Resonant tunneling versus thermally activated transport through strained $Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x$ quantum wells

Julia A. Berashevich* and Viktor E. Borisenko
Laboratory of Nanoelectronics and Novel Materials, Belarusian State University of Informatics and Radioelectronics, P. Browka 6, Minsk, 220013 Belarus

Jean-Louis Lazzari† and François Arnaud D'Avitaya
Centre de Recherche en Matière Condensée et Nanosciences, CRMC-N, UPR-CNRS 7251,‡ Campus de Luminy, Case 913, 13288 Marseille cedex 9, France

(Received 17 October 2006; revised manuscript received 22 January 2007; published 30 March 2007)

Electron transport through strained $Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x$ quantum well embedded in relaxed n-$Si_{1-y}Ge_y$/strained Si emitter and collector was analyzed and numerically simulated taking into account the two main processes that are resonant tunneling and thermally activated transfer through the barriers. These processes were modeled with a system of Schrödinger and kinetic equations resolved self-consistently with the Poisson equation. Within the optimum domain of composition ($0.09<y<0.25$, $0.56<x<0.83$) and thickness providing defect free strained Si and $Si_{1-x}Ge_x$ layers, it has been found that resonant tunneling dominates over the transport mediated by the thermally activated charge transfer for low applied voltages. Peak-to-valley ratio reaches 11 at room temperature. At high voltages ($V_{bias}>0.8-1.0$ V), thermally activated transfer determines the electric current passing through the structure.

DOI: 10.1103/PhysRevB.75.115336 PACS number(s): 85.30.Mn, 73.40.Gk, 73.63.Hs, 73.50.Bk

I. INTRODUCTION

In the last decades, nanosized structures in the form of periodic quantum wells and barriers were actively studied due to their prospects for nanoelectronic devices operating at THz range.[1,2] Traditionally, such devices are designed on the basis of $A^{III}B^{V}$ semiconductors of which good lattice matching provides high quality interfaces important for undisturbed charge carrier transport. So far, resonant tunneling diodes (RTD) operating at 2 THz at room temperature have been reported.[3] However, $A^{III}B^{V}$ semiconductors are hardly integrated with the widely spread SiGe BiCMOS technology. That is why the opportunity to use SiGe alloy to design quantum heterostructures appropriate for resonant tunneling devices is of practical importance.

Recently, resonant tunneling effect at room temperature with a 1.3–2.4 peak-to-valley ratio has been experimentally demonstrated with strained $Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x$ quantum well grown on relaxed $Si_{1-y}Ge_y$ (001) virtual substrate.[4] As compared to a conventional RTD, preconfinement Si regions have been introduced between the unstrained $Si_{1-y}Ge_y$ injecting n-type contacts and $Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x$ active part. Suggested in 1991 by Wie and Choi in $A^{III}B^{V}$ systems and by Ismail *et al.* for SiGe/Si,[5] this band structure design ($x>y$) (*i*) reorganizes the three-dimensional (3D) incoming charge carriers into two-dimensional (2D) electron gas before their following transfer in the active device part and (*ii*) increases the overall conduction band discontinuity. To improve an efficiency of the structure proposed for RTD application, an advanced theoretical investigation of the carrier transfer through the double barrier is required. The anisotropy of the carrier effective mass and the different conduction band valleys in Si and SiGe should be accounted for. In addition, the "relaxed type" of the conduction band profile of an emitter formed by the injecting contact and the preconfinement region supports thermally activated carrier transport.[4] In its turn, the latter can significantly modify the current-voltage (I-V) characteristics of the device, as it has been observed for silicon-insulator multiple quantum wells.[6,7]

The goal of this paper is presentation of a comprehensive model of electron transport mechanisms through $Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x$ quantum well and results of a numerical simulation of the I-V curve. The model is based on the realistic band structure of $Si_{1-x}Ge_x$ and Si layers modified by internal strains, i.e., including the different conduction-band valleys with anisotropic effective masses of electrons. Moreover, special attention is paid to the effects of the charge localization in the Si emitter accumulation layer.

II. BAND DIAGRAM DESIGN

The general conditions of a quantum effects appearance in the systems with potential barriers and wells are observance of the charge carrier reflection inside the quantum well that provides the interference of an electron wave, the quantization of a transverse energy, and generates quantum levels. For this, the growth of $Si/Si_{1-x}Ge_x$ heterostructures onto unstrained $Si_{1-y}Ge_y$ (001) templates by means of the molecular beam epitaxy or the chemical vapor deposition techniques have now reached maturity that allows at certain conditions the fabrication of the appropriate and elaborate strained stacks of the nanometer-scaled layers with the abrupt, sharp and defect-free interfaces.[8] Due to the important lattice mismatch between Si and Ge bulk lattice parameters ($a_{Si}=5.4311$ Å, $a_{Ge}=5.6579$ Å), strain versus compositions however imposes the drastic limitations of the achievable thickness of a $Si_{1-x}Ge_x$ layer pseudomorphically grown on the relaxed $Si_{1-y}Ge_y$. Let us define the mismatch as $\varepsilon(x,y)=[a(x)-a(y)]/a(y)$, where the lattice parameters $a(x)$ and

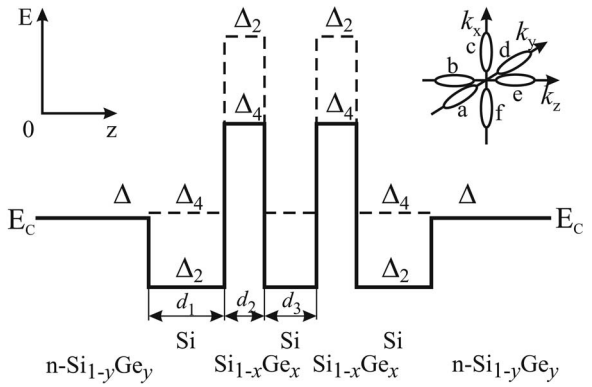

FIG. 1. The conduction band profile along the strained n-Si$_{1-y}$Ge$_y$/Si/Si$_{1-x}$Ge$_x$/Si/Si$_{1-x}$Ge$_x$/Si/n-Si$_{1-y}$Ge$_y$ structure with the strain-induced energy-splitted bands. The inset shows the six isoenergetic ellipsoidal surfaces for the unstrained Si$_{1-y}$Ge$_y$.

