Time-resolved studies of annealed InAs/GaAs self-assembled quantum dots

S. Malik, E. C. Le Ru, D. Childs, and R. Murray*
Centre for Electronic Materials and Devices, Imperial College, London SW7 2BZ, United Kingdom
(Received 4 August 2000; revised manuscript received 6 November 2000; published 29 March 2001)

We have investigated the carrier dynamics in annealed quantum dots where the energy-level separation of the optical transitions can be tuned between 68 and 19 meV. Photoluminescence transients obtained for different excitation densities for the ground and excited states are well described by a random population model, and values of the exciton lifetimes extracted. The ground-state radiative lifetimes are found to vary from 800 ps in the as-grown sample to 490 ps in the sample annealed at the highest temperature. This is attributed to shifts in the emission energy and changes in the electron and hole wave-function overlap as the dot size increases with annealing. We find no evidence of inhibited relaxation of carriers (phonon bottleneck) when the intersublevel energy-level separation is much less than the LO-phonon energies. The rise times of the transients remain fast at low excitation density, where the probability of Coulomb scattering is expected to be low, meaning that another mechanism, possibly multiphonon scattering, is responsible for the fast relaxation observed in self-assembled quantum dots.

DOI: 10.1103/PhysRevB.63.155313 PACS number(s): 78.47.+p, 68.35.Fx, 78.66.Fd, 78.55.Cr

I. INTRODUCTION

There is as yet no consensus regarding the carrier relaxation mechanisms that operate in self-assembled quantum dots (QD’s). A clearer understanding of the carrier dynamics in QD structures with discrete electronic states is of fundamental interest, but may also be helpful in achieving the ultimate QD laser performance. It was originally proposed that a “phonon-bottleneck” inhibits carrier relaxation when the energy-level separation $\Delta E$ is not resonant with a LO-phonon energy (or a multiple of it). Such a mechanism would severely limit the radiative efficiency, but many studies attest to the strong emission from QD’s, casting doubt on the phonon bottleneck process. Alternative relaxation mechanisms such as Auger or carrier-carrier scattering, interactions with point defects, or multiphonon processes involving acoustic phonons have been proposed but not critically assessed.

In this paper we address two issues of importance in QD research. First, what is the main relaxation mechanism operating in QD’s? Our method here is to perform time-resolved measurements on a series of samples where $\Delta E$ can be varied so that it straddles the LO-phonon energies in these systems. This can be achieved simply by subjecting QD samples to post-growth rapid thermal anneals where interdiffusion of the In and Ga atoms results in larger dots with a reduction in the average indium content. A slower relaxation rate and concomitant decrease in the emission intensity should result if $\Delta E$ is not resonant with a LO (or multiple LO) phonon energy. The second issue concerns the degree of confinement in these structures; the relative values of the exciton binding energy and $\Delta E$ would suggest that these systems conform to the intermediate confinement regime, with the exciton diameter being smaller than the dot. After interdiffusion there should be a gradual evolution toward the weak confinement regime, where the exciton lifetime decreases with increasing dot volume. Hence a suitable sample series would allow an investigation of the processes of relaxation and radiative recombination in QD’s of different sizes and compositions.

II. EXPERIMENTAL DETAILS

The samples were grown by solid source molecular-beam epitaxy using As$_4$. A 200-nm GaAs buffer layer and a six period GaAs/AlAs 6×6 monolayers (ML) superlattice (SL) were deposited at a substrate temperature ($T_S$) of 580°C on a (001) GaAs substrate. $T_S$ was then reduced to 500°C during deposition of 25 nm of GaAs followed by 2 ML of InAs. An interesting feature of our growth procedure is to deposit the InAs at a low growth rate of 0.01 ML/s. This results in relatively large QD’s with a small inhomogeneous linewidth of only 24 meV and a density of 1.7×10$^{10}$ cm$^{-2}$. The dots were then capped with 25 nm of GaAs, and $T_S$ is increased to 580°C for deposition of another GaAs/AlAs SL and a final 50-nm GaAs capping layer. All the layers were undoped. 100 nm of SiO$_2$ was deposited onto a central portion of the wafer, which was then cleaved into pieces, and each piece subjected to 10-s rapid thermal anneals in an argon ambient at temperatures between 675 and 780°C. This provided a selection of QD samples (see Table I) having different ground-state emission energies and state separations.

