



Radiation hardness of silicon detectors – a challenge from high-energy physics

G. Lindström*, M. Moll, E. Fretwurst

II. Institut für Experimentalphysik, Universität Hamburg, Luruper Chaussee 149, D-22761 Hamburg, Germany

Abstract

An overview of the radiation-damage-induced problems connected with the application of silicon particle detectors in future high-energy physics experiments is given. Problems arising from the expected hadron fluences are summarized and the use of the nonionizing energy loss for normalization of bulk damage is explained. The present knowledge on the deterioration effects caused by irradiation is described leading to an appropriate modeling. Examples are given for a correlation between the change in the macroscopic performance parameters and effects to be seen on the microscopic level by defect analysis. Finally possible ways are out-lined for improving the radiation tolerance of silicon detectors either by operational conditions, process technology or defect engineering. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Radiation hardness; Silicon detectors; Nonionizing energy

1. Introduction

Silicon microstrip and pixel detectors are widely used for tracking purposes in a variety of present and future HEP experiments. Their unsurpassed virtues as to signal-to-noise ratio for mip, the achievable high spatial resolution, fast direct charge readout and low-voltage operation in ambient atmospheres while employing only minimal material thickness of well below 1% of radiation length have made them the ideal choice for meeting especially the future requirements for LHC [1,2]. However, within these applications the detectors will face extremely intense hadronic radiation flu-

ences of up to 10^{14} cm^{-2} per operational year, causing appreciable deterioration of the detector performance (Section 2). The radiation tolerance of these devices is therefore of supreme interest and had led to a broad line of investigations worldwide. Along with the developments for the LHC experiments these projects were coordinated within the ROSE-collaboration CERN-RD48 [3,4].

It is well known, that the main damage effects in the performance of silicon particle detectors are caused in the silicon bulk, hence the presented overview concentrates on these issues only. As a basis for the needed evaluations the expected hadron fluences are given and the NIEL scaling (Non Ionizing Energy Loss) is discussed, used successfully for the normalization of different irradiation experiments as well as for projecting the results

*Corresponding author. E-mail: gunnar@sesam.desy.de.

to any radiation environment (Section 3). The main macroscopic effects of changes in the detector performance consist of a fluence proportional increase in the leakage current, a dramatic change of the depletion voltage, needed to maintain a full sensitivity of the whole detector thickness and the damage-related decrease of the charge collection efficiency. All these effects have been measured so far in a number of impressive systematic studies, covering the fluence dependence for protons, neutrons and pions of different energies and including also the annealing effects observed after irradiation. A comprehensive modeling had been developed allowing reliable projections for prolonged operations in the forthcoming experiments (Section 4). Such estimates have shown that especially for the areas close to the interaction points in the LHC collider experiments the radiation tolerance of present state of the art detectors may not be sufficient for a 10 yr operation, which is however needed for cost effectiveness. The search for more radiation hard material is therefore inevitable. A successful strategy leading to improved quality in this respect can only be based on the knowledge of defect properties being responsible for the changes in the detector performance followed by dedicated search for those key issues which could result in an adequate defect engineering. Several groups have therefore studied a possible correlation between the macroscopic radiation damage effects and defect analysis using e.g. DLTS (deep level transient spectroscopy) and other microscopic tools (Section 5). Both the obtained reliable knowledge of macroscopic effects and also preliminary findings from microscopic defect analyses have led to projections for possible improvements of the radiation tolerance under LHC operational conditions, to be discussed in the last Section 6.

2. The challenge: high-intensity hadron fluences in future HEP experiments

As a specific example of the challenge presented for the application of silicon particle detectors in a forthcoming high-energy physics experiment the ATLAS experiment is taken, presently under development for the large hadron collider LHC at

CERN [1]. Here for the innermost part of the tracking area around the interaction point pixel detectors are used, exposed to a 1 MeV equivalent neutron fluence of up to $5 \times 10^{14} \text{ cm}^{-2} \text{ yr}^{-1}$ (maximum). For the strip detectors used further outside (semi-conductor tracker SCT) covering an overall area of 60 m^2 the maximum fluence will be $2 \times 10^{13} \text{ cm}^{-2} \text{ yr}^{-1}$. The very high cost for the silicon instrumented tracker has led to the request that the detectors should be fully operational for at least 10 yr and hence it has to be guaranteed that the detectors will have a sufficiently high radiation tolerance.

A few points have to be added in view of the problems arising from the irradiation during the experiments. First of all both the fluence calculations and the results from damage experiments are accurate only within some larger errors, which add an appreciable amount of uncertainty to any operational projection. For example for ATLAS this was accounted for by adding 50% on top of the nominal radiation levels as a rough safety margin.

A substantial cooling of the detectors down to well below freezing point has furthermore proved to be inevitable. This is needed not only during the operational beam periods (10^7 s actual beam during 7 months each year), primarily for reducing the leakage current and thus assuring the required signal-to-noise ratio, but also during both beam and stand by periods, since otherwise the adverse effects of anti-annealing, increasing the needed operational voltage, could not be tolerated (see Section 4). In fact it has been shown that the yearly maintenance period in which detectors have to be warmed up to room temperature should not exceed about 2 weeks. Maintaining such a scenario during the total operational period of 10 yrs is in itself an engineering problem, not easily met.

While operation under these conditions is a necessity for keeping the radiation damage effects small enough within the design goals, it should be clear that no accidental and uncontrolled beam dump may be allowed at any time!

Similar evaluations of damage-connected problems are relevant for the CMS experiment at LHC, but even more so for HERA-B and LHC-B, where silicon tracking detectors are to be placed extremely close to the fixed target facing a huge pion flux

which most likely will lead to detector replacements every operational year [5,6].

3. NIEL scaling of bulk damage

The bulk damage produced in silicon particle detectors by hadrons (neutrons, protons, pions, others) or higher energetic leptons is caused primarily by displacing a Primary Knock on Atom (PKA) out of its lattice site resulting in a silicon interstitial and a left over vacancy (Frenkel defect). Both can migrate through the silicon mono-crystalline lattice and may finally form complex point defects together with impurity atoms being resident in the silicon. However, the original PKA can only be displaced if its energy is higher than the binding energy of ≈ 25 eV. The energy of a recoil silicon PKA or any other residual atom resulting from a nuclear reaction can of course be much higher. Along the paths of these recoils the energy loss consists of two competing contributions, one being due to ionization and the other caused by further displacements. At the end of any heavy recoil range the nonionizing interactions are prevailing and a dense agglomeration of defects (disordered region) is formed. Both, point defects along the particle paths and such clusters at the end of their range may be responsible for the various damage effects in the bulk of the silicon detector but ionization losses will not lead to any relevant changes in the silicon lattice. Hence the bulk damage depends exclusively on the Non Ionizing Energy Loss (NIEL) and it has widely been verified that it is strictly proportional to this value (NIEL scaling hypothesis). As an equivalent expression for the NIEL value the displacement damage function $D(E)$ is used here. It can be calculated employing the individual reaction cross section, the energy distribution of recoils produced by that reaction, the partition between ionizing and nonionizing energy loss of the recoils and finally summing over all different reaction channels possible for the initial particle and its energy. For neutrons a comprehensive description can be found in Ref. [7] and the whole issue of NIEL scaling for all relevant particles will be published shortly [8]. A systematic comparison of the various damage effects observed

experimentally in silicon for different particles as function of energy has been used for selecting the most reliable evaluations of the displacement function. They are displayed in Fig. 1 [9–12] and specifically exclude or supercede older tabulations [13,14]. More details can be found in Refs. [7,8,12] and literature cited there.

