Determination of the composition of light thin films with artificial neural network analysis of Rutherford backscattering experiments

V. Matias, G. Öhl, J. C. Soares, and N. P. Barradas*

Instituto Tecnológico e Nuclear, Estrada Nacional 10, Apartado 21, 2685-953 Sacavém, Portugal
and Centro de Física Nuclear da Universidade de Lisboa, Avenida Professor Gama Pinto 2, 1699 Lisboa Codex, Portugal

A. Vieira

Instituto Superior de Engenharia do Porto, R. António Bernardino de Almeida 431, 4200 Porto, Portugal

S. Cardoso and P. P. Freitas

Instituto Nacional de Engenharia e Sistemas de Computadores, R. Alves Redol 9-1, 1000 Lisboa, Portugal
and Instituto Superior Técnico, Avenida Rovisco Pais, 1096 Lisboa, Portugal

AlO_xN_y ultrathin films are used as insulating layers in advanced microelectronic devices. Structural characterization of these films is often done by the Rutherford backscattering (RBS) analysis. The RBS analysis of these oxinitrides is a difficult task since the relevant signals of the spectrum are washed out by the large substrate background and a considerable time is required for an analyst to characterize the sample. In this work we developed specialized artificial neural networks that are able to perform a fast and efficient analysis of the data. The results are in good agreement with traditional methods.

I. INTRODUCTION

Conventional methods involve the use of interactive or batch-oriented computer codes, such as RUMP [1] and NDF [2], that need the intervention of an analyst. The time needed to perform the analysis is in favorable cases of the same order of magnitude of the Rutherford backscattering (RBS) spectrum acquisition, but with interactive codes and complex spectra it can be very time consuming. On the other hand artificial neural networks (ANNs) are becoming a reliable alternative on the RBS data analysis, achieving the same accuracy of other methods while being faster and much easier to interpret by the analyst.

ANNs have been previously applied to the RBS data analysis with considerable success. We studied systems of increasing complexity, starting with Ge-implanted Si [3], where we have a Ge signal well separated from the single-element background. We then applied it to a more complex two-element background, namely, Er implanted on sapphire [4]. Finally, we used ANN to study thin films of NiTaC deposited on Si and analyzed with protons [5], where several signals are superimposed. Note that in this case the scattering cross section of C and Si varies rapidly with energy, thus increasing the difficulty of spectra interpretation. In this work we push the applicability of ANN to the limit by considering a highly complex system of great technological interest.

Thin films of AlN_xO_y are being used as insulating barriers between the metallic layers of spin-tunnel junctions for advanced read and recording devices [6–9]. The RBS analysis of these films is not an easy task [10–12]. Relevant signals

from N and O are small and superimposed to a large background Si signal, and the Al signal is also partially superimposed to the Si background. Due to these difficulties we have recently organized a round robin experiment in collaboration with the International Atomic Energy Agency to address this problem with different analytical techniques [13].

Here we describe the results of applying ANN to this system. The basis ANN structure was adapted in order to deal with the different data manipulation necessary to tackle this problem, namely, smoothing and differentiation necessary to improve the signal-background ratio. To evaluate the ANN results, we assume as reference the values obtained with the usual formalism for RBS spectra analysis, namely, the peak integration method and the step height analysis [14].

II. EXPERIMENTAL DETAILS

We prepared two sets of samples grown on (100) silicon single crystals. The first set, composed of three samples (1–3), had nominal composition AlN. Small amounts of oxygen present in the chamber were incorporated in the samples during preparation. This set was prepared by reactive magnetron sputtering of an Al target in a Nordiko 7000 system. Samples 1 and 3 have a nominal thickness of 1000×10^{15} at./cm^2 and sample 2 has 500×10^{15} at./cm^2. The second set is composed of two samples, 4 and 5, with nominal composition Al_2O_3 and nominal thicknesses of 200×10^{15} at./cm^2 and 1500×10^{15} at./cm^2. They were prepared by ion beam sputtering of an Al target in a Nordiko3000 system, using an Ar-O beam during the deposition (assisted deposition) [15].

