Rigorous carrier dynamic model of electroluminescent metal-oxide-semiconductor silicon tunneling diodes

Eih-Zhe Liang, Ting-Wei Su, and Ching-Fuh Lin[a]
Graduate Institute of Electro-optical Engineering, National Taiwan University, Taipei 106, Taiwan, Republic of China

The carrier dynamics of electroluminescent metal-oxide-semiconductor silicon tunneling diodes is rigorously modeled in this study. Various tunneling and recombination current densities are formulated without using the Maxwell-Boltzmann approximation for the carrier concentrations. This model satisfactorily explains the current-voltage characteristics in strong accumulation using a self-consistent formulation. It also relates light emission efficiency to the interface-state density and the bulk-trap density in a straightforward manner. The internal radiative recombination efficiency in the bulk Si substrate is estimated to be in tens of percents. The model also explains clearly the strongly influenced small-signal light-current response time with respect to the injection current density. It enables the small-signal method to be useful in extracting the parameters of the interface-state density and the bulk-trap density. © *2006 American Institute of Physics.*
[DOI: 10.1063/1.2345048]

I. INTRODUCTION

Recently the study of Si-based light-emitting diodes has progressed significantly and their light emission efficiency has been improved greatly.[1–4] Among these Si electroluminescent (EL) devices, the metal-oxide-semiconductor tunneling diode (MOS-TD) has the advantage of possible integration with modern complementary metal-oxide-semicondutor integrated circuit (CMOS-IC) technology.[1,2] Much study has been devoted to raising its external light emission efficiency, either by increasing oxide roughness in the rapid thermal oxide[5] (RTO) or by employing SiO_2 nanoparticles as the oxide layer to enhance carrier confinement.[2,6] Through theoretical investigation, the light emission mechanism at the silicon band gap energy in MOS-TDs has been explained as being due to phonon-assisted exciton recombination.[7] However, the study of the influence of nonradiative recombination on EL efficiency[8,9] is either qualitative or semiempirical. In this work, a rigorous carrier dynamic model is provided to quantitatively explain the influence of interface states and bulk traps on light emission efficiency.

The theory of light emission efficiency of the MOS-TD begins with the model of tunneling current densities to the conduction band, the valence band, and the interface states.[10,11] To determine the quasi-Fermi level of the interface states, the nonradiative recombination current densities through the interface states are balanced. As the MOS-TD is operated in strong accumulation, the device behaves very differently as a light emitter from a leaky MOS junction. Therefore, unlike in Ref. 12, the surface band bending is determined by solving the Poisson equation without using the Maxwell-Boltzmann approximation of the carrier concentrations. The recombination lifetime in the bulk region cannot be treated as a constant, as the radiative recombination rate varies significantly with the spatial distribution of both majority and minority carriers. The light emission rate per unit area is calculated from a part of the diffusion current density. After all the equations of current components are formulated, the solution is determined by self-consistent iteration.

This model satisfactorily explains the current-voltage characteristics in strong accumulation; it relates light emission efficiency to the interface-state density and the bulk-trap density in a straightforward manner. The model also explains quantitatively the strongly influenced small-signal light-current response time with respect to the injection current density. It employs the frequency-response method, an important tool for extracting the parameters of interface-state density and bulk-trap density.

In strong accumulation, a high concentration of the majority carriers is present at the semiconductor surface. The surface concentration of the minority carriers is also high because the tunneling current to the minority band has a much larger magnitude than the tunneling currents to the conduction band and the interface states. The accumulation of both carriers leads to a high radiative recombination rate at the silicon surface. From the small-signal method, the internal radiative recombination efficiency in the bulk silicon substrate is estimated to be as high as tens of percents.

II. MODELING OF CURRENT COMPONENTS

The modeling of current routes in the metal-oxide-semiconductor tunneling diode has been investigated in detail without consideration of light emission.[12] Here, two important factors are emphasized in the MOS-TD as an electroluminescent device.

First, the MOS-TD is operated in strong accumulation. Its behavior is different from that of a MOS transistor, which switches between accumulation, depletion, and inversion.

---

[a]Author to whom correspondence should be addressed; also at the Graduate Institute of Electronics Engineering and Department of Electrical Engineering; electronic mail: cflin@cc.ee.ntu.edu.tw
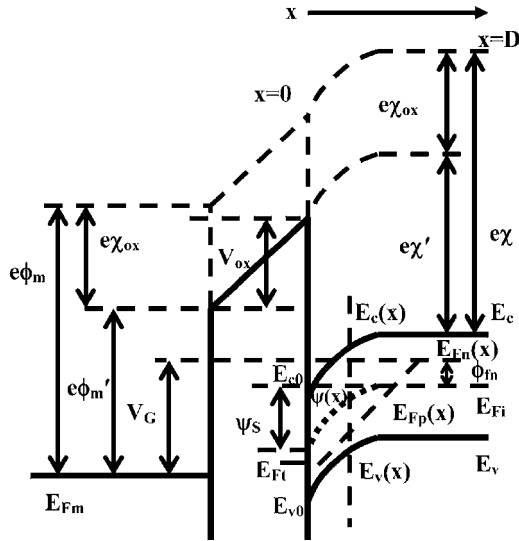
FIG. 1. Energy bands of a metal-insulator-semiconductor silicon tunneling diode.

When operated in strong accumulation, the majority carriers, for example, electrons in the n-type silicon MOS-TD, accumulate near the Si/SiO$_2$ interface and the concentration normally exceeds 10^{19} cm^{-3}. In such a case, the quasi-Fermi level of the majority carriers moves above the conduction band in the n-type MOS-TD or beneath the valence band in the p-type MOS-TD. The electron and hole concentrations, n and p, are determined by the quasi-Fermi level functions $E_{Fn}(x)$ and $E_{Fp}(x)$ in the following equations:

$$n(x) = \frac{4\pi}{h^3}(2m_n^*)^{3/2}\int_{E_c(x)}^{\infty}\sqrt{E - E_c(x)}$$
$$\times\left\{\exp\left[\frac{E - E_{Fn}(x)}{kT}\right] + 1\right\}^{-1}dE, \tag{1}$$

