Towards Simultaneous Improvements of Storage and Erase Times in Quantum Dot-Based Memories

vorgelegt von

M. Sc. Physik
Ismail Firat Arikan
geb. in Istanbul

von der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzende:	Prof. Dr. rer. nat. Michael Lehmann
Berichter/Gutachter:	Prof. Dr. phil. nat. Dieter Bimberg
Berichter/Gutachter:	apl. Prof. Dr. rer. nat. Armin Dadgar

Tag der wissenschaftlichen Aussprache: 6. April 2018

Berlin 2018

Abstract

This work attempts to improve memory devices based on self-organized semiconductor quantum dots (QDs) in two ways. The first aim is to increase the hole storage time at room temperature in such a memory (QD-Flash) to achieve non-volatility. Hence, the electronic properties of QDs, such as the localization energy and the capture cross section, are determined by static capacitance-voltage (C-V) spectroscopy and Deep Level Transient Spectroscopy (DLTS). Using the mentioned two quantities the hole storage time is extrapolated at room temperature. The second aim in this work is to decrease the erase time in the QD-Flash by eliminating the trade-off between storage and erase times through insertion of quantum wells (QWs) into the barrier. For this, a transparency engineering method based on a One-Dimensional Schrödinger Poisson Solver (one-DSPS) and Non-Equilibrium Green's function (NEGF) is developed.

In the first part, the electronic properties of two different material combinations are investigated for the first time: GaSb QDs embedded in GaP and InGaSb QDs embedded in GaP with an additional AlP barrier. A localization energy E_{loc} = 1.18 ($\pm$0.01) eV is determined for the GaSb/GaP system with an associated capture cross section $\sigma_\infty = 1\times10^{-12}$ cm^2 (with an uncertainty of 0.1 orders of magnitude). The corresponding hole storage time at room temperature is 3.9 days (with an uncertainty of 0.04 orders of magnitude), which is the longest storage time ever reported and marks an improvement of 3 orders of magnitude with respect to the previous best result. Moreover, different samples based on the GaSb/GaP QDs are grown under different growth conditions to investigate the effect of growth conditions on the electronic properties of QDs. It is demonstrated that higher temperatures and slower growth rates lead to bigger QDs in size, resulting in larger localization energies. For the InGaSb/GaP QD system a localization energy E_{loc} = 1.15 ($\pm$0.02) eV is determined with an associated capture cross section $\sigma_\infty = 9 \times 10^{-11}$ cm^2 (with an uncertainty of 0.2 orders of magnitude). The corresponding hole storage time at room temperature is obtained to be 3200 s (with an uncertainty of 0.1 orders of magnitude), marking an improvement of 1 order of magnitude in comparison to the structure based on As (InGaAs QDs embedded in GaP).

In the second part, different resonant tunneling structures based on GaAs and GaP are designed and simulated. The calculations show that the transparency of the tunnel barrier increases by 7 orders of magnitude for the GaAs-based structures with one QW due to the creation of resonance throughout the device, while the transparency increases by 10 orders of magnitude for the GaAs-based structures with two QWs. Moreover, the calculations demonstrate that resonant tunneling can be achieved in GaP-based structures, which are promising candidates to achieve non-volatility.

Based on the results of this work, the following recommendations are given to improve QD-Flash: the localization energy can be further increased by altering the growth conditions of the QD layer to achieve non-volatility. Moreover, an additional barrier can be implemented to increase the localization energy, hence the storage time. Alternatively, the capture cross section can be decreased to increase the storage time. For this, techniques such as Sb-soaking or growth interruption (GRI) can be employed. For the resonant tunnelling structures, the epitaxial growth of the samples has to be achieved and the simulation results given in this work have to be confirmed experimentally.

Zusammenfassung

Diese Arbeit verfolgt zwei Wege, um selbstorganisierte quantenpunktbasierte Speicher (QD-Flash) zu verbessern. Das erste Ziel ist die Verlängerung der Lochspeicherzeit im quantenpunktbasierten Speicher bei Raumtemperatur, um Nichtflüchtigkeit zu erreichen. Dafür werden die elektronischen Eigenschaften der selbstorganisierten Halbleiter-Quantenpunkte (QP), d.h. die Lokalisierungsenergie und der Einfangquerschnitt, mittels statischer Kapazitätsspektroskopie (C-V) und zeitaufgelöster Kapazitätsspektroskopie (DLTS) untersucht. Über die Lokalisierungsenergie und den Einfangquerschnitt wird die Lochspeicherzeit bei Raumtemperatur extrapoliert. Das zweite Ziel ist die Reduzierung der Löschzeit des quantenpunktbasierten Speichers durch Auflösen des Konflikts zwischen Speicherzeit und Löschzeit, indem Quantengräben (QWs) in die Barriere eingefügt werden. Hierfür wird eine Transparenzanpassungsmethode auf der Grundlage eines eindimensionalen Schrödinger-Possion-Solvers (one-DSPS) und Nichtgleichgewichts-Greenschen-Funktionen (NEGF) entwickelt.

Im ersten Teil werden die elektronischen Eigenschaften für zwei unterschiedliche Materialsysteme zum ersten Mal untersucht: in GaP eingebettete GaSb-Quantenpunkte und in GaP eingebettete InGaSb-Quantenpunkte mit einer AlP-Barriere. Die gemessene Lokalisierungsenergie des GaSb/GaP-Quantenpunkt-Materialsystems ist $E_{loc} = 1.18$ (± 0.01) eV mit einem Einfangquerschnitt $\sigma_\infty = 1 \times 10^{-12}$ cm^2 (Messabweichung: 0.1 Größenordnungen). Die extrapolierte Speicherzeit bei Raumtemperatur ist 3.9 Tage (Messabweichung: 0.04 Größenordnungen). Dies ist die längste Lochspeicherzeit, die bisher berichtet wurde; sie führt zu einer weiteren Verbesserung der Lochspeicherzeit um 3 Größenordnungen. Zudem werden unterschiedliche Proben bei unterschiedlichen Wachstumsbedingungen für die Quantenpunkte gewachsen, um den Effekt der Wachstumsbedingungen auf die elektronischen Eigenschaften der Quantenpunkte zu untersuchen. Die Messungen bestätigen, dass höhere Temperaturen und langsameres Wachstum größere Quantenpunkte bewirken, also zu einer größeren Lokalisierungsenergie führen. Die gemessene Lokalisierungsenergie der InGaSb/GaP-Quantenpunkt-Materialsysteme ist $E_{loc} = 1.15$ (± 0.02) eV mit einem Einfangquerschnitt $\sigma_\infty = 9 \times 10^{-11}$ cm^2 (Messabweichung: 0.2 Größenordnungen). Die extrapolierte Speicherzeit bei Raumtemperatur ist 3200 Sekunden (Messabweichung: 0.1 Größenordnungen). Dies macht einen weiteren Fortschritt von einer Größenordnung im Vergleich mit dem As-Materialsystem (InGaAs/GaP) aus.

Im zweiten Teil wurden unterschiedliche GaAs-basierte und GaP-basierte resonante Tunnelstrukturen entworfen und simuliert. Die Simulationen bestätigen, dass die Transparenz der GaAs-basierten Strukturen mit einem Quantengraben (QW) eine Steigerung der Transparenz um 7 Größenordnungen ausmachen, während die Steigerung der Transparenz bei den GaAs-basierten Strukturen mit zwei Quantengräben 10 Größenordnungen erreicht. Die Berechnungen bestätigen auch, dass das Konzept des resonanten Tunnelns für GaP-basierte Strukturen anwendbar ist, welche die aussichtsreichsten Kandidaten für die Erreichung von Nichtflüchtigkeit in Quantenpunktspeichern sind.

Basierend auf den Ergebnissen dieser Arbeit werden Vorschläge für eine weitere Verbesserung des quantenpunktbasierten Speichers (QD-Flash) gemacht: die Lokalisierungsenergie kann durch Änderungen der Wachstumsbedingungen der Quantenpunkte weiter erhöht werden, um Nichtflüchtigkeit zu erreichen. Außerdem kann die Lokalisierungsenergie, d.h. die Speicherzeit, durch den Einsatz einer Barriere weiter erhöht werden. Alternativ kann die Speicherzeit durch Verringerung des Einfangquerschnitts verlängert werden. Dafür können Methoden wie Sb-Soaking und Wachstumsunterbrechung

(GRI), verwendet werden. Im Hinblick auf die resonanten Tunnelstrukturen muss epitaktisches Wachstum dieser Strukturen durchgeführt werden und die Simulationsergebnisse dieser Arbeit müssen experimentell bestätigt werden.

Publications

Journal

. L. Bonato, I. F. Arikan, L. Desplanque, C. Coinon, X. Wallart, Y. Wang, P. Ruterana, and D. Bimberg, "Hole localization energy of 1.18 eV in GaSb quantum dots embedded in GaP", *Physica status solidi (b)* 10, pp. 1877-1881 (2016).

https://onlinelibrary.wiley.com/doi/abs/10.1002/pssb.201670563

Accepted

. I. F. Arikan, N. Cottet, T. Nowozin, and D. Bimberg, "Transparency engineering in quantum dot-based memories", *Physica status solidi (a)* (In print).

Submitted

. E. M. Sala, I. F. Arikan, L. Bonato, F. Betram, P. Veit, J. Christen, A. Strittmatter, and D. Bimberg, "MOVPE-growth of InGaSb/AlP/GaP(001) quantum dots for nano-memory applications", *Physica status solidi (b)*.

Conference

. I. F. Arikan, T. Nowozin, D. Bimberg, and N. Öncan, "Resonant tunneling structures to improve the erase time in memory devices based on quantum dots", *DPG Frühjahrstagung*, 2015.

. I. F. Arikan, N. Cottet, T. Nowozin, D. Bimberg, and N. Öncan, "Resonant tunneling structures for reducing the erase time in a quantum dot-based memory", *DPG Frühjahrstagung*, 2014.

Aileme

Contents

Introduction

Storing information in digital media and accessing the stored data (processing, transmitting, retrieving) are crucial requirements in our modern society. The global datasphere was 19 zettabyte[1] in 2017 according to the *Data Age 2025*'s report [1], and it doubles in size every year [2]. Following this trend, it is predicted to reach 163 zettabyte by 2025 [1]. Hence, the greatest challenge in memory market is to meet the excess amount of digital data.

The miniaturization method is used in order to meet the increasing digital data. Accordingly, the size of components is scaled down, so that higher number of components can be packed into a given area. The scaling down is achieved in accordance with the *Moore law* which states that "the number of components per unit area on a chip doubles approximately every 24 months" [3].

However, the miniaturization approach can not be employed indefinitely due to two major limitations. First one is that the fabrication cost for scaling down increases as well [3] because the miniaturization method needs special techniques such as lithography and etching. The second limitation is that quantum effects become significant at very small scales, negatively affecting the functionality of the device. For instance, the fluctuations in interface roughness lead to uncontrollable tunneling leakage. Based on this fact, the scaling down approach is predicted to reach its limit by 2020 [4]. In order to sustain the progress in memory market new functionalities have to be implemented [5].

Nowadays, two different types of semiconductor memory device are extensively used: Dynamic Random Access Memory (DRAM) [6–8] and Flash [9–11]. They both have some individual advantages and disadvantages in terms of volatility, access time, and durability. DRAMs have a fast access time, $\sim$10 ns, and a high durability, $\sim 10^{15}$ write/erase cycles. However, DRAMs are volatile memory devices, which means the stored information is lost after a certain period of time and they need to be re-written, resulting in high energy consumption. On the other hand, Flash can store information at least ten years at room temperature without needing energy, hence they are non-volatile. However, their access time is slow, around ms and their durability is poor, $\sim 10^5$ write/erase cycles.

The goal in memory device market is to combine the advantages of both DRAM and Flash Memory into one device [12], which fulfills non-volatility (10 years) and provides sufficiently fast access time (ns) at high durability (10^{15} times of write/erase cycle), in other words, to have a non-volatile device with a fast access time, the *ultimate memory device* [5].

A memory device based on self-organized quantum dots (QD-Flash) [13–16] is a promising candidate for the ultimate memory device. Self-organized QDs [17, 18] formed by III-V semiconductors act as storage nodes in QD-Flash. Carriers can be confined inside them, stay there for a certain period of time, and finally be emitted. In order to perform operations a modulation doped field effect transistor (MODFET) is used, hence band structure is altered by applied bias. Previous works demonstrated

[1] 1 zettabyte = 1 billion terabytes = 10^{21} bytes

230 s of storage time at room temperature [19], 6 ns of write time [20], 3 ns of read time [21], and 44 ns of erase time [22]. Moreover, the full functionality of QD-Flash is demonstrated at low temperatures [22–24]. However, two problems still have to be solved: The storage time has to be increased above 10 years to achieve non-volatility and the erase time has to be decreased below $\sim$ 10 ns (that of DRAM).

This work attempts to solve these two problems. In the first part, two material combinations which are promising to achieve non-volatility are investigated: GaSb QDs embedded in GaP, and InGaSb QDs embedded in GaP with an additional AlP barrier. The electronic properties of QDs, such as the localization energy and capture cross section, are determined by static capacitance spectroscopy and Deep Level Transient Spectroscopy (DLTS). Finally, the hole storage time at room temperature is extrapolated using the mentioned two parameters. In the second part, a transparency calculation method based on a one-Dimensional Schrödinger-Poisson Solver and Non-Equilibrium Green's Function is developed to decrease the erase time by eliminating the trade-off between storage and erase times. Insertion of quantum wells (QWs) into the structure eliminates the trade-off by creating tunneling resonances as erase mechanism. In such a structure, it is possible to switch transparency between very high and very low values by altering a bias voltage, on which erase time depends. Therefore, different resonant tunneling structures based on various material combinations are designed and simulated.

The thesis is organized in that way:

Chapter 1 describes the fundamental concepts of semiconductors and quantum dots with their fabrication techniques and the electronic properties. Semiconductor memory devices are explained emphasizing DRAM and Flash. Finally, QD-Flash concept is introduced.

Chapter 2 explains the carrier dynamics in quantum dots. The understanding of carrier dynamics in quantum dots is essential to understand all the operations in a QD-Flash.

Chapter 3 presents the band engineering to decrease the erase time. The transparency calculation is explained in this chapter.

Chapter 4 introduces the measurement methods used in this work to determine the localization energies and the capture cross sections of the samples, thus their storage times.

Chapter 5 presents the DLTS results of the GaSb/GaP samples.

Chapter 6 contains the DLTS results of the InGaSb/GaP samples.

Chapter 7 introduces the simulation results of resonant tunneling structures.

Chapter 8 contains the conclusion and the outlook.

Appendix A contains the detailed information of all the samples.

Appendix B explains the processing of the samples.

Appendix C introduces the experimental setups with the detailed illustrations.

Appendix D presents the error margin calculation with the results for each sample.

Chapter 1

Fundamentals

This chapter presents the fundamentals of the work. Initially, semiconductor heterostructures are described, which are the basis of p-n junctions and quantum dots (QDs). Then, QDs are explained with their fabrication techniques and electronic properties. Afterwards, the currently used memory devices are presented with their advantages and disadvantages. Finally, the QD-Flash concept is introduced. At the end, the chapter is summarized.

1.1 Heterostructures

A heterostructure is formed by bringing into contact two or more semiconductor materials. The place where they interact each other is called *heterojunction* and it determines the electronic properties of the whole structure. A huge variety of heterostructures are obtained by using different semiconductor materials, especially III-V and II-VI semiconductor groups [25]. Fig. 1.1 shows the energy gap-lattice constant graphs for some semiconductor compounds. Beside the binary compounds, it is possible to fabricate ternary and quaternary compounds. Therefore, a heterojunction that exhibits some specific electrical or optical behavior can be designed according to the required application. Epitaxial growth techniques such as Molecular Beam Epitaxy (MBE) [26] and Metal-Organic Chemical Vapor Deposition (MOCVD) [27] enable us to grow heterostructures with a high degree of precision. High-electron-mobility transistors (HEMT) [28,29], lasers [30,31], amplifiers [32,33], and LEDs [34] can be named as examples of applications based on heterostructures. One of the most prominent hetetostructure application is the double-heterostructure laser invented by Z.I. Alferov and H. Kroemer in 1963, which was awarded with the Nobel Prize in Physics in 2000 [35,36].

An important parameter for a heterostructure is band alignment, i.e the relative positions of the valence band and conduction band edges. Heterostructures generate 3 different band alignment [37,38] as shown in Fig. 1.2.

- *Type-I*: The bandgap of one semiconductor lies completely inside the bandgap of the other semiconductor. Therefore, confinement occurs for both electrons and holes. (e.g. InAs/GaAs)

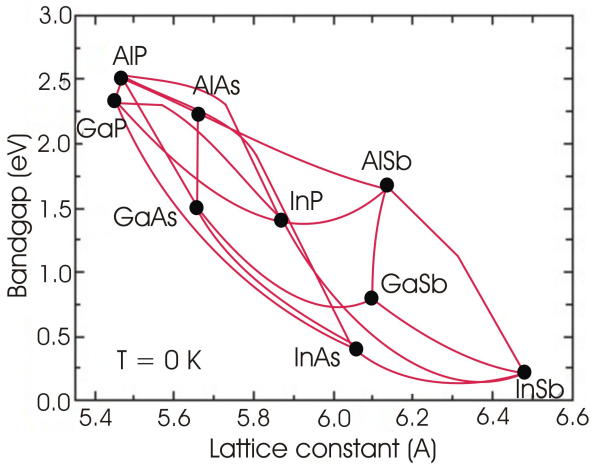

Figure 1.1: Energy gap versus the lattice constant of III-V semiconductor compounds. The black dots indicate the binary compounds, while the red lines indicate the ternary alloys (after [25]).

- *Type-II staggered*: Only the valence (or conduction) band edge of one semiconductor lies inside the band gap of the other. Therefore, confinement exists for only one type of carrier. (e.g. GaSb/GaAs)
- *Type-II broken-gap*: Both conduction and valence band edges lie outside the band gap of other semiconductor. In other words, the band gaps do not overlap at all. Therefore, there is no confinement of either electrons or holes. (e.g. InAs/GaSb)

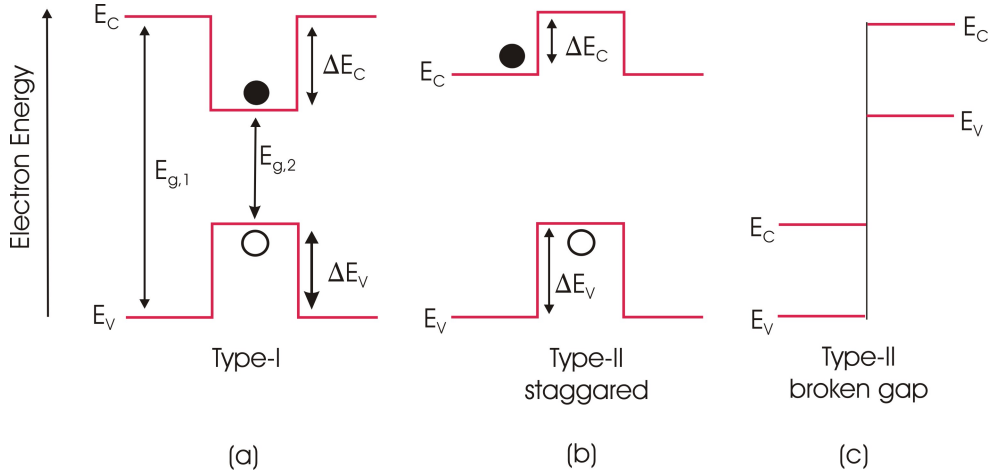

Figure 1.2: Possible band alignments in heterostructures. E_C is the conduction band edge, E_V the valence band edge. (a) Type-I alignment where confinement is possible for both electrons and the holes. (b) Type-II staggered band alignment where the confinement exists only for holes, and (c) Type-II broken gap band alignment where there is no confinement for either electrons or holes.

1.2 Low-Dimensional Structures

Opposed to the bulk materials that allow carriers to freely move in all three spatial dimensions, low dimensional materials restrict the motion of carriers in one, two or all three spatial directions. The confinement on the motion of carriers occurs when the size of the material is comparable to the wave length of carriers, i.e. the *De Broglie wave length*. The De Broglie wavelength is given by

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{3m^* k_B T}} \tag{1.1}$$

where h is the Planck constant, m^* the effective mass, k_B the Boltzmann constant, and T the temperature. The De Broglie wavelength depends on the temperature and the effective mass of the material. Unlike the free mass, the effective mass takes into account the effect of interactions of the charge carriers with the surrounding lattice. Therefore, it varies depending on the material.

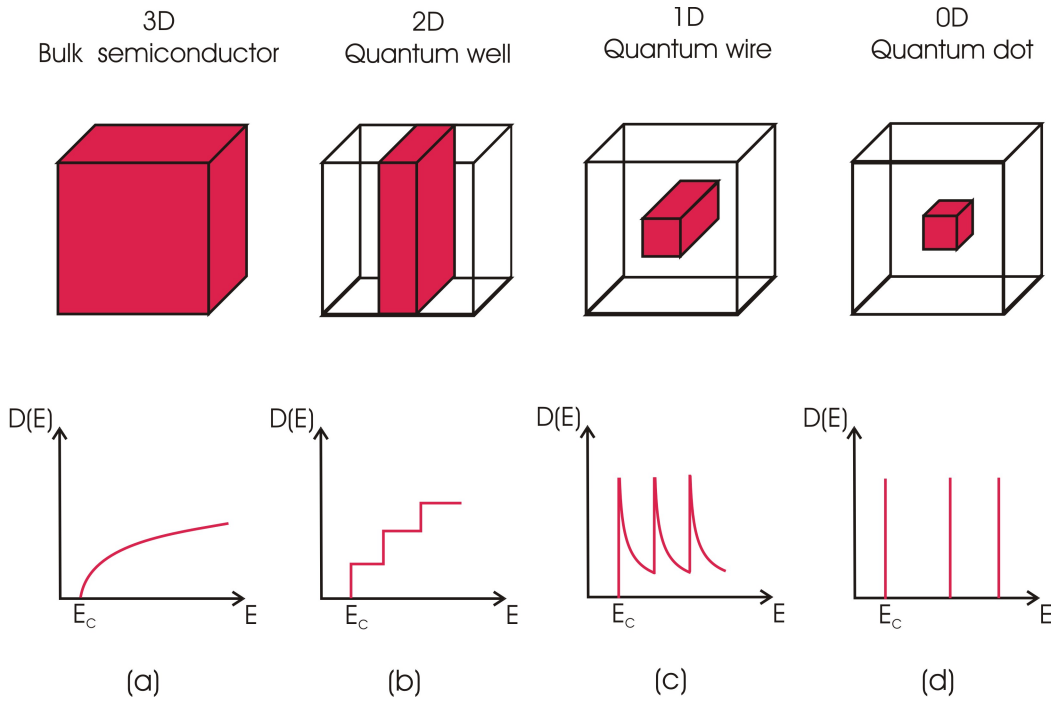

Figure 1.3: The degree of spatial restriction on the motion of carriers (top row) and the resulting density of states (bottom row) for low dimensional structures. (a) No confinement exists for bulk material having the density of states proportional to $\sqrt{E}$. (b) Quantum well where the confinement exists in only one spatial dimension and the density of state is step-like. (c) Quantum wire. The confinement occurs in two dimensions and the density of states is spike-like. (d) Quantum dots where the confinement occurs in all the spatial directions and the density of states is discrete.

Spatial confinement of the carriers induces a modification of the density of states, which describes the number of energy states available to the carriers. The carriers can have discrete energy values, instead of having continuous spectra. Hence, they exhibit different optical and electrical behaviors which are governed by quantum physics.

Low dimensional structures are classified according to the number of restricted spatial direction. A structure is called quantum well (2D) when the restriction occurs in only one direction while carriers can freely move in the remaining two directions. The density of states for a quantum well is step-like (see Fig. 1.3b) and given by [39]

$$D_{2D}(E_i) = \frac{m^*}{\pi\hbar^2} \tag{1.2}$$

where m^* is the effective mass and $\hbar$ reduced Planck constant.

Similarly, we call the structure quantum wire (1D) when the restriction exists in two spatial dimensions and the carriers move freely in the remaining dimension. The density of states for a quantum wire is spike-like (see Fig. 1.3c). Finally, the low dimensional structures where the confinement exists in all three dimensions are called quantum dots (0D). The density of states for QDs is discrete and given by [39]:

$$D_{QD}(E) = \sum_i \delta(E - E_i) \tag{1.3}$$

where E_i represent the internal energy levels of the quantum dot (see Fig. 1.3d).

Quantum wells and quantum dots are of great importance for this work. The DLTS structures contain quantum dots, while the resonant tunnel structures contain both quantum wells and quantum dots.

1.3 Two-dimensional carrier gases

The restriction of the carrier movement in spatial dimensions enables us to design the devices that contain layers with very high mobility. Hence, very fast transistors are fabricated, known as *high electron mobility transistor* (HEMT) or *modulation-doped field-effect transistor* (MODFET) [40].

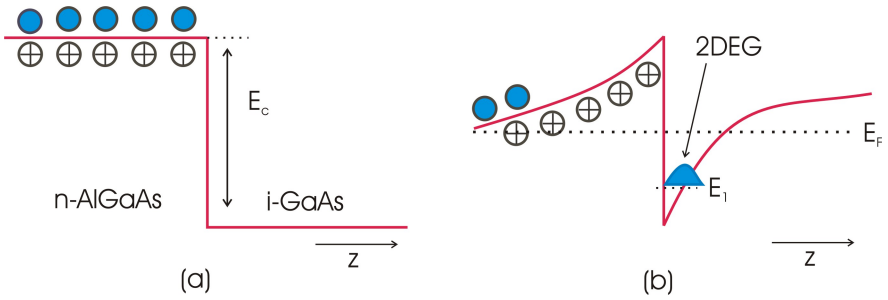

Figure 1.4: Conduction band of n-AlGaAs/GaAs heterojunction before and after the formation of two-dimensional electron gas. (a) Initially, both materials are neutral and the conduction band is flat. Then, the electrons diffuse from n-AlGaAs towards GaAs and are trapped at the interface. (b) An electric field induced by separation of charged carriers tilts the conduction band and electrons are trapped in the triangular well forming the 2DEG, since they do not have enough potential energy to overcome the barrier.

Fig. 1.4 shows the formation of a two-dimensional electron gas (2DEG) for a n-doped AlGaAs and an undoped GaAs system. They form a Type-I heterojunction

where the band gap of GaAs lies completely inside the bandgap of n-AlGaAs [25]. They are initially neutral and the conduction band is flat (see Fig. 1.4a). Then, the electrons in n-AlGaAs layer diffuse towards i-GaAs and they are trapped there since they lose their potential energy needed to overcome the barrier ΔE_C. The separation between the negatively charged electrons and the positively charged donor atoms causes an electric field that expels the electrons from interface. Hence, conduction band tilts creating a field-induced triangular well with a typical width of ∼10 nm. Electrons are trapped in this well (see Fig. 1.4b). Since the width of the well is below the De Broglie wavelength of electrons, only the quantization in z direction is relevant. In the case that only the ground state is occupied, all electrons are in the same quantum state in z-direction, while continue to freely move in x- and y-directions. A 2DEG is thus formed and the total energy for an electron occupying the energy level E_n is expressed as [39]

$$E_n(k_x, k_y) = E_n + \frac{\hbar^2 k_x^2}{2m^*} + \frac{\hbar^2 k_y^2}{2m^*} \tag{1.4}$$

where k_x and k_y are the wave vectors in x and y directions, and m^* the effective mass.

The importance of the modulation doping is that the dopant atoms are placed outside the quantum well. Normally, when the dopant atoms are introduced to the structure in order to further increase the conductance through their surplus charges, they leave ionized donor and acceptor atoms behind that cause some drawbacks such as the reduction in the mobility due to the Coulomb interaction [41], the blur of energy levels, and the disruption of the interference of carrier waves [39]. These problems are solved by separation of the active region and the dopant atoms [42].

Modulation doping can be further improved by insertion of spacer between doped layer and 2DEG yielding electron mobilities higher than 10^7 cm^2/Vs in n-AlGaAs/GaAs [43]. Formation of 2DEG with such high mobilities at low temperature led to the discovery of the Quantum Hall Effect [44, 45].

The use of p-doped AlGaAs instead of n-doped AlGaAs enables the formation of two-dimensional hole gas (2DHG) at valence band. Mobilities up to 10^6 cm^2/Vs are achieved in such structures [46].

1.3.1 Modulation-doped field-effect transistor

Through *the modulation doping technique*, modulation-doped field-effect transistors (MODFET) are fabricated with very high mobility of a two-dimensional carrier gas in their channel. Hence, MODFETs have enormous advantages over traditional MOSFETs [40].

Fig. 1.5 shows the structure of an enhancement mode AlGaAs/GaAs MODFET schematically. The epitaxial structure of the MODFET is depicted in Fig. 1.5a. It consists of an n-doped AlGaAs layer and an undoped GaAs layer. The 2DEG is formed at the interface by transferring the electrons from the n-AlGaAs layer to the i-GaAs layer. Both side of 2DEG layer are highly doped either by ion implantation or by annealing the formerly deposited dopants in order to create Ohmic contacts named *source* and *drain*. Afterwards, a Schottky gate contact is deposited on the n-AlGaAs layer which is used to alter the potential of 2DEG.

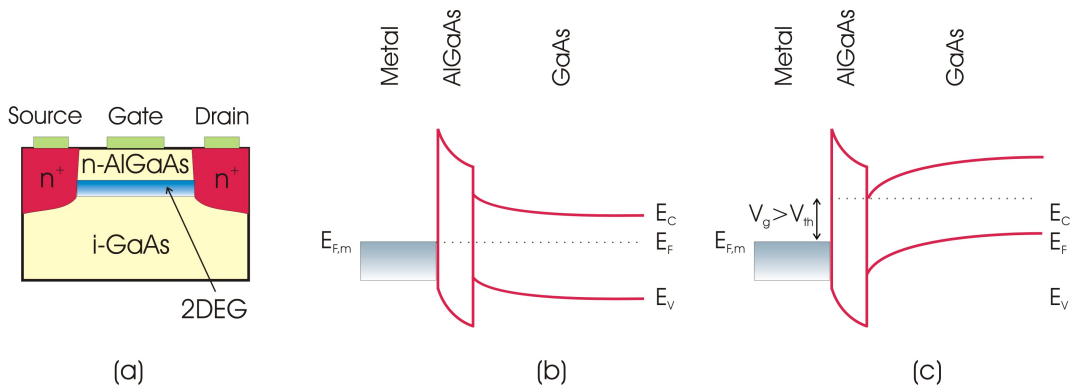

Figure 1.5: MODFET (a) Schematic cross section of an enhancement MODFET. It consist of a n-AlGaAs layer on top of a GaAs layer. (b) The band structure of the MODFET at steady state. The Fermi level is in the same position throughout the device, the channel conductance is low, and the transistor is in OFF-state. (c) A gate voltage higher than threshold voltage is applied making the diffusion of electron from n-AlGaAs to GaAs possible. The channel conductance becomes very high switching the transistor ON-state.

Fig. 1.5b depicts the enhancement MODFET in equilibrium ($V_g = 0$ V). The Fermi level lies at the same energy throughout the device ($E_F = E_{Fm}$), the channel conductance is very low, and the transistor is in OFF-state. When a gate voltage higher than the threshold voltage is applied, the conduction and the valence bands are tilted resulting in the diffusion of electrons from the n-AlGaAs to the GaAs. The accumulated electrons at the interface form a 2DEG in triangular well, as explained in the previous section. In this case, the conductance of the channel becomes very high and the transistor is switched to ON-state.

1.4 Quantum Dots (QDs)

Self-Organized Quantum Dots (QDs) [17,18] are low dimensional heterostructures that confine the motion of charge carriers in all three spatial dimensions. Due to their electronic properties similar to real atoms, they are known as *artificial atoms* [47]. These remarkable electronic and optical properties make QDs especially interesting for numerous applications in electronics such as field-effect transistor [48,49], lasers [50–53], amplifiers [54], single photon sources [55, 56], high frequency optical devices [57, 58], detectors [59, 60] and memory cells [24, 61–64].

1.4.1 Fabrication Techniques

There are mainly two approaches to grow QDs: *top-down* and *bottom-up* approaches. In the top-down method [17], a quantum well with smaller bandgap is embedded into a surrounding material with wider bandgap. Hence, a confinement for the charge carriers is realized within the growth direction. In order to create the lateral size of the QDs, a mask is defined on the surface of the sample by lithographic techniques. Then etching around the mask results in an array of QDs which provide a confinement for carriers

in all three spatial directions. However, this approach has some disadvantages such as insufficient lateral resolution and interface damage by the etching procedure.

In the bottom-up approach, the QDs are grown in a self-organized fashion [17]. All samples used in this work are grown with the bottom-up approach. Two different epitaxial growth methods are used: Molecular Beam Epitaxy (MBE) [26] and Metal-Organic Chemical Vapor Deposition (MOCVD) [27]. In MBE method, the atoms or molecules that are desired to be grown are kept in cells called *effusion cells*. Heating the cells results in a molecular beam, which is directed at the heated substrate. The atoms landing on the substrate condense and build up a thin layer. MBE growth is extremely precise thanks to the low degree of contamination granted by the ultra-high vacuum. In MOCVD the precursor gases containing the chemical element required for the growth are injected into the reactor. The precursor gases react chemically, depositing the required elements on the substrate. The byproducts are removed from the chamber by the vacuum pump.

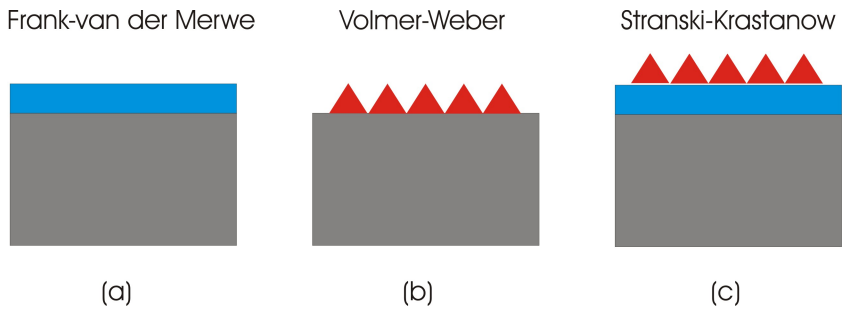

Figure 1.6: The schematic depiction of the three epitaxial growth modes: (a) Frank-van der Merwe where only wetting layer is formed, (b) Volmer-Weber where only islands are formed, and (c) Stranski-Krastanow, a mixed growth mode, where both a wetting layer and islands are formed.

In epitaxial growth different materials with different lattice constants are used (see Fig. 1.1). Different lattice constants lead to the formation of strain. As the growth proceeds, the deposited layer adjusts itself to reduce the strain, leading to coherent strain. If the strain is too large, then dislocations occur.

Coherent strain can be used to fabricate QDs. There are three modes:

- Frank-van der Merwe [65]: The lattice constant difference between the materials is small and the strain can be reduced by the creation of a wetting layer (see Fig. 1.6a).
- Volmer-Weber [66]: The lattice constant difference between the materials is large. Hence, the strain can be reduced by island formations [67] (see Fig. 1.6b).
- Stranski-Krastanow [68]: The mixed growth mode. The total energy is minimized by the formation of both a wetting layer and islands [67] (see Fig. 1.6c).

These small islands can be very regular in size and shape and can confine carriers in all the dimensions [69].

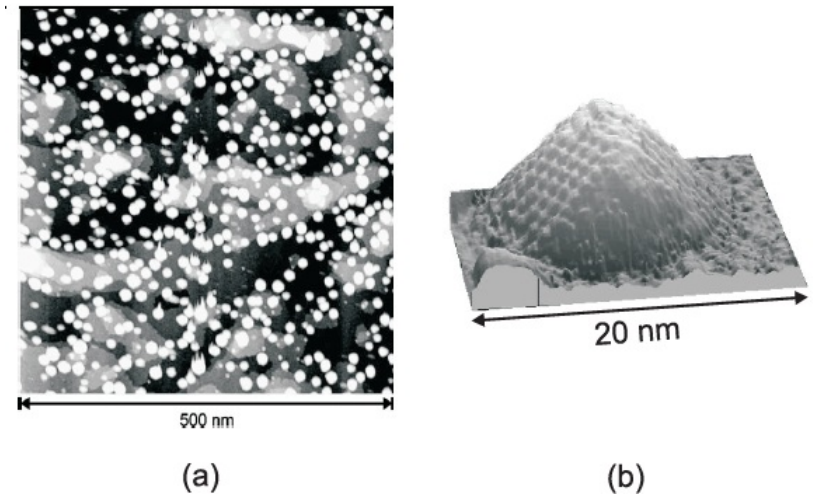

Figure 1.7: (a) STM image of InAs/GaAs QDs with an areal density of 1.9×10^{11} (from [70]). (b) High-resolution STM image of a single uncapped InAs/GaAs QDs (from [71]).

Fig. 1.7a shows a scanning tunneling microscope (STM) image of InAs/GaAs QDs with an areal density of 1.9×10^{11} cm^2, a mean height of ~ 2.2 nm, and a mean lateral dimension of 12.5 ($\pm$1) nm [70]. Fig. 1.7b shows a three-dimensional depiction of a single InAs/GaAs QD [71].

1.4.2 Electronic Properties

A convenient model for QDs is that of a three dimensional quantum potential well with infinite barrier, since they confine the motion of the carrier in all three dimensions. Hence, the QDs have discrete energy levels instead of having continuous energy spectra [72]. The Schrödinger wave equation is solved to determine the energy levels.

$$E_T = \frac{h^2 n^2}{8mL_x} + \frac{h^2 l^2}{8mL_y} + \frac{h^2 q^2}{8mL_z} \tag{1.5}$$

where E is the energy, h the Planck constant, n,l, and q integers, m the mass, and L the length of the well. Fig. 1.8 shows the band structure with internal levels for such a QD system. The distance from the ground state for electrons (holes) to the conductance (valence) band edge is called *localization energy*.

However, this simple model is not useful to determine the exact position of internal levels of the QDs, since it neglects several important parameters such as strains, piezoelectric effects, the chemical composition of the QDs and the surrounding material. It is only useful to qualitatively understand the electronic properties of the QDs. The 8-band $\mathbf{k} \cdot \mathbf{p}$ model [73–75], on the other hand, takes all mentioned parameter into consideration, yield much more precise results.

In 8-band $\mathbf{k} \cdot \mathbf{p}$ model a strain map is generated by using structural properties of the QDs and the matrix material. The calculated strain is incorporated in Hamiltonian of the system. The eigenstates of the Hamiltonian are calculated by taking into account

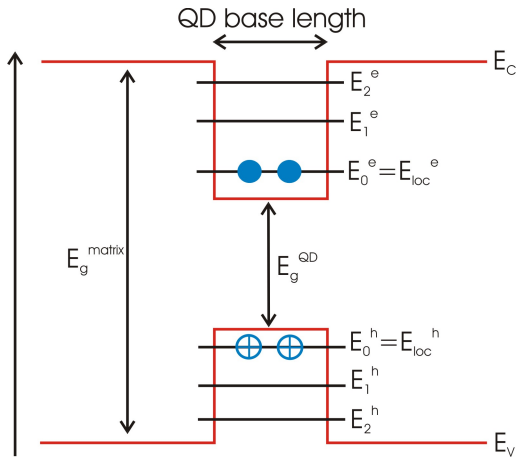

Figure 1.8: The band structure of a finite quantum dot system with the internal energy levels for both electrons and holes. The distance from the ground level state for electrons (holes) to the conductance (valence) band edge is the localization energy of the QD.

the conduction band and the valence bands (light-hole, heavy-hole, and split-off). Each band is counted twice for the electron and hole spin.

A simulation is performed on the GaSb/GaP QD system, which is used in this work, by Andre Schliwa at TU Berlin. The conduction and valence bands of the GaSb/GaP QD system with the QD base length of 9.8 nm and 24.5 nm are shown in Fig. 1.9a and Fig. 1.9b, respectively. The band configuration with strain is depicted as a red line in Fig. 1.9a, while the band configuration without strain as a blue line. The strain results in the transform of the structure from Type-I to Type-II. Hence the electron localization is prevented. On the other hand, strain causes an increase in hole localization.

Fig. 1.9c shows hole localization energy as a function of base length of QDs. It is clearly seen that an increase in base length leads to a higher localization energy. A localization energy of 1184 meV is calculated for QDs with a base length of 9.8 nm. The localization energy increases to 1400 meV with a base length of 24.5 nm. This increase in localization energy drastically increases the storage time.

Shape of QDs also affects the electronic properties of QDs such as localization energy . Based on 8 k.p calculation different localization energy values are obtained for the QDs with pyramidal shape and truncated-pyramidal shape [77].