$a(y)$ are interpolated by $a_{\mathrm{Si}}(1-x)+a_{\mathrm{Ge}}x-c_Bx(1-x)$ with a lattice parameter bowing $c_B=0.027$ Å. The first limitation is the critical thickness h_c at which relaxation occurs by the formation of dislocations, which are known to strongly impede the electronic transport properties. For a high mismatch $\varepsilon(x,y)\geqslant 3.0\%$, the so-called Stranski-Krastanov growth mode is an additional issue. A complex interplay between the surface adatom diffusion and the strain field results in the formation of a dome at a given coverage θ_{SK}. In the following, we only focus on the main limitation in thickness h_c that is related to the generation of threading dislocations at Si$_{1-x}$Ge$_x$/Si$_{1-y}$Ge$_y$ strained and/or relaxed heterointerfaces with moderate mismatch $\varepsilon(x,y)<2.5\%$ (e.g., $x<0.85$ with $y=0.25$). From another viewpoint, strain versus Ge content in the Si$_{1-x}$Ge$_x$ layers controls the band gap, the band splitting, the effective masses, and the band discontinuities at Si$_{1-x}$Ge$_x$/Si$_{1-y}$Ge$_y$ interfaces. Thus, it is the essential engineering tool for designing RTDs and other heterostructures.

Owing to the differences in the energy gaps and electron affinities, the interchange of silicon and SiGe layers forms the Si/Si$_{1-x}$Ge$_x$/Si type II quantum wells.[9] This basic statement has been confirmed by experiment and theory: namely (1) the band-lineup procedure outlined by Ben Zid *et al.* in the framework of the model-solid theory of Van de Walle and Martin (see Ref. 10 and references therein) and (2) the empirical pseudopotential calculations of Rieger and Vogl,[11] both extended to Si$_{1-x}$Ge$_x$/Si$_{1-y}$Ge$_y$ (001) strained and/or relaxed heterointerfaces. Hence, while holes are energetically confined in a strained Si$_{1-x}$Ge$_x$ well of $\approx 0.88x$ (eV) depth, electrons injected from an unstrained Si contact layer must overcome a potential barrier of $\approx 0.38x^2$ (eV).[10] Initially, this insufficient value of the conduction band offset has prompted Ismail *et al.* to develop a RTD structure based on Si$_{1-x}$Ge$_x$/Si/Si$_{1-x}$Ge$_x$ double barrier sandwiched in n-type-Si$_{1-y}$Ge$_y$/Si emitter and Si/n-type-Si$_{1-y}$Ge$_y$ collector.[5]

The energy diagram of the conduction band of this RTD structure is presented in Fig. 1. The Si and Si$_{1-x}$Ge$_x$ $(x>y)$ layers are grown pseudomorphically strained onto a relaxed Si$_{1-y}$Ge$_y$ (001) pseudosubstrate thus having tensile and compressive distortion of their lattice parameter, respectively. Since it is assumed that $y<x<0.85$, the conduction band minimum[10,11] and the band gap[12] are defined by the Δ critical point of the Brillouin zone for each constitutive layer of the stack. The strain splits the sixfold energetically degenerated Δ conduction band of unstrained SiGe into two valleys b and e elongated towards [100] direction denoted as Δ_2 and, four ellipsoidal valleys a, d and c, f denoted as Δ_4, elongated towards [010] and [001] directions of the momentum space, respectively. The energy minimum is Δ_2 for tensile strain $(x<y)$ and Δ_4 for compressively strained layers $(x>y)$. Setting as a reference the energy of the Δ conduction band in unstrained Si$_{1-y}$Ge$_y$, the following first-order approximated analytical laws of the conduction band discontinuities between strain-splitted valleys Δ_2 and Δ_4 has been established using the procedure and parameters of Ref. 10,

$$\Delta E_C^{\Delta_2-\Delta} \cong (x-y)\{0.337+0.206(x+y)+(0.212+0.245x) \times[1+0.138(x+y)]\} \quad \text{in units of eV}, \tag{1}$$

$$\Delta E_C^{\Delta_4-\Delta} \cong (x-y)\{0.337+0.206(x+y)-(0.386-0.261x) \times[1+0.138(x+y)]\} \quad \text{in units of eV}. \tag{2}$$

The rest of it is a rather large value of the strain splitting between the Δ_2 and Δ_4 equivalent valleys that is given by:

$$E^{\Delta_2}-E^{\Delta_4} \cong (x-y)(0.598+0.016x) \times[1+0.138(x+y)] \quad \text{in units of eV}. \tag{3}$$

For unstrained Si$_{1-y}$Ge$_y$, the longitudinal and transverse components of the effective mass tensor for electrons are weakly dependent upon the composition, being $m_l \approx 0.9m_0$ and $m_t \approx 0.2m_0$, respectively, where m_0 is the free electron mass.[11] As the carrier transport takes places along the [001] direction, contributions from valleys b and e with $m_z=m_l$, $m_y=m_x=m_t$ and the four remaining valleys a, d with $m_z=m_t$, $m_y=m_t$, $m_x=m_l$ and c, f with $m_z=m_t$, $m_y=m_l$, $m_x=m_t$ are taken into account in this work. For simplicity sake, the difference of the electron effective masses in the layers having different compositions and strain states is however disregarded, although their changing takes place along the structure.[11]

The design of the "relaxed type" Si$_{1-y}$Ge$_y$/Si injecting contacts is a compromise between the highest $\Delta E_C^{\Delta-\Delta_2}(x=0,y)$ discontinuity and the larger thickness of the Si regions d_1 that collect incoming electrons. Widths in the 8–13 nm range generate the appearance of the quantum levels which are closely distributed at the bottom of the well, thus ensuring a good 2D preconfinement of the charge carriers. $d_3<d_1$ is an obvious condition as the width d_3 of the central well sets up the energy of the resonant levels. Efficient tunneling is provided by the highest Si/Si$_{1-x}$Ge$_x$ potential barrier $\Delta E_C^{\Delta_4-\Delta_2}=\Delta E_C^{\Delta_4-\Delta}+\Delta E_C^{\Delta-\Delta_2}$ (hence the highest $\Delta E_C^{\Delta_4-\Delta}$) and the larger barrier width d_2. For a realistic epitaxy, it is arbitrarily assumed that each elemental layer of the structure has a thickness below the critical thickness and corresponds to a direct growth of a single Si$_{1-x}$Ge$_x$ layer onto relaxed Si$_{1-y}$Ge$_y$. This hypothesis on the thermodynamic stability of each individual strained layer is strictly valid for the first Si layer in the stack. For that purpose, we use the law of h_c given by Mathews and Blakeslee,[13] which is known to be

