Analysis of n^+p silicon junctions with varying substrate doping concentrations made under ultraclean processing technology

Herzl Aharoni[a] and Tadahiro Ohmi
Department of Electronic Engineering, Faculty of Engineering, Tohoku University, Aza-Aoba, Aramaki, Aoba-ku, Sendai 980-77, Japan

Mauricio Massazumi Oka and Akira Nakada
Department of Electronic Engineering, Faculty of Engineering, Tohoku University, Aza-Aoba, Aramaki, Aoba-ku, Sendai 980-77, Japan and Laboratory for Electronic Intelligent Systems, Research Institute of Electrical Communications, Tohoku University, Aza-Aoba, Aramaki, Aoba-ku, Sendai 980-77, Japan

Yukio Tamai
Department of Electronic Engineering, Faculty of Engineering, Tohoku University, Aza-Aoba, Aramaki, Aoba-ku, Sendai 980-77, Japan

(Received 25 July 1996; accepted for publication 16 October 1996)

Using highly controlled ultraclean processing technology, marked improvements in n^+p Si junction quality are achieved presenting a theoretical significance. Boron-doped substrates with various boron doping concentrations N_s were As^+ implanted, forming the n^+ junction sides. The diffusion (I_d) and generation (I_{gen}) currents, as well as the ideality and the generation factors, are significantly reduced, and bulk generation lifetimes are prolonged. Using Shockley–Read–Hall theory it is found that a deviation of the trapping centers energy (E_t) from the midband-gap energy (E_i) is responsible for the improvements. The experimental results show that $|E_t-E_i|$ is a function of N_s, and that the I_{gen}/I_d ratio is significantly low. Accordingly, it is proposed that I_{gen}/I_d ratio should be regarded, under certain conditions, as a figure of merit for junction quality. It is concluded that the $|E_t-E_i|$ deviation is related to the ultraclean processing technology used, due to the formation of new energy levels far from E_i and the suppression of introduction of new energy levels near E_i. Surface generation currents were found experimentally to be significant, and thus not negligible. Surface effects in general demonstrated similar trends to the bulk generation effects. © *1997 American Institute of Physics.* [S0021-8979(97)06602-4]

I. INTRODUCTION

Low value of reverse-bias current (I_R) is an important indicator of a p-n junction quality. Its value has the consequences of the performance of a wide array of semiconductor devices, such as solar cells, transistors, thyristors, etc. Among the four constituents of I_R, namely, diffusion (I_d), tunneling (I_t), bulk and surface generation ($I_{gen,b}$ and $I_{gen,s}$, respectively) current components, the first two result from predesign considerations, for specific devices and applications, mainly through doping concentration. The space-charge region bulk and surface generation current components, on the other hand, which can be controlled to some extent, do not result from predesign considerations. They are parasitic in nature, and as such, their values are in principle, not unique even for a given group of junctions having nominally identical process type, crystal quality, doping concentrations, dimensions, and structure. This is because $I_{gen,b}$ and $I_{gen,s}$ depend on device *process quality*, in a very critical fashion. Essentially, the bulk and surface generation components present loss mechanisms that limit the performance of the above-mentioned devices, and they load the driving source without serving any useful purpose, while at the same time, enhancing circuit operation instabilities and increasing noise level. In nowadays integrated circuit (IC) technology, I_t is negligible at the doping concentrations commonly used. Accordingly, the attention in this work is focused on the I_d, $I_{gen,b}$, and $I_{gen,s}$, where the sum of the last two components represent the overall generation current, I_{gen}. The ratio I_{gen}/I_d is experimentally shown in this work to present a measure of the junction quality.

In the present IC technology, I_{gen} exceeds I_d by far. Thus the I_{gen}/I_d ratio is high and can range from few tens to few hundreds. In such junctions (especially diffused junctions) the reverse-bias current I_R is thought to be practically equal to I_{gen}, and calculations of τ_{gen}, the generation lifetime of carriers within the bulk space charge region is calculated directly from I_R, neglecting the surface current component.[1] While I_d in IC's is mainly determined by doping concentrations, the I_{gen} value can be suppressed in the same junction, in principle, by modifying fabrication quality, yielding greatly reduced I_{gen}/I_d ratios. This is achieved practically in this work by the suppression of foreign undesired contaminants unintentionally introduced into the p-n junction structure during its processing. This suppression approach, which is in a class by itself, is made possible by using precisely controlled fabrication conditions in an ultraclean technology environment, in a high investment facility that was predesigned and built for future submicron ultra-large-scale integration (ULSI) research. In this facility a measurable and drastic reduction in the various contaminants in the laboratory atmosphere, gas storage tanks, gas delivery lines, and fabrication equipment, together with stringent Super Clean

[a]On leave from the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel.

 0021-8979/97/81(3)/1270/19/$10.00

Room (SCR) laboratory practices are carefully maintained. Although all types of unwanted contaminants are maintained at low levels, this publication is concerned mainly with those undesired impurities, which, if introduced during processing into the junction, would form energy levels deep within the band gap, namely heavy metal atoms. Reduction of these atoms content is directly responsible for the very low I_{gen}/I_d ratios achieved in our experiments, together with meaningful improvements in other junction operational parameters, which will be listed in the sequel. The large magnitude of these improvements will be shown to be of theoretical significance. The processing steps and the technological environment that were utilized in fabricating the junctions in this work, will be briefly described in Sec. III.

A series of identical ion (As^+) implanted n^+p junctions were made, with different substrate concentrations N_s (1.6×10^{14}–2.3×10^{18} cm^{-3}), which served as the p side (boron doped) of the junction, in order to investigate this parameter's influence on device performance. The data indicate a strong influence of N_s on the junction parameters. Following the measurements of the current voltage (I-V) relations, fitting of the data to the routinely used double exponential expression for the I-V characteristics of the forward bias recombination and diffusion current components, derived for the case where $E_t=E_i$ and $E_{st}=E_{si}$ (where E_{st} and E_{si} represent the trapping and intrinsic energy levels at the surface respectively), could not be achieved. The expected behavior of the above expression results, according to the Shockley–Read-Hall (SRH) theory,[2,3,4] a power of $1/m$ (where $m=2$ is the generation factor for $E_t=E_i$ and $E_{st}=E_{si}$) in the exponential term of the forward recombination current component. In addition, for the case of $E_t=E_i$ and $E_{st}=E_{si}$, high I_{gen}/I_d is usually obtained. Surface currents are usually considered to be small with respect to overall reverse-bias currents I_R.[1] In such a case, indeed, $I_R \approx I_{\text{gen},b}$ and τ_{gen} can be calculated directly from I_R, and its value can be shown[4,5] to be equal to τ_n the lifetime of the minority carriers outside the space-charge region (i.e., $\tau_{\text{gen}}=\tau_n$ in the case of one sided n^+p junctions). Finally, neglecting surface effects, the temperature dependence of the reverse-bias current should, in the $E_t=E_i$ case, exhibit two slopes, one proportional to $-E_g$ in the high temperature range, and the second, proportional to $-E_g/2$ in the low temperature range, that represent the generation current.[4,5]

Our results exhibited *significant deviations* from all the above expected behavior, for the whole range of substrate concentrations. They can be summarized as follows:

(1) Low I_{gen}/I_d ratios were obtained which were close to unity, the highest ratio being 7.2.
(2) Surface current to junction (bulk) current ratios can reach significant values. Surface effects could not be neglected in the analysis.
(3) Ideality (n) and generation (m) factors were close to unity, all over the current range. The slope in the low current regime, where the generation mechanism dominates in the case of $E_t=E_i$, never reached the value of 1/2.
(4) Generation lifetime of minority carriers within the bulk space charge region is much longer than minority carrier lifetime outside the space charge region. τ_{gen}/τ_n ratios ranged from few tens to few thousands.
(5) Temperature dependence of I_R in the low temperature range (where generation mechanism dominates in the case of $E_t=E_i$) exhibited slopes significantly different from $-E_g/2$.

All the above deviations were found experimentally to be a *function of the substrate doping concentration*.

The objective of this work is to analyze in detail the performance of these super clean room (SCR) fabricated junctions, showing that the theoretical foundation for the above listed deviations can be obtained by the SRH theory using the basic equations where $E_t\neq E_i$ and $E_{st}\neq E_{si}$ conditions exist, with good agreement with the experimental data. This is done by resolving the reverse-bias current into its components, and examining the factors that govern each of them, and their contribution to the overall performance, as a function of N_s. Finally, an explanation is proposed which relates the experimental results to the ultraclean processing technology.

In Sec. II, a brief summary of the theory for $E_t\neq E_i$ case for the bulk, and $E_{st}\neq E_{si}$ case for the surface will be presented. In Sec. III the experimental part is outlined, describing the special features of the ultraclean technology used as well as junction fabrication details and parameters. In Sec. IV the data of the various junction operational parameters as a function of substrate doping concentration is given. In Sec. V those results are discussed in conjunction with the above outlined theory. Finally, conclusions (Sec. VI), followed by a summary (Sec. VII), are given.

II. THEORY

The purpose of the following presentation is to quantitatively determine, directly from dc I-V characteristics of a p-n junction, the influence of the bulk and surface recombination/generation center concentrations N_t and N_{ts}, respectively, and their trapping energy levels E_t and E_{ts}, respectively, on the space-charge region bulk and surface recombination/generation currents, their temperature dependence, their generation factors m, their minority carrier lifetimes, and surface recombination velocities, all as a function of N_s. This is done in accordance with the SRH theory.[2,3,5] Since the large body of data that was generated in the experiments required a detailed treatment, the main expressions that will be used in the analysis are summarized for convenience.

In general, the bulk recombination current $I_{\text{rec},b}$ of a p-n junction operating under forward bias condition and the net recombination rate per unit volume per unit time U_b within the bulk space charge region are related by:

$$I_{\text{rec},b}=qA\int_0^W U_b(x)\,dx, \tag{1}$$

where q is the electron charge, A is the junction cross section, W the space charge region width and:[4–6]

$$U_b(x) = \frac{\sigma_p \sigma_n v_{\rm th} N_t [np - n_i^2]}{\sigma_n \left[n + n_i \exp\left(\frac{E_t - E_i}{kT} \right) \right] + \sigma_p \left[p + n_i \exp\left(\frac{E_i - E_t}{kT} \right) \right]}. \tag{2}$$

n and p are the free carrier concentrations for electrons and holes, respectively, n_i is the intrinsic carrier concentration, $v_{\rm th}$ is the free carrier thermal velocity, T is the absolute temperature, and k is the Boltzmann constant. The simplifying assumptions used here are that within the bulk all the trapping centers (N_t) possess the same energy level E_t, the capture cross sections for electrons (σ_n) and holes (σ_p) are both equal to σ, the recombination centers are evenly distributed all over the crystal volume, and that the pn product is constant all over the space-charge region[4] for a given forward bias condition. Band-gap narrowing is neglected as well. These approximations influence only the magnitude of the calculated results but not their trends. It can be shown that under the above approximations, the maximum value of U_b is given by:

$$U_{b,\max} = \frac{1}{\tau_0} \frac{n_i (e^{qV/kT} - 1)}{2 \left[e^{qV/2kT} + \cosh\left(\frac{E_t - E_i}{kT} \right) \right]}, \tag{3}$$

where from Eq. (2):

$$\tau_0 = \frac{1}{\sigma v_{\rm th} N_t}. \tag{4}$$

In this work the influence of N_t and E_t is of major interest. In our n^+p junctions, the data obtained demonstrate that in the p side of the junctions, $p_p \gg n_p$ and $p_p \gg n_i \exp[|E_t - E_i|/kT]$ situation exists. In such a case, Eq. (2) can be approximated as $U_b \simeq \sigma_n v_{\rm th} N_t (n_p - n_{p0})$.[4,7] Accordingly, by definition, τ_0 represents in these junctions the minority carrier lifetime on the p side. Thus $\tau_0 = \tau_n = 1/\sigma v_{\rm th} N_t$. Finally, under the approximation that maximum generation rate conditions (where $n = p$) exist all over the space-charge region,[4] Eq. (1) becomes:

$$I_{{\rm rec},b} = qAW_F U_{b,\max} = \frac{qAW_F n_i}{2\tau_n} \frac{e^{qV/kT} - 1}{e^{qV/2kT} + \cosh\left(\frac{E_t - E_i}{kT} \right)}. \tag{5}$$

Equation (5), which presents the I-V relation of the net bulk recombination component, was derived for forward bias conditions, where $V = V_F > 0$, $W = W_F$, and $np > n_i^2$, within the bulk space-charge region. In reverse bias situation, where $V = V_R < 0$, $W = W_R$ and $pn \ll n_i^2$, within the bulk space charge region, the recombination is negligibly small and generation is dominant, yielding from Eq. (5) the reverse bulk generation current:

$$I_{{\rm gen},b} = -\frac{qAW_R n_i}{2\tau_n \cosh\left(\frac{E_t - E_i}{kT} \right)}. \tag{6}$$

Using $I_{{\rm gen},b} = Q_{\rm gen}/\tau_{\rm gen}$, where $Q_{\rm gen} = qAW_R n_i/2$, it follows from Eq. (6) that the generation lifetime of minority carriers within the bulk space charge region is given by:

$$\tau_{\rm gen} = \tau_n \cosh\left(\frac{E_t - E_i}{kT} \right) \tag{7a}$$

and

$$I_{{\rm gen},b} = -\frac{qn_i W_R A}{2\tau_{\rm gen}}. \tag{7b}$$

From Eq. (7b) independent determination of $I_{{\rm gen},b}$ leads to the calculation of $\tau_{\rm gen}$, and by Eq. (7a) to the calculation of $|E_t - E_i|$. This is important since $I_{{\rm gen},b}$ decreases as $|E_t - E_i|$ increases [Eq. (6)], a fact that is of major interest in this work. Ideally, $\tau_{\rm gen}$ is expected to be independent of doping concentration,[8,9] since the space-charge region becomes depleted under reverse bias. Our data show, however, such a dependence, although a weak one. This is because E_t and N_t were experimentally found to be in this work, doping dependent for reasons which will be given in the discussion. Finally, according to Eq. (7a) large $\tau_{\rm gen}/\tau_n$ ratios indicate large deviations of E_t from E_i. This is confirmed by our data.

