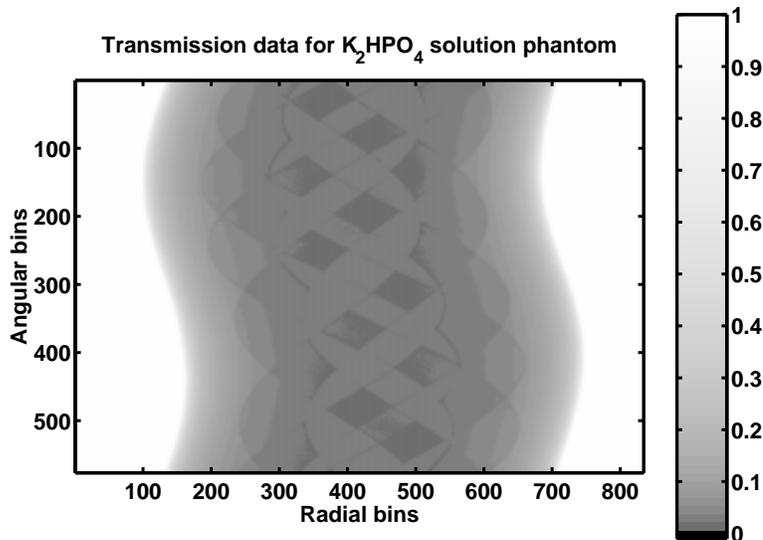
Figure 5.16: Measured sinograms. Note that the display range has been warped to better illustrate details.

gram is shown in Fig. 5.16 with a narrow display window to illustrate fine details. We estimate scatter at about 1% of the overall measured signal.

The shape and value of the estimated scatter sinogram are similar to those in earlier reports [37, 51]. This is not a fully developed scatter estimation/correction technique. Rather, it is a heuristic approach to help verify that scatter is the likely source of the observed artifact and to correct for it in the absence of any information on the scatter properties of the system.

Results and Comparisons

Fig. 5.18 shows the polyenergetic iterative reconstructed image. We do not show the FBP reconstruction because this phantom has no high density regions that cause severe beam hardening streaking, and the images are visually comparable. We do, however, compare the quantitative accuracy of our approach with water-corrected FBP in determining the density of the different K_2HPO_4 concentrations in Fig. 5.19. We also compare the accuracy of the displacement and solution object models in
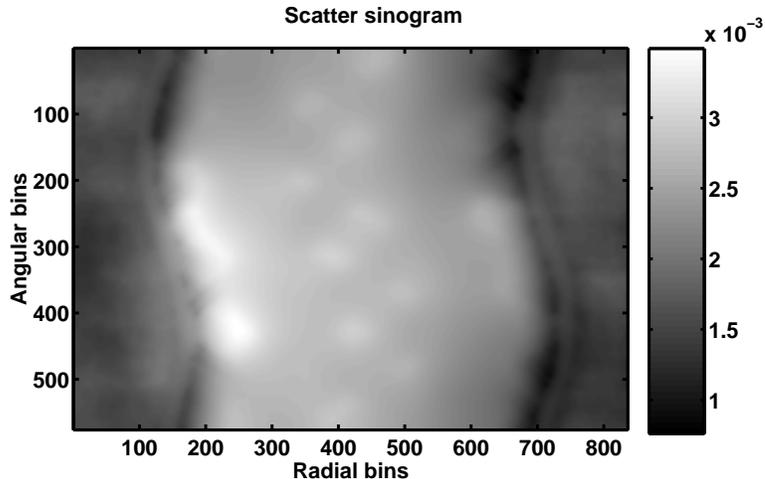
Figure 5.17: Estimated scatter sinogram obtained by smoothing the difference between an idealized sinogram and the real sinogram.

Fig. 5.19. The apparent roughness of the edges of the phantom image in Fig. 5.18 is due to the algorithm attempting to reconstruct the walls of the plastic container in the phantom.

Although both the displacement and solution models give good visual results, the solution model has significantly higher quantitative accuracy. The solution model polyenergetic statistical reconstruction has the best accuracy of all methods, with maximum error of -1.2% (2.7% error before post-correction for water displacement). The values in the Figure 5.19 are the averages of uniform regions inside each cylinder. Trying to address the apparent bias present in the potassium phosphate cylinders in the displacement model reconstruction was the motivation behind deriving the solution model.

5.5.3 MicroCT Reconstruction

MicroCT scanners, which offer micron-level resolution, are being promoted and used to image small animals for research and drug development purposes. We ac-
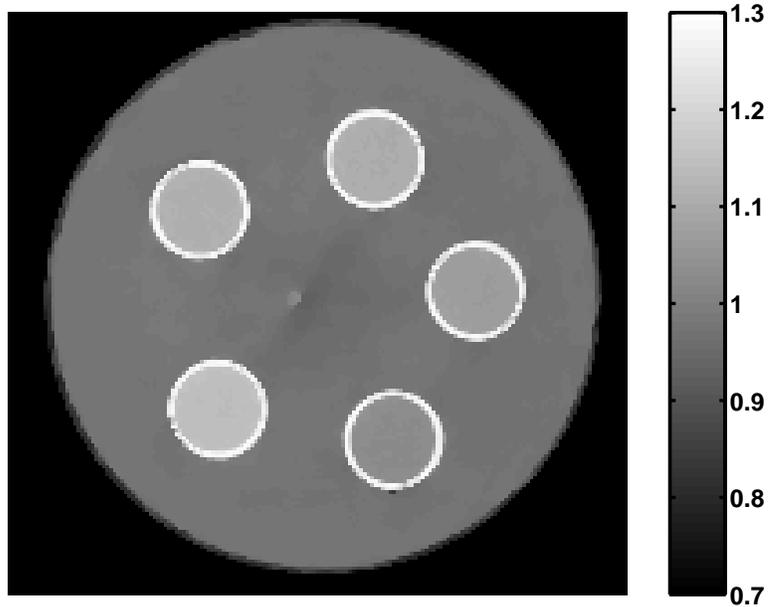
Figure 5.18: Iterative reconstruction of real data using 10 iterations and 26 subsets with an edge-preserving penalty.

quired cone beam data on a microCT system[4] (EVS Corp., now part of General Electric Medical Systems) that is routinely used for bone density measurements on small animals. We extracted and reconstructed fan-beam data. The scanner tube emitted 50 kVp X-rays and the detector collected 278 views over a range of 180°. There were 875 detector channels at a resolution of 25μm each. The scanner did not have a scatter rejection grid.

The scanned phantom was a mouse-sized cylinder containing smaller cylinders of water, air, fat mimic and bone mimic. We are uncertain as to the exact density values of the bone and fat mimic. Since the phantom is used for calibration purposes for bone density measurements, we reconstructed it using the solution model. We ran the algorithm for 10 iterations with 20 subsets with an edge-preserving penalty function. We compared the statistical reconstruction to the FBP algorithm available

---

[4]Courtesy of Pfizer Research and Development BioImaging Center, Ann Arbor, MI.
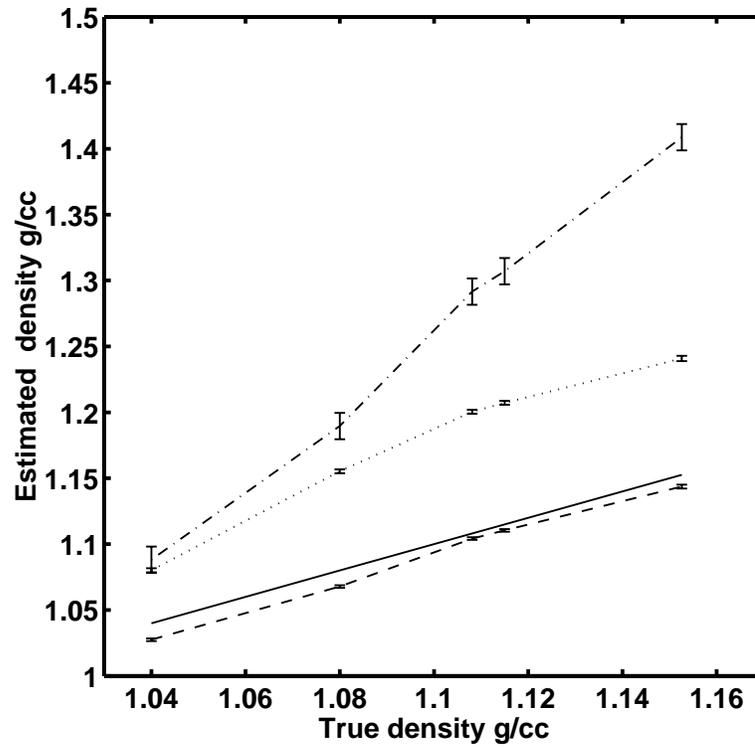
Figure 5.19: Accuracy of the different reconstruction techniques. Figure key: true values (—-); solution model (‗ ‗ ‗); displacement model (...); water-corrected FBP (‗ . ‗ .). The solution model reconstruction values are very close to the true values (max. error -1.2%). The displacement model has errors on the order of 8%. The solution model image was post-corrected for the displacement effect. The error bars represent the width of one standard deviation of each measurement.

on the scanner. As Fig. 5.20 illustrates, the FBP reconstruction suffers from beam hardening artifacts that are mostly absent in the statistical reconstruction. Ring artifacts appear in the iterative image. Close examination reveals similar artifacts in the FBP image, obscured by the higher noise level. These artifacts are most likely caused by some detector effect that is not accounted for in the reconstruction. In the absence of accurate information about the densities of the bone and fat mimics, we compared the density of the water insert to the known value of 1.0 g/cc, and found the statistical algorithm to be within 3%.

5.6   Conclusions

We have introduced a statistical iterative reconstruction algorithm for energy dependent X-ray attenuation that produces images with significantly reduced beam hardening artifacts, and we have demonstrated its effectiveness for bone and soft tissue objects. Unlike most other transmission CT iterative algorithms, our algorithm is based on a realistic polyenergetic model and has been extended to allow for mixed pixels and mineral solutions. The algorithm we present requires knowledge of the incident spectrum or tabulated measurements of the line integrals of two base substances. The statistical algorithm minimizes a separable paraboloidal surrogate to the Poisson likelihood and is parallelizable and fairly simple to implement. We use ordered subsets and pre-computed surrogate curvatures to accelerate convergence and reduce computation. It is possible to derive curvatures such that the algorithm monotonically decreases the cost function with one subset [25, 32, 82]. This is about the most that can be said about convergence since the cost function is not convex.

When compared with the post-processing technique of Joseph and Spital [49], the statistical algorithm yielded fewer artifacts. The JS method estimates the line-

microCT data FBP reconstruction
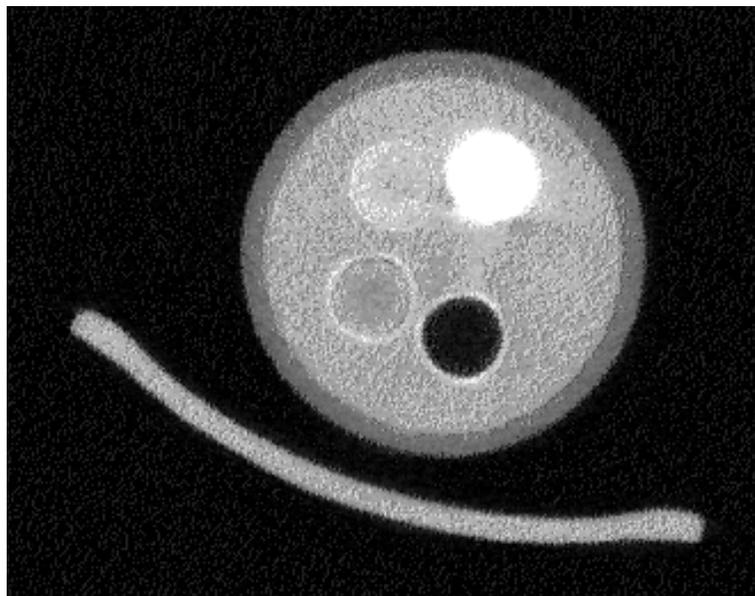
microCT data iterative reconstruction
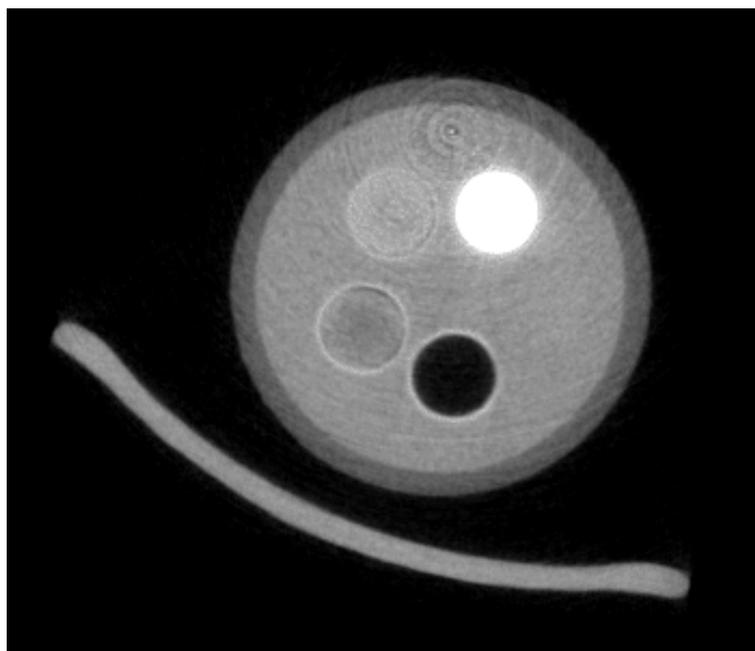
Figure 5.20: Phantom reconstruction from microCT data. Top: FBP reconstruction. Bottom: Statistical reconstruction. Window level=-250 HU and window width=1500 HU.

integral dependent nonlinearity for each sinogram bin and then recalculates the line integrals. The statistical method needs no such post-processing since it inherently accounts for the nonlinearities. This is likely the reason for its higher accuracy. Displacement model reconstruction gives images with reduced artifacts in anatomical objects. The solution model is promising for quantitative applications where estimating the density of mineral solutions is needed. One possible clinical application of the solution model is in bone mineral density studies. Many structures within bone can be characterized as mineral solutions and quantitative estimates of their densities are useful in understanding the progress of certain diseases.

