Deconvolution of SIMS Depth Profiles of Boron in Silicon

Brice Gautier, ' R Prost, G Prudon, J C Dupuy

SURFACE AND INTERFACE ANALYSIS, VOL. 24, 733–745 (1996)

Deconvolution of SIMS Depth Profiles of Boron in Silicon

B. Gautier,*[1] R. Prost,[2] G. Prudon[1] and J. C. Dupuy[1]

[1] Institut National des Sciences Appliquées de Lyon, Laboratoire de Physique de la Matière, UMR CNRS 5511, Bât. 502, 20 Avenue A. Einstein, 69621 Villeurbanne Cedex, France

[2] Institut Natonal des Sciences Appliquées de Lyon, CREATIS, URA CNRS 1216, Bât. 502, 20 Avenue A. Einstein, 69621 Villeurbanne Cedex, France

We have measured the depth resolution function of the SIMS analysis of boron in silicon for different experimental conditions and fitted this function with an analytical expression initially proposed by Dowsett *et al.* We use this analytical depth resolution function for the implementation of an iterative deconvolution algorithm, taking into account several properties of the signal, such as positivity and regularity. This algorithm is described precisely.

The algorithm is tested on several theoretical structures and then implemented for the deconvolution of real structures of boron-doped silicon layers in silicon. In particular, a sample constituted by six consecutive delta layers and a 75 Å thick layer are deconvolved. It is shown that the asymmetry of the profiles is completely removed and that the full width at half-maximum of the deconvolved delta layers can be reduced down to 41 Å. It is also shown that a layer whose real thickness is smaller than the measured width of the resolution function can be easily distinguished from a delta layer, and its thickness estimated.

INTRODUCTION

The microelectronics industry has reached a level where the dimensions of the new devices do not exceed a few tens of angstroms. Therefore, the depth resolution of the characterization techniques must be increased in order to provide reliable and precise information on structures that are sometimes below their power of resolution.

The analysis by secondary ion mass spectrometry (SIMS) is one of the most powerful tools for the characterization of very thin structures, because of its great sensitivity. In a magnetic sector instrument at low primary beam energy the SIMS depth resolution in silicon samples is governed mostly by the collisional mixing, and in the case of an oxygen primary beam by the incorporation of the primary oxygen ions in the matrix of the sample (swelling). Those phenomena, which are responsible for the artificial broadening of the profiles, are inherent in the measurement process and can only be limited by improving the performances of a SIMS instrument. The lowering of the primary ion energies seems to be a good solution and low-energy columns have already been developed, but the price to pay for the improvement of the depth resolution is a great increase in the measurement time. Moreover, the development of more and more sophisticated instrumentation might give rise to an increase in the skills of SIMS operators, which is already very high.

This is why the development of alternative solutions, like the deconvolution of the profiles, is needed. Deconvolution will not replace either a good experiment nor instrumental improvements (which are obviously the only way toward the best resolution), but it allows better depth resolutions to be reached from a careful analysis. We will discuss this point further in the text.

Some attempts have already been carried out using deconvolution procedures based on the maximum entropy method[1] (inverse methods) or on Fourier transforms[2] (direct method). The object of this paper is to present another way of deconvolving the SIMS depth profiles by implementing an iterative algorithm.

The first step of our work has been to find the depth resolution function (DRF) of the SIMS experiment in the case of an Si:B delta layer in silicon. This DRF has been fitted by an analytical expression and the experimental conditions have been changed in order to verify that the parameters that govern the analytical expression vary in a satisfactory way. Then, we have performed the deconvolution of simulated profiles in order to test the possibilities of our algorithm. The last step was the deconvolution of real profiles using an analytical expression of the DRF.

THE RESOLUTION FUNCTION

General properties and analytical expression

In signal processing, the recovery of signals that have been distorted by some linear and shift invariant process is called deconvolution.[3] To restore the initial signal from its output, one has to solve the convolution equation

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)\,\mathrm{d}\tau + v(t)$$

where $y(t)$ is the output signal, $x(t)$ is the input signal, $h(t)$ is the resolution function and $v(t)$ is the additive noise.

* Author to whom correspondence should be addressed.

CCC 0142–2421/96/110733–13

Received 26 February 1996
Accepted 7 June 1996

The resolution function is a specific function that has to be determined for each set of experimental conditions. In the case of a SIMS experiment, the DRF changes each time the primary beam energy, the impact angle, the matrix or the impurity under investigation changes.

First of all, in order to prove the correctness of this model in the case of the SIMS experiment, one must verify that the SIMS analysis is a linear and a shift invariant process. In the case of the analysis of boron in silicon, there is no degradation of the depth resolution due to the formation of ripples or uneven crater bottom. It has been verified in the case of a bombardment of Si by an oxygen beam, that the DRF can be considered as constant once the steady state is reached, until the eroded depth is at least 8000 Å.[2] The linearity of the process is verified so long as the concentration of the boron-doped layer stays below the dilute limit.[4] Thus, the assumption on which SIMS analysis can be described in terms of the convolution of an initial profile with a DRF that depends on the instrument and the analysis conditions seems to be valid, as has already been pointed out by Dupuy *et al.*[5]

It is easy to see that if the input signal is a delta function, then $y(t) = h(t)$. The DRF can thus be found from the measurement of a delta function. This function has to be determined for each specific combination of matrix, dopant and experimental condition. From a practical point of view, the elaboration of delta-doped structures of uniformly doped layers with ultrasharp interfaces is better controlled,[5] so that it is possible to get the DRF from the measurement of a rapid thermal chemical vapour deposition (RTCVD)-grown sample, which can be supposed to be very abrupt and very thin.

The deconvolution process can be performed with a numerical DRF, experimentally determined. Another way is to use an analytical function that perfectly fits the experimental data. This method leads to some advantages:

(1) The noise in the DRF is smoothed out. (We can assume that the noise is a consequence of the measurement, and not an inherent parameter of the DRF.) Thus some of the artifacts in the result of the deconvolution can be avoided.
(2) The sampling interval of the analytical DRF can be easily adjusted to match that of other experimental profiles to be deconvolved, especially where the erosion rate is not exactly the same as in the initial DRF measurement.
(3) If the DRF is experimentally measured for a lot of energies, it is possible to determine the variation of the fitting parameters with respect to the energy: a better determination of these parameters for a given energy can be obtained by checking a continuity criterion.
(4) The possibility of describing the depth resolution with an analytical expression, via its moments, makes the performance of the analysis more comparable for different users working with different apparatus.
(5) One is free to choose the extent of the analytical DRF so that the fitting covers only the experimental data (in that case, the dynamic range of the analytical DRF is the same as the experimental one) or to extrapolate the DRF to simulate a very large dynamic range. In the first case, where the use of an analytical function is just a smoothing of the analytical data, there is no reason why the analytical form should not be implemented in a deconvolution scheme, as claimed by Dowsett,[6] if we assume that our sample is a real delta layer (MBE-grown samples are known to be very abrupt) and that the entire response is due to the measurement process. Comments on the physical validity of the DRF can be found in Refs 6, 17 and 26.

According to Zalm,[7] many analytical expressions can be chosen that fit equally well the result of the analysis, and care must be taken so that the parameters that govern the analytical model function have a satisfactory behaviour when varying the experimental conditions.[8] Actually, because a variety of input functions lead to the same output function when convolved with the response function with experimental error, the deconvolution process would result in any of these initial functions. This leads to the conclusion that implementing a DRF whose behaviour is not in agreement with the physical truth will lead to a result that cannot be considered to be physically valid.

When looking at a SIMS profile of a delta-doped layer of Si:B in silicon, we notice that it comprises an exponential trailing edge and a Gaussian-like rounded top. The rising part of the profile is often exponential too, although it is not always evident. Nevertheless, our DRF will have to take into account these features of the profiles, which is why we have chosen to implement an analytical DRF, initially proposed by Dowsett *et al.*,[1] which is constituted by the convolution of a double exponential with a Gaussian

$$D\exp(z) = A\begin{cases}\exp\left(\dfrac{z-z_0}{\lambda_{up}}\right) & z < z_0\\ \exp\left(\dfrac{-(z-z_0)}{\lambda_{down}}\right) & z > z_0\end{cases}$$

and

$$\text{Gauss}(z) = \frac{B}{\sqrt{2\pi}\,\sigma_{gauss}}\exp\left(\frac{-z^2}{2\sigma_{gauss}^2}\right)$$

The result of this convolution is given by the normalized expression

$$\begin{aligned}\text{Re } s(z) = \frac{1}{2(\lambda_{up}+\lambda_{down})}\Bigg\{&\exp\left(\frac{z-z_0}{\lambda_{up}}+\frac{\sigma^2}{2\lambda_{up}^2}\right)\\ &\times\left[1+\text{erf}\left(\frac{-(z-z_0)}{\sigma}-\frac{\sigma}{\lambda_{up}}\right)\frac{1}{\sqrt{2}}\right]\\ &+\exp\left(\frac{-(z-z_0)}{\lambda_{down}}+\frac{\sigma^2}{2\lambda_{down}^2}\right)\\ &\times\left[1+\text{erf}\left(\frac{z-z_0}{\sigma}-\frac{\sigma}{\lambda_{down}}\right)\frac{1}{\sqrt{2}}\right]\Bigg\}\end{aligned} \quad (1)$$

This DRF can be described by three parameters λ_{up}, λ_{down} and σ_{gauss}, and a fourth parameter Z_0 that represents the position of the cusp of the double exponential. We will give more precisions on this parameter later.

Experimental determinaton of the resolution functon

To have an analytical function at one's disposal allows the properties of this function to be characterized, thanks to its moments. The first-order moment represents the mean value of Res (z) and is equal to $\lambda_{down} - \lambda_{up}$. The second-order moment is

$$\sigma^2_{Res} = \sigma^2_{Gauss} + 2(\lambda^2_{down} + \lambda^2_{up} - \lambda_{up}\lambda_{down})$$

This expression will be very useful to characterize, in an objective manner, the depth resolution of the SIMS experiment. Published results for the depth resolution always mention the length of the exponential trailing or/and leading part of the profile without taking care of the flat top of the function that is not negligible when the energy of the primary beam is getting low. On the other hand, talking about full width at half-maximum (FWHM) does not take correctly into account the exponential part of the profile.

