

Optimization of X-Ray Diffraction Imaging of Medical Specimens by Monte Carlo
Methods

by

Matthew Japzon

Medical Physics Graduate Program
Duke University

Date: _____

Approved:

Anuj J. Kapadia, Supervisor

Joel A. Greenberg

Robert E. Reiman

Thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science in the Department of
Medical Physics in the Graduate School
of Duke University

2019

ABSTRACT

Optimization of X-Ray Diffraction Imaging of Medical Specimens by Monte Carlo
Methods

by

Matthew Japzon

Medical Physics Graduate Program
Duke University

Date: _____

Approved:

Anuj J. Kapadia, Supervisor

Joel A. Greenberg

Robert E. Reiman

Thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science in the Department of
Medical Physics in the Graduate School
of Duke University

2019

Copyright by
Matthew Charles Japzon
2019

Abstract

Our research group has previously described the development and testing of a coherent-scatter spectral imaging system for identification of cancer using surrogate phantoms, formalin-fixed pathology tissues and, more recently, surgically resected breast tumor. Here we present the implementation of a Monte-Carlo simulation tool for optimization of the imaging system.

MC-GPU, a GPU-enabled Monte Carlo software was modified and used to simulate X-ray diffraction experiments for combinations of X-ray spectra (tungsten and molybdenum anode), kV (15-150), filtration (material and thickness) and phantom geometry and material (normal, adipose, fibroglandular, and cancerous breast tissue). For each combination, a simulated measurement of contrast-to-noise (CNR), signal strength and object detectability were assessed.

Examination of Monte Carlo simulations showed optimal spectrum characterization strategies that exploit spectral and filter characteristics to increase material identification probabilities via momentum transfer measurement. Increased detectability was shown with molybdenum energy spectra, and a higher CNR metric was observed to show better pathological assessments and findings of cancer.

This work demonstrates the utility of Monte Carlo methods and MCGPU in optimizing coherent scatter imaging systems and can be used to provide insightful information regarding the design of coherent scatter imaging systems for material classification breast tissue types.

Dedication

To my wife, Jessica: you are an incredible gift. Thank you for your encouragement; thank you for your lasting support and love. Without you, this work would not be possible. My motivation to learn, grow and “succeed” is because of you. I hope that I make you proud.

Micah, Jacob, Emma and Lyla, it’s my prayer that you learn from my experiences and that I can impart a desire in you to love people and to learn. You are my greatest blessing.

I’m blessed to know, without doubt, that there are too many more people to formally acknowledge for my academic achievements. My family, my friends, my mentors and my teachers, this work is dedicated to you.

I would be remiss if I didn’t acknowledge the sacrifices of my parents. Mom and dad, you’ve instilled in me a desire to learn and more importantly, to serve. I will always be grateful for you.

Contents

Abstract	iv
List of Figures	x
Acknowledgements	xiii
1. Introduction	1
1.1 Theory	5
1.2 Implementation of XRD Imaging.....	7
1.3 Motivation	11
1.4 X-Ray Transport via Monte Carlo Methods	19
1.5 Coherent Scattering in MC-GPU.....	23
2. Measurements and Metrics for XRD Optimization	28
3. MC-GPU Modification and Comparison	33
3.1 Experimental ADXRD/EDXRD Measurements using MultiX ME100	35
3.2 Simulated ADXRD/EDXRD Measurements using Modified MC-GPU Kernel	37
4. Maximum Signal Energy Bin Characterization.....	42
4.1 Source Spectrum.....	44
4.2 Rapid Simulation via MC-GPU Wrapper	46
4.3 Simulation Set Up.....	47
4.4 Results	49
4.5 Discussion.....	51
5. Cancer Detectability Among Spectra	53

5.1 Embedded Cancer Simulation Details	55
5.2 Results.....	56
5.3 Results.....	59
6. Quantitative Comparison of Tungsten and Molybdenum Anode Spectra.....	60
6.1 Fibroglandular and Cancerous Tissue Simulation and Measures	60
6.2 Results and Discussion.....	61
7. Discussion and Conclusion	64
8. Improvements	67
10. Future Work	68
Appendix A	69
A.1 MC-GPU Modification: MultiX ME100 Emulating Kernel	69
A.2 MC-GPU Modification: 1 keV Energy Bin One-Dimensional Detector	69
A.3 MC-GPU Modification: Energy Discriminating Two-Dimensional Detector	70
References	72

List of Tables

Table 1: X-ray diffraction simulation speed (including initialization) in a simple X-ray diffraction scheme for different architectures and compilation modes.....	22
Table 2: Scanned Material Modeling and Compound Mixture References	37
Table 3: MC-GPU Tungsten Anode Simulation Command Line Arguments	48
Table 4: MC-GPU Molybdenum Anode Simulation Command Line Arguments.....	49
Table 5: Embedded cancer voxels within tissues – XRD measurements made via tungsten anode spectra classified as cancerous tissue (True Positives).	58
Table 6: Embedded cancer voxels within tissues – XRD measurements made via tungsten anode spectra classified as non-cancerous tissue (False Negatives).....	58
Table 7: Embedded cancer voxels within tissues – XRD measurements made via molybdenum anode spectra classified as cancerous (True Positives).	58
Table 8: Embedded cancer voxels within tissues – XRD measurements made via molybdenum anode spectra classified as non-cancerous tissue (False Negatives).	59

List of Figures

Figure 1: XRD form factor shown by measurement of graphite power on a Bruker Phase II diffractometer. Note distinct “Bragg” peaks.....	3
Figure 2: Diffraction patterns in XRD breast tissues [Kidane <i>et al.</i> , 1999] [8].....	4
Figure 3: Rayleigh or Coherent scattering [12]. The incident and exiting photon energy are conserved.....	5
Figure 4: Simplified depiction of ADXRD imaging – A monochromatic spectrum X-ray beam is incident on an object. Diffraction intensity is measured at variable angles that correspond to momentum transfer.	8
Figure 5: Simplified depiction of EDXRD imaging - A variable energy X-ray beam is incident on an object. Diffraction intensity is measured at fixed angles.	9
Figure 6: Simplified depiction of EDXRD and ADXRD used in combination. A polychromatic source is used. X-rays, in the shape of a pencil beam, interact with the object causing diffraction whose diffraction X-rays are detected by a two-dimensional energy-sensitive detector.	10
Figure 7: XRD classification assessment of excised breast tissue following lumpectomy of cancerous tissue. A comparative view is shown between pathological assessment (left), neural network classification and cross-correlation algorithm [22].....	13
Figure 8: Momentum transfer spectrum of fibroglandular and cancerous breast tissue presented to highlight the similarities between the two [8].	14
Figure 9: Attenuation coefficients for soft tissues ($Z \approx 7.5$) are plotted as a function of energy [24]	17
Figure 10: Normalized depiction of a tungsten-filtered tungsten anode X-ray spectrum at 160 kV highlighting two energy bins around the $K_{\alpha 1}$ -edge and the k-edge of tungsten at 59.3 keV and 69.5 keV, respectively.	18
Figure 11: Two-dimensional XRD signal detection showing signal (single-Rayleigh X-rays in red) and signal contamination (Compton and multiple-scatter X-rays) in green, in two different energy bins: 59-61 (left) and 68 to 70 keV (right).	19

Figure 12: MC-GPU single-Rayleigh comparison and normalized form factor reconstruction (showing reference IAA or MIF momentum transfer function) from two identical simulations of XRD using the IAA model (top) and incorporating molecular interference (bottom) for diffraction on 1 cm spheres of water [33] [16].	24
Figure 13: Analytical approximation of F_{IAA} compared to F_{MIF} of water showing alignment at high momentum transfer [16].	28
Figure 14: Diffraction patterns in XRD breast tissues [Kidane <i>et al.</i> , 1999] [8].	31
Figure 15: Breast tissue classification algorithm.	33
Figure 16: MC-GPU output with modified kernel designed to emulate a one-dimensional, energy-sensitive detector. X-rays incident upon the one-dimensional detector array are binned into energies and displayed on the row according to energy bin on the output images. Transmission X-ray not shown.	35
Figure 17: XRD measurement system utilizing. A) Three vertically mounted MULTIX ME100 detectors; B) Beam Block; C) Coded Aperture; D) Scanned object; E-F) Collimators; G) X-Ray Tube.	36
Figure 18: Coherent scattering measurement setup for adipose tissue on the Bruker D2 Phaser	38
Figure 19: An example of smoothing, interpolating a measured momentum transfer function (left) and normalizing the square root of the smoothed and interpolated momentum transfer function of Adipose (right).	39
Figure 20: Experimental MC-GPU comparison – Experimental MultiX ME 100 measurements (left) and MC-GPU simulated MultiX ME 100 measurements (right) for Bone, Water, and Adipose.	39
Figure 21: Reconstruction of the momentum transfer function for Adipose, Bone and Water based of experimentally measured XRD data and MC-GPU Monte Carlo simulation.	41
Figure 22: Percent deviation between simulated and experimentally measure momentum transfer values for MC-GPU comparison.	41
Figure 23: Raw Bruker D2 Phaser measurements made by an ADXRD scan of bone.	42

Figure 24: Tungsten (top) and Molybdenum (bottom) spectra modeled with increasing filter thickness from left to right. Tungsten anode spectra are filtered with 0, 0.25 and 0.5 mm of tungsten and molybdenum anode spectra are filtered with 0, 0.015 and 0.3 mm of molybdenum.	45
Figure 25: Flow Map of Simulation Setup. All files are sent to a common folder vis SSH and the executable MC-GPU file is run with the Input File as the command line argument.	47
Figure 26: Homogenous tissue scans detailing cross-correlation to ground truth and CNR for the highest-signal energy bin using tungsten anode spectra.	50
Figure 27: Homogenous tissue scans detailing cross-correlation to ground truth and CNR for the highest-signal energy bin using molybdenum anode spectra.	50
Figure 28: Simulated XRD scans of homogenous tissues. High CNR outcomes were from the 15 kV molybdenum anode spectra.	52
Figure 29: Subset of momentum transfer functions reconstructed from low CNR measurements.	53
Figure 30: Reconstructed momentum transfer function from simulated XRD measurement of adipose tissue with embedded cancer at various volumetric fractions (left). Reconstructed momentum transfer function of 7/9 cancer/adipose shown with reference momentum transfer functions for fibroglandular and cancer tissues [Kidane <i>et al.</i> , 1999] [8].	55
Figure 31: Voxel combinations used in this experiment. Red voxels are malignant breast tissue and the anterior and posterior gold voxels shown are adipose, normal or fibroglandular tissue.	56
Figure 32: Cross-Correlation based on XRD scans of fibroglandular (top) and cancerous (bottom) tissues. Molybdenum spectra (right) show more differentiation between cancer and fibroglandular tissue when fibroglandular tissue is scanned.	62
Figure 33: CNR observed during XRD scans of fibroglandular and cancer tissue scans.	62

Acknowledgements

My time at Duke University has been incredible. I could not have asked for better colleagues and for better mentors. I came here almost ten years removed from academia and, from day one, I had tremendous support and I've made great friends. Thank you to anyone who took the time to sit down with me and to invest in my success here!

Dr. Anuj Kapadia, my advisor and friend, encouraged me and provided me with unrivaled mentorship. I still can't believe that I accomplished some of the things that he set out for me to accomplish. He patiently explained difficult concepts to me, sometimes more than once. He answered my questions and guided me any day of the week at almost any hour of the day. Dr. Kapadia, thank you.

Dr. Joel Greenberg is incredible resource. He responded to all of my queries and questions and I cannot be more thankful for him. I'm also thankful for the team of scholars that Dr. Kapadia and Dr. Greenberg put together. Dean Hazineh, Camen Royse and Josh Carpenter work so well together and went out of their way to facilitate my experiments. You all will do incredible thing and make significant contributions to this area of research. Thank you!

I must acknowledge my friends and peers who are working or have worked in this field: David Nacouzi, Caley Buxton, Shobhit Sharma and James Spencer. I would be lost without you. Shobhit, you got me started with MC-GPU and sat down with me with some of the most basic concepts. James, you certainly didn't have to, but you took time

to invest in me, probably more than you should have. Caley Buxton, you are an inspiration; your passion to help others less fortunate than you and your friendship propelled me further than you know. David Nacouzi, you are a true friend and you're going to be a world changer! Thank you all for the opportunity to work with you and further XRD research at Duke.

Lastly, Dr. Andreu Badal-Soler, and Dr. Bahaa Ghammraoui, I appreciate your time and your expertise with MC-GPU. It's open source products like MC-GPU that will make this world a better place. Thank you for taking the time to work with me and Dr. Kapadia and thank you for working so hard on such an impactful tool.

1. Introduction

Breast cancer is the most common cancer among women in the United States [1]. More than one in every ten women will receive a breast cancer diagnosis sometime in their lifetime. Furthermore, in 2018, it was estimated that 40,920 women lost their lives as a result of breast cancer [2]. In contrast, we have made significant improvement in patient outcomes throughout the years: patients who were diagnosed with breast cancer in 1975 to 1977 had a 74.8% five-year survival rate whereas patients diagnosed with breast cancer between 2005 and 2011 now have a 90.7% five-year relative survival rate [3]. This increase in patient five-year survivals is largely due to advances in early cancer detection, through the development of advanced imaging techniques and imaging modalities [4].

Many traditional transmission-based ionizing imaging techniques, such as mammography, obtain image contrast through the exploitation of the tissue's inherent ability to attenuate X-rays. These modalities are limited in their ability to differentiate between tissue types because, often times, body tissues that are inherently different may have similar attenuation properties [5] [6]. In fact, healthy and diseased breasts exhibit only a 3% difference between linear attenuation coefficients at mammography energies [7] [8]. Optimization of mammographic imaging techniques aims to highlight these small differences in tissue attenuation; however, limits to attenuation-based imaging of these tissues arise because of this natural phenomenon.

Our research group has focused much of our interest around breast cancer and breast cancer detection; however, we have explored an X-ray imaging modality that aims to provide additional detail and perhaps differentiation between tissues of similar attenuation properties: X-ray diffraction (XRD). XRD is a unique and non-invasive method to interrogate objects for material identification and material characterization. This technique has been historically used in crystallography, in which X-rays have been shown to interact with crystalline structures such that their diffraction gives insight into the crystal's material composition, inter-atomic and inter-molecular structure [9]. X-rays incident on crystalline structures interact, via diffraction, within the crystal lattice giving rise to X-ray diffraction patterns known as scatter form factors, where peaks are observed at clearly defined angles based in the crystal lattice structure and spacing. Figure 1 shows XRD measurement of graphite powder; here X-rays diffract off the graphite powder and show significant intensity at 26.7° , a unique characteristic of graphite XRD measurements.

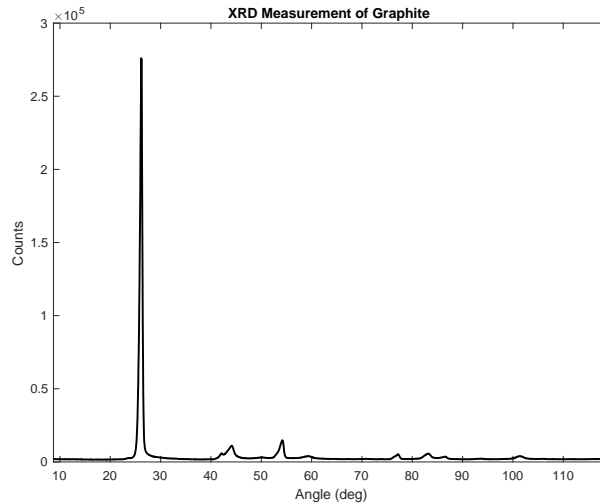


Figure 1: XRD form factor shown by measurement of graphite power on a Bruker Phase II diffractometer. Note distinct “Bragg” peaks.

In more recent years, X-ray diffraction has been expanded, beyond its uses in crystallography, to the interrogation of liquids and amorphous material. XRD has been used in security applications as a method to identify hazardous and explosive materials in baggage. It has been shown to be useful in the classification and detection of illicit drugs and because of its non-destructive nature, it has shown to also be useful in environmental and industrial settings [10].

As each of these areas pursue their own tangible uses for X-ray diffraction imaging, our research group, among others, has focused our efforts to study X-ray diffraction of biological tissues. While it is true that diffracted X-rays in current medical imaging technologies are commonly ignored or actively suppressed, it has been shown that X-ray diffraction in biological tissues gives rise to form factor measurements that are unique. In the context of breast imaging, XRD has shown promise in differentiating

common breast tissue types, and even breast cancers. This is apparent as each breast tissue type, when measure via XRD, produces diffraction patterns with unique peak location and defining characteristics. [8] [11].

