R1-C.3: Characterizing, Modeling, and Mitigating Texturing in X-Ray Diffraction Tomography

I. PARTICIPANTS INVOLVED FROM JULY 1, 2019 TO JUNE 30, 2020

Faculty/Staff			
Name	**Title**	**Institution**	**Email**
Joel Greenberg	PI	Duke University	joel.greenberg@duke.edu
Anuj Kapadia	Co-PI	Duke University	anuj.kapadia@duke.edu
Dean Hazineh	Research Associate	Duke University	dean.hazineh@duke.edu
Graduate, Undergraduate and REU Students			
Name	**Degree Pursued**	**Institution**	**Month/Year of Graduation**
Stefan Stryker	PhD (Med Phys)	Duke University	5/2022
Oluwadamilola Fasina	MS (Med Phys)	Duke University	5/2021
Jordan Houri	MS (Med Phys)	Duke University	5/2021
Jeffrey Xiao	MS (Med Phys)	Duke University	5/2020

II. PROJECT DESCRIPTION

A. Project Overview

The detection of contraband and explosives concealed within a large volume of confounding items is a challenging task. X-ray based imaging has been central to this effort for nearly the last two decades, but conventional transmission-based methods lack the material specificity required to accurately detect the target items. X-ray diffraction tomography (XRDT) has the potential to dramatically improve the sensitivity and specificity of X-ray–based explosives detection systems because of its ability to identify concealed materials based on their microscopic atomic and/or molecular structures. As such, it has begun appearing on DHS roadmaps as a future solution for reducing false alarm rates while increasing the threat list. While this sensitivity to molecular composition is necessary for accurately assessing a host of benign and threat materials, it may at times be too sensitive: the details of the measured X-ray diffraction scatter signatures can depend on a variety of factors relating to environmental history and conditions. This can lead to a variation in the X-ray diffraction (XRD) "signatures" that complicate classification and reduce the efficacy of the XRD system. A well-known example of this effect is the dependence on the degree that crystallinity, orientation of crystalline grains, size of the grains, and other related factors result in seemingly random modulations of the powder XRD patterns, called texture. In this project, we took a multipronged approach to investigating the role of texture as it relates to aviation security in order to understand its impact as well as provide methods for mitigating its effects. In particular, we addressed the following topics, with the goal of improving the performance of current and future XRD systems:

- **Empirical testbed:** Built a novel experimental system to explore the prevalence of texture in both stream of commerce and explosive materials.

- **Database creation:** Acquired scatter data for multiple instantiations (e.g., different orientations, grain sizes, and processing history) of a broad range of materials of interest and incorporated the measurements into a shareable database (hosted by the University of Rhode Island [URI]).
- **Simulation tools:** Developed an empirically informed approach to modeling texture in arbitrary XRD systems and implemented it in both a deterministic and GPU-enhanced Monte Carlo (MCGPU) simulation approach. The MCGPU tool is available for sharing with interested parties.
- **Detection algorithms:** Explored machine learning and nonlinear dimension reduction techniques for robust detection performance in the presence of texture. We showed that support vector machine, random forest, convolutional neural net algorithms trained on simulated data lead to improved probability of error when applied to experimental data, as compared to the conventional matched-filter or correlation-based classifier approaches. These approaches generalize nicely to the case of mixtures (e.g., homemade explosives) and to situations when the library is incomplete.
- **System design:** Performed system-level studies that demonstrate measurement approaches that minimize the impact of texturing on detection and imaging performance.

The tools and findings of this project have already been adopted by other members of the aviation security industry and will be further applied and refined via separate DHS-funded efforts (e.g., the current BAA 14-02 T/Si program led by Duke in collaboration with Smiths, University of Arizona, and Quadridox) aimed at developing and deploying XRD systems at the checkpoint and in checked baggage.

B. State of the Art and Technical Approach

The topics of texturing, numerical simulation methods, and XRDT architectures have all been studied previously; however, none of these previous investigations have been unified. As a result, there exist critical gaps in quantifying and overcoming the effects of texturing that make accurate evaluation of XRDT impossible. We address these gaps and evaluate the utility of XRDT in aviation security applications.

B.1. X-Ray Diffraction Tomography

X-ray diffraction has been shown to greatly increase the material specificity of X-ray-based inspection; as a result, it can lead to significantly reduced false alarm rates [1]. XRDT involves determining the XRD signature associated with each voxel throughout an object. Unfortunately, the physics of XRD conflates the location and material composition of an object. While several methods for performing the XRDT task have been studied previously, the two methods most applicable to aviation security are direct tomography (DT) [2] and coded aperture XRDT (CA-XRDT) [3-6]. In DT, one uses collimators on the source and detector side of the object, such that each detector pixel images only a single voxel within the object. With the signal localized in this way, the XRD signature is recorded through the use of energy-sensitive detection at each pixel. This technique forms the basis of the Smiths HDX 10065 and Smiths XDi machines [7]. While DT has a well-defined spatial resolution and low computational overhead, the heavy collimation results in extreme photon starvation in the imaging system, making the scanning system slow and/or requiring expensive, complicated components [8]. In addition, the requirement of viewing an object voxel from only a single angle may result in the system measuring incorrectly or missing altogether the scatter in the case of a highly textured material.

B.2. Characterization of Texturing

For more than one hundred years, the fundamental science behind XRD has been developed and applied to understanding the structure and properties of materials. It is well known that the XRD signal depends on the molecular structure of the material at the microscopic scale. Because of this, XRD yields a signature for

distinguishing materials from one another (i.e., performing material identification); however, details in the microscopic structure of the material can impact the XRD signature. A particular example of this is known as texturing, in which the type, degree, and relative orientation of the crystal structure of a sample impacts the details of the XRD scatter signal (see Figure 1). While these details are useful for investigating the properties of a known material, they pose challenges (and therefore are usually ignored) in the case of performing material identification and/or threat detection of an unknown material.

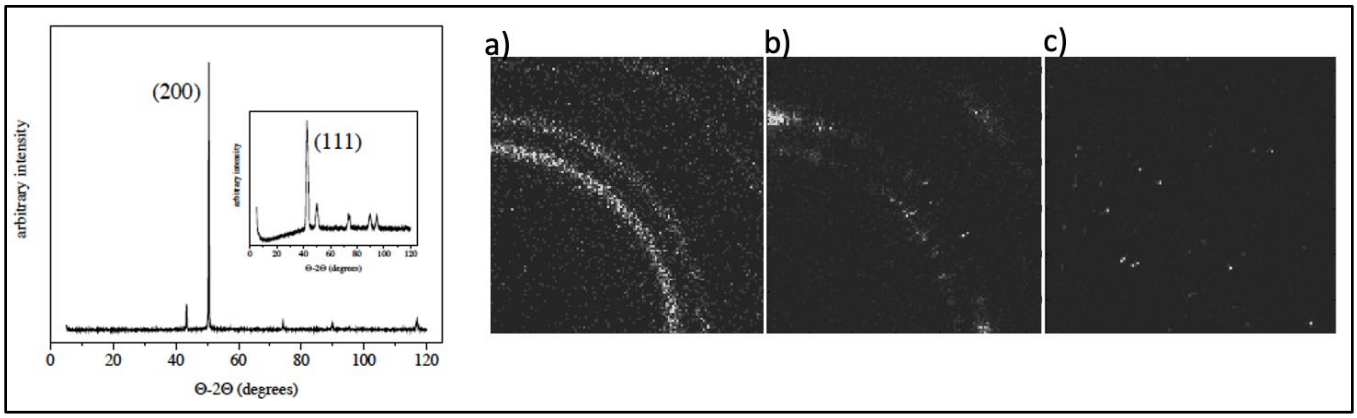


Figure 1: Degree of texturing; (left image) XRD form factor for the same copper sheet with two different orientations relative to the x-ray beam; (right three images, in order of increasing texturing from left to right) XRD scatter at 60 keV using a pencil-beam CAXSI system and 2D energy-sensitive detector spanning π/2 for (a) Al powder, (b) Al sheet, and (c) ammonium nitrate powder. Texture modulates the scatter intensity in both angle and energy in an unpredictable way.

A variety of XRD databases exist; however, they are of limited value to aviation security for several reasons. First, they typically involve measurements made with carefully prepared powder samples (as opposed to solid or highly crystalline samples), which guarantees that texturing is absent. Furthermore, they typically only cover specific classes of materials (e.g., biomolecules or metal alloys) which do not adequately represent many common objects found in luggage (such as toothpaste, cheese, paper, and water). Finally, existing databases usually include only a few instantiations for each unique material and do not included various orientations of textured materials. Without an adequate understanding of the inherent variability of the XRD signatures, one cannot evaluate the performance of any real-world XRDT system.

In addition, the conventional XRD measurements are made using a standard X-ray diffractometer operated in reflection mode at a single, low (e.g., 8 keV) X-ray energy to record the scatter form factors; however, for aviation security applications, measurements are made in transmission mode and require high energies (20–180 keV) in order to penetrate bags. Because the scatter cross-section changes as a function of energy and angle, it is important to record the complete energy- and angle-dependent scatter cross-sections. While a synchrotron system has recently been built to measure the angle- and energy-dependent XRD signal over a limited range of energies [9], this method is hardly practical for use in real labs and too expensive to use to generate large databases worth of data.

Without the aforementioned databases (which include stream of commerce materials measured at appropriate energies and including material variations), it is extremely difficult to design XRDT systems or develop appropriate detection algorithms, as material signatures and their associated variability are central to these tasks. We are therefore building a comprehensive database relevant to threat detection and developing a statistical description of texturing in order to improve the performance of existing systems and properly evaluate the performance of future scanners.

B.3. XRD Simulation Tools

Numerical tools that enable simulation of X-ray physics as well as dependencies on component behavior (e.g., source, detector, and collimator) and geometry are critical for understanding and evaluating the benefits of XRDT systems. The two dominant simulation methods are (1) deterministic numerical models for calculating mean quantities (which typically consider only first order scatter phenomena) and (2) Monte Carlo (MC) techniques (which are exhaustive but time-consuming). The advantage of a deterministic numerical model is that it is fast to run and can therefore be used to perform design studies over a broad trade space. A robust model of this sort is also necessary for performing model-based reconstruction for an XRDT system (e.g., recovering the threat status of an object based on the raw measurements). On the other hand, MC is the most physically accurate simulation technique and naturally accommodates stochastic processes (such as multiple scatter and noise). Although time-consuming, Monte Carlo enables realistic and accurate determination of system performance with all relevant sources of noise.

