2015-12-01

Advanced Data Analysis Tools and Multi-Instrument Material Characterization

Bhupinder Singh
Brigham Young University

Advanced Data Analysis Tools and Multi-Instrument Material Characterization

Bhupinder Singh

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Matthew R. Linford, Chair
Milton L. Lee
Jaron C. Hansen
Daniel E. Austin
Robert C. Davis

Department of Chemistry and Biochemistry

Brigham Young University

December 2015

ABSTRACT

Advanced Data Analysis Tools and Multi-Instrument Material Characterization

Bhupinder Singh
Department of Chemistry and Biochemistry, BYU
Doctor of Philosophy

My dissertation focuses on (i) the development of new analysis tools and methodologies for analyzing X-ray photoelectron spectroscopy (XPS) and time-of-flight secondary ion mass spectrometry (ToF-SIMS) data, and (ii) the comprehensive characterization of materials (nanodiamonds) using a multi-instrument approach. Chapter 1 contains (i) a discussion of the two techniques I focused on most in my work: XPS and ToF-SIMS, (ii) a discussion of the common chemometrics techniques used to analyze data from these methods, and (iii) the advantages/rationale behind the multi-instrument characterization of materials. Chapter 2 describes various good practices for obtaining reasonable peak fits in XPS, which can also be applied to peak fitting data from different techniques. To address the issue of user subjectivity/bias in XPS peak fitting, I introduce two less biased mathematical functions for characterizing XPS narrow scans, namely the equivalent width (EW) and the autocorrelation width (AW). These functions are discussed in Chapters 3 and 4. In Chapter 5, I then introduce uniqueness plots as simple and straightforward graphical tools for assessing the quality of XPS peak fits and for determining whether fit parameters are correlated. This tool is extensively used in spectroscopic ellipsometry, and the mathematics behind it is known in XPS. However, to the best of my knowledge, this graphical tool has never been applied to XPS. ToF-SIMS data analysis is somewhat challenging due to the enormous amounts of data that are collected, and also the matrix effect in SIMS. This amount of information is significantly increased when depth profiles are performed on samples. Chapter 6 discusses a new chemometrics tool that I introduce for analysis of complex data sets, with emphasis on XPS and ToF-SIMS depth profiling data. The new approach is called the Information Content (IC) or entropy, which is adapted from Claude Shannon's work on Information Theory. Chapter 7 then contains a presentation of the comprehensive characterization of five nanodiamond samples used to manufacture particles for liquid chromatography. The advantages of a multi-instrument approach for material characterization and the lack of comprehensive material characterization in the literature are emphasized. To the best of my knowledge this is the most comprehensive characterization of nanodiamonds that has been reported in the literature. Chapter 8 presents conclusions of my work and future work. This thesis also contains six appendices. Appendix 1 contains an article from a scientific magazine that I wrote to highlight the importance and applications of the EW and AW to characterize XPS narrow scans. Appendices 2-5 are application notes I wrote on separations I performed on a nanodiamond based HPLC column. Finally, Appendix 6 describes the ToF-SIMS analysis of the tungsten species in the nanodiamond samples characterized in Chapter 7.

Keywords: XPS, ToF-SIMS, equivalent width, autocorrelation width, uniqueness plot, information content, multi-instrument material characterization

ACKNOWLEDGEMENTS

The journey through graduate life hasn't been a solo one. I surely had considerable help from the people around me. I would like to extend my greatest appreciation and gratitude to my PhD advisor Dr. Matthew R. Linford for his excellent guidance and support throughout my graduate work. I feel privileged to have worked with him during my graduate studies.

I would also like to thank my committee members: Drs. Milton Lee, Jaron Hansen, Daniel E. Austin and Robert C. Davis for providing their valuable suggestions and insights throughout my graduate work.

I have had the opportunity to work with some old friends and make new friends during my stay at BYU. I would like to acknowledge my friends and lab-mates: Dr. Chuan-Hsi Hung, Dr. Landon Wiest, Hao Wang, Cody Cushman, Tuhin Roychowdhury, and Varun Jain. I would like to specifically thank Dr. Supriya Kanyal, Anubhav Diwan, and Dr. David S. Jensen for their help and support.

At BYU, I met many helpful and nice people. The administrative staff of the Department of Chemistry and Biochemistry is awesome. I would like to thank Janet Fonoimoana who has always been helpful and caring. I would also like to acknowledge the help of Sue Mortensen.

In addition, I would like to thank Diamond Analytics Corporation (Andrew Dadson and Michael Vail), Moxtek, Inc., and the Roland K. Robins fellowship for helping my research financially. I am deeply thankful to the Department of Chemistry and Biochemistry for giving me the opportunity to pursue my graduate studies at Brigham Young University.

One thing that I have missed the most during my graduate journey has been my family and it was tough living away from them. However, their continuous support and encouragement kept me going. I would like to thank my parents, Mr. Pritpal Singh and Mrs. Hardeep Kaur, and my uncle, Mr. Iqbal Singh, and aunt, Mrs. Pardeep Kaur. I also acknowledge with love my brothers, Gurjeet Singh and Kulwant Singh, for all the fun talks and being there for me.

Last, but not least, I would like to thank *The Almighty* for providing me this opportunity and helping me through my education.

I would like to dedicate my dissertation to my paternal grandfather the late Shri Piara Singh and my maternal grandfather the late Shri Natha Singh, who would have been very happy about my achievement

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2,4-D	2,4-dichlorophenoxyacetic acid
AA	Advanced Abrasives
AW_{XPS}	Autocorrelation width of XPS narrow scans
BET	Brunauer-Emmett-Teller
CNTs	Carbon nanotubes
DRIFT	Diffuse reflectance infrared Fourier transform
EELS	Electron energy loss spectroscopy
EW_{XPS}	Equivalent width of XPS narrow scans
HPLC	High performance liquid chromatography
IC	Information content
ICP-MS	Inductively coupled plasma-mass spectrometry
ITC	International technology center
MCPA	2-methyl-4-chlorophenoxyacetic acid
NDs	Nanodiamonds
PCA	Principal components analysis
PSD	Particle size distribution
SE	Spectroscopic ellipsometry
TEM	Transmission electron microscopy
ToF-SIMS	Time-of-flight secondary ion mass spectrometry
XPS	X-ray photoelectron spectroscopy
XRD	X-ray diffraction

Chapter 1: Introduction to XPS and ToF-SIMS, including their data analysis, and the multi-instrument characterization of materials

1.1 Introduction

Material characterization and analysis play a central role in the advancement of a wide range of materials, including semiconductor devices,[1-4] materials for separation science,[5-13] data storage devices,[14-18] biomaterials,[19, 20] and hydrophobic coatings.[13, 21-25] In this regard, a comment from the National Materials Advisory Board (NMAB) on materials characterization, which was written some time ago, is still relevant and insightful: "*Characterization describes those features of the composition and structure (including defects) of a material that are significant for a particular preparation, study of properties, or use, and suffice for the reproduction of the material*."[26, 27] For many materials, the most important area of interest is its *surface* because the surface interacts directly with its surroundings vis-à-vis its physical and chemical properties.[26, 28-31] For example, the outer monolayer (0.1 – 1 nm) of a material often plays a central role in catalysis, biological signaling and transport, separation science, and sensors. For other applications, such as semiconductor manufacturing, it is often critical to know the distribution of elements in a material as a function of depth into it.[32-34] A knowledge of surfaces also aids in our understanding of corrosion, adhesion, wettability, and device failure.[35] Fortunately, there is a wide suite of surface analytical techniques available to scientists, which have become available as a result of technological improvements in the fields of vacuum technology, optics, lasers, electronics, and computer science.[29]

A thorough understanding of materials leads to the development of desired surface chemistries suited to specific applications, as well as to other advances in material properties. It

is nearly impossible to completely characterize a material using just one characterization technique. Each of the commonly used methods *views* a material in a somewhat limited way, shedding light on a certain property/characteristic of the same. Therefore, comprehensive material characterization requires a multi-instrumental approach. This has been one of the major focuses of the Linford group at Brigham Young University.[5, 6, 14, 18, 36]

Some of the analytical tools that are extensively used in the Linford group at BYU include X-ray photoelectron spectroscopy (XPS), time-of-flight secondary ion mass spectrometry (ToF-SIMS), spectroscopic ellipsometry (SE), water contact angle goniometry, atomic force microscopy (AFM), scanning electron microscopy (SEM), transmission electron microscopy (TEM), and diffuse reflectance infrared Fourier transform spectroscopy (DRIFT). This dissertation focuses primarily on two of these methods, which are two of the most commonly used surface analytical techniques: XPS and ToF-SIMS. My major contribution as a graduate student has been in developing new tools for the analysis of XPS and ToF-SIMS data. I have also carefully characterized a set of different nanodiamond samples using these and other instrumental techniques. In the remainder of this Introduction, I will now discuss XPS and ToF-SIMS in detail. I will then describe various data processing tools that can be used for XPS and ToF-SIMS data analysis. This will be followed by a discussion of the multi-instrument characterization of materials.

1.2 X-ray photoelectron spectroscopy (XPS)

XPS is one of the most commonly used surface analytical techniques,[28, 37-40] as evidenced by the fact that it receives in excess of 10,000 citations in the literature per annum.[41] A basic schematic of XPS is given in Figure 1.1. An XPS instrument consists of an electron source, e.g.,

a tungsten filament at ca. 10 kV relative to a metal anode, e.g., Al, Mg, or Ag, a monochromator, a flood gun source to compensate for charging of non-conductive samples, and a concentric hemispherical analyzer. XPS works by illuminating samples with X-rays, which eject core level electrons from elements near the surface of the material by the photoelectron effect. The kinetic energies of these ejected electrons are measured by a concentric hemispherical analyzer. These kinetic energies are then converted to binding energies, and XPS spectra are most often plotted as the number of counts obtained at each binding energy. XPS data can be collected either as survey scans (lower resolution over a wide binding energy range) or narrow scans that are collected at higher resolution (over a narrower binding energy range). In general, each element has a unique binding energy for its core level electrons and, hence, XPS helps to identify and quantify the elemental composition of surfaces. [37-39]

In XPS, electrons ejected (formed) in a material lose energy as they travel through it. Beyond three mean free paths, the electrons have a high probability of experiencing an inelastic scattering event, by which they lose energy. Therefore the maximum depth that XPS can probe is ca. three mean free paths, which is generally 5 – 10 nm. The following equation is the fundamental relationship in XPS:[42]

$$h\nu = E_{K.E.} + E_{B.E.} + \phi \tag{1.1}$$

which could be rewritten as

$$E_{B.E.} = h\nu - E_{K.E.} - \phi \tag{1.2}$$

where $h\nu$ is the photon (X-ray) energy, $E_{K.E.}$ and $E_{B.E.}$ are the kinetic and binding energies of the electrons, respectively, and ϕ is the spectrophotometer work function. The most commonly used photon energy is from an Al source: the Al kα line at 1486.7 eV.

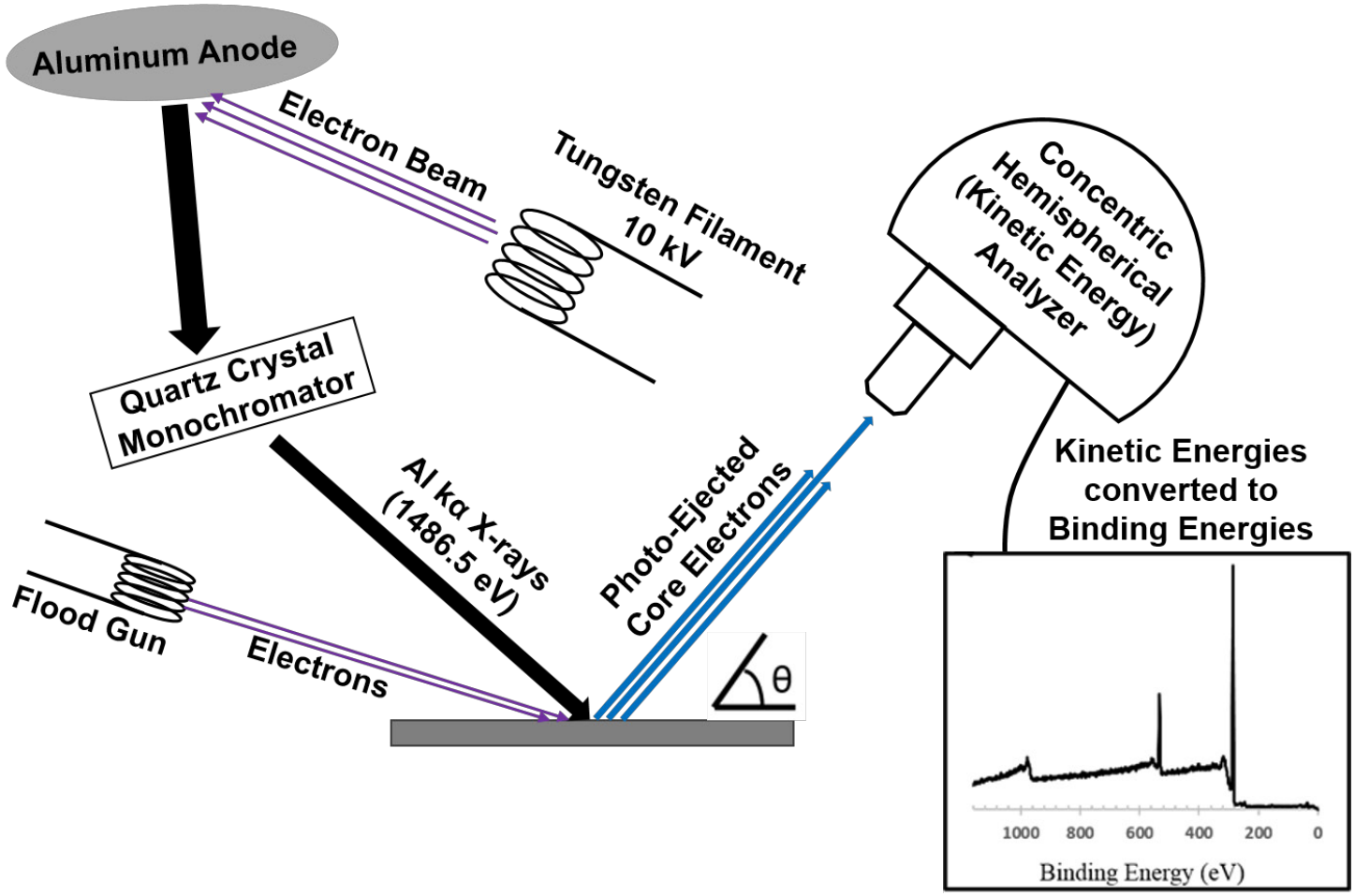


Figure 1.1 Schematic of an X-ray photoelectron spectrometer (XPS). Here, θ is the take-off angle of the electrons.

1.3 Time-of-flight secondary ion mass spectrometry (ToF-SIMS)

Static ToF-SIMS was developed by Alfred Benninghoven in the late 1960's.[43] His initial work on SIMS started at the University of Cologne (Köln), Germany, and further development of a dedicated instrument for static ToF-SIMS took place in the 1970's when Benninghoven moved to the University of Münster.[43] Mass spectrometry is one of the most important techniques for surface chemical analysis, providing elemental and isotopic compositions of materials, and aiding in structural elucidation of compounds.[44-46] As Benninghoven stated: "*Static SIMS supplies chemical information, which cannot be obtained by any other surface analytical technique. This is the most important reason for the wide and rapid growth in the analytical applications of Static SIMS.*"[43] Thus, SIMS is now widely used for the surface analysis of materials.

Figure 1.2 shows a general schematic of a static ToF-SIMS instrument. It consists of a primary ion source, which generates cluster ions, e.g., Bi_3^{2+} or massive clusters of argon atoms, or monatomic ions like Ga^+ or Cs^+. For minimal energy and time dispersion of secondary ions, the primary ion source is pulsed to sub-nanosecond widths, with a pulse separation of ca. 20 ns.[44] The primary ion beam is focused using electromagnetic lenses. A buncher is used to accelerate the lagging ions in a pulse so that all of the ions arrive at a surface within ca. 1 ns of each other. SIMS works by bombarding a sample with energetic *primary* ions, which initiates a cascade of collisions on the surface that results in sputtering of material from the sample. The sputtered material is composed of neutral species, electrons, and a small number of charged species, which are the *secondary* ions. To prevent sample charging, a flood gun illuminates the sample with short pulses of low energy electrons between the primary ion pulses. In the case of a ToF mass analyzer, the secondary ions are accelerated by applying a bias, which feeds them into a field-

free drift tube. This application of bias helps ensure that all the ions initially have the same kinetic energy. Of course there is a spread in the initial kinetic energies of the ions, which gives some deviation in their energies when they enter the mass analyzer. Therefore, a reflectron (an ion mirror) is used to focus the ions that have the same m/z values but different kinetic energies so that they strike the detector at the same time. The time taken by these secondary ions to reach the detector is measured, and this time is converted into an m/z ratio using the following equations. First, the energy of the extracted ion is given as:

$$E = qV \tag{1.3}$$

where E refers to the energy of the ion, and q (or z) is its charge, and V is the voltage with which it is pulsed. Of course this energy is equal to the kinetic energy of the ion, which gives:

$$\frac{1}{2}mv^2 = zV \tag{1.4}$$

where m is the mass of the ion and v is its velocity. Rearranging this equation gives:

$$\frac{m}{z} = \frac{2V}{v^2} \tag{1.5}$$

velocity, v, is distance (d) divided by time (t). Inserting this relationship into Equation 1.5 then yields:

$$\frac{m}{z} = \frac{2Vt^2}{d^2} \tag{1.6}$$

This equation demonstrates that the mass-to-charge ($\frac{m}{z}$) ratio of the ion is directly proportional to the square of the time it takes to reach the detector. Lighter species reach the detector earlier and heavier species later.

Ion detection can be performed with a microchannel plate, a scintillator, and a photomultiplier tube (PMT) in series. The microchannel plate is usually made from a highly resistive material, and is used to convert the ion beam into a pulse of electrons. It consists of a parallel array of ca. 10 μm diameter microchannels separated by ca. 15 μm. The thickness of the microchannel plate is around 2 mm. A single particle arriving at the surface produces significant gain – it generates ca. 1000 electrons at the other face of the microchannel plate (see Figure 1.3). Thus, this device allows for single particle detection. This pulse of electrons hits the scintillator, which absorbs its energy and emits a flash of light. These photons strike the PMTs, which produces an electrical pulse.

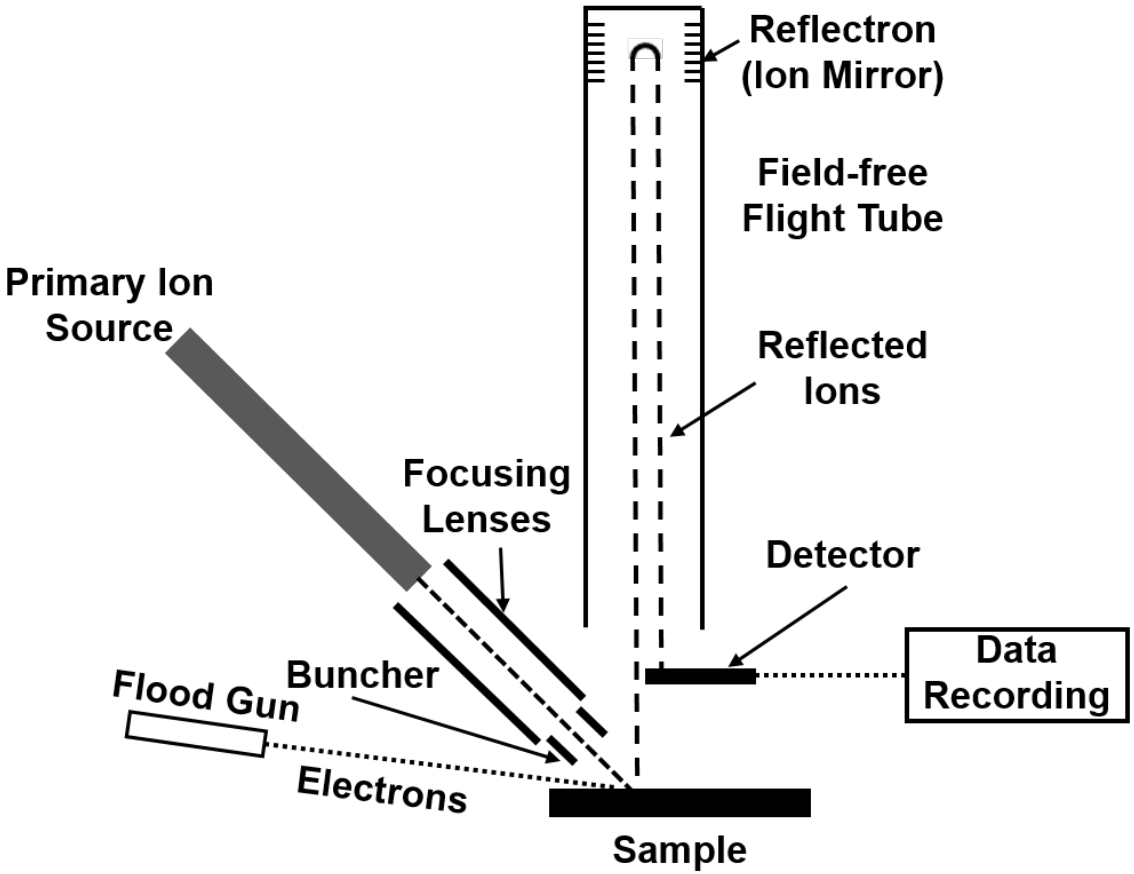


Figure 1.2. Schematic of a static ToF-SIMS instrument.

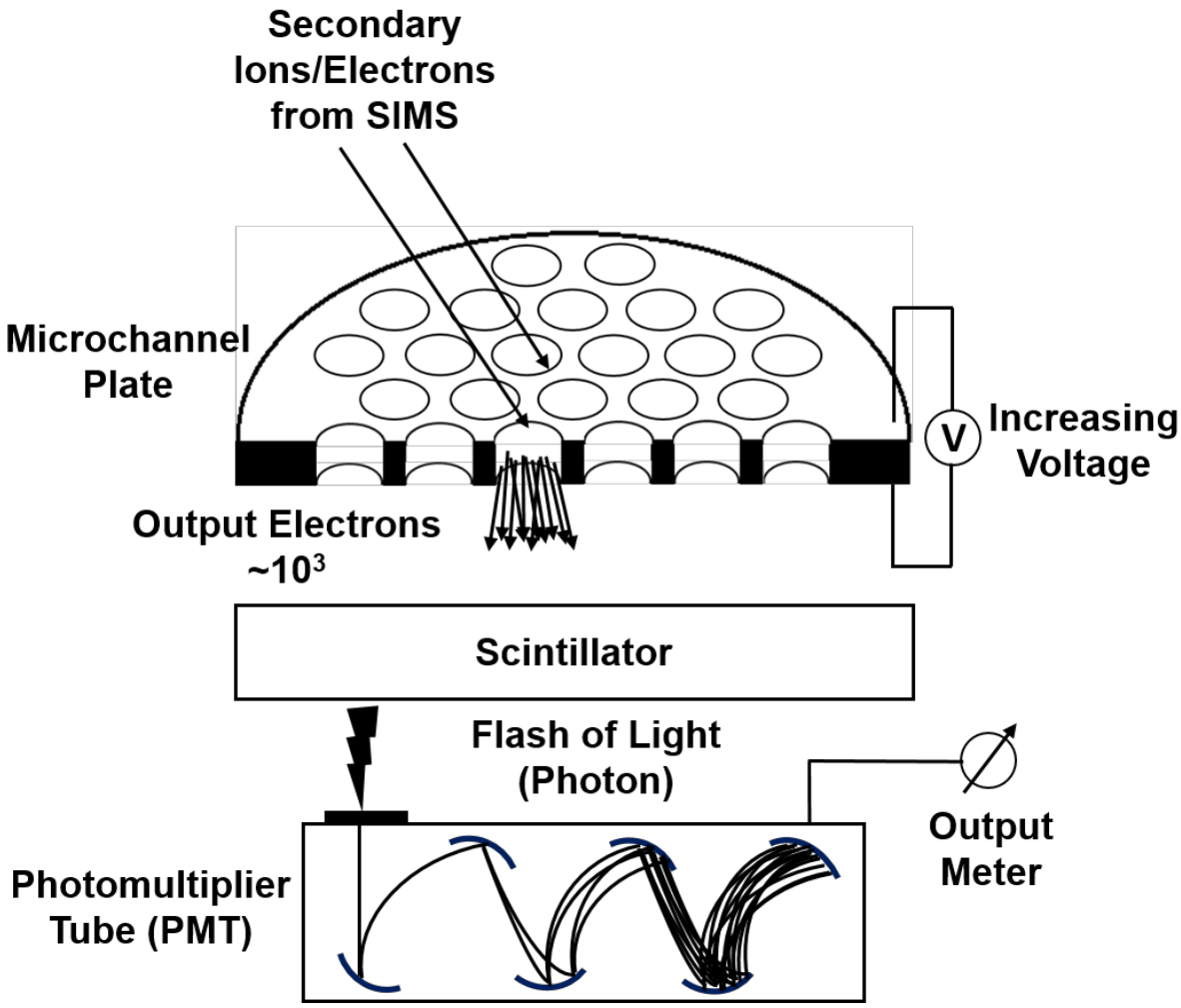


Figure 1.3. Schematic and functioning of a SIMS detector.

The ion detection scheme discussed above assumes one electrical impulse for each ion that strikes the detector. This scenario becomes complicated when more than one ion strikes the detector at the same time, or if an ion strikes the detector before the detector can recover from a previous strike. In these cases, the ions produce a single electrical pulse, or no signal at all, which lead to an underestimation of the signal. Of course, this problem is insignificant for very low signals, but it becomes increasingly important for larger secondary ion yields. Under these conditions, it may be desirable to correct the data, which is typically done by applying a *Poisson* correction. Siméon-Denis Poisson was a French mathematician who derived the distribution that bears his name. This distribution gives the probability of a given number of discrete events happening in a defined time interval, provided that these events occur with a known average rate and are independent of the time of the previous event.[47] The Poisson distribution is given by:

$$f(k,\lambda) = \frac{e^{-\lambda}\lambda^{k}}{k!} \tag{1.7}$$

Here, k is the number of discrete events in a given time interval, e.g., the number of ions striking the detector in a given time interval, and λ is the expected number of events in the time interval, which is simply the total number of events divided by the total time, i.e., in ToF-SIMS, λ is the total number of counts divided by the data acquisition time. Thus, using the Poisson distribution, one can determine the distribution of random events (number of ions predicted to strike the detector surface), i.e., how often one would expect no ion, one ion, two ions, etc. This correction can be applied to intense SIMS peaks that in all likelihood have been underestimated. It goes without saying that it is always better to deal with uncorrected data than to change it mathematically. Accordingly, if one is working in a high signal regime that requires the Poisson

correction, it may be better to lower the primary ion current, which will result in a lower secondary ion yield that does not require correction.

Benninghoven and coworkers were among the very first that worked on developing static SIMS for surface analysis.[48, 49] Before that, dynamic SIMS was the widely used variant of the technique. Static and dynamic SIMS differ in the primary ion dose delivered to the surface, and therefore the depth of surface analyzed. In dynamic SIMS, the surface receives in access of 10^{13} ions/cm^2 which results in significant surface sputtering, i.e., depth profiling. Dynamic SIMS has found immense application in the depth profiling of semiconductors to measure their dopant profiles.[43, 50, 51] Obviously, dynamic SIMS is both destructive and insensitive to surfaces.[45] In static SIMS, the surface is bombarded with a relatively low dose of primary ions (10^{11} - 10^{13} ions/cm^2), which is known as the static regime. Under static conditions, the probability of the same spot being reanalyzed (struck twice) by the beam is negligible, "*no spot on the surface should receive more than one primary ion strike*."[43] This helps to keep the majority of the surface unaffected by the effects of sputtering. Hence, the surface over which a spectrum is collected in static SIMS is essentially pristine. With regards to the depth sensitivity of static SIMS, Benninghoven noted that: *"...over 95 % of the secondary particles originate from the top two layers of the solid.*"[43] This capability of static SIMS has made it the technique of choice for comprehensive chemical and structural elucidation of the top few atomic layers of a sample. In the early 80's, static SIMS began to find immense application in the analysis of organic and polymer materials, especially after the development of charge neutralization techniques.[52, 53]

In spite of its remarkable versatility and importance, SIMS has some drawbacks. One of them is that ion yields are very strongly affected by surface chemistry and, hence, SIMS is not generally a quantitative technique. The strong effect of the surface/matrix on ion yield is known

as the matrix effect of SIMS. For example, the yield of Na^+ from a sample of sodium chloride (NaCl) will be much higher than the Na^+ signal from pure sodium metal because the sodium is already in an ionic state in NaCl. Therefore, SIMS and XPS, which is quantitative, are highly complementary techniques. Second, SIMS, in its various modes, can yield *too much information* about a sample.[54] This is especially true in its imaging modes where it collects a complete mass spectrum per pixel. A typical scan might consist of 128 x 128 pixels, yielding ca. 16,000 spectra. Obviously, this quantity of data could require weeks to comprehensively analyze.

1.4 Chemometrics techniques

1.4.1 XPS Data Analysis

The literature shows various methodologies for XPS data analysis,[55] which include traditional peak fitting of the XPS peak envelope, the use of the Fourier transform,[56] and taking derivatives of the spectra.[57, 58] In particular, derivatives of XPS spectra have been used by Pavlath and Millard to efficiently determine peak positions in a curve fitting analysis.[57] However, in general, taking the derivative of a spectrum increases the noise in it and often requires some type of smoothing. Indeed, the derivative approach fails to provide peak positions for complex peak envelopes that consist of numerous overlapping peaks. The Fourier transform and derivative approaches are not widely used by the community. By far the most commonly used data analysis approach in XPS is peak fitting.

Apart from determining surface elemental compositions, XPS is important for finding the chemical states of the elements it detects, and this analysis is done through peak fitting. For example, a carbon 1s peak envelope can often be peak fitted to five different chemical states attributable to C-C (C(0)), C-O (C(I)), C=O (C(II)), O=C-O (C(III)), along with shake-up peaks

($\pi - \pi^*$ transitions).[59-62] This information about functional groups is vital for understanding a surface, and therefore for devising new surface and material chemistries for specific applications.[41] Peak fitting involves (i) subtracting/fitting the background caused by inelastically scattered electrons, and (ii) resolving the various chemical states that constitute the peak envelope.[63] I will now discuss the challenges associated with each of these steps:

Background approximation for XPS peak envelopes

[**]The choice of background approximation poses a challenge for attaining good quality peak fits.[63, 64] For simpler spectra, a linear background approximation can be a reasonable first choice. However, XPS spectra often show greater complexity because the ejected photoelectrons lose energy while traveling through the substrate, causing significant backgrounds. These backgrounds depend on the substrates that are analyzed, and care should be taken to choose the appropriate background approximation. To address the issue of complex background shapes, several more sophisticated approximation methods are frequently employed. The most common of these are the Shirley[65] and Tougaard[66-68] backgrounds. There are several reports in the literature that describe modified versions of these backgrounds,[63, 69, 70] which speaks of a general lack of agreement on their merits and deficiencies. As noted in the CasaXPS software manual, "*The truth is that none of the background types on offer are correct and therefore selection of one background type over another is essentially chosen as the least wrong rather than the most right.*"[71] In summary, one is never entirely sure of one's choice of background, although with some care and experience, quite good peak fitting results can often be obtained.

[**]This paragraph has been reproduced with permission from (Bhupinder Singh, Ronald Hesse, and Matthew R. Linford), *Vacuum Technology and Coating*, December, 2015: p. 22-29

Resolving (overlapping) chemical states in the XPS peak envelope

Once a background correction has been applied, peak fitting involves specifying the number of constituent peaks in the overall peak envelope, and then placing reasonable and well informed constraints on their positions, heights, widths, asymmetries, and Gaussian : Lorentzian (G:L) ratios. A reasonable knowledge of both the sample being tested and its preparation, along with a thorough literature search, helps to place reasonable constraints on the above-mentioned parameters. Once reasonable constraints are applied, one allows the software to optimize the fit parameters. Of course, the fewest possible parameters should be chosen for any given fit. The resulting fit envelope is the sum of the constituent components (peaks), which correspond to the various chemical states of the material. Modern software is capable of identifying the quality of a fit through a comparison of the experimental and fit envelopes. The difference between the two, i.e., the inability of the fit envelope to match the experimental peak envelope, is determined as a fit error. The chi-squared (χ^2) values, the Abbe criterion, and the peak residuals can be used to determine and even quantify the goodness of a fit.

A low error implies a good fit, by which we mean close agreement between one's experimental data and the corresponding fit, e.g., a low χ^2 value. However, a close fit between model and experimental results does not guarantee that a fit is reasonable. For example, one might allow a large number of parameters to float in one's model, which should give a low χ^2 value, but a high degree of correlation between fit parameters. Accordingly, one might ask: how does one know if one's peak fitting is correct? As noted by Sherwood, there is not, in general, a unique solution to fitting XPS narrow scans.[72] Of course, very simple scans may show good uniqueness, but peak fitting for more complex spectra is challenging. Thus, one proceeds using one's experience, reliance on sound principles, the experience of others (literature research), and

common sense. Often, naive users and, on occasion, even more experienced practitioners allow numerous parameters to float, which may, again, yield, a low χ^2 value, but a completely unreasonable fit. I discuss various good practices for performing XPS peak fits in Chapter 2 of this thesis.

To address the issue of subjectivity in XPS peak fitting, I have introduced mathematical functions to characterize XPS narrow scans. These include the equivalent width and the autocorrelation width, which are compared to the variance and (of course) the standard deviation. These new XPS data analysis tools were published as two separate papers in the *Journal of Electron Spectroscopy and Related Phenomena*, [59, 60] and are discussed in Chapters 3 and 4. I also wrote an article in *Vacuum Technology & Coating* magazine that illustrates the mathematics of equivalent width and autocorrelation width, and their applicability to XPS narrow scans. This is discussed in Appendix 1.

Chapter 5 of this thesis discusses uniqueness testing applied to XPS narrow scans. Uniqueness testing has been extensively used in spectroscopic ellipsometry to determine the quality of a fit, and to see if fit parameters are correlated. However, to the best of my knowledge, this graphical tool has never been used in XPS. Of course, the mathematical analysis and calculation of fit parameter correlation is known and carried out in SE and XPS data analysis.[73, 74] However, as noted, the uniqueness plot, which is a useful and powerful graphical tool that can be very easily interpreted has not been applied to XPS data analysis. There are many examples of poor XPS peak fits in the literature. This type of tool could rather easily prevent some of these mistakes from occurring.

1.4.2 ToF-SIMS Data Analysis

The problem of dealing with massive quantities of data/information, especially in ToF-SIMS, can be approached by using chemometrics techniques,[6, 54, 75-80] which are pattern recognition tools that help to identify and group similar spectra, e.g., samples or pixels. Two of the most commonly used chemometrics tools for ToF-SIMS data analysis are principal components analysis (PCA)[81] and cluster analysis.

In PCA, each data set, e.g., spectrum, can be considered to be a single point in a hyperspace, where, in effect, PCA rotates the axes of the coordinate system to capture the largest amount of the variation in the data points (spectra). The new (rotated) axes are termed principal components (PCs). The axis that captures the greatest degree of variation in the data is the first principal component, PC1. PC2 then captures the next largest amount of variation in the data, etc. The projections of the original data points onto the new, rotated axes (PCs) are called scores, and the contributions of the original axes (variables) to the PCs are called loadings. Scores plots often identify trends and clustering of spectra, while loadings plots reveal the chemical bases for the differences between the spectra observed in scores plots. Obviously, PCA maintains the orthogonality of the original axes. When there is some degree of correlation between the peaks/information within spectra, which occurs for many data sets, PCA often results in a significant reduction in the number of variables (PCs) required to describe a data set.[82]

PCA is widely employed. It is often the first chemometrics tool used to analyze a data set, and the combination of scores and loadings plots can be powerful. Nevertheless, PCA scores and loadings plots, which often involve negative values (peaks), can be difficult for the uninitiated to interpret. Cluster analysis is also used to identify similarities and differences between spectra. Cluster analysis uses a different algorithm than PCA, where, fundamentally, it involves finding

the distances between data points (vectors) in a hyperspace, connecting (clustering) the data points based on the distances between them, and then representing these distances in the form of a plot called a dendrogram. While a PCA model will generally include most of the information in a data set (the remainder of it is often noise), cluster analysis considers all of the data. Dendrograms can be easier to interpret than scores and loadings plots, although they lack the chemical information provided by loadings plots. Nevertheless, in the case of complex, varied spectra, they too can be challenging to interpret. Thus, because both techniques may struggle to fully and easily explain the variation in complex data sets, analyses based on different algorithms that 'view' the data differently can be important.[75, 79, 80] And of course, different data sets will often be better suited to different chemometrics methods.

It follows from this discussion that there is room for additional statistical analysis tools based on new algorithms that may 'perceive' data in new ways. Chapter 6 discusses the important issues concerning the analysis of complex data, using examples of ToF-SIMS and XPS depth profiles. There, I focus on three chemometrics techniques that can be used to analyze whole spectra. The first tool I discuss is principal components analysis (PCA).[81] The second tool is cluster analysis. The third data analysis tool I discuss is new. It is the information content (IC), or entropy, of a spectrum. This approach comes out of the groundbreaking work of Shannon,[83, 84] who applied the statistical thermodynamics definition of entropy to signal/communication theory. Here I take Shannon's approach back into the physical sciences, calculating the IC of spectra and using the resulting values to determine where changes are taking place in spectra. Each chemometrics method is applied to the same four data sets: the ToF-SIMS depth profile of ca. 100 nm of plasma polymerized C_3F_6 on Si, the ToF-SIMS depth profile of ca. 100 nm of plasma polymerized PNIPAM (poly(N-isopropylacrylamide)) also deposited on a Si substrate,

the XPS depth profile of a film of SiO_2 on Si, and the XPS depth profile of a film of Ta_2O_5 on Ta.

1.5 Multi-instrument material characterization of nanodiamonds used in high pressure liquid chromatography (HPLC)

Silica-based particles have dominated the HPLC market for decades, showing high separation efficiencies within a relatively limited range of mobile phase pH values and temperatures.[85] Under acidic conditions (pH < 2), siloxane bonds between silane adsorbents in bonded phases and the underlying silica are prone to hydrolysis, and under basic conditions (pH > 8) the silica support dissolves.[85-90] Elevated temperatures accelerate these degradation modes.[87, 88, 91-93] Considering the advantages of separations performed at elevated temperatures and extremes of pH, and the inherent instability of silica under these conditions, it has become important to evaluate alternative stationary phases for liquid chromatography.

[***]Diamond exhibits exceptional mechanical stability, even at high pressures, [12, 94-98] is hydrolytically stable over the entire pH range,[12, 85, 94-96] shows excellent chemical inertness[12, 85, 98, 99] and outstanding thermal stability,[12, 95, 97, 98] has a low coefficient of thermal expansion,[12] does not shrink or swell in aqueous or organic solvents,[94-97] and provides a solid platform for chemical modification at its surface.[94, 97, 98] Accordingly, various researchers have become interested in diamond as an alternative stationary phase and/or support for liquid chromatography,[96, 97, 100] including those in the Linford group at Brigham Young University.[36, 85, 99, 101] The superficially porous particles developed at BYU are made by the alternating deposition of poly(allylamine) (PAAm) and nanodiamond (ND) onto carbon core particles in a layer-by-layer (LbL) fashion.[85, 101] Each LbL assembly is terminated with a layer of PAAm that is functionalized and cross-linked with a mixture of 1,2-epoxyoctadecane and 1,2,7,8-

diepoxyoctane. The resulting mixed-mode particles show weak anion exchange (WAX) and hydrophobic properties, where the dominant mode of these particles depends on the mobile phase pH.[36, 101] These particles have been steadily improved, showing efficiencies up to 112,000 N/m with a reduced plate height, *h* of 2.22.[36] As expected, the resulting columns are stable under harsh conditions: 120 °C and pH 11.3, and have good selectivity, e.g., pull apart critical pairs of acidic herbicides at pH 2 and separate basic pharmaceuticals at pH 12.[102-104]

[***]Extensive characterization of a material increases one's understanding of it in a way that should lead to product improvement.[5, 18] As discussed earlier, there is no stand-alone technique that can comprehensively characterize a material and, therefore, comprehensive material characterization involves a multi-instrument approach. Chapter 7 discusses the multi-instrument characterization of five different nanodiamond (ND) samples that were used to make core shell particles for HPLC. The nanodiamonds were characterized by a suite of techniques including X-ray photoelectron spectroscopy (XPS), time-of-flight secondary ion mass spectrometry (ToF-SIMS), diffuse reflectance infrared Fourier transform spectroscopy (DRIFT), transmission electron microscopy (TEM), electron energy loss spectroscopy (EELS), X-ray diffraction (XRD), inductively coupled plasma-mass spectrometry (ICP-MS), and Brunauer-Emmett-Teller (BET) surface area analysis. Chemometrics techniques were employed to better understand the ToF-SIMS data. To the best of my knowledge, this is the most comprehensive characterization study of nanodiamonds undertaken to date. This chapter emphasizes the lack of comprehensive characterization of nanomaterials reported in the literature, and presents this study as a model/exemplary approach to nanomaterial characterization.

[***]These paragraphs are reproduced from (Bhupinder Singh, Stacey J. Smith, David S. Jensen, Hodge F. Jones, Andrew E. Dadson, Paul B. Farnsworth, Richard Vanfleet, Jeffrey K. Farrer, and Matthew R. Linford) *Analytical and Bioanalytical Chemistry*, accepted for publication, 2015.

I also worked on developing applications, understanding the mechanisms of retention and probing reproducibility of nanodiamond-based HPLC columns. Appendix 2 presents the application of these diamond-based mixed mode columns to separate acidic herbicides. Appendix 3 shows the comparison of these diamond-based columns to their commercial counterparts, especially for separation of critical pairs of acidic herbicides. Appendix 4 reports on their retention mechanism, and Appendix 5 presents the reproducibility and stability of these diamond-based columns, respectively. Appendix 2-5 were published as commercial application notes on diamond-analytics.com.

1.6 References

1. D. K. Schroder, *Semiconductor Material And Device Characterization*, Third ed. (John Wiley & Sons, Inc., New Jersey, 2006).

2. C. B. Murray, D. J. Norris and M. G. Bawendi, J. Am. Chem. Soc. **115** (19), 8706-8715 (1993).

3. B. Ferguson and X.-C. Zhang, Nat Mater **1** (1), 26-33 (2002).

4. O. Jani, I. Ferguson, C. Honsberg and S. Kurtz, Appl Phys Lett **91** (13), 132117 (2007).

5. D. S. Jensen, S. S. Kanyal, N. Madaan, J. M. Hancock, A. E. Dadson, M. A. Vail, R. Vanfleet, V. Shutthanandan, Z. H. Zhu, M. H. Engelhard and M. R. Linford, Surf Interface Anal **45** (8), 1273-1282 (2013).

6. C. H. Hung, B. Singh, M. G. Landowski, M. Ibrahim, A. J. Miles, D. S. Jensen, M. A. Vail, A. E. Dadson, S. J. Smith and M. R. Linford, Surf Interface Anal **47** (8), 815-823 (2015).

7. M. Mesa, L. Sierra, B. López, A. Ramirez and J.-L. Guth, Solid State Sciences **5** (9), 1303-1308 (2003).

8. A. Krueger, J Mater Chem **18** (13), 1485-1492 (2008).

9. I. I. Kulakova, Phys Solid State+ **46** (4), 636-643 (2004).

10. V. F. Loktev, V. I. Makal'skii, I. V. Stoyanova, A. V. Kalinkin, V. A. Likholobov and V. N. Mit'kin, Carbon **29** (7), 817-819 (1991).

11. D. F. Mitev, A. T. Townsend, B. Paull and P. N. Nesterenko, Carbon **60**, 326-334 (2013).

12. P. N. Nesterenko and P. R. Haddad, Anal Bioanal Chem **396** (1), 205-211 (2010).

13. G. Saini, K. Sautter, F. E. Hild, J. Pauley and M. R. Linford, J. Vac. Sci. Technol., A **26** (5), 1224-1234 (2008).

14. G. L. Jiang, F. Rivera, S. S. Kanyal, R. C. Davis, R. Vanfleet, B. M. Lunt, V. Shutthanandan and M. R. Linford, Opt Eng **50** (1) (2011).

15. B. Bhushan, *Tribology and Mechanics of Magnetic Storage Devices*. (Springer-Verlag, New York, 1990).

16. B. M. Lunt, M. R. Linford, R. C. Davis, S. Jamieson, A. Pearson and H. Wang, Archiving Conference **2013** (1), 132-136 (2013).

17. H. Wang, N. Madaan, J. Bagley, A. Diwan, Y. Liu, R. C. Davis, B. M. Lunt, S. J. Smith and M. R. Linford, Thin Solid Films **569**, 124-130 (2014).

18. H. Wang, B. M. Lunt, R. J. Gates, M. C. Asplund, V. Shutthanandan, R. C. Davis and M. R. Linford, ACS Appl. Mater. Interfaces **5** (17), 8407-8413 (2013).

19. D. B. Warheit, Toxicological Sciences **101** (2), 183-185 (2008).

20. H. Zhuang and J. A. J. Gardella, MRS Bulletin **21** (01), 43-48 (1996).

21. A. M. Almanza-Workman, S. Raghavan, S. Petrovic, B. Gogoi, P. Deymier, D. J. Monk and R. Roop, Thin Solid Films **423** (1), 77-87 (2003).

22. W. A. Daoud, J. H. Xin and X. Tao, Appl Surf Sci **252** (15), 5368-5371 (2006).

23. S. Brandriss and S. Margel, Langmuir **9** (5), 1232-1240 (1993).

24. P. Favia, G. Cicala, A. Milella, F. Palumbo, P. Rossini and R. d'Agostino, Sur. Coat. Technol. **169–170**, 609-612 (2003).

25. S. Veeramasuneni, J. Drelich, J. D. Miller and G. Yamauchi, Progress in Organic Coatings **31** (3), 265-270 (1997).

26. P. F. Kane and G. B. Larrabee, *Characterization of Solid Surfaces*, First ed. (Plenum Press, New York, 1974).

27. D. Groves and J. Wachtman, Jr., in *Materials Characterization for Systems Performance and Reliability*, edited by J. McCauley and V. Weiss (Springer US, 1986), Vol. 26, pp. 425-443.

28. J. M. Beck, R. T. Farouki and J. K. Hinds, Computer Graphics and Applications, IEEE **6** (12), 18-36 (1986).

29. L. A. Casper and C. J. Powell, *Industrial Applications of Surface Analysis*. (American Chemical Society, New York, 1982).

30. J. M. Walls, *Methods of Surface Analysis: Techniques and Applications*. (Cambridge University Press, New York, 1989).

31. D. Briggs and M. P. Seah, *Practical Surface Analysis*. (John Wiley & Sons, New York, 1990).

32. D. J. O'Connor, B. A. Sexton and R. S. C. Smart. (Springer-Verlag, Berlin, Germany, 1992).

33. D. R. Baer, D. J. Gaspar, P. Nachimuthu, S. D. Techane and D. G. Castner, Anal Bioanal Chem **396** (3), 983-1002 (2010).

34. B. E. Erickson, in *Chem Eng News* (2008), Vol. 86, pp. 25-26.

35. *Surface Characterization Methods: Principles, techniques, and Applications*. (Marcel Dekker Inc., New York, 1999).

36. C.-H. Hung, L. A. Wiest, B. Singh, A. Diwan, M. J. C. Valentim, J. M. Christensen, R. C. Davis, A. J. Miles, D. S. Jensen, M. A. Vail, A. E. Dadson and M. R. Linford, J Sep Sci **36** (24), 3821-3829 (2013).

37. C. S. Fadley, R. J. Baird, W. Siekhaus, T. Novakov and S. Å. L. Bergström, J Electron Spectrosc **4** (2), 93-137 (1974).

38. C. J. Powell and P. E. Larson, Applications of Surface Science **1** (2), 186-201 (1978).

39. R. S. Swingle, Anal Chem **47** (1), 21-24 (1975).

40. M. P. Seah, Vacuum **34** (3–4), 463-478 (1984).

41. B. Singh and M. R. Linford, in *Vacuum Technology & Coating* (October, 2015).

42. J. F. Watts, *An Introduction to Surface Analysis by Electron Spectroscopy*. (Oxford Science Publications, New York, 1990).

43. *ToF-SIMS: Surface Analysis by Mass Spectrometry*. (IM Publications and Surface Scince Spectra, West Sussex and Manchester, 2001).

44. A. M. Belu, D. J. Graham and D. G. Castner, Biomaterials **24** (21), 3635-3653 (2003).

45. P. Bertrand and W. Lu-Tao, in *Microbeam and Nanobeam Analysis*, edited by D. Benoit, J.-F. Bresse, L. Van't dack, H. Werner and J. Wernisch (Springer Vienna, 1996), Vol. 13, pp. 167-182.

46. A. Benninghoven, Surf Interface Anal **43** (1-2), 2-11 (2011).

47. F. A. Haight, *Handbook of the Poisson Distribution*. (John Wiley & Sons, New York, 1967).

48. A. Benninghoven, Surf Sci **35**, 427-457 (1973).

49. A. Benninghoven, Journal Name: Z. Phys. 230: 403-17(1970).; Other Information: Orig. Receipt Date: 31-DEC-70, Medium: X (1970).

50. L. Van Vaeck, A. Adriaens and R. Gijbels, Mass Spectrometry Reviews **18** (1), 1-47 (1999).

51. G. Gillen, L. King, B. Freibaum, R. Lareau, J. Bennett and F. Chmara, J. Vac. Sci. Technol., A **19** (2), 568-575 (2001).

52. D. Briggs and A. B. Wootton, Surf Interface Anal **4** (3), 109-115 (1982).

53. A. Brown and J. C. Vickerman, Surf Interface Anal **6** (1), 1-14 (1984).

54. C. Brüning, S. Hellweg, S. Dambach, D. Lipinsky and H. F. Arlinghaus, Surf Interface Anal **38** (4), 191-193 (2006).

55. A. Proctor and P. M. A. Sherwood, Anal Chem **54** (1), 13-19 (1982).

56. P. Griffiths, in *Transform Techniques in Chemistry*, edited by P. Griffiths (Springer US, 1978), pp. 1-9.

57. A. E. Pavlath and M. M. Millard, Appl Spectrosc **33** (5), 502-509 (1979).

58. T. Inouye, T. Harper and N. C. Rasmussen, Nuclear Instruments and Methods **67** (1), 125-132 (1969).

59. B. Singh, D. Velázquez, J. Terry and M. R. Linford, J Electron Spectrosc **197** (0), 56-63 (2014).

60. B. Singh, D. Velázquez, J. Terry and M. R. Linford, J Electron Spectrosc **197** (0), 112-117 (2014).

61. T. I. T. Okpalugo, P. Papakonstantinou, H. Murphy, J. McLaughlin and N. M. D. Brown, Carbon **43** (1), 153-161 (2005).

62. G. Zhang, S. Sun, D. Yang, J.-P. Dodelet and E. Sacher, Carbon **46** (2), 196-205 (2008).

63. M. Aronniemi, J. Sainio and J. Lahtinen, Surf Sci **578** (1–3), 108-123 (2005).

64. J. E. Castle, H. Chapman-Kpodo, A. Proctor and A. M. Salvi, J Electron Spectrosc **106** (1), 65-80 (2000).

65. D. A. Shirley, Phys Rev B **5** (12), 4709-4714 (1972).

66. S. Tougaard, Phys Rev B **34** (10), 6779-6783 (1986).

67. S. Tougaard, Surf Sci **216** (3), 343-360 (1989).

68. M. P. Seah, Surf Sci **420** (2–3), 285-294 (1999).

69. J. Végh, J Electron Spectrosc **151** (3), 159-164 (2006).

70. J. Végh, Surf Sci **563** (1–3), 183-190 (2004).

71. (CasaXPS, http://www.casaxps.com/help_manual/manual_updates/peak_fitting_in_xps.pdf, 2006).

72. P. M. A. Sherwood, J. Vac. Sci. Technol., A **14** (3), 1424-1432 (1996).

73. V. A. Yakovlev and E. A. Irene, J Electrochem Soc **139** (5), 1450-1455 (1992).

74. R. Hesse, T. Chasse, P. Streubel and R. Szargan, Surf Interface Anal **36** (10), 1373-1383 (2004).

75. L. Pei, G. Jiang, B. J. Tyler, L. L. Baxter and M. R. Linford, Energ Fuel **22** (2), 1059-1072 (2008).

76. X. V. Eynde and P. Bertrand, Surf Interface Anal **25** (11), 878-888 (1997).

77. O. D. Sanni, M. S. Wagner, D. Briggs, D. G. Castner and J. C. Vickerman, Surf Interface Anal **33** (9), 715-728 (2002).

78. L. Yang, R. Bennett, J. Strum, B. Ellsworth, D. Hamilton, M. Tomlinson, R. Wolf, M. Housley, B. Roberts, J. Welsh, B. Jackson, S. Wood, C. Banka, C. Thulin and M. Linford, Anal Bioanal Chem **393** (2), 643-654 (2009).

79. L. Yang, Y. Y. Lua, G. L. Jiang, B. J. Tyler and M. R. Linford, Anal Chem **77** (14), 4654-4661 (2005).

80. F. Zhang, R. J. Gates, V. S. Smentkowski, S. Natarajan, B. K. Gale, R. K. Watt, M. C. Asplund and M. R. Linford, J Am Chem Soc **129** (30), 9252-9253 (2007).

81. I. Jolliffe, in *Wiley StatsRef: Statistics Reference Online* (John Wiley & Sons, Ltd, 2014).

82. S. Raychaudhuri, J. M. Stuart and R. B. Altman, Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing, 455-466 (2000).

83. C. E. Shannon, At&T Tech J **27** (4), 623-656 (1948).

84. E. T. Jaynes, Phys Rev **106** (4), 620-630 (1957).

85. G. Saini, D. S. Jensen, L. A. Wiest, M. A. Vail, A. Dadson, M. L. Lee, V. Shutthanandan and M. R. Linford, Anal Chem **82** (11), 4448-4456 (2010).

86. J. J. Kirkland, M. A. van Straten and H. A. Claessens, J Chromatogr A **691** (1–2), 3-19 (1995).

87. H. A. Claessens, M. A. van Straten and J. J. Kirkland, J Chromatogr A **728** (1–2), 259-270 (1996).

88. T. Teutenberg, K. Hollebekkers, S. Wiese and A. Boergers, J Sep Sci **32** (9), 1262-1274 (2009).

89. J. M. Miller, *Chromatography: Concepts and Contrasts*, 2 ed. (John Wiley & Sons, Hoboken, NJ, 2005).

90. J. Nawrocki, C. Dunlap, A. McCormick and P. W. Carr, J Chromatogr A **1028** (1), 1-30 (2004).

91. B. Yan, J. Zhao, J. S. Brown, J. Blackwell and P. W. Carr, Anal Chem **72** (6), 1253-1262 (2000).

92. E. Kikta Jr and E. Grushka, Anal Chem **48** (8), 1098-1104 (1976).

93. H. Colin, J. C. Diez-Masa, G. Guiochon, T. Czjkowska and I. Miedziak, J Chromatogr A **167** (0), 41-65 (1978).

94. P. N. Nesterenko and O. N. Fedyanina, J Chromatogr A **1217** (4), 498-505 (2010).

95. P. N. Nesterenko, O. N. Fedyanina and Y. V. Volgin, Analyst **132** (5), 403-405 (2007).

96. P. N. Nesterenko, O. N. Fedyanina, Y. V. Volgin and P. Jones, J Chromatogr A **1155** (1), 2-7 (2007).

97. O. N. Fedyanina and P. N. Nesterenko, Russ J Phys Chem a+ **84** (3), 476-480 (2010).

98. G. Yushin, E. N. Hoffman, M. W. Barsoum, Y. Gogotsi, C. A. Howell, S. R. Sandeman, G. J. Phlllips, A. W. Lloyd and S. V. Mikhalovsky, Biomaterials **27** (34), 5755-5762 (2006).

99. G. Saini, L. A. Wiest, D. Herbert, K. N. Biggs, A. Dadson, M. A. Vail and M. R. Linford, J Chromatogr A **1216** (16), 3587-3593 (2009).

100. V. S. Bondar', I. O. Pozdnyakova and A. P. Puzyr', Phys Solid State+ **46** (4), 758-760 (2004).

101. L. A. Wiest, D. S. Jensen, C. H. Hung, R. E. Olsen, R. C. Davis, M. A. Vail, A. E. Dadson, P. N. Nesterenko and M. R. Linford, Anal Chem **83** (14), 5488-5501 (2011).

102. L. A. Wiest, D. S. Jensen, A. J. Miles, A. Dadson and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/%CE%B22-Agonists%20and%20Amphetamines.pdf, 2013).

103. L. A. Wiest, D. S. Jensen, A. J. Miles, A. E. Dadson and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/FlareMixedModeColumn-TriazineHerbicides.pdf, 2013).

104. B. Singh, D. S. Jensen, A. J. Miles, A. E. Dadson and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/Acidic%20Herbicides.pdf, 2013).

Chapter 2: Good Practices for XPS (and Other Types of) Peak Fitting. Use Chi Squared, Use the Abbe Criterion, Show the Sum of Fit Components, Show the (Normalized) Residuals, Choose an Appropriate Background, Estimate Fit Parameter Uncertainties, Limit the Number of Fit Parameters, Use Information from Other Techniques, and Use Common Sense [*]

2.1 Introduction

Some form of peak fitting or data modeling is used to analyze the results from most surface and materials characterization methods. For example, most spectroscopic ellipsometry studies involve at least some data modeling.[1-6] In this column we will focus on peak fitting in X-ray photoelectron spectroscopy (XPS), which is an indispensable part of XPS narrow scan analysis.[7] As we have discussed previously, XPS is a widely used surface characterization technique that probes the upper ca. 5 – 10 nm of a surface.[8-12] XPS yields the elemental compositions of materials, along with information about the oxidation states of the elements in them. Peak fitting in XPS often helps determine the fraction of an element in a certain oxidation state.[13] This information can be vital for understanding a surface or a material. However, peak fitting in XPS generally suffers from at least some user bias/subjectivity. Accordingly, it is important to apply sound principles in one's peak fitting efforts. Unfortunately, the literature contains quite a few examples of poor fitting of XPS narrow scans. On a more positive note, there are also good examples of peak fitting from a number of research groups. It is useful to be able to discern between the better fits and the less reasonable ones.

[*]This chapter has been reproduced with permission from (Bhupinder Singh, Ronald Hesse, and Matthew R. Linford), *Vacuum Technology and Coating*, December, 2015: p. 22-29

In this article we will discuss the following principles for good peak fitting:

(i) Use chi squared

(ii) Use the Abbe criterion

(iii) Show the sum of your fit components with your original data

(iv) Show the (normalized) residuals

(v) Choose an appropriate background

(vi) Estimate the fit-parameter uncertainties (if available in your software)[14]

(vii) Constrain one's fit parameters reasonably, and also limit the number of fit parameters

(viii) Consider information from other techniques, and also other narrow scans within ones' analysis

(ix) Use common sense

2.2 Theory

We introduce here the mathematical formulas for chi squared, the (normalized) residuals, and the Abbe criterion. These peak-fit quantities have been discussed in detail in a series of publications by Hesse and coworkers, and others.[7, 15-17]

2.2.1 Chi squared (χ^2)

The chi squared (χ^2) quantity gives a measure of the goodness of fit between a set of experimental data points, *M(i)*, and the sum of the fit components/peaks used to model that data, $S(i, \vec{p})$, where the $\vec{p}$ vector is the set of parameters upon which the fit components depend. As we can see in Equation 2.1, χ^2 is calculated in a point-by-point fashion by squaring the difference between each data point, *M(i),* and the value of the fit at that same point: $S(i, \vec{p})$:

$$\chi^2(\vec{p}) = \sum_{i=1}^{N} \frac{\left[M(i) - S(i,\vec{p})\right]^2}{M(i)} \tag{2.1}$$

Here, N is the number of experimental data points collected and fit. Obviously, we want the fit at each point to be as close as possible to the experimental data, i.e., we want $\frac{\left[M(i) - S(i,\vec{p})\right]^2}{M(i)}$ to be as small as possible at each value of i. Thus, smaller values of χ^2 correspond to better fits to the experimental data. Note that (i) $M(i) - S(i,\vec{p})$ is squared in the formula for χ^2 to ensure that positive and negative deviations between the fit and the original data do not cancel each other in the summation (we account for all the differences between the original data and the fit to it by squaring them), and (ii) the $[M(i) - S(i,\vec{p})]^2$ term is normalized by the magnitude of the experimental data point: $M(i)$, i.e., a difference between the data point and the fit: $M(i) - S(i,\vec{p})$ will matter more in χ^2 if $M(i)$ is small.

2.2.2 The reduced chi squared (χ^{2*}) value

The reduced chi squared value (χ^{2*}) may also be used to assess goodness of fit:

$$\chi^{2*}(\vec{p}) = \frac{\chi^2(\vec{p})}{N - P} \tag{2.2}$$

This is simply χ^2 divided by $N - P$, where P is the number of independent fit parameters. Clearly, $N - P \approx N$ for $N >> P$.

2.2.3 *The Residual*

Technically, the residual is simply the difference between the fit and the experimental data point at the same point:

$$R(i) = S(i, \vec{p}) - M(i) \tag{2.3}$$

Normalized residuals, $R_N(i)$, are normalized to the value of $M(i)$ as follows:

$$R_N(i) = \frac{S(i, \vec{p}) - M(i)}{\sqrt{M(i)}} \tag{2.4}$$

Note that Equation 2.4 is essentially the square root of the argument in the summation in Equation 2.1. The percent residual has also been used by various researchers:

$$\%R(i) = \frac{S(i,p) - M(i)}{M(i)} * 100 \tag{2.5}$$

2.2.4 *The Abbe Criterion*

The Abbe criterion is another figure of merit that helps assess the quality of a peak fit. It is given by:

$$Abbe = \frac{1}{2}\frac{\sum_{i=1}^{N-1}[R(i+1) - R(i)]^2}{\sum_{i=1}^{N}[R(i)]^2} \tag{2.6}$$

Here, *R(i)* and *R(i + 1)* refer to the residuals of the fit at the i^{th} and $(i + 1)^{th}$ data points, respectively. In essence, the Abbe criterion tells a user how the residuals of a fit are distributed. Let's consider two possibilities. First, imagine that the residuals of a fit consistently have the same value. This would occur if the calculated fit were always above or below the measured data

points by a fixed amount. Clearly this situation is undesirable. Now, if $R(i) = R(i + 1)$ for the $N - 1$ data points considered in Equation 2.6, $Abbe = 0$. Next, let's consider the possibility that $R(i) = -R(i + 1)$ for the $N - 1$ data points in Equation 2.6. In other words, each subsequent residual has the same magnitude, but a sign opposite to the previous residual. A little math shows that Equation 2.6 approaches a value of two here. This situation is also undesirable. In summary, we would not want the fit to our data to be consistently above or below it ($Abbe = 0$), and we should be suspicious if we find that our residuals are anticorrelated, i.e., of the same magnitude, but alternating in sign ($Abbe \rightarrow 2$). It turns out that for random noise and statistically distributed residuals, $Abbe \approx 1$. This value of unity for the Abbe criterion is the desired value for this parameter, and significant deviations from it suggest issues with a data fit.

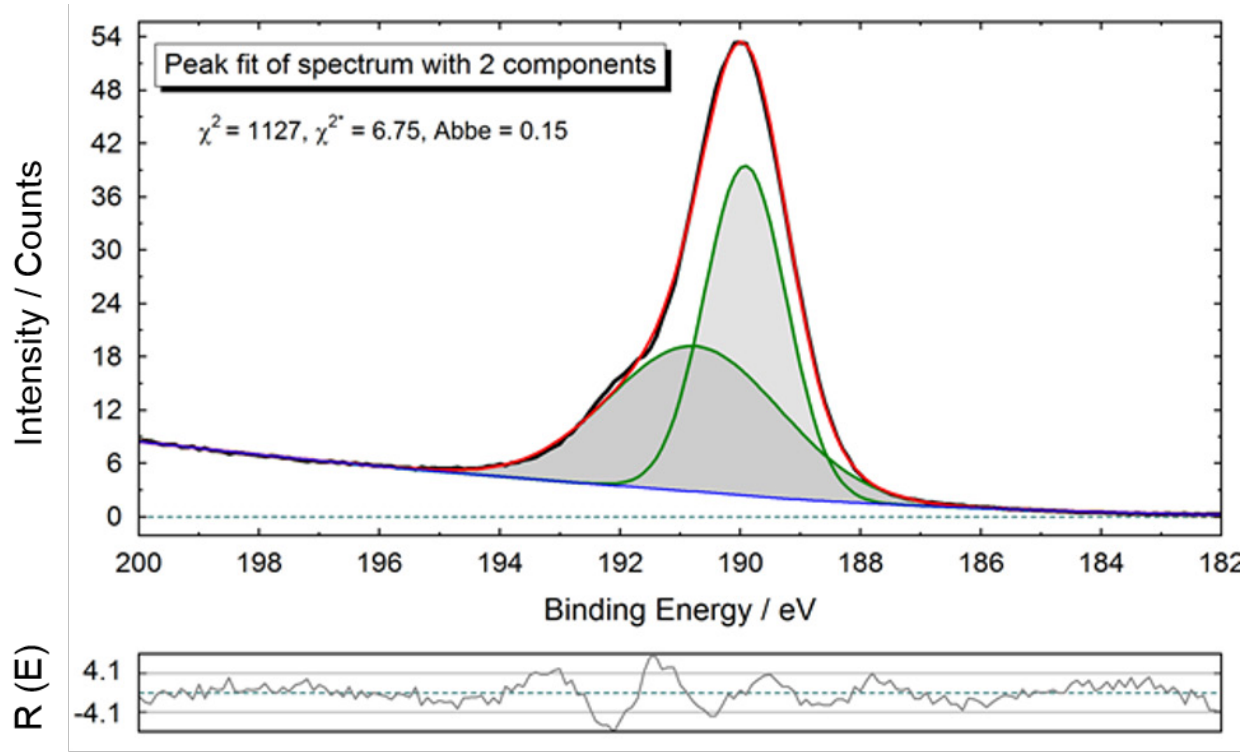


Figure 2.1 Two-component peak fit of a simulated spectrum on a polynomial background. Black: spectrum, red: sum of curves, green: individual components, blue: background, grey: residuals (below the fitted spectrum). The fit components are the product of Gaussian and Lorentzian peaks. The background function was calculated in parallel to the peak fit.

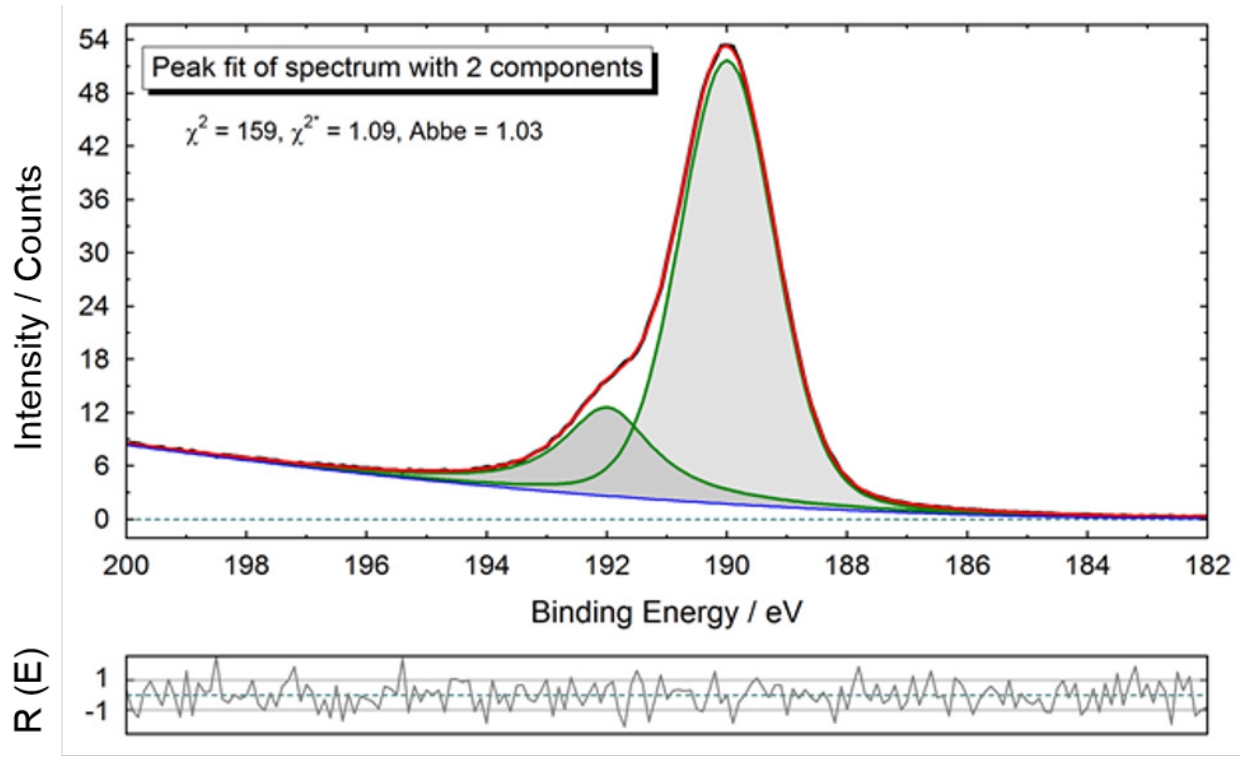


Figure 2.2 Two-component peak fit of a simulated spectrum on a polynomial background. Black: spectrum, red: sum of curves, green: individual components, blue: background, grey: residuals (below the fitted spectrum). The fit components are convolutions of Gaussian and Lorentzian peaks (Voigt functions). The background function was calculated in parallel to the peak fit.

2.3 Good practices for peak fitting

2.3.1 Use chi squared

Sherwood et al. mentioned that reporting chi squared values for a single spectrum is often not very informative.[7] However, reporting chi squared values for different fits applied to the same spectrum, or for fits to different, but related, spectra, *can* provide useful information. Figures 2.1 and 2.2 show two different simulated fits to a single simulated spectrum. Figure 2.1 shows a rather poor fit to the data – there are clear deviations between the original 'spectrum' (black) and the sum of the fitted peaks (red). Figure 2.2 shows much better peak fitting. The goodness of these fits (or lack thereof) in Figures 2.1 and 2.2 is clearly reflected in their χ^{2*} values, which are 6.75 (higher) and 1.09 (much lower), respectively. In summary, chi squared values can be useful for assessing the goodness of a fit.