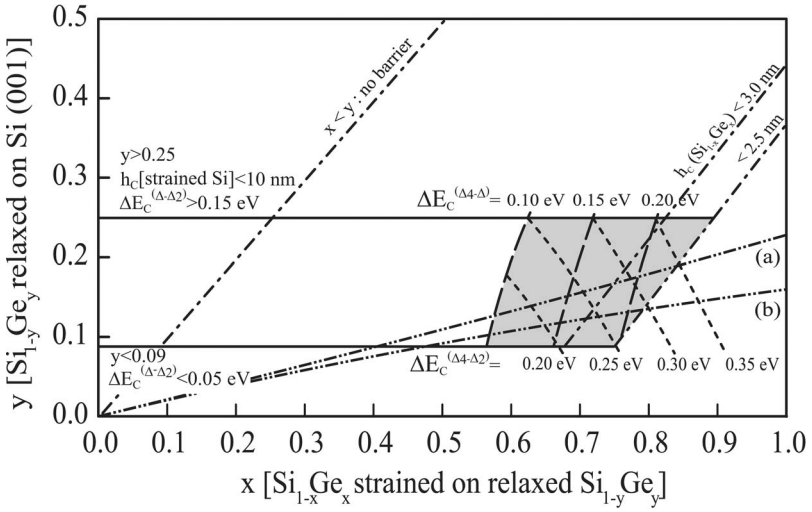

FIG. 2. The boundary conditions (gray area) defined by the Si and $Si_{1-x}Ge_x$ critical layer thickness and band offsets setting up the germanium fractions (x,y) for the design of the relaxed-strained-relaxed $n-Si_{1-y}Ge_y/Si/Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x/Si/n-Si_{1-y}Ge_y$ symmetrical stack. The optimum domain is cross-hatched by the isoenergetic solutions of the band discontinuities $\Delta E_C^{\Delta_4-\Delta_2}$ and $\Delta E_C^{\Delta_4-\Delta}$. $Y=f(x)$ solutions of Eq. (5) for strain-compensated stacks with $d_1=10$ nm, $d_2=3$ nm, and $d_3=3$ nm: (a) with $M=0$, $C_B=0.027\,45$ Å and assuming no difference in the elastic strain energy parameter σ_i; (b) with $M=2$, $C_B=0$ and taking a compositional dependence of $\sigma_{Si_{1-x}Ge_x}$.

pessimistic regarding other models and experimental data in semiconductor lattice-mismatched cubic systems. Briefly, for a 60° threading dislocation with a Burger vector magnitude $B=a(x)\sqrt{2}/2$, the latter is given in Å units by

$$h_c = \frac{B}{2^N \pi \varepsilon(x,y)} \frac{1-\nu_P/4}{1+\nu_P} \ln\left(1+\frac{h_c}{B}\right), \tag{4}$$

where $\nu_P=\frac{C_{12}}{C_{11}+C_{12}}$ is the Poisson ratio for (001) orientation that is calculated assuming a compositional linear interpolation of the elastic stiffness coefficients $C_{ij}(x)$ given in Ref. 10.

On the basis of these conditions, numerical solutions of the equations (1), (2), and (4) (with $N=2$) were used to find out an (x,y) appropriate composition domain for the structure design. As one can see in Fig. 2, a pseudosubstrate composition $y>0.25$ limits the critical thickness of the Si preconfinement strained layer to $d_1=10$ nm. Hence, the $\Delta E_C^{\Delta-\Delta_2}$ offset cannot exceed 0.15 eV (horizontal solid line). In addition, for a preconfinement of the incoming electrons, we do not allow the Si emitter well depth to be lower than $\Delta E_C^{\Delta-\Delta_2}=0.05$ eV $(y<0.09)$. Another reason for this is to satisfy an overall barrier height $\Delta E_C^{\Delta_4-\Delta_2}$ of 0.15 eV, at least. $\Delta E_C^{\Delta_4-\Delta}>0.1$ eV is then achieved for $x>0.56$. Within the assumption of attainable h_c for a single $Si_{1-x}Ge_x$ layer, it ends up that the counterpart for larger values of $\Delta E_C^{\Delta_4-\Delta}$ and $\Delta E_C^{\Delta_4-\Delta_2}$ is a thinner d_2 barrier width. This is graphically illustrated in Fig. 2. The appropriate composition domain of the strained $Si_{1-x}Ge_x$ layers on relaxed $Si_{1-y}Ge_y$ contacts lies within the gray area marked out in Fig. 2. For instance, with $y=0.25$, x should be kept below 0.83 to guarantee a $d_2=3$ nm minimal thickness of the barrier.

The assumption made is again pessimistic as strain balancing occurs at each interface of the stack allowing the next thin layer to keep out an additional excess of stress. This strain compensation results from the successive tensile and compressive strain state of the individual layers. In a general way, the strain compensation is supposed to be partial in this stack. This means that each layer adds to the overall stress and leads to an additional critical thickness for the whole structure. For a periodic superlattice built on a $Si/Si_{1-x}Ge_x$ bilayer, setting $N=0$ in Eq. (4) would give a valuable law with a predicted value of $h_c \sim 4$ times greater than the one given for a single layer (see Ref. 13). Nevertheless, if the RTD stack is symmetrical, it is nonperiodic. Therefore, one cannot use or extract an extra critical thickness for the whole structure because it is not possible to predict at which interface a dislocation should be generated. In order to put an additional constraint on the RTD parameters (x, y, d_1, d_2, d_3), one can arbitrarily assume the special case of a full strain compensation. Obviously, the critical thickness of the overall stack is not a design-limiting factor in such a case. From the three strain-balanced criteria given by Ekins-Daukes *et al.*,[14] we have derived the following common formula:

$$y = \frac{x}{1+\frac{\sigma_{Si}}{\sigma_{Si_{1-x}Ge_x}}\frac{2d_1+d_3}{2d_2}\left(1+x\frac{a_{Ge}-a_{Si}-c_B(1-x)}{a_{Si}}\right)^M \times\left(\frac{a_{Ge}-a_{Si}-c_B(1-x)}{a_{Ge}-a_{Si}-c_B(1-y)}\right)}, \tag{5}$$

where $\sigma_i^{(001)}=\left(\frac{E_Y}{1-\nu_P}\right)=\frac{(C_{11}-C_{12})(C_{11}+2C_{12})}{C_{11}}$ is the (001) elastic strain-energy parameter for Si and $Si_{1-x}Ge_x$; E_Y being the Young modulus. In the present formula, the exponent M stands for the chosen criteria, e.g., $M=0$ for the average-lattice method, $M=1$ for the thickness-weighted method, and $M=2$ for the zero-stress method. Equation (5) has been solved for the three criteria setting $d_1=10$ nm, $d_2=3$ nm, and $d_3=3$ nm. It is worth noting that the numerical difference between the slope of the curves obtained by the three methods is small (±0.6%). As shown in Fig. 2, the main discrepancy (±3.1%) arises when either the lattice parameter bowing or the difference in the elastic stiffness coefficients is taken into account. Within the optimum domain of compositions $(0.09<y<0.25, 0.56<x<0.83)$ set above, this calculation shows that strain-compensated stacks can be designed as far as $0.17<\frac{2d_2}{2d_1+d_3}<0.59$.