Low temperature (14 K) time integrated photoluminescence (TIPL) and PL decay (time-correlated single-photon counting) measurements were made using an argon-ion-
spectrally resolving

30 ps. Time resolved PL measurements were made by
of the detection system. spectra were normalized to account for the spectral response
non resonances are present in the spectra. All the optical
relatively high spectral resolution is necessary if sharp pho-

pumped Ti:sapphire laser delivering 1.2-ps pulses at 82 MHz
exciting into the GaAs barrier (≈1.60 eV). The laser was
focused to a ≈200-μm-diameter spot. The luminescence
was dispersed with a SPEX 1404 double-grating monochro-
mator and detected with a Hamamatsu microchannel plate
with an extended S1 response. The temporal resolution is
≈30 ps. Time resolved PL measurements were made by
spectrally resolving (resolution of 1 meV) the signal con-
tained within a ±50-ps window with an adjustable delay. A
relatively high spectral resolution is necessary if sharp pho-
on resonances are present in the spectra. All the optical
spectra were normalized to account for the spectral response
of the detection system.

III. EXPERIMENTAL RESULTS

A. Time-integrated spectra

Figure 1 shows the low-temperature time integrated pho-
luminescence spectra obtained at high excitation
(347 W cm⁻²) from the unannealed QD sample (A) and
samples cut from the same wafer and annealed for 10 s at
675 °C (B), 700 °C (C), 750 °C (D), and 780 °C (E). Annealing shifts the ground-state (GS) emission energy E₀
from 1.047 to 1.393 eV corresponding to a maximum blue-
shift of ≈350 meV. Each sample exhibits strong level-
filling effects, and the separation of the optical transitions,
∆E, decreases from 68 meV for the as-grown sample to 19
meV for sample E. Since all the samples exhibit similar
level-filling behaviors with equally spaced excited states, we
conclude that the QD-like properties are retained after an-
nealing. Under lower excitation conditions, when only
ground-state emission is significant, similar integrated PL in-
tensities are observed for each sample. This suggests that the
quality of the samples does not deteriorate upon annealing,
and implies that they are also free from significant nonradi-
active recombination in the dots. The optical transitions are
well resolved, making them ideally suited for time resolved
measurements. Table I lists the important parameters derived
from the TIPL measurements for all the samples investigated
here. The full width at half maximum (FWHM) of the GS
peak exhibits a monotonic decrease with the annealing tem-
perature, although the as-grown sample does not fit in with
this trend. This behavior was already reported. The im-
portant point for this work is that the overlap of the emission
peaks remains small. The confining/wetting layer (CL/WL) also
exhibit a slight shift with annealing, similar to the interdiffusion effects reported for quantum wells.

B. Time-resolved measurements

Photoluminescence decay transients were obtained for all
the samples at excitation powers of 347, 12.1, and
1.1 W cm⁻², designated high, medium, and low-
respectively. Figure 2(a) shows the results for sample C. Al-
though there have been many similar studies performed on
as-grown QD samples, in this paper we would like to
demonstrate the importance of normalizing the transients in
order to extract the maximum information. The data shown
in Fig. 2(a) need to be normalized to account not only for the
detector spectral response but also for the relative emission
intensities of the transitions. In principle this can be achieved
simply by measuring the detector counts at constant time for
each transition. An alternative method is to measure time-
resolved photoluminescence (TRPL) spectra at different de-
lays after the laser pulse. The spectrally resolved signal then
provides a good illustration of the temporal evolution of the
carrier distribution among the QD levels. Figure 3 shows the
TRPL spectra for delays up to ≈3 ns for sample C. It
should be noted that similar PL transients and TRPL spectra
were obtained for all the samples, and differences between
the samples will be discussed in Sec. IV. The PL transients
of each transition shown in Fig. 2(a) have been scaled ac-
cording to the TRPL spectra shown in Fig. 3 for all delays.
This process is repeated for medium and low excitation
power.

IV. DISCUSSION

The series of samples investigated here were specifically
designed to investigate two outstanding issues relating to
QD’s: the existence of a phonon bottleneck that inhibits car-
rier relaxation, and the influence (if any) of changes in the
dot size and composition on the radiative decay rates. There
have been many previous studies that have touched on these
issues, but we believe that their conclusions were limited
primarily by sample range and quality. In what follows we
will attempt to clarify the situation although, as will become
clear, a complete resolution will require further experiments.

