
Article:
Evans, C.A., Jovanovic, V.D., Indjin, D. et al. (2006) Investigation of thermal effects in quantum-cascade lasers. IEEE Journal of Quantum Electronics, 42 (9). pp. 859-867. ISSN: 0018-9197

https://doi.org/10.1109/JQE.2006.880116

Investigation of Thermal Effects in Quantum-Cascade Lasers

Craig A. Evans, Vladimir D. Jovanović, Dragan Indjin, Zoran Ikonić, and Paul Harrison, *Senior Member, IEEE*

***Abstract*—The development of a thermal model for quantum cascade lasers (QCLs) is presented. The model is used in conjunction with a self-consistent scattering rate calculation of the electron dynamics of an InGaAs–AlAsSb QCL to calculate the temperature distribution throughout the device which can be a limiting factor for high temperature operation. The model is used to investigate the effects of various driving conditions and device geometries, such as epilayer down bonding and buried heterostructures, on the active region temperature. It is found that buried heterostructures have a factor of eight decrease in thermal time constants compared to standard ridge waveguide structures in pulsed mode and allow a ∼78% increase in heat sink temperature compared to epilayer down mounted devices in continuous-wave mode. The model presented provides a valuable tool for understanding the thermal dynamics inside a quantum cascade laser and will help to improve their operating temperatures.**

***Index Terms*—Electron dynamics, quantum-cascade lasers (QCLs), thermal modeling.**

I. Introduction

In general, quantum-cascade lasers (QCLs) suffer from large threshold currents and voltages which lead to strong local heating effects inside the device active regions [1]. These internal heating effects have significant implications on device performance independent of the material system used and the wavelength range covered. Above threshold, and especially for continuous-wave (CW) and high duty-cycle operation, the device active region can be at a considerably higher temperature than the heat sink. At these higher temperatures, more carriers in the device are thermally activated into higher energy continuum-like states and also back into the lower laser level. These factors combine to reduce the population inversion between the laser levels and, hence, the laser gain decreases. This makes room temperature CW operation of QCLs difficult to achieve without special attention paid to the device thermal management.

Despite these difficulties, room-temperature CW operation has been achieved as a result of both optimized active region designs and improved laser packaging. The first report of room temperature CW operation was a $\lambda \sim 9.1$ μm InGaAs–AlInAs QCL utilizing a buried heterostructure configuration mounted epilayer side down on a diamond heat sink [2]. The use of a buried heterostructure allows heat to be dissipated in all directions from the active region compared to a standard ridge or double channel waveguide in which the heat is primarily removed in the direction perpendicular to the active region layers. By mounting the device epilayer down, the heat generating active region can be in closer proximity to the heat sink than in equivalent substrate side mounted devices; this improves heat dissipation.

Other room temperature CW QCLs emitting in the range $\lambda \sim 4$ to 6 μm have been reported employing strain-balanced InGaAs–AlInAs active regions ($\Delta E_c \approx 800$ to 900 meV) [3]–[7]. The use of strain increases the conduction band offset and reduces electron leakage into the continuum at shorter wavelengths. These devices have employed a variety of packaging techniques with the aim of improving thermal performance including the use of exclusively InP waveguides (due to InP having a higher thermal conductivity than InGaAs) [7], thick gold layers electroplated on top of the laser ridge to help remove waste heat [5], the use of buried heterostructures [3] and epilayer down mounting [6]. This had led to QCLs operating with high power at room temperature in CW mode (480 mW) and at temperatures as high as 363 K [7].

Recently InGaAs–AlAsSb QCLs lattice matched to InP substrates have been reported [8]. The InGaAs–AlAsSb heterostructure has a very large conduction band offset of ∼1.6 eV which is almost double that of the strained InGaAs–AlInAs heterostructure and makes this material system inherently less sensitive to temperature. Also, unlike early GaAs-based QCLs which required a new type of waveguide to be developed in order to improve performance [9], InGaAs–AlAsSb QCLs are compatible with the well-established InP-based waveguides intially developed for use in InGaAs–AlInAs QCLs. These factors led to the rapid development of high temperature InGaAs–AlAsSb QCLs, albeit in pulsed mode. An InGaAs–AlAsSb QCL emitting at $\lambda \sim 4.5$ μm up to at least 400 K in pulsed mode has been reported [10]. The device was processed into a standard double channel waveguide and mounted substrate side down with no particular attention paid to optimizing the structure to improve the thermal management. This led to threshold currents approximately three times higher at 300 K than those in a similar wavelength QCL in the InGaAs–AlInAs material system (~ 1.96 kA/cm^2) with an exclusive InP waveguide mounted epilayer down [6], a factor which prevented high temperature CW operation being achieved.

Manuscript received March 23, 2006; revised May 24, 2006. This work was supported in part by the Engineering and Physical Sciences Research Council, U.K.

C. A. Evans, D. Indjin, Z. Ikonić, and P. Harrison are with the Institute of Microwaves and Photonics, School of Electronic and Electrical Engineering, University of Leeds, Leeds LS2 9JT, U.K. (e-mail: een0cae@leeds.ac.uk; d.indjin@leeds.ac.uk; z.ikonic@leeds.ac.uk; p.harrison@leeds.ac.uk).

V. D. Jovanović was with the Institute of Microwaves and Photonics, School of Electronic and Electrical Engineering, University of Leeds, Leeds LS2 9JT, U.K.. He is now with Lehman Brothers, London E14 5LE, U.K. (e-mail: vladimir.jovanovic@lehman.com).

Digital Object Identifier 10.1109/JQE.2006.880116

Although to date there has been no report of CW emission from InGaAs–AlAsSb QCLs, CW emission has been observed from QCLs in the similar InGaAs–AlGaAsSb material system. Incorporation of Ga into the barrier material reduces the conduction band offset to ∼1–1.2 eV, depending on the exact composition of the quarternary layer. This increases the electron tunnelling probability and relaxes the requirement for the growth of ultrathin layers while still providing sufficient quantum confinement for short wavelength emission. Following the report of laser emission from a $\lambda \sim 4.9$ μm InGaAs–AlGaAsSb QCL up to at least 400 K in pulsed mode [11], CW operation up to 94 K from the same active region design was achieved by covering the top and sides of the laser ridge in 2.5-μm-thick electroplated gold in order to promote heat removal in the lateral direction [12].

It is clear that in order to improve the temperature performance of QCLs, the thermal dynamics inside the device must be understood. Active region heating is inevitable, but can be reduced by lowering the threshold current densities and improving heat dissipation by both optimizing the active region design and by implementing one of (or a combination of) the thermal management techniques outlined above. In this work we outline the development of a two-dimensional thermal model of a QCL and use it in conjunction with our self-consistent carrier transport calculation [13]–[15] to investigate the effects of device geometry and drive conditions on the thermal dynamics inside the QCL. The investigation will center on a $\lambda \sim 3$ μm InGaAs–AlAsSb QCL, which was the focus of a previously reported simulation [15] of an experimentally measured device [16].