$$p(x) = \frac{4\pi}{h^3}(2m_p^*)^{3/2}\int_{-\infty}^{E_v(x)}\sqrt{E_v(x) - E}$$
$$\times\left\{\exp\left[\frac{E_{Fp}(x) - E}{kT}\right] + 1\right\}^{-1}dE, \tag{2}$$

where $m_n^* = 1.08m_0$ and $m_p^* = 0.56m_0$ are the effective masses of electrons and holes, respectively (m_0 is the free-electron mass).[13] In strong accumulation, using the Maxwell-Boltzmann approximation for computing the majority carrier concentration from Eqs. (1) and (2) is not appropriate. Moreover, the MOS-TD conducts the greatest tunneling current to the minority band in strong accumulation, not to the majority band. This results in a high concentration of minority carriers. Therefore, using the Maxwell-Boltzmann approximation to calculate the minority carrier concentration is not suitable either. Equations (1) and (2) are used to calculate the carrier concentrations for other equations as well. The x axis is taken as the direction from the surface to the silicon, as shown in Fig. 1. The conduction band and the valence band are written as $E_c(x)$ and $E_v(x)$ to describe the band bending. Constants h and k are the Plank constant and the Boltzmann constant, respectively, and T is the temperature.

Second, in the EL MOS-TD, the tunneling current is mainly dissipated by radiative recombination. The radiative recombination rate R varies spatially with the majority carrier concentration n and the minority carrier concentration p. Therefore, the recombination coefficients such as τ_n or τ_p are not constant. This is conventionally written as

$$R = Bnp + A(n + p), \tag{3}$$

where B is the radiative recombination coefficient (or radiative band-to-band coefficient) and A is the nonradiative recombination coefficient. Because the Maxwell-Boltzmann approximation of the occupation probability of electrons f_n and holes f_p is not appropriate for use in strong accumulation, the product Bnp in Eq. (3) is replaced by[14]

$$Bnp = \int \beta_0(\omega)d\omega\left[\frac{4\pi(2m_n^*)^{3/2}}{h^3}\right]\left[\frac{4\pi(2m_p^*)^{3/2}}{h^3}\right]$$
$$\times\int_{E_c - (\hbar\omega + \hbar\Omega)}^{E_v}\sqrt{E - [E_c - (\hbar\omega + \hbar\Omega)]}\sqrt{E_v - E}$$
$$\times f_p(E)f_n[E + (\hbar\omega + \hbar\Omega)]dE. \tag{4}$$

This formula is derived in order to model the emission spectrum of phonon-assisted radiative recombination of electrons and holes in bulk silicon with TO/LO phonon energy $\hbar\Omega$ without the exciton effect mentioned in Ref. 7. The exciton effect simply results in a redshift of the spectrum and increases the B coefficient for a higher density of states of excitons. It does not change the dependence of R on E_{Fn} and E_{Fp}. The spectral shape function $\beta_0(\omega)$ is proportional to the absorption coefficient in bulk silicon.

A schematic diagram of the n-type silicon MOS-TD operated in strong accumulation is shown in Fig. 1. A p-type MOS-TD is similar, but the applied voltage V_G has opposite polarity. V_{ox} is the voltage drop across the oxide barrier, ψ_S is the voltage drop across the silicon substrate (or band bending), ϕ_m is the work function of the metal, χ is the electron affinity of silicon, and χ_{ox} is the electron affinity of the oxide. The band bending at the back electrode is $\psi(D) = 0$, and at the Si/SiO$_2$ interface it is defined as $\psi(0) = \psi_S$. The conduction band and the valence band energies inside the substrate are given as E_c and E_v, respectively. The two energies vary across the substrate and are given as $E_c(x) = E_c - \psi(x)$ and $E_v(x) = E_v - \psi(x)$, respectively. At the Si/SiO$_2$ interface, they are written as E_{c_0} and E_{v_0}. The quasi-Fermi level of the interface states is E_{Ft}. The applied voltage can be obtained from the following equations:

$$V_G = \phi_{ms} + V_{ox} + \psi_S, \tag{5}$$

$$\phi_{ms} = \phi_{m\prime} - \chi\prime - \left(\frac{E_g}{2e} - \phi_{Fn}\right), \tag{6}$$

where $\phi_{Fn} = [E_{Fn}(D) - E_v - E_g/2]$.

The voltage drop across the oxide is given by

$$V_{ox} = \frac{t_{ox}}{\varepsilon_{ox}}[Q_S(\psi_S) + Q_{it} + Q_X], \tag{7}$$

where t_{ox} is the oxide thickness and ε_{ox} is the corresponding dielectric constant. The surface charge density of the semi-
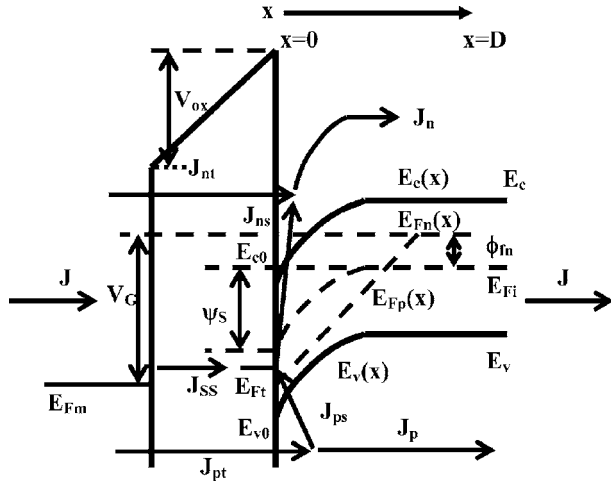
FIG. 2. Current density components of a metal-insulator-semiconductor tunneling diode in strong accumulation.

conductor, Q_S (C cm^{-2}), is given by $Q_S(\psi_S) = \varepsilon_S(d\psi/dx)|_{x=0}$. The fixed charge density in the oxide, Q_X (C cm^{-2}), is related to the fixed charge density N_X (cm^{-2}) by $Q_X = eN_X$. The interface charge density Q_{it} (C cm^{-2}) is written as $Q_{it} = eD_{it}(E_{Fn} - E_{Ft})$, where the interface-state density function D_{it} (cm^{-2} eV^{-1}) is assumed to be uniform across the silicon band gap.

