THE UNIVERSITY OF CHICAGO

SYSTEM OPTIMIZATION AND IMAGE RECONSTRUCTION FOR SPECTRAL COMPUTED TOMOGRAPHY

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

COMMITTEE ON MEDICAL PHYSICS

BY
DAVID S. RIGIE

CHICAGO, ILLINOIS
JUNE 2016

“When I write, I feel like an armless, legless man with a crayon in his mouth..”

— Kurt Vonnegut

Table of Contents

List of Figures

List of Tables

Acknowledgments

Without my advisor, Patrick La Rivière, this entire dissertation would be full of grammatical errors and punctuation mistakes — it would all just be incoherent babbling, probably. I mean, this is the one section that he didnt revise, if that gives you an idea. Patrick is the kind of guy who has an inexplicable thirst for knowledge of all things: cancer biology, Chicago architecture, shipping containers, How the States Got Their Shapes, home heating and cooling systems, to name a few. Its easy to imagine that he might have gone on to become a field expert in something else entirely, but I am glad that he didnt. His patience, enthusiasm and good attitude have set an example that I will always try to live up to. I cant imagine a better mentor.

I am also grateful for committee members Dr. Xiaochuan Pan, Dr. Emil Sidky, and Dr. Chien-Min Kao. They hold themselves and others to a very high scientific standard, and their mere presence has made me a better critical thinker. Our conversations over the years have molded my perspective as a researcher.

I have a lot of colleagues to thank. We've had many fun adventures, and without your friendship, I certainly would have gone insane. Adrian, I still think about that time we pushed through crowds of people, ducked under security ropes, and shook President Obama's hand at the rally on the Midway. I think there's probably a life lesson embedded in that story — something about determination and overcoming obstacles — either way, it was pretty cool.

Molly, its been three years — no, I mean EXACTLY three years as I type this sentence — Happy Anniversary! Based on the success of our relationship, I take it I was a pretty

good TA. Seriously, we have been through a lot together; you have seen every side of me, and you are still here with unwavering love and support. I dont know yet what the future will bring, but as long as we're in it together, I know I can count on years of happiness. I am also very proud of all of the hard work and dedication you have put into getting your own Ph.D.

Finally, I owe a great deal to my parents. In an irrational act of love, they emptied their bank accounts to pay for my Cornell education, which I convinced them was more valuable than a free education at Rutgers. I have no idea if this turned out to be true, but I am grateful for their sacrifice. They put my ambitions ahead of everything else and remain a source of continual love and support.

Abstract

Spectral computed tomography (CT) refers to a variety of imaging techniques, whereby projection data are recorded at multiple energies. This extra spectral information can be leveraged to classify and quantify materials based on their chemical properties and may improve dose efficiency for many clinical protocols. A number of potential applications have been demonstrated and FDA approved; however, widespread clinical adoption has not yet been achieved. We believe that this is due, in part, to the fact that current spectral CT systems are not achieving their full potential, which requires refinement to every step of the image formation process. In this work we propose several potential improvements to different parts of the spectral CT imaging chain, including data acquisition, pre-processing, and image reconstruction. The broad goal is to improve dose efficiency and overall image quality. The proposed methods are applicable to any spectral CT system, including commercially available dual-energy scanners as well as multi-energy photon-counting systems, which are being actively researched.

First, we consider a task-based framework that can be used to determine optimal acquisition parameters and to guide hardware design. The proposed metric provides a mechanism for quantifying and predicting the relationship between imaging parameters and material classification performance. Our simulation studies support the fact that this rapidly computable metric accurately predicts optimal system configurations. After the projection data are acquired, an important prereconstruction step called "basis-material decomposition" is performed. We present a projection-space basis-material decomposition method that makes

near-optimal use of the spectral information, while requiring only a few iterations. This should make it possible to compute accurate basis-material sinograms in clinically reasonable amounts of time. Lastly, spectral CT involves reconstructing multiple data channels, which could correspond to different energies or basis materials. Since these channels all represent the same basic anatomy, there are likely to be strong inter-channel correlations and shared edges. We propose a technique for jointly reconstructing multi-channel images by generalizing the popular total variation (TV) penalty to vector functions. Our results indicate that this approach has benefits over sequentially processing each energy channel, which is how commercially available dual-energy scanners work, today. We believe the developed techniques may also be applicable to problems outside of CT, such as PET/MRI and other situations that involve reconstructing multiple images with shared anatomical features.

1

Introduction

"It is possible that this technique may open up a new chapter in X-ray diagnosis."

– Sir Godfrey N. Hounsfield, *1973 [2]*

The history of spectral computed tomography is nearly as old as computed tomography (CT) itself, dating back to Hounsfield's earliest published experiments in 1973 [2]. An astoundingly complete theoretical description was provided by Alvarez and Macovski in 1976 [3], which was followed by a flurry of publications and even a clinical product from Siemens in the mid 1980s [4]. However, due to the technical limitations of the time and a lack of compelling applications, interest in spectral CT waned and development stagnated for a few decades [5].

Spectral CT has experienced a huge revival in the past ten years. In this chapter, we discuss the conditions that caused this initial failure as well as its current resurgence and why, we believe, widespread clinical adoption is now inevitable. Then, we will discuss some of the key remaining challenges and the approaches that we have taken toward solving them.

1.1 History of X-ray Computed Tomography

The Advent of CT

X-rays have played a large role in medical imaging since almost immediately after their discovery by physicist Wilhelm Roentgen in the late nineteenth century [6]. A kind of electromagnetic radiation, x-rays are able to penetrate deeply into human tissues, making it possible to form high-resolution images of internal structures, non-invasively. Fewer x-rays are able to pass through dense tissues, such as bone, resulting in less exposure to the film (or digital detector). Hence, a shadow-like 2D image is formed, and the gray values are determined by all of the various structures that the x-rays pass through.

In the 1960s, Hounsfield realized that by acquiring x-ray projections at many different angles around an object, it was possible to solve for a 3D map of the x-ray attenuation properties with a computer program [7]. This technique came to be known as "computed tomography." By comparison, conventional radiographs inherently form 2D images, projecting a 3D scene into a single plane causing a superposition of structures in depth that can cause confusion to the radiologist. CT allows for patient anatomy to be viewed as a series of thin cross-sectional images, eliminating superposition problems, and greatly improving sensitivity [2]. It is commonly accepted that radiographs are only sensitive enough to be able to appreciate about five general levels of contrast: air/lung, fat, soft-tissue, bone, and metal/implants [8]. Due to the cross-sectional nature of CT images, minute differences in tissue densities can be resolved, making it possible to clearly distinguish different organs and pathologies such as cysts and tumors [9, 5].

These developments led to Hounsfield receiving the Nobel Prize for Physiology or Medicine in 1979, which he shared with Allan McLeod Cormack [9, 10]. By this time, CT scanners had been installed in about 1000 hospitals worldwide with more than 10,000 scanners in use [9, 5]. Today, CT is used in a huge variety of clinical applications. For example, it plays a role in the evaluation of cancer [11], cardiovascular disease [12], trauma [13], and musculoskeletal

disorders [14], as well as in guiding surgical procedures [15] and radiotherapy [16].

Early Efforts in Dual-Energy Imaging

The physical quantity that CT aims to measure is called the "linear attenuation coefficient," commonly denoted by μ. The μ value for a given material is determined, primarily, by two factors: its density and its atomic number (Z). Additionally, μ depends on the energy of the x-rays, with higher energy generally leading to lower μ. Most tissue types share roughly the same energy dependence with the exception of bone, due to its high concentration of calcium (Z=20). For higher-atomic-number substances, μ increases even more dramatically at lower energies.

Hounsfield realized early on that this principle could be used to estimate the atomic number of a substance by taking CT scans at two different energies and examining the amount of enhancement that occurs as the energy is decreased. Specifically, he writes about distinguishing iodine (Z=53) from calcium (Z=20), which may be difficult to tell apart in a single-energy image [2].

This idea of dual-energy imaging was greatly refined by Alvarez and Macovski in 1976 [3]. They pointed out that nearly all x-ray interactions occuring at clinically relevant energies (0-150 keV) are due two physical phenomena, known as Compton scattering and photoelectric absorption. The former is weakly dependent on energy, while the latter has an approximately inverse cubic relationship. They demonstrated that by imaging at two distinct energies, one could solve for the precise contribution of each of these two mechanisms to μ, which we denote by σ and τ. Thus, each pixel would be described by two values, instead of one, and, together, σ and τ constitute a virtually complete description of the energy dependence of μ. Furthermore, they proposed a pre-reconstruction technique for decomposing μ into σ and τ and provided a rigorous theoretical analysis of how their estimation would be impacted by measurement noise. As a side effect, their method inherently models the x-ray tube spectrum, thus eliminating common "beam-hardening" artifacts; this is discussed further in

Chapter 2.

In the late 70s and early 80s, there were numerous publications on dual-energy CT (DECT) speculating about different acquisition methods, algorithms, and applications. The most obvious method for acquiring dual-energy CT data was to simply perform two separate scans, changing the tube potential, i.e. the average energy, in between. However, it was quickly realized that this approach was highly susceptible to motion occuring between scans. Slight changes in attenuation caused by patient movement would be misinterpreted as being due to spectral changes and lead to incorrect results [5].

One proposed solution to this problem was to simultaneously acquire dual-energy projection data using a "dual-crystal" detector [17]. The idea was that since an x-ray tube spectrum is naturally polychromatic, one could try to separate out low- and high- energy components on the detection side. Typical designs consisted of two phosphor layers: a thin front layer that preferentially recorded lower-energy x-rays and a thicker rear layer that detected the remaining higher-energy x-rays. It is conceptually similar to how color filter arrays (CFA's) allow color CCD's to parse white light into red, green, and blue components. However, filtering x-rays is more challenging, and this approach suffered because the energy separation between the detector layers was inadequate.

Yet another proposed solution is to rapidly change the tube potential, thus altering the average beam energy, every subsequent projection view. This results in two almost-registered sets of dual-energy projection data, obtained within the timeframe of a single scan. This technique, commonly called "rapid kV switching," is the only one that made it into a commercial product in the early days of dual-energy CT. It was offered by Siemens in the mid 1980s [4, 5]. One considerable challenge for this design is achieving adequate x-ray flux at lower energies. For example, dropping the tube potential from 140 to 80 kVp, results in a corresponding 3× decrease in flux. Unfortunately, it is not physically possible to increase the tube current fast enough to compensate for this [5].

In the late 1980s, breakthroughs in helical scanning shifted the focus of the industry

almost entirely toward increasing imaging speed. Rapid kV switching was somewhat incompatible with these goals. Imaging times decreased by orders of magnitude, and it was no longer possible to acquire an adequate number of projections at both energies [5]. Furthermore, one of the primary applications of Siemens' dual-energy scanner was bone mineral density analysis; however cheaper technologies, such as DXA (dual-emission x-ray absorptiometry), became available, limiting the need for DECT [18]. Due to these factors, the technology was discontinued. Then, in the late 1990s, attention shifted toward "multidetector CT" (MDCT), and development was largely focused on adding more detector rows and increasing rotation speeds [5]. By this time, interest in dual-energy CT had almost completely dissipated.

1.2 Modern Spectral CT Scanners

The Revival of DECT

About ten years ago, interest in dual-energy CT was renewed by the introduction of Siemens' dual-source CT (DSCT) scanner [19]. The primary motivation for the DSCT scanner appears to be increasing temporal resolution for cardiac examinations, such as coronary CT angiography (CTA) [19, 20]. At that time, state-of-the-art MDCT scanners could achieve maximum rotation speeds of about 3 revolutions per second (rev/s). It was hypothesized that scan speeds exceeding 5 rev/s would enable robust clinical performance over a wide range of heart rates without the need for heart rate control or multi-segment reconstruction; however, this would subject gantry components to mechanical forces exceeding 75G, which was technically infeasible at the time [19]. Instead, Siemens built the DSCT scanner with two x-ray tubes, offset by 90 degrees, each facing its own detector array. Thus, a complete set of projection data could be acquired in approximately a quarter of a turn, cutting the scan time in half.

In addition, this DSCT scanner also represented an interesting solution to dual-energy

CT, solving several of the limitations of earlier designs. For example, since each tube operates independently, there is complete freedom to adjust the kVp and mAs to ensure adequate flux at both energies. Because of the simultaneous operation of both sources and the high rotation speeds, dual-energy data are acquired nearly simultaneously, minimizing the possibility for inconsistencies between low- and high-energy images caused by patient motion. Furthermore, additional filtration can be used on the high-kVp source to increase the spectral separation between datasets, which can significantly improve performance [17, 21].

It was soon realized that this DSCT scanner could be particularly useful for scans enhanced by iodinated contrast agents. Dual-energy data could be acquired quickly enough to freeze movement and changes in contrast enhancement, and dual-energy processing methods could reliably isolate iodine from bone and soft-tissue, due to its high atomic number (Z=53). Conventionally, separate scans are often needed before and after the injection of contrast media. However, with DECT, a single scan could produce a contrast-enhanced image, an iodine image, and a virtual non-contrast (VNC) image, all perfectly co-registered. In principle, this could improve visualization, while cutting the radiation dose to the patient in half [22].

Over the last several years, dual-energy CT has been applied to a variety of clinical problems with varying degrees of success, including the identification of gout [23, 24], evaluation of lung perfusion defects [25], visualization of tendons and ligaments [26], and differentiation of iodine from calcified plaques [27]. Since then, GE and Philips have made their own dual-energy CT scanners commercially available [28, 29]. Interestingly, both have resorted to designs which were discussed earlier that were not previously successful. Philips' design is based on the dual-crystal detector, while GE has resurrected the rapid-kV switching paradigm, overcoming some of its previous limitations thanks to advances in detector and x-ray tube/generator technology. Presently, every major vendor of diagnostic CT scanners offers at least one solution for dual-energy imaging. These concepts are discussed in more detail, with an overview of their relative merits in Chapter 2.

Spectral CT

In addition to the recent commercial availability of DECT diagnostic scanners, there has also been significant progress in the development of direct-conversion detectors with energy-discriminating capabilities [30]. Conventional energy-integrating detectors (EID's) consist of a scintillating crystal that converts x-rays into visible light, which is subsequently recorded by a photodiode, ultimately measuring the total energy deposited in the crystal layer. The contributions from x-rays spanning a continuous range of different energies are collapsed into a single measurement and all spectral information is lost.

This new type of detector has a single semiconductor layer that is capable of measuring individual x-ray photon interactions, directly converting them into charge pulses. If the pulse is larger than a predefined threshold (measured by a comparator) it is registered as an actual event; otherwise, it is rejected as noise. The detector registers the number of photons impinging on it, hence why they are called "photon counting detectors" (PCD's). Furthermore, the charge pulse is proportional to the energy of the photon, so multiple thresholds can be used to bin all detected x-rays into different energy windows. Thus, the PCD can leverage the natural polychromaticity of the x-ray source to obtain multi-energy CT data. Hence, the term "spectral CT" has become popularized to refer broadly to the aforementioned DECT techniques as well as multi-energy photon-counting CT (PCCT).

Though two energies are generally sufficient for most clinical scenarios, since μ is dominated by only two interaction types (σ and τ), this two-parameter model does not account for "k-edges." The term k-edge refers to an abrupt increase in $\mu(E)$, which is unique to each element. For most clinically relevant materials, the k-edge occurs at a low energy and can be safely ignored; however, it would allow certain high-Z metals to be detected and quantified with remarkable sensitivity using PCD's [31]. This has lead to interest in developing new contrast agents using materials such as gold and gadolinium [30]. This is discussed further in Chapter 2. However, even without the introduction of k-edge materials, having more than two energy bins can improve the accuracy of decomposition methods, and allows for

thresholds to be optimized for each task [32, 33].

Besides being a very attractive solution for spectral CT, PCD's have several other key advantages, which will accelerate their adoption into clinical scanners. Firstly, they can eliminate virtually all electronic noise by appropriately configuring the lowest energy threshold [34]. Secondly, their geometric efficiency is nearly 100%, whereas conventional EID's are closer to 80%, due to inactive regions between scintillator elements [35]. Because of these architectural differences, PCD's can also be manufactured to much smaller pixel sizes [5], which may be beneficial for specialized applications requiring very high spatial resolution (e.g. breast CT). Lastly, counting photons, as opposed to measuring energy deposition, inherently results in more efficient use of the x-ray information and could lead to higher contrast-to-noise ratios for similar dose levels [36, 37]. Therefore, even if spectral information is not needed, PCD's can help to achieve significant improvements in dose efficiency and spatial resolution.

Despite all of these benefits, PCD's have only been used in prototype CT scanners so far, due to a variety of technical challenges and non-idealities. Still, all of the major vendors recognize the importance of the technology and are working on prototypes. Siemens has recently begun testing their PCCT system on human patients [38], so it is likely that a product will become commercially available in the next few years. Unlike in the 1970s and 1980s, vendors are more invested in dual- and multi-energy scanners now, and more recent designs address some of the early limitations. Furthermore, PCCT and DSCT can justify their existence for reasons beyond spectral imaging (i.e dose efficiency and fast cardiac imaging), and, for these reasons, we believe will continue to gain traction in the clinic.

1.3 Current Challenges

Many of the remaining challenges in spectral CT revolve around refining the data acquisition, image reconstruction, and postprocessing steps so as to make best use of the spectral

information.

Multi-energy imaging protocols are inherently more complicated than conventional ones because of the larger number of parameters that can be selected. For example, dual-source scanners allow for the independent selection of the kVp (i.e. average energy), mAs, and filtration of each x-ray tube. A photon-counting CT system with 6 adjustable energy thresholds could realize millions of different configurations. The optimal settings for all of these parameters will depend entirely on the patient and imaging task. Similarly, different vendors have taken different approaches to dual-energy CT, e.g. dual-source, rapid kV switching, dual-crystal, and understanding the relative merits of these designs, too, will depend upon the imaging task. Therefore, it will be important to establish task-based metrics for predicting the performance of various spectral imaging configurations. Preferably, they should be efficient to compute so that large parameter spaces can be rapidly explored and imaging protocols could be adapted, situationally.

The next step of the imaging chain that can be optimized is the decomposition of the projection data into either physical interaction mechanisms, τ and σ, or into basis materials. This important step is what enables the discrimination and quantification of materials based on their atomic composition; however, it is well documented that it is ill-conditioned and tends to amplify noise [39]. Energy-discriminating PCD's typically acquire more energy channels than are required to perform this decomposition; however, appropriate use of this redundant spectral information can actually improve the estimation of τ and σ, combatting this noise amplification [32]. Therefore, it would be desireable, especially for PCCT systems, to have statistical estimation methods that optimally utilize the spectral information during the decomposition, that are also computationally efficient.

Finally, it will be important to develop multi-channel image processing schemes for reconstruction and postprocessing. The algorithms implemented on current dual-energy scanners are largely designed around fitting into legacy image reconstruction pipelines, which were designed for single-energy CT. Thus, each energy channel is typically processed one at a time.

It is likely that noise suppression and overall image quality could be improved by formulating a model for jointly reconstructing (or restoring) multi-energy CT data and leveraging the inherent similarities and anatomic correlations across channels. This is conceptually similar to observations that have been made about color images; noise removal algorithms that consider all RGB (red, green, blue) channels simultaneously tend to be more robust than ones that treat them separately [40, 41]. Furthermore, a joint model could reduce the number of parameters in the reconstruction, since each channel wouldn't require individual tuning [42].

1.4 Organization of Thesis

In this thesis, we address each of the aforementioned topics, providing potential improvements to different parts of the spectral CT imaging chain. We begin with an overview of spectral imaging, including an expanded discussion of current implementations and clinical applications. In Chapter 3, we present a formalism for optimizing protocol and hardware design, by extending the classical signal detection theory to material classification tasks. This yields a rapidly computable metric that can be used to evaluate spectral imaging parameters for different material classification tasks. In Chapter 4, we present a framework for projection-space, material decomposition that achieves similar accuracy to the gold-standard, maximum-likelihood estimation approach while requiring only a few iterations. Then, in Chapter 5, we propose a technique for jointly reconstructing multi-energy CT data by generalizing the popular total variation (TV) regularizer to multi-channel images. This "vectorial" TV is formulated to leverage the anatomical correlations and common edge structure across energy or material channels. We further validate the utility of this VTV penalty in Chapter 6 on real scanner data by demonstrating its use for post-reconstruction, denoising. Finally, Chapter 7 concludes with a summary and discussion of potential future directions.

2

Review of Spectral CT Imaging

2.1 Physics of X-ray Attenuation

X-rays are a powerful tool for imaging the internal structures of objects due to their penetrative nature. In particular, human tissues are highly radiotranslucent, making it possible to peer deep into the human body. Though a significant fraction of impinging x-rays will pass completely through, many will be scattered or absorbed — the particular number depending upon the properties of the tissue. These interactions that remove x-ray photons from the primary beam are responsible both for creating image contrast and depositing radiation dose into the patient (Therefore, we cannot have x-ray absorption imaging without radiation dose). In this section, we will review the physical interactions pertinent to diagnostic x-rays, which range in energy from about 0 to 150 keV.

Coherent scattering, also known as Rayleigh scattering, occurs when an x-ray interacts with a whole atom without imparting or giving up any of its kinetic energy. We will not comment much on Rayleigh scattering because it contributes very little to the overall x-ray attenuation and does not result in any dose deposition. However, we note that it does occur in all kinds of x-ray imaging, especially at lower energies.

Incoherent (or Compton) scattering occurs when the incident x-ray inelastically scatters

with an inner-shell electron. Since energy is imparted to an atomic electron, radiation dose is imparted to the patient. Compared with photoelectric absorption, the energy dependence is relatively flat. The precise formula (ignoring constant scale factors) is given by the Klein-Nishina function [3]:

$$\psi_{\mathrm{kn}}(\epsilon) = \frac{1+\epsilon}{\epsilon^2}\left[\frac{2(1+\epsilon)}{1+2\epsilon} - \frac{1}{\epsilon}\ln(1+2\epsilon)\right] + \frac{1}{2\epsilon}\ln(1+2\epsilon) - \frac{1+3\epsilon}{(1+2\epsilon)^2}, \tag{2.1}$$

where $\epsilon \equiv E/m_e c^2$.

Photoelectric absorption is an interaction whereby an atomic electron is dislodged by an impinging x-ray photon if it exceeds a certain threshold energy. The vacancy left by the ejected electron is filled by one of the outer shell electrons; upon making this transition, another photon may be emitted with kinetic energy precisely determined by the energy-gap size of this shell transition. The photoelectric cross section is approximately proportional to Z^3/E^3, where Z is the effective atomic number of the material, and E is the incident photon energy. For tissues it is dominant only at very low energies, but for high-Z materials (such as iodine) it can be the dominant scattering mechanism across the entire energy range of the x-ray source spectrum.

The energy dependencies of these scattering mechanisms are derived from physical models considering single electrons or atoms. However, x-ray interactions are largely determined by the atomic nuclei or inner-shell electrons, whereas the chemical bonds that hold molecules and compounds together involve the outer-shell electrons. Therefore, determining the x-ray scattering properties of bulk materials is simply a matter of averaging the contributions of constituent atoms, weighted according to their fractional abundance.

Figure 2.1 plots the attenuation coefficients for the aforementioned scattering mechanisms in water, based on the NIST Photon Cross Sections Database.

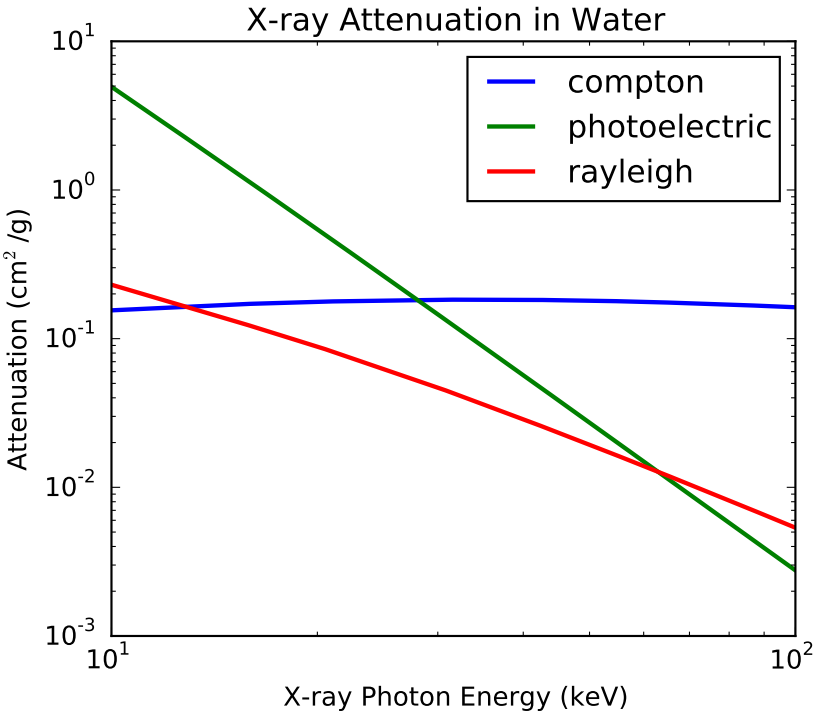


Figure 2.1: A plot of the various x-ray attenuation coefficients in water based on the NIST Photon Cross Sections Database.

2.2 Conventional Single-Energy CT

In x-ray computed tomography, the quantity we wish to reconstruct is the linear attenuation coefficient (LAC), commonly denoted by μ. It quantifies the probability that an x-ray photon will undergo an interaction per unit length of material. The particular interaction event could be any one of the ones mentioned previously: Rayleigh scattering, Compton scattering, or photoelectric absorption. It is assumed that once a photon is scattered or absorbed, it is removed from the primary beam and will not contribute to the measurement along that ray. In reality, this photon may go on to be detected by another detector element. This scattered radiation is an undesirable contribution to the transmission data and is typically corrected prior to image reconstruction.

Beer's Law

An x-ray source and detector array are placed on opposing sides of the patient in order to determine what fraction of x-rays are transmitted through the patient along a fan of different rays (see figure 2.2). If an x-ray flux of I_0 is incident on the object, along a ray L, then the transmitted flux I is given by Beer's law:

$$I = I_0 \, e^{-\int_L \mu(\boldsymbol{r}) \, d\ell}, \tag{2.2}$$

where $\mu(\boldsymbol{r})$ is the linear attenuation coefficient at location $\boldsymbol{r}$. The line integral,

$$\int_L \mu(\boldsymbol{r}) d\ell,$$

is shorthand for

$$\int_{-\infty}^{\infty} \mu\left(\boldsymbol{r}_0 + t\hat{\theta}\right) dt,$$

where $\hat{\theta}$ is a unit vector specifying a direction and $\boldsymbol{r}_0$ is an arbitrary point along the line.

Linear Tomography Problem

Next we define the quantity $g(L) \equiv -\log(I/I_0)$, which allows us to write

$$g(L) = \int_L \mu(\boldsymbol{r})\, d\ell. \tag{2.3}$$

If projections are collected along all possible lines L that intersect the object, then the problem of recovering $\mu(\boldsymbol{r})$ from $g(L)$ becomes one of inverting the x-ray transform. The element-wise logarithm of the normalized measurements, $g(L)$, is referred to as the *sinogram*. Recovering an image function from its sinogram is the fundamental problem of all tomographic imaging methods. In practice, only a finite number of such projections can be acquired, and the image is represented by a finite, discrete set of values (e.g. a pixel matrix). By defining sinogram data vector g, image vector f, and a discretized version of the x-ray transform, X, one can write down the following linear inverse problem:

$$Xf = g. \tag{2.4}$$

Many numerical algorithms exist for approximating solutions to this inverse problem. The most widely used one is filtered backprojection, which is based on a discretization of the closed-form x-ray transform inversion formula. One can also tackle (2.4) directly with iterative reconstruction algorithms, which allows greater flexibility in the data sampling requirements as well as the incorporation of prior knowledge. We will discuss an iterative reconstruction scheme that is tailored to multi-energy CT in Chapter 5.

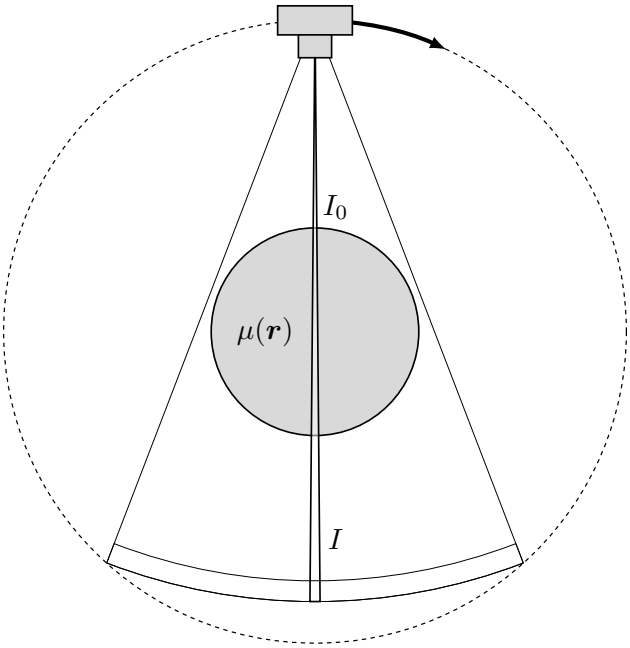


Figure 2.2: 2D schematic of CT measurement model.

Limitations of Standard Model

Note that it was necessary to assume a monochromatic beam of radiation in order to obtain a log-linear relationship between the desired image and the projection data. In practice, x-ray tubes emit x-rays across a broad energy spectrum. Since μ is an aggregation of energy-dependent scattering mechanisms, it also varies continuously with x-ray energy, thus complicating the relationship between the incident radiation and μ. Ignoring these effects causes nonlinear shading artifacts in the reconstructed images, known as *beam-hardening* artifacts. Figure 2.3 illustrates how beam-hardening causes misleading shading artifacts in this dual-energy CT phantom (Gammex Inc., Middleton, WI), consisting of a uniform, water-equivalent background and inserts made of calcium- and iodine-equivalent materials.

In all human tissues the attenuation is strongly dominated by two interaction mechanisms discussed above, Compton scattering and photoelectric absorption (Rayleigh scattering is an extremely small contribution to the measured signal), which have different dependencies on x-ray energy and atomic number of the absorber. The standard imaging model provides no way to separately probe the contribution of each of these mechanisms, since they are

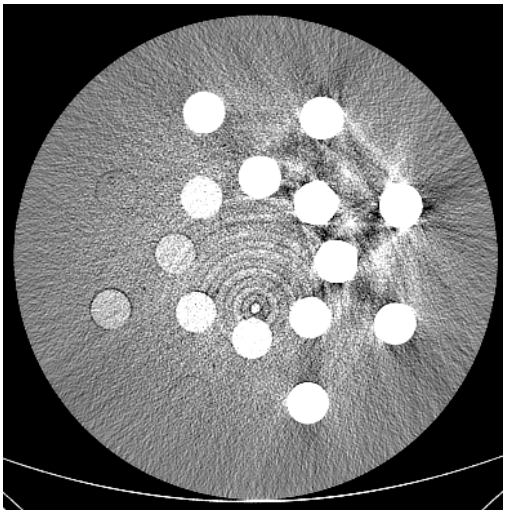


Figure 2.3: 80 kVp image of a dual-energy CT phantom (Gammex Inc., Middleton, WI) consisting of a uniform, water-equivalent background and inserts made of calcium- and iodine-equivalent materials. The light and dark streaks seen around certain inserts are due to nonlinear beam-hardening artifacts. Gray-level window: [-200, 200] HU.

entangled in the single parameter μ. This can lead to problematic ambiguities when two chemically distinct materials appear isointense in the reconstructed image. For example, iodine-based contrast agents may be indistinguishable from calcified plaques.

2.3 Dual-Energy CT

Dual-energy CT refers to a family of x-ray CT imaging methods that attempt to reconstruct the entire energy dependence of the LAC throughout the object by measuring projections at two different average beam energies and untangling the effects of Compton scattering and photoelectric absorption. This would make it possible to eliminate the aforementioned beam-hardening artifacts and to gain additional information for characterizing materials based on their atomic composition (in particular, their effective atomic number, Z_{eff}). This generally requires specialized hardware and additional image processing steps.

Firstly, the assumption of monochromatic illumination is relaxed, and instead we model the incident radiation by the spectrum $S(E)$, expressed in units of photons per unit area, per unit time, per energy (e.g. $\text{mm}^{-2} \cdot \text{s}^{-1} \cdot \text{keV}^{-1}$). Figure 2.4 depicts typical x-ray tube

spectra at 80 and 140 kVp, which clearly have different average energies. Additionally, we explicitly include the energy-dependence of the LAC in the measurement model, $\mu(\boldsymbol{r}, E)$. A measurement along a ray L is given by

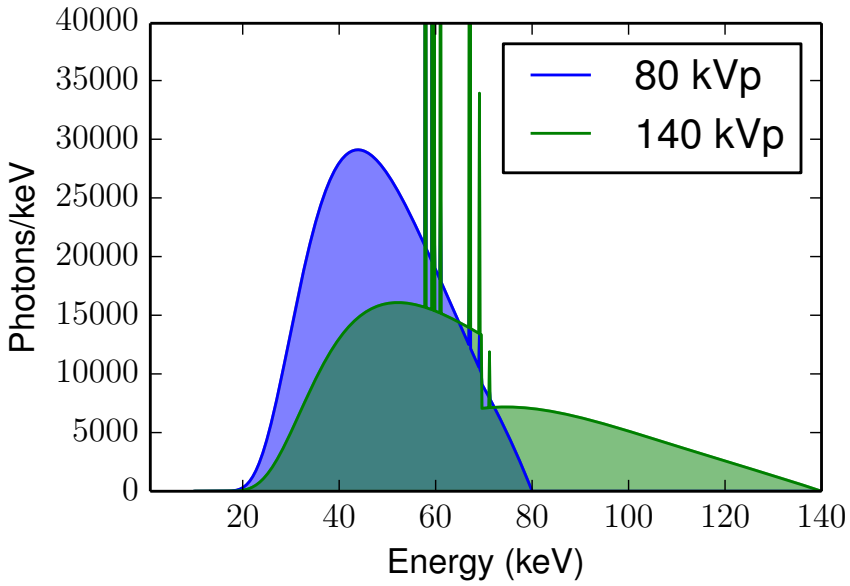


Figure 2.4: example plot of x-ray tube spectrum

$$y(L) = \int_{-\infty}^{\infty} S(E) e^{-\int_L \mu(\boldsymbol{r}, E)\, d\ell}\, dE. \tag{2.5}$$

Many detectors will also have an energy-dependent gain, but for notational simplicity, we absorb this term into $S(E)$. Because of the energy integral, we can no longer relate these transmission measurements to the x-ray transform of $\mu(\boldsymbol{r}, E)$ by normalizing and taking the logarithm.

Theoretically, if one could use monochromatic illumination, so that $S(E) = \delta(E - E_0)$, one could sweep E_0 across the desired energy range and recover $\mu(\boldsymbol{r}, E)$. This can be achieved approximately at certain synchrotron x-ray sources, but unfortunately, standalone scanners all use x-ray tubes that emit broad-spectrum radiation. Therefore, we need to employ an alternative strategy.

Parameterizing the Attenuation with Basis Materials

The first thorough, theoretical description of spectral CT comes from the highly influential work of Alvarez and Macovski [3]. Their key insight is that the energy-dependent LAC of all tissues can be accurately parameterized by two basis functions, corresponding to the energy dependencies of photoelectric absorption and Compton scatter:

$$\mu(\boldsymbol{r}, E) = c_1(\boldsymbol{r})\psi_{\text{pe}} + c_2(\boldsymbol{r})\psi_{\text{kn}}(E). \tag{2.6}$$

They used $\psi_{\text{pe}}(\boldsymbol{r}) = 1/E^3$ and ψ_{kn} as defined in (2.1). The quantities $c_1(\boldsymbol{r})$ and $c_2(\boldsymbol{r})$ are unknown coefficient maps. This expansion greatly reduces the number of unknown parameters and decouples the energy and spatial dependence of $\mu(\boldsymbol{r}, E)$.

If the LAC of any material can be represented by linear combinations of $\psi_{\text{pe}}(E)$ and $\psi_{\text{kn}}(E)$, then the converse is also true: $\psi_{\text{pe}}(E)$ and $\psi_{\text{kn}}(E)$ can each be represented by linear combinations of any two unique materials. Therefore, we could just as well use the LAC's of two *basis materials* as our expansion functions:

$$\mu(\boldsymbol{r}, E) = c_1(\boldsymbol{r})\mu_1(E) + c_2(\boldsymbol{r})\mu_2(E). \tag{2.7}$$

or similarly,

$$\mu(\boldsymbol{r}, E) = \rho_1(\boldsymbol{r})(\mu_1(E)/\rho_1) + \rho_2(\boldsymbol{r})(\mu_2(E)/\rho_2), \tag{2.8}$$

where the quantity μ/ρ is the mass-attenuation coefficient. The MAC is often used as a basis function instead of the LAC, μ, because it is the more fundamental quantity; for example, water and ice have different LAC's but the same MAC, since the MAC is density-independent. With this decomposition, the unknown spatially varying coefficients correspond to the physical density of each material, so we denote them by $\rho_1(\boldsymbol{r})$ and $\rho_2(\boldsymbol{r})$.

In more recent works, the basis-material decomposition is used more often than the physics decomposition. One advantage of the former representation is that the basis materials

can be chosen so that the coefficient maps, $\rho_j(\boldsymbol{r})$, are physically meaningful. For example, if $\psi_1(E)$ and $\psi_2(E)$ are bone and soft-tissue MAC's, then the $\rho_j(\boldsymbol{r})$ correspond to bone and tissue density maps. For the remainder of this dissertation, it is assumed that the basis functions, $\psi_1(E)$ and $\psi_2(E)$, refer to the MAC's of basis materials, unless explicitly stated otherwise.

Validity of Two-Material Decomposition

The previous argument that the attenuation can be represented as a linear combination of just two known functions of energy is ubiquitous in the literature on spectral CT; however, it is worth examining the underlying assumptions. Firstly, we have ignored Rayleigh scattering, which is a small, but non-zero, contribution to the total attenuation. Secondly, we have implicitly assumed that the energy-dependence of Compton scattering and photoelectric absorption are independent of the material (up to a constant scale factor), which is only approximately true. In particular, materials will have abrupt increases in their photoelectric absorption cross-sections corresponding the k-shell binding energies that are unique to each atom. For all materials that are naturally abundant in the human body, these so-called "k-edges" occur at very low energies and have practically no impact within the energy range that is relevant to medical applications. However, one might still wonder: is it possible to decompose the attenuation into more than two unique basis materials?

For high-Z materials that may be introduced into the patient (e.g. iodine), the answer is certainly "yes". This is due to the fact that they have k-edges at higher energies which fall within the detectable range of the x-ray tube spectrum; this is discussed further in section 2.4. However, for the lower-Z materials that comprise human tissues, the answer is less clear. Bornefalk [43] performed a thorough analysis of the intrinsic dimensionality, N, of the attenuation of all materials with $Z < 20$ within diagnostically relevant energies (25-120 keV) in order to determine a theoretical upper bound on how many materials could be distinguished and concluded that $N = 4$. However, this not necessarily mean that 4

dimensions can actually be measured in practice, due to noise and systematic uncertainties in the measurement process.

Alvarez[44] attempted to characterize the relationship between noise and dimensionality by modeling idealized photon-counting detectors and simulating two- and three-material decomposition schemes. For example he compared a bone and soft-tissue decomposition to a bone, soft-tissue, and adipose-tissue decomposition. While he was able to successfully decompose simulated data into three tissue compartments, he found that there was an enormous noise penalty (more than 10^3-fold increase in variance), regardless of the available energy resolution. In practice, the two-material model is likely to be totally sufficient in all imaging scenarios that don't involve introducing exogenous contrast materials.

Projection Space Decomposition

Once the LAC is parametrized by known basis functions, one must still solve for the unknown coefficient maps. Plugging equation 2.6 into equation 2.5 yields

$$y(a_1, a_2) = \int_{-\infty}^{\infty} S(E) e^{-a_1 \psi_1(E) - a_2 \psi_2(E)} \, dE, \tag{2.9}$$

where

$$a_{1,2}(L) \equiv \int_L \rho_{1,2}(\boldsymbol{r}) \, d\ell. \tag{2.10}$$

For a given ray L, the transmission measurement, y, is a function of the unknown parameters a_1 and a_2. In order to be able to solve for both a_1 and a_2, we need to make at least two measurements with different average energies. Suppose that every projection ray, L, is measured with spectra $S_1(E)$ and $S_2(E)$. Thus, we would have two complete sets of

transmission measurements given by

$$y_1(a_1, a_2) = \int_{-\infty}^{\infty} S_1(E) e^{-a_1\psi_1(E) - a_2\psi_2(E)}\, dE \tag{2.11}$$

$$y_2(a_1, a_2) = \int_{-\infty}^{\infty} S_2(E) e^{-a_1\psi_1(E) - a_2\psi_2(E)}\, dE. \tag{2.12}$$

Equations 2.11 and 2.12 form a system of two nonlinear equations with two unknowns, and as long as $S_1(E)$ and $S_2(E)$ are unique, we can solve for a_1 and a_2 along each ray L. We discuss some specific numerical methods in Chapter 4.

The quantities $a_1(L)$ and $a_2(L)$ can be interpreted as basis-material sinograms (see eq. 2.10). Once they are known, sinogram data can be synthesized at any desired energy via equation 2.6:

$$\int_L \mu(\boldsymbol{r}, E)\, d\ell = a_1\psi_1(E) + a_2\psi_2(E). \tag{2.13}$$

These virtual monoenergetic (VM) sinograms can then be reconstructed using standard CT reconstruction algorithms. Alternatively, one could apply the same algorithms directly to $a_1(L)$ and $a_2(L)$ to yield material density images, $\rho_1(\boldsymbol{r})$ and $\rho_2(\boldsymbol{r})$. Once these coefficient maps are known, we have complete information about the LAC at all energies and locations (see eq. 2.6). All of the same logic follows if more than two basis functions are used, and in general the number of energy measurements must be greater than or equal to the number of basis materials.

