Resonant Tunnelling Diode – Photodetectors for spiking neural networks

J Lourenço[1*], Q R Al-Taai[2], A Al-Khalidi[2], E Wasige[2] and J Figueiredo[1]

[1]Centra-Ciências and Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Portugal.
[2]High Frequency Electronics Group, James Watt School of Engineering, University of Glasgow, UK

*jplourenco@fc.ul.pt

Abstract. Spike-based neuromorphic devices promise to alleviate the energy greed of the artificial intelligence hardware by using spiking neural networks (SNNs), which employ neuron like units to process information through the timing of the spikes. These neuron-like devices only consume energy when active. Recent works have shown that resonant tunnelling diodes (RTDs) incorporating optoelectronic functionalities such as photodetection and light emission can play a major role on photonic SNNs. RTDs are devices that display an N-shaped current-voltage characteristics capable of providing negative differential conductance (NDC) over a range of the operating voltages. Specifically, RTD photodetectors (RTD-PDs) show promise due to their unique mixture of the structural simplicity while simultaneously providing highly complex non-linear behavior. The goal of this work is to present a systematic study of the how the thickness of the RTD-PD light absorption layers (100, 250, 500 nm) and the device size impacts on the performance of InGaAs RTD-PDs, namely on its responsivity and time response when operating in the third (1550 nm) optical transmission window. Our focus is on the overall characterization of the device optoelectronic response including the impact of the light absorption on the device static current-voltage characteristic, the responsivity and the photodetection time response. For the static characterization, the devices I-V curves were measured under dark conditions and under illumination, giving insights on the light induced I-V tunability effect. The RTD-PD responsivity was compared to the response of a commercial photodetector. The characterization of the temporal response included its capacity to generate optical induced neuronal-like electrical spike, that is, when working as an opto-to-electrical spike converter. The experimental data obtained at each characterization phase is being used for the evaluation and refinement of a behavioral model for RTD-PD devices under construction.

1. Introduction

Over the past decade the computational needs for nearly every industry have been rising at galloping rate mainly caused by pervasiveness of machine learning methods and the potential competitive advantage that these methods offer. All these advantages come with the price of high energy consumption which can be attributed to the computing architecture used by these computational resources [1].

Novel architectures have been explored in the past few years to tackle these issues of high energy consumption by computational resources. Of the many explored computing architectures, the photonic neuromorphic approach seems to offer one of the best solutions for next generation computational resources as it promises to be very energy efficient [2]. Photonic neuromorphic architectures, in particular Spiking Neural Networks (SNNs), take inspiration from neural structures found in nature and attempts to mimic or emulate their computational process by making use of spiking generation or

processing while harnessing the advantages of optical computing. Due to their novelty, the best technology as to implement these architectures is still a matter of discussion.

Optoelectronic devices incorporating Resonant Tunnelling Diodes (RTDs) like structures have been considered among the best candidate technologies [2] for these low power consuming computing architectures. The RTD is an optoelectronic device that is formed by low energy bandgap semiconductor surrounded by thin barriers of a higher energy bandgap material both sandwiched by lower energy band gap materials forming a double barrier quantum well (DBQW). A depiction of the structure of a typical RTD-Photodetector (RTD-PD) and its band diagram are presented in as Figure 1 in the a) and c) panels, respectively.

Under this configuration the DBQW acts as an electron wave filter controlling the transmission of carriers through the device in a non-linear manner. When the energy of incoming carriers is close or equal to the energy levels - in resonance - of the quasi-bound states of the DBQW that transmission is close to unity; when the energy is not close to these levels – off resonance – the transmission is greatly reduced. These unique carrier transmission dynamics leads to the presence of a region negative differential conductance (NDC) in the RTD's I-V, as depicted in the b) panel of Figure 1.

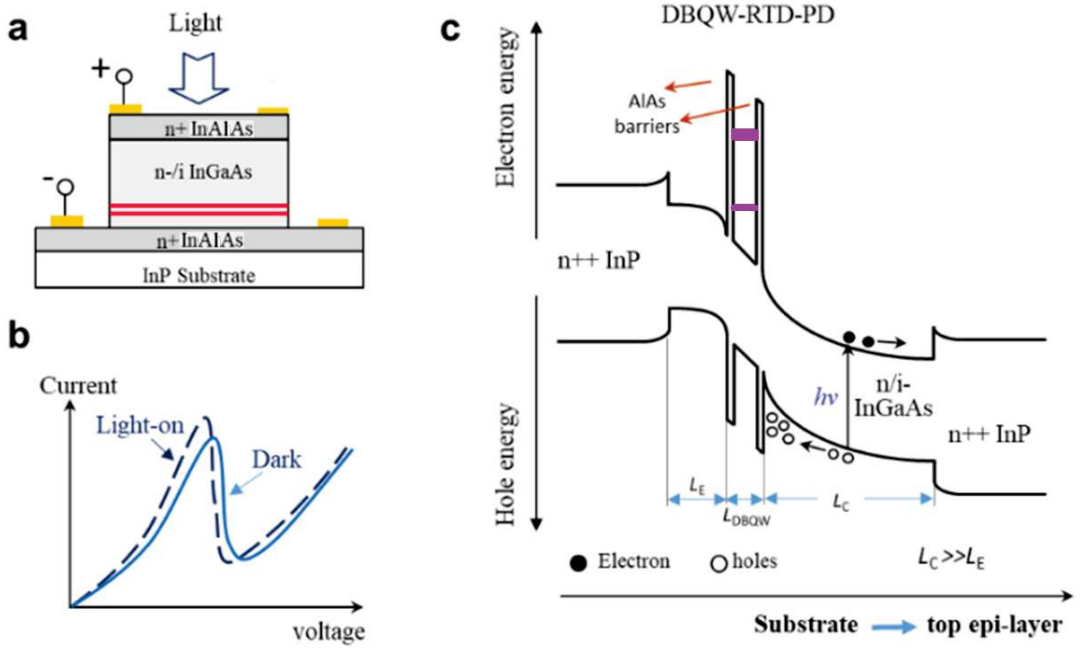

Figure 1 a) Structure of an InGaAs/InP RTD-PD; b) I-V curve of RTD-PD under dark and illumination; c) Band diagram of an RTD-PD at resonance.

