Semester Report

Nuno C., Dr. Lev G.
Department of Materials Science and Engineering, The University of Texas at Dallas

December 18, 2022

Abstract

X-ray photoelectron spectroscopy (XPS) is perhaps the most ubiquitous method of surface characterization (barring use with light elements); providing quantitative insight into material composition, chemical state, and electronic structure. However, XPS spectra bears the detriment of being clouded by surface contaminants, and differences in instrumentation. As a result, human analysis of spectra can be very limited, if not purely unproductive. We train multiple machine-learning (ML) algorithms to simplify and guarantee the process of XPS analysis. In particular, using NIST's SESSA application as a generator for our training data-sets, we train algorithms for the purposes of quantitatively analyzing layer thickness, determining what material/s compose each layer, and for the novel application of inelastic-scattering background removal. This paper serves as a summary of the work done for MSEN 4V95 during the Fall '22 semester.

1 Introduction

XPS is a method that takes advantage of the photoelectric effect. As per its namesake, x-ray light impinges upon a reference material, exciting atoms, and leading to the ejection of photo-electrons. This procedure is performed in ultra high-vacuum ($< 10^{-7}$ [1]), as a means to increase the IMFP. Otherwise, scattering distorts kinetic-energy measurements. A fundamental limitation of this procedure is its inability to detect very light elements. In particular, XPS provides no usable data for hydrogen and helium.

XPS can identify most elements in a material, their relative composition and chemical state, and the electronic structure of the irradiated medium. In practice, XPS is used for any medium that can remain in a solid state; in particular, semiconductors [2, 3], thin films [4, 5], and polymers [6, 7] have a deep catalogue in tandem with XPS. We illustrate a prototypical XPS spectrum in Fig. 1. XPS spectra plot the electron counts vs. the binding energy of the incoming electron. This energy can be obtained from the Einstein equation [1]:

$$E_{binding} = h\nu - KE, \tag{1}$$

where $h\nu$ is the energy of the x-ray source, and KE is the kinetic energy of the incoming electron (which is measured by a spectrometer). Binding energy will vary with the type of atom/material, the net charge of the material, and the addition of other inter-molecular interactions. This means that different isotopes provide near identical XPS spectra. XPS, in practice, provides chemical information associated with the ionic/covalent bonds between atoms.

X-rays can penetrate very deep into a sample, having a maximum depth in the range of a few microns. However, the escape depth of electrons - due to the high density of inelastic-scattering collisions when traveling through the material - is only on the order of $\approx 10\ nm$. This contributes to a background signal throughout the spectra. The loss of electron energy due to scattering allows for calculation of sample thickness by Beer's law:

$$\begin{cases} I_f = I_0[1 - \exp(\frac{-d}{\lambda cos\theta})] & \textbf{Overlayer}. \\ I_f = I_0 \exp(\frac{-d}{\lambda cos\theta}) & \textbf{Substrate}, \end{cases} \tag{2}$$

where $I_{0,f}$ are the initial and final intensities, d is the thickness of the propagation medium, λ is the IMFP for the electron in that material, and θ is the angle by which the electron leaves the material. IMFP values can be calculated as a function of electron energy, and type of material [8].

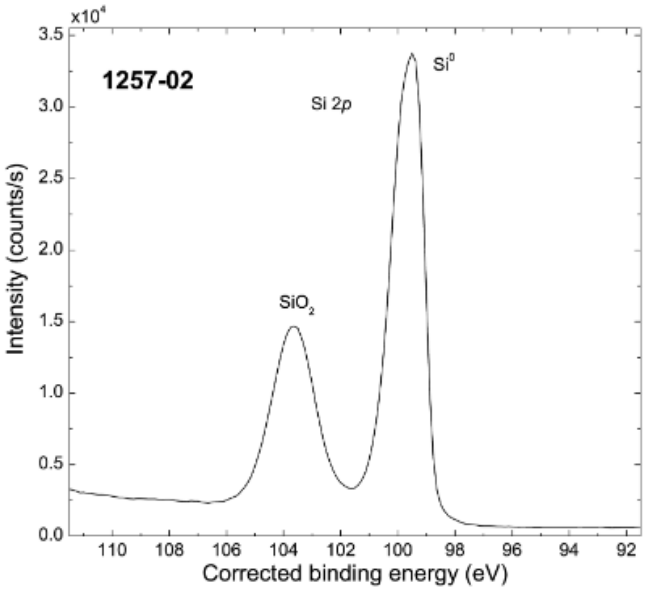


Figure 1: Spectra in the vicinity of Si 2p peak [9].