A few details should however be mentioned. The minimum energy of neutrons transferring enough energy for displacements by elastic scattering is 190 eV. Yet, as can be seen from Fig. 1, the damage cross section rises below that value with decreasing energy. This is entirely due to neutron capture, for which the emitted gamma rays result in a recoil energy of about 1 keV, much higher than the displacement energy of 25 eV (see above). Therefore this part of the displacement damage cross section, misleadingly described as the “disappearance part” in the original paper published by Lazo [14], can by no means be neglected in general. For spectra to be expected in the inner detectors of the LHC experiments it does however not play a significant role. For neutrons with energies in the MeV range an increasing number of nuclear reactions opens up adding to the displacement function. Up to about 20 MeV an accurate and almost complete data basis does exist which was used as a reliable source for damage calculations [9]. At higher energies theoretical approaches had to be used [10]. The proton damage function is on the other hand

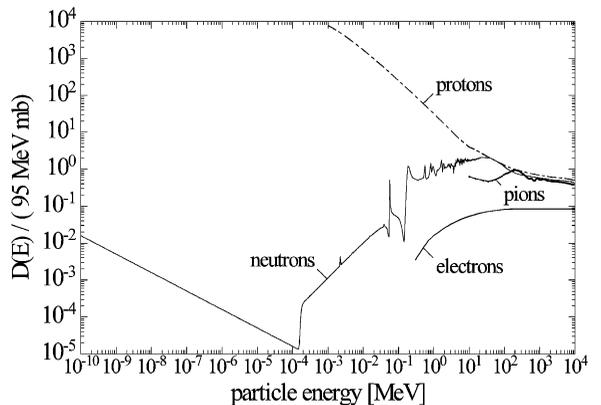


Fig. 1. Most reliable evaluation of the displacement damage functions for neutrons, protons, pions and electrons as recommended by the ROSE collaboration [15] (see text for details).

dominated by Coulomb interaction at lower energies and is therefore much larger than that for neutrons [11]. For very high energies in the GeV range both damage functions approach a common value. Here the coulomb contribution becomes almost negligible and the nuclear reactions are practically the same for neutrons and protons. Pion interaction is largely influenced by the delta resonance around a few hundred MeV but in the high energy limit the damage function tends to be about 2/3 (quark weighting factor) that of neutrons resp. protons [12]. Finally it should be mentioned that the damage displacement cross section for protons, as given in Fig. 1, is now undoubtedly confirmed even at high energies where the hardness factor for 24 GeV/c protons had been experimentally measured as being 0.51 in good agreement with the calculations [12,16].

The damage efficiency D_{eff} of any particular particle source with an energy distribution of its fluence given by $\Phi(E)$ can then be calculated using the appropriate damage displacement function $D(E)$ by the following formalism. It is common practice to relate D_{eff} , resp., the total fluence Φ_{tot} to the displacement damage cross section and the equivalent fluence of 1 MeV neutrons, producing the same damage, employing a “hardness factor” κ which is characteristic for the source

$$D_{\text{eff}} = \int D(E)\Phi(E) dE = D_{\text{neutron}}(1 \text{ MeV}) \cdot \Phi_{\text{eq}} \quad (1)$$

with

$$\Phi_{\text{eq}} = \kappa \Phi_{\text{tot}} = \kappa \int \Phi(E) dE \quad (2)$$

According to an ASTM standard the displacement damage cross section for 1 MeV neutrons is taken to be 95 MeVmb [17]. Following the above equations the hardness factor for a monoenergetic source is just given by $\kappa = D(E)/D_{\text{neutron}}(1 \text{ MeV})$ and in the general case we can derive κ from the following folding equation:

$$\kappa = \frac{\int D(E) \cdot \Phi(E) dE}{D_{\text{neutron}}(1 \text{ MeV}) \cdot \int \Phi(E) dE} \quad (3)$$

For any reliable comparison of irradiation experiments carried out with different sources an

exact knowledge of the different energy distributions is therefore essential. The same applies of course to the projections for the LHC environment, for which the spectral distribution of various particle components can only be obtained by Monte Carlo calculations, which is certainly one possible source of error for any final damage estimates.

4. Deterioration effects of the detector performance and modeling

4.1. Experimental considerations

Test diodes are normally quite small ($< 1 \text{ cm}^2$), and have – contrary to the microstrip detectors used for tracking – a simple pad structure with one or a few guard rings surrounding the central electrode. Such structures are fabricated by various manufacturers in addition to prototype strip or pixel detectors or exclusively for test purposes. It has been shown, that for all measurements the proper use of the guard ring is essential, since only this way the depleted volume is precisely defined by the geometrical area of the pad electrode and lateral effects contributing to both the capacitance and the leakage current are practically excluded [16]. Although that is of course true for the original p^+n diodes with the pad and guard ring structure on the p^+ side and the electric field starting below that electrode, it is even more relevant for diodes after high irradiation levels, when the original n-type bulk material had effectively changed to p-type. In this case the junction side is switched to the rear electrode and the electric field starts to grow from there. Normally the rear electrode is completely metallized all over the total area and hence the volume of the depleted zone is much larger. Only when the electric field extends completely to the front electrode, we get a homogeneous field distribution across both the pad and guard ring extension and hence the guard ring will act again as such, defining the active volume exclusively. The resulting effects can be clearly seen in Fig. 2, where for the same situations measurements are compared with floating and connected guard ring [16].

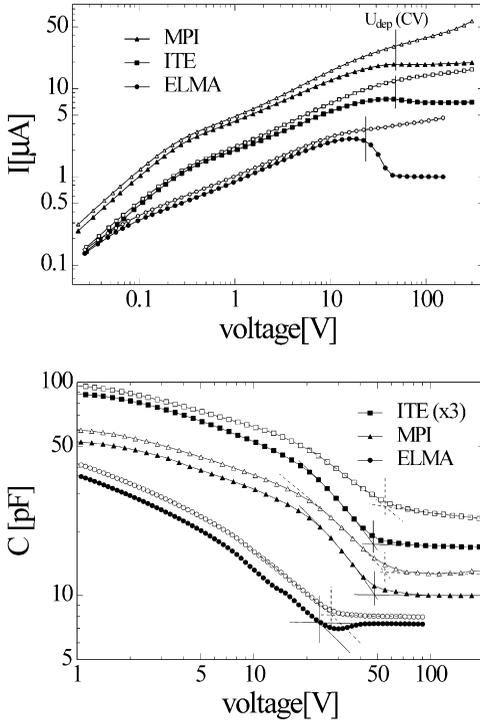


Fig. 2. Comparison between I/V resp. C/V measurements with and without connected guard ring. The filled symbols denote the measurements with connected guard ring.

