Peak Fitting in XPS

Small and sometimes not so small differences between the initial and final state of an atom when a core level electron is excited by an x-ray is fundamental to XPS as an analytical technique. These variations in the initial and final state energy are due to the environment in which an atom is found and results in electrons ejected from the same element emerging with kinetic energy characteristic of the chemical state in which the element exists. In terms of an XPS spectrum, the increase in counts as a function of kinetic energy associated with the excitation of a core level electron appears as an ensemble of peaks rather than a single peak. These chemically shifted peaks offer information about the chemistry of the surface. Consider the XPS spectrum (Figure 1) of the carbon 1s electrons measured from a nylon sample; the chemical formula for nylon [1] indicates four chemical environments for a carbon atom. The C 1s spectrum clearly contains two chemically shifted C 1s peaks; however the more subtle shifts associated with the peaks labelled CH2 are the reason peak fitting is an important tool in XPS.
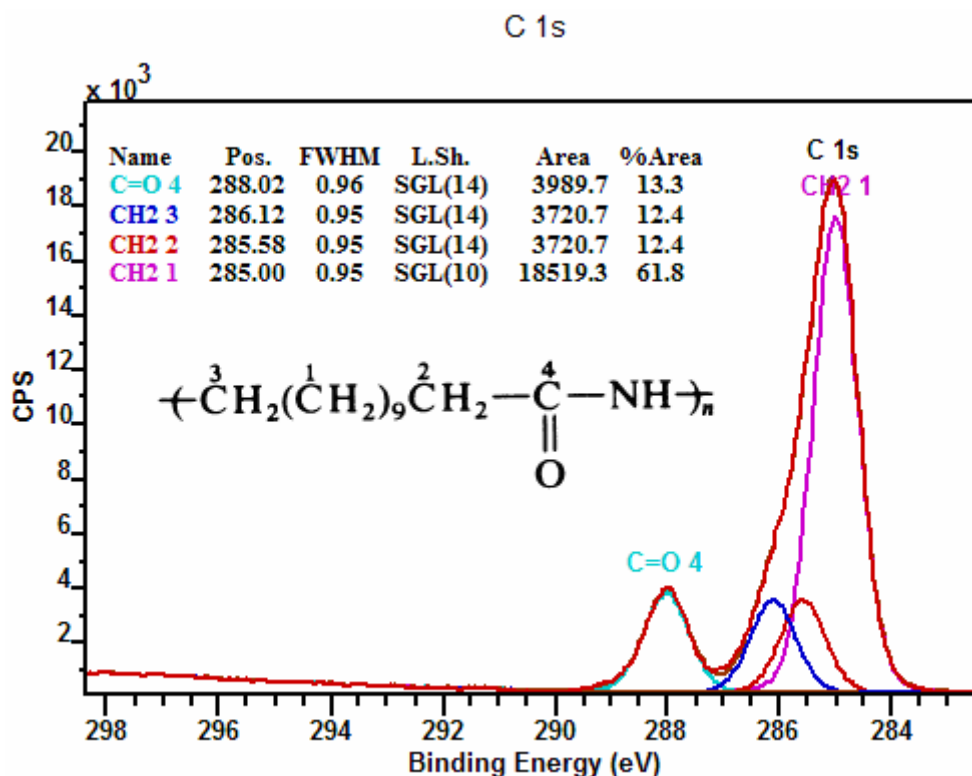

Figure 1: C 1s region measured using a Kratos Axis 165 from a nylon sample.

Sadly, while central to XPS, peak fitting of line-shapes to spectra is far from simple and if treated as a black-box tool will almost always yield incorrect results. The problem is

that a good fit is always achieved by a sufficient number of Gaussian-Lorentzian curves when optimized without constraints. The fit in Figure 1 is guided by the chemical formula for nylon. Understanding the chemistry is important as it suggests the number of chemical states and therefore number of peaks in this example is four; introducing parameter constraints to restrict the peak widths and relative intensities of the peaks force the peak model to obey the chemistry. Without these inputs any model designed purely on the spectral envelope would be a cause for concern.

When peak-fitting XPS spectra a further issue is the nature of the background signal on top of which the synthetic peaks must sit. The data in Figure 1 represents a relatively simple case, where most analysts would use a linear background approximation, however in general, the background to XPS peaks are far from simple. Figure 2 illustrates the rapid changes to the background resulting from energy loss processing occurring as the photoelectrons are ejected from the surface material. These background shapes are dependent on the material under analysis and significant variation occurs in practice. As a consequence backgrounds other than simple linear interpolation of the intensities at either end of an energy interval are required.
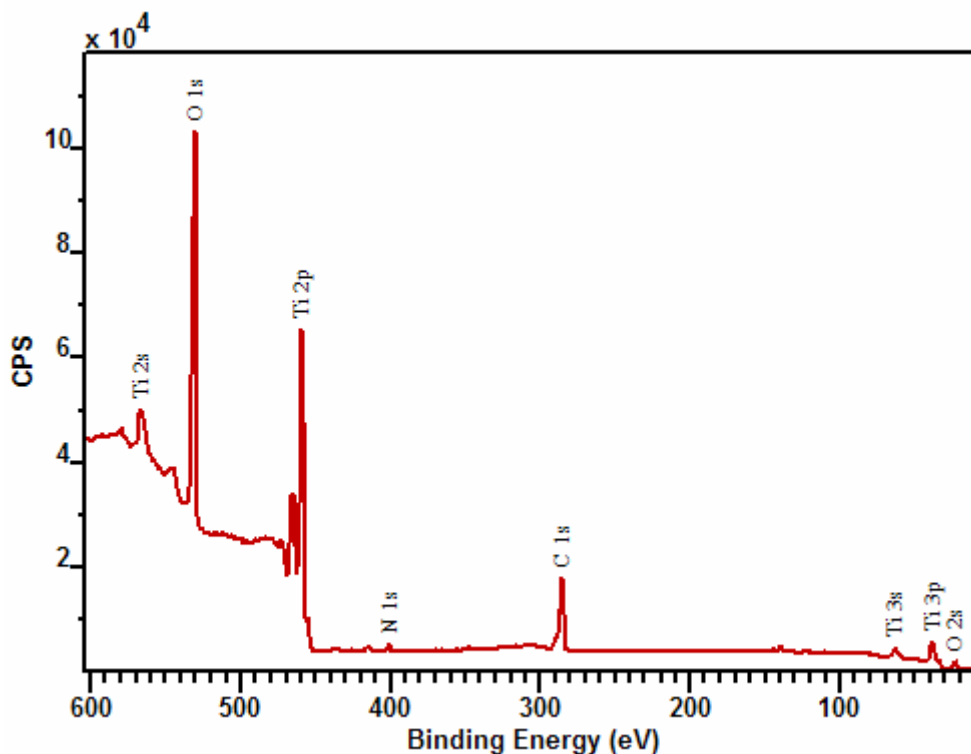

Figure 2: Ti spectrum showing rapid background variation. Data supplied by Elise Pegg, Bioengineering Group, University of Nottingham.

All the peaks modeled in Figure 1 are due to chemical shifts. Electrons ejected from s-subshells generally appear as chemically shifted *primary* peaks. Photoelectric lines

resulting from the excitation of non s-subshell electrons appear in pairs and are related in their intensities. Electrons ejected from core levels with symmetries defined by p, d, f, … angular momenta may leave the core excited ion in one of two (XPS observable) states. These doublet states are characterized by the $j = l \pm \frac{1}{2}$ quantum number, which define the multiplicity of the state, namely $2j+1$. The relative intensities of these doublet pairs are therefore $(2(l- \frac{1}{2})+1) : (2(l+ \frac{1}{2})+1)$, thus for p electrons ($l = 1$) the relative intensities are $1:2$, while for d electrons the doublet pairs are in the proportion $2:3$ and for f electrons the ratio is $3:4$. The separation of these doublet pairs varies with angular momentum and principal quantum number, which is illustrated in Figure 3 for those photoelectrons ejected from clear gold metal. The energy separation of these doublet pairs depends on both the principal and angular momentum quantum numbers of the core level electrons, and can result in widely separated peaks such as the Au 4p doublet pair indicated in Figure 3. Equally, these doublets can be overlapping such as the Au 4f lines. It should be pointed out that not all peaks in XPS spectra are due to simple photoelectric transitions. *Secondary* peaks may appear in a spectrum due to Auger transitions, plasmon structures, shake-up and shake-off energy loss processes.
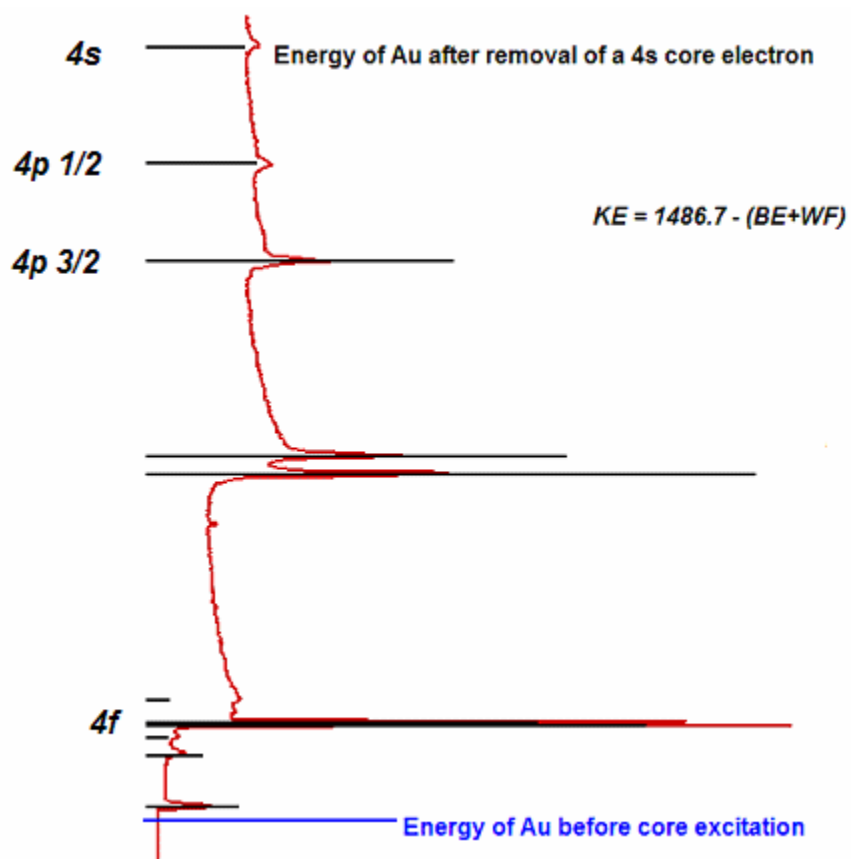

Figure 3: Alternative perspective of an Au XPS spectrum acquired using an aluminium x-ray anode.

Backgrounds to spectra containing both doublet pairs coupled with a variety of chemical shifts represent the greatest challenge to modeling XPS spectra, since without a proper

description of the background a good fit to the data obeying the chemistry and physics involved is hard to achieve. Backgrounds to XPS spectra are therefore an important part of peak fitting XPS data.

Backgrounds to XPS Spectra

There are numerous backgrounds on offer in CasaXPS, however for most analysts the basic linear, Shirley and universal cross-section Tougaard backgrounds are the tools of choice. Nevertheless, the wide variety of backgrounds available is a measure of the dissatisfaction often felt about using these basic shapes. 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.