Models of both types have been developed previously by several groups [10-12]. We have previously developed state-of-the-art simulation models of both types: a deterministic numerical model of Compton scatter and XRD [8], as well as an MC model based on GEANT4 [13] that includes empirically measured XRD signatures in place of the conventional Hubble form factors. We have previously validated our models against experimental measurements and used them to study various XRDT configurations (see Figure 2) [3, 13]. In addition, the FDA has developed a GPU-accelerated MC approach (MCGPU) that is intended for simulating a limited range of materials at X-ray energies [14]. While useful, none of the models developed to date include the energy and angular dependencies exhibited by textured materials and are therefore constrained to describing a limited set of materials. We are therefore extending the current models to include a more complete description of texturing in order to faithfully represent real-world materials present in luggage.

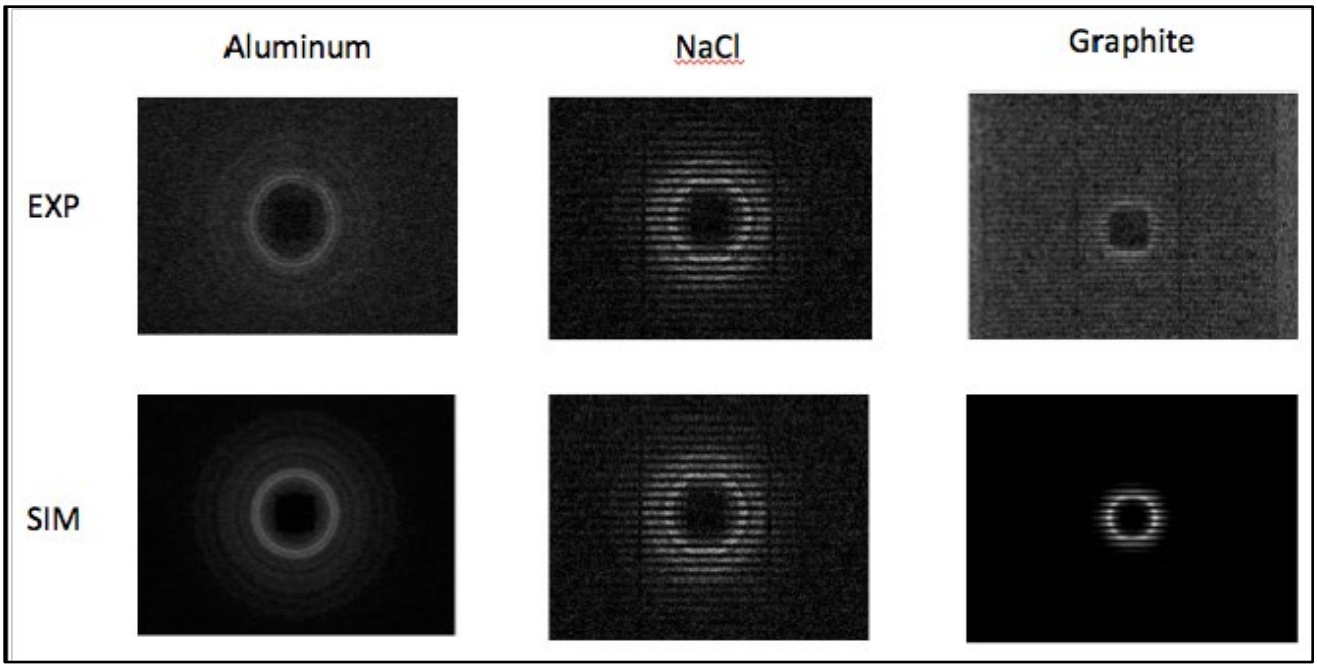


Figure 2: Comparison of raw experiment (top) and MC simulation (bottom) XRD data for Al, NaCl, and graphite powders (left to right).

B.4. Data-centric Analysis of XRD Signatures

Traditional XRD threat classification methods usually involve two separate steps in order to assign a class to a particular measured form factor. In the first step, one performs material identification, which is a multi-class classification problem, by comparing the unknown form factor to a premeasured library of form factors and determining the best match. Since X-ray diffraction is most often performed with crystalline materials (either naturally occurring or artificially crystallized) whose form factors consist of a multitude of narrow peaks, it is common to simply compare the locations (although not the amplitudes) of the prominent peaks against the locations of peaks of the form factors in the library. If enough of the peaks coincide within some predetermined accuracy, then a match is made. This approach, however, does not work for amorphous and liquid materials, in which scatter occurs at a broad range of momentum transfer values and the relative peak intensities are important in distinguishing materials. In this case, one can instead calculate a similarity metric, such as the correlation (which we use in this work), between the unknown form factor and all elements of the library. Whichever material is most similar to the unknown material is then matched with the unknown material. After the material identification process is complete, the second step is then to assign the class of the matched material to the previously unknown material. While this process can work in some situations, there are several challenges associated with it. For example, the entire process fails if the unknown material is not already in the library, which is a significant issue given the breadth and evolution of explosive-related threats. Furthermore, it can be difficult to accurately associate a given form factor with the correct library element in the presence of noise, measurement uncertainty, and material variability.

The problem of the traditional identify-then-classify approach stems from the fact that it does not make use of similarities between members of a given class (or, more specifically, does not take into account class differences). If one considers correlation-based classifiers, for example, it is clear that each feature (e.g., momentum transfer value or principal component) receives a uniform weight when performing the intermediate material identification step. In this case, a missing peak (as in the case of texturing) or a noise spike at a location that happens to coincide with that of a different material could produce an incorrect identification and, correspondingly, a classification error.

In order to improve the detection performance relative to traditional approaches, in this program, we turn to machine learning as a method to automatically generate a set of feature weights that are optimized for performing the specific task of interest. Note that this approach does not necessarily aid in the intermediate problem of identifying the unknown material or in obtaining any auxiliary data about the material (e.g., whether it is a solid or a crystalline material). Instead, it eliminates the need for the intermediate identification step entirely: the traditional process collapses to a single step that simply uses a previously-trained classier to identify directly whether a material is an explosive or not. With sufficient training data, the use of machine learning therefore provides the potential for increased robustness to noise and better performance in the face of "new" form factors (i.e., those on which the classier has not been trained). Some previous studies have investigated support vector machines (SVMs), principal component analysis, or hybrid discriminant analysis to identify classes of interest; however, these studies have always been limited to narrow categories such as the evaluation of only liquid materials, only considered linear dimension reduction methods, or did not account for XRD signature variability arising from intrinsic effects (such as texture).

C. Major Contributions

C.1. Construction of an Energy-Dispersive Laue Diffraction System (Year 4)

We built an experimental testbed X-ray system capable of collecting the XRD signal over a large range of energies and angles relevant to aviation security [15]. This allows us to collect the 3D scatter data (in terms

of 2D scatter angle and energy) that will be included in the XRD database. All components are commercial off-the-shelf (Figure 3), and a high SNR scan of a sample can take between 5 and 45 minutes, depending on the material.

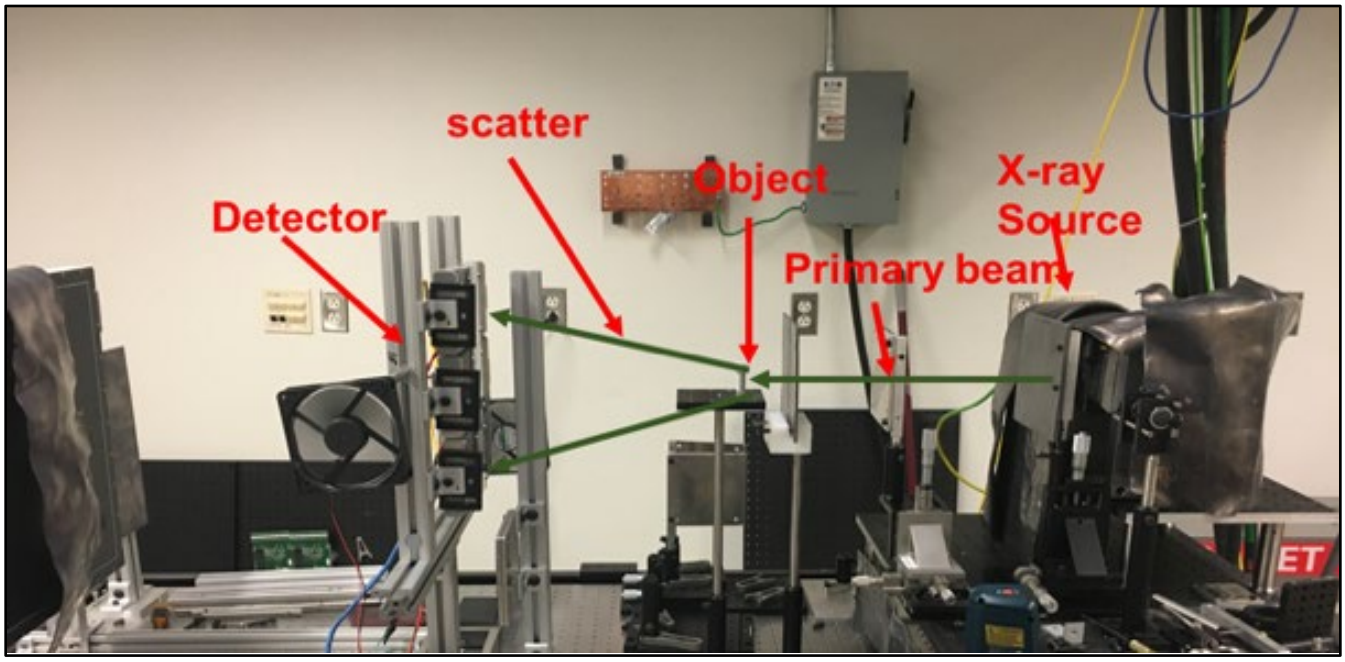


Figure 3: Experimental setup of our energy-dispersive Laue diffractometer. From right to left, we show the X-ray source, series of pinhole collimators, object, and energy-resolving, pixelated detectors. Both the detectors and object are placed on rotation and translation stages to enable flexible illumination and detection.

C.2. Commercial XRD Database Creation (Years 4–7)

Using a conventional, commercial diffractometer system (Bruker D2 Phaser, operated in reflection mode at the copper k-Alpha line), we acquire high-resolution scans of the XRD form factors. To record a wide range of materials in different orientations, we have modified the system to enable title and azimuthal rotation of the sample, to vibrate the sample to promote randomness, and to heat or cool the same to vary the sample temperature. Using this setup, we measured over four hundred different substances, including both benign and energetic materials. For explosives, we visited Professor Jimmie Oxley's (Projects R1-A.1, R1-B.1, and R1-C.2) lab at URI and scanned a variety of recipes of explosives (in powder and slurry form) at a range of different orientations on both her Rigaku XRD system and our Bruker system. Examples of the recipe dependence and orientation dependence of the XRD signatures are shown in Figure 4. Since it is extremely difficult to accurately predict the XRD signature from first principles for complicated materials, this empirical approach is necessary for understanding and simulating XRD scatter. In Year 6, we added more liquids to the database, provided additional material metadata (such as whether a material is solid/liquid, crystalline/amorphous), and included the energy-dependent attenuation coefficient for the materials in addition to their XRD signatures). In Year 7, we included additional stream of commerce materials of particular relevance to discrimination of organic materials with some degree of texturing.

