Title

Illuminating Heterogeneous Catalyst Encapsulation with Electron Energy Loss Spectroscopy and Unsupervised Machine Learning Methods

Permalink

https://escholarship.org/uc/item/2pb9j7p9

Author

Blum, Thomas Frederick

Publication Date

2020

UNIVERSITY OF CALIFORNIA,
IRVINE

Illuminating Heterogeneous Catalyst Encapsulation with Electron Energy Loss Spectroscopy and Unsupervised Machine Learning Methods

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Physics

by

Thomas Frederick Blum

Dissertation Committee:
Professor Xiaoqing Pan, Chair
Dr. Miaofang Chi
Assistant Professor Huolin Xin

2020

DEDICATION

To

my partner Hilary
without whom this thesis would not have been possible
and to my parents and friends who have provided their love and support

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

First and foremost I would like to express appreciation to my committee chair, Professor Xiaoqing Pan, who continuously pushed me to improve my work and strive for novel and impactful contributions to science and who shaped the ways I conducted research.

I would also like to thank my committee members: Dr. Miaofang Chi and Professor Huolin Xin. Dr. Chi supervised my work at Oak Ridge National Laboratory and oversaw the majority of the work presented in this thesis. Without her insight, guidance, and support this work would not have been possible. I am also grateful to Professor Xin, whose work inspired me to explore novel applications of machine learning in transmission electron microscopy.

I would like to thank Dr. Karren More, Director of the Center for Nanophase Materials Sciences at Oak Ridge National Laboratory, and Dr. Ramakrishnan Kannan of the Computer Science and Mathematics Division at Oak Ridge National Laboratory. Their support and guidance has been invaluable throughout my appointment at ORNL and in producing this research.

I would also like to thank those who contributed to the research reported in this thesis. In addition to the oversight, support, and expertise provided by Professor Pan, Dr. Chi, and Dr. Kannan, Jeffery Graves implemented the core algorithms in Python; Dr. Michael Zachman provided guidance in data analysis; and Dr. Felipe Polo-Garzon and Dr. Zili Wu provided the catalyst material used in this work.

Finally I want to thank my partner Hilary Blum, to whom this work is dedicated, for her patience, support, and editing expertise. I could not have done this without you.

This research was funded by a Lab Directed Research and Development project at Oak Ridge National Laboratory, a U.S. Department of Energy facility managed by UT-Battelle, LLC. Electron microscopy was conducted at the Center for Nanophase Materials Sciences, which is a DOE Office of Science User Facility. This research used resources of the Compute and Data Environment for Science (CADES) at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. This research is sponsored by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division, Catalysis Science Program.

ABSTRACT OF THE THESIS

Illuminating Heterogeneous Catalyst Encapsulation with Electron Energy Loss Spectroscopy and Unsupervised Machine Learning Methods by

Thomas Frederick Blum

Master of Science in Physics

University of California, Irvine, 2020

Professor Xiaoqing Pan, Chair

Encapsulating a permeable layer of oxide support on metal catalysts, as a type of strong metal-support interaction (SMSI), is often adapted to design heterogeneous catalysts with enhanced reactivity and catalytic activity. The success of such a design highly relies on the thickness of the encapsulation layer, often only one or two atomic layers, and its chemical composition and structure, which is highly delicate given its thickness and complex chemical environment. Precisely detecting such a trace layer and determining its chemistry however, is challenging. Scanning transmission electron microscopy (STEM), the most commonly used technique for such analysis, also suffers from a low signal strength due to the thickness and the electron beam sensitivity of the surface encapsulation layer. This can lead to the potential misinterpretation or overlooking of the encapsulation signal. Here, using Pd-TiO_2 as a prototype system, we develop and demonstrate an unsupervised machine learning method that allows us to reveal the presence and chemical information of the SMSI encapsulation layer that is otherwise hidden in STEM-electron energy loss spectroscopy (EELS) datasets. This method not only provides a robust tool for the analysis of trace SMSI in catalysts, but is generally applicable to any materials and spectroscopy datasets of any material systems where revealing a trace signal is critical.

INTRODUCTION

For decades machine learning algorithms have been widely used in numerous scientific fields, such as geospatial remote sensing, computer vision, and graph network analysis and have repeatedly proven their usefulness. Only recently, however, has the field of material science begun to incorporate these promising tools. To best make use of these techniques, these algorithms must be refined so that they are more applicable to materials applications, and it is important that the results the algorithms produce be interpretable with a physical meaning. Extracting trace signals from experimental data, which are not easily detectable by conventional data analysis, is a research topic that can significantly benefit from the integration of machine learning.

Heterogeneous catalysts are the subject of intense scientific interest. They are critical to the development of fuel cell, energy storage, and emissions control technologies, for example. There are multiple types of SMSIs including metal-support charge transfer, metal-support interphase layer formation, and the encapsulation metal-support interaction among others.[1] In particular, the encapsulation interaction shows promise as a potential mechanism for immobilizing oxide-supported nanoparticle catalysts and controlling their selectivity.[2–4] Reactions on heterogeneous catalysts occur at surfaces and are highly dependent on the atomic configuration of the surface, which can vary drastically between different crystal faces and interfaces between the catalyst and the support. These reactions are also greatly affected by the chemical bonding environment present at these surfaces. Understanding the encapsulation SMSI will lead to more control over crucial catalytic properties such as selectivity and activity.

Surface scientists have studied this phenomenon with scanning tunneling microscopy (STM) and atomic force microscopy and have provided important information about the surface states of the catalytic materials.[5] However these techniques cannot probe chemistry and structure simultaneously at the necessary spatial resolution, and often are limited to specific sample configurations. For example, scanning probe microscopes require extremely flat surfaces, and STM further requires that the support be conductive and under ultra high vacuum. Transmission electron microscopy (TEM) offers the flexibility to work with catalysts synthesized for use in actual chemistry and even perform in situ studies without significantly sacrificing spatial resolution.[6] TEM is also the only technique that allows examining the thin SMSI layer in the context of the structure of the small catalyst particles and their local oxide support.

Scanning transmission electron microscopy (STEM) combined with electron energy loss spectroscopy (EELS) has been the primary technique for investigating the structure and chemistry of encapsulation SMSIs in recent years. The high resolution of aberration-corrected (AC) STEM and the chemical sensitivity of EELS allows researchers to identify thin encapsulation layers, differentiating between carbon and thin oxide layers. One of the limiting factors in the investigation of SMSIs with STEM and EELS, however, is the potential for complicated datasets where the EELS spectral signatures overlap or the signal strength is low, and thus cannot be analyzed using standard techniques. Traditional EELS analysis is also ineffective at displaying the spatial distribution of energy-loss near-edge structure (ELNES) changes, which can be used to gather information about the nature of the oxidation state of the SMSI layer. In one form of traditional EELS analysis the chemical composition is examined by integrating the scattered intensity of the EELS edge after the removal of the pre-edge background. For *L* edges, the white line ratio may also be characterized for information about the chemical composition, but background removal requires spectrum fitting and quantitative results require Fourier-ratio deconvolution of the zero-loss peak (ZLP) which often increases the weight of noise in the dataset hindering sensible fitting.

It is widely believed that the chemistry of the encapsulation layer, and the transition metal in the encapsulation layer, if present, may have a different oxidation state than that in the bulk support. The encapsulation layer is very thin and its composition and chemical characteristics vary significantly among different metal-oxide pairs, their treatment conditions, as well as the surface structure of the encapsulated metal catalyst particle and its atomic species.[7] The spatial distribution of the encapsulation layer is important for understanding the active sites of the catalyst, and the local bonding configurations that can be revealed by EELS carry valuable information in understanding catalytic reaction mechanisms. The thin encapsulation layer produces a weak signal in EELS that, because of the necessity for a low electron dosage to avoid significant damage, can be difficult to detect. This leads to low signal intensities which make fitting algorithms impractical. In addition, the imaging conditions can vary between microscopes, affecting the spectral resolution and the spectral response. While thickness variations can be approximately removed by Fourier-ratio deconvolution, microscope variations effectively require that reference spectra be taken under the same imaging conditions as the region of interest for an accurate comparison. Moreover, core-hole induced multi-electron interactions make accurate theoretical modeling of L_{23} spectra challenging.[8]

While not a panacea, machine learning can help with some of these challenges. Unsupervised machine learning techniques can help characterize changes in the EELS spectrum but do not require reference information as the spectrum is not compared to models or external empirical examples. It is instead self-comparative, in that the spectral data is only compared to itself. This is useful as this technique does not require the use of a model or a reference, and therefore can be used to compare spectra that do not have a model or reference or spectra for materials where available models are not useful or inapplicable. Another key benefit is that this algorithm requires significantly less operator input than other methods, thus reducing the risk for operator bias. Here, we demonstrate the advancement of unsupervised machine learning methods to study encapsulation phenomena, one SMSI in heterogeneous catalysis, using palladium supported on titanium dioxide (Pd/TiO_2) as a model system.

Pd on TiO_2 with an encapsulation SMSI serves as an ideal model heterogeneous catalyst system to explore the capabilities of our new machine learning algorithm. While not its only application, Pd/TiO_2 with an encapsulation SMSI is used in the hydrogenation of carbon dioxide to form methanol[9], an important reaction for the production of alternative fuels. In addition to its importance in chemistry, this model system combines several of the complex problems faced in the context of EELS spectrum image unmixing including overlapping edges, changes in fine structure, and weak signals, which will be described in detail in the results section. It is also ideal because the process for generating an SMSI in this material system has been demonstrated consistently and the preparation method can be verified by CO adsorption measurements.

Multivariate machine learning algorithms can be used to deconstruct a spectrum image and retrieve a defined set of spectral signatures, or components, that are present in the spectrum image as well as their spatial distribution. Among the most popular algorithms are principle component analysis (PCA) and non-negative matrix factorization (NMF). PCA produces abstracted representations of spectra because of the orthogonality requirement of the algorithm, which is not necessarily reflected in the spectra. Each physical component is then a linear combination of abstract spectral components which are difficult to interpret individually[10]. NMF, however, has been shown to produce more interpretable results with its requirement that the components and abundance maps be positive.[11] NMF suffers though from a sensitivity to local minima and its minimization function is under-defined so that mathematically, it can produce many decompositions with the same numerical error. More constraints are, therefore, needed to produce a set of components that more uniquely fit the data.

