Phonon- and Auger-assisted tunneling from a quantum well to a quantum dot

Shu-Wei Chang, Shun-Lien Chuang,* and Nick Holonyak, Jr.
Department of Electrical and Computer Engineering, University of Illinois at Urbana—Champaign, Urbana, Illinois 61801, USA
(Received 15 December 2003; revised manuscript received 3 May 2004; published 20 September 2004)

We investigate phonon- and Auger-assisted tunnelings from an adjacent quantum well coupled to a quantum dot by Fermi's golden rule. The filling of quantum-well states is included, and we obtain an average tunneling lifetime depending on the concentration in the quantum well. Depending on the barrier width between the quantum dot and quantum well, both mechanisms can result in a tunneling time of a few to several hundred picoseconds. For the condition of high concentration of carriers in a quantum well, the typical time scale of the net increase of carriers in a quantum dot due to phonon-assisted tunneling can be in sub-picosecond range, which agrees with the recent experimental observation.

DOI: 10.1103/PhysRevB.70.125312 PACS number(s): 73.63.Kv, 73.63.Hs, 73.40.Gk, 72.10.Di

I. INTRODUCTION

Phonon-assisted tunneling was observed by Holonyak *et al.* in heavily-doped silicon *p-n* junctions at 4.2 K.[1] This was the first example of inelastic tunneling which became a powerful tool for inelastic electron tunneling spectroscopy.[2–4] In recent years, quantum dots (QDs) in semiconductor material systems are of great interest not only due to their atomic-like properties but also the potential applications in quantum computation (information)[5–7] and high-performance semiconductor lasers.[8–12] However, carrier capture by the QDs for semiconductor lasers is a problem because of the slow phonon relaxation of the carriers to the ground state of the QDs, which reduces the efficiency of lasing.[13–15] A tunneling-injection quantum well (QW) coupled QD laser structure has recently been proposed.[16–19] The QW serves as a collection layer of carriers, followed by lateral diffusion and tunneling injection into the QDs. Experimentally, reduced threshold current density has been achieved by the QW coupled to the QD.[17] Coherent tunneling injection from the QW to the QD has been investigated by Chuang and Holonyak.[20] In the QW-coupled-QD structure, various mechanisms such as photon-, phonon-, and Auger-assisted processes can also contribute to the tunneling. Chuang *et al.* analyzed the possibility of photon-assisted tunneling.[21] The recent experimental observation of the differential transmission spectrum in similar structures has shown a phonon-assisted tunneling time of 1.7 ps.[8]

For QDs grown in a QW wetting layer, a few carrier relaxation processes have been researched.[22–26] Among them, the relaxations caused by longitudinal-optical-phonon (LO) scattering and Auger scattering receive the most attention. Magnusdottir *et al.*[22] calculated the Auger relaxation rate from the wetting layer to the QD due to the Coulomb interaction of two carriers in the wetting layer. They also calculated the one-phonon and two-phonon assisted processes for spherical QDs embedded in three-dimensional continuum.[23] In their calculation, relaxation times of the order of picosecond are possible for both mechanisms.

In this paper, we analyze phonon-assisted and Auger-assisted tunneling rates in this QW coupled QD structure. We present our theoretical model for the QW and QD wave functions, followed by the phonon- and Auger-assisted capture from the QW to the QD.

II. MODEL FOR QW AND QD WAVE FUNCTIONS

The model is depicted in Fig. 1. The QW layer serves as the collection layer of electrons. The electrons in the QW can tunnel into the QD by emitting an LO phonon or transferring energy via Coulomb repulsion to another electron in the QW.

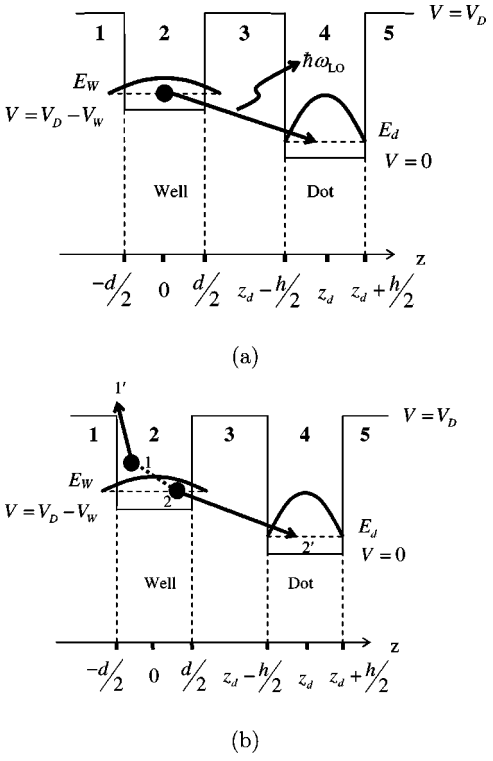


FIG. 1. (a) The QW to QD LO-phonon-assisted tunneling. The carriers in the QW emit an LO phonon and tunnel into the QD. (b) The QW to QD Auger-assisted tunneling. Two electrons 1 and 2 in the QW interact by Coulomb repulsion. One makes a transition to a QD state labeled as $2'$ while the other is excited to a higher QW-energy state labeled as $1'$.

In the following formulation, we use the effective mass approximation and neglect the difference between the Bloch periodic parts of the single-particle wave functions in the QW and QD. For LO-phonon-assisted tunneling, the Fröhlich Hamiltonian using bulk modes will be adopted since it has been shown that in QD-related scattering problems the difference between confined and bulk LO-phonon modes is not so significant.[27] For Auger-assisted tunneling, the screening effect will be neglected to simplify the situation.

The QD is modeled as a quantum disk with a height h and radius ρ_0. The distance from the center of the QW to that of the QD is denoted as z_d. We set the band edges of the barrier, QD, and QW regions as V_D, 0, and $V_D - V_w$, where V_D and V_w are the potential depths of QD and QW, respectively. For simplicity, we assume that the QD envelope wave function can be approximated by the product of the in-plane and growth-direction parts, namely,

$$\phi_{\bar{N}}(\mathbf{r}) = \phi_{lmn}(\mathbf{r}) = \Psi_{lm}(\boldsymbol{\rho})\varphi_n(z) = R_{lm}(\rho)\frac{e^{im\phi}}{\sqrt{2\pi}}\varphi_n(z), \quad (1)$$

where $\bar{N}$ represents the quantum numbers (l,m,n), which are in-plane radial quantum number, magnetic quantum number, and the growth-direction quantum number, respectively. This approximation works for the QD states whose eigenenergies are not close to the barrier band edge.

The explicit forms of the approximate growth-direction wave function and in-plane wave function of QDs are as follows:

$$\varphi_n(z) = \begin{cases} B_n^z e^{i\theta_n} e^{\nu_n(z-z_d)}, & \text{in regions } \mathbf{1},\mathbf{2},\mathbf{3}, \\ \dfrac{A_n^z}{2}(e^{-i\theta_n/2}e^{ik_n(z-z_d)} + e^{i\theta_n/2}e^{-ik_n(z-z_d)}), & \text{in region } \mathbf{4} \\ B_n^z e^{-\nu_n(z-z_d)}, & \text{in region } \mathbf{5}; \end{cases} \quad (2)$$

$$\theta_n = \begin{cases} 0, \text{ if even parity for QD Hamiltonian in growth direction}, \\ \pi, \text{ if odd parity for QD Hamiltonian in growth direction}; \end{cases} \quad (3)$$

$$R_{lm}(\rho) = \begin{cases} A_{lm}^{xy} J_m(\kappa_{lm}\rho), & \rho \leqslant \rho_0, \\ B_{lm}^{xy} K_m(\gamma_{lm}\rho), & \rho \geqslant \rho_0; \end{cases} \quad (4)$$

where B_n^z and A_n^z are the normalized constants of the growth-direction wave function; ν_n and k_n are the parameters describing the penetration into barrier and standing-wave behavior in the QD region; A_{lm}^{xy} and B_{lm}^{xy} are the in-plane normalized constants; κ_{lm} as well as γ_{lm} play similar roles as k_n and ν_n for in-plane wave functions; and J_m as well as K_m are the m_{th}-order Bessel function of the first kind and modified Bessel function of the second kind, respectively. The growth-direction normalized constants B_n^z and A_n^z as well as the in-plane ones A_{lm}^{xy} and B_{lm}^{xy} satisfy the following two equations due to the continuity of the wave function across the boundaries $z = z_d \pm h/2$ and $\rho = \rho_0$,

$$B_n^z = A_n^z e^{\nu_n h/2} \cos\left(k_n \frac{h}{2} - \frac{\theta_n}{2}\right), \quad (5)$$

$$B_{lm}^{xy} = \frac{J_m(\kappa_{lm}\rho_0)}{K_m(\gamma_{lm}\rho_0)} A_{lm}^{xy}. \quad (6)$$

For QW states, we model them as two-dimensional plane waves:

$$\phi_{\mathbf{k}}(\mathbf{r}) = \frac{e^{i\mathbf{k}\cdot\boldsymbol{\rho}}}{\sqrt{A}}\Phi_w(z), \quad (7)$$

where $\mathbf{k}$ is the wave vector of the plane wave, A is the area of the QW, and $\Phi_w(z)$ is the quantized wave function in the growth direction.

