PHYSICAL REVIEW B **81**, 205311 (2010)

Simplified density-matrix model applied to three-well terahertz quantum cascade lasers

E. Dupont,* S. Fathololoumi, and H. C. Liu
Institute for Microstructural Sciences, National Research Council, Ottawa, Ontario, Canada K1A 0R6

(Received 17 December 2009; published 12 May 2010)

A simplified density-matrix model describing the population and coherence terms of four states in a resonant phonon scattering based terahertz quantum cascade laser is presented. Despite its obvious limitations and with two phenomenological terms—called the pure dephasing time constants in tunneling and intersubband transition—the model agrees reasonably well with experimental data. We demonstrate the importance of a tunneling leakage channel from the upper lasing state to the excited state of the downstream phonon well. In addition, we identify an indirect coupling between nonadjacent injector and extractor states. The analytical expression of the gain spectrum demonstrates the strong broadening effect of the injection and extraction couplings. The gain is decomposed into three terms: a linear gain and two nonlinear components related to stimulated anti-Stokes scattering processes. The nonlinear gain is not negligible at high temperature. Under certain approximations, analytical forms of population and coherence terms are derived. This model is well suited for structures with only a few states involved. This model can simplify the optimization process for new laser designs; it is also convenient for experimentalists to adopt.

DOI: 10.1103/PhysRevB.81.205311 PACS number(s): 73.40.Gk, 73.23.−b, 42.55.Ah, 72.20.Dp

I. INTRODUCTION

The temperature performance of terahertz (THz) quantum cascade lasers (QCL) has been continuously improved to the point that the empirical limit of the maximum thermal energy, $k_B T_{\max}$, to the photon energy[1] has been slightly surpassed in devices emitting from 1.2 to 3 THz,[2,3] which makes these photonic devices unique. The maximum operating temperature reached so far is 186 K,[4] and it is now only a step away from temperatures achievable by thermoelectric coolers. This will then open a broad range of applications from biological sensing, pharmaceutical sciences, THz wave imaging to hazardous materials detection, to cite just a few.[5] These high-temperature devices are based on resonant longitudinal-optical (LO) phonon scattering for depopulation, except for the structures emitting close to 1 THz. These structures offer the advantage of consisting of only a few quantum wells (QWs), and therefore only a few relevant states per period are involved—which simplifies the model presented in this work.

QCLs have attracted the attention of many theoreticians as well, and several useful models have been developed to predict the transport, the electron distribution, the gain, and the relative importance of different scattering mechanisms. We refer to the self-consistent rate equation model,[6] the Monte Carlo simulations,[7–10] the hybrid density-matrix Monte Carlo model,[11] and the quantum theories based on nonequilibrium Green's function[12–16] or on density matrix.[17–20] Some groups have reported a remarkable agreement between the theoretical electrical characteristics and experiments.[14–16,20] However, the implementation of such models is difficult and also computationally very demanding.

Alternatively, Scalari *et al.* have employed a simplified density-matrix model for a five-well QCL emitting at 3.7 THz. This work, after a fitting procedure of the optical and electrical characteristics, estimated the upper lasing state lifetime and relaxation time constant in the injector region.[21] These parameters are useful for subsequent optimization. Kumar *et al.* developed a similar model which is used as a guideline tool for designing resonant LO-phonon scattering based QCLs,[22,23] as well as an analyzing tool of the laser performance.[24] With the same model, they also simulated the gain profile for different range of laser intensities.[22]

Recently, our group has introduced the three-well THz QCL (Ref. 25) and carried out several optimization studies, such as the effect of the injection[26] and extraction[27] barrier thicknesses. This design consists of placing an active double well with a vertical intersubband transition in a sandwich between two identical LO-phonon resonant wells. The tunneling interactions between these three wells occur through the extraction and injection tunneling barriers. Here, the active double well and its two phonon wells for injection and extraction form the so-called *three-well structure*, even though there are four quantum wells. Within the main four states of this laser structure, five tunneling processes occur and they all participate in the diagonalization when calculating the adiabatic wave functions of this system. The simplified density-matrix model of Ref. 21 cannot be immediately applied to our three-well design because of the numerous tunneling and leakage processes taking place. In this work, we propose to extend the simplified density-matrix model to this case, which involves four states and five tunneling processes. In essence, our model is simply a system of rate equations for population and coherence terms, in which the electrons in each subband behave the same (i.e., like a single particle), regardless of their kinetic energy. We study two cases: first, a simple and very convenient case in which the gain profile is Lorentzian type and its strength is proportional to population inversion; and second, a more rigorous approach where the oscillating coherence terms between different states are directly related to the laser field. Near threshold, an analytical form of the gain can be derived; and it shows the strong broadening effect by the injection and extraction tunnelings. From this expression, three components can be identified: a main linear term proportional to population inversion and two nonlinear terms that are not negligible, particularly at high temperature. We apply this model to the experimental results of our injection barrier thickness study[26] and find estimates for the electron temperature and tunneling pure dephasing time constant.

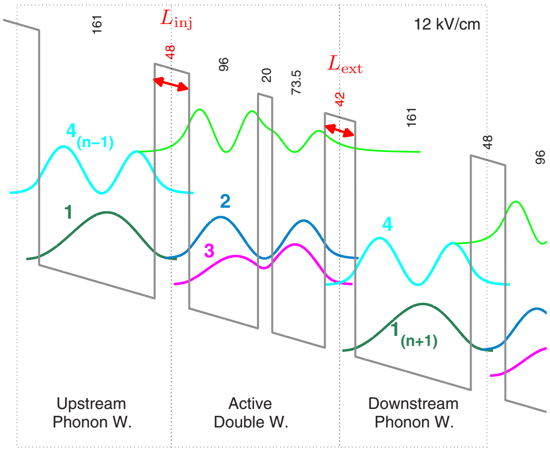

FIG. 1. (Color online) Conduction-band diagram of the three-well QCL structure of Ref. 25 at 12 kV/cm and the square modulus of the wave functions of the active double well and the upstream/downstream phonon wells when taken isolated from the adjacent quantum wells. The thickness in angstrom of each layer is given in vertically oriented font. In Ref. 25, the centered 50 Å of the phonon wells are Si doped at 7.2×10^{16} cm^{-3} for an equivalent two-dimensional carrier concentration $N_{2\mathrm{D}}=3.6\times10^{10}$ cm^{-2}. The subscripts $(n+1)$ and $(n-1)$ are referring to the downstream and upstream set of four states, respectively.

The paper is organized as follows. In Sec. II, we describe the quantum system under study. In Sec. III, we point out the principal limitations of our model. In Sec. IV, we solve the problem assuming the laser field only acts on the population of the lasing states; and we show that a *wrong extraction* channel is detrimental for continuous-wave (cw) operation. In Sec. V, the laser-induced coherence terms are added in the model, and an analytical expression for the gain close to threshold is derived and discussed. In Sec. VI, numerical applications of the model are performed to fit the experimental results of the injection barrier study of Ref. 26. As a result, pure dephasing time constants as well as electron heating temperature are estimated. Finally, the last section summarizes our conclusions. For readers who wish to use this model, more equations are given in the Appendix.

II. DESCRIPTION OF THE QUANTUM SYSTEM UNDER STUDY

Initially, the model requires calculating the wave functions of different quantum wells that are coupled to each other by a slow tunneling process. From now, we look at the three-well QCL design as an optically active double well interacting with an upstream phonon well for the carrier injection and a downstream phonon well for the extraction. The various wave functions of these wells are computed separately; this means they are isolated from the neighboring QWs. These wave functions are then appropriately shifted in the growth direction and finally plotted as shown in Fig. 1. This figure represents the isolated wave functions of the three-well QCL reported in Ref. 25 at 12 kV/cm. The relevant four states are named 1 for the injector level, 2 for the upper lasing state, 3 for the lower lasing state, and 4 for the excited state of the downstream phonon well acting like an extractor level. These levels are part of the nth set of the states and the states belonging to the upstream and downstream sets are noted as $i_{(n-1)}$ and $i_{(n+1)}$, respectively. At this electric field, the oscillator strength between the lasing states is 0.677. The expected values of position for the lasing states, Z_{ii} $(i=2,3)$, differ only by 1.4 nm, which denotes a high overlap between the wave functions. This is the reason why we say this design uses vertical intersubband transitions.

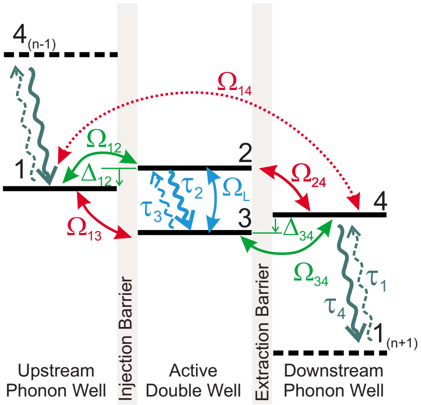

FIG. 2. (Color online) Schematic of the interactions considered in the model between the four relevant states in a three-well THz QCL. The Ω_{ij} are the tunneling coupling strengths. The injection (Ω_{12}) and extraction (Ω_{34}) are represented in green as opposed to the not so desirable tunneling channels like Ω_{13} for the wrong injection channel, and Ω_{24} for the wrong extraction channel. A parasitic and negligible channel Ω_{14} between 1 and 4 can also occur.

The interactions considered in this model between the four states are schematically represented in Fig. 2. The useful coupling strengths represented by green double-ended arrows are the injection, Ω_{12}, and extraction, Ω_{34}, tunnelings. These channels are optimum at the design electric field of the laser, E_{dsg}, i.e., close to $\sim$12.1 kV/cm for the structure of Ref. 25. However, at low electric fields, the injector level can also interact with the lower lasing state, a process called the *wrong injection* channel, Ω_{13}. Similarly, carriers injected to the upper lasing state can directly tunnel toward the extractor state, a process called wrong extraction, Ω_{24}. Less importantly, there is a weak direct *parasitic coupling* between the injector and extractor states, Ω_{14}, which is three orders of magnitude smaller than the other four tunneling couplings. The last three couplings are not desirable for the laser, and therefore, are represented by double-ended red arrows. The laser field, $\mathcal{E}_L$, as schematically represented by its Rabi frequency $\Omega_L=qZ_{23}\mathcal{E}_L/\hbar$ in Fig. 2, causes the lasing states to interact. The reason is that only this pair of states is in close resonance with the laser frequency and has a significant dipole moment. The laser field not only induces an oscillating coherence between 2 and 3, but also between 1–3, 2–4, and 1–4, as will be discussed in Sec. V. From a simplistic point of view, the laser field just reduces the upper lasing state lifetime via stimulated emission. This approximation is treated in Sec. IV. The inelastic scattering of the carriers in various states are represented by wavy arrows. The main relaxation processes for achieving laser operation are drawn with solid wavy downward arrows and processes related to phonon absorption are drawn by dashed wavy upward arrows. The upper lasing state time constant, τ_2, is estimated from the 2–3 intersubband unscreened impurity and bulk LO-phonon scattering rates.

From these isolated wave functions, the tunneling coupling strengths between the four states are derived by a simple tight-binding approach.[28] Each $i \rightarrow j$ tunneling is characterized by a detuning energy, $E_{ij}=E_i-E_j$, and a coupling strength, $\hbar\Omega_{ij}$. Figure 3 plots the detunings and couplings for all five tunnelings, versus the electric field. At the design electric field, the injection coupling is $\Omega_{12}=0.92$ meV and the extraction coupling is $\Omega_{34}=1.92$ meV. Ω_{34} is designed to be higher than Ω_{12} to ensure population inversion even at high temperature. The wrong injection coupling, $\Omega_{13}=0.56$ meV, is lower due to the high confinement energy of the two states 1 and 3. This observation already suggests that this leakage path will not play a major role around the design electric field. On the other hand, the wrong extraction coupling, $\Omega_{24}=2.39$ meV, is stronger than the extraction coupling. This might be a competing leakage path, particularly at low temperature, when the upper lasing time constant, τ_2, is long.

III. LIMITATIONS OF THE MODEL

Our model consists of solving Bloch equations of four interacting states independently of the kinetic energy of the carriers. This means that the carriers on a particular subband behave the same, like a single particle, regardless of their kinetic energy. Therefore, our solution is not resolved in the $\mathbf{k}_\parallel$ space of subbands. With this model, tunnelings conserve the in-plane wave vector $\mathbf{k}_\parallel$ but are not subject to Pauli's exclusion principle, which can be a severe limitation if the structures are heavily doped. We also ignore Pauli's exclusion principle for the intersubband lifetime of the states. Since the THz QCLs are lightly doped, we do not think this simplification will cause substantial errors.

This model—as it is implemented here—concentrates the thermal effects *only* on the upper lasing time constant, τ_2. It is important to underline the fact that this model omits any thermal effects on the intrasubband component of the dephasing time constants between the states. This component contributes strongly to the linewidth of the optical intersubband transition.[29] From this point on, this component is called the pure dephasing time constant, τ_{ij}^*. Simulations have suggested a strong temperature effect on intersubband transition linewidth since the electron screening length (Debye length) of the impurity Coulomb potential increases with the electron temperature.[30] However, simple models for the intrasubband scattering rate do not result in the same optical gain linewidth as the one computed in programs that use nonequilibrium Green's function.[31–33] Moreover, these sophisticated techniques should better predict the temperature dependence of the linewidth. Since in our model, τ_2 carries most of the responsibility for the temperature degradation of the laser, our fitting exercise in Sec. VI may lead to an underestimation of τ_2, or equivalently, to an overestimation of the electronic temperature. We think that this approximation may be a serious one which simplifies the model to the great extent.

