Saturation and noise properties of quantum-dot optical amplifiers

Berg, Tommy Winther; Mørk, Jesper

Published in:
IEEE Journal of Quantum Electronics

Link to article, DOI:
10.1109/JQE.2004.835114

Publication date:
2004

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Abstract—Based on extensive numerical calculations, quantum-dot (QD) amplifiers are predicted to offer higher output power and lower noise figure compared to bulk as well as quantum well amplifiers. The underlying physical mechanisms are analyzed in detail, leading to the identification of a few key requirements that QD amplifiers should meet in order to achieve such superior linear characteristics. The existence of a highly inverted wetting layer or barrier region, acting as a carrier reservoir, is central to this performance enhancement. It is shown that amplified spontaneous emission acts to decrease the inversion of the wetting layer states, thus helping to quench the gain of these states, which might otherwise dominate.

Index Terms—Numerical modeling, optoelectronics, quantum dots (QDs), semiconductor optical amplifiers.

I. INTRODUCTION

Optical amplifiers play a crucial role in communication networks of the present day. The applications range from high quality linear amplifiers in long haul systems, to booster amplifiers in access and metro networks. Furthermore, research has shown the potential of applying semiconductor optical amplifiers (SOAs) as gating elements in switching nodes. Fiber amplifiers are today the standard choice for high-end amplifiers due to their high saturated output power (SOP) and lower noise compared to SOAs [1]. On the other hand, the compactness, flexibility, efficiency, and potential low price of SOAs render them interesting candidates for future communication networks if the performance issues mentioned above are improved.

In the past decade, the potential advantages in using quantum-well (QW) rather than bulk active regions in SOAs have been investigated thoroughly. It has been shown that in many cases QW SOAs exhibit higher SOP [2] and lower noise figure [3] due to the low dimensionality and smaller confinement factor of the QW material. It should be noted, however, that until recently the record saturation output power (SOP) of +17 dBm was achieved in a SOA with a 50-nm-thick bulk active region [4] and not in a QW device.

The technological realization within recent years of high-quality quantum-dot (QD) material offers many new possibilities for both semiconductor lasers [5], [6] and amplifiers [7]–[10]. In particular, QD lasers have been investigated and advantages such as low threshold current density [11], small linewidth enhancement factor [12], [13], high output power [14], [15], and good temperature stability [16] have been experimentally demonstrated. Recently, also QD SOAs have received some attention, mainly focused on their non-linear properties. Main results include the demonstration of ultrafast gain recovery [17], signal processing and pattern free amplification up to 40 Gbit/s [18] and unique four-wave mixing properties [19]. With respect to their application as linear amplifiers, QD SOAs were predicted to exhibit low noise figure and high saturation power [20], [21]. Recently, Akiyama et al. [22] demonstrated 23-dBm penalty-free output power in an InAs–InP QD SOA, thus verifying the predictions of high saturation power.

In this work, we theoretically investigate the amplification and saturation properties of QD amplifiers and compare to bulk and QW SOAs. Since the detailed properties of SOAs vary significantly depending on dimensions, structure, and material composition, it is difficult to perform a quantitative comparison on a general level. We have, therefore, chosen three specific SOA structure: QD, QW, and bulk, which may each be considered as typical representatives, as the basis for the simulations and comparisons. However, in our analysis, we will focus on the qualitative differences between the amplifier types and the physical explanation of these.

The paper is organized as follows. After introducing the amplifier model in Section II, the general amplification properties of the three different amplifier types are compared in Section III. Section IV analyzes the physical mechanisms leading to a high QD saturation power. Section V discusses the important role of amplified spontaneous emission, and finally Section VI concludes.

II. MODEL

A. Device Structure

The QD SOA considered is illustrated schematically in Fig. 1(a). It is a ridge waveguide device with an active region of length $L$, width $W$, height $H$, and volume $V_a$. This region contains a number of QD layers, $N_{QW}$, each including a wetting layer (WL), which is typical to the Stranski–Krastanov growth mode. The WL is modeled as a narrow QW of width $W$ and volume $V_W$. The QDs and WL are surrounded by a barrier material, which is assumed to separate the QD layers to the extent that these do not couple directly. Furthermore, tunneling between dots within the same layer is neglected.

Fig. 1(b) illustrates the conduction band density of states (DOS) of the device. Each dot contains a spin degenerate ground state (GS) and a fourfold degenerate (including spin) excited state (ES). Due to size and composition fluctuations...
among the dots, these states are inhomogeneously broadened with a full-width at half-maximum (FWHM) of \( \gamma_G \) and \( \gamma_E \), respectively. Above the QD states, the step-like DOS typical of a QW is visible. Due to the narrowness of the WL, typically 1–2 nm [23], only a single electronic subband is assumed to be present. The barrier material is treated as a bulk-like region, leading to a square root dependence of the DOS at high energies. Due to the vastly different volumes of the three different regions (QD, WL, and barrier), the absolute number of states in the QDs is much smaller than that of the WL, which in turn is much smaller than that of the barrier. The relative number of states in these regions has profound consequences for the saturation properties of the device [24], [25].

Amplified spontaneous emission (ASE) spectra from SOAs are typically very broad and the spectrum is discretized into a number of photon populations across the range of transition energies of the QDs, the WL, and the barrier. The central photon energy of population \( p \) is denoted by \( \tilde{\omega}_p \) and the width of the population by \( \Delta \tilde{\omega}_p \), which is taken equal to the homogeneous linewidth (FWHM), \( 2\gamma_H \). Within each interval, the spectral density is assumed constant.

### B. Carrier Dynamics

The electrons in the barrier and WL are assumed to be in mutual equilibrium at all times and are thus described by one common carrier density. Due to the larger effective mass of holes compared to electrons, and resulting smaller level spacing, holes are expected to relax faster than electrons [26] and, therefore, electrons are assumed to limit the carrier dynamics. Due to fast intradot carrier relaxation and the large energy separation between the GS and the WL bandedge, carrier capture directly into the GS is neglected. Holes are assumed to be in equilibrium at all times over the entire valence band (including WL and barrier) and only described in terms of the evolution of the effective quasi-Fermi level of the valence band, cf., Appendix I.

The inhomogeneous broadening is included by assuming a Gaussian broadening of the transition energies and splitting the ensemble of QDs into a number of populations depending on bound state energies and the homogeneous linewidth. The GS and ES energies of QD population \( n \) are denoted \( E_{Gn} \) and \( E_{En} \), as illustrated in Fig. 1(b), and similar to the photon populations we assign a width \( 2\gamma_H \) to each population. Dots within the same population are assumed identical and the carrier dynamics may thus be described by rate equations for the electron densities of the ground state \( N_{Gn} \), excited state, \( N_{En} \) for each dot population, and one for the combined wetting layer and barrier (reservoir) \( N_R \).

\[
\frac{\partial N_R}{\partial t} = R_{\text{pump}} - \sum_n \frac{V_n}{V_0} \left( R_{\text{cap}}^n - R_{\text{esc}}^n \right) - R_{\text{spont}}^R - R_{\text{stim}}^R 
\]

\[
\frac{\partial N_{Gn}}{\partial t} = R_{\text{cap}}^n - R_{\text{rel}}^n + R_{\text{esc}}^n - R_{\text{spont}}^G - R_{\text{stim}}^G 
\]

\[
\frac{\partial N_{En}}{\partial t} = R_{\text{rel}}^n - R_{\text{esc}}^n - R_{\text{spont}}^E - R_{\text{stim}}^E 
\]

\( N_R \) is normalized here with respect to the active region volume \( V_0 = \text{WHL} \), while \( N_{E}^n \) and \( N_{G}^n \) are normalized with respect to the total volume of all dots in that particular \( n \)th population \( V_0 \). In the following, each of the terms in (1)–(3) are described in detail.