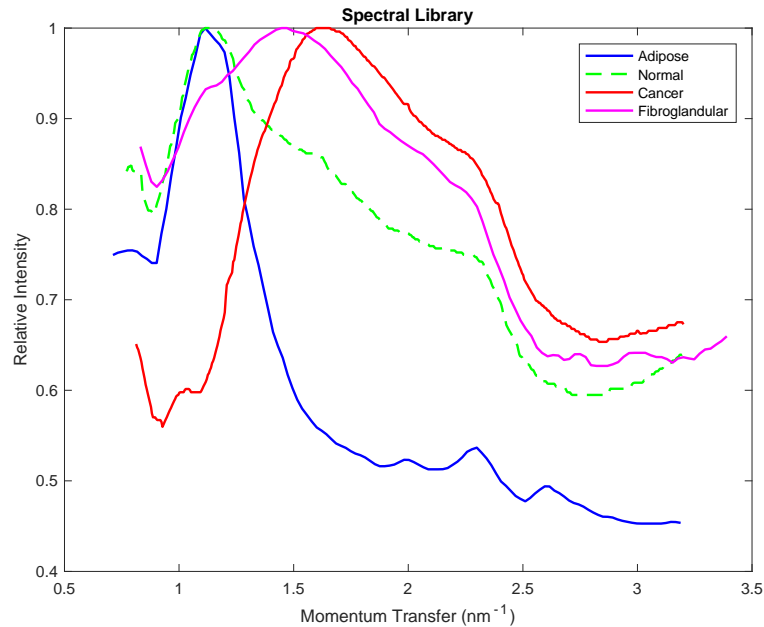


Figure 2: Diffraction patterns in XRD breast tissues [Kidane *et al.*, 1999] [8]

Figure 2 shows X-ray diffraction patterns in breast tissues. Our research group is dedicated to the imaging of these types of materials, among others, and, although clear differentiation is shown, it is important to note how similar these diffraction patterns can be. This study asks how we can optimize our imaging techniques to highlight the differences in these form factor patterns and develop tools that are sensitive enough to detect and exploit them. We have chosen the use of computer simulation to carry out this study and aim to optimize one of the most critical elements in XRD imaging systems: the X-ray source.

1.1 Theory

XRD imaging hinges on the detection of coherent scatter X-rays. Among the different X-ray scattering methods, coherent scatter, also known as Rayleigh scatter, is most prevalent at low energies (below 30 keV) and, in crystalline structures, is governed by Bragg's Law:

$$n\lambda = 2d \sin\left(\frac{\theta}{2}\right) \quad (1)$$

Here, X-rays are incident upon an atomic structure within the crystalline lattice. Should Rayleigh scatter occur in a crystal, the X-ray, of wavelength λ , will diffract at a Bragg angle $\left(\frac{\theta}{2}\right)$ according to the "nth" order diffraction and the inverse interplanar lattice distance d [9] [10]. Because of this strict adherence to Bragg's law, we observe unique peaks in crystalline XRD imaging that serve as the "fingerprint" for crystalline material identification.

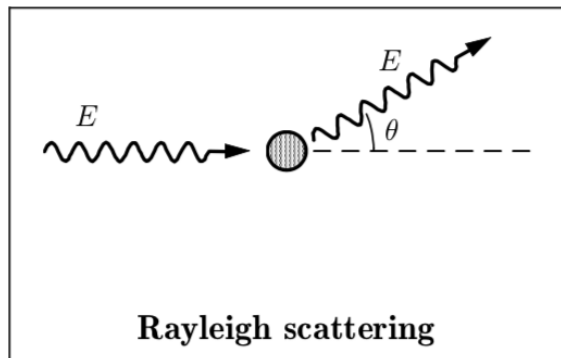


Figure 3: Rayleigh or Coherent scattering [12]. The incident and exiting photon energy are conserved.

It is important to recognize that coherent scattering is an elastic scattering event; kinetic energy is conserved and the scatter X-ray energy will not be affected by it.

Because of this, we also define momentum transfer, q , which is a measure of the momentum transferred to the scatter X-ray. This is derived from Bragg's law and yields a measure that is both energy and angle dependent, by solving for the inverse of twice the crystalline lattice structure distance $\frac{1}{2d}$:

$$q = \frac{1}{2d} = \frac{1}{\lambda} \sin\left(\frac{\theta}{2}\right) = \frac{E}{hc} \sin\left(\frac{\theta}{2}\right) \quad (2)$$

Just as crystalline structures show unique diffraction patterns, we can measure form factor patterns of amorphous materials as a function of momentum transfer and observe unique and characterizing shapes. Amorphous material, each unique in atomic composition and molecular spacing, show unique X-ray diffraction that gives rise to a distinguishable scatter form factor.

For an isolated atom, XRD is shown to give rise to patterns described by the atomic form factor (F_{IAA}). This factor is a function of the atomic number, Z , and of the momentum imparted to the scatter X-ray, again, momentum transfer, q . This function, $F(q, Z)$, which can be expressed as the Fourier transform of the atomic electron density, has been shown to monotonically decrease as a function of momentum transfer, where $F(0, Z) = Z$ and $F(\infty, Z) = 0$ [13] [14]. As we consider molecules comprised of different atoms, the diffractions concerning interatomic spacing is governed by the additivity rule:

$$F_{IAA}^2 = \sum n_i F^2(q, Z_i) \quad (3)$$

where n_i is the weight fraction and Z_i is the atomic number of element i .

We discuss the atomic form factor F_{IAA} , not only because amorphous X-ray diffraction is dependent on it, but also because it has been shown that the high momentum region of XRD profiles is dominated it, after which, molecular interference effects can be ignored [15] [6] [16]. In other words, measured amorphous material X-ray diffraction approaches the atomic form factor at high momentum transfer values, after which, molecular spacing does not contribute to changes in the diffraction pattern. This will play an important role as we, later, discuss coherent scattering via computer simulation.

1.2 Implementation of XRD Imaging

Knowing that X-ray diffraction is a function of both X-ray energy and scatter angle, leads to the implementation of two distinct X-ray diffraction imaging techniques: angular-dispersive X-ray diffraction (ADXRD) and energy-dispersive X-ray diffraction (EDXRD). The principle behind these imaging techniques is to hold one of the two diffraction variables constant during X-ray measurement, while adjusting the other.

ADXRD requires a monochromatic incident X-ray source. X-rays of a fixed energy are directed towards the surface of an object and X-ray diffraction measurements are taken at fixed angles, which correlate to fixed momentum transfer values. This method is favorable for specimen analysis and is used in most commercial

diffractometers because they can achieve excellent momentum transfer resolution through the use of long wavelength sources. Our lab uses a Bruker D2 Phaser ADXRD diffractometer, specifically for specimen analysis. The process to measure sufficient signal with these types of diffractometers is generally time consuming; the detector is held at each angle for a fixed period of time. A sufficiently sampled momentum transfer for our work is sometimes in excess of 40 minutes.

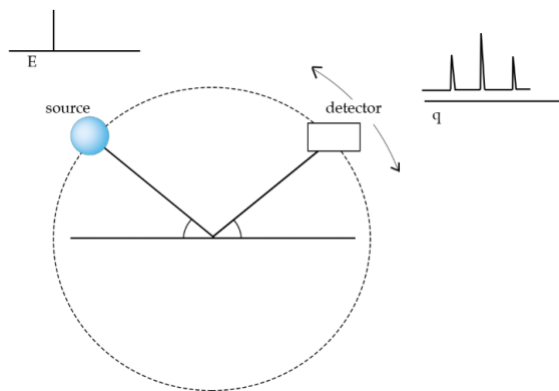


Figure 4: Simplified depiction of ADXRD imaging – A monochromatic spectrum X-ray beam is incident on an object. Diffraction intensity is measured at variable angles that correspond to momentum transfer.

EDXRD imaging systems vary the incident X-ray source energy and measure the intensity of the diffracted X-rays at fixed angles. As the source energy is increased, the intensity measurements at these fixed angles are directly linked to the intensity observed at higher momentum transfer. The benefit of these systems is that conventional X-ray tubes may be used for these types of measurements and that a broadband spectral source coupled with energy-sensitive detectors can be used to simultaneously sample q -space.

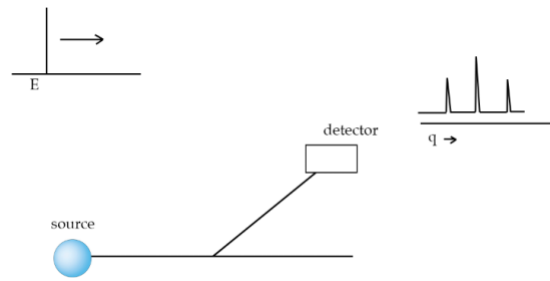


Figure 5: Simplified depiction of EDXRD imaging - A variable energy X-ray beam is incident on an object. Diffraction intensity is measured at fixed angles.

Our lab has adopted a hybrid detection method to measure the diffraction scatter pattern. Many of our systems use energy discriminating detectors at fixed locations. Because large-area energy sensitive detector costs can be high, we have also built systems that translate small energy-discriminating detectors (often one-dimensional detector arrays) to sample large ranges of momentum transfer. By using a detector with the ability to resolve the energy of the detected signal, we can realize the scatter form factor by isolating a detector energy bin and observing the diffraction intensity as a function of the radial distance of the detector pixels orthogonal to the X-ray beam. This allows us to adapt more traditional X-ray sources and use sources and detectors commonly found in our area of medical research. This combination of multiple-fixed-angle detector bins and polychromatic source is beneficial in that we can simultaneously sample different ranges of q -space at different momentum transfer resolutions, with each measurement.

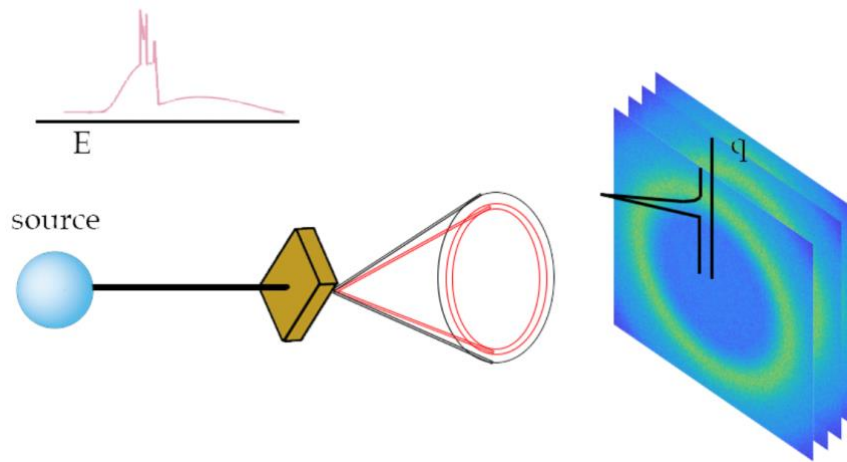


Figure 6: Simplified depiction of EDXRD and ADXRD used in combination. A polychromatic source is used. X-rays, in the shape of a pencil beam, interact with the object causing diffraction whose diffraction X-rays are detected by a two-dimensional energy-sensitive detector.

Figure 6 shows a simplified schematic of an EDXRD/ADXRD imaging system. In addition to the use of a polychromatic source and energy-sensitive detector, this depiction shows a two-dimensional X-ray diffraction scheme. In this scenario momentum transfer is a function of energy, radial distance from the primary beam and the object to detector distance. Both beam divergence and the use of a flat detector give rise to the effects of the inverse square law and obliquity which reduces the intensity at higher values of momentum transfer [17].

It is important to understand that many X-ray diffraction experimental setups used powdered crystals to force randomly oriented crystal patterns to be exposed to the primary beam. This practice is done to observe scatter from all possible lattice spacings consistent with Bragg's law in crystal analysis and due to the random orientation of biological cells and molecular structures. It is important to understand two-dimensional

ring patterns are also observed during XRD of amorphous materials when considering XRD imaging fundamentals.

1.3 Motivation

To this point, we have explored how XRD is used in many different fields to non-destructively probe material specimen to better understand their material properties. The physics of the coherent scatter phenomenon should be understood and methods to measure diffraction have been shown. With a keen interest in breast cancer and breast cancer detection, our research group believes that XRD should be considered to have significant application in medicine, especially in breast cancer research.

Consider the clinical scenario where a breast cancer patient undergoes breast conservation surgery (BCS) with the purpose of removing cancerous tissue found in her breast. Following surgery, the excised tissue is examined and a determination is made whether or not the margins surrounding the excised tissue contain cancerous cells – this yields information about whether or not all intended incidences of cancer were removed. The excised tissue is sent to pathology and the pathologists' finding, days after surgery, indicates positive margins; breast conservation surgery did not completely remove the patient's cancerous tissue. The patient is then told that the breast conservation surgery was incomplete and that the patient must return for an attempted re-excision. One study showed that over 10% of the retrospective cases studies studied

had focally positive resection margins and more than 6% had more than focally positive resection margins [18], an indication that this scenario may not be a rare occurrence.

Our research has been dedicated to helping to counteract this problem. In 2015, Lakshmanan et al. described an X-ray coherent scatter imaging technique for intra-operative margin detection that was shown to have greater accuracy and speed than current techniques found in practice [19]. Proposed systems could be clinically utilized to conduct non-invasive XRD “spot scans” and characterize excised breast tissue with high resolution in real time [19]. Spencer et al. showcased a system designed for this task – a proposed system that could be used in real time to provide indications during survey about whether or not cancer was probable at the excised tissue margins [20].

In addition to the characterization and implementation of such XRD imaging systems, we have shown that measurements from these systems are unique enough to be automatically classified using algorithms [21]. Momentum transfer functions of common breast tissue types have features that may be explored to differentiate these tissues based on XRD scans, alone.

In a recent study, Nacouzi et al. obtained excised breast tissue of invasive ductal carcinoma resected during a performed lumpectomy, which was preserved in formalin [22]. Our experimental coded aperture coherent scatter spectral imaging (CACSSI) system, an XRD imaging system similar to the system described in detail in Section 3.1 of this work, was used to examine X-ray diffraction from 651 separate locations on the

preserved lumpectomy specimen. Each location was probed via an X-ray pencil beam produced by a tungsten/rhenium anode X-ray tube tuned to 110 kV and 100 mAs.

XRD measurement was examined via two classifier algorithms: a weighted cross-correlation-based algorithm previously developed in our group and a neural network classification protocol implemented in MATLAB [21]. Both classifier systems, when compared to a pathologist's assessment, showed promising results (greater than 79% accuracy) in differentiating healthy and cancerous tissues found throughout the specimen. The study concluded that these classification methods may have been even more accurate, as the back side of the specimen, which may have shown indications of cancer, was not considered in the pathology analysis, whereas it was included in the XRD measurement [22].

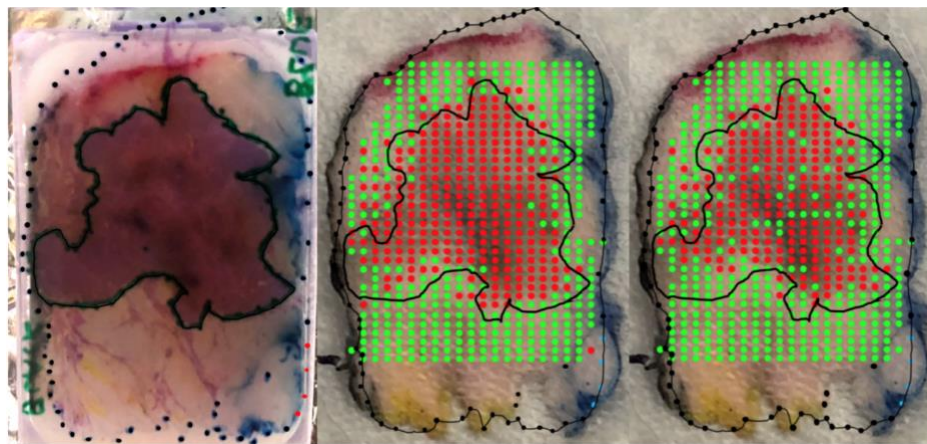


Figure 7: XRD classification assessment of excised breast tissue following lumpectomy of cancerous tissue. A comparative view is shown between pathological assessment (left), neural network classification and cross-correlation algorithm [22].

All this to say, much thought went into the possibility of making pathological assessments of breast tissue specimen via XRD measurements. These classification methods are only effective, however, through the acquisition of cleanly measured and accurate data. Consider, for example, fibroglandular and cancerous breast tissue, whose coherent scatter form factors are similar above $q=1.5 \text{ nm}^{-1}$ (Fig 8). Distinguishing these two materials is a challenge due to the similarity.

