# Optimization of XPS Inelastic Background using Tougaard's Formula - Recent Progress

Masatoshi Jo
*National Metrology Institute of Japan*
*National Institute of Advanced Industrial Science and Technology*

*AIST Tsukuba Central 3-9, Umezono 1-1-1, Tsukuba, Ibaraki 305-8563, Japan*
(m-jo@aist.go.jp)

(Received March 19, 2002; accepted May 30, 2002)

Recent development of background optimization is described. Assumed peak intensities and boundaries for integration that were fixed in the former version can be treated as optimization parameters. The 'effective' loss function is allowed to take negative value to analyze non-uniform material. In addition to peak ratio and tail intensity, peak intensity is treated as a new optimized functional. The Al XPS spectrum is analyzed.

## 1.Introduction

According to the 3-step model of X-ray photoelectron spectroscopy (XPS), a photoelectron ejected from the original atom upon photon absorption travels through the solid, escapes from the surface and finally enters into the analyzer. During the travel in the solid, the photoelectron loses its kinetic energy probabilistically through inelastic interactions with surrounding electrons. The photoelectrons that lost kinetic energy form a broad and large inelastic background on the spectrum, which has long been a serious obstacle against precise evaluation of peak intensity and peak shape. Since the phenomenon is a result of complicated multiple scatterings, it had not been possible to be solved until Tougaard [1] gave the analytical formula treating this effect correctly.

In the case of uniform material, Tougaard's formula is reduced to

$$j(E) \propto \frac{j_0(E)}{\lambda(E)} - \int_E^\infty j_0(E')K(E',E)dE' \quad (1)$$

$j_0(E')$ is the observed photoelectron energy distribution (spectrum) with kinetic energy $E'$, $K(E',E)$ is the probability of inelastic scattering of photoelectron from $E'$ to $E$, $\lambda(E)$ is the inelastic mean-free path of a photoelectron of kinetic energy $E$ inside the solid, and $j(E)$ is the spectrum after background removal. The integral represents the generated inelastic background. $K$ is called the "loss function".

In the above formula, $j_0$ and $\lambda$ are known but $K$ and $j$ are generally not known. Therefore eq.(1) cannot be solved unless additional conditions are considered. Tougaard calculated the background using $K$ obtained from an electron energy-loss spectrum or from a concise analytical form called the universal function. It is very convenient to use the universal function because it has only two or three adjustable parameters [2].

In 1995 Jo[ 3 ] proposed another approach to solve eq.(1) using optimization. It was shown that a realistic loss function and an inelastic background were simultaneously obtained by making use of some assumptions described below, even if the measured spectrum was smooth and looks to have lost all the detailed information of inelastic scattering. Though this was a unique feature, it also had problems.

The method's application was limited to only uniform material, such as a pure metal. Further, strictly speaking, even a pure metal is not uniform for the photoelectrons because of surface excitation. Thus the formalism cannot be used for practical analysis.