The pump term in (1) is given as \( R_{\text{pump}} = I/V_0 q \), with \( I \) being the bias current and \( q \) the magnitude of the electronic charge. Current is assumed to be injected directly into the barrier region and transport phenomena, such as drift or diffusion, are not explicitly included in the model.

The spontaneous recombination term for the reservoir is \( R_{\text{spont}}^R = N_R/\tau_B \), where \( N_R(N_B) \) is the carrier density of the WL (barrier) and \( \tau_B(\tau_B) \) the corresponding spontaneous recombination time. \( N_W \) and \( N_B \) are given as

\[
N_i = \int_{-\infty}^{\infty} \rho_C(E) \left( 1 + \exp\left( \frac{E - E_{\text{RF}}}{k_B T} \right) \right)^{-1} dE, 
\]

where \( \rho_C(E) \) is the conduction band DOS, and \( k_B T \) is the thermal energy, \( E_{\text{RF}} \) is the conduction band Fermi energy of the reservoir and is related to \( N_R \) through an integral similar to (4) except that \( \rho_C(E) \) is replaced by \( \rho_W(E) + \rho_B(E) \). The spontaneous recombination times, \( \tau_W \) and \( \tau_B \), contain contributions from non-radiative, radiative, and Auger recombination and are given as

\[
\frac{1}{\tau_i} = A_i + B_i N_i + C_i N_i^2, \quad i = W, B. 
\]

\( R_{\text{cap}}^n \) is the rate of capture of electrons from the WL into the ES in dot population \( n \). The general form of this term is

\[
R_{\text{cap}}^n = f_{n}^{W} \epsilon_{E}^n \frac{1 - f_{E}^{W}}{\tau_C} 
\]

where \( f_{n}^{W} \) is the occupation probability at the bandedge of the WL, \( \epsilon_{E}^n \) is the degeneracy of the ES level including spin, \( f_{E}^{W} = N_{E}^{n}/V_0 \) is the electronic occupation probability at the ES, and \( \tau_C \) is the effective capture time. Carrier capture is mediated by phonon and Auger processes [27], [28] and the capture time is parameterized as \( 1/\tau_C = Z_G^E + Z_A^E f_{E}^{W} \), where \( Z_G^E \) and \( Z_A^E \) are coefficients governing the phonon and Auger parts, respectively.
This form of the capture term assumes that capture occurs only from the WL bandedge and, consequently, the capture time does not decrease further once the WL bandedge is fully occupied. The capture time is thus defined as the refilling time of the ES level when the WL bandedge is completely inverted and the ES is completely empty (corresponding to the rise time of the photoluminescence) [29].

The rate of carrier escape from the QD ES to the WL bandedge is described by

\[ R_{\text{esc}}^m = N_E \left( 1 - \frac{f_l^m}{\tau_{\text{esc}}^m} \right) \tag{7} \]

where \( \tau_{\text{esc}}^m \) is the characteristic escape time. Under conditions of thermal equilibrium \( R_{\text{esc}}^m = R_{\text{cap}}^m \), which leads to the requirement: \( \tau_{\text{esc}}^m = \eta_E \tau_{\text{cap}}^m \), where \( \eta_E = \exp(\Delta E_E/k_B T) \), with \( \Delta E_E = E_C^m - E_E^m \) denoting the energy difference between the ES and the WL bandedge.

Intradot relaxation is assumed to occur through the same mechanisms as the capture process and consequently the two terms appear similar. We thus have

\[ R_{\text{rel}}^m = N_E \left( 1 - \frac{f_l^m}{\tau_0} \right) \tag{8} \]

with \( f_l^m = N_D^m V_0/\varepsilon_G \) being the occupation probability in the QD GS, \( \varepsilon_G \) the degeneracy of the GS including spin, and \( \tau_0 = 1/\left( Z_R^m + Z_R^h R_l^m \right) \) the characteristic relaxation time, where \( Z_R^m \) and \( Z_R^h \) are the coefficients governing the phonon and Auger assisted intra dot relaxation time.

The intradot excitation terms are defined as

\[ R_{\text{exc}}^m = N_E \left( 1 - \frac{f_l^m}{\tau_{\text{exc}}^m} \right). \tag{9} \]

The intradot excitation time \( \tau_{\text{exc}}^m \) is related to the relaxation time as \( \tau_{\text{exc}}^m = \eta_E^m \tau_0 \) with \( \eta_E^m = \varepsilon_G/\varepsilon_E \exp(\Delta E_E/k_B T) \), and \( \Delta E_C^m = E_C^m - E_E^m \) being the energy splitting between the GS and ES of dot population \( m \).

The stimulated emission term appearing in the equation for the reservoir carrier density includes separate contributions from the WL and the barrier and takes the form

\[ R_{\text{stim}}^R = \frac{1}{WH} \sum_p \left( g_p^W + g_p^B \right) P_p \tag{10} \]

where \( g_p^W \) (\( g_p^B \)) is the modal gain from the WL (barrier) for photon population \( p \), and \( P_p \) is the corresponding photon flux (in units of photons/second).

The modal gains are calculated as

\[ g_p^R = \Gamma_i \alpha_i \rho_{ie}^i \left( f_{ie}^p + f_{he}^p - 1 \right), \quad i = W, B. \tag{11} \]

Here \( \Gamma_i \) is the confinement factor, \( \alpha_i \) is the gain coefficient, \( \rho_{ie}^i \) is the reduced density of states, and \( f_{ie}^p \) and \( f_{he}^p \) are the electron and hole occupation probabilities.

The occupation probabilities are evaluated using the Fermi energy of the reservoir and the carrier energies corresponding to the photon energy \( h \omega_p \) [30]

\[ F_{E_{\text{c}(h),-p}}^n = F_{C(V)}^n + \left( h \omega_p - E_{0}^n \right) \frac{m_i^p}{m_i^0(V)}, \quad i = W, G \tag{12} \]

with \( E_{0}^n \) (\( E_{1}^n \)) denoting the conduction (valence) bandedge energy, \( E_{0}^n \) is the bandgap, and \( m_i^p \), \( m_i^0 \), and \( m_i^0 \) are the reduced, electron, and hole masses, respectively. \( \rho_{ie}^i \) is an effective DOS, where the homogeneous broadening is included by convoluting the traditional reduced DOS with the homogeneous broadening function. In agreement with [30], we use for the reservoir states a \( \text{sech}^2 \) broadening function. This minimizes the influence of the slowly decaying tails of the often used Lorentzian lineshape, which in the present case could dominate over the QD transitions to an unrealistic degree.

