

High-depth-resolution Rutherford backscattering data and error analysis of SiGe systems using the simulated annealing and Markov chain Monte Carlo algorithms

N. P. Barradas, A. P. Knights, and C. Jeynes

School of Electronic Engineering, Information Technology and Mathematics, University of Surrey, Guildford, Surrey GU2 5XH, United Kingdom

O. A. Mironov,* T. J. Grasby, and E. H. C. Parker

Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom

(Received 28 September 1998)

The thickness and composition of SiGe/Si quantum wells has been studied by high-depth-resolution Rutherford backscattering (RBS) analysis. While a depth resolution at the surface of 8 nm is achieved, the degradation of resolution with depth must be taken into account for a correct analysis of the data. Fully automated analysis incorporating the depth resolution as a function of depth was performed using the simulated annealing algorithm. Bayesian inference using the Markov chain Monte Carlo method was also employed, and confidence limits on the SiGe depth profiles obtained from the RBS data are established. The results obtained are used to show that the interface sharpness in SiGe/Si multilayers depends on the temperature of growth. Finally, it is shown that standard RBS using a normal angle of incidence is sufficient to obtain narrow limits of confidence in the thickness and stoichiometry of single quantum wells, but not of multiple quantum wells, for which high-depth-resolution experiments are necessary. [S0163-1829(99)01107-8]

I. INTRODUCTION

Si-Si_{1-x}Ge_x multiple quantum wells have many possible applications in optoelectronics, with the potential for monolithic integration with Si technology.¹ These include infrared photodetectors based on a number of *p*-type strained Si_{1-x}Ge_x wells with thicknesses around 3–10 nm, separated by Si barriers with thicknesses around 30–50 nm, capped with *p*-type Si for electrical contact.² The layer quality, stoichiometry, thickness, and interface sharpness have a strong influence on the optical properties of Si/SiGe systems, and, therefore, their careful assessment is necessary.

Rutherford backscattering spectrometry (RBS) is a well-established and powerful technique for determining the elemental composition of thin films using an energetic light ion beam, normally H or He, in the 1–2-MeV range. Because the data analysis is based on Newtonian mechanics, it is fully quantitative without the need to recur to standards (see, e.g., Ref. 3). RBS can be therefore used to obtain quantitative and traceable information about homogeneous films. It has been used in a round robin exercise organized by the National Physical Laboratory⁴ to characterize the 100-nm Ta₂O₅ standard for electron spectroscopy, and also to quantify a secondary implanted standard by the Institute for Reference Materials and Measurements in Geel.⁵ Additionally we recently demonstrated 1% accuracy in the determination of the In content of In_xGa_{1-x}As films.⁶

However, most samples of interest, such as the ones studied in this work, are not homogeneous. In such a case it is usually difficult to devise a computational method transparent enough to obtain depth profiles from the spectrum whose accuracy is traceable. Moreover, in many cases the spectra obtained from real samples are sufficiently complex to preclude manual extraction of accurate depth profiles in a reasonable time. Addressing these problems, we have recently

demonstrated that the general inverse RBS problem (obtaining the depth profile from the spectrum) can be efficiently solved using a simulated annealing computer algorithm.⁷

Another relevant problem is that the study of samples containing 3–10-nm-thick SiGe layers is limited by the depth resolution of the RBS technique. The problem is still more severe when several layers are present, and the individual thickness and composition of each is to be determined. In standard RBS, a H⁺ or He⁺ beam in the 1–2-MeV energy range is employed, which strikes in the sample normal to its surface. Backscattered nuclei are detected with a solid-state detector. The depth resolution is generally not good enough to study systems such as those of interest here. An alternative to the use of expensive magnetic spectrometers⁸ is to perform RBS at the grazing angle of incidence, in which the increased path length of the beam in each layer leads to an enhancement of the depth resolution.⁹ However, the depth resolution degrades as the beam enters deeper into the sample, due to different processes such as the statistical nature of the process of energy loss through collision with the electrons (energy straggling), or to multiple small-angle collisions with the sample nuclei (multiple scattering). These effects become more important as the angle of incidence becomes shallower, and when analyzing the results of grazing angle of incidence experiments they must be accounted for, otherwise the composition of the layers will be incorrectly evaluated. Szilágyi, Pászti, and Amsel presented a computer code, DEPTH,¹⁰ that calculates the depth resolution as function of depth with an error smaller than 10% by taking into account the detection system, the physical properties of the sample, and the physical effects that lead to degradation of depth resolution.

In this paper, we present a fully automated analysis, with the simulated annealing algorithm of SiGe quantum wells using high-resolution RBS, accounting for the correct values

of the depth resolution as a function of depth. Furthermore, we extend this work to show how the computational method is able to generate information about the errors involved in the depth profiles obtained, by applying Bayesian inference using the Markov chain Monte Carlo integration method. This allows us not only to derive the SiGe layer thickness and stoichiometry automatically from the RBS data, but also to establish confidence limits on the solutions obtained.

II. EXPERIMENTAL DETAILS

Samples with a nominal composition Si(001)/Si 300 nm/ (Si 30 nm/Si_{0.78}Ge_{0.22} 5 nm)₅/Si 50 nm were grown by solid source molecular-beam epitaxy, at different substrate temperatures ranging from 550 to 810 °C. A Si(001)/Si_{0.7}Ge_{0.3} 7-nm/Si 70-nm sample was also produced. The experimental details are described in detail in Ref. 11. The thickness of the SiGe layers was much below the equilibrium critical thickness for the substrate temperatures used, and cross-section transmission electron microscopy results show no strain-induced dislocations in the epitaxial region. The samples were fully-strained as verified by measurements of x-ray rocking curves and Raman spectra.¹²