Defining the background parameters

Backgrounds are, in general, computed to ensure the background meets the data at the limits of the energy interval defining a set of peaks. Sometime the intensity from a single data channel is susceptible to noise or perhaps not truly representative of the relationship between the background and the data. Under these circumstances the intensities I_1 and I_2 of the background limits at the end points E_1 and E_2 (Figure 4) can be modified using the **Av Width** parameter and the **St Offset** and **End Offset** parameters on the Regions property page. When the **Av Width** is "0", the values I_1 and I_2 are set equal to the intensity of the spectrum at the data channel closest to the energies E_1 and E_2. If the **Av Width** is greater than zero, the number of channels specified in the **Av Width** to the left and right of the data bin otherwise used are averaged to determine the intensities used to compute the background. These intensities can be further adjusted using the **St Offset** and **End Offset** parameters to reduce the intensities required to calculate the background beneath the peaks. The **St Offset** and **End Offset** parameters are percentage offsets from the original intensities before any offset is applied. A value of "0" means no offset while "100" means the background at the end point is zero.

The backgrounds in CasaXPS will now be described.

Linear Background: BG Type "linear" or "l"

Polymers and other materials with large band-gaps tend to have a relatively small step in the background over the energy range covered by the peaks. The nylon spectrum in Figure 1 is an example where the background to the C 1s peaks can be approximated by a linear background type:

$$L(E) = I_1 \frac{(E_2 - E)}{(E_2 - E_1)} + I_2 \frac{(E - E_1)}{(E_2 - E_1)}$$

where E_1 and E_2 are two distinct energies and I_1 and I_2 are two intensity values usually chosen to cause the background to merge with the spectral bins at E_1 and E_2 (Figure 4).
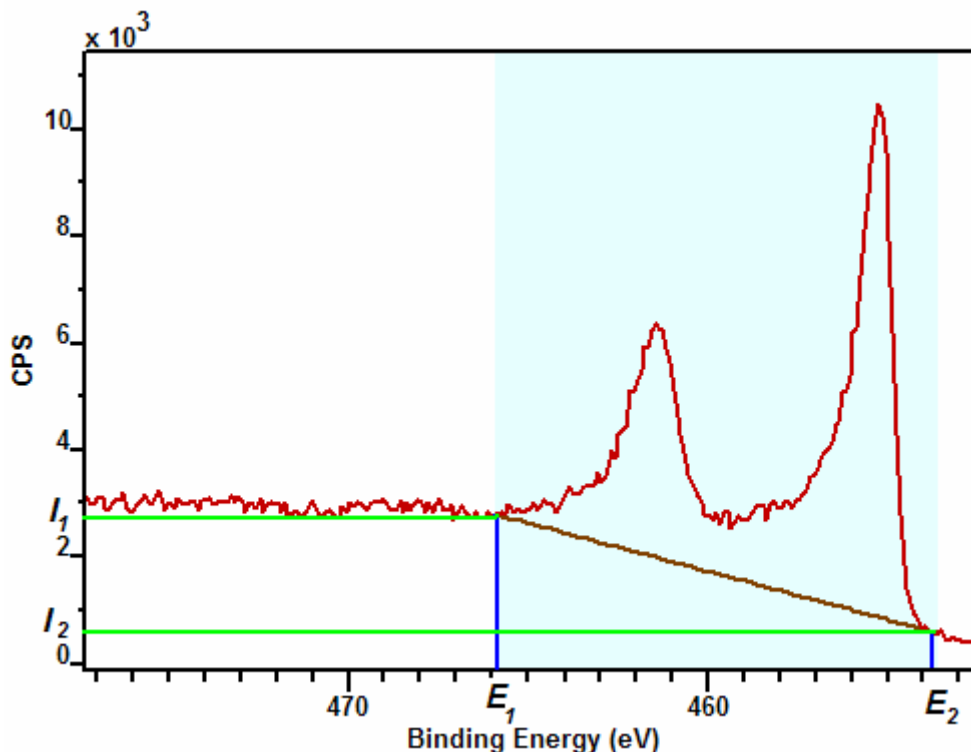

Figure 4: Example of a linear background type applied to a Ti 2p doublet pair. Data supplied by Elise Pegg, Bioengineering Group, University of Nottingham.

Shirley Background: BG Type "Shirley" or "s"

The Shirley algorithm [2] is an attempt to use information about the spectrum to construct a background sensitive to changed in the data. The essential feature of the Shirley algorithm is the iterative determination of a background using the areas marked *A1* and *A2* in Figure 5 to compute the background intensity *S(E)* at energy *E*:

$$S(E) = I_2 + \kappa \frac{A2(E)}{(A1(E) + A2(E))}$$

where κ defines the step in the background and is typically equal to $(I_1 - I_2)$. Clearly the quantities *A1(E)* and *A2(E)* are known provided the background *S(E)* is already known and since *S(E)* is initially unknown, the calculation of a Shirley background from spectral data is an iterative procedure. That is, the integrated areas *A1(E)* and *A2(E)* for each point on the background *E* must initially be computed using an approximation to *S(E)*, then

refined using the background computed from the first approximation as input to improve the values computed for *A1(E)* and *A2(E)*.
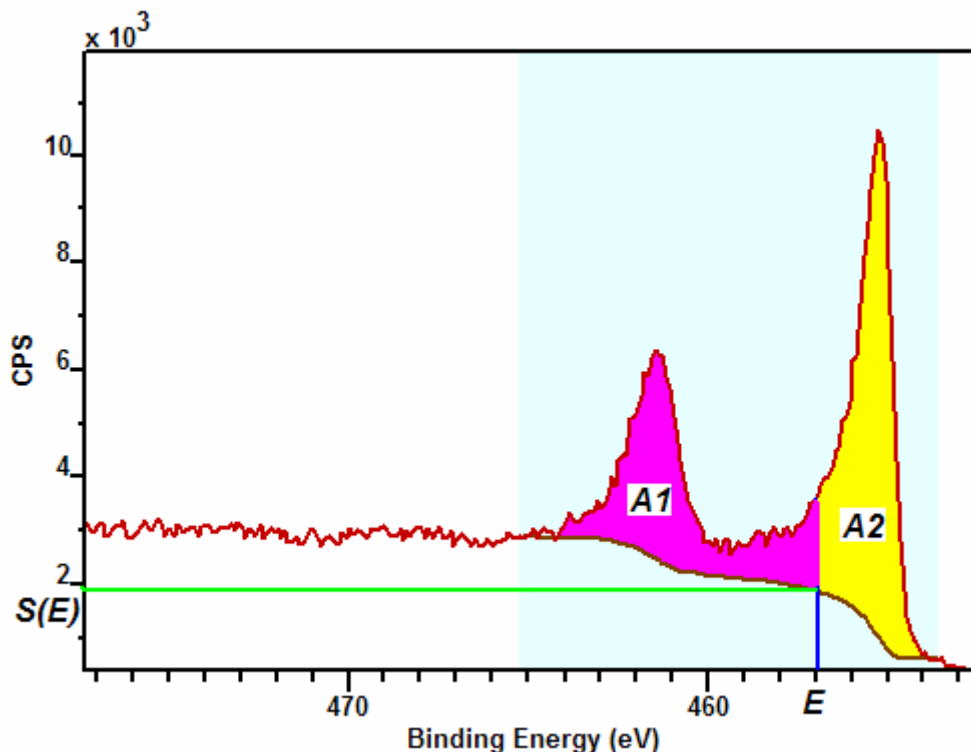

Figure 5: A Shirley background computed from a Ti 2p spectrum.

The background in Figure 5 accommodates a Ti 2p doublet pair using a single energy interval and therefore a single background calculation for both peaks. There are situations where the separation of a doublet pair makes it inappropriate to use a Shirley background to span both peaks using a single background definition. Figure 6 shows a Ti 2p spectrum in which two regions are used to compute Shirley backgrounds independently for each peak. The area beneath the peaks is clearly different for these two strategies where the difference between computing the background for the individual peaks and performing the calculation for the combined peaks is marked in Figure 6 with the letter *A*. The relative intensity of these two peaks must be 1:2, which is indeed the case for the integrated areas *p1:p2* measured using the two regions and displayed in the table over the data in Figure 6. If the ratio of the peaks is wrong, then the background is probably wrong too, but sadly if the ratio is correct, this does not necessarily imply the background is correct. It is entirely feasible that with appropriate asymmetric line-shapes, the background calculated for the combined peaks could yield similarly correct intensity ratios for the doublet pair.

The situation becomes more involved when a Shirley background is used for spectra resulting from chemically shifted doublet pairs as illustrated in Figure 7. A common

problem with using a Shirley background over such an extended range is that the algorithm produces a background curve that cuts through the data. Under these circumstances the Shirley background in clearly nonphysical and therefore wrong. Nevertheless, there are situations where an estimate for the peak areas is required and so variations on a theme are offered as a solution.
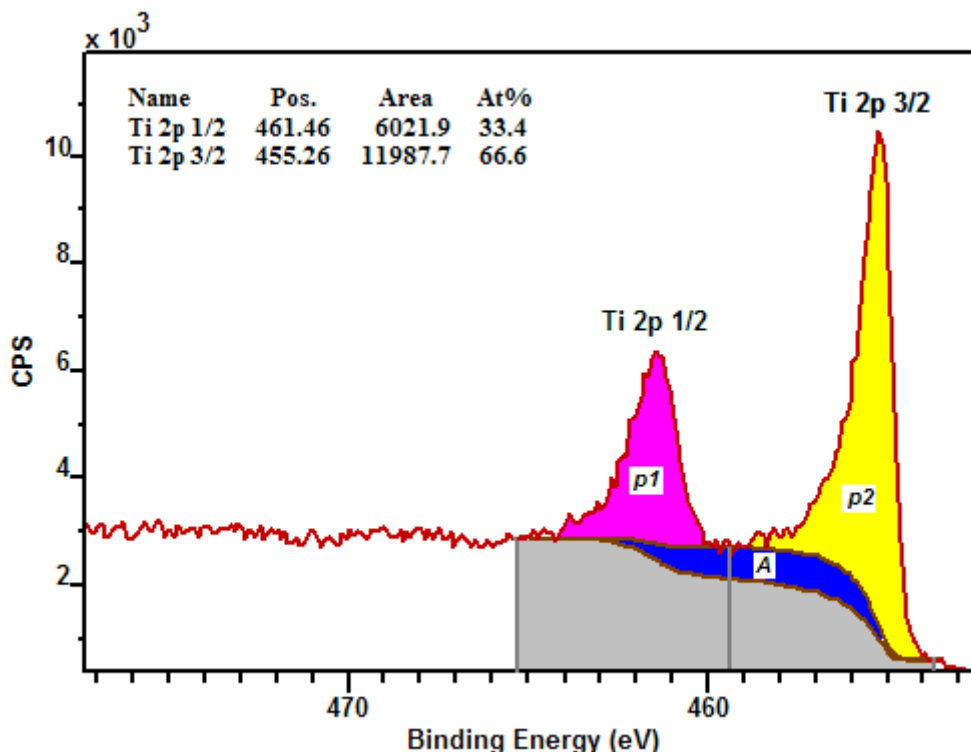

Figure 6: Metallic Ti modeled using two regions each defining a Shirley background (upper curve). The lower background curve is the Shirley background computed using the combined peaks. The At% column is computed using an RSF of unity for both peaks in the doublet pair, hence the 1:2 ratio in peak areas *p1* and *p2*.

Blending Shirley and Linear Backgrounds: "Offset Shirley" or "OS"

The background to the data in Figure 7 is calculated from a blend of a Shirley and linear backgrounds:

$$OS(E : \lambda, \delta) = S(E - \delta)(1 - \lambda) + L(E)\lambda$$

The parameters λ and δ represent a linear blend between the Shirley background corresponding to $\lambda = 0$ and a linear background $\lambda = 1$, where the Shirley curve is offset by an energy of δ eV. These parameters are entered using the last two fields in the **Cross section** entry on the Regions property page.