The stimulated recombination terms are slightly more complex for the QD states since they involve the influence of both homogeneous and inhomogeneous broadening. They are given as \( R_{\text{stim}}^n = (L/V_n) \sum_p g_p^m P_p \), with \( I = E, G \), where \( g_p^m \) (\( g_{E}(E_{0}^n) \)) is the modal gain from the GS (ES) transition of dot population \( n \) for the photon population \( p \). Following Blood [31], but modifying the approach to include the inhomogeneous broadening, we describe the modal gain as

\[ g_{E}^m = \frac{D_D N_{\text{kv}} \sigma_D \varepsilon_{\text{inc}}} {\omega_{\text{inc}}} Q_{E}^n \left( f_l^m + f_h^m - 1 \right), \quad i = E, G \tag{13} \]

Here, \( D_D \) is the two-dimensional (2-D) density of dots in a single layer, \( N_{\text{kv}} \) is the number of QD layers, \( \sigma_D \) is the gain coefficient of a single QD state (assumed the same for all dots), \( \rho_{E}(E_{0}^n) \) is the hole occupation probability evaluated for the transition energy corresponding to the electronic GS (ES) transition with energy \( h \omega_p \) (see Appendix I), and \( \omega_{\text{inc}} \) is the effective width of the optical mode. \( Q_{E}^n \) denotes the coupling between dot population \( n \) and photon population \( p \).

\[ Q_{E}^n = \int \int \int \int \int \frac{B(\gamma, \omega, \omega_p, \varepsilon)} {2\gamma H} dE d(\omega_p d) \tag{14} \]

where \( B(\gamma, \omega, \omega_p, \varepsilon) \) is the central GS (ES) transition energy of population \( n \). \( \varepsilon_I(E_{\varepsilon}) \) is the normalized, Gaussian inhomogeneous distribution function of the QD GS (ES) transition with FWHM \( \gamma_{\varepsilon}(\varepsilon_I) \) and centered at transition energy \( E_{\varepsilon,I}(E_{\varepsilon,I}) \), \( B(\gamma, h \omega_p) \) is the normalized, Lorentzian homogeneous broadening function used for the dot transitions.

C. Propagation Effects

Propagation effects are accounted for by discretizing the SOA into a number of sections. Using the Einstein relation between gain and spontaneous emission into the guided waveguide mode the generation and propagation of ASE is included [32]. The
propagation equation for the total optical intensity (signal and ASE) over a single section is thus

\[ P_{p,s}^\pm = P_{p,s}^{\pm+1} G_{p,s} + L_s \frac{G_{p,s} - 1}{\ln G_{p,s}} \Delta \omega \left( \sum_j g_{p,s,j}^t n_w^j \right) \]  

(15)

where the first term describes the gain of the optical intensity entering from the previous section and the second term describes the amount of spontaneously emitted photons within the section. In (15), \( P_{p,s}^\pm \) is the photon flux within the frequency interval \( \Delta \omega \) traveling in the forward (+) or backward (−) direction at the front of section \( s, G_{p,s} = \exp(\sum_j g_{p,s,j}^t - \alpha_{int})L_s \) is the total gain for the photon population with \( \alpha_{int} \) being the internal waveguide loss, and \( L_s \) the section length. The summation \( j \) extends over all contributions to the modal gain, i.e., the QD GS and ES from all dot populations, the WL and the barrier. Note, that in principle all carrier densities, occupation probabilities, and gain factors described previously should be specified also by a section index, \( s \), but this has been omitted for simplicity. Finally, \( n_w^j \) is the inversion factor

\[ n_w^j = \frac{f_l^j f_h^j}{f_l^j + f_h^j - 1} \]  

(16)

where \( f_l^j(f_h^j) \) is the electron (hole) occupation factor for the various levels, i.e., \( f_l = f_{l,G(E)}(f_h = f_{h,G(E)}) \) for the QD states and \( f_l = f_{l,p}^W(B)(f_h = f_{h,p}^W(B) \) for the reservoir states. The photon flux used in the carrier equations, \( P_p \), equals the sum of the flux from the front and back edge of the section, i.e., \( P_p = P_{p,s-1}^+ + P_{p,s}^- \).

The steady state ASE distribution is found by iteration. First, the forward ASE intensity is propagated through the device, while keeping the backward ASE to each section constant. The backward ASE is then propagated back through the amplifier with the forward ASE intensity being held constant. This process is repeated until a steady state is reached. When a signal is injected into the device, the iteration process is repeated for all input powers in order to allow the ASE distribution to adjust to the new steady state.

D. Bulk and QW SOAs

Bulk and QW devices are modeled using the same approach, but omitting the QD terms and using appropriate parameter values. The bulk SOA model thus corresponds to the model described above for the barrier material and a QW SOA is modeled as the WL plus barrier in the QD case. This means that the effect of a finite capture time of carriers from the barrier region into the QWs is neglected for the QW SOA.

E. Validation

Throughout the paper one specific device of each type (QD, QW, and bulk) will be used as starting point for the discussion of the qualitative differences. These devices are referred to as the “reference device” and the parameter values for these are listed in Appendix II together with a short description of each device.

The numerical model has been validated by comparison to both theory and experiment. The bulk part of the model has been compared to other bulk SOA models [33], [34] and agrees well with these in all aspects. Furthermore, the QD model has been tested against pump-probe experiments [17] and is able to reproduce these. A final indication of the validity of the models is the good qualitative agreement between the numerical results reported here and the experimental results published by Fujitsu [22], regarding SOP, device gain, and noise figure.

III. AMPLIFICATION PROPERTIES

The modal gain spectrum for the reference QD SOA including waveguide loss is shown in Fig. 2 for three different currents. Maximum modal gain values of 12.5 cm\(^{-1}\) at the centre of the GS transition and of 22 cm\(^{-1}\) at the ES peak are obtained. In comparison, the maximum gain of the WL reaches a much higher modal gain of 56 cm\(^{-1}\) under full inversion, and the barrier an even higher value.

Two important parameters of an amplifier are the small signal gain, i.e., the gain in the absence of a saturating signal, and the SOP. The definitions of these terms are illustrated in Fig. 3, where three specific examples of saturation curves, one for each SOA type, are shown. For the bulk case, which we will use as an example, the gain approaches its unsaturated, small-signal, value \( G_0 \), for small optical powers; in the present case \( G_0 = 25 \) dBm. As the input power (and consequently
also the output power) is increased, the gain starts to saturate and eventually the amplifier is forced toward transparency, where the net gain of the device is completely determined by the total waveguide loss, \( A = \exp(-\alpha L) \). The traditional definition of the SOP is the output power at which the gain has decreased by 3 dB, i.e., at \( G_0/2 \). However, in some of the cases investigated here (see for example the saturation curve for the QD device in Fig. 3), the small signal gain can be very small and in some cases even smaller than 3 dB, which means that the standard definition cannot be used. We will, therefore, use a more general definition, where we define the saturated gain as \( G_{sat} = (G_0 + A)/2 \), which corresponds to the mid-point between the unsaturated and the completely saturated gain. At gain values significantly higher than 3 dB, which is usually the case, this definition agrees well with the traditional one. The advantage of the more general definition is that it can be used for all devices and that the saturation power is well defined as soon as the device operates above transparency.

