

Multimode regimes in quantum cascade lasers: From coherent instabilities to spatial hole burning

Ariel Gordon,¹ Christine Y. Wang,² L. Diehl,³ F. X. Kärtner,^{1,*} A. Belyanin,⁴ D. Bour,⁵ S. Corzine,⁵ G. Höfler,⁵ H. C. Liu,⁶ H. Schneider,⁷ T. Maier,⁷ M. Troccoli,³ J. Faist,⁸ and Federico Capasso^{3,†}

¹*Department of Electrical Engineering and Computer Science and Research Laboratory of Electronics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA*

²*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*

³*School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA*

⁴*Department of Physics, Texas A & M University, College Station, Texas 77843, USA*

⁵*Agilent Technologies, Palo Alto, California 94306, USA*

⁶*Institute for Microstructural Science, National Research Council, Ottawa, Ontario, K1A R6 Canada*

⁷*Fraunhofer Institute for Applied Solid-State Physics, D-79108 Freiburg, Germany*

⁸*ETH Zürich, Institute for Quantum Electronics, Wolfgang-Pauli-Strasse 16 8093 Zürich, Switzerland*

(Received 10 December 2007; published 7 May 2008)

A theoretical and experimental study of multimode operation regimes in quantum cascade lasers (QCLs) is presented. It is shown that the fast gain recovery of QCLs promotes two multimode regimes: One is spatial hole burning (SHB) and the other one is related to the Risken-Nummedal-Graham-Haken instability predicted in the 1960s. A model that can account for coherent phenomena, a saturable absorber, and SHB is developed and studied in detail both analytically and numerically. A wide variety of experimental data on multimode regimes is presented. Lasers with a narrow active region and/or with metal coating on the sides tend to develop a splitting in the spectrum, approximately equal to twice the Rabi frequency. It is proposed that this behavior stems from the presence of a saturable absorber, which can result from a Kerr lensing effect in the cavity. Lasers with a wide active region, which have a weaker saturable absorber, do not exhibit a Rabi splitting and their multimode regime is governed by SHB. This experimental phenomenology is well-explained by our theoretical model. The temperature dependence of the multimode regime is also presented.

DOI: [10.1103/PhysRevA.77.053804](https://doi.org/10.1103/PhysRevA.77.053804)

PACS number(s): 42.55.Px, 42.65.Sf, 42.60.Mi

I. INTRODUCTION

Many physical mechanisms can drive a laser from a single-mode to a multimode regime. Common examples are spatial and spectral hole burning (SHB), saturable absorption, and self-phase modulation [1–3]. Understanding these mechanisms is of key importance to laser science and technology, whether one is interested in single-mode behavior or in a particular multimode operation such as mode locking.

The multimode regimes listed above are well-understood and documented both theoretically and experimentally in many types of lasers. However, the understanding of multimode regimes in quantum cascade lasers (QCLs) is still in its infancy, as these lasers were only demonstrated in 1994 [4], and studies of their multimode regimes commenced even more recently [5–8].

As it was shown recently [9] multimode dynamics in QCLs is different from that of more common lasers. This is mainly due to the unusually fast gain recovery of QCLs, which occurs on a picosecond scale. While a saturable absorber triggers mode locking in lasers with slow gain recovery (relative to the round-trip time), in lasers with fast gain recovery a saturable absorber triggers a mechanism [9] similar to the Risken-Nummedal-Graham-Haken (RNGH) instability [10,11]. While in standard semiconductor lasers carrier

diffusion eliminates spatial hole burning, in QCLs the gain recovery process is faster than carrier diffusion, and spatial hole burning is dominant, favoring multimode operation.

This paper presents a detailed study of multimode regimes in QCLs. The results of Ref. [10,11] are substantiated and extended. The first part of the paper is theoretical and the second one is experimental. In the theoretical section, the laser Maxwell-Bloch equations are introduced and analyzed for both a ring laser cavity and a Fabry-Perot cavity. This way one can study the interplay of coherent phenomena and spatial hole burning. A saturable absorber is added to the model as well. The model is studied analytically and numerically. The stability region of a continuous wave (cw) solution is found. It is shown that in a ring cavity, the presence of a saturable absorber lowers the threshold of the RNGH instability from about nine times above laser threshold [10,11] to arbitrarily low above laser threshold, depending on the strength of the absorber. However, the nature of the instability remains the same: the population inversion begins to oscillate at the Rabi frequency, modulating the gain in the laser. The result is sidebands around the original cw mode, separated from it by roughly the Rabi frequency. A Rabi splitting in the spectrum is the primary signature of the RNGH instability.

It is then shown that the ring-cavity Maxwell-Bloch model with a saturable absorber can explain the Rabi splitting and the lowering of the threshold, but cannot explain a key feature in the experimental spectra. The latter exhibit two relatively equal groups of modes with a gap in between,

*kaertner@mit.edu

†capasso@seas.harvard.edu

whereas the ring-cavity Maxwell-Bloch model predicts a large central mode in the spectrum with two sidebands. In view of this discrepancy, the Maxwell-Bloch equations were extended to include coupling between counterpropagating modes in a Fabry-Perot cavity, which is of course a more adequate model for QCLs. This model takes into account the development of SHB. The inclusion of SHB generates theoretical spectra with a Rabi splitting and without a central peak in the spectrum, in agreement with experiments.

After presenting the theoretical results, a thorough study of the experimental phenomenology is presented. It is shown that QCLs with narrower active regions tend to exhibit a more pronounced Rabi splitting than lasers with wider active regions. Lasers with a wider active region tend to exhibit multimode spectra that are governed by SHB. The explanation we propose to this behavior is that narrow QCLs have a stronger saturable absorption effect than wider ones because the optical intensity (and thus any nonlinear effect) is enhanced, and as a result of an enhanced Kerr-lensing effect [3,5]. In this case a nonlinear index enhancement in the waveguide core gives rise to an increased overlap of the transverse laser mode with an active region and a reduced overlap with lossy sidewalls, leading to an additional enhancement of the saturable absorption. The idea of saturable absorption by Kerr lensing is further supported by the fact that QCLs with metal coating on the sides of the ridge have a stronger RNGH behavior than lasers without metal coating: The metal coating enhances saturable losses that originated from Kerr lensing.

In the last part of the paper, the temperature dependence of the multimode behavior in QCLs is studied experimentally. It is found that at higher temperature the multimode behavior tends to be governed by the RNGH instability, whereas at lower temperatures it is governed by SHB. This behavior may be explained by the fact that at higher temperatures hot carriers populate more states in the injector superlattice, creating additional quaresonant absorption transitions between ground and excited minibands. This leads to an additional saturable absorption of laser radiation.

This paper is organized as follows: Section II gives a brief survey of prior work on multimode regimes in QCLs, Secs. III–V study theoretically the Maxwell-Bloch equations in a ring cavity and a Fabry-Perot cavity, Secs. VI–VIII summarize the experimental study, and Sec. IX is a brief summary.

II. PRIOR WORK ON MULTIMODE REGIMES IN QCLS

Multimode regimes in QCLs were observed in a series of recent works [5–8]. In Ref. [5], for example, it was observed that at a certain pumping current above lasing threshold, QCLs cease to operate in cw and develop a multimode regime. This multimode regime was characterized by a broadband optical spectrum and a narrow (less than 100 kHz) radio frequency (rf) beat note in the power spectrum at the cavity round-trip frequency.

The narrow beat note, whose width is $1/10^5$ of the central frequency, shows that the wave form of the electric field circulating in the laser cavity was stable over approximately

10^5 round-trips. In other words, the phase relationships between the longitudinal modes were stable for about 10^5 round-trips. The modes were therefore locked. However, in order to characterize the wave form which is circulating in the laser and to see if it is indeed an isolated pulse, as in traditional mode locking, one has to apply pulse characterization techniques, such as second order autocorrelation.

At the time when the experiment in Ref. [5] was performed, no second-order autocorrelation apparatus was available. However, second harmonic generation from QCLs [6] provided some information for pulse characterization. This measurement showed an increase by more than a factor of 5 in the second harmonic signal as the multimode behavior set in. This increase indicates that the duty cycle of the pulses was roughly 5. Since the number of modes in the spectrum is more than 5, we conclude that not all the modes were locked into a pulse. Without better pulse characterization data, one can infer that the laser could have had more than one pulse per round-trip.

It should also be noted that traditional mode locking, with a single pulse per round-trip, requires that the gain recovery time be longer than the cavity round-trip [3]. In QCLs this condition is violated, and therefore according to the existing theory one cannot expect mode locking with a single pulse per round-trip in QCLs. One can expect multiple pulses per round-trip.

In view of the above, one can see that the nature of various multimode regimes in QCLs requires further elucidation. This is the main goal of the present work.

III. MAXWELL-BLOCH EQUATIONS IN A FABRY-PEROT CAVITY

In this section we derive the Maxwell-Bloch equations in a Fabry-Perot cavity. We model the gain medium of QCLs as a two level system, described by the Bloch equations

$$\dot{\rho}_{ab} = i\omega\rho_{ab} + i\frac{\mu E}{\hbar}\Delta - \frac{\rho_{ab}}{T_2}, \quad (1)$$

$$\dot{\Delta} = -2i\frac{\mu E}{\hbar}(\rho_{ab}^* - \rho_{ab}) - \frac{\Delta_p - \Delta}{T_1} + D\frac{\partial^2 \Delta}{\partial z^2}, \quad (2)$$

where ρ_{ab} is the off-diagonal element of the density matrix, $\Delta \equiv \rho_{bb} - \rho_{aa}$ is the population inversion, ω and μ are the resonant frequency and the dipole matrix element of the lasing transition, T_1 and T_2 are the longitudinal and transverse relaxation times, and Δ_p is the steady-state inversion at $E = 0$, which characterizes the pumping rate. The last term in Eq. (2) is added phenomenologically and accounts for spatial diffusion of the inversion due to carrier diffusion. D is the diffusion coefficient. E is the electric field, which is assumed to satisfy the wave equation

$$\partial_z^2 E - \frac{n^2}{c^2} \partial_t^2 E = \frac{N\Gamma\mu}{\epsilon_0 c^2} \partial_t^2 (\rho_{ab} + \rho_{ab}^*). \quad (3)$$

N is the number of two-level systems per unit volume, which equals the average doping density in the active regions. Γ is the overlap factor between the optical mode and the active

region, and n is the background refractive index.

Equations (1) and (2) describe an open two-level system [12]: The total number of electrons in the system is not conserved, but rather can flow in and out, and vary depending on the bias conditions. In general the two levels, the upper one b and the lower one a , can have different relaxation times. In this situation Eq. (2) should be replaced by two equations, each with its own relaxation times. In QCLs indeed the upper and lower levels have different relaxation times. However, we neglect this difference for simplicity's sake and assign to them the same value T_1 . The generalization of the model (1) and (2) to a model with two different relaxation times is straightforward, as is also the generalization to a model with more than two levels.

We now make the following set of ansatzs:

$$E(z,t) = \frac{1}{2}[E_+^*(z,t)e^{-i(\omega t - kz)} + E_+(z,t)e^{i(\omega t - kz)}] + \frac{1}{2}[E_-^*(z,t)e^{-i(\omega t + kz)} + E_-(z,t)e^{i(\omega t + kz)}], \quad (4)$$

$$\rho_{ab}(z,t) = \eta_+(z,t)e^{i(\omega t - kz)} + \eta_-(z,t)e^{i(\omega t + kz)}, \quad (5)$$

$$\Delta(z,t) = \Delta_0(z,t) + \Delta_2(z,t)e^{2ikz} + \Delta_2^*(z,t)e^{-2ikz}, \quad (6)$$

where $k \equiv \omega n/c$. E_{\pm} , η_{\pm} , Δ_0 , and Δ_2 are assumed to vary slowly in time and space, on the scale defined by ω and k , respectively. The quantities with a + (−) subscript represent waves traveling to the positive (negative) z direction. Equation (6) allows taking SHB into account, Δ_2 being the envelope of the inversion grating. Note that Eq. (6) can be extended by adding terms proportional to e^{4ikz} , e^{6ikz} , etc. Neglecting these terms means that higher spatial frequencies on the inversion are neglected. Due to the nonlinearity of the gain medium such frequencies can appear if the gain is heavily saturated, but they are neglected in our analysis for simplicity's sake and since the pumping in our system is never much higher than the laser threshold.

Substituting Eqs. (4)–(6) into Eqs. (1)–(3) and making the slowly varying envelope approximation, we obtain the following set of equations:

$$\frac{n}{c}\partial_t E_{\pm} = \mp \partial_z E_{\pm} - i \frac{kN\mu\Gamma}{2\epsilon_0 n^2} \eta_{\pm} - \frac{1}{2}\ell(E_+, E_-)E_{\pm}, \quad (7)$$

$$\partial_t \eta_{\pm} = \frac{i\mu}{2\hbar}(\Delta_0 E_{\pm} + \Delta_2^{\pm} E_{\mp}) - \frac{\eta_{\pm}}{T_2}, \quad (8)$$

$$\partial_t \Delta_0 = \frac{\Delta_p - \Delta_0}{T_1} + \frac{i\mu}{\hbar}(E_+^* \eta_+ + E_-^* \eta_- - \text{c.c.}), \quad (9)$$

$$\partial_t \Delta_2^{\pm} = \pm i \frac{\mu}{\hbar}(E_+^* \eta_- - \eta_+^* E_-) - \frac{\Delta_2^{\pm}}{T_1} - 4k^2 D \Delta_2^{\pm}, \quad (10)$$

where we have introduced the notation $\Delta_2^{\pm} \equiv \Delta_2$, $\Delta_2^{\mp} \equiv \Delta_2^*$, in order that Eqs. (7)–(10) can be written more compactly. The last term in Eq. (7) has been added, and represents loss. The

loss ℓ is allowed to depend on the field to represent phenomena such as optical saturation.

