UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY. CABLE: CENTRATOM TRIESTE

SMR.550 -17

SPRING COLLEGE IN MATERIALS SCIENCE ON
"NUCLEATION, GROWTH AND SEGREGATION IN MATERIALS SCIENCE AND ENGINEERING"
(6 May - 7 June 1991)

BACKGROUND MATERIAL

for lectures on

"STATIC AND DYNAMIC SIMS"

M.G. DOWSETT
Advanced Semiconductor Group
Department of Physics
University of Warwick
Coventry CV4 7AL
United Kingdom

These are preliminary lecture notes, intended only for distribution to participants.

Springer-Verlag 19

The application of surface analytical techniques to silicon technology

M. G. Dowsett

Advanced Semiconductor Group, Department of Physics, University of Warwick, Coventry CV4 7AL, UK

Received February 25 1991

Summary. Most modern semiconductor device engineering takes place in the top micron of the host wafer and involves the creation of regions whose composition varies over lateral dimensions which may be less than 0.5 μm. Normal to the wafer surface, large changes in matrix and impurity composition may occur in the space of a single atomic plane. In future, the fabrication of quantum dots and wires will result in active device features a few nm in extent. Such material developments need to be supported by parallel development in surface analytical techniques with high spatial resolution. There are, however, fundamental limitations to what can be achieved directly. For a destructive technique such as secondary ion mass spectrometry (SIMS), the analytical sensitivity and spatial resolution are determined by the analyte volume which needs to be consumed to achieve the necessary statistical precision. Moreover, the type of information obtained depends on the details of the interaction between the primary probe and the sample surface. In order to combine high spatial resolution with high sensitivity, special sample structures and modified instruments are required. Techniques need to be developed for accurately compensating for the effects on the analysis of large localized changes in conductivity in the materials. A multi-technique approach to semiconductor analysis is required both to investigate the limitations of the techniques themselves, and to fully describe the material properties.

1 Introduction

Surface analysis techniques may be divided into two broad categories: those in which consumption and alteration of the sample is intrinsic to the process (destructive techniques), and those in which the sample is not deliberately consumed (non-destructive techniques). Mass spectrometric techniques (e.g. secondary ion neutral mass spectrometry – SIMS or SNMS) fall into the former category, whilst electron spectroscopies (AES, XPS) and Rutherford backscattering (RBS) fall into the latter. For semiconductor analysis, there is often little difference between the two, either because a sectioning technique (e.g. sputtering) is a necessary adjunct to analysis in order to measure a spatial distribution, or because the accumulated probe does is sufficient to disrupt the electrical or structural properties of the material anyhow. In all cases it is important to determine the effect that any disruption of the sample has on the quality of the data and the limitations of the technique. To this end it is necessary to use a range of analytical techniques in a mutual characterization exercise with respect to the material system of interest.

In SIMS and related techniques [1], a primary beam of ions or neutrals is used to sputter material from a solid surface. The sputtered products consist of atoms, clusters, and molecules a fraction of which are themselves ionised in the sputtering process. SIMS makes use of these secondary ions only, whilst other techniques, e.g. SNMS and resonant ionisation mass spectrometry (RIMS), post ionize some fraction of the sputtered neutrals [2, 3]. In either case, the ions produced are collected and mass analysed using double focussing magnetic sector (DFMS) [4], time-of-flight (TOF) [5], or quadrupole (QMS) mass spectrometers [6] depending on the exact purpose of the instrument. Primary beams in common use are oxygen, caesium, gallium, and argon, although many others have been reported [7–9]. The reactive species are used to enhance the ionization coefficient for SIMS, with caesium (or other electropositive species) being used to enhance negative ion yields, and oxygen (or other electronegative species) being used to enhance positive ion yields [10, 11, 31, 39]. However, this bland statement masks an extremely complex situation where bombardment with either ion type can enhance the emission of both charge states and also specific cluster ions [12]. The type of analysis acquired (viz. micro-bulk, image, depth profile, surface) depends on the parameters of the primary probe, the ion optics of the mass spectrometry system, and the accumulated dose per data point.

2 Limitations of destructive techniques

For a destructive analytical technique such as SIMS, the spatial resolution required in a measurement will determine the size and shape of the analyte volume. This may vary between extremes of a few $mm^2 x 0.01$ monolayers (static SIMS [13]) to 20 nm diameter x ≈ nm's deep (high lateral resolution imaging [14]). Nevertheless, the amount of material in the volume will directly determine the analytical precision and hence (in the absence of an interfering background) the ultimate useful sensitivity. The fractional atomic concentration C_x of an atomic species X in a measure-

Analytische Chemie Ms -No 5452 Author Dowsett Ms 1-15 Pages 1-11

Springer-Verlag 6900 Heidelberg 1 / Wiesbadener Graphische Betriebe GmbH 6200 Wiesbaden

Provisorische Seitenzahl / Provisional page number 1 Korr Date

ment which consumes N sample atoms is determined from a SIMS or similar analysis from

$$C_x = n_x N / \Gamma_x \alpha_x \qquad (1)$$

where n_x is the number of secondary particles (positive or negative ions, or neutrals) of X detected; Γ_x is the product of the collection, transmission, and detection efficiencies of the spectrometer for species X; α_x is the emission or production probability for the charge state detected.

$$\alpha_x^r = \frac{\text{no. of X produced in charge state } c}{\text{total no. of X sputtered}}$$

It is important to note that α implicitly contains factors due to impurity migration occurring during analysis (e.g. segregation effects) and similar phenomena. It also includes the effects of post-ionisation processes, even in SIMS [15]. Γ_x and α_x are difficult to measure independently (for SIMS). Their product $Y_x = \Gamma_x \alpha_x$, known as the useful yield of X is easily obtained, and is used as a measure of instrumental quality. Useful yields are a strong function of the analytical conditions, especially the primary probe species and bombardment angle, and the field of view of the ion optical system used to collect the secondary products. By virtue of the Helmholtz-Lagrange equation [16], a SIMS instrument optimised for a small field of view (e.g. 10 – 30 μm for DFMS or TOF and 100 – 300 μm for QMS) will collect a larger fraction of the energy and angular distribution sputtered from that area than one optimised for a larger field of view, hence increasing Γ[1]. Similarly, the use of reactive primary ions (e.g. oxygen, caesium) at normal (oxygen) and near normal (caesium) incidence leads to enhanced values for α [10, 11]. Typical useful yields, and the conditions under which they may be obtained are given in [17] and [18].

Equation 1 gives a simple guide to whether a particular analytical task is possible. For example, typical SIMS depth profiling conditions involve a 10^4 μm^2 field of view. If $n \geq 10$ is taken as the condition for acceptable measurement precision, and for a high useful yield of 0.01, a detection limit of $\approx 10^{13}$ atoms cm^{-3} (i.e. $C_x = 2\times10^{-11}$) could be achieved for the erosion of a 10 nm depth. In an image with a lateral resolution of 30 nm, assuming the same useful yield could be obtained, the detection limit would be $\approx 4\times10^{19}$ atoms cm^{-3} for a sampled depth of 30 nm.

A mass spectrum is discrete, and inherently background free, whereas electron and other spectra are continua and the analytical information is contained in changes superimposed thereon. There should be no problem in obtaining an adequate abundance sensitivity for a mass M in the presence of mass M ± 1 in a well set up instrument [19]. However, detection limits are generally worse than the purely statistical limits described above because of memory effects, residual gas adsorption, and mass interferences [20 – 23].

3 Effects of the primary probe

3.1 The altered layer

In order of increasing near surface damage per unit primary dose, probes used in common surface chemical analysis techniques may be listed as follows: UV, X-rays, electrons, high energy (> 100's keV) ions, low energy (keV) neutrals, low energy (keV) ions. However, mass spectrometric techniques using the latter two probe types provide a very efficient analytical yield per unit dose because of the discrete nature of a mass spectrum. When a solid is bombarded by a beam of ions (or neutrals), the coincident processes of sputtering, beam induced mixing, and primary beam incorporation give rise to an altered layer in the surface of the material [24 – 26]. As sputtering progresses, it is frequently the case that the altered layer comes into a steady state for a primary dose density of $\approx 10^{17}$ cm^{-2}. This phenomenon makes dynamic SIMS and allied techniques feasible.

