Gain and recombination dynamics in photodetectors made with quantum nanostructures: The quantum dot in a well and the quantum well

B. Movaghar, S. Tsao, S. Abdollahi Pour, T. Yamanaka, and M. Razeghi
Center for Quantum Devices, Electrical Engineering, and Computer Science, Northwestern University, Evanston, Illinois 60208, USA
(Received 1 May 2008; revised manuscript received 30 July 2008; published 23 September 2008)

We consider the problem of charge transport and recombination in semiconductor quantum well infrared photodetectors and quantum-dot-in-a-well infrared detectors. The photoexcited carrier relaxation is calculated using rigorous random-walk and diffusion methods, which take into account the finiteness of recombination cross sections, and if necessary the memory of the carrier generation point. In the present application, bias fields are high and it is sufficient to consider the drift limited regime. The photoconductive gain is discussed in a quantum-mechanical language, making it more transparent, especially with regard to understanding the bias and temperature dependence. Comparing experiment and theory, we can estimate the respective recombination times. The method developed here applies equally well to nanopillar structures, provided account is taken of changes in mobility and trapping. Finally, we also derive formulas for the photocurrent time decays, which in a clean system at high bias are sums of two exponentials.

DOI: 10.1103/PhysRevB.78.115320 PACS number(s): 85.35.Be, 07.57.Kp, 73.21.La

I. INTRODUCTION

In a previous paper, we considered the problem of gain in quantum-dot infrared photodetector (QDIP) devices.[1] In this paper, we continue the analysis and consider a well established quantum well infrared photodetector (QWIP) (Ref. 2) structure and a related quantum dot/quantum well infrared photodetector (QDWIP). The two device structures are shown in Fig. 1. The QDWIP detectors are based on $InAs/In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ structures and differ strongly from both typical QDIPs and dot in a well detectors because the band offset between the InAs quantum dot (QD) and the InGaAs quantum well (QW) is only about 100 meV giving rise to one to two shallow bound states in the QD. These devices are a hybrid between the traditional QWIP and QDIP technologies and their performance is discussed in detail in Ref. 3. The InAs/InGaAs/InAlAs type QDWIPs with InP injection contacts have been reported in Ref. 3. QWIPs with the same material but with an InGaAs injection contact have been reported recently in Ref. 4. QD in a well devices with deeper bound states using $In_{0.15}Ga_{0.85}As$ have been considered, for example, in Refs. 5 and 6.

In this paper, we present a more general and a more quantum-mechanical approach to the problem of gain than what is typical in the literature. The dynamics of recombination in QWIPs (Refs. 7–11) are treated using a rate equation approach. The theory derived here is equally applicable to *e*-beam fabricated nanopillars of semiconductors,[12] superlattices of type I and type II materials, and nanowires. The traditional QDIP class of device requires a different methodology, as shown in Ref. 1.

Because they are based on rigorous random-walk theory, the methods we develop here and in a previous paper dealing with QDIPs (Ref. 1) are well suited to nanostructures in general. For the purpose of understanding gain in QW physics, one can think in terms of individual wells and barrier regions that are connected to each other. The total system is then built up from individual building blocks, which although interacting, preserves some of their own identity. Starting from each unit, one can see how the coupling gives rise to the overall properties of the system. The states in the plane of the QW can be treated as effective-mass plane wavelike, and superlattices can then be treated like a quasi-one-dimensional Kronig-Penney model. The same holds true for QWIP nanopillars[12] and QD layers buried in QW layers.[3]

We begin in Sec. II by defining the basic quantities that enter the analysis of the photoconductivity and gain. Then in

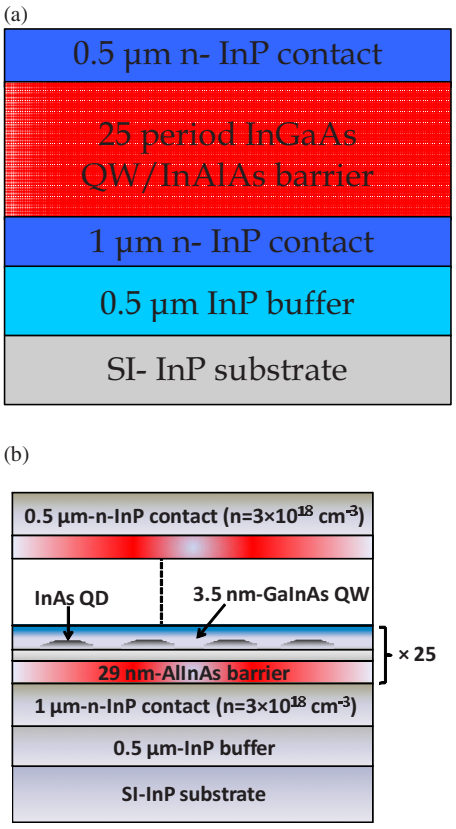

FIG. 1. (Color online) Schematic structures of (a) the QWIP and (b) the QDWIP.

Sec. III, we consider the problem of recombination using discretized diffusion theory (rate equations). In Sec. IV, we specialize to the drift regime, which is the most common situation in experiment. Here we touch upon the usual semiclassical results for the gain. The effect of traps is included in Sec. V. In Sec. VI we compare our theory with experiment and discuss the limitations of the generation-recombination (GR) noise formulation. Finally in Sec. VII, we consider the photocurrent and time dependence of the photocurrent decay. The summary and conclusions are presented in Sec. VIII.

The results can be summarized as follows: (1) A simple formula for the photoconductive gain as a function of bias and temperature, which includes trapping and velocity saturation in the drift limited regime; (2) The time decay of the pulsed photoconductivity in QWIP and QDWIPs; (3) Comparison with experiment and the conclusion/observation that the noise in the saturated velocity regime is most likely due to hot carrier phonon emission and intervalley crossing rather than generation-recombination noise. Although Monte Carlo work[13,14] has clearly demonstrated that these mechanisms are the cause of velocity saturation, the corresponding noise caused by time-dependent space-charge effects has not been evaluated. We suggest that this mechanism dominates the noise at high bias. Using only the generation-recombination formula leads to absurdly high values of gain in these devices.

II. BASIC QUANTITIES

Let us start with some definitions. The trap limited band drift velocity v_d can be written in a form that includes both the hot carrier saturation at high fields and the trapping,[1]

$$v_d = \frac{\mu_o F}{1 + \sum_l \frac{W_{bl}(1-n_l)}{W_{lb}}} \cdot \frac{1}{\left\{1 + \left(\frac{\mu_0 F}{v_s}\right)^2\right\}^{1/2}}, \tag{1}$$

where μ_0 is the pure ordered band mobility, F the electric field, W_{bl} is the trapping rate from a free band state "b" to a trap state "l," W_{lb} is the detrapping rate from l to b, v_s is the saturation velocity, and n_l is the occupation probability of the trap state l. The gain g is here defined as the ratio of recombination time to transit time,

$$g = \frac{v_d}{LC_{be}}, \tag{2}$$

where L is the device length and C_{be} is the inverse carrier recombination time from a free band state b to a capturing level "e." When the carrier has a problem getting into the device because there is a high injection barrier, then the system will have a low gain. The injection barrier or other dark current barriers serve to lengthen the transit time. The band mobility in Eq. (1) on the other hand includes only the trapping processes in the barrier layers (conduction band and electron transport only). Here it is assumed that the injection time is negligible. If this is not the case, then it should be included as an additional resistance. The responsivity R is normally written as

$$R = \frac{e}{\hbar\omega}\eta g, \tag{3}$$

where η is the internal quantum efficiency (QE), ω is the light frequency, and e is the magnitude of the charge.