It should be also noted that there is no only one QD in a sample, instead, high number of QDs. Hence, they have to be considered as a statistical ensemble. During the growth of self-organized QDs, different QDs in size, shape, and composition are formed due to the thermodynamic effects. These structural properties affect the electronic properties of QDs. Therefore, a Gaussian broadening of the energy levels occurs. In the case that the broadening of the energy levels is large, some energy levels can not be distinguished.

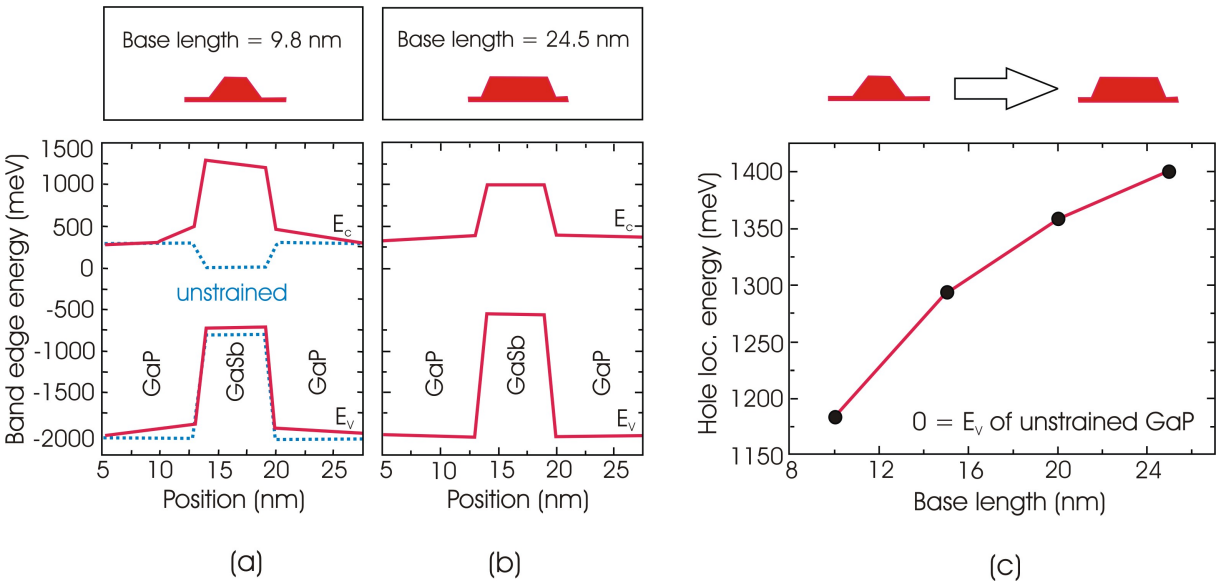

Figure 1.9: 8 k.p simulation for the GaSb/GaP QD system. (a) Band structure for QD with the base length of 9.8 nm (blue line) with strain (blue line) and without strain (red line) (b) Band structure for QD with the base length of 24.5 nm. (c) Hole localization energy versus QD base length (from [76])

1.5 Charged-based Memories

Modern information technology is based on the binary system in which all information is encoded and represented by two logic states namely *logic-1* and *logic-0* in a storage node [7,8]. These are the smallest stored information in a storage node and called *bit*. Electronic devices recognize and interpret these two cases, e.g. in a transistor logic-1 represents the *ON-state*, while logic-0 the *OFF-state*.

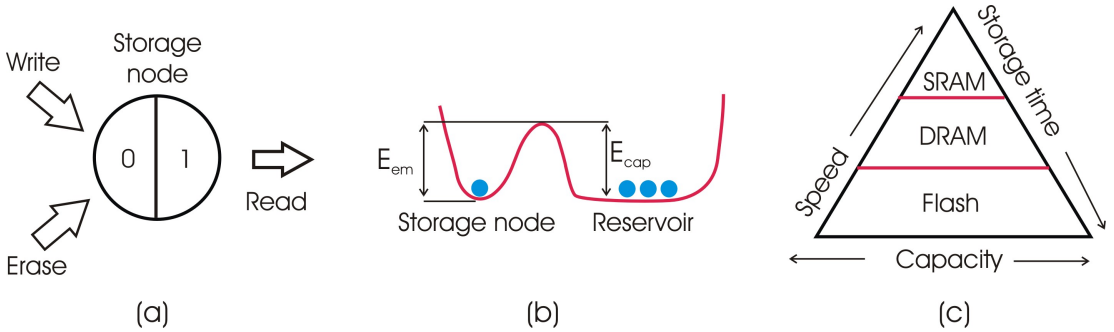

Figure 1.10: Basics of a charge-based memory (a) the storage node is the component where a single bit with the logic states 1 or 0 is stored. Also the operations such as write, erase, and read out are carried out in storage node. (b) The simple model of a storage node. A potential minimum, where electrons are confined, is separated from electron reservoir by a potential barrier. (c) The hierarchy of most used semiconductor memory devices with their advantages and disadvantages in terms of storage time, speed, and capacity

The basic operations in a memory device are write, erase, and read-out of the

stored information (see Fig. 1.10a). In order to achieve the operations, a storage node is needed. The storage node is realised by a potential minimum separated from an electron reservoir by a potential barrier (see Fig. 1.10b). The presence of an electron in the storage node represents the logic-1, while the absence of an electron in the storage node represents the logic-0. Write operation is the process that an electron in the reservoir is captured by the storage node turning the logic-0 into logic-1. Erase operation is opposite to the write operation that the electron inside the storage node leaves the storage node turning the logic-1 into logic-0. Storage operation represent the retention of the current state, either it is logic-1 or logic-0, and it is a non-equilibrium process. Therefore the performance of the operations is determined by the energetic height and the shape of the barrier.

Let us assume that an electron is stored in the storage node corresponding to the logic-1. The electron can leave the storage node with a certain probability either by thermally overcoming the barrier or by tunneling through the barrier. The probability for the electron to thermally overcome the barrier depends on the height of the barrier E_{em} according to the Boltzman's distribution (proportional to $exp(E_{em}/kT)$). The probability for the electron to tunnel through the barrier depends on the barrier's height E_{em}, shape, and width w. In order to have a long storage time, the device needs to be designed in such a way that it has a high and wide barrier. Similarly, to retain logic-0, the capture of an electron from the reservoir into the storage node has to be prevented. The capture of an electron from the reservoir into the storage node can be realised that either the electron can overcome the barrier or tunnel through the barrier. To avoid capture of an electron, a high and wide barrier is needed. On the other hand, to achieve the write and erase operations, the electron has to be captured by storage node or emitted from the storage node either by thermally overcoming the barrier or tunneling the barrier. Hence, a low and thin barrier is needed to ensure the fast write and erase operations. This dilemma is known as the *trade-off* in charge-based memory devices.

In order to circumvent the problem, different types of memories are employed [10]. The memories are designed focusing on only some of the features such as storage time, speed, capacity. The most used semiconductor memory devices are SRAM (Static Random Access Memory), DRAM (Dynamic Random Access Memory), and Flash (see Fig. 1.10c). SRAM is remarkably fast, hence it is used for CPU cache. However it is volatile and its capacity is low. On the other hand, Flash is non-volatile, hence it is useful for storage. Its disadvantage is that its access time is slow. DRAM is in between the mentioned two. It is used as working memory in computers.

1.5.1 Dynamic Random Access Memory (DRAM)

A DRAM is a volatile memory device, which means that a constant energy supply is needed to store information. It consist of a transistor and a capacitor per unit cell, i.e. 1T-1C design [7]. The transistor provides access to the capacitor, while the capacitor is used as a storage node. The schematic equivalent circuit of a single unit cell of DRAM is depicted in Fig. 1.11a. Applying a voltage to the *word line* activates the capacitor, switching it to ON-position, resulting in a connection between the capacitor and the

bit line.

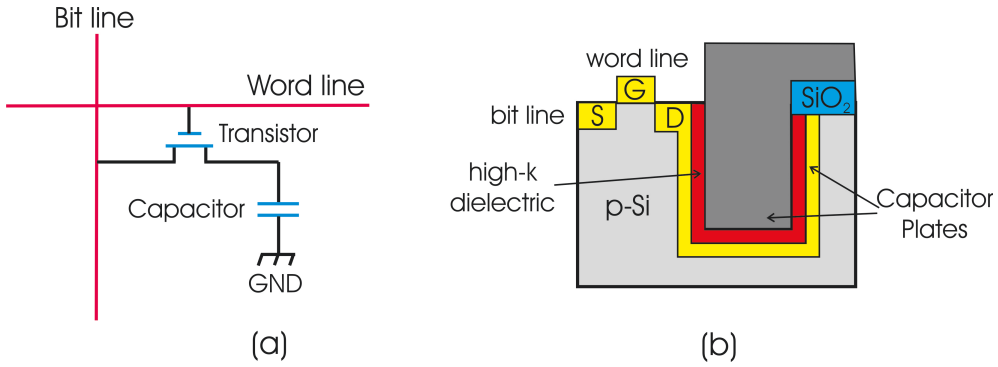

Figure 1.11: Schematic structure of a DRAM's unit cell (a) circuit representation. It consists of a transistor and a capacitor, i.e. 1T-1C design. The transistor provides access to the capacitor through the word line and the capacitor stores the information. (b) Cross section of a DRAM cell. The transistor is controlled by source, gate, and drain contacts. The capacitor has deep trench design to ensure the formation of big plates in size.

- In read operation: If the capacitor is already charged by carriers, when the connection between the capacitor and the bit line is established, carriers from the capacitor begin to flow towards the bit line, discharging the capacitor. If a voltage is measured on the bit line, a *logic-1* is read. Otherwise a *logic-0* is read.

- In write operation: After the connection between the capacitor and the bit line is established by activating the transistor, a voltage is applied to the bit line. Carriers flow to the capacitor and they charge the plate of the capacitor. Erase operation is the inverse of the write process.

The main advantages of DRAMs are their fast access time and their durability. Their access time is ∼10 ns and they can sustain up to $\sim 10^{15}$ write/erase cycles. The main disadvantage of DRAMs is that, being a volatile memories, their storage time is in the order of 10 ms. The charged capacitor loses the carriers because of the leakage. The capacitor, hence, needs to be refreshed in a certain period of time. The refreshing time specifies the storage time. In order to increase the storage time, the plate of the capacitor should be as big in size as possible ensuring to store large number of carriers. Although this method is not sufficient to provide non-volatility, it prolongs the storage time. In Fig. 1.11b the sketch of a single unit cell of a DRAM is shown. The capacitor goes very deep, called *trench cell*, in a substrate material to ensure the formation of big plates of capacitor in size.

Another critical disadvantage of DRAMs is that the reading process is destructive. As it is explained above, the capacitor releases the carriers and the reading process is achieved by comparing the voltage value of the bit line and the initial voltage value of the capacitor. Thus, the capacitor has to be charged again after reading the information. Also, large number of carriers participate to carry out the operations in DRAMs resulting in high energy consumption.

1.5.2 Flash Memory

A Flash is a non-volatile memory device that stores information for at least ten years at room temperature without needing power supply. Due to its non-volatility, speed, compactness and the energy efficiency it is suitable for numerous mobile applications such as MP3 players, digital cameras, smartphones, tablets etc [9, 10].

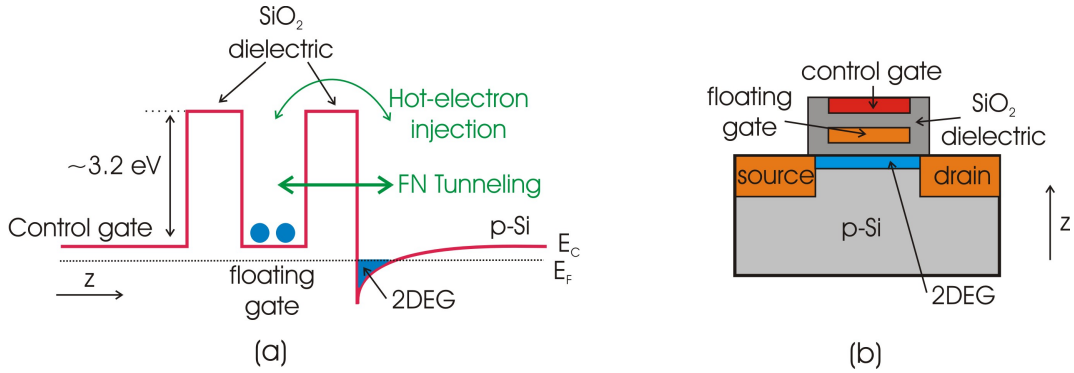

Figure 1.12: The schematic structure of a Flash cell. (a) The conduction band of the Flash cell. A floating gate used as a storage node is created by sandwiching a Si layer between two SiO_2 layers resulting in an energy barrier of 3.2 eV. The presence of carriers in floating gate corresponds to logic-0, while empty floating gate corresponds to logic-1. Write and erase operations are achieved by hot electron injection and Fowler-Nordheim tunneling. (b) The cross section of the Flash cell. It is controlled by source, gate, and the drain contacts.

The cross section of a Flash memory cell is depicted in Fig. 1.12a. It consists of a field-effect transistor with an additional gate, called *floating gate* made of Si, between the control gate and the channel. The floating gate, which is used as a storage node, is isolated by dielectric materials, i.e. SiO_2, creating an energy barrier of 3.2 eV [78] (see Fig. 1.12b).

- Read Operation: The read operation is performed by applying a voltage between source and drain, thus measuring the resistance of the channel. Charge carriers trapped in the floating gate would deplete the channel, resulting in a higher resistance (*logic-1*). A lower resistance represents the *logic-0*.

- Write/Erase Operation: In write operation, the carriers need to be injected into the floating gate (see Fig. 1.12b). By applying a positive high voltage to the control gate, the carriers in the substrate are forced to jump into floating gate by hot electron injection [79]. In erase operation, opposed to the write operation, the occupied floating gate needs to be depleted from carriers. A negative high voltage is applied to the control gate to force the carriers to leave the floating gate by Fowler-Nordheim tunneling [80].

The main advantages of the Flashes are that they are non-volatile and they consume less energy. Also, the read process is not destructive opposite to the DRAMs. They, however, have some disadvantages too. The most significant one is that their access time is slow, around ms, due to the fixed high energy barrier. The carriers have

to either overcome the barrier or tunnel through the barrier. Both processes need long time to be carried out. Also, both mechanisms, the hot electron injection and the Fowler-Nordheim tunneling, are harmful for the device causing poor endurance as they need high electric fields to be induced.

1.5.3 QD-Flash

Both DRAM and Flash have some advantages and disadvantages in terms of non-volatility, speed, durability, and charge consumption. Table. 1.1 shows the performance of the DRAM and the Flash alongside with the desired values for a memory device. Accordingly, an ideal memory device [5], i.e. the *ultimate memory device*, has to fulfill the following requirements:

	DRAM	Flash	Ultimate Memory
Storage time	30 ms	10 y	10 y
Read time	$<$ 10 ns	0.1 ms	10 ns
Write time	$<$ 10 ns	1 ms	10 ns
Erase time	$<$ 10 ns	0.1 ms	10 ns
Write/erase cycle	10^{16}	10^{5}	10^{16}
Number of carriers	100000	1000	1000

Table 1.1: The comparison of the performance of the DRAM and the Flash alongside the desired values for an ideal memory device. The DRAM has remarkably fast access time, but is volatile. The Flash is non-volatile, but has slow access time. The ultimate memory has to be non-volatile with fast access time. Additionally its durability has to be high, while consuming little energy

.

- **Storage time** has to be longer than 10 years at room temperature, as per definition of non-volatility.
- **Data access time** (write, erase, and read) has to be faster than 10 ns.
- **Durability** (writing/erasing cycle) has to be over 10^{15} times.
- **Charge consumption** (number of the carriers that participate to perform the operations) has to be less than 1000 per bit to avoid high energy consumption.

In order to combine the advantages of the DRAM and the Flash, the QD-Flash concept was developed at TUB [13–16]. QD-Flash is a quantum dot-based memory architecture that might fulfill all the requirements for the ultimate memory device. Previous works demonstrated 230 s of storage time at room temperature [19], 6 ns of write time [20], 3 ns of read time [21], and 44 ns of erase time [22]. Moreover, the full functionality of QD-Flash is demonstrated at low temperatures [22–24].

The QD-Flash relies on following principles:

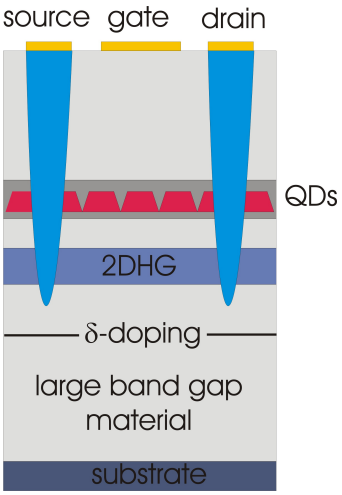

Figure 1.13: The cross-section of a QD-Flash cell. A 2DHG layer is used as a hole reservoir which is populated by the holes released by underlying δ-doping layer and it is connected to the source and the drain contacts. The quantum dot layer is placed above the 2DHG layer [24].

- III-V semiconductors are used instead of fixed Si/SiO_2 storage node. Therefore a flexible barrier is formed. The band structure can be altered by replacing one material with another, creating alloys, changing the composition of the alloys, and changing the doping of a certain layer. Hence, a flexible band engineering is possible. Moreover, III-V semiconductors are suitable to create self-organized quantum dots since they have wide range of lattice mismatch and resulting strain patterns.

- Nanostructures, specifically quantum dots, are used. The quantum dots yield high localization energy. Thus, there is no need for an additional barrier to store carriers opposed to the Flash where additional barriers are needed to store carriers in the floating gate. Besides yielding high localization energy, the QDs yield large capture cross sections that lead to fast capture time, hence fast erase time in the order of picoseconds [81, 82]. Self-organized growth generates higher quality interfaces as accumulated strain is released by the formation of islands instead of dislocations and defects.

- Holes are used as carriers instead of electrons. Therefore, larger localization energy is obtained for a given material leading to a longer storage time, as the energy levels of QDs for holes are deeply located due to their heavier effective mass in comparison to electrons.

- Modulation doped field effect transistor (MODFET) [40] is used. The QDs are embedded into a MODFET structure which provides the infrastructure for the memory operations.

The Fig. 1.13 shows the cross-section of a QD-Flash cell. The QDs are embedded into a MODFET between the gate and the 2DHG layer that acts as a channel. The

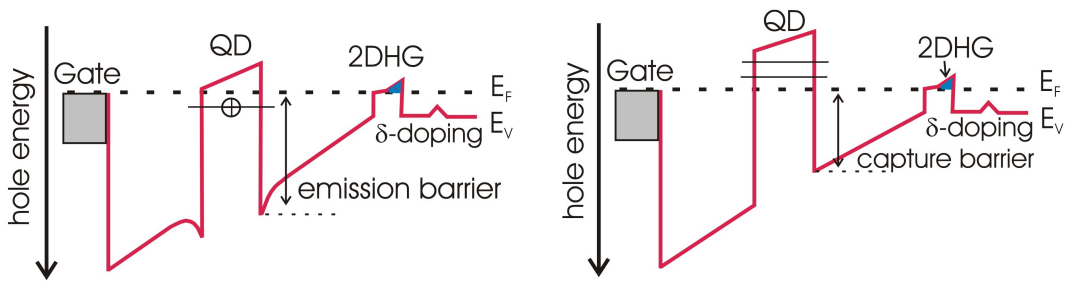

Figure 1.14: The depiction of valence band of QD-Flash in storage operation. (a) The storage of logic-1. The QD is occupied by holes and an emission barrier is needed to prevent holes from emission. (b) The storage of logic-0. The QD is empty and a capture barrier is needed to prevent holes from being captured by the QD.

δ-doping layer releases carriers into the 2DHG layer which is contacted to the source and the drain. The operations are achieved by applying voltage to the contacts.

- **Storage operation** The storage operation includes two different cases. The first case is that the QDs are filled by holes, which corresponds to the logic-1 (see Fig. 1.14a). To store logic-1, an emission barrier is needed that prevents hole emission. The emission barrier is created by band discontinuity. The second case is that there is no holes in the QDs, which corresponds to the logic-0 (see Fig. 1.14b). To store logic-0, a capture barrier is needed that prevents holes from being captured by the QD. The capture barrier is created by the band bending induced by the Schottky junction at the gate contact.

  It should be note that storage is due exclusively to band bending without any voltage application.

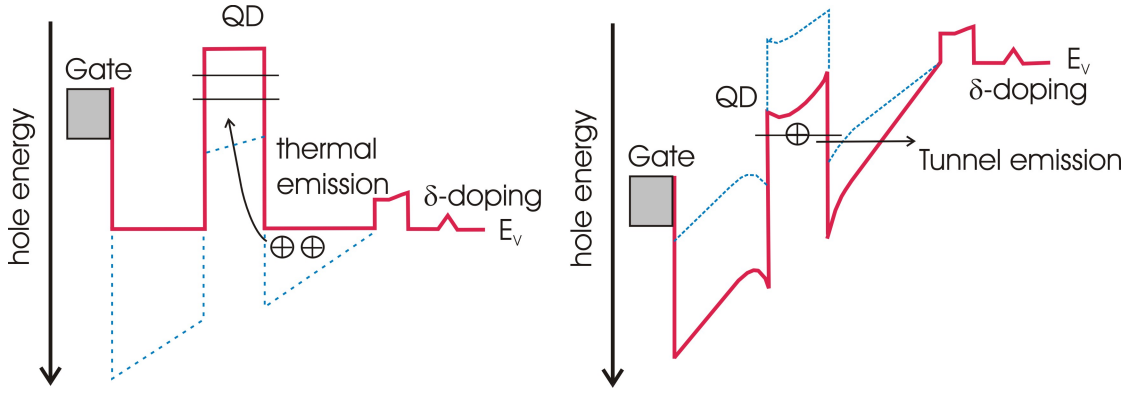

Figure 1.15: The depiction of valence band of QD-Flash in (a) write operation, where the capture barrier is completely eliminated by forward bias resulting in flatband condition, so that holes can be captured by QDs turning logic-0 into logic-1, and (b) erase operation, where the holes leave the QD by tunnelling through the narrowed barrier by reverse bias turning the logic-1 to logic-0.

The storage time has been investigated for various material systems [83–88] and

the longest value is reported to be 230 s [19]. One of the aim of this work is to improve the storage time.

- **Write operation** In write operation a forward bias is applied to the gate in order to eliminate the capture barrier (see Fig. 1.15a). The flatband condition is realized when the capture barrier is completely eliminated. The holes, then, are captured by QDs and they relax thermally down into the lowest energy level extremely fast, on the order of picoseconds [81, 82]. A write time of 6 ns is demonstrated [20], which is in the range of that of DRAM.

- **Erase operation** In erase operation, a reverse bias is applied to the gate to narrow the emission barrier increasing the tunneling probability (See Fig. 1.15b). When the emission barrier is thin enough, the holes can leave the QD by tunneling through the barrier. The fastest erase time is demonstrated [22] to be 44 ns.

  It should be, however, noted that increasing the localization energy to further increase the storage time causes an increase in the width of the barrier as well. A wider barrier reduces the tunneling probability leading to longer erase times. This problem, the trade-off between storage and erase times, has also to be considered. One of the aim of this work is to solve the mentioned problem.

- **Read operation** Read operation is achieved by conductance measurement of 2DHG layer underneath the QD [89, 90]. In the case that holes are located in QDs, they decrease the mobility and the charge density of 2DHG layer by remote impurity scattering and field effect [91] (see Fig. 1.14). The effect can be directly measured in the source/drain current and translated into the logic state. The shortest read time is demonstrated to be 3 ns [21], which is in the same order of magnitude of DRAM.

Summary

- Heterostrcutures are formed by bringing two or more semiconductor materials into a contact. Therefore, the devices that exhibit some specific electrical and optical properties can be designed for required applications. The epitaxial growth techniques such as MOCVD and MBE enable the formation of such structures.

- Low dimensional structures restrict the movement of the carriers opposed to the bulk materials in which carriers can freely move in all three spatial dimensions. The restriction of the movement of carriers arise, when the size of material is as small as the *De Broglie* wavelength of the carriers.

- The restriction of the carrier movement in spatial dimension enables us to design the devices that contain of layers with very high mobilities. Hence, very fast transistors such as HEMT and MODFETs are fabricated.

- A quantum dot (QD) is a low-dimensional structure that confine the motion of carrier in all three spatial dimensions. They are also known as *artificial atoms* due to their electronic properties similar to real atoms. Therefore, they are used for numerous applications in electronics such as field-effect transistor, lasers, amplifiers, single photon sources, and memory cells.

- Modern information technology is based on *binary system* in which all information is encoded by either *logic-1* or *logic-0* in a storage node. The storage node is formed by a potential minimum separated from a carrier reservoir by potential barrier. A long storage time is obtained by having a height and a wide barrier to avoid from emission of carriers. On the other hand, a fast write/erase operation is needed a small and a thin barrier. This dilemma is known as *trade-off*.

- The trade-off is circumvented by designing different memory devices focusing on the applications. The most extensively used semiconductor memory devices are DRAM and Flash. Both have individual advantages in terms of volatility, speed, endurance. DRAM has very fast operation time, $\sim$10 ns, but is volatile. On the other hand, Flash is non-volatile, but its operation time is considerable slow, $\sim$ms.

- The aim in memory market is to combine the advantages of both DRAM and Flash into one device, an *ultimate* memory device. QD-Flash is a promising candidate as an ultimate memory device due to its unique design.

- The storage operation is achieved relying on the band bending of the structure due to the Schottky contact. In write operation, a forward bias is applied to eliminate the capture barrier. In erase operation, a reverse bias is applied to narrow the barrier, resulting in the tunneling emission of the carriers through the barrier. Read-out operation is carried out by performing a conductance measurement of two-dimensional carrier gas.

Chapter 2

Carrier Dynamics

In this chapter, the carrier dynamics in quantum dots are explained, since operations in the QD-Flash memory such as storage, write, and erase are directly specified by the capture and emission of charge carriers from quantum dots [22,92]. Charge carriers are captured by quantum dots and trapped there for a certain amount of time depending on the localization energy of the quantum dot. Therefore, an understanding of carrier dynamics in quantum dots is essential to understand QD-Flash concept. This, also, provides better understanding of the experimental methods employed in this work.

A description to understand the carrier dynamics in deep trap levels in semiconductor is already developed [93, 94]. This concept with some modifications is applied to quantum dots to understand the carrier dynamics in the energy levels of quantum dots due to the similarities between the deep levels in semiconductors and the energy levels of quantum dots [24, 92, 95, 96].

At first, the rate equation will be introduced, which defines the occupancy of an available charge carrier state over time. Then, the capture and emission processes will be explained. Afterwards, some effects that influence the carrier dynamics in quantum dots will be discussed.

2.1 Rate equation

The occupation of an electron state over time in a quantum dots can be changed by the following four processes [94] (see top part of Fig. 2.1):

- Electron capture from the conduction band to the quantum dot: c_n
- Electron emission from the quantum dot to the conduction band: e_n
- Electron-hole pair generation (Electron capture from the valence band to the quantum dot)
- Electron-hole pair recombination (Electron emission from the quantum dot to the valence band)

Similarly, the occupation of a hole state over time in a quantum dot can be changed by following four processes (see bottom part of Fig. 2.1)

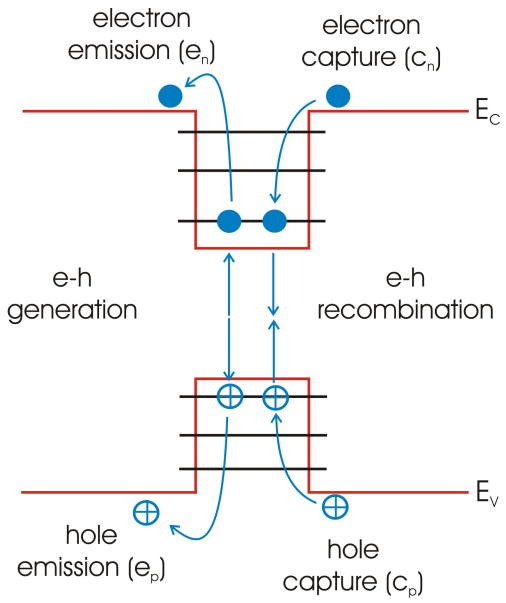

Figure 2.1: Capture and emission processes.

- Hole capture from the valence band to the quantum dot: c_p
- Hole emission from the quantum dot to the valence band: e_p
- Electron-hole pair generation (Hole capture from the conduction band to the quantum dot)
- Electron-hole pair recombination (Hole emission from the quantum dot to the conduction band)

The generation of electron-hole pair can be neglected, as the measurements in this work are performed in dark and the band gaps of the samples used in this work are large compared to kT. Moreover, the recombination of electron-hole pair can also be neglected as the majority carrier concentration is many orders of magnitude than the minority carrier concentration of samples used in this work. Therefore, only capture and emission processes will be considered in the following.

The rate equation approach is used to describe the carrier dynamics in quantum dots. The density of quantum dots whose energy levels are assumed to be same is represented by N_{QD}. $n_n(t)$ is the number of occupied electron states in QDs, while $n_p(t)$ the number of occupied hole states. In this case the number of the unoccupied electron states in QDs is written as $N_{QD} - n_n(t)$, while the number of the unoccupied hole states in QDs is written as $N_{QD} - n_p(t)$. The net emission rate for electron, which describes the change in the number of electrons emitted from QDs to conduction band, is written as follows:

$$R_n^e(t) = e_n n_n(t) \tag{2.1}$$

The net capture rate for electrons, which describes the change in the number of electrons captured from conduction band to the QDs, is expressed as follows:

$$R_n^c(t) = c_n[N_{QD} - n_n(t)] \tag{2.2}$$

These net emission and net capture rates are time-dependent. The time-independent emission and capture rates for electrons are e_n and c_n, and they represent the rate per occupied electron state and the rate per unoccupied electron state, respectively.

The net change in occupation of the electron states over time is the difference between the net capture and emission rates:

$$\frac{dn_n(t)}{dt} = R_n^c(t) - R_n^e(t) = c_n[N_{QD} - n_n(t)] - e_n n_n(t) \tag{2.3}$$

This equation is a first-order inhomogeneous differential equation. The general solution is obtained as follows with the initial condition $n_n = n_n(0)$ for $t = 0$:

$$n_n(t) = n_n(\infty) - [n_n(\infty) - n_n(0)] \cdot \exp[-(e_n + c_n)t] \tag{2.4}$$

At thermodynamic equilibrium the occupation of electron states does not change $\frac{dn_n(t)}{dt} = 0$ for $t = \infty$. The equilibrium occupation is written:

$$n_n(\infty) = N_{QD}\left(1 + \frac{e_n}{c_n}\right)^{-1} \tag{2.5}$$

Eq. 2.4 can be rewritten

$$n_n(t) = n_n(\infty) - [n_n(\infty) - n_n(0)] \cdot \exp\left(-\frac{t}{\tau}\right) \tag{2.6}$$

where the time constant τ is given by

$$\frac{1}{\tau} = e_n + c_n \tag{2.7}$$

Eq. 2.6 implies that the density of occupied electron states eventually relaxes into its equilibrium states $n_n(\infty)$ mono-exponentially with time constant introduced above, when it is changed from the equilibrium states $n_n(\infty)$ to an occupation state $n_n(0)$. Also, it should be noted that the total rate for capture and emission is given by the combination of all the emission and capture processes such as thermal, tunneling, and thermally-assisted tunnelling.

Similarly, the number of occupied holes states is given by

$$n_p(t) = n_p(\infty) - [n_p(\infty) - n_p(0)].\exp\left(-\frac{t}{\tau}\right) \tag{2.8}$$

where $1/\tau = e_p + c_p$.

2.2 Capture

The capture process represents the transition of a charge carrier from the valence or conduction band into one of the lowest available level of a quantum dot. It takes place in two steps: first, a carrier is captured into the quantum dot and then it relaxes

towards lower energy levels by dissipating its excess energy. Fig. 2.2a shows the capture mechanisms of an electron by a QD. The electron can overcome the barrier with the energy gained thermally. Another mechanism is tunneling where the electron tunnels through the barrier. The last mechanism is the combination of the mentioned two. First, the electron is thermally excited and then it tunnels through the remaining barrier. In the case that no emission barrier exist, electrons are captured on the timescale of picoseconds [81, 82].

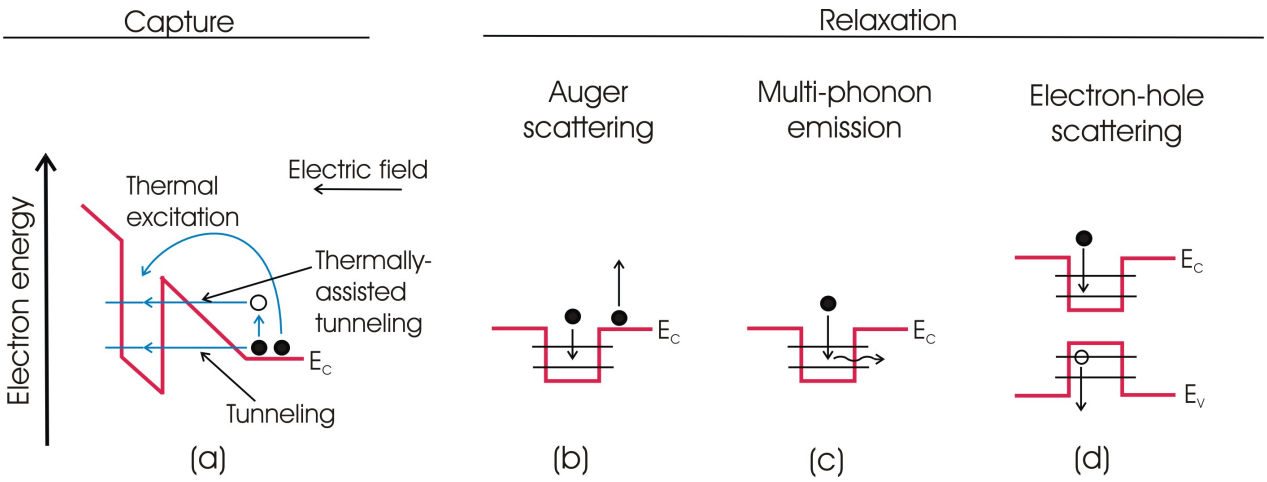

Figure 2.2: The capture and relaxation of an electron in QDs. (a) Capture mechanism is realized by thermal excitation, tunneling, and thermally-assisted tunneling. (b) Auger scattering: the excess energy is transferred to another electron in the conduction band. (c) Multi-phonon-emission: the excess energy is transferred to several LO phonons. (d) Electron-holes scattering: the excess energy is dissipated by exciting a confined hole in the QD.

After the electron is captured by the QD, it relaxes towards the lower energy levels by dissipating its excess energy by the mechanisms [81, 97]

- Auger Scattering (Electron-electron or hole-hole): A captured electron relaxes into lower energy levels by giving its excess energy to another electon in the conduction band (Fig. 2.2b).
- Multiple LO-phonon emission: A captured electron transfers its excess energy to LO-phonons resulting in the relaxation into a lower energy level (Fig. 2.2c).
- Electron-Hole scattering. A captured electron relaxes into a lower state by transferring its excess energy to a confined hole in the QD. This mechanism takes place only if the heterostructure is Type-I in which QDs are formed for both electrons and holes (Fig. 2.2d).

Capture process is described by capture rate.

$$c_n = \sigma_\infty^n < v_n > n_{QD} \tag{2.9}$$

where σ_∞^p is the capture cross section, which is an intrinsic property of the material and describes the probability that a charge carrier in the vicinity of a QD be captured

into it. $< v_p >$ is the thermal velocity and describes how fast a carrier moves due to the temperature. It is given by

$$< v_n >= \sqrt{\frac{3k_BT}{m_n^*}} \tag{2.10}$$

The net capture rate is given by the multiplication of the capture rate by unoccupied electron states of the QD as expressed in Eq. 2.9

All the mentioned capture and relaxation mechanism are analogous for holes. Accordingly, the hole capture rate is given by

$$c_p = \sigma_\infty^p < v_p > p_{QD} \tag{2.11}$$

with the thermal velocity of holes

$$< v_p >= \sqrt{\frac{3k_BT}{m_p^*}} \tag{2.12}$$

2.3 Emission

A quantum dot can localize a carrier for a certain amount of time depending on its localization energy. The localization time is the inverse of the probability of the electron being emitted either thermally or by tunneling. Fig. 2.3 shows the four possible emission mechanisms. The overall emission will be the combination of four different processes depending on physical conditions such as temperature, electric field, and photon absorption.

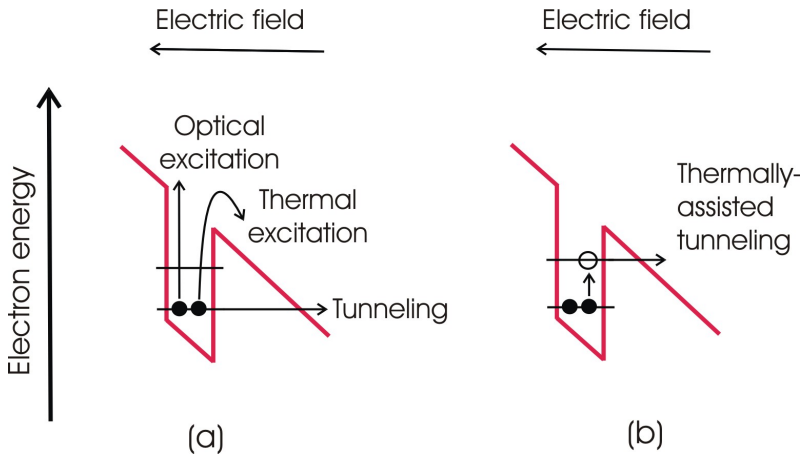

Figure 2.3: The emission processes of an electron in QDs. (a) The confined electron can leave the QD either by thermal or optical excitation above the barrier, or tunneling through the barrier. (b) The thermally-assisted tunneling is a combination of thermal and tunnel emissions. The confined carrier is first excited thermally, then it tunnels through the remaining barrier.

- Thermal emission: In this process the confined electron leaves the QD by thermal excitation above the barrier (Fig. 2.3a).
- Tunnel emission: In this process the confined electron leaves the QD by tunneling through the barrier. The tunneling probability depends on the height and the shape of the barrier. By exerting an electric field the barrier height can be altered, so that the tunneling probability can be further increased (Fig. 2.3a).
- Thermally-assisted tunneling: This process is a combination of the two emission processes above. It takes place in two steps: At first the confined electron is excited to intermediate energy level, and it subsequently tunnels through the remaining barrier (Fig. 2.3b).
- Optical emission: This process is almost identical to the thermal emission with the distinction that the confined electron gains optical energy to overcome the barrier (Fig. 2.3a). As the measurements in this work are performed in dark, this process is neglected.

As all the emission processes depend on different physical parameters, the emission rate for each process has to be expressed differently.

2.3.1 Thermal Emission

To express thermal emission rate, we cannot use an approach similar to the capture cross section, since the emission mechanism is not an intrinsic feature to the material. Instead, we need to use the rate equation.

