Simulated annealing analysis of Rutherford backscattering data

N. P. Barradas,[a] C. Jeynes, and R. P. Webb
School of Electronic Engineering, Information Technology and Mathematics, University of Surrey, Guildford, Surrey GU2 5XH, United Kingdom

(Received 24 March 1997; accepted for publication 14 May 1997)

The combinatorial optimization simulated annealing algorithm is applied to the analysis of Rutherford backscattering data. The analysis is fully automatic, i.e., it does not require time-consuming human intervention. The algorithm is tested on a complex iron-cobalt silicide spectrum, and all the relevant features are successfully determined. The total analysis time using a PC 486 processor running at 100 MHz is comparable to the data collection time, which opens the way for on-line automatic analysis. © *1997 American Institute of Physics.* [S0003-6951(97)04528-2]

Rutherford backscattering (RBS) is a standard technique in the analysis of materials, and is used to determine what are the elements present in a given sample, their stoichiometry, and their depth distribution.[1] Its main advantages are that it is fully quantitative, i.e., the use of external standards is not necessary, and that a precision better than 1% can be achieved with careful analysis (see e.g., Ref. 2). Computer-aided data interpretation has been common practice for over two decades,[3] and several programs have been developed to facilitate the task of analyzing the data, the best known being RUMP.[4] While a recent review can be found in Ref. 5, new programs are still regularly developed (see e.g., Ref. 6). A common trait of the codes reported is that they all require a considerable degree of input from the user, who must normally define the depth profile of all the elements, and then compare a theoretical function generated from that layer structure with the data. Although there is research in new directions (e.g., using maximum entropy methods[7]), there is at the moment no general algorithm available for the solution of the inverse RBS problem, i.e., to calculate from the data the depth profile of the elements present.

The aim of this letter is to present a computer program that performs automatic analysis of RBS data without the need of time-costly human involvement. The only inputs necessary are the experimental conditions in which the experiment was done, and the elements present in the sample. This is in many cases known *a priori*, and takes only a few seconds to feed into the computer. The combinatorial optimization simulated annealing[8] algorithm was used, due to its two main features: first, the solution is independent of the initial guess chosen, and therefore a human-input initial layer structure is not needed. Second, it tends asymptotically to the absolute minimum rather than to a local minimum as in conventional minimization algorithms, and hence high quality solutions can be achieved. It has been applied in fields as diverse as very large scale integrated (VLSI) design,[9] Biophysics,[10] or Natural Language Processing.[11] The inverse RBS problem can be defined as a discrete combinatorial optimization problem; the solution space S containing all possible RBS spectra (ignoring the differences in the experimental setups, which are not relevant) is indeed finite, due to the finite sensitivity and energy resolution of the technique. If s_i is the sensitivity for element i, only $1/s_i$ different concentrations of element i in a given layer are distinguishable. Further, an energy resolution of 1% limits the number of separable layers n_1 to about 100, and the size of the solution space is then limited by $\#S < n_1^{\Pi_i 1/s_i}$.

Simulated annealing is based on an analogy with annealing, i.e., removing defects from a crystal by melting it and subsequently cooling it down very slowly. If enough time is allowed for the crystal to reach thermodynamical equilibrium at each temperature, and if the cooling is slow enough, then at $T=0$ K the crystal will be in a state of minimum energy. This is opposed to quenching, where a fast rate of cooling ensures that the crystal will freeze in a highly defective metastable state. The analogy with simulated annealing is as follows: the possible structures of the sample (i.e., number of layers, layer thickness, and stoichiometry) are equivalent to states of a physical system; the chi-square (χ^2, defined in the usual way as a weighted distance between the proposed theoretical function and the data) is equivalent to the energy of a state; and a control parameter T is equivalent to the temperature. Further, a transition of state is defined as the generation of a new structure from the previously calculated one (e.g., by increasing or decreasing the number of layers, or altering the layers thickness and stoichiometry by some amount). The transitions are generated randomly, as is the case in a real material. The probability P of accepting a transition is

$$P = 1 \quad \text{if } \Delta\chi^2 < 0$$
$$= \exp(-\Delta\chi^2/T) \quad \text{if } \Delta\chi^2 > 0. \tag{1}$$

