CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL IPN

UNIDAD QUERÉTARO

Internal Report
(Created: 8/2011. Last Update: 2/2012)

The Peak-Shirley Background

Alberto Herrera-Gomez
CINVESTAV- Queretaro

Content

Abstract

In this Internal Report it is introduced the Peak-Shirley background. This is employed for fitting photoelectric spectra, such as in X-ray (XPS) or Ultraviolet (UPS) Photoelectron Spectroscopy. In contrast with the standard Shirley approach, the peak-Shirley background does not require iterations and, more importantly, neither choosing two ending points. It makes possible to assign different scattering strengths to different regions of the spectrum. It also provides better fits than the standard Shirley.

I. Introduction

I.1. Active versus static background

In another document (*AnActiveTreatmentOfTheBackground.pdf*) it is described the importance of employing algorithms that treat the background in an active way while fitting photoemission (XPS, UPS) spectra.[1] By "active way" it is meant that the intensity of the background is optimized during the fitting process, at the same time that the intensity of the peaks is assessed. This contrast with the standard or "static" method, in which the background is first subtracted from the original data and the modified, backgroundless spectrum, is fitted with peaks. The advantages of the optimized method are many:

- In the static method it is necessary to choose two points, one at side of the spectrum features or peaks. The operator has to guaranty or to assume that there is no appreciable contribution to the signal at those two points from any of the peaks. The method does not allow for considering peaks that extend beyond any of the chosen points. In contrast, the active

treatment does not restrain this option since the strength of the background can be varied to account for contributions from extended peaks. The optimized method should to be employed when the data was acquired with a too-short energy range, when one of the peaks extends beyond the spectrum limits. The extent of the errors incurred by using the static method in the quantification of the peak intensity is illustrated in Reference 1.

- The active treatment allows for a combination of backgrounds: the strength of a lineal background employed simultaneously with a Shirley background can be decided by the optimization process. In contrast, the static method requires that the operator chooses among the background types available (lineal, Shirley, Tougaard, etc.). The relative strength of combined backgrounds cannot be decided through optimization since any background type can be forced to exactly go through the two points.
- When the background is subtracted before peak-fitting, the Poisson-character of the data is not longer applicable, so the uncertainty on the peak intensities cannot be calculated from the covariant matrix.[2]

There are other issues, such that the operator tends to forget about the background once it has been subtracted. In the dynamic method, the changing role of the background is accounted as the operator considers different options during peak-fitting. The dynamic role of the background is illustrated in Figure 1.

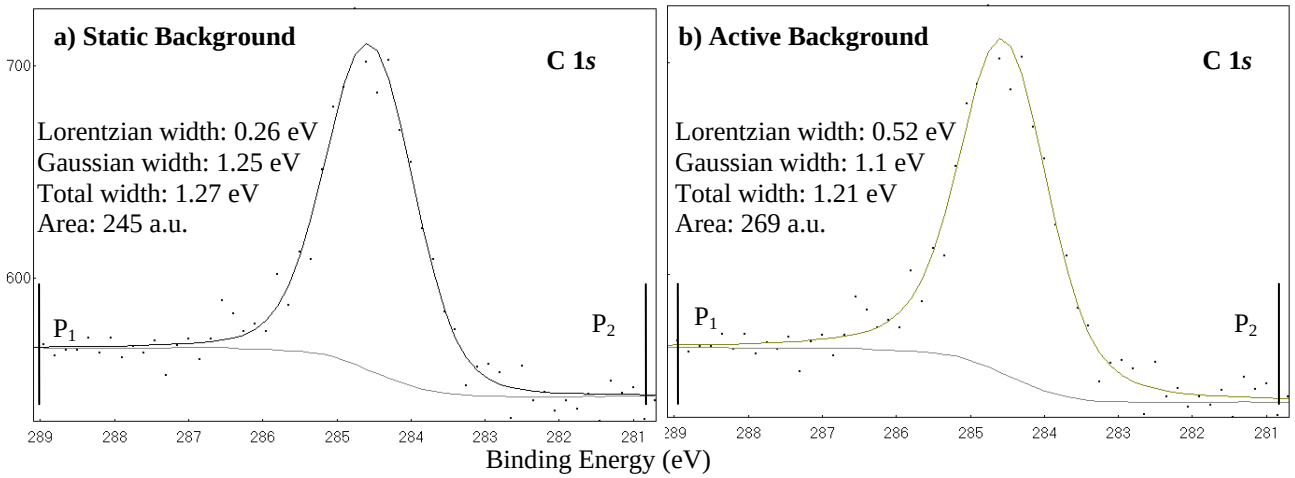

Figure 1. Fitting of a C 1s peak with the background treated a) in a static way (first subtracted and then fitted) and b) active (optimized during the fitting process). Although both fits looks equally good, the change in the area is 10%. The main difference comes from the background. In the second case, the background is allowed to be below the signal at the chosen points for the Shirley background (288.8 and 280.9 eV). This freedom allowed for larger values of the Lorentzian width during the optimization process.

All the calculations shown in this Internal Report employ the active method for the background. It should be mentioned that the importance of employing the active method has been mentioned in other reports.[3,4] However, the algorithms employed in those references for applying the active method required approximations that reduced its applicability[3] or just propose that an appropriate software should be developed.[4] In contrast, the algorithms encompassed in AAnalyzer®,[5] which is the software employed for the analysis, allow for a very simple application of the method without involving extra approximations.