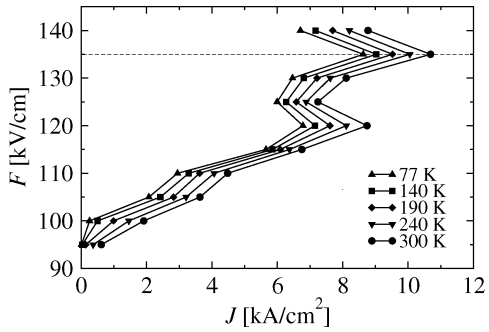


Fig. 1. Electric field versus current density characteristics for a range of active region temperatures. The operating bias point (135 kV/cm) is denoted by the dotted line.

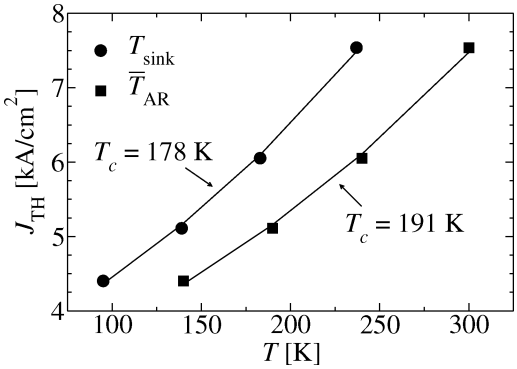


Fig. 2. Calculated threshold current density as a function of temperature. The squares denote the active region temperature; the circles the corresponding heat sink temperatures for 100-ns pulses at 10 kHz. The given characteristic temperatures (T_c) are calculated using an exponential fit of the form $J_{\rm TH} = J_0 + J_1 \exp(T/T_c)$.

II. Theoretical Background

A. Electron Dynamics Simulation

The electron dynamics inside the device active region were calculated using a self-consistent scattering rate approach [15]. The electron energies and wavefunctions were firstly obtained by solving the Schrödinger equation within the envelope function and effective mass approximations with the energy dependent effective mass taken into account via Kane's two-band model. The wavefunctions were then used to evaluate all the principal electron–electron and electron–longitudnal optical (LO) phonon intra- and inter-period scattering rates [17]. The resulting rate equations must be solved self-consistently since the subband populations depend upon both the initial and final populations, and, hence, the process was repeated until the subband populations converged. The self-consistent solution yields the nonequilibrium electron density and the lifetime of a carrier in each of the subbands, from which the total current density J and the local gain g can be calculated. By repeating the above procedure at different values of electric field strength F and at different lattice temperatures, F–J curves can be generated for each temperature. The self-consistent procedure was run for the $\lambda \sim 3$ μm InGaAs–AlAsSb QCL and the F–J curves generated for several temperatures are plotted in Fig. 1. In order to calculate the modal gain G_M, the local gain g was calculated from the population inversion and the waveguide parameters calculated. The waveguide parameters (the confinement factor Γ, waveguide absorption losses $\alpha_{\rm wg}$ and mirror losses α_m) were calculated using a transfer-matrix technique [18]. In order to include the effect of temperature on the waveguide parameters, a temperature dependent mobility [19] was used to calculate the complex refractive index of the waveguide layers. The effect of temperature on the waveguide losses was found to be described by $\alpha_{\rm wg} = 3.62 + 0.39\exp(T/194.59)$ cm^{-1} and the mirror losses estimated to be 4.16 cm^{-1} at all temperatures. The modal gain was calculated using $G_M = \Gamma g$. The threshold current density $J_{\rm TH}$ is given by the value of current at which G_M exceeds the total losses, $\alpha_{\rm wg} + \alpha_m$. The threshold current density as a function of active region temperature is shown in Fig. 2. An exponential fit of the usual form $J_{\rm TH} = J_0 + J_1 \exp(T/T_c)$ was applied to the data and a characteristic temperature of $T_c = 191$ K estimated (denoted by the squares in Fig. 2). The calculations show that the InGaAs–AlAsSb QCL has enough gain to achieve laser emission up to a temperature of 300 K.

B. Development of the Thermal Model

The thermal model is based upon the heat diffusion equation in two dimensions

$$\rho(x,y,T)c_p(x,y,T)\frac{\partial T(x,y,t)}{\partial t} = \nabla\cdot[\kappa(x,y,T)\nabla T(x,y,t)] + S(x,y,t) \quad (1)$$

where ρ is the density of the material $(\mathrm{kg/m^3})$, c_p is the specific heat capacity of the material (J/kg K), T is the temperature (K), κ is the thermal conductivity (W/m K), and S is the power density of the heat source $(\mathrm{W/m^3})$. Since ρ, c_p, and κ are temperature dependent (it will be shown later that S is also temperature dependent), (1) is nonlinear and is solved using a finite-difference approach. In this method, the QCL structure is meshed into a nonoverlapping grid which defines the boundaries between the individual control volumes and the temperature is found at the center (or node) of each of these control volumes. Equation (1) is expanded and both the spatial and time derivatives are approximated by finite-differences:

$$\begin{aligned}\rho_{i,j}c_{p_{i,j}}&\left(\frac{T_{i,j}^{m+1}-T_{i,j}^{m}}{\Delta t}\right)\\&=\frac{\kappa_{i+\frac{1}{2},j}\left(T_{i+1,j}^{m}-T_{i,j}^{m}\right)}{\Delta x^2}-\frac{\kappa_{i-\frac{1}{2},j}\left(T_{i,j}^{m}-T_{i-1,j}^{m}\right)}{\Delta x^2}\\&\quad+\frac{\kappa_{i,j+\frac{1}{2}}\left(T_{i,j+1}^{m}-T_{i,j}^{m}\right)}{\Delta y^2}-\frac{\kappa_{i,j-\frac{1}{2}}\left(T_{i,j}^{m}-T_{i,j-1}^{m}\right)}{\Delta y^2}\\&\quad+S_{i,j}^{m}\end{aligned}\tag{2}$$

where i and j are the indexes to the control volumes (of size Δx by Δy) in the x and y direction, respectively, and m is an integer counter of the time step Δt. The midpoint values of the thermal conductivity were calculated as the average of the conductivities in the appropriate control volumes. Equation (2) is solved using Euler's forward method, which gives the nodal temperature at a point (i, j). Since QCL experimental measurements are typically made *in vacuo*, Neumann (zero-derivative) boundary conditions are applied to the surfaces exposed to the vacuum, with a constant temperature boundary condition applied to the surface in contact with the heat sink. The temperatures at each node were found in every time step to give a complete time-dependent picture of the thermal dynamics inside a QCL structure.

The source power density S is applied to the device active region and is given by $S = FJ$, with F and J calculated from the electron dynamics simulation as described in Section II-A. From Fig. 1 it can be seen that for a given value of F, J is increasing with temperature. By extracting the values of J at the operating bias point (dashed line in Fig. 1) for each temperature, S was calculated as a function of temperature $(= \ 1000\exp(T/1023)\ [\mathrm{GW/m^3}])$. This "quantum source term" can then be used in the thermal simulation to succinctly describe the QCL active region heating effects due to quantum mechanical processes. Both pulsed and CW operating conditions can be simulated by varying the time dependency of S accordingly.

