THE APPLICATION OF DYNAMIC SIMS IN SILICON SEMICONDUCTOR TECHNOLOGY

by P.C. ZALM

Philips Research Laboratories, P.O. Box 80 000, 5600 JA Eindhoven, The Netherlands

Abstract
After a short introduction of the basic characteristics of secondary ion mass spectrometry (SIMS) as used in sputter depth profiling, a few selected applications of this technique are discussed. The distribution of the various dopants in a standard npn transistor illustrates its analytical potential. Next the role of SIMS in elucidating the mechanisms underlying impurity diffusion/migration is addressed. Finally an example is given of how the design of dedicated test structures may help to overcome the inherent limitations associated with dynamic SIMS, in the particular case of scaling down lateral dimensions of MOSFET structures into the submicron realm. The paper concludes by making an inventory of the suggested possibilities for future improvements in depth resolution within the constraints of the method.

Keywords: depth profiling, impurity, mass spectrometry, MOSFET structures, SIMS.

1. Introduction and outline

Modern low-temperature (hetero-) epitaxial growth techniques, that enable formation of extremely well-confined dopant profiles and multilayer structures, have matured to the level that these become a viable production option. In a parallel, but otherwise independent, development, in-depth dimensions in integrated circuit technology have become progressively smaller. Both trends require an increasingly precise process control which leads to more stringent demands on the supporting analytical techniques. Of the available characterization methods, dynamic secondary ion mass spectrometry (SIMS) is perhaps the only one that can potentially provide the detection limits (typically down to the ppm level or better) and depth resolution (presently at best about a few nanometers) that are currently needed. Others, such as high resolution X-ray diffraction and transmission electron microscopy, may provide informa-
tion about crystal quality and majority particle confinement but definitely lack the sensitivity at a sub-percent level of impurity atoms. Ultimately, of course, the electronic and/or optical properties determine the feasibility of an emerging technology but their investigation requires a near-completed device. In the research and early development stage SIMS is an indispensable tool.

Unfortunately, in spite of recent instrumental advances, there are fundamental limitations to SIMS. These originate from the inherent disruption created by the sputtering (i.e. removal of target atoms by energetic ion bombardment) process itself as well as by the limited lateral resolution. This necessitates the construction of model experiments to obtain an approximate solution for many practical problems which push the technique up to its limits. In this paper we will present a few typical examples from our own daily routine work to illustrate these remarks. First, however, a fairly condensed description of dynamic SIMS and the routine difficulties encountered will be given. For a detailed account of background, history, instrumentation, implementation etc. the interested reader is referred to the extensive textbooks available). Finally we will briefly indicate where room is left for improvement within the constraints of the technique.

2. A concise description of SIMS depth profiling

In SIMS ionized species liberated ("sputtered") from a target surface by bombarding it with a primary ion beam are separated according to their mass and a proportion are detected. As the exiting particles stem from the outermost few atomic layers, they carry information about the (local instantaneous) composition of the near-surface region. By monitoring the signal intensity for one or more mass(es) as a function of time during continuous erosion an in-depth distribution is obtained. This roughly summarizes the basic principle, but there are multiple practical snags that have to be taken into account as will be discussed now.

Usually only a minute fraction of the outgoing species (typically \(< 10^{-3}\)) leaves the surface in a charged state. By saturating the topmost layers with respectively oxygen or cesium, however, the formation/survival probability for positive or negative secondary ions can be greatly enhanced. The usual approach is to use an O\(^+\) or Cs\(^+\) primary ion beam. Further a very flat erosion front is needed in order to avoid smearing/averaging the depth information. To this end the primary beam is finely focused and rastered over part of the target. Contributions from the side walls of the sputtered crater, which carry information on shallower depths, are suppressed by allowing only secondary ions from the center of the crater to be analyzed. This can be done by either activating the detector exclusively when the primary beam passes this central
Dynamic SIMS in Si semiconductor technology