It is the second condition that allows escape from local minima of the χ^2. Clearly, at sufficiently high values of the control parameter T, practically all the transitions are accepted, corresponding in the analogy to a liquid state with high entropy. As T decreases, the probability of transitions with a related high increase of the χ^2 becomes smaller, and at very small values of T only transitions that lead to a χ^2 decrease are accepted. The main elements of the program are:[8]

(a) A cooling schedule, i.e., the initial value T_0 of the control parameter, the number L of proposed transitions at

[a]Permanent address: Centro de Física Nuclear da Universidade de Lisboa, Av. Prof. Gama Pinto 2, 1699 Lisboa Codex, Portugal. Electronic mail: n.barradas@surrey.ac.uk

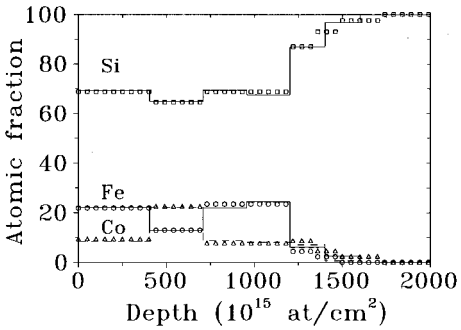

FIG. 1. The depth profile used to generate the theoretical spectrum fitted is shown as the symbols (squares: Si; circles: Fe; triangles: Co). The final depth profile as determined by the simulated annealing fit is shown as lines (solid: Fe; dashed: Co).

each value of T, and the rate of cooling. T_0 is determined by calculating $\langle \Delta\chi^2 \rangle$, i.e., the average chi-square increase for χ^2 increasing transitions, and requiring that the probability of the average transition be superior to a certain high value P_0, typically around 0.95. This is expressed as $T_0 = -\langle \Delta\chi^2 \rangle / \ln(P_0)$. The number of proposed transitions should be large enough to ensure that equilibrium is reached, i.e., at least a certain number of them should be accepted at each value of T. The cooling rate is defined by $T_{i+1} = kT_i$, where k is a positive constant smaller than 1.

(b) A transition of state procedure, i.e., to calculate a new state from the current one. For RBS, this amounts to varying randomly the thickness and stoichiometry of the layers, as well as creating randomly new layers and merging existing ones. Finally, the program tests only the elements specified by the user.

(c) The calculation of the χ^2 value. This amounts basically to calculating a theoretical RBS spectrum $Y_{\text{the}}(E)$ corresponding to the current layer structure, to then compare it with the data $Y_{\text{exp}}(E)$, where E is the energy of the backscattered particles. In this implementation, several simplifications were done on the generation of the RBS spectra, in order to reduce the calculation time. The 512 channels of the experimental data were compressed by adding four channels each; the energy straggling is not taken into account, which leads to overestimation of the thickness of any interface present; and the average mass of the elements is used instead of the full isotopic distribution. The stopping cross sections are calculated, as usual, from the tabulated elemental ones[12] using Bragg's rule. Further, there is no simple way of calculating the effect of plural and multiple scattering which would distort the spectral shape at low E values,[13] and hence it was not included in the calculations. The influence of pileup was fully taken into account following the procedure given in Ref. 2.

(d) The acceptance criterion given by Eq. (1). The probability P of a transition is weighted against a random number r between 0 and 1. If $P \geqslant r$, the transition is accepted.

The algorithm was tested by first generating a theoretical spectrum from a given layer structure. The algorithm is then applied to the inverse problem, i.e., from the spectrum it attempts to determine the layer structure. A multilayered iron-cobalt silicide test structure, shown in Fig. 1, was chosen because, first, to separate the neighbor Fe and Co elements is within the limits of the RBS technique, and second, it is a system of actual interest.[14] The resulting spectrum is shown in Fig. 2, and the superposition of the Co and Fe signals leads to a complex structure. Random statistical error was added to the original theoretical spectrum in order to test the robustness of the solution found by the algorithm.