By analogy, a similar set of equations having identical form as Eqs. (1)–(6), describing the surface behavior can be written.[4,10,11] In these equations the same notations as in Eqs. (1)–(6) are used, possessing in addition the subscript s for surface. Accordingly, E_{st}, E_{si}, E_{sg}, σ_{sn}, σ_{sp}, N_{st}, n_s, p_s, and U_s have the same meanings for the surface, as the parallel notations have for the bulk. As a result the surface current component due to the contribution of the surface generation centers at the silicon–oxide (S–O) interface, occupied by the junction periphery space-charge region area A_s, can be shown[4,10,11] to be

$$I_{{\rm gen},s} = q|U_s|A_s = \tfrac{1}{2} q n_i s A_s, \tag{8a}$$

where

$$U_s = -\tfrac{1}{2} n_i s \tag{8b}$$

and assuming $\sigma_{sn} = \sigma_{sp} = \sigma_s$:

$$s = \frac{s_0}{\cosh\left(\frac{E_{st} - E_{si}}{kT} \right)} \tag{8c}$$

and

$$s_0 = \sigma_s v_{\rm th} N_{st}. \tag{8d}$$

In Eq. (8) U_s is the surface generation rate per unit area per unit time, N_{st} is the trapping centers concentration per unit area, and s is the surface recombination velocity (in cm/s) of minority carriers, all at the S–O interface. According to Eq. (8c), s becomes smaller as E_{st} deviates from E_{si}. From our data it can be inferred that similarly to the situation in the bulk $|E_{st} - E_{si}|$ and N_{st} are substrate doping (N_s) dependent. This will be discussed in Sec. V. s_0 is the surface recombination velocity at the S–O interface for the $E_{st} = E_{si}$ case, which results in [from Eq. (8c)] the highest value for s. According to Eq. (8a), independent measurement of $I_{{\rm gen},s}$ can lead to the determination of s as a function of N_s. Using Eq. (8c), this dependence can provide a measure to the de-

viation of E_{st} from E_{si} as a function of N_s. This is important because as E_{st} deviates from E_{si}, $I_{\text{gen},s}$ decreases, according to Eq. (8).

The total reverse generation current I_{gen} is given by the sum of the bulk and surface generation components of Eqs. (7b) and (8a), respectively. The forward surface recombination current $I_{\text{rec},s}$ due to recombination centers at the S–O interface is dependent exponentially on the forward voltage in a similar fashion as the bulk recombination current [Eq. (5)]. Accordingly, the total reverse generation and forward recombination currents I_{gen} and I_{rec}, respectively, are given by:[4]

$$|I_{\text{gen}}|=\frac{1}{2}\,qn_i\left(\frac{W_R A}{\tau_{\text{gen}}}+sA_{sR}\right), \tag{9a}$$

$$I_{\text{rec}}=\frac{1}{2}\,qn_i\left(\frac{W_F A}{\tau_{\text{gen}}}+sA_{sF}\right)(e^{qV_F/mkT}-1). \tag{9b}$$

Changes in W in the bulk for forward and reverse bias are accompanied by parallel changes in A_s at the S–O interface. Accordingly $A_s=A_{sR}$ for reverse bias and $A_s=A_{sF}$ for forward bias. In Eq. (9b), m is the generation factor. Comparing Eq. (9b) to Eq. (5), and to expressions similar to Eq. (5) that can be written for the surface current using Eq. (8) it can be shown that in general m is dependent on E_t, E_{st}, N_t, N_{st}, σ_n, σ_{sn}, σ_p, and σ_{sp}. The generation factor equals 2[4] for the $E_t=E_{st}=E_{si}=E_i$, $\sigma_n=\sigma_p=\sigma$, and $\sigma_{sn}=\sigma_{sp}=\sigma_s$ case as can be seen from Eq. (5), and from a similar expression that can be written for the surface current. Otherwise $m<2$, as is the case in this work. The expression in the square brackets in Eq. (9b), multiplying the exponential term is denoted I_{gf}, the generation current under forward bias conditions.

Generally, the N_t dependence on N_s can be described, in principle, by a power series. In this work only the first two terms are taken, as a practical approximation.[12–15] Accordingly:

$$N_t=N_{ti}+BN_s, \tag{10}$$

where B is a constant, and N_s is the substrate (p type) dopant concentration, which in our series of junctions constitute the low doped side of the junctions. N_t in Eq. (10) is the effective value of trapping centers concentration. That is, in addition to technological factors that induce certain amount of defects as described in the discussion, N_t is also determined by the position of the Fermi level enabling more of these defects to participate in the recombination process as N_s increased.[16] This is further discussed in Sec. V. Substituting Eq. (10) into Eq. (4) yields a very familiar experimental expression for the dependence of minority carrier lifetime on N_s, which was used by several researchers,[12–14,17–19] for both τ_p and τ_n, namely:

$$\tau_n=\frac{\tau_{ni}}{1+\left(\frac{B}{N_{ti}}\right)N_s}, \tag{11}$$

where $\tau_{ni}=1/\sigma v_{\text{th}}N_{ti}$, and N_{ti} are the lifetime and recombination centers concentration, respectively, at very low doping concentrations. For high doping concentrations τ_n decreases proportionally to N_s^{-1}, due to the increase in the active defect concentration [Eq. (10)].

The temperature dependence of I_{gen} is determined mainly by $n_i(T)$ and the hyperbolic expression of Eq. (6), and the parallel relation for the surface component of Eq. (8):

$$I_{\text{gen}}=-q\sqrt{N_v N_c}\left[\frac{AW_R}{2\tau_n}\frac{e^{-(E_g/2kT)}}{\cosh\left(\frac{E_t-E_i}{kT}\right)}+\frac{A_s s_0}{2}\frac{e^{-(E_{sg}/2kT)}}{\cosh\left(\frac{E_{st}-E_{si}}{kT}\right)}\right], \tag{12}$$

where N_V and N_C are the effective density of states in the valence and conduction bands, respectively. To get the general trend, then, for the case where $E_t=E_{st}$, $E_i=E_{si}$, $E_g=E_{sg}$, and $|E_t-E_i|\gg kT$, where T is the experimental temperature, yields the approximated result that the slope of I_{gen} in log scale versus $1/T$ is proportional to $-[E_g/2+|E_t-E_i|]$.[5] It can be seen that for $E_t=E_i$ the expression is reduced to $-E_g/2$.

The diffusion component of the junction current used in this work is:[4,5]

$$I_D=I_d(e^{qV/kT}-1), \tag{13a}$$

where

$$I_d=qAn_i^2\left(\sqrt{\frac{D_n}{\tau_n}}\frac{1}{N_A}+\sqrt{\frac{D_p}{\tau_p}}\frac{1}{N_D}\right). \tag{13b}$$

D_n is the electron diffusion constant and τ_n is the electron lifetime in the p side. D_p and τ_p are the same for holes in the n side. N_A and N_D are the p and n sides doping concentrations, respectively. The $1/T$ dependence of I_d yields a slope proportional to $-E_g$ at the high temperature through $n_i^2(T)$.

According to the above, the double exponential expression for the overall I-V characteristics used in this work is:[20]

$$I=\frac{V-IR_s}{R_p}+I_{gf}(e^{q(V-IR_s)/mkT}-1)+I_d(e^{q(V-IR_s)/kT}-1). \tag{14}$$

R_s and R_p in Eq. (14) are the equivalent series and parallel resistances of the junction, respectively, where R_p is connected directly across the ideal junction and R_s is the resistance between the junction and the terminals. The R_s/R_p ratio is assumed to be negligible when deriving Eq. (14) as verified by the experimental results. Finally, it is noted that Eq. (14) takes the standard form once $E_t=E_i$ and $E_{st}=E_{si}$ is substituted into the above equations, i.e., it results in $m=2$, $\tau_{\text{gen}}=\tau_n=\tau_0$, $I_{\text{gen},b}=qAW_Rn_i/2\tau_0$, $s=s_0=\sigma_s v_{th}N_{st}$, and the slope of the I_{gen} (log) vs $1/T$ becomes proportional to $-E_g/2$.

Under reverse bias, R_p further increases and Eq. (14) becomes:

$$I_R=-[I_d+I_{\text{gen}}]=-[I_d+(I_{\text{gen},b}+I_{\text{gen},s})]. \tag{15}$$

Accordingly, calculating the junction overall reverse current density J_R by simply dividing the terminal current I_R

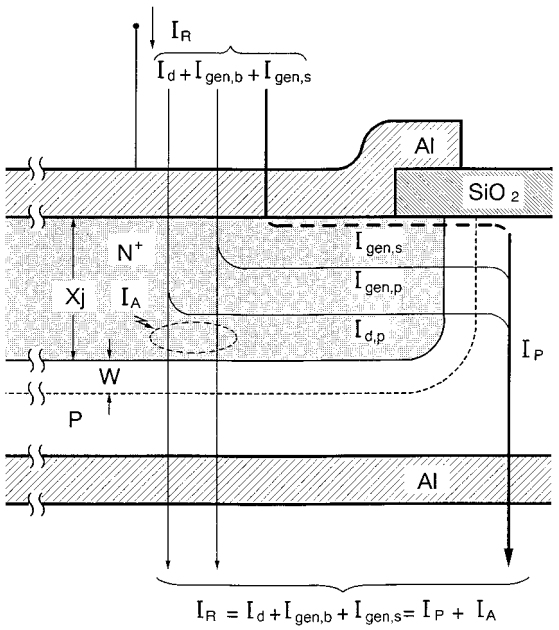

FIG. 1. A schematic description of various junction current paths.

by A, as sometimes done by various authors, is physically meaningless and can introduce a serious error to the analysis of this type. This is because I_R contain a peripheral component I_p (Fig. 1) which includes the surface current $I_{\text{gen},s}$ which flows via A_s and not through A. In addition I_p includes small fraction of the diffusion ($I_{d,p}$) and bulk generation-current ($I_{\text{gen},p}$) of the junction as shown in Fig. 1, which schematically shows the various current paths. In this figure all the currents are assumed to flow at one side of the periphery, and they are shown as lumped components, for simplicity, while in reality they flow all around the periphery, in a distributed fashion. Similarly, I_{gen} [Eq. (9a)] and I_{rec} [Eq. (9b)] the overall generation and recombination currents, respectively, cannot be divided by A in order to find their current densities and for the same reasons, i.e., both contain surface components, which do not flow through A. The current I_A that flows through the junction cross section A is composed of the sum of I_d and $I_{\text{gen},b}$. Accordingly the overall reverse current can be also expressed as the sum of the cross section (area) current and the peripheral current, i.e., $I_R = I_A + I_p$. The peripheral current increases with L the length of the junction perimeter, at the periphery of the junction at the surface:

$$I_R = J_A A + J_p L. \tag{16}$$

The peripheral current density J_p in Eq. (16) has a dimension of current per unit length (A/cm). I_p and specifically its surface component $I_{\text{gen},s}$ can compose a significant portion of the overall reverse current as experimentally shown in this work and therefore cannot always be neglected. This problem becomes more severe as the junction area is decreased, since the L/A ratio increases, resulting an increase of the I_p fraction in I_R at the expense of I_A [Eq. (16)]. This bears a consequence on future ULSI devices since the scaling down of transistors dimensions can reduce the portion of the active (area) current that participates, for example, in the amplification or switching action, in bipolar transistors, due to losses resulting from the increase in the surface generation current, which does not participate in the transistor action. Similar relations and remarks that were given for I_R apply for forward bias situation with the reversal of the current directions in Fig. 1. In this work, the contribution of I_p is investigated by producing on the same wafers a series of additional test junctions with varying L and keeping their area A_p the same. Dividing Eq. (16) by $A = A_p$ yields:

$$I_R/A_p = J_A + J_p(L/A_p). \tag{17}$$

Plotting I_R/A_p vs L/A_p, for the above junctions, the resulting slope is J_p and the intersection with the vertical axis is J_A. The values of J_p and J_A obtained for the test junctions with the area A_p are the same for the junctions under investigation with area A, since both series were made simultaneously on the same wafers. This will be dealt with in detail in Sec. IV.

III. EXPERIMENT

Ultraclean technology has been developed very extensively at Tohoku University and it is based on the conception that the process quality needed for future submicron ULSI manufacturing will require the complete elimination of contamination and process variations. The whole philosophy behind the ultraclean technology can be summarized in a few concepts: environmental clean, energy clean, magnetic clean, thermal clean, vibration clean, and process-variation clean.[21,22] In order to fulfill all these requirements, a number of technologies have been developed, some of the key ones being detailed below. Ultraclean gas system[23–27] has been achieved using tubing systems, gas cylinders, valves, and other components made of stainless steel with internal wall surface-electropolishing treatment that causes drastic reduction in the internal area and thus reduces outgassing as well as contamination adsorption and desorption (especially O_2 and H_2O moisture). Most important, it drastically reduces the generation of particles, mainly heavy metals. The tubing lines and employed components were also carefully designed to eliminate deadzones that become source of particles due to the corrosion caused by the remaining gas. As a result, for the first time impurity levels of the order of ppb or even sub-ppb, instead of the usually reported ppm level, were attained. Table I compares the impurity levels obtained in the presently described gas system in comparison to typical values, taking N_2 as an example.[27] Ultrapure water supply systems were designed not only to improve the water quality (reducing particles, bacteria, total organic content (TOC), dissolved oxygen, and SiO_2) but also to continuously guarantee the same water quality obtained in the system output, at the user point.[28] Vibration control has been achieved by the use of pneumatically controlled high-performance active vibration removal unit and by minimizing vibration sources, which are motors used in the water and air circulation systems, by reducing their size and improving their efficiency, and the building itself, predesigned specifically for this laboratory by the use of appropriate construction techniques. The vibration measurement on the floor shows amplitudes of less than 0.4 μm and less than 0.15 μm for frequencies larger

TABLE I. Comparison of the impurity levels obtained in the gas system of the super clean room (SCR) of the Tohoku University compared to typical values.