We had already obtained an expression for the mean occupation of the electron states in equilibrium (Eq. 2.5):

$$n_n(\infty) = N_{QD}\left(1 + \frac{e_n}{c_n}\right)^{-1} \tag{2.13}$$

On the other hand, the mean occupation of the electron states in equilibrium has to be described by Fermi-Dirac distribution function [94]:

$$n_n(\infty) = N_{QD}\frac{1}{1 + \frac{g_0}{g_1}\exp\left(\frac{E_0 - E_F}{kT}\right)} \tag{2.14}$$

where E_0 is the energy level in the QD, E_F the Fermi level, and g_0 and g_1 the spin-degeneracy of the unoccupied and the occupied energy states, respectively. By equating Eq. 2.5 and Eq. 2.14, we obtain an expression for the thermal emission rate:

$$e_n^{th} = c_n\frac{g_0}{g_1}\exp\left(\frac{E_0 - E_F}{kT}\right) \tag{2.15}$$

By inserting Eq. 2.9, the expression for the capture rate, into Eq. 2.15, we obtain following expression for the thermal emission:

$$e_n^{th} = \sigma_n < v_n > n_{QD}\frac{g_0}{g_1}\exp\left(\frac{E_0 - E_F}{kT}\right) \tag{2.16}$$

Free electron density N_{QD} is approximated with the Boltzmann distribution, which is applicable for a non-degenerate n-type semiconductor ($n_{QD} << N_C$)

$$n_{QD} = N_C \exp\left(-\frac{E_C - E_F}{kT}\right) \tag{2.17}$$

By inserting Eq. 2.17 into Eq. 2.16 the expression for thermal emission rate for electrons is rewritten:

$$e_n^{th} = \sigma_n < v_n > N_C \frac{g_0}{g_1} \exp\left(-\frac{E_C - E_0}{kT}\right) \tag{2.18}$$

Similarly, the thermal emission rate for holes is written:

$$e_h^{th} = \sigma_p < v_p > N_V \frac{g_1}{g_0} \exp\left(-\frac{E_0 - E_V}{kT}\right) \tag{2.19}$$

where N_V is the effective density of states of holes at the valence band edge, E_V the energy of the valence band edge, and E_0 the energy of hole state in the quantum dots. The energetic differences $E_C - E_0$ and $E_0 - E_V$ are the emission barriers for electrons and holes, respectively, and known as *activation energies* E_a : $E_a^e = E_C - E_0$ for electrons and $E_a : E_a^h = E_0 - E_V$ for holes. Although the capture cross section of the QDs is temperature-dependent due to the fact the processes involved in carrier capture are temperature dependent [98,99], the dependence is negligible [92]. Hence, $\sigma(T)$ can be replaced with σ^∞ for $T \to \infty$. Moreover, Eq. 2.18 can be rewritten by insertion of the thermal velocity and the effective density of electrons at the conduction band edge:

$$e_n^{th}(T, E_a^e) = \gamma_n T^2 \sigma_n^\infty \exp\left(-\frac{E_a^e}{k_B T}\right) \tag{2.20}$$

where γ_n is the temperature-independent constant ($\gamma_n = \sqrt{12(2\pi)^3}\frac{k_B^2}{h^3} N_C m_n^*$). Similarly, the thermal emission rate for holes is rewritten

$$e_p^{th}(T, E_a^h) = \gamma_p T^2 \sigma_p^\infty \exp\left(-\frac{E_a^h}{k_B T}\right) \tag{2.21}$$

where γ_n is the temperature-independent constant ($\gamma_p = \sqrt{12(2\pi)^3}\frac{k_B^2}{h^3} M_V m_p^*$).

2.3.2 Tunneling

A confined carrier can be emitted by tunneling through the barrier with a certain probability. The tunneling probability, also known as *transparency*, strongly depends on the height, the width, and the shape of the barrier. Depending on the structure the barrier, and therefore its transparency, can be altered by electric field.

The samples used in this work contain quantum dots embedded in a pn junction and the effective barrier for the confined carriers has a triangular shape. The transparency for this special case is given by [100]:

$$T(E_B, F) = \exp\left[-\frac{4}{3}\sqrt{\frac{2m^*}{\hbar^2}}.\frac{E_B^{3/2}}{eF}\right] \tag{2.22}$$

where F is the electric field and E_B the height of the barrier. In order to obtain more general expression for transparency, Eq. 2.22 can be extended to barriers with any arbitrary shape by using the WKB method (the Wentzel-Kramers-Brillouin method) [39]. Accordingly, the transparency is rewritten [101]

$$T(E_B, F) \approx \exp\left[-\frac{2}{\hbar}\int_{z_0}^{z_1}\sqrt{2m^*(V(z)-E_b)}dz\right] \tag{2.23}$$

where z_0 and z_1 are the start and the finish points of the tunneling path, $V(z)$ the energetic height of the barrier. The tunneling emission rate e^{tun} for a Dirac well with a triangular emission barrier is given by [100]

$$e^{tun}(E_B, F) = \frac{eF}{4\sqrt{2m^* E_B}}\exp\left[-\frac{4}{3}\sqrt{\frac{2m^*}{\hbar^2}}.\frac{E_B^{3/2}}{eF}\right] \tag{2.24}$$

The exponential part of Eq. 2.24 represents the transparency. Eq. 2.24 is valid for any type of carriers. It can be rewritten for only electrons or holes by choosing the proper barrier height E_B and effective mass m^*

2.3.3 Thermally-assisted Tunneling

This process is a combination of thermal activation and tunneling. It takes place in two steps: First, a carrier is excited by thermal activation to an intermediate state, then it tunnels through the remaining barrier. Therefore, the emission rate of thermally-assisted tunneling depends on the temperature, the barrier height, the barrier shape, and the electric field that alters the barrier.

In order to express the emission rate of thermally-assisted tunneling, the semi-classical approach is used where the emission rate is represented by the product of the thermal emission rate and the transparency of the barrier [100].

$$e_n^{tat} = \gamma_n T^2 \sigma_n . \exp\left(-\frac{E_A - E_i}{kT}\right) . \exp\left[-\frac{4}{3}\sqrt{\frac{2m^*}{\hbar^2}}.\frac{E_i^{3/2}}{eF}\right] \tag{2.25}$$

where E_A is the initial energy level, E_i the intermediate energy level. If all possible emission paths are included, Eq. 2.25 has to be integrated over all the possible intermediate states E_i

$$e_n^{tat} = \gamma_n T^2 \sigma_n . \int_0^{E_A} \exp\left(-\frac{E_A - E_i}{kT}\right) . \exp\left[-\frac{4}{3}\sqrt{\frac{2m^*}{\hbar^2}}.\frac{E_i^{3/2}}{eF}\right] .dE_i \tag{2.26}$$

2.3.4 Total Emission

The total emission rate is expressed by the sum of all the emission processes. Although it is suggested by Vincent *et al* [100] that the total emission is given by the sum of the thermal emission and the thermally-assisted tunneling emission, calculations [102] and measurements [164] demonstrated that tunneling also contributes to the total emission

rate at low temperatures. Based on the fact, the total emission rate for electrons is given by [22]

$$e_n^{tot} = e_n^{th}(T, E_A) + e_n^{tun}(F, E_B) + e_n^{tat}(T, F, E_A) \tag{2.27}$$

Similarly, the total emission for holes is given by

$$e_p^{tot} = e_p^{th}(T, E_A) + e_p^{tun}(F, E_B) + e_p^{tat}(T, F, E_A) \tag{2.28}$$

2.4 The Effects that Influence the Emission Rates

Some important effects such as the confinement of multiple carriers in the quantum dots and the presence of an electric field have an influence on the emission rates. In this section, these effects will be explained.

2.4.1 Many-particle effects

In the case that multiple carriers are confined inside a quantum dot, the many particle problem arises, hence, the interaction between the carriers has to be taken into account.

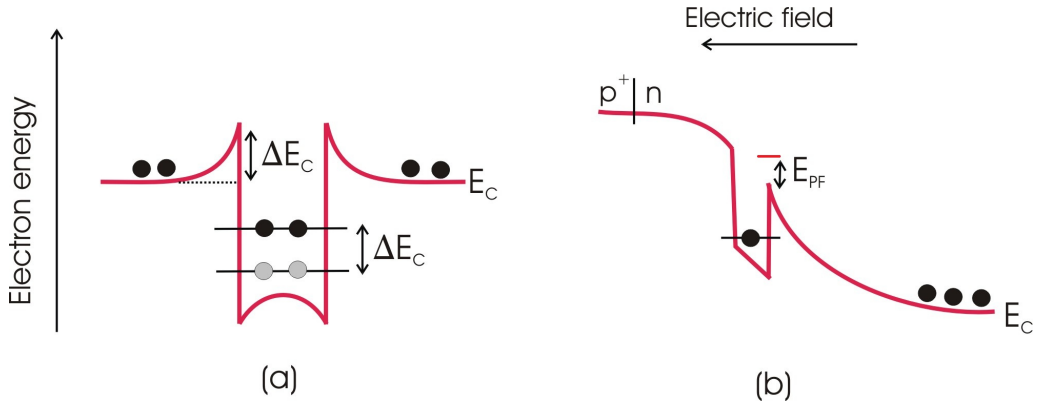

Figure 2.4: The conduction band of a quantum dot (a) Coulomb interaction due to the multiple electrons. The conduction band is increased by the same amount of that the ground states are raised due to Coulomb interaction between the electrons. (b) The Poole-Frenkel effect. A high electric field leads to a reduction in the emission barrier

Fig. 2.4a shows the effect of the Coulomb interaction for the electrons [103]. Due to the Coulomb interaction between the confined electrons inside the QD, the degeneracy of the ground states is lifted and its potential energy is increased by the amount of ΔE_C. On the other hand, the band edge of the conductance band is increased by the same amount due to the Coulomb interaction between the confined electrons inside the QD and the electrons in the conduction band. Globally, the total emission barrier remains unchanged since the increases in the band edge and the ground state are same. Hence, the Coulomb interaction cannot be seen by the DLTS experiment. Based on this fact, we neglect the Coulomb interaction in this work.

Since electrons are fermions and the fermions are not allowed to occupy the same energy level based on the Pauli exclusion principle unless their spins are opposite, the

exchange interaction has also to be considered. Both the Coulomb interaction and the exchange interaction can be calculated by 8-band $\mathbf{k} \cdot \mathbf{p}$ theory [104]. However, both effects are not detectable by DLTS, such calculations, hence, are not presented here.

2.4.2 Poole-Frenkel effect

In the presence of an electric field, the emission barrier in a quantum dot system is lowered resulting in an increase in the thermal emission rate. For such a case, the thermal emission rate is given by

$$e_n^t h(T, E_a^e) = \gamma_n T^2 \sigma_\infty^n \exp\left(-\frac{E_a^e - E_{PF}}{k_B T}\right) \tag{2.29}$$

Since the QDs are located in the vicinity of a pn junction and the operations are achieved by voltage application in the samples used in this work, the Poole-Frenkel effect [105] has to be taken into account. However, the applied electric fields in the measurements are relatively small, hence, the contribution of the Poole-Frenkel effect to the emission rate is small and can be neglected [92].

Summary

- The occupation of a state in a quantum dot is changed by the capture and the emission processes and represented by the rate equation.
- Capture process takes place in two steps: First, a carrier is captured by QDs either by thermal activation, tunneling or thermally-assisted tunneling, and then it relaxes into lower energy states by dissipating its excess energy either via Auger scattering, multiple LO-phonons, or electron-hole scattering.
- Emission is the process that confined carriers leave a quantum dots. There are four mechanisms of emission: the thermal emission in which a confined carrier overcomes the barrier by gaining thermal energy, the tunneling in which a confined carrier tunnel through the barrier, the thermally-assisted tunneling that is the combination of the mentioned two processes in which a confined carrier is thermally excited to an intermediate level and subsequently tunnel through the remaining barrier, and the optic emission in which a confined electron overcomes the barrier by gaining optical energy. The total emission is the sum of the all mentioned processes.
- Some effects influence the carrier dynamics in a quantum dot such as the Coulomb interaction due to the multiple carriers and the Poole-Frenkel effect due to the electric field. However, their effects are negligible in the scope of this work.

Chapter 3

Band Engineering

This chapter describes the resonant tunnel structures [106–109] and their usage to improve the erase time in QD-Flash by eliminating the trade-off between storage and erase times. At first the trade-off problem is introduced and then a solution is proposed, which is the insertion of a superlattice into the structure in order to create a resonance throughout the device as erase mechanism. At the resonance voltage the tunneling probability reaches the maximum value –ideally 1– resulting in a fast erase time. For this, the electronic properties of the structures of interest are calculated by One-Dimensional Schrödinger Poisson Solver [110,111] and then the transparency is calculated by Non-Equilibrium Green's function [112, 113]. Therefore, One-Dimensional Schrödinger Poisson Solver and the Non-Equilibrium Green's function are briefly explained.

3.1 The Trade-off between Storage and Erase Times

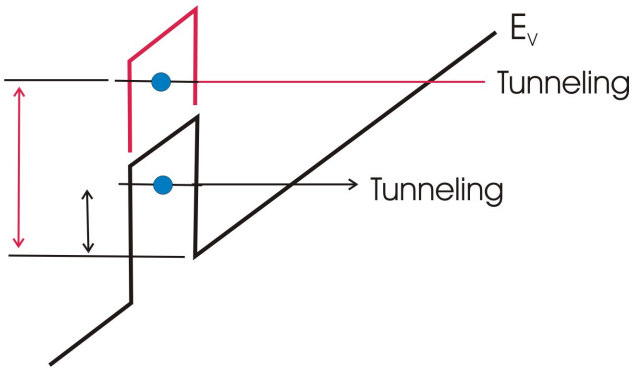

Figure 3.1: The valence band of a QD system. The black line shows a QD with a smaller localization energy, while the red line shows a QD having larger localization energy which results in reduction in the tunneling probability due to the wider barrier.

Since the desired storage time for a non-volatile memory device should be at least 10 years at room temperature [22–24], significant efforts in the QD-Flash study have

been directed to increase the storage time [83–88], which depends strongly on localization energy and capture cross section. In order to increase the storage time, either the localization energy should be increased or the capture cross section should be decreased. Due to lack of knowledge on capture cross section and its complexity, the studies focuses on the localization energy. In a QD system, an additional barrier is inserted into the structure to increase the localization energy making the total localization energy the sum of the sole QD localization energy and the height of the barrier, thus the storage time increases. However, when the height of barrier (in this case the sum) increases, the width of the barrier increases as well. A wider barrier results in a reduction in the tunneling probability leading to the slower erase time [22], since the erase mechanism depends on the tunneling (see Fig. 3.1). In other words, when the height of the barrier is increased in order to increase the storage time, the barrier width also increases causing the slower erase time. Therefore, it is concluded that a longer storage time comes with a slower erase time. This phenomenon is called *trade-off* between the storage time and erase time, and it has to be solved if QD-flash is to compete with state-of-the-art memories.

3.2 Resonant Tunnel Structures

The insertion of a superlattice into the structure is proposed to eliminate the trade-off between the storage and the erase times by creating the resonant tunneling effect at a given voltage as erase mechanism [114]. The idea behind the insertion of superlattice is to replace the thick barrier with several thinner barriers which causes a reduction in the erase time by increasing the tunneling probability. Two different states can be

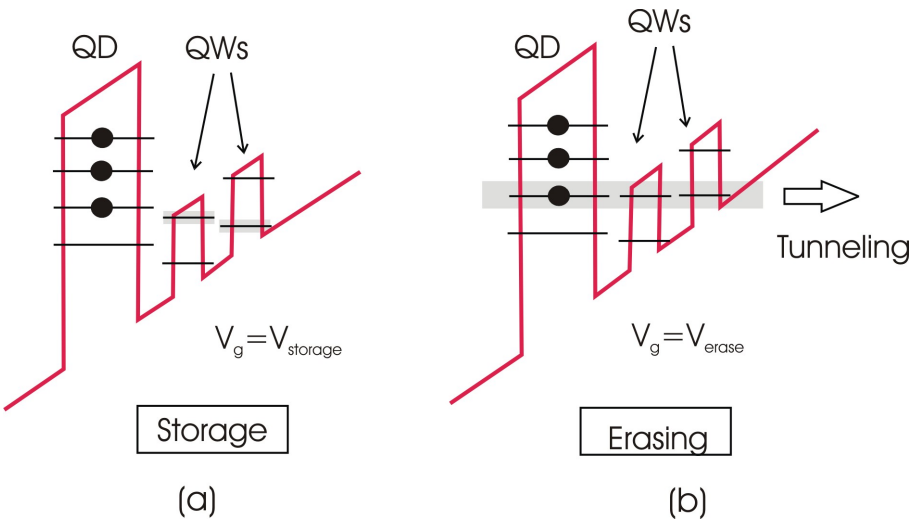

Figure 3.2: Valence band diagram of a resonant tunneling device (a) at the storage state where the energy levels of the QD and the QWs are not aligned resulting in a long storage time due to the low tunneling probability, and (b) at erasing state where the energy levels of the QD and the QWs are aligned throughout the device causing a fast erase time due to the high tunneling probability.

considered for a resonant tunneling device. Fig. 3.2 shows the valence band diagram of the structure for these two states. For the initial case, i.e. storage state, the energy levels of the QD and the QWs lie in different positions due to having different well widths resulting in the low tunneling probability –ideally 0–. Therefore, the holes stay in the QD, which provide us a longer storage time (see Fig. 3.2a). On the other hand a resonance can be created by aligning the energy levels of the QD and the QWs throughout the device at a given voltage (see Fig. 3.2b) [115]. For this, an overlap of the wavefunctions of a state on the left side of the superlattice with the state in the quantum well and a state on the right side of the superlattice is required. As far as they share at least one eigenvalue, a resonance can be created through multiple QW systems. Thanks to the resonant tunneling effect, the tunneling probability will be very high –ideally 1– resulting in the fast erase time.

To sum up, the device operates in two different states: at the storage state the device yields a long storage time due to the low transparency as a result of the absence of the alignment of the energy levels throughout the device, and at the erase state the device yields a fast erase time due to the high transparency as a result of the resonant tunneling effect.

Resonant tunneling structures were first used in resonant tunneling diodes [108, 116]. Transparency engineering in opto-electronic devices was successfully performed in the development of Quantum Cascade Lasers by Faist *et al* [117].

3.3 Band Engineering

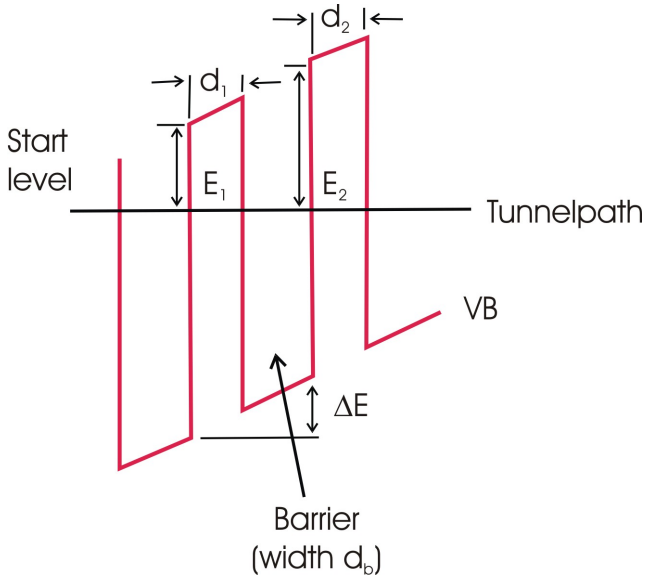

Figure 3.3: The recursive rule enables us to calculate the widths of arbitrary number of QWs.

Band Engineering is needed to design the resonant tunneling structures. The resonance through the multiple QW systems depends on well number, well width, and

barrier height [118–120]. The resonant tunneling structures are designed by altering the mentioned parameters .

A *simple model* might be helpful to understand the resonance through the multiple QW system. For simplicity the QWs are considered as infinitely deep one-dimensional quantum wells. The energy of the ground state of a QW is given by

$$E = \frac{h^2}{8m^* w^2} \tag{3.1}$$

where h is the Planck constant, m^* the effective mass for holes, and w the width of the well. After knowing the ground state of the first QW, we can calculate the width of the second QW whose ground state lies in the same position as the first QW's ground state does. For this, we use the recursive rule. We consider two QWs with the widths w_1 and w_2, and the energy levels E_1 and E_2 separated by a barrier whose width w_b (see Fig. 3.3). When a voltage is applied, it results in a band bending BB expressed by

$$BB = \frac{\Delta E}{w_1 + w_2} \tag{3.2}$$

where $\Delta E = E_2 - E_1$. The resonance condition for the second QD are obtained using $E_2 = E_1 + \Delta E$, Eq.3.1, and Eq.3.2.

$$w_2 = \sqrt{\frac{h^2}{8m^*(E_1 + BB(w_1 + w_b))}} \tag{3.3}$$

By extending this formula the widths of an arbitrary number of the QWs that provide resonance can be calculated.

Unfortunately this simple model is not able to calculate the exact positions of the energy levels due to its simplifications such as that QWs have infinite barriers. However, it gives an insight to grasp an idea of how resonances are induced and which parameters have to be altered in order to design the suitable structures. For the accurate calculation, we chose to represent the evolution of the energy levels of the quantum wells, obtained by diagonalization of the Hamiltonian, as a function of the applied voltage. When the energy levels cross, the resonance occurs.

3.4 Calculation Cycle

The final aim of the calculation is to show the decrease in erase time at resonance voltage. Therefore, the transparency on which the erase time depends has to be calculated. The calculation cycle is shown in Fig. 3.4. As first step, a resonant tunneling structure is proposed and the simple model gives an insight for that. Then the electronic properties of the suggested structure such as the valance and conduction bands, the electron and hole concentrations are calculated at different gate voltages by one-Dimensional Schrödinger Poisson Solver (1DSPS). Afterwards the active region, i.e. the tunneling resonance region, is extracted from the whole device in order to save the computation time by reducing the number of the dash points. Finally, the transparency is calculated by Non-Equilibrium Green's Function (NEGF). Since the erase time is inversely

proportional to the tunneling probability, the increase in the transmission corresponds to the decrease in the erase time in the same order of magnitude.

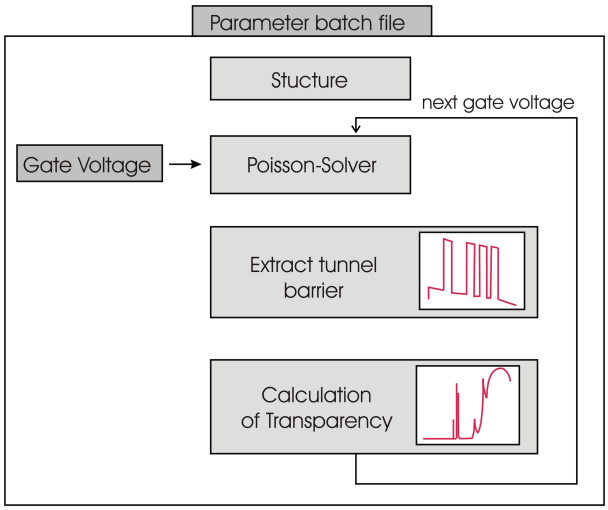

Figure 3.4: The calculation cycle followed in this work. After calculation of the electronic properties of the structure at different gate voltage by one-Dimensional Schrödinger Poisson Solver, the transparency is calculated by Non-Equilibrium Green's Function. The increase in the transparency corresponds to the decrease in erase time.

3.5 One-Dimensional Schrödinger-Poisson Solver (1DSPS)

The electronic properties of the suggested structure such as the conduction and valence bands, and the hole and electron concentrations are calculated by One-Dimensional Schrödinger Poisson Solver [121, 122].

The one dimensional Schrödinger equation is expressed:

$$-\frac{\hbar^2}{2}\frac{d}{dx}\left(\frac{1}{m^*(x)}\frac{d}{dx}\right)\psi(x) + V(x)\psi(x) = E\psi(x) \tag{3.4}$$

where ψ is the wavefunction, E the energy, V the potential energy, $\hbar$ the Planck's constant divided by 2π, and m^* the effective mass. The one-dimensional Poisson equation is written:

$$\frac{d}{dx}\left(\epsilon_s(x)\frac{d}{dx}\right)\phi(x) = \frac{-q[N_D(x) - n(x)]}{\epsilon_0} \tag{3.5}$$

where ϵ_s is the dielectric constant, ϕ the electrostatic potential, N_D the ionized donor concentration, and n the electron density distribution.

In order to solve these two equations, we should know the potential energy V and the electron concentration n. The potential energy V is given by

$$V(x) = -q\phi(x) + \Delta E_c(x) \tag{3.6}$$

where ΔE_c is the pseudopotential energy due to the band offset at the heterointerface. The electron density n is written as:

$$n(x) = \sum_{k=1}^{m} \psi_k^*(x)\psi_k(x)n_k \tag{3.7}$$

where m is the number of the bound states, and n_k the electron occupation for each state. It is expressed following:

$$n_k = \frac{m^*}{\pi\hbar^2} \int_{E_k}^{\infty} \frac{1}{1 + exp^{(E-E_F)/kT}} dE \tag{3.8}$$

where the E_k is the eigenenergy [121, 122].

The only way to solve the Schrödinger and the Poisson equations is to employ an iteration procedure, hence we can obtain a self-consistent solution. We have to start choosing a trial potential. Thus we can solve the wave equation using Eq. 3.4. Afterwards we can calculate the electron density using Eq. 3.7 and Eq. 3.8. Then, we calculate the electrostatic potential using Eq. 3.5 . Now we can calculate a new potential using Eq. 3.6. After the calculation is done, we compare the initial potential energy and the calculated potential energy. If they are consistent, the calculated potential energy is considered as the real potential energy. Otherwise, the initial potential energy is modified and the mentioned steps are repeated until obtaining a consistent potential energy [121–123].

The numerical solution of the Schrödinger equation has been obtained by *the finite difference method* (FDM) [121, 124]. In this method, the real space is divided into discrete mesh points and the wavefunction is solved within these discrete spacings. For the Poisson Equation, the numerical calculation is carried out by Newton Equation [121].

In this work, a software called *Band Diagram Calculator* by Gregory Snider is used to calculate the band diagram of the interested structures and the carrier concentrations [125].

3.6 Non-Equilibrium Green's Function (NEGF)

The transparency is calculated by Non-Equilibrium Green's Function (NEGF). Fig. 3.5 shows the model used for the transparency calculation [126]. It is assumed that a channel is sandwiched between two carrier reservoirs, i.e. contacts, labelled as *source* and *drain*, respectively. Each contact seeks to establish an equilibrium with the channel. The source pumps carriers into the channel, while the drain pulls the carriers out. However, no equilibrium is established throughout the device due to the differences in the actions of the contacts. At steady state, two different local equilibriums arise with two different electrochemical potentials, namely μ_1 and μ_2, attributed to the source and the drain respectively. Thus, two distinct Fermi functions exist [127]

$$f_1 \equiv f_0(E - \mu_1) = \frac{1}{exp[(E - \mu_1)/k_B T] + 1} \tag{3.9}$$

$$f_2 \equiv f_0(E-\mu_2) = \frac{1}{exp[(E-\mu_2)/k_BT]+1} \tag{3.10}$$

The source seeks to equalize the number of electrons occupying the level to $f_1(\epsilon)$ while the drain tries to equalize this number to $f_2(\epsilon)$. The actual steady-state number of electrons N lies between the two. The different actions of these two contact results in a current flows in the external circuit.

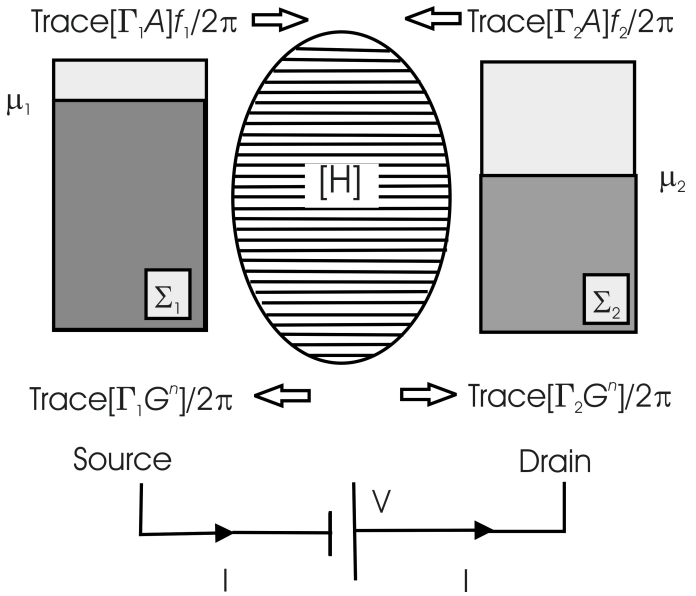

Figure 3.5: The model used for the transparency calculation. It is assumed that a channel is sandwiched between two contacts named source and drain respectively, which act as carrier reservoirs. Each contact seeks to establish an equilibrium with channel. The source pumps carriers into the channel, while the drain pulls the carriers out. The number of the electrons transmitted from the source to the drain is calculated by NEGF.

The NEGF formalism enables us to calculate the number of the carriers that flow from the source to the drain. What we need is to know the Hamiltonian of the channel $[H]$ and its coupling to the contacts described by self-energy matrices $\Sigma_{1,2}$ [128]. After describing the mentioned parameters it is straightforward to write the Green's function of the system, which is expressed:

$$G = [EI - H - \Sigma_1 - \Sigma_2]^{-1} \tag{3.11}$$

where E is the carrier energy, I the identity matrix, H the Hamiltonian, and $\Sigma_{1,2}$ the self-energy matrices of the contacts.

The transparency is easily calculated by Green's function:

$$T = Trace[\Gamma_1 G \Gamma_2 G^+] = Trace[\Gamma_2 G \Gamma_1 G^+] \tag{3.12}$$

where $\Gamma_{1,2}$ are the broadening matrices that define the broadening of the energy levels of the channel due to the coupling between the channel and the contacts [126]. They are expressed:

$$\Gamma_{1,2} = i[\Sigma_{1,2} - \Sigma_{1,2}^+] \tag{3.13}$$

The external current from the source to the drain can be expressed by using the transparency and the Fermi functions of the contacts as follows:

$$I_i = (q/h) \int_{-\infty}^{+\infty} dET(E)(f_1(E) - f_2(E)) \tag{3.14}$$

It is convenient to use this model to investigate the resonance for QD-Flash memory, since the structures used in this work contain one QW (or more based on the structure) sandwiched between a QD and a 2DHG layer that act as reservoirs. In the erase process, holes in the QD have to leave the QD for 2DHG layer through the QW, and vice versa in the write process. What we do is to find the Hamiltonian of the QW (channel) and its coupling to the QD and the 2DHG layer (the source and the drain), then calculate the transparency. The increase in the transparency is translated into the decrease in the erase time.

The transparency calculations are carried out by MATLAB. The MATLAB codes used in the work are adapted from Supriyo Datta's book [126].

3.6.1 Hamiltonian of the Channel

The calculation of the Hamiltonian of the system is straightforward: it is the sum of the free-particle Hamiltonian and the potential profile U calculated by the Poisson Solver. With the effective mass m^* of the holes, it is obtained

$$H = -\frac{\hbar^2}{2m^*}\frac{\delta^2}{\delta x^2} + U \tag{3.15}$$

In Green's function calculation, using matrix representation rather than differential equations is more convenient. The finite difference method (FDM) [121, 124] is used to convert the differential equation into a matrix equation. The discrete lattice approach provides following conversions:

$$\frac{\delta^2\psi}{\delta x^2} \rightarrow \frac{1}{a^2}[\psi(x_{n+1}) - 2\psi(x_n) + \psi(x_{n-1})] \tag{3.16}$$

$$U(x)\psi(x) \rightarrow U(x_n)\psi(x_n) \tag{3.17}$$

Accordingly the Hamiltonian is written by

$$[H_{op}\psi]_{x=x_n} = (U + 2t_0)\psi_n - t_0\psi_{n-1} - t_0\psi_{n+1} \tag{3.18}$$

where $t_0 = \hbar^2/2m^*a^2$. Finally, we obtain a matrix for the Hamiltonian given by

$$H = \begin{bmatrix} 2t_0 + U_1 & -t_0 & \dots & 0 & 0 \\ -t_0 & 2t_0 + U_2 & -t_0 & \dots & 0 \\ \dots & -t_0 & 2t_0 + U_{N/2} & -t_0 & \dots \\ 0 & \dots & -t_0 & 2t_0 + U_{N-1} & -t_0 \\ 0 & 0 & \dots & -t_0 & 2t_0 + U_N \end{bmatrix}$$

The potential is obtained through the One-Dimensional Schrödinger Poisson Solver and it is actually the valence band of the structure. What we need to do in order to set the Hamiltonian matrix up is to choose only a proper lattice spacing a, the rest of the parameters in t_0 are constants. As far as the lattice spacing a is smaller than the spatial extension of the wave function, the Hamiltonian matrix yields accurate solutions [126, 129].

3.6.2 Self Energy Matrices

For a contact between continuum and an element with a discrete energy spectrum, the self-energy matrix is

$$\Sigma(E) = \sum_j \frac{|\tau_j|^2}{E - \varepsilon_j + i0^+} \tag{3.19}$$

where $|\tau_j|$ coupling energy between the continuum and the energy level ε_j and 0^+ a regularization constant [126]. Applied to the QD/superlattice contact, where the QD is considered as a one-level element with the energy ϵ_{QD}, it becomes

$$\Sigma_{QD}(E) = \sum_j \frac{|\tau_j|^2}{E - \varepsilon_{QD} + i0^+} \tag{3.20}$$

The self-energy of the 2DHG/superlattice contact is obtained by integration over all the states of the 2DHG. Assuming only one subband to be occupied and using the two-dimensional density of states $D_{2D} = m^*/(\pi\hbar^2)$ we get:

$$\Sigma_{2DHG} = \int D_{2D} \frac{|\tau_\varepsilon|^2}{E - \varepsilon + i0^+} \tag{3.21}$$

Assuming τ_ε independent of the energy ε, we obtain the imaginary part of the self-energy:

$$\xi(\Sigma_{2DHG}) = -2\pi D_{2D}|\tau|^2 \tag{3.22}$$

The integration process for the real part does not converge, but its effect on the transparency factor is negligible such that the following approximate expression will be used: $\Sigma_{2DHG} \approx i\xi(\Sigma_{2DHG})$. Unfortunately, the values of the coupling energies are unknown, but we can tolerate this problem as far as we are interested only in the relative variations of the transparency (i.e. between a case with very low transparency and a case with high transparency) and not the absolute values. In the simulations we take $\tau_{QD} = \tau_{2DHG} = 0.05eV$.

Summary

- We increase the localization energy of the QD system by inserting an additional barrier in order to increase the storage time. However, the additional barrier causes an increase in the width of the barrier beside its height, which results in reduction in tunneling probability on which erase time strongly depends. In short, a longer storage time comes with a slower erase time.

- In order to eliminate the trade-off between storage and erase times, we suggested the use of superlattice to create a resonance throughout the device as erase mechanism. At the resonance the tunneling probability will be high due to the alignment of the energy levels of the superlattice and the contacts resulting in a fast erase time. Out of the resonance voltage, the tunneling probability will be low –ideally 0– since the energy levels of the superlattice and the contacts are not aligned resulting in a long storage time.

- In order to design such device, band engineering is needed. The energy levels of the superlattice are adjusted by altering the number of wells, the width of the wells, and the height of the barrier.

- The electronic properties of the resonant structures such as conduction and valence bands, and the electron and hole concentrations are calculated by One-Dimensional Schrödinger Poisson Solver.

- The transparency calculation is carried out by Non-Equilibrium Green's Function. In the model it is assumed that a channel is sandwiched between two carrier reservoirs whose actions are different. The first contact named as the source seeks to establish an equilibrium with the channel by pumping carriers in the channel, while the second contact named as the drain seeks to establish an equilibrium with the channel by pulling the carriers out. The difference in the actions results in an external current throughout the device.

- At the resonance voltage, the transparency is expected to be higher than the one out of the resonance. Since the erase time inversely proportional to the transparency, the increase in the transparency is translated into the decrease in erase time with the same order of magnitude. A comparison between the structures with and without superlattice demonstrates the improvement of erase time.

Chapter 4

Experimental Methods

In this chapter the measurement techniques used in the work will be described. In this work, capacitance-voltage spectroscopy and time-resolved capacitance measurements are employed in order to investigate the electronic properties of QDs. Therefore, the capacitance-voltage measurement and its applications will be presented, then time-resolved capacitance measurement used to determine the localitazion energies and capture cross sections of QDs will be explained.

4.1 Static Capacitance Spectroscopy

Capacitance-voltage spectroscopy is used to investigate the electronic properties of semiconductors [40, 94, 130]. In the case that QDs are embedded into a pn junction, the QDs contribute to the capacitance of the device, as QDs have energy levels which charge carriers can be captured in and emitted from. Therefore the electronic properties of QDs can be examined via the static capacitance measurement [131].

4.1.1 pn Junctions

Intrinsic semiconductors conduct electricity either very poorly or not at all, depending on the temperature, and the number of electrons and holes in them are equal. The conductance of a semiconductor can be increased by doping. Depending on the dopant type n- or p-type semiconductors are formed. If the dopant is donor, n-type semiconductor material is formed whose majority carriers are electron, while holes are minority carriers. Similarly, p-type semiconductor is formed, where holes are majority carriers, if the dopant is acceptor. Essential microelectronic components such as diodes, transistors are formed by combining doped semiconductors in a certain sequence. The simplest device is the pn junction, also known as diode, and it is formed by bringing a p-type and an n-type semiconductor into electrical contact [40].

Depletion Region

When a pn junction is formed by bringing a p-type and an n-type semiconductors into electric contact, electrons from the n-side diffuse towards p-side due to the difference in

the carrier concentration on either side of the junction. When electrons move to p-side, they leave behind uncompensated ionized donor atoms and they lead to a positively charged region inside the n-side of the junction. Similarly, holes from p-side diffuse towards n-side leaving ionized uncompensated acceptor atoms behind. Acceptor atoms cause a negatively charged region inside p-side of the junction. An electric field is induced by the oppositely charged region that leads to a drift current on either side. Consequently, a region depleted from free carriers is formed, called *depletion region* or *free-charge region*.

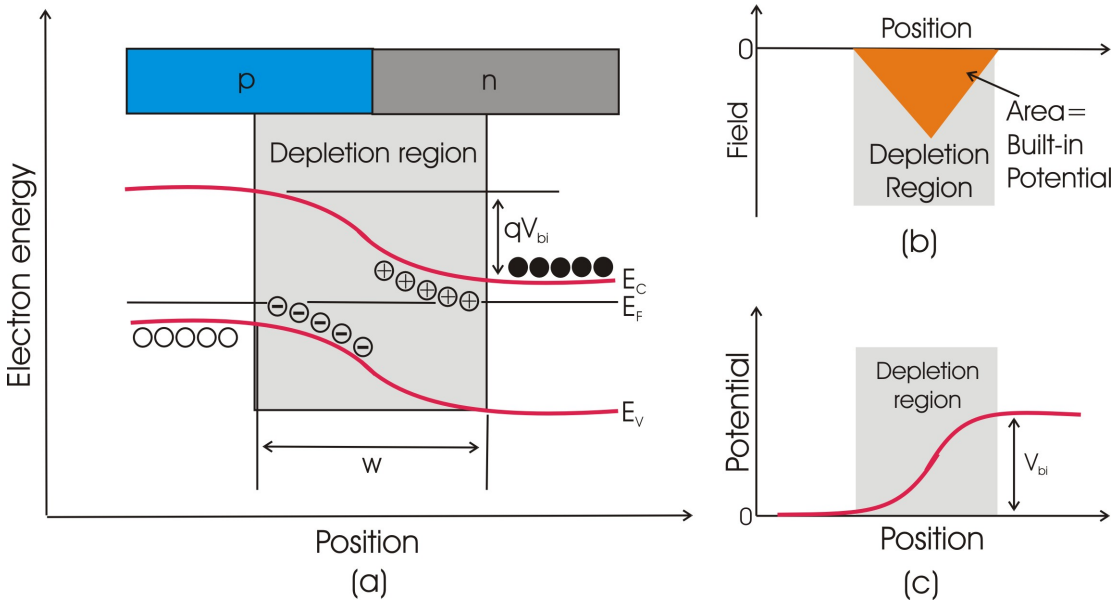

Figure 4.1: (a) Schematic representation of the band structure of a pn junction. A depletion region extends to $-z_p$ on the p side and to $-z_n$ on the n side. A built-in voltage is formed at equilibrium where the Fermi level is aligned throughout the device. (b) Density of charge on either side. (c) Electric field in the depletion region.

In thermal equilibrium the diffusion and drift currents are equal and the Fermi level is aligned throughout the whole device resulting in a band bending of the conduction and valence bands. A potential barrier V_{bi} for free carriers is formed as a result of the band bending (see Fig. 4.1) and it is given by [40]

$$V_{bi} = \frac{k_B T}{q} \ln\left(\frac{N_A N_D}{n_i^2}\right) \tag{4.1}$$

where N_A is the number of acceptor atoms on the p-side, N_D the number of the donor atoms on n-side, and n_i the intrinsic carrier concentration.

If a forward bias is applied, the depletion region shrinks and V_{bi} is reduced by the amount of the applied voltage. When the barrier is completely eliminated, the flat-band condition is achieved and carriers can freely move throughout the device. If a reverse bias is applied, the depletion region becomes wider, V_{bi} increases and carriers cannot flow. Therefore a pn junction exhibits rectifying behavior.

The width of the depletion region is given by [40]

$$w = \sqrt{\frac{2\epsilon_s}{q}\left(\frac{N_A + N_D}{N_A N_D}\right) V_{bi}} \tag{4.2}$$

where ϵ_s is the permittivity of semiconductor material, q the elementary charge, V_{bi} the build-in potential.