region or by secondary ion optical imaging. Most of the impinging primary ions are implanted in the target and it takes some time (the pre-equilibrium regime) before steady-state conditions build up as erosion proceeds. The incoming beam not only disrupts the target by this embedding of primary species, but as these come to rest in a sequence of collisions with target atoms the latter will be redistributed. This form of in-depth information smearing, called ion beam mixing, depends on the impact energy $E_i$ and angle of incidence $\theta_i$ (relative to the surface normal) of the primary ions. It is, very approximately, proportional to $E_i^{1/2} \cos \theta_i$. Thus, lowering of $E_i$ and/or more grazing $\theta_i$ will lead to a more truthful representation of the compositional depth distribution. Unfortunately the sputtering yield decreases with diminishing $E_i$ and also primary beam handling becomes difficult at low $E_i$ owing to space charge blow-up (Coulomb repulsion). The sputtering yield increases considerably with $\theta_i$, but this leads to a reduced incorporation of primary beam species which adversely affects the secondary ion information/survival probability. Also focusing and rastering becomes more awkward. So, effectively, all attempts to improve depth resolution lead to a degradation in signal intensity and, thus, in the capability to detect low concentrations. In addition the total time required for analysis is prolonged and this puts more stringent demands on stability. Consequently for each problem the analyst is forced to make a trade-off for the problem at hand between the various possibilities and limitations forced upon him by the degrees of freedom in the selection of primary beam parameters.

As for the secondary ions, one of the drawbacks of the technique lies in the fact that polyatomic cluster ions and or multiply charged elemental ions will be emitted, leading to severe mass interferences. Famous examples for dopants in silicon are the $^{30}\text{SiH}^{31}\text{P}$ and $^{29}\text{Si}^{30}\text{Si}^{16}\text{O}^{75}\text{As}$ interferences, while the detection limit for $^{10}\text{B}^+$ may ultimately be limited by $^{30}\text{Si}^{3+}$. There are two ways to circumvent this particular problem. One could strive for very high mass resolution, but signal intensity loss and instrumental stability demands limit this in practice to $M/\Delta M < 5000$, which is not always enough. Alternatively one may restrict the analysis to those secondary species that were ejected with a sufficiently high kinetic energy in the sputtering process. Clusters have a strongly reduced survival probability under those conditions, since when so much energy is transferred to them they will probably dissociate, and their contribution is thus suppressed (note: this trick fails with highly asymmetric-mass polyatomic ions). Also with this solution signal intensities will be reduced, however.

The most important aspect of SIMS is that the departing secondary particles can still exchange electrons with the receding surface when they are sufficiently close (\textless 1 nm above). The efficiency of this process strongly depends on their electronic configuration and that of the (local) target environment. Thus their
formation, survival and detection probabilities vary enormously with species and matrix type. A high signal intensity for a specific mass consequently does not necessarily imply a high concentration of the corresponding species. Further, a linear relationship between perceived signal intensity and actual (instantaneous, near-) surface concentration is only valid up to about a 1 at% impurity level, since otherwise the nearby presence of electronic-configuration-altering species will affect their detection efficiency. All this implies that calibration is cumbersome and requires standards. The best solution is to use a gauge implantation of the impurity under investigation into the same matrix as for the unknown sample to be calibrated and to measure its depth profile under identical conditions. Afterwards the perceived intensity is integrated over time, to get the total number of counts in the implant profile, and this is equated to the implanted fluence which is well known. This gives a direct conversion from counts to areal density (i.e. impurity atoms cm$^{-2}$). Time-to-depth conversion follows by determining the sputtered crater depth $d$ after termination of the measurement at time $t_{\text{stop}}$ and assuming a constant erosion rate (i.e. depth at time $t$ equals $dt/t_{\text{stop}}$). Experimental conditions must be chosen such that errors introduced by pre-equilibrium erosion effects are small. An additional advantage of gauge implants is that these immediately inform you about attainable detection limits, since in principle these extend only to finite depth and so the signal intensity should drop to zero. Less tedious concentration calibration schemes exist, but these lack the potential accuracy of the one presented here. No first principle calculations of sensitivities are available with a precision of better than an order of magnitude.