Impurity	H_2O (ppb)	CO_2 (ppb)	O_2 (ppb)	NO (ppb)	THC[a] (ppb)	Others (ppb)
SCR	1.8	<0.1	<0.1	<0.1	<0.1	<0.1
typical	<200	<10	1200	···	<50	···

[a]Total Hydro Carbon.

than 3 Hz and for frequencies between 3 and 10 Hz, respectively.[25,29] Elimination of static electricity and magnetic fields in the fabrication environment has been pursued.[30,31] Static electricity is kept below 6 V which is at least less than half of the average values usually reported, and much less than the maximum values that can reach locally to a typical value of 100 V or higher in usual clean rooms. Magnetic fields variations are kept below 1 mG eliminating their variation effects on processes that make use of electron or ion beams.[30,31] The presently described experiments were totally carried out in a facility built using all the above technologies, where the cleanliness defined as the number of particles with radii greater than 0.17 μm is less than $2\times10^{-6}/\text{ft}^3$, in comparison to values of around $1–10/\text{ft}^3$ usually used in today's technology.[32]

The ultrahigh vacuum medium-current ion implanter employed in this experiment has a background pressure of 10^{-10} Torr, and uses an ultraclean gas delivery system[33,34] for the ion source. Wafers are held by electrostatic chucks[35] from their back side. The vacuum chamber as well as the internal components of the system are made almost entirely of aluminum alloys and the inner surfaces are TiN coated to minimize outgassing. Plates made of Si are strategically positioned inside the implanter in order to minimize the amount of metal contaminants sputter-generated by the high energy ion beam allowing the concentration of metal contamination to be kept lower than 10^{10} atoms/cm^2 [36] as shown in Table II. It is important to notice that the detection limit of the total reflection x-ray fluorescence analysis (TRXRFA) equipment used in Table II is of the order of 10^{10} atoms/cm^2. In other words, there is no detectable amount of metallic contaminant even after ion implantation.

Si wafers of p type doped with B, (100) orientation, grown by Cz and FZ methods were employed. The series includes FZ wafers with substrate concentrations of 1.6×10^{14}, 2.5×10^{15}, and 2.8×10^{16} cm^{-3}, and Cz wafers with concentrations of 6.3×10^{14}, 2.5×10^{15}, 1.3×10^{17}, and 2.3×10^{18} cm^{-3}. These concentrations were determined by measuring the capacitance-voltage (C-V) characteristics of metal-oxide-semiconductor (MOS) capacitors fabricated using these wafers. After performing the RCA cleaning, 600-nm-thick field oxide was initially formed by pyrogenic oxidation at 1000 °C. This oxide was etched from the wafer backside by buffered HF (BHF) and 200-nm-thick borosilicate glass (BSG) was deposited by atmospheric pressure chemical vapor deposition (APCVD) (SiH_4, O_2, and B_2H_6) at 400 °C. This BSG film was used as a source for the B diffusion, carried out at 1000 °C, in order to facilitate the ohmic contact at the wafer backside. After removing the BSG film, the active area was patterned on the top surface, forming a square geometry with area of $A=1\times1$ mm^2. As^+ was implanted into the exposed bare silicon in the ultraclean ion implantation machine[33,34,36] at 25 keV with a dose of 2×10^{15} cm^{-2}. Post-implantation annealing was carried out at 1000 °C in N_2 ambient for 30 min, yielding x_j ranging from 0.245 to 0.362 μm depending on N_s. 100-nm-thick oxide was deposited on the top surface by APCVD at 400 °C and 250-nm-thick Al was evaporated over it. Aluminum guard rings situated at a distance of 50 μm from the junction periphery having a width of 40 μm were patterned around the diodes in order to electrically isolate the individual diodes by forming a accumulation region through MOS capacitance action. 1-μm-thick oxide was deposited at the top surface by APCVD at 400 °C and contact holes for the n^+ side of the junctions, and for the guard rings were patterned over it. An alloy of Al/Si was then deposited by sputtering to the frontside and Al was evaporated to the backside of the wafer. Contact pads for the n^+ side of the junctions and for the guard rings were patterned on the frontside. Finally, sintering in forming gas environment was carried out at 400 °C for 20 min. Additional set of junctions for the determination of peripheral current component I_p was made simultaneously on the same substrate, with identical processing conditions and sequence, including guard rings. Four different junctions were employed in each substrate, all having the same area ($A_p=0.0016$ cm^2) but different perimeter L. The geometries used were square (400×400 μm) for the lowest perimeter ($L=1600$ μm) and comb-shaped patterns for the larger perimeters. As a result four ratios of L/A_p were obtained for these junctions, namely, 100, 200, 300 and 400 cm^{-1}. They will be referred to as the ''peripheral junctions.'' Accordingly, in the following, a distinction should be made between

TABLE II. Metal concentration measured by TRXRFA (total reflexion x-ray fluorescence analysis) over the Si wafer before and after ion implantation (I/I).

Metal		Fe	Cu	Ni	Mn	Cr	Co	Zn	Ti
Concentration $\times10^{10}$ (atoms/cm^2)	Before I/I	3.5	0.0	1.0	0.0	1.9	0.0	0.0	0.0
	After I/I	5.1	1.7	1.0	0.0	2.1	0.0	3.2	0.0

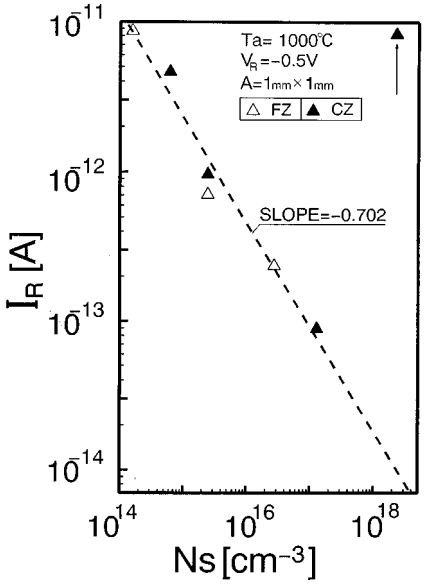

FIG. 2. Reverse bias current I_R at $V_R=-0.5$ V, as a function of substrate dopant concentration, N_s (1×1 mm^2 junctions).

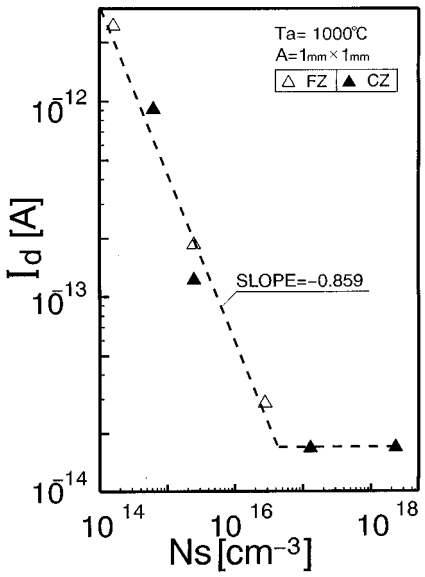

FIG. 3. Diffusion current I_d as function of substrate dopant concentration, N_s (1×1 mm^2 junctions).

the main series of junctions with area of 1×1 mm^2 which are the subject of this work, and the ''peripheral junctions'' used only to determine J_p and J_A, according to Eq. (17).

The data presented in the next section was measured in the computer controlled HP-4061 test system, under dark conditions, at 300 K, with -5 V applied to the guard ring with respect to the substrate. The 1×1 mm^2 junctions selected were the best in each wafer (i.e., the lowest I_R). Special care was taken in determining the time interval between measurements of I-V points at the very low current levels near the origin. This is needed due to the large charging time of the junction capacitance via the high impedance which exists at low currents (down to 10^{-15} A). Our experience showed that otherwise large errors in the junction parameters information extracted from this current region may result. This is important in this work since R_p as well as generation parameters such as I_{gf}, and m are mainly determined from the low current region. As mentioned, these errors were all eliminated. The data presented below was fitted by the least square method.

IV. RESULTS

Figure 2 shows the reverse bias current I_R of the 1×1 mm^2 junctions as a function of the substrate concentration $N_s(=N_A)$, measured at reverse bias of 0.5 V. Good fit to a straight line is obtained except for the 2.3×10^{18} cm^{-3} substrate, where the current significantly deviates to a high value (indicated by an arrow). This is due to a high tunneling current component as verified by the low breakdown voltage and its negative temperature coefficient. Accordingly, this point deviates also in other relevant graphs in this work and will be indicated in each case. By fitting the measured forward bias characteristics of the 1×1 mm^2 junctions to Eq. (14), I_d, I_{gf}, m, R_s, and R_p were obtained, all as function of N_s as shown in Figs. 3 to 7. Two slopes appear in Fig. 3 for I_d. The straight line fitting for the low doping concentration range (up to 10^{17} cm^{-3}) indicates that the junctions can be regarded as one sided, where the As$^+$ doped n^+ side is a dominant current injector (electrons) under forward bias. Its slope indicates that τ_n in the p side is decreasing with increasing N_s, as will be shown later. The slope at the high substrate concentration portion indicates that the p side injects, under forward bias, significant currents as well, in those junctions (holes to the n^+ side), and that, accordingly, should be regarded as two sided junctions. I_{gf} in Fig. 4 exhibits a straight line fitting except for the tunneling junction. This is meaningful because it is noted that this point deviates only in the I_{gf} case, and not in the I_d case. This is because both I_{gf} and I_{tf}, the tunneling current, are the result of mechanisms that take place within the high field depletion

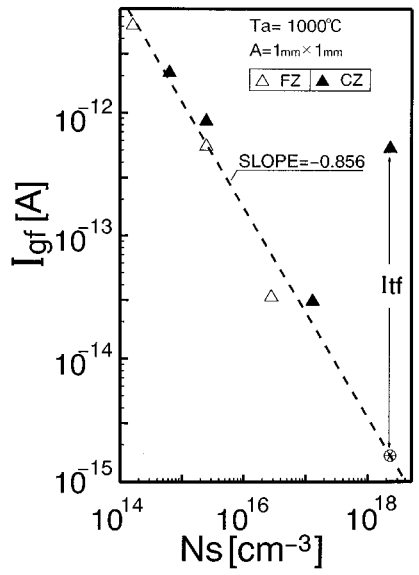

FIG. 4. Generation current component I_{gf} under forward bias, as a function of substrate dopant concentration (1×1 mm^2 junctions).

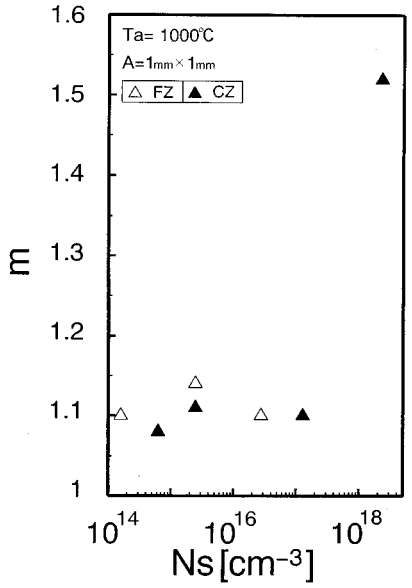

FIG. 5. Generation factor m, as a function of substrate dopant concentration N_s (1×1 mm^2 junctions).

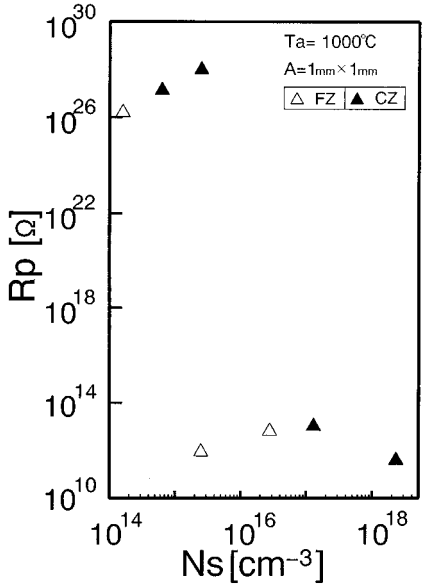

FIG. 7. Equivalent parallel resistance R_p as a function of substrate dopant concentration N_s (1×1 mm^2 junctions).

region, while I_d is determined by the injection in the low field quasi-neutral zones, out of the junction space-charge region. In our case it is possible to resolve the generation and the tunneling components for this junction. The generation current for this substrate is approximated to lie on the straight line of I_{gf}, as indicated by the crossed circle in Fig. 4. In the sequel, this point, and not the deviated point will represent I_{gf} for this concentration in the data manipulation. The vertical distance from this point to the measured point represents I_{tf}, the tunneling current. The tunneling behavior in this junction will be described in detail separately.[37] The m values in Fig. 5 are close to unity (typically 1.1). No dependence of m on N_s could be observed. Similar behavior was found for the ideality factor n (not shown). The low values obtained for I_{gf} and m, hint that $E_t \neq E_i$ and $E_{st} \neq E_{si}$ situation might exist for these junctions. This will be addressed later on. Figure 6 shows that R_s is decreasing monotonically with increasing N_s. The slope indicates that R_s includes not only bulk resistivity, but other effects as well, such as contact and possibly high injection effects. While R_s exhibits monotonic behavior, R_p (Fig. 7), in contrast, exhibits a significant sharp drop in its value for $N_s \geqslant 2.5\times10^{15}$ cm^{-3}. The high values of R_p for $N_s \leqslant 2.5\times10^{15}$ cm^{-3} do not represent any physical significance except indicating that R_p is very high in this region. This is due to the fact that the currents in this R_p region are much below the 10^{-15} A which is the measurement limit. The actual numbers for this region are higher. However, the sharp transition to the lower values, as well as the lower values of R_p (10^{11}–10^{13} Ω) represent real effects. Such values are usually neglected, but as for the junctions in this work they will be shown to be of importance for the determination of the correct values of I_{gf} and I_d. Possible reasons for this R_p behavior will be given in the discussion.