In this work, we demonstrate the advantage of using machine learning data analytics in the study of SMSIs. We introduce a new algorithm, automatic target generation process (ATGP) pre-conditioned Joint-NMF, that leverages de-noising and clustering properties to recover electron energy loss spectrum image information that would otherwise be obscured by noise. This new algorithm combines two existing algorithms, ATGP and Joint-NMF. The results from this new algorithm are compared with those from scikit Non-Negative Matrix Factorization (NMF) and Principle Component Analysis (PCA) and show significant advantages over these pre-existing algorithms. These advantages include the removal of the orthogonality condition while retaining non-negativity, and utilization of a pre-conditioning step to aid in convergence, which are important to the analysis of STEM-EELS spectrum imaging datasets with weak and complex signals, such as the encapsulation strong metal-support interaction in the Pd/TiO_2 system.

CHAPTER 1:
Background

Scanning Transmission Electron Microscopy

The physics of electron scattering off of atomic potentials is at the core of electron microscopy. In STEM there are many different ways of forming an image, each providing unique information about the material being imaged. These techniques are largely grouped into two types of physical interactions: elastic electron scattering and inelastic electron scattering.

One of the most common STEM imaging techniques is high-angle annular dark field imaging, also called Z-contrast imaging. In this method, a beam of electrons is focused to a point and is generally raster-scanned across a region to be examined. An annular detector is placed around the transmitted electron beam in order to intercept elastically scattered electrons that have been scattered to a relatively high solid angle. The large scattering angle of these electrons can generally be attributed to electrons that have passed very close to the nucleus of a given atom in the sample. Thus, the scattering of these electrons closely follows that of Rutherford scattering of an electron off an unscreened nuclear potential. High-angle annular dark field imaging is also called Z-contrast imaging because of the relationship between the intensity of the scattered electrons and the atomic number of the material being imaged. Depending on the electron beam energy and the convergence angle of the electron beam, the number of electrons scattered to a high angle is roughly proportional to the atomic number of the atoms squared, referring to the atoms with which the beam is interacting. It is also roughly proportional to the thickness of the sample, or is at least a monotonically increasing function of the thickness of the sample. Bright-field imaging is a form of phase contrast imaging where elastically scattered electrons interact with un-scattered electrons coherently on a circular detector. All of the electrons on this detector are integrated into one signal and the image we see is a result of the interference of the scattered and un-scattered electrons. This phase contrast is particularly useful for imaging thin materials with low atomic numbers and in general when ignoring interference effects, the thicker and denser the material is the darker the image will be in that area.

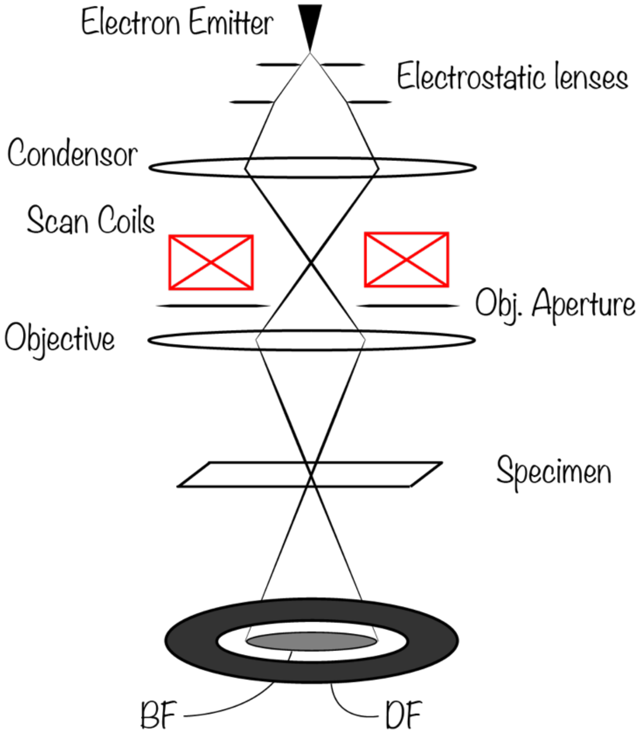


Figure 1:
Schematic diagram illustrating the beam-forming electron-optics and electron detectors in a STEM.

Z-Contrast STEM imaging was introduced by Crewe in 1970 and 1975 and was re-addressed by Pennycook in 1988 and Yamashita in 2018 and was recently reviewed by Treacy in 2011. The principle comes from the expected scattering amplitude of an electron by an unscreened atomic nucleus, or Rutherford scattering, which is proportional to the atomic number Z, an expression for which is given below. In high-angle annular dark-field imaging we detect the intensity of this scattering rather than the amplitude so the contrast is approximately proportional to the square of Z. In practice, the exponent varies depending on the detector geometry and the accelerating potential of the microscope in addition to the effect of electron cloud screening. This mode of imaging is an incoherent mode of imaging as outlined in the textbook from Pennycook *et. al.* in 2011 and thus, unlike CTEM and the BF STEM, it does not show a contrast reversal with changing defocus, and thus produces an image like that of a traditional camera. It is for these reasons that Z-Contrast STEM imaging has become one of the most popular methods in high-resolution electron microscopy. Mathematically these differences come about from two different ways to convolve the probe function of the microscope with the response function of the sample. If $\varphi (R)$ is the response function of the sample and $P (R)$ is the probe function of the microscope, then the expected intensity of the detected signal for bright-field $I_{BF} (R)$ and the dark-field $I_{DF} (R)$ are as follows:

$$I_{BF} (R) = |\varphi (R) \otimes P (R)|^2 \qquad I_{DF} (R) = |\varphi (R)|^2 \otimes | P (R)|^2$$

Thus I_{BF} is a coherent interaction between the probe and the sample, so the image inherently contains phase contrast but is not directly interpretable. Conversely, I_{DF} is the intensity of the dark field image and is a convolution of the intensity of the phase function and the intensity of the probe. The resulting image is then a result of an incoherent interaction between the beam and the sample and is thus directly interpretable. A useful derivation of these relationships is given in the text *Scanning Transmission Electron Microscopy* by Pennycook *et. al.* in 2011.

In a scanning transmission electron microscope the resolution is primarily dependent on three different factors: the image diameter of the electron source just after the gun lens, the diffraction limit from the accelerating potential, and spherical aberration in the electromagnetic lenses. The diffraction limit is simply the Rayleigh criterion $d = 1.22\ \lambda/\alpha$. The probe size, or gun image diameter, depends on several factors: the emission current, the brightness factor β, and the divergence angle α. In practice, we do not calculate the beam diameter but instead use imperial methods to measure it—either by adjusting the lenses to project an image of the beam intensity onto a camera or by using a knife edge where the edge is atomically sharp and there is full transmission of the beam on one side and full blocking of the beam just after the edge in the ideal case.

The improvement of the electron microscope's resolving power has been a longstanding goal in the physics and materials science fields. Richard Feynman, in his well-known address, "There is Plenty of Room at the Bottom," emphasized "the importance improving the electron microscope by a hundred times"[12] Since then, scientists have continued this work and have made several advancements that have revolutionized the field of materials study, especially at the nanometer and sub-nanometer level.

Recent advances in the design and control of electron lenses[13–15] has enabled improvements in the resolving power of the STEM,[16] and it is now common to see atomic resolution images. The first of these was the development of electron ptychography[17–21] and holography[22,23] which uses the principles of electron interference to retrieve the phase component of transmitted electron beam having been affected by the thin specimen under observation. With the ability to determine both the phase and amplitude of the transmitted beam rather than just the recorded intensity, the ability to compute the precise spherical aberration coefficients of the electron lenses would allow an increase in the resolving power of transmission electron microscopes.[16] Further development of electron lens technology has also produced pre- and post-specimen spherical aberration correctors which have enabled straight forward atomic resolution first in parallel beam transmission electron microscopes then in converged beam scanning transmission electron microscopes.

Electron Energy Loss Spectroscopy

Developments in STEM lens technology have benefited other measurement techniques such as electron energy loss spectroscopy and energy dispersive X-ray spectroscopy allowing atomic resolution chemically sensitive spectrum imaging.[24] The use of inelastically scattered electrons is another essential imaging technique. This method reveals more information about the chemistry and electronic structure and properties of the material being imaged.

In understanding the way a beam of electrons interacts with a material, it is important to remember that interactions at an atomic scale can have both local/short range interactions and long range collective interactions. Dark-field STEM imaging makes use of the short range interactions between electrons and nuclei. EELS, however, is dominated by both short range electron-electron interactions and intermediate to long range electron-electron cloud interactions. Thus, EELS provides information about the different elements that may be present in a given material and where they are and it also can provide information about local bonding behavior and the behavior of an electron gas in the material.

In a scanning transmission electron microscope the electrons in the beam have predominantly a single energy in the range of 60-300 keV. The spread of the illumination energy is usually between 0.6-2 eV. When one of these high energy electrons passes through the sample it can interact with one of the electrons around a given atom and excite that electron to an empty higher energy state. The amount of energy that the incident electron loses as it excites this electron is the signal that we record.

These excitations have strict transition rules and the available transitions reflect information about the elemental bonding with surrounding elements and the local electronic structure. This can give us information about the chemical and electronic properties of the material that we are investigating.