Since, this model inherently incorporates the x-ray spectrum and the energy-dependence of the x-ray attenuation, the synthesized images do not suffer from beam-hardening artifacts. Note, that we can also convert the basis-material images into Compton and photoelectric images by taking the appropriate linear combinations of $\rho_1(\boldsymbol{r})$ and $\rho_2(\boldsymbol{r})$. Thus, in addition to eliminating beam-hardening artifacts, we have disentangled the contributions of photoelectric absorption and Compton scattering to the total attenuation, which allows us to estimate Z_{eff} of different image regions. This provides a more complete description of the object and improves material differentiation. For example, calcium (Z=20) can be unambiguously

differentiated from iodine (Z=53).

Image Space Decomposition

The data-domain material decomposition method has the advantage of fully accounting for the polychromaticity of the spectrum and attenuation and eliminating beam-hardening artifacts, but it also has several limitations. Firstly, multi-energy projection data must be acquired along each ray, so the sinogram data of all energy channels need to be perfectly coregistered. Furthermore, solving the nonlinear system of equations for a_1 and a_2 requires a very accurate characterization of the incident spectra, $S_1(E)$ and $S_2(E)$. Measuring these spectra accurately can be difficult in practice, and even small errors can result in large errors in the decomposition [45].

An alternative approach that alleviates these complications is to separately reconstruct the CT data at each energy and to perform the material decomposition in the image domain. This approach is much simpler because completely standard algorithms can be used, and all of the additional spectral information is leveraged only through post-processing algorithms. The image domain (ID) approach assumes that the beam-hardening effects are either small enough that they can be ignored or pre-corrected with heuristic algorithms [46].

To demonstrate how the image domain (ID) approach works, we consider the case where we've acquired projection data with two different spectra, and assume that the polychromaticity of the beam can be ignored. We denote these spectra by

$$S_1(E) = I_1\,\delta(E - E_1) \tag{2.14}$$

$$S_2(E) = I_2\,\delta(E - E_2). \tag{2.15}$$

The delta function collapses the energy integral and yields the following transmission mea-

surements along each ray L:

$$y_1(L) = I_1 e^{-\int_L \mu(\boldsymbol{r},E_1)\,d\ell} \tag{2.16}$$

$$y_2(L) = I_2 e^{-\int_L \mu(\boldsymbol{r},E_2)\,d\ell}. \tag{2.17}$$

As with the standard CT imaging model we define $g_{1,2} \equiv -\log(y_{1,2}/I_{1,2})$ yielding dual-energy sinogram data

$$g_1(L) = \int_L \mu(\boldsymbol{r}, E_1)\,d\ell \tag{2.18}$$

$$g_2(L) = \int_L \mu(\boldsymbol{r}, E_2)\,d\ell. \tag{2.19}$$

By applying standard CT reconstruction algorithms, separately to g_1 and g_2, we can recover dual-energy images $\mu(\boldsymbol{r}, E_1)$ and $\mu(\boldsymbol{r}, E_2)$. Finally, we can use the basis-material expansion on $\mu(\boldsymbol{r}, E)$ to set up the following linear system:

$$\mu(\boldsymbol{r}, E_1) = \rho_1(\boldsymbol{r})\psi_1(E_1) + \rho_2(\boldsymbol{r})\psi_2(E_1) \tag{2.20}$$

$$\mu(\boldsymbol{r}, E_2) = \rho_1(\boldsymbol{r})\psi_1(E_2) + \rho_2(\boldsymbol{r})\psi_2(E_2). \tag{2.21}$$

This linear system of equations is separable over pixel location, $\boldsymbol{r}$, so we can decompose each pixel separately into its contributions from each basis material. It is trivial to show that

$$\rho_1(\boldsymbol{r}) = \frac{\psi_2(E_2)\mu(\boldsymbol{r}, E_1) - \psi_2(E_1)\mu(\boldsymbol{r}, E_2)}{\psi_1(E_1)\psi_2(E_2) - \psi_1(E_2)\psi_2(E_1)} \tag{2.22}$$

$$\rho_2(\boldsymbol{r}) = \frac{-\psi_1(E_2)\mu(\boldsymbol{r}, E_1) + \psi_1(E_1)\mu(\boldsymbol{r}, E_2)}{\psi_1(E_1)\psi_2(E_2) - \psi_1(E_2)\psi_2(E_1)}. \tag{2.23}$$

We can write this more compactly by defining

$$\boldsymbol{\mu}(\boldsymbol{r}) = (\mu(\boldsymbol{r}, E_1), \mu(\boldsymbol{r}, E_2))^T \tag{2.24}$$

$$\boldsymbol{\rho}(\boldsymbol{r}) = (\rho_1(\boldsymbol{r}), \rho_2(\boldsymbol{r}))^T \tag{2.25}$$

$$\mathrm{F} = \begin{pmatrix} \psi_1(E_1) & \psi_2(E_1) \\ \psi_1(E_2) & \psi_2(E_2) \end{pmatrix}. \tag{2.26}$$

Then we can write

$$\boldsymbol{\mu}(\boldsymbol{r}) = \mathrm{F}\boldsymbol{\rho}(\boldsymbol{r}) \tag{2.27}$$

$$\boldsymbol{\rho}(\boldsymbol{r}) = \mathrm{F}^{-1}\boldsymbol{\mu}(\boldsymbol{r}). \tag{2.28}$$

If monoenergetic illumination is assumed, the image domain approach leads to the same results as the data-domain decomposition. However, when polychromatic illumination is used, as is the case in practice, the image domain method may suffer from biased image values due to inaccurate compensation of beam hardening artifacts.

One-Step Reconstruction

Another approach to spectral CT reconstruction is to develop a model that directly relates basis-material images to the measured data, thus combining reconstruction and material decomposition into a single step. This sort of unified reconstruction model is attractive, theoretically, because it can be adapted to handle very general acquisition strategies. In essence, it combines the flexibility of the image domain approach with the accuracy of the data domain method.

On the other hand, one-step reconstruction models are quite computationally expensive and lead to large, non-convex, nonlinear optimization problems that may be difficult to solve in practice. Consequently, one-step reconstruction models have not been studied as

extensively as the projection space and image domain decomposition methods and will likely require the development of new algorithmic tools. Despite the additional complexity, this approach may be useful for atypical geometries and very low dose imaging scenarios. In this work, we have focused mainly on the other two workflows, so we will not elaborate further on one-step reconstruction.

Dual-Energy Implementations

In principle, dual-energy CT (DECT) data can be acquired on an ordinary scanner by performing sequential scans with two different tube potentials. While this may be sufficient for some applications, this technique is not robust to motion artifacts. In the last ten years, several specialized DECT scanners have become clinically available, and these can acquire a complete set of dual-energy data in a single scan. In this section, we will review some of these configurations and briefly discuss the relative advantages and disadvantages of each technique.

Dual-Source CT. In 2006, Siemens unveiled the first dual-source scanner, comprising two independently operating sources, each facing its own detector array, in a single gantry [19]. The sources are offset by 90 degrees, so that two perpendicular views are simultaneously measured at each gantry angle. Figure 2.5 depicts a schematic of the dual-source scanner.

The dual-source system has several advantages over other DECT designs. Firstly, one can apply additional filtration to the higher-energy source to increase its average energy and reduce spectral overlap between the low- and high-kVp data. Secondly, one has complete flexibility in adjusting the mAs and kVp of each tube independently, ensuring adequate flux at both energies. Finally, the gantry rotation speed is similar to a single-source scanner, so one can either acquire dual-energy data in an ordinary scan time or set both tubes to the same potential to acquire a double-speed single-energy scan.

This configuration also has several disadvantages. Firstly, a design compromise is made

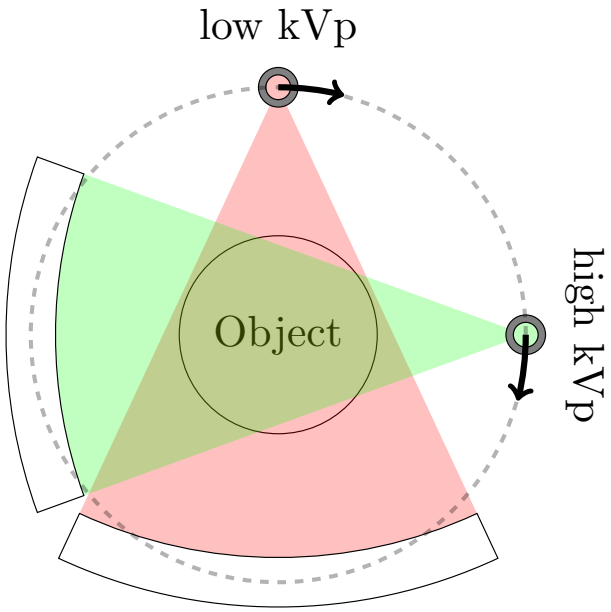


Figure 2.5: Schematic illustrating principal of dual-source CT. Low and high energy sources each face their own detector and simultaneously rotate around the patient, collecting their own projection views.

due to insufficient space in the gantry: the secondary x-ray source faces a smaller detector array, resulting in a significantly smaller field-of-view (FOV) for dual-energy scans. Table 2.1 gives the FOV size for each source in the three existing generations of Siemens' dual-source scanner. Furthermore, due to the presence of two active detector arrays, there is additional cross-scatter contaminating the projection data at both energies. Lastly, since the dual-energy projection data are not registered, material decomposition analysis is typically performed post-reconstruction, which may be less accurate than projection-domain methods, since polychromatic effects cannot be fully accounted for [47].

Scanner	**Tube A FOV(cm)**	**Tube B FOV(cm)**
SOMATOM Definition (1st Gen.)	50	26
SOMATOM Definition Flash (2nd Gen.)	50	33
SOMATOM Force (3rd Gen.)	50	35

Table 2.1: FOV restrictions on dual-energy scans for different scanner generations.

Fast kV-Switching. Another approach to dual-energy CT is to rapidly alternate the tube potential as the source rotates around the gantry. This results in two nearly registered sinograms at two different energies. This is illustrated in Figure 2.6. However, this scheme cannot be easily realized on a standard CT scanner. Successful implementation requires a fast generator and a short scintillator response time. For this purpose GE has developed a proprietary "Gemstone" detector with fast primary decay and low afterglow [48].

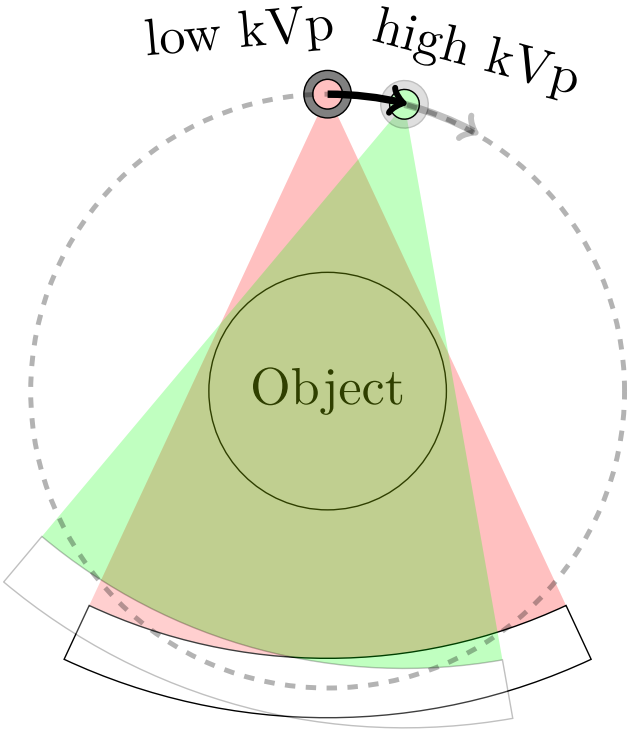


Figure 2.6: Schematic illustrating principal of rapid kV switching. A single source-detector pair rotate around the patient, alternating between low and high energies by switching the tube potential.

This design affords several benefits over the dual-source approach. Since there is only a nominal angular offset between projection views at each energy, accurate data-domain material decomposition methods can be applied. Also, since this configuration requires only a single source and detector array, both energies have a full FOV, and there is no potential for cross-scatter.

Rapid kV switching has several disadvantages as well. Spectral separation is worse than

the dual-source system, since additional filtration cannot be applied to the high-kVp projections. Furthermore, there is less flexibility in balancing the flux between low- and high-energy spectra. When the tube potential is dropped, there is a corresponding (roughly quadratic) decrease in flux. This is partially compensated by dynamically changing the dector integration time, which puts additional constraints on the gantry rotation speed in dual-energy mode.

Dual-Layer CT. Since the x-ray tube spectrum is naturally polychromatic, it is also possible to collect multi-energy data with an energy sensitive detector. This is the principle behind the dual layer detector, sometimes referred to as a dual-crystal or sandwich detector. Two scintillator materials are sandwiched together, each separately coupled to photodiodes, so that the energy deposited in each layer can be separately read out. The front layer preferentially absorbs lower-energy x-rays, while the rear layer stops the remaining higher-energy x-rays. This is illustrated in figure 2.7.

Unlike the dual-source and rapid kV switching methods, the projection data are perfectly registered between energies, both spatially and temporally, since they are collected simultaneously. This allows for accurate data-domain decomposition methods. Furthermore, there is no potential for cross-scatter, and both energies have a full field-of-view. An interesting additional benefit is that no special scanning protocols are needed. Spectral information is always available in every scan, and conventional images can be recovered by aggregating the data from both layers. The previously mentioned methods alter the source energy, so it must be determined in advance if spectral information is needed. Despite these advantages, the spectral separation of dual-layer CT is much worse than the other methods, since the average energy detected in each layer is quite similar [49].

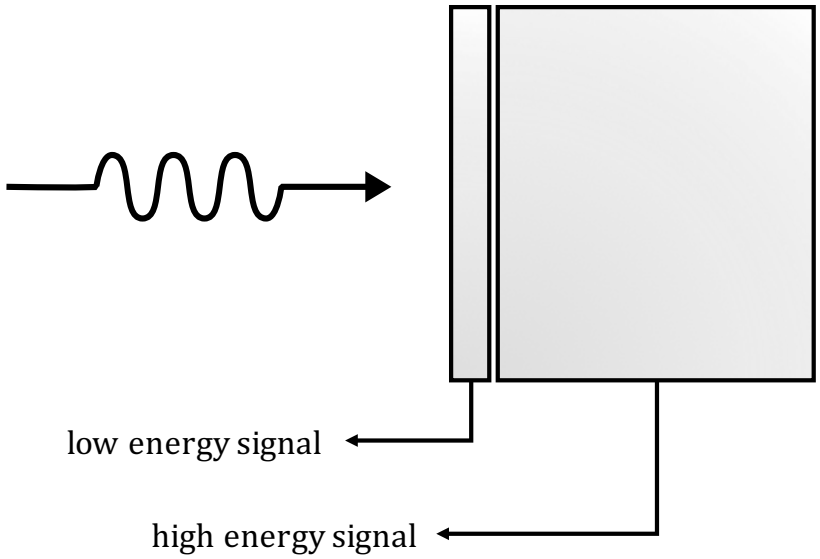


Figure 2.7: Schematic illustrating principal of dual layer (sandwich) detector. The front layer stops, primarily, the lower energy x-rays, while the rear layer detects the remaining, higher-energy x-rays.

Clinical Applications

Beam Hardening and Metal Artifact Reduction - Compared with conventional CT, dual-energy CT recovers the complete energy-dependence of the linear attenuation coefficient. Therefore virtual monochromatic images can be synthesized at any arbitrary energy. This also leads to a dramatic reduction in polychromatic beam hardening artifacts that plague conventional CT; these are especially problematic in patients with metal implants. Figure 2.8 depicts virtual monochromatic images showing screws in the spine (top) and tibia (bottom) reconstructed at various keV's [50] [1].

Virtual Non-Contrast - Besides correcting for polychromatic artifacts, dual-energy CT can also be used to quantify the distribution of a particular material throughout the patient. For scans enhanced by iodinated contrast agents, the iodine distribution can be obtained. Consequently, if a pre- and post-contrast scan are required, as is the case for many clinical

1. "This figure is reprinted from European Radiology, Metal Artifact Reduction by Dual Energy Computed Tomography Using Monoenergetic Extrapolation, Vol. 21, 2011, p. 1426, Bamberg, Fabian et al. With permission of Springer."

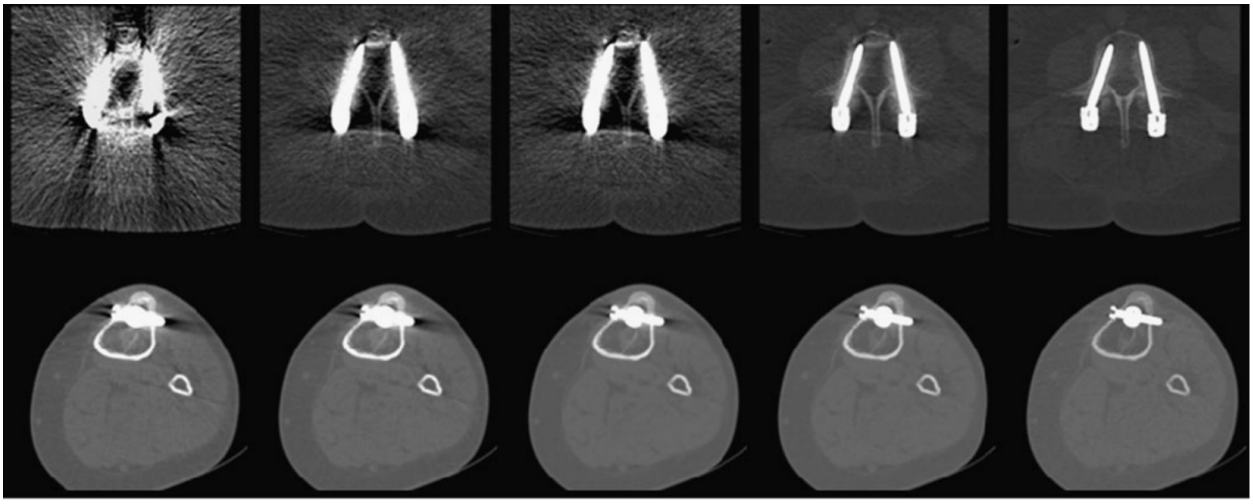


Figure 2.8: Two sets of images showing screws in the spine and in the tibia reconstructed at 64, 69, 88, 105 and an optimal keV setting (left to right). Note that the spinal canal the thin layer of bone covering the left screw are only discernible in the two reconstructions at the highest energy. Similarly, the screw in the tibia is optimally depicted in the rightmost image.

protocols, one can acquire a single contrast-enhanced scan and virtually subtract out the iodine after the fact. This virtual contrast subtraction is shown in Figure 2.9 [51][2].

Compared with the usual method of acquiring two separate scans before and after administering contrast, VNC protocols reduce the amount of radiation dose and ensure that the pre- and post-contrast images are perfectly registered.

Kidney Stone Classification - In addition to material quantification, dual-energy CT may be used to classify different kinds of materials, e.g. based on their effective atomic number. One interesting application is for distinguishing uric-acid (UA) kidney stones from non-UA stones [52]. This task is clinically significant because uric acid kidney stones can often be simply dissolved by administering the patient oral medications. On the other hand, non-UA stones may require invasive surgery if they cannot be passed safely. Figure 2.10 depicts an image of porcine kidneys that have been artificially implanted with many kidney

2. Reprinted from Radiologic Clinics of North America, 47, Joel G. Fletcher, Naoki Takahashi, Robert Hartman, Luis Guimaraes, James E. Huprich, David M. Hough, Lifeng Yu, Cynthia H. McCollough, Dual Energy and Dual Source CT: Is There a Role in the Abdomen and Pelvis?, pp. 41-47, 2009, with permission from Elsevier."

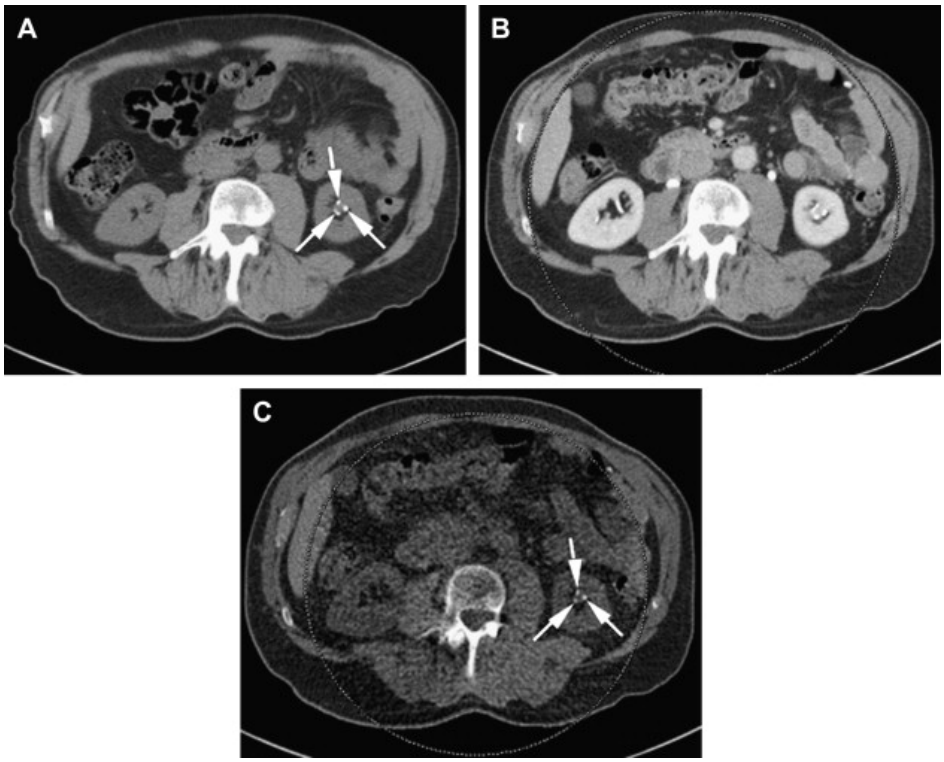


Figure 2.9: Iodine removal using virtual noncontrast, as demonstrated by three small stones in a left inferior renal calyx. Unenhanced CT image shows the stones (arrows, A). During the pyelographic phase (B) obtained as part of a dual-energy CT urogram, the mixed-kV image shows high attenuation, which is a mixture of iodinated contrast and the stones, within the left inferior pole calices. Using the high- and low-kV images obtained from the dual-energy examination, iodine classification and removal was performed to yield a virtual noncontrast image (C), which also demonstrates the three small stones in the inferior pole calices (C, arrows).

stones, and dual-energy post-processing software has been used to label them as UA or non-UA stones [52] [3].

3. Reprinted from Academic Radiology, 14, Andrew N. Primak, Joel G. Fletcher, Terri J. Vrtiska, Oleksandr P. Dzyubak, John C. Lieske, Molly E. Jackson, James C. Williams, Cynthia H. McCollough, Noninvasive Differentiation of Uric Acid versus NonUric Acid Kidney Stones Using Dual Energy CT, pp. 1441-1447, 2007, with permission from Elsevier.

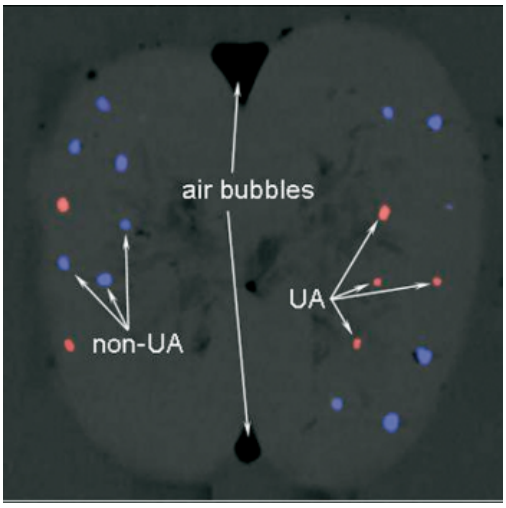


Figure 2.10: Examples of the color-coded images produced by the dual-energy software tool (Kidney Stones, Syngo DE Viewer, Siemens). Uric-acid (UA) and non-UA kidney stones were implanted in a porcine phantom, and dual-energy images were used to classify and label them accordingly.

2.4 Multi-Energy Photon Counting CT

Recently, direct-conversion, photon counting detectors (PCDs) with energy-discriminating capabilities have been developed for x-ray imaging. Though this technology is currently found only in research and prototype systems so far, it offers several attractive benefits over the aforementioned dual-energy imaging techniques. In particular PCDs have the ability to bin detected photons according to their energy, resulting in a plurality of spectral measurements with excellent energy separation. For this reason, the term "spectral CT" has been coined to refer to both dual-energy CT and PCD-based systems, which can have more than two energy channels.

Photon Counting vs. Energy Integrating Detectors

Figure 2.11 offers a side-by-side comparison of a PCD (left) and a conventional energy-integrating detector (right). In an EID, photons interact in a phosphor layer, where they are converted, first, to visible light and then to charge via a photodiode. What is ultimately measured is the total energy deposited in the scintillator. Hence, all of the spectral information

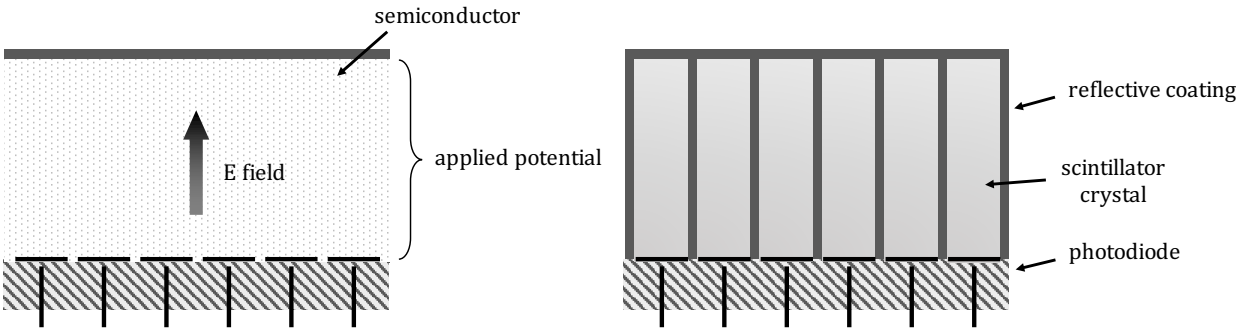


Figure 2.11: Side-by-side comparison of a conventional energy integrating detector (EID) (left) and a photon counting detector (PCD) (right).

is collapsed together into a single measurement.

On the other hand, the PCD consists of a monolithic semiconductor layer that directly converts x-ray photons into charge. Each x-ray results in a charge pulse that is proportional to its kinetic energy, and the readout speed is generally high enough to capture individual interaction events. A single comparator can be used to distinguish actual events from electronic noise. Additionally, multiple comparators can be used to parse all of the detected photons into multiple energy bins, resulting in a coarse histogram of the detected x-ray photons. This concept is illustrated in figure 2.12.

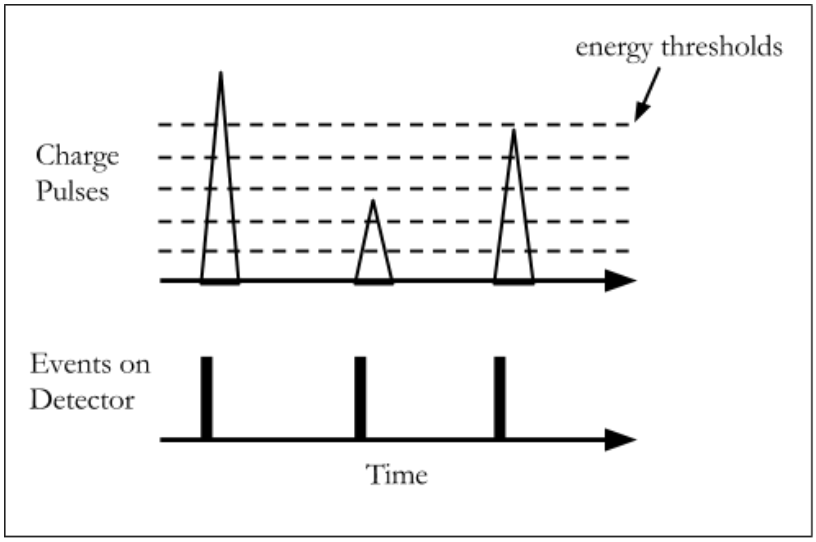


Figure 2.12: schematic of pulse height analysis

k-Edge Imaging

All tissues in the human body can be accurately modeled as a linear combination of water and bone [53], so dual-energy imaging is sufficient for complete characterization of their x-ray attenuation properties. However, heavy metals have a discontinuity in their attenuation coefficients, referred to as a "k-edge," which gives them a unique spectral signature. One can set the energy thresholds in the PCD strategically (e.g. straddling the k-edge) in order produce highly sensitive maps of this material's concentration [54]. Thus for k-edge imaging it is advantageous to measure more than two energy channels and tune the energy thresholds according to the metal.

One such k-edge material is iodine, which is already commonly used as a contrast agent in diagnostic imaging. However, not many photons below its k-edge (33 keV) will be transmitted through a patient. For this reason, there is interest in using other metals to develop alternative contrast agents, such as gadolinium and gold. Figure 2.13 depicts the mass-attenuation coefficients of iodine, gadolinium, and gold beside a typical x-ray tube spectrum.

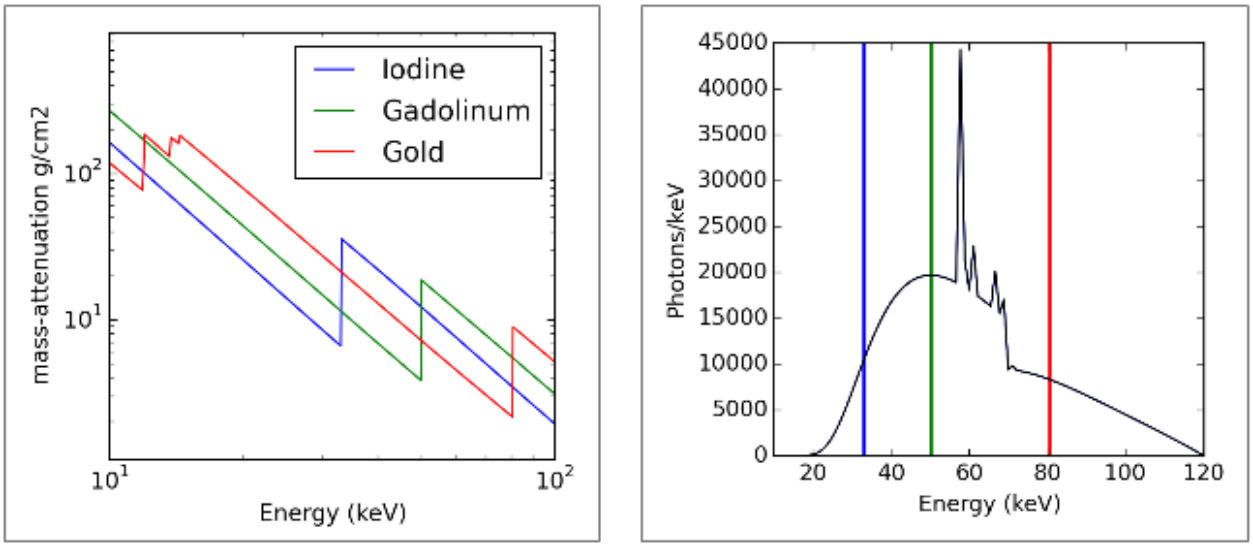


Figure 2.13: This figure shows the location of the k-edges of iodine, gadolinium, and gold in relation to a typical x-ray tube spectrum. Note that gold and gadolinium have k-edges in a portion of the spectrum where many more photons may be detected.

Additional Benefits

Besides the potential for multi-energy and k-edge imaging, PCDs have several other benefits for diagnostic CT applications as well. Firstly, because photon interactions are counted discretely with a comparator, virtually all electronic noise can be rejected; this is particularly useful for low-dose imaging tasks, where electronic noise can significantly contribute to signal degradation [35].

Besides the rejection of electronic noise, PCD's can be more easily manufactured to smaller pixel sizes than EID's and have a substantially higher geometric efficiency [5]. Most diagnostic CT EID arrays consist of individually cut and polished scintillating crystals, coupled to photodiodes, and surrounded by reflective material to prevent crosstalk between detector elements. Usually, the detectors are $\sim$1mm, while the reflective material is on the order of 100μm. Therefore, the geometric efficiency is only about 80% (in two dimensions) and would be even worse for smaller detector sizes [35]. However, PCD's consist of a monolithic slab of an active semiconductor material. Since the x-rays directly create electron-hole pairs, guided by an applied electric field, there is no need for reflectors. As a result the geometric efficiency is nearly 100% and the detector size can be reduced dramatically; this is especially important for specialized applications requiring very high resolution, such as breast CT.

Furthermore, conventional EIDs have a response function that is approximately linear with energy. Hence they give more weight to higher-energy photons, which contribute less contrast information than lower-energy photons. It has been shown that this inherent energy weighting is undesirable and that PCDs may achieve a much higher contrast-to-noise ratio (CNR), even for tasks that don't require spectral information [37].

Hence, PCD's offer a number of benefits besides their spectral capabilities which are likely to lead to more dose-efficient, higher-resolution CT scanners in the future. However, currently PCD's have not supplanted EID's in diagnostic CT due to their higher cost and numerous technical hurdles. We will now discuss some of these issues.

Non-Idealities and Technical Challenges

Numerous physical effects can degrade the energy resolution of PCD's or cause them to record the incorrect number of photons. One of the biggest challenges is *pulse-pileup*, which occurs when two photons are detected in quick succession, such that they are registered as a single higher-energy photon. This is especially problematic at high fluxes, which are required in diagnostic CT in order to achieve fast gantry rotation (i.e. low scan times). Furthermore, CT detectors are subjected to a very wide dynmamic range. Elements may see anywhere from zero to $\sim 10^6$ photons per view, making it very difficult to make PCD's that are robust to pulse-pileup problems.

Another problem that occurs is *charge sharing*, whereby an x-ray interacts in one detector element, and the resulting charge cloud diffuses partially into a neighboring element. Both elements incorrectly register a count, and the energy is skewed due to incomplete collection of charge. In a similar vein, photons may spread their energy among several detector elements via multiple Compton scatter or via the escape of secondary fluorescent x-rays. The latter effect is known as *k-escape*.

Despite these and other technical issues, PCD's are likely to make their way into commercial CT scanners soon. Several manufacturers are already testing advanced prototypes [55].

3

Optimal Parameters for Material Classification Tasks

The first step in any spectral CT workflow is to acquire projection data at two or more different average energies. In the previous chapter, we saw that many different techniques exist for achieving this, including tube-potential switching (e.g. dual source and rapid kV switching), dual-layer detectors, or energy-resolving PCD's. Each has many more tunable parameters than a conventional scanner. For example, a dual-source system allows independent control of the kVp, mAs, and filtration of each tube, and PCD's can have several adjustable energy thresholds, allowing for millions of potential configurations. The optimal settings for these parameters are seldom obvious and always task dependent. In this chapter, we extend the classical signal detection theory to material classification tasks, to develop an objective metric to guide the selection of imaging parameters and hardware design.

3.1 Background

Many clinical applications of spectral CT involve material-classification tasks. For example, Primak et al. showed that dual-source CT can reliably differentiate uric acid (UA) from non-UA kidney stones [52]. This distinction is important because the former can be

treated with oral medication while the latter may require more invasive therapies. Similar algorithms have also been used to identify gout in the hands and feet directly from reconstructed images [23]. This image analysis could eventually replace the gold-standard test: a painful joint aspiration followed by optical analysis of the synovial fluid [56]. Additionally, Thomas et al. demonstrated that dual-energy, bone-and-plaque-removal algorithms can enhance CT angiography (CTA) by using the spectral information to identify calcium- and iodine-containing voxels [57]. This can improve the assessment of arterial stenoses and eliminate the need for extra precontrast scans. Thus, spectral-CT-enabled material-classification algorithms may lead to lower-dose imaging protocols and valuable, new diagnostic tests.

Several prior works have interrogated the effect of imaging parameters on sinogram and image noise. Alvarez and Macovski's foundational work on dual-energy CT [3] describes a method to estimate the variance of the basis-material sinograms using the Cramer-Rao lower bound (CRLB). Roessl generalized this technique with a more general noise model and demonstrated its use for optimizing detector parameters to achieve minimal noise-variance in material density maps [32]. This analysis is useful for characterizing material *quantification* tasks but does not provide a mechanism for measuring material *classification* performance.

In [58], a linear cascaded systems model is presented for analyzing material-classification tasks. This model is powerful in that it can track task-performance metrics through a "cascade" of image processing and reconstruction steps[1]; however, it relies upon several restrictive assumptions: (1) the basis-material decomposition is linear and performed post-reconstruction, (2) the mass-densities of the contrast materials are known exactly. On the contrary, the most accurate basis-material decomposition algorithms are necessarily nonlinear [3] and account for varying noise levels in projection space. The latter fact is especially important in photon-counting CT, where the number of spectral bins often exceeds the number of basis materials. Also, one usually seeks to classify many image pixels spanning

1. Linear cascaded systems models implicitly assume that each step in the can be represented by a circulant linear operator. In practice this is never strictly true but may be approximately valid within a region-of-interest.

a range of mass-densities (or concentrations), such as iodine-enhanced regions comprising a continuous range of water-iodine mixing.

The goal of this chapter is to provide a simple metric for determining how best to acquire spectral imaging data, independent of reconstruction technique. Specifically, we seek imaging parameters that maximize material-classification performance while minimizing radiation dose. We extend the classical signal detection theory to material-classification tasks and derive a figure-of-merit (FOM) based on the Hotelling Observer (HO) SNR. As in [59] the mathematical observer model is computed post basis-material decomposition, and noise properties of the basis-material decomposition are estimated via the Fisher information matrix. Signal variability, such as variations in the concentration of contrast media, are incorporated with a stochastic object model. This stands in contrast to the signal-known-exactly/background-known-exactly approach of [58]. Furthermore, the proposed model considers merely a single ray, which serves as a surrogate to the complete projection data yet captures the relevant physics for profiling the tasks of interest. The result is an analytic FOM that directly relates SCT imaging parameters to material classification performance. It can be computed very rapidly, making it possible optimize multiple system parameters efficiently.

3.2 Theory and Methods

Material discrimination in dual-energy CT

Due to varying photoelectric and Compton cross-sections, as well as element-specific k-edges, different materials have unique spectral signatures related to their effective atomic numbers (Z_{eff}).

Figure 3.1 illustrates this principle with a disk phantom containing calcium and iodine inserts of varying concentration. The Hounsfield Unit (HU) values of the disks overlap significantly at a single-energy, so simple thresholding cannot successfully distinguish these two elements. On the other hand, if HU values at different energies (e.g. 53 keV and 71 keV)

are plotted simultaneously in two dimensions, the calcium and iodine pixels cluster along unique rays in the resulting 2D vector space, making classification possible. However, due to image noise and the small angle of separation, some misclassifications are inevitable.

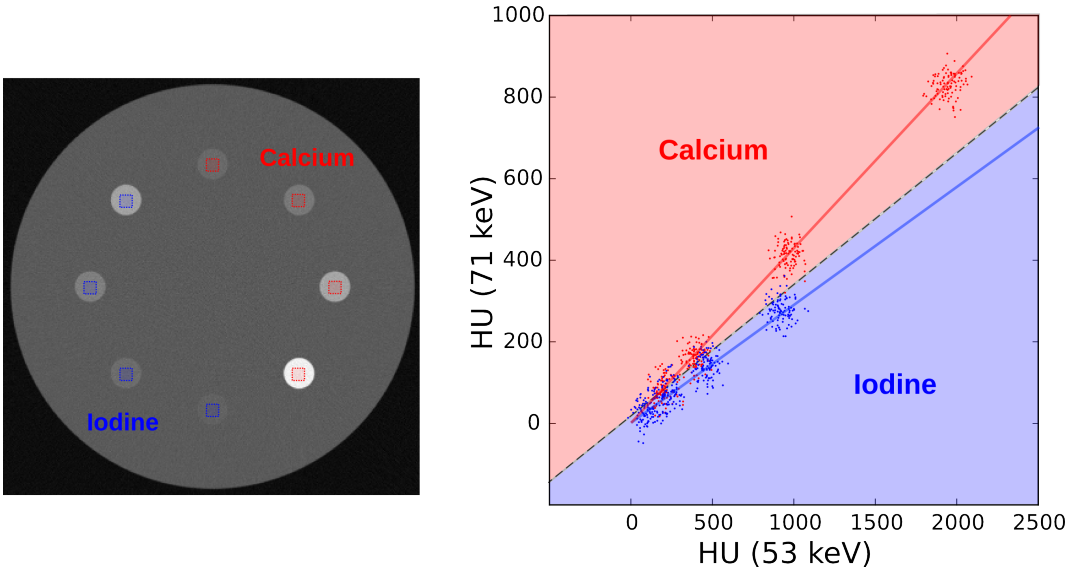


Figure 3.1: Example of calcium vs. iodine classification with numerical disk phantom (left). Pixels of indicated ROI's are plotted in 2D vector space, with HU values at two different energies as x-y coordinates. One possible linear decision boundary is shown.

One can attempt to improve the separation between these material clusters by altering parameters of the x-ray source(s) or detectors (e.g. kVp, filtration, energy-thresholds), but there is a complex trade off between spectral separation and image noise that makes the optimal choices non-obvious. Some metric of material classification performance that can be related back to spectral parameters is needed. The schematic in Figure 3.2 illustrates this general concept for tuning a single imaging parameter. We will refer to this type of plot as a *parameter optimization curve* (POC). In general a POC is an N-dimensional function relating spectral CT hardware/acquisition parameters to some metric of material classification performance. The optimal configuration corresponds to a global maximum of the POC.

In the following section, we present our single-ray Hotelling Observer (SRHO) model for computing analytic POC's.