4.2. Current-related damage rate α

The reverse current of an irradiation damaged silicon detector is proportional to the active volume V and the equivalent 1 MeV fluence $\Phi_{\text{eq}} = \kappa \Phi_{\text{tot}}$. For a fully depleted detector it can be written as

$$I = \alpha \Phi_{\text{eq}} V \quad (4)$$

and it has been shown that the current-related damage rate α is independent of the specific silicon material used for production of the detectors, the process technology used for manufacturing and the particular particles used for irradiation. This fact is nicely demonstrated in Fig. 3, taken from Ref. [16]. As was emphasized in Section 2, in addition to the generation of the immediate damage produced by the irradiation, long-term annealing effects play a major role for the application of the final detectors in any future experiment. It has therefore

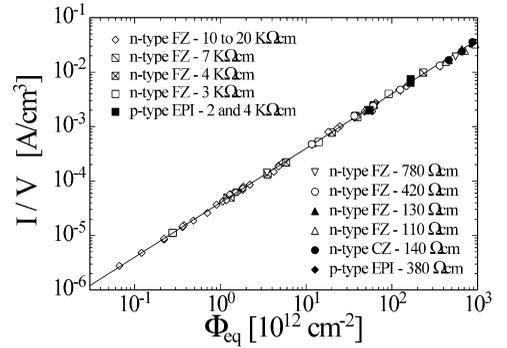


Fig. 3. Fluence dependence of leakage current for silicon detectors produced by various process technologies from different silicon materials [16].

been vital to show that one can indeed compress the annealing expected during 10 yr of LHC operation to only days or weeks in a proper laboratory experiment. This can be achieved by increasing the temperature moderately in such a way that no additional effects are to be considered. The following empirical function, although not yet fully understood, had been shown to describe the complex dependence of the damage rate α on the annealing temperature T and annealing time t in an adequate way [16]:

$$\alpha(T, t) = \alpha_1 \exp(-t/\tau_1(T)) + [\alpha_0 - \beta \ln(\theta(T)t/t_0)]. \quad (5)$$

Here all α -values are normalized to a measuring temperature of 20°C by using an appropriate scaling factor for the reverse current I taking mainly the effective energy gap of 1.12 eV into account [18]

$$I(T_{\text{meas}}) \propto T_{\text{meas}}^2 \exp\left(\frac{-E_g}{2k_B T_{\text{meas}}}\right). \quad (6)$$

It has meanwhile been shown, that Eq. (5) describes all data within the range of annealing temperatures between 20°C and 100°C with an activation energy of 1.11 eV responsible for τ_1 [19]. For lower temperatures and annealing times up to several weeks additional exponential terms have to be taken into account [20,21]. Here the dependence on the annealing temperature can be described by an activation energy of 1.09 eV [21]. A normalized annealing at 60°C, for which the acceleration factor

with respect to 20°C is ≈ 200 , has proven to be quite useful. A moderately short annealing of e.g. 80 min at that temperature (equivalent to about 10 days at room temperature) has hence been proposed to be used as an intrinsic way of monitoring the irradiation fluence employing a universal α -value [16]

$$\alpha_{80/60} = 4.0 \times 10^{-17} \text{ A cm}^{-1} \pm 5\%. \quad (7)$$

4.3. Effective doping concentration and depletion voltage

For a given detector thickness d with pad area A and an effective doping concentration N_{eff} the voltage necessary to extend the electric field through a depth w is given by

$$C(V) = \epsilon\epsilon_0 \frac{A}{w(V)} \quad (8)$$

with

$$w(V) = \sqrt{\frac{2\epsilon\epsilon_0 V}{q_0 N_{\text{eff}}}} \quad (9)$$

and hence the depletion voltage (necessary for reaching $w = d$) is related to N_{eff} by

$$|N_{\text{eff}}| = \frac{2\epsilon\epsilon_0 V_{\text{depl}}}{q_0 d^2}. \quad (10)$$

A measurement of V_{depl} , easily obtainable from a capacitance voltage characteristic is therefore the normal method to investigate the radiation induced change in N_{eff} . Again, a proper determination of the electric field geometry (by use of a guard ring, see above) is very important. Also the used frequency and temperature at which the measurement is performed plays a role especially in the case of damaged silicon diodes in which deep levels are affecting the C/V behavior. This has recently been shown again in Ref. [22] and a thorough discussion can be found in Ref. [23]. Finally it should be mentioned that the detector could be either looked at as a parallel or serial circuit of a capacitance and a resistor. All these different methods could result in slightly different values for the depletion voltage resp. the effective doping concentration. In addition what really counts is the full sensitivity of the de-

tor thickness to ionizing radiation and hence instead of relying on C/V measurements one could extract the full depletion voltage as that for which a saturation of charge collection can be obtained [21,24,25]. Most measurements have however been done with the normal C/V method and at a frequency of 10 kHz. The whole matter had been the subject of several papers with somewhat slightly different approaches for parameterization of the results. The following description uses mainly the results of the Hamburg group. Their findings are not only based on very systematic investigations but have also included a number of other results in a universal comparison and modeling [21,26,27].

The observed change in the effective doping concentration N_{eff} of the bulk material as measured immediately after irradiation, is displayed in Fig. 4 [20].

For the starting n-type material an exponential decrease of the effective impurity concentration N_{eff} had been observed at lower fluences as obvious also in Fig. 4 and this behavior had been interpreted as an apparent “donor removal”. Alternatively the low fluence reduction of N_{eff} had been accounted for by a compensation model, which however would not explain the observed exponential decrease [28]. There is so far no clear understanding of the underlying physics. On top of the “donor removal part” acceptor-like states are being generated leading finally to inversion of the conduction type and a further fluence proportional increase of N_{eff} .

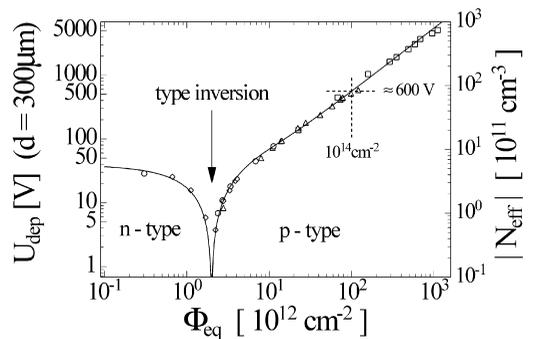


Fig. 4. Change in the bulk material as measured immediately after irradiation [20].