The displacement model was less accurate in estimating the density of fat and breast tissue. As seen in Fig. 5.1, fat spectral properties deviate most from water, when compared to other soft tissues. This deviation may explain the inaccuracy in estimating fat density. In the case of breast tissue, its density is very close to that of water and this contributes to the difficulty in visually distinguishing it from water in the reconstructed images. The linear attenuation coefficient of breast is also very close to that of water at clinical kVp settings. In applications where fat must be estimated accurately, a modified object model is necessary. One possibility, for example, is to use the mass attenuation coefficient of fat as a base substance instead of water. Another alternative would be to segment the fat regions of an image and model the beam hardening nonlinearity they introduce in the projection. It may also be possible to generalize the tissue fraction functions to include the known spectral properties of different tissues.

The solution model did not incorporate the nonlinear dependence of solution density on mineral concentration. This dependence can be expressed in terms of a low-order polynomial fit to empirical measurements of solution density and con-

centration. This polynomial fit can be used to post-correct the image, or can be incorporated in the design of the solution model fraction functions.

Future work will include applying the polyenergetic approach to objects with three or more tissue types. The bone/soft-tissue model is sufficient for most cases, but a three-class model is necessary when contrast agents such as Iodine are introduced [48] and possibly when metallic implants are present. A three-substance model may also help accurately estimate the density of fat.

Future work will also address some of the challenges posed by the penalized-likelihood approach. Penalized likelihood is attractive because it improves the conditioning of the problem and enables one to choose penalty functions that control desired properties such as edge preservation. One drawback, however, is the absence of an intuitive method for choosing the values of the regularization parameters, which is often done by trial and error. This is inefficient and time consuming, and there is a need for a more systematic method for choosing the parameters. Another undesirable property of penalized-likelihood image reconstruction is its non-uniform spatial resolution [35]. A remedy for quadratic penalties exists [68]. For transmission imaging, we prefer to use the edge-preserving Huber penalty, for which the non-uniform resolution problem has not been addressed.

Recent advances with dual-energy statistical reconstruction [33, 71] have offered another approach for polyenergetic image reconstruction. Dual energy provides an additional set of measurements that enable reconstruction of separate bone and water images. It will be interesting in the future to compare the performance of single-energy and dual-energy statistical reconstruction.

CHAPTER VI

System Model for the GE LightSpeed Scanner

6.1 Introduction

The ultimate test of the usefulness of the techniques proposed in this dissertation is their clinical utility. The use of ordered subsets with transmission X-ray CT image reconstruction algorithms [1, 3, 26] significantly accelerates algorithm convergence to a solution. In addition, the development of statistical algorithms for CT that account for the polyenergetic source spectrum and other physical effects has advanced the potential for practical use of statistical methods in X-ray CT.

This chapter reports on the progress of a collaboration with GE Medical Systems (Milwaukee, WI) to evaluate the potential of statistical image reconstruction methods on real X-ray CT data acquired on the GE LightSpeed CT scanner (GE Medical Systems, Milwaukee, WI). The LightSpeed is a third generation scanner capable of multislice and helical scanning. For purposes of the preliminary phases of this collaboration, only single slice data has been acquired.

The LightSpeed scanner employs a highly optimized version of FBP for image reconstruction. Since raw CT data almost never conform to the assumptions used to derive FBP, data conditioning and pre-correction are necessary. Although it is possible to apply iterative reconstruction algorithms to preprocessed data, such an

approach does not capture the full utility of these algorithms. Data preprocessing alters the statistical properties of the data, and determining the statistical distribution of the processed data is challenging. Moreover, one of the advantages of iterative algorithms is their ability to incorporate the physics and system parameters involved in data formation in the forward CT model to give superior image quality.

An accurate forward model for a CT scanner needs to account for such things as the polyenergetic source spectrum, detector response and source effects, etc., which are considered deviations from ideal conditions in FBP image reconstruction. In statistical reconstruction we do not precorrect data for 'spurious' effects, but rather accommodate them in the forward CT transmission model. In general terms, given a data set, we seek to find a solution (reconstructed image), that best fits the data according to a model that incorporates physical and system effects.

To formulate such a model may require extensive experiments and empirical verification. At the initial stages of this investigation, we are seeking to model the various physical effects that preprocessing addresses by essentially implementing the 'inverse' of the preprocessing steps used by GE Medical Systems. Although this approach is suboptimal, it is a good starting point and can provide an initial assessment of the usefulness of statistical reconstruction. An understanding of the preprocessing steps is therefore important, and we discuss some preprocessing steps in the next section.

6.2   Data Preprocessing Steps

In this section we outline a partial list of data correction steps typically applied to raw CT data. Much of the detailed information about preprocessing is proprietary in nature. We will by necessity keep our discussion general, and restrict it to what is public information. We rely heavily on information provided in [45].

6.2.1 Noise Reduction

Noise in CT measurements arises from two sources: X-ray quantum noise and electronic additive noise. Since the X-ray photons are approximately Poisson distributed, signal variation increases with diminished photon flux reaching the detector. Excessive noise can lead to severe streaking. In fact, it can be shown that after the log operation, signal variance is inversely proportional to photon flux. In addition, the filtering operation in FBP is essentially a derivative operator, further increasing the adverse effects of noise. The backprojection operation maps highly fluctuating noise samples to dark and bright streaks in the image [45]. Fig. 6.15-Fig. 6.17 illustrate the effect of different levels of SNR on image quality.

Noise reduction schemes in the data domain involve nonlinear and adaptive signal processing techniques that adjust noise-reduction filtering depending on the signal level [41].

6.2.2 Offset and Gain Correction

Offset correction addresses the effect of electronic dark current inevitably arising in the detector electronics. Detector electronics are designed to keep dark current to a minimum. Artifacts arise, however, when there are significant channel-to-channel variations in the offset current of the detector. Such variations can lead to ring or band artifacts [45]. To correct for the dark current offset, the dark current is measured right before a scan. The average offset is then subtracted from the projection measurements prior to image reconstruction.

Another important detector effect that is simple to compensate for is the nonuniform gain factor in the detector channels. It is impossible to manufacture identical detector channels, and gain variations can lead to ring or band artifacts. Most CT

scanners use air scans to correct for this effect [45]. Typically, a number of air scans are taken and their average used for gain correction. Projection data are divided by the gain vectors to produce gain-normalized projection set. Fig. 6.1 illustrates transmission data from one view before and after air scan correction.
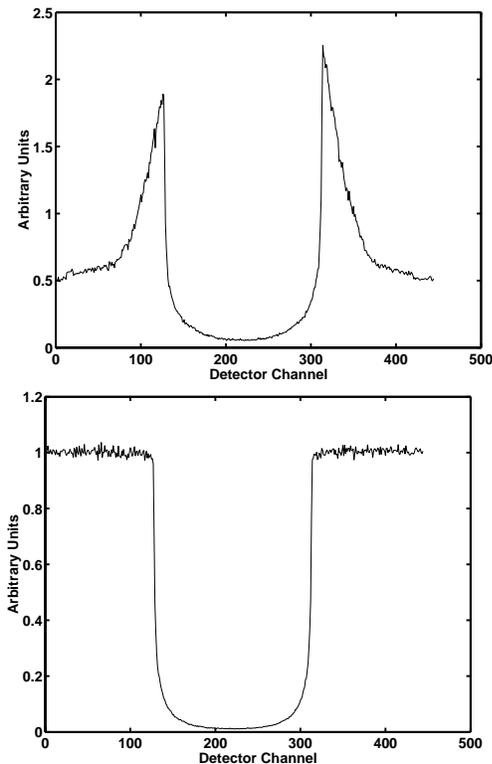
Figure 6.1: Left: one view of transmission data before air calibration. Right: same view after air calibration

6.2.3 Primary Speed and Afterglow

Solid-state scintillating detectors exhibit exponential signal decay once the X-ray flux is turned off. Depending on the characteristics of the detector, a residual signal persists from a few microseconds to milliseconds. In CT, this effect causes the measurements acquired at any particular view angle to be contaminated by residual signals from earlier views. This effect results in resolution loss and other artifacts.

The impulse response of this phenomenon can be characterized by a multi-term exponential [43, 45]:

$$h(t) = \sum_{n=1}^{N} \frac{\alpha_n}{\tau_n} e^{-\frac{t}{\tau_n}} U(t), \tag{6.1}$$

where $U(t)$ is the unit step function, α_n represents the relative strength of each decay term, and τ_n is the time constant for each decay term. Often times it is sufficient to use $N = 2$ terms.

Based on linear system theory and (6.1), a recursive relationship between ideal and corrupted (sampled) data provides the data correction algorithm for afterglow [43]:

$$y_k^t = \frac{y_k^c - \sum_{n=1}^{N} \beta_n e^{-\frac{\Delta t}{\tau_n}} S_{nk}}{\sum_{n=1}^{N} \beta_n}, \tag{6.2}$$

where

$$S_{nk} = x_{k-1} + e^{-\frac{\Delta t}{\tau_n}} S_{n(k-1)}$$

and

$$\beta_n = \alpha_n \left(1 - e^{-\frac{\Delta t}{\tau_n}} \right),$$

where y_k^t is the true measurements at sample k, y_k^c is the corrupted measurement, and Δt is the data sampling period. The details of the derivation are in [43].

6.2.4 Off-focal Radiation

In Chapter II we briefly mentioned off-focal radiation in our discussion of X-ray tubes. Off-focal radiation in the X-ray tube is caused mainly by secondary electrons [45]. High-speed electrons emitted from the cathode generate secondary electrons upon impact with the anode target. These secondary electrons can in turn return to the target at points outside the focal spot and produce X-rays. This results in a beam profile consisting of a high-intensity center spot surrounded by a

low-intensity halo. This effect can cause degradation in low-contrast detectability as well as shading artifacts.

Off-focal radiation can be reduced by source collimation, but it cannot be completely eliminated if the system is to retain a reasonable signal level. Proprietary software correction algorithms are used to eliminate off-focal radiation or reduce its effects to acceptable levels. The algorithm is complicated by the fact that off-focal radiation has a non-stationary point spread function. Fig. 6.2 illustrates one view of the GE QA phantom data (Fig. 6.4) before and after off-focal radiation correction.

6.2.5   Beam Hardening Correction

Beam hardening has already received a lot of attention in this dissertation The standard beam hardening algorithm available on most commercial scanners is water correction applied in the data domain. This approach is based on the fact that 80% of the human body is water, and therefore water correction is adequate in most cases. When the object of interest contains dense bones or metallic implants, water correction is not sufficient, and more elaborate schemes are necessary. See for example [49].

With a polyenergetic X-ray spectrum, the relationship between tissue thickness and the log signal is nonlinear. Fig. 6.3 illustrate the deviation of log-processed data from linearity with respect to the path length of X-rays in water. The relationship between the measured signal and the path length is a smooth, monotonically increasing concave function, with a one-to-one relationship to the ideal straight line. It is therefore possible to map the nonlinear curve onto the straight line. Following the notation of Chapter V, denote by F_i the nonlinear curve, which can be mapped to a
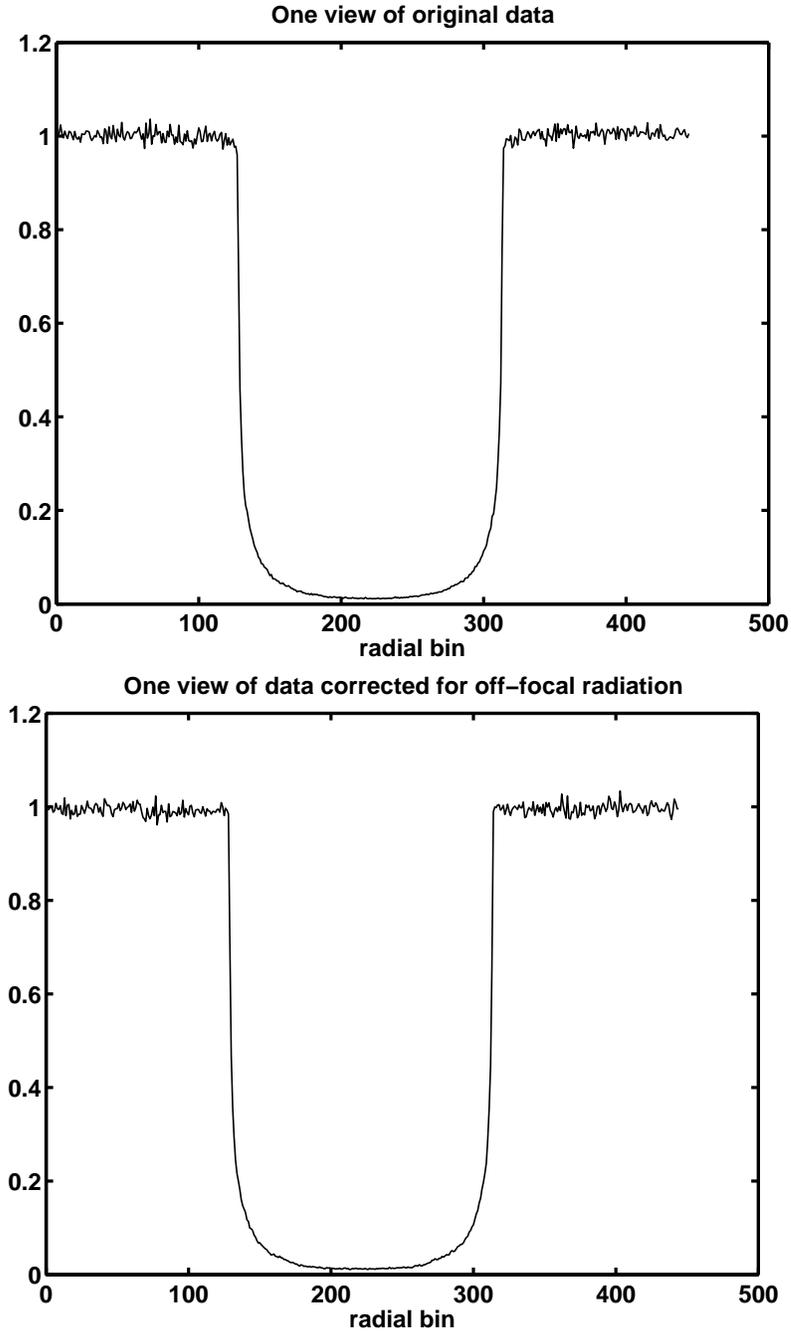
Figure 6.2: Effect of off-focal radiation on CT measurements. Top: data prior to correction (note blured edges). Bottom: Data after correction.