The target itself is a further factor in dynamic SIMS. Obviously flatness is important. A corrugated surface gives rise to considerable smearing of in-depth information. But even originally very smooth surfaces may rapidly develop topography under ion bombardment. Polycrystalline metals, and other ductile materials, do so excessively because of preferential sputtering of some crystallographic orientations which leads to facetting. This makes meaningful depth profiling beyond a few tenths of a micrometer virtually impossible (without further precautions, see Sec. 4) and restricts the technique to brittle materials (that amorphize readily under ion bombardment) such as semiconductors or oxides, nitrides etc., although even on these texturing can occur at considerable (i.e. of the order of micrometers) eroded depth. Also surface contamination may act as a local mask preventing erosion and thereby creating roughness, or, for instance in the case of a native oxide layer, delay the onset of steady-state erosion. Finally special precautions are needed when examining insulating targets. Flooding of the target with thermal electrons or concurrent bombardment with an electron beam can help. Sometimes it is
Dynamic SIMS in Si semiconductor technology

enough to coat the surface with a thin (~25 nm) conductive layer (often gold). Also the use of an O\(^{-}\) primary beam instead of the more common O\(_2^+\) ions alleviates the problem.

Modern equipment, combining many (if not all) of the above-mentioned facilities and options is commercially available. At our laboratory we use a CAMECA ims3f and an ims4f secondary ion spectrometer annex secondary ion microscope. The heart of this instrument consists of a doubly focused electrostatic sector energy/magnetic sector mass analyser. Characteristic for this type of analyser is the large extraction/bias voltage on the target (of the order of kilovolts). The primary beam comes in at off-normal angles and is decelerated (or accelerated depending on polarity) towards the sample. \(E_i\) and \(\theta\) are coupled (i.e. \(\theta\) becomes more glancing at lower \(E_i\)). The transmission (i.e. the accepted fraction of secondary ions generated) is very high but the sample size is restricted (of the order of cm\(^2\)). The extraction geometry allows for a secondary ion optical imaging system (i.e. position-sensitive detection with a lateral resolution down to \(~1\mu\text{m}\) irrespective of primary beam focus, so it is truly a microscope). The alternative widespread, commercially available, type of instrument design combines a band-pass energy filter with a quadrupole mass analyser (which limits mass resolution to about 1 amu). Here the target bias is small or zero and the transmission is low (owing to the narrow secondary ion ejection energy window acceptable to the quadrupole and the limited opening angle of the energy filter). \(E_i\) can be varied independently of \(\theta\) (often \(\theta \approx 0^\circ\), but the sample may be placed on a slanted holder or a tilting stage can be used). Potentially, large wafers (up to 6 in) can be examined. The set-up can be used as an add-on facility to an existing vacuum chamber, or in adverse situations there is room to install additional characterization techniques in the SIMS instrument.

With such modern equipment it is routinely possible to obtain (sub-) ppm detection limits for many impurity/matrix combinations. Species that do not ionize readily, such as the noble gases and, for instance, gold remain notoriously difficult. Poor detection limits must also be expected for those elements that abound in the residual, background, UHV contaminants such as (foremost) hydrogen, but also carbon, nitrogen and oxygen. Further, problems are to be expected with those materials used in primary-beam-defining apertures in the instrument (e.g. tungsten or tantalum), which will end up on the target under investigation. Last but not least, it should be remarked that in practice the dynamic range (i.e. the change in magnitude over which a concentration distribution can be followed) is ultimately restricted to less than about six decades in secondary ion intensity owing to redeposition onto the
analyzed area of particles that were first sputtered onto the surround (i.e. chamber walls and/or extraction lenses).

3. Selected application examples

3.1. Studies on bipolar transistors

The various processing steps that determine the distribution of the respective dopant atoms in standard very large-scale integrated circuit technology for the formation of a bipolar transistor can be summarized as follows. Here we take the example of a typical npn transistor.

I. Starting material is a commercial boron-doped ($\sim 1 \cdot 10^{15}$ B atoms cm$^{-3}$) silicon wafer. Antimony- (or arsenic-) doped areas with a sheet resistance of about 35 $\Omega/\square$ are introduced by implantation (to a top concentration of $1 \cdot 10^{18}$ Sb atoms cm$^{-3}$), before an epitaxial Si layer 0.5–1.5 $\mu$m thick is deposited. During subsequent high temperature processing steps, substantial indiffusion of Sb into this layer will occur.