This non-linear behaviour displayed by the RTD is what propels this device to one best candidates for the neuromorphic computing architectures for several reasons. First and foremost, the RTD can exhibit excitable dynamics[3,4] like biological neurons. Excitable dynamics or excitability is the capacity of a system to respond strongly to a weak stimulus if it is over a specific threshold before relaxing down to the same state it was previously to the stimulus [5], if the stimulus does not push the device over the threshold only a small response is produced. Second, the RTD is the highest bandwidth room-temperature device capable of operating at frequencies up to the Terahertz range [6]. Finally, these devices when designed as Photodetectors (RTD-PD) have intrinsic electrical gain and act as highly sensitive photodetectors.

The ultimate goal of RTD-PD design is to obtain a device that is capable of operating at high speeds and highly sensitive. Within the context of neuromorphic photonics, such device would also display excellent photonic neuron-like behavior such as optical induced spike generation. Here we present preliminary results on the characterization of the InGaAs/InP RTD-PDs in the third (1550 nm) optical communication window, with different light absorption layer thickness, in terms of responsivity and optical induced electrical spike generation.

This paper is organized as follows, in the first section the characterized devices and samples are presented alongside the criteria for selection of the best devices for neuromorphic applications. The next section presents the results of the Static (DC) and Dynamic (AC) responsivity characterization of RTD-PDs. The next section contains the pulse and spiking responsivity characterization, and neuromorphic like behavior of the selected devices. Finally, the last section presents the conclusions obtained from this work and the prospects to move this study forward.

2. Devices

The devices characterized in this study were InGaAs/InP devices with different absorption layers length, 100 nm, 250 nm and 500 nm. The devices with 100 nm and 250 nm spacers had an asymmetric design with absorption layers (spacers) only on the emitter side of the DBQW, with InAlAs cladding layers (see Fig. 1a) adjacent to the InGaAs highly doped contacts (not shown in Fig. 1a)). The devices with 500 nm absorption layer have also an asymmetric design but lack the InAlAs cladding layers.

Furthermore, three mesa sizes were characterized 20x20, 15x15 and 10x10 um^2. Not all mesa sizes were available for all absorption length layer thicknesses. For the devices with 500 nm, only 15x15 and 20x20 um^2 mesa sizes were available; for the 250 nm only the 10x10 and 20x20 um^2 were available; for the 100 nm all the mesa sizes were available. All these devices were characterized in terms of their static and dynamic responsivity. From the results of these characterizations two devices were selected for the pulse and spiking response characterization. The criteria used for the selection were: the peak current – which affects the magnitude of the response; static responsivity – capacity for photogeneration of carriers; and dynamic responsivity – which is related to the bandwidth of the device.

3. Static and Dynamic response characterization

When illuminated, an RTD-PD with an intentional added light absorption layer of thickness L operates in a similar way as a photoconductive device, producing an excess current due to photogeneration of excess carriers, holes in the valence band and electrons in the conduction band. When subjected to an applied bias, these excess photogenerated carriers (electron-holes) are swept towards the contacts and hence into the external circuit, giving rise to a photocurrent that adds to the device current when no light is present. The magnitude of the photocurrent is proportional to the incident flux of photons and to $[1 - \exp(\alpha \cdot L)]$, where α is the absorption coefficient of the light absorption layer thickness L, assuming that the absorption in any other layers of the device is negligible. The optical window on top surface of the devices under study is on direct contact with air (there is no antireflection coating), which leads to a wasteful reflection of around 30% of the incident photons. The responsivity of the RTD-PD is the ratio of the difference of the currents under illumination and dark conditions to the optical power.

We have evaluated the devices photo-response for continuous-wave illumination (from now on referred static responsivity) and for sinusoidal modulated light illumination (dynamic responsivity). As referred above, the static responsivity characterization consisted of measuring the I-V curve in dark conditions and under illumination – in this case the Thorlabs S3F1550 1550 nm laser source was used - for different optical powers. The static characterization also gave us insight on the tunability of the I-V curve, i.e., how the I-V curve changes when the device is under illumination, and the overall responsivity of each device.

For the dynamic responsivity – which gives information about the bandwidth of each device - we used the same 1550 nm laser, modulated at several distinct frequencies, which was then split using a 50:50 optical power splitter. Half was coupled to the RTD-PD and the other half to a reference high

bandwidth photodetector with a known responsivity (band DC-32 GHz). By using this reference photodetector, we computed the dynamic responsivity of the RTD-PD directly. A schematic of the setup used for the dynamic responsivity characterization is presented in Figure 2.