Fig. 3 shows an example of a typical temperature-time profile for a standard ridge waveguide operating in pulsed mode. It can be seen that at the end of the first period, the active region has not regained the heat sink temperature and so the active region temperature increases further during the next period. This heat accumulation effect causes the active region to progressively rise in temperature during each successive pulse until saturation occurs and the device reaches a steady state. In the following simulations, the active region temperatures are extracted once this steady state has been reached. The active region of the QCL is made up from a multitude of thin semiconductor layers and due to extra phonon scattering channels at the interfaces between these layers [20], [21] the thermal conductivity perpendicular to the layers $(\kappa_\perp)$ is found to be approximately an order of magnitude smaller than the thermal conductivity in the plane of the layers $(\kappa_{||})$ [22]. Fig. 4 shows the effect of the different ratios of $\kappa_{||}$ to $\kappa_\perp$ on the average active region temperature (T_{AR}) of the InGaAs–AlAsSb QCL at the center of this work. A heat sink temperature (T_{sink}) of 200 K with 100-ns pulses at a frequency of 100 kHz is used in the simulation. It can be seen from Fig. 4 that with an isotropic thermal conductivity, the average T_{AR} is 222 K and increases with the ratio before starting to saturate at $\kappa_{||}/k_\perp \approx 10$. Since the ratio is not known exactly, in this work $\kappa_{||}$ will be taken to be that of bulk InGaAs, and $\kappa_\perp$ a tenth of the bulk value in order to simulate the worst case scenario.

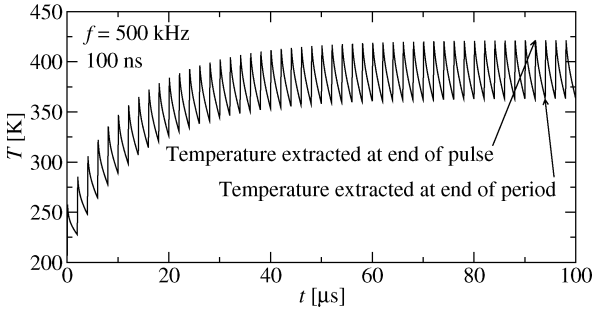


Fig. 3. Temperature-time profile for standard ridge waveguide showing the average active region temperature as a function of time. $T_{\mathrm{sink}} = 200$ K and the repetition frequency is 500 kHz with 100-ns pulses. In the following simulations, the active region temperatures are extracted once they have reached saturation and the device is in steady state.

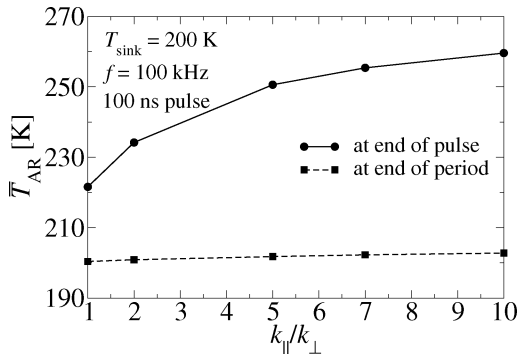


Fig. 4. Average T_{AR} for different levels of anisotropic active region thermal conductivity. For the case of $\kappa_{||}/\kappa_\perp = 1$, the thermal conductivity is isotropic. $T_{\mathrm{sink}} = 200$ K and the pulsewidth is 100 ns with a 100-kHz repetition rate.

III. Numerical Results and Discussion

Using the thermal model outlined in the previous section, the influence of various aspects of the QCL device design and driving conditions on the temperature performance of the device can be investigated. As discussed in Section II-A, the maximum operating temperature of the $\lambda \sim 3\ \mu$m InGaAs–AlAsSb QCL active region is predicted to be 300 K and the thermal model can be used to find the corresponding T_{sink} for various geometries and driving conditions. In the following sections the QCL active region and waveguide cladding layers are as described in [15]

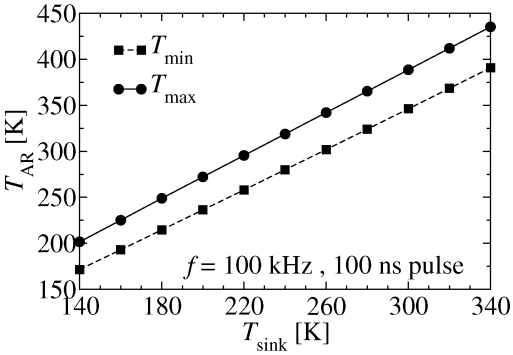


Fig. 5. Spatial minimum and maximum values of $T_{\rm AR}$ for different values of $T_{\rm sink}$ at 100 kHz with 100-ns pulsewidths. The values of $T_{\rm AR}$ are extracted at the end of the pulse.

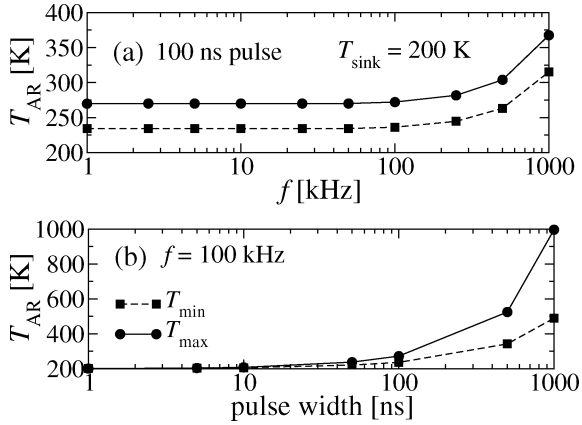


Fig. 6. (a) $T_{\rm AR}$ for different pulse frequencies with a pulsewidth of 100 ns. (b) $T_{\rm AR}$ for different pulsewidths with $f = 100$ kHz. In both cases $T_{\rm sink} = 200$ K and the spatial minimum and maximum values of $T_{\rm AR}$ are shown and are extracted at the end of the pulse.

with the device itself considered to be a standard ridge waveguide of width 10 μm and etched to a depth of 9 μm, mounted substrate side down.

A. Effect on Threshold Current Density

The thermal model was used with 100-ns pulses at 10 kHz to calculate the corresponding $T_{\rm sink}$ for each $T_{\rm AR}$ point in Fig. 2 (denoted by circles in Fig. 2). The low repetition frequency was used in order to let the device active region regain the heat sink temperature before the arrival of the next pulse and avoid the heat accumulation effects depicted in Fig. 3. Since the temperature profile of a QCL active region is nonuniform, the circles in Fig. 2 denote the heat sink temperatures at which the calculated *average* active region temperature is equal to the temperature used in the electron dynamics simulation [as denoted by the squares in Fig. 2)]. The same exponential function as before was fitted to the heat sink points and a characteristic temperature of $T_c = 178$ K found, compared to 191 K previously (see Fig. 2). This shows that the threshold current density increases more quickly with the $T_{\rm sink}$ rather than $T_{\rm AR}$ and this could be the reason why calculated threshold current densities are often underestimated at higher temperatures [23], [24].