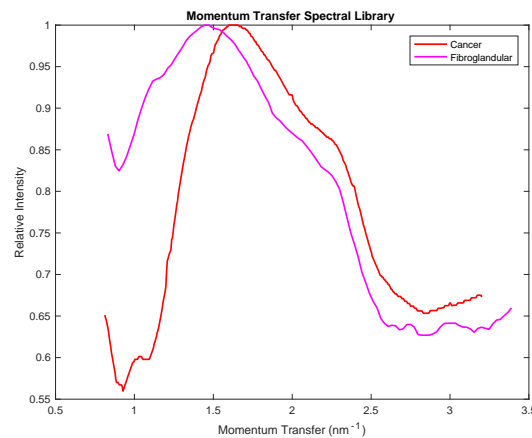


Figure 8: Momentum transfer spectrum of fibroglandular and cancerous breast tissue presented to highlight the similarities between the two [8].

If we consider the material properties of these two tissue types, it should be apparent why their diffraction patterns would be similar. They are both human tissues, likely with comparable concentrations of hydrogen, oxygen, carbon and nitrogen [23]. Based on their atomic make-up, alone, and considering what we know about the F_{IAA} model, they exhibit similar atomic form factors and their deviations in diffraction patterns are primarily due to small deviations in their molecular interference properties. Nonetheless, under poor measurement conditions, differentiation between these two

tissue types could easily lead to classification error, no matter how robust the classification algorithm might be. As a result, there is a need to generate large amounts of data to train and test classification algorithms. Given the complexity of the overall acquisition system, there is also a strong need to optimize the various acquisition parameters and identify the best combination for a given detection task.

This work is dedicated to building a simulation tool to model virtual X-ray diffraction acquisition experiments and generate data to optimize the imaging system and improve classification outcomes and material. As discussed previously, we have employed the use of fixed energy-sensitive detectors and a hybrid ADXRD/EDXRD measurement system. Our detection method of combining ADXRD and EDXRD is beneficial in that we can sample a large range of momentum transfer at varying resolution through the use of a polychromatic source and energy-sensitive detector.

The use of a polychromatic source spectrum gives rise to “signal contamination” in polychromatic-source XRD measurements. One prevalent X-ray interaction at coherent scattering energies that gives rise to signal contamination is Compton scatter. Compton scatter is an incoherent scattering event where a photon of energy E_γ interacts with an atomic electron. Here, the electron re-emits a secondary Compton photon of energy $E_{\gamma'}$ in a trajectory that is preferentially forward, but likely with a non-zero scatter angle relative to the incident angle:

$$E_{\gamma'} = \frac{E_{\gamma}}{1 + \left(\frac{E_{\gamma}}{m_e c^2}\right)(1 - \cos \theta)} \quad (3)$$

By using a polychromatic beam, signal contamination from Compton scatter X-rays in a particular detector energy bin window is dependent on the source spectrum intensity of the energies above our energy bin window. This means that regardless of the energy bin selected to reconstruct the form factor measurement, signal contamination from energies higher than that energy bin have the opportunity to affect the measurement that may subsequently lead to a classification error.

Similarly, multiple scatter events could contaminate the measurement. Multiple-scatter events are those that have undergone more than one scatter interaction (coherent, Compton, and any combination of these), causing these X-rays to lose their most valuable piece of information, which is their first coherent scattering angle. This factor is not explicitly linked to the use of a polychromatic source, but it is linked in to the source selection in that coherent scatter interactions are energy dependent.

The probabilities that govern these interactions for soft tissues are shown in Figure 9. Here, we see the probability of Compton scattering overtake Rayleigh scattering just above 10 keV and overshadow all other interactions in the 100s of keV. Therefore, it is expected that single-Compton signal contamination occurs with higher keV spectra whereas multiple-scatter via multiple-Rayleigh events occurs at lower energies. Multiple-scatter, in general, will increase as specimen size increases.

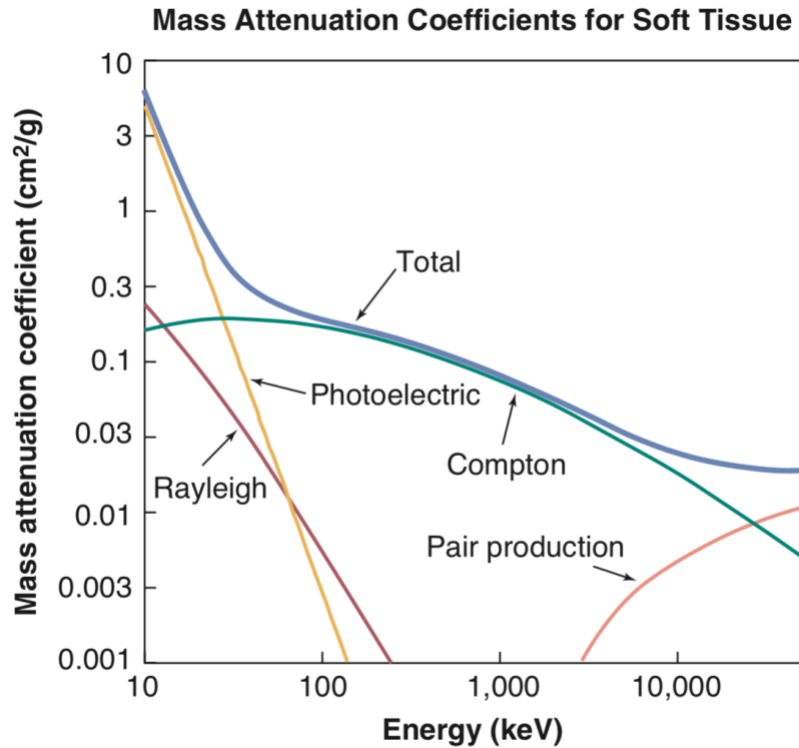


Figure 9: Attenuation coefficients for soft tissues ($Z \approx 7.5$) are plotted as a function of energy [24]

Typically, our lab has designed XRD experiments such that most of our form factor reconstruction is based on the detector energy bin that is expected to receive the most signal. For instance, a tungsten anode spectrum at 160 keV is expected to show characteristic peaks at 59.3 keV and 69.5 keV, which correspond to the $K_{\alpha 1}$ and k-edge of tungsten [25]. If we were to place a tungsten filter between the X-ray source and the target object, much of the X-ray spectrum above the 69.5 keV k-edge would be attenuated in the filter before it reached the object for X-ray interrogation. However, our strongest signal would still be expected around the 59.3 $K_{\alpha 1}$ -edge, as shown in Figure 10.

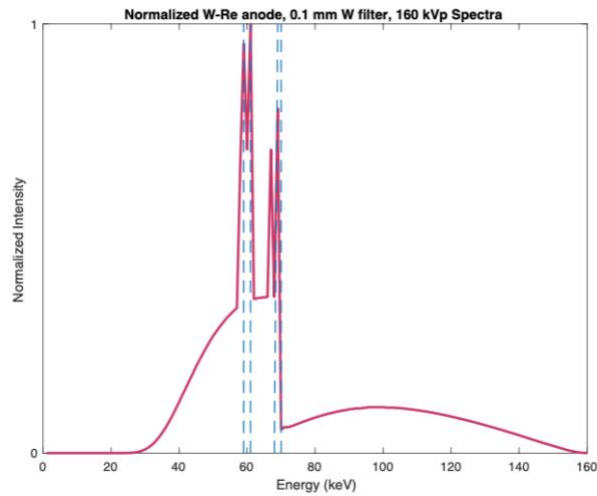


Figure 10: Normalized depiction of a tungsten-filtered tungsten anode X-ray spectrum at 160 kV highlighting two energy bins around the $K_{\alpha 1}$ -edge and the k-edge of tungsten at 59.3 keV and 69.5 keV, respectively.

Basing our form factor reconstruction on data observed in the energy bin with the maximum signal (i.e., 59-61 keV) may not be optimal due to potential signal contamination in this energy bin. For example, consider the spectrum between 59.3 and 70 keV. Here, the signal is relatively strong, and it is possible that X-rays produced in this range could undergo Compton scatter whose secondary X-rays would fall within our 59-61 keV binning window.

This binning form factor reconstruction strategy is in contrast to a binning strategy that considers the X-rays at 69.5 keV. In this higher window, Compton scatter X-rays are likely less prevalent, simply because there are fewer source X-rays present above 69.5 keV to experience a Compton scatter event. Figure 11 shows how the detected signal (single-Rayleigh X-rays) and signal-contamination (Compton-scatter and

multiple scatter X-rays) may look from a two-dimensional energy-resolving detector with these two binning strategies from an XRD scan of powdered graphite:

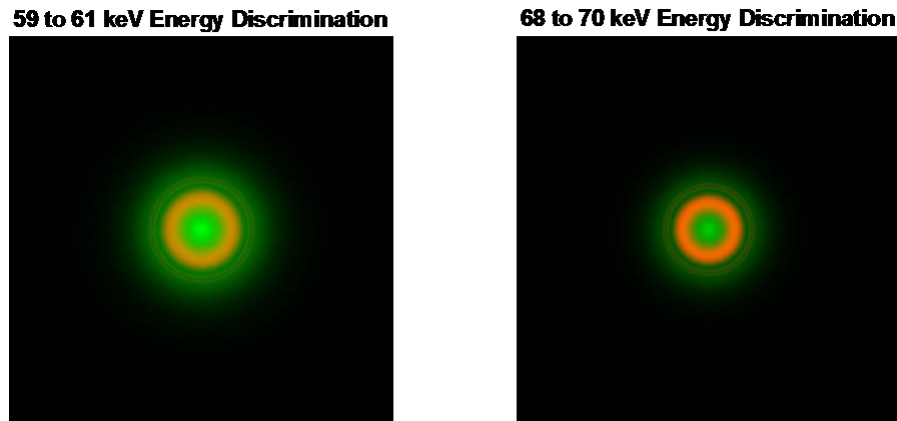


Figure 11: Two-dimensional XRD signal detection showing signal (single-Rayleigh X-rays in red) and signal contamination (Compton and multiple-scatter X-rays) in green, in two different energy bins: 59-61 (left) and 68 to 70 keV (right).

Our research surrounding pathological XRD breast imaging has experimentally shown that form factor reconstruction and sensitivity is greatly dependent on X-ray source strength and potential [21]. However, little is known about how the X-ray source spectrum and how unwanted X-ray interactions could affect our signal detectability, form factor reconstruction and ultimately our classification estimations.

This study asks aims to utilize Monte-Carlo simulations to uncover these effects and develop a platform for optimization of our X-ray diffraction imaging approach.

1.4 X-Ray Transport via Monte Carlo Methods

Medical X-ray imaging has a unique history in which computer simulations guide the optimization of imaging systems. In 2012 Cunham et al. defined a figure of

merit (FOM) aimed at maximizing contrast-to-noise and minimizing and average glandular dose to the breast that a wide variety of Monte Carlo simulations. Their work showed the efficacy of Monte Carlo in mammography by allowing them to simulate typical mammographic studies and confirm that many techniques commonly used in the field could be quantifiably attributed to best practices that promote optimal image quality and patient safety [5]. Similar studies were done to estimate dose in CT and assess dose image quality in chest radiography [26] [27].

Because the physical effects, radiation interactions and nuclear processes involved in X-ray imaging are so interdependent, Monte Carlo is a proven effective tool in this area of research. Our lab has validated Monte Carlo methods during its pursuit to design and test our previously mentioned CACCSI XRD imaging system [19]. The Monte Carlo suite previously used to model our CACCSI system was GEANT4, which, among other common Monte Carlo X-ray suites like PENELOPE and MCPNX, was not originally designed to model coherent scattering events. However, Lakshmanan et al. described modifications to model coherent scattering effects accurately in GEANT4.

This work intends to build on the accuracy of modeling X-ray diffraction in Monte Carlo simulation and also improve the speed of our X-ray transport simulations. While accurate and validated for modeling of coherent scatter effects, our modifications to GEANT4 require these simulations to be run on CPU. Since the advent of GPU processing, we have sought to model X-ray transport in a cost effective and time efficient

manner and has done so, to estimate CT dose, with a MC toolkit called MC-GPU [28]. MC-GPU provides an ideal platform to develop the proposed X-ray diffraction simulations for optimization of our CACSSSI system.

MC-GPU is a Monte Carlo X-ray transport framework used to generate X-ray and CT radiographic images by modeling X-ray transport within materials organized in a voxelized phantom [29] [30] [31] [32]. Material properties modeled in MC-GPU have been adapted from the PENELOPE 2006 MC Tool Kit. The processing behind MC-GPU is facilitated by GPU through the CUDA framework, a parallel computing platform and application programming interface model created by NVIDIA which boasts increased simulation speeds between 20 and 40 times faster compared to a single CPU core [29] [33].

To emphasize the reduced time costs MC-GPU offers, simple X-ray diffraction simulations were compared across different architectures and compilation schemes for a setup involving a polychromatic spectrum of X-rays, $1e8$ histories in total, interacting with a 2.5 cm tall cylinder of water with a 0.65 cm radius. The results of these comparisons are detailed in Table 1:

Table 1: X-ray diffraction simulation speed (including initialization) in a simple X-ray diffraction scheme for different architectures and compilation modes.

Architecture	Total Run Time (s)	X-Rays/s
GEANT 4 (4 core)	13,501.82	7,406.40
MC-GPU: CPU (32 core)	32.49	3,077,870.11
MG-GPU: GPU (no fast math)	2.65	37,738,264.15
MC-GPU: GPU (fast math)	2.00	50,003,200.00

The results of the CPU simulations show a significant speed up by use of MC-GPU.

Even after normalizing to the number of cores, we see an increase in simulation speed 51.9X faster. It should also be noted that Badal et al. showed an 8.2 times speed increase between simulations run on the PENELOPE 2006 architecture and MC-GPU run on CPU [29]. Compilation of MC-GPU using NVIDIA’s nvcc compiler is compliant with the IEEE standard for binary floating-point representation, with few exceptions. When the compilation option of “fast math” is used, some mathematical operations computed on the host system are completed using function calls designed to increase speed at the cost of accuracy [34]. MC-GPU simulations from code compiled with the “fast math” option have been shown to be qualitatively equivalent with a maximum pixel difference of 2.5% [29].

Whereas diffraction modeling in the coherent scattering implementations of MC simulations is typically ignored, MC-GPU is designed to incorporate coherent scattering effects. When compared to GEANT4, PENELOPE, and MCPNX, it has the benefits of both modeling coherent scattering and GPU speed and, as such, will be used in this work.

1.5 Coherent Scattering in MC-GPU

Initial versions of the PENELOPE 2006 Monte Carlo Tool Kit and MC-GPU modeled coherent scattering in a material using an independent atomic approximation [33]. The independent atomic approximation, while accurate for modeling large angle scatter, is inaccurate for the purposes of this work.

In 2012, Bahaa Ghamraoui and Andreu Badal introduced an update to MC-GPU and the PENELOPE 2006 toolkit [33]. These updates allow the MC-GPU simulation model to account for the effect of molecular interference in coherent scattering.

Experimentally measured scatter form factors (that incorporate both the IAA and molecular interference), $F_{MIF}(x)$, can now be processed and modeled in the X-ray transport of coherent scatter X-rays. The user is asked to provide experimentally measured values for $F_{MIF}(x)/\sqrt{MW}$ or $s(x)$:

$$F_{MIF}^2(x) = F_{IAA}^2(x) \times s(x) \quad (4)$$

where MW is the molecular weight and $s(x)$ is the molecular interference function.

Figure 12 shows coherent scattering modeling from two identical experiments. One simulation shows coherent scatter under the IAA model and the other shows incorporation of molecular interference via MC-GPU. A reconstruction of momentum transfer is shown for each model and will be discussed in the following section.