Further commentary is out of the scope of this report. However, the importance of XPS for material characterization can't be overstated. It is the most popular and widely available technique for surface analysis, and is able to provide a trove of chemical, electronic, and quantitative information about a given material. Because of this, we are motivated to pursue methods of interpreting XPS spectra accurately and efficiently. In particular, we investigate the use of ML for analysis of XPS spectra. For the remainder of this paper, we will refer to the GitHub repository for the resources, work done during the semester, and the particular code written for this task.

2 Methods

Training data-sets are produced with NIST's SESSA [10]. SESSA produces a data-set with immutable array dimensions. For a given spectrum, arrays are of size [2048, 2]; where the second dimension represents the energies in eV (index [:, 0]), and the counts for each energy (index [:, 1]). In practice, for each "variable" iterated over (material, thickness, both, etc.), one of these spectra is appended to an array with overall dimensions/sizes [(N,M,...), 2048, 2] where N, M, etc. are the # of parameter values for each variable. Take for example, an array generated for the layer-distinguishing algorithm summarized in sec. 3.2. Here, we iterate over 20 different material combinations at 40 Å (a single thickness). One can check the materials, dimensions, and parameters of this example in the output text file "Master_ref.txt", along with the actual array "Master_list.npy".

We perform this process with two differing scripts. In particular, "Parallel Spectra" generates the individual arrays, threading over each material iteration. "array creation" outputs a .npy array with the aforementioned dimensions for training/testing, based on the spectra generated from "Parallel Spectra". Both scripts have been commented on for an overview of what each line in the code is doing. Additional settings can be appended. In particular, one can change the number of surface collisions (with the hopes of simulating zero-scattering) along with allowing surface excitations. Furthermore, additional parameters can be modified to simulate differing instruments. We generate our spectra for a Si substrate with differing materials/thicknesses. The algorithms, as mentioned prior, can be divided into methods of distinguishing layer thickness, material composition, and removing inelastic-scattering backgrounds (although this is done in practice by simulating the minimum allowed number of particle collisions, which is 1 due to a bug with SESSA).

Notebooks with each prototypical algorithm can be found in the same repository. All training is done on neural network algorithms. We utilize the same base-layers for each model. In this case, we utilize 4 nonlinear activation layers followed by a linear output layer to replicate input dimensions with our output. This is illustrated in Fig. 2. The algorithm for distinguishing layer-composition was trained with cross-entropy loss, along with encoding to a matrix of probability densities for use in a classification task. The other two models were trained with mean-square error losses. This process is well-documented, with comments in the code present in the above link.

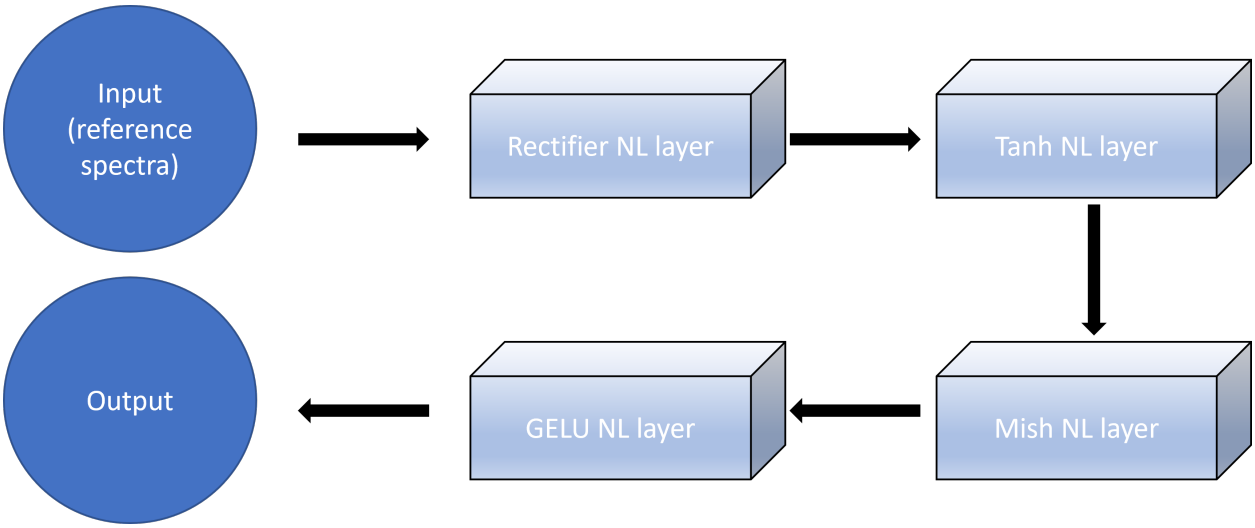


Figure 2: Trained model and respective layers. Inputs are spectra. Output may be affiliated thickness, material probabilities, or spectra sans inelastic-scattering background.