The altered layer is not expected to be homogeneous, and its composition will in general be a disordered mixture of the original matrix elements and probe atoms. All the processes which give rise to distortions in SIMS data take place in this layer or at its boundaries. It follows that a better understanding of the processes by which the layer forms and maintains its equilibrium will lead to improved ability to deal empirically and theoretically with atomic mixing.

With others, we have carried out a detailed study of the altered layer produced by normal incidence oxygen bombardment on silicon using SIMS, transmission electron microscopy (TEM), Auger depth profiling (AES), X-ray photoelectron spectroscopy (XPS), Rutherford backscattering (RBS), and other techniques [27 – 30]. Other studies in this area include those of Wittmaack et al. [12, 31 – 34], von-Criegern and co-workers [35, 36], and Kilner and co-workers [37, 38]. The latter used ^{18}O isotope tracer techniques to observe the substitution and migration of oxygen in a buried altered layer, and developed a model to explain the observations.

Figure 1a shows an off-focus TEM cross section of an altered layer formed on silicon by 4.4 keV O_2^+ bombardment at normal incidence. The back interface and the surface are flat to about 1 nm which has important consequences for the ultimate depth resolution of the technique. A Fresnel fringe is consistently observed towards the back of the layer. This fringe is not observed on similar layers made well away from normal incidence, or in layers made with inert gas bombardment. This can be seen in Fig. 1b where the altered layer due to the Ar^+ ions used in the TEM specimen preparation intersects with that due to O_2^+. Figure 2 shows a TEM lattice image of a similar altered layer. The Si(111) planes are clearly visible, and although the altered layer is obviously not crystalline, there is a clear change in contrast in the region corresponding to that between the Fresnel fringe and the back interface in Fig. 1a. Close examination of the region corresponding to the fringe shows the existence of an interdigitated boundary or region of patchy contrast suggesting chemical inhomogeneity.

Figure 3a shows XPS data taken in-situ as the primary dose was increased from zero, and the altered layer evolved towards steady state. Initially, the elemental Si 2p peak at 102.1 eV, and a peak at 106.4 eV due to the native oxide are observed. As bombardment progresses, a complex structure develops between the two peaks, corresponding to random bonding of new oxygen. The centroid of the system shifts progressively to 106 eV, until at steady state the spectrum closely resembles that from thermally grown SiO_2. (Thin chain dotted line for comparison.) The ion beam grown material however, exhibits a somewhat wider peak, possibly because of radiation induced bond strains.

[1] The ultimate useful yield in a SIMS instrument is obtained by optimizing the extraction optics for an area a little larger than the analytical probe, and then electrostatically scanning the area in synchronism with the probe to achieve the desired field of view. The process has been termed dynamic emittance matching by Liebl [59]

Figure 3b shows an AES profile through a 4.4 keV altered layer. The composition is apparently constant until a depth corresponding the fringe in Fig. 1a.

These results for normal incidence bombardment are somewhat surprising. Over a wide primary ion energy range (R_p) the steady state altered layer consists of a uniform surface layer with the stoichiometry of SiO_2 extending to about 2.5 R_p, followed by a layer with stoichiometry graded smoothly from SiO_2 to Si about 0.5 R_p thick and a relatively sharp back interface with the silicon. The TEM data show the boundaries between these regions clearly and indicate that the transition zone is not laterally homogeneous. The overall thickness of the layer scales with energy and for O_2^+ ions amounts to 4.5 nm keV^{-1}. The thickness as a function of primary ion energy, and comparisons with other work are shown in Fig. 4.

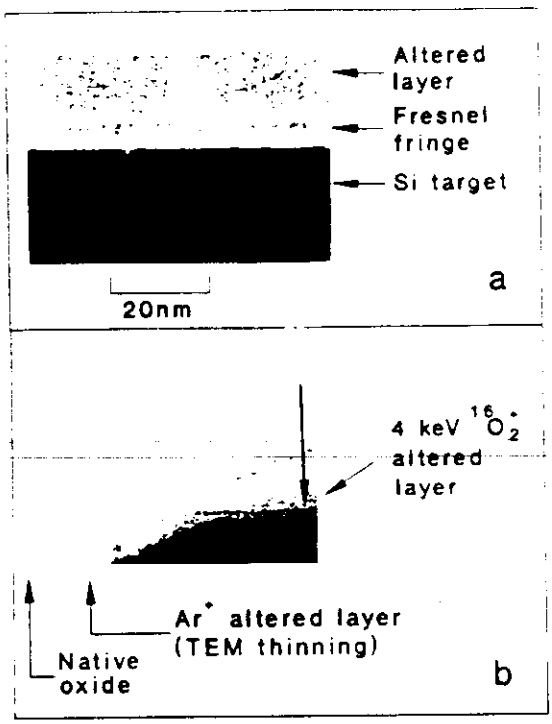


Fig. 1. a Transmission electron micrograph (TEM) of the altered layer produced on Si (100) by bombardment to steady state with 4.4 keV $^{16}O_2^+$ ions. Both the surface and the interface between the altered layer and the Si are flat to better than 1 nm (although the TEM will exaggerate the flatness because it is sampling through several superimposed planes in the section). The Fresnel fringe is consistently observed at about 3/4 the total altered layer depth in a through focus sequence, and is believed to delineate the boundary between stoichiometric SiO_2 and a sub-oxide in the vicinity of the interface. **b** Intersection (fortuitous) of the oxygen altered layer with and Ar$^+$ altered layer formed during sectioning for TEM. Note the native oxide on the surface of the Ar$^+$ layer (formed on removal from the thinning kit) and the behaviour of the Fresnel fringe, which is not seen in the Ar$^+$ altered layer

3.2 *Mechanism of formation*

It is interesting to note that to grow a thermal oxide at the rate observed in the altered layer requires a temperature around 950 C. The real temperature of the surface is nowhere near this level (probably only a few degrees above ambient). It is therefore certain that the oxygen migration and reaction are radiation and or electric field enhanced. It is also clear that a substantial fraction of the input primary oxygen atoms merely escape from the surface, probably by effusion. This can be concluded from equilibrium arguments originally presented by Williams [39]. In the present case, at the measured sputter rate for the 4.4 keV bombardment, oxide must have been created at a rate of 1.73×10^{-10} g s. The dose rate was 1.28×10^{-10} g s. Thus the unreacted fraction = 0.28. An oxygen atom implanted into a previously formed altered layer can therefore effectively migrate freely until it either escapes, or finds some silicon to react with. Collins [40] has recently developed a theoretical model of the altered layer formation based on this observation.

The breakdown strength of thermally grown SiO_2 is about 13 MV cm^{-1}, and an SiO_2 layer 10's nm thick can support a large electric field. However, the XPS measurements made in situ showed no sign of charging. Ex-situ measurements of I – V characteristics of the altered layers [41] showed that they were extremely leaky, and passed current at low voltages. However, the slopes of the I – V curves would permit the sustaining of a few volts across the altered layer at current levels used in SIMS and the possibility that a surface electric field contributes to the altered layer formation mechanism cannot currently be ruled out