The experiments have been carried out with a Cameca IMS 3f/4f apparatus with O_2^+ primary ions, collecting the positive second ions. With this kind of instrument the energy ranges theoretically from 1.5 to 13 keV, from 3 to 13 keV in practice (because of the instrumental tuning being difficult, and the craters being less flat and more difficult to measure at very low energy), and the impact angle ranges from 35.5° to 64.3°. The fitting of the experimental curves is very good over two orders of magnitude for the leading edge, and on more than three for the trailing edge. Figure 1 shows the result of the fitting for two different energies, with a linear and a logarithmic intensity scale.[9]

We have changed the experimental conditions in order to study the behaviour of the fitting parameters. The results are summarized in Fig. 2. We have verified that the set of parameters that fit the resolution function for a given energy is unique.

The results are not expressed with respect to the primary ion energy but with respect to the primary ion range, R_p, given by the formula (derived from TRIM simulations using O^+ ions with the same impact angles as the O_2^+ ions and an energy equal to $E_p/2$)[5]

$$R_{p(Å)} = 50.46E_p^{0.665}\cos\theta \qquad (2)$$

where θ is the implicit angle of the beam and E_p is the primary energy per incident oxygen ion. This representation has been chosen in order to have synthetic results taking into account both the energy and the angle of

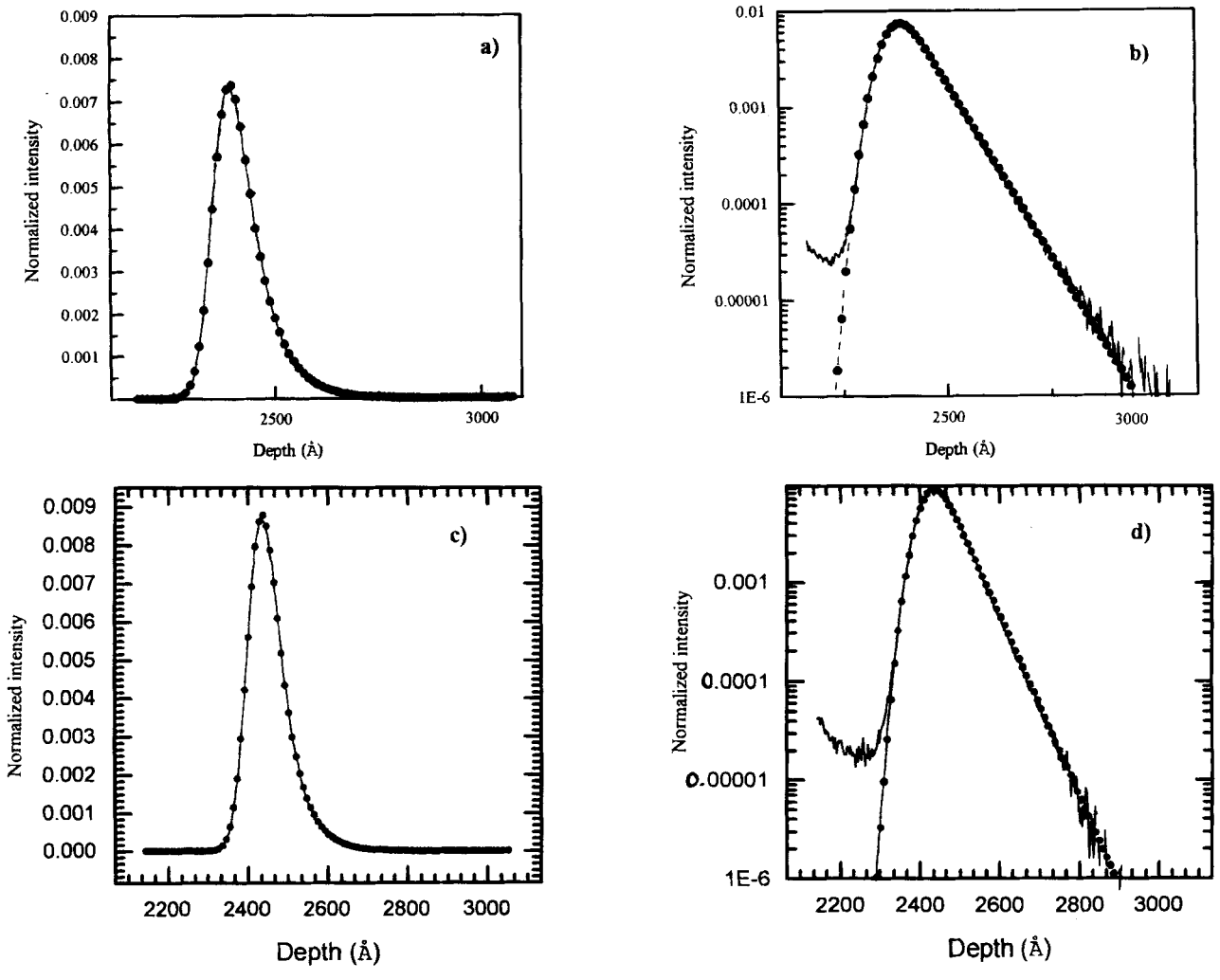


Figure 1. Fitting of the resolution function by the analytical expression: (a, b) 9.5 keV/O_2^+; (c, d) 6.5 keV/O_2^+ (normalized functions). Full lines: experimental curve; circles: fitting.

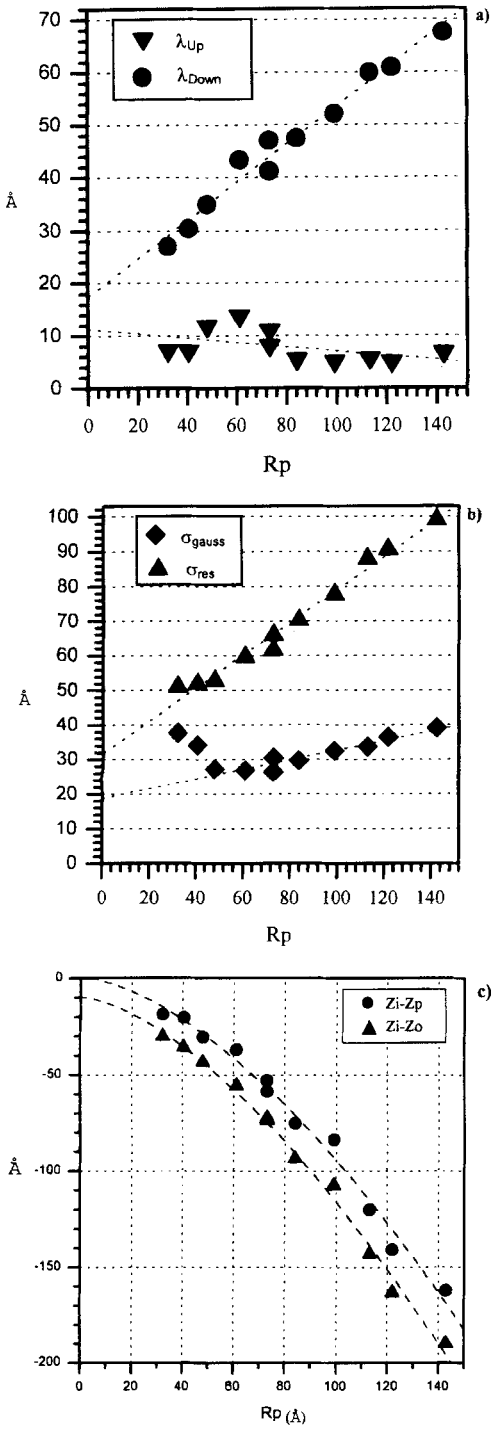


Figure 2. Evolution of the fitting parameters with the primary ion range: (a) evolution of the exponential coefficients λ_{up} and λ_{down}; (b) evolution of the Gaussian parameter σ_{gauss} and of σ_{res}; (c) evolution of the position Z_p of the apparent peak (maximum of experimental resolution function) and of the parameter Z_0 that represent the position of the double exponential, with respect to the real depth of the layer: Z_i. We have also represented (dashed lines) the fitting of the curves by a power law: (●) $Z_i - Z_p = 0.52 - 0.055 * x^{1.62}$; (●) $Z_i - Z_0 = -9.52 - 0.078 * x^{1.56}$

impact, which are not independent in a magnetic sector instrument.

The variation of the fitting parameters when changing the experimental conditions conforms with the behaviour of their physical homologues. We notice in particular the increase of λ_{down} with R_p, which has already been reported,[1,8,10–13] and the increase of σ_{gauss} which confirms the results from Dowsett.[1] However, when the impact energy is lowered below 3.5 keV/O_2^+, this parameter increases again. This may not be surprising considering the fact that the lower the energy, the poorer the focusing of the beam and the higher the impact angle. The increase of this parameter could be due to the fact that at very low energy, the craters can become asymmetric and inclined because of the variation of the incidence angle along the crater at grazing impact angles.

The rising exponential coefficient λ_{up} seems to decrease slightly with R_p. We suspect that this coefficient is linked to the instrumental conditions and may vary, among other parameters, with the shape of the beam.

The well-documented differential shift of the profile[13,14] is present in our results: the maximum of the peak (which we will call Z_p) moves slightly toward the surface when the energy of the primary ions increases, and so does the centroid of the DRF. This may be surprising, considering the fact that at an oblique angle and in the absence of oxygen bleed-in there should be no shift due to the formation of an SiO_2 layer.[27] We think that the shift we have measured is mostly due to the collisional mixing and to the anticipation of the response that becomes greater when the thickness of the altered layer is increased, i.e. when the energy is increased (the diffusion in the altered layer is supposed to be fast compared to the erosion of the sample). This phenomenon can be amplified by the diminution of the angle of incidence when the energy is increased, which could produce a slight shift due to a small variation of the initial erosion rate with the augmentation of the incorporation of oxygen.