III. MODEL OF ELECTRON TRANSPORT

The model of the carrier transport is developed for a n-$Si_{1-y}Ge_y/Si/Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x/Si/n$-$Si_{1-y}Ge_y$ device

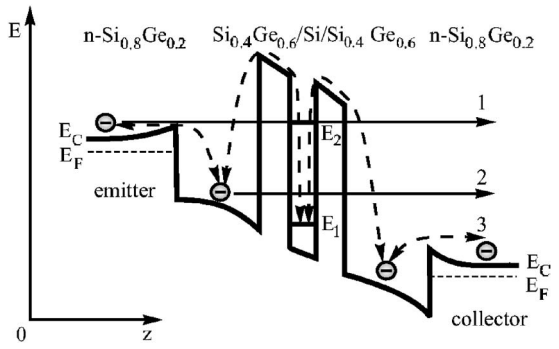

FIG. 3. The energy band diagram of the n-$Si_{1-y}Ge_y$/Si/$Si_{1-x}Ge_x$/Si/$Si_{1-x}Ge_x$/Si/n-$Si_{1-y}Ge_y$ (y=0.20, x=0.60) stack under an external bias, and carrier behaviors inside the structure: 1 is the tunneling transfer from extended states, 2 is the tunneling transfer from the localized states in preconfinement region, and 3 is the thermally activated carrier transport.

in which external n-$Si_{1-y}Ge_y$/Si and Si/n-$Si_{1-y}Ge_y$ regions act as an electron injector and a collector for the internal $Si_{1-x}Ge_x$/Si/$Si_{1-x}Ge_x$ quantum well. The lower energy conduction band is solely considered, that corresponds to the sixfold Δ, fourfold Δ_4, and twofold Δ_2 degenerated valleys for the $Si_{1-y}Ge_y$ injecting contacts, the $Si_{1-x}Ge_x$ potential barriers, and the Si preconfinement regions and the central well, respectively [see Eq. (3)]. Availing that the injector is a strong electron source, we consider only electron transport in the device neglecting a minority-hole component.

The energy diagram of the device under an external applied voltage V_{bias} and possible mechanisms of the electron transport are shown schematically in Fig. 3. The elastic tunneling turning to the resonant tunneling at the coincidence of the energy of the injected electrons and the energy levels in the silicon central quantum well is supposed to be a general carrier transfer mechanism (mechanism 1 in Fig. 3). The preconfinement Si regions before and after the $Si_{1-x}Ge_x$/Si/$Si_{1-x}Ge_x$ stack accumulate electrons. Electrons fill the states possessing the energy close to the bottom of the silicon well that can be identified as localized states. Electron tunneling from these states through the stack provides an additional component to the charge flow through the device (mechanism 2 in Fig. 3). The relatively low potential barriers for electrons and "relaxed type" of the injecting contacts open a possibility for thermally activated emission over the barrier which can be very efficient depending on temperature, doping, and interface quality (mechanism 3 in Fig. 3). The competition of the forward and the reverse thermally activated electron transfers, which are under nonequilibrium conditions for an applied external bias that direct charge carrier flow toward field vector, is observed.

One of the peculiarities of the model is to take the thermally activated transfer through the structure into consideration. The presence of the opposite components of this transfer provides a predominate charge carrier flow from the preconfinement region back to injecting contact over the flow of these charge carriers through the $Si_{1-x}Ge_x$/Si/$Si_{1-x}Ge_x$ well. Their contribution depends on the band offset between Si/$Si_{1-y}Ge_y$ and Si/$Si_{1-x}Ge_x$ paths.

The height of the potential barriers, doping, and quality of interfaces mainly define a prevalence of one of the above mechanisms. For modeling purpose, the electron charges are supposed to flow in the z direction that is normal to the layer planes, as it is shown in Fig. 3. The charge distribution in the injecting contacts is defined within the Boltzmann approximation of the Fermi-Dirac statistic,

$$n(z) = n_{ie} \exp\left(\frac{(\phi(z)q - E_F)}{k_B T}\right), \tag{6}$$

where n_{ie} is the effective equilibrium concentration of electrons at the contact at zero bias, q is the elementary charge, k_B is the Boltzmann constant, T is the temperature, $\phi(z)$ is the electrostatic potential extension along the z direction, and E_F is the Fermi energy. The charge injection occurs from the emitter (E) and collector (C) of the device. The Fermi energy is constant $E_{F(E)}=E_F$ at the emitter contact and $E_{F(C)}=E_F - qV_{\text{bias}}$ at the collector contact. At zero bias, $E_{F(E)}=E_{F(C)}=E_F$ and the charge injections from emitter and collector contacts are identical. Otherwise, the redistribution of the higher electron density close to collector contact takes place and equilibrium between the two current flows is broken.

Electrons fill all available states in the active part of the device according to the Fermi-Dirac $f(E)$ distribution.[6] The electron density in the active part of the device is described in the following way:

$$n(x,y,z) = 2\sum_{v}\sum_{\kappa} |\psi_v(x,y,z)|^2 f(E), \tag{7}$$

where the factor of 2 stands for the spin degeneracy and a summation is performed over all conduction band valleys v and all available values of the wave vector κ; $\psi_v(x,y,z)$ being one electron wave function. Because of the spatial invariance in the xy plane, the wave function of the electron is separable $\psi_v(x,y,z)=\psi_v(z,k_z)\chi(x,y)$, where the transverse wave function is described by plane waves as $\chi(x,y)=\exp(ik_x x+ik_y y)$. The longitudinal wave function $\psi_v(z,k_z)$ along the z direction corresponding to wave vector value k_z is one solution of the Schrödinger equation,

$$-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\left(\frac{1}{m_z}\frac{\partial}{\partial z}\right)\psi_v(z) + V(z)\psi_v(z) = E_z\psi_v(z), \tag{8}$$

where $\hbar$ is the reduced Plank's constant; m_z is the effective electron mass of electron along the z direction. $V(z)$ is the potential energy of electron defined as follows:

$$V(z) = -q\phi(z) + \Delta E_C(z), \tag{9}$$

where $\Delta E_C(z)$ is the conduction band profile along the structure.

Assuming a parabolic dispersion relation in the case of isoenergetic elliptic surface for the different conduction valleys along the whole structure, and after performing integration over the transverse component of the momentum for the total electron density, the charge distribution expression is found to be

$$n(z) = \frac{k_B T}{\pi \hbar^2} \left\{ \sum_{k=1}^{m} \psi_l^*(z,k_l)\psi_l(z,k_l) g_2 m_t \times \ln\left[1 + \exp\left(\frac{E_F - E(k_l)}{k_B T}\right)\right] + \sum_{k=1}^{m} \psi_t^*(z,k_t)\psi_t(z,k_t) g_4 \sqrt{m_t m_l} \times \ln\left[1 + \exp\left(\frac{E_F - E(k_t)}{k_B T}\right)\right] \right\}, \quad (10)$$

where the first term describes the contribution from the two valleys b, e; whereas the second term corresponds to the contribution from the four valleys a, d and f, c (see inset in Fig. 1 and Sec. I). In the above expressions, the summation over momentum has been replaced by the integration index. Factors $g_2=2$ and $g_4=4$ stand for the valley degeneracy.