To determine the surface band bending ψ_S, the Poisson equation is solved numerically:[15]

$$\frac{\partial^2}{\partial x^2}\psi(x) = -\frac{e}{\varepsilon_S}(p - n + N_D), \tag{8}$$

where N_D is the doping concentration of the silicon substrate and ε_S is the dielectric constant of silicon.

Various current density components in a n-type MOS-TD are shown in Fig. 2. J is the total current density flowing from the metal electrode to the silicon back electrode, J_{nt} is the tunneling current density to the conduction band, J_{pt} is the tunneling current density to the valence band, J_{SS} is the tunneling current density to Si/SiO$_2$ interface states, J_{ns} is the recombination current density from the conduction band to the interface states, J_{ps} is the recombination current density from the valence band to the interface states, J_n is the drift current density of electrons, and J_p is the diffusion current density of holes.

The node equations include Eq. (9) at the metal electrode/back electrode, Eq. (10) at the interface states, Eq. (11) at the conduction band, and Eq. (12) at the valence band:

$$J = J_{nt} + J_{pt} + J_{SS} = J_n + J_p, \tag{9}$$

$$J_{SS} = J_{ns} - J_{ps}, \tag{10}$$

$$J_{nt} = J_n - J_{ns}, \tag{11}$$

$$J_{pt} = J_p + J_{ps}. \tag{12}$$

On combining Eqs. (9)–(12), Eq. (13) is obtained, which is used to calculate the total current density J at different applied voltages V_g from J_{nt}, J_{SS}, J_{ps}, and J_p,

$$J = J_{nt} + J_{SS} + J_{ps} + J_p. \tag{13}$$

The tunneling current densities from the metal electrode to the conduction band and to the valence band, J_{nt} and J_{pt}, respectively, are given in Ref. 10 and rewritten here as Eqs. (14) and (15), without employing the Maxwell-Boltzmann approximation of the occupation probability. The energy difference of each term in Eqs. (14) and (15) is $E_c(0) - E_{Fn}(0) = E_c + e\psi_S - E_{Fn}$, $E_c(0) - E_{Fm} = eV_{ox} - e(\phi_m' - \chi')$, $E_{Fm} - E_v(0) = E_g - eV_{ox} + e(\phi_m' - \phi_s')$, and $E_{Fp}(0) - E_v(0) = E_{Fp}(0) - (E_v + e\psi_S)$, respectively. When the n-type MOS-TD is operated in strong accumulation, the quasi-Fermi level of electrons at the silicon surface, $E_{Fn}(0)$, is normally above the metal Fermi level E_{Fm}, as the voltage drop across the oxide is positive. Therefore the electron tunneling current is from the metal electrode to the conduction band. The quasi-Fermi level of holes E_{Fp} is also well above the metal Fermi level E_{Fm} even if it is lower than E_{Fn}. Thus the tunneling current of holes flows from the metal electrode to the valence band. The tunneling barrier height of electrons χ_n and holes χ_p is assumed to be 0.7 eV.[12] The unit used for oxide thickness t_{ox} is angstrom, for brief formulation.[10] The transverse effective mass m_t is taken to be the same as m_n^* and m_p^* for the corresponding current density:

$$J_{nt} = \frac{em_t}{\pi\hbar^3}\exp(-\sqrt{\chi_n}t_{ox})\int_0^\infty \left(\left\{\exp\left[\frac{E + E_c(0) - E_{Fn}(0)}{kT}\right]\right.\right.$$
$$\left.+ 1\right\}^{-1} - \left\{\exp\left[\frac{E + E_c(0) - E_{Fm}}{kT}\right] + 1\right\}^{-1}\right)EdE, \tag{14}$$

$$J_{pt} = \frac{em_t}{\pi\hbar^3}\exp(-\sqrt{\chi_p}t_{ox})\int_0^\infty \left(\left\{\exp\left[\frac{E_{Fm} - E_v(0) + E}{kT}\right]\right.\right.$$
$$\left.+ 1\right\}^{-1} - \left\{\exp\left[\frac{E_{Fp}(0) - E_v(0) + E}{kT}\right] + 1\right\}^{-1}\right)EdE. \tag{15}$$

To treat the current densities related to the interface states[11] without an approximation of the occupation probability, J_{SS}, J_{ns}, and J_{ps} are derived as shown below:

$$J_{SS} = \int_{E_v}^{E_c}\frac{eD_{it}}{\tau_t}(f_t - f_m)dE$$
$$= \frac{eN_{it}}{\tau_tE_g}\int_0^{Eg}\left(\left\{\exp\left[\frac{E + E_v(0) - E_{Ft}}{kT}\right] + 1\right\}^{-1}\right.$$
$$\left. - \left\{\exp\left[\frac{E + E_v(0) - E_{Fm}}{kT}\right] + 1\right\}^{-1}\right)dE, \tag{16}$$

$$J_{ns} = eN_{it}v_{th}\sigma_n[(1 - f_t)n_S - f_tn_1], \tag{17}$$

$$J_{ps} = eN_{it}v_{th}\sigma_p[f_tp_S - (1 - f_t)p_1], \tag{18}$$

where v_{th} is the thermal velocity, σ_n and σ_p are the capture cross sections of electrons and holes by the interface states, and τ_t is the tunneling time constant of electrons from the metal electrode to the interface states. $n_S = n(x=0)$ and p_S

$=p(x=0)$ are the surface concentrations of electrons and holes, respectively. n_1 and p_1 are expressed as

$$n_1 = \frac{4\pi}{h^3}(2m_n^*)^{3/2} \int_{E_{c0}}^{\infty} \sqrt{E - E_{c0}} \left[\exp\left(\frac{E - E_{Ft}}{kT} \right) + 1 \right]^{-1} dE,$$

(19)

$$p_1 = \frac{4\pi}{h^3}(2m_p^*)^{3/2} \int_{-\infty}^{E_{v0}} \sqrt{E_{v0} - E} \left[\exp\left(\frac{E_{Ft} - E}{kT} \right) + 1 \right]^{-1} dE.$$

(20)