In QCLs the laser cavity is formed by the two cleaved facets, one located at $z=0$ and the other one at $z=L$. At each facet the Fresnel reflection law dictates the following relations:

$$E_-(L,t) = \frac{n-1}{n+1}E_+(L,t), \quad (11)$$

$$E_+(0,t) = \frac{n-1}{n+1}E_-(0,t). \quad (12)$$

In what follows we study analytically and numerically the model introduced in this section. We begin with a simplified case, namely the standard Maxwell-Bloch equation in a ring cavity. In addition to briefly reviewing known results about the RNGH instability, we study the effect of a saturable absorber on the latter.

IV. RING CAVITY

In this section we consider a ring cavity, where SHB does not exist because standing waves cannot form. The aim is to understand the interplay of coherent effects and a saturable absorber alone, while avoiding complications due to SHB. We shall see that without SHB, the qualitative agreement between theory and the experiments on QCLs is not complete. After introducing SHB in the next section, the agreement is much more satisfactory.

A. RNGH instability with a saturable absorber

Dropping all the quantities with a “−” subscript from Eqs. (7)–(10), one arrives at the standard Maxwell-Bloch equations, with a saturable absorber added.

$$\frac{n}{c}\partial_t E = -\partial_z E - \frac{i\mu\eta}{\hbar\ell_0\Delta_{\text{th}}} - \frac{1}{2}(\ell_0 - \gamma|E|^2)E, \quad (13)$$

$$\partial_t \eta = \frac{i\mu}{2\hbar}\Delta E - \frac{\eta}{T_2}, \quad (14)$$

$$\partial_t \Delta = \frac{\Delta_p - \Delta}{T_1} + \frac{i\mu}{\hbar}(E^* \eta - \text{c.c.}). \quad (15)$$

The saturable absorber is approximated to the lowest order in E , and is characterized by γ , which is often referred to as the self-amplitude modulation coefficient [3]. ℓ_0 is the linear loss. Δ_{th} is the lasing threshold value of Δ_p for $\gamma=0$, given by (see Appendix A)

$$\Delta_{\text{th}}^{-1} = \frac{kN\mu^2\Gamma T_2}{2\hbar\ell_0\epsilon_0 n^2}. \quad (16)$$

Linear stability analysis of Eqs. (13)–(15) is carried out in detail in Appendix A. The gain of a perturbation at the frequency Ω (relative to the resonance frequency ω) is approximately given by

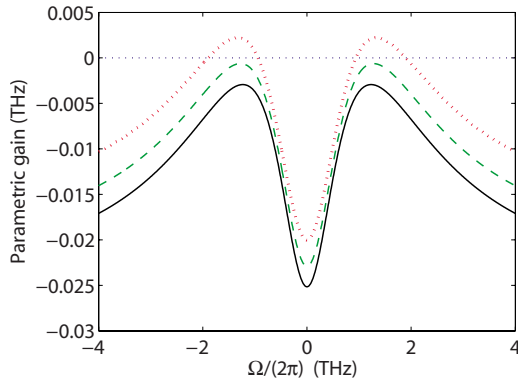


FIG. 1. (Color online) $g(\Omega)$ for $p=2$ and the parameters in Table I, apart from γ . The latter is $\gamma=0$ (solid), $\gamma=10^{-9}$ cm/V² (dashed), and $\gamma=2 \times 10^{-9}$ cm/V² (dotted).

$$g(\Omega) = -\frac{c}{2n} \text{Re} \left[\ell_0 \frac{(\Omega T_1 + i)\Omega T_2 - 2(p-1)}{(\Omega T_1 + i)(\Omega T_2 + i) - (p-1)} + \frac{\gamma \hbar^2 (p-1)}{\mu^2 T_1 T_2} \frac{(\Omega T_1 + i)(3\Omega T_2 + 2i) - 4(p-1)}{(\Omega T_1 + i)(\Omega T_2 + i) - p + 1} \right]. \quad (17)$$

The approximations made in the derivation of Eq. (17) are discussed in detail in Appendix A and mainly include assuming that the photon lifetime in the empty cavity is much longer than T_1 and T_2 . This approximation is excellent for QCLs. p is the pumping above lasing threshold (for $\gamma=0$):

$$p \equiv \frac{\Delta_p}{\Delta_{\text{th}}}. \quad (18)$$

Figure 1 shows $g(\Omega)$ for different values of γ at $p=2$. The parameters used in all examples in this paper, unless specified otherwise, are given in Table I. The effect of γ is to increase $g(\Omega)$ more or less uniformly across the frequency domain. In particular, it can bring $g(\Omega)$ above zero, thereby triggering an instability, even when the laser is not pumped as high above threshold. The reason why a saturable absorber lowers the RNGH threshold is that a saturable absorber itself always favors a multimode regime to a single mode one. It introduces a frequency-independent parametric gain. The latter is added to the RNGH parametric gain from Fig. 1, bringing it above threshold. Note, however, that the instability still

TABLE I. The parameters used in all calculations and simulations in this paper, unless indicated otherwise.

Quantity	Symbol	Value
Gain recovery time	T_1	0.5 ps
Dephasing time	T_2	0.067 ps
Linear cavity loss	ℓ_0	5 cm ⁻¹
Transition dipole element	μ	2.54 nm \times e
Background refractive index	n	3.3
Cavity length	L	3 mm
Saturable absorber coefficient	γ	10 ⁻⁸ cm/V ²

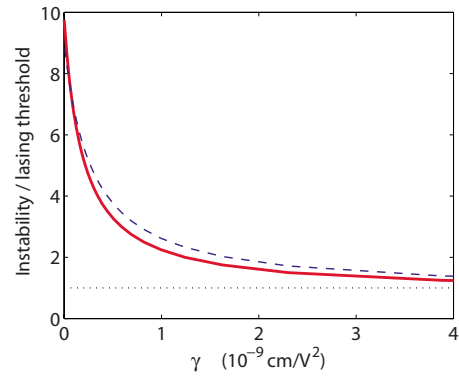


FIG. 2. (Color online) The pumping ratio p at which the RNGH instability sets in as a function of the saturable absorber coefficient γ . Solid line: exact result from linear stability analysis. Dashed line: approximate result [Eq. (21)]. The parameters, apart from γ , are given in Table I.

starts by developing Rabi sidebands around the cw lasing frequency. In this sense, it can be interpreted as a modified version of the original RNGH instability, rather than as a modulation instability caused by the saturable absorber alone.

The dependence of the threshold for instability on γ is shown in Fig. 2. For $\gamma=0$ we recover the standard RNGH instability, which occurs when pumping of slightly above 9 times the lasing threshold. For $\gamma \rightarrow \infty$, the instability threshold approaches the lasing threshold.

The results in Figs. 1 and 2 were obtained from numerical solutions of the algebraic equations involved in the stability analysis (see Appendix A). Since even the approximate expression of $g(\Omega)$ [Eq. (17)] is not very simple, it is useful to derive some approximate simple expressions for the properties of the instability. This is done in detail in Appendix A, and here we only give the results.

$g(\Omega)$ (e.g., in Fig. 1) has a local minimum at $\Omega=0$, and peaks at approximately

$$|\Omega_{\text{max}}| = \Omega_{\text{Rabi}} \sqrt{\frac{2p}{p-1}}. \quad (19)$$

The Rabi frequency $\Omega_{\text{Rabi}} = \mu E / \hbar$, where μ is the electron charge times the matrix element of the laser transition. The position of the peak depends weakly on γ within the parameter range of interest to our system. Note that for $\gamma=0$ one has

$$\Omega_{\text{Rabi}} = \sqrt{\frac{p-1}{T_1 T_2}}. \quad (20)$$

The dependence of $|\Omega_{\text{max}}|$ on p is shown in Fig. 3.

The instability threshold is approximately given by

$$p_{\text{th}} = 1 + 8 \left[4 \left(\frac{\hbar^2 \gamma}{\mu^2 T_2^2 \ell_0} \right)^2 + 12 \frac{\hbar^2 \gamma}{\mu^2 T_2^2 \ell_0} + 1 \right]^{-1}. \quad (21)$$

Figure 2 shows that Eq. (21) fairly well approximates the exact threshold condition.

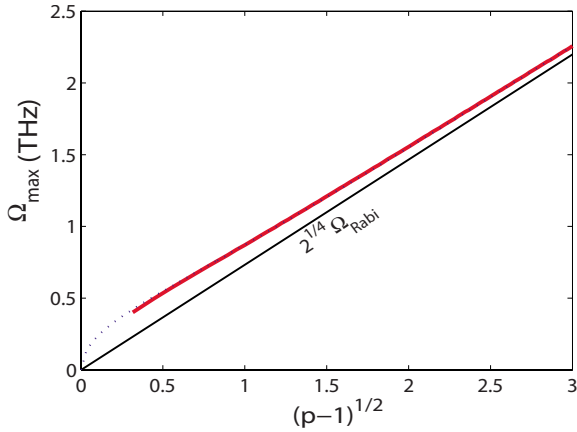


FIG. 3. (Color online) Ω_{\max} as a function of the pumping. Thick line: exact, and dotted line: Eq. (19). The parameters are from Table I with the only difference that $\gamma=0$.

B. Numerical results

The linear stability analysis given above can only find the threshold condition, and does not describe the behavior of the laser above the point of instability. Here we therefore present the results of a numerical solution of Eqs. (13)–(15).

According to Ref. [13], the energy in the Rabi sidebands can change either discontinuously or continuously at the RNGH instability threshold. In the language of phase transitions, this would be a first- or second-order phase transition, respectively. The criterion that determines the order of the transition involves the laser cavity length, but has not been given explicitly in the literature known to us.

According to the numerical simulations, for the parameters corresponding to our QCLs, including the saturable absorber, the transition is continuous (second order). Below the instability threshold the lasing is cw, and as the threshold is crossed, the Rabi sidebands around the central cw mode grow continuously. This behavior is demonstrated in Fig. 4. Note that standard mode locking, with a saturable absorber and slow gain recovery, is a discontinuous (first-order) transition [14,15].

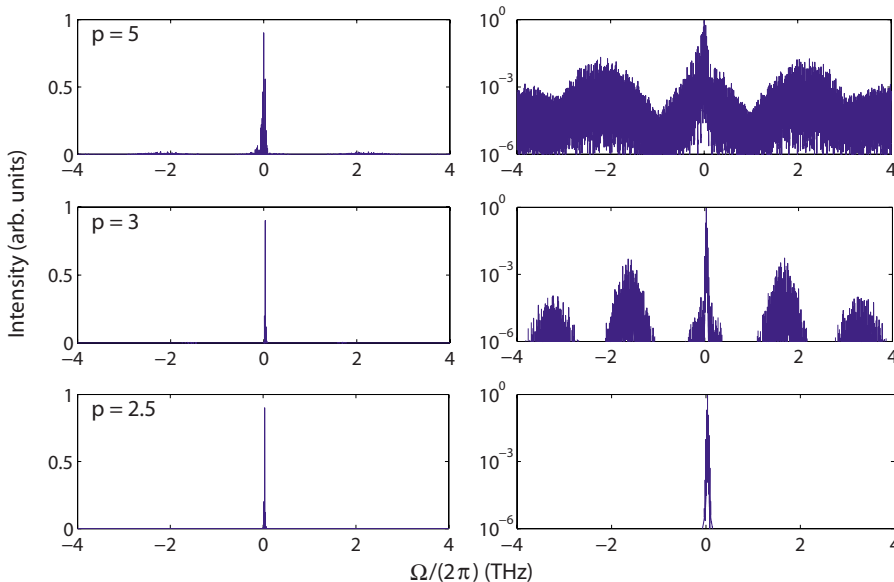


FIG. 4. (Color online) Spectra obtained from a numerical solution of the Maxwell-Bloch equations [Eqs. (13)–(15)]. The parameters are given in Table I, with the following exceptions: $\gamma=1 \times 10^{-9}$ cm/V², $L=6$ mm (to give the same round-trip time as in a Fabry-Perot 3 mm long cavity), and $\ell_0=7$ cm⁻¹ (to account for the mirror losses too).

The spectra in Fig. 4 have two groups of modes separated by roughly twice the Rabi frequency. In this respect they resemble the experimental spectra. However, they have a strong cw peak in between, a feature which is not shared by the experimental spectra. As we show in the next section, in a Fabry-Perot cavity with SHB, the central cw peak disappears.

V. FABRY-PEROT CAVITY

Spatial hole burning is associated with Δ_2 in Eq. (10). Intuitively, Δ_2 is the amplitude of the grating that couples the two propagation directions in the laser. The parameter that controls the strength of SHB is D : in the limit of $D \rightarrow \infty$, Δ_2 approaches zero. In order to better understand the interplay between SHB and the RNGH instability, we present now the results of analytical and numerical studies of Eqs. (7)–(10). We start with linear stability analysis. The calculation is shown in detail in Appendix A, and here we only give the results.