For an analysis of the thermally activated carrier transport, the electron transfer components in two opposite directions from the preconfinement silicon regions are included. The rates of electron n_i concentrations in the ith point of mesh are written as[6]

$$\frac{dn_i}{dt} = g_{i-1\to i} n_{i-1} - n_i g_{i\to i+1} - n_i g_{i\to i-1} + n_{i+1} g_{i+1\to i}, \quad (11)$$

where n_i is the electron concentration in the ith mesh point ($i=1,\ldots,N$, where N is the number of mesh points) and the $g_{i\to i+1}, g_{i+1\to i}$ are the rates of carrier transfer from i to $i+1$ mesh point and in the opposite direction. This rate that we have defined within the thermally activated transport theory includes the image-force effect in the following way:[6,15]

$$g_{i\to i+1} = \omega_0 \exp\left(\frac{q(\phi_i - \phi_{i+1})}{k_B T}\right) \times \exp\left[-\left(\Delta E_{Ci} - \sqrt{\frac{q^3(\phi_i - \phi_{i+1})}{16\pi\varepsilon_i\varepsilon_0 hz_i}}\right) \Big/ k_B T\right], \quad (12)$$

where ε_0 is the permittivity in vacuum, ε_i is the relative permittivity of the materials, hz_i is the step of the mesh, and ω_0 is the frequency of the local carrier oscillations in the quantum well [$\omega_0 = 10^{12}$ s^{-1} (Ref. 16)]. All variables $\phi(z)$, $\Delta E_C(z)$, and $\varepsilon(z)$ have been attached to discrete mesh. The system of kinetic equations has been solved for the stationary case.

The charge accumulation in the quantum wells is nonuniform and defines the potential distribution in the structure. To evaluate the potential $\phi(z)$ drop along the z direction, the Poisson equation is solved,

$$\frac{\partial}{\partial z}\left(\varepsilon(z)\frac{\partial}{\partial z}\right)\phi(z) = \frac{q[N_D(z) - n(z)]}{\varepsilon_0}, \quad (13)$$

where $N_D(z)$ is the ionized donor doping concentration.

The finite-difference method has been applied for the Schrödinger and the Poisson equation solutions. The real space along the stack has been divided into discrete uniform mesh points and equations have been solved within those discrete spacings. Within this method, the discretization of the differential equation (8) and the Poisson equation (13) has been performed by using the three-point finite difference scheme.[17] The interaction procedure is used to obtain a self-consistent solution of the Schrödinger equation (8), the kinetic equation system for the thermally activated carrier transport (11), and the Poisson equation (13).

After extracting the self-consistent solutions of the Schrödinger and the Poisson equations providing the potential distribution to be known, the distribution of the charge carrier density along the structure is again defined. Finally, the probability of the resonant tunneling has been determined by a system of wave function equations which consider the incident and reflected waves in every point of the discrete mesh and include the boundary conditions at points 1 and N,[18] according to

$$\psi_{l(t),1} = \exp[ik_{l(t)}(1)z] + A\exp[-ik_{l(t)}(1)z],$$

$$\psi_{l(t),N} = B\exp[ik_{l(t)}(N)z], \quad (14)$$

where A, B are the amplitude coefficients. Thus, the transmission coefficient for tunneling is given by

$$T(E_{x(z)}, k_{l(t)}) = \frac{k_{l(t)}(N)}{k_{l(t)}(1)}|B|^2. \quad (15)$$

This transmission coefficient includes 3D to 2D resonant tunneling, namely from injecting contacts to the active part of the structure, where quantum effects are predominant. Yet, calculation of the potential profile along the whole structure within this method misses the current from localized states in preconfinement regions, which are occupied according to the Fermi-Dirac distribution. The contribution of the localized states to the total current is supposed to be significant in the case of the positive energy gap between the injecting contact and the bottom of the quantum well in the preconfinement region. For this reason the procedure of the probability calculation has been used for the active part of the device excluding the potential profile at the contact regions that allows to account the contribution of 2D-2D current from localized states in the preconfinement regions.

The formula after Tsu and Esaki[19] generalized for the case of the several conduction valleys and the effective mass anisotropy is used for the tunneling current calculation. For the parabolic dispersion relations with integration over the transverse component, the current density from extended and localized states has the form

$$J_{\text{ext,loc}} = \frac{qk_B T}{2\pi^2\hbar^3}\int [m_t g_2 T(E_z,k_l) + \sqrt{m_t m_l} g_4 T(E_z,k_t)] \times [f(E_z) - f(E_z + qV_{\text{bias}})]dE_z. \quad (16)$$

The currents related to extended and localized states are calculated with different integration procedures. For the localized carriers, the conduction band of injecting contacts restricts the upper limit of the integration. The calculation of the thermally activated carrier transport is performed with the same expression, but the total valley degeneracy $g_6=6$

TABLE I. Parameters of the n-$Si_{1-y}Ge_y/Si/Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x/Si/n$-$Si_{1-y}Ge_y$ devices.

Parameter	Value
Conduction band energy of $Si_{1-x}Ge_x$ barrier, $\Delta E_C^{\Delta_4-\Delta}$	0.03–0.25 eV
Conduction band energy at $Si/Si_{1-y}Ge_y$ injecting contact, $\Delta E_C^{\Delta-\Delta_2}$	0.05–0.15 eV
Transverse effective mass, m_t	$0.2m_0$ (Ref. 12)
Longitudinal effective mass, m_l	$0.9m_0$ (Ref. 12)
Thickness of Si preconfinement region, d_1	8–13 nm
Thickness of potential $Si_{1-x}Ge_x$ barrier, d_2	2.5–3.5 nm
Thickness of Si middle quantum well, d_3	2.5–4 nm
Temperature, T	150–300 K

and an effective mass $m_{\text{eff}}=(m_t^2 m_l)^{1/3}$ are used.