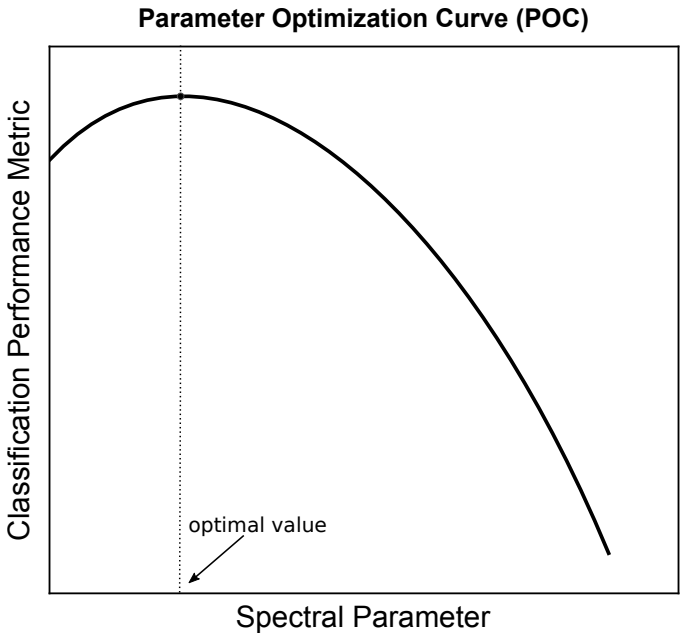


Figure 3.2: Example parameter optimization curve (POC). Optimal spectral parameters, with respect to material classification performance, are identified by the maximum value of the POC.

Single-ray Hotelling Observer Model (SRHO)

The model we develop for characterizing material-classification performance is based on classical signal-detection theory, which has been applied to numerous medical imaging tasks [60, 61, 62]. Typically, a mathematical observer, privy to certain knowledge of the object and imaging physics, must decide whether a signal (e.g. cancer) is present or not [63]. The same notion can be applied to material classification by redefining the hypotheses: signal is material 1 (hypothesis 1), signal is material 2 (hypothesis 2).

In the same spirit as [32] and [36], we use a single ray as a surrogate to the complete projection data, leading to a one-dimensional model specified by the mass-thicknesses of the background and contrast materials. The path lengths used in the 1D model correspond a ray passing through the center of a circularly symmetric object. The diameters should be chosen to reflect the imaging task, e.g. based on the size of the patient and contrast objects.

SKE/BKE

First, we describe a common model-observer paradigm referred to as a "signal-known-exactly/background-known-exactly" (SKE/BKE) task [64], which serves as a natural segue into our signal-known-statistically/background-known-exactly (SKS/BKE) model. For SKE/BKE it is assumed that the contribution of the background (e.g. water) is known exactly. Additionally, the shape and size of the signals are presumed known; the only unknown parameter is whether the contrast object is made of material 1 or 2. This is illustrated in Figure 3.3.

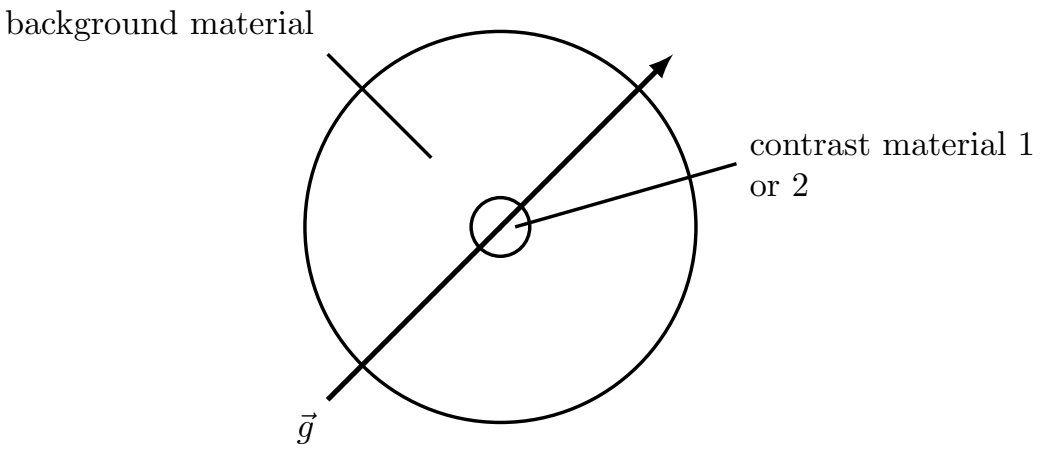


Figure 3.3: The task is determining if the contrast object is material 1 or 2, and the background is presumed to be known.

The task of the mathematical observer is to make this determination based on the projection data $\boldsymbol{g} = [g_1, g_2, \ldots g_N]$, where $\boldsymbol{g}$ are line-integrals of the object's total attenuation along a single ray at energies $E_1, E_2, \ldots, E_N$. The two hypotheses are explicitly defined by the following equations:

$$\mathrm{E}\left\{g_i|H_1\right\} = a_b f_b(E_i) + a_1 f_1(E_i) \tag{3.1}$$

$$\mathrm{E}\left\{g_i|H_2\right\} = a_b f_b(E_i) + a_2 f_2(E_i), \tag{3.2}$$

where $\mathrm{E}\left\{g_i|H_{1,2}\right\}$ represents the expected value of g_i under H_1 or H_2. The density of the background material is denoted by ρ_b. The quantities $a_b \equiv \int_L \rho_b \, d\ell$ and $f_b(E) \equiv \mu_b(E)/\rho_b$ are its "mass-thickness" and "mass-attenuation coefficient (MAC)," respectively. Similar definitions hold for contrast materials 1 and 2, indicated by the subscripts.

We assume that the data, $\boldsymbol{g}$, are mono-energetic with known mean and covariance. This appears to be a very restrictive assumption, since real spectral CT systems use polychromatic illumination. However, virtual monoenergetic (VM) data can be synthesized from arbitrary spectral CT measurements, via nonlinear, basis-material decomposition algorithms. In this case, the VM data are considered to be the input to the HO, and the first- and second-order statistics of $\boldsymbol{g}$ are estimated via the Fisher information matrix. The details of these computations are given in Appendix 3.6.

In this chapter, we specifically consider the Hotelling Observer (HO), which is the optimal *linear* observer; i.e. the HO constructs a test statistic, t, by performing linear operations on the data,

$$t = \boldsymbol{w}^T \boldsymbol{g}. \tag{3.3}$$

The weight-vector $\boldsymbol{w}$ is chosen to maximize

$$\mathrm{SNR}^2_{\mathrm{HO}} \equiv \frac{\left[\mathrm{E}\left\{t|H_1\right\} - \mathrm{E}\left\{t|H_2\right\}\right]^2}{\frac{1}{2}\mathrm{var}\left\{t|H_1\right\} + \frac{1}{2}\mathrm{var}\left\{t|H_2\right\}}, \tag{3.4}$$

and the quantity, $\mathrm{SNR}_{\mathrm{HO}}$, describes how easily separable are the conditional distributions $\Pr(t|H_1)$ and $\Pr(t|H_2)$. The maximizer of (3.4), $\boldsymbol{w}_{\mathrm{HO}}$, is given by

$$\boldsymbol{w}_{\mathrm{HO}} \equiv K_g^{-1} \Delta \bar{\boldsymbol{g}}, \tag{3.5}$$

where $\Delta\bar{\boldsymbol{g}}$ is shorthand for

$$\Delta\bar{\boldsymbol{g}} \equiv \mathrm{E}\left\{\boldsymbol{g}|H_1\right\} - \mathrm{E}\left\{\boldsymbol{g}|H_2\right\}, \tag{3.6}$$

and K_g is the covariance of $\boldsymbol{g}$, averaged over each hypothesis:

$$K_g \equiv \frac{1}{2}\left(K_{g|H_1} + K_{g|H_2}\right). \tag{3.7}$$

Furthermore, it can be shown that

$$\mathrm{SNR}^2_{\mathrm{HO}} = \boldsymbol{w}^T_{\mathrm{HO}}\Delta\bar{\boldsymbol{g}}. \tag{3.8}$$

These relationships are all derived in [65].

Hypothetically, one could construct non-linear observers that may outperform the HO. For example, the Ideal Observer (IO) utilizes complete knowledge of the statistics of $\boldsymbol{g}$ to minimize Bayes' risk. However, the HO is commonly used in place of the IO because it is far simpler to compute and requires knowledge of only the second-order statistics of the data [65]. Furthermore, the HO often serves as an excellent approximation of the IO and may be a better indicator of human performance [66].

SKS/BKE

A crucial component of spectral CT material-classification tasks that we have not yet modeled is signal variability. Note that in Figure 3.1, classifying the calcium and iodine disks is difficult because both types of inserts span a range of concentrations. It is this variability in mass-density that necessitates the use of multi-energy CT rather than conventional CT. If the task were merely to classify two different materials, each of known density, one could always select a single beam energy that would sufficiently separate the two signals. In fact, for an SKE/BKE material classification task, there may be no benefit at all to a multi-energy

imaging technique over single-energy imaging [67, 68]. We will demonstrate this effect later in the context of optimizing a dual-kVp CT system.

In the SKS/BKE model, signal variability is incorporated via a stochastic object model, where the mass-thicknesses, a_1 and a_2 are presumed to follow known distributions, $\Pr(a_1)$ and $\Pr(a_2)$. In this work, we assume that a_1 and a_2 are uniformly distributed on intervals $\left(a_1^{\min}, a_1^{\max}\right)$ and $\left(a_2^{\min}, a_2^{\max}\right)$, but any other probability distribution function can be used just as easily. In practice, the distribution functions characterizing the object variability would have to be obtained from prior knowledge. For example, if the task of interest is a kidney stone classification task, then population statistics can be used to determine the range or distribution of possible stone sizes.

In the SKS formalism, we start by expressing the hypotheses, conditional on a_1 and a_2:

$$\mathrm{E}\left\{g_i|H_1, a_1\right\} = a_b f_b(E_i) + a_1 f_1(E_i) \tag{3.9}$$

$$\mathrm{E}\left\{g_i|H_2, a_2\right\} = a_b f_b(E_i) + a_2 f_2(E_i) \tag{3.10}$$

Computing the Mean Signal. To compute $\mathrm{E}\left\{\boldsymbol{g}|H_1\right\}$ we simply average equation 3.9 and over all possible objects, yielding

$$\mathrm{E}\left\{g_i|H_1\right\} = a_b f_b(E_i) + \bar{a}_1 f_1(E_i) \tag{3.11}$$

where

$$\begin{aligned} \bar{a}_1 &\equiv \int \Pr(a_1)\, a_1\, da_1 \\ &= \frac{1}{2}\left(a_1^{\min} + a_1^{\max}\right), \end{aligned} \tag{3.12}$$

for the uniform distribution. The computation of $\mathrm{E}\left\{\boldsymbol{g}|H_2\right\}$ is completely analogous, and the quantity $\Delta\bar{\boldsymbol{g}}$ is given by equation 3.6.

Computing the Total Covariance. To compute $K_{g|H_1}$ we invoke the "law of total covariance" [65]. This matrix comprises two terms,

$$K_{g|H_1} = \bar{K}^{\text{noise}}_{g|H_1} + K^{\text{obj}}_{\bar{g}|H_1}, \tag{3.13}$$

where $\bar{K}^{\text{noise}}_{g|H_1}$ represents the noise-covariance averaged over all possible objects, and $K^{\text{obj}}_{\bar{g}|H_1}$ represents the variation in the conditional means. For a given mass-thickness a_1, the covariance of $\boldsymbol{g}$, under H_1 is denoted by $K_{g|H_1,a_1}$. The steps for computing $K_{g|H_1,a_1}$ for arbitrary spectral CT systems are detailed in appendix 3.6. Then the noise covariance term is given by

$$\bar{K}^{\text{noise}}_{g|H_1} = \int \Pr(a_1)\, K_{g|H_1,a_1}\, da_1. \tag{3.14}$$

Likewise, the expected measurement, g_i, for a given mass-thickness, a_1, is denoted by $\mathrm{E}\{g_i|H_1, a_1\}$. If we define $g_i' \equiv \mathrm{E}\{g_i|H_1, a_1\} - \mathrm{E}\{g_i|H_1\}$, then the second covariance term is given by

$$K^{\text{obj}}_{\bar{g}|H_1} = \int \Pr(a_1) \cdot \boldsymbol{g'}\boldsymbol{g'}^T\, da_1. \tag{3.15}$$

After computing $K_{g|H_2}$ in a similar fashion, the total covariance matrix K_g is given by equation 3.7. Finally, equations 3.5 and 3.8 are used to calculate the HO SNR, our figure-of-merit for classification tasks. By sweeping out different system parameters and recalculating the HO SNR, a curve like figure 3.2 can be generated, in order to determine which settings result in the optimal classification performance.

3.3 Simulation Studies

The SRHO model provides a rapid method for discovering optimal spectral imaging parameters with respect to material-classification performance. In this section we present a method for validating its predictions by directly measuring material classification performance in simulation studies.

Kidney Stone Classification Task

We created a numerical kidney stone phantom consisting of a water background and 100 randomly located kidney-stone inserts (see figure 3.4). This phantom is motivated by a kidney stone classification task [52] where the goal is to distinguish uric acid (UA) stones from non-UA stones in a known background (e.g. urine). The stone objects are a mixture of water and varying densities of either uric acid or calcium oxalate. Physiologically, the observed differences in stone density are due to the varying porosity of different calculi.

In figure 3.4 the stone materials are indicated in red (uric acid) and blue (calcium oxalate), and each stone's opacity is proportional to its physical density. In our simulations, the contrast objects are selected randomly from a uniform distribution to match the signal variability condition of our SRHO model. The size and composition of the contrast inserts can be modified to simulate other material classification tasks as well.

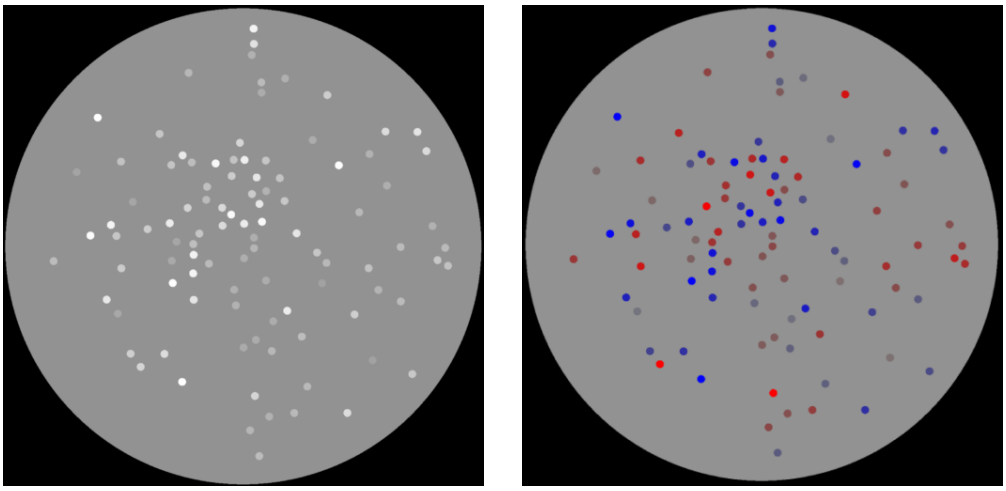


Figure 3.4: "Material classification phantom" containing a water background with different concentrations of contrast material 1 (red) and contrast material 2 (blue). The color opacity is proportional to the concentration. The grayscale image (left) depicts the phantom at 65 keV. Note there is a significant amount of overlap in HU values. The display window is [0 - 400] HU.

Plaque Removal Task

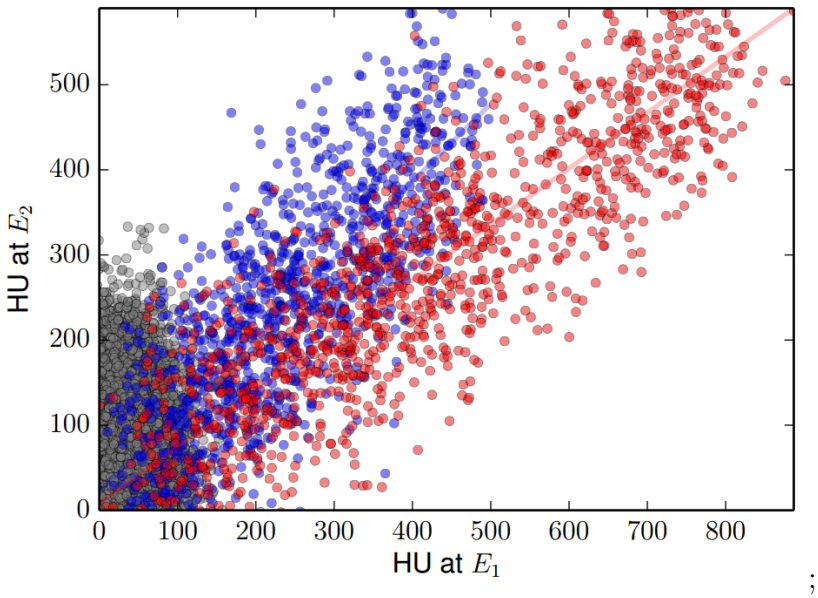
The second task is motivated by plaque-removal classification algorithms that have been developed for dual-energy CT (DECT) [57]. Occasionally, arteries can become occluded by calcified plaques that are indistinguishable from flowing contrast agent on conventional CT scans. Dual-energy measurements can be leveraged to identify and remove the pixels that are made of calcium, rather than iodine, resulting in a clean depiction of the arterial tree. To measure calcium vs. iodine classification performance, we use the same phantom as in the kidney-stone stask, now with calcium and iodine as the insert materials.

Simulation study processing chain

Using these phantoms, we can simulate various spectral CT system configurations and measure their material classification performance. First, raw transmission data are generated using an ellipse-based projector and realistic x-ray tube spectra [1]. Noise is added using a Poisson model for counting detectors and a Gaussian model for energy-integrating detectors [32]. A maximum-likelihood basis-material decomposition is performed, using calcium and water as the basis materials, and a pair of VM images, at energies E_1 and E_2, is reconstructed using FBP. The pixels belonging to the contrast inserts are plotted in the same fashion as in figure 3.1, resulting in the scatter plot shown in figure 3.5. Each point represents a pixel in the reconstructed image pair, and the colors indicate either contrast material 1 (red) or contrast material 2 (blue).

As expected, the two different materials naturally cluster along straight lines, with some dispersion due to noise. We characterize the class separability by defining

$$\mathrm{SNR}^2_{\mathrm{LDA}} = \max_{\boldsymbol{w}} \; 2\frac{\left(\boldsymbol{w}^T\bar{\boldsymbol{x}}_1 - \boldsymbol{w}^T\bar{\boldsymbol{x}}_2\right)^2}{\boldsymbol{w}^T(\Sigma_1+\Sigma_2)\boldsymbol{w}}. \tag{3.16}$$



;

Figure 3.5: A sample scatter plot illustrating the vector space representation of the VM images created from the numerical material classification phantom. The two contrast materials are color coded red and blue, while the background pixels are shown in gray. Fisher's linear discriminant provides an SNR-like measure of how well separated these point clouds are.

This ratio is known as Fisher's linear discriminant [65] (FLD) and is maximized at

$$\boldsymbol{w} = (\Sigma_1 + \Sigma_2)^{-1} (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2). \tag{3.17}$$

FLD can be thought of as an analog to the HO SNR, replacing the population means and covariances with sample statistics. The quantities $\bar{\boldsymbol{x}}_j$ and Σ_j represent the sample mean and covariance, respectively, of material class j, and the quantity $\boldsymbol{w}$ is the normal vector to the best separating hyperplane.

The quantity, SNR_{LDA}, provides a measure of how well the two contrast materials in the stone phantom can be separated. We can simply repeat this entire process while iterating through different imaging parameters in order to generate an empirical parameter-optimization curve, as in figure 3.2. This procedure is summarized by the chart in figure 3.6. In this work, we also repeated each simulation 100 times to reduce uncertainty in the measured POC's, so each point constitutes a considerable amount of computational effort.

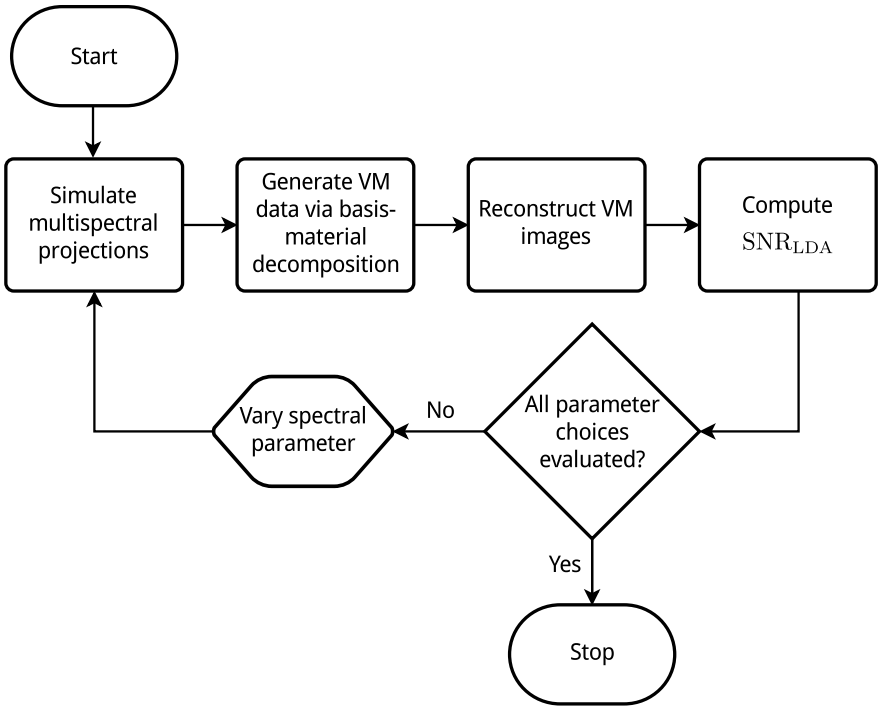


Figure 3.6: This chart describes the processing chain for measuring material classification performance from simulation data. Simulated spectral CT projection data are ultimately converted into VM images, and linear discriminant analysis (LDA) is used to derive a classification SNR based on *Fisher's Linear Discriminant*

Though computationally intensive, the POC's we generate from simulations provide an accurate reference with which to validate our SRHO model and further demonstrate the need for fast, analytic methods.

3.4 SRHO Model Validation

In this section we directly compare the POC's generated by the SRHO model to those derived from numerical simulations. As a proof of concept, we will optimize a dual-kVp system and a photon-counting system with respect to the kidney stone classification task and the plaque removal task. The parameters used for each are given in table 3.1. By sweeping out different spectral CT parameters and computing both $\mathrm{SNR}_{\mathrm{HO}}$ and $\mathrm{SNR}_{\mathrm{LDA}}$ we generate parameter-optimization curves (POC) that reveal which system configurations are optimal for each task.

Geometric Scale Factor

In order to compare POCs between the SRHO model and the simulation studies, we need to introduce a correction factor to account for the acquisition and phantom geometries used in the simulation studies.

The SRHO model is inherently one-dimensional, so a scaling factor is required to account for the geometry of the acquisition and the resolution of the reconstructed image. The mean number of transmitted photons along a given ray is

$$\bar{y}_i = I_0 \int s_i(E) e^{-\int_L \mu(\boldsymbol{r},\boldsymbol{E})\, d\ell}\, dE, \tag{3.18}$$

where we have explicitly indicated the number of incident photons I_0 by defining the normalized spectrum $s_i(E) = S_i(E)/\int S_i(E)\, dE$. In the simulations, the classification performance is sensitive to the total number of views, N_θ, and the reconstructed pixel-size, Δx. Specifically, increasing N_θ or Δx results in less noise per pixel and, thus, better classification performance.

One might think that these factors would only impact the overall scale of the POC and not the shape. However, both the absolute performance and the shape of the POC are impacted because the task is limited by both the measurement noise and the stochastic object model. Increasing N_θ or Δx can only help to reduce the measurement noise, thus altering the trade-off between these two sources of uncertainty. We introduce the scaling factor Ω in the SRHO model, such that

$$I_0' \leftarrow \Omega I_0 \tag{3.19}$$

and

$$\Omega \equiv \frac{N_\theta}{A} I_0, \tag{3.20}$$

where N_θ is the number of projection views and A is the area of the contrast object (e.g.

parameter	kidney stone classification	plaque removal
phantom diameter	30 cm	30 cm
background material	water	water
insert (stone) diameter	5mm	5mm
material 1	uric acid	Ca
material 1 concentration	200 to 550 mg $\cdot$ cm^{-2}	150 to 350 mg $\cdot$ cm^{-2}
material 2	calcium oxalate	iodine
material 1 concentration	75 to 950 mg $\cdot$ cm^{-2}	5 to 30 mg $\cdot$ cm^{-2}

Table 3.1: Phantom parameters used for kidney stone classification and plaque removal tasks

kidney stone) in # of pixels.

Case Study 1: Dual kVp dose allocation

First, we consider the task of optimizing the flux balance in a dual kVp system, operating at peak tube voltages of 80 and 140 kV. The goal is to determine what fraction of the total flux should come from the 80 kVp tube to maximize material classification performance. Specifically, we enforce that along each line integral through our object, a total of 10^6 photons are incident on the phantom:

$$\int r\, S_{80}(E) + (1-r) S_{140}(E)\, dE = 10^6. \tag{3.21}$$

We assume that $0 < r < 1$. The spectra used in our simulations are shown in figure 3.7.

Parameter optimization curves are computed for a range of r values between 0 and 1 using both the SRHO model and via simulations with the numerical phantom; this is shown in figure 3.8. Since SNR_{LDA} is based on image realizations, 100 averages were taken for each r value. The error bars from the simulation studies correspond to two standard deviations.

The POC curves derived from the SRHO model cannot be compared to those from the simulations in terms of absolute scale, since they correspond to fundamentally different things: one is based on a one-dimensional, data-domain model, and the other is measured from reconstructed images. However, since the goal of the SRHO model is to select opti-

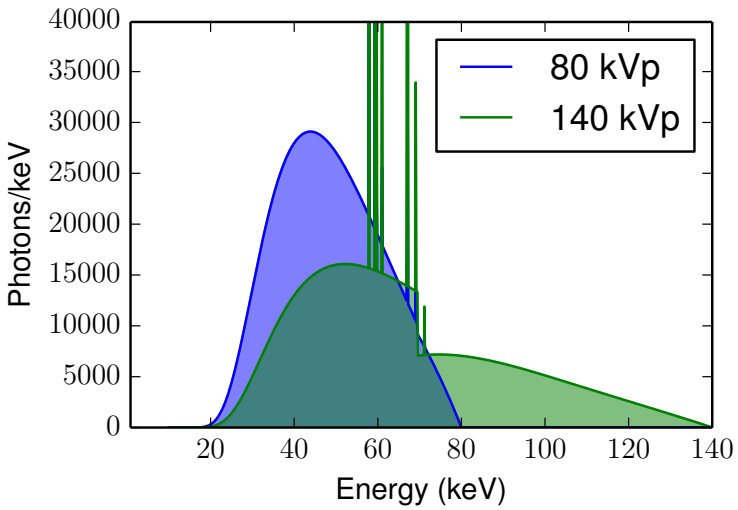


Figure 3.7: Dual kVp spectra used for the simulation study and HO SNR computations are computed using Tucker's model [1].

mal parameters, we are only interested in the shape of these curves, so we can rank order different spectral CT configurations. Therefore, we quantify their agreement with the Pearson product-moment correlation coefficient [69] (PPMCC) because it is insensitive to these absolute scale differences. Note that the analytic model and simulations yield very similar curves, indicating that the SRHO model can accurately predict good operating parameters. For the kidney-stone task, the PPMCC is 0.994. Figure 3.9 shows an identical comparison for the plaque removal (Ca vs. I) task (PPMCC = 0.998).

Case Study 2: PC detector threshold optimization

In the second case study, we consider optimizing a 3-bin photon counting detector. For simplicity, we model an ideal detector with perfect energy resolution, as shown in figure 3.10.

This is a two-parameter optimization problem, as we must select two thresholds given by t_1 and t_2. Figure 3.11 depicts the SNR values for the kidney stone classification task from both the SRHO model (left) and the simulation studies (right) over a wide range of possible threshold values (PPMCC = 0.999). The results from the plaque removal task are shown in

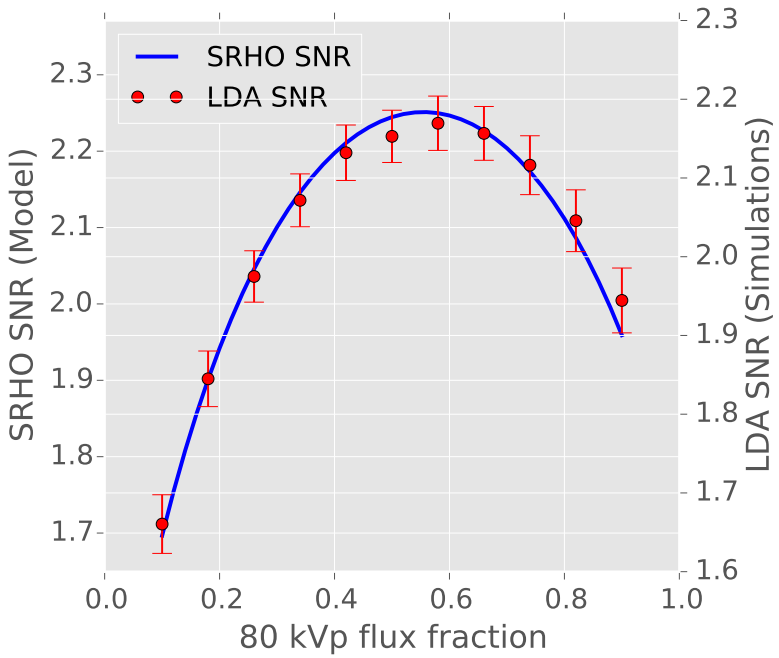


Figure 3.8: Kidney stone classification task: Comparison of classification SNRs derived from SRHO model and 2D simulations with numerical material classification phantom containing uric acid (UA) and calcium oxalate inserts. For the simulation study, 100 averages are used at each r value, and error bars represent two standard deviations. PPMCC = 0.994.

figure 3.12 (PPMCC = 0.981). Again, 100 averages are taken at each combination of (t_1, t_2) when computing SNR_{LDA}. For this task, we also find that HO model accurately predicts the trends in material classification performance for varying threshold values.

In Figure 3.12, agreement is not as good as for the kidney stone classification task. One possible explanation is that the simulation study utilized a two-material decomposition (calcium and water), despite the presence of three materials in the phantom (water, calcium and iodine). We relied on the approximation that attenuation is dominated by Compton scatter and photo-electric absorption, since very few photons are detected below the k-edge of iodine. For configurations where $t_1 \neq t_2$ it should technically be possible to perform a three-material decomposition, but we only used two because the three material-decomposition was numerically unstable. However, we also point out that the SNR tends to vary slowly around the optimum. Even if the HO model does not exactly predict the optimal operating point, it is still very likely to provide a configuration that achieves nearly optimal performance.

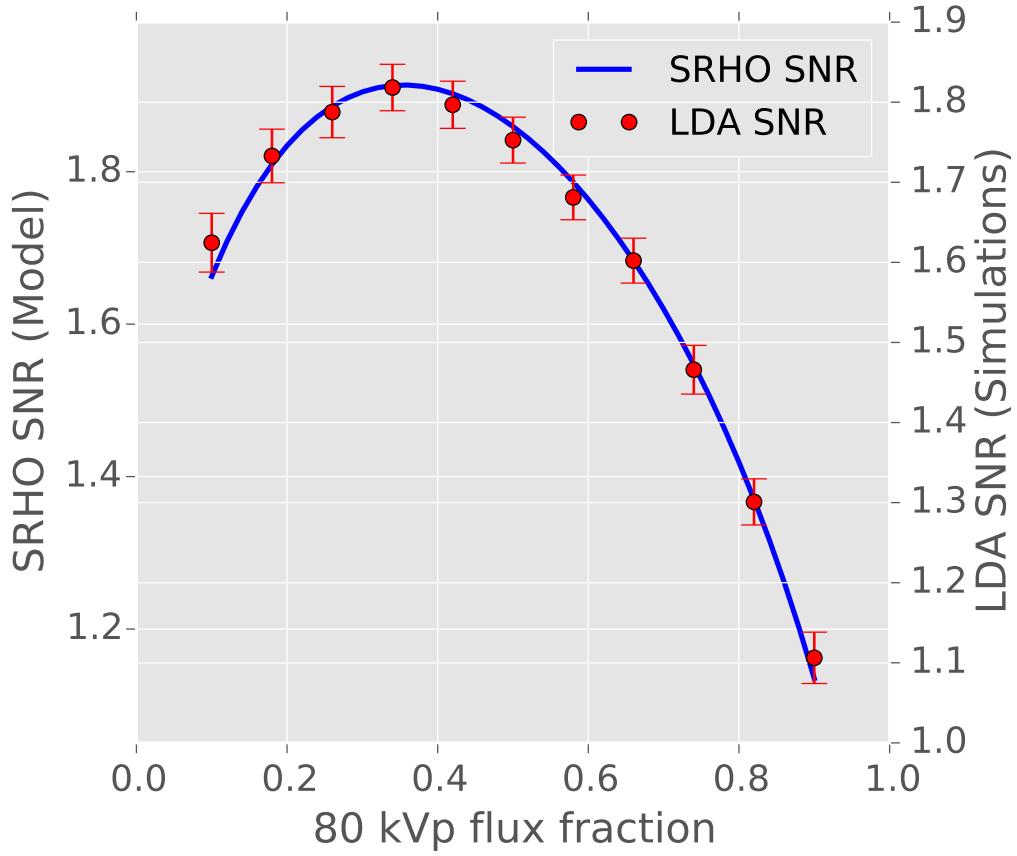


Figure 3.9: Plaque removal task: Comparison of classification SNRs derived from SRHO model and 2D simulations with numerical material classification phantom containing iodine and calcium inserts. For the simulation study, 100 averages are used at each r value, and error bars represent two standard deviations. PPMCC = 0.998.

FloatBarrier

Impact of signal variability

One important component of the aforementioned SRHO model is the inclusion of signal variability. Physically, this may arise due to the mixing of contrast agents during uptake and perfusion (e.g. iodine and blood) or natural variations in tissue composition (e.g. bone density). Mathematically, this enters the SRHO model by specifying a stochastic object model; the density of objects being classified are presumed to follow a known distribution function. In [58], the objects are, instead, presumed to be of a fixed density, corresponding

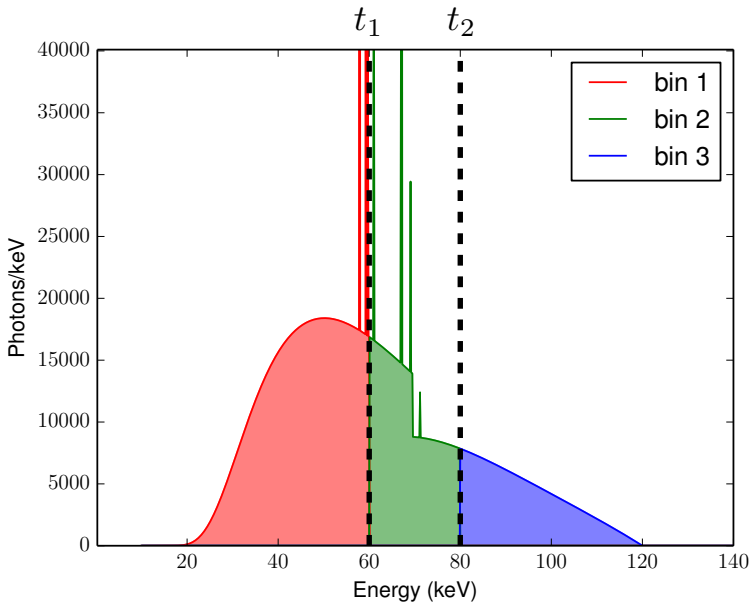


Figure 3.10: Ideal photon counting detector with two adjustable energy thresholds, t_1 and t_2.

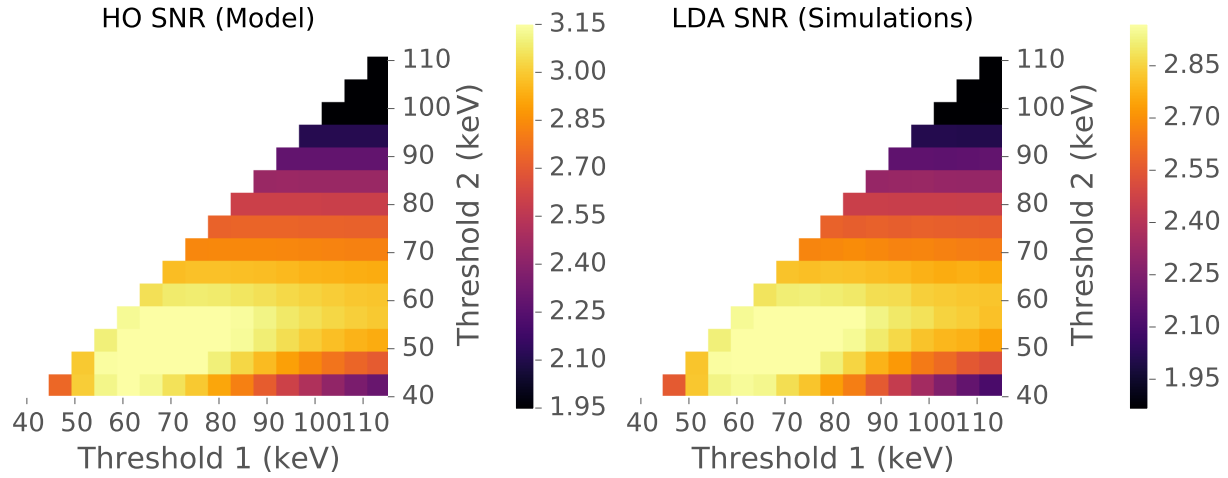


Figure 3.11: Kidney stone classification task: Comparison of classification SNRs derived from SRHO model (left) and 2D simulations (right) with numerical material classification phantom containing uric acid and calcium oxalate inserts. The upper diagonal is omitted because it is redundant with the lower diagonal. PPMCC = 0.999.

to the signal-known-exactly/background-known-exactly (SKE/BKE) task.

In our experience, the SKE/BKE model is too simple to guide the selection of imaging parameters, which is the goal of this work. In fact, we find that in many situations, conventional, single-energy CT performs no worse than spectral CT in material classification tasks

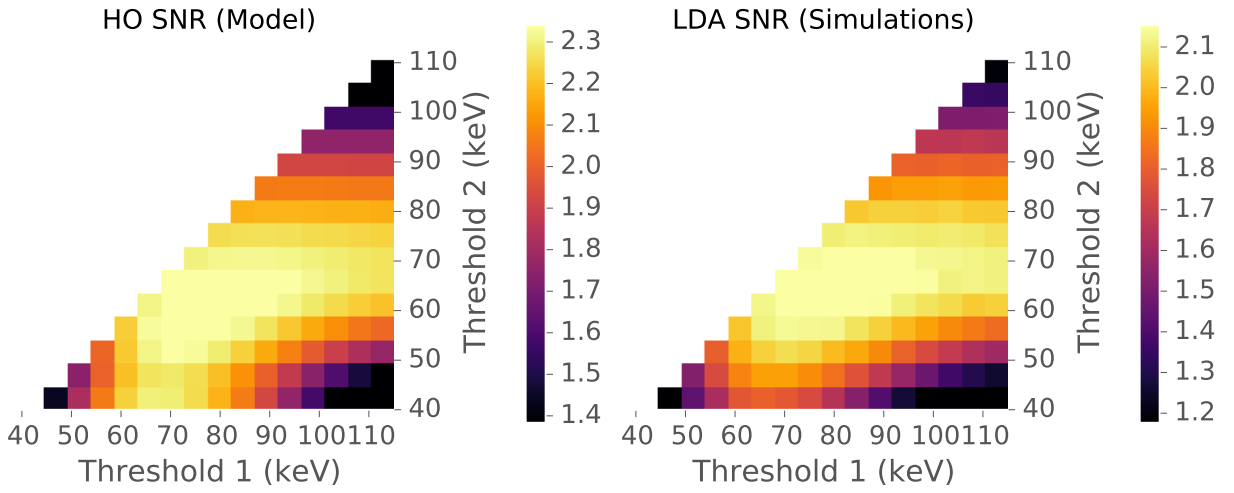


Figure 3.12: Plaque removal task: Comparison of classification SNRs derived from SRHO model (left) and 2D simulations (right) with numerical material classification phantom containing Ca and I inserts. The upper diagonal is omitted because it is redundant with the lower diagonal. PPMCC = 0.981.

with low signal uncertainty. We can illustrate this phenomenon by, again, considering the dual-kVp, flux-balancing problem while shrinking the amount of signal variability toward zero.

In the previous results, we modeled the density of the contrast objects (calcium, iodine, uric acid, or calcium-oxalate) with a uniform distribution,

$$\rho \in \mathcal{U}\left(\rho_{\min}, \rho_{\max}\right), \tag{3.22}$$

where the particular values of $\rho_{\min}$ and $\rho_{\max}$ are specified in table 3.1. In order to investigate the impact of the *amount* of signal variability, we define

$$\begin{aligned} \rho'_{\min} &= \bar{\rho} - k\Delta\rho \\ \rho'_{\max} &= \bar{\rho} + k\Delta\rho \end{aligned} \tag{3.23}$$

where $\bar{\rho} = (\rho_{\min} + \rho_{\max})/2$, $\Delta\rho = (\rho_{\max} - \rho_{\min})/2$, and $k \in (0, 1)$.

The amount of signal variability is controlled by the fraction k, where $k = 0$ corresponds to *no* variability (SKE/BKE), $k = 1$ yields the same distributions as in the previous studies,

and $0 < k < 1$ specifies a range of intermediate scenarios.

Figure 3.13 demonstrates what happens to material classification performance as k varies. The top row illustrates that when k is small enough, the POCs are monotonic, indicating the optimal classification performance is achieved when all of the flux is allocated to a single kVp.

This has two important consequences: (1) For SKE/BKE material classification tasks, dual-energy CT may actually perform *worse* than single-energy CT, and (2) any reasonable material classification model must account for signal variability if it is to be used for parameter tuning or hardware optimization. Figure 3.13 also shows that these observations are confirmed by both the SRHO model and the numerical simulations.