Before proceeding, from Eq. (10) we define the lifetime of the gain grating T_g as

$$T_g^{-1} = T_1^{-1} + 4k^2D. \quad (22)$$

T_g is the parameter that determines the strength of spatial hole burning. T_g can therefore range from zero (no SHB) to T_1 (strongest of SHB). The diffusion coefficient D can be estimated from the Einstein relation. With an electron mobility of 7000 (cm²/s)/V at room temperature, one has $D = 180$ cm²/s, $k=3.7 \times 10^4$ cm⁻¹, which roughly corresponds to a vacuum wavelength of 5 μ m, and we obtain $4k^2D \approx 1$ THz. With $T_1^{-1} \approx 2$ THz we find that $T_g \approx 0.3$ ps. Note that the mobility used here was relatively high, and the wavelength was on the short side of the scale. Therefore in reality T_g is closer to T_1 . It therefore follows that due to the fast gain recovery of QCLs, carrier diffusion does not eliminate spatial hole burning. This is in contrast with diode lasers [16,17].

In single mode operation, the standing wave associated with the lasing mode imprints a grating in the medium which

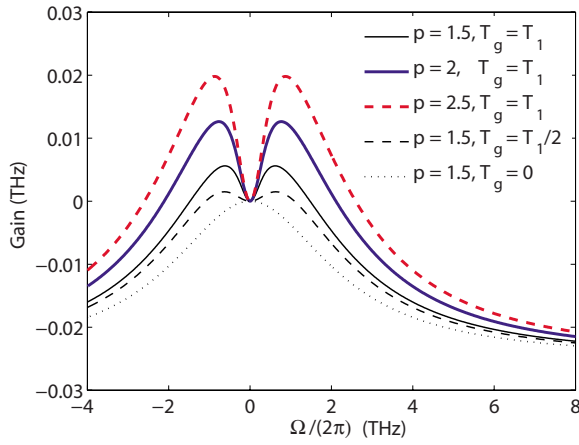


FIG. 5. (Color online) $g(\Omega)$ of the perturbation, associated with spatial hole burning. p and T_g corresponding to every curve are denoted in the legend.

has, to first approximation, a sinusoidal profile. This grating causes the lasing mode to experience heavier gain saturation than any other mode would experience. When the gain grating becomes strong enough, the single mode regime becomes unfavorable, and additional modes are excited.

A. Linear stability analysis

Linear stability analysis of Eqs. (7)–(10) gives two families of unstable modes. One is associated with the RNGH instability, and the other one with spatial hole burning. The first family is very similar to the case of a ring cavity studied earlier, with small differences that are discussed in Appendix A. The second one is derived in Appendix A, and the result is shown in Fig. 5.

Figure 5 shows the gain of a perturbation around a cw solution. The shorter is T_g and the smaller is the pumping, the smaller is the gain of the instability. For parameters typical to our QCLs, the SHB instability occurs a few percents above lasing threshold.

The gain curve in Fig. 5 exhibits two peaks and a dip at $\Omega=0$. For $T_2 \ll T_1$ and $p-1 \ll 1$, the location of the peaks is given by (see Appendix A)

$$\Omega_{\max}^2 \approx \frac{1}{T_1} \sqrt{\frac{p-1}{3T_1T_2}}. \quad (23)$$

Note that Eq. (23) gives a smaller frequency than Eq. (19). In addition, the splitting in Fig. 5 scales like the square root of the Rabi frequency. The cw solution is destabilized when the cavity admits a mode for which $g(\Omega)$ in Fig. 5 is positive.

B. Numerical results

The moment the cw solution is destabilized, studying Eqs. (7)–(10) requires a numerical simulation. The results of such a simulation are shown in Fig. 6. The parameters are given in Table I, with the only difference that $\gamma=0$ was used.

Figure 6 shows a clear pattern in the spectrum. This pattern appears only after very long averaging (order of a microsecond) of the spectrum. Such averaging is appropriate,

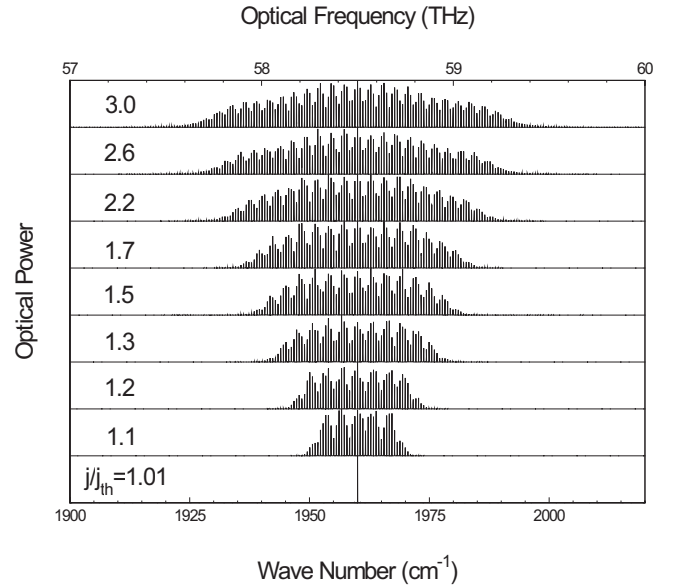


FIG. 6. Results of numerical simulations of the spectra based on the Maxwell-Bloch equations including a saturable absorber and spatial hole burning for different values of the current density normalized to the threshold value.

since similar averaging occurs when the spectra are measured. The pattern depends on T_2 and ℓ_0 , but very weakly depends on T_1 . A similar pattern occurs in the experimental spectra. However, we were not able to trace its origin.

The combined effect of SHB and a saturable absorber is demonstrated in Fig. 7. One can observe two effects. First is spectral broadening due to SHB, similarly to Fig. 6. Second, however, is the appearance of a splitting in the spectrum. This splitting is roughly equal to twice the Rabi frequency. Note, however, that in contrast to a ring laser, in a Fabry-

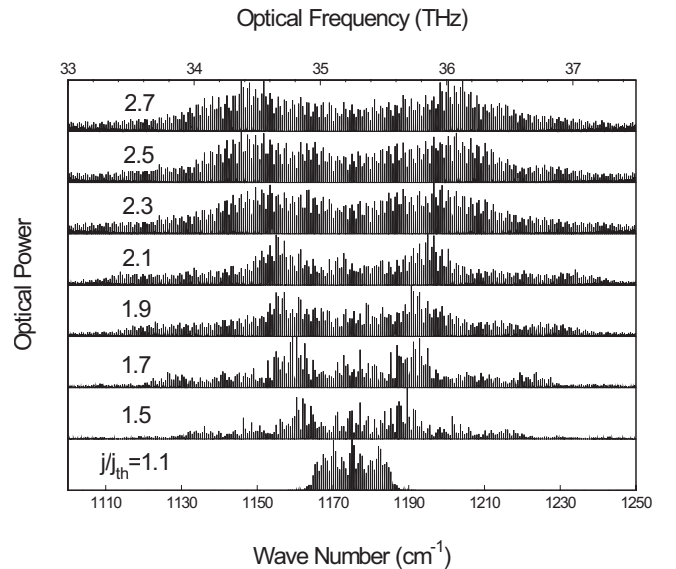


FIG. 7. Results of numerical simulations of the spectra based on the Maxwell-Bloch equations including a saturable absorber and spatial hole burning for different values of the current density normalized to the threshold value.

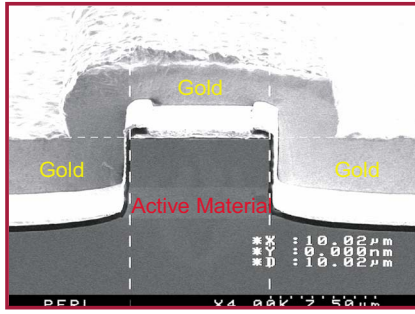


FIG. 8. (Color online) Scanning electron microscope (SEM) image of the cross section of a 10 μm wide ridge QC laser.

Perot laser “the Rabi frequency” is not a perfectly well-defined quantity. Since a standing wave is formed in the cavity, the electric field and thus the Rabi frequency depend on the position in the cavity. This dependence is even stronger when the mirrors have a relatively low reflection coefficient, since the field amplitude even more strongly depends on position.

In the previous section we saw that a saturable absorber lowers the threshold of the RNGH instability. Here we see that SHB suppresses the central peak seen in Fig. 4 and replaces it by a minimum in the spectrum. In the next sections we see that Fig. 7 agrees well with experimentally measured spectra, at least for devices where the RNGH behavior was dominant.

VI. EXPERIMENTAL RESULTS: RIDGE LASERS

We first study the multimode regimes in standard ridge QCLs, in which the sidewalls of the laser ridges are covered by a thick layer of electrically plated gold contact, see Fig. 8 for a typical laser cross section. The active region of the samples tested is based on a three-quantum-well design emitting at a wavelength $\lambda \sim 8 \mu\text{m}$ [18]. The wafer was grown by metalorganic vapor phase epitaxy (MOVPE). Figure 9(a) shows the voltage-current (V - I) and light-current (L - I) characteristics of a 10 μm wide laser operated in cw at 200 K, and Fig. 9(b) shows the corresponding optical spectra. The laser was cleaved into a 2 mm long bar and soldered with indium onto a copper heat sink. The optical power was measured by an thermal head power meter with a collection efficiency of nearly 100%. The spectra were measured by a Nicolet Fourier transform infrared spectrometer (FTIR) equipped with a deuterated triglycine sulfate (DTGS) detector.

As shown in Fig. 9(b), the laser spectrum is single mode close to laser threshold. It broadens and splits into two separated humps as the pumping current increases. The separation between the two peaks of the two humps increases linearly with the square root of the collected output power from one facet, as shown in Fig. 9(c). The Rabi angular frequency can be calculated from the collected output power, using the formula $\Omega_{\text{Rabi}} \equiv \mu E / \hbar = \mu \sqrt{2nI_{\text{ave}} / (c\epsilon)} / \hbar$, where μ is the electron charge times the matrix element of the laser transition ($\approx 1.9 \text{ nm}$ for this particular device). I_{ave} is the average intracavity intensity in the gain region, which can be derived

from the measured output power [19]. For all values of intensity corresponding to the spectra reported in Fig. 9(b), $\Omega_{\text{Rabi}}/2\pi$ was calculated, multiplied by a factor of 2 and then added to Fig. 9(c) (solid line). Reasonably good agreement is found between the experimental splitting and twice the estimated Rabi frequency. The error bars of the spectrum splittings come from the uncertainty in determining the exact position of the peaks, which is the full width at half maximum (FWHM) of the humps. As mentioned previously in the theoretical section, the RNGH instability predicts that large intracavity intensity will result in parametric gain at frequencies detuned from the maximum of the gain curve by the Rabi frequency. The measured spectra thus show clear indication of the RNGH instability in ridge QCLs.

The lowering of the RNGH instability threshold in our QCLs is due to the presence of a saturable absorber. This phenomenon is demonstrated analytically in the theoretical section. Such a saturable absorption mechanism in our experiments is likely to come from Kerr lensing, caused by a nonlinear (i.e., intensity dependent) refractive index n_2I in the active region [5]. As the light intensity increases, the mode becomes more confined in the plane transverse to the propagation direction, and the corresponding net modal gain also increases. The reason is twofold: First, the mode overlaps more with the active region, leading to a larger modal gain (this mechanism is often called “soft Kerr-lensing”). Second, the overlap with the metal contacts is reduced, leading to smaller losses. Thus the metal acts like a saturable absorber. A detailed analysis is given in Appendix C.

The same RNGH splitting in spectra is observed in many different devices, from wafers grown by both molecular beam epitaxy (MBE) and MOVPE. Figure 10 shows the spectra of a laser fabricated from a MBE-grown wafer with the same active region design as in Fig. 9, taken in continuous wave at 77 K. The spectrum starts from single mode close to the laser threshold, and the mode hops at a pumping ratio of $j/j_{\text{th}} = 1.66$, where there is a corresponding kink in the (L - I) curve at 330 mA. At 590 mA, there is another kink in the (L - I) curve, and a corresponding abrupt shift in the spectrum associated with mode hopping. Mode hopping and its relationship to kinks in the (L - I) curves are common phenomena in semiconductor lasers [20–22], but its cause in QC lasers has not yet been studied and is beyond the focus of this paper. After the spectrum broadens, it forms two separated humps whose peaks shift apart with increased optical power, similar to the MOVPE-grown lasers.