For some applications, semiconductor materials might be differently doped. For instance all the samples used in this work have pn junctions whose n-side are highly doped in comparison to the their p-side (several orders of magnitude). Such structures are referred as n^+p junction where acceptor concentration is several orders of magnitude than donor concentration, and as p^+n in the opposite case. The depletion region, then, extends towards the lowly doped side of the junction similar to a metal-semiconductor junction. The depletion region width can be modified for such structures as follows:

$$w = \sqrt{\frac{2\epsilon_s(V_{bi} - V_{ext})}{qN_B}} \tag{4.3}$$

where N_B is the ionized atoms concentration on the lowly-doped side and V_{ext} the external bias.

Capacitance of a pn junction

The presence of spatially-separated fixed ionized donors and acceptors in a pn junction generates a capacitance. The expression for the differential capacitance per unit area $C = dQ/dV$ has to be used due to the voltage dependence of the depletion region width. In our case, dQ is the the incremental change in depletion region charge per unit area, while dV is the incremental change in the applied voltage.

When the voltage is increased by an amount of dV, the charge on either side of depletion region increases too by an amount of dQ. The incremental space charge on both sides of the depletion region is the same with opposite charge polarities. Therefore, charge neutrality is maintained. An increase in charge by dQ causes an increase in electric field by the amount $C = dQ/\epsilon_s$, which is obtained from Poisson's equation. Incremental voltage change can be written as $dV = W \cdot dE = W \cdot dQ/\epsilon_s$. The capacitance of the depletion region of a pn junction is obtained by substituting this expression for dV into the definition of capacitance [40]:

$$C = \frac{dQ}{dV} = \frac{dQ}{w\frac{dQ}{\epsilon_s}} = \frac{\epsilon_s}{w} \tag{4.4}$$

This expression corresponds to the capacitance of a parallel-plate capacitor where the spacing between the two plates is the depletion layer width. The expression for a pn junction can be obtained by plugging depletion region equation (Eq. 4.2) into Eq. 4.4:

$$C = \sqrt{\frac{\epsilon_s}{2(V_{bi} - V_{ext}} \cdot \frac{N_A N_D}{N_A + N_D}} \tag{4.5}$$

This equation is rewritten for the pn junctions where the doping is much higher on one side than on the other side.

$$C = \frac{\epsilon_s}{w} = \sqrt{\frac{q\epsilon_s N_B}{2(V_{bi} - V_{ext})}} \tag{4.6}$$

Static Capacitance spectroscopy

Static capacitance spectroscopy can be used to determine the doping concentration of semiconductors. The doping concentration can be obtained from the measured capacitance as a function of voltage with the assumption that all doping atoms are ionized, which is reasonable at room temperature. The charge carrier fluctuation induced by the alteration of the external bias occurs at the edge of the depletion region of the pn junction in the Schottky approximation. In an abrupt n^+p junction doping concentration on the n-side is much higher than on the p-side resulting the formation of the depletion region only on p-side, which is a good approximation for the structures used in this work. Therefore, the doping concentration can be extracted from Eq. 4.6 by plotting $1/C^2$ versus V_{ext}:

$$\frac{1}{C(V_{ext})^2} = \frac{2(V_{bi} - V_{ext})}{q\epsilon_0\epsilon_r A^2 N_B} \tag{4.7}$$

If the doping concentration is constant, the plot of $1/C^2$ over V_{ext} will be a straight line. Therefore, the doping concentration can be acquired from its slope, while the built-in voltage is derived from the intercept point with the x-axis. If the doping concentration is not constant throughout the device, then $1/C^2$ must be expressed in differential form with respect to the V_{ext}. The doping profile becomes

$$N_B(z) = -\frac{1}{q\epsilon_r\epsilon_0 A^2}\left[\frac{1}{\frac{d}{dVext}\left(\frac{1}{C^2}\right)}\right] \tag{4.8}$$

Quantum Dots inside a pn junction

If quantum dots are embedded in a pn junction, carriers in the QDs change the capacitance of the pn junction by compensating the space-charge induced by ionized dopants. Carriers in the QDs change the electric field resulting in a change of depletion region width, hence in the capacitance of the junction.

In Fig. 4.2 a n^+p junction in which a quantum dot layer is embedded is depicted. In such a structure, the depletion region exists on p-side of the junction and the majority carriers are holes. The left-hand side of the figure portrays the case of unoccupied QDs, while the right-hand side that of occupied QDs. Fig. 4.2a shows a sketch of the junction. The formation of the depletion region on the p-side of the junction is clearly seen. Fig. 4.2b shows the charge distribution. In the case that holes are located in the QD, an extra charge distribution appears inside the depletion region. This charge distribution leads to a reduction in electric field by compensating it (see Fig. 4.2c). The decrease in electric field results in the extension of depletion region (see Fig. 4.2d). Finally, the capacitance of the junction is decreased.

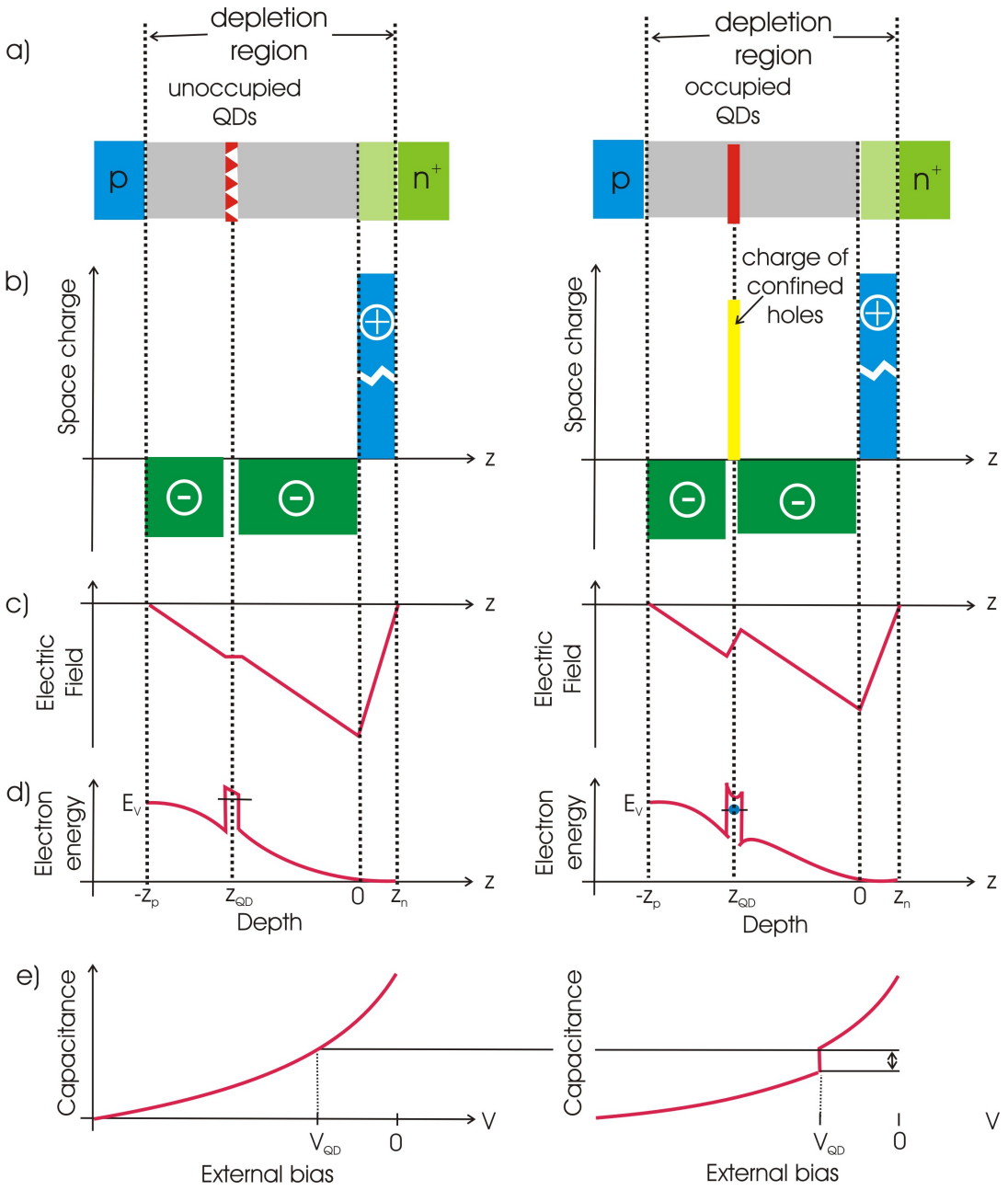

Figure 4.2: Quantum dots inside the p-doped region of a n^+p junction. The left hand-side depicts non-occupied QDs, while right hand-side depicts the occupied QDs. (a) Fixed space charge distribution, (b) electric field, (c) valence band, and (d) capacitance (Adapted from [102]).

The capacitance for the quantum dot layer is given by

$$C_{QD} = \epsilon_s \sqrt{\frac{qN_B}{2(\epsilon_s(V_{bi} - Vext) + qn_h N_{QD} x_{QD})}} \tag{4.9}$$

where n_h is the number of confined holes, N_D the density of the QDs, and x_{QD} the distance of the QD layer from the pn junction. The relative change in the capacitance

of the junction due to the charged QDs is expressed as follows

$$\Delta C = \frac{C}{C_{QD}} = \sqrt{1 + \frac{q n_h N_{QD} x_{QD}}{\epsilon_s (V_{bi} - V_{ext})}} \tag{4.10}$$

The change in the capacitance due to the QDs can be seen as a hysteresis in the capacitance characteristic of the pn junction. The capacitance of a pn junction is inversely proportional to square root of voltage as shown in Fig. 4.3 as a dashed line ($\sqrt{1/V}$). In the presence of QDs in pn junction a capacitance-voltage profile deviated from the ideal behavior is obtained. The capacitance of the pn junction in which QDs are embedded is depicted in Fig. 4.3 as a blue line for the downward voltage sweep (from 0 V towards higher reverse voltages). Initially, the QDs are occupied by holes and the capacitance profile of the junction is the same as a simple pn junction. When the QDs enter the depletion region with increasing reverse voltage, the confined holes in QDs are emitted and compensate the negatively charged ions in the junction. After all the holes are emitted, capacitance profile returns to the one of simple pn junction.

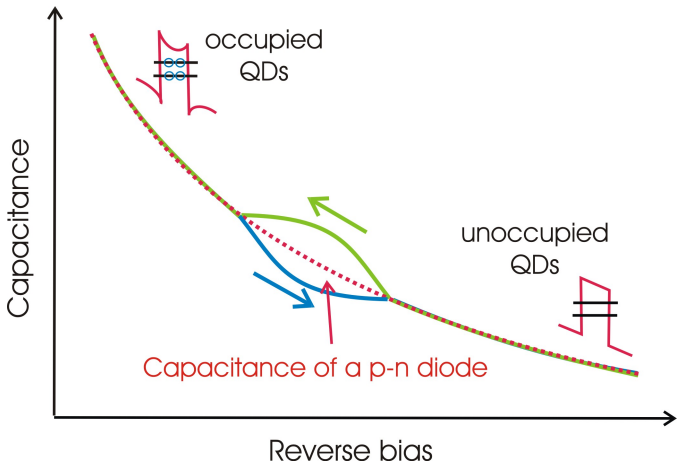

Figure 4.3: The ideal capacitance-voltage characteristic of a pn junction is depicted as a dashed line. The capacitance decreases with an increasing voltage. In the presence of the QDs in the pn junction, the capacitance characteristic deviates from that of the ideal case. In downward sweep, the holes emitted from the fully charged QDs lower the capacitance. In upward sweep, the holes captured by the empty QDs increase the capacitance. These effects can be observed via the capacitance measurement as a hysteresis opening.

When the direction of the voltage sweep is reversed, a similar behavior is observed as depicted in Fig 4.3 as a green line. Initially, no holes are confined inside the QDs and the capacitance exhibits the ideal characteristic of a pn junction without QDs. In the case that the barrier is eliminated, holes are captured by QDs resulting in an increase in the capacitance due to the thinner depletion region. It can be seen in the curve as a deviation from the ideal behavior. After all the QDs are filled, the deviation disappears. The difference between these two curves caused by voltage sweeping is known as the *hysteresis opening* and its magnitude depends on the QD density and the distance of QDs from the junction.

4.2 Time-resolved capacitance spectroscopy

Eq. 4.10 implies that the capacitance of the junction is sensitive to any charges trapped in the deep levels of the depletion region. Hence, the carrier dynamics of deep levels can be examined by the time-resolved capacitance spectroscopy method.

Deep Level Transient Spectroscopy (DLTS) was developed by Lang [93] in 1974 to determine the electronic properties of deep levels in semiconductors such as activation energy, capture-cross section, and carrier concentration [94, 98, 132, 133]. This method is based on the time-resolved measurement of the capacitance of a pn junction under varying bias voltages and temperatures. Quantum dots can capture and emit carriers similar to the deep levels in semiconductors. Therefore, the same method can be used to investigate the carrier dynamics of QDs [134, 135]. Indeed, DLTS has been successfully used to study various QD systems [136–142].

In the above, we derived an expression for the capacitance of a on junction in the static (i.e. not time-dependent) case. To account for the time-dependent case, Eq. 4.10 must be rewritten:

$$\Delta C(t) = \frac{C}{C_{QD(t)}} = \sqrt{1 + \frac{q n_h(t) N_{QD} x_{QD}}{\epsilon_s (V_{bi} - V_{ext})}} \tag{4.11}$$

4.2.1 Measurement principle

In this section, the DLTS measurement principle for hole emission is described. All the processes are analogous for electron emission. The sketch of a DLTS sample used in the work is depicted in Fig. 4.4. It is a n^+p diode into which a single QD layer is embedded. As explained in previous sections, the depletion region occurs mostly on the p-side of the sample.

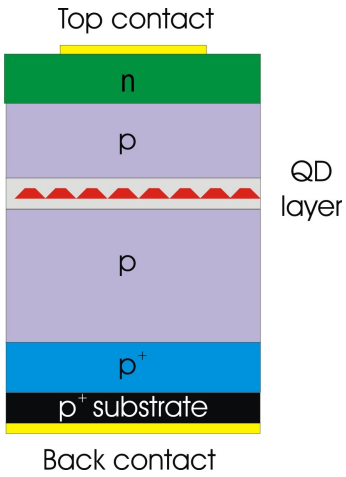

Figure 4.4: Schematic representation of a DLTS sample used in this work. A single quantum dot layer is embedded in n^+p diode where the depletion region extends towards p-side. The charge state of the QDs can be altered by applying a bias voltage. The electronic properties of the QDs can be examined by time-resolved measurement.

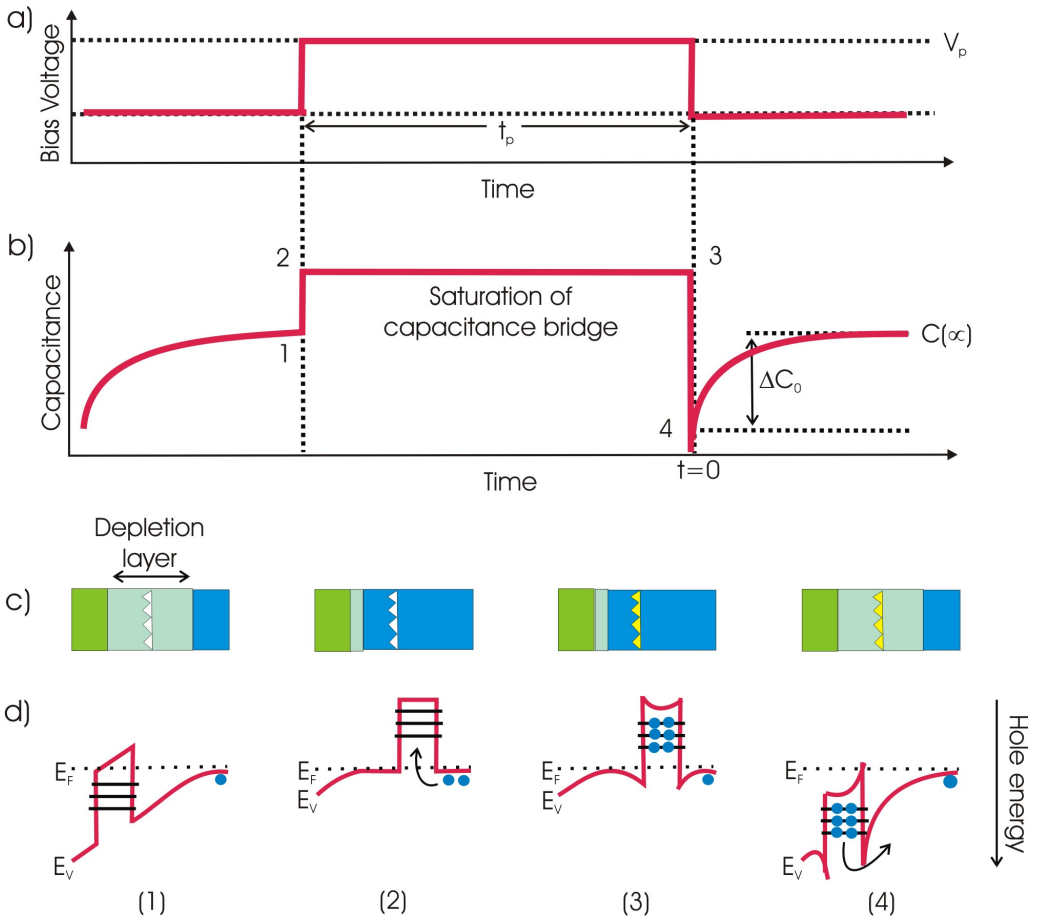

Figure 4.5: The DLTS measurement cycle for the hole emission from a single QD layer embedded into n^+p diode. (a) The voltage is set to V_m such that the QD is in depletion region. The Fermi level is above all the energy levels of QD leading to non-occupied QD. (b) The voltage is set to V_p where the QD is out of depletion region. The Fermi level is under all the energy levels of QDs. The holes are captured. (c) The QD is fully occupied by holes. (d) The voltage is re-set to the V_m voltage such that the QD enters the depletion region resulting in thermal emission of the holes. During the thermal emission, the capacitance is recorded.

Fig. 4.5 depicts the work cycle on the DLTS sample. The measurement is performed in following order: in initial state, the voltage is set to measurement voltage V_m at which QD sits in the depletion region. Therefore, the Fermi level lies above all the energy levels of the QD, the QD, hence, is unoccupied (see Fig. 4.5(1)). Then the voltage is set to the pulse voltage V_p such that the QD layer lies outside the depletion region. The Fermi level shifts below all the energy levels of the QD. Now, holes are captured by the QD and subsequently relax into lowest energy level by dissipating their excess energy with the mechanism explained in Chapter 2 (see Fig. 4.5(2)). At steady-state the QD is completely filled by holes (see Fig. 4.5(3)). In the final step, the voltage is re-set to the V_m. The QD enters again in the depletion region and the Fermi level shifts back above all the energy levels of the QD. However, the QD is still occupied by holes. They are emitted by thermal emission (see Fig. 4.5(4)). From this moment (t_0) until the thermal equilibrium where the QD is completely depleted from

holes, the capacitance is recorded yielding a capacitance transient curve. According to the Eq. 4.11 the capacitance transient should be mono-exponential with a specific time constant τ. Therefore the capacitance transient is given by

$$C(t) = C(\infty) - \Delta C_0 \exp\left(-\frac{t}{\tau}\right) \tag{4.12}$$

where ΔC_0 is the amplitude depending on the number of carriers involved and $C(\infty)$ is the steady state capacitance. From the recorded transient the time constant is extracted. The emission rate e_p for holes is defined as the reciprocal of the time constant ($e_p = 1/\tau$). Therefore, the activation energy and capture cross section are derived from the emission rate described in Chapter 2.

For the measurements carried out in this work, the same transient is recorded approximately 20-30 times in order to enhance the signal-noise ratio [142]. All the process is repeated at different temperature steps sweeping a temperature range generally from 30 K to 400 K.

4.2.2 Rate window and double-boxcar method

The emission transient from a single energy level should be mono-exponential and τ should be easily derived from it, as expressed Eq. 4.12. However, in general, multi-exponential behavior is obtained for the emission process from QDs instead of mono-exponential behavior due to ensemble broadening, multiple emission paths, and many particle effects [143]. Therefore, the linear fit of the emission transient is not possible. The *rate window (boxcar) method* can be hence used in order to extract the time constant from a multi-exponential transient.

Fig. 4.6a depicts the protocol of the boxcar method. Two different time points are chosen for each capacitance transient to define the edges of the boxcar and the difference in the capacitance across the rate window is calculated:

$$S(T, t_1, t_2) = C(T, t_2) - C(T, t_1) = \Delta C_0 \left[\exp\left(-\frac{t_2}{\tau(T)}\right) - \exp\left(-\frac{t_1}{\tau(T)}\right)\right] \tag{4.13}$$

The resulting curve $S(T, t_1, t_2)$ is plotted versus T as shown in Fig. 4.6b. The maximum of the peak is obtained by differentiating 4.13 with respect to τ and occurs for:

$$\tau(T) = \frac{t_2 - t_1}{\ln\left(\frac{t_2}{t_1}\right)} = \tau_{ref} \tag{4.14}$$

τ_{ref} is the *reference time constant* of the rate window and it represents the time constant of the system at the temperature at which the maximum occurs. Basically, in the boxcar method, the whole temperature range is swept to find the temperature at which the reference time constant occurs instead of obtaining the time constant at a given temperature. However, it should be mentioned that the maximum in the DLTS spectra occurs only if holes are emitted thermally , because the tunnel emission is independent of temperature producing only constant DLTS signals.

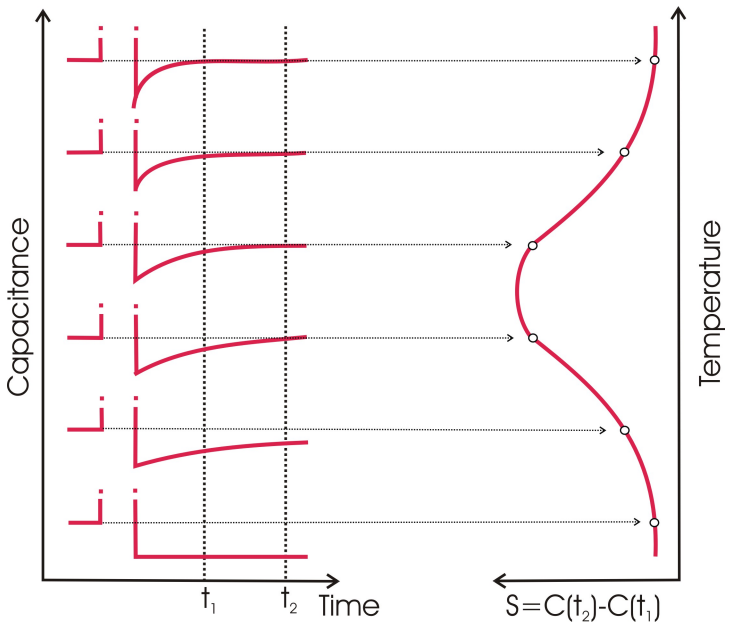

Figure 4.6: Schematic depiction of the boxcar method (a) Two different time points are chosen for each transient at different temperatures to define the edges of the boxcar. (b) The difference in capacitance at two different time points are plotted against the temperature. The position of the peak yields the reference time.

For an ensemble of self-organized QDs such as the structure used in this work, the DLTS signal will be broadened inhomogeneously, representing the mean emission process of all thermally activated emission processes involved. The energy levels of such systems are affected by Gaussian braodening. In this case, the derived time constant is the time constant of the levels at the maximum of the Gaussian [143].

The time constant is equal to the reciprocal of the thermal emission:

$$\frac{1}{\tau_{ref}} = \gamma_p T^2 \sigma_p^{\infty} \exp\left(-\frac{E_a^h}{k_B T}\right) \tag{4.15}$$

Eq. 4.15 is rewritten by taking the natural logarithm of both sides:

$$\ln(\tau_{ref} T^2 \gamma_p) = \frac{E_a^h}{k}\frac{1}{T} - \ln(\sigma_{\infty}) \tag{4.16}$$

By plotting the $\ln(\tau_{ref} T^2 \gamma_p)$ at different τ_{ref} against $1/T$ *Arrhenius plot* is obtained. Then, the activation energy E_a^h is derived from the slope of the linear fit, and the capture cross section σ_p^{∞} from the intercept with the y-axis of the linear fit.

4.2.3 Charge selective DLTS measurements

In DLTS measurement the QDs are fully charged with carriers by the pulse voltage V_p and the emission of all bound states is observed. The DLTS spectrum is hence broadened and a *mean activation energy* and a *mean capture cross section* can be derived. In order to investigate the electronic properties of the samples more accurately,

the charge selective DLTS method was developed by Geller *et al.* at TUB [92]. This method was successfully applied to study the electronic properties of various material systems [83–85]. The aim of this method is to observe the emission of ideally only one charge carrier per QD. This method enables the investigation of the energy levels separately and the determination of the activation energy and the capture cross section of the ground state. The pulse voltage V_p and the measurement voltage V_m are divided into smaller portions to charge and discharge the energy levels one by one.

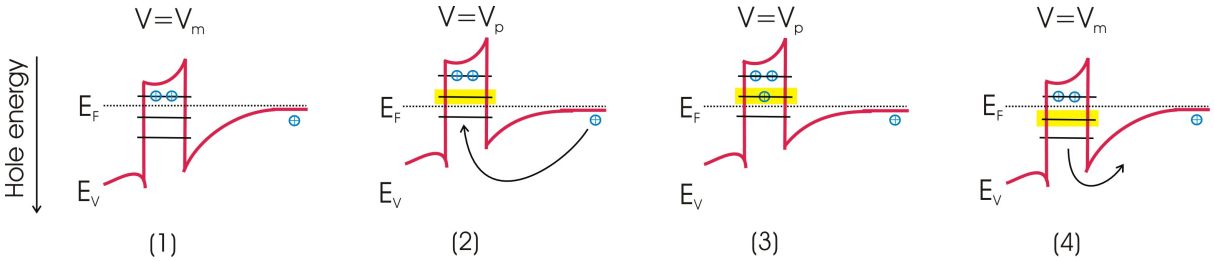

Figure 4.7: The charge selective DLTS measurement cycle for the hole emission from a single QD layer embedded into n^+p diode. (a) The voltage set to the V_m such that the Fermi level is under the first energy level of the QD (b) The voltage is set to the V_p shifting the Fermi level under the energy level of the QD. (c) The holes occupy the second energy level. (d) The voltage is re-set to the V_m voltage such that the Fermi level returns to its initial position resulting in the emission of the holes from the second energy level. By this method, all the energy levels of the QD can be examined one by one by choosing the proper voltage interval.

Fig. 4.7 depicts the work cycle of the charge selective DLTS. During the charge selective DLTS measurement, the QDs are always located in the depletion region opposed to the conventional DLTS. Initially, a measurement voltage V_m is chosen such that the QD is partially occupied (see Fig. 4.7a). Then a pulse voltage V_p is chosen moving the Fermi level towards the higher energy level (see Fig. 4.7b). Now, a hole from the valence band is captured by the QD due to the presence of an available state (see Fig. 4.7c). Finally, voltage is re-set to the V_m shifting the Fermi level to its initial position. The captured hole is now emitted by thermal emission. During emission process, the capacitance transient is recorded. After the measurement is done, the V_m and the V_p are increased by ΔV to investigate the next energy level. By choosing different voltage values all the energy levels of the QD can be swept. Finally, the activation energy of the ground state, called *localization energy*, and its capture cross section are obtained. The storage time is then calculated from localization energy and capture cross section.

As in conventional DLTS, the capacitance transient is recorded 20-30 times to enhance the signal-noise ratio and the measurement is repeated at different temperature steps varying mostly from 30 K to 400 K.

Summary

- A pn junction is formed by bringing the oppositely doped semiconductors into an electric contact. At equilibrium, a depleted region from the free carriers is formed and the either sides of the junction are oppositely charged giving a rise to a capacitance.
- The capacitance of the junction is sensitive to any charge carriers in the depletion region, hence, the QDs in the junction contribute to the total capacitance.
- In order to determine the activation energy and the capture cross section of a QD, time-resolved capacitance measurement is used. In this method, the QDs are filled by a pulse voltage, the pulse voltage is then removed. The capacitance is recorded during the thermal emission of the carriers.
- The double boxcar method is used for analysis, Two different time points are chosen for each transient at different temperatures and the difference in capacitance is calculated. The reference time constant represents the time constant for the temperature at which maximum occurs. Thermal emission is equalized to the reciprocal of the reference time constant. The activation energy and the capture cross section are respectively derived from the slope and the intercept with y-axis of the linear fit of the Arrhenius plot obtained through the time constant and the thermal emission.
- The transients recorded by the DLTS consist of the contribution of all the energy levels of an ensemble of quantum dots. The DLTS spectra is thus broadened. In order to investigate the electronic properties of the samples more accurately the Charge Selective DLTS method is employed, which enables us to investigate ideally a single charge carrier per QD.

Chapter 5

DLTS Results - GaSb/GaP

In this chapter the electrical characterization of the samples containing GaSb QD layers embedded in GaP matrix is presented. The localization energies and the capture cross sections of each sample are determined by time-resolved capacitance measurement. By using the mentioned parameters their storage times at room temperature are calculated.

The importance of the GaSb/GaP material system is that it is a completely novel material system which was grown for first time in 2012 [144, 145], and it is a good candidate to achieve non-volatility due to its high projected localization energy [25, 146,147]. The samples were grown by MBE by the group of X. Wallart at the University of Lille, France.

Parts of this chapter have been published [148].

5.1 Growth of the samples

Five different samples are grown with slightly different growth conditions to examine the effect of growth condition on localization energy and capture cross section. The schematic structures of the samples are depicted in Fig. 5.1. Four samples out of five (named QD-1, QD-2, QD-3, and QD-4) contain GaSb QD layers, whereas the fifth one is a reference sample without a QD layer (named Ref). During the growth of the samples of QD-1, QD-2, and QD-3, the QD thicknesses and the growth rates are kept constant to 1.2 ML and 0.2 ML/s, respectively, while temperature is increased in 20 °C steps for each sample. The deposition temperature for QD-1 is 430 °C, 450 °C for QD-2, and 470 °C for QD-3. For QD-4 the growth rate is halved, while keeping the rest of parameters constant. Tab. 5.1 shows the growth conditions for all the samples.

For all the samples, a p-doped GaP (001) substrate (nominal $p = 1.3 \times 10^{18}$ cm^{-3}) is used. A 500 nm thick highly p-doped GaP contact layer ($p = 2 \times 10^{18}$ cm^{-3}) is grown at 600 °C at a growth rate of 1 ML/s, whereas a 700 nm thick lowly p-doped GaP layer ($p = 2 \times 10^{16}$ cm^{-3}) and a 7 nm undoped GaP layer are grown at 560 °C. The QD layers are grown as reported above. After QD formation, the growth is interrupted for 50 s to obtain bigger QDs by ripening, hence higher localization energy. Then, the QDs are capped with 7 nm of undoped GaP at 430 °C at a growth rate of 0.2 ML/s to stop the ripening of the QDs. Afterwards, a lowly p-doped layer ($p = 2 \times 10^{16}$ cm^{-3}) is

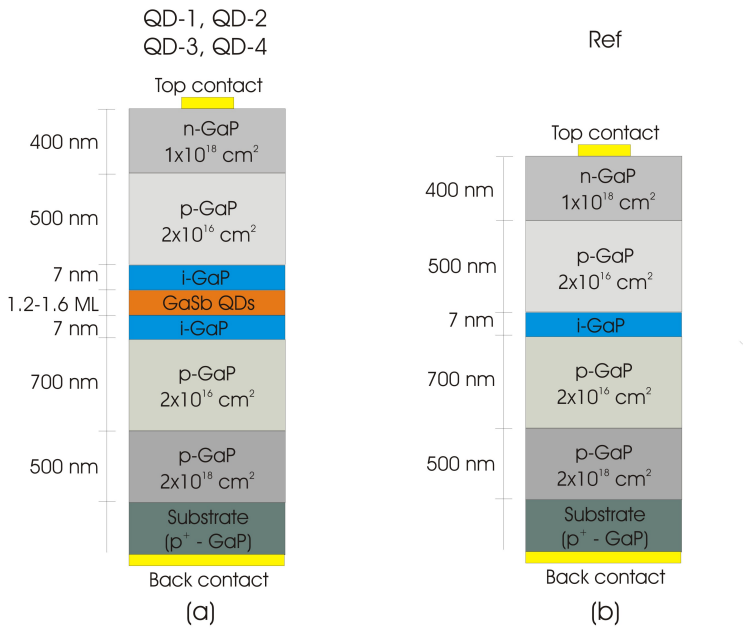

Figure 5.1: Epitaxial structure of the samples containing GaSb QDs embedded in GaP. (a) The structures of the samples QD-1, QD-2, QD-3, and QD-4. Only their growth conditions are different. (b) The structure of the reference sample without a QD layer.

grown at 560 °C. As a final step, a highly n-doped layer ($n = 1 \times 10^{18}$ cm^{-3}) is grown at 600 °C to ensure that the top contact is Ohmic. Be atoms for p layers and Si atoms for n layers are used as dopants.

All the samples are processed by standard optic lithography and dry etching methods, so that 400 μm and 800 μm diameters mesas are formed. The Ohmic contacts are thermally evaporated, where Ni/Au-Ge/Au is used on the n side and Ni/Zn/Au on the p side. After deposition of the contacts, the samples are annealed at 400 °C for 3 minutes in a nitrogen atmosphere.

Fig. 5.2a shows an Atomic Force Microscope (AFM) micrograph of a sample

Sample	Temperature (°C)	Thickness (ML)	Rate (ML/s)
Ref	n/a	n/a	0.2
QD-1	430	1.2	0.2
QD-2	450	1.2	0.2
QD-3	470	1.2	0.2
QD-4	470	1.2	0.1

Table 5.1: The growth conditions of the samples. The QD thickness and the growth rate are kept constant during the growth of QD-1, QD-2, and QD-3, while the growth temperature is increased in 20 °C steps. For the QD-4 the growth rate is halved and the remaining parameters are kept constant.

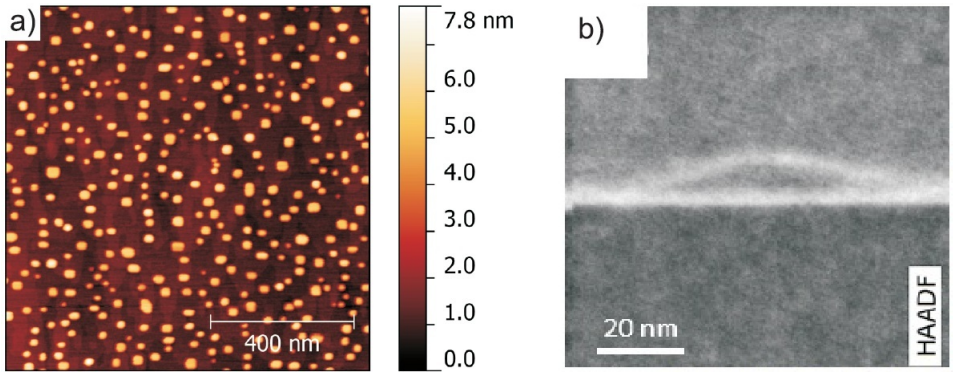

Figure 5.2: (a) The AFM picture of a GaSb/GaP QDs sample, and (b) the HAADF-SEM picture of a single QD from the same sample.

identical to QD-4 up to the QD layer. The QDs can be clearly seen. Their average lateral size is ∼36 nm, their height ∼8 nm, and their density 9×10^9 cm^{-2}. Fig. 5.2b shows the high annular dark field anaylsis (HAADF) - scanning electron microscope (STEM) image of a single QD from the same sample [149].

5.2 Capacitance-Voltage (CV) Characterization

Capacitance-voltage (CV) measurement were performed on all the samples to identify the voltage values where the QDs are completely filled by holes and depleted from holes, which are used for the DLTS measurement. CV measurements were performed at four different frequencies: 1 kHz, 10 kHz, 100 kHz, and 1 MHz. The voltage was swept from $V = -1$ V (forward bias) to $V = 10$ V (reverse bias) in 0.05 V step. The temperature was swept from 100 K to 400 K in 50 K steps.

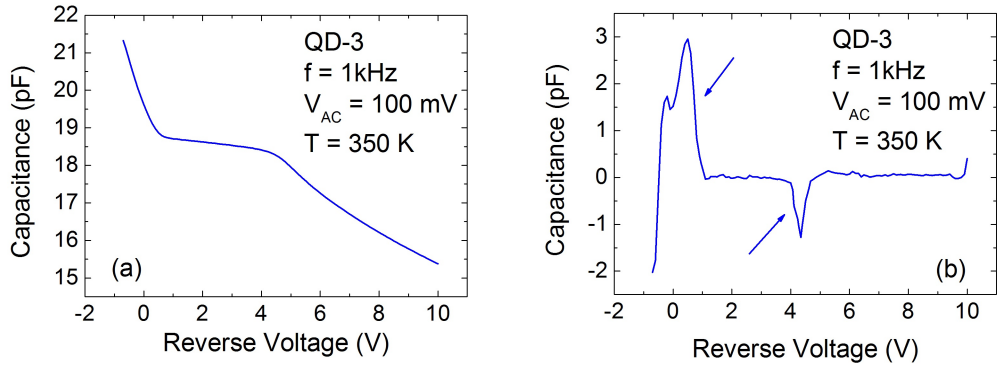

Figure 5.3: (a) CV profile of QD-3 at 1 kHz at 350 K. A plateau attributed to the QDs can be seen. (b) Second derivative of the CV profile of QD-3. The maximum and the minimum voltages become more visible.

Fig. 5.3a shows the CV profile of QD-3 at 1 kHz at 350 K. A plateau attributed to the QDs can be clearly seen between -0.5 V (forward bias) and 4.5 V (reverse

bias). These values represent the voltage values where the QDs are fully filled and completely depleted, and they are needed to specify the voltage interval in the DLTS measurement. Fig. 5.3b shows the second derivative of the CV curve. The maximum and the minimum values can be clearly seen.

5.3 Conventional DLTS

Conventional DLTS was performed on all the samples to determine the mean activation energy and the mean capture cross section. The voltage values used in the DLTS measurement were specified by the CV measurement for each sample. The rest of the parameters are the same for all the samples. The pulse voltage last 5 s, whereas the measurement voltage lasts 95 s. The temperature was swept from 200 K to 400 K in 2 K steps. The capacitance transients were measured at a frequency of 1 MHz and an AC voltage of 100 mV. The transients were recorded 10 times to enhance the signal-to-noise ratio. Results were analyzed with the double-boxcar method. t_1 ranges from 1 s to 27 s, whereas $t_2 = 3t_1$ and $t_{av} = 0.5t_1$. These parameters yield reference times τ_{ref} between 2 s and 56 s.

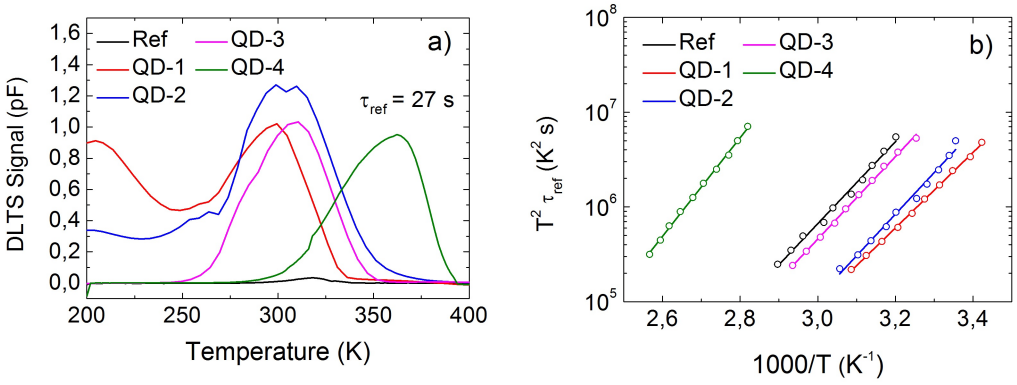

Figure 5.4: (a) Conventional DLTS results of all GaSb/GaP samples at $\tau_{ref} = 27$ s. (b) Arrhenius plots of the same measurements.

Fig. 5.4 shows the DLTS signal for all the samples for τ_{ref} =27 s. The reference sample presents a very small peak centered around 320 K shown as a black line. The samples QD-1, QD-2, and QD-3 have similar DLTS signals drawn as red, blue and pink lines, respectively. All of them represent a peak centered around 300 K suggesting that they have similar localization energies. The sample QD-4, however, has a different behavior. A peak appears at a higher temperature, around 365 K. This implies that QD-4 has the highest localization energy.

The maximum points of DLTS peaks with the different reference times are chosen and plotted in the Arrhenius plot. The slope of the linear fit of the Arrhenius plot yields the mean activation energy, whereby its intercept with the y-axis yields the mean capture cross section. The activation energies and the capture cross section for each sample are listed in Tab. 5.2 with their error margins.