II. Next, locally, the wafer is implanted with boron (typically at energies of the order of 25 keV to fluences of about $5 \cdot 10^{13}$ $B$ atoms cm$^{-2}$). Subsequently the wafer is annealed to restore the radiation damage and activate the B. This leads to redistribution of B and Sb.

III. Then, a polycrystalline Si layer is deposited onto the wafer surface (thickness $\sim 0.1$–$0.3$ $\mu$m) which is implanted with arsenic at low energies ($< 100$ keV to a fluence of $\sim 10^{16}$ As atoms cm$^{-2}$) followed by a heating step (at about 900–1100 $^\circ$C). This leads to an extremely rapid redistribution of As in the poly-Si and a much slower outdiffusion of the As into the mono-Si. Ideally the latter only affects the shallow-depth part of the $^{11}$B distribution.

In the above, emphasis has been placed on those processing steps that affect dopant profiles. In addition there are numerous (photo-) lithographic masking steps, which define the lateral confinement of the various implantations, etching steps (e.g. to allow contacting of the Sb-doped collector) and metallization steps (to contact the B-base and As-emitter) which have been omitted for simplicity. Usually the lateral confinement in an individual device structure is such that it does not allow for immediate inspection by SIMS (we will elaborate this aspect in Sec. 3.3). However, for characterization/analysis purposes a sufficiently large area can be prepared for inspection by leaving out some (or all) of these extra and intermediate steps.

An example of the resultant depth distributions of the dopants is given in fig.
Dynamic SIMS in Si semiconductor technology

Fig. 1. Relevant dopant profiles in an npn transistor. The approximate base width $W_B$ is indicated.

1. Roughly speaking the intercept of $^{11}$B and Sb profiles determines the position of the base/collector junction and the crossing of the As and $^{11}$B profiles the emitter/base junction position. The depth difference between those two points is, approximately, the base width $W_B$. Of course determination of the electrical properties of a device, or rather the whole array of those on the wafer, is essential to assess the quality, reliability and reproducibility of the various technological steps. Yet, in the early stages of process evaluation and also for failure analysis, SIMS is an invaluable tool. Errors in parameter selection can be signalled when the penalty in terms of time and investment losses is still low. To name just one issue of practical importance: the extent of outdiffusion of Sb and the remaining depth of undoped epi-Si. These parameters are of critical importance for the electrical characteristics of high-speed transistors with very small vertical dimensions.

Further, it was found that an imperfect combination of poly-Si thickness and As implantation energy led to considerable redistribution of the $^{11}$B to larger depths. This is caused by creation of crystal damage by energetic As penetrating the mono-Si. Upon thermal processing this damage anneals out releasing Si interstitials which will rapidly diffuse into the crystal and replace part of the substitutional $^{11}$B atoms which then start to move (we come back to this in more detail in the next subsection). This example is illustrated in fig. 2 where, for clarity, only the $^{11}$B relocation is shown. Note that the emitter/base junction has stayed in place but that the base/collector junction shifts to considerably greater depth. As a byproduct of this particular study we would
like to draw attention to the very-near surface $^{11}$B “tail” in the poly-Si after As implantation but prior to anneal. Although unimportant for the functioning of the device this spurious tail was reason for some concern. After a detailed investigation of this phenomenon we could prove that it was not simply surface contamination but originated from so-called cross-contamination of the incident As ions in the accelerator. That is, the implantation machine had previously been used for $^{11}$B doping and apparently some of this ended up on the inside of the implanter from which it became liberated later during the As implantation and was deposited onto and implanted (at, on average, low energy because it can stem from anywhere in the accelerator tube) into the poly-Si.