The interface-state density function D_{it} (cm^{-2} eV^{-1}) is assumed to be uniformly distributed over the band gap, so the total interface-state density is given by $N_{it} = D_{it}E_g$ (cm^{-2}). The occupation probability f_t can be determined from Eqs. (10) and (16)–(18). Then E_{Ft} is retrieved from

$$f_t = E_g^{-1} \int_0^{E_g} \left[\exp(E + E_v - E_{Ft}) + 1 \right]^{-1} dE.$$

(21)

The device is in quasiequilibrium, so the majority carrier concentration $n(x)$ and the minority carrier concentration $p(x)$ are determined by the quasi-Fermi levels in Eqs. (1) and (2). The value of $E_{Fn(x)}$ can be determined from the majority drift current density by $J_n = \mu_n n^{(dE_{Fn}/dx)}$. A small amount of variation in the quasi-Fermi level of the majority carriers can result in a large current flowing through the bulk region due to the large majority carrier concentration in strong accumulation.[10] To a good approximation, it can be assumed that E_{Fn} remains constant throughout the bulk region and the majority carrier concentration $n(x)$ does not differ greatly from the value in thermal equilibrium $\bar{n}(x)$.

The minority carrier concentration $p(x) = \bar{p}(x) + \Delta p(x)$ as well as $E_{Fp}(x)$ can be determined from the ambipolar carrier transport equation, or the diffusion equation

$$D_p \frac{d^2}{dx^2} p + \bar{g} - R = \frac{dp}{dt} = 0,$$

(22)

where D_p is the diffusion coefficient in the n-type silicon substrate. The hole concentration $p(x)$ is many orders larger than its thermal equilibrium value $\bar{p}(x)$ when the diffusion current flows to the accumulation region. The generation rate per volume density in thermal equilibrium $\bar{g}$ is given by

$$\bar{g} = B\bar{n}\bar{p} + A(\bar{n} + \bar{p}).$$

(23)

The product $B\bar{n}\bar{p}$ still gives a good approximation to the radiative recombination rate in thermal equilibrium in a MOS junction without the tunneling current. $A(\bar{n} + \bar{p})$ is the nonradiative recombination rate per unit volume by bulk traps and is expressed as[11]

$$A(\bar{n} + \bar{p}) = \frac{N_{bt}v_{th}(\sigma_n \bar{n} + \sigma_p \bar{p})}{\tau_{bt}},$$

(24)

where N_{bt} is the bulk-trap density (cm^{-2}) and τ_{bt} is the bulk-trap time constant. The recombination rate R in quasiequilibrium is expressed as Eq. (3). Through combining Eqs. (3) and (22)–(24), the diffusion equation (22) becomes Eq. (25).

The diffusion current density of the minority carrier J_p is defined as

$$D_p \frac{d^2}{dx^2}(\Delta p) - [B\bar{n}(\Delta p) + A(\Delta p)] = 0,$$

(25)

$$J_p = -eD_p \left. \frac{d}{dx}(\Delta p) \right|_{x=0} = -e \int_0^H [B\bar{n}(\Delta p) + A(\Delta p)]dx.$$

(26)

J_p in Eq. (26) can be written as the integral form if the radiative recombination-generation current density is defined as $J_{rg} = e\int_0^x B\bar{n}(\Delta p)dx$ and the nonradiative recombination current density from bulk traps is defined as $J_{bt} = e\int_0^x A(\Delta p)dx$. The diffusion equation (26) is then readily written as Eq. (27). The current densities J_p, J_{rg}, and J_{bt} are used to determine the internal carrier recombination efficiency and the carrier recombination lifetime,

$$J_p = J_{rg} + J_{bt} = e\int_0^H B\bar{n}(\Delta p)dx + e\int_0^H A(\Delta p)dx.$$

(27)

To obtain the relation of the total current density through the device J to the given applied voltage V_G, the quasi-Fermi level of the majority carriers E_{Fn}, for example, electrons in a n-type MOS-TD, is assumed to be constant within silicon. The variation of the electron concentration is calculated by the band bending function $\psi(x)$. To reach a consistent solution, the algorithm is described in the following steps:

(a)   Quasi-Fermi level of holes is the same as that of electrons $E_{Fp}(x) = E_{Fn}(x)$ and $\Delta p(x) = 0$.
(b)   Surface band bending ψ_s is obtained by solving the Poisson equation, Eq. (8).
(c)   Balance equation for J_{ns}, J_{ps}, and J_{SS} in Eq. (10) is solved to give E_{Ft}.
(d)   Diffusion equation for holes in Eq. (27) is solved to give Δp and J_p.
(e)   If $|J_{pt} - J_{ps} - J_p| <$ tolerance, then J is given by Eq. (9). Otherwise it repeats from (b) to (d), with $p(x) = \bar{p}(x) + \Delta p(x)$.

Figure 3 shows both the experimental J-V curve from a fabricated n-type MOS-TD device (solid line) and the theoretical curves (dotted line) obtained from the above algorithm. The total current density J is the sum of the tunneling current density to the majority band J_{nt} in Eq. (14), to the minority band J_{pt} in Eq. (15), and to the interface states J_{SS} in Eq. (16). The kink of the J-V curve at the applied voltage of 1.0 V indicates that the tunneling current density to the valence band J_{pt} raises orders of magnitudes larger than the current density to the conduction band J_{nt} and the interface states J_{SS}. Therefore, J_{pt} is the major part of the total tunneling current density in a n-type MOS-TD operated in strong accumulation. The same situation happens to the p-type MOS-TD in forward bias, where J_{nt} is the major part of the total tunneling current density.[11] However, this does not mean that most tunneling current density becomes the minority diffusion current density J_p. Only part of J_{pt} becomes the diffusion current density and contributes to radiative recom-
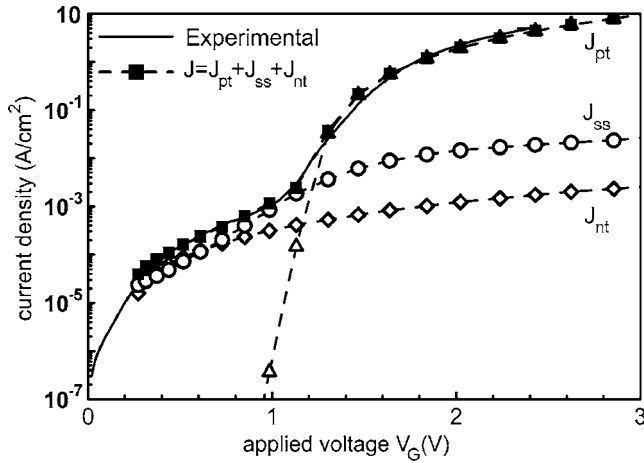
FIG. 3. Comparison of the tunneling current density vs voltage relation of a *n*-type MOS-TD (solid line) and the fitted simulation curves of the tunneling current density (t_{ox}=1.5 nm). The total current density J (square dotted) is the sum of tunneling current densities to the minority band J_{pt} (triangle dotted), to the majority band J_{nt} (diamond dotted), and to the interface states J_{SS} (circle dotted).

bination. Another part becomes the recombination current density J_{ps} through the interface states. This is because f_t in Eq. (18) is pinned near the conduction band edge for a very low f_m in strong accumulation.

