Surface science insight note: A linear algebraic approach to elucidate native films on Fe3O4 surface

P. Bargiela, V. Fernandez, D. Morgan, N. Fairley, J. Baltrusaitis

Surface Science Insight Note: A Linear Algebraic Approach to Elucidate Native Films on Fe_3O_4 Surface

Pascal Bargiela,[1] Vincent Fernandez,[2] David Morgan,[3] Neal Fairley[4] and Jonas Baltrusaitis[5*]

[1]The Institute for Research on Catalysis and the Environment of Lyon (*IRCELYON)*, 2 Avenue Albert Einstein, 69626 Villeurbanne, France

[2]Nantes Université, CNRS, Institut des Matériaux Jean Rouxel, IMN, F-44000 Nantes, France

[3]School of Chemistry, Translational Research Hub, Cardiff University, Maindy Road, Cardiff CF24 4HQ, United Kingdom

and

HarwellXPS – EPSRC National Facility for Photoelectron Spectroscopy, Research Complex at Harwell (RCaH), Didcot, Oxon, OX11 0FA, United Kingdom

[4]Casa Software Ltd, Bay House, 5 Grosvenor Terrace, Teignmouth, Devon TQ14 8NE, United Kingdom

[5]Department of Chemical and Biomolecular Engineering, Lehigh University, 111 Research Drive, Bethlehem, PA 18015, USA

Abstract

Standard materials are often used to obtain spectra that can be compared to those from unknown samples. Spectra measured from these known substances are also used as a means of computing sensitivity factors to allow quantification by X-ray Photoelectron Spectroscopy (XPS) of less well-defined materials. Spectra from known materials also provide line shapes suitable for inclusion in spectral models which, when fitted to spectra, permit the chemical state for a sample to be assessed. Both types of information depend on isolating photoemission signals from the inelastically scattered signal. In this *Insight* note, technical issues associated with the use of XPS of as received Fe_3O_4 powder sample surface are discussed. The *Insight* note is designed to show how linear algebraic techniques applied to data collected from a sample marketed as pure Fe_3O_4 powder are used to verify that XPS has been performed on chemistry representative of the sample. The methods described in this *Insight* note can further be utilized in elucidating complex XPS data obtained from thin films formed or evolved during cyclic/non-steady use of complex (electro)catalyst surfaces, especially in the presence of contaminants.

*corresponding author: job314@lehigh.edu; +1-610-758-6836

Keywords: Linear Algebra, PCA, XPS, iron oxide, abstract factor

Introduction

The merit of X-ray Photoelectron Spectroscopy (XPS) is the ability to measure the information from the upper few nanometers of a sample while providing elemental and chemical state composition. However, the accuracy of quantification obtained by XPS is dependent on the availability of scale factors to transform the raw peak area to the amount of substance [1]. While, in principle, these scale factors are prescribed, in practice for samples of interest such factors can be difficult to obtain and consequently, empirical sensitivity factors, calculated by performing XPS on standard materials [2], are often used. However, a problem encountered when deriving empirical sensitivity factors is that samples presented as standard materials, when measured by XPS are far from the composition of the material specified. The strength of XPS is the root cause of this issue; namely, XPS samples the signal from the topmost few nanometers. While a standard sample may be pure at depths measured in µm, the surface composition often differs from the bulk. The case considered in this *Insight* note is commercially available Fe_3O_4.The X-ray Diffraction Pattern (XRD) and black color of the sample may reflect the expected color for Fe_3O_4, but the chemistry of the surface available to XPS is very different from Fe_3O_4. Consequently, XPS of Fe_3O_4 powder, unless great care is taken, returns spectra that are not representative of this oxide and therefore unsuitable for computing empirical sensitivity factors.

In this *Insight* note, analysis of XPS data acquired from Fe_3O_4 powder surface based on Principal Component Analysis (PCA) [3–5], linear algebra [6] and fitting of peak models to data [4] confirms the chemistry of the surface is not that of Fe_3O_4, yet underneath the surface layers Fe_3O_4 stoichiometry does exist. Helium and argon ion beams are used to remove surface contaminants allowing XPS of these modified surfaces to obtain spectra with characteristics closer to the expected material. An extensive discussion of PCA and linear algebra methods is provided in the Appendix. Terminology related to PCA, used throughout this Insight note, is described in the Appendix. The meaning for PCA abstract factors (AFs), expressed through the mathematical origins for PCA, is explained and placed in the context of XPS. In contexts other than XPS, AFs are described as PCA scores, but AFs are preferred within this insight note for reasons explained within the Appendix.

Sample Analysis Using Ion Beam and Initial Data Processing to Derive PCA Abstract Factors

Commercial Fe_3O_4 (Iron (II, III) oxide nanopowder, 97% from Alfa Aesar, 50-100 nm APS Powder, S.A 8.5-11.5m^2/g) was used as received. Ideally, XPS does not alter sample chemistry but, in practice, changes in surface chemistry do occur, as also demonstrated in this *Insight*. Some samples can materially be altered by interventions necessary during measurement by XPS [4,7,8]. When a sample is relatively stable during the XPS measurement, changes to a sample surface can be induced by changing the temperature of the sample [5] or by interaction with ion beams [9]. When ion beams are used to induce changes to the sample, the extent to which changes occur, amongst other factors, is governed by the mass of the ion in use, the energy of ions and the duration the sample is exposed to the ion beam. The spectra shown in Figure 1 represent the XPS of a sample that evolves in surface chemistry through 1 keV helium ion-beam action. The objective in acquiring these data in Figure 1 is to create a sequence of spectra that evolve from the as-received surface but do so in gradual steps that facilitate the use of PCA and linear algebraic techniques, as implemented in CasaXPS 2.3.26 [10], to identify the chemistry of a Fe_3O_4 sample surface.

The fact that an experiment involving an ion beam causes a change in the sample is not surprising. In fact, as a rule, the role of an ion beam in XPS measurements is to either remove adventitious material from a sample before performing XPS or to etch the surface with an ion beam to uncover buried layers of the material – an option not only commonly used in XPS but also routinely used in Low Energy Ion Scattering (LEIS). The most obvious change to the sample, which can be observed from the spectra in Figure 1, is the attenuation of the C 1s signal. Thus, even for helium ions with an energy of 1 keV, the material is removed from the sample surface. However, Fe 2p, O 1s and valence band spectra are also modified due to the ion beam action. These less obvious changes can be observed and quantified using linear algebra. The use of PCA is illustrated by application to the data in Figure 1. Spectra in Figure 1 are transformed into PCA AFs, whereby the signal is extracted from spectra and presented as AFs that are ordered concerning a statistic that measures variance. The first AF is a curve that fits all spectra in the data set such that the sum of the square differences between the first AF and each spectrum in the data set is a minimum. For

this reason, the first AF for each set of spectra corresponding to C 1s, O 1s, Fe 2p and valence band shown in Figure 2 have the appearance of an average spectrum. Computing PCA AFs is a recursive procedure. Once an AF is computed, the computed AF can be removed (illustrated by the deflation step in NIPALS (Appendix 2, Equation A2.7)) from the set of spectra. The new data set is in a state where the same steps used to calculate the AF can be applied to the new data set to compute the next AF for the original spectra. Figure 2 includes the first four PCA AFs corresponding to the greatest variance of spectra overlaid in Figure 1. The results in Figure 2 are obtained by PCA calculations performed independently for C 1s, O 1s, Fe 2p and valence band data. Each PCA AF with variation in intensity is different from variations expected for noise indicating changes to spectra occurred during the experiment. In particular, the PCA AFs shown in Figure 2 of Fe 2p spectra in Figure 1 include two AFs with shapes distinct from noise. Hence, based on PCA applied to Fe 2p spectra in Figure 1, it is reasonable to assume the surface composition experienced by iron changed throughout the experiment. Further, these changes can be characterized by two distinct shapes within Fe 2p spectra. PCA does not directly provide spectral shapes that would be required to identify physically meaningful spectral forms for Fe 2p spectra but it does show that changes of significance did occur during the experiment.