Grazing angle of incidence RBS experiments were performed using a 1.5-MeV He⁺ beam on the Sacavém 3-MV van de Graaff accelerator.¹³ A surface barrier detector with 20-keV energy resolution was located under the beam at 160° to the beam direction (Cornell geometry). For each sample, spectra were collected from normal incidence to an 11° grazing incidence tilt angle ϑ , defined as the angle between the beam direction and the sample surface with a vertical tilt axis. The beam was 0.2–0.4 mm wide and 0.6 mm high. A 1.5×5.0-mm² detector aperture was used to reduce geometrical dispersion. The detector-sample distance was 75 mm. The 5-nA beam current was measured with a transmission Faraday cup with precision $\approx 2\%$.¹⁴ The pressure during the experiments was 8×10^{-8} mbar. Standard RBS was also done using the Surrey 2-MV van de Graaff.¹⁵ The beam was circular with a 1-mm diameter. Two detectors, A and B, were used simultaneously. Detector A was located at a 165° scattering angle in the same plane as the beam and the normal to the samples (IBM geometry), and had a 16-keV full width at half maximum (FWHM) and 6.5 msr solid angle. Detector B was located at a 133° scattering angle in the IBM geometry, and had 14-keV FWHM and 4.3-msr solid angle. Both detectors were circular and no aperture was used. The energy calibration was determined with a 0.5% error in the gain and a three-channel error in the offset, and the beam fluence was measured with a 2% precision.¹⁵ Complementary x-ray diffraction (XRD), transmission electron microscopy, and secondary-ion-mass-spectrometry (SIMS) measurements were also carried out on some samples.

III. SIMULATED ANNEALING AND MARKOV CHAIN MONTE CARLO ALGORITHMS

Simulated annealing (SA) is a global optimization algorithm designed to find the absolute minimum (or maximum) of any given function.^{16–18} It is completely general in the sense that it entails in principle no restrictions on the function to be minimized. It has solved previously intractable

problems such as the NP-complete (i.e., insoluble in polynomial time) traveling salesman problem,¹⁹ and is widely applied in different fields, from ion-beam analysis to natural language processing.²⁰ SA is based on an analogy with annealing, i.e., removing defects from a crystal by melting it and subsequently cooling it down very slowly. Given a current state i of the system with energy E_i , another state j with energy E_j is generated by a random process, in which state i is slightly changed. If the energy difference $\Delta E = E_j - E_i$ is negative, the system is taken to be in the new state j . If ΔE is positive, that is, if the energy of the system increases, the transition can still occur with a certain probability given by the Boltzmann factor

$$P(i \rightarrow j) = \exp(-\Delta E/k_B T), \quad (1)$$

where T is the temperature of the system and k_B is the Boltzmann constant. If at each temperature time enough is allowed for the crystal to reach thermodynamical equilibrium, and if the cooling is slow enough, then at $T = 0$ K the crystal is guaranteed to be in a state of minimum energy. Note that if the system is degenerate, that is, if more than one state corresponds to the minimum energy, at $T = 0$ K the system will be in any one of those states.

In the analogy with simulated annealing, E becomes any objective function $f \equiv f(\mathbf{x})$ to be minimized, where the state of the system is defined as $\mathbf{x}$. In ion-beam analysis, $\mathbf{x}$ is the depth profile of all elements. A state transition is defined as the generation of a new structure given the previously calculated one by randomly changing the depth profiles by some amount. The probability P of accepting a transition from state $\mathbf{x}$ to state $\mathbf{y}$ is given by the so-called Metropolis criterion²¹

$$P(\mathbf{x} \rightarrow \mathbf{y}) = \min\{\exp(-\Delta\chi^2/T), 1\}, \quad (2)$$

where $f \equiv \Delta\chi^2$ is the χ^2 change due to the transition, and T is a control parameter (not a temperature). At 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 related high increase of χ^2 becomes smaller, and at very small values of T only transitions that lead to a decrease in χ^2 are accepted. In SA, T is initialized at some high value T_0 that allows practically all transitions to be accepted. Then T is decreased slowly, according to $T_{i+1} = kT_i$, where k is a positive constant smaller than 1. At each value of T , L_M transitions are proposed; the succession of all the accepted states is called a Markov chain, so during the SA process several Markov chains are computed, one for each value of the control parameter. The values of T_0 , k , and L_M define what is called the cooling schedule. For sufficiently high values of these three parameters, it can be mathematically proven¹⁶ that the absolute minimum of the objective function is found. In practical terms, that would lead to extremely long calculation times, and a reasonable cooling schedule, that leads to a high-quality solution as opposed to the best one, must be used.

We have previously described the implementation of the simulated annealing algorithm to RBS, elastic (non-Rutherford) backscattering^{7,22} and elastic recoil detection analysis,²³ and demonstrated its application in complex prac-

tical cases that would have been very hard to analyze otherwise.^{24–26} The resulting computer program, NDF,²⁷ is to the best of our knowledge currently the only tool that allows automatic extraction of depth profiles from raw data. Other approaches, such as the Bayesian inference and maximum entropy methods of Fischer and co-workers,^{28–30} have been extremely successful in improving the depth resolution of RBS by effectively deconvoluting the apparatus response function from the data, but they have not applied it to extract depth profiles from data.

Simulated annealing, while very successful in a wide range of problems, has one main shortcoming: it returns one state of the system (i.e., one depth profile) corresponding to the minimum χ^2 found, without any indication of fit error. Further, RBS is many times ambiguous, with more than one depth profile producing the same spectrum,³¹ and as stated above simulated annealing would just find, at random, one of those solutions. In many cases extra information [that can be input in NDF (Ref. 22)] is available to constrain the solution, and many ambiguities can be solved by measuring more than one spectrum under different experimental conditions³² (and NDF can fit simultaneously up to ten spectra collected from the same sample²³). However, it would be highly desirable to be able to calculate both confidence intervals on the final depth profile and also to obtain all the possible solutions.

One possibility is Bayesian Inference (BI).^{33,34} Suppose we wish to have information on $\mathbf{x}$, about which some *a priori* information I exists which can be expressed in terms of the conditional probability $p(\mathbf{x}|I)$, known as the *prior distribution*, i.e., it is the knowledge we have about $\mathbf{x}$ before any experiment is done. Suppose also that we have some experimental observations $d \equiv d(\mathbf{x})$ which depend on the parameters $\mathbf{x}$ in a known way. The knowledge of the dependence of the observations upon $\mathbf{x}$ is then the conditional probability $p(d|\mathbf{x})$, the so-called *likelihood function*. It describes how probable it is to obtain a certain experimental result given well-known parameters; in RBS, this corresponds to calculating a theoretical spectrum from a known depth profile, and adding the apparatus response function and statistical fluctuations. The question becomes, how much do the experimental observations alter our original beliefs about the parameters $\mathbf{x}$? In other words, given the experimental RBS spectra and any *a priori* information, what can we say about the depth profile? The answer is given by Bayes' theorem