linear function P_i by a polynomial [45]:

$$P_i = \sum_n a_n F_i^n \tag{6.3}$$

where a_n is the polynomial coefficient for the nth term and can be computed using a minimum least-squares fit.
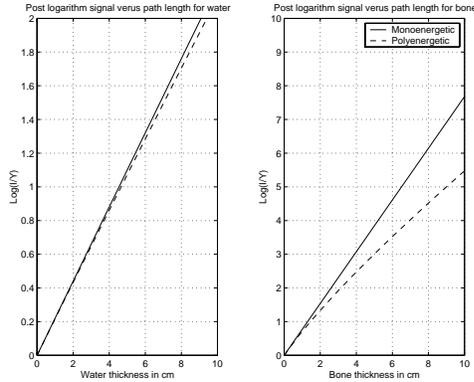
Figure 6.3: Water log transmission data deviation from a linear relationship with tissue thickness as a result of polyenergetic X-ray propagation and attenuation.

6.3   System Model

We utilized the information provided by GE Medical Systems about the data correction steps to formulate a forward model for the measurement process. During each step of the iterative algorithm, the system model is used to generate CT data from the current image, and that data is compared to the raw measurements.

The system model includes an accurate two-dimensional software forward projector based on the LightSpeed geometry. The projector computes the area of intersection between each pixel and a fan X-ray beam defined by the source and a detector element. This projector is more geometrically accurate than necessary and is used at this stage to eliminate any error inadvertently introduced by an inaccurate projector.

At this stage, the system model also includes water beam hardening effects, off-focal radiation, primary speed and afterglow, and blank scan and offset corrections. In the context of statistical reconstruction, noise reduction is accomplished by basing

the algorithm on a statistical distribution that closely reflects the measurements statistical properties. In this work, we assume Poisson statistics. In the algorithm implementation, we use a penalized likelihood approach to reduce image noise.

Mathematically, let $\bar{Y}_i$ be the mean of the data measured at detector i. The current system model is a discretized version of the following:

$$\bar{Y}_i = \int \int \int g_i(\beta) k_i(\gamma) I_i(\mathcal{E}) e^{-F_i(s(\beta,\gamma,\mathcal{E});\gamma)} d\mathcal{E} \, d\beta \, d\gamma + r_i, \tag{6.4}$$

where β and γ represent the source and fan angles, respectively, and $g_i(\beta)$ and $k_i(\gamma)$ represent the effects of afterglow and off-focal radiation, respectively. The function $F_i(\cdot)$ denotes beam hardening distortion of the line integral, and $I_i(\mathcal{E})$ is the energy-dependent intensity. We denote the line integral with $s(\beta, \gamma, \mathcal{E})$ and the background measurements with r_i.

6.3.1 Iterative Algorithm

We will not present a detailed derivation of the iterative algorithm, since it is very similar to what was presented in earlier chapters and in related work [33, 82]. The algorithm optimizes a cost function that is a surrogate of the Poisson log likelihood, given in (4.11), where $E[Y_i]$ is replaced with the expression in (6.4). The algorithm is a gradient descent algorithm, where the iterative updated depends on the ratio of the first and second derivatives of the cost function. Since we use ordered subsets to accelerate convergence, a precomputed denominator suffices.

6.4 Preliminary Results

In this section we discuss the preliminary results of iterative reconstruction of data from the GE LightSpeed scanner. We do not conduct an extensive bias-variance analysis and comparison of the performance of statistical reconstruction compared

to FBP. We try to compare noise performance at equivalent resolutions and assess the visual quality of the images.

6.4.1 GE QA Phantom

Fig. 6.4 shows an image of the GE QA phantom. This phantom is cylindrical, consisting mostly of water with some high resolution structures. The image shown in Fig. 6.4 is of a FBP standard reconstruction with 190mAs data. This high SNR makes it possible to discern the fine structure so clearly. We will compare FBP and iterative reconstruction performance at a much lower mAs.
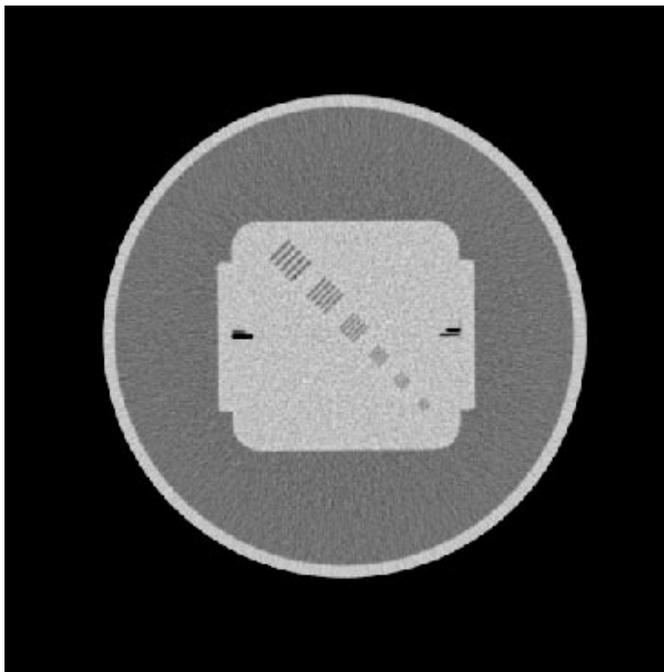
Figure 6.4: GE QA Phantom

6.4.2 Reconstruction Algorithms

We reconstruct a single slice of the GE QA phantom data acquired with a 10mAs scan. This is a fairly low signal level, and the reconstruction generally suffers from high noise that significantly decreases visibility of fine structures. We reconstruct

using the following algorithms:

- In-house fan beam FBP (with a Hanning window). Although an initial image with the better optimized GE FBP reconstruction is available, we use the in-house FBP image to initialize the statistical algorithm. This avoids any errors introduced by subtle geometric differences between the two algorithms (since the in-house FBP uses the same geometric model as the iterative algorithm). The FBP was applied to preprocessed data.

- Statistical reconstruction with edge-preserving penalty applied to raw data. The algorithm incorporates the full forward model. We use 10 iterations with 82 subsets and then reduce the number of subsets to 41 for 30 more iterations to avoid any subset-induced noise amplification.

- Statistical reconstruction with edge-preserving penalty applied to preprocessed data. The data is preprocessed with all steps except for beam hardening correction (which is applied in the log domain). The statistical algorithm includes the beam hardening model. We use the same number of iterations and subsets as above.

- GE standard FBP reconstructions were provided to us for comparison.

Fig. 6.5 shows the FBP reconstruction, and Fig. 6.7 and Fig. 6.8 show the 20- and 40-iteration images of the iterative reconstruction using raw data. The outlined regions (A and B) in Fig. 6.11 were used to compute variance, plotted in Fig. 6.12. Fig. 6.6 shows GEMS FBP reconstruction of the data and Fig. 6.10 shows the 40-iteration image of the iterative reconstruction using preprocessed data. Fig. 6.9 is the GE FBP reconstruction with some blur introduced to approximately match the resolution of the iterative reconstruction image.

UM FBP of 1 View
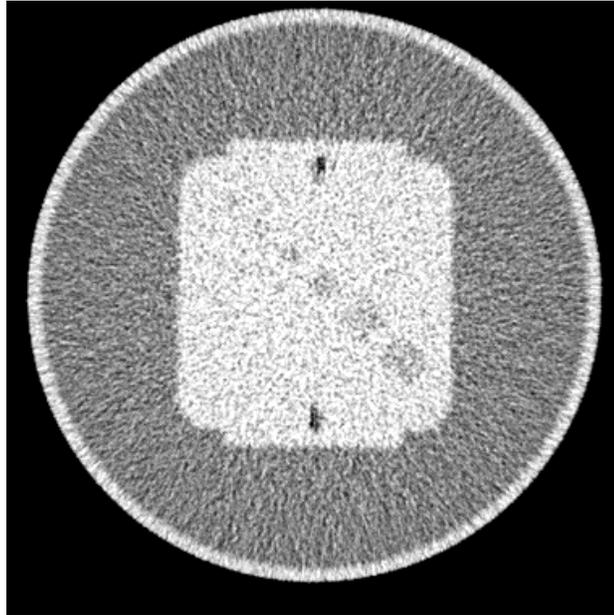
Figure 6.5: Filtered Backprojection Reconstruction

GE FBP of 1 View
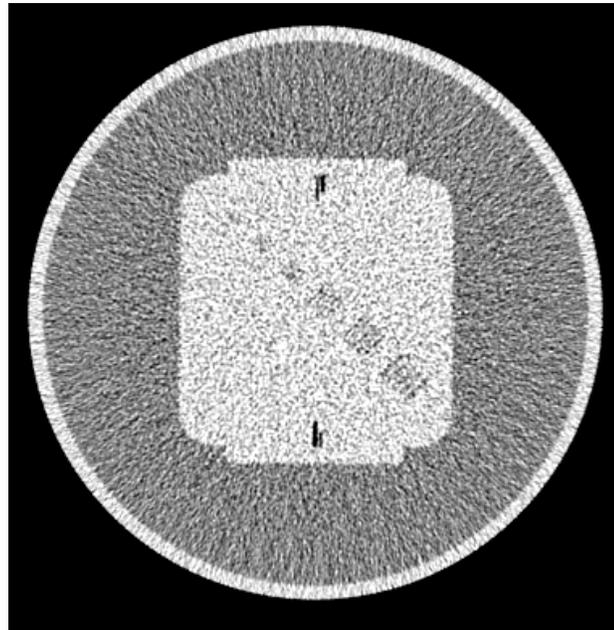
Figure 6.6: GE Filtered Backprojection Reconstruction

124
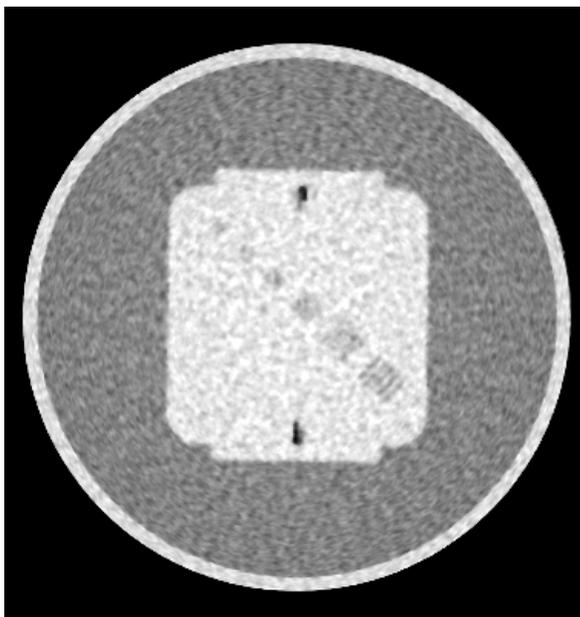
PL 20 Iterations

Figure 6.7: Iterative Reconstruction after 20 Iterations with Raw Data

PL 40 Iterations
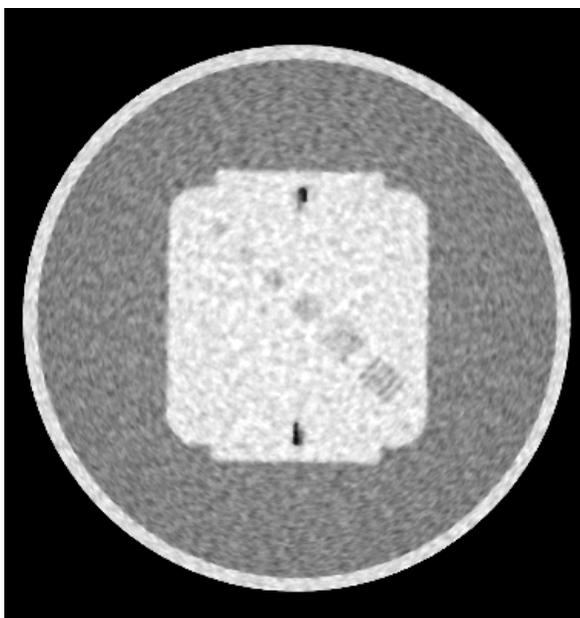
Figure 6.8: Iterative Reconstruction after 40 Iterations with Raw Data
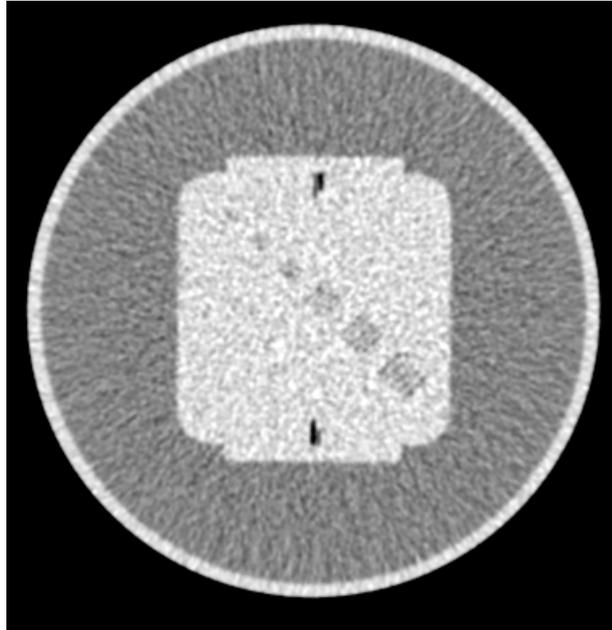
Figure 6.9: Blurred GE Filtered Backprojection Reconstruction