The amount of energy lost by an electron as it passes through a material generally separates the signals into different spectral regions, each of which tells us something different about the material. The core-loss region is generally above 50eV and represents energy lost due to the excitation of an electron from the core levels of an atom into its available energy levels. Typically, these kinds of interactions are used to identify chemicals by their atomic number. As the energy decreases, below 50eV, we see more interaction from valence electrons which are easier to excite, and this reveals information about the electronic structure and states in the material. This region also reflects electronic resonance in the valence electrons, such as surface plasmon-polaritons and bulk plasmons. Below 1eV, we start to see interactions analogous to optical IR spectroscopy which reveal information about the sample related to the strength of the chemical bonds as indicated by vibrations in the atomic structure. These low-energy excitations have only recently been enabled by the improvement in electron mono-chromation with the best energy resolution as measured by full-width half-maximum now at approximately 5meV.[25]

Core-loss EELS is generally used to track chemical concentrations in materials. The shape of the core-loss edge close to its onset is referred to as the near-edge structure. It generally arises as a result of bonds between different atoms, either of the same atomic species or different atomic species, and acts as a kind of fingerprint that can give us information about the local bonding environment of the atoms present. The most well-known atomic energy approximations are the hydrogenic model and the Hartree-Slater model which both assume single atoms without the presence of any bonding. In order to explain the structure at the low end/beginning of the energy-loss edge, we have to examine the change in the electronic structure that results from a chemical bond and allowed electronic transitions from occupied states to unoccupied states. Bonding forms hybrid orbitals between atoms and this structure can be further affected by these bonds being in a repeating structure such as an atomic lattice. These available states are affected by these kinds of local changes in the bonding environment. The theory required to accurately describe this behavior is still an active field of investigation and accurate simulations of this behavior are still rare. The most useful theoretical frameworks that we have to describe this kind of bonding and its effect on electron energy loss spectra is multiplet theory.[8,26–28]

Though the methods described later in the paper are applicable to spectrum imaging at all of these different ranges in the electron energy-loss spectrum, and even beyond electron spectroscopy to other forms of spectroscopy, in this thesis I will focus on energy loss near-edge structure which exists in the core-loss range, and specifically I will focus on the transition metal titanium and its L_{23} edge.

Interaction of Fast-Electrons with Solids

EELS uses the interaction of a relativistic electron with the atomic potentials in a given material. When a fast-electron scatters off of an electron in a given state, it can excite that electron into an empty state according to the transition rules governed by the Pauli exclusion principle. This scattering probability can be represented by the following double differential expression,

$$\frac{d^2\sigma}{d\,\Omega\,dE} = \frac{1}{\pi\varepsilon_0}\left(\frac{Z\,m_{red}\,e^2}{\hbar^2}\right)^2\left(\frac{k'}{k}\right)\frac{1}{q^4}\times\left|\sum_{j=1}^{N}\langle f|e^{i\,\boldsymbol{q}\cdot\boldsymbol{r}_j}|i\rangle\right|^2\rho(E)$$

Where **k** is the momentum vector of the incident electron and **k**' is the scattered electron momentum vector and **q** is the transferred momentum imparted on the excited electron. Z is the atomic number, m_{red} is the reduced mass of the incident electron, and rho is the density of states function.

We separate the electrons with given energies with a magnetic prism. A simplified schematic of the electron spectrometer is given below.

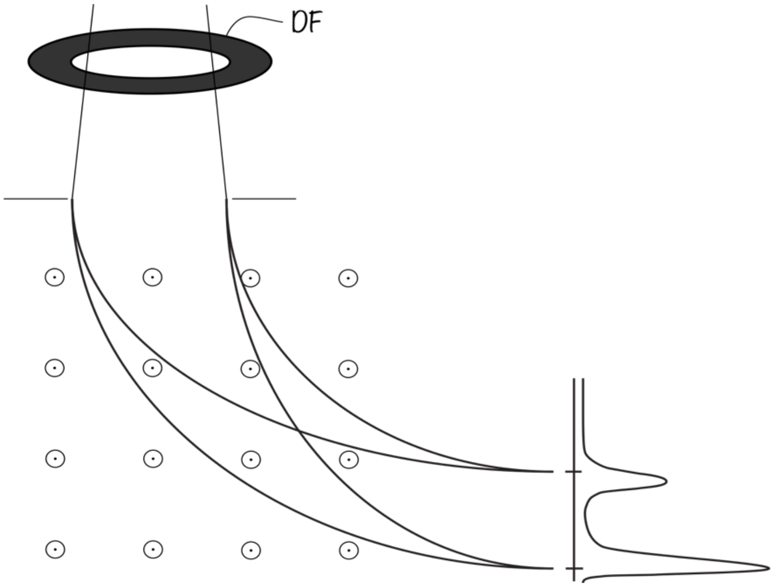


Figure 2:
Schematic diagram illustrating the working principle of the magnetic prism in an electron energy-loss spectrometer.

The electrons enter the spectrometer at the entrance aperture and then pass through a mostly uniform magnetic field inside a tube called the drift tube. The electrons' paths are bent by the Lorentz force, and because they all have the same mass we can select and record an electron flux as a function of kinetic energy according to the position in a plane where the focal point has been dispersed by the different radii of curvature of their respective paths. This much is common to all the commonly used transmission electron spectrometers but there are a few different types which differ by how many channels are in the detector. Essentially there are two types of spectrometers. There are serial spectrometers in which the detector is scanned through the various energy loss channels. Then there are the parallel spectrometers which measure several channels of energy loss at one time. In this work all the data has been acquired with parallel spectrometers which use a cooled scintillator and camera to detect the incident electrons.

CHAPTER 2:

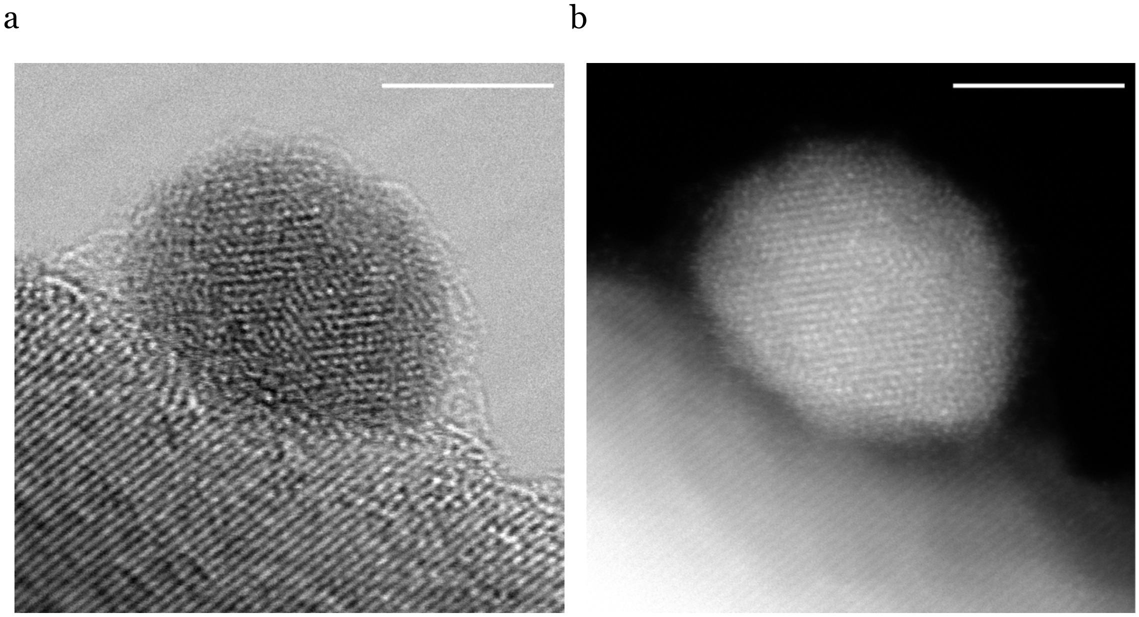


Figure 3:

Scanning Transmission Electron micrographs of Pd nanoparticle on TiO2 support. a.) Bright-field image showing thin encapsulation layer. b.) Simultaneous dark-field Z-contrast image. Note the lack of sufficient contrast to identify the light element encapsulation layer. Scale Bar: 5nm

Figure 3 shows bright-field and Z-contrast AC-STEM images of a typical Pd nanoparticle (~10nm) on a TiO_2 support, as used in this study. While AC-STEM is well known for producing atomic resolution Z-contrast dark-field images, it is not always sufficient for analysis of an encapsulation layer. In Figure 1, though what appears to be an encapsulation structure is visible in the bright-field image, bright-field imaging cannot provide any information about the chemical structure. The dark-field Z-contrast image has chemical information but the thin encapsulation layer is not resolvable. EELS spectrum imaging, however, should be able to provide both the required resolution and chemical sensitivity. Further, the Ti $L_{2,3}$ edge and the O K edge ELNES both exhibit changes encoding information on the local coordination of the O and Ti bonds. Variations in this near edge structure can be caused by a change in structural symmetry, as exhibited by different Magnelli phases of Ti oxides such as anatase and rutile.[29,30] Changes can also result from an altered Ti oxidation state, which in the case of an encapsulation SMSI of Pd/TiO_2 has been used to indicate the growth of TiO_x layers on the surface of the Pd particles. Such TiO_x layers are often one or two monolayers in thickness and thus produce a very small signal in EELS compared to the TiO_2 support or the Pd nanoparticle.[31]

