Artificial neural networks for automation of Rutherford backscattering spectroscopy experiments and data analysis

N. P. Barradas*

Instituto Tecnológico e Nuclear, Reactor, Estrada Nacional 10, Apartado 21, 2686-953 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

R. Patrício

Instituto Tecnológico e Nuclear, Reactor, Estrada Nacional 10, Apartado 21, 2686-953 Sacavém, Portugal

We present an algorithm based on artificial neural networks able to determine optimized experimental conditions for Rutherford backscattering measurements of Ge-implanted Si. The algorithm can be implemented for any other element implanted into a lighter substrate. It is foreseeable that the method developed in this work can be applied to still many other systems. The algorithm presented is a push-button black box, and does not require any human intervention. It is thus suited for automated control of an experimental setup, given an interface to the relevant hardware. Once the experimental conditions are optimized, the algorithm analyzes the final data obtained, and determines the desired parameters. The method is thus also suited for automated analysis of the data. The algorithm presented can be easily extended to other ion beam analysis techniques. Finally, it is suggested how the artificial neural networks required for automated control and analysis of experiments could be automatically generated. This would be suited for automated generation of the required computer code. Thus could RBS be done without experimentalists, data analysts, or programmers, with only technicians to keep the machines running.

I. INTRODUCTION

Rutherford backscattering (RBS) is a standard technique to study the elemental composition and depth profile of samples, and as such it is widely used [1]. It belongs to the cluster of techniques known as ion beam analysis (IBA). One of the advantages of RBS is that it is based on Newtonian physics, and hence it is fully quantitative without the need to recur to external standards. In fact, it is often used to calibrate standards for other techniques [2–4].

Given a known elemental depth profile, it is easy to calculate the corresponding RBS spectrum for given experimental conditions. The inverse problem is, however, considerably more complex. Until recently, data analysis was done either manually for simple cases, or with the aid of interactive computer programs as reviewed in, e.g., Ref. [5]. The analyst would input a given depth profile, let the program calculate the spectrum, compare it to the data, and iterate until sufficiently good agreement was found. It was the analyst's task to decide what agreement was sufficiently good.

A computer program, called NDF, for automated data analysis has been presented [6]. It is based on the simulated annealing algorithm [7,8], and requires the user to input the data, the experimental conditions, and the elements present, with the analysis being done automatically without user in-

terference. This code, being general, is not optimized for any given system, and in complex cases it can take a long time to run.

We presented recently an artificial neural network (ANN) to analyze data from a specific simple system, namely Ge-implanted Si [9]. We also presented ANNs for more complex systems, such as Er-implanted sapphire (bielemental substrate) [10], and elastic backscattering of Ni-Ta-C thin films deposited on Si [11]. The advantage of ANNs is that, while each one is dedicated to a single system, they are optimized for that system, and the analysis is instantaneous, which opens the door to analysis of large batches of samples with RBS and other IBA techniques, something that is currently not done due to the inefficiency in the data analysis.

However, the experimental conditions must often be optimized for each given sample. NDF, like other codes presented in the literature (a review of existing codes can be found in Ref. [5]), is not able to perform this task. Analysis of large batches of samples would require the presence of a skilled experimentalist, which again is inefficient and costly.

The purpose of this work is to present an algorithm, based on ANNs, which is able, for a given system, namely, Ge-implanted Si, to optimize the experimental conditions for each sample, and then analyze the final spectrum collected. The algorithm is easily extendable to other systems. Once this algorithm is implemented in a code connected to an experimental setup with automated sample loading, this will lead to the performance of RBS experiments entirely without the assistance of humans.

_____

*Author to whom correspondence should be addressed. Electronic address: nunoni@itn.pt

II. ARTIFICIAL NEURAL NETWORKS

A. Basic principles

ANNs can, in principle, approximate an arbitrary unknown function [12]. An ANN consists of an array of input nodes connected to an array of output nodes through successive layers of intermediate nodes. Each connection between nodes has a weight, initially random, which is adjusted during a training process. The output of each node of a specific layer is a function, usually a sigmoid, of the sum on the weighted signals coming from the previous layer. Hence, for given inputs (for instance, RBS spectral data and the corresponding experimental parameters), the outputs (for instance, the dose and depth of a given implant) are directly calculated.

To train the network we chose the backpropagation algorithm [12]. A large number of examples, called the training set, for which the outputs are known are consecutively presented to the input layer and propagated to the output layer. For each example in succession, the corresponding output y_n is compared with the evaluated output o_n and the error is computed. The weights of the last layer are adjusted in order to reduce the error corresponding to each output node. This adjusting process is then backpropagated through the successive layers. Several iterations are performed, each iteration consisting of presenting to the ANN the whole training set.

The results provided by the ANN so obtained are then compared with known examples not used in the training, called the test set. The training process is stopped when a minimum error is found for the test set. The ANN can then be used to evaluate examples for which the parameters of interest (the outputs) are unknown. The error should be about the same as for the test set.

Thousands of training data are often required, and it is not feasible to collect so many RBS spectra. We hence used constructed spectra to train the ANNs used in this work. This means that, although the generated data were as realistic as feasible, in principle the ANN error for real data will increase slightly. We previously described the algorithm used to generate realistic RBS spectra [9].

B. Definition of spectral classes

In our previous work three distinct classes of spectra were identified [9]. They were (1) spectra with well separated Si and Ge signals and a large Ge peak compared to the background; (2) spectra with superimposed Si and Ge signals; and (3) spectra with separated Si and Ge signals, but with a very small Ge peak. Spectra belonging to classes (2) and (3) are difficult to analyze, involving an error larger than that for spectra belonging to class (1).

The distinction between classes is not always clear, due to different reasons. There may be partial superposition between the Si and Ge signals. While the statistical error on the Ge peak is a measure of how small the peak is, it is not possible to define an exact threshold separating classes (1) and (3). Furthermore, a spectrum with superimposed signals can also have a very small Ge peak. That is to say, there must be a certain fuzziness in the definition of the classes, which

to some extent must be *ad hoc*. This is defined for each spectrum as a weight of belonging to each class, that is, membership probabilities. When one needs to decide in a clear-cut way to which class a given spectrum belongs, we take it to be the class with the largest membership probability.

We considered that a pure class (1) spectrum is one where no superposition exists and the statistical error at the Ge peak is smaller than 3%. A pure class (2) spectrum is defined as having the Ge peak in a channel smaller than the Si edge, with a Ge dose larger than 3×10^{15} Ge/cm^2. A pure class (3) spectrum requires signal separation and a statistical error in the Ge peak larger than 12%.

Simultaneous membership of classes (1) and (2) occurs when the Ge peak is in a channel larger than the Si edge channel, but with some degree of superposition between the two signals. We define the degree of overlap to be the fraction of Si counts relative to Ge counts in the region of the Ge peak. It is 1 when there are at least as many Si counts as Ge counts in that region. The class (2) weight is the degree of overlap.