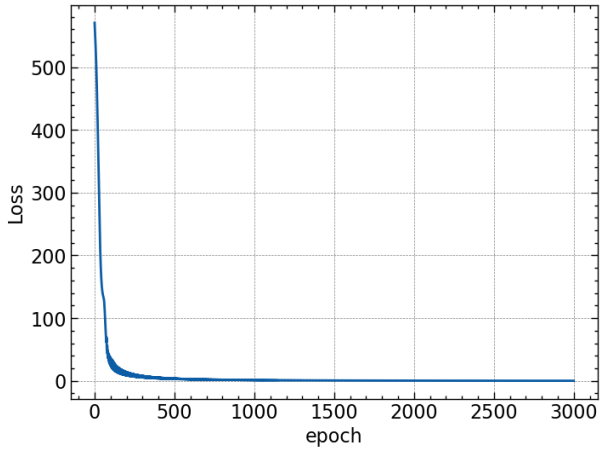
3 Model Performance

Models will be trained as indicated above. Afterwards, models are tested with training data as inputs (along with plotting of their respective losses) to indicate successful "base" training of the model. It is important to note that a few of these models fail when testing data even slightly extraneous to the training data-set. Examples of this will be shown below. Model performance is always at record-lows when input test data bears even negligible noise. As a result, in the future, training for the below algorithms shall be implemented both with "clean" data, as well as with spectra with Poisson-noise added.

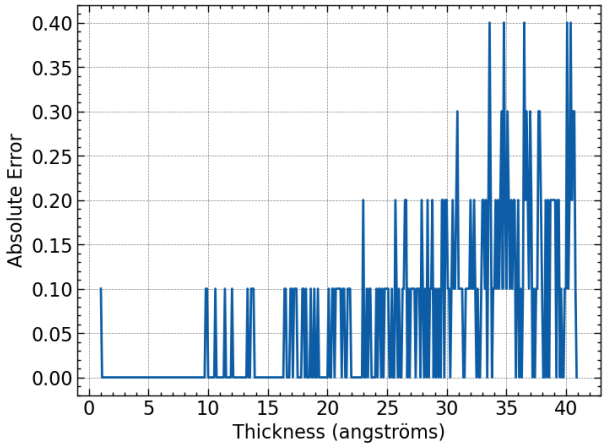
3.1 Distinguishing Material Thicknesses

We first trained an algorithm to distinguish the particular thickness of a spectrum **given** the actual material. This is perhaps the least useful, but easiest-to-implement model, due to it's high accuracy. For each material, our model is trained on an evenly-spaced set of material thicknesses. We perform this calculation on thicknesses between 1-40 Å for Al, with a 0.1 Å step-size. Loss and performance of this particular model can be seen in Fig. 3.

In practice, using thicknesses that surpass the maximum escape depth of electrons (>10 nm [11]) results in a higher absolute error (see Fig. 4), as there should be no difference between spectra at higher depths. In general, this was the best-performing algorithm, and test data within or slightly outside the training boundaries for thicknesses performed similarly well. I.e, this model is trained faithfully for each material, and doesn't require great changes (besides the addition of Poisson-noise to simulate the use of instrumentation). Needless to say, using a different material for testing will cloud the prediction model and nearly never provide accurate results.

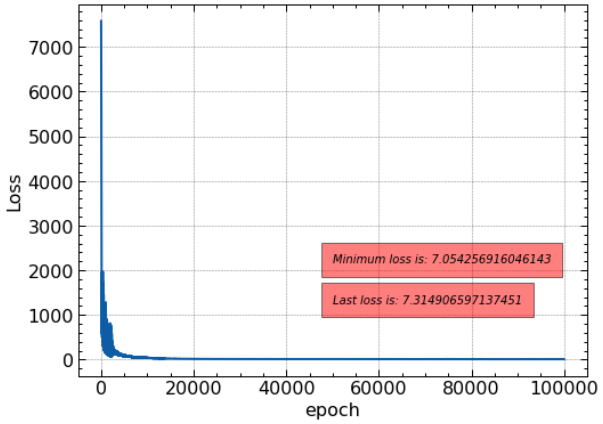


(a) Loss for thickness-distinguishing (1-40 Å) model as a function of training epoch.