IV. RESULTS AND DISCUSSION

The experimental I-V characteristics of the n-$Si_{1-y}Ge_y/Si/Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x/Si/n$-$Si_{1-y}Ge_y$ devices ($x=0.60$, $y=0.20$, $d_1=10$ nm, $d_2=2$ nm, $d_3=3$ nm) (Ref. 4) have broad resonant peaks with the low current peak-to-valley contrasts at room temperature (1.3–2.4), that is insufficient for digital applications. To establish the conditions for an efficient resonant-tunneling effect, the I-V curves were calculated within the above model. The parameters used are summarized in Table I. The donor concentration in the n-type $Si_{1-y}Ge_y$ with $0.09<y<0.25$ was fixed to $N_D=3\times10^{18}$ cm^{-3}. At room temperature, this value sets the Fermi level position 0.05 eV below the conduction band of the $Si_{1-y}Ge_y$ emitter and the collector. In order to reduce the direct electron transmission from injector to collector contact, the thickness of the preconfinement regions was set large enough in the 8–13 nm range, keeping it lower than the critical thickness of a Si layer under tensile strain on relaxed $Si_{1-y}Ge_y$ with $0.09<y<0.25$. The Si middle quantum well was varied in the range of 2.5–4 nm width that should provide the resonant-tunneling effect. The potential barrier widths were chosen in the range of 2.5–3.5 nm. As seen in Fig. 2, the critical thickness of these $Si_{1-x}Ge_x$ layers compressively strained on the relaxed $Si_{1-y}Ge_y$ limits their achievable composition, thus the $\Delta E_C^{\Delta_4-\Delta}$ band offset, in a drastic way.

Calculations of the transmission probability have shown the presence of two peaks for the transverse effective mass of electrons and three well-resolved peaks for the longitudinal one. The results are demonstrated in Fig. 4 where the transmission coefficient is plotted. The large value of the longitudinal effective mass of electrons m_l causes the related low transmission coefficient and the three levels appearance. That is the reason of the significant domination of the transmission component relative to the transverse mass. The vertical dotted line separates the energy range that is not considered for the modeling of the 3D-2D resonant tunneling transfer. For parameters of Fig. 4, it corresponds to a conduction band offset at injecting contacts $\Delta E_C^{\Delta-\Delta_2}=0.12$ eV. This confirms the requirements of the additional calculations of 2D-2D tunneling from localized states in the preconfinement quantum well for adequate current simulations. A reduction of the thickness of the middle quantum well d_3 shifts the second resonant level to the higher energy range and increases the peak transmission coefficient. For $d_3\leqslant 2.5$ nm, this level is actually delocalized over the $Si_{0.20}Ge_{0.80}$ barrier of height $\Delta E_C^{\Delta_4-\Delta_2}=0.32$ eV.

The contribution of electrons from localized and extended states to the total current and the role of thermally activated transport have been estimated. The simulation results are presented in Fig. 5. As it has been suggested, the current due to the charge transfer from 3D extended states is the dominant current provided by the transverse mass component. This current includes only one resonant peak due to the coincidence of the Δ conduction band of the injector with the second level in the middle quantum well. The transverse and longitudinal components of 2D-2D current related to localized states are significantly smaller. The potential barrier height $\Delta E_C^{\Delta_4-\Delta_2}=\Delta E_C^{\Delta_4-\Delta}+\Delta E_C^{\Delta-\Delta_2}$ seen by localized electrons is obviously larger than the potential barrier $\Delta E_C^{\Delta_4-\Delta}$ for 3D carriers in the injecting contact. For the highest achievable potential barrier $\Delta E_C^{\Delta_4-\Delta_2}=0.32$ eV preserving 3-nm-thick $Si_{0.20}Ge_{0.80}$ layer in pseudomorphic strain configuration, the peak current density reaches 1.64×10^{11} A/m^2 at $V_{\text{bias}}=0.36$ V and the peak-to-valley ratio equals 11. One should note that the calculated value of the peak current density is

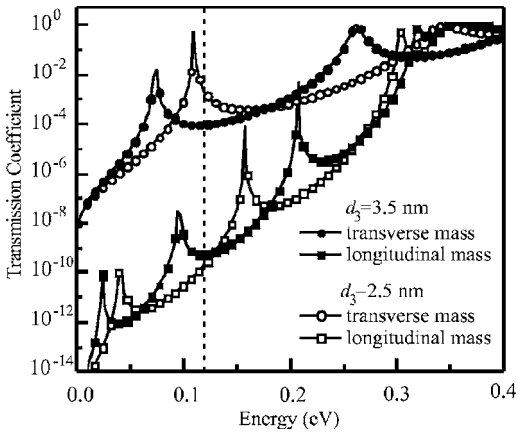

FIG. 4. Transmission coefficient vs electron energy for the transverse and longitudinal effective masses for two well widths $d_3=3.5$ nm and $d_3=2.5$ nm ($d_1=10$ nm, $d_2=3$ nm, $y=0.2$, and $x=0.8$).

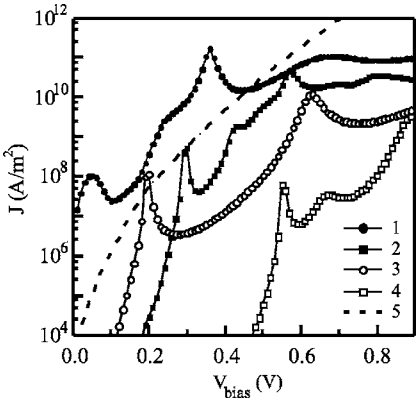

FIG. 5. Current-voltage characteristic of the n-$Si_{0.8}Ge_{0.2}$/Si/$Si_{0.2}Ge_{0.8}$/Si/$Si_{0.2}Ge_{0.8}$/Si/n-$Si_{0.8}Ge_{0.2}$ RTD including: 1, 2, extended states related current for transverse and longitudinal components, respectively; 3, 4, localized states related current for transverse and longitudinal components, respectively; 5, thermally activated current (d_1=10 nm, d_2=3 nm, d_3=3 nm, T=300 K).