2.3.2 Use the Abbe criterion

The Abbe criterion is also useful for determining the quality of a fit. It is particularly helpful for identifying fits that are consistently above or below their corresponding experimental data. Figure 2.1 shows a fit that entirely misses the experimental data (either above or below it) in various places. Accordingly, its Abbe criterion is quite low (0.15). In contrast, Figure 2.2 shows a fit that follows its experimental data quite well. It has an Abbe criterion close to unity (1.03), which, again, is the most desirable value of this parameter for a good peak fit.

2.3.3 Show the sum of your fit components with your original data

We will begin with a quote from Sherwood, who observed: "*Curve fitting ... has the advantage that an observer can clearly see how the summation of the fitted data compares to the original experimental data. It is very unfortunate that this advantage is often negated by authors who do not show the summation of the fitted data, but only the component peaks, thus depriving the reader of their work of any means of assessing the quality of their fit.*"[7] We echo his sentiment. From time to time we also see fits to XPS narrow scans in the Literature that only show the individual components (peaks) of the fit without their summation. Again, this leaves the reader without adequate information to judge the quality of the fit. Figures 2.1 and 2.2 follow the preferred course. Each figure shows two fitted peaks, and also the sum of these peaks, which can be directly and easily compared to the 'experimental' data. Also shown in the fits are the baselines that were used, which is clearly advisable as well. The fit shown in Figure 2.1 is inadequate, which is reflected in the sum of the peaks in Figure 2.1. In contrast, the sum of the peaks in Figure 2.2 shows that they fit the experimental data much better.

2.3.4 Show the (normalized) residuals

Figures 2.1 and 2.2 contain plots of the residuals to the fits. In Figure 2.1, the residuals, and also the discrepancy between the experimental data and the fit envelope, clearly show why the Abbe criterion is close to zero. In Figure 2.2, the residuals reveal why the Abbe criterion is close to unity. Note that (i) in many cases it probably doesn't matter too much whether one shows R, R_N, or $\%R$ (what matters, is that the residuals be presented in some form), and (ii) when the fit is quite good, it may be difficult to see the discrepancy between the experimental data and the sum of the fit components – the residuals can be helpful here.

2.3.5 Choose an appropriate background

This has been discussed in Chapter 1, under *section 1.4.1*.

2.3.6 Estimate the fit-parameter uncertainties

Hesse and coworkers have emphasized the calculation of uncertainties in fit parameters as an error estimation tool to understand the robustness of fit model parameters.[14, 15] The calculation of uncertainty (Δp_i) of a parameter (p_i) stems from the approximation that near the minimum of the plot of $\chi^2(\vec{p})$ vs. p_i, the curve has a parabolic shape. To generate this plot, one, in essence, selects different values of p_i, where for each of these values the other fit parameters are optimized such that an optimal value for $\chi^2(\vec{p})$ is obtained. Under these circumstances, we accept the following definition of Δp_i:[18]

$$\chi^2\ (p_i + \ \Delta p_i) = \ \chi^2(p_i) + 1 \tag{2.7}$$

This equation, again a definition, states that a change of p_i by Δp_i, leads to a change in $\chi^2(p_i)$ of 1. Taylor series expansions are extensively used in numerical analysis.[19] Accordingly, the left side of Equation 2.7 is expanded as a truncated (three-term) Taylor series as follows:

$$\chi^2\ (p_i + \ \Delta p_i) = \ \chi^2(p_i) + \ \frac{\partial \chi^2}{\partial p_i}|_{p_0}\Delta p_i + \ \frac{1}{2}\frac{\partial^2 \chi^2}{\partial {p_i}^2}|_{p_0}(\Delta p_i)^2 \tag{2.8}$$

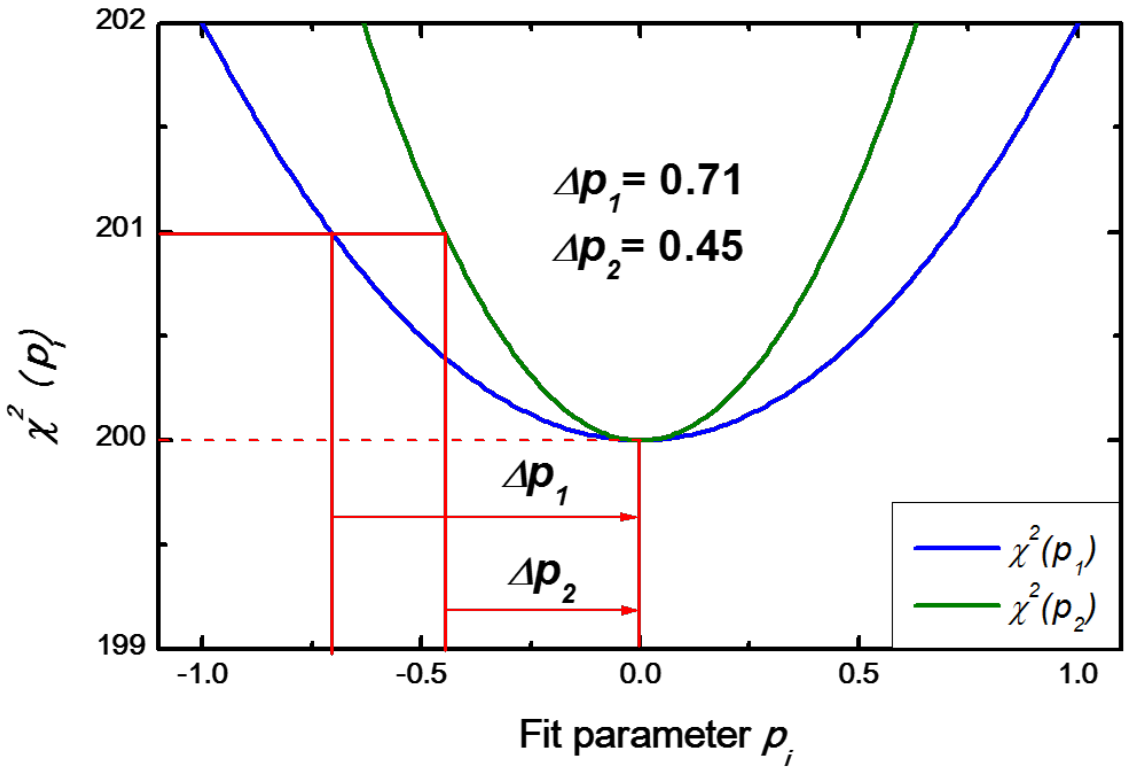


Figure 2.3 Plots of $\chi^2(\vec{p})$ *vs.* p_i showing different curvatures. The uncertainty in parameter p_2 is less than the uncertainty in parameter p_1 because its corresponding parabola is narrower (has a higher degree of curvature). This figure has been adapted from R. Hesse, T. Chasse, P. Struebel, R. Szargan. "Error estimation in peak-shape analysis of XPS core-level spectra using UNIFIT 2003: how significant are the results of peak fits?" Surf. Interface Anal. 2004; 36; 1373-1383.

Because the curve in question is assumed to be parabolic in form, its first derivative ($\frac{\partial \chi^2}{\partial p_i}$) is zero at its minimum. Thus, Equation 2.8 reduces to:

$$\chi^2(p_i + \Delta p_i) = \chi^2(p_i) + \frac{1}{2}\frac{\partial^2 \chi^2}{\partial {p_i}^2}|_{p_0}(\Delta p_i)^2 \tag{2.9}$$

Now, combining Equations 2.7 and 2.9 we obtain:

$$\chi^2(p_i) + 1 = \chi^2(p_i) + \frac{1}{2}\frac{\partial^2 \chi^2}{\partial {p_i}^2}|_{p_0}\sigma^2(p_i) \tag{2.10}$$

where we have now equated $(\Delta p_i)^2$ to $\sigma^2(p_i)$, which is the variance of the fit parameter p_i. Cancelling identical terms in Equation 2.10 and performing a little algebra then leads to:

$$\sigma(p_i) = \sqrt{\frac{2}{(\frac{\partial^2 \chi^2}{\partial {p_i}^2})}} \tag{2.11}$$

Thus, in Equation 2.11 we see that the curvature of $\chi^2(p_i)$, which is $\frac{\partial^2 \chi^2}{\partial {p_i}^2}$, is inversely related to the standard deviation of p_i. In other words, a small degree of curvature in $\chi^2(p_i)$ (a flat parabola) leads to a large uncertainty ($\sigma(p_i)$) in p_i. Conversely, a high degree of curvature in $\chi^2(p_i)$ (a narrow parabola) gives lower uncertainty to p_i. Another way of looking at this is to say that if we change the parameter p_i, and $\chi^2(p_i)$ rises sharply, there is little uncertainty in the value of p_i. Figure 2.3 illustrates these concepts. Here we see two plots of χ^2 vs. p_i that share the same minimum value. Both are parabolic in shape near their minima, and one (the curve for p_2) is narrower than the other. Also marked on this plot is the point at which $\chi^2(p_i)$ has increased in value by 1. Clearly, this change in $\chi^2(p_i)$ takes place with a smaller uncertainty (Δp_2) for the

narrower curve. In other words, the uncertainty associated with p_2 is less than that associated with p_1 because the parabola associated with it has a higher curvature.

In this analysis of uncertainties, it is also important to understand if there is any correlation between fit parameters. For this determination, a *Hessian* matrix (**H**) is constructed. The Hessian matrix contains the partial derivatives of $\chi^2(p_i)$ as follows:

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2\chi^2}{\partial {p_1}^2} & \frac{\partial^2\chi^2}{\partial p_1 \partial p_2} & \cdots \frac{\partial^2\chi^2}{\partial p_1 \partial p_N} \\ \frac{\partial^2\chi^2}{\partial p_2 \partial p_1} & \frac{\partial^2\chi^2}{\partial {p_2}^2} & \cdots \frac{\partial^2\chi^2}{\partial p_2 \partial p_N} \\ \vdots & \vdots & \cdots \vdots \\ \frac{\partial^2\chi^2}{\partial p_N \partial p_1} & \frac{\partial^2\chi^2}{\partial p_N \partial p_2} & \cdots \frac{\partial^2\chi^2}{\partial {p_N}^2} \end{bmatrix} \tag{2.12}$$

We note that the diagonal elements of this matrix are of the form $\frac{\partial^2\chi^2}{\partial {p_i}^2}$, while the off-diagonal elements are of the form $\frac{\partial^2\chi^2}{\partial p_i \partial p_j}$, with i ≠ j. Here, the off-diagonal elements of this matrix indicate the degree of correlation between the fit parameters p_i and p_j. One way to see this is to write $\frac{\partial^2\chi^2}{\partial p_i \partial p_j}$ as $\frac{\partial}{\partial p_i}(\frac{\partial\chi^2}{\partial p_j})$. That is, we see that if the derivative $\frac{\partial\chi^2}{\partial p_j}$ is independent of p_i, i.e., the two parameters are uncorrelated, then $\frac{\partial^2\chi^2}{\partial p_i \partial p_j}$ vanishes.

Now, assuming the off-diagonal elements of the Hessian matrix (Equation 2.12) are zero, we can calculate the uncertainties ($\sigma(p_i)$) by inverting **H** as follows:

$$\boldsymbol{B} = \boldsymbol{H}^{-1} \tag{2.13}$$

where **B** is referred to as the error matrix. The inverse of a diagonal matrix, i.e., a matrix that has non-zero elements only along its diagonal, can be calculated by inspection – in the inverse, each element along the diagonal is the reciprocal of the corresponding element in the original matrix. Thus, from Equations 2.11 and 2.13 we obtain:

$$\sigma(p_i) = \sqrt{2b_{ii}} \tag{2.14}$$

The reader who is interested in detailed calculations and additional analysis of these uncertainties is referred to Hesse's paper on this topic.[15]

2.3.7 Constrain one's fit parameters reasonably, and also limit the number of fit parameters

Inexperienced XPS users sometimes make a common mistake, which is to allow too many parameters to float (vary) in their modeling. This reminds us of a short dialogue in the Pixar movie "The Incredibles". Mrs. Parr (Elastigirl) is trying to explain to her son Dash why he can't go out for track (run on the track team at his school). In the course of the conversation, she says to him: "Everyone's special, Dash." He replies by saying: "Which is another way of saying no one is." By analogy, if you try to make every parameter in your fit special, i.e., you let every parameter float/vary in the fit, you may find out that none of your parameters is special, i.e., has any real meaning. Figures 2.4a and 2.4b illustrate these principles nicely. In Figure 2.4a, we show a reasonable fit of a C 1s narrow scan that was taken from some lightly oxidized carbon nanotubes. These oxidized carbon nanotubes have been used as scaffolds in the preparation of microfabricated thin layer chromatography plates.[20-22] The peaks in Figure 2.4a corresponding to reduced carbon, C(0), and carbon in its various oxidation states, C(I), C(II), and C(III),[13] were constrained to have the same width, and the width of the peak corresponding to the higher

binding energy shake-up peak was constrained relative to the other peak widths. The spacings between the peaks were also constrained relative to the C(0) peak, where the distances between these peaks were based on literature values. The Gaussian-Lorentzian ratios were the same for all the peaks.[9, 23] This same set of parameters/approach could be applied to a series of similar C 1s spectra to study the differences between them. The fit in Figure 2.4a shows a χ^{2*} value (as reported by our instrument software) of 2.11 and a value for the Abbe criterion (we calculated this separately) of 0.15.

As evidenced by the reduced chi squared and Abbe criterion values, there is probably some room for improvement in the fit to the data in Figure 2.4a, i.e., the peak positions, widths, and/or Gaussian:Lorentzian ratios might be tweaked somewhat to improve things. Of course, for all of these parameters in this optimization we would want to stay within chemically reasonable limits. Figure 2.4b shows the wrong way to approach this problem. Here, no constraints were placed on the peaks – all of the parameters floated. While this extra flexibility allowed the software to create a peak envelope that more closely matched the experimental data (the χ^{2*} value decreased to 0.85 and the Abbe criterion value increased to 0.76), this fit is entirely inappropriate and unphysical. For example, one would never expect the different peaks for reduced and oxidized carbon to have such radically different widths. In addition, the spacings between these peaks are inconsistent with literature values.

In summary, it is a bad idea to allow a large number of fit parameters to float in a peak fitting exercise. Obviously, another important way to show restraint in a peak fit is to use the fewest possible number of parameters that can be reasonably justified.

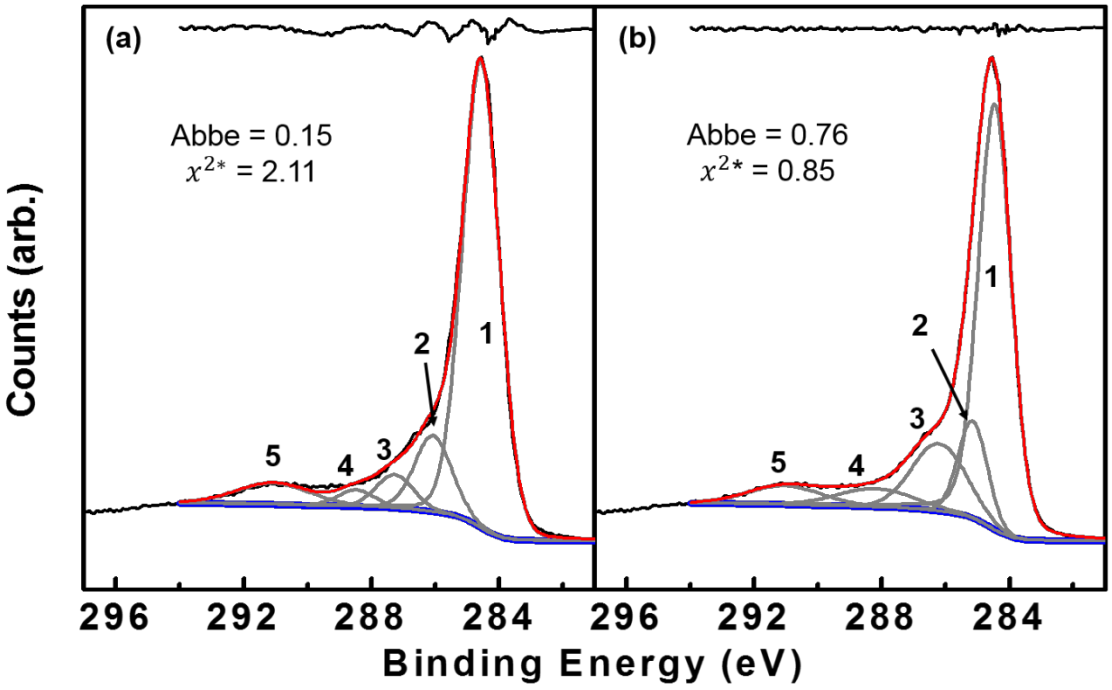


Figure 2.4 C 1s narrow scans from an oxidized CNT forests peak fitted to 5 components: (a) using reasonable constraints on their peak heights, widths, positions, and their Gaussian: Lorentzian ratios, and (b) letting all the peak parameters float. Black: experimental data, red: sum of fit components. The residuals to the fits are shown at the top of the panels.

2.3.8 Consider information from other techniques, and also other narrow scans within ones' analysis

Other information about a material can help guide XPS peak fitting. For example, one should know, at least approximately, how many peaks to place in a fit to a C 1s narrow scan, as well as their positions, if one knows the chemical structure of the polymer one is analyzing. Of course, in the case of block copolymers, one may find preferential migration of one block to the surface, which clearly creates an exception to the previous statement. But, in general, good prior knowledge about the structure of a material can be very helpful in guiding a peak fit. Note that this prior knowledge may include an understanding of how the sample was prepared/processed.

Useful information can also come from other narrow scans in the same analysis. For example, the component peaks in fits to C 1s spectra from materials that are rich in sp^2 carbon are expected to show asymmetry if the carbonaceous material is conductive. Without understanding this, it might be tempting to assume that this tail to higher binding energies is due to oxygen in one's sample. That is, if one sees asymmetry/tailing on a C 1s signal, but there is no corresponding oxygen (or other electronegative element) in the sample, it would be inappropriate to introduce peaks corresponding to oxidized carbon.

We have repeatedly emphasized the importance of the multi-instrument analysis of materials.[24, 25] For example, we regularly analyze materials by XPS, time-of-flight secondary ion mass spectrometry (ToF-SIMS), spectroscopic ellipsometry (SE), wetting, scanning electron microscopy (SEM), and atomic force microscopy (AFM), combining the results from these techniques to create a coherent and complete picture of a surface or material. An analysis becomes particularly compelling when results from the different techniques reinforce and agree

with each other. Thus, information from a different analytical technique can help guide a peak fit, and vice versa.

2.3.9 Use common sense

In the peak fits in Figures 2.1 and 2.2, we used a lower chi squared value and an Abbe criterion value closer to unity as evidence for an improved peak fit. However, in Figures 2.4a and 2.4b we did exactly the opposite – the chi squared value dropped and the Abbe criterion value moved closer to one when we removed the constraints on our peak fit parameters, creating a physically unreasonable peak fit. Is there a contradiction here? Shouldn't a decrease in reduced chi squared and an increase in the Abbe criterion towards unity always indicate the same thing (either a better or a worse peak fit)? The answer is no. Quite simply, we have to use our common sense here. We need to realize that within the context of each of these two analyses (that of Figures 2.1 and 2.2, and that of Figures 2.4a and 2.4b) the change in chi squared and Abbe criterion values mean different things. Peak fitting is not the type of exercise that should be undertaken without some careful thought, and probably consultation with the literature.

2.4 Conclusions

We have discussed a series of reasonable principles that, if followed, should lead to improved peak fits. While it may not be appropriate to apply every tool to every fit, many peak fits could benefit from a little more rigor.

We end with a brief anecdote that reemphasizes the importance of thinking carefully about one's work. My PhD advisor Dr. Matthew R. Linford is an editor for *Applied Surface Science*. In a recent meeting he attended with Henrik Rudolph, the Editor-in-Chief of the journal, he heard Henrik say (in reference to manuscripts submitted to the journal): "If something seems

fishy, it probably is." Similarly, we should trust our intuition, looking broadly and holistically at any peak fit we perform, so that if we see something that appears inappropriate or out of place we realize that it probably is, and then act to fix it.

Author Contributions: A good fraction of this chapter is based on R.H.'s previous publications.[14- 17] B.S. and M.R.L. wrote most of this document, also contributing their own thoughts and ideas. R.H. provided Figures 2.1 - 2.3.

2.5 References

1. A. A. Khan, D. Mathine, J. A. Woollam and Y. Chung, Phys Rev B **28** (12), 7229-7235 (1983).
2. J. A. Woollam, W. A. McGaham and B. Johs, Thin Solid Films **241** (1–2), 44-46 (1994).
3. J. A. Woollam and P. G. Snyder, Mater. Sci. Eng., B **5** (2), 279-283 (1990).
4. R. A. Synowicki, B. D. Johs and A. C. Martin, Thin Solid Films **519** (9), 2907-2913 (2011).
5. J. N. Hilfiker, C. L. Bungay, R. A. Synowicki, T. E. Tiwald, C. M. Herzinger, B. Johs, G. K. Pribil and J. A. Woollam, J. Vac. Sci. Technol., A **21** (4), 1103-1108 (2003).
6. G. Leibiger, V. Gottschalch, V. Riede, M. Schubert, J. N. Hilfiker and T. E. Tiwald, Phys Rev B **67** (19), 195205 (2003).
7. P. M. A. Sherwood, J. Vac. Sci. Technol., A **14** (3), 1424-1432 (1996).
8. M. R. Linford, in Vacuum Technology & Coating (August, 2014), pp. 27-35.
9. M. R. Linford, in Vacuum Technology & Coating (July, 2014), pp. 27-33.
10. M. R. Linford, in Vacuum Technology & Coating (March, 2014), pp. 31-35.
11. M. R. Linford, in Vacuum Technology & Coating (February, 2014), pp. 27-33.

12. H. Wang and M. R. Linford, in Vacuum Technology & Coating (March, 2015), pp. 27-32.

13. V. Gupta, H. Ganegoda, M. H. Engelhard, J. Terry and M. R. Linford, J Chem Educ **91** (2), 232-238 (2013).

14. Unifit Scientific Software GmbH, http://www.unifit-software.de.

15. R. Hesse, T. Chasse, P. Streubel and R. Szargan, Surf Interface Anal **36** (10), 1373-1383 (2004).

16. R. Hesse, T. Chasse and R. Szargan, Fresen J Anal Chem **365** (1-3), 48-54 (1999).

17. R. Hesse, P. Streubel and R. Szargan, Surf Interface Anal **39** (5), 381-391 (2007).

18. P. B. Bevington, *Data Reduction and Error Analysis for the Physical Sciences*, Second ed. (McGraw-Hill, New York, 1992).

19. R. L. Burden and J. D. Faires, *Numerical Analysis*, 9th ed. (Cengage, 2011).

20. D. S. Jensen, S. S. Kanyal, V. Gupta, M. A. Vail, A. E. Dadson, M. Engelhard, R. Vanfleet, R. C. Davis and M. R. Linford, J Chromatogr A **1257**, 195-203 (2012).

21. S. S. Kanyal, D. S. Jensen, A. E. Dadson, R. R. Vanfleet, R. C. Davis and M. R. Linford, JPC-J Planar Chromat **27** (3), 151-156 (2014).

22. S. S. Kanyal, T. T. Habe, C. V. Cushman, M. Dhunna, T. Roychowdhury, P. B. Farnsworth, G. E. Morlock and M. R. Linford, J Chromatogr A **1404**, 115-123 (2015).

23. M. R. Linford, in *Vacuum Technology & Coating* (June, 2014), pp. 25-31.

24. D. S. Jensen, S. S. Kanyal, N. Madaan, J. M. Hancock, A. E. Dadson, M. A. Vail, R. Vanfleet, V. Shutthanandan, Z. H. Zhu, M. H. Engelhard and M. R. Linford, Surf Interface Anal **45** (8), 1273-1282 (2013).

25. G. L. Jiang, F. Rivera, S. S. Kanyal, R. C. Davis, R. Vanfleet, B. M. Lunt, V. Shutthanandan and M. R. Linford, Opt Eng **50** (1) (2011).

Chapter 3: The Equivalent Width as a Figure of Merit for XPS Narrow Scans*

3.1 Abstract

X-ray photoelectron spectroscopy (XPS) is a widely used surface analytical tool that provides information about the near surface regions of materials. And while indispensable for XPS data analysis, peak fitting of narrow scans is often a fairly subjective exercise. Herein we introduce the equivalent width (EW) as an additional and less subjective figure of merit for XPS narrow scans. We believe that this parameter will prove particularly useful for analyzing series of similar or nominally identical spectra, perhaps as a component of an expert software system for the machine interpretation of spectra. It also appears to be useful, shedding light on the chemical state of materials, when additional information about a sample is known. The EW_{XPS} is simply defined as the area of a narrow scan divided by the height of the maximum of its peak envelope. To limit any ambiguity in EW_{XPS} for a series of spectra, we may also list the peak position of the maximum of the envelope (PE_{max}). The potential usefulness and limitations of the EW_{XPS} and PE_{max} parameters are demonstrated by their application to the narrow scans of: (i) four sets of ozone-treated carbon nanotubes (EW_{XPS} ~ 2.11 – 2.16 eV for a Shirley background, and up to 2.88 eV for no background, PE_{max} ~ 284.4 – 284.5 eV), (ii) a series of silicon wafers with different oxide thicknesses (EW_{XPS} ~ 1.5 – 2.8 eV, PE_{max} ~ 99 – 103 eV), (iii) hydrogen-terminated silicon before and after derivatization with pentyl groups, and after annealing of the pentyl-modified material (EW_{XPS} ~ 0.7 – 1.0 eV, PE_{max} ~ 25.9 – 26.1 eV), and (iv) five nanodiamond samples, where three of the spectra showed charging (EW_{XPS} ~ 2.6 – 4.9 eV, PE_{max} ~ 272.7 – 293.9 eV). In this final example, EW_{XPS} was plotted against PE_{max} to identify

*This chapter has been reproduced with permission from (Bhupinder Singh, Daniel Velázquez, Jeff Terry, and Matthew R. Linford), *Journal of Electron Spectroscopy and Related Phenomena*, 2014. **197**(0): p. 56-63

the region corresponding to the materials that showed the least charging. EW_{XPS} and PE_{max} appear to correlate with the expected chemistries of all the systems studied. We calculate EW_{XPS} using a Shirley baseline and with no baseline at all. In setting the baseline limits for EW_{XPS}, we consider the derivative of C 1s narrow scans. We also show the application of EW_{XPS} to single, fitted components within a narrow scan.

3.2 Introduction

XPS is a quantitative, surface sensitive technique that is extremely important for understanding surface chemistries.[1] It is a core electron spectroscopy that functions by illuminating a sample with X-rays and then measuring the kinetic energies of the ejected photoelectrons. These kinetic energies are then converted into binding energies that are plotted as survey (lower resolution) or narrow (higher resolution) spectra. Peak fitting of XPS narrow scans often plays a central role in revealing chemical information about a surface or material. However, peak fitting almost always involves at least some degree of user bias/subjectivity. Fortunately, this is not a significant issue in a number of cases. Well-understood and/or relatively simple materials often yield narrow scans that can be well fit and interpreted, especially by experienced practitioners and when additional information, such as process knowledge, is available. However, narrow scans of more complex materials can be difficult to fit, and these problems become particularly severe when inexperienced users apply too many fit parameters to their data without having a solid rationale for their choices. But even competent practitioners appear to struggle with challenging peak fitting problems. As Sherwood emphasized in his paper on peak fitting XPS narrow scans: "there is never a unique solution to fitting the data".[2] As an additional example, Wepasnick and co-workers fitted the same C 1s narrow scan of oxidized carbon nanotubes using peak parameters from two previously published fits.[3-5] They showed that

the *overall* fits to the signals were good in both cases. However, in one fit the signal due to carboxyl groups was 5.9% and in the other 11.0%. Clearly this is a substantial discrepancy that significantly changes one's understanding of this material.

The signal processing literature contains different methods for characterizing functions.[6] One of the more straightforward of these is the equivalent width (EW) of a function (see Figure 3.1a), which is simply defined as the total area of a function divided by its central ordinate:

$$EW = \frac{\int_{-\infty}^{\infty} f(x)dx}{f(0)} \tag{3.1}$$

That is, one can represent a function, f(x), as a rectangle with the same area, $\int_{-\infty}^{\infty} f(x)dx$, and height, *f(0)* (the central ordinate of the function), as the original function, where the width of this rectangle is the equivalent width (EW) of the function. Of course, the EW would be undefined for a function that passes through the origin, and it would have limited meaning for a function that is substantially shifted from the origin such that *f(0)* is small compared to other values of *f(x)*. While we are not aware of the application of the EW to XPS, it has been applied to other spectroscopic techniques. As Bracewell notes: "… the equivalent width of a spectral line is defined as the width of a rectangular profile which has the same central intensity and the same area as the line."[6] In particular, the concept of the EW has been widely used by astronomers.[7-9] That is, many astronomical signals will show a continuous background with intermittent absorption lines that are described by their EWs. There are also other definitions of the widths of functions,[6] but Equation 3.1 appears to offer the simplest definition with the most straightforward interpretation.

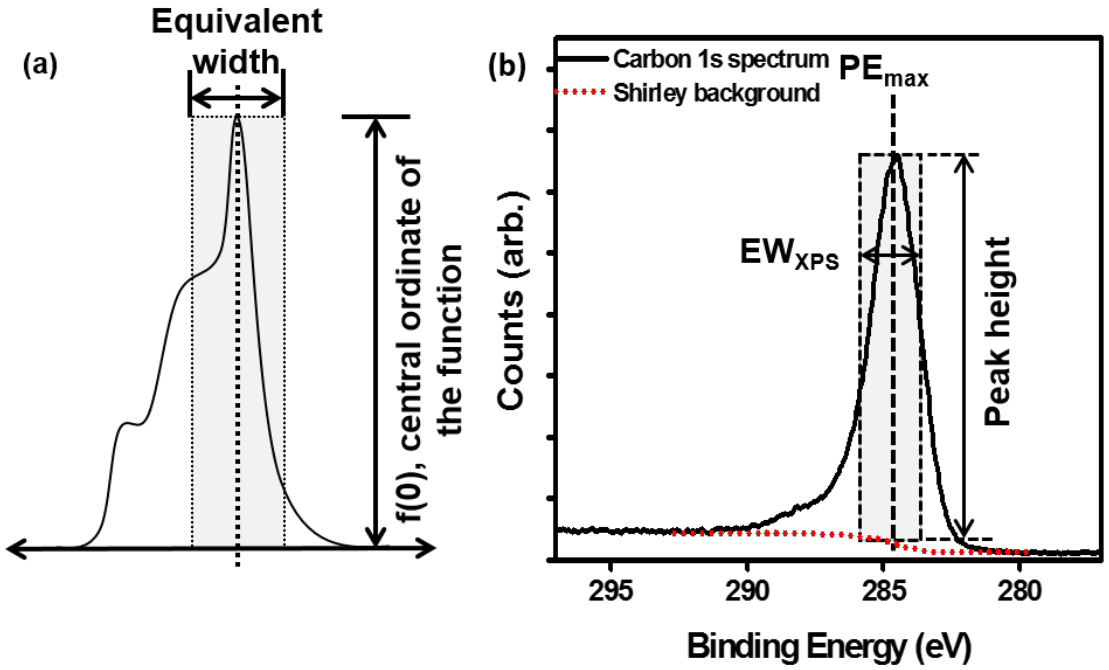


Figure 3.1 Graphical illustrations of (a) the EW of a general function, *f(x)*, and (b) the EW_{XPS} of a C 1s XPS narrow scan calculated with respect to a user defined baseline.

Herein we propose a slightly modified form of the EW as a figure of merit for XPS narrow scans (EW_{XPS}):

$$EW_{XPS} = \frac{Peak\ Area}{Peak\ Height} \tag{3.2}$$

where 'Peak Area' in Equation 3.2 is the area of an XPS peak above a user-defined baseline, which will typically be a Shirley,[10] Tougaard,[11] or linear background, and the 'Peak Height' is the height of the maximum point in the peak envelope, as measured from the background or other reference point. We also consider here the calculation of EW_{XPS} from narrow scans that have not been corrected for their backgrounds. The definition for EW_{XPS} in Equation 3.2 differs only modestly from the definition of EW in Equation 3.1 in that (i) EW uses the entire area of a function, both positive and negative regions, while EW_{XPS} only considers the area of a peak in question above a user-defined background or between user-defined limits, where this area will always be positive, and (ii) the denominator of the EW function is the central ordinate of the function, while that of EW_{XPS} is the height of the maximum of the peak envelope, as measured from its background.

In our exploration of EW_{XPS} as a figure of merit for XPS narrow scans, we found that there are some cases in which there may be some ambiguity in it. Accordingly, we recommend that the energy at the maximum of the peak envelope (PE_{max}) also be recorded (see Figure 3.1b). Both EW_{XPS} and PE_{max} should be sensitive to changes in the chemical states of surface species, and they should involve less user bias than traditional peak fitting. Additional insight into a series of materials may be obtained by plotting the EW_{XPS} and PE_{max} parameters against each other.

Here we apply EW_{XPS} and PE_{max} to four problems that illustrate the usefulness of this methodology. The first application is to four C 1s narrow scans of ozone-treated carbon nanotubes, which, like the spectrum discussed by Wepasnick and coworkers,[3] are challenging to peak fit unambiguously. We show a correlation between the EW_{XPS} of these narrow scans and their XPS oxygen-to-carbon ratios. As expected, the EW_{XPS} increases as the CNTs are oxidized, and the PE_{max} values stay nearly constant. We next follow the oxidation of silicon via Si 2p narrow scans, i.e., from silicon shards with different oxide thicknesses. In this example, starting with the EW_{XPS} of the bulk Si peak, and with increasing oxidation, we first observe an increase and then an asymptotic decrease in the EW_{XPS} of the Si narrow scans. Here the EW_{XPS} is also applied to the peak fitted bulk silicon and oxide signals in these spectra. While EW_{XPS} is not a single-valued function for this process, any ambiguity is removed by referencing the PE_{max} values for these spectra, which changes substantially from the signal of the silicon substrate to that of a heavily oxidized surface. In the third case, we consider Si 2p narrow scans collected with synchrotron radiation of hydrogen-terminated Si(111) before and after its modification with pentyl groups, and then after annealing of the pentyl-terminated surface. Same EW_{XPS} values were obtained for the Si 2p narrow scans of hydrogen-terminated Si (Si-H) and pentyl-terminated silicon. The pentyl-terminated silicon is then annealed, and new signals from the resulting 7x7 reconstruction appear that broaden the Si 2p signal. This change is reflected in the greater EW_{XPS} for this surface. However, as expected, the PE_{max} values for all three materials remain similar because these spectra are dominated by their bulk Si signals. Finally, we discuss five C 1s narrow scans from five different nanodiamond samples. Our interest in nanodiamond stems from the work some of us have done in making nanodiamond-containing core-shell particles for liquid chromatography.[12-14] Three of these five samples showed significant

charging. EW_{XPS} vs. PE_{max} were plotted for these spectra. In this plot, the 'good' C 1s narrow scans corresponded to the lowest EW_{XPS} values and the most reasonable PE_{max} values.

There have been many efforts to understand and fit XPS data. In this vein, we are trying to provide another means of rapidly comparing XPS narrow scans. For example, Tougaard has studied changes in baselines in XPS spectra and how they relate to surface structure.[15] However, the purpose of this article is *not* to encourage scientists to peak fit their XPS data less. Peak fitting is an indispensable part of XPS data analysis and will continue to be so for the foreseeable future. Rather, we are proposing another methodology that: (i) should be complementary to traditional peak fitting, (ii) will often involve much less user bias than traditional peak fitting, (iii) should be helpful in comparing and understanding different XPS narrow scans, (iv) is mathematically straightforward, and (v) may have unique applications, such as in quality control.

3.3 Experimental

3.3.1 Samples

Four sets of XPS narrow scans were evaluated.

1) *Carbon 1s narrow scans from four ozone-treated carbon nanotube (CNT) forests*. These spectra were previously taken as part of a study on infiltrated, CNT-templated thin layer chromatography plates.[16-18]
2) *Silicon surfaces with different oxide thicknesses*. Native oxide terminated silicon shards (ca. 1.5 x 1.5 cm^2) were plasma cleaned for 2 minutes in a Harrick plasma cleaner (PDC-32G) and air oxidized at 900 °C in a Thermolyne™ benchtop muffle furnace from Thermo Scientific for varying amounts of time. The different oxide thicknesses created in this manner were determined by spectroscopic ellipsometry (SE) (M-2000 instrument

from the J.A. Woollam Co., Lincoln, NE, ca. 200 – 1000 nm spectral range). Optical constants for silicon and silicon oxide from the instrument software were used to model the silicon substrate and the oxide. In this analysis, a normal fit was performed to get the thickness of the oxide layer. XPS analyses of all five samples were performed using a Surface Science SSX-100 X-ray photoelectron spectrometer (serviced by Service Physics, Bend, OR) with a monochromatic Al Kα source (1486.7 eV) and a hemispherical analyzer. Narrow scans were recorded with a spot size of 500 μm x 500 μm, resolution: 3, number of scans: 20 and step size of 0.065 eV. Peaks were referenced to the O 1s peak taken at 532.0 eV. This signal appeared as a single, shoulderless peak.

3) *Si 2p narrow scans from hydrogen-terminated Si(111) before and after modification with pentyl groups, and after annealing of the pentyl-modified material.* These spectra have been previously published.[19-21] These spectra were originally collected at the Stanford Synchrotron Radiation Laboratory on Beam Line 8-1 (BL 8-1) using a Toroidal Grating Monochromator (TGM). BL 8-1 was equipped with a low energy electron diffraction analyzer (LEED) and a double pass cylindrical mirror analyzer (CMA). The photon energy was 130 eV. As the original data were not readily available, Plot Digitizer (http://plotdigitizer.sourceforge.net) was used to generate it from a pdf copy of the data as it appeared in Jeff Terry's dissertation.[20]

4) *C 1s narrow scans of five different types of nanodiamond.* Sample preparation consisted of preparing a slurry of the nanodiamond in ultrapure water, placing a few droplets of this slurry on a clean piece of a Si wafer, and placing this surface over a hot plate to evaporate the water. This procedure resulted in a thin, uniform film of nanodiamond. The XPS instrument and settings of these analyses were the same as in Example 2.

3.3.2 Calculation of EW_{XPS}

The EW_{XPS} was calculated as follows (see Figure 3.2):

1) A background, e.g., linear, Shirley, or no background at all, was defined. If a series of narrow scans from related materials is to be compared, it will obviously be advantageous to choose the backgrounds and limits of integration in the same way. Some of the C 1s narrow scans described herein were differentiated numerically using Microsoft Excel to better determine the limits for integration/peak fitting. For all the data collected at BYU, Shirley backgrounds were determined using the instrument software. For the data previously obtained at the synchrotron, Shirley backgrounds were calculated in Microsoft Excel.
2) The height of the function was calculated as the height of the maximum of the peak envelope above the background.
3) The peak was numerically integrated, where each data point contributed an area equal to its height above the background multiplied by the step size in eV between data points. While this approach is adequate in this paper for demonstrating the concepts outlined herein, there are obviously more sophisticated methods for numerically integrating functions.[22] If the spectra are particularly noisy it may be advantageous to smooth them before performing this analysis.
4) The EW_{XPS} was calculated as the quotient of the peak area and the peak height.

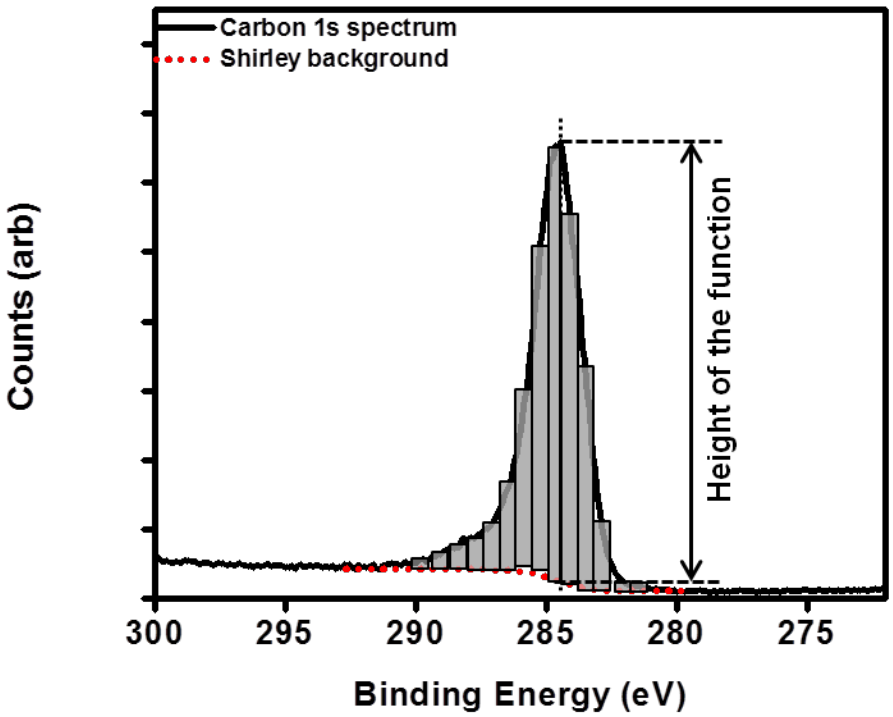


Figure 3.2 Illustration of the numerical integration of a peak for the calculation of the EW_{XPS} parameter of an XPS narrow scan. The actual spacing between data points is much smaller than suggested here.

3.4 Results and Discussion

3.4.1 Example 1. Ozone-treated CNT forests

Carbon nanotube (CNT) forests were treated with ozone (O_3) as part of a study to make microfabricated, CNT-templated thin layer chromatography plates.[16] Prior to O_3 treatment, the CNTs showed no oxygen by XPS.[23] After ozone treatment, an O 1s signal became apparent in the survey scans, and the C 1s narrow scans broadened in a way that suggested multiple chemical states for carbon. As was the case for the Wepasnick work noted above,[3] these scans are difficult to fit well. For example, there is a natural asymmetry to the main carbon peak, there is a shake-up signal, and the carbon appears to be in multiple oxidation states, probably as C-O, C=O, and carboxyl groups.[4, 5, 24-28] The rather low intensities of all of these components, except the main peak, add to the complexity of these spectra, i.e., accurately determining the concentrations of all of these low intensity components using only the narrow scans would be very challenging.

Figures 3.3a – c shows three of the four C 1s narrow scans of ozone-treated CNT forests considered in this work. The C 1s spectrum of the fourth sample was essentially the same as these spectra and is not shown. As noted above, the C 1s narrow scans of these materials are complex – any peak fitting is expected to be challenging. They are also quite similar to each other so differentiating between them is expected to be challenging. By XPS, these samples showed ca. 3.7 – 5.0 at. % oxygen. In a first attempt to apply the EW_{XPS} to these C 1s spectra, the beginning and ending points for Shirley baselines were chosen as the points on either side of the peak envelope where the spectra appear to flatten. Figure 3.3d shows a plot of the EW_{XPS} values of these C 1s narrow scans plotted against their O/C ratios, which were determined from the corresponding XPS narrow scans using Shirley backgrounds. In spite of the complexity of

these C 1s spectra and their obvious similarity, there is good correlation between the EW_{XPS} values and the amount of oxygen in the samples. As expected, the PE_{max} values for these spectra remain nearly constant, i.e., the relatively low amounts of oxygen in these materials suggests that the main carbon peak should remain in a roughly constant position.

In addition to the analysis with EW_{XPS} and PE_{max} that led to Figure 3.3, we also undertook a more sophisticated analysis of the data to decrease the degree of user bias, although we believe that any bias in the use of EW_{XPS} and PE_{max} is already quite low. This analysis also demonstrates the robustness of the EW_{XPS} parameter, i.e., we show that essentially the same trends are obtained when different approaches to data processing are utilized. We approached these analyses by reasoning that while the spectra would be expected to differ on their higher binding energy sides because of different amounts of chemically shifted carbon, the lower energy regions of the spectra, especially to the lower energy side of the peak maximum, would be expected to be very similar because they should be almost exclusively a result of unreacted carbon in the CNTs. In these approaches, the spectra were first numerically differentiated using both the definition of the derivative and a well-known five-point formula[29]:

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0)}{h} \tag{3.3}$$

$$f'(x_0) = \frac{1}{12\,h}[f(x_0 - 2h) - 8f(x_0 - h) + 8f(x_0 + h) - f(x_0 + 2h)] \tag{3.4}$$

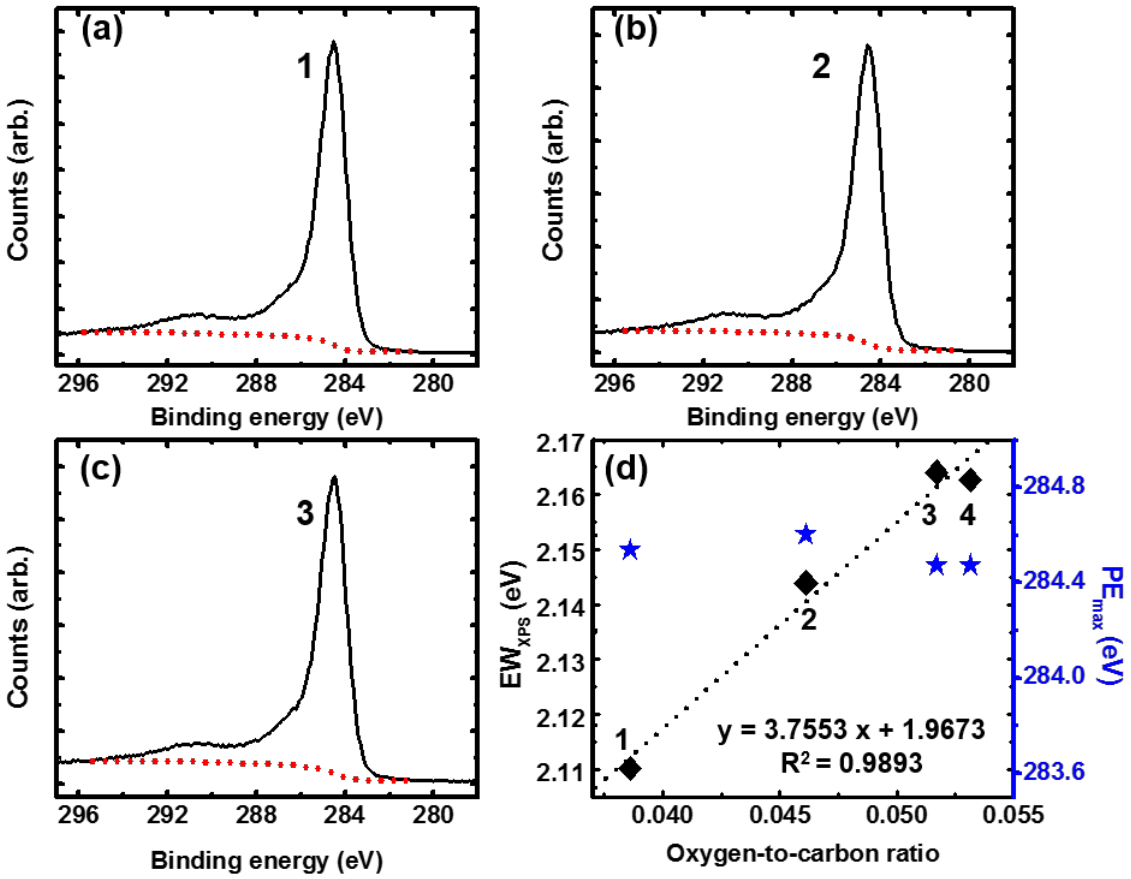


Figure 3.3 (a) – (c) C 1s XPS narrow scans of three carbon nanotube (CNT) samples: 1, 2, and 3. (d) Plot of EW_{XPS} (◆) and PE_{max} (★) as a function of the oxygen-to-carbon ratio for these and one other similar sample.

In these equations, x_0 is the point in question and h is the distance between equally spaced points in a spectrum, e.g., $x_0 + h$ and $x_0 - h$ are the points around the point x_0. In addition to differentiating a function, the five-point formula has the effect of smoothing it (see Figure 3.4). This is clearly advantageous because the two-point formula is more strongly affected by noise in the data. Obviously there are many ways to handle these spectra – the data might be smoothed before differentiation and/or other functions for their differentiation might be considered – the purpose of this particular analysis is not to suggest a best method for identifying reference points in XPS spectra. The minimum in these derivative spectra are expected to correspond to the inflection points on the lower energy side of the maxima in the C 1s spectra (see Figure 3.5a).

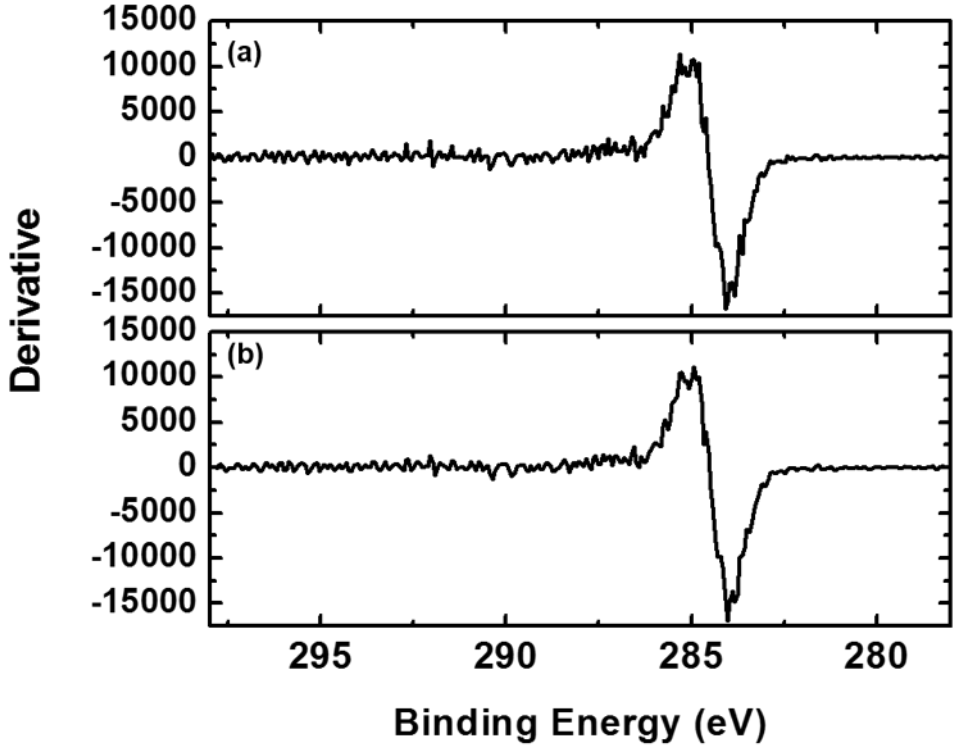


Figure 3.4 Derivative of the C 1s narrow scan of CNT Sample 1 using (a) the definition of derivative and (b) the five-point formula given in the text.

After differentiating the spectra, either the minima in the resulting derivative curves, e.g., Point A in Figure 3.5a, or the points half way between the minimum and the zero level, e.g., Point B on Figure 3.5a, were used as references. Because Points A and B were rather close together, ca. 0.6 eV apart, the beginning point of the background was taken -3.0 eV from the reference point, and the ending point +11.0 eV from the same point, regardless of which reference point was employed (A or B). These limits were chosen so that the baseline contructed from them would reasonably capture the area under the C 1s curve. Using these beginning and ending points, two types of backgrounds were considered: a Shirley background and no background at all, i.e., in this latter approach, all the counts between the end points were used. This 'no background' approach will allow changes in the background to be better accounted for in the EW_{XPS} function. Overall, this latter approach is expected to be the least biased, and the plot of the EW_{XPS} vs. the oxygen-to-carbon ratio in the 'no background' approach showed the highest degree of correlation (see Figure 3.5b). However, the correlations are still very good for the Shirley backgrounds, whether used in the standard background fitting methodology (Figure 3.3d) or derivative (Figure 3.5c) approaches. The data shown herein (Figure 3.5) are obtained with reference to the minimum in the derivative curve (Point A in Figure 3.5a). The results obtained from the halfway point, e.g., Point B in Figure 3.5a, are almost as good as those obtained when the minimum point (Point A) was the reference point. The raw data are shown in Table 3.1. Again, it speaks to the robustness of the EW_{XPS} parameter that very similar correlations to the O/C ratio are obtained for different baselines and different limits of integration.

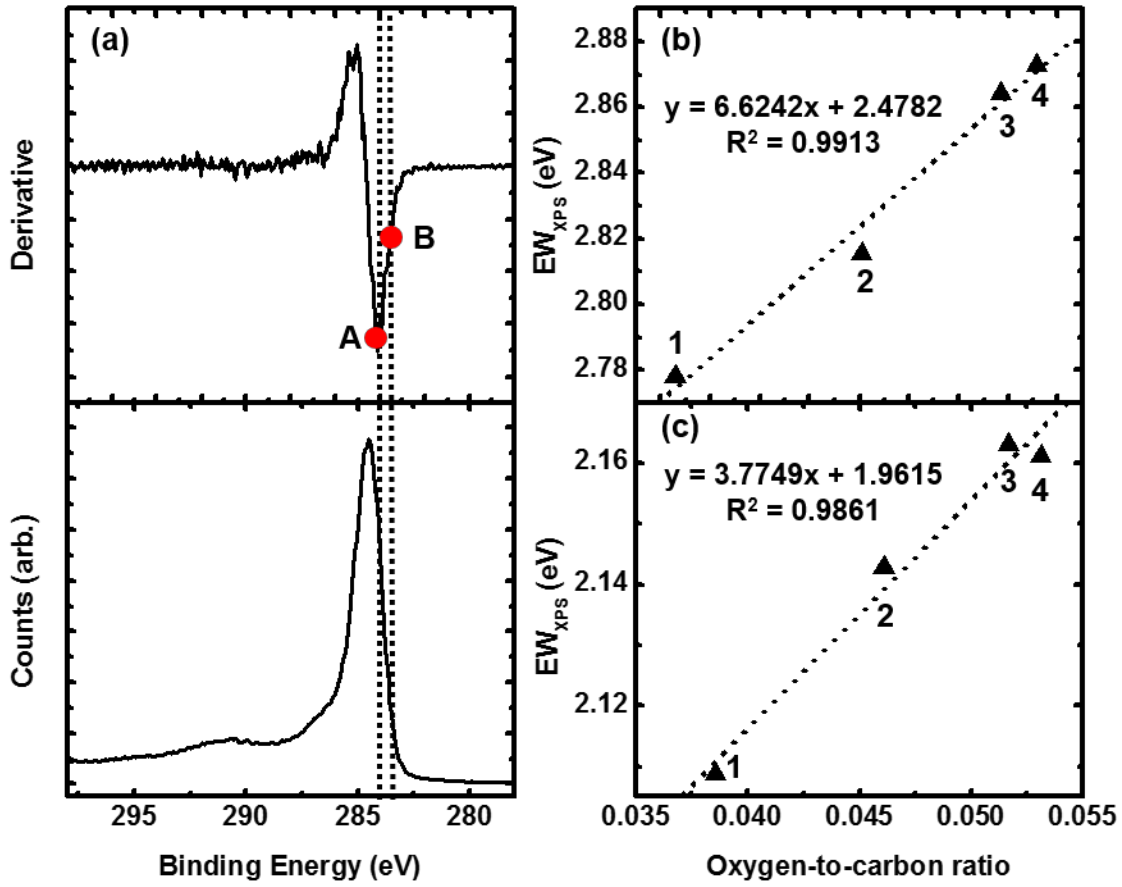


Figure 3.5 (a) Derivative obtained with a five-point formula (Equation 3.4) of the C 1s narrow scan below it, where this scan corresponds to CNT Sample 1 in Figure 2. (b – c) EW_{XPS} values plotted against their corresponding oxygen-to-carbon ratios from XPS analyses. The minimum in each derivative curve, e.g., Point 'A' in panel (a), was used as the reference point for integrating each peak envelope. In (b), the data were analyzed without a background, and in (c), a Shirley background was employed.

Table 3.1 EW_{XPS} (eV) values of CNT samples calculated under various conditions to test the robustness of the method.

Sample #	O/C	Ref. Pt. 'B' No background	Ref. Pt. 'A' No background	Standard fitting method	Ref. Pt. 'B' Shirley background	Ref. Pt. 'A' Shirley background
1	0.03858	2.738	2.778	2.110	2.105	2.109
2	0.04611	2.775	2.816	2.144	2.140	2.143
3	0.05170	2.824	2.864	2.164	2.159	2.163
4	0.05314	2.832	2.873	2.163	2.157	2.161

Note on the significant figures in this document. We certainly do not claim that the numbers reported herein have significance to the number of significant figures reported in most cases below. We simply report here the output from our calculations.

3.4.2 Example 2. Thermally grown oxide on silicon shards

Native oxide terminated silicon wafers were plasma cleaned and then air oxidized at 900 °C for 1, 2, 3, 4, and 5 minutes. The samples were labeled Si_{NO} (untreated, native oxide-terminated silicon wafer) and 1 – 5 corresponding to the number of minutes they were oxidized. Spectroscopic ellipsometry (SE) showed a reasonable correlation between the increase in oxide thickness and the oxidation time; the fits to a simple oxide-silicon model, as determined by the mean squared error (MSE), were reasonable (see Table 3.2).

Figures 3.6 (a-c) shows plots of the Si 2p narrow scans of samples Si_{NO}, 3, and 5, respectively. These spectra show a peak at ca. 99 eV that corresponds to bulk Si and another at ca. 103 eV attributed to oxidized silicon. The EW_{XPS} parameters for these spectra were calculated using Shirley backgrounds. As expected, the EW_{XPS} values are highest for the narrow scans that showed relatively intense Si *and* SiO_2 peaks, i.e., the presence of two nearly equal peaks gives a lower value for the height of the EW_{XPS} function, compared to the total area of the peak envelope. As the samples were increasingly oxidized, and a single SiO_2 peak became dominant, the EW_{XPS} parameter appeared to drop asymptotically (see Figure 3.6d). To explore the use of the EW_{XPS} parameter, we also peak fitted the six narrow scans to two unconstrained Gaussian: Lorentzian (G:L) peaks. For both components (the bulk Si peak and the SiO_2 peak), the best G:L ratio was nearly the same: 85:15. Accordingly, this ratio was used for the peak fitting reported herein. The EW_{XPS} was calculated for each of these components. For samples Si_{NO} and 1 – 4, the average EW_{XPS} for the bulk Si peak was 1.54 ± 0.06 eV, showing good sample-to-sample consistency (the fifth scan was not considered because of the significant attenuation of the Si signal through the thicker oxide layer above it). For samples Si_{NO} and 1 – 5, the average EW_{XPS} value of the SiO_2 peak was 2.13 ± 0.08 eV, also showing good sample-to-

sample consistency for this parameter (see Table 3.3). These values are plotted in Figure 3.6d for comparison to the other data points. It might be expected that EW_{XPS} values for oxides, such as SiO_2, would be greater than for the bulk elemental Si peak. Sherwood similarly found that the oxide peak was wider than the metal peak on superficially oxidized aluminum, perhaps because of greater vibrational broadening in the oxide.[2] Peak broadening can also be an indication of structural disorder in materials.[30-32]

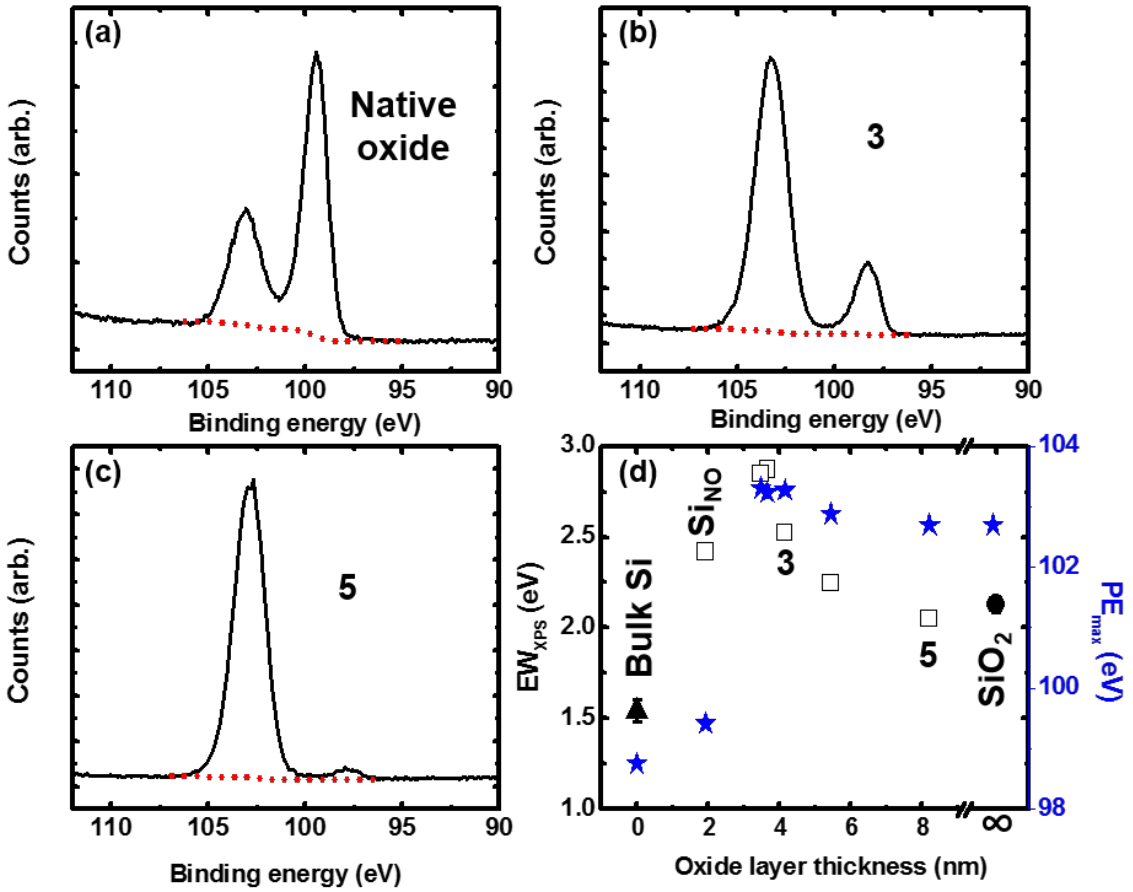


Figure 3.6 (a) Si 2p narrow scan of a native oxide terminated Si wafer. (b – c) Representative Si 2p narrow scans of silicon shards with different thicknesses of oxide. (d) EW_{XPS} values plotted as a function of oxide layer thickness (nm) of complete peak envelopes (□), and also the average EW_{XPS} values from peaks corresponding to the Si substrate (▲) and the SiO_2 overlayer (●). Error bars are the standard deviations of these measurements. Also shown are corresponding PE_{max} values (★). Si_{NO} refers to native oxide terminated silicon.

Table 3.2 Oxide growth on silicon shards as measured by spectroscopic ellipsometry.

Sample #	Duration of oxidation	Initial oxide thickness (nm)	MSE	Final oxide thickness (nm)	MSE	Growth (nm)
Native oxide	-	1.922 ± 0.0101	2.909	-	-	-
1	1 min	1.825 ± 0.0139	3.891	3.667 ± 0.0126	3.584	1.842
2	2 min	1.810 ± 0.0138	3.894	3.481 ± 0.0141	3.847	1.671
3	3 min	1.833 ± 0.0135	3.988	4.158 ± 0.0141	3.712	2.325
4	4 min	1.809 ± 0.0135	3.801	5.441 ± 0.0141	3.765	3.632
5	5 min	1.849 ± 0.0134	3.479	8.217 ± 0.0134	3.509	6.368

Table 3.3 EWXPS and PE_{max} values of entire Si 2p envelopes and the individual bulk Si and SiO_2 peaks from peak fitting the Si 2p envelopes of silicon samples with different oxide thicknesses.

Sample #	EW_{XPS} Bulk Si	PE_{max} Bulk Si	EW_{XPS} SiO_2	PE_{max} SiO_2	EW_{XPS} Si 2p envelope	PE_{max} Si 2p envelope
Native oxide	1.594	99.49	2.266	103.05	2.421	99.42
1	1.541	98.82	2.080	103.29	2.875	103.23
2	1.587	98.44	2.141	103.30	2.849	103.30
3	1.548	98.34	2.141	103.26	2.524	103.27
4	1.442	97.63	2.107	102.82	2.248	102.88
5	1.167	98.56	2.039	103.62	2.049	102.69

Unlike the previous example, where we observed a steady increase in EW_{XPS} with an increase in the oxygen-to-carbon ratio, the EW_{XPS} of the Si 2p envelope first increases as the silicon is initially oxidized, and then decreases asymptotically with increasing oxide layer thickness (see Figure 3.6d and Table 3.3).[33] According, EW_{XPS} is not expected to be a single-valued function for this process. However, the combination of the PE_{max} and EW_{XPS} parameters in the analysis of this system clearly improves this situation. For example, we observe the expected, large difference between the PE_{max} values for spectra dominated by the bulk Si signal (PE_{max} ~ 99 eV) and for spectra dominated by the oxide signal (PE_{max} ~ 103 eV).

3.4.3 Example 3. Hydrogen-terminated Si(111) and its derivatives

Spectra from three different modified Si(111) surfaces were considered: (i) hydrogen terminated silicon (Si(111)-H), (ii) pentyl-terminated silicon (Si(111)-C_5) that was derived from Si(111)-H, and (iii) Si(111)-C_5 that had been annealed at 700 ºC, (Si(111)-C_5/700º). The EW_{XPS} values for the Si(111)-H and Si(111)-C_5 surfaces were both 0.72 eV. This result is consistent with Terry and coworkers' previous report of these spectra.[19, 21] However, after the Si(111)-C_5 surface was annealed, the resulting Si(111)-C_5/700º surface showed a significantly increased EW_{XPS} value of 1.04 eV. This value is consistent with the shoulder on the peak envelope at higher KE attributable (see Figure 3.7c) to surface states from the Si(111)-(7 x 7) reconstruction.[34] The PE_{max} values of the three spectra differed by only ~ 0.1 eV, which was again consistent with the prior work (see Figure 3.7), i.e., the spectra are dominated by the bulk Si signal.[19]

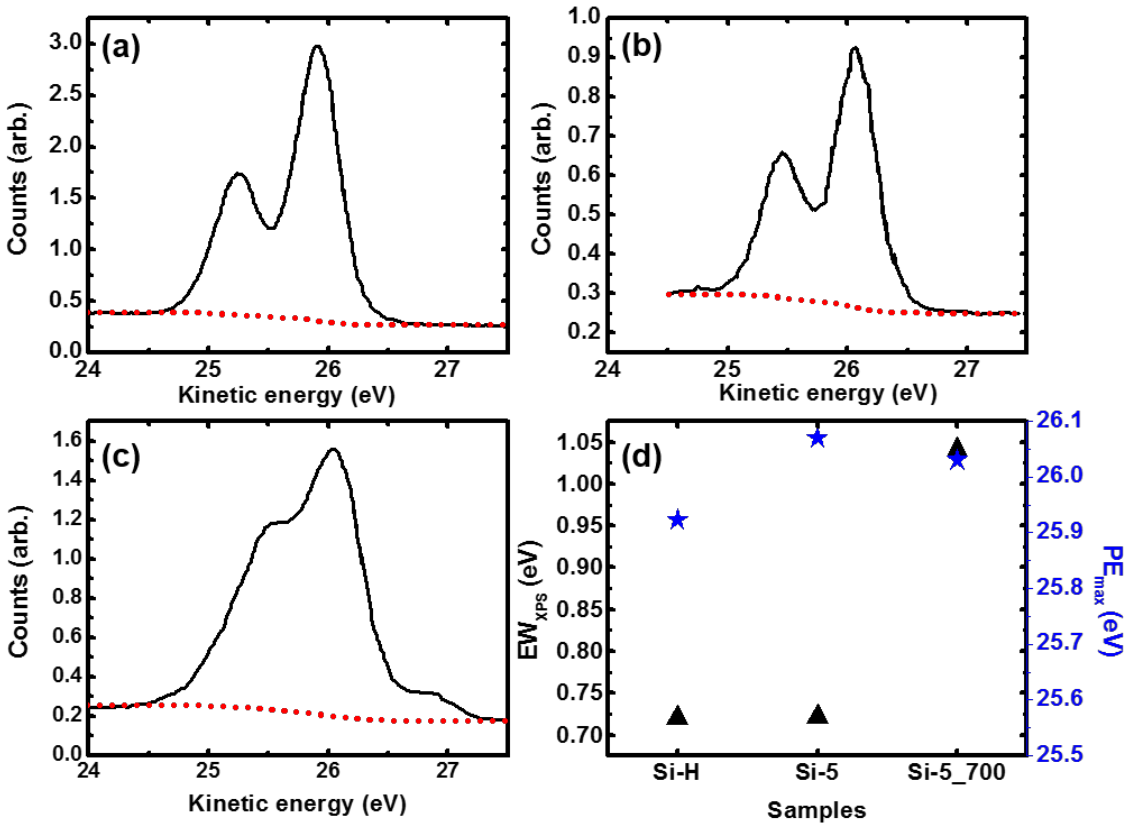


Figure 3.7 XPS narrow scans of (a) hydrogen terminated silicon (Si-H), (b) pentyl-terminated silicon (Si-5), and (c) Si-5 annealed at 700 ºC (Si-5_700). (d) graph of EW_{XPS} (▲) and PE_{max} (★).

3.4.4 Example 4. Sample charging

Five different nanodiamond samples were analyzed by XPS. Three of the resulting C 1s narrow scans are presented in Figure 3.8. Good spectra, by which we mean C 1s narrow scans that do not show significant broadening and/or artifacts, e.g., Figure 3.8a, were only obtainable from two of the samples. Figures 3.8b and 3.8c show C 1s narrow scans from samples that charged, where the same unusual/distorted shapes, i.e., the low energy shoulders on the peaks, were also found in the other peaks in the corresponding spectra. The EW_{XPS} and PE_{max} values for the five uncorrected C 1s narrow scans from these samples were determined, and these two parameters were plotted against each other (see Figure 3.8d). The dotted lines in this figure are guides to the eye for the following observations. The two points inside the dotted line correspond to good, narrow, C 1s spectra that did not show artifacts. They have the lowest EW_{XPS} values, and their PE_{max} values are not grossly shifted from the value expected for a hydrocarbon (~285.0 eV). The remaining three points in the plot correspond to spectra that showed artifacts. Their EW_{XPS} values were either substantially higher than those of the good samples and/or they showed extreme PE_{max} values because of either sample charging or failed attempts at charge compensation. Like the other data shown in this work, these results are not intended to represent an exhaustive study of this material system. Rather, they suggest that the EW_{XPS} and PE_{max} parameters may be useful in identifying charged or otherwise outlying spectra, where the 'space' created by plotting these two parameters against each other may show regions corresponding to 'good' and 'bad' spectra.