In the most recent extensive study on npn bipolar transistors, the arsenic in step III was replaced by phosphorus. For details of the device aspects of this swap the interested reader is deferred to ref. 5. Here it suffices to say that it results in a more efficient emitter while the series resistance of the emitter remains sufficiently low. Here, the complexity for dynamic SIMS lies in the accurate determination of the P depth distribution, both in terms of resolution as well as detection limit. Traditionally the $^{30}$SiH/$^{31}$P interference problem is tackled by employing Cs$^+$ primary ions in combination with negative secon-
dynamic ion detection together with high mass resolution \((M/\Delta M \simeq 3500)\) in the case of non-UHV instruments (i.e. where background pressures are above \(10^{-9}\) Torr such that the residual gas consists mainly of hydrogen which adsorbs on the surface of the target under investigation). Unfortunately, efficient Cs\(^+\) extraction, as well as acceleration to the target in some types of instrument (see Sec. 2), implies considerable incident ion energies and consequently quite some ion-beam-induced dopant atom relocation (mixing). On the basis of theoretical considerations we were able to show\(^6\) that the broadening by mixing is, very approximately, proportional to \(E_i^{1/2}\cos \theta_i\) (with \(E_i\) and \(\theta_i\) impact energy and angle of incidence, respectively) with a prefactor largely independent of primary ion and impurity type but quite sensitive to the matrix (majority species) in good qualitative agreement with experimental observations. We therefore attempted to use low energy O\(^+\) primary ions at glancing \(\theta_i\) and found that this, in combination with concurrent oxygen flooding of the target to enhance the positive secondary ion yield (and, of course, high mass resolution) generated excellent detection limits and adequate depth resolution\(^7\). Next the extent to which oxygen bleed-in affects the absolute depth resolution had to be established. Theoretical and semi-empirical evidence was conflicting in that it was on the one hand argued that the reduction in erosion rate upon oxidation necessitated an increased flux of primary ions to attain the same depth whereas on the other hand the associated swelling upon oxidation would seem to indicate that more of the mixing takes place outside the depth-of-interest. We found\(^8\) that the relation between depth resolution and the parameters \(E_i\), \(\theta_i\) and impurity type is very complicated, but that around \(E_i \simeq 2-3\) keV and \(\theta_i \simeq 50^\circ-60^\circ\) reliable and reproducible results can be achieved in conjunction with oxygen exposure. Thus SIMS depth profiles for the P-emitter transistor could be realized that were sufficiently reliable to be used directly in a device modeller to predict its electrical properties accurately. (Note that this necessitates measurements under optimal conditions for each particular dopant as well as a “good depth-defining/poor detection limit” quality determination of all three (P or As, \(^{11}\)B and Sb) distributions simultaneously to establish as accurately as possible their relative positions).

3.2. Studies on dopant mobility

In the previous subsection it was seen (fig. 2) that the creation of Si interstitials (and vacancies) invokes a redistribution of dopant atoms upon anneal. This is a well-known phenomenon that manifests itself when for instance a dopant implantation/thermal activation cycle is applied to a single-crystal Si wafer. The as-implanted depth distribution changes dramatically on a very
Fig. 3. SIMS depth profiles of a series of B "delta"-doped layers in epitaxial Si grown by atmospheric pressure CVD at 750°C. Thin full line, as grown; broken line, after 20 min anneal at 850°C in N₂; thick line, after 20 min anneal at 850°C in O₂.

short time scale (of the order of minutes) when the sample is heated. This so-called transient diffusion is caused by relocation of dopant atoms when the fast-moving Si interstitials sweep through the lattice. Only after this migration has progressed to well outside the doped region does the dopant diffusion take on its normal and much slower, thermally activated, character. It is a particularly severe manifestation of the consequences of radiation-induced damage that limits the attainable depths of shallow junctions made via implantation. It is further an established fact that self-interstitials are also injected during (local) oxidation of crystalline Si leading to a similar broadening of dopant depth distributions. An example of this so-called oxidation-enhanced diffusion (OED) is given in fig. 3. Here SIMS depth profiles are shown of an Si sample that contains six B-doped "delta" layers (i.e. confined to almost a single atomic plane), both as grown by low temperature (~750°C) atmospheric pressure chemical vapour deposition, and after a 20 min anneal at 850°C in either a pure nitrogen or a pure oxygen ambient. In N₂ the broadening is small and solely due to thermal diffusion, but in O₂ the reduction in modulation owing to OED is considerable.