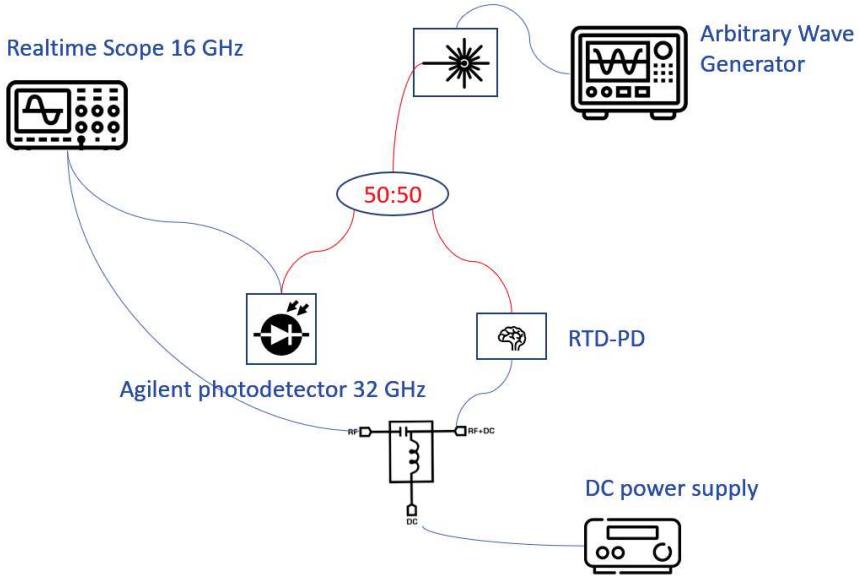

Figure 2 Dynamic Responsivity experimental setup

We started by measuring the I-V curves of all the devices for the static characterization in dark conditions. The I-V curves of the 20x20 um^2 devices for every absorption layer thickness is presented in Figure 3 and the I-V curves of the 10x10 um^2 devices is presented in Figure 4.