This shift is not in disagreement with the work from Barlow *et al.*,[28] which measures a mean shift equal to ~70 Å at all energies and at a 45° angle of incidence. The variation of Barlow's shift is chaotic, and we notice from Barlow's figure 6f that the experimental point comparable to our experiments (4 keV/O_2^+, 45° incidence) shows a ~130 Å shift.

We notice from Fig. 2(c) that the parameter that we have called Z_0 moves significantly toward the surface too. This may not be surprising due to the fact that this parameter is linked to the centroid of the DRF by means of $\langle z \rangle$ and will move toward the surface if the centroid is moved because of the differential shift. One must not expect this parameter to give the exact position of the delta layer. In order to have an idea of what the true depth of the delta layer could be, an extrapolation of the position of the peak when R_p tends to zero might be a good solution, as already pointed out.[15] Figure 2(c) represents the difference between Z_p (or Z_0) and the real position of the delta layer, Z_i, which we have found to be equal to 2506 Å. A power law fit $Z_p = aR_p^n + b$, with $n \sim 1.5$, seems to be the best. Such a behaviour corresponds to a fairly good linear fit when the data are expressed with respect to E_p instead of R_p. Nevertheless, there is no physical reason why this parameter should change linearly with the primary beam energy.

We notice also that the parameters Z_p and Z_0 tend to very close values when R_p tends to zero, although they are very different at high energy. This behaviour seems

to be linked to the asymmetry of the curve and one can think that both parameters would be equal if the DRF was symmetric.

Figure 3(a) makes clear the difference between the parameter Z_0 and the real position of the delta layer by showing all the measured DRF. Figure 3(b) summarizes the different position parameters that we are dealing with.

In conclusion, the behaviour of the fitting parameters can be related to the experimental conditions. Just as it can be observed, for example, that the exponential decay of a SIMS profile gets slower when the primary ion energy is increased, so the parameter λ_{down} increases. Those parameters are not only mathematical entities, like the coefficients of a polynomial fitting would have been, but they correspond to independent aspects of the resolution function's shape and can be used to describe the depth resolution in an unambiguous manner. From this point of view, we can say that their behaviour is satisfactory, as they are directly correlated to the physical changes in the shape of the DRF when the experimental conditions are changed. Moreover, with this kind of fitting it is possible to extrapolate the shape of the resolution function by extrapolating the values of the fitting parameters, which would not have been possible with a purely mathematical fitting like a polynomial fitting.

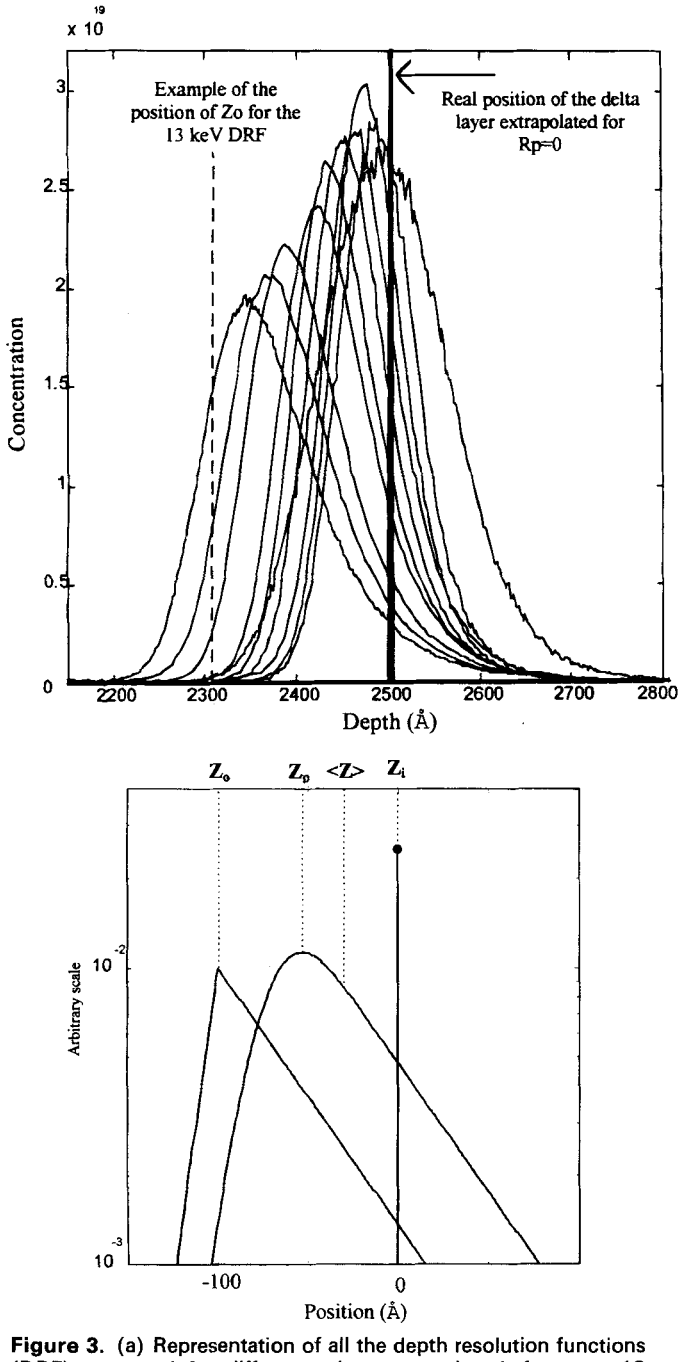


Figure 3. (a) Representation of all the depth resolution functions (DRF) measured for different primary energies. Left curve: 13 keV/O_2^+; right curve: 2.5 keV/O_2^+. The last two DRFs at very low energy are less high because of the increase of σ_{gauss}; (b) Position parameters Z_p, Z_i and Z_0. Z_p is the maximum of the DRF, Z_i the initial (real) position of the delta layer and Z_0, the position of the double exponential.

THE DECONVOLUTION PROCEDURE

The next step is the implementation of our analytical resolution in a deconvolution procedure. Many deconvolution schemes are available, which are usually separated into two groups: 'inverse' and 'forward' method[13] or 'direct' and 'indirect' method,[16] depending on whether they use a convolution procedure or a deconvolution procedure. Our method is close to what is commonly called a 'forward' method, because the deconvolution is not achieved through a division in the Fourier space. It is close to a maximum entropy method, and the main difference between the methods is the manner in which the problem of the deconvolution is regularized.

We have chosen this algorithm because it is very flexible and allows the constraints applied to the signal to be chosen and more constraints to be added when more information on the signal is available. Moreover, it allows a confidence level to be defined under which the deconvolution is no more reliable.

The algorithm

Discretizing the convolution equation results in a problem of solving a system of linear equations

$$y(m\Delta) = \sum_{k=0}^{N_h-1} h(k\Delta)x[(m-k)\Delta] + v(m\Delta)$$

$$m = 0, \ldots, N_y = N_x + N_h - 1 \quad (3)$$

where N_x, N_h and Δ are the length of the true profile, the length of the resolution function array and the depth step, respectively.

In this equation, the DRF is assumed to be causal, i.e. it corresponds to a delta layer located at $z = 0$. This does not correspond to the physical situation, and will shift the result of the profile. The correct positioning of the result of the deconvolution will be achieved *a posteriori* by means of extrapolation of the position of the layer when R_p tends to zero.

Equation (3) can then be rewritten as

$$y = Hx + v \quad (4)$$

where the vectors y, x and v represent the observed profile, the true profile and the noise, respectively. The operator H is a Toeplitz matrix, of size $N_x N_y$, constructed with the discretized resolution function. From Eqn (3) the length of y is greater than that of x. As a result, the system of linear Eqn (4) for $v = 0$ (noise free) is overdetermined. Thus, there will be only a least-squares solution. The noisy Eqn (4) has an infinite number of solutions. Then, an admissible solution from the physical meaning is the minimum normal least-squares solution[18]

$$x_{ls}(H^T H)^{-1} H^T y \tag{5}$$

where H^T denotes the transposed matrix.

However this solution is usually unacceptable because the normal matrix $H^T H$ is most often ill conditioned due to small singular values. Thus, a small error in the data results in a large error in the solution. It can be noted that the finer the discretization, the worse the conditioning of $H^T H$.

Considering a derivation of Eqn (4) where $x(n)$ is padded with zeros from N_x up to $N_y - 1$ and H is a circulant square Toeplitz matrix of size $N_y N_y$, the resulting vector $Hx + v$ is identical to the observed y. The singular values of the circulant matrix $H^T H$ are the modulus of the discrete Fourier transform (DRF) coefficients of the resolution function $h(m\Delta)$ padded with zeros from N_h up to $N_y - 1$. Figure 4(a) shows that some values of this spectrum vanish. This result clearly proves the ill conditioning of $H^T H$ in the problem of interest. Even in the case where an analytical form of the resolution function is used, the DRF falls very rapidly to very small values [Fig. 4(b)].

Regularization is a remedy to both the lack of stability and uniqueness of the problem. The basic feature is to introduce a compromise between fidelity to the data and fidelity to some prior knowledge about the solution. A regularized solution is defined as

$$x(y, \beta) = \text{argmin } \{J_1(y, Hx) + \beta J_2(x)\} \tag{6}$$

where $J_1(y, Hx)$, $J_2(x)$ and β are the measure of the distance between the convolved solution and the data (the observed profile), the measure of the closeness to some *a priori* information and the mixing factor, respectively. 'Argmin' denotes the argument that minimizes the expression in the brackets. Perfect fidelity to the data is achieved for $\beta = 0$, whereas perfect matching with *a priori* knowledge is achieved for $\beta = \infty$.