The apparent broadening of the deeper lying B deltas in fig. 3 in the as-grown sample (thin full line) is probably a SIMS artefact, although the possibility that diffusion is somewhat higher during Si deposition cannot be ruled out. Clearly one would like to know the true reason. There are, however,
Dynamic SIMS in Si semiconductor technology

limits as to what can be done with SIMS profiling. Under the most optimal circumstances, i.e. a primary $\mathrm{O}_2^+$ ion beam with $E_i \approx 1.5$ keV and $\theta_i \approx 75^\circ$, the best result to date for the (apparently) most ideal delta yielded a full width at half-maximum of 2.5–3.0 nm and exponential leading and trailing slopes with characteristic lengths of 0.4 nm and 1.0 nm respectively. Such values cannot be maintained to eroded depths in excess of about 0.1 $\mu$m, owing to erosion inhomogeneities caused by the poor focusability of the slow primary beam. Also detection limits are not very good since the signal intensity is correspondingly low. At somewhat higher $E_i$ ($\approx 3$ keV) and $\theta_i$ ($\approx 50^\circ$) beam handling is far more easy and resolution can be kept constant to depths of around 1 $\mu$m in Si. Unfortunately ion-beam-induced mixing will then approximately have doubled the above figures. Furthermore, at larger depths, surface roughening may become the dominant factor and dictate the perceived resolution. This has probably occurred in fig. 3.

Obviously it is extremely important to understand the mechanisms underlying the anomalous diffusion behaviour as it will affect the electrical properties of future, and perhaps already present, device generations. The realization of near-perfect delta-doped layers has proved to be an indispensable aid in this research, because these enable visualization of even very small effects. On the basis of B deltas in Si grown by molecular beam epitaxy in our laboratory, which were subjected to a variety of heat and low energy (Si-) implantation treatments and then profiled with SIMS, it was verified that 9):

(i) substitutional B is knocked out of its lattice position by interstitial Si atoms and becomes highly mobile,

(ii) although the generation probability decreases at low temperatures the B migration step length (the average relocation distance before it is trapped again) increases.

These and other findings will provide the input parameters necessary for a proper modelling of (anomalous) boron diffusion. Of course one would have to repeat the above-sketched experiments for all other dopant types used in Si technology, which may diffuse by (slightly) different mechanisms. At this moment it is not yet feasible to fabricate delta layers of, for example, P with the same quality as those shown in fig. 3. Yet such samples are highly desirable since, in order to elucidate the processes contributing to dopant migration, it is often essential to evoke minute alterations of the depth distributions by fairly extreme experimental conditions (low temperatures, brief thermal excursions, low implant fluences etc.) to enable separation of the various contributing and often competing mechanisms of diffusion.
3.3. The challenge of small lateral dimensions

One of the key elements in failure analysis and also in the characterization of realistic devices is the ability meaningfully to address individual features of minute lateral dimensions. For rather trivial reasons dynamic SIMS scores poorly in this respect. First of all it is virtually impossible to focus a sufficiently high density (to get a fair erosion rate) primary ion beam at low enough energy (to obtain good depth resolution). But even with this problem solved, there is still simply not enough material available. A minimal requirement for a depth profile would be one data point per 10 nm. For a $1 \times 1 \mu m^2$ analyzed area this implies that only a few times $10^8$ matrix atoms are liberated for each data point. Since the dopant levels of interest rarely exceed 0.1–0.01 at% and for typical ionization/collection efficiencies of the order of $< 10^{-3}$ for most secondary ions, even in favourable cases a dynamic range of only two decades and a detection limit of $> 5 \times 10^{17}$ impurities cm$^{-3}$ (i.e. well above the common junction level) may at best be hoped for. Considerable effort has gone into the development of post-ionization methods to improve statistics, but even when these eventually become successful only half of the problem is remedied. Practical devices often exhibit considerable height differences in the area of interest. As we have demonstrated conclusively\(^\text{10}\), this leads to distortions of a profile on a depth scale commensurate with the minimum of height or lateral dimensions.