Further analysis of the beam-modified spectra of Fe_3O_4

The experiment yielding the spectra in Figure 1 was designed so that changes to the sample were limited, yet could still be identified and monitored. Despite XRD indicating the sample is Fe_3O_4, PCA AFs shown in Figure 2 suggest that the XPS of the Fe_3O_4 pellets does not yield spectra typical of pure Fe_3O_4, indicating the composition of both bulk and surface is not identical. In fact, the Fe 2p spectrum measured from the as-received surface (Figure 1, iteration 0) does not have the appearance of Fe 2p photoemission from Fe_3O_4, with the as-received surface exhibiting a satellite at ~719 eV, characteristic of Fe_2O_3 [11,12]. Moreover, iterations of XPS-only measurements using short acquisition times yield one abstract factor (Figure 3) for Fe 2p suggesting the surface oxide can be measured by XPS without alteration. When PCA is applied to the Fe 2p spectra in Figure 1, the existence of two abstract factors shown in Figure 2 is potentially due to a side-effect of the ion beam removing adventitious material, causing alteration of the iron oxidation state. However, ion beam damage is thought not to be entirely responsible for the changes in the spectra since it is possible to

observe a change in the Fe 2p photoemission by XPS measurement alone. An identically prepared Fe_3O_4 pellet was measured iteratively using a Kratos Axis Nova using electron-only charge compensation and X-ray power of 300 W with a measurement duration totaling 6 hours. Similar changes to Fe 2p spectra were observed in the absence of ion beam damage to the sample, as shown by the number of abstract factors (three) computed for these data in Figure 4. The mechanisms responsible for these changes in surface chemistry for iterations of XPS are different from the mechanism causing alteration to surface chemistry by helium ions. While the helium ion beam appears to behave similarly to XPS measurements, changes to the Fe_3O_4 pellet are induced by the use of argon ions. PCA applied to data recorded on a ThermoFisher NEXSA G2, where 200 eV argon ions are used rather than 1 keV helium ions (Figure 5), results in three AFs for Fe 2p spectra. Clearly, in the case of these Fe_3O_4 pellets, argon ions cause greater damage to the chemistry of iron than either excessive use of XPS or normal use of XPS combined with helium ions. The conclusion from these observations is that linear algebra applied to data collected using helium ions offers a means of transforming the spectra to two component spectra capable of reproducing, via linear least squares fitting, all spectra measured except for the argon irradiated sample. In particular, the decomposition into two component spectra provides insight into the chemistry of the as-received Fe_3O_4 pellet.

Implications of the PCA analysis of beam-modified spectra of Fe_3O_4

The evidence described to this point suggests the spectra measured from an as-received Fe_3O_4 pellet are not representative of Fe_3O_4. Using an ion beam to alter the surface may also result in chemistry different from Fe_3O_4. However, given the data shown in Figure 1, it is possible to compute component spectra that cast light on the chemistry measured by XPS. Figure 6 presents three component spectra computed from the data in Figure 1 using C 1s, Fe 2p and O 1s spectra. These three spectral regions are merged to form one spectrum per XPS measurement. Combining narrow scan spectra from C 1s, Fe 2p and O 1s to form a single spectrum per measurement allows the construction of the difference spectra. These difference spectra provide intensity distributions some of which take on shapes that are compatible with physical properties. Three examples of difference spectra computed from the data in Figure 1 are shown in Figure 6. The difference spectrum labeled Phase 1 was chosen because this shape included the Fe 2p signal where the lower-binding-energy onset

of the Fe 2p signal is at an energy above the lower binding energy onset Fe 2p signal in the original spectra. The heuristic used in making this selection was that a signal from iron in a higher oxidation state is expected to occur at higher binding energies than in lower oxidation states. Therefore Phase 1 represents a signal that can be considered from iron in a higher oxidation state, such as the initial Fe_2O_3, likely with some traces of FeOOH, as observed from the O 1s region shoulder at 532.0 eV. The difference spectra labeled Phase 2a and Phase 2b are two alternative options for Fe 2p in a lower oxidation state than Phase 1. Of these two options, the difference spectrum labeled Phase 2b is considered more useful in terms of understanding the proposed composition of the sample. Phase 2a is interesting in the sense that the Fe 2p shape includes narrow, better-defined peak shapes but these shapes occur only when the C 1s signal is distinctly non-physical in shape. A possible interpretation of Phase 2a is that the distribution of the Fe 2p signal is characteristic of ion beam-damaged iron oxide. That is, the process by which the sample evolves is the initial chemistry is systematically changed through growth within the oxide of the material identified by Phase 2a Fe 2p. Phase 2b Fe 2p signal correlates with O 1s and an absence of C 1s signal. That is, Phase 2b, is chosen to minimize the C 1s signal and has the merit that the ratio of Fe 2p to O 1s (Figure 6 d) equals 3:3.9, which is close to the expected 3:4 ratio for Fe_3O_4. The ratio for Phase 1 for iron and oxygen is 2:5 and is therefore clearly very different from the expected ratio for Fe_3O_4. Even accounting for carbon-oxygen chemistry that is included in the difference spectrum in Figure 6 c, the ratio 2:5 implies higher oxidation states for iron than is expected for Fe_3O_4.

The chosen component spectra to a peak model, capable of fitting all spectra from Figure 1, are difference spectra Phase 1 and Phase 2b. Two examples of fits to data based on these two component spectra are shown in Figures 6 e and f. Data used in these example-fits are the as-received Fe_3O_4 pellet and the Fe_3O_4 pellet after 3,507 seconds of irradiation with 1 keV helium ions. It should be noted that similar component spectra to Phase 1 and Phase 2b were obtained through the analysis of data corresponding to PCA AFs shown in Figure 4, which are computed from data collected without the use of an ion beam during the experiment. This observation is intended to emphasize that while helium ions cause an evolution in the Fe_3O_4 pellet, the types of changes in spectra caused by helium ions are very similar to changes induced by XPS alone. The implication is that helium ions are delivering

energy to the sample without excessive impact damage. Hence Phase 2b is chosen for the peak model rather than Phase 2a.

Conclusions

The objective of using PCA is to understand the nature of these changes in a mathematical sense. That is, PCA applied to a set of spectra returns mathematical shapes ordered by variance which convey visually how the signal is partitioned into abstract shapes such that these mathematical shapes are both necessary and sufficient to permit each spectrum in the original data set to be precisely defined by a *unique* linear-combination of these abstract factors. The uniqueness property of PCA AFs (in terms of reconstruction of spectra by forming linear combinations of AFs) is noteworthy since fitting curves to data in the sense of fitting measured spectra from standard materials to data, or when fitting mathematically-defined components in a peak model to XPS data while has the advantage of linking component-spectra to chemical state in the sample, also has the disadvantage of lack of uniqueness. Whenever we construct a set of curves with the intention of fitting data there are choices made in terms of line shapes, background curves and constraints applied to fitting parameters [13]. These choices represent ambiguity and guesswork on the part of the analyst that seldom matches the mathematical rigor offered by PCA decomposition of data into AFs. The consequence of fitting curves to data that do not include a complete set of shapes necessary to reconstruct the data is that least-squares-fitting allows fits to data that may change when the line shapes change [14], even slightly. The disadvantage of PCA AFs is that all AFs are mathematical and do not, in general, identify individual component spectra that represent distinct chemical states for a sample. Nevertheless, PCA does offer insight into how many different curves are required to make sense of a data set.

XPS analysis of a sample with the bulk composition of Fe_3O_4 was performed through iterations of spectral regions. Comparing results from XPS-only experiments and equivalent XPS measurements interleaved with irradiating the sample with ions demonstrates that, depending on ion beam, ion-beam energy and duration of ion-beam action, the as-received sample evolves in composition. PCA of data sets is used to show similarities and differences in data sets, which help to evaluate the sample and create component spectra that describe the evolution of sample chemistry from, the uncertain initial composition to, a composition

more representative of Fe_3O_4. Results obtained when exploring data sets presented herein suggest samples presented as standard material, which in theory should be suitable for computing empirical sensitivity factors or providing XPS line shapes, require careful analysis before concluding these empirical items can be used to understand the chemistry of unknown samples.

Acknowledgments

This work by JB was supported as part of Understanding & Controlling Accelerated and Gradual Evolution of Materials for Energy (UNCAGE-ME), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award # DE-SC0012577. The CNRS is acknowledged for financial support to the Thematic Workshop (N° 1317144) held at the Station Biologique, Roscoff, France.

Author contributions

Pascal Bargiela: Investigation (lead); Writing – review and editing (equal). Vincent Fernandez: Investigation (lead); Writing – review and editing (equal). David Morgan: Methodology (supporting); Writing – review and editing (equal). Neal Fairley: Conceptualization (lead); Investigation (supporting); Methodology (lead); Writing – original draft (equal); Writing – review and editing (equal). Jonas Baltrusaitis: Conceptualization (supporting); Methodology (supporting); Supervision (lead); Writing – original draft (equal); Writing – review and editing (equal).

Competing interests

The authors declare no competing interests.