I.2. **The Standard Shirley background**

As mentioned above, the standard Shirley background requires choosing two points, one at a kinetic energy above $\left(E_{\text{right}}, I_{\text{right}}\right)$ and another below $\left(E_{\text{left}}, I_{\text{left}}\right)$ the peak. These correspond to P_1 and P_2 in Figure 1. The construction of the background involves an iterative process:[6,7]

Equation 1

$$B_1(E) = k_1 \int_E^{E_{\text{right}}} dE' \left[I(E') - I_{\text{right}} \right],$$

$$B_2(E) = k_2 \int_E^{E_{\text{right}}} dE' \left[I(E') - I_{\text{right}} - B_1(E') \right], \qquad \text{for} \quad E_{\text{left}} \le E < E_{\text{right}}$$

$$\ldots$$

$$B_n(E) = k_n \int_E^{E_{\text{right}}} dE' \left[I(E') - I_{\text{right}} - B_{n-1}(E') \right],$$

where $B_n(E)$ is the n-th iteration of the Shirley background, E is the kinetic energy, k_n is the iterative value of the scattering factor,[3] and $I(E)$ is the photoelectric signal. At each iteration, the value of the scattering factor is chosen in such way that the total background goes through the elected point on the left $\left(E_{\text{left}}, I_{\text{left}}\right)$:

Equation 2

$$k_n = \frac{I_{\text{left}} - I_{\text{right}}}{\int_{E_{\text{left}}}^{E_{\text{right}}} dE' \left[I(E') - I_{\text{right}} - B_{n-1}(E') \right]}$$

The total background is the sum of a constant term, equal to I_{right}, and the Shirley background. Outside the range indicated in the equation above, the Shirley background is equal to 0 for $E_{\text{right}} \le E$, and to I_{left} - I_{right} for $E \le E_{\text{left}}$. This construction guaranties that the total background also passes through $\left(E_{\text{right}}, I_{\text{right}}\right)$ since $B_n\left(E_{\text{right}}\right) = 0$ for every n.

It should be pointed out that the scattering factor is defined in Reference 3 by normalizing with respect to the area of the spectrum after background subtraction:

$$B_n(E) = \frac{k_n' \int_E^{E_{\text{right}}} dE' \left[I(E') - I_{\text{right}} - B_{n-1}(E') \right]}{\int_{E_{\text{left}}}^{E_{\text{right}}} dE' \left[I(E') - I_{\text{right}} - B_{n-1}(E') \right]}$$

This definition is not useful because its value scales with the intensity of the peak. That is, a peak with the double of intensity would require a scattering factor of the double of its value to reproduce adequately the background intensity. In contrast, the definition of the scattering factor used in Equation 1 is independent of the vertical scale.

The iterative process ends when the difference between B_n and B_{n-1} is negligible, which in turn guaranties that the scattering factor k has also converged. Figure 4 shows the background of

Figure 1b but for different number of iterations. The background is very similar to each other from the first iteration, and the difference between B_n and B_{n-1} is negligible for $n \geq 2$.

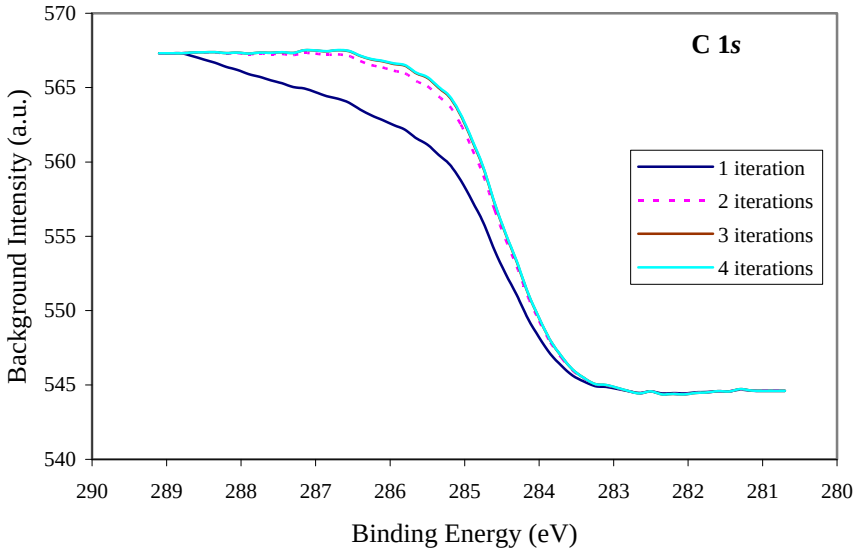

Figure 2. Shirley background for different iterations for the fit shown in Figure 1a. For this simple spectrum, the background does not change significantly after n = 2. The fit looks almost equally good from n = 1 since the change on the background is compensated by the peak intensity.

As it is illustrated in Figure 3, the convergence is slower when the spectrum contains many peaks.