The RBS analysis was performed using a He$^+$ beam at the ITN 3 MV Van de Graaff accelerator. We used beam energies between 1.0 and 1.9 MeV, and angles of incidence θ_{inc} between 0° and 30°. We used an annular surface barrier

---

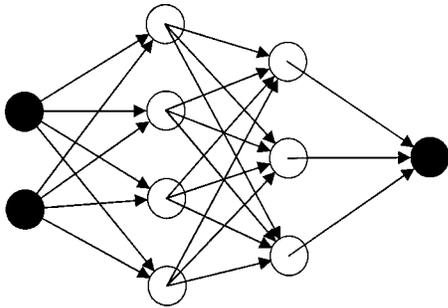*Corresponding author. Fax: +351-219941049. Email address: nunoni@itn.pt

FIG. 1. Schematic representation of an ANN with two nodes in the input layer, one node in the output layer, and two hidden layers.

detector in the IBM geometry placed at about 180° to the incident beam direction. The energy resolution of the system is 26 keV at full width half maximum (FWHM). The product of the detector solid angle with the analyzing beam fluence was between 2.6 and 315.6 μC msr.

III. ARTIFICIAL NEURAL NETWORKS

A. Basic principles

An artificial neural network is a simplistic electronic model inspired by the intricate web of neurons that compose the brain. In this organ, neural cells are highly connected through synapses exchanging excitatory or inhibitory signals. Feedforward ANNs try to capture the brain computational efficiency by considering several layers of artificial neurons that are fully connected by the set of weights, see Fig. 1. The ANN architecture is specified by the number of layers and the number of nodes in each layer. A signal is presented to the input layer and the corresponding result to the output layer. Training is performed by adjusting the connection weights in order to minimize the difference between the ANN result and the desired output. This procedure tries to mimic the massive parallel capabilities of the brain.

These networks are versatile systems that are able to recognize recurring patterns in data, which make them suitable tools to analyze the RBS spectra. Note, however, that after the network is trained the information that codifies the map between input and output is delocalized all through the connections weights array. ANN is thus a black box almost impossible to decipher by the user.

The size of the network depends on the problem complexity and the amount and quality of the training data. Since we have a reliable numerical code to generate pseudoexperimental spectra the amount of data is not a relevant problem. This allows us to work with relatively large networks without warring with the usual training problems like being trapped on a local minimum.

In this work we used backpropagation supervised learning [16] where a large set of examples containing a set of inputs, in our case the spectra and the experimental parameter, and the corresponding outputs, implanted elemental doses, are presented to the network.

B. Network architecture

The network architecture is defined by the size of the input layer, the number of hidden layers and the output layer.

The input layer receives the RBS spectrum data to be analyzed along with some experimental parameters related with the spectrum, namely, the beam energy, angle of incidence, and deposited charge. At the output layer, containing four nodes, the three elements concentrations and the film thickness are presented. Although the RBS spectra used for training have 512 channels, we only used 207 channels, where the Al, O, and N signals are relevant. The input layer has, therefore, 210 nodes considering the three experimental parameters. All ANNs have three hidden layers containing 80, 40, and 20 nodes.

We used four ANNs, one dedicated to raw data analysis (ANN$_{raw}$), another dedicated to the smoothed data (ANN$_{smoothed}$), one dedicated to the analysis of the smoothed and differentiated data (ANN$_{differentiated}$), and finally one equal to the previous, but without information of the experimental conditions (ANN$_{no\ expt.\ param.}$). The architecture of the networks is (N, 80, 40, 20, 4) with N=210 for ANN$_{raw}$, ANN$_{smoothed}$, and ANN$_{differentiated}$, and N=207 for ANN$_{no\ expt.\ param.}$

C. Training and test sets

Analysis of AlO_xN_y thin films is a rather complex problem, since the signals from N and O are small and very close to each other. Furthermore, since the Rutherford cross section depends on the square of the atomic number, signals from these elements are washed up by the stronger signals of heavier elements contained in the films. Finally, a small signal is received in the multiple scattering background region of the spectrum due to the Si buffer. Multiple scattering occurs due to the considerable number of secondary deflections the beam suffers, in addition to the Rutherford backscattering events, on its way in and out of the sample, increasing the number of counts for certain channels in the RBS spectrum. To make the analysis task even more difficult, some of our spectra were channeled, due to the channeling effect in the crystalline Si buffer.