References

1. Shard, A.G. (2020) Practical guides for x-ray photoelectron spectroscopy: Quantitative XPS. *J. Vac. Sci. Technol. A*, **38** (4), 041201.
2. Wagner, C.D., Davis, L.E., Zeller, M. V, Taylor, J.A., Raymond, R.H., and Gale, L.H. (1981) Empirical atomic sensitivity factors for quantitative analysis by electron spectroscopy for chemical analysis. *Surf. Interface Anal.*, **3** (5), 211–225.
3. Fairley, N., Bargiela, P., Huang, W.-M., and Baltrusaitis, J. (2023) Principal Component Analysis (PCA) unravels spectral components present in XPS spectra of complex oxide films on iron foil. *Appl. Surf. Sci. Adv.*, **17**, 100447.
4. Fernandez, V., Morgan, D., Bargiela, P., Fairley, N., and Baltrusaitis, J. (2023) Combining PCA and nonlinear fitting of peak models to re-evaluate C 1s XPS spectrum of cellulose. *Appl. Surf.*

Sci., **614**, 156182.

5. Garland, B.M., Fairley, N., Strandwitz, N.C., Thorpe, R., Bargiela, P., and Baltrusaitis, J. (2022) A study of in situ reduction of MoO3 to MoO2 by X-ray Photoelectron Spectroscopy. *Appl. Surf. Sci.*, **598**, 153827.

6. Bauer, F.L., Householder, A.S., Wilkinson, J.H., and Reinsch, C. (2012) *Handbook for Automatic Computation: Volume II: Linear Algebra*, Springer, Berlin Heidelberg.

7. Morgan, D.J., and Uthayasekaran, S. (2022) Revisiting degradation in the XPS analysis of polymers. *Surf. Interface Anal.*

8. Baltrusaitis, J., Mendoza-Sanchez, B., Fernandez, V., Veenstra, R., Dukstiene, N., Roberts, A., and Fairley, N. (2015) Generalized molybdenum oxide surface chemical state XPS determination via informed amorphous sample model. *Appl. Surf. Sci.*, **326**, 151–161.

9. Greczynski, G., and Hultman, L. (2021) Towards reliable X-ray photoelectron spectroscopy: Sputter-damage effects in transition metal borides, carbides, nitrides, and oxides. *Appl. Surf. Sci.*, **542**, 148599.

10. Fairley, N., Fernandez, V., Richard-Plouet, M., Guillot-Deudon, C., Walton, J., Smith, E., Flahaut, D., Greiner, M., Biesinger, M., Tougaard, S., Morgan, D., and Baltrusaitis, J. (2021) Systematic and collaborative approach to problem solving using X-ray photoelectron spectroscopy. *Appl. Surf. Sci. Adv.*, **5**, 100112.

11. Bagus, P.S., Nelin, C.J., Brundle, C.R., Crist, B.V., Lahiri, N., and Rosso, K.M. (2021) Combined multiplet theory and experiment for the Fe 2p and 3p XPS of FeO and Fe2O3. *J. Chem. Phys.*, **154** (9).

12. Bagus, P.S., Nelin, C.J., Brundle, C.R., Crist, B.V., Lahiri, N., and Rosso, K.M. (2022) Origin of the complex main and satellite features in Fe 2p XPS of Fe 2 O 3. *Phys. Chem. Chem. Phys.*, **24** (7), 4562–4575.

13. Major, G.H., Fernandez, V., Fairley, N., Smith, E.F., and Linford, M.R. (2022) Guide to XPS data analysis: Applying appropriate constraints to synthetic peaks in XPS peak fitting. *J. Vac. Sci. Technol. A*, **40** (6), 063201.

14. Moeini, B., Linford, M.R., Fairley, N., Barlow, A., Cumpson, P., Morgan, D., Fernandez, V., and Baltrusaitis, J. (2021) Definition of a new (Doniach-Sunjic-Shirley) peak shape for fitting asymmetric signals applied to reduced graphene oxide/graphene oxide XPS spectra. *Surf. Interface Anal.*

15. Walton, J., and Fairley, N. (2005) Noise reduction in X-ray photoelectron spectromicroscopy by a singular value decomposition sorting procedure. *J. Electron Spectros. Relat. Phenomena*, **148** (1), 29–40.

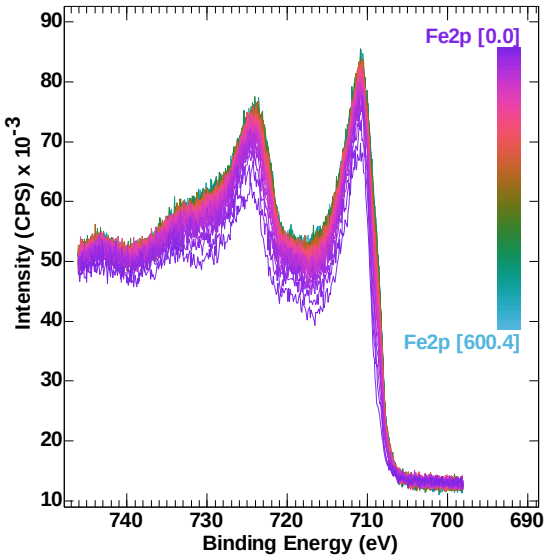
16. Béchu, S., Richard-Plouet, M., Fernandez, V., Walton, J., and Fairley, N. (2016) Developments in numerical treatments for large data sets of XPS images. *Surf. Interface Anal.*, **48** (5), 301–309.

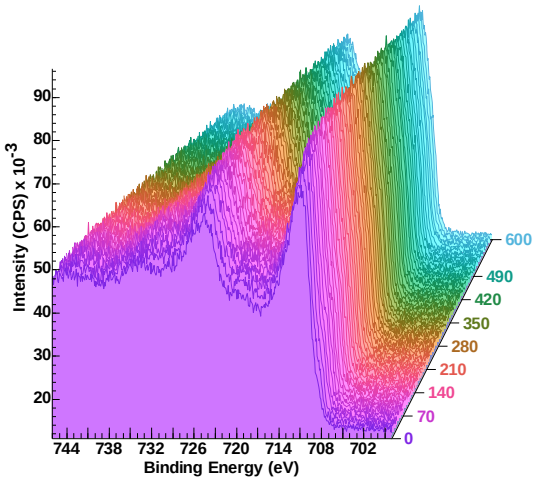
17. E.R. Malinowski (2002) *Factor analysis in chemistry*, John Wiley & Sons, Ltd.

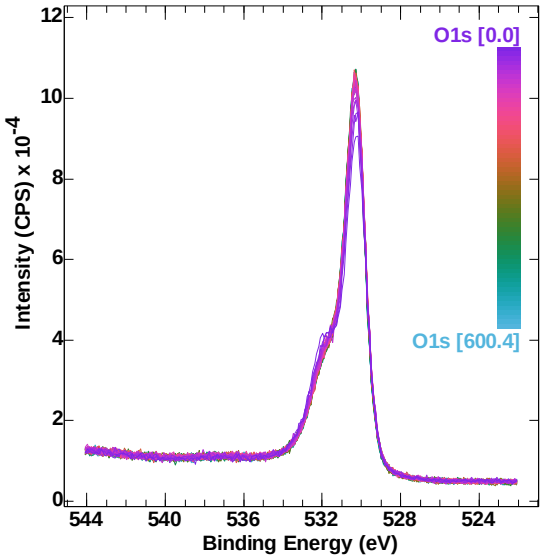
18. Major, G.H., Fairley, N., Sherwood, P.M.A., Linford, M.R., Terry, J., Fernandez, V., and Artyushkova, K. (2020) Practical guide for curve fitting in x-ray photoelectron spectroscopy. *J. Vac. Sci. Technol. A*, **38** (6), 061203.

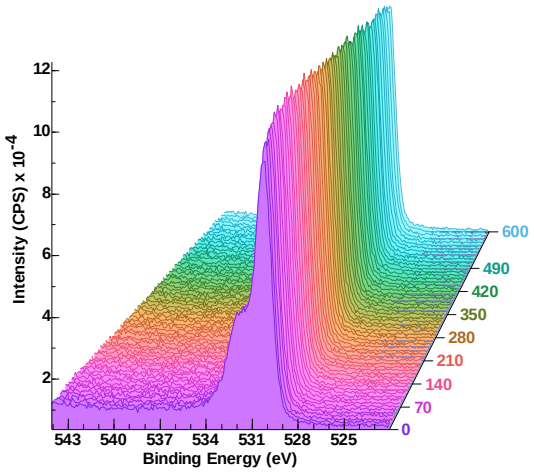
19. Walton, J., and Fairley, N. (2009) Data scaling for quantitative imaging XPS. *Surf. Interface Anal.*, **41** (2), 114–118.