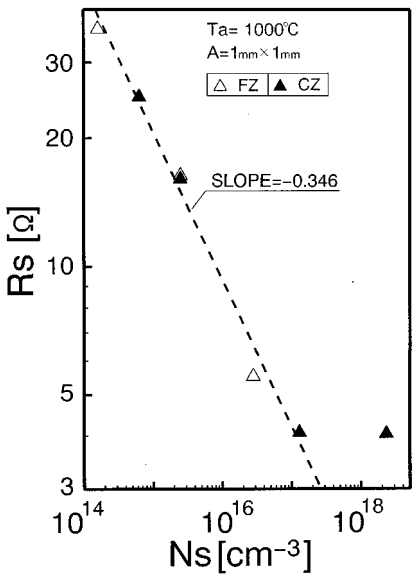

FIG. 6. Equivalent series resistance R_s as a function of substrate dopant concentration N_s (1×1 mm^2 junctions).

In order to determine the surface current, the reverse current I_{Rp} (at $V_R=-0.5$ V) of the four test structures termed ''peripheral junctions'' in the experimental section was measured at reverse bias of 0.5 V, and -5 V applied to their guard rings. Figure 8(a) shows I_{Rp} as a function of N_s for the largest perimeter junction as an example. Figure 8(b) shows the plot of I_{Rp}/A_p vs L/A_p according to Eq. (17), for the junctions with $N_s=6.3\times10^{14}$ cm^{-3} (Cz wafer) as an example. A good fitting to a straight line is obtained and the values of the peripheral current density (J_p) and the cross section (area) current density (J_A) are obtained for this junction from the slope and from the point where the dashed line cross the vertical axis, respectively. An increase in scattering of the data was noted as N_s increased. Figures 9 and 10 show the values of J_p (A/cm) and J_A (A/cm^2) obtained by this way

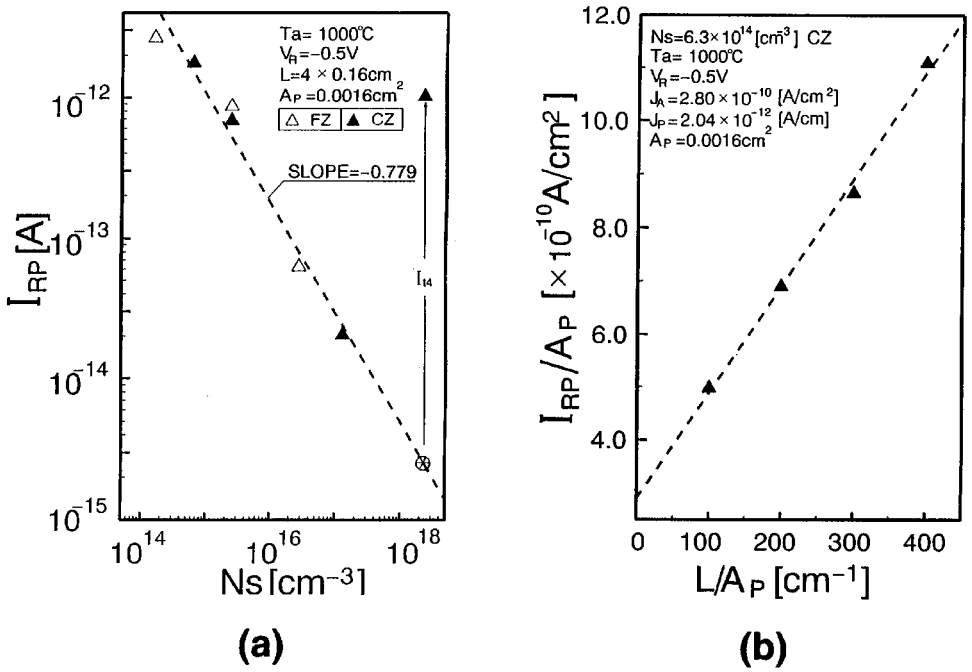

FIG. 8. (a) The reverse current I_{Rp} ($V_R=-0.5$ V) of the largest perimeter ($L=4\times0.16$ cm) peripheral test diode, as a function of substrate dopant concentration. (b) The I_{Rp}/A_p ratio as a function of L/A_p for the four peripheral test diodes for the $N_s=6.3\times10^{14}$ cm^{-3} Cz diodes.

for the whole range of N_s, of the peripheral junctions. These values of J_p and J_A, however, are the same for the 1×1 mm^2 junctions as well, since as mentioned in Sec. III both junctions are fabricated on the same substrates and measured under the same conditions. Accordingly, the results obtained above for J_p and J_A are used in the following to determine I_p and I_A in the 1×1 mm^2 junctions by multiplying J_p and J_A by $L=0.4$ cm, and $A=0.01$ cm^2, respectively. Accordingly all the following figures (Figs. 11–22) are referred to the 1×1 mm^2 junctions. In Figs. 9 and 10, it is noted that the data points for the 2.3×10^{18} cm^{-3} junction deviate from the fitted lines. This is mainly due to surface and bulk tunneling, respectively, as will be shown later on. The same point in Fig. 8(a) includes the combined tunneling current (I_{t4}). Figure 11 presents the ratio of the peripheral currents I_p to the area currents I_A as a function of N_s. With the exception of the tunneling junction, the data separation of the Cz and FZ substrates in Fig. 11 clearly reveals the fact that could not be explicitly seen in Figs. 9 and 10, that is, in junctions made of

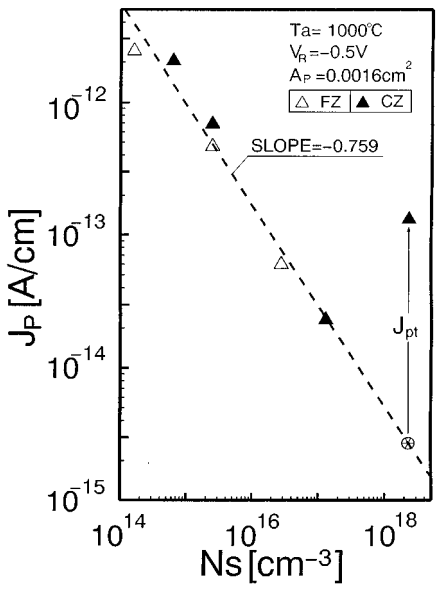

FIG. 9. The peripheral current density J_p obtained from the various slopes of the I_{Rp}/A_p vs L/A_p of the peripheral test junctions as a function of N_s.

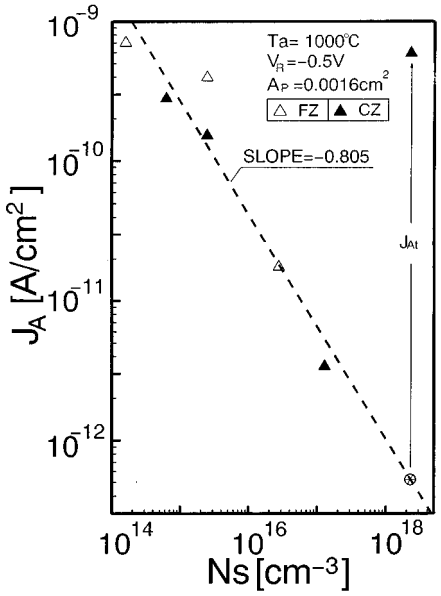

FIG. 10. The junction cross section (area) current density J_A, obtained from the intersection of the lines with the I_{Rp}/A_p axis (vs L/A_p) of the peripheral test junctions as a function of N_s.

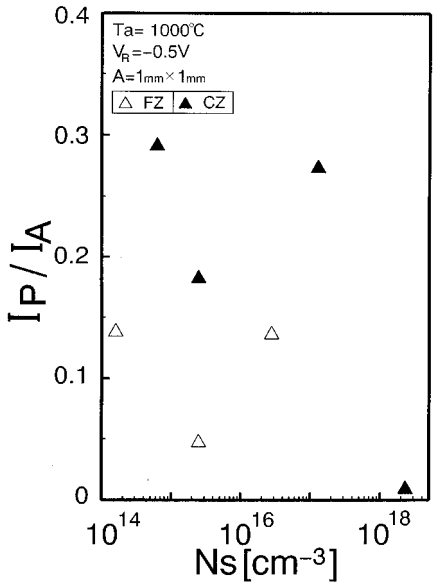

FIG. 11. The ratio of the peripheral current to the cross section current as a function of N_s (1×1 mm^2 junctions).

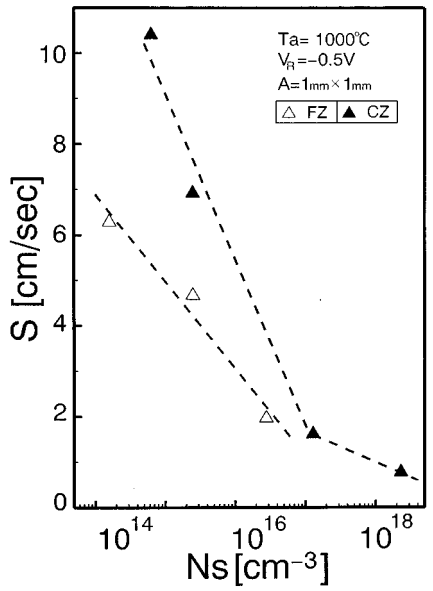

FIG. 13. Surface recombination velocity as a function of substrate doping concentration (1×1 mm^2 junctions).

Cz crystals the I_p/I_A ratio (20%–30%) is significantly higher than those made in FZ crystals (5%–15%), depending on N_s. This is despite the fact that both surfaces were treated simultaneously by the same process. This result can be understood, bearing in mind that the quality of FZ crystals are superior to those of Cz crystals due to the lower oxygen content. As discussed in Sec. V, oxygen gives rise to generation/recombination centers. This is true for both bulk and surface. Such high values of I_p cannot be neglected if τ_{gen} and $|E_t-E_i|$ as well as $J_{\text{gen},b}$ needed to be determined, even for junctions as large as 1×1 mm^2. Paradoxically, the above high ratios arise due to the higher quality of device processing which improves the bulk properties more than the surface properties which, in the junctions under discussion, are largely determined by the oxide–silicon interface properties. As mentioned, the increase in I_p/I_A ratio becomes a severe problem in ultra-large-scale integrated (ULSI) devices where L/A increases due to the scaling down of device areas. The peripheral currents are recently discussed in the literature[38–40] where the authors attribute the large I_p found in their junctions to field emission effects.

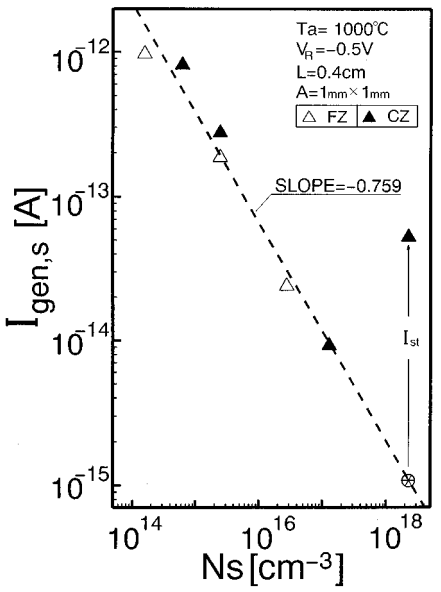

FIG. 12. The surface generation current as a function of N_s (1×1 mm^2 junctions).

I_p is composed of (Fig. 1) $I_{\text{gen},s}$, $I_{d,p}$, and $I_{\text{gen},p}$. The relative contributions of each of these components to I_R can be estimated by a simplifying assumption that $J_{d,p}$ and $J_{\text{gen},p}$ are the same as in the rest of the area A, then, $I_R \cong (J_d+J_{\text{gen},b})A+(J_d+J_{\text{gen},b})Lx_j+I_{\text{gen},s}$. The last two terms represent I_p and the first, I_A. Since the second term in our junctions (where $A \gg Lx_j$ situation exists) is much smaller than the first one, and since I_p/I_A is high (Fig. 11) it implies that in this case practically $I_p \cong I_{\text{gen},s}$. This will be quantitatively verified later on. Figure 12 represents $I_{\text{gen},s}$ ($=J_p\times0.4$ cm) as a function of N_s. Caution should be exercised when, unlike in this case, very small area junctions are analyzed where A becomes closer to the junction side wall area Lx_j. In such a case the above approximation may not be valid. Two points arise from the above. The first regards the deviation of the 2.3×10^{18} cm^{-3} data point in Figs. 9 and 12. Since $I_p \cong I_{\text{gen},s}$, this deviation in both figures is mainly due to surface tunneling resulting from the high doping concentration in both sides of this junction. The value of the surface tunneling current I_{st} can be estimated by assuming that the value of the surface generation current density for this junction is on the fitted line of Fig. 12 as indicated by the crossed circle. In the following this point will represent $I_{\text{gen},s}$ for this junction. The vertical distance between this point to the measured data point is the surface tunneling current I_{st}. The second point regards the calculation of s by Eq. (8a) assuming $I_{\text{gen},s} = I_p$, where A_s is the area of the surface depletion region around the periphery of the 1×1 mm^2 junctions. Fig-

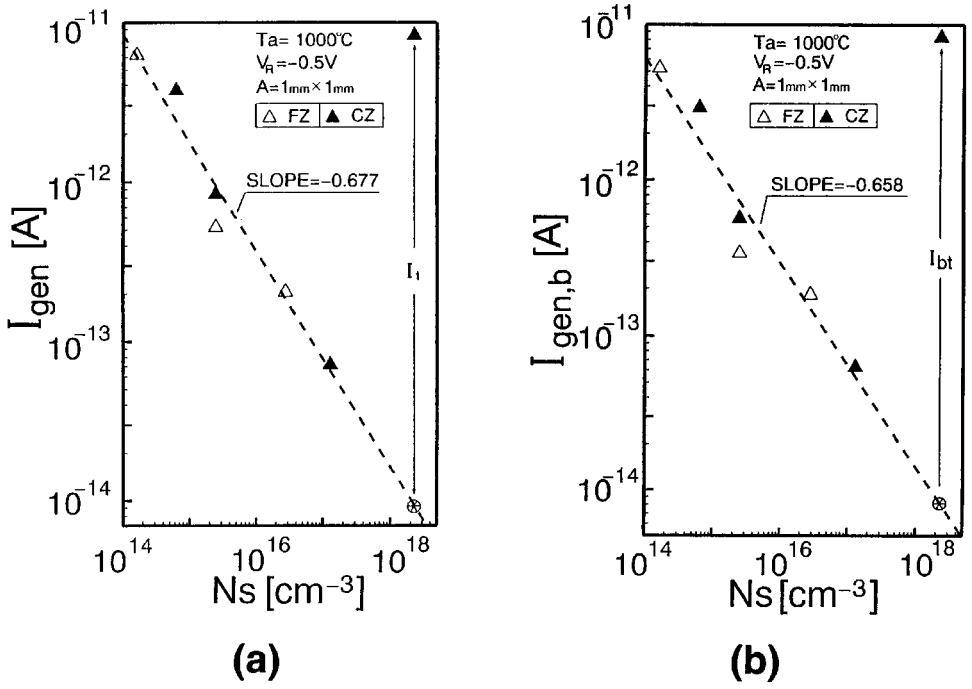

FIG. 14. (a) Overall reverse generation current I_{gen} as a function of N_s (1×1 mm^2 junctions). (b) Bulk generation current component $I_{\text{gen},b}$ as a function of N_s (1×1 mm^2 junctions).

ure 13 presents the surface recombination velocity as a function of N_s. Again, as in the case of Fig. 11, the values of s for the FZ and Cz crystals are different for the same reasons. It is noted that s decreases with increasing N_s, unlike previously published data, where surface recombination velocity for p-type silicon demonstrated only moderate changes up to the 10^{16} cm^{-3} doping concentrations and then an increase at higher concentrations.[11] This trend reversal of the surface recombination velocity is of principal significance and will be referred to in the discussion.