Comparing the three devices in Fig. 3, we see that the QW SOA gives only a slightly smaller gain than the bulk component even though biased with a significantly lower current density. This is due to the combination of smaller carrier masses and the step-like DOS for the QW SOA, which leads to a significantly higher inversion at lower currents. The QD SOA in turn has even fewer active states than the QW device and is, therefore, even easier to invert. However, since the maximum modal gain of the QD is limited to 12.5 cm\(^{-1}\) the device gain of such a short device, even under complete inversion is very small (2.5 dB in this case). The saturation power, however, is much larger for the QD device (20 dBm) despite the lower bias current density used. In comparison, the SOP for the QW is 12 and 10.5 dBm for the bulk device.

To make a more general comparison between the devices, Fig. 4 shows the SOP versus current for the three different SOAs. Looking first at the QD curve (solid line) the SOP is seen to increase with current from a value close to transparency of 4 dBm to a maximum value of 25 dBm for current densities above 10 kA/cm\(^2\). The increase with current is caused by the increased degree of bandfilling of the ES, WL, and barriers, and these carriers can be consumed by the energetically lower GS transition without significant reduction of the gain. Eventually, the inversion becomes sufficiently high that the limiting mechanism for the SOP becomes the transport of carriers into the active QD states and not the availability of carriers in the device. The maximum SOP is in this case limited by the minimum capture time of electrons from the WL to the ES, 2 ps in the present case. In this regime the quasi-equilibrium between the QD and the reservoir electron states breaks down and strong spectral hole burning (or barrier-dot hole burning) occurs.

It should be noted, that in the regime of operation, where the SOP is completely limited by the capture time, the assumption of equilibrium among all hole states is expected to break down. This means that in reality the gain saturation would occur through a reduction of both the electron and hole inversion, instead of only the electron inversion as in the present model. Most likely the degree of nonequilibrium introduced to the hole distribution will be smaller than for the electrons, due to the expected faster relaxation process of the holes. The maximum achievable SOP seen in Fig. 4 is, therefore, expected to be slightly overestimated. If we use instead the assumption that the dynamics of the holes follow the electrons, which is a common assumption for QD devices [35], [36], we would find a reduction by approximately 3 dB of the maximum SOP, since in this case the reduction of the inversion is shared evenly between the conduction and valence bands. This method would probably underestimate the SOP, and we expect the realistic value to lie somewhere in between these two bordering cases.

An important point to note about the capture limited maximum SOP is, that it represents a limit for both the average as well as the peak signal power. If a pulsed signal injected into the QD SOA has an average power below this limit, e.g., 10 dBm, but a peak power above the maximum SOP, the amplifier will not be able to deliver the full device gain to the signal. The reason is that the high peak power leads to a momentary depletion of the gain, since the capture process is not able to keep up with the rate of stimulated emission, and a nonequilibrium situation is created, corresponding to dot-barrier hole-burning. However, the presence of gain saturation in this regime of operation will not lead to pattern effects, contrary to the case of bulk and QW SOAs, where operation close to the saturation point commonly leads to strong pattern effects for pulsed signals [37]. The reason for the difference is that in this regime the gain recovers on the timescale of the capture time, i.e., within picoseconds, whereas the gain in bulk and QW SOAs recovers on the timescale of the effective carrier lifetime, i.e., hundreds of picoseconds. As a result, QD SOAs can be operated well into the saturation regime for high currents, where the capture time limits the output power, with only little distortion of the signal, as long as the data rates are below the inverse of the capture time. This feature was recently demonstrated experimentally [22].

In comparison with the QD device, the QW SOA (dashed line in Fig. 4) starts out at a smaller value of 0 dBm for the SOP around transparency and increases steadily with current to a value slightly below 16 dBm for a current density of 20 kA/cm\(^2\). The increase with current is in this case caused by a combination of increased inversion of the QW as well as the barrier states and a decrease of the effective carrier lifetime in the device. The decrease of the carrier lifetime is caused by the polynomial dependence of the carrier lifetime on carrier density, describing the
influence of radiative and Auger-assisted recombination. The small value of the SOP close to transparency is, in contrast, explained by a long carrier lifetime and a high differential gain at this point. The same mechanisms cause a slight increase of the SOP of the bulk device from a value of 5 dBm at transparency to 12 dBm at 20 kA/cm². When compared at the same current density, it is clear that the QD SOA outperforms both the bulk and the QW SOAs regarding SOP.

A most important parameter for amplifiers is of course the magnitude of the small-signal gain. Since the modal gain of QD SOAs usually is significantly smaller than for QW and bulk devices, it is necessary to use significantly longer devices in order to get a reasonable device gain. The dependence of the small signal peak gain on device length is shown in Fig. 5(a). For the bulk and QW devices, the gain is seen to increase linearly with length for up to 1 mm in the bulk case and 2 mm in the QW case, but then saturates at a value of 40 dB for longer devices. Note that since the gain peak shifts with length, different signal wavelengths are used for the different lengths. The smaller modal gain of the QD device is observed to lead to a slow increase of the device gain with length. For the QD SOA the optimum signal wavelength is nearly independent of device length, due to the nature of bandfiling in the QDs. For devices longer than 6 mm, the gain increases sublinearly with length. For devices longer than the 14 mm shown here, the gain does eventually saturate completely as for the bulk and QW case.

In the absence of an input signal, gain saturation is caused by the presence of strong ASE intensity in long devices [38]. The optical power of the ASE at any given wavelength is proportional to the gain at that wavelength [see (15)]. Consequently, the ASE power increases with length until the point where it becomes comparable to the saturation power of the device and the gain, as a result, starts to saturate. The maximum gain obtainable thus scales linearly with the saturation power of the device. Note, however, that other parameters, such as the waveguide dimensions for example, also influence the maximum gain value.

The saturated gains of 40 dB for the bulk and QW amplifiers shown in Fig. 5(a) are approximately 4 dB larger than the highest values for the chip gain reported experimentally [39], [40]. The explanation for this difference is probably a combination of experimental and modeling issues. First of all it is not trivial to realize a long amplifier (in this case more than 2 mm) with AR coatings on both facets with reflections below $10^{-1}$, which is required in order to achieve such high gains. Also these high gain values are obtained in the regime where the amplifier is strongly saturated by ASE, which leads to a high noise figure as can be seen in Fig. 5(b). The maximum gain regime may thus not be interesting from an application point of view. In contrast the numerical model ignores reflections from the facets and any length desired can be considered. Another point is that the model only takes ASE in the guided mode into account and it might be expected that ASE in nonguided modes could also deplete the gain to a certain extent resulting in a lower saturated gain. However, since these points apply to all three types of amplifiers in similar ways, it is important to note that the high saturation power of QD SOAs should lead to higher saturated gains of QD SOAs compared to bulk and QW amplifiers at the same current density.

The final figure of merit of amplifiers that we consider is the noise figure, which we define as follows:

$$F(\eta_{\text{opt}}) = \frac{1}{G_p} + 2 \sum_s \frac{G_{p,s}}{\prod_{i=1}^{n} G_{p,i} - \alpha + \sum_j g_{s,j} n_{s,j}}$$

where $G_p = \prod_s G_{p,s}$ is the total device gain at the photon energy $\eta_{\text{opt}}$. In the last two summations, $j$ extends over all carrier populations, including WL and barrier. The first term in (17) describes shot noise and the second describes signal-spontaneous beat noise, which dominates when the device gain is above a few decibels. We have neglected spontaneous-spontaneous beat noise, which is a good approximation for narrow filter bandwidths [41]. Note that the definition does not include coupling loss, implying that any input coupling loss should be added to the actual noise figure of the device. The noise figures are in all cases calculated in the absence of a signal, i.e., without saturation due to the signal.