A. Carrier capture and relaxation in annealed quantum dots

The phonon bottleneck was originally proposed to explain the
poor luminescence efficiency of QD’s produced by pho-
tolithographic methods. However, as the results presented here and in other publications show, self-assembled QD’s exhibit relatively high luminescence efficiencies. This implies that the relaxation is not inhibited in these structures. Relaxation may of course occur via other mechanisms such as Auger scattering, interactions with impurities, or multiphonon processes. A mechanism involving the presence of a quasi-continuum in the QD’s was also proposed.

In principle, time-resolved experiments can quantify carrier relaxation rates. However, as will be pointed out in Sec. V, the shape of the decays is fairly insensitive to the exact values of the relaxation times between states when these times are much smaller than the radiative lifetimes. Therefore, only the rise times of the PL transients can provide information on carrier capture and relaxation in most cases. Given the temporal resolution of our system, we are not able to detect fast rise times which could be due to phonon resonances or to strong Auger scattering. However, we can detect a decrease of the rise times if they become longer than 40 ps. The phonon bottleneck effect predicts such a decrease when the interlevel spacing $\Delta E$ is not resonant with a LO-phonon energy. Many experiments have shown that the rise times are fast, but these are usually performed at very high injection rates that would favor Auger scattering, which could hide the phonon bottleneck effect. Moreover, they do not span a wide range of $\Delta E$ values. Heitz et al. and Morris et al. reported relaxation lifetimes of around 40 ps, using measurement systems with better temporal resolutions (5 ps), and any slowdown could therefore be detected by our system. Therefore, we have performed experiments at different excitation densities, particularly low densities (1.1 W cm$^{-2}$) where Auger scattering should be weak, on samples where $\Delta E$ varies from 68 to 19 meV. Although the LO-phonon energies are not well known in QD samples, they are usually quoted to lie in the range 28–38 meV. We can therefore expect a significant slowdown of the rise times for samples A and B for which $\Delta E$ is larger than any LO-phonon energy. A similar effect can be expected for sample E, where $\Delta E$ is well below the LO-phonon energies.

The measured values of the rise times for all the annealed samples (B–E) are indistinguishable, and close to the temporal resolution of our system. In addition, there are no measurable differences over the range of excitation levels employed (high, medium, or low). This implies that the relaxation between QD states for this range of $\Delta E$ (19–43 meV) is relatively fast (less than 40 ps). If the relaxation was much faster for sample C, where $\Delta E$ is close to the bulk LO-phonon energy, we could not detect it. However, we do see that the relaxation remains fast (below 40 ps) for samples B and E, where longer times are predicted, and this is also true under low excitation. Therefore, there must be another
B. Radiative lifetimes in annealed quantum dots

An aspect of self-assembled QD’s, which has not been extensively studied is the influence of dot size, and composition on the radiative lifetime. The exciton Bohr radius is comparable with the as-grown dot size and this corresponds to the intermediate confinement regime where the exciton binding energy is less than $\Delta E$. Estimates of the exciton binding energy are in the range 10–20 meV, and are comparable with $\Delta E$ in the samples annealed at the highest temperatures. Therefore we might expect annealed QD’s to tend toward weaker confinement and the exciton lifetime to decrease with the increase in dot volume.

The GS exciton lifetime $\tau_0$ can be estimated from the PL transients at long times, where feeding from higher-lying states is negligible. The values for sample C are listed in Table II, and appear to be independent of the excitation level. Extracting values of the lifetime of the first excited state, $\tau_1$, from the PL transients is a more difficult task, since there is feeding from higher-lying states and relaxation to the ground state. Here we present an independent method of estimating $\tau_1$ based on the PL transients at high excitation, where state blocking of the ground and first excited states leads to plateaus in the detected signal. For the duration of the plateau ($\approx 1.5$ ns), carriers emitted from these states are immediately replaced by carriers from higher-lying states, and these states can be considered full during this time. Assuming that the radiative lifetimes of the ground and first excited states are unaffected by the presence of carriers in the upper states (therefore neglecting coulombic interactions), the ratio of the PL intensities is given by