The surface band bending and the oxide voltage drop with respect to different applied voltages are shown in Fig. 4. The interpretation is as follows: at an applied voltage of higher than 0.2 V in the *n*-type MOS-TD, the voltage almost falls on the oxide as shown, so V_{ox} changes linearly with V_G. The voltage drop across the region of the semiconductor band bending is only a few hundred meV. This voltage drop cannot increase much more. When the quasi-Fermi level of electrons is above the conduction-band edge, the band bending and the electron concentration do not increase significantly with applied voltage.

The surface concentrations of electrons are depicted in Fig. 5. The value of the surface concentration of electrons n_s reaches 2×10^{19} cm^{-3} when the quasi-Fermi level of electrons E_{Fn} is lifted above E_c, for the applied voltage of less than 1 V. The value of the surface concentration of holes p_s also approaches 7×10^{18} cm^{-3}, when the quasi-Fermi level
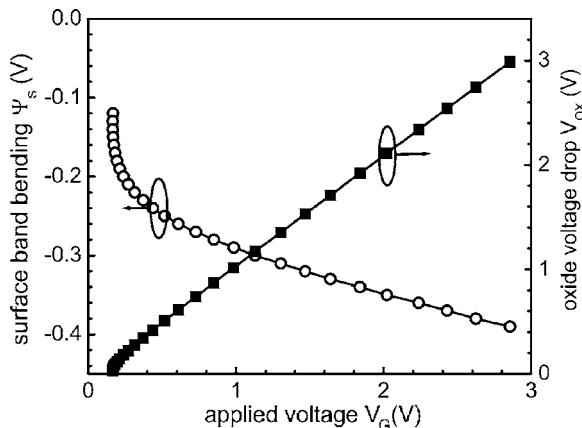

FIG. 4. Simulation of the surface band bending and the oxide voltage drop of a *n*-type MOS-TD (t_{ox}=1.5 nm).
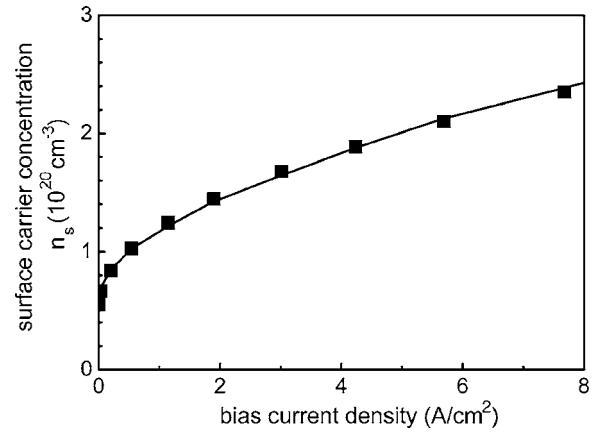

FIG. 5. Simulation of the surface carrier concentration of electrons in a *n*-type MOS-TD with respect to different current densities in the small-signal light-current response.

of holes E_{Fp} reaches E_v and the applied voltage exceeds 2 V. The minority tunneling current density J_{pt} has a larger magnitude than the other two tunneling current densities, J_{nt} and J_{SS}, as shown in Fig. 3. To dissipate such a large current, the surface concentration of holes is raised. In such a case, the strong accumulation of both electrons and holes leads to an intense radiative recombination rate at the semiconductor surface.

III. LIGHT CONVERSION EFFICIENCY AND ITS IMPROVEMENT

After the carrier concentrations and all current densities are formulated at the given applied voltage, light emission efficiency is then derived from the model in Sec. II. The externally detected photon emission flux $L/\hbar\omega$ (cm^{-2}) of the MOS-TD operated as a light emitter is expressed as follows:

$$L/\hbar\omega = \eta_c(J_{rg}S_A/e) = \eta_c\eta_i(JS_A/e), \quad (28)$$

where J_{rg} is the recombination current density, η_c is the light extraction efficiency from a MOS-TD light emitter to a photodetector, η_i is the internal quantum efficiency, and S_A is the total area of injection current. The external quantum efficiency is expressed as

$$\eta_{ext} = \left(\frac{e}{\hbar\omega}\right)\left(\frac{L}{JS_A}\right) = \eta_c\eta_i. \quad (29)$$

The injection efficiency η_{inj} can be obtained from Eq. (13) and written as

$$\eta_{inj} = \frac{J_p}{J_{nt} + J_{SS} + J_{ps} + J_p}. \quad (30)$$

The internal bulk recombination efficiency η_{rec} can be obtained from Eq. (27) and explicitly expressed as

$$\eta_{\mathrm{rec}} = \frac{J_{\mathrm{rg}}}{J_{\mathrm{rg}} + J_{\mathrm{bt}}} = \frac{e \int_0^H B \bar{n}(\Delta p) dx}{e \int_0^H B \bar{n}(\Delta p) dx + e \int_0^H A(\Delta p) dx}. \quad (31)$$

Combining Eqs. (30) and (31), the internal quantum efficiency becomes $\eta_i = \eta_{\mathrm{rec}} \eta_{\mathrm{inj}}$.