According to Eq. (15) the value of the total reverse generation current I_{gen} is calculated by subtracting I_d from I_R.

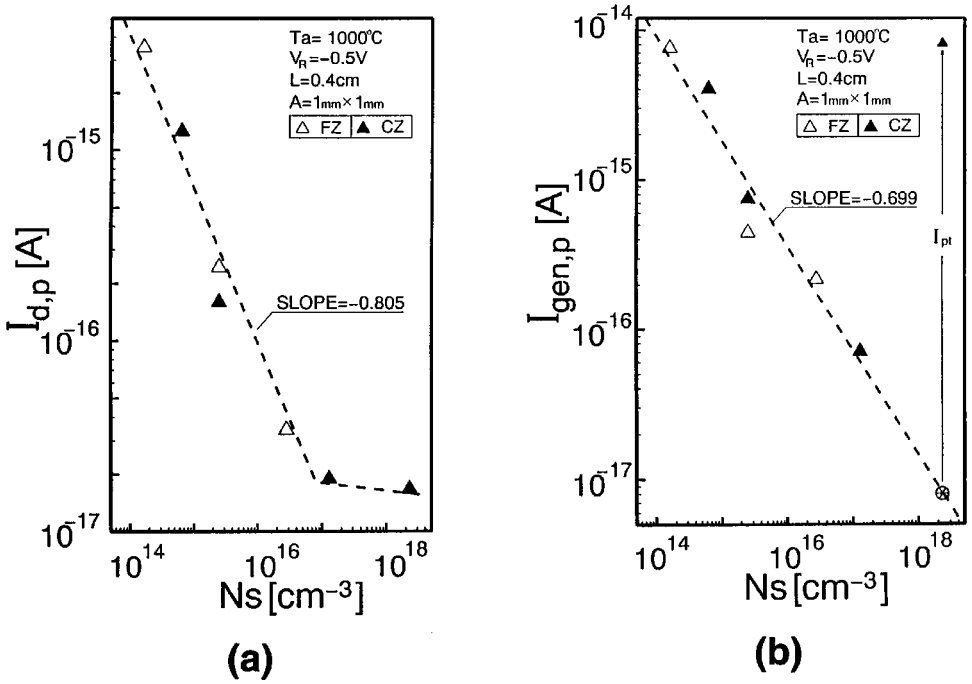

FIG. 15. (a) The peripheral (lateral) diffusion current component $I_{d,p}$ as a function of N_s (1×1 mm^2 junctions). (b) The peripheral (lateral) generation current component $I_{\text{gen},p}$, as a function of N_s (1×1 mm^2 junctions).

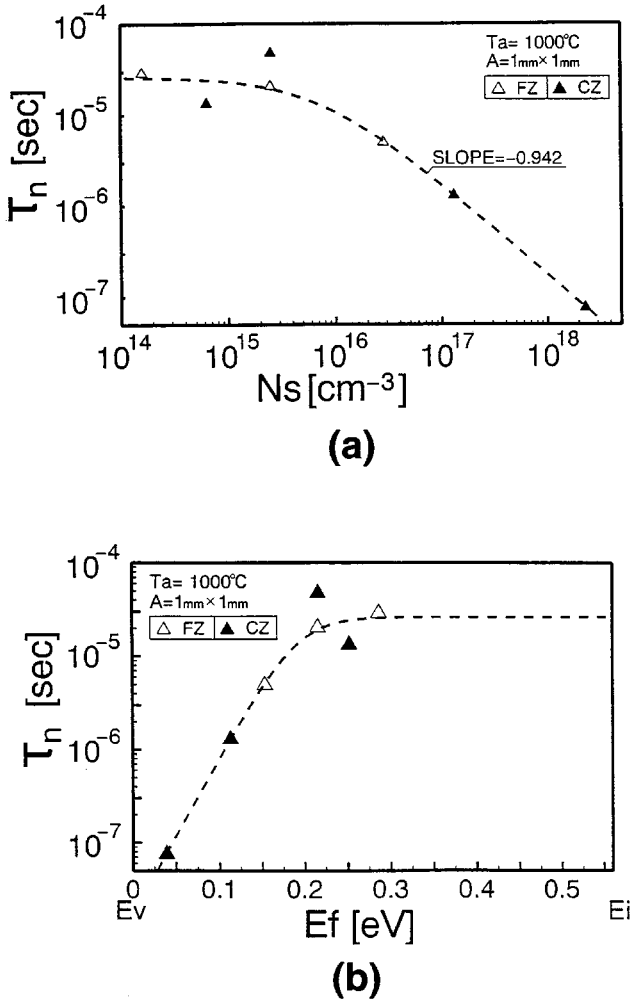

FIG. 16. (a) Minority carrier lifetime τ_n of electrons in the p substrate as a function of substrate dopant concentration, N_s (1×1 mm^2 junctions). (b) Minority carrier lifetime τ_n as a function of substrate Fermi level (1×1 mm^2 junctions).

The value of the reverse bulk generation current component $I_{\text{gen},b}$ is obtained by subtracting $I_{\text{gen},s}$ from I_{gen}. Figures 14(a) and 14(b) present I_{gen} and $I_{\text{gen},b}$, respectively, as a function of N_s. Again, in both cases the value of the 2.3×10^{18} cm^{-3} data deviates from the fitted straight line. This deviation in Fig. 14(b) is due to the existence of bulk tunneling effect. The data point in Fig. 14(a) deviated due to the combined surface (Fig. 12) and bulk [Fig. 14(b)] tunneling components $I_t=I_{bt}+I_{st}$. In both cases the values of I_{gen} and $I_{\text{gen},b}$ for the tunneling junction are approximated to lie on the fitted straight lines (crossed circles). In the following these points will represent I_{gen} and $I_{\text{gen},b}$ for the respective 2.3×10^{18} cm^{-3} junctions. The tunneling current in each case (I_t and I_{bt}, respectively), are determined as the vertical distance from the fitted straight line to the measured data point for this concentration. The tunneling current of Fig. 14(b) is responsible for the deviation of the 2.3×10^{18} cm^{-3} data point in Fig. 10. The area current density J_A is composed (Fig. 1) of J_d and $J_{\text{gen},b}$. Since I_d (Fig. 3) does not have any tunneling component the only contribution of tunneling to J_A in Fig. 10 comes from Fig. 14(b). Finally, the contribution of the two peripheral current components, namely $I_{d,p}$ and $I_{\text{gen},p}$ are crudely estimated assuming for simplicity that their current densities are the same as in the junction area. Figure 15(a) presents the lateral diffusion component (J_dLx_j) and Fig. 15(b) represent the lateral generation component ($J_{\text{gen},b}Lx_j$) both as a function of N_s. Comparing the results to Fig. 12 it can be seen that $I_{\text{gen},s}\gg I_{d,p}$, and $I_{\text{gen},s}\gg I_{\text{gen},p}$ situation exists, for about two orders of magnitude, verifying that in these junctions $I_p\cong I_{\text{gen},s}$. In Fig. 15(b) it can be seen that tunneling component exists, as expected.

According to Eq. (7a), in order to find $|E_t-E_i|$, τ_n and τ_{gen} should be determined. This is addressed now. The values of τ_n for the five one sided junctions of the high slope portion of Fig. 3 are obtained by the first term in Eq. (13b) which dominates over the second term in this N_s range. The values of the respective N_t were then calculated from Eq. (4) for these five junctions and were fitted to Eq. (10), yielding $N_t=3.8\times10^{12}+5.64\times10^{-4}N_s$ [cm^{-3}]. For estimation purpose, the N_t values for the remaining two sided junctions ($N_s=1.3\times10^{17}$ and 2.3×10^{18} cm^{-3}) were assumed to be on the same straight line and their N_t values were calculated accordingly. Then, the respective τ_n values for these two junctions were calculated from Eq. (4). The validity of using Eq. (4) for our junctions is dealt with in the sequel. The above procedure is based on the fact that Eq. (4) and Eq. (10) lead to Eq. (11), which is a well-established experimental dependence for τ_n.[12–14,17–19] Figure 16(a) shows the τ_n dependence on N_s with $\tau_{ni}=2.63\times10^{-5}$ s, which gave the best fit. The slope at the high concentration region of Fig. 16(a) indicates that τ_n is inversely proportional to N_s as given by Eq. (11). This equation provides the lifetime due to phonon assisted recombination only, and does not include Auger recombination process that may take place in the vicinity of 10^{18} cm^{-3} dopant concentration. This will be referred to in the discussion. Figure 16(b) shows τ_n as a function of the respective substrate Fermi levels which dominates τ_n[16] as discussed in Sec. V.

τ_{gen} is calculated from $|I_{\text{gen},b}|$ for all the junctions using Eq. (7b). The results are plotted in Fig. 17(a) which shows a weak dependence of τ_{gen} on N_s, namely $\tau_{\text{gen}}=1.5\times10^{-6}N_s^{0.186}$ s. The ratio τ_{gen}/τ_n is shown in Fig. 17(b). This ratio is needed in order to determine $|E_t-E_i|$ by Eq. (7a). It is interesting to note that despite the scattering in τ_{gen} [Fig. 17(a)], τ_{gen}/τ_n [Fig. 17(b)] is relatively ''smooth'' as a function of N_s because the scattering in τ_n and τ_{gen} are canceled out. The τ_{gen}/τ_n ratio increases significantly with N_s, ranging from 25 to 1746 for dopant concentrations from 1.6×10^{14} cm^{-3} to 1.3×10^{17} cm^{-3}. The value of this ratio obtained for the tunneling junction is exceptionally high (5.54×10^4). Remembering that for $E_t=E_i$ junctions, $\tau_{\text{gen}}/\tau_n=1$ [Eq. 7(a)], the result clearly shows that the influence of the deviation of E_t from E_i. $|E_t-E_i|$ is now calculated for each substrate concentration junctions from this ratio by Eq. 7(a) and the results are plotted in Fig. 18. Again, only a small scattering is obtained despite the scattering in τ_{gen}, due to the well behaved τ_{gen}/τ_n ratio. Figure 18 shows that as N_s increases, E_t deviates further from E_i. Possible explanation for this will be given in Sec. V. Similar arguments can be made for $|E_{st}-E_{si}|$ at the S–O interface. Using the data of Fig. 13 it

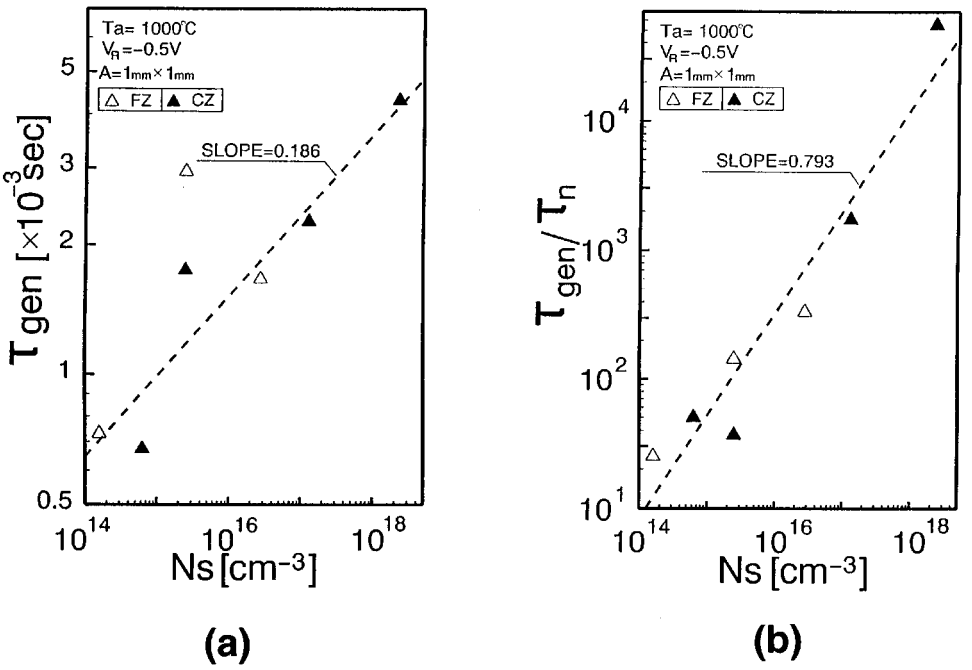

FIG. 17. (a) Generation lifetimes of the reverse biased junctions as a function of N_s (1×1 mm^2 junctions). (b) The ratio of depletion region lifetime to minority carrier lifetimes at the neutral zones of the junctions as a function of N_s (1×1 mm^2 junctions).

will be qualitatively shown in the discussion that E_{st} also deviates from E_{si}, as N_s increases.