Figure 6.10: Iterative Reconstruction after 40 Iterations with Preprocessed Data
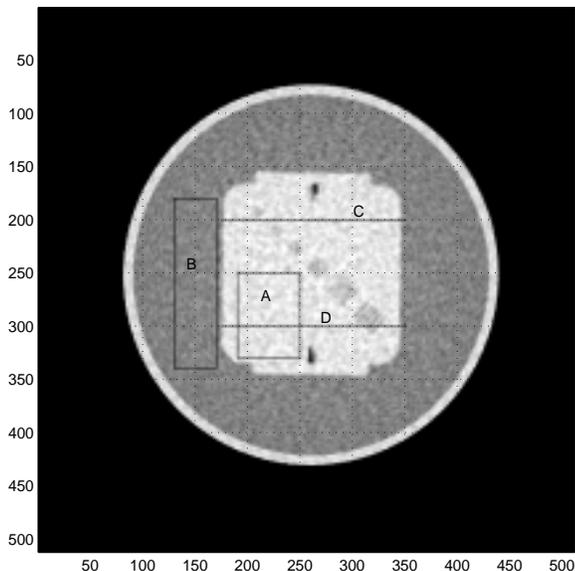
Figure 6.11: ROI A and B were used to compute variance, and rows between lines
C D were averaged to examine resolution

Reconstruction Method	Average FWHM	% Standard Deviation in Region A	% Standard Deviation in Region B
FBP	4.69	4.27	3.98
GE FBP	2.63	5.40	4.92
PL at 20 iterations	3.60	1.64	1.65
PL at 40 iterations	3.52	1.54	1.61
GE FBP blurred	3.57	2.04	1.82

Table 6.1: Noise performance comparison between GE FBP and statistical reconstructions of GE QA phantom at approximately equivalent resolution.

We compare the resolution of the different reconstructions by averaging 100 image rows between lines C and D in Fig. 6.11. We then compute the full-width half-max (FWHM) of the first derivative of the average row at the rightmost and leftmost edges of the center region of the phantom. An averaged FWHM is obtained from these results. We also compute the percent standard deviation in the two uniform regions A and B, inside and outside the center of the phantom (each region had 100×50 pixels). The first derivatives are illustrated in Fig. 6.13, and the results of

the FWHM and standard deviation comparisons are outlined in table 6.4.2. Note that the FWHM measurements are in units of pixels and the percent standard deviation is the ratio of the standard deviation to the mean.

At roughly equivalent resolution, iterative reconstruction gives better noise performance than (blurred) GE standard FBP reconstrucion. This gives us confidence that we can determine regularization parameters which will give equivalent resolution to FBP, with better noise reduction.
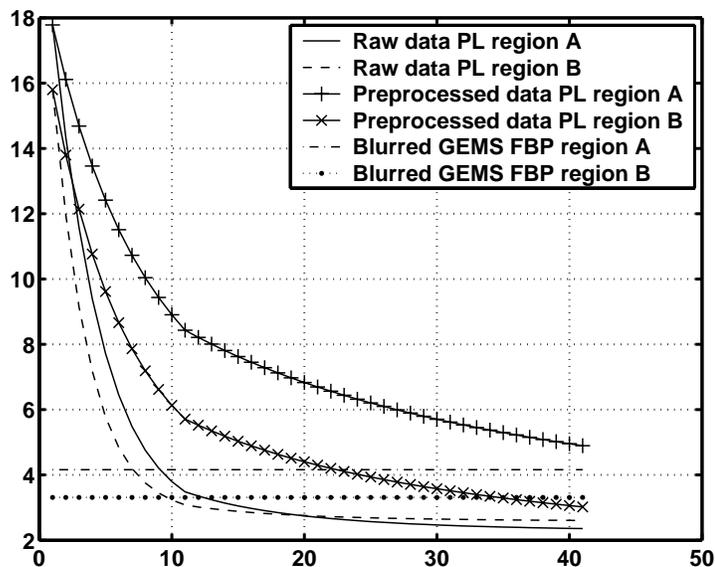
Figure 6.12: Variance of regions A and B as a function of iteration

These preliminary results indicate that while the iterative algorithm resolution performance is equivalent to FBP, it outperforms FBP in terms of noise reduction. Fig. 6.12 shows a promising decrease in variance with iteration. With statistical iterative reconstruction, the variance reduction is significantly larger with raw data than with preprocessed data. This illustrates the noise-reduction potential of statistical techniques.
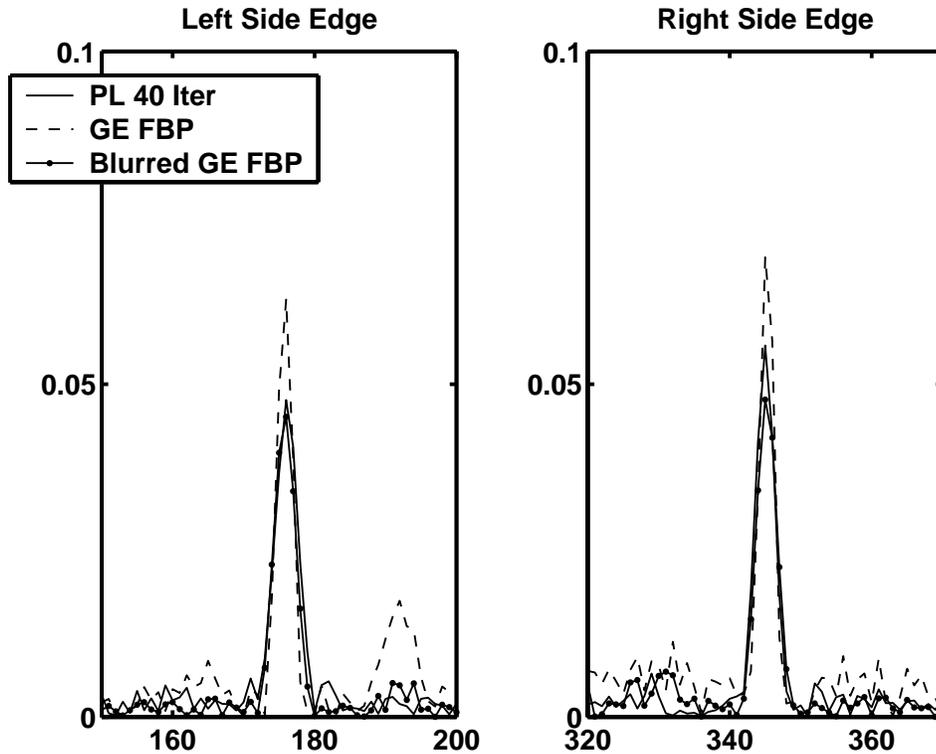
Figure 6.13: First derivatives of the average of 100 rows within the center region of the QA phantom

The noise texture in iterative reconstruction is different from the 'conventional' appearance of CT images. Whereas the FBP noise is correlated radially, the penalized likelihood noise is correlated tangentially. This may be due to the regularization penalty, and is an area of future investigation.

Less iterations may be necessary if more of the initial iterations are run with a larger number of subsets. Also, less iterations are needed if the iterative algorithm is initialized with a better image, such as the GEMS FBP reconstruction. Determining the optimal iteration/subset combination is important for considerations of computational load, and is a topic of future work.

% Standard Deviation		
Scan	Statistical Algorithm	GE Standard FBP
10 mAs	20.5	23.6
50 mAs	4.0	5.4
200 mAs	2.1	2.2

Table 6.2: Noise performance of GE FBP and statistical reconstruction of shoulder phantom.

6.4.3 Shoulder Phantom

We next investigate an anthropomorphic shoulder phantom. This phantom contains bone-like structures that are more attenuating than water. It also has fine details that can help assess the resolution performance of the different algorithms on data that are more similar to clinical practice. The asymetry of this phantom is also a good test of the accuracy of the system model.

We reconstruct 120 kVp data at three different mAs values: 10mAs, 50mAs and 200mAs. Fig. 6.15 - Fig. 6.17 illustrate the reconstruction results. The left hand side of each figure is the statistical reconstruction of raw data, and the right hand side is the GE FBP standard reconstruction[1] The upper and lower rows of each figure display soft tissue and lung window levels, respectively. The iterative algorithm parameters are similar to those described for the QA phantom in the previous section.

Statistical reconstruction gives lower noise, as expected. However, the improvement in noise performance is modest, as table 6.4.3 shows. In addition, the trained eye can observe subtle shading artifacts in the statistical reconstructions.

The cause of the shading artifact is possibly some degree of mismatch between

---

[1]Unfortunately, the GE FBP and statistical reconstruction images are not for the same slice. Fig. 6.14 is a reconstruction of the 200mAs data set with the UM inhouse FBP, provided for reference.
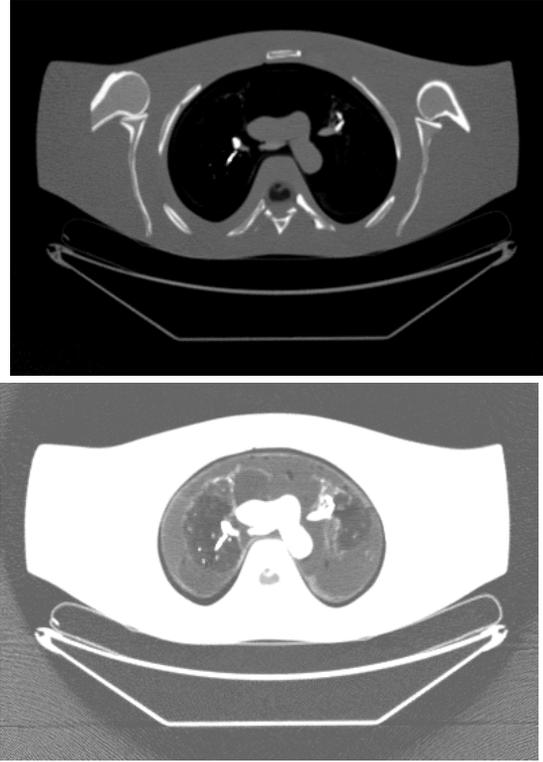
Figure 6.14: UM FBP reconstruction of 200mAs shoulder phantom data.

the system model and the actual data acquisition process. The artifact may also be caused by an implementation error. The next step in our collaboration with GE Medical Systems will be to determine and eliminate the cause of this shading artifact.