As shown in the next section, the

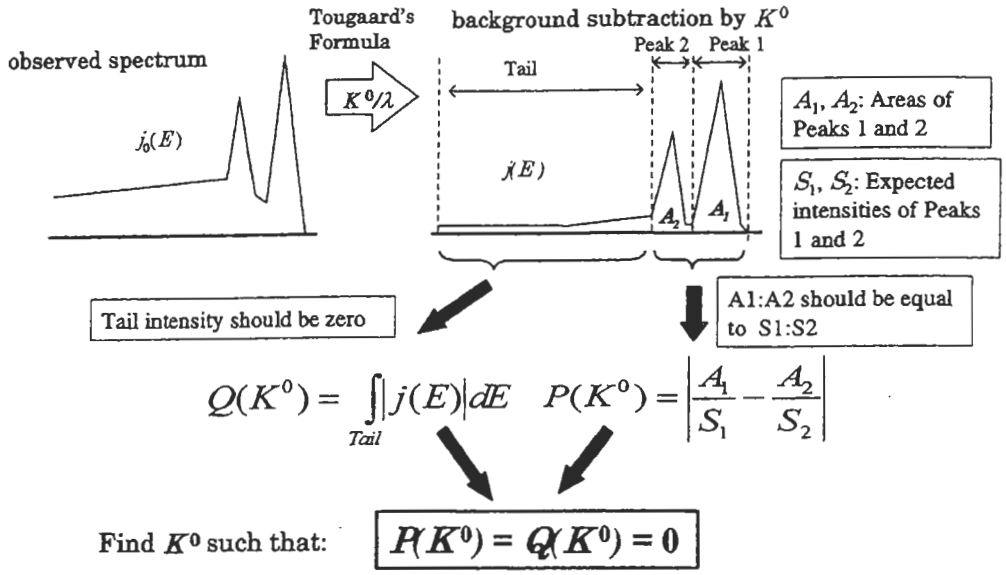


Fig.1 Principles of optimization process using Tougaard's Formula.

assumptions used in the formalism include some simplified constants about the system that is to be analyzed, in order to maintain universality of the method. Peak intensity and the extent of a peak are such quantities. On the other hand, the real material's complexity could change their values. In general, the purpose of analysis is not to analyze a known material with known values but to analyze an unknown one, whose chemical composition and structure are also unknown. Therefore one should not expect that sample-specific constants are known prior to analysis. Exhaustive trials by changing the values would not be realistic when there are several such uncertainties.

In the present paper, efforts to overcome these difficulties are reported. The contents of this paper are as follows. The general formalism, problems, and improvements are described in sections 2, 3 and 4, respectively. In the final section, the first results for aluminum by application of the new optimization method are shown.

## 2. Background optimization

The assumptions used in the background optimization method are as follows. The primary excitation spectrum, defined as a hypothetical energy distribution of the photoelectron at the very moment of ejection from the atom, is determined only by the nature of the original atom and its neighborhood. The peak intensity is expected to be proportional to the photoelectron excitation probability of an isolated atom, though intensity redistribution may occur due to final state effects such as shake-up satellites. By definition, the primary excitation spectrum has no inelastic background.

The algorithm was designed so that satisfying the above assumptions is equivalent to solving an optimization problem as defined below. (Fig.1)

Two functions for this purpose are defined.

1) The peak intensity ratio of a selected pair of the same element, such as 2p3/2 and 2p1/2, is equal to the corresponding ratio of photoelectron excitation probabilities if the correct background is removed using the correct $K$. This is interpreted as minimizing $P(K^0)$.

$$P(K^0) = \left| \frac{A_1}{S_1} - \frac{A_2}{S_2} \right|$$

$A_1$ and $A_2$ are the measured peak intensities after background subtraction; $S_1$ and $S_2$ are the expected intensities. $K^0$ is the part that depends only on the amount of energy loss, $K(E',E)=[1/\lambda(E')]*K^0(E'-E)$ according to Tougaard [1]. $P$=0 if $A_1:A_2=S_1:S_2$. For Al $K\alpha$ X-rays, the photoelectron excitation cross sections calculated by Scofield are used as $S_i$'s [4]. The expression $P(K^0)$ means that $P$ is a functional of $K^0$.In practice $K^0$ is represented by a set of values $K_i^0$ at loss energy $E_i$, i.e. $P(K^0)=P(K_1^0, K_2^0,...)$. The step of loss energy $(E_{i+1} - E_i)$ is typically the same as that of the original spectrum. The number of $K_i^0$'s is 200 - 300 for a typical spectrum.

In calculating a peak area $A_i$, the pair of end points between which integration is done is intuitively assumed prior to calculation.

2) Far from the peak (denoted as *Tail*), there should be only a background contribution and no peak intensity should be observed if the correct background is removed using the correct $K$. This corresponds to minimizing $Q(K^0)$.

$$Q(K^0) = \int_{Tail} |j(E)| dE$$

Again, the integration boundaries for *Tail* are given prior to the calculation.

During the optimization, $K$ is assumed to be positive, because it is a probability.