The explicit expression of the growth-direction wave function of QW is as follows:

$$\Phi_w(z) = \begin{cases} B_w^z e^{i\theta_w} e^{\nu_w z}, & \text{in region } \mathbf{1}, \\ \dfrac{A_w^z}{2}(e^{-i\theta_w/2}e^{ik_w z} + e^{i\theta_w/2}e^{-ik_w z}), & \text{in region } \mathbf{2}, \\ B_w^z e^{-\nu_w z}, & \text{in regions } \mathbf{3},\mathbf{4},\mathbf{5}; \end{cases} \quad (8)$$

$$\theta_w = \begin{cases} 0, \text{ if even parity for QW Hamiltonian in growth direction}, \\ \pi, \text{ if odd parity for QW Hamiltonian in growth direction}; \end{cases} \quad (9)$$

where B_w^z and A_w^z are the normalized constants of the growth-direction wave function, and ν_w and k_w are the parameters describing the penetration into barrier and standing-wave behavior in QW region. The normalized constants B_w^z and A_w^z satisfy the following equation due to the continuity of the wave function at the boundary $z=\pm d/2$:

$$B_w^z = A_w^z e^{\nu_w d/2} \cos\left(k_w \frac{d}{2} - \frac{\theta_w}{2}\right). \tag{10}$$

In principle, the tunneling from excited subbands in QW can also be considered. However, for simplicity, we will not consider this kind of process in this paper.

III. THEORY FOR QW TO QD PHONON-ASSISTED TUNNELING

Resonant (coherent) tunneling transfer (without phonon or Auger assistance) does occur when $E_w+\hbar^2k^2/2m_w^{xy*}=E_d$, or $E_d>E_w$. We refer to Ref. 20 for this fast coherent tunneling process into the excited state of the dot. In this paper, we are mainly interested in $E_d<E_w$ (the ground state of the dot), when the phonon- or Auger-assisted process may be important. For the 25 Å barrier, we have calculated the QD ground state wave function in the presence and absence of the QW, and found out they differ by only 2% in the overlap integral. For a thin barrier width of about 10 Å or less, the QW and QD may be strongly coupled. Improvements of the QW and QD wave functions will be required but are beyond the scope of this paper if we can keep these phonon- and Auger-assisted formulations tractable.

Fermi's golden rule based on these single-particle states are applied. For LO-phonon scattering, the tunneling rate, or the net capture rate of the QD state $|\bar{N},\sigma\rangle$, where σ is the spin of the state, is as follows:

$$W_{w\to\bar{N},\sigma}^{\text{phonon}} = \frac{2\pi}{\hbar}\sum_{\mathbf{k},\mathbf{p}} |M_{\bar{N},\mathbf{k}}^{\sigma}(\mathbf{p})|^2[(n_{LO}+1)f_k(1-f_{\bar{N}}) - n_{LO}(1 - f_k)f_{\bar{N}}]\delta\left(E_w + \frac{\hbar^2k^2}{2m_w^{xy*}} - E_{\bar{N}} - \hbar\omega_{LO}\right), \tag{11}$$

where $M_{\bar{N},\mathbf{k}}^{\sigma}(\mathbf{p})$ is the Fröhlich matrix element; $\mathbf{p}$ is the wave vector of the LO phonon; $n_{LO}=1/[\exp(\hbar\omega_{LO}/k_BT)-1]$ is the Bose-Einstein occupation number of the LO phonon; $f_{\mathbf{k}}$ and $f_{\bar{N}}$ are the occupation numbers of the QW state labeled by the wave vector $\mathbf{k}$ and QD state labeled by quantum number $\bar{N}$, respectively; and E_w is the subband edge energy of the QW, i.e., the energy of the QW state with a zero wave vector. The LO phonon is modeled as dispersionless. The spin indices have been neglected in the labeling of the occupation number of both QW and QD states since they are spin-independent in the current case

For LO-phonon scattering, the interaction matrix element is given by the Fröhlich Hamiltonian:

$$M_{\bar{N},\mathbf{k}}^{\sigma}(\mathbf{p}) = i\left(\frac{e^2\hbar\omega_{LO}}{2\varepsilon_0 V_{ol}}\right)^{1/2}\left[\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_s}\right]^{1/2}\frac{1}{p} < \bar{N}|e^{i\mathbf{p}\cdot\mathbf{r}}|\mathbf{k}>, \tag{12}$$

where V_{ol} is the volume of the bulk; $\mathbf{p}$ and ω_{LO} are the wave vector and angular frequency of an LO phonon; ϵ_∞ and ϵ_s are the high-frequency and static dielectric constants, respectively; and σ is the spin of both initial and final states. Because LO-phonon scattering does not change the spin of the carrier in the conduction band, we do not need to label both the spins of the initial and final states. Specification of the final spin state is enough.

Define a new vector $\mathbf{q}=\mathbf{p}+\mathbf{k}$. Because the wave vector $\mathbf{k}$ of the state $|\phi_{\mathbf{k}}\rangle$ only has the component in the QW plane, the growth-direction component of the wave vector $\mathbf{q}$ is the same as that of the wave vector $\mathbf{p}$:

$$q_z = p_z. \tag{13}$$

We can rewrite Eq. (12):

$$M_{\bar{N},\mathbf{k}}^{\sigma}(\mathbf{p}) \equiv \hat{M}_{\bar{N}}^{\sigma}(\mathbf{p}+\mathbf{k},\mathbf{k}) = i\left(\frac{e^2\hbar\omega_{LO}}{2\varepsilon_0 V_{ol}}\right)^{1/2}\left[\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_s}\right]^{1/2}\frac{1}{|\mathbf{q}-\mathbf{k}|}\int_{-\infty}^{\infty} dz\varphi_n^*(z)e^{iq_zz}\Phi_w(z)\int_0^\infty d\rho\rho\frac{R_{lm}^*(\rho)}{\sqrt{A}}\int_{-\pi}^{\pi} d\phi\frac{e^{-im\phi+iq_\perp\rho\cos(\phi_q-\phi)}}{\sqrt{2\pi}}$$
$$= i\left(\frac{e^2\hbar\omega_{LO}}{2\epsilon_0 V_{ol}}\right)^{1/2}\left[\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_s}\right]^{1/2}\frac{1}{|\mathbf{q}-\mathbf{k}|}\times X(q_z)\frac{e^{-im\phi_q}(\sqrt{2\pi}i^m)}{\sqrt{A}}\int_0^\infty d\rho\rho R_{lm}^*(\rho)J_m(q_\perp\rho), \tag{14}$$

where the function $X(q_z)$ is defined as

$$X(q_z) = \int_{-\infty}^{\infty} dz\varphi_n^*(z)e^{iq_zz}\Phi_w(z). \tag{15}$$

To simplify the expression above, we define two functions, $Y(q_z)$ and $U(q_\perp)$:

$$Y(q_z) = |X(q_z)|^2, \tag{16}$$

$$U(q_\perp) = \left|\int_0^\infty d\rho\rho R_{lm}^*(\rho)J_m(q_\perp\rho)\right|^2. \tag{17}$$

Also, define a variable Q due to the conservation of particle and phonon energy:

$$Q = \frac{\sqrt{2m_w^{xy*}(E_{\bar{N}} + \hbar\omega_{LO} - E_w)}}{\hbar}. \tag{18}$$

By substituting Eqs. (16)–(18) into Eq. (11), and utilizing the following identity of integration:

$$\int_{-\pi}^{\pi} \frac{1}{a - b\cos t} = \frac{2\pi}{\sqrt{a^2 - b^2}}, \tag{19}$$

where a and b are two real numbers and $|a| > |b|$, we can simplify Eq. (11) as follows:

$$W_{w\to\bar{N},\sigma}^{\text{phonon}} = \frac{m_w^{xy*}(2\pi)^2}{(2\pi\hbar)^3}\frac{e^2\hbar\omega_{LO}}{2\epsilon_0}\left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_s}\right) \times[(n_{LO}+1)f_Q(1-f_{\bar{N}}) - n_{LO}(1-f_Q)f_{\bar{N}}] \times \int_0^\infty dq \int_{-1}^{1} dt \frac{q^2 Y(qt)U(q\sqrt{1-t^2})}{\sqrt{(q+Q)^2(q-Q)^2 + (2Qqt)^2}}. \tag{20}$$

Equation (20) can then be evaluated numerically to give the tunneling rate. When $q=Q$ and $t=0$, the integrand contains a singularity. This singularity should be specifically treated in numerical simulations. Also, under the current model, the functions $Y(qt)$ and $U(q\sqrt{1-t^2})$ both have analytical expressions. We will discuss the derivation of these functions in Appendix A.

IV. THEORY OF QW TO QD AUGER-ASSISTED TUNNELING

As shown in Fig. 1(b), we consider the process in which two electrons 1 and 2 in the QW interact with each other through Coulomb interaction. One of them acquires enough energy to be excited to a high-energy state 1' in the QW while the other tunnels into the QD state $|\bar{N},\sigma\rangle$, which is denoted as 2' in Fig. 1(b). The net capture rate of the state $|\bar{N},\sigma\rangle$ caused by Auger-assisted tunneling is as follows:

$$W_{w\to\bar{N},\sigma}^{\text{Auger}} = \frac{2\pi}{\hbar}\sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_1'}(|M_{\mathbf{k}_1',\bar{N};\mathbf{k}_1,\mathbf{k}_2}^{\sigma,\sigma}|^2 + |M_{\mathbf{k}_1',\bar{N};\mathbf{k}_1,\mathbf{k}_2}^{\sigma,\bar{\sigma}}|^2) \times[(1-f_{\bar{N}})(1-f_{k_1'})f_{k_2}f_{k_1} - (1-f_{k_2})(1-f_{k_1})f_{\bar{N}}f_{k_1'}] \times \delta(E_{\bar{N}} + E_{k_1'} - E_{k_1} - E_{k_2}), \tag{21}$$

where $\bar{\sigma}$ is the opposite projection of σ; $M_{\mathbf{k}_1',\bar{N};\mathbf{k}_1,\mathbf{k}_2}^{\sigma,\sigma}$ is the Auger matrix element when the states labeled by the wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ both have spins σ; and $M_{\mathbf{k}_1',\bar{N};\mathbf{k}_1,\mathbf{k}_2}^{\sigma,\bar{\sigma}}$ is the one when they have opposites spins.