In Sec. II τ_n was equated to τ_0 [$=1/\sigma v_{\text{th}} N_t$, Eq. (4)] under a given approximation. The validity of the above approximation is shown now for our junctions. Figure 19(a) shows $n_i \exp[|E_i-E_t|/kT]$ as a function of $N_s(=p_p)$. It is evident that for all the junctions under discussion, a situation where $N_s \gg n_i \exp[|E_i-E_t|/kT]$ exists, as was required in Sec. II,[4,7] when τ_n was equated to $1/\sigma v_{\text{th}} N_t$ [from Eq. (2)]. Depending on the doping concentration, p_p exceeds the concentration of holes in the trapping centers by orders of magnitude. This is more conveniently demonstrated by Fig. 19(b) where the results show that the $|E_i-E_F|>|E_i-E_t|$ situation exists for all the junctions, indicating that most of the trapping centers above E_F in the p side (substrate) are filled with holes. Accordingly, τ_n in these junctions is dominated by the position of E_F,[16] which is closer to E_v than E_t, which therefore, do not determine τ_n for the junctions under discussion. As a result τ_n is determined by the dependence of N_t on N_s only [Eq. (10)] which leads to Eq. (11). This will be further discussed in Sec. V.

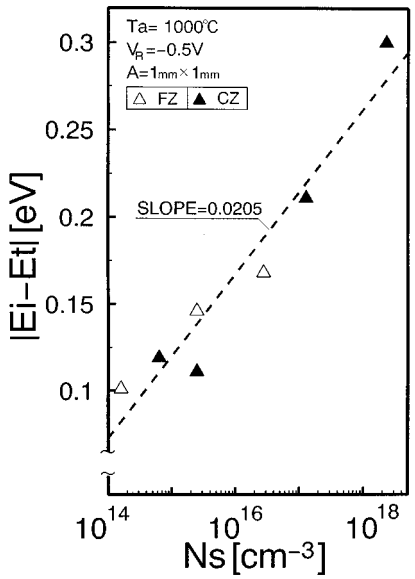

FIG. 18. The absolute value of the deviation of the trapping centers energy E_t, from mid-band-gap energy E_i, as a function of substrate dopant concentration N_s.

The ratio of I_{gen}/I_d is given in Fig. 20. Two things are noted from the measured data. The first is the very small ratios obtained, ranging from less than unity to less than 7.2, and the second, that a maximum exists. This ratio is significantly smaller than that obtained for junctions where $E_t=E_i$ and $E_{st}=E_{si}$ situation exist (mainly in diffused junctions made using current technology clean room), which exhibit ratios ranging from few tens to few hundreds. These small ratios in Fig. 20 therefore present marked improvement which was achieved by the deviation of E_t from E_i and E_{st} from E_{si}, a fact that reflects on the junction quality. The shape of this dependence on N_s can be understood from Figs. 3 and 14(a). At the low N_s region, I_d decreases at a faster rate then I_{gen} while maintaining $I_d<I_{\text{gen}}$, resulting an increase in this ratio with N_s. At the high N_s region, this ratio decreases because I_{gen} continues to decrease, while I_d changes the slope due to the addition of the significant p side injection, causing an almost constant I_d. The contribution of bulk generation current $I_{\text{gen},b}$ and the surface generation current $I_{\text{gen},s}$ to the overall ratio of Fig. 20 are presented separately in Figs. 21(a) and 21(b), respectively. The results clearly show that the ratio of Fig. 20 is determined mainly by

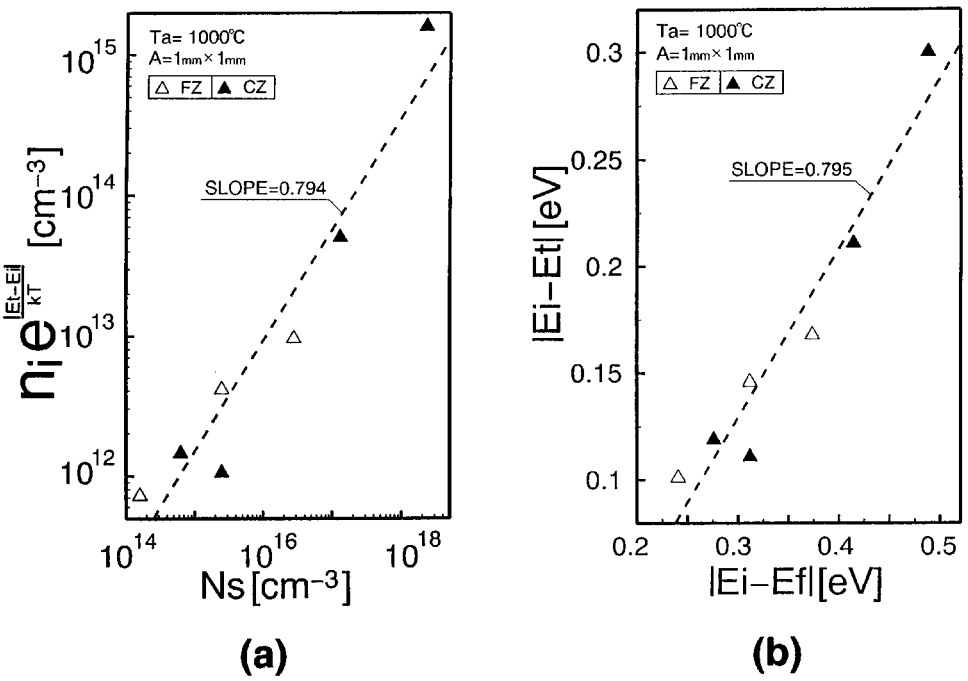

FIG. 19. (a) $n_i \exp[|E_i-E_t|/kT]$ as a function of N_s (1×1 mm^2 junctions). (b) $|E_i-E_t|$ as a function of $|E_i-E_F|$. An indication that in these junctions τ_n is dominated by E_F and not by E_t (1×1 mm^2 junctions).

the bulk generation component. Both curves of Fig. 21 preserve the shape of Fig. 20, for similar reasons. The dashed lines of Figs. 20 and 21 represent the fitting done by using Eqs. 7(b), 8(a), 9(a), and 13(b). The values of s, τ_{gen}, and τ_n in those equations were taken from the fitted values of Figs. 13, 16, and 17(a), respectively. D_n as a function of N_s was obtained from mobility data using $D/\mu=kT/q$.

A representative example for reverse current temperature dependence is given in Fig. 22 for junction made on Cz substrate with $N_s=2.5\times10^{15}$ cm^{-3} boron concentration. Two slopes are observed, however neither exhibits the expected activation energy of $-E_g$ and $-E_g/2$ which would occur if $E_t=E_i$ and $E_{st}=E_{si}$ situation existed. This is probably because the slopes in these junctions do not purely represent the respective diffusion and generation regimes, since in these junctions relatively high surface current exists. According to Eq. (12) this can give rise to the deviation from $-E_g/2$ as noted in the low temperature region where $1000/T=3$–4, where $I_{\text{gen},b}$ usually dominates (in the $E_t=E_i$ case). The absolute value of the activation energy for the higher temperatures in the case of Fig. 22 is only about the band-gap energy. From Eq. (12) it is clear why $|E_t-E_i|$ cannot be evaluated in our case by the slope of I_R/n_i as a function of $1000/T$ as was suggested by Schroder.[1] In his junctions as well in other published cases this approach was possible, since I_{gen}/I_d was very high, and surface current was neglected so that the low temperature regime presented practically the pure bulk generation current ($I_R \simeq I_{\text{gen},b}$). As a result a new method should be devised for the determination of $|E_t-E_i|$ by temperature measurements in future high quality junctions which is capable to resolve $I_{\text{gen},b}$ as a function of T. Finally, as $|E_t-E_i|$ increases, the slope of the low temperature region should approach that of the high temperature region. This was indeed observed for both the Cz and FZ junctions at the low substrate concentrations. However, this trend was reversed for $N_s>2.8\times10^{16}$ cm^{-3} for both Cz and FZ substrates. This phenomena is not understood and it is still under investigation.

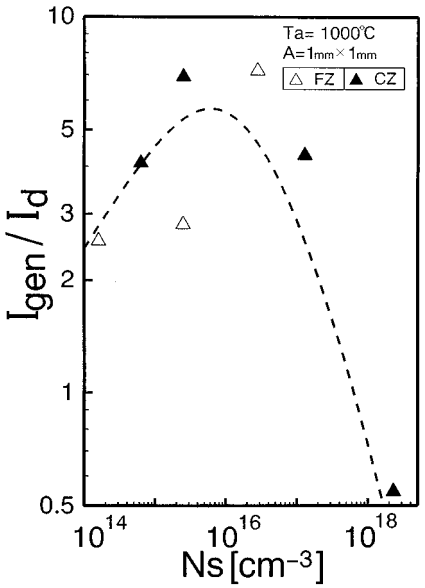

FIG. 20. Ratio of reverse overall generation current I_{gen} to the diffusion current I_d, as a function of substrate dopant concentration N_s (1×1 mm^2 junctions).

V. DISCUSSION

The experimental results clearly demonstrated that the bulk depletion region generation parameters, namely, τ_{gen},

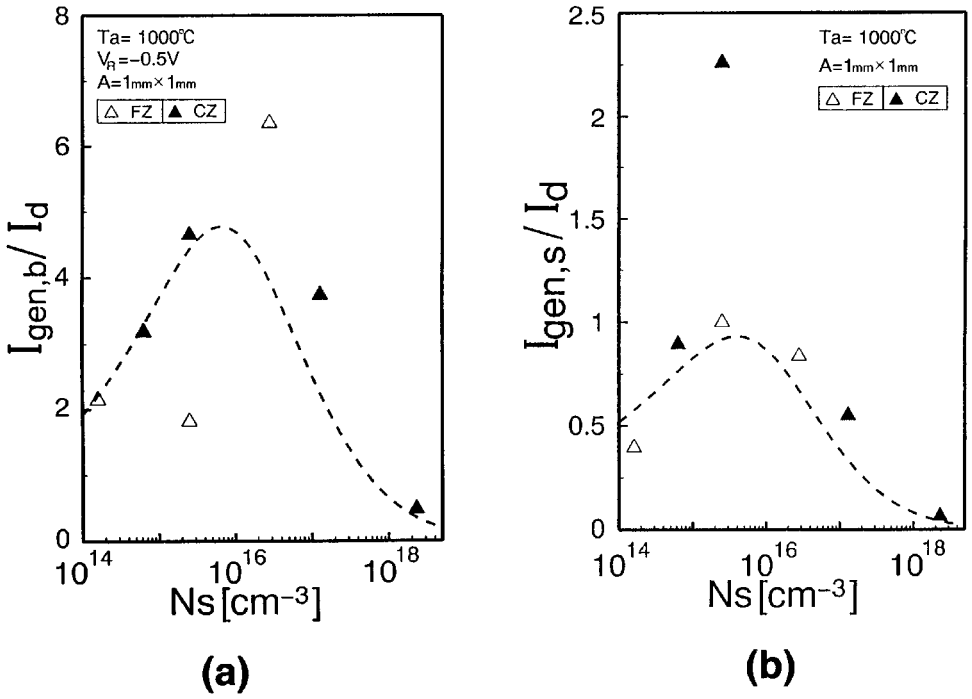

FIG. 21. (a) The ratio of the bulk generation current to the diffusion current as a function of N_s (1×1 mm^2 junctions). (b) The ratio of the surface generation current to the diffusion current as a function of N_s (1×1 mm^2 junctions).

$I_{gen,b}$, and the resulting I_{gen}/I_d and $I_{gen,b}/I_d$ ratios, as well as n and m are all function of $|E_t-E_i|$ and N_t which were in turn found experimentally to be dependent on N_s. As a starting point for the discussion it is recalled that all the formulas presented in Sec. II assumed as a simplifying assumption that in the bulk only one energy level (E_t) exists within the band gap. The real situation however is that there are several levels distributed within the band gap, each of them contributing to the generation/recombination process according to its position within the band gap [Eq. (2)]. E_t in this case only symbolizes an ''effective'' or ''equivalent'' value which represent the weighted influence of all these energy levels. The $|E_t-E_i|$, N_s relation (Fig. 18) is determined by an interplay among three main sources of energy levels. The first is the existence of initial energy levels within the substrate itself, prior to device processing. The second is the unintentional introduction of energy levels by the addition of undesired foreign atoms contamination during processing to the active parts of the junction. Finally, the introduction of energy levels by the junction fabrication process itself (ion implantation, diffusion, etc), its parameters (implantation energy, diffusion temperature, etc), and its quality. The overall behavior of the junction performance depends on the relative contribution of the above sources through $|E_t-E_i|$, N_t, $|E_{st}-E_{si}|$, and N_{st}. The contribution of each source and its possible correlation with N_s is now discussed.

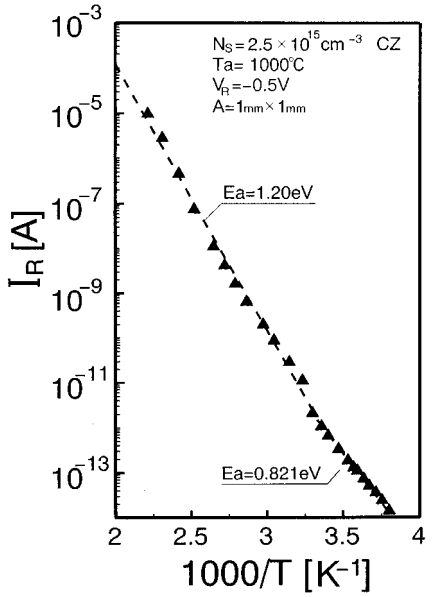

FIG. 22. The reverse current as a function of $1000/T$ of the $N_s=2.5\times10^{15}$ cm^{-3} Cz substrate (1×1 mm^2 junctions).