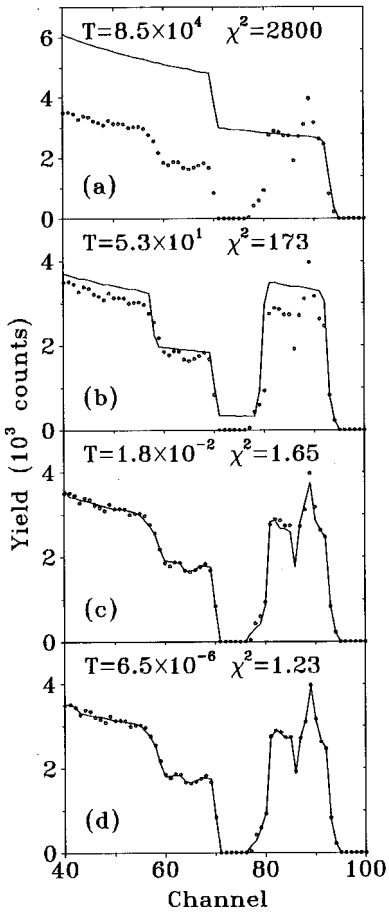

FIG. 2. Different steps of the fit, from (a) the initial guess to (d) the final result. The values of the control parameter T and χ^2 at each step are shown. On real data, $\chi^2 = 1$ would mean that the fit is within the statistical error in all data points.

Figure 2(a) shows the initial guess used in the simulated annealing fit. It assumes simply a single homogeneous layer of Si, Fe, and Co. The initial stoichiometry was determined from the height of the detected edges; as the Fe and Co edges are not separated, equal initial concentrations are assigned to them. The cooling schedule was defined by $T_0 = 8.5 \times 10^4$, $L = 2200$, and $k = 0.567$. These values were determined automatically by the program; T_0 as given above, and L and k from an empirical procedure optimized by trial and error on a number of real spectra with different characteristics. They depend on the number of elements in the sample and on the number of edges and peaks in the spectrum, so that increased data complexity leads to slower cooling.

After reducing the value of T to 5.3×10^1, the best solution obtained so far, shown in Fig. 2(b), already consists of a silicide layer with the correct thickness, on top of bulk Si. The stoichiometry of the silicide layer is still approximately homogeneous, and the smooth interface between the silicide

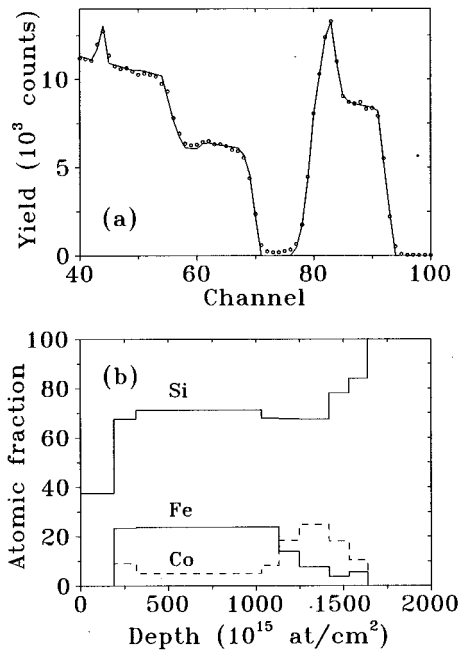

FIG. 3. (a) Final result of a simulated annealing fit of an Fe–Co silicide, including a surface oxide. (b) Depth profile determined from the fit. The O profile is not shown, for clarity. The depression in the Si concentration between about 1000 and 1500×10^{15}/cm^2 corresponds to a stoichiometric cubic silicide, while below about 1000×10^{15}/cm^2 the layer has a stoichiometry similar to the Si-rich α-$FeSi_2$ phase.