Characterization of the pulse emission was done using the method of second-harmonic interferometric autocorrelation. It is based on a Michelson interferometer in which the input beam is split into two and one of them is delayed by τ . In the standard setup, once recombined, the two pulses are sent collinearly first into a nonlinear crystal, and then a filter, which allows only the second-harmonic generation (SHG) component to be detected. One can test if it originates from a stationary isolated pulse train from the ratio between the interference maximum and the background (see Appendix B). The pulse duration can also be determined. However, due to the extremely low SHG conversion efficiency of nonlinear crystals in the mid-IR, the conventional setup is not feasible. To overcome this problem, we used a two-photon quantum

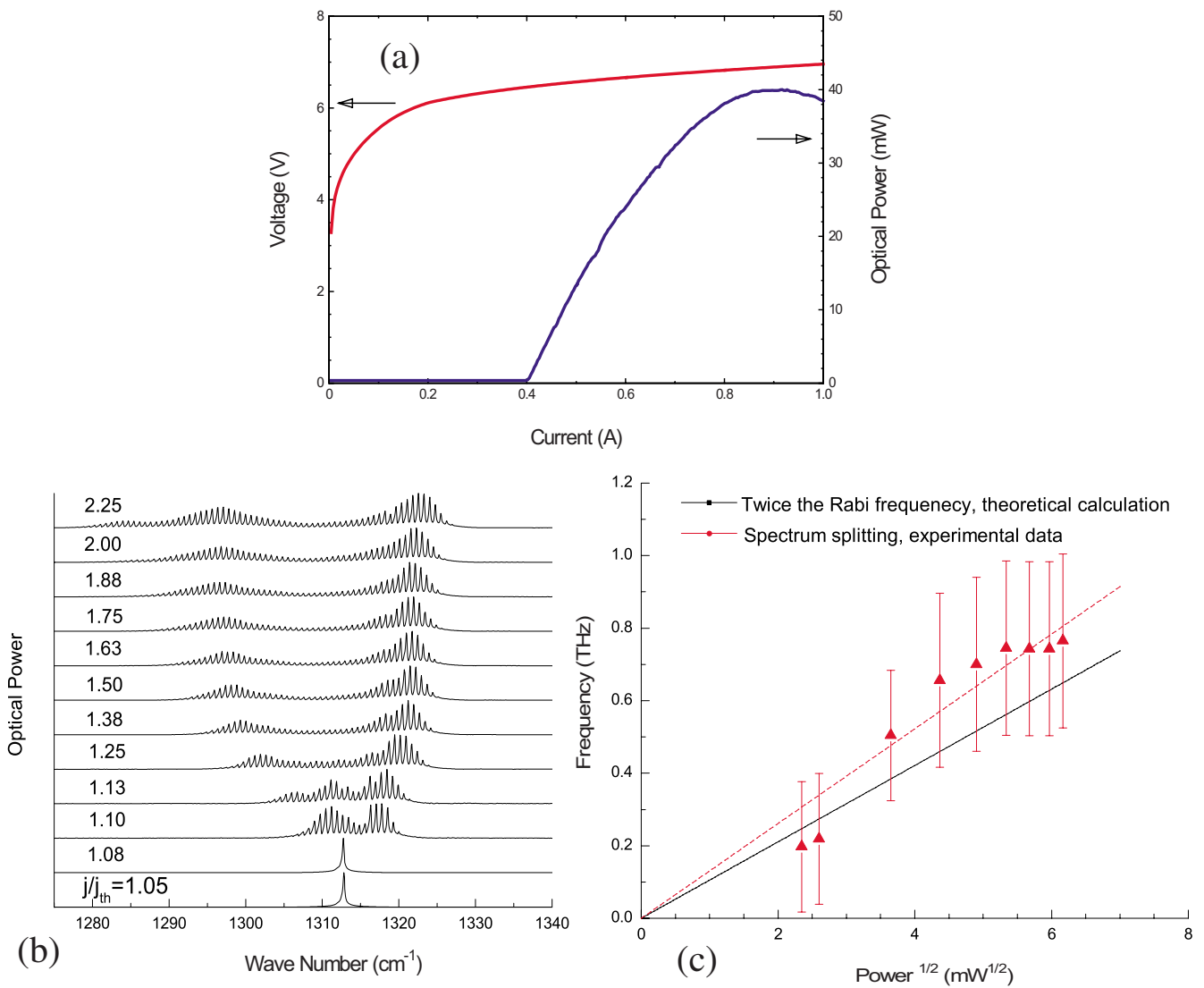


FIG. 9. (Color online) (a) V - I and L - I curves and (b) optical spectra vs pumping ratio (j/j_{th}) above threshold obtained in cw at 200 K of a 10 μm wide ridge laser emitting at 8 μm (wafer No. 2721). (c) Spectral splitting and twice the Rabi frequency $\Omega_{Rabi}/(2\pi)$ vs square root of output power collected from a single laser facet. The different quantities reported on the graph were deduced from the experimental data shown in (b). The dashed line is a least-squares linear fit of the data.

well infrared photodetector (QWIP) [23,24] instead of using a nonlinear crystal plus a linear detector. The energy diagram of one period of the multiquantum well detector under bias is shown in the inset of Fig. 11. The first three electronic states are nearly equidistant in energy. When electrons in the doped quantum wells absorb two photons simultaneously and the detector is biased (1–3 V), a photocurrent is generated and the signal can be detected by use of a preamplifier and a lock-in amplifier. This experimental setup is shown in Fig. 11. The second-order interferometric autocorrelation trace of the MBE-grown device mentioned above is shown in Fig. 12. Interference fringes are observed when the delay time from one arm of the autocorrelator is equal to multiples of the cavity round-trip time. The ratio between the maximum of the interference fringes and the background is smaller than 8, and the autocorrelation trace has some features between the cavity round-trip times, indicating that the multimode regime observed in this device is not stable mode locking

with a single pulse per round-trip. The meaning of the autocorrelation traces is discussed in Appendix B, and simulated traces are given therein.

In addition to second-order autocorrelation, the microwave spectrum of the laser output was also measured with an ultrafast QWIP [25] whose bandwidth is 52 GHz. The laser output is sent directly to the ultrafast QWIP, and the resulting photocurrent is displayed in a spectrum analyzer. Figure 12 shows the beat note signal of the MOVPE-grown sample at pump current 800 mA at 77 K. A steady peak with a FWHM of 13 MHz at 22.01 GHz, which corresponds to the cavity round-trip frequency of the 2 mm long laser (background refractive index $n=3.3$), is observed on the spectrum analyzer. It indicates a modulation of the laser output at the cavity round-trip frequency, and thus at least partial phase locking between the longitudinal modes: The phase relationships between the modes are stable for about 10^3 round-trips, as inferred from the peak frequency-to-FWHM ratio.

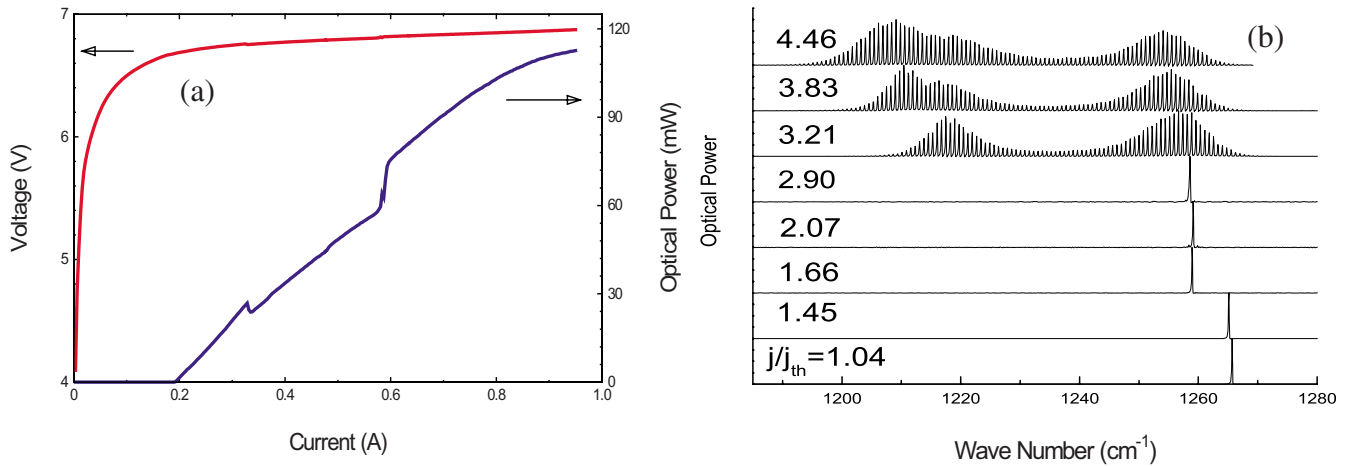


FIG. 10. (Color online) (a) V - I and L - I curves and (b) the optical spectra vs pumping ratio (j/j_{th}) above threshold obtained from a 10 μm wide ridge QCL operated in cw at 77 K emitting at 8 μm (wafer No. 2743). Note that mode hopping occurs while there is a jump in the L - I curve.

As the ridge width is increased, the RNGH instability and the Rabi splitting in the spectrum are suppressed. Figure 13 shows the spectra of a device processed from the same wafer as the one in Fig. 9, with the difference of its ridge width being increased to 15 μm . The spectra do not broaden much even at very high pumping currents, and do not show the Rabi splitting. As discussed in Appendix C, the effect of Kerr lensing decreases significantly when the active region width is increased. Thus this is strong evidence that the saturable absorption needed for lowering the RNGH threshold is provided by the Kerr-lensing mechanism.

VII. EXPERIMENTAL RESULTS: BURIED HETEROSTRUCTURE LASERS

The second type of lasers we have tested are the so-called buried heterostructure lasers, in which an insulating Fe-doped InP layer is regrown after etching of the ridges. A

thick layer of electrically plated gold is deposited on top of the ridges as the top contact layer after the InP regrowth. Figure 14 shows the cross section of such a device.

The active region of the buried heterostructure lasers tested is based on a four-quantum-well design, which relies on a double phonon resonance to achieve population inversion [26]. Figure 15(a) shows the V - I and L - I characteristics of a laser with an active region width of 3 μm and wavelength 8.38 μm operated in cw at room temperature, and Fig. 15(b) shows its optical spectra. The spectra show Rabi splitting similar to the ridge laser case, indicating the RNGH instability in this narrow buried heterostructure laser. The spectral splitting and twice the Rabi frequency are plotted against the square root of the collected output power in Fig. 15(b). Good agreement is found between the experimental splitting and twice the calculated Rabi frequency.

The second-order interferometric autocorrelation trace of the device when it is pumped five times above threshold at 80 K is shown in Fig. 16. The ratio between the maximum of the interference fringes and the background is close to 8 to 3, similar to the ridge laser case. In addition, there are smaller interference fringes within one cavity round-trip period. Both features indicate no stable pulsation from the laser (see Appendix B). The microwave spectrum of the laser output shows a steady peak with FWHM of 700 kHz at 15.018 GHz. This is shown in the inset of Fig. 16. This indicates that the phase relationship between the modes are maintained for about 10^4 cavity round-trips.

As in the ridge laser case, the lowering of the threshold of RNGH instability can also be attributed to the Kerr-lens type saturable absorption. In order to better support this idea experimentally, the spectra of another device processed from the same wafer as in Fig. 15 but with a wider active region (7.5 μm) are measured. Two-dimensional waveguide simulations indicate a much weaker Kerr-lensing effect in these QCLs, due to the much larger ratio of active region width to wavelength (see Appendix C). The measured optical spectra obtained at 300 K in cw mode are shown in Fig. 17(a). The envelopes of the spectra consist of multiple peaks whose separation is independent of the pumping current. The spec-

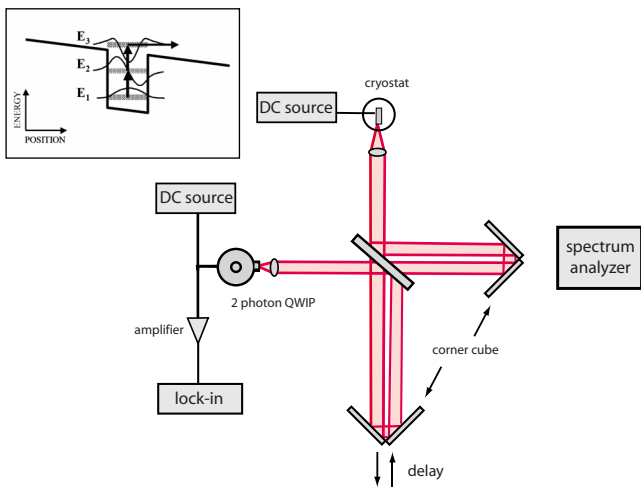


FIG. 11. (Color online) The experimental setup of a two-photon autocorrelation measurement. (Inset) Conduction band diagram of the two-photon QWIP showing three equidistant energy levels.