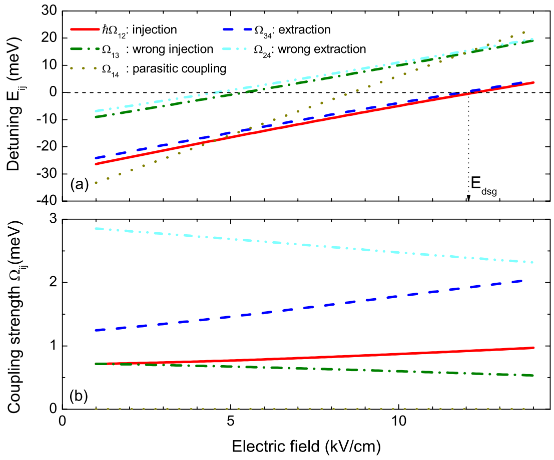

FIG. 3. (Color online) Panel (a) shows the detunings E_i-E_j for the different tunneling processes between the four states. Panel (b) shows the coupling strengths $\hbar\Omega_{ij}$. The same color and line style code for the different tunneling channels applies to both panels. The horizontal dashed line at zero detuning indicates the electric field for which the different tunnelings are in resonance. The vertical dashed line indicates the design electric field of the QCL, E_{dsg}. The 1–4 coupling strength is only 0.2–0.3 μeV.

The upper lasing lifetime, τ_2, is thermally activated and, in this work, is simply modeled as

$$\tau_2^{-1} = \frac{1+n_{\mathrm{LO}}}{\tau_{23}^{(\mathrm{LO\ emi,\ hot}\ e^-)}} \exp\left[\frac{E_{23}-E_{\mathrm{LO}}}{k_B(T+\Delta T_e)}\right] + \frac{n_{\mathrm{LO}}}{\tau_{23}^{(\mathrm{LO\ abs})}} + \frac{1}{\langle \tau_{23}^{(\mathrm{imp})} \rangle}, \tag{1}$$

where $\langle \tau_{23}^{(\mathrm{imp})} \rangle$ is the intersubband 2–3 unscreened impurity scattering time constant averaged over the Maxwell-Boltzmann distribution of carriers on level 2, E_{LO} is the LO-phonon energy, $\tau_{23}^{(\mathrm{LO\ emi,\ hot}\ e^-)}$ is the LO-phonon scattering time for an electron on level 2 with a kinetic energy equal to $E_{\mathrm{LO}}-E_{23}$, $\tau_{23}^{(\mathrm{LO\ abs})}$ is the LO-phonon absorption scattering time for an electron at the bottom of subband 2 to level 3, n_{LO} is the Bose-Einstein factor, k_B is the Boltzmann constant, and finally, $T+\Delta T_e$ is the electron temperature on level 2, T being the lattice temperature and ΔT_e the electron heating temperature. It is obvious that the electron heating temperature depends on input electrical power as well as the laser intensity.[34–36] However, in order to simplify our numerical applications, its value was kept independent of the electric field and lattice temperature, although we are aware that this simplification can distort the simulated current density versus electric field characteristic. Microphotoluminescence experiments on resonant phonon scattering based QCLs have demonstrated that the temperature of lasing subbands can be $\Delta T_e \approx 100$ K higher than the lattice.[35]

The lifetime of lower lasing, extractor, and injector states are written in a similar manner

$$\tau_3^{-1} = \frac{n_{\mathrm{LO}}}{\tau_{23}^{(\mathrm{LO\ emi,\ hot}\ e^-)}}, \tag{2a}$$

$$\tau_4^{-1} = \frac{1}{\langle \tau_{41'}^{(\mathrm{imp})} \rangle} + \frac{1 + n_{\mathrm{LO}}}{\tau_{41'}^{(\mathrm{LO\ emi})}} + \frac{n_{\mathrm{LO}}}{\tau_{41'}^{(\mathrm{LO\ abs})}}, \tag{2b}$$

$$\tau_1^{-1} = \frac{n_{\mathrm{LO}}}{\tau_{41'}^{(\mathrm{LO\ emi})}}. \tag{2c}$$

The terms with n_{LO} as a coefficient are related to phonon absorption processes and are not negligible at high temperatures. The phonon absorption processes from levels $1_{(n+1)}$ [which is denoted $1'$ in Eq. (2)] to 4, and from 3 to 2, are the reverse of the phonon emission processes $4 \rightarrow 1_{(n+1)}$ and $2 \rightarrow 3$. This is the reason why the same term $\tau_{41'}^{\mathrm{LO\ emi}}$ ($\tau_{23}^{\mathrm{LO\ emi,\ hot\ } e^-}$) appears in the expression of τ_4^{-1} (τ_2^{-1}) and τ_1^{-1} (τ_3^{-1}). Among the absorption terms in Eqs. (1) and (2), the most important term is τ_1^{-1}, because the phonon absorption from the heavily occupied state $1_{(n+1)}$ to state 4 acts as a backfilling process at high temperatures. The electron-phonon scattering rates are slightly overestimated since they are computed at resonance, i.e., for a vanishing initial or final kinetic energy. We use Harrison's numerical method to calculate them.[37]

In the model, states belonging to different wells are only coupled to each other through coherent tunneling, mitigated by dephasing processes. Interwell inelastic scattering between the different adjacent states was omitted because of the reduced overlap between these wave functions. Although these scattering terms should normally be included in the model, their omission keeps the model simple, and yet accurate. For instance, the interwell LO phonon scattering times are very long, except for $4_{(n-1)} \rightarrow 2$ where $\tau_{4_{(n-1)}2} \approx 7$ ps. We confirmed that the influence of this latter time is marginal. Since this model only focuses on the four main states within one period, it cannot predict the leakage to more remote states. For instance, at high electric fields, level 2 could be coupled to the second excited state of the next period active double well which is displayed by a light green wave function in Fig. 1.

With these obvious limitations in mind, it is clear that the fitted parameters derived with our model should be treated with caution. Nevertheless, we believe this simplified model can give useful insights for the effects of the different couplings on the gain. It can also be extended to structures consisting of more, or fewer, wells per period. Its simplicity makes the implementation very easy by experimentalists and can be used as a first-order optimization tool for new designs. In the next section, we will model the steady-state populations and the static coherence terms without including the laser Rabi frequency which induces oscillating coherence terms between states.

IV. MODEL WITHOUT LASER-INDUCED COHERENCE TERMS

When the laser-induced coherence terms are excluded, the time evolution of the density matrix, ρ, for the four-level system shown in Fig. 2, can be written as

$$\frac{d\rho}{dt} = \frac{1}{\jmath\hbar}\left[\begin{pmatrix} E_1 & \hbar\Omega_{12} & \hbar\Omega_{13} & \hbar\Omega_{14} \\ \hbar\Omega_{12} & E_2 & 0 & \hbar\Omega_{24} \\ \hbar\Omega_{13} & 0 & E_3 & \hbar\Omega_{34} \\ \hbar\Omega_{14} & \hbar\Omega_{24} & \hbar\Omega_{34} & E_4 \end{pmatrix}, \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & \rho_{22} & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & \rho_{33} & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & \rho_{44} \end{pmatrix}\right]$$
$$- \begin{pmatrix} \tau_1^{-1}\rho_{11} - \tau_4^{-1}\rho_{44} & \tau_{\parallel 12}^{-1}\rho_{12} & \tau_{\parallel 13}^{-1}\rho_{13} & \tau_{\parallel 14}^{-1}\rho_{14} \\ \tau_{\parallel 12}^{-1}\rho_{12} & \tau_2^{-1}\rho_{22} - \tau_3^{-1}\rho_{33} + \tau_{\mathrm{sti}}^{-1}\Delta\rho & \tau_{\parallel 23}^{-1}\rho_{23} & \tau_{\parallel 24}^{-1}\rho_{24} \\ \tau_{\parallel 13}^{-1}\rho_{31} & \tau_{\parallel 23}^{-1}\rho_{32} & \tau_3^{-1}\rho_{33} - \tau_2^{-1}\rho_{22} - \tau_{\mathrm{sti}}^{-1}\Delta\rho & \tau_{\parallel 34}^{-1}\rho_{34} \\ \tau_{\parallel 14}^{-1}\rho_{41} & \tau_{\parallel 24}^{-1}\rho_{42} & \tau_{\parallel 34}^{-1}\rho_{43} & \tau_4^{-1}\rho_{44} - \tau_1^{-1}\rho_{11} \end{pmatrix}, \tag{3}$$

where

$$\tau_{\mathrm{sti}}^{-1} = \frac{c}{n_g}\sigma_{\mathrm{op}}\Theta. \tag{4}$$

The variable Θ is the photon density in the cavity and $\Delta\rho$ is the population inversion, $\rho_{22}-\rho_{33}$. The left matrix in the commutator of Eq. (3) represents the Hamiltonian of the two phonon wells and the active double well interacting with each other by tunneling. We verified that the diagonalization of this matrix gives the same eigenfunctions as the adiabatic solution of the three-well structure. The last matrix on the right-hand side represents the natural decay of population and coherence terms when these terms are placed in nonequilibrium condition. The decay of the populations, ρ_{ii}, is given by the diagonal terms. For the lasing states, (i=2,3), these terms include the stimulated emission rate, τ_{sti}^{-1}, the expression of which is given by Eq. (4). In Eq. (4), σ_{op} represents the optical intersubband cross section in square centimeter and reads as

$$\sigma_{\mathrm{op}} = \frac{\pi q^2 n_g}{\epsilon_0 \epsilon_r c}|Z_{23}|^2 \omega \mathcal{L}(\hbar\omega - E_{23}), \tag{5}$$

where c/n_g is the group velocity in the optical waveguide, Z_{23} is the matrix dipole moment between the lasing states, ω

is the laser angular frequency, $\mathcal{L}(\cdots)$ is a normalized Lorentzian function, q is the elementary charge, and $\epsilon_0\epsilon_r$ is the permittivity of the active region. The model presented in this section does not calculate the laser-induced coherence terms; instead, it assumes a voltage-independent Lorentzian function for the gain profile. This simplification is convenient to quickly estimate the total stimulated emission lifetime inside the cavity, the emitted optical power, and hence the change in electric conductivity when lasing occurs. Analytical expressions of these quantities are given in the Appendix with some approximations.

The exponential decay of the coherence terms, ρ_{ij} $(i \neq j)$, is characterized by the total phase coherence time constant, $\tau_{\parallel ij}$, which describes how long the phase correlation is conserved during the tunneling transport from state i to j. The total phase loss rate between two interacting states i and j is half the sum of population decay rates via intersubband processes, $\tau_{i,j}^{-1}$, plus the "pure" phase randomization rate during tunneling transport or optical intersubband transition, $\tau_{ij}^{*^{-1}}$.[11] The higher $\tau_{\parallel ij}^{-1}$, the more relaxed becomes the $\mathbf{k}_\parallel$-conservation law. For the sake of simplicity, we assumed the same pure dephasing time constant for the five different tunneling channels (ij=12, 13, …, etc.): $\tau_{ij}^{*}=\tau^{*}$. The assumption is that the scattering potentials and matrix scattering elements for different pairs of states would be similar. We recall that the intrasubband component of the linewidth at a particular kinetic energy, $\epsilon_{\mathbf{k}_\parallel}$, and for a scattering potential, $V^{(\mathrm{sc})}$, can be written[38] as

$$\tau_{ij}^{*^{-1}}(\epsilon_{\mathbf{k}_\parallel}) = \pi \sum_{\mathbf{q}_\parallel} |\langle i,\mathbf{k}_\parallel + \mathbf{q}_\parallel | V^{(\mathrm{sc})} | i,\mathbf{k}_\parallel\rangle - \langle j,\mathbf{k}_\parallel + \mathbf{q}_\parallel | V^{(\mathrm{sc})} | j,\mathbf{k}_\parallel\rangle|^2 \times \delta(\epsilon_{\mathbf{k}_\parallel + q_\parallel} - \epsilon_{\mathbf{k}_\parallel}). \quad (6)$$

As an example, in the following discussion we consider the ionized impurity scattering from Si dopants introduced in the phonon well. With this scattering potential, $V^{(\mathrm{sc})}=V^{(\mathrm{imp})}$, we see that the correlated intraband scattering rate between two states should substantially change whether the states are spatially separated, as in a tunneling process, or overlapping each other, as in an optical intersubband transition in the active double well [see Eq. (6)]. In the latter case, the remoteness of both wave functions from the ionized impurity potential and the strong overlap of the two wave functions imply that the two matrix elements in Eq. (6) tend to cancel each other. In the former case, the matrix element relative to a state in the phonon well, such as $\langle 1,\mathbf{k}_\parallel + \mathbf{q}_\parallel | V^{(\mathrm{imp})} | 1,\mathbf{k}_\parallel\rangle$, is compensated, to a lesser degree, by the matrix element relative to a state in the active double well, such as $\langle 2,\mathbf{k}_\parallel + \mathbf{q}_\parallel | V^{(\mathrm{imp})} | 2,\mathbf{k}_\parallel\rangle$. Therefore, a different pure dephasing time constant is used between the two lasing states: $\tau_{23}^{*} \neq \tau^{*}$. From the above discussion, with Eq. (6) and the ionized impurity scattering example, we predict $\tau_{23}^{*} > \tau^{*}$. The different total phase loss rates read

$$\tau_{\parallel ij}^{-1} = (2\tau_i)^{-1} + (2\tau_j)^{-1} + (1 - \delta_{ij}^{23})/\tau^{*} + \delta_{ij}^{23}/\tau_{23}^{*}, \quad (7)$$

where δ_i^j stands for the Kronecker delta. Having a temperature-independent pure dephasing time constant means that the broadening of tunneling or optical intersubband transition resonances with temperature will be minimal due to the thermal activation of τ_2.