6.5   Conclusion

In this chapter we reported on the results of applying statistical reconstruction to data acquired on the GE LightSpeed scanner. Only single slice data has been considered so far. An iterative algorithm based on a system model that incorporates significant physical effects and system parameters has been shown to give promising results in terms of noise performance. Iterative reconstruction of raw data outperforms iterative reconstruction of preprocessed data, further confirming the notion that reconstruction based on accurate forward models and raw data produces supe-
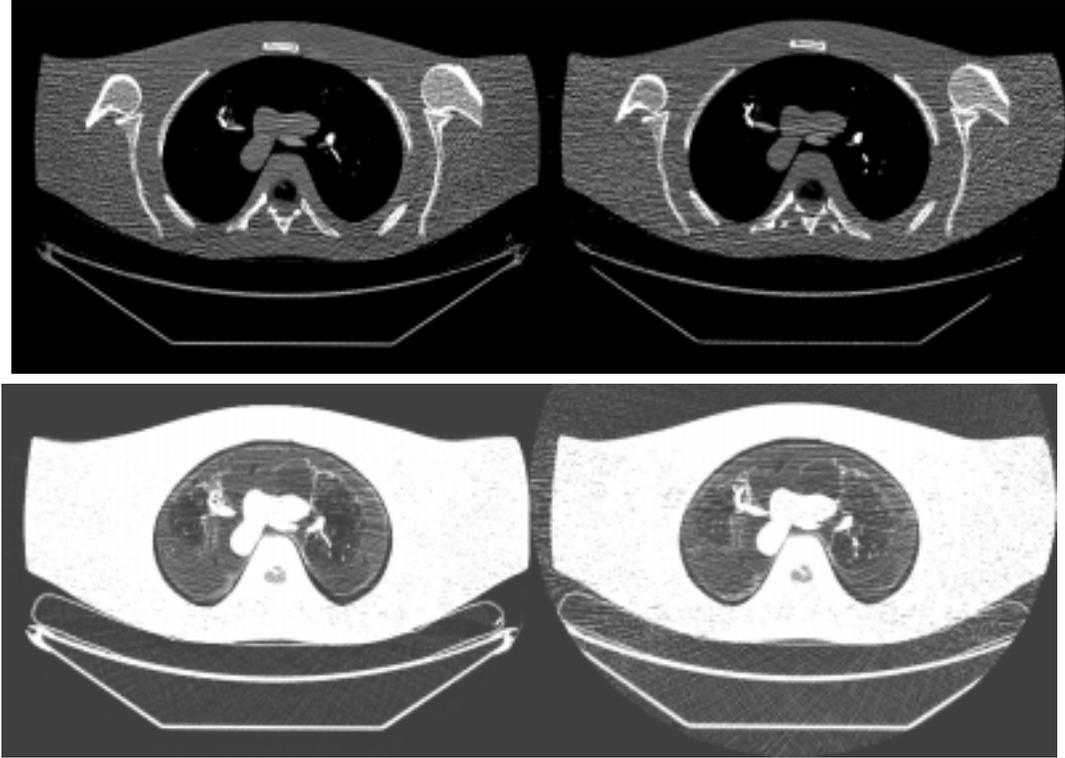
Figure 6.15: Statistical (left) and GE standard FBP (right) reconstructions of 10mAs shoulder phantom data.

rior image quality.

Extensions of this collaboration will include investigating and eliminating the cause of shading artifacts and confirming the accuracy of the system model, followed by extensive and systematic evaluation to quantify the advantages of iterative reconstruction.
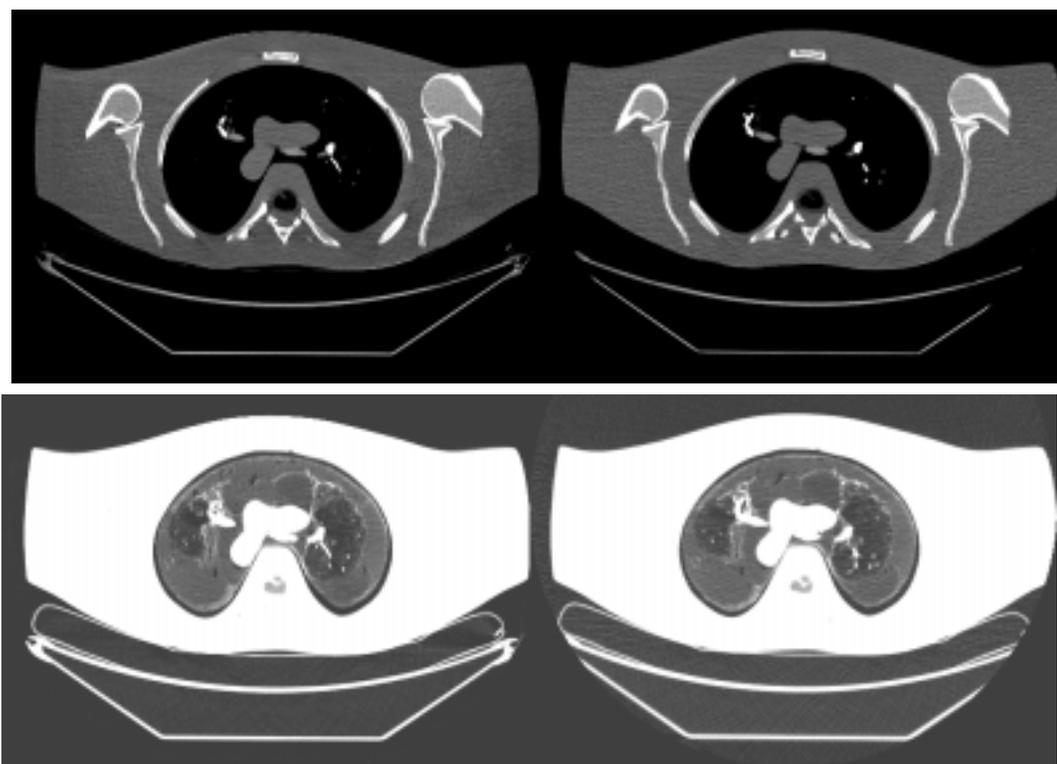
Figure 6.16: Statistical (left) and GE standard FBP (right) reconstructions of 50mAs shoulder phantom data.
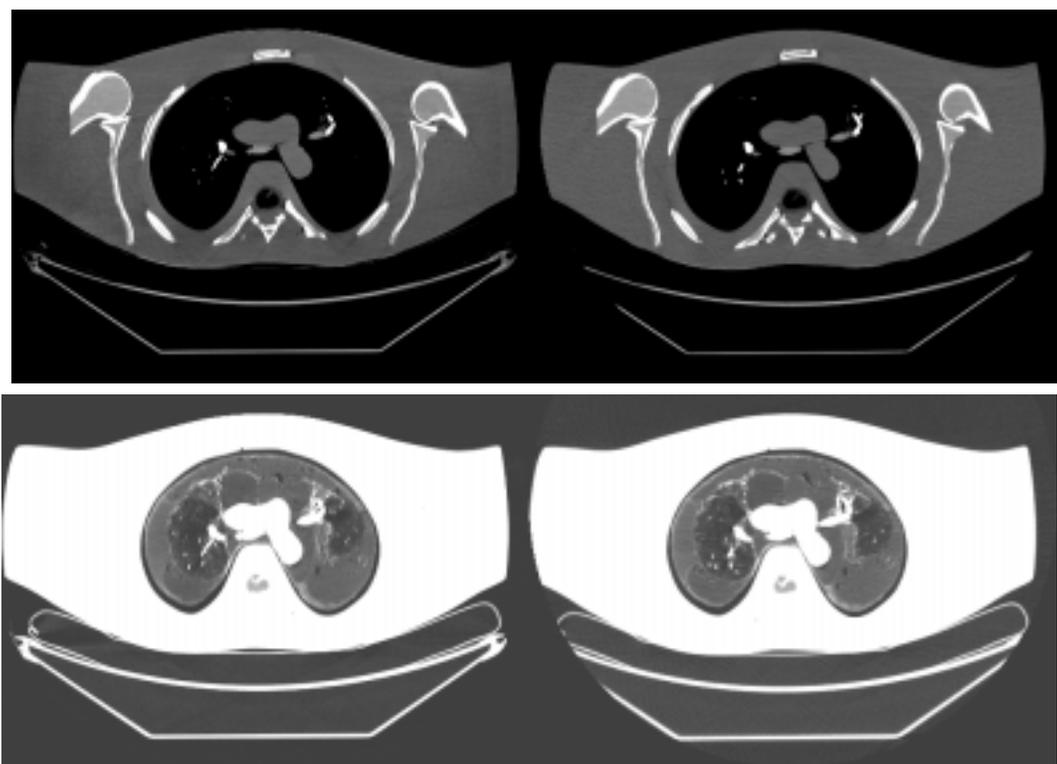
Figure 6.17: Statistical (left) and GE standard FBP (right) reconstructions of 200mAs shoulder phantom data.

CHAPTER VII

Conclusions and Future Work

7.1 Conclusions

This thesis introduces statistical reconstruction algorithms for X-ray CT within the general framework of formulating accurate data forward models for the CT problem. In addition to statistical and system models and algorithm derivations, we present simulation and real data reconstruction results and incorporate detector and X-ray tube effects in the system model for the GE lightSpeed.

In Chapter III we developed a compound Poisson statistical model for X-ray CT measurements that is more sophisticated and more realistic than the Poisson model. The model accounted for energy-dependent statistics, random light (or electron-hole pair) generation in detector crystals and additive readout noise. We also derived a practical approximation to the exact compound Poisson likelihood using a generalization of the saddle-point integration method. Comparisons between the exact compound Poisson likelihood, the saddle point approximate likelihood and the ordinary Poisson likelihood showed that the saddle point approximate likelihood may have some utility in image reconstruction in very low X-ray flux situations.

We introduced a statistical iterative reconstruction algorithm for energy dependent X-ray attenuation that produces images with significantly reduced beam hard-

ening artifacts in Chapter IV. The algorithm is based on polyenergetic X-ray attenuation, and is derived by successive applications of the optimization transfer principle. The algorithm is parallelizable and fairly simple to implement and can be accelerated with ordered subsets and a pre-computed surrogate curvature. The algorithm is applicable for an arbitrary number of non-overlapping materials in the object, and we demonstrated its effectiveness for bone and soft tissue objects. The algorithm requires prior knowledge of the incident X-ray spectrum and pre-segmented bone and soft tissue maps. In this regard it is comparable that the classical beam hardening correction method of Joseph and Spital.

In Chapter V we develop object models that relieve the polyenergetic algorithm of its pre-segmentation requirement. This enables the algorithm to accommodate mixed pixels that arise at tissue boundaries and to accommodate more tissues without requiring more scans. Two object models are proposed in Chapter V. The *displacement model* works well for anatomical tissues and objects with distinct tissues where mixed pixels arise mostly at tissue boundaries. The model is based on the fact that most soft tissues in the body have densities and attenuation properties similar to those of water. The *solution model* is designed for quantitative determination of the density of mineral solutions, and was shown to give more accurate results that conventional water correction.

In addition to algorithm validation with simulated and real data in Chapter IV and Chapter V, we applied one version of the algorithm to the data acquired on the GE LightSpeed scanner. The algorithm was based on a system model that included some of the more important 'spurious' effects that are routinely corrected for on the LightSpeed such as beam hardening, detector afterglow and tube off-focal radiation. The algorithm gave promising results in terms of noise performance. Iterative re-

construction of raw data outperformed iterative reconstruction of preprocessed data, further confirming the notion that reconstruction based on accurate forward models and raw data produces superior image quality.

7.2 Future Work

Like any research endeavor, the work presented in this thesis has generated some answers *and* questions. Realizing the impossibility of getting tenure as a student, I leave the answers to the following open problems for future researchers:

- The value of the compound Poisson likelihood derived in Chapter III for image reconstruction is still to be determined. For most clinical CT situations, signal levels are likely to be high enough for the regular Poisson model to be adequate. Are there situations (inefficient detectors, low signal, etc.) in which the more sophisticated compound Poisson model is useful? How about microCT? In addition to comparing the compound Poisson and regular Poisson likelihoods, the utility of the compound Poisson likelihood will be better understood by deriving and evaluating an algorithm based on it. The algorithm derivation will involve application of the optimization transfer principle and derivation of surrogates to the compound Poisson likelihood. There are also additional possible refinements of the statistical model, such as incorporating the wavelength dependence of the detector photodiodes and the nonlinearities in the detector electronics.

- The iterative algorithms of Chapter IV and Chapter III were tested based on a two-material categorization of tissues, namely bone (or bone mineral) and the general category of soft tissue (or water). One of the interesting problems arises when a third highly attenuating material such as a contrast agent is introduced.

How can the displacement model be generalized to include tissue fractions for a third base substance? How can issues relating to the temporal change in contrast agent concentration in the blood stream be resolved? Also, is there a possible hybrid object model that incorporates properties of the displacement and solution models that better captures the complexities of human tissues?

- For higher accuracy at mixed pixels, one can design the fraction functions of the displacement model to better match the mixture mass fractions of soft tissues and bone in the range of tissue densities between 1.0 g/cc and 1.92 g/cc.

- The solution model can easily incorporate the displacement effects in solution by reparameterizing the fraction functions in terms of mineral concentration.

- In this work we adopted a penalized-likelihood reconstruction approach. Penalized likelihood is attractive because it improves the conditioning of the problem and enables one to choose penalty functions that control desired properties such as edge preservation. However, we have suffered through the process of determining the values of regularization parameter in the absence of systematic methods for choosing them. Will it be possible to generalize the results of [68] to edge-preserving penalties? Noting the circularly correlated noise of Fig. 6.8, one may be inclined to consider whether the CT reconstruction problem will require a unique penalty design.

- In Chapter V, single-slice data obtained from cone beam systems were contaminated by scatter. Short of measuring or simulating a scatter estimate, we applied an ad-hoc method for scatter estimation from the real sinogram. Better scatter estimation is certainly one of the first things that can be done to improve the accuracy of the proposed algorithms.

- In addition to scatter, there are several other effects that are important for the system model to account for. These include partial volume and exponential edge-gradient effects [42, 50]. These effects occur when a scanned object is partially intruded into the scan plane. Addressing the partial volume effect may require revising our projector. The current projector calculates areas of overlap between X-ray beams and pixels. To account for partial volume may require a more refined projector of thin rays traversing pixels.

- The relationship in (5.7) can be determined through measurements made on wedge or step phantoms of known thicknesses and densities, and can be mathematically described by a low-order polynomial. Assuming the beam hardening polynomial is slowly varying and independent of detector channel, it is certainly feasible to derive an image reconstruction algorithm that estimates the polynomial coefficients simultaneously with the image. This algorithm can take the form of an alternating maximization algorithm.