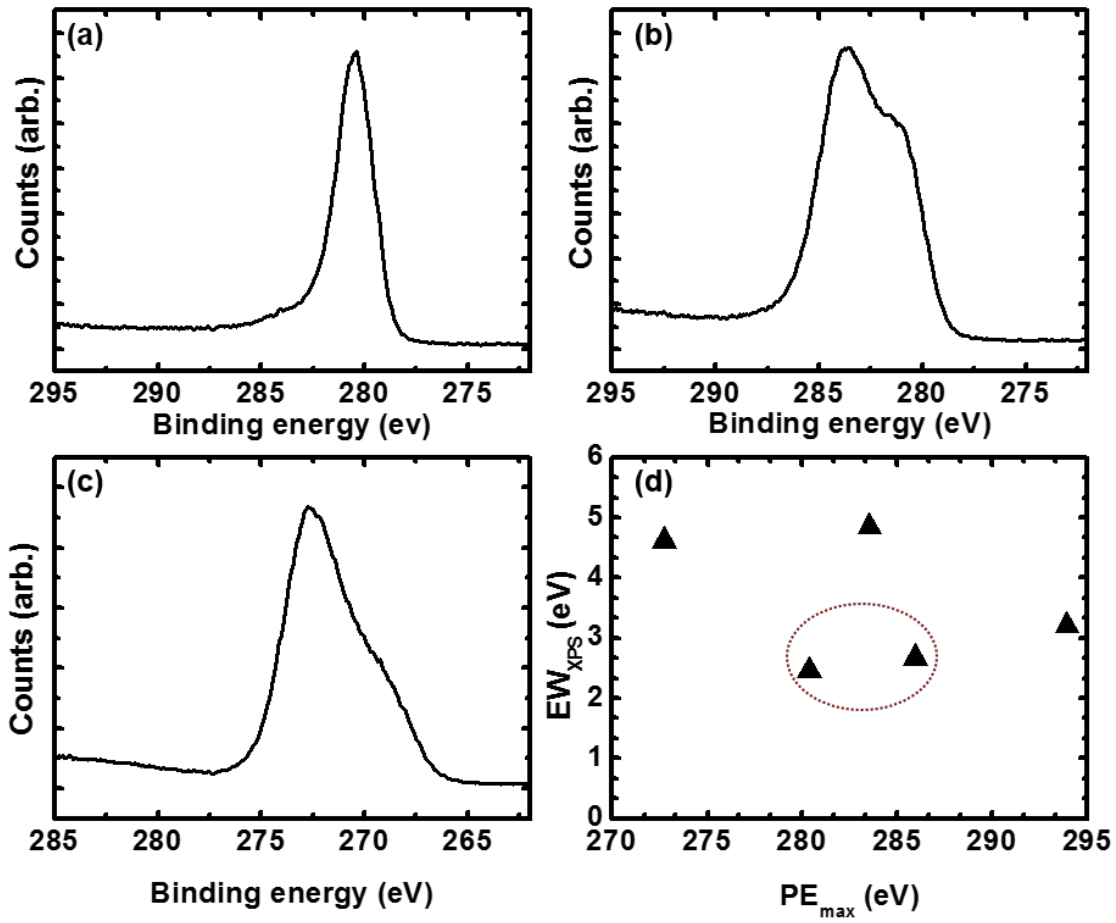


Figure 3.8 XPS C1s narrow scans of nanodiamonds that (a) showed a reasonable peak envelope, and (b – c) showed artifacts due to non-uniform charge compensation. (d) Graph of the EWXPS vs. PE_{max} values from the C 1s narrow scans of five nanodiamond samples. The red dotted line is a guide to the eye.

3.5 Conclusions

In the four examples presented herein, EW_{XPS} and PE_{max} behave in chemically/physically reasonable ways; EW_{XPS} and PE_{max} are responsive to surface chemistry and sample charging. They appear to play a complementary role to peak fitting in XPS narrow scan analysis, i.e., traditional peak fitting is richer in chemical information than EW_{XPS} and PE_{max}, but it also appears to be more susceptible to user bias. Consider, for example, the dozen or so parameters (peak positions, peak widths, peak heights, Gaussian-Lorentzian ratios, the Gaussian-Lorentzian fit function, the asymmetry of the C-C peak, etc.), that would need to be selected and fit, beyond those needed to define a baseline, for the complex C 1s spectra of the oxidized CNTs discussed herein. Unlike peak fitting, where one attempts to break up a signal into its components, EW_{XPS} and PE_{max} characterize entire peak envelopes. The resulting values for a series of spectra can be plotted against each other. In addition to their usefulness in comparing entire narrow scans, they can also be applied to individual, peak-fitted components within a spectrum. The EW_{XPS} can be determined without a user-defined baseline, where one only needs to select two endpoints, which further reduces any user bias. We suggest that EW_{XPS} and PE_{max} should be considered as additional tools for understanding, comparing, and validating XPS narrow scans. For example, these parameters could be used in place of the qualitative analysis commonly performed on narrow photoemission scans. They may also prove valuable as components of an expert software system for the machine interpretation of spectra.

3.6 Acknowledgement

We thank the Department of Chemistry and Biochemistry and the College of Physical and Mathematical Sciences at Brigham Young University for their support of this work. DV was

partially supported by the U.S. Department of Education through the GAANN Fellowship Program.

Author Contributions: B.S. grew the different oxide thicknesses for the Si/SiO_2 samples used herein, and did XPS measurements on them, along with nanodiamond samples. B.S. did all the data analysis, made figures, helped in writing the document, and had significant discussions with M.R.L. on this topic, making important intellectual contributions to the project. D.V. crosschecked all the calculations. J.T. gave his feedback and technical inputs to the project. M.R.L. suggested that EW might be useful for characterizing XPS narrow scans, and helped write the original manuscript.

3.7 References

1. V. Gupta, H. Ganegoda, M. H. Engelhard, J. Terry and M. R. Linford, J Chem Educ **91** (2), 232-238 (2013).

2. P. M. A. Sherwood, J. Vac. Sci. Technol., A **14** (3), 1424-1432 (1996).

3. K. A. Wepasnick, B. A. Smith, J. L. Bitter and D. H. Fairbrother, Anal Bioanal Chem **396** (3), 1003-1014 (2010).

4. M. Li, M. Boggs, T. P. Beebe and C. P. Huang, Carbon **46** (3), 466-475 (2008).

5. H. Wang, A. Zhou, F. Peng, H. Yu and J. Yang, J Colloid Interf Sci **316** (2), 277-283 (2007).

6. R. N. Bracewell, *The Fourier Transform and its Applications*, Third ed. (McGraw Hill, 2000).

7. T. R. Bedding, H. Kjeldsen, J. Reetz and B. Barbuy, Mon Not R Astron Soc **280** (4), 1155-1161 (1996).

8. S. Malhotra and J. E. Rhoads, Astrophys J **565** (2), L71-L74 (2002).

9. U. Munari and T. Zwitter, Astron Astrophys **318** (1), 269-274 (1997).

10. D. A. Shirley, Phys Rev B **5** (12), 4709-4714 (1972).

11. S. Tougaard, Surf Sci **216** (3), 343-360 (1989).

12. C.-H. Hung, L. A. Wiest, B. Singh, A. Diwan, M. J. C. Valentim, J. M. Christensen, R. C. Davis, A. J. Miles, D. S. Jensen, M. A. Vail, A. E. Dadson and M. R. Linford, J Sep Sci **36** (24), 3821-3829 (2013).

13. G. Saini, D. S. Jensen, L. A. Wiest, M. A. Vail, A. Dadson, M. L. Lee, V. Shutthanandan and M. R. Linford, Anal Chem **82** (11), 4448-4456 (2010).

14. L. A. Wiest, D. S. Jensen, C. H. Hung, R. E. Olsen, R. C. Davis, M. A. Vail, A. E. Dadson, P. N. Nesterenko and M. R. Linford, Anal Chem **83** (14), 5488-5501 (2011).

15. S. Tougaard, J Vac Sci Technol A **14** (3), 1415-1423 (1996).

16. D. S. Jensen, S. S. Kanyal, N. Madaan, A. J. Miles, R. C. Davis, R. Vanfleet, M. A. Vail, A. E. Dadson and M. R. Linford, J Vac Sci Technol B **31** (3) (2013).

17. J. Song, D. S. Jensen, D. N. Hutchison, B. Turner, T. Wood, A. Dadson, M. A. Vail, M. R. Linford, R. R. Vanfleet and R. C. Davis, Adv Funct Mater **21** (6), 1132-1139 (2011).

18. D. S. Jensen, S. S. Kanyal, V. Gupta, M. A. Vail, A. E. Dadson, M. Engelhard, R. Vanfleet, R. C. Davis and M. R. Linford, J Chromatogr A **1257**, 195-203 (2012).

19. J. Terry, M. R. Linford, C. Wigren, R. Cao, P. Pianetta and C. E. D. Chidsey, J Appl Phys **85** (1), 213-221 (1999).

20. J. Jeffrey Harve Terry, Stanford University, 1997.

21. J. Terry, R. Mo, C. Wigren, R. Y. Cao, G. Mount, P. Pianetta, M. R. Linford and C. E. D. Chidsey, Nucl Instrum Meth B **133** (1-4), 94-101 (1997).

22. R. L. Burden and J. D. Faires, *Numerical Analysis*, 9th ed. (Cengage, 2011).

23. D. S. Jensen, S. S. Kanyal, N. Madaan, M. A. Vail, A. E. Dadson, M. H. Engelhard and M. R. Linford, Surface Science Spectra **20** (1) (2013).

24. H. Ganegoda, D. S. Jensen, D. Olive, L. D. Cheng, C. U. Segre, M. R. Linford and J. Terry, J Appl Phys **111** (5) (2012).

25. C. K. Liu, J. M. Wu and H. C. Shih, Sensor Actuat B-Chem **150** (2), 641-648 (2010).

26. V. Datsyuk, M. Kalyva, K. Papagelis, J. Parthenios, D. Tasis, A. Siokou, I. Kallitsis and C. Galiotis, Carbon **46** (6), 833-840 (2008).

27. Y. Shao, G. Yin, J. Zhang and Y. Gao, Electrochimica Acta **51** (26), 5853-5857 (2006).

28. T. I. T. Okpalugo, P. Papakonstantinou, H. Murphy, J. McLaughlin and N. M. D. Brown, Carbon **43** (1), 153-161 (2005).

29. R. L. Burden and J. D. Faires, *Numerical Analysis*, Eighth ed. (Thomson Brooks/Cole, Belmont, CA, 2005).

30. R. J. Cole and P. Weightman, Journal of Physics: Condensed Matter **10** (25), 5679 (1998).

31. W. Olovsson, T. Marten, E. Holmström, B. Johansson and I. A. Abrikosov, J Electron Spectrosc **178–179** (0), 88-99 (2010).

32. T. Marten, W. Olovsson, S. I. Simak and I. A. Abrikosov, Phys Rev B **72** (5), 054210 (2005).

33. F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi, J. A. Yarmoff and G. Hollinger, Phys Rev B **38** (9), 6084-6096 (1988).

34. J. A. Carlisle, M. T. Sieger, T. Miller and T. C. Chiang, Phys Rev Lett **71** (18), 2955-2958 (1993).

Chapter 4: Comparison of the Equivalent Width, the Autocorrelation Width, and the Variance as Figures of Merit for XPS Narrow Scans*

4.1 Abstract

X-ray photoelectron spectroscopy (XPS) is widely used in surface and materials laboratories around the world. It is a near surface technique, providing detailed chemical information about samples in the form of survey and narrow scans. To extract the maximum amount of information about materials it is often necessary to peak fit XPS narrow scans. And while indispensable to XPS data analysis, even experienced practitioners can struggle with their peak fitting. In our previous publication, we introduced the equivalent width (EW_{XPS}) as both a possible machine automated method, one that requires less expert judgment for characterizing XPS narrow scans, and as an approach that may be well suited for the analysis of complex spectra. The EW_{XPS} figure of merit was applied to four different data sets. However, as previously noted, other width functions are also regularly employed for analyzing functions. Here we evaluate two other width functions for XPS narrow scan analysis: the autocorrelation width (AW_{XPS}) and the variance (σ^2_{XPS}). These widths were applied to the same four sets of spectra studied before: (a) four C 1s narrow scans of ozone-treated carbon nanotubes (CNTs) (EW_{XPS}: ~2.11 – 2.16 eV, AW_{XPS}: ~3.9 – 4.1 eV, σ^2_{XPS}: ~5.0 – 5.2 eV, and a modified form of σ^2_{XPS}, denoted σ^{2*}_{XPS}: ~ 6.3 – 6.8 eV), (b) silicon wafers with different oxide thicknesses (EW_{XPS}: ~1.5 – 2.9 eV, AW_{XPS}: ~2.28 – 4.9, and σ^2_{XPS}: ~ 0.7 – 4.9 eV), (iii) hydrogen-terminated silicon surfaces, before and after modification with pentyl groups, and after annealing of the pentyl-terminated surface (EW_{XPS}: ~ 0.7 – 1.0 eV, AW_{XPS}: ~1.2 – 1.6 eV, and σ^2_{XPS}:

*This chapter has been reproduced with permission from (Bhupinder Singh, Daniel Velázquez, Jeff Terry, and Matthew R. Linford), *Journal of Electron Spectroscopy and Related Phenomena*, 2014. **197**(0): p. 112-117

~0.12 – 0.19 eV), and (iv) C 1s narrow scans from five different nanodiamond samples, three of which showed charging (EW_{XPS}: ~2.6 – 4.8 eV, AW_{XPS}: ~3.8 – 6.9 eV, and σ^2_{XPS}: ~1.6 – 4.2 eV). All three of the width functions showed similar trends, except in the case of the C 1s spectra of the CNT samples, which were the most complex spectra evaluated, where σ^2_{XPS} showed poor correlation with the corresponding O/C ratios. Accordingly, we favor EW_{XPS} and AW_{XPS}. EW_{XPS} is advantageous because it is conceptually simple, giving the most intuitive results. AW_{XPS} has the advantage of not requiring the user to specify the height of the function at its maximum, which will be affected by noise. Because these functions are based on different mathematical operations/algorithms, best practices may involve the calculation of both widths for a set of narrow scans. The standard deviation, σ_{XPS}, i.e., the square root of the variance, was also examined. As expected, it gave results similar to σ^2_{XPS}.

4.2 Introduction

X-ray photoelectron spectroscopy (XPS) is an indispensable analytical tool for surface/material characterization. Indeed, an ISI Web of Science search on the terms 'X-ray photoelectron spectroscopy' or 'XPS' for papers published in 2013 identified more than 11,000 publications.[1, 2] XPS is a quantitative, near surface characterization tool that operates by illuminating a sample with X-rays.[3] The kinetic energies of the photoelectrons generated in the technique are measured by the instrument and converted into binding energies that identify the elements in the sample. XPS spectra are obtained as either lower resolution survey scans or higher resolution narrow scans. Peak fitting is a central part of the analysis of XPS narrow scans because, as noted by Sherwood, the widths of the peaks and the chemical shifts observed for different oxidation states of many elements are often quite similar.[4]

In our previous publication,[5] we discussed some of the limitations of traditional XPS peak fitting. We noted that while many analysts practice peak fitting with care and competence, the literature also contains many examples of poorly fit narrow scans. But even when peak fitting is practiced in a reasonable way, some ambiguity remains. And of course this ambiguity increases when a large number of adjustable parameters, sometimes of necessity, are used in a fit. With regard to peak fitting, we again quote Sherwood who said: "there is never a unique solution to fitting the data."[4] For example, Wepasnick and coworkers fitted a C 1s narrow scan of oxidized carbon nanotubes (CNTs) to the fitting parameters of similar spectra reported previously in two different papers.[6-8] Both sets of parameters provided good, overall fits to the data. However, with one set of parameters the fraction of carboxyl groups in the material was 5.9%, and in the other 11.0%. This is more than a trivial difference.

To help address these issues, we introduced the equivalent width for XPS narrow scans (EW_{XPS}) as a figure of merit that would potentially require less user intervention/be less biased than traditional peak fitting. We showed the effectiveness of the EW_{XPS} in the analysis of complex C 1s spectra from oxidized CNTs, which were similar to the data considered by Wepasnick. The equivalent width (EW) has previously been used in other areas of spectroscopy to characterize spectral lines. For example, the astronomy community regularly uses it to characterize peaks in their spectra.[9-11] The classical definition of the EW is the area of a function, *f(x)*, divided by its central ordinate, *f(0)*:

$$EW = \frac{\int_{-\infty}^{\infty} f(x)dx}{f(0)} \tag{4.1}$$

Thus, the EW models a function as a rectangle with the same area, $\int_{-\infty}^{\infty} f(x)dx$, and height, f(0), as the original function. Of course, Equation 4.1 is undefined for functions that pass through the origin, and makes little sense for functions that are well separated from the origin such that f(0) is small. Accordingly, we introduced a slightly modified form of Equation 4.1 for application to XPS narrow scans, which we named EW_{XPS}:

$$EW_{XPS} = \frac{Peak\ Area}{Peak\ Height} \tag{4.2}$$

where the 'Peak Area' is the area of an XPS peak above a user-defined background and 'Peak Height' is the height of the peak envelope maximum, also measured from the background. We also recommended that the binding energy at the maximum of the peak envelope (PE_{max}) be recorded with EW_{XPS} to avoid any ambiguity in the interpretation of the data. We used EW_{XPS} and PE_{max} to characterize four sets of XPS narrow scans,[5] which will again be used in this study. All of the results of the previous analyses were consistent with the expected chemistries of these materials.

Our previous publication focused exclusively on the equivalent width of an XPS narrow scan (EW_{XPS}) as a rather unbiased measure of its width. However, we noted that the EW is not the only measure of the width of a function.[5] Indeed, two other important widths of functions are the autocorrelation width and the variance.[12] In this work we describe these two widths, which we refer to herein as AW_{XPS} and σ^2_{XPS}, respectively. We apply these widths to the same four sets of XPS narrow scans that were analyzed in our previous publication, and compare the results obtained with these widths to those from the EW_{XPS}. As was the case for EW_{XPS}, AW_{XPS} performed well in all cases. σ^2_{XPS} performed well for three of the sample sets, but failed for the most complex set. Accordingly, we favor EW_{XPS} and AW_{XPS}. EW_{XPS} is advantageous in that

both its mathematical definition and results are straightforward – easier to understand. AW_{XPS} is advantageous because it does not require the user to calculate the height of the maximum of the peak envelope. The standard deviation, σ, was also evaluated as a width function, but it was found to be even less useful than the variance.

4.2.1 The Autocorrelation Width and the Variance

The autocorrelation of a real function, f, is defined as:

$$h(x) = f \star f = \int_{-\infty}^{\infty} f(u) f(u - x) du \tag{4.3}$$

For those not familiar with this mathematics, it is worth emphasizing that, on the left side of Equation 4.3 'x' plays the role of a variable (in h(x)), and that on the right side of Equation 4.3 it plays the role of a constant, i.e., within the integral. In an autocorrelation, one considers every possible shift, *x*, of a copy of a function with respect to the function itself, where for each shift of amount *x*, the two functions, *f(u)* and *f(u – x)*, are multiplied together and the area of the product function is recorded as the value of *h(x)*. Finally we note in passing that (i) all autocorrelations are even functions, i.e., whether a function is shifted a certain amount to the right or the same amount to the left with respect to itself the answer is the same, and (ii) the maximum in an autocorrelation function is at $x = 0$, i.e., at no shift.

The autocorrelation width of a real function is given by the following equation:[12]

$$W_{f*f} = \frac{\int_{-\infty}^{\infty}(f \star f)dx}{f \star f\,|_0} = \frac{(\int f dx)^2}{\int f^2 dx} = \frac{\int_{-\infty}^{\infty} h(x)dx}{h(0)} = AW_{XPS} \tag{4.4}$$

Here, the numerator ($\int_{-\infty}^{\infty}(f \star f)dx$) is just the integral, the area, of the complete autocorrelation function, and the denominator ($f \star f\,|_0$) is the autocorrelation function evaluated at $x = 0$, which is just the integral of the square of the function. That is, W_{f*f} is really just the equivalent width of the autocorrelation function, *h(x)*, i.e., it is the area of *h(x)* divided by its height, *h(0)*. Thus, we define AW_{XPS} according to Equation 4.4 in terms of the area and squared area of a narrow scan, i.e., $\frac{(\int f dx)^2}{\int f^2 dx}$. There are, however, some important differences between the EW and W_{f*f} (AW_{XPS}) of a function. First, in its classical definition, EW is dependent on the position of a function along the x-axis, while W_{f*f} is independent of that position. Second, there is no need to specially define the height of the integral of the autocorrelation function, i.e., the height of the function is given by $f \star f\,|_0 = \int f^2 dx$ = *h(0)*, while the height of EW_{XPS} needs to be specified. Third, one would in general expect the autocorrelation width of a function to be larger than its EW because the autocorrelation of a function should be broader than the original function. As we will show, AW_{XPS} can be directly applied to an XPS narrow scan, where as before, a background may or may not be applied to the data.

The spread in a set of data points is often measured by its variance, σ^2, given by:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N - 1} \tag{4.5}$$

where the x_i values here represent the data points, $\bar{x}$ is the mean value of the data set, and *N* is the number of data points. Of course the square root of the variance is the standard deviation, σ. To calculate σ^2 of an XPS narrow scan, we use the following equation for $\bar{x}$:

$$\bar{x} = \frac{\sum(count_i * x_i)}{\sum count_i} \tag{4.6}$$

where x_i refers to the binding energy (eV) at a particular point and $count_i$ refers to the number of counts at that binding energy. The following formula can then be used to calculate the variance of an XPS peak envelope, σ^2_{XPS}:

$$\sigma^2_{XPS} = \frac{\sum(count_i * (\bar{x} - x_i)^2)}{\sum count_i - 1} \tag{4.7}$$

Note that for the large numbers of counts typically encountered in an XPS narrow scan, $\sum count_i - 1 \approx \sum count_i$.

The EW_{XPS} and other width functions considered herein are utilized to measure the breadth of the envelope of a photoelectron spectrum. This envelope can change with sample changes (physical and chemical) and systematic problems such as charging. These functions are not designed to determine the components of the spectrum as in peak fitting[13] or principle component analysis.[14] Indeed, the EW and related functions are to be utilized in concert with the elemental analysis obtained from the core levels, and perhaps with regard to data from other characterization tools, to understand the nature of the peak width changes that are observed. Thus, these envelope functions allow the user to develop an understanding of chemical changes without the need for fitting. In particular, this approach should be valuable in the analysis and comparison of sets of very complex spectra where peak fitting is challenging. In addition, the width functions can identify spectral changes that show systematic errors, such as for charging.

4.3 Experimental

The same narrow scans used in our previous publication[5] on the EW_{XPS} are also used in the present study (see Chapter 3.3).

4.4 Results

In this work we consider the application of two new width functions to XPS narrow scans: the autocorrelation width (AW_{XPS}) and the variance (σ^2_{XPS}). To evaluate these figures of merit, we apply them to the same four sets of narrow scans that were studied with the EW_{XPS} in our previous publication.[5]

4.4.1 Data Set 1. Ozone-treated CNT forests

Figure 4.1a-c shows the EW_{XPS}, AW_{XPS}, and σ^2_{XPS} values of the C 1s spectra of four different samples of ozone-treated carbon nanotubes that had oxygen contents of 3.7 – 5.0 at. %. Shirley backgrounds[15] were first determined for all four spectra after which the different widths were calculated. As discussed previously,[5] with increased sample oxidation, a general broadening (increase in width) of the C 1s peak envelopes is expected. As previously reported, EW_{XPS} correlates quite well to the oxygen contents of the samples (correlation coefficient, R, of 0.989, see Figure 4.1a). AW_{XPS} followed the same trend and gave a similar correlation coefficient of 0.982 (see Figure 4.1b). Surprisingly, the σ^2_{XPS} values (see Figure 4.1c) appeared to be poorly correlated to the oxygen contents of the samples, i.e., they showed a substantially lower correlation to the O/C ratio (R = 0.661).

In an effort to understand the unexpected σ^2_{XPS} values for this set of CNT samples, we took a closer look at the equation for σ^2_{XPS} (Equation 4.7). Due to the natural asymmetry of the C

1s peak, we found that the $\bar{x}$ values for the four samples were different from their corresponding PE_{max} values (see Table 4.1). The value of $\bar{x}$ in Equation 4.7 is expected to have a significant effect on the final result. However, the value of $\bar{x}$ should also depend on how the data are processed, e.g., how the baseline is determined. Accordingly, we took the maximum of the peak envelope (PE_{max}) as the reference point, obtaining a modified form of σ^2_{XPS} that we denote as σ^{2*}_{XPS} in which we replace $\bar{x}$ with PE_{max}:

$$\sigma^{2*}_{XPS} = \frac{\sum(count_i * (PE_{max} - x_i)^2)}{\sum count_i - 1} \tag{4.8}$$

Figure 4.1d shows the graph of σ^{2*}_{XPS} plotted against the oxygen-to-carbon ratios for the CNT samples. The correlation with the XPS O/C ratio is a little better (R = 0.7493), but the results are still not nearly as good as those obtained with the EW_{XPS} and AW_{XPS}. Again, it is quite clear that the variance is sensitive to the value of $\bar{x}$ (PE_{max}) in Equations 4.7 and 4.8, and that determining the appropriate value of this parameter may be challenging.

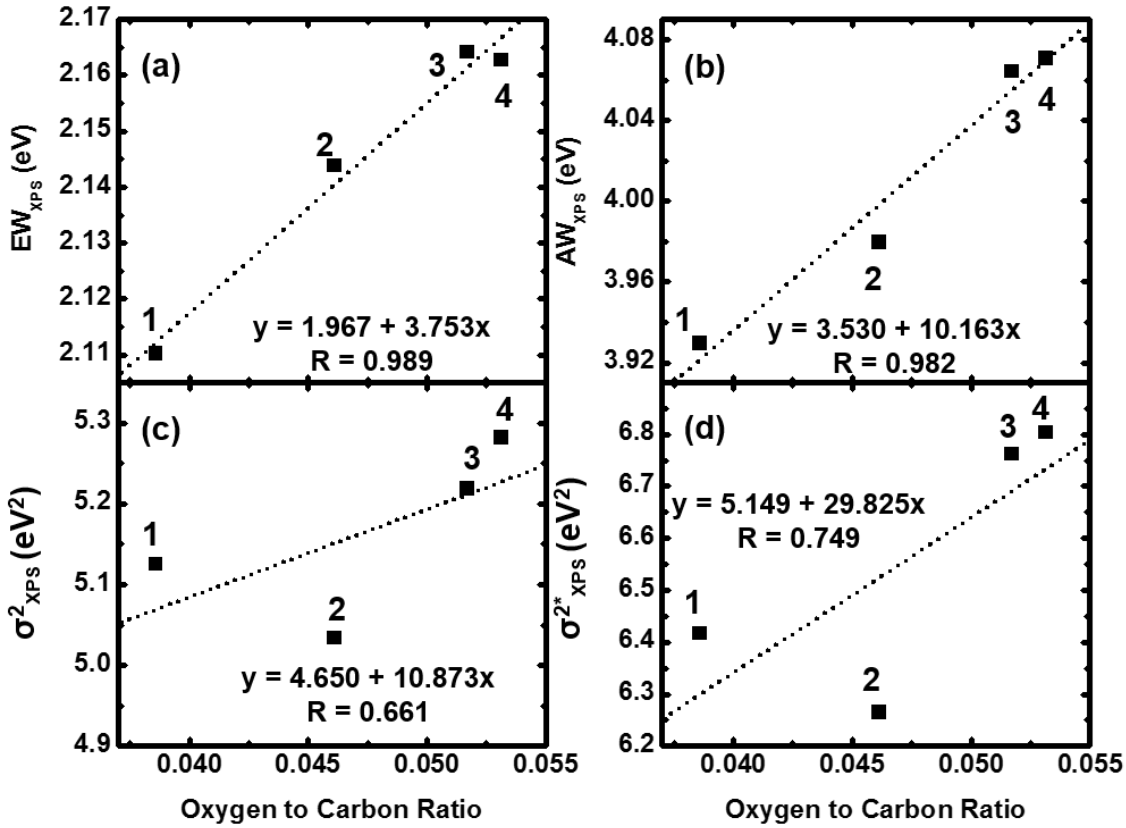


Figure 4.1 (a) EW_{XPS}, (b) AW_{XPS}, (c) σ^2_{XPS}, and (d) σ^{2*}_{XPS} widths of C 1 s narrow scans of four CNT samples plotted as a function of their oxygen-to-carbon ratios.

Table 4.1 EW_{XPS}, AW_{XPS}, σ^2_{XPS}, σ^{2*}_{XPS}, PE_{max}, and $\bar{x}$ values for four different ozone-treated CNT samples.

CNT sample	**O/C**	**EW_{XPS} (eV)**	**AW_{XPS} (eV)**	**σ^2_{XPS} (eV)**	**σ^{2*}_{XPS} (eV)**	**PE_{max}**	**$\bar{x}$**
1	0.03858	2.1103	3.9295	5.1250	6.4176	284.5	285.67
2	0.04611	2.1439	3.9797	5.0337	6.2652	284.6	285.71
3	0.0517	2.1641	4.0638	5.2185	6.7620	284.4	285.71
4	0.05314	2.1627	4.0707	4.2819	6.8037	284.4	285.71

Note on the significant figures in this document. We certainly do not claim that the numbers reported herein have significance to the number of significant figures reported in most cases below. We simply report here the output from our calculations.

4.4.2 Data Set 2. Silicon samples with different oxide thicknesses

The second set of spectra consisted of native oxide terminated silicon (Si_{NO}) and five different Si_{NO} surfaces that had been heated in the air at 900 °C for 1, 2, 3, 4, and 5 min. The samples are labeled according to the number of minutes they were oxidized. As before,[5] the Si 2p narrow scans were peak fitted to two signals, one attributed to the silicon substrate and the other to the surface oxide layer, and the AW_{XPS} and σ^2_{XPS} values of these 'Bulk Si' and 'SiO_2' component peaks were also calculated. Figure 4.2 shows the EW_{XPS}, AW_{XPS}, and σ^2_{XPS} values for the complete narrow scans and their average 'Bulk Si' and 'SiO_2' components as a function of the oxide thicknesses of the samples, which was measured by spectroscopic ellipsometry. The numerical values for these widths are given in Table 4.2. Qualitatively, all three measures of width gave similar results. That is, the widths of the average 'Bulk Si' and 'SiO_2' peaks were low, and the widths of the samples with substantial 'Bulk Si' and 'SiO_2' contributions were higher. Ultimately, as the oxide thicknesses increased, the width values decreased asymptotically to that of the 'SiO_2' sample. In all cases, the widths of the Si_{NO} narrow scans were greater than the widths of 'SiO_2' peak, which was consistent with the previous (EW_{XPS}) results. The σ^2_{XPS} arguably provides the best results. As the oxide layer thickness increases, σ^2_{XPS} shows a minimum at the 'bulk' SiO_2 value. In contrast, for the other two width functions, the 'bulk' SiO_2 value is a little higher than those for the oxidized samples.

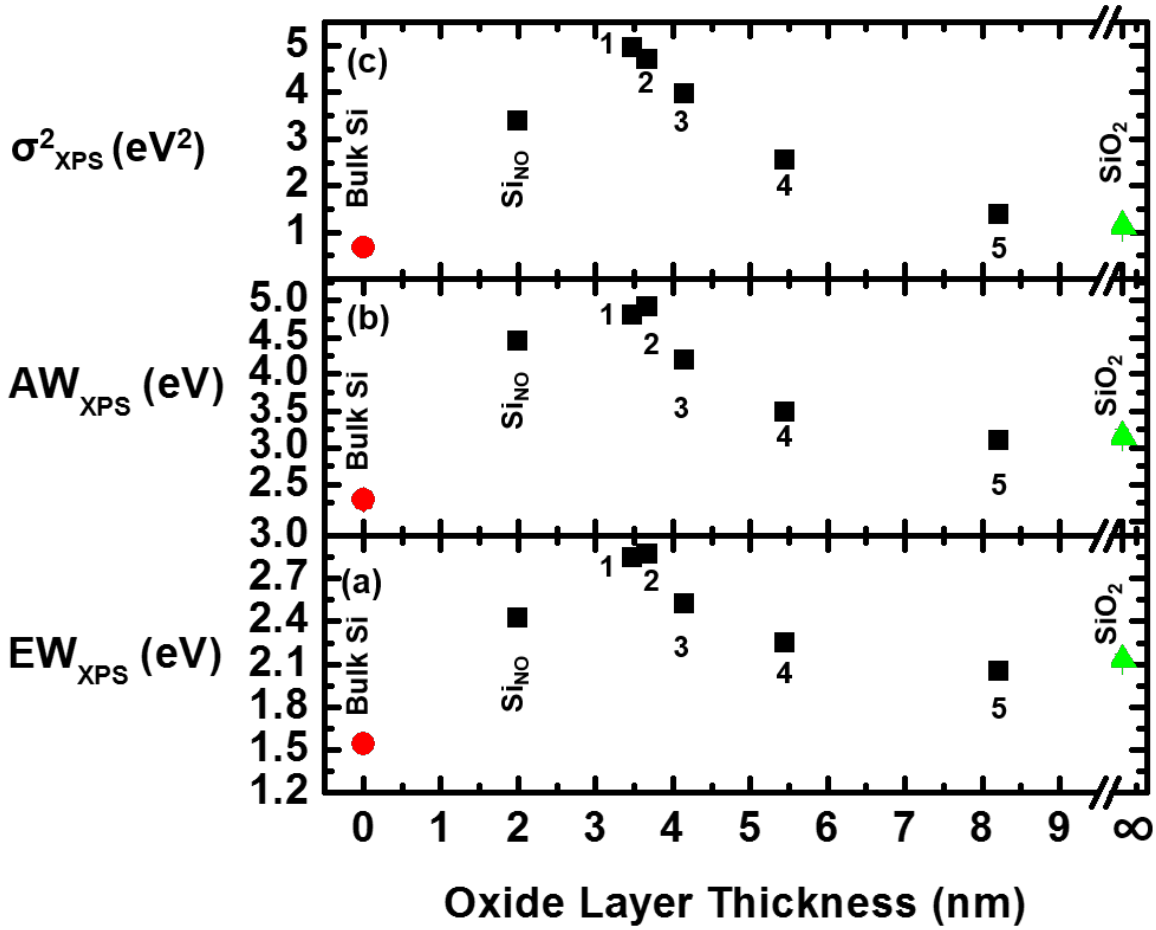


Figure 4.2 (a) EW_{XPS}, (b) AW_{XPS}, and (c) σ^2_{XPS} of Si 2p spectra from Si wafers with varied oxide thicknesses. The 'Bulk Si' peaks (circles) were calculated as the average width of the 'Bulk Si' peak component of the SiNO sample and samples 1 – 4. Sample 5 was excluded due to significant attenuation of the bulk Si signal. The 'SiO_2' peaks (triangles) were calculated as the average from the 'SiO_2' peak components of all the samples.

Table 4.2 EW_{XPS}, AW_{XPS} and σ^2_{XPS} values of the bulk Si and SiO_2 peaks obtained from peak fitting the Si 2p envelopes of silicon samples that had been oxidized for different amounts of time.

Si Sample	EW_{XPS} Si	EW_{XPS} SiO_2	AW_{XPS} Si	AW_{XPS} SiO_2	σ^2_{XPS} Si	σ^2_{XPS} SiO_2
Si_{NO}	1.5940	2.2660	2.3707	3.3327	0.7423	1.1999
1	1.5410	2.0800	2.2875	3.0818	0.7369	1.1250
2	1.5870	2.1410	2.3538	3.1655	0.7518	1.1495
3	1.5480	2.1410	2.2944	3.1664	0.6958	1.1417
4	1.4420	2.1070	2.1270	3.1219	0.5183	1.1232
5	1.1670	2.0390	1.7108	3.0218	0.2790	1.0497
Average	1.5424	2.1290	2.2867	3.1483	0.6890	1.1310
Standard deviation	0.0608	0.0775	0.0963	0.1056	0.0978	0.0488

4.4.3 Data Set 3. Hydrogen-terminated Si(111) and its derivatives

As in our previous work, Si 2p narrow scans of three different modified Si(111) surfaces were analyzed: (a) hydrogen terminated silicon (Si(111)-H), (b) Si(111)-H modified with pentyl groups (Si(111)-C_5), and (c) Si(111)-C_5 annealed at 700 °C (Si(111)-C_5/700°). All three widths (EW_{XPS}, AW_{XPS}, and σ^2_{XPS}) were calculated and showed similar results (see Figure 4.3). In all cases, (i) the widths of the narrow scan from the Si(111)-H and Si(111)-C_5 samples were nearly the same, and (ii) the width of the narrow scan from Si(111)-C_5/700° was substantially greater. Again, the annealing of the Si(111)-C_5 surface led to the formation of the Si(111)-(7 x 7) reconstruction, which produced a shoulder in the spectrum that broadened the narrow scan. In summary, these width functions 'see' the two similar spectra as similar and the broadened spectrum as wider.

4.4.4 Data Set 4. Sample charging

The C 1s spectra from five different nanodiamond samples were analyzed. Two of the narrow scans looked very reasonable, while three of them showed significant broadening and/or artifacts due to sample charging. Figure 4.4 shows the EW_{XPS}, AW_{XPS}, and σ^2_{XPS} values for these spectra plotted against their PE_{max} values, where these PE_{max} values were obtained after the attempted instrumental charge compensation, but before any correction of the binding energy scale. In all three cases, the 'good' spectra (circled in the figure) showed lower width values. They also showed PE_{max} values in a range that is more reasonable for a C 1s spectrum from an uncharged sample. Overall, the results from these three width functions are quite similar, although the discrimination between the 'good' and 'bad' samples/regions appears to be somewhat greater for EW_{XPS} and AW_{XPS} than for σ^2_{XPS}.

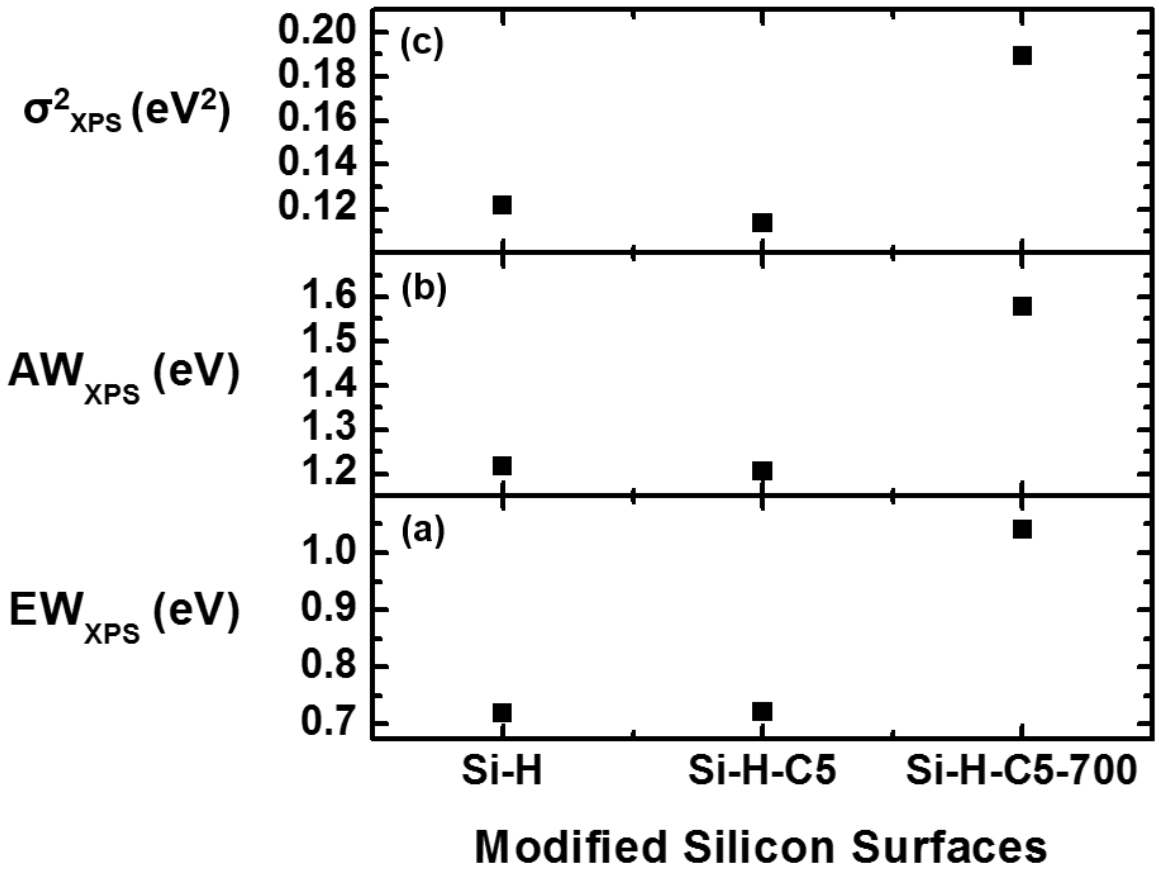


Figure 4.3 (a) EW_{XPS}, (b) AW_{XPS}, and (c) σ^2_{XPS} widths of Si 2p spectra of three different modified Si(111) surfaces: hydrogen terminated silicon (Si(111)-H, Si-H), Si(111)-H modified with pentyl groups (Si(111)-C5, Si-H-C5), and Si(111)-C5 annealed at 700 °C (Si(111)-C5/700°, Si-H-C5-700).

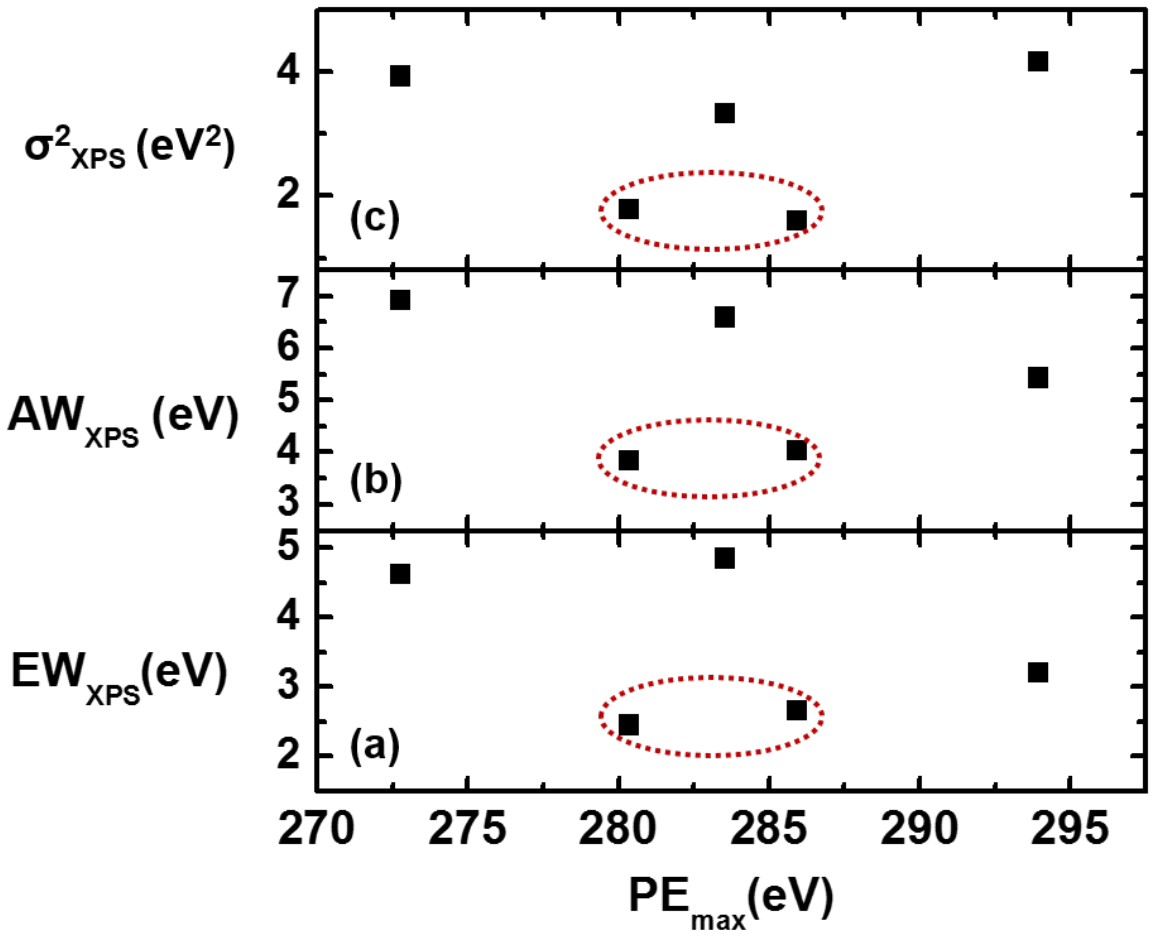


Figure 4.4 (a) EW_{XPS}, (b) AW_{XPS}, and (c) σ^2_{XPS} values of C 1s spectra of five different nanodiamond samples. The dotted red lines are guides to the eye to distinguish the two samples that did not charge and that showed good peak shapes from the samples that charged and showed artifacts.

4.4.5 The Standard Deviation

We also explored the standard deviation (σ) as a possible figure of merit for XPS narrow scans (σ_{XPS}). These functions are given by the following equations:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}} \tag{4.9}$$

$$\sigma_{XPS} = \sqrt{\frac{\sum(count_i * (\bar{x} - x_i)^2)}{\sum count_i - 1}} \tag{4.10}$$

where x_i, $\bar{x}$, N, and $count_i$ have the same meanings given them in Equations 4.6 – 4.8. In general, and as expected, σ_{XPS} performed similarly to σ^2_{XPS}. It worked well with Examples 2 and 3, failed for Example 1, and did not perform quite as well as σ^2_{XPS} in Example 4. We conclude that σ^2_{XPS} is a better analytical tool than σ_{XPS}. Figure 4.5 compares results for σ_{XPS} and σ^*_{XPS}, which is the square root of $\sigma^2{}^*_{XPS}$ (Equation 4.8).

4.5 Discussion

The EW_{XPS} values for the four sets of spectra studied here were all reasonable – this was reported previously. The EW_{XPS} figure of merit has the advantage of being easy to understand and explain. It is also computationally simple – just the area of a function divided by its height. The AW_{XPS} figure of merit does not provide the approximate width of the original function as the EW_{XPS} does. Rather it is the equivalent width of the autocorrelation of the function, which will be broader than the function from which it is derived. Thus it is a somewhat more abstract measure of width. Fortunately there is a simple formula for AW_{XPS} (see Equation 4.4), and it has

the advantage of not requiring the user to specify the height of the function, where this specification will be sensitive to noise. In this sense, it is arguably less biased than EW_{XPS}. In the examples reviewed herein, AW_{XPS} and EW_{XPS} show essentially the same effectiveness.

Variances are so widely used in science and technology that even though the σ^2_{XPS} width may be a little more difficult to calculate than the other two widths, there should be no barrier to its adoption. σ^2_{XPS} performed a little better than EW_{XPS} and AW_{XPS} in Example 2, about the same in Example 3, arguably not quite as well in Example 4, and very poorly in Example 1. Fixing the center point of the σ^2_{XPS} function improved the results in Example 1, but not nearly to the level of the EW_{XPS} and AW_{XPS} results. Therefore, σ^2_{XPS} appears to be a less robust width parameter than EW_{XPS} and AW_{XPS}. Overall, σ_{XPS} appears to be even less effective for analysis of XPS narrow scans.

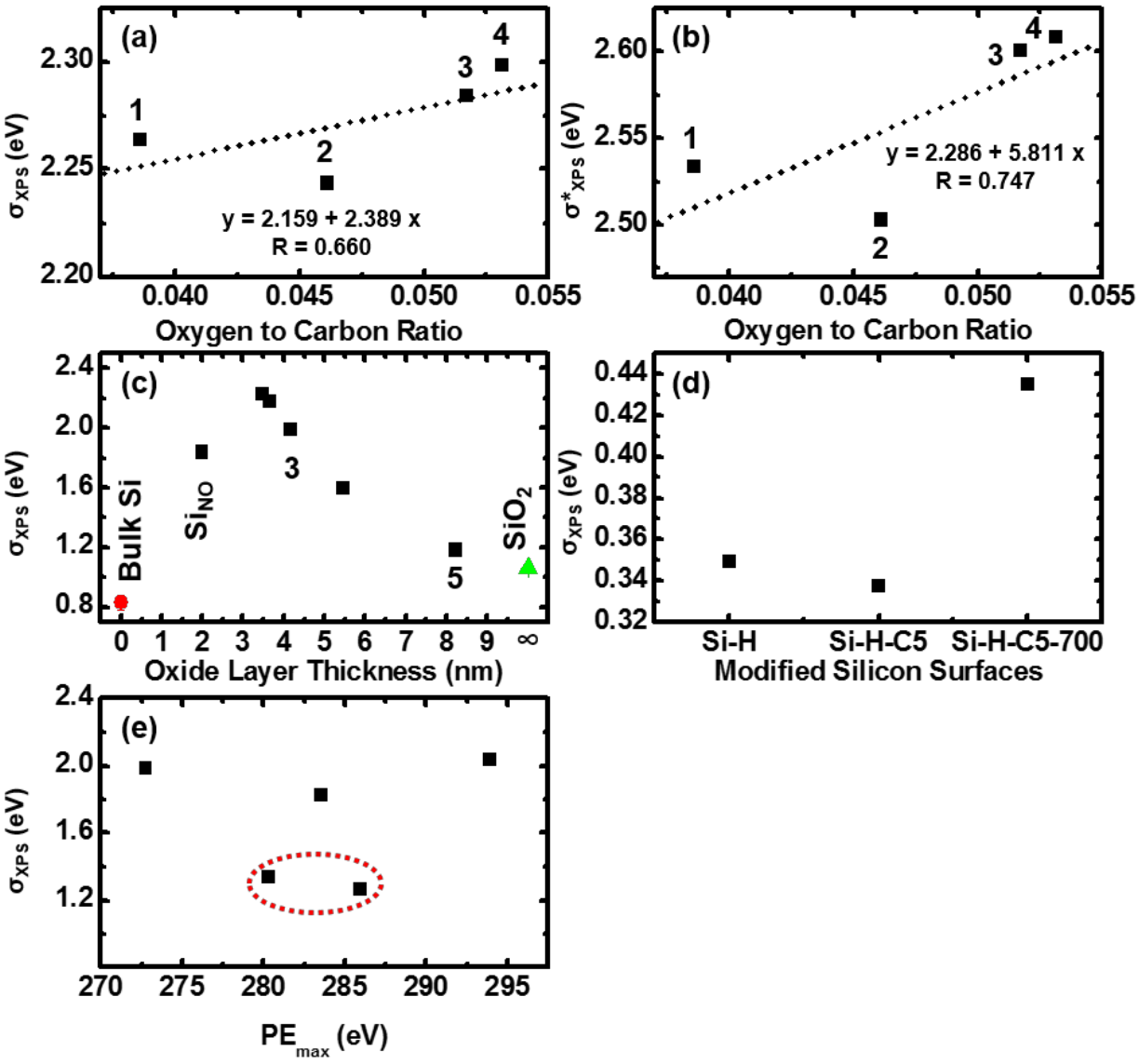


Figure 4.5 (a) σ_{XPS} and (b) σ^*_{XPS} values of C 1s spectra from oxidized CNT samples as a function of their oxygen-to-carbon ratios; σ_{XPS} value of (c) Si 2p narrow scans collected from a series of oxidized silicon surfaces, plotted as a function of their oxide layer thickness, (d) Si 2p narrow scans from various modified silicon surfaces, and (e) σ_{XPS} values plotted as a function of PE_{max} for C 1s spectra collected from various nanodiamond samples.

We have emphasized that EW_{XPS}, AW_{XPS}, and σ^2_{XPS} should constitute a less biased approach to XPS narrow scan analysis. However, we again acknowledge conventional peak fitting of XPS narrow scans as indispensable for gathering chemical information about a sample. Indeed, conventional peak fitting will remain important in spite of the fact that it is sometimes done poorly and/or an analysis may require a relatively large number of adjustable parameters. Conventional peak fitting breaks XPS peak envelopes into constituent peaks. Our width function approaches characterize the entire peak envelope. Recording PE_{max} with these widths also helps one to understand chemical changes taking place in a material. Possible applications/uses of EW_{XPS}, AW_{XPS} and σ^2_{XPS} include:

a) Capturing the similarity and variation between series of narrow scans.
b) Quality control of materials.
c) Being part of an expert/automated system for peak/spectrum analysis. We have included a flow chart in Figure 4.6 that might be used, at least in part, in such cases.
d) Analysis of very complex samples. EW_{XPS} and AW_{XPS} were able to pick up subtle differences in the complex spectra of oxidized CNTs, nicely correlating the results to their oxygen-to-carbon ratios.
e) As a complementary method to conventional peak fitting that could provide a rather unbiased check on results.

Some of us have spent years performing chemometrics analyses of moderately complex data sets – mostly positive and negative ion static SIMS spectra.[16-22] After a reasonable amount of practice, we have concluded that it is generally advisable to apply more than one chemometrics tool to a data set. That is, we will often perform at least two of the following analyses: principle components analysis (PCA), cluster analysis (CA), and/or multivariate curve

resolution (MCR) when we wish to study/explore the chemical variation between samples. Our experience has been that the different analyses almost always separate/differentiate our spectra in similar ways even though they are based on considerably different mathematical algorithms. When we see reasonable agreement between different analyses, we have more confidence in our conclusions. Clearly we do not have as much experience with the three width functions proposed in this paper. However, the results described herein suggest that it may also be wise to apply more than one width function to a set of narrow scans. As noted, our preferred functions are the EW_{XPS} and AW_{XPS} functions.

4.6 Conclusions

We have evaluated three width functions: the equivalent width (EW_{XPS}), the autocorrelation width (AW_{XPS}), and the variance (σ^2_{XPS}), for their abilities to analyze/characterize/differentiate XPS narrow scans. In three out of the four examples studied, all three of the width functions seemed to capture and correlate with the chemical and charging variations between samples in a reasonable way. In the case of the quite challenging C 1 s narrow scans of lightly oxidized CNTs, σ^2_{XPS} failed to deliver reasonable results. The standard deviation, σ_{XPS}, appeared to be even less effective than σ^2_{XPS}. EW_{XPS} stood out as being straightforward, robust, sensitive, and easy to calculate and understand. AW_{XPS} seemed to give very similar results to EW_{XPS} – like EW_{XPS} it worked for all four materials. While a little more abstract, AW_{XPS} has the advantage of being insensitive to noise in a spectrum, which can affect EW_{XPS} results. It may be advantageous to use more than one width function in the analysis of a set of XPS narrow scans to validate results. Here we would recommend the EW_{XPS} and AW_{XPS}.

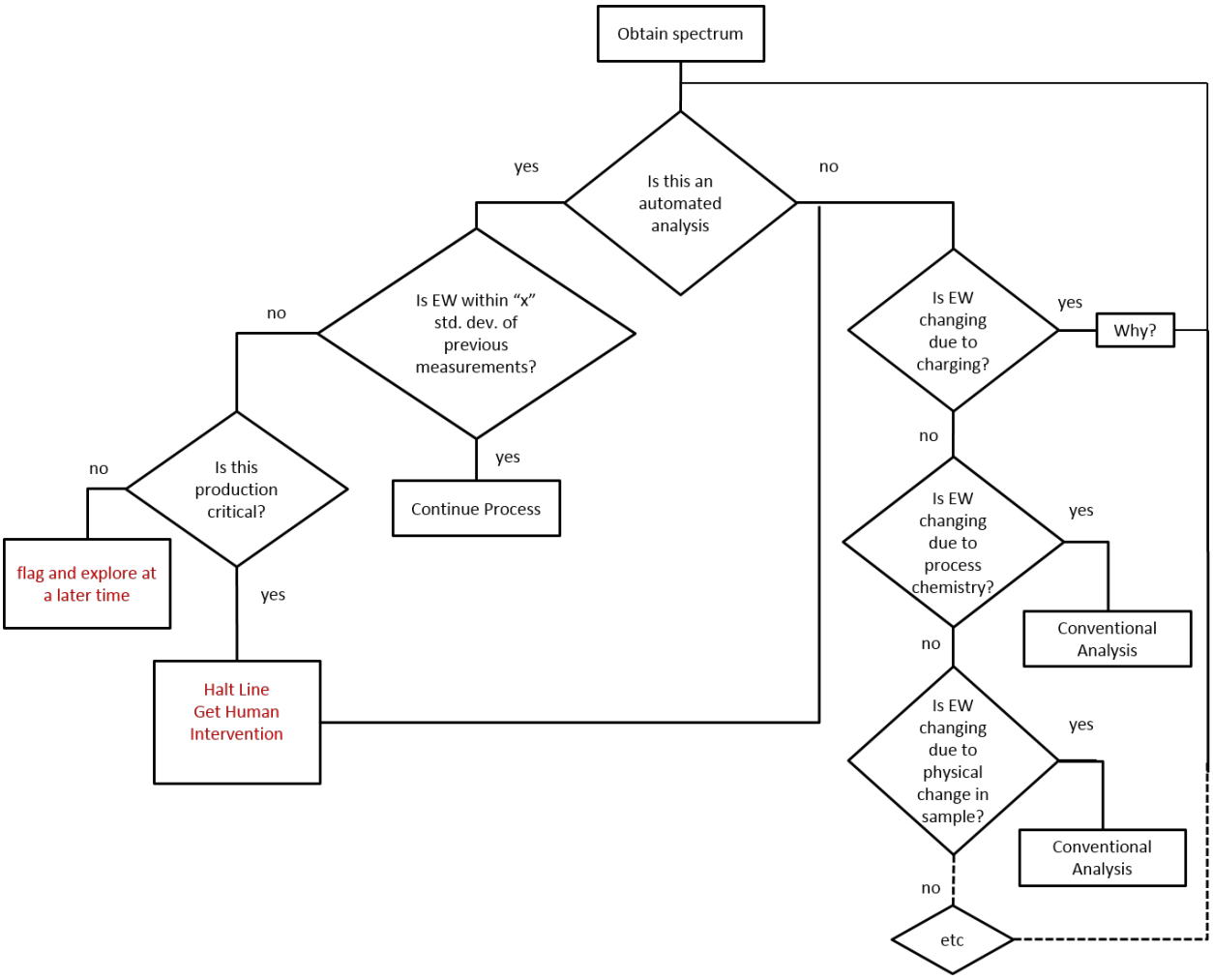


Figure 4.6 Flow chart depicting a possible analysis approach for using the width functions.

4.7 Acknowledgments

We thank the Department of Chemistry and Biochemistry and the College of Physical and Mathematical Sciences at Brigham Young University for their support of this work. DV was partially supported by the U.S. Department of Education through the GAANN Fellowship Program.

Author Contributions: M.R.L. suggested looking at other mathematical functions to characterize XPS narrow scans. B.S. did all the data analysis and comparison of different width functions mentioned in the manuscript, and made all the figures. D.V. crosschecked all the calculations. J.T. gave his feedback and technical inputs to the project. The manuscript was written by B.S and M.R.L.

4.8 Data Archiving

The raw data of the sets of narrow scans analyzed in this contribution will be submitted to Surface Science Spectra for archiving for the community.

4.9 References

1. C. J. Powell, J. Vac. Sci. Technol. A **21** (5), S42-S53 (2003).

2. C. J. Powell, J. Surface Analysis **20** (2014).

3. V. Gupta, H. Ganegoda, M. H. Engelhard, J. Terry and M. R. Linford, J Chem Educ **91** (2), 232-238 (2013).

4. P. M. A. Sherwood, J. Vac. Sci. Technol. A **14** (3), 1424-1432 (1996).

5. B. Singh, D. Velázquez, J. Terry and M. R. Linford, J Electron Spectrosc **197** (0), 56-63 (2014).

6. K. A. Wepasnick, B. A. Smith, J. L. Bitter and D. H. Fairbrother, Anal Bioanal Chem **396** (3), 1003-1014 (2010).

7. M. Li, M. Boggs, T. P. Beebe and C. P. Huang, Carbon **46** (3), 466-475 (2008).

8. H. Wang, A. Zhou, F. Peng, H. Yu and J. Yang, J Colloid Interf Sci **316** (2), 277-283 (2007).

9. T. R. Bedding, H. Kjeldsen, J. Reetz and B. Barbuy, Mon Not R Astron Soc **280** (4), 1155-1161 (1996).

10. S. Malhotra and J. E. Rhoads, Astrophys J **565** (2), L71-L74 (2002).

11. U. Munari and T. Zwitter, Astron Astrophys **318** (1), 269-274 (1997).

12. R. N. Bracewell, *The Fourier Transform and its Applications*, Third ed. (McGraw Hill, 2000).

13. B. Hellsing, A. Eiguren, F. Reinert, G. Nicolay, E. V. Chulkov, V. M. Silkin, S. Hufner and P. M. Echenique, J Electron Spectrosc **129** (2-3), 97-104 (2003).

14. J. Terry, R. K. Schulze, T. G. Zocco, J. D. Farr, J. Archuleta, M. Ramos, R. Martinez, R. Pereyra, J. Lashley, S. Wasserman, M. Antonio, S. Skanthakumar and L. Soderholm, AIP Conference Proceedings **532** (1), 364-366 (2000).

15. D. A. Shirley, Phys Rev B **5** (12), 4709-4714 (1972).

16. L. Yang, N. Shirahata, G. Saini, F. Zhang, L. Pei, M. C. Asplund, D. G. Kurth, K. Ariga, K. Sautter, T. Nakanishi, V. Smentkowski and M. R. Linford, Langmuir **25** (10), 5674-5683 (2009).

17. L. Yang, R. Bennett, J. Strum, B. B. Ellsworth, D. Hamilton, M. Tomlinson, R. W. Wolf, M. Housley, B. A. Roberts, J. Welsh, B. J. Jackson, S. G. Wood, C. L. Banka, C. D. Thulin and M. R. Linford, Anal Bioanal Chem **393** (2), 643-654 (2009).

18. V. Gupta, N. Madaan, D. S. Jensen, S. C. Kunzler and M. R. Linford, Langmuir **29** (11), 3604-3609 (2013).

19. L. Pei, G. Jiang, B. J. Tyler, L. L. Baxter and M. R. Linford, Energ Fuel **22** (2), 1059-1072 (2008).

20. M. V. Lee, M. T. Hoffman, K. Barnett, J. M. Geiss, V. S. Smentkowski, M. R. Linford and R. C. Davis, J Nanosci Nanotechno **6** (6), 1639-1643 (2006).

21. L. Yang, Y. Y. Lua, G. L. Jiang, B. J. Tyler and M. R. Linford, Anal Chem **77** (14), 4654-4661 (2005).

22. F. Zhang, R. J. Gates, V. S. Smentkowski, S. Natarajan, B. K. Gale, R. K. Watt, M. C. Asplund and M. R. Linford, J Am Chem Soc **129** (30), 9252-9253 (2007).

Chapter 5: Using a Simple Graphical Tool, Uniqueness Plots, to Avoid Peak Fit Parameter Correlation in X-ray Photoelectron Spectroscopy

5.1 Abstract

Data modeling and peak fitting are essential tools for the analysis of X-ray photoelectron spectroscopy (XPS) narrow scans. However, in general, peak fitting is subject to at least some measure of user bias. This problem deepens, especially in the case of novice practitioners who tend to employ too many fit parameters and/or fail to apply reasonable constraints to them. Under these conditions, fit parameters are often correlated, and therefore lacking in statistical meaning. The mathematical determination of correlation among XPS fit parameters has been reported in the literature. This statistical analysis is also used in the modeling of spectroscopic ellipsometry (SE) data. However, the SE community also employs a useful graphical tool, the uniqueness plot, to identify correlation between fit parameters. In these plots, the error of a fit is plotted as a function of the value of a particular fit parameter that is systematically fixed to values about its optimal value. Where results are unique, a parabola-like curve with significant curvature is obtained. On the flip side, a completely unreasonable peak fit, with correlated fit parameters, yields a horizontal line, i.e. the same (low) error is obtained for all values of the fit parameters in question. Herein, we introduce the uniqueness plot to the XPS community as a simple and straightforward graphical tool for assessing peak fitting results. We explore this approach with two data sets: a C 1s narrow scan from ozone-treated carbon nanotube forests and an Si 2p narrow scan from an air oxidized silicon wafer. We consider various scenarios for each of these data sets in which different numbers of fit parameters are varied. The resulting uniqueness plots are as expected, showing parabolic curves when more reasonable numbers of fit parameters and constraints are applied, and giving horizontal lines where too many fit parameters

are used without appropriate constraints. This latter case implies statistically meaningless results. We advocate that the ability to generate uniqueness plots be part of XPS instrument software packages, which will help XPS practitioners validate their peak fitting results.

5.2 Introduction

X-ray photoelectron spectroscopy (XPS) is a very widely used surface characterization technique, as evidenced by the fact that, as of late, the technique has been receiving approximately 10,000 citations per year.[1-3] XPS quantifies the elemental compositions of the near surface regions of samples, revealing the chemical states of the elements in them, often via peak fitting.[4, 5] However, peak fitting can be a somewhat subjective and challenging exercise.[6] To address the issue of peak fit uncertainty, we recently introduced the equivalent width and autocorrelation width as relatively less biased figures of merit for XPS narrow scans.[7, 8] These width functions were shown to be sensitive to chemical changes in a variety of materials. As an example of peak fit uncertainty, Wepasnick and coworkers peak fitted a C 1s narrow scan from oxidized carbon nanotubes (CNTs) using the parameters from two previously published fits.[9-11] In both cases, the overall fits were good. However, the fraction of the carboxyl, C(III),[12] signal in the fits differed substantially: ca. 6% and 11.0%. This difference, nearly a factor of 2, could certainly change one's understanding of this material.

There are a number of possible pitfalls in XPS peak fitting/data analysis. These include (i) using an inappropriate background, (ii) adding in too many peaks and/or fit parameters, and (iii) failing to evaluate the quality of one's fit by using tools such as chi squared (χ^2), the reduced chi squared (χ^{2*}), the summation of individual fit components, the Abbe criterion, etc.[13, 14] A particularly common error, especially among inexperienced practitioners, is the addition of

too many peaks/fit parameters and/or the failure to constrain them appropriately. When too many parameters can vary (float) in a fit, correlation is often present between them, which means that they lack meaning, i.e., have high standard deviations. The degree of correlation can be identified mathematically through the Hessian matrix, **H**

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 \chi^2}{\partial {p_1}^2} & \frac{\partial^2 \chi^2}{\partial p_1 \partial p_2} & \cdots \frac{\partial^2 \chi^2}{\partial p_1 \partial p_N} \\ \frac{\partial^2 \chi^2}{\partial p_2 \partial p_1} & \frac{\partial^2 \chi^2}{\partial {p_2}^2} & \cdots \frac{\partial^2 \chi^2}{\partial p_2 \partial p_N} \\ \vdots & \vdots & \cdots \vdots \\ \frac{\partial^2 \chi^2}{\partial p_N \partial p_1} & \frac{\partial^2 \chi^2}{\partial p_N \partial p_2} & \cdots \frac{\partial^2 \chi^2}{\partial {p_N}^2} \end{bmatrix} \tag{5.1}$$

where parameters, e.g., p_1 and p_2, are correlated if the mixed partial derivatives of χ^2 (the off-diagonal elements of this matrix) are non-zero.[13]

Peak fitting, per se, is not performed in the analysis of spectroscopic ellipsometry (SE) data. However, the work up of SE and XPS data are similar in that both involve the fitting of multiple parameters, and that when a large number of fit parameters are present in a modeling exercise they can be correlated. In practice, 'correlation' means that when a particular parameter changes, the other parameters in the fit can change in a compensatory way, giving the same error to the fit. Thus, correlated parameters lack meaning. The SE and XPS communities are well aware of the mathematical calculation of parameter correlation, including the parabolic relationship that is often found between chi squared and a given fit parameter.[13, 15, 16] However, in the analysis of SE data, correlation is also identified through the use of uniqueness plots, which are graphical tools that can be easily interpreted.[17-19]

In a uniqueness plot, the error to the fit is plotted as a function of one of the variables (fit parameters), where this parameter is repeatedly fixed to different values about its optimal value, while the remaining fit parameters are allowed to vary (float) in the model. That is, for each fixed value of the parameter, the error of the fit is calculated and plotted. If the error in the fit rises as the parameter in question is varied, in general in a parabolic fashion,[13] the fit is deemed to be unique. However, if the fit parameter can be systematically varied and the error of the fit stays constant, there is no 'uniqueness' to the fit. This means that any value of the parameter will give the same result (the same error), which clearly means that the fit parameter is meaningless. Figure 5.1 shows an example of an SE uniqueness plot. It consists of SE data collected from a ca. 500 nm film of silicon dioxide (SiO_2) on silicon. The data were modeled in standard fashion[20, 21] using four layers: (i) the silicon substrate (it was assumed to be optically opaque), (ii) an Si/SiO_2 interface layer (the optical constants from the software were used), (iii) an SiO_2 layer modeled with a Cauchy dispersion relationship, and (4) a roughness layer that was modeled using a Bruggeman effective medium approximation (BEMA).

In the first approach, reasonable constraints were applied to the model, i.e. the optical constants of the Si substrate were taken from the instrument software and fixed, the thickness of Si/SiO_2 interface layer was fixed at 1 nm and its optical constants were also taken from the software, and the SiO_2 layer was modeled by fitting the 'A' and 'B' terms of the Cauchy dispersion relationship. The thickness of the SiO_2 layer was also allowed to vary. For the BEMA, the void and SiO_2 fractions were fixed at 50 % each and the thickness of this roughness layer was allowed to float. Using this approach, the optimum thickness of the SiO_2 layer was determined to be 511 nm. Accordingly, we probed the uniqueness of our results by varying the SiO_2 thickness from 500 – 530 nm, recording the mean square error (MSE) of the fit at each value of the

thickness. Figure 5.1 shows the resulting uniqueness plot (plot of MSE vs. SiO_2 thickness), which showed a parabola form. This plot implies a unique result because the errors increase as we vary the thickness of the SiO_2 layer from its optimal value. For the second approach, the model was unconstrained, i.e. the optical constants of the Si substrate, the optical constants of the Si/SiO_2 interface and its thickness, the void fraction of the roughness layer, and its optical constants, were all allowed to float, along with all the parameters in the first approach. As expected, Figure 5.1 shows a non-unique result here. That is, the error of the fit does not change when the thickness of the SiO_2 layer is varied, i.e., the thickness of the SiO_2 layer has no meaning. Clearly, the uniqueness plots in Figure 5.1 constitute a simple, straightforward, and valuable graphical tool that can aid both the novice and experienced SE practitioner to pinpoint unreasonable peak fitting efforts.