The initial state is a two-particle state $|\mathbf{k}_1,\mathbf{k}_2\rangle$ in the QW. The final state is composed of one particle in the QD state and the other in the high-energy state in the QW, $|\mathbf{k}_1',\bar{N}\rangle$. The *direct* matrix element is

$$M_{\mathbf{k}_1',\bar{N};\mathbf{k}_1,\mathbf{k}_2} = \frac{e^2}{4\pi\epsilon_s}\frac{1}{A^{3/2}}\int_{-\infty}^{\infty} dz_2 dz_1 \varphi_n^*(z_2)\Phi_w(z_2)\Phi_w^*(z_1)\Phi_w(z_1) \times \int d\boldsymbol{\rho}_2 d\boldsymbol{\rho}_1 \frac{\Psi_{lm}^*(\boldsymbol{\rho}_2)e^{-i\mathbf{k}_1'\cdot\boldsymbol{\rho}_1}e^{i\mathbf{k}_2\cdot\boldsymbol{\rho}_2}e^{i\mathbf{k}_1\cdot\boldsymbol{\rho}_1}}{\sqrt{|\boldsymbol{\rho}_2-\boldsymbol{\rho}_1|^2 + (z_2-z_1)^2}}. \tag{22}$$

Define the two vector variables for integration: the relative coordinate $\mathbf{x}=\rho_1-\rho_2$ as well as the center-of-mass coordinate $\mathbf{y}=(\boldsymbol{\rho}_1+\boldsymbol{\rho}_2)/2$. We also introduce the Fourier transformation of the in-plane wave function, namely, $\tilde{\Psi}_{lm}(\mathbf{p})$. Equation (22) is rewritten as

$$M_{\mathbf{k}_1',\bar{N};\mathbf{k}_1,\mathbf{k}_2} = \frac{e^2}{4\pi\epsilon_s}\frac{1}{A^{3/2}}\int_{-\infty}^{\infty} dz_2 dz_1 \varphi_n^*(z_2)\Phi_w(z_2)\Phi_w^*(z_1)\Phi_w(z_1) \times \int \frac{d\mathbf{p}}{(2\pi)^2}\tilde{\Psi}_{lm}^*(\mathbf{p}) \int d\mathbf{x}\frac{e^{i(\mathbf{k}_1-\mathbf{k}_1'-\mathbf{k}_2+\mathbf{p})\cdot\mathbf{x}/2}}{\sqrt{x^2+(z_1-z_2)^2}} \times \int d\mathbf{y} e^{i(\mathbf{k}_2-\mathbf{p}+\mathbf{k}_1-\mathbf{k}_1')\cdot\mathbf{y}}. \tag{23}$$

The integration over the center-of-mass coordinate results in a delta function. In this way, the integration over the variable $\mathbf{p}$ can be eliminated.

By using the following two identities of the zeroth-order modified Bessel function of the second kind $K_0(qx)$:

$$\frac{1}{\sqrt{x^2+(z_1-z_2)^2}} = \frac{2}{\pi}\int_0^\infty dq\, \cos[q(z_2-z_1)]K_0(qx), \tag{24}$$

$$\int d\mathbf{x} K_0(qx)e^{i(\mathbf{k}_1-\mathbf{k}_1')\cdot\mathbf{x}} = \frac{2\pi}{|\mathbf{k}_1-\mathbf{k}_1'|^2+q^2}, \tag{25}$$

we can rewrite the matrix element $M_{\mathbf{k}_1',\bar{N};\mathbf{k}_1,\mathbf{k}_2}$ as

$$M_{\mathbf{k}_1',\bar{N};\mathbf{k}_1,\mathbf{k}_2} = \frac{e^2\tilde{\Psi}_{lm}^*(\mathbf{k}_2+\mathbf{k}_1-\mathbf{k}_1')}{4\epsilon_s A^{3/2}}F(|\mathbf{k}_1-\mathbf{k}_1'|), \tag{26}$$

where the function $F(|\mathbf{k}_1-\mathbf{k}_1'|)$ is related to the form factors between QW states and between QW and QD states:

$$F(|\mathbf{k}_1-\mathbf{k}_1'|) \equiv \int_{-\infty}^{\infty} dq \frac{<\varphi_n|e^{iqz}|\Phi_w><\Phi_w|e^{-iqz}|\Phi_w>}{|\mathbf{k}_1-\mathbf{k}_1'|^2+q^2}. \tag{27}$$

The form factor between QW states will be evaluated in Appendix B.

For the *exchange* matrix element, we just switch $\mathbf{k}_1$ and $\mathbf{k}_2$ in the expression of the direct one. With these matrix elements, we can consider the scattering of two electrons with the same or opposite spins. If the electrons in the $\mathbf{k}_1$ and $\mathbf{k}_2$ states happen to have identical spin, say σ, the Auger matrix element is

$$M_{\mathbf{k}_1',\bar{N};\mathbf{k}_1,\mathbf{k}_2}^{\sigma,\sigma} = \frac{e^2\tilde{\Psi}_{lm}^*(\mathbf{k}_2+\mathbf{k}_1-\mathbf{k}_1')}{4\epsilon_s A^{3/2}} \times\{F(|\mathbf{k}_1-\mathbf{k}_1'|) - F(|\mathbf{k}_2-\mathbf{k}_1'|)\}. \tag{28}$$

On the other hand, if the spins of the two particles in the

initial states happen to be opposite, the matrix element is then

$$M^{\sigma,\bar{\sigma}}_{\mathbf{k}'_1,N;\mathbf{k}_1,\mathbf{k}_2} = \frac{e^2\tilde{\Psi}^*_{lm}(\mathbf{k}_2+\mathbf{k}_1-\mathbf{k}'_1)}{4\epsilon_s A^{3/2}}F(|\mathbf{k}_1-\mathbf{k}'_1|). \quad (29)$$

We then consider the total tunneling rate from QW states to QD state $|\bar{N},\sigma\rangle$. From Eq. (21), the net capture rate of the QD state $|\bar{N},\sigma\rangle$ is

$$W^{\text{Auger}}_{w\to\bar{N},\sigma} = \frac{2\pi}{\hbar}\frac{A^3}{(2\pi)^6}\int d\mathbf{k}'_1 d\mathbf{k}_1 d\mathbf{k}_2 (|M^{\sigma,\sigma}_{\mathbf{k}'_1,N;\mathbf{k}_1,\mathbf{k}_2}|^2 + |M^{\sigma,\bar{\sigma}}_{\mathbf{k}'_1,N;\mathbf{k}_1,\mathbf{k}_2}|^2)[(1-f_{\bar{N}})(1-f_{k'_1})f_{k_2}f_{k_1} - (1-f_{k_2})(1-f_{k_1})f_{\bar{N}}f_{k'_1}]\delta(E_{\bar{N}}+E_{k'_1}-E_{k_1}-E_{k_2}). \quad (30)$$

Equation (30) is the integral form of the Auger-assisted tunneling rate in this QW-QD system. The simplification and approximation of this six-dimensional integration will be presented in Appendix C. The final result of this net capture rate for each spin σ then can be written in a three-dimensional integration:

$$W^{\text{Auger}}_{w\to\bar{N},\sigma} \simeq \frac{3m^{xy*}_w}{2}\left(\frac{1}{4\pi\hbar}\right)^3\left(\frac{e^2}{\epsilon_s}\right)^2\int_0^\infty dk_1 k_1\int_{\hat{p}(k_1)}^\infty dk'_1 k'_1\int_{-\pi}^{\pi}d\hat{\phi} \times\{<|F(|\mathbf{k}_1-\mathbf{k}'_1|)|^2>|\tilde{\Psi}_{lm}(\mathbf{k}_1-\mathbf{k}'_1)|^2 \times[(1-f_{\bar{N}})(1-f_{k'_1})f_{\beta(k'_1,k_1)}f_{k_1} - (1-f_{\beta(k'_1,k_1)})(1-f_{k_1})f_{\bar{N}}f_{k'_1}]\}, \quad (31)$$

where $\beta(k'_1,k_1)$ and $\hat{p}(k_1)$ are functions resulting from energy conservation and will be defined in Appendix C.