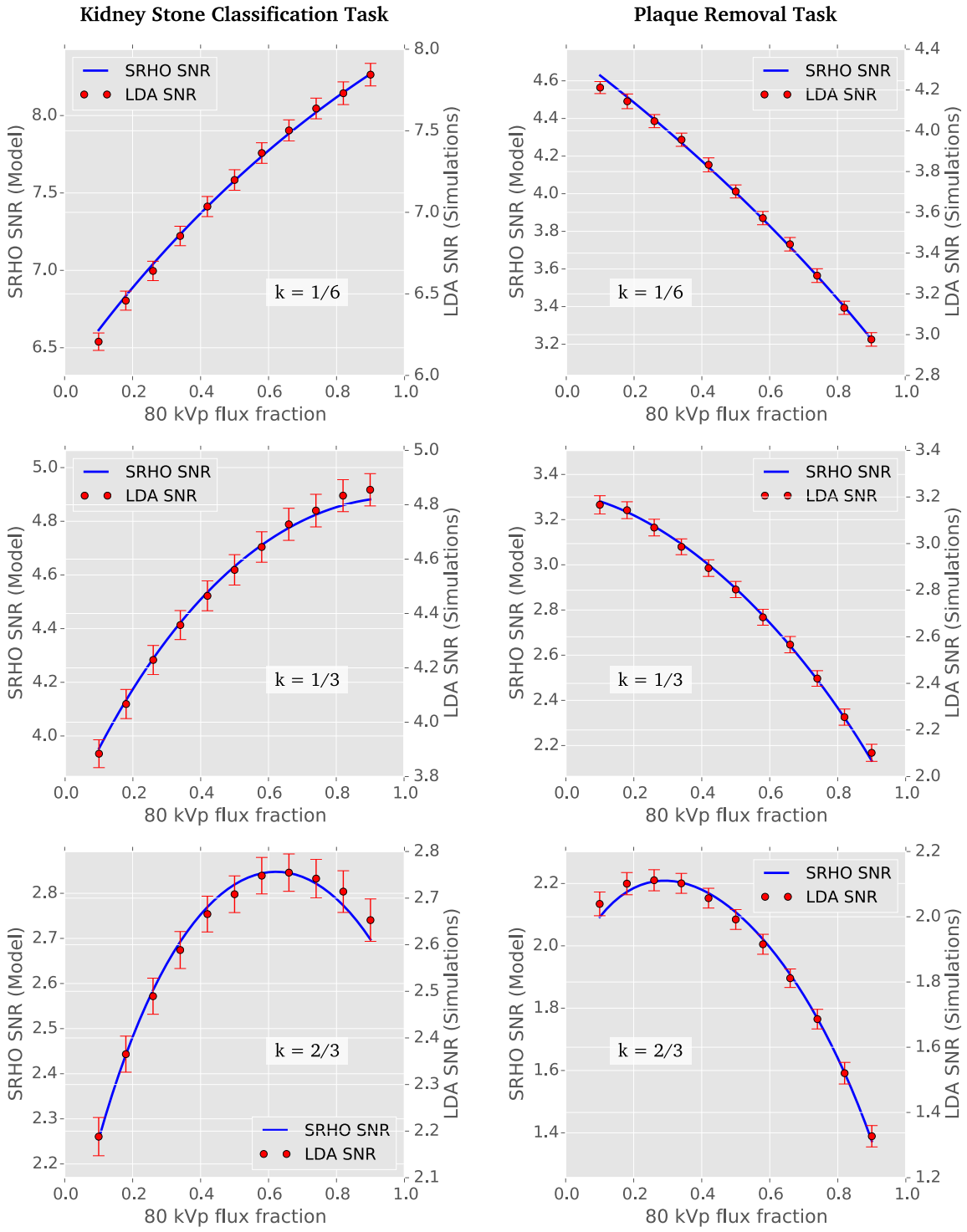


Figure 3.13: The shape of the POC changes significantly with different amounts of signal uncertainty. The parameter k refers back to equation 3.23. For small values of k, the POC is monotonic, indicating that optimal performance is achieved when 100% of the radiation dose is allocated to a single kVp. Kidney stone classification task (from top to bottom): PPMCC = 0.999, 0.999, 0.995. Plaque removal task (from top to bottom): PPMCC = 0.999, 0.999, 0.998.

Impact of background shape

Another concern that arises when comparing the SRHO model to numerical simulations is how to handle asymmetric backgrounds. For example, cross-sections of the human body may span a wide range of apparent diameters, depending upon the projection view angle.

Since the SRHO is inherently one-dimensional, we need a procedure for extracting a single *effective diameter* that allows for a reasonably fair comparison to our 2D simulations. For this reason, we have considered both circular and elliptical kidney stone phantoms. In particular, we considered an elliptical phantom with major and minor axis lengths of 40cm and 24cm, respectively, shown in Figure 3.14. For an elliptical phantom, the intersection

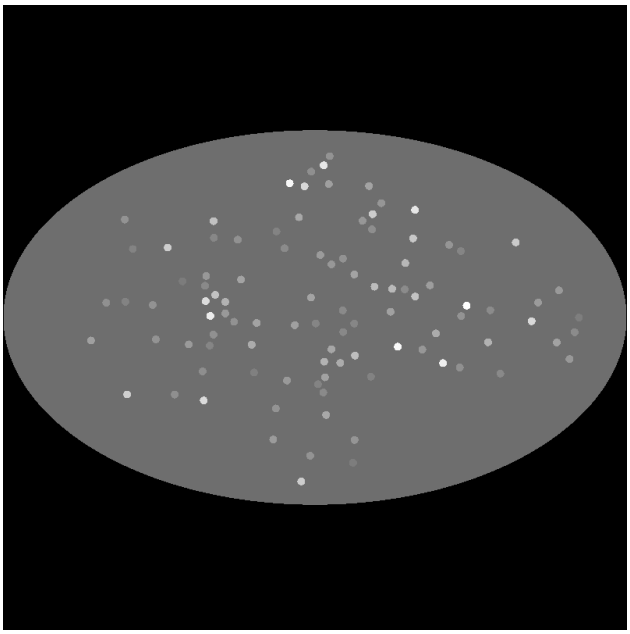


Figure 3.14: An elliptical kidney stone phantom with major and minor axis lengths of 40cm and 24cm respectively. The display window is [0 - 400] HU.

length of a line passing through its center is given by

$$d(\theta) = 2\sqrt{a^2 \cos^2\theta + b^2 \sin^2\theta}. \tag{3.24}$$

Next, we define the quantities

$$V_{\text{ellipse}}(\theta) \equiv \left(\int S(E) e^{-\mu_b(E) d(\theta)} \, dE \right)^{-1} \tag{3.25}$$

and

$$V_{\text{circ}} \equiv \int S(E) e^{-\mu_b(E) d_{\text{eff}}} \, dE. \tag{3.26}$$

The desired quantity d_{eff} is found numerically by equating

$$\left\langle V_{\text{ellipse}}(\theta) \right\rangle_\theta = V_{\text{circ}} \tag{3.27}$$

where $\left\langle V_{\text{ellipse}}(\theta) \right\rangle_\theta$ is an average of V_{ellipse} over a discrete set of projection angles, $0 \leq \theta \leq 2\pi$. In this work, used 1200 equally spaced angles matching those of our simulation geometry and apply the bisection method to find a value d_{eff} that lies somewhere in between the minor and major axis lengths of the ellipse phantom. More details on the effective diamter computation are found in Appendix 3.7.

Once d_{eff} is computed, the SRHO model can be applied to elliptical backgrounds. Figures 3.15-3.18 demonstrate this, with comparisons to POC's derived from numerical simulations. The agreement between the 1D SRHO model and the simulations is slightly worse than with the circular phantom (see Figure 3.15) but still quite good. Most importantly, we can see that the model still accurately predicts near-optimal imaging parameters for these material classification tasks.

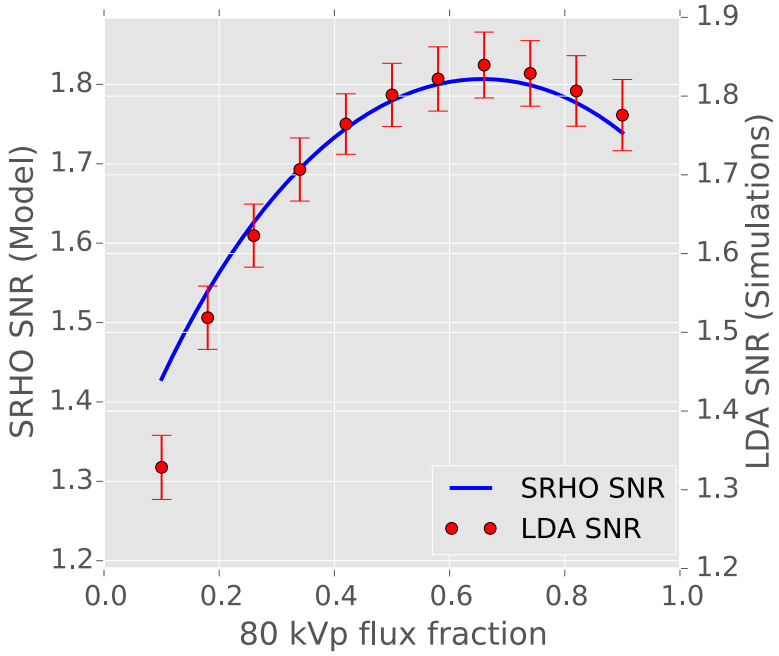


Figure 3.15: Kidney stone classification task with ellipse phantom: Comparison of classification SNRs derived from SRHO model and 2D simulations with elliptical material classification phantom containing uric acid (UA) and calcium oxalate inserts. PPMCC = 0.997.

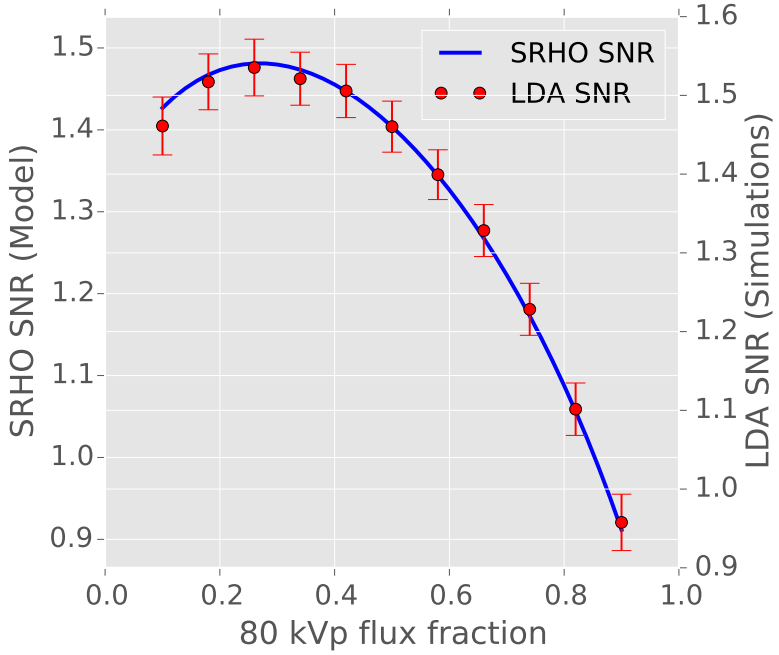


Figure 3.16: Plaque removal task with ellipse phantom: Comparison of classification SNRs derived from SRHO model and 2D simulations with elliptical material classification phantom containing Ca and I inserts. PPMCC = 0.999.

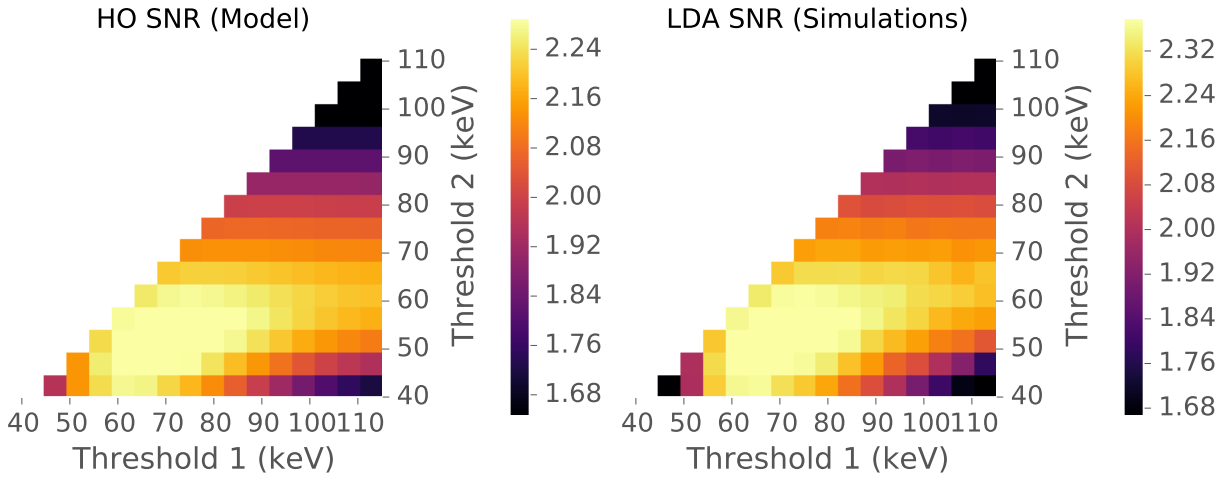


Figure 3.17: Kidney stone classification task with ellipse phantom: Comparison of classification SNRs derived from SRHO model (left) and 2D simulations (right) with numerical material classification phantom containing uric acid and calcium oxalate inserts. The upper triangle is omitted because it is redundant with the lower triangle. PPMCC = 0.990.

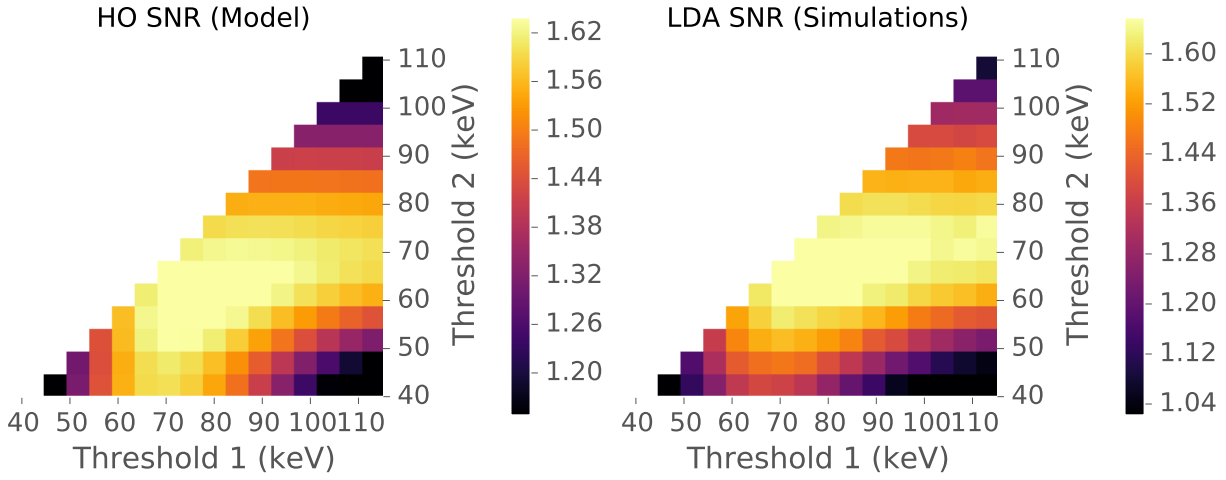


Figure 3.18: Plaque removal task with ellipse phantom: Comparison of classification SNRs derived from SRHO model (left) and 2D simulations (right) with numerical material classification phantom containing Ca and I inserts. The upper triangle is omitted because it is redundant with the lower triangle. PPMCC = 0.973.

3.5 Discussion

In this chapter, we have presented a simple, analytic model for characterizing spectral CT, material-classification tasks that is suitable for optimizing image acquisition parameters and hardware design. Specifically, we have adapted the Hotelling Observer formalism in order to derive an unambiguous, task-specific metric that is independent of the reconstruction method.

Furthermore, the proposed model does not make restrictive assumptions about the basis-material decomposition being linear or the contrast objects having an exactly known density/concentration. We demonstrated that modeling the variability of the contrast objects, e.g. due to uptake and perfusion of contrast agents or natural variations in tissue composition, is essential to material classification tasks. For dual-kVp scans, our model even showed that for low signal variability, the spectral information is not needed, and it is better to deliver all of the dose at a single optimized kVp. This trend was confirmed by our simulations, and echoes a similar sentiment to that expressed in [68] about CNR limited tasks; specifically, they report that for tasks that do not require material classification or quantification, it is better to deliver all of the radiation dose at a single, optimized tube potential, rather than performing a dual-energy scan. Based on this work, we find that *trivial* material classification tasks, where the density of each material is precisely known, also do not benefit from multi-energy scanning.

The proposed model could be especially useful for exploring large parameter spaces because it is analytic and rapidly computable. For example, a photon counting detector with five adjustable thresholds [54] can realize millions of different configurations, making it immensely impractical to explore them all experimentally.

Although this work only considers binary classification tasks, the Hotelling Observer can be easily extended to the L-class problem, where $L > 2$. A detailed discussion of the L-class HO can be found in Chapter 13 of [65]. The linear discriminant analysis applied in our simulation studies can also be generalized in similar manner [70]. This makes it possible

to address classification tasks involving more than two different materials with minimal modification to the methods presented here.

One potential weakness of the proposed approach is the inability to compensate for highly inhomogeneous backgrounds. The SRHO reduces the classification task to a one-dimensional model, where the background and signal are represented by a single path length through each material. This may not pose significant problems in relatively homogeneous slices, such as kidney stone classification in the abdomen, since few bony structures exist. On the other hand, in slices where different projection views may see vastly different attenuation, e.g. due to the skull, shoulders, or hip bones, it is unclear if this simple 1D model would still predict useful operating parameters.

One possible solution to this problem would be to consider multiple rays in the Hotelling Observer model. In principle, one could use the complete sinogram data to derive the HO SNR, but this would significantly increase the computational cost. To balance this trade off, a small subset of rays (e.g. 10) could be used as a surrogate to the entire projection data, in order to capture some of the heterogeneities. However, one typically does not have much a priori knowledge of the object anyway. For these reasons, we have focused on developing a simple metric that is rapidly computable that can be used to guide the search for optimal imaging parameters. Extending the proposed model to include multiple projection rays and prior knowledge about heterogeneous background anatomy is an interesting subject for a future investigation, but it is beyond the scope of this work.

We believe that the simulation studies presented here do demonstrate some robustness to heterogeneities; however, since the projection data comprises rays passing through many different path lengths of water and contrast materials (see Figure 3.4), yet the 1D SRHO model still accurately predicted parameter optimization curves. Furthermore, we showed that even an elliptical phantom, which has a highly view-dependent attenuation profile, could be characterized well with the proposed metric.

In this study, the parameter-optimization curves inherently frame the problem as one

of achieving maximal material-classification performance at a fixed dose. An alternative, equally valid viewpoint is one of achieving clinically acceptable material-classification performance at the lowest possible radiation dose. Therefore, as new spectral CT imaging protocols and hardware become available, it is of increasing importance that careful consideration is given to parameter selection.

3.6 Appendix A: Statistical Properties of Virtual Monoenergetic Data

Maximum Likelihood Material Decomposition

For a general spectral CT system we denote a measurement along a ray $\mathcal{L}$ in the i^{th} spectral channel by

$$\bar{y}_i = \int S_i(E) e^{-\int_{\mathcal{L}} \mu(\boldsymbol{r},E)\, d\ell}\, dE. \tag{3.28}$$

The quantity $S_i(E)$ encompasses the energy-dependence of the incident beam and the detector response. A key insight of spectral CT is that the attenuation can be parametrized by a finite number of basis functions:

$$\mu(\boldsymbol{r}, E) = \sum_{j=1}^{M} \rho_j(\boldsymbol{r}) f_j(E) \tag{3.29}$$

where the $f_j(E)$ may correspond to physical interaction cross-sections or to the mass-attenuation functions of basis-materials [4]. In the latter case, the spatially varying coefficient, $\rho_j(\boldsymbol{r})$, corresponds to the density distribution of material j. This parametrization allows us to write the transmission measurements as

$$\bar{y}_i(\boldsymbol{a}) = \int S_i(E) e^{-\boldsymbol{a}^T \boldsymbol{f}(E)}\, dE, \tag{3.30}$$

where $\boldsymbol{a} = (a_1, \ldots, a_M)^T$, $\boldsymbol{f}(E) = (f_1, \ldots, f_M(E))^T$, and a_j is the line integral of $\rho(\boldsymbol{r})$:

$$a_j \equiv \int_{\mathcal{L}} \rho(\boldsymbol{r})\, d\ell. \tag{3.31}$$

When the number of spectral measurements, $N \geq M$, we can estimate $\boldsymbol{a}$ from $\bar{\boldsymbol{y}}$. Specifically, the maximum likelihood estimate of $\boldsymbol{a}$ is given by

$$\hat{\boldsymbol{a}} = \arg\max_{\boldsymbol{a}} \; L(\boldsymbol{a}), \tag{3.32}$$

where $L(\boldsymbol{a})$ is the log-likelihood function. For Poisson data,

$$L(\boldsymbol{a}) = \sum_i -\bar{y}_i(\boldsymbol{a}) + y_i \log\left(\bar{y}_i(\boldsymbol{a})\right). \tag{3.33}$$

Given an estimate of $\boldsymbol{a}$, virtual monoenergetic (VM) projections can be synthesized at any energy:

$$g(E) = \hat{\boldsymbol{a}}^T \boldsymbol{f}(E). \tag{3.34}$$

Therefore, basis decomposition can be used to transform any arbitrary spectral CT measurements into virtual-monoenergetic (VM) data. By defining $\boldsymbol{g} = (g_1, g_2, \ldots, g_M)^T$ and $F_{ij} = f_j(E_i)$, we get

$$\hat{\boldsymbol{g}} = F\hat{\boldsymbol{a}}. \tag{3.35}$$

Typically, material classification is performed on VM images, since unique materials naturally cluster along straight lines in this space. Since the VM data vector, $\hat{\boldsymbol{g}}$, is related to $\hat{\boldsymbol{a}}$ by a linear transformation, we know that the covariance matrices are related by

$$K_{\hat{g}} = F K_{\hat{a}} F^T. \tag{3.36}$$

Therefore, the HO SNR will depend upon how noise is propagated through the basis-material decomposition.

Noise in the basis-material decomposition

In general, the ML basis decomposition is an implicit, nonlinear operation, so we cannot write down a closed form expression for the covariance of the basis coefficients, $K_{\hat{\boldsymbol{a}}}$. However, we can derive a useful lower-bound from the Fisher Information matrix, as in [3] and [32].

A general lower bound on the variance of any estimator is

$$\mathrm{var}(\hat{\theta}_i) \geq \mathcal{F}_{ii}^{-1}, \tag{3.37}$$

where $\mathcal{F}$ is known as the Fisher information matrix and is related to the curvature of the log-likelihood function. Specifically, it is given by

$$\mathcal{F} = -\mathrm{E}\left\{\nabla^2 L(\boldsymbol{\theta})\right\}. \tag{3.38}$$

This inequality is widely known as the Cramér-Rao lower bound (CRLB). A slightly more general form of the CRLB is

$$K_{\hat{\boldsymbol{\theta}}} \succeq \mathcal{F}^{-1}, \tag{3.39}$$

where the notation $A \succeq B$ means $(A-B)$ is positive semidefinite. When the above holds with equality, the estimator $\hat{\boldsymbol{\theta}}$ is said to saturate the CRLB. All maximum-likelihood estimators (MLE) asymptotically saturate the CRLB, and under many circumstances, it is an excellent approximation to assume

$$K_{\hat{\boldsymbol{\theta}}} \approx \mathcal{F}^{-1}. \tag{3.40}$$

Likewise, if $\hat{\boldsymbol{a}}$ is an MLE, it is asymptotically unbiased, so with a sufficient number of counts, $\mathrm{E}\{\hat{\boldsymbol{a}}\} \approx \boldsymbol{a}$. For the Poisson case, it can be shown that the Fisher information matrix is given by

$$\mathcal{F}_{\alpha\beta} = \sum_{i=1}^{N} \frac{1}{\bar{y}_i} \frac{\partial \bar{y}_i}{\partial a_\alpha} \frac{\partial \bar{y}_i}{\partial a_\beta}. \tag{3.41}$$

For CT systems with energy-integrating detectors, a Gaussian noise model may be more

appropriate, due to energy-dependent detector gain. The detector response is approximately linear with x-ray energy, so the mean and variance are sometimes approximated by

$$\begin{aligned} \bar{y}_i &= \int E\, S_i(E) e^{-\boldsymbol{a}^T \boldsymbol{f}(E)}\, dE \\ v_i &= \int E^2\, S_i(E) e^{-\boldsymbol{a}^T \boldsymbol{f}(E)}\, dE, \end{aligned} \tag{3.42}$$

and the resulting Fisher information matrix is given by the following equation [32]:

$$\mathcal{F}_{\alpha\beta} = \sum_i \frac{1}{v_i} \frac{\partial \bar{y}_i}{\partial a_\alpha} \frac{\partial \bar{y}_i}{\partial a_\beta} + \frac{1}{v_i^2} \frac{\partial v_i}{\partial a_\alpha} \frac{\partial v_i}{\partial a_\beta}. \tag{3.43}$$

Using these formulas we can compute an accurate approximation of $K_{\hat{\boldsymbol{a}}}$ for arbitrary spectral CT systems. Combined with equation 3.36, one can compute the covariance of synthetic VM data. Therefore, we have all of the inputs we need for the SRHO model.

For the SRHO model, the Fisher information matrix is used to compute the conditional covariance matrices under each hypothesis (see equation 3.14). In the Poisson case, the mean transmission measurements can be expressed in terms of the known background and possible contrast materials,

$$\bar{y}_i(a_1, a_2) = \int S_i(E) e^{-a_b f_b(E) - a_1 f_1(E) - a_2 f_2(E)}\, dE, \tag{3.44}$$

Then using equations 3.35 and 3.41, we have

$$K_{g|H_1,a_1} = F \left[\mathcal{F}(a_1, 0)\right]^{-1} F^T \tag{3.45}$$

$$K_{g|H_2,a_2} = F \left[\mathcal{F}(0, a_2)\right]^{-1} F^T \tag{3.46}$$

where partial derivatives of $\bar{y}_i$ are only taken with respect to a_1 and a_2 (and not a_b) since the background is presumed to be exactly known. For the gaussian case, the same process is carried out with equations 3.42 and 3.43. The specific energies used to compute VM

projections are unimportant, since the HO is invariant to invertible linear transformations.

3.7 Appendix B: Effective Diameter Computation for Elliptical Phantoms

One limitation of the SRHO model is that it is inherently one-dimensional, so it is not clear how to choose the appropriate lengths for the background and contrast materials. In reality, the projection data will consist of many different path lengths through both signal and background objects.

When comparing the SRHO to simulation studies, we simply used the diameters of the circular phantom and contrast inserts as the path length for the 1D model. This choice resulted in excellent agreement between the predicted and observed POC curves. However, a modification is required if the aspect ratio of the object is not nearly circular, such as an ellipse with eccentricity $e \not\approx 0$. For these cases, we introduce a method for computing an *effective* diameter.

Consider an elliptical phantom with background attenuation $\mu_b(E)$, defined parametrically by

$$\begin{aligned} x(\theta) &= a\,\cos\theta \\ y(\theta) &= b\,\sin\theta. \end{aligned} \tag{3.47}$$

The intersection length of a ray passing through its centroid (see Figure 3.19) is defined by

$$d(\theta) = 2\sqrt{a^2\cos^2\theta + b^2\sin^2\theta}. \tag{3.48}$$

If the attenuation of the signal is small compared to the background, then we can approxi-

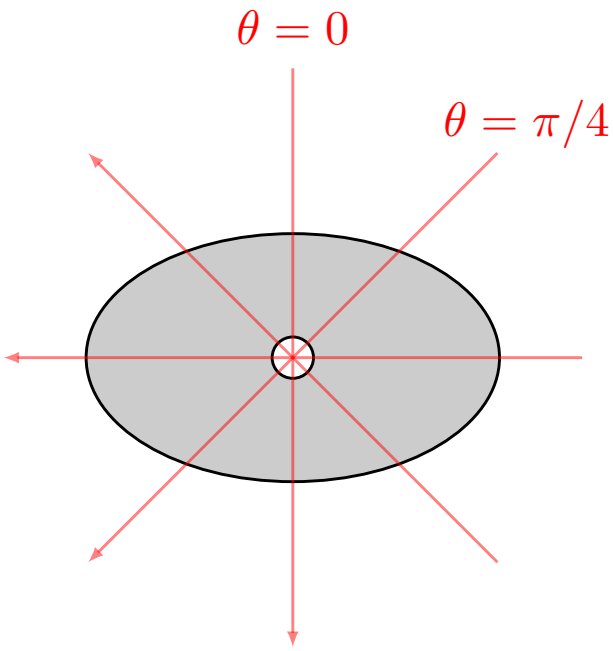


Figure 3.19: For elliptical phantoms, the apparent diameter depends upon the angle, θ, of the inersecting ray. Four different angles are shown here, for illustrative purposes.

mate the number of transmitted photons along angle θ by

$$I(\theta) = \int S(E) e^{-\mu_b(E) d(\theta)} \, dE, \tag{3.49}$$

where $S(E)$ is the incident x-ray spectrum. If $I(\theta)$ is approximately Poisson, then the logarithm of the data, $g(\theta) \equiv -\log\left(I(\theta)/\int S(E)\, dE\right)$ has variance

$$V(\theta) \approx \left[\int S(E) e^{-\mu_b(E) d(\theta)} \, dE\right]^{-1}. \tag{3.50}$$

If we consider the summation of N independent measurements along different angles, the total variance is

$$\sum_{k=1}^{N} V(\theta_k) \tag{3.51}$$

For a circular phantom, with diameter $d = d_{\text{eff}}$, the variance along any ray is given by

$$V_{\text{circ}} \approx \left[\int S(E) e^{-\mu_b(E) d_{\text{eff}}} \, dE \right]^{-1}, \tag{3.52}$$

and the total variance of N independent measurements is

$$N \cdot V_{\text{circ}} \tag{3.53}$$

The effective diameter, d_{eff} is found by equating equations 3.53 and 3.51, which results in

$$\langle V(\theta_k) \rangle_\theta = V_{\text{circ}}, \tag{3.54}$$

where $\langle . \rangle_\theta$ indicates a discrete average over $\theta_1, \ldots, \theta_N$.

Solving for d_{eff}, numerically, can be accomplished by applying any scalar, root-finding method to the function

$$f(d_{\text{eff}}) = V_{\text{circ}}(d_{\text{eff}}) - \langle V(\theta_k) \rangle_\theta \, . \tag{3.55}$$

Specifically, we are looking for a root of $f(d_{\text{eff}})$ on the interval $[2b, 2a]$, where we have assumed that $b < a$; d_{eff} must lie somewhere between the minor and major axes lengths of the ellipse. In this work, we have used the well-known bisection method.

4

Efficient Data-Domain Material Decomposition

4.1 Background

As we reviewed in Chapter 2, one of the primary advantages of spectral CT is the ability to extract complete energy-dependence of the attenuation by decomposing the projection data into basis functions. In their seminal paper [3], Alvarez and Macovski present a polynomial estimator for extracting basis sinograms from dual-energy projections. This non-iterative method works well for dual-energy data, where the estimation essentially involves inverting a function, f: $\mathbb{R}^2 \mapsto \mathbb{R}^2$. However, for photon counting CT systems involving $N \geq 2$ energy channels, this mapping is generally not one-to-one, and an iterative, maximum-likelihood estimation (MLE) scheme may be preferred [54]. The MLE approach elegantly handles noisy and inconsistent data and can be shown to be asymptotically unbiased and "efficient." However, Alvarez has pointed out recently that computational issues may impede the clinical utilization of the MLE method [71]. In particular he mentions that such iterative algorithms may fail to converge or have long or unpredictable computation times. In this chapter, we present an optimization-based approach to material decomposition that achieves errors

similar to the MLE approach and converges to accurate solutions in just a few iterations.

4.2 Theory

Basis-material model

Consider a photon-counting CT (PCCT) system with N energy bins, where the product of the spectrum and the detector response function for bin i is given by $S_i(E)$. In accordance with (2.5), the transmission measurement for the i^{th} bin is given by

$$y_i = \int S_i(E) e^{-\int_L \mu(\boldsymbol{r},E) d\ell} \, dE. \tag{4.1}$$

Next, the linear attenuation coefficient, $\mu(\boldsymbol{r}, E)$, is parametrized by a finite number of basis-materials, as detailed in Chapter 2:

$$\mu(\boldsymbol{r}, E) = \sum_{j=1}^{M} \rho_j(\boldsymbol{r}) \psi_j(E), \tag{4.2}$$

where $\psi_1(E), \psi_2(E), \ldots, \psi_M(E)$ are the mass-attenuation functions associated with some set of M basis materials, and $\rho_1(\boldsymbol{r}), \ldots, \rho_M(\boldsymbol{r})$ are their corresponding density distributions. This allows us to rewrite equation 4.1 as

$$y_i(a_1, \ldots, a_M) = \int S_i(E) e^{-\sum_{j=1}^{M} a_j \psi_j(E)} \, dE, \tag{4.3}$$

where

$$a_j \equiv \int_L \rho_j(\boldsymbol{r}) \, d\ell \tag{4.4}$$

as in (2.10). Note that a_j implicitly refers to a particular ray in the sinogram data, defined by the line integral over L. However, we are building toward a data-domain material decomposition model that is ray-by-ray separable, so we will proceed to describe the model for a

single arbitrary ray; in practice the decomposition algorithm is applied to all rays.

It is useful to define the vector quantites

$$\boldsymbol{a} = (a_1, \ldots, a_M)^T$$

and

$$\boldsymbol{\psi}(E) = (\psi_1(E), \ldots, \psi_M(E))^T$$

in order to write (4.3) more compactly as

$$y_i(\boldsymbol{a}) = \int S_i(E) e^{-\boldsymbol{a}^T \boldsymbol{\psi}(E)}\, dE. \tag{4.5}$$

Though we have referred to a PCCT system, equation 4.5 applies to any dual-energy or spectral CT system with consistent rays, i.e. the multi-energy sinogram data must comprise a geometrically consistent set of projections.

Maximum-Likelihood Estimation

If electronic noise can be ignored, then the measurements $y_1, \ldots, y_N$ are generally Poisson distributed with mean values given by (4.5):

$$\bar{y}_i(\boldsymbol{a}) = \int S_i(E) e^{-\boldsymbol{a}^T \boldsymbol{\psi}(E)}\, dE. \tag{4.6}$$

Thus, the problem of recovering $\boldsymbol{a}$ from the measurements $y_1, \ldots, y_N$ can be viewed as a statistical estimation problem. In particular, the maximum-likelihood estimator (MLE) is defined by

$$\hat{\boldsymbol{a}}_{\text{MLE}} = \arg\max_{\boldsymbol{a}} \; L(\boldsymbol{a}) \tag{4.7}$$

where

$$L(\boldsymbol{a}) = \sum_{i=1}^{N} y_i(\boldsymbol{a}) \log\left(\bar{y}_i(\boldsymbol{a})\right) - \bar{y}_i(\boldsymbol{a}). \tag{4.8}$$

The quantity y_i represents the actual measured data from channel i, and $L(\boldsymbol{a})$ is the Poisson log-likelihood function (omitting terms that do not depend on $\boldsymbol{a}$).

The MLE defined by (4.7) has many nice properties, such as being asymptotically efficient and unbiased [72]. One nice feature of the MLE is that it elegantly handles the case where the estimation problem is overdetermined (i.e. $N > M$). In this chapter, we consider the MLE to be the gold-standard because it very nearly achieves the theoretical bound for an unbiased estimator in many practical situations [54].

Newton Update Step for MLE Maximization. The quantity $\hat{\boldsymbol{a}}$ must be found numerically by iteratively approximating the solution to (4.7). This is a very low-dimensional optimization problem: typically $M \leq 3$, so fast-converging, second-order algorithms can be applied. In particular, we use Newton's Method, which has the following update step:

$$\boldsymbol{a}_{n+1} = \boldsymbol{a}_n - \left[\nabla^2 L\left(\boldsymbol{a}_n\right)\right]^{-1} \nabla L(\boldsymbol{a}_n) \tag{4.9}$$

where $\nabla^2 L(\boldsymbol{a})$ is the Hessian of $L(\boldsymbol{a})$. Conceptually, Newton's Method replaces the function to be maximized by a local quadratic approximation and maximizes this quadratic at each iteration. When $\nabla^2 L$ is 2×2 or 3×3 (as is usually the case for spectral CT material decomposition) an analytic formula can be used for the matrix inversion. In the rare cases where $M > 3$, the Newton update can be solved efficiently using Cholesky decomposition [73]. In our experience, this material decomposition algorithm typically reaches machine precision in 20-50 iterations. This is the fastest method that we are aware of for solving the MLE problem. We will hereafter refer to this scheme of maximizing the Poisson Likelihood function via Newton's method as MLE-Newton.

Logged Nonlinear Least Squares (LNLS)

We have observed that an alternative formulation can give nearly the same estimates as the aforementioned MLE-Newton scheme, while being somewhat easier to solve numerically. It

involves working with the "logged" data, referring to the transmission data after taking an element-wise natural logarithm.

This approach involves casting the material-decomposition as a nonlinear, least-squares problem. First we define the logged, normalized data $g_i = -\log y_i/I_0$ where $I_0 = \int S_i(E)\,dE$. Then we can define a basis-material estimator $\hat{\boldsymbol{a}}$ by

$$\hat{\boldsymbol{a}} = \arg\min_{\boldsymbol{a}} \quad \frac{1}{2}\sum_{i=1}^{N} \left(\bar{g}_i(\boldsymbol{a}) - g_i\right)^2, \tag{4.10}$$

where $\bar{g}_i(\boldsymbol{a}) \equiv -\log \bar{y}_i/I_0$. Next, we define $r_i(\boldsymbol{a}) = \bar{g}_i(\boldsymbol{a}) - g_i$, which form the *residual* vector $\boldsymbol{r} = (r_1, r_2, \ldots, r_N)^T$. This transforms the nonlinear least-squares problem in (4.10) into the canonical form [73]:

$$\hat{\boldsymbol{a}} = \arg\min_{\boldsymbol{a}} \ \frac{1}{2}\|\boldsymbol{r}(\boldsymbol{a})\|_2^2.$$

To better approximate the Poisson MLE, we could replace the euclidean norm by the following weighted norm:

$$\hat{\boldsymbol{a}} = \arg\min_{\boldsymbol{a}} \ \frac{1}{2}\|\boldsymbol{r}(\boldsymbol{a})\|_W^2,$$

where $\|x\|_W^2 \equiv x^T W x$ and W is a diagonal weighting matrix. Specifically, we use the weights $W_{ii} = y_i$. This scheme is commonly referred to as *data-weighted least-squares* and can be viewed as a quadratic approximation to the Poisson likelihood [74]. This weighting scheme can also be viewed as an approximation to inverse-variance weighting, since

$$\mathrm{var}(g_i) = \mathrm{var}(\log y_i) \approx 1/y_i.$$

We will refer to the weighted version as the weighted, logged, nonlinear least-squares (WLNLS) estimator.

Gauss-Newton Update for LNLS We have cast the problem into this form for two main reasons: Firstly, normalizing and taking the logarithm leads to a residual $\boldsymbol{r}(\boldsymbol{a})$ that is approximately linear in $\boldsymbol{a}$. Secondly, we can apply the well-known Gauss-Newton (GN) algorithm [73]. The GN algorithm — not to be confused with the previously mentioned Newton's Method — is a first-order algorithm that applies specifically to nonlinear least-squares problems, and it exhibits rapid convergence when $\boldsymbol{r}(\boldsymbol{a})$ is nearly linear.

We will derive the Gauss-Newton update and show how it can be interpreted as a quasi-Newton method. First, we define the function $\Phi(\boldsymbol{a}) = \frac{1}{2}\|\boldsymbol{r}(\boldsymbol{a})\|_2^2$. By the chain rule, it can be shown that its gradient and Hessian are given by

$$\nabla\Phi = J_r^T \boldsymbol{r} \tag{4.11}$$

$$\nabla^2\Phi = J_r^T J_r + \sum_{i=1}^{N} r_i \nabla^2 r_i, \tag{4.12}$$

where J_r is the Jacobian, defined by $[J_r]_{ij} = \frac{\partial r_i}{\partial a_j}$. Note that all of these terms are functions of $\boldsymbol{a}$, but we have left out this explicit dependence in order to make the notation more concise. The special structure of the nonlinear least-squares problem allows us to make the following approximation to the Hessian of $\Phi(\boldsymbol{a})$:

$$\nabla^2\Phi(\boldsymbol{a}) \approx J_r^T J_r, \tag{4.13}$$

where we have simply ignored the second term. Hence, when the residual $\boldsymbol{r}$ is almost linear in $\boldsymbol{a}$, then $\nabla^2 r_i(\boldsymbol{a}) \approx 0$, and this becomes an excellent approximation. This suggests the quasi-Newton update

$$\boldsymbol{a}_{k+1} = \boldsymbol{a}_k + \left(J_r^T J_r\right)^{-1} J_r^T \boldsymbol{r}, \tag{4.14}$$

where J_r and $\boldsymbol{r}$ are evaluated at the current iterate $\boldsymbol{a}_k$. Carrying out the same steps for the

WLNLS formulation yields

$$\boldsymbol{a}_{k+1} = \boldsymbol{a}_k + \left(J_r^T W J_r\right)^{-1} J_r^T W \boldsymbol{r}, \tag{4.15}$$

or equivalently,

$$\boldsymbol{a}_{k+1} = \boldsymbol{a}_k + \underset{\boldsymbol{x}}{\arg\min}\ \frac{1}{2}\|J_r\boldsymbol{x} - \boldsymbol{r}\|_W^2.$$

We will refer to this scheme combining the WLNLS estimator and the Gauss-Newton solver as the WLNLS-GN method. The motivation for this scheme is that we find that the WLNLS estimator is an excellent approximation to the MLE, and yet the WLNLS-GN scheme converges to sufficiently accurate solutions in as few as two iterations. Furthermore, the per-iteration cost of WLNLS-GN is less than MLE-Newton, since it is a first-order method and does not require the computation of any second derivatives. We point out that the GN algorithm cannot be applied to the original Poisson-likelihood maximization problem, since it only applies to nonlinear least squares; hence the formulation of an alternative estimator.