It has been well observed that peak fitting is indispensable for XPS narrow scan analysis.[3, 7, 8] In this study we apply the concept of the uniqueness plot to the fitting of XPS narrow scans. The purpose of this study is to illustrate the use of a graphical tool that can aid XPS practitioners in identifying when their peak fitting approaches are inappropriate, i.e., when their peak fit parameters are correlated and therefore lacking in meaning. We emphasize that the uniqueness plot is most useful in identifying poor peak fitting efforts, as opposed to better ones. Indeed, because there are (i) quite a few examples of questionable peak fitting on XPS data in the literature, and (ii) the popularity of XPS has increased dramatically over the recent years,[2, 3] we believe that this tool will be of use to inexperienced and experienced practitioners alike. Those that are inexperienced will quickly identify unconstrained, inappropriate models, and those who are more experienced will quickly confirm that their models are more reasonable. To demonstrate this approach, we fit various narrow scans, including the C 1s spectrum of an ozone-

treated carbon nanotube forest (CNT) and the Si 2p spectrum from an air-oxidized silicon substrate. The uniqueness approach performed as expected – parabolic uniqueness plots were obtained when a limited number of fit parameters were employed and more shallow curves, or horizontal lines, were observed when larger numbers of unconstrained parameters were employed. Both reduced chi squared (χ^{2*}) and Abbe criterion values were explored in these plots. The uniqueness plot is considered to be a standard feature of SE data analysis software. We recommend that this feature similarly be added to XPS fitting programs to aid the user in understanding the validity of his/her peak fitting.

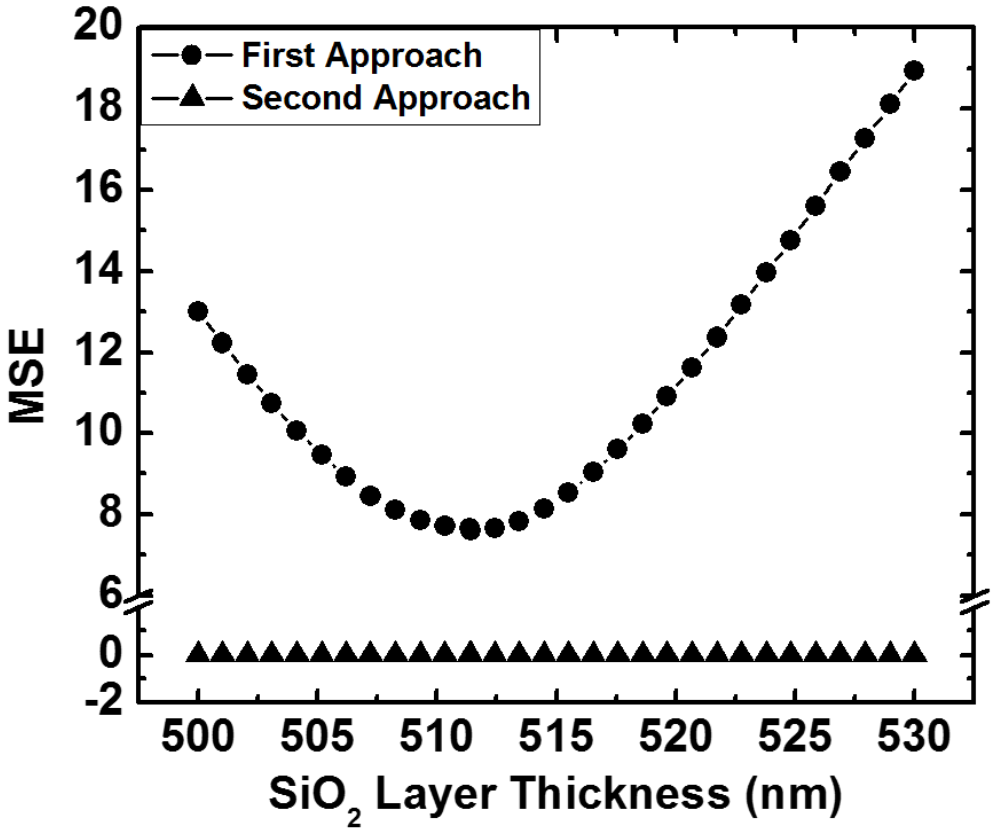


Figure 5.1 Uniqueness fit showing the MSE as a function of the fixed SiO_2 layer thickness for spectroscopic ellipsometry data collected of an Si/SiO_2 sample. The first approach involves reasonable constraints on the fit parameters. In the second approach, the fit parameters are unconstrained.

5.3 Experimental

Two different sets of XPS data were employed in the present study:

(a) A C 1s narrow scan from an ozone-treated carbon nanotube (CNT) forest, with an oxygen content of 3.7 at. % by XPS. This CNT spectrum was previously collected as part of a study on infiltrated, CNT-templated thin layer chromatography plates.[22]

(b) An Si 2p narrow scan of a native-oxide terminated silicon shard[23] (ca. 1.5 x 1.5 cm^2) that was air oxidized for one minute at 900 °C in a Thermolyne™ benchtop muffle furnace from Thermo Scientific. The resulting oxide layer was 3.58 nm by SE. The optical constants of Si and SiO_2 for the SE modeling of this material were obtained from the instrument software.

Instrument settings:

XPS analyses of the CNT sample and silicon substrate were performed using a Surface Science SSX-100 X-ray photoelectron spectrometer (serviced by Service Physics, Bend, OR) with a monochromatic Al Kα source (1486.7 eV) and a hemispherical analyzer. Narrow scans were recorded with a spot size of 500 µm x 500 µm, resolution: 3 (nominal pass energy 100 eV), number of scans: 20 and step size of 0.065 eV. The CNT spectrum was referenced to the main C 1s peak in the narrow scan, which was taken at 284.6 eV, and the silicon spectrum was referenced to the O 1s peak taken at 532.0 eV. This latter signal appeared as a single, shoulderless peak. Peak fitting was done in our instrument software, and fits were optimized based on a minimization of the χ^{2*}. All of the backgrounds used in the fits reported in this work are of the Shirley type.[24]

SE data were collected on an M-2000 instrument from the J.A. Woollam Co., Lincoln, NE, over a wavelength range of ca. 200 – 1700 nm. For Figure 5.1, the data were range selected from 300 – 400 nm. When a wider wavelength window was used with an unconstrained model the time necessary for the software to generate the uniqueness plot was extremely long.

5.4 Theory of Error Estimation Tools

The error estimation tools employed herein are the reduced chi squared (χ^{2*}) and Abbe criterion values. These error estimation tools help determine the quality of a fit vis-à-vis the difference between the experimental data and the modeled (fit) envelope. The fit envelope is the sum of the individual fit components in a model.

The chi squared value is defined by the following equation:

$$\chi^2(\vec{p}) = \sum_{i=1}^{N} \frac{[M(i) - S(i,\vec{p})]^2}{M(i)} \tag{5.2}$$

Where, *M(i)* and $S(i,\vec{p})$ refer to the values of the experimental data and fit at a given point. The vector, $\vec{p}$, contains the fit parameters used in the model. The difference between the experimental data and fit envelope is squared at each data point and then normalized to the value of the experimental data point. The reduced chi squared value is defined as:

$$\chi^{2*}(\vec{p}) = \frac{\chi^2(\vec{p})}{N - P} \tag{5.3}$$

where *N* is the number of data points and *P* is the number of fit parameters employed.

Obviously, a lower value of χ^2 or χ^{2*} implies a better fit. However, as Sherwood clarified,[4] χ^2 is of greatest value when it is reported for a series of peak fits on the same spectrum, or when applied to similar spectra.

The residuals, *R(i)*, of a fit are given by $S(i,\vec{p}) - M(i)$.

The Abbe criterion is defined as:

$$Abbe = \frac{1}{2}\frac{\sum_{i=1}^{N-1}[R(i+1) - R(i)]^2}{\sum_{i=1}^{N}[R(i)]^2} \tag{5.4}$$

Where *R(i + 1)* and *R(i)* refer to the residuals at the $(i + 1)^{th}$ and i^{th} data points, respectively. An Abbe criterion of zero indicates that the fit envelope consistently misses the experimental peak envelope by a fixed amount, i.e., *R(i) = R(i + 1)*. An Abbe criterion value of two is obtained for anti-correlated residuals. When the residuals are statistically distributed as random noise, the Abbe criterion is unity. This latter scenario is the ideal one.

Finally, it is vital to use common sense when peak fitting XPS narrow scans.[4, 14] For example, the Abbe criterion and χ^{2*} values can be excellent in the case of a meaningless fit where too many parameters are allowed to vary.

5.5 Results and Discussion

To show the usefulness of uniqueness plots in XPS narrow scans analysis, we consider narrow scans that are peak fitted with and without a reasonable number of constraints. As will be shown below, as the number of unconstrained fit parameters increases, the uniqueness of the fits decreases.

5.5.1 Ozone-treated CNT sample (varying C-C peak width)

In this section we will describe:

a. A reasonable peak fit to a C 1s narrow scan from ozone-treated CNTs
b. The systematic variation of a parameter in this fit (the width of the C-C peak) to generate uniqueness plots, where the number of parameters constrained in the fits was gradually decreased until no constraints were applied to it.

The reasonable peak fit of the narrow scan was obtained as follows. According to theory, the C 1s narrow scan was fit to a series peaks that represent: (i) carbon bonded to carbon (and obviously hydrogen), denoted C-C (this peak was initially placed at ca. 284.6 eV, but then allowed to float – we designate its variable position as *x* eV), (ii) carbon bonded to one oxygen atom C-O at *x* + 1.5 eV, (iii) carbonyl carbon C=O at *x* + 2.7 eV, and (iv) carboxyl carbon O-C=O at *x* + 3.9 eV. These peaks could also be designated by the oxidation states of the carbon they represent: C(0), C(I), C(II), and C(III), respectively.[7] In addition, as would be appropriate for this aromatic material, a shake-up peak was included at x + 6.5 eV. All peak widths were constrained to have the same value, except the shake-up peak, which had double the width of the other peaks. All the peaks were symmetric and modeled as Gaussian-Lorentzian (G:L) sum functions,[5] where the G:L ratio of the peaks varied in the fits, but the same G:L ratio was maintained for all the peaks. The heights of all the peaks were allowed to float. Of course these assignments, and all those considered in this work, are semi-empirical in nature. Figure 5.2a shows the optimal fit obtained through this approach, which gave a C-C peak width of 1.37 eV and a G:L ratio of 93%.

In a uniqueness fit, one fit parameter is systematically fixed to different values about its optimal value, while the other parameters are allowed to float to obtain the best possible fit. This is illustrated graphically in Figures 5.2b – d. Here, the width of the C-C peak (C(0)) is given

values of 1.0, 1.2, and 1.9 eV. In this particular case, the only parameter that was allowed to float was the G:L ratio of the C-C peak – obviously one would not expect changes in the G:L ratio to fully compensate for such large changes in the C-C peak width. The lack of fit in Figures 5.2b – d, is clearly shown in the residuals. That is, in Figure 5.2a, the best fit, the residuals suggest a reasonable fit, but substantially larger residuals are observed for the non-optimal peak widths.

A more quantitative way of describing the quality of the fits in Figure 5.2 is through χ^{2*} and/or Abbe criterion values (see section on 'Theory of Error Estimation Tools' above). Figure 5.3a shows a plot of the χ^{2*} values for the fits in Figure 5.2 plotted as a function of the width of the C-C peak (the black squares). Also included in this plot are the χ^{2*} values corresponding to C-C peaks with widths that were systematically varied from 0.8 to 2.0 eV. This plot is a uniqueness plot. Its parabolic shape indicates that there is uniqueness, or meaning, to the C-C peak width in the modeling approach taken for Figure 5.2. As expected, the minimum error (χ^{2*} value) is obtained at the optimal value of the C-C peak width. Figure 5.3b also shows the Abbe criterion values as a function of the peak width for the fits in Figure 5.2 (see again black squares). Here, we see a maximum value for the Abbe criterion values (values closer to unity) at the optimal C-C peak width. Then, as the peak width deviates from its best value, the Abbe criterion values drop, which is consistent with the more pronounced residuals in Figure 5.2 for the less optimal C-C peak widths.

To demonstrate the use of uniqueness plots in XPS peak fitting, the constraints on the fit parameters for the fit represented in Figure 5.2 were increasingly relaxed (see Table 5.1). The resulting uniqueness plots for each set of conditions were then plotted in Figure 5.3. The general result here was that as the constraints on the fit were removed, the curves in the χ^{2*} vs. C-C peak width plot became increasingly flat – they fan out and show less curvature, while the minimum

value in χ^{2*} drops somewhat. Finally, when all the constraints are removed from the fit (Approach 8), a horizontal line is obtained (dark green hexagons). This flat line means that the C-C peak width can be varied over a wide range of values, and that for each of these values the other parameters in the fit can compensate for any change in it, always yielding the same, low error. In other words, for this unconstrained approach, the C-C peak width, and presumably all the other fit parameters, lack any statistical meaning.

Figure 5.4 shows four examples of the peak fitting from Approach 8 (the unconstrained approach). Here we see four fits that show equally low χ^{2*} values (low residuals), but with significantly different peak/fit components, e.g., the fit components/peaks obviously have different positions and widths (the G:L ratios of the peaks also differed within each fit). The positions and widths of the peaks in these fits are clearly unreasonable. For example, it is hard to imagine how the light blue and dark blue peaks in Figure 5.4a, which have very different widths, could be justified chemically. In addition, the shake-up signals have become excessively broad in all the spectra. The high Abbe criterion values in Figure 5.3b for Approach 8 are consistent with the low residuals for the fits in Figures 5.4a – d. Indeed, the Abbe criterion (Equation 5.4) is very sensitive to small, consistent deviations between the fit envelope and experimental data. To some degree, the traces in Figure 5.3b are a reversal of those in Figure 5.3a. Where the χ^{2*} vs. C-C peak width traces are parabolic, the Abbe criterion vs. C-C peak width plots are Gaussian in form. Also, the horizontal line at the bottom of Figure 5.3a corresponds to the approximately horizontal line at the top of Figure 5.3b. In summary, both plots appear to give similar information, although the χ^{2*} vs. fit parameter plot appears to be more useful.

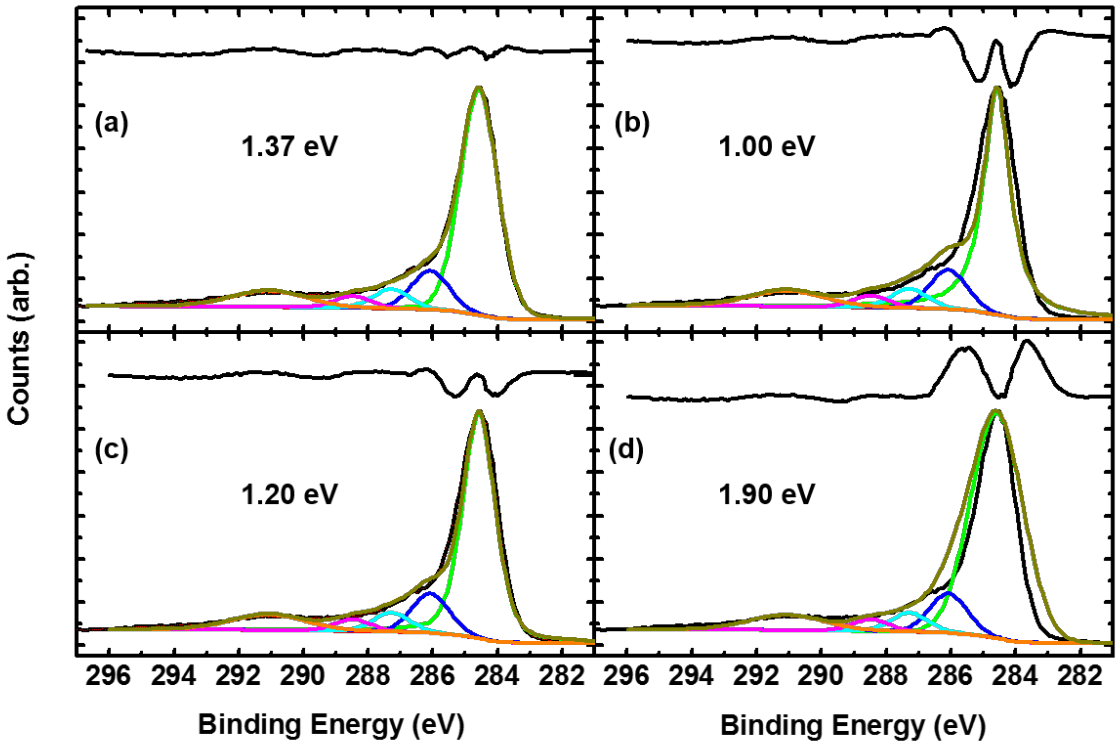


Figure 5.2 (a) Peak fitting of a C 1s narrow scan of ozone-treated CNTs obtained using Approach 1 (the most constrained approach considered, see Table 5.1) showing the optimal peak width value of 1.37 eV. (b – d) Poorer peak fits for constrained C-C peak widths of 1.00, 1.20, and 1.90 eV, respectively. The black traces at the top of each panel represent the residuals to the fits.

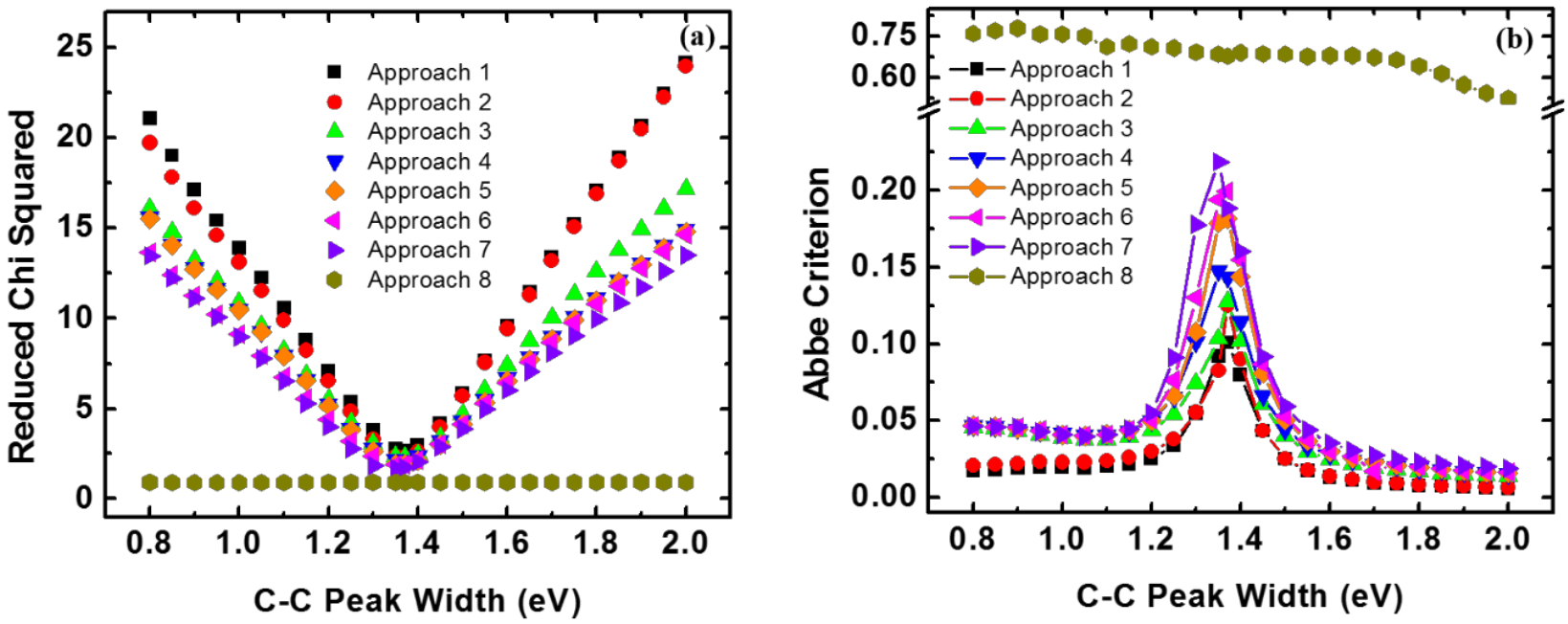


Figure 5.3 Uniqueness plots for the fits to a C 1s narrow scan of ozone-treated CNTs using the eight approaches described in Table 5.1. See also Figures 5.2 and 5.4. Plots of (a) reduced chi squared and (b) the Abbe criterion value vs. a fit parameter (the C-C peak width) that was fixed to specific values.

Table 5.1 Approaches used to generate the eight uniqueness plots in Figure 5.3. The number of fixed parameters was decreased proceeding from Approach 1 to Approach 8. All fits were to the C 1s spectrum of ozone-treated CNTs shown in Figures 5.2 and 5.4, and are based on changing the width of the C-C peak.

Approach	**Fixed**	**Floated in the Fit**	**Varied in the Uniqueness Test**
1	All peak positions, peak widths, peak heights and Gauss % (except C-C Gauss %)	Gauss % of C-C peak	The peak width of the C-C peak
2	All peak positions, peak widths, and peak heights	Gauss % of all peaks	The peak width of the C-C peak
3	All peak positions, peak widths, and peak heights (except C-C peak)	Gauss % of all peaks, and peak height of C-C peak	The peak width of the C-C peak
4	All peak positions, and peak widths	Gauss % and peak heights of all peaks	The peak width of the C-C peak
5	All peak positions, and peak widths (except shake-up peak)	Gauss %, peak heights of all peaks, and peak width of shake-up peak	The peak width of the C-C peak
6	All peak positions	Gauss %, peak heights, and peak widths of all peaks	The peak width of the C-C peak
7	All peak positions (except C-C peak)	Gauss %, peak heights, and peak widths of all peaks, peak position of C-C peak	The peak width of the C-C peak
8	Nothing	All Gauss %, peak heights, all peak positions, and all peak widths	The peak width of the C-C peak

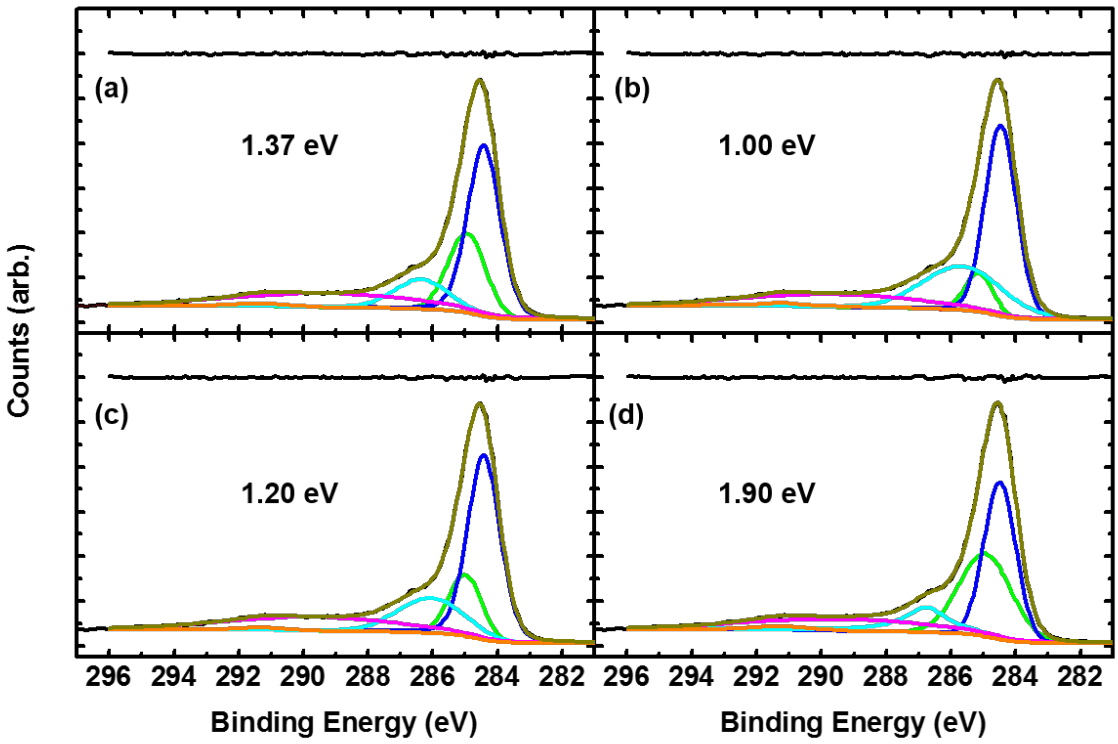


Figure 5.4 (a) Peak fitting of a C 1s narrow scan of ozone-treated CNTs obtained using Approach 8 (the least constrained approach considered, see Table 5.1) showing the optimal peak width value of 1.37 eV. (b – d) Equally unreasonable peak fits for constrained C-C peak widths of 1.00, 1.20, and 1.90 eV, respectively. The black traces at the top of each panel represent the residuals to the fits.

5.5.2 Ozone treated CNT sample (varying C-O peak position)

To further illustrate the properties of uniqueness plots, we peak fitted the same C 1s narrow scan of oxidized CNTs used in the previous example (see Figures 5.2 and 5.4), while varying the C-O peak position instead of the C-C peak width. The optimal C-O peak position had been determined to be 286.07 eV (see Figure 5.2a). Accordingly, the C-O peak position was systematically varied (and fixed at each point) at 285.50 – 286.50 eV in increments of 0.05 eV, and the χ^{2*} and Abbe criterion values were calculated at each of these values. Eight peak fitting approaches were considered (see Table 5.2), where they again ranged from a most constrained (Approach 1) to a least constrained (Approach 8) approach. As expected, in the plot of χ^{2*} vs. the C-O peak position, the most constrained approach yielded the curve (parabola) with the highest curvature. Then, as the constraints on the fit were relaxed, the uniqueness plots showed less curvature (they fanned out), and the unconstrained approach showed the lowest χ^{2*} value (see Figure 5.5a) with extremely low curvature. This latter result again illustrates that for a significantly unconstrained model the fit parameters are statistically meaningless. Figure 5.5b shows a uniqueness plot with the Abbe criterion value as the figure of merit. As before (see Figure 5.3), it shows an approximately inverse relationship to the plot in Figure 5.5a.

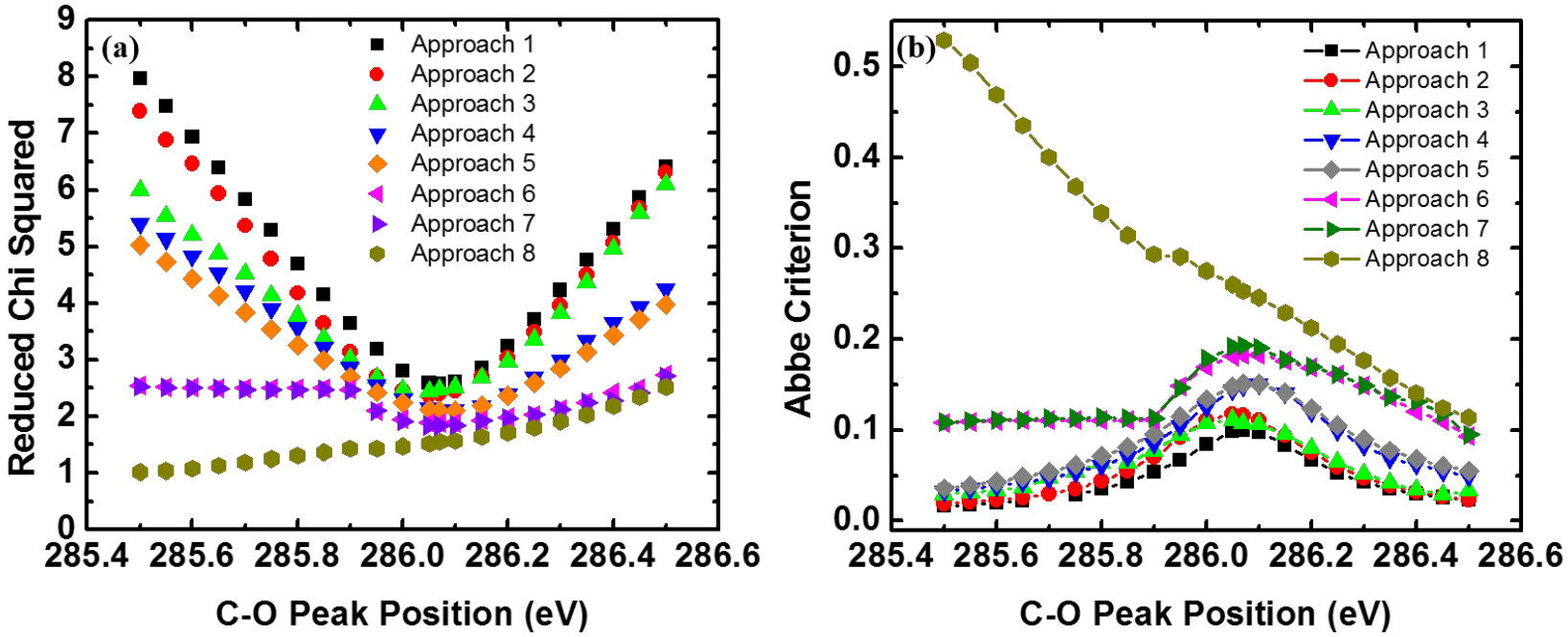


Figure 5.5 Uniqueness plots for the fits to a C 1s narrow scan of ozone-treated CNTs using the eight approaches described in Table 5.2. See also Figures 5.2 and 5.4. Plots of (a) reduced chi squared and (b) the Abbe criterion value vs. a fit parameter (the C-O peak position) that was fixed to specific values.

Table 5.2 Approaches used to generate the eight uniqueness plots in Figure 5.5. The number of fixed parameters was decreased proceeding from Approach 1 to Approach 8. All fits were to the C 1s spectrum of ozone-treated CNTs shown in Figures 5.2 and 5.4, and are based on changing the position of the C-O peak.

Approach	Fix	Float	Varied
1	All peak positions, peak widths, peak heights and Gauss % (except C-O Gauss%)	Gauss % of C-O peak	The peak position of the C-O peak
2	All peak positions, peak widths, and peak heights	Gauss % of all peaks	The peak position of the C-O peak
3	All peak positions, peak widths, and peak heights (except C-C peak)	Gauss % of all peaks, and peak height of C-C peak	The peak position of the C-O peak
4	All peak positions, and peak widths	Gauss % and peak heights of all peaks	The peak position of the C-O peak
5	All peak positions, and peak widths (except C-C)	Gauss %, peak heights of all peaks, and peak width of C-C peak	The peak position of the C-O peak
6	All peak positions	Gauss %, peak heights, and peak widths of all peaks	The peak position of the C-O peak
7	All peak positions (except shake-up peak)	Gauss %, peak heights, and peak widths of all peaks, peak position of shake-up peak	The peak position of the C-O peak
8	Nothing	All Gauss %, peak heights, all peak positions, and all peak widths	The peak position of the C-O peak

5.5.3 Air-oxidized silicon substrate

In this example we generate uniqueness plots by fitting a Si 2p narrow scan obtained from an air-oxidized silicon surface that had a final oxide thickness of 3.58 nm. We consider two different scenarios: peak fitting to two and to five peaks.[25]

Si 2p narrow scan fit to two peaks: varying the SiO_2 peak width

The Si 2p spectrum was initially fit to two peaks: one corresponding to the bulk Si material at ca. 99 eV, and the other corresponding to the SiO_2 film at ca. 103 eV (see Figure 5.6). The only constraint on this fit was that the G:L ratios of the peaks were constrained to have the same value. The optimal peak width of the SiO_2 peak (1.84 eV) was 1.35 times the width of bulk Si peak, which is in accord with literature reports. Again, the purpose of this initial fit was to obtain values about which the fit parameters could be varied in uniqueness plots. Seven uniqueness plots were generated using the approaches described in Table 5.3 that were based on varying the SiO_2 peak width from 1.00 – 2.20 eV in increments of 0.05 eV. Figure 5.7 shows the resulting uniqueness plots (χ^{2*} and Abbe criterion vs. SiO_2 peak width). Interestingly, but not surprisingly, uniqueness is obtained for all seven approaches. That is, fitting a Si 2p XPS peak envelope with two obvious components to two peak components significantly limits the chances of getting unreasonable results. For both the χ^{2*} and Abbe criterion vs. SiO_2 peak width plots, the two most constrained approaches (Approaches 1 and 2) give the most unique results (the tightest parabola-like or Gaussian-like curves, respectively), while the less constrained approaches (Approaches 3 – 7) give the broader curves.

Table 5.3 Approaches used to generate the seven uniqueness plots in Figure 5.7. The number of fixed parameters was decreased proceeding from Approach 1 to Approach 7. All fits were to the Si 2p spectrum from a sample of air-oxidized silicon (see Figure 5.6), and are based on changing the width of the SiO_2 peak.

Approach	Fix	Float	Varied
1	All peak positions, peak widths, peak heights and Gauss % (except SiO_2 peak Gaussian %)	Gauss % of SiO_2 peak	The peak width of SiO_2 peak
2	All peak positions, peak widths, and peak heights	Gauss % of all peaks	The peak width of SiO_2 peak
3	All peak positions, peak widths, and peak heights (except SiO_2 peak peak)	Gauss % of all peaks and peak height of SiO_2 peak	The peak width of SiO_2 peak
4	All peak positions, and peak widths	Gauss % and peak heights of all peaks	The peak width of SiO_2 peak
5	All peak positions, and peak widths (except Si peak)	Gauss % and peak heights of all peaks, and peak width of Si peak	The peak width of SiO_2 peak
6	Peak position of Si peak	Gauss % and peak heights of all peaks, peak width of Si peak, and peak position of SiO_2 peak	The peak width of SiO_2 peak
7	Nothing	All Gauss %, peak heights, all peak positions and peak widths	The peak width of SiO_2 peak

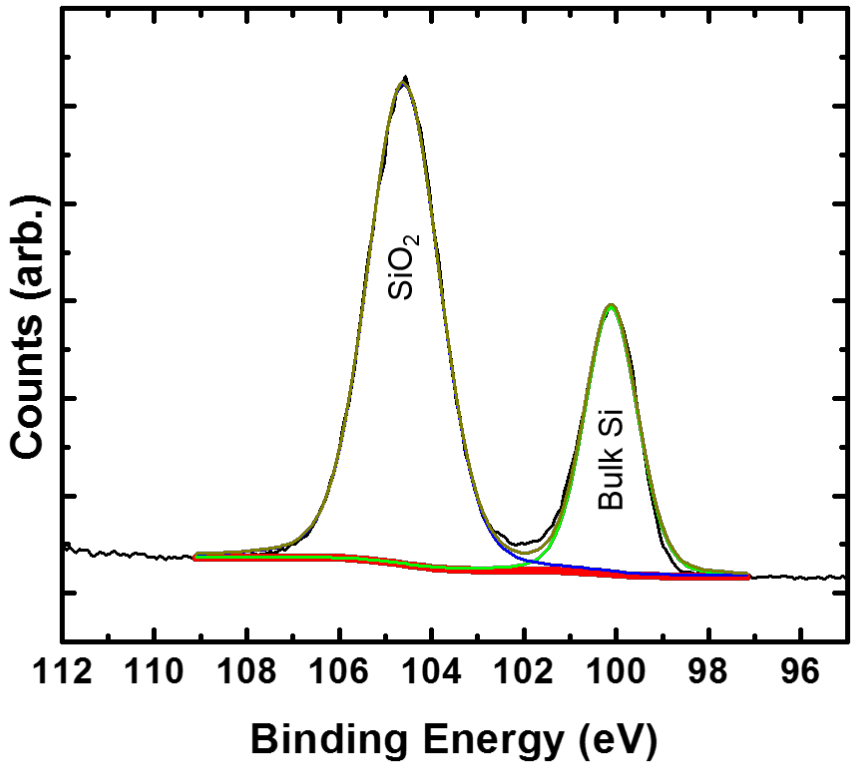


Figure 5.6 Optimal fit to an Si 2p narrow scan using two components: a bulk Si and an oxidized Si peak.

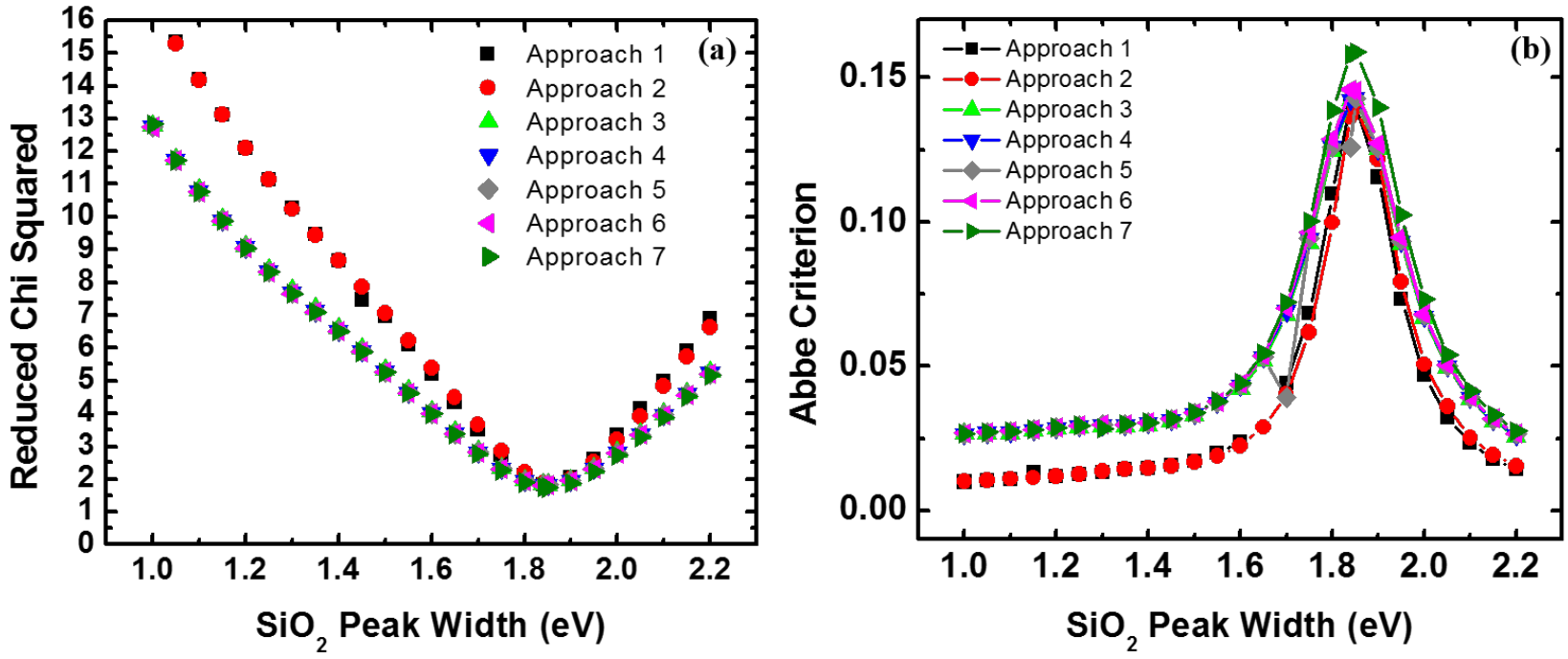


Figure 5.7 Uniqueness plots for the fits to an Si 2p narrow scan of an oxidized silicon wafer (see Figure 5.6). Plots of (a) reduced chi squared and (b) the Abbe criterion value vs. a fit parameter (the SiO_2 peak width) that was fixed to specific values.

Si 2p envelope peak fit to five peaks: varying the SiO_2 peak width

In this example, the Si 2p spectrum in Figure 5.6 was fit to five peaks (see Figure 5.8). These corresponded to a bulk Si peak, and four peaks that represented various forms of oxidized silicon, i.e., SiO_x with x ranging from 1 to 4, where x represents the average number of oxygen atoms around each silicon atom and not the bulk stoichiometry of the material.[25] In the initial, reasonable, fit to the narrow scan, the bulk Si and SiO_1 peak positions were allowed to float. The peak positions of the SiO_2, SiO_3 and SiO_4 components were constrained with respect to the SiO_1 peak position, where a 0.6 eV binding energy shift was applied for each oxygen atom attached to Si. This approach was previously taken by Alexander et al.[25] The widths and G:L ratios of the bulk Si and SiO_1 peaks were allowed to float, and the widths and G:L ratios of the SiO_2, SiO_3 and SiO_4 peaks were constrained to be the same as that of the SiO_1 peak. All the peak heights floated. From this fit, the optimal peak width of the SiO_2 peak was 1.32 eV.

Using the eight approaches described in Table 5.4, which varied from most constrained/fewest number of fit parameters (Approach 1) to least constrained/largest number of fit parameters (Approach 8), uniqueness plots were obtained based on χ^{2*} and Abbe criterion values (see Figure 5.9). In these plots, the width of the SiO_2 peak was systematically varied from 0.8 to 2.0 eV. The trends are similar to what was observed in the previous examples. Again, the uniqueness plots based on the χ^{2*} value fanned out and flattened as the number of fit parameters increased (Figure 5.9a), with the least constrained approach showing nearly a horizontal line. These changes appeared in an approximately inverted fashion for the plot of the Abbe criterion (Figure 5.9b).

Table 5.4 Approaches used to generate the eight uniqueness plots in Figure 5.9. The number of fixed parameters was decreased proceeding from Approach 1 to Approach 8. All fits were to the Si 2p spectrum from a sample of air-oxidized silicon (see Figures 5.6 and 5.8), and are based on changing the width of the SiO_2 peak.

Approach	**Fix**	**Float**	**Varied**
1	All peak positions, peak widths, peak heights and Gauss % (except SiO_2 peak Gaussian %)	Gauss % of SiO_2 peak	The peak width of SiO_2 peak
2	All peak positions, peak widths, and peak heights	Gauss % of all peaks	The peak width of SiO_2 peak
3	All peak positions, peak widths, and peak heights (except SiO_2 peak height)	Gauss % of all peaks and peak height of SiO_2 peak	The peak width of SiO_2 peak
4	All peak positions, and peak widths	Gauss % and peak heights of all peaks	The peak width of SiO_2 peak
5	All peak positions, and peak widths (except SiO_3 peak)	Gauss % and peak heights of all peaks, and peak width of SiO_3 peak	The peak width of SiO_2 peak
6	All peak positions	Gauss % and peak heights, and peak widths of all peaks	The peak width of SiO_2 peak
7	All peak positions (except SiO_2 peak position)	Gauss % and peak heights of all peaks, peak width of Si peak, and peak position of SiO_2 peak	The peak width of SiO_2 peak
8	Nothing	All Gauss %, peak heights, all peak positions and peak widths	The peak width of SiO_2 peak

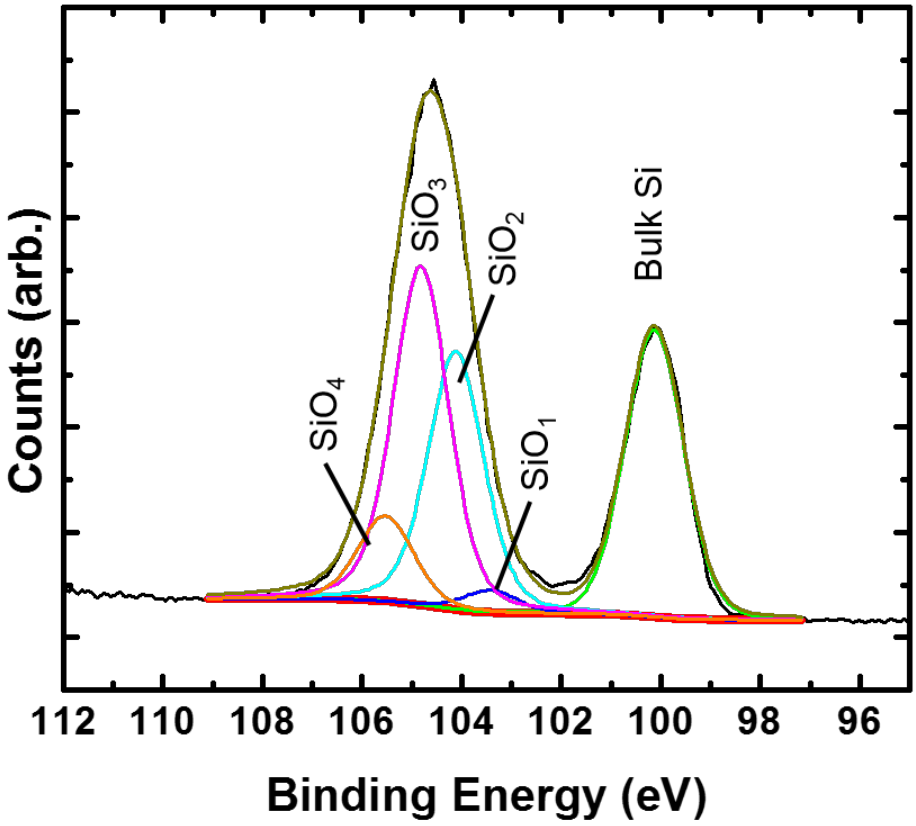


Figure 5.8 Optimal fit to an Si 2p narrow scan using five components: a bulk Si peak and four oxidized Si peaks denoted SiO_x, with x = 1 to 4.

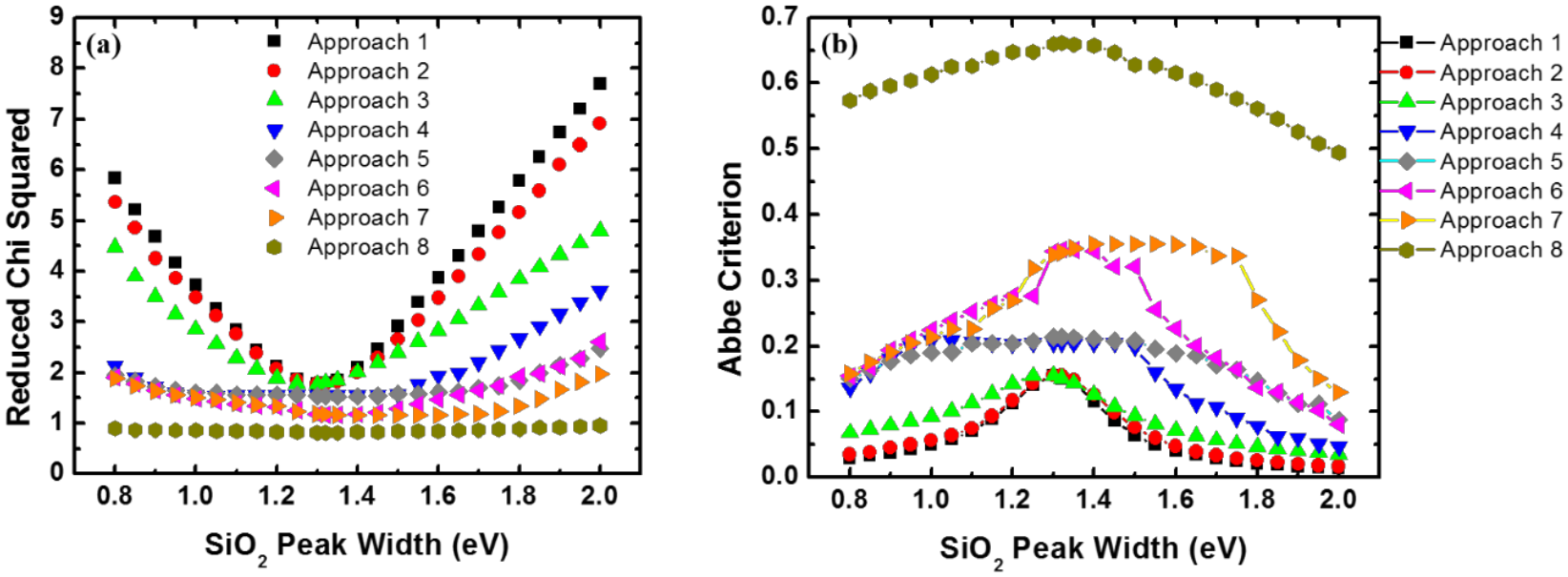


Figure 5.9 Uniqueness plots for the fits to a Si 2p narrow scan from an oxidized silicon wafer using the eight approaches described in Table 5.4. See also Figure 5.8. Plots of (a) reduced chi squared and (b) the Abbe criterion value vs. a fit parameter (the SiO_2 peak width) that was fixed to specific values.

5.6 Conclusions

XPS has become increasingly common and important over the past few years. However, its popularity appears to have attracted some novice practitioners that may lack necessary training to adequately perform peak fits. Herein, we have demonstrated the usefulness of uniqueness plots in the analysis of XPS narrow scans using diverse sets of data and fitting approaches. The usefulness of uniqueness plots is well established in the spectroscopic ellipsometry community. Uniqueness plots are simple and straightforward graphical tools that XPS practitioners can use to identify inappropriate peak fits, and also help confirm more acceptable ones. These plots show that for relatively simple narrow scans that can be fit to a few components/peaks, the chances of getting unreasonable peak fits is rather low. Of the two types of uniqueness plots discussed herein, the χ^{2*} vs. fit parameter plot seems to be the more straightforward to interpret. In these plots, the degree of curvature of the parabola-like curves is proportional to the uniqueness of the fit – a high degree of curvature implies a unique fit, and a lower degree of curvature (at its extreme, a horizontal line) implies a fit in which there is a high degree of fit parameter correlation. That is, a uniqueness plot without significant curvature, i.e., a horizontal line, is a strong indication of an inappropriate peak fit, whereas uniqueness plots *with* significant curvature are expected for good fits, but do not necessarily confirm them. Uniqueness plots based on the Abbe criterion bear an approximately inverse relationship to those based on χ^{2*}. We recommend that, in the future, XPS peak fitting software packages contain the capability of generating these plots.

5.7 Acknowledgement

We acknowledge the Department of Chemistry and Biochemistry and College of Physical and Mathematical Sciences at Brigham Young University for their support of this work.

Author Contributions: Taking leads from SE, M.R.L. suggested to explore the use of uniqueness plots in XPS. B.S. selected the samples sets to be employed in the study, finalized various fit approaches based on number of fit parameters, performed the data analysis, and made all the figures. M.R.L. and B.S. wrote the manuscript.

5.8 References

1. C. J. Powell, J. Vac. Sci. Technol., A **21** (5), S42-S53 (2003).

2. C. J. Powell, J. Surface Analysis **20** (2014).

3. B. Singh and M. R. Linford, in *Vacuum Technology & Coating* (October, 2015).

4. P. M. A. Sherwood, J. Vac. Sci. Technol., A **14** (3), 1424-1432 (1996).

5. M. R. Linford, in *Vacuum Technology & Coating* (July, 2014), pp. 27-33.

6. B. V. Crist, Journal of Surface Analysis **4** (3), 428-434 (1998).

7. B. Singh, D. Velázquez, J. Terry and M. R. Linford, J Electron Spectrosc **197** (0), 56-63 (2014).

8. B. Singh, D. Velázquez, J. Terry and M. R. Linford, J Electron Spectrosc **197** (0), 112-117 (2014).

9. K. A. Wepasnick, B. A. Smith, J. L. Bitter and D. H. Fairbrother, Anal Bioanal Chem **396** (3), 1003-1014 (2010).

10. M. Li, M. Boggs, T. P. Beebe and C. P. Huang, Carbon **46** (3), 466-475 (2008).

11. H. Wang, A. Zhou, F. Peng, H. Yu and J. Yang, J Colloid Interf Sci **316** (2), 277-283 (2007).

12. V. Gupta, H. Ganegoda, M. H. Engelhard, J. Terry and M. R. Linford, J Chem Educ **91** (2), 232-238 (2013).

13. R. Hesse, T. Chasse, P. Streubel and R. Szargan, Surf Interface Anal **36** (10), 1373-1383 (2004).

14. B. Singh, R. Hesse and M. R. Linford, in *Vacuum Technology and Coating* (December, 2015).

15. M. M. Ibrahim and N. M. Bashara, J Opt Soc Am **61** (12), 1622-1629 (1971).

16. G. H. Bu-Abbud, N. M. Bashara and J. A. Woollam, Thin Solid Films **138** (1), 27-41 (1986).

17. J. N. Hilfiker, N. Singh, T. Tiwald, D. Convey, S. M. Smith, J. H. Baker and H. G. Tompkins, Thin Solid Films **516** (22), 7979-7989 (2008).

18. G. Droulers, A. Beaumont, J. Beauvais and D. Drouin, J Vac Sci Technol B **29** (2), 021010 (2011).

19. J. M. Zhao and P. Yang, Microsystem Technologies **18** (9-10), 1455-1461 (2012).

20. C. M. Herzinger, B. Johs, W. A. McGahan, J. A. Woollam and W. Paulson, J Appl Phys **83** (6), 3323-3336 (1998).

21. V. A. Yakovlev and E. A. Irene, J Electrochem Soc **139** (5), 1450-1455 (1992).

22. D. S. Jensen, S. S. Kanyal, N. Madaan, A. J. Miles, R. C. Davis, R. Vanfleet, M. A. Vail, A. E. Dadson and M. R. Linford, J Vac Sci Technol B **31** (3) (2013).

23. D. S. Jensen, S. S. Kanyal, N. Madaan, M. A. Vail, A. E. Dadson, M. H. Engelhard and M. R. Linford, Surface Science Spectra **20** (1), 36-42 (2013).

24. D. A. Shirley, Phys Rev B **5** (12), 4709-4714 (1972).

25. M. R. Alexander, R. D. Short, F. R. Jones, W. Michaeli and C. J. Blomfield, Appl Surf Sci **137** (1-4), 179-183 (1999).

Chapter 6: Information Content / Entropy as a Statistical / Chemometrics Tool for Analyzing Complex Data Sets, with Application to XPS and ToF-SIMS Depth Profiles

6.1 Abstract

Advanced scientific research often involves the measurement and analysis of multiple variables related to a system. This information may include complex spectra, often from different instruments. Various multivariate analysis tools, for example principal components analysis (PCA) and cluster analysis are frequently employed to interpret and compare complex data sets. While indispensable for advanced statistical analysis, these multivariate tools are somewhat arcane, and difficult for a naive user to understand. This leaves room for exploration of alternate statistical tools based on different algorithms. Information theory was first put forward by Shannon Claude in 1948, which revolutionized the digital world. Herein, we employ the information theory to calculate the information content (IC) of complex data sets by treating them as probability distributions. IC analysis is undertaken on complex, yet interesting depth profiling data collected via X-ray photoelectron spectroscopy (XPS) and time-of-flight secondary ion mass spectrometry (ToF-SIMS). It helps to differentiate the information rich interfacial layer, which is composed of signals from the superficial and substrate layer. PCA and cluster analysis was also performed on the data sets and results compared to that from IC analysis. It was observed that although PCA and cluster analysis tend to show similar groupings/separations of various data sets, it was the IC analysis that gave the most straightforward interpretation. IC analysis provides a way to quantify information in complex spectra in units of bits. Another advantage of IC analysis is that we could compare data from various different techniques as IC is in the same units. This would help to evaluate materials in multi-dimensions.

6.2 Introduction

Two of the most popular and important surface analytical techniques, which are central to the chemical analysis of materials, are X-ray photoelectron spectroscopy (XPS)[1, 2] and time-of-flight secondary ion mass spectrometry (ToF-SIMS).[3, 4] XPS is currently receiving in excess of 10,000 citations in the literature per annum,[1] and for complex, patterned materials, static and dynamic ToF-SIMS is unsurpassed in its ability to provide near surface and bulk information, respectively.[5-8] Both, XPS and ToF-SIMS can provide significant amounts of information about surfaces, where this amount of information is increased substantially when depth profiling is performed.

In a traditional analysis of an XPS or ToF-SIMS depth profile, one might plot the signal from one element or perhaps the ratio of two signals as a function of sputtering time (depth into a material) to determine the compositions of layers. For example, in a depth profile through a thin film of SiO_2 on Si, one might plot the oxygen signal from the corresponding O 1s XPS narrow scans, or perhaps the O/Si ratios determined from the O 1s and Si 2p narrow scans, as a function of sputter time. Certainly, at some level, this is not unreasonable, and useful insights into a material can be gained in this way. However, this approach will typically miss important chemical information about a material. For example, it overlooks any evolution in the chemical (oxidation) states of the probed materials as a function of depth.[9] In the case of ToF-SIMS, simply focusing on a peak or two in the spectra in a depth profile omits the vast majority of the information collected.

Various multivariate, statistical analytical tools are frequently used to simplify and understand these complex data sets. Two of the most widely used are principal components analysis (PCA) and cluster analysis.[10, 11] PCA is perhaps the most widely used

chemometrics/bioinformatics tool.[12-14] It has been extensively used for comparing spectra. In this study, we discuss the important issue of the analysis of complex data, using examples of XPS and ToF-SIMS depth profiles. Apart from PCA and cluster analysis, we explore a new data analysis tool. It is the information content (IC), or entropy, of a spectrum. This approach comes out of the groundbreaking work of Shannon,[15] who applied the statistical thermodynamics definition of entropy to signal/communication theory. Here we take Shannon's approach back into the physical sciences, calculating the IC of spectra and using the resulting value to determine where changes are taking place in a series of them. Chapter 1 (section 1.3.2) contains a brief introduction to PCA and cluster analysis.

Information theory was powerfully influenced by Claude Shannon, who developed a mathematical description for communication through the application of the entropy function from statistical thermodynamics.[15] Based on his proposal, the information in a signal is quantified and referred to as *entropy*, or equivalently as *information content (IC)*, which is a measure of the ultimate (data) compression a signal can undergo. Over the past six decades, researchers have appreciated the potential and power of his theory, applying it to data storage, data compression, digital communication, and more specifically for MP3s, JPEGs, mobile phones, optical communications, satellite communication, space exploration programs, plagiarism detection, pattern recognition/detection, etc.[16-21] Shannon's theory profoundly influenced our world, as evidenced by the fact that his initial paper has been cited nearly 75,000 times.

In statistical thermodynamics, the entropy of mixing for an ideal solution is given by:

$$\Delta S_{mix} = -R \sum_{k=1}^{c} X_k \ln X_k \tag{6.1}$$

where R is the gas constant, and the X_k values are the mole fractions of the components to be mixed. Alternatively, the Gibbs formula for entropy is:

$$\Delta S = -k_B \sum_{i} p_i \ln(p_i) \tag{6.2}$$

Where k_B is Boltzman's constant, and p_i is the probability that a microstate will be sampled. Taking his lead from these types of equations, Shannon proposed the following formula to describe the entropy of a communication signal:

$$H(x_i) = -\sum_{i=1}^{n} p(x_i) * log_2 p(x_i) \tag{6.3}$$

where the signal is considered to be a probability distribution with $p(x_i)$ the probability of observing a signal at a certain channel, and the sum of the probabilities for a signal summing to unity. Equation 6.3 uses the base 2 logarithm, and, accordingly, the final result is given in units of 'bits'.

A classic example of Shannon's formula comes from the toss of a fair coin, which, of course, has a probability, *p*, of ½ for occurrences of either heads or tails. Inserting these probabilities into Shannon's formula gives:

$$H_{fair\ coin} = -\left(\frac{1}{2} log_2 \frac{1}{2} + \frac{1}{2} log_2 \frac{1}{2}\right) = 1\ bit \tag{6.4}$$

However, in the case of an unfair coin that would always give either heads or tails, the p_i values are 1 and 0, which, in Shannon's formula, give:

$$H_{unfair\ coin} = -(0log_2 0 + 1log_2 1) = 0\ bit \qquad (6.5)$$

where we have taken: 0 log 0 = 0, which is true in the limit of $p_i \rightarrow 0$. The entropies (information contents) for all such possibilities (through Equation 6.3), ranging from a completely fair coin to an entirely unfair one, yield the classic plot shown in Figure 6.1. Clearly, the entropy (uncertainty) in the measurement is greatest for a fair coin with $p_{heads} = p_{tails} = ½$, and it goes to zero for an entirely unfair coin for which the outcome of the coin toss is certain. Alternatively, we can interpret the entropies of the coin tosses in Equation 6.3 as saying that it requires one bit of information (a zero or a one) to represent the outcome of a fair coin toss (Equation 6.4), while it does not require any bit of information to represent the outcome of the entirely unfair coin for which the outcome is always known (Equation 6.5).

In this work we take Shannon's formula and bring it back into science as a tool for data analysis. In particular, we consider a spectrum or a data set to be a probability distribution that can be inserted into Equation 6.3. We have chosen to use 'IC' (information content) to refer to the results from Equation 6.3, and not H, which can be confused with enthalpy. We provide multiple examples of the use of the IC function for data analysis. In particular, it seems well suited for analyzing depth profiles in surface and interface analysis, as well as for clustering/analyzing samples from data taken from different instruments. For these IC analyses, spectra, such as those from X-ray photoelectron spectroscopy (XPS) and time-of-flight secondary ion mass spectrometry (ToF-SIMS), are first normalized by dividing each energy or m/z value by the sum of all the values in the corresponding spectrum. The resulting

'probabilities' are then entered into Equation 6.3. This procedure also works well with selected peaks from spectra. As is evident from Equation 6.3, a larger entropy value typically points to a more complex data set, i.e., one that has a larger number of peaks of comparable size, while a spectrum that may be strongly dominated by one peak, such that p_i approaches 1, will have a lower IC. Obviously, Shannon's use of the base 2 logarithm in Equation 6.3 is not necessary here – other bases should work equally well for differentiating spectra. However, for consistency with information theory, we will use the base 2 logarithm here, i.e., the units of all entropy (IC) calculations made herein will be 'bits'.

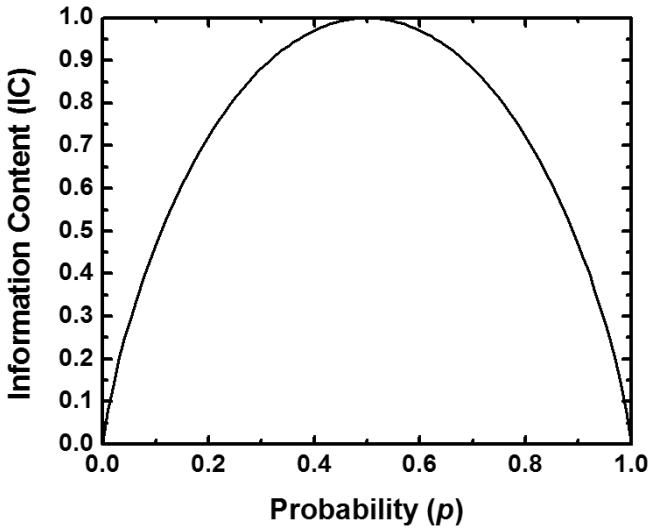


Figure 6.1 Information content for probabilities from coins that will never yield a desired outcome (p = 0), to a fair coin that gives even odds of heads or tails (p = ½), to an entirely unfair coin that has complete certainty associated with the outcome (p = 1).

6.3 Experimental

The following data sets are employed in the study:

6.3.1 XPS depth profiling of Si/SiO_2 and Ta/Ta_2O_5 samples

XPS measurements were performed with a Physical Electronics Quantera scanning X-ray microprobe instrument. This system uses a focused monochromatic Al Kα X-ray (1486.7 eV) source for excitation and a spherical section analyzer. The instrument has a 32 element multichannel detection system. A 70 W, 200 μm X-ray spot was used for the sputter depth profiles. The X-ray beam is incident normal to the sample and the photoelectron detector is at 45° off-normal. High energy resolution spectra were collected using a pass-energy of 69.0 eV with a step size of 0.125 eV. For the Ag $3d_{5/2}$ line, these conditions produced a FWHM of 0.91 eV. The sample experienced variable degrees of charging. Low energy electrons at ca. 1 eV, 20 μA, and low energy Ar^+ ions were used to minimize this charging. Ar^+ ions at 2kV were used for the depth profiles with an ion gun incident angle of 45°, polar angle of 60°, and azimuthal angle of 90°. A total of 54 and 37 sputter cycles were performed on the Si/SiO_2 and Ta/Ta_2O_5 samples respectively.

6.3.2 ToF-SIMS depth profiling

Two ToF-SIMS depth profiling samples were employed for the IC analysis: (i) ca. 100 nm of C_3F_6 and (ii) ca. 100 nm of PNIPAM (poly(N-isopropylacrylamide)) deposited on Si substrates. Samples were prepared by plasma deposition of C_3F_6 and poly(N-isopropylacrylamide) (PNIPAM) onto clean 1 cm x 1 cm shards of silicon. Silicon wafers were cleaned by soaking in DI water overnight, followed by sonicating twice in acetone, methanol, and dichloromethane. Plasma deposition was carried out in a custom built RF plasma system.

Deposition of C_3F_6 was at 150 mTorr at 5 W for 20 min, and deposition of PNIPAM was at 140 mTorr at 5 W for 15 min. ToF-SIMS depth profiling was carried out in non-interlaced mode using a ToF-SIMS V instrument (ION-TOF GmbH, Münster, Germany) with 25 kV Bi^{3+} as the analytical beam and 20 keV C_{60}^{++} as the sputter beam. Spectra were acquired from a 100 μm x 100 μm area in the center of the sputter crater using a Bi^{3+} current of 0.09 pA, with a total dose per analysis cycle of 1 x 10^{11} ions/cm^2. Sputtering was done over a 500 μm x 500 μm area with a current of 0.62 nA, with a dose of 7.7 x 10^{12} ions/cm^2 per sputter cycle. A total of 24 spectra were collected during the depth profiling study. Charge neutralization was carried out using an electron flood gun that was pulsed after each analysis/sputter cycle.

6.3.3 Nanodiamond ToF-SIMS and ICP characterization

Five nanodiamond samples: unwashed, doubly, and triply washed Advanced Abrasives (AA) (50 nm), International Technology Center (ITC) (50 nm), and Adamas Nanotechnologies (5 nm) were characterized via ToF-SIMS and ICP.[22] The acquisition parameters for the instruments are the same as those used in Chapter 7.

6.3.4 Chemometrics analyses

Principal components analysis (PCA) and cluster analysis of the ToF-SIMS and XPS data were performed using the PLS Toolbox, version 7.9.3 (Eigenvector Research, Inc., Wenatchee, WA, USA) in MATLAB, version R2014b (Natick, MA). The raw peak areas were organized row-wise in a data matrix. The data were preprocessed by mean centering, a column operation. As the data were collected using the same instrument settings during a depth profile, there was no need for normalization. Also, the data were not autoscaled, as this would give equal weights to the noisy regions of the spectra and the peaks of interest. For the IC analyses, the data were

treated as a probability distribution and arranged column-wise in a matrix. The probability of occurrence of signal (count) at each m/z (for ToF-SIMS), and binding energy (eV for XPS) was calculated as the respective signal divided by the sum of all the signals in the respective column. Using these probabilities, the IC of each spectrum was calculated using Equation 6.3.

6.4 Results and Discussion

6.4.1 Understanding the Entropy Function: ICs of Fabricated Spectra

To illustrate the properties of the IC function, Equation 6.3, it was applied to a series of mock spectra. Figure 6.2a shows the ICs of three flat baselines. Here, these functions (horizontal lines) have constant values. Accordingly, the probability of any point occurring must be the same, i.e., it is 1/N, where N is the number of data points, and the sum of the probabilities is 1. (The bottom line in Figure 6.2a at $y = 0$ is, admittedly, a somewhat artificial scenario. If the probabilities here are zero everywhere, then IC = 0. However, technically speaking, Equation 6.3 should not be applied to this, or any, series of zeros because the sum of these probabilities is not unity – there is no probability distribution here. In general, real measurements show signals and/or have noise/uncertainty.)

$$IC = -\sum_{i=1}^{N} \frac{1}{N} * log_2 \frac{1}{N} = 9.9658 \ (for\ N = 1000\ ) \tag{6.6}$$

Figure 6.2b shows the IC values for three functions that contain one, two, or three spikes of equal amplitude. Normalization gives the first function a probability of 1 for the spike, and probabilities of zero everywhere else. These values yield $IC_i = 0$, which is analogous to the situation for the entirely unfair coin considered above. For the second spectrum, normalization

gives two spikes of equal magnitude, i.e., equal probability of ½, which gives $IC_{ii} = 1$. This situation is analogous to that of the fair coin toss. For the three spikes of equal probability in Figure 6.2b, we obtain, as expected, a higher information content of 1.585:

$$IC = -\sum_{i=1}^{N} \frac{1}{N} * log_2 \frac{1}{N} = 1.585 \ (for\ N = \ 3) \tag{6.7}$$

Thus, increasing the number of features in a spectrum generally increases its IC value. Figure 6.2c shows two noisy baselines, which differ only by a multiplicative constant. When normalized, they have the same values (probabilities), therefore they have the same IC values. Figure 6.2d shows three functions. The first (i) is a rather narrow peak on a flat baseline. This spectrum has a relatively low IC value. The second function (ii) is broader, and, as expected, it has a larger IC value. The third function is the same as a second, except that it has a shoulder at its left side. As expected, this shoulder increases the IC value of the function. Figure 6.2e (i) shows a peak with a relatively low amount of noise, and its IC value. The peak above it is identical to the first, with the exception that it is even noisier. Accordingly, this spectrum has a higher IC value. Finally, a third spectrum (Figure 6.2e (iii)) was created by taking every third data point in the second spectrum (Figure 6.2e (ii)). As expected, this spectrum shows a lower IC value. That is, in general (i) noisier spectra have higher IC values (this is admittedly counterintuitive – as analytical chemists we think of an increase in noise as a loss of information), and (ii) there is, in general, an increase in IC with an increasing number of data points (this follows from Equation 6.3). Finally, Figure 6.2f shows two spectra and their IC values. As expected, the spectrum represented by the dashed line shows broader peaks and has a somewhat higher IC value. This figure also indicates that the IC value of a combined spectrum

containing the peaks from both spectra would be higher than the IC values of the individual spectra. This final panel segues into our discussion of the IC function in depth profiling.