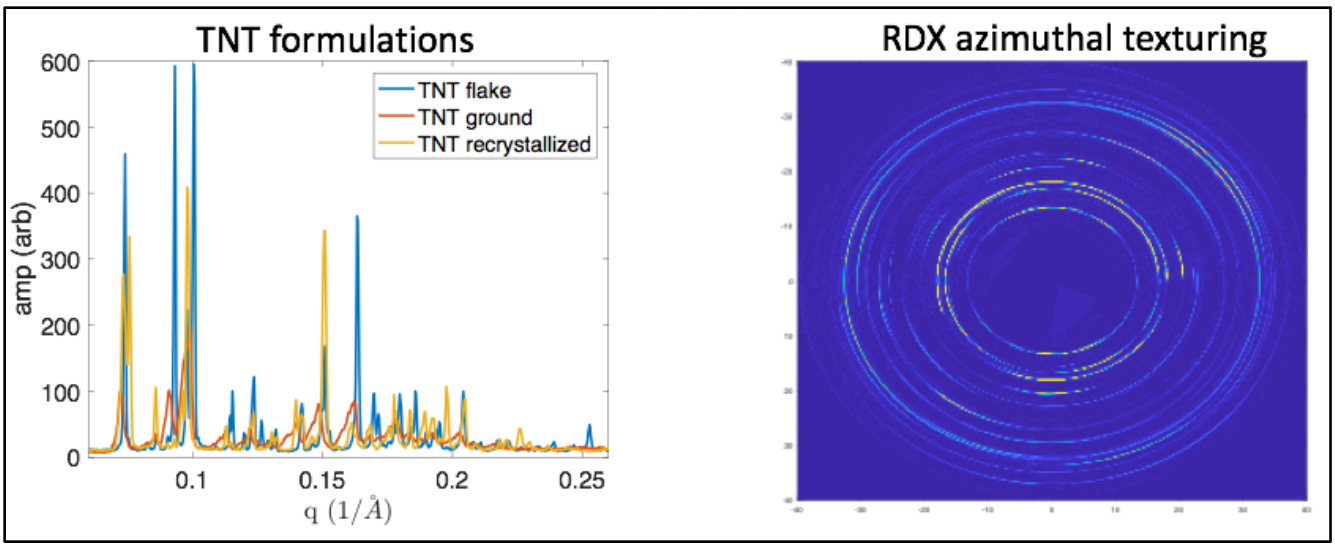
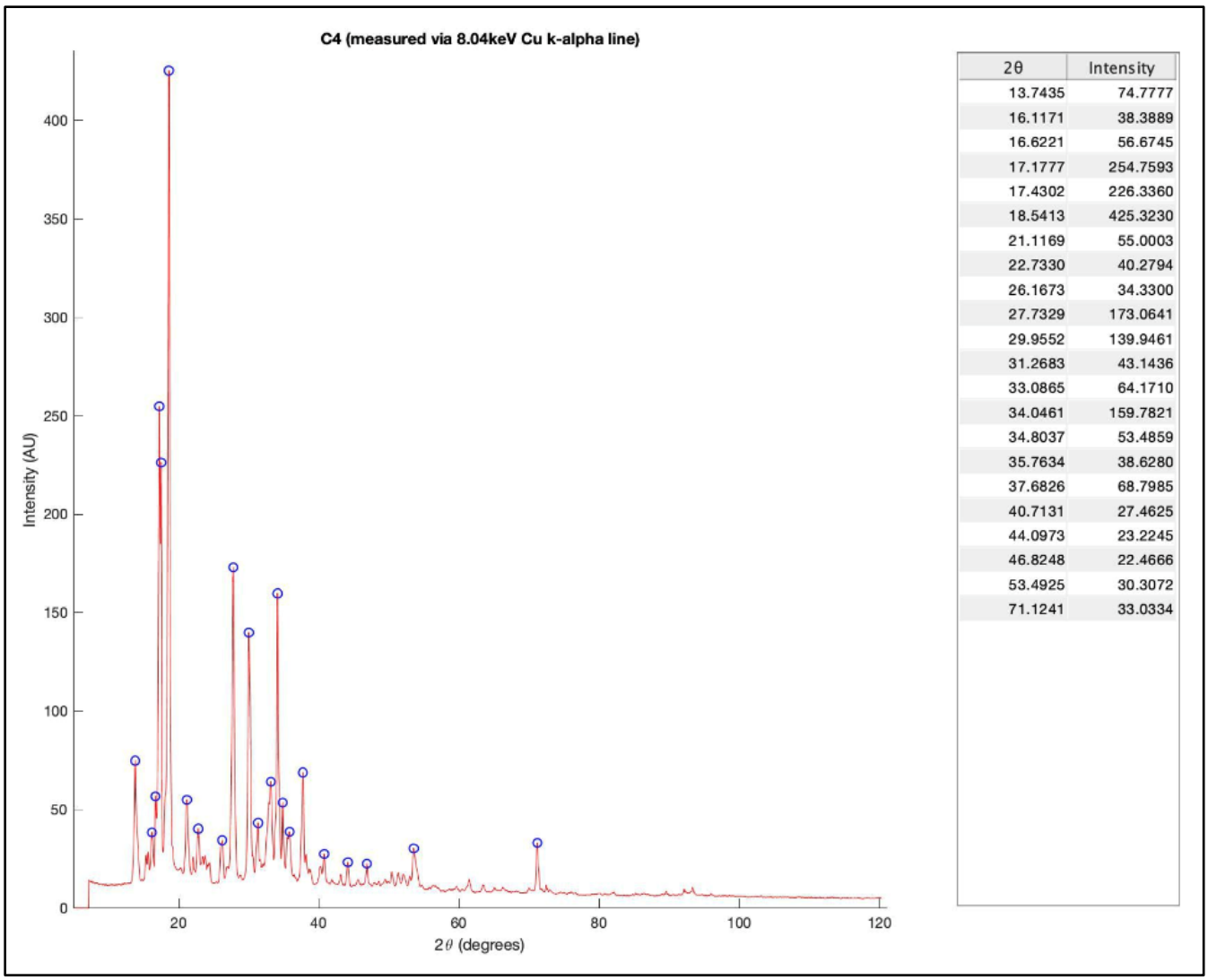


Figure 4: (Left) X-ray diffraction form factors for TNT in flake, ground powder, and recrystallized crystalline forms. The peaks occur at nearly identical locations, but the relative amplitudes vary significantly (despite the identical molecular composition of the samples). (Right) An image of the scatter pattern for a sample of RDX as it is rotated over 360 degrees in the sample holder. For all measurements, we use a Bruker D2 phaser system.

In Year 7, we formalized and organized the data and formatted it for inclusion in Jimmie Oxley's Explosives Database. An example of such a file is shown in Figure 5—each material has an associated PDF file showing a plot of the spectra and a list of the locations of the dominant peaks. The full data for scatter intensity versus diffraction angle is included explicitly in table form on the succeeding pages. The files were shared with Jimmie and her team.



2θ	Intensity
13.7435	74.7777
16.1171	38.3889
16.6221	56.6745
17.1777	254.7593
17.4302	226.3360
18.5413	425.3230
21.1169	55.0003
22.7330	40.2794
26.1673	34.3300
27.7329	173.0641
29.9552	139.9461
31.2683	43.1436
33.0865	64.1710
34.0461	159.7821
34.8037	53.4859
35.7634	38.6280
37.6826	68.7985
40.7131	27.4625
44.0973	23.2245
46.8248	22.4666
53.4925	30.3072
71.1241	33.0334

Figure 5: Example cover page image for C4 as part of the URI Explosives Database.

C.3. Fundamental Study of the Relationship between Grain Size and Texturing (Year 5)

Powder material attributes of crystallinity, crystal faceting, and powder packing collectively influence texture and the corresponding diffraction spectra are shown in Figure 6. We have analyzed how the X-ray aperture size and crystal grain size influence the XRD signatures. Smaller powder size relative to the X-ray aperture (1 mm × 16 mm) showed diffraction features characteristic of randomly configured powder since many particles and their configuration were analyzed in the sampling volume. Figure 6 shows more spectral variability, or texturing, exists for larger powder grain size relative to the fixed aperture size due to the fact that fewer grains were included in the illuminated volume. A new finding was that the manner in which the powder is packed into the holder impacts texturing, as shown in Figure 6.

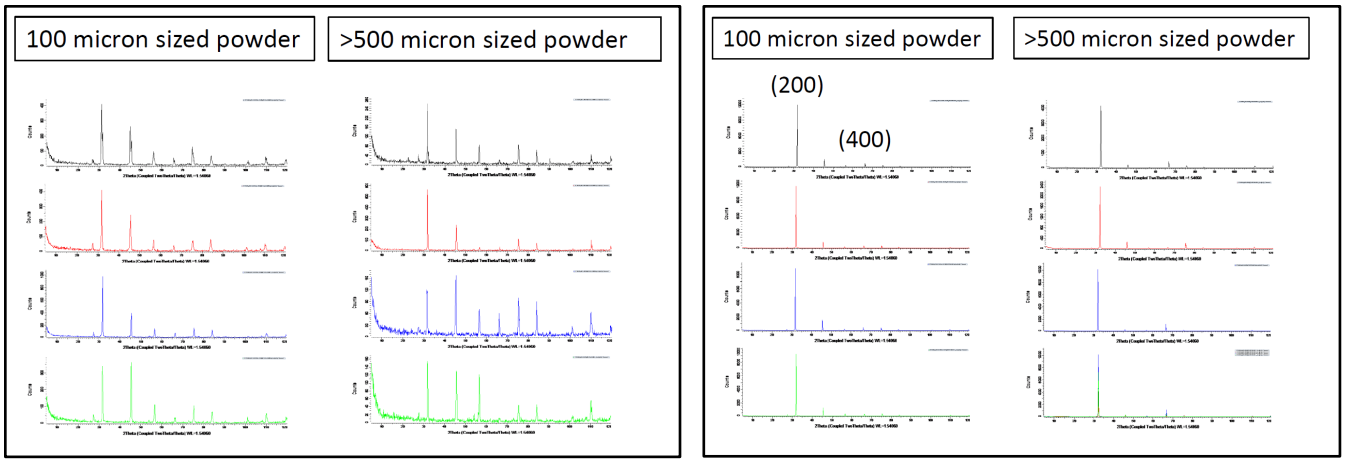


Figure 6: X-ray diffraction spectra for 100 mm and >500 mm NaCl powder for material loosely placed (left) and pressed into the sample holders (right).