For optimization, the successive quadratic programming method (SQP) written by Fukushima [5] was used. SQP can solve a non-linear optimization problem of function $F$ of $n$ variables with $m$ equality and/or inequality constraints $c_j \geq 0$ or $c_j = 0$, where $n$ and $m$ are arbitrary. When $P$ is chosen as $F$, $Q$ is included in a constraint as $c_Q = (Q_0\text{-}Q) \geq 0$, which means $Q$ must not exceed a limit $Q_0$, and vice versa. To choose which of them ($P$ or $Q$) as $F$ depends on the particular situation. Also an area $A_{neg}$, negative portion of spectrum $j(E)$ after background removal, is treated as an additional constraint $c_{neg} = -|A_{neg}|$, which means $A_{neg}$ must be zero because negative $j(E)$ is meaningless. SQP searches for appropriate $K_i^0$'s considering above conditions.

## 3. Problems

In order to analyze real materials in various situations, the above formula and assumptions are not satisfactory due to the following reasons.

1) Measured peak intensity $A_i$ for a given integration interval may not become necessarily the same as $S_i$ because elastic scattering inside the solid might have changed the angular distribution of the photoelectron from that predicted by the asymmetry factor. Also if two peaks lie close together, the change by the overlapped area may be not negligible (Fig.2). Further, the photoabsorption cross sections have been calculated only for a few selected discrete photon energies[4, 6], and are not available for experiments using nontraditional photon source, e.g., synchrotron radiation or anodes other than Al or Mg. This limitation means that the formalism assuming fixed $S_i$'s and integration boundaries is not sufficient.

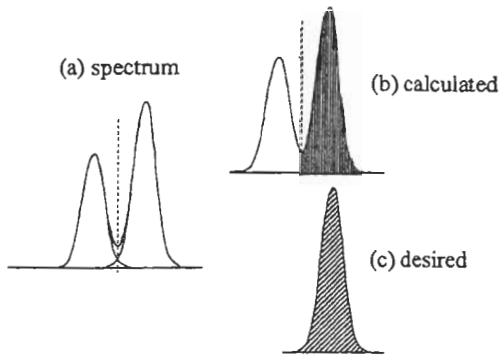


Fig.2 Calculated and desired intensities are different due to overlapping.

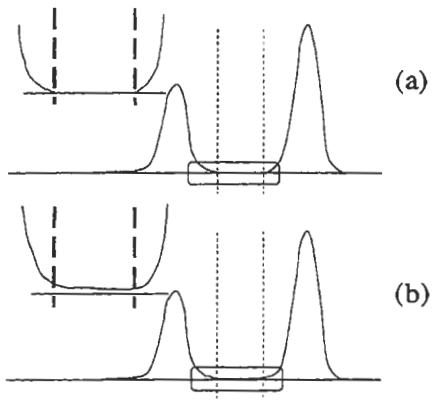


Fig.3 (a) No intensity is assumed outside the peak. (b) Small intrinsic contribution may exist.

2) The boundary separating the peak from other regions may not be clearly defined due to the presence of (generally continuous) intrinsic satellites (Fig.3). Therefore, a wider interval should be ascribed to a peak, and/or $Q_0$ may not be very small. This requires a more robust and stable algorithm.

3) The material is generally not uniform. A structure change, composition change, or surface excitation may exist. These make eq.(1) invalid.

The general formulae are written as [7,8],

$$F(E,\Omega) \propto J(E,\Omega) - \int dE' J(E',\Omega) K_{General,\theta}(E,E')$$

$$K_{General,\theta}(E,E') = \frac{1}{2\pi}\int ds \exp[-is(E-E')]\left[1-\frac{P_{1,\theta}}{P_\theta(s)}\right]$$

$$P_{1,\theta} = \int_0^\infty f(x)\exp[-\int_0^x \frac{1}{\lambda(z)\cos\theta}dz]dx$$

$$P_\theta(s) = \int_0^\infty f(x)\exp(\frac{-\Sigma(x,s)}{\cos\theta})dx$$

$$\Sigma(x,s) = \int_0^x\int_0^\infty K(T,z)[1-\exp(-isT)]dTdz$$

where $K(T, z)$ is the depth-dependent loss function for energy loss $T$ and depth $z$. $f(x)$ is the number of atoms per Å of signal element at depth $x$, $\lambda(E)$ is the inelastic mean-free path, and $\theta$ is the angle between the photoelectron detection direction and the surface normal. These results show that eq(1) is still valid but the obtained loss function should be interpreted as $K_{General}(E,E')$, which is related to the true loss function $K$ via the above complicated equations. Since $K_{General}$ is not strictly a probability, it can take negative values. Therefore, the algorithm should be changed to allow negative values.