$$R = \frac{I_1}{I_0} = \frac{g_1}{g_0} \frac{\tau_0}{\tau_1}. \quad (1)$$

A similar method can also be used in cw experiments. The ratio of the degeneracies for the two levels is usually quoted as $g_1/g_0 = 2$, and it is then possible to evaluate $\tau_1$ knowing $\tau_0$ and $R$. However, the method requires an accurate estimate of $R$ and depends on an accurate normalization of the transients using the TRPL spectra. $R$ is measured to be 1.6, giving $\tau_1 = 855$ ps for sample C [Figs. 2(a) and 3]. Comparing this with $\tau_0$, we see that the lifetime of the first excited state is only slightly longer than the GS value. This contrasts with previous work, which deduced much larger differences between $\tau_1$ and $\tau_0$ for QD’s with similar emission energies. The method presented here to derive the radiative lifetimes does not rely on a measurement of the decay rate of the first excited state, and we therefore believe it is more reliable.

As pointed out in Sec. III, the relaxation times are much faster than the radiative lifetimes. The random population (RP) model then predicts that the measured decay $\tau_1^{\text{meas}}$ of the first excited state at long times is given by

$$\frac{1}{\tau_1^{\text{meas}}} = \frac{1}{\tau_1} + \frac{g_0}{\tau_0}, \quad (2)$$

where $g_0$ is the degeneracy of the GS. Values of $\tau_1^{\text{meas}}$ for the excited states are extracted from Fig. 2(a), and are listed in Table II. Again, there are no significant differences in the values obtained for the different excitation conditions. Inserting the values for $\tau_0$ and $\tau_1^{\text{meas}}$ into Eq. (2) yields a negative value for $\tau_1$ when $g_0 = 2$, which is clearly unrealistic. The same applies to the other annealed samples. Other authors have ignored (without justification) the factor of 2 in this equation. In our case, the measured values for the decays show that the assumption $g_0 = 2$ cannot be correct. However, assuming $g_0 = 1$ leads to a value of $\tau_1$, in good agreement with the one derived earlier. This implies that the dots be-
TABLE III. Values of the radiative lifetimes (in ps) of the transitions used in the modeling for each sample. The fits are quite sensitive to the exact values for the ground (GS), first excited (X1), and second excited (X2) states. However, they are only slightly influenced by the exact values of the higher-lying states, which are consequently less reliable. These states still need to be included because they act as a reservoir under high excitation.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS</td>
<td>800</td>
<td>650</td>
<td>600</td>
<td>550</td>
<td>490</td>
</tr>
<tr>
<td>X1</td>
<td>920</td>
<td>783</td>
<td>720</td>
<td>611</td>
<td>563</td>
</tr>
<tr>
<td>X2</td>
<td>889</td>
<td>903</td>
<td>812</td>
<td>689</td>
<td>653</td>
</tr>
<tr>
<td>X3</td>
<td>1143</td>
<td>1083</td>
<td>941</td>
<td>917</td>
<td>563</td>
</tr>
<tr>
<td>X4</td>
<td>1256</td>
<td>1300</td>
<td>1100</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

have as if the degeneracy of the GS is only 1. One possible explanation is that relaxation from the first excited state to the ground state becomes long when the GS is only half-filled (one electron-hole pair). This observation is also consistent with phonon-assisted relaxation, where the spin should be conserved. To test for self-consistency of our data, we have applied the RP model to fit the decay transients. As Bimberg and co-workers pointed out,31 the results of the models are insensitive to values of relaxation times smaller than 100 ps. We therefore assume the relaxation times to be infinitely fast. The measured decay from the CL/WL is taken to reflect the temporal evolution of the population of the reservoir from which carriers are captured into the dot. A factor (similar to a generation rate $G$) is applied to this signal to allow for changes in the intensity of the signals at different excitation levels. The only other input parameters are the radiative lifetimes for each transition. The model is then able to predict the shape of the decays and the relative intensities of the signals from different states. Good fits to the data were obtained for all the samples, but only when assuming degeneracies of 1, 2, 3... for the ground, first, second, etc. excited states. For each sample, it is also important to include in the model the exact number of excited states (four for samples A, B, and C, three for sample D, and two for sample E) to predict the correct duration of the plateaus. This is reasonable since the duration of the plateau of the GS under high excitation will depend strongly on the number of carriers that can be accommodated in the higher excited states of a dot. The fits shown in Fig. 2(b) are extremely good, and the lifetimes for each state extracted from the model are listed in Table III. For the model to be self-consistent, the data obtained at medium and low excitation density must also be fitted by changing $G$ only. These fits are shown in Fig. 2(b). The agreement is again very good for all the samples. In particular, the model predicts the correct changes in the relative intensities of the states and the duration of the plateaus which disappear for low excitation. Comparing the results from the three methods—from the decays in the transients, from the plateaus in the PL transients and from the RP model—we see that the agreement is good between the latter methods, placing greater confidence in the values obtained.