C.4. Energy-Dispersive Laue Database Creation (Years 4–7)

Using the energy-dispersive (ED) Laue system, we scanned over one hundred materials and recorded the scatter signal from 20–170 keV and for a deflection angle between 1 and 30 degrees. As shown in Figure 7, each material was placed in the system with multiple orientations, and many in multiple forms (e.g., sheet, powder, different vendors, and different processing histories). We find that fine powders, such as Al, give symmetric rings with no texturing at all energies and orientations. Aluminum sheet, however, has textured rings that vary around the ring (azimuthally) and depend on the orientation of the sample relative to the beam, but have no variation or structure in energy. For polycrystalline powders with large grain sizes, such as salt in a shaker, scatter only arises at particular angles and energies, and depends sensitively on the configuration of the individual polycrystalline grains. This indicates the challenges associated with XRDT on highly textured materials, as both collecting and interpreting the signal can be very challenging (compared to an untextured material). In Year 6, we added fifty materials (in multiple orientations) to the database, focusing on materials with some degree of crystallinity that existed already in the XRD database (including some single crystals) [15].

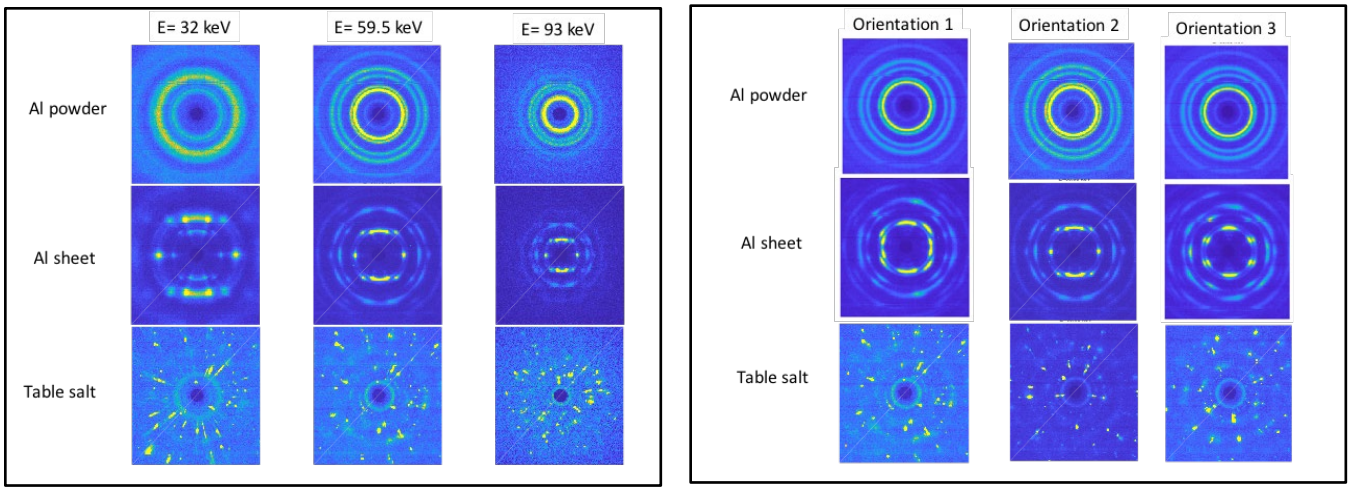


Figure 7: Scatter signals measured across different energies using the ED Laue system (left). The top, middle, and bottom rows show the scatter for aluminum powder, aluminum sheet, and large grain sodium chloride, respectively. From left to right, we show the scatter at 32, 59.5 and 93 keV. Together, this data shows examples of different ways in which the XRD signal may deviate from the Debye cones described by Bragg's law. Scatter signals measured for different orientations of the same samples in the ED Laue system (right). The top, middle, and bottom rows show the scatter for aluminum powder, aluminum sheet, and large grain sodium chloride, respectively. From left to right, we show the scatter for three different orientations of the sample in the system (with the rest of the parameters held fixed). This data shows examples of the orientation dependence for materials with different degrees of texturing.

In Year 7, we have continued scanning more materials to add to the database. These materials were chosen to both broaden the stream of commerce breadth of the database (e.g., including liquids and foods) and to include additional materials with "interesting" texture patterns (e.g., single crystals and formed metal objects). Due to COVID-19-related lab closures, this work is still ongoing.

C.5. Texturing Taxonomy (Year 5–6)

Using the ED Laue database results, we quantified the type and degree of texturing by developing a custom metric that characterizes the variation of the scatter signal around the ring (azimuthal texturing coefficient) and across different energies (energy texturing coefficient). By plotting each material in this space, we can visualize the distribution of materials (see Figure 8). A new result in Year 6 is that our metric places materials with similar structures and processing histories near each other and emphasizes these attributes over molecular structure. For example, copper sheet and aluminum sheet appear nearer to one another and farther away from copper powder and aluminum powder, despite having the exact same chemical and molecular composition. This shows that, rather than a nuisance, texture can be used to determine the processing history of a material. In addition, we find that only part of the texture space is filled—materials with a high degree of energy texturing generally also have a high degree of angular texturing. Finally, we can define a single parameter, referred to as the degree of texturing (DOT), which quantifies how textured the material is. For the materials in our library (which represent common, stream of commerce materials), we find that over half have at least some degree of texturing. These findings are important for simulating ensembles of objects with appropriate degrees of texturing, as well as giving insight into the texturing "categories" into which materials fall, so that we can get a representative understanding of the range of possible scenarios while studying only a limited number of cases [15].

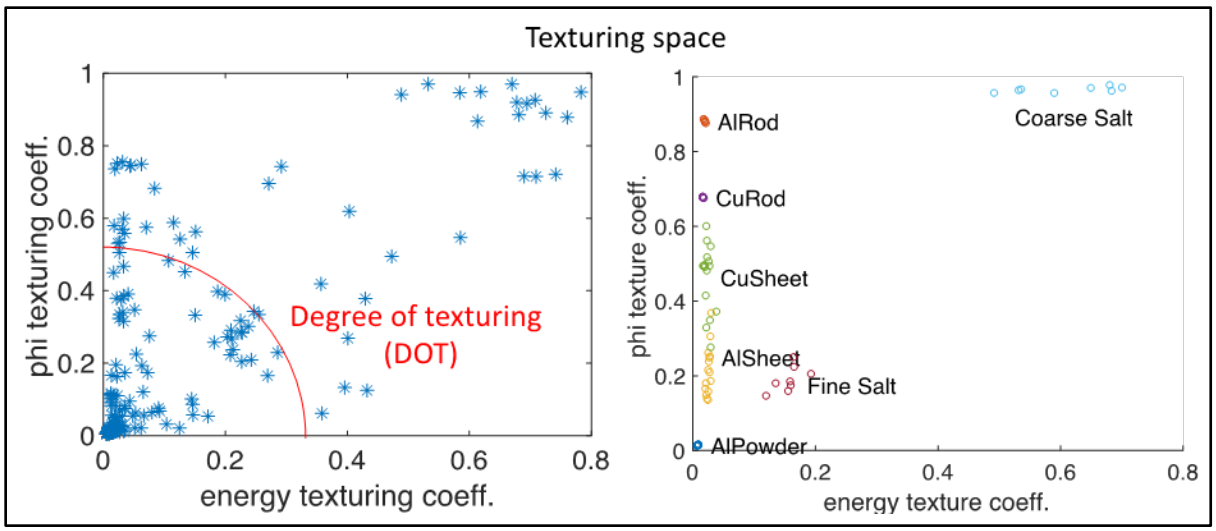


Figure 8: Texturing taxonomy showing the distribution of materials in the ED Laue database in terms of the degree of texturing of each material and in terms of energy and azimuthal scatter angle (left). The materials are clustered in the upper left hand corner of the plot. The degree of texturing is defined as the distance from the origin. Highlighting materials whose processing history can be distinguished based on texture alone (right).

C.6. Using Texture as a Feature for Material Identification (Year 7)

Based on the characterization of materials in texture space from the previous year, we investigated whether texture could be used as a useful feature (rather than a problem) in the XRDT detection paradigm. We focused on the case of aluminum, for which we had access to a single crystal, sheet, rod, and fine powder. Figure 9 shows that using texture features related to the intensity variation around a Debye ring (c_phi) and as a function of energy (c_E) allows one to separate powders from solid sheets and rods much more accurately than by using the conventional XRD form factor (fq) alone.

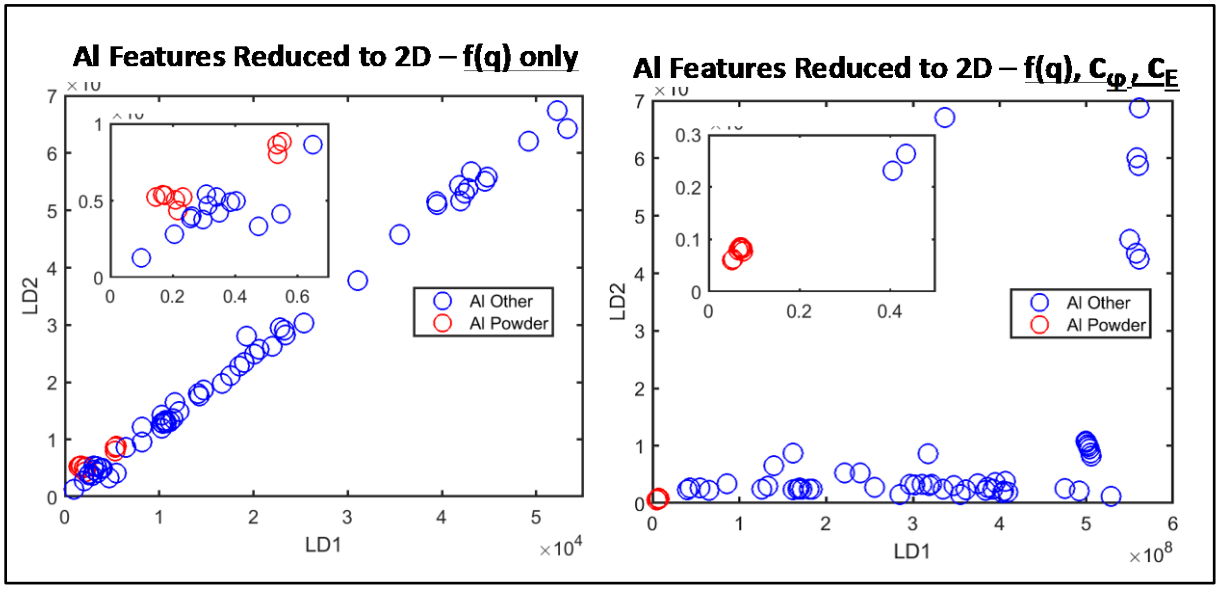


Figure 9: Plot showing Al powder (blue) and other forms of Al (red) in linear discriminant analysis space. When only f(q) is used, there is substantial overlap between the powder and other forms of Al, whereas when the texture features are used, the different forms of the same substance are easily separable.