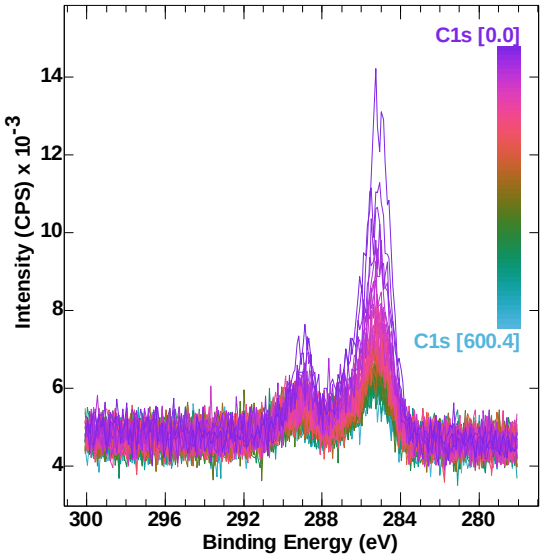
20. Golub, G.H., and Reinsch, C. (1970) Singular value decomposition and least squares solutions. *Numer. Math.*, **14** (5), 403–420.

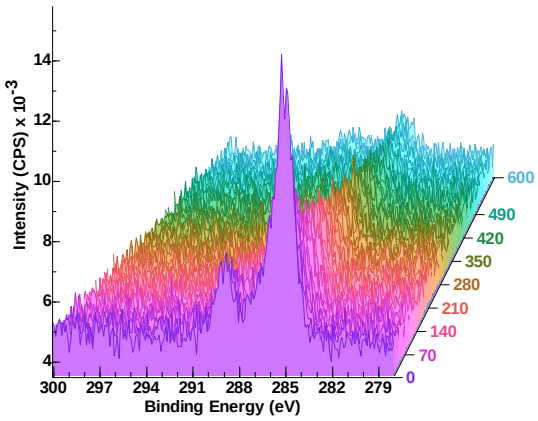

Fe2p [0.0]
Fe2p [600.4]
Intensity (CPS) x 10^-3
Binding Energy (eV)


Intensity (CPS) x 10^-3
Binding Energy (eV)
600
490
420
350
280
210
140
70
0


O1s [0.0]
O1s [600.4]
Intensity (CPS) x 10^-4
Binding Energy (eV)


Intensity (CPS) x 10^-4
Binding Energy (eV)
600
490
420
350
280
210
140
70
0


C1s [0.0]
C1s [600.4]
Intensity (CPS) x 10^-3
Binding Energy (eV)


Intensity (CPS) x 10^-3
Binding Energy (eV)
600
490
420
350
280
210
140
70
0

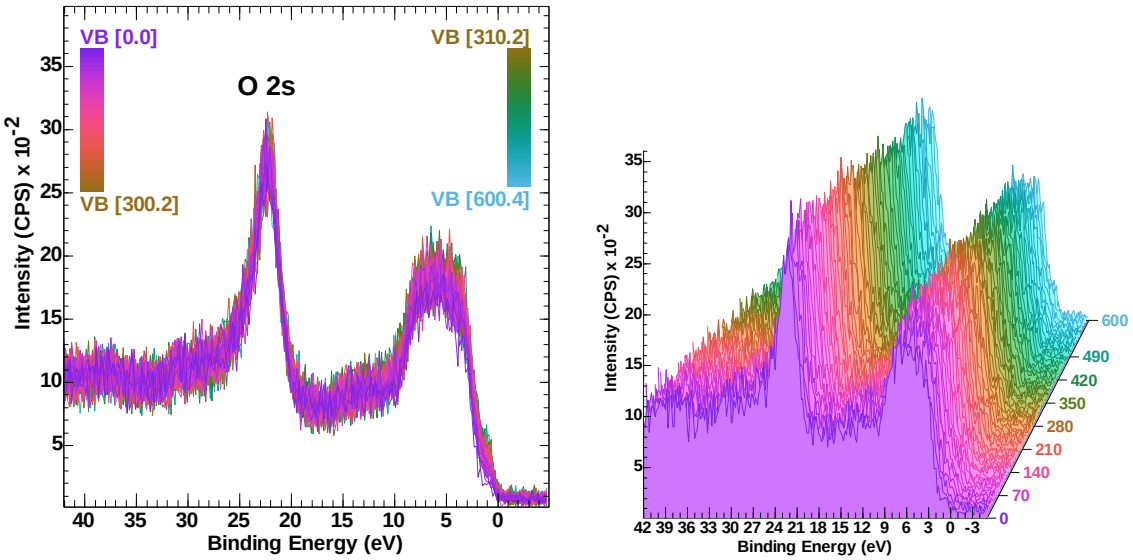


Figure 1. XPS spectra were measured from the sample with bulk composition Fe_3O_4 using ThermoFisher NEXSA G2. These spectra represent iterations of identical XPS interleaved with 1keV helium ion-beam sputter cycles. The variation in color for spectra is representative of the increasing time (seconds) the sample is exposed to helium ion sputtering.

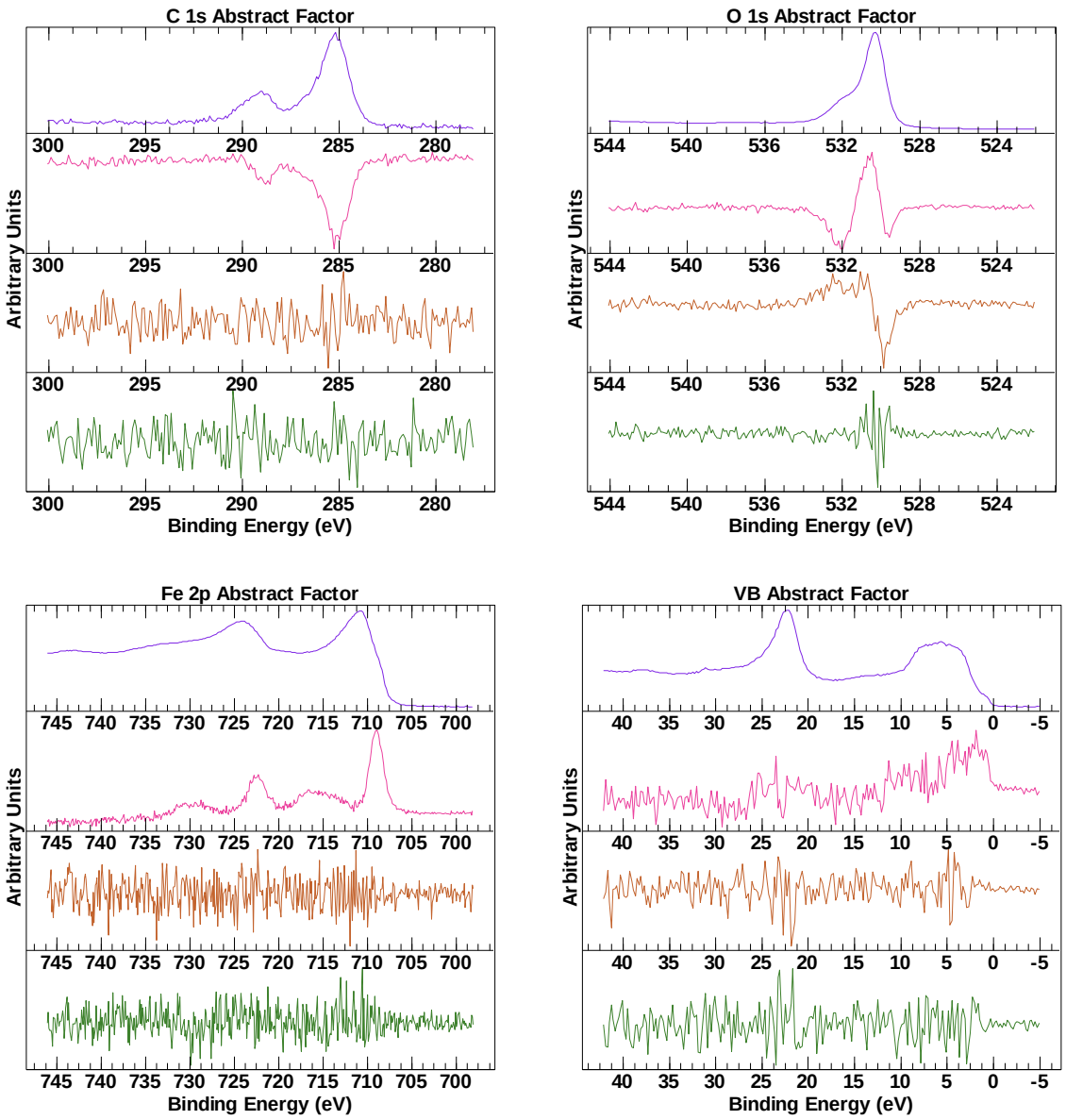


Figure 2. PCA AFs computed for XPS measurements interleaved with helium 1 keV energy ion-beam cycles of 10 seconds per cycle from Fe_3O_4 pellet performed using ThermoFisher NEXSA G2. The number of abstract factors of significance is as follows: C 1s (2), O 1s (3), Fe 2p (2) and VB (2).