The internal QE is itself proportional to the absorbance αL or, more generally, $\{1-\exp[-\alpha L]\}$ multiplied by the escape probability out of the photoexcited state into the continuum band,

$$\eta = \frac{W_{eb}}{W_{eb} + W_{eg}}\{1 - \exp[-\alpha L]\}, \tag{4}$$

where W_{eb} is the total escape rate out of the excited state into the continuum and includes all the pathways out,[1] and W_{eg} is the recombination rate back down into the lower QD levels. This process may be a multistep process. For the temperature and field dependence of R, the proximity of the extended state to the continuum is critical.[3,15]

Let us now consider the carrier dynamics from a time-dependent point of view and consider the noise in a photodetector. When the noise in the structure is dominated by GR noise,[16] then the measured noise current I_n is related to the gain g via the dark current I_D through the equation,

$$I_n = \sqrt{4eI_D g \Delta\nu}, \tag{5}$$

where e is electron charge and $\Delta\nu$ is the frequency band.[16] In a real QDIP system, for example, that in Ref. 1, one normally finds that the measured noise is more complex. At very low bias, the noise is frequency dependent and behaves as $1/f$ noise. Then as we go up in frequency, and depending on the quality of the QDIP layers, the noise behavior becomes constant in frequency, which is a sign that we have reached the GR white noise situation. At very high bias, the carrier velocity saturates due to optic phonon emission and intervalley crossing.[13,14] For the AlInAs barriers, the Γ to L and X energies are roughly 0.13 and 0.2 eV, respectively. In this limit, we expect the GR noise to disappear and the intervalley scattering induced space-charge noise and optic phonon emission hot carrier noise (an extension of Johnson noise)[17] to take over. The experimental results indicate that there is only a small range of biases over which we can talk about GR noise dominating. The problem at high bias was pointed out by Levine.[8] One of the objectives of this paper is to identify and model the GR noise gain region. To do this, let us start by opening with the problem of carrier recombination in a QWIP-like structure.

III. CALCULATION CONCEPT IN THE QWIP OR QWIP NANOPILLARS

The equation guiding the dynamics of charge carriers in a semiconductor is the usual diffusion recombination equation,[18–20]

$$\frac{\delta n(r,t)}{\delta t} = G(r,t) + D\nabla^2 n(r,t) + \frac{\delta n(r,t)}{\delta x}\mu_b F_x - \sum_i n(r,t)V(r-r_i), \quad (6)$$

where $G(r,t)$ is the local carrier generation rate, D is the diffusivity, $V(r)$ is the capture potential, and $n(r,t)$ is the carrier density at point r and time t. The diffusion recombination equation in a discrete form suitable for a superlattice is

$$\frac{dn_i(t)}{dt} = G_i - \sum_l n_i(t)W_{\mathrm{il}} + \sum_l n_l(t)W_{\mathrm{li}} - \sum_l \delta_{li}\varepsilon_l n_l(t), \quad (7)$$

where ε_l are the recombination rates at site l, W is the diffusion rate, and n_l is the occupation density at site l.

In a multilayer QWIP, semiconductors such as GaAs/AlAs or InGaAs/InP, for example, as discussed in Ref. 8, one has to deal with a quasi-one-dimensional diffusion and recombination problem. One can identify the "diffusion sites" with finite extended "regions" so that, for example, the site "A" is the quantum well region with length a and site "B" is the barrier layer with length b, as follows:

$$b = ma. \quad (8)$$

As a special case m can be an integer. The $m=1$ limit is the simple A-B-A-B… one-dimensional (1D) alloy problem in the analogous quantum-mechanical problem. So one can use the same methodologies and solve the analogous tight-binding model.

An "A" layer (QW) has a ground state "g" and an excited state e, which may or may not be in the region of the free band state, see Fig. 1(a). Let us assume first that the state e is a free (band) continuum state. Then it follows that the barrier layer B only has free band states (no recombination) and possible trapping and detrapping centers within the ensuing band trajectory. From the state e in the well region A, the carrier can only recombine into the ground state g, or transfer into the bandlike state in the barrier region B with rate W_{AB}.

In the other case, the excited level e is also bound. Then it follows that the carrier in e can recombine into the ground state or escape into the barrier as before, but now it can also excite up vertically into the continuum.

The carrier dynamics in a QWIP nanopillar obey essentially the same dynamic except that the $k\|$ states are now the discrete box eigenstates. The real difference is the mobility along the pillar μ_p, which is likely to be subject to many more surface traps than the mobility in a bulk QWIP μ_b. We can take this into account using the trapping function as discussed later.

The situation of the QD in QW can in most cases of interest be treated like a QWIP because when the carrier falls into the QW, it falls first into the QW excited state. It does not really matter how far the excited-state trap is from an actual QD recombination level. This is because the time to escape from the QW is usually much longer than the time it takes to relax deeper into the well and find a deeper state. Capture into excited states from which the carrier can escape again before recombining, and which are therefore number conserving, can be treated as traps and can be included in the mobility, as will be shown later in Eq. (27). Trapping levels lead to an overall increase in the band transit time, which may be bias and temperature dependent as in Eq. (27). Thus the QD in QW, the QWIP, and the QWIP-nanopillar problems differ by the value of the recombination parameters, the density-of-states structure, and the trapping delay function. There could be of course also other effects such as the strength of the electron-phonon couplings or polaron energy shifts, but these are secondary effects in this theory and are subject to a different line of investigation.

Let us now consider a superlattice with A (QW) and B (barrier) regions. If the excited level in A is free and continuously joins the band in the barrier region B, then we can put $W_{AB}=W_{BA}=W$, in field direction $W^+=(1+\eta_F)W$, and in opposite direction $W^-=(1-\eta_F)W$; with $\eta_F=eaF/2k_bT$. One can connect the domain transfer rate W to the usual diffusivity via $Wa^2=W_i a_i^2=D$, where a_i is the atomic lattice length, a is the width of the region A, and W_i is the corresponding site to site random-walk jump rate. The W can be related to the band mobility as shown below. The capture process, and then recombination, is described by the combined rate ε_A.

The occupation probability of the carriers $n_A(t)$ or $n_B(t)$ obey the standard rate equation in Eq. (7). Assuming the carrier started at "i" at time $t=0$, then the probability of finding it on "j" at time t is given by $n_{ij}(t)$ or $n(k,t)$ in the Fourier space.

In p-Laplace space we have, given that the particle started on a domain A or B, the survival fraction $n_{A,B}(p,k;p=0;k=0)$. We solve the A-B one-dimensional lattice problem for the situation $m=1$ so that $a=b$, for the special case that the length of the QW region is the same as the length of the barrier region. The problem is analogous to the 1D tight-binding $m^*=0.085m_0$ alloy, where W plays the role of the orbital overlap and ε_l the local orbital energies.