Simultaneous membership in classes (1) and (3) was considered to exist for a statistical error at the Ge peak between 3% and 12%. The class (3) weight is larger for larger errors. The class (1) and class (3) weights were defined to be equal for an error of 7.5%.

When the Si and Ge peaks are superimposed but the Ge dose is smaller than 3×10^{15} Ge/cm^2, there is overlap between classes (2) and (3). The class (2) membership probability is 1 for that value of the dose, and 0.6 for 1×10^{15} Ge/cm^2, which is the minimum Ge dose value considered in the class (2) data set.

A spectrum with low Ge dose and partial signal superposition will belong, with a certain probability, to all three classes.

For each class we generated 5000 spectra to build the training set and 1000 for the test set. The three training and test sets together constituted the training and test sets for the classification ANN. One should note that the ANNs developed for each class should be able to analyze correctly border cases, that is a spectrum belonging 50% to class (1) and 50% to class (3) should be correctly analyzed by the ANNs developed either for class (1) or for class (3). In order to ensure this, each training and test set included border cases, for instance, the training and test sets for class (1) included spectra with class (2) and class (3) weights up to 0.5.

The distribution of the depth and dose in the training and test sets are given in Fig. 1 for the three classes. It is clear that there is extensive superposition in the range convered by each of the classes. The minimum dose for class (2) was cut at 1×10^{15} Ge/cm^2, as it was considered that a spectrum with superimposed signals and an even smaller dose is impossible to analyze. Even for doses smaller than around 1×10^{16} Ge/cm^2, the task is considerably difficult.

The spectra were calculated for a very broad range of realistic experimental conditions [9]. The beam was ^{4}He, with energy between 1 and 2 MeV, scattered at an angle between 130° and 180° and detected with a resolution between 11 and 40 keV full width at half maximum (FWHM).
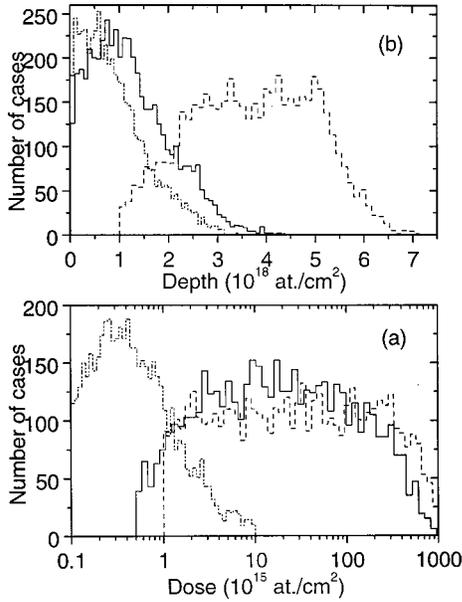
FIG. 1. Distribution of implanted (a) dose and (b) depth in the generated data.

The normal of the sample was tilted at an angle between $-30°$ and $30°$ with the direction of incidence. The collected charge (beam fluence) was between 0.2 and 250 μC for a solid angle of 1 msr. We added the contribution of pulse pileup to the theoretical spectra [13], as well as Poisson noise in order to simulate experimental data as closely as possible.

C. ANNs for calculation of implant parameters and for classification

We developed four different ANNs: one for each class, dedicated to determining the implanted dose and depth, and one to classify the spectra. They were all multilayer perceptions (MLP), trained with backpropagation [14].

The individual class ANNs, called $(ANN)_i$ for class (1), all had the same architecture, previously optimized for the problem, consisting of three hidden layers, with 100, 80, and 50 nodes. The optimization procedure has been previously described [9]. The inputs are the yield x of 128 channels, normalized to the charge–solid-angle product C, the beam energy E and energy resolution, the angle of incidence θ_{inc}, the scattering angle α_{scatt}, and the pulse pileup factor. The outputs y_n are the implanted dose and depth, to be compared with the evaluated output o_n. The figure of merit to be minimized was the mean square error,

$$E = \frac{1}{2} \sum_{n=1}^{N} (y_n - o_n)^2. \qquad (1)$$

For class (3) the errors were consistently very high, due to the fact that the Ge peak is almost unnoticeable as the doses are very small for that class, so that relative errors are very high. In order to enhance performance we decided to apply the following transformation:

$$\tilde{x} = \frac{1}{\ln x + a}, \qquad (2)$$

where a is a small number (we used $a = 5$). This transformation has the advantage of eliminating the unimportant signal from the Si background and enhance the yield from the Ge peak.

The inputs of the classification network, that we name ANNC, were the same as those of ANN1 and ANN2. We chose the 1-of-c coding scheme [15], and therefore there are three outputs, one for each class. We used a unipolar activation function.

Concerning the error function, it is known that the mean square error (mse) is not the most appropriate for classification. It is derived from the maximum likehood assumption of Gaussian distributed target data [16]. However, the 1-of-c coding scheme is binary and hence far from having a Gaussian distribution. A better alternative is to use a cross-entropy error function such as [12]

$$E = -\sum_{n} [t^n \ln y^n + (1 - t^n)\ln(1 - y^n)], \qquad (3)$$

where t^n are the target values (0 or 1 for pure-class spectra that is belonging to a single class) and y^n the actual outputs of the network. By using a sigmoid for the output activation function, the error term applied to the output layer is just

$$\delta^n = (y^n - t^n)^2, \qquad (4)$$

which is equal to the sum of squares error with a linear output activation function.

The advantage of using cross entropy is that it performs better than the sum of squares error function, especially when estimating small membership probabilities. Since we used the logistic activation function for the outputs, we can interpret their values as probabilities of the presence of the corresponding classes [17].

In classification problems we can interpret the role of the hidden layers to be the determination of the weights for the last layer so that an optimum discrimination of the classes of input vectors is achieved in the last layer by means of a linear transformation. Minimizing the error of this linear discriminate requires that the input data undergo a nonlinear transformation into the space spanned by the activations of the hidden units in such a way as to maximize a given discriminant function. It is thus desirable that the ANN should have at least two hidden layers in order to discriminate arbitrary complex decision regions.

We tested the performance of several architectures. Those with a single hidden layer performed very poorly, while increasing from two to three hidden layers improved the error considerably. An increase to four hidden layers did not lead to a further decrease of the error, and introduces unnecessary complexity.

The ANNC chosen had thus three hidden layers with about 25 000 weights. This confirms our previous findings where we found that, for RBS, given a certain number of weights it is preferable to have a MLP with multiple hidden

TABLE I. Summary of the actions taken by the feedback algorithms $F1$, $F2$, and $F3$ corresponding to classes (1), (2), and (3). $E_{\text{min. no. superp.}}$ is the minimum beam energy (between 1 and 2 MeV) for which there is no signal superposition.