$$p(\mathbf{x}|dI) = p(d|\mathbf{x})p(\mathbf{x}|I)/p(d|I). \quad (3)$$

The probability density function $p(\mathbf{x}|dI)$ is called the *posterior distribution*, as it is the knowledge we have about $\mathbf{x}$ taking into account the experiments done and any other information. It therefore contains all the information about the sample given the data. With it, it is possible to calculate the mean solution $\langle \mathbf{x} \rangle$ as well as confidence intervals given by the standard deviation $\sigma(\mathbf{x})$. As $p(d|I)$ is independent of $\mathbf{x}$, we can treat it as a normalization constant. In the general case, any depth profile is possible, so $p(\mathbf{x}|I)$ could be a constant for all atomic functions between 0 and 1, and zero outside that interval. However, in many cases previous information about the system is available from other sources, and this term represents those constraints as imposed by the user in NDF; for instance, the user may specify that in a

SiO₂-based system the maximum O concentration allowed is $\frac{2}{3}$. In this case, it is not convenient to use maximum entropy's uninformative prior.²⁸ Finally, $p(d|\mathbf{x})$ can be taken to depend on the χ^2 , e.g., through

$$p(d|\mathbf{x}) = \exp(-\chi^2/2), \quad (4)$$

with the χ^2 function defined in the usual way:

$$\chi^2 \equiv \sum_i [(Y_{\text{expt}}^i - Y_{\text{the}}^i(\mathbf{x}))/\sigma_i]^2, \quad (5)$$

where Y_{expt}^i and Y_{the}^i are the experimental and theoretical spectra respectively, and $\sigma_i = \text{sqrt}(Y_{\text{expt}}^i)$ is the experimental error in the i th channel of the spectrum. Other sources of error, such as uncertainty in the beam fluence or in the energy calibration, can also be included. The problem now is to evaluate Eq. (3), which we will do with the Markov chain Monte Carlo (MCMC) sampling method, in particular the Metropolis Hastings algorithm²¹ based on the Metropolis criterion given in Eq. (1). Take a Markov chain $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n, \dots, \mathbf{x}_m$, such that $p(\mathbf{x}_i|\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{i-1}) = p(\mathbf{x}_i|\mathbf{x}_{i-1})$, that is, the probability that the i th member of the chain be $\mathbf{x}_i$ depends only on the previous element of the chain, and is determined by the random distribution $q(\mathbf{x}_i, \mathbf{y})$. The Markov chain is then generated by proceeding from $\mathbf{x}_i$ to $\mathbf{x}_{i+1}$ by considering a candidate point $\mathbf{y}$ generated with the random distribution $q(\mathbf{x}_i, \mathbf{y})$. The candidate $\mathbf{y}$ is then accepted (that is, it becomes $\mathbf{x}_{i+1}$) with probability $P(\mathbf{x} \rightarrow \mathbf{y})$ according to the generalized Metropolis criterion

$$P(\mathbf{x} \rightarrow \mathbf{y}) = \min\{[\pi(\mathbf{y})q(\mathbf{y}, \mathbf{x})]/[\pi(\mathbf{x})q(\mathbf{x}, \mathbf{y})], 1\}, \quad (6)$$

where the acceptance function $\pi(\mathbf{x})$ is a function of $\mathbf{x}$ to be chosen, as seen below. After running the Markov chain for a period of n iterations until equilibrium has been reached [that is, until the probability of the system being in the state $\mathbf{x}$ be given by $\pi(\mathbf{x})$], we continue for a further $m - n$ iterations; the Markov chain $\mathbf{x}_n, \dots, \mathbf{x}_m$ then constitutes a sample from the acceptance distribution $\pi(\mathbf{x})$. This sample is an empirical distribution which, if large enough, reflects all the properties of π without need for direct evaluation of π .

The transition distribution q can be chosen such that it is symmetric, i.e., $q(\mathbf{x}, \mathbf{y}) = q(\mathbf{y}, \mathbf{x})$. All that is finally needed is to set $\pi(\mathbf{y}) = p(d|\mathbf{x})p(\mathbf{x}|I)$ as in Eq. (3), and taking Eq. (4) and the fact that $p(\mathbf{x}|I)$ is incorporated in the generation of transitions of state into account, the acceptance criterion to generate the Markov chain becomes

$$P(\mathbf{x} \rightarrow \mathbf{y}) = \min\{\exp(-\Delta\chi^2/2), 1\}. \quad (7)$$

The Markov chain so generated is then a sample of $p(\mathbf{x}|dI)$, that is, it reflects all the information over $\mathbf{x}$ that can be obtained from the experimental data, taking into account any previous information on the system. Not only averages and standard deviations of the concentration of any element at any depth can be calculated, but ambiguous problems can also be conveniently treated: if there is more than one solution that fits the data correctly, $p(\mathbf{x}|dI)$ will be multimodal, that is, it will have one maximum per possible solution.

IV. ENERGY RESOLUTION AS A FUNCTION OF DEPTH

In RBS analysis the energy resolution degrades with depth, faster at grazing angles of incidence. The result is an artificial broadening of any features in the spectra obtained, and to determine the real depth profile it is necessary to know the depth resolution precisely. This is particularly critical when high-resolution experiments are performed by using a grazing angle of incidence, as in this work. In this case, the geometrical straggling and the multiple scattering are the main factors that lead to a fast degradation of the energy resolution with depth.

The computer code DEPTH (Refs. 10 and 35) calculates the depth resolution as a function of depth with a precision better than 10%. It has been validated for many different systems, including pure Si,³⁶ SiGe,¹¹ and Co/Re multilayers.⁹ It takes into account the detector resolution, the energy and angular spread of the beam, the geometric spread caused by the finite beam size and by the detector's solid angle, and energy straggling and multiple scattering in the sample. It does not, however, allow for the small low-energy tails due to plural scattering from high-Z target elements.