In the following section, we will compare MLE-Newton and WLNLS-GN, empirically, in terms of both accuracy and convergence rates. Though WLNLS-GN usually reaches machine precision slightly faster than MLE-Newton in our experiments, we pay particular attention to how quickly it reaches medium-accuracy solutions in the very early iterations.

4.3 Results

Comparison of LNLS, WLNLS, and ML Estimators

In this section, we compare the LNLS, WLNLS, and MLE estimators by allowing all of these schemes to converge to machine precision and analyzing the sample mean and covariance of each estimator. We will consider both the fully-determined ($M = N$) and overdetermined ($N > M$) cases as well as two- and three-material decompositions.

Fully-Determined Material Decomposition. To test out the case where $M = N$, we consider a water-and-calcium material decomposition. First, we simulated a single line integral, according to (4.5), using an idealized, two-bin photon-counting detector model with 10^6 incident photons. Next, Poisson noise was added to y_1 and y_2. These noise-corrupted measurements were fed into the MLE-Newton, LNLS-GN, and WLNLS-GN material decomposition schemes and iterated until machine precision was reached. This entire sequence was repeated for one million realizations in order to obtain accurate statistics on each estimator.

The resulting means and covariances are visualized in Figure 4.1 as error-ellipses. The major and minor axes of these error ellipses are proportional to the singular values of the sample covariance matrix. They can be thought of as a two-dimensional generalization of error bars, since a bivariate normal distribution has elliptical confidence intervals. Note

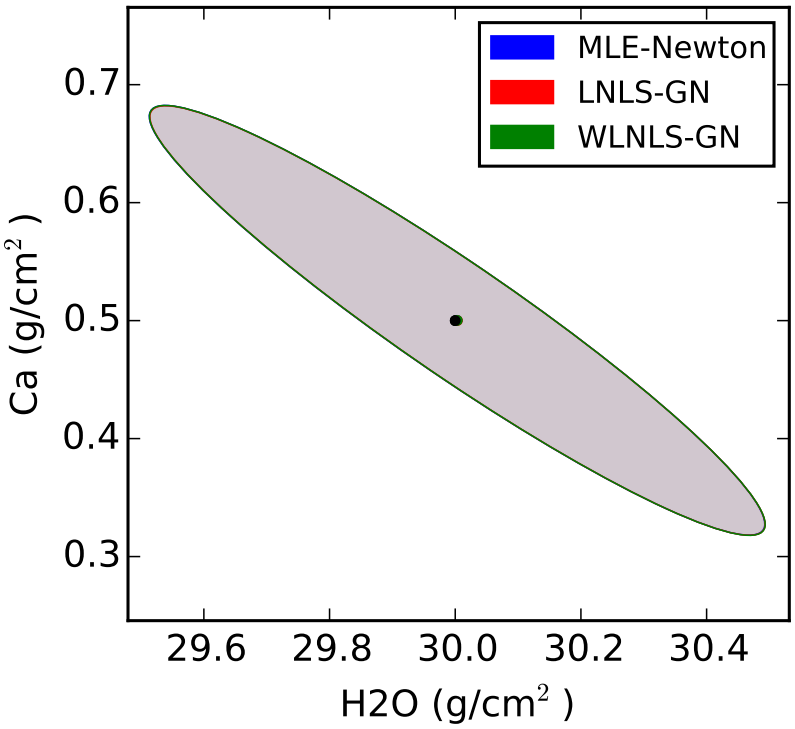


Figure 4.1: Empirical error ellipses for ML, LNLS, and WLNLS estimators based on 10^6 realizations of a water-calcium material decomposition with 2 energy bins. The black dot indicates the true value of $\boldsymbol{a}$.

that for the fully-determined case, all three estimators yield almost identical results with no discernible bias. Since there is a one-to-one mapping between $\boldsymbol{a}$ and $\boldsymbol{y}$, the choice of estimator is not particularly important. All three schemes are, essentially, just finding the same solution to a locally invertible system of nonlinear equations.

Overdetermined Material Decomposition The choice of estimator becomes more important when the number of spectral measurements exceeds the dimension of $\boldsymbol{a}$, i.e. $N > M$. In this case, the system of nonlinear equations is inconsistent, so some criterion must be used to select the desired solution, such as the Poisson-likelihood function.

We carried out numerical simulations similar to the previous section but with a 6-bin photon-counting detector model, instead of only 2 bins. Figure 4.2 shows the resulting error ellipses for this case. This time we find that each estimator yields slightly different results

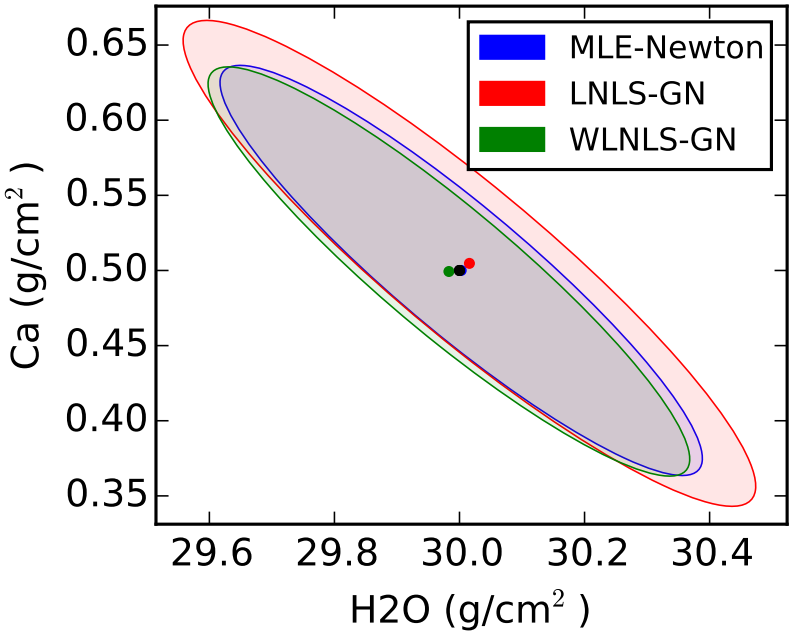


Figure 4.2: Empirical error ellipses for ML, LNLS, and WLNLS estimators based on 10^6 realizations of a water-calcium material decomposition with 6 energy bins. The black dot indicates the true value of $\boldsymbol{a}$.

with LNLS having the largest errors. The data-weighted scheme (WLNLS) improves the

estimate of $\boldsymbol{a}$, yielding a similar covariance to the MLE. Conceptually, the data-weighting helps to account for the fact that certain spectral measurements have larger errors than others. Interestingly, the MLE appears to exhibit virtually no bias, while the WLNLS has a small but discernible bias.

With 6 energy bins, it is also possible to perform a three-material decomposition on certain objects containing k-edge materials. To test out this scenario, we carried out simulations with water, gadolinum (Gd) and iodine (I) in the forward model and performed a three-material decomposition. In this case, the three-dimensional analogue of the error ellipses would be error ellipsoids. However, for easier visualization, we simply plotted 2D error ellipses for two components of $\boldsymbol{a}$ at a time. Figure 4.3 depicts the sample mean and covariance for the Gd and I components of $\boldsymbol{a}$ for all three estimators. Here, it is clear that

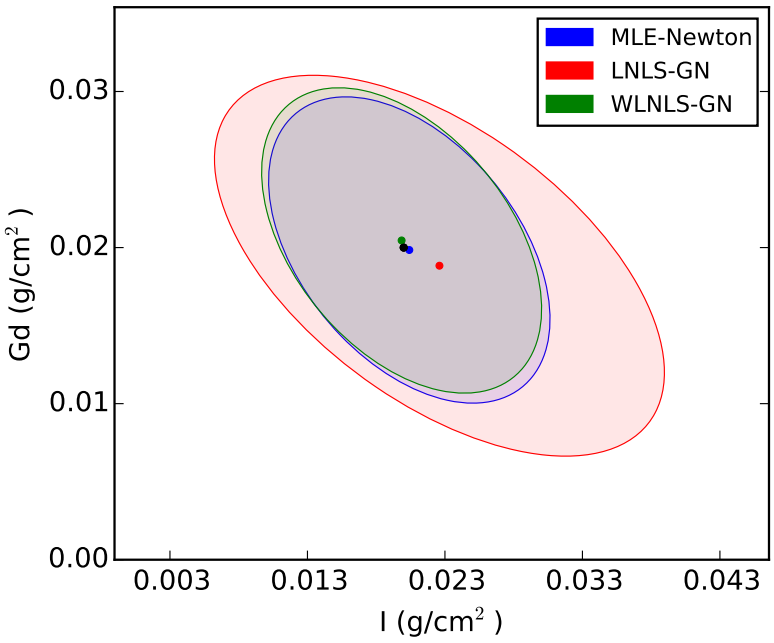


Figure 4.3: Empirical error ellipses for ML, LNLS, and WLNLS estimators based on 10^6 realizations of a water-Gd-I material decomposition. Only Gd and I components are shown here. The black dot indicates the true value of $\boldsymbol{a}$.

the LNLS estimator is inferior to WLNLS; however, WLNLS compares quite favorably to MLE. Both have very similar error magnitudes, and each exhibits a slight bias in different

directions.

These experiments indicate that the WLNLS formulation yields an estimator comparable to the MLE. It also demonstrates the importance weighting the data in the overdetermined case.

Empirical Comparison of Convergence Rates

In this section, we focus on the convergence rate of the various basis-material estimation schemes. The results presented here are derived from the same simulation studies used in the previous section, but we track two convergence metrics at each iteration.

Convergence to the Optimum. The first metric that we look at is the distance between the current iterate and the fully-converged estimator. We define this metric as

$$\text{RMSD}_{\boldsymbol{a}^*} \equiv \|\boldsymbol{a}_k - \boldsymbol{a}^*\|_2. \tag{4.16}$$

Note that the quantity $\boldsymbol{a}^*$ is the optimal point that a particular iterative algorithm reaches as $k \to \infty$. To compute $\boldsymbol{a}^*$, we let each scheme (MLE-Newton, LNLS-GN, and WLNLS-GN) converge to machine precision. Since each scheme defines a unique estimator, there will be a different $\boldsymbol{a}^*$ for each one. Hence, we are only comparing how fast each algorithm approaches its own optimum. The quantity $\text{RMSD}_{\boldsymbol{a}^*}$ should theoretically approach zero as $k \to \infty$, but in practice it will oscillate around a small value once machine precision is saturated, due to rounding errors.

We observe that both LNLS-GN and WLNLS-GN reach machine precision in fewer iterations than MLE-Newton, though not much faster in the three-material decomposition. See figures 4.4 and 4.5. More interestingly, the early convergence is much faster in both GN schemes than MLE-Newton. Thus, if only moderate accuracy is required, these schemes may save many iterations.

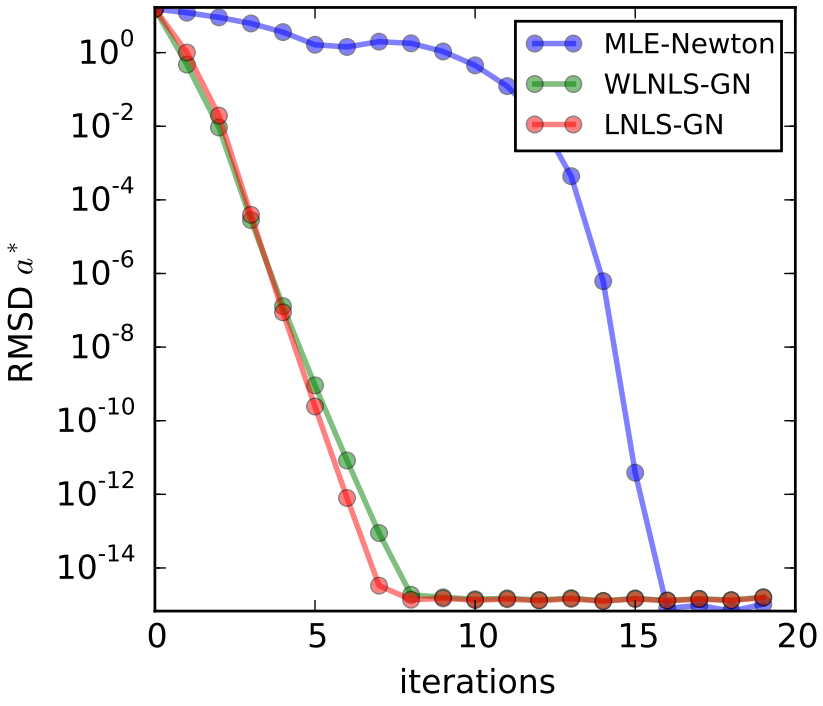


Figure 4.4: Empirical convergence plot of MLE-Newton, LNLS-GN, and WLNLS-GN for a water and calcium material decomposition. This plot illustrates the convergence of each scheme to its own optimizer as a function of iteration number.

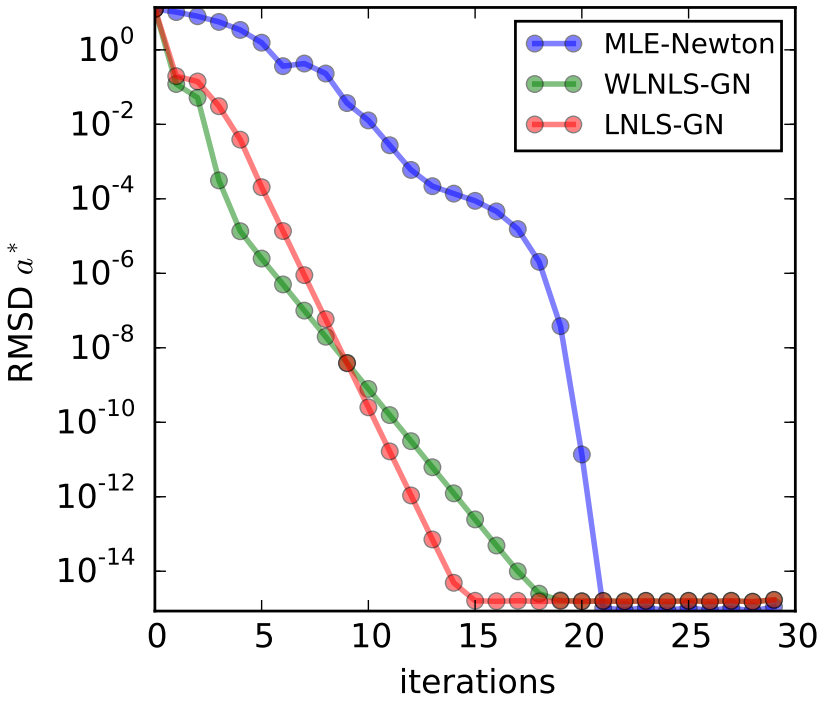


Figure 4.5: Empirical convergence plot of MLE-Newton, LNLS-GN, and WLNLS-GN for a water, gadolinum, and iodine decomposition. This plot illustrates the convergence of each scheme to its own optimizer as a function of iteration number.

Convergence to the True Material Vector. In practice, it is rarely useful to converge such estimation problems to machine precision due to the presence many other confounding error sources, such as detector/spectrum model error, scatter, and quantum noise. To illustrate this, we define a second convergence metric that measures the distance between the current iterate and the true basis-material vector. This quantity is defined as

$$\text{RMSD}_{\boldsymbol{a}_{\text{true}}} = \|\boldsymbol{a}_k - \boldsymbol{a}_{\text{true}}\|_2 \tag{4.17}$$

The quantity $\boldsymbol{a}_{\text{true}}$ comprises the actual inputs used in the forward model to generate the data, prior to adding noise. It can be viewed as the ground truth. Since we are estimating $\boldsymbol{a}$ from noisy data, it should not be expected that $\text{RMSD}_{\boldsymbol{a}_{\text{true}}}$ will tend toward zero. Instead it will converge to some constant minimum value, which may be different for each estimator.

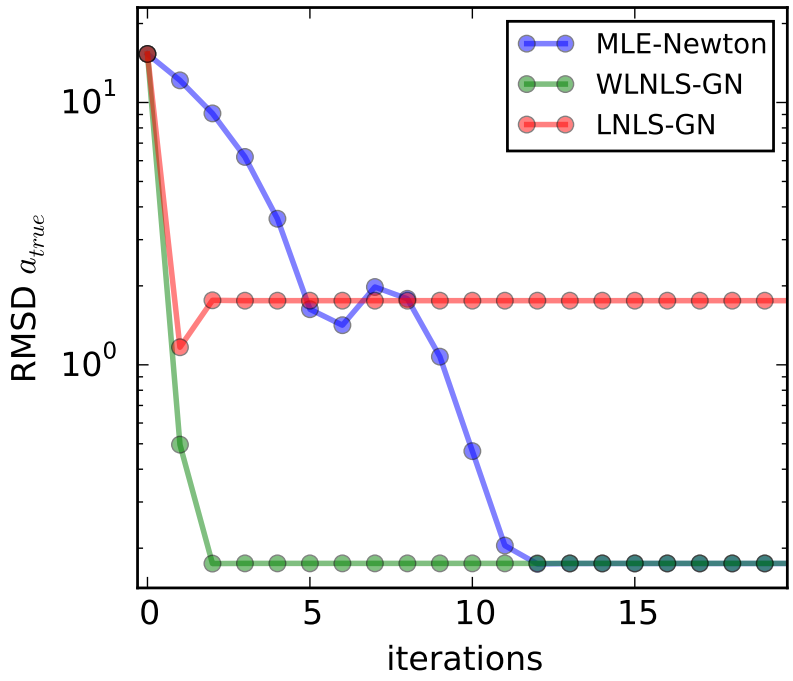


Figure 4.6: Empirical convergence plot of MLE-Newton, LNLS-GN, and WLNLS-GN for a water and calcium decomposition. This plot illustrates the convergence of the estimator error, with respect to the true material vector

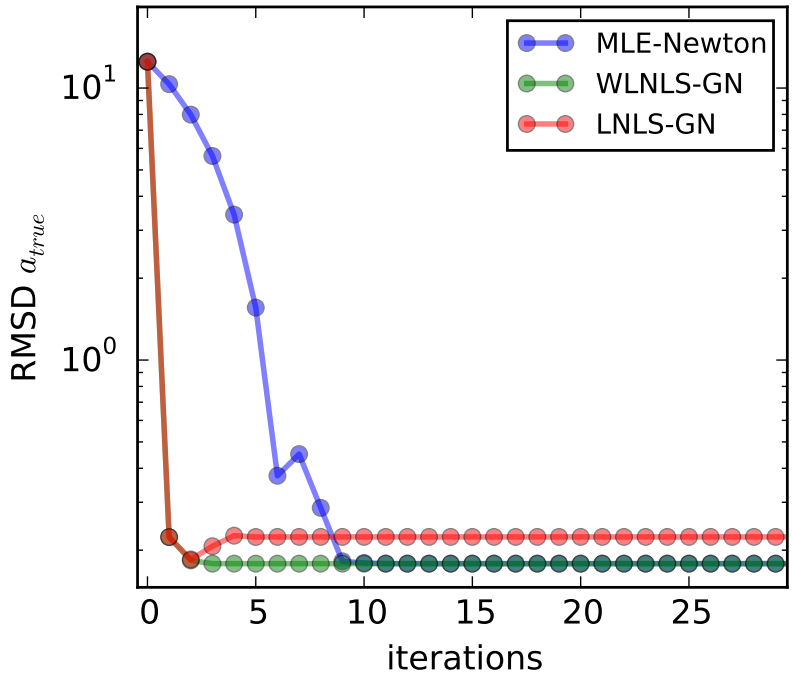


Figure 4.7: Empirical convergence plot of MLE-Newton, LNLS-GN, and WLNLS-GN for a water, gadolinium, and iodine decomposition. This plot illustrates the convergence of the estimator error, with respect to the true material vector

Figures 4.6 and 4.7 illustrate the decay of $\text{RMSD}_{\boldsymbol{a}_{\text{true}}}$ for each scheme. Remarkably, the WLNLS and LNLS schemes reach the noise floor after about 2 iterations. It is also interesting to note that the LNLS scheme plateaus to a larger final error than WLNLS and MLE, which supports our previous observations that it is an inferior estimator.

To better illustrate the fast early convergence of WLNLS-GN, compared to MLE-Newton, we apply both decomposition schemes to sinogram data simulated from a disk phantom, containing water, gadolinium, and iodine. We visualize the progress of the material decomposition by performing FBP reconstructions of different iterates. This allows us to, qualitatively, assess the progress of each material decomposition scheme.

Figure 4.8 depicts the reconstructed gadolinium image from a three-material water, gadolinum, and iodine decomposition. The result from two iterations of WLNLS-GN looks very similar to the fully-converged, gold-standard MLE. However, even after 9 iterations,

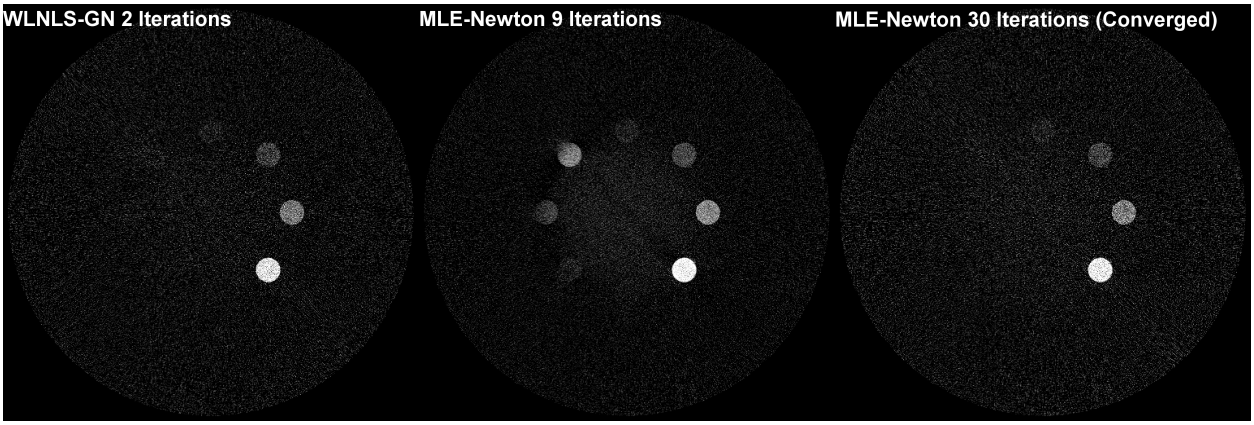


Figure 4.8: FBP reconstructions of WLNLS-GN and MLE-Newton iterates for gadolinium channel from water, gadolinum and iodine decomposition.

the MLE-Newton scheme still shows substantial errors (middle).

4.4 Discussion

By casting the material decomposition task as a weighted, nonlinear least squares problem and applying the Gauss-Newton algorithm, we are able to achieve similar estimation performance to the maximum likelihood approach but reaching accurate solutions in only a few iterations. Unlike Newton's method, the GN algorithm is a first-order method, so the complexity of each iteration is also reduced. Thus, we have presented a data-domain material decomposition scheme that seems to achieve similar accuracy to the gold-standard MLE, while being only slightly more computationally expensive than direct, non-iterative methods.

5

Multi-Channel Image Regularization

In Chapter 3, we discussed strategies for optimizing imaging parameters for spectral CT material classification tasks, and in Chapter 4, we presented a fast, accurate algorithm for performing data-domain material decomposition. In this chapter, we shift our focus to image reconstruction, specifically addressing how to formulate a cost function for the simultaneous reconstruction of multi-channel images, consisting of a plurality of energies or basis materials.

The dual-energy CT systems mentioned in Chapter 2 all utilize analytic algorithms (e.g. FBP and FDK), which act separately on each energy or material channel. This approach is attractive because it allows for the straight-forward application of well-understood reconstruction methods to DECT data and is fast enough to meet the demands of the clinical environment. However, in recent years, great strides have been made in the development of iterative reconstruction algorithms for x-ray CT, which are known to offer a number of advantages [75, 76, 77, 78], including the ability to incorporate prior knowledge about the object into the reconstruction model [79, 80]. In particular, for spectral CT, iterative algorithms offer the possibility of formulating a joint reconstruction model that exploits the structural similarities shared across image channels, which could lead to improved noise suppression and feature preservation. In this chapter, we design such a model by defining a multi-channel generalization of the popular total variation (TV) semi-norm.

5.1 Background

Typically, iterative reconstruction methods identify the desired image as the minimizer of a cost function comprising a data-fidelity term and some sort of regularization. The latter term can help prevent overfitting to noisy data, select a unique solution to an ill-posed inverse problem, and encourage desired properties in the image (e.g. smoothness). The total variation (TV) is one of the most widely used regularizers in imaging inverse problems. It was originally proposed by Rudin, Osher, and Fatemi for noise removal [81], but it has since been applied to a multitude of other problems, including deblurring [82], demosaicking [83], and tomographic reconstruction [84]. The success of the TV is due to several reasons. It was designed with the explicit intention of preserving sharp discontinuities while suppressing noise. Additionally, others have shown that TV regularized reconstruction algorithms can sometimes yield accurate reconstructions from highly undersampled measurement data [84]. This phenomenon may be partially explained in CT by the fact that images tend to be approximately piecewise constant, and the TV promotes sparsity in the image gradients.

A topic that has received more attention recently is how to generalize the total variation to multi-channel images [85, 86, 87, 41]. A multi-channel image may be viewed as a stack of scalar images or as a vector field, where each point in space corresponds to multiple contrast values. The most straightforward way to generalize the total variation to multi-channel images is to sum up the total variation of the individual channels. This technique, which we refer to as "channel-by-channel" TV, has been employed for spectral CT reconstruction previously [80, 88]. A major theoretical disadvantage of this approach is that it penalizes each image channel independently, despite the fact that there are generally large inter-channel correlations. For example, a simple visual inspection of a pair of low- and high-kVp images from the same phantom sugggests that there is a high degree of similarity (See Figure 5.1). Contrast changes appear to follow the same pattern, edges are in the same locations, and textures are likely to be similar. If some of these properties could be captured by a generalized, TV-like prior, then perhaps it would be advantageous to reconstruct both

channels in tandem.

It turns out that there are many ways to generalize the TV to multi-channel images, but we will focus on a specific choice, which has been referred to as the "total nuclear variation" ($\mathrm{TV_N}$) [41]. We had previously suggested the $\mathrm{TV_N}$ prior [89] for spectral CT, unaware that at least two earlier works had already discussed it in the context of color image restoration [87, 41]. The $\mathrm{TV_N}$ penalty encourages all of the image channels in a multi-channel reconstruction to have common edge locations and for their gradient vectors to point in a common direction. Equivalently, we can say that it has a bias for image channels that have parallel level sets [90]. The notion that multi-channel images should have gradient fields that share a common direction has been discussed in several other works on color image processing [91, 92, 90]. These multi-channel extensions of the TV are also commonly referred to as vectorial TV penalties since they are formulated to regularize vector-valued images.

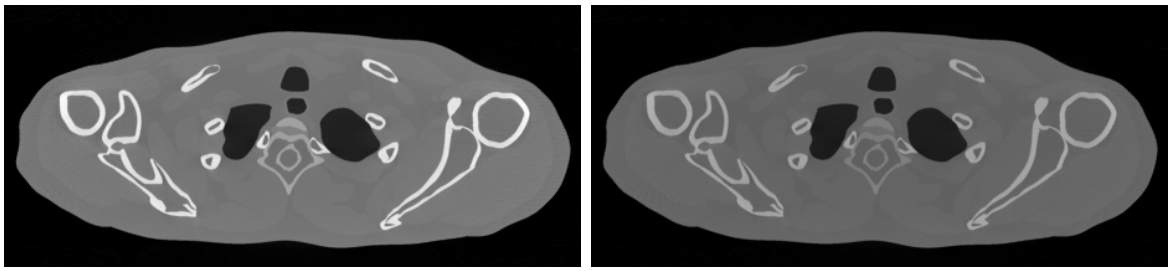


Figure 5.1: These 80 kVp (left) and 140 kVp (right) reconstructions of the numerical XCAT phantom suggest that images derived from common projection data acquired with different energy spectra are correlated. In particular they appear to have a very similar edge structure.

The development of effective regularization strategies is particularly important for spectral CT for a variety of reasons. Since the detected photons are divided into multiple energy bins, the SNR of each channel is immediately reduced. Further, the multi-energy data is typically used to derive quantitative, basis-material images (see Chapter 2), which may occur before [3], during [93], or after [94] the image reconstruction process. Unfortunately, the decomposition is usually quite ill-conditioned, which results in a substantial amplification of noise. This already poses significant challenges for protocols involving a two-material decomposition and will be even worse when trying to handle three or more basis materials [18].

Therefore, whether the BMD is performed in the sinogram domain or the image domain, it is important to mitigate as much noise as possible during the reconstruction.

In this chapter, we develop an iterative reconstruction model for jointly reconstructing multi-channel spectral CT data with the proposed TV_N prior. Using Chambolle and Pock's primal-dual algorithm (CPPD) [95], we derive update equations for two different optimization programs. The first minimizes the proposed TV_N subject to a constraint on the euclidean data-divergence, and the second minimizes the more primitive channel-by-channel TV with the same data constraint. This approach allows us to fairly compare these two generalized TV penalties over a variety of different smoothing levels, by adjusting a single parameter. We believe that in practice there may be good reasons for performing the basis-material decomposition either before or after the image reconstruction process, so we test the TV_N in both domains. In the first simulation study, we consider a 5-bin photon counting system and jointly reconstruct the corresponding five sinograms. We will refer to the energy-bin data hereafter as the "energy basis." In the second study, we investigate the impact of the TV_N in the "material basis." We simulate dual kVp data and perform a maximum-likelihood material decomposition to obtain a pair of bone and soft-tissue sinograms. We simultaneously reconstruct all four sinograms (low, high, bone, and soft-tissue) using both the channel-by-channel TV and TV_N optimization programs. We hypothesize that this hybrid approach of co-reconstructing the dual kVp data with the decomposed sinograms may help control noise in the material images by coupling them to the much higher SNR energy basis images.

5.2 Theory

Notation and definitions

In this section, we outline the notation that we will use for the remainder of the document. We adopt a somewhat unconventional set of indexing rules to maintain as much clarity as

possible in describing algebraic operations on vectors with multiple dimensions of spatial and spectral information.

Single-channel images. Consider a discretized 2D image $u \in I$ where $I = \mathbb{R}^{M \cdot N}$ is a finite dimensional vector space equipped with an inner product

$$\langle u, v \rangle_I = \sum_{i,j} u(i,j)v(i,j), \quad u, v \in I. \tag{5.1}$$

We use the convention $u(i,j)$ to refer to a particular pixel, where i and j specify the row and column indices and $u(i,j) \in \mathbb{R}$ is a discretized version of some continuous image function $u(x,y)$. The quantity ∇u is a vector in the space $G = I \times I$, where the operator $\nabla u : I \mapsto G$ represents a discrete approximation to the gradient. At each pixel location we define the quantity $(\nabla u)(i,j) \in \mathbb{R}^2$ as

$$(\nabla u)(i,j) = \begin{pmatrix} (\nabla u)^x(i,j) \\ \\ (\nabla u)^y(i,j) \end{pmatrix}, \tag{5.2}$$

where

$$(\nabla u)^x(i,j) = \begin{cases} u(i+1,j) - u(i,j) & j < N \\ \\ 0 & j = N \end{cases} \tag{5.3}$$

$$(\nabla u)^y(i,j) = \begin{cases} u(i,j+1) - u(i,j) & i < M \\ \\ 0 & i = M \end{cases}. \tag{5.4}$$

We also define an inner product in G,

$$\langle \boldsymbol{\nu}, \mathbf{z} \rangle_G = \sum_{i,j} \nu^x(i,j)z^x(i,j) + \nu^y(i,j)z^y(i,j), \quad \boldsymbol{\nu}, \mathbf{z} \in G. \tag{5.5}$$

Note that we use bold font to indicate that each spatial location (i, j) maps to a vector of values. Further, we will need a discrete divergence operator $\operatorname{div} \mathbf{z} : G \mapsto I$, which is chosen to be the negative transpose of the gradient operator, defined by

$$\langle \nabla u, \mathbf{z} \rangle_G = - \langle u, \operatorname{div} \mathbf{z} \rangle_I . \tag{5.6}$$

Lastly, we define the mixed ℓ_1/ℓ_2-norm in G as

$$\|\mathbf{z}\|_{1,2} = \sum_{i,j} \|\mathbf{z}(i,j)\|_2 \quad \mathbf{z} \in G, \tag{5.7}$$

indicating that we take an ℓ_1-norm over the spatial indices (i, j) and an ℓ_2-norm of each 2-vector $\mathbf{z}(i, j)$. This mixed-norm notation is often used in the literature on sparse regression and occasionally to compactly define the isotropic total variation [96],

$$\mathrm{TV}(u) = \|\nabla u\|_{1,2}. \tag{5.8}$$

Multi-channel images Now, we consider a discrete image $\mathbf{u} \in \mathcal{I}$ with L spectral or information channels,

$$\mathbf{u}(i,j) = \begin{pmatrix} u_1(i,j) \\ u_2(i,j) \\ \vdots \\ u_L(i,j) \end{pmatrix}, \tag{5.9}$$

where $\mathcal{I} = \mathbb{R}^{L \cdot M \cdot N}$. The quantity $\mathbf{u}$ is a discretized version of some continuous vector field $\mathbf{u}(x, y)$, such as an RGB color image. We also define an inner product in $\mathcal{I}$,

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathcal{I}} = \sum_{\ell} \sum_{i,j} u_\ell(i,j) v_\ell(i,j) \quad \mathbf{u}, \mathbf{v} \in \mathcal{I}, \tag{5.10}$$

where the subscript ℓ denotes a particular image channel. Since $\mathbf{u}$ is the discrete analog of a vector field, we can also define the discrete Jacobian, $J\mathbf{u} : \mathcal{I} \mapsto \mathcal{G}$, which generalizes the gradient operator to vector fields. In particular we have

$$(J\mathbf{u})(i,j) = \begin{pmatrix} (\nabla \mathbf{u}_1)^x (i,j) & (\nabla \mathbf{u}_1)^y (i,j) \\ (\nabla \mathbf{u}_2)^x (i,j) & (\nabla \mathbf{u}_2)^y (i,j) \\ \vdots & \\ (\nabla \mathbf{u}_L)^x (i,j) & (\nabla \mathbf{u}_L)^y (i,j) \end{pmatrix}. \tag{5.11}$$

At every pixel (i,j), the $L \times 2$ sub-matrix $(J\mathbf{u})(i,j)$ fully characterizes the first-order derivatives of $\mathbf{u}$, with each row consisting of the gradient vector of one of the L image channels. The quantity $J\mathbf{u}$ is in vector space $\mathcal{G} = \mathcal{I} \times \mathcal{I} = \mathbb{R}^{(L\times 2)\cdot(M\cdot N)}$ with an inner product

$$\langle V, Z \rangle_{\mathcal{G}} = \sum_{\ell} \sum_{i,j} V_\ell^x(i,j) Z_\ell^x(i,j) + V_\ell^y(i,j) Z_\ell^y(i,j) \quad V, Z \in \mathcal{G}. \tag{5.12}$$

A element $V \in \mathcal{G}$ is a discretized version of a tensor field $V(x,y)$, so we use uppercase font to indicate that every spatial location (i,j) maps to a matrix. Again, we need an analog of the divergence operator $\mathrm{Div}\, Z : \mathcal{G} \mapsto \mathcal{I}$ that is the negative transpose of J. It is constructed to satisfy

$$\langle J\mathbf{u}, Z \rangle_{\mathcal{G}} = - \langle \mathbf{u}, \mathrm{Div}\, Z \rangle_{\mathcal{I}}. \tag{5.13}$$

We define the mixed ℓ_1/nuclear-norm as

$$\|Z\|_{1,\star} = \sum_{i,j} \|Z(i,j)\|_\star \quad Z \in \mathcal{G}, \tag{5.14}$$

where $\|Z(i,j)\|_\star$ is referred to as the "nuclear norm" of matrix $Z(i,j)$ and is equal to the sum of its singular values.

Linear imaging model In this work we will assume that the reconstructed image $\mathbf{u}$ is related to the data vector $\mathbf{g}$ by the linear equation

$$\mathbf{g} = A\mathbf{u}, \tag{5.15}$$

where $\mathbf{g}$ is a vector in the vector space $\mathcal{D} \in \mathbb{R}^{L \cdot K}$, and K represents the total number of line integrals in our sinogram. The quantity $\mathbf{g}$ is formed by concatenating the sinograms of each channel, and the operator $A : \mathcal{I} \to \mathcal{D}$ is similarly formed by concatenating the projector model of each of the L channels,

$$\mathbf{g} = \begin{pmatrix} A_1 u_1 \\ A_2 u_2 \\ \vdots \\ A_L u_L \end{pmatrix}. \tag{5.16}$$

We define an inner product in $\mathcal{D}$,

$$\langle \mathbf{g}, \mathbf{q} \rangle_{\mathcal{D}} = \sum_{\ell} \sum_{k} g_\ell(k) q_\ell(k) \quad \mathbf{g}, \mathbf{q} \in \mathcal{D}. \tag{5.17}$$

Just as with the image domain, we use subscripts to index over spectral channels and the index k in parenthesis to identifies a specific ray in the projection space.

The Scalar Total Variation

Total variation regularized CT reconstruction algorithms have been studied extensively due to the approximate gradient sparsity of CT images. Several works have demonstrated that such schemes may yield accurate reconstructions from highly undersampled projection data [97, 84, 98, 99, 80]. However, the success of these sparse-view reconstruction methods is highly task-dependent.

The anisotropic TV The "anisotropic" TV is simply defined as the ℓ_1-norm of the derivative of the image u,

$$\mathrm{TV}_a(u) = \|\nabla u\|_1 = \sum_{i,j} \|(\nabla u)(i,j)\|_1, \tag{5.18}$$

where it is useful to think of the $\|.\|_1$ as a "sparsity-inducing" norm because it is a convex relaxation of $\|.\|_0$.

The isotropic TV Somewhat more common is the "isotropic" TV, which is defined in terms of the gradient-magnitude image,

$$\mathrm{TV}(u) = \|\nabla u\|_{1,2} = \sum_{i,j} \|(\nabla u)(i,j)\|_2. \tag{5.19}$$

This penalty function has similar properties to the anisotropic TV but is sometimes favored because it is rotationally invariant in the continuous setting. The anisotropic TV tends to favor horizontal and vertical edges [100].

Vectorial Total Variation

Now, we consider generalizing the TV to the $\mathcal{M}$-channel image $\mathbf{u}$.

The channel-by-channel TV. The simplest way to extend the definition of the TV to this vector-valued image, $\mathbf{u}$, would be to take the summation of the total-variation of each channel. This is given by

$$\mathrm{TV}_S(\mathbf{u}) = \sum_{\ell=1}^{L} \mathrm{TV}\left(u_\ell\right). \tag{5.20}$$

We will refer to this approach as the "channel-by-channel" TV because it imposes no coupling between image channels.

The Total Nuclear Variation. The proposed multi-channel generalization of the TV induces a tight coupling between image channels through a pixel-wise penalty on the rank of the Jacobian, $J\mathbf{u}$. This particular form, which we will call the "total nuclear variation" [41] is given by

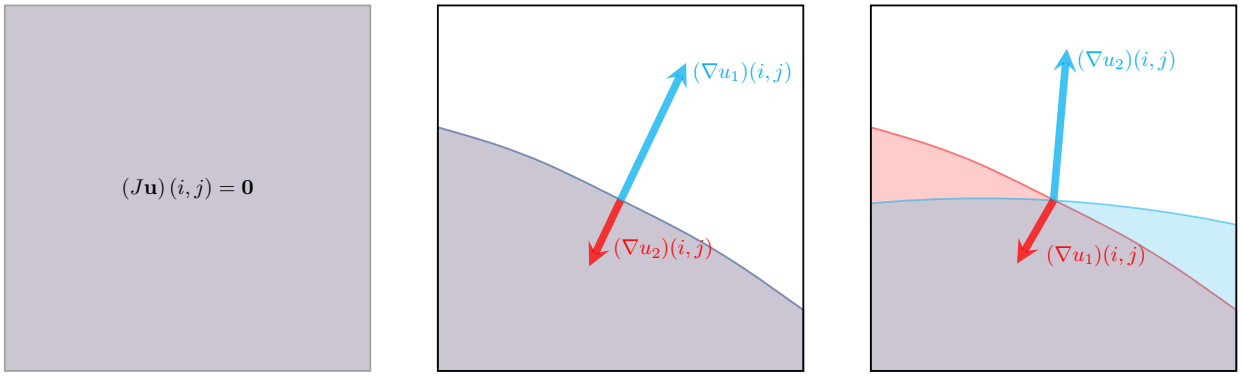
$$\mathrm{TV}_N(\mathbf{u}) = \|J\mathbf{u}\|_{1,\star} = \sum_{i,j} \| \left(J\mathbf{u}\right)(i,j)\|_\star, \tag{5.21}$$

where $\|(J\mathbf{u})_\ell\|_\star = \|\boldsymbol{\sigma}\|_1$, and $\boldsymbol{\sigma}$ is a vector of the singular values of $J\mathbf{u}$. This same form has been independently proposed by at least three different authors since 2013 [87, 41, 89]. These authors have pointed out several nice properties of the TV_N. Similarly to the channel-by-channel TV, it is convex and rotationally invariant when defined on the space of continuous images. Additionally, for a single-channel image, it reduces exactly to the usual isotropic TV.

However, unlike the channel-by-channel TV, the TV_N couples the different image channels by encouraging common edge locations and alignment of their gradient vectors. We also point out the following interesting equivalence: two images have gradient fields that point in the same direction (up to a sign difference) if and only if they share the same set of level curves [90].