## 4. Improvements

### 4.1 Coefficients $S_i$

Since the present optimization routine SQP [5] accepts any number of variables as long as they are independent, variables $s_i$'s that describe $S_i$'s can be added as follows.

$$S_i = [1 + \frac{s_i}{\sqrt{1+s_i^2}}] \times S_i^0$$

$-\infty < s_i < \infty$, $0 < S_i < 2S_i^0$

In SQP, it is convenient not to restrict the range of the variable $s_i$. On the other hand, $S_i$ must not take physically meaningless values. $S_i$ also must be a smooth function of $s_i$ (Fig.4). The above function was chosen in order to fulfill these requirements. Any other function will work as long as the same requirements are satisfied.

Initial guesses of $s_i$'s are calculated from the peak areas using convenient background methods, e.g. linear, Shirley etc.

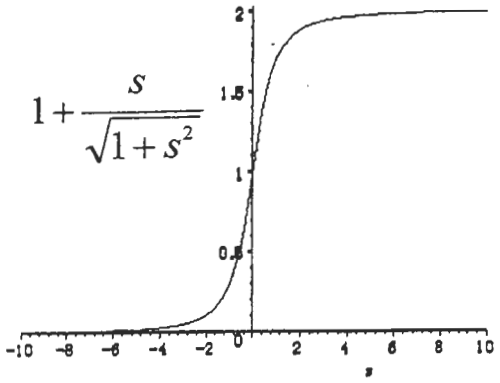


Fig.4 Relation between conefficient $S_i$ and corresponding internal parmeter $s_i$.

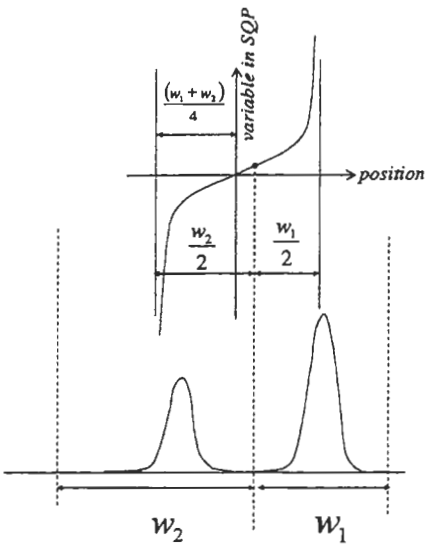


Fig.5 Initial position and the range of variable.

It is noted, for each pair $S_i$ and $S_j$, one of which is fixed and the other is allowed to vary since only the ratio $S_i/S_j$ is important in $P$.

It is also noted that $P$=0 is instantly attained if $S_i$ is varied, however it is not the solution unless other constraints are simultaneously small.

### 4.2. Position of Boundary

It is convenient to make peak boundaries movable. A boundary must not pass through its neighbors. This requirement suggests the form similar to $S_i$. (Fig.5) $w_1$ and $w_2$ are widths of peaks or tails on both sides of the boundary. When the internal variable moves from $-\infty$ to $\infty$, the boundary cannot exceed beyond half of the neighboring intervals.

### 4.3. Another function $R(K^0)$

In addition to $P$ and $Q$, the peak intensity has been introduced as a third functional $R$

$$R(K^0) = \int_{A_1} j(E)dE$$

in order to enhance robustness and stability. It is mainly used as a constraint to be kept within lower and/or upper limits. $R$ is, like $P$, a condition on the low energy-loss part of the loss function, which is reinforcement to $P$.

4.4. Negative $K$

There are two reasons that prohibited $K$ from becoming negative in the previous formalism. One is because $K$ was considered as a probability. The other is to avoid divergence. It is known that an integral equation like (1) is not well-behaved because an arbitrary function that changes sign rapidly and therefore gives a vanishing contribution to the integral can be added to the solution $K$. In order not to restrict the internal variable $k_i$ in SQP, this was done by
$K_i^0 = k_i^2$
where $k_i$ is the $i$-th internal variable describing $K^0$ in SQP. In order to allow $K_i^0$ to become negative, this is simply replaced by
$K_i^0 = k_i^2 - B$, $B > 0$.
Therefore $K_i^0 \geq -B$.
Generally, this transformation would not be necessary when the algorithm becomes more robust and stable.