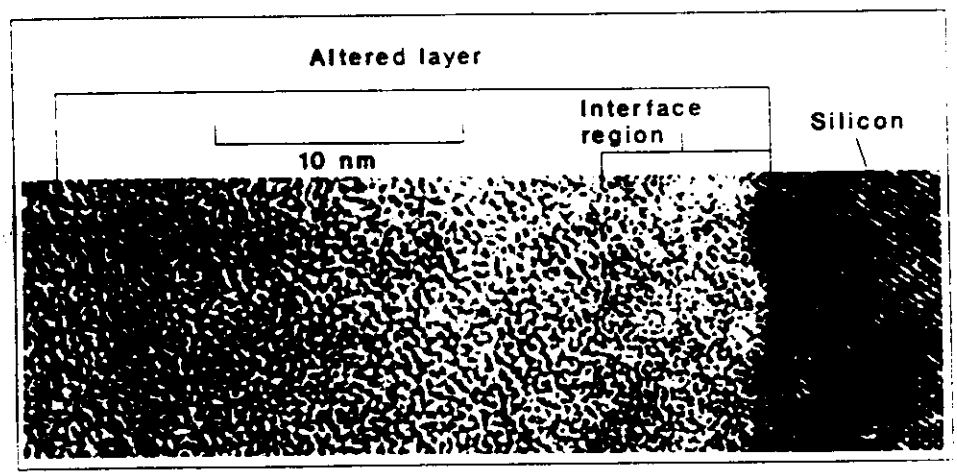
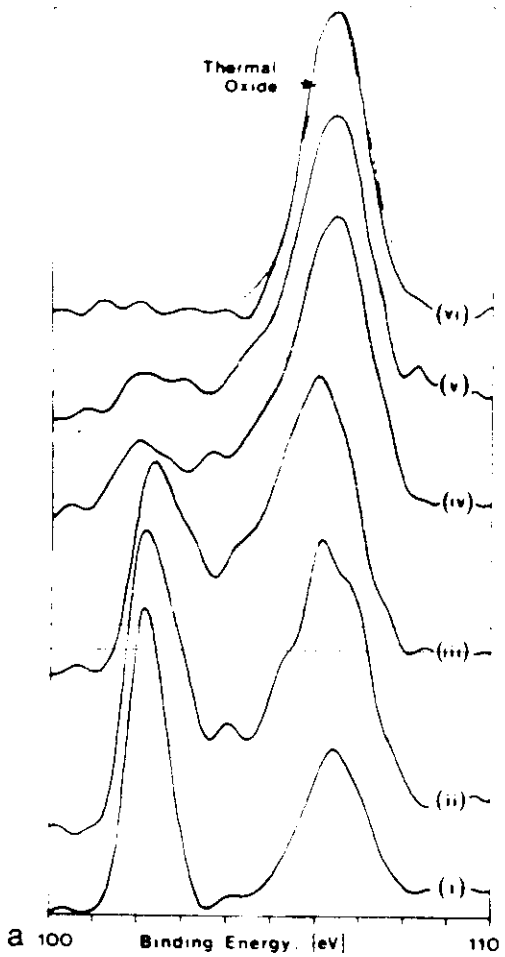


Fig. 2. TEM lattice image of the 4.4 keV altered layer. The abrupt termination of the Si(110) planes at the interface can be seen clearly. The interface appears very flat (about 1 nm waviness). The patchy contrast in the interface region may indicate chemical or structural inhomogeneity – e.g. silicon rich regions. This changes to a uniform grainy effect at depths shallower than the Fresnel fringe in Fig. 1a



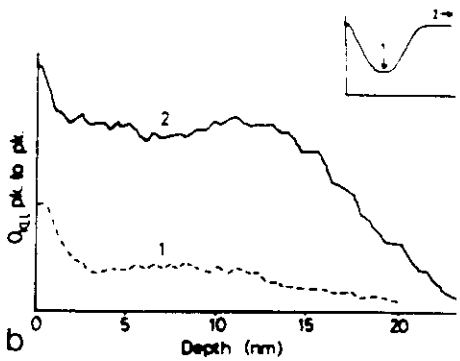


Fig. 3. **a** XPS spectra in the Si 2p region showing the effect of accumulation 4 keV O_2^+ dose on the near surface chemistry. Spectrum *i* shows the initial state of the Si(100) surface with a metallic Si peak at 102.1 eV and a peak due to native oxide at 106.4 eV. Spectra *ii* to *vi* were taken after accumulated in-situ doses of 1.3×10^{16}, 2.7×10^{16}, 4×10^{16}, 1×10^{17}, and 3×10^{17} O_2^+ cm^{-2} respectively. The final state of the surface is very similar to thermal SiO_2 (thin chain dotted line) with some broadening possibly due to radiation-induced bond strain. **b** AES depth profile (1.5 keV Ar^+ ions, defocussed e^- beam) of the 4.4 keV altered layer. Profile *i* is through a dose corresponding to the minimum of the Si^+ secondary ion emission, profile *ii* is from a crater dosed to steady state. By comparison with similar profiles through thin thermal oxides, the plateau region of *ii* is found to correspond to SiO_2. The surface peak is due to atmospheric contaminants and oxide reduction in the Ar^+ pre-equilibrium region

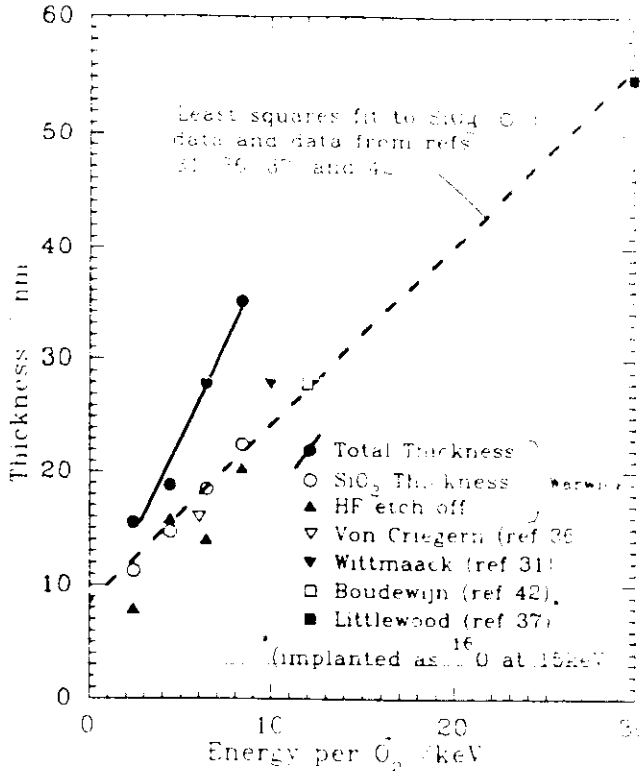


Fig. 4. Altered layer thickness plotted as a function of incident O_2^+ energy. Total thickness (*filled circles*) is estimated from TEM data (similar to Fig. 1a. SiO_2 thickness (*open circles*) is estimated from the depth to the Fresnel fringe. Etch data (*filled triangles*) is the thickness removed from the altered layer by a 5 min etch in 5% hydrofluoric acid

3.3. Implications for SIMS analysis

The formation of the SiO_2 altered layer is intimately linked to the observations of chemical segregation for dopants such as copper [42], and the anomalous redistribution of arsenic and germanium [43, 44], during analysis. There are additional beneficial implications for atomic mixing. The altered layer consists of matrix atoms which are on average farther apart than in the original material, interspersed with primary beam atoms. A layer of matrix silicon is redistributed in the altered layer into an expanded and diluted band approximately twice as thick. If cascade atomic mixing only takes place towards the top of the altered layer where the energy deposition is greatest, and if the expansion effect takes place "gently", then because mixing is occurring in a feature that has already been stretched, it will have a relatively lesser effect on a depth profile. More generally, it might be the case that mixing effects due to the presence of a chemical boundary (Eg. Gibbsian segregation) might be spatially decoupled from cascade mixing effects.

Secondary ion yields will be determined by the composition of the top 1 nm or so of the altered layer. The stoichiometry of the bulk of the altered layer cannot therefore be determined directly from ion yields. In addition this surface layer exists only in equilibrium with the primary beam, and cannot therefore be examined ex situ.

The SiO_2 altered layer may be removed together with the damage it contains by simply etching with hydrofluoric acid. This observation is the basis for the technique of bevel etch imaging [60].