Previously,¹¹ we analyzed the RBS data collected for the SiGe/Si multilayers with a cumbersome, although correct, procedure: The RBS spectra were first converted to a set of apparent depth profiles, using Si and SiGe stopping powers and densities.³⁷ A test depth profile with the nominal sample composition was then convoluted with the depth resolution function, as calculated with DEPTH. The different spectra obtained, one for each angle of incidence, were then compared with the corresponding experimental Ge profiles. An automatic fit process followed, where the fit parameter was the Ge depth profile. There are several problems with this method: (1) It is not general, i.e., only the signal coming from one of the elements (Ge) is used, and the information on the Si signal is ignored. (2) Also, the signal of the element of interest is required to be fully separated from all other elements present. (3) It does not take the isotopic distribution of the elements into account (because generally the different isotopes cannot be separated). (4) The method is sensitive to statistical fluctuations on the data, and extra information has to be introduced (e.g., that the samples were multilayers with five SiGe layers). (5) The code developed was not general, as it was written for the particular system under study. It was therefore desirable to develop an alternative, general, method, unique in that no other of the computer codes presented so far (see Ref. 38 for a recent review) could perform automated fits including the effect of energy resolution degradation.

To ensure that the depth profiles obtained with simulated annealing (or any other algorithm) are correct, the theoretical test functions generated must be as realistic as possible. In the papers quoted above, the degradation of energy resolution with depth was explicitly not taken into account, which led to an overestimation of the thickness of interfaces. This was done for reasons of simplicity of the code as well as running time. Also, in most cases (e.g., normal angle of incidence) the energy resolution changes only slowly with depth and full calculations are not necessary. When grazing angle of incidence [as well as in other cases, e.g., in elastic recoil detection analysis (Ref. 39)] is used as in the present

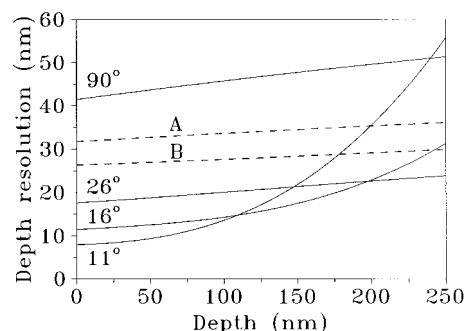


FIG. 1. Depth resolution as a function of depth, as calculated with DEPTH for the two setups used. For the Sacavém high-resolution setup, the angle of incidence ϑ for each curve is shown. For the Surrey standard set up, the depth resolution values for the two detectors A and B used are shown.

case, however, the effect of energy resolution degradation is crucial and must be taken into account for a correct interpretation of the data. The widely distributed code RUMP (Ref. 40) calculates only the effect of Bohr straggling, which makes analysis of complex samples such as multilayers a tedious procedure,⁴¹ and the results are not always reliable. We therefore decided to allow NDF to take advantage of the state-of-the-art calculations available with DEPTH.

To calculate the depth resolution as a function of depth (either by incorporating DEPTH into NDF or by rewriting the code) each time a theoretical function is computed would increase the calculation time unacceptably because in the simulated annealing process thousands or tens of thousands of functions are computed, while each DEPTH calculation takes a few seconds to run on a PII processor running at 300 MHz. The alternatives are either to calculate the depth resolution as function of depth beforehand using the nominal sample composition, or to recalculate it at each value of the control parameter, that is, once per Markov chain. We opted for the first method as it allowed us to use DEPTH directly without the need to incorporate it into NDF. If the final fitted depth profile is significantly different from the nominal one, the depth resolution should be recalculated and the fit repeated. This was not necessary in any of the examples shown in this work.

Finally, the depth resolution values calculated for the experimental conditions in the two setups used in this work are shown in Fig. 1. It is clear that, in the Sacavém high-resolution system, an improvement by a factor of 5 in the depth resolution at the surface is reached by using a grazing angle of incidence $\vartheta = 11^\circ$. However, at that angle the degradation with depth is very fast, due to the increased probability of multiple scattering. Deeper than 100 nm, multiple scattering is the largest contribution to the energy resolution. On the other hand, at normal incidence ($\vartheta = 90^\circ$) or at $\vartheta = 26^\circ$ the depth resolution at the surface is much worse than at a grazing angle of incidence, but degrades much slower with depth, and deeper than 150 nm a better depth resolution is obtained using a $\vartheta = 26^\circ$ angle of incidence than using a $\vartheta = 11^\circ$ grazing angle. The differences observed in depth resolution at normal incidence for the Sacavém and Surrey (detectors A and B) setups are due mostly to the difference in the resolution of the detectors.

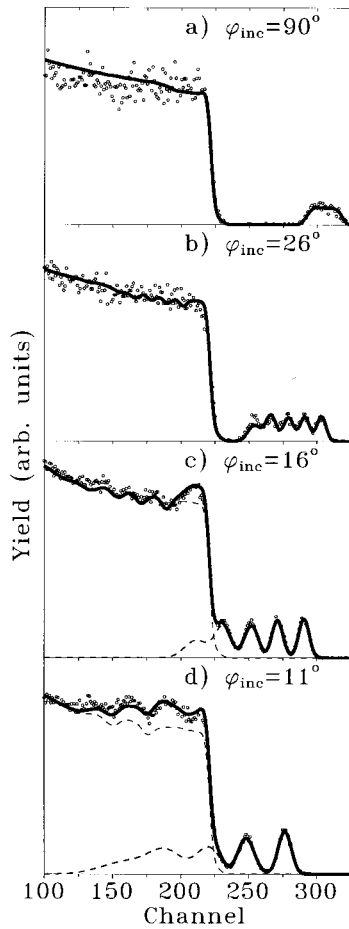


FIG. 2. Fits obtained for the $(\text{Si } 30\text{-nm}/\text{Si}_{0.78}\text{Ge}_{0.22} \text{ 5-nm}) \times 5$ multilayer grown at 550°C . The same depth profile was used to fit the four spectra simultaneously. The fitted partial spectra corresponding to Si and Ge are also shown for the spectra collected at angles of incidence of 16° and 11° .