To teach the ANNs, we fed them with a large set of the RBS examples, called the training set, and another smaller independent set to test their performance—the test set. We take particular care to build a training set composed by a sufficiently large and representative set of examples that adequately represent all possible experimental situations. We used a computational model [2,3] to obtain 20 000 simulated experimental data, to which we added realistic Poisson noise in order to simulate the statistical fluctuation associated to the experimental data. The obtained spectra were initially supposed to cover all the experimental conditions we used to analyze our samples, namely, beam energy between 1 and 2 MeV, angle of incidence between −40° and 40°, charge-solid angle product between 52.4 and 393 μC msr, film thickness between 150 and 850×10^{15} at./cm^2, Al atomic concentration between 19% and 77%, O and N atomic concentrations between 0% and 68%. In fact, sample 5 is thicker than the upper limit in these training data, and in a few cases the analyzing beam fluence was lower than the lower limit used. This provides us with the cases where the limitations of the ANNs developed can be tested.
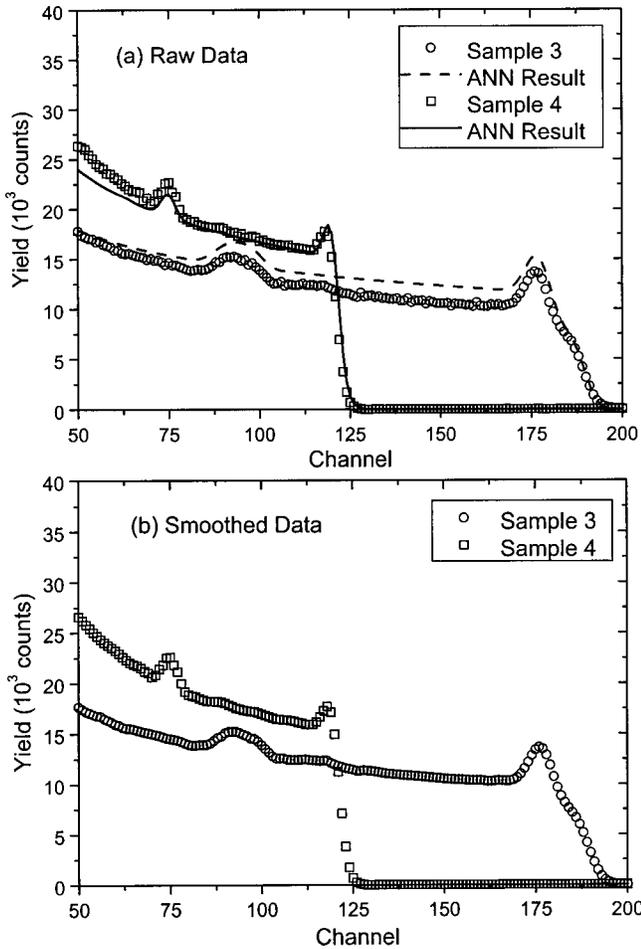
FIG. 2. The RBS spectra of samples 3 and 4. (a) Raw data. The solid lines are results obtained with $ANN_{differentiated}$. (b) Smoothed data.