Fig. 5(b) shows the calculated noise figures corresponding to the gain curves in Fig. 5(a). For the QD device the noise figure is seen to increase from below 3 dB for devices shorter than 2 mm and then stabilizes slightly above 3 dB with a slight increase for the longest devices shown. Notice that the often-considered lower limit of 3 dB for the noise figure only applies in the high-gain limit [42]. For the short QD SOAs the gain is very low and the noise figure is consequently below 3 dB. The reason for the very low noise figure even for devices longer than 2 mm is that the active electronic QD states are nearly completely inverted; the inversion factor $n_{\text{opt}}$ then approaches unity and the noise figure becomes close to the 3-dB limit. This property of QD devices has previously been predicted by Komori et al. [20]. For long devices, the ASE starts to saturate the device, i.e., the inversion decreases, and the noise figure starts to increase. Similar behavior is seen for the bulk and QW devices, except that the increase in noise figure occurs at much shorter device lengths due to the higher gain of these devices and consequently faster
ASE saturation. The QW device is seen to have a noise figure close to 3 dB for devices shorter than 1 mm, indicating that the active electronic states of this device are also nearly completely inverted. In contrast, the bulk device remains far from full inversion at the signal wavelengths, despite the higher bias current density, resulting in higher noise figures for all device lengths. It should be noted that in order to perform a fair comparison of the amplifier types, the noise figure should be compared at the same device gain. If we thus compare for a device gain of 20 dB corresponding to $L = 0.4$ mm for the QW device and $L = 4$ mm for the QD, we see that the QW amplifier actually has a slightly smaller noise figure of 3.4 compared to 3.7 for the QD SOA. Since the electronic states of both devices are nearly completely inverted at these current densities, the main reason for this difference is the ratio of the modal gain to the waveguide loss, which enters the expression for the noise figure [see (17)]. This ratio is highest for the QW device for the device parameter values used here and, as a result, the noise figure under full inversion is slightly lower compared with the QD SOA.

### IV. MECHANISMS OF HIGH-SATURATION POWER

To understand the reason for the higher SOP of the QD SOA compared with the QW and bulk devices, we take advantage of approximate analytical results for a linear amplifier [43], [44]. In this case, the gain is assumed to have a linear dependence on the carrier density, which leads to an expression for the internal or local saturation power: $P_{\text{sat}} = \eta \omega \Gamma / \tau_{\text{eff}}$, where $\Gamma$ is the effective cross section, $\eta$ is the differential modal gain, and $\tau_{\text{eff}}$ is the effective carrier lifetime. $P_{\text{sat}}$ corresponds to the power, where the modal gain is reduced by a factor of two compared with the unsaturated value. Despite the fact that, under high inversion, the gain is not well described as linearly dependent on carrier density, it is fundamentally still the same factors, which determine the saturation power in the QD case, albeit in a more complex and nonlinear fashion. The differential gain and the effective carrier lifetime thus still play crucial roles and we will in the following try to illustrate the importance of these parameters.

Fig. 6 shows the normalized modal gain averaged over the length of the device versus current for the three different amplifier types. The modal gain is normalized with respect to the maximum modal gain and thus corresponds to an average filling factor at the signal wavelength. For the standard values given in section two, the QD SOA (solid line) is seen to approach full inversion more rapidly and saturate more sharply with current than the QW device, which in turn saturates more sharply than the bulk device. The origin of this behavior is partly the smaller absolute number of states of lower dimensional systems and the consequent reduction in current required to reach full inversion and partly the energy separation between the QD states and the continuum of states in the WL.

The slopes of the curves shown in Fig. 6 are proportional to the differential gain values. It is clear that for low currents, the differential gain of QD and QW SOAs is higher than for a bulk device, but for high currents, where the QD gain is saturated, the differential gain becomes much smaller. As the inversion of the dot states starts to saturate, most of the carriers further added to the system will remain in the ES, the WL and the barrier. When the optical signal starts to deplete carriers from the system, the inversion of these energetically higher states will be reduced before the inversion and thus gain of the GS starts reducing, i.e., the higher states act as a reservoir of carriers for the QD GS. Therefore, a strong saturation of the QD GS gain also indicates an efficient filling of the reservoir and, hence, a high SOP.

Due to the large confinement energies, the conduction band QD GSs are well separated from the reservoir energetically and the system can in this regard be considered as a three-level system. In comparison, the carrier transitions in QW and bulk SOAs are located within their own carrier reservoir and these devices can, therefore, be considered as two-level systems in comparison. In this respect, the QD SOA can be considered similar to the erbium doped fiber amplifier, which is also a three-level system. Indeed the properties of QD SOAs, in some respects, resemble more those of EDFAs than traditional bulk SOAs. In particular a high degree of inversion, low noise figure, and high SOP are characteristics seen for EDFAs.

The dash–dotted line in Fig. 6 illustrates the result when a device with larger carrier masses of the QD and WL is considered. The masses used in this case are $m_e^W = 0.07 m_0$ and $m_h^W = 0.3845 m_0$. In particular, the large hole mass means that there are significantly more states to populate in the WL before the QD states are completely inverted and as a result the rate of filling with current is reduced. For larger carrier masses and smaller energy separation between the active QD states and the carrier reservoir, the QD amplifier behaves more like a QW device operated close to the bandedge, where the modal gain is small. The difficulty in obtaining full inversion of the active states at room temperature has been observed experimentally [24] and was attributed to a strong coupling between the QD states and the WL with a much higher DOS. It is clear that small carrier masses are as essential for QD SOAs as for bulk and QW SOAs. In the case illustrated in Fig. 6, the smaller degree of gain saturation results in an increase of the noise figure for the current densities shown. It does, however, not influence the maximum achievable SOP of the device, as illustrated later.

The length dependence of the SOP is depicted in Fig. 7 for the three devices. The SOP of the QW and bulk devices is seen to increase gradually with length, whereas the SOP of the QD
SOA decreases slightly with length in the 10-kA/cm² case and increases in the 2-kA/cm² case. For the bulk and QW devices the main mechanism behind the increase is that the ASE intensity scales with length, and consequently the stimulated recombination time decreases and the SOP increases. For the QD device, the effective carrier lifetime also decreases with length giving rise to the slight increase in the 2-kA/cm² case. For the higher current of 10-kA/cm², however, the SOP is not determined by the carrier lifetime, but completely limited by the capture time into the dot. As a result, the internal saturation power, $P_{\text{sat}}$, cannot be increased any further. The slight decrease of the SOP with length observed in Fig. 7 is caused by the increased influence of the waveguide losses in long devices when the ASE reduces the modal gain toward the ends of the device. A single section analysis of an amplifier yields the approximate ratio between the SOP, $P_{\text{sat}}^\text{eff}$, and the internal saturation power, $P_{\text{sat}}$ [45]

$$P_{\text{sat}}^\text{eff} = \frac{P_{\text{sat}} \ln 2}{1 + \alpha_{\text{int}}/g}$$  \hspace{1cm} (18)

where $g$ is the modal gain. For a device with $g \gg \alpha_{\text{int}}$, the denominator is very close to unity and the ratio becomes constant. For the QD device this is, however not the case. As the device length is increased, the device gain saturates gradually (see Fig. 5) and the average modal gain of the device decreases. Therefore, the ratio $\alpha_{\text{int}}/g$ increases and $P_{\text{sat}}^\text{eff}$ is reduced relative to $P_{\text{sat}}$. Since $P_{\text{sat}}$ is limited by capture and cannot increase any further in response to the reduced carrier lifetime, the result is a slight reduction with length of $P_{\text{sat}}^\text{eff}$. This effect also influences the bulk and QW devices for lengths longer than 1 mm, where the device gain is constant and the average modal gain thus reduces linearly with device length, but for these devices the decrease of the effective carrier lifetime dominates over the effect of decreasing average modal gain.