We refer the reader to the Appendix for the explicit formulas of the approximate solutions of Eq. (3). The solutions depend on different tunneling times between states i and j, T_{ij}, which are defined as[21]

$$T_{ij} = \frac{1 + \Delta_{ij}^2 \tau_{\parallel ij}^2}{2\Omega_{ij}^2 \tau_{\parallel ij}}, \quad (8)$$

where $\Delta_{ij}=E_{ij}/\hbar$ is the detuning in per picosecond. After some algebra, the expression of current density, J, becomes

$$\frac{J}{qN_{\mathrm{2D}}} = 2\Omega_{12}\Im(\rho_{12}) + 2\Omega_{13}\Im(\rho_{13}) + 2\Omega_{14}\Im(\rho_{14}) + 2\frac{Z_{22} - Z_{33}}{L_{\mathrm{sp}}}[\Omega_{24}\Im(\rho_{24}) - \Omega_{12}\Im(\rho_{12})] - 2\frac{Z_{23}}{L_{\mathrm{sp}}}\tau_{\parallel 23}^{-1}\Re(\rho_{32}), \quad (9)$$

where N_{2D} is the two-dimensional carrier density, $L_{\mathrm{sp}} \approx Z_{11} - Z_{44}$ is the superperiod of the active region, $Z_{22} - Z_{33}$ is the static dipole between the lasing states, $\Im(\cdots)$ and $\Re(\cdots)$ are the imaginary and real components of a complex entity. The first three terms represent the three tunneling currents departing from the injector state 1: injection Ω_{12}, wrong injection Ω_{13}, and the negligible parasitic Ω_{14} channels. The fourth term is only a small correction which takes into account the different centroids of the two lasing states in the active double well. In the design of Ref. 25, the static dipole is small, and ranges from −5 nm at 1 kV/cm to 2.6 nm at 14 kV/cm, because the optical intersubband transition is vertical. The physical origin of last term which is directly related to the tunneling induced coherence between the lasing states is not as clear. In the next paragraph, we will see that this coherence is induced by the leakage paths and the so-called cross terms. At most, the last two terms of Eq. (9) contributes to ~5% of the total current. The first two terms are dominant in this expression.

When writing the time evolution of the coherence terms in Eq. (3), we notice the presence of cross terms between coupling strengths and coherence terms, i.e., terms in such a form $\Omega_{ij}\rho_{nm}$, where $nm \neq ij$ or ji. In the final steady-state solution, these terms are responsible for two non-negligible coherence terms between states 2 and 3 and between 1 and 4, despite the fact the parasitic coupling, $\hbar\Omega_{14}$, is in the microelectron volt range. This is because several tunneling trajectories bridge these states, which otherwise would not be coupled if leakage paths were absent. For instance, states 2 and 3 develop a static coherence by the tunneling trajectories $2 \rightarrow 4 \rightarrow 3$ or $2 \rightarrow 1 \rightarrow 3$. These cross terms are responsible for the prethreshold shoulder in current-voltage characteristic that is observed in three-well designs (at ~8–9 V).[4,26,39] The tunneling times of the four main channels are comparable, on the order of 1–10 ps, when levels 1 and 4 are both approximately equidistant in energy from the two lasing states. For instance, at 8.5 kV/cm the adiabatic solution of the three-well structure shows a strong mixing between levels 1 and 4 with an anticrossing energy of 0.7 meV. This

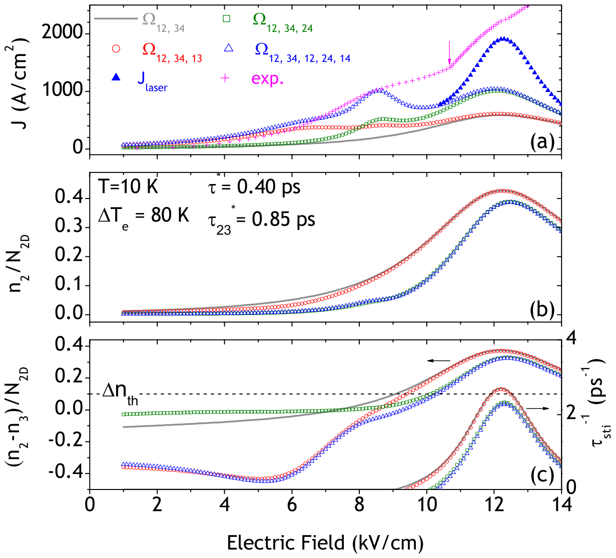

FIG. 4. (Color online) Simulation of (a) current density; (b) upper laser state population ρ_{22}; (c) population inversion $\rho_{22}-\rho_{33}$ (left vertical axis); and stimulated emission rate τ_{sti}^{-1} (right vertical axis) without the laser-induced coherence terms in the model. The lattice temperature is 10 K, the electron heating temperature, ΔT_e, is set constant at 80 K, the pure dephasing time constant in tunneling is $\tau^*=0.4$ ps, and in optical intersubband transition is $\tau_{23}^*=0.85$ ps. The different leakage tunneling paths are artificially turned on or off to visualize their effects. In the three panels, the set of gray solid lines corresponds to the situation of no leakage paths, so only Ω_{12} and Ω_{34} are nonzero. The red lines with circles include the correct injection, extraction, and effect of the wrong injection tunneling, Ω_{13}. The green lines with squares include the correct injection, extraction and effect of the wrong extraction tunneling, Ω_{24}. Finally, the set of blue lines with empty triangles includes the five tunneling paths. The blue line with solid triangles in panel (a) represents the simulated current density of a laser after threshold for a threshold population inversion of 10%. The magenta curve with crosses represents the experimental data at 4.2 K of a metal-metal waveguide lasing device which structure is similar to the one reported in Ref. 25. The kink (pointed by an arrow) of the current density at 10.7 kV/cm and 1.4 kA/cm^2 corresponds to lasing threshold.

strong *indirect coupling* between the injector and extractor states will be a source of a leakage path. However, since the carriers falling on the extractor level are meant to be scattered efficiently to the next period injector level, the indirect leakage at 8.5 kV/cm does not alter the population inversion significantly.

Simulated current densities, upper lasing state populations, population inversions, and stimulated emission rates in the cavity are plotted in Fig. 4. It shows different cases, where leakage tunneling paths are artificially turned on or off, in order to visualize their impact. The lattice temperature is set at 10 K. The values used in the simulations for the electron heating temperature (80 K), pure dephasing time constant in tunneling (0.4 ps) and optical intersubband transition (0.85 ps) were inferred by the fitting exercise discussed later in Sec. VI. Panel (a) shows that the current density around the design electric field increases significantly by the wrong extraction channel, 2–4, whereas the effect of the wrong injection, 1–3, is not as severe. However, below threshold, the presence of all tunneling channels is important to model the large current increase by the aforementioned 1–4 indirect coupling. For comparison, the electrical characteristics at 4.2 K of a 0.1×1 mm^2 metal-metal ridge waveguide laser with a structural design identical to that reported in Ref. 25 is plotted. Both contacts of this device were Ohmic by using an e-beam sputter deposition of Pd/Ge/Ti/Pt/Au with the thickness sequence of 550/1000/250/550/5000 Å, followed by a 15 s, 380 °C rapid thermal annealing. It is unfortunate that we could not report an electrical measurement with Ohmic contacts from Ref. 25's wafer—as this material was exhausted—because the later "clone" wafer shows a higher threshold and a lower maximum operating temperature. Nevertheless, our simulations with this simplified model are consistent with the experimental data: the order of magnitude for the peak current density is well predicted and the current density at the shoulder (~8.5 kV/cm) is close to experimental value.

Figure 4(b) shows the depletion of the upper lasing state by the wrong extraction path, 2–4, which results in a 14% decrease in the maximum population inversion at the design electric field, as shown in Fig. 4(c). The last panel shows that the wrong injection channel, 1–3, is very active at its resonance (5.4 kV/cm), but loses its strength at the design electric field. The consequence of this channel is to increase the transparency electric field from 7.4 without leakage to 8.6 kV/cm. The rapid decrease in strength of this channel at higher electric field causes the laser threshold electric field to be increased slightly from 9.1 to 9.4 kV/cm. The maximum population inversion and laser emission are not affected by this channel. If only the wrong extraction channel, 2–4, is turned on, the maximum population inversion is affected. The laser threshold electric field in this case increases significantly to 10.2 kV/cm. However at low fields, below transparency, the population inversion is improved compared to the zero leakage situation, because the carriers injected to level 2 are quickly diverted to level 4 rather than been scattered to the lower lasing state: only few carriers accumulate on level 3. When the two wrong channels are included, the transparency and the laser threshold electric fields are even more increased as a consequence of the depletion of the upper lasing state to the collector state. The maximum population inversion is affected and hence the laser emission. Before the transparency, i.e., below 9.3 kV/cm, the population inversion is strongly negative, mainly by the wrong injection channel. This is consistent with the observation of strong THz absorption lines from the lower lasing state measured, below threshold, by time-domain spectroscopy.[40]

Without leakage channels, the population inversion is simply given by

$$\Delta\rho|_{\Omega_{13}=\Omega_{24}=0} = \frac{\tau_2 - T_{34} - \tau_4}{T_{34} + T_{12} + 2\tau_2 + 2\tau_4} \tag{10}$$

and is represented by a green line in Fig. 4(c). As mentioned in the discussion of Eq. (1), the upper lasing state lifetime, τ_2, is strongly reduced at higher temperature. On the other side, the "effective lifetime" of the lower lasing state,[23] $\tau_{\mathrm{eff},3}=T_{34}+\tau_4$, is weakly temperature dependent by the nature of the resonant phonon depopulation scheme. Therefore, it cannot compensate for the thermally activated τ_2. In other words, the depopulation scheme cannot sustain the increasing flow of carriers originating from the shortening of the upper lasing lifetime with temperature. This bottleneck at the depopulation stage explains the degradation of population inversion with temperature. In addition, this simple expression shows that the population inversion becomes less temperature sensitive as $\tau_{23}^{(\mathrm{LO\ emi,\ hot}\ e^-)}$ increases. This can be achieved with diagonal 2–3 transitions.

From the above discussion, we concluded that the wrong injection channel, in this particular design, does not severely limit the laser operation. We also showed that the 1–4 indirect coupling is only a significant source of leakage at its resonance; and that the wrong extraction channel increases the current density and slightly reduces the maximum population inversion. Therefore, the latter leakage channel is detrimental for cw operation and explains why the three-well design, with vertical transitions, does not work in this mode.[41] The channel 2–4 is more to blame than channel 1–3, because of the high coupling strength Ω_{24}, and also, because of the high total phase loss rate of this tunneling, $\tau_{\parallel 24}^{-1}$, which is limited by the phonon resonant scattering τ_4 [see Eq. (7)]. The tunneling resonance T_{24} will be broader than T_{13} and reasonably fast at the design electric field, which means that the depletion $2 \rightarrow 4$ will be still active. In other words, at the design electric field, the upper lasing state population does not recover fully from the 2–4 resonance taking place at 4.6 kV/cm. Of course, the effect of this shunt channel is more pronounced at higher upper lasing state lifetime, τ_2, i.e., at low temperature.

Figure 5(a) shows the four main tunneling times versus electric field. At the design electric field, T_{24} is about ~4 ps, while T_{13} is larger than 100 ps and hence cannot have a large influence. Around 8.5 kV/cm, the same panel shows that the four tunneling times are comparable within an order of magnitude, which makes the building of a coherence between 1–4 possible. The range of electric fields for which these tunneling times are comparable is narrow, and this explains the small finesse (± 1.5 kV/cm) of the simulated indirect resonance of $|\rho_{14}|$. Panel (b) of Fig. 5 shows the subband populations in the nonlasing and lasing conditions. In the lasing condition, the threshold population inversion, $\Delta\rho_{th}$, is fixed at 10% of the total number of carriers per period. We see that the population of the injector state, ρ_{11}, quickly recovers from the resonance of the wrong injection, T_{13}, shown in Fig. 5(a) at ~5.4 kV/cm. Near the resonance of the wrong extraction channel (4.6 kV/cm), level 4 is more populated than level 2, due to a combination of a strong resonance (high Ω_{24}) and other channels that also supply the extractor state. We see that all four populations fluctuate around 8.5 kV/cm, where all four states are coupled to each other to form the 1–4 indirect resonance.

To illustrate the effect of the wrong extraction path near the design electric field, we give approximate formulas that are derived from Eqs. (A2)–(A6) in the Appendix, with the assumption of an infinitely long tunneling time T_{13}. The transit time reads as

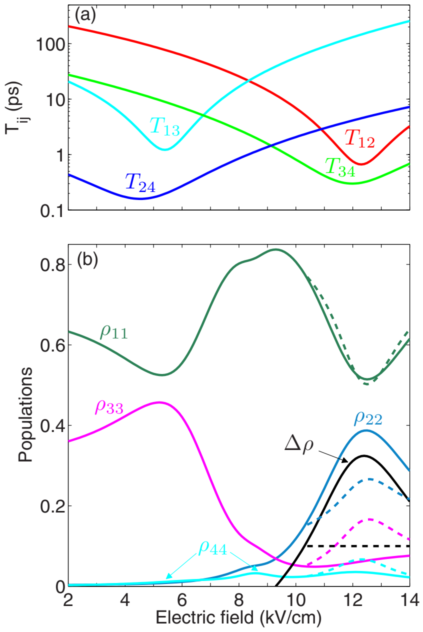

FIG. 5. (Color online) Panel (a) shows the four main tunneling times T_{ij} as defined by Eq. (8). Panel (b) shows the populations of all the states and the population inversion on a nonlasing device (solid lines) and on a lasing device with a threshold population inversion of 10% (dashed lines). Simulations are performed with the same parameters as Fig. 4.

$$\tau_{\mathrm{transit}} = \frac{T_{24}}{T_{24}+\tau_2}(T_{34}+T_{12}+2\tau_2+2\tau_4) + \frac{(T_{12}\tau_2+T_{34}\tau_4+4\tau_2\tau_4)}{T_{24}+\tau_2} \tag{11}$$

and the product between the population inversion and the transit time becomes

$$\Delta\rho\tau_{\mathrm{transit}} = \frac{T_{24}}{T_{24}+\tau_2}(\tau_2-T_{34}-\tau_4) - \frac{T_{34}\tau_4}{T_{24}+\tau_2}. \tag{12}$$

The product $\Delta\rho\tau_{\mathrm{transit}}$ can be viewed as the current efficiency for population inversion, in which the factor $T_{24}/(T_{24}+\tau_2)$ represents the detrimental shunting effect by the wrong extraction channel, and should be minimized, if possible. This observation brings an additional argument for a three-well QCL design based on *diagonal* transitions which demonstrated the record $T_{\max}$, a rather low threshold current density[4] and even cw operation.[23]

The objective of this section was to visualize the transfer of charges versus the electric field with and without the laser field. The model has used a voltage-independent Lorentzian gain profile [$\mathcal{L}$ in Eq. (5)]. In the next section, we will see how the gain profile is strongly modified by the insertion of the laser interaction, Ω_L, in the Hamiltonian.