We find, assuming that the carrier started at a site B at $t=0$, the exact solution,

$$n^B_{\vec{k}=0}(p) = \frac{p+\varepsilon_A+W^++W^-}{(p+\varepsilon_A+W^++W^-)(p+\varepsilon_B+W^++W^-)-(W^++W^-)^2}, \quad (9)$$

$$n^A_{\vec{k}=0,B} = n^B_{\vec{k}=0}\left\{\frac{W^++W^-}{p+W^++W^-+\varepsilon_A}\right\}. \quad (10)$$

The survival fraction in the Laplace p-space, which will give us the time dependence, is the sum of the two terms,

$$n_B(p) = n^A_{\vec{k}=0,B} + n^B_{\vec{k}=0}, \quad (11)$$

$$n_B(p) = \frac{p+\varepsilon_A+2(W^++W^-)}{(p+\varepsilon_A+W^++W^-)(p+\varepsilon_B+W^++W^-)-(W^++W^-)^2}, \quad (12)$$

$$n_B(p=0) = \tau_B = \frac{\varepsilon_A + 2(W^+ + W^-)}{(\varepsilon_A + W^+ + W^-)(\varepsilon_B + W^+ + W^-) - (W^+ + W^-)^2}. \tag{13}$$

The solution for A is similarly obtained. The $p=0$ limit in Eq. (13) gives us the average recombination time τ_B, which is what we need to evaluate the gain. Unfortunately, the exact result is not easily generalized analytically to the situation of arbitrary b with $b=ma$ and electric bias. The exact result for $m=2$ is given in the Appendix.

IV. SOLUTIONS IN THE PURE DRIFT REGIME

A. Pure drift exact results for the probability

In the limit of pure drift, the carrier never goes back and has no memory of previous sites. The result is simple, and the recombination time exactly solvable for any quasi-one-dimensional situation, even with arbitrary disorder of both recombination and transfer rates in every well. The reason is that if $n_{ij}(p)$ is defined as the time Laplace or Fourier transform of the probability of going from i to j having started at $t=0$ at i, $n_{ij}(p)$ has no back diffusion terms in pure drift, i.e., $W^-=0$, and is a simple product,[18,19]

$$n_{1j}(p) = \frac{\delta_{1j}}{p + \varepsilon_1 + W_{12}^+} + \frac{1}{p + \varepsilon_1 + W_{12}^+} W_{12}^+ \frac{1}{p + \varepsilon_2 + W_{23}^+} + \ldots \tag{14}$$

If there is no recombination at all on the B sites, $\varepsilon_B=0$, and the drift velocity is given by $v_D=W^+a$, we have (pure drift) from Eq. (13),

$$n_B(k=0,p=0) = \tau_B = \frac{\varepsilon_A + 2(W^+)}{W^+\varepsilon_A} = \frac{2}{\varepsilon_A} + \frac{1}{W^+}. \tag{15}$$

On the other hand, if the particle starts at A, then we have the exact result,

$$n_A(p) = G^B_{\vec{k}=0,A} + G^A_{\vec{k}=0}, \tag{16}$$

$$n_A(p=0) = \tau_A = \frac{\varepsilon_B + 2(W^+ + W^-)}{(\varepsilon_A + W^+ + W^-)(\varepsilon_B + W^+ + W^-) - (W^+ + W^-)^2}, \tag{17}$$

$$n_A(p=0) = \tau_A = \frac{2}{\varepsilon_A}. \tag{18}$$

In other words, if the particle starts at A, it does not need to move to recombine, but from B, it clearly has to move at least one step.

We now allow the barrier layer to be thicker than the well region. The barrier thickness $b=ma$ is the interval of sites from A to A (one well to the next well). Now only $1/(m+1)$ of all sites are A sites, i.e., recombination sites, then the result of Eq. (18) is easily generalized and becomes

$$n_A(p=0,x) = \frac{1}{\tau_A} = \frac{m+1}{\varepsilon_A}, \tag{19}$$

when starting at A, and

$$n_B(p=0,b=ma) = \tau_B = \frac{m+1}{\varepsilon_A} + \frac{m}{W^+}, \tag{20}$$

when starting at B. The first term is the recombination term and is the inverse of the fraction of time the carrier is occupying a trap, multiplied by the trapping rate. The second term is the time to reach the first trap.

B. Gain in the drift limited regime

The limit of pure drift is extreme and it is reasonable to replace in Eq. (20),

$$W^+ \to \frac{v_d}{a}. \tag{21}$$

The band drift velocity can be trap limited and can be subject to velocity saturation as given by Eq. (1).

The gain is then by definition

$$g = \frac{\tau_B}{L} v_d = \frac{\mu(p=0)F}{L}\left\{\frac{m+1}{\varepsilon_A} + \frac{am}{v_d}\right\} = \frac{v_d}{aN(m+1)}\left\{\frac{m+1}{\varepsilon_A} + \frac{am}{v_d}\right\}, \tag{22}$$

where $L=N(m+1)a$ and N is the number of repeat periods. The second term would not be there if the carrier is assumed to start its journey at A. Note that if we included the time to travel to the middle of the A (quantum well), as in classical physics, then the second term would also be $(m+1)$ and not m. Equation (22) implies that as soon as the transit time per repeat unit of the superlattice (SL) is shorter than the capture rate, the constant gain is controlled by the capture/recombination rate and grows with the drift velocity. If the drift velocity is linear in bias V, then so is the gain. On the other hand, if the drift is trap controlled, then it can rise up faster than V, as will be discussed later under the assumption of a constant value of ε_A. If the capture rate decreases with bias, as shown in the Monte Carlo work of Refs. 13 and 14 at high biases, then the increase in gain with bias can be even stronger. The problem, however, is that the noise, we believe, is no longer dominated by GR processes at high biases. We suggest that in the saturation regime hot electron processes determine the noise. This idea needs detailed theoretical consideration, which is beyond the scope of this paper. It seems intuitive though that at high bias carrier velocities are only defined up to an optic phonon energy of ~30 meV and slowed down carriers trapped in X-L valleys will seriously disturb newly arriving carriers until they themselves relax back down into Γ. We will come back to this point later in the paper.

In the papers by Beck,[21] Liu,[22] Choi,[23] and Levine *et al.*,[8] one used a semiclassical description and wrote for the gain g, again defined as the ratio of recombination to transit time,

$$g = \frac{1}{N}\frac{1-p_c}{p_c}, \tag{23}$$

$$p_c = 1 - \exp(-t_p/\tau_H), \tag{24}$$

where t_p is the transit time across one period and τ_H is the recombination time from an extended state back into the well. When $t_p \ll \tau_H$, we have

$$g = \frac{1}{N}\frac{\tau_H}{t_p}, \tag{25}$$

where $N \cdot t_p$ is just the total transit time, and Eq. (25) is the familiar result if the carrier starts its journey from the quantum well with $1/\tau_H = \varepsilon_A$ and $t_p = a/\mu F$.

Equation (25) has a quantum-mechanical definition. It is the ratio of the effective recombination time from an excited extended state to the ground state to the transit time. It involves therefore the quantum-mechanical mobility. Here we are not counting the time it takes to cross the well itself since it is considered to be one coherent eigenstate.

The present derivations in Eq. (22) agree with the usual result in this limit but they are more general, see the Appendix for $m=2$. In particular, they show that, for small applied fields, diffusion can play an important role. The present method also gives us the time dependence of the carrier recombination kinetic [see Eqs. (19) and (20)].

When $t_p \gg \tau_H$, then Ref. 23 predicts

$$g \sim \frac{e^{-t_p/\tau_H}}{1} = e^{-a\varepsilon_A/\mu F}, \tag{26}$$

a result which is purely classical and not verified in this formalism. It says that the gain decreases exponentially as the drift slows down or alternatively as the recombination rate increases.