To simulate the channeling in Si that leads to a reduction in the yield due to the Si crystal, we used a zeroth-order approximation, which consisted of multiplying the Si signal with a constant between 0 and 1. This ignores the fact that the yield reduction due to channeling depends on the beam energy, on the probed depth, and on the exact orientation of the crystal relative to the incident beam [14]. There is no simple and accurate way of calculating these effects, which is a problem in data analysis using conventional methods, where the simple approximation used would lead to wrong results. However, ANNs are, in principle, able to abstract these features, and this zeroth-order approximation should suffice, particularly when combined with the methods to reduce the influence of the background in the analysis.

D. Data preprocessing

Since these spectra are composed of small signals on top of a large background signal, we had to develop numerical transformations to better expose the relevant signals. First, we smoothed the raw data in order to reduce statistical fluctuations. This proved to be ineffective as can be seen by comparing the raw and smoothed data in Fig. 2.
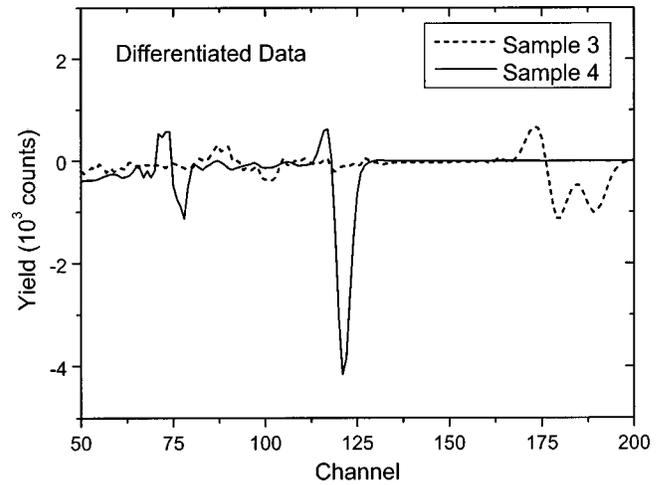

FIG. 3. Spectra from the differentiated raw data for samples 3 and 4.

Since the O, N, and Al peaks change rapidly, while the Si background signal is nearly constant, the relevant signals should be enhanced by smoothing and differentiating the raw data spectra. Figure 3 shows two examples of differentiated spectra. We used smoothing and differentiating routines with cubic fitting splines and adaptive smoothing weights prescribing third end point derivatives $f'''(x_0)=f'''(x_n)=0$ [17]. This procedure leads to differentiated data with minimized fluctuations due to statistical uncertainties.

E. Training and overfitting

From the set of simulated experimental RBS spectra, 18 000 examples were used to train the network, while the other 2000 were used to test the network, not being used for training. For each example, the corresponding output is compared with the output evaluated by the ANN. The calculated error, a mean-squared error (MSE), between the ANN result and the expected result is used as a control tool to adjust the weights between the nodes during the feedforward backpropagation learning process. During the training and test stages, after a few iterations, all the cases with a MSE larger than 40% are excluded from both sets, as can be seen in Fig. 4 around iteration 50. The MSE evolution with the iterations is shown in this figure for the training and test sets.

The learning process should stop to avoid overfitting. This point is achieved when the test set error starts to increase, as can be seen in Fig. 4 after iteration 3000. The final weights taken are those for which the test set error is smallest. Table I shows the final MSE for the train and test sets for each ANN. In all cases they are rather small, indicating that the networks could generalize well the patterns in the data. The errors are smallest for ANN_{raw}, comparable for $ANN_{smoothed}$ and $ANN_{differentiated}$, and largest for $ANN_{no\ expt.\ param}$. Note that these errors are only an indication of the performance of ANNs on the trained data. The real test is their performance on *real* experimental data