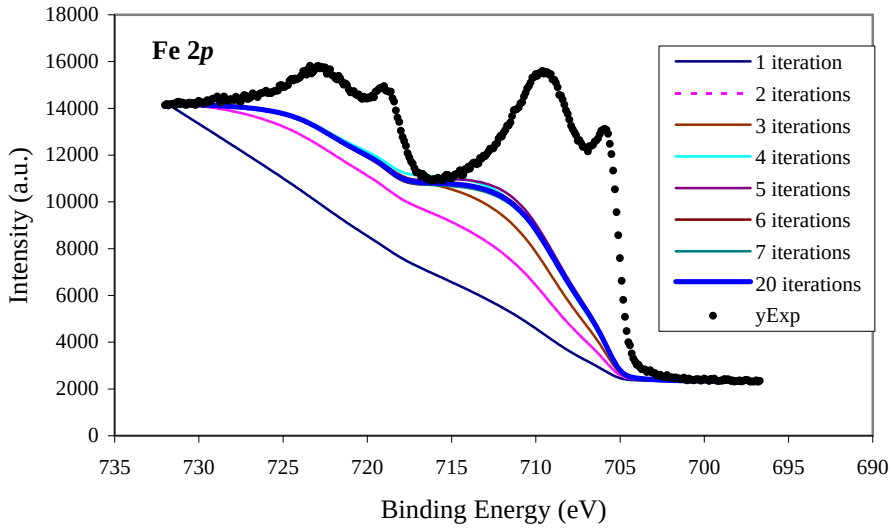

Figure 3. Traditional Shirley background with different iterations for a spectrum containing many peaks. The convergence is slower than for simpler spectra such as that shown in Figure 1.

I.3. **The active Shirley background**

The active Shirley background is calculated in the same way as in Equation 1 but without the restriction of Equation 2. The value of k is allowed to be optimized during the fitting process. Figure 4 shows the background calculated in an active way for various iterations for the fit shown in Figure 1b. The height of the ending point on the left, I_{left}, varies with the number of iterations. There is an oscillation between even and odd iterations. The convergence is fast.

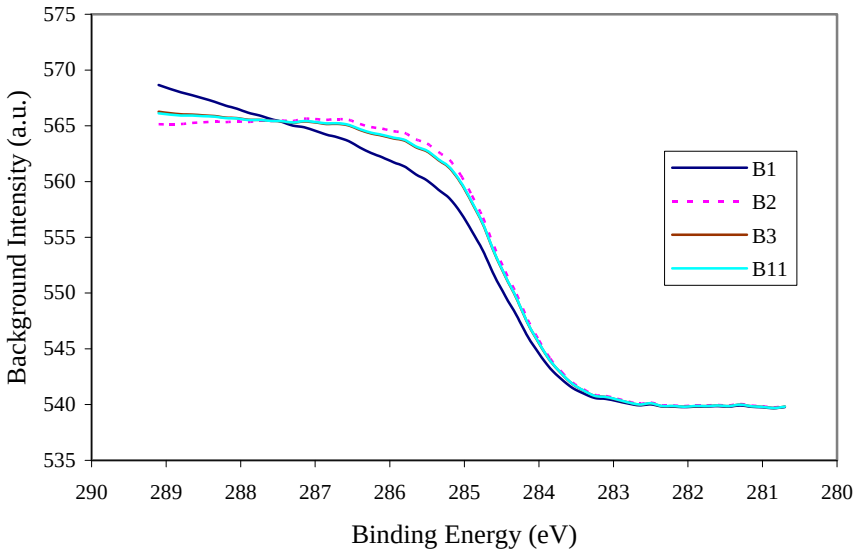

Figure 4. Active Shirley background for different iterations for the fit shown in Figure 1b. For this simple spectrum, the background does not change significantly after $n = 3$. The fit looks almost equally good from $n = 1$ since the change on the background is compensated by the peak intensity.

The convergence of the optimized Shirley background is also slower for complex spectra than for simple spectra. However, it does it faster than for the traditional Shirley background.

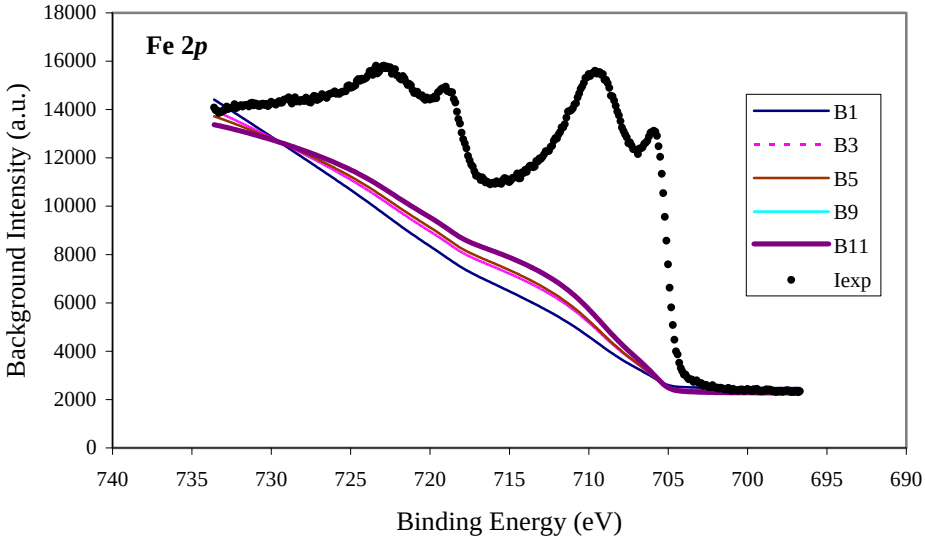

Figure 5. Optimized Shirley background for different iterations for a spectrum containing many peaks.