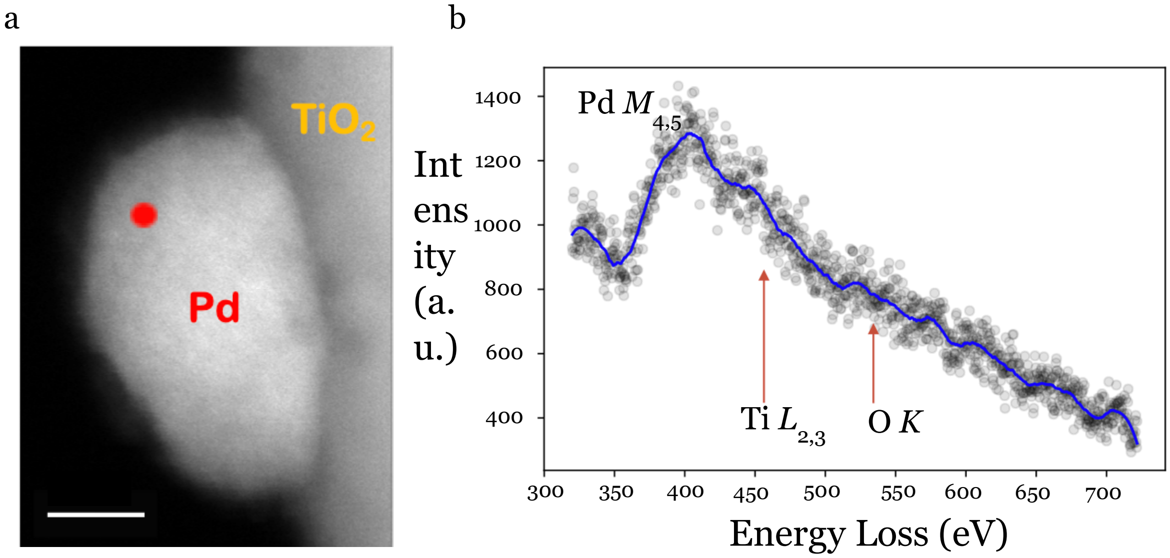


Figure 4:

EELS Spectrum (b) from single point showing a strong Pd signal and approximate locations for Ti and O signals. HAADF image (a) shows location from which the point was taken. The convergence semi-angle was about 30.0 mrad and the pixel exposure time was 0.1 s . The EELS acceptance semi-angle was about 48.0 mrad. Scale Bar: 5 nm

Figure 4 shows a typical single point EELS spectrum from a spectrum imaging dataset. The Pd $M_{4,5}$ edge is clearly seen while the Ti $L_{2,3}$ edge and O K edge are practically invisible on the back of the Pd edge. These edges are essentially indistinguishable from the noise, but their presence is suggested after applying heavy spectral smoothing. This is an indication that the signal is present but obscured by Poisson statistical noise. The noise level and Pd edge distortion of the power-law in the pre Ti-edge region and the small Ti signal drastically increases the error of power-law background subtracted integration methods. For these reasons, we cannot use model fitting. For example, Figure S1 shows the extraction of elemental maps from a typical EELS-SI dataset using the traditional data analysis method, power-law background subtraction with no local averaging, provided in GMS (Digital Micrograph).[24] Because of the limited EELS signal on the particle surface here, even the background subtraction introduced too many artifacts for the analysis to be reliable. As shown, the Ti *L* signal was only resolved on the TiO_2 support, while the O map shows a signal on the particle but is comparable to the noise level. Machine learning methods adapted in this work are designed to address these challenges, enabling analysis where noise and signal strength cause issues for standard techniques.

CHAPTER 3:

Data Pre-Treatment: Skipping Conventional Background Subtraction and Applying a Single Input Parameter

Generally in the analysis of EELS spectra, pre-edge background subtraction is a standard procedure, after any known outliers such as X-ray or gamma-ray spikes in the spectrum image have been removed. Here we can depart from this convention. This is because while power-law background subtraction can be beneficial as a way to limit the analysis to the known elements and signals, the presence and close proximity of other EELS edges, such as the Pd and Ti edges mentioned above, can make the power-law approximation invalid for the energy loss range in front of the edge of interest. Also, in cases where the beam intensity and dwell time are small due to dose limitations, backgrounds can be challenging to fit accurately and stably, and thus background subtraction methods can introduce artifacts. Furthermore, in SIs with the presence of a vacuum interface, the datasets include an extremely small but crucial interface signal within the spectral range of interest—at or near the noise level of the detection camera — rendering background subtraction methods ineffective because of instability in the background fitting and the mixing of the Gaussian camera noise and Poissonian detection noise. Not removing the background leaves open the possibility of the decomposition component signals having values below the background which maintains the overall positivity of the reconstructed signals but breaks the constraint that the components themselves be positive in the same way that the signals we are detecting are positive.

The machine learning algorithms we are using are unsupervised, and thus no pre-existing labels are needed. Here, in order to allow for a reasonable computation time, a single parameter was taken as an input. Here, we selected the number of components (k) as the input parameter, which is determined using PCA. From a plot of the explained variance ratio produced by PCA shown in Figure S2 , using the elbow method,[32] we determined k to be 6.

CHAPTER 4:
Applying Existing NMF Machine Learning To EELS-SI Datasets

The NMF algorithm, was first applied to the analysis of the SMSI datasets. NMF has been recently used for analysis of EELS[33,34] and EDS SI datasets,[33–35] focusing on differentiating overlapping signals. Its ability to reveal trace signals in spectroscopic datasets, however, has not yet been reported. In Figure 5, we show the decomposition from a standard NMF algorithm[36] which features non-rigid orthogonality constraints applied to the full range of the spectrum image. Per common practice, the data analyzed through NMF was not scaled and X-ray and gamma-ray spikes were removed. Figure 5 shows the result of analyzing the full spectrum image using NMF given an input of 6 components. The default implementation of NMF in sklearn includes assumptions of orthogonality and sparsity which can introduce artifacts in the decomposition results. These artifacts present as dips in the spectral components. Examples of these artifacts particularly related to sparsity have been annotated in Figure 5 with colored bars. This is analogous to the way PCA spectral components include negative values in order to satisfy the orthogonality condition.

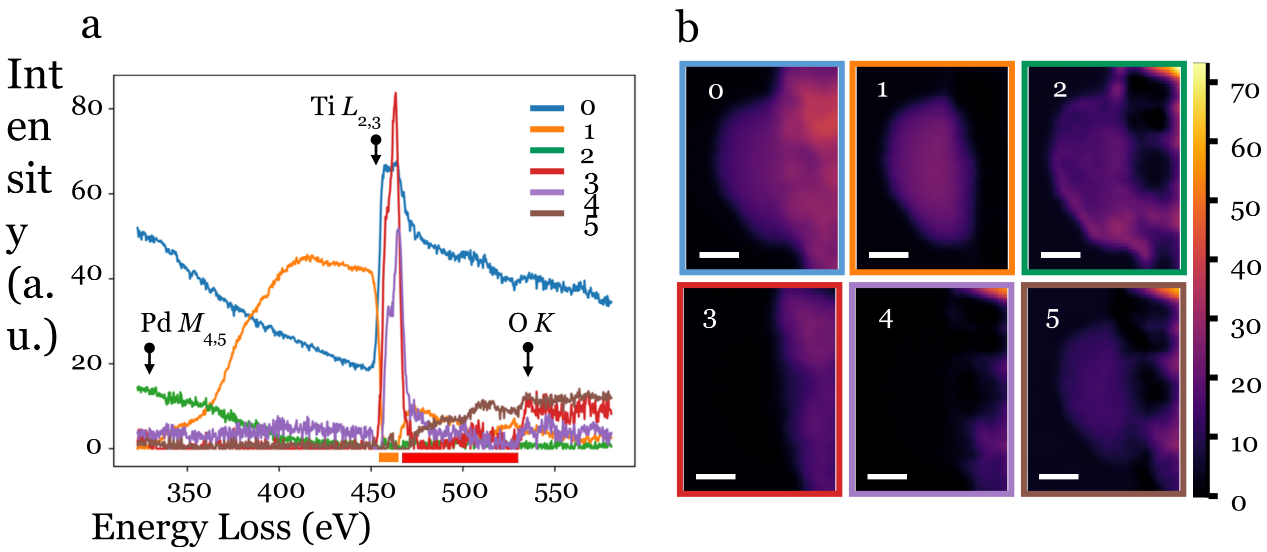


Figure 5:

Scikit-learn NMF matrix decomposition of unscaled EELS spectrum image. a.) endmembers showing locations of artifacts identified with orange and red bars under the spectra. b.) abundance maps showing the difficulty of the algorithm to differentiate between the oxide support and the metal nanoparticle. Arrows indicate approximate onset energies. Scale Bar: 5 nm

This method however is inadequate for the analysis of the chemical structure of the encapsulation layer, because it does not reveal the fine structure in the Ti L_{23} edge, particularly the edge onset, which is critical to the identification of the valence state of the oxide layer on the metal nanoparticle and in differentiating it from the oxide support. Additionally, NMF pulls out variations in the pre Pd background which could be the tail of the C K edge, obscuring the encapsulation layer. NMF also appears to be focusing on variations in the substrate fine-structure which could be due to the quickly changing thickness of the material as opposed to the fine structure difference between the Ti on the particle and in the support.

To address the inadequacies of existing methods, including NMF, we developed the ATGP Pre-Conditioned Joint NMF algorithm. ATGP is an algorithm for determining component spectra and comes from geo-remote sensing research where simulation of component spectra is exceedingly difficult. As explained in the article by Ren and Chang in 2004, ATGP uses orthogonal geometric projections to find pixels that form the vertices of a polygon in the orthogonal coordinate space.[37] Or, in other words, the ATGP algorithm finds the pixels that are the most different from within the dataset. ATGP is sensitive to large numerical values in the dataset and can generate spurious components particularly when there exist spectra with high valued outliers such as those introduced by X-rays generated in the spectrometer and cosmic rays which can occasionally hit the sensor during acquisition. Therefore, we first removed the largest and most obvious outliers from the dataset before processing by fitting a spline curve to the region before and after the spike, or outlier, and interpolating the missing value. To catch any small outliers which were less obvious, and to increase the signal-to-noise ratio, we applied a weak Gaussian blurring in the spatial dimensions (local averaging) and weak spectral smoothing in the spectral dimension to suppress some of these high intensity noise values. After finding the maximally different pixels (spectra) through ATGP, we use these as initial guesses for our decomposition algorithm. The spectra returned by ATGP are directly used as spectral component guesses, and each of these is fit to the spatial pixels to form initial guesses for the distribution maps. With these initial guesses, we use Joint-NMF to decompose the data matrix.[38] While blurring and smoothing were applied to the data for initial ATGP analysis, they were not for analysis by Joint-NMF.