Sample	E_a (meV)	σ_∞ (cm^2)
Ref	852 ($\pm$9)	9×10^{-15} (0.1)
QD-1	790 ($\pm$10)	6×10^{-15} (0.2)
QD-2	850 ($\pm$10)	9×10^{-14} (0.2)
QD-3	860 ($\pm$20)	2×10^{-14} (0.4)
QD-4	1030 ($\pm$10)	5×10^{-14} (0.2)

Table 5.2: The mean localization energies and the mean capture cross sections of the samples obtained by conventional DLTS.

The mean activation energy for QD-1, QD-2, and QD-3 are all close to each other varying from 790 meV to 860 meV. The mean activation energy for QD-4 is considerably higher than the others, which is E_a =1.03 ($\pm$0.01) eV with an associated mean capture cross section of 5×10^{-14} cm^2 with an uncertainty of 0.2 orders of magnitude.

5.4 Charge Selective DLTS

As explained in Chapter 4, the DLTS measurement detects the signal from the different QDs and the different energy levels, so that it yields the mean activation energy and the mean capture cross section. By charge selective DLTS the signal from ideally only one hole per QD can be detected enabling us to investigate the energy levels of the QDs separately. Hence, the localization energy and its associated capture cross section can be determined, which are expected to be higher than the mean activation energy and the mean capture cross section.

Charge selective measurements were performed on all the samples. The pulse and measurement voltages are determined using the CV measurement for each sample. The rest of the parameters are same for all the samples. The pulse time lasts 5 s, whereas the measurement pulse lasts 95 s. The temperature was swept from 200 K to 400 K in 2 K steps. The transients were recorded 10 times to improve the signal-to-noise ratio. DLTS signals were analyzed with double-boxcar method. t_1 ranges from 1 s to 27 s, whereas $t_2 = 3t_1$ and $t_{av} = 0.5t_1$. These parameters yield reference times τ_{ref} between 2 s and 56 s.

5.4.1 Sample Ref

Measurement voltage V_m ranges from -0.5 V to 3 V in 0.5 V steps, whereas pulse voltage $V_p = V_m - 0.5$ V. The DLTS signals for the reference time of $27s$ are shown in Fig. 5.5. There is no single peak at any of the measurement voltages, which is the expected behavior since the sample does not have a QD layer, thus any internal energy level. However, a small peak was detected by the conventional DLTS for the same sample. That signal is attributed to deep defect levels and the charge selective DLTS is suitable for only self-organized QDs, not for deep defect levels. Because deep defects are distributed uniformly in the bulk of material and the small range of applied voltage in charge selective DLTS is not sufficient to detect deep defects.

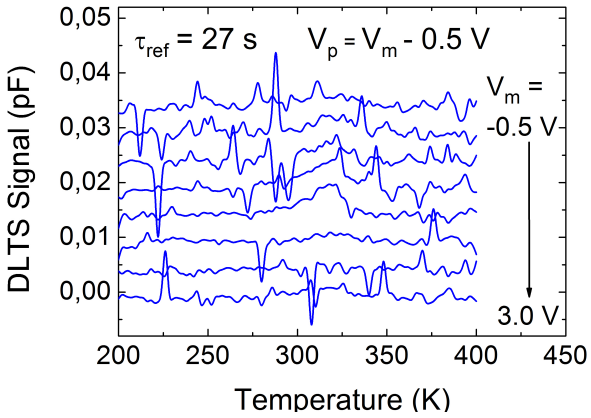

Figure 5.5: Charge Selective DLTS result of the reference sample for $\tau_{ref} = 27$ s. No peak can be seen due to the absence of QDs.

5.4.2 Sample QD-1

Measurement voltage V_m ranges from 0 V to 10 V in 1 V steps, whereas pulse voltage $V_p = V_m - 1$ V. The DLTS signals for $\tau_{ref} = 27$ s are shown in Fig. 5.6a. A single peak can be seen for all the measurement voltages. The peak is centered around 200 K at $V_m = 0$ V, and shifts towards higher temperature with increasing bias. A peak is centered around 300 K at $V_m = 10$ V.

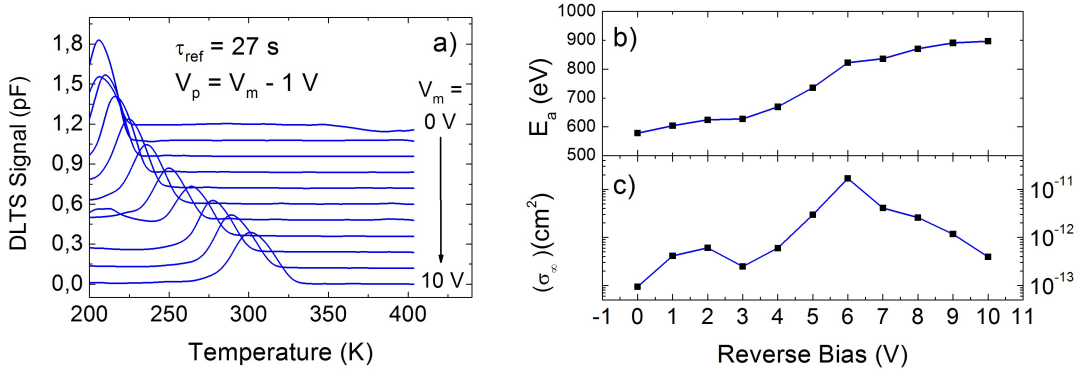

Figure 5.6: (a) Charge Selective DLTS result for $\tau_{ref} = 27$ s, from $V_m = 0$ V to $V_m = 10$, $V_p = V_m - 1$ V, (b) activation energies, and (c) capture cross sections.

For each measurement voltage, an activation energy and a capture cross section are determined from the Arrhenius plot. Accordingly, the activation energy starts from 578 (±6) meV at $V_m = 0$ V and increases gradually until 900 (±20) meV at $V_m = 10$ V as shown in Fig. 5.6b. The last activation energy represents the localization energy of the sample. The capture cross section starts from 9×10^{-14} cm^2 at $V_m = 0$ V (with an uncertainty of 0.1 orders of magnitude) and increases with increasing bias. At $V_m = 10$ V the capture cross section is determined to be 4×10^{-13} cm^2 (with an uncertainty of

0.3 orders of magnitude) shown in Fig. 5.6c.

To conclude, the localization energy of the sample QD-1 is determined to be 900 (±20) meV with the associated capture cross section of 4×10^{-13} cm^2 (with an uncertainty of 0.3 orders of magnitude).

5.4.3 Sample QD-2

Measurement voltage V_m ranges from 0 V to 10 V in 1 V steps, whereas the pulse voltage $V_p = V_m - 1$ V. The DLTS spectra for the reference time of 27 s are shown in Fig. 5.7a. A single peak can be seen for all the measurement voltages except $V_m = 0$ V. A peak is centered around 250 K at $V_m = 1$ V. Then the peak shifts towards higher temperatures until 350 K at $V_m = 8$ V. It should be noted that the DLTS spectra of QD-2 is more noisy than the previous sample, which originates from either the structural integrity of the sample or the contact quality.

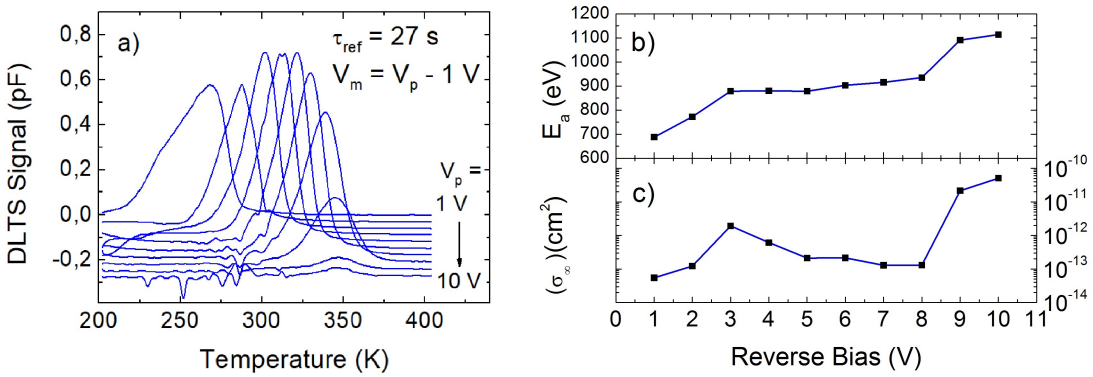

Figure 5.7: (a) Charge Selective DLTS result for $\tau_{ref} = 27$ s, from $V_m = 1$ V to $V_m = 10$ V, $V_p = V_m - 1$ V, (b) activation energies, and (c) capture cross sections.

The activation energy starts from 687 (±7) meV at $V_m = 1$ V and increases gradually until 1.10 (±0.03) meV at $V_m = 10$ V shown in Fig. 5.7b. The last parameter represents the localization energy of the sample. The capture cross section starts from 5×10^{-14} cm^2 at $V_m = 1$ V (with an uncertainty of 0.2 orders of magnitude) and increases gradually until 4×10^{-11} cm^2 at $V_m = 10$ V (with an uncertainty of 0.5 orders of magnitude) shown in Fig. 5.7c.

To conclude, the localization energy of the sample is determined to be 1.10 (±0.03) eV with the associated capture cross section of 4×10^{-11} cm^2 with an uncertainty of 0.5 orders of magnitude.

5.4.4 Sample QD-3

Measurement voltage V_m ranges from 0 V to 6 V in 1 V steps, whereas pulse voltage $V_p = V_m - 1$ V. The DLTS signals for the reference time of 27 s from $V_m = 2$ V to $V_m = 5$ V are shown in Fig. 5.8a. At the other measurement voltage the sample presents no signal. A single peak can be seen for all the mentioned measurement voltages. A peak

appears around 300 K at $V_m = 2$ V and shifts slightly towards higher temperatures until 325 K at $V_m = 5$ V. For this sample less number of peaks appear than the other samples. This could imply that the sample has less bound states.

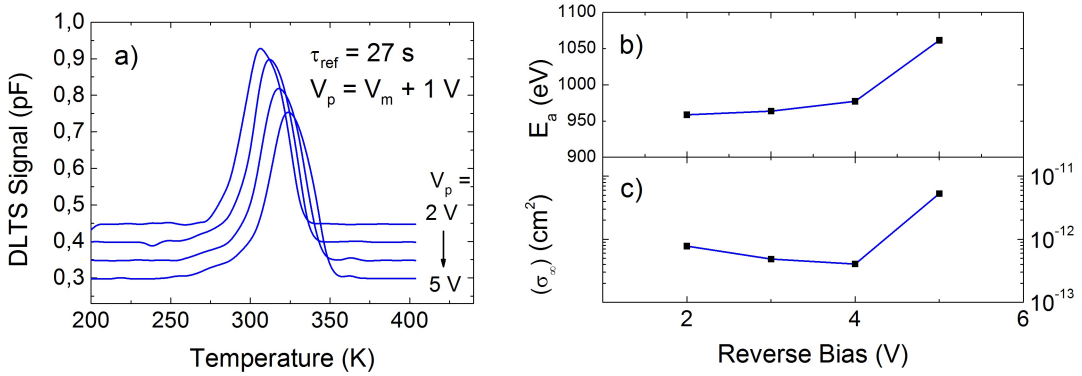

Figure 5.8: (a) Charge Selective DLTS result for $\tau_{ref} = 27$ s, from $V_m = 2$ V to $V_m = 5$, $V_p = V_m - 1$ V (b) activation energies, and (c) capture cross sections.

The activation energy starts from 960 ($\pm$6) meV at $V_m = 2$ V and remains almost same for the next two measurement voltages. Finally, it increases to 1.07 ($\pm$0.01) eV at $V_m = 5$ V. This pattern confirms the argument above that the sample has less bound states. The capture cross section starts from 8×10^{-13} cm^2 at $V_m = 2$ V (with an uncertainty of 0.1 orders of magnitude) and remains almost unchanged until the last measurement voltage $V_m = 5$, which is 7×10^{-12} cm^2 (with an uncertainty of 0.2 orders of magnitude).

To conclude, the localization energy of the sample is determined to be 1.07 ($\pm$0.01) eV with the associated capture cross section of 7×10^{-12} cm^2 with an uncertainty of 0.2 orders of magnitude.

5.4.5 Sample QD-4

Measurement voltage V_m ranges from -1.5 V to 4 V in 0.5 V steps, whereas pulse voltage $V_p = V_m - 0.5$ V. The DLTS signals for the reference time of 27 s for the measurement voltage from $V_m = -1.5$ V to $V_m = 3$ V are shown in Fig. 5.9a, while the remaining spectra are excluded due to having small signals assumed that they originate from the defect levels. A single peak can be seen for all the measurement voltages. A peak appears around 290 K for $V_m = -1.5$ V and shifts towards higher temperatures until 370 K at $V_m = 3$ V.

The activation energies and the capture cross sections are extracted from the Arrheinus plot. Accordingly, the activation energy starts from 735 ($\pm$7) meV at $V_m = -1.5$ V and increases gradually until 1.18 ($\pm$0.01) eV at $V_m = 3$ V. The capture cross section starts from 2×10^{-15} cm^2 at $V_m = -1.5$ V (with an uncertainty of 0.1 orders of magnitude) and increases gradually until 1×10^{-12} cm^2 at $V_m = 3$ V (with an uncertainty of 0.1 orders of magnitude).

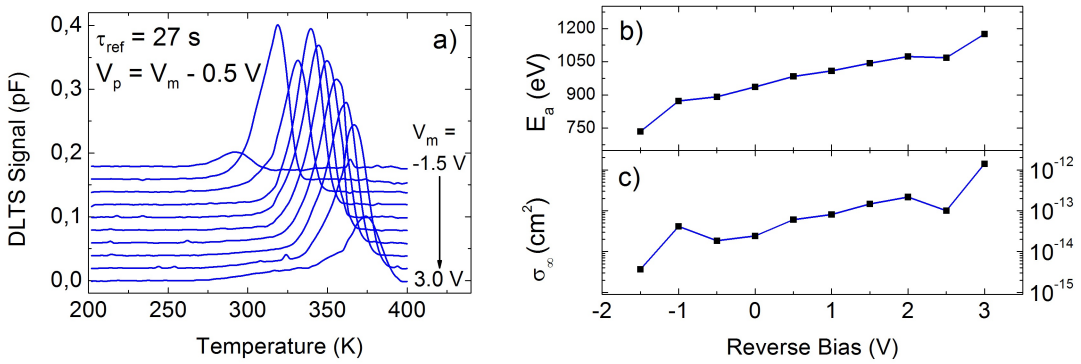

Figure 5.9: (a)Charge Selective DLTS result of the sample 1-QD-4, (b)activation energies, and (c) capture cross sections

To sum up, the localization energy of the sample is determined to be 1.18 (±0.01) eV with the associated capture cross section of 1×10^{-12} cm^2 with an uncertainty of 0.1 orders of magnitude.

5.5 Storage Time

Storage time is calculated using localization energy and capture cross section [1]. Tab. 5.3 shows the activation energies, the capture cross sections, and the storage times for holes at room temperature for all the samples.

Storage time was not calculated for the reference sample, since it does not contain a quantum dot layer. The localization energy of QD-1 is determined to be 900 (±20) meV with the associated capture cross section of 4×10^{-13} cm^2 with an uncertainty of 0.3 orders of magnitude. Accordingly, the storage time for holes at room temperature is calculated to be 30 s (with an uncertainty of 0.1 orders of magnitude). For QD-2, the localization energy is determined to be 1100 (±30) meV with the associated capture cross section of 4×10^{-11} cm^2 with an uncertainty of 0.5 orders of magnitude, yielding the storage time of 850 s (with an uncertainty of 0.07 orders of magnitude) for holes at room temperature. The increase in storage time in comparison to the previous sample originates from the fact that the latter sample has larger localization energy. The storage time for QD-3 is calculated to be 1100 s (with an uncertainty of 0.04 orders of magnitude) at room temperature with the localization energy of 1.07 (±0.01) eV with the associated capture cross section of 7×10^{-12} cm^2 with an uncertainty of 0.2 orders of magnitude. Although the sample has smaller localization energy than the previous sample, its storage time is longer based on the fact that its capture cross section is smaller than the previous sample. The last sample, QD-4, has the largest localization energy, 1.18 (±0.01) with the associated capture cross section of 1×10^{-12} cm^2 with an uncertainty of 0.1 orders of magnitude, since it was grown with the most optimized parameters, thus, it has also the longest storage time, 3.9 days (with an uncertainty

[1] Storage time is the reciprocal of the thermal emission rate

Sample	E_a (eV)	σ_∞ (cm^2)	τ (at 300 K)
Ref	n/a	n/a	n/a
QD-1	0.90 ($\pm$0.02)	4×10^{-13} (0.3)	30 (0.1) s
QD-2	1.10 ($\pm$0.03)	4×10^{-11} (0.5)	850 (0.07) s
QD-3	1.07 ($\pm$0.01)	7×10^{-12} (0.2)	1100 (0.04) s
QD-4	1.18 ($\pm$0.01)	1×10^{-12} (0.1)	3.9 (0.04) d

Table 5.3: Summary of the all localization energy E_{loc}, capture cross sections σ_∞, and the storage times at room temperature for holes τ of all the samples.

of 0.04 orders of magnitude) at room temperature. This value is the longest value which has been ever reported and represents an increase by 3 orders of magnitude in comparison to the last reported structure [19].

5.6 Discussion

In this chapter, the results of the GaSb/GaP structures grown under slightly different conditions to investigate the effect of the growth condition on localization energy and capture section were presented. The importance of the GaSb/GaP material system is that it is a completely novel material system which was grown for first time in 2012 [144,145], and it is a good candidate to achieve non-volatility due to its projected high localization energy of 1.4 eV [25,146,147]. The results clearly show that there is a relation between the growth condition and the electronic properties of QDs.

For the first three sample, QD-1, QD-2, and QD-3, the QD layer thicknesses and the growth rates are fixed to 1.2 ML and 0.2 ML/s, respectively, while the growth temperature is increased in 20 K step for each sample. Accordingly, QD-1 is grown at 430 K and it has the smallest localization energy. The temperature is increased to 450 K for QD-2 resulting in an increase of 200 meV in localization energy. The temperature is again increased to 470 K for the QD-3, however, there is no further increase in localization energy regarding the error margins. For the QD-4, the temperature is kept constant in 470 K beside the QD thickness of 1.2 ML/s, while the growth rate is halved. As a result, the localization energy is increased by additional 100 meV.

Higher temperature has two different effects: first one is a higher surface mobility, which leads to larger QDs. Second effect is the lower defect density leading to the QDs with higher structure quality. A reduction in growth rate increases the growth time. Therefore, the QDs have much more time to ripen causing the formation of the larger QDs. Both higher temperature and the slower growth rate lead to larger QDs. According to the 8-band $\mathbf{k} \cdot \mathbf{p}$ calculations the larger QDs lead to larger localization energies (see Sec. 1.4.2). It is concluded that the results are compatible to the theoretical predictions.

It should be noted that there is a discrepancy between the projected and the obtained localization energy of the GaSb/GaP system. The projected value is reported to be 1.4 eV [146] and the largest localization energy is determined to be 1.18 ($\pm$0.01) eV. This discrepancy originates from the intermixing of P and Sb atoms in QDs, which

was not taken into account in the simulations.

Similar to the localization energy, the capture cross section also depends on the growth condition since the size of QDs changes based on the growth condition. However, the capture cross sections determined in this study do not show any clear trend opposed to the localization energy. The capture cross section is a complicated parameter and requires much more study to be well understood. Unfortunately, it is out of the scope of this work.

Based on the results, storage time has a trend. For the first sample it is calculated to be 30 s (with an uncertainty of 0.1 orders of magnitude). Then, the storage time for QD-2 is calculated to be 850 s (with an uncertainty of 0.07 orders of magnitude). The increase in storage time originates from the increase in localization energy. The storage time for QD-3 is further increased to 1100 s (with an uncertainty of 0.04 orders of magnitude) due to the decrease in capture cross section. Finally, the storage time of QD-4 is calculated to be 3.9 days (with an uncertainty of 0.04 orders of magnitude), which is the longest storage time value ever reported.

Although the reported storage time value is the longest one, it is still far from the desired value of 10 years. In order to increase the storage time, either localization energy has to be increased or capture cross section has to be decreased. This study shows that different growth conditions change the electronic properties of QDs such as localization energy and capture cross section. Therefore, further alterations in growth condition result in larger localization energy as a result of having larger QDs, thus a longer storage time. Regarding this fact and the presented results, an increase in the localization energy or a decrease in the capture cross section of the presented structure can yield the desired storage time of 10 years. For instance, the localization energy of QD-4 is determined to be 1.18 ($\pm$0.01) eV with an associated capture cross section of 1×10^{-12}. In order to achieve non-volatility with this sample, decreasing its capture cross section to 1×10^{-15} could be sufficient with the same localization energy. Alternatively, its localization energy could be increased to 1.35 eV with the same capture cross section.

To conclude, based on both theoretical calculation [25, 146, 147] and presented experimental results, GaSb/GaP is a promising material system to achieve non-volatility. In order to further increase the storage time either the localization energy has to be increased or capture cross-section engineering has to be developed in order to fine-tune the capture cross-section.

Summary

- The GaSb/GaP QD system is a novel system, which was grown for first time in 2012. This material combination is a good candidate to achieve non-volatility, the 10 years of storage time at room temperature, due to its projected high localization energy of 1.4 eV.
- Different samples were grown under slightly different growth conditions to examine the effect of growth condition on the electronic properties of QDs such as localization energy and capture cross section. Accordingly, the growth temperature is increased in 20 C steps for the first three samples changing from 430 C to 470 K, while keeping the growth rate and the QD thickness constant. For the fourth sample the growth rate is halved, while the remaining parameters are kept constant.
- Localization energy has a trend. When the temperature is increased by 20 C, the localization energy is increased by 200 meV. However, the localization energy is not further increased (within the error margin) by further increase in temperature. On the other hand, the localization energy is increased by 100 meV by halving the growth rate. Higher temperature has two effect on QDs: It yields larger QDs as a result of higher surface mobility and QDs with higher structure quality as a result of lower defect density. Larger quantum dots lead to higher localization energy.
- Capture cross section also changes based on growth condition. However, no clear trend is apparent for capture cross section in this study. Capture cross section is a complex parameter and much more studies are needed to understand it well.
- Storage time is increased constantly for each sample. For QD-2 the storage time is increased due to the increase in localization energy. For QD-3 a decrease in capture cross section leads to longer storage time. For QD-4 the longest storage time is obtained due to its large localization energy.
- Although the longest storage time is reported, it has to be improved to achieve non-volatility. An increase in localization energy or a decrease in capture cross section could lead to non-volatility. Such structure can be obtained by altering growth conditions. It is concluded based on the results that the GaSb/GaP material combination is a promising candidate for non-volatility.

Chapter 6

DLTS Results - InGaSb/GaP

This chapter presents the results of the samples containing of InGaSb quantum dot layers embedded in GaP matrix. The localization energies and the capture cross sections of the samples are determined by time-resolved capacitance measurement. Their storage times at room temperature are calculated using the determined localization energies and the capture cross sections.

InGaSb/GaP is a novel material combination. This work reports on the first ever characterization of this material system. This material combination is also a good candidate to achieve non-volatility according to its estimated high localization energy [25, 150] and the experimental works performed on similar structures[1] [19, 87]. Besides, for first time, Carbon atoms are used as dopant in GaP to prevent some of the issues caused Zn doping, which is the standard doping for p-GaP. The structures are grown by MOCVD method at Technische Universität Berlin by E. Sala.

Parts of this chapter have been submitted to *Phyica Status Solidi (a)*[2].

6.1 Growth of the Samples

Four different samples having different quantum dot layer and barrier layer are grown by MOCVD method: one sample with only an InGaSb QD layer in GaP (named Only-QD), two samples with an InGaSb QD layer in GaP and an additional AlP barrier (named QD-Bar1 and QD-Bar2), and the last one, the reference sample, with only an AlP barrier (named Only-Bar). The schematic of the samples are depicted in Fig. 6.1. All the structures are in the form of n^+p diode grown on a p-doped GaP (001) substrate (nominal $p = 1.15 \times 10^{18}$ cm^{-3}). After the substrate, a 300 nm p-doped GaP layer (nominal $p = 5 \times 10^{16}$ cm^{-3}) is deposited at 620 °C for all the sample. The layer is doped with C atoms by using *carbon trichloro-bromide* (CCl_3Br) as precursor with an input flux of ∼1.1 μmol/min and V/III ratio of 13. Then, an AlP barrier is grown at 800 °C for the samples of Only-Bar, QD-Bar1 and QD-Bar2. The additional barrier acts as the emission barrier for charge carriers in QDs and is used to further increase

[1] Structures based on $In_xGa_{1-x}As$/GaP

[2] E. M. Sala, I. F. Arikan, L. Bonato, F. Betram, P. Veit, J. Christen, A. Strittmatter, and D. Bimberg, "MOVPE-growth of InGaSb/AlP/GaP(001) quantum dots for nano-memory applications."

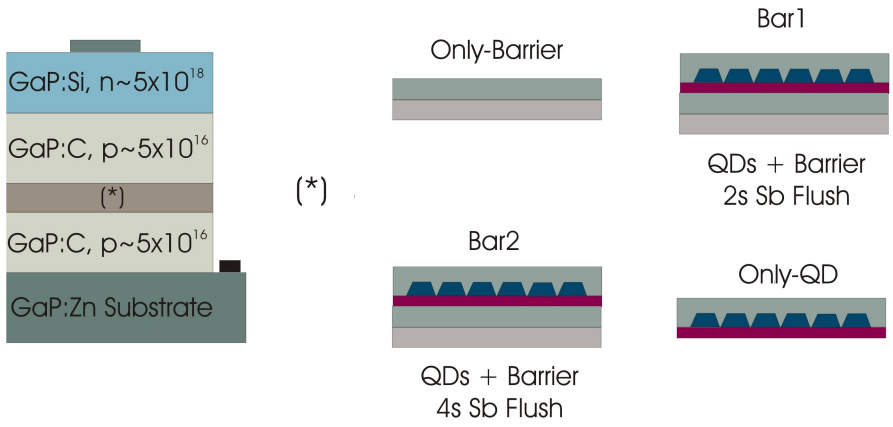

Figure 6.1: Epitaxial structure of the samples consisting of InGaSb quantum dot layer embedded in GaP. (a) Only-Bar. A reference sample containing only AlP barrier layer without a quantum dot layer. (b) QD-Bar1 with an InGaSb QD layer embedded in GaP with an additional AlP barrier. (c) QD-Bar2, identical to the QD-Bar1 with the distinction that its Sb flush is 4 s. It was 2s for QD-Bar1. (d) Only-QD, without AlP barrier.

the localization energy of the QD, hence the storage time. The AlP barrier is capped with a GaP layer to prevent AlP from adsorbing the impurities during the cooling down step before the growth of QDs. Afterwards, a GaAs interlayer, to enable the formation of QD layer, and an InGaSb QD layer are grown at 500 °C. The QD layer is then capped with a 6 nm undoped GaP to stop ripening of the QDs. Thereafter a 500 nm p-doped GaP layer (nominal $p = 5 \times 10^{16}$ cm^{-3}) is deposited at 620 °C. Finally, the structure is completed with a 400 nm highly n-doped layer (nominal $n = 5 \times 10^{18}$ cm^{-3}) at 620 °C to insure the Ohmic top contact. Si atoms are used as dopant for the last segment.

The difference between QD-Bar1 and QD-Bar2 is the exposure time of Sb. It is 2 s for the former, while 4 s for the latter. Sb-soaking is performed prior to the QD formation with a *Triethyl-Antimony* (TESb) input flux kept constant at 2.6 μmol/min. The properties of the quantum dots such as localization energy and capture cross section on which storage time depend can be altered through Sb-soaking [151, 152]. In order to examine the effect of Sb-flush on localization energy and capture cross section, thus on the storage time, two identical sample exposed to Sb for different time of period are grown.

The p-doping of GaP in the structures is carried out by using C atoms for first time. Zn atoms has been so far used as dopants [153, 154], however, it causes some important side effects such as the memory effect [155, 156], a high diffusivity [157, 158], and the high ionization energy of Zn [159]. The use of C atoms as dopants does not present these issues [159], yielding better structures, thus better results.

The samples are processed via standard optical lithography and dry etching methods. Ni/Au-Ge/Au top contacts with diameters of 400 μm and 800 μm are thermally evaporated on the highly n-doped layer. The reason to have the contacts in two different sizes is to have flexibility in the measurements. Because the larger mesas yield the more intense DLTS signal while the smaller mesas yield the smaller RC low-pass

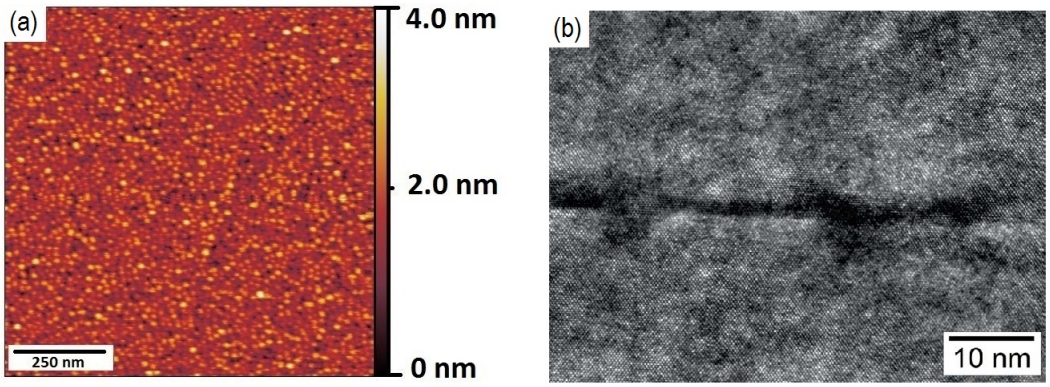

Figure 6.2: (a) The AFM picture of an InGaSb/GaP QDs sample, and (b) the TEM picture of a cross section of the quantum dot layer.

filter effect. On the back side of the structures (on the p-doped substrate) Ni/Zn/Au back contacts are deposited. After deposition the samples are annealed at 400 °C for 3 minutes in Nitrogen atmosphere.

Fig. 6.2a shows the AFM picture of QD-Bar1 up to the QD layer. The QDs have the literal size of 28 (±4) nm, and the height of 2.3 (±0.4) nm with density of $\sim 1.5 \times 10^{11}$ cm^2. Fig. 6.2b shows the transmission electron microscopy (TEM) picture of QD-Bar1. The cross section of the QD layer is clearly visible.

6.2 Conventional DLTS

Conventional DLTS have been performed on all the samples in order to obtain the mean activation energies and the mean capture cross sections of the samples. For all samples, the pulse voltage V_p was set to 0 V and the measurement voltage V_m to 6 V (reverse direction). The choice of the voltages is based on the CV measurements. The pulse length is 0.2 s and the measurement time is 1.8 s for Only-QD, while they are 0.5 s and 2.5, respectively, for Only-Bar, and for QD-Bar1 and QD-Bar2, 2 s and 8 s, respectively. The pulse length and the measurement time are increased for QD-Bar1 and QD-Bar2, because much more time is needed to fill and deplete the QDs as they are expected to have larger activation energy due to the presence of additional barriers. The temperature was swept from 50 K to 400 K in 5 K steps for all the samples. The capacitance transients were measured at a frequency of 1 MHz and an AC voltage of 100 mV. The transients were recorded 20 times to enhance the noise-signal ratio. Results were analyzed with the double-boxcar method. The reference times τ_{ref} for all the samples ranges between 0.5 s and 5 s.

The DLTS signals of the samples for $\tau_{ref} = 1.6$ s are shown in Fig. 6.3. Only-QD sample presents one single peak centered around 110 K (black line). For the Only-Bar a single peak is centered around 230 K shown (red line). QD-Bar1 and QD-Bar2 have similar DLTS signals shown (green and blue lines), respectively. A single peak is centered around 350 K for QD-Bar1, while a single peak is centered around 360 K for QD-Bar2. These DLTS spectra suggest that QD-Bar1 and QD-Bar2 have larger

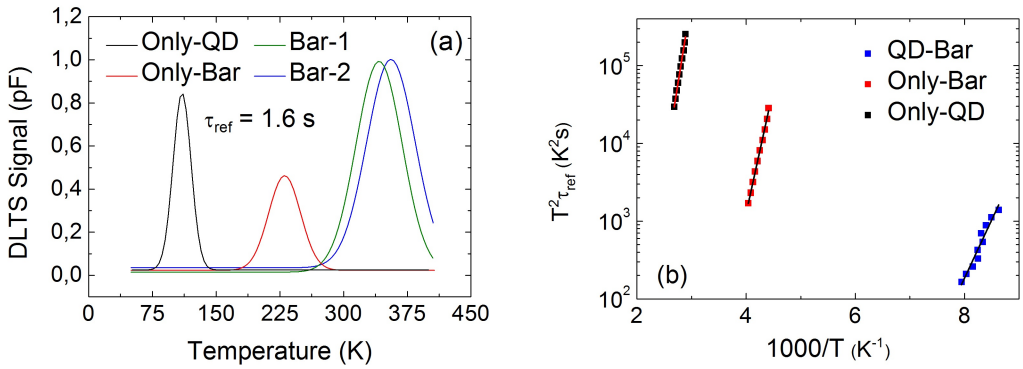

Figure 6.3: (a) Conventional DLTS results of all InGaSb/GaP samples for $\tau_{ref} = 1.6$ s. (b) Arrhenius plots of the same measurements.

localization energy in comparison to Only-QD and Only-Bar samples as expected due to the additional barriers.

The activation energies and the capture cross sections are extracted from the Arrhenius plot and listed in Tab. 6.1 with their error margins. Only-QD has the smallest activation energy, $E_a = 0.290$ (± 0.006) eV with an apparent capture cross section $\sigma_\infty = 5 \times 10^{-12}$. Only-Bar has no QD layer, however, its activation energy represents the height of the energy barrier formed by AlP on GaP. It is determined to be $E_a = 0.63$ (± 0.01) eV and this value is in agreement with values available in literature, which range from 0.34 eV to 0.69 eV [25]. The activation energies for QD-Bar1 and QD-Bar2 samples are determined to be $E_a = 0.86$ (± 0.01) eV and $E_a = 0.96$ (± 0.02) eV, respectively. Their capture cross sections are $\sigma_\infty = 1 \times 10^{-14}$ and $\sigma_\infty = 5 \times 10^{-13}$, respectively. It should be noted that the activation energies for QD-Bar samples represent the sum of the activation energy of the QD and the barrier. Therefore, the activation energies of QD-Bar samples are consistent with the sum of the activation energies of Only-QD and Only-Bar samples (i.e. 0.92 (± 0.01) eV).

6.3 Charge Selective DLTS

The charge selective DLTS measurement were performed on all the samples to determine their localization energies and capture cross sections, except Only-Bar sample since it does not have a QD layer, thus no internal energy levels. The pulse and the

Sample	E_a (eV)	σ_∞ (cm^2)
Only-QD	0.290 (± 0.006)	5×10^{-12} (0.1)
Only-Bar	0.63 (± 0.01)	7×10^{-12} (0.1)
QD-Bar1	0.86 (± 0.01)	1×10^{-14} (0.1)
QD-Bar2	0.96 (± 0.02)	5×10^{-13} (0.1)

Table 6.1: Conventional DLTS results of the samples.

measurement voltages are chosen based on the CV measurement for each sample.

6.3.1 Only-QD

The measurement voltage V_m ranges from 1 V to 6 V in 1 V steps, whereas pulse voltage $V_p = V_m - 1$ V. The pulse time is 0.5 s and transients are recorded for 2.5 s. Each transient is recorded 30 times to increase the signal-noise ratio. The temperature is swept from 50 K to 400 K in 5 K steps. The analysis is performed via double-boxcar method with the setting of t_1 varies from 30 ms to 300 ms, $t_2 = 10.t_1$ and t_{ave} =0.5 yielding reference time τ_{ref} between 0.1 s and 1.2 s.

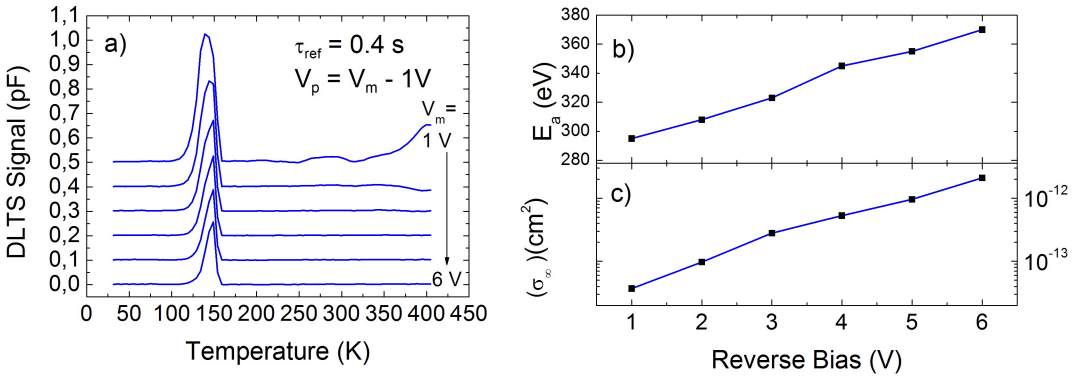

Figure 6.4: (a) Charge selective DLTS result of only-QD for $\tau_{ref} = 0.4$ s, (b) activation energies, and (c) capture cross sections

Fig. 6.4a shows the DLTS spectra for $\tau_{ref} = 0.4$ s. A single peak for each measurement voltage can be seen. The peak is centered around 140 K at $V_m = 1$ V and slightly shifts towards higher temperature with increasing measurement voltage. At $V_m = 6$ V the peak is centered around 150 K. It should be noted that the shift in temperature depending on the measurement voltage is small. It can be justified that the internal energy levels of the sample are localized very closely.

An activation energy and a capture cross section are determined for all the measurement voltages. Accordingly, the activation energy starts from 0.295 ($\pm$0.006) eV at $V_m = 1$ V and increases with the increasing measurement voltage. At the last measurement voltage (at $V_m = 6$ V), which represents the localization energy of the sample, it is determined to be 0.370 ($\pm$0.008) as shown in Fig. 6.4b. The capture cross section starts from 4×10^{-14} cm^2 (with an uncertainty of 0.1 orders of magnitude) at $V_m = 1$ V and increases until 2×10^{-12} cm^2 (with an uncertainty of 0.1 orders of magnitude) at $V_m = 6$ V as shown in 6.4c.

To conclude, the localization energy of the sample is determined to be 0.370 ($\pm$0.008) with an associated capture cross section of 2×10^{-12} cm^2 with an uncertainty of 0.1 orders of magnitude.

6.3.2 QD-Bar1

The measurement voltage V_m ranges from 1 V to 6 V in 1 V steps, whereas pulse voltage $V_p = V_m - 1$ V. The pulse time lasts 0.5 s, while measurement time lasts 8 s. Each transient is recorded 30 times to enhance the signal-noise ratio. The temperature is swept from 250 K to 400 K in 5 K steps. The setting in double-boxcar method is that t_1 varies from 50 ms to 500 ms, $t_2 = 10.t_1$ and $t_{ave} = 0.5$. The reference time τ_{ref} ranges between 0.2 s and 2.5 s.

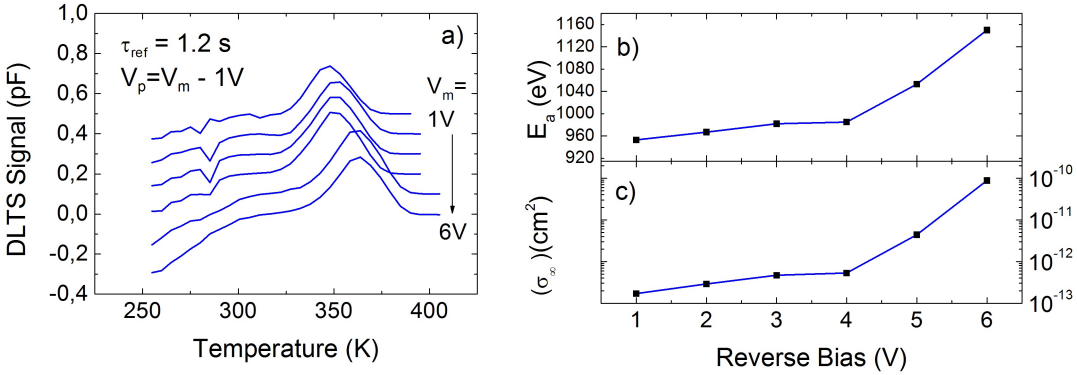

Figure 6.5: (a) Charge Selective DLTS result of the QD-Bar1 sample for the reference time $\tau_{ref} = 1.2$ s, (b) activation energies, and (c) capture cross sections

Fig. 6.5a shows the DLTS spectra for $\tau_{ref} = 1.2$ s. A single peak for each measurement voltage can be seen. The peak is centered around 340 K at $V_m = 1$ V and remain almost unchanged until $V_m = 5$ V, then it shifts towards higher temperature. At $V_m = 6$ V the peak is centered around 360 K. This behavior is similar to the previous sample's behavior that the DLTS signal slightly shifts towards higher temperature due to the closely-spaced internal energy levels.