Equation (4) only defines a set of admissible solutions. *A priori* knowledge should be used in order to obtain an acceptable solution. In our problem we will assume that the unknown profile and the additive noise lie in the set of regular profiles with bounded roughness and in the set of bounded energy noise, respectively. In addition, the unknown profile lies both in the set of positive profiles and bounded support (finite extent) profiles. This idea is not new and has been addressed by a variety of mathematical methods.[18] For example, if the distances J_1 and J_2 are, respectively, the squared Euclidian distance and the Kullback distance,[18] then it results in the maximum entropy method of Gull and Skilling.[15,29] We will consider the quadratic approach[19,20] based on the estimation theory for unknown but bounded vectors.[21] It has an interesting physical meaning for the depth profile deconvolution

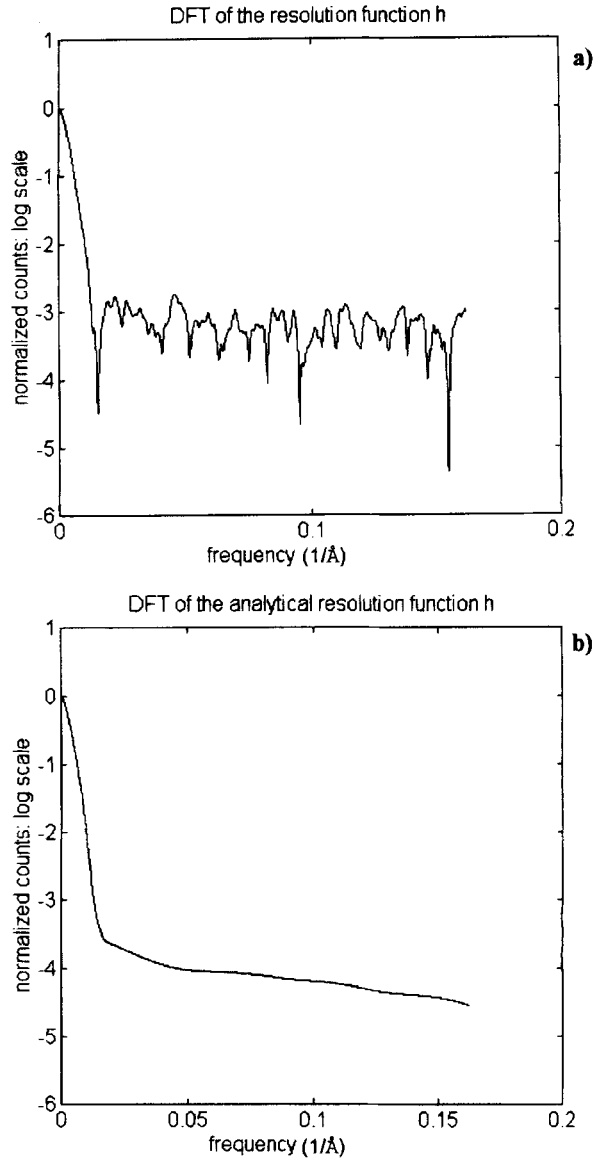


Figure 4. Direct Fourier transform (DFT) of the resolution function: (a) experimental; (b) analytical.

problem. The vector sets considered in our case are often ellipsoids. (Quadratic distance measure results in general in ellipsoidal sets [see Eqns (7), (9), (12) and (14)], however Ω_{est} in the following derivation is not an ellipsoid.) The equation for an ellipsoid is given by

$$\Omega_x = \{x : (x - x_c)^T \Gamma^{-1} (x - x_c) \leqslant 1\} \tag{7}$$

where x_c is the centre of the ellipsoid and Γ is a positive definite matrix whose eigenvectors and eigenvalues determine the orientation and the lengths of the semi-major axes of the ellipsoid.

The deconvolution problem may be stated as follows. The noise vector v lies in the set Ω_v of bounded energy noise

$$\|v\|^2 \leqslant n^2 \tag{8}$$

The set Ω_v is a sphere and its centre is the null vector

$$\Omega_v = \left\{ v : v^T \frac{I}{n^2} v \leqslant 1 \right\} \tag{9}$$

where $\boldsymbol{I}$ is the unity matrix.

The observed profile $\boldsymbol{y}$ specifies a set $\Omega_{x/y}$ which must contain $\boldsymbol{x}$

$$\boldsymbol{x} \in \Omega_{x/y} \quad \Omega_{x/y} = \{\boldsymbol{x} : \boldsymbol{y} - \boldsymbol{H}\boldsymbol{x} \in \Omega_v\} \tag{10}$$

From Eqn (4) we have

$$\|\boldsymbol{v}\|^2 = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 \tag{11}$$

From Eqns (8) and (11) the set $\Omega_{x/y}$ is an ellipsoid

$$\Omega_{x/y} = \left\{\boldsymbol{x} : (\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x})^{\mathrm{T}} \frac{\boldsymbol{I}}{n^2} (\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}) \leqslant 1\right\} \tag{12}$$

The solution $\boldsymbol{x}$ lies in the set Ω_x of the regular profiles with bounded roughness

$$\|\boldsymbol{D}\boldsymbol{x}\|^2 \leqslant r^2 \tag{13}$$

The Toeplitz matrix $\boldsymbol{D}$ is a differential operator or a high-pass filter with impulse response $d(n\Delta)$, $n = 0, \ldots, N_d - 1$. The scalar $\|\boldsymbol{D}\boldsymbol{x}\|^2$ increases with the roughness of the true profile. A bounded roughness means that the profile has a limited sharpness. However, Eqn (13) allows sharp discontinuities for high values of r^2. The choices of this upper bound for the sharpness should be consistent with the available data $\boldsymbol{y}$. Both the properties of D and the value of r^2 will be discussed in detail later.

The set Ω_x is an ellipsoid and its centre is the null vector

$$\Omega_x = \left\{\boldsymbol{x} : \boldsymbol{x}^{\mathrm{T}} \frac{\boldsymbol{D}^{\mathrm{T}}\boldsymbol{D}}{r^2} \boldsymbol{x} \leqslant 1\right\} \tag{14}$$

Consider both the set Ω_x and $\Omega_{x/y}$. Each set contains $\boldsymbol{x}$. Thus $\boldsymbol{x}$ must lie in their intersection $\Omega_{\mathrm{est}} = \Omega_x \cap \Omega_{x/y}$. The intersection is the smallest set which must contain $\boldsymbol{x}$ and which can be calculated from the available data. This set is the best estimate set.[21] In contrast with usual estimation theory, the estimate is a set of vectors and not a single vector. In order to solve the problem of lack of uniqueness, a reasonable choice for a vector estimate $\boldsymbol{x}_{\mathrm{est}}$ is to define $\boldsymbol{x}_{\mathrm{est}}$ at the centre of Ω_{est}. Unfortunately Ω_{est} is not an ellipsoid. As a result, the centre cannot be easily computed. However, a bounding ellipsoid can be found which contains the intersection Ω_{est}. This bounding ellipsoid is defined by

$$\Omega_{\mathrm{b}} = \left\{\boldsymbol{x} : (1-\rho)\boldsymbol{x}^{\mathrm{T}} \frac{\boldsymbol{D}^{\mathrm{T}}\boldsymbol{D}}{r^2} \boldsymbol{x} + \rho(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x})^{\mathrm{T}} \frac{\boldsymbol{I}}{n^2} (\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}) \leqslant 1\right\} \quad 0 \leqslant \rho \leqslant 1 \tag{15a}$$

or

$$\Omega_{\mathrm{b}} = \{\boldsymbol{x} : (\boldsymbol{x} - \boldsymbol{x}_{\mathrm{b}})^{\mathrm{T}} \Gamma_{\mathrm{b}}^{-1} (\boldsymbol{x} - \boldsymbol{x}_{\mathrm{b}}) \leqslant 1\} \tag{15b}$$

The centre $\boldsymbol{x}_{\mathrm{b}}$ of Ω_{b} is the solution of the normal equation

$$\boldsymbol{A}\boldsymbol{x}_{\mathrm{b}} = \boldsymbol{z} \tag{16}$$

where

$$\boldsymbol{A} = \rho\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H} + (1-\rho)\alpha\boldsymbol{D}^{\mathrm{T}}\boldsymbol{D} \tag{17}$$

$$\boldsymbol{z} = \rho\boldsymbol{H}^{\mathrm{T}}\boldsymbol{y} \tag{18}$$

with

$$\alpha = n^2/r^2 \tag{19}$$

The matrix Γ_{b} is defined by

$$\Gamma_{\mathrm{b}} = (1 - \delta^2) n^2 \boldsymbol{A}^{-1} \tag{20}$$

where[20]

$$\delta^2 = \frac{1}{n^2} \rho(\|\boldsymbol{y}\|^2 - \boldsymbol{y}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{x}_{\mathrm{b}}) \tag{21}$$

There is a whole family of bounding ellipsoid estimate sets Ω_{b} depending on the choice of ρ. Setting $\rho = 1$, then $\boldsymbol{x}_{\mathrm{b}}$ becomes the minimum normal least-squares solution of Eqn (4), defined by Eqn (5). Setting $\rho = 0.5$, then $\boldsymbol{x}_{\mathrm{b}}$ becomes the well-known Miller[22] solution.

Clearly, if $1 - \delta^2$ is negative the matrix Γ_{b} becomes definitely negative and the intersection Ω_{est} is empty. It is proved[19] that $\|\boldsymbol{y}\|^2 - \boldsymbol{y}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{x}_{\mathrm{b}}$ is positive and is the minimum of the functional in Eqn (15a). The minimum occurs for $\boldsymbol{x} = \boldsymbol{x}_{\mathrm{b}}$. The non-emptiness of Ω_{est} can be verified *a posteriori*. It can be noted that if the computation of the estimate $\boldsymbol{x}_{\mathrm{b}}$ requires only the knowledge of $\alpha = n^2/r^2$, the non-emptiness requires the knowledge of n^2 in Eqns (20) and (21).

Denoting $\lambda^2_{k,H}$ and $\lambda^2_{k,D}$ the singular values of $\boldsymbol{H}$ and $\boldsymbol{D}$, respectively, then the condition number for $\boldsymbol{A}$ is

$$C_A = \frac{[\rho\lambda^2_{k,H} + (1-\rho)\alpha\lambda^2_{k,D}]_{\max}}{[\rho\lambda^2_{k,H} + (1-\rho)\alpha\lambda^2_{k,D}]_{\min}} \tag{22}$$

In order that the matrix $\boldsymbol{A}$ is better conditioned than $\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}$, the operator $\boldsymbol{D}$ should be chosen with large singular values (or small singular values) when singular values of $\boldsymbol{H}$ are small (or large). As the operator $\boldsymbol{A}$ is a low-pass filter, the operator $\boldsymbol{D}$ should be a high-pass filter.[19] Figure 5 shows the frequency response of the filter used in the experiments. The deconvolution results are robust in terms of the choice of the high-pass operator $\boldsymbol{D}$.