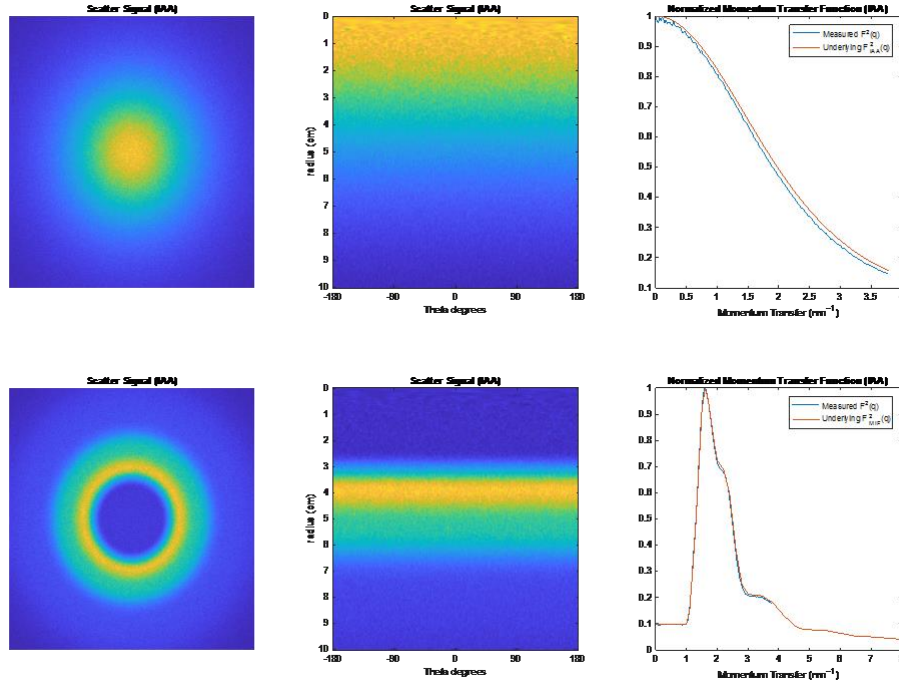


Figure 12: MC-GPU single-Rayleigh comparison and normalized form factor reconstruction (showing reference IAA or MIF momentum transfer function) from two identical simulations of XRD using the IAA model (top) and incorporating molecular interference (bottom) for diffraction on 1 cm spheres of water [33] [16].

The scatter signals shown in Figure 12 are the result of identical simulated XRD experiments using MC-GPU. A monoenergetic X-ray source was oriented towards a 1cm sphere of water. The first simulation (top) modeled coherent scattering using the IAA, while the second simulation (bottom) shows the incorporation of molecular interference into the coherent scatter modeling. A 400x400 pixel, 20x20cm, energy-integrating detector was placed 19 cm from the beam exit of the water sphere and single-coherent scatter X-rays were tallied in units of eV/cm².

The momentum transfer function, $F^2(q)$, was constructed, first by converting these scatter intensities to a polar view and summing the intensities across each row.

The effects of the inverse square law and obliquity are compensated for and the intensity of the scatter signal is increased in each pixel according to its radial distance from the primary beam:

$$I_r = I_{d,r} \cos^{-3} \theta \quad (5)$$

where $I_{d,r}$ is the detected intensity at a detector pixel at a radius, r ; θ is the angle between the primary beam and the detector pixel x and I_r is the measured scatter intensity at a radius r . Finally, the radius was mapped onto an axis of momentum transfer by algebraic manipulation of Equation (2):

$$q = \frac{E}{hc} \sin(\tan^{-1}(\frac{r}{ODD}) \times \frac{1}{2}) \quad (6)$$

where ODD is the object-to-detector distance.

Because the measured object is relatively thin, and the location of the object (i.e. the z -location of our scatter events) is limited, these methods to reconstruct the material form factor are sufficient. However, with large specimens, where the scatter location is unknown, it is possible to reconstruct and determine the molecular structure of an object at each voxel throughout its volume with advanced techniques such as X-ray diffraction tomography (XRDT), or coded aperture XRDT (CA-XRDT). These techniques make multiple or parallel multiplexed XRD measurements that give rise to a more precise angle of diffraction and therefore scatter location when de-multiplexed [10] [35].

Incorporating molecular interference into MC-GPU and using experimentally measured form factors for Monte Carlo does require additional manipulation. Although,

MC-GPU offers the user the opportunity to feed in data files listing $F_{MIF}(x)/\sqrt{MW}$ or $s(x)$, in experimental setups, it should be apparent that the magnitude of XRD measurements are a function of both source intensity and a function of scan time. Because of this, measured XRD data must be appropriately scaled to accurately represent $F_{MIF}(x)$ or be used to calculate $s(x)$. This is important because the coherent scatter interaction probabilities determined in PENELOPE physics are keenly reliant on the magnitude of $F_{MIF}(x)$:

$$\frac{d\sigma_{Ra}}{d\Omega} = \frac{d\sigma_T}{d\Omega} [F(q, z)]^2 \quad (7)$$

where

$$\frac{d\sigma_T}{d\Omega} = r_e^2 \frac{1 + \cos^2 \theta}{2} \quad (8)$$

where r_e is the classical electron radius and θ is the angle of deflection angle of the exiting X-ray [12].

In order to properly scale an experimentally measured molecular form factor, it is important to understand that the F_{IAA} and F_{MIF} converge at high momentum transfer [11] [16]. We accomplish this by analytical approximation. First, $F_{IAA}(q, Z_i)$, also known as the Hubbell form factor, is calculated for momentum transfer values, q , from 0 to 8 nm^{-1} :

$$F(q, Z) = \begin{cases} f(x, Z) \equiv Z \frac{1 + a_1 x^2 + a_2 x^3 + a_3 x^4}{(1 + a_4 x^2 + a_5 x^4)^2} \\ \max\{f(x, Z), F_k(q, Z)\} \text{ if } Z > 10 \text{ and } f(x, Z) > 2 \end{cases} \quad (9)$$

where

$$F_k(q, Z) \equiv \frac{\sin(2b \arctan Q)}{bQ(1 + Q^2)^2} \quad (10)$$

with

$$Q = \frac{q}{2m_e c a'}, \quad b = \sqrt{1 - a^2}, \quad a \equiv \alpha \left(Z - \frac{5}{16} \right) \quad (11), (12), (13)$$

which is described by Baró et al. and was adapted from PENELOPE 2006 to MATLAB [14] [12]. The inverse fine-structure constant, α^{-1} , has been updated to reflect the value reported in the NIST database containing the 2014 CODATA recommended values of fundamental physical constants [36].

Finally, the additivity rule is applied according to the weight fraction of a compound or mixture and the values of the weighted Hubbell form factor at high momentum transfer are used to normalize XRD measured form factors. Figure 13 shows the analytical F_{IAA} for water compared with an experimentally measured form factor for water.

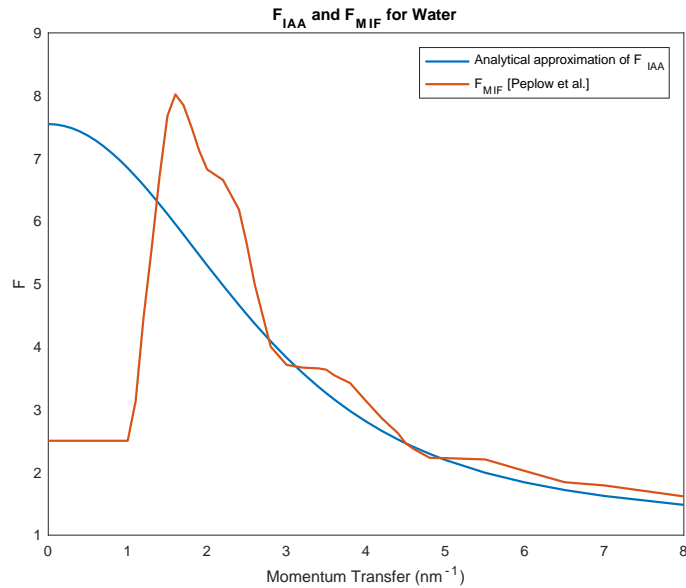


Figure 13: Analytical approximation of F_{IIAA} compared to F_{MIF} of water showing alignment at high momentum transfer [16].

Scaling experimentally measured form factors to analytically approximated F_{IIAA} is done to calculate and input the molecular interference function $S(x)$ into our Monte Carlo model without the need to measure the molecular weight.

2. Measurements and Metrics for XRD Optimization

MC-GPU was designed to model PENELOPE physics and meant to be used to generate realistic CT and radiographic images. The output of MC-GPU is standard and is in the form of two identical files used to store the same information in two separate file types for various post-processing steps.

The first, an ASCII data file, lists information regarding the simulation study included as header information above and below the main body: focal spot, pixel size,

number of pixels, simulated X-rays, simulation speed/time, fraction of energy incident on the detector and the maximum energy detected.

This study considers the information stored in the data file. The body of the data file contains the energy fluence (eV/cm^2) incident upon the detector and it segregates this energy fluence into four categories depending on the type and number of interactions experienced by the X-rays during simulation. X-rays are grouped into four categories: single-Rayleigh, single-Compton, transmission, and multiple-scatter. As mentioned previously, a duplicate file containing identical information is provided by MC-GPU. This file is a raw image file that mirrors the detector information in the data file, with the addition of a fifth image showing a simulated radiograph of all incident X-rays, which is the sum of the original four.

In light of this, it is possible to understand the amount of pure signal that is available for form factor reconstruction and it is possible to identify the extent to which other X-rays contribute to the contamination of this signal. This work differentiates between signal (single-Rayleigh X-rays incident on the detector) and noise (any X-ray that is detected and detracts from the accurate reconstruction of the true scatter form factor of the material; i.e. Compton and multiple-scatter X-rays). Because practical implementations of X-ray diffraction purposefully attenuate transmission X-rays with the use of a high Z beam block, transmission X-rays are not considered.

Signal is measured as the sum of the single-Rayleigh signal incident on the detector. At a basic level, an estimation of true signal, or an estimation of the relative signal under theoretically perfect conditions, is useful to understand how variation in detected (total) signal may affect form factor reconstruction. As we intend to use an energy discriminating detector, it is expected that the maximum detected signal is located at one of the k-edge peaks of our spectrum, or at the bremsstrahlung peak at lower kV.

Signal-to-noise (SNR), in this context of this work, is considered to be the ratio of detected single-Rayleigh X-rays (signal) to the Compton and multiple-scatter X-rays (noise) detected, but this measurement is considered to be an absolute measurement. Contrast-to-noise (CNR), however, is a proposed metric that will be evaluated:

$$CNR = \frac{Signal - Noise}{Noise} \quad (14)$$

This metric is preferred over SNR because it removes bias by normalizing the difference between signal to noise with the noise.

Finally, it is useful to observe how variations in signal and noise affect classifier performance on a given measurement. The goal of this optimization study is to increase the confidence that the signal detected is as close as possible to the scatter form factor of the object. A measure of cross-correlation, which is a measure of how similar two functions are as a function of displacement, is proposed to assess how closely the measured signal is to ground truth. As the form factor measurement matches more

closely to the reference form factor used in the Monte Carlo model, the cross-correlation will approach unity:

$$Correlation = \frac{\sum_i^N (x_i - \bar{x})(y_i - \bar{y})}{(\sum (x_i - \bar{x})^2)^{\frac{1}{2}} (\sum (y_i - \bar{y})^2)^{\frac{1}{2}}} \quad (15)$$

Our work has used cross-correlation based algorithms to classify tissue, liquids and synthetic fluoropolymer and we has done so with great success [21]. A cross-correlation based classification algorithm we have used for breast tissue types, first, discriminates tissues based on the form factor peak location on the momentum transfer axis. Revisiting the breast tissue form factors by Kidane et al. in Figure 14, it is seen that normal and adipose tissue can be differentiated between cancer and fibroglandular tissues at a momentum transfer value of approximately 1.3 nm^{-1} .

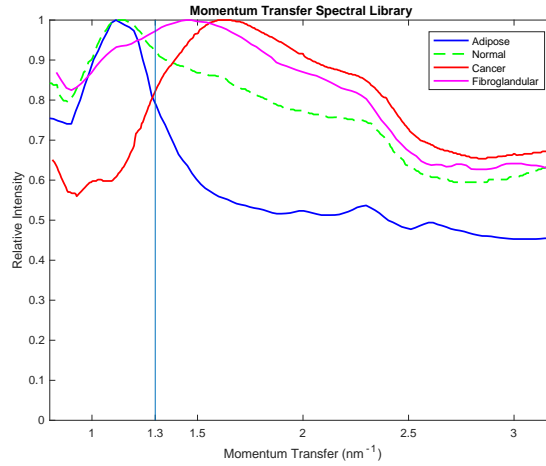


Figure 14: Diffraction patterns in XRD breast tissues [Kidane *et al.*, 1999] [8]

After differentiating between these two groups of tissues, a weighted cross-correlation metric is used to identify specific tissue types. Here, form factor

measurements are segmented into two distinct measurements based on how similar or different the remaining two tissues are within each segment region.

Considering adipose and normal breast tissue form factors, the two show significant similarities below 1.4 nm^{-1} . Above 1.4 nm^{-1} the two signals show stronger deviation from each other. Therefore, cross-correlation of the two tissues' form factors above 1.4 nm^{-1} is considered to provide more of a distinction between the two and is weighted more heavily than the signal measured below 1.4 nm^{-1} . Fibroglandular tissue and cancerous tissues show more similarities above 1.6 nm^{-1} than below 1.6 nm^{-1} . For this reason, segmented regions above 1.6 nm^{-1} are weighted less than the segmented regions below it.

Finally, if the measured form factor does not correlate with any tissue with significant confidence (above 0.7), no classification is made and it is assumed that there is not enough signal present to make a clear determination about which tissue type is present, or the signal has been contaminated appreciably, and a decision cannot be made. Figure 15 shows the algorithm in its entirety.

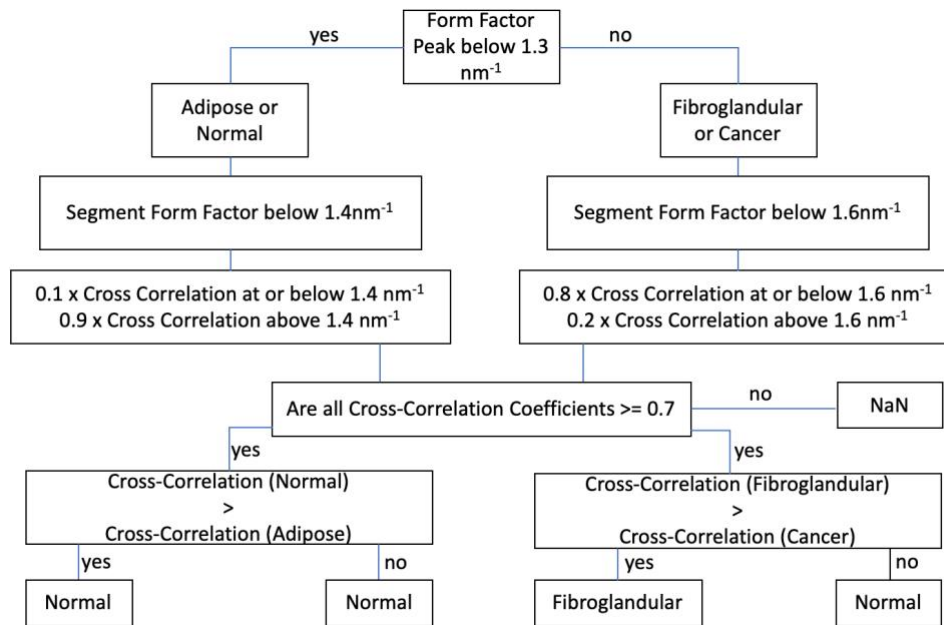


Figure 15: Breast tissue classification algorithm.

3. MC-GPU Modification and Comparison

MC-GPU's has been previously validated for use as an EDXRD simulation tool but has been modified to emulate an ADXRD/EDXRD hybrid XRD detection scheme, similar to the one used in our lab [33]. This ADXRD/EDXRD combination signal detection method will be compared against experimental measurements performed on a testbed prototype being developed in our group.

Off the shelf, MC-GPU can output radiographic images that represent energy-integrated X-rays tallied in units of eV/cm². Due to the need for energy-sensitive X-ray detection in the proposed technology, the MC-GPU source code was modified to model a one-dimensional, energy-binning, energy-sensitive detector. This modification to the MC-GPU kernel, among two others used in this work, is detailed in Appendix A.

The motivation for this modification was to emulate the MultiX ME100 X-ray-sensitive detector (MultiX, France). A detector that facilitates ADXRD/EDXRD hybrid measurements using a polychromatic source, the MultiX ME100 is a spectrometric detector that combines semiconductor crystal (CdTe/CZT) detection with unique ASICs and high-speed front-end electronics capable of precisely measuring the energy of each incident X-ray photon [37]. This detector is a one-dimensional detector.

To accomplish this emulation, changes were made to the GPU kernel. The main MC-GPU file stores detector pixel information as a global variable, an image matrix. The image matrix is generated after the user specifies the number of rows and columns desired in the imaging plane. During simulation, the GPU kernel identifies the X-rays incident on the detector plane; our modification directed the GPU (or CPU, if compiled to run on CPU) to consider only the X-rays incident on the center row of the detector plane and store the information about the detected X-ray (energy and lateral detection location) in the image matrix according to the energy of the incident X-ray.

Because the MultiX ME100 only considers X-rays from 20 to 160keV and typically sorted into 128 energy bins, only incident X-rays with energies above 20.48 keV were considered and binned into 1.1keV energy bins. This emulating detector, although modeled after the ME100, is a theoretical one-dimensional detector with 100% detective quantum efficiency and perfect energy-sensitivity. Figure 16 depicts how the modified

MC-GPU kernel detects X-rays along the center row (shown via horizontal red line) and how it stores the X-ray information in the image matrices.