In contrast to what was seen for the damage rate α the time dependence of the irradiation effect in N_{eff} is not only subject to a beneficial annealing but also to an adverse effect, called anti-annealing or reverse annealing. An example of the whole complex behavior is given in Fig. 5 [19]. Here ΔN_{eff} is the irradiation induced change in the effective doping concentration with respect to its initial value before irradiation:

$$\Delta N_{\text{eff}}(\Phi_{\text{eq}}, t(T)) = N_{\text{eff}0} - N_{\text{eff}}(\Phi_{\text{eq}}, t(T)). \quad (11)$$

As function of time and fluence ΔN_{eff} can be described as

$$\Delta N_{\text{eff}} = N_{\text{C}}(\Phi_{\text{eq}}) + N_{\text{a}}(\Phi_{\text{eq}}, t(T)) + N_{\text{Y}}(\Phi_{\text{eq}}, t(T)). \quad (12)$$

Here it has been explicitly denoted that the time dependence is in itself subject to the annealing temperature T . N_{C} is not depending on annealing and therefore called the stable damage part. According to the Hamburg model, used here throughout, it consists of an “incomplete donor removal”, depending exponentially on the fluence with a final value of $N_{\text{C}0}$ and in addition a fluence proportional introduction of stable acceptors:

$$N_{\text{C}} = N_{\text{C}0}(1 - \exp(-c \cdot \Phi_{\text{eq}})) + g_{\text{C}} \cdot \Phi_{\text{eq}}. \quad (13)$$

It should be emphasized that the parameters $N_{\text{C}0}$ and g_{C} are of considerable importance for the operability of detectors in intense radiation fields

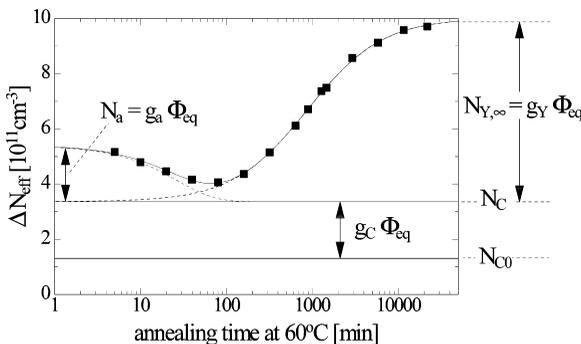


Fig. 5. An example for the annealing behavior of the radiation-induced change in the effective doping concentration ΔN_{eff} . The sample was irradiated with a neutron fluence of $1.4 \times 10^{13} \text{ cm}^{-2}$ and annealed at a temperature of 60°C [19].

like for the LHC applications (compare Section 6). N_{a} is a short-term annealing component, which may be represented by a sum of several exponentials [20]. Its value has only to be accounted for during each beam period every year, for which the time constants are appreciably prolonged due to cooling, but will completely anneal out during the yearly maintenance period of about 2 weeks with a time constant of only 2 days at room temperature. Hence for the 10 yr operational scenario in LHC experiments this part is not significant. The last component N_{Y} in Eq. (11) is called the “reverse annealing”, as its behavior is opposite to the beneficial annealing: N_{Y} starts from zero at $t = 0$ and saturates to a fluence proportional value $N_{\text{Y}\infty}$ for very large times with a time constant of about 350 days at room temperature. The responsible activation energy for this temperature dependence is 1.31 eV [21]. The exact function of the annealing time had been a matter of some discussion in the past. While in the beginning of the investigations it was modeled as being the result of a second order effect, both experimental evidence and a more physical interpretation of the inherent kinetics have now led to a first order description with a likely normal exponential behavior:

$$N_{\text{Y}} = N_{\text{Y}\infty} \cdot (1 - \exp(-t/\tau)) \approx N_{\text{Y}\infty} \cdot \left(1 - \frac{1}{1 + t/\tau}\right) \quad (14)$$

with $N_{\text{Y}\infty} = g_{\text{Y}} \cdot \phi_{\text{eq}}$

and a temperature-dependent time constant $\tau(T)$. Other evaluations have analyzed the reverse annealing as a genuine first order process (see e.g. Refs. [29,30]). However, it has to be emphasized that the results of the Hamburg group had taken an apparent relaxation effect into account which was observed at room temperature storage following each high-temperature annealing step leading to a slightly different time dependence [31]. Although with these measurements the first order model does not allow the same excellent fit as by the second order approach, it can easily be seen that for moderate annealing times the description is almost identical and would lead to the same time constant $\tau(T)$. Also as to the operational projections for LHC applications an annealing of 1000 min at 60°C is already more than equivalent to what is

foreseen as the overall storage at room temperature during 10 yr of operation which is appreciably less than 1 yr. Therefore, even without solving the open question about the first or second order time dependence, the description given above is certainly a reliable parameterization for any long-term projections.

It is of course very important to demonstrate that the relevant parameters for the model description are indeed independent of the particle source (if properly normalized by NIEL scaling) and independent of the material used, because only then one could rely on using such a model for the projection to any given operational scenario like e.g. for LHC. As to the independence on the particles used for irradiation, a dedicated experiment was carried out, in which detectors from the same material and process technology have been used for irradiations performed with the Be (d,n) neutron source at the PTB Braunschweig, the 24 GeV/c proton beam at CERN-PS and a 236 MeV π^+ beam at PSI Villigen [21]. Using the proper hardness factors for these sources the generation rate g_c , responsible for the formation of stable acceptors differs only by $\pm 5\%$ and g_Y , accounting for the reverse annealing by not more than $\pm 10\%$. The independence of damage effects on the starting material used so far, is demonstrated in Fig. 6 for the example of the reverse

annealing amplitude $N_{Y\infty}$. Also in this case the results given are those from a dedicated experiment in which all detectors have been subjected to damage using the same irradiation source and annealing measurements were done under identical conditions such that again systematic errors could be largely avoided [19]. In addition to the detector grade n- and p-type silicon with a resistivity between 2 and 20 k Ω cm measurements are included using a low-resistivity n-type silicon grown by the normal float-zone (FZ) technique and also from Czochralski (Cz) grown crystals to be commented in Section 6. Apart from the unconventional Cz silicon all results can obviously be represented by one straight line giving a universal value of $g_Y \approx 5 \cdot 10^{-2} \text{ cm}^{-1}$

An up to date analysis of a number of different experiments, performed with detectors produced from normal FZ high-resistivity silicon (1–20 k Ω cm n-type) has revealed the parameters listed in Table 1, relevant for the “Hamburg model” [19,27].

4.4. Charge collection efficiency

In addition to the formation of generation/recombination centers, being responsible for the increase of the leakage current and the creation of charged defects leading to the dramatic effects for the depletion voltage, damage-induced defects can also act as trapping centers. The trapping would then affect the charge collection efficiency giving rise to an irradiation-induced decrease of the signal height. However, for silicon detectors this is not nearly as problematic as the other two deterioration causes, given that experimentally an operational voltage could be established which is high enough to guarantee a sufficiently high electric field throughout the detector thickness. The minimum voltage needed is of course the depletion voltage but it proves to be advisable to maintain some overbias, since otherwise the electric field would drop to zero at the end of the depleted zone. Measurements have been performed with minimum ionizing particles, e.g. using an appropriate β -source showing that even after an equivalent fluence of 10^{14} cm^{-2} the mip signal height had dropped only to 90% [32]. Thus even for higher

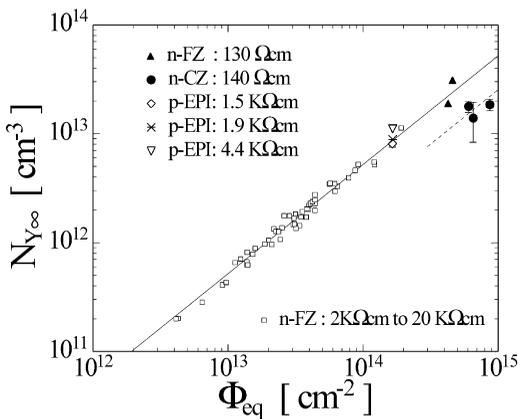


Fig. 6. Reverse annealing amplitude $N_{Y\infty}$ versus fluence for various silicon materials. The solid line represents the average introduction rate $g_Y = (5.16 \pm 0.09) \cdot 10^{-2} \text{ cm}^{-1}$ for the high resistivity n-type FZ silicon and the dashed one the introduction rate of $(2.53 \pm 0.27) \times 10^{-2} \text{ cm}^{-1}$ for the Cz silicon [19].