An activation energy and a capture cross section are determined for all the measurement voltages. The activation energy starts from 0.953 ($\pm$0.01) eV at $V_m = 1$ V and increases with the increasing measurement voltage. At $V_m = 6$ V, which represents the localization energy of the sample, is determined to be 1.15 ($\pm$0.02) (see Fig. 6.5b). The capture cross section shows the similar behavior. It starts from 2×10^{-13} cm^2 (with an uncertainty of 0.2 orders of magnitude) at $V_m = 1$ V and ends up of 9×10^{-11} cm^2 (with an uncertainty of 0.2 orders of magnitude) at $V_m = 6$ V as shown in 6.5c.

To sum up, the localization energy of the sample is determined to be 1.15 ($\pm$0.02) with the associated capture cross section with 9×10^{-11} cm^2 with an uncertainty of 0.2 orders of magnitude.

6.3.3 QD-Bar2

The same parameters used for QD-Bar1 are used for QD-Bar2. Accordingly, the measurement voltage V_m ranges from 1 V to 6 V in 1 V steps, whereas pulse voltage $V_p = V_m - 1$ V. The pulse time is 0.5 s, while measurement time is 8 s. Each transient

is recorded 30 times to improve the signal-noise ratio. The temperature range is from 250 K to 400 K in 5 K steps. The setting in double-boxcar method is that t_1 varies from 50 ms to 500 ms, $t_2 = 10.t_1$ and t_{ave} =0.5. The reference time τ_{ref} ranges between 0.2 s and 2.5 s.

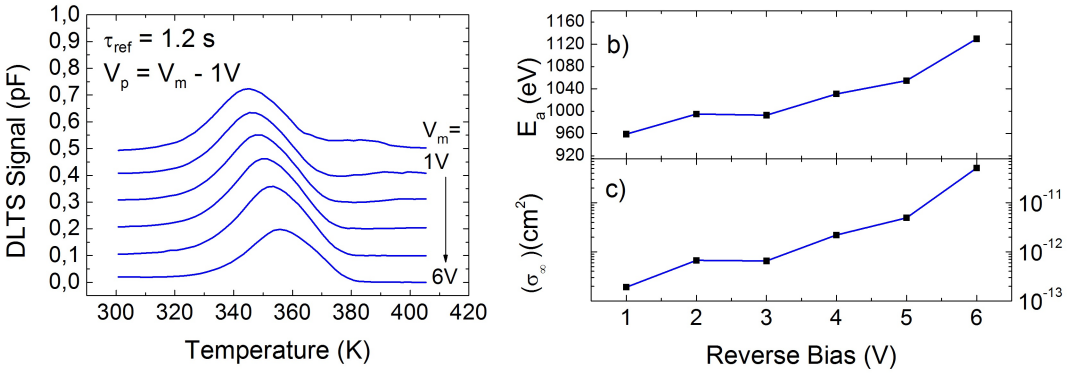

Figure 6.6: (a) Charge Selective DLTS result of the QD-Bar2 sample for the reference time $\tau_{ref} = 1.2$ s, (b) activation energies, and (c) capture cross sections

Fig. 6.6a shows the DLTS spectra for $\tau_{ref} = 1.2$ s. A single peak for each measurement voltage can be seen. The peak is centered around 345 K at $V_m = 1$ V and increases with increasing measurement voltage towards higher temperature. At $V_m = 6$ V, the peak is centered around 355 K. Similar to the trend, amount of the shift in temperature range is small justified with having closely-spaced internal levels. It is concluded that this behavior originates from the material itself based on the results of all the samples and the results of the similar structures [19, 87].

An activation energy and a capture cross section are determined for all the measurement voltages. The activation energy starts from 0.959 ($\pm$0.01) eV at $V_m = 1$ V and increases with the increasing measurement voltage. At $V_m = 6$ V it is determined to be 1.13 ($\pm$0.02) as shown in Fig. 6.6b. The capture cross section shows a similar behavior. It starts from 2×10^{-13} cm^2 (with an uncertainty of 0.2 orders of magnitude) at $V_m = 1$ V and ends up of 5×10^{-11} cm^2 (with an uncertainty of 0.2 orders of magnitude) at $V_m = 6$ V as shown in 6.5c.

To sum up, the localization energy of the sample is determined to be 1.13 ($\pm$0.02) with the associated capture cross section of 5×10^{-11} cm^2 with an uncertainty of 0.2 orders of magnitude.

6.4 Storage Time

Tab. 6.2 shows the localization eneries E_a, the capture cross sections σ_∞, and the storage times τ for holes at room temperature for Only-QD, QD-Bar1, and QD-Bar2 with their error margins.

Localization energy of Only-QD is determined to be 0.370 ($\pm$0.008) with an associated capture cross section of 2×10^{-12} cm^2 (with an uncertainty of 0.1 orders of

Sample	E_a (meV)	σ_∞ (cm^2)	τ (s) (at 300K)
Only-QD	0.370 ($\pm$0.008)	2×10^{-12} (0.1)	8×10^{-9} (0.1)
QD-Bar1	1.15 ($\pm$0.02)	9×10^{-11} (0.2)	3200 (0.1)
QD-Bar2	1.13 ($\pm$0.02)	5×10^{-11} (0.2)	2500 (0.1)

Table 6.2: Charge Selective DLTS and the storage times

magnitude). Using these values, the storage time for holes at room temperature is calculated to be 8×10^{-9} s (with an uncertainty of 0.1 orders of magnitude). For QD-1, the localization energy is determined to be 1.15 ($\pm$0.02) with the associated capture cross section with 9×10^{-11} cm^2 (with an uncertainty of 0.2 orders of magnitude), yielding a storage time of 3200 s ($\sim$ 1 hour) (with an uncertainty of 0.1 orders of magnitude). For QD-Bar2, the localization energy is determined to be 1.13 ($\pm$0.02) with the associated capture cross section with 5×10^{-11} cm^2, yielding a storage time of 2500 s (with an uncertainty of 0.1 orders of magnitude).

To sum up, the longest storage time for holes at room temperature is calculated for QD-Bar1 sample. Considering that this is the first attempt ever at QD-based storage in InGaSb/GaP, a storage time of $\sim$ 1 hour is a promising result.

6.5 Discussion

In this chapter, the results of the InGaSb/GaP structures were presented. This material combination is novel and reported for first time in this work. During the growth of the samples, C atoms are used as dopants for GaP, for first time, to prevent us from the drawbacks induced by Zn atoms, which is most used dopant for GaP. The another importance of the structure is that this material combination is a good candidate to achieve non-volatility due to its projected high localization energy [25, 150] and the experimental results on similar structures [19, 87].

Four different structures were investigated. First one contains only an InGaSb QD layer embedded in GaP matrix, named Only-QD. Two other structures, named QD-Bar1 and QD-Bar2, contain InGaSb QD layers embedded in GaP matrix with additional AlP barriers. The aim to use an additional barrier is to increase the total localization energy of the QD for holes (in this case the total localization energy equals to the sum of the QD localization energy and the barrier height), thus, to further increase the storage time. The difference between the mentioned two samples is the Sb-flush time, which takes place prior to the QD formation to alter the optical and electrical properties of QDs. The Sb-flush time is 2 s for QD-Bar1, while 4 s for QD-Bar2. The last sample, named Only-Bar, contain only AlP barrier without a QD layer serving as a reference sample.

The localization energy of Only-QD is determined to be 0.370 ($\pm$0.008) with an associated capture cross section of 2×10^{-12} cm^2 (with an uncertainty of 0.1 orders of magnitude). Although no experimental result was reported on the InGaSb/GaP quantum dot system before, the IFIGS-and-electronegativity method can provide a good estimate. Accordingly, it yields values of $\sim$ 0.3 and $\sim$ 0.6 eV for the valence

band offset for the QD systems of $In_{0.5}Ga_{0.5}As/GaP$ and $In_{0.5}Ga_{0.5}Sb/GaP$, respectively [160, 161]. Therefore the measured value of $E_{loc} = 0.370(\pm 0.008)$ eV is within the predicted limit. The storage time of the sample for holes at room temperature is then calculated to be 8×10^{-9} s (with an uncertainty of 0.1 orders of magnitude). For Only-Bar the localization energy is determined to be 0.63 ($\pm$0.01) eV with the associated capture cross section of 7×10^{-12} (with an uncertainty of 0.1 orders of magnitude). The energy barrier for InGaSb/GaP structure is available in literature varying from 0.34 eV to 0.69 eV [25]. Therefore, our result is in agreement with the projected values.

The localization energy for QD-Bar samples represents the sum of the Only-QD and Only-Bar samples. Localization energy of QD-Bar1 is determined to be 1.15 ($\pm$0.02) with the associated capture cross section of 9×10^{-11} cm^2 (with an uncertainty of 0.2 orders of magnitude). This localization energy is consistent with the results of Only-QD and Only-Bar (the sum of their localization energies is 1.00 ($\pm$0.01) eV). The storage time of QD-Bar sample at room temperature is calculated to be 3200 s ($\sim$ 1 hour) (with an uncertainty of 0.1 orders of magnitude), marking an improvement of one order of magnitude to respect with the reported similar structures[3] whose storage time was obtained to be 230 s [19]. The localization energy for said structure was 1.14 ($\pm$0.04) eV with an associated capture cross section of 8×10^{-10}. Regarding the fact that their localization energies are identical within their error margin, the increase in the storage time originates from the reduction (one order of magnitude) in the capture cross section of the sample in this work, due to the use of Sb instead of As. Localization energy of QD-Bar2 is determined to be 1.13 ($\pm$0.02) with the associated capture cross section with 5×10^{-11} cm^2 (with an uncertainty of 0.2 orders of magnitude), which is consistent too with Only-QD and Only-Bar samples. The storage time of the sample at room temperature is calculated to be 2500 s (with an uncertainty of 0.1 orders of magnitude). The only difference between QD-Bar1 and QD-Bar2 is the Sb-flush time prior to the QD formation to alter its optical and electrical features. However, there is no significant difference in their localization energy and capture cross sections regarding their error margins. Thus, the storage time could not be further improved for QD-Bar2.

Highest localization energy for this batch is obtained to be 3200 s ($\sim$ 1 hour) (with an uncertainty of 0.1 orders of magnitude). Although this value is far from the desired storage time of 10 years, it makes an improvement of one order of magnitude in comparison to the similar structure [19, 87]. Moreover, this material system is completely novel and grown for first time. Almost one hour of storage time is promising result for such a novel system. With some alterations, this value can be further improved. To achieve non-volatility with this sample, it would be sufficient to decrease its capture cross section to 1×10^{-15} cm^2 with the same localization energy of 1.15 eV. Alternatively, having the same capture cross section of 9×10^{-11}, the localization energy has to be increased to 1.4 eV.

In Chapter 5 the GaSb/GaP structures grown under different growth conditions were presented. There, we demonstrated that the electronic properties of QDs such localization energy and capture cross section depend strongly on growth conditions

[3] MOCVD-grown $In_xG_{1-x}aAs/GaP$ QDs

and they can be altered by changes in growth conditions such as growth temperature and growth rate. Besides, Sb-soaking [151, 152] and growth interruption [162, 163] methods can be employed for localization energy and capture cross section engineering. The InGaSb/GaP structures similar to the DLTS structures were grown and photo-luminescence measurements were performed on them. Based on the results, it is concluded that the Sb-soaking prior to QD deposition reduces the QD dimensions, while increasing their density. Hence, smaller capture cross sections can be achieved leading to longer storage times. Results also suggest that growth interruption increases the size of QDs as they gain time to ripen, leading to higher localization energies, thus longer storage times.

To conclude, the presented storage time of 3200 s (with an uncertainty of 0.2 orders of magnitude) is promising considering the novelty of the structure. Therefore, non-volatility can likely be achieved with this material combination.

Summary

- The InGaSb/GaP QD system is a novel system, which was reported for first time in this work. This material combination is a good candidate to achieve non-volatility, at least 10 years of storage time at room temperature, due to its high localization energy.
- A sample consisting of a InGaSb QD layer in GaP matrix with an additional AlP (QD-Bar1) is grown and investigated electrically. To compare, a sample containing only a InGaSb QD layer (Only-QD) in GaP matrix and another sample containing only an AlP barrier (Only-Bar) are grown. Additionally, an identical structure to the first one with longer Sb-flush time (QD-Bar2) is grown to examine the effect of Sb-soaking on electronic properties of QDs such as localization energy and capture cross section.
- The localization energy values determined for Only-QD and Only-Bar samples are in good agreement with the values in literature. For QD-Bar samples the localization energy represents the sum of the QD localization energy and the barrier height. The localization energy values obtained for QD-Bar samples equal the sum of the localization energy of Only-QD and Only-Bar making all the results self-consistent.
- Highest storage time is obtained to be 3200 ($\pm$300) s for QD-Bar1. The storage time of approximately 1 hour is a promising value for such a novel system. Moreover, this value marks an improvement of one order of magnitude in comparison to the reported similar structures.
- Non-volatility can be achieved with this material combination either by increasing its localization energy or by decreasing its capture cross section. For this, different growth conditions or techniques such as Sb-soaking and growth interruption can be employed.

Chapter 7

RT Results

In this section simulation results of the resonant tunnel structures suggested to solve the trade-off between storage and erase times are presented. First, the results for the basic structures with only quantum wells, and then for the realistic structures with quantum dots, quantum wells, and 2DHG layers are presented. The aim of designing a simple structure is to demonstrate that resonances can be induced and observed. Realistic structures are fully optimized structures which can carry out all memory operations.

Parts of this chapter have been accepted by *Physical Status Solidi A*[1].

7.1 Simple Structures

Fig. 7.1a shows the structural depiction of the sample. A 300 nm highly p-doped ($N_a = 10^{18}$ cm^{-3}) GaAs layer is deposited on the substrate to ensure that the back contact is Ohmic. After a 10 nm p-doped GaAs layer, the quantum well structure is grown. The quantum well structure consists of two undoped GaAs quantum wells separated by three $Al_{0.9}Ga_{0.1}As$ barriers. The width of the quantum wells are 4 nm and 3 nm, respectively. Above the superlattice a 10 nm layer of p-doped GaAs layer is grown. Finally, the structure is completed with a 100 nm highly p-doped ($N_a = 10^{18}$ cm^{-3}) GaAs layer to realize an Ohmic top contact.

The sample is processed by standard optical lithography and wet etching methods in a clean room forming 400 μm and 800 μm mesas. Ni, Au/Ge, and Au are used to achieve the top contact; Ni, Zn, and Au for the back contact. After the contacts are deposited, the sample is annealed at 400 C$^\circ$ for 3 minutes in a nitrogen atmosphere.

7.1.1 Simulation Results

The final aim in the simulation is to calculate the current density versus the voltage. The peaks attributed to the resonances can be observed in the current density-voltage profile. The effective mass used for GaAs are 0.48 m_e for heavy holes and 0.082 m_e for

[1] I. F. Arikan, N. Cottet, T. Nowozin, and D. Bimberg, "Transparency engineering in quantum dot-based memories."

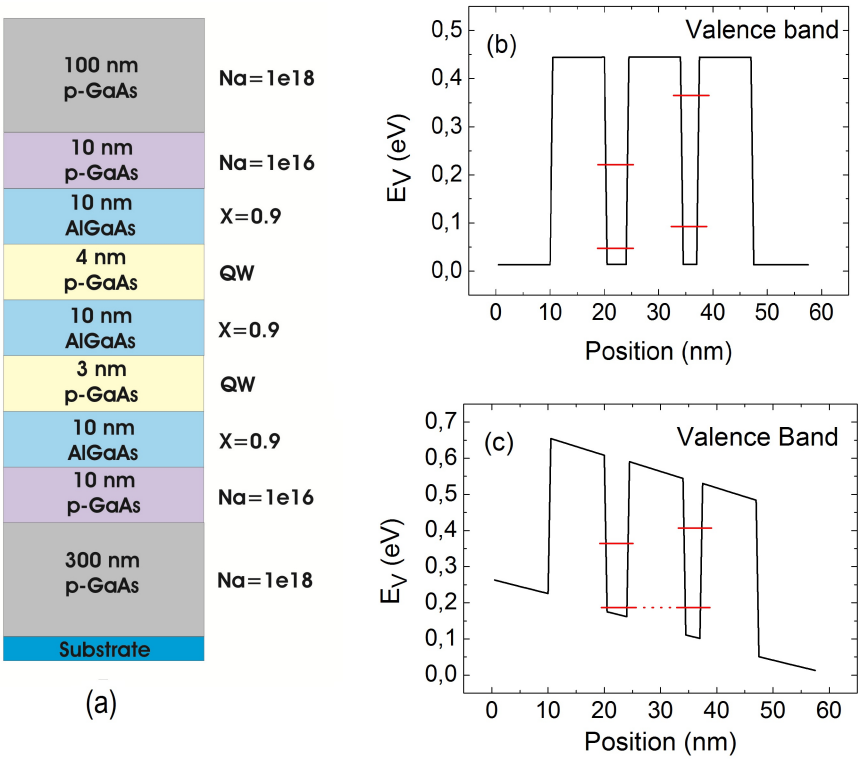

Figure 7.1: Simple structure with two undoped GaAs QWs separated by three $Al_{0.9}Ga_{0.1}As$ barriers. (a) Epitaxial structure. (b) Inverted valence band of the structure out of the resonance voltage. As the widths of the QWs are different, the energy levels lie in different positions. (c) Inverted valence band of the structure at resonance voltage where the ground states of the QWs are aligned.

light holes [125].

First, we calculate the electronic properties of the structure such as the conduction and valence bands, and the doping profile by one-Dimensional Schrödinger-Poisson Solver (see Chapter 3). Fig. 7.1b shows the inverted valence band of the structure calculated by one-DSPS out of resonance voltages. As the QWs have different well widths, the energy levels are located in different positions. The valence band of the structure can be bent by the applied voltage resulting in the shift of the energy levels. At the resonance voltage the energy levels of the QWs are aligned as shown in Fig. 7.1c.

After the band calculation of the structure, the current density versus the applied voltage is calculated via Non-Equilibrium Green's Function. Based on the transparency calculation model explained in Chapter 3, a channel sandwiched between two carrier reservoirs is assumed and the current from the first contact to the second contact is calculated. It should be noted that current flows from one contact to another, only if the chemical potential of the contacts are different [126]. The calculated energy levels of the structure versus the applied voltage for light and heavy holes are depicted in Fig. 7.2a and Fig. 7.2b, respectively. There is one intersection point for light holes

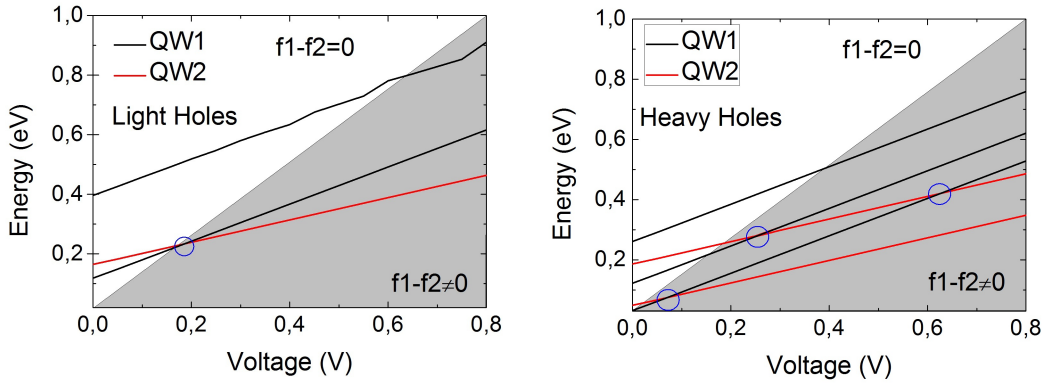

Figure 7.2: (a) Calculated energy levels versus the voltage for light holes. There is one intersection point between the energy levels of the QWs attributed to the resonance. (b) Calculated energy levels versus the gate voltage for heavy holes. There are three different intersection points between the energy levels of the QWs attributed to the resonance.

between the energy levels of the QWs attributed to the resonance, while there are three intersection points for heavy holes since a heavier mass reduces the spacing between energy levels, hence higher number of energy levels in a given volume. Fig. 7.3 shows the calculated current density versus applied voltage for both light (top) and heavy (bottom) holes at 40 K. The peaks can be clearly seen at the intersection voltages attributed to the resonance. There is only one peak for light holes, while there are three peaks for heavy holes.

The simulation demonstrates that the current density increases at resonance voltage due to the higher transparency. This can be used to reduce the erase time in QD-Flash.

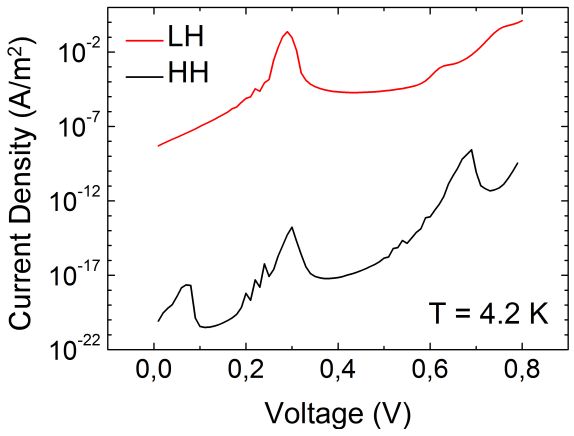

Figure 7.3: Calculated current density versus applied voltage for both light and heavy holes. The peaks can be seen at the resonance voltages.

7.1.2 Experimental Results

The current-voltage measurement (IV) is performed on the heavy hole sample at different temperature to demonstrate the resonances experimentally. Fig. 7.4a shows the IV profile of the sample at different measurement temperatures. Three peaks can be clearly seen at resonance voltages. The smaller shifts in voltage compared to the simulation can be attributed to the simplifications in the calculation and the differences between the nominal structure and the actual structure of the grown samples. However, deviation of the thickness or composition of the QWs does not change the calculated transmission coefficient. Small changes of the mentioned parameters result in a small shift of the bias voltage where the resonance takes place. The results also show the temperature dependence of the process. The peaks attributed to the resonance shrink gradually with increasing temperatures and completely disappear after 200 K, as thermal emission becomes dominant at moderate and higher temperatures.

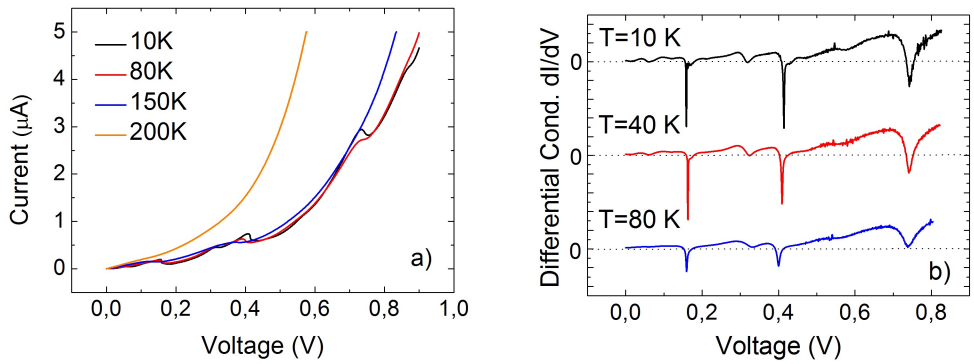

Figure 7.4: Result of the current-voltage measurement for the heavy holes. (a) The current voltage profile of the sample at different temperatures. Three peaks appear, indicating the resonance voltages. The peaks shrink with increasing temperature and completely disappear at 200 K. (b) The first derivative of the conductance. The change in the conductance at resonance voltage are more visible.

The position of the peaks can be determined more precisely using the first derivative of the IV profile, as shown in Fig. 7.4b.

7.2 Realistic Structures

Realistic structures are fully optimized structures containing a quantum dot layer and a two-dimensional carrier gas layer in addition to the quantum wells. All memory operations can be carried out in the realistic structures.

Two different material combinations are designed: the first group is based on GaAs, while the second group is based on GaP. Although it has been demonstrated that GaAs-based material systems are not suitable for memory applications because of their small localization energies [134], the first batch of samples is based on GaAs

because it has been studied extensively before and is therefore very well understood. Moreover, the many years of experience with the material warrant a straightforward and reliable growth. The second batch of samples was grown on GaP because, even though the growth is much more challenging, GaP-based materials have the potential to achieve non-volatility due to their projected high localization energy [25, 146, 147] and previous experimental results [19, 87].

Simulations were run for both light and heavy holes. The effective masses used for GaAs are 0.082 m_e for light holes and 0.48 m_e for heavy holes. However, only the results for light holes are presented in following. The process is same for heavy holes. Only difference is that energy levels of QDs are closely spaced due to their heavier masses, resulting in a shift in the position of resonance voltages.

Samples based on these structures were grown epitaxially at TU Berlin, but unfortunately the high defect density prevented electrical characterization. Further growth runs were not possible due to time constraints. Therefore, only simulation results are presented for the realistic structures.

7.2.1 GaAs-based structures

Sample Structures

Four samples have been designed based on GaAs: a sample with one QW, a sample with two QWs, and their reference samples with additional barriers instead of QWs. A schematic of the one-QW-sample is shown in Fig. 7.5a. The basic structure is a modulation-doped field-effect transistor (MODFET). The QW is introduced between the QD layer and the 2DHG. On top of an undoped substrate and a 1000 nm nominally undoped GaAs layer, a 40-nm-wide p-doped layer ($p = 1 \times 10^{18}$ cm^{-3}) is introduced to provide holes. After a spacer layer of 7 nm undoped GaAs, the 2DHG is formed in an 8-nm-wide $In_{0.25}Ga_{0.75}As$ QW, on top of which the QW sandwiched between two $Al_{0.9}Ga_{0.1}As$ barriers is placed. The InAs QD layer is placed on top of the QW, separated by 5 nm of undoped GaAs, since the formation of InAs QDs directly on top of an $Al_{0.9}Ga_{0.1}As$ surface would be difficult due to surface roughness. The localization energy for holes for InAs QDs embedded in a GaAs matrix is assumed to be 210 meV, which was experimentally determined [134]. After the QD layer, the device is completed by another 180 nm of undoped GaAs. The 2DHG is contacted via two Ohmic source and drain contacts and the energy of the QD hole levels can be tuned by applying a bias voltage to the gate contact.

For the two-QW-sample the superlattice has a similar design, just adding an additional QW and a barrier shown in Fig. 7.5d. To compare the performance with regular samples which do not have a superlattice structure, one reference sample is also simulated for each case. In the reference sample, the GaAs QWs are replaced by $Al_{0.9}Ga_{0.1}As$ barriers, so that a single barrier is formed, shown in Fig. 7.5c and Fig. 7.5e.

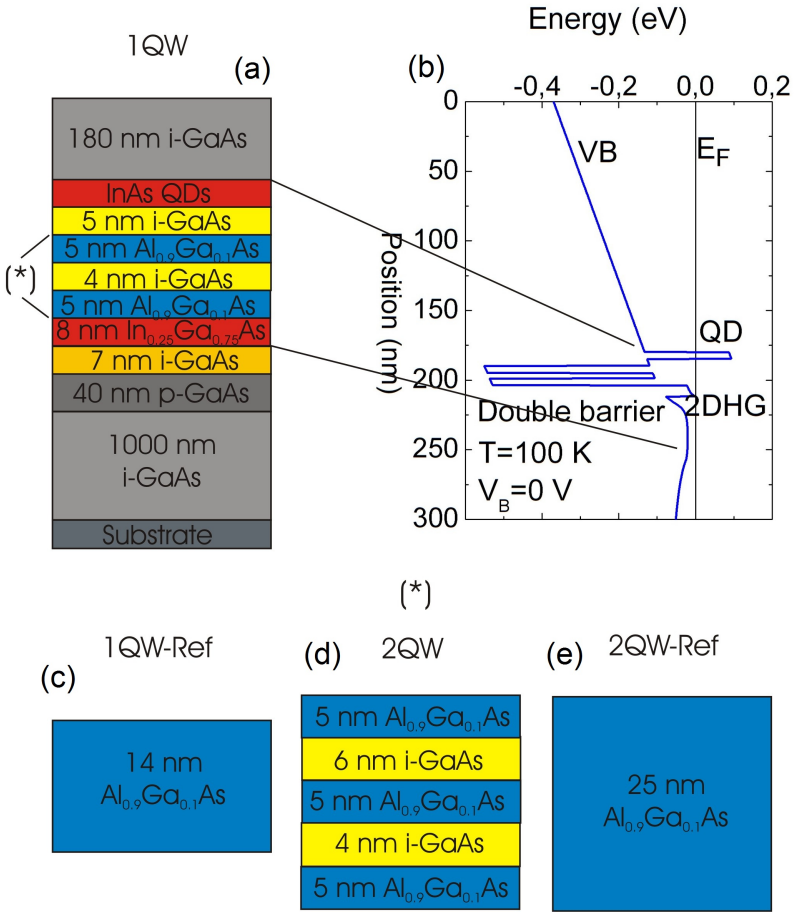

Figure 7.5: (a) Schematic of the one-QW-sample based on GaAs and (b) its valence band diagram. (c) Reference sample of one-QW-sample. (d) The two-QW-sample with an additional QW and a barrier. (e) Reference sample of two-QW-sample. In reference samples QWs are replaced with extra barriers.

Results

1-QW

Fig. 7.6a shows the calculated energy levels of the QD and the QW as function of applied voltage for the one-QW-sample. An intersection point between the energy levels of the QD and the QW can be seen at 3.94 V, which represents the alignment between the energy levels throughout the device. At this resonance voltage the transparency increases by 7 orders of magnitude as shown in Fig. 7.6b as a black line. The same figure also shows the transparency for the reference sample as a blue line. No additional increase in transparency is found for the reference sample. Its transparency increases smoothly with larger reverse bias due to the reduction in barrier height.

The results demonstrate that an increase in transparency can be achieved by replacing a wide barrier by a QW, leading to a faster erase time since the erase mechanism relies on tunneling.

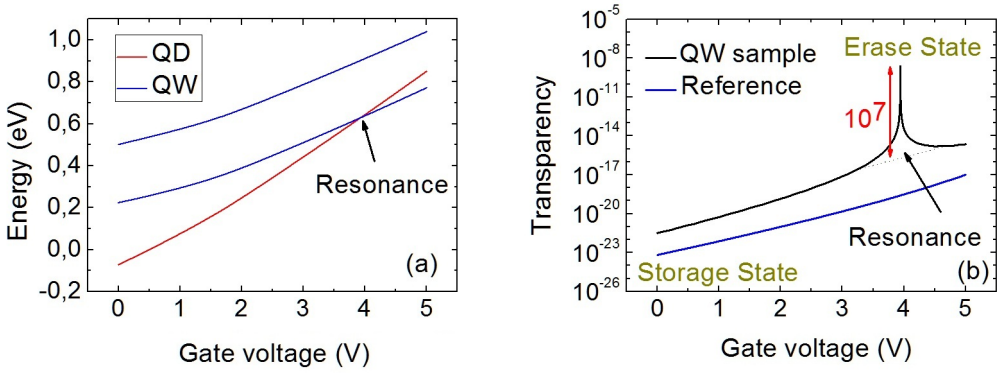

Figure 7.6: (a): The energy levels of the QD and the QW versus gate voltage for the one-QW-sample. There is an intersection point at 3.94 V attributed to the resonance. (b): The transparency versus gate voltage for both the one-QW-sample and the reference sample. The transparency increases by 7 orders of magnitude at the resonance voltage for the one-QW-sample, while there is no increase in transparency for the reference sample.

2QWs

Similar structure is designed with an additional quantum well and barrier layers. The advantage of multiple quantum well structures is that they yield higher storage/erase time ratio, since the tunnel emission cannot take place out of resonant voltage resulting in a longer storage time. Additionally, they provide the advantage that a thick barrier can be divided into several thin barriers of a given length, hence larger transparency coefficient is obtained, only if all the quantum well share at least one eigenvalue, as the transparency depends strongly on the barrier width. However, the disadvantage carried by additional quantum wells is that they induce more energy levels, which have to be all aligned in order to produce resonant tunneling.

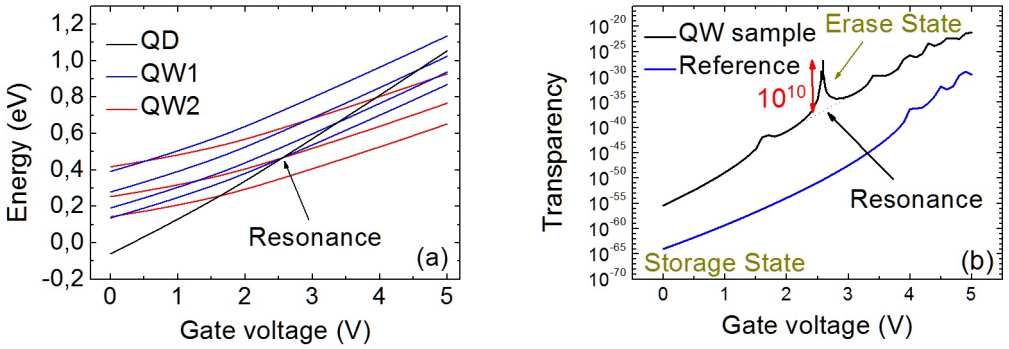

Figure 7.7: (a) The energy levels of the QD and the QWs versus gate voltage for the two-QW-sample. There is one intersection point at 2.56 V attributed to the resonance. (b) The transparency versus gate voltage for both the two-QW-sample and its reference. The transparency increases by 10 orders of magnitude at resonance voltage for the QW-sample, while there is no increase in transparency for the reference sample.

The calculated energy levels of the sample versus gate voltage is shown in Fig. 7.7a. An intersection point attributed to resonance between the energy levels of QD and QWs can be seen at 2.56 V. The transparency for the sample increases by 10 orders of magnitude at resonance voltage as depicted in Fig. 7.7b as a black line, implying that the erase time can be decreased by 10 orders of magnitude at this specific voltage as a result of resonance. The same graph also shows the transparency for reference sample having an additional barrier instead of QW versus gate voltage. There is no any additional increase in transparency for the reference sample.

7.2.2 GaP-based Structures

Sample Structures

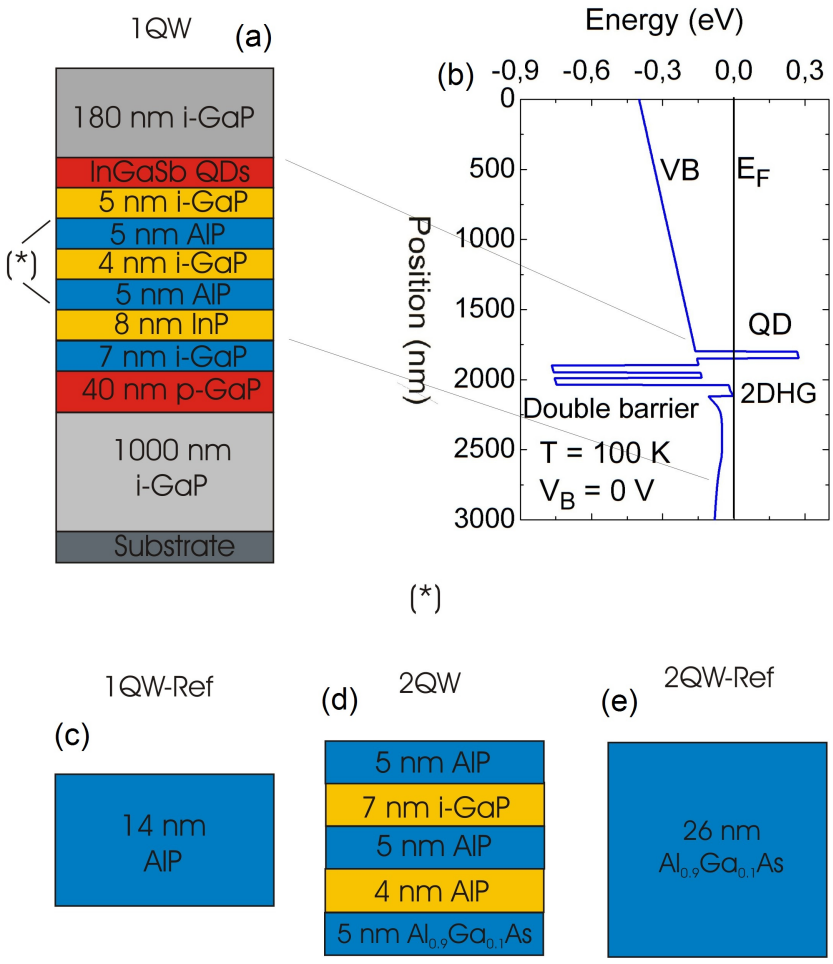

Figure 7.8: (a) Schematic of the one-QW-sample based on GaP and (b) its valence band diagram. (c) Reference sample of one-QW-sample. (d) The two-QW-sample with an additional QW and a barrier. (e) Reference sample of two-QW-sample. In reference samples QWs are replaced with extra barriers.

Four samples have been designed based on GaP. The structures of the GaP samples

are identical to those of GaAs with the difference that GaAs is replaced by GaP and $Al_{0.9}Ga_{0.1}As$ is replaced by AlP. Also, an InGaSb QD layer is used instead of the InAs QD layer and the 2DHG layer is formed by InP. More specifically, after a 1000 nm unpodep GaP layer on substrate, a 40 nm highly p-doped layer ($p = 1 \times 10^{18}$ cm^{-3}) is realized to provide holes. A 7 nm GaP spacer and the 2DHG of InP are introduced prior to the superlattice structure. Afterwards, the structure is completeed with the InGaSb QD layer and a 180 nm undoped GaP layer. The localization energy of the InGaSb QDs embedded in GaP matrix and the energy barrier between GaP and AlP are measured by DLTS and presented in Chapter 6. Accordingly, the localization energy of InGaSb QDs used in the simulations is 400 meV, whereas the energy barrier between GaP and AlP is set to 600 meV. It should be also noted that the InGaAs/GaP QD system can be also used as an alternative to the InGaSb/GaP QD system. Stracke *et al.* demonstrated that localization energy for $In_{0.25}Ga_{0.75}As$/GaP is 490 meV [87] and Bonato *et al.* reported that the localization energy $In_{0.5}Ga_{0.5}As$/GaP is 600 meV [19].

Parameters for the resonant tunnel region for GaP samples are the following: a 5 nm quantum well is sandwiched between two 5 nm wide barriers for the one-QW-sample; two quantum wells whose widths are respectively 6 and 4 nm are sandwiched between three 5 nm wide barriers for two-QW-sample. In the reference samples, the GaP QWs are replaced by AlP barriers.

Results

1QW

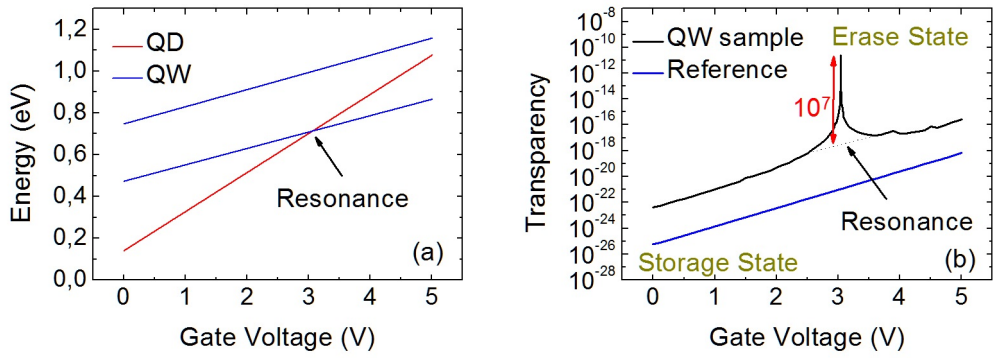

Figure 7.9: (a): The energy levels of the QD and the QW versus gate voltage. There is one intersection point at 3.04 V attributed to the resonance. (b): The transparency versus gate voltage for both the one-QW-sample and its reference. The transparency increases by 7 orders of magnitude at resonance voltage for the QW-sample, while there is no peak in transparency for the reference sample.