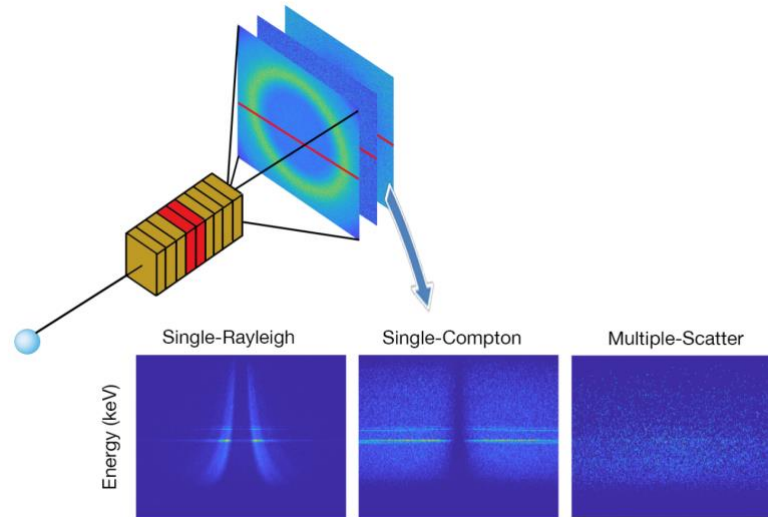


Figure 16: MC-GPU output with modified kernel designed to emulate a one-dimensional, energy-sensitive detector. X-rays incident upon the one-dimensional detector array are binned into energies and displayed on the row according to energy bin on the output images. Transmission X-ray not shown.

3.1 Experimental ADXRD/EDXRD Measurements using MultiX ME100

An experimental XRD imaging system was used to scan samples of adipose, water and bone. The imaging system is comprised of the following components: a tungsten anode X-ray tube source, collimators, object stand, beam stop, a coded aperture, and three MULTIX ME100 detectors. Figure 17 shows this experimental setup.

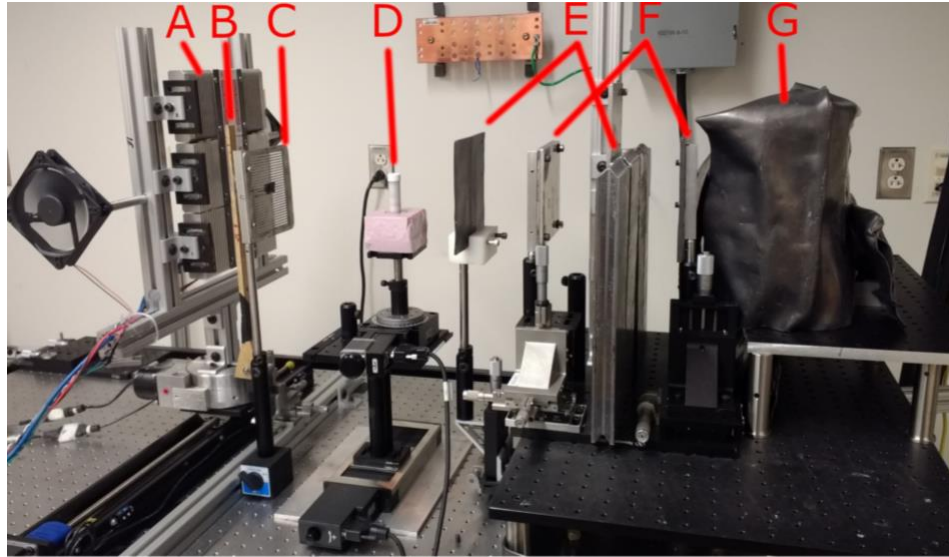


Figure 17: XRD measurement system utilizing. A) Three vertically mounted MULTIX ME100 detectors; B) Beam Block; C) Coded Aperture; D) Scanned object; E-F) Collimators; G) X-Ray Tube.

The MultiX ME100 detectors are placed on a horizontally translating motor. During image acquisition, the 1-Dimensional set of detectors translates horizontally with a total travel length of 14.78 cm; this feature, with a continuously applied X-ray source, uses a step-and-detect method to produce a two-dimensional image with a one-dimensional detector array. Specimens were scanned using this translational feature; however, only the centerline data acquisition was used during data processing. This is because MC-GPU orients the source completely perpendicular and centered to the detector and the capability to mimic this detector motion is not yet achieved.

XRD measurements were run on the experimental XRD imaging system by placing specimens that were approximately 1.3 cm in diameter, 28cm from the detector along the primary beam path. The tungsten X-ray source was set at 160 kV and 11.25 mA

applied for 3 seconds. The centerline measurement was taken and 180 pixels centered on the origin of the pencil beam were considered for reconstruction.

3.2 Simulated ADXRD/EDXRD Measurements using Modified MC-GPU Kernel

MC-GPU material files were created for each material to be scanned: adipose, water and bone. The first step in this process is to define the material composition of each object to be modeled. The compounds used in the experiment and the references to the compound mixture weighting used to create the MC-GPU material files are found in Table 2.

Table 2: Scanned Material Modeling and Compound Mixture References

Material Scanned	Modeled Material	Reference
Water	Water	NIST
Crisco	Adipose Tissue	ICRP report 110 Appendix B.2 [38]
Bone	Bone Compact	ICRU report 46 Appendix A [23]

Because molecular interference is considered in this work, we then measured the momentum transfer function that would be used in Monte Carlo simulation. Momentum transfer functions were measured using a Bruker D2 Phaser ADXRD X-ray diffractometer. This system uses a Ni filtered Cu K α ($\lambda=1.5406 \text{ \AA}$) ceramic sealed tube and was used carry out the structural characterization with measurement angles between 8.70166° to 119.039° which, correspond to momentum transfer values of 0.49 and 5.60 nm^{-1} , respectively. These measurements were made at $.0234 \text{ nm}^{-1}$ intervals.

Figure 18 shows the Bruker D2 Phaser setup used to measure diffraction of adipose tissue.

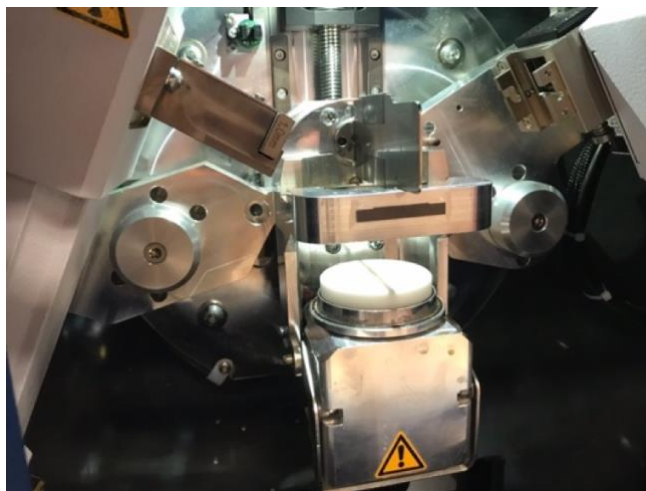


Figure 18: Coherent scattering measurement setup for adipose tissue on the Bruker D2 Phaser

The XRD measurements were then smoothed using a robust quadratic regression over a 0.02 \AA^{-1} window. The data was interpolated and the analytical approximation of F_{IAA}^2 for each compound was calculated for normalization, a process detailed in Section 1 of this work. The molecular interference function was calculated for each of the materials to be modeled and was used to complete the material file creation process. Figure 19 shows the measured, smoothed and interpolated momentum transfer function of adipose (left) and the normalized form factor data at high momentum transfer values (right).

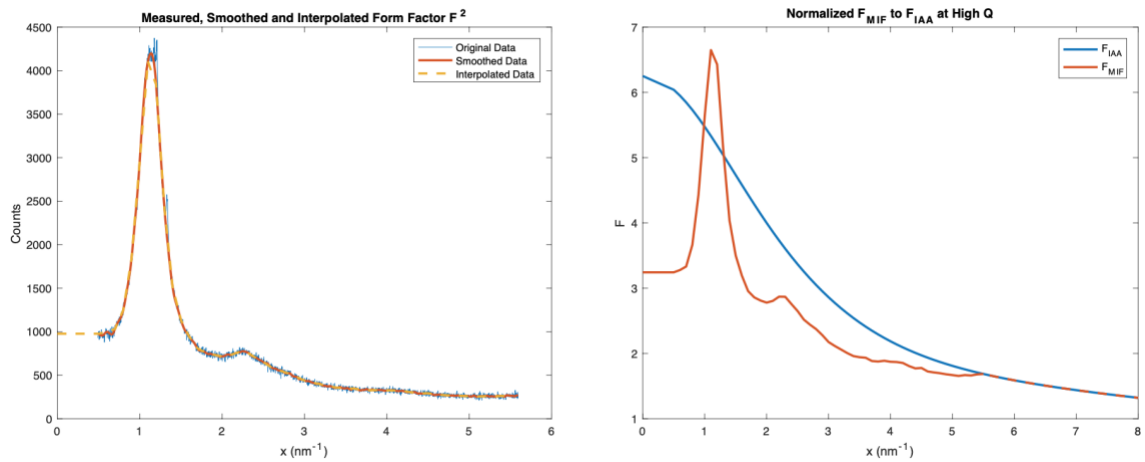


Figure 19: An example of smoothing, interpolating a measured momentum transfer function (left) and normalizing the square root of the smoothed and interpolated momentum transfer function of Adipose (right).

For each experimental measurement, identical experiments were recreated in MC-GPU using a modified MultiX ME100 emulating executable file, and the experimental measurements were compared to the simulated image acquisition:

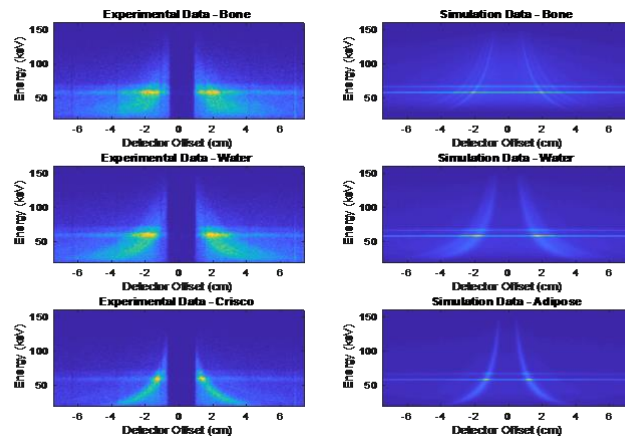


Figure 20: Experimental MC-GPU comparison – Experimental MultiX ME 100 measurements (left) and MC-GPU simulated MultiX ME 100 measurements (right) for Bone, Water, and Adipose.

The reconstructed images in Figure 20 are windowed and leveled to show a consistent dynamic range which normalizes the inherent discrepancies between number of simulated and experimentally incident X-rays.

Some similarities are noted. Both show significant intensity around the tungsten $k_{\alpha 1}$ edge at 59.31 keV and the tungsten k-edge at 69.52 keV. However, experimental measurements show higher signal in almost all other energy bins, relative to the peak intensity observed in the peak energy bin. Also, in experimental measurements, non-transmission X-rays were removed by a physical beam block whereas transmission X-rays are simply not displayed from the Monte Carlo simulation. This leads to the absence of transmission signal in the experimental reconstruction across all energies near the centerline pixels.

A reconstruction of the momentum transfer function, $F^2(q)$, for both image sets shows alignment in relative magnitude to the peak and width along the momentum transfer axis:

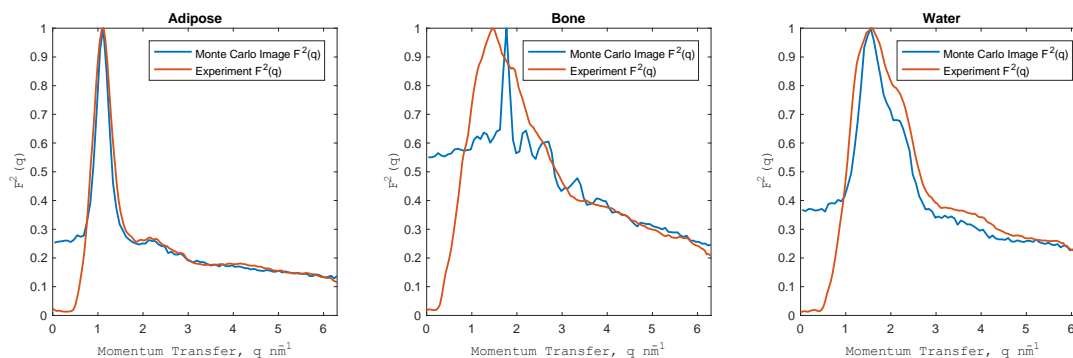


Figure 21: Reconstruction of the momentum transfer function for Adipose, Bone and Water based of experimentally measured XRD data and MC-GPU Monte Carlo simulation.

The percent deviation between the simulated and experimentally measured momentum transfer function values between 0.85 and 6 nm⁻¹ was calculated for each tissue and is shown in Figure 22.

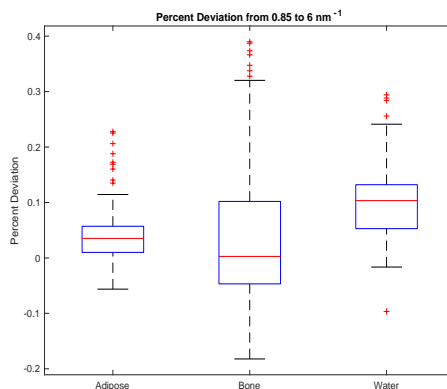


Figure 22: Percent deviation between simulated and experimentally measure momentum transfer values for MC-GPU comparison.

It should be noted that experimental measurements and simulation measurements of momentum transfer in bone show the largest deviation in percent

error. An assessment of the momentum transfer function measurement by the Bruker D2 Phaser, shown in Figure 23 shows a turbulent function with defined peaks and valleys.

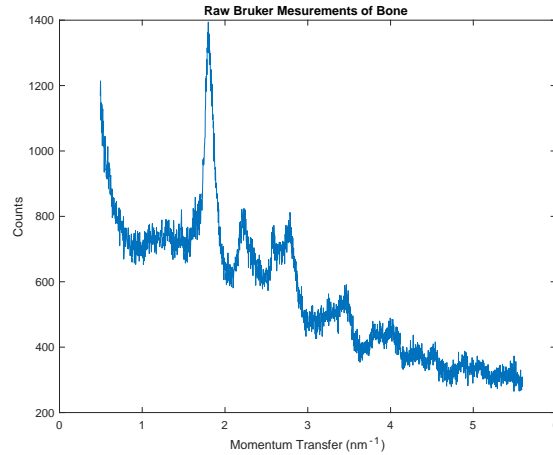


Figure 23: Raw Bruker D2 Phaser measurements made by an ADXRD scan of bone.

Because the MultiX ME100 energy-resolution is expected to follow Gaussian law, it should be expected the MultiX measurement show significant signal blurring, whereas the detector in MC-GPU is designed to have a theoretically perfect energy-resolution and can be used to more accurately reconstruct momentum transfer function with large variation. Regardless, comparison of experimentally measured signal with Monte Carlo shows great alignment and demonstrates suitability for this work.

4. Maximum Signal Energy Bin Characterization

During form factor reconstruction, we have typically observed diffraction via the energy channel with the strongest signal. A reconstruction of the form factor using only signal from the strongest signal energy bin has been satisfactory for our work, even when additional data via other channels is available. While it is true that a combination

ADXRD/EDXRD measurement system allows us to measure at different momentum transfer ranges and resolution with each energy bin, a careful consideration of the effects of specimen placement (ODD) and Energy, along with an understanding of the lateral pixel displacement of each pixel from the beam has historically allowed for sufficient reconstruction of the measured form factor for classification purposes, using only the strongest signal bin. The peak signal energy bin, for most spectrum kV settings, occur in the detector energy bin that corresponds to one of the characteristic peaks of the anode material.

Many of our sources are tungsten anode sources, often at energies above 100 kV. However, we also consider the use of a lower kV molybdenum anode spectra which, because of the increase in Rayleigh scattering and decrease in Compton scattering probabilities at low keV, should promote more coherent scatter interactions and reduce the effect of Compton scatter on our measurements.

The aim of the experiments in this section is to observe the effects of signal contamination from XRD measurements on common breast tissue types as well as cancerous breast tissue. Here we simulate XRD experiments on homogenous tissues using a variety of source spectra. The form factor measurements are made using the signal acquired in most prominent energy bin and the measured form factors are compared against ground truth provided by Kidane et al [8].

4.1 Source Spectrum

Both tungsten and molybdenum anode spectra were modeled in this section.