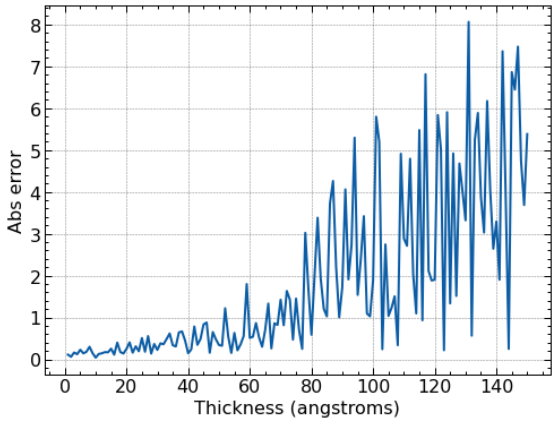


(b) Absolute error from model rounded to one significant figure.

Figure 3



(a) Loss for thickness-distinguishing (1-150 Å) model as a function of training epoch.



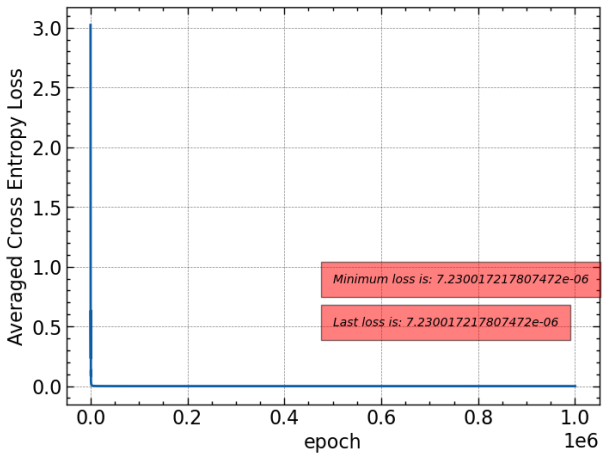
(b) Absolute error from model. Compounding error can be seen after XPS photo-electron depth-limit. This is likely due to spectra at thicknesses > 10 nm being functionally the same.

Figure 4

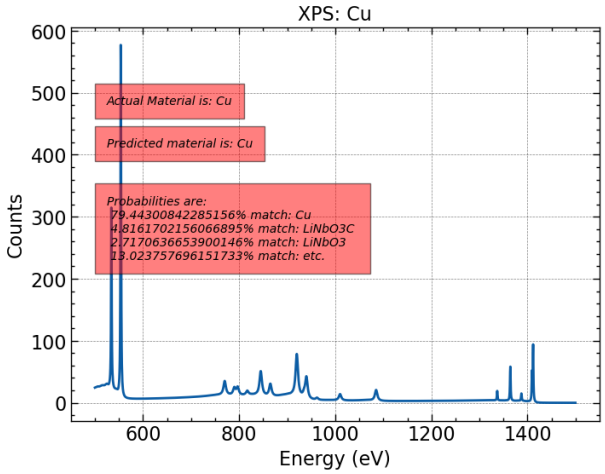
3.2 Distinguishing Layer Material

We also train a model to predict material compositions (with a Si substrate) at a particular thickness (in this case, 10 Å). One can of course use a myriad of thicknesses for this model. However we will use this model at a particular thickness to test inputs with differing thicknesses. This shall speak to how robust the model is with regards to the training thickness. Material/layer prediction can be done reasonably with human analysis upon inspection, but nevertheless there is some use to this model in both general cases (where a human could reasonably analyze the spectrum), and especially in fringe cases (where human analysis is limited). This task performs relatively well on the training data. We report 50-80% probability matches for the corresponding samples on training data, along with correct predictions for 100% of training scenarios (i.e, all materials are correctly identified). Probability matches for Cu (the best performing material) can be seen in Fig. 5.

Additionally, we also report correct predictions of the actual material for 70% of test data when thicknesses for test data are changed from 10 to 40 Å (e.g, Fig. 6). This implies a relative high fidelity if training



(a) Loss for material-distinguishing model as a function of epoch.



(b) Predictions on Cu (training data). Model successfully predicts the actual material with high certainty under the choice between 20 base classes/materials.

Figure 5

is performed on multiple thicknesses (even with few representative points in the training thicknesses). Additionally, in all cases with erroneous predictions, the second most likely match was the correct material. This result, in tandem with the results in section 3.1, implies a low-requisite resolution for training data with regards to input thicknesses. Since a class for materials outside the data-set has not been defined, this model (as expected) cannot provide predictions on extraneous materials.