The only justification for using the Offset Shirley background with the data in Figure 7 is that the ratio of the doublet pairs satisfies the 1:2 ratio dictated by the degeneracy of the two states responsible for the doublet peaks. Use of a pure linear or a pure Shirley background over the same interval would require the intensity ratio to be violated and as such represents a less wrong solution used when peak fitting the data.
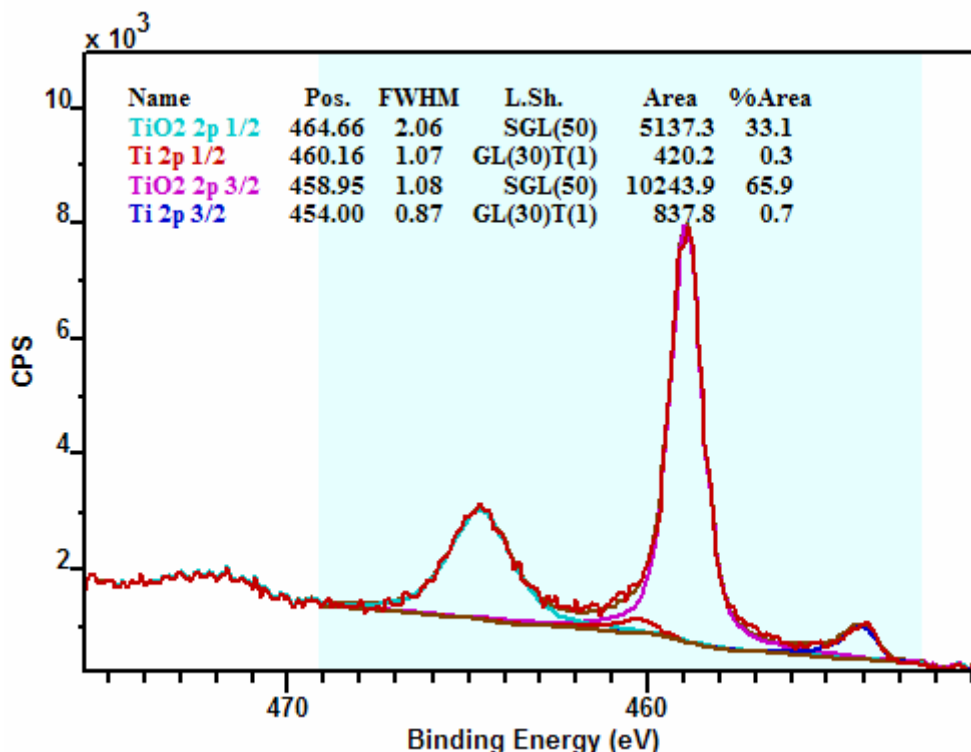

Figure 7: Chemically shifted Ti 2p doublet pairs. Data supplied by Elise Pegg, Bioengineering Group, University of Nottingham.

Tougaard Background: BG Type "Tougaard" or "t"

The general formulation of a Tougaard background is based on the existence of an energy loss cross section $F(x)$ representing the probability that an electron at energy offset x undergoes a loss event and therefore appears as a contribution to the background. The background is computed from the measured spectrum $S(E)$ using the integral:

$$T(E) = \int_E^\infty F(E' - E)S(E')dE' .$$

A number of such cross-sections are offered in CasaXPS, where the functional form is determined for various material properties; however the form typically associated with the name *Tougaard background* is the so called universal cross-section:

$$F(x) = \frac{Bx}{(C + x^2)^2}$$

The *B* parameter is typically adjusted in CasaXPS to ensure the computed background meets the data at the region limits. A second variation of the universal cross-section is the "**W Tougaard**" or "**w**" background, where the *B* parameter in influenced by the intensities I_1 and I_2 described above.

While conceptually appealing, the universal cross-section was designed as a practical background for general use and as such is almost certainly as incorrect as the Shirley and linear background types (Figure 8). Nevertheless, there are situations were the results from peak-fitting using a Tougaard background are more plausible than for other background types.
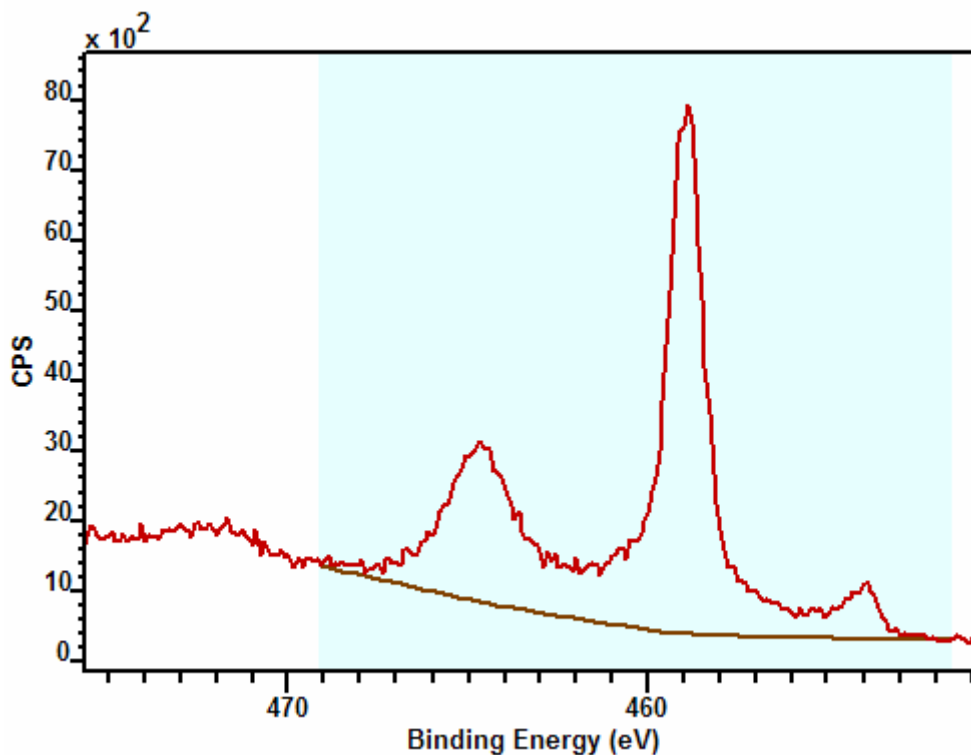

Figure 8: Universal Cross Section Tougaard background.

Three Parameter Tougaard Backgrounds

The universal cross-section represents, in some sense, an average shape for the loss function for a range of material. Tougaard also offers more specific cross-sections determined for individual material and these are also offered in the list of background

types available in CasaXPS. The functional from for the cross section involves three parameters B, C and D, plus an offset T_0:

$$F(x) = \begin{cases} \dfrac{Bx}{(C - x^2)^2 + Dx^2} & x > T_0 \\ 0 & x \leq T_0 \end{cases}$$

The parameters for Si, SiO$_2$, Al, Ge and polymers published by Tougaard [3] are implemented in CasaXPS. In addition to these published values, the four parameters may be customised using the Cross Section field on the Regions property page, which can be adjusted in conjunction with the "**U 4 Tougaard**" or "**U 4**" background type. The parameters are entered into the Cross section field in the order B, C, D, T_0 as a sequence of comma separated numbers.

A problem with the Tougaard backgrounds is that the background calculation assumes the spectrum for which the background is determined is without instrumental artefacts. Few XPS spectra are refined to the extent required to properly use the Tougaard approach, which is why the linear and Shirley algorithms persist in practical XPS. The acceptance of the lack of general applicability of the Tougaard approach but the fundamental correctness of determining the background from the data leads to the set of additional cross-section based background types in CasaXPS. The "U 4 Tougaard" background is offered with the awareness that its use is limited to cases where a custom background can be crafted, within the Tougaard framework, by simply matching the background to the data over a range of energies by adjustments to the four parameters.

Given an acceptance of this *ad hoc* approach to creating backgrounds, a natural step is to simplify the description of the loss function. The "E Tougaard" or "E" background type is such a simplification, where the loss functional form is a Gaussian. The parameters in the cross-section list are the Gaussian height, full-width-half-maximum, exponential tail parameter and an offset to the Gaussian. The exponential tail is introduced in an identical way to the exponential tail to the GL line-shape. The background in Figure 9 is created using a broad Gaussian with an exponential tail designed to raise the background to high binding energy values. The similarity with a Shirley background should not be surprising given the *A2(E)* in the equation for the Shirley background is a special case of the Tougaard integral where the loss function is a constant.

Flexible Background Shapes based on Cubic Spline Polynomials

CasaXPS offers backgrounds based upon the existing CasaXPS background types Shirley, Linear, Tougaard and (so called) None, but where the curve is defined by a set of cubic spline polynomials rather than the underlying functional forms. These background types "**Spline Linear**" (abbreviation "**sl**"), **Spline Shirley** ("**ss**"), **Spline Tougaard** ("**st**") and **Spline None** ("**sn**") are flexible backgrounds and can be adjusted under mouse control or optimised as part of a synthetic peak model.

A *spline* is a mechanical device used by draftsman to draw aesthetically pleasing curves using pen and paper: the draftsman fixes a set of points (knots) on the drawing, then bends a flexible strip of plastic or wood (the spline) around the knots and traces the shape onto the paper. Spline interpolation is the mathematical equivalent of this process and is achieved using a piece-wise cubic polynomial approximation to replace the mechanical spline. The smooth curve is achieved by requiring that adjacent cubic polynomials take on the value of the knots and that the first derivatives of these polynomials at the knots are equal. These conditions together with two end point conditions, such as assuming the second derivative at the two extreme knots is zero, allow a unique solution for the set of cubic polynomials for a given set of knots.

CasaXPS uses six knots to define five cubic polynomials cross a region which determine the background shape for that region. The knots are evenly spaced between the start and end points of the region, while the intensity at a knot may be adjusted under mouse control. This is achieved on the Components Property Page, where selecting a point on the background and dragging the mouse to a new position repositions the nearest knot and therefore adjusts the shape of the spline background. If the knot is an internal knot then the spline is bent to a new shape, while a knot corresponding to the end-points, the start and end offset region-parameters, is adjusted resulting in the background gradient altering.

A new option on the Components Property Page allows the internal knots within a spline background to be optimised. The chi-square goodness-of-fit knows nothing about the physical world and therefore optimising a background, while improving the goodness-of-fit will probably have nothing to do with the true mechanisms unless the model optimised is perfectly defined and there is no noise in the data. Well-defined rigid models are the only models worth optimising. Backgrounds are the least-well understood shapes in XPS and as such optimisation is not recommended. The inclusion of the spline background type in CasaXPS is intended for use where all else fails: for example a peak positioned on a strong plasmon loss feature is not open to a Shirley background approximation, however a Spline Linear with a few judicious modifications may allow an analysis to proceed.