V. GAIN PROFILE MODIFIED BY FOUR LASER INDUCED COHERENCE TERMS

A. Derivation of the gain

When introducing the laser Rabi frequency, Ω_L, in the time evolution equation of the density matrix, we obtain

$$\frac{d\rho}{dt} = \frac{1}{\jmath\hbar}\left[\begin{pmatrix} E_1 & \hbar\Omega_{12} & \hbar\Omega_{13} & \hbar\Omega_{14} \\ \hbar\Omega_{12} & E_2 & \Omega_L\cos(\omega t) & \hbar\Omega_{24} \\ \hbar\Omega_{13} & \Omega_L\cos(\omega t) & E_3 & \hbar\Omega_{34} \\ \hbar\Omega_{14} & \hbar\Omega_{24} & \hbar\Omega_{34} & E_4 \end{pmatrix}, \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & \rho_{22} & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & \rho_{33} & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & \rho_{44} \end{pmatrix}\right]$$
$$- \begin{pmatrix} \tau_1^{-1}\rho_{11} - \tau_4^{-1}\rho_{44} & \tau_{\|12}^{-1}\rho_{12} & \tau_{\|13}^{-1}\rho_{13} & \tau_{\|14}^{-1}\rho_{14} \\ \tau_{\|12}^{-1}\rho_{12} & \tau_2^{-1}\rho_{22} - \tau_3^{-1}\rho_{33} & \tau_{\|23}^{-1}\rho_{23} & \tau_{\|24}^{-1}\rho_{24} \\ \tau_{\|13}^{-1}\rho_{31} & \tau_{\|23}^{-1}\rho_{32} & \tau_3^{-1}\rho_{33} - \tau_2^{-1}\rho_{22} & \tau_{\|34}^{-1}\rho_{34} \\ \tau_{\|14}^{-1}\rho_{41} & \tau_{\|24}^{-1}\rho_{42} & \tau_{\|34}^{-1}\rho_{43} & \tau_4^{-1}\rho_{44} - \tau_1^{-1}\rho_{11} \end{pmatrix}. \tag{13}$$

The laser field which oscillates at the angular frequency ω, can potentially induce oscillating coherence terms between states that have energy spacing close to $\hbar\omega$. Provided the quality factor of the 2–3 resonance is high enough and the injection and extraction detunings are small, or in other words, if $\Delta_{12,34}^2 + \tau_{\|23}^{-1} \gg \Delta_{23} + \omega$, the rotating wave approximation is justified. Keeping in mind that in Sec. IV, we found static coherence terms between any combination of states in the absence of laser field, we look for solutions in the form of

$$\rho_{ij} = [\omega \approx \Delta_{ij}]\tilde{\rho}_{ij}e^{-\jmath\omega t} + \rho_{ij}^{(0)}, \tag{14}$$

where $\rho_{ij}^{(0)}$ is the static tunneling induced coherence between states i and j, $\tilde{\rho}_{ij}$ is the laser-induced coherence between the same states and the bracket is a logical operator which equals 1 (0) if the expression inside is true (false). Here, the population and coherence terms are calculated for a vanishing Rabi frequency, which implies that our solutions are valid below threshold and slightly above. As expected, when we bring Ω_L to zero, we find that the static terms of the density matrix, $\rho_{ij}^{(0)}$, fulfill the Eq. (3), with a vanishing stimulated emission rate, $\tau_{\text{sti}}^{-1} = 0$. As will be briefly discussed in the Appendix, if the laser intensity is not negligible, the right-hand side of Eq. (A1) (which models the static terms $\rho_{ij}^{(0)}$) is changed and no simple analytical solution can be derived. Kumar[22] showed numerical solutions of gain broadening by a strong stimulated emission rate in a three-level system with one injection tunneling channel. The four oscillating coherence terms fulfill a simple system of linear equations

$$\frac{\Omega_L}{2}\begin{pmatrix} \rho_{12}^{(0)} \\ 0 \\ \rho_{22}^{(0)} - \rho_{33}^{(0)} \\ -\rho_{34}^{(0)} \end{pmatrix} = \begin{pmatrix} \Gamma_{13} & -\Omega_{34} & \Omega_{12} & 0 \\ -\Omega_{34} & \Gamma_{14} & 0 & \Omega_{12} \\ \Omega_{12} & 0 & \Gamma_{23} & -\Omega_{34} \\ 0 & \Omega_{12} & -\Omega_{34} & \Gamma_{24} \end{pmatrix} \times \begin{pmatrix} \tilde{\rho}_{13} \\ \tilde{\rho}_{14} \\ \tilde{\rho}_{23} \\ \tilde{\rho}_{24} \end{pmatrix}, \tag{15}$$

where Γ_{ij} is the complex detuning of states i and j with respect to the laser frequency

$$\Gamma_{ij} = (\Delta_{ij} - \omega) - \jmath\,\tau_{\|ij}^{-1}. \tag{16}$$

The solution of Eq. (15) is found as

$$\begin{aligned} \frac{\tilde{\rho}_{23}}{\Omega_L/2} = {} & (\rho_{22}^{(0)} - \rho_{33}^{(0)})\frac{\Gamma_{13}\Gamma_{14}\Gamma_{24} - \Omega_{12}^2\Gamma_{13} - \Omega_{34}^2\Gamma_{24}}{(\Omega_{34}^2 - \Omega_{12}^2)^2 - \Omega_{34}^2(\Gamma_{13}\Gamma_{14} + \Gamma_{23}\Gamma_{24}) - \Omega_{12}^2(\Gamma_{13}\Gamma_{23} + \Gamma_{14}\Gamma_{24}) + \Gamma_{13}\Gamma_{14}\Gamma_{23}\Gamma_{24}} \\ & + \rho_{12}^{(0)}\Omega_{12}\frac{\Omega_{12}^2 - \Omega_{34}^2 - \Gamma_{14}\Gamma_{24}}{(\Omega_{34}^2 - \Omega_{12}^2)^2 - \Omega_{34}^2(\Gamma_{13}\Gamma_{14} + \Gamma_{23}\Gamma_{24}) - \Omega_{12}^2(\Gamma_{13}\Gamma_{23} + \Gamma_{14}\Gamma_{24}) + \Gamma_{13}\Gamma_{14}\Gamma_{23}\Gamma_{24}} \\ & + \rho_{34}^{(0)}\Omega_{34}\frac{\Omega_{34}^2 - \Omega_{12}^2 - \Gamma_{13}\Gamma_{14}}{(\Omega_{34}^2 - \Omega_{12}^2)^2 - \Omega_{34}^2(\Gamma_{13}\Gamma_{14} + \Gamma_{23}\Gamma_{24}) - \Omega_{12}^2(\Gamma_{13}\Gamma_{23} + \Gamma_{14}\Gamma_{24}) + \Gamma_{13}\Gamma_{14}\Gamma_{23}\Gamma_{24}}, \end{aligned} \tag{17}$$

from which the gain/absorption coefficient, G_{op}, can be derived as

$$G_{\text{op}} = \frac{N_{3\text{D}}q^2|Z_{23}|^2 n_g\omega}{\epsilon_0\epsilon_r c\hbar}\Im\left(\frac{\tilde{\rho}_{23}}{\Omega_L/2}\right). \tag{18}$$

B. Numerical applications of gain spectra

To illustrate the effects of injection and extraction tunnelings on the gain, we computed the maximum gain versus the thicknesses of injection and extraction barriers, for two cases. First, we consider a voltage-independent Lorentzian

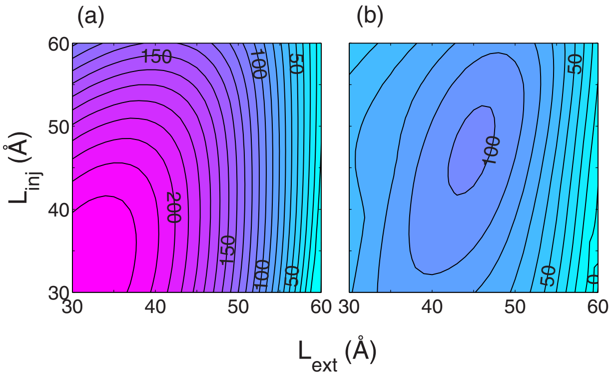

FIG. 6. (Color online) Simulation results for the maximum gain (in per centimeter) as a function of injection and extraction barrier thicknesses with $\tau^*=0.4$ ps, $\tau_{23}^*=0.85$ ps, $T=50$ K, and $\Delta T_e=80$ K. In panel (a), the gain spectrum is assumed to be a voltage-independent Lorentzian with a $(\pi\tau_{23}^*)^{-1}=0.375$ THz full width at half maximum. In panel (b) the complete gain model of Eq. (17) is used. For the sake of comparison, the same color scale is used in both panels.

gain profile with a $\hbar\tau_{\parallel 23}^{-1}$ half width at half maximum; this means that the coherence $\tilde{\rho}_{23}$ depends only on population inversion, $2\tilde{\rho}_{23}/\Omega_L=(\rho_{22}^{(0)}-\rho_{33}^{(0)})/\Gamma_{23}$. Second, we use the complete model summarized in Eq. (17). The results of this calculation are plotted in Fig. 6 for $\tau^*=0.4$ ps, $T=50$ K, $\Delta T_e=80$ K, and $\tau_{23}^*=0.85$ ps. Panel (a) shows that the maximum population inversion is obtained for ~ 33 Å thick barriers for the first case; however, the complete model [panel (b)] suggests further isolating the active double well from the phonon wells, with ~ 45 Å barriers. The full model recommends limiting the tunneling interferences by increasing the barriers, even if these larger barriers decrease the population inversion [as in Eq. (10)]. Getting the correct extraction barrier is more critical than having the correct injection barrier, as a thick extraction barrier can damage the depopulation mechanism. We also see that the peak gain with the complete model is significantly smaller than the one predicted with the Lorentzian model. The full gain spectra at 10 K for different electric fields, as well as the peak position and the two half width at half maximum points, are displayed in Fig. 7. The broadest spectra are obtained around the design electric field. At this bias, they show a 5 meV full width at half maximum, a value between $2\hbar\Omega_{34}$ and $2\hbar(\Omega_{12}+\Omega_{34})$. Below the design electric field, the gain shows a peak at a photon energy higher than E_{23}, and above the design electric it is redshifted relatively to E_{23}.

We are going to use a simple tight-binding model to illustrate the evolution of the photon energy at peak gain versus electric field. This very simplified model considers four possible transitions between the two doublets, also called dressed states, formed by the coupling of 1–2 and 3–4 states. The phase between these states is considered uncorrelated. Of course, the density-matrix formalism takes into account the phase relation between the states, which can induce interferences in the gain/absorption spectrum. For long phase coherence time constants, narrow linewidth symmetric S_{ij} and antisymmetric A_{ij} dressed states are formed by the mixing of states 1–2 on the injection side and states 3–4 on the extraction side. They can be written as

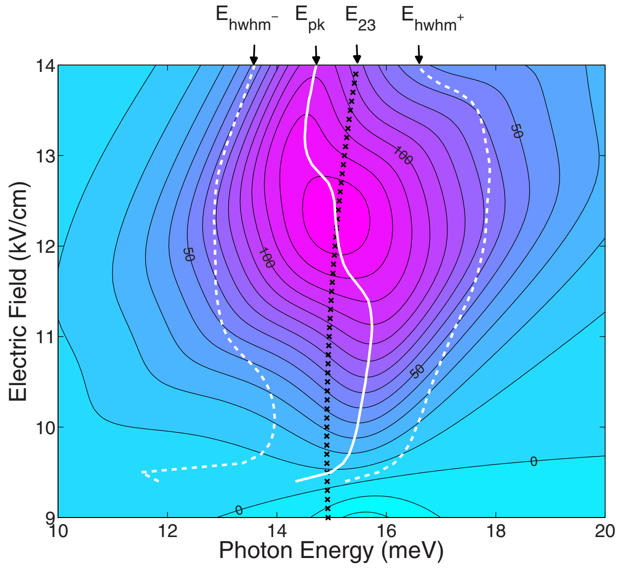

FIG. 7. (Color online) Contour plot of gain spectra for different electric fields. The lattice temperature is 10 K, the electron heating temperature, ΔT_e, is fixed at 80 K, the pure dephasing time constant in tunneling is $\tau^*=0.4$ ps, and in optical intersubband transition is $\tau_{23}^*=0.85$ ps. The line with crosses represents the position of E_{23} as a function of electric field (quadratic Stark effect). The white solid line represents the position of the peak gain. Relative to E_{23}, the peak gain is blueshifted before the design electric field, and redshifted after. The white dashed lines represent the position the two points at half width at half maximum. The full width at half maximum is 5 meV at 12.5 kV/cm. The unit of gain is per centimeter.