Tungsten anode spectra were generated using Spektr 3.0 and tungsten anode spectral model using interpolating polynomials (TASMIP) (Boone et. Al) [39] [40]. As Spektr 3.0 does not model molybdenum anode spectra, molybdenum anode spectral model using interpolating polynomials (MASMIP) was adapted by importing the MASMIP coefficients presented by Boone et al. Spectrum filtration was implemented in both approaches with Spektr 3.0, which samples tables of linear attenuation coefficients for different materials at precise energies and applies Beer's Law appropriately. Figure 24 shows the spectra considered in this work.

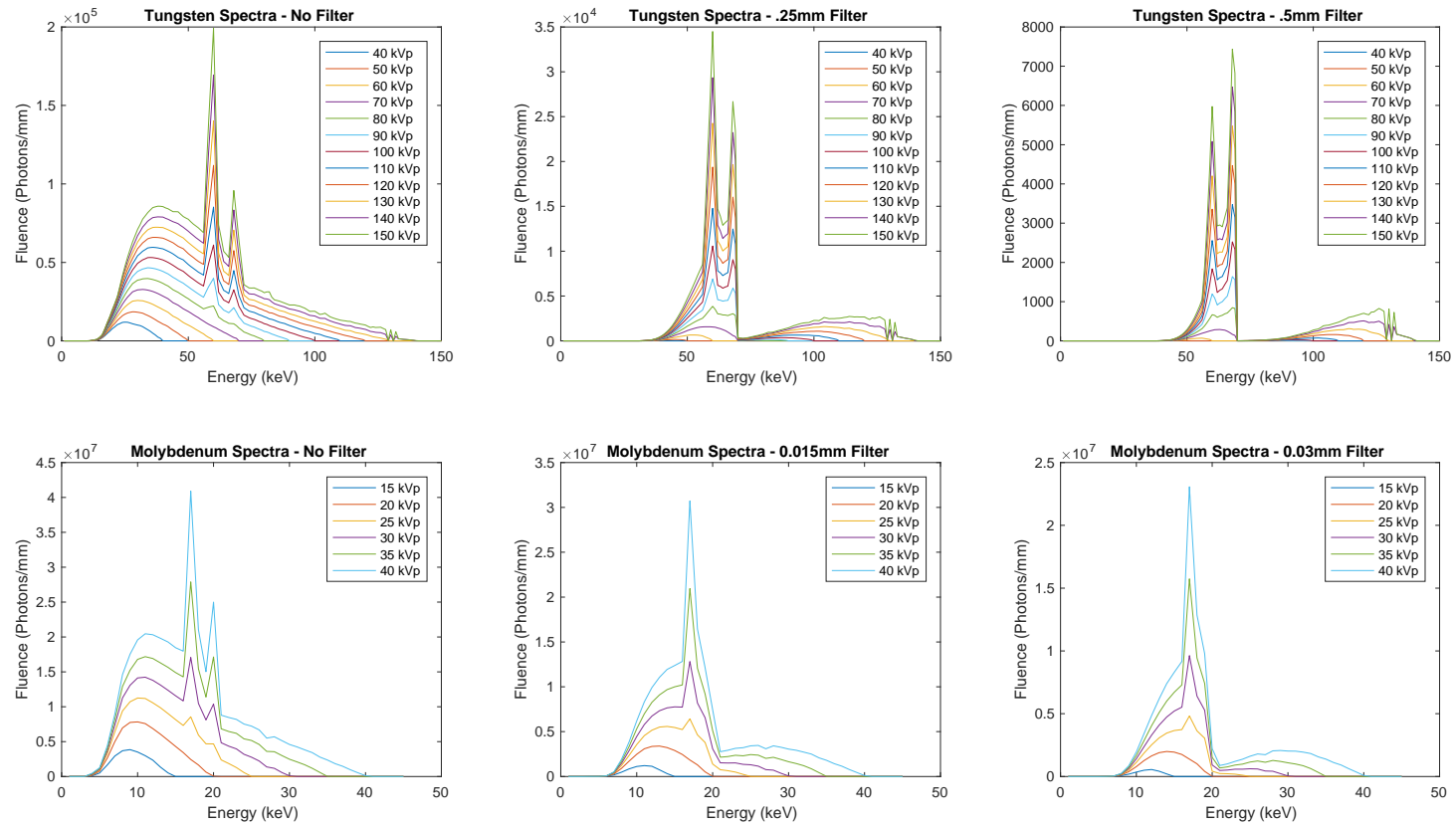


Figure 24: Tungsten (top) and Molybdenum (bottom) spectra modeled with increasing filter thickness from left to right. Tungsten anode spectra are filtered with 0, 0.25 and 0.5 mm of tungsten and molybdenum anode spectra are filtered with 0, 0.015 and 0.3 mm of molybdenum.

4.2 Rapid Simulation via MC-GPU Wrapper

MC-GPU simulations were run on an NVIDIA GPU cluster via Secure Shell (SSH), which is a cryptographic network protocol for operating network services securely. In order to run multiple experiments in rapid succession, a MATLAB wrapper was designed to generate MC-GPU input files and utilize SSH to not only transfer the appropriate spectra, material, voxel geometry and input files to a specified directory but also to run a specified command line script with variable command line arguments.

The MC-GPU wrapper is used to design and create the voxelized material phantom to be used in the MC-GPU simulation. A three-dimensional matrix in MATLAB was used as a 3D representation of the voxelized environment to be used in MC-GPU. The 3D matrix contains unique integers, each used to represent a material meant to occupy that particular voxel. The three-dimensional matrix along with a user-specified voxel size are programmatically translated into a one-dimensional voxel geometry file, which can be read by MC-GPU.

The MATLAB wrapper prepares the MC-GPU input file to be used as a command line argument. A random ten-digit seed, source spectral files, material files, voxel geometry files, object dimensions and simulation parameters are amended into the input file and copied to the GPU cluster via Secure Copy. The desired modified MC-GPU executable is called and the output image is returned to the local computer via Secure Copy for analysis.

Alternatively, if more than one MC-GPU simulation is run, the MATLAB wrapper has the functionality to send a MATLAB script to the GPU cluster. The MATLAB script is designed to open all output data files in a remote folder, parse through the output file and reorient each output image into the intended dimension of the image. The output images are then converted to a sparse matrix and saved to a single MATLAB structure array.

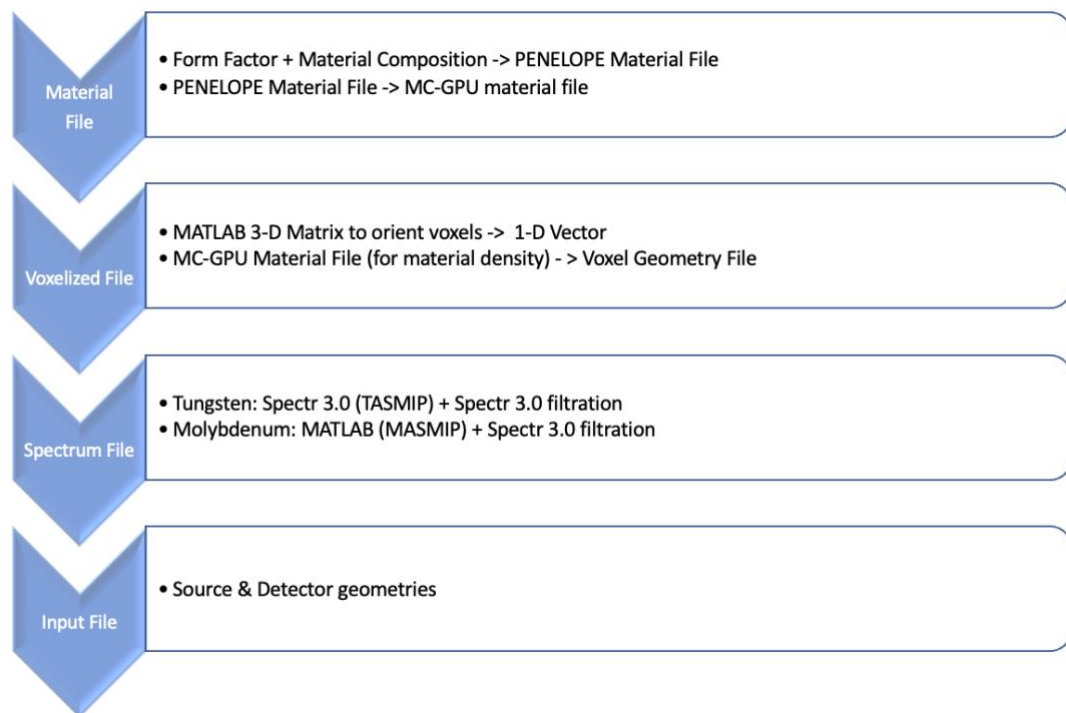


Figure 25: Flow Map of Simulation Setup. All files are sent to a common folder via SSH and the executable MC-GPU file is run with the Input File as the command line argument.

4.3 Simulation Set Up

Homogenous phantoms of adipose, normal, fibroglandular and cancerous tissues were modeled in this experiment. Material files were created using the material recipes

found in ICRU 46 [23] and the molecular interference functions, $s(x)$, were derived from the scatter form factors presented by Kidane et al [8]. MC-GPU material files were produced by considering energies from 1 keV to 161 keV sampled 25005 times, the maximum number of samples allowed in MC-GPU.

For each experiment, a 10x10x1cm homogenous specimen was placed between an ideal pencil beam source with the source oriented to the tissue's center. A MultiX emulating modified MC-GPU executable was called with the command line arguments detailed in Table 3 for tungsten spectra.

Table 3: MC-GPU Tungsten Anode Simulation Command Line Arguments

Histories	1.00E+10
Seed	Random 10-digit seed
Source Position	Object X,Y Center
Pixels (X)	400
Energy Bins (Z)	128
Image Width (cm)	20
Pixel Width (cm)	.05
Pixel Height (cm)	0.155
Object to Detector Distance (from object center)	28

Because molybdenum spectra are in the 15-40 kV range, the MultiX-emulating MC-GPU modification is not suitable for molybdenum anode simulations. The MultiX-emulating modification discards 20.48keV X-rays incident on the detector plane. A modified MC-GPU executable was created that retained all X-ray signal at the detector and binned into 1.0 keV bins. Details regarding this modified MC-GPU executable can be found in Appendix A. For molybdenum anode spectra, this modified version of MC-

GPU was used and the executable was called with the command line arguments listed in Table 4.

Table 4: MC-GPU Molybdenum Anode Simulation Command Line Arguments

Histories	1.00E+10
Seed	Random 10-digit seed
Source Position	Object X,Y Center
Pixels (X)	400
Energy Bins (Z)	40
Image Width (cm)	20
Pixel Width (cm)	.05
Pixel Height (cm)	0.155
Object to Detector Distance (from object center)	10

A total of 144 tungsten anode simulations were run: 12 different kV settings, 3 filter combinations and 4 homogenous tissues were scanned. A total of 72 molybdenum anode simulations were run: 6 different kV settings, 3 filter combinations, 4 homogenous tissues in total.

The measured form factor was reconstructed for the strongest signal energy bin. The form factor was run through the classification algorithm detailed in this work and the weighted-correlation results, energy bin used, and CNR were recorded.

4.4 Results

Figure 26 details the figures of merit as a function of spectrum kV setting for tungsten anode spectra.

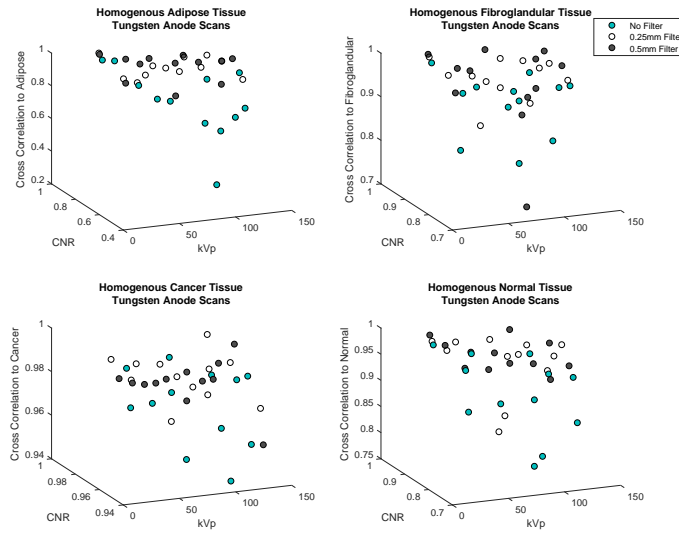


Figure 26: Homogenous tissue scans detailing cross-correlation to ground truth and CNR for the highest-signal energy bin using tungsten anode spectra.

Figure 27 details the figures of merit as a function of spectrum kV setting for molybdenum anode spectra.

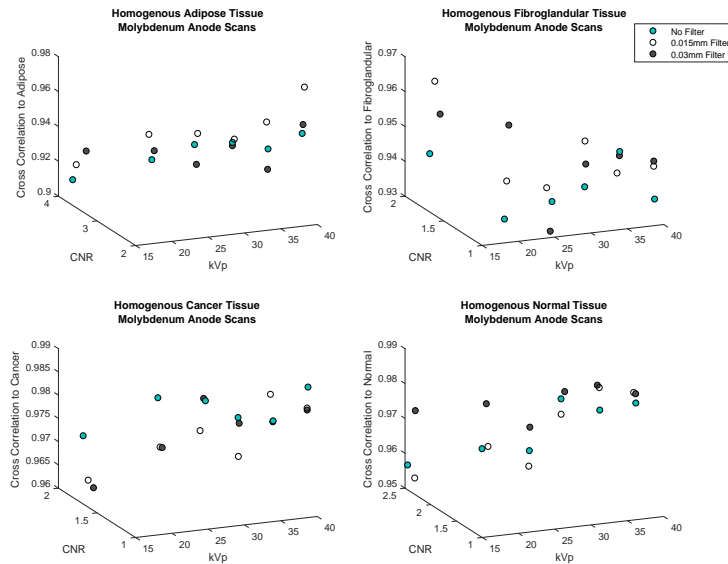


Figure 27: Homogenous tissue scans detailing cross-correlation to ground truth and CNR for the highest-signal energy bin using molybdenum anode spectra.

4.5 Discussion

The top left datasets shown in Figure 26 and Figure 27 are from a homogenous XRD scan of adipose tissue. These results are from 36 simulations in which adipose tissue was scanned with various tungsten anode spectra (Figure 26) and 18 simulations in which adipose tissue was scanned with various molybdenum anode spectra. The measured form factors were compared to the reference form factor for adipose tissue [Kidane *et al.*, 1999] [8] via weighted cross-correlation algorithm, and mostly showed cross-correlation scores in excess of 0.9, with few exceptions. Additional datasets for XRD simulations of fibroglandular, cancerous and normal breast tissues are presented in the remaining scatter plots.

The weighted cross-correlation classification algorithm was the most consistent across molybdenum anode spectra and showed increasing trends with increase in kV for these spectra, unlike the tungsten anode spectra which showed little correlation between cross-correlation score and kV. The cross-correlation measurements when compared to the ground truth across all tissues scanned averaged .903 with a standard deviation of 0.088 for measurements made via the tungsten anode spectra, whereas the molybdenum spectra made cross-correlation measurements against the ground truth with an average result of 0.962 and a standard deviation of 0.028.

The unfiltered molybdenum anode 15 kV spectrum showed the highest CNR measurements. A reconstruction of the form factor for simulations showing the highest CNR values are shown in Figure 28.

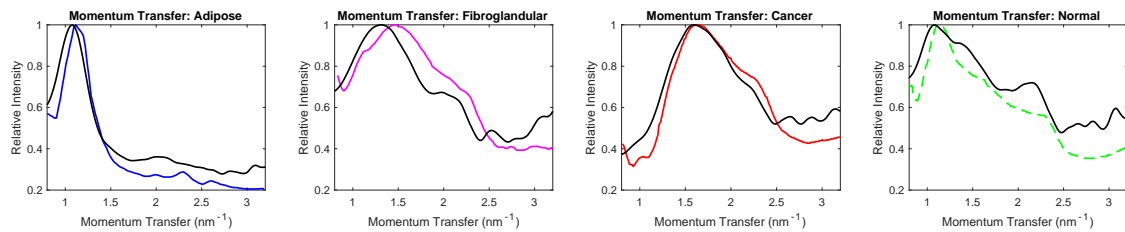


Figure 28: Simulated XRD scans of homogenous tissues. High CNR outcomes were from the 15 kV molybdenum anode spectra.