4.5. Objective function and constraints

As mentioned above, one of $P$, $Q$, $R$ is chosen as an objective function that is to be minimized. Others are treated as additional constraints. For example, if $P$ is chosen, the constraints are

$c_Q = Q_0 - Q$,
$c_R = R_0 - R$, and
$c_{neg} = -|A_{neg}|$.

$Q_0$ and $R_0$ are given adjustable constants representing upper limits of $Q$ and $R$. Also, lower limit can be chosen for $R$ by defining $c_R = R - R_0$.

During the calculation, a solution that minimizes the objective function is searched for, while checking the violation of constraints. Therefore the convergence is controlled by these limit values. Thus the complexity of giving various material-specific constants prior to calculation has been greatly reduced in the present formalism.

## 5. Analysis

The present program is running on dual Pentium II Xeon (2MB cache) PC with Windows NT4.0. The number of variables is between several tens and several hundreds. The typical calculation time is between a few seconds and a few minutes, depending on the number of variables and the data size, to reach the converged solution from the initial function form that simulates Tougaard's universal function with B=3047, C=1100. It is noted that convergence is not always attained for arbitrary limit values.

The data analyzed here is the same Al spectrum [9].used before [3], which is a part of a wide-scan spectrum. In the previous study [3] the 2s and 2p intensities were assumed to occur only in the main peaks, and regions where plasmon satellites appeared were treated as the *Tail*. Also, the loss function was restricted to be non-negative. In the present analysis, the main plus first and second plasmon satellites are treated as one peak (Fig.6), i.e., the total integrated area is considered and no assumptions on the shape, intensity and the position of satellite peaks are made within the interval. This means that the optimization condition became much looser. $P$ is chosen as the objective function, and $c_R = R_0 - R$. For the inelastic mean-free path, the calculation by Tanuma et al. is used.[10]

For the expected intensities, $S_{2s}$ is fixed and $S_{2p}$ is varied. The ratio of initial values is tentatively given for $S_{2p}/S_{2s} = 0.71$. $K^0$ is allowed to be negative. However the peak boundary remains fixed because convergence failure occurred. Examples of adjusting peak boundaries will be reported in a separate paper. The variables $K_i$ are allowed to change up to the cutoff loss energy $E_c$, beyond which $K_i$ is fixed to zero. In principle $E_c$ should be sufficiently large. However, this will cause another instability as shown below.

Fig. 6 shows the measured spectrum, the calculated background (dotted curve) and the spectrum after background removal for the