To improve external quantum efficiency, one may increase the light extraction efficiency η_c by applying an anti-reflection pattern to the Si surface,[4] increase the injection efficiency by reducing the interface states, or increase the recombination efficiency by reducing the bulk traps. In the modern IC industry, the silicon wafer has a high crystal quality that results in a carrier lifetime as high as several milliseconds. However, the interface-state density can only be reduced to the order of 10^{10} cm^{-2}. The tunneling current density to the valence band J_{pt} consists of the recombination current density to the interface state J_{ps} and the diffusion current density J_p. The portion of the diffusion current density contributed by J_{pt} is simulated to be several tens of percents according to different interface-state densities. Therefore, the internal recombination efficiency η_i in bulk silicon is estimated to be a few tens of percents. This estimation can be also deduced from the small-signal light-current response and *is* discussed in Sec. IV. The light extraction efficiency, however, can be very low if the generated light is blocked by the thick metal pad. The external quantum efficiency from the periphery of the metal pad is around 10^{-4} in the SiO$_2$ nanoparticle-modified MOS-TD.[2]

The arguments about injection efficiency also predict another possibility of raising the external quantum efficiency. If the radiative recombination coefficient B is raised by the benefit of a carrier confinement structure, the diffusion current density will be enhanced due to the fast dissipation of the injected current. The injection efficiency can then be further increased without facing the difficulty of reducing the interface-state density.

IV. SMALL-SIGNAL LIGHT-CURRENT RESPONSE

To further broaden the carrier dynamic model, the small-signal response of light emission by a modulation current is investigated. A strong influence of the injection current density on the response time, or the carrier lifetime, is found in the MOS-TD. This effect exists in the rapid-thermal-oxide MOS-TD at a very high injection current density and is measurable at a lower injection current density in the SiO$_2$ nanoparticle-modified MOS-TD.[16] In discussions of this effect, the frequency response of the current with respect to the voltage modulation in a MOS junction, or the small-signal capacitance-voltage (*CV*) method, is reviewed. It is a common technique used to characterize the influence of the interface states on electrical performance, and has been extended to the MOS tunneling diode, which involves the tunneling current to the conduction band, the valence band, and the interface states.[11] The response time of the current by small-signal voltage modulation is limited by the majority tunneling current density J_{nt}, the tunneling current density to interface states J_{SS}, and the minority tunneling current density J_{pt}. The time constant τ_{CV} can be expressed as in Eq. (32), where G_{CB} and G_{VB} are the tunneling conductances of the tunneling currents to the conduction band and the valence band, respectively, G_p and C_p are the conductance and capacitance of the recombination current to interface states, respectively, G_T is the conductance of the tunneling current to interface states, C_{SC} is the capacitance due to the surface charge, and C_{ox} is the oxide capacitance. The typical value of τ_{CV} is lower than 1 μs in the MOS-TD with an active area of 0.01 cm^2,

$$\tau_{CV}^{-1} = 2\pi[C_{\mathrm{ox}}/(G_T + G_{\mathrm{CB}} + G_{\mathrm{VB}}) + (C_{\mathrm{ox}} + C_{\mathrm{SC}} + C_P)/G_p]. \quad (32)$$

The small-signal response of light with respect to the modulation current is considered here because the MOS-TD is now operating as a light emitter. The small-signal light-current response is normally around $10-100$ μs, slower than τ_{CV}. From Eqs. (30) and (31), the modulation speed of light generation with respect to the different current densities is mainly controlled by two factors: the nonradiative recombination rate through the interface states and the recombination rate in the bulk region. Therefore, the response of light emission to the small-signal current modulation, i.e., the ratio of the modulated light amplitude to the modulated current amplitude, is expressed as

$$\frac{L(i\omega)}{JS_A} = \frac{\hbar\omega}{e}\mathrm{EQE}\frac{1}{1 + i\omega\tau_L}, \quad (33)$$

where the response time τ_L is expressed as $\tau_L^{-1} = \tau_0^{-1} + \tau_{\mathrm{rec}}^{-1}$. EQE is the external quantum efficiency under DC bias. The response time can be derived from the light-current amplitude-frequency plot by the frequency response measurement and is explained later. The parameter τ_{rec} is the response time of the diffusion equation (27). The parameter $\tau_0 = (2\pi C_p/G_p)^{-1}$ is the response time of the recombination current density through interface states J_{SS} and J_{ps}. Capacitance C_p and conductance G_p are frequency dependent; they vary significantly with the surface concentration of the majority carriers when the applied frequency is approaching the inverse of interface recombination time $\tau_s \approx (n_s v_{\mathrm{th}} \sigma_n)^{-1}$.[12] In the small-signal light-current response of the MOS-TD, because the operation frequency of the modulation current is much less than τ_s, τ_0 can be assumed to be independent of the total injection current density J.

In the small-signal capacitance-voltage response, the interface-state conductance (G_p) of a MOS-TD, even acting as a light emitter, is less influential than the direct tunneling conductance in accumulation.[17] However, in the small-signal light-current response, its contribution to the response time of light generation is not negligible. This is because the response time of light generation is mainly limited by both radiative and nonradiative recombination rates, not the current injection rate. The response time is a few tens of microseconds and the nonradiative recombination time τ_0, involving G_p, is a few hundreds of microseconds. In such a case, to better estimate the radiative recombination lifetime, τ_0 must be included. Therefore, G_p is considered here.

The variation of light emission response time with the different bias current densities is mostly related to the recombination time τ_{rec} in the bulk region. The response time τ_L is explicitly expressed in the following equation:

$$\tau_L^{-1} = \tau_0^{-1} + \tau_{rec}^{-1} = \tau_0^{-1} + \frac{\int_0^H [(B\bar{n} + A)\delta\Delta p]dx}{\int_0^H \delta\Delta p dx}$$

$$\approx \tau_0^{-1} + (B\bar{n} + A)\frac{L_n}{L_p}. \tag{34}$$

The parameter L_n is the accumulation depth of electrons. It is defined as the distance from the surface to the point where the electron concentration reduces to 1%. Its value is around 0.3 nm when the injection current is several A/cm². The parameter L_p is the accumulation depth of holes. It is defined as the distance from the surface to the point where the hole concentration reduces to 1%. Its value is around 5 nm, according to the solution of Eq. (27). The product L_n/L_p is proportional to the concentration of the excess minority carriers Δp, as the higher Δp results in higher recombination rate in the bulk region when the device is operated in strong accumulation.