$$S_{ij}=\alpha_{ij}|i\rangle+\beta_{ij}|j\rangle, \tag{19}$$

$$A_{ij}=\beta_{ij}|i\rangle-\alpha_{ij}|j\rangle. \tag{20}$$

Defining θ_{ij} by $\tan(\theta_{ij})=2\Omega_{ij}/|\Delta_{ij}|$, the component α_{ij} will be $\sin(\theta_{ij}/2)$ [or $\cos(\theta_{ij}/2)$] and the component β_{ij} will be $\cos(\theta_{ij}/2)$ [or $\sin(\theta_{ij}/2)$] for positive (or negative) detuning. The electric field dependence of the lifetime of a dressed state, D_{ij}, where D is either A (antisymmetric) or S (symmetric), is modeled as

$$\tau_{D_{ij}}^{-1}=\alpha_{ij}\beta_{ij}\tau^{*^{-1}}+(\delta_D^S\beta_{ij}^2+\delta_D^A\alpha_{ij}^2)(2\tau_j)^{-1}. \tag{21}$$

This equation suggests that the tunneling pure dephasing rate, $\tau^{*^{-1}}$, contributes largely to the total broadening when the states are strongly mixed near resonance ($\alpha_{ij}\approx\beta_{ij}$). This model for the lifetime of the hybridized states is only qualitative. The broadening of the optical intersubband transition between states $D=A_{12},S_{12}$ and $D'=A_{34},S_{34}$ would be estimated like $\tau_{\parallel DD'}^{-1}=\tau_D^{-1}+\tau_{D'}^{-1}+\tau_{23}^{*^{-1}}$, and the oscillator strength would be proportional to the modulus square of the product of the components along states 2 and 3, $|\langle D|2\rangle\langle 3|D'\rangle|^2$. Figure 8 gives a simple picture of the mechanism explaining the voltage dependence of the peak frequency of the gain. Around the design electric field [panel (b)], the four transitions have comparable strengths and the resonance energies are on both sides of E_{23}: the total gain is centered around E_{23} and is broad. For large positive injection and extraction de-

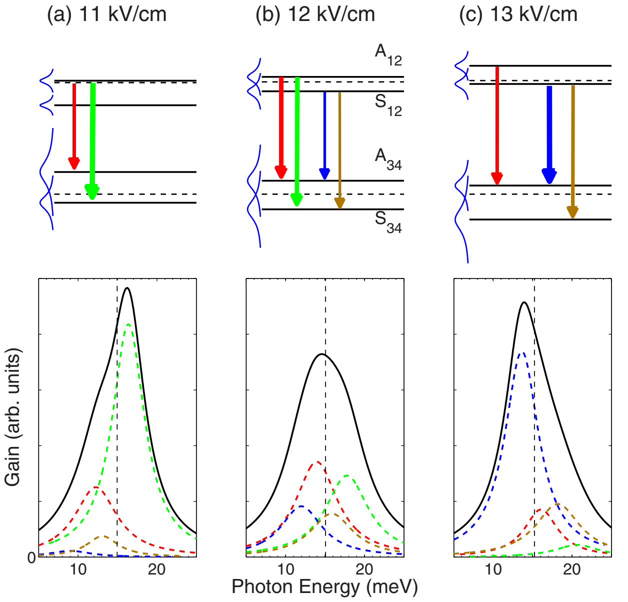

FIG. 8. (Color online) Simplistic tight-binding model of the gain spectrum (a) before, (b) around, and (c) after the design electric field. The lattice temperature is 10 K, the electron heating temperature, ΔT_e, is fixed at 80 K, the pure dephasing time constant in tunneling is $\tau^*=0.4$ ps, and in optical intersubband transition is $\tau_{23}^*=0.85$ ps. On the upper part of the figure, the symmetric and antisymmetric states are represented by horizontal solid lines and the estimated linewidth of these states are represented by vertical oriented Lorentzian peaks. The uncoupled states 2 and 3 are represented by horizontal dashed lines. On the upper and lower parts of the figure, the same color code scheme is used: red for $A_{12}\rightarrow A_{34}$, green for $A_{12}\rightarrow S_{34}$, blue for $S_{12}\rightarrow A_{34}$, and brown for $S_{12}\rightarrow S_{34}$. On the lower part, the contribution of each transition is represented by a dashed colored Lorentzian line and the sum of four transitions by a solid black line. The vertical dashed line shows the position of E_{23}.

tunings [panel (c)], the oscillator strength is concentrated mainly on $S_{12}\rightarrow A_{34}$ transition which is slightly redshifted with respect to E_{23}. For large negative detunings [panel (a)], the transition $A_{12}\rightarrow S_{34}$ is dominant and is blueshifted. A large positive (negative) extraction detuning decouples the final state A_{34} (S_{34}) from the short-lifetime state 4, and this explains why the spectrum is narrower below or above the design electric field.

C. Different components of the gain

Equation (17) can be decomposed into several terms that are associated with different physical processes. From this point on, the first term on the right-hand side of Eq. (17) (proportional to population inversion) is called the *linear* component of the gain. It is the dominant component and represents a strongly modified Lorentzian gain profile under the effect of coherent injection and extraction. The two other terms, which are not linked to the population inversion, are called the *nonlinear* components of the gain.

If the extraction coupling Ω_{34} and the extractor state 4 are ignored in Eq. (17), and if the nondesirable tunneling couplings, $\Omega_{13,24,14}$, are neglected, one finds the gain of a three-state system with one injection tunneling as demonstrated in Ref. 42,

$$\frac{\tilde{\rho}_{23}}{\Omega_L/2}=\frac{(\rho_{22}^{(0)}-\rho_{33}^{(0)})-(\rho_{11}^{(0)}-\rho_{22}^{(0)})\dfrac{\Omega_{12}^2}{\Gamma_{12}\Gamma_{13}}}{\Gamma_{23}-\dfrac{\Omega_{12}^2}{\Gamma_{13}}}, \tag{22}$$

where Γ_{12} is the complex detuning at the injection side, $\Delta_{12}-\jmath\,\tau_{\parallel 12}^{-1}$. It is interesting to note that this latter expression looks similar to the gain of a three-level system in the p configuration, in which the coherence between the two highest levels, 1 and 2, is caused by an external pump laser field [see Fig. 9(a)]. This system has been extensively studied for its potential in amplification without population inversion.[43] Recently, lasing by an intersubband Stokes Raman process was demonstrated inside a midinfrared (mid-IR) QCL acting like a pump.[44] The gain of this Raman laser is modeled as in Eq. (22). In our lasers, the only difference comes from the injection tunneling Rabi frequency which drives the coherence between states 1 and 2. The second term in the numerator in Eq. (22) (proportional to $\rho_{11}^{(0)}-\rho_{22}^{(0)}$) is associated with a stimulated anti-Stokes scattering process. For positive population difference, $\rho_{11}^{(0)}-\rho_{22}^{(0)}>0$, this nonlinear gain term is at maximum around $\hbar\omega\approx E_{13}$ and is highly dispersive. Similar to Ref. 42, the second term of Eq. (17) can be identified as a third-order process, which resonantly scatters carriers from the injector level 1 to lower lasing state 3 in presence of the "injection field" Ω_{12}.

Also, when ignoring Ω_{12} and the injector level 1 in Eq. (17), and neglecting the same nondesirable tunneling couplings, an expression very similar to Eq. (22) is derived as

$$\frac{\tilde{\rho}_{23}}{\Omega_L/2}=\frac{(\rho_{22}^{(0)}-\rho_{33}^{(0)})-(\rho_{33}^{(0)}-\rho_{44}^{(0)})\dfrac{\Omega_{34}^2}{\Gamma_{34}\Gamma_{24}}}{\Gamma_{23}-\dfrac{\Omega_{34}^2}{\Gamma_{24}}}, \tag{23}$$

where Γ_{34} is the complex detuning at the extraction side, $\Delta_{34}-\jmath\,\tau_{\parallel 34}^{-1}$. This expression reminds us the gain of a three-level system in the h configuration,[43] in which the coherence between the two lowest levels, 3 and 4, is prepared by a pump laser [see Fig. 9(b)]. For positive population difference, $\rho_{33}^{(0)}-\rho_{44}^{(0)}>0$, the nonlinear gain term is at maximum around $\hbar\omega\approx E_{24}$. The third term of Eq. (17) represents the stimulated scattering from the upper lasing state 2 to the extractor level 4 in presence of the "extraction field" Ω_{34}.

The expression of Eq. (17) is, therefore, simply modeling the "sum" of p and h configurations, where two coherence terms between two highest and two lowest energy states are electrically driven by tunneling [see Fig. 9(c)]. The effects of the nondesirable tunneling couplings, $\Omega_{13,24,14}$, on the gain are hidden in $\rho_{22}^{(0)}-\rho_{33}^{(0)}$, $\rho_{12}^{(0)}$, and $\rho_{34}^{(0)}$.

We find it instructive to show the relative strength of the three terms in Eq. (17). In panels (b)–(d) of Fig. 10, we display the contribution of each term to the total gain of panel (a). The input parameters used during the simulation are $T=10$ K, $\tau^*=0.4$ ps, $\tau_{23}^*=0.85$ ps, and $\Delta T_e=80$ K. If

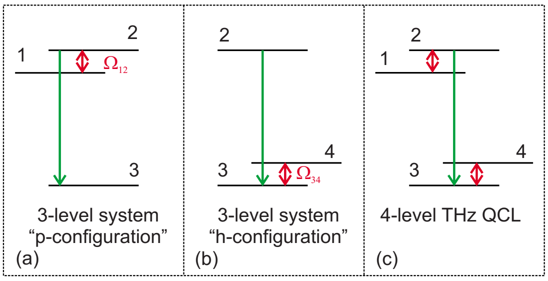

FIG. 9. (Color online) (a) Schematic representation of a three-level system in p configuration, where the coherence between the two highest states is conditioned by a field (laser, tunneling) with a coupling strength Ω_{12}. (b) Schematic of h configuration, where the coherence between the two lowest states is conditioned by a field Ω_{34}. (c) Schematic representation of a four-level system, such as the three-well THz QCL which can be viewed as the "sum" of p and h configurations.

we compare the total gain [panel (a)] and the linear gain [panel (b)], we see that the nonlinear gain of the second term [panel (c)] has an impact. This term gives an additional anti-Stokes gain at the photon energy around E_{13} and is strongly dispersive. The maximum value of the sum of two nonlinear terms is about $\sim 10\%$ of the maximum linear gain. This is small, but the relative contribution of nonlinear gain is actually higher below or above the design electric field, i.e., when the population inversion is not maximum. At the design electric field, the peaks of the linear and nonlinear gains coincide in energy, but the linear gain has a slight negative Stark effect (as in Fig. 8), while the nonlinear gain shows a strong positive Stark effect. This observation suggests that the nonlinear components can potentially change the total gain spectrum (shoulder or shift of the peak frequency) when levels are not aligned, i.e., below or above the design electric field. Our numerical results confirm that the nonlinear terms of the gain are more active in THz than in mid-IR QCLs because of the relatively long dephasing time. Unlike the mid-IR QCLs, the dephasing time is not limited by interface roughness scattering in THz devices.[42] In THz QCLs, the coupling strengths $\Omega_{12,34}$ are comparable to the dephasing rates $\tau_{\parallel ij}^{-1}(ij=12,13,34,24)$ and therefore, this enhances the contribution of the nonlinear gain. At 10 K, the third term is about one third of the second term because of the smaller coherence $\rho_{34}^{(0)} < \rho_{12}^{(0)}$, which comes from $\Omega_{34} > \Omega_{12}$ and $\tau_4 \ll \tau_2$ and (as shown in the Appendix).

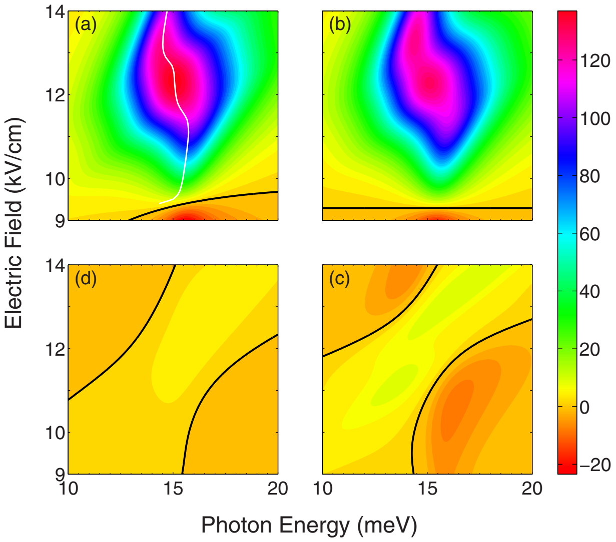

FIG. 10. (Color online) Four contour plots of the total gain (a) at 10 K, showing all three components as decomposed in Eq. (17). The first term depending on $(\rho_{22}^{(0)} - \rho_{33}^{(0)})$ is displayed in panel (b), the second term depending on $\rho_{12}^{(0)}$ in panel (c), and the third term depending on $\rho_{34}^{(0)}$ in panel (d). The dispersive nonlinear gain in panel (c) is strong enough to change the linear gain contour [panel (b)] into a different total gain contour [panel (a)]. The constant gain lines at 0 cm^{-1} are displayed by a solid black line. The same parameters as in Fig. 7 are used in the simulations. The thin white line in panel (a) shows the position of the peak gain versus electric field. Generally, the total gain is characterized by a negative Stark effect, i.e., a decrease in peak frequency with electric field. The unit of gain is per centimeter.