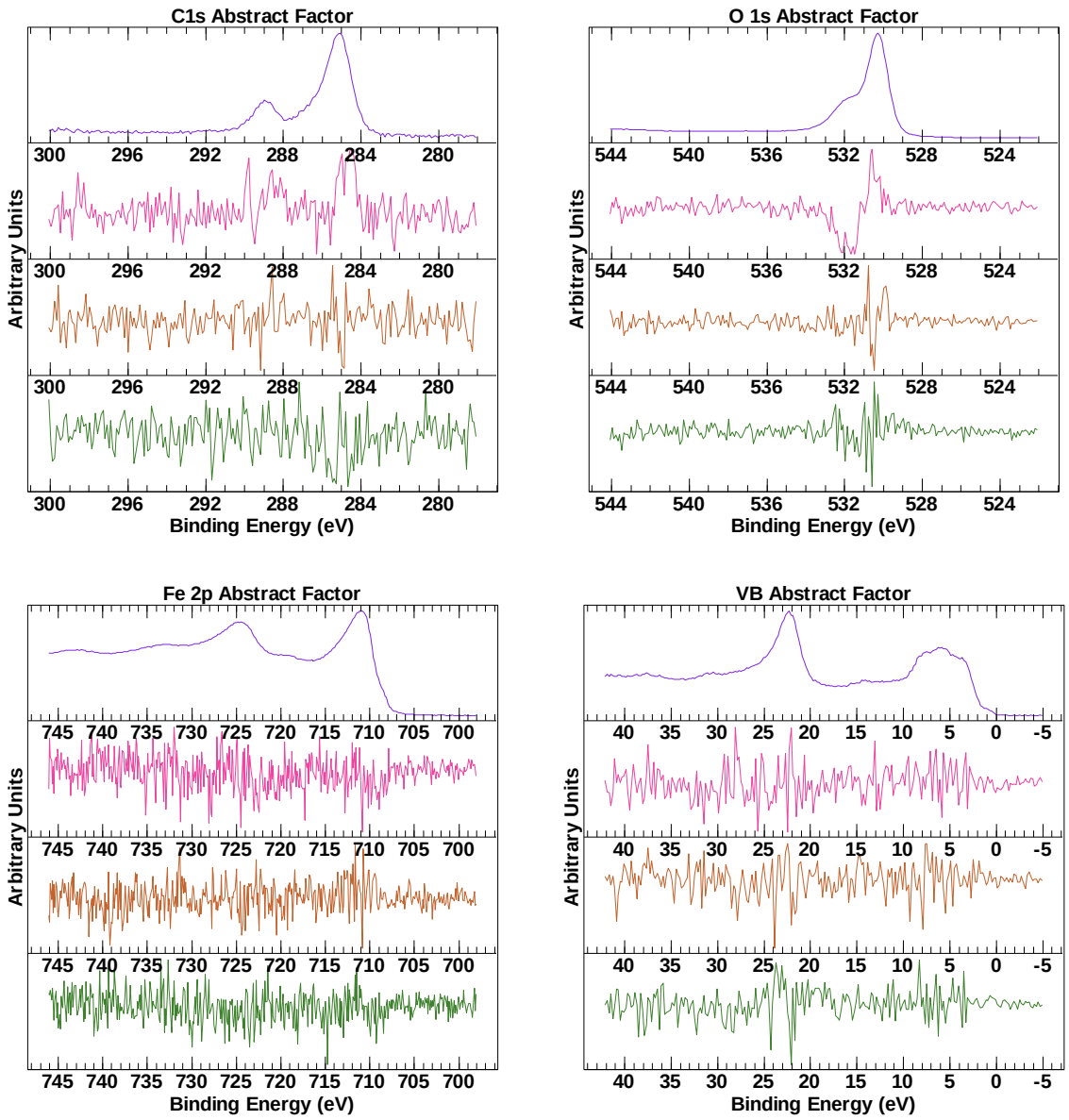


Figure 3. PCA AFs computed for XPS-only measurements from Fe_3O_4 pellet performed using ThermoFisher NEXSA G2. The number of abstract factors of significance are as follows: C 1s (1), O 1s (2), Fe 2p (1) and VB (1).

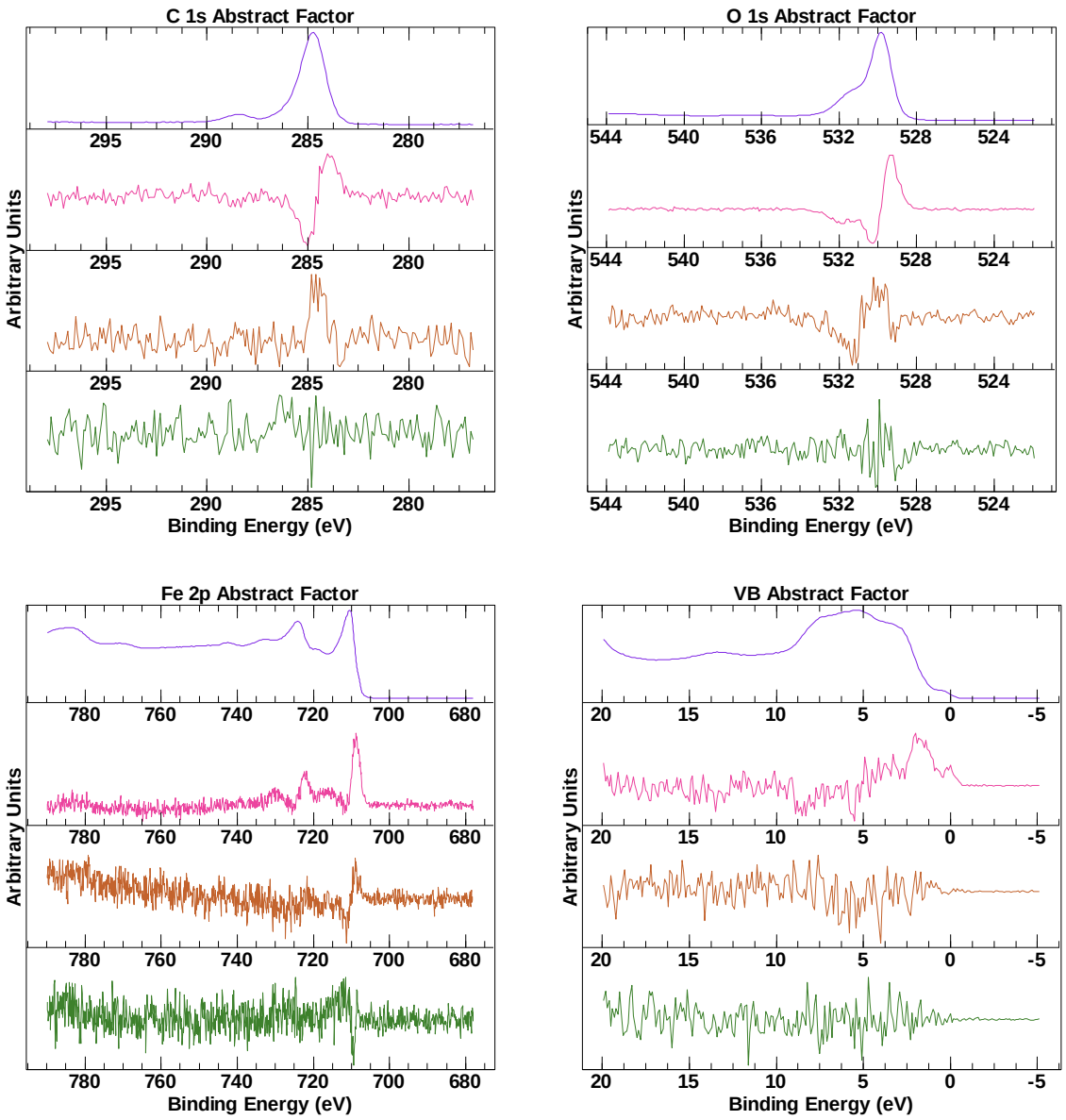


Figure 4. PCA AFs computed for XPS-only measurements from Fe_3O_4 pellet performed using Kratos Axis Nova. The number of abstract factors of significance is as follows: C 1s (3), O 1s (3), Fe 2p (3) and VB (2).