C.7. Comparing Transmission Versus Diffraction for Material Identification (Years 4 and 7)

A natural question to ask is how well explosives and nonexplosives can be distinguished from one another based on their X-ray properties. We considered the class separation (average distance between explosives and nonexplosives) using three different feature sets: density and effective atomic number, the XRD signature, and a combination of the two. We find that the XRD features provide significantly better class separation, J, for reasonably high-quality XRD measurements (indicated in Figure 10 as Δq) as compared to the transmission-related features (density and the effective atomic number). Also, combining these features gives rise to better separation than either feature set alone. This shows that classifiers that combine transmission and XRD data should outperform those that use only one measurement approach, and that XRD fundamentally has the potential to outperform transmission measurements.

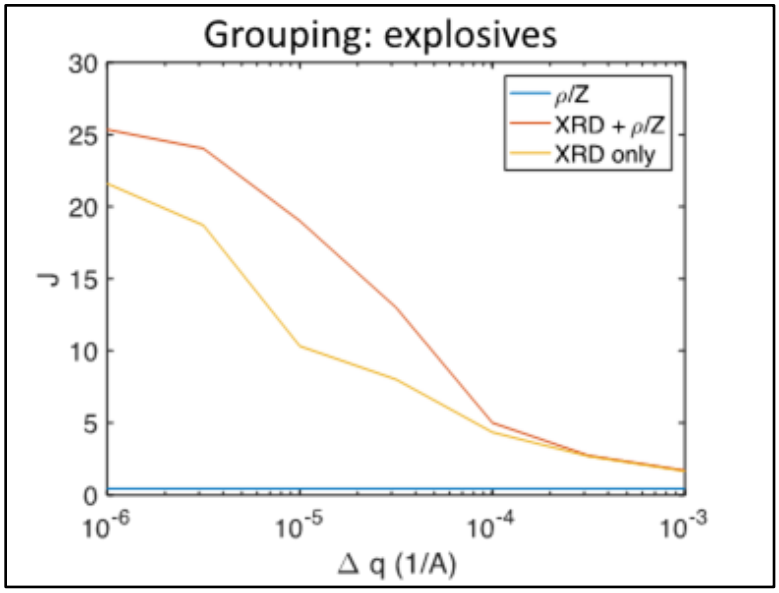


Figure 10: Class separation, J, as a function of XRD momentum transfer resolution (i.e., quality of the XRD measurement) when transmission (density and effective atomic number), XRD, or transmission + XRD features are used to discriminate between explosives and nonexplosives.

C.8. GPU-Based Monte Carlo: Fast and Accurate Simulation (Year 6–7)

We have built and implemented a Monte-Carlo simulation for X-ray diffraction using the software MCGPU. MCGPU is a freely available software developed and maintained by the US Food and Drug Administration. It is designed to model X-ray interactions at energies of relevance to our diffraction-based modeling and detection approach. The software utilizes GPU frameworks and can be fully parallelized across multiple GPU processors, either within a single computer or across multiple computers. Based on our early experiments to measure GPU-computing efficiency, we have found speed-ups of up to 50X compared to CPU-based computing, such as in GEANT4. We have created our own MATLAB wrapper for simple end-user interaction with the software and modified the default coherent scattering models in MCGPU to incorporate XRD signatures derived from our diffractometer database. Using this modified code, we have demonstrated preliminary validation against data generated via experiment and GEANT4 (see Figure 11).

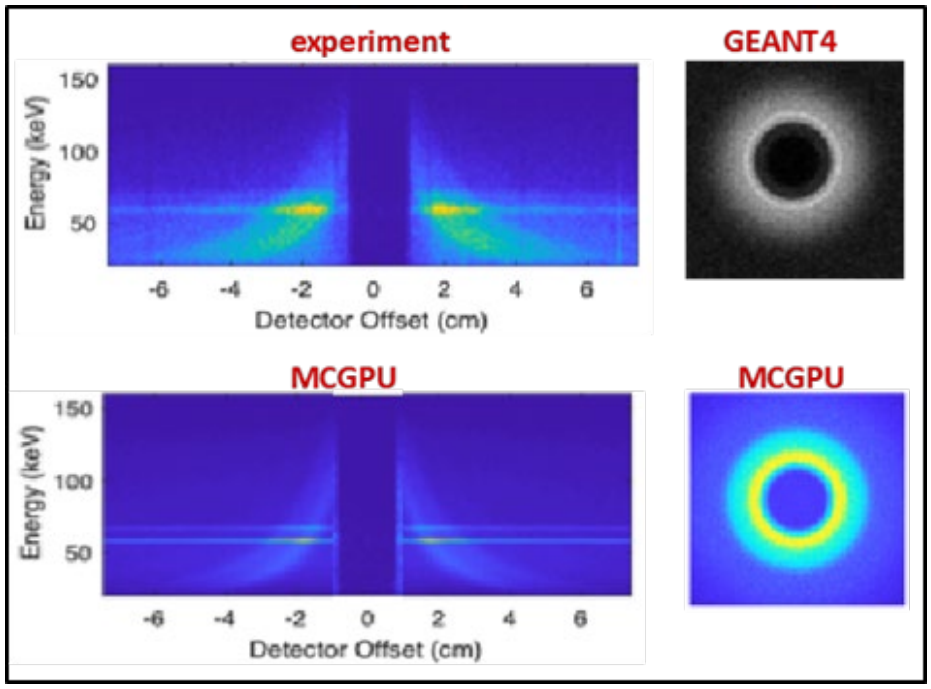


Figure 11: Validation of MCGPU against experiment (left) and GEANT4-based Monte Carlo (right).

In Year 7, the MCGPU model has been tested and validated against experimental data acquired using the ED Laue testbed system. Shown in Figure 12 are simulated versus experimental spectra from water, graphite, and aluminum. The total scatter signal was accurate to within 5% of the values obtained from literature. Furthermore, the peak locations and spectral shape are consistent with the experimental measurements. Additional validation of the accuracy and speed benchmarking against GEANT4 (the Monte Carlo "gold standard") show that MCGPU is over 100 times faster than GEANT4 while producing nearly identical results.

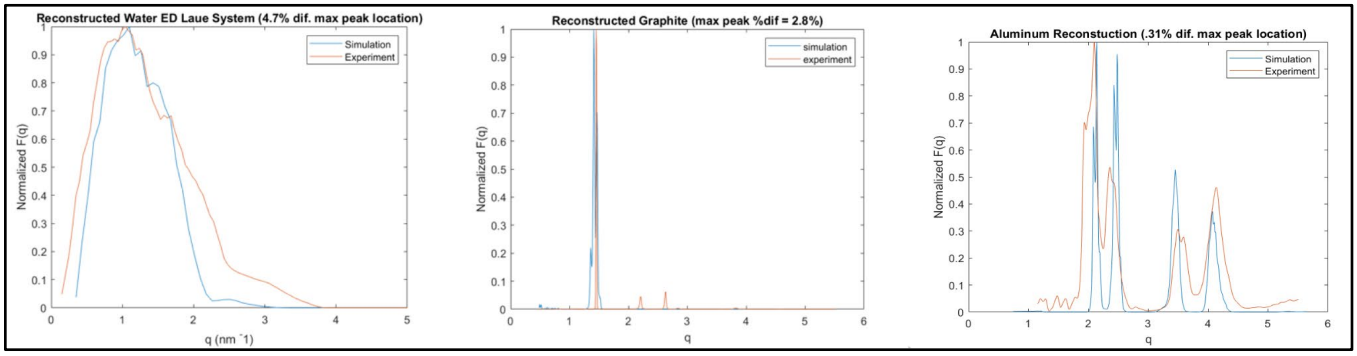


Figure 12: Simulated (blue) and experimental (orange) XRD signals for water, graphite, and aluminum powder (left to right, respectively).

C.9. Implementing Texturing in Simulation: Deterministic Model (Years 4–6)

We previously developed a fast, deterministic, first-order scatter model of XRD and Compton scatter for untextured materials. Because the details of the texture pattern are impossible to predict ab initio but the details of their structure are critical to system performance, we implemented an empirical approach to XRD texture modeling in which we use the ED Laue scatter measurements as a texture "mask" in the simulation system. By modulating the scatter at specific energies and angles in the simulation, we can create realistic

scatter distributions for materials whose texturing ranges from negligible to highly textured. We have validated our scatter simulation against experimental data. The simulation enables us to simulate arbitrary system architectures by changing component parameters (e.g., detector type), geometry, and optics (including coded apertures and collimators) for objects with spatial and material descriptions limited only by the extent of our material databases. The general description of this method is shown in Figure 13. In Year 6, we quantified this performance and showed that we can use the model to perform accurate recovery of the XRD form factor from experimental data despite the presence of texturing.

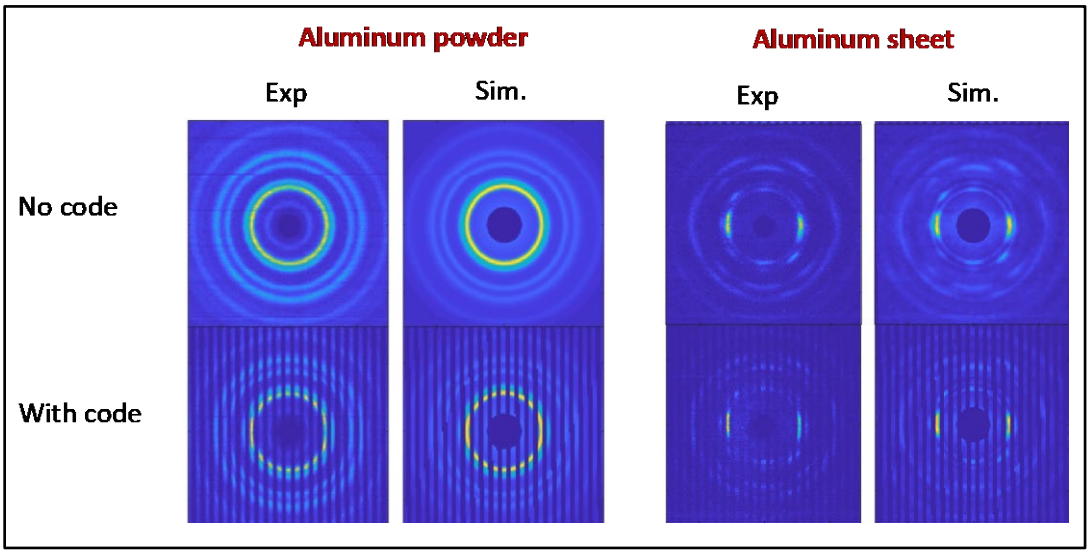


Figure 13: Example of experimental and simulated scatter from aluminum power and a textured sheet of aluminum at 60 keV (without and with a coded aperture).