For QD devices longer than those shown here, the strong ASE intensity starts to deplete the inversion, i.e., empty the carrier reservoir, and the differential gain therefore increases and the SOP decreases. However, this happens in the regime where the device gain saturates and the noise figure sharply increases [see Fig. 5(b)] and this range of device lengths is thus not very interesting from the point of view of linear amplification.

In order to illustrate the influence of some of the key parameters on the QD performance, Fig. 8 shows the SOP versus current for a variety of devices deviating slightly from the reference QD device considered until now. The first example corresponds to the case shown in Fig. 6, where the carrier masses are increased substantially beyond the strained masses used for the reference device. In Fig. 6 this was seen to give rise to a decrease of the differential gain at low current and an increase at high currents. In agreement with the change of the differential gain, the SOP is in Fig. 8 seen to be slightly higher for current densities below 3 kA/cm² and lower for current densities above. It is important to note, however, that the maximum achievable SOP is still the same, although a larger current density is required to reach that point (note the logarithmic current scale). Other drawbacks of the device with larger carrier masses are a smaller device gain, for the same device length, and a higher noise figure, resulting from the smaller degree of inversion of the active states.

The second example (dotted line) shows the effect of lowering the capture time. In this case, we use Auger and phonon assisted capture times, which are ten times higher than for the reference device, giving rise to a minimum capture time of 20 ps. Since the ultimate limitation to the SOP of the reference device is determined by the capture time, the further reduction of this characteristic time has a severe effect on the maximum achievable SOP. There is a one-to-one correspondence between the rate of capture into the dot and the rate of stimulated emission it can support before the gain is saturated. Consequently, the maximum achievable SOP decreases inversely with the capture time. Another important factor is the gain from a single dot, i.e., the oscillator strength. If the oscillator strength is small, a higher optical power is required before the rate of stimulated emission becomes comparable to the capture rate. It was shown that the presence of large strain fields in QDs may significantly reduce the overlap between electron and hole wavefunctions [46]. This reduction of modal gain would require a larger number of QDs in the device in order to achieve a reasonable modal gain and, therefore, a higher current density would be required to reach the same level of inversion. On the other hand, it would also lead to a larger maximum SOP. It is central to the performance of QD
devices that the capture time into the QD is fast and, therefore, that phonon bottleneck effects are minimized.

The intradot relaxation time plays an equally important role, which is also illustrated in Fig. 8 (triangles). In this case the intradot relaxation time is set equal to the capture time, i.e., a minimum relaxation time of 2 ps under full inversion. The effect is seen to be a reduction by 4 dBm of the maximum SOP compared with the reference QD SOA (solid line). The reason behind this is again a reduction of the transport time into the active states. In the case of a large rate of intradot relaxation compared with the capture rate, a large fraction of the carriers in the ES of the QD can be consumed before the gain saturates. The ES thus acts as small reservoir for the GS transition. If, on the other hand, the rate of intradot relaxation is small compared with the rate of capture, the availability of the carriers in the ES is reduced, corresponding to a reduction of the size of the available carrier reservoir. If the relaxation time is increased even further, it becomes the limiting factor for the SOP instead of the capture time and as a result the maximum SOP will scale inversely with the relaxation time.

The final example (marked with solid squares) in Fig. 8 illustrates the effect of doubling the number of QD layers from 3 to 6. This effectively doubles the modal gain, but at the same time a higher bias current density is required in order to reach the same level of inversion, hence, the same SOP. In the figure this is seen as a 3-dBm increase of the SOP compared with the reference device at a fixed current density. However, since the modal gain is larger a shorter device can be used in order to reach the same device gain and the total bias current needed to obtain the same SOP remains unchanged.

There are other parameters influencing the performance of QD SOAs, but not as critically as the ones discussed previously. The waveguide loss degrades both the noise figure and the device gain and should be minimized. However, as long as the modal gain is significantly larger than the waveguide losses, the effect of a slight increase of the losses should be marginal to the overall performance. Waveguide losses, which are comparable with the modal gain of a device, will on the hand lead to a very severe degradation of the noise figure.

Another factor is the confinement energy of the bound states in the QDs. Again this should be as large as possible (as long as the capture and relaxation times are not reduced), but as long as it is significantly larger than the thermal energy there is not much to be gained by increasing it further. The effect of small confinement energy resembles the effect of having large carrier masses, illustrated in Figs. 6 and 8, i.e., the device becomes more bulk-like and more difficult to invert fully.

V. THE ROLE OF ASE

As previously described, the build-up of a large ASE intensity has important consequences for the behavior of long amplifiers. For such long amplifiers, it is thus the ASE that saturates the device gain and causes the noise figure to increase sharply. It is also the ASE, which causes an increase of the SOP with length (see Fig. 7).

It might be anticipated that for a long QD device under high inversion, the device gain of the barrier and WL would become very high, requiring an unrealistically good antireflection in order to avoid lasing from these states. Fig. 9 shows the maximum gain of the barrier (dotted line), the WL (dashed line), the ES (dash–dotted line) and the GS (solid line) versus device length for a bias current of 4 kA/cm$^2$. As expected, the higher modal gain of the barrier is causes the gain to increase faster with device length than for any of the other carrier populations. The gain of the barrier is thus dominant for devices up to 2 mm, but then starts to decrease. This picture repeats for the gain of the WL and ES, which dominate for devices between 2–2 mm and finally the GS gain becomes dominant for devices longer than 12 mm. It is important to note that neither the barrier nor the WL reaches a device gain, which is higher than the maximum (saturated) GS gain. As a result, the requirements to the anti-reflection coating of a long QD device would not be any stricter than for a high gain bulk or QW device. In the present case an AR-coating with a reflection coefficient of $10^{-5}$, which is comparable to the best coatings realized today [47], would thus be sufficient for any device length.

The decrease of the gain of the higher states is again caused by the presence of ASE. As the device length increases, more ASE is generated and the inversion is depleted. It is thus ASE generated at lower transition energies that eventually reduces the gain of the energetically higher levels. In this respect, the energetically lower states always win the competition for carriers over the higher states.

Fig. 10(a) shows the ASE output spectral density for different device lengths, corresponding to Fig. 9. The shift of the peak of the ASE spectrum, toward the low energy side (longer wavelength) with length is clearly visible. The total ASE emitted (the integral over the whole spectrum) increases with length, reflecting the fact that for constant current density more carriers are injected in total for a longer device and, as a result, more photons can be emitted. The shift of the peak of the ASE spectrum with increasing length is a feature of QD SOAs, which should be feasible to investigate experimentally. Its verification would indicate the saturation and reduction of the gain of the barrier and WL states with increasing length.