V. RESULTS AND DISCUSSION

Using simulated annealing and following the procedure outlined in Sec. III, we fitted the RBS spectra taken from the $\text{Si}(001)/\text{Si } 300\text{-nm}/(\text{Si } 30\text{-nm}/\text{Si}_{0.78}\text{Ge}_{0.22} \text{ 5-nm}) \times 5/\text{Si } 50\text{-nm}$ multilayers grown at 550 and 810°C , obtained in the Savacém high-depth-resolution setup. Four spectra were collected from each sample, from normal incidence (90°) to a 11° grazing angle of incidence. All the experimental profiles are fitted simultaneously with the same depth profile, to increase the sensitivity and ensure consistency between the fits. The results are shown in Figs. 2 and 3 for the samples grown at 550 and 810°C , respectively. It is interesting to note that there is some unintentional channeling in the spectra collected at normal angle of incidence, in the region corresponding to the bulk Si substrate. Consequently the region of interest in the fit for these spectra is set from 200 to 350 to avoid the channeled part of the data. The fits obtained are very good. The effect of the degradation of depth resolution is clearly observed in the Ge signal at $\vartheta = 16^\circ$ and 11° , with the signal from each successive Ge peak being broader and worse defined than the previous. The fact that the width of the Ge peaks could be correctly reproduced for all angles of incidence using the same depth profile proves that DEPTH calculates the depth resolution correctly. Otherwise, the fit

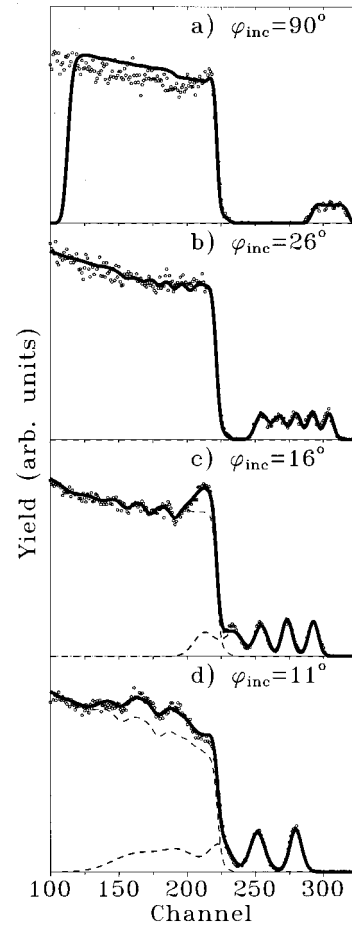


FIG. 3. Fits obtained for the $(\text{Si } 30\text{-nm}/\text{Si}_{0.78}\text{Ge}_{0.22} \text{ 5-nm}) \times 5$ multilayer grown at 810°C . The same depth profile was used to fit the four spectra simultaneously. The fitted partial spectra corresponding to Si and Ge are also shown for the spectra collected at angles of incidence of 16° and 11° .

could not be simultaneously good for all spectra.

The depth profiles obtained are shown in Fig. 4 for Ge (as there are only two elements present the Si and Ge profiles are complementary). Note that a depth of $10^{15} \text{ at}/\text{cm}^2$ corresponds to 0.2 nm , taking into account the density of Si and SiGe, $5 \times 10^{22} \text{ at}/\text{cm}^3$. That the sample grown at 810°C has sharper SiGe layers than the sample grown at 550°C has been confirmed with high-resolution SIMS,¹² and is ex-

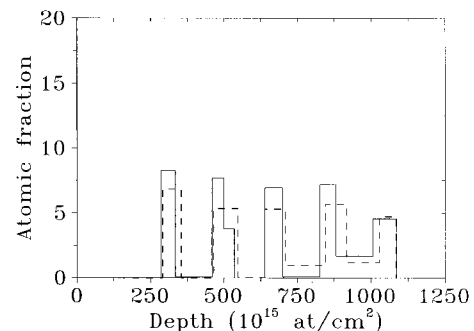


FIG. 4. Fitted Ge depth profiles for the $(\text{Si } 30\text{-nm}/\text{Si}_{0.78}\text{Ge}_{0.22} \text{ 5-nm}) \times 5$ multilayers grown at 550°C (dashed line) and 810°C (solid line).

TABLE I. Depth resolution at the position of the SiGe layers in the $(\text{Si } 30\text{ nm}/\text{Si}_{0.78}\text{Ge}_{0.22} 5\text{ nm})_{\times 5}$ multilayers obtained with the different angles of incidence used, for the Sacavém setup.

Layer	Depth (nm)	Depth resolution (nm)			
		$\vartheta = 11^\circ$	$\vartheta = 16^\circ$	$\vartheta = 26^\circ$	$\vartheta = 90^\circ$
	surface	7.5	10.7	16.9	37.8
1	50	8.9	11.7	17.9	39.0
2	85	11.3	12.7	18.6	40.0
3	120	15.3	14.4	19.4	40.6
4	155	21.3	16.7	20.2	41.4
5	190	29.1	19.6	21.2	42.1

plained because the higher temperature corresponds to the equilibrium regime of Ge segregation near the Si/SiGe interface,⁴² with correspondingly better defined interfaces. These results correlate with the observation by Raman spectroscopy of zone-folded acoustic modes¹² (which depend on the crystalline and interface quality⁴³) only on the sample grown at 810 °C. These modes are not observed for lower growth temperatures due to the increased smearing of the Ge profile. More interesting is the fact that the two Ge peaks closer to the surface seem in the fitted depth profiles to be sharper than the deeper peaks, while we know from the SIMS data that this feature is not real as no such large difference was observed. It is instead due to the depth resolution at the different SiGe layers. The depth resolution values at the position of the five SiGe layers obtained from each of the spectra are given in Table I. The depth resolution is in all cases worse than the 5-nm thickness of the SiGe layers; as any SiGe layer thickness thinner than the depth resolution value is consistent with the data, the simulated annealing algorithm will determine randomly one of the possible values, which however have different probabilities depending on the density of states in the parameter space. At the first SiGe layer the depth resolution at $\vartheta = 11^\circ$ and 16° is still very good, and close to the thickness of the SiGe layer, and the probability of obtaining a solution with a relatively sharp first SiGe layer, thinner than 10 nm, is high. For deeper layers the depth resolution is worse, and that probability is smaller; for instance, for the third layer any thickness value below 15 nm leads to a good fit. On the other hand, the third layer is superimposed on the Si signal at $\vartheta = 11^\circ$, which decreases the sensitivity of the experiment, and in fact the useful information at that depth is taken mostly only from the spectrum collected at $\vartheta = 16^\circ$, which further reduces the sensitivity. Finally, the depth profiles obtained for the multilayer deposited at 810 °C by using the depth resolution values as calculated with DEPTH, and just the 20-keV system resolution without taking into account the degradation of resolution with depth, are compared in Fig. 5. As expected, not to consider the degradation of energy resolution leads to overestimation of the width of the SiGe layers, in order to reproduce the broadening of the Ge signal with an underestimated FWHM below the surface.