V. BEHAVIOR OF TRAPS

A. Case when the band mobility is trap limited

The band mobility can and will in general be limited by trapping in randomly positioned defects and even the QW excited states themselves. As long as the particle escapes again, the defect acts only as a trap. There are of course many different way to treating the trapping function depending on the trap. One elegant way, which encompasses many scenarios, is to treat the escape process as a sum of paths that span the direct vertical thermal path to the pure adiabatic Fowler Nordheim limit.[1]

The mobility function to be inserted into Eq. (1) is

$$\mu = \frac{\mu_0 \left\{ \dfrac{e^{-E_{\rm tc}/kT} - e^{-\varsigma E_{\rm tc}^{3/2}/eFa_0} e^{-\varsigma E_{\rm tc}^{1/2}} e^{eFa_0/kT}}{1 - e^{-\varsigma E_{\rm tc}^{1/2}} e^{eFa_0/kT}} \right\}}{x + \left\{ \dfrac{e^{-E_{\rm tc}/kT} - e^{-\varsigma E_{\rm tc}^{3/2}/eFa_0} e^{-\varsigma E_{\rm tc}^{1/2}} e^{eFa_0/kT}}{1 - e^{-\varsigma E_{tc}^{1/2}} e^{eFa_0/kT}} \right\}}, \tag{27}$$

and this gives the trap and saturation limited drift velocity in the barrier layers,

$$v_d = \frac{\mu F}{\left\{1 + \left(\dfrac{\mu_0 F}{v_s}\right)^2\right\}^{1/2}}, \tag{28}$$

where $\varsigma = a_0(2m_e^*/\hbar^2)^{1/2}$ and a_0 the lattice constant ~ 0.5 nm, and we have used $\Sigma_l[W_{bl}(1-n_l)/p + W_{lb}]$ as the inverse of the path sum shown in the curly brackets in Eq. (27). Also, E_{tc} is the trap energy measured with respect to the conduction band, μ_0 is the pure trap free band mobility, and x is the volume trap concentration. There may also be distributions of trap energies, in particular when we have nanopillars or nanowires with incomplete passivation of surface states. In this case, one has to carry out another average over Eq. (27). The form in Eq. (27) is a good starting point for fitting data, but many other trap functions can be used depending on the nature of the trap. Accurate information on the trapping process and trap energy distributions and escape can only be obtained by looking at the time decay of the photocurrent. Such experiments are however scarce to come by in QWIPs.

B. Capture step

Under the assumption that the carrier re-emission rate from the excited state back to the band W_{eb} is slow compared to the recombination rate to the ground state W_{eg}, it follows that the entire recombination process denoted by ε_A, is, in general, a combined process of capture with rate W_{be} and then recombination from the excited state to the ground state W_{eg}. We can write $1/\varepsilon_A = 1/W_{be} + 1/W_{eg}$. The capture step is in reality one which involves a three-dimensional (3D) extended conduction-band state falling into a semiextended two-dimensional (2D) eigenstate. In the absence of momentum scattering in the (x,y) plane, we can assume that the k state in the plane does not change.

C. Band motion in the plane of the QW with traps

In the drift limited regime and in a very high quality system, the carrier recombination is controlled by the motion in the z direction. However, as the bias is lowered, and/or the defect density increases, recombination can become diffusion controlled and this will also involve the motion in the QW plane. The general quasi-3D case with $m=1$ has been treated theoretically in Ref. 24 in some detail. We recall that the diffusion controlled decay with deep traps of concentration n_t in 2D systems, with in-plane diffusivity $D_\parallel$ and radius R_o in 2D, follows a long-time decay law of the form,

$$n(r_{qw},z,t) \sim \exp\left[-2\sqrt{\pi}\left(\frac{n_t R_0 D_\parallel t}{a}\right)^{1/2}\right] n_z(t). \tag{29}$$

The zero-bias survival fraction in z direction is n_z, for general values of barrier width m, is given by the solution of the diffusion/Kronig-Penney model. The zero-bias relaxation dynamic in a superlattice is discussed in Ref. 24.

VI. CALCULATION AND MEASUREMENT OF THE GAIN

We now have enough information to plot the gain as given by Eqs. (2) or (22) in various scenarios. The capture

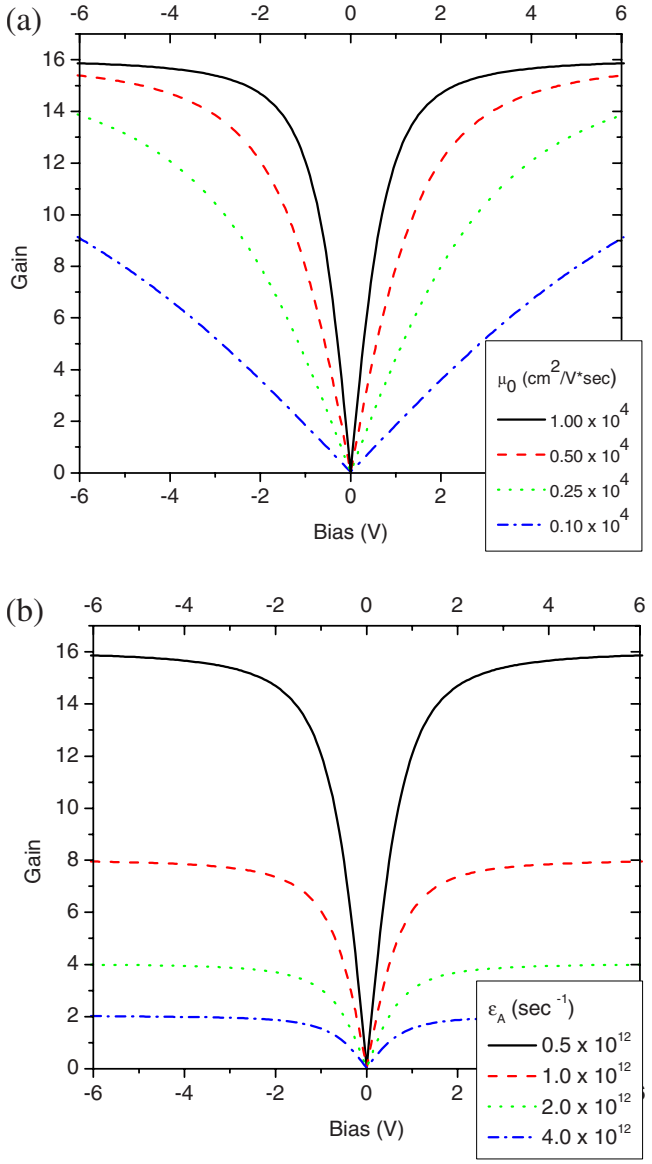

FIG. 2. (Color online) Theoretically calculated gain as a function of a bias with $a=5$ nm, $m=6$, $N=25$, $m^*=0.085m_0$, and $v_s=1.0\times10^8$ cm/s, assuming no traps, (a) for different μ_0 with $\varepsilon_A=5.0\times10^{11}$ s^{-1} and (b) for different ε_A with $\mu_0=1.0\times10^4$ cm^2/V s.