C.10. Implementing Texturing in Simulation: Monte Carlo (Years 4–5)

We previously developed an accurate Monte Carlo tool (based on GEANT4) for simulating arbitrary scatter orders for untextured materials. As in the deterministic approach previously described, we have extended this MC tool to include texturing by modifying the scatter probabilities across different energies and angles according to empirical measurements made in the ED Laue system. The model has been validated against experiment.

C.11. Simulation-Based XRDT System Analysis (Year 6–7)

A number of methods have been proposed for performing XRDT at the airport. Among them, direct tomography (DT) has been implemented and deployed by Smiths and Morpho, and coded aperture XRDT (CA-XRDT) was invented at Duke and is currently under development by Rapiscan and Smiths. While these systems have been studied separately, they have never been compared head-to-head in a controlled analysis. We developed a method for analyzing these systems in a common framework and evaluated the performance as a function of the number of incident (or detected) photons as well as the size of the target object under analysis (in the absence of texturing). We found that, while the DT system has more information per detected photon, the CA-XRDT system can outperform the DT system overall by virtue of being able to measure orders of magnitude more photons (see Figure 14). Coded aperture XRDT is therefore a promising approach for aviation security [16].

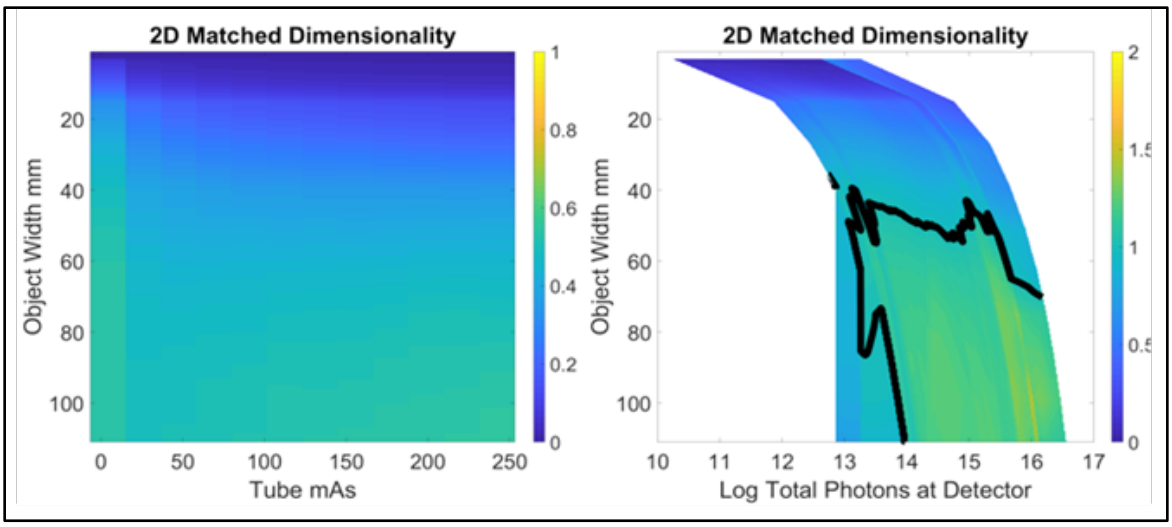


Figure 14: ratio of the error for the CA-XRDT system relative to the DT system as a function of object extend and (left) incident and (right) detected flux. The black line shows the transition between DT and CA-XRDT performing better.

In Year 7, we extended this analysis to include joint DT + CA-XRDT systems. This analysis showed that, while pure coding or pure collimation have individual strengths and weaknesses as the incident photon flux goes down or the object complexity goes up, the coding + collimation system (shown in green in Figure 15 for different degrees of collimation) is able to maintain the strengths of both symptoms and therefore outperform both alternatives.

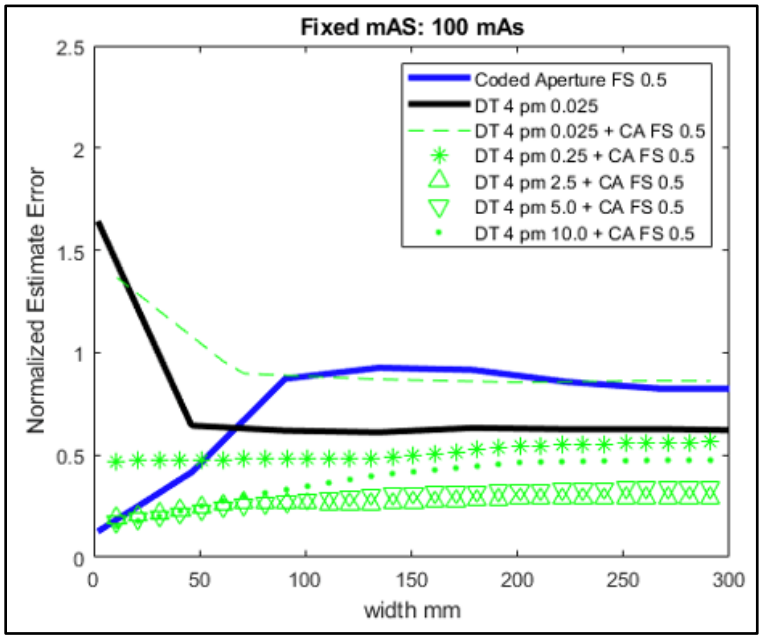


Figure 15: Reconstructed image error as a function of object extent for a coded aperture system (blue), DT system (black), and a CA+DT system (green) with varying degrees of collimation.

C.12. XRDT Performance Analysis with Texturing (Year 5)

We have performed an initial study comparing the impact of texturing on DT and CA-XRDT systems. We found that, while the imaging performance of both systems degrades with increased texturing, DT is somewhat more robust. For both systems, the effect of texturing can be greatly mitigated by using a higher dimensional detector that records both a range of scatter angles and energies [15].

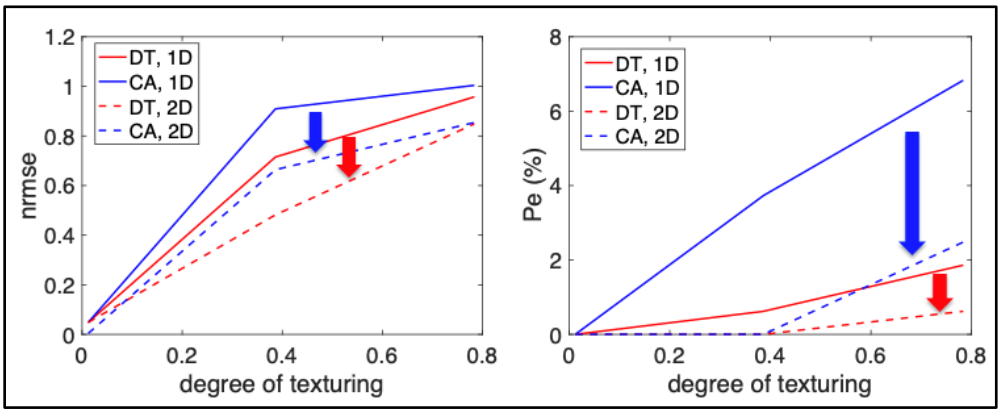


Figure 16: Comparison of reconstructed image error and probability of error in threat/nonthreat classification for a DT and CA system using a 1D and 2D detector as a function of the degree of texturing in the scatter signal.

C.13. Nonlinear Dimension Reduction of XRD Data (Year 6)

Our XRD databases are rich with information; however, considering that each material has multiple texture instantiations (each of which are high-dimensional in their own rights), this data volume can be difficult to interpret. To reduce the data dimension in a manner that preserves the features of interest, we have begun to investigate nonlinear dimension reduction approaches. In particular, we have identified UMAP (Figure 17) as a particularly promising approach—by projecting the high-dimensional XRD signal into a 2D space, for example, we find that individual materials appear as "islands," and clusters with meaningful properties naturally emerge. This is a promising approach both for data-centric analysis of material properties as well for feature engineering in classifier development (as discussed in the next section) [17].

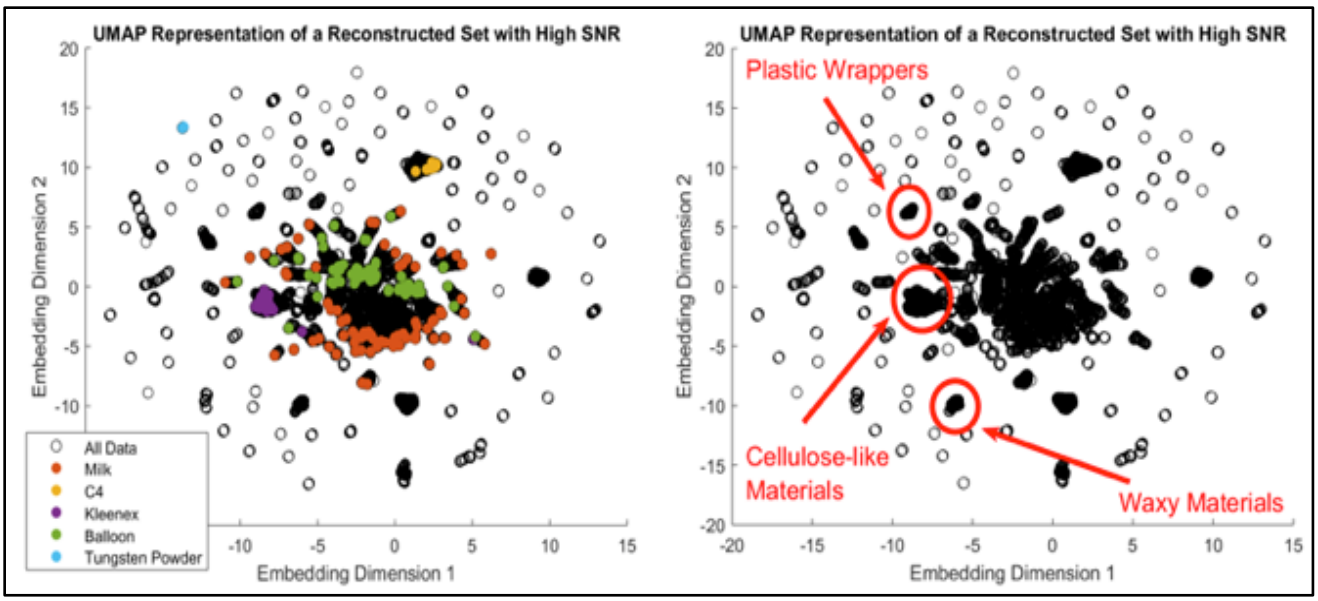


Figure 17: UMAP representation of the reconstructed XRDT data set color-coded to show certain materials (left) and groups (right). The islands represent well-separated materials and the clusters show harder-to-separate materials.