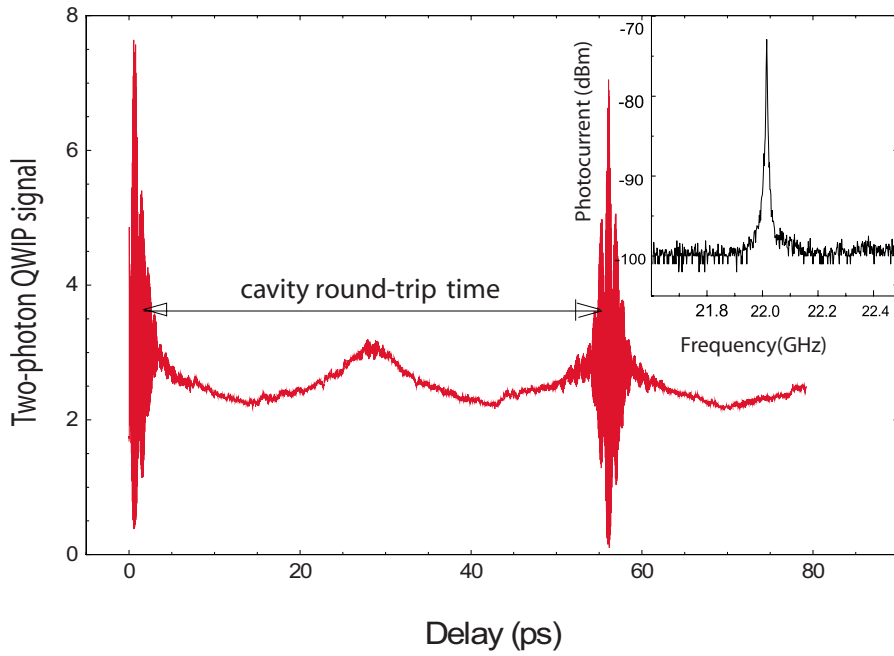


FIG. 12. (Color online) A second-order interferometric autocorrelation trace of a 8 μm wavelength ridge QC laser (wafer No. 2743) under the condition of RNGH instability. (Inset) Microwave spectrum of photocurrent generated by a similar laser (wafer No. 2721) under the condition of RNGH instability (measured with a 68 kHz resolution bandwidth).

tral signatures qualitatively agree with the numerical simulations in a Fabry-Perot cavity without a saturable absorber ($\gamma=0$) (Fig. 6).

Further increasing the active region width tends to suppress both the instabilities caused by RNGH and spatial hole burning. Figure 17(b) shows the spectra of a device also processed from the same wafer but with an even wider active region (10 μm) at 77 K. The spectra do not broaden much even at very high pumping currents, as in the ridge laser case. Since both RNGH instability and SHB stem from nonlinear effects, they are suppressed when the intensity of the field in the cavity is lower. Moreover, when the active region width increases, higher transverse modes are excited. Different transverse modes have different propagation constants β , and thus form different gain gratings which tend to wash out the effect of spatial hole burning.

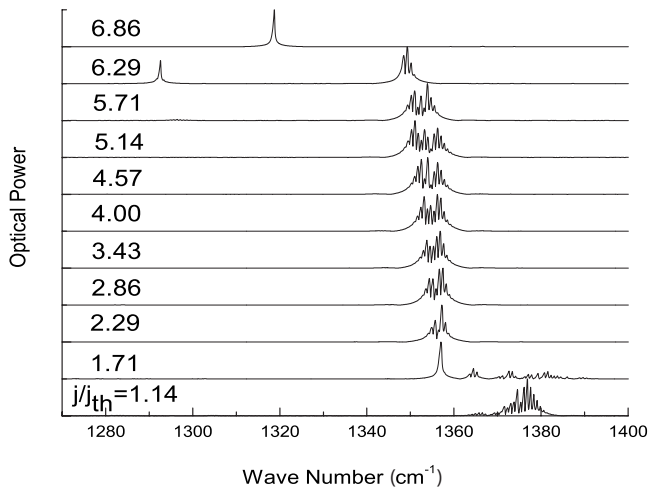


FIG. 13. Optical spectra vs pumping ratio (j/j_{th}) above threshold obtained in cw at 77 K with a 15 μm wide ridge laser emitting at 8 μm (wafer No. 2721).

VIII. TEMPERATURE EFFECTS

Temperature also plays an important role in the shape of the spectrum. To illustrate this point, we now present the spectra at different temperatures for the $\lambda=8.38 \mu\text{m}$ buried heterostructure laser with 3 μm active region width (the same as in Fig. 15).

From the spectra in Fig. 18, it is clear that at lower temperatures the spectra are dominated by spatial hole burning, showing multiple peaks independent of pumping and no significant Rabi splitting. As the temperature increases, the Rabi splitting becomes more evident and finally the RNGH instability takes over.

The effect of temperature on the nature of the multimode regime in QCLs can be explained in part by carrier diffusion, however, this effect alone seems to be too weak to provide a complete explanation. The lifetime of the gain grating T_g is defined as $T_g^{-1} = T_1^{-1} + 4k^2D$. The diffusion coefficient D of the gain grating is proportional to the temperature, $D = \mu k_B T / q$, where μ here is the carrier mobility, k_B is the Boltzmann's constant, T is the temperature, and q is the carrier charge. Therefore at higher temperature carrier diffusion would re-

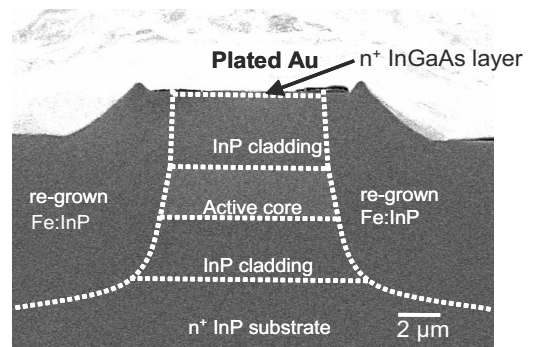


FIG. 14. SEM image of the cross section of a buried heterostructure QC laser.

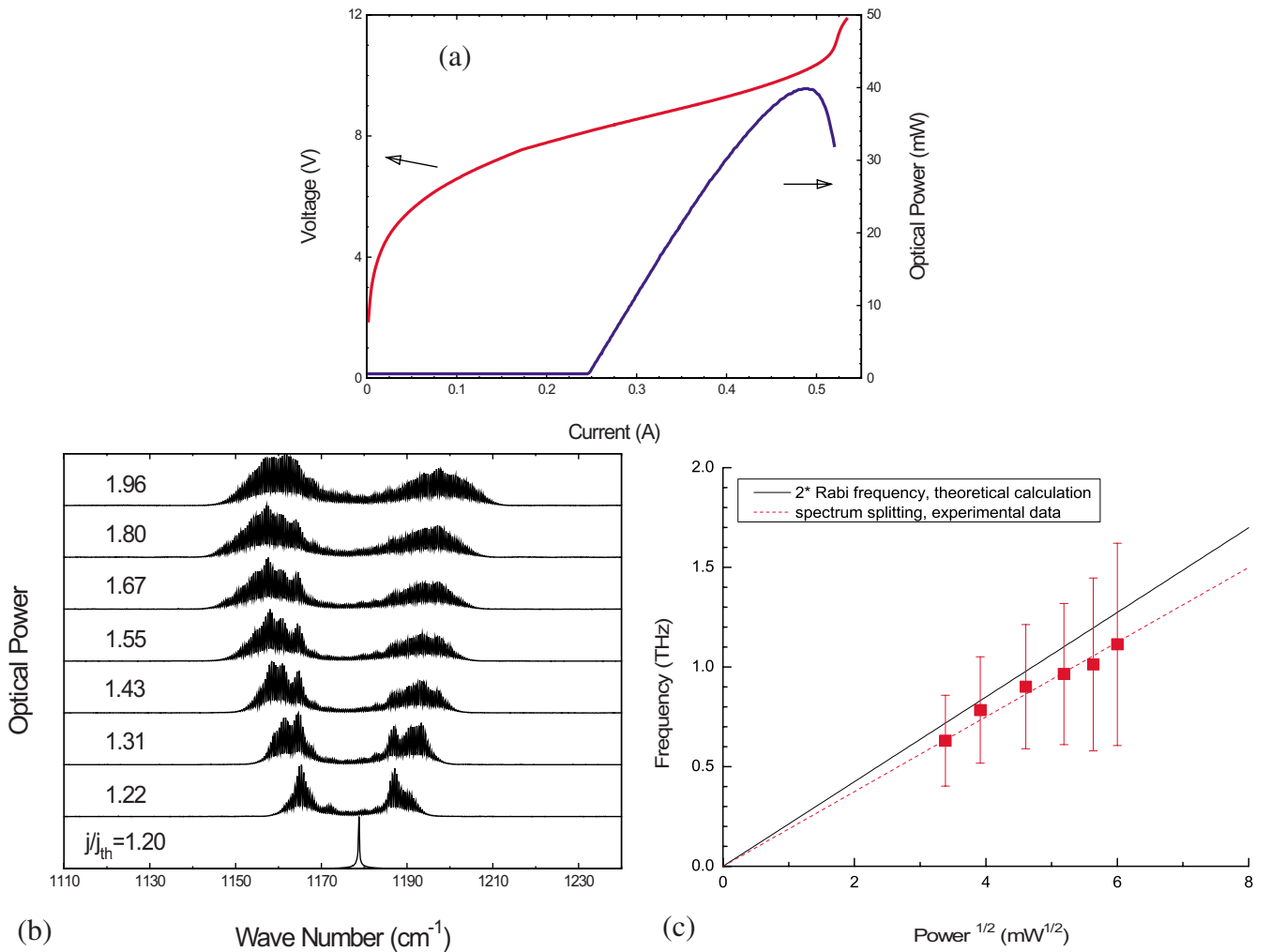


FIG. 15. (Color online) (a) V - I and L - I curves and (b) optical spectra vs pumping ratio (j/j_{th}) above threshold obtained in cw at 300 K with a 3 μm wide ridge laser emitting at 8.38 μm (wafer No. 3251). (c) Spectral splitting and twice the Rabi frequency $\Omega_{\text{Rabi}}/(2\pi)$ vs square root of output power collected from a single laser facet. The dashed line is a least-squares linear fit of the data.

duce spatial hole burning and thus reveal the RNGH instability. However, with an upper limit for the mobility of $7000 \text{ (cm}^2/\text{s)}/\text{V}$ and $k=2.25 \times 10^4 \text{ cm}^{-1}$ (which corresponds to a wavelength in vacuum of 8.38 μm), $4k^2D \approx 0.09 \text{ THz}$ at 77 K and $4k^2D \approx 0.4 \text{ THz}$ at 300 K, both significantly smaller than $T_1^{-1} \approx 0.6 \text{ THz}$ at 77 K and $T_1^{-1} \approx 2 \text{ THz}$ at 300 K. Thus carrier diffusion is unlikely to be the reason for the temperature dependence. Although the temperature effect is not entirely understood to us at this point, one possibility is temperature-dependent saturable absorption. The QCL injector consists of many energy levels which can be thermally populated. It is not surprising that if any two higher levels in the injector are nearly resonant with the laser transition, they will form a two-level saturable absorber. Therefore in this case saturable absorption is stronger at higher temperature and makes the RNGH instability more easily observable.

IX. CONCLUSION

This paper provides a thorough account of different multimode regimes in QCLs. It was found that two key mecha-

nisms which govern the multimode regimes in QCLs are a coherent instability similar to the RNGH instability and spatial hole burning. The former is enhanced due to the large dipole moment μ of the laser transition, which results in a large Rabi frequency compared to the relaxation rates. Thus the Rabi splitting can be resolved by the comb of modes supported by the cavity. SHB is enhanced because carrier diffusion is slower than the gain recovery, and thus leaves the gain grating intact. Note that in conventional semiconductor lasers the RNGH instability is not observed because typical Rabi frequencies are much smaller than the phase relaxation rate $1/T_2$. SHB in diode lasers is not so readily observable because diffusion occurs on a time scale comparable to the recombination time, i.e., the second term on the right-hand side of Eq. (22) is of the same order or greater than the first term $1/T_1 \approx 10^9 \text{ s}^{-1}$.

Due to the fast gain recovery, conventional mode locking, with one pulse per round-trip, is suppressed. In order to achieve conventional mode locking in QCLs, one needs to design a QCL with a slower gain recovery, such that T_1 becomes longer than or comparable to the cavity round-trip. Efforts in this direction are currently underway.

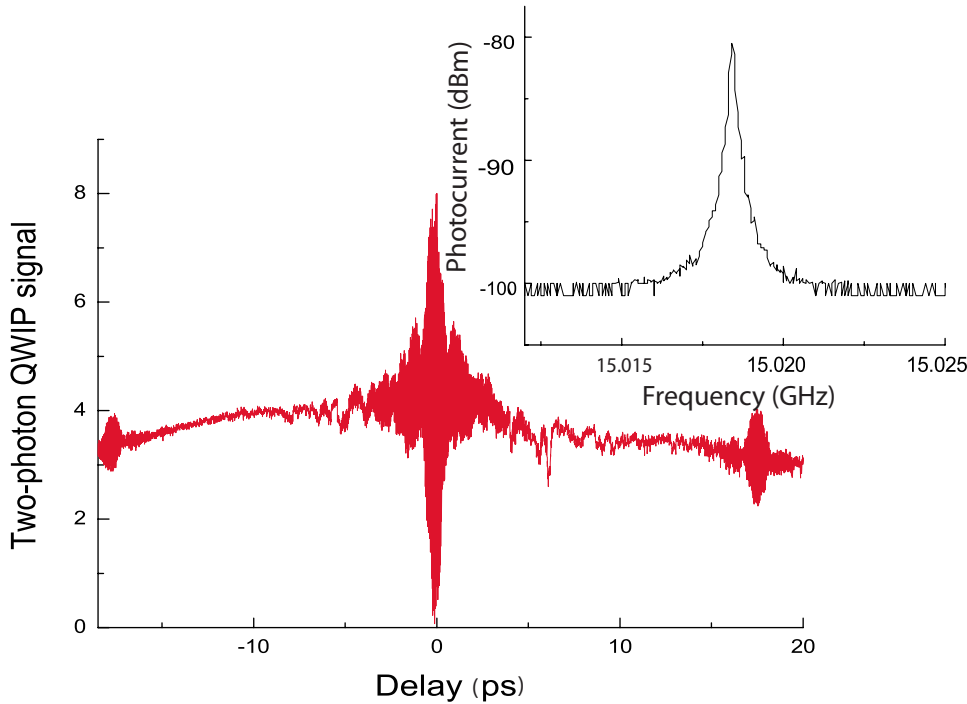


FIG. 16. (Color online) A second-order interferometric auto-correlation trace of a 8 μm wavelength buried heterostructure QC laser (wafer No. 3251) under the condition of RNGH instability. (Inset) Microwave spectrum of the photocurrent generated by the same laser (measured with a 68 kHz resolution bandwidth).