IV. RESULTS AND DISCUSSION

We trained and tested the ANNs to analyze the RBS data from our samples. In all cases the results were compared
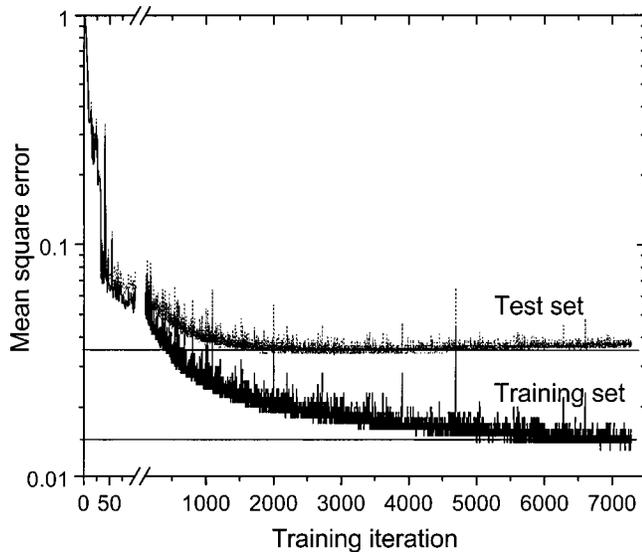
FIG. 4. Mean-square error evolution with the training iteration. Overfitting starts around iteration 2000, where the error on the training set continues slowly to improve, but the error on the test set settles around the same value and starts to worsen after some iterations, as can be verified around iteration 5000.

with reference values obtained by applying the peak integration method [14]. A comparison between the four networks is presented in Table II. To evaluate the quality of each network, the ratio between the ANN result and the reference value is calculated for each sample. The ANN result was taken to be the average of the results obtained for all the

TABLE I. Mean-squared errors in the training and test sets for all the studied networks.

Network	Train set MSE (%)	Test set MSE (%)
ANN_{raw}	2.2	2.8
ANN_{smooth}	2.9	3.7
$ANN_{differentiated}$	2.4	3.4
$ANN_{no\ expt.\ param.}$	3.1	4.3

spectra measured from each given sample. The standard deviation of the ratio values is also shown in the table since it reflects the stability of the ANN against the variation of the experimental conditions. Even for an average ratio of 1, a high standard deviation would mean that the ANN performs poorly.

We see that the results from all ANNs are in general good and fairly similar. The most important exception is sample 5, where all ANNs performs poorly. This can be explained by the fact that the film thickness (nominally 1500 $\times 10^{15}$ at./cm^2) is significantly above the maximum value used to train the ANNs (850×10^{15} at./cm^2), which means that this sample is outside the range where the generalizations capabilities of the ANNs are valid. This shows that the representativeness of the training data is a determinant factor of the reliability of ANN predictions.

The similarity of results obtained with ANN_{raw} and $ANN_{smoothed}$ indicates that the network analysis did not improve much by smoothing the raw data. Smoothing could reduce slightly some noise, but since the relevant signals

TABLE II. Average and standard deviation σ (between parentheses) for the ratio between the ANN values and the reference values for the film thickness, Al, N, and O concentration for all the samples from different spectra, as obtained with the four different ANNs developed. For samples 4 and 5 the values are based on two spectra only.