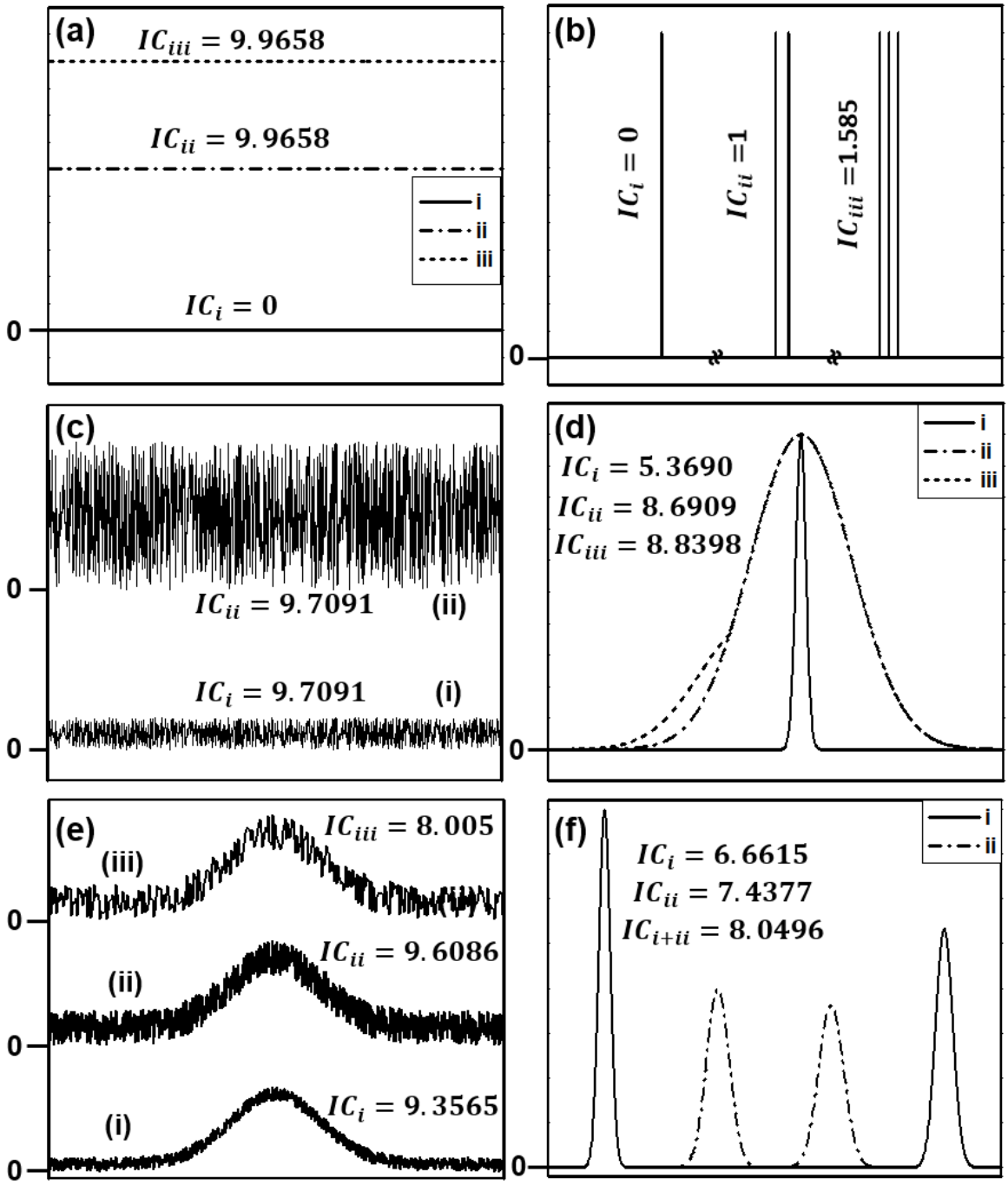


Figure 6.2 Information content tutorial. Understanding IC values and trend using mock spectra.

6.4.2 Applications of the IC Function to Depth Profiling

Depth profiling is an important method for understanding solid materials. For example, it is extensively used to determine the profiles of dopants in semiconductors,[23-26] which is critical information for the integrated circuit industry. In a typical depth profile, a surface is first analyzed by a surface sensitive technique, such as XPS or ToF-SIMS. A small amount of the material is then removed with an ion beam, which can be applied in such a way that it sputters/ablates the material in a rather homogeneous fashion. The surface sensitive analysis is then repeated, which is followed by removal of another small amount of material from the surface. This process can be continued, usually in an automated fashion. If the surface (really the bottom of the crater) is imaged in two dimensions after each ablation, a three-dimensional map of the material can be created. Alternatively, and very commonly, a single measurement of the material is made at each depth. Figure 6.2f is reminiscent of a common situation one encounters in depth profiling through layered materials. Here, to a first approximation, one might expect a series of nearly identical spectra from a film followed by a different series of nearly identical spectra from the next layer or substrate. In general, one would expect the IC values of these two sets of spectra to differ, e.g., XPS peaks from oxidized metals are generally wider than those from reduced metals.[27] A plot of IC value vs. spectrum number for such a depth profile would then indicate the position of the interface between the film and the substrate at the change in the IC value. However, interfaces are rarely perfectly smooth, and the act of depth profiling (sputtering) typically roughens/broadens interfaces. Accordingly, an interface will often show spectral features from both of the materials that surround it (see Figure 6.2f (i + ii)), and therefore a higher IC value than those of the materials around it. Thus, an increase in IC value

can indicate the presence of the interface. We will encounter this scenario in the examples that follow, as well as a few others.

Example 1. Depth profiling through SiO_2 on Si by XPS: Si 2p and O 1s narrow scans

Figure 6.3a shows a series of XPS Si 2p narrow scans from a depth profile of an SiO_2 film on an Si substrate. The first two scans appear to be a little different from those that follow them. It is not unusual for the first scan or two in a depth profile to be different from subsequent scans because of surface contamination or other differences between the surface and bulk of a material. These first two scans are followed by a series of very similar scans that, based on their peak position, clearly correspond to SiO_2.[28] The interface between the SiO_2 and bulk Si is then reached. As suggested above, this region appears to show signals due to both SiO_2 and Si. This interface region is then followed by a series of uniform signals from the Si substrate.

A PCA analysis of the data in Figure 6.3a was performed. And while almost all of the variation in the data was captured by PC1 (ca. 99%), it became apparent after examining the plot of the scores on PC1 and PC2 (Figure 6.3b) that this second PC would also be helpful in explaining the data. Accordingly, Figure 6.3b shows the first two points (spectra) clustered together (they would not form a unique cluster if only the scores on PC1 were considered), a cluster of the next spectra (3 – 20) that correspond to the SiO_2 overlayer, a series of separated points that correspond to spectra 21-29, which seem to follow a trajectory through the plot and come from the interface between the two materials, and a cluster of points (spectra 30 – 54) that correspond to bulk Si. A cluster analysis was then performed on the data, which showed the same general features as the PCA: two fairly large clusters for the bulk SiO_2 and Si spectra that the first two spectra are different from the subsequent spectra, and a series of spectra that do not

fit into either cluster that correspond to the interface region. As an additional guide to these analyses, as well as those discussed below, the individual Si 2p spectra in Figure 6.3a are plotted in Figure 6.4.

Figures 6.3d shows the plot of the IC values of the spectra in Figure 6.3a vs. spectrum number. This plot is dominated by a peak that corresponds to the interface region between the SiO_2 and Si that has flat regions to the left and right of it with slightly different IC values that correspond to the SiO_2 film and Si substrate, respectively. As expected, the first two narrow scans have different IC values than those that follow. And while it is certainly true that PCA and cluster analyses allow an interpretation of the data that is consistent with the depth profile of SiO_2 on Si, it is, at least in our opinion, the IC analysis that shows this most clearly. Indeed, in this work we consistently found it easier to interpret depth profile results with IC plots, and to then apply those results to PCA and cluster analyses. Finally, Figure 6.3e shows the finite difference of each point in Figure 6.3d. The finite difference is closely related to the derivative and will be helpful in revealing trends in some of the IC plots below. Note that, technically speaking, the entity plotted in the IC plots shown herein is the backward difference: $\nabla f(x) = f(x) - f(x-h)$, i.e., $\nabla f(x) = IC_n - IC_{n-1}$ or the value of the current point minus the value of the previous point, and not the finite difference: $\Delta f(x) = f(x+h) - f(x)$, i.e., $\Delta f(x) = IC_{n+1} - IC_n$, which is the value of the subsequent point minus the value of the current point. Ultimately, the backward difference was used here because the peaks in the resulting plots appeared to better correlate with the transitions in the IC data. Nevertheless, we expect that any reasonable difference or derivative-like function, including numerical derivatives that involve more than two data points and that smooth the data, will also work equally well or better in some circumstances.[29]

Figure 6.5a shows a plot of the O 1s narrow scans from the same Si/SiO_2 depth profile considered in Figures 6.3 and 6.4. These spectra come sequentially from the outer surface of the material, the SiO_2 film, the interface region, and the bulk Si. Beyond the interface, the O 1s signal has disappeared – there is no oxide in the bulk and the spectra contain only noise. By considering two PCs, PCA (see Figure 6.5b) again shows a progression through the outer surface region of the material, the SiO_2, the interface region, and finally the bulk. The cluster analysis can be similarly interpreted. However, the IC and finite difference IC analyses again appear to provide a more straightforward description of the process. The IC analysis (Figure 6.5d) shows a small difference between the first two data points and those that follow, a steady increase in the IC values of the next set of points from about 5.7 – 5.9, an abrupt jump from ca. IC = 5.9 – 6.8, which occurs at the Si/SiO_2 interface, and a plateau at high IC that corresponds to a lack of signal (just noise) from the O 1s spectra of the bulk Si. The IC and finite difference IC plots suggest that some change is taking place in the oxygen signals over the course of the depth profile through the SiO_2. This is consistent with a steady decrease in the O 1s/Si 2p area ratio in this depth profile (see Figure 6.6). We notice that there is no peak in this spectrum that signals the presence of the interface, which is a result of the noisy O 1s narrow scans from the bulk Si representing a maximum in the IC for this system (vide supra). Accordingly, the finite difference of the IC plot is even more helpful in identifying the transition regions in this material. It clearly identifies the air-SiO_2 and SiO_2-Si interfaces. Overall, the same conclusions can be drawn here. With some skill/practice in chemometrics and manipulation of the raw data, the PCA and cluster analyses reveal the expected trends in the spectra. However, they appear to be more obvious in the IC analyses (Figures 6.5d – e).

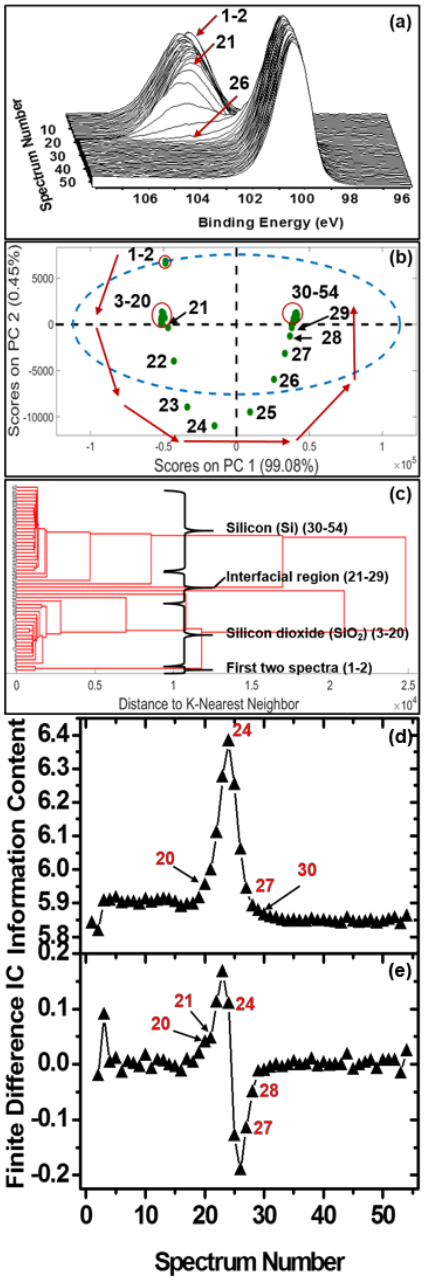


Figure 6.3 XPS depth profiling study of SiO_2/Si substrate: (a) 3-D graph of all 54 Si 2p spectra, (b) scores plot of PCA analysis on the data, (c) cluster analysis on the data set, and (d) IC values and (e) finite difference IC as a function of spectrum number.

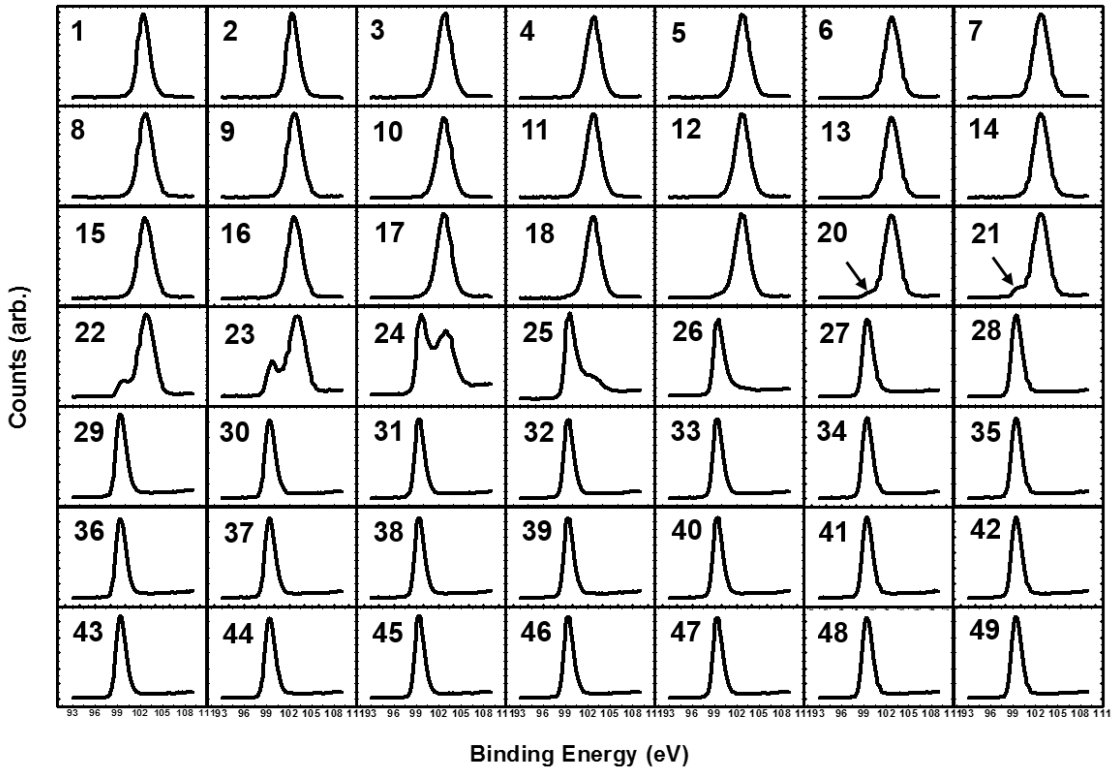


Figure 6.4 Si 2p spectra collected during the XPS depth profiling of SiO_2 on an Si substrate. Spectra 50-54 are very similar to spectra 43-49 and, hence, not included here.

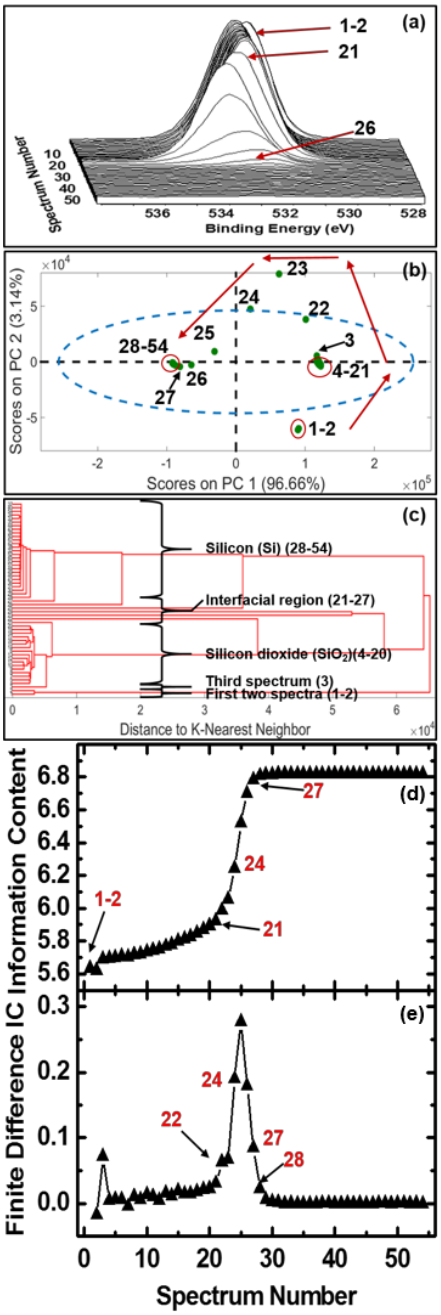


Figure 6.5 XPS depth profiling study of SiO_2/Si substrate: (a) 3-D graph of all 54 O 1s spectra, (b) scores plot of PCA analysis on the data, (c) cluster analysis on the data set, and (d) IC values and (e) finite difference IC as a function of spectrum number.

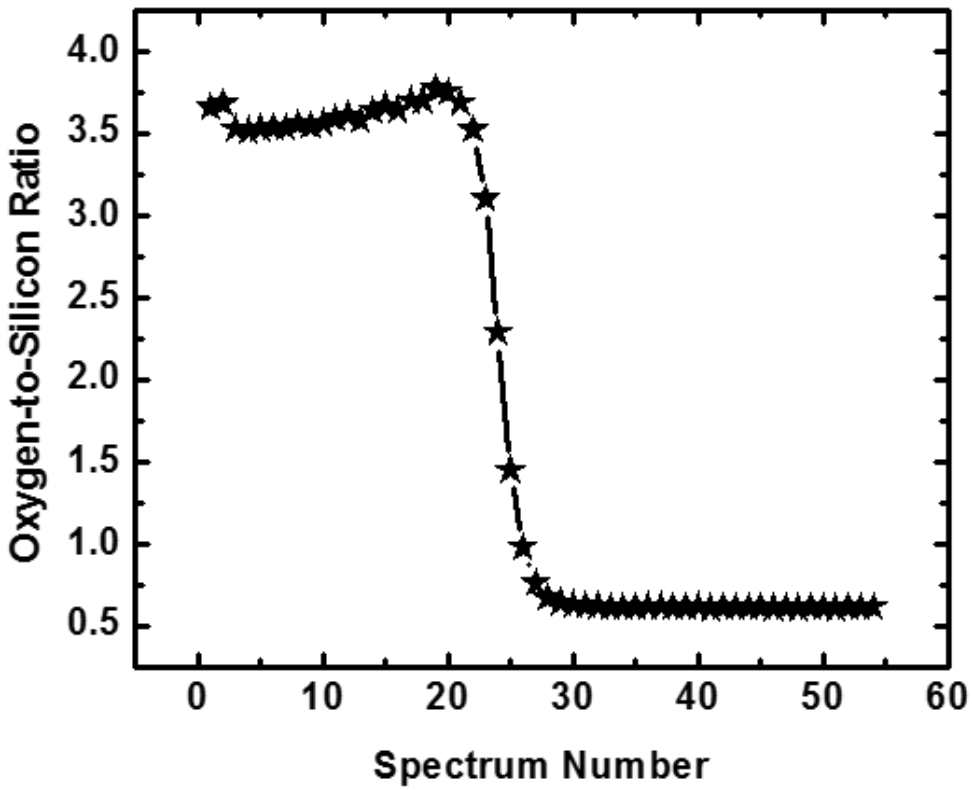


Figure 6.6 Oxygen-to-silicon ratio as a function of spectrum number during the XPS depth profiling study of SiO_2/Si substrate.

Example 2. Depth profiling through Ta_2O_5 on Ta by XPS: Ta 4f and O 1s Narrow Scans

Figure 6.7 shows a depth profile through a film of Ta_2O_5 on Ta. The series of spectra in Figure 6.7a show a known phenomenon that is not observed in Figure 6.3, which is that the Ar^+ sputtering of this oxide reduces it. Accordingly, there is a substantial broadening of the oxide peaks towards lower binding energy. The first two spectra in this series show the oxide most prominently. They are then followed by spectra with broad signals that show both oxidized and metallic components. We then come to the interface region, and finally the signals from the Ta substrate. Once again, with some work and interpretation, the PCA and cluster analysis provide interpretations that are consistent with these results. However, the IC analysis again appears to provide this information in a more straightforward way. The first two scans, and to a small degree the third one, have lower IC values than those of the oxide that follow, which show high IC values because of their breadth. There is then a decrease in the IC value at the interface region, which is followed by the lower IC values for the bulk Ta. The combination of the changes in the IC values (Figure 6.7d) and the peaks in the finite difference IC plot (Figure 6.7e) clearly indicate the presence of two interfaces in this depth profile.

Figure 6.8 shows the O 1s narrow scans that correspond to the Ta 4f spectra in Figure 6.7. The PCA and cluster analyses of these narrow scans are (again) somewhat complex to interpret, although arguably consistent with a depth profiling through the material. Once again, the IC plot in Figure 6.8d and finite difference IC plot in Figure 6.8e are easier to interpret. They suggest two main transitions – one at the air-Ta_2O_5 interface and the other at the Ta_2O_5-Ta interface. As was the case for the O 1s signals in Figure 6.5, the O 1s signals that show only noise have the highest IC values. In addition, as observed in Figures 6.5 and 6.7, the IC values of the film

change somewhat over the course of the depth profile, indicating that some change in the material is taking place.

Example 3. Depth profiling through a plasma polymerized film of C_3F_6 on an Si substrate by ToF-SIMS

A film of C_3F_6 was deposited onto a silicon substrate by plasma polymerization, and then depth profiled by ToF-SIMS. Figure 6.9a shows the resulting positive ion spectra from m/z 0 to 150. These spectra show a considerable amount of complexity through about scan 20, after which they become less complex. PCA and cluster analysis of these spectra suggest that the first spectrum is different from the subsequent ones, that spectra 2 – 16 cluster (with subclusters containing spectra 2 – 9 and 10 – 16), and that there is then, perhaps, a transition region followed by spectra from the substrate. The IC plots again appear to be easier to interpret. There is a transition at the beginning of the depth profile, a region of nearly constant IC values that corresponds to the polymer, and finally a transition region in which the IC value initially increases and then drops towards the IC value of the substrate. The lower IC value of the substrate is expected because of the decreased complexity of its spectrum. Note that (i) the increased IC value at the polymer-substrate interface suggests peaks from both entities in the spectra, and (ii) the IC analysis did not pick up the clustering of the spectra from the polymer, i.e., employing more than one chemometrics tool is often important for thoroughly understanding a data set. This problem was also approached by considering a set of 19 selected peaks from the spectra, instead of the entire spectra. The results obtained were very similar to those in Figure 6.9d-e and are shown in Figure 6.10a-b.

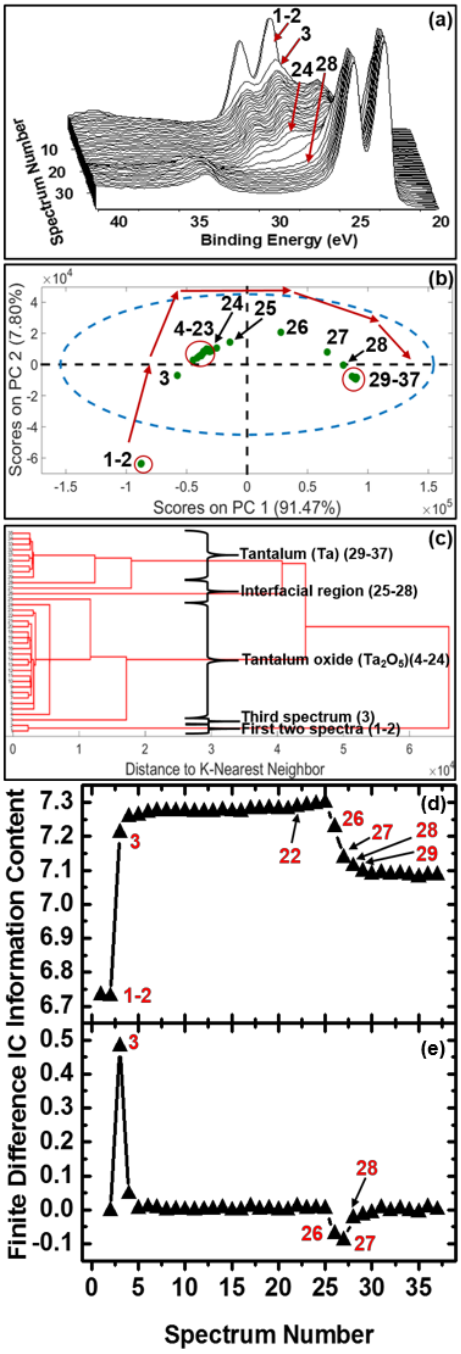


Figure 6.7 XPS depth profiling study of Ta_2O_5/Ta substrate: (a) 3-D graph of all 37 Ta 4f spectra, (b) scores plot of PCA analysis on the data, (c) cluster analysis on the data set, and (d) IC values and (e) finite difference IC as a function of spectrum number.

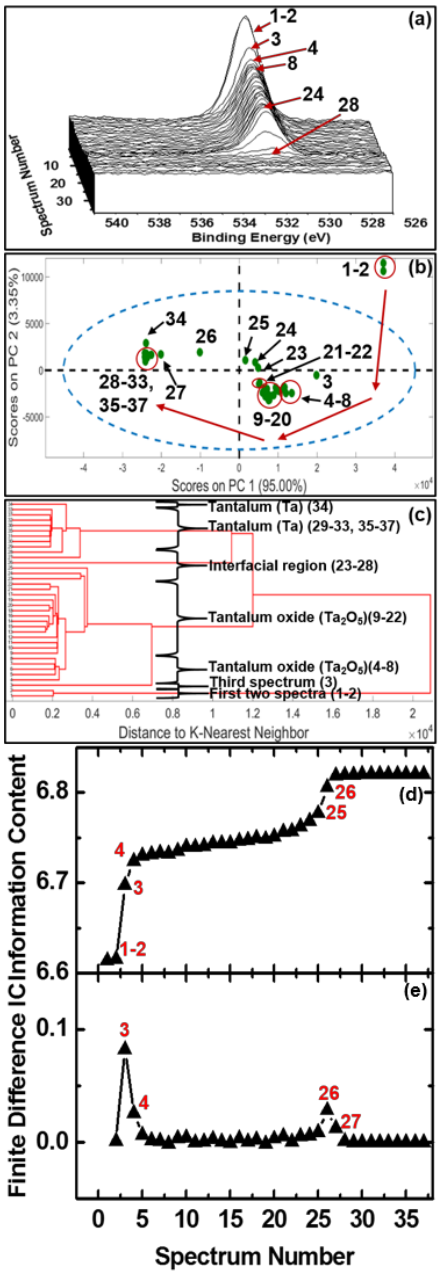


Figure 6.8 XPS depth profiling study of Ta_2O_5/Ta substrate: (a) 3-D graph of all 37 O 1s spectra, (b) scores plot of PCA analysis on the data, (c) cluster analysis on the data set, and (d) IC values and (e) finite difference IC as a function of spectrum number.

Example 4. Depth profiling through a plasma polymerized film of PNIPAM *on an Si substrate by ToF-SIMS*

Figure 6.11 describes a ToF-SIMS depth profile through a plasma polymerized film of PNIPAM on silicon. Here, 19 peaks were selected from the spectra and plotted in Figure 6.11a. The qualitative view of the selected peaks is that they contain a fair amount of complexity at the beginning of the depth profile, and less towards the end of it. A PCA analysis of these spectra is rather challenging to interpret – it does not show clear clustering/transitions that would indicate a depth profile. The cluster analysis appears to give results that are more consistent with a depth profile, although the resulting dendrogram remains fairly complex. In contrast, the IC analysis again seems to more effectively identify the key regions of the depth profile. In particular, it shows (i) a small difference between scans 1 and 2 and those that follow it (these differences are more apparent in the PCA and cluster analyses), (ii) a flat region that corresponds to the polymer, (iii) a rising transition region beginning around scan 13 that identifies the polymer-silicon interface, and (iv) a decrease in the IC values as they approach those of the substrate. Changes in peak intensity are consistent with these chemical changes. For example, peaks at m/z 44.05 and 44.98 are attributable to the polymer layer and substrate layer ($SiOH^+$), respectively, and in spectra 14 – 17, the intensity of the m/z 44.05 peak, which has a mass excess and should be organic, decreases while the intensity of the m/z 44.98 peak, which has a mass deficit and should be inorganic, increases. The transition region in Figure 6.11d is also clearly present in the finite difference IC plot. Finally, this analysis was repeated with the entire spectra from m/z 0 – 150 (not just selected peaks). The results were similar to those shown in Figures 6.11d-e, and are shown in Figures 6.12a-b.

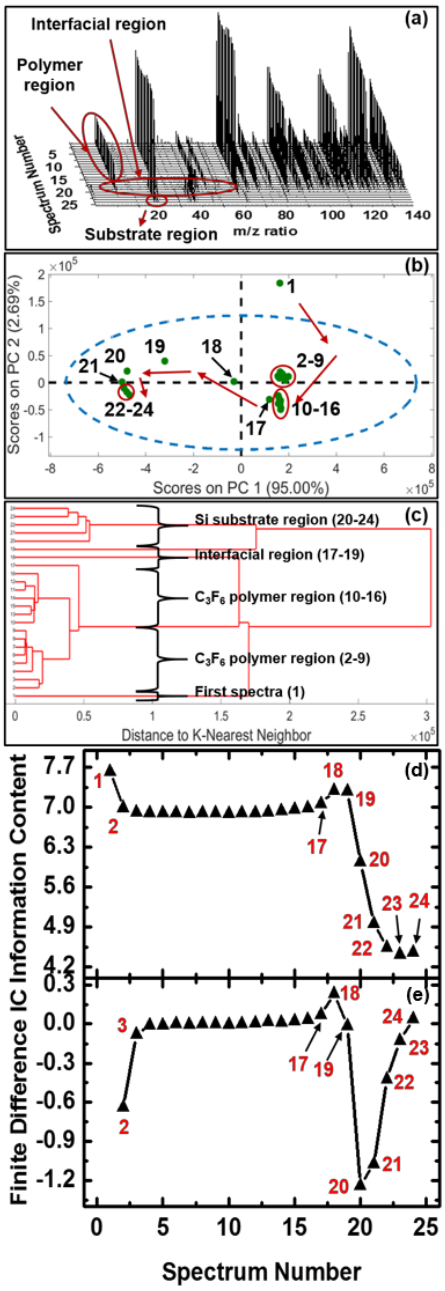


Figure 6.9 ToF-SIMS depth profiling study of C_3F_6/Si substrate: (a) 3-D graph of all 24 ToF-SIMS positive ion spectra, (b) scores plot of PCA analysis on the data, (c) cluster analysis on the data set, and (d) IC values and (e) finite difference IC as a function of spectrum number.

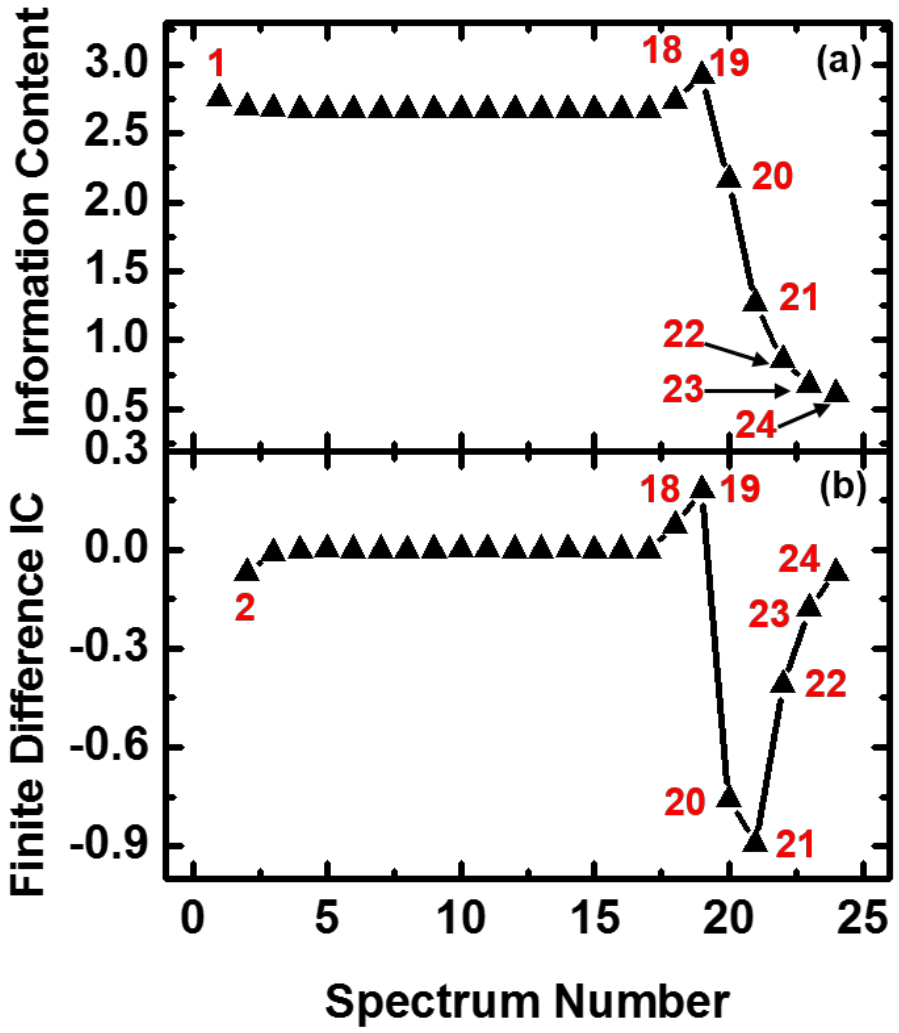


Figure 6.10 (a) IC values and (b) finite difference IC as a function of spectrum number for ToF-SIMS depth profiling of C_3F_6 polymer layer over Si substrate. A total of 19 peaks were selected from each spectrum.

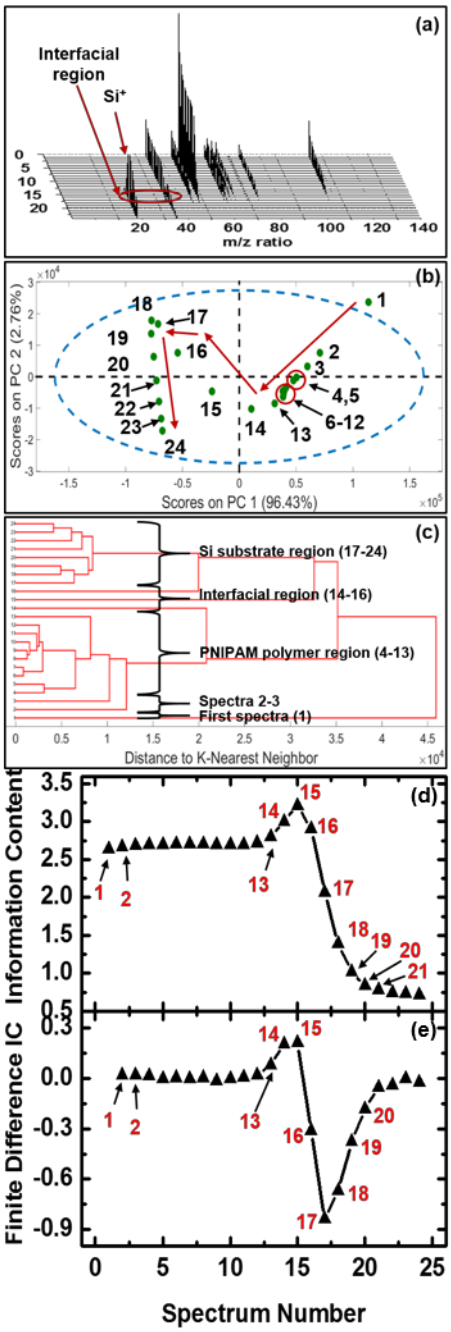


Figure 6.11 ToF-SIMS depth profiling study of PNIPAM/Si substrate: (a) 3-D graph of all 24 ToF-SIMS positive ion spectra, (b) scores plot of PCA analysis on the data, (c) cluster analysis on the data set, and (d) IC values and (e) finite difference IC as a function of spectrum number.

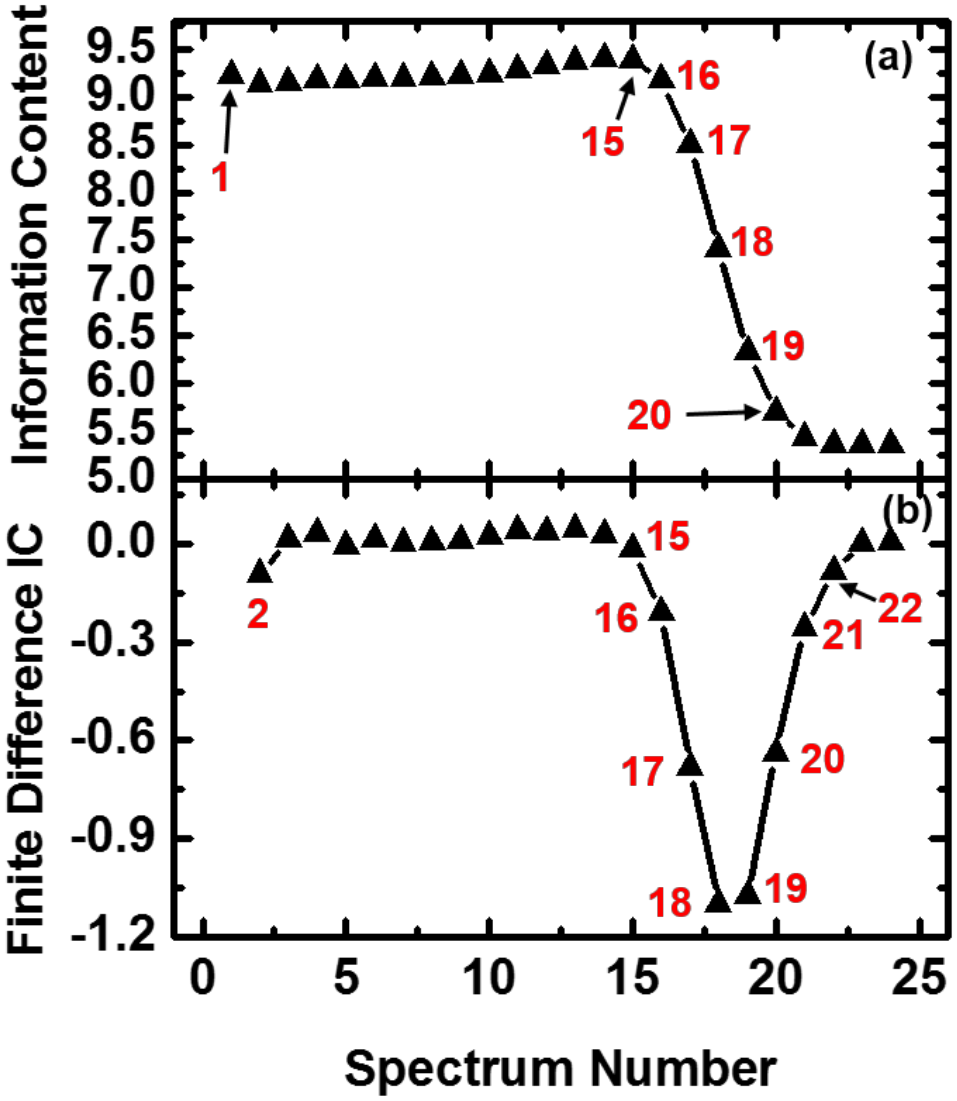


Figure 6.12 (a) IC values and (b) finite difference IC as a function of spectrum number for ToF-SIMS depth profiling of PNIPAM polymer layer over Si substrate. Entire spectra from m/z 0 till 150 were considered for the analysis.

Other uses of the IC function

The full characterization of a surface or a material usually requires the application of multiple analytical techniques,[30-34] and IC analysis appears to be well suited for clustering the resulting materials based on their spectra. For example, we recently characterized a set of nanodiamond samples from different vendors (Adamas, ITC, and AA (Advanced Abrasives)) using a suite of analytical techniques, which included positive ion ToF-SIMS, negative ion ToF-SIMS, and inductively coupled plasma (ICP).[22, 35] A series of peak areas was selected from each of these spectra/analyses, and their IC values were determined. Figure 6.13a shows a plot of the IC values for the positive and negative ion ToF-SIMS spectra for the nanodiamond samples. This plot, which is similar in format to a PCA scores plot, shows a clear difference between the Adamas samples, the AA uncleaned and ITC samples (clustered), and the AA double wash and AA triple wash samples (clustered). When a third dimension is included (the IC of the ICP results), the ITC samples are seen to be significantly different from the AA samples (see Figure 6.13b). Figure 6.13a shows very low IC values for the positive ion spectra of the Adamas samples (< 0.25 bits), while the other samples showed IC values of ca. 2.5 – 3.0 bits. This suggests that the positive ion spectra of the Adamas samples are not as complex as the other positive ion spectra. Analysis of the Adamas positive ion ToF-SIMS spectra shows that they are dominated by a peak from sodium (Na^+, $p = 0.98$), which explains their low IC values. Similarly, the AA samples have very low IC values (see Figure 6.13b). The ICP analyses of these materials showed large amounts of tungsten (~ 1000 ppm; $p = 0.83 – 0.94$), which explains their small IC values. Of course different methods for scaling such sets of peaks could be investigated to give them greater complexity. We have applied this type of IC analysis to data from other materials

and obtained similar separations between spectra. When more than three variables (IC values) are considered in this type of analysis, a radar chart might be employed.

We recently described the equivalent width (EW_{XPS}) and autocorrelation width (AW_{XPS}) as figures of merit for XPS narrow scans (see Chapters 3 and 4).[2, 36] We compare the EW_{XPS} and AW_{XPS} values to IC values for a series of C 1s narrow scans collected from four oxidized CNT samples. As shown in Figure 6.14, IC analysis gives similar results to these other functions.

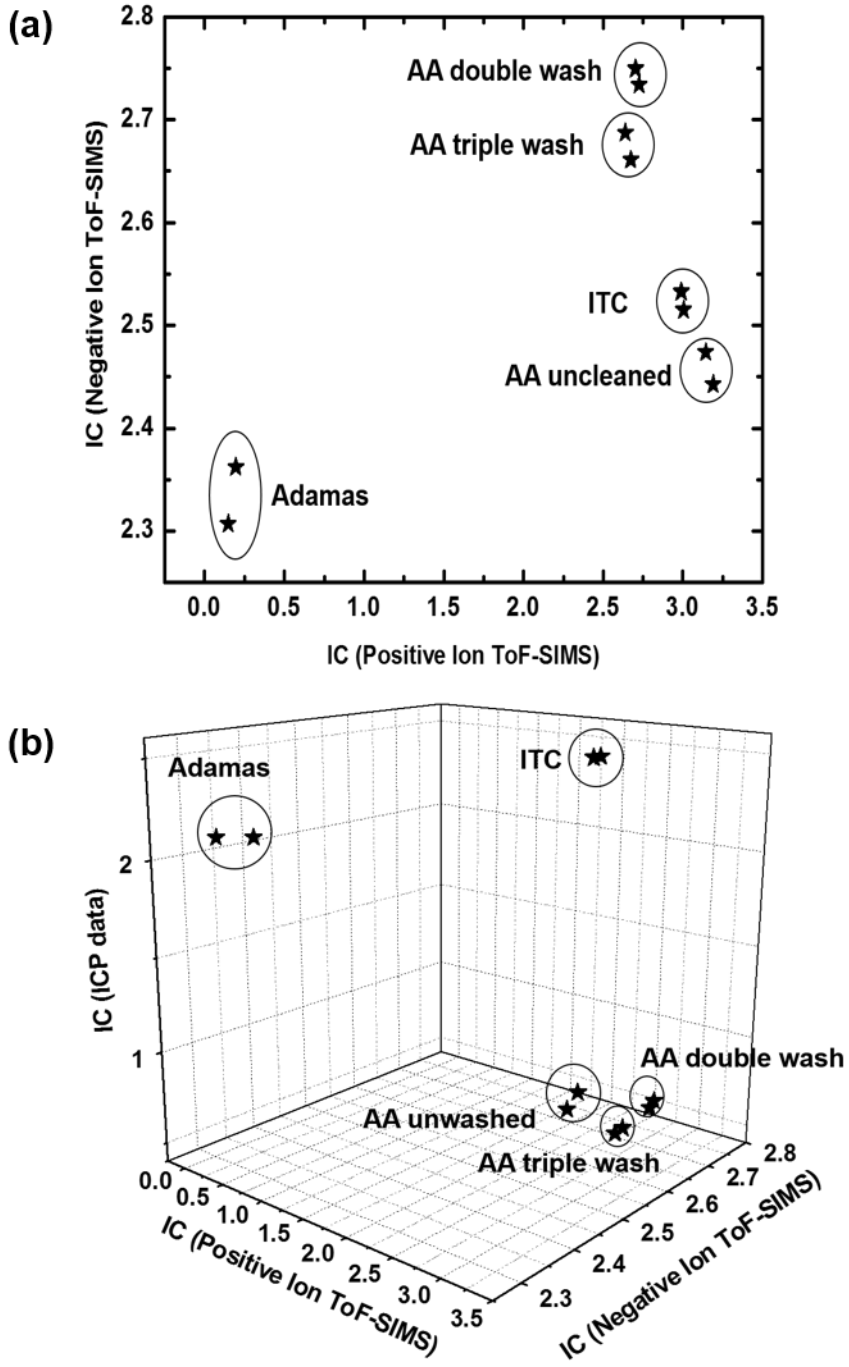


Figure 6.13 (a) Combining IC values from positive and negative ion mode ToF-SIMS analysis of five different nanodiamond samples. (b) Combining IC values from ICP, and positive and negative ion mode ToF-SIMS.

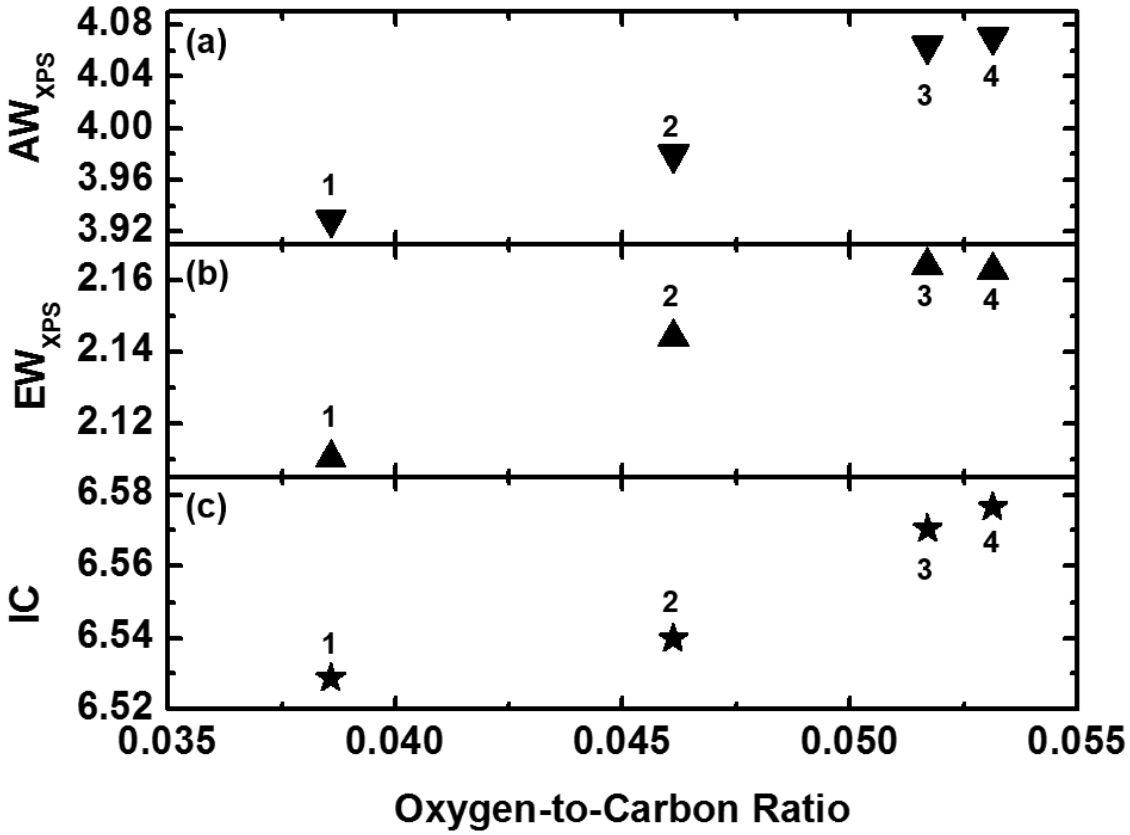


Figure 6.14 Comparison of (a) AW_{XPS} (eV), (b) EW_{XPS} (eV), and (c) IC (bits) values as a function of oxygen-to carbon ratio of a set of ozone primed CNT forests

6.5 Conclusions

The information content has been explored as a tool for classifying and clustering spectra and analytical data. In its direct and derivative forms, it appears to more easily show transitions in depth profiles. It also appears to cluster/separate data in a manner similar to classical PCA scores plots. Further work on the IC function will include application to additional complex data sets, and a more fundamental, statistical analysis of it.

6.6 Acknowledgment

We acknowledge the Department of Chemistry and Biochemistry and College of Physical and Mathematical Sciences at Brigham Young University for their support of this work. A portion of the research was performed using EMSL, a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research and located at Pacific Northwest National Laboratory. We gratefully acknowledge support of the ToF-SIMS data collection by NIH grant EB-002027 and NESAC/BIO.

Author Contributions: Dan Graham and Mark Engelhard provided the ToF-SIMS and XPS data sets respectively. B.S. performed PCA, cluster analysis, and IC analysis on these data sets, and made all the figures. B.S. had important discussions with M.R.L. leading to intellectual contributions to the project. One of the exciting things that came out of one of these discussions was to combine IC of a sample from various techniques and hence allowing the separation/grouping of samples in multi-dimensions. M.R.L. took leads from Information theory and suggested the use of Information content to complex data sets. M.R.L. and B.S. wrote the manuscript.

6.7 References

1. B. Singh and M. R. Linford, in *Vacuum Technology & Coating* (October, 2015).

2. B. Singh, D. Velázquez, J. Terry and M. R. Linford, J Electron Spectrosc **197** (0), 56-63 (2014).

3. A. Benninghoven, Surf Interface Anal **43** (1-2), 2-11 (2011).

4. A. Benninghoven, Surf Sci **35**, 427-457 (1973).

5. A. M. Belu, D. J. Graham and D. G. Castner, Biomaterials **24** (21), 3635-3653 (2003).

6. P. Bertrand and W. Lu-Tao, in *Microbeam and Nanobeam Analysis*, edited by D. Benoit, J.-F. Bresse, L. Van't dack, H. Werner and J. Wernisch (Springer Vienna, 1996), Vol. 13, pp. 167-182.

7. D. Briggs and A. B. Wootton, Surf Interface Anal **4** (3), 109-115 (1982).

8. A. Brown and J. C. Vickerman, Surf Interface Anal **6** (1), 1-14 (1984).

9. V. Gupta, H. Ganegoda, M. H. Engelhard, J. Terry and M. R. Linford, J Chem Educ **91** (2), 232-238 (2013).

10. N. Christophersen and R. P. Hooper, Water Resources Research **28** (1), 99-107 (1992).

11. A. W. F. Edwards and L. L. Cavalli-Sforza, Biometrics **21** (2), 362-375 (1965).

12. D. J. Graham and D. G. Castner, Biointerphases **7** (1), 49 (2012).

13. D. J. Graham, M. S. Wagner and D. G. Castner, Appl Surf Sci **252** (19), 6860-6868 (2006).

14. N. B. Gallagher, J. M. Shaver, E. B. Martin, J. Morris, B. M. Wise and W. Windig, Chemometr Intell Lab **73** (1), 105-117 (2004).

15. C. E. Shannon, At&T Tech J **27** (4), 623-656 (1948).

16. M. Ohya, Reports on Mathematical Physics **27** (1), 19-47 (1989).

17. R. W. Yeung and Z. Zhang, Information Theory, IEEE Transactions on **45** (4), 1111-1120 (1999).

18. E. T. Jaynes, Phys Rev **106** (4), 620-630 (1957).

19. J. E. Cohen, Behavioral Science **7** (2), 137-163 (1962).

20. C. Xin, B. Francia, L. Ming, B. McKinnon and A. Seker, Information Theory, IEEE Transactions on **50** (7), 1545-1551 (2004).

21. D. Guiqiang, P. Yangyang, X. Ningde, C. Varanasi and Z. Tong, Very Large Scale Integration (VLSI) Systems, IEEE Transactions on **20** (9), 1705-1714 (2012).

22. B. Singh, S. J. Smith, D. S. Jensen, H. F. Jones, A. E. Dadson, P. B. Farnsworth, R. Vanfleet, J. K. Farrer and M. R. Linford, submitted to Analytical and Bioanalytical Chemistry (2015).

23. D. K. Schroder, *Semiconductor Material And Device Characterization*, Third ed. (John Wiley & Sons, Inc., New Jersey, 2006).

24. C. B. Murray, D. J. Norris and M. G. Bawendi, J Am Chem Soc **115** (19), 8706-8715 (1993).

25. B. Ferguson and X.-C. Zhang, Nat Mater **1** (1), 26-33 (2002).

26. O. Jani, I. Ferguson, C. Honsberg and S. Kurtz, Appl Phys Lett **91** (13), 132117 (2007).

27. P. M. A. Sherwood, J. of Vac. Sci. Technol., A **14** (3), 1424-1432 (1996).

28. D. S. Jensen, S. S. Kanyal, N. Madaan, M. A. Vail, A. E. Dadson, M. H. Engelhard and M. R. Linford, Surface Science Spectra **20** (1), 36-42 (2013).

29. R. L. Burden and J. D. Faires, *Numerical Analysis*, 9th ed. (Cengage, 2011).

30. D. R. Baer, J. E. Amonette, M. H. Engelhard, D. J. Gaspar, A. S. Karakoti, S. Kuchibhatla, P. Nachimuthu, J. T. Nurmi, Y. Qiang, V. Sarathy, S. Seal, A. Sharma, P. G. Tratnyek and C. M. Wang, Surf Interface Anal **40** (3-4), 529-537 (2008).

31. D. R. Baer, M. H. Engelhard, G. E. Johnson, J. Laskin, J. Lai, K. Mueller, P. Munusamy, S. Thevuthasan, H. Wang, N. Washton, A. Elder, B. L. Baisch, A. Karakoti, S. V. N. T. Kuchibhatla and D. Moon, J. Vac. Sci. Technol., A **31** (5), 050820 (2013).

32. D. S. Jensen, S. S. Kanyal, N. Madaan, J. M. Hancock, A. E. Dadson, M. A. Vail, R. Vanfleet, V. Shutthanandan, Z. H. Zhu, M. H. Engelhard and M. R. Linford, Surf Interface Anal **45** (8), 1273-1282 (2013).

33. G. L. Jiang, F. Rivera, S. S. Kanyal, R. C. Davis, R. Vanfleet, B. M. Lunt, V. Shutthanandan and M. R. Linford, Opt Eng **50** (1) (2011).

34. H. Wang, B. M. Lunt, R. J. Gates, M. C. Asplund, V. Shutthanandan, R. C. Davis and M. R. Linford, ACS Appl. Mater. Interfaces **5** (17), 8407-8413 (2013).

35. D. F. Mitev, A. T. Townsend, B. Paull and P. N. Nesterenko, Carbon **60**, 326-334 (2013).

36. B. Singh, D. Velázquez, J. Terry and M. R. Linford, J Electron Spectrosc **197** (0), 112-117 (2014).

Chapter 7: Multi-Instrument Characterization of Five Nanodiamond Samples: A Thorough Example of Nanomaterial Characterization*

7.1 Abstract

Here we report the most comprehensive characterization of nanodiamonds (ND) yet undertaken. Five different NDs from three different vendors were analyzed by a suite of analytical techniques, including X-ray photoelectron spectroscopy (XPS), time-of-flight secondary ion mass spectrometry (ToF-SIMS), inductively coupled plasma – mass spectrometry (ICP-MS), diffuse reflectance infrared Fourier transform (DRIFT) spectroscopy, X-ray diffraction (XRD), transmission electron microscopy (TEM), electron energy loss spectroscopy (EELS), Brunauer-Emmett-Teller (BET) surface area measurements, and particle size distribution (PSD) measurements. XPS revealed the elemental compositions of the ND surfaces (83 – 87 at. % C, and 12 – 14 at. % O) with varying amounts of N (0.4 – 1.8 at. %), Si (0.1 – 0.7 at. %), and t (0.3 at. % only in samples from one vendor). ToF-SIMS and ICP showed metal impurities (Al, Fe, Ni, Cr, etc. with unexpectedly high amounts of W in one vendor's samples: ca. 900 ppm). PCA were performed on the ToF-SIMS and ICP data. DRIFT showed key functional groups (–OH, C=O, C-O, and C=C). BET showed surface areas of 50 – 214 m^2/g. XRD and TEM revealed PSD (bimodal distribution and a wide PSD, 5 – 100 nm, for one vendor's samples). XRD also provided particle sizes (2.7 – 27 nm) and showed the presence of graphite. EELS gave the sp^2/sp^3 contents of the materials (37 – 88 % sp^3). PSD measurements were performed via differential sedimentation of the particles (mean particle size ca. 17 – 50 nm). This comprehensive understanding should allow for improved construction of nanodiamond-based materials.

*This chapter is reproduced from (Bhupinder Singh, Stacey J. Smith, David S. Jensen, Hodge F. Jones, Andrew E. Dadson, Paul B. Farnsworth, Richard Vanfleet, Jeffrey K. Farrer, and Matthew R. Linford) *Analytical and Bioanalytical Chemistry*, accepted for publication, 2015

7.2 Introduction

Nanoparticles have emerged as important building blocks for many advanced materials. Indeed, for years, all things 'nano' have been hot topics in science and engineering. For example, a Web of Science search showed that the terms 'nanomaterial', 'nanoparticle', or 'nanostructure' were mentioned less than 100 times in publications in 1990, but that these terms appeared in almost 45,000 publications in 2011.[1] This is a remarkable increase. With this large number of publications has come an enormous number of schematics/cartoons that illustrate and explain the use of these nanoparticles/nanomaterials. Accordingly, in a slightly derogatory, but also humorous, way, some have referred to areas of nanotechnology as 'cartoonland'. At least in part, this is because the schematics accompanying these studies often show idealizations of nanocomponents that are combined into nanodevices and nanostructures, while the underlying complexity of these materials and their connectivity is ignored. Indeed, it is becoming increasing recognized that many nanomaterials are extremely complex. For example, their surfaces are often highly heterogeneous, composed of multiple crystal faces, defects, impurities, and chemical functional groups. This complexity also extends to their interiors. It follows that nanoparticles and nanomaterials from different research groups or vendors can be dramatically different from each other, where the exact compositions of these materials can have profound implications for nanomedicine/toxicology/environmental applications, etc.[2, 3] With the increasing recognition of the complexity of many, if not most, nanomaterials, a call has come for their more exhaustive characterization.[1, 4, 5] Because current surface and material analytical techniques generally have specific strengths and capabilities, as well as weaknesses and deficiencies, full characterization of a material will rarely be done with a single instrument; it will almost always require a suite of techniques employed at a high level. At present there are many studies in the

literature that only show, at best, moderate characterization of the nanomaterials described in them, which may limit the concrete conclusions that can be drawn from them.

Extensive, multi-instrument characterization of a material increases one's understanding of it in a way that should lead to improvements in the devices that contain it.[6, 7] Accordingly, we have undertaken the present study to improve the performance of some of our nanodiamond-based superficially porous particles for high performance liquid chromatography (HPLC). To the best of our knowledge, this is the most comprehensive study of nanodiamond that uses the broadest array of relevant analytical tools that has been undertaken to date. We believe that this as an exemplary/model study that might be viewed and used as a benchmark for nanomaterial characterization. Nanodiamond samples from multiple vendors were analyzed. To obtain their elemental compositions, surface areas, particle size distributions, and to understand the functional groups present at their surfaces, the materials were characterized by X-ray photoelectron spectroscopy (XPS), time-of-flight secondary ion mass spectrometry (ToF-SIMS), diffuse reflectance infrared Fourier transform spectroscopy (DRIFT), transmission electron microscopy (TEM), electron energy loss spectroscopy (EELS), X-ray diffraction (XRD), inductively coupled plasma-mass spectrometry (ICP-MS), and Brunauer-Emmett-Teller (BET) surface area analysis. To better understand the ToF-SIMS data, multivariate analyses (principal component and cluster analyses) of the spectra were performed. The unusual perspective that is obtained through these multiple techniques has led to unexpected findings and insights into these materials.

The mechanical and chemical properties that make diamond a material of choice for liquid chromatography are discussed in Chapter 1 (section 1.5). Apart from Dr. Linford's group at BYU, Nesterenko and coworkers have also worked extensively on nanodiamond based liquid

chromatography phases. [8, 9, 12, 24] As part of their latest work, they have emphasized the issue of metal impurities in nanodiamonds and the lack of robust/reliable analytical methods for their determination. In particular, they recently developed a direct infusion ICP-MS protocol for reliable determination of metal impurities in nanodiamond samples.[25] The ICP work described herein is based on their approach. Nanodiamond has also shown promise in the biological arena due to its ease of surface functionalization, and ability to conjugate with biomolecules, where nanodiamonds are generally considered to be biocompatible.[26, 27] These properties make nanodiamonds suitable for applications that include biosensors, cell imaging, fluorescent bio-labeling, drug delivery, protein purification, and separations.[26, 28, 29] For many of these purposes, functionalized diamond particles have been prepared.[14, 15, 30-37]

In 2011, Shenderova et al. noted that the "Detailed and unambiguous characterization of the surface structure of detonation nanodiamond (DND) particles remains one of the most challenging tasks for the preparation of chemically functionalized nanodiamonds."[38] Thus, it is no wonder that there have been reports on the characterization of diamond nanoparticles in the literature. Nevertheless, those we are aware of rely on relatively few analytical techniques.[39-43] Thus, we emphasize again that (i) the present study appears to be the most comprehensive undertaken to date, (ii) we believe it will serve as an exemplary study for the characterization of other nanomaterials, and (iii) the use of nanomaterials is increasing in electronics, health, medicine, energy production, etc., where the use and application of these materials will certainly be advanced by their more thorough characterization.

7.3 Materials and Methods

7.3.1 Reagents and Materials.

Table 7.1 Nanodiamonds characterized in the present study.

Sample #	Vendor	Size (nm)	Abbreviation
1	Advanced Abrasives (AA) Corporation, Pennsauken, NJ 08110	50	AA 50 nm unwashed
2	Advanced Abrasives Corporation, Pennsauken, NJ 08110	50	AA 50 nm double wash
3	Advanced Abrasives Corporation, Pennsauken, NJ 08110	50	AA 50 nm triple wash
4	International Technology Center (ITC), Raleigh, NC -27709	50	ITC 50 nm
5	Adamas Nanotechnologies, Inc., Raleigh, NC 27617	5	Adamas 5 nm

Table 7.1 identifies the sources, sizes, and abbreviations used herein for the nanodiamonds characterized in the present study. Three different samples were procured from Advanced Abrasives Corporation (AA). These samples were nominally identical in size, but differed in the degree to which they had been leached with acid to remove metal impurities. In particular, the 'AA 50 nm unwashed' sample was not leached, the 'AA 50 nm double wash' sample was treated twice with a proprietary acid wash, and the 'AA 50 nm triple wash' sample was treated thrice with the same wash. We have previously used nanodiamond from Advanced Abrasives to make core-shell particles.[14, 17, 18] As a comparison to the 50 nm Advanced Abrasives samples, a different 50 nm sample was procured from ITC. And finally, to study a material that should be substantially different from the 50 nm samples, a 5 nm sample from Adamas Technologies was also analyzed. All nanodiamond samples were procured as liquid suspensions. The AA and ITC samples are high pressure high temperature (HPHT) nanodiamonds, which were milled, whereas the Adamas diamonds are detonation nanodiamonds. High purity water (18 MΩ resistance) was obtained from a Milli-Q Water System (Millipore, Billerica, MA).

7.3.2 XPS

X-ray photoelectron spectroscopy (XPS) measurements were performed using a Surface Science SSX-100 X-ray photoelectron spectrometer (serviced by Service Physics, Bend, OR), with a monochromatic Al K_α source (1486.7 eV) and a hemispherical analyzer. An electron flood gun was employed for charge compensation. Survey scans were recorded as follows: spot size: 800 x 800 μm^2, resolution: 4 (pass energy 150 eV), number of scans: 4, and step size: 1 eV. Narrow scans were recorded as follows: spot size: 500 x 500 μm^2, resolution: 3 (pass energy 100 eV), number of scans: 20, and step size: 0.065 eV. Peaks were referenced to the C 1s hydrocarbon signal taken at 284.6 eV. The instrument software provided atomic compositions

based on peak areas from narrow scans. Samples for XPS were prepared on silicon wafers. The wafers were rinsed with ultrapure water, but no effort was made beyond this to clean the adventitious hydrocarbon layer on them that made them moderately hydrophobic. Droplets of slurries of nanodiamond particles beaded up to a small extent, but did not fully wet these surfaces or run off them. Wafers with droplets on them were placed on a hot plate to evaporate the solvent, leaving a thin, uniform film of nanodiamonds.

7.3.3 ToF-SIMS

Samples for ToF-SIMS and XPS were prepared in the same manner. Static time-of-flight secondary ion mass spectrometry (ToF-SIMS) was performed on a TOF-SIMS IV instrument (ION-TOF GmbH, Münster, Germany) with a 25 keV Ga^+ source over a 150 x 150 μm^2 sample area. An electron flood gun was used for charge compensation. The analyses were performed within the static limit, i.e., analyzed regions saw 5 - 7 x 10^{11} ions/cm^2. For each sample, two spots were analyzed in positive ion mode and two spots in the negative ion mode. Care was taken not to analyze the same area twice.

7.3.4 ICP-MS

Inductively coupled plasma – mass spectrometry (ICP-MS) was performed using an Agilent Technologies 7800 ICP-MS, equipped with an Agilent SPS-4 autosampler (Santa Clara, CA). As ICP-MS is very sensitive, the nanodiamond slurries were diluted 1000 times with ultrapure water to attain final concentrations of ca. 0.005 wt. %. The metal concentrations at these dilutions were expected to be in the *ppb* regime. These diluted slurries were directly injected into the ICP and ultrapure water was used as the blank.[25] The slurry samples were well shaken just before they were injected into the ICP-MS to homogenize them. Standard solutions

(in the ppb regime) of various metals were used to generate calibration curves. All concentrations were solution based (vol/vol). Details concerning the standard solutions are in Table 7.2. The instrument parameters were: RF power: 1540 W, peristaltic pump: 0.1 rpm, plasma gas: argon (15 L/min), integration time: 0.1 sec.

7.3.5 DRIFT

Diffuse reflectance infrared Fourier transform (DRIFT) spectroscopy measurements were performed on a Thermo-Scientific Nicolet 6700 FTIR spectrometer (Waltham, MA). The following were the analysis parameters used: resolution: 4 cm^{-1}, number of scans: 512, apodization: Happ-Genzel, and final format: absorbance. Samples for FTIR were prepared by evaporating the solvent from the respective nanodiamond slurries to obtain chunks/flakes of nanodiamonds, which were ground in a mortar and pestle to yield a fine powder. The nanodiamond powders were diluted with two parts of potassium bromide (KBr) by weight per part nanodiamond and analyzed.

7.3.6 TEM / EELS

Imaging and diffraction patterns were obtained using an FEI Tecnai F30 transmission electron microscope (TEM, Hillsboro, Oregon) operated at 300 keV. Electron energy loss spectroscopy (EELS) was performed on an FEI Tecnai F20 TEM, equipped with an FEG source operated at 200 keV. To prepare the samples, the nanodiamond slurries were diluted 50 times with ultrapure water and a drop of the diluted slurry was placed on a TEM grid and allowed to evaporate.

Table 7.2 Compositions and dilutions of the standard solutions used to make ICP calibration curves.

Standard solution #	**Composition**	**Dilutions used**
1	10000 ppm W, 100 ppm of Mo, Nb, Sb, Si, Ti	10, 1000, 5000 and 10000 ppb of W 0.1, 1, 10, 50, and 100 ppb of Mo, Nb, Sb, Si, and Ti in ultrapure water
2	10000 ppm Co, 100 ppm Cu, Fe, K, Mg, Mn, Na, Ni, Pb, Se, Sr, Tl, Zn, Ag, Al, As, B, Ba, Ca, Cd, Cr	1, 10, 100, and 1000 ppb of Co 0.1, 1, 10, 50, and 100 ppb of other metals in ultrapure water
3	1000 ppm Fe	1, 10, 50, 100, and 1000 ppb of Fe in ultrapure water
4	1000 ppm Zn	1, 10, 100, 1000 ppb of Zn in ultrapure water

7.3.7 XRD

X-ray diffraction (XRD) data were collected using a PANalytical X'Pert Pro diffractometer (Westborough, MA) with a Cu X-ray source (λ = 1.54 Å) and an X'Celerator detector. Data were collected from 15 – 150° 2θ using a step size of 0.017° 2θ with an 80 s/step collection time. Automatic slits were used to maintain an 8 mm beam footprint on the thin layer (0.1 mm) of each sample throughout the entire angular range, thereby fulfilling the constant volume requirement for quantitative analysis. XRD data for a LaB_6 standard (NIST) were collected using the same experimental conditions in order to characterize the instrumental peak broadening. Profile fits of the data were performed using the PANalytical Highscore Plus software. The integral breadths of the peaks determined from these fits (corrected for the instrumental peak broadening) were used in average Williamson-Hall plots (Langford method) to estimate the crystallite size and microstrain for each sample.

7.3.8 BET

BET was performed on a Micromeritics (Norcross, GA) TriStar II Surface Area and Porosity System. The nanodiamond slurries were dried and chunks/flakes of the resulting material were ground into a fine powder with a mortar and pestle. Prior to analysis, the samples were degassed under nitrogen for 24 hours at 200 ºC.

7.3.9 Particle size distribution

Particle size distributions (PSD) were determined using a disc centrifuge system, Model DC24000 (CPS Instruments, Prairieville, LA). This technique operates on the principle of differential sedimentation of particles in a fluid. The operating parameters were: disc speed: 24000 rpm; calibration standard diameter: 0.377 μm; calibration standard density: 1.385 g/mL;

nanodiamond (ND) density: 3.51 g/mL; ND refractive index: 2.41; fluid density: 1.055 g/mL; fluid refractive index: 1.36; and number of data points: 1338-1354.

7.3.10 Chemometrics techniques

Principal components analysis (PCA) and cluster analysis of the ToF-SIMS data were performed using the PLS_Toolbox version 7.9.3 (Eigenvector Research, Inc., Wenatchee, WA, USA) in MATLAB version R2014b (Natick, MA). The raw peak areas were organized row-wise in a data matrix. The data were preprocessed by normalization (a row operation in which each peak area from a sample was divided by the sum of the peak areas for that sample) and autoscaling (a column operation in which the data corresponding to a certain m/z value were mean centered and then divided by the standard deviation of the values).

7.4 Results and Discussion

7.4.1 XPS

XPS probed the surface elemental compositions of the nanodiamonds. As expected, the samples were comprised mostly of carbon and oxygen. Figures 7.1a-e show survey scans from AA 50 nm unwashed, AA 50 nm doubly washed, AA 50 nm triply washed, ITC 50 nm, and Adamas 5 nm nanodiamond samples. All five spectra show strong C 1s and O 1s signals, suggesting oxidized carbon at the nanodiamond surfaces. Table 7.3 summarizes the elemental compositions of the samples as obtained from XPS narrow scans. Along with carbon and oxygen, low levels of nitrogen and silicon are present to varying degrees in all the samples. The AA samples also showed relatively high levels of tungsten (W), where the presence of W was unexpected, and its source is unknown. Only the Adamas 5 nm material showed Na (ca. 1 at. %).