The internal efficiency $\eta_i = \eta_{rec}\eta_{inj}$ can also be analyzed from Eqs. (30), (31), and (34). It is expressed as $\eta_i = \tau_{rec}^{-1}/(\tau_0^{-1} + \tau_{rec}^{-1})$. To estimate the value of the nonradiative recombination coefficient A, the optically pumped carrier lifetime is measured by laser-excitation microwave photoconductivity decay (MW-PCD) of the silicon wafer in 0.8 M iodine/ethanol.[18] The laser excitation creates a lower injection level, compared with electrical injection. The excess carrier concentration n_{PCD} is in the range of 10^{16}–10^{17} cm⁻³. The float-zone growth Si wafer used has a MW-PCD lifetime of 1–3 ms. The internal photoluminescent (PL) efficiency is estimated to be 10% with the consideration of a lifetime of 30 ms in the purest wafers.[19] In such a case, the internal efficiency is modified to

$$\eta_{PCD} \approx \frac{mB\bar{n}_s}{mB\bar{n}_s + A}, \tag{35}$$

where m is the ratio of n_{PCD} to the surface concentration of electrons n_s in a normally operated MOS-TD. Supposed the laser-pumped excess electron concentration is 10^{16} cm⁻³, then the value of A is estimated to be 270 s⁻¹ in a sample with a 3 ms MW-PCD lifetime. This means that in the normal operation of a MOS-TD, the internal EL efficiency in bulk silicon is nearly 100%. In comparison, the light emission efficiency is only several percent below the usual PL or MW-PCD measurement. The high internal EL efficiency is due to the large carrier concentration at the surface when the MOS-TD is operated in strong accumulation.

Frequency response is used to measure the light-generation response time, as the response of the output light in Eq. (33) and the response time in Eq. (34) give information on the radiative and nonradiative recombination rates. This method shows how the carrier lifetime changes with different bias current densities. Its principle is to modulate
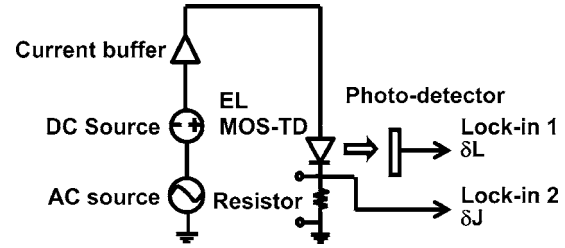

FIG. 6. The measurement setup of the small-signal light-current response of EL MOS-TDs.

the biased light-emitting MOS-TD with a small-signal sinusoidal current. The measurement setup of frequency response is illustrated in Fig. 6. The modulation source (Agilent 33210A) is fed to the device under a dc bias. The dc bias is provided by Keithley 2400. After the voltage excitation is applied to the EL MOS-TD, the current is measured through a 1 Ω resistance. The modulation amplitude of the current density δJ is maintained at 0.5 A/cm² over all bias current densities. The modulated light amplitude δL is detected by an InGaAs detector with a response time smaller than 100 ns and extracted from the output light through a lock-in amplifier (Stanford Research Systems SR830). The modulated current density δJ is detected by another lock-in amplifier.

Figure 7 shows how the modulated output light amplitude decreases as the modulation frequency increases in a n-type MOS-TD. The response of light, i.e., the ratio of the modulated light amplitude δL to the modulated current density amplitude δJ, is proportional to $1/\sqrt{1 + (2\pi f \tau_L)^2}$ according to Eq. (33), where f is the modulation frequency and τ_L is the response time. The response time τ_L can be derived from the curve fitting of the response of light in Fig. 7.

The tunneling current density significantly affects the majority carrier density as well as the carrier lifetime. In the previous measurement of temporal response,[8,9] the input modulation is a current pulse that varies from zero to a large value. As the carrier lifetime strongly varies with the different current levels, the dependence of lifetime on the current density might not be simply extracted. In the frequency response measurement here, the modulation current density is only a few percent of the bias current density. Therefore the
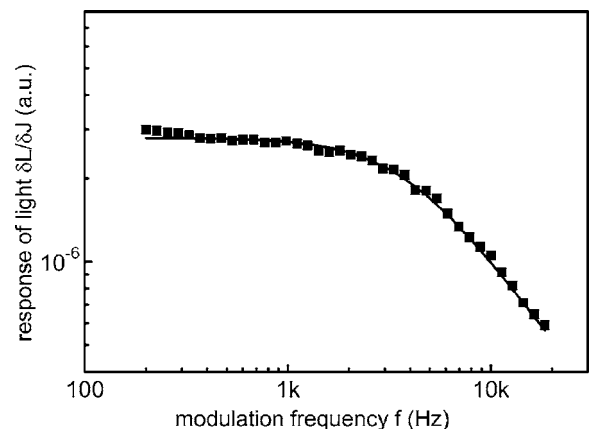

FIG. 7. The response of light, i.e., the ratio of the modulated light amplitude δL to the modulated current density amplitude δJ, of a n-type silicon MOS-TD at the bias current density of 5 A/cm².
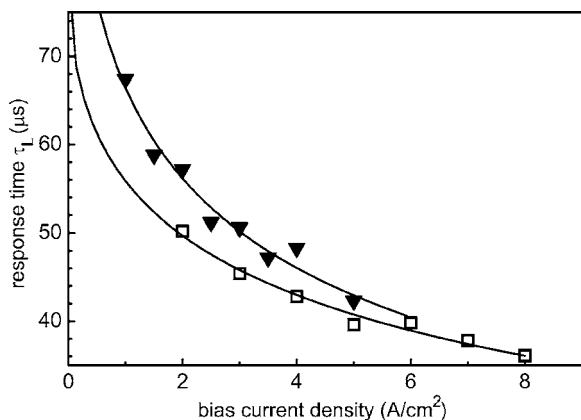
FIG. 8. The small-signal light-current response times with respect to different bias current densities (triangle, device A; square, device B) and the simulation curves (solid).

modulation has less effect on the lifetime, i.e., the response time. As a result, this method shows more clearly how the response time is related to the bias current density as well as the majority carrier density.