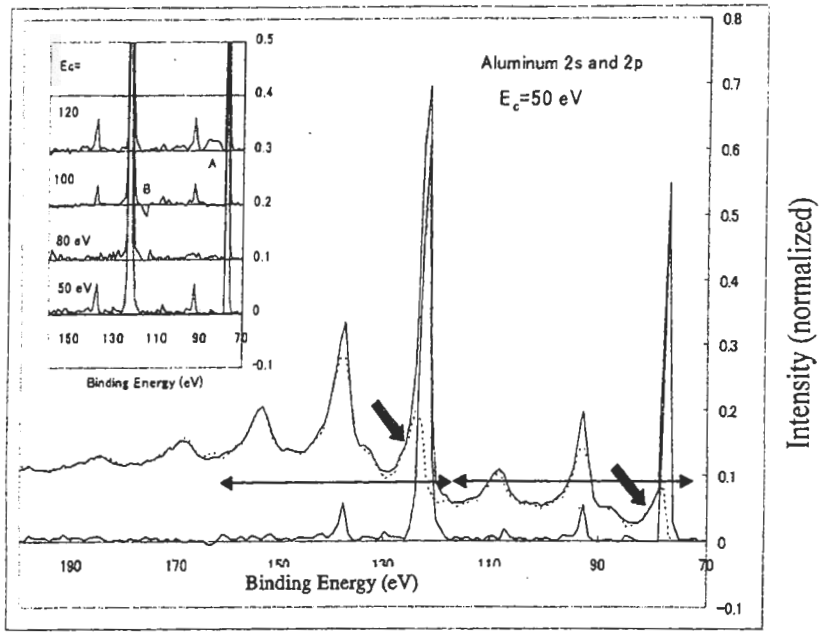


Fig.6 Analyzed Al 2p and 2s peaks with satellites for $E_c$= 50 eV. Dotted curve is the calculated background. Inset: Magnified views of background-removed spectra including other conditions. See text for features A and B, and inclined arrows. Horizontal arrows indicate the intervals used for intensity evaluation.

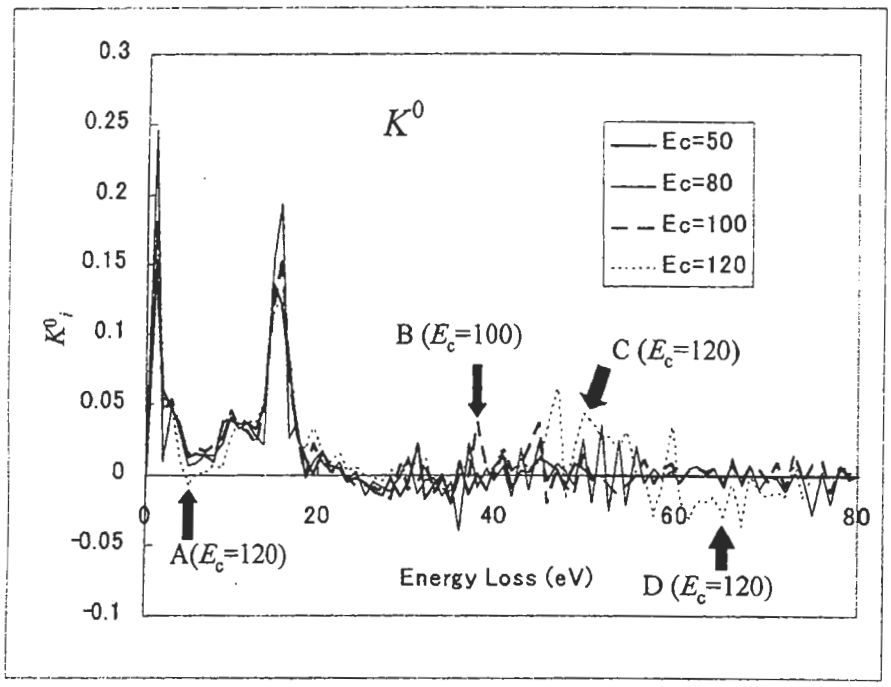


Fig.7 $K^0$'s obtained from different cutoff $E_c$. Features A and B produce bump and dip in the spectrum (inset of Fig.6), respectively. Feature C cancels the bump to appear beside 2s peak for $E_c$=120 due to feature A. Feature D shows the instability occurred.

case $E_c$=50 eV with magnified views for all $E_c$'s in the inset.

Though not very clear especially for $E_c$=80 eV, a residual satellite peak located at 15 eV below the Al 2s and 2p peaks is observed. This may correspond to the 'intrinsic' plasmon satellite suggested earlier [11]. When $R_0$ is reduced until $c_R < 0$, a solution without a plasmon satellite is reached. In this condition, the bulk plasmon and the spike (described below, Fig.7) at 1 eV become very large, which looks unphysical. On the other hand, when $R_0$ is increased, the result does not change very much. This will be reported elsewhere.

The final ratio of $S_{2p}/S_{2s}$ and $A_{2p}/A_{2s}$ are always equal but dependent on $E_c$, and are 0.60, 0.64, 0.73 and 0.82, for $E_c$=50, 80,100, and 120 eV. respectively. It is noted that the value for $E_c$=120 eV includes bumps (marked as A in the inset) beside the main and the satellite peaks. These bumps (decrease of background) are produced by the decrease of $K^0$ at ca. 5 eV loss energy (A in Fig.7). Bump beside 2s main peak is cancelled by the contribution of the energy loss from 2p, which is caused by the structure about 50 eV (C in Fig.7) in $K^0$. A dip (B) in the inset is due to corresponding features in $K^0$ (Fig.7). The data scatter of the ratio is ascribed to the 'spurious' structures thus generated. In fact, the ratios $A_{2p}/A_{2s}$ using the areas which are sufficiently narrow and only includes main and 1st plasmon satellite are 0.61, 0.62, 0.65, and 0.60 for $E_c$=50, 80,100, and 120 eV, respectively. Thus the results which are

considered physically meaningful are not strongly affected by the apparent difference of $K^0$.