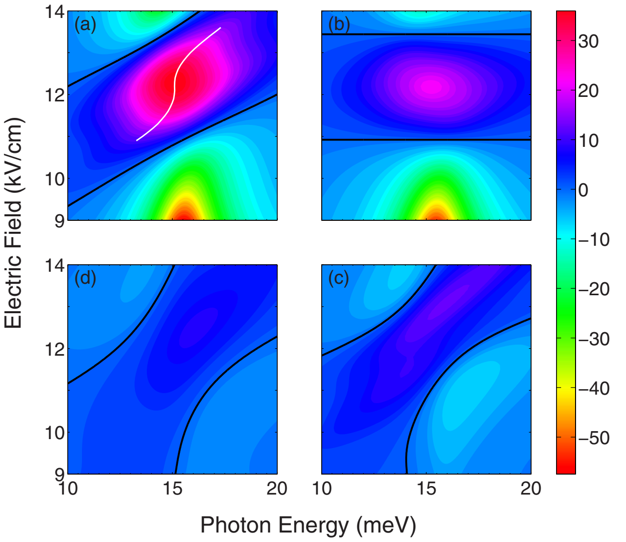

FIG. 11. (Color online) Contour plots of the total gain and its three components, as in Fig. 10, but at $T=140$ K. The white line on panel (a) shows the positive Stark effect of the peak gain, i.e., an increase in peak frequency with electric field.

In Fig. 11 we show that, as the temperature is raised to 140 K and the population inversion is lowered, the relative contribution of the two nonlinear gain terms increases significantly. The same figure shows that below (above) 10.9 (13.5) kV/cm and for photon energies lower (higher) than E_{23}, the gain is positive, while the population inversion is negative. Close to the maximum operating temperature, the THz QCL works in a regime where the nonlinear gain plays an important role. At 140 K, the peak gain is blueshifted with the electric field [Fig. 11(a)] because of the stronger contribution of nonlinear terms. At 10 K, the opposite situation is expected, and the peak is mostly redshifted [Fig. 10(a)] by the large contribution of the linear component. Therefore, measurements of the voltage dependence of spectra for different temperatures might be able to confirm the contribution of the nonlinear gain.

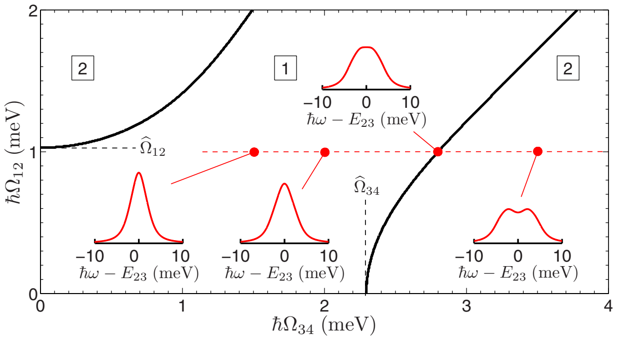

FIG. 12. (Color online) The "phase" diagram of number of peaks in the spectrum of the imaginary part of the first term of Eq. (17) versus the extraction and injection couplings. The number of peaks expected are indicated in square boxes. The calculation is performed for perfect alignment of the states at the injection and extraction sides, $\Delta_{12}=\Delta_{34}=0$. The parameters used are $\tau^{*}=0.4$ ps, $\tau_{23}^{*}=0.85$ ps, and $\tau_2=2$ ps. Four examples of these spectra are given at $\Omega_{12}=1$ meV for different extraction couplings $\Omega_{34}=1.5$, 2, 2.8, and 3.5 meV. At $\Omega_{34}=2.8$ meV, the linear gain is at the boundary of having one or two peaks. For comparison, the sketches in the insets are plotted with the same vertical scale, and broadening of the gain by the extraction coupling is apparent. The limit values of the coupling strengths to obtain a double-peaked gain when the other coupling is absent, $\hat{\Omega}_{ij}$, are also shown [Eqs. (24) and (25)].

D. Criterion for double-peaked gain

As Ref. 23 explained, the large gain linewidth mainly originates from the extraction tunneling mechanism. More precisely, it comes from the anticrossing between 3 and 4, and from the fast "effective lifetime" for the lower lasing state, $\tau_{\text{eff},3}=T_{34}+\tau_4$, which also depends on the extraction coupling. In the same reference, a convenient criterion was given for a double-peaked gain around the design electric field. Several groups have observed dual-wavelength operation of THz QCLs with modes 2.5–3 meV apart.[23,45] This has been attributed to the anticrossing on the injection side[45] or on extraction side.[23]

Our analytical model can predict the conditions for a multipeaked gain spectrum. Figure 12 shows the number of peaks of the imaginary part of the first term of Eq. (17) (which represents the linear gain) as a function of the injection and extraction coupling strengths. This calculation was performed for perfect injection and extraction resonance conditions, i.e., $\Delta_{12}=\Delta_{34}=0$. This "phase diagram" shows that a double-peaked gain behavior occurs if at least one of the coupling strengths, Ω_{ij}, is larger than a limit value, $\hat{\Omega}_{ij}$. The higher the coupling on the injection or extraction side, the higher the ratio $\Omega_{ij}/\hat{\Omega}_{ij}$ on the other side is required to obtain a double-peaked linear gain. When two tunneling couplings from injector and extractor sides have comparable strengths, the gain spectrum shows a single peak. Qualitatively, this can be explained by the large broadening of the transitions and the small photon energy difference, $2|\Omega_{12}-\Omega_{34}|$, between the two dressed state transitions $A_{12}\rightarrow A_{34}$ and $S_{12}\rightarrow S_{34}$ [Fig. 8(b)]. However, as the coupling strengths become unbalanced as to favor a side, the double-peak behavior associated with either the injection or the extraction appears. This means that the coherent injection or extraction is sufficiently strong to reveal the existence of its associated dressed states in the gain spectrum.

In the case where the resonance conditions are not perfectly met, $\Delta_{12}\neq 0$ and/or $\Delta_{34}\neq 0$, we intuitively expect the double-peak behavior to appear for larger $\hat{\Omega}_{ij}$. However, the simulation results for nonaligned states show that the phase diagrams become more complex than just an increase in $\hat{\Omega}_{ij}$.

To understand this behavior, we further investigate the linear gain of a three-level system in the p and h configurations. In either of these cases, only the tunneling coupling at one of the barriers is considered, in resonance condition. One can demonstrate that the imaginary part of the linear gain in Eqs. (22) and (23) will show a double-peak behavior when

$$\Omega_{12} > \hat{\Omega}_{12} = \frac{\tau_{\parallel 13}^{-1}}{\sqrt{2+\dfrac{\tau_{\parallel 13}}{\tau_{\parallel 23}}}} \tag{24}$$

and

$$\Omega_{34} > \hat{\Omega}_{34} = \frac{\tau_{\parallel 24}^{-1}}{\sqrt{2+\dfrac{\tau_{\parallel 24}}{\tau_{\parallel 23}}}}. \tag{25}$$

These equations impose a strong coupling between the interacting states to overcome the total phase loss rate $\tau_{\parallel 13}^{-1}$ and $\tau_{\parallel 24}^{-1}$. As expected, the coupling limit on the extraction side, $\hat{\Omega}_{34}$, is larger than that of injection, $\hat{\Omega}_{12}$, due to the resonant phonon scattering from level 4. The strong coupling ensures a bigger splitting between dressed states A_{ij} and S_{ij}; and hence the transitions between each of these states to the third level can be distinguished. From simplistic arguments, in a four-level system like the one represented in Fig. 8(b) or 9(c), the double-peaked gain should then appear when $|\Omega_{12}-\Omega_{34}|$ reaches $\hat{\Omega}_{12}+\hat{\Omega}_{34}$. This criterion is not accurate, as Fig. 12 shows that a double-peaked gain occurs before this simplistic criterion is met. The plots displayed as insets of Fig. 12 show the spectra of the imaginary part of the linear term of $\tilde{\rho}_{23}$ at $\Omega_{12}=1$ meV, and for four different extraction couplings $\Omega_{34}=1.5$, 2, 2.8, and 3.5 meV. The first two have a single-peaked behavior, the third is marginally double peaked, and the last one shows a double-peaked gain. These insets illustrate the broadening of linear gain and the decrease in the peak gain by the extraction coupling strength. Typically, our three-well QCLs are designed for $\Omega_{12}\sim 1$ meV and $\Omega_{34}\sim 2$ meV, therefore they should show a single-peaked linear gain.

We have here estimated populations, coherence terms, and gain, within the framework of a simplified density matrix model; and in the next section we will compare our model with published experimental data on three-well THz QCLs.

VI. COMPARISON OF THE MODEL WITH EXPERIMENTAL DATA

In order to estimate the pure dephasing time constant and the electron heating temperature, we compare the results of our density-matrix model with the experimental data on laser performance versus injection barrier thickness.[26] We could have chosen a comparison with the data versus extraction barrier thickness[27] but this latter study is not as comprehensive as for the injection barrier. In addition, the thin active region of the wafer used in the extraction barrier study could enhance the sensitivity of the fit to the THz permittivity of the lossy metals used in the surface plasmon waveguides.

Here, the 200-μm-wide and 1-mm-long ridge lasers reported in the injection barrier study have been remeasured with shorter current pulses (200 ns) and with better control of the electrical reflections. Compared to our data published in Ref. 26, we observed an increase in the maximum operating temperature on the samples with a thin injection barrier.

We fitted the experimental maximum operating temperature with our density-matrix model and compared the predicted threshold current density with the experiment. For various pure dephasing time constants for the intersubband transition, τ_{23}^*, tunneling pure dephasing time constants, τ^*, and electron heating temperatures, ΔT_e, the maximum gain was computed for the different injection barrier thicknesses and maximum operating temperatures of the six series of devices studied in Ref. 26. The fitting procedure needs the total waveguide loss coefficient, which was estimated as $\alpha_M + \alpha_W \sim 40\ \text{cm}^{-1}$. This high loss is due to the high doping concentration in the active region, the thick 400 nm top n^+-GaAs contact layer doped to $3 \times 10^{18}\ \text{cm}^{-3}$, and the non-optimized metallic mirror of the double surface plasmon waveguide. In Ref. 26, we used 25 nm Ti/55 nm Pt/300 (1000) nm Au for the top (bottom) mirrors. Titanium being a poor conductor, the Ti/Pt layers were modeled in the waveguide calculation as one metal with a 70 000 cm^{-1} plasma frequency and a short 12 fs electron lifetime. At 40 cm^{-1}, the losses might be overestimated. As a result, the fitting procedure might lead to an overestimation of upper lasing state lifetime, and therefore to an underestimation of electron heating temperature. In the fitting procedure, τ^* and ΔT_e are adjusted to minimize the standard deviation of the maximum gain with respect to the estimated total waveguide loss. The fit is not perfect but acceptable. For a given τ_{23}^*, the minimum standard deviation with the fitted parameters τ^* and ΔT_e is about 9–10 % of the estimated waveguide loss.

Figure 13 shows the summary of the fitting procedure. The results are given for 40 and 35 cm^{-1} total waveguide loss. The tunneling pure dephasing time constant, τ^*, is centered at ~0.4 ps and depends weakly on the choice of τ_{23}^*. The electron heating temperature is more affected by the choice of τ_{23}^*, because a decrease in this parameter has to be compensated by a longer upper lasing state lifetime. Since the electron temperature of the lasing subbands can be as much as 100 K higher than the lattice,[35] our fitting exercise suggests an intersubband pure dephasing time constant on the order of $\tau_{23}^* \sim 1$ ps.

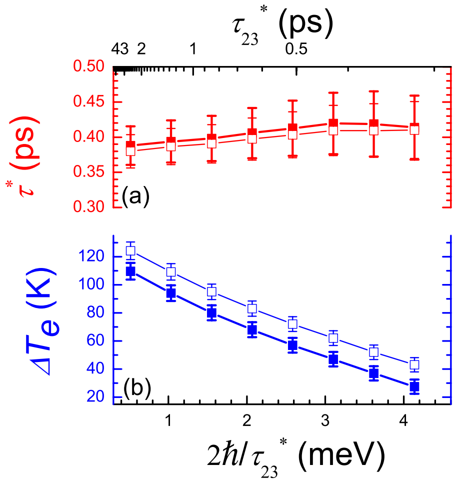

FIG. 13. (Color online) Fitted tunneling pure dephasing time constant, τ^*, (panel a) and electron heating temperature, ΔT_e, (panel b) as functions of the pure dephasing time constant in optical transition, τ_{23}^*. The vertical errors bars indicate the deviation of τ^* and ΔT_e which induce a 10% degradation of the fit. Filled (open) symbols are obtained with a 40 (35) cm^{-1} waveguide loss.

We recall that from the beginning, our model assumes a homogeneous broadening of the optical and tunneling transitions. As a result, if there is a period-to-period inhomogeneity in our structure, the fitted pure dephasing time constants are forced to adapt themselves to this increase in gain linewidth and hence, might be underestimated. With the fitted values of the pure dephasing time constants, the gain linewidth is ~1.2 THz at the design electric field (as in Fig. 7). Williams[46] has measured the electroluminescence linewidth on resonant-phonon depopulation devices and found it comparable to our predicted value. On the other hand, Jukam *et al.*[40] measured a 0.6 THz gain linewidth on the same design as the one under study in this paper. This disagreement could be due to different quality of growth in each of the samples.

With our fitted parameters, the simulation applied to the diagonal structure in Ref. 4 confirms, like in Ref. 23, the double-peak spectra at specified voltages. In this diagonal design, the nonlinear gain is weaker than in the vertical design (as in Fig. 1), due to its high coupling strengths and small current.