While optimisation of peaks and background is not recommended, optimising a background for a region without peaks does have a useful benefit. The spline background, for a well-chosen interval, will create a smooth approximation to the signal and so provides a means of estimating the noise in the data. The estimate for the noise will appear along with the residual trace in the form of the standard deviation in the normalised residual. If Poisson statistics are at work, then the value for the noise should be close to unity. With knowledge of this statistic for a given piece of data the target value for a peak-fit can be established.
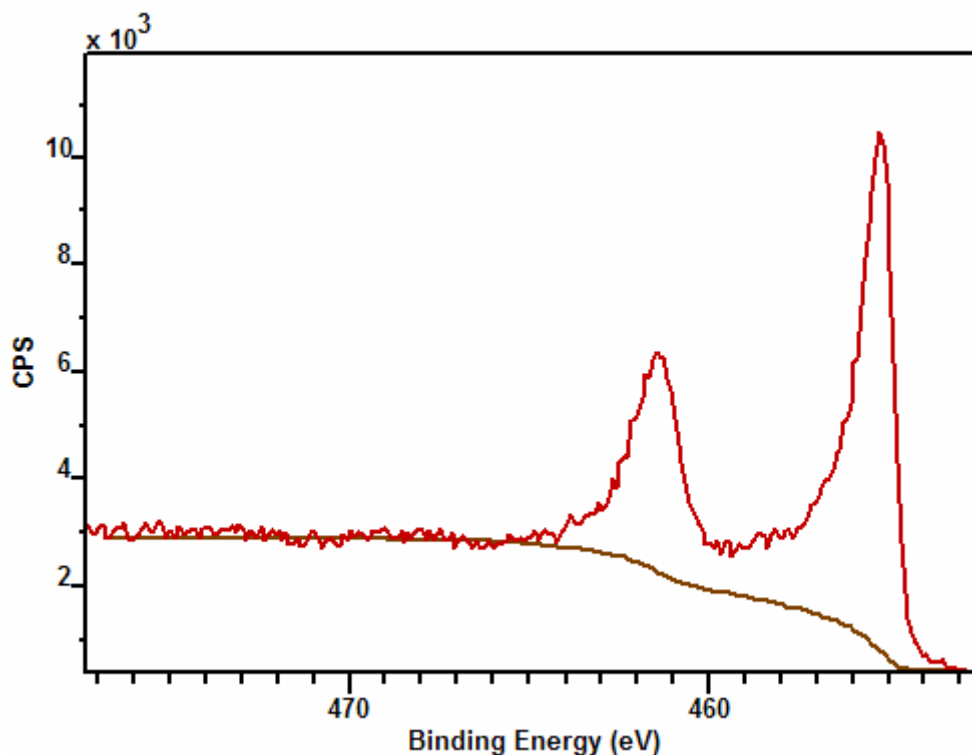
Figure 9: An E Tougaard background showing remarkable similarity to a Shirley background shape. The background shape is achieved by adopting a very broad Gaussian loss structure.

Line Shapes

Core level electrons ejected by x-rays appear in spectra with a variety of shapes. These shapes arise due to a combination of the physics involved in the ionisation process and also distortions due to the measurement mechanism. An idealisation of these influences is the specification of an observed peak as a convolution of a Gaussian and Lorentzian function; where in principle the Gaussian described the measurement process (instrumental response, x-ray line-shape, Doppler and thermal broadening) while the Lorentzian models the lifetime broadening (Natural broadening) due to the uncertainty principle relating lifetime and energy of the ejected electrons. The realities of observed spectral features are complex shapes which, in some cases, are shown theoretically to include asymmetry in the nature of the photoelectric peaks. Further, the line-shape chosen when fitting a given data envelope are as much dictated by the uncertainty in the background shape as any theoretical considerations. For these reasons, when fitting XPS spectra with synthetic line-shapes a variety of functional forms (Figure 10) are required.
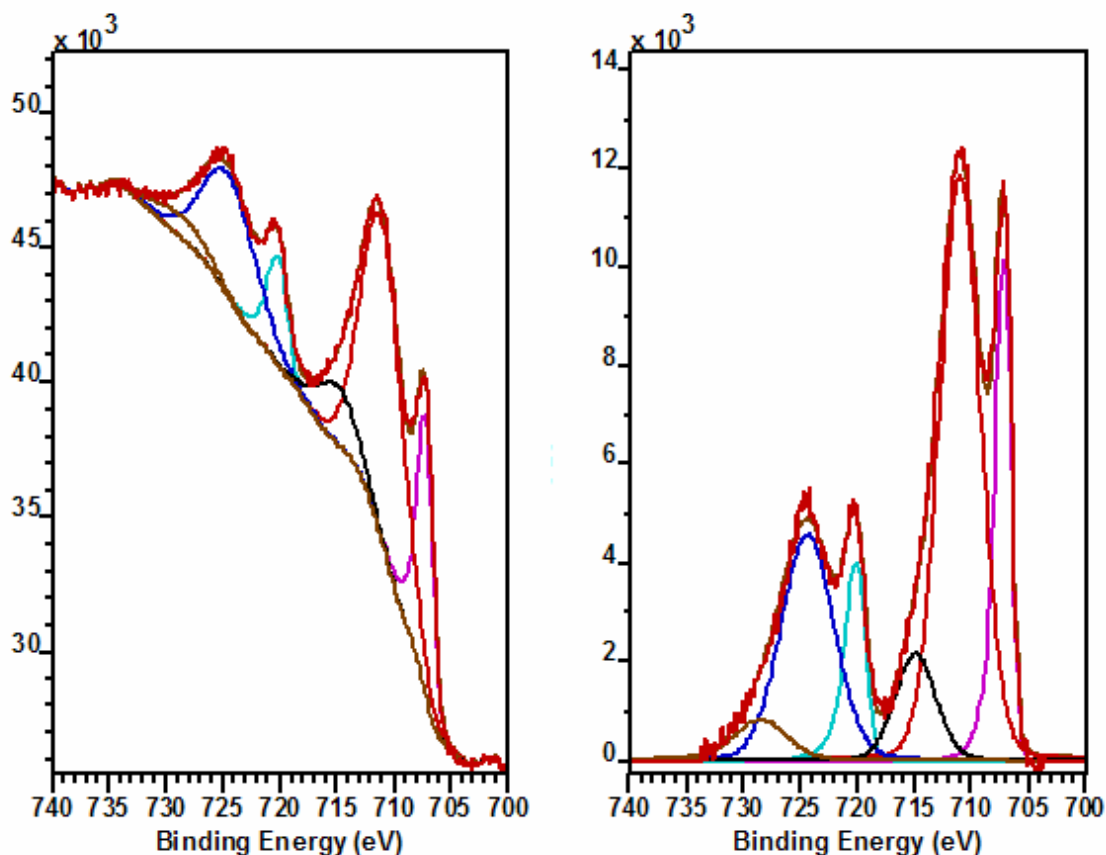
Figure 10: Fe 2p peak model based on an Offset Shirley background type and using a combination of SGL and asymmetric GL tail function line-shapes.

Gaussian Lorentzian Line Shapes

While CasaXPS includes a wide range of synthetic line shapes including Doniach-Sunjic and several variations of the Doniach-Sunjic asymmetric behaviour, for routine analysis of XPS spectra the most popular and least prone to inappropriate use are a class of functional forms based on the Gaussian and Lorentzian functions. The reason these line shapes are appropriate for measuring peak intensities is that both functions, when integrated from minus infinity to plus infinity are finite. Indeed, for all *normal* uses of the Gaussian-Lorentzian line shapes, the intensity as measured by fitting these functions to the data are very localised to the position of the peak. This is in contrast to almost all asymmetric line shapes, where introduction of asymmetry into the line shape extends the energy range of the line shape and alters the ability of the resulting shape to measure intensity without some form of calibration procedure.
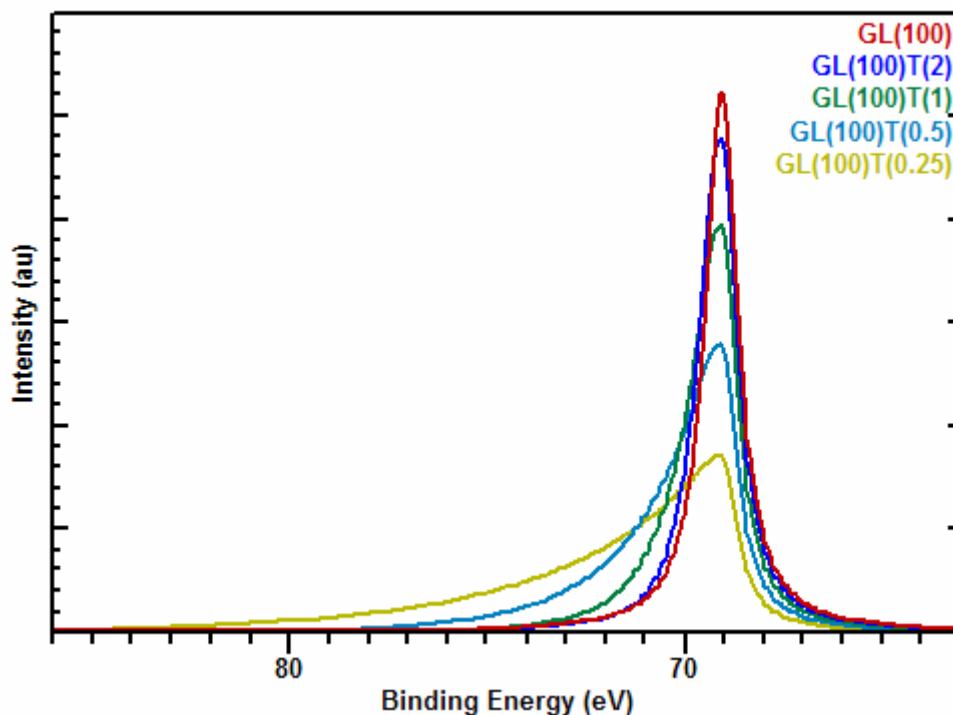
Figure 11:Lorentzian line shape modified using an exponential tail function. All line shapes have identical area beneath the curves.

Gaussian:

$$G(x : E, F, m) = \exp\left[-4\ln 2(1 - \frac{m}{100})\left(\frac{x - E}{F}\right)^2\right]$$

Lorentzian:

$$L(x : E, F, m) = \frac{1}{1 + 4\dfrac{m}{100}\left(\dfrac{x - E}{F}\right)^2}$$

Tail Modifier:

$$T(x : s, k, E, F) = \begin{cases} s\exp\left[-k\left(\dfrac{x - E}{F}\right)\right] & x \le E \\ 1 & otherwise \end{cases}$$

Line Shape parameter on the Components property page:

Symmetric Line Shapes

Product of a Gaussian with a Lorentzian
GL(m): $\qquad\qquad GL(x : E, F, m) = G(x : E, F, m)L(x : E, F, m)$

14

Sum or linear combination of a Gaussian with a Lorentzian

SLG(m): $$SGL(x:E,F,m) = (1 - \frac{m}{100})G(x:E,F,0) + \frac{m}{100}L(x:E,F,100)$$

Gaussian-Lorentzian Asymmetric Line Shapes

GL(m)T(k): $$GL(x:E,F,m) + (1 - GL(x:E,F,m))T(x:1,k,E,F)$$

SGL(m)T(k): $$SGL(x:E,F,m) + (1 - SGL(x:E,F,m))T(x:1,k,E,F)$$

GL(m)SL(s,l): $$GL(x:E,F,m) + (1 - GL(x:E,F,m))T(x:s,13.8/l,E,F)$$

SGL(m)SL(s,l): $$SGL(x:E,F,m) + (1 - SGL(x:E,F,m))T(x:s,13.8/l,E,F)$$

For small values of the asymmetry parameter k, the asymmetric line-shapes spread the intensity to higher binding energies (Figure 11) and, when used to quantify chemical states, the need to calibrate the intensity of the chosen line shape becomes important.

Asymmetric Line Shapes

The following asymmetric line shapes are not suitable for quantification without intensity calibration. The problem with asymmetric line shapes is that the asymmetric tail moves intensity away from the peak maximum and depending on the extent of the tail, significant intensity can be moved outside the energy region over which the background to the spectrum is defined. Under these circumstances, simply integrating the line shape over the energy interval means that the measured intensity is sensitive to the decision of where to place the background limits.