As said above, SA only finds, randomly, one of the solutions consistent with the data. BI/MCMC, on the other hand, takes into account the density of solutions in the parameter space and therefore is able to calculate the mean value and

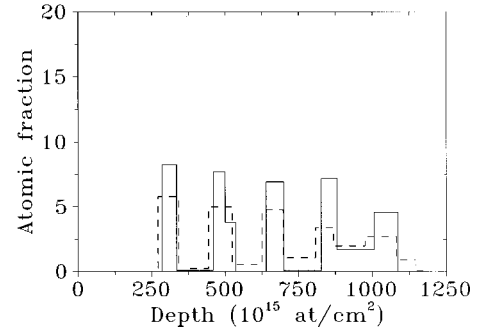


FIG. 5. Fitted Ge depth profile for the $(\text{Si } 30\text{-nm}/\text{Si}_{0.78}\text{Ge}_{0.22} 5\text{-nm})_{\times 5}$ multilayers grown at 810 °C, with the depth resolution as calculated with DEPTH (solid line), and using only the system resolution (dashed line).

standard deviation of the depth profiles that are consistent with the data within the experimental errors. The results obtained with BI/MCMC for the Si(001)/Si 300-nm/(Si 30-nm/Si_{0.78}Ge_{0.22} 5-nm)_{×5}/Si 50-nm multilayers grown at 550 and 810 °C are shown in Figs. 6(a) and 6(b), respectively. The lines indicate the confidence limits (± 1 standard deviation) of the posterior probability distribution defined in Eq. (3). The relatively high ambiguity on the Ge concentration is chiefly due to the limited depth resolution as discussed above. That can be also inferred from the increased error margin at deeper layers, where the depth resolution is worse, and the two deeper layers can not be fully resolved.

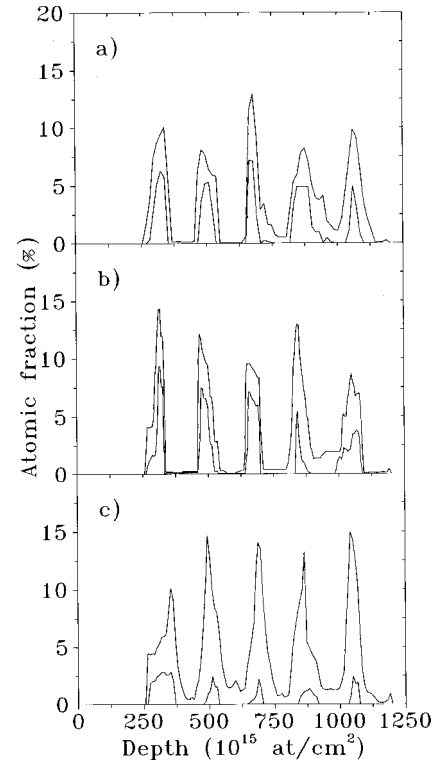


FIG. 6. Confidence limits (± 1 standard deviation) of the posterior probability distribution obtained with BI/MCMC for the Ge depth profiles, for the $(\text{Si } 30\text{-nm}/\text{Si}_{0.78}\text{Ge}_{0.22} 5\text{-nm})_{\times 5}$ multilayers grown at (a) 550 °C and (b) 810 °C. (c) Results obtained for the multilayer grown at 810 °C using only the low resolution spectra collected at normal incidence and at $\vartheta = 26^\circ$.

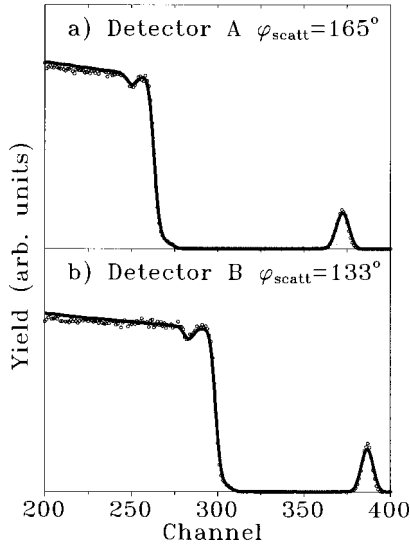


FIG. 7. Fits obtained for the RBS spectra of the Si(001)/Si_{0.7}Ge_{0.3} 7-nm/Si 70-nm sample, taken at normal angle of incidence with detectors A and B in the Surrey standard resolution system. The same depth profile was used to fit the two spectra simultaneously.

The results obtained for the multilayer grown at 810 °C, but derived only from the low resolution spectra collected at normal incidence and at $\vartheta = 26^\circ$ (that is, without using the spectra collected at $\vartheta = 11^\circ$ and at 16°) are shown in Fig. 6(c), and as expected the definition of the layer sharpness is much worse and the error bars increased.

The simulated annealing fits to the data collected with the Surrey standard depth resolution setup for the Si(001)/Si_{0.7}Ge_{0.3} 7-nm/Si 70-nm sample are shown in Fig. 7. There is some unintentional channeling in the spectra, in the region corresponding to the bulk Si substrate. The region of interest in the fit was set from 230 to 400 and from 260 to 400 in the spectra collected with detectors A and B respectively, to avoid the channeled region. The two spectra were fitted simultaneously with the same depth profile to ensure a consistent data treatment. Notice the dip in the Si signal due to the reduced Si concentration in the SiGe layer. The error estimates (± 1 standard deviation) obtained with BI/MCMC are shown in Fig. 8. The expectation value and standard deviation of the Ge concentration and thickness of the layer are 29(8) at. % and 6(2) nm, respectively. These values compare reasonably well with the values obtained with XRD, TEM, and SIMS, 32 at. % and 9(1) nm, respectively.