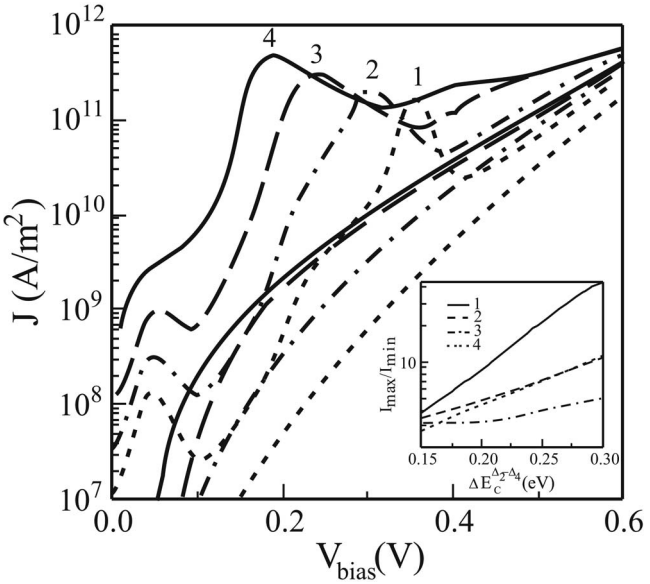

FIG. 6. The room temperature current-voltage characteristics of n-$Si_{1-y}Ge_y$/Si/$Si_{1-x}Ge_x$/Si/$Si_{1-x}Ge_x$/Si/n-$Si_{1-y}Ge_y$ devices and their corresponding thermal current components with d_1=10 nm, d_2=3 nm, d_3=3 nm, and y=0.20 ($\Delta E_C^{\Delta_2-\Delta_2}$=0.12 eV) for different $\Delta E_C^{\Delta_4-\Delta_2}$: 1, 0.30 eV ($x$=0.76); 2, 0.25 eV ($x$=0.67); 3, 0.20 eV ($x$=0.56); and 4, 0.15 eV ($x$=0.40). The inset gives the evolutions of the peak-to-valley ratios J_{max}/J_{min} as a function of $\Delta E_C^{\Delta_4-\Delta_2}$ for different values of barrier d_2 and well d_3 widths. 1, d_2=3.0 nm, d_3=3.5 nm; 2, d_2=3.0 nm, d_3=3.0 nm; 3, d_2=2.5 nm, d_3=3.0 nm; 4, d_2=3.5 nm, d_3=3.0 nm.

not absolute because the integration procedure depends on the overall stack thickness (see Sec. III). The occurrence of a small resonant peak having a maximum at V_{bias}=0.05 V cannot proceed from Fig. 4. It may be explained by the influence of the electrostatic charge distribution and accumulation across the structure modifying the potential profile, which is taken into account via the Poisson equation. Without considering the electrostatic potential extension $\phi(z)$ in the Poisson equation, this low bias peak clearly disappears and, in addition, the peak-to-valley ratio of the current component related to 3D extended states reaches 15 at V_{bias}=0.36 V.

The contribution of the thermally activated current component becomes significant and strongly competes with the tunneling current at large applied voltages. The parity between extended states related—and thermally activated—currents occurs at $V_{bias}\sim$0.47 V. It is worth noting that the fast growth of the thermally activated current with the applied voltage is not observed when one discounts the charge distribution in the structure. A 300 K to 100 K temperature drop (not shown) exponentially decreases the thermally activated current density [see Eq. (12)] of ~1 order of magnitude below the one due to the component resulting from the competition of the two opposite carrier flows through the central well, e.g., from the emitter contact towards the collector and vice versa. Meanwhile, the resonant tunneling process is almost linearly dependent on the temperature [see Eq. (16)].

The peak current in resonant tunneling devices is known to be very sensitive to the barrier height. In a general way, the larger the overall potential offset, the higher is the value of the J_{max}/J_{min} peak-to-valley ratio. Figure 6 presents the results of the I-V characteristics of devices having different barriers $\Delta E_C^{\Delta_4-\Delta_2}$ but thickness kept identical to those of Fig. 5. A reduction of $\Delta E_C^{\Delta_4-\Delta_2}$ pulls down the resonant level of the central well, the energy of which is becoming closer to the conduction band minimum in the injecting contact. This results in a shift of the peak current to smaller voltages with a corresponding increase of the peak current density. However, the significant dispersion of the resonant level with the lowering of the potential barrier height leads to a broadening of the resonant peak and a decrease of the J_{max}/J_{min} peak-to-valley ratio. On the other hand, for applied voltages above 0.40 V or barrier height greater than 0.30 eV, the thermally activated carrier transfer prevails in all cases.

Within the limit of the structure transmission, one of the ways to improve J_{max}/J_{min} ratio is an extension of the potential barrier width d_2 that provides dispersion of resonant levels in quantum well constriction. This is shown in the inset of Fig. 6 where J_{max}/J_{min} ratios are plotted as a function of $\Delta E_C^{\Delta_4-\Delta_2}$ for d_2=2.5 nm, d_2=3.0 nm, and d_2=3.5 nm. Large thickness of $Si_{1-x}Ge_x$ ($x>y$) limit tensile strain accumulation in the stack that is an additional advantage. Nevertheless, d_2 should be kept below the critical thickness providing the pseudomorphic strain state of a $Si_{1-x}Ge_x$ layer grown onto a relaxed $Si_{1-y}Ge_y$ template. As an example, assuming a critical thickness of 3.5 nm for y=0.20, the barrier composition should not exceed x=0.72. As mentioned above, the counterpart is the limitation of the achievable band offset $\Delta E_C^{\Delta_4-\Delta_2}$ to 0.28 eV.

A reduction of the thickness of the middle quantum well d_3 shifts the energy of the resonant level towards the upper limit of the $Si_{1-x}Ge_x$ barrier height. This moves the condition of resonance to the side of larger applied bias where the resonant peak is weaker pronounced, as a result of the quantum level broadening. The resonant peak vanishes at d_3=2.5 nm. Otherwise, increasing d_3 significantly enhances the peak-to-valley ratio, as it is illustrated in the inset of Fig. 6

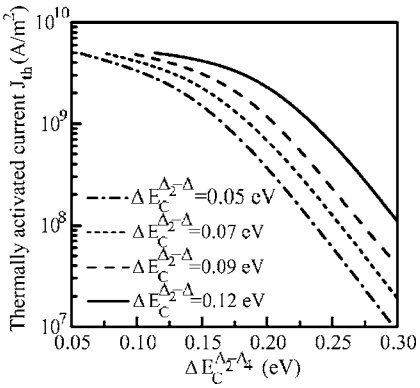

FIG. 7. The thermally activated current density as a function of the potential barrier height $\Delta E_C^{\Delta_4-\Delta_2}$ for n-$Si_{1-y}Ge_y$/Si/$Si_{1-x}Ge_x$/Si/$Si_{1-x}Ge_x$/Si/n-$Si_{1-y}Ge_y$ RTD structures at $T=300$ K and $V_{bias}=0.23$ V, with $d_1=10$ nm, $d_2=3$ nm, $d_3=3$ nm, $x<0.83$ and having different offsets $\Delta E_C^{\Delta-\Delta_2}$: 0.12 eV ($y=0.20$), 0.09 eV ($y=0.15$), 0.07 eV ($y=0.12$), and 0.05 eV ($y=0.09$).

for $d_3=3.5$ nm. Further enlargement of the well, however, turns off the participation of the second resonant level as its location is shifted down in energy below the Fermi level position in the emitter (see Fig. 3). Moreover, the well is limited in thickness as silicon layers introduce excessive tensile strain in these stacks. Such strong sensitivity of the conditions of resonance to the structural parameters can explain the small J_{max}/J_{min} ratio obtained with the experimental devices, which have far from optimized parameters. Indeed, the J_{max}/J_{min} ratio calculated for the set of parameters identical to the experimentally obtained devices[4] ($x=0.60$, $y=0.20$, $d_1=10$ nm, $d_2=2$ nm, $d_3=3$ nm, $N_D=3\times10^{18}$ cm^{-3}) is 2.8, that is in good agreement with the experimental values $J_{max}/J_{min}=1.3-2.4$.