- One inherent limitation of our polyenergetic reconstruction methods lies in the fact that we are reducing a problem with $K \times p$ unknowns to one with p unknowns only. This limitation may be overcome (at least when $K = 2$) with dual-energy reconstruction. Recent advances with dual-energy statistical reconstruction [33, 71] may enable enough dose reduction for the two energy scans needed such that the cost and patient dose concerns are eliminated.

- The system model so far has assumed single-slice fan-beam geometry. One of the advantages of iterative reconstruction is its applicability to more arbitrary geometries such as cone beam and helical CT, and that is certainly an area in which the work presented in this thesis may be extended.

APPENDICES

APPENDIX A

Segmentation-free Polyenergetic Algorithm Surrogate Function

In this appendix we outline details of the derivation of the statistical algorithm for polyenergetic CT described in Chapter V. The goal of the iterative algorithm is to find the minimizer in (5.16). We use the optimization transfer principle [15,16,22,25] which has proven to be a very effective tool in simplifying such problems. Under certain conditions, a simpler and easier to minimize surrogate function can replace the likelihood in the iterative algorithm.

We apply successive surrogates to the penalized likelihood cost function, ultimately resulting in a cost function that is quadratic and separable in ρ. Quadratic functions are more easily minimized, and separability allows the algorithm to update all pixels simultaneously. The derivation is complicated by the fact that tissue fractions are now functions of the independent variable ρ and that the polyenergetic dependence is expressed indirectly through $F_i(\underline{s}_i(\rho))$.

The goal of the algorithm is to minimize the Poisson negative log likelihood:

$$-L(\rho) = \sum_{i=1}^{N} h_i(\bar{Y}_i(\underline{s}_i(\rho)) + r_i) \qquad (A.1)$$

$$h_i(t) \triangleq -Y_i \log t + t, \qquad (A.2)$$

where $\bar{Y}_i(\underline{s}_i(\rho))$ is given in (5.6). Although the algorithm includes a regularization term, in this appendix we focus on algorithm derivation with the likelihood term only. Generalizing the result to include regularization is straightforward.

Define the new function g_i^n that has a surrogate q_i^n

$$g_i^n(l) \quad \triangleq \quad h_i\left(I_i e^{-l} + r_i\right)$$

$$\leq \quad q_i^n(l), \tag{A.3}$$

where

$$q_i^n(l) = g_i^n(l_i^n) + \dot{g}_i^n(l_i^n)(l - l_i^n) + \frac{1}{2}C_i^n(l - l_i^n)^2. \tag{A.4}$$

It is possible to design curvatures such that q_i^n satisfies the conditions of optimization transfer. One possibility is the maximum curvature give by [25, 82]:

$$C_i^n = \left[I_i - \frac{Y_i r_i}{(1 + r_i)^2}\right]_+ \leq [I_i]_+. \tag{A.5}$$

Note that q_i^n is a quadratic in l, but not in the variable of interest, ρ. Combining (A.4) and (A.2), we arrive at the following surrogate function:

$$Q_1(\rho; \rho^n) = \sum_{i=1}^{N} q_i^n(F_i(\underline{s}_i(\rho))). \tag{A.6}$$

Next we derive a quadratic surrogate in $\underline{s}_i$. Let

$$\tilde{q}_i^n(\underline{s}_i) \quad \triangleq \quad q_i^n(F_i(\underline{s}_i)) \tag{A.7}$$

$$\leq \quad p_i^n(\underline{s}_i) \tag{A.8}$$

where

$$
\begin{aligned}
p_i^n(\underline{s}_i) &= \tilde{q}_i^n(\underline{s}_i^n) + \nabla \tilde{q}_i^n(\underline{s}_i^n) \cdot (\underline{s}_i - \underline{s}_i^n) + \frac{1}{2}(\underline{s}_i - \underline{s}_i^n)' \cdot \tilde{\mathbf{C}}_i^n \cdot (\underline{s}_i - \underline{s}_i^n) \\
&= \tilde{q}_i^n(\underline{s}_i^n) + \sum_{k}^{K} \left.\frac{\partial \tilde{q}_i^n(\underline{s}_i)}{\partial s_i^k}\right|_{\underline{s}_i = \underline{s}_i^n} (s_i^k - s_i^{k,n}) + \frac{1}{2}\sum_{k}^{K} \tilde{C}_i^{n,k}(s_i^k - s_i^{k,n})^2.
\end{aligned}
$$

The surrogate thus far is:

$$Q_2(\rho; \rho^n) = \sum_{i=1}^{N} p_i^n(\underline{s}_i(\rho)).$$
(A.9)

Next we derive a quadratic surrogate in ρ. Let

$$\tilde{p}_i^n(\rho) \triangleq p_i^n(\underline{s}_i(\rho)) \leq \zeta_i^n(\rho)$$
(A.10)

where

$$
\begin{aligned}
\zeta_i^n(\rho) &= \tilde{p}_i^n(\rho^n) + \nabla \tilde{p}_i^n(\rho^n) \cdot (\rho - \rho^n) + \frac{1}{2}(\rho - \rho^n)' \cdot \hat{\mathbf{C}}_{\mathbf{i}}^{\mathbf{n}} \cdot (\rho - \rho^n) \\
&= \tilde{p}_i^n(\rho^n) + \sum_{j}^{p} \left.\frac{\partial \tilde{p}_i^n(\rho)}{\partial \rho_j}\right|_{\rho=\rho^n} (\rho_j - \rho_j^n) + \frac{1}{2}\sum_{j}^{p} \hat{C}_i^{n,j}(\rho_j - \rho_j^n)^2.
\end{aligned}
$$
(A.11)

We again assume that $\hat{C}_i^n$ is a curvature that ensures ζ_i^n is a sound surrogate. By construction, this surrogate is also separable since all of the pixels are decoupled in the sums over j. The final surrogate is:

$$Q(\rho; \rho^n) = \sum_{j=1}^{p} \sum_{i=1}^{N} \zeta_i^n(\rho).$$
(A.12)

Tedious algebra shows that the first and second derivatives of the surrogate evaluated at ρ^n are:

$$
\begin{aligned}
\left.\frac{\partial Q}{\partial \rho_j}\right|_{\rho=\rho^n} &= \sum_{i=1}^{N}\sum_{k=1}^{K} a_{ij} \left.\frac{\partial(\rho_j f_j^k(\rho_j))}{\partial \rho_j}\right|_{\rho=\rho^n} \dot{g}_i^n(F_i(\underline{s}_i^n))\frac{\partial F_i(\underline{s}_i^n)}{\partial s_i^k} \\
&= \sum_{i=1}^{N}\left(\frac{Y_i}{\bar{Y}_i(\underline{s}_i^n) + r_i} - 1\right) a_{ij}\sum_{k=1}^{K}\left.\frac{\partial\left(\rho_j f_j^k(\rho_j)\right)}{\partial \rho_j}\right|_{\rho=\rho^n}\frac{\partial F_i(\underline{s}_i^n)}{\partial s_i^k}I_i e^{-F_i(\underline{s}_i^n)} \\
&= \sum_{i=1}^{N}\sum_{k=1}^{K} a_{ij}\left.\frac{\partial\left(\rho_j f_j^k(\rho_j)\right)}{\partial \rho_j}\right|_{\rho=\rho^n}\left(1 - \frac{Y_i}{\bar{Y}_i(\underline{s}_i^n) + r_i}\right)\nabla_k \bar{Y}_i(\underline{s}_i^n) \quad (A.13) \\
&= -\left.\frac{\partial L}{\partial \rho_j}\right|_{\rho=\rho^n} \quad (A.14) \\
\left.\frac{\partial^2 Q}{\partial \rho_j^2}\right|_{\rho=\rho^n} &= \sum_{i=1}^{N} \hat{C}_i^{n,j}. \quad (A.15)
\end{aligned}
$$

The resulting algorithm is a diagonally-preconditioned gradient descent method of the following form:

$$\rho^{n+1} = \left[\rho^n - \mathbf{D}^{-1} \nabla Q(\rho^n) \right]_+ , \tag{A.16}$$

where $\nabla Q(\rho^n)$ is the gradient of the cost function, $\mathbf{D}$ is a diagonal curvature matrix that influences the rate of convergence and monotonicity of the algorithm, and $[\cdot]_+$ enforces the nonnegativity constraint.

The entries of $\mathbf{D}$ are determined by our choices for the various curvature terms in the above derivation ($\hat{\mathbf{C}}_{\mathbf{i}}^{\mathbf{n}}$, $\tilde{\mathbf{C}}_{\mathbf{i}}^{\mathbf{n}}$, C_i^n). For a monotonic algorithm, the elements of $\mathbf{D}$ must be such that a surrogate satisfies the conditions of optimization transfer. To derive a version of $\mathbf{D}$ that gives a monotone algorithm, it is sufficient to assume a finite upper bound on the density ρ and to require the fraction functions of the object model to be twice continuously differentiable. Since monotonicity will be compromised anyway by using ordered subsets to accelerate the algorithm, a precomputed approximate $\mathbf{D}$ will suffice. By making certain assumptions about the object [22], we derive an approximate precomputed "curvature" matrix using second derivatives of the surrogates used at different stages of the derivation.

Using the second derivatives of the functions in the second-order terms of their surrogates, we can show that the second derivative of the surrogate function $Q(\rho; \rho^n)$ evaluated at $\rho = \rho^n$ is:

$$
\begin{aligned}
\left. \frac{\partial^2 Q}{\partial \rho_j^2} \right|_{\rho=\rho^n} &= \sum_{i=1}^{N} \hat{C}_i^{n,j} \\
&= \sum_{i=1}^{N} \sum_{k=1}^{K} \left\{ \dot{g}_i^n(F_i(\underline{s}_i^n)) \frac{\partial^2 F_i(\underline{s}_i^n)}{\partial s_i^{k^2}} + C_i^n \left(\frac{\partial F_i(\underline{s}_i^n)}{\partial s_i^{k^2}} \right)^2 \right\} \left(\left. \frac{\partial s_i^k(\rho_j)}{\partial \rho_j} \right|_{\rho=\rho^n} \right)^2 \\
&= \sum_{i=1}^{N} \sum_{k=1}^{K} a_{ij}^2 \left(\left. \frac{\partial}{\partial \rho_j} (\rho_j f_j^k(\rho_j)) \right|_{\rho=\rho^n} \right)^2 \left\{ \dot{g}_i^n(F_i(\underline{s}_i^n)) \frac{\partial^2 F_i(\underline{s}_i^n)}{\partial s_i^{k^2}} + Y_i \left(\frac{\partial F_i(\underline{s}_i^n)}{\partial s_i^k} \right)^2 \right\} ,
\end{aligned}
$$

where we took advantage of (4.34). This expression is iteration-dependent, and can be simplified even further. Assuming that the object is mostly water, we can make the following approximate:

$$
\frac{\partial}{\partial \rho_j}(\rho_j f_j^k(\rho_j)) = f_j^k(\rho_j) + \rho_j \dot{f}_k^j(\rho_j)
$$

$$
\approx \begin{cases} 1 & \text{for } k = 1 \text{ (water)} \\ 0 & \text{for } k = 2 \text{ (bone)} \end{cases}
$$

where we assume that the density of bone and the first derivative of the water fraction function are zero when the object is mostly water. We can also approximation

$$
\begin{aligned}
F_i(\underline{s}_i) &\approx m_w(\mathcal{E}_{\text{eff}})s_i^1, \\
\frac{\partial F_i(\underline{s}_i)}{\partial s_i^1} &\approx m_w(\mathcal{E}_{\text{eff}}), \\
\frac{\partial^2 F_i(\underline{s}_i)}{\partial s_i^{1^2}} &\approx 0,
\end{aligned}
$$

where s_i^1 is the line integral of the density of water, and

$$
\mathcal{E}_{\text{eff}} \triangleq \frac{\displaystyle\int \mathcal{E}I_i(\mathcal{E})d\mathcal{E}}{\displaystyle\int I_i(\mathcal{E})d\mathcal{E}}
$$

is the effective X-ray energy. With these approximations, we arrive at a precomputed version of the curvature:

$$
d_j = m_w^2(\mathcal{E}_{\text{eff}}) \sum_{i=1}^{N} a_{ij}^2 Y_i,
$$

where d_j is the jth entry of the diagonal matrix $\mathbf{D}$

In summary, the algorithm takes the following shape:

- Precompute d_j

- Tabulate $F_i(\cdot)$ and its gradient over an appropriate range of break points

- Initialize with $\hat{\rho}$ and choose displacement or solution model

- **for** each iteration $n = 1, ..., \text{niter}$

  - **for** each subset $S = 1, \ldots, M$

    * compute $\hat{s}_i^k = \sum_{j=1}^{p} a_{ij} f_k^j(\hat{\rho})\hat{\rho}_j$ for $k = 1, \ldots, K$, set $\underline{\hat{s}}_i = [\hat{s}_i^1, ..., \hat{s}_i^K]$

    * compute $\bar{Y}_i(\underline{\hat{s}}_i)$ and $\nabla \bar{Y}_i(\underline{\hat{s}}_i) = [\nabla_1 \bar{Y}_i(\underline{\hat{s}}_i), ..., \nabla_K \bar{Y}_i(\underline{\hat{s}}_i)]$ using interpo-

      lation from the precomputed tables of F_i and its gradient

    * evaluate $\hat{N}_j = \sum_{i=1}^{N} \sum_{k=1}^{K} a_{ij} \left. \frac{\partial \left(\rho_j f_j^k(\rho_j) \right)}{\partial \rho_j} \right|_{\rho=\hat{\rho}} \left(1 - \frac{Y_i}{\bar{Y}_i(\underline{\hat{s}}_i) + r_i} \right) \nabla_k \bar{Y}_i(\underline{\hat{s}}_i)$

    * compute

$$
\hat{\rho}_j = \left[\hat{\rho}_j - \frac{M\hat{N}_j + \beta \left. \frac{\partial \tilde{R}}{\partial \rho_j} \right|_{\rho=\hat{\rho}}}{d_j + \beta \left. \frac{\partial^2 \tilde{R}}{\partial \rho_j^2} \right|_{\rho=\hat{\rho}}} \right]_+ , \, j = 1, \ldots, p \qquad \text{(A.17)}
$$

  - **end**

- **end**

where $\tilde{R}$ is an appropriate surrogate for the regularization penalty.