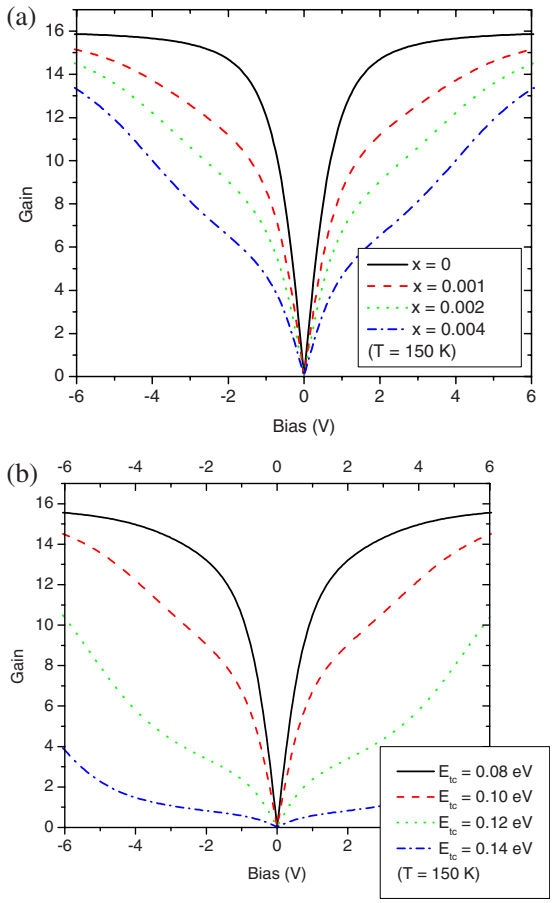

FIG. 3. (Color online) Theoretical evaluation of the trap effects for gains at $T=150$ K with $a=5$ nm, $m=6$, $N=25$, $m^*=0.085m_0$, $v_s=1.0\times10^8$ cm/s, $\mu_0=1.0\times10^4$ cm^2/V s, and $\varepsilon_A=5.0\times10^{11}$ s^{-1}, (a) for different x with $E_{tc}=0.10$ eV and (b) for different E_{tc} with $x=0.002$.

rate ε_A, the band mobility μ_0, the volume trap concentration x, and the trap energy E_{tc} are fit parameters. The rest are known parameters. Figure 2 presents plots of the calculated gain as a function of bias from Eq. (22) for the pure drift limit with $a=5$ nm, $m=6$, $N=25$, $m^*=0.085m_0$, and $v_s=1.0\times10^8$ cm/s, assuming no traps and v_d taken from Eq. (1). Figure 2(a) is a plot for different values of μ_0 with $\varepsilon_A=5.0\times10^{11}$ s^{-1}.

The works in Refs. 13 and 14 shows that the capture rate ε_A can be assumed a constant with bias until we have very high biases. They also show that capture rate is strongly well width dependent going up with a in our parameters. Because the QD layer serves as an additional active region, the QDWIP will have a capture rate 3~4 times as large as the QWIP with otherwise the same parameters. Figure 2(b) shows the calculated gain for different ε_A with $\mu_0=1.0\times10^4$ cm^2/V s.

Figure 3 depicts the effects of trapping on the gain. With $\mu_0=1.0\times10^4$ cm^2/V s, $\varepsilon_A=5.0\times10^{11}$ s^{-1}, and $a_0=0.5$ nm, Fig. 3(a) is a plot of gain for different values of the trap concentration "x" with $E_{tc}=0.10$ eV and; Fig. 3(b) is for different E_{tc} with $x=0.002$ at $T=150$ K. The trapping effects are temperature dependent. Figure 4 is a corresponding plot for different values of E_{tc} with $x=0.002$ at $T=150$, 200, and 250 K. It is clearly seen that the effect of trapping becomes weak as we go up in temperature.

Let us now consider the experimental gain extracted from the GR noise formula. The device structures for QWIP and QDWIP samples are shown in Figs. 1(a) and 1(b), and the fabrication procedures are as explained in Ref. 3. From the fabricated devices, dark and noise current are measured, and gain plots are obtained using Eq. (5). Figure 5 is a plot of the gain as a function of bias V extracted from an InGaAs/InAlAs (3.5 nm) QWIP at $T=120$ and 150 K.

The gain extracted using Eq. (5) around zero bias is an artifact because the noise curve remains bias independent in this limit at low temperatures below $T=150$ K, whereas the dark current depends on bias as can be seen in Fig. 6. The

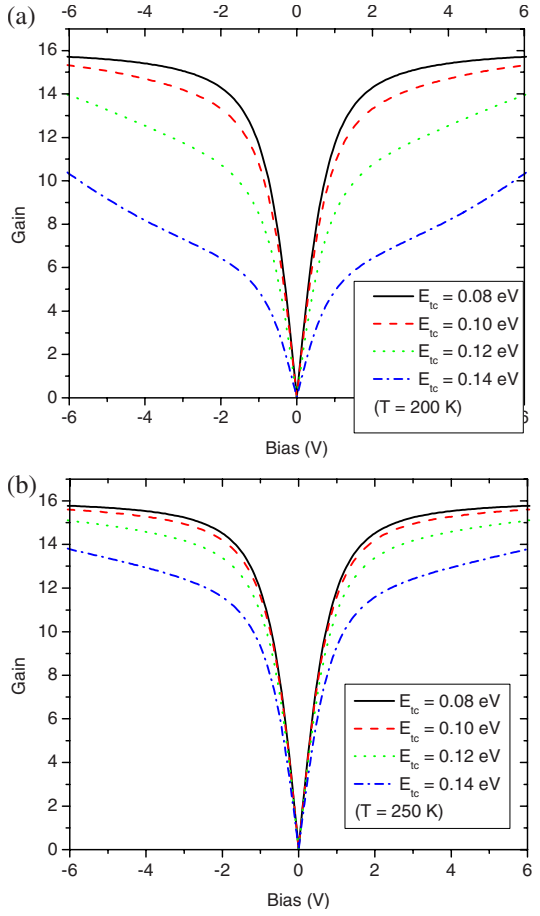

FIG. 4. (Color online) Temperature dependence of the theoretically calculated gain with $a=5$ nm, $m=6$, $N=25$, $m^*=0.085m_0$, $v_s=1.0\times10^8$ cm/s, $\mu_0=1.0\times10^4$ cm^2/V s, $\varepsilon_A=5.0\times10^{11}$ s^{-1}, $x=0.002$, and $E_{tc}=0.10$ eV.

fact that the noise current does not rise with the dark current in this regime suggests that it is not GR limited and therefore the GR noise formula should not really be used in this voltage range. Johnson noise can be bias independent and is

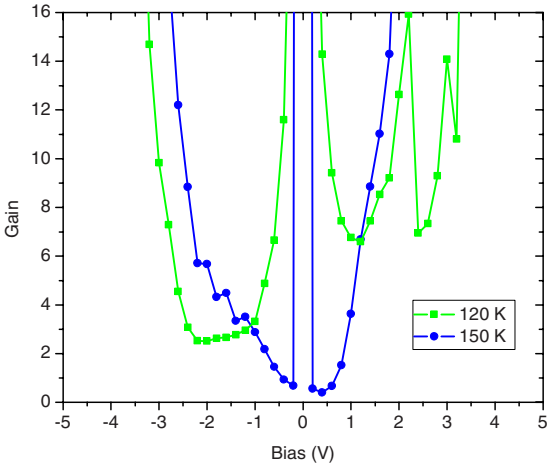

FIG. 5. (Color online) Plot of the gain in the InAlAs/InGaAs QWIP as a function of a bias obtained from the experimental results.