Fig.7 shows $K^0$'s with different cutoffs $E_c$=50, 80, 100, and 120 eV. The peak positions of the lower-loss-energy parts are very similar for all cases though their heights are slightly different. A spike at *ca.* 1 eV, and surface and bulk plasmon peaks are observed. Beyond the bulk plasmon, the loss function becomes slightly negative, with a peak at double the bulk plasmon loss.

The negative value is understood as follows. Roughly speaking, the larger energy loss corresponds to the longer travelling path of the photoelectron considering multiple scatterings. Thus background far from the main peak includes larger contribution from deeper depth. If the loss function at deeper depth is smaller, the observed background far from the peak becomes smaller than the value predicted by the multiple scatterings using the uniform loss function. The 'apparent' loss function calculated in the present method becomes negative to compensate for the difference.

A solution that is without a spike at 1 eV and still shows reasonably good background has been never obtained. This peak appeared to adjust the shape of the lower foot of the main peak (inclined arrows in Fig. 6). Simonsen et al. [12] reported that the effective loss function of aluminum shows a divergence at 0 eV due to sudden hole creation, which seems to be in agreement with the present result.

For larger loss energies, the loss function stays close to zero. However, for larger $E_c$, instability develops in the higher energy loss region (feature D in Fig.7 ). Feature D is considered to have appeared in order to cancel the effect of features C and A. This might also partly be due to the 'constant' background subtraction usually done prior to the present analysis. Since the true background to be thus subtracted may not be perfectly constant, the intensity far from the peak is not reliable.

In summary, improvements of background optimization toward practical analysis and result of Al have been shown. It is very much interesting that peak intensity can be an adjustable parameter. This suggests that intensity information is hidden in the spectrum itself and restorable as long as the analysis is done according to Tougaard's formula, in other words as long as it is caused by multiple scatterings.

**Acknowledgments**

This work was supported by the Special Coordination Fund (*Photoelectron Microscopy*) of the Ministry of Education, Culture, Sports, Science and Technology, and also by *the Material Nanotechnology Program* of New Energy and Industrial Technology Development Organization (NEDO). Also, stimulating discussions with Dr. I. Kojima (AIST), Dr. T. Tomie (AIST), Dr. S. Tanuma (NIMS), Dr. A. Tanaka (ULVAC-PHI), and Dr. S. Ichimura (AIST) are gratefully acknowledged.

**References**

[1] S. Tougaard, J. Electron Spectrosc. Relat. Phenom., **52**, 243 (1990).
[2] S. Tougaard, Surface and Interface Analysis, **25**, 137 (1997).
[3] M. Jo, Surf. Sci., **320**, 191 (1994).
[4] J. H. Scofield, J. Electron Spectrosc. Relat. Phenom. **8**, 129 (1976)
[5] M. Fukushima, Math. Programming, **35**, 253 (1986).
[6] S. M. Goldberg, C. S. Fadley and S. Kono, J. Electron Spectrosc. Relat. Phenom. **21**, 285 (1981).
[7] S. Tougaard and H. S. Hansen, Surf. Interface Anal. **14**, 730 (1989).
[8] S. Tougaard and P. Sigmund, Phys. Rev. B **25**, 4452 (1982)
[9] Spectrum in the database by B. V. Crist (XPS International) is used.
[10] S. Tanuma, C. J. Powell and D. R. Penn, Surf. Interface Anal., **11**, 577 (1988).
[11] P. Steiner, H. Höchst and S. Hüfner, Z. Phys. B **30**, 129 (1978).
[12] A. C. Simonsen, F. Yubero and S. Tougaard, Phys. Rev. B**56**, 1612 (1997).