The lowest CNR measurements came from tungsten anode spectra. The 130 kV 0.5mm filtered tungsten anode spectrum showed the lowest CNR measurement of 0.68 for an XRD scan of adipose tissue; the 140 kV spectrum with .5mm tungsten filter showed an XRD scan of fibroglandular tissue and cancerous tissues with CNR measurements of 0.28 and 0.11, respectively; the 150 kV spectrum with .5mm tungsten filter showed an XRD scan of normal breast tissue with a CNR measurement of 0.32. These suggest low CNR measurements are found using high kV spectra. Figure 29 shows the momentum transfer measurements associated with low CNR measurements.

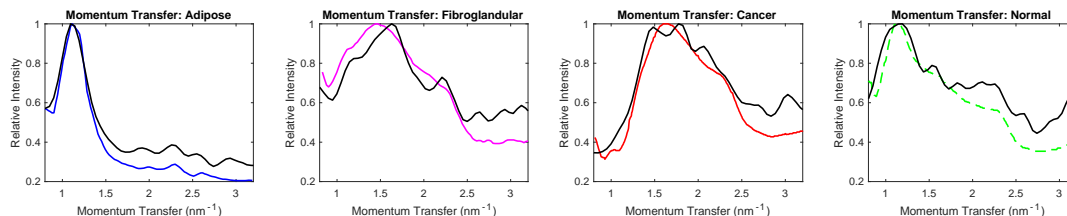


Figure 29: Subset of momentum transfer functions reconstructed from low CNR measurements.

XRD scans with low CNR showed relatively poor weighted-cross correlation when compared with their constituent tissue reference form factors. Adipose, fibroglandular and normal tissues showed weighted cross-correlation values of 0.82, 0.74 and 0.73. Although quantitatively noisy, the low CNR value extracted from the cancer scan showed a weighted cross-correlation value of 0.92.

5. Cancer Detectability Among Spectra

In previous work, Nacouzi et al. showed favorable results when comparing XRD measurement and subsequent classification of cancerous tissue against a pathological indication of cancer on a 5 mm thick excised breast tissue sample following BCS lumpectomy [22]. It should be noted that a 5 mm sample can potentially contain more than one tissue type and that it is very possible that cancerous spindle (thin outgrowths) invade local regions of otherwise healthy tissue that are virtually undetectable due to their size. In this section, we explore the figures of merit of our molybdenum and tungsten anode spectra with specimens that are non-homogenous and have varying amounts of cancer embedded in them.

It is expected that these experiments prove the difficulty in distinguishing cancer from an XRD breast tissue measurement of mixed tissue with no axial resolution.

Consider the momentum transfer function of an adipose/cancer mixed tissue. The peak position of the homogenous adipose tissue is 1.11 nm^{-1} whereas the peak position of cancerous tissue is 1.6 nm^{-1} .

Increasing concentrations of cancerous tissue take an otherwise undisturbed adipose form factor and broaden it into regions of higher momentum transfer as the signals from each tissue type are aggregated. As the cancerous tissue increases in concentration, the broadening becomes more severe; the peak location of the form factor measurement shifts to a higher q , and the resulting form factor closely resembles the form factor for cancer, and could very well be mistaken for fibroglandular tissue. Figure 30 shows this effect (left) and shows a comparison of a 7/9 cancer tissue scan with fibroglandular and cancerous tissues (right) - note that the weighted cross-correlation of this measurement compared better with reference fibroglandular tissue (0.974) than with cancer (0.866).

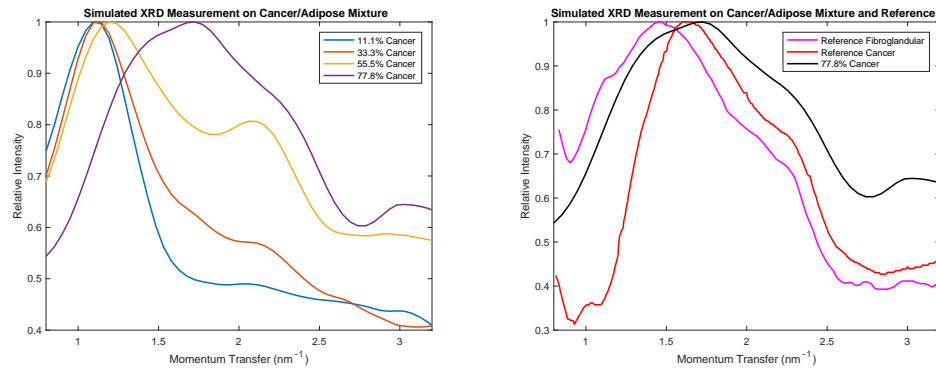


Figure 30: Reconstructed momentum transfer function from simulated XRD measurement of adipose tissue with embedded cancer at various volumetric fractions (left). Reconstructed momentum transfer function of 7/9 cancer/adipose shown with reference momentum transfer functions for fibroglandular and cancer tissues [Kidane *et al.*, 1999] [8].

5.1 Embedded Cancer Simulation Details

To simulate cancerous tissue embedded in normal breast tissue, the specimens will be modeled in layers and cancerous tissue are placed in the middle layers of the voxelized phantom. Figure 31 shows our voxel orientation. Here, non-cancerous tissues of adipose, normal or fibroglandular breast tissues occupy the posterior and anterior voxel spaces with a cancerous voxel(s) embedded in the middle.

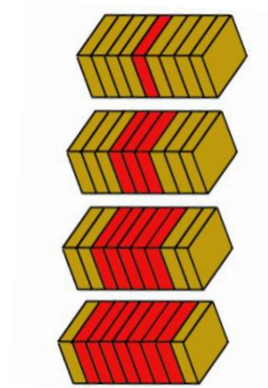


Figure 31: Voxel combinations used in this experiment. Red voxels are malignant breast tissue and the anterior and posterior gold voxels shown are adipose, normal or fibroglandular tissue.

The simulation was engineered to programmatically build each combination of voxel and tissue type for a total of 12 different combinations of tissues to be scanned. To accomplish this, a 1x1x9 voxelized phantom was created with voxel size dimensions of 10x10x1cm. The x and y components were made larger than the z-component to prevent diffracted X-rays from interaction with the outer edges of the phantom before exiting the posterior of the phantom.

The material files, spectra, voxel geometry files and command line arguments detailed in Section 4 were used in these simulation sets.

5.2 Results

432 tungsten anode XRD simulations were run from a combination of twelve different kV settings, three filter combinations, four tissue combinations. Three different tissues types were simulated to be embedded with cancer. A total of 101 positive

correlations to cancer were noted. Table 5 shows the positive findings of cancer noted in this work from tungsten anode spectra.

216 molybdenum anode XRD simulations were run with six different kV settings, three filter combinations, four tissue combinations and of which three different tissues were simulated to be embedded with cancer. A total of 31 positive correlations to cancer were noted. Table 7 shows the positive findings of cancer noted in this work from molybdenum anode spectra.

Table 5 and Table 6 should be viewed together. Both tables, summarize the outcomes of the cross-correlation algorithm with cancerous voxels embedded between common breast tissue types using tungsten anode spectra. Data listed in the “Adipose” column in Table 5 indicates that when 1 cancerous voxel exists in between eight other adipose voxels, 0 of 36 tungsten anode spectra would classify the form factor as cancer. However, when seven cancerous voxels exist in between two adipose voxels, 7 of the 36 tungsten anode spectra classified the diffraction pattern as cancer. The CNR observed during these 7 simulations ranged between 0.3 and 1.3; a boxplot of the CNR from the energy bins used to construct the data is provided for these seven spectra. When the non-cancerous voxels were fibroglandular tissue, Table 5 indicates that more tungsten anode spectra could positively detect cancerous cells based on cross-correlation. These results are mirrored in Table 7 where molybdenum anode spectra positively identified cancerous voxels.

Table 5: Embedded cancer voxels within tissues – XRD measurements made via tungsten anode spectra classified as cancerous tissue (True Positives).

Cancer	Adipose	CNR	Fibro	CNR	Normal	CNR
1/9 Cancer	0/36		7/36		0/36	
3/9 Cancer	0/36		8/36		1/36	
5/9 Cancer	0/36		24/36		3/36	
7/9 Cancer	7/36		30/36		23/36	

Table 6: Embedded cancer voxels within tissues – XRD measurements made via tungsten anode spectra classified as non-cancerous tissue (False Negatives).

Cancer	Adipose	CNR	Fibro	CNR	Normal	CNR
1/9 Cancer	36/36		29/36		36/36	
3/9 Cancer	36/36		28/36		35/36	
5/9 Cancer	36/36		12/36		33/36	
7/9 Cancer	29/36		6/36		13/36	

Table 7: Embedded cancer voxels within tissues – XRD measurements made via molybdenum anode spectra classified as cancerous (True Positives).

Cancer	Adipose	CNR	Fibro	CNR	Normal	CNR
1/9 Cancer	0/18		0/18		0/18	
3/9 Cancer	0/18		4/18		0/18	
5/9 Cancer	0/18		9/18		0/18	
7/9 Cancer	0/18		13/18		13/18	

Table 8: Embedded cancer voxels within tissues – XRD measurements made via molybdenum anode spectra classified as non-cancerous tissue (False Negatives).

Cancer	Adipose	CNR	Fibro	CNR	Normal	CNR
1/9 Cancer	18/18		18/18		18/18	
3/9 Cancer	18/18		14/18		18/18	
5/9 Cancer	18/18		9/18		18/18	
7/9 Cancer	18/18		5/18		5/18	

5.3 Results

The results of this study, generally speaking, confirm some of the suggestions we explored in Section 5.1 and illustrated in Figure 30. Cancerous lesions embedded in tissues are more likely to be detected if they are imbedded in normal breast tissue types with peak location at relatively high momentum transfer values, the most likely is fibroglandular, then, normal breast tissue.

Fatty breasts generally diffract coherent scatter X-rays at lower momentum transfer values; at a fixed energy, diffraction occurs at lower angles. It isn't until the breast tissue are denser and less fatty (normal breast tissue is 50% adipose and 50% fibroglandular) that we can detect cancerous lesions with this current setup. As breasts become denser and resemble tissue closer to fibroglandular breast tissues, a cancer classification is more likely.

Regarding findings concerning X-ray spectra, we noted that XRD scans of cancerous tissues were generally classified as cancer when CNR was high. Table 6 and

Table 8 summarize XRD scans of these cancerous tissues that were not classified as cancer. A comparison of this data against data showing the accurate prediction of cancer (Table 5 and Table 8) shows that measurements that were likely not to be used to classify cancer were low CNR measurements. A study of the false negative cases in which cancerous tissue grew in adipose tissue shed light on an interesting observation that as cancerous voxels multiplied, the CNR inherently exhibited lower values.

6. Quantitative Comparison of Tungsten and Molybdenum Anode Spectra

Our final experiment is an interrogation of fibroglandular and cancerous tissues for optimization of the differentiation between these two groups. We have described in section 5 the difficulties involved in differentiation of fibroglandular and cancerous tissue types by XRD. As such, the purpose of this simulation ensemble is to assess how well these two tissue types may be differentiated under identical source spectrum XRD measurements.

6.1 Fibroglandular and Cancerous Tissue Simulation and Measures

To accomplish this task, simulations similar to those detailed in Section 4 were used; the only exception to this simulation setup was that normal and adipose tissue types were not included. As such, a total of 216 simulations were executed.

The form factor is reconstructed via the maximum signal energy bin and, for each tissue type, our classification algorithm was used to classify each tissue type.

Weighted-cross correlation values were compared and the tissues' weighted cross-correlation values between reference fibroglandular and reference cancer form factors were plotted. The goal is to assess which measurement conditions promote the maximum separation between the cross-correlation of the true tissue's form factor and its competitor.

6.2 Results and Discussion

The cross-correlation values are shown in Figure 32 and CNR is shown in Figure 33.

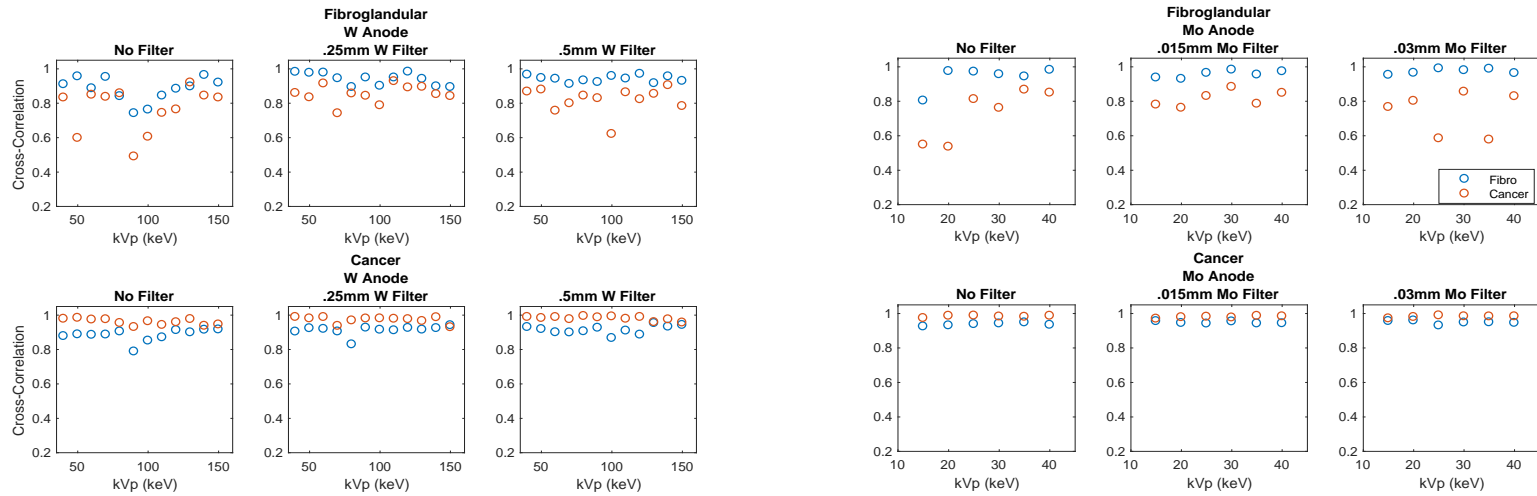


Figure 32: Cross-Correlation based on XRD scans of fibroglandular (top) and cancerous (bottom) tissues. Molybdenum spectra (right) show more differentiation between cancer and fibroglandular tissue when fibroglandular tissue is scanned.

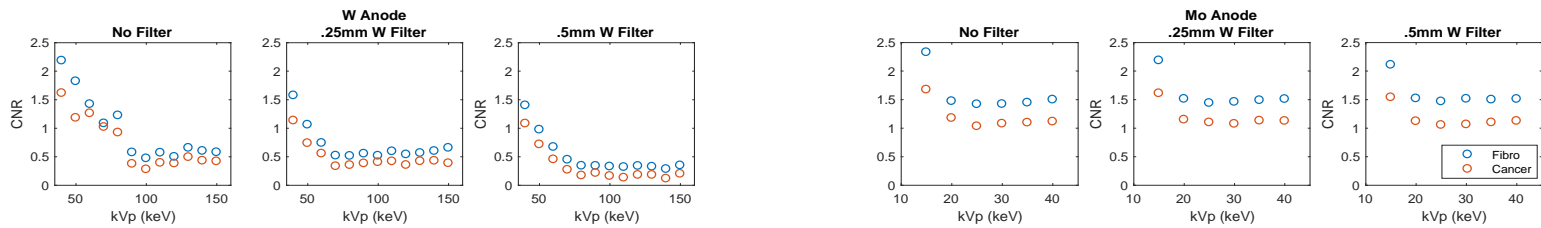


Figure 33: CNR observed during XRD scans of fibroglandular and cancer tissue scans.

The top row in Figure 32 shows homogenous scans of fibroglandular tissue. The measured form factor of the fibroglandular scans are compared to cancerous and fibroglandular reference form factors via the cross-correlation algorithm showing that in all except two instances, two non-filtered tungsten anode scans, the measured form factor classified as fibroglandular tissue. Molybdenum anode scans (right) showed larger differentiation between cross correlation between reference cancer and reference adipose.

The second row is similar, but this represents data from homogenous cancerous tissue XRD scans. In all instances, it became more difficult to accurately classify cancer because the cross correlation with reference fibroglandular tissue was almost as high as that with reference cancerous tissue.

The final and third row depicts the CNR measured during these simulations. The data shows a consistent trend: CNR is relatively high at low keV and drops as kV increases. In addition, molybdenum anode spectra show higher CNR values compared to tungsten anode spectra.

It is important to note that Figure 32 and Figure 33 maintain a consistent y-axis to aid the reader in making an assessment about how spectrum selection affects differentiability. Although XRD scans of cancerous tissue could be identified, the measured form factor correlated closely with fibroglandular tissue. However, we see differentiation between cross-correlation values of scanned fibroglandular tissue with

both types of spectra where molybdenum anode spectra show measured form factor cross correlations that are the most unique.