Algorithm	E	θ_{inc}	α_{scatt}	C
$F1$		$0°$		$2.25F_{\text{old}}$
		If partial		If low statistics
		Superposition		
$F2$	2 MeV	$0°$	$180°$	
	(Third step)	(First step)	(Second step)	
$F3$	$1.1E_{\text{min. no. superp.}}$			$4F_{\text{old}}$
	If $C>300$ μC msr			

layers in a triangular shape than a single one. Although two hidden layers should be sufficient, when enough training data are available as is the case, the inclusion of more hidden layers may achieve better results at the cost of a slower training and local minimum trapping.

III. THE FEEDBACK ALGORITHM

A. Basic principles

One can devise an algorithm to obtain optimal or near-optimal experimental conditions to determine any given property that one wishes to derive from RBS data from any given system (in the present case, implanted Ge dose and depth in Si). Such an algorithm could be based on ANNs: one could examine the results for the test set, and use principal component analysis or clustering algorithms [18] to determine the set of experimental conditions (more accurately, a space of experimental conditions) that leads to the smallest error in the determination of the property of interest.

However, for the simple system studied in this work, it is possible to define a set of heuristic rules derived from physical principles to determine the optimal experimental conditions for each of the classes defined. This is the method used here, and it could also be used for any system composed of an element implanted into a substrate made of lighter elements only.

The optimal experimental conditions are different for each class. Hence three different feedback algorithms are implemented, one for each class. We call them $F1$, $F2$, and $F3$ for classes (1), (2), and (3). ANNC gives as output, membership probabilities for each class. In the general case, the three weights are nonzero. The spectrum is thus sent to all the feedback algorithms. Each one will return different suggested modified values for the experimental conditions. An average, weighed by the class membership probabilities, is then taken to obtain the new experimental parameters. A new spectrum is measured for those parameters, leading to the next iteration.

Finally, one should note that it is experimentally easier to change some parameters than others. For instance, changing the beam energy can be a complicated process. Changing the scattering angle requires having a movable detector, which is not always available. On the other hand, changing the angle of incidence is normally trivial as the samples are almost always located on a goniometer. And increasing the collected statistics requires only increasing the beam fluence, that is, measuring for a longer time. In the algorithm developed, the easier alternatives are adopted first, and other options are considered only when that is not enough to obtain the desired effect.

The actions taken by the feedback algorithm for each of the classes are summarized in Table I.

B. Class (1)

For class (1), no optimization is normally required, as it corresponds to those spectra for which the analysis error is smallest. In this case, algorithm $F1$ should introduce no changes in the experimental conditions. However, the definition of the classes is fuzzy, and there may be some partial superposition, or the statistics may be relatively low. In the first case, algorithm $F1$ will set the suggested angle of incidence to normal incidence, which leads to a better signal separation: $\theta_{\text{inc}}(F1,\text{partial superposition})=0°$. The algorithm $F1$ will not try to change other experimental conditions such as the beam energy or the scattering angle because those parameters are more complicated to change, and in principle a change in the angle of incidence should be, in most cases, enough to solve partial signal superposition.

When the counting statistics are relatively low, algorithm $F1$ will try to reduce the statistical error by 50%. As the statistical error is proportional to the square root of the yield, this corresponds to a suggested increase in the charge–solid-angle product C of $C_{\text{new}}(F1,\text{low statistics})=2.25C_{\text{old}}(F1)$.

The question arises as to how the feedback algorithm knows whether there is signal superposition or not, or what the counting statistics in the Ge signal is. This cannot be determined accurately unless the Ge dose and depth are known exactly, which, obviously, they are not. The alternative is to calculate a spectrum for the dose and depth determined by ANN1 (or by ANN2 and ANN3 for algorithms $F2$ and $F3$, respectively), and evaluate signal superposition and counting statistics for that spectrum. This procedure is as good as ANN1, that is, if the error in the ANN1 calculation is small, then the procedure is valid.

C. Class (2)

For class (2), the problem is to separate the signals, that is, the energy difference between the Ge peak and the edge of the Si signal should become larger than zero and as large

as possible. This energy difference depends, on one hand, on the depth of the Ge implant, which cannot be changed. On the other hand, it depends on the analyzing beam energy and on the scattering angle via the kinematic factor, that is, the ratio between the beam energy before and after scattering, and on the angle of incidence via the length of the beam path inside the sample [19].

In simple terms, the signal separation is better for higher beam energy, angle of scattering close to 180°, and angle of incidence 0° (normal incidence). However, the beam energy cannot be increased indefinitely, first because of the maximum energy accessible with a given accelerator, and second because above a given threshold (which is element dependent), nuclear reactions distort the signal. We limit the energy range to the commonly used 1–2 MeV.

Algorithm $F2$ will change θ_{inc}, α_{scatt}, and E in successive iterations. However, if the angle of incidence is close to 0° (smaller than 7.5°) to start with, then α_{scatt} will be changed in the first iteration. Similarly, if α_{scatt} is close to 180° (larger than 165°), E will be increased already in the first iteration. The maximum value allowed for E is 2 MeV, the maximum value used to construct the training and test sets.

D. Class (3)

For class (3), the aim is to increase the counting statistics. This can be achieved in two ways. One is to measure for a longer time, which is limited by the amount of time one reasonably accepts to measure one single sample. The other is to decrease the beam energy, as the Rutherford cross section is proportional to $1/E^2$ [19].

Algorithm $F3$ will try to double the counting statistics: $C_{new}(F3) = 4 C_{old}(F3)$. This is, however, limited to a maximum value of the charge–solid-angle product of 350 μC msr, the maximum value used to construct the training and test sets, and a large value in RBS. If that limit is reached after a given number of iterations, the $F3$ algorithm

TABLE II. Results of the classification ANN for the spectra before optimization of the experimental conditions.

Real ANNC	(1)	(2)	(3)
(1)	91.7	5.9	2.4
(2)	2.7	90.0	7.3
(3)	1.6	33.1	65.3

will try to decrease the beam energy E. However, a very small beam energy could lead to signal superposition. Hence, the algorithm checks what the minimum beam energy is that does not lead to beam superposition (once again, for the currently determined Ge depth), and returns as a suggested new value a beam energy 10% larger than that.

E. Estimation of implanted dose and depth

When none of the feedback algorithms is able to suggest changed values for the experimental parameters (or when the change is smaller than given thresholds), it is considered that the final optimal experimental conditions have been found. Note, however, that only feedback algorithms for ANNs with weight larger than 0.1 are considered. For instance, if the weight of ANN2 is 0.05, it is not considered relevant to the spectrum at hand and algorithm $F2$ is ignored.