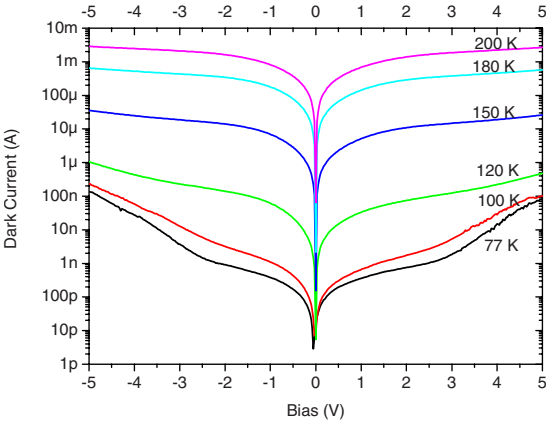

FIG. 6. (Color online) Dark current in the InAlAs/InGaAs QWIP.

probably the reason for this noise behavior at low bias and low temperatures. On the other hand, as we increase the bias, the noise becomes bias dependent and the gain starts to behave in a sensible manner and increases toward a high value of ten and higher. The dark current is saturating with bias, as shown in Fig. 6, but the noise (not shown) is not. In the high bias region, it is expected that the noise is dominated by hot carrier phonon emission and intervalley scattering, a generalization of Johnson noise with local space-charge oscillations and is no longer limited by GR noise. Equation (5) becomes invalid in this limit. As a consequence, the extracted high gain at high bias is unrealistic. This high gain will be discussed again later.

The temperature dependence of the gain in QWIP was also checked. However, because the dark current is too low or the noise is too high or unstable, most of the data seems to be meaningless. So, only the results at 120 and 150 K are shown in Fig. 5.

Figure 7 is the corresponding data for a QDWIP whose structure is shown in Fig. 1(b). The growth conditions of the QD layers are slightly modified from those discussed in Ref. 3, and the data presented in Fig. 7 is the best data to date.

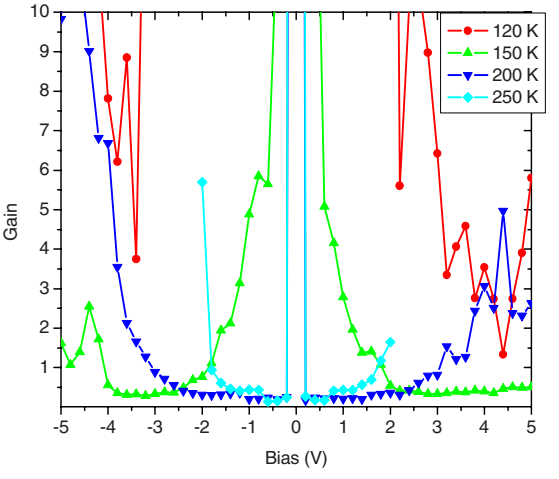

FIG. 7. (Color online) Plot of the gain in the InAlAs/InGaAs-QW/InAs-QD QDWIP as a function of a bias obtained from the experimental results.

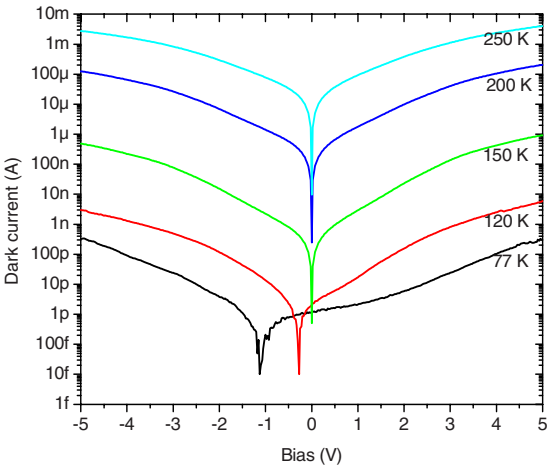

FIG. 8. (Color online) Dark current in the InAlAs/InGaAs-QW/InAs-QD QDWIP.

Figure 8 is a plot of the dark current measured in the QDWIP for several temperatures. Extremely high gain appears again around zero bias and also in the high bias region. Both these limits are unrealistic for the same reasons as in the QWIP. The gain in the physically sensible region of bias, at roughly −2 V, is lower in the QDWIP by a factor of 4. For example, at −2 V, the gain reaches roughly 4 in the QWIP, on the other hand, the gain keeps staying less than 1 in the QDWIP. This is consistent with the fact that the width of the active region is now ~2 nm larger, as shown by the Monte Carlo work in Refs. 13 and 14. In addition, one expects that the QD layers have added additional pathways for capture and recombination.

Figure 9 is a fit of the QWIP data. It is assumed that the QWIP is of good crystal quality and that there are no traps. This implies that the gain curve has no dependency on temperature (apart from what might be due to the saturation velocity). The fit curve in Fig. 9 is obtained using $v_s=1.0\times10^8$ cm/s, $\mu_0=0.10\times10^4$ cm^2/V s, and $\varepsilon_A=3.5\times10^{11}$ s^{-1}.

Figure 10(a) is a fit to the QDWIP data. In this case, the effects of traps are taken into account, the fit parameters are

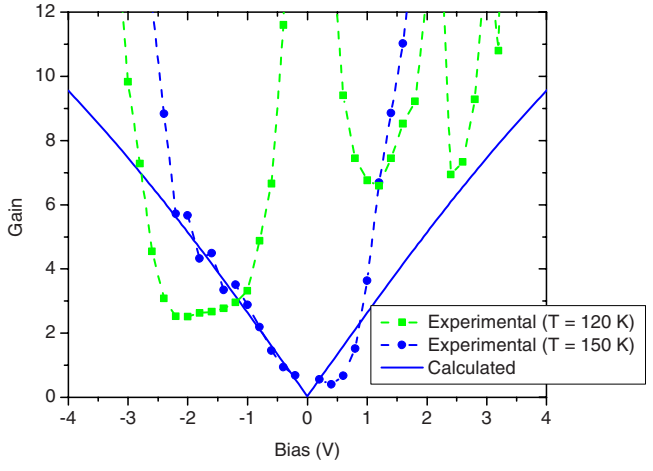

FIG. 9. (Color online) Fitting result for the QWIP gain with $a=5$ nm, $m=6$, $N=25$, $m^*=0.085m_0$, $v_s=1.0\times10^8$ cm/s, $\mu_0=0.10\times10^4$ cm^2/V s, and $\varepsilon_A=3.5\times10^{11}$ s^{-1}.

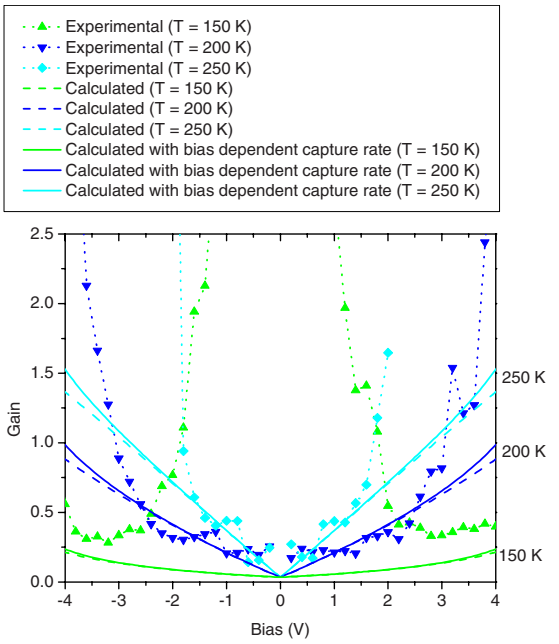

FIG. 10. (Color online) Fitting result for the QDWIP gain with $a=5$ nm, $m=6$, $N=25$, $m^*=0.085m_0$, $v_s=1.0\times10^8$ cm/s, $\mu_0=0.10\times10^4$ cm^2/V s, $\varepsilon_A=2.0\times10^{12}$ s^{-1}, $x=0.004$, and $E_{tc}=0.13$ eV. The dashed lines show the fitting without effect of decreasing ε_A due to the bias. The solid lines show the fitting with the effect included.