The Doniach-Sunjic profile (Figure 12), while theoretically based, is probably the most difficult functional form to use. The factor $\Gamma(1-\alpha)$ included by Doniach-Sunjic [4] in their expression for the photoelectron yield is simply a constant for a given α and therefore not included in the profile below, nevertheless this factor becomes increasingly large as the degree of asymmetry increases and, is a good indicator that the area beneath the Doniach-Sunjic profile is ill-defined.

$$DS(x:\alpha,f,e) = \frac{\cos\left[\frac{\pi\alpha}{2} + (1-\alpha)\tan^{-1}\left(\frac{x-e}{f}\right)\right]}{\left(f^2 + (x-e)^2\right)^{(1-\alpha)/2}}$$
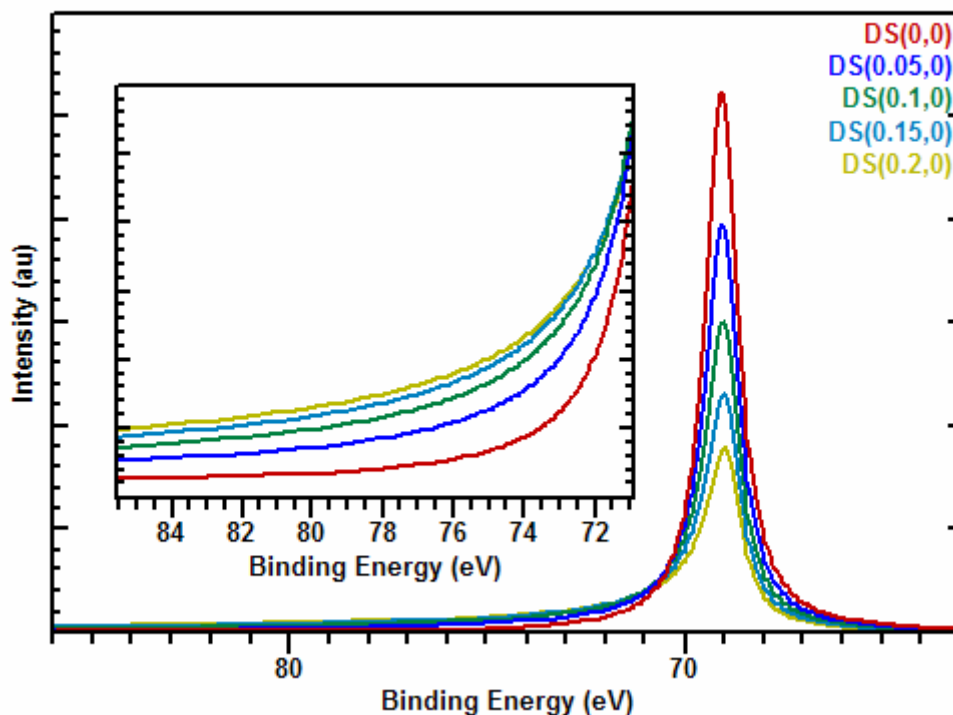
15

Figure 12: Doniach-Sunjic profiles for α between 0 and 0.2.

The Doniach-Sunjic profile is equivalent to the use of a Lorentzian to model the Natural line shape and similarly must be convoluted with a Gaussian (Figure 13) to create an approximation to the observed XPS peaks.

DS(α,n):   Doniach-Sunjic profile convoluted with a Gaussian with width characterized by an integer $0 \le n \le 499$.

Since, for whatever reason, not all asymmetric peaks are as Lorentzian in nature to lower binding energies, modified Doniach-Sunjic shapes are included as options. These line shapes are blends between a Doniach-Sunjic curve to higher binding energy and Gaussian-Lorentzian curves to lower binding energy.

DS(α,n)GL(m)

DS(α,n)SGL(m)

Further flexibility in the asymmetric line shape can be introduced by adding parameters to blend the asymptotic behaviour of the Doniach-Sunjic profile with a Gaussian-Lorentzian line shape as follows:

$$Voigt(x:F,E,m) = \begin{cases} SGL(x;F,E,m) & or \\ GL(x;F,E,m) \end{cases}$$

$$F(x:\alpha,F,E,m) = \begin{cases} \dfrac{m/100}{\left(F^2 + (x-E)^2\right)^{(1-\alpha)/2}} + (1 - m/100)Voigt(x;F,E,m) & x \le E \\ Voigt(x;F,E,m) & otherwise \end{cases}$$

The implementation of the blended line-shape is:

F(α,m$_1$,n)GL(m$_2$)
F(α,m$_1$,n)SGL(m$_2$)

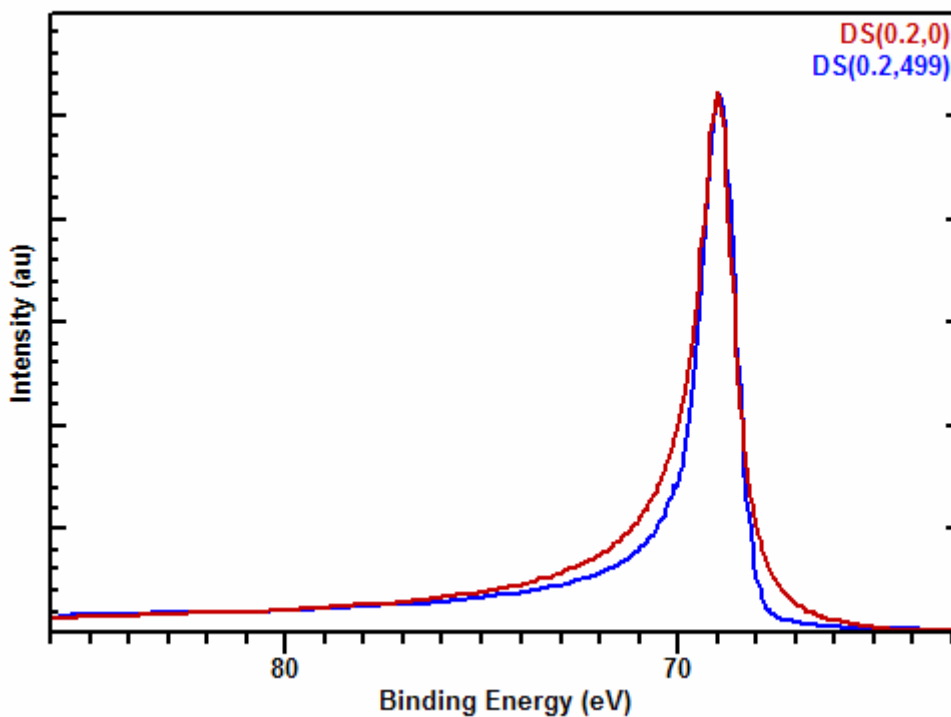The two functional forms in the blended result are convoluted with a Gaussian characterized by the width parameter n.

Figure 13: Comparison between a pure Doniach-Sunjic profile and the same profile convoluted with a Gaussian using the maximum width for the Gaussian function.

Remarkably, even for spectra as simple as the poly(ethylene) spectrum in Figure 14, where the only XPS peaks are the C 1s, C 2s and C 2p (XPS can not detect hydrogen), to understand the structure of apparently a single C 1s peak in Figure 15, a peak model is required. The reason for the need of a peak model is the C 1s spectrum exhibits asymmetry due to quantised molecular vibrational states of the poly(ethylene) which may be excited at the instants a core-level electron is ejected by a photon. Energy absorbed by these vibrational states is not available to the photoelectron in the form of kinetic energy and therefore the C 1s spectrum is actually a sequence of symmetric lines offset by the quantum of energy (0.39 eV [1]) characteristic of these molecular vibrations.
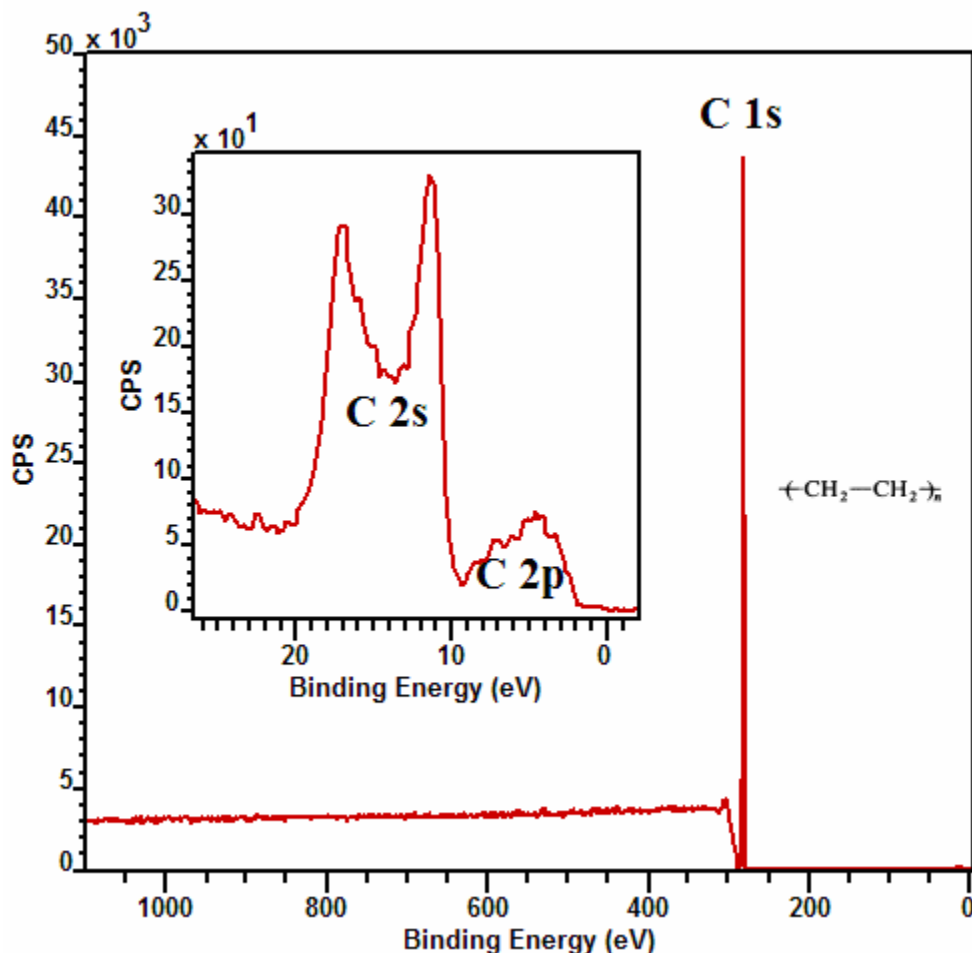

Figure 14: Poly(ethylene) Survey Spectrum.