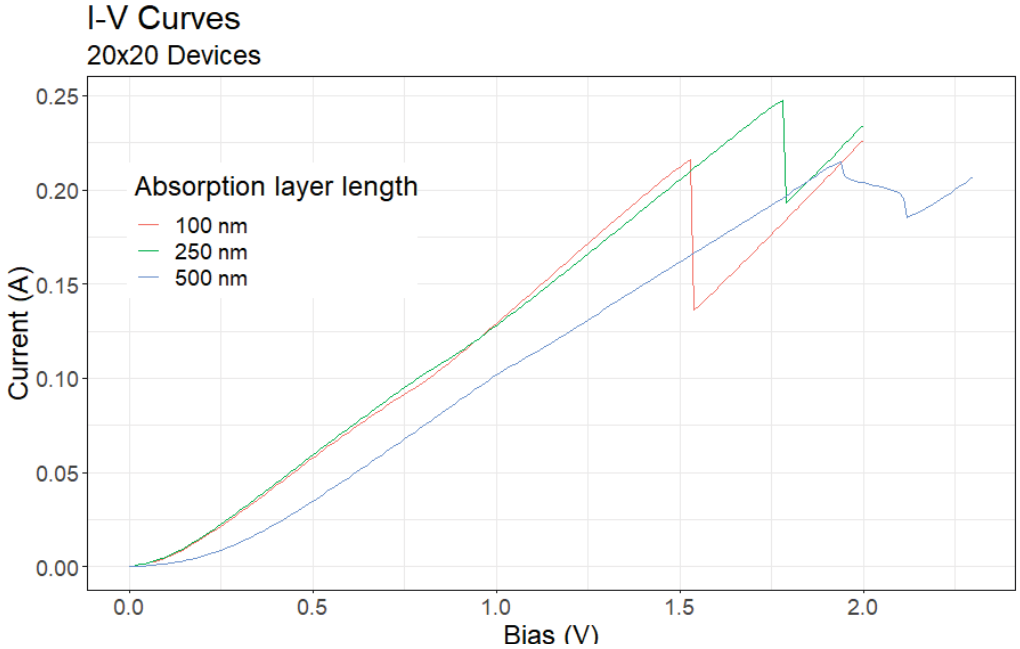

Figure 3 I-V curves of 20x20 um2 devices

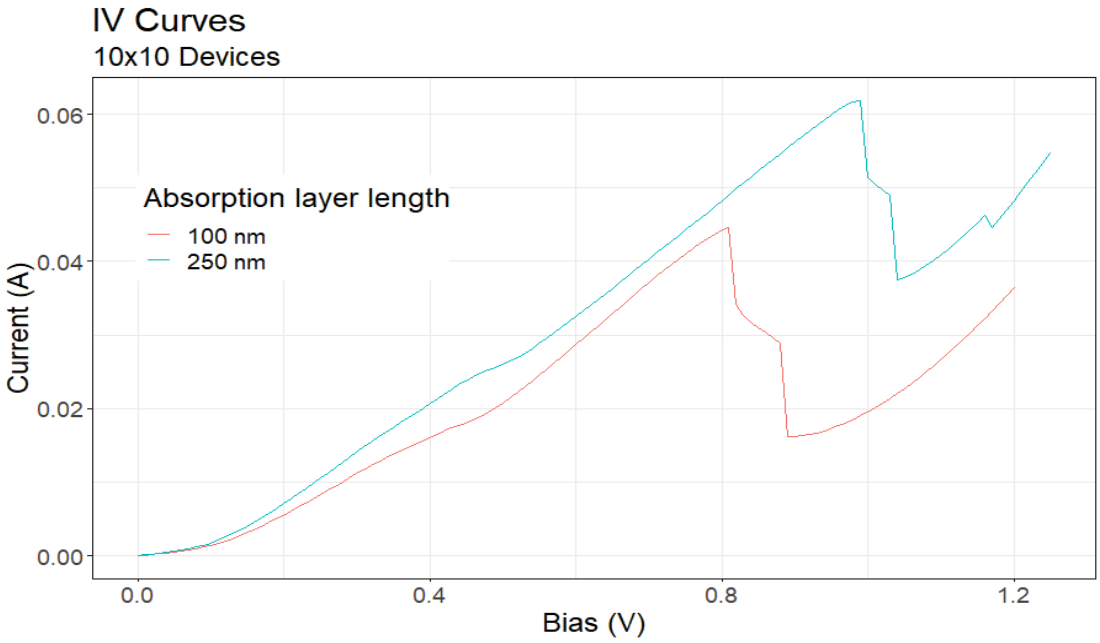

Figure 4 I-V curves of 10x10 um2 devices

Although both classes of devices show an I-V with a negative differential conductance, allowing the device to display non-linear behavior, it can be clearly seen that these regions are quite different for each of the devices and so the non-linear behavior of each of device will also be quite different.

The two main non-linear behaviors displayed by RTDs are the rapid transition between the peak and valley in a single shot manner, and the anharmonic-harmonic oscillation, usually in the form of relaxation oscillation.

The 100 nm and 250 nm absorption layer length devices with 20x20 um^2 area have regions of NDC with negligible width, with transitions between the peak and the valley region of the I-V curve very sharp. This limits the non-linear behavior of this device just to applications switching applications, since in order for self-sustained oscillations to occur it is required that the bias point of the device to be located within the region of NDC. On the other hand, the 10x10 um^2 area devices displays an I-V with a relatively large region of NDC and a smooth transition between the peak and the valley region.

Unlike the previous device, this one is not restricted to switching application and is able to act as an oscillation. Since the neuron-like dynamics in these devices, in particular spike generation, is tied to this behavior, these devices are expected to be better suited for spiking neural networks applications. This discrepancy between the I-V and shape of the NDCs of between the 10x10 and the 20x20 um^2 is explained by the shift of the peak caused by the additional resistivity of the larger device.

We now turn to the photonic properties of these devices. As it can be seen, the devices with 250 nm absorption layer shows a much higher responsivity in both cases (represented by the blue line in the Figure 6 and by the green line in Figure 5).

The reason why devices with an absorption layer length of 250 nm have a higher responsivity than devices with 100 nm is since in the latter device has shorter absorbing layer for the incoming light to be absorbed and thus produces less photocurrent.

The reason for the higher responsivity displayed by the devices with an absorption layer length of 250 nm than the for the devices with a length of 500 nm is slightly different. As it was previously explained, the devices with 100 nm and 250 nm absorption layer length had an asymmetrical design

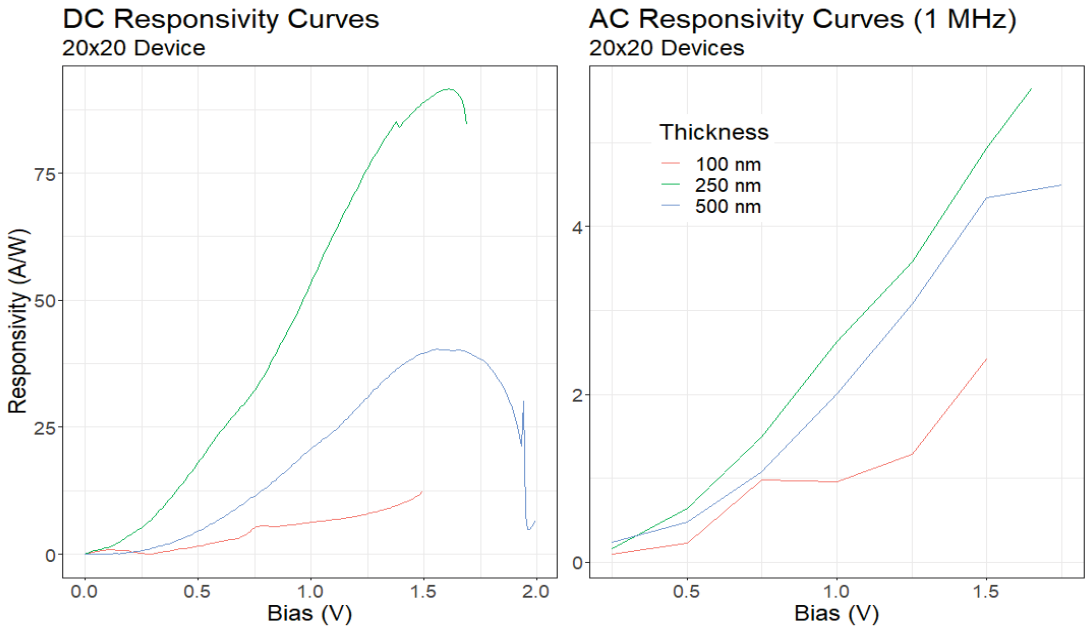

Figure 5 DC and AC Responsivity of 20x20 um2 devices

with the absorption layer only on the emitter side incorporate InAlAs cladding layers, but the devices with 500 nm absorption layer do not. The InGaAs highly doped contact layers are adjacent to the InGaAs light absorption layers adjacent to the DBQW, lead to photogenerated carriers diffusion that does not contribute to the deice overall photocurrent: As light is absorbed by the device, electron-hole pairs are generated and will diffuse until reaching regions of high electric field and then to be collected by the contacts. However, charge accumulation around DBQW will hinder this diffusion increasing the carriers transit time. As the transit time increases and becomes closer to the carrier lifetime in this material, the number of carriers that are collected by the contacts decreases but instead take part in other recombination mechanisms which leads to a reduction in the responsivity.