Figure 14 illustrates the fit with the parameters $\tau^* = 0.4$ ps, $\tau_{23}^* = 0.85$ ps, and $\alpha_M + \alpha_W \sim 40\ \text{cm}^{-1}$. If the simulation is performed as a function of the *extraction* barrier thickness with the same input parameters, the experimental maximum operating temperature points of Ref. 27 align on the $55 \pm 2.5\ \text{cm}^{-1}$ constant gain line. This should be compared with 64 cm^{-1}, the simulated total waveguide loss of a thin 100 period structure. Provided all parameters but the number of periods are the same between the two wafers used in the injection and extraction studies, the discrepancy between 64 cm^{-1}, the simulated waveguide loss with a 4.4 μm thin active region, and 55 cm^{-1}, the maximum gain using the fitted parameters derived form the injection barrier study, suggests that the model of the permittivity of the metals used in the waveguide calculation overestimates the loss. The effect of backfilling by phonon absorption from $1_{(n+1)}$ to 4 is not negligible. Indeed, if the phonon absorption is arbitrarily removed from the model, the fitted electron heating temperature would be ~10 K higher and the pure dephasing time constant would be ~0.33 ps.

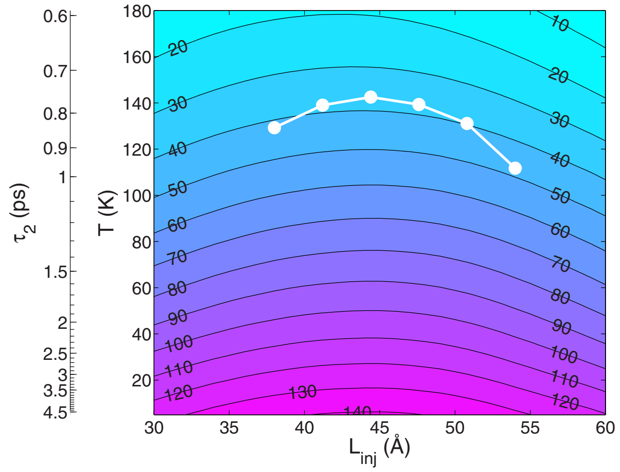

FIG. 14. (Color online) Contour plot of the maximum gain (in cm^{-1}) versus the thickness of injection barrier and lattice temperature for $\tau_{23}^*=0.85$ ps, $\Delta T_e=80$ K, $\tau^*=0.4$ ps. The white dots plot the redetermined maximum operating temperature of the six series of devices reported in Ref. 26, as explained in the text. The equivalent upper lasing state lifetime when the maximum gain is reached is given on the left-most vertical axis. At the six experimental points, the standard deviation of the maximum gain from the expected total waveguide loss 40 cm^{-1} is 3.7 cm^{-1}.

Finally in Fig. 15, we compare the simulated threshold current density at 10 K, and at $T_{\max}$ for 40 cm^{-1} total waveguide loss, with the experimental values. The agreement at 10 K is more satisfactory than at $T_{\max}$. At 10 K, the laser threshold is reached at around 10 kV/cm, an electric field for which the current is mainly limited by the injection tunneling time, rather than the upper lasing state lifetime, τ_2, or the choice of ΔT_e. At $T_{\max}$, the laser barely operates at the design electric field, i.e., close to negative differential resistance where the current depends greatly on τ_2, and therefore on the electronic temperature of subband 2. As the disagreement between experiment and simulation is more pronounced for thick barriers, the assumption of a constant ΔT_e for the six series of devices, independently of the current, might be too far from reality. We think that the electron-heating temperature could be lower for the thicker injection barrier devices. We are conscious there are limits to which this simplified model can be pushed to fit experimental data; however, we think it gives an indication of the high electronic temperature and the range of the pure dephasing time constants.

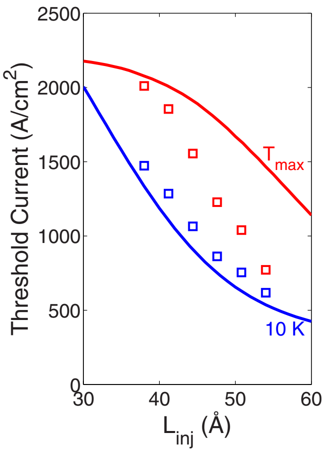

FIG. 15. (Color online) Comparison between theoretical threshold current densities (solid lines) and experimental points (open squares) at 10 K (blue line and squares) and at the simulated maximum operating temperature, $T_{\max}$, (red line and squares). The simulations are performed for $\tau_{23}^*=0.85$ ps, $\Delta T_e=80$ K, $\tau^*=0.4$ ps, and total waveguide loss $\alpha_M+\alpha_W=40$ cm^{-1}.

VII. CONCLUSION

We have developed a simplified density-matrix model for a four-state resonant phonon scattering based THz QCL. It describes the populations of different subbands and the coherence terms between them, as functions of electric field. Applied to a three-well design with vertical transitions at 3.6 THz, the model concludes that the wrong injection channel is not a major issue, while the wrong extraction channel from the upper lasing state to the extractor state is undesirable, particularly for cw operation. If possible, this leakage path should be minimized, for instance, by using diagonal transitions.[4] Under certain approximations, analytical solutions can be found, the gain profile equation being probably the most useful. Equation (17) shows that the gain spectrum is significantly broadened by both tunnelings, injection and extraction. This problem could be minimized by removing one tunneling process, for instance, by using a resonant phonon scattering mechanism for injection[47] or a two-well design.[24,45] In the expression of the optical gain, two nonlinear terms associated with stimulated scattering processes in a tunneling driven system are also identified. Their contribution to the total gain is not negligible, especially as the temperature is raised.

When used to fit experimental data of an injection barrier thickness study, the model showed its limits because of the large number of fitted parameters and its numerous simplifications. One of them is the fact that the thermal effect is mainly concentrated on the upper lasing state lifetime. The other limitation is the assumption of a constant electron-heating temperature, regardless of voltage or barrier thicknesses. However, we think the model indicates the range of different pure dephasing time constants and gives the correct order of magnitude of current density. The attractiveness of this model lies in its simplicity and quick computational implementation. We think it can be a very useful tool for performing a first-order optimization of THz QCL designs.

ACKNOWLEDGMENTS

The authors would like to thank M. Buchanan for a careful reading of the manuscript and J. B. Khurgin for fruitful discussions.

APPENDIX: SOLUTIONS OF EQ. (3)

Equation (3), representing the time evolution of the density matrix without laser-induced coherence terms, can be rewritten as a 16×16 system of linear equations, $\mathcal{A}[\rho]=\mathcal{C}$,

$$\begin{pmatrix}
\jmath\Omega_{12} & 0 & 0 & 0 & -\jmath\Omega_{24} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -(\tau_{\mathrm{sti}}^{-1}+\tau_2^{-1}) & \tau_{\mathrm{sti}}^{-1}+\tau_3^{-1} & 0 \\
-\Omega_{12} & 0 & -\Omega_{13} & 0 & \Omega_{24} & 0 & \Omega_{34} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \jmath\Omega_{24} & 0 & \jmath\Omega_{34} & 0 & \jmath\Omega_{14} & 0 & 0 & 0 & \tau_1^{-1} & 0 & 0 & -\tau_4^{-1} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
\jmath/\tau_{\parallel 12} & \Delta_{12} & 0 & 0 & 0 & \Omega_{14} & 0 & 0 & 0 & -\Omega_{24} & 0 & \Omega_{13} & -2\Omega_{12} & 2\Omega_{12} & 0 & 0 \\
\Delta_{12} & \jmath/\tau_{\parallel 12} & 0 & 0 & -\Omega_{14} & 0 & 0 & 0 & -\Omega_{24} & 0 & -\Omega_{13} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \jmath/\tau_{\parallel 13} & \Delta_{13} & 0 & 0 & 0 & \Omega_{14} & 0 & -\Omega_{34} & 0 & \Omega_{12} & -2\Omega_{13} & 0 & 2\Omega_{13} & 0 \\
0 & 0 & \Delta_{13} & \jmath/\tau_{\parallel 13} & 0 & 0 & -\Omega_{14} & 0 & -\Omega_{34} & 0 & \Omega_{12} & 0 & 0 & 0 & 0 & 0 \\
0 & -\Omega_{14} & 0 & 0 & \jmath/\tau_{\parallel 24} & \Delta_{24} & 0 & 0 & 0 & \Omega_{12} & 0 & -\Omega_{34} & 0 & -2\Omega_{24} & 0 & 2\Omega_{24} \\
\Omega_{14} & 0 & 0 & 0 & \Delta_{24} & \jmath/\tau_{\parallel 24} & 0 & 0 & \Omega_{12} & 0 & -\Omega_{34} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\Omega_{14} & 0 & 0 & \jmath/\tau_{\parallel 34} & \Delta_{34} & 0 & \Omega_{13} & 0 & -\Omega_{24} & 0 & 0 & -2\Omega_{34} & 2\Omega_{34} \\
0 & 0 & \Omega_{14} & 0 & 0 & 0 & \Delta_{34} & \jmath/\tau_{\parallel 34} & \Omega_{13} & 0 & \Omega_{24} & 0 & 0 & 0 & 0 & 0 \\
0 & -\Omega_{24} & 0 & -\Omega_{34} & 0 & \Omega_{12} & 0 & \Omega_{13} & \jmath/\tau_{\parallel 14} & \Delta_{14} & 0 & 0 & -2\Omega_{14} & 0 & 0 & 2\Omega_{14} \\
-\Omega_{24} & 0 & -\Omega_{34} & 0 & \Omega_{12} & 0 & \Omega_{13} & 0 & \Delta_{14} & \jmath/\tau_{\parallel 14} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\Omega_{13} & 0 & \Omega_{12} & 0 & -\Omega_{34} & 0 & \Omega_{24} & 0 & 0 & \jmath/\tau_{\parallel 23} & \Delta_{23} & 0 & 0 & 0 & 0 \\
\Omega_{13} & 0 & \Omega_{12} & 0 & -\Omega_{34} & 0 & -\Omega_{24} & 0 & 0 & 0 & \Delta_{23} & \jmath/\tau_{\parallel 23} & 0 & 0 & 0 & 0
\end{pmatrix}$$

$$\times\begin{pmatrix}
\rho_{21}-\rho_{12} \\
\rho_{21}+\rho_{12} \\
\rho_{31}-\rho_{13} \\
\rho_{31}+\rho_{13} \\
\rho_{42}-\rho_{24} \\
\rho_{42}+\rho_{24} \\
\rho_{43}-\rho_{34} \\
\rho_{43}+\rho_{34} \\
\rho_{41}-\rho_{14} \\
\rho_{41}+\rho_{14} \\
\rho_{32}-\rho_{23} \\
\rho_{32}+\rho_{23} \\
\rho_{11} \\
\rho_{22} \\
\rho_{33} \\
\rho_{44}
\end{pmatrix}=\begin{pmatrix}
0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{pmatrix}. \tag{A1}$$

If the laser-induced coherence terms were included, this 16×16 matrix, $\mathcal{A}$, would be taken for zero stimulated emission rate. The vector $[\rho]$ containing the different elements of the density matrix would be related only to the static terms, $\rho_{ij}^{(0)}$. Moreover, five terms of right-hand side vector, $\mathcal{C}$, would be changed: $\mathcal{C}_1=\Omega_L\Im(\tilde{\rho}_{23})$; $\mathcal{C}_5=\Omega_L\Re(\tilde{\rho}_{13})$; $\mathcal{C}_6=-\jmath\,\Omega_L\Im(\tilde{\rho}_{13})$; $\mathcal{C}_{11}=-\Omega_L\Re(\tilde{\rho}_{24})$; and $\mathcal{C}_{12}=\jmath\ \Omega_L\Im(\tilde{\rho}_{24})$. This new vector $\mathcal{C}$ would be used to simulate the gain broadening and saturation under intense laser field.[22]

If, in the matrix $\mathcal{A}$, we neglect the cross terms, $\Omega_{nm}\rho_{ij}$ with $nm\neq ij$ or ji, which explain the indirect resonance between states 1 and 4 at 8.5 kV/cm, then analytical formulas are easily achieved. This approximation can be justified by the fact that the indirect resonance is narrow due to the limited range of electric field for which the four main tunneling times are comparable. The formulas are given for infinite T_{14} because of the very small direct coupling Ω_{14} between levels 1 and 4. The first two terms of Eq. (9) for the current density give

$$J\approx qN_{2\mathrm{D}}\frac{\hat{c}+\hat{d}\tau_2\tau_{\mathrm{sti}}^{-1}}{\hat{a}+\hat{b}\tau_2\tau_{\mathrm{sti}}^{-1}}, \tag{A2}$$

where

$$\begin{aligned}\hat{c} + \hat{d}\tau_2\tau_{\text{sti}}^{-1} = {} & T_{13}T_{24}[1 + \tau_2\tau_{\text{sti}}^{-1}] + T_{24}[T_{12} + \tau_2 + (\tau_2\tau_{\text{sti}}^{-1})T_{12}] \\ & + T_{13}[\tau_2 + (\tau_2\tau_{\text{sti}}^{-1})T_{34}] + (T_{34} + T_{12})\tau_2 + (\tau_2\tau_{\text{sti}}^{-1})T_{12}T_{34}\end{aligned} \tag{A3}$$

and

$$\begin{aligned}\hat{a} + \hat{b}\tau_2\tau_{\text{sti}}^{-1} = {} & T_{13}T_{24}[T_{34} + T_{12} + 2\tau_2 + 2\tau_4 + (\tau_2\tau_{\text{sti}}^{-1})(3T_{34} + 4\tau_4 + T_{12})] \\ & + T_{24}[2T_{12}T_{34} + 3T_{34}\tau_2 + 3T_{12}\tau_4 + 4\tau_2\tau_4 + (\tau_2\tau_{\text{sti}}^{-1})T_{12}(3T_{34} + 4\tau_4)] \\ & + T_{13}[T_{12}\tau_2 + T_{34}\tau_4 + 4\tau_2\tau_4 + (\tau_2\tau_{\text{sti}}^{-1})T_{34}(T_{12} + 4\tau_4)] \\ & + 2T_{12}T_{34}(\tau_2 + \tau_4) + 4(T_{34} + T_{12})\tau_2\tau_4 + 4(\tau_2\tau_{\text{sti}}^{-1})T_{12}T_{34}\tau_4.\end{aligned} \tag{A4}$$