Joint-NMF is an algorithm which combines the standard NMF algorithm and the more recent symmetric-NMF factorization algorithm to incorporate the spectral denoising and clustering properties of each technique in one algorithm. All NMF algorithms suffer from the occurrence of multiple local minima during the algorithmic minimization. It is for this reason that we use ATGP to generate an initial approximation. The Joint-NMF algorithm has been chosen over the standard scikit-learn NMF implementation because the scikit-learn implementation of NMF can introduce artifacts in the resulting spectral components and distribution maps, see detailed the description in the methods section.

CHAPTER 5:

Application of ATGP Pre-Conditioned Joint-NMF to the EELS-SI Dataset

Figure S3 shows the result of applying ATGP to the raw EELS dataset without blurring after the removal of anomalous X-ray and gamma ray spikes in the data. However, the resulting spectra are noisy, and multiple components are dominated by the TiO_2 signal from the support. Such results indicate that the noise in a single spectrum and the strong TiO_2 signal affect the pre-conditioning step with ATGP, and illustrate the importance of the blurring step prior to analyzing the raw data in the final version of the algorithm. The results of processing the data further with the ATGP Pre-conditioned Joint-NMF algorithm are plotted in Figure S4. The algorithm has done a fairly good job at identifying the Pd particle in component 1 but has unfortunately split the TiO_2 support into many components. One likely reason for the algorithm splitting the TiO_2 support into so many components is the strong signal produced by the support TiO_2 ELNES. ATGP is influenced by intensity in the spectral domain so the high intensity of the TiO_2 support signal and the variation in it will bias the algorithm towards variations in the support. Further, the algorithm also fails in revealing the TiO_2 encapsulation layer. These results show that in order to identify the presence of the encapsulation layer through Ti-L edges, we need to scale the data properly so that the intensity change between the TiO_2 on the particle and the TiO_2 of the support is not so drastic.

Importance of a Proper Data-Scaling

In the application of machine learning algorithms for materials science, an often ignored pre-treatment step is the proper scaling of the data. In the case of an EELS spectrum image where there are large changes in thickness, the intensity of the scattered electrons can vary significantly between thin and thick areas of the specimen. These changes in thickness can also affect the shape of the EELS edges through multiple-scattering induced broadening. These large changes in intensity can be a hindrance to an accurate data fitting. Normalization or scaling must be properly performed to allow the algorithm to fit these variations. This is especially important for the study of ultra-thin, potentially sub-nanometer thick, encapsulation layers on nanoparticles and supports with a thickness that is 2-3 orders of magnitude higher. In this experiment, two assumptions can be made: the noise floor of the detector was approximately normally distributed with a mean at a fixed non-negative value and the electron energy loss signal followed Poisson counting statistics which, to a good approximation, will also appear normally distributed, about some offset, for sufficient intensities in the energy loss spectrum. Thus, offsetting the data such that the minimum value is at least 1.0 and scaling the data by taking the logarithm, element-wise, should continue to allow the algorithms to find components and retain the non-negativity of the dataset. The distributions of the noise floor and the EELS signal should also remain approximately normally distributed about their expected values.[39]

Application of ATGP Pre-Conditioned Joint-NMF to the Dataset

In order to see the small variations in the Ti L_{23} edge we used a smaller energy window view (420-569eV) of the same dataset, which only includes the Ti *L* edges and O *K*-edge. As mentioned in the previous section, we took the natural logarithm of the dataset, element-wise, after shifting the data such that none of the channels have an intensity below 1.0. This offset is consistent with a uniform background signal from the detector and a log scaling helps prevent the large intensity variations between the encapsulation layer and the Pd nanoparticle from drowning out the encapsulation layer's signal. We verified the expected number of components, k = 6, from a PCA explained variance ratio plot via the elbow method as previously described and is plotted in Figure S5 We also applied NMF to scaled data, as mentioned above, which improved the results produced by NMF as expected. These results can be found plotted in Figure S6 in the supplementary information. Still, the results were inferior to those produced by pre-conditioned Joint-NMF.

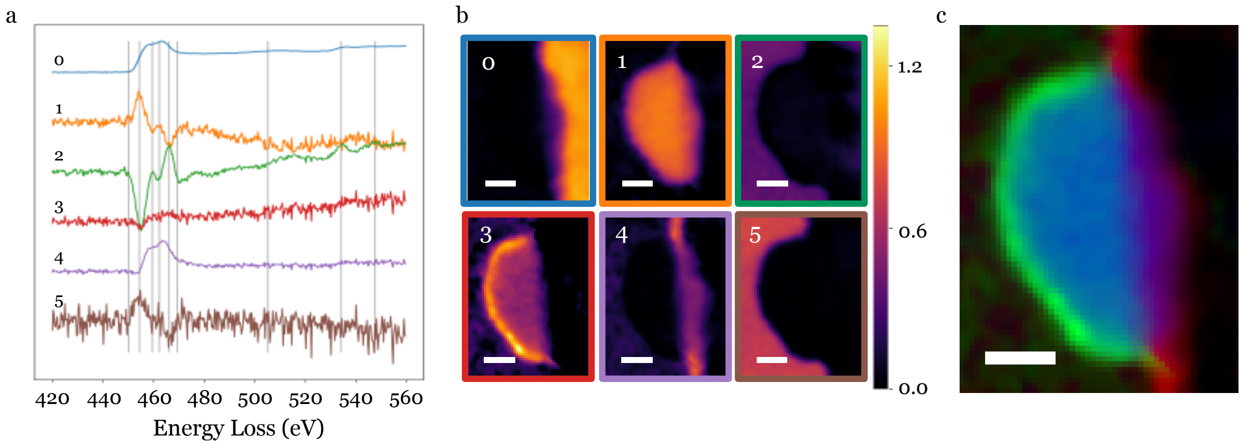


Figure 6:

Background subtracted spectral components a) and weight distribution maps b) from ATGP pre-conditioned joint NMF processing of the log-scaled spectrum image; where the energy loss has been cropped to the range between 420 & 560 eV. c) False-color composite of components 1 (red), 3 (green), 4 (blue) clearly showing how component 3 encapsulates component 1. The six components which can be explained as: (0) the oxide support, (1) the Pd M_{45} background and a part of the Ti L_{23} ELNES signal of the SMSI , (2) modification of the TiO2 support, (3) the SMSI encapsulation layer, (4) the TiO_2 surface / Pd interface (5) the vacuum region and detector noise. With the power-law background removed and scaled to better view the endmember structure related to the Ti L_{23} and O K ELNES. The structure of component 1 (orange) in a) clearly shows an onset at about 450 eV

Scale Bar: 5 nm;

Energy loss (eV) at grey lines, left to right: 450.119, 454.432, 459.507, 462.299, 466.105, 469.403, 505.181, 534.108, 547.557.

Figure 6b and S7 show the results of running the log-scaled data through the pre-conditioned Joint-NMF algorithm. It is immediately apparent from the abundance maps that the algorithm has separated the Pd nanoparticle and the TiO_2 support, components 1 and 0 respectively. It is also apparent that the algorithm has captured the vacuum region on the left side of the spectrum image as indicated in the spectral endmembers in component 5 with nearly uniformly distributed noise.

Figure 6c shows a false-color representation of component maps 1, 3, and 4. Component 3 appears to cover the surface of the Pd nanoparticle, indicating that it has captured information about the SMSI encapsulation layer. Since the Ti at the support surface should be reduced with respect to the bulk,[40,41] component 4 therefore appears to result from this reduced surface material. Component 2 appears to be variation in the signal from the TiO_2 support, with the intensity in the vacuum region being an artifact that is balanced by the peak that appears in the vacuum noise signal captured in component 5. It is interesting to note that the Ti signal at the Pd/TiO_2 interface is represented by a combination of two components, component 1, which includes the Pd particle, and component 4, the support surface layer. It should also be noted that since the Pd nanoparticle is always co-located with some Ti, there is information from the Ti surface encapsulation layer contained in the component that has been identified as the Pd particle, component 1. Aside from relying on PCA for the number of components in order to reduce operator bias, it is possible that the splitting of the Ti surface encapsulation layer EELS signature into two components is a result of the presence of an orientation effect on the Ti ELNES or potentially indicating a real but small spatial separation in the Ti and O signals at the surface.

A closer analysis of the Ti edge onsets and their spatial correlations allows us to assign each spectra to the constituents in the system, and is shown in Figure 4a, where these spectra were obtained by fitting and subtracting a power-law background function from the energy range just before the Ti L_{23} edge. Each spectrum was then normalized to the maximum value in the spectral range shown.

From the plots in Figure 4a, we can see that the Pd nanoparticle identified in component 1 has a peak in the onset region of the Ti L_{23} edge which is consistent with a reduction in the oxidation state of Ti as compared to component 0. Component 2 appears to have captured a variation in the thickness and fine-structure of the Ti L_{23} edge within the TiO_2 support. It should be noted that in addition to the energy resolution, one reason why the Ti L_{23} splitting is not easily resolved in the components is that the Ti L_{23} peaks shift along the energy loss axis and continuous shifts are exceedingly difficult for multilinear decomposition algorithms, including ATGP Pre-conditioned Joint-NMF, to resolve.

While we cannot use the background subtracted plots in Figure 4a to quantitatively characterize the electronic states from their ELNES peak-height ratios in the different pieces of this heterogeneous catalyst system, we can use the differences in the characteristic spectra, particularly their edge onsets, to draw conclusions about changes within the electronic states present in the system. For example, as shown in Figure 6a, the Ti L_{23} ELNES spectrum is encoded with information from the electronic structure of the Ti atoms and its associated bonds. From the spatial maps we can see that components 1 and 3 correspond to the TiO_x encapsulation layer. We can also see that component 1 indicates a lower onset energy of the Ti L_{23} peak on the surface of the Pd particle. This indicates a reduction in the Ti oxidation state on the surface of the particle due to the loss of an electron and the ELNES signatures are significantly different on the surface of the particle compared with the bulk of the material.