According to Eqn (15b) the larger length of the semi-major axes of the ellipsoid of the estimate set Ω_{b} is

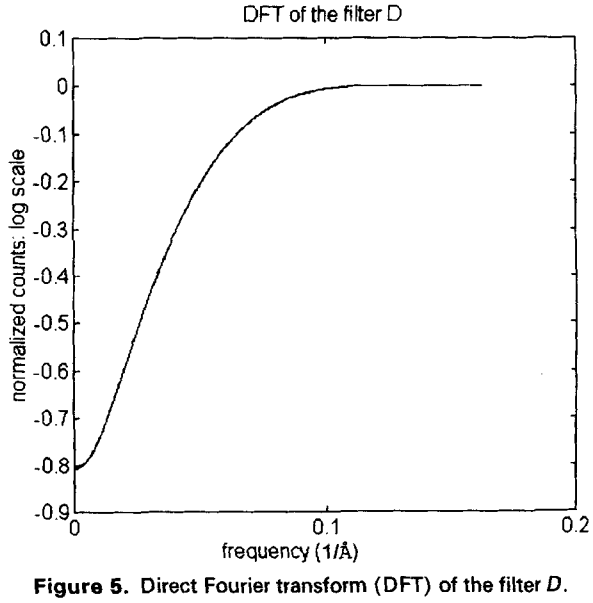


Figure 5. Direct Fourier transform (DFT) of the filter D.

defined by the square root of the largest eigenvalue of the matrix Γ_b, denoted $(\lambda^2_{k,\Gamma_b})_{max}$. Then, from Eqns (20) and (17) we have

$$(\lambda^2_{k,\Gamma_b})_{max} = \frac{(1-\delta^2)n^2}{[\rho\lambda^2_{k,H} + (1-\rho)\alpha\lambda^2_{k,D}]_{max}} \quad (23)$$

This original result allows the estimation of the size of the bounding ellipsoid Ω_b. It can be noted that if the minimum value of $\lambda^2_{K,H}$ is very small, the size of the set Ω_b is only reduced by $\alpha\lambda^2_{K,D}$. We use Eqn (23) for the definition of an estimate for the upper bound of the sum of squares of the errors on the solution x_b.

Let x_a be a vector at the intersection of the surface of the ellipsoid Ω_b with its larger major axis. Then the sum of squares of the errors between the two admissible solutions x_a and x_b is defined by

$$\|x_a - x_b\|^2 = (\lambda^2_{k,\Gamma_b})_{max} \quad (24)$$

This is the upper bound of the error for the solution x_b. Clearly, from Eqn (23) no error occurs when the data are noise free. The squares of the errors increase both with the noise-bound n^2 and with the smallest eigenvalue of the matrix A. As the local error bound cannot be known, we consider that the error is uniformly distributed on the N_x elements of x_b. Then, according to this hypothesis, the upper bound of the error on each element is

$$\varepsilon = \frac{\sqrt{(\lambda^2_{k,\Gamma_b})_{max}}}{N_x} \quad (25)$$

From Eqn (11) a consistent estimate for the noise-bound n^2 is the square of the residual

$$n^2 \# \|y - Hx_b\|^2 \quad (26)$$

We use this estimate both in the evaluation of δ^2 in Eqn (21) for checking that the intersection Ω_{est} is not empty, and in Eqn (23). It will also be useful to define a confidence level below which the restoration cannot be considered as reliable.

The regularization parameter α can be estimated by a variety of techniques.[23,24] We used the generalized cross-validation because it applies for Gaussian white noise and, in addition, is both fast and easily implemented in the DFT domain using the Fast Fourier Transform (FFT) algorithm. In order to prove the whiteness of the noise, we have calculated the power spectrum of the signal at the end of the profile according to the well-known Welch estimator,[30] in a region where the boron level was constant but not too low. In addition we have considered the generalized Gaussian probability distribution function (pdf) as a candidate for the model. We have estimated the shape parameter according to the estimator proposed in Ref. 25. A shape parameter of 1 corresponds to a Laplacian pdf. In contrast, a shape parameter of 2 corresponds to a Gaussian distribution. The experimental result was 2.06, which proves the correctness of the Gaussian model for the noise. In fact, this assumption on the Gaussian nature of the noise has already been used by Allen *et al.*[15] At the bottom of the profile, where the boron concentration is very low, we have verified that the noise was Poissonian, but its level is very low (a few counts per minute). At high count rates, i.e. when a boron profile is being measured, this Poissonian noise is negligible with respect to the total noise, probably constituted by several contributions.

The solution of the normal Eqn (16) is not positive. Figure 6 illustrates this solution for the circulant matrix H constructed with the experimental profile. The Miller deconvolution is implemented in the DFT domain, according to Eqns (16)–(18) as follows

$$x(n\Delta) = \mathrm{DFT}^{-1}\left\{\frac{\rho\mathrm{DFT}[h(n\Delta)]^*\mathrm{DFT}[y(n\Delta)]}{\rho\mathrm{DFT}[h(n\Delta)]^*\mathrm{DFT}[h(n\Delta)] + (1-\rho)\alpha\mathrm{DFT}[d(n\Delta)]^*\mathrm{DFT}[d(n\Delta)]}\right\} \quad (27)$$

where $\mathrm{DFT}[\cdot]^*$ is the conjugate of $\mathrm{DFT}[\cdot]$; $y(n\Delta)$, $h(n\Delta)$ and $d(n\Delta)$ are padded with zeros from $N_y - 1$ to N, $N_h - 1$ to N and $N_d - 1$ to N, respectively. The number of samples N should be greater or equal to Ny.

In order to incorporate the positivity constraint, we solve the normal equations using a constrained iterative procedure of the form[3,20]

$$x^{(n)} = C_x x^{(n-1)} + \mu^{(n)} V^{(n)} \quad (28)$$

where C_x is a non-linear projection operator onto the convex set of positive profiles Ω_p.[31] We consider that the intersection $\Omega_p \cap \Omega_b$ is not empty.

The operator C_x replaces the negative elements of the vector $x^{(n-1)}$ by zeros and leaves the positive ones. This operator is non-expansive, i.e.

$$\|C_x f - C_x g\| \leqslant \|f - g\|, \quad f, g \in \mathbb{R} \quad (29)$$

The distance between two constrained signals never exceeds the distance between two original signals.

Both the relaxation parameter $\mu^{(n)}$ and the vector $V^{(n)}$ are computed in order to minimize the following posi-

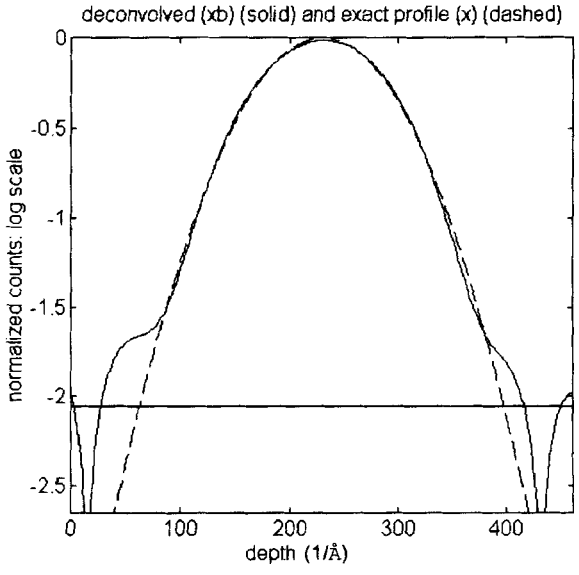


Figure 6. Non-iterative deconvolution of a theoretical Gaussian ($\sigma = 50$ Å). The signal is shown with the (better) dynamic of the iterative deconvolution. The horizontal line represents the upper bound of the error reached by the non-iterative deconvolution, compared to the error reached by the iterative deconvolution (−2.6). Peaks that appear at the right and at the left side of the Gaussian are artifacts caused by the folding up of the negative component of the solution.

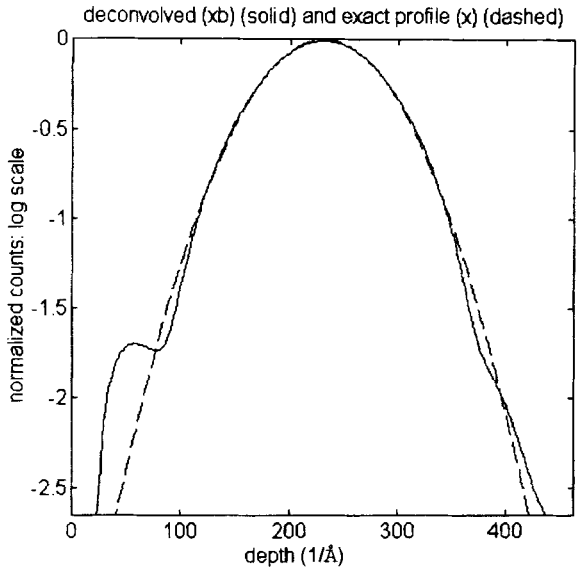


Figure 7. Iterative deconvolution of a theoretical Gaussian ($\sigma = 50$ Å, signal/noise ratio = 25 dB).

tive functional of the error at each step

$$J(n) = \varepsilon^{(n)T} A \varepsilon^{(n)} \tag{30}$$

where the error is defined by

$$\varepsilon^{(n)} = x_b - x^{(n)} \tag{31}$$

Then

$$\mu^{(n)} = V^{(n)T}[z - AC_x x^{(n-1)}]/[V^{(n)T} A V^{(n)}] \tag{32}$$

$$J^{(n)} = J_c^{(n-1)} - \{V^{(n)T}[z - AC_x x^{(n-1)}]\}^2/[V^{(n)T} A V^{(n)}] \tag{33}$$

where

$$J_c^{(n-1)} = \varepsilon_c^{(n-1)T} A \varepsilon_c^{(n-1)} \tag{34}$$

$$\varepsilon_c^{(n-1)} = x_b - C_x x^{(n-1)} \tag{35}$$

From Eqn (33) a sufficient condition for convergence at each step, e.g. $J^{(n)} \leqslant J^{(n-1)}$ is $J_c^{(n-1)} \leqslant J^{(n-1)}$. This