From the results obtained from the static and dynamic characterization it was decided that the devices that met most of our criteria pulse and spiking response were the devices with 250 nm absorption layer length with 10x10 and 20x20 um2 mesa areas.

4. Pulse response and light induced electrical spiking characterization

In this section we present pulse and spike characterization of the selected devices. For the pulse response the same experimental setup as for the dynamic characterization.

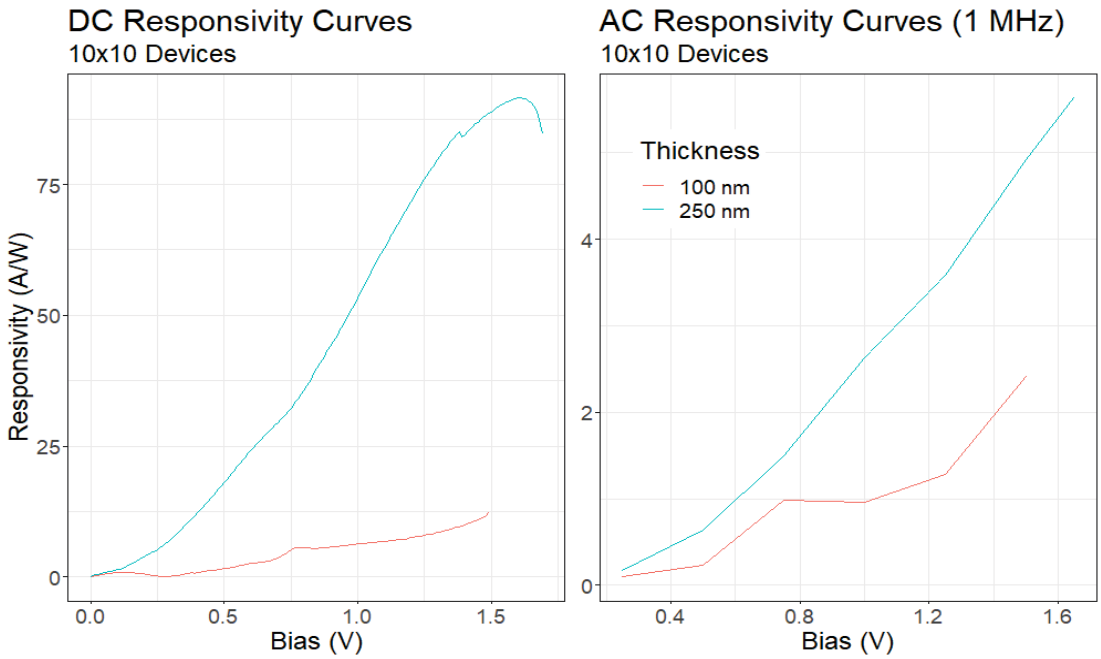

Figure 6 DC and AC responsivity of 10x10 um2 devices

Each device was biased near the peak of its I-V curve (1.65 V for the 20x20 um2 device and 0.85 V for the 10x10 um2 device). Then, an optical pulse with a width of 1 nanosecond was injected into device and its response was measured. The typical response to the optical pulse for the 20x20 and the 10x10 um^2 device is presented respectively in Figure 7 and Figure 8.

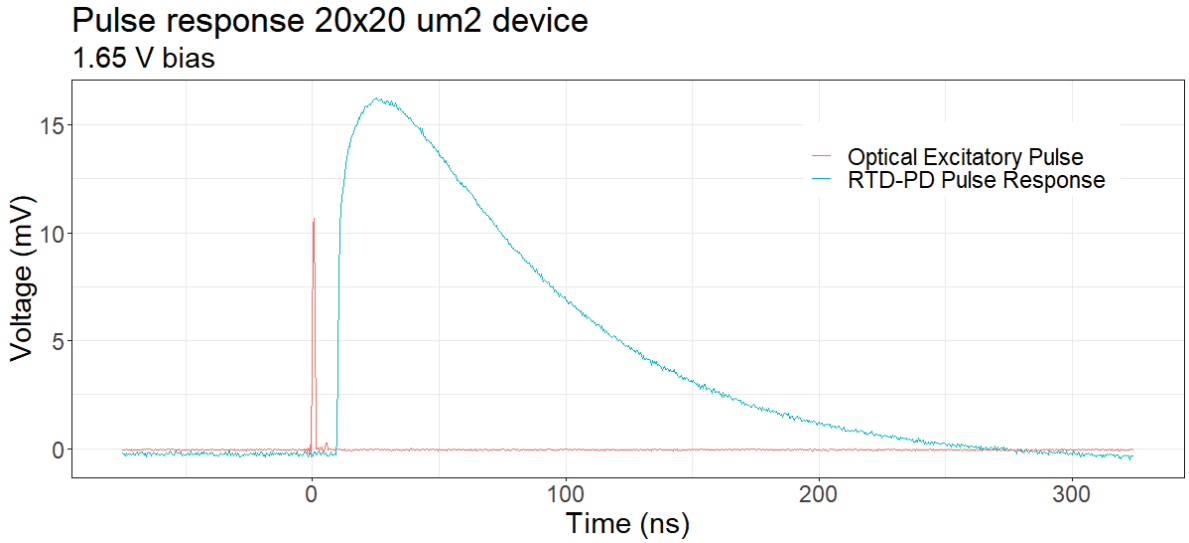

Figure 7 Pulse response of 20x20 um2 device with 250 nm absorption layer length

Several optical pulses with powers ranging from 300 uW up to 900 uW were injected. The device produced an exponentially decaying photocurrent signals response with the same fall-time (~145 ns for the 20x20 device and ~50 ns for the 10x10 device) regardless of the power of the input optical pulse.