V. THEORETICAL RESULTS FOR PHONON- AND AUGER-ASSISTED CAPTURE RATES

Equations (11) and (21) are aimed at the occupation of the QD state $|\bar{N},\sigma\rangle$. For the carriers in the QW, we are more interested in the average time constant for them to tunnel into the QD. Denote the surface density of QDs as N_D. The surface density of carriers in QW is given explicitly as follows:

$$n_w = \frac{2}{A}\sum_{\mathbf{k}} f_k = \frac{m^{xy*}_w k_B T}{\pi\hbar^2}\ln[1+e^{(E_{Fw}-E_w)/k_BT}], \quad (32)$$

where E_{Fw} is the quasi-Fermi level of the carriers in the QW. If the initial occupation of the QD state is zero, the conservation of particle numbers for these tunneling processes enables us to define the net capture rates $W^{\text{phonon}}_{w\to\bar{N}}$, $W^{\text{Auger}}_{w\to\bar{N}}$ as well as the average tunneling time constants $\tau^{\text{phonon}}_{\text{Av}}$, $\tau^{\text{Auger}}_{\text{Av}}$ of the two mechanisms and Auger coefficient C^{Auger}, which all depend on the surface carrier density in the QW:

$$W^{\text{phonon}}_{w\to\bar{N}} \equiv \sum_\sigma W^{\text{phonon}}_{w\to\bar{N},\sigma}, \quad (33)$$

$$W^{\text{Auger}}_{w\to\bar{N}} \equiv \sum_\sigma W^{\text{Auger}}_{w\to\bar{N},\sigma}, \quad (34)$$

$$\frac{n_w}{\tau^{\text{phonon}}_{\text{Av}}(n_w)} \equiv N_D W^{\text{phonon}}_{w\to\bar{N}}\Big|_{f_{\bar{N}}=0}, \quad (35)$$

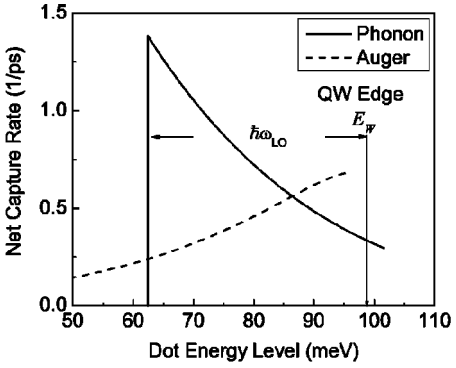
$$\frac{n_w}{\tau^{\text{Auger}}_{\text{Av}}(n_w)} \equiv N_D W^{\text{Auger}}_{w\to\bar{N}}\Big|_{f_{\bar{N}}=0} \equiv C^{\text{Auger}}(n_w)n_w^2. \quad (36)$$

In this paper, our definition of Auger coefficient includes the surface density of QDs and is different from that defined in the paper by Magnusdottir *et al.*[22]

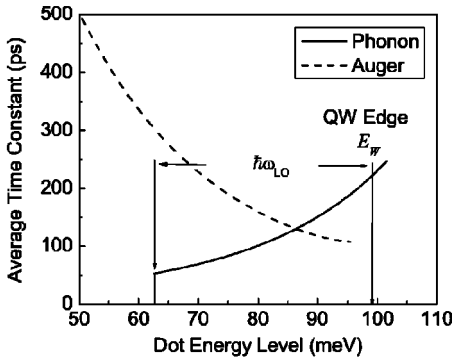
For the calculations of these two tunneling mechanisms, the following parameters for $In_xGa_{1-x}As$ QW and QD are used. Compositions and band offsets can be found in the experiments of Refs. 8 and 28. Here, the Indium composition of QD and QW are 0.175 and 0.202, respectively. In the barrier, QW, and QD regions, the electron effective masses are 0.067, 0.0525, and 0.0508 m_0, respectively, where m_0 is the free electron mass. The potential depths of the QW (V_w) and the QD (V_D) are 170.5 meV and 195.0 meV at room temperature, respectively. The first quantized energy of the QW state is 73.4 meV above the bottom of the QW potential. The width of the QW region and the height of the QD region are 50 and 100 Å, respectively. The radius of the QD is 100 Å, which will give rise to an energy difference of one phonon energy between the QW subband edge and QD state if there is no bias electric field. The LO phonon energy is 35.9 meV. The QD density is set as 10^{10} cm^{-2} unless otherwise mentioned. Only the ground state of the QD in conduction band will be considered. The occupation of this state will be assumed to be empty initially.

Figure 2(a) shows the net capture rate of the QD ground state as a function of the QD ground-state energy. The QD ground energy is shifted by the bias electric field and is referenced from the bottom of the unperturbed QD potential. Figures 2(b) and 2(c) are the corresponding average tunneling time constants $\tau^{\text{phonon}}_{\text{Av}}$, $\tau^{\text{Auger}}_{\text{Av}}$ and Auger coefficients. The quasi Fermi level in the QW is set at 25 meV above the QW subband edge energy. The temperature and the surface QW carrier density n_w are 300 K and 7.3×10^{11} cm^{-2}, respectively. The barrier width is set at 25 Å. Due to the quantized energy of the LO phonon, if the energy difference between the QW subband edge and QD state is larger than one LO phonon energy, there will be no *first-order* phonon-assisted tunneling process. The QD energy has to be high enough so that the difference between the subband edge energy of the QW and the QD energy is within one phonon energy. On the other hand, Auger-assisted tunneling does not have this restriction because the conservation of energy can easily be satisfied by the two-particle process.

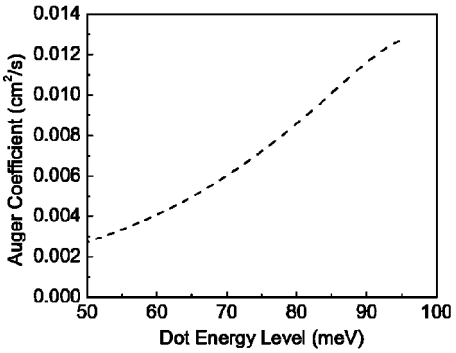
When the QD energy is low, the large difference between the QW subband edge and the QD state makes the effective momentum transfer between two carriers in the QW high. The Coulomb matrix element decreases as the transferred momentum increases. The net capture rate due to Auger-assisted tunneling thus becomes higher as the QD energy increases, which is also pointed out in the paper by Magnus-



(a)



(b)



(c)

FIG. 2. (a) The net capture rate due to phonon- and Auger-assisted processes, $W^{\text{phonon}}_{w\to\bar{N}}$ and $W^{\text{Auger}}_{w\to\bar{N}}$ using Eqs. (33) and (34), respectively, as a function of the QD energy level. (b) The average tunneling time constants, $\tau^{\text{phonon}}_{\text{Av}}(n_w)$ and $\tau^{\text{Auger}}_{\text{Av}}(n_w)$ using Eqs. (35) and (36), respectively. (c) The Auger coefficient $C^{\text{Auger}}(n_w)$. The parameters used are $d=50$ Å; $h=100$ Å; barrier width $=25$ Å; $\rho_0=100$ Å; $N_D=10^{10}$ cm^{-2}; $n_w=7.3\times10^{11}$ cm^{-2}; and $T=300$ K.

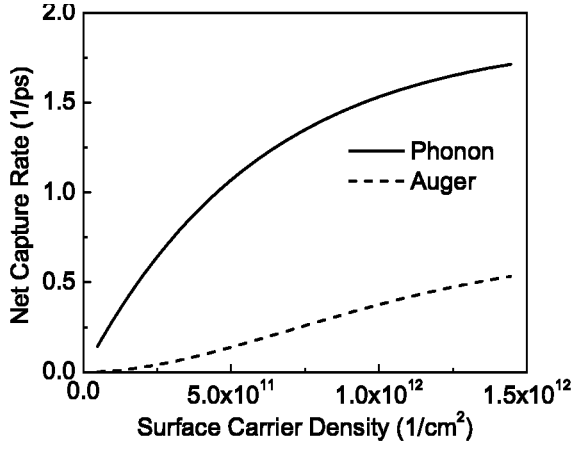
dottir *et al.*[22] On the other hand, for the phonon case, it has the maximum capture rate when the energy difference between the QW subband edge and the QD state is exactly one-phonon energy. The decrease of the net capture rate for the phonon-assisted process results from two contributions. First, the phonon-assisted process actually senses the carrier distribution in the QW. As the QD energy increases, it senses the carrier occupation of the high-energy states. The carrier distribution in the QW is modeled as a Fermi-Dirac distribution with a 25 meV-quasi-Fermi level above the QW subband edge energy, and thus reflects the decreasing trend. This, however, does not apply to the Auger process. It is always the low-energy carrier that is most probable to tunnel into the QD no matter how close are the QW subband edge energy and QD energy. Second, high-energy states in the QW have higher momentum and thus smaller form factor with the ground state of the QD. The smaller form factor causes the decrease of the matrix element. Both conditions result in the decrease of the net capture rate as the QD energy gets higher. In addition, there is a cross-over of the tunneling rates from two different mechanisms where the Auger-assisted process gradually takes over. This cross-over depends on the carrier density in the QW. For lower carrier density, cross-over will happen at a higher QD-energy level since the Auger-assisted tunneling will be weaker.

From Fig. 2(b), the average tunneling lifetime constants also reflect the trend from Fig. 2(a). The Auger assisted-process is usually not as efficient as the phonon-assisted process. If the energy difference is large or the carrier density is not high enough in the QW, the Auger process is usually too slow to be used for the injection of the carriers into the QD. Figure 2(c) shows the calculated Auger coefficient. Due to the presence of the barrier, the calculated Auger coefficient is much smaller than that calculated by Uskov *et al.*[24] because the relaxation from the wetting layer to the QD is faster without the barrier. Our calculation shows the same trend as that by Magnusdottir *et al.*[22] The increase of the Auger coefficient can be explained as the result of the decrease of the transferred momentum in the matrix element.