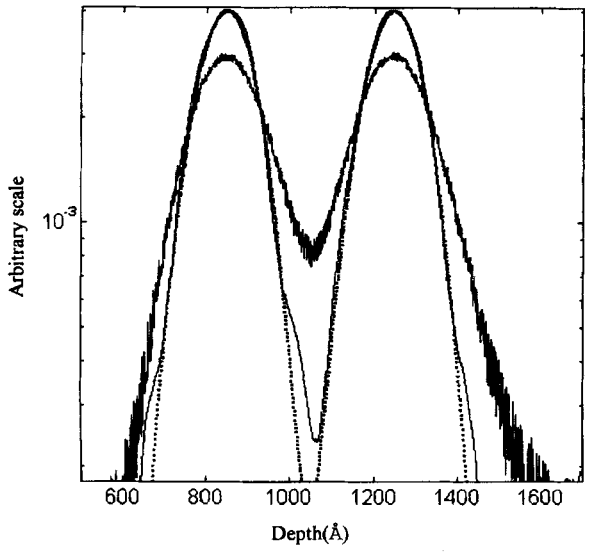
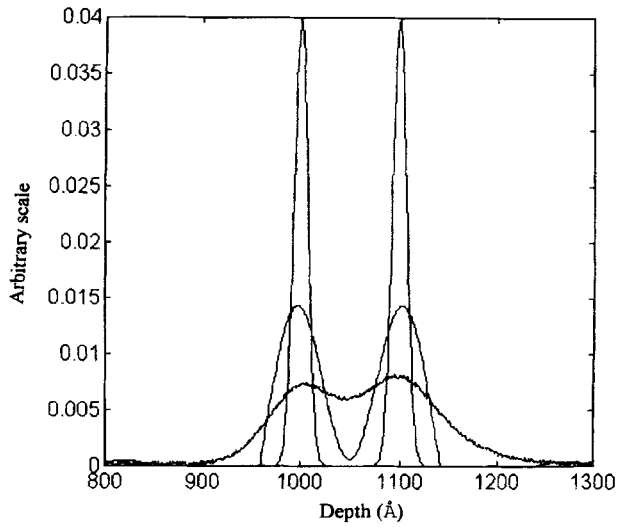


Figure 8. Result of the deconvolution of two Gaussians ($\sigma = 70$ Å, $d = 250$ Å) blurred by the 13 keV/O_2^+ DRF. Noisy line: measured profile, dotted line: initial profile, thin solid line: deconvolved profile.



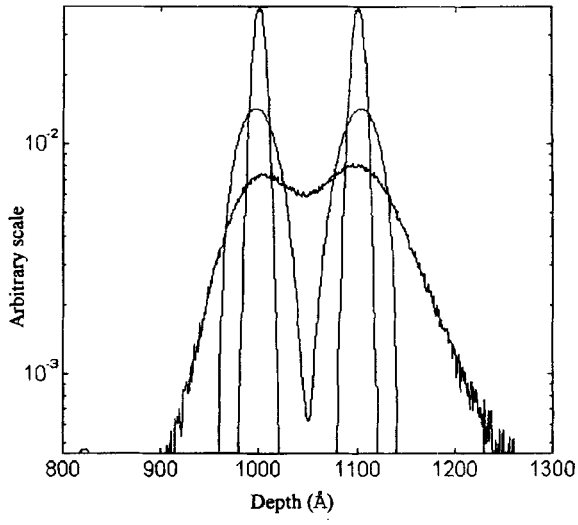


Figure 9. Deconvolution of a theoretical Gaussian profile with $\sigma = 10$ Å and $d = 100$ Å shown on a linear and a semilog scale. The sharpest profile is the initial theoretical input. The convolution of the initial profile with the resolution function (profile to be deconvolved) is also shown for comparison (noisy signal). The intensities on the y-axes of Figs 8–11 are all arbitrary intensities.

condition is fulfilled if C_x is non-expansive. According to Eqn (33) the rate of convergence is maximized by taking

$$V^{(n)} = z - AC_x x^{(n-1)} \tag{36}$$

From Eqn (28) the elements of the vector $x^{(n)}$ are not necessarily positive at each step but converge to positive values.

The iterative procedure is stopped when the residual $\|z - AC_x x^{(n)}\|^2$ becomes almost constant. Then $C_x x^{(n)}$ is on the surface of Ω_p at the point of minimum distance from the centre x_b of Ω_b.

It can be noted that if we choose a constant relaxation parameter $\mu^{(n)}$, this parameter should satisfy

$$0 < \mu^{(n)} < \frac{2}{[\rho\lambda_{k,H}^2 + (1-\rho)\alpha\lambda_{k,D}^2]_{\max}} \tag{37}$$

This results in a slower convergence rate.

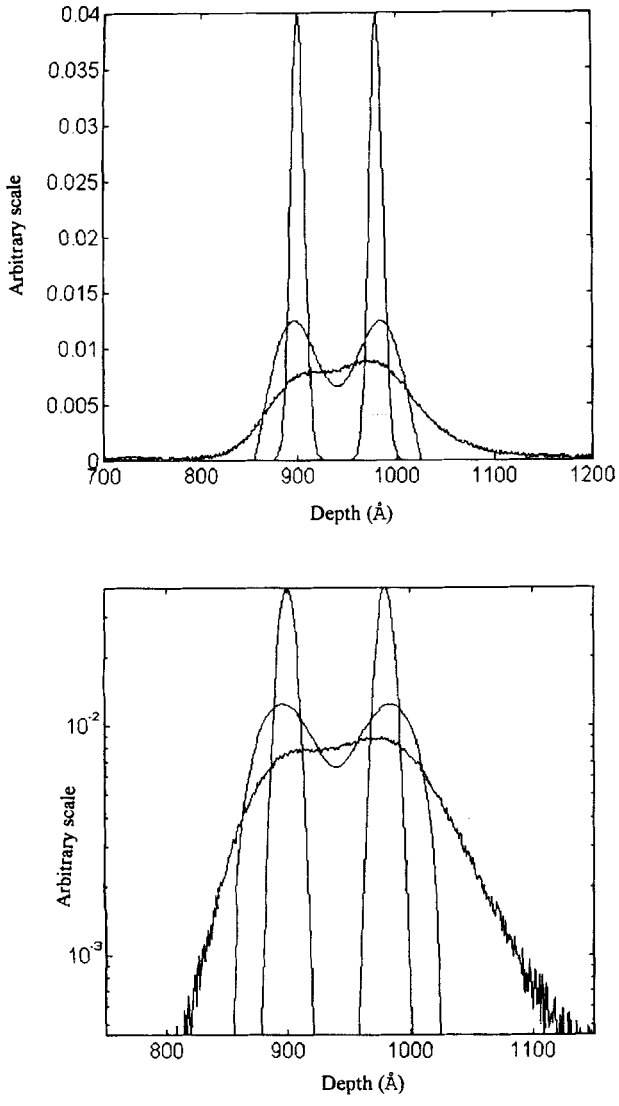


Figure 10. Deconvolution of a theoretical Gaussian profile with $\sigma = 10$ Å and $d = 80$ Å shown on a linear and a semilog scale. Signal-to-noise ratio = 40 (16 dB).

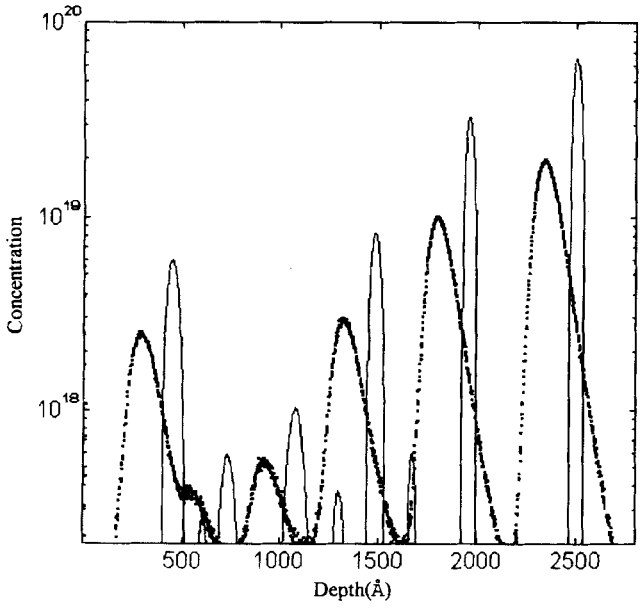


Figure 11. Deconvolution of a 13 keV/O_2^+ profile taking the differential shift into account: straight line: deconvolution, dotted line: initial profile.

Simulated deconvolutions

Deconvolution of a single Gaussian. Before we implement our algorithm to some real SIMS profile, it is interesting to study the theoretical behaviour of the deconvolution. Thus, we have constructed theoretical profiles (noisy or noise-free) that we have convolved with our analytical DRF. Figures 6 and 7 show the result of this simulation for a single Gaussian profile with a 25 dB noise. The normal Eqn (16) is first solved in the DFT domain [see Eqn (27)] [Fig. 6(a)]. The resulting solution is not positive, creating more artifacts. We have verified that the intersection of the ellipsoids $\Omega_x \cap \Omega_{x/y}$ was not empty ($\delta^2 < 1$). The upper bound of the error [Eqn (25)] corresponds to the horizontal line.

The iterative algorithm allows a positive solution to be obtained with a lower upper bound for the error (see the lower bound of the logarithmic y-axis in Fig. 7).

Test of separability: deconvolution of two Gaussians. We have tested the deconvolution scheme on another Gaussian profile, that is two Gaussian functions with a standard deviation σ separated by a distance that we will call d. We have also added a white noise to the perfect analytical convolution of the initial profile with the DRF (both analytical), in order to have a signal-to-noise ratio of 40 (in energy ratio, or 16 dB), which is rather pessimistic compared to a real SIMS experiment.