B. Influence of Heat Sink Temperature

Fig. 5 shows the effect of $T_{\rm sink}$ on the $T_{\rm AR}$ of the InGaAs–AlAsSb QCL for 100-ns pulses at 100 kHz. Both the minimum and maximum active region temperatures at the end of a pulse are shown in the figure. It can be seen from Fig. 5 that $T_{\rm AR}$ increases linearly with $T_{\rm sink}$. The relationship is of the form $T_{\rm AR} = T_1 + \alpha_{\rm s-AR} T_{\rm sink}$, where $\alpha_{\rm s-AR}$ is a coupling constant between the heat sink and active region and is extracted from straight line fits to the data. For the case of $T_{\rm AR}^{\rm max}$, $\alpha_{\rm s-AR} = 1.167$, while for $T_{\rm AR}^{\rm min}$ it is equal to 1.097. Hence, as $\alpha_{\rm s-AR}^{\rm max} > \alpha_{\rm s-AR}^{\rm min}$, the temperature range inside the active region increases with increasing $T_{\rm sink}$ which could cause damaging thermal stress effects at high sink temperatures. The cause of the temperature range increase is probably due to the thermal conductivity of the materials in the devices decreasing with temperature. The hotter areas of the active region have a lower thermal conductivity than the cooler areas and are less able to dissipate heat leading to an increase in temperature, while the cooler areas are better able to dissipate heat and, hence, the temperature rise is not as great.

For the case of the InGaAs–AlAsSb QCL in question with 100-ns pulses at 100 kHz (1% duty cycle) and $T_{\rm sink} = 220$ K, the entirety of the active region is less than 300 K, meaning laser action would be expected for these driving conditions. With $T_{\rm sink} = 240$ K, the active region has a temperature range of $\sim$260–325 K, meaning that only the active region periods at $T < 300$ K would contribute to lasing resulting in a decrease in output power. Even at $T_{\rm sink} = 255$ K, some of the active region are still less than 300 K and so laser action may still be present at this temperature.

The fact that the temperature gradient inside the active region increases with heat sink temperature could help to explain the abrupt decrease in slope efficiency at higher temperatures which has been measured experimentally [25]. When the device is being operated at close to its maximum temperature, not all of the active region periods may be contributing to laser action. As the heat sink temperature is increased, the temperature gradient is increased further and less periods will contribute to laser action. This process will cause the output power to decrease further (and, hence, the slope efficiency) until eventually all of the active region periods will cease lasing and the device reaches cut-off. In addition, if each of the active region periods are at a slightly different temperature, the amount of red shift in the emission wavelength will vary, leading to a broadening of the linewidth.

C. Influence of Pulse Frequency

The duty cycle can be increased by increasing either the pulse repetition frequency or the pulsewidth (or a combination of both). The same device as described in the previous section was simulated with $T_{\rm sink} = 200$ K and a pulsewidth of 100 ns at different repetition rates. The effect of an increased pulse repetition frequency on $T_{\rm AR}$ is shown in Fig. 6(a). For frequencies of less than 50 kHz, the temperature range of the active region is independent of the pulse frequency. This is due to the relatively long time period between pulses, which gives the active region enough time to entirely dissipate the heat accumulated during the pulse and recover the heat sink temperature.

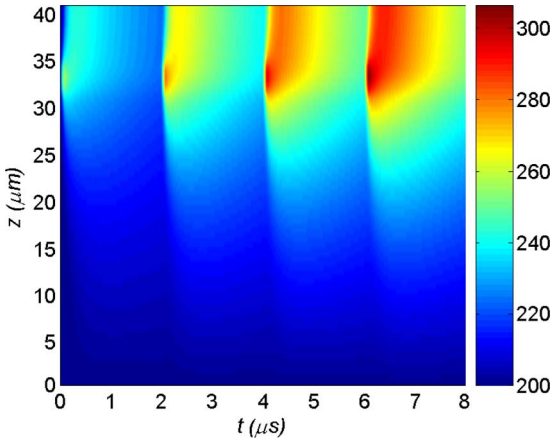


Fig. 7. Time evolution of the QCL cross-sectional temperature profile. $T_{\rm sink} = 200$ K and the time period between the 100-ns pulses is 2 μs (500 kHz). (Color version is available online at http://ieeexplore.ieee.org.)

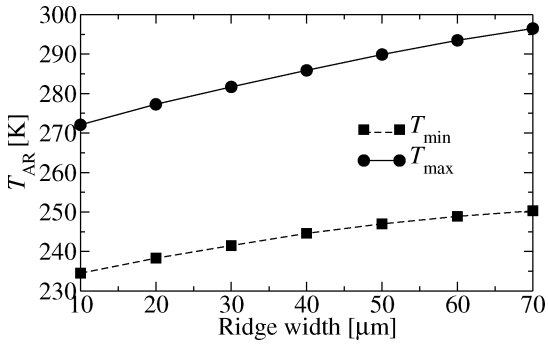


Fig. 8. Spatial minimum and maximum values of $T_{\rm AR}$ for different ridge widths with an etch depth of 9 μm. $T_{\rm sink} = 200$ K and the pulsewidth is 100 ns with a 100-kHz repetition rate. Values of $T_{\rm AR}$ are extracted at the end of the pulse.

For these cases, the only factors which affect the active region temperature range are the heat sink temperature (which determines the thermal conductivities of the materials and the source power density) and the pulsewidth (determines how much the active region heats up). These are constant in this simulation and, hence, the temperature profiles in Fig. 6(a) are flat for frequencies below 50 kHz. Above this value of frequency the active region temperatures start to rise rapidly due to heat accumulation effects. This heat accumulation effect is outlined in Fig. 7 which shows the device cross section temperature profile [taken along the line of symmetry in Fig. 9(a)] evolving with time at a frequency of 500 kHz (period length 2 μs). It can clearly be seen how the heat in the device builds up during each successive pulse. For the InGaAs–AlAsSb QCL in question, with 100-ns pulses and $T_{\rm sink} = 200$ K, a repetition frequency of $\sim$600 kHz (6% duty cycle) is possible without the entire active region being above 300 K and, hence, laser action could occur. An advantage of intersubband devices over interband devices is the ultrafast carrier lifetimes (on the order of picoseconds) meaning extremely high modulation frequencies in gigahertz should be possible [26]. Figs. 6 and 7 highlight the fact that to achieve high frequency modulation (and indeed CW operation) in QCLs, careful thermal management is required to improve the heat dissipation away from the active region and reduce the heat accumulation effects.

D. Influence of Pulse Width

The effect of increasing the duty cycle by increasing the pulsewidth on $T_{\rm AR}$ is shown in Fig. 6(b). For pulsewidths of less than 10 ns, very little active region heating occurs. Above this value there is a considerable increase in active region temperature with pulsewidth. This is due in principal to the same heat accumulation effects that are outlined in Fig. 6(a) but on a more pronounced scale. In Fig. 6(a) for a 100-ns pulse, 1000 kHz corresponds to a 10% duty cycle and the maximum active region temperature is $\sim$370 K. In the case of Fig. 6(b), a 10% duty cycle is equivalent to a 1000-ns pulse and the maximum active region temperature is $\sim$1000 K. This is caused by the rate of temperature increase during the pulse being far greater than the rate of temperature decrease during cooling phase. Therefore, it is better to increase the duty cycle by increasing the pulse repetition frequency rather than increasing the pulsewidth. In terms of the InGaAs–AlAsSb QCL for a 100-kHz repetition frequency at a 200 K heat sink temperature, a maximum pulsewidth of $\sim$200 ns (2% duty cycle) is possible without the entire active region temperature rising above 300 K.