A final spectrum is calculated for the optimal experimental conditions, and analyzed by ANNC to classify it, and by ANN1, ANN2, and ANN3 to determine the implanted dose and depth. The final values are obtained by weighting the individual ANN values by the belonging weights to the third power. The weights as such are not used to reduce the error: suppose a given spectrum has a small weight for class (2), but ANN2 calculates a depth overestimated by a very large factor (which it always does for small real depths). This would lead to a very large error in the final weight-averaged
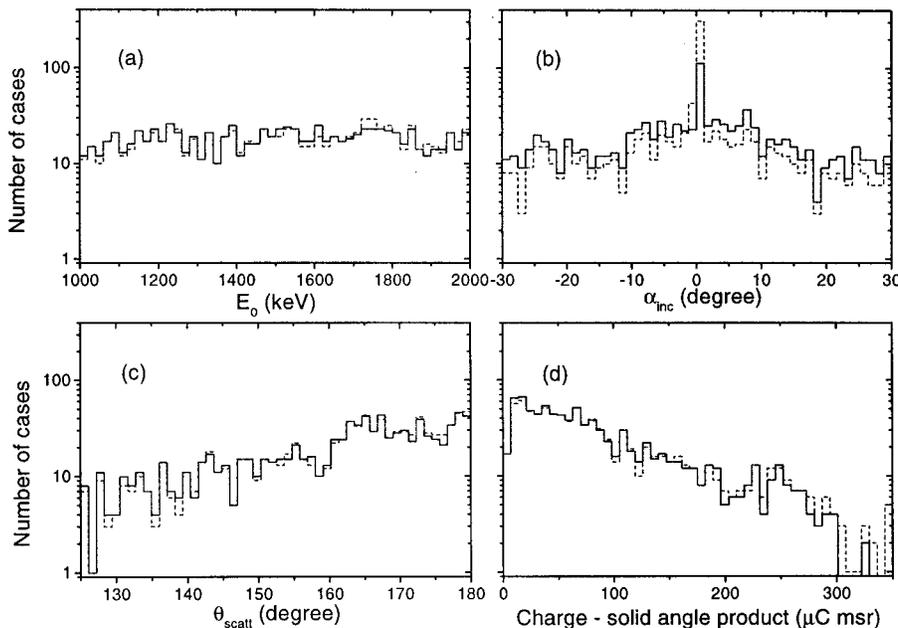

FIG. 2. Distribution of the experimental parameters of class (1) spectra before and after optimization.

TABLE III. Results of the classification ANN for the spectra after optimization of the experimental conditions.

Real ANNC	(1)	(2)	(3)
(1)	98.9	1.1	0
(2)	30.8	65.2	4.0
(3)	62.6	7.5	29.9

depth. By using the weights to the third power, this problem is avoided.

The depth and dose estimated by the ANN cluster are then used as an initial guess in an automated fit with the code NDF, which then provides final values for the implanted depth and dose. Note, however, that NDF does not constrain the Ge profile to any given shape, and the final profile may deviate significantly from Gaussian. In this case, the values determined with NDF may have a large error. This problem could be easily solved by introducing a small change in NDF, forcing the fitted profiles to be Gaussian. In any case, as NDF performs a least-squares fit starting in the values given by the ANN, it will always improve the results. One can thus combine the automation features of ANNs with the accuracy of NDF.

IV. RESULTS AND DISCUSSION

A. Classification

We applied the algorithms developed to 3000 simulated spectra, 1000 of each class, constructed in the same way, and within the same parameter space, as the training and test sets. The results of the classification ANN are shown in Table II. The easiest spectra to classify are those of class (1), where ANNC correctly classifies 91.7% of the spectra. Of the 5.9% that are wrongly assigned to class (2), inspection of the spectra reveals that most have partially superimposed Ge and Si signals, with a nonzero class (2) membership probability. In a similar way, most of the 2.4% wrongly assigned to class (3) have nonzero class (3) weight, that is, they have relatively poor statistics.

The performance of ANNC is also very good for class (2), with 90% of the spectra being correctly classified. 2.7% and 7.3% are classified as belonging to classes (1) and (3), respectively. Most of those that are classified as class (3) have relatively low doses, and the Ge peak is barely visible on top of the Si signal. In class (3) spectra with very low doses, the Ge peak (which in that case is not on top of the Si signal but separated) is also barely visible, that is, low dose class (2) and (3) spectra look rather similar, and hence most of the class (2) classification errors correspond to spectra assigned as class (3). One should note that, if the dose distribution of class (2) allowed for doses as low as those of class (3), the classification error would be still larger.

The performance of ANNC is considerably worse for class (3), reflecting the fact that very small dose values lead to Ge peaks that are almost invisible, that is, almost impossible to analyze no matter what the method employed. Only 65.3% of the spectra are correctly classified, with 33.1% be-

ing incorrectly classified as class (2), although in most cases with relatively high class (3) weight, for the reason pointed out above. The reason why there are more class (3) spectra being classified as class (2) than class (2) spectra being classified as class (3) is that there are more class (3) than class (2) spectra prone to misclassification, that is, with very small or almost invisible Ge signal.

B. Class (1)

Class (1) spectra are not expected to need any optimization of the experimental conditions, aside occasional small increases of the collected charge or setting the angle of incidence to 0°. The distribution of the experimental parameters before and after optimization is shown in Fig. 2, confirming that expectation. After optimization, the number of spectra wrongly identified as class (2) or (3) decreases from 8.3% to 1.1%, as shown in Table II and Table III.

The average and standard deviation of the ratio between the determined depth and dose values (with the ANN for the initial and optimized experimental conditions and by NDF), and the real values, are shown in Table IV. The depth and dose values determined are

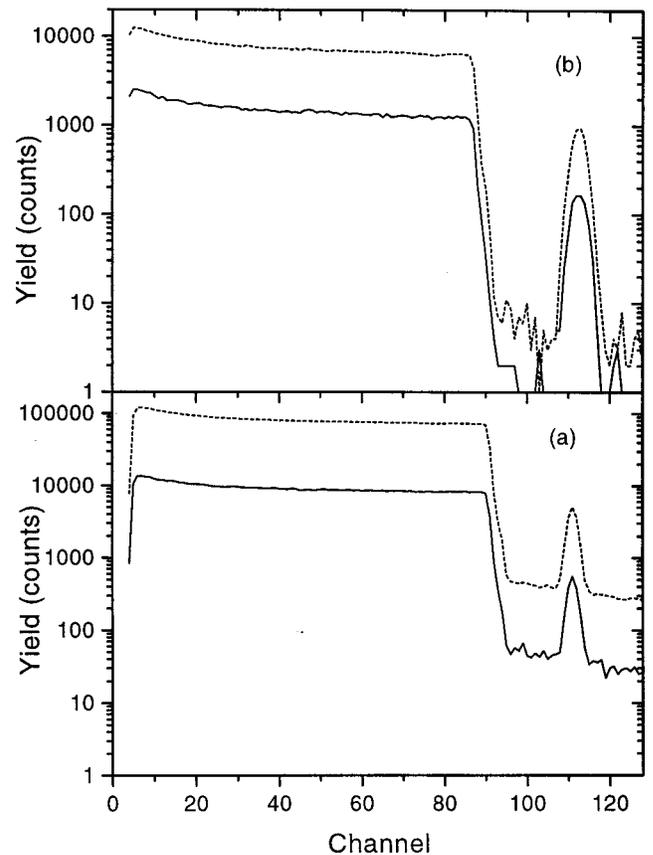$$(\text{ANN})_x = \sum_{i=1}^{3} w_i \text{ANN}_i, \tag{5}$$

FIG. 3. Class (1) spectra. (a) Spectrum was initially wrongly classified as class (3). (b) Spectrum was initially correctly classified as class (1). In both cases the feedback algorithm increased the solid-angle–charge product to improve the statistics.