The fact that these devices display similar fall time responses which were independent of both the

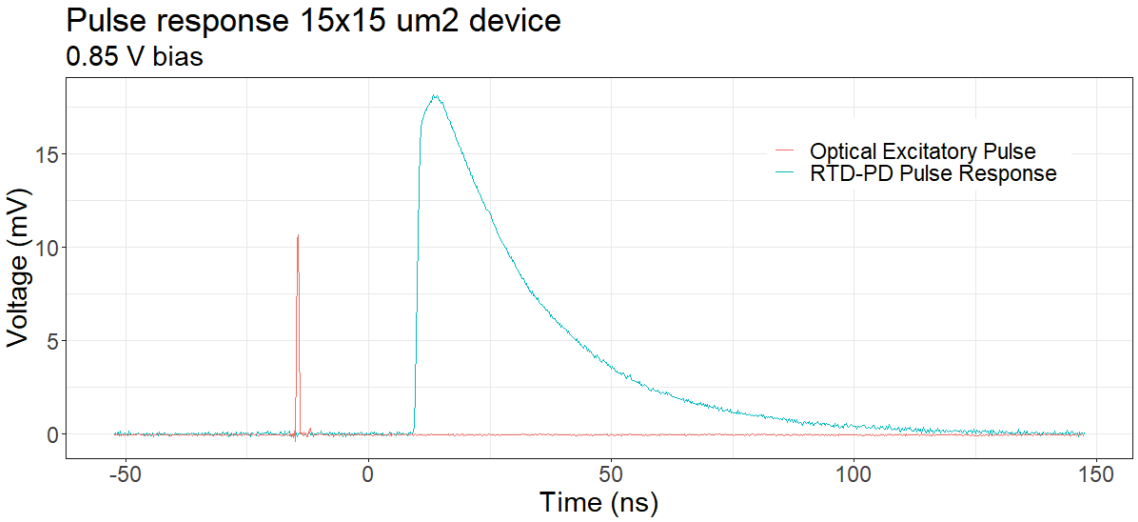

Figure 8 Pulse response of 10x10um2 device with 250 nm absorption layer length

input optical power and bias point – this fall time was measured with the device at several bias points - is most likely related to material properties such as the diffusion of carriers across the device. The photogenerated carriers are generated in a region of the device, in this case in the emitter absorption layer, where the electric field is low Thus the carrier motion mostly by diffusion, resulting in the exponential decaying behavior of the pulse response. Ideally, the difference between the measured fall times of the 10x10 and 20x20 um^2 devices should be close to zero, but a difference of 90 ns was measured. The reason for this discrepancy is because different mesa areas have different electric field distributions which has a non-negligible impact on the motion of the photogenerated carriers.

Next, we present and discuss the optically induced electrical spike generation characterization. For this characterization the same setup was as for the dynamic and the pulse response characterization was used but with small tweak, the addition of a resonant tank circuit.

Due to the interplay between the RTD and the inductance of the resonant circuit, allows charge to be transferred from one circuit element to another with any losses being compensated by the RTD intrinsic gain, thus effectively transforming the RTD into a relaxation oscillator capable of displaying excitable dynamics.

The 20x20 um2 was not able to sustain self-oscillation when the tank circuit was added, nor it was able to generate usable spikes. In fact, its fall time in this configuration increased ten-fold when compared with configuration without the resonant circuit. There are several possible reasons why this device is not able to sustain self-oscillations nor able to generate usable spikes. The fact that this device is not able to oscillate is most likely due to the shape of the I-V curve of this device (Green line presented in Figure 3) which has a very sharp, nearly non-existent region of NDC: self-oscillation usually requires a device that can be biased in the NDC.

On the other hand, the 10x10 um2 device displayed a whole range of dynamic behavior when the resonant tank circuit is added: It was able to sustain self-oscillation at the frequency of 125 MHz (8 ns period) ; displayed neuromorphic features such as excitability and optical to electrical spike conversion, in the sense that a single optical perturbation can generate an electrical spike response, and sub-threshold signal integration.

As it was previously mentioned these devices can display excitable dynamics, which means that when the power of the of an external stimulus signal is enough to push the device's state over a certain threshold, the device produces a large response. The threshold is defined by the system itself, in the case of RTDs/RTD-PDs its threshold is the boundary of the region of NDC in the I-V, so any signal that can move the operating point of the RTD over to that region will trigger a large excitable response by the RTD. If the signal is not enough to push the device's operating point to that region, then a small slow decaying response is produced.

The typical excitable response of the RTD is characterized by two fast rising sharp peaks separated by a slow decay – similar to a biological neuron response – usually this response is referred to as spiking. The features of the spike are impacted by the device's I-V – the sharp peaks – and by the resonant circuit electrical characteristics such as its inductance and capacitance – the slow decays.

Figure 9 shows the device's response to a 1 ns long optical pulse. The device is biased at 0.975 V (near the peak), with an input optical power of 350 uW. This is a sub-threshold signal as it fails to produce a spike and just a small decaying response – like the response of the device without resonant circuit- is produced. Despite being similar in form to the case without resonant circuit, the fall time in this case is significantly smaller, this is a consequence of the presence of the resonant circuit.