3.3 Inelastic Scattering Removal

Perhaps most novel in implementation is the ability to remove scattering backgrounds from input spectra. Background removal is a highly useful methodology for quantitative analysis of spectra, allowing for the user to determine the composition of the material on an atomic level from the intensity of each peak [1]. There is an inherent ambiguity on subtraction of scattering backgrounds, with regards to both functions for subtracting, and choice of points of interpolation for this task [12]. We train a model with input reference spectra, and output background-removed spectra. One can see results for a representative material in Fig. 7. However, data is trained by setting SESSA's number of collision events to exactly 1. This is due to a bug with the original application, disallowing the choice of 0 scattering events. As such, the models provided serve only to attest to the effectiveness of a **future** algorithm to remove backgrounds. The high fidelity of the model below provides a hopeful indication of the ease with which a future algorithm (with 0 scattering collisions) would be implemented. One can see relevant differences in 1 vs. 2 collisions for Si in Fig. 8. Some materials bear no difference when transitioning from 1 to 2 scattering collisions, but some, like Si, bear new peaks due to this change in parameters.

This model is somewhat robust with regards to layer thickness, as are prior algorithms. Despite this, error is substantial at high-differences in thickness from training data. More substantial, however, is the error from differences in material. This is to be expected, and can only be dealt with by adding different materials to the training data. As a result, irregardless of the algorithm used, materials to be used for testing must be in the training data. However, it appears as if the use of ML models - given a large enough representative sample of spectra - can reasonably be used to remove scattering backgrounds from XPS spectra, even with the somewhat brute model illustrated in Fig. 2.

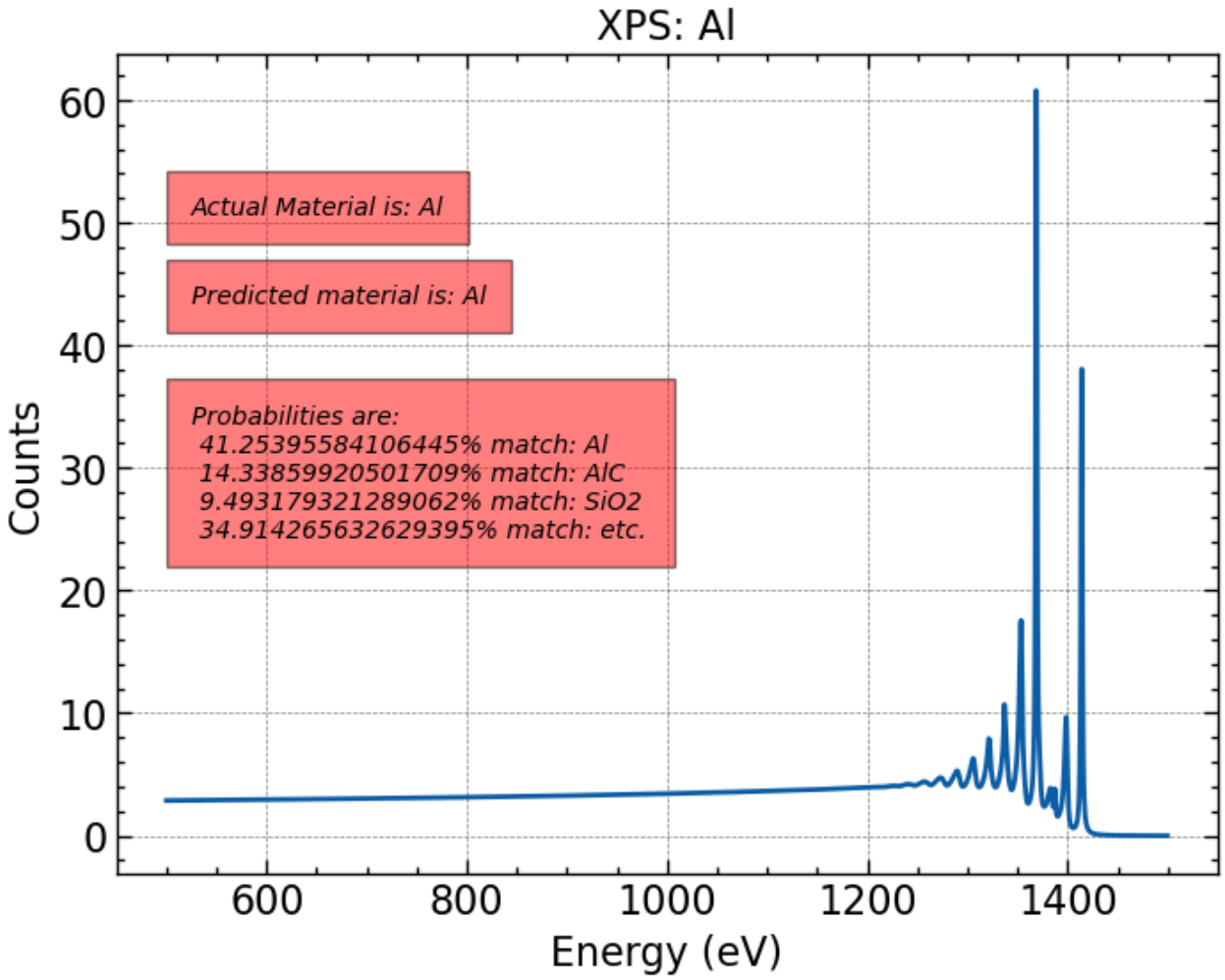
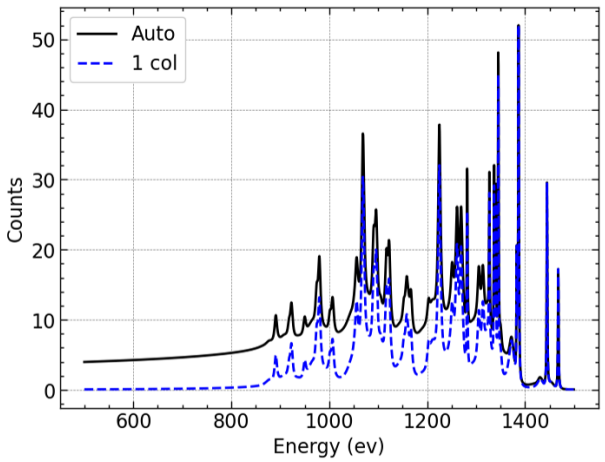
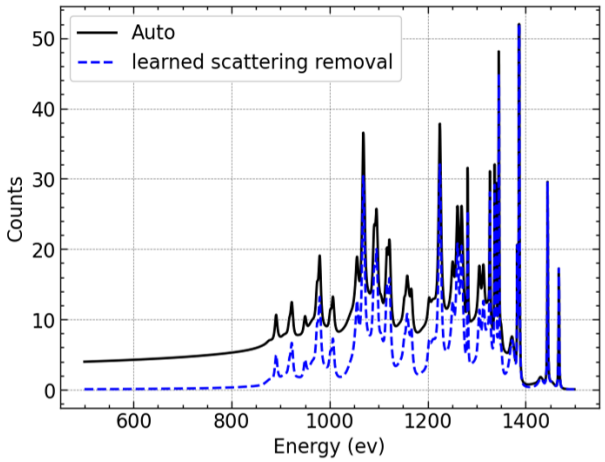