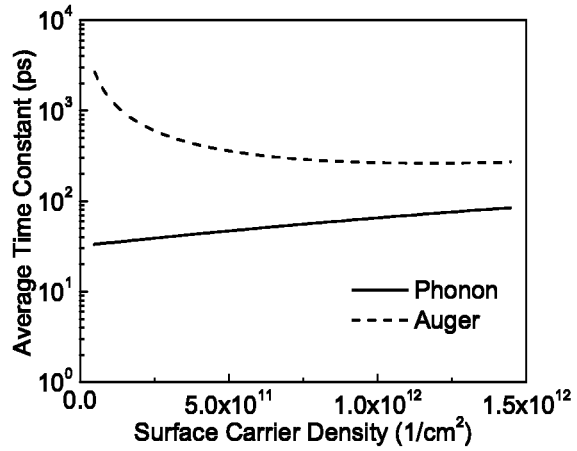
From Figs. 3–6, we consider the dependence of the tunneling rates on the QW surface carrier density. The bias electric field is set to zero. Figure 3(a) shows the net capture rate of the QD ground state as a function of the surface carrier density in the QW. Figures 3(b) and 3(c) are the corresponding average tunneling time constants and Auger coefficient. The barrier width is the same as that in the previous calculation (25 Å). As the QW carrier density increases, the phonon-assisted process gradually saturates due to the saturation of the state occupation participating in the process constrained by energy conservation. At low carrier density, the net capture rate of the Auger-assisted process shows quadratic dependence of the QW surface carrier density. However, at high carrier density, it also saturates because the filling of high-energy QW states prevents the carriers at low energy states from being excited, as predicted by Pauli's exclusion principle.

From Fig. 3(b), the average tunneling time constant for the phonon process gradually increases while that of the Auger assisted-process gradually stops decreasing. Because the number of QDs is fixed, only a limited number of carriers in the QW can tunnel into them. If the number of carriers in the QW keeps on increasing, more and more electrons will just accumulate in the QW without tunneling into QDs. Effectively, the average tunneling lifetime of the carriers in the QW will be longer at the high-density limit. The saturation at the high carrier-density limit for the Auger process due to Pauli's exclusion principle results in the decrease of Auger coefficients at high QW carrier density, as shown in Fig. 3(c).

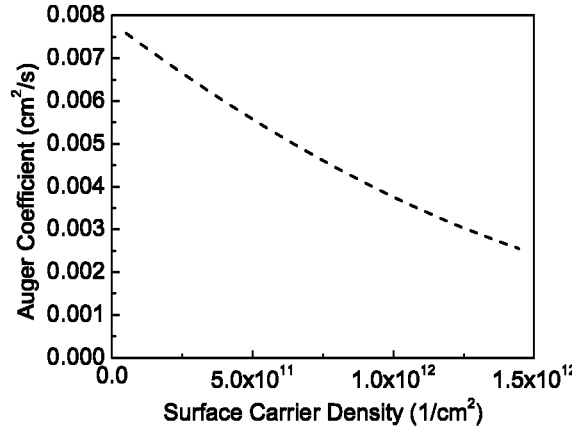
Figure 4(a) shows the phonon-assisted net capture rate of the QD ground state as a function of the surface carrier den-



(a)

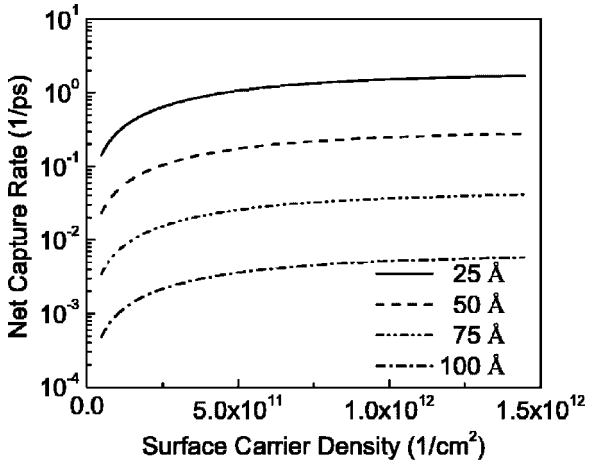


(b)

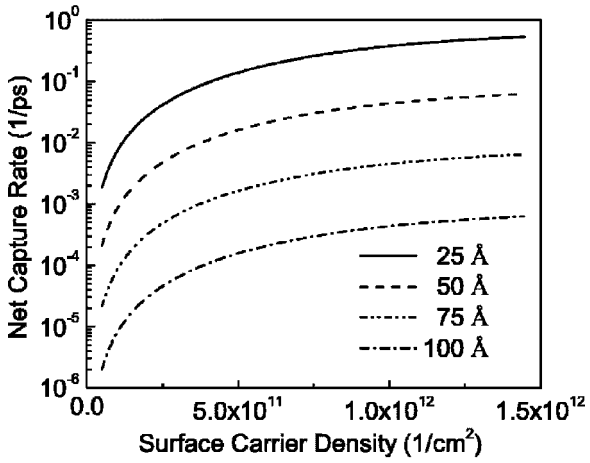


(c)

FIG. 3. (a) The net capture rate due to phonon- and Auger-assisted processes, $W^{\text{phonon}}_{w\to\bar{N}}$ and $W^{\text{Auger}}_{w\to\bar{N}}$ using Eqs. (33) and (34), as a function of QW surface carrier density. (b) The average tunneling time constants $\tau^{\text{phonon}}_{\text{Av}}(n_w)$ and $\tau^{\text{Auger}}_{\text{Av}}(n_w)$ using Eqs. (35) and (36), respectively, and (c) the Auger coefficient $C^{\text{Auger}}(n_w)$. The parameters used are $d=50$ Å; $h=100$ Å; barrier width=25 Å; $\rho_0=100$ Å; $E_w-E_{\bar{N}}\approx 35.9$ meV; $N_D=10^{10}$ cm^{-2}; and $T=300$ K.



(a)

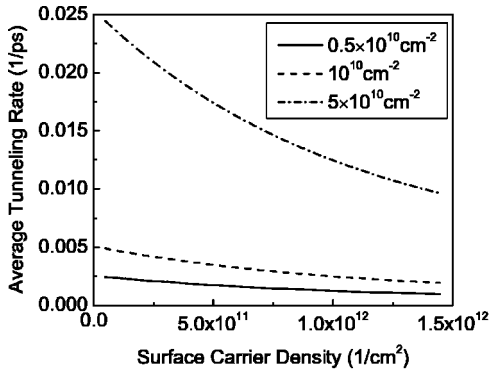


(b)

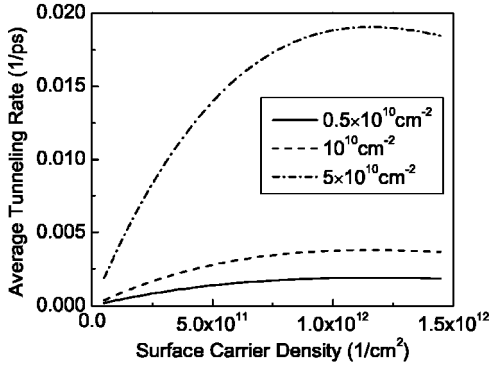
FIG. 4. (a) The net capture rate due to the phonon-assisted process $W^{\text{phonon}}_{w\to\bar{N}}$ using Eq. (33), as a function of QW surface carrier density for different barrier widths. (b) The net capture rate due to the Auger-assisted process, $W^{\text{Auger}}_{w\to\bar{N}}$ using Eq. (34), as a function of the QW surface carrier density for different barrier widths. The parameters used are $d=50$ Å; $h=100$ Å; $\rho_0=100$ Å; $E_w-E_{\bar{N}}\approx 35.9$ meV; $N_D=10^{10}$ cm^{-2}; and $T=300$ K.

sity in the QW for different barrier widths. Figure 4(b) shows the net capture rate due to Auger-assisted tunneling. All the parameters are the same as those used in the simulations for Fig. 3 except the barrier widths. The increase of barrier width decreases the overlap of the QW states and QD states, and thus the matrix elements in all the processes. For both phonon- and Auger-assisted processes, the net capture rates decrease as the barrier width increases.

Figures 5(a) and 5(b) show the average phonon- and Auger-assisted tunneling rates of the carriers in the QW for different surface QD densities. The inverse of these quantities are the average life constants of the QW carriers due to these two mechanisms. All the parameters are the same as those used in the simulations for Fig. 3 except the QD surface density. The QD surface densities are set at 5×10^9, 10^{10}, and 5×10^{10} cm^{-2}. If the number of the QDs is in-



(a)

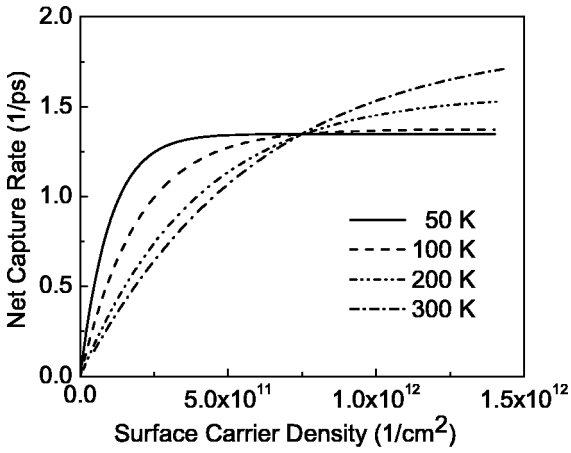


(b)

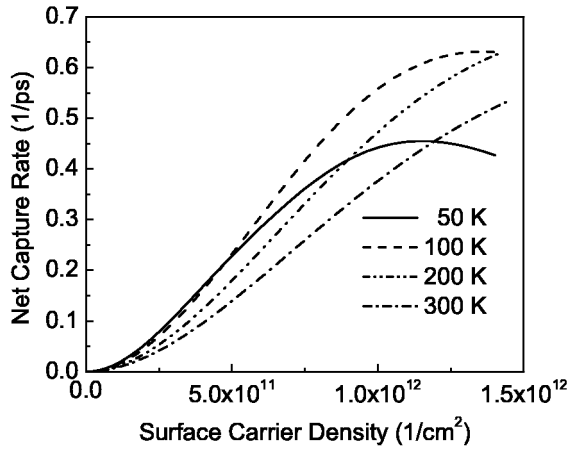
FIG. 5. (a) The average tunneling rate of the QW carriers due to the phonon-assisted process $1/\tau_{\mathrm{Av}}^{\mathrm{phonon}}(n_w)$ using Eq. (35), as a function of the QW surface carrier density for different QD densities. (b) The average tunneling rate of the QW carriers due to the Auger-assisted process $1/\tau_{\mathrm{Av}}^{\mathrm{Auger}}(n_w)$ using Eq. (36), as a function of the QW surface carrier density for different QD densities. The parameters used are $d=50$ Å; $h=100$ Å; barrier width=25 Å; $\rho_0=100$ Å; $E_w-E_{\bar{N}}\approx 35.9$ meV; $N_D=10^{10}$ cm^{-2}; and $T=300$ K.

creased, there will be more QD states for the QW carriers to tunnel into. Thus, the tunneling of the QW carriers will become faster.