A peak model such as the one used in Figure 15 for poly(ethylene) represents the individual peaks via an SGL(13) line-shape, however without imposing parameter constraints on the peak positions and *fwhm* , optimisation of the area, *fwhm* and position parameters is unstable with respect to noise and typically results in percentage areas for the component peaks different from those displayed in the quantification table in Figure 15. Imposing constraints for position relative to the component in column **A** of the

18

components table on the Components property page takes the form a string such as "A+0.39" entered into the position constraints field in column B (say). Similarly, the *fwhm* in constraint field in column B could be set to "A * 1". Provided all the peaks are appropriately offset and the *fwhm* are forced to have the same value, the model becomes less sensitive to noise in the data and returns a consistent answer for the relative intensities of the four component model.
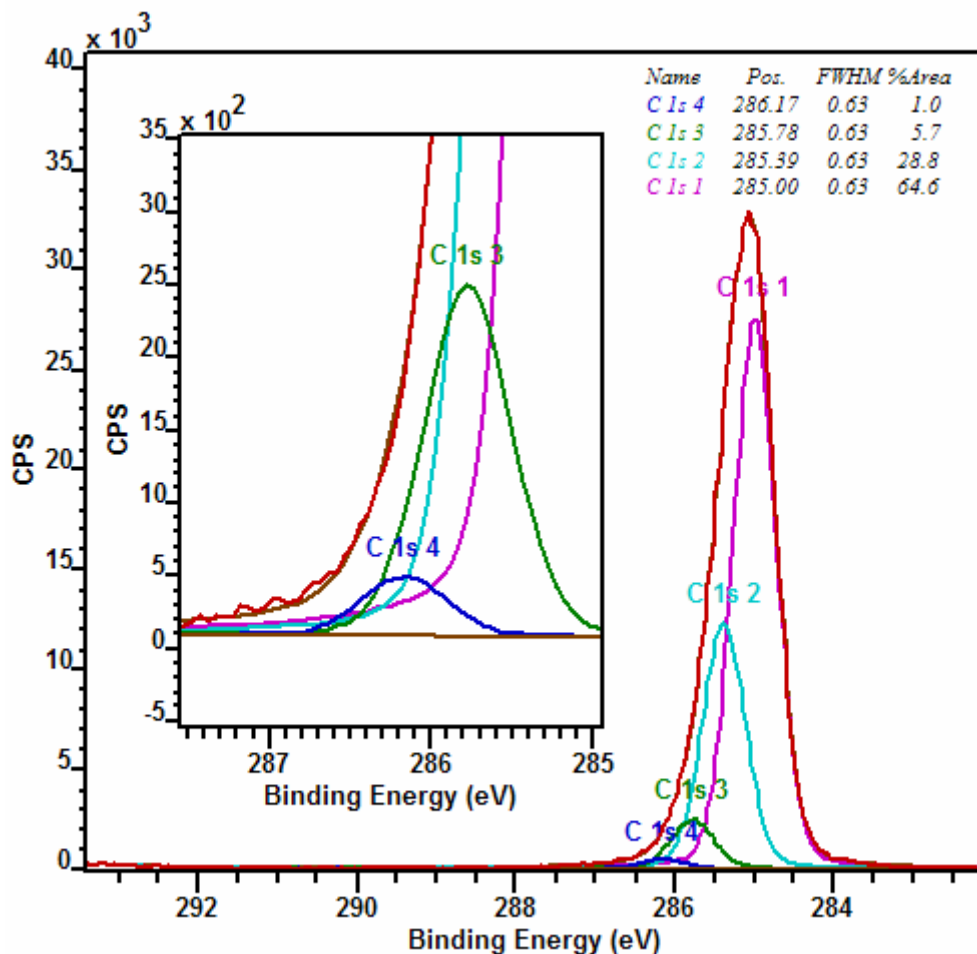

Figure 15: Poly(ethylene) C 1s spectrum fitted using four SGL synthetic peaks.

While the vibrational structure of poly(ethylene) C 1s spectra explains the asymmetry observed in the data envelope, the need to measure the intensity from the vibrational states is often not required and a single asymmetric line-shape is all that is required when measuring the intensity of such a data envelope. For example the C 1s spectrum for poly(propylene), like poly(ethylene), is characterized by a single peak; however the chemistry of the polymer [1] suggests vibrational states for each of the two primary lines are responsible for the observed spectrum (Figure 16), therefore the peak model requires suitable asymmetric line-shapes to describe the vibrational asymmetry.
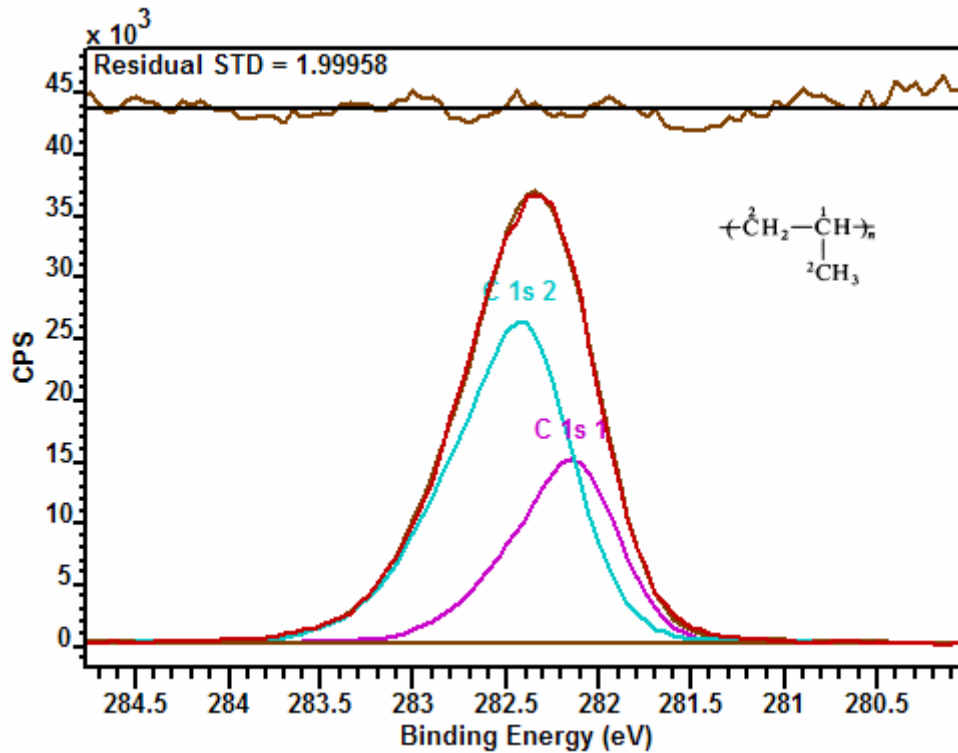
Figure 16: Poly(propylene) C 1s spectrum fitted using Gelius line-shapes to approximate the asymmetry in a two-component model.

The asymmetric line-shape used to model the two peaks in Figure 16 is a Gelius functions defined as follows:

$$A(x;a,b,F,E,m) = \begin{cases} Voigt(x;F,E,m) + w(a,b)\left[AW(x;a,F,E) - GL(x;F,E,0)\right] & x \le E \\ Voigt(x;F,E,m) & x > E \end{cases}$$

where,

$$w(a,b) = b\left[0.7 + \frac{0.3}{(a+0.01)}\right]$$

$$AW(x;a,F,E) = \exp\left[-\left(\frac{2\sqrt{\ln 2}(x-E)}{F - a2\sqrt{\ln 2}(x-E)}\right)^2\right]$$

The characteristics of the Gelius line-shape are different from the Doniach-Sunjic asymmetry and were designed for use with polymers. The Doniach-Sunjic profile, in contrast, is the consequence of a mathematical treatment of the intrinsic losses associated

20

with the ejection of an electron from a metallic surface; an essentially different process from the vibrational asymmetry observed in poly(ethylene) and poly(propylene).

The line-shape parameter in CasaXPS is of the form:

A(a,b,n)GL(m): Gelius profile convoluted with a Gaussian with width characterized by an integer $0 \leq n \leq 499$.

A(a,b,n)SGL(m): Gelius profile convoluted with a Gaussian with width characterized by an integer $0 \leq n \leq 499$.

The line-shape used in fitting the data displayed in Figure 16 is A(0.15,0.7,20)SGL(12) and was determined from the asymmetry found in the poly(ethylene) data. In fact, determination of the asymmetric line-shape is one of the arts of modeling data. The use of the poly(ethylene) shape in modeling the poly(propylene) is actually a measure of how difficult it is to identify the correct line-shape from data such as the poly(propylene) C 1s spectrum alone. There is certainly insufficient information in the data for poly(propylene) to uniquely determine the underlying line-shapes from these highly correlated peaks.

A VERY IMPORTANT NOTE: It is critical to appreciate the issue of *intensity calibration* associated with asymmetric line-shapes. This is particularly true for the Gelius function since the choice of the parameters used to approximate the vibrational state of a transition can cause intensity from the nominal Gaussian-Lorentzian region near the peak maximum to move beyond the region defining the background. This subject is dealt with in detail elsewhere.

The vibrational structure is an example of where part of the photon energy is absorbed in ways other than purely ejecting an electron from the surface. Other absorption modes, which include so called shake-up peaks, are possible where the loss structures appear at lower kinetic energy to the primary peaks with magnitude larger than the vibrational energy losses observed in poly(ethylene) and poly(propylene). Poly(ether ether ketone) (PEEK) is an example of a spectrum in which three primary C 1s peaks are assigned to chemical states of the carbon atoms, while secondary shake-up loss peaks appear to higher binding energy values (Figure 17). Since these shake-up peaks are often significant (5%-12% of the primary intensity [5]) and also overlap with the primary peaks, all must be included in the peak model. The background type is an important part of such a peak model and in the case of the PEEK data in Figure 17, a Tougaard background was used. The advantage of a Tougaard background over a simple linear form is that the shake-up peaks appear at a point in the spectrum where the extrinsic loss mechanism starts to cause the background to rise more steeply compared to the region beneath the primary peaks. The backgrounds in Figure 17 and Figure 18 are examples of a "U Poly Tougaard".
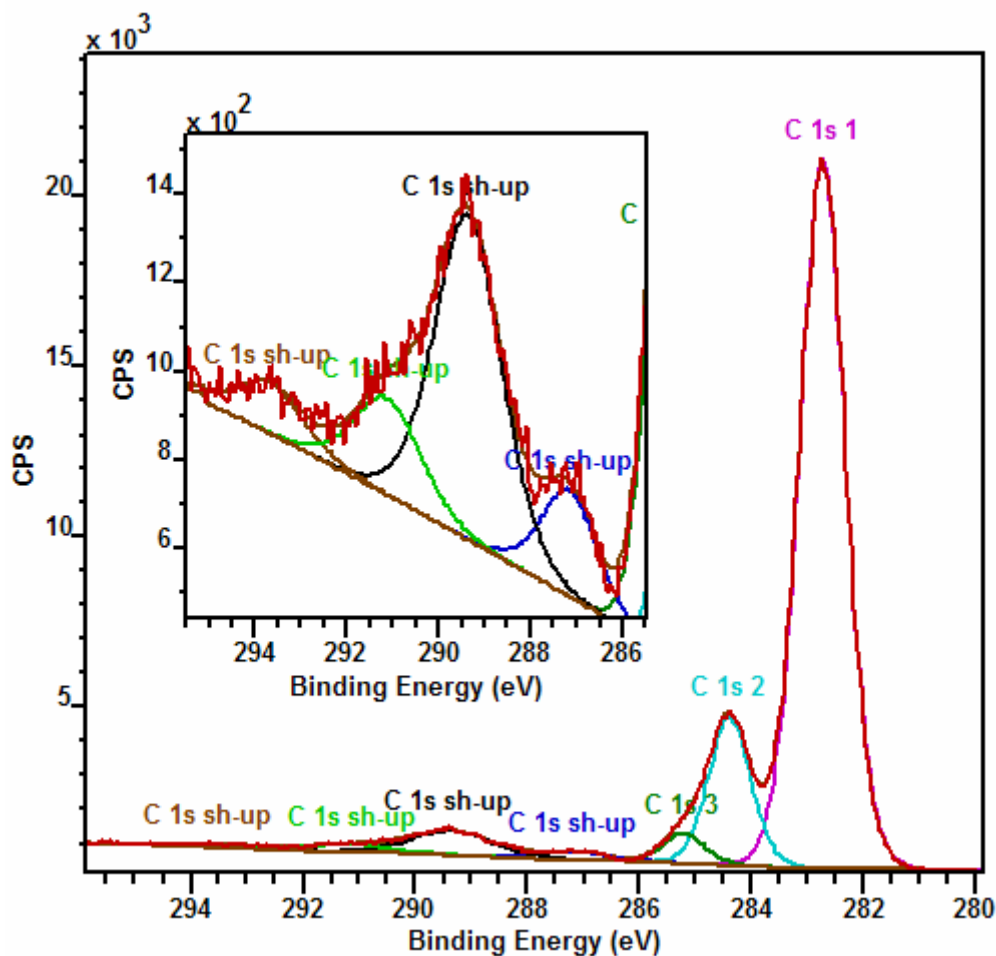
Figure 17: Poly(ether ether ketone) or PEEK showing shake-up peak structures.