CHAPTER 6:

Discussion & Conclusion

We first tested the existing NMF and PCA algorithms for the analysis of EELS-SI datasets. While these methods have been shown to effectively differentiate overlapping peaks in EELS, they struggle in revealing trace signals in a complicated data environment when abrupt change in signal intensity is present among spectra, e.g. revealing the ultra-thin encapsulation layer that is adjacent to a thick high-signal oxide support.

To address the challenges with using NMF in these cases, we proposed a new algorithm, ATGP Pre-conditioned Joint-NMF, which combines the advantages of the complementary algorithms ATGP and NMF. Our results show that while ATGP Pre-conditioned Joint-NMF presents advantages over NMF and PCA in terms of revealing more of the weak signal information from the Pd particle surface, it fails in revealing components that can be straightforwardly associated to specific physical meaning.

We discovered that proper data scaling is critical to maximize the advantages of ATGP Pre-conditioned Joint-NMF. By incorporating log-scaling before the application of ATGP Pre-conditioned Joint-NMF, we successfully revealed not only the presence, but also the chemistry and the oxidation states of Ti in the encapsulation layer along with those of the particle/support pair, information critical to the understanding of catalytic reaction mechanisms and potentially help develop better site-specific reaction models. The individual components of the decomposed spectrum image, as shown in Figure 6b and S7, qualitatively display the distribution of the different oxidation states of Ti and the location of the Pd particle. A closer analysis of the Ti and O edges, shown in

Figure 6a, reveals that the Ti *L* edge of the SMSI components has a lower onset energy relative to the TiO_2 support, indicating the species of the oxide on the particle is in a reduced state, TiO_x, compared to the TiO_2 support congruent with prior reports.[29] Further, a comparison between the individual components shows qualitatively that differences in the peak height ratio have been identified by the algorithm. As shown in Figure 6a, there is a difference in the peak height ratio between components 2, 0, and 4, all representing changes within the TiO_2 support and components 1 and 3 show distinct changes in the TiO_2 ELNES as well.

The algorithm, ATGP Pre-conditioned Joint-NMF, demonstrated here is expected to largely benefit the characterizations of heterogeneous catalysts, given the fact that a reliable method of detecting an encapsulation layer has been lacking and its ability of clearly revealing the hidden signals in SI datasets and quantifying their spatial distributions in a complicated signal environment. In fact, the application of this method can be readily extended to any STEM-EELS or EDX datasets, therefore, detecting challenging signals from trace dopants, impurities, as well as the distribution of chemical coordination at surfaces and interfaces in heterogeneous catalysts. Further, applying ATGP Pre-conditioned Joint-NMF is expected to benefit the analysis of in situ STEM spectrum imaging, as it would enable an effective separation of the signals related to the support membrane and the dispersed catalysts in large multi-dimensional datasets.

Pre-conditioned Joint-NMF could have numerous benefits for the field at large, and it is generally applicable to datasets acquired from other spectroscopy techniques. It can track the relative presence of different components even if the spectral features overlap and can reveal changes in fine structures even in spectra where background fitting is challenging. It can significantly reduce data acquisition time because of its high tolerance to low signal-to-background spectra, especially if it is combined with compressive sensing algorithms for sparse data acquisition. Further, since this method can be run completely unsupervised, the application of this method may allow us to discover new features or phenomena that were neglected previously. Besides the numerous advantages discussed above, it should be emphasized that this algorithm is new and further optimization could be performed, especially regarding its ability to separate components that are connected in the spatial domain. Potential future studies may include modifying the objective function to incorporate more physical constraints such as conserved quantities and known relationships. Largely, machine learning methods for materials science data analysis is still in its infancy. The largest challenge, which is also the key goal in advancing the algorithm in this work, is to extract components that are physically meaningful with minimum descriptors. Through this work, we also learned that proper pre-conditioning for narrowing the number of solutions to the minimization problem is critical to the extraction of meaningful components and preventing overfitting. Further work regarding discovering and incorporating new minimization constraints based on physical conservation rules should be explored.[42]

In summary, we demonstrated a robust new machine learning algorithm, ATGP pre-conditioned Joint-NMF, for the analysis of unmixing spectrum imaging datasets, particularly for the study of the strong metal-support interaction encapsulation in heterogeneous catalysts, which has been challenging to conventional data analysis methods. This algorithm successfully extracts components associated with specific physical meanings which reveal the presence and chemical composition and structure of the trace encapsulation layer. This information would otherwise be hidden in the raw EELS dataset. This new algorithm was built on two existing algorithms, Joint-NMF and ATGP, which individually are not sufficient to reveal trace signals in a complex data environment in the spatial domain, e.g. the case of SMSI encapsulation. Its advancements were achieved by combining careful selection of data scaling methods, the introduction of the clustering effects of Joint-NMF, and the initialization from ATGP. The machine learning method we introduced here is generally applicable to any other spectra-imaging datasets for various material systems, for example, targeting trace or hidden signals when drastic changes in spectra are present in the spatial domain, which can occur when boundaries, interfaces, and surfaces are present in the region of interest.

REFERENCES

1. Liu, J. J. Advanced Electron Microscopy of Metal-Support Interactions in Supported Metal Catalysts. *ChemCatChem* **3**, 934–948 (2011).

2. Shaikhutdinov, S. Strong Metal–Support Interaction and Reactivity of Ultrathin Oxide Films. *Catal. Letters* **148**, 2627–2635 (2018).

3. Tauster, S. J., Fung, S. C., Baker, R. T. K. & Horsley, J. A. Strong interactions in supported-metal catalysts. *Science (80-.).* (1981). doi:10.1126/science.211.4487.1121

4. Tang, H. *et al.* Classical strong metal–support interactions between gold nanoparticles and titanium dioxide. *Sci. Adv.* **3**, e1700231 (2017).

5. Diebold, U. The surface science of titanium dioxide. *Surf. Sci. Rep.* **48**, 53–229 (2003).

6. Sharma, R. Design and Applications of Environmental Cell Transmission Electron Microscope for In Situ Observations of Gas-Solid Reactions. *Microsc. Microanal.* (2001). doi:10.1007/S10005-001-0015-1

7. Tauster, S. J., Fung, S. C. & Garten, R. L. Strong Metal-Support Interactions. Group 8 Noble Metals Supported on TiO2. *J. Am. Chem. Soc.* (1978). doi:10.1021/ja00469a029

8. Stavitski, E. & de Groot, F. M. F. The CTM4XAS program for EELS and XAS spectral shape analysis of transition metal L edges. *Micron* (2010). doi:10.1016/j.micron.2010.06.005

9. Fan, L. & Fujimoto, K. Hydrogenation of Carbon Dioxide to Methanol by Titania-Supported Palladium Catalyst: Promotive SMSI Effect. *Bull. Chem. Soc. Jpn.* **67**, 1773–1776 (1994).

10. Lichtert, S. & Verbeeck, J. Statistical consequences of applying a PCA noise filter on EELS spectrum images. *Ultramicroscopy* **125**, 35–42 (2013).

11. Dobigeon, N. & Brun, N. Spectral mixture analysis of EELS spectrum-images.

Ultramicroscopy **120**, 25–34 (2012).

12. Feynman R. Plenty of room at the bottom? *APS Annu. Meet.* (1959). doi:10.1511/2012.96.226

13. Krivanek, O. L., Dellby, N. & Lupini, A. R. Towards sub-Å electron beams. *Ultramicroscopy* **78**, 1–11 (1999).

14. Dellby, N., Krivanek, O. L. & Murfitt, M. F. Optimized quadrupole-octupole C3/C5 aberration corrector for STEM. *Phys. Procedia* **1**, 179–183 (2008).

15. Krivanek, O. L., Delby, N., Lupini, A. R., Dellby, N. & Lupini, A. R. Autoadjusting charged-particle probe-forming apparatus. (2003).

16. Nellist, P. D. Direct Sub-Angstrom Imaging of a Crystal Lattice. *Science (80-.).* **305**, 1741–1741 (2004).

17. Rodenburg, J. M., Hurst, A. C. & Maiden, A. High resolution transmission imaging without lenses. *J. Phys. Conf. Ser.* **241**, (2010).

18. Lupini, A. R., Oxley, M. P. & Kalinin, S. V. Pushing the limits of electron ptychography. *Science (80-.).* **362**, 399–400 (2018).

19. Yang, H. *et al.* Electron ptychographic phase imaging of light elements in crystalline materials using Wigner distribution deconvolution. *Ultramicroscopy* **180**, 173–179 (2017).

20. Rodenburg, J. M. Ptychography and related diffractive imaging methods. *Adv. Imaging Electron Phys.* **150**, 87–184 (2008).

21. Jesse, S. *et al.* Big Data Analytics for Scanning Transmission Electron Microscopy Ptychography. *Sci. Rep.* **6**, (2016).

22. Midgley, P. A. & Dunin-Borkowski, R. E. Electron tomography and holography in materials science. *Nat. Mater.* **8**, 271–280 (2009).

23. Midgley, P. A. Midgley Holography Review. **32**, 1–18 (2000).

24. Egerton, R. F. *Electron Energy-Loss Spectroscopy in the Electron Microscope.*

56, (Springer US, 2011).

25. Krivanek, O. L. *et al.* Progress in ultrahigh energy resolution EELS. *Ultramicroscopy* **203**, 60–67 (2019).

26. de Groot, F. & Kotani, A. *Core level spectroscopy of solids*. *Core Level Spectroscopy of Solids* (2008). doi:10.1201/9781420008425

27. de Groot, F. M. ., Grioni, M. & Fuggle, J. . Oxygen 1s x-ray-absorption. *Phys. Rev. B* **40**, 5715–5723 (1989).

28. de Groot, F. M. F. X-ray absorption and dichroism of transition metals and their compounds. *J. Electron Spectros. Relat. Phenomena* **67**, 529–622 (1994).

29. Stoyanov, E., Langenhorst, F. & Steinle-Neumann, G. The effect of valence state and site geometry on Ti L 3,2 and O K electron energy-loss spectra of Ti x O y phases. *Am. Mineral.* **92**, 577–586 (2007).