Fig. 9. Maximum gain of the barrier, WL, ES, and GS versus device length. The bias current density is 4 kA/cm$^2$. 
In Fig. 10(b) the wavelength dependencies of the noise figure, corresponding to the ASE spectra shown in Fig. 10(a), are shown. The smaller degree of inversion for the upper states is clearly seen to give rise to very high noise figures at the short wavelength side. In general, a longer device exhibits higher noise figures at all wavelengths for the same current density due to the reduction of the inversion caused by the increase of the ASE intensity. The development of the noise figure at the centre of the GS transition corresponds to the case shown in Fig. 5(b).

Since the small modal gain and high inversion of the active states are essential to achieve a high SOP of the QD SOAs described here, it could be speculated to instead simply use a bulk SOA operated close to the band edge. Similar to the QD device, this would result in a small modal gain and strong inversion of the active states, which would lead to a high SOP for a short device where the ASE saturation is limited. However, similar to a QD SOA, such a device would need to be long in order to obtain a reasonable device gain and in this case the strong ASE quickly reduces the inversion, leading to a reduction of the SOP. For a bulk device operated below the gain peak there will always be a region with significantly higher modal gain energetically close to the operating point, which will generate a large amount of ASE that depletes the inversion with length. It is thus the large confinement energy of the QD states, i.e., the energy separation from the reservoir, which is key in obtaining a higher SOP for QD SOAs than bulk and QW SOAs.

The normalized electron distribution along the length of the amplifier is shown in Fig. 11 for the same four device lengths as in Fig. 10. The length coordinates are normalized with respect to the total device length and only the front half of the device is shown due to symmetry. The carrier distribution of the reservoir (combined WL and barrier) is shown at the top, the ES population of the central dot population (located at the ES peak) in the middle, and the corresponding GS population at the bottom. The reservoir carrier densities are normalized with respect to the value at centre of the 1-mm device and the ES and GS carrier densities are normalized with respect to their respective maximum carrier densities, i.e., $\varepsilon_E/V_0$ and $\varepsilon_G/V_0$. In all cases, the carrier population is seen to be highest at the middle of the devices and lower toward the edges. The reason for this is that both the forward and backward ASE grows close to exponentially toward the edges and as a result the sum of these two components have a minimum at the center. As a consequence, the carrier populations, which are depleted by ASE, have a maximum at this point. For the reservoir population, the carrier depletion toward the edges is seen to be very substantial for all devices except the 1-mm-long device. The strong depletion reflects both the high generation rate of ASE in the reservoir itself, which saturates the gain, but also reflects the diminishing inversion with length caused by the ASE generated at the energetically lower states. Eventually, the reservoir is forced below transparency for all wavelengths due to the consumption of carriers in the QD states. In comparison, the depletion with length is seen to be much less severe for the ES and GS population, which reflects the slower gain saturation with length of these states.

In the presence of an optical signal the ASE and carrier distributions are redistributed similarly to the case of bulk and QW SOAs [34], [48]. Thus the consumption of carriers through stimulated emission with the signal reduces the population inversion toward the rear of the device, which results in a smaller amount of backward traveling ASE, especially at the barrier and WL transitions. The change in ASE distribution causes a redistribution of the carriers such that the point of maximum inversion shifts from the center of the device toward the front end. At high signal powers, i.e., comparable to the saturation input power, the carrier redistribution is effective at all transitions including the QD GS. The main effect of the redistribution is to convert backward-traveling ASE into signal power.

The noise figure is most sensitive to the inversion of the front section of the device as evident from (17). The carrier depletion toward the edges visible in Fig. 11 thus influences the noise
figure of the device strongly and if the depletion toward the front edge could be reduced this would improve the noise figure considerably for long devices.

VI. CONCLUSION

We have theoretically investigated the properties of linear QD amplifiers and compared to those of bulk and QW SOAs. We find that QD SOAs hold promise for high saturation power, high device gain and low noise figure. The key requirements for realizing these properties are: large confinement energies (WL to QD GS separation) of both the electron and hole states, small effective masses of especially the valence band, effective and rapid carrier capture and intradot relaxation, and low waveguide losses compared with the modal gain. It is, furthermore, found that the small modal gain is in itself not a problem (as long as it is significantly larger than the waveguide loss) and that a small oscillator strength actually allows for a higher SOP. We find that when the QD states are well separated energetically from the WL and barrier this allows for the WL and barrier to act effectively as a carrier reservoir for the QD states, which means that the SOP can be increased significantly over that of bulk and QW devices. Another key finding of this paper is that the device gain from neither the barrier nor the WL are expected to exceed that of normal high gain devices and the demands to anti reflection coating of a long QD device are, therefore, not higher than for other SOAs.

APPENDIX I

A. Hole Occupation Probabilities

The calculation of the QD hole occupation probabilities \( f_{h,G} \) and \( f_{h,E} \) entering the QD gain terms, requires in principle three steps:

1. an estimation of the valence band DOS;
2. a calculation of the Fermi energy of the total valence band;
3. a determination of a valence band energy at which to evaluate the Fermi function.

An accurate determination of the QD contribution to the valence band DOS would require a detailed numerical calculation of the eigenstates, which in turn requires knowledge of the precise shape, material composition and strain of the dots. Since such calculations are very complex and time consuming and dependent on a number of parameter values that are hardly known, an alternative approach inspired by the standard theory of bulk devices is chosen. The approach is illustrated schematically in Fig. 12. First the bound conduction band states of the QD, which we assume known at this point, are fitted with a parabolic band (similar to a bulk DOS) with the curvature determined by the effective electron mass. Two equations determining the effective conduction band edge \( E^b_\text{C} \), and the effective single dot volume \( V^b_0 \), are obtained by integrating over the bulk-like DOS from \( E^b_0 \) to the bound energy of either the GS or ES, multiplying by the effective volume of a single dot \( V^b_0 \), and equating it with either the number of states including spin of the GS \( \varepsilon_G \), or that of the combined GS and ES \( \varepsilon_G + \varepsilon_E \).

\[
\varepsilon_G = \frac{V^b_0}{2\pi^2} \left( \frac{2m^b_D}{\hbar^2} (E^b_G - E^b_0) \right)^{3/2}
\]

Fig. 12. Schematic illustration of the scheme used to evaluate the hole occupation probabilities for the QD transitions.

and

\[
\varepsilon_G + \varepsilon_E = \frac{V^b_0}{2\pi^2} \left( \frac{2m^b_D}{\hbar^2} (E^b_G - E^b_0) \right)^{3/2}.
\]

From these we find

\[
E^b_0 = \frac{E^b_C}{(1 + \varepsilon_E/\varepsilon_G)^{2/3}} - E^b_0
\]

and

\[
V^b_0 = \frac{3\pi^2}{2\varepsilon_G} (2m^b_D (E^b_E - E^b_0)^{3/2}).
\]

If we now assume \((E^b_W - E^b_C)/(E^b_W - E^b_V) = \Delta/(1 - \Delta)\), where \(\Delta\) is the relative valence bandgap difference between the two materials, originating from the conduction band, we can determine the effective valence band edge \( E^v_0 \). The contribution to the total effective DOS of the valence band from all dots in the device \( \rho^V_\text{D}(E) \), can now be written as

\[
\rho^V_\text{D}(E) = \sum_n \theta(E - E^v_n) \frac{K^v_0}{V^V_\text{act}} \frac{V^v_n}{2\pi^2} \left( \frac{2m^v_D}{\hbar^2} \right)^{3/2} \sqrt{E - E^v_n}
\]

with \(\theta(E - E^v_n)\) being the Heaviside step function and \(K^v_0\) the total number of dots in each dot population.