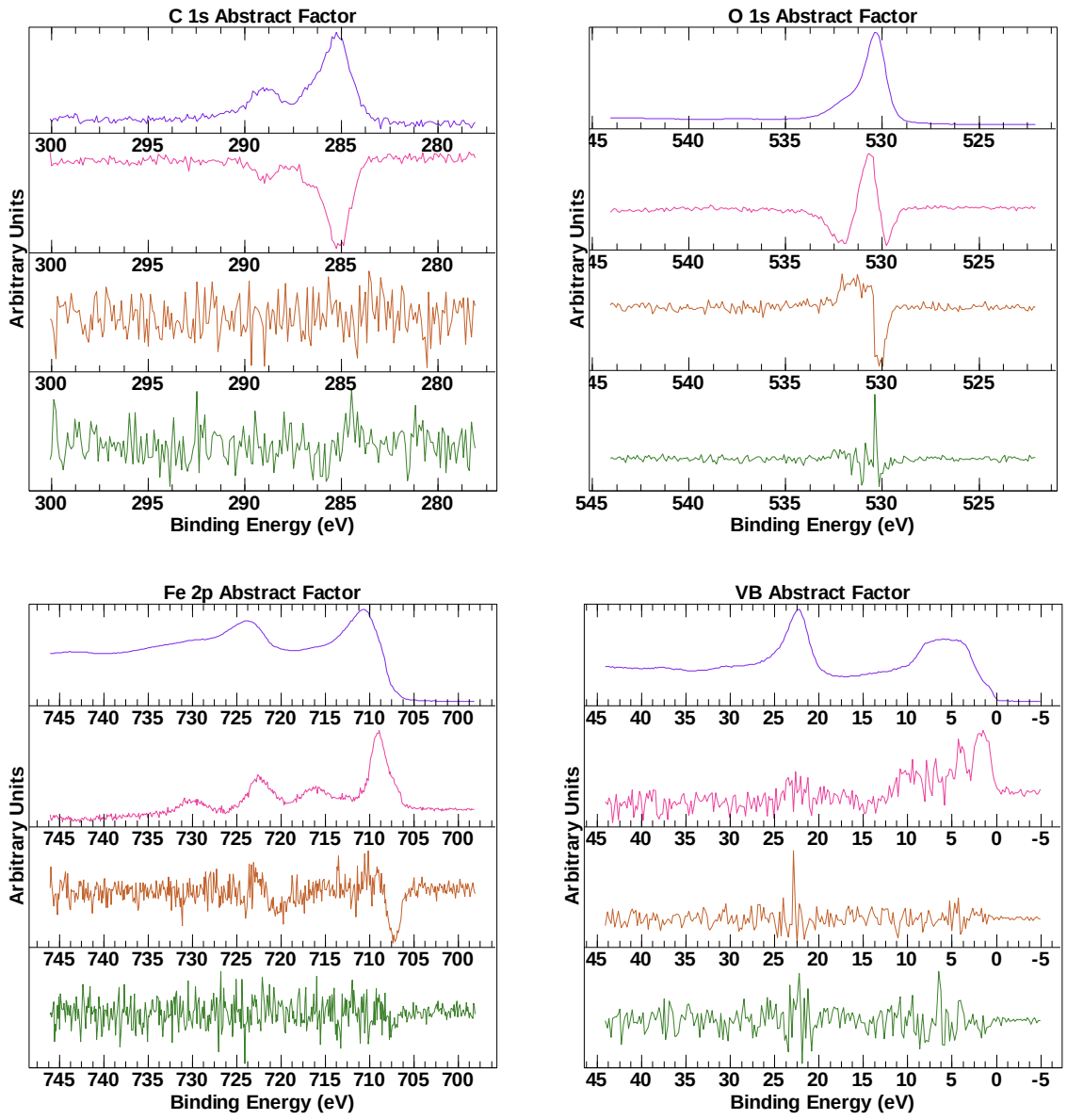
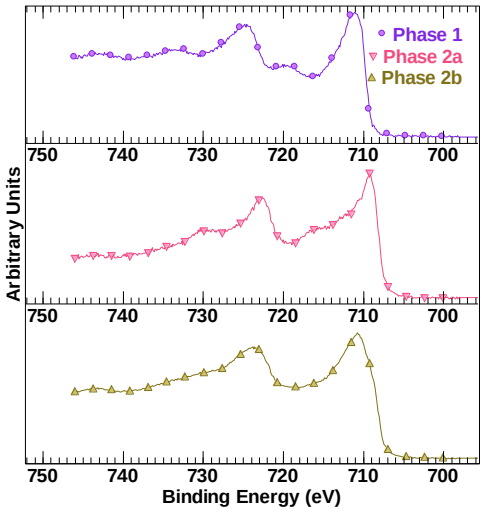
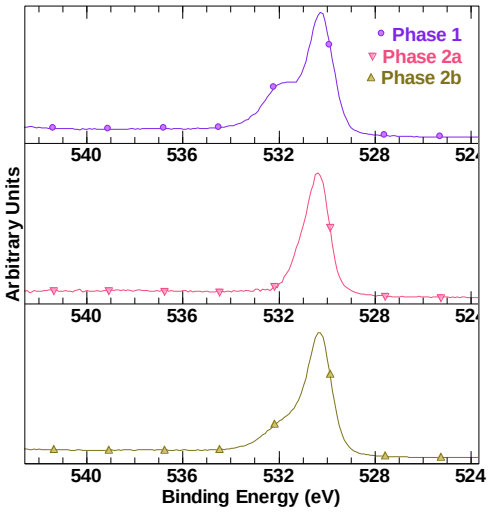


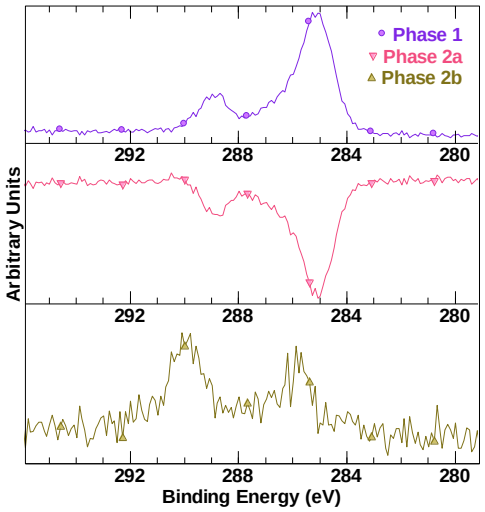
Figure 5. PCA AFs computed for XPS measurements, interleaved with argon 200 eV energy ion-beam cycles (of 10 seconds per cycle), from Fe_3O_4 pellet performed using ThermoFisher NEXSA G2. The number of abstract factors of significance is as follows: C 1s (2), O 1s (3), Fe 2p (3) and VB (2).



(a)



(b)



(c)

Phase 1

Name	At%	Ratio
Fe 2p	22.11	1.00
O 1s	56.03	2.53
C 1s	21.86	0.99

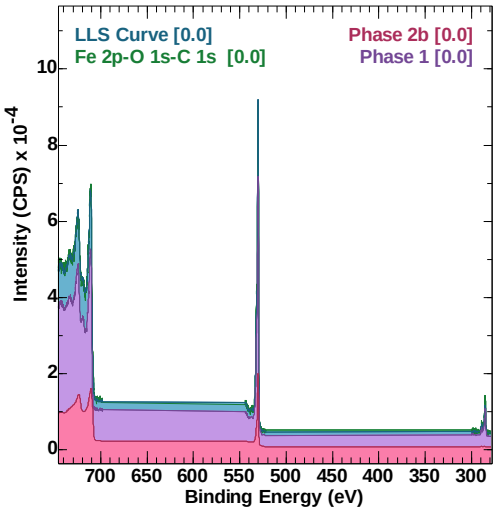
Phase 2a

Name	At%	Ratio
Fe 2p	59.93	1.00
O 1s	40.07	0.67

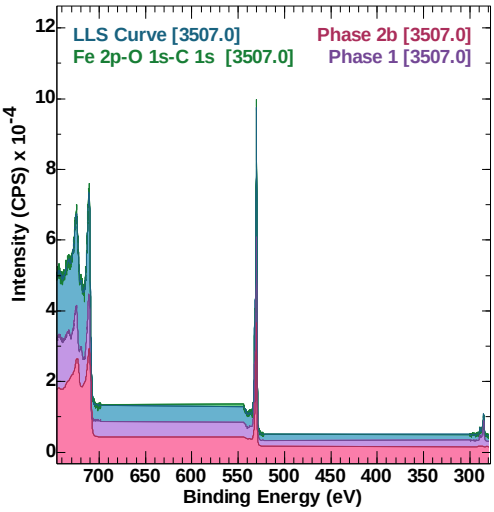
Phase 2b

Name	At%	Ratio
Fe 2p	42.67	1.00
O 1s	55.16	1.29
C 1s	2.18	0.05

(d)



(e)



(f)

Figure 6. Three-component spectra composed of Fe 2p, O 1s and C 1s, computed from data in Figure 1, are presented with atomic concentration tables computed from these component spectra. Two examples of data, where linear least squares fit a model (constructed using two-component spectra) to data, illustrate the quality of fit. a) Component-spectra corresponding to Fe 2p photoemission. b) Component-spectra corresponding to O 1s photoemission. c) Component-spectra corresponding to C 1s photoemission. d) Atomic concentration computed from each of the three component spectra. e) Model spectrum constructed by linear least squares fitting of component spectra Phase 1 and Phase 2b to data that includes signal from Fe 2p, O 1s and C 1s from the as-received surface of the Fe_3O_4 pellet. e) Model spectrum constructed by linear least squares fitting of component spectra Phase 1 and Phase 2b to data that includes signal from Fe 2p, O 1s and C 1s from the surface of the Fe_3O_4 pellet following 3507 seconds of exposure to 1 keV helium ion beam.