Table 1

Best fit parameters for the described damage model extracted from measurements with detectors fabricated from high-resistivity n-type FZ silicon

Beneficial annealing	$g_a = (1.92 \pm 0.05) \times 10^{-2} \text{ cm}^{-1}$ $E_a = (1.09 \pm 0.09) \text{ eV} \Rightarrow \tau_a(60^\circ\text{C}) = 20 \text{ min}, \tau_a(20^\circ\text{C}) = 2 \text{ days}$
Reverse annealing	$g_Y = (5.16 \pm 0.09) \times 10^{-2} \text{ cm}^{-1}$ $E_a = (1.31 \pm 0.04) \text{ eV} \Rightarrow \tau_a(60^\circ\text{C}) = 1000 \text{ min}, \tau_a(20^\circ\text{C}) = 350 \text{ days}$
Stable acceptors	$g_C = (1.49 \pm 0.03) \times 10^{-2} \text{ cm}^{-1}$
Partial donor removal	$N_{\text{co}} = (0.60 - 0.90) N_{\text{eff0}}; c = (1 - 3) \times 10^{-13} \text{ cm}^2$

fluences the charge collection deficiency does not really establish any critical problem. Charge collection has also been the subject of more elaborate experiments in which the electron and hole contributions were measured separately using infrared ultrashort laser pulses incident on either side of the detector and thus allowing to observe the respective charge carrier drift by measuring the time resolved current with a Gigahertz digital oscilloscope [21,24]. These investigations have led to a parameterized description of the electron and hole trapping as function of the equivalent fluence and with this knowledge the overall behavior of the charge collection in the case of mip's had been modeled [21]. It should be emphasized that these results describe the above-mentioned experimental data obtained with minimum ionizing electrons extremely well. Finally, it should be noted that measurements have also been undertaken employing a mono-energetic alpha source. Also in this case the particle range is quite small compared to the detector thickness, in order to allow the above-described separate measurements for both charge carriers. In fact, this method was the first to show undoubtedly the type inversion of the original material, switching from n- to p-type at a given fluence (compare Section 4.3 and Fig. 4) [33,34].

5. Defect analysis and correlation with detector performance

As has been emphasized already, the knowledge on correlation between macroscopic deterioration effects and the microscopic defects being responsible is indispensable if one wants to improve the radiation hardness of silicon detectors. Defect anal-

ysis in radiation-damaged silicon had therefore gained much attention. Although being a well-established field for low-resistivity silicon with applications in electronics, for detector grade high-resistivity material it is a comparatively new area. Using TSC (Thermally Stimulated Current) and DLTS (Deep-Level Transient Spectroscopy) techniques pioneering work had been reported in Refs. [35–37].

The DLTS method is by far the most sensitive tool for obtaining knowledge on specific defect concentrations, their energetic level and capture cross sections. A beautiful example of what has been possible with this method is shown in Fig. 7 [19]. It shows the spectra for electron and hole traps and their development within 170 days of room temperature annealing. Although one cannot deduce a direct correlation of one of these defect levels with the macroscopic effects, the results give an important insight into the involved defect kinetics and thus establish the input for defect modeling and a better understanding. A few points should underline this. Oxygen, being resident in the silicon lattice on interstitial sites and carbon (on substitutional ones) are the most important impurities in detector grade silicon and hence it is not surprising that as a result of the original damage and subsequent annealing many defects are formed with these components. An interstitial silicon Si_i (PKA, see Section 3) could e.g. knock an original substitutional carbon atom C_s out of its lattice site, giving rise to a high concentration of carbon interstitials C_i (acting both as electron and hole trap). As these carbon atoms then migrate through the lattice they could either combine with left over carbon substitutionals, forming C_iC_s or with oxygen interstitials giving rise to C_iO_i complexes (as to carbon-related defects in

high-resistivity silicon see also Ref. [38]). Hence, as is evident in Fig. 7, the C_i concentration decreases with annealing while the formation of the respective complex defects is increasing. One can also see that the vacancies created in the original damage combine either with each other or with the oxygen interstitials. The peak at about 200 K is much broader than would be expected for a single level. As being deduced from a comparison with results from gamma damage the broadening contribution is a unique effect of hadron damage and may be attributed to the formation of clusters in this case [39].

A completely different possibility for defect spectroscopy has however been developed using the TCT-method (Transient Current Technique) for measuring time-resolved current transients, already mentioned in Section 4.4 [40,41]. An example of such an experiment, by which one could not only gain knowledge on the charge collection process but also on defect levels close to the middle of the band gap, is shown in Fig. 8 [41]. Ultrashort light pulses from an infrared laser create electrons and holes in a very shallow region underneath the electrode of incidence. Hence due to their charge only one type of carriers will drift through the detector. If done at a low temperature these carriers will not only be trapped by the respective defects but may even stay frozen there. Therefore, depending on the preceding illumination the concentration of trap-

ped carriers increases and a considerable change of the electric field within the detector is observed, visible as a variation in the slope of the current pulses (see Fig. 8). Another possibility is to measure the detrapping as function of the temperature and thus use the method as a tool for deep-level spectroscopy. It should be added here, that the freezing of charge carriers in deep trapping centers at very low temperatures may result in an improved charge collection, because already filled trapping centers will not be able to act as such any more [42].

In contrast to what had hoped for originally the immense effort being invested in defect spectroscopy had only revealed very little direct correlation with macroscopic observations. Nearly all of the defects observed by DLTS measurements are not responsible for the change in the doping concentration and their introduction rates are by one order of magnitude higher than those expected from macroscopic analysis. Therefore, the relevant defects are most likely hidden as very small contributions in the DLTS spectra and those of other methods. There are however a few exceptions for which such a correlation was established.