APPENDIX B

X-ray CT Measurements Probability Mass Functions

In this appendix we derive the probability mass functions for the different cases of X-ray detection described in chapter III. In each case, we use contour integration to invert the Z-transform of the moment generating function. We make use of the Cauchy integral theorem:

$$
\frac{1}{2\pi j} \oint_c z^{-k} dz = \begin{cases} 1, & k = 1, \\ 0, & k \neq 1, \end{cases} \tag{B.1}
$$

where c is a counterclockwise contour that encircles the origin.

The exact likelihoods follow from the probability mass functions. Except for the simplest case of monoenergetic radiation and deterministic light generation in the detector, the likelihoods have infinite series, and are therefore not practical for image reconstruction, as pointed out in chapter III.

- **Monoenergetic X-rays and Deterministic Light**

  For the idealized case of monoenergetic X-rays and deterministic light scintil-

lation, the moment generating function was shown to be

$$
\begin{aligned}
g_Y(z) &= \exp(-\bar{N}(1 - z^{x_o})) \\
&= e^{-\bar{N}} e^{\bar{N} z^{x_o}} \\
&= e^{-\bar{N}} \sum_{k=0}^{\infty} \frac{1}{k!} (\bar{N} z^{x_o})^k
\end{aligned}
\tag{B.2}
$$

To find the p.m.f of Y, we use the contour integration version of the inverse Z-Transform:

$$
\begin{aligned}
P_Y(y) &= e^{-\bar{N}} \sum_{k=0}^{\infty} \frac{1}{k!} \bar{N}^k \frac{1}{2\pi j} \oint_c z^{-y-1} (z^{x_o})^k dz \\
&= e^{-\bar{N}} \sum_{k=0}^{\infty} \frac{1}{k!} \bar{N}^k \frac{1}{2\pi j} \oint_c z^{kx_o - y - 1} dz \\
&= \frac{e^{-\bar{N}}}{(y/x_o)!} \bar{N}^{\frac{y}{x_o}},
\end{aligned}
\tag{B.3}
$$

when $\frac{y}{x_o}$ is an integer and where we have used the Cauchy integral theorem (B.1).

- **Monoenergetic X-rays and Poisson Light**

  In section 3.4.2 we showed the that moment generating function for the case of monoenergetic X-rays with Poisson light in the detector is:

$$
\begin{aligned}
g_Y(z) &= \exp(-\bar{N}(1 - e^{-x_o(1-z)})) \\
&= e^{-\bar{N}} \exp(\bar{N} e^{-x_o(1-z)}) \\
&= e^{-\bar{N}} \sum_{k=0}^{\infty} \frac{1}{k!} \left(\bar{N} e^{-x_o(1-z)} \right)^k \\
&= e^{-\bar{N}} \sum_{k=0}^{\infty} \frac{1}{k!} (\bar{N} e^{-x_o})^k e^{kx_o z} \\
&= e^{-\bar{N}} \sum_{k=0}^{\infty} \frac{1}{k!} (\bar{N} e^{-x_o})^k \sum_{i=1}^{\infty} \frac{1}{i!} (kx_o)^i z^i.
\end{aligned}
\tag{B.4}
$$

To find the p.m.f of Y, we use the contour integration version of the inverse Z-Transform:

$$
\begin{aligned}
P_Y(y) &= \frac{1}{2\pi j} \oint_c z^{-y-1} g_Y(z) dz \\
&= e^{-\bar{N}} \sum_{k=0}^{\infty} \frac{1}{k!} (\bar{N}e^{-x_o})^k \sum_{i=1}^{\infty} \frac{1}{i!} (kx_o)^i \frac{1}{2\pi j} \oint_c z^{-y-1+i} dz \\
&= e^{-\bar{N}} \sum_{k=0}^{\infty} \frac{1}{k!} (\bar{N}e^{-x_o})^k \frac{(kx_o)^y}{y!},
\end{aligned}
\tag{B.5}
$$

where we used Cauchy's theorem to reduce the sum over i to a single term.

- **Polyenergetic X-rays and Poisson Light**

  The moment generating function for the realistic case of polyenergetix X-rays and Poisson light was derived in section 3.5. We restate it here for convenience and manipulate it to express it in terms of the moment generation functions of each energy l:

$$
\begin{aligned}
g_Y(z) &= \exp\left(-\bar{N}(1 - \sum_{l=1}^{L} \tilde{p}_l e^{-x_l(1-z)})\right) \\
&= e^{-\bar{N}} \exp\{\bar{N} \sum_{l=1}^{L} \tilde{p}_l e^{-x_l(1-z)}\} \\
&= e^{-\bar{N}} \Pi_{l=1}^{L} \exp\{\bar{N}\tilde{p}_l e^{-x_l(1-z)}\} \\
&= e^{-\bar{N}} \Pi_{l=1}^{L} g_Y^l(z),
\end{aligned}
\tag{B.6}
$$

where

$$
\begin{aligned}
g_Y^l(z) &= \exp\{\bar{N}\tilde{p}_l e^{-x_l(1-z)}\} \\
&= \sum_{k=1}^{\infty} \frac{1}{k!} \left(\bar{N}\tilde{p}_l e^{-x_l(1-z)}\right)^k \\
&= \sum_{k=1}^{\infty} \frac{1}{k!} \left(\bar{N}\tilde{p}_l e^{-x_l}\right)^k e^{x_l z k} \\
&= \sum_{k=1}^{\infty} \frac{1}{k!} \left(\bar{N}\tilde{p}_l e^{-x_l}\right)^k \sum_{i=0}^{\infty} \frac{1}{i!} (x_l k)^i z^i.
\end{aligned}
\tag{B.7}
$$

Using the contour integral inverse of the Z-transform, we get

$$
\begin{aligned}
P_Y^l(y) &= \sum_{k=1}^{\infty} \frac{1}{k!} \left(\bar{N} \tilde{p}_l e^{-x_l} \right)^k \sum_{i=0}^{\infty} \frac{1}{i!} (x_l k)^i \frac{1}{2\pi j} \oint_c z^{-y-1+i} dz \\
&= \sum_{k=1}^{\infty} \frac{1}{k!} \left(\bar{N} \tilde{p}_l e^{-x_l} \right)^k \frac{(x_l k)^y}{y!}.
\end{aligned}
\tag{B.8}
$$

Using the convolution property of the Z-transform in (B.6) above, we get the p.m.f. of Y:

$$
P_Y(y) = e^{-\bar{N}} \left[P_Y^1(y) * P_Y^2(y) * \dots * P_Y^L(y) \right].
\tag{B.9}
$$

To compute the p.m.f. of Y requires L convolutions, where L is equal to the number of distinct energies in the spectrum. Each term in the convolution consists of an infinite series. Recall that the number of energies L results from discretizing (for computational purposes) what is essentially a continuous spectrum. Clearly, using the log of (B.9) as a log likelihood is not practical for image reconstruction.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] S. Ahn and J. A. Fessler. Globally convergent ordered subsets algorithms: Application to tomography. In *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, volume 2, pages 1064–8, 2001.

[2] R. E. Alvarez and A. Macovski. Energy-selective reconstruction in X-ray computerized tomography. *Physics in Medicine and Biology*, 21(5):733–744, 1976.

[3] F. J. Beekman and C. Kamphuis. Ordered subset reconstruction for x-ray CT. *Phys. Med. Biol.*, 46(7):1835–55, July 2001.

[4] G. Besson. New classes of helical weighting algorithms with applications to fast CT reconstruction. *Med. Phys.*, 25(8):1521–32, August 1998.

[5] J. Beutel, H. L. Kundel, and R. L. Van Metter, editors. *Handbook of Medical Imaging, Volume 1.* SPIE Press, Bellingham, 2000.

[6] N. Bleistein and R. A. Handelsman. *Asymptotic Expansions of Integrals.* Dover Publications, New York, 1975.

[7] J. M. Boone and J. A. Seibert. An accurate method for computer-generating tungsten anode x-ray spectra from 30 to 140 kv. *Med. Phys.*, 24(11):1661–70, November 1997.

[8] R. A. Brooks and G. D. Chiro. Beam hardening in X-ray reconstruction tomography. *Physics in Medicine and Biology*, 21(3):390–398, 1976.

[9] J. A. Browne and T. J. Holmes. Developments with maximum likelihood X-ray computed tomography. *IEEE Transactions on Medical Imaging*, 11(1):40–52, March 1992.

[10] H. N. Cardinal and A. Fenster. An accurate method for direct dual-energy calibration and decomposition. *Med. Phys.*, 17(3):327–41, May 1990.

[11] D. D. Cody and M. J. Flynn. A technique for measuring regional bone mineral density in human lumbar vertebral bodies. *Med. Phys.*, 16(5):766–772, 1989.

[12] A. M. Cormack. Representation of a function by its line integrals, with some radiological applications. *J. Appl. Phys.*, 34:2722–7, 1963.

[13] B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens. Reduction of metal artifacts in x-ray computed tomography using a transmission maximum a posteriori algorithm. *IEEE Transactions on Nuclear Science*, 47(3):977–981, June 2000.

[14] B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens. An iterative maximum-likelihood polychromatic algorithm for CT. *IEEE Tr. Med. Im.*, 20(10):999–1008, October 2001.

[15] A. R. De Pierro. On the relation between the ISRA and the EM algorithm for positron emission tomography. *IEEE Transactions on Medical Imaging*, 12(2):38–333, June 1993.

[16] A. R. De Pierro. A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. *IEEE Transactions on Medical Imaging*, 14(1):132–137, March 1995.

[17] M. Defrise, F. Noo, and H. Kudo. Rebinning-based algorithms for helical cone-beam CT. *Phys. Med. Biol.*, 46(11):2911–28, November 2001.

[18] A. H. Delaney and Y. Bresler. Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography. *IEEE Transactions on Image Processing*, 7(2):104–221, February 1998.

[19] G Di Chiro, R. A. Brooks, L. Dubal, and E. Chew. The apical artifact: elevated attenuation values toward the apex of the skull. *J. Comp. Assisted Tomo.*, 2(1):65–70, January 1978.

[20] A. J. Duerinckx and A. Mackovski. Nonlinear polychromatic and noise artifacts in x-ray computed tomography images. *J. Comp. Assisted Tomo.*, 3(4):519–26, August 1979.

[21] I. Elbakri and J. A. Fessler. Method for statistically reconstructing a polyenergetic x-ray computed tomography image and image reconstructor apparatus utilizing the method, 2003. US Patent 6,507,633.

[22] I. A. Elbakri and J. A. Fessler. Statistical image reconstruction for polyenergetic X-ray computed tomography. *IEEE Tr. Med. Im.*, 21(2):89–99, February 2002.

[23] I. A. Elbakri and J. A. Fessler. Efficient and accurate likelihood for iterative image reconstruction in x-ray computed tomography. In *Proc. SPIE 4032, Medical Imaging 2003: Image Proc.*, 2003.

[24] I. A. Elbakri and J. A. Fessler. Segmentation-free statistical image reconstruction for polyenergetic X-ray computed tomography with experimental validation. *Physics in Medicine and Biology*, 2003. submitted.

[25] H. Erdoğan and J. A. Fessler. Monotonic algorithms for transmission tomography. *IEEE Transactions on Medical Imaging*, 18(9):801–814, September 1999.

[26] H. Erdoğan and J. A. Fessler. Ordered subsets algorithms for transmission tomography. *Physics in Medicine and Biology*, 44:2835–2851, 1999.

[27] L. A. Feldkamp, L.C. Davis, and J. W. Kress. Practical cone-beam algorithm. *Journal of the Optical Society of America*, 1(6):612–619, June 1984.

[28] W. Feller. *An Introduction to Probability Theory and Its Applications*. Wiley, New York, 1968. p. 190, 245.

[29] J. A. Fessler. ASPIRE 3.0 user's guide: A sparse iterative reconstruction library. Technical Report 293, Comm. and Sign. Proc. Lab., Dept. of EECS, Univ. of Michigan, Ann Arbor, MI, 48109-2122, July 1995. Available from `http://www.eecs.umich.edu/~fessler`.

[30] J. A. Fessler. Hybrid poisson/polynomial objective functions for tomographic image reconstruction from transmission scans. *IEEE Transactions on Image Processing*, 4(10):1439–1449, October 1995.

[31] J. A. Fessler. Statistical image reconstruction methods for transmission tomography. In M. Sonka and J. Michael Fitzpatrick, editors, *Handbook of Medical Imaging, Volume 2. Medical Image Processing and Analysis*, pages 1–70. SPIE, Bellingham, 2000.

[32] J. A. Fessler. Statistical image reconstruction methods for transmission tomography. In M. Sonka and J. Michael Fitzpatrick, editors, *Handbook of Medical Imaging, Volume 2. Medical Image Processing and Analysis*, pages 1–70. SPIE, Bellingham, 2000.