$v_s=1.0\times10^8$ cm/s, $\mu_0=0.10\times10^4$ cm^2/V s, $\varepsilon_A=2.0\times10^{12}$ s^{-1}, $x=0.004$, and $E_{tc}=0.13$ eV. For $T=200$ and 250 K, the fit results look good up to $V=2$ V.

Above 2 V, the experimentally extracted gain rises above the theory fit curve. The work in Refs. 13 and 14 shows that the capture rate can be assumed to be constant under low electric field, but it starts decreasing above a certain value of electric field. The electric field at which the capture rate starts decreasing corresponds to roughly 2 V for our device dimension. Using the Monte Carlo derived function from Refs. 13 and 14, we find that the fit is slightly improved as shown in Fig. 10(b). Although decreasing the capture rate could be one of the reasons for the higher gain, this mechanism on its own does not explain the discrepancy at high gain. The reason is most likely the different noise mechanism.

To understand the high values of the gain in Figs. 5 and 7, one should consider that the dark current has saturated at $V=2$ V but the noise has not. So Eq. (5) gives a gain that is increasing with bias. But this implies that another source of noise is taking over, which is not GR noise but it has to do with hot carriers. At velocity saturation, the standard formulations of noise no longer apply. From Figs. 6 and 8, one can see that I_D has saturated at $V=2$ V but the power is IV and is still going up. The carriers are emitting optic phonons or jump into to a slower velocity upper valley every time they reach the "right" energy. The Monte Carlo work in Ref. 14 clearly shows that this is happening, but the authors have not calculated the corresponding noise. A full Monte Carlo analysis that includes the time-dependent space-charge fluctuations and carrier velocity noise caused by optic phonon

emission would be useful. In a real device, this is further enhanced by percolation effects caused by pinholes, i.e., spatially restricted low resistance pathways.

It is also an interesting point to note in the QWIP results that the gain exceeds 1 and reaches values of $2\sim5$ at $V=-2$ V. This is considered to be still a reasonable regime although larger than the value reported in Ref. 4. In fact, this range of gain is still, in principle, possible as shown by our calculations. However, the gain $g=2$ means that the carriers can "fly" a distance twice as long as the length of the active region, and this implies in our case that the carriers cross the QWs 50 times before recombining. The crossing of the well region does slightly lower the effective band mobility.

In the QDWIP results, on the other hand, the gain is less than 1 around $V=\pm2$ V. When $g=0.4$, for example, the corresponding free flight length is $0.4\times25=10$ period of the active layers, roughly half the distance to the electrode. One of the reasons for the lower gain in the QDWIP is the increased capture rate. The QDWIP has obviously a thicker active region compared to QWIP, and this will increase the capture rate as reported in Ref. 14. In addition, the fit results are an indication that traps could be the origin of the lower gain. In modern growth technology, a planar QWIP generally has good crystal quality, and a QDWIP might have relatively lower quality. This is especially true of our QDWIPs grown by metal-organic chemical-vapor deposition (MOCVD) instead of molecular beam epitaxy (MBE). Thus traps in the structures can reduce the gain and cause a nonlinear bias dependence.

In a situation as in Ref. 25 with detection wavelengths ~8 μm, where the responsivity keeps on going up with bias and there is no apparent saturation, the authors were talking in terms of avalanche processes and avalanche noise contributions.[26,27] This is clearly not the case for our devices because the responsivity saturates above about 2 V. However, hot carrier electron-electron relaxation in the subbands could also be a source of noise and this should be investigated by measuring doping dependence.

The QDWIP introduces a bound excited state into the well and introduces at least one additional ground localized level inside the QD. This increases the number of pathways for recombination into the ground QD/QW bands. The actual band structure of this class of QDs in the well devices (QDWIP) is quite unlike the traditional QDIP,[15,28] or even QD in well devices of Refs. 5 and 6 for example. In the present category of InAs/InGaAs devices, we are dealing with a very low (InAs/InGaAs conduction-band offset) confinement potential of ~90 meV in the QDs. The implication is that we have one bound level and the rest are QW-like bands. Normally at low temperatures, the bound level is doubly occupied and the remaining carriers are in the QW-like bands. Here we have $10^{18}/\mathrm{cm}^3$ doping in a 3.5 nm QW so that there are at least five electrons per QD unit. This system is therefore truly a hybrid, and charges can flow from the QW bands in and out of the QD bound level relatively freely.

VII. TIME DEPENDENCY OF EXCITED CARRIERS IN A QWIP

A. Time decay in the case when $m=1$, i.e., $b=a$

As an illustration of the power of the method, we use Eqs. (12) or (13) to calculate the time dependence of the number of survivors $n_B(t)$ for $m=1$ by elementary algebra and find for pulsed generation, i.e., $G(p)=1$. Here,

$$n_B(t)=\left\{\frac{1}{2}-\frac{\varepsilon_A+W_t-(a_t+b_t)/2}{a_t-b_t}\right\}e^{-a_t t}+\left\{\frac{1}{2}+\frac{\varepsilon_A+W_t-(a_t+b_t)/2}{a_t-b_t}\right\}e^{-b_t t}, \tag{30}$$

$$2a_t=2W_t+\varepsilon_A-[4W_t^2+\varepsilon_A^2]^{1/2}, \tag{31}$$

$$2b_t=2W_t+\varepsilon_A+[4W_t^2+\varepsilon_A^2]^{1/2}, \tag{32}$$

$$W_t=W^+ + W^-, \tag{33}$$

$$n_A(t)=\left\{\frac{1}{2}+\frac{(-\varepsilon_A)+2W_t}{2[4W_t^2+\varepsilon_A^2]^{1/2}}\right\}e^{-a_t t}+\left\{\frac{1}{2}-\frac{(-\varepsilon_A)+2W_t}{2[4W_t^2+\varepsilon_A^2]^{1/2}}\right\}e^{-b_t t}, \tag{34}$$

$$n_B(t)=\left\{\frac{1}{2}+\frac{\varepsilon_A+2W_t}{2[4W_t^2+\varepsilon_A^2]^{1/2}}\right\}e^{-a_t t}+\left\{\frac{1}{2}-\frac{\varepsilon_A+2W_t}{2[4W_t^2+\varepsilon_A^2]^{1/2}}\right\}e^{-b_t t}. \tag{35}$$

The time decay of the particle is not as trivial as one might think. The long-time behavior is in the first term with decay rate constant,

$$a_t=\frac{\varepsilon_A}{m}\left\{\left[\frac{v_d}{\varepsilon_A a_w}+\frac{m}{m+1}\right]-\left[\left(\frac{v_d}{\varepsilon_A a_w}\right)^2+\left(\frac{m}{m+1}\right)^2\right]^{1/2}\right\}. \tag{36}$$

One can see that the first term inside the square bracket is what some people define as the inverse capture probability Eq. (24) when it is >1. The time decay can in principle be measured using a very fast laser pulse or terahertz absorption techniques. The full 3D decay is given by Eq. (29) with Eq. (34) for the z direction. In a low defect sample, Eq. (34) is the dominant process.