30. Gloter, A., Ewels, C., Umek, P., Arcon, D. & Colliex, C. Electronic structure of titania-based nanotubes investigated by EELS spectroscopy. *Phys. Rev. B - Condens. Matter Mater. Phys.* (2009). doi:10.1103/PhysRevB.80.035413

31. Zhang, S. *et al.* Dynamical Observation and Detailed Description of Catalysts under Strong Metal–Support Interaction. *Nano Lett.* **16**, 4528–4534 (2016).

32. Bonnet, N., Brun, N. & Colliex, C. Extracting information from sequences of spatially resolved EELS spectra using multivariate statistical analysis. *Ultramicroscopy* **77**, 97–112 (1999).

33. Shiga, M. *et al.* Sparse modeling of EELS and EDX spectral imaging data by nonnegative matrix factorization. *Ultramicroscopy* **170**, 43–59 (2016).

34. Nomura, Y. *et al.* Quantitative Operando Visualization of Electrochemical Reactions and Li Ions in All-Solid-State Batteries by STEM-EELS with Hyperspectral Image Analyses. *Nano Lett.* **18**, 5892–5898 (2018).

35. Kannan, R. *et al.* Deep data analysis via physically constrained linear unmixing: universal framework, domain examples, and a community-wide platform. *Adv.*

Struct. Chem. Imaging **4**, 6 (2018).

36. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in {P}ython. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

37. Hsuan Ren & Chein-i Chang. Automatic spectral target recognition in hyperspectral imagery. *IEEE Trans. Aerosp. Electron. Syst.* **39**, 1232–1249 (2004).

38. Blum, T. *et al.* Machine Learning for Challenging EELS and EDS Spectral Decomposition. *Microsc. Microanal.* **25**, 180–181 (2019).

39. Salakhutdinov, R. & Mnih, A. Probabilistic matrix factorization. *Adv. Neural Inf. Process. Syst. 20 - Proc. 2007 Conf.* 1–8 (2009).

40. Suzuki, T. & Souda, R. Encapsulation of Pd by the supporting TiO2(110) surface induced by strong metal-support interactions. *Surf. Sci.* **448**, 33–39 (2000).

41. Zhang, S. *et al.* Dynamical observation and detailed description of catalysts under strong metal-support interaction. *Nano Lett.* **16**, 4528–4534 (2016).

42. Jia, X. *et al.* Physics Guided Recurrent Neural Networks For Modeling Dynamical Systems: Application to Monitoring Water Temperature And Quality In Lakes. (2018).

43. Du, R., Drake, B. & Park, H. Hybrid clustering based on content and connection structure using joint nonnegative matrix factorization. *J. Glob. Optim.* 1–17 (2017). doi:10.1007/s10898-017-0578-x

44. Felipe Polo-Garzon, Thomas F. Blum, Zhenghong Bao, Zhennan Huang, Elizabeth E. Bickel, Kristen Wang, Victor Fung, De-en Jiang, Miaofang Chi, Zili Wu. "*In Situ* Strong Metal-Support Interaction (SMSI) is Behind Unusual Catalytic Behavior" 2020. In preparation.

Appendix A: METHODS

Selection of Joint-NMF Over Standard NMF

The Joint-NMF algorithm has been chosen over the standard scikit-learn NMF implementation because the scikit-learn implementation of NMF can introduce artifacts in the resulting spectral components and distribution maps. We believe this occurs due to a step taken to overcome an inherent weakness of the NMF objective function, the difference between the recorded data and the model, which is that the factoring of a single matrix into two matrices is inherently ill-defined as shown below. NMF attempts to optimize the following expression:

$$\min_{W\geq 0,\ H\geq 0} \|X - WH\|^F$$

$X \in \mathbb{R}_+^{m\times n}$, $W \in \mathbb{R}_+^{m\times k}$, $H \in \mathbb{R}_+^{k\times n}$, and k is the number of distinct components in the data. Given the ideal case where $X = WH$, WH can still be expanded so that $X = WA^{-1}AH$ making the solutions W and H not unique. Other constraints to the optimization must then be used to help guide the algorithm to a unique solution. Recently, Du et. al. proposed a new version of NMF termed Joint-NMF which optimizes a linear combination of the NMF objective function and that of Symmetric NMF, which attempts to optimize the following objective function:

$$\min_{H\geq 0} \|S - H^T H\|^F$$

where S is a kernel matrix that represents a pairwise measure of the similarity between the observations in the data matrix. When the symmetric NMF and normal NMF functions are combined, the effect is that the resulting components are inherently clustered.[43] This additional optimization should help the algorithm pick components that are closer to the real components and should help the algorithm find the global minimum and not get stuck in a local minimum.

Scanning Transmission Electron Energy-loss Spectroscopic Imaging

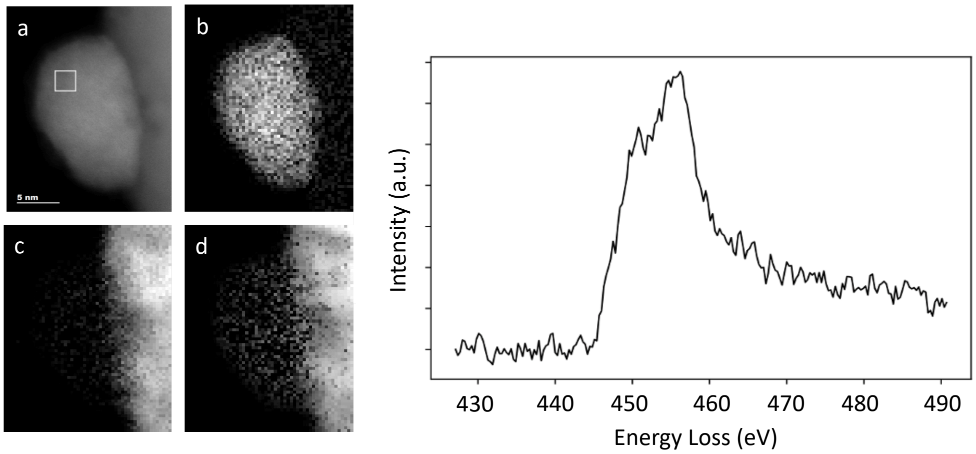
Z-contrast images and EELS spectrum images were acquired with a NiON UltraSTEM aberration-corrected STEM operated with an acceleration potential of 100 kV and outfitted with a Gatan Enfinia EELS spectrometer. The convergence semi-angle was about 30.0 mrad for both STEM images and the EELS spectrum image. The pixel exposure time for the EELS spectrum image was 0.1 s and the acceptance semi-angle was about 48.0 mrad.

Sample Preparation

The Pd/TiO_2 catalyst was prepared by depositing Pd particles on anatase titanium dioxide powder via the deposition-precipitation method with urea. The Pd loading on the sample was 0.35 wt%. The encapsulation SMSI layer was formed by heating the sample in a tube furnace in 83% H_2/Ar at 600 °C for 1h. Diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) of CO adsorption showed that ~70 % of Pd-sites are covered when the sample is treated in 50% H_2/Ar at 600 °C. Details on the sample preparation and CO adsorption measurements can be found in a separate manuscript in preparation by F.G., T.B. M.C., and Z.W..[44] The H_2/Ar treated catalyst powder was loaded onto a standard 3mm lacey carbon grid (Electron Microscopy Sciences) after the treatment by dropping a lacey carbon grid into a glass vial containing a few milligrams of the powder and gently agitating the vial. Then the grid was removed with tweezers and the excess particles gently shaken off. The sample was then loaded immediately into the microscope for imaging.

APPENDIX B: SUPPLEMENTARY FIGURES

Figure S1:



a) Simultaneous Z-contrast image of Pd nanoparticle on TiO_2 support. b-d) Integrated intensity of Pd $M_{4,5}$, Ti $L_{2,3}$, and O K edges respectively. e) summed and background subtracted spectra from inset of a). A demonstration of the traditional analysis technique for EELS core-loss spectrum images. A power law background is subtracted from each core-loss edge and the signal integrated over an energy window placed after the ELNES fine structure.

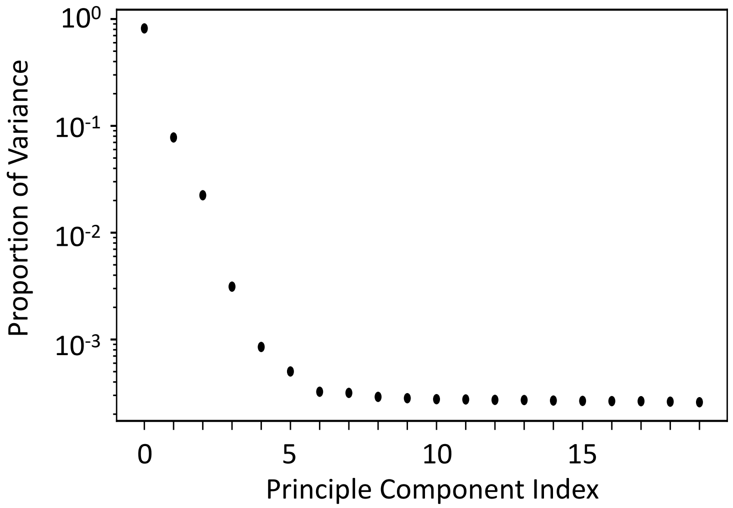


Figure S2:

Explained variance ratio plot versus component index from Principle Component Analysis of the unscaled EELS dataset in this work. Shows the elbow point at component 5 indicating a total number of 6 statistically relevant components.