The depth profile of the sample obtained with BI/MCMC is quite well defined, taking into account that the SiGe layer is very thin and that the experiment has a rather poor

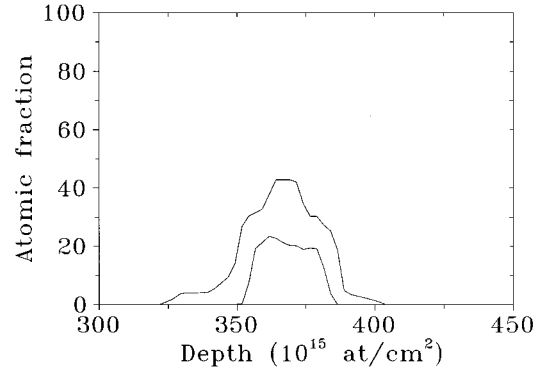


FIG. 8. Confidence limits (± 1 standard deviation) of the posterior probability distribution obtained with BI/MCMC for the Ge depth profile, for the Si(001)/Si_{0.7}Ge_{0.3} 7-nm/Si 70-nm sample.

depth resolution. The depth resolution at 70 nm (350×10^{15} at/cm², taking the Si density into account) for the two detectors is given in Table II. In both cases it is much higher than the $\approx 50 \times 10^{15}$ at/cm² (≈ 10 nm) thickness corresponding to the broader limit determined by BI/MCMC. However, one might expect that any SiGe layer thickness thinner than the depth resolution value would lead to a good fit, and hence the question arises of why does the BI/MCMC calculation produce a significantly less ambiguous confidence interval. The answer lies in the energy loss of the He beam as it crosses the SiGe layer with thickness t , that leads to a finite width $\Delta E_{\text{Ge},t}$ of the Ge (and Si) signal, even not taking into account the broadening due to depth resolution:

$$\Delta E_{\text{Ge},t} = [\varepsilon]t, \quad (8)$$

where $[\varepsilon]$ is the stopping power of He in the SiGe layer. The values of $[\varepsilon]$ calculated for a Si_{0.7}Ge_{0.3} layer and a He energy of 1481 keV (calculated for He after crossing a 350×10^{15} at/cm² thick pure Si layer for an initial energy of 1500 keV) are given in Table II. The energy width of the Ge signal from a delta layer due to the finite depth resolution FWHM is simply

$$\Delta E_{\text{Ge},\text{FWHM}} = \mathcal{F}. \quad (9)$$

Finally, we can determine the thickness t_1 of the SiGe layer that would lead to a one channel broadening of the Ge signal:

$$\Delta E_{\text{Ge},\text{FWHM}}^2 + \{t_1[\varepsilon]\}^2 = \{\Delta E_{\text{Ge},\text{FWHM}} + \Gamma\}^2, \quad (10)$$

where Γ is the gain in keV in the energy calibration. The values obtained for t_1 are given in Table II; a SiGe thicker than 78×10^{15} at/cm² would already lead to a one-channel

TABLE II. Stopping power $[\varepsilon]$ of He in the SiGe layer of the Si(001)/Si_{0.7}Ge_{0.3} 7 nm/Si 70 nm sample, depth resolution at the SiGe layer, energy width $\Delta E_{\text{Ge},\text{FWHM}}$ of the Ge signal due to the depth resolution, energy calibration gain, and layer thickness t_1 necessary to broaden the Ge signal by one channel.

Detector	$[\varepsilon]$ (10^{-15} keV cm ² /at)	Depth resolution (nm)	(10^{15} at/cm ²)	$\Delta E_{\text{Ge},\text{FWHM}}$ (keV)	Gain (keV/channel)	t_1 (10^{15} at/cm ²)
A	0.118	33.0	165	19.5	2.894	93
B	0.147	27.2	136	20.1	3.104	78

broadening in the spectrum collected with detector *B*, and the BI/MCMC calculation rejects that possibility as not being consistent with the data, and hence the confidence interval shown in Fig. 8 is correctly obtained. The same did not happen in the low-resolution BI/MCMC analysis of the multilayer grown at 810 °C [results shown in Fig. 7(c)], where the thickness of each SiGe layer was slightly smaller but the resolution somewhat better, because the peaks coming from the Ge in the five SiGe layers are extensively superimposed.

VI. CONCLUSIONS

We used the simulated annealing and Markov chain Monte Carlo algorithms to derive the composition of SiGe/Si thin films and multilayers from high-resolution Rutherford backscattering experiments, where a depth resolution at the surface of 8 nm was achieved. For the first time, fully automated analysis taking into account the depth resolution as a function of depth was done using the simulated annealing algorithm. This confirmed that the interfaces between the SiGe and Si layers are sharper in samples grown at 810 °C than in samples grown at 550 °C.

Bayesian inference using the Markov chain Monte Carlo method was also employed, and confidence limits on the

SiGe depth profiles obtained from the RBS data were established. We would like to stress that, although it is relatively simple to calculate error bars due to the collection statistics when the signal coming from each layer is well separated, to the best of our knowledge no other method had been previously described able to determine the error in the stoichiometry and thickness for each layer, for the general case when the signals are superimposed as happens in the multilayer samples studied here.

Finally, it is also shown that standard RBS using a normal angle of incidence is sufficient to obtain narrow limits of confidence in the thickness and stoichiometry of single quantum wells, but not of multiple quantum wells, for which high-depth-resolution experiments are necessary.

ACKNOWLEDGMENTS

The authors would like to thank Professor J. C. Soares and Dr. M. F. da Silva for providing access to the high-resolution RBS facility in ITN Sacavém, Portugal, M. Jenkin and Dr. P. K. Marriott for discussions on SA and MCMC, and Dr. Edit Szilágyi for providing the DEPTH code and useful discussions. This work was carried out under EPSRC Contract No. GRL78512.

*On leave from Institute of Radiophysics and Electronics, National Academy of Sciences of Ukraine, Kharkov 310085, Ukraine.

¹R. P. G. Karunasiri, J. S. Park, and K. L. Lang, *Appl. Phys. Lett.* **59**, 2588 (1991).

²D. J. Robins, M. B. Standway, W. Y. Leong, J. L. Glasper, and C. Pickering, *J. Mater. Sci.* **6**, 353 (1995).

³J. A. Leavitt, L. C. McIntyre, and M. R. Weller, in *Handbook of Modern Ion Beam Analysis*, edited by J. R. Tesmer and M. Nastasi (Materials Research Society, Pittsburgh, 1995).

⁴N. 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).

⁵U. Wätjen, H. Bax, and P. Rietveld, *Surf. Interface Anal.* **19**, 253 (1992).

⁶C. Jeynes, Z. H. Jafri, R. P. Webb, M. J. Ashwin, and A. C. Kimber, *Surf. Interface Anal.* **25**, 254 (1997).