The first example is the deconvolution of two Gaussians with $\sigma = 70$ Å and $d = 250$ Å, convolved with the 13 keV resolution (that is the 'worst', in terms of convolution, of our set of measurement: $\lambda_{up} = 6.7$ Å, $\lambda_{down} = 67.5$ Å, $\sigma_{gauss} = 39.2$ Å, $\sigma_{total} = 91.1$ Å). The parameters of the DRF were taken from the results of the analytical fitting. The results of the deconvolution is shown in Fig. 8. We can see that the initial profile is retrieved in a very satisfactory way, although the width of the resolution function is greater than the standard deviation of the Gaussian profiles. Note that the intensities on the y-axis of Figs 8–10 are all arbitrary intensities.

The second example of deconvolution involves two Gaussian profiles with a standard deviation $\sigma = 10$ Å separated by 100 Å (Fig. 9). In this example the simulated profile was convolved with the 3.5 keV primary beam energy DRF ($\lambda_{up} = 11.8$ Å, $\lambda_{down} = 35.0$ Å, $\sigma_{gauss} = 27.2$ Å, $\sigma_{total} = 52.8$ Å). This DRF can be considered as the 'best' in terms of experimental conditions: it appears like a good compromise between collisional mixing and reliability, precision or easiness of the apparatus tuning. It is clear that the Gaussian profile that was severely blurred by the convolution can be retrieved in a very satisfactory way. Figure 10 shows the same exercise with $\sigma = 10$ Å and $d = 80$ Å: this time, the two Gaussians are no more visible when convolved with the DRF. The deconvolution process manages to retrieve partially the original features and shows that two Gaussians are present.

EXPERIMENTAL DECONVOLUTIONS

Deconvolution of six consecutive delta layers

The time has come to implement our deconvolution procedure in the case of real SIMS experiments. This

first sample that we will try to deconvolve is an RTCVD-grown sample containing six deltas of Si:B in silicon. This experimental exercise is a kind of test for our deconvolution procedure. Actually, the DRF used in the deconvolution algorithm is taken from the sixth delta layer of the sample we deconvolve. The result should be a perfect delta if the deconvolution experiment was perfect. Of course, this is not the case and the deconvolution algorithm just leads to an improvement of the measured signal but is not supposed to give the exact original feature.

The result of the deconvolution is given in Fig. 11 (this time the y-axis represents the concentration). This figure takes into account the differential shift: we have positioned the result of the deconvolution by considering that the real position of the delta layer was the extrapolation of Z_p when R_p tends to zero.

Characterization of the deconvolution results. The FWHM of the sixth delta is $\sim$41 Å (41.6 $\pm$ 2 Å at 6.5 keV/O_2^+) and is remarkably constant over the studied range of primary beam energies. This can be considered as a very good result if one takes into account that the original FWHM of the profile is approximately 100 Å at the same energy. The improvement of the measured signal is far better than in the case of the delta layer recovered by the inverse FFT algorithm of Herzel *et al.*[2]

From Fig. 11 it can be seen that the deconvolved delta layers seem to get narrower with depth. This is not surprising, due to the fact that the DRF was taken from the last delta layer, which is supposed to be the best delta layer and where the signal is the highest. Actually, in the case of RTCVD-grown layers, the first deposited layer is a good delta.[5] The quality of the layers deposited later is not guaranteed to be as good as the first. Thus, the result of the deconvolution of the other delta layers is slightly wider and can be related to a physical width and a poorer deconvolution because of their lower intensity.

It is worth pointing out that the deconvolution scheme is responsible for the appearance of artifacts that seem to be more numerous as the primary beam energy is lowered.

It is clear that those artifacts arise from the noise and that a deconvolved peak that is under the level of the noise can be considered as physically meaningless.

It is also important to know whether it is better to deconvolve the SIMS profile rather than perform a very good instrumental analysis, which will require a lot of time or require very expensive instrument, and which is not adapted to a context where the SIMS apparatus is used by many operators who need to make different analyses on different samples with different experimental conditions. This conception of the deconvolution does not mean that deconvolution is a substitute for careful analysis or that a proper analysis cannot be conducted: it is a surrogate way of reaching good depth resolution when time or adapted instruments are not available. Moreover, deconvolution will apply when better instrumental profiles are available, leading to even better depth resolution.

We tried to answer this question by comparing the result of the deconvolution of a 13 keV analysis with the raw data obtained by a careful 3.5 keV analysis. The result can be seen in Fig. 12. It is obvious that the deconvolution procedure has strongly improved the slope and the regularity of the delta layer, even compared to the lowest energy experiment.

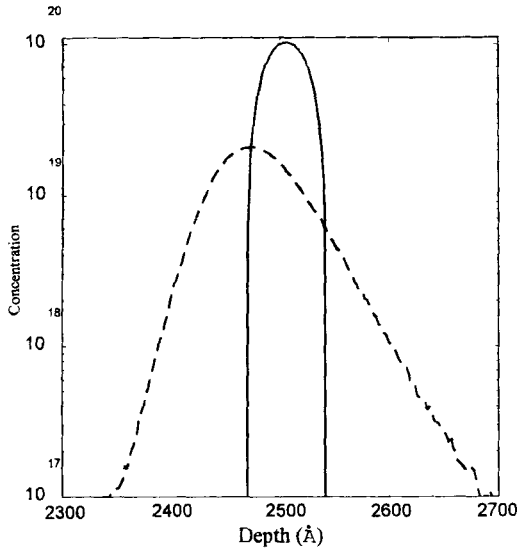


Figure 12. Comparison between a deconvolved 13 KeV/O_2^+ signal (smooth line) with a raw 3.5 keV/O_2^+ data (dashed line).

Deconvolution of the 75 Å thick layer

Structure of the 75 Å layer. This layer was grown under the same conditions as the delta layers, leading to very well-defined interfaces. We have calculated the analytical expression of the convolution of our analytical DRF with an analytical thick layer (thickness $2z_1$) and we have fitted this expression to the experimental data of the analysis of our real layer. The analytical result of the convolution is

$$\begin{aligned}
\mathrm{Conv}(z) = {} & \frac{1}{4z_1(\lambda_{up}+\lambda_{down})} \\
& \times \left[(\lambda_{up}+\lambda_{down}) \left[\mathrm{erf}\left(\frac{z+z_1}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{z-z_1}{\sqrt{2}\sigma}\right) \right] \right. \\
& + \lambda_{up} \exp\left(\frac{z}{\lambda_{up}} + \frac{\sigma^2}{2\lambda_{up}^2}\right) \\
& \times \left\{ \exp\left(\frac{z_1}{\lambda_{up}}\right) \left[1 + \mathrm{erf}\left(\frac{-z}{\sqrt{2}\sigma} - \frac{z_1}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\lambda_{up}}\right) \right] \right. \\
& \left. \left. - \exp\left(\frac{-z_1}{\lambda_{up}}\right) \left[1 + \mathrm{erf}\left(\frac{-z}{\sqrt{2}\sigma} + \frac{z_1}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\lambda_{up}}\right) \right] \right\} \right. \\
& + \lambda_{down} \exp\left(\frac{-z}{\lambda_{down}} + \frac{\sigma^2}{2\lambda_{down}^2}\right) \\
& \times \left\{ \exp\left(\frac{z_1}{\lambda_{down}}\right) \left[1 + \mathrm{erf}\left(\frac{z}{\sqrt{2}\sigma} - \frac{z_1}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\lambda_{down}}\right) \right] \right. \\
& \left. \left. \left. - \exp\left(\frac{-z_1}{\lambda^{down}}\right) \left[1 + \mathrm{erf}\left(\frac{z}{\sqrt{2}\sigma} + \frac{z_1}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\lambda_{down}}\right) \right] \right\} \right]
\end{aligned}$$

The fitting is as good as in the case of the DRF (Fig. 13 shows the fitting of the analysis of the thick layer) for all the range of primary energies we studied.

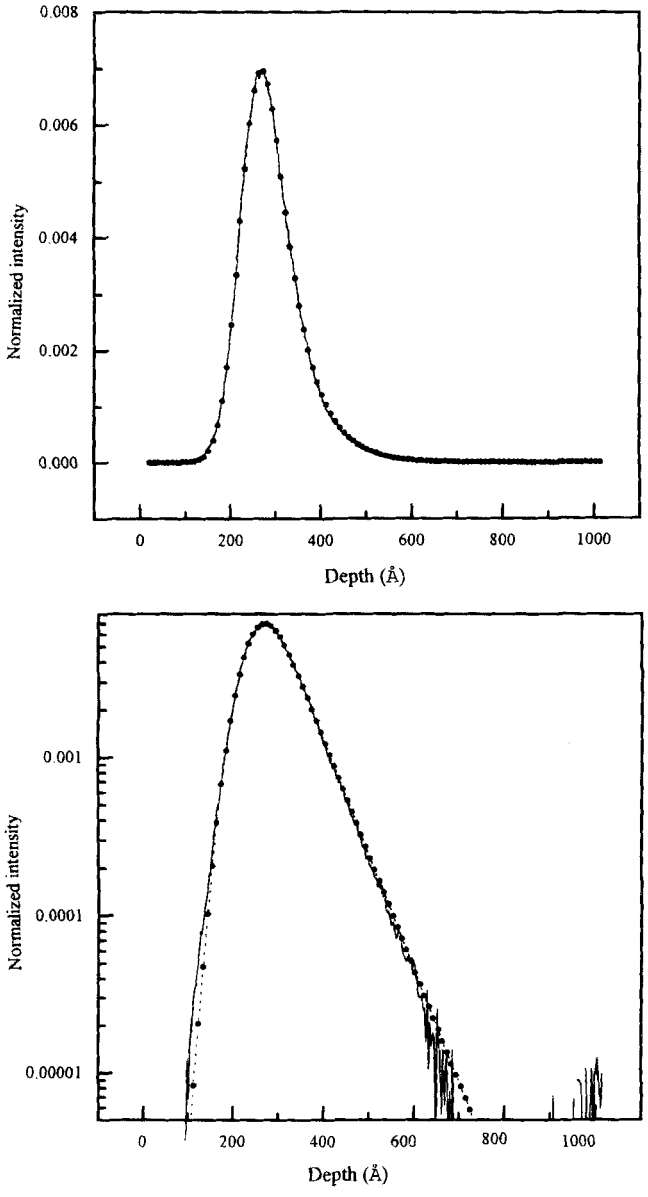


Figure 13. Fitting of a 75 Å thick layer by an analytical expression for 9.5 keV/O_2^+, linear and semilog scale. (● Fitting; (-----) Experimental data.