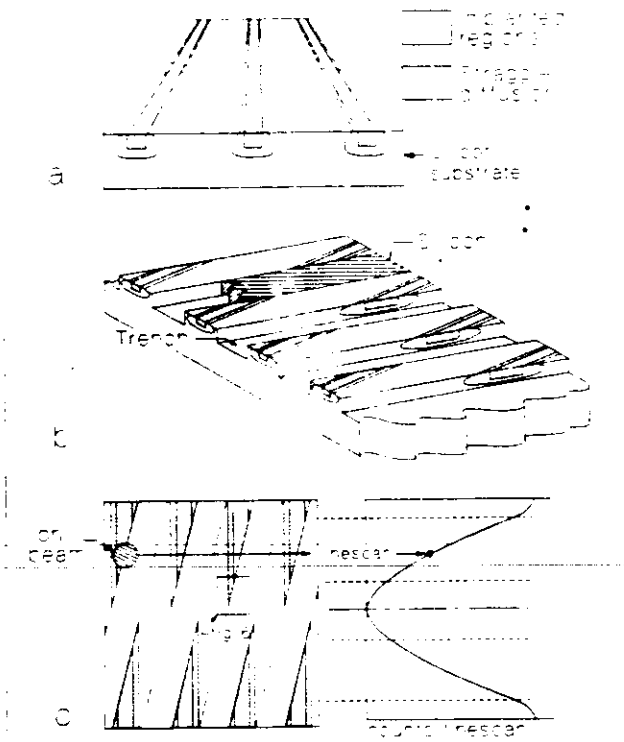


Fig. 5. a 2-D profiling sample: first stage – implantation of dopant stripes 1 – 2 μm wide×2 mm long, separated by 2 – 4 μm, over a 2 mm square pad. **b** Trenches are plasma etched at $\theta = 0.2 - 0.02$ to the implanted stripes. **c** SIMS linescans perpendicular to the stripes produce one data set per complete frame. The slope of the data contains the 1-D lateral distribution. Repeated frames yield 1-D lateral profiles at a series of depths

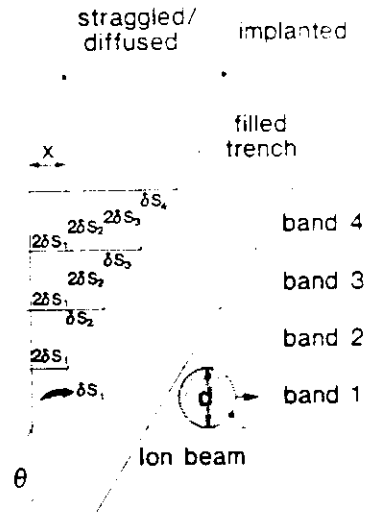


Fig. 6. Detail of apex of triangular region where the stripe intersects with the trench. Each band corresponds to a singe linescan of the SIMS probe. In *band 1*, a set of apex triangles each yielding a signal δS_1 will be analysed. This signal is generated by dopant which has undergone the most lateral migration. *Band 2* yields a signal S_2 from a triangular element δS_2 and a rectangle $2\delta S_1$. δS_2 contains dopant which has laterally migrated a distance $x = d \tan\theta$ less than dopant in δS_1. d is the band width (beam diameter)

4 Two-dimensional profiling

4.1 Problem description

As semiconductor device dimensions are reduced below 1 μm, the control of dopant diffusion during processing becomes progressively more important. Lateral diffusion (i.e. in a direction parallel with the wafer surface) is a significant factor in determining device characteristics and achievable packing densities. The constraints discussed in section 2 show clearly that adequate sensitivity and spatial resolution cannot be achieved directly with a destructive technique. (Detection limits better than 10^{16} atoms cm^{-3} combined with a lateral resolution of at least 30 nm and a depth resolution of at least 10 nm are required.) From Eq. 1, the best possible detection limits which could be achieved for boron, arsenic, and phosphorous in silicon would be $\approx 10^{19}$, 10^{21}, and 10^{20} atoms cm^{-3} respectively. A partial solution is to sacrifice one lateral dimension, and construct samples in the form of a stripe so that the analyte volume may be made very long. However, the practical lateral resolution remains probe diameter limited to about 1 μm since smaller probes contain insufficient current to perform the experiments in a reasonable time.

4.2 A sample geometry

The ideal sample for lateral diffusion studies would divorce the achievable lateral resolution from the beam diameter and permit the integration of n (Eqn. 1) over a large volume without sacrificing resolution. Proceeding by analogy with anodic profiling techniques, Hill and co-workers have devised and tested a range of such samples [45]. The SIMS equipment and analysis technique have been developed in the author's research group [46]. The basic sample structure is as follows (see Fig. 5a, b): Parallel dopant stripes with roughly equal mark/space ratio a few μm wide and ≈ 2 mm long are implanted into a 2 mm×2 mm area. These are then processed in the same way as a prototype device to achieve the same diffusion effects. After processing, channels of approximately the same width and with the same repeat pitch are plasma etched across the stripes at a small angle (≤ 0.2°) to them. The stripes now intersect with the channels producing long, narrow, triangular doped regions. A normal SIMS depth profiling beam say 50 μm in diameter will envelop many of the stripes, and if the beam is line-scanned normally to the direction of the original stripes (Fig. 5c), data corresponding to a 50 μm band at the same height across 100's of triangles may be collected. The sampling depth δz will correspond to the depth sputtered per line. The band containing the apices of the triangles is labelled band 1, a band one beam diameter up the structure band 2, and successive parallel bands 3, 4 ... j etc. For a band containing the apex of the triangles (band 1), the maximum lateral spread x of dopant sampled will be:

$$x = d \tan\theta, \tag{2}$$

where d is the beam diameter and θ is the apex angle. (For $\theta = 0.2$° and $d = 50$ μm, x would be 200 nm.) Figure 6

shows the geometry in more detail. The data from band 1 contains a signal δS_1 from a set of triangular elements of area $(d^2 \tan\theta)/2$ which contain material which has the greatest lateral migration. Band 2 contains signal S_2 from a set of rectangular elements of area $d^2 \tan\theta$ (which contribute twice the signal of the element in band 1) added to the signal from triangular areas a distance x into the stripes. This signal, δS_2, may be recovered by subtracting twice δS_1 from S_2. In general, the signal δS_j for the element set a distance $(j-1)x$ from the (arbitrary) starting position will be:

$$\delta S_j = S_j - 2 \sum_{i=1}^{i=j-1} \delta S_i \qquad (3)$$

4.3 Sensitivity

The elemental area analysed in these experiments is a triangle $(d^2 \tan\theta)/2$. As the beam is scanned laterally, data are accumulated from m triangles, where m is the number of dopant stripes traversed. The analyte volume is therefore

$$V = m(d^2 \tan\theta)\delta z/2. \qquad (4)$$

where δz = the eroded depth per set of linescans.

Once the data have been acquired, bands may be integrated laterally to trade sensitivity for lateral resolution, or successive eroded planes may be integrated vertically to trade sensitivity for depth resolution. If p bands and q planes are integrated, the lateral resolution becomes $p\,d \tan\theta$, the depth resolution becomes $q\,\delta z$, and the total number of sample atoms N consumed per data point is:

$$N = D\,m\,p^2 d^2 \tan\theta\, q\,\delta z/2 M_{ave} \qquad (5)$$

where D is the matrix density and M_{ave} is its average atomic weight. From (1), the minimum detectable impurity concentration becomes:

$$C_x^{min} = \frac{20 M_{ave}}{D\,m\,p^2 d^2 \tan\theta\, q\,\delta z\, \Gamma_x \alpha_x} \qquad (6)$$

For the following easily achieved experimental conditions – $\theta = 0.2$, $d = 50\ \mu m$, $\delta z = 10$ nm, $m = 200$, $p = q = 1$ – and for a useful yield ($\Gamma_x \alpha_x$) of 0.01, an effective lateral resolution of ≈ 200 nm combined with a detection limit of $\approx 10^{14}$ atoms cm^{-3} could be achieved in principle.