This gradient-coupling effect comes from the fact that the nuclear norm encourages rank-sparsity in $(J\mathbf{u})_\ell$. A direct rank penalty on a matrix A with singular values $\boldsymbol{\sigma}$, would be $\|\boldsymbol{\sigma}\|_0$. However, this function is non-convex and numerically difficult to realize. Therefore, it is often replaced by the convex relaxation $\|\boldsymbol{\sigma}\|_1$, which also leads to sparsity in the singular values, under certain conditions [101, 102]. To see why we want to encourage a low-rank Jacobian, we consider the effect of the rank of this matrix on the gradient vectors of each image channel. If pixel ℓ lies within a constant-valued region of all of the image channels then this Jacobian matrix will be entirely null-valued, and all of its singular values will be zero. Further, if all of the gradient vectors of the various image channels are parallel or anti-parallel, such as at a common edge, then the Jacobian will be rank one and thus have only

one non-zero singular value. This is because parallel or anti-parallel vectors are not linearly independent. The fact that the TV_N treats parallel and anti-parallel vectors the same means that it is robust to contrast inversions. The connection between singular-value sparsity and gradient coupling is illustrated in Figure 5.2. In spectral CT, there may be edges that exist



(a) All singular values are zero

(b) One non-zero singular value

(c) Two non-zero singular values

Figure 5.2: This illustrates the relationship between the directionality of the gradient vectors of a 2-channel image and the singular values of its Jacobian. When both channels are constant valued, both singular values of the Jacobian are zero (a). If the gradient vectors are parallel or anti-parallel, one singular value is non-zero (b). Both singular values will be nonzero if the gradients have unique directions (c).

in some channels and not others, such as when imaging flowing contrast agents. However, the TV_N penalty shows no bias toward falsely propagating these edges into other channels where they do not belong. The TV_N would only be increased by inserting a false edge.

Converting to a "noise-balanced" image space In spectral CT imaging it is often the case that certain channels are significantly noisier than others, and we have found that the total nuclear variation regularization is more effective if a noise-balancing transform is applied prior to reconstruction or denoising. Consider a two-channel image, $\breve{\mathbf{u}}$, corrupted by zero-mean, Gaussian noise with variances σ_1^2 and σ_2^2 for channels 1 and 2, respectively. A common approach to denoising is to solve the following optimization problem:

$$\min_{\mathbf{u}} \quad \frac{1}{2}\|(u_1 - \breve{u}_1)/\sigma_1\|_2^2 + \frac{1}{2}\|(u_2 - \breve{u}_2)/\sigma_2\|_2^2 + \lambda\,\mathrm{TV}(\mathbf{u}). \tag{5.22}$$

The quantity $\breve{\mathbf{u}}$ is the original noise corrupted image, and $\mathbf{u}$ is the denoised image. In the data-fidelity term, each channel is inversely weighted by the standard deviation of the noise. This weighting scheme is statistically motivated because it ensures that as λ approaches zero, the likelihood of the resultant image, $\mathcal{L}(\mathbf{u}|\breve{\mathbf{u}})$, is monotonically increasing. We find that it is advantageous when using the TV_N penalty to also incorporate this inverse standard-deviation scaling into the regularization term. Therefore, we can define a new scaled image as $\mathbf{u'} \equiv (u_1/\sigma_1, u_2/\sigma_2)^T$ and solve the following optimization problem:

$$\min_{\mathbf{u'}} \quad \frac{1}{2}\|\mathbf{u'} - \breve{\mathbf{u}}'\|_2^2 + \lambda\,\mathrm{TV}(\mathbf{u'}). \tag{5.23}$$

The scaling ensures that the global noise levels are the same in both image channels, and the modified optimization problem effectively incorporates this scaling into the data-fidelity term as well as the regularization. For image reconstruction, the same type of scaling would be applied to the sinogram data. In general, the projection data will not have a uniform noise level within a single energy channel, so we use some average measure of channel noise to determine the global scale factor. We find that this noise-balancing procedure is an important step, as it improves the noise suppression in multi-spectral images with unequal noise levels. Note we can also write down an equivalent constrained optimization problem,

$$\begin{aligned} \min_{\mathbf{u'}} \quad & \mathrm{TV}_N(\mathbf{u'}) \\ \text{s.t.} \quad & \|\mathbf{u'} - \breve{\mathbf{u}}'\|_2 \leq \epsilon, \end{aligned} \tag{5.24}$$

where some mapping exists between λ and ϵ such that both schemes yield the same family of solutions.

5.3 Materials and methods

General reconstruction model

Assume we have a set of L sinograms, corresponding either to the directly measured energy channels of a spectral CT system or to the basis-material projections. We jointly reconstruct these channels by solving the following data-constrained optimization problem:

$$\begin{aligned} \min_{\mathbf{u}} \quad & \mathrm{TV}(\mathbf{u}) \\ \text{s.t.} \quad & \|A\mathbf{u} - \mathbf{g}\|_W \leq \epsilon, \end{aligned} \tag{5.25}$$

where $\mathbf{g}$ is the multi-channel projection data related to $\mathbf{u}$ by the discrete fan-beam projection operator A. The TV term refers to either the channel-by-channel TV_S or the total nuclear variation TV_N. The norm on the data constraint $\|.\|_W$ is a weighted ℓ_2 norm, defined by

$$\|g\|_W^2 = \langle g, Wg \rangle_{\mathcal{D}} \,. \tag{5.26}$$

This penalized weighted least squares (PWLS) data model is often used in CT reconstruction because when the weights are selected such that $W = K_{\mathbf{g}}^{-1}$, the data fidelity term corresponds to the likelihood function for Gaussian data with covariance $K_{\mathbf{g}}$. The single adjustable parameter ϵ controls the balance between data-fidelity and regularity. By fixing ϵ we can directly compare reconstructions using the naïve channel-by-channel TV_S to our proposed TV_N subject to the same data fidelity constraint.

The Chambolle Pock Primal Dual Algorithm

In this section, we will provide an overview of the first-order, primal-dual algorithm of Chambolle and Pock [95] and demonstrate how to apply it to our reconstruction model. For a tutorial on how to apply the CPPD algorithm to various CT reconstruction schemes, we refer the reader to Sidky et al [103].

The proximal mapping In order to describe the CPPD algorithm, we must first introduce the concept of a proximal mapping. For a function $f(x)$, the proximal mapping is defined by

$$\mathbf{prox}_f(x) = \arg\min_u \; f(u) + \frac{1}{2}\|u - x\|_2^2. \tag{5.27}$$

The proximal mapping can be interpreted as a generalized projection operator because for the special case that $f(x)$ is an indicator function, then it is a euclidean projection. For an extensive overview of the prox operator and applications, refer to Parikh and Boyd [104].

A general saddle-point problem Let X and Y be two finite-dimensional, real vector spaces. The CPPD algorithm is designed to solve the saddle-point problem described by

$$\min_{x \in X} \max_{y \in Y} \; \langle Kx, y \rangle - F(y) + G(x), \tag{5.28}$$

where $F(y)$ and $G(x)$ are convex functions with very simple proximal mappings, and K is a general linear operator. In particular, the proximal mappings associated with F and G should have a closed form or be easily solvable to high precision. It turns out that many interesting image processing and reconstruction problems fit this description, including the reconstruction model described in the previous section. The CPPD update equations are summarized by Algorithm 1.

Algorithm 1 CPPD Algorithm

Initialize: $\theta \in [-1, 1]$, $\sigma\tau\|K\|^2 < 1$, $x^{(0)}, \bar{x}^{(0)} \in X$, $y^{(0)} \in Y$ ▷ $\|K\| \equiv \sigma_{\max}(K)$
repeat
  $y^{(k+1)} = \mathbf{prox}_{\sigma F}\left(y^{(k)} + \sigma K \bar{x}^{(k)}\right)$
  $x^{(k+1)} = \mathbf{prox}_{\tau G}\left(x^{(k)} - \tau K^T y^{(k+1)}\right)$
  $\bar{x}^{(k+1)} = x^{(k+1)} + \theta\left(x^{(k+1)} - x^{(k)}\right)$
until $x^{(k+1)}, y^{(k+1)} = x^{(k)}, y^{(k)}$ ▷ stop when convergence criteria met

The parameters θ, σ, τ can be thought of as step-size parameters that affect convergence speed but not the final solution. In this work, we use $\theta = 1.0$ and $\sigma = \tau = \frac{1}{\|K\|}$. Though

Chambolle and Pock only prove convergence when the strict inequality $\sigma\tau\|K\|^2 < 1$ is satisfied, we find that setting $\sigma\tau\|K\|^2 = 1$ causes no stability problems. A similar observation is made in [103]. The quantity $\|K\|$ is the "spectral norm" of the operator K, which is equivalent to its largest singular value. As in [103] we use the power method to iteratively compute $\|K\|$, which relies only on matrix-vector multiplications and is detailed in Appendix D. The CPPD algorithm has close ties with several other popular algorithms, including ADMM, split-Bregman, and proximal-point. This is further discussed in [95].

Applying the CPPD algorithm to VTV reconstruction Now we will outline how the CPPD algorithm can be applied to the data-constrained optimization problem of (5.25). First we rewrite (5.25) as

$$\min_{\mathbf{u}} \quad \mathrm{TV}(\mathbf{u}) + \delta_\epsilon \left(A\mathbf{u} - \mathbf{g}\right), \tag{5.29}$$

where the indicator function $\delta_\epsilon(\mathbf{x})$ is defined by

$$\delta_\epsilon(\mathbf{x}) = \begin{cases} 0 & \|\mathbf{x}\|_W \leq \epsilon \\ \\ \infty & \|\mathbf{x}\|_W > \epsilon \end{cases}. \tag{5.30}$$

In the subsequent steps we will recast this objective into a saddle-point problem by dualizing the VTV and data-constraint terms. All of the following transformations follow from the definition of the Fenchel conjugate. More details on the Fenchel conjugate, also referred to as the Legendre transform, are in the appendix. First, we introduce an auxiliary variable $\mathbf{q}$ and rewrite the data constraint as

$$\delta_\epsilon(A\mathbf{u} - \mathbf{g}) = \max_{\mathbf{q}} \ \langle \mathbf{q}, A\mathbf{u} - \mathbf{g} \rangle_{\mathcal{D}} - \epsilon \|W^{-1/2}\mathbf{q}\|_2. \tag{5.31}$$

Similarly, we can rewrite the TV penalty function as an optimization over another auxiliary variable, $\mathbf{z}$. For the channel-by-channel TV_S, we have

$$\mathrm{TV}_S(\mathbf{u}) = \max_Z \ \langle J\mathbf{u}, Z\rangle_{\mathcal{G}} - \delta_{\mathcal{S}}(Z), \tag{5.32}$$

where the definition of the set $\mathcal{S}$ is given by

$$\mathcal{S} = \{Z \in \mathcal{G} \ : \ \|Z_\ell(i,j)\|_2 \leq 1 \ \forall \ i,j,m\}\,. \tag{5.33}$$

The proposed TV_N penalty function can be expressed as:

$$\mathrm{TV}_N(\mathbf{u}) = \max_Z \ \langle J\mathbf{u}, Z\rangle_{\mathcal{G}} - \delta_{\mathcal{N}}(Z), \tag{5.34}$$

where the set $\mathcal{N}$ is defined by

$$\mathcal{N} = \{Z \in \mathcal{G} \ : \ \sigma_{\max}(Z(i,j)) \leq 1 \ \forall \ i,j\}\,. \tag{5.35}$$

The quantity $Z(i,j)$ is an $L \times 2$ matrix, with the same dimensions as the Jacobian derivative at pixel (i,j), and $\sigma_{\max}$ is its largest singular value. These transformations, which are detailed in Appendix A, allow us to write our original VTV reconstruction model as a primal-dual, saddle-point problem,

$$\min_{\mathbf{u}} \max_{Z,\mathbf{q}} \ \langle J\mathbf{u}, Z\rangle_{\mathcal{G}} + \langle \mathbf{q}, A\mathbf{u} - \mathbf{g}\rangle_{\mathcal{D}} - \delta_{\mathcal{S}/\mathcal{N}}(Z) - \epsilon\|W^{-1/2}\mathbf{q}\|_2, \tag{5.36}$$

where the indicator function is either $\delta_{\mathcal{S}}$ for the channel-by-channel TV_S or $\delta_{\mathcal{N}}$ for the proposed TV_N. We can directly apply the CPPD update equations of Algorithm 1 by

making the following assignments.

$$x \to \mathbf{u}, \quad y \to \begin{pmatrix} \mathbf{q} \\ Z \end{pmatrix}, \quad K \to \begin{pmatrix} A \\ J \end{pmatrix} \tag{5.37}$$

$$F(\mathbf{q}, Z) = \delta_{\mathcal{S}/\mathcal{N}}(Z) - \epsilon \|W^{-1/2}\mathbf{q}\|_2 - \langle \mathbf{q}, \mathbf{g} \rangle, \quad G(\mathbf{u}) = 0$$

The resulting update equations are given by Algorithm 2. The operator $\Pi_{\mathcal{S}/\mathcal{N}}$ is a Euclidean

Algorithm 2 Data-constrained VTV Update Equations

1: Initialize: $\theta = 1, \;\; \sigma\tau \left\| \begin{pmatrix} A \\ J \end{pmatrix} \right\|_2^2 = 1, \;\; (\mathbf{u}, \bar{\mathbf{u}}, \mathbf{q}, Z) \to \mathbf{0}$
2: **repeat**
3: $\quad Z^{(k+1)} = \Pi_{\mathcal{S}/\mathcal{N}} \left(Z^{(k)} + \sigma J \bar{\mathbf{u}}^{(k)} \right)$
4: $\quad \mathbf{q}^{(k+1)} = \mathbf{prox}_{\sigma\epsilon F_q} \left(\mathbf{q}^{(k)} + \sigma \left(A\bar{\mathbf{u}}^{(k)} - \mathbf{g} \right) \right)$
5: $\quad \mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + \tau \left(\text{Div } Z^{(k+1)} - A^T \mathbf{q}^{(k+1)} \right)$
6: $\quad \bar{\mathbf{u}}^{(k+1)} = \mathbf{u}^{(k+1)} + \theta \left(\mathbf{u}^{(k+1)} - \mathbf{u}^{(k)} \right)$
7: **until** ▷ stop when convergence criteria met

projection onto the set $\mathcal{S}$ or $\mathcal{N}$, and the Div operator is the negative transpose of the discrete Jacobian operator J. Note that the $\mathbf{q}$-update is written in terms of the proximal map of the function F_q, which we define as $F_q = \|W^{(-1/2)}\mathbf{q}\|_2$. While there does not appear to be a closed-form expression for this proximal map, it can be reduced to a very simple scalar optimization problem as long as W is diagonal. This is detailed in the appendix along with an efficient method for performing the Euclidean projection in the primal variable update.

Simulation studies

To investigate the impact of the proposed vectorial TV regularization, we performed two numerical simulation studies with the computerized XCAT phantom [105]. In the first experiment we simulate an ideal, photon-counting system with 5 energy windows and directly reconstruct images corresponding to the log-normalized bin data. We will refer to this image

basis as the "energy" basis because each image channel corresponds directly to the measurements of one energy window.

In the second experiment, we simulate an ideal, dual-kVp scan and perform what we will refer to as a "hybrid" reconstruction. As we will detail later, we first decompose the 80/140 kVp data into a bone/soft-tissue "material" basis and then we co-reconstruct this synthetic data with the raw 80/140 kVp sinograms. The TV_N penalty couples all four image channels so that the relatively noisy basis-material channels may benefit from the higher SNR energy-bin channels.

Data generation

The XCAT shoulder phantom All simulations used the same 2D pixelized shoulder phantom, which was generated from an axial slice of the XCAT phantom and is pictured in Figure 5.3. The XCAT software package was used to generated a set of bone and soft-tissue density maps on a 2048×2048 pixel grid. These density maps were then used as input to a polyenergetic, distance-driven projector model to generate the simulated raw data that would be inputted to our reconstruction algorithm.

Ideal photon counting model To generate the 5-bin photon-counting data, we used a realistic 120 kVp simulated x-ray tube spectrum with hard thresholds at 40,60,80,100, and 120 keV. We did not model any physical factors in the detector response, so our bin response functions are perfect rect functions. Our simulated spectrum had virtually no emissions below 20 keV, so we can think of these bins as being evenly spaced. Using a distance-driven projector model, polyenergetic projections were computed with 896 detector elements per view and 400 views, and Poisson noise was added.

Dual kVp hybrid model The other configuration we looked at was an 80/140 dual kVp acquisition with the same number of detectors and views as the photon counting simulation. We simulated realistic 80 and 140 kVp x-ray tube spectra to generate two sets of consistent

projection data. This type of consistent, dual-kVp data can be acquired on many current scanners by performing two back-to-back scans. One could also approximate consistent, dual-kVp data with a fast kV-switching geometry by interpolating the missing views in the low and high kVp sinograms. Gaussian noise was added to approximate a compound Poisson noise model. From the noisy 80 and 140 kVp projection data, we synthesized bone and soft-tissue basis sinograms using a maximum-likelihood material decomposition. Since the material decomposition is performed in the projection space, it is ray-by-ray separable, resulting in a series of small optimization problems that we solve to high precision using Newton's method. In the "hybrid" study, we will co-reconstruct these 4 channels of data, consisting of our log-normalized, dual kVp data (energy basis) and the synthesized, material-basis data.

(a) 60 keV reference image (b) bone density map (c) soft-tissue density map

Figure 5.3: The 2D XCAT shoulder phantom used for all simulations studies, depicted at 60 keV (a) for reference. The projection data were generated using bone (b) and soft-tissue (c) density maps with the appropriate mass-attenuation curves.

Preprocessing and reconstruction

Both the 5-channel photon-counting data and the 4-channel dual kVp hybrid data were reconstructed using the reconstruction model outlined in (5.25). For reconstruction, we used a 512×512 grid of 1 mm pixels and a matched projector/backprojector pair based on Joseph's method [106].

Computing data-weights For the raw-basis projection data, we will use the data-weighting approach described by Bouman [76], which results in a diagonal W matrix, where the diagonal elements equal the raw projection data (before taking the logarithm). This is a quadratic

approximation to the log-likelihood. For the synthetic, bone/soft-tissue sinograms, we estimate the variances using the Cramer-Rao lower bound [32] and weight by their inverse. This is similar to the approach described by Schirra [107] and Sawatzky [88], but we ignore the off-diagonal terms in this work for simplicity. The reason for omitting these terms is because the CPPD algorithm relies on the functions F and G in (5.28) having simple proximal maps, and we require this simplification to meet that criteria. This problem only arises in the data-constrained form of the optimization problem. When solving the equivalent unconstrained optimization problem (as in [107, 88]), including the inter-channel noise correlations is straightforward. Generally, fewer algorithmic tools are available for solving the data-constrained form, but we favor it for this work because it provides a mechanism for fairly comparing the channel-by-channel TV to the $\mathrm{TV_N}$.

Tuning the data-constraint parameter The only parameter that we will vary in our reconstruction model (5.25) is ϵ, which controls the trade-off between data fidelity and smoothness. In general, smaller values of ϵ will result in noisier images, while higher values allow the VTV term to find a smoother solution. At the extreme case of $\epsilon = 0$, only images with projections that exactly match the measured data are allowed. There is also some finite value $\epsilon = \epsilon_{\max}$ for which the algorithm will return an image of all zeros. In order to determine an interesting range for ϵ, we first compute a ground truth image $\mathbf{u}_{\text{true}}$ by performing FBP on noiseless, non-sparse projection data. Then we define a new parameter ϵ^* which is defined by equation 5.38,

$$\epsilon^* \equiv \|A\mathbf{u}_{\text{true}} - \mathbf{g}\|_W \tag{5.38}$$

where $\mathbf{g}$ are the noisy projection data we wish to reconstruct. In this study we will select values of ϵ that satisfy $\epsilon = \alpha\epsilon^*$, where $\alpha \in (0, 1)$. We subjectively selected a range of α values that represent a range of solutions from under-smoothed to over-smoothed to demonstrate how the TV_N compares to the channel-by-channel TV in various regimes. As described earlier, the projection data were re-scaled prior to reconstruction, so that the average noise

levels were approximately the same in every spectral channel.

5.4 Results

In the following section, we present the resulting images from our simulation study, comparing directly the channel-by-channel TV and the proposed TV_{N}. In all of our comparisons, we refer to the reference image u_{true}, which is obtained by performing FBP on noiseless full-view data (1200 views).

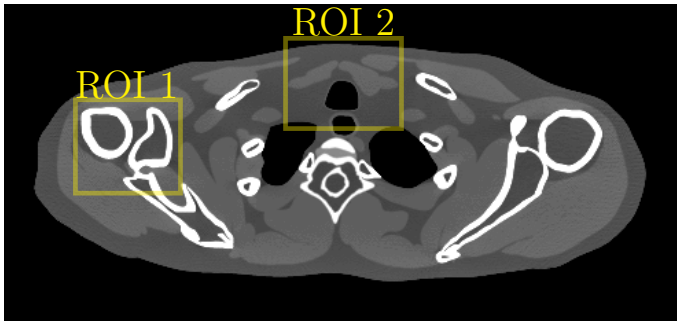
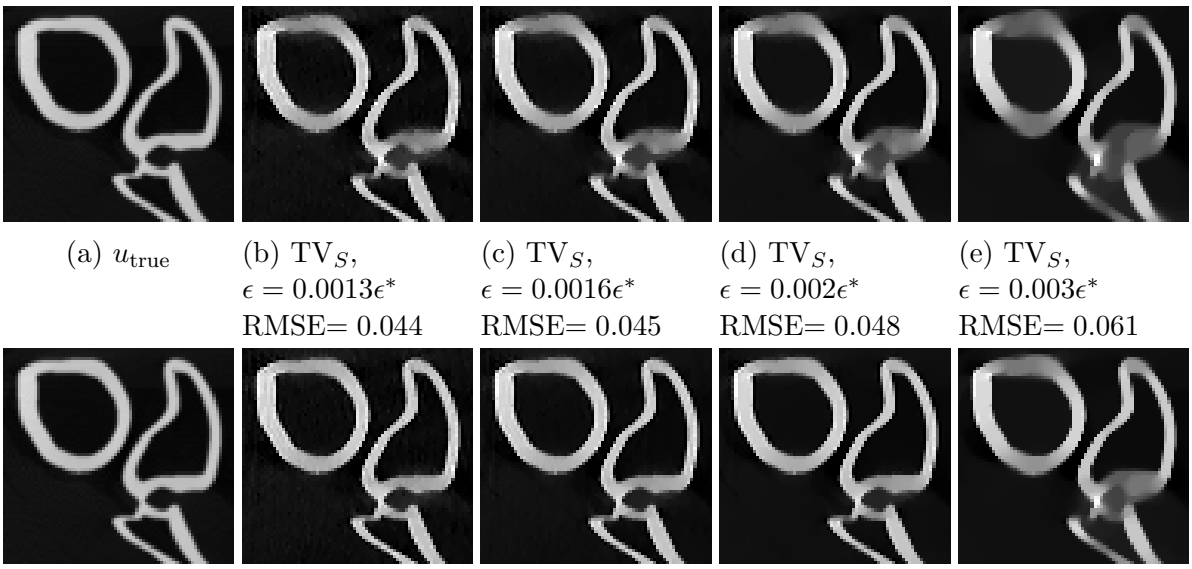


Figure 5.4: XCAT reference image, with ROI's indicated in yellow

Photon counting study

Here we present the results of performing a simultaneous reconstruction of the five photon-counting bins after log-normalization, using both the channel-by-channel TV_S and the proposed TV_N. This is an example of performing a joint reconstruction on data in the energy basis because we did not perform a material decomposition. In this setup, bin 1 (0-40 keV) was the noisiest, due to significant attenuation below 40 keV. We find that the inter-channel coupling introduced by the total nuclear variation has the greatest impact on the noisiest channels, so we present reconstructed images from this energy bin, focusing on the ROI's indicated in Figure 5.4. The resulting images are depicted in figures 5.5 through 5.6. The image window was manually adjusted for each ROI to highlight the relevant structures but is fixed within a particular figure. In general, we find that the images reconstructed with

TV_N regularization suffer from less edge blurring as the smoothing parameter ϵ is increased. The profile plot in Figure 5.7 gives a closer look at how the TV_N better preserves bony structures in the lowest energy bin image. We confirmed that these profiles were extracted from images of similar noise levels by measuring the sample variance in a nearby, uniform muscle ROI. This is expected because our data-constrained, reconstruction model allows us to select the noise level directly by tuning the ϵ parameter.



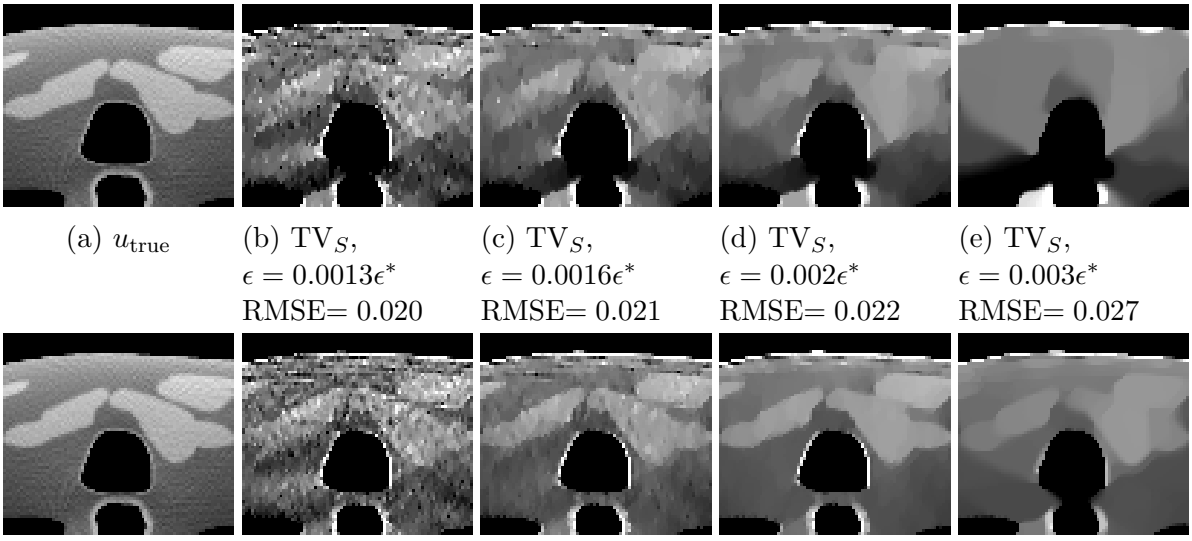
(a) u_{true} | (b) TV_S, $\epsilon = 0.0013\epsilon^*$ RMSE= 0.044 | (c) TV_S, $\epsilon = 0.0016\epsilon^*$ RMSE= 0.045 | (d) TV_S, $\epsilon = 0.002\epsilon^*$ RMSE= 0.048 | (e) TV_S, $\epsilon = 0.003\epsilon^*$ RMSE= 0.061

(f) u_{true} | (g) TV_N, $\epsilon = 0.0013\epsilon^*$ RMSE= 0.035 | (h) TV_N, $\epsilon = 0.0016\epsilon^*$ RMSE= 0.033 | (i) TV_N, $\epsilon = 0.002\epsilon^*$ RMSE= 0.033 | (j) TV_N, $\epsilon = 0.003\epsilon^*$ RMSE= 0.046

Figure 5.5: Bin 1 (0-40keV) image, bone ROI/window comparison with channel-by-channel TV (top) and TV_N (bottom). Grayscale window in cm^{-1} $[0.30, 0.85]$. The reference image u_{true} is an FBP reconstruction of the fully sampled (1200 views) noiseless data.

Dual kVp Hybrid Study

In this study, we co-reconstruct the synthetic soft-tissue/bone image channels with the raw 80/140 kVp data. In general, the material decomposition greatly amplifies noise, so that the soft-tissue and bone images tend to have a much lower CNR than the raw 80 and 140 kVp images. This noise amplification is due to the ill-conditioning of the inversion step in the basis change, which is caused by the relatively poor spectral separation between



(a) u_{true} (b) TV_S, $\epsilon = 0.0013\epsilon^*$ RMSE= 0.020 (c) TV_S, $\epsilon = 0.0016\epsilon^*$ RMSE= 0.021 (d) TV_S, $\epsilon = 0.002\epsilon^*$ RMSE= 0.022 (e) TV_S, $\epsilon = 0.003\epsilon^*$ RMSE= 0.027

(f) u_{true} (g) TV_N, $\epsilon = 0.0013\epsilon^*$ RMSE= 0.018 (h) TV_N, $\epsilon = 0.0016\epsilon^*$ RMSE= 0.018 (i) TV_N, $\epsilon = 0.002\epsilon^*$ RMSE= 0.019 (j) TV_N, $\epsilon = 0.003\epsilon^*$ RMSE= 0.018

Figure 5.6: Bin 1 (0-40keV) image, soft-tissue ROI comparison with channel-by-channel TV (top) and TV_N (bottom). Grayscale window in cm^{-1} $[0.30, 0.35]$. The reference image $*u_{\text{true}}$ is an FBP reconstruction of the fully sampled (1200 views) noiseless data.

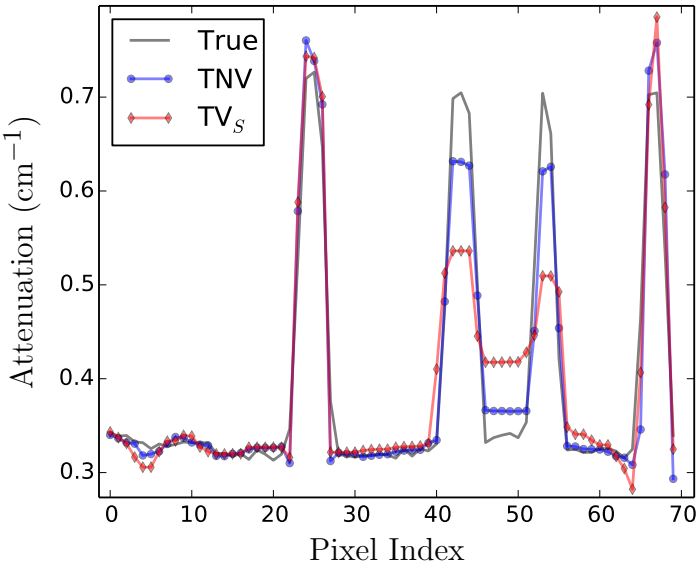
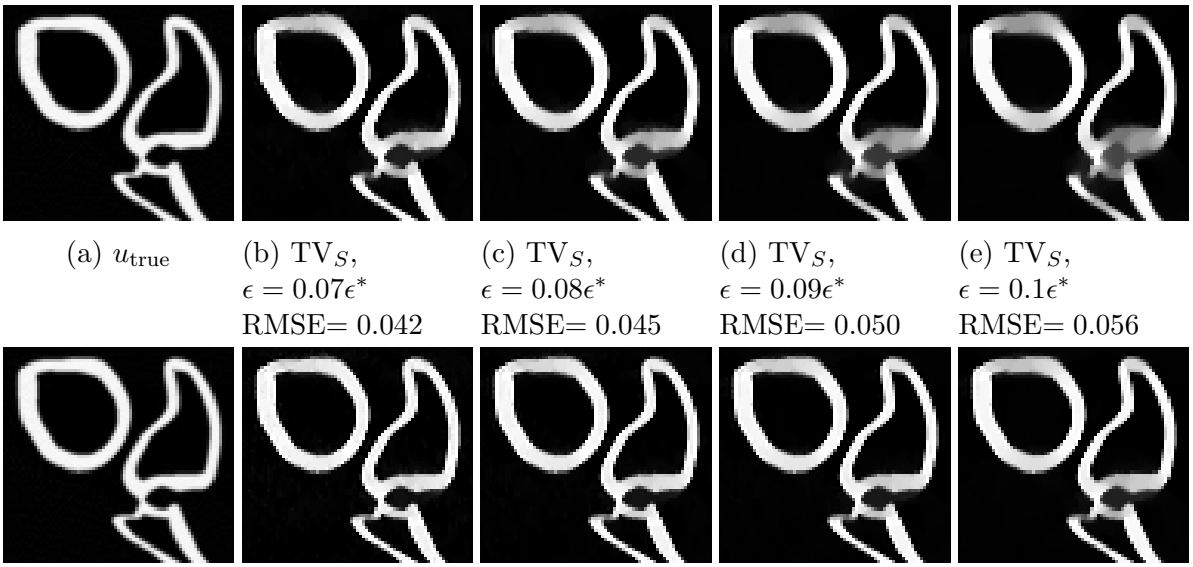


Figure 5.7: This is a vertical profile through the 0-40 keV, bone-ROI image with $\epsilon = 0.0016\epsilon^*$. The TV_N regularized reconstruction shows slightly better preservation of bony structures. The noise levels were estimated from a nearby ROI in a uniform muscle region ($\sigma_{TV_S} = 0.0025$, $\sigma_{\text{TV}_N} = 0.0022$).

basis materials. This poor spectral separation also explains the strong negative correlation between the synthesized material channels. We hypothesize that by coupling the high CNR raw image channels and the low CNR synthetic image channels, we may be able to improve noise suppression in the synthetic data. We call this technique of reconstructing the synthetic and raw data simultaneously "hybrid" reconstruction. We present ROI's from both the bone density image and the soft-tissue density image over a range of different ϵ values in Figures 5.8 and 5.9. We find that using the TV_N allows for a high degree of noise suppression (high

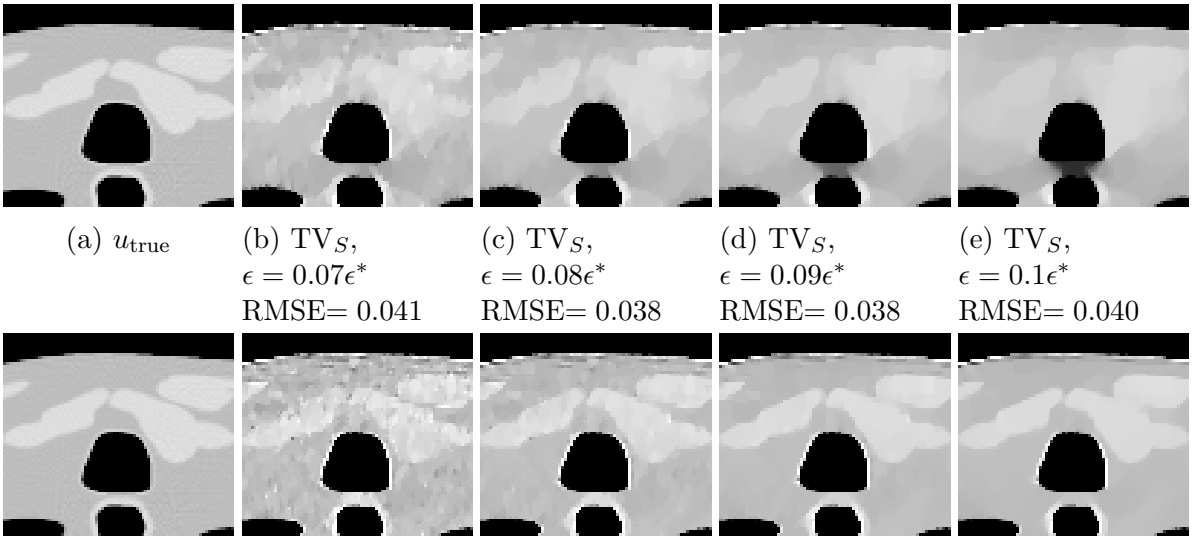


(a) u_{true} | (b) TV_S, $\epsilon = 0.07\epsilon^*$ RMSE= 0.042 | (c) TV_S, $\epsilon = 0.08\epsilon^*$ RMSE= 0.045 | (d) TV_S, $\epsilon = 0.09\epsilon^*$ RMSE= 0.050 | (e) TV_S, $\epsilon = 0.1\epsilon^*$ RMSE= 0.056

(f) u_{true} | (g) TV_N, $\epsilon = 0.07\epsilon^*$ RMSE= 0.040 | (h) TV_N, $\epsilon = 0.08\epsilon^*$ RMSE= 0.039 | (i) TV_N, $\epsilon = 0.09\epsilon^*$ RMSE= 0.039 | (j) TV_N, $\epsilon = 0.1\epsilon^*$ RMSE= 0.041

Figure 5.8: Bone basis image, bone ROI comparison with channel-by-channel TV (top) and TV_N (bottom). Grayscale window in g/mL^{-1} $[0.00, 0.52]$. The reference image u_{true} is an FBP reconstruction of the fully sampled (1200 views) noiseless data.

values of ϵ) without significantly deteriorating bone or soft-tissue structures. However, when the naïve channel-by-channel TV is used, these same ϵ values eventually lead to significant blurring artifacts. We also point out that even though the soft-tissue, 80 kVp, and 140 kVp images have edges that are not present in the bone density image, the TV_N coupling does not falsely propagate these edges into the bone channel.



(a) u_{true} (b) TV_S, $\epsilon = 0.07\epsilon^*$ RMSE= 0.041 (c) TV_S, $\epsilon = 0.08\epsilon^*$ RMSE= 0.038 (d) TV_S, $\epsilon = 0.09\epsilon^*$ RMSE= 0.038 (e) TV_S, $\epsilon = 0.1\epsilon^*$ RMSE= 0.040

(f) u_{true} (g) TV_N, $\epsilon = 0.07\epsilon^*$ RMSE= 0.040 (h) TV_N, $\epsilon = 0.08\epsilon^*$ RMSE= 0.039 (i) TV_N, $\epsilon = 0.09\epsilon^*$ RMSE= 0.039 (j) TV_N, $\epsilon = 0.1\epsilon^*$ RMSE= 0.037

Figure 5.9: soft-tissue basis image, soft-tissue ROI comparison with channel-by-channel TV (top) and TV_N (bottom). Grayscale window in g/mL^{-1} $[0.80, 1.07]$. The reference image u_{true} is an FBP reconstruction of the fully sampled (1200 views) noiseless data.

5.5 Discussion

We have described a framework for jointly reconstructing multi-channel spectral CT images using a generalized vectorial regularizer. Specifically, we presented a multi-channel generalization of the total variation which couples the different image channels by encouraging their gradient vectors to point in a common direction. Preliminary results suggest that this coupling may allow for greater noise suppression with less blurring of image structures compared to the channel-by-channel TV. This regularization strategy can be used to reconstruct energy bin images, basis-material images, or even both simultaneously, such as in our hybrid reconstruction study. This hybrid approach may allow for better noise suppression in the synthetic material images, which typically suffer from elevated noise levels.

In this study, we only considered the case of multi-energy data with consistent rays to isolate the impact of the inter-channel coupling in the TNV. However, we suspect that by coupling the edge structure of the image channels, there may be additional benefits for data

containing inconsistent rays. For example, in a fast kV switching system each energy channel contains slightly different geometric information about the object, due to the fact that the projection views are interleaved, and we hypothesize that the TV_N penalty may allow some of this information to be shared between image channels during the reconstruction. In future work we will investigate the impact of the TV_N penalty on configurations with inconsistent rays.

A final important note about this study is that the numerical XCAT phantom we used is piece-wise constant. We expect our proposed TV_N penalty to suffer from many of the same artifacts and limitations as the conventional TV penalty when applied to more realistic data, such as "staircasing" [108] and loss of contrast [109] . However, other works have demonstrated that TV can still perform somewhat well on data with low frequency structures and complex textures when balanced appropriately with the data-fidelity term [84, 110]. We investigate the application of TV_N to more realistic, non-piecewise CT data in Chapter 6 for the purposes of denoising, rather than image reconstruction.

5.6 Appendix A: Deriving the saddle-point problem

To derive the saddle-point formulation of our optimization problem, we used two fundamental results of convex analysis to "dualize" the data fidelity constraint as well as the VTV term. The transformation of the VTV term follows straightforwardly from the definition of the so-called "dual-norm," which is defined by

$$\|x\|' = \sup_z \ \langle x, z \rangle \ \ \text{s.t.} \ \|z\| \leq 1. \tag{5.39}$$

Every norm $\|.\|$ has an associated dual-norm $\|.\|'$, that obeys this relationship. The ℓ_2 norm utilized in the scalar TV is self-dual, while the nuclear norm and spectral norm form a dual pair.

Dualizing the channel-by-channel TV The channel-by-channel TV can be written as

$$\mathrm{TV}_S(\mathbf{u}) = \sum_{\ell} \sum_{i,j} \| \left(J\mathbf{u}\right)_\ell (i,j)\|_2. \tag{5.40}$$

The definition of the dual-norm allows us to rewrite this expression as

$$\begin{aligned} \mathrm{TV}_S(\mathbf{u}) &= \sum_{i,j} \sum_{\ell} \sup_{Z} \ \langle (J\mathbf{u})_\ell (i,j), Z_\ell(i,j)\rangle \ \text{ s.t. } \|Z_\ell(i,j)\|_2 \leq 1 \\ &= \sup_{Z} \ \langle J\mathbf{u}, Z\rangle_{\mathcal{G}} \ \text{ s.t. } \|Z_\ell(i,j)\|_2 \leq 1 \ \ \forall \ i,j,m \end{aligned} \tag{5.41}$$

which is equivalent to (5.32).

Dualizing the proposed TV_N The proposed TV_N is

$$\mathrm{TV}_N(\mathbf{u}) = \sum_{i,j} \|(J\mathbf{u})(i,j)\|_\star. \tag{5.42}$$

Substituting the definition of the dual-norm, we get

$$\begin{aligned} \mathrm{TV}_N(\mathbf{u}) &= \sum_{i,j} \sup_{Z} \ \langle (J\mathbf{u})(i,j), Z(i,j)\rangle \ \text{ s.t. } \sigma_{\max}(Z(i,j)) \leq 1 \\ &= \sup_{Z} \ \langle J\mathbf{u}, Z\rangle_{\mathcal{G}} \ \text{ s.t. } \sigma_{\max}\left(Z(i,j)\right) \leq 1 \ \ \forall\, i,j\,, \end{aligned} \tag{5.43}$$

which is equivalent to (5.34). We have used the fact that the spectral norm (maximum singular value) is dual to the nuclear norm.

Dualizing the data-fidelity term using the Fenchel conjugate To dualize the data-fidelity term, we make use of the Fenchel conjugate. For a convex, lower semi-continuous function f, the conjugate f^* is defined by

$$f^*(x) = \sup_{x'} \ \left\langle x, x' \right\rangle - f(x'), \tag{5.44}$$

where it is also true that $(f^*)^* = f$. Using this definition, it is easy to show that the following functions form a conjugate pair,

$$\delta_\epsilon(x) \Leftrightarrow \epsilon\|W^{-1/2}x\|_2 \tag{5.45}$$

where $\delta_\epsilon(x)$ is the indicator function, defined by (5.30). This along with the definition of the Fenchel conjugate leads directly to Eqn. 5.31.