An alternative approach, adopted independently in our laboratory and other R&D facilities\(^\text{11}\), is to incorporate dedicated test modules in the standard mask package that will allow for inspection by SIMS. Flat $200 \times 200 \mu m^2$ patches exposing the equivalent of the fully processed emitter, base and collector in our bipolar transistor/integrated circuit production line are routinely made available for a precision analysis in this way. Unfortunately it turns out that sometimes the lateral confinement itself plays a crucial role in the functioning of a device. In such cases the only reliable road to success, albeit a cumbersome one, is to design model structures that enable extraction of the desired parameters/dopant profiles in some indirect form. An example is depicted in fig. 4, which derives from a study to scale down MOSFET dimensions into the deep submicron level\(^\text{12}\)). The essential processing steps, from the viewpoint of dopant distribution (again all masking steps will be ignored), are:

I. Starting point is a uniform implantation of $^{11}$B in an Si wafer;
II. Growth of a thin ($\sim 10–15$ nm) gate oxide, which will be removed locally later to allow source and drain contacting;
Dynamic SIMS in Si semiconductor technology

Fig. 4. $^{11}$B depth profiles (left) in test structures (right) used to elucidate the influence of poly-Si gate sidewall oxidation on channel dopant distribution when scaling down the lateral dimensions of a MOSFET device.

III. A dopant activation/damage restoration anneal cycle;
IV. Deposition and structuring of a polycrystalline Si gate;
V. Sidewall oxidation of this gate (its top surface will likewise become oxidized, but this is removed prior to contacting).

It was observed that the last step had a profound effect on the electrical characteristics when lateral dimensions became small ($\sim 0.25 \mu m$) and led to anomalous short-channel behaviour. The OED for the shallow $^{11}$B implantations used was apparently different from that reported in the literature for deep implantations on which the original modelling was based. It resulted in a merging of the dopant atom distributions in the source and drain regions across the channel. This example stresses once more the necessity of improving our understanding of the dopant migration mechanism as discussed in the previous subsection. In the meantime a solution had to be sought for the particular problem at hand to enable correct modelling of the gate-source/drain overlap in process simulations.

In a close collaboration of process modeller, IC technologist and SIMS analyst a test structure was proposed that would allow for a mimicking of the dopant redistribution associated with poly-Si gate sidewall oxidation involving some computer-aided deconvolution scheme. The structure made consisted of a series of gratings, each $0.8 \times 0.8 \text{mm}^2$, formed by an array of poly-Si stripes (of the appropriate height) deposited onto a Si wafer. Into this wafer beforehand a uniform shallow $^{11}$B implant had been applied and activated and also a thin (gate-) oxide had been grown. Per grating the poly-Si stripe width

Philips Journal of Research Vol. 47 Nos. 3-5 1993
S and their spacing $S$ were fixed, but in between gratings $S$ varied from 0.1 to 2 $\mu$m in steps of 0.1 $\mu$m. A schematic representation of a cross-section of the structure is given on the right-hand-side of fig. 4 (A). Next the poly-Si stripes were oxidized (B) and all oxide and remnant poly-Si were removed by a combination of mechanical abrasion and wet chemical etching (C). Finally, on the flat Si surface, the $^{11}$B depth distributions underlying each former grating were determined with SIMS. A clear dependence of the magnitude of diffusion on lateral feature size could be established (cf. left-hand-side of fig. 4 for some typical values of $S$). From the whole data array, as well as another one with asymmetric stripe/separation-width combinations, the OED effect in a true single MOSFET configuration can be retrieved. Thus, admittedly in a substantially circumferential way, it may be possible to evaluate future devices that are no longer directly accessible for SIMS analysis. Fairness dictates that others too have proposed\textsuperscript{13}) to tackle this issue in a conceptually similar manner (i.e. essentially to replace a single feature by a repetitive array to overcome the problem of poor counting statistics).

4. Future prospects

In the previous section a few examples have been given of where and how dynamic SIMS can contribute in (silicon) semiconductor research and technology. The common denominator in these case studies, and in many others left unmentioned, is the problem associated with the ongoing drive for smaller dimensions. It has been shown that this no longer allows for routine analysis, but that often intelligent and complex solutions have to be sought to circumvent the inherent limitations of the technique. Below we will discuss a few developments, which are largely still in the laboratory stage, that may alleviate part of the SIMS constraints. Hopefully some of these will prove sufficiently successful to carry this characterization method (deep) into the next century.