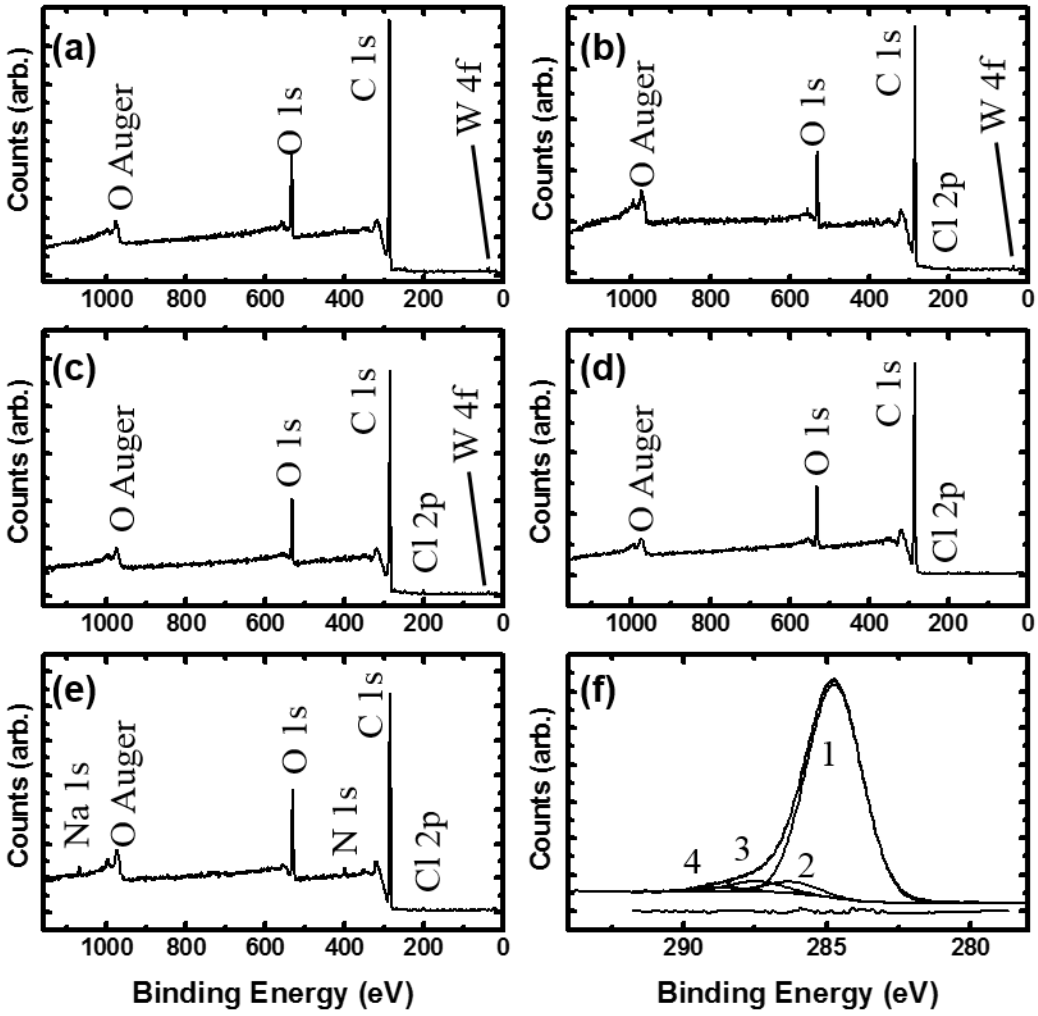


Figure 7.1 XPS survey scans of nanodiamond samples: (a) AA 50 nm unwashed, (b) AA 50 nm double washed, (c) AA 50 nm triple washed, (d) ITC 50 nm, (e) Adamas 5 nm and (f) peak fitting of AA 50 nm triple washed C 1s spectrum. Note: The residuals shown in Fig 1f are calculated as the counts in the experimental peak envelope minus the counts in the peak fitting envelope.

Table 7.3 Elemental compositions (at. %) of the nanodiamond samples, as determined by XPS.

Metals	AA 50 nm unwashed	AA 50 nm double washed	AA 50 nm triple washed	ITC 50 nm	Adamas 5 nm
C 1s	82.7	83.4	84.7	86.7	82.6
O1s	13.2	12.1	11.8	12.1	14.0
N1s	1.5	1.3	0.9	0.4	1.8
Si 2p	0.3	0.3	0.3	0.7	0.1
W 4f	0.4	0.3	0.3	0	0
Cl 2p	0	1.0	0.5	0.2	0.5
Na 1s	0	0	0	0	1.0
O/C ratio	0.1596	0.1447	0.1388	0.1391	0.1692

While some samples charged, good C 1s narrow scans were obtained from the AA unwashed and triply washed samples. These spectra were peak fitted to four signals that corresponded to (1) aliphatic C-H (284.6 eV, C(0)), (2) C-O (286.1 eV, C(I)), (3) C=O (287.3 eV, C(II)) and (4) C(O)=O (288.5 eV, C(III)).[44] In general, there are many possible adjustable parameters in the fitting of a complex XPS peak envelope. Accordingly, care should be taken to ensure that the final results are reasonable/believable.[45] Thus we fitted both narrow scans in the same way, constraining all the peak widths to be the same in each scan. In addition we used 100% Gaussians line shapes. Peaks with Lorentzian character gave poorer fits, which suggested a fair degree of heterogeneity in the surfaces[46]. In addition, we constrained the peak positions to the values given above, and used only symmetric peaks. Under these conditions, good fits (low chi-squared values/low residuals) were obtained. The peak fitting of the C 1s spectrum from the unwashed AA sample is shown in Figure 7.1f, and the corresponding results from both AA samples are summarized in Table 7.4. This table indicates that the carbon in these nanodiamond samples is 11 – 15% oxidized, where these oxidized moieties presumably correspond to alcohols and/or ethers, aldehydes and/or ketones, and carboxylic acids and/or esters. The triple washed particles show more oxidized carbon than the unwashed material, i.e., they show significantly more carbon in its +1 and +3 oxidation states and comparable amounts of carbon in its +2 oxidation state.[44] These results will be discussed below vis-à-vis the analysis of the DRIFT spectra.

Table 7.4 Peak fitting results of C 1s narrow scans of AA unwashed and triple washed samples.

Sample	Aliphatic C-H (%)	C-O (%)	>C=O (%)	-O-C=O- (%)
AA unwashed 50 nm	88.8	4.8	4.6	1.8
AA triple washed 50 nm	84.9	8.0	3.9	3.1

The C 1s narrow scans of the other samples showed artifacts due to charging and were not considered for peak fitting.[47] Nevertheless, all of the narrow scans, including those that charged, were of sufficient quality to conclude that no shake-up signals ($\pi \rightarrow \pi^*$) were present. Turgeon *et al*. studied mixtures of polystyrene and polyethylene. They showed that no shake-up signal was observable unless the amount of aromatic carbon in the mix was more than ca. 30%.[48] If their results can be generalized, we can conclude that there is less than 30% aromatic (sp^2) carbon in our nanodiamond materials. This analysis is relevant because XRD suggests that some of the nanodiamond samples contain some graphitic carbon (vide infra).

7.4.2 ToF-SIMS

ToF-SIMS is complementary to XPS, showing higher sensitivity: ppm to ppb for some species vs. 0.1-1.0 at. % for XPS. In general, it also provides more chemical information than XPS. Both of these capabilities were used in the SIMS analysis of our nanodiamonds. Results are summarized in Table 7.5. Because the matrix effect of SIMS makes quantitation difficult, the entries in Table 7.5 are based on raw peak areas. A strong iron signal was present in the AA 50 nm unwashed sample. As expected, the iron peak then decreased in the double and then triple washed samples. The ITC 50 nm sample also showed a large iron peak. Consistent with the XPS results, the Adamas 5 nm sample showed a strong Na^+ signal. Significant chromium and aluminum signals were present in the Adamas 5 nm and ITC 50 nm samples, respectively. Smaller peaks corresponding to other elements, e.g., silicon, potassium, calcium, titanium, etc., were also observed in the samples.

Table 7.5 Summary of the metals detected by ToF-SIMS (comparison based on raw peak areas).

Metals	AA 50 nm unwashed	AA 50 nm double washed	AA 50 nm triple washed	ITC 50 nm	Adamas 5 nm
Sodium (Na^+)	+	+	+	+	++
Magnesium (Mg^+)	-	-	-	+	-
Aluminum (Al^+)	+	+	+	++	+
Silicon (Si^+)	+	+	+	+	-
Potassium (K^+)	+	+	+	+	+
Calcium (Ca^+)	-	+	+	+	-
Titanium (Ti^+)	-	-	-	+	-
Chromium (Cr^+)	+	+	+	+	++
Manganese (Mn^+)	-	-	-	+	-
Iron (Fe^+)	++	+	+	++	+
Tungsten (W^+)	-	-	-	-	-

('+' refers to element being present; '++' refers to relatively high peak area; '-' denotes element being absent)

Interestingly, the samples showing tungsten by XPS (see Figure 7.1), showed no W^+ signal by positive ion SIMS. However, the corresponding negative ion analysis revealed substantial signals at m/z 230 – 235 and at m/z 246 – 251. Based on the natural abundances and intensities of the different isotopes of tungsten, these values and peak intensities appeared to correspond to WO_3^- and WO_4H^- (see Figure 7.2). However, this assignment could not explain either the peaks at m/z 233 and 235 in the proposed WO_3^- envelope, or the peaks at m/z 246 and 250 in the proposed WO_4H^- envelope. A much better fit to the data was obtained when these unexplained peaks were assumed to be due to WO_3H^- and WO_4^-, respectively. Thus, the two series of peaks in Figure 7.2 at ca. 230 – 235 and 246 – 251 appear to correspond to combinations of signals from WO_3^- and WO_3H^-, and WO_4^- and WO_4H^-, respectively. Details of this analysis are included in Appendix 6. Similar to the results for W, small signals from Se and Mo in their oxidized forms, e.g., SeO_4^- and MoO_3^-, were observable in the AA samples.

The XPS narrow scans of W (see Figure 7.3) provide additional evidence for the presence of oxidized tungsten. Here, the W 4f peaks are shifted by ca. +4 eV compared to the position expected for metallic tungsten (31 eV). This result is clearly consistent with tungsten bonded to multiple oxygen atoms.[49] These XPS and SIMS results again reveal the complementary nature of these techniques.

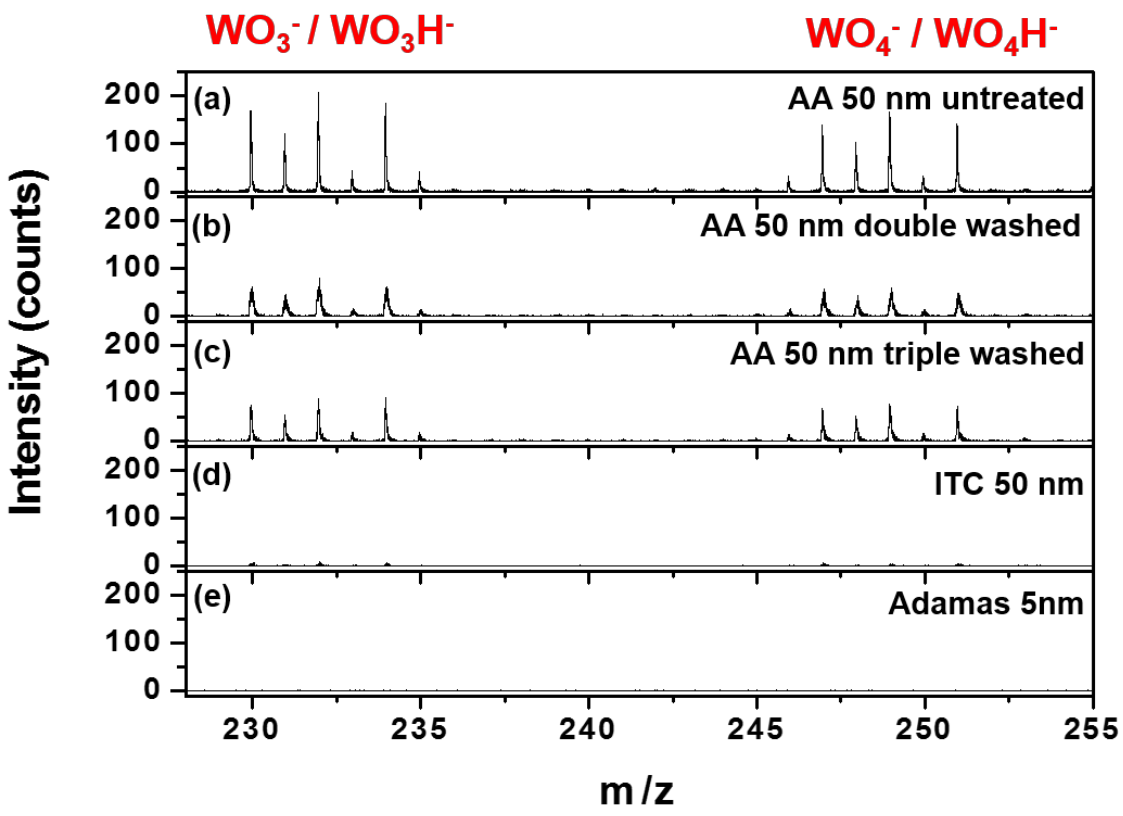


Figure 7.2 Negative mode ToF-SIMS spectra of oxidized tungsten species in the nanodiamond samples (a) AA 50 nm unwashed, (b) AA 50 nm double washed, (c) AA 50 nm triple washed, (d) ITC 50 nm and (e) Adamas 5 nm.

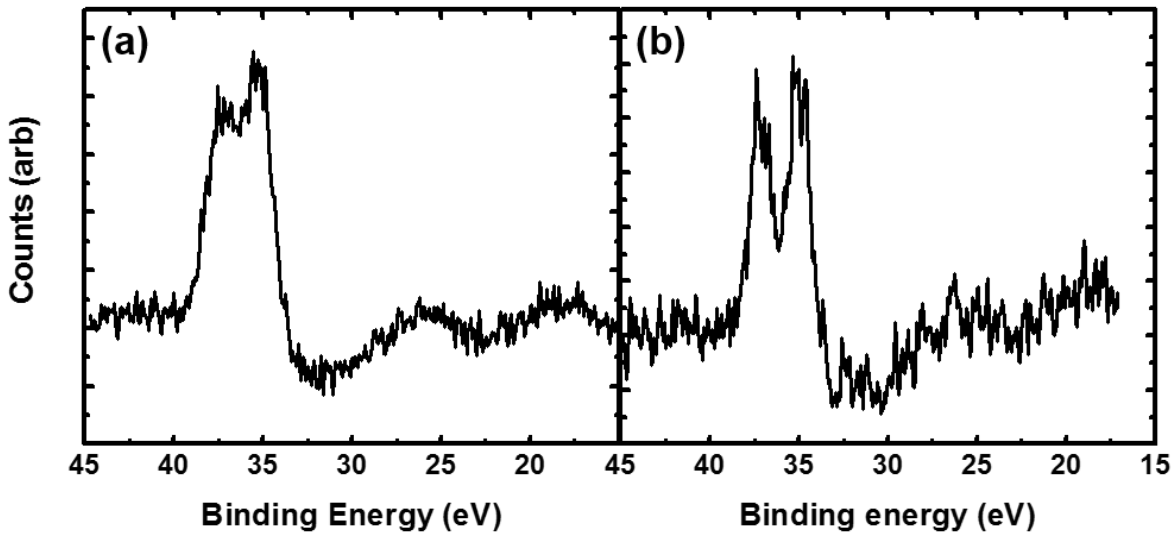


Figure 7.3 XPS W 4f narrow scans of (a) AA 50 nm untreated and (b) AA 50 nm triple washed samples.

SIMS spectra can be very complex, which can make them difficult to compare. For this reason, chemometrics methods, i.e., advanced statistical/pattern recognition tools, have often been applied to SIMS spectra.[50-53] Two frequently used tools are principal components analysis (PCA) and cluster analysis. As previously noted, two spots were analyzed by SIMS on each sample at each polarity. Accordingly, from the positive ion analysis, 10 data sets were subjected to PCA. The first two principal components (PCs) in this analysis, PC1 and PC2, capture 53.4% and 27.4% of the variation in the data, respectively. On the resulting two-dimensional plot of the scores on PC1 vs. PC2 (Figure 7.4a), the data separate well into three groups, each of which corresponds to one of the vendors of the samples. This analysis suggests that the materials from the different vendors are quite different from each other. The loadings plots for PC1 and PC2 (see Figure 7.5a-b) help reveal the source of chemical variation between these data points (spectra). On the average, the materials that have negative scores on PC1, i.e., the Adamas material and to a lesser extent the ITC nanodiamonds, are richest in Na^+, Mg^+, Al^+, Ti^+, and Mn^+, while those with the most positive scores, i.e., the AA samples, are richer in a series of light hydrocarbon fragments, as well as H^+, NH_4^+, H_3O^+, Si^+, K^+, Ca^+, Cr^+, and Fe^+. The ITC samples had positive scores on PC2 (see Figure 7.4a), which suggested that they were, on the average, richer in Mg^+, Al^+, Ti^+, and Mn^+. In contrast, the Adamas samples showed negative scores on PC2, indicating that they were richer in Na^+ and Cr^+. PC 3 also captured a moderately large fraction of the variation in the data (15.5%), and this PC nicely separated the unwashed AA samples from the AA double and triple washed materials (see Figure 7.4b). As expected, the loadings on PC3 showed more Fe^+ for the unwashed (unleached) AA samples (see Figure 7.5c). Clearly this PCA analysis can separate the spectra into four distinct groups that correspond to vendor and chemical treatment. A plot of the Q residuals vs. Hotelling T^2 for this data showed

that all of the spectra fell within 95% confidence limits. A PCA analysis of the negative ion spectra was similarly performed. Figure 7.6a shows the scores plot of PC1 vs PC2 which is able to separate the samples into three groups corresponding to samples from three different vendors. With three PCs, the PCA analysis of the negative ion SIMS data separated the data into five distinct groups. This is quite well shown in the scores plot of PC1 vs PC3 (see Figure 7.6b). The loadings plot for PC1 (see Figure 7.7a) indicates that the ITC and Adamas spectra show strong O^- and Cl^- signals, while the AA samples show stronger signals from almost all of the other ions considered in the analysis, which included WO_3^-, WO_4H^-, SeO_4^-, and MoO_3^-. In general agreement with these results, XPS showed no chlorine for the AA untreated sample, which had the highest positive value in the scores plot. The loadings plot on PC3 (see Figure 7.7c) was helpful for understanding the source of variation between the different AA samples. The AA double and triple washed samples were richer in WO_3^-, WO_4H^-, Cl^-, F^-, etc., whereas the AA untreated sample was richer in SeO_4^-, MoO_3^-, CN^-, etc. Similar results were obtained from XPS, which showed chlorine in the AA double and triple washed samples. XPS also showed decreasing N 1s signals from the AA untreated to the AA double and triple washed samples. This was elegantly picked up by PCA, which showed a maximum signal for CN^- in the uncleaned AA sample.

When two or more chemometrics techniques that depend on different algorithms yield the same result, one can be more certain in that finding.[52-54] Accordingly, cluster analysis was also used to compare the spectra. The resulting dendrograms for the positive and negative ion ToF-SIMS data (see Figure 7.8) showed clustering of the different samples. The fact that the PCA and cluster analyses show similar clustering/differentiation of the samples is a confirmation of the application of chemometrics to this problem.

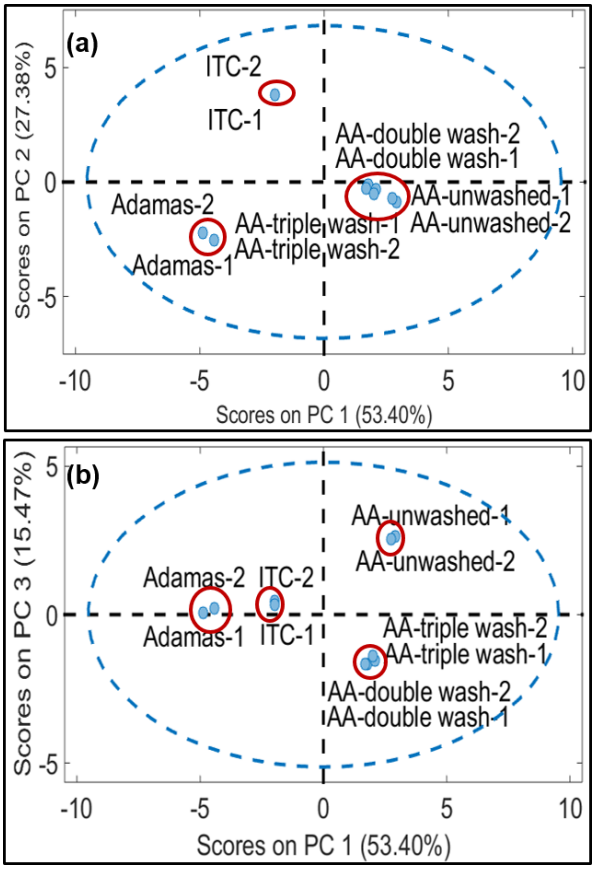


Figure 7.4 PCA scores plots of positive ion SIMS data of nanodiamonds, **(a)** PC 1 vs. PC 2 and **(b)** PC 1 vs. PC 3. The dashed blue line represents a 95% confidence limit for the distribution of the data points in the plane. The red circles are guides to the eye.

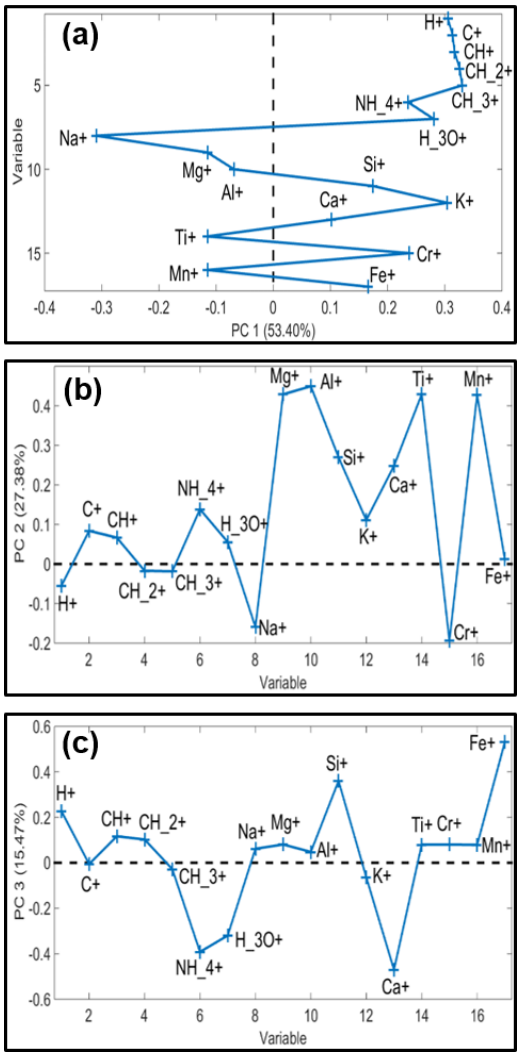


Figure 7.5 Loadings plots from the positive ion ToF-SIMS analysis of nanodiamonds of **(a)** PC 1, **(b)** PC 2, and **(c)** PC 3.

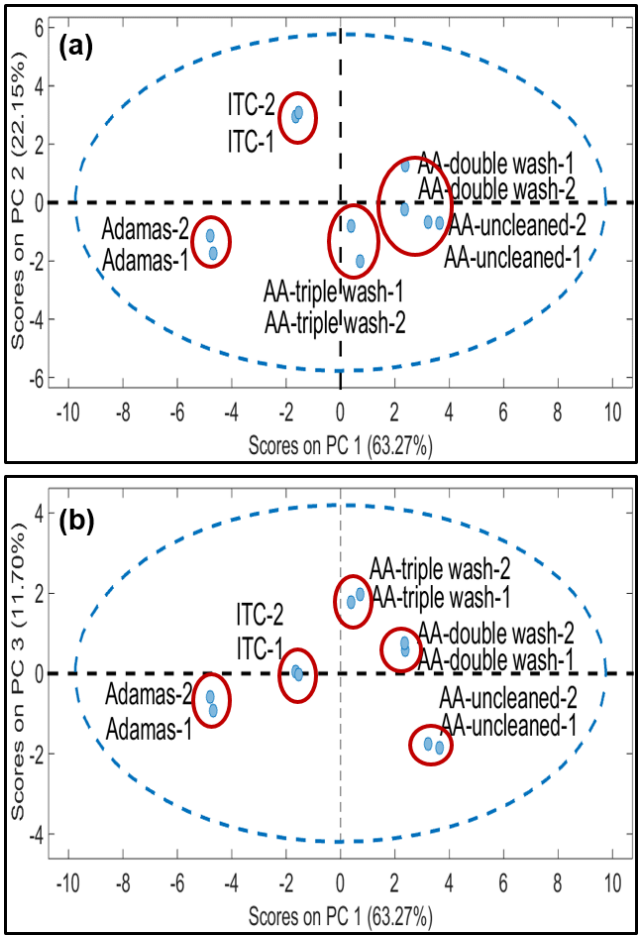


Figure 7.6 Scores plot of negative ion ToF-SIMS analysis of nanodiamond samples: (a) PC1 vs. PC2, and (b) PC1 vs. PC3.

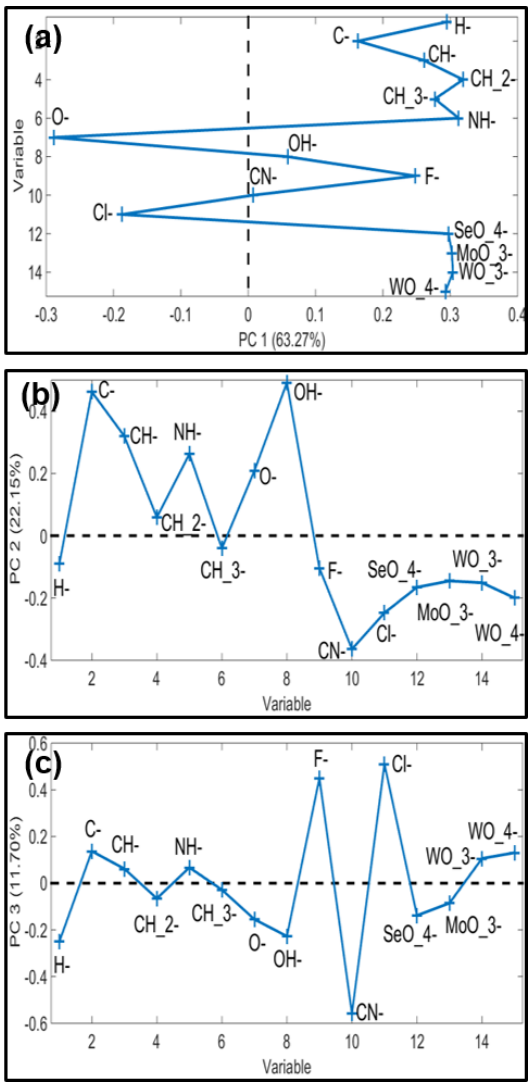


Figure 7.7 Loadings plot of (a) PC1, (b) PC2, and (c) PC3 of negative ion ToF-SIMS analysis of the nanodiamond samples.

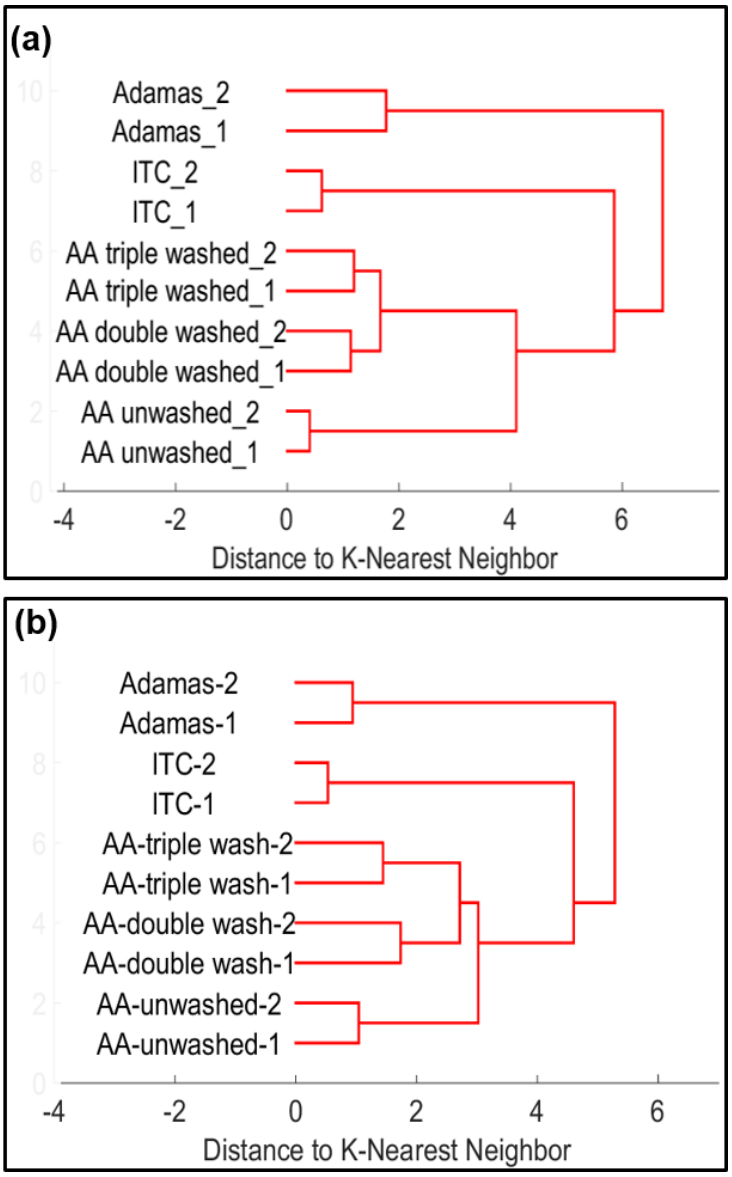


Figure 7.8 Dendrograms representing cluster analyses of ToF-SIMS (a) positive and (b) negative ion spectra obtained for nanodiamonds.

7.4.3 ICP-MS

Unlike XPS and ToF-SIMS, ICP is a bulk analytical technique. It allows quantitative elemental concentrations of a material to be determined. Since metal ions are known to cause tailing in liquid chromatography,[6, 55-57] and because NDs have high surface areas and are present in high abundance in the shells of our core-shell HPLC particles, an understanding of the metal content of these materials is important. Table 7.6 shows the results for 12 elements from ICP analyses of the nanodiamond samples. Consistent with the XPS results that indicated significant tungsten (see Table 7.3 and Figure 7.1), ICP shows very high signals for this element in the AA samples (ca. 900 ppm). In contrast, in the ITC and Adamas samples, W is present at the ca. 8 ppm level. Iron was present at quite a high level in the unwashed AA sample. As expected, the Fe signal in the AA samples decreases substantially as they are washed. SIMS and ICP give somewhat different results for Fe in the ITC and Adamas NDs. Both techniques predict that a moderate amount of Fe is present in these materials, however, by ICP, the Fe signal is quite a bit higher for the Adamas ND, while its SIMS signal appears to be weaker. Two possible explanations for these phenomena are that (i) the ITC shows more iron at its surface than in the bulk, i.e, results from a bulk technique are being compared to those from a surface technique, or (ii) the matrices of the two NDs are different enough that their ToF-SIMS signals cannot be well compared. In contrast to the other NDs, the ITC sample showed a considerable amount of aluminum, and the Adamas sample showed a great deal of chromium. In general, the AA samples have the highest concentrations of metal impurities (see Figure 7.9). As noted above, Mo containing ions are present in the negative ion ToF-SIMS spectra of the AA samples which is present at the ca. 10 ppm level in the AA samples and at a much lower level in the ITC and

Adamas materials. As expected, the concentrations of Fe, and Mo decrease as the AA ND is washed, where the most significant drop in concentration occurs after the first wash.

Interestingly, washing causes the concentrations of some elements, including Ni, Ti, and W, to *increase*. These results suggest that the proprietary acid wash may be a source of contamination to the NDs. The radar plot of these ICP results (Figure 7.9) shows similar patterns for the three AA samples, and different patterns for the other two types of NDs. PCA was also performed on the ICP data, where nearly 98% of the variation in the data was captured by the first two PCs. Figure 7.10a represents PCA data on the scores plot of PC 1 vs. PC2 which shows three distinct groups. AA unwashed, double and triple washed samples were in a single group. Adamas and ITC samples were well separated from all the AA samples. The loadings plot of PC1 (see Figure 7.10b) indicates that the ITC sample, and to a smaller extent the Adamas material are rich in Ti, Mn, Ni, K, Ca, Co and Al, therefore having negative scores on PC1. From this analysis, the AA samples seem to be rich in W and Mo. The loadings on PC 2 reveal that the Adamas sample is separated from the AA and ITC samples due to higher amounts of Fe, Cr and Cu (see Figure 7.10c). All the samples (spectra) fell within 95 % confidence limits on the plot of the Q residuals vs. Hotelling T^2 for this data.

The limits of detection (LOD) were automatically calculated by the software using background equivalent concentrations. The concentrations of the 12 elements listed in Table 7.6 are greater than their respective quantitation limits (see Table 7.7). Because we were performing direct infusion of nanodiamond slurries into the ICP-MS, there were no sample preparation steps involved, e.g., acid digestion. Therefore no significant loss of elements was expected. We also performed spike and recovery determinations under the same operating conditions while spiking

the nanodiamond samples with 100 ppb of Zn. The different samples showed recoveries of 93 - 98 % (see Table 7.8).

Table 7.6 Amounts of various metal impurities (ppm) with standard deviations (S.D.) in nanodiamond samples quantified via ICP-MS.

Sample	AA 50 nm unwashed		AA 50 nm double washed		AA 50 nm triple washed		ITC 50 nm		Adamas 5 nm	
	ppm	S.D.	ppm	S.D.	ppm	S.D.	ppm	S.D.	ppm	S.D.
Aluminum (Al)	0.85	0.03	0.72	0.07	0.93	0.02	15.54	0.38	0.852	0.002
Chromium (Cr)	1.87	0.07	1.19	0.03	1.31	0.02	0.25	0.01	38.9	0.6
Copper (Cu)	0.59	0.02	0.343	0.005	0.393	0.003	0.048	0.006	1.32	0.07
Iron (Fe)	52.0	0.9	11.7	0.5	10.13	0.16	4.5	0.3	22.3	0.5
Manganese (Mn)	0.25	0.01	0.192	0.005	0.24	0.01	1.616	0.009	0.163	0.004
Molybdenum (Mo)	11.09	0.15	7.49	0.03	5.59	0.02	0.139	0.009	0.0435	0.0017
Nickel (Ni)	1.11	0.02	1.34	0.02	1.86	0.02	0.579	0.009	0.20	0.01
Potassium (K)	6.15	0.59	6.98	0.16	7.39	0.46	6.47	0.13	6.31	0.19
Titanium (Ti)	0.79	0.07	1.15	0.02	1.79	0.11	3.01	0.06	0.42	0.02
Tungsten (W)	832.9	4.2	895.2	17.3	964.2	9.4	9.14	0.05	6.95	0.04
Calcium (Ca)	10.2	0.48	8.8	1.13	8.94	0.89	8.07	0.26	8.32	0.47
Cobalt (Co)	0.54	0.02	0.34	0.01	0.419	0.009	0.18	0.01	0.151	0.006

Table 7.7 Limits of Quantitation (LOQ) for elements detected via ICP-MS

Elements	Limit of Quantitation (ppb)
Aluminum (Al)	0.157
Potassium (K)	4.40
Calcium (Ca)	5.71
Titanium (Ti)	0.089
Chromium (Cr)	0.068
Manganese (Mn)	0.007
Iron (Fe)	1.69
Cobalt (Co)	0.013
Nickel (Ni)	0.049
Copper (Cu)	0.030
Molybdenum (Mo)	0.029
Tungsten (W)	0.001

Table 7.8 Spike and recovery data for Zinc using same experimental conditions. The LOQ for Zinc was 1.13 ppb.

Sample	Recovery (ppb)
AA 50 nm unwashed	97.48 ± 3.03
AA 50 nm double washed	97.97 ± 5.25
AA 50 nm triple washed	92.46 ± 0.75
ITC 50 nm	97.65 ± 4.72
Adamas 5 nm	97.28 ± 1.78

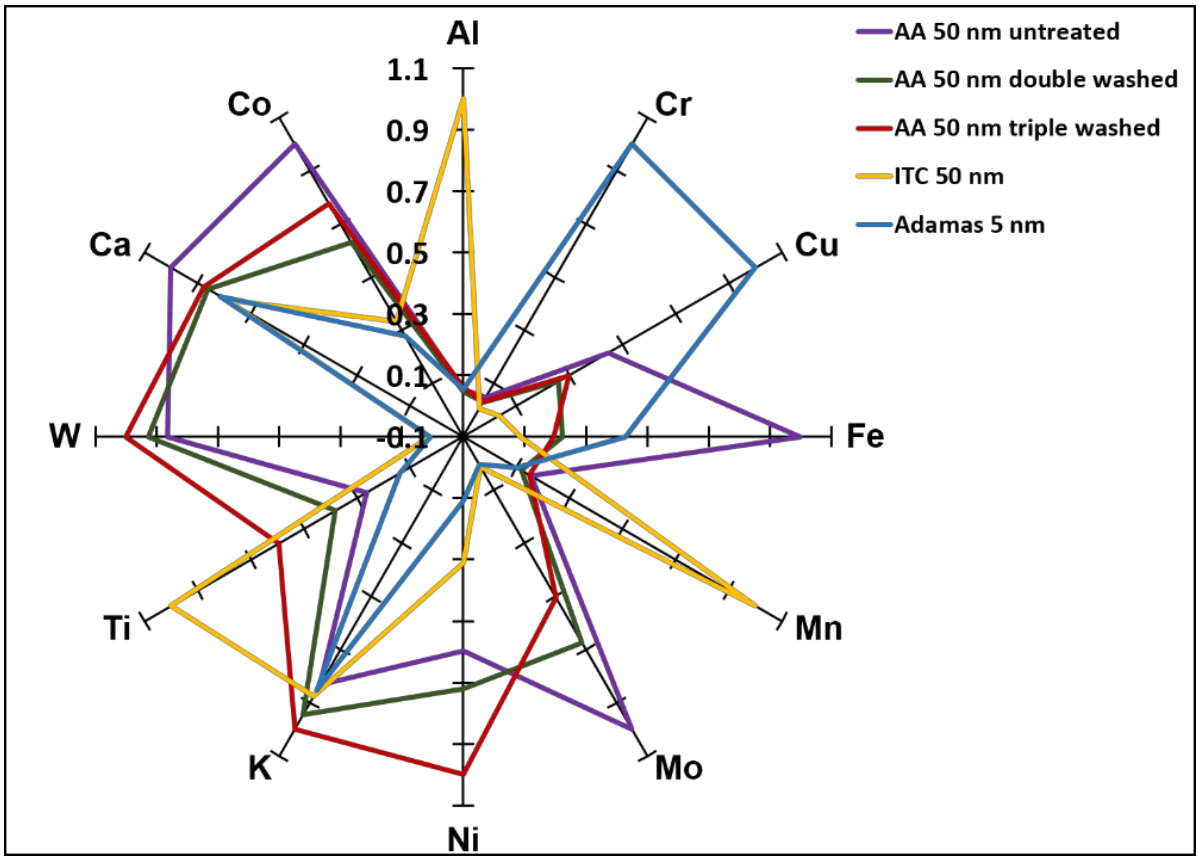


Figure 7.9 Radar plotshowing normalized, linearly scaled quantities of various metal impurities in the nanodiamond samples, as determined by ICP-MS. A value of unity in this figure corresponds to 15.54 ppm for Al, 38.9 ppm for Cr, 1.32 ppm for Cu, 52.0 ppm for Fe, 1.616 ppm for Mn, 11.09 ppm for Mo, 1.86 ppm for Ni, 7.39 ppm for K, 3.01 ppm for Ti and 964.2 ppm for W, 10.2 ppm for Ca, and 0.54 ppm for Co.

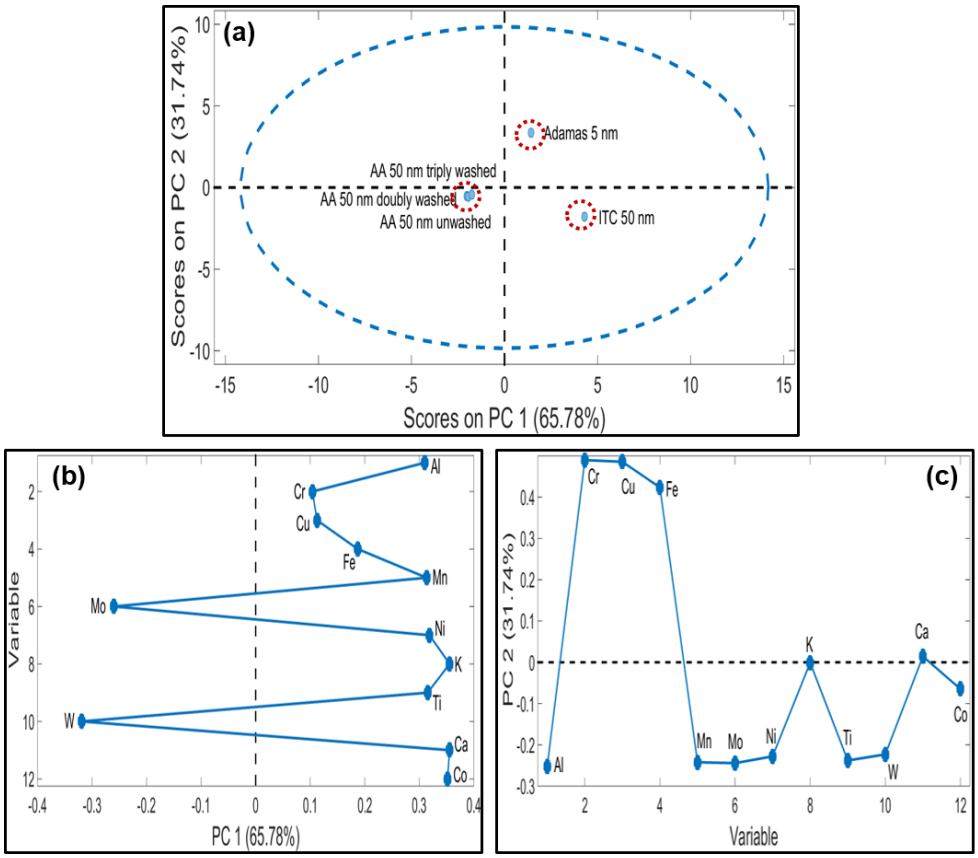


Figure 7.10 PCA of ICP data (except tungsten) from nanodiamond samples. (a) Scores plot, Loadings on PC 1 (b) and PC 2 (c).

7.4.4 DRIFT

Diffuse reflectance infrared Fourier transform (DRIFT) spectroscopy was performed on all the ND samples to understand the organic functional groups on/in them. The XPS analyses of these materials (vide supra) indicated the presence of oxidized carbon. Consistent with these results, DRIFT (see Figure 7.11) showed absorption bands attributable to –OH stretches from alcohols and/or carboxylic acids at ca. 2500 – 3600 cm^{-1}, stretches attributable to C=O groups at around 1750 cm^{-1}, and less well defined C-O stretches at 1000 – 1250 cm^{-1}.[33, 42, 43, 58, 59] The spectra also showed peaks at approximately 1630 cm^{-1}, which may be due to C=C stretches from aromatic carbon.[43] These latter stretches would be consistent with the graphitic carbon observed by TEM (vide infra). This understanding is useful because complex surface chemistries like those suggested by this analysis should contribute to band broadening in HPLC.

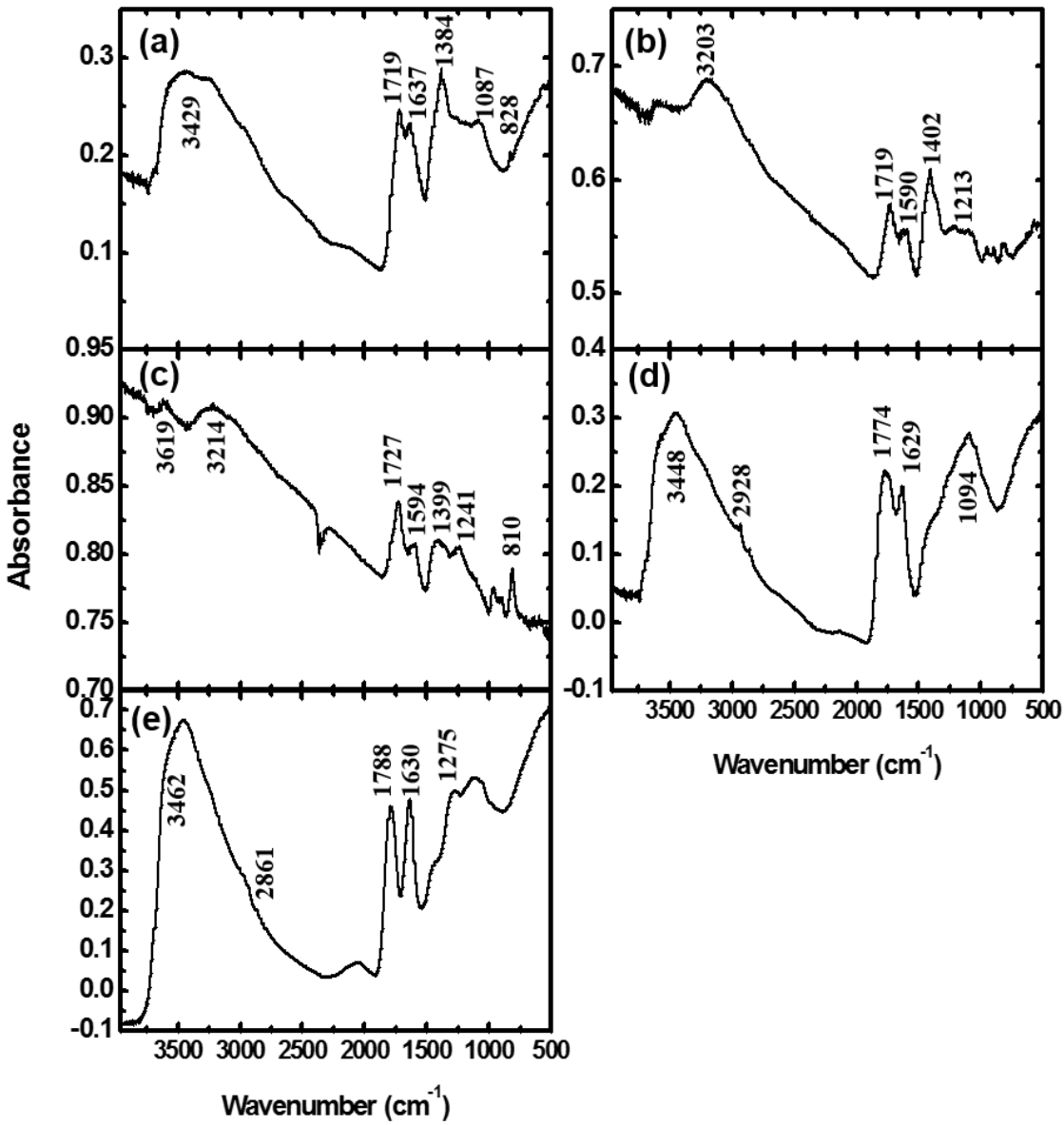


Figure 7.11 FTIR spectra of (a) AA 50 nm unwashed, (b) AA 50 nm double washed, (c)AA 50 nm triple washed, (d) ITC 50 nm and (e) Adamas 5 nm samples.

7.4.5 XRD

Figure 7.12a-e shows the XRD patterns of the AA unwashed 50 nm, AA 50 nm double washed, AA 50 nm triple washed, ITC 50 nm, and Adamas 5 nm samples, respectively. The standard pattern for diamond included for comparison with each sample confirms the presence of crystalline diamond (sp^3 carbon) in the samples. A small amount of graphite (sp^2 carbon) was also found in each of the samples (see Figure 7.13). In the 50 nm AA and ITC samples, the relative amounts of crystalline diamond and crystalline graphite could be quantified through Rietveld refinement. Accordingly, Table 7.9 shows the relative percentages of diamond and graphite in each of the AA and ITC samples. The XRD pattern for each sample was also analyzed quantitatively via profile fitting to gain insight into the size and crystallinity of each sample. Both small crystallite sizes and strain in the crystal lattice (microstrain) tend to broaden diffraction peaks. Fortunately, these effects manifest different angular dependencies, and Williamson-Hall plots (which combine the Scherrer formula for size broadening and the Stokes and Wilson equation for strain broadening) provide a method for deconvolving the contributions of size and strain to the peak broadening. Table 7.9 also gives the size and strain contributions for the different samples under study. The strain contributions were negligible compared to the size contributions, but the plots were greatly improved by including both. As expected from the inverse relationship between crystallite size and peak width, the Adamas samples showed broader peaks (see Figure 7.12e) than the AA and ITC samples (see Figures 7.12a-d). However, for all samples, the XRD size estimates are noticeably smaller than the anticipated 50 nm and 5 nm sizes. This means that the coherence length of the crystalline lattice is smaller than the particle size, which indicates a degree of imperfection in the crystallinity of the samples. Because the microstrain parameters are so small, this may indicate multiple domains per particle.

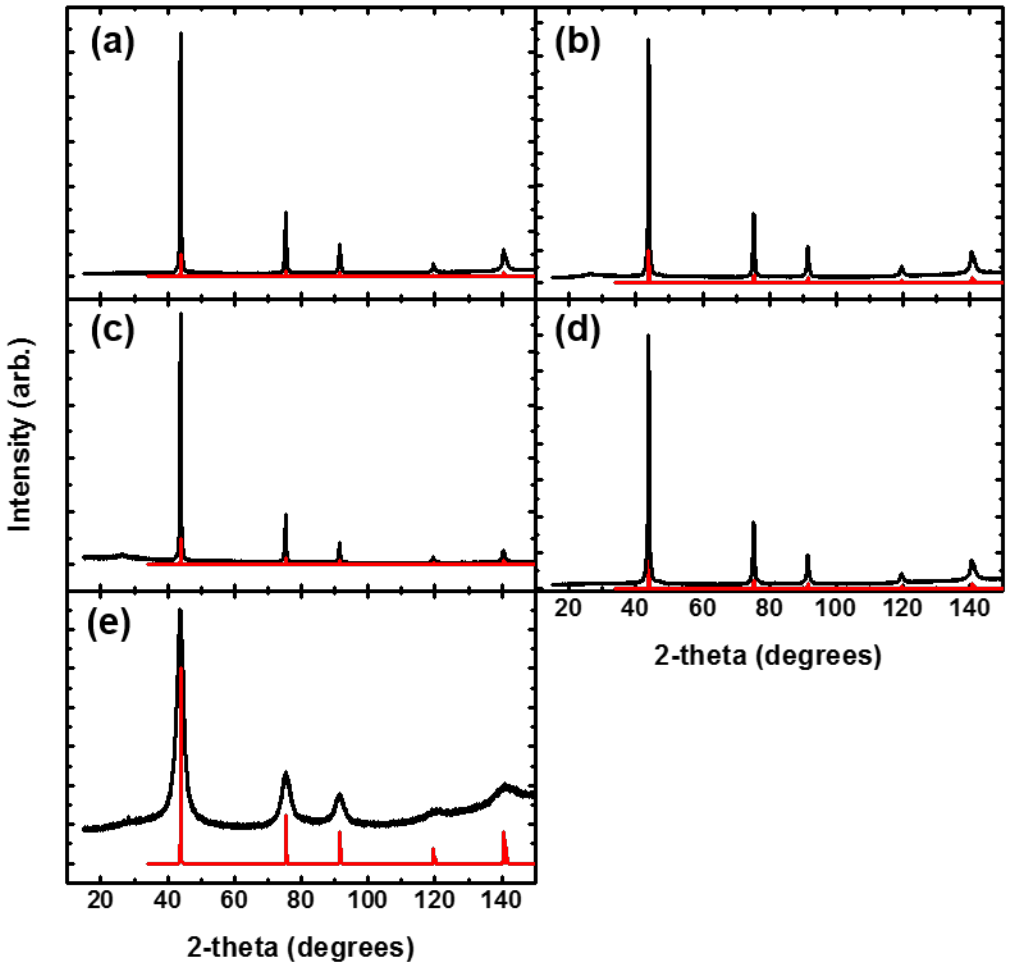


Figure 7.12 XRD diffraction patterns of nanodiamond samples: (a) AA 50 nm unwashed, (b) AA 50 nm double washed, (c) AA 50 nm triple washed, (d) ITC 50 nm, and (e) Adamas 5 nm. Patterns in red are the diamond reference.

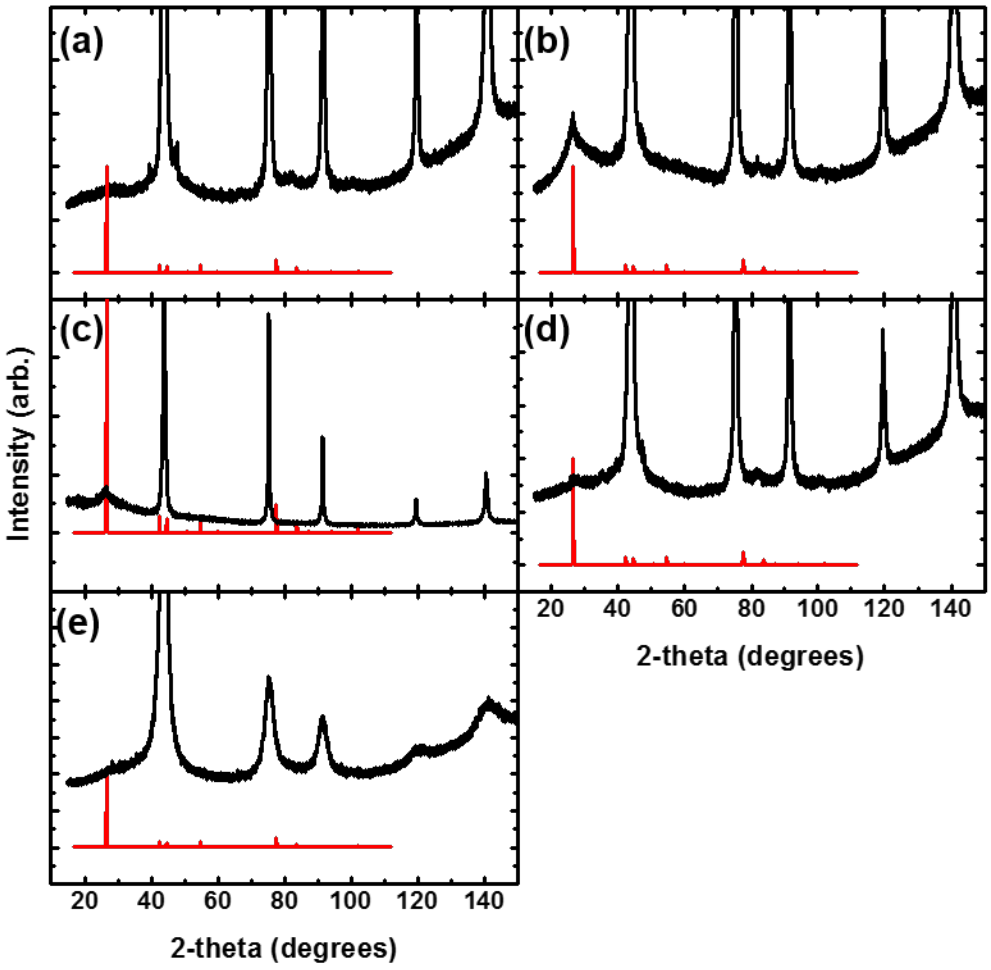


Figure 7.13 XRD diffraction patterns of nanodiamond samples: (a) AA 50 nm unwashed, (b) AA 50 nm double washed, (c) AA 50 nm triple washed, (d) ITC 50 nm, and (e) Adamas 5 nm. Patterns in red are the graphite reference.

In generating profile fits of the XRD patterns for quantitative analysis, we encountered some noteworthy challenges. First, the peak shapes of all of the nanodiamond samples were unusual; they were not initially well modeled by pseudo-Voigt functions (a combination of Gaussian and Lorentzian functions) typically used to model the shape of diffraction peaks. In the Adamas 5 nm sample, large undulations in the baseline around 45 and 80° 2θ further complicated the peak shape analysis. We attributed these undulations to amorphous graphite in the samples, and by including very broad 'amorphous' peaks at these locations we greatly improved the profile fits, though the peak shapes were still unusually Lorentzian. For the other larger nanodiamond samples, the Lorentzian shape of the peaks was even more pronounced. After much consideration, we concluded that these extreme Lorentzian peak shapes could be due to a bimodal distribution of crystallite sizes; it appeared that each peak consisted of contributions from a group of smaller particles (creating a broad base at each peak) and a group of larger particles (creating an unusually tall and sharp maximum for the breadth of the base). Accordingly, the profile fits for all the AA and ITC samples were markedly improved by using 2 sets of peaks to model the data. In essence, each diffraction peak was modeled as a combination of one tall/sharp peak and one short/broad peak constrained to be centered at the same angle. This approach did not drastically alter the sizes estimated for the larger particles (compared to those obtained from the single peak model), but it did drastically improve the profile fit. TEM analyses later confirmed the bimodal size distribution of these samples, validating the models used in the XRD analyses.

Table 7.9 Size and strain contributions, and relative sp^2 and sp^3 content of nanodiamonds calculated via XRD.

Sample	Size (nm)	Strain (%)	sp^3 content (diamond) %	sp^2 content (graphite) %
AA 50 nm unwashed	27.0(7)	0.00(4)	99.4	0.6
AA 50 nm double washed	21.0(3)	0.00(4)	92.7	7.3
AA 50 nm triple washed	25.0(2)	0.00(2)	84.7	15.3
ITC 50 nm	19.0(3)	0.00(5)	98.8	1.2
Adamas 5 nm	2.7(1)	0.0(5)	-	-

Note: The relative diamond (sp^3) and graphite (sp^2) ratios were calculated using Rietveld analysis. The Rietveld analysis couldn't be performed on the Adamas sample owing to its smaller size.

7.4.6 TEM

TEM is an important tool for characterizing nanoparticles. In particular, it provides information about the crystallinity, shapes, and sizes of nanomaterials. A significant disadvantage of TEM is the small sample size analyzed, which might call into question the generality and interpretation of the results. Therefore, in the present study efforts were made to apply very homogeneous sample slurries to TEM grids, after which multiple spots were analyzed. Representative micrographs and diffraction patterns are shown in Figure 7.14. The first row in Figure 7.14 shows TEM diffraction patterns, which in all cases, according to a published report from Gruen et al. correspond to crystalline diamond.[60] The Adamas sample shows hints of amorphous carbon. The micrographs in the second row of Figure 7.14 show a wide size distribution of the nanodiamonds in all the AA samples. In particular, there are 'primary' nanodiamond particles (~ 5 nm), particles that are a step larger than these (ca. 10 – 50 nm), and even some very large particles (ca. 100 nm). These results help to validate the hypothesis from XRD of a bimodal distribution in the nanodiamond samples. TEM indicates that the ITC sample contains very irregularly shaped nanodiamond with sharp edges, and significant amorphous contributions. The Adamas material appears as stacks of flat amorphous sheets, with probable embedded crystallites. The Adamas sample will be discussed in greater detail in the EELS section (vide infra). In order to examine the crystallites more closely, high-resolution TEM micrographs were obtained at ~620k magnification. This view revealed a prominent amorphous covering around the AA unwashed and double washed samples. The magnitude of this amorphous sheath was less for the AA triple washed material. It is important to emphasize again that these micrographs are representative of the general pattern that we observed with many images from the respective samples. The ITC sample appeared to be very amorphous with small

patches of crystallites embedded in the amorphous bulk material. The Adamas sample showed prominent graphite chains running through the sample with d-spacings of 0.1840 Å, which are very close to the reported d-spacing of the 102 plane of graphite (0.1799 Å). For all the samples, d-spacings for 111, 220 and 311 planes were recorded, which are well within 4 % of the reported d-spacing values (by Gruen et al.) for the corresponding planes of crystalline diamond (see Table 7.10). The presence of graphitic carbon in a nanodiamond particle has implications for HPLC because graphitic carbon, e.g., Hypercarb/porous graphitic carbon,[61] often causes peak tailing.

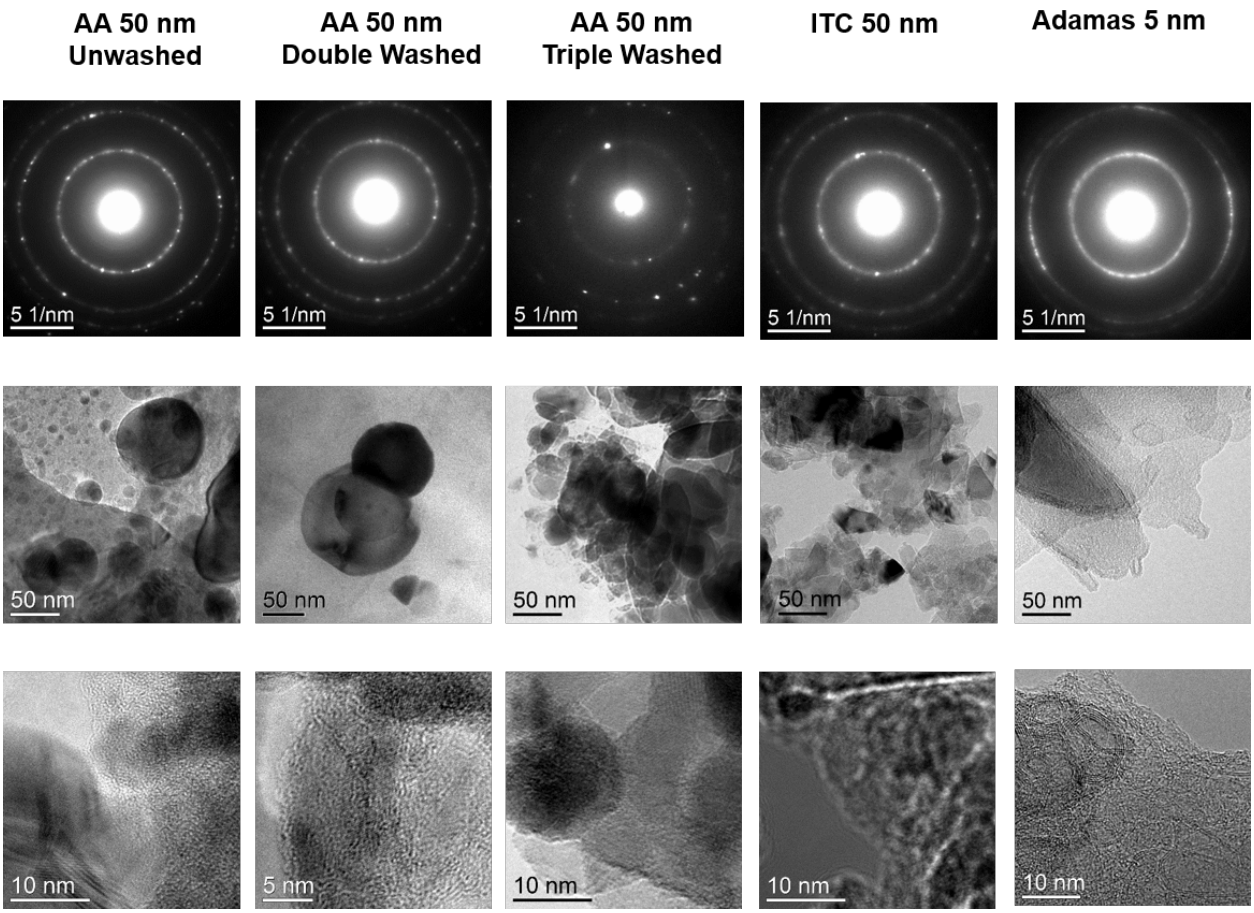


Figure 7.14 Diffraction patterns (first row), TEM micrographs (second row) and high resolution TEM micrographs (third row) of nanodiamond samples.

Table 7.10 Comparing the d-spacings of nanodiamond samples to those reported in the literature.

Planes	{111}	Graphitic {102}	{220}	{311}
Literature values	0.2059	0.1799	0.1261	0.1075
AA unwashed 50 nm	0.2055		0.1258	0.1054
AA double wash 50 nm	0.2074		0.1266	0.1087
AA triple wash 50 nm	0.2139		0.1243	0.1066
ITC 50 nm	0.2075		0.1275	0.1083
Adamas 5 nm	0.2102	0.1840	0.1288	0.1101

7.4.7 EELS

Once TEM and XRD indicated the presence of graphitic and amorphous regions in the samples, it became necessary to quantify the sp^3 carbon content in these samples, which often serves as a measure of the quality of nanodiamonds. Electron energy loss spectroscopy (EELS) is an additional TEM based tool that probes the energy losses of transmitted electrons caused by excitations or energy transitions in the sample. Excitation peaks from the sample are broadened by the incoming energy spread and thus there is great advantage to using an FEG (field emission gun) source. Electrons lost in inner shell ionization events can transition to empty electronic states in a material (subject to appropriate selection rules) with a corresponding loss of energy from the beam of electrons. Thus, the energy loss spectrum for a given inner shell transition maps the local density of empty states in the sample. The empty state energies are a fingerprint of the bonding configuration and thus, for carbon, the hybridization. In carbon, the $1s - \pi^*$ and $1s - \sigma^*$ transitions are separated by several eV. The $1s - \pi^*$ transitions are only available in sp^2 material, and both the sp^2 and sp^3 materials show $1s - \sigma^*$ transitions. Thus, while the presence of π^* transitions indicates some sp^2 material, it is the relative size of the π^* and σ^* transitions that indicate the sp^2 and sp^3 fraction.

Various models are used to quantify the sp^3 and sp^2 content of carbonaceous materials from EELS data.[62-64] We used a *"two-window method"* where the windows are centrally positioned on the major peaks (TWC), namely the $1s - \pi^*$ and $1s - \sigma^*$ peaks.[63] Following previously published reports, we used two models to calculate the sp^3 composition of the nanodiamond samples being studied.[62, 64] The Berger *et al.* approach was named *Model 1* and the Cuomo *et al.* approach was *Model 2*. For both models, an *integral ratio* (I_n) is derived from the EELS data, which is a ratio between the counts in the $1s - \pi^*$ transition peak to counts in another

region of the data. This integral ratio would be zero for a pure sp^3 material and at a maximum for a pure sp^2 material. Thus, the sp^2 fraction (F_{sp2}) is given by a comparison of the integral ratios of the unknown sample (n = u, unknown sample) to that of a 100 % sp^2 carbon reference sample (n = r, reference sample).

For our calculations, energy alignment was made by shifting all the 1s – π^* peaks to be centered at 285 eV. The reference sample was the amorphous support carbon film with the reference spectra collected under similar conditions as the unknowns. The background was subtracted using a standard approach of interpolating from a pre-edge region of the spectra. The zero loss peak showed that the samples were sufficiently thin, hence, no efforts were made to deconvolute the multiple scattering from the carbon edge. We understand that the sp^2 contribution in the nanodiamond samples will be overestimated because of the underlying carbon support film. We never saw sp^3 content greater than ~ 90 %. Thus, we estimate that our EELS measurements were overestimating the sp^2 content by <10 % for the nanodiamond samples. The approach and conditions for all measurements were kept constant and hence the results should provide reasonable insight into our samples. The EELS spectra of all the nanodiamond samples and sp^2 reference sample are given in Figure 7.15.

Model 1: In Model 1, the integral ratio I_n is the area under 1s – π^* (281-287.5 eV) divided by the sum of areas under 1s – π^* and 1s – σ^* (281-310 eV).

$$I_n = \frac{I_{\pi n}*}{I_{\pi n}* + I_{\sigma n}*} \tag{7.1}$$

The sp^2 fraction (F_{sp2}) is simply the ratio of the unknown and reference integral ratios.

$$F_{sp}2 = \frac{I_u}{I_r} \tag{7.2}$$

The sp^3 fraction is 1 minus the sp^2 fraction.

Model 2: In Model 2, the integral ratio and sp^2 fraction formulas change. The integral ratio is calculated as the ratio of the area under the 1s – π^* (283-287 eV) peak to that of the 1s – σ^* (288-303 eV) peak.

$$I_n = \frac{I_{\pi n}\,*}{I_{\sigma n}\,*} \tag{7.3}$$

The sp^2 fraction is given by the following equation,

$$F_{sp}2 = \frac{4\frac{I_u}{I_r}}{3 + \frac{I_u}{I_r}} \tag{7.4}$$

where Equation 7.4 is based on the theoretical ratio of π^* to σ^* orbitals in sp^2 and sp^3 carbon.[64] Results are summarized in Table 7.11. The standard deviations of the values obtained using Models 1 and 2 were within 1.3 %. This speaks to the robustness of the models used. There is an increase in the sp^3 content of the nanodiamonds as we go from the AA 50 nm unwashed to the double and triply washed samples. This might be attributed to the removal of the amorphous covering on the nanodiamond particles by the leaching step. The ITC 50 nm sample had the least sp^3 contribution: ca. 36 %. The Adamas 5 nm sample had ca. 78 % sp^3 content. This is somewhat contradictory to TEM analysis where the Adamas 5 nm micrographs appeared as distinct

contours of amorphous stacked flat sheets (see Figure 7.14). We believe that the Adamas sample contains large numbers of defects with tiny crystallites embedded in the bulk. Therefore, while the TEM images look amorphous, EELS analysis shows crystalline sp^3 content in the Adamas samples.

The results from EELS include crystalline (sp^3) *and* amorphous (sp^2) carbon, and are different from the ratios obtained via XRD Rietveld refinement analysis (see Table 7.9), i.e., the values reported in Table 7.9 are relative ratios of crystalline diamond (sp^3 carbon) and crystalline graphite (sp^2 carbon). For example, HR-TEM images indicated that the AA unwashed and double washed samples were covered in an amorphous sheath. This sheath would not be considered in the XRD analysis in part because it would appear as a broad diffusive peak and would be difficult to analyze. On the other hand, for the AA 50 nm triple washed sample, which had the thinnest covering of the amorphous sheath, the diamond (sp^3) content via XRD (~ 85 %) was in reasonable agreement with the EELS determination (~ 88 %). Thus, EELS should provide more realistic sp^2/sp^3 values for our nanodiamond samples. As noted, XPS analysis showed no shake-up peaks in the C 1s narrow scans of the nanodiamonds, i.e., their sp^2 contributions should be less than 30 %. However, this generalization probably only holds for a homogeneous material, not for a highly heterogeneous material like our nanodiamonds – it is possible that any shake-up peaks from sp^2/aromatic carbon will be too broad and diffuse to be discernable. This could explain why the ITC 50 nm sample had a ca. 63 % sp^2 contribution but did not show any shake-up peak in its XPS narrow scan.