Sample	ANN	Thickness	[Al]	[N]	[O]
1	raw	0.92 (0.12)	0.87 (0.08)	1.17 (0.08)	0.75 (0.5)
	smoothed	0.92 (0.12)	0.85 (0.08)	1.2 (0.08)	0.78 (0.5)
	differentiated	0.94 (0.16)	0.94 (0.1)	1.07 (0.1)	1.1 (0.3)
	no expt. param.	0.84 (0.19)	0.83 (0.08)	1.14 (0.09)	1.66 (0.78)
2	raw	1.14 (0.18)	0.76 (0.17)	1.19 (0.09)	1.5 (1.31)
	smoothed	1.4 (0.18)	0.77 (0.2)	1.17 (0.07)	1.59 (1.24)
	differentiated	1.06 (0.08)	1.01 (0.08)	1.02 (0.08)	0.79 (0.34)
	no expt. param.	1.09 (0.15)	0.98 (0.11)	1.04 (0.11)	0.79 (0.26)
3	raw	0.91 (0.17)	0.95 (0.13)	1.12 (0.13)	0.51 (0.88)
	smoothed	0.9 (0.15)	0.93 (0.11)	1.1 (0.15)	0.85 (0.86)
	differentiated	0.88 (0.15)	0.99 (0.12)	1.04 (0.1)	0.86 (0.3)
	no expt. param.	0.86 (0.12)	0.93 (0.14)	1.09 (0.11)	0.92 (0.51)
4	raw	1.14 (0.21)	0.83 (0.09)	7.16 (0.64)	0.75 (0.15)
	smoothed	1.1 (0.2)	0.89 (0.1)	6.91 (0.63)	0.73 (0.16)
	differentiated	0.82 (0.17)	1.13 (0.1)	1 (0.27)	0.98 (0.07)
	no expt. param.	0.82 (0.16)	0.95 (0.09)	1.3 (0.31)	1.09 (0.08)
5	raw	0.08 (0.03)	1.27 (0.12)	16.82 (1.72)	0.01 (0.04)
	smoothed	0.11 (0.07)	1.09 (0.1)	19.1 (1.83)	0.01 (0.05)
	differentiated	0.1 (0.03)	1.38 (0.11)	15.21 (1.64)	0.04 (0.01)
	no expt. param.	0.1 (0.01)	1.7 (0.19)	10.29 (4.59)	0.09 (0.02)

TABLE III. Results from the ANN$_{differentiated}$ analysis of all the experimental spectra obtained from our sample sets, with different experimental conditions and the calculated reference values for the same samples. The shaded parameters are outside the range used in the training.

Sample	Spectrum	$Q \times \Omega$ (μC msr)	θ_{inc}	Energy (keV)	Thickness (10^{15} at./cm^2)		[Al] (at. %)		[N] (at. %)		[O] (at. %)	
					ANN	Reference	ANN	Reference	ANN	Reference	ANN	Reference
1	aaln0	314.4	0°	1900	708.46	790.1	47.18	51.3	49.1	44.8	3.71	3.9
	aaln1	157.2	30°	1900	712.18		49.31		46.7		4	
	aaln2	157.2	0°	1900	768.58		48.21		48.08		3.71	
	aaln21	2.57	7°	1900	439.75		46.79		47.63		5.57	
	aaln22	15.72	7°	1900	867.45		57.36		38.29		4.35	
	aaln23	78.6	7°	1900	951.11		53.99		42.79		3.22	
	aaln24	62.88	30°	1900	683.69		49.41		46.51		4.08	
	aaln3	157.2	7°	1500	699.12		42.91		52.2		4.89	
	aaln4	157.2	0°	1500	802.14		44.56		50.81		4.62	
	aaln5	157.2	30°	1500	721.38		45.96		48.57		5.47	
	aaln6	157.2	7°	1100	808.72		44.49		52.06		3.45	
	aaln7	157.2	7°	1100	808.7		44.49		52.05		3.45	
	aaln8	62.88	30°	1100	702.43		49.35		45.49		5.16	
2	aaln17	157.2	7°	1900	436.06	413.65	47.56	47.6	49.49	47.5	2.95	4.9
	aaln18	78.6	30°	1900	394.82		41.66		54.34		3.99	
	aaln13	157.2	7°	1500	487.23		49.04		46		4.96	
	aaln14	78.6	30°	1500	456.83		49.67		45.6		4.74	
	aaln9	157.2	7°	1100	426.34		51.18		45.4		3.42	
	aaln10	78.6	30°	1100	425.16		48.19		48.51		3.29	
3	aaln19	157.2	7°	1900	856.07	792.5	49.8	48.2	44.19	45.1	6.01	6.7
	aaln20	78.6	30°	1900	558.24		46.48		48.22		5.3	
	aaln15	157.2	7°	1500	752.57		46.7		47.4		5.9	
	aaln16	78.6	30°	1500	696.52		40.45		53.73		5.82	
	aaln11	157.2	7°	1100	715.08		56.71		39.88		3.41	
	aaln12	78.6	30°	1100	618.7		44.82		47.2		7.98	
4	aaln25	91.7	7°	1000	175.41	209	42.3	38.1	3.48	3.9	54.22	53.9
	aaln26	62.09	7°	1000	166.5		43.91		4.29		51.79	
5	aaln27	151.96	7°	1000	200.37	1989	51.03	37	46.99	3.1	1.98	55.4
	aaln28	90.65	7°	1000	195.77		50.74		47.29		1.97	

remain small, the improvement should not be significant.