Figure 6: Material-prediction model trained on 10 Å data, tested with 40 Å Al spectra.

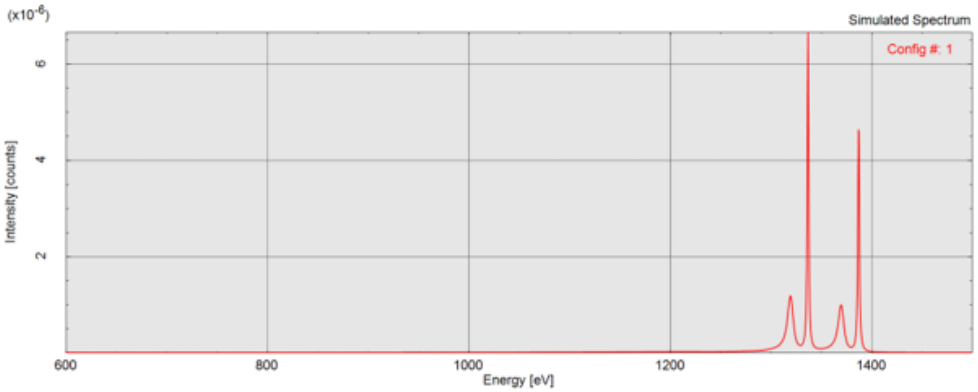


(a) Spectrum for 10 Å GaAs layer on Si, along with spectrum simulated with 1 inelastic-scattering collision.

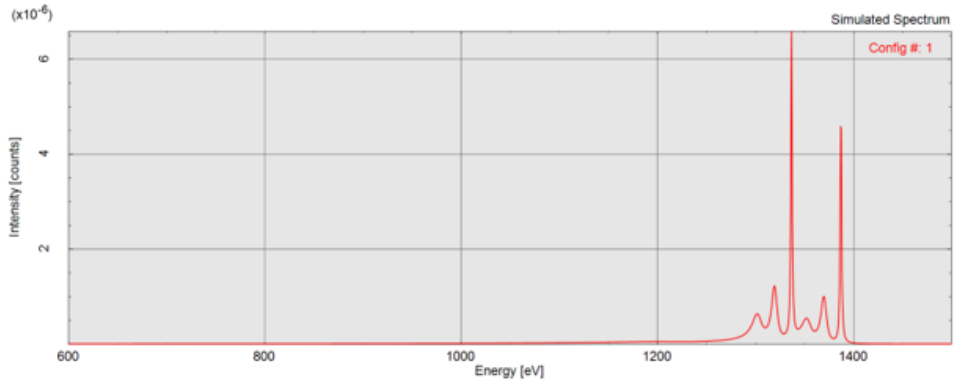


(b) Spectrum for 10 Å GaAs layer on Si, along with spectrum generated from learnt model (with training data as input).