ACKNOWLEDGMENTS

Support from the U.S. Army Research Laboratory and the U.S. Army Research Office under Grant No. W911NF-04-1-0253 is gratefully acknowledged. Part of the device processing was done at the Center for Nanoscale Systems (CNS) at Harvard University. Harvard-CNS is a member of the National Nanotechnology Infrastructure Network (NNIN).

APPENDIX A: LINEAR STABILITY ANALYSIS

This appendix elaborates on the linear stability analysis of the Maxwell-Bloch equations in a ring cavity [Eqs. (13)–(15)] and in the Fabry-Perot cavity [Eqs. (7)–(10)].

1. Ring cavity

We begin with a linear stability analysis of Eqs. (13)–(15), with $\ell(E) = \ell_0 - \gamma|E|^2$. In order to keep the expressions from becoming too cluttered we define

$$\tilde{E} \equiv \frac{E\mu}{\hbar},$$

$$\tilde{\gamma} \equiv \hbar^2 \gamma / \mu^2,$$

$$\tilde{\eta} \equiv \frac{\ell_0 \eta}{\Delta_{\text{th}} T_2},$$

$$\tilde{\Delta} \equiv \frac{\ell_0 \Delta}{\Delta_{\text{th}} T_2}. \tag{A1}$$

Equations (13)–(15) are then rewritten as

$$\frac{n}{c} \partial_t \tilde{E} = -\partial_z \tilde{E} - i\tilde{\eta} - \frac{1}{2}(\ell_0 - \tilde{\gamma}|\tilde{E}|^2)\tilde{E}, \tag{A2}$$

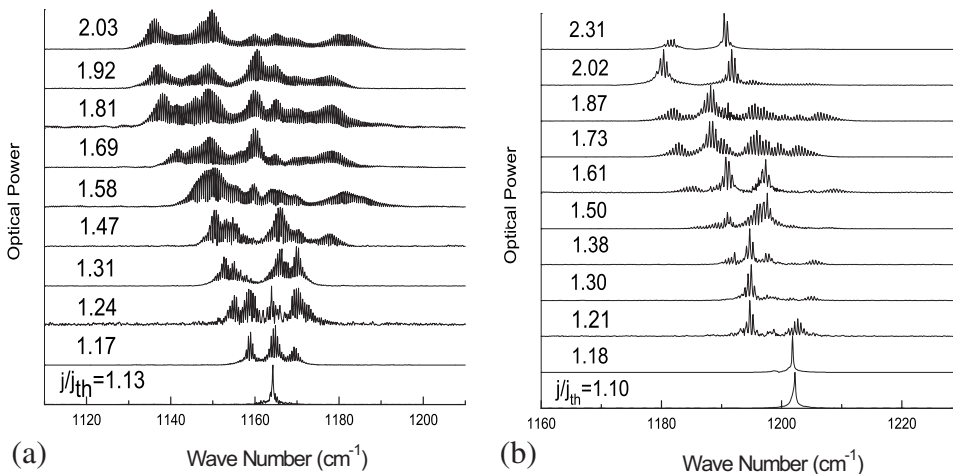


FIG. 17. Optical spectra vs pumping ratio (j/j_{th}) above threshold obtained in cw at 300 K with (a) a 7.5 μm wide and (b) a 10 μm buried heterostructure laser emitting at 8 μm (wafer No. 3251).

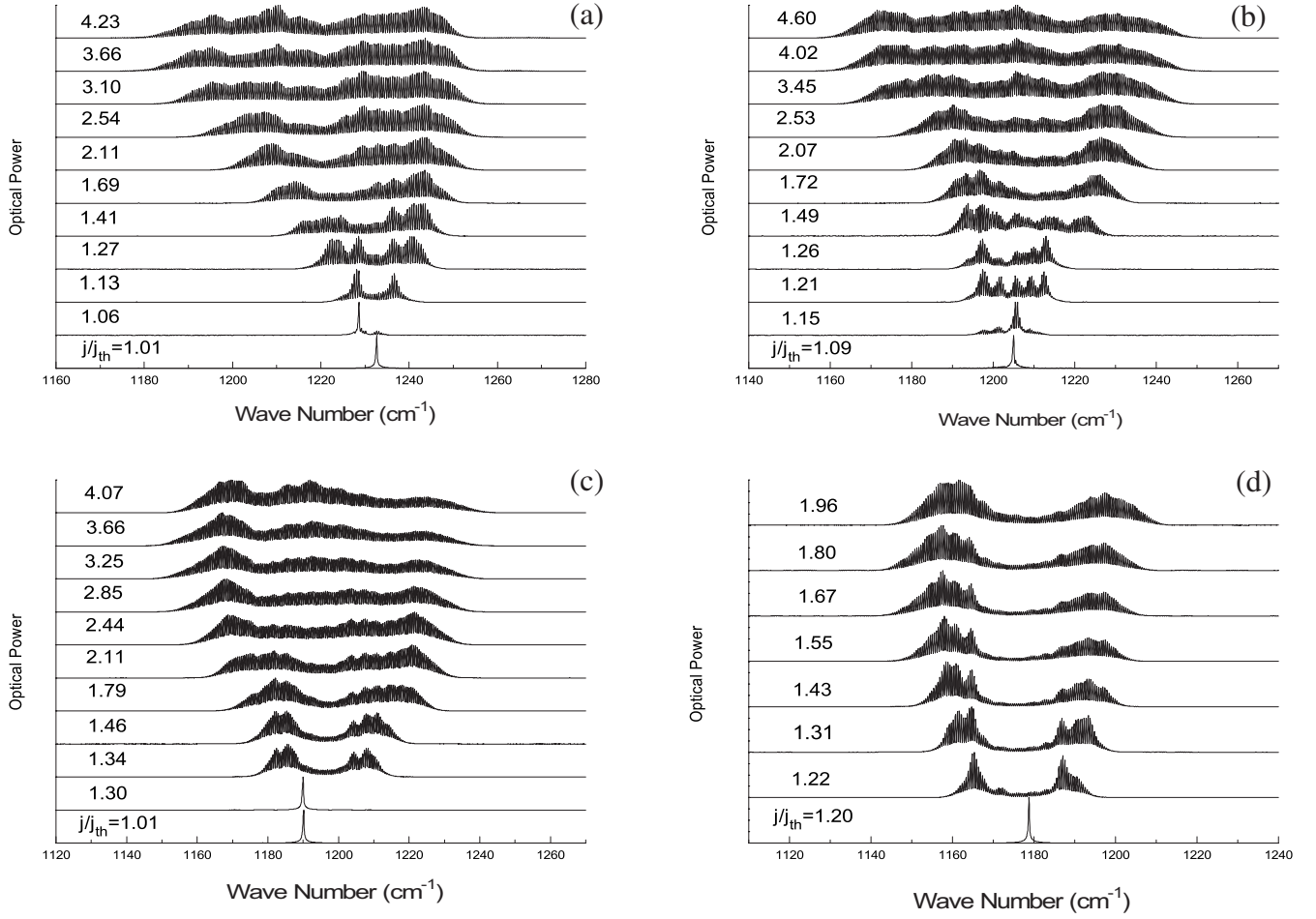


FIG. 18. Optical spectra vs pumping ratio (j/j_{th}) above threshold obtained in cw of a buried heterostructure laser emitting at $8.38 \mu\text{m}$ at different temperatures: (a) 80 K, (b) 150 K, (c) 200 K, and (d) 300 K.

$$\partial_t \tilde{\eta} = \frac{i\tilde{\Delta}}{2} \tilde{E} - \frac{\tilde{\eta}}{T_2}, \quad (\text{A3})$$

$$\partial_t \tilde{\Delta} = \frac{p\ell_0}{T_1 T_2} - \frac{\tilde{\Delta}}{T_1} + i(\tilde{E}^* \tilde{\eta} - \text{c.c.}). \quad (\text{A4})$$

Setting the left-hand sides of Eqs. (A2)–(A4) to zero we find that they admit a steady state solution of the form $\tilde{E} = \bar{E}$, $\tilde{\eta} = \bar{\eta}$, $\tilde{\Delta} = \bar{\Delta}$. \bar{E} , $\bar{\eta}$, $\bar{\Delta}$ are constants in time and space satisfying

$$\bar{\Delta} = \frac{\ell_0}{T_2} - \frac{\tilde{\gamma} \bar{E}^2}{T_2}, \quad (\text{A5})$$

$$\bar{\eta} = \frac{i}{2} (\ell_0 - \tilde{\gamma} \bar{E}^2) \bar{E}, \quad (\text{A6})$$

$$p + 1 = \left(1 - \frac{\tilde{\gamma} \bar{E}^2}{\ell_0}\right) (1 + \bar{E}^2 T_1 T_2). \quad (\text{A7})$$

\bar{E} was assumed real, since it can be always chosen so without loss of generality.

Adding perturbations $\delta \tilde{E}$, $\delta \tilde{\eta}$, $\delta \tilde{\Delta}$ to the steady state solution, and linearizing Eqs. (A2)–(A4) with respect to the perturbations, one obtains a set of linear equations:

$$\partial_t \delta \tilde{\eta}_I = \frac{1}{2} (\bar{\Delta} \delta \tilde{E}_R + \delta \tilde{\Delta} \bar{E}) - \frac{\delta \tilde{\eta}_I}{T_2}, \quad (\text{A8})$$

$$\partial_t \delta \tilde{\Delta} = -T_2 \bar{\Delta} \bar{E} \delta \tilde{E}_R - 2\bar{E} \delta \tilde{\eta}_I - \frac{\delta \tilde{\Delta}}{T_1}, \quad (\text{A9})$$

$$\frac{n}{c} \partial_t \delta \tilde{E}_R = -\partial_z \delta \tilde{E}_R + \delta \tilde{\eta}_I - (\ell_0 - 3\tilde{\gamma} \bar{E}^2) \frac{\delta \tilde{E}_R}{2} \quad (\text{A10})$$

and

$$\partial_t \delta \tilde{\eta}_R = -\frac{1}{2} \bar{\Delta} \delta \tilde{E}_I - \frac{\delta \tilde{\eta}_R}{T_2}, \quad (\text{A11})$$

$$\frac{n}{c} \partial_t \delta \tilde{E}_I = -\partial_z \delta \tilde{E}_I - \delta \tilde{\eta}_R - (\ell_0 - \tilde{\gamma} \bar{E}^2) \frac{\delta \tilde{E}_I}{2}. \quad (\text{A12})$$

The two sets of equations, (A8)–(A10) and (A11) and (A12), are decoupled, and can be thus studied independently. Since

both sets of equations are translationally invariant, plane waves are their eigenfunctions. For Eqs. (A8)–(A10) we therefore choose $\delta\tilde{\eta}_l(z, t) = \delta\tilde{\eta}_l(t)e^{ikz}$, and similarly for $\delta\tilde{\Delta}$ and $\delta\tilde{E}_R$. The stability of the cw solution is thus determined by the eigenvalues of the matrix

$$M = \begin{pmatrix} -T_2^{-1} & \frac{1}{2T_2}(\ell_0 - \tilde{\gamma}\bar{E}^2) & \frac{1}{2}\bar{E} \\ \frac{c}{n} & \frac{c}{n}\left(-\frac{1}{2}\ell_0 + \frac{3}{2}\tilde{\gamma}\bar{E}^2 - ik\right) & 0 \\ -2\bar{E} & \tilde{\gamma}\bar{E}^3 - \ell_0\bar{E} & -T_1^{-1} \end{pmatrix}. \quad (\text{A13})$$

If all eigenvalues have a negative real part, the cw solution is stable.

The eigenvalues of M can be easily found numerically. However, it is always more enlightening to have an analytical expression. To this aim we observe that of the three frequency parameters in M [T_1^{-1} , T_2^{-1} , and $\frac{c}{n}(\ell_0 + \tilde{\gamma}\bar{E}^2)$], the latter, which is the inverse cavity photon lifetime, is often the slowest one in QCLs. According to Table I, $\frac{c}{n}\ell_0 = 0.05$ THz, $T_1^{-1} = 2$ THz, and $T_2^{-1} = 15$ THz. We thus derive expressions for the eigenvalues which are correct to first order in ℓ_0 and $\tilde{\gamma}$.