and the substrate is not reproduced. This is the point where traditional minimization algorithms would stop (as does the one implemented in RUMP,[4] after some further optimization of the exact silicide thickness and composition), because a local minimum has been reached, corresponding to quenching in the analogy with annealing. In fact, a large reduction of the χ^2 has already taken place, or using the analogy with annealing, a low-energy, low-entropy, highly probable solution has already been found. However, the simulated annealing algorithm does not remain trapped in this metastable state; instead, it proceeds to find the real minimum, and at $T=1.8\times10^{-2}$ the solution found has the correct shape, that is due to the varying Fe and Co composition in the silicide. Then it adjusts the details, reproducing the features of the data with a smaller influence on the χ^2 value, and the final solution is reached at $T=6.5\times10^{-6}$. The total number of solutions tested was 9×10^4. A very good fit is obtained. Aarts and Korst[8] show that when alternative traditional methods can solve a particular minimization problem, they generally perform better than simulated annealing. Those methods, however, cannot be used to solve the general inverse RBS problem, due to the different, overlapping, depth scales of the different elements, leading to the existence of strong local minima.

The final sample structure (i.e., the final state) determined is shown in Fig. 1, and it agrees remarkably well with the original one, which demonstrates the power of the simulated annealing algorithm. All the features are correctly determined. Given infinite time, the algorithm would find the best possible solution, as opposed to the high quality solution found. The whole procedure took 6 min 19 s on a 486 processor running at 100 MHz. This is comparable to the data collection time, which means that fully automated on-line analysis of RBS data becomes possible.

The successful use of the algorithm to recover the original structure from a theoretical test spectrum shows that a solution to the inverse RBS problem exists. Errors in the spectrum simulator due to multiple scattering or inaccurate stopping powers will lead to errors in the solution obtained. These problems, avoided by the use of simulated input data, affect all RBS data analysis and are outside the scope of this letter. However, real data are also successfully analyzed by our algorithm, including the data shown in Figs. 5, 9, 10, and 11 of Ref. 14. As an example, Fig. 3 shows the analysis of Fig. 9 of Ref. 14. This is actually a harder case since a surface oxide is present, but an excellent fit is obtained. The depth profiles derived by simulated annealing agree in all cases remarkably well with the x-ray photoemission spectroscopy (XPS) and cross-section transmission electron microscopy (XTEM) results presented by Harry *et al.* The simulated annealing algorithm has also been successfully used in the study of many other systems, including O implanted SiC and the ion beam mixing of Fe/Si and Ta/Si with As and Xe implantation.

[1] J. R. Tesmer and M. Nastasi, *Handbook of Modern Ion Beam Materials Analysis* (MRS, Pittsburgh, 1995).
[2] C. Jeynes, Z. H. Jafri, R. P. Webb, A. C. Kimber, and M. J. Ashwin, Surf. Interface Anal. **25**, 254 (1997).
[3] J. F. Ziegler and J. E. E. Baglin, J. Appl. Phys. **42**, 2031 (1971).
[4] L. R. Doolittle, Nucl. Instrum. Methods Phys. Res. B **9**, 291 (1985).
[5] E. Kotái, Nucl. Instrum. Methods Phys. Res. B **85**, 588 (1994).
[6] B. Rajchel, Nucl. Instrum. Methods Phys. Res. B **113**, 300 (1996).
[7] V. M. Prozesky and J. Padayachee, presented at the 14th International Conference on the Application of Accelerators in Research and Industry, Denton, USA, 1996 (unpublished).
[8] Emile Aarts and Jan Korst, *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing* (Wiley, Chichester, 1989).
[9] H. Murata, K. Fujijoshi, S. Nakatake, and Y. Kajitani, IEEE Trans. Comput.-Aided Des. **15**, 1518 (1996).
[10] Z. Qi and M. Sokabe, Biophys. J. **71**, TH300 (1997).
[11] Y. A. Wilks, B. M. Slator, and L. M. Guthrie, *Electric Words: Dictionaries, Computers, and Meanings* (MIT Press, Cambridge, 1996).
[12] J. F. Ziegler, J. P. Biersack, and U. Littmark, *Stopping and Ranges of Ions in Solids* (Pergamon, New York, 1985).
[13] P. Bauer, E. Steinbauer, and J. P. Biersack, Nucl. Instrum. Methods Phys. Res. B **64**, 711 (1992).
[14] M. A. Harry, G. Curelli, M. S. Finney, K. J. Reeson, and B. J. Sealy, J. Phys. D **29**, 1822 (1996).