As widely described in Reference 1, the optimized Shirley background can be quite different from the traditional Shirley background. The main reason is that the optimized background is not forced to go through $(E_{\text{left}}, I_{\text{left}})$ or $(E_{\text{right}}, I_{\text{right}})$. This can be seen by comparing the backgrounds in Figures 3 and 5.

An important characteristic of the active approach is that it is compatible with more than one background. It is not necessary to choose between, e.g., either Shirley or Tougaard, or either lineal or Shirley. The software employed for the analysis presented in this Report allows for combinations of backgrounds.[5]

II. The Peak Shirley Background

There seems to be confusion in the literature differentiating between intrinsic and extrinsic background. Some authors portray the Shirley background as a special case of the Tougaard background.[8] This cannot be the case since the scattering can occur only once for the intrinsic case (Shirley), when the electron is leaving the atom. The scattered electrons cannot scatter again in the same way. The extrinsic background, however, is generated by any electron traveling through the solid. In fact, the extrinsic background requires accounting for the possibility of multiple scattering events.[9] The intrinsic and the extrinsic scattering takes place at different times and are independent events. The total background has contributions from both extrinsic and intrinsic scattering.

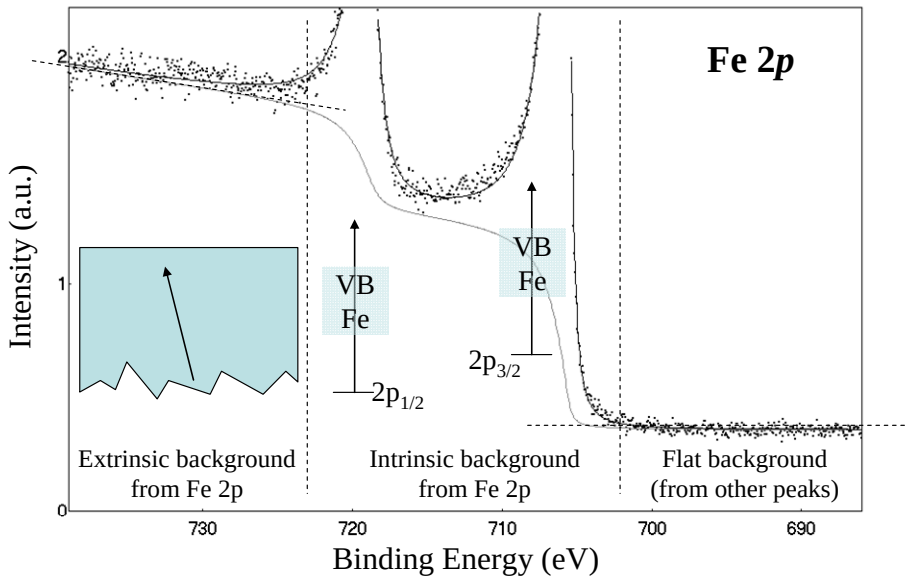

Figure 6. The background has different relative contributions depending on the energy region.

The peak-Shirley background is constructed by assigning a Shirley-type background to individual peaks. This makes complete physical sense since only those electrons coming from a core level can contribute to the intrinsic background. It not only makes physical sense, but provides much better fits than the standard or optimized Shirley background. This is illustrated in Figure 7.