Fig. 7.9a shows the energy levels of the QD and the QW as a function of gate voltage for the one-QW-sample. An intersection point between the energy levels of QD and QW attributed to the resonance is located at 3.04 V. At the corresponding voltage

the transparency increases by 7 orders of magnitude as shown in Fig. 7.9(b) as a black line. In the same figure the transparency for reference sample containing only barrier versus gate voltage is depicted as the blue line. No additional increase in transparency can be seen.

The results clearly show that the transparency increases at resonance voltage as a result of resonance, hence the erase time decreases.

2QWs

The simulation was carried out also for the GaP sample with two QWs. The calculated energy levels of QD and QWs as a function of gate voltage for two-QW-sample is depicted in Fig. 7.10a. A resonance can be seen at 2.56 V. At the same resonance voltage the transparency increases by 10 orders of magnitude as shown in Fig. 7.10b as the black line. The same figure shows the transparency for the reference sample as the blue line. No extra increase is visible.

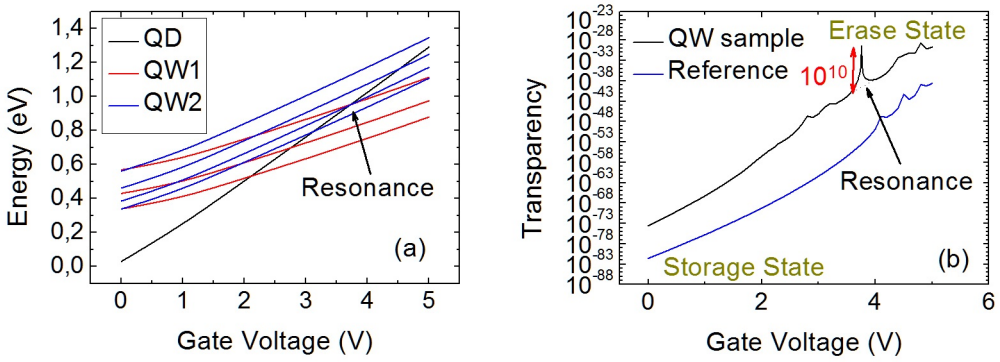

Figure 7.10: (a): The energy levels of the QD and the QWs versus gate voltage for the two-QW-sample. There is one intersection point at 2.56 V attributed to the resonance. (b): The transparency versus gate voltage for both two-QW-sample and its reference. The transparency increases by 10 orders of magnitude at the resonance voltage for the two-QW-sample, while there is no peak in transparency for the reference sample.

7.3 Discussion

In this chapter the simulation results of the resonant tunnel structures were presented. Resonant tunnel structures are proposed in this work to solve the trade-off problem between storage and erase times for QD-Flash. In QD-Flash an additional barrier is inserted into the structure to further increase the localization energy, hence the storage time. However, the additional barrier also decreases the tunneling probability. As the erase process relies on tunneling, a wider barrier results in an increase in erase time. In other words, longer storage time comes with longer erase time, i.e. trade-off. However, in a resonant tunnel structure, a very high transparency –ideally 1– can be obtained by

aligning the energy levels of the quantum wells throughout the device at a given voltage, thus a short erase time is achieved. Out of that resonance voltage, the transparency will be very low –ideally 0–, as the energy levels of the quantum wells are not aligned. Therefore, a long storage time is achieved.

We started the calculation with a simple structure since resonances can be easily observed in it. The simple structure consists of two GaAs quantum wells sandwiched between three $Al_{0.9}Ga_{0.1}As$ barriers. The widths of the QWs are 4 nm and 3 nm, respectively. The simulation was performed for both light and heavy holes. At the end their energy levels versus applied voltage and their conductance-voltage profiles are calculated. For light holes one intersection point between energy levels is observed, which is attributed to the resonance. There are three intersection points for heavy holes. Since a heavy mass leads to a deeper localization, hence many more bound states in a given volume, much more resonance for heavy holes occur. In the conductance-voltage profile a peak is observed at the resonance voltages. The result implies that transparency increases at resonance voltage leading to a higher conductance. The same structure is grown epitaxially for only holes and current-voltage measurement is performed on it. Three peaks attributed to resonance are also observed experimentally. Besides, it is demonstrated that the peaks shrinks gradually with increasing temperature and finally disappear, since the thermal emission becomes the dominant emission process and tunneling emission becomes negligible.

After demonstrating the resonance on the simplified structures, realistic structures are designed. The realistic structures contain a quantum dot layer, quantum wells, and a 2DHG layer. Two different material combinations are used in simulation: the GaAs-based and GaP-based materials. The GaAs-based material combination is designed since it is well understood and easily grown. The GaP-based material combination is designed, because non-volatility can be only achieved with them due to their high localization energy (see Chapter 5 and Chapter 6). For both material combinations four different samples are designed: a sample with one QW, a sample with 2 QWs and their reference samples with additional barriers instead of QWs.

The calculations show that the transparency increases by 7 orders of magnitude for the GaAs one-QW sample at the resonance voltage, where the energy levels of the QDs and QWs are aligned. For the two-QW-structures of GaAs the transparency increases 10 orders of magnitude at resonance voltage. Multiple-QW-structures yield the highest storage/erase time ratio, however, it is harder to achieve resonance due to the presence of much more energy levels to align.

In similar structures without quantum wells an erase times of a few ms is achieved [164]. The simulation implies that the implementation of QWs in QD-based memory structures will cause a shortening by 7 orders of magnitude in erase time. In this case, 1 ns of erase time will be achieved, which is faster than a typical DRAM erase time (10 ns). It should be noted, however, that experimental structures and simulated structures might not exactly be the same, e.g. broadening of the QD energy levels due to a Gaussian distribution of the QD sizes has not been taken into consideration. Nevertheless the increase in the transparency and the decrease in erase time should be of the same order of magnitude.

The simulations yielded similar results for GaP-based samples. An increase of 7 or-

ders of magnitude in transparency at resonance voltage is obtained for one-QW-sample, while 10 orders of magnitude increase is obtained for two-QW-sample. These results prove that the erase time can be decreased by insertion of QWs into the structures. Moreover, resonant tunneling can be achieved in GaP-structure, which is a promising material combination for non-volatility as shown by DLTS results in last two chapters.

The calculations also demonstrate that small changes of the thickness or composition do not affect the transparency coefficient, instead, result in a small shift of the bias voltage, where the resonance takes place. Therefore, the technology of the resonant tunneling memory is stable.

Summary

- Conductance-voltage profile is calculated for simple structure with two GaAs QDs sandwiched between three $Al_{0.9}Ga_{0.1}As$ barriers. At the resonance voltage the peaks attributed to resonances are shown.
- The simple structure is grown epitaxially and current-voltage measurement is performed on it. The peaks at resonance voltage are observed also experimentally.
- Realistic structures with a quantum dot layer and a two-dimensional hole gas layer besides quantum wells sandwiched between barriers are designed for GaAs and GaP structures. The importance of GaAs-based structures is that they are known well and their epitaxial growths are easy. On the other hand, GaP-based structures are good candidate to achieve non-volatility.
- The transparency increases by 7 orders of magnitude at resonance voltage for the GaAs sample including one QW, while it increases by 10 orders of magnitude for the GaAs sample with 2 QWs. This implies that the 10 ms erase time of actual structures without QWs can be decreased to nanoseconds.
- For GaP-based structures the transparency increases by 7 orders of magnitude for one-QW-sample and by 10 orders of magnitude for two-QW-structure. This demonstrates that resonant tunneling can be achieved in GaP-based materials, so that 10 years of storage time and nanosecond of erase time can be simultaneously realised.

Chapter 8

Conclusions and Outlook

This work had two goals, both aimed at improving the memory devices based on quantum dots: The first goal was to increase the storage time in QD-Flash, and the second goal was to decrease the erase time in said structures.

The key electronic properties of QDs such as localization energy and capture cross section are determined by static capacitance measurement and time-resolved measurement. Using the said parameters, the storage time for holes at room temperature is extrapolated. Since storage time depends strongly on localization energy and capture cross section, these parameters can be engineered to increase storage time. This study focuses on localization energy engineering based on the fact that our understanding of capture cross section is limited, and it is a complex parameter depending on many factors. The simplest way to increase localization energy is to add an additional barrier to the structure or to build QD systems based on new material combinations. Accordingly, the GaP-based material combinations emerge as promising candidates to achieve non-volatility due to their projected high localization energies [25, 146, 147]. Stracke *et al.* demonstrated a 3 μs of storage time in the $In_{0.25}Ga_{0.75}As$/GaP QD system [87], and Bonato *et al.* reported a 230 s of storage time in the $In_{0.5}Ga_{0.5}As$/GaP QD system [19], as shown in Fig. 8.1. Therefore, all the structures used in this work are the GaP-based material combinations.

The first batch investigated in this work is the GaSb/GaP material combination due to its projected high localization energy, 1.4 eV, based on the 8-band $\mathbf{k} \cdot \mathbf{p}$ calculation [146]. A storage time of 3.9 days is determined for the GaSb/GaP structure at room temperature. This storage time is the longest storage time ever reported and it marks an improvement of 3 orders of magnitude in comparison to the previous best result for holes in self-organized quantm dots [19]. Beside new material combination, the alterations in growth conditions can also increase the storage time. To investigate the effect of growth conditions on the electronic properties of QDs such as localization energy and capture cross section, hence on storage time, five different samples are grown under slightly different growth conditions. Two different growth parameters are varied: the growth temperature and the growth rate. Accordingly, it is demonstrated that an increase in the growth temperature has two different effects on QDs: a higher surface mobility leading to larger QDs, and a lower defect density leading to the QDs with higher structure quality. The results also show that slower growth rates allow

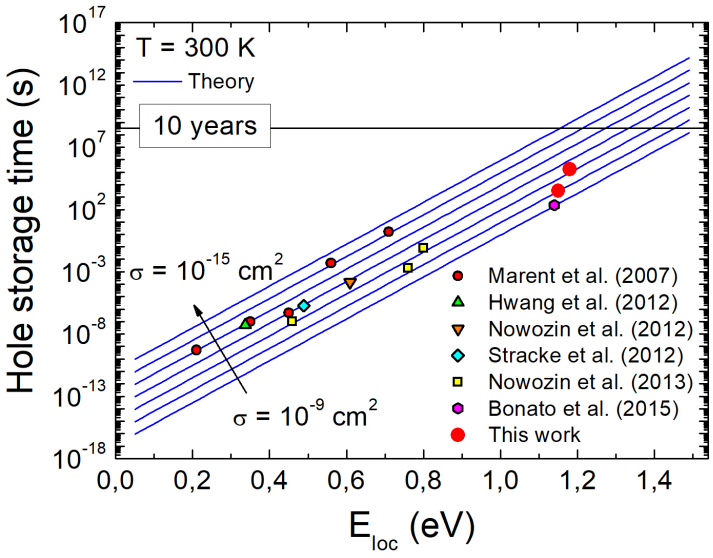

Figure 8.1: Storage time

enough time for QDs to ripen resulting in the formation of larger QDs. According to 8-band $\mathbf{k} \cdot \mathbf{p}$ model larger QDs yield higher localization energies, thus longer storage times (see Sec. 1.4.2). The sample which yielded the storage time of 3.9 days was grown with the most optimized parameters among the whole batch, i.e. the highest growth temperature and the slowest growth rate.

The second batch investigated in this work is the InGaSb/GaP QD system. Although no simulation is performed on the InGaSb/GaP QD system, its localization energy is estimated to be high [25, 150]. This material combination is hence grown and reported for first time in this work. A 3200 s storage time is obtained, marking an improvement of one order of magnitude with respect to similar structures, which are the MOCVD-grown InGaAs/GaP QD systems. The improvement of one order of magnitude in storage time originates from the decrease in capture cross section.

In addition to the alterations in growth conditions, techniques such as Sb-soaking [151, 152] and growth interruption (GRI) [162, 163] can be also used for localization energy and capture cross section engineering. To investigate the effect of Sb-soaking on electronic properties of the QDs, two similar structures exposed to Sb for different periods of time are grown. However, no significant difference, either in localization energy or in capture cross section, is demonstrated.

Although the obtained storage times in this work are far from the ultimate storage time of 10 years, they are promising and mark an improvement of several orders of magnitude in comparison to the previous reported values. Non-volatility can be achieved with the materials investigated in this work either by increasing localization energy or by decreasing capture cross section. To achieve this aim following recommendations are given:

- Alterations in growth conditions: The electronic properties of QDs can be changed for our gain by altering the growth conditions. An increase in growth

temperature leads to higher surface mobility resulting in larger QDs, thus longer localization energies. Similarly, slow growth rates provide time for QDs to ripen resulting in larger QDs, hence larger localization energies.

- Sb-soaking and growth interruption (GRI) techniques: These methods can be used to change the electronic properties of QDs. Based on some photoluminescence measurements performed on the structures similar to the DLTS structures, it is demonstrated that Sb-flush causes smaller capture cross sections leading to longer storage times due to the formation of smaller QDs. It should be noted that, however, smaller QDs lead to smaller localization energies, hence shorter storage times. The Sb-soaking, thus, has to be performed carefully. GRI method, on the other hand, enables the formation of bigger QDs by providing them time to ripen. Larger localization energies can be obtained via GRI method.

- Additional Barrier: An additional barrier can be implemented to further increase the localization energy of QDs, hence storage time. In this case the total emission barrier for holes confined in QDs becomes the sum of the QD localization energy and the barrier height.

- Capture cross section engineering: The efforts to increase the storage time are devoted to increase the localization energy. However, the localization energies determined in this work are already high. To achieve non-volatility with those localization energies, having a capture cross section of $\sim 10^{15}$ cm^2 would be sufficient. Unfortunately, a capture cross section engineering has not been achieved so far due to our limited knowledge of this complex parameter. However, more study on capture cross section can solve the problem.

The second aim in this work is to decrease the erase time. As explained above, the localization energy is increased to further increase the storage time. However, an increase in barrier height causes an increase in barrier width as well resulting in a slower erase time, as the erase mechanism relies on tunnel emission. To solve this trade-off between storage and erase times, the insertion of superlattice structures into QD-Flash is proposed to create a resonance throughout the device as erase mechanism. At the resonance voltage, where the energy levels of quantum wells and quantum dot are aligned throughout the device, a high transparency is achieved leading to a fast erase time, i.e. erase state, while a low transparency is maintained out of the resonance voltage, i.e. storage state.

To demonstrate that the use of the resonant tunnel structure decreases the erase time for QD-Flash, a transparency calculation method is developed. The electronic properties of the resonance structures such as the conduction and valence bands, the electron and hole concentrations are calculated by one-Dimensional Schrödinger Poisson Solver. Afterwards, the transparency is calculated by Non-Equilibrium Green's Function. It is expected that the transparency increases at the resonance voltage where the energy levels of quantum wells and quantum dot intersect. As the erase time is inversely proportional to transparency, the increase in transparency corresponds to the decrease in erase time.

First, we designed a simple structure consisting of only quantum wells sandwiched between barriers because of its simplicity. Its energy levels as a function of applied voltage is calculated. The intersection points between the energy levels of quantum wells attributed to resonances are observed. Afterwards the current density of the structure is calculated as a function of applied voltage. At the resonance voltages increases in current density are demonstrated. Same structure is grown epitaxially and current-voltage measurement is performed on it. The peaks at the same resonance voltages are shown also experimentally. Moreover, the temperature-dependency of the process is proved. The peaks shrink gradually with increasing temperature and are finally wiped out.

Then, realistic structures are designed with quantum dot layers, quantum wells, and 2DHG layers. Two different material combinations are used in the simulation: the GaAs-based materials due to our vast knowledge on them and the relative simplicity of their epitaxial growth, and the GaP-based materials due to their potential to achieve non-volatility. For both GaAs-based and GaP-based materials the calculations show that the transparency increases by 7 orders of magnitude at the resonance voltage, where the energy levels of the QDs and QWs are aligned. For the two-QW-structures transparency increases 10 orders of magnitude at resonance voltage. The multiple-QW-samples yield higher storage/erase time ratio, however, an alignment throughout the device is not easy due to the high number of energy levels.

In similar structures without quantum wells an erase time of a few ms is achieved [164]. The simulation implies that the implementation of QWs in QD-based memory structures will cause a shortening by 7 orders of magnitude in erase time. In this case, 1 ns of erase time can be achieved, which is faster than a typical DRAM erase time (10 ns). It should be noted, however, that experimental structures and simulated structures might not exactly be the same, e.g. broadening of the QD energy levels due to a Gaussian distribution of the QD sizes has not been taken into consideration. Nevertheless the increase in the transparency and the decrease in erase time should be of the same order of magnitude. Furthermore, GaP-based results confirm that we can implement quantum wells into the structure with which non-volatility can be achieved. Therefore, erase times as short as 1 ns can be realized with 10 years of storage time at the same time. The calculation also suggest that small changes in thickness or composition do not affect the transparency coefficient, instead, result in a small shift of the bias voltage, where the resonance takes place. Therefore, the technology of the resonant tunneling memory is stable.

All the presented results for resonant tunneling structures are based on calculation. Definitively, they have to be confirmed experimentally. Unfortunately, the samples needed to verify these results experimentally could not be successfully grown before the submission of this work. This would constitute the natural next step in the investigation of resonant tunneling structures for memory applications.

Appendix A

Samples

The samples investigated in this work are listed with their real names and some details on the growth and structural characterization.

A.1 GaSb/GaP QDs

Ref	1-G150606-A
QD-1	1-G150607-A
QD-2	1-G150608-A
QD-3	1-G150612-A
QD-4	1-G150613-A

Growth: L. Desplanque, C. Coinon, X. Wallart; Intitute for Electronics, Microelectronics and Nanotechnology, Lille, France.

AFM characterization: L. Desplanque, C. Coinon, X. Wallart; IEMN, CNRS Lille, France.

HAADF-STEM and strain maps: Y. Wang and P. Ruterana; Centre of Research on Ions, Materials and Photonics (CIMAP), CNRS- ENSICAEN- CEA- UCBN Caen, France.

EDX maps: G. Patriarche; Centre for Nanosciences and Nanotechnology (C2N), Marcoussis, France.

A.2 InGaSb/GaP QDs

Only-QD	TU12149
Only-Bar	TU12143
QD-Bar1	TU12144
QD-Bar2	TU12145

Growth: E. Sala; D. Bimberg's group; Institute for Solid State Physics, Technical University of Berlin.

AFM characterization: E. Sala; D. Bimberg's group; Institute for Solid State Physics, Technical University of Berlin.

HAADF-STEM: F. Betram, P. Veit and J. Christen; Otto von Guericke University Magdeburg, Germany.

Appendix B

Setups

In this appendix the experimental setups used for CV measurement and DLTS measurement are explained.

B.1 C-V Measurement

Fig. B.1 shows the schematic setup for CV measurement. Two different cryostat systems are used during the course of the work: A helium flow open-cycle system and a closed-cycle system. An Oxford Instruments ITC503S temperature controller is used for the open-cycle system, while a Lakeshore LS325 temperature controller is used for the close-cycle system. The temperature range between 20 K and 425 K can be swept with both systems. The capacitance measurement is achieved with an HP4248A LCR meters. This LSR meter provides measurement frequencies between 22 Hz and 1 MHz and AC amplitudes between 5 mV to 1V. Moreover, the LSR meter has an internal voltage source. All the device used in the setup are controlled by a computer. The measurements and the analysis of data are achieved with LabView software.

B.2 DLTS setup

Fig. B.2 shows the schematic setup for DLTS measurement. The sample is placed inside cryostat. The open-cycle system with the Oxford Instruments ITC503S temperature controller and the close-cycle system with the Lakeshore LS325 temperature controller are used to regulate the temperature as used in CV measurement. The capacitance is measured with a Boonton 7200 capacitance bridge. It operates at frequency of 1 MHz with an AC amplitude of 100 mV. The capacitance is adjusted by a stepper motor which either adds or subtracts a constant offset (varying from 3 pF to 240 pF) to the diode capacitance, in order to increase the resolution of capacitance bridge. The voltage is generated by an HP 8115A pulse generator. All the devices used in the measurement are controlled by a computer. The measurement and analysis of data are achieved LabView software.

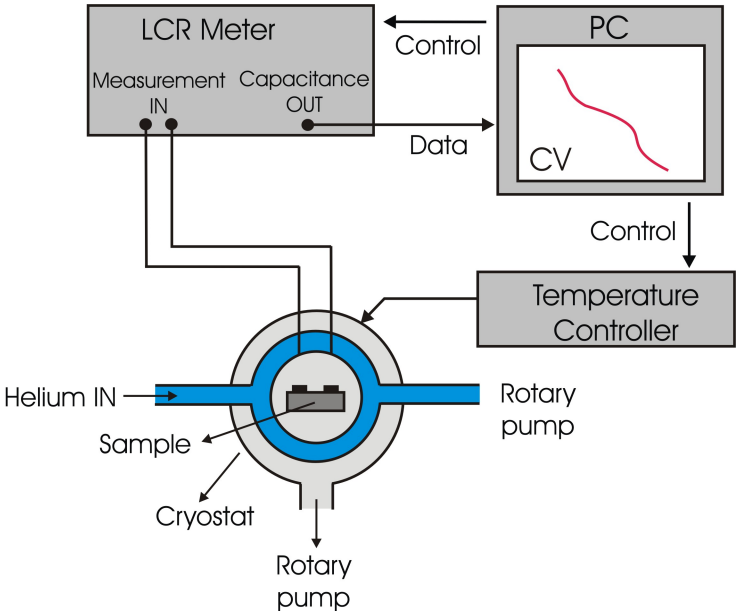

Figure B.1: CV-setup

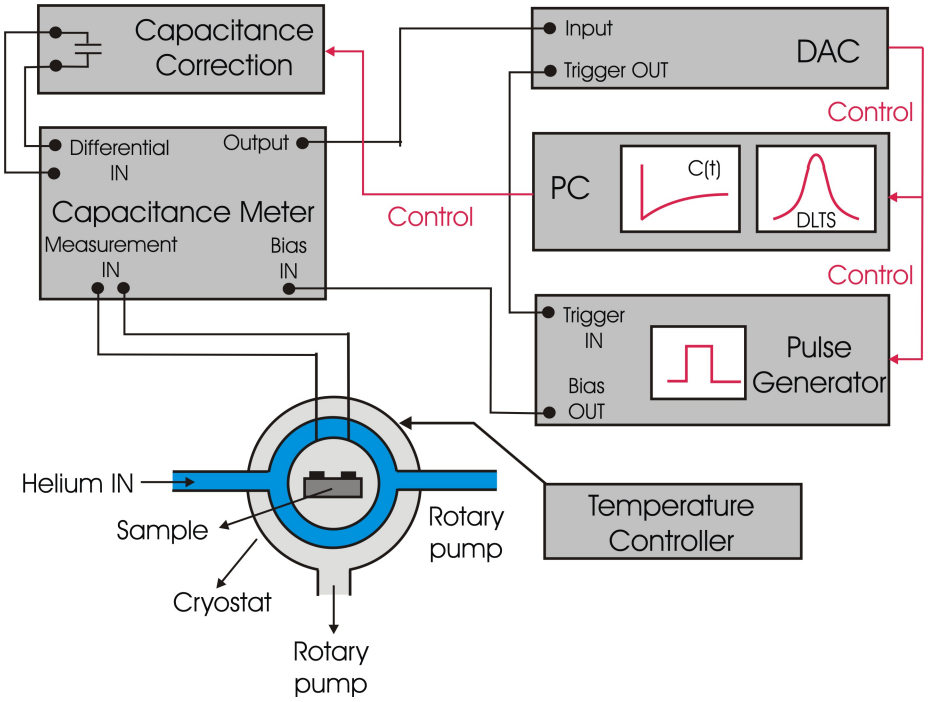

Figure B.2: DLTS-setup

Appendix C

Processing

Standard processing is achieved in three steps: Deposition of the top contact, etching, and deposition of the back contact. In every step, a specific pattern is transferred onto the structure by ultraviolet optical lithography. The top and back contacts are thermally evaporated (step-1 and step-3), where Ni/Au-Ge/Au metals are used for the n-type top contact, while Ni/Zn/Au metals are used for the p-type back contact. The etching is achieved by inductively coupled plasma reactive ion etching (ICP-RIE), creating mesas with 400 μm and 800 μm (step-2). After the deposition of contacts the samples are annealed at 400 °C for 3 minutes in nitrogen atmosphere. Then, the samples are fixed to a ceramic chip carrier using conductive epoxy glue (Epo-Tek H20E). Finally, the devices are contacted using gold wire and a manual ball-wedge bonder (FEK Delvotec 5410).

For some samples p-type semiconductor substrates are used instead of insulator substrates. For these samples, the back contacts are deposited on the back side of the structures. In this case etching and the lithography for the back contact become useless, therefore they are skipped.

Processing in this work were performed at the Center of Nanophotonics (Nanophotonik-Zentrum, NPZ) of the Technische Universität Berlin.

C.1 The top contact

1. Sample cleaning:
    - Rinse in aceton on a hotplate at 76 °C for 2 minutes (2x).
    - Rinse in isopropanol on a hotplate at 76 °C for 2 minutes (2x)
    - Dry with N_2.
2. Lithography of top contact:
    - Heat on hotplate at 120 °C for 5 min.
    - Spin coat with photoresist MaN-440 (at 2500 rpm for 35 s).
    - Pre-bake on hotplate at 90 °C for 5 min.

- Let sample cool down for 5 min.
- Exposure with mask at 6 mW/cm^2 for 10 s.
- Develop with MaD-532-S for ca. 90 s.
- Stop development with H_2O dip (2x) and DI water cascade.
- Remove remaining photoresist with plasma oxidation (150 W, 50 Pa, 76 ml/min O_2, 3 min).

3. Evaporation of top contact:

- Dip in 20:80 HCl/H_2O solution for 30 s.
- Clean in DI water cascade.
- Dry with N_2.
- Evaporate 8 nm Ni, 100 nm Au-Ge (88:12 mixture), 300 nm Au.

4. Lift-off:

- Lift off remaining photoresist with N-Methyl-2-pyrrolidone (NMP) at 76 °C.
- Rinse in aceton on a hotplate at 76 °C for 2 minutes (2x).
- Rinse in isopropanol on a hotplate at 76 °C for 2 minutes (2x).
- Dry with N_2.

C.2 Dry-etching of mesas

1. Sample cleaning:

- Rinse in aceton on a hotplate at 76 °C for 2 minutes (2x).
- Rinse in isopropanol on a hotplate at 76 °C for 2 minutes (2x)
- Dry with N_2.

2. Lithography of mesa structure:

- Heat on hotplate at 120 °C for 5 min.
- Spin coat with photoresist AZ-MIR701 (at 3000 rpm for 35 s).
- Pre-bake on hotplate at 90 °C for 5 min.
- Let sample cool down for 5 min.
- Exposure with mask at 6 mW/^{2} for 35 s.
- Develop with AZ351B for ca. 25 s.
- Stop development with 2x H_2O dip and DI water cascade.
- Remove remaining photoresist with plasma oxidation (150 W, 50 Pa, 76 ml/min O_2, 3 min).

3. Dry-etching of mesas:

   - Use a dummy first to determine the etch rate, and calculate the estimated etching time.
   - Etch sample using the following recipe[1]

1. (Info)
2. Filetype=SENTECH Instruments SI Systems recipe file
3. Anlage=SI 500/D-2M TUB
4. Reaktor=2
5. (Lines)
6. 0=DIM ZEIT AS INTEGER = 300
7. 1=Druck Wafer Rücks.kühlung 1000,0 Pa
8. 2=Wartezeit 10
9. 3=Temperatur Elektrode 20,0 °C
10. 4=Wartezeit 10
11. 5=Gas ein MFC 3 5,0 sccm ' Ar
12. 6=Gas ein MFC 1 20,0 sccm ' BCl3
13. 7=Reaktordruck 1,000 Pa
14. 8=Quellen-Matching manuell 55,9
15. 9=HF-Matching manuell 74,0
16. 10=Wartezeit 10
17. 11=Quelle ein 300,0 W
18. 12=Wartezeit 10
19. 13=HF-Generator ein Bias -200,0 V
20. 14=Quellen-Matching automatisch
21. 15=HF-Matching automatisch
22. 16=Wartezeit ZEIT
23. 17=HF-Generator aus
24. 18=Quelle aus
25. 19=Gas aus MFC 3 ' Ar
26. 20=Gas aus MFC 1 ' BCl3
27. 21=Reaktordruck 0,000 Pa
28. 22=Wartezeit 20
29. 23=Wafer ausschl. ohne Belüftung
30. (Internal Lines)
31. 0=Dim ZEIT As Integer = 300
32. 1=OutAA AA HED 1000.0 Pa
33. 2=WZeit 10

[1]Developed by M. Stubenrauch at TU Berlin

34. 3=OutAA AA TEL 20.0 °C
35. 4=WZeit 10
36. 5=GasEin MFC 3 5.0 sccm ' Ar
37. 6=GasEin MFC 1 20.0 sccm ' BCl3
38. 7=RxDruck 1.000 Pa
39. 8=QMBMan 55.9
40. 9=HFMBMan 74.0
41. 10=WZeit 10
42. 11=Qein 300.0 W
43. 12=WZeit 10
44. 13=HFein Bia -200.0 V
45. 14=QMBAuto
46. 15=HFMBAuto
47. 16=WZeit ZEIT
48. 17=HFaus
49. 18=Qaus
50. 19=GasAus MFC 3 ' Ar
51. 20=GasAus MFC 1 ' BCl3
52. 21=RxDruck 0.000 Pa
53. 22=WZeit 20
54. 23=WafAusOB

1. Lift-off:
    - Lift off remaining photoresist with N-Methyl-2-pyrrolidone (NMP) at 76 °C.
    - Rinse in aceton on a hotplate at 76 °C for 2 minutes (2x).
    - Rinse in isopropanol on a hotplate at 76 °C for 2 minutes (2x).
    - Dry with N_2.

C.3 The back contact

1. Sample cleaning:
    - Rinse in aceton on a hotplate at 76 °C for 2 minutes (2x).
    - Rinse in isopropanol on a hotplate at 76 °C for 2 minutes (2x)
    - Dry with N_2.
2. Lithography of back contact:
    - Heat on hotplate at 120 °C for 5 min.

- Spin coat with photoresist MaN-440 (at 2500 rpm for 35 s).
- Pre-bake on hotplate at 90 °C for 5 min.
- Let sample cool down for 5 min.
- Exposure with mask at 6 $mW/^2$ for 10 s.
- Develop with MaD-532-S for ca. 90 s.
- Stop development with H_2O dip (2x) and DI water cascade.
- Remove remaining photoresist with plasma oxidation (150 W, 50 Pa, 76 ml/min O_2, 3 min)..

3. Evaporation of top contact:

   - Dip in 20:80 HCl/H_2O solution for 30 s.
   - Clean in DI water cascade.
   - Dry with N_2.
   - Evaporate 8 nm Ni, 100 nm Zn, 300 nm Au.

4. Lift-off:

   - Lift off remaining photoresist with N-Methyl-2-pyrrolidone (NMP) at 76 °C.
   - Rinse in aceton on a hotplate at 76 °C for 2 minutes (2x).
   - Rinse in isopropanol on a hotplate at 76 °C for 2 minutes (2x).
   - Dry with N_2.

Appendix D

Error Analysis

In this section the calculation of the error margins on localization energy, capture cross section, and storage time are explained.

DLTS spectra obtained by DLTS measurement are analyzed manually by specifying the temperature at which the maximum of peak occurs for different reference time constants τ_{ref}. The time constant is equal to the reciprocal of the thermal emission. Activation energy and capture cross section are derived from the Arrhenius plot which is plotted by reference time constant. Finally, storage time is calculated using localization energy and capture cross section. The errors on mentioned parameters originate from the determination of the maximum point of DLTS spectra. Therefore, DLTS spectra are analyzed several times and a set of data is collected. Accordingly, a mean, a minimum and a maximum values and a standard deviation are determined. The main idea was taken from [165].

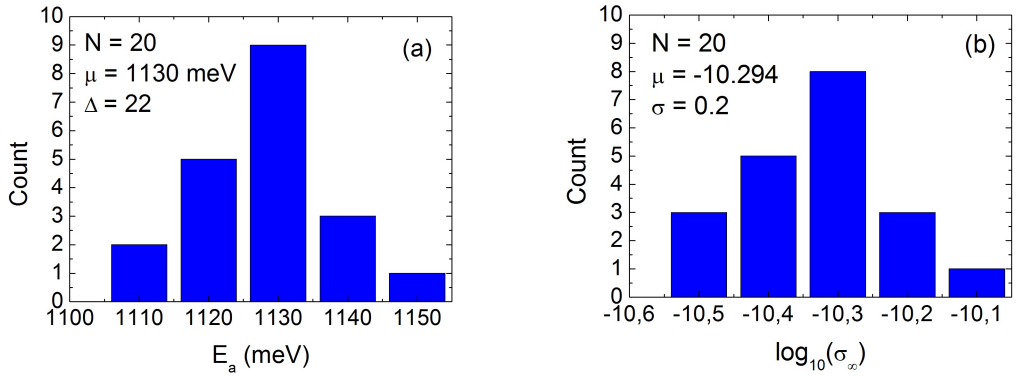

Figure D.1: (a) Histogram of activation energies E_a and (b) histogram of capture cross sections σ_∞ obtained from graphical analysis of DLTS graph for $V_m = 5$ V.

The error analysis is presented for the charge selective DLTS data for QD-Bar2. The measurement voltage was $V_m = 6$ V (reverse direction) and the pulse voltage was $V_m = 5$ V (reverse direction). The graphical analysis is repeated 20 times. Fig. D.1a shows the histogram of the activation energy. The values determined for activation energy follow a Gaussian distribution justifying the use of statistical approach

Parameter	Mean value μ	St. dev. σ	Min. value	Max. value
E_a	1130 meV	22 meV	1109 meV	1152 meV
$\log_{10}\sigma_\infty$	-10.294	0.2	-10.551	-10.052
$\log_{10}\tau$	3.417	0.12	3.333	3.535

Table D.1: Mean value, minimum value, maximum value, and standard deviation from $N = 20$ evaluation for a measurement voltage $V_m = 5$ V for sample QD-Bar2.

to determine the error margins. The mean value for activation energy is determined to be 1130 meV with a standard deviation of 22 meV. The minimum value obtained for activation energy is 1131 meV, while the maximum activation energy is 1154 meV (see Tab. D.1). Fig. D.1b shows the histogram of capture cross section. Similarly, the values obtained for capture cross section follow a Gaussian distribution. The mean value is determined to be -10.294 (5.08×10^{-11}) with a standard deviation of 0.2 in logarithmic scale. The minimum and maximum values of capture cross section are obtained to be -10.551 and -10.052 in logarithmic scale, respectively (see Tab. D.1).

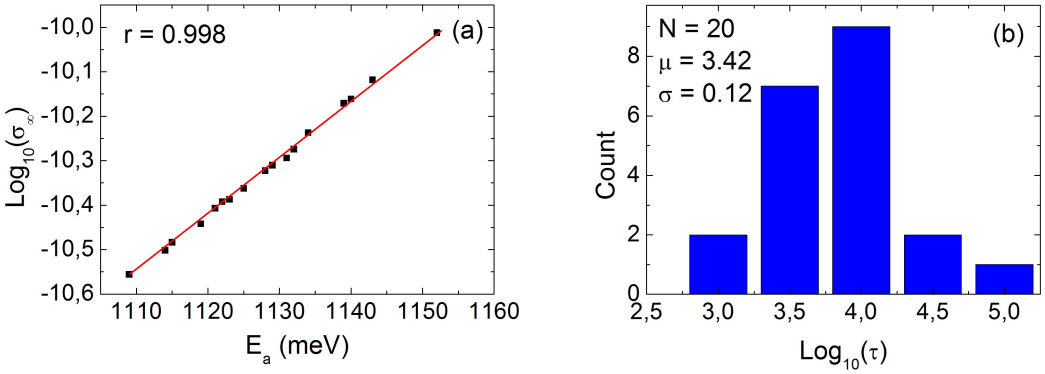

Figure D.2: (a) Logarithmized capture cross section σ_∞ versus corresponding activation energy E_a from $N = 20$ evaluation for $V_m = 5$ V (b) histogram of storage time τ obtained from graphical analysis of DLTS graph from $N = 20$ evaluation for $V_m = 5$ V.

Storage time is calculated using localization energy and capture cross section. Therefore, the uncertainties on storage time originates the uncertainty on localization energy and capture cross section. It should be, however, noted that when the localization energy increases, the capture cross section increases too, as they are derived from the slope of Arrhenius plot. Therefore, they partially cancel each other out and the uncertainty on storage time is hence smaller than the uncertainties on localization energy and on capture cross section. Fig. D.2a shows the correlation between localization energy and capture cross section. The rate is 0.997 and 1 represents the complete correlation. Fig. D.2b shows the hologram of storage time. The mean value is determined to be 3.417 and the standard deviation is determined to be 0.12.

The final values and uncertainties are rounded to significant digits, and listed as follow:

- Localization energy $E_{loc} = 1130(\pm 20)$ meV
- Capture cross section $\sigma_\infty = 5. \times 10^{-11}$ cm^2 (with an uncertainty of 0.2 orders of magnitude)
- Storage time $\tau = 2500$ s (with an uncertainty of 0.1 orders of magnitude)

For the rest of the samples, only uncertainties will be given without any detail.

D.1 GaSb/GaP QDs

D.1.1 QD-1

The results of the error analysis of the charge selective DLTS measurement of QD-1 at a measurement voltage $V_m = 10$ V are listed Tab. D.2.

Parameter	Mean value μ	St. dev. σ	Min. value	Max. value
E_a	900.3 meV	22 meV	875.5 meV	922.8 meV
$\log_{10}\sigma_\infty$	-12.405	0.317	-12.804	-11.772
$\log_{10}\tau$	1.404	0.120	1.213	1.509

Table D.2: Mean value, minimum value, maximum value, and standard deviation from $N = 10$ evaluation for a measurement voltage $V_m = 10$ V for sample QD-1.

- Localization energy $E_{loc} = 900(\pm 20)$ meV
- Capture cross section $\sigma_\infty = 4 \times 10^{-13}$ cm^2 (with an uncertainty of 0.3 orders of magnitude)
- Storage time $\tau = 30$ s (with an uncertainty of 0.1 orders of magnitude)

D.1.2 QD-2

The results of the error analysis of the charge selective DLTS measurement of QD-2 at a measurement voltage $V_m = 10$ V are listed Tab. D.3.

Parameter	Mean value μ	St. dev. σ	Min. value	Max. value
E_a	1107.0 meV	34.7 meV	1025.8 meV	1152.8 meV
$\log_{10}\sigma_\infty$	-10.395	0.518	-11.623	-9.721
$\log_{10}\tau$	2.930	0.066	2.793	3.025

Table D.3: Mean value, minimum value, maximum value, and standard deviation from $N = 10$ evaluation for a measurement voltage $V_m = 10$ V for sample QD-2.

- Localization energy $E_{loc} = 1.11(\pm 0.03)$ eV
- Capture cross section $\sigma_\infty = 4 \times 10^{-11}$ cm^2 (with an uncertainty of 0.5 orders of magnitude)
- Storage time $\tau = 850$ s (with an uncertainty of 0.07 orders of magnitude)

D.1.3 QD-3

The results of the error analysis of the charge selective DLTS measurement of QD-3 at a measurement voltage $V_m = 6$ V are listed Tab. D.4.

Parameter	Mean value μ	St. dev. σ	Min. value	Max. value
E_a	1070.2 meV	14.3 meV	1037.1 meV	1085.1 meV
$\log_{10}\sigma_\infty$	-11.142	0.202	-11.599	-10.928
$\log_{10}\tau$	3.058	0.039	2.959	3.095

Table D.4: Mean value, minimum value, maximum value, and standard deviation from $N = 10$ evaluation for a measurement voltage $V_m = 6$ V for sample QD-3.

- Localization energy $E_{loc} = 1.07(\pm 0.01)$ eV
- Capture cross section $\sigma_\infty = 7 \times 10^{-12}$ cm^2 (with an uncertainty of 0.2 orders of magnitude)
- Storage time $\tau = 1100$ s (with an uncertainty of 0.04 orders of magnitude)

D.1.4 QD-4

The results of the error analysis of the charge selective DLTS measurement of QD-4 at a measurement voltage $V_m = 4$ V are listed Tab. D.5.

Parameter	Mean value μ	St. dev. σ	Min. value	Max. value
E_a	1175.1 meV	9.4 meV	1166.9 meV	1193.9 meV
$\log_{10}\sigma_\infty$	-11.849	0.122	-11.959	-11.604
$\log_{10}\tau$	5.528	0.036	5.497	5.598

Table D.5: Mean value, minimum value, maximum value, and standard deviation from $N = 10$ evaluation for a measurement voltage $V_m = 4$ V for sample QD-4.