Figure 7

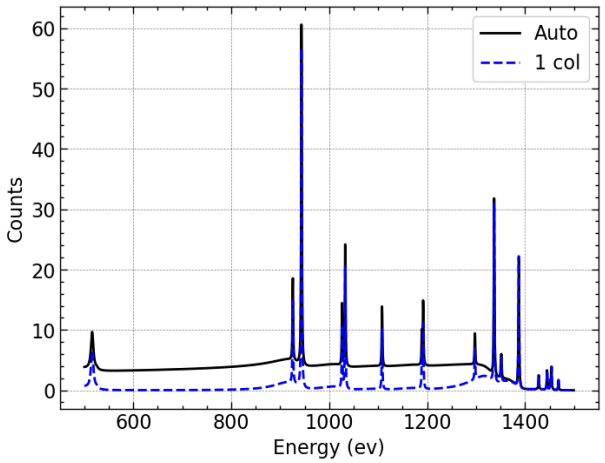


(a) Si on Si simulated with 1 scattering collision.

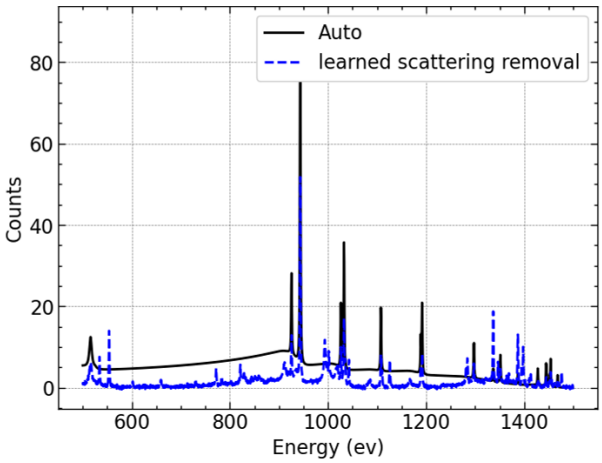


(b) Same as Fig. 8a, with 2 collisions.

Figure 8

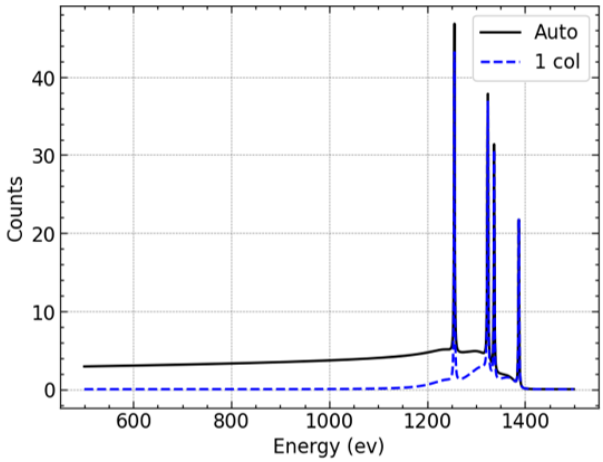


(a) Spectrum for 40 Å $KTiOPO_4$ layer on Si, along with spectrum simulated with 1 inelastic-scattering collision.

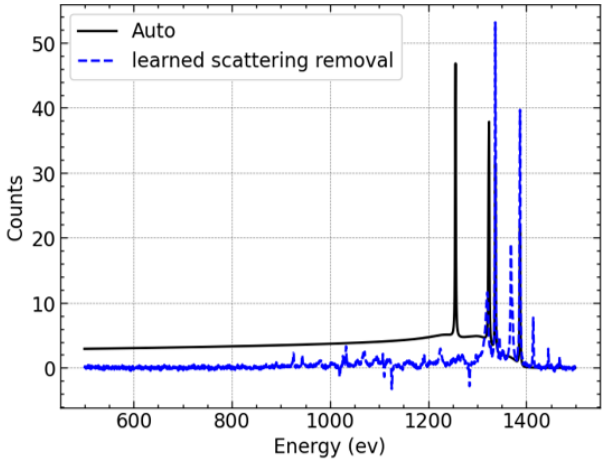


(b) Spectrum for 40 Å $KTiOPO_4$ layer on Si, along with spectrum generated from model trained on 10 Å training data. Error is present, but is likely to be mitigated greatly with more training data at different thicknesses.