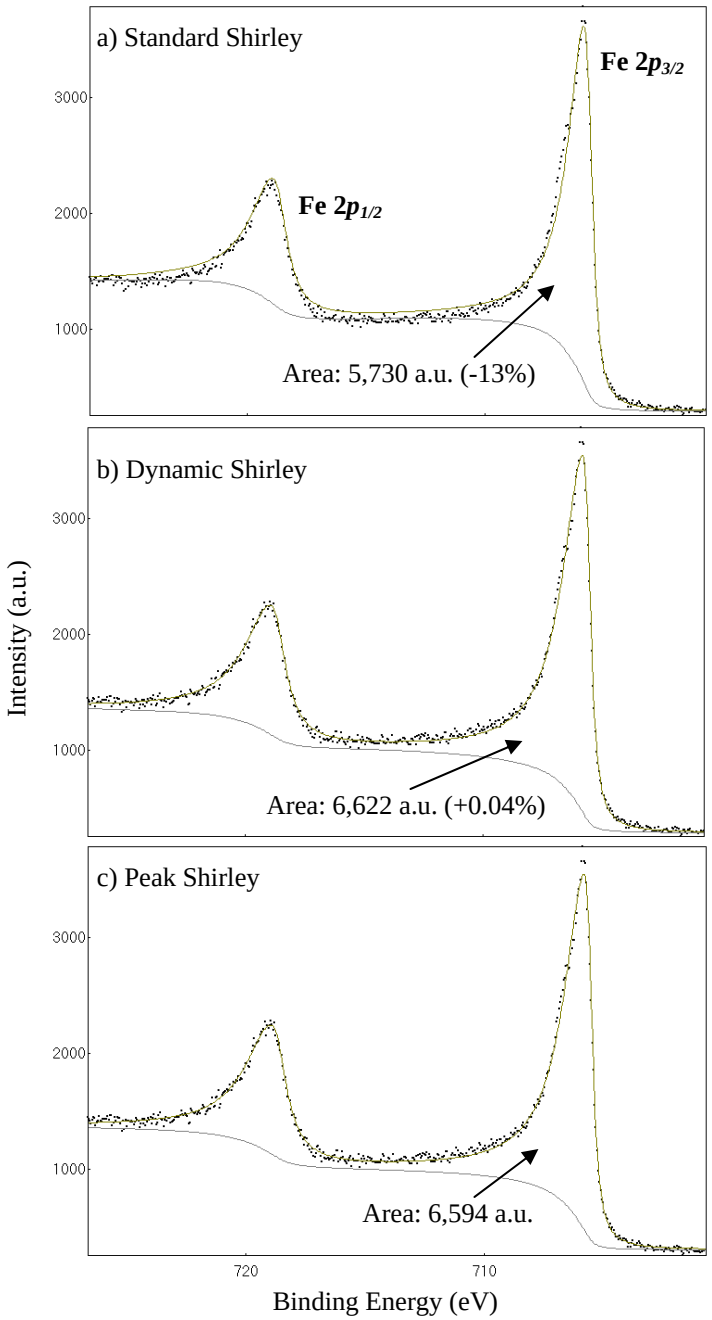

Figure 7. Fe 2p XPS spectrum fitted with a) standard, b) optimized, and c) optimized peak-Shirley background. The area of the $2p_{1/2}$ component was restricted to be one half of the corresponding to $2p_{3/2}$. The fit with the peak-Shirley method is superior even though some extra restrictions were imposed, such that the Gaussian width is the same for both branches. The area for the standard-Shirley method is subestimated because the background is forced to go through the end points. The active Shirley method provides approximately the same answer because both branches have the same scattering factor; this is not the case when the peaks have different scattering factors, as it is shown in Section III.

An additional advantage of the peak-Shirley background is that it does not require iterations.

III. The Peak-Shirley Background for Overlapping Peaks

The most important application for the peak-Shirley background is for overlapping peaks with different scattering factors. Figure 8 shows a Si 2*p* spectrum for which the substrate and oxide peaks are well separated and there is no overlapping. The fitting can be done separately for each of the two regions containing one, the oxide peak, and two, the bulk peak. Figure 8a shows the fit for the oxide peak; a Shirley background was employed, but the scattering factor came out as zero. The size of the step is directly related to the scattering factor. The presence of a signal from suboxide peaks in the region between the Si^{4+} and Si^{0} peaks would affect the step size. This issue is treated later on. The fit for the bulk peak was also done with a Shirley background (Figure 8b), and the value of the scattering factor derived by the optimization process was 0.03 eV^{-1}.

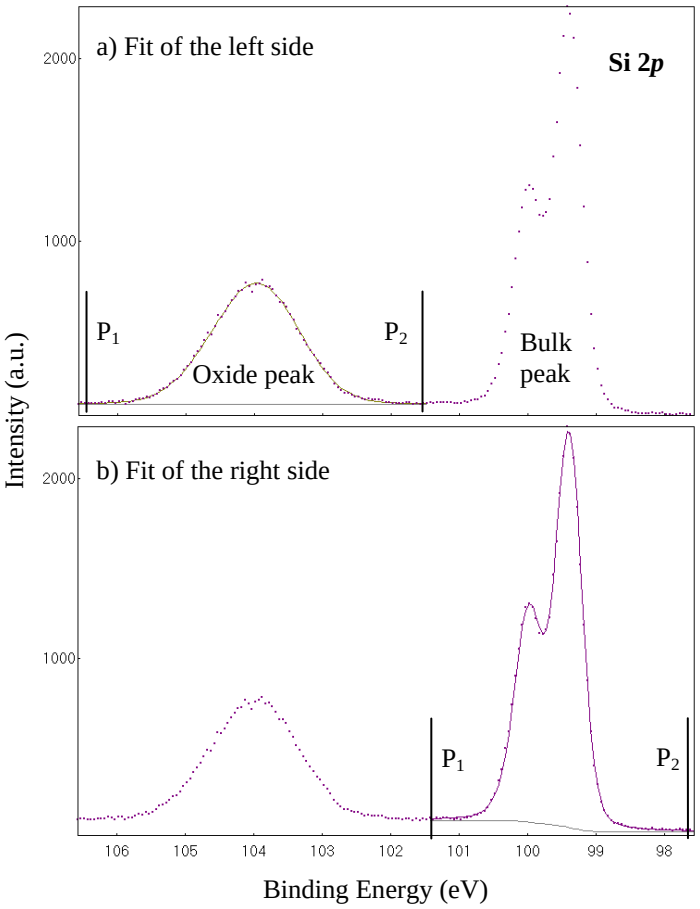

Figure 8. Si 2p spectrum fitted in two regions, each one with its own Shirley background. The peaks are not overlapping and each region can be fitted separately. The scattering factors came out as a) zero for the oxide peak, and b) 0.03 eV^{-1} for the bulk peak. The possibility of suboxide peaks affecting the strength of the scattering factor is discussed below.