B. Time dependence in the drift limited regime

If we consider $m>1$, and pure drift, then we make the corresponding replacements $W_t\to W^+\to v_d/am$ and replace $\varepsilon_A/2$ with $\varepsilon_A/(m+1)$ in Eq. (34). The long-time behavior is determined by the first term in Eq. (36).

VIII. SUMMARY AND CONCLUSIONS

We have presented a more rigorous theory of transport and recombination, which is applicable to QWIPs, QDWIPs, and nanopillar QWIPs. Let us summarize the findings.

In QWIPs we have considered the usual QWIP model and our result agrees with the standard results in the (physically relevant) high-field drift limited Eq. (23) regime. The effect of diffusion is neglected in the standard treatments of gain.

This is reasonable when we have order, but not so when we have disorder or more complex QWIP structures. The present method always refers to quantum-mechanical rates and not to classical probabilities. Our method, which includes trapping, shows more clearly how the gain depends on electric field and temperature. Diffusion adds far more complexity to the results as can be seen from Eq. (A1). We have solved the $m=2$ case exactly to demonstrate this. The time dependence of the photoconductivity in the drift limited limit has been solved. This is an important result for the study of the time dependence of the pulsed photocurrent and will be used later when data are available.

The theory was applied to experiments on QWIP and QDWIPs. The gain was extracted using the GR noise formalism. The GR noise formula Eq. (5) only has a limited region of validity and this must be checked each time. In the present system, the low-temperature noise is dominated by Johnson noise and does not follow the GR formula in the low bias region. The gain data for the QWIP are linear in bias but the QDWIP gain data have nonlinear bias dependence. This is proof that traps must have been introduced in the QDWIP structures. In our devices, deviations from the standard GR noise at high bias, we believe, is due to intervalley crossing and optic phonon emission. Unfortunately there is no theoretical formalism to describe this type of noise at present. We suggest that one good way is to include it in a Monte Carlo simulation. When avalanching contributes to the photoconductivity as in Ref. 26, the additional noise has been considered in Ref. 27. Intersubband avalanche contributions to the noise depends on the doping level and can be studied by varying the degree of doping in the active layers.

APPENDIX: THE RECOMBINATION TIME FOR $m=2$

If $m>1$, the exact result becomes very complicated in comparison to the pure drift limit. For example, when the barrier region is twice as long, $m=2$, we have instead in Eq. (11) the exact result with back diffusion and for $m=2$ we have

$$\tau_B = \frac{(1+S)(W^+ + W^- + \varepsilon_A) + (W^- + W^+ S)}{(W^+ + W^-)(W^+ + W^- + \varepsilon_A) - W^+ W^- - S[(W^+)^2 + W^-(W^+ + W^- + \varepsilon_A)]}, \tag{A1}$$

$$S = 1 - \frac{(W^- \varepsilon_A)}{(W^+ + W^-)[(W^+ + W^- + \varepsilon_A)] - W^+ W^-}, \tag{A2}$$

which gives back the very simple result $\tau_B = 3/\varepsilon_A + 2/W^+$, when $W^-=0$ in Eq. (26).

The complexity occurs because the solution now takes into account the fact that there is diffusion and drift.

[1] H. Lim, B. Movaghar, S. Tsao, M. Taguchi, W. Zhang, A. A. Quivy, and M. Razeghi, Phys. Rev. B **74**, 205321 (2006).
[2] S. D. Gunapala and S. V. Bandara, in *Semiconductors and Semimetals: Intersubband Transitions in Quantum Wells: Physics and Device Applications I*, edited by H. C. Liu and F. Capasso (Academic, New York, 1999), p. 197.
[3] H. Lim, S. Tsao, W. Zhang, and M. Razeghi, Appl. Phys. Lett. **90**, 131112 (2007).
[4] S. Ozer, U. Tumkaya, and C. Besikci, IEEE Photon. Technol. Lett. **19**, 1371 (2007).
[5] P. Aivaliotis N. Vukmirovic, E. A. Zibik, J. W. Cockburn, D. Indjin, P. Harrison, C. Groves, J. P. R. David, M. Hopkinson, and L. R. Wilson, J. Phys. D **40**, 5537 (2007).
[6] P. Aivaliotis, E. A. Zibik, L. R. Wilson, J. W. Cockburn, M. Hopkinson, and N. Q. Vinh, Appl. Phys. Lett. **92**, 023501 (2008).
[7] A. Rogalski, J. Appl. Phys. **93**, 4355 (2003).
[8] B. F. Levine, A. Zussman, S. D. Gunapala, M. T. Asom, J. M. Kuo, and W. S. Hobson, J. Appl. Phys. **72**, 4429 (1992).
[9] H. C. Liu, R. Dudek, A. Shen, E. Dupont, C. Y. Song, Z. R. Wasilewski, and M. Buchanan, Int. J. High Speed Electron. Syst. **12**, 803 (2002).
[10] C. Jelen, S. Slivken, T. David, M. Razeghi, and G. J. Brown, IEEE J. Quantum Electron. **34**, 1124 (1998).
[11] S. D. Gunapala, K. M. S. V. Bandara, B. F. Levine, G. Sarusi, J. S. Park, T. L. Lin, W. T. Pike, and J. K. Liu, Appl. Phys. Lett. **64**, 3431 (1994).
[12] A. Gin, B. Movaghar, M. Razeghi, and G. J. Brown, Nanotechnology **16**, 1814 (2005).
[13] O. O. Cellek and C. Besikci, Semicond. Sci. Technol. **19**, 183 (2004).
[14] S. Memis, O. O. Cellek, U. Bostanci, M. Tomak, and C. Besikci, Turk. J. Phys. **30**, 335 (2006).
[15] J. Szafraniec, S. Tsao, W. Zhang, H. Lim, M. Taguchi, A. A. Quivy, B. Movaghar, and M. Razeghi, Appl. Phys. Lett. **88**, 121102 (2006).
[16] E. Rosencher and B. Vinter, *Optoelectronics* (Cambridge University Press, Cambridge, 2002).
[17] A. Matulionis, IEICE Trans. Electron. **E89-C**, 913 (2006).
[18] B. Movaghar, G. W. Sauer, D. Wurtz, and D. L. Huber, Solid State Commun. **39**, 1179 (1981).
[19] B. Movaghar, B. Pohlmann, and D. Wurtz, Phys. Rev. A **29**, 1568 (1984).
[20] K. K. Ghosh, L. H. Zhao, and D. L. Huber, Phys. Rev. B **25**, 3851 (1982).
[21] W. A. Beck, Appl. Phys. Lett. **63**, 3589 (1993).
[22] H. C. Liu, Appl. Phys. Lett. **60**, 1507 (1992).
[23] K. K. Choi, Appl. Phys. Lett. **65**, 1266 (1994).
[24] B. Movaghar, Semicond. Sci. Technol. **4**, 95 (1989).
[25] S. Ozer, U. Tumkaya, B. Asici, and C. Besikci, IEEE J. Quantum Electron. **43**, 709 (2007).
[26] R. Rehm, H. Schneider, M. Walther, P. Koidl, and G. Weimann, Appl. Phys. Lett. **82**, 2907 (2003).
[27] H. Schneider, Appl. Phys. Lett. **82**, 4376 (2003).
[28] H.-C. Lim, S. Tsao, M. Taguchi, W. Zhang, A. A. Quivy, and M. Razeghi, Proc. SPIE **6127**, 61270N (2006).