Figure 9



(a) Spectrum for 10 Å S layer on Si, along with spectrum simulated with 1 inelastic-scattering collision.



(b) Spectrum for 10 Å S layer on Si, along with spectrum generated from model trained on 10 Å training data **without** S. Error is far greater than change in thickness. It's likely the model attempts to remove background for whatever material most closely matches the S spectrum.

Figure 10

4 Discussion and Outlook

We report on the possibility of the above algorithms for use in XPS analysis. Distinguishing layer thicknesses appears to be the easiest task from the above. This model likely only requires further training, and application of varied training data (Poisson-noise applied, materials varied, etc.) for proper use. Another way to implement this would be by using the model for distinguishing material composition, in tandem with a targeted, specific algorithm for distinguishing the thickness of that particular material. In practice, this method would likely be optimally fast, and would be relatively easy to implement. The model for distinguishing layer composition is also particularly robust towards thicknesses, and can likely also be trained with varied data to provide relatively accurate predictions of the materials with which it is trained. The background-removal algorithm is less robust towards a change in material and thickness, and is unfortunately only able to be trained (at the moment) with 1 scattering collision.

Future plans include accounting for the above deviations in spectra. In particular, each model should be trained on an indeterminate N materials (which should represent all materials with realistic use cases in XPS), at M thickness values (for which we have proven the models are relatively robust), with experimental considerations (instrumentation, noise, etc.). The latter is a problem less illustrated in the above figures. There is a lack of accuracy when input data is noisy, or when the data is changed due to instrumentation. Since realistic use bears these characteristics, it is integral to train the above models for these scenarios. Furthermore, a model for distinguishing materials on the basis of stoichiometry may perform better than the existing classification model, and if successful, can provide predictions outside of the existing training data with regards to material composition.

With these goals in mind, this project will be continued in the 2023 Spring semester, with the eventual hope of providing a novel method of predictive analysis on XPS data, and perhaps with implementation to a common-use package for the Python ecosystem.

References

[1] S. Tougaard, "Practical guide to the use of backgrounds in quantitative xps," *Journal of Vacuum Science & Technology A* **39**(1), p. 011201, 2021.

[2] C. Brundle, G. Conti, and P. Mack, "Xps and angle resolved xps, in the semiconductor industry: Characterization and metrology control of ultra-thin films," *Journal of Electron Spectroscopy and Related Phenomena* **178-179**, pp. 433–448, 2010. Trends in X-ray Photoelectron Spectroscopy of solids (theory, techniques and applications).

[3] F. A. Stevie and C. L. Donley, "Introduction to x-ray photoelectron spectroscopy," *Journal of Vacuum Science & Technology A* **38**(6), p. 063204, 2020.

[4] H. Bluhm, "4 - x-ray photoelectron spectroscopy (xps) for in situ characterization of thin film growth," in *In Situ Characterization of Thin Film Growth*, G. Koster and G. Rijnders, eds., *Woodhead Publishing Series in Electronic and Optical Materials*, pp. 75–98, Woodhead Publishing, 2011.

[5] H. Iwai, J. Hammond, and S. Tanuma, "Recent status of thin film analyses by xps," *Journal of Surface Analysis* **15**, pp. 264–270, 01 2009.

[6] E. D. Giglio, N. Ditaranto, and L. Sabbatini, *3. Polymer surface chemistry: Characterization by XPS*, pp. 73–112. De Gruyter, Berlin, Boston, 2014.

[7] G. Beamson and D. R. Briggs, "High resolution xps of organic polymers: The scienta esca300 database," 1992.

[8] M. P. Seah and W. A. Dench, "Quantitative electron spectroscopy of surfaces: A standard data base for electron inelastic mean free paths in solids," *Surface and Interface Analysis* **1**(1), pp. 2–11, 1979.

[9] D. S. Jensen, S. S. Kanyal, N. Madaan, M. A. Vail, A. E. Dadson, M. H. Engelhard, and M. R. Linford, "Silicon (100)/sio2 by xps," *Surface Science Spectra* **20**(1), pp. 36–42, 2013.

[10] W. S. M. Werner, W. Smekal, and C. J. Powell, "Simulation of electron spectra for surface analysis (SESSA) version 2.1 user's guide," tech. rep., Dec. 2017.

[11] B. D. Ratner and D. G. Castner, *Electron Spectroscopy for Chemical Analysis*, ch. 3, pp. 47–112. John Wiley Sons, Ltd, 2009.

[12] H. Tokutaka, N. Ishihara, K. Nishimori, S. Kishida, and T. Takabuchi, "The comparison of the background removal methods in xps spectra," *Japanese Journal of Applied Physics* **29**, p. 2512, nov 1990.