Since the peaks are well separated in the spectrum shown in Figure 7, it was possible to define the points P_1 and P_2 for each peak. This is not longer the case for the spectrum shown in Figure 9.

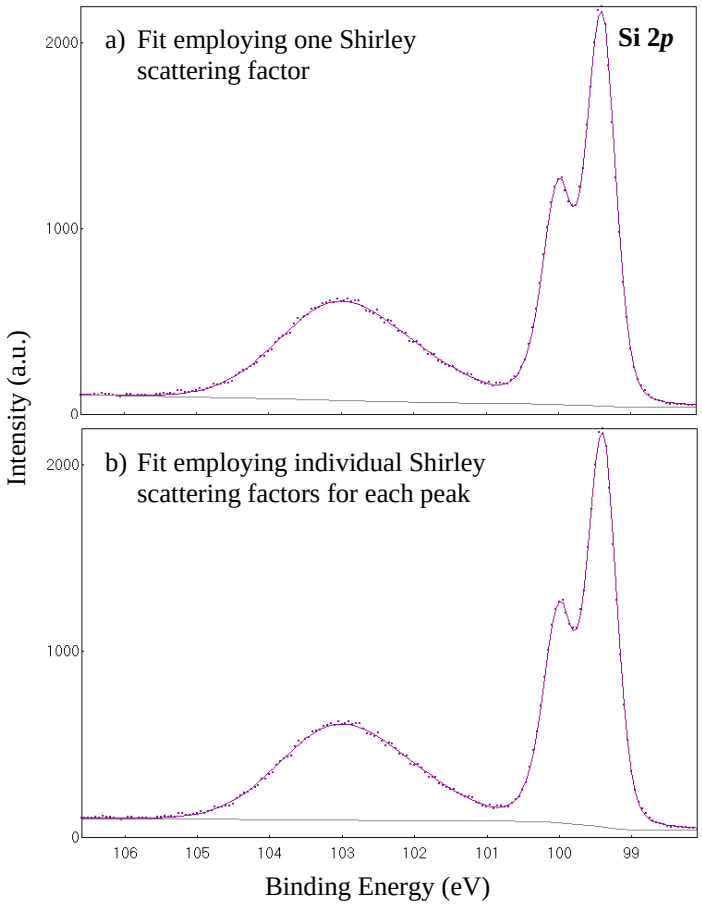

Figure 9. Si 2*p* spectrum of silicon oxide on Si. The Si^{4+} and Si^{0} peaks are closer together due to nitridation. The fit was done a) with one Shirley scattering factor and b) with predetermined scattering factors for each peak.

This issue was rose by Salvi and Castle in 1998.[4] They propose the use of "shape parameter" to quantify the Shirley scattering strength for each peak. Their parameter is not appropriate for quantification since it depends on the spectrum's resolution: the same peak take with different pass energy would yield to a different value (the higher the resolution, the smaller the shape parameter). They mention that

> *For these outcomes to be achieved it will be necessary for the program developers to devise a means by which the shape parameter can be entered and held constant as a fitting parameter.*

The software employed for the analysis presented in this Internal Report is capable to employ predetermined values of the scattering factor defined in Equation 1. It does it by simply clicking on the option of peak-Shirley background for any of the peaks employed in the analysis, and entering the value of the parameter.[5] The scattering factor can be held constant or allowed for optimization. This is illustrated in Figures 10 and 11.