If we increase the power of input optical stimulus to 430 uW, while keeping everything else equal, we can push the device's operation point over its threshold. This results in a device response that is completely different. It has significantly larger amplitude when compared with the previous weak response (4 mV vs 300 mV) and a completely different shape (exponential decay vs Spiking signal), as can be seen in Figure 10. This is an especially important feature of spiking neural networks and neuromorphic processing, as it shows that it is possible to convert an optical pulse into a similar electrical response.

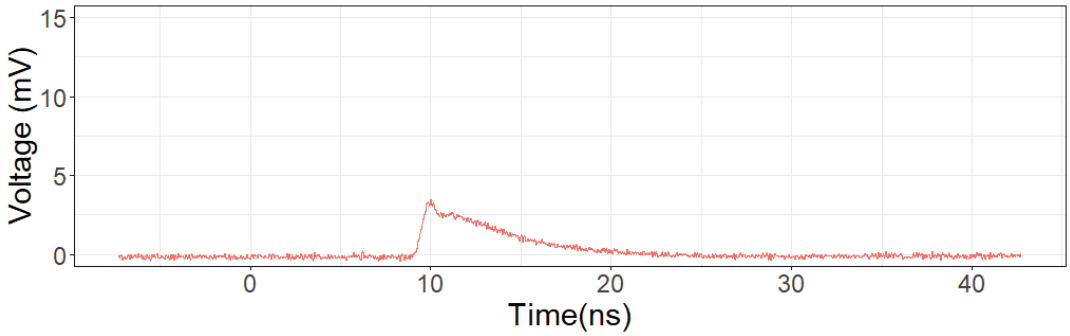

Figure 9 Sub-Threshold optical signal injection

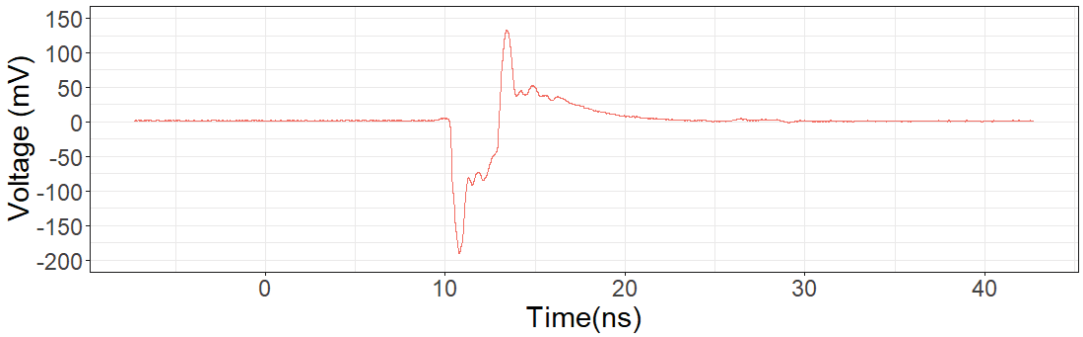

Figure 10 Optical Spike Generation - Optical to Electrical spike conversion

Finally, the last neuromorphic feature displayed by these devices during this work was the sub-threshold signal integration. While sub-threshold signals do not generate on their own trigger a spike response by the device, when several sub-threshold signals arrive within short intervals of one another the system is able to produce a spike response as if it was triggered by over-threshold signals.

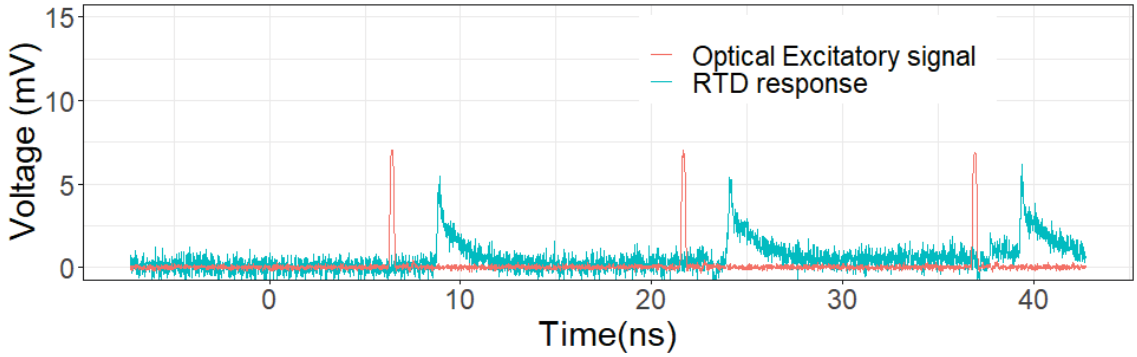

Figure 11 Sub-threshold response to multiple pulses 60 ns apart

Figure 1111 shows the device's response to three pulses with 60 ns interval between them. These pulses are sub-threshold and thus are not able to generate any sort of spike response. But as their arrival interval is reduced the device starts to generate a spike response (350 mV peak to peak signal with a 8 nano second separated between). Such scenario is depicted in Figure 122, where the three optical pulses have an inter arrival interval of only 4 ns - a value lower than the device's oscillation period (8 ns) - three electrical spike signals are generated.

It is noteworthy that in this case the signals are not fully time correlated, as the arrival time of the optical pulses is much shorter than the produced electrical spikes by the device. Time dilation of the signals occurs, or to put it in another way the system shows some sort of memory and "remembers" that it received three optical pulses with a given interval – in this case 4 ns - and is only able to generate electrical spikes at the maximum rate as set by its oscillation period – in this case 8 ns. This is the first time this phenomenon has been reported experimentally.