However, differentiation of the data proved to be very effective in outlining the O and N peaks. This explains the better results obtained with ANN$_{differentiated}$ in most cases. This is particularly true for the smallest signals, which are the O concentration in samples 1–3 and the N concentration in sample 4, where ANN$_{differentiated}$ outperforms ANN$_{raw}$ and ANN$_{smoothed}$ both in the average result and in the standard deviation. This shows clearly that the differentiation was highly successful in enhancing small signals and reducing the influence of the background.

We show in Table III the results obtained with ANN$_{differentiated}$ for all the individual spectra. This table contains the information from several spectra acquired for each sample, varying the incident beam energy, the incidence angle, and the deposited charge. In this table are also shown the calculated reference values for the thickness and composition of each sample. Again, some of the ANN results do not agree well with the reference values, for instance, the thickness results obtained from the ANN analysis of spectrum

aaln21 are almost a factor of 2 smaller than the reference values. This is due to the low charge in which this spectrum was acquired.

The results obtained with ANN$_{no expt. param.}$ are comparable with those obtained with ANN$_{differentiated}$, but less accurate. This indicates that, even without the information on the experimental parameters, the network structure is versatile enough to recognize the relevant patterns for the analysis. The experimental parameters that are introduced in the other three networks are the beam energy, the deposited charge, and the sample incidence angle. The beam energy is responsible for the peak positions, the deposited charge influences the yield, and consequently the height of the peaks, and the incidence angle plays a role in the thickness of the peaks due to the larger path of the beam in the sample for a same range.

To understand why ANN$_{no expt. param}$ was successful in analyzing the data, we consider the following factors. First, the energy calibration was the same in all spectra. This means that the position of Al (which is the signal that appears most to the right in the data) is an effective measure of

the beam energy. Second, if one assumes that only N, O, and Al are present, then the collected charge does not need to enter the calculations, which can be based on the ratio between the three signals. As the neural network was only fed with data that indeed only had N, O, and Al in the film, then it was implicitly supplied that information. Third, in first order, the effect of tilting the sample is simply to increase the number of counts of each element by the same factor (this ignores possible cross section and stopping power effects, which are small for these films), and, again, does not change the ratio between them. In this case, the thickness cannot be estimated from the total number of counts, but the tilt leads to a shift in the position of the Si edge that can be detected and used by $ANN_{no\ expt.\ param}$.

A comparison between the raw data and a simulation done with the $ANN_{differentiated}$ results for samples 3 and 4 is shown in Fig. 2(a). For sample 4, the results from the ANN are very close to the raw data meaning that the analysis performed by the network was successful. For sample 3, at first sight the analysis was less successful, because the simulated yield is higher than the experimental yield. However, this is due to channeling in the Si crystal, which occurred in the experiment, but was not taken into account in the simulation. This does not affect the signal of Al, N, and O, or the shape and height of those peaks. Results from ANN are very close to the raw data, which indicates that the analysis is correct. This is also confirmed by the excellent values of thickness and concentration obtained with $ANN_{differentiated}$ for this spectrum.

V. CONCLUSIONS

We studied the performance of artificial neural networks on a difficult problem of the RBS analysis of spectra from AlN_xO_y thin films deposited on Si substrates. We developed a convenient preprocessing method to increase the signal-to-noise background ratio: smoothing the raw data to eliminate noise, and then differentiating. In order to test further the generalization capabilities of the artificial neural networks, we trained a neural network without information about the experimental conditions.