The population inversion, $\rho_{22}-\rho_{33}$, is

$$\Delta\rho = \frac{\hat{e}}{\hat{a} + \hat{b}\tau_2\tau_{\text{sti}}^{-1}}, \tag{A5}$$

where

$$\hat{e} = T_{13}T_{24}[\tau_2 - T_{34} - \tau_4] - T_{24}T_{12}[T_{34} + \tau_4] - T_{13}T_{34}\tau_4 - T_{12}T_{34}(\tau_4 + \tau_2). \tag{A6}$$

To estimate the gain with Eq. (17), it is also useful to get the expressions of the injection $\rho_{12}^{(0)}$ and extraction $\rho_{34}^{(0)}$ coherence terms, in the absence of laser field $(\Theta=0)$,

$$\rho_{12}^{(0)} = \frac{(\jmath + \Delta_{12}\tau_{\|12})/\Omega_{12}}{2\tau_{\text{transit}}|_{\Theta=0}} \frac{T_{13}T_{24} + T_{24}(T_{34} + \tau_4) + T_{13}\tau_2 + T_{34}(\tau_2 + \tau_4)}{T_{13}T_{24} + T_{24}(T_{12} + \tau_2) + T_{13}\tau_2 + (T_{34} + T_{12})\tau_2}, \tag{A7}$$

$$\rho_{34}^{(0)} = \frac{(\jmath + \Delta_{34}\tau_{\|34})/\Omega_{34}}{2\tau_{\text{transit}}|_{\Theta=0}} \frac{T_{13}T_{24} + T_{24}(T_{12} + \tau_2) + T_{13}\tau_4 + T_{12}(\tau_2 + \tau_4)}{T_{13}T_{24} + T_{24}(T_{12} + \tau_2) + T_{13}\tau_2 + (T_{34} + T_{12})\tau_2}. \tag{A8}$$

Interestingly, in the last two equations, the modulus of the first numerator, $(\jmath+\Delta_{ij}\tau_{\|ij})/\Omega_{ij}$, can be interpreted as the lifetime of the wave packet between states i and j. From these expressions, several characteristics of the laser can be easily derived. The ratio, $\eta_J=\hat{e}/\hat{c}$, is the product between population inversion and transit time in the absence of laser field, and it can be viewed as the current efficiency for population inversion [see Eq. (12)]. If we assume a cavity mirror loss coefficient, α_M, an optical waveguide loss, α_W, and an optical-mode overlap with the active region, Γ, the threshold population inversion, the threshold current, the stimulated emission rate, and the total optical power P_{op} easily become

$$\Delta\rho_{th} = \frac{\alpha_M + \alpha_W}{N_{3\text{D}}\Gamma\sigma_{\text{op}}}, \tag{A9}$$

$$J_{th} = qN_{2\text{D}} \left.\frac{\Delta\rho}{\eta_J}\right|_{th} = q\frac{L_{\text{sp}}}{\sigma_{\text{op}}\Gamma} \left.\frac{\hat{c}}{\hat{e}}\right|_{th} (\alpha_M + \alpha_W), \tag{A10}$$

$$\tau_{\text{sti}}^{-1} = \frac{\hat{a}}{\hat{b}\tau_2}\left(\frac{\Delta\rho|_{\Theta=0}}{\Delta\rho_{th}} - 1\right), \tag{A11}$$

$$\begin{aligned}P_{\text{op}} & = \hbar\omega V_c\alpha_M\frac{\hat{a}}{\hat{b}\sigma_{\text{op}}\tau_2}\left(\frac{\Delta\rho|_{\Theta=0}}{\Delta\rho_{th}} - 1\right) \\ & = \hbar\omega V_c\alpha_M\frac{\hat{c}}{\hat{d}\sigma_{\text{op}}\tau_2}\left(\frac{\eta_J J}{qN_{2\text{D}}\Delta\rho_{th}} - 1\right),\end{aligned} \tag{A12}$$

where V_c is the volume of the cavity mode. The transit time before and after laser threshold reads

$$\left(\frac{1}{\tau_{\text{transit}}}\right)_{I<I_{th}} = \frac{\hat{c}}{\hat{a}}, \tag{A13}$$

$$\left(\frac{1}{\tau_{\text{transit}}}\right)_{I>I_{th}} = \left(\frac{\hat{c}}{\hat{a}} - \frac{\hat{d}}{\hat{b}}\right)\frac{\Delta\rho_{th}}{\Delta\rho|_{\Theta=0}} + \frac{\hat{d}}{\hat{b}}. \tag{A14}$$

In a two-well QCL with one tunneling process (injection), Kumar *et al.*[23] have shown that the slope discontinuity of the current-voltage characteristic at threshold is simply related to the population inversion and the ratio of two intersubband scattering times. With our four-level system, where tunneling leakage channels cannot be neglected, it is not possible to derive a simple analytical expression for the differential resistance at threshold. One can only write a general expression for the discontinuity of the differential conductance at threshold,

$$\frac{1}{\widetilde{R}_{I=I_{th}^{+}}}-\frac{1}{\widetilde{R}_{I=I_{th}^{-}}}=\frac{1}{L_{\mathrm{ar}}}\left(qN_{2\mathrm{D}}\Sigma\left.\frac{\hat{d}}{\hat{b}}\right|_{E_{th}}-I_{th}\right)\left.\frac{d(\Delta\rho/\Delta\rho_{th})}{dE}\right|_{th}, \tag{A15}$$

where $\widetilde{R}$ is the differential resistance of the device, Σ is the area of the laser, and L_{ar} is the thickness of the active region. This expression involves the derivative of population inversion with respect to the electric field and the difference between the threshold current and the current at the same threshold electric field, E_{th}, of a hypothetical zero threshold device (the first term inside the brackets).

The maximum operating temperature of the laser, $T_{\max}$, can be approximatively given by this condition,

$$\left.\frac{\hat{e}}{\hat{a}}\right|_{E_{\mathrm{dsg}},T_{\max}}=\Delta\rho_{th}, \tag{A16}$$

and the differential wall-plug efficiency is approximatively given by

$$\frac{dP_{\mathrm{op}}}{dP_{\mathrm{elec}}}\approx\frac{\hbar\omega}{\hbar\omega+E_{\mathrm{LO}}}\frac{\alpha_M}{\alpha_M+\alpha_W}\frac{1}{\tau_2}\times\left[\frac{\hat{e}}{\hat{d}}+\frac{\widetilde{R}I}{L_{\mathrm{ar}}}\frac{d(\hat{e}/\hat{d})}{dE}-qN_{2\mathrm{D}}\Sigma\Delta\rho_{th}\frac{\widetilde{R}}{L_{\mathrm{ar}}}\frac{d(\hat{c}/\hat{d})}{dE}\right]. \tag{A17}$$

The internal quantum efficiency of the laser is represented by the bracket on the left-hand side of Eq. (A17) divided by the upper lasing state lifetime. It can be roughly estimated by $\hat{e}/\hat{d}\tau_2$ which can be calculated by Eqs. (A3) and (A6) and by approaching T_{13} toward infinity.

---

*emmanuel.dupont@nrc-cnrc.gc.ca

[1] B. S. Williams, Nat. Photonics **1**, 517 (2007).

[2] C. Walther, M. Fischer, G. Scalari, R. Terazzi, N. Hoyler, and J. Faist, Appl. Phys. Lett. **91**, 131122 (2007).

[3] B. S. Williams, S. Kumar, and Q. Hu, Opt. Express **13**, 3331 (2005).

[4] S. Kumar, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **94**, 131105 (2009).

[5] A. G. Davies, E. H. Linfield, and M. Pepper, Philos. Trans. R. Soc. London, Ser. A **362**, 197 (2004) for a review of the principal applications of THz.

[6] D. Indjin, P. Harrison, R. W. Kelsall, and Z. Ikonić, Appl. Phys. Lett. **82**, 1347 (2003).

[7] H. Callebaut, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **83**, 207 (2003).

[8] H. Callebaut, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **84**, 645 (2004).

[9] C. Jirauschek, G. Scarpa, P. Lugli, M. S. Vitiello, and G. Scamarcio, J. Appl. Phys. **101**, 086109 (2007).

[10] C. Jirauschek and P. Lugli, Phys. Status Solidi C **5**, 221 (2008).

[11] H. Callebaut and Q. Hu, J. Appl. Phys. **98**, 104505 (2005).

[12] S.-C. Lee and A. Wacker, Phys. Rev. B **66**, 245314 (2002).

[13] S.-C. Lee, F. Banit, M. Woerner, and A. Wacker, Phys. Rev. B **73**, 245320 (2006).

[14] A. Wacker, Phys. Status Solidi C **5**, 215 (2008).

[15] T. Kubis, C. Yeh, P. Vogl, A. Benz, G. Fasching, and C. Deutsch, Phys. Rev. B **79**, 195323 (2009).

[16] T. Schmielau and M. F. Pereira, Appl. Phys. Lett. **95**, 231111 (2009).

[17] R. C. Iotti and F. Rossi, Phys. Rev. Lett. **87**, 146603 (2001).

[18] I. Waldmueller, W. W. Chow, E. W. Young, and M. C. Wanke, IEEE J. Quantum Electron. **42**, 292 (2006).

[19] I. Savić, N. Vukmirović, Z. Ikonić, D. Indjin, R. W. Kelsall, P. Harrison, and V. Milanović, Phys. Rev. B **76**, 165310 (2007).

[20] R. Terazzi, T. Gresch, A. Wittmann, and J. Faist, Phys. Rev. B **78**, 155328 (2008).

[21] G. Scalari, R. Terazzi, M. Giovannini, N. Hoyler, and J. Faist, Appl. Phys. Lett. **91**, 032103 (2007).

[22] S. Kumar, Ph.D. thesis, Massachusetts Institute of Technology, 2007.

[23] S. Kumar and Q. Hu, Phys. Rev. B **80**, 245316 (2009).

[24] S. Kumar, C. W. I. Chan, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **95**, 141110 (2009).

[25] H. Luo, S. R. Laframboise, Z. R. Wasilewski, G. C. Aers, H. C. Liu, and J. C. Cao, Appl. Phys. Lett. **90**, 041112 (2007).

[26] H. Luo, S. R. Laframboise, Z. R. Wasilewski, and H. C. Liu, Electron. Lett. **43**, 633 (2007).

[27] H. Luo, S. R. Laframboise, Z. R. Wasilewski, H. C. Liu, and J. Cao, Electron. Lett. **44**, 630 (2008).

[28] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Les Éditions de Physique, Les Ulis, France, 1990).

[29] T. Unuma, M. Yoshita, T. Noda, H. Sakaki, and H. Akimaya, J. Appl. Phys. **93**, 1586 (2003).

[30] R. Nelander and A. Wacker, Appl. Phys. Lett. **92**, 081102 (2008).

[31] F. Banit, S.-C. Lee, A. Knorr, and A. Wacker, Appl. Phys. Lett. **86**, 041108 (2005).

[32] A. Wacker, R. Nelander, and C. Weber, in *Novel In-Plane Semiconductor Lasers VIII*, edited by A. A. Belyanin and P. M. Smowton (SPIE, Bellingham, WA, 2009), Vol. 7230, p. 72301A.

[33] R. Nelander and A. Wacker, J. Appl. Phys. **106**, 063115 (2009).

[34] P. Harrison, D. Indjin, and R. W. Kelsall, J. Appl. Phys. **92**, 6921 (2002).

[35] M. S. Vitiello, G. S. V. Spagnolo, B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **86**, 111115 (2005).

[36] M. S. Vitiello, G. S. V. Spagnolo, T. Losco, R. P. Green, A. Tredicucci, H. E. Beere, and D. A. Ritchie, Appl. Phys. Lett. **88**, 241109 (2006).

[37] P. Harrison, *Quantum wells, wires and dots*, 2nd ed. (Wiley, Chichester, England, 2005).

[38] T. Ando, J. Phys. Soc. Jpn. **54**, 2671 (1985).

[39] M. A. Belkin, J. A. Fan, S. Hormoz, F. Capasso, S. P. Khanna, M. Lachab, A. G. Davies, and E. H. Linfield, Opt. Express **16**, 3242 (2008).

[40] N. Jukam, S. S. Dhillon, D. Oustinov, J. Madéo, J. Tignon, R. Colombelli, P. Dean, M. Salih, S. P. Khanna, E. H. Linfield, and

A. G. Davies, Appl. Phys. Lett. **94**, 251108 (2009).

[41] S. Fathololoumi, E. Dupont, D. Ban, M. Graf, S. R. Laframboise, Z. R. Wasilewski, and H. C. Liu, IEEE J. Quantum Electron. **46**, 396 (2010).

[42] J. B. Khurgin, Y. Dikmelik, P. Q. Liu, A. J. Hoffman, M. D. Escarra, K. J. Franz, and C. F. Gmachl, Appl. Phys. Lett. **94**, 091101 (2009).

[43] O. Kocharovskaya, P. Mandel, and Y. V. Radeonychev, Phys. Rev. A **45**, 1997 (1992).

[44] M. Troccoli, A. Belyanin, F. Capasso, E. Cubukcu, D. L. Sivco, and A. Y. Cho, Nature (London) **433**, 845 (2005).

[45] G. Scalari, M. I. Amanti, C. Walther, R. Terazzi, M. Beck, and J. Faist, Opt. Express **18**, 8043 (2010).

[46] B. S. Williams, Ph.D. thesis, Massachusetts Institute of Technology, 2003.

[47] H. Yasuda, T. Kubis, P. Vogl, N. Sekine, I. Hosako, and K. Hirakawa, Appl. Phys. Lett. **94**, 151109 (2009).