TABLE IV. Average and standard deviation (given in parentheses) of the ratio between the depth and dose values determined with the ANN for the initial and optimized experimental conditions and by NDF, for all classes.

Real class	Final class	Analysis	$(ANNC)_i$: Correct		$(ANNC)_i$: Wrong	
			Depth	Dose	Depth	Dose
(1)	(1)	$(ANN)_i$	1.03(13)	0.87(14)	1.07(11)	1.08(12)
		$(ANN)_f$	1.01(12)	0.85(14)	0.98(4)	0.81(10)
		NDF	1.00(2)	0.99(2)	1.00(1)	1.00(1)
(2)	(1)	$(ANN)_i$	1.08(13)	1.17(38)	0.66(35)	0.53(47)
		$(ANN)_f$	0.98(3)	0.87(11)	0.72(36)	0.55(35)
		NDF	1.00(1)	0.98(3)	1.00(5)	0.93(17)
	(2) or (3)	$(ANN)_i$	0.98(11)	1.08(56)	0.37(14)	0.14(17)
		$(ANN)_f$	0.98(06)	1.22(51)	0.73(33)	0.94(99)
		NDF	0.97(3)	0.96(15)	0.96(13)	0.88(24)
(3)	(1)	$(ANN)_i$	1.27(65)	1.56(71)	5.80(5.69)	17.1(15.1)
		$(ANN)_f$	1.08(54)	1.08(46)	1.35(80)	1.39(47)
		NDF	1.05(31)	0.93(24)	1.05(25)	0.94(18)
	(2) or (3)	$(ANN)_i$	1.37(71)	2.47(1.34)	8.94(8.21)	33.2(21.0)
		$(ANN)_f$	2.10(1.91)	3.57(4.75)	4.41(5.06)	7.64(9.55)
		NDF	1.27(82)	0.94(25)	2.67(3.00)	1.03(51)

where the w_i are the normalized weights to the third power, and $x \equiv \{i,f\}$ stands for initial and final values (before and after optimization).

We show the values separately for the cases in which the ANNC initially (that is, for the initial experimental conditions) correctly and incorrectly identified the spectrum as belonging to class (1). We call the initial and final (before and after optimization) classification $(ANNC)_i$ and $(ANNC)_f$, respectively. The initial and final determined dose and depth values are called $(ANN)_i$ and $(ANN)_f$, respectively. Further, we considered in the table only those cases in which the final spectrum (after optimization of the experimental parameters) was classified as class (1) by the ANNC. This is of course the vast majority, with only 11 spectra being classified as class (2) or (3), as shown in Table III.

The optimization does not lead to improvements, while NDF is able to retrieve the correct values with high precision. This is true also for the spectra that were initially misclassified, the reason being that, as discussed above, most of those cases have nonzero weights for classes (2) or (3), and hence ANN2 and ANN3 are able to analyze them correctly.

A class (1) spectrum that was initially wrongly classified as class (3) is shown in Fig. 3(a). The correct class weights were (0.77, 0, 0.23), and the initial classification, for a solid-angle–charge product of 30.1 μC msr, was (0.28, 0, 0.71). The statistical error at the Ge peak is 5%. In a first step the solid-angle–charge product was increased to 104.9 μC msr, corresponding to a statistical error in the Ge peak of 2.5%, which led ANNC to correctly classify the spectrum as class (1). As the statistical error was still too high, the feedback algorithm increased further the solid-angle–charge product, obtaining a final statistical error of 1.5%.

A class (1) spectrum that was initially correctly classified is shown in Figure 3(b). The initial statistical error was 7.4% [remember that the transition between mostly class (1) to

mostly class (3) is at 7.5%]. The feedback algorithm correctly decided this level was still too high and correspondingly increased the solid-angle–charge product, reaching a final statistical error of 3%.

C. Class (2)

Class (2) cases are expected to undergo major changes of all the experimental conditions, except for the collected charge, which is expected to change only slightly. The distribution of the experimental parameters before and after optimization is shown in Fig. 4, confirming that expectation.

As shown in Table II and Table III, the feedback algorithm is able to increase the number of spectra classified as class (1) from 2.7% to 30.8%. This number seems to be small; however, there are samples for which it is impossible to separate the Ge peak from the Si signal, no matter what the experimental conditions. This is the case when the implanted depth is very large. The exact depth depends on the Ge dose and signal shape, and on the energy resolution (FWHM) of the detected system. The most favorable case, for which separation is easiest, is that of no energy dispersion and a delta-shaped small Ge dose. In that case, the maximum depth for which the signals can be separated is 4.4×10^{18} atoms/cm^2. For a finite FWHM and a Gaussian-shaped implant, that number can be as low as 3×10^{18} atoms/cm^2. The vast majority of the spectra that remain classified as class (2) after optimization have a depth larger than 3×10^{18} atoms/cm^2, and indeed belong to class (2).

The average and standard deviation of the ratio between the determined depth and dose values (with the ANN for the initial and optimized experimental conditions and by NDF), and the real values, are shown in Table IV. We show the values separately for correct and incorrect initial classification. Further, we also separate those cases in which the final class was (1), and those in which the final spectrum remains
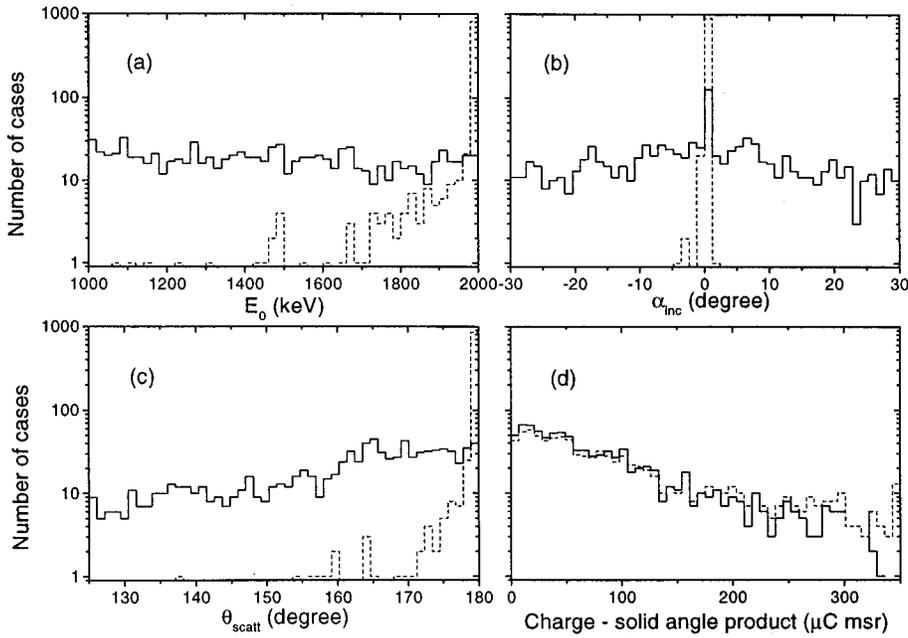
FIG. 4. Distribution of the experimental parameters of class (2) spectra before and after optimization.

class (2) (the majority) or even class (3) (a very small number).