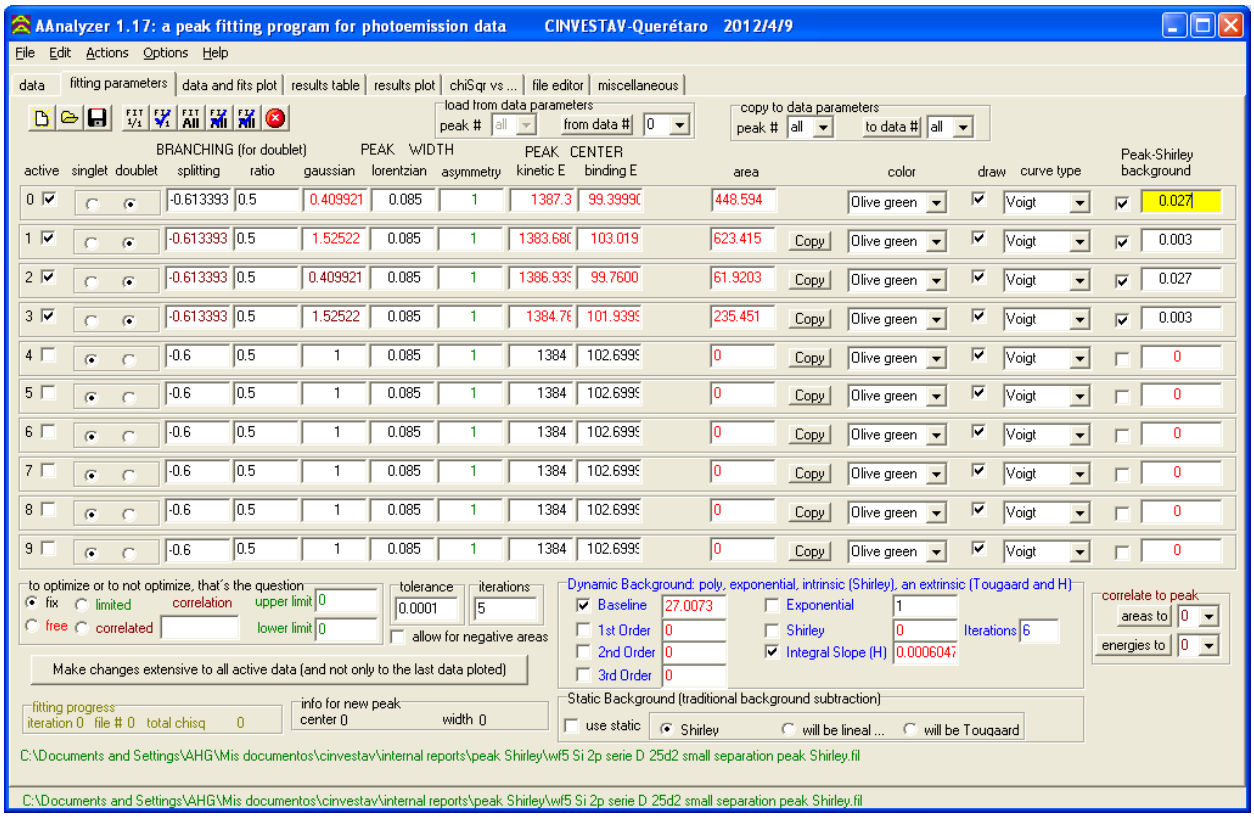

Figure 10. Parameters employed in the software AAnalyzer for the fit shown in Figure 9b. The scattering factor for the peaks associated to the bulk was held constant to a value of 0.027 eV^{-1}. For the peaks associated to the oxide, the value of the parameter was 0.003 e V^{-1}. These values come from the analysis shown in Figure 12.

Figure 11 shows the case where the scattering factor is optimized during the fitting process (the red color indicates that the parameter is allowed to vary).