4.4 Results and discussion

Preliminary experiments on the structures described, revealed the need to fill the trenches with a passive material (e.g. undoped silicon) to avoid rapid erosion of the trench sidewalls. Figure 5c shows typical raw data from one complete linescan set (i.e. each data point is obtained from a single gated linescan across the structure). Figure 7 shows the first 2-D dopant profile reconstructed from a boron implanted silicon sample in comparison with a TITAN[2] simulation. Selective epitaxy was used to fill the trenches (Fig. 8). In this case, the lateral boron distribution is due to lateral straggle occurring during implantation. In the experimental data, the concentration contours are somewhat shallower than in the simulation. This is because the selective epitaxy requires a ≈ 5 nm oxide coating on the trench walls. Under the bombardment conditions used (4 keV O_2^+ ions, normal incidence), the narrow oxide stripes in the initial

[2] CNET CNS Process Simulator TITAN-V

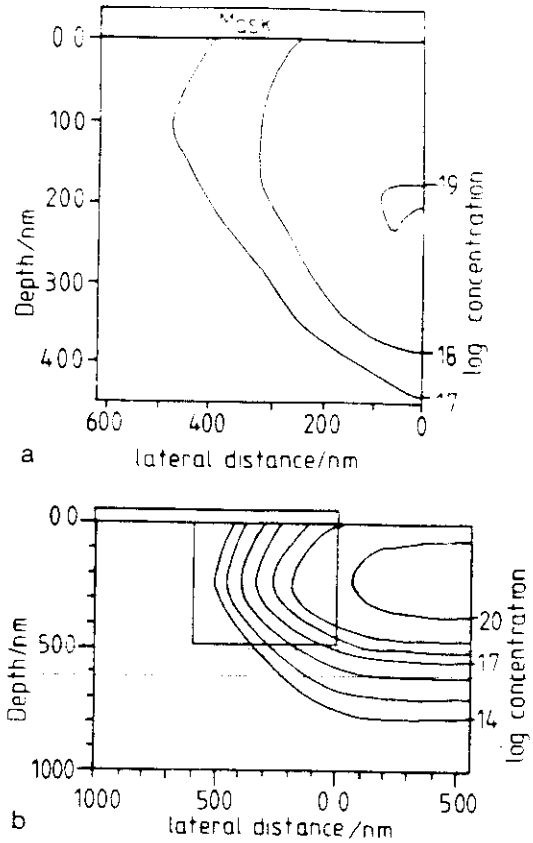


Fig. 7. **a** The first numerical reconstruction of a 2-D profile using Eq. 3 directly. **b** A simulation of straggle under a mask edge using TITAN for comparison

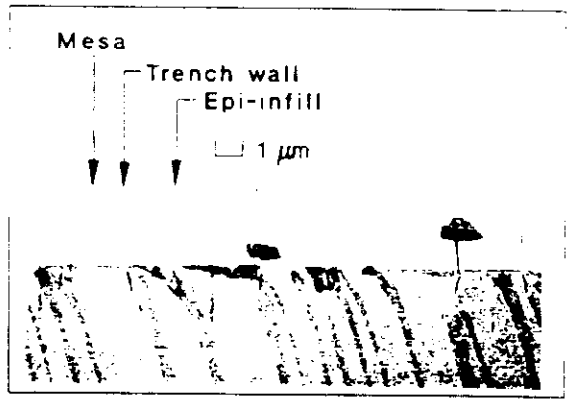


Fig. 8. SEM section of a 2-D sample showing trenches infilled with selective epi-silicon

surface, erode about twice as rapidly as the surrounding silicon. Eventually, this leads to lateral and well as vertical erosion of the overall structure, causing underestimation of the depth scale. A satisfactory infill technique has only been developed recently, and will be reported elsewhere [47].

5 Insulating and resistive layers

5.1 Surface charging and some solutions

The electrical conductivity of different regions in semiconducting materials and devices varies from a few $\mu\Omega$ cm

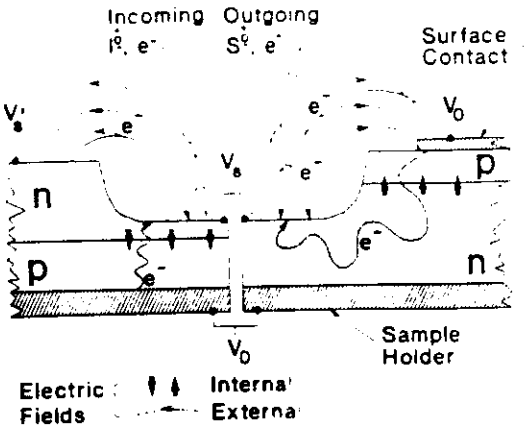
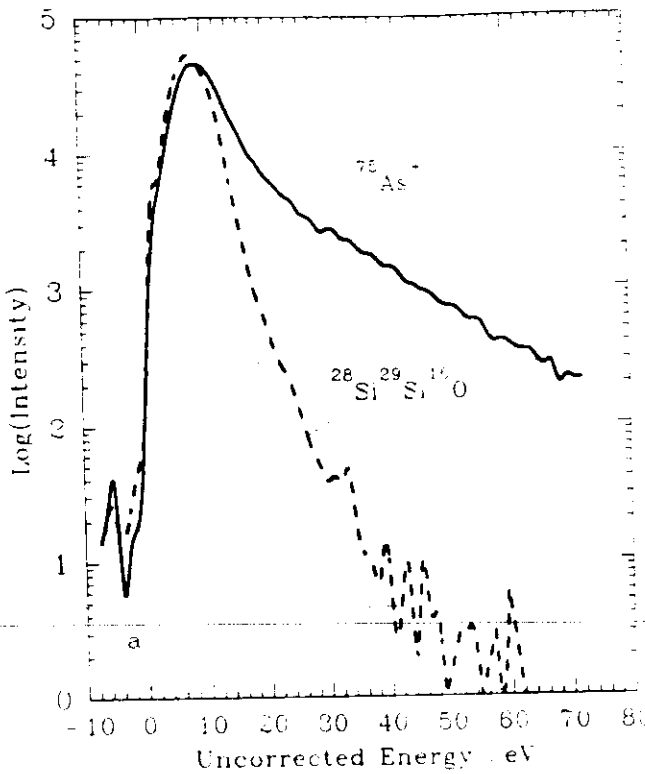


Fig. 9. Factors influencing surface potential during SIMS analysis in a complex multi-layer semiconductor structure. Note particularly the possibility of reverse biased p – n junctions with charge flowing through the material to compensate for probe charge

(very highly doped material or metal regions) to $\approx 10^{14}$ Ω cm (insulators). When positive ions are used as the primary probe, it is essential that electrons can flow from the system common, through the material to the region of impact (to neutralize accumulated surface charge). Otherwise they must be supplied by coincident bombardment at an appropriate energy [48, 54], or by other means [58]. For negative ions, electrons must be able to escape from the region of impact, by flowing through the material, or escaping from the surface. Under some circumstances, it is possible to use negative ion bombardment in a self compensating way on insulators [49]. For fast atoms, a small amount of charging takes place through secondary electron loss [50], but this is not generally serious. In all cases, the instantaneous surface of the material will arrive at a surface potential which is dictated by the nett charge balance flowing into the surface, the conductivity of the escape path, and – most important for a semiconductor and usually forgotten – the reverse or forward bias potentials of any junctions across which charge flows in the neutralization process [51] (see Fig. 9). The surface potential will either stabilize when it can draw sufficient compensating charge from all sources, or will change until the material breaks down or the primary beam(s) cannot reach the surface.

Even with good charge compensation, or in quite well conducting samples, the surface potential can change unexpectedly by 10's V during analysis [51 – 53], especially when profiling across interfaces where large changes in conductivity occur. In addition, using an electron beam for charge compensation may be impossible where phenomena such as electron stimulated desorption occur [53]. Figure 10a shows typical SIMS secondary ion energy spectra for an atomic ion – $^{75}As^+$, and a molecular ion – $^{28}Si_2O^+$. The difference between the surface potential and the potential of the entrance plane to the mass spectrometer will determine the energy with which ions emitted from the surface enter the system. An incorrect or varying surface potential will have the effect of shifting or sweeping the energy spectrum across the instrumental energy window, thus producing unstable transmission. The effect will be most serious in a QMS, or a DFMS operating at high mass resolution. A DFMS operating at nominal resolution can