As a preliminary exercise before we attempt to deconvolve this thicker profile, we have tried to find out the thickness of the layer by fitting the experimental data for a given primary ion energy using the parameters λ_{up}, λ_{down} and σ taken from the analysis of the delta layer analysed in exactly the same conditions. The only parameters that were allowed to vary were the thickness and the position of the layer.

The result is extremely good for energies ranging from 5.5 to 13 keV/O_2^+: the calculated thickness is close to the nominal thickness of 75 Å.[9] We have verified that for given λ_{up}, λ_{down} and σ_{gauss} there was only one possible fitting of the experimental data and that there were single values for the thickness and the position minimizing the quadratic distance between the fitted and the experimental curve.

The results obtained lead us to think that this layer can be considered as a correct thick layer because it can be described by our analytical treatment. It will be used to test our deconvolution procedure.

Deconvolution of the 75 Å layer. The result of the deconvolution for a 6.5 keV/O_2^+ primary energy is given by Fig. 14. The depth resolution has been greatly improved by the deconvolution process, as can be seen in Fig. 14(a), where a deconvolved profile is compared to the raw data. Moreover, the deconvolution algorithm has removed the asymmetry of the profile and has led to a Gaussian-like profile, without any exponential behaviour. The result of the deconvolution for energies higher than 6.5 keV/O_2^+ (up to 13 keV) is almost exactly the same.

In order to compare these results with those obtained in the previous section, we have superimposed on Fig. 13(b) the result shown in Fig. 14(a) and an artificial rectangular profile, which has the same dose as the deconvolved profile and a thickness of 75 Å (which is the nominal thickness; this result we found in the previous section). From Fig. 14 we can conclude that the result of the deconvolution is very consistent with the results of

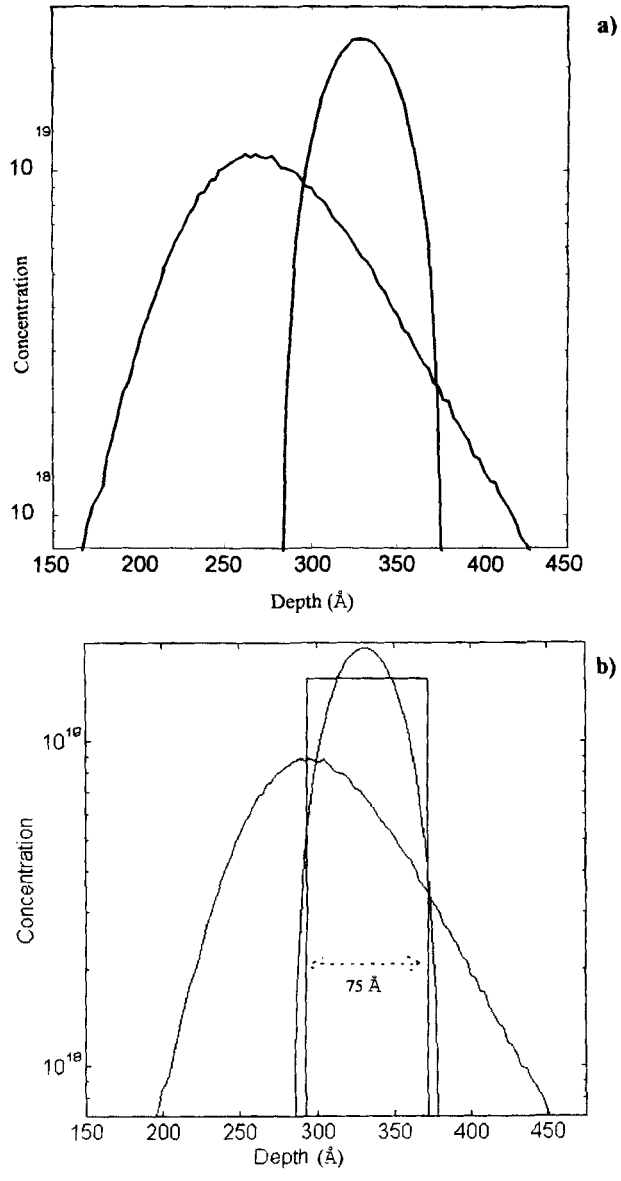


Figure 14. Deconvolution of the thicker layer (thickness 75 Å, approximately located between 300 and 375 Å): (a) comparison with the raw data for a 6.5 keV/O_2^+ primary beam; (b) comparison of a rectangular profile with the deconvolved profile and the measured profile (6.5 keV/O_2^+).

the fitting (where a perfect rectangular profile is assumed), although the result of the deconvolution cannot lead to the restoration of a perfect rectangular profile.

CONCLUSION

We have implemented a deconvolution procedure for SIMS depth profiling, using an analytical form of the DRF. The deconvolution scheme takes into account some of the properties (positivity, smoothness) of the signal in order to obtain a physically acceptable result. The deconvolution procedure has been conducted with an analytical DRF, the behaviour of whose parameters is in agreement with physical reality.

We have shown that a great improvement of the signal can be performed in a short time (deconvolution times are of the order of a few minutes).

Nevertheless, the deconvolution procedures have limits that one must keep in mind in order to perform an improvement of the signal that is not artificial. Those limits, including the use of the noise estimate defined in Eqn (25) as a confidence level, and the ultimate performances that can be expected from the algorithm will be the subject of another paper, in preparation. A comparative study with the maximum entropy algorithm would also be judicious.

Acknowledgements

Thanks go to the French CNET for supporting this work and to P. Warren and D. Dutartre for delta layer elaboration.

REFERENCES

1. M. G. Dowsett, G. Rowlands, P. N. Allen and D. Barlow, *Surf. Interface Anal.* **21**, 310 (1994).
2. F. Herzel, K. H. Ehwald, B. Heinemann, D. Krüger, R. Kurps, W. Röpke and H. P. Zeindl, *Surf. Interface Anal.* **23**, 764 (1995).
3. R. Prost and R. Goutte, *Signal Process.* **7**, 209 (1984).
4. J. B. Clegg and I. G. Gale, *Surf. Interface Anal.* **17**, 190 (1991).
5. J. C. Dupuy, G. Prudon, C. Dubois, P. Warren and D. Dutartre, *Nucl. Instrum. Methods, Phys. Res.* **B85**, 379 (1994).
6. M. G. Dowsett, *Surf. Interface Anal.* **23**, 900 (1995).
7. P. C. Zalm, in *Proc. SIMS International Conference, SIMS X*, Münster (1995).
8. M. G. Dowsett and D. Barlow, *Anal. Chim. Acta* **297**, 253 (1994).
9. B. Gautier, J. C. Dupuy, J. P. Vallard, G. Prudon and C. Dubois, in *Proc. SIMS International Conference, SIMS X*, Münster (1995).
10. M. Meuris and W. Vandervorst, *J. Vac. Sci. Technol.* **A9**(3), 1482 (1991).
11. M. G. Dowsett, D. Barlow, H. S. Fox, R. A. A. Kubiak and R. Collins, *J. Vac. Sci. Technol.* **B10**(1), 336 (1992).
12. K. Wittmaack, *Surf. Interface Anal.* **21**, 323 (1994).
13. M. G. Dowsett, D. Barlow and P. N. Allen, *J. Vac. Sci. Technol.* **B12**(1), 186 (1994).
14. K. Wittmaack and W. Wach, *Nucl. Instrum. Methods* **191**, 327 (1981).
15. P. N. Allen, M. G. Dowsett and R. Collins, *Surf. Interface Anal.* **20**, 696 (1993).
16. R. Collins, M. G. Dowsett and P. N. Allen, in *Proc. SIMS International Conference, SIMS VIII*, Amsterdam (1991).
17. V. V. Makarov, *Surf. Interface Anal.* **20**, 821 (1993).
18. G. Demoment, *IEEE Trans. Acoust. Speech Signal Process.* **37**(12), 2024 (1989).
19. A. K. Katsaggelos, J. Biemond, R. W. Shafer and R. M. Mersereau, *IEEE Trans. Signal Process.* **39**(4), 914 (1991).
20. R. Prost, J. L. Burdeau, S. Guilloteau and R. Goutte, *Eupsico 94, Proc. Signal Processing VII: Theories and Applications*, Vol. I, pp. 307–310 (1994).
21. F. C. Schweppe, *Uncertain Dynamic Systems*. Prentice-Hall, Englewood Cliffs, NJ (1986).
22. K. Miller, *SIAM J. Math. Anal.* **1**, 52 (1970).
23. A. M. Thompson, J. C. Brown, J. W. Kay and D. M. Titterington, *IEEE Trans. Pattern Anal. Machine Intell.* **11**(7), 674 (1989).
24. N. P. Galatsanos and A. K. Katsaggelos, *IEEE Trans. Image Process.* **1**(3), 322 (1992).
25. S. Mallat, *IEEE Trans. Pattern Anal. Machine Intell.* **13**(4), 326 (1991).
26. V. V. Makarov, *Surf. Interface Anal.* **23**, 899 (1995).
27. B. G. Svensson, M. K. Linarsson, B. Mohadjeri, M. Petravic and J. S. Williams, *Nucl. Instrum. Methods Phys. Res.* **B85**, 363 (1994).
28. D. Barlow, M. G. Dowsett, H. S. Fox and R. A. A. Kubiak, *Nucl. Instrum. Methods Phys. Res.* **B72**, 442 (1992).
29. S. F. Gull and J. Skilling, *IEEE Proc.* **131**, 646 (1984).
30. L. R. Rabiner and B. Gold, *Theory and Application of Digial Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ (1975).
31. D. C. Youla and H. Webb, *IEEE Trans. Med. Imaging* **MI-1** (2), 81 (1982).