The minority carrier lifetime is influenced by the radiative recombination of electrons and holes and nonradiative recombination through bulk traps and interface states. With a high level current density larger than 1 A/cm^2, it is operated in strong accumulation and the quasi-Fermi level E_{Fn} is above the conduction band edge. A further injection of electrons increases the surface concentration of electrons slowly by raising E_{Fn}. From Fig. 5, the surface concentration of electrons n_s can be approached by the relation $n_s = 0.55 + 0.65J^{1/2}$, where n_s is in units of 10^{20} cm^{-3} and J is in units of A/cm^2. Moreover, at this level of current density, the nonradiative recombination rate through interface states does not increase further, as the quasi-Fermi level of the interface states remains the same in strong accumulation.

Figure 8 shows the relation between response time and the bias current density of two n-type MOS-TDs. Devices A and B are both SiO_2 nanoparticle-modified MOS-TDs, fabricated by spin coating a 120 nm thick layer of oxide nanoparticles onto Si. The particle size is 8–10 nm. The wafer is treated with KOH solution prior to the spin coating of oxide nanoparticles to remove the polish damage.[2] Both devices have a metal pad area of 0.01 cm^2. The external light emission efficiency η_{ext} is 10^{-5} and 5×10^{-6} (W/A) for A and B, respectively. To fit the response time curve in Fig. 8, the radiative recombination coefficient is needed. In EL MOS-TDs, the physical mechanism of light emission is phonon-assisted radiative recombination of electrons and holes in bulk silicon. This recombination mechanism is the same as the photoluminescence in bulk silicon in Ref. 20, so the reported value of $B = 2 \times 10^{-15}$ cm^3/s can be used. The data fitted are $t_0 = 364$ μs, $Ln/Lp = 0.06$ for device A and $t_0 = 124$ μs, $Ln/Lp = 0.04$ for device B. The higher value of t_0 in device A corresponds to a lower interface-state density N_{it}, so the light emission efficiency of device A is higher than that of device B. When the device has a lower interface-state density, it also has a larger surface concentration of holes, and thus a higher Ln/Lp value, according to Eq. (34). The

theoretical internal efficiency (η_i) can be calculated to be 90% and 72% for devices A and B, respectively. Other factors such as the leak current through oxide pores may reduce the injection efficiency and is not measurable in the small-signal light-current response.

V. CONCLUSION

This work derives in detail the carrier dynamic model of metal-oxide-semiconductor silicon tunneling diodes. The model satisfactorily explains the current-voltage characteristics, light emission efficiency, and the dependence of the small-signal light-current response time on the bias current density. A large majority carrier concentration is accumulated at the semiconductor surface in strong accumulation. The tunneling current density to the minority band is found to be much larger than the tunneling current densities to the majority band and the interface states. Part of the minority tunneling current density becomes the diffusion current and most of the diffusion current produces light by radiative recombination in the bulk region. Several percent of the minority tunneling current density recombine nonradiatively through the interface states. The internal efficiency is mainly determined by the interface-state density and also influenced by the bulk-trap density.

Information on interface states can be retrieved from the small-signal light-current plot versus the injection current density. The result is explained by the relation of the surface carrier concentration with respect to the injection current density. The device with a higher initial response time has lower interface states and thus higher light emission efficiency. The relation of the surface majority carrier concentration with the injection current density is modeled. It is used to explain the small-signal light-current response. The theoretical internal efficiency is estimated to be several tens of percents from the extracted interface-state response time.

[1] C. W. Liu, M. H. Lee, M. J. Chen, I. C. Lin, and C. F. Lin, Appl. Phys. Lett. **72**, 1516 (2000).
[2] C. F. Lin, P. F. Chung, M. J. Chen, and W. F. Su, Opt. Lett. **27**, 713 (2002).
[3] W. L. Ng, M. A. Lourenco, R. M. Gwilliam, S. Ledain, G. Shao, and K. P. Homewood, Nature (London) **410**, 192 (2001).
[4] M. A. Green, J. Zhao, A. Wang, P. J. Reece, and M. Gal, Nature (London) **412**, 805 (2001).
[5] M. H. Lee, K. F. Chen, C. C. Lai, C. W. Liu, W. W. Pai, M. J. Chen, and C. F. Lin, Jpn. J. Appl. Phys., Part 2 **41**, L326 (2002).
[6] C. F. Lin, T. W. Su, P. F. Chung, E. Z. Liang, M. J. Chen, and C. W. Liu, Mater. Chem. Phys. **77**, 430 (2002).
[7] M. J. Chen, E. Z. Liang, S. W. Chang, and C. F. Lin, J. Appl. Phys. **90**, 789 (2001).
[8] M. J. Chen, C. F. Lin, M. H. Lee, and S. T. Chang, Appl. Phys. Lett. **79**, 2264 (2001).
[9] M. J. Chen, J. F. Chang, J. L. Yen, C. S. Tsai, E. Z. Liang, C. F. Lin, and C. W. Liu, J. Appl. Phys. **93**, 4253 (2003).
[10] H. C. Card and E. H. Rhoderick, J. Phys. D **4**, 1602 (1971).

[11]L. R. Freeman and W. E. Dahlke, Solid-State Electron. **13**, 1483 (1970).

[12]M. Y. Doghish and F. D. Ho, IEEE Trans. Electron Devices **39**, 2771 (1992).

[13]C. D. Thurmond, J. Electrochem. Soc. **122**, 1133 (1975).

[14]V. Alex, S. Finkbeiner, and J. Weber, J. Appl. Phys. **79**, 6943 (1996).

[15]J. Sune, P. Olivo, and B. Ricco, J. Appl. Phys. **70**, 337 (1991).

[16]C. F. Lin, T. W. Su, E. Z. Liang, H. H. Hsieh, and W. P. Huang, Proc. SPIE **4996**, 4996–07 (2003).

[17]E. Nicollian, A. Goetzberger, and A. Lopez, Solid-State Electron. **12**, 937 (1969).

[18]E. Yablonovitch, D. L. Allara, C. C. Chang, T. Gmitter, and T. B. Bright, Phys. Rev. Lett. **57**, 249 (1986).

[19]A. W. Stephens and M. A. Green, Sol. Energy Mater. Sol. Cells **45**, 255 (1997).

[20]G. Augustine and A. Rohatgi, IEEE Trans. Electron Devices **39**, 2395 (1992).