After training the networks we applied them to the real experimental RBS spectra obtained from four samples measured in different experimental conditions. The results were very close to the reference values obtained with traditional time-consuming methods. The ANN for the differentiated data performed better than the other ones because the signal from the slowly varying Si background is minimized. The ANN with no knowledge on the experimental conditions also performed well, although slightly worse, because it could extract from the data itself the information required to do a valid analysis.

Finally, in a few cases the samples analyzed had parameters not within the range in which the ANNs were trained. In those cases, all the ANNs performed very poorly. This shows that the way the network is trained is determinant for a successful analysis of real experimental data, and the training and test sets must be as complete as possible, covering all the possible experimental situations.

[1] L.R. Doolitle, Nucl. Instrum. Methods Phys. Res. B **9**, 344 (1985).

[2] N.P. Barradas, C. Jeynes, and R.P. Webb, Appl. Phys. Lett. **71**, 291 (1997).

[3] N.P. Barradas and A. Vieira, Phys. Rev. E **62**, 5818 (2000).

[4] N.P. Barradas, A. Vieira, and E. Alves, Nucl. Instrum. Methods Phys. Res. B **175-177**, 108 (2001).

[5] A. Vieira and N.P. Barradas, Nucl. Instrum. Methods Phys. Res. B **174**, 367 (2001).

[6] S. Cardoso, P.P. Freitas, C. de Jesus, P. Wei, and J.C. Soares, Appl. Phys. Lett. **76**, 610 (2000).

[7] S. Tehrani, J.M. Slaughter, E. Chen, M. Durlam, J. Shi, and M. de Herrera, IEEE Trans. Magn. **35**, 2814 (1999).

[8] A.E.T. Kuiper, M.F. Gillies, V. Kottler, G.W. 't Hooft, J.G.M. van Berkum, C. van der Marel, Y. Tamminga, and J.H.M. Snijders, J. Appl. Phys. **89**, 1965 (2001).

[9] H. Boeve, E. Girgis, J. Schelten, J. de Boeck, and G. Borghs, Appl. Phys. Lett. **76**, 1048 (2000).

[10] B. Brijs, J. Deleu, T. Conard, H. De Witte, W. Vandervorst, K. Nakajima, K. Kimura, I. Genchev, A. Bergmaier, L. Goergens, P. Neumaier, G. Dollinger, and M. Döbeli, Nucl. Instrum. Methods Phys. Res. B **161−163**, 429 (2000).

[11] L.L. Cheng, Y.H. Yu, B. Sundaravel, E.Z. Luo, S. Lin, Y.M. Lei, C.X. Ren, W.Y. Cheung, S.P. Wong, J.B. Xu, and I.H. Wilson, Nucl. Instrum. Methods Phys. Res. B **169**, 94 (2000).

[12] S.R. Walker, J.A. Davies, P. Mascher, S.G. Wallace, W.N. Lennard, G.R. Massoumi, R.G. Elliman, T.R. Ophel, and H. Timmers, Nucl. Instrum. Methods Phys. Res. B **170**, 461 (2000).

[13] *A round robin characterization of the thickness and composition of thin to ultrathin* AlN(O) *layers*, International Atomic Energy Agency Research Contract No. 11317/R0/Regular Budget Fund.

[14] *Handbook of Modern Ion Beam Analysis*, edited by J.R. Tesmer and M. Nastasi (Materials Research Society, Pittsburg, 1995).

[15] V. Gehanno, P.P. Freitas, A. Veloso, J. Ferreira, B. Almeida, J.B. Sousa, A. Kling, J.C. Soares, and M.F. de Silva, IEEE Trans. Magn. **35**, 4361 (1999).

[16] Christopher M. Bishop, *Neural Networks for Pattern Recognition* (Oxford University Press, Oxford, 1995).

[17] Gisela Engeln-Müllges and Frank Uhlig, *Numerical Algorithms with Fortran* (Springer-Verlag, Berlin, 1996).