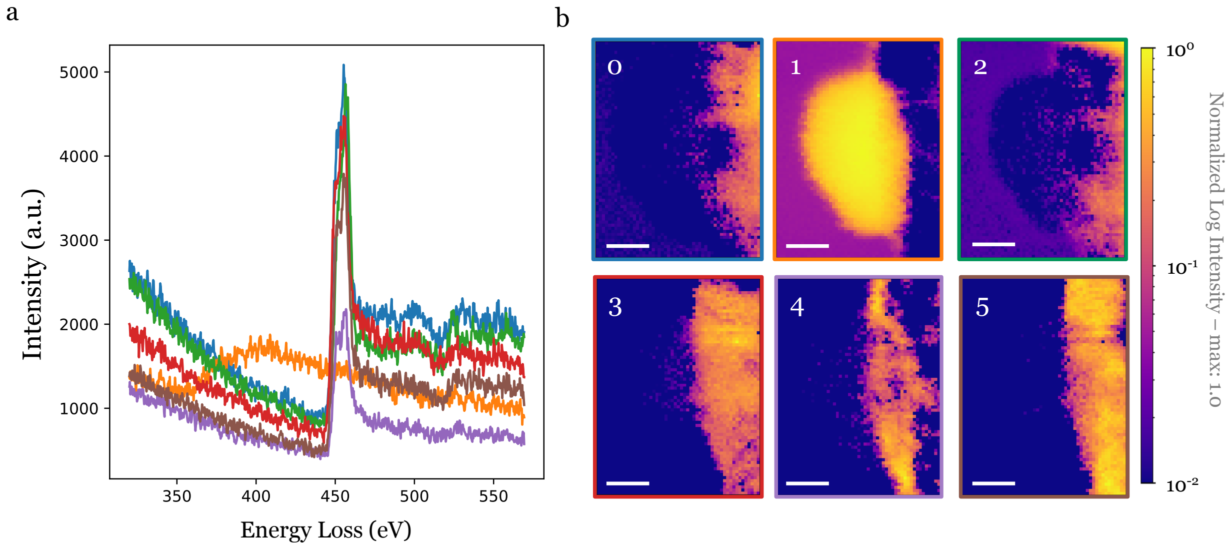


Figure S3:
ATGP spectral components a) and non-negative least squares fitted distribution maps b) used for preconditioning. Each spectrum of a) corresponds to the entire map of b) Note the noise level in the spectra. Scale Bar: 5nm

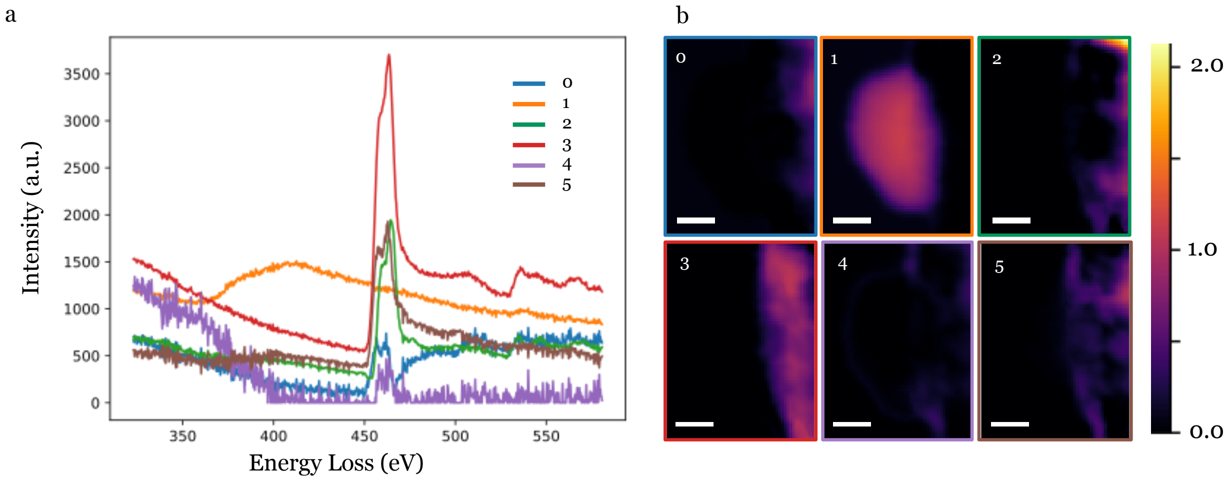


Figure S4:
Spectral components a) and weight maps b) from ATGP Pre-conditioned Joint-NMF decomposition of the unscaled and uncropped data set. These figures show poor contrast from the encapsulation layer and splitting of the support signal potentially due to outliers and thickness variations. Scale bar: 5nm

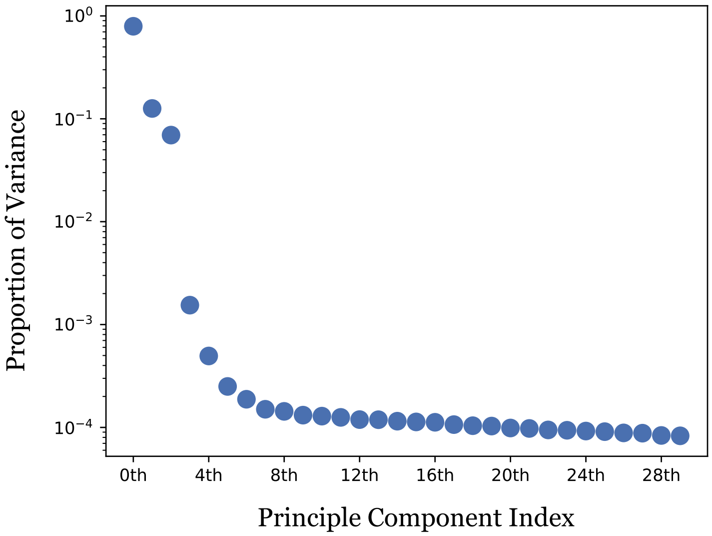


Figure S5:
Explained variance ratio plot from Principle Component Analysis of the log-scaled dataset. The elbow at component 5 indicates 6 six statistically relevant components.

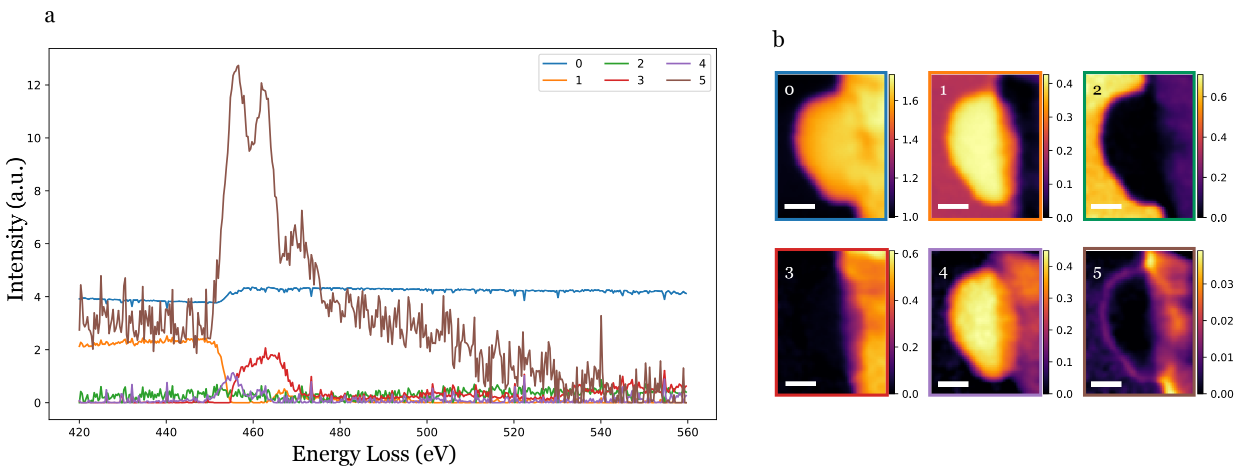


Figure S6:

NMF Decomposition of log-scaled data. The log scaling has improved NMF's ability to separate out what appears to be the encapsulation layer, as evidenced by the maps in b), but the effects of the sparsity constraint in the component spectra is apparent in a). Scale Bar: 5nm

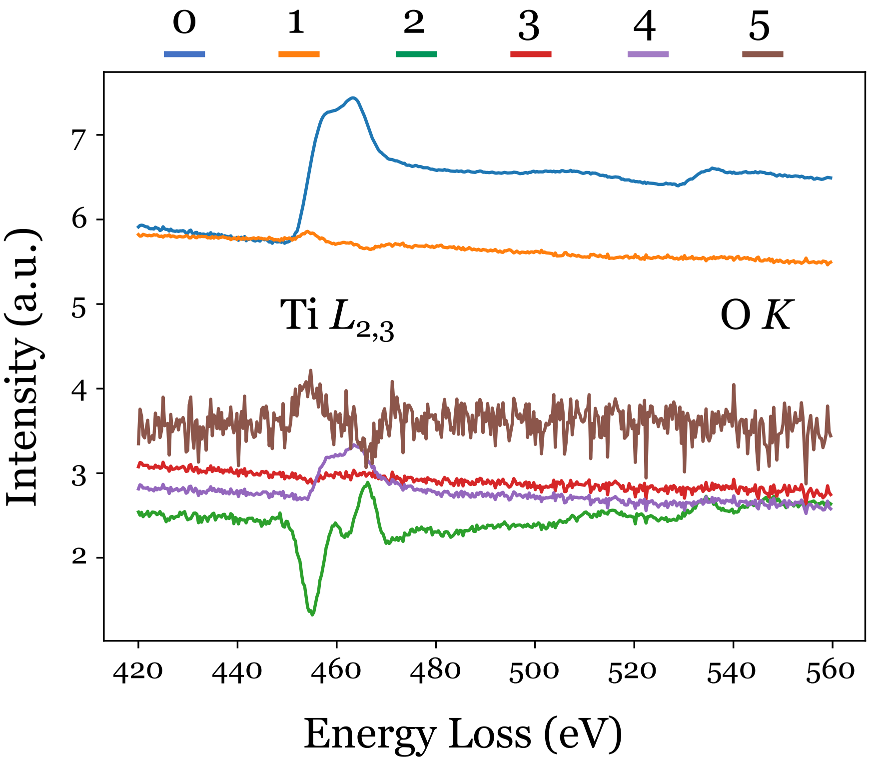


Figure S7:
Spectral components from ATGP Pre-conditioned Joint-NMF decomposition of the log-scaled SI dataset before power-law background removal.