One of the abilities to regulate the thermally activated current in this RTD device is provided by an appropriate choice of the relation between the overall potential barrier height $\Delta E_C^{\Delta_4-\Delta_2}$ and the conduction band offset at the "relaxed type" injecting contact. Results on such computer simulation are presented in Fig. 7. An increase of the potential barrier height $\Delta E_C^{\Delta_4-\Delta_2}$ is followed by exponential drop of the thermally activated current [see Eq. (12)]. For a fixed value of $\Delta E_C^{\Delta_4-\Delta_2}$, however, higher band offset $\Delta E_C^{\Delta-\Delta_2}$ between injecting contact and preconfinement region lead to slightly higher thermally activated current. These results demonstrate the possibility for a realistic tuning of the thermally activated component in this RTD device.

In addition, the performed calculations have good correlation with experimental results[4] and show a weak dependency of the shift of the resonant current peak position to the region of large bias versus $\Delta E_C^{\Delta_4-\Delta_2}$ (see Fig. 6) with a lowering of the conduction band offset $\Delta E_C^{\Delta-\Delta_2}$. That is simply explained by the need of an additional energy for an electron to achieve the corresponding resonant condition with the level in the middle quantum well. The simultaneous decrease of the thermally activated current density with the reduction of $\Delta E_C^{\Delta-\Delta_2}$ compensates this shift of the resonant current peak position. Moreover, the thermally activated current moves to the side of large bias with an increase of $\Delta E_C^{\Delta_4-\Delta_2}$ (see Fig. 6) and a decrease of $\Delta E_C^{\Delta-\Delta_2}$. Thus, it never exceeds the tunneling current. The main advantage of the investigated "relaxed type" design is a stable relationship between the tunneling current and the thermally activated one, which is not much sensitive to the variations of the energy profile.

V. CONCLUSION

The carrier transport in resonant tunneling devices has been modeled on the basis of a self-consistent solver of the Schrödinger, Poisson, and kinetics equations. We have found the range of the structural parameters, thickness, and composition, providing a realistic epitaxy of the optimized resonant tunneling heterostructures built-on strained $Si_{1-x}Ge_x$/Si/$Si_{1-x}Ge_x$ quantum well embedded in n-$Si_{1-y}Ge_y$/Si injector and Si/n-$Si_{1-y}Ge_y$ collector that acts as preconfinement regions of the electronic charge. It has been shown that the tunneling transfer mainly occurs via 3D extended states at the coincidence of the electron energy in the injector with the second resonant level in the middle quantum well that is located above the conduction band of the n-$Si_{1-y}Ge_y$ contact. The computer simulation has shown insignificant contribution of the 2D-2D electron transfer from states localized in the preconfinement silicon well due to the high potential barrier at Si/$Si_{1-x}Ge_x$ interface, although the resonant conditions for such mechanism appears more effective compared to those for extended states.

The central well and barrier widths are the key parameters setting the resonant energy level and the conditions controlling the second current peak in the I-V characteristics, which relates to resonant tunneling of electrons from extended states in the injecting contact. The thermally activated current is found to prevail at high-applied voltages only. The relationship between resonant tunneling—and thermally activated—currents is rather stable with the parameter variations, namely the band offsets. This allows well-resolved resonant current peak at low voltage with expected peak-to-valley ratio as high as 11 at room temperature.

ACKNOWLEDGMENTS

This study was carried out in the framework of the INTAS Contract No. YSF 2002-139 of the East-West Scientific Cooperation Program (http://www.intas.be/) and Contract No. T04-4047 of the Foundation for Basic Research of the Republic of Belarus.

*Electronic address: julia@physics.umanitoba.ca
†Electronic address: lazzari@crmcn.univ-mrs.fr
‡Laboratory associated with the Université de la Méditerranée, Aix Marseille II and the Université Paul Cézanne, Aix Marseille III.

[1] L. L. Chang, L. Esaki, and R. Tsu, Appl. Phys. Lett. **24**, 593 (1974).
[2] V. P. Gribkovskii, Prog. Quantum Electron. **19**, 41 (1995).
[3] Technology Roadmap for Nanoelectronics, European Commission, IST program, Future and Emerging Technology, edited by R. Compaño (Office for Official Publications of the European Communities, Luxembourg, 2000).
[4] D. J. Paul, P. See, I. V. Zozoulenko, K.-F.Berggren, B. Kabius, B. Hollaender, and S. Mantl, Appl. Phys. Lett. **77**, 1653 (2000); D. J. Paul, P. See, I. V. Zozoulenko, K.-F. Berggren, B. Hollaender, S. Mantl, N. Griffin, B. P. Coonan, G. Redmond, and G. M. Crean, Mater. Sci. Eng., B **89**, 26 (2002).
[5] C. R. Wie and Y. W. Choi, Appl. Phys. Lett. **58**, 1077 (1991); K. Ismail, B. S. Meyerson, and P. J. Wang, *ibid.* **59**, 973 (1991).
[6] J. A. Berashevich, A. L. Danilyuk, A. N. Kholod, and V. E. Borisenko, Mater. Sci. Eng., B **101**, 111 (2003).
[7] J. A. Berashevich, A. L. Danilyuk, and V. E. Borisenko, Semiconductors **36**, 718 (2002).
[8] D. J. Paul, Thin Solid Films **336**, 130 (1998).
[9] V. E. Borisenko and S. Ossicini, *What is What in Nanoworld* (Wiley-VCH, Weinheim, 2004).
[10] N. Sfina, J.-L. Lazzari, J. Derrien, F. Arnaud d'Avitaya, and M. Said, Eur. Phys. J. B **48**, 151 (2005).
[11] Martin M. Rieger and P. Vogl, Phys. Rev. B **48**, 14276 (1993).
[12] J. Weber and M. I. Alonso, Phys. Rev. B **40**, 5683 (1989).
[13] J. W. Matthews and A. E. Blakeslee, J. Cryst. Growth **27**, 118 (1974).
[14] N. J. Ekins-Daukes, K. Kawaguchi, and J. Zhang, Cryst. Growth Des. **2**, 287 (2002).
[15] S. M. Sze, *Physics of Semiconductor Devices* (Academic, New York, 1981).
[16] E. H. Nicollian, J. Vac. Sci. Technol. **14**, 1112 (1977).
[17] I.-H. Tan, G. L. Snider, L. D. Chang, and E. L. Hu, J. Appl. Phys. **68**, 4071 (1990).
[18] K. Araki, J. Appl. Phys. **62**, 1059 (1987).
[19] R. Tsu and L. Esaki, Appl. Phys. Lett. **22**, 562 (1973).