The presence of sp^2 carbon can lead to unwanted interactions and hinder the performance of nanodiamond-based HPLC particles. Therefore, EELS serves as a vital tool to determine the quality of the nanodiamonds being used.

Table 7.11 sp^3 contribution in the nanodiamonds under study as calculated by EELS.

Sample	Model 1	Model 2
AA 50 nm unwashed	73.56	71.67
AA 50 nm double washed	84.97	84.13
AA 50 nm triple washed	88.16	87.29
ITC 50 nm	37.41	35.92
Adamas 5 nm	77.9	77.60

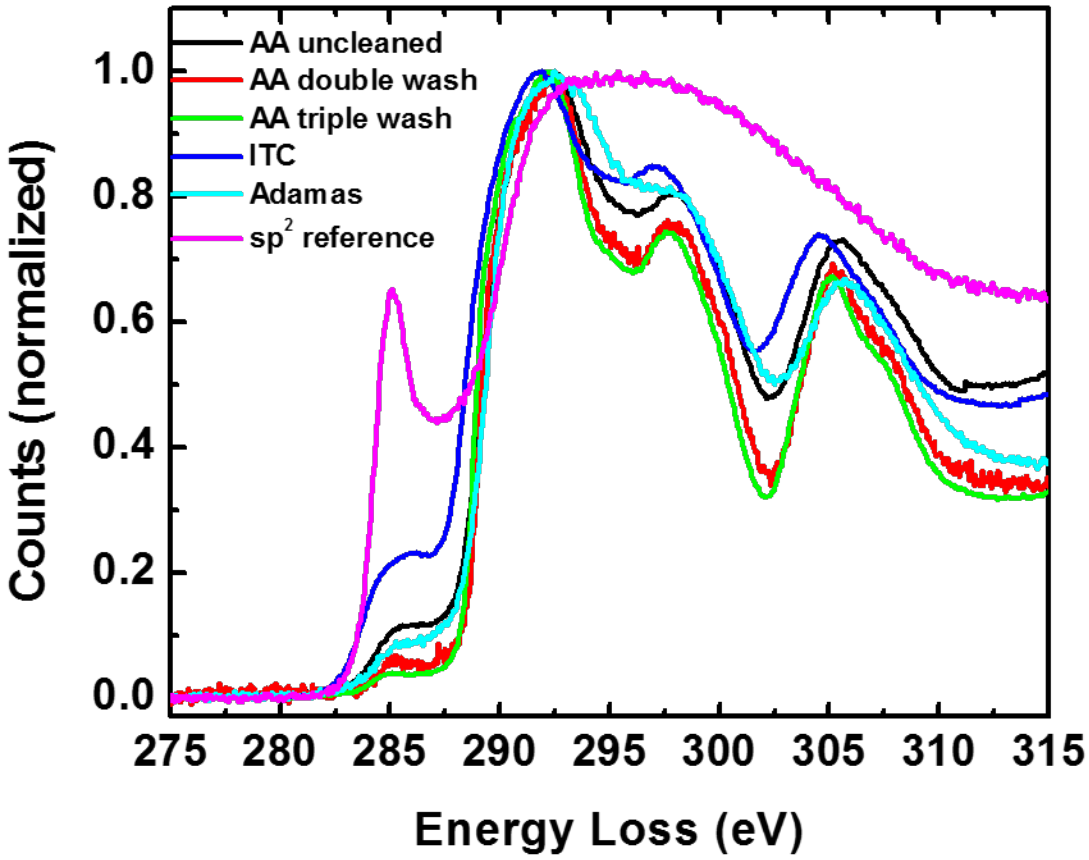


Figure 7.15 EELS spectra of all five nanodiamond and sp^2 reference samples.

7.4.8 BET

BET analysis was performed on all of the powdered nanodiamond samples, the results of which are summarized in Table 7.12. Based on the particle sizes provided by XRD, theoretical surface areas of the nanodiamonds were calculated using the following equation [10]

$$S = \frac{6}{r\rho} \tag{7.5}$$

where S is the surface area of the nanodiamonds, r is the diameter of the diamond particles, and ρ is the density of the nanodiamonds. In these calculations, the density of the nanodiamonds was taken as 3.51 g/cm^3.[10] The above-mentioned equation does not take into account any aggregation of nanodiamond particles. The surface areas obtained from BET were within ca. 20 % of theoretical surface areas, except for the Adamas 5 nm particles. This result suggests a high degree of aggregation for the Adamas material, as was expected because of its small size.

Table 7.12 Theoretical surface areas of nanodiamonds being compared to the surface areas calculated via BET.

Sample	Surface area (m^2/g)	Theoretical surface area (m^2/g)
AA 50 nm unwashed	50.24	63.3
AA 50 nm double washed	70.65	81.4
AA 50 nm triple washed	54.27	68.4
ITC 50 nm	110.51	89.96
Adamas 5 nm	214.43	633.11

7.4.9 PSD

PSD analysis was performed on the undiluted nanodiamond slurries, as they were received from the vendors. The mean particle sizes were determined to be 40.1 nm, 47.1 nm, 50.2 nm, 43.3 nm, and 16.6 nm for AA 50 nm unwashed, AA 50 nm double washed, AA 50 nm triple washed, ITC 50 nm, and Adamas 5 nm samples, respectively. These mean values are based off of relative weight calculations from the plot of weight of particles vs. their diameters. An elaborate distribution of size ranges of the nanodiamond samples are given in Table 7.13. The results from AA and ITC samples align well with the TEM findings, while the Adamas 5 nm results do not align well with the TEM and XRD findings. For Adamas, XRD gives a crystallite size of ~ 2.7 nm; TEM shows stacked sheets, possibly with large numbers of defects, and hence appearing to be amorphous in the micrographs; PSD gives an average particle size of 16.6 nm. PSD does pick up ~ 0.13 % of Adamas particles in the size range of 155 – 495 nm (which is a very small fraction). It may be that the smaller Adamas particles didn't adhere to the TEM mesh, and/or that bigger particles selectively adhered to the TEM mesh. These kinds of somewhat contradictory results are expected in the multi-instrument characterization of a material, where each technique will shed light on different properties of a sample. For example, in the particular case of the Adamas sample, XRD measures a crystallite size (2.7 nm) that doesn't necessarily represent the particle size. That is, the crystallites may be embedded in the sheets observed in TEM, and/or surrounded by amorphous material and/or have high defects. However, PSD, contrary to these findings, gave an average particle size of 16.6 nm for the Adamas samples, with a tiny fraction of larger particles (probably the stacked sheets observed by TEM). Without the PSD measurements, we would have concluded that the Adamas particles were entirely present as crystallites embedded in stacked sheets. Now it appears that the Adamas sample is composed

almost entirely (~ 98.8 %) of particles in the size range of 13.2 - 27.4 nm (see Table 7.13), with a tiny fraction of them in large stacked sheets, possibly containing embedded, defective crystallites.

Table 7.13 PSD of the five nanodiamond samples under study. The columns represent the fraction (%) of particles within a specific size range.

Size range (nm)	AA 50 nm Unwashed	AA 50 nm Double Washed	AA 50 nm Triple Washed	ITC 50 nm	Adamas 5 nm
10.0 - 15.0	4.62	2.88	2.44	(13.2 - 19.0 nm) 5.77	29.8
15.0 - 22.0	13.46	8.79	8.05	(19.0 - 27.4 nm) 16.44	70.08
22.0 - 33.0	21.86	16.92	15.38	(27.4 - 39.4 nm) 24.94	0
33.0 - 48.0	28.97	25.88	24.41	(39.4 - 56.6 nm) 29.61	0
48.0 - 71.0	24.67	30.91	30.43	(56.6 - 81.3 nm) 20.31	0
71.0 - 105	6.35	14.18	18.02	(81.3 - 116.9 nm) 2.94	0
105 - 155	0.07	0.42	1.21	0	0
155 - 229	0	0.02	0.05	0	(155-227 nm) 0.01
229 - 338	0	0.01	0.01	0	(227-336 nm) 0.05
336 - 495	0	0	0	0	0.07

7.5 Conclusions

The results from this study are summarized in the following table:

Technique	AA 50 nm unwashed	AA 50 nm doubly washed	AA 50 nm triply washed	ITC 50 nm	Adamas 5 nm
XPS	All samples showed significant C 1s and O 1s peaks, implying oxidized carbon at their surfaces Low levels of nitrogen and silicon present in all the samples				
	Tungsten present in all AA samples; narrow scans show oxidized W species				Sodium present
ToF-SIMS*	Na^+, Al^+, Si^+, K^+, Cr^+, Fe^+	Na^+, Al^+, Si^+, K^+, Ca^+, Cr^+, Fe^+	Na^+, Al^+, Si^+, K^+, Ca^+, Cr^+, Fe^+	Na^+, Mg^+, Al^+, Si^+, K^+, Ca^+, Ti^+, Cr^+, Mn^+, Fe^+	Na^+, Al^+, K^+, Cr^+, Fe^+
	No W+ signal present. Negative mode SIMS showed peak envelopes corresponding to WO_3^-/WO_3H^-, and WO_4^-/WO_4H^-. MoO_3^- also present in all AA samples				
ICP-MS	~ 900 ppm of tungsten present in all AA samples Concentrations of Ni, Ti, and W increase with washing Concentrations of Fe, and Mo decrease with washing			Maximum aluminum (~16 ppm)	Rich in chromium (~39 ppm) and iron (~22 ppm)
	Iron 52 ppm	Iron ~12 ppm	Iron ~10 ppm		
DRIFT	Complementary to XPS results, DRIFT shows stretches corresponding to –OH, C=O, and C-O groups Presence of C=C stretches hint at graphitic carbon				
XRD	27 nm crystallite size	21 nm crystallite size	25 nm crystallite size	19 crystallite size	2.7 crystallite size
	Possible bimodal distribution of nanodiamonds				
	All samples showed broader peaks corresponding to graphitic peaks				
TEM	Wide size distribution confirmed hypothesis from XRD of bimodal size distribution; Amorphous sheath/covering present in all AA samples that seems to decrease with washing			Irregular shaped, sharp edges	Sheet-like structure; 102 plane of graphite present graphitic chains present in the sample
	Diffraction pattern of all samples show 111, 220, and 311 planes of crystalline diamond. Hints of amorphous carbon in all of the samples				
EELS	~74 sp^3 carbon	~85 sp^3 carbon	~88 sp^3 carbon	~38 sp^3 carbon	~78 sp^3 carbon
BET	The theoretical surface area of all samples were compared with the surface areas obtained by BET. The results indicated excessive agglomeration of the Adamas sample because of their miniature size (~ 2.7 nm via XRD)				
PSD (Average particle size)	40.1 nm	47.1 nm	50.2 nm	43.3 nm	16.6 nm

*Comparison based on raw peak area; red color denotes greater peak areas as compared to others

As is often the case in an extensive, multi-instrument analysis of a material, we have uncovered some unexpected compositions/properties/comparisons. Clearly these insights were gained as a result of a combination of analytical methods. No single technique can determine all of the properties of a material. Therefore, thorough material characterization consists of carefully choosing a suite of techniques that probe different aspects of a material, which is followed by appropriate data analysis and interpretation to draw reasonable conclusions. Based on the number and scope of the analytical techniques employed in the current study, we believe that this is the most comprehensive nanodiamond characterization study undertaken to date and that it can serve as an exemplary report for the community. The findings from this study are aimed at improving nanodiamond-based core-shell particles for HPLC.

7.6 Acknowledgement

This work was supported financially by Diamond Analytics, a US Synthetic Company, Orem, UT, USA. We also acknowledge the Department of Chemistry and Biochemistry and College of Physical and Mathematical Sciences at Brigham Young University for their support of this work. We would like to thank Cody C. Frisby (US Synthetic Corporation) and Anna Nielsen (Brigham Young University) for analyzing the ICP samples.

Author Contributions: B.S. initially showed that the nanodiamonds are exposed in the shells of diamond-based HPLC particles, which eventually led M.R.L. and B.S. to conceive the idea of characterizing these nanodiamond samples to improve the performance of particles containing them. B.S. collected data from XPS, DRIFT, TEM, and EELS, and analyzed data from all the techniques mentioned in the manuscript except for XRD data. B.S. came up with different methods of sample preparation of ND slurries to aid their analyses by various

techniques. D.J., H.J. and S.S. collected BET, PSD, and XRD data, respectively. J.F., R.V., and P.F. provided their insights and technical assistance on TEM, EELS, and ICP-MS sections, respectively. A.D. provided technical inputs to the project. B.S. managed the entire project, working with all of the collaborators to bring it together. M.R.L. was the principal investigator and helped with the data analysis along with B.S. M.R.L. and B.S. wrote the manuscript.

7.7 Conflict of Interest

Some of the authors on this paper stand to benefit financially from the sale of ND-containing core-shell HPLC particles that are produced and marketed by Diamond Analytics.

7.8 References

1. D. R. Baer, M. H. Engelhard, G. E. Johnson, J. Laskin, J. Lai, K. Mueller, P. Munusamy, S. Thevuthasan, H. Wang, N. Washton, A. Elder, B. L. Baisch, A. Karakoti, S. V. N. T. Kuchibhatla and D. Moon, J. Vac. Sci. Technol., A **31** (5), 050820 (2013).

2. B. E. Erickson, in *Chem Eng News* (2008), Vol. 86, pp. 25-26.

3. S. V. N. T. Kuchibhatla, A. S. Karakoti, D. R. Baer, S. Samudrala, M. H. Engelhard, J. E. Amonette, S. Thevuthasan and S. Seal, The Journal of Physical Chemistry C **116** (26), 14108-14114 (2012).

4. D. R. Baer, J. E. Amonette, M. H. Engelhard, D. J. Gaspar, A. S. Karakoti, S. Kuchibhatla, P. Nachimuthu, J. T. Nurmi, Y. Qiang, V. Sarathy, S. Seal, A. Sharma, P. G. Tratnyek and C. M. Wang, Surf Interface Anal **40** (3-4), 529-537 (2008).

5. D. R. Baer, D. J. Gaspar, P. Nachimuthu, S. D. Techane and D. G. Castner, Anal Bioanal Chem **396** (3), 983-1002 (2010).

6. D. S. Jensen, S. S. Kanyal, N. Madaan, J. M. Hancock, A. E. Dadson, M. A. Vail, R. Vanfleet, V. Shutthanandan, Z. H. Zhu, M. H. Engelhard and M. R. Linford, Surf Interface Anal **45** (8), 1273-1282 (2013).

7. H. Wang, B. M. Lunt, R. J. Gates, M. C. Asplund, V. Shutthanandan, R. C. Davis and M. R. Linford, ACS Appl. Mater. Interfaces **5** (17), 8407-8413 (2013).

8. P. N. Nesterenko and O. N. Fedyanina, J Chromatogr A **1217** (4), 498-505 (2010).

9. P. N. Nesterenko, O. N. Fedyanina and Y. V. Volgin, Analyst **132** (5), 403-405 (2007).

10. P. N. Nesterenko and P. R. Haddad, Anal Bioanal Chem **396** (1), 205-211 (2010).

11. P. N. Nesterenko, O. N. Fedyanina, Y. V. Volgin and P. Jones, J Chromatogr A **1155** (1), 2-7 (2007).

12. O. N. Fedyanina and P. N. Nesterenko, Russ J Phys Chem a+ **84** (3), 476-480 (2010).

13. G. Yushin, E. N. Hoffman, M. W. Barsoum, Y. Gogotsi, C. A. Howell, S. R. Sandeman, G. J. Phlllips, A. W. Lloyd and S. V. Mikhalovsky, Biomaterials **27** (34), 5755-5762 (2006).

14. G. Saini, D. S. Jensen, L. A. Wiest, M. A. Vail, A. Dadson, M. L. Lee, V. Shutthanandan and M. R. Linford, Anal Chem **82** (11), 4448-4456 (2010).

15. G. Saini, L. A. Wiest, D. Herbert, K. N. Biggs, A. Dadson, M. A. Vail and M. R. Linford, J Chromatogr A **1216** (16), 3587-3593 (2009).

16. V. S. Bondar', I. O. Pozdnyakova and A. P. Puzyr', Phys Solid State+ **46** (4), 758-760 (2004).

17. L. A. Wiest, D. S. Jensen, C. H. Hung, R. E. Olsen, R. C. Davis, M. A. Vail, A. E. Dadson, P. N. Nesterenko and M. R. Linford, Anal Chem **83** (14), 5488-5501 (2011).

18. C.-H. Hung, L. A. Wiest, B. Singh, A. Diwan, M. J. C. Valentim, J. M. Christensen, R. C. Davis, A. J. Miles, D. S. Jensen, M. A. Vail, A. E. Dadson and M. R. Linford, J Sep Sci **36** (24), 3821-3829 (2013).

19. L. A. Wiest, D. S. Jensen, A. J. Miles, A. Dadson and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/%CE%B22-Agonists%20and%20Amphetamines.pdf, 2013).

20. L. A. Wiest, D. S. Jensen, A. J. Miles, A. E. Dadson and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/FlareMixedModeColumn-TriazineHerbicides.pdf, 2013).

21. B. Singh, D. S. Jensen, A. J. Miles, A. E. Dadson and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/Acidic%20Herbicides.pdf, 2013).

22. B. Singh, D. S. Jensen, A. J. Miles, A. Dadson and M. R. Linford, in *Diamond Analytics Application Note: DA 1000-C* (http://diamond-analytics.com/uploads/Retention%20of%20Acidic%20Herbicides%20vs%20pKa%20Values.pdf , 2013).

23. C. H. Hung, A. A. Kazarian, A. Dadson, B. Paull, P. N. Nesterenko and M. R. Linford, in *Diamond Analytics Application Note* (http://diamond-analytics.com/uploads/Diamond%20Analytics%20Application%20Guidelines.pdf, 2013).

24. O. N. Fedyanina and P. N. Nesterenko, Russ J Phys Chem a+ **85** (10), 1773-1777 (2011).

25. D. F. Mitev, A. T. Townsend, B. Paull and P. N. Nesterenko, Carbon **60**, 326-334 (2013).

26. E. Perevedentseva, P. J. Cai, Y. C. Chiu and C. L. Cheng, Langmuir **27** (3), 1085-1091 (2011).

27. M. Baidakova and A. Vul', J Phys D Appl Phys **40** (20), 6300-6311 (2007).

28. W. Yang, O. Auciello, J. E. Butler, W. Cai, J. A. Carlisle, J. E. Gerbi, D. M. Gruen, T. Knickerbocker, T. L. Lasseter, J. N. Russell, L. M. Smith and R. J. Hamers, Nat Mater **1** (4), 253-257 (2002).

29. A. M. Schrand, S. A. C. Hens and O. A. Shenderova, Critical Reviews in Solid State and Materials Sciences **34** (1-2), 18-74 (2009).

30. B. V. Spitsyn, J. L. Davidson, M. N. Gradoboev, T. B. Galushko, N. V. Serebryakova, T. A. Karpukhina, I. I. Kulakova and N. N. Melnik, Diam Relat Mater **15** (2–3), 296-299 (2006).

31. W. S. Yang, O. Auciello, J. E. Butler, W. Cai, J. A. Carlisle, J. Gerbi, D. M. Gruen, T. Knickerbocker, T. L. Lasseter, J. N. Russell, L. M. Smith and R. J. Hamers, Nat Mater **1** (4), 253-257 (2002).

32. G. N. Yushin, S. Osswald, P. VI, G. P. Bogatyreva and Y. Gogotsi, Diam Relat Mater **14** (10), 1721-1729 (2005).

33. S. Osswald, G. Yushin, V. Mochalin, S. O. Kucheyev and Y. Gogotsi, J Am Chem Soc **128** (35), 11635-11642 (2006).

34. O. Shenderova, A. Koscheev, N. Zaripov, I. Petrov, Y. Skryabin, P. Detkov, S. Turner and G. Van Tendeloo, J Phys Chem C **115** (20), 9827-9837 (2011).

35. G. Saini, L. Yang, M. L. Lee, A. Dadson, M. A. Vail and M. R. Linford, Anal Chem **80** (16), 6253-6259 (2008).

36. L. Yang, D. S. Jensen, M. A. Vail, A. Dadson and M. R. Linford, J Chromatogr A **1217** (49), 7621-7629 (2010).

37. L. Yang, M. A. Vail, A. Dadson, M. L. Lee, M. C. Asplund and M. R. Linford, Chem Mater **21** (19), 4359-4365 (2009).

38. O. Shenderova, A. M. Panich, S. Moseenkov, S. C. Hens, V. Kuznetsov and H. M. Vieth, J Phys Chem C **115** (39), 19005-19011 (2011).

39. A. E. Aleksenskii, M. V. Baidakova, A. Y. Vul' and V. I. Siklitskii, Phys Solid State+ **41** (4), 668-671 (1999).

40. V. F. Loktev, V. I. Makal'skii, I. V. Stoyanova, A. V. Kalinkin, V. A. Likholobov and V. N. Mit'kin, Carbon **29** (7), 817-819 (1991).

41. A. Kruger, Y. J. Liang, G. Jarre and J. Stegk, J Mater Chem **16** (24), 2322-2328 (2006).

42. L. C. L. Huang and H. C. Chang, Langmuir **20** (14), 5879-5884 (2004).

43. O. Shenderova, I. Petrov, J. Walsh, V. Grichko, V. Grishko, T. Tyler and G. Cunningham, Diam Relat Mater **15** (11-12), 1799-1803 (2006).

44. V. Gupta, H. Ganegoda, M. H. Engelhard, J. Terry and M. R. Linford, J Chem Educ **91** (2), 232-238 (2013).

45. P. M. A. Sherwood, J. Vac. Sci. Technol., A **14** (3), 1424-1432 (1996).

46. B. V. Crist, Journal of Surface Analysis **4** (3), 428-434 (1998).

47. B. Singh, D. Velázquez, J. Terry and M. R. Linford, J Electron Spectrosc **197** (0), 56-63 (2014).

48. J. Feng, M. Sun, H. Liu, J. Li, X. Liu and S. Jiang, J Chromatogr A **1217** (52), 8079-8086 (2010).

49. W.-L. Dai, M.-H. Qiao and J.-F. Deng, Appl Surf Sci **120** (1–2), 119-124 (1997).

50. X. V. Eynde and P. Bertrand, Surf Interface Anal **25** (11), 878-888 (1997).

51. O. D. Sanni, M. S. Wagner, D. Briggs, D. G. Castner and J. C. Vickerman, Surf Interface Anal **33** (9), 715-728 (2002).

52. L. Yang, R. Bennett, J. Strum, B. Ellsworth, D. Hamilton, M. Tomlinson, R. Wolf, M. Housley, B. Roberts, J. Welsh, B. Jackson, S. Wood, C. Banka, C. Thulin and M. Linford, Anal Bioanal Chem **393** (2), 643-654 (2009).

53. L. Yang, Y. Y. Lua, G. L. Jiang, B. J. Tyler and M. R. Linford, Anal Chem **77** (14), 4654-4661 (2005).

54. L. Pei, G. Jiang, B. J. Tyler, L. L. Baxter and M. R. Linford, Energ Fuel **22** (2), 1059-1072 (2008).

55. D. Lewis, R. J. Cole and P. Weightman, Journal of Physics: Condensed Matter **11** (43), 8431 (1999).

56. H. Wang, B. M. Lunt, R. J. Gates, M. C. Asplund, V. Shutthanandan, R. C. Davis and M. R. Linford, Acs Appl Mater Inter **5** (17), 8407-8413 (2013).

57. C. H. Hung, L. A. Wiest, B. Singh, A. Diwan, M. J. C. Valentim, J. M. Christensen, R. C. Davis, A. J. Miles, D. S. Jensen, M. A. Vail, A. E. Dadson and M. R. Linford, J Sep Sci **36** (24), 3821-3829 (2013).

58. I. I. Kulakova, Phys Solid State+ **46** (4), 636-643 (2004).

59. R. M. Silverstein, F. X. Webster and D. J. Kiemle, *Spectrometric Identification Of Organic Compounds*, Seventh ed. (John Wiley & Sons Inc., 2005).

60. D. M. Gruen, Annu Rev Mater Sci **29**, 211-259 (1999).

61. C. West, C. Elfakir and M. Lafosse, J Chromatogr A **1217** (19), 3201-3216 (2010).

62. S. D. Berger, D. R. McKenzie and P. J. Martin, Philosophical Magazine Letters **57** (6), 285-290 (1988).

63. N. Bernier, F. Bocquet, A. Allouche, W. Saikaly, C. Brosset, J. Thibault and A. Charaï, J Electron Spectrosc **164** (1–3), 34-43 (2008).

64. J. J. Cuomo, J. P. Doyle, J. Bruley and J. C. Liu, Appl Phys Lett **58** (5), 466-468 (1991).

Chapter 8: Conclusions

8.1 Conclusions

My graduate work was focused primarily in two areas: (a) the development of data analysis tools for X-ray photoelectron spectroscopy (XPS) and time-of-flight secondary ion mass spectrometry (ToF-SIMS), and (b) the multi-instrument characterization of materials.

To better analyze XPS narrow scans, I have introduced the equivalent width and autocorrelation width. These approaches have the potential to act as a tool that is less biased/subjective, but also complementary, to traditional peak fitting practices in XPS. Continuing the effort to improve XPS data analysis, I introduced the concept of uniqueness testing to XPS. Uniqueness tests are routinely performed in spectroscopic ellipsometry to confirm the absence of fit parameter correlation. However, to the best of my knowledge, no one has ever used this concept in XPS. I also discuss various reasonable practices for obtaining good peak fits in XPS.

ToF-SIMS can generate enormous amounts of data, and the analysis of this data can be challenging. Therefore, various pattern recognition tools, i.e., chemometrics techniques, such as principal components analysis, cluster analysis, etc., have been used to help deal with this problem. In this dissertation, I have also introduced a new chemometrics technique known as the Information Content (IC). This technique has been derived from previous work by Claude Shannon on information theory. I show the applicability of IC to complex data sets from ToF-SIMS and XPS.

Dr. Linford's group at BYU has worked extensively to manufacture nanodiamond-based particles for high performance liquid chromatography. The second part of my research work involved a comprehensive multi-instrument characterization of these nanodiamonds with the aim

of exploiting our understanding of these materials for product improvement. A diverse suite of relevant analytical tools was employed, including XPS, ToF-SIMS, inductively coupled plasma-mass spectrometry (ICP-MS), scanning electron microscopy (SEM), transmission electron microscopy (TEM), diffuse reflectance infrared Fourier transform spectroscopy (DRIFT), etc.

8.2 Future Work

In this work, I showed application of mathematical functions (the equivalent width (EW) and the autocorrelation width (AW)) for the characterization of XPS narrow scans using four different data sets. Future work in this area should be focused on using more *complex* data sets, both from XPS and also from other techniques to expand/demonstrate the use of these functions.

I described here the first use of uniqueness plots in XPS. The manufacturers of XPS instruments, and others with XPS data analysis programs, should be contacted about the possibility of integrating this tool, and also the EW and AW into their software.

I introduced Information Content (IC) as a new chemometrics technique for data analysis, specifically to characterize spectra. This tool should also be tested on more advanced/complex data sets, which might include ToF-SIMS imaging/depth profiling data.

A primary goal of my dissertation work was to comprehensively understand materials via their multi-instrument characterization. I presented the nanodiamond characterization study in Chapter 7 as an exemplary study for nanomaterial characterization. The literature contains many examples of the under-characterization of materials that may limit the understanding of them in a way that prevents diagnosis of failure mechanisms and/or product improvement. Therefore, future work should focus on the multi-instrument characterization of other materials that have been developed for chromatography and sample preparation in the Linford lab.

Appendix A1:. An Introduction to the Equivalent Width and the Autocorrelation Width. Their Possible Applications in XPS Narrow Scan Analysis.*

A1.1 Introduction

The primary focuses of our research group at BYU are (i) the development of new materials for separation science, data storage, and surface protection, e.g., hydrophobic coatings, and (ii) the comprehensive characterization of these new materials using a suite of surface and bulk analytical tools, which includes the analysis of the data generated by these methods. Surface and data analysis help us understand the properties of our new materials and therefore explore possible applications for them. We are a little unusual in our research group because of our interest in very practical things, along with the science behind them.

The primary purpose of this article is to explore two mathematical functions in the context of their applications to data analysis. These are the equivalent width and autocorrelation width functions. We will show their application to X-ray photoelectron spectroscopy (XPS) narrow scan analysis, although the use of these width functions has wider applicability. We have written a number of times in the past about XPS, and for good reason.[1-3] XPS is being cited in excess of 10,000 time per year in the literature, which is far more citations than either Auger electron spectroscopy (AES) or time-of-flight secondary ion mass spectrometry (ToF-SIMS) is currently receiving.[3] One might say that XPS is currently the king of surface chemical analysis. However, we emphasize that ToF-SIMS becomes increasingly relevant and important as material complexity increases, and many materials, especially biological ones, are very complex. AES is extremely important for its small spot analytical capability (down to 10 nm!), which is roughly

*This appendix has been reproduced with permission from (Bhupinder Singh and Matthew R. Linford), Vacuum Technology and Coating, October 2015: p. 22-29

three orders of magnitude smaller than the lateral resolution of most XPS systems. In other words, ToF-SIMS and AES play and will continue to play extremely important roles in surface and material analysis, each possessing strengths and capabilities that XPS does not.

We wish to frame the problem of XPS narrow scan analysis in terms of a recent article by Wepasnick and coworkers.[4] In their work, they peak fitted XPS C 1s narrow scans they had taken from oxidized carbon nanotubes (CNTs). As might be expected, they fitted their peak envelopes to components corresponding to C-C, C-O, C=O, and O-C=O, and also a shake up (π-π^*) signal. One can think about this series of carbon signals, which show increasing degrees of oxidation, as carbon taking the following oxidation states: C(0), C(I), C(II), and C(III), respectively.[1] For their peak fitting, Wepasnick and coworkers employed two different sets of fit parameters they had taken from the literature.[5, 6] Two things are significant from their analysis. First, they obtained a good fit in both cases; both sets of peaks adequately fit their experimental data. Second, however, the O-C=O or C(III) signal was 5.9% of the total peak area in one fit and 11% of the fit in the other. This is more than a subtle difference between results. If one is interested in this functional group at a surface, where O-C=O, or C(III), presumably corresponds to surface carboxyl (–COOH) groups, one is not sure within a factor of two what the fraction of this group is. The carboxyl group is one of the most important in bioconjugate chemistry.[7] For example, Figure A1.1 shows the activation of a carboxyl group using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), which is a common water soluble reagent used in bioconjugate chemistry, and it subsequent coupling to an amine. This particular reaction of an amine with this activated carboxyl group (or with any activated carboxyl group for that matter) is extremely important in bioconjugate and organic chemistry.

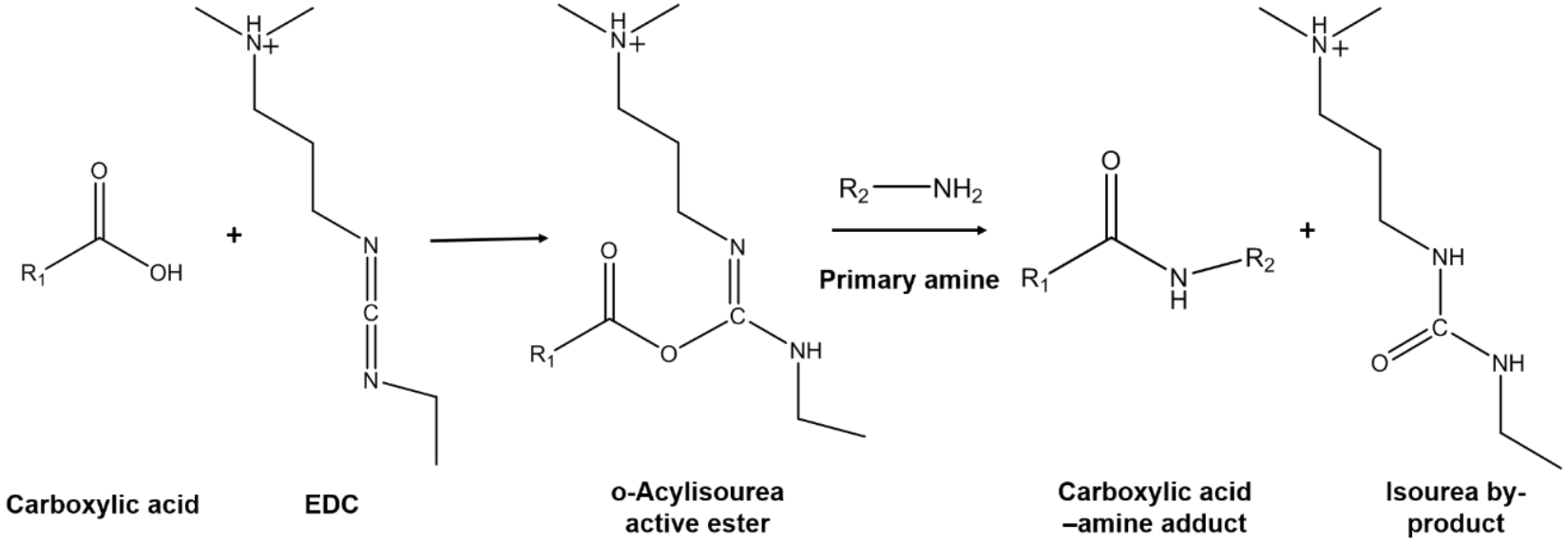


Figure A1. 1 Activation of a carboxylic acid with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and its subsequent reaction with a primary amine to form an amide. In this reaction, a carboxylic acid and an amine are coupled together.

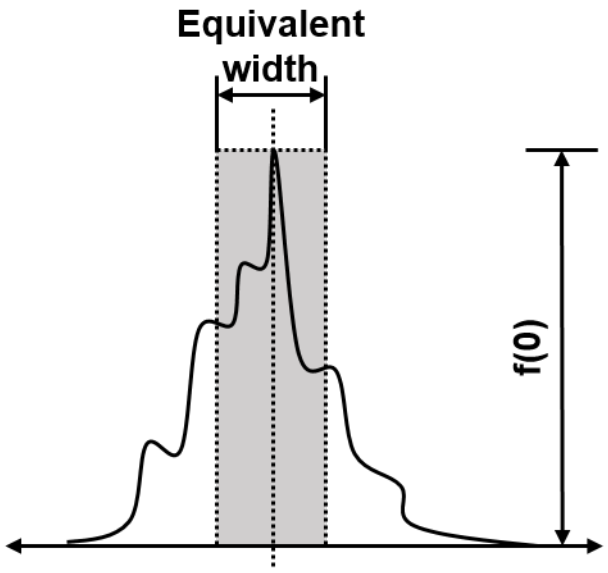


Figure A1. 2 Graphical illustration of the equivalent width (EW) of a general function showing the height of the function at the origin, f(0), and the resulting EW of the function, which is the area of the function divided by f(0).

As we have noted, there is almost always some subjectivity associated with any XPS narrow scan fit.[8] To address this issue, we recently published two papers that represent an unconventional approach to this problem. Before discussing this new direction, we wish to emphasize that we believe in peak fitting, we use it, and we will continue to employ it in our work. Peak fitting, when appropriately done, will remain the mainstay of XPS narrow scan analysis. But again, to address the issue of the subjectivity that is always present in these analyses, especially when there may be a considerable amount of complexity in a material, which may make it difficult to identify the peaks to be chosen for a fit, other approaches may be appropriate. In addition, a tool that can quickly spot differences between narrow scans may be of value. Accordingly, we have proposed that two mathematical functions: the equivalent width (EW) and the autocorrelation width (AW) may prove to be useful for some XPS narrow scan analyses.[2, 3]

A1.2 Theory

A1.2.1 The Equivalent Width (EW)

The equivalent width (EW) is a mathematical function that is used to describe the width of a function. In its classic definition, it is simply the area of a function divided by its height at the origin (its central ordinate), as follows:

$$EW = \frac{\int_{-\infty}^{\infty} f(x)dx}{f(0)} \qquad \text{(A1.1)}$$

If you think about this definition, you will see that the EW models a function as a rectangle, where the height of this rectangle is the value of the function at the origin, f(0). Figure A1.2 illustrates this definition for an arbitrary function.

Clearly, the EW is undefined for a function that passes through the origin, i.e., for f(0) = 0, and it is also sensitive to any shifts of the function along the x-axis. To make this function more applicable to XPS narrow scans, we redefined it somewhat as follows:

$$EW_{XPS} = \frac{Peak\ Area}{Peak\ Height} \tag{A1.2}$$

Where 'Peak Area' refers to the area of an XPS narrow scan over a user-defined baseline (or no baseline at all) and between user-defined limits. 'Peak height' refers to the highest point in the peak envelope, as measured from the baseline.[2] To remove any ambiguity in the EW_{XPS} results, we further recommended that the user report the maximum of the peak envelope (PE), which we designated as PE_{max}. As we showed in our paper, both EW_{XPS} and PE_{max} are quite sensitive to chemical changes in materials. In some cases, the use of these metrics may allow for a less biased analysis of narrow scans than traditional peak fitting. Figure A1.3 shows a C 1s narrow scan from oxidized carbon nanotubes that shows its EW_{XPS} and PE_{max} values.

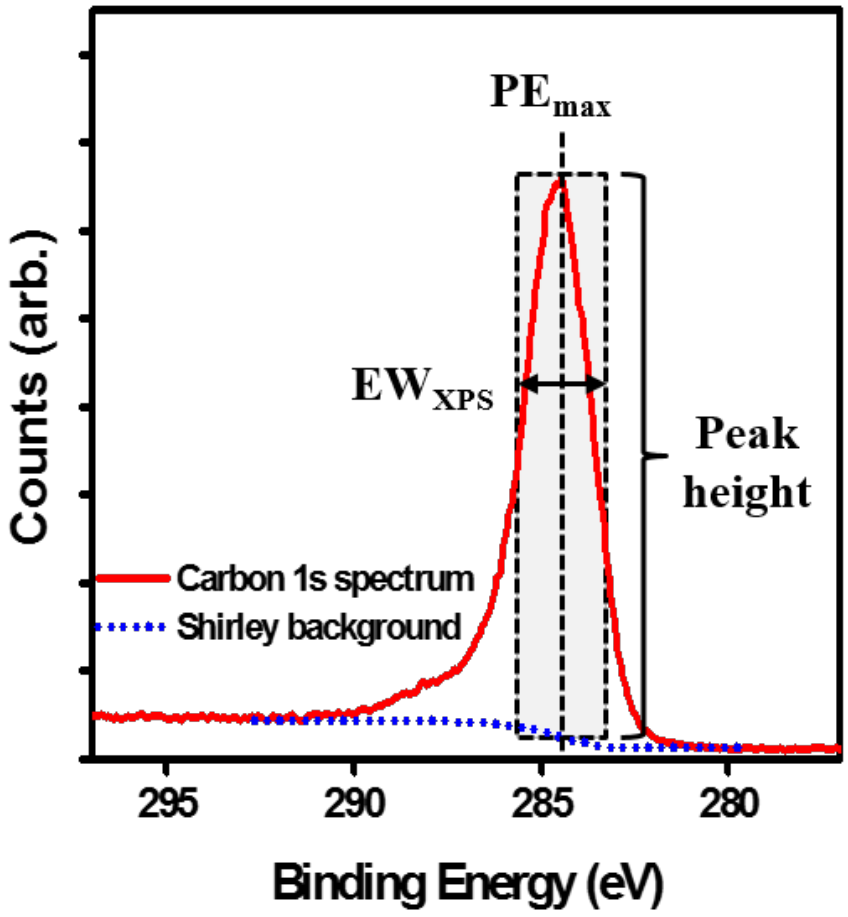


Figure A1. 3 Graphical illustration of the EW_{XPS} and PE_{max} values of a C 1s XPS narrow scan. Figure adapted from Bhupinder Singh, Daniel Velásquez, Jeff Terry, Matthew R. Linford. 'The Equivalent Width as a Figure of Merit for XPS Narrow Scans.' Journal of Electron Spectroscopy and Related Phenomena. **2014**, 197, 56 – 63. http://dx.doi.org/10.1016/j.elspec.2014.06.008.

A1.2.2 Convolution and (Auto)correlation

We explained convolution in some detail in one of our previous VT&C columns.[9] Here is a brief review of it again. The convolution of two functions, f(x) and g(x), is defined as:

$$h(x) = f(x) * g(x) = \int_{-\infty}^{\infty} f(u)g(x-u)du \tag{A1.3}$$

There are several interesting things in this definition. First, the symbol for convolution is '*'. This same symbol often denotes multiplication in computer science, but it obviously means something different here. Second, it is important to recognize that 'x' acts as a variable in h(x) and f(x)*g(x), while it is a constant in the integral on the right – we are integrating over the variable 'u', not 'x'. One finds a similar situation in the Fourier integral. Third, the function g(x) is reversed, i.e., flipped around or reflected, in the integral in Equation A1.3. This is because 'u' is the variable in the integral, and we have '-u' in g(x – u). Fourth, the constant 'x' in the integral acts to shift the reflected function g(x – u) to every possible position along the x-axis. (Mathematically, we might write: $x \epsilon \mathcal{R}$, where ϵ means 'is an element of' and $\mathcal{R}$ represents the real numbers.) Fifth, in a convolution we deal with two functions. We flip one of them around and shift it to every possible position along the x-axis, while leaving the other function where it was. At each of these shifts we multiply the two functions together and determine the resulting areas (integral). The value of the integral for a given shift, x, is the value of the convolution at that same point, i.e., h(x). It should be clear that calculating a convolution integral numerically can be a computationally expensive endeavor.

Conveniently, it turns out that it does not matter which function is reversed and shifted in a convolution. The same answer is obtained either way. In other words:

$$) = f(x) * g(x) = g(x) * f(x) = \int_{-\infty}^{\infty} f(u)g(x-u)du = \int_{-\infty}^{\infty} g(u)f(x-u)du \qquad \text{(A1.4)}$$

Of course, a function can also be convolved with itself, which gives:

$$h(x) = f(x) * f(x) = \int_{-\infty}^{\infty} f(u)f(x-u)du \qquad \text{(A1.5)}$$

Here, again, one copy of the function, f(u), remains unchanged from the original function, f(x). The same function, f(u), is then reversed and shifted to every possible point, x, along the x-axis, which gives us f(x – u) in Equation A1.5.

We now discuss another operation/integral, which is similar to the one in Equation A1.5. This operation is also the result of a function acting on itself in an integral. It is the self-correlation of a function. As we can see in Equation A1.6, the self-correlation (sc) is the same as a convolution, with the exception that the shifted function is not flipped around/reflected. That is, we multiply the original function, f(u), by a shifted copy of the same function, f(u – x), and record the integral of the product of the two functions as a function of the shift, x:

$$sc(x) = f(x) \star f(x) = \int_{-\infty}^{\infty} f(u)f(u-x)du \qquad \text{(A1.6)}$$

Notice that we have a different symbol here for a correlation. Instead of the '*' symbol we are now using the pentagram: ⋆. Again, we see that 'x' acts as a constant in the integral in Equation A1.6, while it is a variable in sc(x) and f(x)⋆f(x). We can abbreviate f(x)⋆f(x) as $f \star f$.

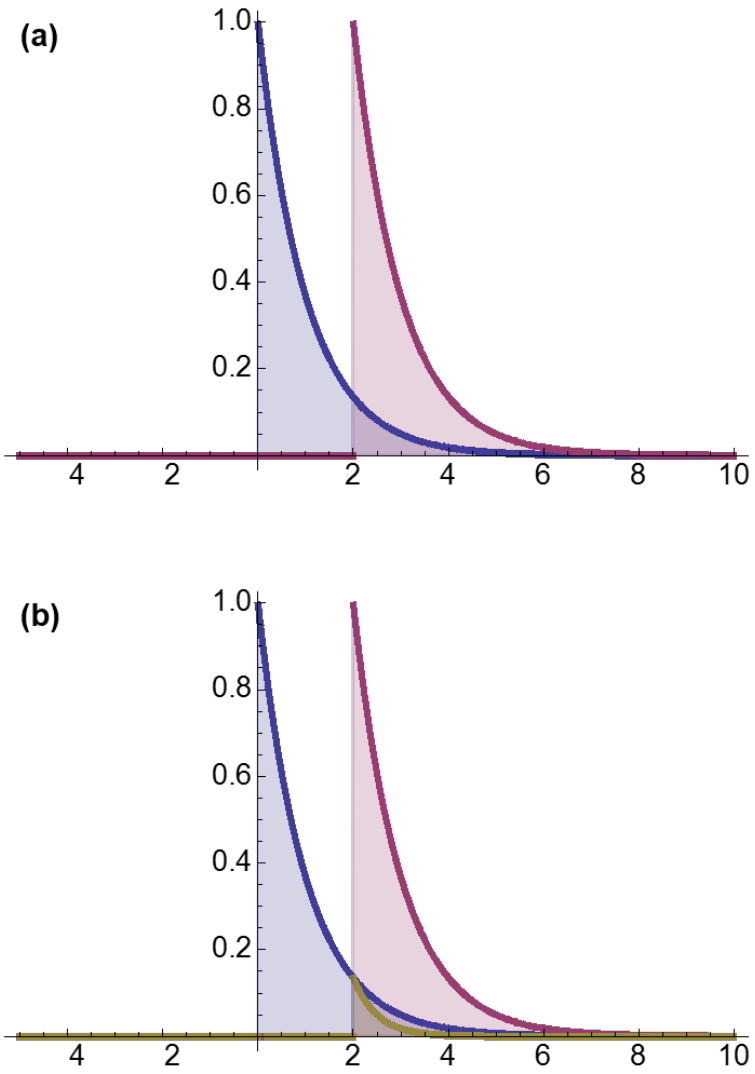


Figure A1. 4 (a) The function $\theta(u)\, e^{-u}$ (blue line) and the same function shifted two units to the right: $\theta(u - 2)\, e^{-(u-2)}$ (red line). $\theta(u)$, the unit step function, is defined in the text. (b) The same two functions as in (a) and their product (yellow-green line).

Figure A1.4 shows an example of multiplying a function by a shifted copy of itself. Obviously this represents a small part of the process that occurs in the self-correlation of a function. This function, g(u), (the blue line) is the product of a decaying exponential, e^{-u}, and the unit step function: θ(u), where the unit step function is defined as:

$$\theta(u) = \begin{cases} 0 \; for \; u < 0 \\ 1 \; for \; u \geq 0 \end{cases} \qquad \text{(A1.7)}$$

We also see in Figure A1.4 a copy of the original function that has been shifted two units to the right: $g(u - 2) = \theta(u - 2)\, e^{-(u - 2)}$ (red line). The product of these two functions, g(u)g(u – 2), is shown as the yellow-green line. The area under this yellow-green line represents the value of the self-correlation of the function g(u) at a shift of 2, i.e., sc(2). Once we have considered all possible shifts between the original function and its unreflected copy we will have generated its self-correlation function.

An interesting property of the self-correlation function is that it is an even function. An even function is a function that is symmetric about the y-axis so that f(x) = f(-x). Thus, from Equation A1.6, sc(x) = sc(-x) so that:

$$sc(x) = sc(-x) = \int_{-\infty}^{\infty} f(u)f(u - x)du = \int_{-\infty}^{\infty} f(u)f(u + x)du \qquad \text{(A1.8)}$$

This seems reasonable. That is, the definition of the self-correlation of a function (Equation A1.6) calls for a copy of the function to be shifted to every point along the x-axis. That is, since $x \in \mathcal{R}$, it does not matter whether we write f(u – x) or f(u + x) in Equations A1.6 and A1.8.

Yet another interesting property of the self-correlation function is that it is a maximum where there is no shift between the functions. That is, we can write:

$$sc_{max} = sc(0) = \int_{-\infty}^{\infty} f(u)f(u)du \tag{A1.9}$$

That is, sc(x) ≤ sc(0) for all x.

Ronald Bracewell's masterful book on the Fourier transform contains a thorough discussion of all of these topics.[10] A significant fraction of the notation in this article is the same as his.

A1.2.3 The Autocorrelation Width (AW)

The autocorrelation width (AW) of a function is defined as:

$$AW = W_{f*f} = \frac{\int_{-\infty}^{\infty}(f \star f)dx}{f \star f\,|_0} = \frac{(\int f dx)^2}{\int f^2 dx} = \frac{\int_{-\infty}^{\infty} h(x)dx}{h(0)} \tag{A1.10}$$

This math may look pretty intense, but we think it can be explained in a pretty straightforward manner. In this equation, $f \star f$ represents the self-correlation function. Accordingly, $\int_{-\infty}^{\infty}(f \star f)dx$ represents the area of this function. The denominator in Equation A1.10, $f \star f\,|_0$, then represents the value of the self-correlation function at x = 0 (where there is no shift between the function and its copy). Thus, the equivalent width and the autocorrelation width bear an important similarity. In both cases, they are simply the integral of a function, f(x) or $f \star f$, divided by the value of the function at the origin: f(0) or $f \star f\,|_0$.

As noted, Equation A1.1 for the EW is undefined for a function that passes through the origin, so the value of the EW function can change if the function shifts along the x-axis. In contrast, the value of the autocorrelation width (Equation A1.10) is the same whether the

function is centered at the origin or shifted to some other location. The two functions in Figure A1.5 have different EW values but the same AW values.

A little thought should convince one that the self-correlation of a function should produce a new function that is broader than the original one – think about how the tails of the two functions in Figure A1.5 would begin to overlap in a self-correlation some distance before the functions overlap (at x = 0 in Equation A1.6). Thus, according to the definitions of the EW and AW, the AW for a function should be broader than its EW.

Finally, we note from Equation A1.10 that the autocorrelation width of a function can be easily calculated simply by dividing the square of the area of the function ($(\int f dx)^2$) by the area of the function squared ($\int f^2 dx$). Additional details for performing these calculations are provided in our paper on this subject.[2, 3]

A1.3 Applications of the EW and AW to XPS Data Analysis

A1.3.1 Disclaimer

The use of width functions, e.g., the EW and AW, is a somewhat new and radical approach for understanding and studying XPS narrow scans. Some purists out there might not like it too much. However, we think that this approach has a place at the table in XPS peak analysis. Our intention is not to supplant traditional peak fitting. Only moments before writing these words, we (Singh and Linford) were working on the traditional peak fitting of the C 1s spectra of some sputtered carbon films we had prepared. This fit had a Shirley background, a main carbon peak, multiple, smaller oxidized components, and peak positions and peak widths that were constrained relative to each other. We considered the Gaussian-Lorentzian ratios for

our peaks and compared the amount of oxygen predicted by the oxidized components of the C 1s narrow scan to the amount of oxygen predicted by the O 1s signal. This analysis led us to introduce some asymmetry into the peaks of one of the fits – in the narrow scan from the material we expected to show the highest degree of conductivity. When we were done we plotted the sum of the fit components with the original data and also the residuals to the fit. The results from this analysis were important for our understanding of this material. It is certainly true that many researchers could get more out of their current peak fitting efforts if they thought harder about them, reading the literature better, and applying more of the sound principles and lessons that have been outlined there. We recognize, however, after years of XPS data analysis, that not every set of narrow scans can be easily fit, especially in the case of complex materials with elements in multiple oxidation states and/or when insufficient information is available about one's samples so that one does not know exactly how many peaks to include in a fit. We have also found that sometimes spectra look quite similar to each other, even if they differ in subtle ways that might be important. So perhaps the first use of the EW and AW functions could be to identify spectra that are different from each other, after which a traditional peak fitting could be performed. It seems like this type of analysis could also be used in automated spectral analysis for quality control. We are strong advocates in the multi-instrument characterization of surfaces and materials,[11-13] as are a number of other researchers. The use of width functions might fit nicely into these types of analyses. For example, a polymeric surface that contains a diversity of functionality will often show a complex C 1s narrow scan. If this surface is treated with an oxygen plasma for increasing amounts of time, it, along with its C 1s narrow scan, will in general become increasingly more complex. (Air/oxygen plasma treatment of polymers is a well-known way to increase their surface energies and therefore the adhesion to them.) In conjunction with

other analytical techniques, which might include wetting, ToF-SIMS, Fourier transform infrared spectroscopy (FTIR), and atomic force microscopy (AFM), perhaps one might provide EW or AW values for the C 1s peak envelopes, noting, for example, that the change in the width of the peak envelope is as expected, e.g., the envelopes become wider as the sample is increasingly oxidized. It has been noted that chi squared values are often most meaningful for a set of peak fits applied to the same narrow scan.[8] Similarly, the EW and AW may be particularly useful for comparing the peak envelopes of a series of related narrow scans.

A1.3.2 Application of the EW_{XPS} and AW_{XPS} to the C 1s spectra of oxidized CNTs

Figure A1.6 shows an example of the use of the EW function to the analysis of four C 1s narrow scans of ozone-treated CNT forests (three of the narrow scans are shown). These narrow scans are complex. They were collected as a part of study on microfabricated, CNT-templated thin layer chromatography (TLC) plates.[14] The oxygen content in these four CNT samples varied from ca. 3.7 – 5.3 at. %. They clearly contain multiple fit components, some at fairly low levels. In addition, they are very similar to each other. A priori, one would expect that an increasing degree of oxidation of these materials would lead to increasingly wide C 1s envelopes. This is what was observed. As shown in Figure A1.6d, the EW_{XPS} values for these narrow scans increase progressively with increasing O/C ratios in the materials. It should be emphasized that even though the oxygen content of these materials only changes to a relatively small degree and the resulting narrow scans are very similar, the EW_{XPS} values for these scans change in the expected way. In addition, because most of the changes in these spectra are expected to occur on the higher binding energy side of the main peak – we are introducing oxidized carbon into the materials – the PE_{max} values of the samples are expected to be nearly constant. This is what was observed (see Figure A1.6d). As expected, the AW_{XPS} values for these peak envelopes are larger

than the corresponding EW_{XPS} values, while showing the same trend. (For consistency with the EW_{XPS}, we will refer to the AW as AW_{XPS}.) In summary, Figure A1.6d clearly shows that the EW_{XPS} and AW_{XPS} pick up subtle changes in the peak envelopes that correlate to the materials' oxygen-to-carbon ratios.

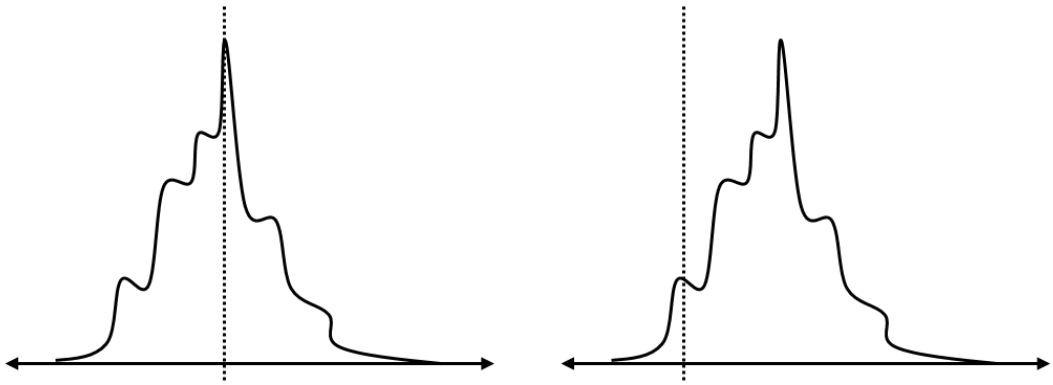


Figure A1. 5 A function at different positions along the x-axis, where the y-axis is indicated by the vertical dashed line.

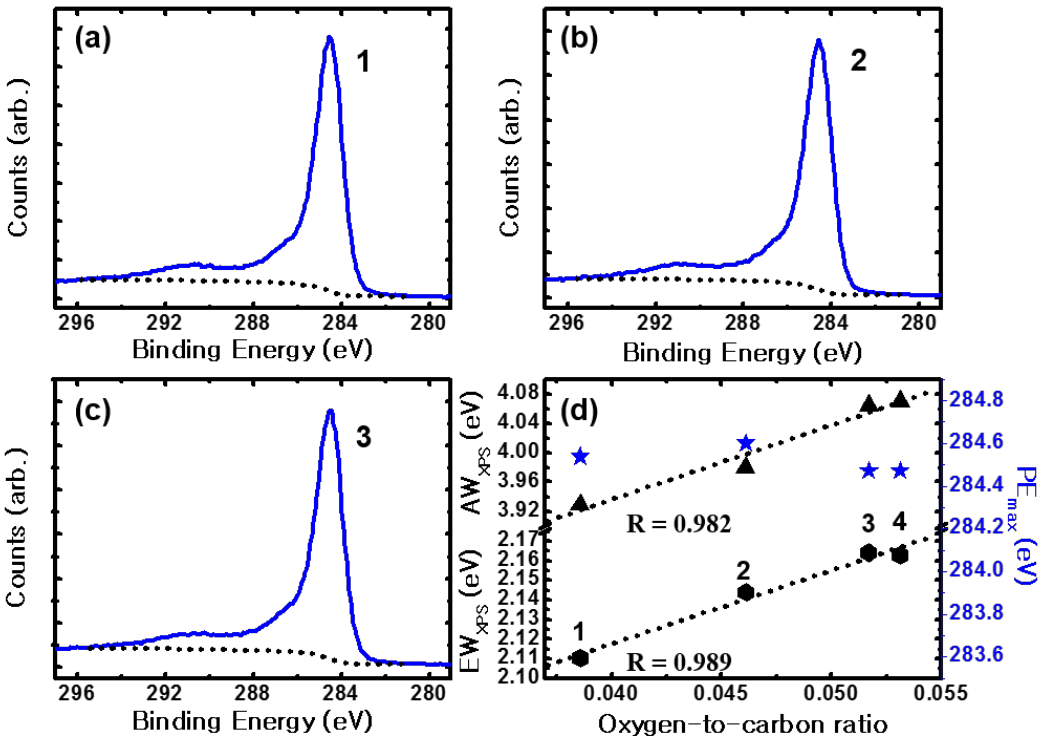


Figure A1. 6 C 1s narrow scans from oxidized CNT samples with (a) 3.7 at. % oxygen, (b) 4.4 at. % oxygen, and (c) 4.9 at. % oxygen. (d) EW_{XPS}, AW_{XPS}, and PE_{max} values of four C 1s spectra of oxidized CNTs (three of which are shown in this figure) as a function of their XPS oxygen-to-carbon ratios. Figure adapted from Bhupinder Singh, Daniel Velásquez, Jeff Terry, Matthew R. Linford. 'The Equivalent Width as a Figure of Merit for XPS Narrow Scans.' Journal of Electron Spectroscopy and Related Phenomena. **2014**, 197, 56 – 63. http://dx.doi.org/10.1016/j.elspec.2014.06.008.

A1.3.3 Application of the EW_{XPS} and AW_{XPS} to the Si 2p spectra of silicon surfaces with different oxide thicknesses

Five silicon shards (ca. 1.5 x 1.5 cm^2) were air plasma cleaned and then air oxidized for 1, 2, 3, 4, or 5 minutes at 900 ºC to increase the thicknesses of their oxides. The new oxide thicknesses were measured by spectroscopic ellipsometry (SE). A native oxide-coated silicon wafer (Si_{NO}) was also considered in this study. Other experimental details are given in our recent publications.[2, 3] Figure A1.7 shows the Si 2p narrow scans of the Si_{NO} sample, as well as for the silicon samples after 3 and 5 minutes of high temperature oxidation. While somewhat cluttery, Figure A1.7d shows the EW_{XPS}, AW_{XPS}, and PE_{max} values obtained from this study. Once again, in accord with theory, the AW_{XPS} values for the silicon samples are greater than their corresponding EW_{XPS} values. Interestingly, the EW_{XPS} and AW_{XPS} rise and then fall as the oxide thickness on the samples increases. This seems reasonable because (i) the Si 2p peak from the bulk silicon at ca. 99 eV will have a certain width (the width of this peak and the oxide peak were estimated by peak fitting the Si 2p envelopes to two peaks), (ii) one would expect the Si 2p peak from the oxide at ca. 103 eV to be broader than the bulk silicon peak, cf., in superficially oxidized aluminum the Al oxide peak is broader than the corresponding bulk peak, probably as a result of vibrational broadening,[8] and (iii) when the bulk silicon and oxide peaks have comparable intensities one would expect the width of this peak envelope to be greater than the width of either component. Finally, there is a shift in the PE_{max} value for the samples. At lower oxide thickness, the PE_{max} value is closer to the peak maximum for the bulk silicon Si 2p peak, but at thicker oxide thickness the PE_{max} value is closer to the peak maximum for the oxide Si 2p peak signal. Again, the EW_{XPS} and AW_{XPS} analysis is consistent with the expected surface chemistry.

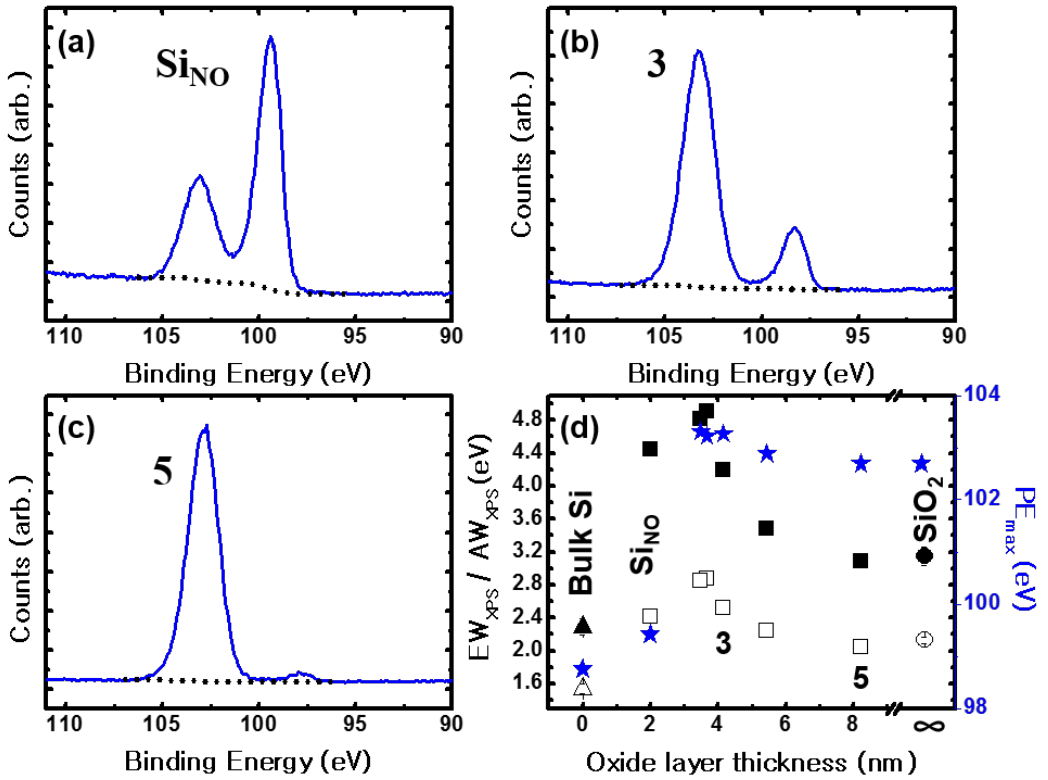


Figure A1. 7 Si 2p XPS narrow scans from (a) native oxide terminated silicon (Si_{NO}), and silicon surfaces oxidized for (b) 3 minutes, and (c) 5 minutes. (d) EW_{XPS}, AW_{XPS}, and PE_{max} values of Si 2p spectra as a function of their oxide layer thickness. Figure adapted from Bhupinder Singh, Daniel Velásquez, Jeff Terry, Matthew R. Linford. 'The Equivalent Width as a Figure of Merit for XPS Narrow Scans.' Journal of Electron Spectroscopy and Related Phenomena. **2014**, 197, 56 – 63. http://dx.doi.org/10.1016/j.elspec.2014.06.008.

A1.3.4 Application of the EW_{XPS} and AW_{XPS} to the C 1s spectra from five different nanodiamond samples

The C 1s narrow scans in Figure A1.8 were collected as part of a multi-instrument characterization of various nanodiamond samples. We have used these nanodiamonds to make the shells of pellicular (core-shell) particles for liquid chromatography.[15, 16] We could only obtain good C 1s scans from two of the samples (see, for example, Figure A1.8a). The other three samples charged, which resulted in artifacts in the spectra (see, for example, Figure A1.8b). In other words, we were unable to obtain good peak shapes for these samples, in spite of our attempts to deal with the charging issues by changing both our sample preparation and acquisition parameters. EW_{XPS} and AW_{XPS} values were plotted as a function of their PE_{max} values for the five samples that were analyzed. The red lines (ovals) are intended as a guide to the eye. Plotting these width functions vs their PE_{max} values provided a quick check on the quality of the spectra. Here, the best narrow scans (encapsulated by the red lines) have reasonable PE_{max} values (close to theoretical value of 284.6 eV), and lower AW_{XPS} and EW_{XPS} values – the samples that showed charging generally showed PE_{max} values that were significantly higher or lower than the expected values and/or higher EW_{XPS} and AW_{XPS} values. These results are in accord with the shapes of the C 1s peak envelopes (charged or uncharged) obtained for the samples.

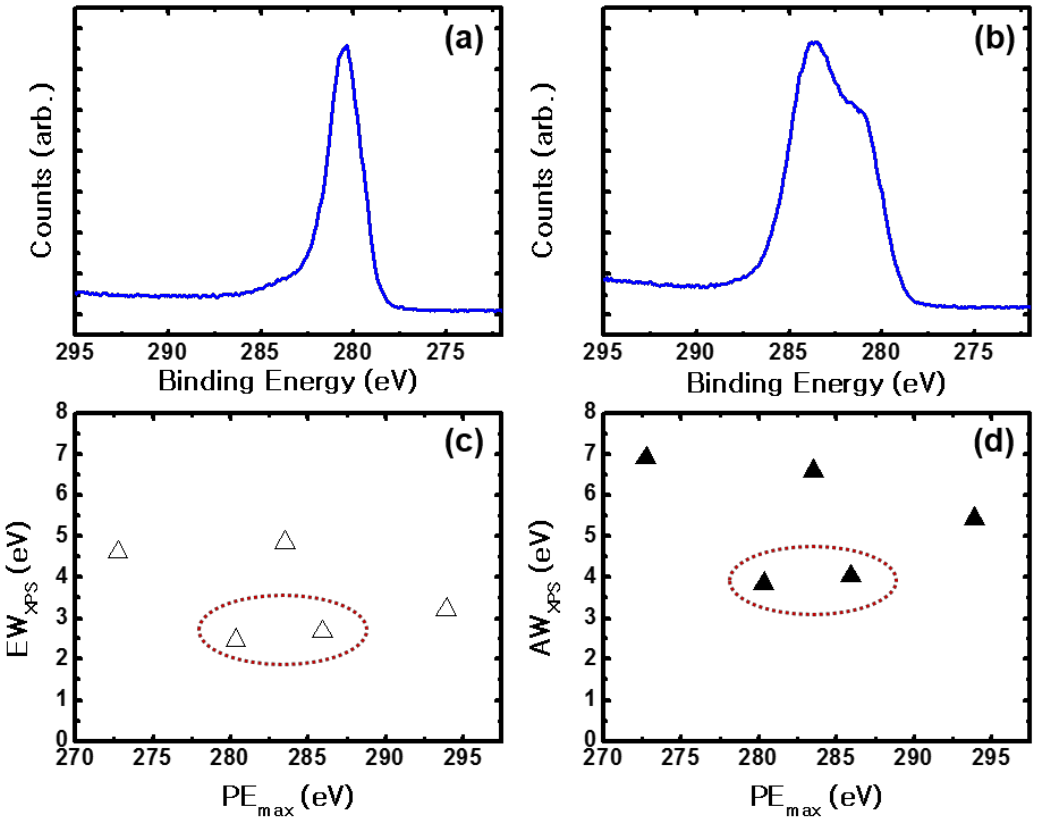


Figure A1. 8 (a) A good C 1s narrow scan from a nanodiamond sample. (b) A bad C 1s narrow scan from a nanodiamond sample showing artifacts due to charging. EW_{XPS} (c) and AW_{XPS} (d) values of five different C 1s spectra collected from five different nanodiamond samples as a function of their PE_{max} values. Figure adapted from Bhupinder Singh, Daniel Velásquez, Jeff Terry, Matthew R. Linford. 'The Equivalent Width as a Figure of Merit for XPS Narrow Scans.' Journal of Electron Spectroscopy and Related Phenomena. **2014**, 197, 56 – 63. http://dx.doi.org/10.1016/j.elspec.2014.06.008.

A1.4 Conclusions

We have discussed herein two functions that we believe may be useful for XPS narrow scan analysis: the equivalent width (EW_{XPS}) and the autocorrelation width (AW_{XPS}). Both appear to be sensitive to surface/material chemistry and/or sample charging. We recommend the EW_{XPS} and AW_{XPS} as complementary tools to traditional peak fitting practices, as they may involve less bias in XPS narrow scan analysis. These functions may ultimately be used in XPS narrow scan quality control and automated spectrum analysis.

A1.5 References

1. V. Gupta, H. Ganegoda, M. H. Engelhard, J. Terry and M. R. Linford, J Chem Educ **91** (2), 232-238 (2013).

2. B. Singh, D. Velázquez, J. Terry and M. R. Linford, J Electron Spectrosc **197** (0), 56-63 (2014).

3. B. Singh, D. Velázquez, J. Terry and M. R. Linford, J Electron Spectrosc **197** (0), 112-117 (2014).

4. K. A. Wepasnick, B. A. Smith, J. L. Bitter and D. H. Fairbrother, Anal Bioanal Chem **396** (3), 1003-1014 (2010).

5. M. Li, M. Boggs, T. P. Beebe and C. P. Huang, Carbon **46** (3), 466-475 (2008).

6. H. Wang, A. Zhou, F. Peng, H. Yu and J. Yang, J Colloid Interf Sci **316** (2), 277-283 (2007).

7. G. T. Hermanson, *Bioconjgate Techniques*, Third ed. (Academic Press, Boston, 2013).

8. P. M. A. Sherwood, J. Vac. Sci. Technol., A **14** (3), 1424-1432 (1996).

9. M. R. Linford, in *Vacuum Technology & Coating* (2014), pp. 25-31.

10. R. N. Bracewell, *The Fourier Transform and its Applications*, Third ed. (McGraw Hill, 2000).

11. D. S. Jensen, S. S. Kanyal, N. Madaan, J. M. Hancock, A. E. Dadson, M. A. Vail, R. Vanfleet, V. Shutthanandan, Z. H. Zhu, M. H. Engelhard and M. R. Linford, Surf Interface Anal **45** (8), 1273-1282 (2013).