E. Influence of Ridge Width

The previous sections have investigated the effects of the QCL driving conditions on the active region heating. The thermal model also allows the influence of changing the device structure on the heating effects inside the QCL to be investigated. All the previous simulations have been carried out using a standard ridge waveguide 10 μm wide and etched to a depth of 9 μm. Fig. 8 shows the influence of increasing the ridge width on $T_{\rm AR}$. It can be seen from Fig. 8 that both the active region temperature and the temperature range inside the active region increase with ridge width which agrees with experiment [27]. This is explained by a wider laser ridge having a greater active region volume meaning higher power and therefore more heating. Narrow ridges would therefore be advantageous at high temperatures from a thermal viewpoint although decreasing the ridge width would also decrease the maximum power due to the smaller emitting volume.

F. Comparison of Different Device Mountings

In the previous sections, the investigations have been carried out on a standard ridge waveguide mounted substrate side down. In this configuration [which is shown schematically in Fig. 9(a)], heat must escape from the active region through the substrate to the heat sink. Efficient heat transfer in this direction is prevented by the small $\kappa_\perp$ and consequently substantial active region heating occurs. Alternate devices geometries are possible which improve the heat dissipation from the active region and the schematic cross sections of these are shown in Fig. 9. For the standard ridge waveguide Fig. 9(a), etching defines the laser ridge after growth by molecular beam epitaxy (MBE), the sidewalls of the ridge are coated in a layer of insulating Si_3N_4 and an opening on top of the ridge is made for the top contact. For the case of the gold covered double-channel ridge waveguide Fig. 9(b), the growth process is similar to that for the standard ridge waveguide except that the laser ridge is defined

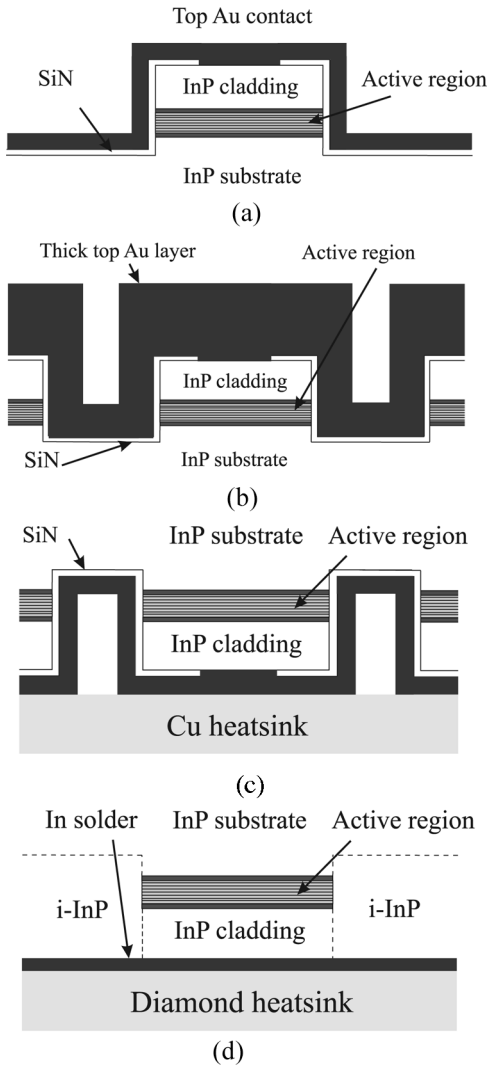


Fig. 9. Schematic cross sections of different InP-based QCL geometries (not to scale). (a) Standard ridge waveguide mounted substrate side down. (b) Double-channel ridge waveguide mounted substrate side down with a thick gold layer electroplated on top of the laser ridge to help dissipate the heat. (c) Double channel waveguide mounted epilayer down on a copper heat sink. (d) Buried heterostructure mounted epilayer side down on a diamond heat sink.

by the etching of two channels either side of the ridge. After the Si_3N_4 insulating layer and the top contact are deposited, a thick gold layer is electroplated on top of the laser ridge to efficiently spread the current and heat on the laser surface and the device mounted substrate side down [5]. The growth of the epilayer down mounted double-channel ridge waveguide Fig. 9(c) is in principle the same as for the gold covered double-channel structure. After growth, the device is bonded epilayer down on a copper heat sink using indium solder. In the following simulations, it is assumed that the indium solder between the heat sink and the top contact is pressed out during mounting, as described in [12]. The fabrication of buried heterostructures Fig. 9(d) is a more complex task compared to the previously mentioned structures. After molecular beam epitaxy (MBE) growth of the QCL active region and n-doped InP cladding layers. The Si_3N_4 insulating layer is replaced by the lateral regrowth of insulating InP by metal–organic chemical vapor deposition (MOCVD) [28].

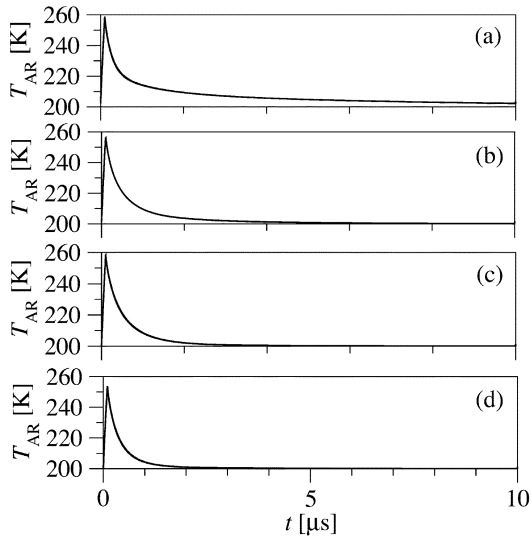


Fig. 10. Temperature-time profiles (average T_{AR}) for (a) a standard ridge waveguide, (b) a double channel waveguide mounted substrate side down with a 5-μm-thick gold layer deposited on top, (c) a double channel waveguide mounted epilayer down on a copper heat sink, and (d) buried heterostructure mounted epilayer down on a diamond heat sink. All simulations are at $T_{\mathrm{sink}} = 200$ K, with 100-ns pulses at a repetition rate of 100 kHz. The circles represent the points at which the cross section temperature profiles in Fig. 11 are extracted.

The use of i-InP improves performance compared to Si_3N_4 due to its higher thermal conductivity and it allows heat to escape from the active region in the plane of, as well as perpendicular to, the active region layers. In order to further improve the performance of the devices, buried heterostructures have been used bonded epilayer down on diamond heat sinks. This configuration led to the first demonstration of QCL operating in CW mode at room temperature [2].