In order to determine the Fermi energy at any given time, charge neutrality for the device as a whole, i.e., not for each separate QD, is assumed. The total number of holes thus equals the number of electrons, which is known from the rate equations (1)–(3). From this carrier number, or carrier density if normalized with respect to volume, we find the Fermi energy from an equation similar to (4).

Finally, the energy at which to evaluate the Fermi function needs to be identified. For this purpose, the effective parabolic band approximation found before is used again and ‘k-conservation’ in the transition is assumed (see Fig. 12). In this way we get

\[
E^v_{h,i} = E^v_i + (E^v_i - E^v_0) \frac{m^D_{i}}{m^d}, \quad i = G, E.
\]

The result of this approximation is that an effective transition energy is identified. It has been shown that due to strain and
other effects the selection rules in QDs are relaxed and the electron GS may have a significant overlap with other states than the hole GS [49], [50]. In this respect \( \Phi_{\text{Hg}} \) acts as an effective transition energy substituting for all the various transitions.

The approach outlined here is one of several, which could have been used. The advantages of this approach are that it requires only little detailed knowledge of the exact state structure of the device. The only prerequisites for the calculations are the knowledge of the bound conduction band states [which can be estimated from photoluminescence (PL) spectra for example] and the effective carrier masses (which can be estimated if the material composition is known). The approximation agrees fairly well with more detailed approaches regarding the number of hole states predicted in each dot (typically 3–5 times the number of electron states depending on dot size) and the transition energies.

APPENDIX II

The values of the parameters used in modeling the QD SOA are listed in Table I. We consider the QD device to consist of InAs QDs grown on GaAs substrate and use parameter values for this material system. The width of the inhomogeneous broadening and level splitting correspond to [17] and [51]. The capture and relaxation factors \( Z_{\text{IR}}^F, Z_{\text{IR}}^P, Z_{\text{IR}}^R, \) and \( Z_{\text{IR}}^A \) have been estimated from [17] and results in a maximum relaxation time of 150 fs and a maximum capture time of 2 ps under full inversion of the WL bandedge. The waveguide loss of only 2 cm\(^{-1}\) was reported for QD devices [15]. Effective masses are calculated based on the material compositions using the approach and parameter values outlined in [52]. A high value for the temperature of 380 K corresponding to a worst-case scenario is used in order to partially compensate for the fact that heating is not included in the model. Carrier heating has been pointed out to be present both in bulk [48] SOAs and QD lasers [24], [53]. and is also expected to be an issue in the context of QD SOAs.

The QW device is assumed to consist of three InAs, \( n = 0.026m_0 \), and GaP, \( m^p = 0.062m_0 \), all device parameters used correspond to those used for the barrier material of the QW SOA.

\begin{table}[h]
\centering
\caption{Parameter Values Used for QD SOA}
\begin{tabular}{|c|c|c|c|c|}
\hline
Symbol & Value & Symbol & Value & \\
\hline
\( W \) & 2 \( \mu \)m & \( \Delta \) & 0.65 & \\
\( H \) & 250 nm & \( m \uparrow^w = m \downarrow^w \) & 0.026\( m_0 \) & \\
\( H \) & 1 nm & \( m \uparrow^p = m \downarrow^p \) & 0.0742\( m_0 \) & \\
\( D \) & 4\( \times \)10\(^9\) cm\(^{-2}\) & \( m \uparrow^G = m \downarrow^G \) & 0.062\( m_0 \) & \\
\( n_{\text{sox}} \) & 3 & \( m \) & 0.1628\( m_0 \) & \\
\( \theta_{\text{sox}} \) & 230 nm & \( \alpha_{\text{sox}} \) & 2 cm\(^{-1}\) & \\
\( \Gamma \) & 0.3 & \( \sigma \) & 1\( \times \)10\(^{-3}\) m\(^2\)eV & \\
\( T \) & 380 K & \( \alpha \) & 2\( \times \)10\(^{-3}\) m\(^2\)eV & \\
\( \gamma \) & 3.3 meV & \( Z \) & 0.5\( \times \)10\(^{-3}\) s\(^{-1}\) & \\
\( \gamma \) & 67 meV & \( Z \) & 2.5\( \times \)10\(^{-3}\) s\(^{-1}\) & \\
\( \gamma \) & 80 meV & \( Z \) & 6.7\( \times \)10\(^{-3}\) s\(^{-1}\) & \\
\( \epsilon \) & 2 & \( \epsilon \) & 2.5\( \times \)10\(^{-3}\) s\(^{-1}\) & \\
\( \epsilon \) & 4 & \( \epsilon \) & 1\( \times \)10\(^{-3}\) s\(^{-1}\) & \\
\( E_{\text{IR}}^G \) & 0.48 eV & \( B \) & 1.2\( \times \)10\(^{-10}\) m\(^3\)s\(^{-1}\) & \\
\( E_{\text{IR}}^p,0 \) & 1.3 eV & \( C \) & 5\( \times \)10\(^{-11}\) m\(^3\)s\(^{-1}\) & \\
\( E_{\text{IR}}^p \) & 1.42 eV & \( B \) & 2.35\( \times \)10\(^{-10}\) m\(^3\)s\(^{-1}\) & \\
\( E_{\text{IR}}^G,0 \) & 1.1 eV & \( B \) & 2.5\( \times \)10\(^{-10}\) m\(^3\)s\(^{-1}\) & \\
\( E_{\text{IR}}^F \) & 1.21 eV & \( C \) & 5.24\( \times \)10\(^{-11}\) m\(^3\)s\(^{-1}\) & \\
\hline
\end{tabular}
\end{table}

REFERENCES


BERG AND MØRKK. SATURATION AND NOISE PROPERTIES OF QUANTUM-DOT OPTICAL AMPLIFIERS

1539


Tommy W. Berg received the M.Sc. and Ph.D. degrees from the Technical University of Denmark, Lyngby, Denmark, in 2000 and 2004, respectively. His Ph.D. dissertation was on the modeling of low-dimensional opto-electronic devices.

He is presently an Assistant Professor at Research Center COM, Technical University of Denmark. His current research interests include carrier dynamics in quantum dots and properties of quantum-dot amplifiers.

Jesper Mørk received the M.Sc., Ph.D., and Dr.Tech degrees from the Technical University of Denmark, Lyngby, Denmark, in 1986, 1988, and 2003, respectively.

Since 2002, he has been Professor in semiconductor devices for optical communication systems and is responsible for modeling and theory in the opto-electronics competence area at the Research Center COM, Technical University of Denmark. His current research interests are in the area of device physics, particularly ultrafast devices for optical signal processing, quantum-dot devices, and noise in nonlinear devices.