Figure 6(a) shows the phonon-assisted net capture rate of the QD ground state as a function of the surface carrier density in the QW under different temperatures. Figure 6(b) shows the net capture rate due to Auger-assisted tunneling. The temperature will affect the carrier distribution due to the thermal Fermi-Dirac distribution. The energy difference between the QW subband edge energy and the energy of the QD ground state is still about one phonon energy. For phonon-assisted processes, under the condition of the same carrier density, carriers tend to occupy lower-energy states at low temperature. Therefore, at the low carrier-density limit, the dense carrier occupations at lower-energy states of QW guarantee that the net tunneling rate is higher than that at high temperature. However, as the density increases, the differences between these occupations at low temperature and room temperature become smaller. The emission process at



(a)



(b)

FIG. 6. (a) The net capture rate due to the phonon-assisted process $W_{w\to\bar{N}}^{\mathrm{phonon}}$ using Eq. (33), as a function of the QW surface carrier density for different temperatures. (b) The net capture rate due to the Auger-assisted process $W_{w\to\bar{N}}^{\mathrm{Auger}}$ using Eq. (34), as a function of the QW surface carrier density for different temperatures. The parameters used are $d=50$ Å; $h=100$ Å; barrier width=25 Å; $\rho_0=100$ Å; $E_w-E_{\bar{N}}\approx 35.9$ meV; and $N_D=10^{10}$ cm^{-2}.

high temperature is also more efficient at room temperature due to the spontaneous emission of an LO phonon. Over a critical density, the phonon-assisted capture rate at room temperature exceeds that at low temperature. For the Auger-assisted process, a similar argument of carrier occupations also applies when the QW density is low. However, when the carrier density is increased, the occupations of the higher-energy states below the quasi-Fermi level also become denser for the low-temperature case. From Pauli's exclusion principle, the occupations of high-energy states will not favor the Auger-assisted tunneling processes. Thus, the net capture rate due to Auger-assisted process at low temperature will saturate and decrease faster than that at higher temperature in the high-density limit.

Bhattacharya *et al.*[8] measured a 1.7-ps phonon-assisted capture time for the occupation of the conduction-band ground state. They designed the QW and QD sizes so that the QD ground state in the conduction band is about one phonon energy below the QW subband edge energy. Although their structure is not identical to ours, the corresponding net capture rates extracted from the experimental results, which correspond to one half of the quantity defined in Eq. (33), are within the same order of magnitude as our theoretical results.

VI. CONCLUSION

We have calculated the phonon-assisted and Auger-assisted tunneling rates. The typical capture times from phonon-assisted tunneling, depending on the barrier width between the QD and barrier, range from less than a picosecond to a few hundred picoseconds. For laser or detector applications, a thin barrier between the QD and QW can assist the efficient capture of the carriers into QD states. Under the low-density limit, the Auger-assisted process is usually much weaker than the phonon-assisted one. However, unlike phonon-assisted tunneling, it does not have the restriction on the energy difference between the QW subband edge energy and QD energy. At the high-density limit, the Auger-assisted tunneling will dominate the tunneling process.

ACKNOWLEDGMENTS

One of the authors (N. H.) wishes to thank the Sony Corporation for the support of the John Bardeen Chair. We thank the support of DARPA Grant No. AFSA3631-22549, DARPA No. MDA972-00-1-0020, and U.S. Army Research Office Grant No. DAAD19-01-1-0951.

APPENDIX A: FORM FACTOR BETWEEN QW AND QD STATES IN PHONON-ASSISTED TUNNELING

For the numerical integration in Eq. (20), the numerical values of two functions $Y(q_z)$ and $U(q_\perp)$ for arbitrary q have to be known first. From Eq. (16), we can divide the evaluation of the function $Y(q_z)$ into five regions:

$$Y(q_z) = \left| \int_{-\infty}^{\infty} dz \varphi_n^*(z) e^{iq_z z} \Phi_w(z) \right|^2 = \left| \sum_{I=1}^{5} \int_I dz \varphi_n^*(z) e^{iq_z z} \Phi_w(z) \right|^2, \tag{A1}$$

where regions 1–5 are defined in Fig. 1. Using the explicit expressions of the growth-direction component of the wave functions for QW and QD states, we can carry out the integrations in different regions explicitly as

$$\int_1 dz \varphi_n^*(z) e^{iq_z z} \Phi_w(z) = \frac{B_n^{z*} B_w^z}{\nu_w + \nu_n + iq_z} \times e^{i(\theta_w - \theta_n)} e^{-\nu_n z_d - (\nu_w + \nu_n) d/2} e^{-iq_z d/2},$$

$$\int_2 dz \varphi_n^*(z) e^{iq_z z} \Phi_w(z) = \frac{B_n^{z*} A_w^z}{2} e^{-i\theta_n} e^{-\nu_n z_d} \left\{ e^{-i\theta_w/2} \frac{e^{[i(q_z + k_w) + \nu_n] d/2} - e^{-[i(q_z + k_w) + \nu_n] d/2}}{i(q_z + k_w) + \nu_n} + e^{i\theta_w/2} \frac{e^{[i(q_z - k_w) + \nu_n] d/2} - e^{-[i(q_z - k_w) + \nu_n] d/2}}{i(q_z - k_w) + \nu_n} \right\},$$

$$\int_3 dz \varphi_n^*(z) e^{iq_z z} \Phi_w(z) = \frac{B_n^{z*} B_w^z}{2} e^{-i\theta_n} e^{-\nu_n z_d} \times \left\{ \frac{e^{[iq_z + (\nu_n - \nu_w)](z_d - h/2)} - e^{[iq_z + (\nu_n - \nu_w)] d/2}}{iq_z + (\nu_n - \nu_w)} \right\}$$

$$\int_4 dz \varphi_n^*(z) e^{iq_z z} \Phi_w(z) = \frac{A_n^{z*} B_w^z}{2} e^{(iq_z - \nu_n) z_d} \left\{ e^{i\theta_n/2} \frac{e^{[i(q_z - k_w) - \nu_n] h/2} - e^{-[i(q_z - k_w) - \nu_n] h/2}}{i(q_z - k_w) - \nu_n} + e^{-i\theta_n/2} \frac{e^{[i(q_z + k_n) - \nu_w] h/2} - e^{-[i(q_z + k_n) - \nu_w] h/2}}{i(q_z + k_n) - \nu_w} \right\},$$

$$\int_5 dz \varphi_n^*(z) e^{iq_z z} \Phi_w(z) = \frac{B_n^{z*} B_w^z}{\nu_w + \nu_n - iq_z} e^{\nu_n z_d} e^{[-(\nu_w + \nu_n) + iq_z](z_d + h/2)}, \tag{A2}$$

The function $Y(q_z)$ is then obtained by squaring the magnitude of the sum of the above equations.

For the function $U(q_\perp)$, from Eq. (4), we can get the explicit expressions of the analytical wave functions for this two-dimensional quantum disk model:

$$\int_0^\infty d\rho \rho R_{lm}^*(\rho) J_m(q_\perp \rho) = \frac{A_{lm}^{xy*} \rho_0}{\kappa_{lm}^2 - q_\perp^2} [q_\perp J_m(\kappa_{lm}\rho_0) J_{m-1}(q_\perp \rho_0) - \kappa_{lm} J_{m-1}(\kappa_{lm}\rho_0) J_m(q_\perp \rho_0)] + \frac{B_{lm}^{xy*} \rho_0}{\gamma_{lm}^2 + q_\perp^2} [q_\perp K_m(\gamma_{lm}\rho_0) J_{m-1}(q_\perp \rho_0) + \gamma_{lm} K_{m-1}(\gamma_{lm}\rho_0) J_m(q_\perp \rho_0)]. \tag{A3}$$

If the magnitude of the wave vector $q_\perp$ happens to be identical to κ_{lm}, the following alternative equation can be used:

$$\lim_{q_\perp \to \kappa_{lm}} \int_0^\infty d\rho \rho R_{lm}^*(\rho) J_m(q_\perp \rho) = \frac{A_{lm}^{xy*} \rho_0}{2\kappa_{lm}} \{-2mJ_{m-1}(\kappa_{lm}\rho_0) J_m(\kappa_{lm}\rho_0) + \kappa_{lm}\rho_0 [J_{m-1}^2(\kappa_{lm}\rho_0) + J_m^2(\kappa_{lm}\rho_0)]\}$$
$$+ \frac{B_{lm}^{xy*} \rho_0}{\gamma_{lm}^2 + \kappa_{lm}^2} [\kappa_{lm} K_m(\gamma_{lm}\rho_0) J_{m-1}(\kappa_{lm}\rho_0) + \gamma_{lm} K_{m-1}(\gamma_{lm}\rho_0) J_m(\kappa_{lm}\rho_0)]. \quad \text{(A4)}$$

Equation (A3) or (A4) is then substituted into Eq. (17) to give the numerical values of the function $U(q_\perp)$.