C.14. Machine Learning Analysis of XRD Signals (Year 5–7)

Over the last two years, we have tackled the challenge of building a classifier capable of distinguishing between threats and nonthreats in XRD space. This is a difficult problem because XRD signatures can be highly material-specific, often to the point that having a complete library against which to compare an unknown measurement is impossible: material variability and the large dimensionality of the signature create an infinite number of signatures. We have pioneered the use of machine learning techniques in the XRD space for performing "classification without identification": determining the threat status of an unknown sample without first having to identify what specific material it might be. In Year 6, we found that ML-based approaches, like a support vector machine (SVM) or convolutional neural net (CNN), can outperform traditional classifiers by several percentage points (see Figure 18 for the case of explosives detection, top, and cancer detection, bottom). In general, we find that the advantage of ML approaches exists when the material variability or measurement uncertainty is largest, as is the case of real-world XRDT system run for high throughput (and especially when texturing is present). In addition, we have shown that training these classifiers using simulated data generalizes well to experimental data and that our ML algorithms outperform the correlation-based classifier on empirical data [17, 18].

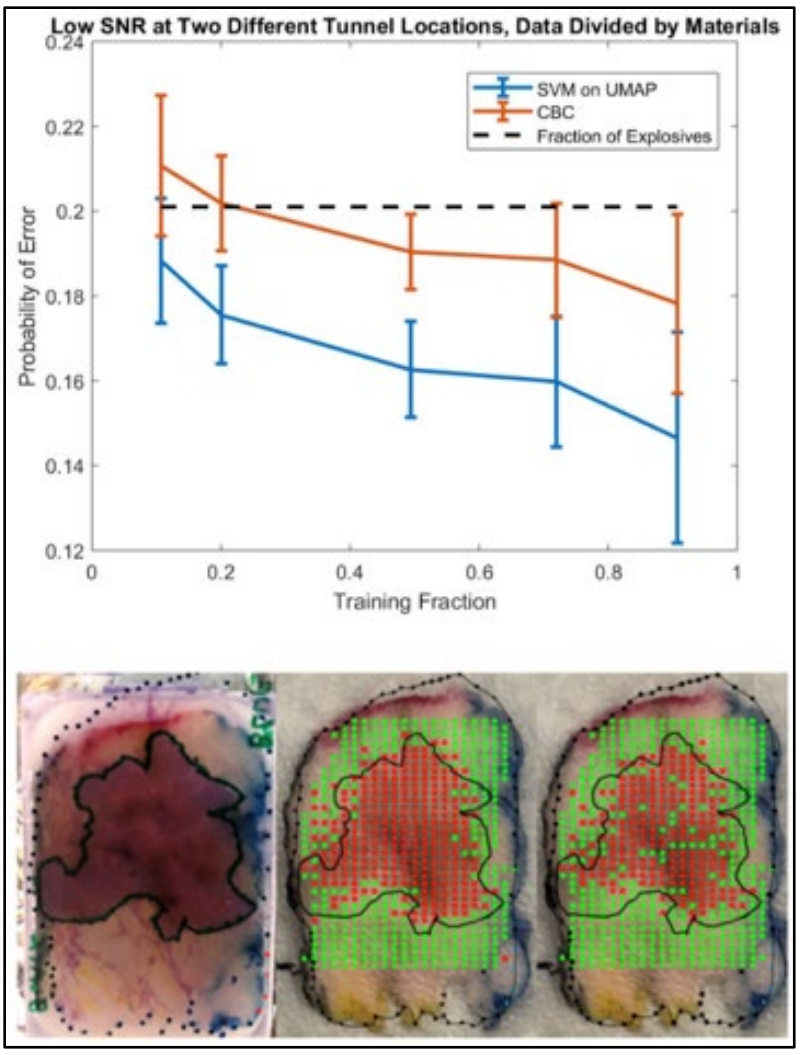


Figure 18: Chart showing the improved error performance of a ML algorithm (blue) versus the traditional algorithm (orange) as a function of the training fraction (top). Images showing cancer classification in tissue (bottom) based on a CNN (middle), traditional classifier (right), and ground truth (left).

In Year 7, we have focused on the problem of spectral unmixing—because the spatial resolution of an XRDT system is typically worse than the sizes of many objects of interest, a single image voxel may end up containing a weighted sum of multiple materials. This mixing is, however, well-described as linear. As such, one must include an additional step between recovering the spectrum at a given voxel and performing classification—namely, one must first recover the abundances and associated spectra of the separate materials present at that location and then perform classification. While we have implemented this successfully with simple cases, we plan to improve spectral unmixing in the constrained case where endmembers are known with providing those and an anomaly class to unmix each pixel. We also plan to look into using synthetic data for pretraining of unsupervised methods of unmixing for case where potential endmembers/materials are too large beyond having a complete database.

D. Milestones

Year 6 Milestone?	Milestone	Description	Status
Yes	XRD database extension	Measure at least 10 instantiations for at least 10 more materials (spanning SoC and prohibited materials) and include transmission information and relevant metadata. Database will enable system-level simulations and algorithm development for XRDT systems.	In process—COVID-19 closed the lab. Duke reopened it, and now we will complete this data acquisition.
Yes	Explosives database fusion	Successful incorporation into URI database. Single-format database available to interested parties from a single access point.	Complete
Yes	MCGPU test report	<10% error between simulation and experiment and time benchmarking versus GEANT4. Validated, rapid, user-friendly X-ray simulation tool available for the community.	Validation work is complete, but timing analysis is pending. Documentation and application to a use case of interest is in progress.
Yes	Classifier test report	Statistically significant performance improvement (in terms of Pe) relative to correlation-based classifier. Algorithm will be incorporated into the Smiths prototype XRDT scanner.	This work is in progress.

E. Final Results at Project Completion (Year 7)

- Year 7
  - XRD database of explosive materials, incorporated in the URI Explosives Database
  - Validated MCGPU simulation tool for simulating X-ray diffraction for real materials
  - New data processing approaches for X-ray diffraction identification and classification
  - Publication of research results
- Over the life of the program
  - First tabletop ED Laue diffractometer system built capable of investigating texture
  - First measurements of XRD texture in a broad range of stream of commerce materials/objects
  - First simulation framework for simulating XRD texture in a meaningful way
  - Demonstration of measurement configurations that mitigate the effect of texture
  - First application of nonlinear dimension reduction and machine learning algorithms for improved XRD-based classification
  - Demonstration of the value of combining coding and collimation for XRD system
  - Demonstration of measurement strategies for mitigating the effects of texture in material identification
  - Demonstration that texture can be a valuable feature for identifying materials

III. RELEVANCE AND TRANSITION

A. Relevance of Research to the DHS Enterprise

This work was directly relevant to the goals of DHS S&T and the TSA, which include seeking new techniques for improving airport security. An integral part of this effort included new methods for identifying threats through a combination of hardware and software, as well as developing tools for understanding and predicting the dependence of the measured signals on the material properties of an object. We proposed to develop these tools and extend the fundamental science behind XRDT systems. Our work benefitted ongoing and future DHS-funded efforts, such as:

- The BAA 13-05 programs at the University of Arizona, Duke University, and SureScan Corp. on information theoretic analysis of X-ray imaging systems
- The Rapiscan Labs, Smiths Detection, and HALO X-ray Technologies Ltd. programs on the development of checkpoint and checked baggage XRDT systems and associated algorithms
- LRBAA 14-02, LRBAA 18-01, and BAA 17-R-03 studies on related technology (e.g., hybrid XRDT and AT/CT scanners, next-generation detectors, and deep-learning classification algorithms)
- Other ALERT task orders (such as the Opioid Detection task order)
- Founding of Quadridox, Inc. to transition technology and expertise to a group capable of working at higher technology readiness levels and make use of sensitive information (e.g., Quadridox is participating in the Air Cargo task order)

The results of this work will continue to aid the TSA, TSL, and NIST in developing standards, calibration phantoms, simulants, and robust classification methods for XRD applications that have not previously been explored.

A.1. Facilitating Next-Generation Scanners

XRDT offers a way to reduce false alarms (Pfa) while maintaining detection (Pd) in baggage screening systems. This ALERT project is playing an important role in furthering the development of these XRDT systems by generating necessary insights and information to be used by OEMs as they develop specific commercial products. Reduced detection errors (Pe) at the checkpoint can save DHS millions of dollars annually while improving throughput, reducing divestment, and ensuring passenger safety.

- Affect XRD imaging system design/determine requirements on coding/sampling approach.
    - Metric: Provided Pd/Pfa improvement resulting from choice of illumination structure and coding/collimation (e.g., CAXI versus XDi system tradeoffs).
- Improve threat detection / provide an understanding of to what degree XRD systems can improve the ability to identify and localize threats. This will also provide a connection between current and next generation technologies (e.g., helping to connect AT/CT with XRDT capabilities and opportunities).
    - Metric: Identified how well materials that are difficult to resolve using transmission (AT/CT) can be identified using XRD.

A.2. Testing and Evaluation

The fast MCGPU simulation tools that we have developed and will refine through Year 7 will allow for the testing and evaluation of deployed and in-development AT, CT, and XRDT scanners. In addition, the XRD

database is essential for performing simulation, as XRD form factors (along with relevant variability) for complex materials can only be obtained empirically and are necessary for accurate simulation of transmission or scatter signals. These tools enable quantitative assessment of the possible performance of any X-ray-based scanner, thus helping determine the future roadmap for checkpoint solutions and component development.

- Creation of tools (the database, simulation, data processing) to enable government and OEMs to evaluate current and future technologies.
    - Metric: Material library contains scans of over 50 explosives and 100 stream of commerce materials (with >5 instantiations per material). MCGPU was shown to be 100 times faster than GEANT4 while providing nearly identical results.

A.3. Standards/Certification

Knowledge about the XRD form factors and inherent material variability can help agencies such as TSL and NIST develop standards, phantoms, and stimulants for evaluating the performance of and certification requirements for next-generation scanners.

- Stakeholders understand luggage from an XRD perspective (e.g., appreciate that confounders, interferents, and simulants are well-established for CT/AT systems, but wholly unknown in XRD) and develop solutions parallel to those in the CT/AT space.
    - Metric: Analyzed and disseminated XRD signature database, texturing taxonomy, and assessment of performance impact of material texturing. This information has been shared with the TSL and is informing some of their XRD testbed design and analysis.

B. Status of Transition at Project End

- The simulation tools and algorithms are currently in use by Smiths, Duke, and Quadridox on other projects (both for DHS and in the medical sector).
- The XRD database is available for use via the URI web interface but has been directly shared with and is being used by Rapiscan, Halo, and Smiths.
- Findings and insights about texturing and its prevalence in stream of commerce materials have been shared with TSL and are playing a role in their investigations of XRD for false alarm reduction.