In Ref. [43] the authors report about TSC measurements revealing a bistable hole trap at $E_v + 0.32$ eV responsible for macroscopic changes in the depletion voltage. The defect is produced in heavily irradiated diodes during heating procedures above 80°C and accompanied by an increase of negative space charge. At room temperature the defect signal decays with a time constant of about 500 min concurrent with a decrease of depletion voltage in type inverted detectors (decrease of negative space charge). Furthermore it was shown in Ref. [44] that the same defect level or a level very close to the one mentioned above is also correlated with the increase of negative space charge during the reverse annealing. Another indication for the microscopic basis of the reverse annealing comes from photoluminescence measurements [45]. Here it has been shown that the so-called W-line is emerging in linear correlation with the preceding reverse annealing. This line belongs to an isoelectronic interstitial-related defect center. Therefore the defect can not directly be responsible for the increase of negative space charge. However, it is strongly suggested that the reverse annealing is

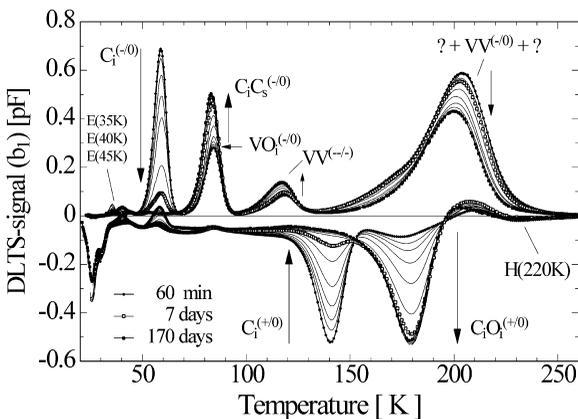


Fig. 7. DLTS-spectrum transformation at room temperature within 170 days after neutron irradiation with a neutron fluence of $2 \times 10^{11} \text{ cm}^{-2}$ [19].

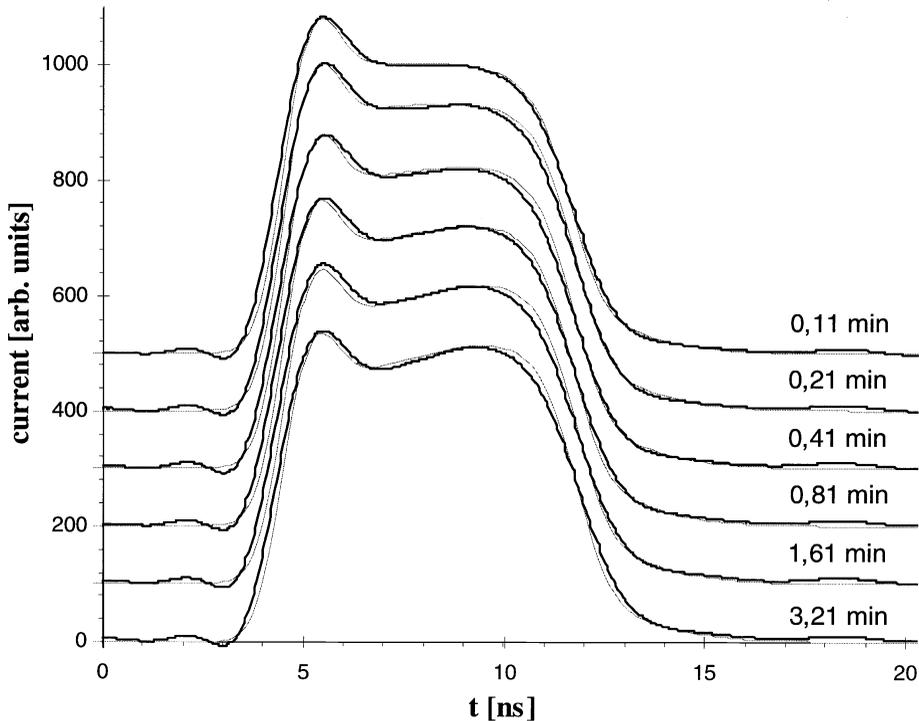


Fig. 8. Current pulse shapes measured at a temperature of 160 K. The thick lines present the waveforms observed after the indicated illumination period from the rear contact (hole injection), the thin lines are the corresponding simulations ($f = 100$ Hz; $V = 150$ V; $\phi_{\text{eq}} = 3.4 \times 10^{13}$ n/cm²) [41].

coming along with a rearrangement of interstitial-related defects.

It should be added at this point that reverse annealing had only been observed following hadron irradiations. While in gamma-induced damage predominantly only point defects are being formed, hadronic interactions lead to cluster formation at the end of each high energetic PKA range. An obvious attempt for explaining the reverse annealing behavior is therefore the assumption that these clusters will dissolve on a large time scale and the evading interstitials and/or vacancies would then migrate through the silicon lattice and form point defects responsible for the observed increase in the space charge [29]. As compelling this model looks, so far no direct proof of such a correlation had been possible. This is mainly due to the fact that clusters formed by hadronic damage have so far evaded direct observation [46].

Furthermore, there is an obvious correlation between the annealing behavior of the reverse current and the time dependence of the broad DLTS-peak at 200 K (see Fig. 7) which is partly attributed to the divacancy level $VV^{(-/0)}$ [19,47,48]. But it has to be stated that the absolute current values cannot be reproduced using the standard Shockley–Read–Hall statistics with the measured level parameters. One explanation for this discrepancy is proposed in Ref. [49]. Since most of the divacancies are produced inside disordered regions and are therefore located close together an interdefect charge transfer could be possible. This would lead to an enhanced current generation sufficient to explain the macroscopic measurements. The observed annealing of the bulk current could then again be explained by the time-dependent disintegration of clusters as mentioned above.

A modeling of macroscopic effects on the basis of a microscopic description had been undertaken and shown acceptable agreement for gamma-induced damage but only little success in the case of hadrons [45].

6. Possible improvements for radiation tolerance

6.1. Process technology

From Section 4.3 it is quite obvious that the detectors should be operable at high voltages up to several hundred volts. The use of a quite elaborate guard ring structure had been shown to result in quite appreciable improvements. As an example, the MPI semiconductor laboratory in Munich has shown that a tailored geometry ensuring a moderate potential drop between adjacent guard rings has resulted in detectors, tolerating an operational voltage of up to 500 V even after an equivalent fluence of 10^{14} cm^{-2} [50]. Although technologically this does not seem to be a big problem, the manufacturers would not easily guarantee such a performance since after the unavoidable-type inversion the detectors may indeed behave quite differently than what could be tested upon delivery of the nonirradiated devices. In many cases the maximum possible voltage proves to be even higher after irradiation than before. A possible explanation had e.g. been given in Ref. [51].

6.2. Operational conditions

As has been stated already above, an appreciable cooling of the detectors is essential to keep the reverse current at a tolerable level. For the ATLAS-SCT detectors an operating temperature of -7°C is foreseen. This reduces the diode current by more than a factor 10, enough to guarantee the design signal-to-noise ratio of $S/N = 12:1$ for mip's even at high fluences. On the other hand, cooling at this level has an additional beneficial effect, as the reverse annealing remains largely frozen. In fact, the room temperature time constant of ≈ 350 days is enormously enlarged by a factor of ≈ 190 , resulting only in 3×10^{-3} of the reverse amplitude to appear within the yearly beam time.

With the exception of an as short as possible maintenance period (e.g. ≈ 2 weeks for ATLAS-SCT), that low temperature has to be maintained continuously for the full 10 years operation. Keeping the detectors at room temperature for the beam off period of 5 months would otherwise lead to $\approx 35\%$ of the reverse annealing amplitude every year. As this is a cumulative process, the large effect on the depletion voltage would not be tolerable. On the basis of the model description given in Section 4, realistic projections of the detector performance can be obtained for any foreseen scenario. Specific calculations of this sort have been used for the design of forthcoming experiments [1,2,5,6,21].