We shall now analyze in detail the different possibilities presented in Table IV.

For cases that were initially well classified by ANNC as belonging to class (2) [i.e., $(ANNC)_i$ correct classification], and for which the final spectrum was class (1), there is a difference in the values of $(ANN)_i$ and $(ANN)_f$. This is because $(ANN)_i$ uses mostly ANN2 (because the corre-

sponding weight w_2 dominates), and $(ANN)_f$ uses mostly ANN1.

For the $(ANNC)_i$ wrong classification and final class (1), the optimization leads to modest improvements only, with the $(ANN)_f$ results being rather poor. This is due to a misclassification problem, as shown in Fig. 5. The average depth and dose are skewed to small values by a few cases that remain class (2) after optimization but are wrongly classified as class (1) due to the relatively small doses. In almost all
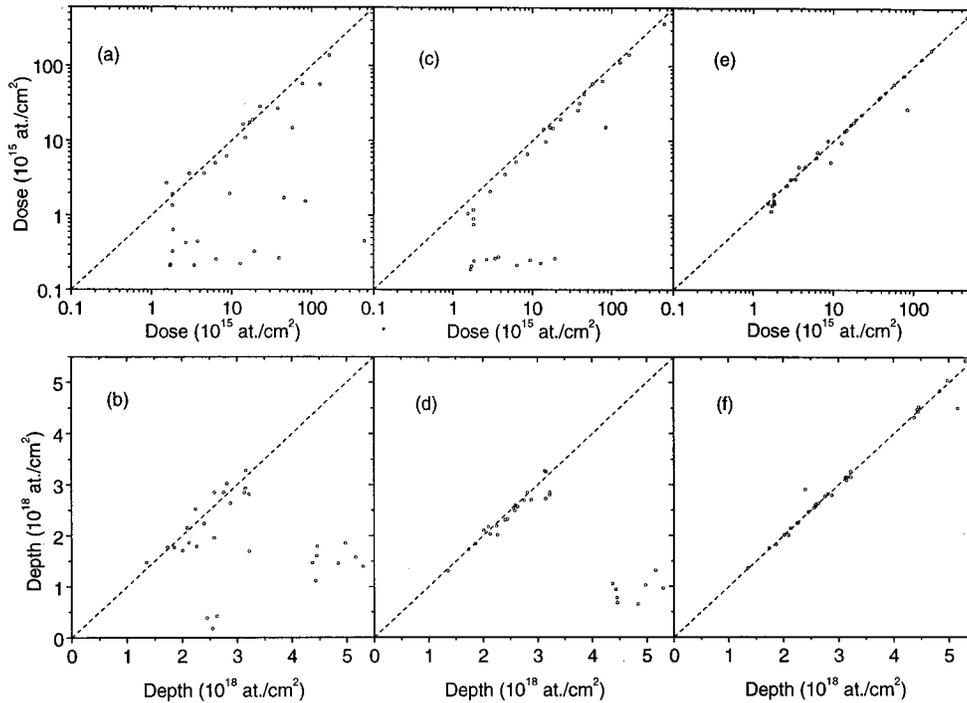

FIG. 5. ANN and NDF results for class (2) spectra initially wrongly classified as class (1) or (3) and with final classification (1). (a) and (b), initial ANN values. (c) and (d), final ANN values. (e) and (f), NDF values.
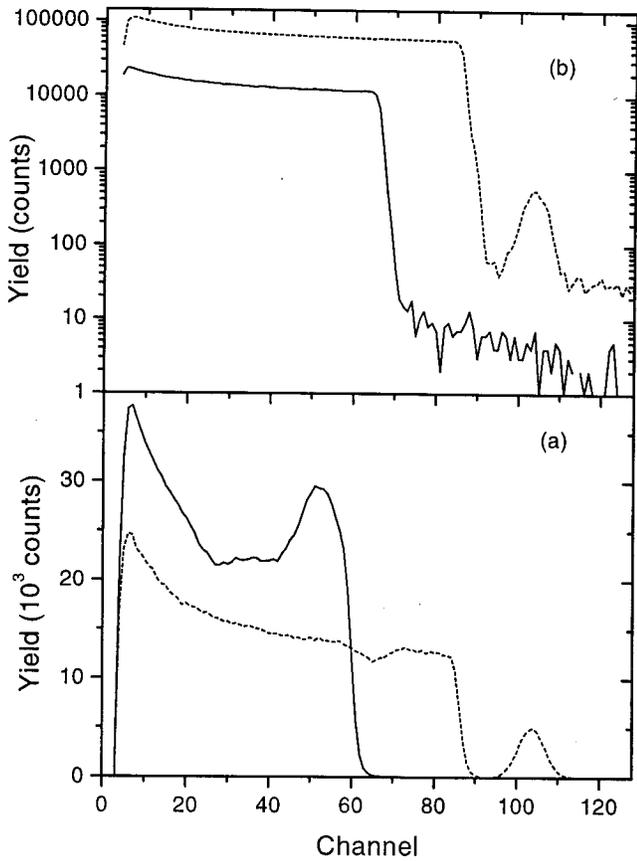
FIG. 6. Class (2) spectra. (a) Spectrum was initially correctly classified as class (2). (b) Spectrum was initially wrongly classified as class (3). In both cases the final class is (1).

depth and dose values, albeit with a larger average error than when the initial classification was correct.

For the $(ANNC)_i$ correct classification and final class (2) or (3) [almost all the cases are class (2)], the results are also good. They become very poor only when $(ANNC)_i$ is incorrect, since a classification error, when the spectra are already difficult to analyze to start with, makes the task even more difficult. NDF can nevertheless provide a reasonable approximation to the correct depth and dose values.

We show in Fig. 6(a) a typical example of a class (2) spectrum that the feedback algorithm is able to lead into class (1), with a large and separated Ge peak. In Fig. 6(b) the initial spectrum was wrongly classified as belonging to class (3), with weights (0, 0.45, 0.55). In the first iteration the charge was duly increased, but the angle of incidence also became closer to normal incidence (0°), which led the weights to (0, 0.84, 0.16). Further optimization resulted in the final spectrum with a completely separated Ge peak, with a statistical error of 4.5%. Note that for the initial charge value, the error would have been (for the same Ge peak separation) 11.6%, which justifies the initial ambiguous classification.