Other than the approach laid down in Sec. 3.3, little can be done about the problems associated with small lateral dimensions (i.e. lack of counting statistics). A beautiful tool to obtain high-resolution images is presently being constructed by Winograd and coworkers of Pennsylvania State University, although other groups progress along similar lines. He combines a finely focused ($\phi \sim 50$ nm) Ga\textsuperscript{+} primary beam from a liquid-metal ion source with laser-induced multiphoton (resonant) ionization to (potentially state selectively) convert ejected neutrals to detectable ions in a position-sensitive way. But this method is, loosely speaking, only a static SIMS technique in that the sub-surface disruption is so severe that meaningful depth information cannot be obtained.

The central issue that is still open for further improvement is that of depth
resolution. A dynamic SIMS instrument has just been marketed (the S1030, manufactured by Kratos Ltd., Manchester, UK) that allows for a further reduction of primary ion impact energy to about 0.5 keV at angles of incidence of around 60°, where the problems of the considerable increase in beam handling difficulties seem to have been overcome. A further improvement introduced recently\(^4\) is that of sample rotation during analysis. This considerably reduces resolution degradation associated with bombardment-induced surface roughening. It has paved the path for meaningful SIMS depth profiling of polycrystalline metal (multilayer) structures, a possibility hitherto considered to lie beyond the horizon. A third promising candidate appears to be the use of reactive gases to flood the target and enhance the sputter yield\(^5\) so that reduced primary ion fluxes (i.e. less extensive radiation damage and beam mixing) are needed to erode to a given depth. The major drawback of this solution is that the appropriate gases (mostly halogens) attack not only the target, but e.g. vessel walls and vacuum pumps as well. Thus a complete redesign of instruments will be necessary. Of course all three propositions can in principle be combined. All in all an improvement by (optimistically estimated) a factor of 2–5 in resolution for shallow features and certainly better for deep impurity distributions may be hoped for.

Finally one other line of approach is worth mentioning. It has been proposed\(^6\) to take the experimental result for a delta-function distribution as the instrumental response function. This is then used to estimate the influence of the measurement parameters on an unknown depth profile of the same impurity/matrix combination determined subsequently under identical conditions. To this end one has to convolute the empirical response function with a plausible trial input distribution to mimic the actually observed result and employ some matching criterion. The success of this scheme is rather limited for two reasons. First, only for a few selected dopant/semiconductor pairs have near-perfect delta-like distributions been realized. And, secondly, convolution will not be able to provide reliable information about features on a scale comparable with the dimensions measured for the alleged delta. So resolution improvement is limited to less than a factor of 2. A much more universal, semi-theoretical, approach has been suggested by a group at Salford University\(^7\). They applied a sophisticated mathematical model that directly simulates the depth profiling process as it proceeds by taking into consideration the combined effects of sputtering, mixing and primary ion incorporation. The relevant input quantities needed to solve the integro-differential equations numerically, are derived from Monte-Carlo type calculations using the binary-collision approximation which are fitted to agree with experimental observables such as the (partial) sputter yields. On the outside it looks as if there are so many adjustable
parameters that reproduction of the experimentally observed profiles becomes a triviality, but this appearance is utterly misleading. In fact only a few (about 3) suffice to account for a large variation in measurement conditions for any given impurity/matrix combination. Once fully operational, and provided the promised generality and flexibility come indeed true, this may well turn out to be a most valuable aid in SIMS depth profile correction.

REFERENCES


Author

P.C. Zalm: M.Sc. (mathematics, 1973; physics, 1974), Utrecht State University, The Netherlands; Ph.D. Utrecht State University, 1977; Philips Research Laboratories, Eindhoven, The Netherlands, 1978-. In his thesis work he was concerned with electromagnetic moments of ultrashort-lived excited nuclear states. At Philips he first participated in the high-definition television research programme, then worked on low-energy ion–solid interactions, later on silicon molecular beam epitaxy and subsequently on high temperature superconductivity. Since 1988 he has been active in the Structural Analysis department where he carries out secondary ion mass spectrometry investigations in a supportive role to the laboratory programme.