6.3. Choice of starting material

Several possibilities have been proposed for increasing the radiation tolerance of silicon detectors by using a specific type of the starting material. These include the proper choice of the resistivity (doping concentration), a modification of the carbon and oxygen concentration as being the main responsible impurities in damage kinetics (see Section 5) or even enriching silicon with Sn, which has been shown to act as vacancy trap [52]. Investigations mainly undertaken by the ROSE collaboration have meanwhile covered an appreciable part of the available phase space regarding the C- and O-concentration and work has been started on Sn-enriched silicon as well. From all the experimental evidence obtained so far as well as by theoretical considerations the oxygen content seems to be a key factor for an improved radiation hardness. Measurements have been performed using standard FZ silicon with a normal concentration of below 10^{16} cm^{-3} , epi-material grown on O-rich Czochralski substrates reaching a value of 10^{17} cm^{-3} by in-diffusion from the substrate and O-enriched FZ material with $2 \times 10^{17} \text{ cm}^{-3}$ [15,53]. By far the maximum O-enrichment of $\approx 10^{18} \text{ cm}^{-3}$ is present in Cz-silicon itself. However, this material is normally only available with a low resistivity of up to $100 \Omega \text{ cm}$, not suitable for real detectors. Nevertheless a demonstration experiment was performed for comparing the radiation hardness of this material with that of FZ silicon with the same resistivity [19,27]. The most

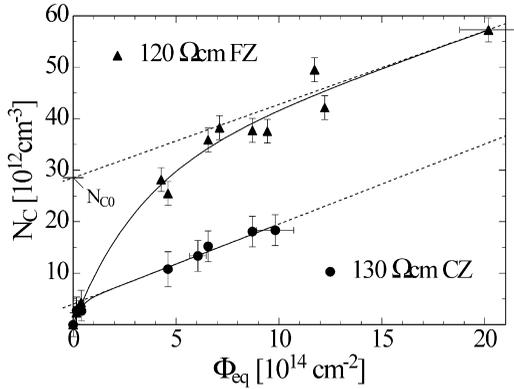


Fig. 9. Damage parameter N_c plotted against the fluence. Each point represents one sample that was subjected to an isothermal heat treatment at 60°C. The corresponding annealing data were fitted according to Eq. (12) to extract the parameter N_c (see also Fig. 5). The intercept of the dashed lines with the ordinate represented the damage parameter N_{co} (see text).

important result is shown in Fig. 9. From this it is quite evident that the parameter N_{co} , depicting the maximum donor removal at high fluences, is much smaller for Cz silicon than for FZ-grown crystals. In fact, for the Cz diode only about 10% of the initial donors were “removed” whereas for the FZ-ones 80–90% had been obtained. It has to be recounted here that for the real ATLAS-SCT scenario the N_Y -contribution (reverse annealing) is practically frozen to a very low value and N_a is effectively annealing out during the short maintenance period every year. The dominating effect for the change of the effective doping concentration and hence for the resulting depletion voltage is therefore the “stable” contribution N_c in Eq. (12), which for higher fluences, i.e. for $\exp(-c\phi_{eq}) \Rightarrow 0$, is given only by $N_{co} + g_c \cdot \phi_{eq}$. Reducing N_{eff0} to a practically negligible value of $0.1 \cdot N_{co}$ would therefore be very efficient, if starting with a comparatively low-resistivity material. According to the simplified description one could then fully benefit from the initial n-type doping which would reduce the necessary depletion voltage also after high fluences by the preirradiation value. It should however be kept in mind that this substantial improvement can only be obtained if material of about 1 k Ω cm and very high oxygen concentration becomes available. Presently work is under way to

ensure this goal on moderate cost levels which would not be prohibitive for LHC applications [54]. A possible explanation for the oxygen effect was proposed using the vacancy sharing model, in which it is assumed that migrating vacancies can either form vacancy-oxygen complexes (*A-centers*) or lead to donor removal by formation of phosphorus-vacancy defects (*E-centers*). Calculations of this model on the basis of capture radii deduced from DLTS studies on defect kinetics (Section 5) have led to a surprisingly good agreement with the observed donor removal [19]. It should be added that for the O-rich Cz-silicon also the generation rate for reverse annealing is reduced by at least a factor of 2 (see Fig. 6) but the other parameters, mainly the stable acceptor generation rate g_c and the current-related damage rate α are unaffected. Finally Fig. 10 shows the result of an operational projection for ATLAS-SCT in which a moderate value of $N_{co} = 0.30 N_{eff0}$ was assumed [4]. In the given example the low donor removal effect, only to be expected for O-rich material, leads to a stable donor concentration saturating at a level of 70% of the initial one at high fluences, i.e. for the latest LHC years. Looking at the extreme cases of Fig. 10 the starting depletion voltages are 34 V (for $N_{eff} = 4 \times 10^{12} \text{ cm}^{-3}$) resp. 274 V (for $N_{eff} = 0.5 \times 10^{12} \text{ cm}^{-3}$). The final difference in depletion voltage after 10 yr of operation is 170 V corresponding to 70% of the starting difference which is a sizeable effect for any practical considerations. In contrast,

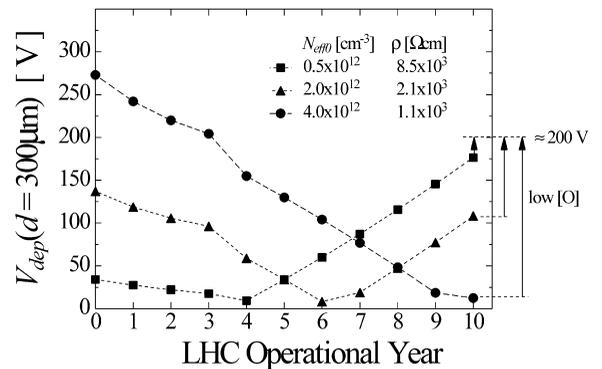


Fig. 10. Calculated depletion voltage as function of LHC operational year for the first layer in the ATLAS SCT barrel ($R = 30 \text{ cm}$, $Z = 0 \text{ cm}$, $\phi_{eq} = 1.75 \times 10^{13} \text{ cm}^{-2}$ per year at full luminosity) [4].

by using standard silicon with a low O-concentration results in $N_{\text{co}} = N_{\text{eff0}}$ and all curves would finally lead to the same depletion voltage as indicated in Fig. 10 without any advantage provided by low-resistivity material. Granted that high O-rich material can readily be used for manufacturing the detectors, one may then even want to start with as low-resistivity material as possible with respect to the needed depletion voltage (see curve for 1.1 k Ω cm). As during the prolonged irradiation the depletion voltage would drop continuously while the reverse current would increase proportional to the accumulated fluence, the power dissipation could be kept approximately constant during the whole 10 yr period, which would indeed be a very nice feature.

Possible advantages of O-enriched silicon – although on a much lower scale than available in the Czochralski material – had also been discussed in Refs. [55].