5.7 Appendix B: The proximal map of $\epsilon\sigma\|W^{-1/2}\mathbf{q}\|_2$

The update equations for Algorithm 2 involve evaluating the proximal map of $\epsilon\sigma F_q \equiv \epsilon\sigma\|W^{-1/2}\mathbf{q}\|_2$, which does not admit a closed form. However, we will now show how it can still be efficiently evaluated using a simple root-finding procedure. First we explicitly write out the proximal mapping as

$$\mathbf{prox}_{\epsilon\sigma F_q}(\mathbf{q}') = \arg\min_{\mathbf{q}} \ \|W^{-1/2}\mathbf{q}\|_2 + \frac{1}{2\epsilon\sigma}\|\mathbf{q} - \mathbf{q}'\|_2^2. \tag{5.46}$$

Next, we compare this optimization problem to a slightly easier problem,

$$\arg\min_{\mathbf{q}} \ \frac{1}{2}\|W^{-1/2}\mathbf{q}\|_2^2 + \frac{1}{2\lambda\sigma}\|\mathbf{q} - \mathbf{q}'\|_2^2, \tag{5.47}$$

and note that for some choice of ϵ and λ, these objectives both have the same set of level curves. We can find the solution to this problem by setting the gradient equal to zero, and for a symmetric weighting matrix W, it is given by

$$\mathbf{q}^* = \left(I + \sigma\lambda W^{-1}\right)^{-1}\mathbf{q}'. \tag{5.48}$$

If the gradient of this easier problem is equal to that of the original problem, then $\mathbf{q}^*$ must be an optimizer of the original problem as well. Momentarily, we ignore the non-differentiable

point $\mathbf{q} = \mathbf{0}$, and set the gradients equal, which yields

$$\begin{aligned} \epsilon &= \lambda \|W^{-1/2}\mathbf{q}^*\|_2. \\ &= \lambda \|W^{-1/2}\left(I + \sigma\lambda W^{-1}\right)^{-1}\mathbf{q}'\|_2. \end{aligned} \tag{5.49}$$

In this work, we will only consider the case where W is a diagonal matrix, which allows us to simplify this expression.

$$\epsilon^2 - \lambda^2 \sum_{j=1}^{J} \frac{W_{jj}}{\left(W_{jj} + \sigma\lambda\right)^2} [\mathbf{q}']_j^2 = 0. \tag{5.50}$$

The index j is a linear index over every element of $\mathbf{q}$. Now, we just need to perform a 1D search for a positive value of λ that satisfies this equation, which can be done very efficiently using Newton's method or a variety of other algorithms. Once this root, λ^*, is found, we simply plug it back into (5.49):

$$\mathbf{prox}_{\epsilon\sigma F_q}(\mathbf{q}') = \begin{cases} \left(I + \sigma\lambda^* W^{-1}\right)^{-1}\mathbf{q}' & \|W^{1/2}\mathbf{q}'\|_2 > \sigma\epsilon \\ 0 & \|W^{1/2}\mathbf{q}'\|_2 \leq \sigma\epsilon \end{cases}. \tag{5.51}$$

Note that because of the non-differentiability of $\|W^{-1/2}\mathbf{q}\|_2$ at $\mathbf{q} = 0$, there will not always be a solution to equation 5.50. Specifically, if $\sigma\epsilon > \|W^{1/2}\mathbf{q}'\|_2$, then there is no viable root. In this case, we need not perform the root-finding procedure, because the optimum must occur at $\mathbf{q} = \mathbf{0}$.

5.8 Appendix C: Implementation of the projection operators $\Pi_{\mathcal{S}/N}$

$\mathbf{\Pi_{\mathcal{S}}}$ Now, we will describe the Euclidean projection onto the set $\mathcal{S}$ defined in (5.33). Consider projecting $Z \in \mathcal{G}$ onto $\mathcal{S}$. We can define this operation element-wise on each vector

$Z_\ell(i,j) \in \mathbb{R}^2$ corresponding to the ℓ^{th} spectral channel and the pixel-location (i,j). The projection of this element is given by

$$\Pi_{\mathcal{S}}\left(Z_\ell(i,j)\right) = \begin{cases} Z_\ell(i,j) & \|Z_\ell(i,j)\|_2 \leq 1 \\ \frac{Z_\ell(i,j)}{\|Z_\ell(i,j)\|_2} & \text{otherwise} \end{cases}. \tag{5.52}$$

For every image channel ℓ and pixel location (i,j), we simply project the vector $Z_\ell(i,j)$ onto the unit ball.

$\mathbf{\Pi_N}$ The projection onto set N defined by (5.35) is slightly more complicated. We define this operation element-wise on each $\mathcal{M} \times 2$ matrix $Z(i,j)$. This time we need to threshold the singular values of $Z(i,j)$. This projection can be succinctly described by

$$\Pi_N\left(Z(i,j)\right) = U\Sigma_p V^T, \tag{5.53}$$

where $U\Sigma V^T$ is the SVD of $Z(i,j)$, and

$$\Sigma_p = \text{sgn}\left(\Sigma\right)\min\left(|1|, |\Sigma|\right) \tag{5.54}$$

is the thresholded version of Σ. To form Σ_p the singular values on the diagonal of Σ are simply thresholded so that their magnitude does not exceed 1. An equivalent projection formula that leads to a much more computationally efficient solution is given by

$$\Pi_N\left(Z(i,j)\right) = Z(i,j)V\Sigma^\dagger\Sigma_p V^T. \tag{5.55}$$

The quantity $\Sigma^\dagger$ is the pseudo-inverse of Σ. Since we are only working with 2 spatial dimensions the matrix V will by a 2×2 square matrix. Therefore, to compute the projection according to (5.55) we only need to compute the eigenvalues and eigenvectors of a 2×2 matrix, for which a very simple closed form is available. This can also be done efficiently for

3D images, where V will be 3×3. Because of this, the update equations that result from the proposed vectorial TV are only trivially more expensive than the channel-by-channel TV, and the projection/backprojection operations are likely to swamp this difference.

5.9 Appendix D: Power Method for Estimating the Maximum Singular Value of a Linear Operator

Algorithm 3 is a basic version of the power method, which can be used to numerically estimate the maximum singular value of a linear operator K. In practice, 10 to 20 iterations are usually sufficient, and this quantity only needs to be computed once for a given system geometry. The vector x is typically initialized in some random fashion, e.g. with samples of the standard normal distribution.

Algorithm 3 N steps of the power method for determining $s = \|K\|$

Initialize: $n \leftarrow 0$
repeat
    $x_{n+1} \leftarrow \frac{K^T K x_n}{\|K^T K x_n\|_2}$
    $s \leftarrow \|K x_{n+1}\|_2$
    $n \leftarrow n + 1$
until $n = N$

6

Assessment of VTV Regularization on Realistic CT Data

6.1 Background

In the previous chapter, we introduced a vectorial extension of the total variation penalty to multi-channel images, which we referred to as the total nuclear variation ($\mathrm{TV_N}$). By jointly processing several image channels with edge-coupling regularization, improved reconstructions were demonstrated from the same noisy data.

In this chapter, we will build on those findings by investigating whether the same benefits can be demonstrated on more realistic CT data. In doing so, we aim to address several limitations of the previous study. Firstly, we will evaluate the TV_N on images that have a realistic texture. Although the numerical XCAT phantom, utilized in the previous study, is based closely on real CT images of a human cadaver, its contrast is piecewise. This could be problemetic for evaluating any sort of TV regularization, since it is known to favor this kind of structure [111]. Furthermore, one of the alleged benefits of the $\mathrm{TV_N}$ is that it promotes aligned gradient vectors across energy channels, but due to its piecewise nature, the XCAT phantom also trivially satisfies this condition. Other authors have applied a very similar

regularizer to the joint reconstruction of PET-MR patient data with favorable results [112], but to our knowledge, no studies have reported on TV_N for real multi-energy CT images.

To address the question of whether the previous findings are translatable to realistic multi-energy CT data, we will investigate the proposed regularizer on dual-energy images of a turkey acquired on a diagnostic, clinical scanner. While the previous study focused on image reconstruction, the $\mathrm{TV_N}$ can also be applied to other image restoration problems [41, 87]. Since we did not have access to the projection data or geometry parameters of our scanner, we focus solely on denoising in this work. Because denoising is much less computationally expensive than reconstruction, this enabled us to conduct thorough exploration of the parameter space which was not feasible in the previous study. In addition to assessing the $\mathrm{TV_N}$ on realistic data, we also address several other issues of practical interest. Firstly, in order to isolate the importance of edge-coupling and gradient alignment, we also study another multi-channel variant of the TV[82] which encourages joint sparsity of the gradient magnitude images but is ambivalent to the alignment of the gradient fields; we call this $\mathrm{TV_F}$. Secondly, we demonstrate a "smoothed" variant of the multi-channel TV, analgous to the "Huber penalty" used in other image reconstruction papers [78, 113, 114, 115, 116]. The Huber penalty is an interesting regularizer for CT because it has some of the edge-preserving properties of the TV but with less of bias toward piecewise images, which may lead to a more natural looking texture [95]. Thus, a vectorial extension of the Huber penalty to multi-channel images may be of practical interest. Lastly, we demonstrate that the $\mathrm{TV_N}$ can be efficiently extended to 3D image volumes, which could improve denoising by allowing regularization along the axial direction and is crucial for reconstructing multidector and cone beam CT data.

6.2 Theory

In this section, we will describe a generic denoising model as well the variants of the TV that we investigated. We use the same notation as in the previous chapter unless otherwise indicated.

Denoising Model

Our denoising model is specified by the following optimization problem:

$$\arg\min_{\boldsymbol{u}} \quad \frac{1}{2}\|\boldsymbol{u}-\boldsymbol{g}\|_2^2 + \lambda\,\mathrm{TV}(\boldsymbol{u}) \tag{6.1}$$

where $\boldsymbol{g}$ is the noisy image, λ controls the tradeoff between data-fidelity and smoothness, and TV refers generally to some variant of the total variation from $\mathrm{TV_S}$, $\mathrm{TV_F}$, or $\mathrm{TV_N}$. Both $\mathrm{TV_S}$ and $\mathrm{TV_N}$ were defined in Chapter 5, and $\mathrm{TV_F}$ is defined as

$$\mathrm{TV_F}(\boldsymbol{u}) = \|J\boldsymbol{u}\|_{1,F} = \sum_{i,j} \|J\boldsymbol{u}(i,j)\|_F, \tag{6.2}$$

where $\|.\|_F$ is the Frobenius norm. For a matrix X the Frobenius norm is defined as

$$\|X\|_F = \left(\sum_{i,j} X_{ij}^2\right)^{1/2}. \tag{6.3}$$

All VTV variants studied are summarized in Table 6.1.

Primal-Dual Saddlepoint Problem As in Chapter 5, the denoising model can be expressed as a saddlepoint problem, which is given by

$$\min_{\boldsymbol{u}} \max_{Z} \ \langle J\boldsymbol{u}, Z\rangle_{\mathcal{G}} - \delta_{\mathcal{K}}(Z) + \frac{1}{2\lambda}\|\boldsymbol{u}-\boldsymbol{g}\|_2^2, \tag{6.4}$$

Defintion	Edge Coupling	Gradient Alignment
$\mathrm{TV_S}(\boldsymbol{u}) = \sum_{\ell=1}^{L} \mathrm{TV}(u_\ell)$		
$\mathrm{TV_F}(\boldsymbol{u}) = \|J\boldsymbol{u}\|_{1,F}$	✓	
$\mathrm{TV_N}(\boldsymbol{u}) = \|J\boldsymbol{u}\|_{1,\star}$	✓	✓

Table 6.1: Three multi-channel variants of the TV studied in this work. Note, $\mathrm{TV_S}$ induces no inter-channel coupling and is equivalent to separately denoising each image channel.

where $\delta_{\mathcal{K}}$ is a generic set indicator function defined by

$$\delta_{\mathcal{K}}(\boldsymbol{X}) \equiv \begin{cases} 0 & \boldsymbol{X} \in \mathcal{K} \\ \\ \infty & \boldsymbol{X} \notin \mathcal{K} \end{cases} . \tag{6.5}$$

The set $\mathcal{K}$ is equal to $\mathcal{S}$, $\mathcal{F}$, or $\mathcal{N}$ for $\mathrm{TV_S}$, $\mathrm{TV_F}$, and $\mathrm{TV_N}$, respectively. The sets $\mathcal{S}$, $\mathcal{F}$, and $\mathcal{N}$ are defined below:

$$\begin{aligned} \mathcal{S} &= \{Z \in \mathcal{G} \ : \ \|Z_\ell(i,j)\|_2 \leq 1 \ \forall \ i,j,\ell\} \\ \mathcal{F} &= \{Z \in \mathcal{G} \ : \ \|Z(i,j)\|_F \leq 1 \ \forall \ i,j\} \\ \mathcal{N} &= \{Z \in \mathcal{G} \ : \ \sigma_{\max}\left(Z(i,j)\right) \leq 1 \ \forall \ i,j\}. \end{aligned} \tag{6.6}$$

Algorithm Updates

Just as in the previous chapter, we use Chambolle and Pock's first-order, primal-dual algorithm (CPPD). However, unlike in Chapter 5, this saddle point problem is strongly convex in the primal variable $\boldsymbol{u}$, with convexity constant λ^{-1}, so we can use their accelerated algorithm. This variant achieves a convergence rate of $O\left(1/N^2\right)$ rather than $O(1/N)$. The updates are given by Algorithm 4.

The primal and dual stepsizes τ and σ must be initialized so that $\tau\sigma\|J\|^2 \leq 1$, where $\|J\|$ is the spectral radius (also the maximum singular value) of the Jacobian operator J. It can be shown in that $\|J\| \leq \sqrt{8}$ [117]. We find that convergence is always fastest when $\sigma\tau = 1/8$;

however there still exists freedom to choose the ratio of primal to dual stepsizes. The observed rate of convergence can be dramatically impacted by this choice, but we are unaware of any good method to automatically select these parameters. We have used $\tau = 0.05$ for everything in this chapter. We iterate until the primal-dual gap falls below 10^{-6}, which we have found, empirically, to be a reasonable stopping point where no additional pixel value changes can be observed in any of the windows we looked at.

Algorithm 4 TV Denoising Update Equations

1: Initialize: $\sigma_k \tau_k = 1/8$

2: **repeat**

3: $\quad Z^{(k+1)} = \Pi_{\mathcal{K}} \left(Z^{(k)} + \sigma_k J \bar{\mathbf{u}}^{(k)} \right)$

4: $\quad \mathbf{u}^{(k+1)} = (\lambda \mathbf{u}^{(k)} + \tau_k \boldsymbol{g} + \lambda \tau_k \text{Div } Z^{(k+1)})/(\lambda + \tau_k)$

5: $\quad \theta_k = \sqrt{\lambda/(\lambda + 2\tau_k)}, \quad \tau_k = \theta_k \tau_k, \quad \sigma_k = \sigma_k/\theta_k$

6: $\quad \bar{\mathbf{u}}^{(k+1)} = \mathbf{u}^{(k+1)} + \theta_k \left(\mathbf{u}^{(k+1)} - \mathbf{u}^{(k)} \right)$

7: **until** ▷ stop when convergence criteria met

Smoothing the VTV

One feature of the TV that may be undesirable for certain tasks is the tendency to bias images toward a piecewise constant structure. This may cause artificially blocky transitions between contrast values, rather than smooth gradations. These artifacts are sometimes called "staircasing" artifacts[108] and become more significant for higher values of λ. One approach that has been suggested to mitigate this effect is to replace the TV with a smooth approximation, utilizing the Huber loss function[95]. This results in a penalty with behavior that falls somewhere between the TV and a quadratic roughness penalty. This sort of penalty has been used many times, previously, for CT reconstruction and sinogram restoration and is sometimes referred to as a "Huber penalty."

We point out that this same principle can be extended to both of the vectorial TV

variants studied in this work, resulting in a vectorial Huber penalty, suitable for processing multi-channel images. For example, $\mathrm{TV_F}$ can be approximated by the smooth function

$$\mathrm{TV}_{\mathrm{F},\mu}(\boldsymbol{u}) = \sum_{i,j} h_\mu(\|J\boldsymbol{u}(i,j)\|_F), \tag{6.7}$$

where we have simply replaced the Frobenius norm $\|.\|_F$ with the $h_\mu(\|.\|_F)$, where $h_\mu(x)$ is the Huber loss function, defined as

$$h_\mu(x) = \begin{cases} |x|^2/(2\mu) & |x| \leq \mu \\ |x| - \mu/2 & |x| > \mu \end{cases}. \tag{6.8}$$

The smoothed approximation of $\mathrm{TV_N}$ is defined by replacing the nuclear norm, $\|X\|_\star = \sum_k \sigma_k$, with $\sum_k h_\mu(\sigma_k)$. Both of these smoothed vectorial TV variants can be summarized by the following saddlepoint problem:

$$\min_{\boldsymbol{u}} \max_{Z} \ \langle J\boldsymbol{u}, Z\rangle_{\mathcal{G}} - \delta_{\mathcal{K}}(Z) + \frac{1}{2\lambda}\|\boldsymbol{u} - \boldsymbol{g}\|_2^2 + \frac{\mu}{2}\|Z\|_2^2, \tag{6.9}$$

where $\|Z\|_2^2 = \langle Z, Z\rangle_{\mathcal{G}}$. Note the only modification from the original saddlepoint problem is the last term, and we can recover the original TV denoising model by setting $\mu \to 0$.

The algorithm updates are also very similar, requiring only a modification to the dual variable update. However, the addition of extra smoothing term makes the saddlepoint problem strongly convex in both $\boldsymbol{u}$ and Z, so a linearly convergent (i.e. $O(\omega^N)$, $\omega < 1$) variant of the CPPD algorithm can be used, referred to as ALG3 in [95].The updates are given in Algorithm 5.

Algorithm 5 Smoothed TV Denoising Update Equations

1: Initialize: $t = \sqrt{\mu/(2\lambda)}, \quad \tau = \lambda t/2, \quad \sigma = t/(2\mu), \quad \theta = 1$

2: **repeat**

3: $\quad Z^{(k+1)} = \Pi_{\mathcal{K}}\left((Z^{(k)} + \sigma J\bar{\mathbf{u}}^{(k)})/(1+\sigma\mu)\right)$

4: $\quad \mathbf{u}^{(k+1)} = \left(\lambda\mathbf{u}^{(k)} + \tau\boldsymbol{g} + \lambda\tau\text{Div } Z^{(k+1)}\right)/(\lambda+\tau)$

5: $\quad \bar{\mathbf{u}}^{(k+1)} = \mathbf{u}^{(k+1)} + \theta\left(\mathbf{u}^{(k+1)} - \mathbf{u}^{(k)}\right)$

6: **until** ▷ stop when convergence criteria met

6.3 Denoising Experiments

Data Acquisition

A series of scans were acquired of a frozen supermarket turkey at 80, 120 and 140 kVp on a 256-slice MDCT (Brilliance iCT, Philips Healthcare) in axial mode. At each kVp, we acquired both a "low-dose" scan (100mAs) and ten "high dose" scans (410 mAs for 80 kVp and 425 mAs for 120/140 kVp). The high-dose scans were repeated so that the resulting images could be averaged to create a very low-noise "ground truth" image. All 33 scans were acquired with the exact same table position and no other parameters were altered between acquisitions besides the kVp and mAs.

All of the scans were reconstructed with FBP (high resolution mode, Filter E) via the manufacturer's software and using the exact same parameters: $512\times512\times128$ image volumes with 0.562mm pixels and a 0.625mm slice thickness. For the 2D denoising studies, a single slice was extracted from the midpoint of the axial stack (see Figure 6.1). For the 3D denoising studies, linear interpolation was used along the z-axis (axial direction) to synthesize a volume with isotropic resolution (0.562mm) to avoid complications with the discrete gradient and divergence operators. Since the scan was acquired in axial mode, the slices toward the axial extremes of the acquired data volume exhibited severe cone beam artifacts. A smaller volume consisting of only the middle 32 slices was extracted in order to minimize these effects.

Figure 6.1: The middle slice used as the ground-truth for the 2D denoising studies at 80, 120, and 140kVp. These images were formed by averaging the 10 high-dose scans.

Comparison of $\mathrm{TV_S}$, $\mathrm{TV_F}$, and $\mathrm{TV_N}$

In order to compare these vectorial TV variants, we synthesized a dual kVp image pair by extracting the same 2D slice from the 80 and 140 kVp low-dose (100mAs) datasets. Then we denoised this dual-energy data over a wide range of λ values with each of $\mathrm{TV_S}$, $\mathrm{TV_F}$, and $\mathrm{TV_N}$; we first identified $\lambda_{\min}$ and $\lambda_{\max}$, corresponding to images that appeared under- and over-smoothed and sampled 1000 equally spaced values within this interval. Each result was then compared to the corresponding 10X averaged high-dose image, and the normalized Euclidean distance was computed as an error metric. This is defined as

$$\epsilon(\lambda) = \frac{\|u^*(\lambda) - u_{\mathrm{ref}}\|_2}{\|u_{\mathrm{ref}}\|_2}, \tag{6.10}$$

where $u^*(\lambda)$ is the denoised image with parameter λ, and u_{ref} is the corresponding 10X high-dose reference image. In the case of $\mathrm{TV_F}$ and $\mathrm{TV_N}$, the dual-energy data is processed jointly due to the interchannel coupling of the regularizer, but we report the error metrics individually for each channel. This is shown in Figure 6.2.

In general, the error curves should not be compared at any one specific common value of λ because the functions, $\mathrm{TV_S}$, $\mathrm{TV_F}$, and $\mathrm{TV_N}$ have inherently different scales. Instead the curves should be evaluated in terms of the possible family of solutions they each represent and, most importantly, their minimum-error solution. As can be seen in Figure 6.2, the

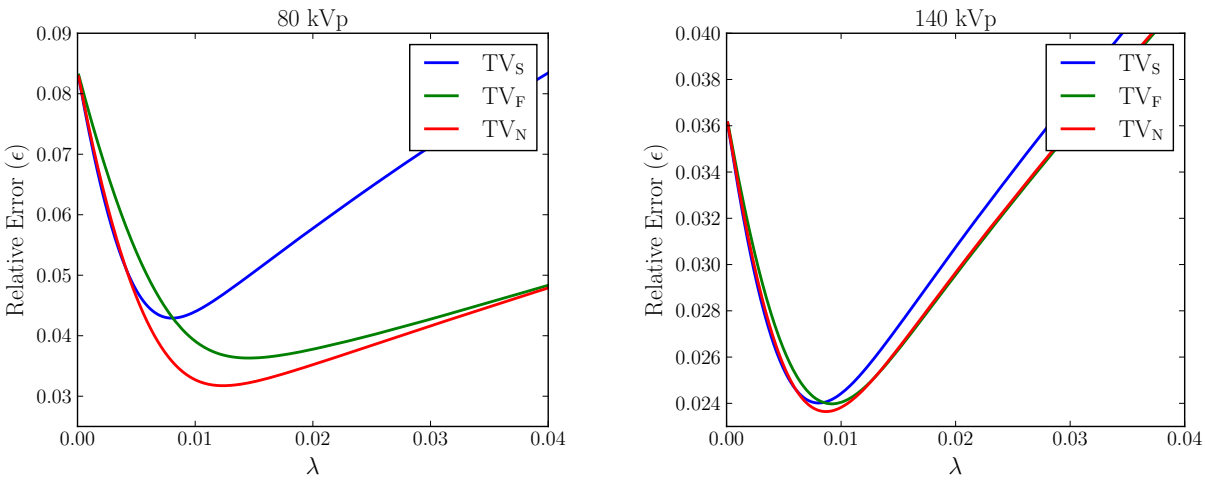


Figure 6.2: Relative error curves based on (6.10) are shown for the three different VTV variants studied for the low-dose (100mAs) study. Optimal values of λ are identified by the minima of these curves.

channel-coupling of the VTV has little effect on the 140 kVp image; however, it does improve the accuracy of the noisier 80 kVp image, where it is clear that $\mathrm{TV_N}$ outperforms $\mathrm{TV_F}$, which in turn outperforms $\mathrm{TV_S}$. The minima of the various error curves is reported for both the 80 and 140 kVp channels for all regularizers in Table 6.2.

Regularizer	**80 kVp**	**140 kVp**
$\mathrm{TV_S}$	4.291e-2	2.401e-2
$\mathrm{TV_F}$	3.631e-2	2.398e-2
$\mathrm{TV_N}$	3.174e-2	2.364e-2

Table 6.2: Relative error of each VTV variant at its respective optimal λ value. These correspond to the minima of the error curves in Figure 6.2, $\epsilon(\lambda_{\mathrm{opt}})$.

Figure 6.3 shows a comparison between the denoised 80 kVp images between different VTV variants, each shown at its optimal lambda value (the minimizer of the error curve); Figures 6.4 and 6.5 shows the same comparison for two different ROI's. It can be seen that the channel-coupling does slightly improve the preservation of details and suppression of noise. The differences between $\mathrm{TV_F}$ and $\mathrm{TV_N}$ are quite subtle with $\mathrm{TV_N}$ arguably having slightly

cleaner edges, but the improvement over channel-by-channel TV (TV_S) is somewhat more pronounced. The 140 kVp channel is not shown because the differences between regularizers are practically indiscernible. This observation is consistent with our previous findings that the channel-coupling effect of these VTV penalties is primarily helpful to the noisier image channels because it allows edge information to be transferred from the less noisy channel.

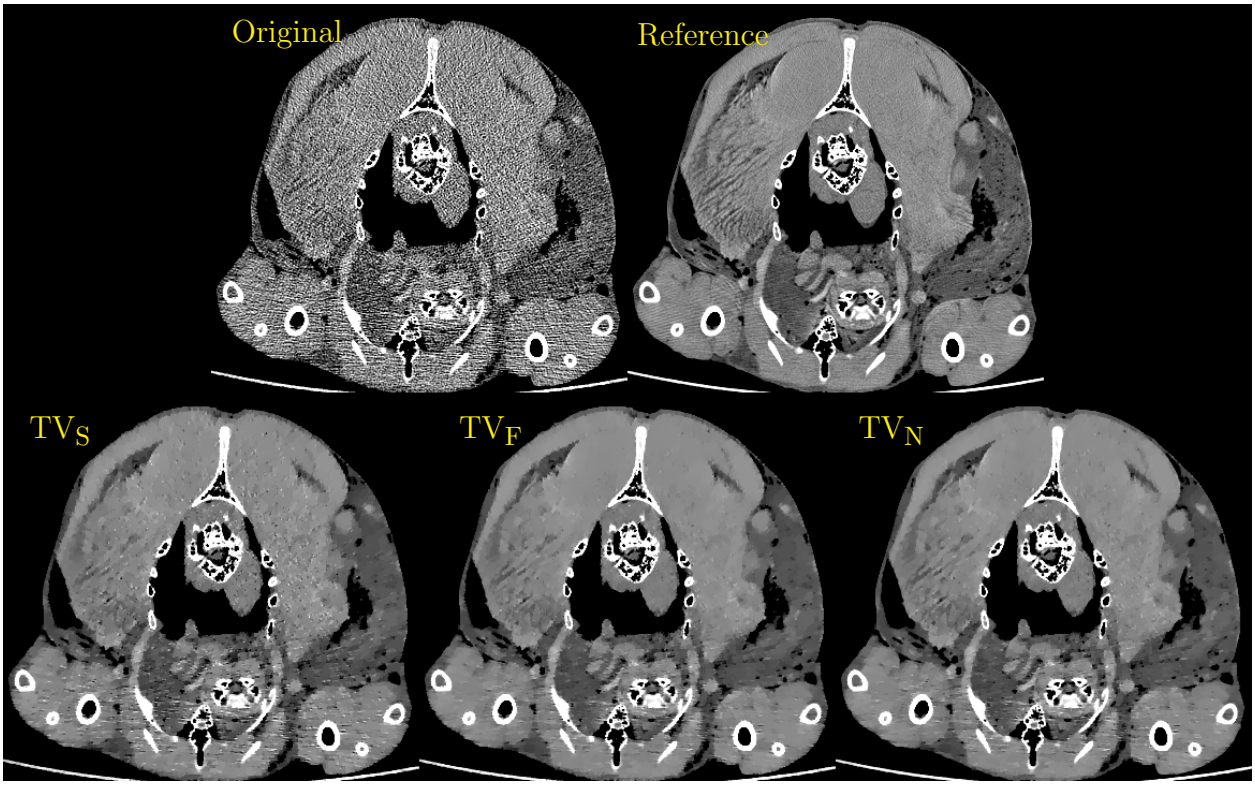


Figure 6.3: Comparison of denoised 80 kVp channel for different VTV variants. Each regularizer is shown at its optimal value, corresponding to the minimum error distance, relative to the reference image. Window (HU) = [-200, 200].

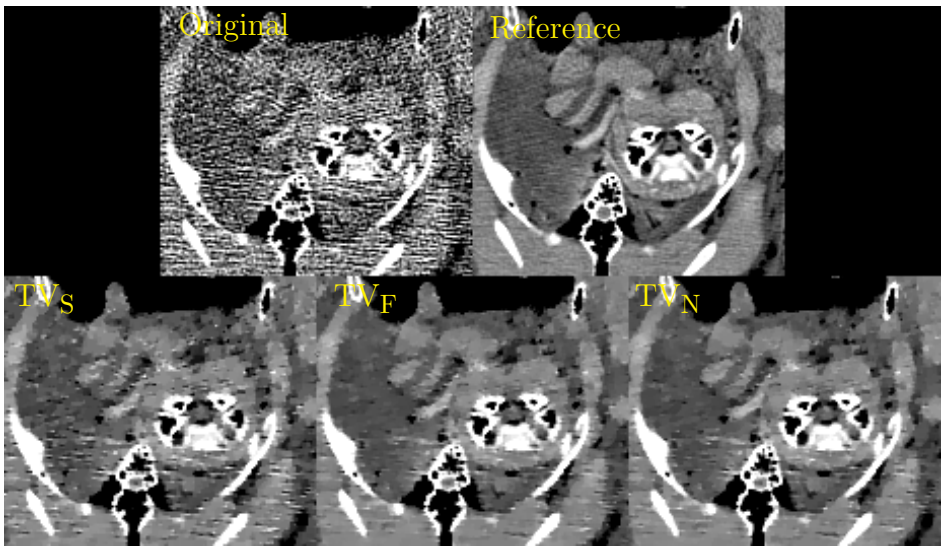


Figure 6.4: Comparison of denoised 80 kVp channel for different VTV variants. Each regularizer is shown at its optimal value, corresponding to the minimum error distance, relative to the reference image. Window (HU) = [-200, 200].

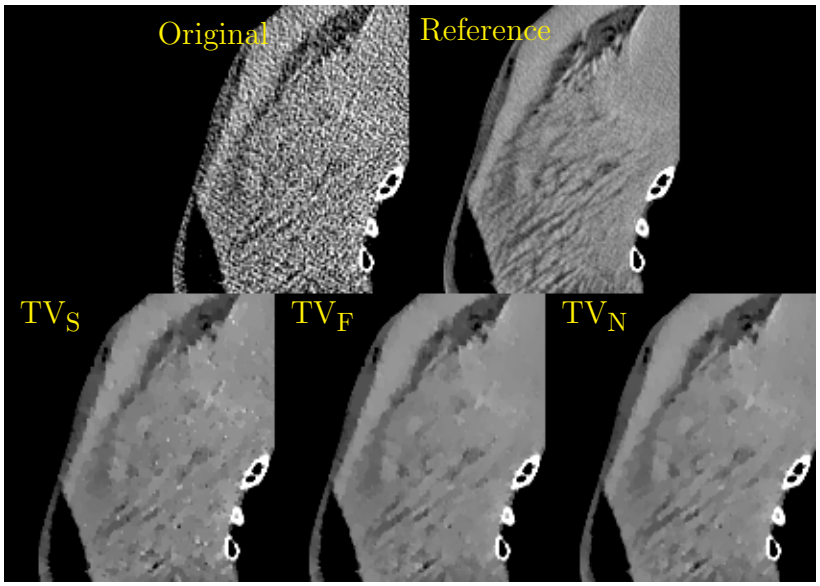


Figure 6.5: Comparison of denoised 80 kVp channel for different VTV variants. Each regularizer is shown at its optimal value, corresponding to the minimum error distance, relative to the reference image. Window (HU) = [-200, 200].

Mismatched Tube Currents

Commonly, dual-kVp CT images can have significantly mismatched noise levels due to the relatively high attenuation of lower-energy x-rays and the reduced flux at lower kVp's (for a fixed mAs). For this reason, 100/140 kVp may be used in certain scenarios [118], rather than 80/140 kVp, despite the inferior spectral separation. This can be particularly problematic in rapid kV switching systems because physical limitations prevent the tube current from being elevated quickly enough to compensate for the decreased tube output as the kVp drops from 140 to 80 kVp. Instead, this effect is partially compensated by increasing the detector integration time for the lower kVp, but this strategy puts undesireable constraints on the gantry rotation speed and scan time.

These noise mismatch issues were not as prominent in our turkey phantom due to its relatively small diameter (<20cm) and lack of large bones or iodinated contrast. However, to emulate the conditions of a severe noise mismatch, we created another dual-energy image pair, combining an 80 kVp low-dose (100 mAs) image with a 140 kVp, high-dose (425mAs) image. Using this mixed-kVp data, we conducted the same denoising experiments as in the previous section to see if the channel-coupling of VTV penalties could improve the quality of the 80 kVp images.

We present the results of the mismatched mAs denoising study in the same manner as the previous section. The error curves are depicted in 6.6, with the lowest error achieved by each TV variant reported in Table 6.3. The error is smaller for all regularizers for the 140 kVp channel because we used a higher-dose image to start with. However, the vectorial TV penalties ($\mathrm{TV_F}$ and $\mathrm{TV_N}$) were also able to improve the relative error in the 80 kVp image by virtue of the channel-coupling. The channel-by-channel TV ($\mathrm{TV_S}$), of course, did no better than before because it is unable to transfer any information from the higher-mAs 140 kVp channel into the lower-mAs 80 kVp channel.

Image comparisons of the same slice and ROI's are shown in Figures 6.7, 6.8, and 6.9 for the 80 kVp channel. Again we can observe that details are best preserved by $\mathrm{TV_N}$, but the

discrepancy is more apparent than with the matched-mAs study.

Regularizer	**80 kVp**	**140 kVp**
TV_S	4.291e-2	1.388e-2
TV_F	3.476e-2	1.392e-2
TV_N	2.714e-2	1.386e-2

Table 6.3: Relative error of each VTV variant at its respective optimal λ value. These correspond to the minima of the error curves in Figure 6.2, $\epsilon(\lambda_{\text{opt}})$.

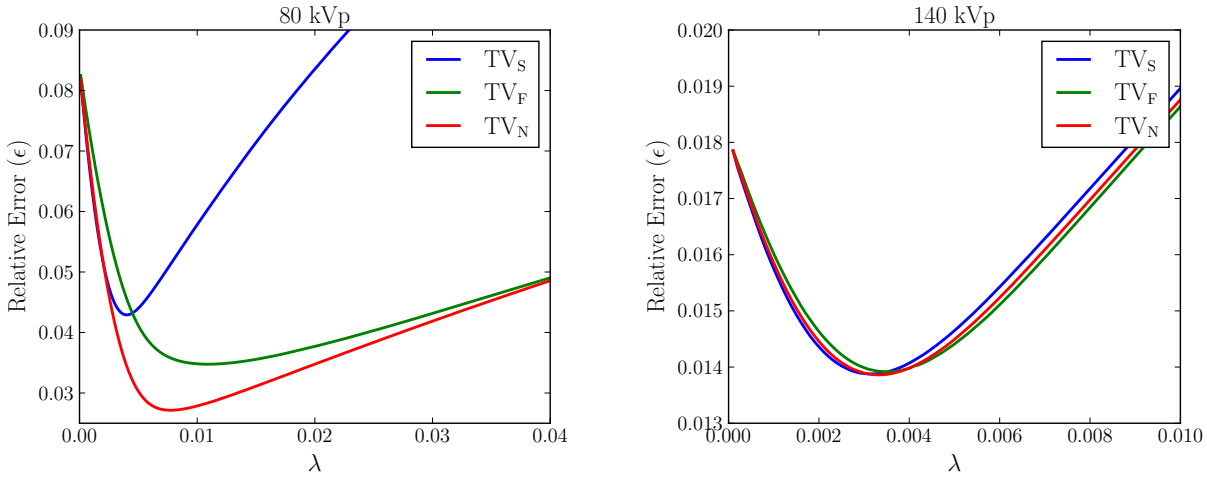


Figure 6.6: Relative error curves based on (6.10) are shown for the three different VTV variants studied for the mixed tube current study (100/425mAs) study. Optimal values of λ are identified by the minima of these curves.

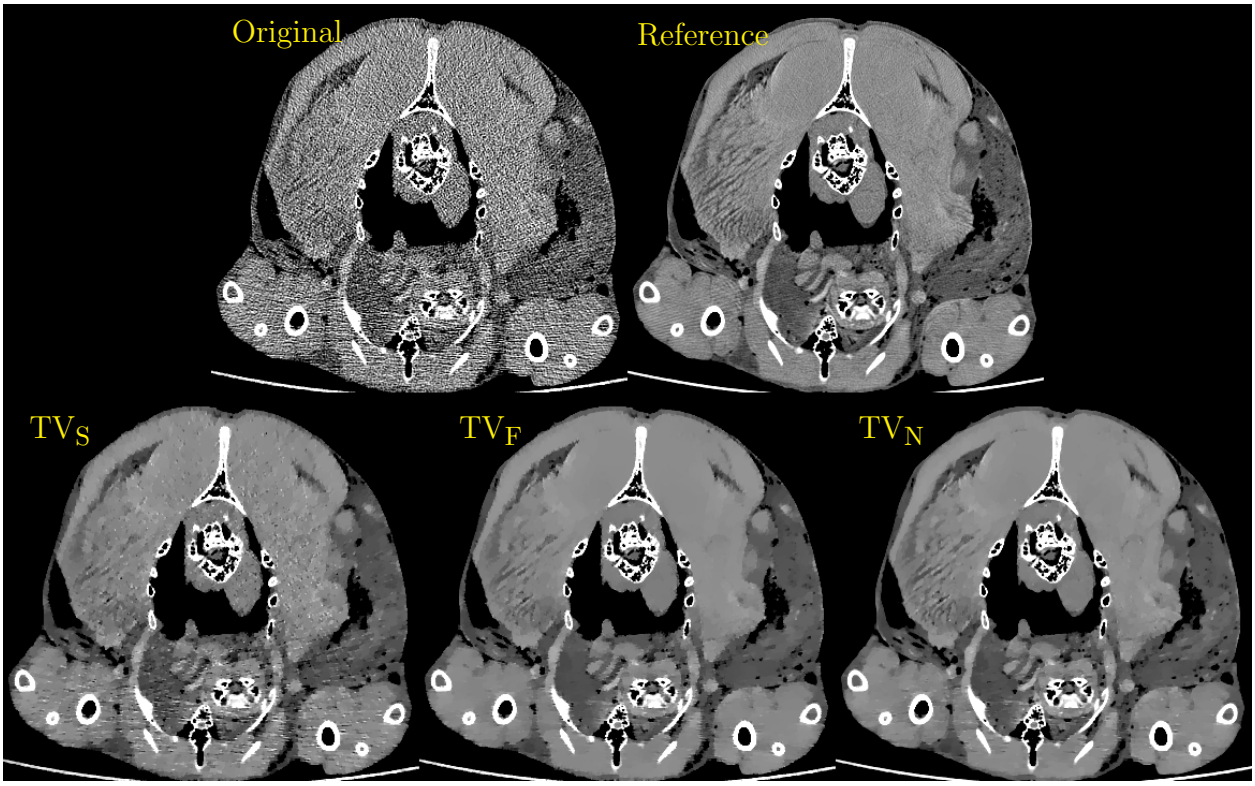


Figure 6.7: Comparison of denoised 80 kVp channel for different VTV variants with mismatched tube currents. Each regularizer is shown at its optimal value, corresponding to the minimum error distance, relative to the reference image. Window (HU) = [-200, 200].

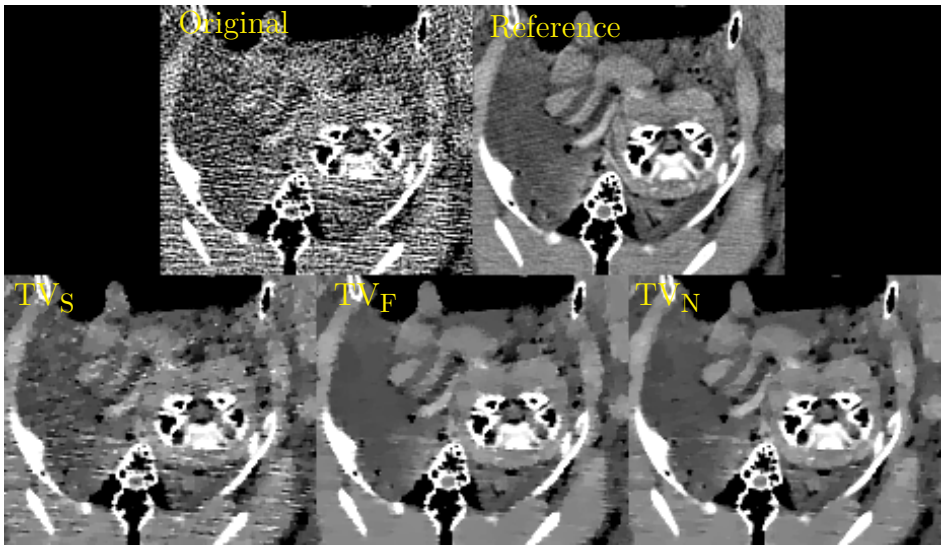


Figure 6.8: Comparison of denoised 80 kVp channel for different VTV variants with mismatched tube currents. Each regularizer is shown at its optimal value, corresponding to the minimum error distance, relative to the reference image. Window (HU) = [-200, 200].