D. Class (3)

Class (3) cases are expected to undergo major changes in the collected charge, which should increase in order to improve statistics, and the other experimental conditions should change only little. The distribution of the experimental parameters before and after optimization is shown in Fig. 7, confirming that expectation.

The average and standard deviation of the ratio between the determined depth and dose values (with the ANN for the initial and optimized experimental conditions and by NDF), and the real values, are shown in Table IV. We show the values separately for correct and incorrect initial classification. Further, we also separate those cases in which the final

those cases, ANNC passed through an intermediate phase during optimization where the spectra were classified as class (3), then increased the charge, and finally misclassified them as class (1). NDF was nevertheless able to derive the correct
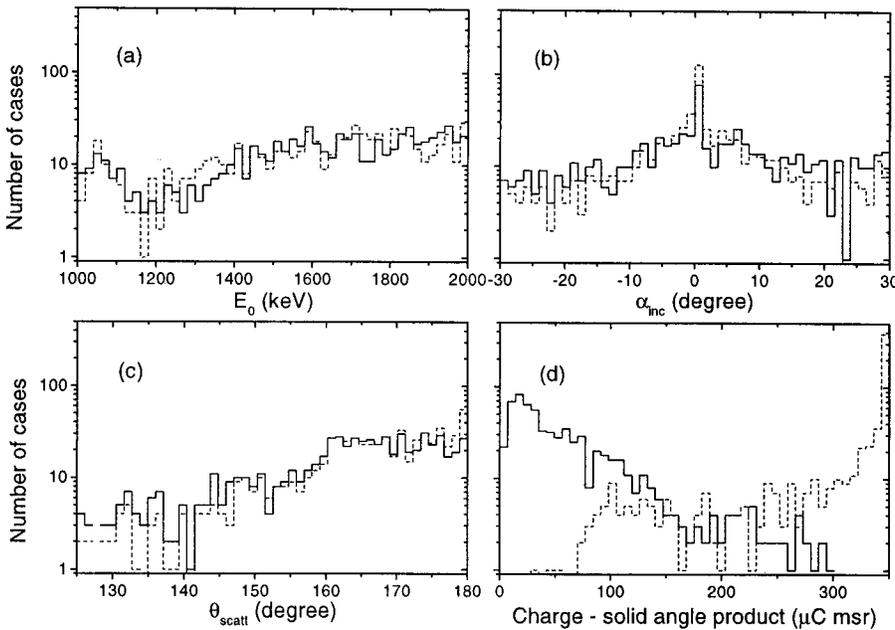

FIG. 7. Distribution of the experimental parameters of class (3) spectra before and after optimization.
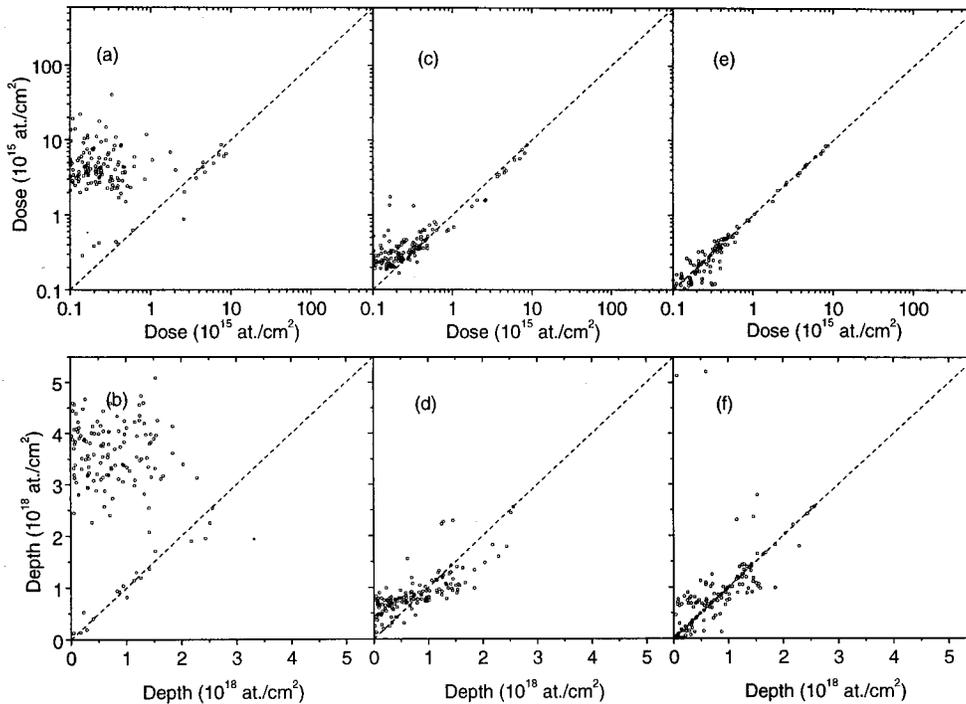
FIG. 8. ANN and NDF results for class (3) spectra initially wrongly classified as class (1) or (2) and with final classification (1). (a) and (b), initial ANN values. (c) and (d), final ANN values. (e) and (f), NDF values.

class was (1) from those in which the final spectrum remains class (3) or even class (2).

For the $(ANNC)_i$ correct classification and final class (1), the improvement in the results is basically due to the transition from the ANN values being initially mainly calculated with ANN3, which is not very precise since it is the hardest class to analyze by any means, and the final ones mainly with ANN1.

For the $(ANNC)_i$ wrong classification and final class (1), the initial results are extremely bad, and optimization leads to dramatic improvements. This is due to the cases that are initially misclassified as class (2), with consequent large values of the dose and depth as can be seen in Fig. 8.

For the $(ANNC)_i$ correct classification and final class (2) or (3), the feedback procedure actually leads to worse results. The reason is that while it was impossible to lead the initially class (3) spectra into class (1), the change in experimental conditions led several of the spectra closer to class (2). While a decrease of beam energy E_0 leads to improved statistics, it also brings the Ge and Si signals closer together. NDF was nevertheless able to find reasonable values for the depth and dose, albeit with a rather large error. For the $(ANNC)_i$ wrong classification, the results are bad, as the initial wrong classification leads to essentially wrong actions being taken by the feedback algorithm, for spectra that are extremely difficult to analyze to start with.