7. Conclusions

The immense work, which had been invested in the study of damage-induced deterioration effects in silicon particle detectors mainly by the ROSE collaboration had led to a reliable model description which allows the prediction of device performance during prolonged operation in high-intensity radiation environments. The values for the employed parameters are a function of the hadronic fluence only, provided NIEL scaling is applied, and independent on the material properties for normal float zone silicon. Microscopic studies on defect characterization and kinetics have paved the way for a better understanding of the real physics processes behind. Operation at low temperatures and using a process technology allowing high-voltage operation are conventional methods for guaranteeing the design criteria during an LHC period of 10 yr but may not prove to be sufficient for critical applications. Recent results with oxygen-rich silicon have been a first success towards improved radiation hardness, especially if combined with a moderately low resistivity of the starting material. This path as well as other possibilities will be vigorously pursued.

References

- [1] ATLAS Technical Proposal, CERN/LHCC/94-43, 1994; ATLAS Inner Detector Technical Design Report, vol. 1&2, CERN/LHCC/97-16&17, 1997.
- [2] CMS Technical Proposal, CERN/LHCC/94-38, 1994.
- [3] ROSE-collaboration CERN/RD-48, Proposal for further work on radiation hardening of silicon detectors, CERN/LHCC/96-23, 1996, see also <http://www.brunel.ac.uk/research/rose>.
- [4] ROSE Status Report, CERN/LHCC/97-39, 1997.
- [5] HERA-B Design Report, DESY-PRC 95/01, 1995.
- [6] LHC-B Technical Proposal, CERN/LHCC/98-4, 1998.
- [7] A. Vasilescu, ROSE Internal Note ROSE/TN/97-2, 1997.
- [8] A. Vasilescu, G. Lindström, Fluence normalization for radiation damage studies of silicon detectors, to be published.
- [9] P.J. Griffin et al., SAND92-0094, SANDIA Natl. Lab., Nov. 93, neutron cross sections taken from ENDF/B-VI, ORNL, unpublished, data available from G. Lindström (gunnar@sesam.desy.de).
- [10] A.Yu. Konobeyev et al., J. Nucl. Mater. 186 (1992) 117.
- [11] G.P. Summers et al., IEEE. Nucl. Sci. NS- 40 (1993) 1372.
- [12] M. Huhtinen, P.A. Aarnio, NIMA 335 (1993) 580; and HU-SEFT R 1993-02, 1993.
- [13] A. Van Ginneken, Fermilab National Accelerator Laboratory FN-522, 1989.
- [14] M. Lazo et al., Technical Report SAND87-0098, vol. 1, 1987.
- [15] A. Vasilescu, Fluence normalization based on the NIEL scaling hypothesis, 3rd ROSE Workshop on Radiation Hardening of Silicon Detectors, DESY Hamburg 12–14 February 1998, DESY-PROCEEDINGS-1998-02.
- [16] M. Moll et al., Nucl. Instr. and Meth. A 426 (1999) 87.
- [17] ASTM E722-85 and ASTM E722-93 (revision 1993).
- [18] S.M. Sze, Physics of Semiconductor Devices, 2nd Ed., Wiley, New York, 1981.
- [19] M. Moll, Ph.D. Thesis, University Hamburg, in preparation.
- [20] R. Wunstorf, Ph.D. Thesis, University Hamburg, see also DESY FHIK-92-01, 1992.
- [21] H. Feick, Ph.D. Thesis, University Hamburg, see also DESY F35D-97-08, 1997.
- [22] L. Beattie et al., Rose Technical Note ROSE/TN/97-4, 1997.
- [23] A.G. Milnes, Deep Impurities in Semiconductors, Wiley-Interscience, New York.
- [24] V. Eremin, Nucl. Instr. and Meth. A 372 (1996) 388.
- [25] G. Casse et al., Nucl. Instr. and Meth. A 426 (1999) 140.
- [26] A. Chilingarov et al., Nucl. Instr. and Meth. A 360 (1995) 432.
- [27] E. Fretwurst et al., A comprehensive model for the effective doping concentration of irradiated silicon detectors, paper presented at The 3rd Int. (Hiroshima) Symp. on the Development and Application of Semiconductor Tracking Detectors at Melbourne, December 9–12, 1997, Nucl. Instr. and Meth. A and literature cited there, to be published.

- [28] Z. Li et al., Nucl. Instr. and Meth. A 409 (1998) 180.
- [29] Z. Li, IEEE Trans. Nucl. Sci. NS-42 (4) (1995) 224.
- [30] H.-J. Ziöck et al., Nucl. Instr. and Meth. A 342 (1994) 96.
- [31] M. Moll et al., Nucl. Phys. B 44 (proc. Suppl.) (1995) 468.
- [32] F. Lemmeilleur et al., Nucl. Instr. and Meth. A 360 (1995) 438.
- [33] R. Wunstorf et al., Nucl. Phys. B (proc. Suppl.) A 23 (1991) 324.
- [34] H.W. Kraner et al., Nucl. Instr. and Meth. A 326 (1993) 350.
- [35] U. Biggeri et al., IEEE Trans. Nucl. Sci. NS-41 (1994) 964.
- [36] C.J. Li et al., Nucl. Instr. and Meth. A 364 (1995) 108.
- [37] E. Fretwurst et al., Nucl. Instr. and Meth. A 377 (1996) 258.
- [38] B. Schmidt et al., J. Appl. Phys. 76 (7) (1994) 4072.
- [39] M. Moll et al., Nucl. Instr. and Meth. A 388 (1997) 335.
- [40] Z. Li et al., Nucl. Instr. and Meth. A 388 (1997) 297.
- [41] E. Fretwurst et al., Nucl. Instr. and Meth. A 388 (1997) 356.
- [42] V. Palmieri et al., Nucl. Instr. and Meth. A 413 (1998) 475.
- [43] H. Feick, M. Moll, Solid State Phenomena, vols. 57, 58, p. 233, Scitec Publications, Switzerland, 1997.
- [44] M. Moll et al., Nucl. Instr. and Meth. A 409 (1998) 194.
- [45] B.C. MacEvoy, Ph.D. Thesis, Imperial College London, RAL-TH-97-003, 1997.
- [46] Z. Li et al., IEEE Trans. Nucl. Sci. NS-43 (3) (1996) 1590.
- [47] S.J. Watts, ROSE Technical Note ROSE/TN/96-2, (1996).
- [48] V. Eremin et al., IEEE Trans. Nucl. Sci. NS-42 (4) (1995) 387.
- [49] K. Gill et al., CERN-PPE/97-11, 1997.
- [50] L. Andricek et al., Nucl. Instr. and Meth. A 409 (1998) 184.
- [51] V. Eremin et al., Nucl. Instr. and Meth. A 388 (1997) 350.
- [52] B.G. Svensson, J.L. Lindström, J. Appl. Phys. 72 (1992) 5616.
- [53] 3rd ROSE Workshop on Radiation Hardening of Silicon Detectors, DESY Hamburg 12–14 February 1998, DESY-PROCEEDINGS-1998-02.
- [54] F. Limeilleur, G. Lindström, S. Watts, ROSE NOTE 98-1.
- [55] Z. Li et al., IEEE Trans. Nucl. Sci. NS-42 (4) (1995) 219.