Figure 4 shows values of $\tau_0$ and $\tau_1$ plotted against the GS emission energy for all the samples. As expected, there is a reduction in the radiative lifetimes as the dot volume increases. Extrapolating the trend to the energy corresponding to the CL/WL, where we expect a two-dimensional layer of uniform composition, would lead to a decay time of $\approx 400$ ps, in good agreement with values measured for In$_x$Ga$_{1-x}$As/GaAs quantum wells.32,33 The radiative lifetime is frequently quoted as $E_x$, and Eq. (3) gives the predicted variation of the GS radiative lifetime $\tau_0$ from Eq. (3), $\tau_1$ plotted against the GS emission energy ($E_\text{v}$). The solid line is the predicted variation of the GS radiative lifetime $\tau_0$ from Eq. (3), $\tau_1$ plotted against the GS emission energy ($E_\text{v}$). The inset gives the values of the overlap factor $|\langle F_c | F_\text{v} \rangle|^2$ deduced from the experimental values of $\tau_0$ and Eq. (3). Dotted lines are a guide for the eyes.

\[ \frac{1}{\tau} = \frac{e^2 E_p n \omega}{2 m_0 c_0} |\langle F_c | F_\text{v} \rangle|^2, \]  

where $E_p$ is the Kane energy (taken to be 22 eV), $n$ the refractive index of the dot material (taken to be 3.6), $\omega$ the emission frequency, and $|\langle F_c | F_\text{v} \rangle|^2$ the overlap of the electron and hole wave functions. Assuming a constant overlap $\tau_0$ is inversely proportional to the GS energy $E_0$, and the solid line in Fig. 4 shows the predicted decrease in $\tau_0$ for the annealed samples. However, the agreement with the experimental values is not perfect, and the predicted values are up to 20% higher for the samples subjected to the highest-temperature anneals. The difference can be explained by a slight increase in the overlap $|\langle F_c | F_\text{v} \rangle|^2$ with annealing, and the inset to Fig. 4 shows the predicted variation of $|\langle F_c | F_\text{v} \rangle|^2$ using the data for $\tau_0$ and Eq. (3). The size of the dots increases with annealing and the diameter can be doubled for the samples annealed at the highest temperature.15 If the exciton Bohr radius is comparable with the size of the as-grown dots and there is penetration into the barrier, an increase in $|\langle F_c | F_\text{v} \rangle|^2$ as the dot increases in size is plausible. Finally, $\tau_1$ is slightly larger than $\tau_0$ in all the samples, con-
consistent with a smaller overlap of the electron and hole wave functions in the excited states.

V. CONCLUSION

Time-integrated and time-resolved measurements have been made on a series of annealed QD samples, where the separation of the optical transitions lies in the range 19 \( \approx \Delta E \approx 68 \) meV. The discrete nature of the electronic states is not lost as a result of annealing, and the samples exhibit clear optical transitions from excited states. Using a combination of PL decay measurements and TRPL spectra, the PL transients can be normalized, and values extracted for the lifetime of the GS excitons. The lifetime of the first excited state can be deduced from measurements made at high excitation densities where the PL transients exhibit flat plateau regions. The values obtained are in good agreement with the results of the random population model, and the effect of the degeneracies is discussed. Both the ground- and first-excited-state lifetimes decrease with annealing. This decrease is mainly explained by the blueshift of the transitions. The slight deviation from this can be interpreted by an increase in the oscillator strengths as the volume of the dots increases. The relaxation times between states in all the annealed QD’s, as deduced from the PL rise times, are fast, with an upper limit of 40 ps, even under low excitation densities, implying that there is no strong phonon bottleneck effect in self-assembled QD’s. It is likely that multiphonon scattering on the scale of a few tens of ps is responsible for relaxation under these conditions. Only an as-grown sample with a relatively wide intersublevel spacing exhibits a slightly slower relaxation.