APPENDIX B: FORM FACTORS BETWEEN WELL STATES

The form factor of the growth-direction component of the QW states is relatively easy to carry out. In this paper, only one QW quantized band is considered. Similar to Eq. (A1), we divide the integration into different parts corresponding to regions 1–5:

$$\langle \Phi_w | e^{iqz} | \Phi_w \rangle = \sum_{I=1}^{5} \langle \Phi_w | e^{iqz} | \Phi_w \rangle_I. \quad \text{(B1)}$$

The individual integrations in different regions are as follows:

$$\langle \Phi_w | e^{iqz} | \Phi_w \rangle_{\mathbf{1}} = \frac{|B_w^z|^2}{2\nu_w + iq} e^{-(2\nu_w + iq)d/2},$$

$$\langle \Phi_w | e^{iqz} | \Phi_w \rangle_{\mathbf{2}} = \frac{|A_w^z|^2}{2} \left\{ e^{i\theta_w} \frac{\sin[(q-2k_w)d/2]}{q-2k_w} + \frac{2\sin(qd/2)}{q} + e^{-i\theta_w} \frac{\sin[(q+2k_w)d/2]}{q+2k_w} \right\},$$

$$\langle \Phi_w | e^{iqz} | \Phi_w \rangle_{\mathbf{3,4,5}} = \frac{|B_w^z|^2}{2\nu_w - iq} e^{-(2\nu_w - iq)d/2}. \quad \text{(B2)}$$

APPENDIX C: FORMULATION OF AUGER-ASSISTED TUNNELING

Consider the following matrix elements:

$$A^3 |M_{\mathbf{k}_1', N; \mathbf{k}_1, \mathbf{k}_2}^{\sigma, \sigma}|^2 = \left(\frac{e^2}{4\epsilon_s}\right)^2 |\tilde{\Psi}_{lm}^*(\mathbf{k}_2 + \mathbf{k}_1 - \mathbf{k}_1')|^2 \times \{|F(|\mathbf{k}_1 - \mathbf{k}_1'|)|^2 + |F(|\mathbf{k}_2 - \mathbf{k}_1'|)|^2 - 2\,\mathbf{Re}[F(|\mathbf{k}_1 - \mathbf{k}_1'|) * F(|\mathbf{k}_2 - \mathbf{k}_1'|)]\} \quad \text{(C1)}$$

$$A^3 |M_{\mathbf{k}_1', N; \mathbf{k}_1, \mathbf{k}_2}^{\sigma, \bar{\sigma}}|^2 = \left(\frac{e^2}{4\epsilon_s}\right)^2 |\tilde{\Psi}_{lm}^*(\mathbf{k}_2 + \mathbf{k}_1 - \mathbf{k}_1')|^2 |F(|\mathbf{k}_1 - \mathbf{k}_1'|)|^2. \quad \text{(C2)}$$

The matrix element with identical initial spins has an interference term. If both the Fourier transformation $\tilde{\Psi}_{\mathrm{lm}}^*(\mathbf{k}_2 + \mathbf{k}_1 - \mathbf{k}_1')$ and the function $F(|\mathbf{k}_1 - \mathbf{k}_1'|)$ have to be significant, the arguments of these functions have to be small. The interference term will be significant only when all the three wave vectors are small. However, energy conservation will not allow this if E_N^- is much below the subband edge energy of the QW. We can safely neglect the interference term. Equation (30) is rewritten as follows:

$$W_{w \to N, \sigma}^{\mathrm{Auger}} = \frac{3}{\hbar} \left(\frac{1}{2\pi}\right)^5 \left(\frac{e^2}{4\epsilon_s}\right)^2 \int d\mathbf{k}_1' d\mathbf{k}_2 d\mathbf{k}_1 \times \{|\tilde{\Psi}_{lm}(\mathbf{k}_2 + \mathbf{k}_1 - \mathbf{k}_1')|^2 |F(|\mathbf{k}_1 - \mathbf{k}_1'|)|^2 \times [(1 - f_{\bar{N}})(1 - f_{k_1'}) f_{k_2} f_{k_1} - (1 - f_{k_2})(1 - f_{k_1}) f_{\bar{N}} f_{k_1'}] \times \delta(E_N^- + E_{k_1'} - E_{k_1} - E_{k_2})\}. \quad \text{(C3)}$$

Equation (C3) is still too complicated to be used for a numerical analysis. To proceed, we will set $\mathbf{k}_2$ in the expression as zero since the relaxation of a QW electron with wave vector $\mathbf{k}_2$ into the QD state usually occurs near the subband edge of the QW, where $\mathbf{k}_2 \simeq \mathbf{0}$.

Define a function $\beta(k, k_1)$:

$$\beta(k_1', k_1) = \sqrt{\frac{k_1'^2 - k_1^2 - 2m_w^{xy*}(E_w - E_{\bar{N}})}{\hbar^2}}. \quad \text{(C4)}$$

In this way, the delta function due to energy conservation can be rewritten in terms of the magnitudes of the wave vectors k_1', k_1, and k_2:

$$\delta(E_N^- + E_{k_1'} - E_{k_1} - E_{k_2}) = \frac{m_w^{xy*}}{\hbar^2} \frac{\delta(k_2 - \beta(k_1', k_1))}{k_2}. \quad \text{(C5)}$$

With the aid of Eqs. (C4) and (C5) and that the integration over the azimuthal angle of the vector $\mathbf{k}_2$ only gives rise to a factor of 2π, the net capture rate is rewritten as

$$W^{\text{Auger}}_{w\to\overline{N},\sigma} = \frac{3m_w^{xy*}}{2\pi}\left(\frac{1}{2\pi\hbar}\right)^3\left(\frac{e^2}{4\epsilon_s}\right)^2 \oint d\phi_{k_1}\int_0^\infty dk_1 k_1 d\phi_{k_1'} \times \int_{\hat{p}(k_1)}^\infty dk_1' k_1' \{|\tilde{\Psi}_{lm}(\mathbf{k}_1-\mathbf{k}_1')|^2 |F(|\mathbf{k}_1-\mathbf{k}_1'|)|^2 \times [(1-f_{\overline{N}})(1-f_{k_1'}) f_{\beta(k_1',k_1)} f_{k_1} - (1-f_{\beta(k_1',k_1)})(1-f_{k_1}) f_{\overline{N}} f_{k_1'}]\}, \tag{C6}$$

where $\hat{p}(k_1)=\sqrt{k_1^2+(2m_w^{xy*}/\hbar^2)(E_w-E_{\overline{N}})}$ is the lower bound because $\beta(k_1',k_1)$ must be positive as indicated in Eq. (C4).

Let $\hat{\phi}$ be the angle between $\mathbf{k}_1$ and $\mathbf{k}_1'$. The integration over the azimuthal angle $\phi_{k_1'}$ is the same as that integrated over the angle $\hat{\phi}$ and can be approximated by the product of two integrations:

$$\oint d\phi_{k_1'} |\tilde{\Psi}_{lm}(\mathbf{k}_1-\mathbf{k}_1')|^2 |F(|\mathbf{k}_1-\mathbf{k}_1'|)|^2 \simeq \oint d\hat{\phi} |\tilde{\Psi}_{lm}(\mathbf{k}_1-\mathbf{k}_1')|^2 \frac{1}{2\pi}\oint d\hat{\phi} |F(|\mathbf{k}_1-\mathbf{k}_1'|)|^2 = \oint d\hat{\phi} |\tilde{\Psi}_{lm}(\mathbf{k}_1-\mathbf{k}_1')|^2 \langle |F(|\mathbf{k}_1-\mathbf{k}_1'|)|^2\rangle, \tag{C7}$$

where $\langle |F(|\mathbf{k}_1-\mathbf{k}_1'|)|^2\rangle$ is the angular average of the function $|F(|\mathbf{k}_1-\mathbf{k}_1'|)|^2$:

$$\langle |F(|\mathbf{k}_1-\mathbf{k}_1'|)|^2\rangle = \frac{1}{2\pi}\oint d\hat{\phi} |F(|\mathbf{k}_1-\mathbf{k}_1'|)|^2. \tag{C8}$$

This approximation will be valid if either the function $F(|\mathbf{k}_1-\mathbf{k}_1'|)$ or the Fourier transformation $\tilde{\Psi}_{\text{lm}}(\mathbf{k}_1-\mathbf{k}_1')$ varies slowly as the azimuthal angle $\phi_{k_1'}$ changes. We can manipulate the integration of the function $F(|\mathbf{k}_1-\mathbf{k}_1'|)$ explicitly:

$$\oint d\hat{\phi} |F(|\mathbf{k}_1-\mathbf{k}_1'|)|^2 = \int_{-\infty}^\infty dq_1 \langle\varphi_n|e^{iq_1 z}|\Phi_w\rangle\langle\Phi_w|e^{-iq_1 z}|\Phi_w\rangle \times \int_{-\infty}^\infty dq_2 \langle\Phi_w|e^{iq_2 z}|\Phi_w\rangle\langle\Phi_w|e^{-iq_2 z}|\varphi_n\rangle \times \oint d\hat{\phi} \frac{1}{|\mathbf{k}_1-\mathbf{k}_1'|^2+q_1^2}\frac{1}{|\mathbf{k}_1-\mathbf{k}_1'|^2+q_2^2}. \tag{C9}$$