C. Transition Pathway and Future Opportunities

- **Quadridox, Inc**.: Cofounded by Profs. Greenberg and Kapadia (as well as two other university colleagues), Quadridox is a university spin-out company whose focus is to transition the tools, technologies, and expertise generated at the university into real-world solutions for DHS. Quadridox's main purpose is to work with the government and OEMs to further technology in the security space. Quadridox is currently working with DHS S&T to build a prototype XRDT scanner and, as such, is seriously interested in using (and potentially commercializing) the tools generated in this ALERT program. In addition, Quadridox is developing a suite of commercial tools for virtual testing and evaluation of X-ray systems, and both the simulation tools and XRD databases developed in this ALERT program play a critical role in these activities. Thus, Quadridox represents a key transition vehicle for this ALERT program.

- **Related government contracts**: Drs. Greenberg and Kapadia are involved in other contracts with DHS through Duke University that rely and build directly on the results generated in this program. Examples include SureScan's BAA 13-05 project on fixed gantry CT analysis, Duke's BAA 14-02 project focused on design a hybrid AT-XRDT or CT-XRDT system, and Kromek's BAA 17-03 project focused on developing detector modules for coded aperture XRDT systems. We are also using the tools and methods developed here as part of the ALERT Opioid Detection Task Order to analyze the detectability of opioids for CBP applications. Work on this ALERT effort therefore benefits both DHS S&T and its partners and serves as a natural method for transitioning our results.
- **Direct interaction with major OEMs**: In addition to simply making our results available, we are actively working with several major OEMs to help them answer scientific and engineering questions that most interest them. Rapiscan, Quadridox/Smiths, and HALO are building prototype XRDT systems and have received data from and provided input to us throughout this project that target specific questions of interest.
- **Creation of knowledge products and training**: We documented our findings by publishing papers and books, attending conferences, and more. In addition, we are training students in the problems and techniques relevant to the overall DHS enterprise.

D. Customer Connections

- Kristofer Roe, Souleymane Diallo, Christopher Gregory, Matthew Mertzbacher, Smiths Detection (weekly communication, database shared, collaborating on Duke/Smiths joint BAA 14-02 program for transmission + XRDT study)
- Adam Grosser, Kris Iniewski, Redlen Tech (monthly communication, joint proposals in place to TSA, DHS S&T, OEMs)
- Simon Godber/Keith Rogers, Paul Evans, Halo Tech (monthly communication, database shared)
- Dan Strellis/Ed Franco, Rapiscan (collaboration under development, database shared)
- Robert Klueg, Ron Krauss, TSL (potential collaboration discussed)
- Jack Glover/Larry Hudson, NIST (potential collaboration discussed)
- David Lockley, DSTL (Monte Carlo code shared, annual licensing agreement in progress)
- Brian Harris, Kromek (discussion of detector specifications needed, based on findings herein)
- Alex DeMasi, contractor supporting TSL (shared texture results and ALERT ED Laue testbed design details)

IV. PROJECT ACCOMPLISHMENTS AND DOCUMENTATION

A. Education and Workforce Development Activities

1. Student Internship, Job, and/or Research Opportunities
    a. Summer research opportunities afforded to three high school students
    b. Hosted three REU students for the summer

B. Peer Reviewed Journal Articles

Pending –

1. Carpenter, J., Hazineh, D., & Greenberg, J.A. "Application of Energy-Dispersive Laue Diffractometry to Texture-Based Material Identification." *Journal of Applied Crystallography*, submitted.
2. Hazineh, D., & Greenberg, J.A. "Empirically-Informed Approach to Modeling Texture in an X-Ray Diffraction Tomography System." *Nuclear Instruments and Methods in Physics Research B*, submitted.
3. Hazineh, D., & Greenberg, J.A. "Coding and Collimation for X-ray Diffraction Tomography" *Applied Optics*, in preparation.
4. Japzon, M., & Kapadia, A. "Rapid Monte-Carlo Simulation of X-Ray Diffraction Imaging Systems Using MC-GPU." *Nuclear Instruments and Methods in Physics Research B*, in preparation.
5. Royse, C., Wolter, S., & Greenberg, J.A. "Application of Machine Learning to X-ray Diffraction-Based Explosives Detection." *NIMA*, in preparation.
6. Hazineh, D., & Greenberg, J.A. "Comparison of Direct and Coded Aperture X-Ray Diffraction Tomography in the Absence of Texturing." *Applied Optics*, in preparation.
7. Zhao, B., Yuan, S., Wolter, S., & Greenberg, J.A. "Application of Machine Learning to X-Ray Diffraction-Based Explosives Detection." *Machine Learning and Data Mining in Pattern Recognition*, in preparation.
8. Hazineh, D., Yue, J., & Greenberg, J.A. "Quantifying the Degree of Texture Using an Energy-Dispersive Laue Diffractometer." *NIMB*, in preparation.
9. Greenberg, J.A., Hazineh, D., & Gehm, M. "Energy-Angle Correlations in Energy-Dispersive Laue Diffraction." *Optics Express*, in preparation.
10. Hazineh, D., MacGibbon, C., & Greenberg, J.A. "Comparison of Direction and Coded Aperture X-Ray Diffraction Tomography in the Presence of Texturing." *Applied Optics*, in preparation.
11. Hazineh, D., MacGibbon, C., Kapadia, A.K., & Greenberg, J.A. "Simulation of Anisotropic X-Ray Diffraction Patterns Using Empirically-Measured Texture Masks." *NIMB*, in preparation.

C. Other Publications

Pending –

1. Greenberg, J.A., & Carpenter, J. "Application of X-Ray Diffraction to Explosives Detection." *Counterterrorist Detection Techniques of Explosives*, 2020, in preparation. Edited by Avi Kagan and Jimmie Oxley.

D. Peer Reviewed Conference Proceedings

1. Japzon, M., Nacouzi, N., Sharma, S., & Kapadia, A.J. "Optimization of X-Ray Diffraction Images of Medical Specimens by Monte Carlo Methods." *AAPM Annual Meeting 2019*, San Antonio, Texas, July 2019.
2. Kapadia, A.J., Stryker, S., McCall, S.J., & Greenberg, J.A. "The Role of X-Ray Diffraction Imaging in Digital Pathology." *SPIE Medical Imaging*, virtual, February 2020 [11320-37].

E. Student Theses or Dissertations Produced from This Project

1. Xiao, J. "X-Ray Diffraction Imaging for Breast Tissue Characterization." MS Thesis, Duke University, April 2020.

F. Databases

1. XRD diffractometer database incorporated into the URI Explosives Database (XRD signatures acquired in reflection mode at 8 keV for >400 materials, including explosives and benign materials). Continually updated. Shareable upon request
2. ED Laue database updated (angle- and energy-dependent XRD scatter signal in transmission mode for 510 measurements of >50 materials, including only benign materials commonly found in luggage). Continually updated. Shareable upon request.

G. Models

1. Updated Monte Carlo simulation in GEANT4 for texture-based modeling of X-ray diffraction (available to end users at the end of Year 7). Shareable upon request.

V. REFERENCES

[1] Shanks, N. E. L., & Bradley, A. L. W. Handbook of Checked Baggage Screening: Advanced Airport Security Operation. (John Wiley & Sons, 2005).

[2] G. Harding, M. Newton, & J. Kosanetzky, "Energy-dispersive X-ray diffraction tomography." Phys. Med. Biol, 35(1), 1990, 33.

[3] K. MacCabe, K. Krishnamurthy, A. Chawla, D. Marks, E. Samei, & D. Brady. "Pencil Beam Coded Aperture X-ray Scatter Imaging." Opt. Express 20, 2012, 16310-16320.

[4] S. O. Diallo, K. Tadlock, C. Gregory, S. Wolter, J. A. Greenberg, & K. Roe. "Towards an x-ray-based coded aperture diffraction system for bulk material identification." Proc. SPIE 10632, Anomaly Detection and Imaging with X-Rays (ADIX) III, 1063209 (27 April 2018).

[5] Edward Franco, Dan A. Strellis, Rapiscan Systems Labs. (United States); Kenneth P. MacCabe, Rapiscan Systems Ltd. (United States). "Application of coded aperture x-ray scatter imaging to checkpoint screening." SPIE ADIX 2018 (Orlando, FL).

[6] J.A. Greenberg, K. Krishnamurthy, & D. Brady. "Snapshot molecular imaging using coded energy-sensitive detection." Optics Express 21, 2548, 2013.

[7] Kosciesza, D. et al. in 2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC), 1-5.

[8] Harding, G. "Compact Multi-focus X-ray Source, X-ray Diffraction Imaging System, and Method for Fabricating Compact Multi-focus X-ray Source." U.S. Patent (2009).

[9] Send et al. "Application of a pnCCD for energy-dispersive Laue diffraction with ultra-hard X-rays." Journal of Applied Crystallography, 49(222) (2016).

[10] Venkatesh Sridhar, Sherman J. Kisner, Sondre Skatter, & Charles A. Bouman. "Model-Based Reconstruction for X-ray Diffraction Imaging." Proc. SPIE, vol. 9847, pp. 98470K-98470K-11, April 17, 2016.

[11] Ke Chen, & David A. Castañón. "Architectures and algorithms for x-ray diffraction imaging." Proc. SPIE 9020, Computational Imaging XII, 902006, March 7, 2014.

[12] Qian Gong, David Coccarelli, Razvan-Ionut Stoian, Joel Greenberg, Esteban Vera, & Michael Gehm. "Rapid GPU-based simulation of x-ray transmission, scatter, and phase measurements for threat detection systems." Proc. SPIE 9847, Anomaly Detection and Imaging with X-Rays (ADIX), 98470Q, May 12, 2016. DOI: 10.1117/12.2223244.

[13] Lakshmanan MN, Kapadia AJ, Sahbaee P, Wolter SD, Harrawood BP, Brady DJ, & Samei E. "An x-ray scatter system for material identification in cluttered objects: a Monte Carlo simulation study." Nuclear Inst. and Methods in Physics Research, B, Volume 335, 2014, pp. 31-38.

[14] https://www.fda.gov/about-fda/cdrh-offices/monte-carlo-simulation-x-ray-transport-gpu-cuda-mc-cpu

[15] Joel A. Greenberg, Chris MacGibbon, Dean Hazineh, Jesse Yue, Brian Keohane, & Scott Wolter. "The role of texturing in x-ray diffraction tomography." Proc. SPIE 1063210, Anomaly Detection and Imaging with X-Rays (ADIX) III (2018)

[16] D. Hazineh, & J. A. Greenberg. "Coding versus collimation in pencil-beam X-ray diffraction tomography." Proc. SPIE 10999-8, Anomaly Detection and Imaging with X-Rays (ADIX) IV (2019).

[17] C. Royse, S. D. Wolter, & J.A. Greenberg. "Emergence and distinction of classes in XRD data via machine learning." Proc. SPIE 10999-12, Anomaly Detection and Imaging with X-Rays (ADIX) IV (2019).

[18] D. Nacouzi. "Smarter Cancer Detection Through Neural Network Classification of High-Resolution X-ray Diffraction Tissue Scans." Master's Thesis, Duke Medical Physics Program, 2019.