In order to compare the different device configurations, the temperature-time profiles and device cross section temperature profiles were calculated for each. The average active region temperatures as a function of time for the different configurations are shown in Fig. 10. For the ridge waveguide (a), the average T_{AR} at the end of the pulse is 262 K. Both the double-channel ridge waveguide mounted epilayer-down (b) and the gold covered double-channel ridge waveguide mounted substrate side down (c) have the same T_{AR} of $\sim$ 258 K at the end of the pulse. For the case of the buried heterostructure (d), this is reduced slightly to $\sim$252 K. In this instance, due to the small pulsewidth, the effect of the different configurations on the maximum active region temperature is relatively small, however it does play a large part in the ability of the device to dissipate heat away from the active region. In order to characterize the ability of the devices to dissipate heat, the cooling phase of the temperature-time profiles were fitted using a second-order exponential decay function [29]. Table I shows the fitted equations for each structure.

It can be seen that for the case of the devices mounted substrate side down [(a) and (b)], a_0 is slightly higher than T_{sink} due to the slight heat accumulation effects in these configurations. However, for the epilayer down mounted devices, no heat accumulation effects occur and so a_0 is very close to T_{sink}.

TABLE I
RESULTS OF FITTING THE COOLING PHASES OF EACH STRUCTURE (FIG. 10) TO A SECOND-ORDER EXPONENTIAL DECAY FUNCTION

$T_{AR} = a_0 + a_1 e^{-t/\tau_1} + a_2 e^{-t/\tau_2}$ $(t > 100$ ns)					
Device	a_0	a_1	$\tau_1(\mu s)$	a_2	$\tau_2(\mu s)$
a	203.0	47.5	0.19	32.1	2.67
b	203.1	75.5	0.17	13.7	1.98
c	200.1	32.4	0.14	52.3	0.47
d	200.1	43.7	0.12	45.3	0.33

The second-order exponential fit indicates that there are two distinct cooling stages (denoted by the two thermal time constants τ_1 and τ_2). τ_1 represents heat initially escaping from the active region into the waveguide cladding and insulation layers primarily in the vertical direction. The substrate side mounted devices are similar in this respect and, hence, the values of τ_1 for these devices are relatively similar. Although device (c), the double-channel ridge waveguide mounted epilayer down, has a similar structure to devices (a) and (b), the upper cladding layer is in closer proximity to the heat sink and, hence, its value of τ_1 is smaller. In the case of device (d), the buried heterostructure, heat diffusion can take place in both the vertical and lateral directions and in combination with the epilayer down mounting causes the buried heterostructure to have the smallest value of τ_1. The time constant τ_2 represents the second stage of cooling in which the heat diffuses towards the heat sink through the substrate and cladding regions. It appears from the data that this mechanism dominates the cooling process and is the major difference between the different device mountings. τ_2 is much longer in devices (a) and (b), compared to devices (c) and (d), since they are mounted substrate side down and so the heat generated in the active region has to diffuse through the substrate to reach the heat sink. τ_2 is smaller in device (b) compared to device (a) since heat can spread into the thick gold layer on top of the laser ridge as well as into the substrate which is the only heat channel in device (a). Devices (c) and (d) have values of τ_2 several times smaller than devices (a) and (b) due to the epilayer down mounting and, hence, the active region is in closer proximity to the heat sink. τ_2 is smaller in the buried heterostructure and this is likely due to both the effect of heat escaping both vertically and laterally, and the fact that diamond has a higher thermal conductivity than copper. Fig. 11 shows the cross section temperature profiles for each of the devices at $t = 1$ μs.

Devices (a) and (b) have similar temperature profiles, with heat primarily escaping through the substrate, although the thick gold layer causes a reduction in the average active region temperature at $t = 1$ μs from 225 K in the standard ridge to 210 K in the gold covered device. Fig. 11(d) shows the heat escaping from the buried heterostructure active region in all directions compared to the primarily vertical heat channel in the epilayer down mounted double-channel ridge waveguide Fig. 11(c), leading to a lower average active region temperature.

Table I and Fig. 11 highlight the fact that a buried heterostructure is the best solution in terms of thermal management of QCLs, especially in conjunction with epilayer down mounting.

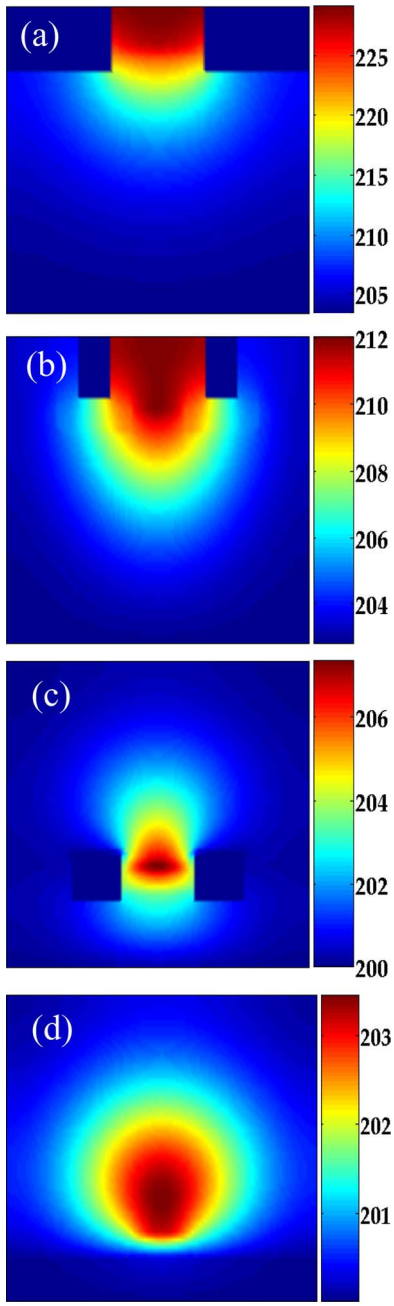


Fig. 11. Cross sectional temperature profiles for each of the devices in Fig. 10 1 μs into the period (0.99 μs into the cooling stage). Note the different temperature scales. (Color version is available online at http://ieeexplore.ieee.org.)

As well as the reduced active region temperature, the temperature profile inside the active region is much more uniform and so the majority of the periods will be at the same temperature. Despite this improvement, the increased performance must be balanced against the much more complicated processing techniques required.

G. Investigation Into CW Operation

The possibility of CW operation was investigated for structures (c) and (d). As in the case of Fig. 5, a linear relationship is found between the heat sink and active region temperature under CW operating conditions. For the case of the epilayer

down mounted device (c), $T_{\mathrm{AR}} = 139.1 + 3.9T_{\mathrm{sink}}$ and for the buried heterostructure, $T_{\mathrm{AR}} = 73.9 + 3.2T_{\mathrm{sink}}$. It can be seen that the active region temperature in the buried heterostructure is much less sensitive to the heat sink temperature than in the epilayer down mounted device, further emphasizing the fact that buried heterostructures offer the best thermal management solution for QCLs. For the InGaAs–AlAsSb QCL active region with a maximum active region temperature of 300 K that has been at the center of this work, the use of an epilayer down configuration would allow CW operation up to a heat sink temperature of ∼40 K, while incorporating a buried heterostructure would allow the maximum heat sink temperature to increase by ∼78% to ∼71 K.