In Eq. (C9), the integration over the angle $\hat{\phi}$ can be simplified as

$$\oint d\hat{\phi} \frac{1}{|\mathbf{k}_1-\mathbf{k}_1'|^2+q_1^2}\frac{1}{|\mathbf{k}_1-\mathbf{k}_1'|^2+q_2^2} = \frac{1}{q_1^2-q_2^2}\int_{-\pi}^{\pi} d\hat{\phi}\left[\frac{1}{q_2^2+k_1^2+k_1'^2-2k_1'k_1\cos(\hat{\phi})} - \frac{1}{q_1^2+k_1^2+k_1'^2-2k_1'k_1\cos(\hat{\phi})}\right]. \tag{C10}$$

Equation (C10) can be evaluated by using the same procedure as that for equation (19). The final result is

$$\oint d\hat{\phi} \frac{1}{|\mathbf{k}_1-\mathbf{k}_1'|^2+q_1^2}\frac{1}{|\mathbf{k}_1-\mathbf{k}_1'|^2+q_2^2} = 2\pi \frac{2(k_1^2+k_1'^2)+(q_2^2+q_1^2)}{\sqrt{[(k_1+k_1')^2+q_2^2][(k_1-k_1')^2+q_2^2]}+\sqrt{[(k_1+k_1')^2+q_1^2][(k_1-k_1')^2+q_1^2]}} \times \frac{1}{\sqrt{[(k_1+k_1')^2+q_2^2][(k_1-k_1')^2+q_2^2]}}\frac{1}{\sqrt{[(k_1+k_1')^2+q_1^2][(k_1-k_1')^2+q_1^2]}}. \tag{C11}$$

If we substitute Eq. (C11) into (C9), the integration for the Auger-assisted tunneling rate is still hard to carry out due to the dimension of integration.

There is a trick to further approximate Eq. (C11). Given a nonsingular function $f(x,y)$ around the origin, if the following conditions are valid:

$$\left.\frac{\partial f}{\partial x}\right|_{(x,y)=(0,0)} \neq 0, \tag{C12}$$

$$\left.\frac{\partial f}{\partial y}\right|_{(x,y)=(0,0)} \neq 0, \tag{C13}$$

then the following approximation will be correct up to the first order in Taylor's expansion:

$$f(x,y) \simeq \sqrt{f(2x,0)f(0,2y)}. \tag{C14}$$

Further, if $f(x,y)$ has an asymptotic value when any of its arguments tends to infinity, this approximation will result in the same asymptotic value as both the arguments tend to infinity.

Let $x=q_1^2$, $y=q_2^2$. We identify the function $f(q_1^2,q_2^2)$ as the nonseparable part of Eq. (C11):

$$f(q_1^2,q_2^2)=\frac{2(k_1^2+k_1'^2)+(q_2^2+q_1^2)}{\sqrt{[(k_1+k_1')^2+q_2^2][(k_1-k_1')^2+q_2^2]}+\sqrt{[(k_1+k_1')^2+q_1^2][(k_1-k_1')^2+q_1^2]}}. \tag{C15}$$

With the aid of Eq. (C14), the nonseparable part can be approximated as

$$f(q_1^2,q_2^2)\simeq\sqrt{\frac{2(k_1^2+k_1'^2+q_1^2)}{|k_1^2-k_1'^2|+\sqrt{[(k_1+k_1')^2+2q_1^2][(k_1-k_1')^2+2q_1^2]}}}\sqrt{\frac{2(k_1^2+k_1'^2+q_2^2)}{|k_1^2-k_1'^2|+\sqrt{[(k_1+k_1')^2+2q_2^2][(k_1-k_1')^2+2q_2^2]}}}. \tag{C16}$$

The angular average of the function $|F(|\mathbf{k}_1-\mathbf{k}_1'|)|^2$ is

$$\langle|F(|\mathbf{k}_1-\mathbf{k}_1'|)|^2\rangle\simeq\left|\int_{-\infty}^{\infty}dq\langle\varphi_n|e^{iqz}|\Phi_w\rangle\langle\Phi_w|e^{-iqz}|\Phi_w\rangle\sqrt{\frac{2(k_1^2+k_1'^2+q^2)}{|k_1^2-k_1'^2|+\sqrt{[(k_1+k_1')^2+2q^2][(k_1-k_1')^2+2q^2]}}}\times\frac{1}{\sqrt{[(k_1+k_1')^2+q^2][(k_1-k_1')^2+q^2]}}\right|^2. \tag{C17}$$

In Eq. (C6), the integration over the azimuthal angle ϕ_{k_1} will only result in a factor of 2π. We substituted Eq. (C17) into (C6) and thus obtain Eq. (31).

*Electronic address: s-chuang@uiuc.edu

[1] N. Holonyak, Jr., I. A. Lesk, R. N. Hall, J. J. Tiemann, and H. Ehrenreich, Phys. Rev. Lett. **3**, 167 (1959).
[2] L. Esaki and Y. Miyahara, Solid-State Electron. **1**, 13 (1960).
[3] L. Esaki, L. L. Chang, P. J. Stiles, D. F. O'Kane, and N. Wiser, Phys. Rev. **167**, 637 (1968).
[4] C. J. Adkins and W. A. Phillips, J. Phys. C **18**, 1313 (1985).
[5] M. Friesen, P. Rugheimer, D. E. Savage, M. G. Lagally, D. W. van der Weide, R. Joynt, and M. A. Eriksson, Phys. Rev. B **67**, 121301 (2003).
[6] X.-Q. Li and Y. Yan, Phys. Rev. B **65**, 205301 (2002).
[7] A. Imamoglu, D. D. Awschalom, G. Burkard, D. P. DiVincenzo, D. Loss, M. Sherwin, and A. Small, Phys. Rev. Lett. **83**, 4204 (1999).
[8] P. Bhattacharya and S. Ghosh, Appl. Phys. Lett. **80**, 3482 (2002).
[9] L. Harris, D. J. Mowbray, M. S. Skolnick, M. Hopkinson, and G. Hill, Appl. Phys. Lett. **73**, 969 (1998).
[10] H. Shoji, Y. Nakata, K. Mukai, Y. Sugiyama, M. Sugawara, N. Yokoyama, and H. Ishikawa, Appl. Phys. Lett. **71**, 193 (1997).
[11] D. Bimberg, M. Grundmann, and N. N. Ledentsov, *Quantum Dot Heterostructures*, 1st ed. (Wiley, New York, 1999).
[12] M. Sugawara, J. C. Bean, and X. Willardson, *Self-Assembled* InGaAs/GaAs *Quantum Dots*, 1st ed. (Academic, New York, 1999).
[13] R. Heitz, H. Born, F. Guffarth, O. Stier, A. Schliwa, A. Hoffmann, and D. Bimberg, Phys. Rev. B **64**, 241305 (2001).
[14] J. Urayama, T. B. Norris, J. Singh, and P. Bhattacharya, Phys. Rev. Lett. **86**, 4930 (2001).
[15] K. Mukai, N. Ohtsuka, H. Shoji, and M. Sugawara, Phys. Rev. B **54**, R5243 (1996).
[16] G. Walter, N. Holonyak, Jr., J. H. Ryou, and R. D. Dupuis, Appl. Phys. Lett. **79**, 1956 (2001).
[17] G. Walter, N. Holonyak, Jr., J. H. Ryou, and R. D. Dupuis, Appl. Phys. Lett. **79**, 3215 (2001).
[18] T. Chung, G. Walter, and N. Holonyak, Jr., Appl. Phys. Lett. **79**, 4500 (2001).
[19] G. Walter, T. Chung, and N. Holonyak, Jr., Appl. Phys. Lett. **80**, 1126 (2002).
[20] S. L. Chuang and N. Holonyak, Jr., Appl. Phys. Lett. **80**, 1270 (2002).
[21] S. L. Chuang, P. Littlewood, G. Walter, and N. Holonyak, in *Conference on Lasers and Electro-Optics/Quantum Electronics and Laser Science Conference*, Baltimore Convention Center, Baltimore, 2003.
[22] I. Magnusdottir, S. Bischoff, A. V. Uskov, and J. Mørk, Phys. Rev. B **67**, 205326 (2003).
[23] I. Magnusdottir, A. V. Uskov, S. Bischoff, B. Tromborg, and J. Mørk, J. Appl. Phys. **92**, 5982 (2002).
[24] A. V. Uskov, J. McInerney, F. Adler, H. Schweizer, and M. H. Pilkuhn, Appl. Phys. Lett. **72**, 58 (1998).
[25] A. V. Uskov, F. Adler, H. Schweizer, and M. H. Pilkuhn, J. Appl. Phys. **81**, 7895 (1997).
[26] R. Ferreira and G. Bastard, Appl. Phys. Lett. **74**, 2818 (1999).
[27] U. Bockelmann and G. Bastard, Phys. Rev. B **42**, 8947 (1990).
[28] G. Walter, T. Chung, and N. Holonyak, Jr., Appl. Phys. Lett. **80**, 3045 (2002).