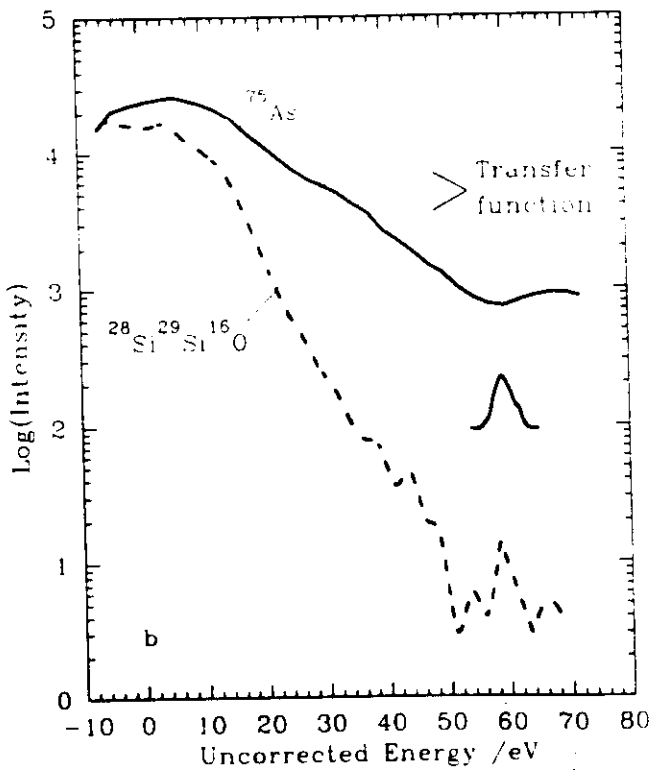


Fig. 10. a Atomic ($^{75}As^+$) and molecular ($^{28}Si_2O^+$) energy spectra for ions emitted from highly As doped polysilicon. b Energy spectra for the same species from a polysilicon on oxide region taken with electron beam charge compensation running. The way in which the Si_2O^+ signal from Figure 11b is used to generate a rapid charging correction function for As^+ (Eq. 7) and SI is depicted schematically

have an energy bandwidth of up to 150 eV and is relatively insensitive to the phenomenon. It is possible to measure the average surface potential during analysis, essentially by interleaving fast energy spectrum acquisition and ordinary data capture. Our EVA 2000 instrument has an effective system for doing this [51].

5.2 *Rapid or localized surface charging*

Samples containing very thin insulating layers can exhibit changes in surface potential which are too rapid for any stabilization technique to deal with.

Figure 11a shows raw SIMS data from a partially recrystallized highly arsenic doped polysilicon layer on 20 nm of silicon dioxide on silicon. The behaviour of the arsenic during recrystallization, and in particular, the existence of arsenic pile-up at the interface are to be determined from the analysis. The data were obtained in EVA 2000 using a 300 nA, 4 keV $^{16}O_2^+$ beam, normally incident, with a 5 μA, 800 eV e^- beam focussed into the SIMS crater for charge compensation. The As^+ channel was run using a negative target bias offset to suppress the $^{29}Si^{30}SiO^+$ interference at 75 daltons [55] and Si_2O^+ channels at 73 and 76 daltons with the same target bias were included as a diagnostic [52]. All channels exhibit spikes 1 – 2 orders of magnitude in height in the vicinity of the oxide layer. The behaviour of the Si_2O^+ channels shows immediately that a charging phenomenon is occurring [52]. The oxide layer should exhibit a lower Si_2O^+ yield than the surrounding silicon. A positive sweep in surface potential is occurring as the oxide is traversed, and this is moving the maxima in the energy spectra into the instrumental pass band. The behaviour of the As^+ signal is therefore spurious to some unknown extent. Figure 11b shows As^+ and Si_2O^+ signals where the charge compensation was better. The spike in the Si_2O^+ channel has reduced to a factor of 2 change.

The spikes in the Si_2O channels (believed at present to be wholly due to charging) have a different structure and a slightly different peak position from the that in the As channel. The As spike may well be composed of two merged features: a charging response, aligned in depth with the Si_2O^+, and a true concentration peak superimposed on this. This is suggested by Fig. 11a where there is a sharp increase due to charging at the point marked (i), followed by a decrease in slope at (ii), and then a doubling of the slope at (iii). This type of behaviour was observed consistently. From other work [56], involving the profiling of As implants straddling SiO_2 Si interfaces we know that the As response at this interface, for oxygen bombardment at normal incidence, is a doubling of slope with no discontinuity in the profile (unlike the case for boron [57]). This implies that any change in the degree of ionization of arsenic due to the change of matrix is balanced by the increase of sputter rate in SiO_2. For the present, we tentatively associate the start of the oxide with position (iii) in the profile, and calculate its width on the basis of the measured SiO_2 sputter rate with respect to silicon.

5.3 *A rapid charging correction procedure*

(i) It is assumed that all the variation in the Si_2O signals (Fig. 11a, b) is due to a variation in the local sample surface potential as the oxide is approached.

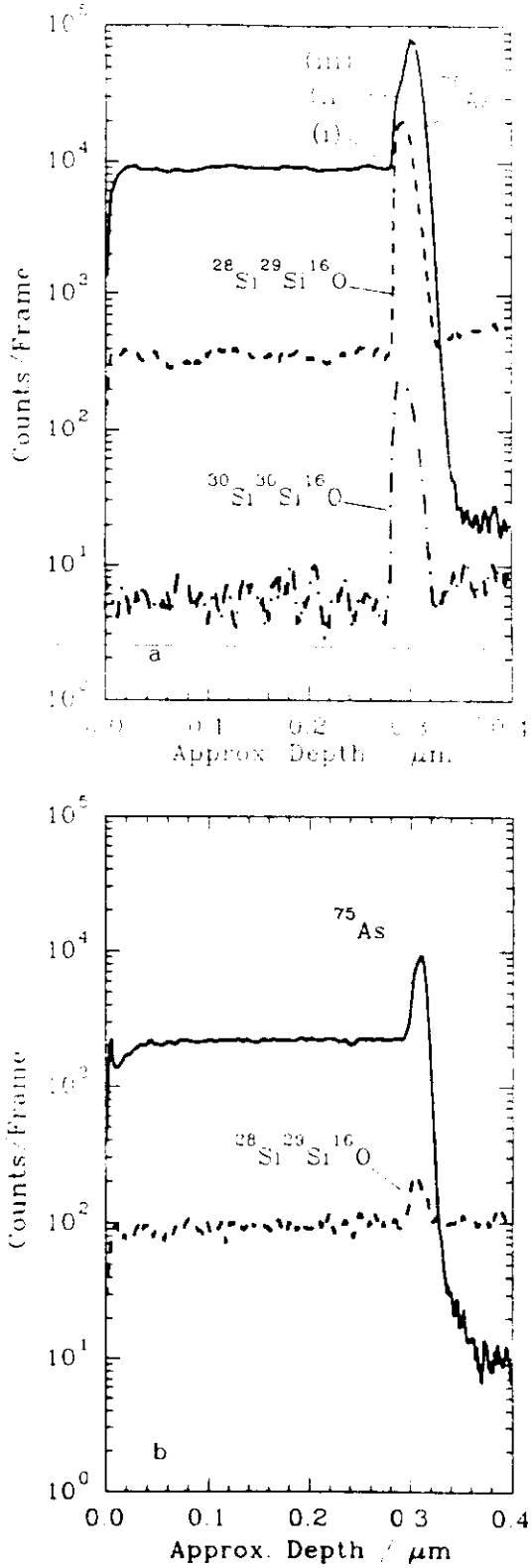


Fig. 11. a Raw SIMS profile (approximate depth calibration) for As in polysilicon on silicon dioxide. The two Si_2O^+ channels at 73 and 76 daltons show clear evidence of charging at the poly SiO_2 interface, despite careful charge compensation with an e^- beam. However, the shapes of the Si_2O peaks and the As peak are different, with three clear changes of slope occurring on the latter (see text). **b** The same sample profiled under best charge compensation conditions. The swing in the Si_2O^+ channel has diminished from a factor of 80 to a factor of 2.