The initially existing energy levels within the substrate prior to device fabrication are formed mainly by undesired foreign atoms, the dominant being heavy metals, which constitute energy levels near E_i, energy levels produced by inherent crystal defects, and energy levels induced by the existing substrate dopant atoms themselves, apart from the acceptor level. The latter are responsible for part of the N_s dependence of $|E_t-E_i|$. The exact mechanism of formation of the dopant related energy levels is not clear, however two possible mechanisms are speculated for their formation. One is the small amount of local stress caused by the misfit in the covalent radii of the Si and the substrate dopant (B) atoms. Once this stress is above a critical value, an interaction with some extended defect (dislocation, oxide precipitate) is developed and hence an energy level results. The existence of

defect levels in Czochralski grown silicon due to oxide precipitates, and their relation to prior thermal history have been shown.[16,41–43] Their intentional formation has been demonstrated as well.[44] The added stress σ_1 caused by the dopants is directly proportional to their concentration[45] (see the Appendix) a fact that is well correlated with the N_s dependence of N_t [Eq. (10)]. The other mechanism, is a possible weakening of the bonds between the host Si atoms caused by the dopant (B) atoms, due to the difference in their valence electrons. It is speculated that electron leaving the Si–Si covalent bond to fill the empty B–Si bond cause weakening of the neighboring Si–Si bonds. This can also induce interaction with extended defects. Both mechanisms are dopant concentration dependent. In addition, the substrate contains dopant atoms hosting lattice sites, that possess local stress (or bond weakening), resulting from the above mechanisms, which are not yet enough to create or interact with extended defects and form energy levels. Statistically, however, some of these sites bear the potential of creating recombination centers by such an interaction once enough energy is added to them. This will be referred as an ''activation.'' Their concentration is also N_s dependent. In the following, they will be referred to as the ''weak sites.''

The merit of the ultraclean technology used here is that the second source, namely, the introduction of new energy levels near E_i to the junction during device fabrication is *almost eliminated*, due to the marked reduction of foreign atoms originated within the fabrication facility, as described in Sec. III. As a result, the main cause for the deviation of E_t from E_i is the *different nature and density* of the initially existing (mainly those close to E_i) energy levels and newly induced energy levels by the ion implantation irradiation, which are far from E_i. This is an important point with respect to other processing technologies, that do introduce deep levels through the unintentional introduction of unwanted heavy metals atoms to the junction during processing. In that case, the density of the energy levels near E_i increases due to the additive nature of the process, and they become the dominant ones for the overall bulk generation/recombination process. As mentioned, this is largely eliminated in the present case. The increase in $|E_t-E_i|$ with N_s in our case occurs due to the residual vacancies or divacancies[46] created by the implantation, left after annealing, which rapidly develop into complex lattice defects, and act as recombination centers, and due to the ''activation'' of some of the ''weak sites'' by the irradiation. These newly induced energy levels contain, accordingly, an inherent N_s dependent component. They are far from E_i, and due to their large density (with respect to the initially existing ones) become increasingly dominant. As a result the balance is now gradually shifted towards the capturing of more and more carriers in the newly added levels as N_s increases. This is expressed in the experimental results as a deviation of E_t from E_i (Fig. 18).

The addition of the newly introduced energy levels with higher $|E_t-E_i|$ causes a decrease in the generation/recombination rate [Eq. (2)] and an increase in the generation lifetime, τ_{gen}, with N_s [Fig. 17(a)]. This is essentially expressed in Eq. 7(a). Note, that this equation indicates that τ_{gen} is determined by two opposing trends, since as N_t increase with N_s, τ_{gen} tends to decrease, however, $|E_t-E_i|$ is the dominating factor in this case causing an overall increase in τ_{gen} with N_s. Examining for example the junction made in the 2.8×10^{16} cm^{-3} substrate, which is a representative one for IC circuits, show that for $E_t=E_i$ case $\tau_{\text{gen}}=\tau_n=4.95\times10^{-6}$ s. Since however in this case $|E_t-E_i|=0.168$ eV, τ_{gen} increases to 1.67×10^{-3} s. This in turn will be shown to decrease $I_{\text{gen},b}$ for this substrate junction, as happens for all the other junctions with varying N_s [Fig. 14(b)]. This reduction in $I_{\text{gen},b}$ with N_s comes in addition to its reduction by N_s through W_R [Eq. 7(b)], as implied by the slope of Fig. 14(b).

Similar arguments can be made for results obtained for the surface behavior. The trend reversal of the surface recombination velocity in our data (Fig. 13) in comparison to earlier data[11] can be correlated to the ultraclean technology used, by the aid of Eqs. 8(c) and 8(d). The earlier data cited dealt with diffused junctions where $E_t=E_i$ and $E_{st}=E_{si}$ situation existed as indicated by the authors. This means $s=s_0$ situation for these junctions. Unlike the values of N_t in the bulk, N_{st} on the surface could not be obtained from our data. Assuming however that like N_t in the bulk, N_{st} increases as well with N_s, this can explain the fact that the surface recombination rate in the earlier data[11] exhibited an increasing trend with N_s [Eq. 8(d)] since s_0 is directly proportional to N_{st}. The reduction of the surface recombination velocity as a function of N_s (Fig. 13) in our case can be explained assuming that similarly to the bulk, $|E_{st}-E_{si}|$ increases as well with N_s. According to Eq. 8(c) this can cause a trend reversal of s in the case where the denominator increases much faster with N_s than the numerator. This in turn explains the reduction of $I_{\text{gen},s}$ (Fig. 12) with N_s, with relation to Eq. 8(a). The slope of Fig. 12 indicates that this reduction comes in addition to its reduction by A_s, as a function of N_s.

The results show (Fig. 20) that I_{gen}/I_d ratios, obtained for these junctions, range from values close to unity, up to about 7, at the maximum point. These significantly small ratios (with respect to ratios of tens to hundreds for junctions with $E_t=E_{st}=E_i$) originate from the fact that I_d and I_{gen} are dependent in a different fashion on their respective lifetimes τ_n, τ_{gen}, and s. While τ_n practically depend only on N_t [Eq. (4)] and its dependence on E_t is negligible [$N_A \gg n_i \exp[|E_t-E_i|/kT]$, or $|E_i-E_F|>|E_i-E_t|$, see Figs. 19(a) and 19(b)], τ_{gen} and s on the other hand are dependent on N_t and N_{st}, respectively, and dominantly on $|E_t-E_i|$ and $|E_{st}-E_{si}|$, respectively. Accordingly, from Eq. 13(b) and Eq. 9(a), it is evident that the I_{gen}/I_d ratio is sensitive to $|E_t-E_i|$ and $|E_{st}-E_{si}|$ variations. Examine again for example, the junction produced on the 2.8×10^{16} cm^{-3} substrate, with $|E_t-E_i|=0.168$ eV and $I_{\text{gen}}/I_d=7.2$ (Figs. 18 and 20, respectively). This ratio is composed of $I_{\text{gen},b}/I_d$ [Fig. 21(a)] and $I_{\text{gen},s}/I_d$ [Fig. 21(b)]. Since the first ratio is the dominant, it is enough to examine it in order to characterize I_{gen}/I_d for this substrate. From about 6.3 this ratio increases drastically by factor of 337–2022 if $E_t=E_i$ situation would exist, increasing I_{gen}/I_d accordingly. Therefore I_{gen}/I_d can be regarded as a ''figure of merit'' for dc junction quality, in cases where $|E_i-E_F|>|E_i-E_t|$ condition exists. This

point will be further addressed in the Sec. VI. The advantage of this figure of merit is that it produces a number, rather than referring to a ''low reverse current,'' and that it is easy to determine by simple dc measurements. It also provides a real indication for process ''cleanness.'' Finally, the low $I_{\rm gen}/I_d$ ratios mean that for these junctions, $\tau_{\rm gen}$ cannot be calculated directly from I_R as is routinely done[1] in the case where $E_t=E_i$ or E_t is close to E_i (and hence $I_{\rm gen}\gg I_d$) or by neglecting the surface current. Instead, $\tau_{\rm gen}$ should be calculated from $I_{{\rm gen},b}$, which is found by first obtaining I_d from the $n=1$ region (taking into account the R_s effects) and subtracting it from I_R in order to find $I_{\rm gen}$. Then $I_{{\rm gen},b}$ should be found by subtracting $I_{{\rm gen},s}$ from $I_{\rm gen}$. $\tau_{\rm gen}$ is then calculated from $I_{{\rm gen},b}$. This is specially true, according to Figs. 20 and 21(a), for low and high substrate concentrations junctions, where these ratios are particularly low.

The dependence of N_t on N_s given in Eq. (10) for As irradiation of B-doped silicon in this case is similar to the N_t dependence resulting by high energy electron and γ-ray irradiation of silicon.[47,48] This dependence led in both cases to the same dependence of minority carrier lifetime, as in Eq. (11). This indicates that high energy irradiation of Si, in general, creates defects which lead to the reduction of carrier lifetime in the form of Eq. (11). While the form of Eq. (11) is general and routinely used for lifetime characterization, the process parameters, such as implantation energy irradiation, and annealing temperature, determine the magnitude of the newly added recombination centers concentration with conjunction to N_s. This is expressed by the constant $B(=dN_t/dN_s)$ as well as τ_{ni}, which are process dependent parameters.[17,18] The creation of these levels does not automatically imply that all of them participate in the recombination process, but only a fraction of them, which increases with N_s. An explanation of the lifetime dependence of minority carriers in the neutral parts of the junction on doping concentration was given by Meier *et al.*[16] through the Fermi level position. It is based on the fact that not all the energy levels but only those which are filled with majority carriers can contribute to the recombination process with minority carriers under certain conditions.[16] As N_s increases, the Fermi level moves towards the majority carrier band edge (valence band in our case). This opens a ''window'' with a width of $2|E_F-E_i|$ which includes larger fraction of the existing energy levels which are filled with majority carriers, increasing the recombination rate and accordingly decreasing the minority carrier lifetime. This situation is essentially expressed in this work by the condition $|E_i-E_F|>|E_i-E_t|$ given in Fig. 19(b), which means that in the junctions under discussion τ_n is dependent on E_F and not on E_t, as previously mentioned.

As pointed earlier, Eq. (11) provides the lifetime due to phonon assisted recombination process. It does not include the Auger recombination process, which may exist in the vicinity of 2.3×10^{18} cm^{-3} substrate concentration. This can introduce some error that will be addressed later on. However, it should be emphasized that its existence or nonexistence is *irrelevant* as far as the bulk generation parameters are concerned, namely, $I_{{\rm gen},b}$ and $\tau_{\rm gen}$. This is because both are depletion region parameters, and accordingly, require only the phonon assisted lifetime for their determination, as seen from Eqs. (6) and (7), which were derived from Eq. (2). As a result, if the minority carrier lifetime in the neutral parts of the junction is determined, say, by a measurement that indicates that Auger recombination process component exists (at the high doping concentrations, of about 10^{18} cm^{-3} or greater), then it would be necessary to resolve from the measured data the phonon assisted component of the lifetime for the determination of $\tau_{\rm gen}$ and $I_{{\rm gen},b}$. Accordingly, the value of τ_n obtained for the 2.3×10^{18} cm^{-3} junction is the correct one as far as $\tau_{\rm gen}$ and $I_{{\rm gen},b}$ are concerned. However, its value might be larger, as far as the diffusion current I_d is concerned, for the same junction, if Auger process exists, since $1/\tau=1/\tau_n+1/\tau_A$, where τ_A is the lifetime due to the Auger process. Accordingly, the values inferred from τ_n for the diffusion current components calculated for this junction in Figs. 20 and 21(a) for fitting purpose are left in the figures in order to indicate the general trend only.

Two final points regarding N_t and E_t should be noted. The first is what, at the first glance, may be seen as a ''paradox.'' It is noted from the experimental results that as N_t increases (with N_s) the recombination/generation rate decreases as seen by $I_{{\rm gen},b}$ and $\tau_{\rm gen}$, while one would expect the opposite. The solution to the ''paradox'' is that U [Eq. (2)] is determined simultaneously by both N_t and E_t, the latter being the dominant one. While U is linearly dependent on N_t, it is inversely dependent on $\exp(E_t-E_i)$, which in our junction cause an overall reduction in U as N_s increases, despite the fact that N_t increases. This might be explained as follows: as N_s increases, N_t increases as well [Eq. (10)] and the number of generation/recombination events increase, but due to the position of E_t, which becomes far from E_i, the carriers are trapped for shorter time, and then released to the bands, giving rise to larger $\tau_{\rm gen}$, and thus decreasing the recombination rate, U. The second point regarding N_t and E_t is that in Eq. (2) and in the following equations, all the recombination centers (N_t) are assumed to be evenly distributed in the crystal volume, an assumption which is probably correct in the substrate within reasonable limits, before processing. The real situation within the active parts of the n^+p junctions under discussion, however, is different since N_t is not constant (specifically within the depletion region) and deviates by a magnitude which depends on the implantation/annealing parameters in conjunction with N_s. According to Eq. (10) it can be expected that the recombination/generation centers distribution has a profile which is related to the doping profile in both junction sides. This means that N_t and E_t are different not only from junction to junction as a function of N_s, but vary also within each junction. This fact bear two results. The first that like E_t, N_t in the equations used, is some ''effective'' or some weighted average value. The second is that the results obtained for N_t and E_t are dependent on the magnitude and polarity of junction bias. This is because a change in bias results in a change in W which may include higher or lower values of N_t and E_t depending on the direction and magnitude of the change. Therefore, when providing numbers for N_t and $|E_t-E_i|$, the experimental conditions (bias) should be specified. The effect of the above will

somewhat affect the magnitude of N_t and $|E_t-E_i|$, but not their trends with N_s.