The shake-up peaks are part of a class of intrinsic loss events, where after absorption of a photon the final state of the electronic system is left in an excited state. These two electron processes [5] add to the complexity of XPS spectra and their interpretation with respect to the chemical formulae. The relationship between the primary lines and the shake-up lines in PEEK is clearly not one-to-one.

An *intrinsic* energy loss takes place at the same moment the electron is ejected from the surface. An *extrinsic* energy loss occurs following the ejections of the electron from a bound state in an atom and is independent of the photon induced event. X-rays penetrate far into the material compared to the depth from which electrons of a given kinetic energy can escape from the surface without loss. Inelastic scattering of electrons emerging from beneath the surface account for the extrinsic background observed in XPS. Figure 18 illustrates the context in which the C 1s spectrum in Figure 17 is set. The extrinsic background rises rapidly to higher binding energies from the position of the primary C 1s spectrum, on top of which the intrinsic shake-up peaks are located. The

shape of the extrinsic loss structure is dependent on the material under analysis and any overlaying contamination. Extrinsic loss structures can be significant. Figure 19 is an example of a spectrum taken from elemental silicon. The spectrum is an example of a material where the loss structures are almost as pronounced as the primary peaks.
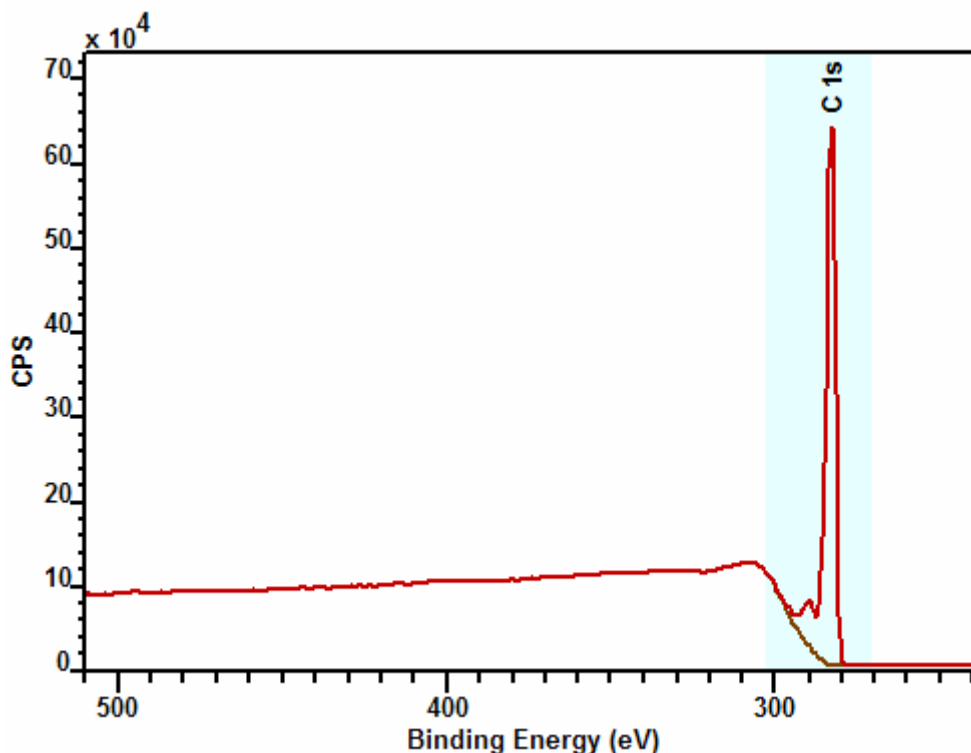

Figure 18: A low resolution spectrum over a wide region about the C 1s transition in poly(ether ether ketone). The extrinsic loss structure due to inelastic scattering of C 1s electrons represents a significant feature of the data and should be compared to the high resolution data in Figure 17.

The nature of the loss structure for silicon is related to the existence of free electrons in the silicon, which may be excited with characteristic frequencies resulting in the sharp plasmon peaks to higher binding energy of the primary photoelectric peaks. Plasmon excitation is possible by both intrinsic and extrinsic mechanisms. Partitioning the intensity between intrinsic and extrinsic counts is difficult due to the fact that counts from both mechanisms contribute to the same plasmon peak in the energy spectrum [5]. A further observation about the data in Figure 18 is that the silicon plasmon losses are not visited on the O 1s electrons (inset plot in Figure 18), where an entirely different loss structure results from the O 1s excitation.
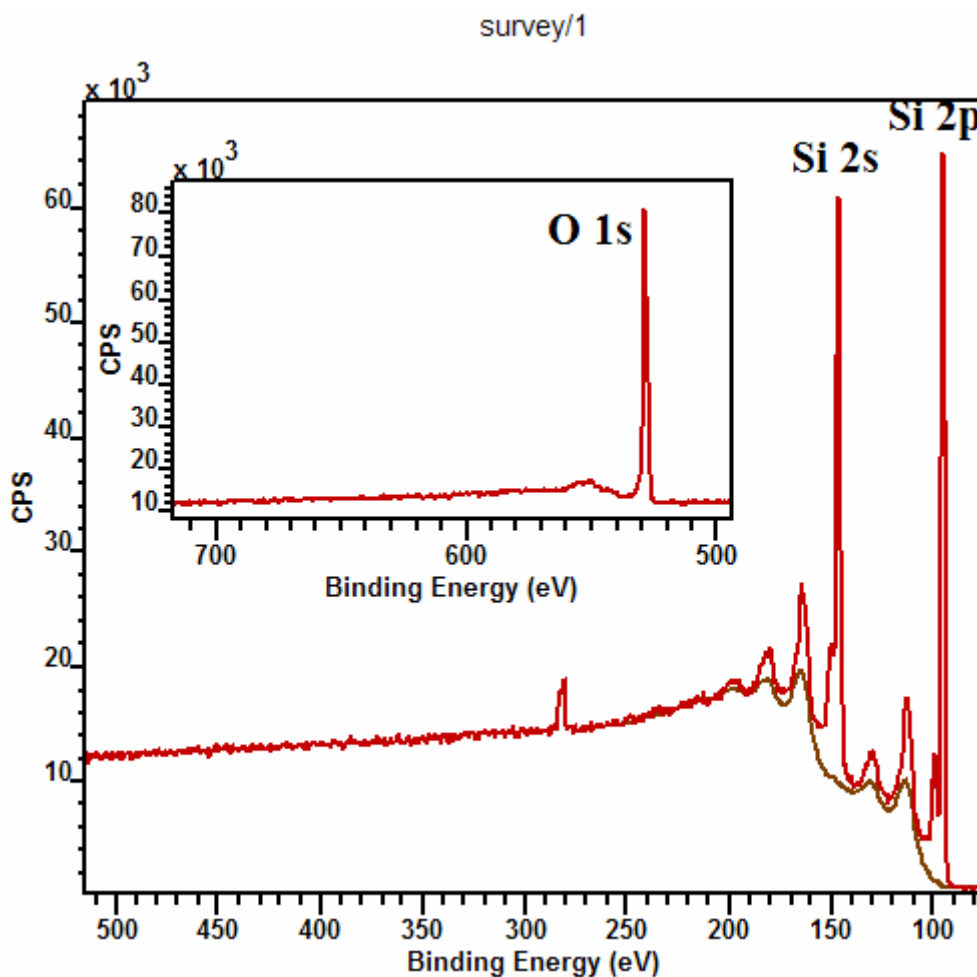
Figure 19: Silicon spectrum showing plasmon loss features.

As a further example of loss behaviour, the metallic aluminium spectrum in Figure 20 was acquired after clearing of the surface metal by argon etching. The Al 2s and Al 2p peaks are accompanied by the sharp plasmon loss peaks associated with metallic aluminium spectra; however similar loss peaks can be identified for the argon implanted as part of the bombardment of the surface by argon ions during the cleaning of the surface. The background type "**U Al Tougaard**" or "**U Al**" is used to illustrate the shape of the extrinsic background expected for metallic aluminium spectra. The shape of the background supports the common practice of using the Al 2p peak to quantify metallic aluminium, since the Tougaard background in Figure 20 predicts a loss structure from the Al 2p peak beneath the Al 2s peak. All things being equal, the peak with the largest photo-ionization cross-section would normally be used to quantify samples, so simply on the basis of optimal signal-to-noise the Al 2s would be the preferred peak of the two.
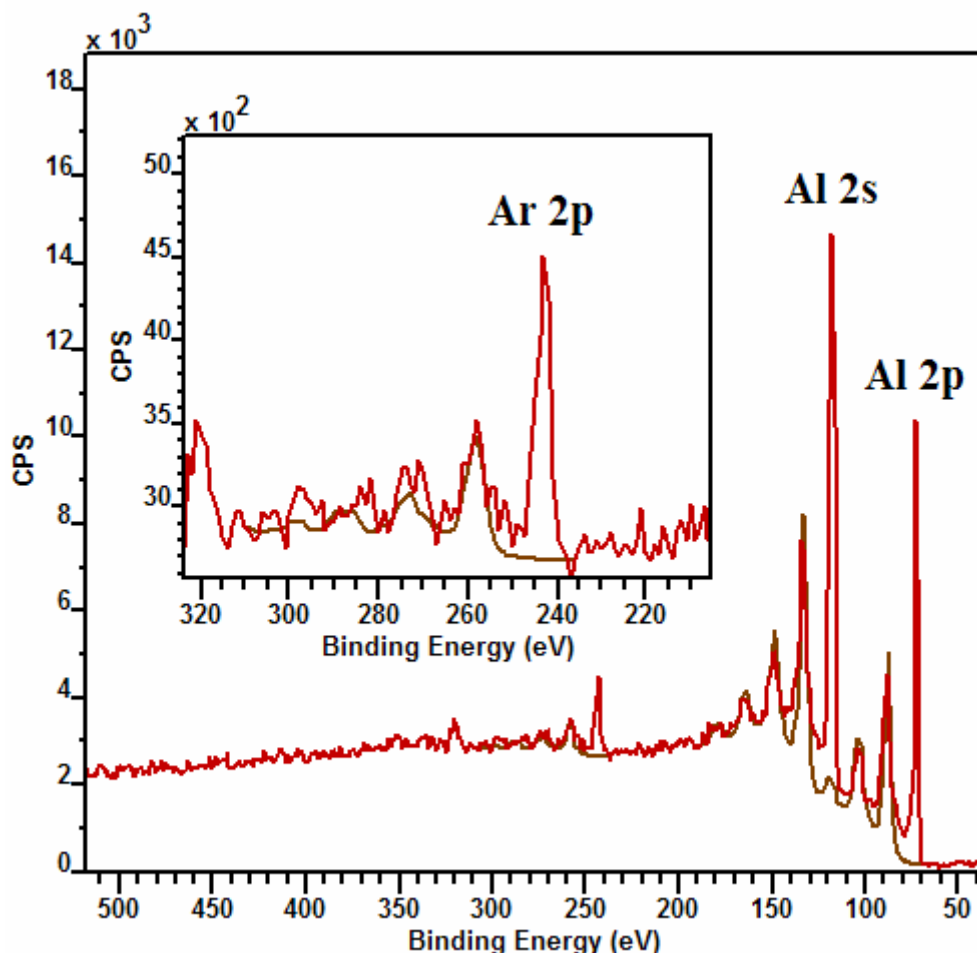
Figure 20: Clean aluminium metal survey spectrum.