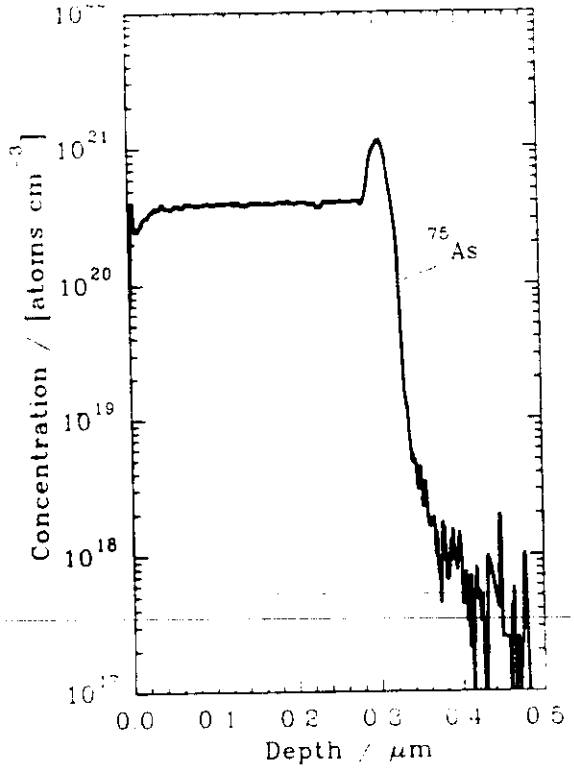


Fig. 12. Fully quantified profile from the data in Fig. 11b. The correction procedure (Eq. 8) has reduced the amplitude of the As^+ interfacial spike from a factor of 10 to a more reasonable factor of 3

(ii) Energy spectra are taken for Si_2O^+ and As^+ ions under conditions as close as possible to those during the profile. Figure 10b shows energy spectra from the poly oxide region with the electron gun on. Note the contrast with Fig. 10a.

(iii) The intensity variation of the Si_2O channel is used to deduce the magnitude of the surface potential change for each data point in the spike by comparison with the energy spectrum. Provided the charge compensation is good (e.g. Fig. 11b) the variation remains in the exponential part of the energy spectrum and is of the form:

$$V = -2.414 \ln(C^{Si_2O}(V)/C^{Si_2O}(V_w)) \quad (7)$$

where V is the surface voltage, and $C^{Si_2O}(V)$ is the corresponding Si_2O intensity. The point to point change in arsenic channel intensity is then deduced by measuring the intensity at voltage V on the As^+ energy spectrum, relative to the intensity at the instrumental working point. Again, for good charge compensation, the variation stays on the exponential part of the As^+ energy spectrum, and a complete digital correction function for the spike region can be found as follows:

$$C^{As}(V_w) = C^{As}(V) \exp\{0.618 \ln(C^{Si_2O}(V)/C^{Si_2O}(V_w))\} \quad (8)$$

where V_w is the working point of the instrument, $C^{As}(V_w)$ is the arsenic signal which would have been obtained if the instrument had stayed on the working point, $C^{As}(V)$ is the arsenic signal actually observed in the profile, and the constants are found from the slopes and intercepts of the energy spectra

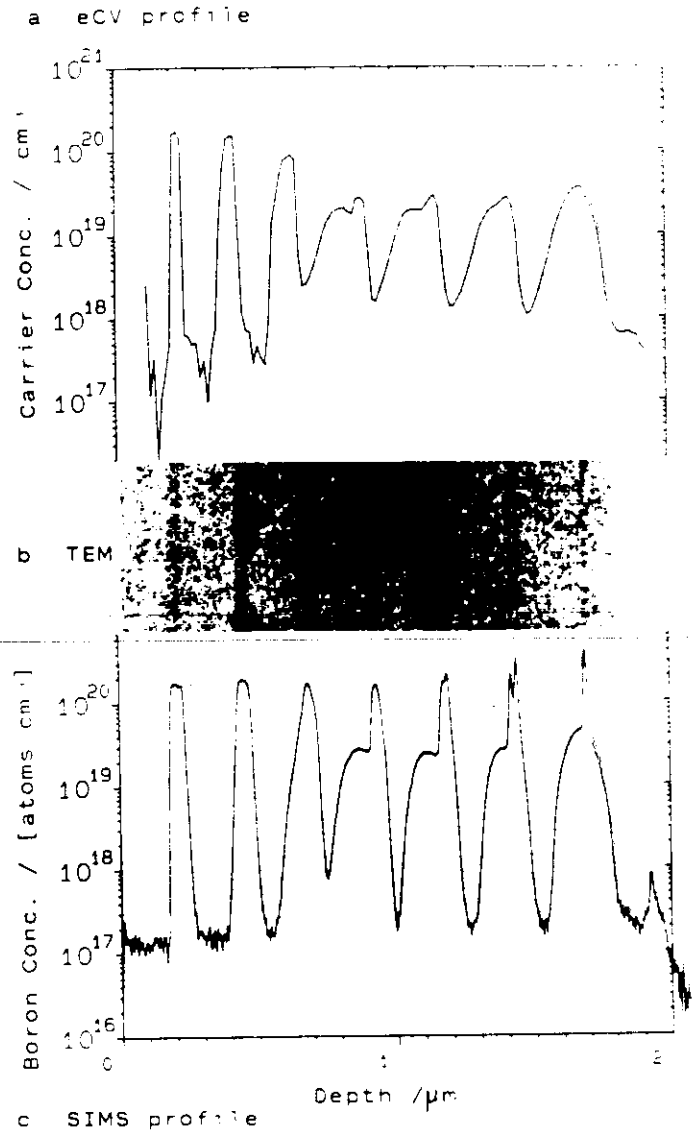


Fig. 13a–c. Boron modulation doped structure grown by MBE viewed with three different techniques. The parameter is the growth temperature °C for each boron period. Each period was grown with identical boron flux and time. **a** The eCV profile shows the carrier concentration decreases and broadens as the growth temperature increases. **b** The TEM shows the presence of the dopant clearly, probably because of the tensile strain, and the large concentration gradients. TEM reveals the presence of very narrow bands of precipitates at high growth temperatures. **c** A comparison between the SIMS and eCV profiles shows that the 450 and 600 °C layers are fully electrically active, whereas the higher temperature layers appear as surface segregation shoulders on very narrow concentration spikes whose fine structure agrees with the TEM data. Note that exact correspondence between SIMS and eCV data does not occur because the eCV depth scale varies non-linearly with the carrier concentration because of variations in the Debye length

Equation (7) has been used correct the processed profile presented in Fig. 12, which has also been corrected for the differential sputter rates in Si and SiO_2. Note that if there is a significant Si_2O interference contribution of 75 daltons, the above procedure will also correct for it implicitly

It is possible, that a modification of the above procedure could be used to correct for charging effects in images from laterally inhomogeneous samples.

6 Combinations of techniques

Complimentary information from chemical, structural and electrical characterisation techniques is required in most semiconductor investigations. Figure 13a shows an electrochemical capacitance-voltage (eCV) profile through a boron doping structure grown by Si-MBE. The structure was produced to investigate the effect of growth temperature on boron incorporation at high doping levels [62]. Each layer was grown using an identical boron flux from an atomic boron source. The carrier profile shows an apparent loss of dopant and an anomalous (i.e. not consistent with boron diffusion) increase in the width of the doped layers with increasing temperature. Figure 13c shows a SIMS profile through the same structure. At low growth temperatures, the boron is incorporated substitutionally into well defined, sharply bounded layers. As the temperature increases, however, the boron appears in increasingly complex sharp structures (see particularly the 800 layer) which appear in the correct growth positions, but which exhibit skirts due to surface segregation during growth. It is these broader regions which are electrically active. Figure 13b shows a cross-sectional TEM of the same sample. Surprisingly, all the doping layers are visible. The TEM and SIMS detail superimpose exactly. It is evident that in the 700–900 structures, most of the boron is in the form of precipitates (around 10 nm across). Very similar behaviour has been observed for high dose implanted boron, where the precipitates were identified as a boron silicide phase [63].

7 Conclusions

The analysis of modern semiconductor materials requires a sophisticated multi-technique approach. Moreover, each technique may need to be applied in a novel way to meet the stringent demands of ever smaller devices. No currently commercially available SIMS instruments are sufficiently flexible in their use of sample volume, their field of view, or their control over the primary probe to take advantage of most of the technique variants described herein without substantial modification. For example, DFMS machines have the highest useful yields, and therefore, potentially the greatest sensitivity. Since this performance is linked irrevocably to a small field of view (for a reasonable magnet radius) it is an essential requirement to scan the sample to perform the 2-D analyses described in section 4, or to acquire data from bevels which can be 10 mm long per μm of depth [60, 61].