It was shown that form factor contrast is better observed through low energy systems and that, among the spectra studied, the contrast to noise decreased as tube potential increased with slight increases after the primary anode k-edge. With these figures of merit in mind, the 0.03mm molybdenum filtered molybdenum anode spectra consistently provided the most differentiation between fibroglandular tissue and cancerous tissue when fibroglandular tissue was scanned. The 25 and 35 kV molybdenum anode spectra with 0.03 mm of Mo filter showed the largest separation. When using these parameters, it is recommended to operate above the characteristic k-edged of molybdenum to increase the CNR of the highest signal energy bin.

7. Discussion and Conclusion

In the design and application of pathological imaging systems, we have often been asked how quickly can we make accurate measurements and how precise can they be. Much of our desire to use polychromatic sources in the XRD imaging scanners is motivated by the ubiquity of these source types in medical research and also the ability to obtain large amounts of information in a single XRD measurement with ADXRD/EDXRD combination imaging tools.

This work has investigated how well we can make XRD measurements with commonly used polychromatic source spectra and what types of tools can be used to

make this assessment. We have chosen MC-GPU as a Monte Carlo suite and modified it to provide the energy-sensitive information we need meet our needs and model systems being designed in our group.

This work made multiple measurements on hundreds of simulations; this was only made possible through the use the X-ray transport toolkit MC-GPU which provides a GPU framework to accelerate the simulations. Because of our reliance on a hybrid ADXRD/EDXRD measurement system made to increase the information density provided with each scan, our modifications to MC-GPU were warranted and compared against experimental measurements to evaluate accuracy and efficiency of the code.

We first defined figures of merit and a classification algorithm that mirrors one used previously in our own work. With these, we conducted 216 simulation XRD measurements over four homogenous tissues and made classifications of these tissues, as we would in practice. We found that most measurements were suitable enough to make accurate classifications of tissues; however, we also observed variability in classification metrics, especially among tungsten anode spectra measurements.

Our second batch of simulations, 648 in all, varied the extent of cancerous tissue in a fixed tissue specimen. As cancerous tissues grew in a fixed adipose specimen, it was difficult to provide an indication of cancer compared to tissue types that inherently were denser and showed momentum transfer spectra in the higher- q range. Moreover, we

noted that CNR played a factor in whether or not a given spectrum was able to show an indication of cancer.

Finally, because differentiation between fibroglandular and cancerous tissues via these simulated measurements was so difficult, simulations were made to observe the classification measurements on homogenous fibroglandular and homogenous cancerous tissues. The idea behind this experiment was to observe the effects of CNR and kV on these measurements and to determine if CNR allowed for more differentiation between the two measurements. No distinct correlation was identified between these figures of merit but molybdenum spectra were observed to produce cross-correlation measurements that were more separable when adipose tissue was scanned.

We have made a concentrated effort to pursue a combination of ADXRD/EDXDR imaging along with different X-ray sources that provide desirable low-energy spectra. This work identifies the quantitative advantage of using a lower energy source over the existing radiography sources employed in our previous work. Lower tube potential increases the probability of coherent scatter and helps to minimize noise contribution from Compton scattering. The work also demonstrates the utility of Monte Carlo in optimizing coherent scatter imaging systems and can be used to provide insightful information regarding the design of coherent scatter imaging systems for material classification in breast tissue.

8. Improvements

This work relied heavily on a modified MC-GPU package that facilitated the simulation of a one-dimensional energy-resolving detector for uses in XRD imaging. We have explored X-ray diffraction onto a plane and it should be apparent that the use of a 1-dimensional detector in XRD imaging is not ideal. The sheer amount of signal lost by reducing our detector plan into a single row is massive and would be made even worse had we incorporated a coded aperture.

In assessing signal degradation and detectability, this work also relied on the cross-correlation metric, which is beneficial in many applications. However, because much of XRD imaging involves the maintenance and creation of momentum transfer spectral libraries and hinges on accurate classification methods, the use of a machine-learning-based classification tool that could better differentiate between the features of cancerous and fibroglandular momentum transfer functions could be more suited than a conventional mathematical approach.

Finally, depth resolution (in the z-dimension) was not considered in this work. Depth resolution might not be necessary when probing small, millimeter sized samples; however, it may be valuable in probing larger objects, particularly as we consider prospective in vivo coherent scatter imaging.

10. Future Work

XRD systems have the potential to be revolutionary. If realized, our pathology XRD system has the potential to be used for ex-vivo pathology assessments that would, in real time, give a pathologist or surgeon insight into the molecular and atomic structures in the tissues being assessed. The current and future direction of our XRD medical imaging is the pursuit of in vivo XRD imaging systems that could provide a practitioner the precise location of a particular type of biological structure such as malignant tissue. The challenges in in-vivo XRD imaging will likely center around the thickness of the specimen to be scanned because of the competing effects of the photoelectric effect and coherent scattering at low kV energies.

Finally, our future work will most definitely be furthered by MC-GPU simulations of current and future XRD imaging systems. The simulations shown here are the first step in many and will continue to help evaluate and optimize XRD cancer imaging systems as we move forward.

Appendix A

A.1 MC-GPU Modification: MultiX ME100 Emulating Kernel

The standard MC-GPU software is used to produce two-dimensional images simulating a two-dimensional detector. A one-dimensional array detector simulation, used to simulate output images similar to the MultiX ME100 image output, was made by changing line 539 of the GPU version 1.3 kernel to

```
“if ((pixel_coord_z>-1)&&(pixel_coord_z<detector_data_SHARED->num_pixels.y)&&>(*energy > 20480)&&(pixel_coord_z == ceilf(detector_data_SHARED->num_pixels.y/2)))”;
```

adding the following before line 545:

```
int energyBin = (((*energy - 20480)/1100) + 1);”;
```

and changing line 548 to

```
“(pixel_coord_x + (detector_data_SHARED->num_pixels.y - energyBin)*(detector_data_SHARED->num_pixels.x)) , // Offset to the corresponding x pixel and energy bin”.
```

Identical changes were made between lines 580 and 586 of the kernel file.

A.2 MC-GPU Modification: 1 keV Energy Bin One-Dimensional Detector

MC-GPU was modified to model a one-dimensional detector with 1keV energy-resolution by changing line 539 of the GPU version 1.3 kernel to

```
“if ((pixel_coord_z>1)&&(pixel_coord_z<detector_data_SHARED-  
>num_pixels.y)&&(pixel_coord_z == ceilf(detector_data_SHARED-  
>num_pixels.y/2)))”;
```

adding the following before line 545:

```
int energyBin = *energy /1000;”;
```

and changing line 548 to

```
“(pixel_coord_x + (detector_data_SHARED->num_pixels.y -  
energyBin)*(detector_data_SHARED->num_pixels.x)) , // Offset to the corresponding x  
pixel and energy bin”.
```

Identical changes were made between lines 580 and 586 of the kernel file.

A.3 MC-GPU Modification: Energy Discriminating Two-Dimensional Detector

A separate executable MC-GPU file was designed to allow the user to simulate a two-dimensional detector plane that allowed for upper and lower bounded energy discrimination.

The source code of MC-GPU version 1.3 kernel was changed: line 539 was changed to

```
“if ((pixel_coord_z>1)&&(pixel_coord_z<detector_data_SHARED-  
>num_pixels.y)&&(*energy >=LBound)&&(*energy <= UBound))”
```

and an identical edit to line 583 was made. Here, LBound and UBound were changed to the lower and upper bounds of the energy discrimination desired in unites of eV.

Simulations were compiled on NVIDIA enabled GPU with using the following shell script:

```
“nvcc -g -DUSING_CUDA MC-GPU_v1.3.cu -o MC-GPU_v1.3.x -O3 -  
use_fast_math -L/usr/lib -I. -I/usr/local/cuda/include -  
I/usr/local/cuda/samples/common/inc -I/usr/local/cuda/samples/shared/inc/ -lz --  
ptxas-options=-v”
```

which specifies an optimization level of 3 and uses the zlib compression library [41].

References

- [1] M. Mohebian, H. Marateb, M. Mansourian, M. Mañanas and F. Mokarian, A hybrid computer-aided-diagnosis system for prediction of breast cancer recurrence (HPBCR) using optimized ensemble learning., *Comput. Struct. Biotechnol. J.*, 2017, pp. 15, 75–85.
- [2] "National Cancer Institute. SEER stat fact sheets: female breast cancer: statistics at a glance.," [Online]. Available: <http://seer.cancer.gov/statfacts/html/breast.html>. [Accessed 18 02 2019].
- [3] "National Cancer Institute. Table 4.13: Cancer of the female breast (invasive): 5-year relative and period survival by race, diagnosis year, age and stage at diagnosis," [Online]. Available: http://seer.cancer.gov/csr/1975_2012/browse_csr.php?sectionSEL=4&pageSEL=sect_04_tab. [Accessed 18 2 2019].
- [4] L. Wang, Early Diagnosis of Breast Cancer, vol. 17, *Sensors*, 2017, p. 1572.
- [5] D. M. Cunha, A. Tomal and M. E. Poletti, "Optimization of x-ray spectra in digital mammography through Monte Carlo simulations," *Physics in Medicine & Biology*, vol. 57, no. 7, 2012.
- [6] B. Ghammraoui and A. Badal, "Monte Carlo simulation of novel breast imaging modalities based on coherent x-ray scattering," *Phys. Med. Biol.*, vol. 59, p. 3501–3516, 2014.
- [7] M. N. Lakshmanan, J. A. Greenberg, E. Samei and A. J. Kapadia, "Accuracy assessment and characterization of x-ray coded aperture coherent scatter spectral imaging for breast cancer classification," *J. Med. Imag.*, vol. 4(1), pp. 013505-1 through 013505-8, 2017.
- [8] G. Kidane, R. Speller, G. Hanby and A. Hanby, "X-ray scatter signatures for normal and neoplastic breast tissues," *Phys. Med. Biol.*, vol. 44, p. 1791, 1999.
- [9] W. W. Bragg and Bragg, "The Reflection of X-rays by Crystals," *Proc. R. Soc. London A*, vol. 88, no. 605, p. 428–438, 1913.
- [10] J. A. Greenberg, *X-Ray Diffraction Imaging Technology and Applications*, Duke University, NC, USA: Taylor and Francis Group, 2018.

- [11] M. E. Poletti, "X-ray scattering from human breast tissues and breast-equivalent materials," *Phys. Med. Biol.*, vol. 47, no. 1, 2002.
- [12] F. Salvat, J. M. Fernández-Varea and J. Sempau, PENELOPE-2006: A Code System for Monte Carlo Simulation of Electron and Photon Transport, Barcelona, Spain: OECD, 2006.
- [13] J. Hubbell, "Atomic form factors, incoherent scattering functions, and photon scattering cross sections," *Journal of Physical and Chemical Reference Data*, vol. 6, no. 615, 1977.
- [14] J. Baro, M. Roteta, J. F. Fernández-Varea and F. Salvat, "ANALYTICAL CROSS SECTIONS FOR MONTE CARLO SIMULATION OF PHOTON TRANSPORT," *Radiat. Phys. Chem.*, vol. 44, p. 531–552, 1994.
- [15] J. D. G. Harding, "Liquids identification with x-ray diffraction," in *Proc. SPIE: Optical Engineering + Applications*, San Diego, CA, 2007.
- [16] D. E. Verghese and K. Peplow, "Measured Molecular Coherent Scattering Form Factors of Animal Tissues, Plastics and Human Breast Tissue," *Physics in Medicine & Biology*, vol. 43, no. 9, pp. 2431-2452, 1998.
- [17] J. Links and J. Prince, *Medical Imaging Signals and Systems*, Upper Saddle River, NJ: Pearson Education, Inc, 2015.
- [18] H. v. Deurzen, "Predictors of Surgical Margin Following Breast-Conserving Surgery: A Large Population-Based Cohort Study," *Annals of surgical oncology*, vol. 23, pp. 627-633, 2016.
- [19] M. N. Lakshmanan, "Dissertation: X-ray Coherent Scatter Imaging for Intra-operative Margin Detection in Breast Conserving Surgeries," *Duke University*, 2015.
- [20] J. Spencer, "X-ray Diffraction Spectral Imaging for Breast Cancer Assessment," *Duke University - Master's Thesis*, 2017.
- [21] J. E.-C. Carter, "Optimization of a Coded Aperture Coherent Scatter Spectral Imaging System for Classification of Breast Cancer," *Thesis*, 2017.
- [22] D. Nacouzi, "Smarter Cancer Detection Through Neural Network Classification of High-Resolution Diffraction Tissue Scans," in *AAPM*, Nashville, TN, 2018.
- [23] I. Report, "Photon, Electron, Proton and Neutron Interaction Data for Body Tissues," vol. 46, 1992.

- [24] J. T. Bushberg, *The Essential Physics of Medical Imaging*, vol. 18, Philadelphia, PA: LIPPINCOTT WILLIAMS & WILKINS, 2012, pp. 457-468.
- [25] N. P. Laboratory, "X-ray absorption edges, characteristic X-ray lines and fluorescence yields," [Online]. Available: http://www.kayelaby.npl.co.uk/atomic_and_nuclear_physics/4_2/4_2_1.html. [Accessed 2 02 2019].
- [26] M. Salvadó, M. López, J. Morant and A. Calzado, "Monte Carlo calculation of radiation dose in CT examinations using phantom and patient tomographic models.," *Radiation Protection Dosimetry*, vol. 114, no. 1-3, p. 364-368, 2005.
- [27] C. SC, S. EM, S. AX and L. R. a. Y. H., "Dose-image quality study in digital chest radiography using Monte Carlo simulation," *Appl Radiat Isot*, vol. 66, no. 9, 2008.
- [28] S. Sharma, A. Kapadia, E. Abadi, W. Fu, W. P. Segars and E. Samei, "A rapid GPU-based Monte Carlo simulation tool for individualized dose estimations in CT," in *SPIE Medical Imaging*, Houston, TX, 2018.
- [29] A. Badano and A. Badal, "Accelerating Monte Carlo simulations of photon transport in a voxelized geometry using a massively parallel Graphics Processing Unit," *Med. Phys.*, vol. 36, pp. 4878-4880, 2009.
- [30] A. Badano and A. Badal, "Monte Carlo Simulation of X-Ray Imaging Using a - Graphics Processing Unit," *IEEE NSC-MIC, Conference Record*, vol. HP3-1, pp. 4081- 4084, 2009.
- [31] A. Badal, I. Kyrianiou, D. Sharma and A. Badano, "Fast cardiac CT simulation using a Graphics Processing Unit-accelerated Monte Carlo code," *Proc. SPIE Medical Imaging Conference 7622*, p. 762231, 2010.
- [32] A. Badano and A. Badal, *Fast Simulation of Radiographic Images Using a Monte Carlo X-Ray Transport Algorithm Implemented in CUDA*, Burlington MA: Morgan Kaufmann (Elsevier), 2010, pp. 813-830.
- [33] B. Ghamraoui and A. Badal, "Tutorial: Including the effect of molecular interference in the coherent x-ray scattering modeling in MC-GPU," 2014.
- [34] N. Corporation, "CUDA programming guide ©version 10.1.105," 26 Month 2019. [Online]. Available: www.nvidia.com/cuda. [Accessed 10 April 2019].

- [35] J. A. Greenberg, "Coded apertures for faster x-ray scatter imaging," *SPIE Newsroom*, pp. 10.1117/2.1201608.006646 1-3, 16 08 2016.
- [36] P. Mohr, D. Newell and B. Taylor, "The 2014 CODATA Recommended Values of the Fundamental Physical Constants," National Institute of Standards and Technology, Gaithersburg, MD 20899, 2014.
- [37] MultiX X-Ray Spectrometric Imaging, "Multix Products: Sensors," MultiX, 2015. [Online]. Available: <http://www.multixdetection.com/product/sensor/>. [Accessed 18 March 2019].
- [38] ICRP, "ICRP Publication 110: Adult Reference Computational Phantoms," *International Commission on Radiological Protection*, vol. 39, no. 2, 2009.
- [39] J. Siewerdsen, "Technical Note: spektr 3.0—A computational tool for x-ray spectrum modeling and analysis," *Am. Assoc. Phys. Med*, vol. 43, no. 8, pp. 4711-4717, 2016.
- [40] J. M. Boone and J. A. Seibert, "An accurate method for computer-generating tungsten anode x-ray spectra from 30 to 140 kV," *Med. Phys.*, vol. 24, no. 11, p. 1997, 1997.
- [41] G. Roelofs, J.-l. Gailly and M. Adler, "Zlib: A Massively Spiffy Yet Delicately Unobtrusive Compression Library," [Online]. Available: <http://zlib.net>. [Accessed 23 2 2019].