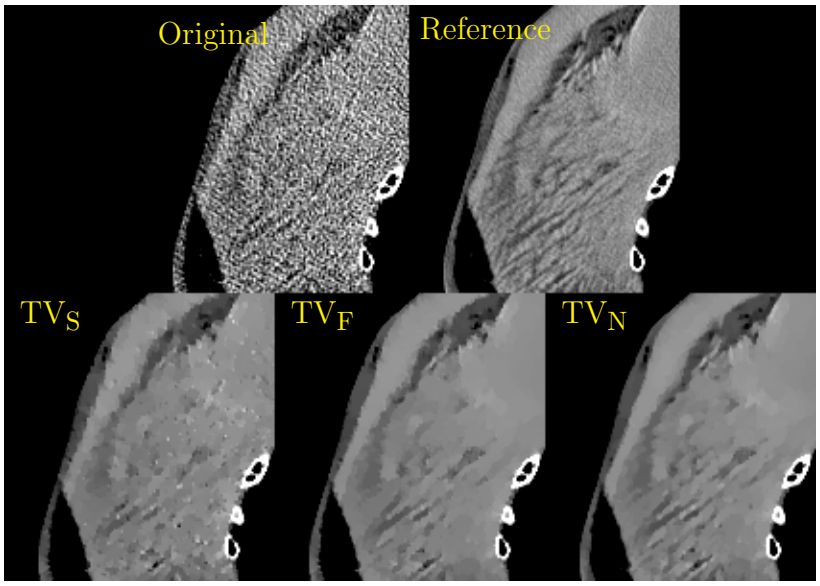


Figure 6.9: Comparison of denoised 80 kVp channel for different VTV variants with mismatched tube currents. Each regularizer is shown at its optimal value, corresponding to the minimum error distance, relative to the reference image. Window (HU) = [-200, 200].

Smoothed (Huber) VTV

We also denoised the 80/140 kVp, low-dose (100 mAs) image pair using a vectorial Huber penalty formulated as a smoothed variant of TV_{N}, which we denote as $\mathrm{TV}_{\mu,\mathrm{N}}$. Figure 6.10 shows a qualitative comparison between the unsmoothed TV_{N} penalty, at two different λ values, and the smoothed $\mathrm{TV}_{\mathrm{N},\mu}$. The images denoised with the vectorial Huber penalty have a somewhat less patchy texture, which may be seen as more natural due to the smoother contrast gradations; however, this comes at the expense of worse edge preservation. This trade-off can be balanced by adjusting μ, where $\mu = 0$ reverts to TV_{N}. The particular parameters selected here likely do not represent a fair comparison between the vectorial Huber penalty and the vectorial TV. We only aim to illustrate, qualitatively, that they can have different effects on image texture and point out that the extension of the Huber penalty to multi-channel images is possible.

Figure 6.10: Low-dose 80 kVp image denoised with TV_{N} at $\lambda = 0.015$ (Left) and $\lambda = 0.01$ (Middle) as well as the vectorial Huber penalty $\mathrm{TV}_{N,\mu}$ at $\lambda = 0.015$, $\mu = 0.003$ (Right). Window (HU) = [-200, 200].

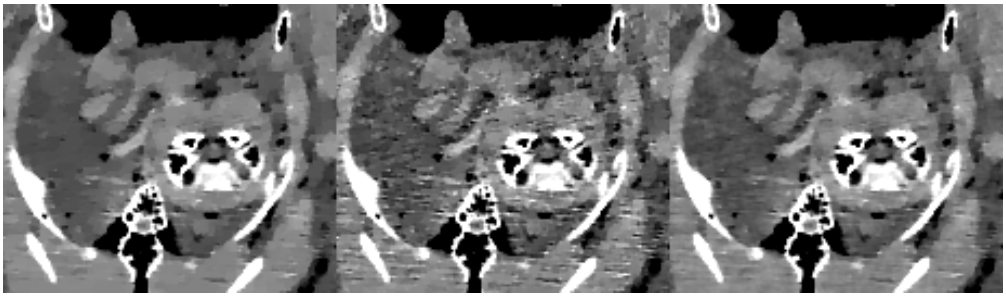


Figure 6.11: An ROI of the low-dose 80 kVp image denoised with $\mathrm{TV_N}$ at $\lambda = 0.015$ (Left) and $\lambda = 0.01$ (Middle) as well as the vectorial Huber penalty $\mathrm{TV}_{N,\mu}$ at $\lambda = 0.015$, $\mu = 0.003$ (Right). Window (HU) = [-200, 200].

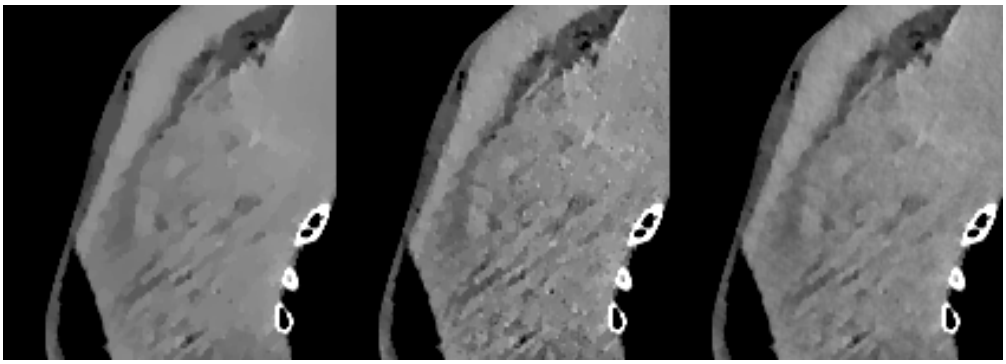


Figure 6.12: An ROI of the low-dose 80 kVp image denoised with $\mathrm{TV_N}$ at $\lambda = 0.015$ (Left) and $\lambda = 0.01$ (Middle) as well as the vectorial Huber penalty $\mathrm{TV}_{N,\mu}$ at $\lambda = 0.015$, $\mu = 0.003$ (Right). Window (HU) = [-200, 200].

Application to 3D Image Volumes

For reconstructing or denoising MDCT or cone beam data, it may be preferable to apply the vectorial TV to a 3D voxel array, which allows regularization along the z-axis as well. To demonstrate that this is feasibile, we also processed an 80/140 kVp pair of image volumes synthesized from the low-dose (100mAs) data. Specifically, we processed volumes consisting of 32 slices. Additional details about how the 3D volumes were pre-processed are outlined in the data acquisition section.

For this 3D study, we focus only on $\mathrm{TV_N}$ because the extension of $\mathrm{TV_S}$ and $\mathrm{TV_F}$ to 3D is trivial. In 2D, the dual-variable update, involving the projection onto the set $\mathcal{N}$, necessitates performing an eigendecomposition of one 2×2 Hermitian matrix per image pixel. Luckily, this has a simple, closed-form solution (as detailed in Chapter 5). In 3D, this requires a

3×3 eigendecomposition for each voxel (regardless of the # of energy channels). Though this computation is slightly more complicated, it turns out that there is still a closed-form solution that can be computed efficiently. In particular, we have used the technique outlined in [119].

Figure 6.13: The middle slice of the low-dose (100mAs) 80 kVp volume before denoising (left) and after denoising (middle) compared to the 10X averaged high-dose reference image (right). Window (HU) = [-200, 200].

Figure 6.14: The middle slice of the low-dose (100mAs) 140 kVp volume before denoising (left) and after denoising (middle) compared to the 10X averaged high-dose reference image (right). Window (HU) = [-200, 200].

In our implementation, the per-iteration time for the vectorial $\mathrm{TV_N}$ was not substantially different than the channel-by-channel $\mathrm{TV_S}$; in all cases it was less than a factor of two longer. We suspect that for 3D image reconstruction, there would be practically no penalty to using $\mathrm{TV_N}$, rather than $\mathrm{TV_S}$, because the vast majority of the computation time would be spent in the (back)projection steps.

6.4 Discussion

In this chapter, we assessed the utility of the vectorial TV for realistic CT images. Consistent with Chapter 5, we found that the edge-coupling properties of $\mathrm{TV_N}$ can help to preserve details and also improve quantitative accuracy, compared with channel-by-channel TV. In this study, the benefits were more subtle than in the reconstruction study utilizing the numerical XCAT phantom. This is not particularly surprising since that phantom is piecewise constant, which makes it possible to operate in a regime of very strong regularization.

In addition, we looked at $\mathrm{TV_F}$ which also encourages common edges but is ambivalent to gradient-vector alignment between image channels. Although $\mathrm{TV_N}$ led to slightly more accurate images (relative to the reference images) and arguably cleaner edges, the differences between $\mathrm{TV_F}$ and $\mathrm{TV_N}$ were far more subtle. This may be of practical interest because the implementation of $\mathrm{TV_F}$ is simpler and slightly cheaper, computationally.

Interestingly, the benefits of the inter-channel coupling were most apparent when there was a very large disparity in the noise levels between the low- and high-energy image channels. Intuitively, this makes sense because both $\mathrm{TV_F}$ and $\mathrm{TV_N}$ allow for edge information to be transferred from the low-noise image into noisy image. This could be useful for making multi-energy CT acquisitions more robust to scenarios where certain channels receive low count rates, such as dual kVp imaging of obese patients.

Additionally, we demonstrated that the vectorial TV can be smoothed to derive a multi-channel analog of the commonly used Huber penalty. This could be beneficial for certain applications, since the Huber penalty allows the user to adjust the trade-off between edge-preservation and bias toward piecewise images. Due to the number of works utilizing the Huber penalty for single-energy CT reconstruction, we think a multi-energy extension may be of practical utility. We have not performed an in-depth comparison between the vectorial Huber penalty and the vectorial TV because it is beyond the scope of this work, but it may be an interesting topic for future studies.

Lastly, we demonstrated that the total nuclear variation ($\mathrm{TV_N}$) can be efficiently ex-

tended to 3D volumes, which is important for several reasons. Firstly, regularizing multi-slice volumes in 3D, rather than processing each slice separately, can improve noise suppression by allowing regularization along the axial direction. Secondly, working with 3D voxel arrays is essential for processing certain kinds of cone-beam and multidetector CT data, which inherently requires 3D modeling.

6.5 Appendix: Projection onto $\mathcal{F}$

In Chapter 6, the projection operators $\Pi_{\mathcal{S}}$ and $\Pi_{\mathcal{N}}$ are defined for the Euclidean projection of a vector $Z \in \mathcal{G}$ onto the sets $\mathcal{S}$ and $\mathcal{N}$. These operations are involved in the algorithm updates for $\mathrm{TV_S}$ and $\mathrm{TV_N}$, respectively. In this section, we define a similar projection operator, Π_F, for $\mathrm{TV_F}$.

We define this operation element-wise on each $\mathcal{M} \times 2$ matrix $Z(i,j)$, where $\mathcal{M}$ is the number of spectral channels, and 2 corresponds to the number of spatial dimensions.

$$\Pi_{\mathcal{F}}\left(Z(i,j)\right) = \begin{cases} Z(i,j) & \|Z(i,j)\|_F \leq 1 \\ \frac{Z(i,j)}{\|Z(i,j)\|_F} & \text{otherwise} \end{cases}. \tag{6.11}$$

For every pixel location (i,j), we simply project the matrix $Z(i,j)$ onto the unit ball defined by the Frobenius norm.

7

Summary and Conclusions

In this thesis we have investigated potential improvements to various parts of the spectral CT imaging chain, including data acquisition, pre-processing, and image reconstruction. These contributions aim to improve the end-to-end dose efficiency and are broadly applicable to any spectral CT system, including commercially available dual-energy scanners as well as multi-energy photon-counting systems, which are being actively researched.

In Chapter 3, we presented a task-based framework for optimizing spectral CT imaging protocols and hardware design. Specifically, we consider material classification tasks by extension of classical signal detection theory. Previous works have studied how imaging parameter choices affect the noise in basis-material sinograms [32] and images [120], but many relevant tasks are better formulated as binary decision problems. The metric we presented is analytic, rapidly computable, and indpendent of the reconstruction technique. We validated our model with simulation studies and close agreement is demonstrated for a variety of different tasks and imaging configurations. Future work could attempt to validate the optimality of the imaging parameters selected by our model with real scanner measurements. However, this would be immensely challenging since it would require extensive knowledge of the x-ray tube spectrum and detector response as well as complex phantoms and an enormous number of image realizations.

In Chapter 4, we presented an iterative method for projection-space material decomposition that achieves an excellent balance between speed and accuracy. Though the proposed model can be applied to any spectral CT system, it is of particular interest for photon-counting CT (PCCT) systems with $N > 2$ energy bins, since the material decomposition is overdetermined, and the noise levels may vary significantly between energy bins. For dual-energy systems, this is not the case, and the choice of data model is not particularly important [121]. By formulating the problem as a weighted, nonlinear least-squares problem in the log-transformed data and applying the Gauss-Newton algorithm, we demonstrated that highly accurate results can be obtained in only a few iterations. Furthermore, this model consistently performs similarly to the maximum-likelihood estimator (MLE), while requiring significantly fewer iterations. This is significant because the MLE has been known to achieve the Cramer-Rao bound in many practical situations [54], implying that no other unbiased estimator can outperform it. Thus, our model allows for near optimal utilization of the spectral information with very low processing times.

Chapter 5 addresses the issue of jointly reconstructing multi-energy spectral CT data with a novel multi-channel regularizer. The proposed regularization function can be viewed as an extension of the widely used total variation (TV) penalty to vector-valued images. We demonstrated that the proposed vectorial TV promotes a common edge structure and aligned gradient vectors across image channels. Our results indicate that these features result in improved noise suppression compared to regularizing each image channel independently and that the VTV can be applied in either the energy or basis-material domain with similar success. We discovered that the vectorial TV is particularly helpful in imaging scenarios where some of the channels are significantly noisier than others. In these cases, the channel-coupling allows for edge information to be transferred from higher-SNR channels into lower-SNR channels. Additionally, we found that the VTV improved quantitative accuracy (compared to ground truth images) and did not transfer false edges into channels where they should not be.

Finally, in Chapter 6 we assessed the utility of our VTV regularizer on real, dual-energy CT data. Since we did not have access to the raw projection data, we evaluated the vectorial TV for denoising images that were reconstructed using the onboard software. A weakness of our work in Chapter 5 is that the results were demonstrated on a piecewise-constant phantom. The total variation has an inherent bias that favors piecewise structures, so this may exaggerate performance. Therefore, the studies in Chapter 6 utilized scans of a frozen supermarket turkey, which has complex textural elements that provides similar challenges to real patient images. The improvements provided by the VTV were more modest than what is observed in Chapter 5; however, we still found that jointly denoising 80/140 kVp images with VTV was advantageous to sequential processing. Additionally, we tested two different variants of the VTV and found that TV_{N}, which promotes aligned gradient vectors, slightly outperformed TV_{F}, which does not. Lastly, we showed that the VTV can be applied efficiently to 3D volumes, utilizing the closed-form, 3×3 matrix eigendecomposition method of Kopp [119]. This is of vital importance for processing multidetector and conebeam CT data.

Though spectral CT has not yet achieved widespread clinical adoption, we believe that this will come through the continued refinement of imaging hardware, protocols, and algorithms. In this thesis, we have touched on all of these areas in some capacity. We hope that some of these techniques might contribute, however minutely, to mainstreaming spectral CT. The methods presented in Chapter 3 could be used to immediately boost the performance of existing dual-energy systems by ensuring that the best imaging parameters are used for each task. The later work on multi-channel image regularization is also applicable to current systems. Though the proposed vectorial TV reconstruction model would be difficult to implement clinically, due to unacceptably long computation times, the VTV could be immediately useful for denoising, as demonstrated in Chapter 6. Furthermore, the concept of multi-channel image regularization is broadly applicable outside of spectral CT, so our work on the VTV may be useful in other areas, such as PET/CT, PET/MRI, or any

other situation involving the reconstruction of multiple images that have shared anatomical features.

8

References

[1] D M Tucker, G T Barnes, and D P Chakraborty. Semiempirical model for generating tungsten target x-ray spectra. *Med. Phys.*, 18(2):211–218, March 1991.

[2] G N Hounsfield. Computerized transverse axial scanning (tomography). 1. description of system. *Br. J. Radiol.*, 46(552):1016–1022, December 1973.

[3] R E Alvarez and A Macovski. Energy-selective reconstructions in x-ray computerised tomography. *Phys. Med. Biol.*, 21(5):733, 1976.

[4] J R Vetter, W H Perman, W A Kalender, R B Mazess, and J E Holden. Evaluation of a prototype dual-energy computed tomographic apparatus. II. determination of vertebral bone mineral content. *Med. Phys.*, 13(3):340–343, May 1986.

[5] Willi A Kalender. *Computed Tomography: Fundamentals, System Technology, Image Quality, Applications*. Wiley, 29 August 2011.

[6] Otto Glasser. Wilhelm conrad roentgen and the early history of roentgan rays. *Am. J. Med. Sci.*, 187(4):566, 1934.

[7] E C Beckmann. CT scanning the early days. *Br. J. Radiol.*, 79(937):5–8, January 2006.

[8] William E Brant and Clyde A Helms. *Fundamentals of diagnostic radiology*. Lippincott Williams & Wilkins, 2012.

[9] G N Hounsfield. Computed medical imaging. *Science*, 210(4465):22–28, 3 October 1980.

[10] A M Cormack. Early two-dimensional reconstruction and recent topics stemming from it. *Science*, 209(4464):1482–1486, 26 September 1980.

[11] National Lung Screening Trial Research Team, Denise R Aberle, Amanda M Adams, Christine D Berg, William C Black, Jonathan D Clapp, Richard M Fagerstrom, Ilana F Gareen, Constantine Gatsonis, Pamela M Marcus, and Jorean D Sicks. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N. Engl. J. Med.*, 365(5):395–409, 4 August 2011.

[12] Matthew J Budoff, Stephan Achenbach, Roger S Blumenthal, J Jeffrey Carr, Jonathan G Goldin, Philip Greenland, Alan D Guerci, Joao A C Lima, Daniel J Rader, Geoffrey D Rubin, Leslee J Shaw, Susan E Wiegers, American Heart Association Committee on Cardiovascular Imaging and Intervention, American Heart Association Council on Cardiovascular Radiology and Intervention, and American Heart Association Committee on Cardiac Imaging, Council on Clinical Cardiology. Assessment of coronary artery disease by cardiac computed tomography: a scientific statement from the american heart association committee on cardiovascular imaging and intervention, council on cardiovascular radiology and intervention, and committee on cardiac imaging, council on clinical cardiology. *Circulation*, 114(16):1761–1791, 17 October 2006.

[13] M J Haydel, C A Preston, T J Mills, S Luber, E Blaudeau, and P M DeBlieux. Indications for computed tomography in patients with minor head injury. *N. Engl. J. Med.*, 343(2):100–105, 13 July 2000.

[14] J S Wilson, M Korobkin, H K Genant, and E G Bovill, Jr. Computed tomography of musculoskeletal disorders. *AJR Am. J. Roentgenol.*, 131(1):55–61, July 1978.

[15] M G Cavalcanti, J W Haller, and M W Vannier. Three-dimensional computed tomography landmark measurement in craniofacial surgical planning: experimental validation in vitro. *J. Oral Maxillofac. Surg.*, 57(6):690–694, June 1999.

[16] David A Jaffray, Jeffrey H Siewerdsen, John W Wong, and Alvaro A Martinez. Flat-panel cone-beam computed tomography for image-guided radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 53(5):1337–1349, 1 August 2002.

[17] F Kelcz, P M Joseph, and S K Hilal. Noise considerations in dual energy CT scanning. *Med. Phys.*, 6(5):418–425, September 1979.

[18] Björn J Heismann, Bernhard T Schmidt, and Thomas Flohr. *Spectral Computed Tomography*. SPIE Press, Bellingham, Washington, 2012.

[19] Thomas G Flohr, Cynthia H McCollough, Herbert Bruder, Martin Petersilka, Klaus Gruber, Christoph Süss, Michael Grasruck, Karl Stierstorfer, Bernhard Krauss, Rainer Raupach, Andrew N Primak, Axel Küttner, Stefan Achenbach, Christoph Becker, Andreas Kopp, and Bernd M Ohnesorge. First performance evaluation of a dual-source CT (DSCT) system. *Eur. Radiol.*, 16(2):256–268, February 2006.

[20] Thorsten R C Johnson, Bernhard Krauss, Martin Sedlmair, Michael Grasruck, Herbert Bruder, Dominik Morhard, Christian Fink, Sabine Weckbach, Miriam Lenhard, Bernhard Schmidt, Thomas Flohr, Maximilian F Reiser, and Christoph R Becker. Material differentiation by dual energy CT: initial experience. *Eur. Radiol.*, 17(6):1510–1517, June 2007.

[21] A N Primak, J C Ramirez Giraldo, X Liu, L Yu, and C H McCollough. Improved dual-energy material discrimination for dual-source CT by means of additional spectral filtration. *Med. Phys.*, 36(4):1359–1369, April 2009.

[22] Anno Graser, Thorsten R C Johnson, Hersh Chandarana, and Michael Macari. Dual energy CT: preliminary observations and potential clinical applications in the abdomen. *Eur. Radiol.*, 19(1):13–23, 2 January 2009.

[23] Savvakis Nicolaou, Charlotte Jane Yong-Hing, Sandro Galea-Soler, Daniel J Hou, Luck Louis, and Peter Munk. Dual-Energy CT as a potential new diagnostic tool in the management of gout in the acute setting. *American Journal of Roentgenology*, 194(4):1072–1078, April 2010.

[24] Sarah C Metzger, Michaela Koehm, Julian L Wichmann, Stefan Buettner, Jan-Erik Scholtz, Martin Beeres, J Matthias Kerl, Moritz H Albrecht, Renate Hammerstingl, Thomas J Vogl, and Ralf W Bauer. Dual-Energy CT in patients with suspected gouty arthritis:: Effects on treatment regimen and clinical outcome. *Acad. Radiol.*, 23(3):267–272, March 2016.

[25] François Pontana, Jean-Baptiste Faivre, Martine Remy-Jardin, Thomas Flohr, Bernhard Schmidt, Nunzia Tacelli, Vittorio Pansini, and Jacques Remy. Lung perfusion with dual-energy multidetector-row CT (MDCT): feasibility for the evaluation of acute pulmonary embolism in 117 consecutive patients. *Acad. Radiol.*, 15(12):1494–1504, December 2008.

[26] Cong Sun, Fan Miao, Xi-Ming Wang, Tao Wang, Rui Ma, Dao-Ping Wang, and Cheng Liu. An initial qualitative study of dual-energy CT in the knee ligaments. *Surg. Radiol. Anat.*, 30(5):443–447, July 2008.

[27] Kensuke Uotani, Yoshiyuki Watanabe, Masahiro Higashi, Tetsuro Nakazawa, Atsushi K Kono, Yoshiro Hori, Tetsuya Fukuda, Suzu Kanzaki, Naoaki Yamada, Toshihide Itoh, Kazuro Sugimura, and Hiroaki Naito. Dual-energy CT head bone and hard plaque removal for quantification of calcified carotid stenosis: utility and comparison with digital subtraction angiography. *Eur. Radiol.*, 19(8):2060–2065, August 2009.

[28] Naveen Chandra and David A Langan. Gemstone detector: Dual energy imaging via fast kvp switching. In *Dual Energy CT in Clinical Practice*, Medical Radiology, pages 35–41. Springer Berlin Heidelberg, 2011.

[29] Alain Vlassenbroek. Dual layer CT. In *Dual Energy CT in Clinical Practice*, Medical Radiology, pages 21–34. Springer Berlin Heidelberg, 2011.

[30] Carsten O Schirra, Bernhard Brendel, Mark A Anastasio, and Ewald Roessl. Spectral CT: a technology primer for contrast agent development. *Contrast Media Mol. Imaging*, 9(1):62–70, January 2014.

[31] E Roessl and R Proksa. K-edge imaging in x-ray computed tomography using multi-bin photon counting detectors. *Phys. Med. Biol.*, 52(15):4679–4696, 7 August 2007.

[32] E Roessl and C Herrmann. Cramér-Rao lower bound of basis image noise in multiple-energy x-ray imaging. *Phys. Med. Biol.*, 54(5):1307–1318, 7 March 2009.

[33] A S Wang and N J Pelc. Sufficient statistics as a generalization of binning in spectral x-ray imaging. *IEEE Trans. Med. Imaging*, 30(1):84–93, January 2011.

[34] Katsuyuki Taguchi and Jan S Iwanczyk. Vision 20/20: Single photon counting x-ray detectors in medical imaging. *Med. Phys.*, 40(10):100901, October 2013.

[35] Norbert J Pelc. Recent and future directions in CT imaging. *Ann. Biomed. Eng.*, 42(2):260–268, February 2014.

[36] Polad M Shikhaliev. Energy-resolved computed tomography: first experimental results. *Phys. Med. Biol.*, 53(20):5595–5613, 21 October 2008.

[37] Taly Gilat Schmidt. Optimal "image-based" weighting for energy-resolved CT. *Med. Phys.*, 36(7):3018–3027, 1 July 2009.

[38] Amir Pourmorteza, Rolf Symons, Veit Sandfort, Marissa Mallek, Matthew K Fuld, Gregory Henderson, Elizabeth C Jones, Ashkan A Malayeri, Les R Folio, and David A Bluemke. Abdominal imaging with contrast-enhanced photon-counting CT: First human experience. *Radiology*, page 152601, 3 February 2016.

[39] Robert Alvarez and Edward Seppi. A comparison of noise and dose in conventional and energy selective computed tomography. *IEEE Trans. Nucl. Sci.*, 26(2):2853–2856, April 1979.

[40] Julien Mairal, Michael Elad, and Guillermo Sapiro. Sparse representation for color image restoration. *IEEE Trans. Image Process.*, 17(1):53–69, January 2008.

[41] Kevin Holt. Total nuclear variation and jacobian extensions of total variation for vector fields. *IEEE Trans. Image Process.*, 23(9):3975–3989, 23 June 2014.

[42] David S Rigie and Patrick J La Rivière. Joint reconstruction of multi-channel, spectral CT data via constrained total nuclear variation minimization. *Phys. Med. Biol.*, 60(5):1741–1762, 7 March 2015.

[43] Hans Bornefalk. XCOM intrinsic dimensionality for low-z elements at diagnostic energies. *Med. Phys.*, 39(2):654–657, February 2012.

[44] Robert E Alvarez. Dimensionality and noise in energy selective x-ray imaging. *Med. Phys.*, 40(11):111909, November 2013.

[45] Hans Bornefalk, Mats Persson, and Mats Danielsson. Allowable forward model misspecification for accurate basis decomposition in a silicon detector based spectral CT. *IEEE Trans. Med. Imaging*, 34(3):788–795, March 2015.

[46] Marc Kachelriess, Katia Sourbelle, and Willi A Kalender. Empirical cupping correction: a first-order raw data precorrection for cone-beam computed tomography. *Med. Phys.*, 33(5):1269–1274, May 2006.

[47] Cynthia H McCollough, Shuai Leng, Lifeng Yu, and Joel G Fletcher. Dual- and Multi-Energy CT: Principles, technical approaches, and clinical applications. *Radiology*, 276(3):637–653, September 2015.

[48] Thorsten Johnson, Christian Fink, Stefan O Schönberg, and Maximilian F Reiser. *Dual Energy CT in Clinical Practice*. Springer Science & Business Media, 18 January 2011.

[49] Daniele Marin, Daniel T Boll, Achille Mileto, and Rendon C Nelson. State of the art: dual-energy CT of the abdomen. *Radiology*, 271(2):327–342, May 2014.

[50] Fabian Bamberg, Alexander Dierks, Konstantin Nikolaou, Maximilian F Reiser, Christoph R Becker, and Thorsten R C Johnson. Metal artifact reduction by dual energy computed tomography using monoenergetic extrapolation. *Eur. Radiol.*, 21(7):1424–1429, July 2011.

[51] Joel G Fletcher, Naoki Takahashi, Robert Hartman, Luis Guimaraes, James E Huprich, David M Hough, Lifeng Yu, and Cynthia H McCollough. Dual-energy and dual-source CT: is there a role in the abdomen and pelvis? *Radiol. Clin. North Am.*, 47(1):41–57, January 2009.

[52] Andrew N Primak, Joel G Fletcher, Terri J Vrtiska, Oleksandr P Dzyubak, John C Lieske, Molly E Jackson, James C Williams, Jr, and Cynthia H McCollough. Noninvasive differentiation of uric acid versus non-uric acid kidney stones using dual-energy CT. *Acad. Radiol.*, 14(12):1441–1447, December 2007.

[53] Paulo R S Mendonca, Peter Lamb, and Dushyant V Sahani. A flexible method for Multi-Material decomposition of Dual-Energy CT images. *IEEE Trans. Med. Imaging*, 33(1):99–116, January 2014.

[54] J P Schlomka, E Roessl, R Dorscheid, S Dill, G Martens, T Istel, C Bäumer, C Herrmann, R Steadman, G Zeitler, A Livne, and R Proksa. Experimental feasibility of multi-energy photon-counting k-edge imaging in pre-clinical computed tomography. *Phys. Med. Biol.*, 53(15):4031–4047, 7 August 2008.

[55] Ralf Gutjahr, Ahmed F Halaweish, Zhicong Yu, Shuai Leng, Lifeng Yu, Zhoubo Li, Steven M Jorgensen, Erik L Ritman, Steffen Kappler, and Cynthia H McCollough. Human imaging with photon Counting-Based computed tomography at clinical dose levels: Contrast-to-Noise ratio and cadaver studies. *Invest. Radiol.*, 25 January 2016.

[56] C A Agudelo and C M Wise. Gout: diagnosis, pathogenesis, and clinical manifestations. *Curr. Opin. Rheumatol.*, 13(3):234–239, May 2001.

[57] C Thomas, A Korn, B Krauss, D Ketelsen, I Tsiflikas, A Reimann, H Brodoefel, C D Claussen, A F Kopp, U Ernemann, and M Heuschmid. Automatic bone and plaque removal using dual energy CT for head and neck angiography: feasibility and initial performance evaluation. *Eur. J. Radiol.*, 76(1):61–67, October 2010.

[58] Grace J Gang, Wojciech Zbijewski, J Webster Stayman, and Jeffrey H Siewerdsen. Cascaded systems analysis of noise and detectability in dual-energy cone-beam CT. *Med. Phys.*, 39(8):5145–5156, August 2012.

[59] Robert E Alvarez. Near optimal energy selective x-ray imaging system performance with simple detectors. *Med. Phys.*, 37(2):822–841, February 2010.

[60] J H Siewerdsen and D A Jaffray. Optimization of x-ray imaging geometry (with specific application to flat-panel cone-beam computed tomography). *Med. Phys.*, 27(8):1903–1914, August 2000.

[61] Lili Zhou and Gene Gindi. Collimator optimization in SPECT based on an LROC ideal observer. In *2008 IEEE Nuclear Science Symposium Conference Record*, pages 5540–5547. IEEE, October 2008.

[62] Adrian A Sanchez, Emil Y Sidky, and Xiaochuan Pan. Task-based optimization of dedicated breast CT via hotelling observer metrics. *Med. Phys.*, 41(10):101917, October 2014.

[63] C E Metz. Basic principles of ROC analysis. *Semin. Nucl. Med.*, 8(4):283–298, October 1978.

[64] H H Barrett, J Yao, J P Rolland, and K J Myers. Model observers for assessment of image quality. *Proc. Natl. Acad. Sci. U. S. A.*, 90(21):9758–9765, 1 November 1993.

[65] Harrison H Barrett and Kyle J Myers. *Foundations of image science*. Wiley-Interscience, 2004.

[66] Adrian A Sanchez, Emil Y Sidky, Ingrid Reiser, and Xiaochuan Pan. Comparison of human and hotelling observer performance for a fan-beam CT signal detection task. *Med. Phys.*, 40(3):031104, March 2013.

[67] David S Rigie and Patrick J La Riviere. Task based characterization of spectral CT performance via the hotelling observer. In *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)*, pages 3729–3731. IEEE, October 2012.

[68] Lifeng Yu, Shuai Leng, and Cynthia H McCollough. Dual-Energy CT–Based monochromatic imaging. *American Journal of Roentgenology*, 199(5_supplement):S9–S15, November 2012.

[69] Joseph Lee Rodgers and W Alan Nicewander. Thirteen ways to look at the correlation coefficient. *Am. Stat.*, 42(1):59–66, 1988.

[70] T Li, S Zhu, and M Ogihara. Using discriminant analysis for multi-class classification. In *Third IEEE International Conference on Data Mining*, pages 589–592. IEEE Comput. Soc, 2003.

[71] Robert E Alvarez. Estimator for photon counting energy selective x-ray imaging with multibin pulse height analysis. *Med. Phys.*, 38(5):2324–2334, May 2011.

[72] Steven G Self and Kung-Yee Liang. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Am. Stat. Assoc.*, 82(398):605–610, 1987.

[73] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[74] K Sauer and C Bouman. A local update strategy for iterative reconstruction from projections. *IEEE Trans. Signal Process.*, 41(2):534–548, February 1993.

[75] Johan Nuyts, Bruno De Man, Jeffrey A Fessler, Wojciech Zbijewski, and Freek J Beekman. Modelling the physics in the iterative reconstruction for transmission computed tomography. *Phys. Med. Biol.*, 58(12):R63–96, 21 June 2013.

[76] C A Bouman and K Sauer. A unified approach to statistical tomography using coordinate descent optimization. *IEEE Trans. Image Process.*, 5(3):480–492, 1996.

[77] Junguo Bian, Jeffrey H Siewerdsen, Xiao Han, Emil Y Sidky, Jerry L Prince, Charles A Pelizzari, and Xiaochuan Pan. Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT. *Phys. Med. Biol.*, 55(22):6575–6599, 21 November 2010.

[78] Idris A Elbakri and Jeffrey A Fessler. Statistical image reconstruction for polyenergetic x-ray computed tomography. *IEEE Trans. Med. Imaging*, 21(2):89–99, February 2002.

[79] J Webster Stayman, Yoshito Otake, Jerry L Prince, A Jay Khanna, and Jeffrey H Siewerdsen. Model-based tomographic reconstruction of objects containing known components. *IEEE Trans. Med. Imaging*, 31(10):1837–1848, October 2012.

[80] Qiaofeng Xu, Alex Sawatzky, Mark A Anastasio, and Carsten O Schirra. Sparsity-regularized image reconstruction of decomposed k-edge data in spectral CT. *Phys. Med. Biol.*, 59(10):N65–79, 21 May 2014.

[81] Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4):259–268, November 1992.

[82] T F Chan and C K Wong. Total variation blind deconvolution. *IEEE Trans. Image Process.*, 7(3):370–375, 1998.

[83] L Condat and S Mosaddegh. Joint demosaicking and denoising by total variation minimization. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 2781–2784. IEEE, 2012.

[84] E Y Sidky, C M Kao, and X Pan. Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. *J. Xray Sci. Technol.*, 14:119–139, 2006.

[85] Xavier Bresson and Tony F Chan. Fast dual minimization of the vectorial total variation norm and applications to color image processing. *Inverse Probl. Imaging*, 2(4):455–484, 2008.

[86] Bastian Goldluecke, Evgeny Strekalovskiy, and Daniel Cremers. The Natural Vectorial Total Variation Which Arises from Geometric Measure Theory. *SIAM J. Imag. Sci.*, 5(2):537–563, 2012.

[87] Stamatios Lefkimmiatis, Anastasios Roussos, Michael Unser, and Petros Maragos. Convex Generalizations of Total Variation Based on the Structure Tensor with Applications to Inverse Problems. In Arjan Kuijper, Kristian Bredies, Thomas Pock, and Horst Bischof, editors, *Scale Space and Variational Methods in Computer Vision*, volume 7893 of *Lecture Notes in Computer Science*, pages 48–60. Springer Berlin Heidelberg, Berlin, Heidelberg, 2 June 2013.

[88] Alex Sawatzky, Qiaofeng Xu, Carsten O Schirra, and Mark A Anastasio. Proximal ADMM for multi-channel image reconstruction in spectral x-ray CT. *IEEE Trans. Med. Imaging*, 33(8):1657–1668, August 2014.

[89] David S Rigie and Patrick J La Rivière. A generalized vectorial total-variation for spectral CT reconstruction. *Proceedings of the Third International Conference on Image Formation in X-ray Computed Tomography*, 1(1):9–12, 2014.

[90] Matthias Joachim Ehrhardt and Simon R Arridge. Vector-valued image processing by parallel level sets. *IEEE Trans. Image Process.*, 23(1):9–18, January 2014.

[91] Daniel Keren and Anna Gotlib. Denoising Color Images Using Regularization and "Correlation Terms". *J. Vis. Commun. Image Represent.*, 9(4):352–365, December 1998.

[92] K M Holt. Angular regularization of vector-valued signals. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 1105–1108. IEEE, May 2011.

[93] Yong Long and Jeffrey A Fessler. Multi-material decomposition using statistical image reconstruction for spectral CT. *IEEE Trans. Med. Imaging*, 33(8):1614–1626, August 2014.

[94] Martin Petersilka, Herbert Bruder, Bernhard Krauss, Karl Stierstorfer, and Thomas G Flohr. Technical principles of dual source CT. *Eur. J. Radiol.*, 68(3):362–368, December 2008.

[95] Antonin Chambolle and Thomas Pock. A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging. *J. Math. Imaging Vis.*, 40(1):120–145, 21 December 2010.

[96] Martin S Andersen and Per Christian Hansen. Generalized row-action methods for tomographic imaging. *Numer. Algorithms*, 67(1):121–144, 1 November 2013.

[97] E J Candes, J Romberg, and T Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory*, 52(2):489–509, February 2006.

[98] Xiao Han, Junguo Bian, Erik L Ritman, Emil Y Sidky, and Xiaochuan Pan. Optimization-based reconstruction of sparse images from few-view projections. *Phys. Med. Biol.*, 57(16):5245–5273, 21 August 2012.

[99] Junguo Bian, Jiong Wang, Xiao Han, Emil Y Sidky, Lingxiong Shao, and Xiaochuan Pan. Optimization-based image reconstruction from sparse-view data in offset-detector CBCT. *Phys. Med. Biol.*, 58(2):205–230, 21 January 2013.

[100] Martin Burger and Stanley Osher. A Guide to the TV Zoo. In *Level Set and PDE Based Reconstruction Methods in Imaging*, volume 2090 of *Lecture Notes in Mathematics*, pages 1–70. Springer International Publishing, Cham, 2013.

[101] Jian-Feng Cai, Emmanuel J Candès, and Zuowei Shen. A Singular Value Thresholding Algorithm for Matrix Completion. *SIAM J. Optim.*, 20(4):1956–1982, 2010.

[102] Emmanuel J Candès, Xiaodong Li, Yi Ma, and John Wright. Robust principal component analysis? *J. ACM*, 58(3):11, 1 May 2011.

[103] Emil Y Sidky, Jakob H Jørgensen, and Xiaochuan Pan. Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm. *Phys. Med. Biol.*, 57(10):3065–3091, 21 May 2012.

[104] Neal Parikh and Stephen Boyd. Proximal Algorithms. *Foundations and Trends® in Optimization*, 1(3):127–239, 2014.

[105] W P Segars, G Sturgeon, S Mendonca, Jason Grimes, and B M W Tsui. 4D XCAT phantom for multimodality imaging research. *Med. Phys.*, 37(9):4902–4915, September 2010.

[106] P M Joseph. An improved algorithm for reprojecting rays through pixel images. *IEEE Trans. Med. Imaging*, 1(3):192–196, 1982.

[107] C O Schirra, E Roessl, T Koehler, B Brendel, A Thran, D Pan, M A Anastasio, and R Proksa. Statistical reconstruction of material decomposed data in spectral CT. *IEEE Trans. Med. Imaging*, 32(7):1249–1257, July 2013.

[108] T Chan, S Esedoglu, F Park, and A Yip. Handbook of Mathematical Models in Computer Vision. In Nikos Paragios, Yunmei Chen, and Olivier Faugeras, editors, *Handbook of Mathematical Models in Computer Vision*, pages 17–31. Springer US, Boston, MA, 2006.

[109] David Strong and Tony Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Probl.*, 19(6):S165, 12 November 2003.

[110] Ludwig Ritschl, Frank Bergner, Christof Fleischmann, and Marc Kachelriess. Improved total variation-based CT image reconstruction applied to clinical data. *Phys. Med. Biol.*, 56(6):1545–1561, 21 March 2011.

[111] Tony Chan, Antonio Marquina, and Pep Mulet. High-Order total Variation-Based image restoration. *SIAM J. Sci. Comput.*, 22(2):503–516, 2000.

[112] Florian Knoll, Thomas Koesters, Ricardo Otazo, Tobias Block, Li Feng, Kathleen Vunckx, David Faul, Johan Nuyts, Fernando Boada, and Daniel K Sodickson. Joint reconstruction of simultaneously acquired MR-PET data with multi sensor compressed sensing based on a joint sparsity constraint. *EJNMMI Phys*, 1(Suppl 1):A26, July 2014.

[113] Daniel F Yu and Jeffrey A Fessler. Edge-preserving tomographic reconstruction with nonlocal regularization. *IEEE Trans. Med. Imaging*, 21(2):159–173, February 2002.

[114] W Chlewicki, F Hermansen, and S B Hansen. Noise reduction and convergence of bayesian algorithms with blobs based on the huber function and median root prior. *Phys. Med. Biol.*, 49(20):4717–4730, 21 October 2004.

[115] Jing Wang, Tianfang Li, and Lei Xing. Iterative image reconstruction for CBCT using edge-preserving prior. *Med. Phys.*, 36(1):252–260, January 2009.

[116] Kevin J Little and Patrick J La Rivière. Sinogram restoration in computed tomography with an edge-preserving penalty. *Med. Phys.*, 42(3):1307–1320, March 2015.

[117] Amir Beck and Marc Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Trans. Image Process.*, 18(11):2419–2434, November 2009.

[118] Benjamin M Yeh, John A Shepherd, Zhen J Wang, Hui Seong Teh, Robert P Hartman, and Sven Prevrhal. Dual-energy and low-kvp CT in the abdomen. *AJR Am. J. Roentgenol.*, 193(1):47–54, July 2009.

[119] Joachim Kopp. Efficient numerical diagonalization of hermitian 3× 3 matrices. *Int. J. Mod. Phys. C*, 19(03):523–548, 2008.

[120] Hans Bornefalk and Mats Persson. Theoretical comparison of the iodine quantification accuracy of two spectral CT technologies. *IEEE Trans. Med. Imaging*, 33(2):556–565, February 2014.

[121] Adam S Wang, Scott S Hsieh, and Norbert J Pelc. A review of dual energy CT: Principles, applications, and future outlook. *CT ???????*, 21(3):368, 2012.