12. H. Wang, B. M. Lunt, R. J. Gates, M. C. Asplund, V. Shutthanandan, R. C. Davis and M. R. Linford, ACS Appl. Mater. Interfaces **5** (17), 8407-8413 (2013).

13. C. H. Hung, B. Singh, M. G. Landowski, M. Ibrahim, A. J. Miles, D. S. Jensen, M. A. Vail, A. E. Dadson, S. J. Smith and M. R. Linford, Surf Interface Anal **47** (8), 815-823 (2015).

14. D. S. Jensen, S. S. Kanyal, N. Madaan, A. J. Miles, R. C. Davis, R. Vanfleet, M. A. Vail, A. E. Dadson and M. R. Linford, J Vac Sci Technol B **31** (3) (2013).

15. C.-H. Hung, L. A. Wiest, B. Singh, A. Diwan, M. J. C. Valentim, J. M. Christensen, R. C. Davis, A. J. Miles, D. S. Jensen, M. A. Vail, A. E. Dadson and M. R. Linford, J Sep Sci **36** (24), 3821-3829 (2013).

16. L. A. Wiest, D. S. Jensen, C. H. Hung, R. E. Olsen, R. C. Davis, M. A. Vail, A. E. Dadson, P. N. Nesterenko and M. R. Linford, Anal Chem **83** (14), 5488-5501 (2011).

Appendix A2: Flare Mixed-Mode Column: Separation of 2,4-D, MCPA, and Dicamba*

A2.1 Introduction

2,4-D (2,4-dichlorophenoxyacetic acid), MCPA (2-methyl-4-chlorophenoxyacetic acid), and dicamba constitute a major portion of the broad leaf herbicides used in agriculture.[1, 2] In addition, major agrochemical companies are planning to introduce 2,4-D and dicamba resistant crops into the market, which will be associated with increased use of these herbicides and may pose health and environmental risks.[3-5] If present in 2,4-D formulations, MCPA, a close chemical analogue of 2,4-D, can ruin the growing seedlings of 2,4-D resistant crops. Unfortunately, 2,4-D and MCPA are a critical pair – difficult if not impossible to separate by conventional LC.[1, 6] Thus, it is important to have a method that can completely separate these analytes. Here we present a rapid, baseline separation of 2,4-D, MCPA, and dicamba on the Diamond Analytics Flare mixed-mode column.

A2.2 Experimental

General: Isocratic elution was used to separate a mixture of 2,4-D, MCPA, and dicamba. A mobile phase of water and acetonitrile had formic acid as an additive. Analytes were purchased from Sigma-Aldrich (St. Louis, MO).

Chromatograph: Agilent 1290 Infinity Binary LC, DAD, ChemStation software

Sample: ca. 2 μg/mL of 2,4-D, MCPA, and dicamba in a mixture of 1:1 (Acetonitrile and H_2O)

Column: Flare Mixed-Mode Column (4.6 × 3.3 mm, 4.0 μm)

*This appendix has been reproduced with permission from (Bhupinder Singh, David S. Jensen, Andrew J. Miles, Andrew E. Dadson, and Matthew R. Linford), *Diamond Analytics Application Note DA1000-A*, 2013.

Injection volume: 2.0 μL

Temperature: 60 °C

Flow rate: 1.0 mL/min

Detection: UV detection at 280 nm

Mobile Phase:

A: 1.5 % formic acid in H_2O (pH 2.0)

B: 1.5 % formic acid in Acetonitrile

Elution: Isocratic: A:B :: 10:90

A2.3 Results and Discussion

2,4-D, MCPA, and dicamba were baseline separated in less than 2 minutes by isocratic elution on the Diamond Analytics Flare column, where 2,4-D and MCPA are a critical pair. Simple UV detection of the analytes at 280 nm was employed. The mobile phase was based on acetonitrile and water with formic acid as an additive. The aqueous mobile phase was at a relatively low pH (2.0). The Flare column can be used in a straightforward manner to separate these analytes.

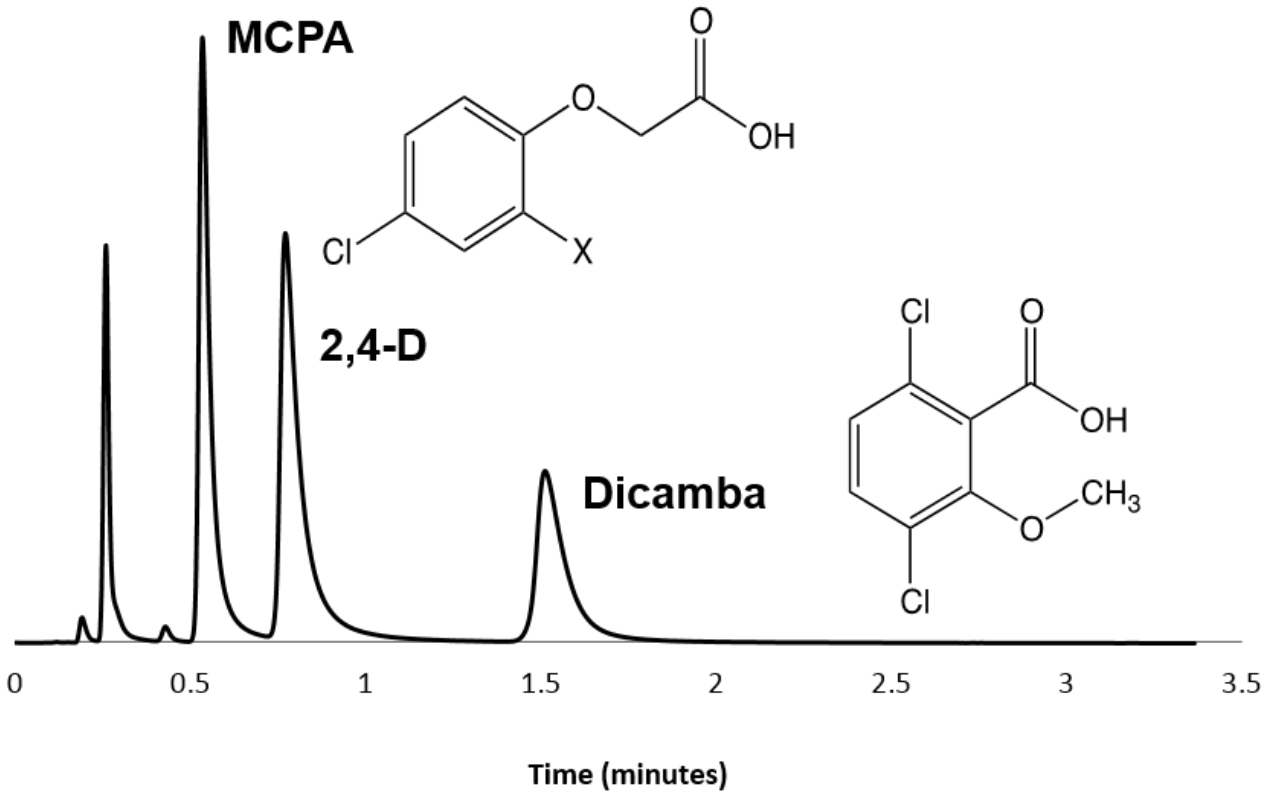


Figure A2. 1 Isocratic separation of MCPA (-X = -CH3), 2,4-D (-X = -Cl), and dicamba (280 nm).

A2.4 References

1. A. Balinova, J Chromatogr A **728** (1-2), 319-324 (1996).

2. R. J. Flanagan and M. Ruprah, Clin Chem **35** (7), 1342-1347 (1989).

3. W. G. H. Johnson, S.G.; Legleiter, T.R.;Whitford, F.; Weller, S.C; Boredelon, B.P.; Lerner, B.R., (2012).

4. B. Freese, (http://www.centerforfoodsafety.org/wp-content/uploads/2012/02/FSR_24-D.pdf, 2012).

5. N. P. I. Center, in *National Pesticide Information Center* (npic.orst.edu/factsheets, http://npic.orst.edu/factsheets/2,4-DTech.pdf).

6. J. S. Kowalski, K; Nolan, L. and Wittrig, R., (Restek Corporation, http://www.restek.com/pdfs/pres-2005-260-13p.pdf).

Appendix A3: Comparison of the Flare Mixed-Mode Column Against Commercial C_{18} and PFP Columns for the Separation of Critical Pairs of Acidic Herbicides*

A3.1 Introduction

Two significant critical pairs[1-3] of acidic herbicides: 2,4-D and MCPA, and dichlorprop and mecoprop were separated with dicamba, which is another important herbicide, on the nanodiamond-based, core-shell, mixed-mode/C_{18} 4 μm Flare column. The structures of these chemicals are given in Figure A3.1, where 2,4-D and MCPA, and dichlorprop and mecoprop differ by a CH_3 group at the 'Y' position of the general structure of the molecules.

The performance of the Flare column was compared to commercial C_{18} and PFP columns, which showed coelution of 2,4-D and MCPA, and of dichlorprop and mecoprop. Dicamba, another important acidic herbicide, was present in some of the test mixtures.

A3.2 Results and Discussion

EPA-Method 555 was employed with a commercial C_{18} column in an attempt to separate mixtures of (a) 2,4-D, MCPA and dicamba, and (b) dichlorprop and mecoprop.
As shown in Figure A3.2, the commercial C_{18} column was unable to resolve 2,4-D from MCPA and dichlorprop from mecoprop, and dicamba eluted before 2,4-D and MCPA.

Because of a previous report that showed the separation of various chlorinated analogues of 2,4-D on a PFP column, but not of 2,4-D and MCPA or dichlorprop and mecoprop,[4] the separation of these critical pairs was attempted on a commercial PFP column.

*This appendix has been reproduced with permission from (Bhupinder Singh, David S. Jensen, Andrew J. Miles, Andrew E. Dadson, and Matthew R. Linford), *Diamond Analytics Application Note DA1000-B*, 2013

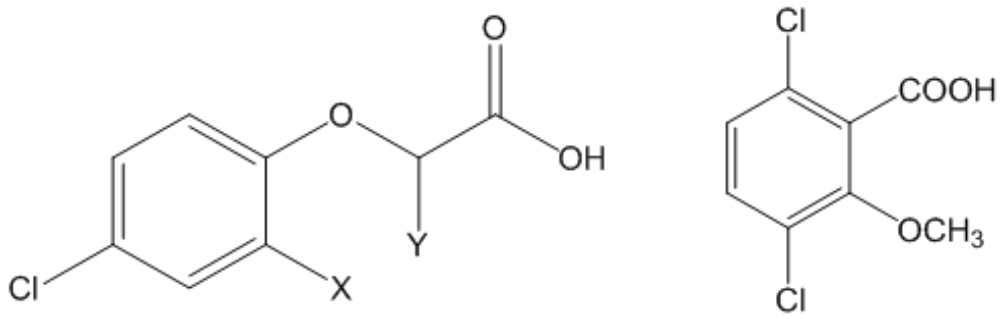


Figure A3. 1 Structures of 2,4-D (X = Cl, Y = H), MCPA (X = CH_3, Y = H), dichlorprop (X = Cl, Y = CH_3), mecoprop 2,4-D (X = CH_3, Y = CH_3), and dicamba (on right).

Under the conditions that resolved the chlorinated analogues of 2,4-D from each other, the PFP column could not separate either 2,4-D from MCPA or dichlorprop from mecoprop (see Figure A3.3). These results provide further evidence for the difficulty associated with separating these compounds. Note that dicamba could be separated from 2,4-D and MCPA and that it again eluted before the critical pair.

The separation of these critical pairs was again attempted on the C_{18} column, but this time using the conditions from the PFP column (see Separation Conditions 2 below). Coelution of the critical pairs again occurred, albeit at lower retention times.

Finally, the Diamond Analytics Flare mixed-mode/C_{18} column was used to probe the analyte mixtures. Note that the Flare column was shorter and had larger particles than either of the comparison commercial columns.

As shown in Figure A3.4, both 2,4-D and MCPA and also mecoprop and dichlorprop could be baseline separated. Both separations occurred in less than two minutes. The Flare column shows different selectivity than the C_{18} and PFP columns – dicamba elutes after 2,4-D and MCPA on the Flare column. The unique selectivity of the Flare mixed-mode column stems from its combination of amines and C_{18} chains. Indeed, as shown herein, neither C_{18} nor PFP groups alone is sufficient to separate 2,4-D and MCPA or mecoprop or dichlorprop, where these critical pairs differ by a single substituent that is *ortho* to and shielded by an acidic moiety. Thus it appears that an interaction between the carboxyl groups in the analytes and amines in the stationary phase facilitates the separation of the critical pairs.

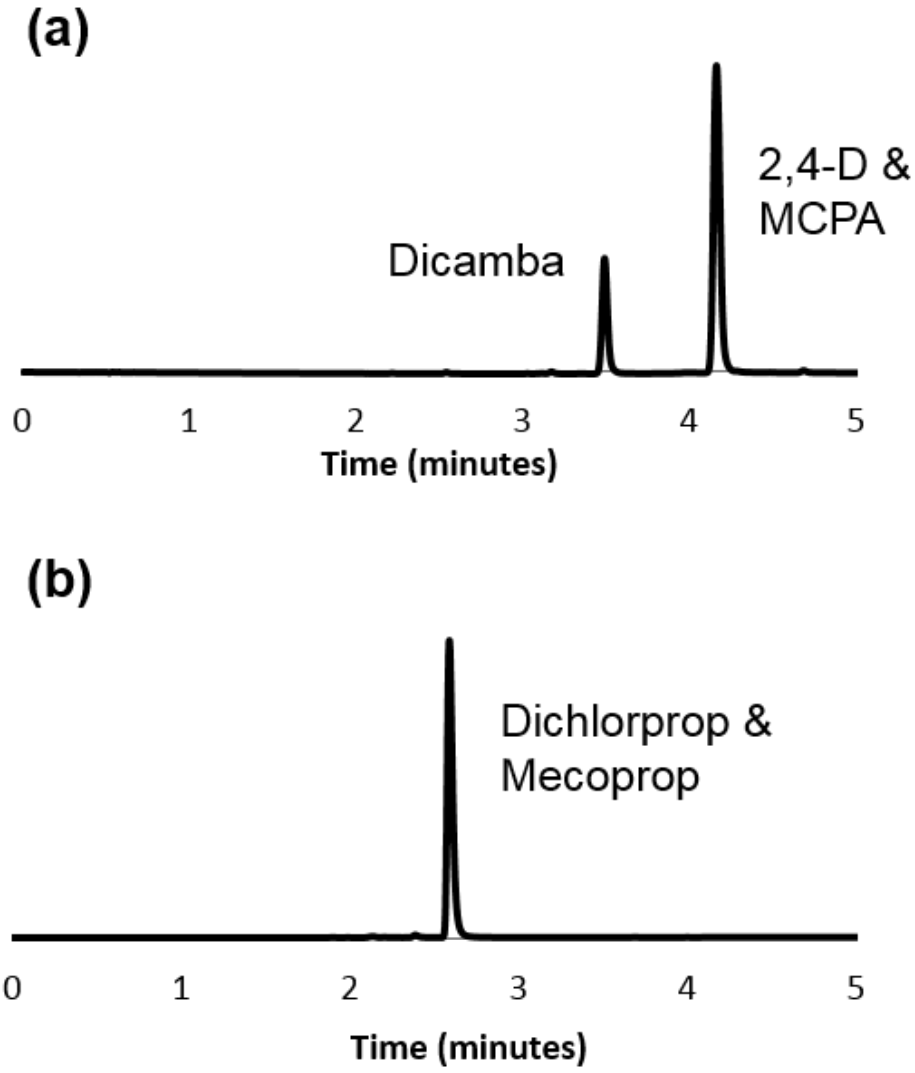


Figure A3. 2 Attempted separations on a commercial C_{18} column of (a) MCPA, 2,4-D, and dicamba, and (b) dichlorprop and mecoprop. See Separation Conditions 1 below.

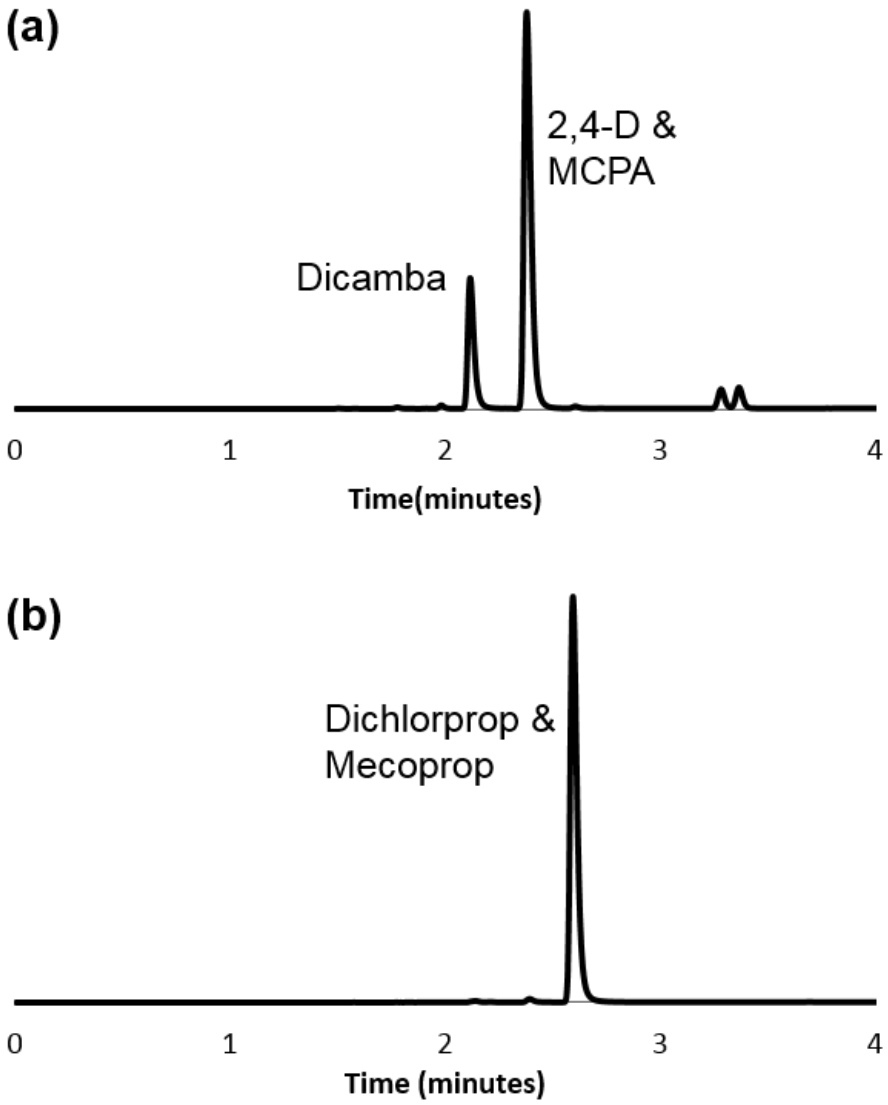


Figure A3. 3 (a) Attempted separation on a commercial PFP column of (a) MCPA, 2,4-D, and dicamba, and (b) dichlorprop and mecoprop. See Separation Conditions 2 below.

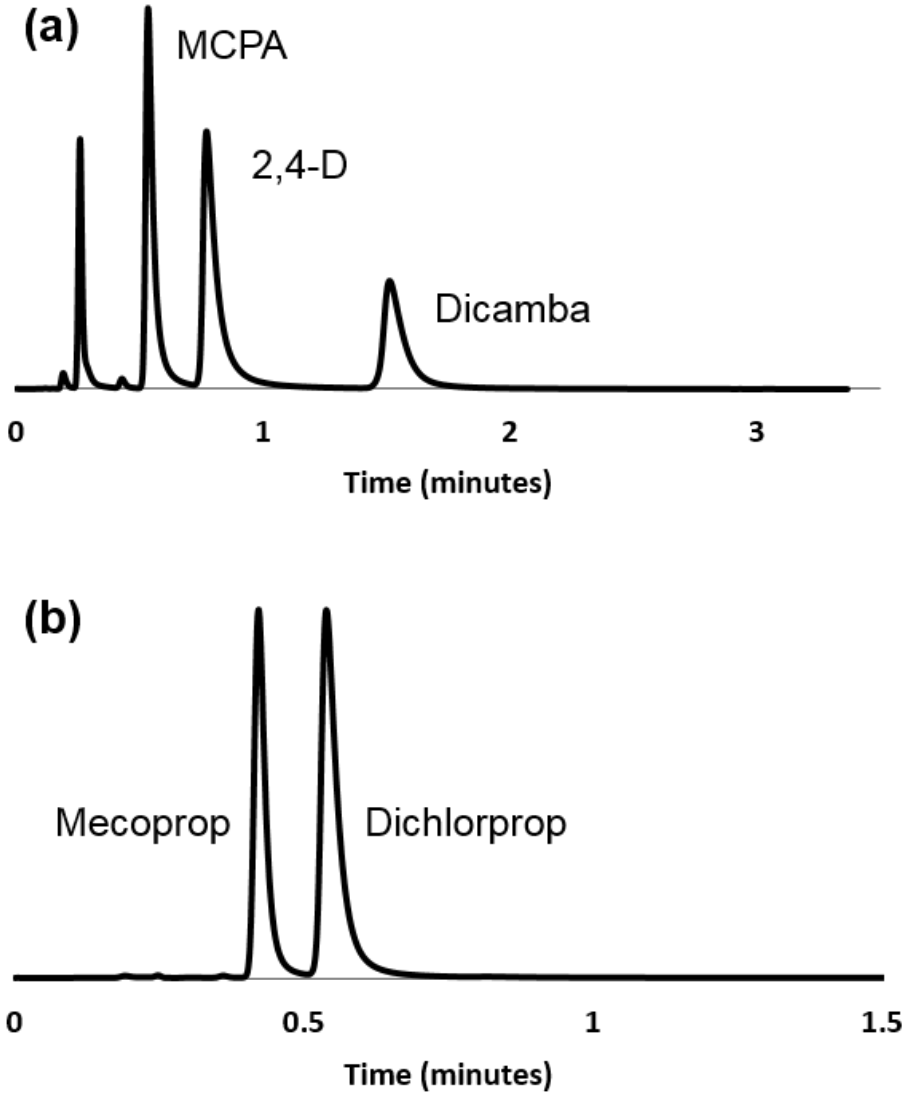


Figure A3. 4 Separation on a Diamond Analytics Flare mixed-mode column of (a) MCPA, 2,4-D, and dicamba, and (b) dichlorprop and mecoprop. See Separation Conditions 3 below.

A3.3 Conclusions

Under two sets of conditions, a commercial C_{18} column failed to separate either 2,4-D and MCPA or mecoprop and dichlorprop. A commercial PFP column was also unable to separate these critical pairs. Both of these columns were longer and had smaller particles than the Diamond Analytics Flare mixed-mode column. In contrast, the Flare column baseline separates 2,4-D and MCPA, and also mecoprop and dichlorprop. The Flare column shows different selectivity than the commercial columns.

A3.4 Experimental

Reagents: Analytes were purchased from Sigma-Aldrich (St. Louis, MO)

Chromatograph: Agilent 1290 Infinity Binary LC, DAD, ChemStation software

Sample: 2 μg/mL of 2,4-D, MCPA, dicamba, dichlorprop and/or mecoprop in 1:1 ACN:H_2O.

Columns: Diamond Analytics Flare Mixed-Mode (4.6 x 33 mm, 4.0 μm), commercial C_{18} (2.6 μm, 100 Å, 4.6 x 50 mm), and commercial PFP (2.6 μm, 100 Å, 4.6 x 150 mm).

Separation Conditions

Separation Conditions 1 (EPA Method 555)

Employed with the commercial C_{18} column.

Injection volume: 2.0 μL

Temperature: 30 °C

Flow rate: 1.0 mL/min

Detection: UV detection at 280 nm

Mobile Phase:

A: 25 mM H_3PO_4

B: Acetonitrile

Elution: Gradient

10 to 90 % B in 10 minutes

Separation Conditions 2

Employed with the commercial PFP and C_{18} columns.

Injection volume: 2.0 μL

Temperature: ambient

Flow rate: 1.0 mL/min

Detection: UV detection at 280 nm

Mobile Phase:

A: 0.1 % Formic acid in H_2O

B: 0.1 % Formic acid in Acetonitrile

Elution: Gradient

55 to 75 % B in 6 minutes

Separation Conditions 3

Employed with the Diamond Analytics Flare mixed-mode column.

Injection volume: 2.0 μL

Temperature: 60 °C

Flow rate: 1.0 mL/min

Detection: UV detection at 280 nm

Mobile Phase:

A: 1.5 % formic acid in H_2O (pH 2.0)

B: 1.5 % formic acid in Acetonitrile

Elution: Isocratic (10:90 :: A:B)

A3.5 References

1. A. Balinova, J Chromatogr A **728** (1-2), 319-324 (1996).

2. J. S. Kowalski, K; Nolan, L. and Wittrig, R., (Restek Corporation, http://www.restek.com/pdfs/pres-2005-260-13p.pdf.

3. C. H. Lindh, M. Littorin, A. Amilon and B. A. G. Jonsson, Rapid Commun Mass Sp **22** (2), 143-150 (2008).

4. Commercial application note: http://www.phenomenex.com/Application/Detail/18520?returnURL=/Search.

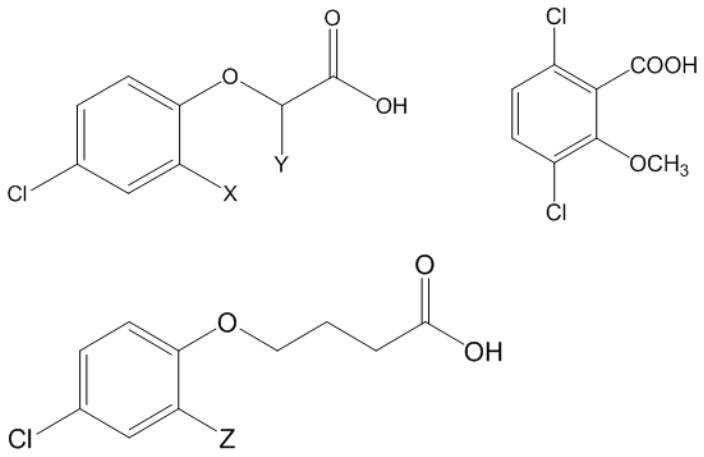
Appendix A4: Probing the Retention Mechanism of the Flare Mixed-Mode Column at Low pH via Acidic Herbicides with Different pK_a Values*

A4.1 Introduction

The Diamond Analytics Flare core-shell column is made by depositing alternating layers of polyallyamine (PAAm) and nanodiamond on solid carbon core particles.[1, 2] The final PAAm layer is reacted with 1,2-epoxyoctadecane and cross linked with 1,2,7,8-diepoxyoctane to produce a mixed-mode weak anion exchange (WAX)/C_{18} column.[2] These diverse functionalities impart unique selectivity to the column. A number of classes of compounds have been successfully analyzed on the column, including the separation of critical pairs of acidic herbicides.[3]

The present work is focused on understanding the retention mechanism of acids (acidic herbicides) as a function of their pK_a values on the Flare column at low pH. In particular, gradient elution was used to separate 2,4-D (2,4-dichlorophenoxyacetic acid), MCPA (2-methyl-4-chlorophenoxyacetic acid), 2,4-DB (4-(2,4-dichlorophenoxybutyric acid), MCPB (4-(4-chloro-2-methylphenoxy) butanoic acid), dichlorprop, mecoprop and dicamba (see Figure A4.1). At high pH the elution order of the analytes is reversed.

*This appendix has been reproduced with permission from (Bhupinder Singh, David S. Jensen, Andrew J. Miles, Andrew E. Dadson, and Matthew R. Linford), *Diamond Analytics Application Note DA1000-C*, 2013



Analyte	X	Y	Z
2,4-D	Cl	H	-
MCPA	CH_3	H	-
Dichlorprop	Cl	CH_3	-
Mecoprop	CH_3	CH_3	-
2,4-DB	-	-	Cl
MCPB	-	-	CH_3

Figure A4. 1 Structures of various acidic herbicides. Dicamba is on the upper right. The X, Y, and Z groups in the table define the other analytes.

A4.2 Experimental

All analytes were purchased from Sigma-Aldrich (St. Louis, MO).

Chromatograph: Agilent 1290 Infinity Binary LC, DAD, ChemStation software

Sample: 2 mg/mL of 2,4-D, MCPA, 2,4-DB, MCPB, dichlorprop, mecoprop and/or dicamba in a mixture of 1:1 H_2O:acetonitrile (ACN)

Column: Flare Mixed-Mode Column (4.6 × 3.3 mm, 4.0 μm)

Injection volume: 2.0 μL

Temperature: 60 °C

Flow rate: 1.0 mL/min

Detection: UV detection at 280 nm

Mobile Phase:

A: 1.5 % formic acid in H_2O, pH 2.0

B: 1.5 % formic acid in ACN

Elution: Gradient

40 to 60 % B in 6 minutes

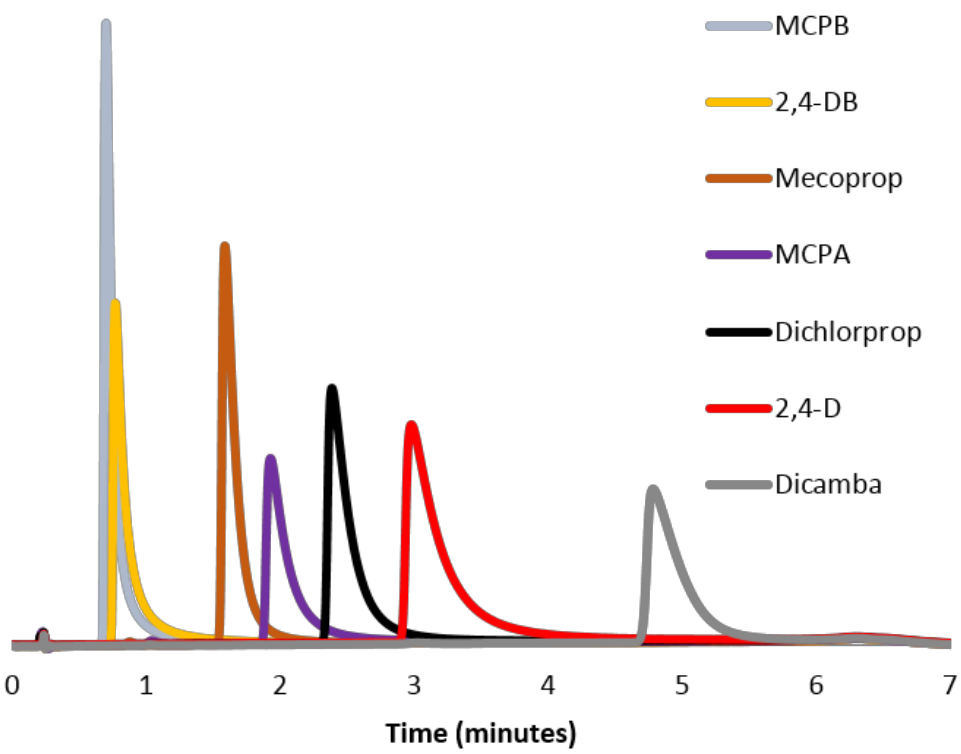


Figure A4. 2 Separation of acidic herbicides on the Flare mixed-mode column.

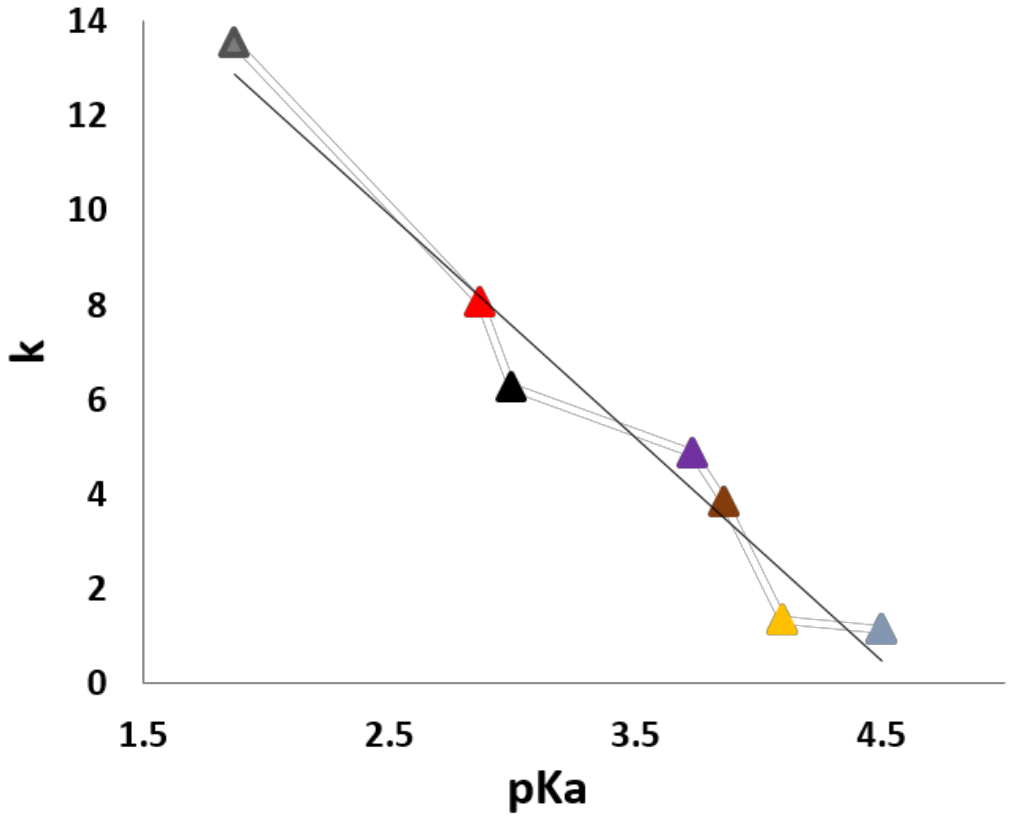


Figure A4. 3 Dependence of retention factor (k) on the pK_a values of the analytes. The colors of the symbols for the analytes correspond to the colors in Figure *A4.*2.

A4.3 Results and Discussion

We first discuss the charge/protonation state of the Flare column and the acidic herbicide analytes at pH 2.0. The state of the column is simple to understand. At this low pH the amino groups on the Flare stationary phase are mostly protonated so the Flare column is strongly positively charged. To help us understand the protonation state of the acidic analytes, we use the Henderson-Hasselbach (H.-H.) equation. A description of the acid-base chemistry that leads to the H.-H. equation and a derivation of the equation itself are given in a Diamond Analytics app note on this topic.[4] The H.-H. equation is:

$$pH = pK_a + \log ([A^-]/[HA])$$

where the pH is that of the mobile phase, the pK_a is that of the analyte, $[A^-]$ is the concentration of the deprotonated analyte, and [HA] is the concentration of the protonated analyte. Of course the H.-H. equation, as presented here, applies more accurately to aqueous solutions. Nevertheless, it should help us estimate the degree of analyte deprotonation in the water/ACN mobile phase and thus analyte retention.

Table A4.1 gives the pK_a values of the acidic analytes considered herein and also the ratio of the deprotonated to protonated form of these analytes at the mobile phase pH as given by the H.-H. equation. The least acidic herbicides, i.e., MCPB and 2,4-DB, are almost entirely protonated at pH 2.0. In contrast, the most acidic herbicides, i.e., dichlorprop, 2,4-D, and dicamba have significant amounts of the analyte in the deprotonated form.

The seven acidic herbicide analytes under consideration here were injected on the Flare column. As shown in Figures A4.2 and A4.3, the retention of these analytes is strongly influenced by their pK_a values. That is, less acidic analytes with higher pK_a values show less

retention – a smaller fraction of these analytes is deprotonated – they are mostly neutral – so they interact less with the positively charged stationary phase. For example, MCPB has the highest pK_a value (4.5) of the analytes considered herein and elutes first. In contrast, the more acidic analytes with lower pK_a values show greater retention – a larger fraction of these analytes is deprotonated so they interact more with the positively charged stationary phase. For example, dicamba has the lowest pK_a value (1.87) of the analytes considered herein and elutes last.

To illustrate that the retention mechanism of the Flare column is strongly pH dependent, dicamba, 2,4-D, and 2,4-DB were injected separately using a pH 12 mobile phase. We have previously shown that the Flare column acts in reversed phase mode (not ion exchange) under these conditions.[5, 6] Accordingly, for these analytes we observed a reversal of the elution order that is in Figures A4.2 and A4.3. Dicamba was least retained (it is the most compact of the three analytes), followed by 2,4-D (it is a structural isomer of dicamba but more extended), followed by 2,4-DB (it has two more methylene units than 2,4-D).

Table A4. 1 pK_a values of acidic herbicide analytes and an estimate of the ratio of deprotonated to protonated analyte ($[A^-]/[HA]$) in the pH 2.0 mobile phase.

Analytes	**pK_a**	**[A-]/[HA]**
MCPB	4.5	0.003
2,4-DB	4.1	0.008
Mecoprop	3.86	0.014
MCPA	3.73	0.019
Dichlorprop	3.0	0.100
2,4-D	2.87	0.134
Dicamba	1.87	1.349

A4.4 References

1. G. Saini, D. S. Jensen, L. A. Wiest, M. A. Vail, A. Dadson, M. L. Lee, V. Shutthanandan and M. R. Linford, Anal Chem **82** (11), 4448-4456 (2010).

2. L. A. Wiest, D. S. Jensen, C. H. Hung, R. E. Olsen, R. C. Davis, M. A. Vail, A. E. Dadson, P. N. Nesterenko and M. R. Linford, Anal Chem **83** (14), 5488-5501 (2011).

3. B. Singh, D. S. Jensen, A. J. Miles, A. E. Dadson and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/Acidic%20Herbicides.pdf, 2013).

4. C. H. Hung, A. A. Kazarian, A. E. Dadson, B. Paull, P. N. Nesterenko and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/Diamond%20Analytics%20Application%20Guidelines.pdf, 2013).

5. L. A. Wiest, D. S. Jensen, A. J. Miles, A. Dadson and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/%CE%B22-Agonists%20and%20Amphetamines.pdf, 2013).

6. L. A. Wiest, D. S. Jensen, A. J. Miles, A. E. Dadson and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/FlareMixedModeColumn-TriazineHerbicides.pdf, 2013).

Appendix A5: A Reproducibility Study with the Flare Mixed-Mode Column of 50 Consecutive Injections at Elevated pH*

A5.1 Introduction

The Diamond Analytics Flare core-shell column offers unique selectivity for an array of analytes. As a mixed-mode (weak anion exchange/C_{18}) column, it can be made more hydrophobic or more charged depending on the pH of the mobile phase.[1] For example, at pH 2.0, the column has weak anion exchange properties, making separation of critical pairs of acidic herbicides possible.[2] Working at the other end of the pH scale (ca. pH 12), the column behaves in reversed phase mode, separating triazine herbicides, amphetamines and tricyclic antidepressants (TCAs).[3-5] The amphetamines and TCAs are bases that are best retained on a reversed phase stationary phase when they are deprotonated, i.e., under elevated pH conditions.

The present work is focused on the stability and reproducibility of the Flare column at extremes of pH. For the high pH separations (pH 12), a complex test mixture (Test Mixture 1, see Figure A5.1) was prepared that contained acidic, basic, and neutral analytes. For the low pH analyses (pH 2), the test mixture contained dicamba and 2,4-dichlorophenoxyacetic acid (2,4-D) (Test Mixture 2, see Figure A5.2). During stability tests, % RSD values were recorded for various peak parameters, including retention times (t_R), retention factors (k), tailing factors (T_f), efficiencies (N/m), and resolutions (R_s).

*This appendix has been reproduced with permission from (Bhupinder Singh, Supriya S. Kanyal, David S. Jensen, Andrew E. Dadson, and Matthew R. Linford), *Diamond Analytics Application Note DA2000-A*, 2013

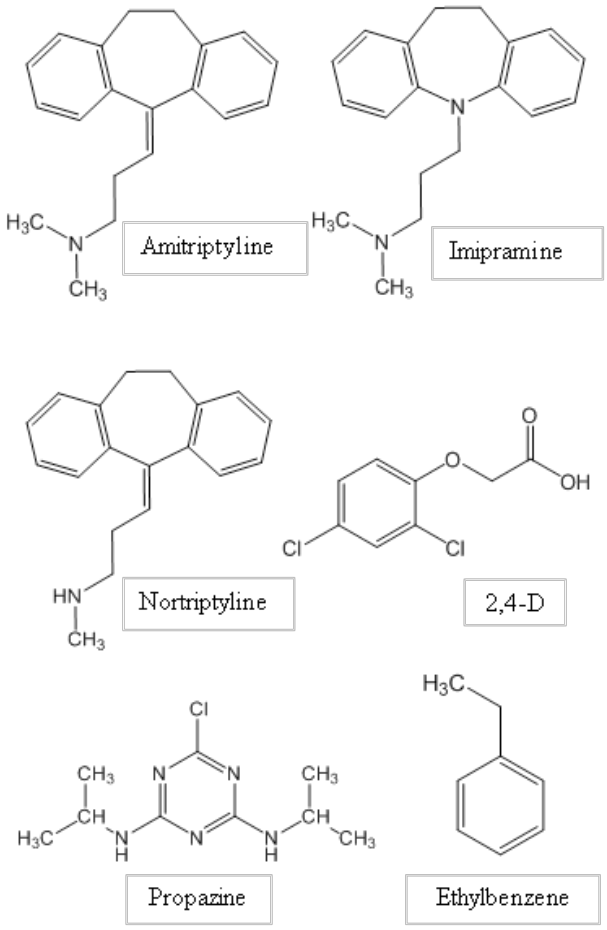


Figure A5. 1 Names and structures of analytes in the test mixture.

.

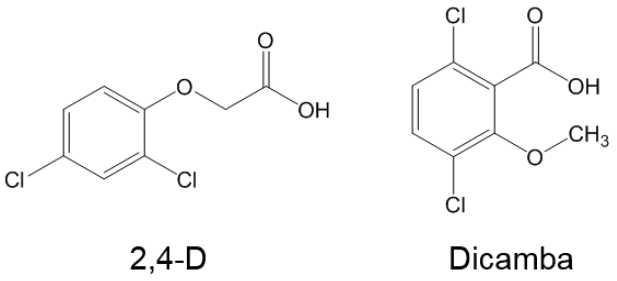


Figure A5. 2 Names and structures of the analytes in Test Mixture 2.

A5.2 Experimental

A5.2.1 Stability and reproducibility of the Flare column at high pH

Fifty consecutive injections of Test Mixture 1 were made on the Flare column on Day 1 to test the intra-day variability of the column (see Separation Conditions 1). After Day 1, 20 injections were made per day for 6 consecutive days under the same conditions on the same Flare column to determine the inter-day variability. Around 2700* column volumes of the pH 12 mobile phase were flushed through the column during the duration of this test.

A5.2.2 Stability and reproducibility of the Flare column at low pH

Approximately 5000* column volumes of a pH 2.0 mobile phase at 60 °C (see Separation Conditions 2) were flushed through the column, during which time Test Mixture 2 was periodically injected to monitor the column stability.

* The column volume referred here is the geometrical column volume, $\pi r^2 h$ (r and h denote the radius and length of the column respectively)
If void volume was to be used instead of geometrical column volume, the figures would have changed to 4000 instead of 2700 column volumes and 7400 instead of 5000 column volumes.

A5.2.3 General Separation Conditions

Analytes: All analytes were purchased from Sigma-Aldrich (St. Louis, MO).

Chromatograph: Waters 1525 Binary HPLC pump; Breeze 3.30 SPA software.

Column: Diamond Analytics Flare Mixed-Mode column (4.6 x 33 mm, 4 μm)

Injection volume: 5.0 μL

Elution: Isocratic

Detection: UV detection at 254 nm

Flow rate: 1.0 mL/min

Separation Conditions 1

To probe column stability at elevated pH

Test Mixture 1: ca. 2 mg/mL each of 2,4-D, propazine, ethylbenzene, nortriptyline, imipramine and amitriptyline in acetonitrile:water (1:1)

Column temperature: 35 °C

Mobile Phase: 70:30 premixed solution of aqueous phosphate buffer at pH 12: acetonitrile, prepared gravimetrically.

Separation Conditions 2

To probe stability at low pH

Test Mixture 2: ca. 0.5 mg/mL of 2,4-D and dicamba in acetonitrile:water (1:1)

Temperature: 60 °C

Mobile Phase: 10:90 (v/v) premixed solution of H_2O:acetonitrile, with 1.5% formic acid as additive.

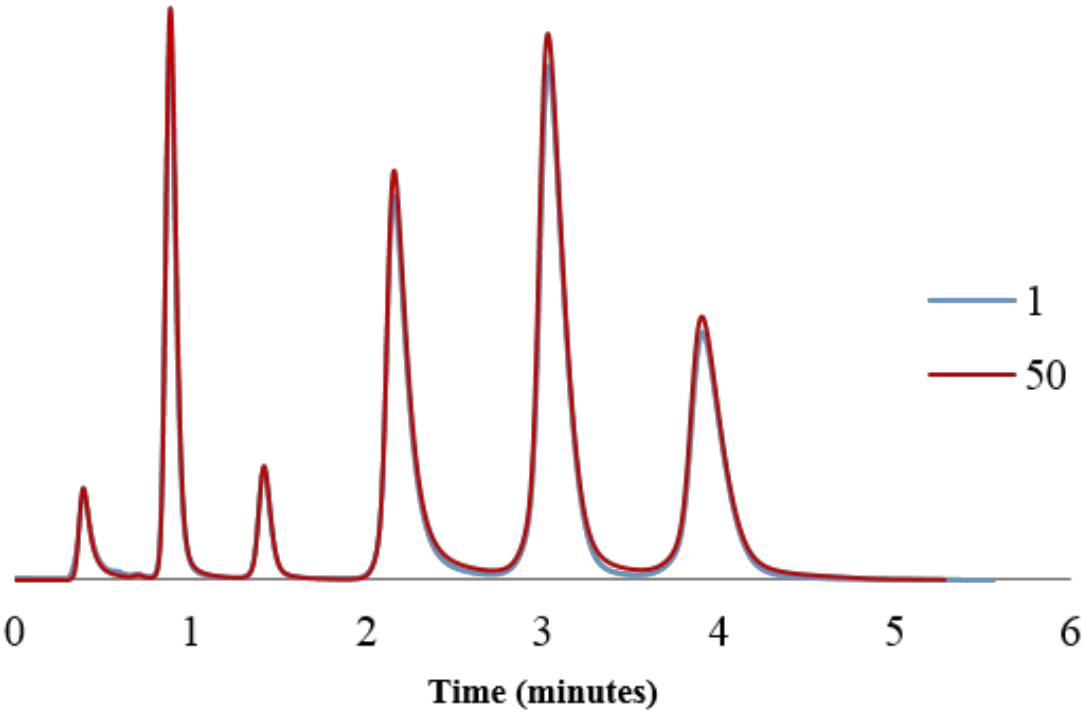


Figure A5. 3 Chromatograms corresponding to the 1st and 50th injections of the test mixture on the Flare column.

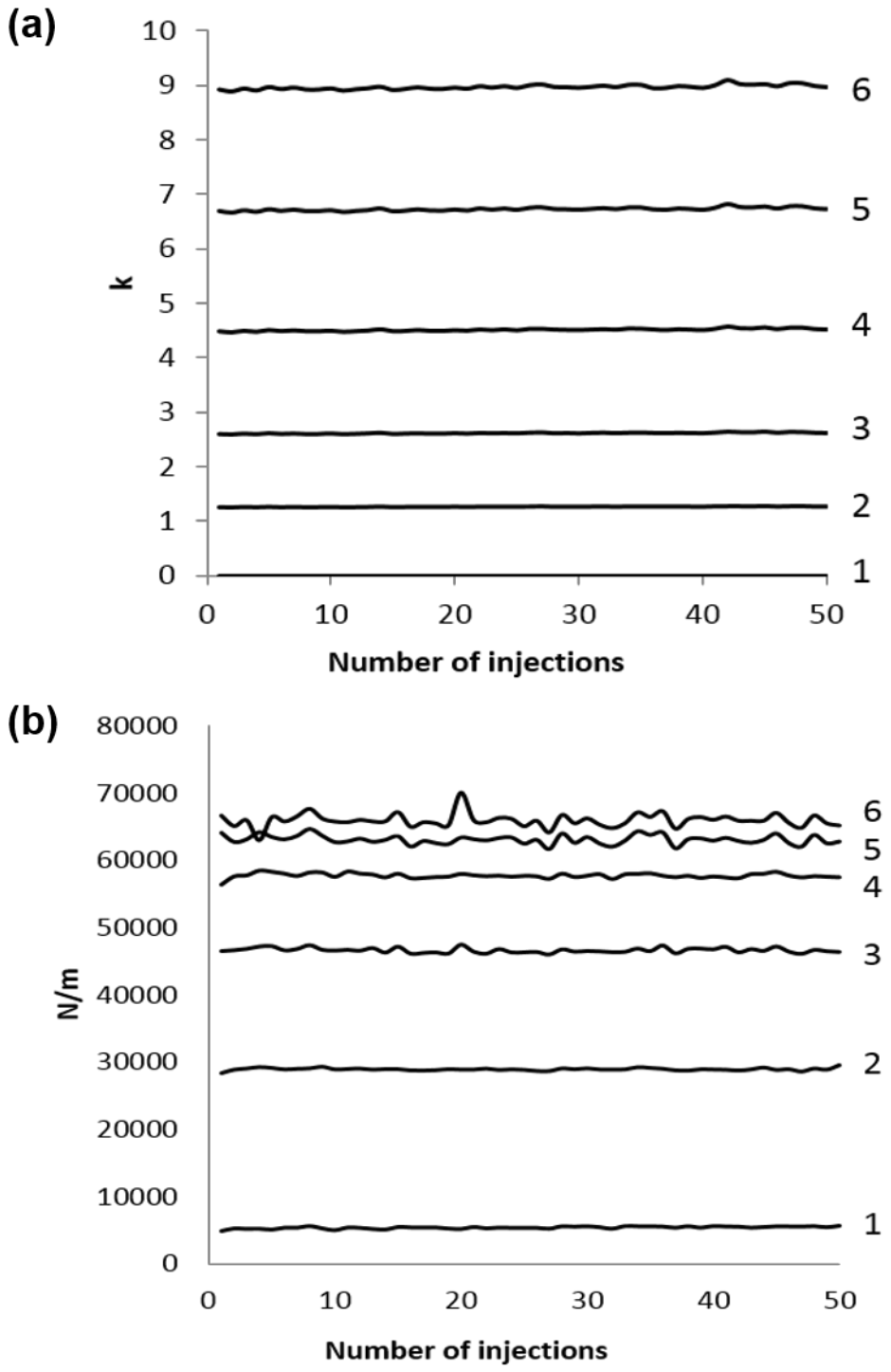


Figure A5. 4 Plot of (a) the retention factor (k) and efficiency (N/m) vs. number of injections for the analytes under study. (1 = 2,4-D, 2 = Propazine, 3 = Ethylbenzene, 4 = Nortriptyline, 5 = Imipramine, 6 = Amitriptyline)

A5.3 Results and Discussion

A5.3.1 Intra-Day Reproducibility at high pH

Figure A5.3 shows chromatograms of the first and fiftieth injections of Test Mixture 1 on day 1 on the Flare column. These chromatograms overlap extremely well and represent 640 column volumes. All of the analytes showed good peak shapes, i.e., for propazine, ethylbenzene, nortriptyline, imipramine, and amitriptyline, the average tailing factors over the 50 injections were 1.31, 1.25, 1.79, 1.34, and 1.27, respectively. % RSD values were calculated for various peak parameters, including retention times, retention factors, tailing factors, efficiencies, and resolutions, all of which point to excellent intra-day reproducibility, i.e., in most cases % RSD values are below unity (see Table A5.1).

Figures A5.4a and A5.4b further show that the retention factors and efficiencies for all the analytes remain essentially constant over the 50 injections. For the later eluting analytes, efficiencies range from ca. 50,000 – 70,000 N/m.

A5.3.2 Inter-Day Reproducibility at high pH

To determine the inter-day variability, the data from the first day (fifty consecutive injections) and from the next six days (twenty consecutive injections per day) were pooled and various peak parameters, including retention times, retention factors, tailing factors, efficiencies, and resolutions were recorded. Table A5.2 shows % RSD values for these different peak parameters, which are generally less than or equal to 5 %, suggesting good inter-day reproducibility for the Flare column.

Table A5. 1 Intra-day reproducibility results;[6] **%** RSD values of various peak parameters for the analytes in Test Mixture 1 over 50 consecutive injections. t_R = retention time, k = retention factor, N/m = plates/meter, R_s = resolution, T_f = tailing factor.

Analyte	**t_R**	**k**	**N/m**	**R_s**	**T_f**
Propazine	0.60	0.72	0.68	1.66	0.91
Ethyl-benzene	0.60	0.57	0.59	0.30	1.37
Nortriptyline	0.62	0.52	0.73	0.47	2.90
Imipramine	0.60	0.46	1.07	0.59	1.82
Amitriptyline	0.60	0.45	1.55	0.75	1.72

A5.3.3 Inter-Day Reproducibility at Low pH

Figure A5.5 shows a separation of Test Mixture 2, using Separation Conditions 2 (at pH 2.0). 5000 column volumes of the mobile phase were flushed through the column at 60 °C and % RSD values for t_R and k are given in Table A5.3, which correspond to excellent stability of the column at low pH. Figure A5.6 shows plots of retention time and retention factor vs. column volume, which again show the solid stability of the Flare column at low pH.

Table A5. 2 Inter-day reproducibility results. **%** RSD values of various peak parameters for the analytes in Test Mixture 1 over a span of 7 days (50 injections on day 1 and 20 injections per day for the following 6 days): t_R = retention time, k = retention factor, N/m = plates/meter, R_s = resolution, T_f = tailing factor.

Analyte	**t_R**	**k**	**N/m**	**R_s**	**T_f**
Propazine	1.48	4.08	0.88	4.80	1.35
Ethyl-benzene	1.10	2.68	1.75	1.65	3.41
Nortriptyline	1.42	2.87	1.65	1.42	5.15
Imipramine	1.65	2.96	2.88	1.16	4.23
Amitriptyline	1.69	3.10	2.97	0.79	4.62

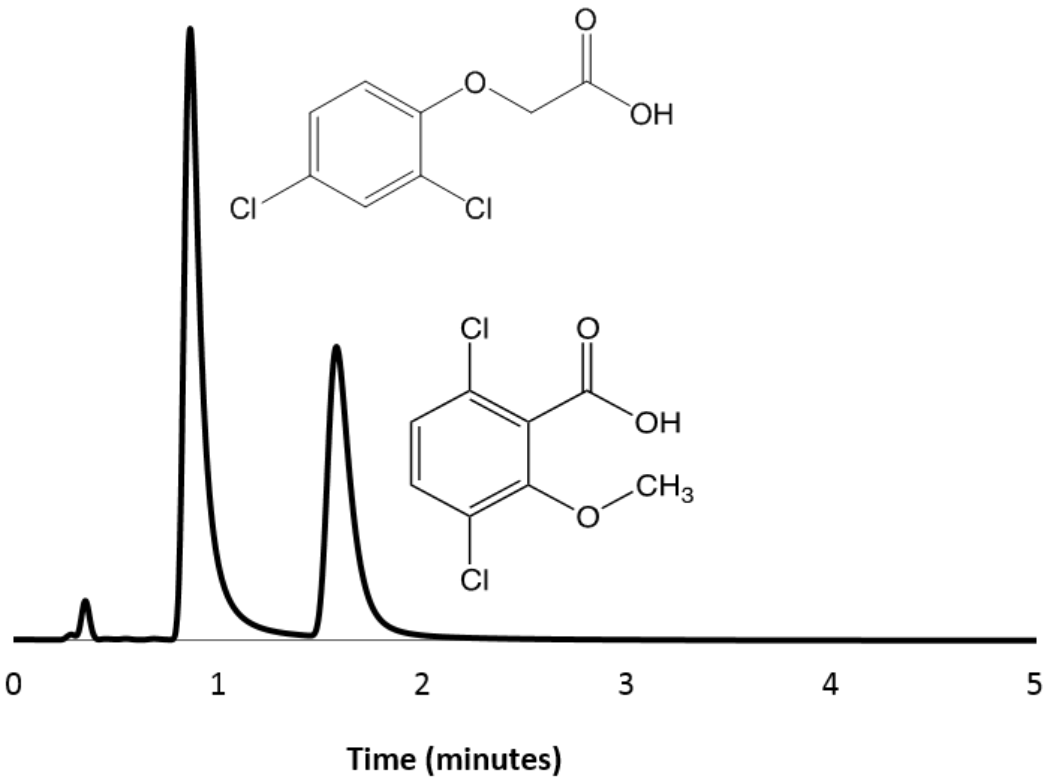


Figure A5. 5 Separation of the analytes in Test Mixture 2, using Separation Conditions 2.

Table A5. 3% RSD values for retention time (t_R) and retention factors (k) of 2,4-D and dicamba for passage of 5000 column volumes of mobile phase through the Flare column using Separation Conditions 2.

Analyte	**t_R**	**k**
2,4-D	0.62	1.28
Dicamba	0.88	1.26

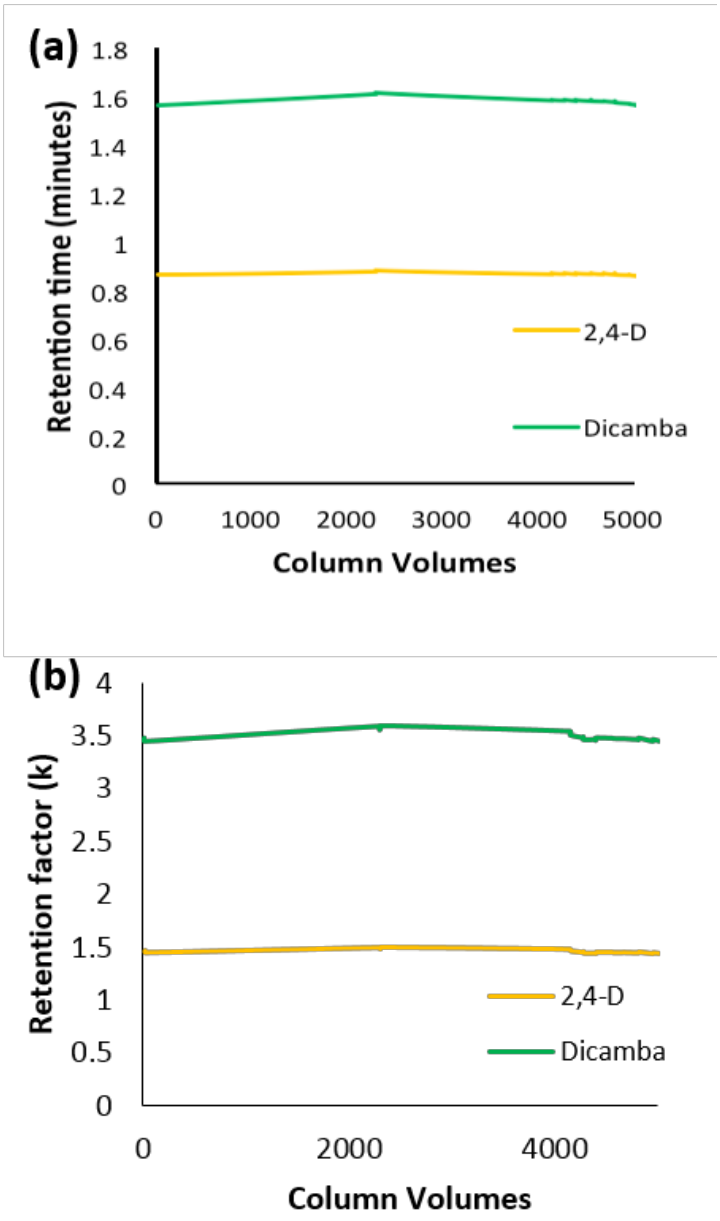


Figure A5. 6 Plots of (a) retention time and (b) retention factor vs. column volumes for 2,4-D and dicamba at pH 2 on the Flare column.

A5.4 References

1. C. H. Hung, A. A. Kazarian, A. E. Dadson, B. Paull, P. N. Nesterenko and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/Diamond%20Analytics%20Application%20Guidelines.pdf, 2013).

2. B. Singh, D. S. Jensen, A. J. Miles, A. E. Dadson and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/Acidic%20Herbicides.pdf, 2013).

3. L. A. Wiest, D. S. Jensen, A. J. Miles, A. Dadson and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/%CE%B22-Agonists%20and%20Amphetamines.pdf, 2013).

4. L. A. Wiest, D. S. Jensen, A. J. Miles, A. E. Dadson and M. R. Linford, (Diamond Analytics, http://diamond-analytics.com/uploads/FlareMixedModeColumn-TriazineHerbicides.pdf, 2013).

5. D. Analytics, (Diamond Analytics, http://diamond-analytics.com/uploads/Tricyclic%20Antidepressants%20(TCAs)1.pdf, 2013).

6. Peak parameters for 2,4-D were not included in Tables A5.1 and A5.2 because: (i) an impurity in the test mixture prevented a straightforward determination of the T_f of 2,4-D and (ii) at elevated pH, 2,4-D is deprotonated and poorly retained on the Flare column so it co-elutes with the dead time marker, making determination of efficiency difficult.

Appendix A6: WO_3^-/WO_3H^- and WO_4^-/WO_4H^- Assignments to Negative ion ToF-SIMS Spectra from Five Nanodiamond Samples Discussed in Chapter 7*

A6.1 Introduction

The negative ion ToF-SIMS analysis of the nanodiamond samples is shown in Figure 7.2 (see Chapter 7). We observe two peak envelopes. The first peak envelope contains six peaks positioned at m/z 230 to m/z 235, and a second envelope has six peaks positioned from m/z 246 to m/z 251. Initially, based on molecular weight estimates, the peak envelopes were attributed to WO_3^- and WO_4H^-, respectively. To confirm that we were making reasonable assignments, we looked at the tungsten isotopes and their ratios. The following table gives the most stable tungsten isotopes.

Tungsten (W) isotopes	^{182}W	^{183}W	^{184}W	^{186}W
Relative Abundance	26.50%	14.31%	30.64%	28.43%

A6.2 Calculations

Looking at the isotopic ratios of tungsten, it becomes clear that the initial assignment to the peak envelopes of WO_3^- and WO_4H^- did not explain the peaks at m/z 233 and 235 in the proposed WO_3^- envelope and the peaks at m/z 246 and 250 in the proposed WO_4H^- envelope. The next hypothesis was that the peak envelopes were a contribution of peaks due to WO_3^- and WO_3H^-, and WO_4^- and WO_4H^-. To prove our hypothesis, we calculated the contributions of each peak using the following equations.

*This appendix is reproduced from the Supporting Information (Bhupinder Singh, Stacey J. Smith, David S. Jensen, Hodge F. Jones, Andrew E. Dadson, Paul B. Farnsworth, Richard Vanfleet, Jeffrey K. Farrer, and Matthew R. Linford) *Analytical and Bioanalytical Chemistry*, accepted for publication, 2015.

$$(I_{230,WO3}) + (0) = I_{Total,230}$$

$$(I_{231,WO3} = I_{230,WO3} * \frac{14.31}{26.50}) + (I_{231,WO3H}) = I_{Total,231}$$

$$(I_{232,WO3} = I_{230,WO3} * \frac{30.64}{26.50}) + (I_{232,WO3H} = I_{231,WO3H} * \frac{14.31}{26.50}) = I_{Total,232}$$

$$(I_{233,WO3} = 0) + (I_{233,WO3H} = I_{231,WO3H} * \frac{30.64}{26.50}) = I_{Total,233}$$

$$(I_{234,WO3} = I_{230,WO3} * \frac{28.43}{26.50}) + (I_{234,WO3H} = 0) = I_{Total,234}$$

$$(I_{235,WO3} = 0) + (I_{235,WO3H} = I_{231,WO3H} * \frac{28.43}{26.50}) = I_{Total,235}$$

For the first peak envelope, the peak areas denoted by $I_{Total,n}$ were calculated based on these equations (where n stands for the m/z ratio). The peak area at each m/z is given as the sum of two contributions: area of peaks corresponding to WO_3^- isotopes and WO_3H^- isotopes.

To prove that the peak envelopes were in fact from contributions of WO_3 and WO_3H, we calculated the values of both peak components using the above equations, and then compared them to the experimental values (see Table A6.1). The errors were minimal for the possibility of both components together.

Table A6. 1 Calculated and experimental raw peak areas of WO_3^- and WO_3H^- peak envelopes.

m/z	230	231	232	233	234	235
Experimental values						
Isotopes	**182.00**	**183.00**	**184.00**	**185.00**	**186.00**	**187.00**
AA 50 nm uncleaned	9726.00	7124.00	12360.00	2245.00	10209.00	2190.00
AA double cleaned	7780.00	5660.00	9673.00	1626.00	8366.00	1561.00
AA triple cleaned	5816.00	4322.00	7189.00	1386.00	6513.00	1356.00
WO_3 contributions calculated						
	182.WO3	**183WO3**	**184WO3**	**185WO3**	**186WO3**	**187WO3**
AA 50 nm uncleaned	9726.00	5252.04	11245.46	0.00	10434.35	0.00
AA double cleaned	7780.00	4201.20	8995.44	0.00	8346.62	0.00
AA triple cleaned	5816.00	3140.64	6724.61	0.00	6239.58	0.00
WO_3H contributions calculated						
AA 50 nm uncleaned	0.00	1941.66	1048.50	2245.00	0.00	2083.07
AA double cleaned	0.00	1406.30	759.40	1626.00	0.00	1508.72
AA triple cleaned	0.00	1198.73	647.31	1386.00	0.00	1286.03
Calculated total peak area (WO_3 + WO_3H contributions)						
	182.00	**183.00**	**184.00**	**185.00**	**186.00**	**187.00**
AA 50 nm uncleaned	9726.00	7193.70	12293.96	2245.00	10434.35	2083.07
AA double cleaned	7780.00	5607.50	9754.84	1626.00	8346.62	1508.72
AA triple cleaned	5816.00	4339.37	7371.93	1386.00	6239.58	1286.03
Errors % (Calculated vs experimental values)						
	182.00	**183.00**	**184.00**	**185.00**	**186.00**	**187.00**
AA 50 nm uncleaned	0.00	-0.98	0.53	0.00	-2.21	4.88
AA double cleaned	0.00	0.93	-0.85	0.00	0.23	3.35
AA triple cleaned	0.00	-0.40	-2.54	0.00	4.20	5.16
Errors (%) considering there is only WO_3 and no WO_3H contributions						
errors	**182.WO3**	**183WO3**	**184WO3**	**185WO3**	**186WO3**	**187WO3**
AA 50 nm uncleaned	0.00	26.28	9.02	100.00	-2.21	100.00
AA double cleaned	0.00	25.77	7.00	100.00	0.23	100.00
AA triple cleaned	0.00	27.33	6.46	100.00	4.20	100.00

Similarly, for the second peak envelope, the following equations were used to calculate the contributions from the WO_4 and WO_4H peaks:

$$(I_{246,WO4}) + (0) = I_{Total,246}$$

$$(I_{247,WO4} = I_{246,WO4} * \frac{14.31}{26.50}) + (I_{247,WO4H}) = I_{Total,247}$$

$$(I_{248,WO4} = I_{246,WO4} * \frac{30.64}{26.50}) + (I_{232,WO4H} = I_{247,WO4H} * \frac{14.31}{26.50}) = I_{Total,248}$$

$$(I_{249,WO4} = 0) + (I_{249,WO4H} = I_{247,WO4H} * \frac{30.64}{26.50}) = I_{Total,249}$$

$$(I_{250,WO4} = I_{246,WO4} * \frac{28.43}{26.50}) + (I_{250,WO4H} = 0) = I_{Total,250}$$

$$(I_{251,WO4} = 0) + (I_{251,WO4H} = I_{247,WO4H} * \frac{28.43}{26.50}) = I_{Total,251}$$

As in the first case, to prove that the peak envelopes were in fact from contributions of WO_4^- and WO_4H^-, we calculated the values of both peak components using the above equations, and then compared them to the experimental values. The errors were minimal when we considered the possibility of both components present together (see Table A6.2).

Table A6. 2 Calculated and experimental raw peak areas of WO_4^- and WO_4H^- peak envelopes.

m/z	246	247	248	249	250	251
Experimental values						
Isotopes	**182**	**183**	**184**	**185**	**186**	**187**
AA 50 nm uncleaned 1	1741	8835	6052	9692	1921	8966
AA double cleaned 1	1270	6739	4457	7070	1401	6636
AA triple cleaned 1	1112	5771	3946	6581	1125	6204
WO_4 contributions calculated						
	182	**183**	**184**	**185**	**186**	**187**
AA 50 nm uncleaned 1	1741	940.14	2012.99	0	1867.797	0
AA double cleaned 1	1270	685.8	1468.408	0	1362.494	0
AA triple cleaned 1	1112	600.48	1285.724	0	1192.987	0
WO_4H contributions calculated						
	182	**183**	**184**	**185**	**186**	**187**
AA 50 nm uncleaned 1	0	8382.441	4526.518	9692	0	8992.936
AA double cleaned 1	0	6114.719	3301.948	7070	0	6560.055
AA triple cleaned 1	0	5691.792	3073.568	6581	0	6106.326
Calculated total peak areas (WO_4 + WO_4H contributions)						
	182	**183**	**184**	**185**	**186**	**187**
AA 50 nm uncleaned 1	1741	9322.581	6539.508	9692	1867.797	8992.936
AA double cleaned 1	1270	6800.519	4770.356	7070	1362.494	6560.055
AA triple cleaned 1	1112	6292.272	4359.291	6581	1192.987	6106.326
Errors % (Calculated vs experimental values)						
	182	**183**	**184**	**185**	**186**	**187**
AA 50 nm uncleaned 1	0	-5.51875	-8.05533	0	2.769528	-0.30042
AA double cleaned 1	0	-0.91289	-7.03065	0	2.748441	1.144432
AA triple cleaned 1	0	-9.03261	-10.4737	0	-6.0433	1.574371
Errors (%) considering there is only WO_4 and no WO_4H contributions						
errors	**182WO4**	**183**	**184**	**185**	**186**	**187**
AA 50 nm uncleaned 1	0	89.35891	66.73843	100	2.769528	100
AA double cleaned 1	0	89.82342	67.0539	100	2.748441	100
AA triple cleaned 1	0	89.59487	67.41704	100	-6.0433	100
Errors (%) considering there is only WO_4H and no WO_4contributions						
	182WO4H	**183**	**184**	**185**	**186**	**187**
AA 50 nm uncleaned 1	100	5.12234	25.20624	0	100	-0.30042
AA double cleaned 1	100	9.263699	25.91545	0	100	1.144432
AA triple cleaned 1	100	1.372522	22.10929	0	100	1.574371