Devices are now becoming so small that surface analytical techniques (with the exception of electron microscopy) can no longer access them directly. It is therefore important that communication between the analyst and the device engineer is established so that appropriate test areas can be incorporated into device wafers. Increasingly, special analytical structure will be required to obtain the necessary information.

Acknowledgements. Much of the material presented here was researched during the UK Alvey and LDS programmes and at the beginning of the Silicon Towards 2000 programmes. The author is grateful to SERC and DTI for their support. He would like to thank Chris Hill (Plessey Research) and Geoff Spiller (British Telecom Research) the Alvey 005 and 029 project managers for permission to use the 2-D profiling and altered layer material; Peter Augustus (Plessey Research) for all the TEM micrographs; Graham Cooke, Harvey Fox, Liz Clark and Bob Barlow (Warwick University) for some of the other figures, and Peter Marsh (STC Technology) for permission to use the As on Polysilicon data

References

1. Benninghoven A, Rudenauer FG, Werner HW (1987) Secondary ion mass spectrometry. Wiley, New York
2. Oechsner H (1983) In: Oechsner H (ed) Thin film and depth profile analysis. Topics in Current Physics. Springer, Berlin Heidelberg New York, pp 63–85
3. Gruen DM, Pellin MJ, Calaway WF, Young CE (1987) Proc SIMS VI. Wiley, New York, pp 789–796
4. Slodzian G (1980) In: Applied charged particle optics. Advances in Electronics and Electron Physics Supplement 13B. Academic Press, New York, pp 1–44
5. Steffens P, Niehuis E, Friese T, Greifendorf D, Benninghoven A (1984) Proc SIMS IV. Springer, Berlin Heidelberg New York Tokyo, p 404
6. Wittmaack K (1982) Vaccum 32:65–88
7. Lareau RT, Williams P (1986) Proc SIMS V. Springer Berlin Heidelberg New York Tokyo, pp 149–151
8. Hervig RL, Williams P (1986) Proc SIMS V. Springer Berlin Heidelberg New York Tokyo, pp 152–154
9. ~~Wittmaack K (1975)~~ Int J Mass Spectrom ~~Ion Phys~~ 17:39–50
10. Magee CW (1979) J Electrochem Soc 126:660–663
11. Williams P (1982) Appl Surface Sci 13:241–259
12. Wittmaack K (1981) Surface Sci 112:168–180
13. Reed NM (1989) In: Vickerman JC, Brown A, Reed NM (eds) Secondary ion mass spectrometry. Clarendon Press, Oxford
14. Levi-Setti R, Crow G, Wang YL (1986) Proc SIMS V. Springer, Berlin Heidelberg New York Tokyo, pp 132–138
15. Bernheim M, Blaise G, Slodzian G (1972/3) Int J Mass Spectrom Ion Phys 10:293–308
16. Klemperer O, Barnett ME (1971) Electron optics. Cambridge Monographs on Physics. Cambridge University Press, Cambridge
17. Clegg JB [redacted] (1984) Surface Interface Anal 6:162–166
18. Zeininger H, Criegern RV (1990) Proc SIMS VII. Wiley, New York, pp 419–424
19. Dowsett MG, Parker EHC, McPhail DS (1986) Proc SIMS V. Springer, Berlin Heidelberg New York Tokyo, pp 340–342
20. Wittmaack K (1985) Appl Phys A 38:235–252
21. Clegg JB (1986) Proc SIMS V. Springer, Berlin Heidelberg New York Tokyo, pp 112–114
22. Scilla GJ (1986) Proc SIMS V. Springer, Berlin Heidelberg New York Tokyo, pp 115–117
23. Homma Y, Ishii Y (1985) J Vac Sci Technol A3(2):356–360
24. Sigmund P, GrasMarti A (1981) NIM 182–183:25–41
25. King BV, Tsong IST (1984) J Vac Sci Technol A2:1443–1447
26. Armour DG (1988) Proc SIMS VI. Wiley, New York, pp 399–407
27. Augustus PD, Spiller GDT, Dowsett MG, Kightley P, Thomas GR, Webb R, Clark EA (1988) Proc SIMS VI. Wiley, New York, pp 485–488
28. Clark EA, Dowsett MG, Spiller GDT, Thomas GR, Augustus PD, Sutherland I (1988) Vacuum TAIP 38:937–941
29. Dowsett MG, Clark EA, Spiller GDT, Augustus PD, Thomas GR, Webb R (1988) Proc UK IT88:512–516
30. Clark EA, Dowsett MG, Augustus PD, Spiller GDT, Thomas GR, Sutherland I, Webb R, Harker A, Turner L, Collins R, unpublished
31. Wittmaack K (1987) Appl Phys Lett 50:815
32. Reuter W, Wittmaack K (1980) Appl Surface Sci 5:221–242
33. Wach W, Wittmaack K (1981) J Appl Phys 52:3341–3352
34. Wittmaack K, Wach W (1981) NIM 191:327–334
35. Treichler R, Cerva H, Hösler W, v Criegern R (1990) Proc SIMS VII. Wiley, New York, pp 259–262
36. v Criegern R ([redacted]) Presented at SIMS V (Abstract only)
37. Littlewood SD, Kilner JA (1988) J Appl Phys 63:2173–2179

38. Jäger HU, Kohler EA, Chater RJ, Hemment PLF, Peart RF, Reeson KJ (1988) Thin Solid Films 162:333 – 340
39. Williams P (1979) Surface Sci 90:588 – ■
40. Collins R (unpublished)
41. Taylor S, private communication
42. Boudewijn PR, Ackerboom HWP, Kemperers MNC (1984) Spectrochim Acta 39B:1567 – 1571
43. Vandervorst W, Remmerie J (1986) Proc SIMS V. Springer, Berlin Heidelberg New York Tokyo, pp 288 – 290
44. Dowsett MG, Jeynes C, Clark EA, Webb R, Newstead SM (1990) Proc SIMS VII. Wiley, New York, pp 615 – 618
45. Hill C (1990) In: Analytical techniques for semiconductor materials and process characterization (eds) Kolbesen BO, McCaughan DV, Vandervorst W. Electrochemical Soc, p 65
46. Cooke GA, Dowsett MG, Hill C, Clark EA, Pearson P, Snowden I, Lewis B (1990) Proc SIMS VII. Wiley, New York, pp 667 – 669
47. Cooke GA, Hill C, Dowsett MG, Pearson P, unpublished
48. Wittmaack K (1979) J Appl Phys 50:493 – 497
49. Andersen CA, Roden HJ, Robinson CF (1969) J Appl Phys 40:3419 – 3420
50. Van den Berg JA (1986) Vacuum 36:981 – 989
51. Dowsett MG, McPhail DS, Fox [illegible] Vacuum 36:913 – 916
52. McPhail DS, Dowsett MG, Parker EHC (1986) J Appl Phys 60:2573 – 2579
53. McPhail DS, Dowsett MG, Parker EHC (1986) Proc SIMS V. Springer, Berlin Heidelberg New York Tokyo, pp 343 – 346
54. Slodzian G (1987) Cameca News
55. Wittmaack K (1976) Appl Phys Letts 29:552 – 554
56. Dowsett MG, unpublished
57. Dowsett MG, King RM, Parker EHC, Mole PJ (1983) J Appl Phys 54:6340 – 6345
58. Werner HW, Morgan AE (1976) J Appl Phys 47:1232 – 1242
59. Liebl H (1978) Inst Phys Conf Ser No 38:260 – 281
60. McPhail DS, Dowsett MG (1987) Proc SIMS VI. Wiley, New York, pp 269 – 272
61. Gerodolle A, Corbex C, Poncet A, Pedron T, Martin S (1989) In: Software tools for process, device and circuit modelling. Crans W (ed) Boole Press, ■, pp 56 – 67
62. Parry CP, Newstead SM, Barlow RD, Augustus P, Kubiak RAA, Dowsett MG, Whall TE, Parker EHC, unpublished
63. Dowsett MG, Clark EA, Lewis MH, Godfrey DJ (1987) Proc SIMS VI. Wiley, New York, pp 725 – 728