[33] J. A. Fessler, I. A. Elbakri, P. Sukovic, and N. H. Clinthorne. Maximum-likelihood dual-energy tomographic image reconstruction. In *Proc. SPIE 4684, Medical Imaging 2002: Image Proc.*, 2002.

[34] J. A. Fessler, E. P. Ficaro, N. H. Clinthorne, and K. Lange. Grouped-coordinate ascent algorithms for penalized likelihood transmission image reconstruction. *IEEE Transactions on Medical Imaging*, 16(2):166–175, April 1997.

[35] J. A. Fessler and W. L. Rogers. Spatial resolution properties of penalized-likelihood image reconstruction methods: Space-invariant tomographs. *IEEE Tr. Im. Proc.*, 5(9):1346–58, September 1996.

[36] A. Ginzburg and C. E. Dick. Image information transfer properties of x-ray intensifying screens in the energy range form 17 to 320 KeV. *Med. Phys.*, 20(4):1013–1021, July 1993.

[37] G. H. Glover. Compton scatter effects in CT reconstructions. *Med. Phys.*, 9(6):860–867, Nov/Dec 1982.

[38] G. T. Gullberg, C. R. Crawford, and B. M. W. Tsui. Reconstruction algorithm for fan beam with a displaced center-of-rotation. *IEEE Tr. Med. Im.*, 5(1):23–9, March 1986.

[39] C. Helstrom. Approximate evaluation of detection probabilities in radar and optical communications. *IEEE Tr. Aero. Elec. Sys.*, 14(4):630–40, July 1978.

[40] G. Hounsfield. A method of apparatus for examination of a body by radiation such as x-ray or gamma radiation, 1972. US Patent 1283915. British patent 1283915, London. Issued to EMI Ltd. Filed Aug. 1968.

[41] J. Hsieh. Adaptive streak artifact reduction in computed tomography resulting from excessive x-ray photon noise. *Med. Phys.*, 25(11):2139–47, November 1998.

[42] J. Hsieh. Nonlinear partial volume artifact correction in helical ct. *IEEE Trans. Nuclear Science*, 46(3):743–747, June 1999.

[43] J. Hsieh, O. E. Gurmen, and K. F. King. Investigation of a solid-state detector for advanced computed tomography. *IEEE Tr. Med. Im.*, 19(9):930–40, September 2000.

[44] J. Hsieh, R. C. Molthen, C. A. Dawson, and R. H. Johnson. An iterative approach to the beam hardening correction in cone beam CT. *Medical Physics*, 27(1):23–29, January 2000.

[45] Jiang Hsieh. *Computed Tomograph: Principles, Design, Artifacts and Recent Advances*. SPIE Press, Bellingham, 2003.

[46] J. H. Hubbell and S. M. Seltzer. Tables of x-ray mass attenuation coefficients and mass energy-absorption coefficients. http://physics.nist.gov/PhysRefData/XrayMassCoef/.

[47] H. M. Hudson and R. S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Transactions on Medical Imaging*, 13(4):601–609, December 1994.

[48] P. M. Joseph and C. Ruth. A method for simultaneous correction of spectrum hardening artifacts in CT images containing both bone and iodine. *Medical Physics*, 24(10):1629–1634, October 1997.

[49] P. M. Joseph and R. D. Spital. A method for correcting bone induced artifacts in computed tomography scanners. *Journal of Computer Assisted Tomography*, 2:100–108, January 1978.

[50] P. M. Joseph and R. D. Spital. The exponential edge-gradient effect in x-ray computed tomography. *Phys. Med. Biol.*, 26(3):473–487, 1981.

[51] P. M. Joseph and R. D. Spital. The effects of scatter in x-ray computed tomography. *Med. Phys.*, 9(4):464–72, July 1982.

[52] A. C. Kak and Malcolm Slaney. *Principles of Computerized Tomographic Imaging.* IEEE Press, New York, 1988.

[53] C. Kamphius and F. J. Beekman. Accelerated iterative transmission CT reconstruction using an ordered subsets convex algorithm. *IEEE Transactions on Medical Imaging*, 17(6):1101–1105, December 1998.

[54] K. Lange, M. Bahn, and R. Little. A theoretical study of some maximum likelihood algorithms for emission and transmission tomography. *IEEE Transactions on Medical Imaging*, MI-6(2):106–114, June 1987.

[55] K. Lange and R. Carson. EM reconstruction algorithms for emission and transmission tomography. *Journal of Computer Assisted Tomography*, 8(2):306–316, April 1984.

[56] A. Macovski. *Medical Imaging Systems.* Prentice Hall, Englewood Cliffs, 1983.

[57] S. H. Manglos, G. M. Gange, A. Krol, F. D. Thomas, and R Narayanaswamy. Transmission maximum-likelihood reconstruction with ordered subsets for cone beam CT. *Physics in Medicine and Biology*, 40:1225–1241, 1995.

[58] W. D. McDavid, R. G. Waggener, W. H. Payne, and M. J. Dennis. Correction for spectral artifacts in cross-sectional reconstruction from x rays. *Med. Phys.*, 4(1):54–7, January 1977.

[59] J. M. Meagher, C. D. Mote, and H. B. Skinner. CT image correction for beam hardening using simulated projection data. *IEEE Transactions on Nuclear Science*, 37(4):1520–1524, August 1990.

[60] O. Nalcioglu and R. Y. Lou. Post-reconstruction method for beam hardening in computerized tomography. *Physics in Medicine and Biology*, 24(2):330–340, 1979.

[61] F. Noo, M. Defrise, and R. Clackdoyle. Single-slice rebinning method for helical cone-beam CT. *Physics in Medicine and Biology*, 44:561–570, 1999.

[62] J. Nuyts, B. De Man, P. Dupont, M. Defrise, P. Suetens, and L. Mortelmans. Iterative reconstruction for helical CT: a simulation study. *Physics in Medicine and Biology*, 43:729–737, 1998.

[63] A. V. Oppenheim and R. W.. Schafer. *Discrete-time signal processing.* Prentice-Hall, New York, 1989.

[64] G. U. Rao, I. Yaghmai, A. O. Wist, and G. Arora. Systematic errors in bone-mineral measurements by quantitative computed tomography. *Med. Phys.*, 14(1):62–9, January 1987.

[65] D. D. Robertson and H. K. Huang. Quantitative bone measurements using x-ray computed tomography with second-order correction. *Med. Phys.*, 13(4):474–9, July 1986.

[66] C. Ruth and P. M. Joseph. Estimation of a photon energy spectrum for a computed tomography scanner. *Medical Physics*, 24(5):695–702, May 1997.

[67] D. L. Snyder, C. W. Helstrom, A. D. Lanterman, M. Faisal, and R. L. White. Compensation for read-out noise in CCD images. *J. Opt. Soc. Am. A.*, 12(2):273–83, February 1995.

[68] J. W. Stayman and J. A. Fessler. Regularization for uniform spatial resolution properties in penalized-likelihood image reconstruction. *IEEE Tr. Med. Im.*, 19(6):601–15, June 2000.

[69] J. P. Stonestrom, R. E. Alvarez, and A. Macovski. A framework for spectral artifact corrections in X-ray CT. *IEEE Tr. Biomed. Engin.*, 28(2):128–41, February 1981.

[70] P. Sukovic and N. H. Clinthorne. Design of an experimental system for dual energy X-ray CT. In *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, volume 2, pages 1021–2, 1999.

[71] P. Sukovic and N. H. Clinthorne. Penalized weighted least-squares image reconstruction in single and dual energy X-ray computed tomography. *IEEE Trans. on Med. Imaging*, 19(11):1075–81, November 2000.

[72] B. M. W. Tsui, J. A. Terry, and G. T. Gullberg. Evaluation of cardiac cone-beam SPECT using observer performance experiments and ROC analysis. *Investigative Radiology*, 28(12), 1993.

[73] Ge Wang, M. W. Vannier, and P. C. Cheng. Iterative x-ray cone beam tomography for metal artifact reduction and local region reconstruction. *Microscopy and Microanalysis*, 5:58–65, 1999.

[74] B. R. Whiting. Signal statistics of x-ray computed tomography. In *Proc. SPIE 4682, Medical Imaging 2002: Med. Phys.*, 2002.

[75] B. R. Whiting, L. J. Montagnino, and D. G. Politte. Modeling X-ray computed tomography sinograms, 2001. submitted to Med. Phys.

[76] J. F. Williamson, B. R. Whiting, J. Benac, R. J. Murphy, G. J. Blaine, J. A. O'Sullivan, D. G. Politte, and D. L. Snyder. Prospects for quantitative computed tomography imaging in the presence of foreign metal bodies using statistical image reconstruction. *Med. Phys.*, 29(10):2404–18, October 2002.

[77] H. Q. Woodard and D. R. White. The composition of body tissues. *The British Journal of Radiology*, 59:1209–1218, December 1986.

[78] C. H. Yan, R. T. Whalen, G. S. Beaupré, S. Y. Yen, and S. Napel. Modeling of polychromatic attenuation using computed tomography reconstructed images. *Medical Physics*, 26(4):631–642, April 1999.

[79] C. H. Yan, R. T. Whalen, G. S. Beaupré, S. Y. Yen, and S. Napel. Reconstruction algorithm for polychromatic CT imaging: application to beam hardening correction. *IEEE Transactions on Medical Imaging*, 19(1):1–11, January 2000.

[80] M. Yavuz and J. A. Fessler. Statistical image reconstruction methods for randoms-precorrected PET scans. *Med. Im. Anal.*, 2(4):369–378, 1998.

[81] S. W. Young, H. H. Muller, and W. H. Marshall. Computed tomography: beam hardening and environmental density artifacts. *Radiology*, 148:279–283, July 1983.

[82] D. F. Yu, J. A. Fessler, and E. P. Ficaro. Maximum likelihood transmission image reconstruction for overlapping transmission beams. *IEEE Tr. Med. Im.*, 19(11):1094–1105, November 2000.

ABSTRACT

X-ray Computed Tomography Statistical Iterative Reconstruction

by

Idris A. Elbakr

Chair: Jeffrey A. Fessler

Statistical reconstruction for transmission tomography is emerging as potential alternative to conventional analytic image reconstruction. To fully realize their potential in noise reduction and image quality improvement, statistical algorithms should be based upon a system model that incorporates measurement statistics, attenuation physics and system parameters.

CT measurements are often assumed to follow Poisson statistics. CT detectors, however, are energy integrators that give rise to more complex compound Poisson statistics. We derive the compound Poisson probability mass function and a practical but approximate likelihood. The likelihood is based on a statistical model that accounts for energy-dependent statistics, Poisson scintillation light and electronic additive Gaussian noise. We compare the approximate likelihood with the ordinary Poisson and exact likelihoods. The approximate likelihood is more accurate than

the ordinary Poisson likelihood in low count situations, and may be useful for image reconstruction in such situations.

We derive a polyenergetic statistical X-ray CT reconstruction algorithm. The algorithm is based on polyenergetic X-ray attenuation physics and has been derived for objects containing an arbitrary number of materials. The algorithm derivation involves successive application of the optimization transfer principle to arrive at a simple and easy to maximize cost function. The algorithm requires knowledge of the X-ray spectrum or related measurements and a pre-segmented map of the distributions of different tissues within the image. Such a map is available from FBP reconstruction. The pre-segmentation map keeps the number of unknowns in the reconstruction problem equal to the number of pixels. In this regard the algorithm is comparable to conventional beam hardening correction methods. The algorithm is a gradient descent algorithm that can be accelerated using ordered subsets and a precomputed curvature. It is also possible to derive a curvature that guarantees monotonicity. We use the algorithm to reconstruct objects that contain materials that can be categorized as bone and (water-like) soft tissue. The iterative algorithm is superior to conventional beam hardening reduction methods in terms of artifact suppression and noise reduction.

To relax the requirement for a pre-segmentation map, we propose object models that parameterize the scanned object in terms of spatial and energy components. The object models keep the number of unknowns equal to the number of pixels, which is necessary if one does not wish to rely on multiple-energy scans. The models are based on the attenuation properties of tissues, and allow pixels to contain mixtures of tissues. This is accomplished by restricting the tissue fractions at each pixel to be (known) functions of the pixel density. We develop models (for two base materials)

suitable for distinct anatomical structures as well as for objects better characterized as mineral solutions in water. The segmentation-free iterative algorithms perform better than the FBP pre-segmented iterative algorithm and conventional beam hardening correction methods. Moreover, the segmentation-free algorithm is not hyper sensitive to mismatch between the model spectrum and the actual tube spectrum.

We also develop a system model for the GE LightSpeed CT scanner (General Electric Medical Systems, Milwaukee, WI) by examining spurious effects for which CT measurements are typically corrected. The system model accounts for first order polyenergetic effects, X-ray tube off-focal radiation and detector afterglow. A penalized likelihood algorithm based on the system model gives promising noise performance results, when compared to FBP. Image reconstruction of a large asymmetric object produces a shading artifact, that may be due to a mismatch in the measurements forward model. Eliminating the shading artifact and improving the forward model are important future work topics.

The performance of the proposed algorithms are examined with simulation and real data. The polyenergetic statistical algorithm is effective in suppressing beam hardening artifacts. Its performance is shown to be comparable to a statistical algorithm based on an idealized oracle segmentation. The algorithm with an object model for mineral solutions is promising for quantitative applications, and is shown to estimate mineral solution density values within 1%.