The increase in N_t with N_s, can also explain the drastic fall in R_p in Fig. 7, as N_s exceeds some critical concentration. The equivalent circuit used for deriving Eq. (14), includes R_p directly parallel to the ideal junction. Therefore R_p in this case does not include any bulk or contact properties related to R_s but it does provide a measure of the ohmic behavior of the depletion region. In the fitting process, R_p is mainly determined by the computer program from the lower current region of the I-V characteristics which overlap part of the I-V region for the simultaneous determination of I_{gf}. I_{gf} is created by the action of discrete recombination centers, which are distinctly separated from each other, and have no ''coupling'' between them. They trap and emit carriers from and to the respective bands, and as such, their action represent *electronic activities*, which is manifested by I_{gf}. As N_t increases above a certain value, some of the defects possibly together with others which were electronically inactive at low doping concentrations, may form a path through which the carriers are moving from one center to another and not through the bands. These current ''channels'' represent an *ohmic* current component I_Ω, which is manifested by the fall in R_p and represent an *electrical activity* which means loss, in this case. It is not clear if the path is continuous, or if some ''jumping'' mechanism of carriers exists between some ''islands'' of defects in the path. In any case, the value of the critical point for R_p drastic reduction, in Fig. 7 (at $N_s \geqslant 2.5\times10^{15}$ cm^{-3}), depend on the technological process which introduces the defects and determine their properties. These depletion region currents, namely I_Ω, I_{gf}, and the diffusion current I_d, are comparable in their values, in the low R_p regime ($N_s > 2.5\times10^{15}$ cm^{-3}) of Fig. 7. Accordingly, neglecting R_p from the equivalent circuit in this regime will result in an error in the determination of the values of I_d and even more in I_{gf}. It is noted that at the same time, R_s (Fig. 6) is decreasing monotonicaly with N_s, in contrast to the drastic fall in R_p. This is due to the different mechanisms of current transport within the low and high field regions. Defects in the neutral bulk region have much weaker effect on the resistance, which is mainly determined by dopant concentration. The drastic fall of R_p on the other hand, is defect dominated, rather than N_s.

VI. CONCLUSION

Using ultraclean process technology and the related suppression of unintentional introduction of heavy metal atoms, as well as other contaminants to the active parts of the junction was shown to cause a principal change in junction behavior resulting marked improvements in its dc performance quality. It is concluded that this is achieved by altering the recombination process itself, which results from a shift of E_t from E_i. This shift is shown to be dependent on N_s, and as such, enables one to control the marked increase in τ_{gen}, and the parallel reduction in I_{gen}, and I_{gen}/I_d ratios. The $|E_t-E_i|$, N_s relation can therefore be used for device design considerations. This is in contrast to the previous situation where I_{gen} as well as other generation parameters were largely determined by parasitic effects.

From Fig. 20 it is seen that low values of I_{gen}/I_d can be attained for both low and high N_s values. Generally speaking, from the point of view of low I_{gen}/I_d considerations, it seems more desirable to choose high N_s values because in such a case both I_{gen} and I_d are low [Figs. 14(a) and 3, respectively] in addition to their low ratio. Obviously, the doping concentrations are determined by other device performance requirements as well. However, a situation may arise, that in future ULSI devices design considerations, a tendency towards choosing high N_s values may be preferred, if related low temperature annealing problems will be solved. This is due to several reasons, some of which are outlined here. The reduction in junction area will result in an increased R_s, and the increased N_s, with conjunction to other structural parameters, can alleviate this problem. In addition, high N_s may reduce surface inversion problems to a degree that might eliminate guarding, altogether, increasing device packaging density. Finally, the scaling down is accompanied by a significant reduction in operation currents, and therefore necessitate a parallel reduction in noise levels, to ensure reasonable signal to noise ratios and reduce information errors. Since noise level is directly proportional to I_R, choosing high N_s values, where I_R is small (Fig. 2), is a viable solution.

The value of τ_n was determined to be $1/\sigma v_{\text{th}} N_t$ for our junctions, under the condition that $|E_i-E_F| > |E_i-E_t|$, which was shown to prevail for this case. If, however, in the future, better substrate quality with greatly suppressed (or eliminated) heavy metal atom content will be realized, the above inequality may be reversed. This can cause three results: an increase of minority carrier lifetime ($\tau_n > 1/\sigma v_{\text{th}} N_t$) according to Eq. (2) in the junction neutral zones, as a function of the deviation of E_t from E_i (similarly to τ_{gen}). Secondly, an improvement in I_d, which will be decreased as a result of increased τ_n, although in a less drastic fashion than $I_{\text{gen},b}$ (and hence I_{gen}). This is because $I_{\text{gen},b}$ is inversely proportional to τ_{gen}, while I_d is inversely proportional to $\tau_n^{1/2}$. As a result, the sensitivity of I_{gen}/I_d ratio to $|E_t-E_i|$ will decrease, a fact that will stabilize its value against the influence of the above mentioned parasitic effects. In any case, the simultaneous reduction in both I_{gen} and I_d as a function of $|E_t-E_i|$, will further reduce I_R, improving dc junction performance quality. As a paradox, it should be noted that the I_{gen}/I_d ratios might increase in such a case (namely, if $|E_i-E_F| < |E_i-E_t|$), despite the reduction in I_R. Therefore, the I_{gen}/I_d ratio can serve as a figure of merit for $|E_i-E_F| > |E_i-E_t|$ situation only, which, at present, was experimentally shown to be a realistic situation.

VII. SUMMARY

It was experimentally demonstrated that the suppression of the introduction of undesired heavy metal atoms and other contaminations to the active parts of ion implanted n^+p junctions during device processing, results in significant improvement in the junction quality, mainly through the generation parameters, namely, I_{gen}, τ_{gen}, n, and m. It is concluded that these improvements are result of the deviation of E_t from E_i, and E_{st} from E_{si} which were found to increase with N_s. Under the ultraclean processing technology used

here, the number of new energy levels introduced near E_i during processing is minimized. Accordingly, the above deviation of E_t and E_{st} is dominantly a result of a balance between initially existing energy levels in the substrate prior to processing (with mainly E_t close to E_i levels) and newly created energy levels by the implantation irradiation and post-annealing, both having an N_s dependent component. These new energy levels are far from $E_i(E_{si})$, and due to their high density which increase with N_s, cause the above deviation. I_{gen}/I_d ratios were found to be extremely small, with respect to their values in the $E_t=E_i$ case, due to the fact that while $I_{\text{gen},b}$ is dependent on both N_t and E_t, I_d in these junctions is practically dependent only on N_t. The above approach can, under these fabrication conditions, obtain better quality control for device performance. It was accordingly proposed that the criteria for quality determination will be defined as I_{gen}/I_d, which serves as a figure of merit for dc junction performance.

ACKNOWLEDGMENT

This work was carried out in the Super Clean Room of the Laboratory for Electronic Intelligent Systems, Research Institute of Electrical Communication, Tohoku University, Japan.

APPENDIX: IMPURITY RELATED STRESS

The reduction in lifetime [Eq. (11)] with $N_s(=N_A)$ was related to the increase in N_t [Eq. (10)]:

$$N_t=N_{ti}+BN_s. \tag{A1}$$

The second term in Eq. (A1) represents the *added trapping centers* due to the presence of dopants. A possible mechanism for the formation of the trapping centers is a local stress σ_1 [dyne/cm^2] in a lattice site that host an impurity atom. This stress develops due to the misfit in the covalent radii of the Si atoms ($r_{\text{Si}}=0.104$ nm) and the Boron atoms ($r_{\text{B}}=0.088$ nm),[49] and its value is calculated by:[45]

$$\sigma_I=\left[1-\left(\frac{r_d}{r_{\text{Si}}}\right)^3\right]\left(\frac{N_s}{N_{\text{Si}}}\right)\frac{Y_1}{1-P}, \tag{A2}$$

where Y_1 is the Young's modules (1.31×10^{11} dyne/cm^2), P is the Poisson's ratio (-0.33), N_{Si} is the concentration of silicon atoms (4.96×10^{22} cm^{-3}), which yields:

$$\sigma_1=AN_s, \tag{A3}$$

where $A=7.83\times10^{-12}$ dyne cm. Equation (A3) shows that the *added stress* due to the presence of dopant atoms is directly proportional to their concentration, a fact that is in accordance with the second term of Eq. (A1).

[1] D. K. Schroder, IEEE Trans. Electron Devices **ED-29**, 1336 (1982).
[2] W. Schockley and W. T. Read, Phys. Rev. **87**, 835–842 (1952).
[3] R. N. Hall, Phys. Rev. **87**, 387 (1952).
[4] A. S. Groove, *Physics and Technology of Semiconductors Devices*, (Wiley, Singapore, 1967).
[5] C. T. Sah, R. N. Noyce, and W. Schockley, Proc. IRE **45,** 1228 (1957).
[6] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, Singapore, 1981).
[7] M. V. Whelan, Solid-State Electron. **12**, 963 (1969).
[8] J. Cornu, R. Sitting, and W. Zimmermann, Solid-State Electron. **17**, 1099 (1974).
[9] S. K. Ghandi, *Semiconductor Power Devices* (Wiley, New York, 1977).
[10] A. S. Grove and D. J. Fitzgerald, Solid-State Electron. **783–806**, 1966.
[11] D. J. Fitzgerald and A. S. Grove, Surf. Sci. **9**, 347 (1968).
[12] D. B. M. Klaassen, Solid State Electron. **35**, 961 (1992).
[13] J. G. Fossum, R. P. Mertens, D. S. Lee, and J. F. Nijs, Solid State Electron. **26**, 569 (1983).
[14] M. S. Tyagi and R. Van Overstraten, Solid State Electron. **26**, 577 (1983).
[15] H. S. Bennet, Solid State Electron. **27**, 893 (1984).
[16] D. H. Meier, J. O. Hwang, and R. B. Campbell, IEEE Trans. Electron Devices **ED-35**, 70 (1988).
[17] D. J. Roulstone, N. D. Arora, and S. G. Chamberlaine, IEEE Trans. Electron Devices **ED-29**, 284 (1982).
[18] S. Bellone, G. Busatto, and C. M. Ransom, IEEE Trans. Electron Devices **ED-8**, 532 (1991).
[19] H. S. Bennett, IEEE Trans. Electron Devices **ED-30**, 920 (1983).
[20] A. Neugroschel, F. A. Lindholm, and C. T. Sha, IEEE Trans. Electron Devices **ED-24**, 662 (1977).
[21] T. Ohmi, N. Mikoshiba, and K. Tsubouchi, *ULSI Science and Technology/1987*, edited by S. Broydo, and C. M. Osburn (The Electrochemical Society, Pennington, 1988), pp. 761–85.
[22] T. Ohmi, Microcontamination **6**, 49 (1988).
[23] K. Sugiyama and T. Ohmi, Microcontamination **6**, 49 (1988).
[24] Y. Kanno and T. Ohmi, Microcontamination **6**, 23 (1988).
[25] K. Sugiyama, T. Ohmi, T. Okumura, and F. Nakahara, Microcontamination **7**, 37 (1989).
[26] K. Sugiyama, F. Nakahara, and T. Ohmi, Microcontamination **7**, 29 (1989).
[27] K. Sugiyama, T. Ohmi, Y. Mizoguchi, and F. Nakahara, *Automated Integrated Circuits Manufacturing*, edited by V. Akins (The Electrochemical Society, Pennington, 1990), Vol. PV90-3, pp. 148–172.
[28] K. Yabe, Y. Motomura, H. Ishikawa, T. Mizuniwa, and T. Ohmi, Microcontamination **7**, 37 (1989).
[29] T. Ohmi, and M. Yasuda, Microcontamination **7**, 23 (1989).
[30] T. Ohmi, H. Inaba, and T. Takenami, Microcontamination **7**, 29 (1989).
[31] T. Ohmi, H. Inaba, and T. Takenami, Microcontamination **7**, 29 (1989).
[32] T. Ohmi and N. Mikoshiba, in Proceedings of the First International Symposium on ULSI Science and Technology, Philadelphia, 1987, Abst. No. 212.
[33] T. Ohmi, K. Masuda, T. Hashimoto, T. Shibata, M. Kato, and Y. Ishihara, in Extended Abstracts of the Conference on Solid State Devices and Materials, Tokyo, 1987, pp. 299–302.
[34] T. Nitta, T. Ohmi, Y. Ishihara, A. Okita, T. Shibata, J. Sugiura, and N. Ohwada, J. Appl. Phys. **67**, 7404 (1990).
[35] T. Ohmi, M. Onodera, G. Sato, T. Shibata, and M. Morita, in Extended Abstracts of the Electrochemical Society Fall Meeting, Chicago, 1988, pp. 596–597.
[36] K. Tomita, T. Migita, S. Shimonishi, T. Shibata, T. Ohmi, and T. Nitta, J. Electrochem. Soc. **142**, 1692 (1995).
[37] H. Aharoni, T. Ohmi, M. M. Oka, A. Nakada, and Y. Tama (to be published).
[38] K. Ohyu, A. Hiraiwa, T. Itoga, J. Yugami, and M. Ohkurai, in Proceedings of the Symposium on ULSI Ultra Clean Technology, Tokyo, Japan, 1994, pp. 356–364.
[39] M. Horiuchi and K. Ohyu, IEEE Trans. Electron Devices **42**, 876 (1995).
[40] K. Ohyu, M. Hiraiwa, and K. Watanabe, IEEE Trans. Electron Devices **42**, 1404 (1995).
[41] J. M. Hwang and D. K. Schroder, J. Appl. Phys. **59**, 2476 (1986).
[42] J. Hwang, D. K. Schroder, and A. M. Goodman, IEEE Electron Device Lett. **EDL-7**, 172 (1986).
[43] A. Bourret, J. Thibault-Desseaux, and D. N. Seidman, J. Appl. Phys. **55**, 825 (1984).
[44] H. Holzlein, G. Pensl, M. Schulz, and N. M. Johnson, Appl. Phys. Lett. **48**, 916 (1986).
[45] H. F. Wolf, *Semiconductors* (Wiley-Interscience, New York, 1971).
[46] S. K. Ghandi, *VLSI Fabrication Principles* (Wiley, New York, 1994).
[47] G. K. Wertheim, Phys. Rev. **105**, 1730 (1957).
[48] P. V. Kuchinski and V. M. Lomako, Solid State Electron. **29**, 1041 (1986).
[49] C. Kittel, *Introduction to Solid State Physics*, 6th ed. (Wiley, New York, 1986).