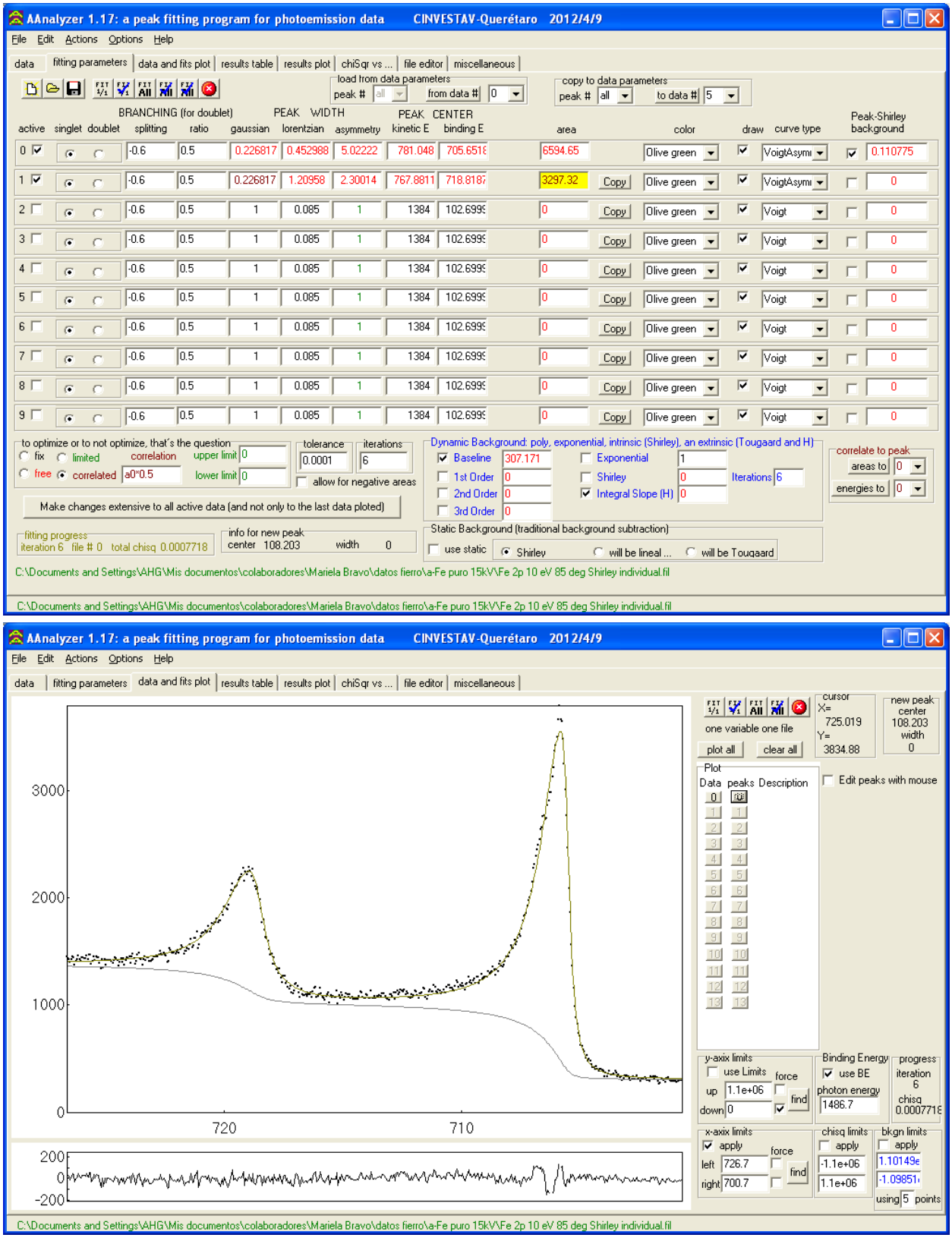

Figure 11. Parameters employed to fit a Fe 2*p* spectrum.

The predetermined scattering factors used for the fit shown in Figure 9b were obtained from a clean Si sample (Figure 12a) and a thick silicon oxide sample (Figure 12b). The size of the Shirley background steps, and therefore the scattering factors, could be reliably assessed since none of the samples shows any sign of silicon suboxide peaks.

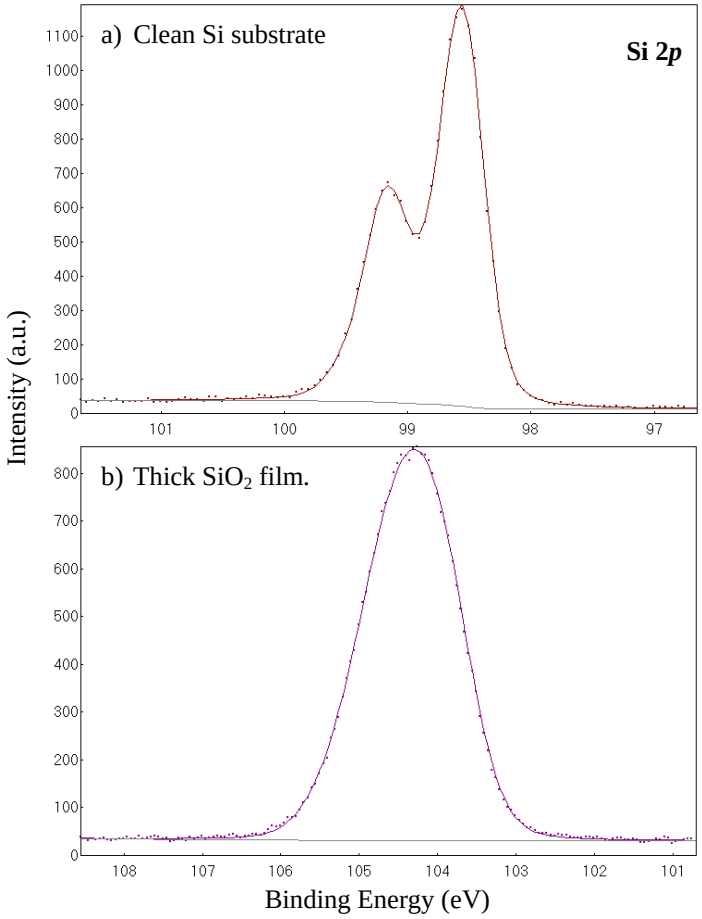

Figure 12. Spectra employed to assess the values of the Si 2*p* Shirley scattering factors for a) clean Si (0.027 eV^{-1}) and b) silicon oxide (0.003 eV^{-1}).

An important advantage of the scattering factor, as defined in Equation 1, is that it is transferable.

IV. Conclusions

The peak-Shirley background has many advantages:

- It makes physical sense since the background is generated only from electrons coming out in a photoelectric peak.
- It is transferable since it comes from an internal process. The presence of overlayers does not affect its value.
- It does not require iterations.
- It can assign different scattering factors to different peaks even when they overlap.

- Last, but not least, it provides better fits than the standard Shirley and even than the active Shirley.

References

[1] A. Herrera-Gomez. Internal Report. “A dynamical treatment of the background.” Internal Report. CINVESTAV-Unidad Querétaro. (Created: 3/2011. Last Update: 1/2011). http://www.qro.cinvestav.mx/~aherrera/reportesInternos/dynamicVsStaticBackground.pdf

[2] J. Muñoz-Flores, A. Herrera-Gomez. “Resolving overlapping peaks in ARXPS data: the effect of noise and fitting method.” J. Elec. Spec. Rel. Phen. 184 (2012) 533– 541, and references therin.

[3] J. Vegh. J. Elec. Spec. Rel. Phen. 46 (1988) 411.

[4] A.M. Salvi and J.E. Castle. J. Elec. Spec. Rel. Phen. 95 (1988) 45.

[5] More information about the software can be found at www.rdataa.com/aanalyzer.

[6] P.M.A. Sherwood, in: D. Briggs, M.P. Seah (Eds.), Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy, Wiley, Chichester, 1983 (Appendix 3).

[7] J. Vegh, J. Elec. Spec. Rel. Phen. 151 (2006) 159–164

[8] J. Vegh, Surf. Sci. 563 (2004) 183.

[9] S. Tougaard, Surf. Interf. Anal. 11 (1988) 453.