We show in Figure 9(a) a case in which the initial weights are (0, 0.36, 0.64), that is, the initial classification was correct. The statistical error in the Ge peak is 10.8%. After increasing the collected charge and reducing the beam energy, the class weights are (0.87, 0, 0.13) corresponding to class (1), for a statistical error of 3.2%.
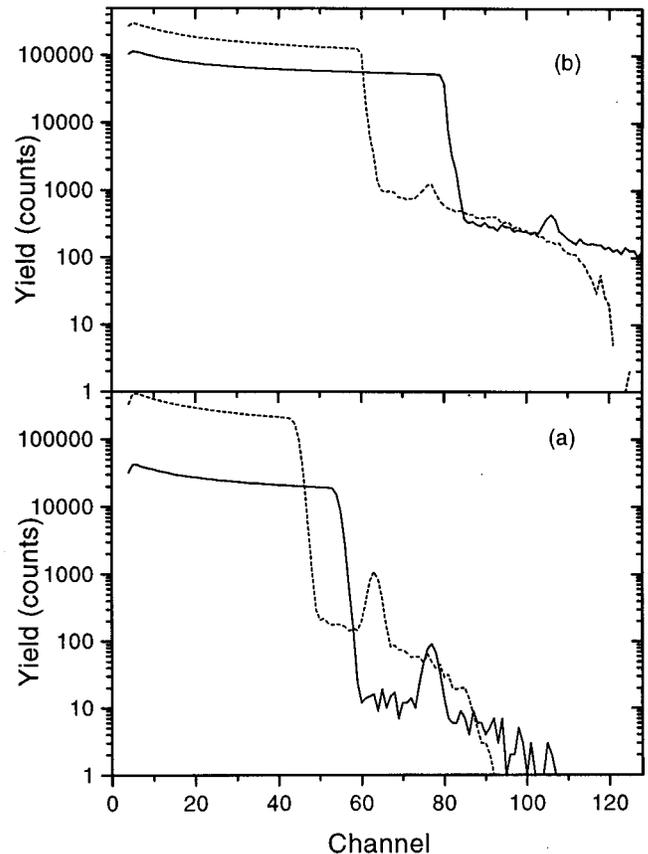

FIG. 9. Class (3) spectra. In both cases the spectrum was initially correctly classified as class (3). The final class is (a) class (1), and (b) class (3).

On the contrary, the feedback algorithm could not lead the spectrum shown in Fig. 9(b) into class (1), even increasing the collected charge to the maximum allowed value and decreasing the beam energy until the Ge peak is very close to the Si signal. The reason is threefold: (a) the implanted dose is very low at 0.3×10^{15} atoms/cm^2; (b) the pileup background is high; and (c), the implanted depth is relatively large at 1320×10^{15} atoms/cm^2, which prevents the feedback algorithm from decreasing the beam energy further, for a shallower depth the signal separation would be larger, giving more room for further beam energy decreases.

V. CONCLUSIONS

We presented an algorithm based on artificial neural networks able to determine optimized experimental conditions for RBS measurements of Ge-implanted Si. The algorithm can be implemented for any other element implanted into a lighter substrate. We have previously shown that ANNs are successful in analyzing data from more complex systems, such as multielemental thin films [11] or implants of any element into a multielemental substrate [20]. It is hence foreseeable that the method developed in this work can be extended to many other systems.

The algorithm presented is a push-button black box, and does not require any intervention from humans. It is thus suited for unsupervised automated control of an experimental setup, given an interface to the relevant hardware. The presence of an experimentalist during the measurements is hence no longer required.

Once the experimental conditions are optimized, the algorithm analyzes the final data obtained, and outputs the relevant desired parameters in a nearly instantaneous way, as opposed to codes that are interactive or that take some time (often long) to run. It is thus suited for unsupervised on-line automatic analysis of the data. A trained physicist or data analyst is hence also no longer required.

For each new system to be analyzed, the required ANNs must be developed and created, which involves a great deal of work by computer programmers with knowledge of IBA.

This could, however, be circumvented if an automatic procedure to generate the required neural networks can be devised. The following scheme is such a procedure.

Suppose that a large batch of samples is to be analyzed. For the first sample, a spectrum would be measured with standard experimental conditions, and then automatically analyzed with NDF. The depth profile determined with NDF may be inaccurate if no previous knowledge of the samples is available. However, a number of different depth profiles, all random variations of the derived one, could be automatically created. For each such profile, a number of theoretical spectra with different experimental conditions can be generated.

In this way, training and test sets can be automatically generated for an ANN dedicated to any given problem. The outputs of the ANN will be the required parameters. These can be, for instance, the implanted depth and dose of a given ion in a given target, or the thickness and composition of a thin film. If one wishes to automatically optimize the experiment, the experimental conditions can be taken as those that lead to the smallest ANN error for the parameters of interest.

Note that these ANNs would have been created in a fully automatic process. They can be used to reanalyze all the spectra previously measured, optimize the experimental conditions and measure them again if necessary, and then take over the whole process for the following samples.

We suggest that all the steps involved can be made automatically. Once the original computer program is written, computer programmers will also no longer be required. Thus could RBS be done without experimentalists, data analysts, or programmers, with only technicians to keep the machines running.

ACKNOWLEDGMENTS

[1] *Handbook of Modern IBA*, edited by J. R. Tesmer and M. Nastasi (MRS, Pittsburgh, 1995).

[2] N. P. Seah, D. David, J. A. Davies, C. Jeynes, C. Ortega, C. Sofield, and G. Weber, Nucl. Instrum. Methods Phys. Res. B **30**, 140 (1988).

[3] U. Wätjen, H. Bax, and P. Rietveld, Surf. Interface Anal. **19**, 253 (1992).

[4] K. H. Ecker, A. Berger, R. Grötzschel, L. Persson, and U. Wätjen, Nucl. Instrum. Methods Phys. Res. B **175-177**, 797 (2001).

[5] E. Kótai, Nucl. Instrum. Methods Phys. Res. B **85**, 588 (1994).

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

[7] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983).

[8] Emile Aarts and Jan Korst, *Simulated Annealing and Boltz-* *mann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing* (Wiley, Chichester, 1989).

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

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

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

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

[13] C. Jeynes, Z. H. Jafri, R. P. Webb, A. C. Kimber, and M. J. Ashwin, Surf. Interface Anal. **25**, 254 (1997).

[14] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, edited by D. E. Rumelharb and J. L. McClelland (MIT Press, Cambridge, MA, 1986), Vol. 1, p. 318.

[15] H. White, Neural Comput. **1**, 425 (1989).

[16] C. M. Bishop, IEE Proc. Vision Image Signal Process. **141**, 217 (1994).

[17] D. E. Rumelhart, R. Durbin, R. Golden, and Y. Chauvin, in *Backpropagation: Theory, Architectures, and Applications*, edited by D. E. Rumelharb and Y. Chauvin (Lawrence Erlbaum, Hillsdale, NY, 1995), pp. 1–34.

[18] P. Baldi and K. Hornik, Neural Networks **2**, 53 (1989).

[19] W.-K. Chu, J. W. Mayer, and M.-A. Nicolet, *Backscattering Spectrometry* (California Institute of Technology, Pasadena, 1978), pp. 25, 29, 62.

[20] Analysis of sapphire implanted with different elements using artificial neural networks, A. Vieria, N. P. Barradas, and E. Alves, Nucl. Instrum. Methods Phys. Res. B **190**, 241 (2002).