Appendix - Review of Principal Component Analysis Concepts

Principal Component Analysis (PCA) is the application of Singular Valued Decomposition (SVD) [6] to experimental data. SVD is a linear algebraic procedure (Appendix 1) that accepts a set of vectors with no ordering metric and transforms them into a set of vectors with a very precise ordering. The computed vectors are mutually orthogonal and one of these vectors points in the direction of greatest variance in the original set of vectors.

When used to process XPS data, the SVD input vectors are either conventional spectra or spectra obtained from images [15]. The vector computed by SVD pointing in the direction of greatest variance in the original set of vectors which, in terms of XPS, is a spectrum that can be described as the principal component of the data set. The precise steps used to compute an SVD are only important so sufficient precision is achieved in an acceptable time. However, some algorithms of SVD are insightful because a sequence of iterations applied to a set of vectors without preprocessing of any kind creates, on termination of the initial sequence of iterations, the principal component. In the process of computing the principal component a set of vectors of reduced dimension is also created and vectors within the newly created set are all orthogonal to the principal component. Therefore, an iterative algorithm separates the initial data set containing *n* vectors into a principal component vector plus a data set of *n-1* vectors. Applying the same iterative steps to the data set containing *n-1* processed vectors returns the principal component for the remaining data plus a data set of dimension *n-2*. Thus, computing SVD iteratively and recursively generates a sequence of vectors in the order of greatest variance concerning the original data set. Hence, SVD generates a new set of *n* vectors that are arranged in order and mathematically describes the original data set since any vector from the original data set can be constructed by forming a linear combination of vectors from the set of vectors constructed by SVD. *Effectively, PCA creates vectors in an order of importance to signal in the data set.*

Altogether, PCA is used to form an alternative perspective of a data set. Principal components are not, in general, physically useful,however, principal components can be of value *en route* to a meaningful interpretation of data. One use of PCA is to enhance signal and suppress noise in a data set. If it is decided that at some point in the SVD algorithm, an iteration returns a principal component that is essentially noise [16], then making use of only the preceding principal components obtained during SVD to form a linear approximation to the original data set, analysis creates vectors that contain mostly signal. The number of principal components deemed to be signals conveys insight into the true composition of spectra or images. That is, the number of principal components indicates the minimum number of spectral shapes required when defining a model with chemical meaning, applicable to all data from an experiment.

The terminology used in this *Insight* note differs from terminology often used relating to PCA. Principal components, as described above, in Factor Analysis [17] are often replaced by the quantities referred to as scores. Scores and loadings are a more general description of relationships found in multivariate problems. XPS use of SVD is more specific. The results of SVD are matrices from which scores are obtained. However, in XPS columns or scores are not independent entities but are governed by correlated photoemission signals. The concept of a principal component in XPS is better understood by describing the $\boldsymbol{U}$ matrix (Appendix 1) as constructed from column vectors that are mathematically equivalent to spectra or images. Hence, rather than talking about scores and loadings, XPS data analysis is better

served by talking about PCA abstract factors (AFs) and coefficients rather than loadings. Loadings are the coefficients in linear combinations of PCA AFs that are used to convert PCA AFs back to the original spectra or images. Thus, loadings in XPS terms are coefficients used to fit PCA AFs to data through least squares optimization. Spectra, curves and fitting of curves to data [18] are more natural to XPS than scores and loadings which lend more to a statistical analysis of data. Hence, PCA is presented here in terms of PCA AFs and how PCA AFs are fitted to spectra to recover the signal from noise.

Accordingly, it is sometimes recommended to perform data scaling and shift preprocessing in preparation for PCA or other algorithms based on linear algebra. The use of these preprocessing steps is not without cost. Namely, the direct connection to spectral shapes is further removed by preprocessing data. In applications other than XPS there may be merit in preprocessing data. However, XPS spectra convey in the shape of peaks and background information used to understand a sample. Even simple preprocessing, such as mean centering of spectra, distorts data from spectroscopic forms and creates a false impression from the outputs of PCA. Mean scaling is a good example of an operation that is performed but serves no real purpose which can be seen by following the mathematical logic for the NIPALS algorithm described in Appendix 2. Mean centering of data is performed by computing the mean average spectrum and then subtracting the mean average spectrum from each of the spectra in the data set. The mean average spectrum approximates the first principal component for a data set. NIPALS computes the first component spectrum and once obtained, via Equation A2.7 removes the first principal component from the data set to resume iterations that compute the second principal component. Effectively, Equation A2.7 is a more sophisticated way to recenter spectra than to subtract off the mean spectrum. The reason one might mean center spectra is based on numerical analysis. Namely, summing data held in floating point precision does not cause a loss of significant digits in the average spectrum. Subtracting the full numerical precision mean average spectrum from all spectra minimizes the loss of significant digits when floating point arithmetic is used during PCA. However, the data set created by mean centering is nothing other than the data set created by the first iteration of NIPALS with the exception that a rather poor approximation to the first principal component was used to reduce the dimension of the data set by unity (Appendix 3). The numerical design of a PCA algorithm is the place to limit the influence of loss of significant digits rather than perturbing the data set from spectroscopic forms before performing PCA. There are, however, some clear advantages to scaling of data that aid in extracting signal from noise that justify distorting spectroscopic or imaging data before an application of PCA [19].

Appendix 1: Singular Valued Decomposition

PCA is an exercise, in constructing a singular value decomposition (SVD) for a data matrix, where a data matrix is an $n \times m$ matrix formed by the n coordinate values for data vectors $\{\boldsymbol{d}_1, \boldsymbol{d}_2, \cdots, \boldsymbol{d}_m\}$. The measure for variance in a data set and the corresponding direction for which variance is shown to be a maximum is encapsulated in the definition of the covariance matrix that is used to construct an SVD for a data matrix $\boldsymbol{D} = [\boldsymbol{d}_1, \boldsymbol{d}_2, \cdots, \boldsymbol{d}_m]$.

Given a set of three data vectors $\{\boldsymbol{d}_1, \boldsymbol{d}_2, \cdots, \boldsymbol{d}_m\}$, the standard procedure for expressing these three vectors, as a set of abstract vectors $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m\}$, is by performing a singular valued decomposition [20] of the data matrix into three matrices in Equation [A1.1].

[A1.1] $$\boldsymbol{D} = \boldsymbol{U}\boldsymbol{W}\boldsymbol{V}^T$$

where

$\boldsymbol{d}_i \in \mathbb{R}^n$, $\boldsymbol{u}_i \in \mathbb{R}^n$, $\boldsymbol{D} = [\boldsymbol{d}_1, \boldsymbol{d}_2, \cdots, \boldsymbol{d}_m]$ and $\boldsymbol{U} = [\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2, \cdots, \hat{\boldsymbol{u}}_m]$

$\boldsymbol{W}$ is a diagonal matrix with diagonal matrix elements equal to the square root of the eigenvalues computed for the covariance matrix

[A1.2] $$\boldsymbol{Z} = \boldsymbol{D}^T\boldsymbol{D}$$

and $\boldsymbol{V}$ is the matrix formed from the normalized eigenvectors of $\boldsymbol{Z}$ ordered concerning the eigenvalues. The columns and rows of the matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ and ordered concerning the magnitude of values that appear along the diagonal of $\boldsymbol{W}$.

Appendix 2: Nonlinear Iterative Partial Least Squares (NIPALS) Algorithm

The NIPALS algorithm is an iterative procedure for generating principal components, where one iterative step calculates one principal component. NIPALS is not the algorithm of choice for computing principal components, but rather has merit in the sense that, the NIPALS procedure is insightful for those wishing to appreciate the logic of PCA without the need to understand the intricacies of linear algebra required by other algorithms.