The nature of the background and choice of background algorithm used to model these contributions to the spectra have a bearing on the line-shape chosen to measure the photoelectron lines. The data in Figure 21 illustrates the variation possible for the same spectrum after background subtraction where different background strategies have been employed. The Ti 2p doublet in Figure 6 is used in Figure 21. Not only does the asymmetry of the peaks alter depending on the background but clearly the area beneath the background subtracted curves changes too.

While the various asymmetric functional forms include parameters used to define the degree of asymmetry, the meaning of these parameters is also dependent on the choice of background type and to some extent should be viewed as simply a means of recreating the appropriate shape for the peak rather than a metric for characterizing the peak. For example, the Gelius line-shape is a good example of a functional form for which the parameters are solely for the purpose of modeling the shape of an ensemble of peaks due to vibrational-states. The Gelius line-shape therefore replaces many peaks by a single

line-shape, a strategy that can be extended by allowing any shape defined in terms of data points to act as a line-shape when modeling complex spectral data.
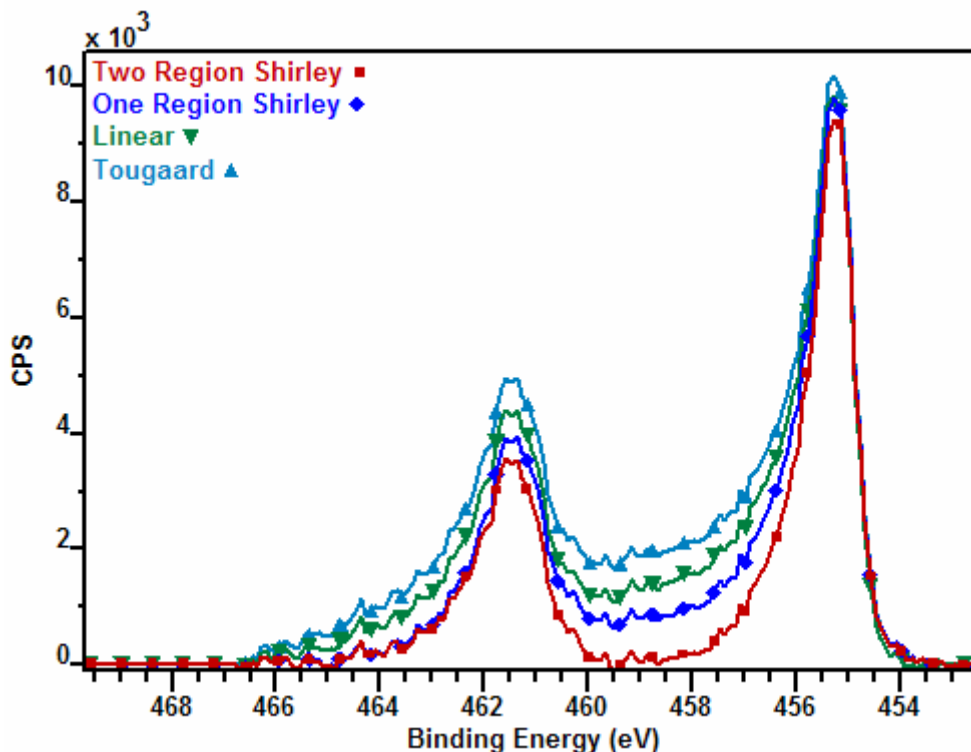

Figure 21: The same Ti 2p Doublet displayed using background subtracted data, where four different background types are used to model the background.

Line Shapes Derived from Data

Rather than using a synthetic line-shape defined by a mathematical formula, it is sometimes easier to define a line-shape using data from known materials. For example, the relative intensities of metal and oxide peaks are often used to measure the oxide thickness over metallic aluminium. Such measurements are typically made using the Al 2p transition for the reason given above and also the superior line resolution of the Al 2p over the Al 2s transition, but nevertheless the chemical shift between the metal and oxide peaks is such that a peak model must be developed in order to separate the contributions from the oxide and the metal. Further, metallic Al 2p peaks exhibit asymmetry which extends into the region beneath the oxide peak as can be seen in Figure 22. Even if the synthetic line-shapes are used, assessing the contribution of the metallic Al 2p peak to the intensity beneath the oxide peak would be difficult without a spectrum from clean aluminium (Figure 23). CasaXPS offers the option of using the data in Figure 23 to directly measure the intensity of the metallic peaks in Figure 22.
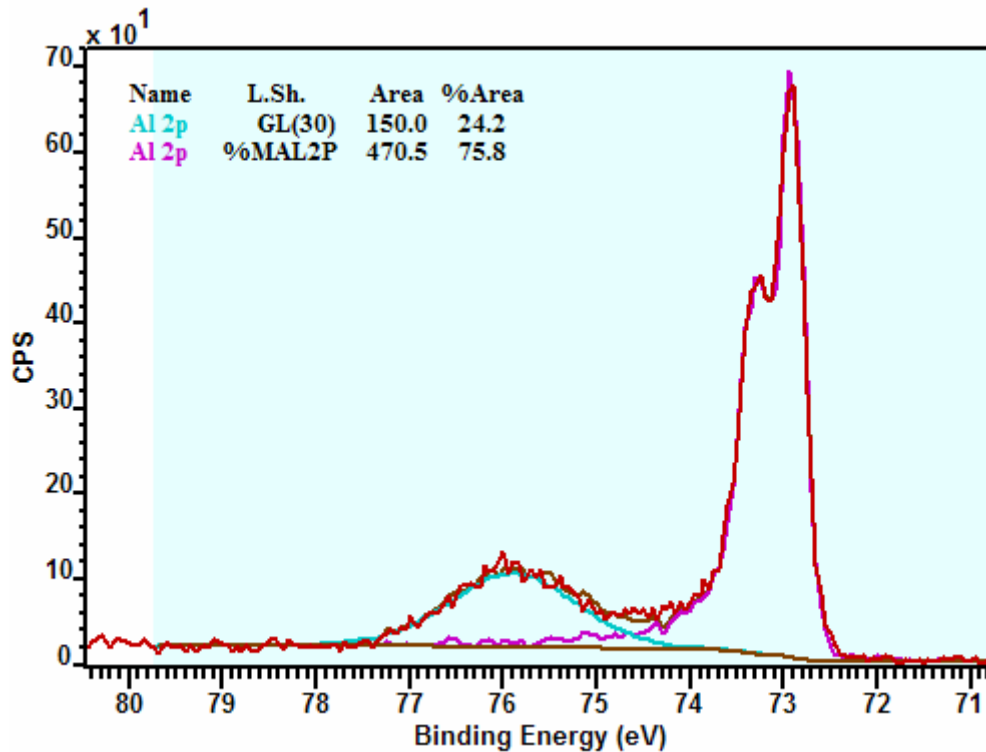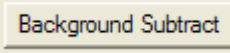
Figure 22: Early stages of oxide growth on a metallic aluminium surface.

The line-shape labeled %MAL2P in Figure 22 is defined from the data in Figure 23. The procedure involves 1) including the VAMAS block containing the data in Figure 23 in the same Experiment Frame as the data requiring the peak model, 2) deciding on how to remove the background from each spectrum in use and 3) changing the block id of the data acting as a line-shape to include a prefix of "%". The steps used to create the peak model in Figure 22 are as follows:

Manipulate the data to be used as the line-shape:
1. Select the VAMAS block containing the metallic spectrum.
2. Add a quantification region to define the background.
3. Using the Test Data property page on the Spectrum Processing dialog window, press the Background Subtract button [Background Subtract]. The processed data is therefore background subtracted data and in a state ready to be transferred to the Experiment Frame containing the target data.

Copy the processed data to a new VAMAS block in the Experiment Frame containing the target data.
4. Select the VAMAS block containing the background subtracted line-shape data, change the focus, if required, to the target Experiment Frame and press the Copy and Paste toolbar button.

5. Tick the check-box labeled "Processed Data Only" ☐ Processed Data Only before pressing the OK button. A copy of the processed data will be added to the Experiment Frame with focus. The use of the Processed Data Only tick-box preserves the processing performed on the line-shape data and avoids future accidents causing the line-shape data to revert back to the raw data.
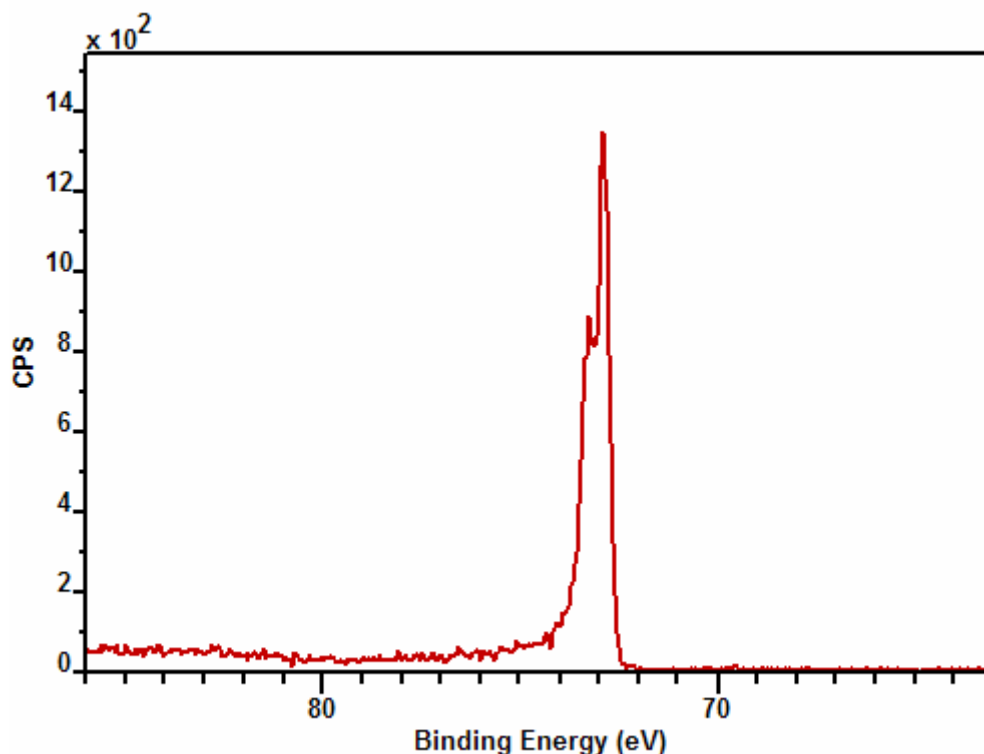

Figure 23: Metallic Al 2p spectrum with characteristic asymmetric tail.

6. Use the block info toolbar button [Block Info] to change the block id to a unique string beginning with the % character, e.g. %MAL2P.

Create the peak model on the target data in Figure 22

7. Display the target spectrum in the Active Tile.
8. Create a region with background consistent with the background type used to background subtract the line-shape data.
9. Add a component using the Components property page on the Quantification Parameters dialog window
10. Enter the unique string used in the block id field of the line-shape data into the Line Shape field of the newly created component.
11. Set the FWHM to unity and fix the parameter by entering a # character into the FHWM constraints field. The line-shape defined from data should not vary in FWHM since this will cause any features in the line-shape, such as doublet pairs, to alter in separation.

12. Add any other components to complete the peak model and press the Fit Components button.

References

[1] Beamson and Briggs, "The XPS of Polymers Database", ISBN 0-9537848-4-3.
[2] D.A. Shirley, *Phys. Rev. B* **5**, 4709-4714 (1972).
[3] S. Tougaard, SIA, vol. 25, 137-154 (1997).
[4] S. Doniach, M. Sunjic, *J. Phys. C* **3**, 285-291 (1970).
[5] Briggs and Grant, "Surface Analysis by Auger and X-ray Photoelectron Spectroscopy" ISBN 1 901019 04 7