- Localization energy $E_{loc} = 1.18(\pm 0.01)$ meV
- Capture cross section $\sigma_\infty = 1 \times 10^{-12}$ cm^2 (with an uncertainty of 0.1 orders of magnitude)
- Storage time $\tau = 337000$ s, 3.9 days, (with an uncertainty of 0.04 orders of magnitude)

D.1.5 Ref

The results of the error analysis of the charge selective DLTS measurement of Ref at a measurement voltage $V_m = 4$ V are listed Tab. D.6.

Parameter	Mean value μ	St. dev. σ	Min. value	Max. value
E_a	852.3 meV	9.3 meV	838.6 meV	864.0 meV
$\log_{10}\sigma_\infty$	-14.042	0.129	-14.223	-13.879
$\log_{10}\tau$	2.297	0.031	2.242	2.338

Table D.6: Mean value, minimum value, maximum value, and standard deviation from $N = 10$ evaluation for a measurement voltage $V_m = 4$ V for sample Ref.

- Localization energy $E_{loc} = 852(\pm 9)$ meV
- Capture cross section $\sigma_\infty = 9 \times 10^{-15}$ cm^2 (with an uncertainty of 0.1 orders of magnitude)

D.2 InGaSb/GaP QDs

D.2.1 Only-QD

The results of the error analysis of the charge selective DLTS measurement of Only-QD at a measurement voltage $V_m = 5$ V are listed Tab. D.7.

Parameter	Mean value μ	St. dev. σ	Min. value	Max. value
E_a	372 meV	8 meV	363 meV	380 meV
$\log_{10}\sigma_\infty$	-11.673	0.09	-11.763	-11.583
$\log_{10}\tau$	-8.115	0.07	-8.180	-8.122

Table D.7: Mean, minimum and maximum values, and standard deviation from $N = 20$ evaluation for a measurement voltage $V_m = 5$ V for sample Only-QD.

- Localization energy $E_{loc} = 370(\pm 8)$ meV
- Capture cross section $\sigma_\infty = 2 \times 10^{-12}$ cm^2 (with an uncertainty of 0.1 orders of magnitude)
- Storage time $\tau = 8 \times 10^{-9}$ s (with an uncertainty of 0.1 orders of magnitude)

Parameter	Mean value μ	St. dev. σ	Min. value	Max. value
E_a	630.4 meV	10.2 meV	621.1 meV	639.8 meV
$\log_{10}\sigma_\infty$	-11.182	0.12	-11.084	-11.292

Table D.8: Mean value, minimum value, maximum value, and standard deviation from $N = 20$ evaluation for a measurement voltage $V_m = 5$ V for sample QD-Bar.

D.2.2 Only-Bar

The results of the error analysis of the charge selective DLTS measurement of Only-Bar at a measurement voltage $V_m = 9$ V are listed Tab. D.8.

- Localization energy $E_{loc} = 630(\pm 10)$ meV
- Capture cross section $\sigma_\infty = 7 \times 10^{-12}$ cm^2 (with an uncertainty of 0.1 orders of magnitude)

D.2.3 QD-Bar1

The results of the error analysis of the charge selective DLTS measurement of QD-Bar1 at a measurement voltage $V_m = 9$ V are listed Tab. D.9.

Parameter	Mean value μ	St. dev. σ	Min. value	Max. value
E_a	1151 meV	22 meV	1134 meV	1173 meV
$\log_{10}\sigma_\infty$	-10.054	0.18	-9.857	-10.253
$\log_{10}\tau$	3.500	0.120	3390	3620

Table D.9: Mean value, minimum value, maximum value, and standard deviation from $N = 20$ evaluation for a measurement voltage $V_m = 5$ V for sample QD-Bar1.

- Localization energy $E_{loc} = 1.15(\pm 0.02)$ meV
- Capture cross section $\sigma_\infty = 9 \times 10^{-11}$ cm^2 (with an uncertainty of 0.2 orders of magnitude)
- Storage time $\tau = 3200$ s (with an uncertainty of 0.1 orders of magnitude)

D.2.4 QD-Bar2

See the first section.

Bibliography

[1] D. Reinsel, J. Gantz, and J. Rydning, "Data Age 2025: The Evolution of Data to Life-Critical, *IDC iView*, (2017).

[2] J. Gantz and D. Reinsel, *The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East*, *IDC iView*, (2012).

[3] G. Moore, "Cramming more components onto integrated circuits", *Electronics* 8, 114 (1965).

[4] *International Technology Roadmap for Semiconductors* 2.0. Tech. rep. 2015.

[5] *International Technology Roadmap for Semiconductors* (ITRS) - 2009 Edition, 2009.

[6] J. T. Clemens, J. D. Cuthbert, F. J. Procyk, and G. M. Trout, "Dynamic random access memory", 4240195 (1980).

[7] R. Waser, *Nanoelectronics and Information Technology*, Wiley-VCH, Berlin, 2003.

[8] R. Waser, *Nanotechnology Volume 3: Information Technology I*, Nanotechnology, Wiley-VCH, Weinheim, 2008.

[9] R. Bez, E. Camerlenghi, A. Modeli, and A. Visconti, "Introduction to Flash Memory", *Proceedings of the IEEE* 91 (4), (2003).

[10] J. E. Brewer and M. Gill, *Non-volatile Memory Technologies with Emphasis on Flash*, John Wiley Sons, New Jersey, 2008.

[11] F. Masuoka, M. Asano, H. Iwahashi, T. Komuro, and S. Tanaka, "A new flash EEPROM cell using triple polysilicon technology", *1984 International Electron Devices Meeting*, Institute of Electrical & Electronics Engineers (IEEE), (1984).

[12] L. Geppert, "The new indelible memories - Its a three-way race in the multibillion dollar memory sweepstakes", *IEEE Spectrum* 40(3), 4854 (2003).

[13] M. Geller, A. Marent, and D. Bimberg, Speicherzelle und Verfahren zum Speichern von Daten (Memory cell and method for storing data), International patent EP/2097904, 2006.

[14] A. Marent, M. Geller, T. Nowozin, and D. Bimberg, Speicherzelle auf Basis von Nanostrukturen aus Verbindungshalbleitern, International patent application PCT/12/970,744, 2010.

[15] A. Marent and D. Bimberg. Speicherzelle. German Patent N. 10 2011 006 782.5. Oct. 2012.

[16] D. Bimberg, M. Geller, A. Marent, and T. Nowozin. Memory.U.S. PatentN. 8331142 B2. Dec. 2012.

[17] D. Bimberg, M. Grundmann, and N. N. Ledentsov, *Quantum Dot Heterostructures*, John Wiley and Sons, Chichester, 1998.

[18] D. Bimberg, editor, *Semiconductor Nanostructures*, Springer, Berlin Heidelberg, 2008.

[19] L. Bonato, E. M. Sala, G. Stracke, T. Nowozin, A. Strittmatter, M. N. Ajour, K. Daqrouq, and D. Bimberg, "230 s room-temperature storage time and 1.14eV hole localization energy in $In_0.5Ga_0.5As$ quantum dots on a GaAs interlayer in GaP with an AlP barrier", *Appl. Phys. Lett.* 106, 042102 (2015).

[20] M. Geller, A. Marent, T. Nowozin, D. Bimberg, N. Akcay, and N. Öncan. "A write time of 6 ns for quantum-dot-based memory structures.", *Appl. Phys. Lett.* 92, p. 092108, (2008).

[21] T. Nowozin, A. Beckel, D. Bimberg, A. Lorke, and M. Geller. "3 ns single-shot read-out in quantum dot-based memory structure.", *Appl. Phys. Lett.* 104, p. 053111, (2014).

[22] A. Marent, "Entwicklung einer neuartigen Quantenpunkt-Speicherzelle", Dissertation, Technische Universität Berlin, 2010.

[23] A. Marent, T. Nowozin, J. Gelze, F. Luckert, and D. Bimberg, "Hole-based memory operation in an InAs/GaAs quantum dot heterostructure", *Appl. Phys. Lett.* 95, 242114 (2009).

[24] A. Marent, T. Nowozin, M. Geller, and D. Bimberg, "The QD-Flash: a quantum dot-based memory device", *Semicond. Sci. Technol.* 26, 014026 (2011).

[25] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for IIIV compound semiconductors and their alloys", *J. Appl. Phys.* 89(11), 58155875, (2001).

[26] J. R. Arthur, "Molecular Beam Epitaxy", *Surface Science*, 500, 189217 (2002).

[27] G. B. Stringfellow, *Organometallic Vapor-Phase Epitaxy: Theory and Practice*, Academic Press, San Diego, 2nd edition, 1999.

[28] T. Mimura, S. Hiyamizu, T. Fujii, and K. Nanbu, "A New Field-Effect Transistor with Selectively Doped GaAs/n-$Al_xGa_{1x}As$ Heterojunctions", *Jpn. J. Appl. Phys.* 19, L225L227 (1980).

[29] D. Delagebeaudeuf, P. Delescluse, P. Etienne, M. Laviron, J. Chaplart, and N. T. Linh, "Two-dimensional electron gas MESFET structure", *Electron. Lett.* 16(17), 667668 (1980).

[30] I. Hayashi, M. B. Panish, P. W. Foy, and S. Sumski, "Junction lasers which operate continuously at room temperature", *Appl. Phys. Lett.* 17, 109 (1970).

[31] F. K. Reinhart, "Mode Reflectivity and Waveguide Properties of Double-Heterostructure Injection Lasers", *J. Appl. Phys.* 42, 1659796 (1971).

[32] D. Schicketanz and G. Zeidler, "GaAs-Double-Heterostructure Lasers as Optical Amplifiers", *IEEE Journal of Quantum Electronics* 11(2) (1975).

[33] C. Y. Chen, "Ultrahigh speed modulation-doped heterostructure field-effect photodetectors", *Appl. Phys. Lett.* 42,1040 (1983).

[34] N. Holonyak and S. F. Bevacqua, "Coherent (visible) light emission from $Ga(As_{1-x}P_x)$ junctions. *Appl. Phys. Lett.* 1.4 , 82, (1962).

[35] Z. I. Alferov and R. F. Kazarinov, "Semiconductor laser with electric pumping", Inventors certificate No. 181737 (in Russian), 1963.

[36] H. Kroemer, "A proposed class of hetero-junction injection lasers", *Proceedings of the IEEE* 51(12), 1782 (1963).

[37] H. Kroemer, "Barrier control and measurements: Abrupt semiconductor heterojunctions", *J. Vac. Sci. Technol. B* 2(3), 433439 (1984).

[38] E. T. Yu, J. O. McCaldin, and T. C. McGill, "Band Offsets in Semiconductor Heterojunctions", *Solid State Physics* 46, 1146 (1992).

[39] J. H. Davies, *The physics of low-dimensional semiconductors*, Cambridge University Press, 1988.

[40] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, John Wiley & Sons, 3rd edition, 2006.

[41] E. Conwell and V. F. Weisskopf, "Theory of Impurity Scattering in Semiconductors", *Phys. Rev.* 77(3), 388390 (1950).

[42] R. Dingle, H. L. Störmer, A. C. Gossard, and W. Wiegmann, "Electron mobilities in modulation-doped semiconductor heterojunction superlattices", *Appl. Phys. Lett.* 33(7), 665 (1978).

[43] C. T. Foxon, "Three decades of molecular beam epitaxy", *Journal of Crystal Growth* 251, 18 (2003).

[44] K. v. Klitzing, G. Dorda, and M. Pepper, "New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance", *Phys. Rev. Lett.* 45(6), 494497 (1980).

[45] K. v. Klitzing, "The quantized Hall effect", *Rev. Modern Physics* 58(3), 519531 (1986).

[46] M. J. Manfra, L. N. Pfeiffer, K. W. West, R. d. Picciotto, and K. W. Baldwin, "High mobility two-dimensional hole system in GaAs/AlGaAs quantum wells grown on (100) GaAs substrates", *Appl. Phys. Lett.* 86, 162106 (2005).

[47] M. A. Kastner, "Artificial Atoms", *Physics Today* 46, 24 (1993).

[48] K. Koike, K. Saitoh, S. Li, S. Sasa, M. Inoue, and M. Yano, "Room-temperature operation of a memory-effect AlGaAs/GaAs heterojunction field-effect transistor with self-assembled InAs nanodots", *Appl. Phys. Lett.* 76(11), 1464 (2000).

[49] H. Kim, T. Noda, T. Kawazu, and H. Sakaki, "Control of Current Hysteresis Effects in a GaAs/n-AlGaAs Quantum Trap Field Effect Transistor with Embedded InAs Quantum Dots", *Jpn. J. Appl. Phys.*, 39, 7100, (2000).

[50] N. Kirstaedter, N. N. Ledentsov, M. Grundmann, D. Bimberg, V. M. Ustinov, S. S. Ruvimov, M. V. Maximov, P. S. Kopev, Z. I. Alferov, U. Richter, P. Werner, U. Gosele, and J. Heydenreich, "Low threshold, large T_0 injection laser emission from (InGa)As quantum dots", *Electron. Lett.*, 30(17), 1416, (1994).

[51] F. Heinrichsdorff, C. Ribbat, M. Grundmann, and D. Bimberg, "High-power quantum-dot lasers at 1100 nm", *Appl. Phys. Lett.*, 76(5), 556-558, (2000).

[52] M. Kuntz, G. Fiol, M. Lämmlin, D. Bimberg, M. G. Thompson, K. T. Tan, C. Marinelli, R. V. Penty, I. H. White, V. M. Ustinov, A. E. Zhukov, Y. M. Shernyakov, and A. R. Kovsh, "35 GHz mode-locking of 1.3 μm quantum dot lasers", *Appl. Phys. Lett.*, 85, 843, (2004).

[53] N. Ledentsov, M. Grundmann, F. Heinrichsdorff, D. Bimberg, V. Ustinov, A. Zhukov, M. Maximov, Z. Alferov, and J. Lott, "Quantum-dot heterostructure lasers", *IEEE J. Select. Topics Quantum Electron.*, 6.3, 439-451, (2000).

[54] M. Lämmlin, G. Fiol, C. Meuer, M. Kuntz, F. Hopfer, A. R. Kovsh, N. N. Ledentsov, and D. Bimberg, "Distortion-free optical amplification of 20-80 GHz modelocked laser pulses at 1.3 μm using quantum dots", *Electronics Lett.*, 42, 697, (2006).

[55] E. Stock, T. Warming, I. Ostapenko, S. Rodt, A. Schliwa, J. A. Töfflinger, A. Lochmann, A. I. Toropov, S. A. Moshchenko, D. V. Dmitriev, V. A. Haisler, and D. Bimberg, "Single-photon emission from InGaAs quantum dots grown on (111) GaAs", *Appl. Phys. Lett.*, 96, 093112, (2010).

[56] W. Unrau, D. Quandt, J. H. Schulze, T. Heindel, T. D. Germann, O. Hitzemann, A. Strittmatter, S. Reitzenstein, U. W. Pohl, and D. Bimberg, "Electrically driven single photon source based on a site-controlled quantum dot with self-aligned current injection", *Appl. Phys. Lett.*, 101(21), 211119, (2012).

[57] F. Hopfer, A. Mutig, M. Kuntz, G. Fiol, D. Bimberg, N. N. Ledentsov, V. A. Shchukin, S. S. Mikhrin, D. L. Livshits, I. L. Krestnikov, A. R. Kovsh, N. D. Zakharov, and P. Werner, "Single-mode submonolayer quantum-dot vertical-cavity surface-emitting lasers with high modulation bandwidth", *Appl. Phys. Lett.* 89, 141106 (2006).

[58] X. Huang, A. Stintz, H. Li, L. F. Lester, J. Cheng, and K. J. Malloy, "Passive modelocking in 1.3 μm two-section InAs quantum dot lasers", *Appl. Phys. Lett.* 78(19), 2825 (2001).

[59] J. Campbell, D. Huffaker, H. Deng, and D. Deppe, "Quantum dot resonant cavity photodiode with operation near 1.3 μm wavelength", *Electron. Lett.* 33 (15), 1337 (1997).

[60] L. Chu, A. Zrenner, M. Bichler, and G. Abstreiter, "Quantum-dot infrared photodetector with lateral carrier transport", *Appl. Phys. Lett.* 79(14,) 2249 (2001).

[61] G. Yusa and H. Sakaki, "Trapping of photogenerated carriers by InAs quantum dots and persistent photoconductivity in novel GaAs/n-AlAs field-effect transistor structures", *Appl. Phys. Lett.* 70(3), 345 (1997).

[62] C. Balocco, A. M. Song, and M. Missous, "Room-temperature operations of memory devices based on self-assembled InAs quantum dot structures", *Appl. Phys. Lett.* 85(24), 59115913 (2004).

[63] D. Nataraj, N. Ooike, J. Motohisa, and T. Fukui, "Fabrication of one-dimensional GaAs channel-coupled InAs quantum dot memory device by selective-area metalorganic vapor phase epitaxy", *Appl. Phys. Lett.* 87, 193103 (2005).

[64] M. Geller, C. Kapteyn, L. Müller-Kirsch, R. Heitz, and D. Bimberg, "Hole storage in GaSb/GaAs quantum dots for memory devices", *Phys. Stat. sol. b* 238 (2003).

[65] F. C. Frank and J. H. v. d. Merwe, "One-Dimensional dislocations. I. Static Theory", *Proc. Roy. Soc. Lond.*, A 198, 205 (1949).

[66] M. Volmer and A. Weber, "Keimbildung in übersättigten Gebilden", *Zeitschr. f. phys. Chem.*, 119, 277 (1926).

[67] V. A. Shchukin and D. Bimberg, "Spontaneous ordering of nanostructures on crystal surfaces", *Rev. Modern Physics* 71, 1125 (1999).

[68] I. N. Stranski and L. Krastanow, "Zur Theorie der orientierten Ausscheidung von Ionenkristallen aufeinander", *Sitzungsber. Akad. Wiss. Wien, Math.-Naturwiss.* Kl., Abt. 2B 146, 797 (1938).

[69] D. J. Mowbray and M. S. Skolnick, "New physics and devices based on self-assembled semiconductor quantum dots", *J. Phys. D: Appl. Phys.* 38 (13),20592076 (2005).

[70] L. Müller-Kirsch, R. Heitz, U. W. Pohl, D. Bimberg, I. Häusler, H. Kirmse, and W. Neumann,"Temporal evolution of GaSb/GaAs quantum dot formation", *Appl. Phys. Lett.* 79(7), 1027 (2001).

[71] K. Jacobi, "Atomic structure of InAs quantum dots on GaAs", *Progress in Surface Science* 71, 185-215 (2013).

[72] C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Quantenmechanik - Band I und II*, Gruyter, 2nd edition, 1999.

[73] M. Grundmann, O. Stier, and D. Bimberg, "InAs/GaAs pyramidal quantum dots: Strain distribution, optical phonons, and electronic structure", *Phys. Rev. B*, 52, 1196911981 (1995).

[74] O. Stier, M. Grundmann, and D. Bimberg, "Electronic and optical properties of strained quantum dots modeled by 8-band-k.p theory", *Phys. Rev. B*, 59, 5688 (1999).

[75] A. Schliwa, M. Winkelnkemper, and D. Bimberg, "Impact of size, shape and composition on piezoelectric effects and the electronic properties of InGaAs/GaAs quantum dots", *Phys. Rev. B*, 76, 205324 (2007).

[76] L. Bonato, "Towards non-volatility in quantum-dot-based memories", Dissertation, Technische Universit¨t Berlin, 2017.

[77] T. Nowozin, "Electronic properties of and carrier dynamics in self-organized quantum dots for memories", Dissertation, Technische Universit¨t Berlin, 2013.

[78] P. Pavan, R. Bez, P. Olivo, and E. Zanoni, "Flash memory cells - An overview", *Proceedings of the IEEE* 85(8), 12481271 (1997).

[79] P. Cottrell, R. Troutman, and T. Ning, "Hot-electron emission in n-channel IGFETs", *IEEE J. Solid-State Circuits* 14(2), 442455 (1979).

[80] M. Lenzlinger, "Fowler-Nordheim Tunneling into Thermally Grown SiO_2". *J. Appl. Phys.* 40, 278 (1969).

[81] T. Müller, F. F. Schrey, G. Strasser, and K. Unterrainer, "Ultrafast intraband spectroscopyof electron capture and relaxation in InAs/GaAs quantum dots", *Appl. Phys. Lett.* 83, p. 3572, (2003).

[82] M. Geller, A. Marent, E. Stock, D. Bimberg, V. I. Zubkov, I. S. Shulgunova, and A. V. Solomonov. "Hole capture into self-organized InGaAs quantum dots", *Appl. Phys. Lett.* 89, p. 232105, (2006).

[83] M. Geller, C. Kapteyn, L. Müller-Kirsch, R. Heitz, and D. Bimberg, "450 meV hole localization energy in GaSb/GaAs quantum dots", *Appl. Phys. Lett.* 82(16), 27062708 (2003).

[84] A. Marent, M. Geller, D. Bimberg, A. P. Vasiev, E. S. Semenova, A. E. Zhukov, and V. M. Ustinov, "Carrier storage time of milliseconds at room temperature in self-organized quantum dots", *Appl. Phys. Lett.* 89(7), 072103 (2006).

[85] A. Marent, M. Geller, A. Schliwa, D. Feise, K. Pötschke, D. Bimberg, N. Akcay,and N. Öncan. "10^6 years extrapolated hole storage time in GaSb/AlAs quantum dots.", *Appl. Phys. Lett.* 91, 242109, (2007).

[86] T. Nowozin, A. Marent, L. Bonato, A. Schliwa, D. Bimberg, E. P. Smakman, J. K. Garleff, P. M. Koenraad, R. J. Young, and M. Hayne, "Linking structural and electronic properties of high-purity self-assembled GaSb/GaAs quantum dots", *Phys. Rev. B* 86, 035305 (2012).

[87] G. Stracke, A. Glacki, T. Nowozin, L. Bonato, S. Rodt, C. Prohl, A. Lenz, H. Eisele, A. Schliwa, A. Strittmatter, U. W. Pohl, and D. Bimberg, "Growth of $In_{0.25}Ga_{0.75}As$ quantum dots on GaP utilizing a GaAs interlayer", *Appl. Phys. Lett.* 101, 223110 (2012).

[88] T. Nowozin, L. Bonato, A. Högner, A. Wiengarten, D. Bimberg, W. -H. Lin, S. -Y. Lin, C. J. Reyner, B. L. Liang, and D. L. Huffaker, "800 meV localization energy InGaSb/GaAs/$Al_{0.3}Ga_{0.7}As$ quantum dots, *Appl. Phys. Lett.* 102, 052115 (2013).

[89] B. Marquardt, M. Geller, A. Lorke, D. Reuter, and A. D. Wieck, "Using a twodimensional electron gas to study non-equilibrium tunneling dynamics and charge storage in self-assembled quantum dots", *Appl. Phys. Lett.* 95, 022113 (2009).

[90] M. Russ, C. Meier, A. Lorke, D. Reuter, and A. D. Wieck, "Role of quantum capacitance in coupled low-dimensional electron systems", *Phys. Rev. B* 73, 115334 (2006).

[91] M. Russ, C. Meier, B. Marquardt, A. Lorke, D. Reuter, and A. D. Wieck, "Quantum dot electrons as controllable scattering centers in the vicinity of a two-dimensional electron gas", *Phase Transitions* 79(9-10), 765770 (2006).

[92] M. Geller, "Investigation of Carrier Dynamics in Self-Organized Quantum Dots for Memory Devices", Dissertation, Technische Universität Berlin, 2007.

[93] D. V. Lang, "Deep-level transient spectroscopy: A new method to characterize traps in semiconductors, *J. Appl. Phys.* 45(7), 3023 (1974).

[94] P. Blood and J. W. Orton, *The Electrical Characterization of Semiconductors: Majority Carriers and Electron States*, Academic Press, London, 1992.

[95] C. Kapteyn, *Carrier Emission and Electronic Properties of Self-Organized Semiconductor Quantum Dots*, Mensch & Buch Verlag, Berlin, 2001, Dissertation, Technische Universität Berlin.

[96] C. Kapteyn, M. Lion, R. Heitz, D. Bimberg, P. Brunkov, B. Volovik, S. Konnikov, A. Kovsh, and V. Ustinov, "Time-Resolved Capacitance Spectroscopy of Hole and Electron Levels in InAs/GaAs Quantum Dots", *Phys. Stat. Sol. b*, 224 (2001).

[97] R. Ferreira and G. Bastard, "Phonon-assisted capture and intradot Auger relaxation in quantum dots, *Appl. Phys. Lett.* 74(19), 2818 (1999).

[98] J. Bourgoin and M. Lannoo, *Point Defects in Semiconductors II - Experimental Aspects*, volume 35 of Springer Series in Solid-State Sciences, Springer, Berlin, 1983.

[99] D. V. Lang and C. H. Henry, "Nonradiative recombination at deep levels in GaAs and GaP by lattice-relaxation multiphonon emission", *Phys. Rev. B* 35(22), 1525 1528 (1975).

[100] G. Vincent, A. Chantre, and D. Bois, "Electric field effect on the thermal emission of traps in semiconductor junctions", *J. Appl. Phys.* 50(8), 5484 (1979).

[101] W. Nolting, *Grundkurs theoretische Physik - Band 5/2: Quantenmechanik*, Springer, Berlin Heidelberg, 2004.

[102] J. Gelze, *Ladungsträgerdynamik in Quantenpunkt-basierten Speicherbausteinen*, Diplomarbeit, TU-Berlin, 2009.

[103] R. J. Warburton, B. T. Miller, C. S. Dürr, C. Bödefeld, K. Karrai, J. P. Kotthaus, G. Medeiros-Ribeiro, P. M. Petroff, and S. Huant, "Coulomb interactions in small charge-tunable quantum dots: A simple mode"l, *Phys. Rev. B* 58(24), 1622116231 (1998).

[104] A. Schliwa, M. Winkelnkemper, and D. Bimberg, "Few-particle energies versus geometry and composition of $In_xGa_{1-x}As$/GaAs self-organized quantum dots", *Phys. Rev. B* 79 (2009).

[105] J. Frenkel, "On pre-breakdown phenomena in insulators and electronic semiconductors", *Phys. Rev.* 54, 647 (1938).

[106] B. Ricco, and M. Ya. Azbel, "Physics of resonant tunneling: The one-dimensional double-barrier case", *Phys. Rev. B* 29 (4), 1970 (1984).

[107] R. Tsu and L. Esaki, "Tunneling in a finite superlattice", *Appl. Phys. Lett.* 22, 562 (1973).

[108] L. L. Chang, L. Esaki, and R. Tsu, "Resonant tunneling in semiconductor double barriers", *Appl. Phys. Lett.* 24, 593 (1974).

[109] L. L. Chang, E. E. Mendez, and C. Tejedor, *Resonant tunneling in semiconductors: Physics and applications*, Springer, 1991.

[110] A. M. C. Serra and H. A. Santos, A one-dimensional, self-consistent numerical solution of Schrödinger and Poisson equations", **Jour. of Appl. Phys.** 70, 2734 (1991).

[111] L. Wang, P. M. Asbeck, Y. Taur, "Self-consistent 1-D SchrödingerPoisson solver for IIIV heterostructures accounting for conduction band non-parabolicity", *Solid-State Electronics* 54, 12571262 (2010).

[112] P. J. Oliver, *Introduction to Partial Differential Equations*, Springer, Switzerland, 2014.

[113] L. E. Henrickson, A. J. Glick, G. W. Bryant, D. F. Barbe, "Non-equilibrium Green's function theory of transport in interacting quantum dots", *Phys. Rev. B* 50, 7 (1994).

[114] A. Marent, M. Geller, T. Nowozin, and D. Bimberg, U.S. Patent 8, 331, 142 (2012).

[115] X. D. Zhao, H. Yamamoto, K. Taniguchi, and Z. M. Chen, "Resonant tunneling in periodic multiple-barrier structures with compound-barrier unit", *Superlattice and Microstrcutures* 19 (4), 279 (1996).

[116] E. E. Mendez, W. I. Wang, B. Ricco, and L. Esaki, "Resonant tunneling of holes in AlAs/GaAs/AlAs heterostructures", *Appl. Phys. Lett.* 47, 415 (1985).

[117] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, "Quantum cascade laser", *Science* 264 (5158), 553 (1994).

[118] O. V. Pupysheva, A. V. Dimitriev, A. V. Farajian, H. Mizuseki, and Y. Kawazoe, "Transition between N- and Z- shaped current-voltage characteristics in semiconductor multiple-quantum-well structures", *Jour. of Appl. Phys.*, 100, 033718 (2006).

[119] M. Tsuchiya, and H. Sakaki, "Dependence of resonant tunneling current on well widths in AlAs/GaAs/AlAs double barrier diode structures", *Appl. Phys. Lett.* 49, 88 (1986).

[120] M. Tsuchiya, and H. Sakaki, "ependence of resonant tunneling current on Al mole fractions in $Al_xGa_{1-x}As$-GaAs- $Al_xGa_{1-x}As$ double barrier structures", *Appl. Phys. Lett.* 50, 1053 (1987).

[121] I. H. Tan, G. L. Snider, L. D. Chang, and E. L. Hu, "A self-consistent solution of Schrödinger-Poisson equations using a nonuniform mesh", *Jour. of Appl. Phys.* 68 (6), 2849 (1990).

[122] G. L. Snider, I. H. Tan, and E. L. Hu, "Electron states in mesa-etched one-dimendional quantum well wires", *J. Appl. Phys.* 68 (8), 4071 (1990).

[123] L. R. Ram-Mohan, K. H. Yoo, J. Moussa, "The SchrödingerPoisson self-consistency in layered quantum semiconductor structures", *J. Appl. Phys.* 95, 6 (2004).

[124] S. E. Laux and F. Stern, "Electron states in narrow gate-induced channels in Si", *Appl. Phys. Lett.* 49 (91), (1986).

[125] G. Snider, 1D Poisson-Schrödinger Solver, https://www3.nd.edu/ gsnider/, 2013.

[126] S. Datta, *Quantum Transport Atom to Transistor*, edited by C. U. Press, Cambridge University Press, 2005.

[127] S. Datta, "Nonequilibrium Green's function method applied to double-barrier resonant-tunneling diodes", *Phys. Rev. B* 45 (12), 253-278 (1992).

[128] S. Datta, "Nanoscale device modeling: the Green's function method", *Superlattice and Microstrcutures* 28 (4), 253-278 (2000).

[129] S. Datta, *Electronic Transport in Mesoscopic Systems*, Cambridge University Press, 1995.

[130] D. K. Schroder, *Semiconductor material and device characterization*, John Wiley & Sons, Hoboken, New Jersey, 3rd edition, 2006.

[131] P. N. Brunkov, S. G. Konnikov, V. M. Ustinov, A. E. Zhukov, A. Yu. Egorov, M. V. Maksimov, N. N. Ledentsov, and S. P. Kop'ev, "Capacitance spectroscopy of electron energy levels in InAs quantum dots in a GaAs ,atrix", *Semiconductors* 30 (5), 1063-7826 (1996).

[132] G. L. Miller, D. V. Lang, and L. C. Kimerling, "Capacitance Transient Spectroscopy", *Ann. Rev. Mat. Sci.* 7, 377 (1977).

[133] H. G. Grimmeiss and C. Ovr'en, "Fundamentals of junction measurements in the study of deep energy levels in semiconductors", *J. Phys. E: Sci. Instrum.* 14, 1032 (1981).

[134] M. Geller, E. Stock, C. Kapteyn, R. L. Sellin, and D. Bimberg, "Tunneling emission from self-organized In(Ga)As/GaAs quantum dots observed via time-resolved capacitance measurements", *Phys. Rev. B* 73(20), 205331 (2006).

[135] O. Engström, M. Kaniewska, M. Kaczmarczyk, and W. Jung, "Electron tunneling from quantum dots characterized by deep level transient spectroscopy", *Appl. Phys. Lett.* 91, 133117 (2007).

[136] S. Anand, N. Carlsson, M. E. Pistol, L. Samuelson, and W. Seifert, "Deep level transient spectroscopy of InP quantum dots", *Appl. Phys. Lett.* 67(20), 3016 (1995).

[137] C. M. A. Kapteyn, F. Heinrichsdorff, O. Stier, R. Heitz, M. Grundmann, N. D. Zakharov, D. Bimberg, and P. Werner, "Electron Escape from InAs Quantum Dots", *Phys. Rev. B* 60(20), 14265 (1999).

[138] C. M. A. Kapteyn, M. Lion, R. Heitz, D. Bimberg, P. N. Brunkov, B. V. Volovik, S. G. Konnikov, A. R. Kovsh, and V. M. Ustinov, "Hole and electron emission from InAs quantum dots", *Appl. Phys. Lett.* 76(12), 1573 (2000).

[139] C. M. A. Kapteyn, M. Lion, R. Heitz, D. Bimberg, C. Miesner, T. Asperger, and G. Abstreiter, "Many-particle effects in Ge quantum dots investigated by time-resolved capacitance spectroscopy", *Appl. Phys. Lett.* 77(25), 4169 (2000).

[140] O. Engström, M. Malmkvist, Y. Fu, H. Ö. Olafsson, and E. Ö. Sveinbjörnsson, "Thermal emission of electrons from selected s-shell configurations in InAs/GaAs quantum dots", *Appl. Phys. Lett.* 83(17), 35783580 (2003).

[141] S. Schulz, S. Schnüll, C. Heyn, and W. Hansen, "Charge-state dependence of InAs quantum-dot emission energies", *Phys. Rev. B* 69(19), 195317 (2004).

[142] D. S. Day, M. Y. Tsai, B. G. Streetman, and D. V. Lang, "Deep-level-transient spectroscopy: System effects and data analysis", *J. Appl. Phys.* 50(8), 5093 (1979).

[143] P. Omling, L. Samuelson, and H. G. Grimmeis, "Deep level transient spectroscopy evaluation of non-exponential transients in semiconductor alloys", *J. Appl. Phys.* 54(9), 5117 (1983).

[144] S. El Kazzi, L. Desplanque, X. Wallart, Y. Wang, and P. Ruterana, "Interplay between Sb flux and growth temperature during the formation of GaSb islands on GaP", *Journal of Applied Physics* 111, 123506 (2012).

[145] D. S. Abramkin, E. A. Emelyanov, M. A. Putyato, A. K. Gutakovskii, A. S. Kozhukhov, B. R. Semyagin, V. V. Preobrazhenskii, and T. S. Shamirzaev, "Formation and crystal structure of GaSb/GaP quantum dots", *Bull. Russ. Acad. Sci. Phys.* 80, 1722 (2016).

[146] S. -H. Wei and A. Zunger, "Calculated natural band offsets of all II-VI and III-V semiconductors: Chemical trends and the role of cation d orbitals", *Applied Physics Letters* 72, 2011 (1998).

[147] T. Nowozin, D. Bimberg, K. Daqrouq, M. N. Ajour, and M. Awedh, "Materials for future quantum dot-based memories", *J. Nanomater.*, 215613 (2013).

[148] L. Bonato, I. F. Arikan, L. Desplanque, C. Coinon, X. Wallart, Y. Wang, P. Ruterana, and D. Bimberg, "Hole localization energy of 1.18 eV in GaSb quantum dots embedded in GaP", *Physica Status Solidi b* 10, 18771881 (2016).

[149] S. Pennycook and D. Jesson, "High-resolution Z-contrast imaging of crystals", *Ultramicroscopy* 37, 14-38 (1991).

[150] E. M. Sala, G. Stracke, S. Selve, T. Niermann, M. Lehmann, S. Schlichting, F. Nippert, G. Callsen, A. Strittmatter, and D. Bimberg, "Growth and structure of $In_{0.5}Ga_{0.5}Sb$ quantum dots on GaP(001)", *Appl. Phys. Lett.* 109, 102102 (2016).

[151] D. Guimard, M. Nishioka, S. Tsukamoto, and Y. Arakawa, "Effect of antimony on the density of InAs Sb:GaAs(100) quantum dots grown by metalorganic chemical-vapor deposition", *Journal of Crystal Growth* 298, 548-552 (2007)

[152] D. Guimard, M. Ishida, L. Li, M. Nishioka, Y. Tanaka, H. Sudo, T. Yamamoto, H. Kondo, M. Sugawara, and Y. Arakawa, "Interface properties of InAs quantum dots produced by antimony surfactant-mediated growth: Etching of segregated antimony and its impact on the photoluminescence and lasing characteristics", *Appl. Phys. Lett.* 94, 103116 (2009).

[153] J. Starkiewicz and J. W. Allen, "Injection electroluminescence at p-n junctions in Zinc-doped Gallium Phosphide", *J. Phys. Chem. Solids* 23, 881-884 (1962).

[154] L. L. Chang and G. L. Pearson, "Diffusion Mechanism of Zn in GaAs and GaP Based on Isoconcentration Diffusion Experiments", *J. of Appl. Phys.* 35, 6 (1964).

[155] H. W. Allison, "Solubility and Diffusion of Zinc in Gallium Phosphide", *J. of Appl. Phys.* 34, (1963).

[156] X. -L. Wang, A. Wakahara, A. Sakaki, "Si and Zn doping of GaP grown by OMVPE using tertiarybutylphosphine", *J. of Crys. Growth* 158, 49-52 (1996).

[157] K. Kaneko, M. Dosen and N. Watanabe, "Efficient Red LEDs of GaP by Vapor Phase Doping of Zinc", *Jpn. J. Appl. Phys.* 12, 1732 (1973).

[158] K. A. Bertness, S. R. Kurtz, D. J. Friedman, A. E. Kibbler, C. Kramer, and J. M. Olson, "29.5%-efficient GaInP/GaAs tandem solar cells", *Appl. Phys. Lett.* 65, 989 (1994).

[159] O. Madelung, U. Rössler, and M. Schulz, "Gallium Phosphide (GaP), energies and capture cross sections of hole traps", *Landolt-Brnstein - Group III Condensed Matter.* 41A2ß. Springer Materials (2016).

[160] W. Mönch, "On the band-structure lineup at Schottky contacts and semiconductor heterostructures", *Mater. Sci. Semicond. Processing* 28, 2 (2014)

[161] W. Mönch, *Electronic Properties of Semiconductor Interfaces*, Springer, Berlin, 2004.

[162] F. Heinrichsdorff, A. Krost, M. Grundmann, D. Bimberg, F. Bertram, J. Christen, A. Kosogov, and P. Werner, "Self organization phenomena of In_xGa_{1-x}AsGaAs quantum dots grown by metalorganic chemical vapour deposition", *J. Cryst. Growth* 170, 568-573 (1997).

[163] K. Pötschke, L. Müller-Kirsch, R. Heitz, R. L. Sellin, U. W. Pohl, D. Bimberg, N. Zakharov, and P. Werner, "Ripening of self-organized InAs quantum dots", *Physica E* 21, 606-610 (2004).

[164] T. Nowozin, A. Marent, M. Geller, D. Bimberg, N. Akcay, and N. Öncan, "Temperature and electric field dependence of the carrier emission processes in a quantum dot-based memory structure", *Appl. Phys. Lett.* 94, 042108 (2009).

[165] J. R. Taylor, *An Introduction to Error Analysis*, University Science Books, 1997.

Acknowledgements

First and foremost, I would like to thank Prof. Dr. Dieter Bimberg for giving me the opportunity to work in his group. This study couldn't have been achieved without his academic guidance and financial support. Besides, I was always welcomed by him, when I need help.

I am grateful to Prof. Dr. André Strittmatter for his financial support.

I am thankful to other members of the commitee: to apl. Prof. Dr. Armin Dadgar for reviewing the thesis and to Prof. Dr. Michael Lehmann for heading the committee.

I would like to thank Dr. Tobias Nowozin for making me familiar with the topic and for sharing the office.

I have to specially thank Dr. Leo Bonato for the collaboration and fruitful discussions over the course of the study, for proof-reading of the thesis, and for sharing the office in a pleasant atmosphere.

I am very thankful to Prof. Dr. Xavier Wallart and Elisa Sala Maddelana for growth of the samples used in this work. Besides, I thank Elisa for the collaboration and helpful discussions.

My gratitudes are also extended to my colleagues and members of the group for being helpful and for a friendly working environment: Dr. Ricardo Rosales, Dr. Tagir Sadeev, Dr. Gunter Larisch and Dr. Jarez Miah. I also thank Stephan Bock and Ronny Schmidt for their help in the clean room.

I am very thankful to Mrs Hella Farell and especially to Mrs Roswitha Koskinas for their assistance with bureaucracy, so that I was able to concentrate on my research.

I am very thankful to Volkan Odabas for supporting me via phone calls from Istanbul.

Finally, I would like to thank my family for their unceasingly support and trust.