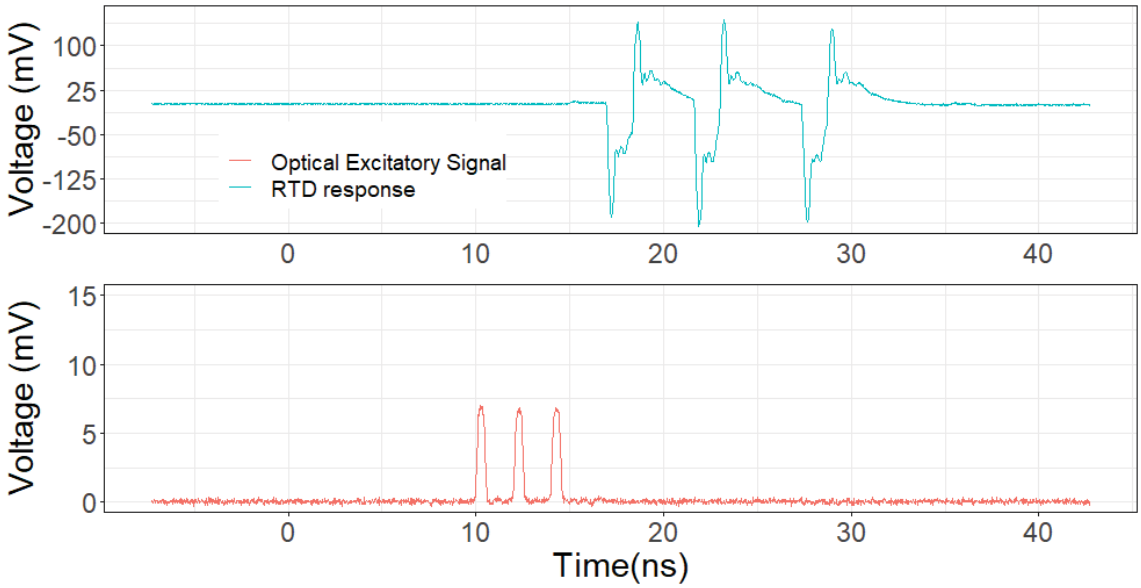

Figure 12 Sub-threshold spike integration for three input optical pulses - 4 ns apart. The system is only to produce response at its period of oscillation of 8 ns.

5. Conclusions and future work

Our goal with this work was to systematically characterize the responsivity, both the static and the dynamic, of samples of RTD-PDs with different absorption layer thickness and mesa sizes. The Static and dynamic characterization gave us insight on how the different layer thickness and mesa sizes affect the performance of the photodetection of the RTD-PD, how the responsivity is not merely a product of available the absorption length area but also how it is distributed within the epitaxial structure of the device.

From the static and dynamic characterization results we selected the best devices for the characterization for pulse and spiking response. The devices were selected using a performance-based criteria that considered, the peak current, static responsivity and dynamic responsivity. The devices which met these criteria were the 10x10 and 20x20 um^2 devices with 250 nm absorption layer length. These devices were then subjected to optical pulse and spiking response characterization. The pulse response showed that the devices have characteristic fall-times which are independent of both its bias and input optical power, hinting that these features fundamentally tied to the semiconductor material properties.

As for the spiking response characterization, only the 10x10 um^2 device displayed any dynamic behavior after the resonant tank circuit was added to the experimental setup. Nevertheless, it has been demonstrated that RTD-PDs can be used as key elements for photonic SNNs as they act as efficient optical to electrical spike conversion and have been shown experimentally for the first time – to the best of our knowledge – memory-like behavior which allow for the integration optical of sub-threshold inputs. An extension of the often-used models that describe the RTD-PD dynamics is being developed as to accommodate this new memory-like behavior reported in this work. As for future experimental work, more adequate resonant tank circuits could be designed as to improve the generation of electrical spike, as the slowest time scales are due to charge and discharge of elements that are part of the resonant tank circuit.

6. Acknowledgements
EC Grant No. 828841 659 ChipAI-H2020-FETOPEN-2018–2020, FCT Grant PD/BD/142830/2018, JWNC nanofabrication centre (UGLA).

References
[1] Nahmias M A, Shastri B J, Tait A N, Ferreira de Lima T and Prucnal P R 2018 Neuromorphic Photonics *Opt. Photonics News* **29** 34
[2] Prucnal P R, Shastri B J and Teich M C 2017 Neuromorphic photonics *Neuromorphic Photonics* (Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742: CRC Press) pp 1–412
[3] Hejda M, Alanis J A, Ortega-Piwonka I, Lourenço J, Figueiredo J, Javaloyes J, Romeira B and Hurtado A 2022 Resonant Tunneling Diode Nano-Optoelectronic Excitable Nodes for Neuromorphic Spike-Based Information Processing *Phys. Rev. Appl.* **17** 024072
[4] Ortega Piwonka J, Hejda M, Alanis J, Lourenco J, Hurtado A, Figueiredo J, Romeira B and Javaloyes J 2022 Spike propagation in a nanolaser-basedoptoelectronic neuron *Opt. Mater. Express*
[5] Romeira B, Figueiredo J M L and Javaloyes J 2017 Delay dynamics of neuromorphic optoelectronic nanoscale resonators: Perspectives and applications *Chaos An Interdiscip. J. Nonlinear Sci.* **27** 114323
[6] Asada M, Suzuki S and Kishimoto N 2008 Resonant tunneling diodes for sub-terahertz and terahertz oscillators *Jpn. J. Appl. Phys.* **47** 4375–84