For $\ell_0 = 0$ and $\tilde{\gamma} = 0$, the eigenvalue with the greatest real part is $\lambda_0(k) = -ick/n$. Putting $\lambda(k) = \lambda_0(k) + \lambda_1(k)$ into the characteristic polynomial of M and equating the parts which are first order in ℓ_0 , $\tilde{\gamma}$, and λ_1 , one arrives at

$$\lambda_{\max} = -i\Omega - \frac{\ell_0 c}{2n} \frac{(\Omega T_1 + i)\Omega T_2 - 2(p-1)}{(\Omega T_1 + i)(\Omega T_2 + i) - (p-1)} + \frac{(p-1)\tilde{\gamma}c}{2nT_1T_2} \frac{(\Omega T_1 + i)(3\Omega T_2 + 2i) - 4(p-1)}{(\Omega T_1 + i)(\Omega T_2 + i) - (p-1)}, \quad (\text{A14})$$

where $p = \Delta_p/\Delta_{\text{th}} - 1$ and $\Omega = kc/n$. Taking the real part of Eq. (A14) one obtains Eq. (17).

The position of the maximum of the gain curve is independent of $\tilde{\gamma}$, and to first order in T_2/T_1 is given by Eq. (19). The gain at that frequency is given by

$$g((\Omega_{\max})) = \frac{\ell_0 c T_2}{2n T_1} [3(p-1) - 2\sqrt{2p(p-1)}] + \frac{c\tilde{\gamma}(p-1)}{T_1 T_2}. \quad (\text{A15})$$

The threshold for the instability is found, to leading order in T_2/T_1 , by equating Eq. (A15) to zero. This yields Eq. (21).

2. Fabry-Perot cavity

Employing the same transformation as in Eq. (A1), Eqs. (7)–(10) take the form

$$\frac{n}{c}\partial_t\tilde{E}_{\pm} = \mp\partial_z\tilde{E}_{\pm} - i\tilde{\eta}_{\pm} - \frac{1}{2}\ell_0\tilde{E}_{\pm}, \quad (\text{A16})$$

$$\partial_t\tilde{\eta}_{\pm} = \frac{i}{2}(\tilde{\Delta}_0\tilde{E}_{\pm} + \tilde{\Delta}_2^{\pm}\tilde{E}_{\mp}) - \frac{\tilde{\eta}_{\pm}}{T_2}, \quad (\text{A17})$$

$$\partial_t\tilde{\Delta}_0 = \frac{\tilde{\Delta}_p - \tilde{\Delta}_0}{T_1} + i(\tilde{E}_+^*\tilde{\eta}_+ + \tilde{E}_-^*\tilde{\eta}_- - \text{c.c.}), \quad (\text{A18})$$

$$\partial_t\tilde{\Delta}_2^{\pm} = \pm i(\tilde{E}_+^*\tilde{\eta}_- - \tilde{\eta}_+^*\tilde{E}_-) - \frac{\tilde{\Delta}_2^{\pm}}{T_g}. \quad (\text{A19})$$

Splitting the variables with the tilde into their real and imaginary parts, Eqs. (A16)–(A19) can be rewritten as

$$\begin{aligned} \partial_t\delta\tilde{\eta}_+^I + \partial_t\delta\tilde{\eta}_-^I &= \frac{1}{2}(\tilde{\Delta}_0 + \tilde{\Delta}_2)(\delta\tilde{E}_-^R + \delta\tilde{E}_+^R) + \bar{E}(\delta\tilde{\Delta}_0 + \delta\tilde{\Delta}_R) \\ &\quad - \frac{\delta\tilde{\eta}_+^I + \delta\tilde{\eta}_-^I}{T_2}, \end{aligned} \quad (\text{A20})$$

$$\begin{aligned} \frac{n}{c}(\partial_t\delta\tilde{E}_+^R + \partial_t\delta\tilde{E}_-^R) &= -\partial_z(\delta\tilde{E}_+^R - \delta\tilde{E}_-^R) + (\delta\tilde{\eta}_+^I + \delta\tilde{\eta}_-^I) \\ &\quad - \frac{\ell_0}{2}(\delta\tilde{E}_+^R + \delta\tilde{E}_-^R), \end{aligned} \quad (\text{A21})$$

$$\partial_t\delta\tilde{\Delta}_0 = 2i\bar{\eta}(\delta\tilde{E}_+^R + \delta\tilde{E}_-^R) - 2\bar{E}(\delta\tilde{\eta}_+^I + \delta\tilde{\eta}_-^I) - \frac{\delta\tilde{\Delta}_0}{T_1}, \quad (\text{A22})$$

$$\partial_t\delta\tilde{\Delta}_2^R = i\bar{\eta}(\delta\tilde{E}_+^R + \delta\tilde{E}_-^R) - \bar{E}(\delta\tilde{\eta}_+^I + \delta\tilde{\eta}_-^I) - \frac{\delta\tilde{\Delta}_2^R}{T_g}, \quad (\text{A23})$$

$$\begin{aligned} \partial_t\delta\tilde{\eta}_+^R - \partial_t\delta\tilde{\eta}_-^R &= -\frac{1}{2}(\tilde{\Delta}_0 - \tilde{\Delta}_2)(\delta\tilde{E}_+^I - \delta\tilde{E}_-^I) + \delta\tilde{\Delta}_2^I\bar{E} \\ &\quad - \frac{\delta\tilde{\eta}_+^R - \delta\tilde{\eta}_-^R}{T_2}, \end{aligned} \quad (\text{A24})$$

$$\begin{aligned} \frac{n}{c}(\partial_t\delta\tilde{E}_+^I - \partial_t\delta\tilde{E}_-^I) &= -\partial_z(\delta\tilde{E}_+^I + \delta\tilde{E}_-^I) - (\delta\tilde{\eta}_+^R - \delta\tilde{\eta}_-^R) \\ &\quad - \frac{\ell_0}{2}(\delta\tilde{E}_+^I - \delta\tilde{E}_-^I), \end{aligned} \quad (\text{A25})$$

$$\partial_t\delta\tilde{\Delta}_2^I = -i\bar{\eta}(\delta\tilde{E}_+^I - \delta\tilde{E}_-^I) - \bar{E}(\delta\tilde{\eta}_+^R - \delta\tilde{\eta}_-^R) - \frac{\delta\tilde{\Delta}_2^I}{T_g}. \quad (\text{A26})$$

We now define the generalized parity operator P , which reflects the space around the center of the cavity and swaps all + and – quantities:

$$PX_{\pm}(z, t) = X_{\mp}(L - z, t),$$

where X is any of the quantities $\tilde{\eta}$, \tilde{E} , or $\tilde{\Delta}$, with an R or I superscript. Equations (A20)–(A26) commute with P , and therefore their eigenmodes can be chosen to be eigenstates of P as well. Since P^2 is the unity operator, P can have only “even” or “odd” eigenstates, with +1 and –1 eigenvalues, respectively.

We now observe that Eqs. (A20)–(A23) are decoupled from Eqs. (A24)–(A26). For Eqs. (A20)–(A23), odd solutions are trivial, whereas for Eqs. (A24)–(A26), even solutions are trivial. Assuming an even solution for Eqs. (A20)–(A23) and an odd solution to Eqs. (A24)–(A26), Eqs. (A20)–(A23) are reduced to

$$\partial_t \delta \tilde{\eta}_+^I = \frac{\bar{\Delta}_0 + \bar{\Delta}_2}{2} \delta \tilde{E}_+^R + \frac{\bar{E}}{2} (\delta \tilde{\Delta}_0 + \delta \tilde{\Delta}_R) - \frac{\delta \tilde{\eta}_+^I}{T_2}, \quad (\text{A27})$$

$$\frac{n}{c} \partial_t \delta \tilde{E}_+^R = -\partial_z \delta \tilde{E}_+^R + \delta \tilde{\eta}_+^I - \frac{\ell_0}{2} \delta \tilde{E}_+^R, \quad (\text{A28})$$

$$\partial_t \delta \tilde{\Delta}_0 = 4i\bar{\eta} \delta \tilde{E}_+^R - 4\bar{E} \delta \tilde{\eta}_+^I - \frac{\delta \tilde{\Delta}_0}{T_1}, \quad (\text{A29})$$

$$\partial_t \delta \tilde{\Delta}_2^R = 2i\bar{\eta} \delta \tilde{E}_+^R - 2\bar{E} \delta \tilde{\eta}_+^I - \frac{\delta \tilde{\Delta}_2^R}{T_g}, \quad (\text{A30})$$

$$\partial_t \delta \tilde{\eta}_+^R = -\frac{\bar{\Delta}_0 - \bar{\Delta}_2}{2} \delta \tilde{E}_+^I + \frac{1}{2} \delta \tilde{\Delta}_2^I \bar{E} - \frac{\delta \tilde{\eta}_+^R}{T_2}, \quad (\text{A31})$$

$$\frac{n}{c} \partial_t \delta \tilde{E}_+^I = -\partial_z \delta \tilde{E}_+^I - \delta \tilde{\eta}_+^R - \frac{\ell_0}{2} \delta \tilde{E}_+^I, \quad (\text{A32})$$

$$\partial_t \delta \tilde{\Delta}_2^I = -2i\bar{\eta} \delta \tilde{E}_+^I - 2\bar{E} \delta \tilde{\eta}_+^R - \frac{\delta \tilde{\Delta}_2^I}{T_g}. \quad (\text{A33})$$

Similarly to the discussion around Eq. (A13), the stability of Eqs. (A27)–(A33) is studied by finding the eigenvalues of the matrix

$$\begin{pmatrix} -T_2^{-1} & \frac{\ell_0}{2T_2} & \frac{\bar{E}}{2} & \frac{\bar{E}}{2} \\ \frac{c}{n} & \frac{c}{n} \left(-\frac{1}{2} \ell_0 - ik \right) & 0 & 0 \\ -4\bar{E} & -2\ell_0 \bar{E} & -T_1^{-1} & 0 \\ -2\bar{E} & -\ell_0 \bar{E} & 0 & -T_g^{-1} \end{pmatrix} \quad (\text{A34})$$

for Eqs. (A27)–(A30), and

$$\begin{pmatrix} -T_2^{-1} & \frac{\ell_0}{2T_2} & \frac{\bar{E}}{2} \\ -\frac{c}{n} & \frac{c}{n} \left(-\frac{1}{2} \ell_0 - ik \right) & 0 \\ -2\bar{E} & \ell_0 \bar{E} & -T_g^{-1} \end{pmatrix} \quad (\text{A35})$$

for Eqs. (A31)–(A33). \bar{E} is now related to p via

$$\bar{E}^2 = \frac{p-1}{2T_1 T_2 + T_2 T_g}.$$

The matrix (A34) is related to the RNGH instability, whereas matrix (A35) is related to spatial hole burning. In the limit of

$T_g \rightarrow 0$, as well as in the case $T_g = T_1$, $g(\Omega)$ is identical to the expression (17) without a saturable absorber ($\gamma=0$):

$$g(\Omega) = -\frac{\ell_0 c}{2n} \text{Re} \left[\frac{(\Omega T_1 + i)\Omega T_2 - 2(p-1)}{(\Omega T_1 + i)(\Omega T_2 + i) - (p-1)} \right]. \quad (\text{A36})$$

The RNGH instability threshold is therefore again around $p=9$. Note, however, that the cw solution destabilizes for much smaller p due to spatial hole burning. Indeed, $g(\Omega)$ obtained from Eq. (A35) is given by

$$g(\Omega) = -\frac{\ell_0 c}{2n} \text{Re} \left[1 + \frac{3i(i + \Omega T_1) - (p-1)}{3(i + \Omega T_2)(i + \Omega T_2) - (p-1)} \right] \quad (\text{A37})$$

for $T_g = T_1$. The peak of the gain curve is obtained at

$$\Omega_{\text{max}}^2 \approx \frac{1}{T_1} \sqrt{\frac{p-1}{3T_1 T_2}}$$

for $T_2 \ll T_1$ and $p-1 \ll 1$. For $T_g=0$ spatial hole burning does not exist, and $g(\Omega)$ obtained from Eq. (A35) is given by

$$g(\Omega) = -\frac{\ell_0 c}{2n} \text{Re} \left[\frac{2i + \Omega T_2}{i + \Omega T_2} \right], \quad (\text{A38})$$

which is never positive.