It might seem odd to use the term nonlinear when naming an algorithm, that involves an application of linear algebra to computing principal components. However, the use of nonlinear in NIPALS is referring to the rate of convergence rather than any nonlinearity in the mathematics. NIPALS is nonlinear in the same sense that the Newton-Raphson method (for finding a root of a function) is nonlinear. Nonlinear, in both cases, relates to the number of significant digits in the value-of-interest, achieved following one iterative cycle of these algorithms. Sadly, in NIPALS like Newton-Raphson, nonlinear convergence is the best-case behavior. Worst-case convergence can be slow. It is worth considering why NIPALS performs well in calculating the first few principal components. One of the most remarkable aspects of the iterative procedure is how, for most data sets, the first eigenvector is obtained with almost no effort. This observation is particularly-true for spectra typical of XPS. The reason for this rapid convergence of NIPALS, when computing the first principal component, is photoemission peaks are superimposed on the background signal, the result of which is the variance in an XPS data set is heavily weighted in favor of the first principal component. The mathematics, described below, explains why weighting in variance, favors or hinders the convergence of NIPALS to a given principal component.

The mathematics of NIPALS is now presented.

Let $\boldsymbol{x}$ be a vector of dimension m, where the dimension for $\boldsymbol{x}$ corresponds to the number of data vectors $\boldsymbol{d}_i$ available to the NIPALS algorithm. A matrix $\boldsymbol{A} = [\boldsymbol{d_1}, \boldsymbol{d_2}, \boldsymbol{d_3}, \ldots, \boldsymbol{d_m}]$ is formed from these data vectors $\boldsymbol{d}_i$, each of which has n acquisition channels.

If it is assumed a set of m eigenvectors $\boldsymbol{p}_i$ exist corresponding to the covariance matrix $\boldsymbol{A}^T\boldsymbol{A}$

[A2.1] $$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{p}_i = w_i\boldsymbol{p}_i$$

and all else being equal, the eigen Equation A2.1 yields m orthonormal eigenvectors. Since these eigenvectors for a real symmetric matrix form a basis set it is possible to write

[A2.2] $$\boldsymbol{x} = \sum_{i=1}^{m} \alpha_i \boldsymbol{p}_i$$

Thus, for any vector $\boldsymbol{x}$, performing matrix multiplication of $\boldsymbol{x}$ by the matrix $\boldsymbol{A}^T\boldsymbol{A}$ can be expressed in terms of eigenvalues and eigenvectors of $\boldsymbol{A}^T\boldsymbol{A}$ as follows.

[A2.3] $$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{A}\left(\sum_{i=1}^{m}\alpha_i\boldsymbol{p}_i\right) = \left(\sum_{i=1}^{m}\alpha_i\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{p}_i\right) = \left(\sum_{i=1}^{m}\alpha_i w_i\boldsymbol{p}_i\right)$$

Therefore, repeated multiplication by $\boldsymbol{A}^T\boldsymbol{A}$ results in a transformation of $\boldsymbol{x}$ as follows.

[A2.4] $$(\boldsymbol{A}^T\boldsymbol{A})^{\boldsymbol{k}}\boldsymbol{x} = \left(\sum_{i=1}^{m}\alpha_i {w_i}^k\boldsymbol{p}_i\right)$$

Thus, by repeatedly multiplying a vector by the matrix $\boldsymbol{A}^T\boldsymbol{A}$, the eigenvalues are raised to a power. If one eigenvalue is larger than all others, that factor will dominate the summation term. Hence the resulting vector due to iterations is the eigenvector corresponding to the largest eigenvalue.

A further consideration derives from Equation A2.2. Since $\boldsymbol{p}_i$ belongs to a set of orthonormal eigenvectors $\boldsymbol{p}_i.\boldsymbol{p}_j = \begin{cases}1 & i = j\\ 0 & i \neq j\end{cases}$ and therefore the coefficients in Equation A2.2 are computed as follows $\alpha_i = \boldsymbol{x}.\boldsymbol{p}_i$.

If $\boldsymbol{x} \rightarrow \boldsymbol{p}_1$ then $\alpha_1 \rightarrow \|\boldsymbol{x}\|$ and $\alpha_{i\neq 1} \rightarrow 0$. These relationships provide an alternative perspective for the converging sequence of vectors. In the event the eigenvalues differ only marginally resulting in a slow movement towards the largest of these similar eigenvalues, a good guess for the initial eigenvector is important to obtain convergence within a reasonable number of iterations.

NIPALS includes iterative steps which rely on these types of transformations. However rather than forming a covariance matrix $\boldsymbol{A}^T\boldsymbol{A}$ the data matrix $\boldsymbol{A}$ is used to transform vectors in a sequence leading to the computation of $\boldsymbol{u}_1$, the abstract factor corresponding to the largest eigenvalue of $\boldsymbol{A}^T\boldsymbol{A}$.

The essential iterative steps performed during a NIPALS are as follows.

The input to an iterative step is a matrix $\boldsymbol{D}_i$, where initially $\boldsymbol{D}_1 = \boldsymbol{A}$. The output from each iterative sequence makes use of $\boldsymbol{D}_i$ is a vector $\boldsymbol{u}_i$ and a matrix $\boldsymbol{D}_{i+1}$. These steps involve selecting an initial vector $\boldsymbol{y}_0$ of dimension n . A vector $\boldsymbol{x}_0$ of dimension m is computed using the following operations.

[A2.5] $$\boldsymbol{D}_i{}^{\boldsymbol{T}}\hat{\boldsymbol{y}}_j = \boldsymbol{x}_j$$

[A2.6] $$\boldsymbol{D}_i\hat{\boldsymbol{x}}_j = \boldsymbol{y}_{j+1}$$

where $\hat{\boldsymbol{a}}$ is the unit vector corresponding to the vector $\boldsymbol{a}$. These steps represent a vector multiplied by a matrix rather than directly constructing a covariance matrix. Separating the action of the covariance matrix into these two intermediate steps is advantageous if convergence to the desired vector is rapid as it is designed to avoid matrix multiplication.

A sequence of vectors $\{\boldsymbol{y}_j\}$ is constructed which converges to the vector $\boldsymbol{u}_i$, corresponding to the largest eigenvalue of ${\boldsymbol{D}_i}^T\boldsymbol{D}_i$. Once the vector $\boldsymbol{u}_i$ is established the ultimate operation for a single iteration is to deflate $\boldsymbol{D}_i = \left[\boldsymbol{d}_1^i, \boldsymbol{d}_2^i, \boldsymbol{d}_3^i, \cdots, \boldsymbol{d}_m^i\right]$ using the computed vector $\boldsymbol{u}_i$

[A2.7] $$\boldsymbol{d}_j^{i+1} = \boldsymbol{d}_j^i - \frac{\boldsymbol{d}_j^i.\boldsymbol{u}_i}{\boldsymbol{u}_i.\boldsymbol{u}_i}\boldsymbol{u}_i$$

resulting in the next matrix in the NIPALS sequence $\boldsymbol{D}_{i+1}$. The new matrix represents a set of vectors all belonging to a subspace of dimension one less than the previous step. The projection operation in Equation A2.7 is found also in Gram-Schmidt, a procedure for constructing a set of mutually orthogonal vectors. Deflating these $\boldsymbol{D}_i$ matrices permit the next iteration of the NIPALS procedure to target the next largest eigenvalue and hence compute the next vector $\boldsymbol{u}_{i+1}$.

Appendix 3: Mean Centre of Data

Given data vectors $\{\boldsymbol{d}_1, \boldsymbol{d}_2, \cdots, \boldsymbol{d}_m\}$, the mean spectrum $\bar{\boldsymbol{d}}$ is computed using Equation A3.1.

[A3.1] $$\bar{\boldsymbol{d}} = \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{d}_i$$

A set of mean centred vectors is obtained by computing $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_m\}$, where $\boldsymbol{c}_i = \boldsymbol{d}_i - \bar{\boldsymbol{d}}$. However, if we consider $\boldsymbol{c}_1$, for example, by definition $\sum_{i=1}^{m} \boldsymbol{c}_i = 0$, therefore $\boldsymbol{c}_1 = -\sum_{i=2}^{m} \boldsymbol{c}_i$. Hence $\boldsymbol{c}_1$is a linear combination of $\{\boldsymbol{c}_2, \boldsymbol{c}_3, \cdots, \boldsymbol{c}_m\}$. Therefore, the set $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_m\}$ is of dimension one less than the dimension of the set of vectors $\{\boldsymbol{d}_1, \boldsymbol{d}_2, \cdots, \boldsymbol{d}_m\}$. Moreover, the set of vectors $\{\bar{\boldsymbol{d}}, \boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_m\}$ is of the same dimension as $\{\boldsymbol{d}_1, \boldsymbol{d}_2, \cdots, \boldsymbol{d}_m\}$, hence any spectrum in the original data set can be constructed from the abstract factors $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m\}$ computed from $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_m\}$ by reversing the offset $\bar{\boldsymbol{d}}$ using Equation A3.2.

[A3.2] $$\boldsymbol{c}_j = \sum_{i=1}^{m} \alpha_{ij}\boldsymbol{u}_i \text{ and } \boldsymbol{d}_j = \boldsymbol{c}_j + \bar{\boldsymbol{d}}$$