IV. Conclusion

In conclusion, a thermal model of a QCL has been presented. The model takes into account the temperature-dependent material parameters and the anisotropic active region thermal conductivity, although in order to improve the thermal modeling of QCLs, experimental results on the temperature behavior and anisotropicity of the thermal conductivity of ternary semiconductor superlattices used in QCL active regions are required. The heating effects in the QCL active region were calculated using a self-consistent scattering rate model and then inserted into the developed thermal model to investigate the influence of various driving conditions and device geometry on the QCL active region temperature and, hence, the performance. Buried heterostructures are found to be the best solution in terms of thermal management due to heat being able to dissipate from the active region in all directions, leading to thermal time constants several times smaller than those found in standard ridge waveguides. Under CW operating conditions, buried heterostructures are found to allow the maximum heat sink temperature to increase by ∼78% compared to epilayer down mounted devices.

Although the thermal model presented here has focused on an InGaAs–AlAsSb QCL emitting in the near-infrared, it is equally applicable to all QCL material systems and emission wavelengths. In particular, the thermal model could be particularly useful for improving the temperature performance of terahertz QCLs. Thermal management is of the up most importance at these longer wavelengths due to the small energy gap between levels and reduced conduction band offset compared to near- and midinfrared QCLs and makes the population inversion much more sensitive to temperature. It could also be useful for closing the gap in temperature performance between InP-based and GaAs-based midinfrared QCLs. InP-based QCLs have been reported operating in CW above room temperature using a variety of thermal management techniques, while GaAs-based QCLs operating at room temperature have been limited to pulsed mode. GaAs has a lower thermal conductivity than InP and so knowledge of the thermal dynamics inside GaAs-based QCLs must be understood in order to improve the temperature performance.

References

[1] C. Faugeras, S. Forget, E. Boer-Duchemin, H. Page, J. Bengloan, O. Parillaud, M. Calligaro, C. Sirtori, M. Giovannini, and J. Faist, "High-power room temperature emission quantum cascade lasers at $\lambda = 9\ \mu$m," *IEEE J. Quantum Electron.*, vol. 41, no. 12, pp. 1430–1438, Dec. 2005.

[2] M. Beck, D. Hofstetter, T. Aellen, J. Faist, U. Oestrele, M. Ilegems, E. Gini, and H. Melchior, "Continuous wave operation of a midinfrared semiconductor laser at room temperature," *Science*, vol. 295, pp. 301–305, 2002.

[3] A. Evans, J. S. Yu, J. David, L. Doris, K. Mi, S. Slivken, and M. Razeghi, "High-temperature, high-power, continuous-wave operation of buried heterostructure quantum-cascade lasers," *Appl. Phys. Lett.*, vol. 84, no. 3, pp. 314–316, 2004.

[4] A. Evans, J. S. Yu, S. Slivken, and M. Razeghi, "Continuous-wave operation of $\lambda \sim 4.8\ \mu$m quantum-cascade lasers at room temperature," *Appl. Phys. Lett.*, vol. 85, no. 12, pp. 2166–2168, 2004.

[5] J. S. Yu, S. Slivken, S. R. Darvish, A. Evans, B. Gokden, and M. Razeghi, "High-power, room-temperature, and continuous-wave operation of distributed-feedback quantum-cascade lasers at $\lambda \sim 4.8\ \mu$m," *Appl. Phys. Lett.*, vol. 87, no. 041104, 2005.

[6] J. S. Yu, S. R. Darvish, A. Evans, J. Nguyen, S. Slivken, and M. Razeghi, "Room-temperature continuous-wave operation of quantum-cascade lasers at $\lambda \sim 4\ \mu$m," *Appl. Phys. Lett.*, vol. 88, no. 041111, 2006.

[7] A. Evans, J. Nguyen, S. Slivken, J. S. Yu, S. R. Darvish, and M. Razeghi, "Quantum-cascade lasers operating in continuous-wave mode above 90 °C at $\lambda \sim 5.25\ \mu$m," *Appl. Phys. Lett.*, vol. 88, no. 051105, 2006.

[8] D. G. Revin, L. R. Wilson, E. A. Zibnik, R. P. Green, and J. W. Cockburn, "InGaAs/AlAsSb quantum cascade lasers," *Appl. Phys. Lett.*, vol. 85, no. 18, pp. 3992–3994, 2004.

[9] C. Sirtori, P. Kruck, S. Barbieri, H. Page, and J. Nagle, "Low-loss Al-free waveguides for unipolar semiconductor lasers," *Appl. Phys. Lett.*, vol. 75, no. 25, pp. 3911–3913, Dec. 1999.

[10] Q. Yang, C. Manz, W. Bronner, C. Mann, L. Kirste, K. Köhler, and J. Wagner, "GaInAs/AlAsSb quantum-cascade lasers operating up to 400 K," *Appl. Phys. Lett.*, vol. 86, no. 131107, 2005.

[11] Q. Yang, C. Manz, W. Bronner, L. Kirste, K. Köhler, and J. Wagner, "GaInAs/AlGaAsSb quantum-cascade lasers," *Appl. Phys. Lett.*, vol. 86, no. 131109, 2005.

[12] Q. Yang, C. Manz, W. Bronner, R. Moritz, C. Mann, G. Kaufel, K. Köhler, and J. Wagner, "Continuous-wave operation of GaInAs–AlGaAsSb quantum-cascade lasers," *IEEE Photon. Technol. Lett.*, vol. 17, no. 11, pp. 2283–2285, Nov. 2005.

[13] D. Indjin, P. Harrison, R. W. Kelsall, and Z. Ikonić, "Mechanisms of temperature performance degradation in terahertz quantum-cascade lasers," *Appl. Phys. Lett.*, vol. 82, no. 9, pp. 1347–1349, 2003.

[14] V. D. Jovanović, D. Indjin, N. Vukmirović, Z. Ikonić, P. Harrison, E. H. Linfield, H. Page, X. Marcadet, C. Sirtori, C. Worrall, H. E. Beere, and D. A. Ritchie, "Mechanisms of dynamic range limitations in GaAs/AlGaAs quantum cascade lasers: influence of injector doping," *Appl. Phys. Lett.*, vol. 86, no. 211117, 2005.

[15] C. A. Evans, V. D. Jovanović, D. Indjin, Z. Ikonić, and P. Harrison, "Design and simulation of InGaAs/AlAsSb quantum-cascade lasers for short wavelength emission," *Appl. Phys. Lett.*, vol. 87, no. 141109, 2005.

[16] D. G. Revin, M. J. Steer, L. R. Wilson, R. J. Airey, J. W. Cockburn, E. A. Zibnik, and R. P. Green, "InGaAs-AlAsSb quantum cascade structures emitting at 3.1 μm," *Electron. Lett.*, vol. 40, no. 14, pp. 874–875, 2004.

[17] P. Harrison, *Quantum Wells, Wires, and Dots: Theoretical and Computational Physics*, 2nd ed. Chichester, U.K.: Wiley, 2005.