⁷N. P. Barradas, C. Jeynes, and R. Webb, *Appl. Phys. Lett.* **71**, 291 (1997).

⁸W. A. Lanford, B. Anderberg, H. Enge, and B. Hjørvarsson, *Nucl. Instrum. Methods Phys. Res. B* **136-138**, 1177 (1998).

⁹N. P. Barradas, J. C. Soares, M. F. da Silva, F. Pászti, and E. Szilágyi, *Nucl. Instrum. Methods Phys. Res. B* **94**, 266 (1994).

¹⁰E. Szilágyi, F. Pászti, and G. Amsel, *Nucl. Instrum. Methods Phys. Res. B* **100**, 103 (1995).

¹¹N. P. Barradas, C. Jeynes, O. A. Mironov, P. J. Phillips, and E. H. C. Parker, *Nucl. Instrum. Methods Phys. Res. B* **139**, 239 (1998).

¹²O. A. Mironov, P. J. Phillips, E. H. C. Parker, M. G. Dowsett, N. P. Barradas, C. Jeynes, M. Mironov, V. P. Gnezdilov, V. Ushakov, and V. V. Eremenko, *Thin Solid Films* **306**, 307 (1997).

¹³M. F. da Silva, M. R. da Silva, E. J. Alves, A. A. Melo, J. C. Soares, J. Winand, and R. Vianden, in *Surface Engineering*, Vol. 85 of *NATO Advanced Study Institute, Series E: Applied Sciences*, edited by R. Kossowsky and S. Singhal (Les Arcs,

Nijhoff, The Hague, 1984), p. 74.

¹⁴F. Pászti, A. Manuaba, C. Hadju, A. A. Melo, and M. F. da Silva, *Nucl. Instrum. Methods Phys. Res. B* **47**, 187 (1990).

¹⁵C. Jeynes, N. P. Barradas, M. J. Blewett, and R. P. Webb, *Nucl. Instrum. Methods Phys. Res. B* **136-138**, 1229 (1998).

¹⁶S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Science* **220**, 671 (1983).

¹⁷Emile Aarts and Jan Korst, *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing* (Wiley, Chichester, 1989).

¹⁸R. H. J. M. Otten and L. P. P. van Ginneken, *The Annealing Algorithm* (Kluwer, Norwell, MA, 1989).

¹⁹W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed. (Cambridge University Press, Cambridge, 1992), p. 438.

²⁰Y. A. Wilks, B. M. Sator, and L. M. Guthrie, *Electric Words: Dictionaries, Computers, and Meanings* (MIT Press, Cambridge, MA, 1996), pp. 203–206.

²¹N. Metropolis, A. W. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).

²²N. P. Barradas, P. K. Marriott, C. Jeynes, and R. P. Webb, *Nucl. Instrum. Methods Phys. Res. B* **136-138**, 1157 (1998).

²³N. P. Barradas, C. Jeynes, R. P. Webb, U. Kreissig, and R. Grötzschel (unpublished).

²⁴N. P. Barradas, C. Jeynes, and M. A. Harry, *Nucl. Instrum. Methods Phys. Res. B* **136-138**, 1163 (1998).

²⁵N. P. Barradas, C. Jeynes, and S. M. Jackson, *Nucl. Instrum. Methods Phys. Res. B* **136-138**, 1168 (1998).

²⁶N. P. Barradas, C. Jeynes, K. P. Homewood, B. J. Sealy, and M. Milosavljevic, *Nucl. Instrum. Methods Phys. Res. B* **139**, 235 (1998).

²⁷The Data Furnace: manual for version v1.0 (October 1997), and update to version v2.1 (March 1998), University of Surrey; <http://www.ee.surrey.ac.uk/Research/SCRIBA/ndf/>.

- ²⁸R. Fischer, M. Mayer, W. von der Linden, and V. Dose, Phys. Rev. E **55**, 1 (1997).
- ²⁹R. Fischer, M. Mayer, W. von der Linden, and V. Dose, Nucl. Instrum. Methods Phys. Res. B **136-138**, 1140 (1998).
- ³⁰V. M. Prozesky, J. Padayachee, R. Fischer, W. von der Linden, V. Dose, and R. A. Weller, Nucl. Instrum. Methods Phys. Res. B **136-138**, 1146 (1998).
- ³¹J. W. Butler, Nucl. Instrum. Methods Phys. Res. B **45**, 160 (1990).
- ³²P. F. A. Alkemade, F. H. P. M. Habraken, and W. F. van der Weg, Nucl. Instrum. Methods Phys. Res. B **45**, 139 (1990).
- ³³N. P. Barradas, C. Jeynes, M. Jenkin, and P. K. Marriott (unpublished).
- ³⁴P. M. Lee, *Bayesian Statistics: An Introduction*, 2nd ed. (Arnold, London, 1997).
- ³⁵E. Szilágyi. DEPTH can be downloaded free of charge from <http://www.kfki.hu/~ionhp/>.
- ³⁶E. Szilágyi and F. Pászti, Nucl. Instrum. Methods Phys. Res. B **85**, 616 (1994).
- ³⁷P. Børgesen, R. Behrisch, and B. M. U. Scherzer, Appl. Phys. A: Solids Surf. **27**, 183 (1982).
- ³⁸E. Kotái, Nucl. Instrum. Methods Phys. Res. B **85**, 588 (1994).
- ³⁹E. Szilágyi, L. S. Wielunski, and F. Pászti, Nucl. Instrum. Methods Phys. Res. B **136-138**, 701 (1998).
- ⁴⁰L. R. Doolittle, Nucl. Instrum. Methods Phys. Res. B **9**, 291 (1985).
- ⁴¹C. J. Tavares, L. Rebouta, E. J. Alves, B. Almeida, J. Bessa e Sousa, M. F. da Silva, and J. C. Soares, Nucl. Instrum. Methods Phys. Res. B **136-138**, 278 (1998).
- ⁴²S. Fukatsu, K. Fujita, P. Yaguchi, and R. Ito, Appl. Phys. Lett. **59**, 2103 (1991).
- ⁴³C. Colvard, T. A. Grant, M. V. Klein, R. Martin, R. Fisher, H. Morkoc, and A. C. Gossard, Phys. Rev. B **31**, 2080 (1985).