APPENDIX B: INTERFEROMETRIC AUTOCORRELATION

If we write the electric field as a function of time as $E(t)e^{i\omega t}$, the two-photon interferometric autocorrelation is given by

$$\begin{aligned} I(\tau) &= \int_{-\infty}^{\infty} |E(t+\tau)e^{i\omega\tau} + E(t)|^4 dt \\ &= \int_{-\infty}^{\infty} [|E(t+\tau)|^4 + |E(t)|^4 + 4|E(t+\tau)E(t)|^2] dt \\ &\quad + 2 \left[\int_{-\infty}^{\infty} dt |E(t)|^2 E^*(t) E(t+\tau) e^{i\omega\tau} + \text{c.c.} \right] \\ &\quad + 2 \left[\int_{-\infty}^{\infty} dt |E(t+\tau)|^2 E^*(t) E(t+\tau) e^{i\omega\tau} + \text{c.c.} \right] \\ &\quad + \left[\int_{-\infty}^{\infty} dt [E(t+\tau)E^*(t)e^{2i\omega\tau}]^2 + \text{c.c.} \right]. \end{aligned} \quad (\text{B1})$$

We define the background as

$$I_b = 2 \int_{-\infty}^{\infty} |E(t)|^4 dt. \quad (\text{B2})$$

From Eq. (B1) one can see that $I(0)=8I_b$. We now assume that $E(t)$ is an isolated pulse, and for simplicity's sake we assume that $E(t)$ is nonzero only over an interval T . Then for $\tau > T$ all terms in Eq. (B1) which include both t and $t+\tau$ vanish. Therefore $I(\tau > T) = I_b$. In other words, for a pulse one has

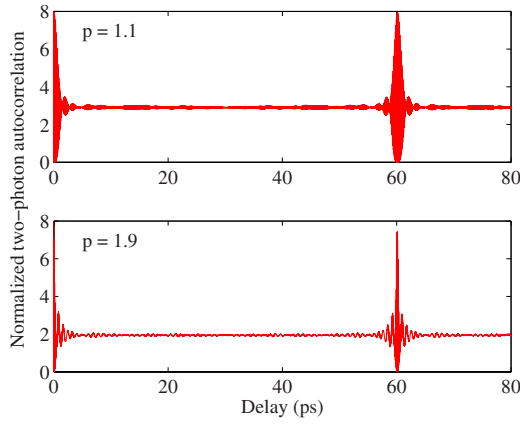


FIG. 19. (Color online) Simulated interferometric autocorrelation, corresponding to the same parameters as in Fig. 7. The value of p is denoted at each plot.

$$\frac{I(0)}{I(\tau > T)} = 8. \quad (\text{B3})$$

Now let us assume that $E(t)$ is a complex stationary random process whose phase at each point in time is uniformly distributed over the interval $[0, 2\pi]$. Let us assume for simplicity that there is T such that for $\tau > T$ $E(t)$ and $E(t + \tau)$ are statistically independent. One can see that the mean value of the third, the fourth, and the fifth line of Eq. (B1) vanish. The second line of Eq. (B1) equals $3I_b$. We therefore obtain that for a random process

$$\frac{I(0)}{I(\tau > T)} = \frac{8}{3}. \quad (\text{B4})$$

The ratio between the peak and the background in the autocorrelation function is therefore 8:1 for an isolated pulse, that is when the phases of all modes are all zero, and 8:3 for modes with completely random phases.

Figure 19 shows simulated interferometric autocorrelation traces for the same parameters as in Fig. 7. For $p=1$ the phases of the modes are random, and the peak-to-background ratio is 8:3. When the saturable absorber is absent, $\gamma=0$, the ratio in the autocorrelation is also 8:3. When the pumping is higher (Fig. 19 below), the saturable absorber induces some phase relationships between the modes, and the peak-to-background ratio is about 8:2. Note that in the autocorrelation in Fig. 12, the ratio is slightly greater than 8:3. The autocorrelation also has a nontrivial structure, with a peak at half the cavity round-trip time. We were not able to reproduce this structure in the simulations.

APPENDIX C: KERR LENSING IN QCLs

This section elaborates on the analysis of the Kerr lensing effect (including soft Kerr lensing) in QCLs. The mechanism of Kerr lensing can be understood as follows: above lasing threshold, there is net gain in the active region which compensates the mirror losses. Assuming a nonlinear refractive index $\Delta n_{NL} = n_2 I$ in the active region, it causes self-focusing of the transverse mode, which results in a net increase in the modal gain due to an increased overlap with the active region and a decreased overlap with lossy waveguide cladding. A stronger intensity leads to stronger Kerr lensing, thus forming an intensity-dependent saturable absorber.

The active region width is expected to play an important role in the Kerr lensing effect. The narrower the active region, the less confined is the transverse mode, which results in a bigger increase of the modal gain due to Kerr lensing. Therefore the saturable absorber coefficient γ is expected to be larger in lasers with a narrower active region.

We performed FDTD waveguide simulations using the commercial software BeamPROP. We simulated buried heterostructure QCLs (wafer No. 3251) with various active region widths. The software allows us to assign complex refractive indices to each layer, and the effective modal index

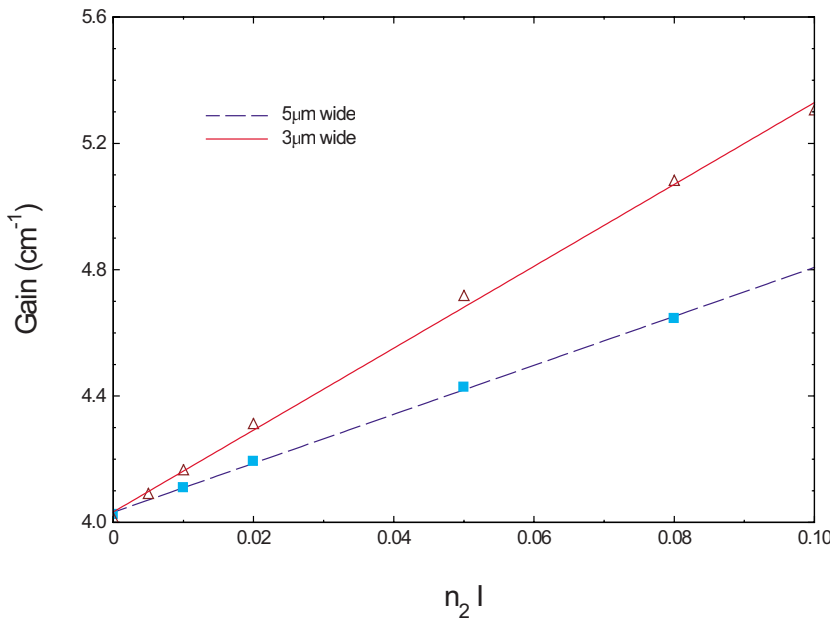


FIG. 20. (Color online) Simulation results of the modal gain vs the nonlinear refractive index $n_2 I$ for a 3 μm wide and a 5 μm wide active region in wafer No. 3251.

is calculated by the software. The detailed steps of our Kerr-lensing simulation are as follows.

(1) We first assign a net gain (i.e., a negative imaginary part of n_{active}) in the active region, making the total modal gain equal to the total mirror loss. In our case, given the reflectivity $r=0.27$ at our laser facets, a 3 mm long laser has a total mirror loss of 4 cm^{-1} .

(2) We assign different values of $\Delta n_{NL}=n_2 I$ to the real part of n_{active} . After running the simulation, we get the net modal gain for each Δn_{NL} .

(3) We plot the modal gain versus Δn_{NL} . For example, Fig. 20 shows the plots for 3 μm wide and 5 μm wide active regions.

(4) In order to relate to our theoretical model, the change in modal gain is attributed entirely to change in modal losses.

(5) The change in modal loss equals $\gamma|E|^2$, which is also proportional to the intensity I . Thus from the slope of the plot of modal gain vs $\Delta n_{NL}=n_2 I$, we can extract the ratio of γ/n_2 .

(6) For a given active region width, we choose the value of γ that brings the threshold of RNGH instability down close to our experimental data. Then we can determine the value of n_2 from the ratio of γ/n_2 .

(7) With the fixed value of n_2 , we repeat the waveguide simulation with different active region widths and plot γ vs the active region width.

For an active region width of 3 μm , the RNGH threshold is dramatically lowered from nine times the lasing threshold to less than twice the lasing threshold (see Fig. 15). From Fig. 2, the required γ is about $2 \times 10^{-9} \text{ cm/V}^2$, from which we obtain that n_2 is about $2 \times 10^{-8} \text{ cm}^2/\text{W}$. The estimated γ vs active region width is plotted in Fig. 21. As the active region width increases, γ decreases, and when the active region is as wide as 7.5 μm , γ decreases to $1 \times 10^{-9} \text{ cm/V}^2$. This pushes the RNGH instability threshold to above twice the lasing threshold. This may explain why

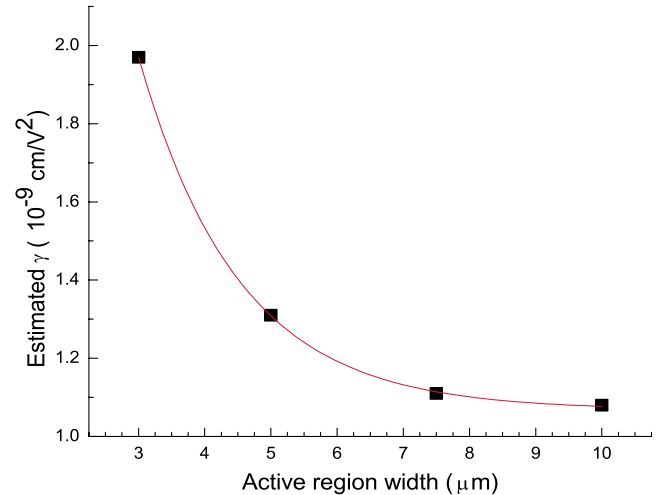


FIG. 21. (Color online) Simulation results of the saturable absorber coefficient γ vs the active region width in wafer No. 3251 assuming $n_2=2 \times 10^{-8} \text{ cm}^2/\text{W}$.

we see only instabilities resulting from spatial hole burning in wider buried heterostructure lasers (see Fig. 17).

For ridge lasers, the lossy gold contact on the sidewalls of the laser ridges will make the saturable absorber even stronger. This may explain the absence of spatial hole burning in ridge lasers.

There are several possibilities which might explain the origin of n_2 in the active region. The complex states in the injector can all contribute to both real and imaginary parts of $\chi^{(3)}$ at the laser wavelength, which results in both n_2 and γ . The laser transition itself can contribute an n_2 as big as $10^{-9} \text{ cm}^2/\text{W}$ [5]. γ should be enhanced with increasing temperature since electrons at higher temperatures populate a larger number of excited states, from which absorption can take place; and although not very significant, there is always bulk n_2 of the material.

-
- [1] A. E. Siegman, *Lasers* (University Science Books, Mill Valley, CA, 1986).
- [2] A. Yariv, *Quantum Electronics*, 3rd ed. (Wiley, New York, 1989).
- [3] H. A. Haus, *IEEE J. Sel. Top. Quantum Electron.* **6**, 1173 (2000).
- [4] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, *Science* **264**, 553 (1994).
- [5] R. Paiella *et al.*, *Science* **290**, 1739 (2000).
- [6] A. Soibel, F. Capasso, C. Gmachl, M. L. Peabody, A. M. Sargent, R. Paiella, D. L. Sivco, A. Y. Cho, and H. C. Liu, *IEEE J. Quantum Electron.* **40**, 197 (2004).
- [7] A. Soibel, F. Capasso, C. Gmachl, M. L. Peabody, A. M. Sargent, R. Paiella, H. Y. Hwang, D. L. Sivco, A. Y. Cho, H. C. Liu, C. Jirauschek, and F. X. Kärtner, *IEEE J. Quantum Electron.* **40**, 844 (2004).
- [8] R. Paiella, F. Capasso, C. Gmachl, H. Y. Hwang, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho, *Appl. Phys. Lett.* **75**, 2536 (1999).
- [9] C. Y. Wang, L. Diehl, A. Gordon, C. Jirauschek, F. X. Kärtner, A. Belyanin, D. Bour, S. Corzine, G. Höfler, M. Troccoli, J. Faist, and F. Capasso, *Phys. Rev. A* **75**, 031802(R) (2007).
- [10] H. Risken and K. Nummedal, *J. Appl. Phys.* **39**, 4662 (1968).
- [11] P. Graham and H. Haken, *Z. Phys.* **213**, 420 (1968).
- [12] R. W. Boyd, *Nonlinear Optics*, 2nd ed. (Academic Press, London, 2003).
- [13] H. Haken and H. Ohno, *Opt. Commun.* **26**, 117 (1978).
- [14] A. Gordon and B. Fischer, *Phys. Rev. Lett.* **89**, 103901 (2002).
- [15] A. Gordon, B. Vodonos, V. Smulakovsky, and B. Fischer, *Opt. Express* **11**, 3418 (2003).
- [16] Y. Silberberg, P. W. Smith, D. J. Eilenberger, D. A. B. Miller, A. C. Gossard, and W. Wiegmann, *Opt. Lett.* **9**, 507 (1984).
- [17] S. Gee, R. Coffie, P. J. Delfyett, G. Alphonse, and J. Connolly, *Appl. Phys. Lett.* **71**, 2569 (1997).
- [18] C. Gmachl *et al.*, *Appl. Phys. Lett.* **72**, 3130 (1998).
- [19] The ratio between the intracavity power and the collected output power can be derived from the reflective coefficients of the laser facets and the collection efficiency of the power meter.

- I_{ave} can be derived from the intracavity power given the overlap factor of the laser mode with the laser active region.
- [20] M. Ohtsu, Y. Otsuka, and Y. Teramachi, *Appl. Phys. Lett.* **46**, 108 (1985).
- [21] N. Chinone *et al.*, *IEEE J. Quantum Electron.* **21**, 1264 (1985).
- [22] M. Ohtsu *et al.*, *IEEE J. Quantum Electron.* **22**, 535 (1986).
- [23] A. Zavriyev *et al.*, *Opt. Lett.* **20**, 1886 (1995).
- [24] T. Maier *et al.*, *Appl. Phys. Lett.* **84**, 5162 (2004).
- [25] For a general review of ultrafast QWIPs, see H. C. Liu *et al.* *et al.*, *IEEE Circuits Devices Mag.* **19** (6), 9 (2003).
- [26] M. Beck *et al.*, *Science* **295**, 301 (2002).