[18] E. Anemogiannis, E. N. Glytsis, and T. K. Gaylord, "Determination of guided and leaky modes in lossless and lossy planar multilayer optical waveguides: Reflection pole method and wavevector density method," *J. Lightw. Technol.*, vol. 17, no. 5, pp. 929–941, May 1999.

[19] M. Sotoodeh, A. Khalid, and A. Rezazadeh, "Empirical low-field mobility model for III–V compounds applicable in device simulation codes," *J. Appl. Phys.*, vol. 87, no. 6, pp. 2890–2900, Mar. 2000.

[20] W. S. Capinski, H. J. Maris, T. Ruf, M. Cardona, K. Ploog, and D. S. Katzer, "Thermal-conductivity measurements of GaAs/AlAs superlattices using a picosecond optical pump-and-probe technique," *Phys. Rev. B*, vol. 59, pp. 8105–8113, 1999.

[21] G. Chen, "Thermal conductivity and ballistic-phonon transport in the cross-plane direction of superlattices," *Phys. Rev. B*, vol. 57, pp. 14958–14973, 1998.

[22] C. Pflügl, M. Litzenberger, W. Schrenk, D. Pogany, E. Gornik, and G. Strasser, "Inferometric study of thermal dynamics in GaAs-based quantum-cascade lasers," *Appl. Phys. Lett.*, vol. 82, no. 11, pp. 1664–1666, 2003.

[23] J. Faist, M. Beck, T. Aellen, and E. Gini, "Quantum-cascade lasers based on a bound-to-continuum transition," *Appl. Phys. Lett.*, vol. 78, no. 2, pp. 147–149, 2001.

[24] A. Mirčetić, D. Indjin, V. Milanović, P. Harrison, Z. Ikonić, R. Kelsall, M. Giehler, R. Hey, and H. Grahn, "GaAs/AlGaAs quantum cascade lasers based on double resonant electron—LO phonon transitions," in *Mat. Science. Forum*, 2005, vol. 494, pp. 25–30.

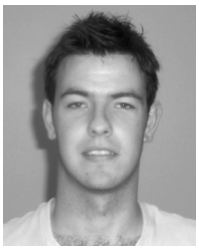
[25] J. S. Yu, A. Evans, S. Slivken, S. R. Darvish, and M. Razeghi, "Short wavelength ($\lambda \sim 4.3\ \mu$m) high-performance continuous-wave quantum-cascade lasers," *IEEE Photon. Technol. Lett.*, vol. 17, no. 6, pp. 1154–1156, Jun. 2005.

[26] F. Capasso, R. Paiella, R. Martini, R. Colombelli, C. Gmachl, T. L. Myers, M. S. Taubman, R. M. Williams, C. G. Bethea, K. Unterrainer, H. Y. Hwang, D. L. Sivco, A. Y. Cho, A. M. Sergent, H. C. Liu, and E. A. Whittaker, "Quantum-cascade lasers: Ultrahigh-speed operation, optical wireless communication, narrow linewidth, and far-infrared emission," *IEEE J. Quantum Electron.*, vol. 38, no. 6, pp. 511–532, Jun. 2002.

[27] S. Slivken, J. S. Yu, A. Evans, J. David, L. Doris, and M. Razeghi, "Ridge-width dependence on high-temperature continuous-wave quantum-cascade laser operation," *IEEE Photon. Technol. Lett.*, vol. 16, no. 3, pp. 744–746, Mar. 2004.

[28] M. Beck, J. Faist, U. Oesterle, M. Ilegems, E. Gini, and H. Melchior, "Buried heterostructure quantum cacade lasers with a large optical cavity waveguide," *IEEE Photon. Technol. Lett.*, vol. 12, no. 11, pp. 1450–1452, Nov. 2000.

[29] C. Zhu, Y. G. Zhang, A. Z. Li, and Y. L. Zheng, "Comparison of thermal characteristics of antimonide and phosphide MQW lasers," *Semicond. Sci. Technol.*, vol. 20, pp. 563–567, 2005.



Craig A. Evans was born in Wigan, U.K., in 1982. He received the M.Eng. degree in electronic and photonic communications from the University of Leeds, Leeds, U.K., in 2004. He is currently working towards the Ph.D. degree at the Institute of Microwaves and Photonics, School of Electronic and Electrical Engineering, University of Leeds, in the theory and design of quantum-cascade lasers.

Vladimir D. Jovanović was born in Belgrade, Yugoslavia, in 1978. He received the B.Sc. degree in electrical engineering from the University of Belgrade, Yugoslavia, in 2002, and the Ph.D. degree from the Institute of Microwaves and Photonics, School of Electronic and Electrical Engineering, University of Leeds, Leeds, U.K., in 2005.

His research interests are in the physical modelling, design, and optimization of quantum-well infrared photodetectors and quantum-cascade lasers in GaN- and GaAs-based materials for nearinfrared, midinfrared, and terahertz applications. He is currently a Senior Associate in the Equity volatility analytics group at Lehman Brothers, London, U.K.

Dr. Jovanović is Laureate of the prestigious IEE Leslie H Paddle Fellowship and received the SPIE Educational Awards in 2004 and also in 2005.

Dragan Indjin was born in Zemun, Yugoslavia, in 1963. He received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the University of Belgrade, Belgrade, Yugoslavia, in 1988, 1993, and 1996, respectively.

Since 1989, he has been with the Faculty of Electrical Engineering, University of Belgrade, where he holds the position of Associate Professor. In 2001, he joined the Institute of Microwaves and Photonics, School of Electronic and Electrical Engineering, University of Leeds, Leeds, U.K., where he has recently obtained the prestigious academic fellowship. His research interests include the electronic structure, optical and transport properties, optimization and design of quantum wells, superlattices, quantum-cascade lasers, and quantum-well infrared photodetectors in the near–far-infrared spectral range.

Zoran Ikonić was born in 1956 in Belgrade, Yugoslavia. He received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the University of Belgrade, Belgrade, Yugoslavia, in 1980, 1984, and 1987, respectively.

Since 1981, he has been with the Faculty of Electrical Engineering, University of Belgrade (Full Professor since 1998). In 1999, he joined the Institute of Microwaves and Photonics, University of Leeds, Leeds, U.K. His research interests include the electronic structure, optical and transport properties of semiconductor nanostructures, and devices based upon them.

Paul Harrison (SM'99) received the B.Sc. degree from the University of Hull, Hull, U.K., in 1988 and the Ph.D. degree from the University of Newcastle-upon-Tyne, Newcastle-upon-Tyne, U.K., in 1991.

He was a Postdoctoral Research Assistant at the University of Hull until 1995, when he obtained a Fellowship at the University of Leeds, Leeds, U.K. Since joining the Institute of Microwave and Photonics, University of Leeds, he has been working on ways to adapt his theoretical and computational experience in semiconductor heterostructures to terahertz sources and detectors. He currently holds a chair in Quantum Electronics and is Head of the School of Electronic and Electrical Engineering. He is author of the book *Quantum Wells, Wires and Dots, 2nd ed.* (Wiley, 2005).